PROBABILISTIC REASONING WITH INCOMPLETE AND INCONSISTENT BELIEFS

DISSERTATION

ZUR ERLANGUNG DES GRADES EINES DOKTORS DER NATURWISSENSCHAFTEN

der Technischen Universität Dortmund
an der Fakultät für Informatik

Matthias Thimm

Technische Universität Dortmund — Fakultät für Informatik — Dortmund, Germany — 2011

Tag der mündlichen Prüfung:	21.06.2011
Dekan:	Prof. Dr. Gabriele Kern-Isberner
Gutachter/Gutachterinnen:	Prof. Dr. Gabriele Kern-Isberner
	Prof. Dr. Christoph Beierle

Matthias Thimm: *Probabilistic Reasoning with Incomplete and Inconsistent Beliefs,* Dissertation, Department of Computer Science, Technische Universität Dortmund, Germany.

ABSTRACT

Reasoning with inaccurate information is a major topic within the fields of artificial intelligence in general and knowledge representation and reasoning in particular. This thesis deals with information that can be incomplete, uncertain, and contradictory. We employ probabilistic conditional logic as a foundation for our investigation. This framework allows for the representation of uncertain pieces of information by using probabilistic conditionals, a specific approach to handle *if-then-rules* within a probabilistic framework. Uncertainty can be expressed by means of probabilities attached to those rules and incompleteness can be handled in this framework by reasoning based on the principle of maximum entropy. This principle is a powerful approach when only incomplete information is at hand as it allows for filling missing pieces of information in the most unbiased way. This principle is realized for reasoning by selecting the one unique probability function that both satisfies the knowledge base and has maximum entropy. In this thesis we focus on two major issues that arise when representing knowledge with probabilistic conditional logic. On the one hand, we look at the problem of contradictory information that, e. g., arises when multiple experts share their knowledge in order to come up with a common knowledge base consisting of probabilistic conditionals. As in classical logic this is a severe problem because inconsistency of a knowledge base forbids application of model-based inductive inference approaches such as reasoning based on the principle of maximum entropy. We investigate this issue by presenting ways to analyze inconsistencies and ways to restore consistency. More specifically, we introduce the concept of measuring inconsistencies in probabilistic conditional logic and present inconsistency measures that are apt for the application in a probabilistic setting. While analyzing inconsistencies is a first step to remove inconsistencies we also have a look at automatic approaches for restoring consistency. On the other hand, we investigate an extension of the syntactical and semantical notions of probabilistic conditional logic to the relational case. Until now, most approaches for reasoning in probabilistic conditional logic only consider propositional logic as the underlying foundation. When switching to the more expressive framework of first-order logic, traditional semantics seem to fail in the interpretation of relational probabilistic conditionals in a commonsensical manner. In particular, the problem of contradictory information is an issue in first-order extensions of probabilistic conditional logic as well and we present novel semantical approaches to probabilistic conditionals that circumvent non-satisfiability in an intuitive and rational manner. We also extend the approach of reasoning based on the principle of maximum entropy to the framework of relational probabilistic conditional logic and investigate its properties.

PUBLICATIONS AND DISCLAIMER

Some ideas and texts in this thesis have been previously published in several papers. As scientific work is never the work of a single researcher, many people participated in the work reported in this thesis. In the following, I want to acknowledge both the published works that led to this thesis and the contributions of several of my co-workers who participated in these works.

The work reported in this thesis on measuring and resolving inconsistencies in probabilistic conditional logic bases on the paper (Thimm, 2009a). Nonetheless, the ideas from (Thimm, 2009a) have been further pursued and the approximations for the MinDev inconsistency measure, the extensions to more expressive frameworks, and the whole investigation on solving conflicts is novel to this thesis.

The work on inference in relational probabilistic conditional logic bases on several works by Christoph Beierle, Marc Finthammer, Jens Fisseler, Gabriele Kern-Isberner, Sebastian Loh, and myself. While the development of averaging semantics in relational probabilistic conditional logic and the properties for reasoning in this framework has been started by myself (Thimm, 2009b), the initial work on aggregating semantics is due to Gabriele Kern-Isberner (Kern-Isberner and Thimm, 2010). Further research in this area is due to Gabriele Kern-Isberner, Jens Fisseler, and myself (Thimm and Kern-Isberner, 2011; Thimm *et al.*, 2011b). Novel to this thesis is a deeper analysis of the averaging and aggregating semantics and the technical elaborations on both standard and lifted inference. Much of this work on relational probabilistic reasoning has been initialized by my participation in the DFG project KReate[1] that "aims at developing a common methodology for learning, modelling and inference in a relational probabilistic framework". While the theoretical research in this project led to the work cited above, a practical aspect of KReate was the development and implementation of the integrated development environment KReator[2] which is supervised by Marc Finthammer and myself. Reports on the work on KReator and general relational probabilistic frameworks are due to Christoph Beierle, Marc Finthammer, Gabriele Kern-Isberner, Sebastian Loh, and myself (Finthammer *et al.*, 2009; Thimm *et al.*, 2010; Beierle *et al.*, 2010a; Loh *et al.*, 2010; Finthammer and Thimm, 2011; Thimm *et al.*, 2011a). Some discussions in this thesis are influenced by those works.

1 http://www.fernuni-hagen.de/wbs/research/kreate/
2 http://kreator.cs.tu-dortmund.de/

ACKNOWLEDGMENTS

There are many people to whom I am very grateful for their support during the past few years. First of all, I want to thank Gabriele Kern-Isberner for being my advisor and guiding me around the various obstacles of scientific research. In particular, I appreciate the freedom she gave me in conducting research and the opportunities to present my work on various conferences. I also like to thank Christoph Beierle for being my second reviewer and our various discussions in Hagen and Dortmund. Furthermore, I thank Joachim Biskup and Christoph Schubert for agreeing to be part of my dissertation committee.

Special thanks go to Patrick Krümpelmann for enduring both my distraction attempts and my bad taste in tea and coffee. Our discussions were an inspiration for me and our joint trips around the globe have been fun and entertaining.

My working environment has always been wonderful and I thank everyone who took a part in that. Among others I thank the people of chairs 1 and 6 in Dortmund, the people of the chair "Wissenbasierte Systeme" in Hagen, and the people from the LIDIA lab in Bahía Blanca. Additionally, I thank Patrick Krümpelmann, Marc Finthammer, and Thomas Zeume for proofreading preliminary versions of this thesis, Christian Eichhorn for help with several figures and tables, and Adam Chachaj for his work on the implementations for reasoning in relational probabilistic conditional logic.

I thank Lora Encheva Asenova for all her support and love, especially during the final months of writing this thesis.

Although not explicitly being included in this thesis I also want to acknowledge the contributions of my co-authors in some other works as some ideas or formalizations in this thesis might be influenced by them. Those other works contain research on argumentation, more specifically, research at the intersection of formal models of argumentation and multi-agent systems which has been started in my diploma thesis (Thimm, 2008) and in (Thimm, 2010), further pursued by Gabriele Kern-Isberner and myself in (Thimm and Kern-Isberner, 2008b,a) and by Alejandro J. García, Gabriele Kern-Isberner, Guillermo R. Simari and myself in (Thimm *et al.*, 2008). Another strain of argumentation related research is on analysis of defeasible argumentation (Thimm and Kern-Isberner, 2008d,c) and is joint work with Gabriele Kern-Isberner. Research on strategical issues of argumentation is due to Alejandro J. García and myself (Thimm and García, 2010). Works on argumentation and law are joint work with Christoph Beierle, Bernhard Freund, and Gabriele Kern-Isberner (Beierle *et al.*, 2010c,b). Recent works on the combination of argumentation and belief revision is due to Diego R.

García, Sebastián Gottifredi, Patrick Krümpelmann, Gabriele Kern-Isberner, Marcelo A. Falappa, Alejandro J. García, and myself (García *et al.*, 2011).

Work on reasoning in multi-agent systems is mostly joint work with Patrick Krümpelmann (Thimm and Krümpelmann, 2009b,a; Krümpelmann and Thimm, 2010) but also joint work with Gabriele Kern-Isberner and Manuela Ritterskamp (Krümpelmann *et al.*, 2008). This work has also been extended within the student project group *Intelligent Cowbots*, resulting in the publication (Hölzgen *et al.*, 2011) which is joint work with Daniel Hölzgen, Thomas Vengels, Patrick Krümpelmann, and Gabriele Kern-Isberner. Research on confidentiality issues in multi-agent systems is joint work with Joachim Biskup and Gabriele Kern-Isberner (Biskup *et al.*, 2008).

Finally, research on qualitative knowledge discovery and data mining is joint work with Gabriele Kern-Isberner, Marc Finthammer, and Jens Fisseler (Kern-Isberner *et al.*, 2008, 2009).

CONTENTS

1 INTRODUCTION

In this chapter we introduce the topic of this thesis and give an overview on the research field where it is located. Further, we motivate the work conducted in this thesis and pose a series of research questions that are addressed. We enumerate the contributions to answer these questions and give an outline on the structure.

1.1 CONTEXT AND MOTIVATION

Knowledge representation (Brachman and Levesque, 2004) is one of the major subfields of artificial intelligence (Russell and Norvig, 2009). This field is concerned with formal representations of knowledge and how these formalizations can be used for *reasoning*, i. e., how new information can be automatically inferred using a formal system. There are many real-world applications for knowledge representation, the most common application being the *expert systems* (Jackson, 1998; Beierle and Kern-Isberner, 2008). An expert system is a piece of software that allows for representation of and reasoning with knowledge of some domain in order to support and guide the user in his or her tasks. For example, one of the earliest expert systems MYCIN (Shortliffe and Buchanan, 1975) was designed to be used in medical diagnosis, i. e., given some information on a patient's symptoms the system would identify infections such as meningitis and recommend a treatment. MYCIN performed well in laboratory experiments but was never used in practice, mainly due to acceptance problems and ethical issues (Beierle and Kern-Isberner, 2008). Nowadays, expert systems are rapidly gaining attention in many fields such as accounting (Vasarhelyi *et al.*, 2005), chemistry (Judson, 2009), and law (Popple, 1996). Furthermore, modern multi-purpose systems such as Drools[1] or Jess[2] are able to cross domain borders and allow for application in arbitrary fields. Another strain of applications for knowledge representation lies in the *semantic web* (Davies *et al.*, 2003). In the futuristic vision of the internet agents search the web, link contents, and perform time-consuming tasks on behalf of humans. Although we are far away from a purely *semantic* web there are already many systems available that make use of semantical information on a very low level like the recommendation system of Amazon[3] or the DBPedia project[4] which allows for a structured access to the contents of Wikipedia[5]. Formal knowledge representation formalisms, that allow for

1 http://www.jboss.org/drools/
2 http://www.jessrules.com/
3 http://www.amazon.com
4 http://dbpedia.org
5 http://www.wikipedia.org

a uniform method to exchange information, lie at the core of the semantic web. For those, research in the field of description logics (Baader *et al.*, 2003) and ontologies (Baader *et al.*, 2005) is applied in technologies like RDF and XML (Geroimenko, 2004).

One of the big issues in knowledge representation is *accuracy*. Usually, the term *"knowledge"* is used to describe *strict* or *objective* knowledge that is considered to be absolutely true in the given frame of reference, i. e. the real world. The counterpart, denoted by *"subjective knowledge"* or *"beliefs"*, is used to describe knowledge that is assumed to be true by the individual under consideration. As some terms like *knowledge base* have been established in the literature we adapt those conventions in this thesis. Consequently, we use the terms *"knowledge"* and *"beliefs"* interchangeably but always assume that information represented in some formalism to be subjective.

Besides being incorrect with respect to the real world, the beliefs of a single human being (and a computer system as well) can be *incomplete*, *uncertain*, or *inconsistent*. That is, some piece of information I might be unknown (*incompleteness*), might be believed only to a certain degree (*uncertainty*), or might be in conflict with another piece of information I' (*inconsistency*). Note that inconsistency of two pieces of information I and I' implies that at least one of them is *incorrect*. However, even without the possibility to compare I and I' with the state of the real world, an inconsistency *can* be detected by a being capable of reasoning, which is not necessarily true for incorrect information in general. These issues also apply to the beliefs of an expert in a field, and when feeding information into an expert system the system has to take these issues into account when inferring new information. Within the field of knowledge representation and reasoning there are several subfields that deal with (some of) these issues like *defeasible reasoning* (Kyburg *et al.*, 1990), *argumentation* (Bench-Capon and Dunne, 2007; Rahwan and Simari, 2009), or *possibilistic and fuzzy reasoning* (Siler and Buckley, 2005). Among the most established logical frameworks for dealing with uncertainty is *probability theory* (Paris, 1994; Pearl, 1998). There have been numerous works on combining probability theory with knowledge representation like Bayesian networks and Markov nets that allow for derivation of uncertain beliefs from other uncertain beliefs. Especially in application areas such as medical diagnosis, where the user has to rely crucially on the certainty of individual recommendations, reasoning using probabilistic models of knowledge serves well (Parmigiani, 2002). Probability theory—or, more precisely, information theory—also provides for a nice solution to the problem of *incomplete information*. Using the principle of maximum entropy (Paris, 1994) one can complete uncertain and incomplete information in order to gain new information that was unspecified before, see also (Kern-Isberner, 2001). The expert system SPIRIT (Rödder and Meyer, 1996) is a working system that employs reasoning based on the principle of maximum entropy and has been applied to various fields of operations research such as project risk management (Ahuja and Rödder, 2002) and portfolio selection (Rödder *et al.*, 2009). Though reasoning based on the principle of maximum entropy allows for dealing with both incomplete and uncertain information it is not suitable for reasoning with *inconsistent* information. But inconsistency is a ubiquitous matter human beings have to deal with all the time: *"Ask five economists and you'll get five different answers – six if one went to Harvard."*[6]. Of course, this phenomenon appears not only in economics but everywhere. This issue becomes most apparent when multiple experts try to build up a common knowledge base which happens regularly in knowledge engineering and expert system design. However, this issue has been dealt with in the literature only little so far, cf. (Rödder and Xu, 2001; Finthammer *et al.*, 2007; Daniel, 2009).

Another issue in knowledge representation is *expressivity*. Most traditional probabilistic methods for reasoning use propositional logic which is not suitable to express most real-world scenarios. Many applications demand for the ability to express *relational* information such as relational databases, social network modeling, or genetics. The relatively young research fields of *probabilistic inductive logic programming* and *statistical relational learning* aim at extending statistical learning and probabilistic reasoning to relational settings (Getoor and Taskar, 2007; De Raedt *et al.*, 2008). The focus of the research in those areas lies in knowledge discovery and data mining problems, i. e., they address the issue of given a (large) set of data samples, how to find "rules" that describe the data? There are also some works that aim at extending reasoning based on the principle of maximum entropy to the relational case and focus on the problem of inference, cf. e. g. (Kern-Isberner and Lukasiewicz, 2004; Fisseler, 2010; Loh *et al.*, 2010). However, inference in those frameworks treat relational formulas as schemas for their instances and base heavily on grounding relational knowledge in order to get a propositional view on the information. However, straightforward extensions of propositional techniques suffer greatly from conflicting pieces of information that arise in the process of grounding and, until now, there have been no works on principled foundations for relational probabilistic reasoning that consider alternative semantical notions.

1.2 RESEARCH QUESTIONS AND CONTRIBUTIONS

The general research question that underlies the work in this thesis can be phrased as follows:

> *How to infer knowledge from incomplete, uncertain, and possibly inconsistent information?*

We handle this question using the framework of *probabilistic conditional logic* (Benferhat *et al.*, 1999; Rödder, 2000; Kern-Isberner, 2001), a general probabilistic framework that allows for a declarative knowledge representation

6 Quote by Edgar R. Fiedler, an American economist who served as assistant secretary of the treasury for economic policy from 1971 to 1975.

based on *conditionals* (*if-then-rules*). In this framework, by employing the principle of maximum entropy one can already reason with both incomplete and uncertain information, cf. e. g. (Paris, 1994). Here, we also address the problem of inconsistency. As understanding inconsistencies is a prerequisite for resolving them we start this investigation by posing the following research question:

> *How to analyze inconsistencies in probabilistic conditional logic and how to measure their severities?*

Understanding both the causes of inconsistencies and their implications also helps in gaining more insight into the semantics of the logic under consideration. There are only few works that consider the problem of analyzing inconsistencies in probabilistic conditional logic (Rödder and Xu, 2001; Finthammer *et al.*, 2007) or generalizations thereof (Daniel, 2009) and we investigate this topic in-depth. The goal of this investigation is to lead to techniques that resolve inconsistencies.

> *How to restore consistency in inconsistent probabilistic knowledge?*

Restoring consistency is a major topic in many fields of artificial intelligence like belief revision (Hansson, 1999) and information fusion (Bloch and Hunter, 2001), and there have been only few works that consider this problem for probabilistic frameworks, see (Finthammer *et al.*, 2007) and (Rödder and Xu, 2001) for some exceptions. By developing techniques that deal with inconsistency in propositional probabilistic conditional logic we reach a position where we are able to reason with uncertain, incomplete, and inconsistent information at the same time. But as discussed before, propositional logic is not expressive enough to represent most real-world scenarios. Consequently, we switch to the more expressive *first-order logic* and discuss the semantical consequences of this extension. As the combination of relational concepts and probabilistic conditional logic is novel we have to start by laying solid syntactical and semantical foundations first. Therefore, our first issue lies in answering the following research question:

> *How to express relational knowledge in probabilistic conditional logic and what is a meaningful interpretation of relational conditionals?*

Previous approaches that aim in answering this question and base on schematic interpretation of relational conditionals can be found in e. g. (Kern-Isberner and Lukasiewicz, 2004; Fisseler, 2010; Loh *et al.*, 2010). Having defined relational probabilistic conditional logic on a firm basis we are interested in reasoning. As for the propositional case we demand a reasoning mechanism that allows for handling incomplete, uncertain, and possibly inconsistent information. Therefore, the generic research question can be phrased as follows:

> *How to infer knowledge from relational probabilistic conditionals?*

Although this question has been phrased rather vaguely it is also intended to address the issues of *principled* and *efficient* reasoning. Previous work in this direction can be found in (Kern-Isberner and Lukasiewicz, 2004; Fisseler, 2010; Loh *et al.*, 2010).

In the following, we list the contributions of this thesis that aim at answering the questions raised above.

1.2.1 *Measuring Inconsistency in Probabilistic Conditional Logic*

Building on results that have been previously published in (Thimm, 2009a) we investigate the problem of analyzing and measuring inconsistencies in probabilistic conditional logic. We do this in a principled way by developing a series of rationality postulates. We extend existing inconsistency measures for classical logics to the probabilistic setting and investigate their properties. As those measures lack adhering to several peculiarities of the probabilistic setting we develop a novel inconsistency measure, compare it to the classical measures, and investigate its properties.

1.2.2 *Solving Inconsistencies*

We present *culpability measures* (Daniel, 2009) as a generalization of inconsistency measures that show how inconsistency is distributed across the individual pieces of information. We develop two culpability measures and employ those for consistency restoration. We also propose a series of rationality postulates for solving inconsistencies and adapt properties from both social choice theory and belief merging which address similar questions.

1.2.3 *Novel Semantical Approaches to Relational Probabilistic Conditional Logic*

We extend the formalism of probabilistic conditional logic to a relational setting by introducing first-order concepts into the syntax of conditionals. We do this by building on previous work (Thimm, 2009b; Kern-Isberner and Thimm, 2010) and propose novel semantical approaches for the interpretation of relational conditionals. We analyze and compare these semantics and discuss their properties.

1.2.4 *Inference in Relational Probabilistic Conditional Logic*

By employing the novel semantical approaches we discuss the problem of inductive inference for relational conditionals and continue previous work (Thimm, 2009b; Kern-Isberner and Thimm, 2010). We propose several rationality postulates for inference and make use of the principle of maximum entropy to define an inference mechanism that allows for a flexible reasoning behavior.

1.2.5 *Lifted Inference in Relational Probabilistic Conditional Logic*

We circumvent computational problems for reasoning in relational probabilistic conditional logic by exploiting structural similarities of probability functions that have been obtained by applying the principle of maximum entropy. We develop a reasoning mechanism that is tractable for unary languages and analyze its properties.

1.3 OUTLINE

This thesis is organized as follows. In Chapter 2 we give the necessary preliminaries for this thesis. These preliminaries consist of classical logic and probabilistic knowledge representation and reasoning. More precisely, we give an introduction to both propositional and first-order logic by formalizing their syntax and semantics. We continue by giving a brief overview on probability theory and probabilistic networks. Afterwards we introduce probabilistic conditional logic and illustrate reasoning based on the principle of maximum entropy. We conclude Chapter 2 with a brief overview on statistical relational learning in general and the two formalisms of Bayesian logic programs and Markov logic networks in particular. The notations introduced in Chapter 2 are used throughout this thesis and a necessary requirement for understanding the rest of this thesis. However, after Chapter 2 we depart in two different but related research strains: while in Chapters 3 and 4 we elaborate on the problem of inconsistencies in propositional probabilistic conditional logic, in Chapters 5, 6, and 7 we investigate relational probabilistic conditional logic.

In particular, Chapter 3 investigates the problem of measuring inconsistency in propositional probabilistic conditional logic. We develop a principled approach for measuring inconsistency and discuss both established approaches for measuring inconsistency from classical logic and novel approaches. In Chapter 4 we discuss the issue of restoring consistency in inconsistent knowledge bases by presenting culpability measures as a means for finding the culprits for creating inconsistencies. We continue by proposing several different approaches for restoring consistency and compare these approaches with respect to their properties and by means of examples. In Chapter 5 we start by discussing the problem of relational probabilistic reasoning and laying the syntactical and semantical foundations of a relational probabilistic conditional logic. We propose different semantics for this new logic and discuss their relationships. In Chapter 6 we discuss the problem of inductive reasoning in relational probabilistic conditional logic using several rationality postulates. Afterwards we introduce probabilistic reasoning based on the principle of maximum entropy for the relational setting and investigate the implications with respect to the different semantics. Chapter 7 develops a novel technique to enhance computational efficiency for reasoning with maximum entropy. We introduce condensed probability functions as compact representations for probability functions and discuss

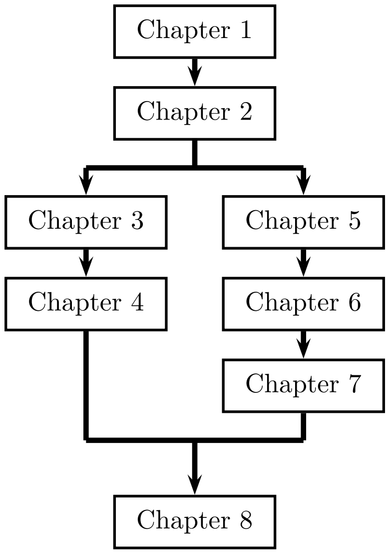

Figure 1: Outline of the thesis

reasoning based on these representations. Finally, in Chapter 8 we bring together the two different strains by summarizing the contributions of this thesis and making some final remarks. Figure 1 gives an overview on the outline of this thesis with respect to the two different logics employed for knowledge representation.

Proofs of statements that are too technical or not relevant for discussion within the main text of this thesis can be found in Appendix A. We give a hint whenever a proof can be found there. Appendix C contains further examples of the techniques developed in Chapters 3 and 4.

2 LOGICAL BACKGROUND AND PROBABILISTIC REASONING

In this chapter we give background information needed for this thesis. This comprises a basic overview on classical logic, including propositional as well as first-order logic, and probabilistic reasoning. In particular, we give a brief introduction to probability theory as far as needed for probabilistic knowledge representation and reasoning. We review existing frameworks for probabilistic reasoning in propositional settings, such as Bayesian networks and Markov nets as representatives for graphical probabilistic models and probabilistic conditional logic with inference based on the principle of maximum entropy as a non-graphical model. Finally, we discuss the field of *statistical relational learning* which investigates probabilistic models for first-order logics.

2.1 CLASSICAL LOGIC

Classical logics are the most widely used representations for formalizing reasoning. A classical logic is defined by a *syntax* that is used to model statements, and a *semantics* that is used to describe relations between statements and give meaning to statements, cf. (Beierle and Kern-Isberner, 2008). The syntax of a classical logic declares the set of symbols and some set of building rules that are used to form complex formulas of the language. The semantics of a classical logic defines *interpretations* and describes when an interpretation *satisfies* a formula. When all interpretations that satisfy some formula ϕ also satisfy another formula ψ then we say that ψ *semantically follows from* ϕ. Usually, classical logics come with a *syntactical calculus* that can be used for testing whether a formula semantically follows from another one without considering all interpretations of the language. For classical logics, these calculi are based on *deductive reasoning* and the inference rule *modus ponens* which is defined as: If "ϕ" and "ψ *follows from* ϕ" then derive "ψ" (Peirce, 1972). In order to implement deduction for a syntactical calculus properly, a set of *derivation rules* such as the *law of excluded middle* or the *De Morgan's laws* (Gabbay, 1994) can be derived.

We go on by discussing the two most important forms of classical logic: propositional logic and first-order logic.

2.1.1 *Propositional Logic*

Propositional logic (Fitch, 1952; Mendelson, 1997) is one of the first formal systems for logical reasoning and goes back to Aristotle (384–322 BC), cf.

(Bocheński, 1961). It is equivalent to Boolean algebra (Boole, 2009) and deals with logical relations between propositional statements. Even nowadays it is broadly used in fields of artificial intelligence like knowledge representation (Brachman and Levesque, 2004) and belief revision (Alchourrón *et al.*, 1985; Hansson, 1999).

Syntax

The syntax of propositional logic is defined by a propositional signature which consists of a set of atomic statements called *propositions*.

Definition 2.1 (Propositional signature, atom, proposition)**.** A *propositional signature* At is a finite set of identifiers, called *atoms* or *propositions*.

Using a propositional signature the language of propositional logic is generated using the connectives $\wedge$ (*and*), $\vee$ (*or*), and $\neg$ (*negation*).

Definition 2.2 (Propositional language)**.** Let At be a propositional signature. The *propositional language* $\mathcal{L}(\mathsf{At})$ for At is the minimal set $\mathcal{L}$ satisfying

1. $\mathsf{At} \subseteq \mathcal{L}$,
2. $\top, \bot \in \mathcal{L}$ (tautology and contradiction), and
3. for every $\phi, \psi \in \mathcal{L}$ it holds that
   a) $\phi \wedge \psi \in \mathcal{L}$ (conjunction),
   b) $\phi \vee \psi \in \mathcal{L}$ (disjunction), and
   c) $\neg\phi \in \mathcal{L}$ (negation).

The special symbols $\top$ and $\bot$ are used to denote *tautology* (a statement that is always true) and *contradiction* (a statement that is never true), respectively. An element $\phi \in \mathcal{L}(\mathsf{At})$ is called a *formula* or *sentence* of the propositional language $\mathcal{L}(\mathsf{At})$. A *literal* is an atom or a negated atom of a propositional signature At. Thus, the set of literals, denoted by $\mathsf{Lit}(\mathsf{At})$, is defined by $\mathsf{Lit}(\mathsf{At}) = \{a, \neg a \mid a \in \mathsf{At}\}$.

We abbreviate $\neg\phi \vee \psi$ by $\phi \Rightarrow \psi$ and $(\phi \Rightarrow \psi) \wedge (\psi \Rightarrow \phi)$ by $\phi \Leftrightarrow \psi$. We also use over-lining to denote the complement, i.e. it is $\overline{\phi} = \neg\phi$, and abbreviate conjunctions $\phi \wedge \psi$ simple by $\phi\psi$. For a set $\Phi = \{\phi_1, \ldots, \phi_n\} \subseteq \mathcal{L}(\mathsf{At})$ we abbreviate

$$\bigvee \Phi =_{def} \phi_1 \vee \ldots \vee \phi_2 \qquad\qquad \bigwedge \Phi =_{def} \phi_1 \wedge \ldots \wedge \phi_2 \quad .$$

If $\Phi = \emptyset$ we define

$$\bigvee \Phi =_{def} \bot \qquad\qquad \bigwedge \Phi =_{def} \top \quad .$$

The syntax of a propositional language restricts the set of words that can be phrased. The actual meaning of a word is defined by a semantics.

Semantics

Semantics to a propositional language is given by *interpretations*. An interpretation describes a specific *world* by assigning to each atom a *truth value*.

Definition 2.3 (Propositional interpretation)**.** Let At be a propositional signature. A *propositional interpretation* I on At is a function

$$I : \mathsf{At} \to \{\mathsf{true}, \mathsf{false}\} \quad .$$

Let $\mathsf{Int}(\mathsf{At})$ denote the set of all propositional interpretations for At.

We denote a propositional interpretation simply by "interpretation" if its type is clear from context. An interpretation can also be written as a *complete conjunction* enumerating all literals that are true in the given interpretation.

Example 2.1. Consider a propositional signature $\mathsf{At} = \{a, b, c\}$. The interpretation I_1 of At given by

$$I_1(a) =_{def} \mathsf{true} \qquad I_1(b) =_{def} \mathsf{false} \qquad I_1(c) =_{def} \mathsf{true}$$

can be fully described by the complete conjunction $a\overline{b}c$.

A complete conjunction is also referred to as a *possible world*. Let $\Omega(\mathsf{At})$ denote the set of all possible worlds with respect to the signature At. An interpretation I *satisfies* an atom $a \in \mathsf{At}$, denoted by $I \models^{\mathrm{P}} a$, if and only if $I(a) = \mathsf{true}$. An interpretation I *falsifies* an atom $a \in \mathsf{At}$, denoted by $I \not\models^{\mathrm{P}} a$, if and only if $I(a) = \mathsf{false}$. The satisfaction relation $\models^{\mathrm{P}}$ is extended to arbitrary sentences recursively as follows. Let $\phi, \psi \in \mathcal{L}(\mathsf{At})$ be some sentences.

- $I \models^{\mathrm{P}} \phi \vee \psi$ if and only if $I \models^{\mathrm{P}} \phi$ or $I \models^{\mathrm{P}} \psi$
- $I \models^{\mathrm{P}} \phi \wedge \psi$ if and only if $I \models^{\mathrm{P}} \phi$ and $I \models^{\mathrm{P}} \psi$
- $I \models^{\mathrm{P}} \neg\phi$ if and only if $I \not\models^{\mathrm{P}} \phi$

Furthermore, for every interpretation I it holds that $I \models^{\mathrm{P}} \top$ and $I \not\models^{\mathrm{P}} \bot$.

Definition 2.4 (Propositional model)**.** Let At be a propositional signature and $I \in \mathsf{Int}(\mathsf{At})$ an interpretation on At. I is a *propositional model* of a sentence $\phi \in \mathcal{L}(\mathsf{At})$ if and only if $I \models^{\mathrm{P}} \phi$.

In the following, we just use the term "model" to refer to a propositional model if its type is clear from context. An interpretation I is a model of a set of formula $\Phi \subseteq \mathcal{L}(\mathsf{At})$ if and only if I is a model of every formula in Φ, i.e., $I \models^{\mathrm{P}} \Phi$ if and only if for every $\phi \in \Phi$ it holds that $I \models^{\mathrm{P}} \phi$. Let $\mathsf{Mod}^{\mathrm{P}}(\Phi) \subseteq \mathsf{Int}(\mathsf{At})$ denote the set of all models of $\Phi \subseteq \mathcal{L}(\mathsf{At})$. If

Φ consists of a single element, i. e. $\Phi = \{\phi\}$, we write $\mathsf{Mod}^{\mathrm{P}}(\phi)$ instead of $\mathsf{Mod}^{\mathrm{P}}(\{\phi\})$. The satisfaction relation $\models^{\mathrm{P}}$ can also be used to describe *semantical entailment* between formulas. A set of formulas Φ_2 *semantically follows* from a set of formulas Φ_1, denoted by $\Phi_1 \models^{\mathrm{P}} \Phi_2$, if and only if $\mathsf{Mod}^{\mathrm{P}}(\Phi_1) \subseteq \mathsf{Mod}^{\mathrm{P}}(\Phi_2)$. If Φ_1 or Φ_2 consist of a single element we omit the curly brackets, e. g., we write $\phi_1 \models^{\mathrm{P}} \phi_2$ instead of $\{\phi_1\} \models^{\mathrm{P}} \{\phi_2\}$ for $\phi_1, \phi_2 \in \mathcal{L}(\mathsf{At})$. Note that if ω_I is the possible world representing the interpretation I it holds that $\omega_I \models^{\mathrm{P}} \Phi$ if and only if $I \models^{\mathrm{P}} \Phi$. Due to this property we use interpretations and possible worlds interchangeably. Two formulas ϕ and ψ are *equivalent*, denoted by $\phi \equiv^{\mathrm{P}} \psi$, if and only if $\phi \models^{\mathrm{P}} \psi$ and $\psi \models^{\mathrm{P}} \phi$, i. e., if and only if $\mathsf{Mod}^{\mathrm{P}}(\phi) = \mathsf{Mod}^{\mathrm{P}}(\psi)$. Note that it holds that $\mathsf{Mod}^{\mathrm{P}}(\top) = \mathsf{Int}(\mathsf{At})$ and $\mathsf{Mod}^{\mathrm{P}}(\bot) = \emptyset$.

If a formula ϕ is *contradictory*, i. e., it holds that $\phi \equiv^{\mathrm{P}} \bot$, then every other formula ϕ' semantically follows from ϕ, i. e. $\phi \models^{\mathrm{P}} \phi'$ for every $\phi' \in \mathcal{L}(\mathsf{At})$, as ϕ has no models. This phenomenon is referred to as *ex falso quod libet* (lat. *from falsehood follows everything*) and is one of the major drawbacks of propositional logic and classical logics in general when considering *commonsense reasoning*.

Checking whether a formula ϕ' semantically follows from a formula ϕ, i. e. whether $\phi \models^{\mathrm{P}} \phi'$ holds, is an NP-complete problem as it is equivalent to the famous satisfiability problem SAT (Garey and Johnson, 1990). However, there exist several syntactical calculi and practical algorithms for propositional logic such as the DPLL algorithm (Davis *et al.*, 1962) that work well for most real-world applications, see e. g. (Strichmann and Szeider, 2010) for recent developments.

2.1.2 *First-Order Logic*

First-order logic (Hilbert and Ackermann, 1928; Neuhaus *et al.*, 2004) extends propositional logic by introducing *relations between objects*, *functions*, and *quantification*. It is strictly more expressive than propositional logic but the problem of satisfaction is undecidable in general (Church, 1936; Turing, 1937). Nonetheless, it is widely used in e. g. automated theorem proving (Fitting, 1996) and knowledge representation (Brachman and Levesque, 2004) and is the foundation for description logics (Baader *et al.*, 2003).

Syntax

In first-order logic, an object is referred to by a single *constant*, e. g. tweety, and is a single entity describing a particular individual in the domain under discourse. A *predicate* can be used to describe either an attribute of a single object, e. g. *flies*(tweety), or a relation between two or more objects, e. g. *chases*(sylvester, tweety). A *functor* is used to describe a functional relationship from one or more objects to another object. For example, the value of *ownerOf*(tweety) refers to the unique owner of Tweety that may be equivalent to the individual described by the constant *granny*.

Definition 2.5 (First-order signature)**.** A *first-order signature* Σ is a tuple $\Sigma = (U, Pred, Func)$ with

- U is a set of constants,
- $Pred$ is a set of predicates, and
- $Func$ is a set of functors.

A first-order signature $\Sigma = (U, Pred, Func)$ is *finite* if and only if U, $Pred$, and $Func$ are finite. We use the notation v/n to denote that the arity of the predicate (or functor) v is n with $n \in \mathbb{N}$. While the set of constants U describes all known individuals in the scenario under discourse, first-order logic also allows statements concerning all or a set of undetermined objects. A *variable* is a descriptor for a specific but unknown or not determined individual in the domain under discourse. To distinguish constants from variables, we usually write the latter with a beginning upper-case letter and the former with a beginning lower-case letter and vectors of these with $\vec{\mathsf{X}}$ and $\vec{\mathsf{a}}$, respectively. Constants, variables, and functional expressions are subsumed by the notion of a *term*.

Definition 2.6 (Term)**.** Let $\Sigma = (U, Pred, Func)$ be a first-order signature and V a set of variables. The set of *terms* $\mathsf{Terms}(\Sigma, V)$ for Σ and V is the minimal set T satisfying

1. $U \subseteq T$,
2. $V \subseteq T$, and
3. for all $f/n \in Func$ and $\mathsf{t}_1, \ldots, \mathsf{t}_n \in T$ it holds that $f(\mathsf{t}_1, \ldots, \mathsf{t}_n) \in T$.

Every term describes a specific but possibly undetermined object. Using terms and predicates one can construct atomic expressions of first-order logic as follows.

Definition 2.7 (First-order atom)**.** Let Σ be a first-order signature, V a set of variables, $p/n \in Pred$ an n-ary predicate, and $\mathsf{t}_1, \ldots, \mathsf{t}_n \in \mathsf{Terms}(\Sigma, V)$. The expression $p(\mathsf{t}_1, \ldots, \mathsf{t}_n)$ is called a *first-order atom* of Σ and V.

We denote a first-order atom simply by "atom" if its type is clear from context. The important difference between a functional term $f(\mathsf{t}_1, \ldots, \mathsf{t}_n)$ and an atom $p(\mathsf{t}_1, \ldots, \mathsf{t}_n)$ is its value. While the value of a functional term describes an object, the value of an atom is either true or false and thus describes whether the specified relation holds or not.

Definition 2.8 (First-order language)**.** Let Σ be a first-order signature and V a set of variables. The *first-order language* $\mathcal{L}(\Sigma, V)$ for Σ and V is the minimal set $\mathcal{L}$ satisfying

1. for every atom $p(\mathsf{t}_1, \ldots, \mathsf{t}_n)$ of Σ and V it holds that $p(\mathsf{t}_1, \ldots, \mathsf{t}_n) \in \mathcal{L}$,

2. $\top, \bot \in \mathcal{L}$ (tautology and contradiction), and
3. for every $\phi, \psi \in \mathcal{L}$ and $\mathsf{X} \in V$ it holds that
   a) $\phi \wedge \psi \in \mathcal{L}$ (conjunction),
   b) $\phi \vee \psi \in \mathcal{L}$ (disjunction),
   c) $\neg\phi \in \mathcal{L}$ (negation),
   d) $\forall \mathsf{X} : \phi \in \mathcal{L}$ (all-quantification), and
   e) $\exists \mathsf{X} : \phi \in \mathcal{L}$ (existential-quantification).

An element $\phi \in \mathcal{L}(\Sigma, V)$ is called a *formula* of the first-order language $\mathcal{L}(\Sigma, V)$. A formula $\phi \in \mathcal{L}(\Sigma, V)$ is *ground* if and only if no variable occurs in ϕ. Let $\mathcal{L}^{\not\forall\not\exists}(\Sigma, V) \subseteq \mathcal{L}(\Sigma, V)$ denote the fragment of $\mathcal{L}(\Sigma, V)$ without quantification. As in propositional logic, cf. Section 2.1.1, a *literal* of a first-order signature Σ and a set of variables V is either an atom of Σ and V or a negated atom. Let $\mathsf{Lit}(\Sigma, V)$ denote the set of all literals of Σ and V.

A variable $\mathsf{X} \in V$ can appear in a formula $\phi \in \mathcal{L}(\Sigma, V)$ *bound* and/or *free*. If X appears in sub-formulas of the form $\forall \mathsf{X} : \phi'$ or $\exists \mathsf{X} : \phi'$ of ϕ then X is *bound* in ϕ. If X appears in ϕ without a corresponding quantifier then X is *free* in ϕ. In order to simplify matters, we assume that all variables in a formula ϕ are *either* bound *or* free in ϕ, but not both. This can be achieved easily by appropriately renaming bound occurrences of a variable. A formula $\phi \in \mathcal{L}(\Sigma, V)$ is called a *sentence* if and only if all variables in ϕ are bound in ϕ. In particular, every ground formula is a sentence.

As for propositional logic we abbreviate

$$
\begin{aligned}
\phi \Rightarrow \psi &=_{def} \neg\phi \vee \psi & \qquad \phi \Leftrightarrow \psi &=_{def} (\phi \Rightarrow \psi) \wedge (\psi \Rightarrow \phi) \\
\overline{\phi} &=_{def} \neg\phi & \phi\psi &=_{def} \phi \wedge \psi \\
\bigvee \Phi &=_{def} \phi_1 \vee \ldots \vee \phi_n & \bigvee \emptyset &=_{def} \bot \\
\bigwedge \Phi &=_{def} \phi_1 \wedge \ldots \wedge \phi_n & \bigwedge \emptyset &=_{def} \top
\end{aligned}
$$

for $\phi, \psi \in \mathcal{L}(\Sigma, V)$ and $\Phi = \{\phi_1, \ldots, \phi_n\} \subseteq \mathcal{L}(\Sigma, V)$. If $\{\mathsf{X}_1, \ldots, \mathsf{X}_n\}$ is the set of free (and not bound) variables of a formula ϕ we also write ϕ as $\phi(\vec{\mathsf{X}})$ with $\vec{\mathsf{X}} = (\mathsf{X}_1, \ldots, \mathsf{X}_n)$ (we assume some arbitrary total order on the free and not bound variables of a formula as given). If $\vec{\mathsf{a}} = (\mathsf{a}_1, \ldots, \mathsf{a}_n)$ is a vector of constants of the same length we denote by $\phi(\vec{\mathsf{a}})$ the formula ϕ' that is the same as ϕ but every free occurrence of X_i is replaced by a_i for $i = 1, \ldots, n$.

An important syntactic restriction of first-order logic is *Horn logic* (Horn, 1951). A *Horn clause* is a disjunction of negated atoms and at most one not-negated atom, i.e., a formula of the form $p \vee \neg q_1 \vee \ldots \vee \neg q_n$ with atoms $p, q_1, \ldots, q_n$. A Horn clause of this form can also be written as an implication $p \Leftarrow q_1 \wedge \ldots \wedge q_n$ which explains the interest in these clauses as they can be used to model rule-like knowledge. In fact, Horn logic is the foundation for logic programming languages such as Prolog (Covington *et al.*, 1996) and answer set programming (Gelfond and Lifschitz, 1991), see also (Gelfond and Leone, 2002). For a Horn clause ϕ of the form $\phi = p \Leftarrow q_1 \wedge \ldots \wedge q_n$ we abbreviate $\mathsf{head}(\phi) = p$ (the *head* or *conclusion* of the clause) and $\mathsf{body}(\phi) = \{q_1, \ldots, q_n\}$ (the *body* or *premise* of the clause).

Semantics

Semantics is given to a first-order language $\mathcal{L}(\Sigma, V)$ by means of *interpretations* and *variable assignments*.

Definition 2.9 (First-order interpretation)**.** Let $\Sigma = (U, Pred, Func)$ be a first-order signature. A *first-order interpretation* I on Σ is a tuple $I = (U_I, f_I^U, Pred_I, Func_I)$ with

1. a non-empty set of objects U_I (the *universe*),
2. a function $f_I^U : U \to U_I$,
3. a set $Pred_I$ of relations $Pred_I = \{p_I \subseteq U_I^n \mid p/n \in Pred\}$, and
4. a set $Func_I$ of functions $Func_I = \{f_I : U_I^n \to U_I \mid f/n \in Func\}$.

An element $\mathsf{c}_I \in U_I$ is called an interpretation of $\mathsf{c} \in U$ if and only if $\mathsf{c}_I = f_I^U(\mathsf{c})$. Similarly, an element $p_I \in Pred_I$ ($f_I \in Func_I$) is also called an interpretation of its corresponding $p \in Pred$ ($f \in Func$). Let $\mathsf{Int}(\Sigma)$ denote the set of all first-order interpretations for Σ.

We denote a first-order *interpretation* simply by "interpretation" if its type is clear from context.

Definition 2.10 (Variable assignment)**.** Let $I = (U_I, f_I^U, Pred_I, Func_I)$ be a first-order interpretation on $\Sigma = (U, Pred, Func)$ and V a set of variables. A *variable assignment* VA on V is a function $VA : V \to U_I$.

A variable assignment VA can be extended to terms with respect to an interpretation $I = (U_I, f_I^U, Pred_I, Func_I)$ as follows. If $\mathsf{c} \in \mathsf{Terms}(\Sigma, V)$ is a constant, i. e. $\mathsf{c} \in U$, then $VA(\mathsf{c}) =_{def} f_I^U(\mathsf{c})$. If $f(\mathsf{t}_1, \ldots, \mathsf{t}_n) \in \mathsf{Terms}(\Sigma, V)$ with $f \in Func$ then $VA(f(\mathsf{t}_1, \ldots, \mathsf{t}_n)) =_{def} f_I(VA(\mathsf{t}_1), \ldots, VA(\mathsf{t}_n))$.

An interpretation $I = (U_I, f_I^U, Pred_I, Func_I)$ on Σ together with a variable assignment VA *satisfies* an atom $p(\mathsf{t}_1, \ldots, \mathsf{t}_n)$ of Σ and V, denoted by $I, VA \models^{\mathrm{F}} p(\mathsf{t}_1, \ldots, \mathsf{t}_n)$, if and only if $(VA(\mathsf{t}_1), \ldots, VA(\mathsf{t}_n)) \in p_I$ where $p_I \in Pred_I$ is the interpretation of p in I. The interpretation I and the variable assignment VA *falsify* $p(\mathsf{t}_1, \ldots, \mathsf{t}_n)$, denoted by $I, VA \not\models^{\mathrm{F}} p(\mathsf{t}_1, \ldots, \mathsf{t}_n)$, if and only if $(VA(\mathsf{t}_1), \ldots, VA(\mathsf{t}_n)) \notin p_I$. The satisfaction relation $\models^{\mathrm{F}}$ is extended to arbitrary formulas recursively as follows. Let $\phi, \psi \in \mathcal{L}(\Sigma, V)$ be some formulas.

- $I, VA \models^{\mathrm{F}} \phi \vee \psi$ if and only if $I, VA \models^{\mathrm{F}} \phi$ or $I, VA \models^{\mathrm{F}} \psi$
- $I, VA \models^{\mathrm{F}} \phi \wedge \psi$ if and only if $I, VA \models^{\mathrm{F}} \phi$ and $I, VA \models^{\mathrm{F}} \psi$

- $I, VA \models^{\mathrm{F}} \neg\phi$ if and only if $I, VA \not\models^{\mathrm{F}} \phi$
- $I, VA \models^{\mathrm{F}} \forall X : \phi$ if and only if for every variable assignment VA' that is the same as VA but possibly $VA'(X) \neq VA(X)$ it holds that $I, VA' \models^{\mathrm{F}} \phi$
- $I, VA \models^{\mathrm{F}} \exists X : \phi$ if and only if for some variable assignment VA' that is the same as VA but possibly $VA'(X) \neq VA(X)$ it holds that $I, VA' \models^{\mathrm{F}} \phi$

Furthermore, for every interpretation I and every variable assignment VA it holds that $I, VA \models^{\mathrm{F}} \top$ and $I, VA \not\models^{\mathrm{P}} \bot$. An interpretation I satisfies a formula $\phi \in \mathcal{L}(\Sigma, V)$, denoted by $I \models^{\mathrm{F}} \phi$, if and only if for every variable assignment VA on V it holds that $I, VA \models^{\mathrm{F}} \phi$.

Definition 2.11 (First-order model)**.** Let $\Sigma = (U, \mathit{Pred}, \mathit{Func})$ be a first-order signature, V be some set of variables, and I an interpretation on Σ. I is a *first-order model* of a formula $F \in \mathcal{L}(\Sigma, V)$ if and only if $I \models^{\mathrm{F}} F$.

A first-order model I is simply referred to as a model if its type is clear from context. An interpretation I is a model of a set of formulas $\Phi \subseteq \mathcal{L}(\Sigma, V)$ if and only if I is a model of every formula in Φ, i. e., $I \models^{\mathrm{F}} \Phi$ if and only if for every $\phi \in \Phi$ it holds that $I \models^{\mathrm{F}} \phi$. Let $\mathsf{Mod}^{\mathrm{F}}(\Phi)$ denote the set of all models of $\Phi \subseteq \mathcal{L}(\Sigma, V)$. If Φ consists of a single element, i. e. $\Phi = \{\phi\}$, we write $\mathsf{Mod}^{\mathrm{F}}(\phi)$ instead of $\mathsf{Mod}^{\mathrm{F}}(\{\phi\})$. As for propositional logic we define *semantical entailment* between sets of formulas Φ_1, Φ_2 via $\Phi_1 \models^{\mathrm{F}} \Phi_2$ if and only if $\mathsf{Mod}^{\mathrm{F}}(\Phi_1) \subseteq \mathsf{Mod}^{\mathrm{F}}(\Phi_2)$. Two formulas ϕ_1 and ϕ_2 are *equivalent*, denoted by $\phi_1 \equiv^{\mathrm{F}} \phi_2$, if and only if $\phi_1 \models^{\mathrm{F}} \phi_2$ and $\phi_2 \models^{\mathrm{F}} \phi_1$, i. e., if $\mathsf{Mod}^{\mathrm{F}}(\phi_1) = \mathsf{Mod}^{\mathrm{F}}(\phi_2)$. Note that $\mathsf{Mod}^{\mathrm{F}}(\top) = \mathsf{Int}(\Sigma)$ and $\mathsf{Mod}^{\mathrm{F}}(\bot) = \emptyset$.

Replacements

We need some further notation that is used throughout this thesis.

Definition 2.12 (Replacement)**.** Let $\Sigma = (U, \mathit{Pred}, \mathit{Func})$ be a first-order signature and V a set of variables. A function $\theta : U \cup V \to U \cup V$ with $\theta(x) \neq x$ only for a finite number of $x \in U \cup V$ is called a *replacement* for Σ and V. Let $\Gamma(\Sigma, V)$ denote the set of all replacements for Σ and V.

Note that replacements are a similar concept like *substitutions*, see e. g. (Russell and Norvig, 2009). The main difference between substitutions and replacements is that the former do not allow the substitution of constants while the latter do not allow for the insertion of functional terms. Replacements are extended to first-order formulas in a straightforward fashion, i. e., if $\phi \in \mathcal{L}(\Sigma, V)$ and $\theta \in \Gamma(\Sigma, V)$ then $\theta(\phi)$ is the same as ϕ except that every occurrence of $x \in U$ is substituted by $\theta(x)$ and every *free* occurrence of $x \in V$ is substituted by $\theta(x)$. Let $\mathsf{Im}\ f$ denote the image of a function f. Then a replacement θ is called a *grounding replacement* if and only if $\mathsf{Im}\ \theta \subseteq U$ and for every $\mathsf{a} \in U$ it holds that $\theta(\mathsf{a}) = \mathsf{a}$. Let $\Gamma^{\text{gnd}}(\Sigma, V) \subseteq \Gamma(\Sigma, V)$ be the set of all grounding replacements. Note that for a grounding replacement $\theta \in \Gamma^{\text{gnd}}(\Sigma, V)$ and a formula ϕ that contains a quantifier the formula $\theta(\phi)$ is not ground. However, for every formula ϕ and grounding replacement $\theta \in \Gamma^{\text{gnd}}(\Sigma, V)$ it follows that $\theta(\phi)$ is a sentence.

For an inline definition of a replacement we use the operator $[\cdot]$. If $x_1, \ldots, x_n, y_1, \ldots, y_n \in U \cup V$ with $x_i \neq x_j$ for $i, j = 1, \ldots, n$ and $i \neq j$ and $\phi \in \mathcal{L}(\Sigma, V)$ then the expression $\phi[y_1/x_1, \ldots, y_n/x_n]$ is defined via

$$\phi[y_1/x_1, \ldots, y_n/x_n] =_{def} \theta(\phi)$$

with $\theta \in \Gamma(\Sigma, V)$ and $\theta(x) =_{def} x$ for all $x \in U \cup V$ except $\theta(x_i) =_{def} y_i$ for $i = 1, \ldots, n$. If $x_i \neq x_j$ and $y_i \neq y_j$ for $i, j = 1, \ldots, n$ and $i \neq j$ and also $x_i \neq y_j$ for $i, j = 1, \ldots, n$ then we abbreviate further

$$\phi[y_1 \leftrightarrow x_1, \ldots, y_n \leftrightarrow x_n] =_{def} \phi[y_1/x_1, x_1/y_1, \ldots, x_n/y_n, y_n/x_n] \quad .$$

For a formula $\phi \in \mathcal{L}(\Sigma, V)$ let $\mathsf{Const}(\phi)$ denote the set of constants appearing in ϕ. For a set Φ of first-order formulas define $\mathsf{Const}(\Phi)$ to be the union of $\mathsf{Const}(\phi)$ for $\phi \in \Phi$. For a formula $\phi \in \mathcal{L}^{\not\forall\not\exists}(\Sigma, V)$ that contains no quantification and a set $D \subseteq U$ of constants with $\mathsf{Const}(\phi) \subseteq D$ we define the *set of ground instances* $\mathsf{gnd}_D(\phi)$ of ϕ with respect to D via

$$\mathsf{gnd}_D(\phi) =_{def} \{\theta(\phi) \mid \theta \in \Gamma^{\text{gnd}}(\Sigma, V) \text{ and } \mathsf{Im}\ \theta \subseteq D\} \quad .$$

That is, $\mathsf{gnd}_D(\phi)$ contains all sentences ϕ' that arise from ϕ by substituting every variable with some constant in D.

Herbrand Interpretations

In this thesis we are mainly working with a special kind of interpretations for first-order languages, the so-called *Herbrand interpretations*. With Herbrand interpretations the semantics is handled in a very simple manner as constants are interpreted by themselves. So Herbrand interpretations interpret the language on an almost syntactical level without the need of a semantical overload and hence are the first-order equivalent of propositional *possible worlds*, cf. page 11. If the first-order signature under discourse contains no functors the use of Herbrand interpretations is more intuitive than the use of ordinary first-order interpretations. The formalism of Herbrand interpretation is based on the notion of the *Herbrand base*.

Definition 2.13 (Herbrand base)**.** Let Σ be a first-order signature. The *Herbrand base* $\mathsf{At}(\Sigma)$ is the set of all ground atoms of $\mathcal{L}(\Sigma, \emptyset)$.

Let $\mathfrak{P}(S)$ denote the power set of a set S.

Definition 2.14 (Herbrand interpretation)**.** A *Herbrand interpretation* ω for a first-order signature Σ is any set $\omega \subseteq \mathsf{At}(\Sigma)$. Let $\Omega(\Sigma)$ denote the set of all Herbrand interpretations, i. e. $\Omega(\Sigma) =_{def} \mathfrak{P}(\mathsf{At}(\Sigma))$.

A Herbrand interpretation ω satisfies a ground atom $a \in \mathsf{At}(\Sigma)$, also written as $\omega \models^{\mathrm{F}} a$, if and only if $a \in \omega$. Everything said before on the relation $\models^{\mathrm{F}}$ (that is defined for general first-order interpretations) applies in the same way to Herbrand interpretations.

For the rest of this thesis let $\mathcal{L}$ denote any of $\mathcal{L}(\mathsf{At})$ or $\mathcal{L}(\Sigma, V)$ for some appropriate signature.

2.1.3 *Beyond Classical Logics*

A major drawback of classical logics for their use in commonsense reasoning is their monotonicity. Monotonicity of a logic ensures that derivations prevail no matter what new information may bring. Classical formulas are strict in the sense that no exceptions are allowed.

Example 2.2. Consider the set of first-order sentences Φ_w given by

$$\begin{aligned}\Phi_w = \{ \quad & \mathit{bird}(\mathsf{tweety}),\\ & \forall \mathsf{X} : \mathit{bird}(\mathsf{X}) \Rightarrow \neg \mathit{speaks}(\mathsf{X}) \quad \} \quad .\end{aligned}$$

The sentences in Φ_w describe a scenario where we have some general beliefs on birds, i. e., birds do not speak, as well as the observation of an actual bird tweety. From Φ_w we can infer that *tweety* does not speak, i. e. $\Phi_w \models^{\mathrm{F}} \neg \mathit{speak}(\mathsf{tweety})$. However, suppose that tweety is an exceptional bird that can indeed speak. Considering the extended knowledge base Φ'_w given by

$$\begin{aligned}\Phi'_w = \{ \quad & \mathit{bird}(\mathsf{tweety}),\\ & \mathit{speaks}(\mathsf{tweety}),\\ & \forall \mathsf{X} : \mathit{bird}(\mathsf{X}) \Rightarrow \neg \mathit{speaks}(\mathsf{X}) \quad \} \quad .\end{aligned}$$

From Φ'_w we can still infer $\neg \mathit{speaks}(\mathsf{tweety})$ because Φ'_w is, as a matter of fact, inconsistent, so it holds that $\Phi'_w \models^{\mathrm{F}} \phi$ for any formula ϕ of the corresponding language.

There are many non-classical frameworks that try to capture this and other properties of commonsense reasoning, both *qualitative* ones and *quantitative* ones. Among the qualitative frameworks the most widespread ones are paraconsistent logics (Belnap, 1976, 1977; Béziau *et al.*, 2007), defeasible logics (Nute, 1994), default logics (Reiter, 1980; Antoniou, 1999), argumentation frameworks (Bench-Capon and Dunne, 2007; Rahwan and Simari, 2009), and answer set programming (Gelfond and Lifschitz, 1991; Gelfond and Leone, 2002). The most common quantitative frameworks are probability theory (Pearl, 1998; Jaynes, 2003), fuzzy logic (Gerla, 2001), and Dempster-Shafer theory (Shafer, 1976), see also (Paris, 1994). In this thesis we focus on logics that base on probability theory.

2.2 PROBABILITY THEORY AND PROBABILISTIC NETWORKS

Probability theory is the oldest formalism for dealing with uncertainty in a quantified manner, the first recorded publication being "*De ratiociniis in ludo aleae*" ("On Reasoning in Games of Chance") by Christiaan Huygens from the year 1657 (Gullberg, 1997). However, the influential works by Kolmogorov (Kolmogorov, 1933) have mainly shaped todays formal treatment of probability and his axioms of probability are the foundation of reasoning with probabilities in artificial intelligence and knowledge representation (Paris, 1994; Pearl, 1998; Jaynes, 2003). In the following we give a short introduction to the basic concepts of probability theory and illustrate its application in knowledge representation using the formalisms of Bayesian Networks and Markov Nets.

2.2.1 *Foundations of Probability Theory*

In this thesis we are dealing with *discrete probability functions*, i. e., probability functions that are defined on countable *sample spaces*. In general, a sample space is a set of mutual exclusive events whose uncertain occurrence is to be measured by the probability function. For example, the set $\mathcal{X}_{\text{dice}} = \{1, 2, 3, 4, 5, 6\}$ can be regarded as the set of possible events occurring as outcomes when throwing a dice. A probability function on $\mathcal{X}$ then assigns to each of the sides of the dice the "probability" that this side turns up when throwing the dice.

Definition 2.15 (Probability function)**.** Let $\mathcal{X}$ be some countable set. A *probability function* P on $\mathcal{X}$ is a function $P : \mathfrak{P}(\mathcal{X}) \to [0, 1]$ that satisfies

1. $P(\mathcal{X}) = 1$ and
2. $P(X_1 \cup X_2) = P(X_1) + P(X_2)$ for $X_1, X_2 \subseteq \mathcal{X}$ and $X_1 \cap X_2 = \emptyset$.

For notational convenience we write $P(x)$ instead of $P(\{x\})$ for $x \in \mathcal{X}$. Conditions 1. and 2. in Definition 2.15 are also referred to as *Kolmogorov axioms of probability* (Jaynes, 2003). A probability function P_0 that assigns the same probability to each $x \in \mathcal{X}$ is called a *uniform probability function*, i. e., it holds that $P_0(x) = 1/|\mathcal{X}|$ for each $x \in \mathcal{X}$ (for finite $\mathcal{X}$). For example, considering the probability function $P_{\text{fair\_dice}}$ of a fair dice, the probability of each $x \in \mathcal{X}_{dice}$ for the example above is defined as $P_{\text{fair\_dice}}(x) =_{def} 1/6$, i. e. $P_{\text{fair\_dice}}$ is a uniform probability function on $\mathcal{X}_{dice}$.

Some immediate observations on properties of a probability function P can be made.

Proposition 2.1. *Let $X_1, X_2 \subseteq \mathcal{X}$. Then it holds that*

1. $P(X_1) = \sum_{x \in X_1} P(x)$,
2. $P(X_1 \cup X_2) = P(X_1) + P(X_2) - P(X_1 \cap X_2)$, *and*
3. $P(\mathcal{X} \setminus X_1) = 1 - P(X_1)$.

The proof of Proposition 2.1 can be found in Appendix A on page 225. Due to property 1.) of the above proposition we often write $P : \mathcal{X} \to [0,1]$ instead of $P : \mathfrak{P}(\mathcal{X}) \to [0,1]$ as only the probabilities for each $x \in \mathcal{X}$ have to be defined.

In knowledge representation and reasoning (Paris, 1994; Pearl, 1998) we are mostly interested in probability functions on *possible worlds* of an underlying logical language $\mathcal{L}$. For what follows we assume a propositional language $\mathcal{L}(\mathsf{At})$ for some propositional signature At and consider probability functions of the form $P : \Omega(\mathsf{At}) \to [0,1]$. Let $\mathcal{P}^{\mathrm{P}}(\mathsf{At})$ denote the set of all these probability functions.

Example 2.3. Consider $\mathsf{At} = \{a, b, c\}$. The set of all interpretations $\mathsf{Int}(\mathsf{At})$ of At is given by $\mathsf{Int}(\mathsf{At}) = \{I_1, \ldots, I_8\}$ with

$$
\begin{array}{lll}
I_1(a) = \mathsf{true} & I_1(b) = \mathsf{true} & I_1(c) = \mathsf{true} \\
I_2(a) = \mathsf{true} & I_2(b) = \mathsf{true} & I_2(c) = \mathsf{false} \\
I_3(a) = \mathsf{true} & I_3(b) = \mathsf{false} & I_3(c) = \mathsf{true} \\
I_4(a) = \mathsf{true} & I_4(b) = \mathsf{false} & I_4(c) = \mathsf{false} \\
I_5(a) = \mathsf{false} & I_5(b) = \mathsf{true} & I_5(c) = \mathsf{true} \\
I_6(a) = \mathsf{false} & I_6(b) = \mathsf{true} & I_6(c) = \mathsf{false} \\
I_7(a) = \mathsf{false} & I_7(b) = \mathsf{false} & I_7(c) = \mathsf{true} \\
I_8(a) = \mathsf{false} & I_8(b) = \mathsf{false} & I_8(c) = \mathsf{false}
\end{array}
$$

or equivalently with the set $\Omega(\mathsf{At}) = \{\omega_{I_1}, \ldots, \omega_{I_8}\}$ of possible worlds with

$$
\begin{array}{llll}
\omega_{I_1} = abc & \omega_{I_2} = ab\overline{c} & \omega_{I_3} = a\overline{b}c & \omega_{I_4} = a\overline{b}\overline{c} \\
\omega_{I_5} = \overline{a}bc & \omega_{I_6} = \overline{a}b\overline{c} & \omega_{I_7} = \overline{a}\overline{b}c & \omega_{I_8} = \overline{a}\overline{b}\overline{c} \quad .
\end{array}
$$

A probability function $P_1 : \Omega(\mathsf{At}) \to [0,1]$ on $\Omega(\mathsf{At})$ can be given via

$$
\begin{array}{ll}
P_1(\omega_{I_1}) =_{def} 0.1 & P_1(\omega_{I_5}) =_{def} 0.15 \\
P_1(\omega_{I_2}) =_{def} 0.2 & P_1(\omega_{I_6}) =_{def} 0.25 \\
P_1(\omega_{I_3}) =_{def} 0.2 & P_1(\omega_{I_7}) =_{def} 0.02 \\
P_1(\omega_{I_4}) =_{def} 0.05 & P_1(\omega_{I_8}) =_{def} 0.03 \quad .
\end{array}
$$

Due to property 1.) of Proposition 2.1 the above assignments completely describe a probability function P_1 as e. g. $P_1(\{\omega_{I_1}\omega_{I_2}\}) = P_1(\omega_{I_1}) + P_1(\omega_{I_2}) = 0.3$.

A probability function P on a set of possible worlds $\Omega(\mathsf{At})$ can be extended to the whole language $\mathcal{L}(\mathsf{At})$ via

$$P(\phi) =_{def} \sum_{\omega \models^{\mathrm{P}} \phi,\ \omega \in \Omega(\mathsf{At})} P(\omega) \tag{2.1}$$

for $\phi \in \mathcal{L}(\mathsf{At})$ (Beierle and Kern-Isberner, 2008). This means, that the probability of a formula ϕ is defined to be the sum of the probabilities of the models of ϕ. If $\Phi \subseteq \mathcal{L}(\mathsf{At})$ is a finite set of formulas with $\Phi = \{\phi_1, \ldots, \phi_n\}$ we abbreviate $P(\Phi) =_{def} P(\phi_1 \wedge \ldots \wedge \phi_n)$.

Example 2.4. We continue Example 2.3 and consider the formula $\phi_1 =_{def} a \wedge b$. The probability of ϕ_1 given P_1 can be computed as

$$P_1(\phi_1) = P_1(\omega_{I_1}) + P_1(\omega_{I_2}) = 0.3$$

as $\omega_{I_1} \models^{\mathrm{P}} \phi_1$ and $\omega_{I_2} \models^{\mathrm{P}} \phi_1$ and these are the only models of ϕ_1.

Some simple properties of probability functions on propositional languages are as follows.

Proposition 2.2. *Let P be a probability function on $\mathcal{L}(\mathsf{At})$ and $\phi, \psi \in \mathcal{L}(\mathsf{At})$.*

1. *If $\phi \models^{P} \perp$ then $P(\phi) = 0$.*
2. *If $\top \models^{P} \phi$ then $P(\phi) = 1$.*
3. *If $\phi \equiv^{P} \psi$ then $P(\phi) = P(\psi)$.*
4. *If $\phi \wedge \psi \models^{P} \perp$ then $P(\phi \vee \psi) = P(\phi) + P(\psi)$.*
5. *It holds that $P(\neg\phi) = 1 - P(\psi)$.*
6. *If $\phi \models^{P} \psi$ then $P(\phi) \leq P(\psi)$.*

The proof of Proposition 2.2 can be found in Appendix A on page 226.

One of the most interesting properties of relationships between propositions in a probabilistic framework is *probabilistic independence*.

Definition 2.16 (Probabilistic independence)**.** Let P be probability function on $\mathcal{L}(\mathsf{At})$ and $A_1, A_2 \subseteq \mathsf{At}$ sets of atoms. Then A_1 and A_2 are *probabilistically independent* with respect to P, denoted by $A_1 \perp\!\!\!\perp_{\mathrm{P}} A_2$, if and only if for all $\omega_1 \in \Omega(A_1)$, $\omega_2 \in \Omega(A_2)$ it holds that $P(\omega_1 \wedge \omega_2) = P(\omega_1)P(\omega_2)$.

Probabilistic independence allows, among other things, for a compact representation of a probability function as only *marginal distributions* of probabilistic independent sets of atoms have to be stored instead of the whole function. Note that $\perp\!\!\!\perp_{\mathrm{P}}$ is a symmetric relation, i. e., it holds that $A_1 \perp\!\!\!\perp_{\mathrm{P}} A_2$ if and only if $A_2 \perp\!\!\!\perp_{\mathrm{P}} A_1$.

An important concept of probability theory and especially of knowledge representation is the *conditional probability*.

Definition 2.17 (Conditional probability)**.** Let P be a probability function on $\mathcal{L}(\mathsf{At})$ and $\phi, \psi \in \mathcal{L}(\mathsf{At})$ with $P(\phi) > 0$. Then the *conditional probability* of ψ given ϕ, written as $P(\psi \mid \phi)$, is defined as

$$P(\psi \mid \phi) =_{def} \frac{P(\phi \wedge \psi)}{P(\phi)}$$

The expression $P(\psi \mid \phi)$ describes the probability that ψ is true when ϕ is already known to be true. If $\Phi_1, \Phi_2 \subseteq \mathcal{L}(\mathsf{At})$ are finite sets of formulas with $\Phi_1 = \{\phi_1^1, \ldots, \phi_n^1\}$ and $\Phi_2 = \{\phi_1^2, \ldots, \phi_m^2\}$ we abbreviate

$$P(\Phi_1 \mid \Phi_2) =_{def} P(\phi_1^1 \wedge \ldots \wedge \phi_n^1 \mid \phi_1^2 \wedge \ldots \wedge \phi_m^2) \quad .$$

If $\Phi_2 = \emptyset$ we define $P(\Phi_1 \mid \Phi_2) = P(\Phi_1)$. For conditional probabilities the concept of probabilistic independence can be generalized as follows.

Definition 2.18 (Conditional independence)**.** Let P be probability function on $\mathcal{L}(\mathsf{At})$ and $A_1, A_2, A_3 \subseteq \mathsf{At}$ sets of atoms. Then A_1 and A_2 are *conditionally independent* given A_3 with respect to P, denoted by $A_1 \perp\!\!\!\perp_P A_2 \mid A_3$, if and only if for all $\omega_1 \in \Omega(A_1)$, $\omega_2 \in \Omega(A_2)$, $\omega_3 \in \Omega(A_3)$ it holds that $P(\omega_1 \wedge \omega_2 \mid \omega_3) = P(\omega_1 \mid \omega_3)P(\omega_2 \mid \omega_3)$.

Note that standard probabilistic independence is equivalent to conditional independence given the empty set. Similar to probabilistic independence conditional independence allows for a *decomposition* of a probability function into manageable *conditional probability functions*, cf. (Pearl, 1998). Conditional independence is symmetric in its first two arguments, i. e., it holds that $A_1 \perp\!\!\!\perp_P A_2 \mid A_3$ if and only if $A_2 \perp\!\!\!\perp_P A_1 \mid A_3$.

Conditional probabilities are often used in knowledge representation to describe causal or diagnostic dependencies (Pearl, 1998). Well-known frameworks which rely heavily on the notions of conditional probability and conditional independence are Bayesian Networks and Markov Nets which we present now.

2.2.2 *Bayesian Networks*

A Bayesian network (Pearl, 1998) is a convenient method to represent a special type of probability functions in a compact way. Bayesian networks have been used for the representation of and reasoning with uncertain beliefs for quite some time as they allow for efficient inference methods (Lauritzen and Spiegelhalter, 1988). Structurally, a Bayesian network is a directed graph on propositions that models probabilistic (in-)dependencies of some underlying probability function.

For a directed graph (V, E) and a node $v \in V$ let $\mathsf{pa}(v)$ denote the set of *parents* of v, i. e nodes $v' \in V$ such that there is a directed edge $(v', v) \in E$. A *directed path* p from a node $v_1 \in V$ to a node $v_n \in V$ is a sequence of nodes $p = (v_1, \ldots, v_n)$ such that $(v_i, v_{i+1}) \in E$ for $i = 1, \ldots, n-1$. Furthermore, let $\mathsf{nd}(v)$ denote the set of non-descendants of v, i. e. nodes $v' \in V \setminus \mathsf{pa}(v)$ such that there is no directed path from v to v'.

Definition 2.19 (Bayesian network)**.** Let At be a set of propositions. A Bayesian network BN for At is a tuple $BN = (\mathsf{At}, E, P)$ such that (At, E) is a directed acyclic graph and P is a probability function that obeys

$$\{a\} \perp\!\!\!\perp_{\mathrm{P}} \mathsf{nd}(a) \mid \mathsf{pa}(a) \quad \text{(for every } a \in \mathsf{At}) \quad . \tag{2.2}$$

Condition (2.2) is also called the *local Markov property*. Due to this property, the probability function P can be decomposed into conditional probability functions for each node $a \in \mathsf{At}$. That is, for $a \in \mathsf{At}$ with $\mathsf{pa}(a) = \{b_1, \ldots, b_n\}$ it is sufficient to store the conditional probabilities $P(a \mid \dot{b}_1 \wedge \ldots \wedge \dot{b}_n)$ for each $\dot{b}_i \in \{b_i, \overline{b}_i\}$ for $i = 1, \ldots, n$. Then, the probability of a possible world $\omega \in \Omega(\mathsf{At})$ can be computed via

$$P(\omega) = \prod_{l \in \mathsf{Lit}(\mathsf{At}),\ \omega \models^{\mathrm{P}} l} P(l \mid \bigwedge\{l' \in \mathsf{Lit}(\mathsf{At}) \mid \omega \models^{\mathrm{P}} l' \text{ and } l' \in \mathsf{pa}(l) \text{ or } \overline{l'} \in \mathsf{pa}(l)\}) \tag{2.3}$$

due to the conditional independence of each atom to their non-descendants given their parents, cf. (Pearl, 1998). Furthermore, the computation of a formula's probability can be simplified as, in general, not the whole Bayesian network has to be considered but only a small portion, see (Lauritzen and Spiegelhalter, 1988; Pearl, 1998) for details.

Example 2.5. We adapt an example on medical diagnosis from (Beierle and Kern-Isberner, 2008), see also (Pearl, 1998). Consider the propositions $\mathsf{At} = \{a, b, c, d, e\}$ with the informal interpretations *cancer* (a), *elevated serum calcium level* (b), *brain tumor* (c), *coma* (d), and *headache* (e). Let $BN_{\mathrm{med}} = (\mathsf{At}, E, P)$ with (At, E) be the Bayesian network depicted in Figure 2. It follows that P has to adhere to the conditional independence $\{b\} \perp\!\!\!\perp_{\mathrm{P}} \{c\} \mid \{a\}$ (among others). Moreover, due to Equation (2.3) the probability of a possible world such as $ab\overline{c}\overline{d}e \in \Omega(\mathsf{At})$ can be written as

$$P(ab\overline{c}\overline{d}e) = P(e \mid \overline{c}) \cdot P(\overline{d} \mid b\overline{c}) \cdot P(\overline{c} \mid a) \cdot P(b \mid a) \cdot P(a) \quad .$$

Therefore P can be completely described by e. g. the following assignments[1]:

$$\begin{array}{ll}
P(a) =_{def} 0.20 & \\
P(b \mid a) =_{def} 0.80 & P(b \mid \overline{a}) =_{def} 0.20 \\
P(c \mid a) =_{def} 0.20 & P(c \mid \overline{a}) =_{def} 0.05 \\
P(e \mid c) =_{def} 0.80 & P(e \mid \overline{c}) =_{def} 0.60
\end{array}$$

1 The numbers have been arbitrarily chosen and may not describe the real world.

$$P(d \mid b \wedge c) =_{def} 0.80 \qquad P(d \mid b \wedge \overline{c}) =_{def} 0.90$$
$$P(d \mid \overline{b} \wedge c) =_{def} 0.70 \qquad P(d \mid \overline{b} \wedge \overline{c}) =_{def} 0.05$$

Note that the probabilities of negated variables derive from the above equations via e. g. $P(\overline{e} \mid c) = 1 - P(e \mid c)$. By only defining the above conditional probabilities the function P can be compactly stored. While by exploiting Equation (2.3) one only needs to specify the above eleven assignments, a full description of P needs $|\Omega(\mathsf{At})| = 32$ assignments.

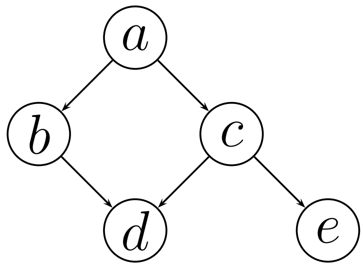

Figure 2: The graph (At, E) from Example 2.5

2.2.3 *Markov Nets*

While Bayesian networks employ directed graphs for modeling conditional (in-)dependencies, Markov nets (Pearl, 1998) build on undirected graphs. The central notion for defining a Markov net is the *graph separation*. Remember that for an undirected graph (V, E) a *path* p from a node $v_1 \in V$ to a node $v_n \in V$ is a sequence of nodes $p = (v_1, \ldots, v_n)$ such that $\{v_i, v_{i+1}\} \in E$ for $i = 1, \ldots, n-1$. If $p = (v_1, \ldots, v_n)$ is a path let $p^\dagger =_{def} \{v_1, \ldots, v_n\}$ denote the set of all nodes contained in the path.

Definition 2.20 (Graph separation)**.** Let $\mathcal{G} = (V, E)$ be an undirected graph and let $A_1, A_2, A_3 \subseteq V$ be pairwise disjoint. Then A_3 *separates* A_1 and A_2 in $\mathcal{G}$, denoted by $A_1 \perp\!\!\!\perp_{\mathcal{G}} A_2 \mid A_3$, if and only if for every $a_1 \in A_1$ and every $a_2 \in A_2$ and every path p from a_1 to a_2 it holds that $p^\dagger \cap A_3 \neq \emptyset$.

The concept of graph separation aims at representing conditional independencies in a graphical manner. Ideally, we want to characterize conditional independence by graph separation, i. e., we want to have

$$A_1 \perp\!\!\!\perp_{\mathcal{G}} A_2 \mid A_3 \quad \textit{iff} \quad A_1 \perp\!\!\!\perp_{\mathrm{P}} A_2 \mid A_3 \tag{2.4}$$

for every $A_1, A_2, A_3 \subseteq \mathsf{At}$. However, the condition represented by Equation (2.4) is not achievable in general as conditional independence is more expressive than graph separation (Beierle and Kern-Isberner, 2008). For defining Markov nets we demand only a weaker condition to hold.

Definition 2.21 (Independence map)**.** Let At be a propositional signature, $\mathcal{G} = (\mathsf{At}, E)$ an undirected graph, and P a probability function on $\Omega(\mathsf{At})$. Then $\mathcal{G}$ is an independence map for P if and only if for every $A_1, A_2, A_3 \subseteq$ At it holds that

$$\textit{if} \quad A_1 \perp\!\!\!\perp_{\mathcal{G}} A_2 \,|\, A_3 \quad \textit{then} \quad A_1 \perp\!\!\!\perp_{\mathrm{P}} A_2 \,|\, A_3 \quad . \tag{2.5}$$

Equation (2.5) is called the *global Markov property*.

Definition 2.22 (Markov net)**.** Let At be a propositional signature, $\mathcal{G} = (\mathsf{At}, E)$ an undirected graph, and P a probability function on $\Omega(\mathsf{At})$. Then $\mathcal{G}$ is a *Markov net* for P if and only if $\mathcal{G}$ is a minimal independence map for P, i. e., if and only if no $\mathcal{G} = (\mathsf{At}, E')$ with $E' \subsetneq E$ is an independence map for P.

Example 2.6. Let $\mathsf{At} = \{a, b, c, d\}$ be a set of propositions and consider the graph $\mathcal{G} = (\mathsf{At}, E)$ in Figure 3. Then $\mathcal{G}$ is an independence map for a probability function P if it holds that

$$\{a\} \perp\!\!\!\perp_{\mathrm{P}} \{d\} \,|\, \{b, c\} \quad \text{and} \quad \{b\} \perp\!\!\!\perp_{\mathrm{P}} \{c\} \,|\, \{a, d\} \quad .$$

Furthermore, $\mathcal{G}$ is a Markov net for P if *none* of the following independencies hold

$$\begin{aligned}
&\{a\} \perp\!\!\!\perp_{\mathrm{P}} \{b\} \,|\, S \quad \text{with } S \subseteq \{c, d\} \\
&\{a\} \perp\!\!\!\perp_{\mathrm{P}} \{c\} \,|\, S \quad \text{with } S \subseteq \{b, d\} \\
&\{a\} \perp\!\!\!\perp_{\mathrm{P}} \{d\} \,|\, S \quad \text{with } S \in \{\{b\}, \{c\}, \emptyset\} \\
&\{b\} \perp\!\!\!\perp_{\mathrm{P}} \{c\} \,|\, S \quad \text{with } S \in \{\{a\}, \{d\}, \emptyset\} \\
&\{b\} \perp\!\!\!\perp_{\mathrm{P}} \{d\} \,|\, S \quad \text{with } S \subseteq \{a, c\} \\
&\{c\} \perp\!\!\!\perp_{\mathrm{P}} \{d\} \,|\, S \quad \text{with } S \subseteq \{a, b\} \quad .
\end{aligned}$$

Having a Markov net $\mathcal{G}$ for a probability function P allows for a compact representation of P as P factorizes over the cliques $\mathsf{cl}(\mathcal{G})$ of $\mathcal{G}$. This means, for each clique C of $\mathcal{G}$ a function $\phi_C : C \to [0, 1]$, also referred to as *clique potential*, can be determined such that

$$P(\omega) = \prod_{C \in \mathsf{cl}(\mathcal{G})} \phi_C(\omega_C)$$

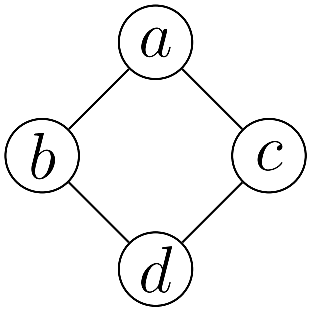

Figure 3: The graph $\mathcal{G} = (\mathsf{At}, E)$ from Example 2.6

with $\omega \in \Omega(\mathsf{At})$ and ω_C is the *projection* of ω on C, e. g. $a\bar{b}c\bar{d}_{\{a,b\}} =_{def} a\bar{b}$. For a more throughout discussion on this topic see e. g. (Pearl, 1998; Beierle and Kern-Isberner, 2008).

As Example 2.6 shows, Markov nets allow for cyclic dependencies of propositions, in contrast to Bayesian networks. However, Markov nets cannot represent induced dependencies like Bayesian networks, cf. (Pearl, 1998).

Both Bayesian networks and Markov nets heavily rely on assumptions regarding conditional independences that can be hardly fulfilled in real-world scenarios. In the following section we describe probabilistic conditional logic as an alternative to these frameworks that makes no such assumptions whatsoever.

2.3 PROBABILISTIC CONDITIONAL LOGIC AND MAXIMUM ENTROPY

Conditional logic (Nute and Cross, 2002) is a knowledge representation formalism that concentrates on the role of *conditionals* or *if-then-rules*. A conditional of the form $(\psi \,|\, \phi)$ connects some detached pieces of information ϕ, ψ and represents a rule "*If ϕ then (usually, probably) ψ*". For the upcoming presentation of the syntactical constructs of probabilistic conditional logic let $\mathcal{L}$ be one of $\mathcal{L}(\mathsf{At})$ or $\mathcal{L}(\Sigma, V)$.

Definition 2.23 (Conditional)**.** Let $\phi, \psi \in \mathcal{L}$. A *conditional* is an expression of the form $(\psi \,|\, \phi)$. Let $(\mathcal{L} \,|\, \mathcal{L})$ denote the set of all conditionals of $\mathcal{L}$.

In this thesis we consider *probabilistic conditionals* (Benferhat *et al.*, 1999; Rödder, 2000; Kern-Isberner, 2001), i. e. rules that are weighted by some probability.

Definition 2.24 (Probabilistic conditional)**.** Let $\phi, \psi \in \mathcal{L}$ be finite[2], and $d \in [0, 1]$. A *probabilistic conditional* is an expression of the form $(\psi \,|\, \phi)[d]$. Let $(\mathcal{L} \,|\, \mathcal{L})^{pr}$ denote the language of all probabilistic conditionals of $\mathcal{L}$.

2 A formula $\phi \in \mathcal{L}$ is finite if and only if it has finite length.

If $\phi \equiv^{\mathrm{P}} \top$ or $\phi \equiv^{\mathrm{F}} \top$ we abbreviate $(\psi \mid \phi)[d]$ simply by $(\psi)[d]$. Such a probabilistic conditional is also called a *probabilistic fact*.

Note that we restrain our attention to "logical" variables, i. e. variables that represent atoms of the logical language. Probabilistic conditional logic also allows for multi-valued variables that take different values than the classical $\{\mathsf{true}, \mathsf{false}\}$. Extending probabilistic conditional logic with multi-valued variables is straightforward (Kern-Isberner, 2001) but those are not needed in this thesis.

Definition 2.25 (Knowledge base)**.** A *knowledge base* $\mathcal{R}$ on $\mathcal{L}$ is a finite subset of $(\mathcal{L} \mid \mathcal{L})^{pr}$.

Note that in general the term *knowledge base* simply refers to a collection of formulas represented in some knowledge representation formalism. As we are only dealing with knowledge bases of probabilistic conditionals we simply use the term *knowledge base* for collections of those. For technical reasons we assume the existence of an arbitrary total order $<_{\mathcal{R}}$ on the probabilistic conditionals of a knowledge base $\mathcal{R}$ and denote by $\langle \mathcal{R} \rangle$ the vector representation of $\mathcal{R}$ with respect to $<_{\mathcal{R}}$. More precisely, if $\mathcal{R} = \{r_1, \ldots, r_n\}$ and $r_i <_{\mathcal{R}} r_j$ if and only if $i < j$ (for all $i, j = 1, \ldots, n$) then $\langle \mathcal{R} \rangle = (r_1, \ldots, r_n)$. The order $<_{\mathcal{R}}$ is meant to have no special meaning and is only used to enumerate the elements of $\mathcal{R}$ in an unambiguous way. For any set X let $\mathfrak{P}^{\mathrm{ord}}(X)$ denote the set of all vectors with elements of X. Then for every knowledge base $\mathcal{R}$ it holds that $\langle \mathcal{R} \rangle \in \mathfrak{P}^{\mathrm{ord}}((\mathcal{L} \mid \mathcal{L})^{pr})$.

Example 2.7. The omnipresent penguin example—see e. g. (Beierle and Kern-Isberner, 2008)—can be represented as a knowledge base $\mathcal{R}_{\text{penguins}}$ with $\mathcal{R}_{\text{penguins}} =_{def} \{r_1, r_2, r_3\}$ on $\mathcal{L} =_{def} \mathcal{L}(\{b, p, f\})$ with

$$r_1 =_{def} (b \mid p)[1.0] \qquad r_2 =_{def} (f \mid b)[0.9] \qquad r_3 =_{def} (f \mid p)[0.01] \quad .$$

In $\mathcal{R}_{\text{penguins}}$ the rule r_1 denotes that every penguin (p) is a bird (b), rule r_2 denotes that "most" birds fly (f), and rule r_3 denotes that "almost no" penguins fly (the probabilities are arbitrary and just for presentation).

The replacement operator $[\cdot]$ is extended on first-order (probabilistic) conditionals and knowledge bases in the usual way, e. g., it holds that

$$\begin{aligned} (\psi \mid \phi)[d][x/y] &= (\psi[x/y] \mid \phi[x/y])[d] \quad \text{and} \\ \{r_1, \ldots, r_n\}[x/y] &= \{r_1[x/y], \ldots, r_n[x/y]\} \quad . \end{aligned}$$

Semantics are given to propositional probabilistic conditionals using conditional probabilities. Giving semantics to first-order probabilistic conditionals is an issue that has not been dealt with in the literature yet, with few exceptions, cf. e. g. (Kern-Isberner and Lukasiewicz, 2004; Fisseler, 2010; Loh *et al.*, 2010). This is one of the major topics of this thesis and is discussed in Chapter 5. For the rest of this section we assume $\mathcal{L}$ to be a propositional language $\mathcal{L}(\mathsf{At})$ for some propositional signature At. Every probability function P on $\mathcal{L}(\mathsf{At})$ is an interpretation for a probabilistic conditional of $\mathcal{L}(\mathsf{At})$. A probability function P on $\mathcal{L}(\mathsf{At})$ is a *probabilistic model* (or just *model*) for a probabilistic conditional $(\psi \,|\, \phi)[d] \in (\mathcal{L}(\mathsf{At}) \,|\, \mathcal{L}(\mathsf{At}))^{pr}$, written $P \models^{pr} (\psi \,|\, \phi)[d]$, if and only if

$$P(\psi \,|\, \phi) = d \text{ or } P(\phi) = 0 \quad . \tag{2.6}$$

Note that the above definition differs from the usual definition used for probabilistic satisfiability—see e. g. (Kern-Isberner, 2001)—which is

$$P \models^{pr} (\psi \,|\, \phi)[d] \quad \text{if and only if} \quad P(\psi \,|\, \phi) = d \text{ and } P(\phi) > 0 \quad . \tag{2.7}$$

For this thesis we choose the definition via Equation (2.6) as this can concisely be characterized by $P(\phi \wedge \psi) = d \cdot P(\phi)$ without any case differentiation for $P(\phi) = 0$. Note that this approach has also been taken in other works on probabilistic reasoning such as (Paris, 1994). It follows, that this definition allows for the knowledge base $\mathcal{R} = \{(b \,|\, a)[d], (a)[0]\}$ to be satisfiable for any value of $d \in [0,1]$. It is arguable whether this is meaningful or not but we stick to this definition and give a remark when this choice has a crucial effect.

Let $\mathsf{Mod}^{\mathrm{Pr}}(r)$ be the set of models of a probabilistic conditional r, i. e. $\mathsf{Mod}^{\mathrm{Pr}}(r) = \{P \mid P \models^{pr} r\}$. A probabilistic conditional $(\psi \,|\, \phi)[d]$ is *self-consistent* if there exists a probability function P with $P \models^{pr} (\psi \,|\, \phi)[d]$, i. e., if $\mathsf{Mod}^{\mathrm{Pr}}((\psi \,|\, \phi)[d]) \neq \emptyset$. For example, the conditional $(\perp)[d]$ is self-consistent for $d = 0$ and not self-consistent for $d \in (0,1]$. A probabilistic conditional $(\psi \,|\, \phi)[d]$ is *tautological* if for all probability functions P on $\mathcal{L}(\mathsf{At})$ it holds that $P \models^{pr} (\psi \,|\, \phi)[d]$. For example, the conditional $(\perp)[0]$ is tautological, as well as $(\phi' \,|\, \phi)[1]$ for every $\phi', \phi \in \mathcal{L}(\mathsf{At})$ with $\phi \models^{\mathrm{P}} \phi'$. Note that if $(\psi \,|\, \phi)[d]$ is tautological then $(\psi \,|\, \phi)[d']$ is not self-consistent for every $d' \neq d$. For the rest of this thesis, as a technical convenience we consider only self-consistent and non-tautological probabilistic conditionals.

A probability function P is a *model* of a knowledge base $\mathcal{R}$ on $\mathcal{L}(\mathsf{At})$, written $P \models^{pr} \mathcal{R}$, if and only if P is a model for every probabilistic conditional $r \in \mathcal{R}$. Let $\mathsf{Mod}^{\mathrm{Pr}}(\mathcal{R})$ be the set of models of $\mathcal{R}$.

Example 2.8. We continue Example 2.7. Consider the probability function P_1 given in Table 1. In P_1 the following holds

$$\begin{aligned} P_1(p) &= 0.05 & P_1(p \wedge b) &= 0.05 \\ P_1(b) &= 0.55 & P_1(b \wedge f) &= 0.495 \\ P_1(p \wedge f) &= 0.005 \quad . \end{aligned}$$

It follows that $P_1(b \,|\, p) = 1.0$, $P_1(f \,|\, b) = 0.9$, and $P_1(f \,|\, p) = 0.01$. Therefore it holds that $P_1 \models^{pr} \mathcal{R}_{\text{penguins}}$.

$\Omega(\{b, p, f\})$	P_1	$\Omega(\{b, p, f\})$	P_1
bpf	0.005	$\overline{b}pf$	0.0
$bp\overline{f}$	0.045	$\overline{b}p\overline{f}$	0.0
$b\overline{p}f$	0.49	$\overline{b}\overline{p}f$	0.2
$b\overline{p}\overline{f}$	0.01	$\overline{b}\overline{p}\overline{f}$	0.25

Table 1: The probability function P_1 for $\mathcal{R}_{\text{penguins}}$

A knowledge base $\mathcal{R}$ is *consistent* if there is at least one probability function P with $P \models^{pr} \mathcal{R}$, otherwise $\mathcal{R}$ is *inconsistent*. Two probabilistic conditionals r_1 and r_2 are *equivalent*, denoted by $r_1 \equiv^{\text{pr}} r_2$, if and only if for every probability function P it holds that $P \models^{pr} r_1$ whenever $P \models^{pr} r_2$, i.e., if and only if $\mathsf{Mod}^{\text{Pr}}(r_1) = \mathsf{Mod}^{\text{Pr}}(r_2)$. The equivalence of probabilistic conditionals can be characterized as follows.

Proposition 2.3. *Let $(\psi \mid \phi)[d]$ and $(\psi' \mid \phi')[d']$ be some probabilistic conditionals. It holds that $(\psi \mid \phi)[d] \equiv^{pr} (\psi' \mid \phi')[d']$ if and only if either*

1. *$\phi \equiv^P \phi'$ and $\psi \wedge \phi \equiv^P \psi' \wedge \phi'$ and $d = d'$ or*
2. *$\phi \equiv^P \phi'$ and $\psi \wedge \phi \equiv^P \overline{\psi'} \wedge \phi'$ and $d = 1 - d'$ or*
3. *both $(\psi \mid \phi)[d]$ and $(\psi' \mid \phi')[d']$ are not self-consistent or*
4. *both $(\psi \mid \phi)[d]$ and $(\psi' \mid \phi')[d']$ are tautological.*

The proof of Proposition 2.3 can be found in Appendix A on page 226.

Two knowledge bases $\mathcal{R}_1$ and $\mathcal{R}_2$ are *kb-equivalent*, denoted by $\mathcal{R}_1 \equiv^{\text{kb}} \mathcal{R}_2$, if and only if for every probability function P it holds that $P \models^{pr} \mathcal{R}_1$ whenever $P \models^{pr} \mathcal{R}_2$, i.e., if and only if $\mathsf{Mod}^{\text{Pr}}(\mathcal{R}_1) = \mathsf{Mod}^{\text{Pr}}(\mathcal{R}_2)$. Due to its symmetric definition $\equiv^{\text{kb}}$ is obviously an equivalence relation, i.e. it is reflexive ($\mathcal{R} \equiv^{\text{kb}} \mathcal{R}$), symmetric ($\mathcal{R}_1 \equiv^{\text{kb}} \mathcal{R}_2$ implies $\mathcal{R}_2 \equiv^{\text{kb}} \mathcal{R}_1$) and transitive ($\mathcal{R}_1 \equiv^{\text{kb}} \mathcal{R}_2$ and $\mathcal{R}_2 \equiv^{\text{kb}} \mathcal{R}_3$ implies $\mathcal{R}_1 \equiv^{\text{kb}} \mathcal{R}_3$). It is also quite clear that all inconsistent knowledge bases lie in the same equivalence class because it holds that $\mathsf{Mod}^{\text{Pr}}(\mathcal{R}) = \emptyset$ for inconsistent $\mathcal{R}$.

Proposition 2.4. *For inconsistent $\mathcal{R}_1$ and $\mathcal{R}_2$ it holds that $\mathcal{R}_1 \equiv^{kb} \mathcal{R}_2$.*

In order to distinguish different types of inconsistent knowledge bases we introduce an alternative concept of equivalence. Two knowledge bases $\mathcal{R}_1$ and $\mathcal{R}_2$ are *cond-equivalent*, denoted by $\mathcal{R}_1 \equiv^{\text{cond}} \mathcal{R}_2$, if and only if for every $r_1 \in \mathcal{R}_1$ there is an $r_2 \in \mathcal{R}_2$ such that $r_1 \equiv^{\text{pr}} r_2$ and for every $r_2 \in \mathcal{R}_2$ there is an $r_1 \in \mathcal{R}_1$ such that $r_1 \equiv^{\text{pr}} r_2$. Note that $\equiv^{\text{cond}}$ is indeed an equivalence relation. It also holds that neither kb-equivalence nor cond-equivalence implies $\mathcal{R}_1$ and $\mathcal{R}_2$ to have the same cardinality. For example, the two knowledge bases $\mathcal{R}_1 = \{(\neg a \vee b)[0.2], (a \wedge \neg b)[0.8]\}$ and $\mathcal{R}_2 = \{(a \Rightarrow b)[0.2]\}$ are both kb- and cond-equivalent. These concepts of equivalence are related as follows.

Lemma 2.1. *If* $\mathsf{Mod}^{Pr}(\mathcal{R}_1) = \emptyset$ *and* $\mathcal{R}_1 \equiv^{cond} \mathcal{R}_2$ *then* $\mathsf{Mod}^{Pr}(\mathcal{R}_2) = \emptyset$.

Proof. Assume $\mathsf{Mod}^{\mathrm{Pr}}(\mathcal{R}_2) \neq \emptyset$ and let $P \in \mathsf{Mod}^{\mathrm{Pr}}(\mathcal{R}_2)$. Then $P \models^{pr} r$ for every $r \in \mathcal{R}_1$ as there exists a $r' \in \mathcal{R}_2$ with $r \equiv^{\mathrm{pr}} r'$ and $P \models^{pr} r'$. It follows $P \models^{pr} \mathcal{R}_1$ contradicting $\mathsf{Mod}^{\mathrm{Pr}}(\mathcal{R}_1) = \emptyset$. □

Proposition 2.5. *If* $\mathcal{R}_1 \equiv^{cond} \mathcal{R}_2$ *then* $\mathcal{R}_1 \equiv^{kb} \mathcal{R}_2$.

Proof. Let P be a probability function with $P \models^{pr} \mathcal{R}_1$ and let $r \in \mathcal{R}_2$. Due to $\mathcal{R}_1 \equiv^{\mathrm{cond}} \mathcal{R}_2$ there is an $r' \in \mathcal{R}_1$ with $r \equiv^{\mathrm{pr}} r'$. As $P \models^{pr} \mathcal{R}_1$ it follows that $P \models^{pr} r'$ and therefore $P \models^{pr} r$. As this is true for all $r \in \mathcal{R}_2$ it follows $P \models^{pr} \mathcal{R}_2$. Hence, for every P with $P \models^{pr} \mathcal{R}_1$ it holds that $P \models^{pr} \mathcal{R}_2$ and with the same argumentation for every P with $P \models^{pr} \mathcal{R}_2$ it holds that $P \models^{pr} \mathcal{R}_1$. It follows that $\mathcal{R}_1 \equiv^{\mathrm{kb}} \mathcal{R}_2$. If $\mathsf{Mod}^{\mathrm{Pr}}(\mathcal{R}_1) = \emptyset$ then $\mathsf{Mod}^{\mathrm{Pr}}(\mathcal{R}_2) = \emptyset$ as well due to Lemma 2.1 and therefore $\mathcal{R}_1 \equiv^{\mathrm{kb}} \mathcal{R}_2$ due to Proposition 2.4. □

The other direction is not true as the following example shows.

Example 2.9. Consider the two knowledge bases

$$\mathcal{R}_1 =_{def} \{(a)[0.7], (a)[0.4]\} \quad \text{and} \quad \mathcal{R}_2 =_{def} \{(b)[0.8], (b)[0.3]\} \quad .$$

Both $\mathcal{R}_1$ and $\mathcal{R}_2$ are inconsistent and therefore $\mathcal{R}_1 \equiv^{\mathrm{kb}} \mathcal{R}_2$. But it holds that $\mathcal{R}_1 \not\equiv^{\mathrm{cond}} \mathcal{R}_2$ as e. g. both $(a)[0.7] \not\equiv^{\mathrm{pr}} (b)[0.8]$ and $(a)[0.7] \not\equiv^{\mathrm{pr}} (b)[0.3]$.

Even if both $\mathcal{R}_1$ and $\mathcal{R}_2$ are consistent the other direction is not true.

Example 2.10. Consider the two knowledge bases

$$\mathcal{R}_1 =_{def} \{(b \mid a)[1.0], (a)[1.0]\} \quad \text{and} \quad \mathcal{R}_2 =_{def} \{(a \wedge b)[1.0]\} \quad .$$

It holds that $\mathcal{R}_1 \not\equiv^{\mathrm{cond}} \mathcal{R}_2$ but $\mathcal{R}_1 \equiv^{\mathrm{kb}} \mathcal{R}_2$.

In this thesis we are especially interested in inconsistent knowledge bases. Due to Proposition 2.4 the relation $\equiv^{\mathrm{kb}}$ is not an appropriate choice for dealing with inconsistent knowledge bases, so we use $\equiv^{\mathrm{cond}}$ instead. We come back to this issue in Chapters 3 and 4.

Usually, one is interested in using a (consistent) knowledge base for reasoning. A simple inference relation $\models^{pr}_{int}$ for probabilistic conditional logic can be given by computing lower and upper bounds for queries with respect to the whole set of models (Lukasiewicz, 1999). Let $\mathcal{R}$ be a knowledge base. Then a *bounded probabilistic conditional* is an expression of the form $(\psi \mid \phi)[l, u]$ with $\psi, \phi \in \mathcal{L}(\mathsf{At})$ and $l, u \in [0, 1]$ with $l \leq u$. The interpretation of a bounded probabilistic conditional $(\psi \mid \phi)[l, u]$ is that the probability of ψ given ϕ lies in between l and u, i. e., a probability function P satisfies $(\psi \mid \phi)[l, u]$, denoted by $P \models^{pr} (\psi \mid \phi)[l, u]$, if and only if $P(\psi \mid \phi) \in [l, u]$ or $P(\phi) = 0$. Let $(\mathcal{L} \mid \mathcal{L})^{pr,pr}$ be the language of all bounded probabilistic conditionals. Then a bounded probabilistic conditional $(\psi \mid \phi)[l, u]$ can be inferred from a knowledge base $\mathcal{R}$ for $\mathcal{L}(\mathsf{At})$, written $\mathcal{R} \models^{pr}_{int} (\psi \mid \phi)[l, u]$, if and only if it holds that

$$l = \inf\{P(\psi \mid \phi) \mid P \models^{pr} \mathcal{R}\} \quad \text{and}$$
$$u = \sup\{P(\psi \mid \phi) \mid P \models^{pr} \mathcal{R}\} \quad .$$

Note that $(\mathcal{L} \mid \mathcal{L})^{pr,pr}$ subsumes $(\mathcal{L} \mid \mathcal{L})^{pr}$ as $(\psi \mid \phi)[d] \in (\mathcal{L} \mid \mathcal{L})^{pr}$ is semantically equivalent to $(\psi \mid \phi)[d, d] \in (\mathcal{L} \mid \mathcal{L})^{pr,pr}$, i.e., it holds that $P \models^{pr} (\psi \mid \phi)[d]$ if and only if $P \models^{pr} (\psi \mid \phi)[d, d]$ for every probability function P.

Given a consistent knowledge base $\mathcal{R}$ the set of models of $\mathcal{R}$ is usually infinite and inference via $\models^{pr}_{int}$ is only of limited value.

Example 2.11. Consider the knowledge base $\mathcal{R} =_{def} \{r_1, r_2\}$ given via

$$r_1 =_{def} (b \mid a)[d_1] \qquad\qquad r_2 =_{def} (b \mid c)[d_2]$$

for a propositional signature $\mathsf{At} = \{a, b, c\}$. There are infinitely many models of $\mathcal{R}$ and it holds that $\mathcal{R} \models^{pr}_{int} (b \mid a \wedge c)[0, 1]$, cf. (Beierle and Kern-Isberner, 2008).

One way to cope with the problem of infinitely many models is to perform *model-based inductive reasoning*, i.e. to select one "suitable" representative and rely on this one for reasoning. Guided by the *principle of maximum entropy* a probability function with maximum entropy is such a "suitable" representative (Paris, 1994).

Definition 2.26 (Entropy)**.** Let P be a probability function on a set $\mathcal{X}$. The *entropy* $H(P)$ of P is defined as

$$H(P) =_{def} - \sum_{x \in \mathcal{X}} P(x) \,\mathsf{ld}\, P(x)$$

with $0 \cdot \mathsf{ld}\, 0 =_{def} 0$.[3]

The entropy measures the amount of indeterminateness of a probability function P. A probability function P_1 that describes absolutely certain knowledge, i.e. $P_1(x) = 1$ for some $x \in \mathcal{X}$ and $P_1(x') = 0$ for every $x' \in \mathcal{X}$ with $x' \neq x$, yields minimal entropy $H(P_1) = 0$. The uniform probability function P_0 with $P_0(x) = 1/|\mathcal{X}|$ for every $x \in \mathcal{X}$ (with finite $\mathcal{X}$) yields maximal entropy $H(P_0) = -\mathsf{ld}\, 1/|\mathcal{X}|$.

By selecting a model of a knowledge base $\mathcal{R}$ that has maximal entropy one gets a probability function that both satisfies all conditionals in $\mathcal{R}$ and adds as less additional information (in the information-theoretic sense) as possible.

3 ld is the binary logarithm.

Definition 2.27 (Maximum entropy model)**.** Let $\mathcal{R}$ be a consistent knowledge base on $\mathcal{L}(\mathsf{At})$. A *maximum entropy model* P^* of $\mathcal{R}$ is a probability function on $\mathcal{L}$ that satisfies

$$P^* = \arg\max_{P \models^{pr} \mathcal{R}} H(P) \tag{2.8}$$

For a consistent knowledge base $\mathcal{R} \subseteq (\mathcal{L}(\mathsf{At}) \,|\, \mathcal{L}(\mathsf{At}))^{pr}$ the maximum entropy model P^* is uniquely determined, cf. (Shore and Johnson, 1980; Goldszmidt *et al.*, 1993; Kern-Isberner, 2001). Therefore, for the rest of this work we refer to the unique maximum entropy model P^* of $\mathcal{R}$ by $\mathsf{ME}(\mathcal{R})$.

Reasoning with $\mathsf{ME}(\mathcal{R})$ satisfies several commonsense properties for inference (Kern-Isberner, 2001). In (Shore and Johnson, 1980) it has been shown that the maximum entropy model is characterized by four simple properties for probabilistic reasoning like *uniqueness* and *irrelevance of syntax*. A similar but constructive characterization has been made in (Kern-Isberner, 2001) where the *principle of minimum cross-entropy* (a generalization of the principle of maximum entropy) has been characterized by four simple properties as well: *the principle of conditional preservation*, *functional concept*, *the principle of logical coherence*, and *the principle of representation invariance*. Reasoning using the principle of maximum entropy also satisfies the system P properties which are regarded as the minimal requirement any reasonable default reasoning machinery should meet (Makinson, 1989), see (Kern-Isberner, 2001) for details.

2.4 RELATIONAL PROBABILISTIC REASONING

Propositional logic has been employed for probabilistic reasoning for decades and many reasonable and practical approaches have been developed so far, see e. g. the previous section and (Paris, 1994; Pearl, 1998). But propositional logic fails to model more complex scenarios that involve reasoning about individuals or reasoning about relationships of individuals. First-order logic extends propositional logic and compensates for this lack of expressivity, cf. Section 2.1.2. When considering probabilistic reasoning the need for a more expressive (object-level) language is present as well. Consider the following example taken from (Friedman *et al.*, 1999).

Example 2.12. The blood type of a person probabilistically depends on the blood types of his or her parents. For example, if the mother has blood type AB and the father has blood type 0 then the person's blood type is either A or B where the distribution is a probabilistic one. If one is interested in representing the probabilistic dependencies of blood types of persons given some pedigree, propositional approaches to probabilistic reasoning are hardly apt for this task. While it is possible to represent a single such dependency in e. g. probabilistic conditional logic via[4]

$$(\mathit{bloodtype\_john\_is\_A} \mid \mathit{bloodtype\_carl\_is\_0} \wedge \mathit{bloodtype\_mary\_is\_AB})[0.7]$$

the generalization to a whole pedigree is cumbersome and results in many similar looking conditionals. However, in this example we are concerned with individuals in some domain that are related via binary relations (*mother* and *father*) and in whose attributes (e. g. *bloodtype*) we are interested. This motivates the need for combining probabilistic reasoning and first-order logic.

Frameworks that combine probabilistic reasoning with *full* first-order logic are rare due to the computational complexity of inference, see (Grove *et al.*, 1996a) for some discussion. In order to avoid confusion we use the term "relational probabilistic framework" to denote frameworks that use first-order elements for probabilistic reasoning. During the past few years the fields of *probabilistic inductive logic programming* and *statistical relational learning* have put forth a lot of proposals that deal with combining traditional probabilistic models of belief like Bayesian networks or Markov nets (see above) with first-order logic, cf. (Getoor and Taskar, 2007; De Raedt *et al.*, 2008). The relational structure of many real-world application domains such as telecommunication networks, citation analysis, human sciences, bioinformatics, and logistics as well as the presence of uncertainty in these domains demand sophisticated reasoning and learning methods employing both these concepts, see e. g. (Lodhi and Muggleton, 2004; Cocura *et al.*, 2006) for some applications. Two of the most prominent approaches for extending propositional approaches to the relational case are Bayesian logic programs (Kersting and Raedt, 2007) and Markov logic networks (Richardson and Domingos, 2006; Domingos and Richardson, 2007), extending Bayesian networks and Markov nets, respectively.

In the following, we give a brief introduction to both Bayesian logic programs and Markov logic networks.

2.4.1 *Bayesian Logic Programs*

Bayesian logic programming (Kersting and Raedt, 2007) is an approach that combines logic programming (Gelfond and Lifschitz, 1991; Gelfond and Leone, 2002) and Bayesian networks. Bayesian logic programs (BLPs) use a standard logic programming language and attach to each logical (Horn) clause a set of probabilities that define a conditional probability distribution of the head of the clause given specific instantiations of the body of the clause.

In contrast to first-order logic, the general framework of Bayesian logic programming employs an extended form of predicates and atoms. In BLPs,

4 The probability in this example is just arbitrary.

Bayesian predicates are predicates that feature an arbitrary set as possible states, i. e. not necessarily the Boolean values $\{\mathsf{true}, \mathsf{false}\}$. However, as for probabilistic conditional logic (see Section 2.3) we focus on "logical" predicates and atoms in this thesis. For the rest of this section let $\Sigma = (U, Pred, \emptyset)$ be a first-order signature with finite U and without functors and let V be a set of variables.

The basic structure for knowledge representation in Bayesian logic programs are *Bayesian clauses* which model probabilistic dependencies between Bayesian atoms. Let $\mathbb{B} = \{\mathsf{true}, \mathsf{false}\}$ denote the Boolean truth values.

Definition 2.28 (Bayesian clause)**.** A *Bayesian clause* c is an expression $c = (h \mid b_1, \ldots, b_n)$ with atoms $h, b_1, \ldots, b_n$ of Σ and V.

Note that Bayesian clauses use the conditional pipe operator "$\mid$" instead of a logical implication "$\Leftarrow$" to highlight the probabilistic interpretation of the clause. As for Horn logic, for a Bayesian clause $c = (h \mid b_1, \ldots, b_n)$ we abbreviate $\mathsf{head}(c) = h$ (the *head* or *conclusion* of the clause) and $\mathsf{body}(c) = \{b_1, \ldots, b_n\}$ (the *body* or *premise* of the clause).

Definition 2.29 (Conditional probability distribution)**.** Let c be a Bayesian clause of the form $c = (h \mid b_1, \ldots, b_n)$. A *conditional probability distribution* cpd_c for c is a function

$$\mathsf{cpd}_c : \mathbb{B}^{n+1} \to [0, 1]$$

that satisfies

$$\mathsf{cpd}_c(\mathsf{true}, x_1, \ldots, x_n) + \mathsf{cpd}_c(\mathsf{false}, x_1, \ldots, x_n) = 1 \quad . \tag{2.9}$$

for every $(x_1, \ldots, x_n) \in \mathbb{B}^n$. Let CPD_p denote the set of all conditional probability distributions for atoms of predicate p, i. e., it holds that

$$\mathsf{CPD}_p =_{def} \{\mathsf{cpd}_{h \mid b_1, \ldots, b_n} \mid h \text{ is an atom of } p\} \quad .$$

For a Bayesian clause $c = (h \mid b_1, \ldots, b_n)$ and a conditional probability distribution cpd_c the expression $\mathsf{cpd}_c(y, x_1, \ldots, x_n) = d$ is to be read as "*the probability of h being y is d given that b_i is x_i for $i = 1, \ldots, n$*". Equation (2.9) is a normalization constraint that ensures that a conditional probability distribution can be used to propagate probabilities in the sense of Kolmogorov, see Definition 2.15 on page 19.

As usual, if the body of a Bayesian clause c is empty we write c as (h) instead of $(h \mid)$ and call c a *Bayesian fact*. A function cpd_c for c expresses the conditional probability distribution $P(\mathsf{head}(c) \mid \mathsf{body}(c))$ and thus partially describes an underlying probability function P.

The following example is adapted from (Pearl, 1998) and can also be found in (Kersting and De Raedt, 2000).

Example 2.13. We consider a scenario where our protagonist James is on the road and gets a call from his neighbor saying that the alarm of James' house is ringing. James has some uncertain beliefs about the relationships between burglaries, types of neighborhoods, natural disasters, and alarms. For example, he knows that if there is a tornado warning for his home place, then the probability of a tornado triggering the alarm of his house is 0.9. He also knows that if a burglary attempt takes place, the alarm will ring with probability 0.9. Further he knows that if you live in a bad neighborhood, then there is 0.6 probability of a burglary, whereas in an average neighborhood, there is 0.4 probability, and in a good neighborhood there is merely a 0.3 probability. We represent this scenario using the predicates $\{alarm/1, burglary/1, tornado/1, lives\_in/2, neighborhood/2\}$ with informal interpretations

$alarm(\mathsf{X})$	The alarm at X's house is ringing
$burglary(\mathsf{X})$	A burglary attempt takes place at the X's house
$tornado(\mathsf{Y})$	There is a tornado warning for city Y
$lives\_in(\mathsf{X}, \mathsf{Y})$	Person X lives in city Y
$neighborhood(\mathsf{X}, \mathsf{Z})$	The neighborhood of X is in condition Z

and define a set $\{c_1, c_2, c_3, c_4, c_5\}$ of Bayesian clauses via

$$c_1 =_{def} (alarm(\mathsf{X}) \mid burglary(\mathsf{X}))$$
$$c_2 =_{def} (alarm(\mathsf{X}) \mid lives\_in(\mathsf{X}, \mathsf{Y}), tornado(\mathsf{Y}))$$
$$c_3 =_{def} (burglary(\mathsf{X}) \mid neighborhood(\mathsf{X}, \mathsf{good}))$$
$$c_4 =_{def} (burglary(\mathsf{X}) \mid neighborhood(\mathsf{X}, \mathsf{average}))$$
$$c_5 =_{def} (burglary(\mathsf{X}) \mid neighborhood(\mathsf{X}, \mathsf{bad}))$$

For each Bayesian clause c_i, we define a function cpd_{c_i} that expresses our subjective beliefs (notice that the probabilities stated in the right column are redundant due to Equation (2.9)):

$$\mathsf{cpd}_{c_1}(\mathsf{true}, \mathsf{true}) =_{def} 0.9 \qquad \mathsf{cpd}_{c_1}(\mathsf{false}, \mathsf{true}) =_{def} 0.1$$
$$\mathsf{cpd}_{c_1}(\mathsf{true}, \mathsf{false}) =_{def} 0 \qquad \mathsf{cpd}_{c_1}(\mathsf{false}, \mathsf{false}) =_{def} 1$$

$$\mathsf{cpd}_{c_2}(\mathsf{true}, \mathsf{true}, \mathsf{true}) =_{def} 0.9 \qquad \mathsf{cpd}_{c_2}(\mathsf{false}, \mathsf{true}, \mathsf{true}) =_{def} 0.1$$
$$\mathsf{cpd}_{c_2}(\mathsf{true}, \mathsf{false}, \mathsf{true}) =_{def} 0 \qquad \mathsf{cpd}_{c_2}(\mathsf{false}, \mathsf{false}, \mathsf{true}) =_{def} 1$$
$$\mathsf{cpd}_{c_2}(\mathsf{true}, \mathsf{true}, \mathsf{false}) =_{def} 0.01 \qquad \mathsf{cpd}_{c_2}(\mathsf{false}, \mathsf{true}, \mathsf{false}) =_{def} 0.99$$
$$\mathsf{cpd}_{c_2}(\mathsf{true}, \mathsf{false}, \mathsf{false}) =_{def} 0 \qquad \mathsf{cpd}_{c_2}(\mathsf{false}, \mathsf{false}, \mathsf{false}) =_{def} 1$$

$$\mathsf{cpd}_{c_3}(\mathsf{true}, \mathsf{true}) =_{def} 0.3 \qquad \mathsf{cpd}_{c_3}(\mathsf{false}, \mathsf{true}) =_{def} 0.7$$
$$\mathsf{cpd}_{c_3}(\mathsf{true}, \mathsf{false}) =_{def} 0.7 \qquad \mathsf{cpd}_{c_3}(\mathsf{false}, \mathsf{false}) =_{def} 0.3$$

$$\mathsf{cpd}_{c_4}(\mathsf{true},\mathsf{true}) =_{def} 0.4 \qquad \mathsf{cpd}_{c_4}(\mathsf{false},\mathsf{true}) =_{def} 0.6$$
$$\mathsf{cpd}_{c_4}(\mathsf{true},\mathsf{false}) =_{def} 0.5 \qquad \mathsf{cpd}_{c_4}(\mathsf{false},\mathsf{false}) =_{def} 0.5$$

$$\mathsf{cpd}_{c_5}(\mathsf{true},\mathsf{true}) =_{def} 0.6 \qquad \mathsf{cpd}_{c_5}(\mathsf{false},\mathsf{true}) =_{def} 0.4$$
$$\mathsf{cpd}_{c_5}(\mathsf{true},\mathsf{false}) =_{def} 0.4 \qquad \mathsf{cpd}_{c_5}(\mathsf{false},\mathsf{false}) =_{def} 0.6$$

For example, cpd_{c_2} expresses that our subjective belief on the probability that the alarm of a person X will go on given that we know that X lives in town Y and there is currently a tornado in Y is 0.9. Furthermore, we believe that the probability that the alarm of X will trigger if we know that X lives in Y and that there is no tornado in Y is 0.01.

Considering clauses c_1 and c_2 in Example 2.13 one can see that it is possible to have multiple clauses with the same head. This means that there may be multiple causes for some effect or multiple explanations for some observation. In order to represent these kinds of scenarios the probabilities of causes or explanations have to be aggregated. BLPs facilitate *combining rules* in order to aggregate probabilities that arise from applications of different Bayesian clauses. A combining rule cr_p for a predicate p/n is a function $\mathsf{cr}_p : \mathfrak{P}(\mathsf{CPD}_p) \to \mathsf{CPD}_p$ that assigns to the conditional probability distributions of a set of Bayesian clauses a new conditional probability distribution that represents the *joint* conditional probability distribution obtained from aggregating the given clauses. Appropriate choices for such functions are *average* or *noisy-or*, cf. (Kersting and Raedt, 2007). For example, the *noisy-or* combination of two values $a, b \in [0, 1]$ is defined as $1 - (1 - a)(1 - b)$ and represents a quantified extension of a disjunctive combination.

Example 2.14. We continue Example 2.13. Suppose *noisy-or* to be the combining rule for *alarm*. Then the joint conditional probability distribution $\mathsf{cpd}_{c'}$ for

$$c' =_{def} (\mathit{alarm}(\mathsf{X}) \mid \mathit{burglary}(\mathsf{X}), \mathit{lives\_in}(\mathsf{X},\mathsf{Y}), \mathit{tornado}(\mathsf{Y}))$$

can be computed via

$$\mathsf{cpd}_{c'}(\mathsf{true}, t_1, t_2, t_3) = 1 - (1 - \mathsf{cpd}_{c_1}(\mathsf{true}, t_1))(1 - \mathsf{cpd}_{c_2}(\mathsf{true}, t_2, t_3))$$
$$\mathsf{cpd}_{c'}(\mathsf{false}, t_1, t_2, t_3) = 1 - \mathsf{cpd}_{c'}(\mathsf{true}, t_1, t_2, t_3)$$

for any $t_1, t_2, t_3 \in \{\mathsf{true}, \mathsf{false}\}$.

A combining rule is a heuristic for estimating the probability of an event e given two causes z_1 and z_2 (Pearl, 1998). It has to be noted that combining probabilities in this manner might remove the probabilistic interpretation of the resulting values as specific relationships between the causes have to be assumed. For example, using the *noisy-or* combining rule assumes *ac-*

countability and *exception independence* of z_1 and z_2, cf. (Pearl, 1998). If the presumed relationships do not hold the results might be unexpected and even unwanted. Consider an extension of the BLP defined in Examples 2.13 and 2.14 with the Bayesian clause $c_6 =_{def} (alarm(\mathsf{X}) \,|\, power\_failure(\mathsf{X}))$. Imagine that the conditional probability distribution of c_6 assigns a probability of 0.001 to $alarm(\mathsf{X})$ if there is a power failure. Given the evidences of both a tornado and a power failure the probability of an alarm should be determined only by considering clause c_6 and not by combining c_2 and c_6. We refer to (Pearl, 1998) for a deeper discussion on this topic.

Now we are able to define Bayesian logic programs as follows.

Definition 2.30 (Bayesian logic program)**.** A *Bayesian logic program* B is a tuple $B = (C, D, R)$ with a finite set of Bayesian clauses $C = \{c_1, \ldots, c_n\}$, a set of conditional probability distributions $D = \{\mathsf{cpd}_{c_1}, \ldots, \mathsf{cpd}_{c_n}\}$ (one for each clause in C), and a set of combining rules $R = \{\mathsf{cr}_{p_1}, \ldots, \mathsf{cr}_{p_m}\}$ (one for each predicate appearing in C) .

Semantics are given to Bayesian logic programs via transformation into the propositional case, i. e. into Bayesian networks, cf. Section 2.2.2. Using the constants in U a Bayesian network BN can be constructed by introducing a node for every grounded atom in B. Using the conditional probability distributions of the grounded clauses and the combining rules of B a (joint) conditional probability distribution can be specified for any node in BN. If BN is acyclic, this transformation uniquely determines a probability function P on the grounded Bayesian atoms of B that permits inference, i. e. P can be used to answer queries.

Example 2.15. Let B be the Bayesian logic program described in Example 2.13 and Example 2.14. Let $Q = (alarm(\mathsf{james}) \,|\, E)$ with

$$E =_{def} \quad \{ \quad lives\_in(\mathsf{james}, \mathsf{yorkshire}), tornado(\mathsf{yorkshire}), \\ neighborhood(\mathsf{james}, \mathsf{average}) \quad \}$$

be a query to B that asks for the probability of an alarm in James' house given that he lives in an average neighborhood in Yorkshire and there is currently a tornado warning for Yorkshire. Therefore, the universe under discourse consists of the constants james and yorkshire. By instantiating properly and combining the conditional probability distributions of c_1 and c_2 yielding the function $\mathsf{cpd}_{c'}$ from Example 2.14 and summing over the alternatives true and false for the uncertain event $burglary(\mathsf{james})$, in applying c_4 we get

$$\begin{aligned} P(Q) &= \mathsf{cpd}_{c_4}(\mathsf{true}, \mathsf{true})\mathsf{cpd}_{c'}(\mathsf{true}, \mathsf{true}, \mathsf{true}, \mathsf{true}) + \\ &\qquad \mathsf{cpd}_{c_4}(\mathsf{false}, \mathsf{true})\mathsf{cpd}_{c'}(\mathsf{true}, \mathsf{false}, \mathsf{true}, \mathsf{true}) \qquad (2.10) \\ &= 0.936 \quad . \end{aligned}$$

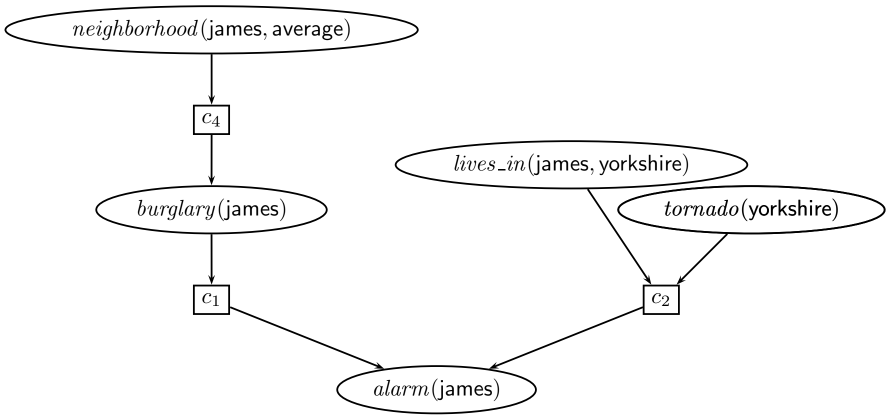

Figure 4: The derivation for the query Q in Example 2.15

Figure 4 illustrates the derivation of the query Q given evidence E. By omitting nodes representing clauses the derivation tree in Figure 4 can be directly transformed into a Bayesian network that can be used to calculate the answer above.

A detailed description of the above declarative semantics and an equivalent procedural semantics which is based on SLD resolution is given in (Kersting and Raedt, 2007).

2.4.2 *Markov Logic Networks*

While Bayesian logic programs extend Bayesian networks to the first-order case Markov logic networks (Richardson and Domingos, 2006; Domingos and Richardson, 2007) extend Markov nets, cf. Section 2.2.3. Another difference is that while BLPs use the syntax of Bayesian networks and incorporate logic programming aspects into it, Markov logic networks use the syntax of first-order logic and incorporate weights.

Definition 2.31 (Markov logic network)**.** Let Σ be a finite first-order signature without functors and let V be a set of variables. A *Markov logic network* (MLN) L on $\mathcal{L}^{\not\forall\not\exists}(\Sigma, V)$ is a finite set of tuples $L = \{(\phi_1, g_1), \ldots, (\phi_n, g_n)\}$ with $\phi_1, \ldots, \phi_n \in \mathcal{L}^{\not\forall\not\exists}(\Sigma, V)$ and $g_1, \ldots, g_n \in \mathbb{R}$.

Note that Markov logic networks only consider the quantifier-free fragment of first-order logic. The weights of an MLN L have no obvious probabilistic interpretation and are interpreted relative to each other when defining the joint probability function for L (see below).

Example 2.16. We represent the scenario from Example 2.13 on page 35 using Markov logic networks. In order to do so, the weights of formulas have to be determined. In (Richardson and Domingos, 2006) it is suggested that weights of formulas have to be learned from data. Nonetheless, in (Richardson and Domingos, 2006; Fisseler, 2008) a heuristic is discussed that determines weights of formulas from probabilities. There, an interpretation of the weight g of a formula ϕ is provided as the log-odd between a world where ϕ is true and a world where ϕ is false, other things being equal. Considering this interpretation one might choose $g =_{def} \ln p/(1-p)$ ($\ln x$ is the natural logarithm of x), see (Fisseler, 2008) for a discussion.

The signature for the Markov logic network L_{alarm} is the same as for the BLP given in Example 2.13 on page 35, i. e., the informal interpretation of the predicates $\{alarm/1, burglary/1, tornado/1, lives\_in/2, neighborhood/2\}$ is the same. We define $L_{\text{alarm}} = \{c_1, c_2, c_3, c_4, c_5\}$ via[5]

$$c_1 =_{def} (\, burglary(\mathsf{X}) \Rightarrow alarm(\mathsf{X})\,, 2.2)$$
$$c_2 =_{def} (\, lives\_in(\mathsf{X}, \mathsf{Y}) \wedge tornado(\mathsf{Y}) \Rightarrow alarm(\mathsf{X})\,, 2.2)$$
$$c_3 =_{def} (\, neighborhood(\mathsf{X}, \mathsf{bad}) \Rightarrow burglary(\mathsf{X})\,, 0.4)$$
$$c_4 =_{def} (\, neighborhood(\mathsf{X}, \mathsf{average}) \Rightarrow burglary(\mathsf{X})\,, -0.4)$$
$$c_5 =_{def} (\, neighborhood(\mathsf{X}, \mathsf{good}) \Rightarrow burglary(\mathsf{X})\,, -0.8)$$

with the heuristically determined weights

$$2.2 = \ln \frac{0.9}{0.1} \qquad 0.4 = \ln \frac{0.6}{0.4} \qquad -0.4 = \ln \frac{0.4}{0.6} \qquad -0.8 = \ln \frac{0.3}{0.7}$$

where the probabilities have been taken from the informal description of the scenario in Example 2.13 on page 35.

Note that the way the scenario of Example 2.13 from page 35 is represented with a Markov logic network in Example 2.16, stems from a direct representation of the informal description and not from transforming the BLP into an MLN. The aim of this representation is to illustrate the formalism of Markov logic networks and *not* to describe the same underlying probabilistic model.

Semantics are given to an MLN L by grounding L appropriately in order to build a Markov net, similar to the approach used for BLPs. For simplicity of presentation we only consider the resulting probability function $P : \Omega(\Sigma) \to [0,1]$ on Herbrand interpretations that can be used to answer queries to L. To this end we need the following notation. Let $\Sigma =_{def} (U, Pred, \emptyset)$, $\phi \in \mathcal{L}^{\not\forall\not\exists}(\Sigma, V)$ and $\omega \in \Omega(\Sigma)$. Define

$$n_\phi(\omega) =_{def} |\{\phi' \in \mathsf{gnd}_U(\phi) \mid \omega \models^{\mathrm{F}} \phi' \}| \quad .$$

5 Note that while it is custom in Markov logic networks to denote variables with a beginning lower-case letter and constants with a beginning upper-case letter, we stick to the notation used throughout this thesis and denote variables with a beginning upper-case letter and constants with a beginning lower-case letter.

The term $n_\phi(\omega)$ denotes the number of instances of ϕ that are satisfied in ω. Then the probability function $P_L \Omega(\Sigma) \to [0,1]$ is defined as

$$P_L(\omega) =_{def} \frac{1}{Z} \exp\left(\sum_{(\phi,g) \in L} n_\phi(\omega) g \right) \tag{2.11}$$

with

$$Z =_{def} \sum_{\omega \in \Omega(\Sigma)} \exp\left(\sum_{(\phi,g) \in L} n_\phi(\omega) g \right)$$

being a normalization constant and $\exp(x) = e^x$ is the exponential function with base e. By defining P_L in this way, worlds that violate fewer instances of formulas are more probable than worlds that violate more instances (depending on the weights of the different formulas). Hence, the fundamental idea for Markov logic networks is that first-order formulas are not handled as hard constraints. Instead, each formula is more or less softened depending on its weight. Hence, a possible world *may* violate a formula without necessarily receiving a zero probability. A formula's weight specifies how strong the formula is, i. e. how much the formula influences the probability of a satisfying world versus a violating world. This way, the weights of all formulas influence the determination of a possible world's probability in a complex manner. One clear advantage of this approach is that Markov logic networks can directly handle contradictions in a knowledge base, since the (contradictory) formulas are weighted against each other.

The probability function P_L can be extended to sentences (ground formulas) of $\mathcal{L}^{\not\forall\not\exists}(\Sigma, V)$ in the same way as in propositional probabilistic conditional logic via

$$P(\phi) =_{def} \sum_{\omega \models^{\mathrm{F}} \phi} P(\omega) \tag{2.12}$$

for ground $\phi \in \mathcal{L}^{\not\forall\not\exists}(\Sigma, V)$.

Determining the probability of a sentence ϕ using Equations (2.11) and (2.12) is merely manageable for very small sets of constants, but intractable for domains of a more realistic size. While $P(\phi)$ can be approximated using Markov chain Monte-Carlo methods (MCMC methods) performance might still be too slow in practice (Richardson and Domingos, 2006). There are more sophisticated and efficient methods to perform approximated inference if ϕ is a conjunction of ground literals, cf. (Richardson and Domingos, 2006). As a side note, for the MLN L_{alarm} from Example 2.16 the probability of the query $Q =_{def} (alarm(\mathsf{james}) \mid E)$ with

$$\begin{aligned} E =_{def} \quad \{ \quad & lives\_in(\mathsf{james}, \mathsf{yorkshire}), tornado(\mathsf{yorkshire}), \\ & neighborhood(\mathsf{james}, \mathsf{average}) \quad \} \end{aligned}$$

from Example 2.15 on page 37 computes to approximately 0.930957 when using the Alchemy system (Kok *et al.*, 2008) with MC-SAT inference algorithm and a maximum of 1000 MCMC sampling steps.

MLNs can handle conflicts between default and specific beliefs quite well, but the determination of appropriate weights is crucial for adequate knowledge representation and inference (Finthammer and Thimm, 2011). It has been noted that MLNs do not allow to express if-then-rules in terms of conditional probabilities (Fisseler, 2008). So the rule-like beliefs in the previous example has to be modeled using material implications. This can be a major drawback, because probabilities of material implications may differ significantly from conditional probabilities and are known to be quite unintuitive in certain cases.

2.5 SUMMARY

In this chapter we gave an overview on both logic and probabilistic reasoning. We presented syntax and semantics for propositional and first-order logic which can be regarded as the foundation for most languages for knowledge representation. Besides classical logical languages we also covered probabilistic ones that allow for the representation of quantified uncertainty. We gave a brief presentation on foundations of probability theory and on the two most prominent representatives of probabilistic networks: Bayesian networks and Markov nets. While probabilistic reasoning is computationally complex in general, probabilistic networks make diverse assumptions regarding probabilistic independencies that render them tractable and allow for an intuitive knowledge representation and high expressivity. However, these assumptions are not justified in many cases and we also presented probabilistic conditional logic which allows for a more declarative knowledge representation of both uncertain and incomplete beliefs. Moreover, probabilistic reasoning using the principle of maximum entropy is well-founded on information-theoretic principles and satisfies several properties for commonsense reasoning. We also briefly covered the field of relational probabilistic reasoning and presented Bayesian logic programs and Markov logic networks. The fields of probabilistic inductive logic programming and statistical relational learning combine probabilistic reasoning with first-order logic (or subsets thereof) in order to be able to reason with uncertainty on relational domains.

In the following two chapters, we continue investigating (propositional) probabilistic conditional logic. While reasoning on consistent knowledge bases has been looked into for quite some time we discuss the problem of inconsistencies from both an analytical and a practical perspective. We pick up on the problem of relational probabilistic reasoning in Chapter 5.

3 MEASURING INCONSISTENCY

Inconsistencies arise easily when experts share their beliefs in order to build a joint knowledge base. Although these inconsistencies often affect only a little portion of the knowledge base or emerge from only little differences in the experts' beliefs, they cause severe damage. In particular, for knowledge bases that use classical logic for knowledge representation, inconsistencies render the whole knowledge base useless, due to the well-known principle *ex falso quod libet*. Therefore reasoning under inconsistency is an important field in artificial intelligence and there are many proposals to deal with inconsistency in classical logic, e. g. (Rescher and Manor, 1970; Konieczny *et al.*, 2005), or in other logical frameworks, e. g. paraconsistent logics (Belnap, 1976, 1977; Béziau *et al.*, 2007), default logics (Reiter, 1980; Antoniou, 1999), defeasible logics (Nute, 1994), and argumentation theory (Bench-Capon and Dunne, 2007; Rahwan and Simari, 2009). Furthermore, there are several approaches to analyze and measure inconsistency in classical frameworks, e. g. (Knight, 2001; Grant and Hunter, 2006; Hunter and Konieczny, 2010), and some in quantitative frameworks (mainly possibilistic frameworks), e. g. (Dubois *et al.*, 1992).

In this and the next chapter we have a closer look on analyzing inconsistencies in a probabilistic framework and in particular measuring inconsistency in conditional probabilistic knowledge bases, cf. Section 2.3. There are very few works on the treatment of inconsistencies in a conditional probabilistic framework, cf. (Rödder and Xu, 2001; Finthammer *et al.*, 2007; Daniel, 2009). For example, the method described in (Finthammer *et al.*, 2007)—see also (Finthammer, 2008)—consists of a set of heuristics that are used to restore consistency in a knowledge base. Although this method is not based on a theoretical elaboration it works well in real-world examples and has been applied successfully to improve fraud detection in management. Other related work (Hansen and Jaumard, 1996; Andersen and Pretolani, 2001) investigates inconsistencies in classical theories enriched with probabilistic semantics but without treatment of conditional probabilities as we do here.

This chapter is organized as follows. In Section 3.1 we begin by discussing the problem of inconsistencies in probabilistic conditional logic in an abstract way in order to motivate the research conducted in the subsequent sections. In Section 3.2 we approach the problem of measuring inconsistency in a principled way by proposing and discussing a series of rationality postulates for inconsistency measures. In Section 3.3 we recall some established approaches for measuring inconsistency in non-probabilistic logics and discuss extensions of those to the probabilistic conditional framework. It will turn out that these inconsistency measures do not allow for an in-depth analysis of inconsistencies in probabilistic settings due to the quantitative nature of probability theory. Hence, in Section 3.4 we propose a novel inconsistency measure for probabilistic conditional logic that bases on the minimal distance to consistency. We give a complete computational account for implementing the measure and discuss several approximations and extensions. We continue in Section 3.5 with a discussion of some related work and conclude the topic of measuring inconsistency in Section 3.6 with a summary and some final remarks.

3.1 INCONSISTENCIES IN PROBABILISTIC CONDITIONAL LOGIC

When specifying beliefs in some knowledge representation formalism one is usually interested in using these beliefs to perform inference, i. e. to derive new beliefs. In propositional probabilistic conditional logic beliefs are specified with probabilistic conditionals of the form $(\psi \mid \phi)[d]$ with $\phi, \psi \in \mathcal{L}(\mathsf{At})$ and $d \in [0,1]$ being a probability. Reasoning on sets $\mathcal{R}$ of probabilistic conditionals (*knowledge bases*) can be performed using the principle of maximum entropy. As inference based on maximum entropy is a model-based inference technique the existence of at least one model of knowledge base is mandatory. If a knowledge base has no model, model-based inference fails and no further beliefs can be obtained. In order to deal with this problem the first task is to detect inconsistency. In classical logics this problem can be solved using SAT tests, cf. Section 2.1. In probabilistic logics the problem is even more challenging as inconsistencies can appear on both a logical level and on the level of probabilities. Consider the following example.

Example 3.1. Let $\mathcal{R}_1$ be a knowledge base on $\mathsf{At} = \{a, b\}$ given via

$$\mathcal{R}_1 =_{def} \{(a)[0.8], (a)[0.2]\} \quad .$$

The knowledge base $\mathcal{R}_1$ is inconsistent as there can be no probability function P with both $P(a) = 0.8$ and $P(a) = 0.2$. In $\mathcal{R}_1$ the inconsistency arises due to a logical error in representing the beliefs on the probability of a. Consider furthermore $\mathcal{R}_2$ given via

$$\mathcal{R}_2 =_{def} \{(b \mid a)[0.8], (a)[0.6], (b)[0.4]\} \quad .$$

The knowledge base $\mathcal{R}_2$ is inconsistent as well but this is quite harder to recognize as for $\mathcal{R}_1$. By just considering $(a)[0.6]$ and $(b \mid a)[0.8]$ one can derive that $P(a \wedge b) = 0.48$ for any P with $P \models^{pr} (a)[0.6]$ and $P \models^{pr} (b \mid a)[0.8]$. As for every P it holds that $P(b) \geq P(a \wedge b)$, there can be no model for $\mathcal{R}_2$. Here, the inconsistency arises on the level of probabilities.

However, there is no clear distinction between logical inconsistencies and inconsistencies on the level of probabilities as even in $\mathcal{R}_1$ the problem actually lies in the probabilities. And the larger the knowledge base the harder it is even for an experienced knowledge engineer to ensure consistency. Hence, for the rest of this section we give a simple computational account for determining whether a given probabilistic conditional knowledge base is consistent.

The problem of determining whether a knowledge base is consistent can be reduced to a constraint satisfaction problem similar to the approaches in (Hansen and Jaumard, 1996; Andersen and Pretolani, 2001) which consider purely propositional knowledge bases without conditionals. Let $\mathcal{R} = \{r_1, \ldots, r_n\}$ with $r_i = (\psi_i \,|\, \phi_i)[d_i]$ for $i = 1, \ldots, n$ be a knowledge base on a propositional signature At. Remember that we require every $r \in \mathcal{R}$ to be self-consistent and non-tautological. In order to verify that $\mathcal{R}$ is consistent we have to find a probability function that satisfies $\mathcal{R}$. For a probability function $P : \Omega(\mathsf{At}) \to [0,1]$ to be a model of $\mathcal{R}$, every r_i imposes $P(\psi_i\phi_i) = d_i P(\phi_i)$ to hold (for $i = 1, \ldots, n$). For every $\omega \in \Omega(\mathsf{At})$ define $\alpha_\omega =_{def} P(\omega)$. If P satisfies r_i this amounts to

$$\sum_{\omega \in \mathsf{Mod}^{\mathrm{P}}(\psi_i\phi_i)} \alpha_\omega = d_i \cdot \sum_{\omega \in \mathsf{Mod}^{\mathrm{P}}(\phi_i)} \alpha_\omega \tag{3.1}$$

which is equivalent to

$$\sum_{\omega \in \mathsf{Mod}^{\mathrm{P}}(\psi_i\phi_i)} (1 - d_i)\alpha_\omega - \sum_{\omega \in \mathsf{Mod}^{\mathrm{P}}(\overline{\psi_i}\phi_i)} d_i\alpha_\omega = 0 \quad . \tag{3.2}$$

Given a knowledge base $\mathcal{R}$ the problem of checking whether $\mathcal{R}$ is consistent then reduces to finding values $\mathcal{A} = \{\alpha_\omega \in [0,1] \mid \omega \in \Omega(\mathsf{At})\}$ that fulfill Equation (3.2) for every conditional $r_i \in \mathcal{R}$. In order to ensure that $\mathcal{A}$ represents a probability function the integrity constraints

$$\sum_{\omega \in \Omega(\mathsf{At})} \alpha_\omega = 1 \quad \text{and} \tag{3.3}$$

$$\alpha_\omega \geq 0 \quad \text{for all } \omega \in \Omega(\mathsf{At}) \tag{3.4}$$

have to be respected as well.

Taken together Equation (3.2) for $i = 1, \ldots, n$ and the equations (3.3) and (3.4), this yields a constraint satisfaction problem $\mathsf{Cons}(\mathcal{R})$ on the variables $\{\alpha_\omega \mid \omega \in \Omega(\mathsf{At})\}$. It is clear that every assignment of values to the variables α_ω, that is legal with respect to the constraint satisfaction problem $\mathsf{Cons}(\mathcal{R})$, directly corresponds to a probability function $P^{\mathsf{Cons}(\mathcal{R})}$ with $P^{\mathsf{Cons}(\mathcal{R})}(\omega) =_{def} \alpha_\omega$. Hence, if there is an assignment for all α_ω the corresponding probability function $P^{\mathsf{Cons}(\mathcal{R})}$ is a model for all conditionals $r_i \in \mathcal{R}$ $(i = 1, \ldots, n)$ and therefore a model for $\mathcal{R}$. Therefore, we have proven the following statement.

Proposition 3.1. *A knowledge base $\mathcal{R}$ is consistent if and only if* $\mathsf{Cons}(\mathcal{R})$ *has a solution.*

As a side note we remark that the problem of deciding whether a given knowledge base is consistent is NP-complete, cf. e. g. (Paris, 1994).

Looking at equations (3.2), (3.3), and (3.4) one also gains a geometrical interpretation of the problem of consistency. Each probabilistic conditional r can be interpreted as a hyperplane H_r in $\mathbb{R}^{|\Omega(\mathsf{At})|}$, i. e. the hyperplane that contains exactly the set of vectors of the form $(\alpha_{\omega_1},\ldots,\alpha_{\omega_m})$ for $\Omega(\mathsf{At}) = \{\omega_1,\ldots,\omega_m\}$ that satisfy Equation (3.2). Furthermore, the normalization constraint (3.3) describes a hyperplane H_0 as well, called the *normalization hyperplane*. Then, if $\mathbb{R}_0^+$ is the set of non-negative real numbers and $\mathcal{R} = \{r_1,\ldots,r_n\}$ is a knowledge base the problem of deciding whether $\mathcal{R}$ is consistent is equivalent to the problem of deciding whether $H_{r_1} \cap \ldots \cap H_{r_n} \cap H_0 \cap (\mathbb{R}_0^+)^{|\Omega(\mathsf{At})|} \neq \emptyset$, i. e., whether the intersection of all hyperplanes of the probabilistic conditionals and the normalization hyperplane is non-empty in $(\mathbb{R}_0^+)^{|\Omega(\mathsf{At})|}$.

Just knowing that knowledge base $\mathcal{R}$ is inconsistent is—in general—not sufficient for knowledge engineering and analyzing. In order to make an inconsistent knowledge base usable it is necessary to remove the inconsistencies, i. e. to modify the modeled beliefs appropriately. Modifying an inconsistent knowledge base adheres for rationality postulates such as *minimal change*, cf. e. g. (Konieczny and Pino-Pérez, 1998; Hansson, 1999). Before coming to the actual issue of restoring consistency (see Chapter 4) we first continue with an investigation on *analyzing* inconsistencies. There is much work on analyzing inconsistency in qualitative frameworks, see e. g. (Knight, 2001; Grant and Hunter, 2006; Hunter and Konieczny, 2010), but there are only few works on analyzing inconsistency in quantitative frameworks, especially in probabilistic frameworks as the one discussed here, cf. (Rödder and Xu, 2001; Daniel, 2009).

We take a formal approach for the analysis of inconsistency by formalizing and developing *inconsistency measures* on probabilistic knowledge bases. In a nutshell, an inconsistency measure is a function Inc that maps a knowledge base $\mathcal{R}$ to a non-negative real number $\mathsf{Inc}(\mathcal{R})$ that quantifies the severity or amount of the inconsistency in $\mathcal{R}$. We go on by stating some desirable properties for inconsistency measures.

3.2 DESIRABLE PROPERTIES FOR AN INCONSISTENCY MEASURE

In this thesis we take a principled approach for the discussion on inconsistency measures on probabilistic conditional knowledge bases. In the following, we propose several properties and argue that any reasonable inconsistency measure should fulfill these properties. Let At be a propositional signature and Inc be a function

$$\mathsf{Inc} : \mathfrak{P}((\mathcal{L}(\mathsf{At}) \mid \mathcal{L}(\mathsf{At}))^{pr}) \to [0,\infty)$$

that maps a knowledge base $\mathcal{R} \subseteq (\mathcal{L}(\mathsf{At}) \mid \mathcal{L}(\mathsf{At}))^{pr}$ onto a non-negative real number. We say that Inc is an *inconsistency measure* and the value $\mathsf{Inc}(\mathcal{R})$ for a knowledge base $\mathcal{R}$ is called the *inconsistency value* for $\mathcal{R}$ with respect to Inc. When talking about inconsistency values we omit the reference to the inconsistency measure when this is clear from context.

Some of the following properties are inspired by similar properties demanded for inconsistency measures on classical knowledge bases and have been modified to fit a probabilistic framework, see e. g. (Hunter and Konieczny, 2006, 2008). Intuitively, we want Inc to be a function on knowledge bases that is monotonically increasing with the inconsistency in the knowledge base. If the knowledge base is consistent, Inc shall be minimal. For the upcoming definitions let $\mathcal{R}, \mathcal{R}_1, \mathcal{R}_2 \subseteq (\mathcal{L}(\mathsf{At}) \mid \mathcal{L}(\mathsf{At}))^{pr}$ be knowledge bases and $r, r' \in (\mathcal{L}(\mathsf{At}) \mid \mathcal{L}(\mathsf{At}))^{pr}$ be probabilistic conditionals.

Our first property relates to the case of a consistent $\mathcal{R}$. In this case Inc should take the minimal value and not distinguish between different consistent knowledge bases. This is essentially different to the basic property of information measures which do not distinguish between different inconsistent knowledge bases, cf. (Cover and Thomas, 2006). Furthermore, for every inconsistent knowledge base a strict-positive inconsistency value is expected.

(Consistency) $\mathcal{R}$ is consistent if and only if $\mathsf{Inc}(\mathcal{R}) = 0$.

Another basic demand for an inconsistency measure is indifference to syntactical variants.

(Irrelevance of Syntax) If $\mathcal{R}_1 \equiv^{\text{cond}} \mathcal{R}_2$ then $\mathsf{Inc}(\mathcal{R}_1) = \mathsf{Inc}(\mathcal{R}_2)$.

Note that we used the relation $\equiv^{\text{cond}}$ to define (Irrelevance of Syntax). As by Proposition 2.4 on page 29 all inconsistent knowledge bases are equivalent with respect to $\equiv^{\text{kb}}$ demanding (Irrelevance of syntax) with respect to $\equiv^{\text{kb}}$ amounts to assigning the same degree of inconsistency to every inconsistent knowledge base.

Next, we consider properties that describe how inconsistency values may change when new information is added. For one, by adding a new piece of information to a knowledge base the inconsistency should not decrease.

(Monotonicity) It holds that $\mathsf{Inc}(\mathcal{R}) \leq \mathsf{Inc}(\mathcal{R} \cup \{r\})$.

A similar demand for a non-decreasing inconsistency measure derives from joining disjoint knowledge bases.

(Super-Additivity) If $\mathcal{R} \cap \mathcal{R}' = \emptyset$ then $\mathsf{Inc}(\mathcal{R} \cup \mathcal{R}') \geq \mathsf{Inc}(\mathcal{R}) + \mathsf{Inc}(\mathcal{R}')$.

Note that (Super-Additivity) is the stronger property, as it can be easily seen that (Super-Additivity) implies (Monotonicity).

Proposition 3.2. *If* Inc *satisfies (Super-Additivity) then* Inc *satisfies (Monotonicity).*

Proof. Let Inc satisfy (Super-Additivity). If $r \in \mathcal{R}$ then $\mathsf{Inc}(\mathcal{R}) = \mathsf{Inc}(\mathcal{R} \cup \{r\})$. If $r \notin \mathcal{R}$ then $\mathsf{Inc}(\mathcal{R} \cup \{r\}) \geq \mathsf{Inc}(\mathcal{R}) + \mathsf{Inc}(\{r\}) \geq \mathsf{Inc}(\mathcal{R})$ due to (Super-Additivity). □

As mentioned before, we assume that every probabilistic conditional r is *self-consistent*, i. e., for a probabilistic conditional r there is a probability function P with $P \models^{pr} r$, cf. Section 2.3. Considering a probabilistic conditional r that is *independent* of the beliefs in $\mathcal{R}$ should not increase the inconsistency. The simplest form of independence are disjoint languages. Let $\mathsf{At}(S) \subseteq \mathsf{At}$ denote the set of atoms appearing in a set S of probabilistic conditionals.

(Weak Independence) If $\mathsf{At}(\{r\}) \cap \mathsf{At}(\mathcal{R}) = \emptyset$ then $\mathsf{Inc}(\mathcal{R}) = \mathsf{Inc}(\mathcal{R} \cup \{r\})$.

The property (Weak Independence) can be generalized by not only considering different language biases but the actual part that probabilistic conditionals play in creating inconsistencies.

Definition 3.1 (Minimal inconsistent subset)**.** Let $\mathcal{R}$ be a knowledge base. A *minimal inconsistent subset* $\mathcal{M}$ of $\mathcal{R}$ is a set $\mathcal{M} \subseteq \mathcal{R}$ such that $\mathcal{M}$ is inconsistent and every $\mathcal{M}' \subsetneq \mathcal{M}$ is consistent. Let $\mathsf{MI}(\mathcal{R})$ denote the set of all minimal inconsistent subsets of $\mathcal{R}$.

Definition 3.2 (Free conditional)**.** A conditional $r \in \mathcal{R}$ is called a *free conditional* in $\mathcal{R}$ if and only if for every $\mathcal{M} \in \mathsf{MI}(\mathcal{R})$ it holds that $r \notin \mathcal{M}$.

Free conditionals feature the following property regarding consistency.

Proposition 3.3. *Let $\mathcal{R}$ be a knowledge base and let $r \notin \mathcal{R}$ be a free conditional in $\mathcal{R} \cup \{r\}$. Then it holds that*

$$\mathcal{R} \text{ is consistent} \quad \text{iff} \quad \mathcal{R} \cup \{r\} \text{ is consistent} \quad .$$

Proof. If $\mathcal{R}$ is consistent and r is a free conditional in $\mathcal{R} \cup \{r\}$ then for every minimal inconsistent subset $\mathcal{M}$ of $\mathcal{R} \cup \{r\}$ it holds that $r \notin \mathcal{M}$. Then $\mathcal{M}$ would also be a minimal inconsistent subset of $\mathcal{R}$ contradicting the premise that $\mathcal{R}$ is consistent. The reverse direction holds as well as any subset of a consistent knowledge base is consistent. □

Using the notion of free conditionals we can strengthen the above property as follows.

(Independence) If r is a free conditional in $\mathcal{R} \cup \{r\}$, then it holds that $\mathsf{Inc}(\mathcal{R}) = \mathsf{Inc}(\mathcal{R} \cup \{r\})$.

Inconsistency measures that satisfy (Independence) feature the following property.

Proposition 3.4. *Let $\mathcal{R}_1, \mathcal{R}_2$ be knowledge bases and let* Inc *be an inconsistency measure that satisfies (Independence). If* $\mathsf{MI}(\mathcal{R}_1) = \mathsf{MI}(\mathcal{R}_2)$ *then* $\mathsf{Inc}(\mathcal{R}_1) = \mathsf{Inc}(\mathcal{R}_2)$.

Proof. Let $\mathcal{R} =_{def} \bigcup_{\mathcal{M} \in \mathsf{MI}(\mathcal{R}_1)} \mathcal{M}$. It holds that $\mathsf{Inc}(\mathcal{R}_1) = \mathsf{Inc}(\mathcal{R})$ due to the facts that $\mathcal{R}_1 \setminus \mathcal{R}$ only contains free conditionals of $\mathcal{R}_1$ and that Inc satisfies (Independence). As the same is true for $\mathcal{R}_2$ it follows $\mathsf{Inc}(\mathcal{R}_1) = \mathsf{Inc}(\mathcal{R}_2)$. □

Hence, inconsistency measures satisfying (Independence) can be defined solely on the set of minimal inconsistent subsets. This resembles the notion of MinInc inconsistency values in (Hunter and Konieczny, 2008).

The set of minimal inconsistent subsets is monotonically increasing with respect to larger knowledge bases.

Proposition 3.5. *Let $\mathcal{R}$ and $\mathcal{R}'$ be knowledge bases. If $\mathcal{R} \subseteq \mathcal{R}'$ then* $\mathsf{MI}(\mathcal{R}) \subseteq \mathsf{MI}(\mathcal{R}')$.

Proof. Let $\mathcal{M} \in \mathsf{MI}(\mathcal{R})$ be a minimal inconsistent subset of $\mathcal{R}$. Then it holds that $\mathcal{M} \subseteq \mathcal{R} \subseteq \mathcal{R}'$. Suppose $\mathcal{M} \notin \mathsf{MI}(\mathcal{R}')$ which is equivalent to stating that either $\mathcal{M}$ is not minimal or not inconsistent. Both cases contradict the assumption $\mathcal{M} \in \mathsf{MI}(\mathcal{R})$. □

Satisfaction of (Independence) implies satisfaction of (Weak Independence) as the following results show.

Lemma 3.1. *Let $\mathcal{R}$ be a knowledge base and let r be a probabilistic conditional with $r \notin \mathcal{R}$. If* $\mathsf{At}(\{r\}) \cap \mathsf{At}(\mathcal{R}) = \emptyset$ *then r is a free conditional in $\mathcal{R} \cup \{r\}$.*

The proof of Lemma 3.1 can be found in Appendix A on page 230.

Proposition 3.6. *If* Inc *satisfies (Independence) then* Inc *satisfies (Weak Independence).*

Proof. Let Inc satisfy (Independence), $\mathcal{R}$ be a knowledge base, and let r be a conditional with $\mathsf{At}(\{r\}) \cap \mathsf{At}(\mathcal{R}) = \emptyset$. As r is a free conditional in $\mathcal{R} \cup \{r\}$ by Lemma 3.1 it follows $\mathsf{Inc}(\mathcal{R} \cup \{r\}) = \mathsf{Inc}(\mathcal{R})$ by (Independence) and, hence, Inc satisfies (Weak Independence). □

The following property is from (Hunter and Konieczny, 2008).

(MinInc Separability) If $\mathsf{MI}(\mathcal{R}_1 \cup \mathcal{R}_2) = \mathsf{MI}(\mathcal{R}_1) \cup \mathsf{MI}(\mathcal{R}_2)$ and $\mathsf{MI}(\mathcal{R}_1) \cap \mathsf{MI}(\mathcal{R}_2) = \emptyset$ then $\mathsf{Inc}(\mathcal{R}_1 \cup \mathcal{R}_2) = \mathsf{Inc}(\mathcal{R}_1) + \mathsf{Inc}(\mathcal{R}_2)$.

The above property states that the inconsistency value of some knowledge base $\mathcal{R}_1 \cup \mathcal{R}_2$ is the sum of the inconsistency values of $\mathcal{R}_1$ and $\mathcal{R}_2$ if the set of minimal inconsistent subsets of $\mathcal{R}_1 \cup \mathcal{R}_2$ is partitioned among $\mathcal{R}_1$ and $\mathcal{R}_2$. The property (MinInc Separability) is an even more general property than (Independence).

Proposition 3.7. *If* Inc *satisfies (*MinInc *Separability) then* Inc *satisfies (Independence).*

Proof. Let Inc satisfy (MinInc Separability), let $\mathcal{R}$ be a knowledge base, and let r be a free conditional in $\mathcal{R} \cup \{r\}$. Note that it holds that $\mathsf{MI}(\{r\}) = \emptyset$ as r is self-consistent. Then it also holds that $\mathsf{MI}(\mathcal{R} \cup \{r\}) = \mathsf{MI}(\mathcal{R}) = \mathsf{MI}(\mathcal{R}) \cup \mathsf{MI}(\{r\})$ and $\mathsf{MI}(\mathcal{R}) \cap \mathsf{MI}(\{r\}) = \emptyset$. By (MinInc separability) it follows that $\mathsf{Inc}(\mathcal{R} \cup \{r\}) = \mathsf{Inc}(\mathcal{R}) + \mathsf{Inc}(\{r\}) = \mathsf{Inc}(\mathcal{R})$. □

The previous properties describe cases where the inconsistency of a knowledge base should remain constant despite the addition of new information. Conversely, the next property describes a case when the inconsistency should increase.

(Penalty) If $r \notin \mathcal{R}$ is not a free conditional in $\mathcal{R} \cup \{r\}$ then $\mathsf{Inc}(\mathcal{R}) < \mathsf{Inc}(\mathcal{R} \cup \{r\})$.

Similar to the motivation for (Independence) we state that if a conditional r contributes to a minimal inconsistent subset of the knowledge base, then the inconsistency must be strictly greater than in the knowledge base without r. As for (Independence) alone we can state a similar property like in Proposition 3.4 in terms of minimal inconsistent sets when the inconsistency measure satisfies both (Independence) and (Penalty).

Proposition 3.8. *Let* $\mathcal{R}_1$ *and* $\mathcal{R}_2$ *be knowledge bases and let* Inc *be an inconsistency measure that satisfies (Independence) and (Penalty). If* $\mathsf{MI}(\mathcal{R}_1) \subsetneq \mathsf{MI}(\mathcal{R}_2)$ *then* $\mathsf{Inc}(\mathcal{R}_1) < \mathsf{Inc}(\mathcal{R}_2)$.

Proof. Let $\mathcal{R} =_{def} \bigcup_{\mathcal{M} \in \mathsf{MI}(\mathcal{R}_1)} \mathcal{M}$. It holds that $\mathsf{Inc}(\mathcal{R}_1) = \mathsf{Inc}(\mathcal{R})$ due to the facts that $\mathcal{R}_1 \setminus \mathcal{R}$ only contains free conditionals of $\mathcal{R}_1$ and that Inc satisfies (Independence). As $\mathcal{R} \subsetneq \mathcal{R}_2$ due to $\mathsf{MI}(\mathcal{R}_1) \subsetneq \mathsf{MI}(\mathcal{R}_2)$ and $\mathcal{R}_2 \setminus \mathcal{R}$ contains at least one conditional r that is not free in $\mathcal{R}_2$—otherwise it would be $\mathsf{MI}(\mathcal{R}_1) = \mathsf{MI}(\mathcal{R}_2)$— it follows that $\mathsf{Inc}(\mathcal{R}_2) > \mathsf{Inc}(\mathcal{R}) = \mathsf{Inc}(\mathcal{R}_1)$ as Inc satisfies (Penalty). □

So far we have not taken into account that we are working in a probabilistic framework. It is hard to grasp in what way the probabilities of the conditionals influence the inconsistency of the whole knowledge base. Consider a knowledge base $\mathcal{R}$ and a conditional $(\psi \,|\, \phi)[d] \in \mathcal{R}$. How should the inconsistency measure Inc behave when increasing (or decreasing) the value d? There is no definite answer to this question as, on the one hand, the inconsistency may vanish because the conditional may become consistent with the rest of the knowledge base, or, on the other hand, the inconsistency may increase because the conditional may remove itself from a "consistent state". But one demand can be made: The change in the inconsistency value should be continuous in d. If one does only slightly change a given knowledge base, the resulting inconsistency value should have only changed slightly as well. We formalize this intuition as follows.

Definition 3.3 (Characteristic function)**.** Let $\mathcal{R}$ be a knowledge base with $\langle\mathcal{R}\rangle = (r_1, \ldots, r_n)$ and $r_i = (\psi_i \mid \phi_i)[d_i]$ for $i = 1, \ldots, n$. The function

$$\Lambda_{\mathcal{R}} : [0,1]^{|\mathcal{R}|} \to \mathfrak{P}((\mathcal{L}(\mathsf{At}) \mid \mathcal{L}(\mathsf{At}))^{pr})$$

with $\Lambda_{\mathcal{R}}(x_1, \ldots, x_n) =_{def} \{(\psi_1 \mid \phi_1)[x_1], \ldots, (\psi_n \mid \phi_n)[x_n]\}$ and $\langle\Lambda_{\mathcal{R}}(x_1, \ldots, x_n)\rangle =_{def} ((\psi_1 \mid \phi_1)[x_1], \ldots, (\psi_n \mid \phi_n)[x_n])$ is called the *characteristic function* of $\mathcal{R}$.

Note that in the above definition we used the vector representation of $\mathcal{R}$ in order to give a sound definition of the characteristic function, cf. Definition 2.25 on page 27.

Definition 3.4 (Characteristic inconsistency function)**.** Let Inc be an inconsistency measure and let $\mathcal{R}$ be a knowledge base. The function

$$\theta_{\mathsf{Inc},\mathcal{R}} : [0,1]^{|\mathcal{R}|} \to [0, \infty)$$

with $\theta_{\mathsf{Inc},\mathcal{R}} =_{def} \mathsf{Inc} \circ \Lambda_{\mathcal{R}}$ is called the *characteristic inconsistency function* of Inc and $\mathcal{R}$.

The above definitions allow us to state the next property in a concise way as follows.

(Continuity) The characteristic inconsistency function $\theta_{\mathsf{Inc},\mathcal{R}}$ is continuous on $[0,1]^{|\mathcal{R}|}$ (with respect to the standard topology on $\mathbb{R}^{|\mathcal{R}|}$).

Our final property concerns normalization of the inconsistency measure.

(Normalization) For every $\mathcal{R}$ it holds that $\mathsf{Inc}(\mathcal{R}) \in [0,1]$.

The above property states that inconsistency values should be bounded from above by one. On the one hand, this demand makes perfect sense as this allows for comparing inconsistency values of different knowledge base in a unified way. On the other hand, this demand is—in general—in conflict with the demand for (Super-Additivity) as the following example shows.

Example 3.2. Consider the probabilistic conditionals

$$r_1^k =_{def} (a_k)[0.6] \qquad\qquad r_2^k =_{def} (a_k)[0.4]$$

on a propositional signature $\mathsf{At}_i =_{def} \{a_1, \ldots, a_i\}$. Obviously, the knowledge base $\{r_1^1, r_2^1\}$ on At_1 is inconsistent and therefore some inconsistency measure Inc satisfying (Consistency) assigns some degree of inconsistency to $\{r_1^1, r_2^1\}$, i.e. $\mathsf{Inc}(\{r_1^1, r_2^1\}) = x > 0$. Furthermore, any knowledge base $\{r_1^i, r_2^i\}$ on $\{a_i\}$ is inconsistent as well and should be assigned the same inconsistency value, i.e. $\mathsf{Inc}(\{r_1^i, r_2^i\}) = x$. It follows that, if Inc satisfies

(Super-Additivity) and does not take the size of signature of a knowledge base into account then there is a natural number $n \in \mathbb{N}$ such that for $\mathcal{R}_n = \{r_1^1, r_2^1, \ldots, r_1^n, r_2^n\}$ it holds that $\mathsf{Inc}(\mathcal{R}_n) \geq \mathsf{Inc}(\{r_1^1, r_2^1\}) + \ldots + \mathsf{Inc}(\{r_1^n, r_2^n\}) \geq nx > 1$. Thus, Inc cannot satisfy (Normalization).

The previous example showed that an inconsistency measure that does not take (the size of) the signature into account cannot satisfy (Consistency), (Super-Additivity), and (Normalization) at the same time. Furthermore, taking (the size of) the signature into account may become unintuitive. As for the case of Example 3.2, in order to allow Inc to satisfy (Consistency), (Super-Additivity), and (Normalization) it has to hold that for $\mathcal{R} =_{def} \{r_1^1, r_1^1\}$ defined on At_1 and $\mathcal{R}' =_{def} \{r_1^1, r_1^1\}$ defined on At_2 it follows that $\mathsf{Inc}(\mathcal{R}) \neq \mathsf{Inc}(\mathcal{R}')$. As $\mathcal{R} = \mathcal{R}'$ this result is obviously very unintuitive. Due to this discrepancy, for each approach for measuring inconsistency we discuss variants that satisfy (Normalization) and variants that do not.

To conclude this section, Table 2 gives an overview on the properties of inconsistency measures that have been discussed above. In the next section we continue with investigating traditional approaches for measuring inconsistency in non-probabilistic frameworks and their applicability for probabilistic conditional logic.

Property	Description
(Consistency)	$\mathcal{R}$ consistent iff $\mathsf{Inc}(\mathcal{R}) = 0$
(Irrelevance of Syntax)	$\mathcal{R}_1 \equiv^{\text{cond}} \mathcal{R}_2$ implies $\mathsf{Inc}(\mathcal{R}_1) = \mathsf{Inc}(\mathcal{R}_2)$
(Monotonicity)	$\mathsf{Inc}(\mathcal{R}) \leq \mathsf{Inc}(\mathcal{R} \cup \{r\})$
(Super-Additivity)	If $\mathcal{R} \cap \mathcal{R}' \neq \emptyset$ then $\mathsf{Inc}(\mathcal{R} \cup \mathcal{R}') \geq \mathsf{Inc}(\mathcal{R}) + \mathsf{Inc}(\mathcal{R}')$
(Weak Independence)	If $\mathsf{At}(\{r\}) \cap \mathsf{At}(\mathcal{R}) \neq \emptyset$ then $\mathsf{Inc}(\mathcal{R}) = \mathsf{Inc}(\mathcal{R} \cup \{r\})$
(Independence)	If r free in $\mathcal{R} \cup \{r\}$ then $\mathsf{Inc}(\mathcal{R}) = \mathsf{Inc}(\mathcal{R} \cup \{r\})$
(MinInc Separability)	If $\mathsf{MI}(\mathcal{R}_1 \cup \mathcal{R}_2) = \mathsf{MI}(\mathcal{R}_1) \cup \mathsf{MI}(\mathcal{R}_2)$ and $\mathsf{MI}(\mathcal{R}_1) \cap \mathsf{MI}(\mathcal{R}_2) = \emptyset$ then $\mathsf{Inc}(\mathcal{R}_1 \cup \mathcal{R}_2) = \mathsf{Inc}(\mathcal{R}_1) + \mathsf{Inc}(\mathcal{R}_2)$
(Penalty)	If $r \notin \mathcal{R}$ not free in $\mathcal{R} \cup \{r\}$ then $\mathsf{Inc}(\mathcal{R}) < \mathsf{Inc}(\mathcal{R} \cup \{r\})$
(Continuity)	$\theta_{\mathsf{Inc},\mathcal{R}}$ is continuous
(Normalization)	$\mathsf{Inc}(\mathcal{R}) \in [0, 1]$

Table 2: Properties of inconsistency measures

3.3 TRADITIONAL APPROACHES FOR MEASURING INCONSISTENCY

In this section we review some existing approaches to measuring inconsistencies in classical theories and adapt them to fit our probabilistic framework.

The simplest approach for an inconsistency measure is the *drastic inconsistency measure*, cf. (Hunter and Konieczny, 2010).

Definition 3.5 (Drastic inconsistency measure)**.** Let Inc^d be the function $\mathsf{Inc}^d : \mathfrak{P}((\mathcal{L}(\mathsf{At}) \mid \mathcal{L}(\mathsf{At}))^{pr}) \to [0, \infty)$ defined as

$$\mathsf{Inc}^d(\mathcal{R}) =_{def} \begin{cases} 0 & \text{if } \mathcal{R} \text{ is consistent} \\ 1 & \text{if } \mathcal{R} \text{ is inconsistent} \end{cases} \quad .$$

Inc^d is called the *drastic inconsistency measure*.

The drastic inconsistency measure allows only for a binary decision on inconsistencies and does not quantify the severity of inconsistencies. Although being a very simple inconsistency measure Inc^d still satisfies several basic properties as the next proposition shows.

Proposition 3.9. *The function* Inc^d *satisfies (Consistency), (Irrelevance of Syntax), (Monotonicity), (Weak Independence), (Independence), and (Normalization).*

The proof of Proposition 3.9 can be found in Appendix A on page 231.

Notice, that Inc^d satisfies neither (Super-Additivity), (Penalty), (MinInc Separability), nor (Continuity).

Example 3.3. Consider the knowledge bases $\mathcal{R}_1 =_{def} \{r_1, r_2\}$ and $\mathcal{R}_2 =_{def} \{r_3, r_4\}$ given via

$$r_1 =_{def} (a)[0.4] \quad r_2 =_{def} (a)[0.6] \quad r_3 =_{def} (b)[0.4] \quad r_4 =_{def} (b)[0.6] \quad .$$

It follows that $\mathsf{Inc}^d(\mathcal{R}_1) = \mathsf{Inc}^d(\mathcal{R}_2) = 1$ but

$$\mathsf{Inc}^d(\mathcal{R}_1 \cup \mathcal{R}_2) = 1 \neq \mathsf{Inc}^d(\mathcal{R}_1) + \mathsf{Inc}^d(\mathcal{R}_2) \quad ,$$

therefore violating both (Super-Additivity) and (MinInc Separability). Furthermore, r_4 is not a free conditional in $\mathcal{R}_1 \cup \mathcal{R}_2$ but $\mathsf{Inc}^d(\mathcal{R}_1 \cup \mathcal{R}_2 \setminus \{r_4\}) = \mathsf{Inc}^d(\mathcal{R}_1 \cup \mathcal{R}_2)$ violating (Penalty). Also, Inc^d fails to satisfy (Continuity) as $\mathrm{Im}\ \mathsf{Inc}^d = \{0, 1\}$.

One thing to note is that Inc_d is the upper bound for any inconsistency measure that satisfies (Consistency) and (Normalization).

Proposition 3.10. *If* Inc *satisfies (Consistency) and (Normalization) then for every* $\mathcal{R}$ *it holds that* $\mathsf{Inc}(\mathcal{R}) \leq \mathsf{Inc}^d(\mathcal{R})$*.*

Proof. If $\mathcal{R}$ is consistent it holds that $\mathsf{Inc}(\mathcal{R}) = \mathsf{Inc}^d(\mathcal{R}) = 0$ as both Inc and Inc^d satisfy (Consistency). If $\mathcal{R}$ is inconsistent then $\mathsf{Inc}(\mathcal{R}) \in (0,1]$ as Inc satisfies (Normalization) and therefore $\mathsf{Inc}(\mathcal{R}) \leq 1 = \mathsf{Inc}^d(\mathcal{R})$. □

The next inconsistency measure employs the set of minimal inconsistent subsets of $\mathcal{R}$ in a simple manner, cf. (Hunter and Konieczny, 2010).

Definition 3.6 (MI inconsistency measure)**.** Let $\mathsf{Inc}^{\mathsf{MI}}$ be the function $\mathsf{Inc}^{\mathsf{MI}} : \mathfrak{P}((\mathcal{L}(\mathsf{At}) \mid \mathcal{L}(\mathsf{At}))^{pr}) \to [0,\infty)$ defined via

$$\mathsf{Inc}^{\mathsf{MI}}(\mathcal{R}) =_{def} |\, \mathsf{MI}(\mathcal{R})\, | \quad .$$

$\mathsf{Inc}^{\mathsf{MI}}$ is called the *MI inconsistency measure*.

The MI inconsistency measure assesses severity of inconsistencies by considering the number of minimal inconsistent subsets. It follows the intuition that the more minimal inconsistent subsets the greater the inconsistency.

Proposition 3.11. *The function $\mathsf{Inc}^{\mathsf{MI}}$ satisfies (Consistency), (Monotonicity), (Super-Additivity), (Weak Independence), (Independence), (*MinInc *separability), and (Penalty).*

The proof of Proposition 3.11 can be found in Appendix A on page 231.

Notice, that $\mathsf{Inc}^{\mathsf{MI}}$ satisfies neither (Irrelevance of Syntax), (Continuity) nor (Normalization).

Example 3.4. Consider again $\mathcal{R}_1$ and $\mathcal{R}_2$ from Example 3.3 on page 53. It holds that $\mathsf{Inc}^{\mathsf{MI}}(\mathcal{R}_1 \cup \mathcal{R}_2) = 2$ violating (Normalization). Also, $\mathsf{Inc}^{\mathsf{MI}}$ fails to satisfy (Continuity) as $\mathsf{Im}\ \mathsf{Inc}^{\mathsf{MI}} = \mathbb{N}_0$. Consider the knowledge bases $\mathcal{R} =_{def} \{r_1, r_2\}$ and $\mathcal{R}' =_{def} \{r_1, r_2, r_3\}$ given via

$$r_1 =_{def} (a)[0.3] \qquad r_2 =_{def} (a)[0.7] \qquad r_3 =_{def} (\neg a)[0.3] \quad .$$

Then $\mathcal{R} \equiv^{\text{cond}} \mathcal{R}'$ but $\mathsf{Inc}^{\mathsf{MI}}(\mathcal{R}_1) = 1 \neq 2 = \mathsf{Inc}^{\mathsf{MI}}(\mathcal{R}_2)$. Hence, $\mathsf{Inc}^{\mathsf{MI}}$ fails to satisfy (Irrelevance of syntax).

A *normalized* MI inconsistency measure can be defined using the following technical lemma.

Lemma 3.2. *Let M be a set and let $S \subseteq \mathfrak{P}(M)$ such that for all $M_1, M_2 \in S$ with $M_1 \neq M_2$ it holds that $M_1 \not\subseteq M_2$. Then*

$$|S| \leq \binom{|M|}{\lceil \frac{|M|}{2} \rceil} \quad . \tag{3.5}$$

Proof. Let $|M|$ be even and let $S = \{M_1, \ldots, M_n\} \subseteq \mathfrak{P}(M)$ such that for all $M_i, M_j \in S$ with $i \neq j$ it holds that $M_j \not\subseteq M_j$ and let S be maximal (with respect to cardinality) with this property. The maximal cardinality of S is achieved when all $M_i \in S$ have the same cardinality as the problem is symmetric in all M_i. If the cardinality of all M_i is m, then

$$n = \binom{|M|}{m} \quad . \tag{3.6}$$

The above term is maximal for $m = {}^{|M|}/_2$. For M with odd $|M|$ both $m = \lfloor {}^{|M|}/_2 \rfloor$ and $m = \lceil {}^{|M|}/_2 \rceil$ maximize (3.6), so (3.5) captures both cases. □

Corollary 3.1. *Let $\mathcal{R}$ be a knowledge base. Then*

$$|\mathsf{MI}(\mathcal{R})| \leq \gamma_{\mathcal{R}} =_{def} \binom{|\mathcal{R}|}{\lceil \frac{|\mathcal{R}|}{2} \rceil}$$

Proof. For any two distinct $M_1, M_2 \in \mathsf{MI}(\mathcal{R})$ it holds that $M_1 \not\subseteq M_2$ and $M_2 \not\subseteq M_1$ by definition. Then Lemma 3.2 is applicable proving the claim. □

Using the result from Corollary 3.1 we can define a *normalized* MI inconsistency measure as follows.

Definition 3.7 (Normalized MI inconsistency measure)**.** Let $\mathsf{Inc}_0^{\mathsf{MI}}$ be the function $\mathsf{Inc}_0^{\mathsf{MI}} : \mathfrak{P}((\mathcal{L}(\mathsf{At}) \mid \mathcal{L}(\mathsf{At}))^{pr}) \to [0,1]$ defined as

$$\mathsf{Inc}_0^{\mathsf{MI}}(\mathcal{R}) =_{def} \begin{cases} 0 & \text{if } \gamma_{\mathcal{R}} = 0 \\ \frac{\mathsf{Inc}^{\mathsf{MI}}(\mathcal{R})}{\gamma_{\mathcal{R}}} & \text{otherwise} \end{cases} \quad .$$

$\mathsf{Inc}_0^{\mathsf{MI}}$ is called the *normalized MI inconsistency measure*.

By construction, $\mathsf{Inc}_0^{\mathsf{MI}}$ satisfies (Normalization) but fails to satisfy several other properties.

Proposition 3.12. *The function $\mathsf{Inc}_0^{\mathsf{MI}}$ satisfies (Consistency) and (Normalization).*

The proof of Proposition 3.12 can be found in Appendix A on page 232.

$\mathsf{Inc}_0^{\mathsf{MI}}$ fails to satisfy (Monotonicity), (Super-Additivity), (Weak Independence), (Independence), (MinInc Separability), and (Penalty) as the following example shows.

Example 3.5. Consider the knowledge base $\mathcal{R} =_{def} \{r_1, r_2\}$ given via

$$r_1 =_{def} (a)[0.4] \qquad\qquad r_2 =_{def} (a)[0.6] \quad .$$

It follows that

$$\mathsf{Inc}_0^{\mathsf{MI}}(\mathcal{R}) = \frac{\mathsf{Inc}^{\mathsf{MI}}(\mathcal{R})}{\gamma_{\mathcal{R}}} = \frac{1}{\binom{|\mathcal{R}'|}{\lceil \frac{|\mathcal{R}'|}{2} \rceil}} = \frac{1}{2} \quad .$$

Now consider $r_3 =_{def} (b)[0.5]$ and $\mathcal{R}' =_{def} \mathcal{R} \cup \{r_3\}$. Notice that r_3 is a free conditional in $\mathcal{R}$ and does not mention any atom appearing in $\mathcal{R}$. But it follows that

$$\mathsf{Inc}_0^{\mathsf{MI}}(\mathcal{R}') = \frac{\mathsf{Inc}^{\mathsf{MI}}(\mathcal{R}')}{\gamma_{\mathcal{R}'}} = \frac{1}{\binom{|\mathcal{R}'|}{\lceil \frac{|\mathcal{R}'|}{2} \rceil}} = \frac{1}{3} \quad .$$

As one can see $\mathsf{Inc}_0^{\mathsf{MI}}$ violates both (Monotonicity) and (Weak Independence). Due to Propositions 3.2 on page 47, 3.6 on page 49, and 3.7 on page 50 the function $\mathsf{Inc}_0^{\mathsf{MI}}$ cannot satisfy (Super-Additivity), (Independence), and (MinInc Separability) as well. Now consider $r_4 =_{def} (b)[0.4]$ and $\mathcal{R}'' =_{def} \mathcal{R}' \cup \{r_4\}$. Note that r_4 is not a free conditional in $\mathcal{R}''$. It follows that

$$\mathsf{Inc}_0^{\mathsf{MI}}(\mathcal{R}'') = \frac{\mathsf{Inc}^{\mathsf{MI}}(\mathcal{R}'')}{\gamma_{\mathcal{R}''}} = \frac{2}{\binom{|\mathcal{R}''|}{\lceil \frac{|\mathcal{R}''|}{2} \rceil}} = \frac{2}{6} = \frac{1}{3} \quad .$$

As $\mathsf{Inc}_0^{\mathsf{MI}}(\mathcal{R}') = \mathsf{Inc}_0^{\mathsf{MI}}(\mathcal{R}'')$ it follows that $\mathsf{Inc}_0^{\mathsf{MI}}$ violates (Penalty). $\mathsf{Inc}_0^{\mathsf{MI}}$ also fails to satisfy (Irrelevance of Syntax) and (Continuity) as $\mathsf{Inc}^{\mathsf{MI}}$ already fails to satisfy (Irrelevance of Syntax) and (Continuity).

Only considering the number of minimal inconsistent subsets is a too simple approach for assessing inconsistencies. Another indicator for the severity of inconsistencies is the size of minimal inconsistent subsets. A large minimal inconsistent subset means that the inconsistency is distributed over a large number of conditionals. The more conditionals involved in an inconsistency the less severe the inconsistency can be seen. Furthermore, a small minimal inconsistent subset means that the participating conditionals strongly represent contradictory information. Consider the following example for classical logic that can be found in e. g. (Hunter and Konieczny, 2008).

Example 3.6. In a lottery there are n lottery tickets and only one of them is the winner ticket. If w_i denotes the proposition that ticket i will win the lottery then the (classical) formula $\phi =_{def} w_1 \vee \ldots \vee w_n$ can be regarded as true. Furthermore, the *belief* of each ticket buyer i is that he will not win the lottery, i. e., the formula $\phi_i =_{def} \neg w_i$ is regarded to be true for each $i = 1, \ldots, n$. Obviously the merged knowledge base $\{\phi, \phi_1, \ldots, \phi_n\}$ is inconsistent as ϕ demands that one ticket has to win and, hence, one ticket owner k is wrong in assuming $\neg w_k$. However, with increasing number of available tickets the inconsistency becomes negligible and each ticket owner is justified in believing that he will not win.

Although the previous example has been formulated for classical logic the argument stands for probabilistic logics as well. The following inconsistency measure is inspired by (Hunter and Konieczny, 2008) and aims at differentiating between minimal inconsistent sets of different size.

Definition 3.8 (MI^C inconsistency measure)**.** Let $\mathsf{Inc}_C^{\mathsf{MI}}$ be the function $\mathsf{Inc}_C^{\mathsf{MI}}$: $\mathfrak{P}((\mathcal{L}(\mathsf{At})\,|\,\mathcal{L}(\mathsf{At}))^{pr}) \to [0,\infty)$ defined as

$$\mathsf{Inc}_C^{\mathsf{MI}}(\mathcal{R}) =_{def} \sum_{\mathcal{M}\in\mathsf{MI}(\mathcal{R})} \frac{1}{|\,\mathcal{M}\,|} \quad .$$

$\mathsf{Inc}_C^{\mathsf{MI}}$ is called the *MI^C inconsistency measure*.

The MI^C inconsistency measure sums over the reciprocal of the sizes of all minimal inconsistent subsets. In that way, a large minimal inconsistent subset contributes less to the inconsistency value than a small minimal inconsistent subset. As the MI inconsistency measure the MI^C inconsistency measure behaves well with respect to most desirable properties.

Proposition 3.13. *The function $\mathsf{Inc}_C^{\mathsf{MI}}$ satisfies (Consistency), (Monotonicity), (Super-Additivity), (Weak Independence), (Independence), (*MinInc *Separability), and (Penalty).*

The proof of Proposition 3.13 can be found in Appendix A on page 232.

Note that $\mathsf{Inc}_C^{\mathsf{MI}}$ satisfies neither (Irrelevance of Syntax), (Continuity) nor (Normalization).

Example 3.7. Consider the knowledge base $\mathcal{R} =_{def} \{r_1,\ldots,r_6\}$ given via

$$\begin{aligned} r_1 &=_{def} (a)[0.4] & r_2 &=_{def} (a)[0.6] \\ r_3 &=_{def} (b)[0.4] & r_4 &=_{def} (b)[0.6] \\ r_5 &=_{def} (c)[0.4] & r_6 &=_{def} (c)[0.6] \quad . \end{aligned}$$

It follows that $\mathsf{Inc}_C^{\mathsf{MI}}(\mathcal{R}) = 1.5$ thus violating (Normalization). $\mathsf{Inc}_C^{\mathsf{MI}}$ also fails to satisfy (Continuity) as $\mathsf{Im}\ \mathsf{Inc}_C^{\mathsf{MI}} = \mathbb{Q}_0^+$. Consider the knowledge bases $\mathcal{R}' =_{def} \{r_1', r_2'\}$ and $\mathcal{R}'' =_{def} \{r_1', r_2', r_3'\}$ given via

$$r_1' =_{def} (a)[0.3] \qquad r_2' =_{def} (a)[0.7] \qquad r_3' =_{def} (\neg a)[0.3] \quad .$$

Then $\mathcal{R}' \equiv^{\mathrm{cond}} \mathcal{R}''$ but $\mathsf{Inc}_C^{\mathsf{MI}}(\mathcal{R}') = 1/2 \neq 1 = 1/2 + 1/2 = \mathsf{Inc}^{\mathsf{MI}}(\mathcal{R}'')$. Hence, $\mathsf{Inc}_C^{\mathsf{MI}}$ fails to satisfy (Irrelevance of Syntax).

As before we can define a normalized MI^{C} inconsistency measure by exploiting the following observation.

Proposition 3.14. *Let $\mathcal{R}$ be a knowledge base. Then*

$$\mathsf{Inc}^{\mathsf{MI}}_{\mathsf{C}}(\mathcal{R}) \leq 1/2\gamma_{\mathcal{R}} \quad .$$

Proof. As we only consider self-consistent conditionals the minimal size of a minimal inconsistent subset is 2. As $\gamma_{\mathcal{R}}$ is an upper bound for the number of minimal inconsistent subsets of $\mathcal{R}$ (see Corollary 3.1 on page 55) the claim follows. □

Definition 3.9 (Normalized MI^{C} inconsistency measure)**.** Let $\mathsf{Inc}^{\mathsf{MI}}_{\mathsf{C},0}$ be the function $\mathsf{Inc}^{\mathsf{MI}}_{\mathsf{C},0} : \mathfrak{P}((\mathcal{L}(\mathsf{At}) \mid \mathcal{L}(\mathsf{At}))^{pr}) \to [0,1]$ defined as

$$\mathsf{Inc}^{\mathsf{MI}}_{\mathsf{C},0}(\mathcal{R}) =_{def} \begin{cases} 0 & \text{if } \gamma_{\mathcal{R}} = 0 \\ \frac{\mathsf{Inc}^{\mathsf{MI}}_{\mathsf{C}}(\mathcal{R})}{\frac{1}{2}\gamma_{\mathcal{R}}} & \text{otherwise} \end{cases} \quad .$$

$\mathsf{Inc}^{\mathsf{MI}}_{\mathsf{C},0}$ is called the *normalized MI^{C} inconsistency measure*.

Proposition 3.15. *The function $\mathsf{Inc}^{\mathsf{MI}}_{\mathsf{C},0}$ satisfies (Consistency) and (Normalization).*

The proof of Proposition 3.15 can be found in Appendix A on page 233.

Example 3.8. We continue Example 3.5 from page 55. As for $\mathsf{Inc}^{\mathsf{MI}}_{0}$, it holds that

$$\mathsf{Inc}^{\mathsf{MI}}_{\mathsf{C},0}(\mathcal{R}) = 1/2 \quad \text{and} \quad \mathsf{Inc}^{\mathsf{MI}}_{\mathsf{C},0}(\mathcal{R}') = 1/3$$

and therefore $\mathsf{Inc}^{\mathsf{MI}}_{\mathsf{C},0}$ fails to satisfy (Monotonicity) and (Super-Additivity), as well as (Weak Independence), (Independence), and (MinInc Separability). It also holds that $\mathsf{Inc}^{\mathsf{MI}}_{\mathsf{C},0}(\mathcal{R}'') = 1/3$ violating (Penalty). $\mathsf{Inc}^{\mathsf{MI}}_{\mathsf{C},0}$ also fails to satisfy (Irrelevance of Syntax) and (Continuity) as $\mathsf{Inc}^{\mathsf{MI}}_{\mathsf{C}}$ already fails to satisfy (Irrelevance of Syntax) and (Continuity).

Comparing $\mathsf{Inc}^{\mathsf{MI}}_{0}$ and $\mathsf{Inc}^{\mathsf{MI}}_{\mathsf{C},0}$ we can make the following observation.

Proposition 3.16. *It holds that $\mathsf{Inc}^{\mathsf{MI}}_{\mathsf{C},0}(\mathcal{R}) \leq \mathsf{Inc}^{\mathsf{MI}}_{0}(\mathcal{R})$ for every knowledge base $\mathcal{R}$.*

Proof. For $\gamma_{\mathcal{R}} = 0$ it clearly holds that $\mathsf{Inc}^{\mathsf{MI}}_{\mathsf{C},0}(\mathcal{R}) = \mathsf{Inc}^{\mathsf{MI}}_{0}(\mathcal{R}) = 0$. Let $\gamma_{\mathcal{R}} > 0$ then it holds that

$$\mathsf{Inc}^{\mathsf{MI}}_{\mathsf{C},0}(\mathcal{R}) = \frac{\mathsf{Inc}^{\mathsf{MI}}_{\mathsf{C}}(\mathcal{R})}{\frac{1}{2}\gamma_{\mathcal{R}}} = \frac{\sum_{\mathcal{M}\in\mathsf{MI}(\mathcal{R})} \frac{1}{|\mathcal{M}|}}{\frac{1}{2}\gamma_{\mathcal{R}}}$$

$$\leq \frac{|\mathsf{MI}(\mathcal{R})|\frac{1}{2}}{\frac{1}{2}\gamma_{\mathcal{R}}} = \frac{|\mathsf{MI}(\mathcal{R})|}{\gamma_{\mathcal{R}}} = \mathsf{Inc}_0^{\mathsf{MI}}(\mathcal{R}) \qquad \square$$

There are other more sophisticated inconsistency measures for classical theories but most of these cannot be applied to a probabilistic setting in a meaningful manner. Nonetheless, we review the particularly interesting approach of η-consistency (Knight, 2001, 2002) in Section 3.5.

The inconsistency measures Inc^d, $\mathsf{Inc}^{\mathsf{MI}}$, and $\mathsf{Inc}_C^{\mathsf{MI}}$ were initially developed for inconsistency measurement in classical theories and therefore allow only for a "discrete" measurement. Hence, all of the above discussed inconsistency measures do not satisfy (Continuity). But satisfaction of (Continuity) is crucial for an inconsistency measure in probabilistic conditional logic in order to assess inconsistencies in a meaningful manner. Consider the following example.

Example 3.9. Consider again the knowledge base $\mathcal{R}_2 = \{r_1, r_2, r_3\}$ from Example 3.1 on page 44 given via

$$r_1 = (b \mid a)[0.8] \qquad r_2 = (a)[0.6] \qquad r_3 = (b)[0.4] \quad .$$

As pointed out in Example 3.1 the knowledge base is inconsistent and the set of minimal inconsistent subsets is given by $\mathsf{MI}(\mathcal{R}_2) = \{\{r_1, r_2, r_3\}\}$. It follows that

$$\mathsf{Inc}^d(\mathcal{R}_2) = 1 \qquad \mathsf{Inc}^{\mathsf{MI}}(\mathcal{R}_2) = 1 \qquad \mathsf{Inc}_C^{\mathsf{MI}}(\mathcal{R}_2) = \frac{1}{3} \quad .$$

Consider a slight modification $\mathcal{R}'_2 =_{def} \{r'_1, r'_2, r'_3\}$ of $\mathcal{R}_2$ given via

$$r'_1 =_{def} (b \mid a)[0.8] \qquad r_{2'} =_{def} (a)[0.6] \qquad r'_3 =_{def} (b)[0.479] \quad .$$

The knowledge base $\mathcal{R}'_2$ is still inconsistent and it holds that $\mathsf{Inc}(\mathcal{R}'_2) = \mathsf{Inc}(\mathcal{R}_2)$ for $\mathsf{Inc} \in \{\mathsf{Inc}^d, \mathsf{Inc}^{\mathsf{MI}}, \mathsf{Inc}_C^{\mathsf{MI}}\}$. Now consider the knowledge base $\mathcal{R}''_2 =_{def} \{r''_1, r''_2, r''_3\}$ given via

$$r''_1 =_{def} (b \mid a)[0.8] \qquad r'_{2'} =_{def} (a)[0.6] \qquad r''_3 =_{def} (b)[0.48] \quad .$$

The knowledge base $\mathcal{R}''_2$ is consistent and it follows that $\mathsf{Inc}^d(\mathcal{R}''_2) = \mathsf{Inc}^{\mathsf{MI}}(\mathcal{R}''_2) = \mathsf{Inc}_C^{\mathsf{MI}}(\mathcal{R}''_2) = 0$. By comparing $\mathcal{R}'_2$ and $\mathcal{R}''_2$ one can discover only a minor difference of the modeled knowledge. Whereas in $\mathcal{R}''_2$ the proposition b is assigned a probability of 0.48 in $\mathcal{R}'_2$ it is assigned a probability of 0.479. From a practical point of view this difference may be of no relevance and whether b has probability 0.48 or 0.479 may not matter for the intended application. Still, a knowledge engineer may not grasp the harmlessness of the inconsistency in $\mathcal{R}'_2$ as $\mathcal{R}_2$ has the same degree of inconsistency with respect to Inc^d, $\mathsf{Inc}^{\mathsf{MI}}$, and $\mathsf{Inc}_C^{\mathsf{MI}}$.

The above example motivates the need for a more graded approach to measure the inconsistencies in $\mathcal{R}_2$, $\mathcal{R}_2'$, and $\mathcal{R}_2''$. This measure should assign $\mathcal{R}_2'$ a much lower inconsistency value than to $\mathcal{R}_2$ in order to distinguish their severities. In the next section, we continue with the development of an inconsistency measure that aims to satisfy those needs by taking the actual probabilities into account.

3.4 MEASURING INCONSISTENCY BY DISTANCE MINIMIZATION

In this section we develop an inconsistency measure that bases on the minimal distance of an inconsistent knowledge base to a consistent one. We show that this measure is a meaningful inconsistency measure and fulfills many of the desirable properties discussed above. We continue with giving upper and lower bounds for this inconsistency measure that can be efficiently computed and discuss extensions of the measure to more expressive frameworks.

3.4.1 *The General Approach*

The problem of applying traditional approaches to inconsistency measurement for probabilistic knowledge bases lies in the crucial role of probabilities, as discussed at the end of the previous section. As we only consider self-consistent conditionals, inconsistencies can only occur because of "wrong" probabilities. Consider the following proposition.

Proposition 3.17. *Let* $\mathcal{R} = \{(\psi_1 \mid \phi_1), \ldots, (\psi_n \mid \phi_n)\} \subseteq (\mathcal{L}(\mathsf{At}) \mid \mathcal{L}(\mathsf{At}))$ *be a set of conditionals. Then there are* $d_1, \ldots, d_n \in [0,1]$ *such that* $\{(\psi_1 \mid \phi_1)[d_1], \ldots, (\psi_n \mid \phi_n)[d_n]\}$ *is consistent.*

Proof. Consider the uniform probability function P_0 on $\mathcal{L}(\mathsf{At})$. For $i = 1, \ldots, n$, if $\phi_i \not\equiv \perp$ then $P_0(\phi_i) > 0$ and assign $d_i =_{def} P_0(\psi_i \mid \phi_i)$. If $\phi_i \equiv \perp$ then $P_0(\phi_i) = P_0(\psi_i \phi_i) = 0$ and $d_i \in [0,1]$ can be arbitrarily chosen and $P_0 \models^{pr} (\psi_i \mid \phi_i)[d_i]$ is satisfied in any case. It follows that $P_0 \models^{pr} \{(\psi_1 \mid \phi_1)[d_1], \ldots, (\psi_n \mid \phi_n)[d_n]\}$ and hence the claim. □

Bearing this observation in mind we define the MinDev *inconsistency measure* as follows.

Definition 3.10 (MinDev inconsistency measure)**.** Let $\mathcal{R}$ be a knowledge base with $\langle \mathcal{R} \rangle = (r_1, \ldots, r_n)$ and $r_i = (\psi_i \mid \phi_i)[d_i]$ for $i = 1, \ldots, n$. Let Inc^* be the function $\mathsf{Inc}^* : \mathfrak{P}((\mathcal{L}(\mathsf{At}) \mid \mathcal{L}(\mathsf{At}))^{pr}) \to [0, \infty)$ defined via

$$\mathsf{Inc}^*(\mathcal{R}) =_{def} \min\{|d_1 - x_1| + \ldots + |d_n - x_n| \mid \Lambda_{\mathcal{R}}(x_1, \ldots, x_n) \text{ is consistent}\}. \tag{3.7}$$

Inc^* is called the MinDev *inconsistency measure* (*minimal deviation*).

The above definition presupposes that the minimum in Equation (3.7) exists. In the following, we show that this is indeed the case.

Probability functions can be identified with points in $[0,1]^{|\Omega(\mathsf{At})|}$, i.e., assuming some arbitrary total order on the elements in $\Omega(\mathsf{At})$, $\{\omega_1, \ldots, \omega_m\} = \Omega(\mathsf{At})$, a probability function $P \in \mathcal{P}^{\mathrm{P}}(\mathsf{At})$ is uniquely determined by $(P(\omega_1), \ldots, P(\omega_m)) \in [0,1]^{|\Omega(\mathsf{At})|}$. Therefore, properties relating to convergence behavior in $\mathcal{P}^{\mathrm{P}}(\mathsf{At})$ are interpreted using the standard norm on $[0,1]^{|\Omega(\mathsf{At})|}$.

Lemma 3.3. *The set $\mathcal{P}^{\mathrm{P}}(\mathsf{At})$ is closed.*

Proof. Let $P_i \in \mathcal{P}^{\mathrm{P}}(\mathsf{At})$ for $i \in \mathbb{N}$ be a sequence of probability functions such that $\lim_{i\to\infty} P_i$ exists. In particular, each P_i satisfies

$$\sum_{\omega \in \Omega(\mathsf{At})} P_i(\omega) = 1 \quad .$$

Let now $Q = \lim_{i\to\infty} P_i$, i.e., it holds that $Q(\omega) = \lim_{i\to\infty} P_i(\omega)$ for every $\omega \in \Omega(\omega)$ (point-wise convergence). Then it follows

$$\begin{aligned}\sum_{\omega \in \Omega(\mathsf{At})} Q(\omega) &= \sum_{\omega \in \Omega(\mathsf{At})} \lim_{i\to\infty} P_i(\omega) \\ &= \lim_{i\to\infty} \sum_{\omega \in \Omega(\mathsf{At})} P_i(\omega) \\ &= \lim_{i\to\infty} 1 \\ &= 1\end{aligned}$$

Therefore, it holds that $Q \in \mathcal{P}^{\mathrm{P}}(\mathsf{At})$ and $\mathcal{P}^{\mathrm{P}}(\mathsf{At})$ is a closed set. □

It also holds that $\mathcal{P}^{\mathrm{P}}(\mathsf{At})$ is bounded by $[0,1]^{|\Omega(\mathsf{At})|}$ and therefore $\mathcal{P}^{\mathrm{P}}(\mathsf{At})$ is *compact*.

Proposition 3.18. *The function* Inc^* *is well-defined.*

Proof. Let $\mathcal{R}$ be a knowledge base with $\langle \mathcal{R} \rangle = (r_1, \ldots, r_n)$ and $r_i = (\psi_i \,|\, \phi_i)[d_i]$ for $i = 1, \ldots, n$. Consider the set $\mathcal{P}_{\mathcal{R}} \subseteq \mathcal{P}^{\mathrm{P}}(\mathsf{At}) \times [0,1]^n$ defined via

$$\mathcal{P}_{\mathcal{R}} =_{def} \{(P, (x_1, \ldots, x_n)) \in \mathcal{P}^{\mathrm{P}}(\mathsf{At}) \times [0,1]^n \mid P \models^{pr} \Lambda_{\mathcal{R}}(x_1, \ldots, x_n)\}$$

We prove now that $\mathcal{P}_{\mathcal{R}}$ is closed. Let $(P_i, (x_1^i, \ldots, x_n^i)) \in \mathcal{P}_{\mathcal{R}}$ for $i \in \mathbb{N}$ be such that $\lim_{i\to\infty}(P_i, (x_1^i, \ldots, x_n^i))$ exists and define

$$(Q, (y_1, \ldots, y_n)) =_{def} \lim_{i\to\infty} (P_i, (x_1^i, \ldots, x_n^i))$$

In particular, it holds that $\lim_{i\to\infty} P_i = Q$ and by Lemma 3.3 it follows $Q \in \mathcal{P}^{\mathrm{P}}(\mathsf{At})$. For $j = 1, \ldots, n$, if $Q(\phi_j) > 0$ then there is some $k \in \mathbb{N}$ such that for all $i > k$ it holds that $P_i(\phi_j) > 0$ as well. Therefore, for $i > k$ it holds that $P_i(\psi_j \mid \phi_j) = x^i_j$ and

$$\begin{aligned} Q(\psi_j \mid \phi_j) &= \frac{Q(\psi_j\phi_j)}{Q(\phi_j)} = \frac{\lim_{i\to\infty} P_i(\psi_j\phi_j)}{\lim_{i\to\infty} P_i(\phi_j)} \\ &= \lim_{i\to\infty} \frac{P_i(\psi_j\phi_j)}{P_i(\phi_j)} = \lim_{i\to\infty} P_i(\psi_j \mid \phi_j) \\ &= \lim_{i\to\infty} x^i_j = y_j \end{aligned}$$

which implies $Q \models^{pr} (\psi_j \mid \phi_j)[y_j]$. If $Q(\phi_j) = 0$ (for $j = 1, \ldots, n$) then trivially $Q \models^{pr} (\psi_j \mid \phi_j)[y_j]$ due to our definition of probabilistic satisfaction. It follows that $(Q, (y_1, \ldots, y_n)) \models^{pr} \Lambda_{\mathcal{R}}(y_1, \ldots, y_n)$ and therefore $Q \in \mathcal{P}_{\mathcal{R}}$, i.e., $\mathcal{P}_{\mathcal{R}}$ is closed. Consider now the projection $\rho : \mathcal{P}_{\mathcal{R}} \to [0,1]^n$ defined via $\rho((P, (x_1, \ldots, x_n))) = (x_1, \ldots, x_n)$ for $(P, (x_1, \ldots, x_n)) \in \mathcal{P}_{\mathcal{R}}$. As $\mathcal{P}^{\mathrm{P}}(\mathsf{At})$ is compact it follows that ρ is a *closed map*, cf. the Tube Lemma[1] (Munkres, 1999). Therefore, ρ maps closed sets to closed sets and it follows that

$$\begin{aligned} \rho(\mathcal{P}_{\mathcal{R}}) &= \{(x_1, \ldots, x_n) \in [0,1]^n \mid \exists P : (P, (x_1, \ldots, x_n)) \in \mathcal{P}_{\mathcal{R}}\} \\ &= \{(x_1, \ldots, x_n) \in [0,1]^n \mid \Lambda_{\mathcal{R}}(x_1, \ldots, x_n) \text{ is consistent}\} \end{aligned}$$

is a closed set. We can write $\mathsf{Inc}^*(\mathcal{R})$ as

$$\mathsf{Inc}^*(\mathcal{R}) = \min\{|d_1 - x_1| + \ldots + |d_n - x_n| \mid (x_1, \ldots, x_n) \in \rho(\mathcal{P}_{\mathcal{R}})\} \quad .$$

As $\rho(\mathcal{P}_{\mathcal{R}})$ is a closed set (and also compact as it is bounded due to $\rho(\mathcal{P}_{\mathcal{R}}) \subseteq [0,1]^n$) and the mapping $(x_1, \ldots, x_n) \mapsto |d_1 - x_1| + \ldots + |d_n - x_n|$ is continuous the set

$$N_{\mathcal{R}} =_{def} \{|d_1 - x_1| + \ldots + |d_n - x_n| \mid (x_1, \ldots, x_n) \in \rho(\mathcal{P}_{\mathcal{R}})\}$$

is closed as well. As $\rho(\mathcal{P}_{\mathcal{R}})$ and therefore $N_{\mathcal{R}}$ are non-empty due to Proposition 3.17 it follows that $\mathsf{Inc}^*(\mathcal{R}) = \min N_{\mathcal{R}}$ is well-defined. □

Determining $\mathsf{Inc}^*(\mathcal{R})$ amounts to finding $(x_1, \ldots, x_n)$ such that $\Lambda_{\mathcal{R}}(x_1, \ldots, x_n)$ is consistent and $(x_1, \ldots, x_n)$ is closest to $(d_1, \ldots, d_n)$ with respect to the 1-norm distance (or *manhattan distance*), cf. (Bourbaki, 1987). Using the 1-norm distance is not mandatory and a whole family of inconsistency measures can be defined by considering other distance measures on $[0,1]^n$, cf. (Picado-Muiño, 2011). Nonetheless, in this thesis we restrain our attention to the MinDev inconsistency measure employing the 1-norm distance.

Let $\mathcal{MD}(\mathcal{R}) \subseteq [0,1]^{|\mathcal{R}|}$ be the set of arguments for $\Lambda_{\mathcal{R}}$ that yield a

1 An equivalent formalization of the Tube Lemma is "If X is Hausdorff and Y is Hausdorff and compact then $p : X \times Y \to X$ with $p(x, y) = x$ is a closed map". Note, that all spaces above are Hausdorff as they are subsets of Euclidean spaces.

consistent knowledge base and minimize the 1-norm distance to $\mathcal{R}$, i. e., for $\mathcal{R}$ with $\langle\mathcal{R}\rangle = (r_1, \ldots, r_n)$ with $(\psi_i \,|\, \phi_i)[d_i]$ for $i = 1, \ldots, n$ we define

$$\begin{aligned}\mathcal{MD}(\mathcal{R}) =_{def} \{(x_1, \ldots, x_n) \in [0,1]^{|\mathcal{R}|} \mid & \\ & |d_1 - x_1| + \ldots + |d_n - x_n| = \mathsf{Inc}^*(\mathcal{R}) \text{ and} \\ & \Lambda_{\mathcal{R}}(x_1, \ldots, x_n) \text{ consistent}\} \quad .\end{aligned}$$

We also define the set of probability functions that satisfy a nearest consistent knowledge base via

$$\begin{aligned}\mathcal{PMD}(\mathcal{R}) =_{def} \{P \in \mathcal{P}^{\mathsf{P}}(\mathsf{At}) \mid P \models^{pr} \Lambda_{\mathcal{R}}(x_1, \ldots, x_n) & \\ \text{for some } (x_1, \ldots, x_n) \in \mathcal{MD}(\mathcal{R})\} \quad .\end{aligned}$$

Example 3.10. Let $\mathcal{R}$ be a knowledge base with $\langle\mathcal{R}\rangle =_{def} \{r_1, r_2, r_3\}$ given via

$$r_1 =_{def} (b \,|\, a)[1] \qquad r_2 =_{def} (a)[1] \qquad r_3 =_{def} (b)[0] \quad .$$

Note that it holds that $\mathsf{Inc}^*(\mathcal{R}) = 1$ and that all three $\Lambda_{\mathcal{R}}(1,1,1)$, $\Lambda_{\mathcal{R}}(1,0,0)$, and $\Lambda_{\mathcal{R}}(0,1,0)$ are consistent, i. e. $(1,1,1), (1,0,0), (0,1,0) \in \mathcal{MD}(\mathcal{R})$.

Before discussing formal properties of Inc^* we first have a look at the computational issue of determining $\mathsf{Inc}^*(\mathcal{R})$ that will be useful for proving some technical results. We do so by extending the constraint satisfaction problem from Section 3.1 and setting up an optimization problem such that the solution of the problem is the value of $\mathsf{Inc}^*(\mathcal{R})$ for some knowledge base $\mathcal{R}$. Let $\mathcal{R}$ be a knowledge base with $\langle\mathcal{R}\rangle = (r_1, \ldots, r_n)$ and $(\psi_i \,|\, \phi_i)[d_i]$ for $i = 1, \ldots, n$. For every $i = 1, \ldots, n$ we introduce a variable $\eta_i \in [-1, 1]$ that measures the minimal deviation of the value of the probabilistic conditional $(\psi_i \,|\, \phi_i)[d_i]$. In order to realize Inc^* we have to modify the probabilistic conditionals in $\mathcal{R}$ in a minimal way, such that the knowledge base $\mathcal{R}'$ with the modified probabilistic conditionals is consistent, i. e., there is a probability function P that is a model for $\mathcal{R}'$. As before, let α_ω denote the probability of a complete conjunction $\omega \in \Omega(\mathsf{At})$. For every conditional $(\psi_i \,|\, \phi_i)[d_i]$ for $i = 1, \ldots, n$ we write

$$\sum_{\omega \in \mathsf{Mod}^{\mathsf{P}}(\psi_i \phi_i)} \alpha_\omega = (d_i + \eta_i) \cdot \sum_{\omega \in \mathsf{Mod}^{\mathsf{P}}(\phi_i)} \alpha_\omega \tag{3.8}$$

or equivalently

$$\sum_{\omega \in \mathsf{Mod}^{\mathsf{P}}(\psi_i \phi_i)} (1 - d_i - \eta_i)\alpha_\omega - \sum_{\omega \in \mathsf{Mod}^{\mathsf{P}}(\overline{\psi_i} \phi_i)} (d_i + \eta_i)\alpha_\omega = 0 \tag{3.9}$$

to comprehend for the fact that the modified conditional $(\psi_i \,|\, \phi_i)[d_i + \eta_i]$

imposes $P(\psi_i \mid \phi_i) = d_i + \eta_i$. To ensure well-formed conditionals we also have to consider the following normalization constraints

$$0 \leq d_1 + \eta_1 \leq 1, \ldots, 0 \leq d_n + \eta_n \leq 1 \tag{3.10}$$

and as before

$$\sum_{\omega \in \Omega(\mathsf{At})} \alpha_\omega = 1 \quad \text{and} \tag{3.11}$$

$$\alpha_\omega \geq 0 \quad \text{for all } \omega \in \Omega(\mathsf{At}) \tag{3.12}$$

We denote with $\mathsf{DevCons}(\mathcal{R})$ the set of constraints (3.9), (3.10), (3.11), and (3.12) for a knowledge base $\mathcal{R}$. In order to determine the minimal necessary deviation of $\mathcal{R}$ from a consistent knowledge base, we formulate an optimization problem by minimizing the function

$$f_{\mathsf{Inc}^*}(\eta_1, \ldots, \eta_n) =_{def} |\eta_1| + \cdots + |\eta_n|$$

and obeying $\mathsf{DevCons}(\mathcal{R})$. As $\mathsf{DevCons}(\mathcal{R})$ and f_{Inc^*} have been defined to characterize Inc^* we obtain the following result.

Proposition 3.19. *If $\eta_1^*, \ldots, \eta_n^*$ minimizes f_{Inc^*} and obeys $\mathsf{DevCons}(\mathcal{R})$ then $\mathsf{Inc}^*(\mathcal{R}) = f_{\mathsf{Inc}^*}(\eta_1^*, \ldots, \eta_n^*)$.*

Note that minimizing f_{Inc^*} with respect to $\mathsf{DevCons}(\mathcal{R})$ is a non-convex optimization problem as Equation (3.9) is non-convex.

Example 3.11. Let $\mathcal{R}$ be a knowledge base with $\langle \mathcal{R} \rangle =_{def} (r_1, r_2, r_3, r_4)$ and

$$r_1 =_{def} (b \mid \overline{a})[0.8] \qquad r_2 =_{def} (b \mid a)[0.6]$$
$$r_3 =_{def} (a)[0.5] \qquad r_4 =_{def} (b)[0.2] \quad .$$

Here, we have $\mathsf{Inc}^*(\mathcal{R}) = 0.5$ and some η_i^* $(i = 1, \ldots, 4)$ that minimize f_{Inc^*} can be determined as

$$\eta_1^* = \eta_2^* = \eta_3^* = 0 \qquad \eta_4^* = 0.5 \quad .$$

Therefore, the fourth conditional $(b)[0.2]$ has to be adjusted to $(b)[0.7]$ in order to restore consistency (this is just one possible adjustment).

If a knowledge base features multiple minimal inconsistent subsets it may not always be just a single conditional, that has to be modified in order to restore consistency.

Example 3.12. Consider the knowledge base $\mathcal{R}$ with $\langle \mathcal{R} \rangle =_{def} (r_1, r_2, r_3, r_4, r_5)$ and

$$r_1 =_{def} (a \mid c)[0.7] \qquad r_2 =_{def} (b \mid \bar{c})[0.8] \qquad r_3 =_{def} (a)[0.2]$$
$$r_4 =_{def} (b)[0.3] \qquad r_5 =_{def} (c)[0.5] \quad .$$

Here, we have $\mathsf{Inc}^*(\mathcal{R}_2) = 0.25$ and

$$\eta_1^* = 0 \qquad \eta_2^* = 0 \qquad \eta_3^* = 0.15$$
$$\eta_4^* = 0.1 \qquad \eta_5^* = 0$$

minimizes f_{Inc^*} and obeys $\mathsf{DevCons}(\mathcal{R})$.

Note also, that there is no other minimal adjustment. Consider the minimal inconsistent subset $M_1 = \{r_1, r_3, r_5\}$ of $\mathcal{R}$ and let $M(x, y, z) = \{(a \mid c)[x], (a)[y], (c)[z]\}$ for $x, y, z \in [0, 1]$. As one can see $M(x, y, z)$ is consistent if $xz \leq y$ as $(a \mid c)[x]$ and $(c)[z]$ together imply that a has to hold with at least a probability of xz. In order to restore consistency in M_1 in a *minimal* way we have to find x', y', z' such that $(0.7 - x')(0.5 - z') \leq 0.2 - y'$ is true and $|x'| + |y'| + |z'|$ is minimal. Due to the product on the left-hand side of this inequality the only minimal modification is setting $x' = 0$, $y' = -0.15$, and $z' = 0$. A similar argumentation holds for the minimal inconsistent subset $M_2 = \{r_2, r_4, r_5\}$. Consequently, $\eta_1^*, \ldots, \eta_5^*$ represent the only minimal adjustment.

We now turn to formal properties of Inc^*.

Theorem 3.1. Inc^* *satisfies (Consistency), (Monotonicity), (Super-Additivity), (Weak Independence), (Independence), and (Continuity).*

The proof of Theorem 3.1 can be found in Appendix A on page 233. Moreover, the examples below and in Appendix C suggest that Inc^* satisfies (MinInc Separability). However, no formal proof has been found yet.

Conjecture 3.1. Inc^* *satisfies (*MinInc *Separability).*

Observe also that Inc^* does satisfy neither (Irrelevance of syntax) nor (Penalty). The latter one has been mistakenly claimed in (Thimm, 2009a). Consider the following counterexample.

Example 3.13. Consider the knowledge base $\mathcal{R} =_{def} \{r_1, r_2\}$ given via

$$r_1 =_{def} (a)[0.7] \qquad r_2 =_{def} (a)[0.3]$$

and the probabilistic conditional $r_3 = (a)[0.5]$. Then r_3 is not free in $\mathcal{R}' =_{def} \mathcal{R} \cup \{r_3\}$ as $\mathsf{MI}(\mathcal{R}') = \{\{r_1, r_2\}, \{r_1, r_3\}, \{r_2, r_3\}\}$. However, it holds that $\mathsf{Inc}^*(\mathcal{R}) = \mathsf{Inc}^*(\mathcal{R}') = 0.4$ and therefore violating (Penalty). Consider also the knowledge base $\mathcal{R}'' =_{def} \{r_1, r_2, r_4, r_5\}$ with r_1 and r_2 as above and

$$r_4 =_{def} (\neg a)[0.3] \qquad\qquad r_5 =_{def} (\neg a)[0.7] \quad .$$

Note that it holds that $\mathcal{R} \equiv^{\text{cond}} \mathcal{R}''$. However, it also holds that $\mathsf{Inc}^*(\mathcal{R}'') = 0.8 \neq 0.4 = \mathsf{Inc}^*(\mathcal{R})$ and therefore Inc^* violates (Irrelevance of syntax).

Note also that Inc^* does not satisfy (Normalization).

Example 3.14. Consider the knowledge base $\mathcal{R} =_{def} \{r_1, \ldots, r_6\}$ given via

$$\begin{array}{lll} r_1 =_{def} (a)[0.0] & r_2 =_{def} (a)[1.0] & r_3 =_{def} (b)[0.0] \\ r_4 =_{def} (b)[1.0] & r_5 =_{def} (c)[0.0] & r_6 =_{def} (c)[1.0] \quad . \end{array}$$

It follows that $\mathsf{Inc}^*(\mathcal{R}) = 1.5$ thus violating (Normalization).

However, the following observation can be made regarding an upper bound of the value of $\mathsf{Inc}^*(\mathcal{R})$ for a knowledge base $\mathcal{R}$.

Proposition 3.20. *It holds that* $\mathsf{Inc}^*(\mathcal{R}) \leq |\mathcal{R}|$ *for every knowledge base* $\mathcal{R}$.

Proof. Let $\mathcal{R}$ be a knowledge base with $\langle \mathcal{R} \rangle = (r_1, \ldots, r_n)$ and $r_i = (\psi_i \,|\, \phi_i)[d_i]$ for $i = 1, \ldots, n$. Due to Proposition 3.17 on page 60 there are $d'_1, \ldots, d'_n \in [0,1]$ such that $\{(\psi_1 \,|\, \phi_1)[d'_1], \ldots, (\psi_n \,|\, \phi_n)[d'_n]\}$ is consistent. Therefore

$$\mathsf{Inc}^*(\mathcal{R}) \leq |d_1 - d'_1| + \ldots + |d_n - d'_n| \leq n = |\mathcal{R}|$$

as $|x - y| \leq 1$ for $x, y \in [0,1]$. □

For a specific knowledge base $\mathcal{R}$ Proposition 3.20 states that the value of $\mathsf{Inc}^*(\mathcal{R})$ is bounded above by the number of conditionals in $\mathcal{R}$. By exploiting this observation one can define a normalized inconsistency measure as follows.

Definition 3.11 (Normalized MinDev inconsistency measure)**.** Let Inc^*_0 be the function $\mathsf{Inc}^*_0 : \mathfrak{P}((\mathcal{L}(\mathsf{At}) \,|\, \mathcal{L}(\mathsf{At}))^{pr}) \to [0,1]$ defined as

$$\mathsf{Inc}^*_0(\mathcal{R}) =_{def} \begin{cases} 0 & \text{if } \mathcal{R} = \emptyset \\ \frac{\mathsf{Inc}^*(\mathcal{R})}{|\mathcal{R}|} & \text{otherwise} \end{cases}$$

with a knowledge base $\mathcal{R}$. Inc^*_0 is called the *normalized* MinDev *inconsistency measure*.

Proposition 3.21. *The function* Inc^*_0 *satisfies (Consistency), (Continuity), and (Normalization).*

The proof of Proposition 3.21 can be found in Appendix A on page 236.

Example 3.15. Consider the knowledge base $\mathcal{R} =_{def} \{r_1, r_2\}$ given via

$$r_1 =_{def} (a)[0.7] \qquad\qquad r_2 =_{def} (a)[0.3] \quad .$$

It holds that $\mathsf{Inc}^*_0(\mathcal{R}) = {}^{\mathsf{Inc}^*(\mathcal{R})}/_2 = 0.2$. Consider now the knowledge base $\mathcal{R}' =_{def} \mathcal{R} \cup \{(b)[0.5]\}$. It holds that

$$\mathsf{Inc}^*_0(\mathcal{R}') = \frac{\mathsf{Inc}^*(\mathcal{R}')}{3} = 0.1\overline{3}$$

and it follows that Inc^*_0 violates (Monotonicity) and therefore (Super-Additivity) (by contraposition of the statement of Proposition 3.2 on page 47). Notice also that r_3 does not mention any atom in $\mathcal{R}$ but $\mathsf{Inc}^*_0(\mathcal{R}') \neq \mathsf{Inc}^*_0(\mathcal{R})$. Hence, Inc^*_0 also violates (Weak Independence) and by contraposition of the statements of Propositions 3.6 on page 49 and 3.7 on page 50, Inc^*_0 violates (Independence) and (MinInc Separability). Inc^*_0 also fails to satisfy (Penalty) as Inc^* already fails to satisfy (Penalty). Consider the knowledge bases $\mathcal{R}' =_{def} \{r'_1, r'_2\}$ and $\mathcal{R}'' =_{def} \{r'_1, r'_2, r'_3\}$ given via

$$r'_1 =_{def} (a)[0.3] \qquad r'_2 =_{def} (a)[0.7] \qquad r'_3 =_{def} (\neg a)[0.3] \quad .$$

Then $\mathcal{R}' \equiv^{\text{cond}} \mathcal{R}''$ and $\mathsf{Inc}^*(\mathcal{R}') = 0.4 = \mathsf{Inc}^*(\mathcal{R}'')$ but $\mathsf{Inc}^*_0(\mathcal{R}') = 0.2 \neq {}^{0.4}/_3 = \mathsf{Inc}^*_0(\mathcal{R}'')$. Hence, Inc^*_0 fails to satisfy (Irrelevance of Syntax).

One last thing to discuss is what changes in the above elaboration when defining $P \models^{pr} (\psi \,|\, \phi)[d]$ if and only if $P(\psi \,|\, \phi) = d$ and $P(\phi) > 0$ instead of $P \models^{pr} (\psi \,|\, \phi)[d]$ if and only if $P(\psi \wedge \phi) = d \cdot P(\phi)$, cf. Section 2.3. Consider the knowledge base $\mathcal{R}$ given by $\mathcal{R} = \{(b \,|\, a)[0.8], (a)[0]\}$. When using the definition of probabilistic satisfaction employed in this thesis we get $\mathsf{Inc}^*(\mathcal{R}) = 0$. But by using the alternative definition of probabilistic satisfaction we get that $\mathcal{R}$ is inconsistent. An appropriate value for this inconsistency that respects our idea of distance as elaborated before would be an infinitesimal value. This approach is followed in (Picado-Muiño, 2011) where the distinction between consistent knowledge bases and inconsistent ones with an infinitesimal inconsistency value is made explicit. In (Picado-Muiño, 2011) our approach of measuring inconsistency is extended by this notion and also a whole family of inconsistency measures is presented based on different p-norms. However, the focus of the discussion of these inconsistency measures in (Picado-Muiño, 2011) is complementary to our discussion. In particular, in (Picado-Muiño, 2011) no investigation on rationality postulates as those presented in Section 3.2 is undertaken.

Determining the value of $\mathsf{Inc}^*(\mathcal{R})$ and $\mathsf{Inc}^*_0(\mathcal{R})$ using non-convex optimization techniques is—in general—a computationally demanding task due to 1.) the exponential many interpretations that have to be considered when setting up the optimization problem $\mathsf{DevCons}(\mathcal{R})$ and 2.) the lack of efficient and reliable solvers for general non-convex optimization problems. As the problem of determining the consistency of a knowledge is NP-complete and a sub-problem of determining $\mathsf{Inc}^*(\mathcal{R})$ we cannot expect feasibility and avoidance of 1.) above. However, in the following section we are going to address the second issue by developing approximations for $\mathsf{Inc}^*(\mathcal{R})$ that avoid the need for non-convex optimization.

3.4.2 *Approximating Distance Minimization*

One of the culprits of the high computational complexity of determining the minimal distance to a consistent knowledge base is the non-convex nature of the constraint set $\mathsf{DevCons}(\mathcal{R})$. In the following, we discuss two simplifications of $\mathsf{DevCons}(\mathcal{R})$ to a set of linear equations in order to benefit from the computational advantages of solving a linear optimization problem, cf. (Boyd and Vandenberghe, 2004).

Consider again Equation (3.8) from page 63

$$\sum_{\omega\in\mathsf{Mod}^{\mathrm{P}}(\psi_i\phi_i)} \alpha_\omega = (d_i+\eta_i)\cdot \sum_{\omega\in\mathsf{Mod}^{\mathrm{P}}(\phi_i)} \alpha_\omega$$

which can also be written as

$$\sum_{\omega\in\mathsf{Mod}^{\mathrm{P}}(\psi_i\phi_i)} \alpha_\omega = d_i\cdot \sum_{\omega\in\mathsf{Mod}^{\mathrm{P}}(\phi_i)} \alpha_\omega + \eta_i\cdot \sum_{\omega\in\mathsf{Mod}^{\mathrm{P}}(\phi_i)} \alpha_\omega \tag{3.13}$$

Obviously, the sole reason of $\mathsf{DevCons}(\mathcal{R})$ being non-linear is the second term on the right-hand side of the above equation. We introduce new variables $\mu_1,\ldots,\mu_n$ and substitute the non-linear terms with these variables, i. e., we define $\mu_i =_{def} \eta_i\cdot\sum_{\omega\in\mathsf{Mod}^{\mathrm{P}}(\phi_i)}\alpha_\omega$ for $i=1,\ldots,n$ and write

$$\sum_{\omega\in\mathsf{Mod}^{\mathrm{P}}(\psi_i\phi_i)} \alpha_\omega = d_i\cdot \sum_{\omega\in\mathsf{Mod}^{\mathrm{P}}(\phi_i))} \alpha_\omega + \mu_i \tag{3.14}$$

The above term is linear in α_ω for $\omega\in\Omega(\mathsf{At})$ and μ_i for $i=1,\ldots,n$. Consider furthermore the constraints

$$-d_i \sum_{\omega\in\mathsf{Mod}^{\mathrm{P}}(\phi_i)} \alpha_\omega \leq \mu_i \leq (1-d_i) \sum_{\omega\in\mathsf{Mod}^{\mathrm{P}}(\phi_i)} \alpha_\omega \tag{3.15}$$

for all $i=1,\ldots,n$ which as well are linear in α_ω for $\omega\in\Omega(\mathsf{At})$ and μ_i for $i=1,\ldots,n$. Equation 3.15 derives directly from Equation (3.10) on page 64 by replacing η_i with $\mu_i/\sum_{\omega\in\mathsf{Mod}^{\mathrm{P}}(\phi_i)}\alpha_\omega$. Let now $\mathsf{DevConsLin}(\mathcal{R})$ denote the set of constraints (3.11) on page 64, (3.12) on page 64, (3.14), and (3.15) for a knowledge base $\mathcal{R}$ and consider minimizing the function

$$g(\mu_1,\ldots,\mu_n) =_{def} |\mu_1|+\ldots+|\mu_n|$$

with respect to $\mathsf{DevConsLin}(\mathcal{R})$. Note that minimizing g with respect to $\mathsf{DevConsLin}(\mathcal{R})$ is a linear optimization problem. The following theorem links solutions of this problem to solutions of the original problem.

Proposition 3.22. *Let $\mu_1^*, \ldots, \mu_n^*, \alpha_{\omega_1}^*, \ldots, \alpha_{\omega_m}^*$ satisfy $\mathsf{DevConsLin}(\mathcal{R})$. Then $\eta_1^*, \ldots, \eta_n^*, \alpha_{\omega_1}^*, \ldots, \alpha_{\omega_m}^*$ with*

$$\eta_i^* =_{def} \frac{\mu_i^*}{\sum_{\omega_k \in \mathsf{Mod}^{\mathrm{P}}(a_i)} \alpha_{\omega_k}^*} \tag{3.16}$$

for $i = 1, \ldots, n$ satisfy $\mathsf{DevCons}(\mathcal{R})$ (with $0/0 =_{def} 0$).

Proof. Let $\mu_1^*, \ldots, \mu_n^*, \alpha_{\omega_1}^*, \ldots, \alpha_{\omega_m}^*$ satisfy $\mathsf{DevConsLin}(\mathcal{R})$ and define $\eta_1^*, \ldots,$ η_n^* via (3.16) for $i = 1, \ldots, n$. We have to show that $\eta_1^*, \ldots, \eta_n^*, \alpha_{\omega_1}^*, \ldots, \alpha_{\omega_m}^*$ satisfy Equations (3.8), (3.10), (3.11), and (3.12) on page 63. Equations (3.11) and (3.12) are satisfied as they are already contained in $\mathsf{DevConsLin}(\mathcal{R})$. For Equation (3.10) it holds that

$$\begin{aligned}
& 0 \le d_i + \eta_i^* \le 1 \\
\textit{iff} \quad & 0 \le d_i + \frac{\mu_i^*}{\sum_{\omega_k \in \mathsf{Mod}^{\mathrm{P}}(a_i)} \alpha_{\omega_k}^*} \le 1 \\
\textit{iff} \quad & -d_i \sum_{\omega \in \mathsf{Mod}^{\mathrm{P}}(\phi_i)} \alpha_\omega^* \le \mu_i^* \le (1 - d_i) \sum_{\omega \in \mathsf{Mod}^{\mathrm{P}}(\phi_i)} \alpha_\omega^*
\end{aligned}$$

which is the same as Equation (3.15) and therefore satisfied. For Equation (3.8) it holds that

$$\begin{aligned}
& \sum_{\omega \in \mathsf{Mod}^{\mathrm{P}}(\psi_i \phi_i)} \alpha_\omega^* = (d_i + \eta_i^*) \cdot \sum_{\omega \in \mathsf{Mod}^{\mathrm{P}}(\phi_i)} \alpha_\omega^* \\
\textit{iff} \quad & \sum_{\omega \in \mathsf{Mod}^{\mathrm{P}}(\psi_i \phi_i)} \alpha_\omega^* = \left(d_i + \frac{\mu_i^*}{\sum_{\omega_k \in \mathsf{Mod}^{\mathrm{P}}(a_i)} \alpha_{\omega_k}^*} \right) \sum_{\omega \in \mathsf{Mod}^{\mathrm{P}}(\phi_i)} \alpha_\omega^* \\
\textit{iff} \quad & \sum_{\omega \in \mathsf{Mod}^{\mathrm{P}}(\psi_i \phi_i)} \alpha_\omega^* = d_i \cdot \sum_{\omega \in \mathsf{Mod}^{\mathrm{P}}(\phi_i))} \alpha_\omega^* + \mu_i^*
\end{aligned}$$

which is the same as Equation (3.14) and therefore satisfied. □

Note that in the definition of η_i^* in Equation (3.16) it cannot be the case that the denominator is zero while the numerator is not zero. This is due to the fact that from $P(a) = 0$ it follows $P(ab) = 0$ as well. So Equation (3.14) becomes $0 = d_i \cdot 0 + \mu_i$ and can only be satisfied if $\mu_i = 0$ as well.

Proposition 3.23. *Let $\eta_1^*, \ldots, \eta_n^*, \alpha_{\omega_1}^*, \ldots, \alpha_{\omega_m}^*$ satisfy $\mathsf{DevCons}(\mathcal{R})$. Then $\mu_1^*,$ $\ldots, \mu_n^*, \alpha_{\omega_1}^*, \ldots, \alpha_{\omega_m}^*$ with*

$$\mu_i^* =_{def} \eta_i^* \cdot \sum_{\omega_k \in Mod(\phi_i)} \alpha_{\omega_k}^* \tag{3.17}$$

for $i = 1, \ldots, n$ *satisfy* $\mathsf{DevConsLin}(\mathcal{R})$.

Proof. The proof that $\mu_1^*, \ldots, \mu_n^*, \alpha_{\omega_1}^*, \ldots, \alpha_{\omega_m}^*$ satisfy $\mathsf{DevConsLin}(\mathcal{R})$ is analogous to the the proof of Proposition 3.22 and involves only replacing μ_i^* with its above definition. □

The observations from Proposition 3.22 and Proposition 3.23 together imply that both sets of constraints $\mathsf{DevCons}(\mathcal{R})$ and $\mathsf{DevConsLin}(\mathcal{R})$ are *equivalent*. However, the equivalence of the constraint sets does not imply that the *optimization problems* are equivalent as minimizing g differs (in general) from minimizing f_{Inc^*}.

Example 3.16. We continue Example 3.10 from page 63 and consider the knowledge base $\mathcal{R} = \{r_1, r_2, r_3\}$ with

$$r_1 = (b \,|\, a)[1] \qquad r_2 = (a)[1] \qquad r_3 = (b)[0] \quad .$$

Let $\omega_1 =_{def} ab, \omega_2 =_{def} a\overline{b}, \omega_3 =_{def} \overline{a}b, \omega_4 =_{def} \overline{a}\overline{b}$ and abbreviate α_{ω_i} by α_i for $i = 1, \ldots, 4$. Then Equation (3.14) becomes (for $i = 1, 2, 3$)

$$\begin{aligned} \alpha_1 &= \alpha_1 + \alpha_2 + \mu_1 \\ \alpha_1 + \alpha_2 &= 1 + \mu_2 \\ \alpha_1 + \alpha_3 &= \mu_3 \end{aligned}$$

Setting

$$\begin{array}{llll} \mu_1 =_{def} 0 & \mu_2 =_{def} 0 & \mu_3 =_{def} 1 & \\ \alpha_1 =_{def} 1 & \alpha_2 =_{def} 0 & \alpha_3 =_{def} 0 & \alpha_4 =_{def} 0 \end{array}$$

fulfills these constraints and the other constraints in $\mathsf{DevConsLin}(\mathcal{R})$ as well. Furthermore, the assignment

$$\begin{array}{llll} \mu_1' =_{def} -\dfrac{1}{3} & \mu_2' =_{def} -\dfrac{1}{3} & \mu_3' =_{def} \dfrac{1}{3} & \\ \alpha_1' =_{def} \dfrac{1}{3} & \alpha_2' =_{def} \dfrac{1}{3} & \alpha_3' =_{def} 0 & \alpha_4' =_{def} \dfrac{1}{3} \end{array}$$

also fulfills these constraints and the other constraints in $\mathsf{DevConsLin}(\mathcal{R})$ as well. Both assignments minimize g with function value 1. However, the values from the first assignment yield $\eta_1^* = \eta_2^* = 0$ and $\eta_3^* = 1$ (thus minimizing f_{Inc^*}) while the values from the second assignment yield $\eta_1^* = -1/2$, $\eta_2^* = -1/3$, and $\eta_3^* = 1/3$ with a function value of $7/6$ for f_{Inc^*}.

Nonetheless, by exploiting Proposition 3.23 we obtain some relationship between the solutions of these problems. Let $V(\mathcal{R})$ be the set of vectors $(\mu_1, \ldots, \mu_n)$ such that $\mu_1, \ldots, \mu_n$ satisfy $\mathsf{DevConsLin}(\mathcal{R})$. Then consider

$$\mathcal{I}^{\leq}(\mathcal{R}) =_{def} \min_{(\mu_1,\ldots,\mu_n)\in V(\mathcal{R})} g(\mu_1,\ldots,\mu_n) \quad . \tag{3.18}$$

By defining $\mathcal{I}^{\leq}$ in this manner we obtain a lower approximation for the value of $\mathsf{Inc}^*(\mathcal{R})$.

Proposition 3.24. *It holds that $\mathcal{I}^{\leq}(\mathcal{R}) \leq \mathsf{Inc}^*(\mathcal{R})$.*

Proof. Let $\eta_1^*,\ldots,\eta_n^*,\alpha_{\omega_1}^*,\ldots,\alpha_{\omega_m}^*$ minimize f_{Inc^*} with respect to the constraints in $\mathsf{DevCons}(\mathcal{R})$. Define $\mu_1^*,\ldots,\mu_n^*$ according to Equation (3.17). By Proposition 3.23 if follows that $\mu_1^*,\ldots,\mu_n^*,\alpha_{\omega_1}^*,\ldots,\alpha_{\omega_m}^*$ satisfy the constraints in $\mathsf{DevConsLin}(\mathcal{R})$ as well. Bearing in mind that it holds that

$$\sum_{\omega_k\in Mod(\phi_i)} \alpha_{\omega_k}^* \leq 1$$

for every $i = 1,\ldots,n$ it follows for $\mu_1^{**},\ldots,\mu_n^{**},\alpha_{\omega_1}^{**},\ldots,\alpha_{\omega_m}^{**}$ that minimize g with respect to $\mathsf{DevConsLin}(\mathcal{R})$ that

$$\begin{aligned}
\mathsf{Inc}^*(\mathcal{R}) &= |\eta_1^*| + \ldots + |\eta_n^*| \\
&= \left|\frac{\mu_1^*}{\sum_{\omega_k\in Mod(\phi_1)} \alpha_{\omega_k}^*}\right| + \ldots + \left|\frac{\mu_n^*}{\sum_{\omega_k\in Mod(\phi_n)} \alpha_{\omega_k}^*}\right| \\
&\geq |\mu_1^*| + \ldots + |\mu_n^*| \\
&\geq |\mu_1^{**}| + \ldots + |\mu_n^{**}| \\
&= \mathcal{I}^{\leq}(\mathcal{R})
\end{aligned}$$

□

For a knowledge base $\mathcal{R}$, determining the value of $\mathcal{I}^{\leq}(\mathcal{R})$ is computationally easier as determining the value of $\mathsf{Inc}^*(\mathcal{R})$ as the former only needs to solve a linear optimization problem. Nonetheless, the exponential blow-up for setting up the optimization problem itself remains.

By exploiting Proposition 3.22 we also gain an upper approximation for $\mathsf{Inc}^*(\mathcal{R})$. Consider the function g^H defined via

$$\begin{aligned}
g^H(\mu_1,\ldots,\mu_n,\alpha_{\omega_1},\ldots,\alpha_{\omega_m}) =_{def} &|\mu_1| \mathsf{ld}\, |\mu_1| + \ldots + |\mu_n| \mathsf{ld}\, |\mu_n| + \\
&\alpha_{\omega_1} \mathsf{ld}\, \alpha_{\omega_1} + \ldots + \alpha_{\omega_m} \mathsf{ld}\, \alpha_{\omega_m}
\end{aligned}$$

and let $V^H(\mathcal{R})$ denote the set of vectors $(\mu_1,\ldots,\mu_n,\alpha_{\omega_1},\ldots,\alpha_{\omega_m})$ such that $\mu_1,\ldots,\mu_n,\alpha_{\omega_1},\ldots,\alpha_{\omega_m}$ minimize g^H with respect to $\mathsf{DevConsLin}(\mathcal{R})$.

Proposition 3.25. *It holds that $|V^H(\mathcal{R})| = 1$ for every knowledge base $\mathcal{R}$.*

Proof. Each constraint in $\mathsf{DevConsLin}(\mathcal{R})$ is linear and therefore convex. As the function $h(\mu_1,\ldots,\mu_n,\alpha_{\omega_1},\ldots,\alpha_{\omega_m}) = -|\mu_1| \mathsf{ld}\, |\mu_1| - \ldots - |\mu_n| \mathsf{ld}\, |\mu_n| - \alpha_{\omega_1} \mathsf{ld}\, \alpha_{\omega_1} - \ldots - \alpha_{\omega_m} \mathsf{ld}\, \alpha_{\omega_m}$ is the same as the entropy of $(\mu_1,\ldots,\mu_n,\alpha_{\omega_1},\ldots,\alpha_{\omega_m})$—which is strictly concave—it follows that the function $g^H = -h(\mu_1, \ldots, \mu_n, \alpha_{\omega_1}, \ldots, \alpha_{\omega_m})$ is strictly convex. Minimizing a strictly convex function over a convex set has a unique solution (Boyd and Vandenberghe, 2004). □

Now consider

$$\mathcal{I}^{\geq}(\mathcal{R}) =_{def} |\eta_1^*| + \ldots + |\eta_n^*| \tag{3.19}$$

for $\eta_1^*, \ldots, \eta_n^*$ stemming from the application of Equation (3.16) for some $\mu_1^*, \ldots, \mu_n^*, \alpha_{\omega_1}^*, \ldots, \alpha_{\omega_m}^*$ that minimize g^H with respect to $\mathsf{DevConsLin}(\mathcal{R})$. This definition is sound as $\mu_1^*, \ldots, \mu_n^*, \alpha_{\omega_1}, \ldots, \alpha_{\omega_m}$ are uniquely determined due to Proposition 3.25.

Note that we could have defined $\mathcal{I}^{\geq}$ using another strictly convex function than g^H. We choose to use the negative entropy only as an example.

Corollary 3.2. *It holds that $\mathcal{I}^{\geq}(\mathcal{R}) \geq \mathsf{Inc}^*(\mathcal{R})$.*

Proof. Let $\mu_1^*, \ldots, \mu_n^*, \alpha_{\omega_1}^*, \ldots, \alpha_{\omega_m}^*$ minimize g^H with respect to the set of constraints $\mathsf{DevConsLin}(\mathcal{R})$. Define $\eta_1^*, \ldots, \eta_n^*$ according to Equation (3.16). By Proposition 3.22 it follows that $\eta_1^*, \ldots, \eta_n^*, \alpha_{\omega_1}^*, \ldots, \alpha_{\omega_m}^*$ satisfy the set of constraints $\mathsf{DevCons}(\mathcal{R})$ as well. Therefore, for $\eta_1^{**}, \ldots, \eta_n^{**}, \alpha_{\omega_1}^{**}, \ldots, \alpha_{\omega_m}^{**}$ minimizing f_{Inc^*} it holds that

$$\mathsf{Inc}^*(\mathcal{R}) = |\eta_1^{**}| + \ldots + |\eta_n^{**}| \leq |\eta_1^*| + \ldots + |\eta_n^*| = \mathcal{I}^{\geq}(\mathcal{R})$$

□

Determining the value of $\mathcal{I}^{\geq}(\mathcal{R})$ is computationally easier as determining the value of $\mathsf{Inc}^*(\mathcal{R})$ due to the same reasons as for $\mathcal{I}^{\leq}$. By defining

$$\mathcal{I}_0^{\leq}(\mathcal{R}) =_{def} \begin{cases} 0 & \text{if } \mathcal{R} = \emptyset \\ \frac{\mathcal{I}^{\leq}(\mathcal{R})}{|\mathcal{R}|} & \text{otherwise} \end{cases}$$

$$\mathcal{I}_0^{\geq}(\mathcal{R}) =_{def} \begin{cases} 0 & \text{if } \mathcal{R} = \emptyset \\ \frac{\mathcal{I}^{\geq}(\mathcal{R})}{|\mathcal{R}|} & \text{otherwise} \end{cases}$$

we also obtain approximations for Inc_0^*.

Proposition 3.26. *For every knowledge base $\mathcal{R}$ it holds that $\mathcal{I}_0^{\leq}(\mathcal{R}) \leq \mathsf{Inc}_0^* \leq \mathcal{I}_0^{\geq}(\mathcal{R})$.*

By replacing $\mathcal{I}_0^{\leq}$, Inc_0^*, and $\mathcal{I}_0^{\geq}$ with their definitions based on $\mathcal{I}^{\leq}$, Inc^*, and $\mathcal{I}^{\geq}$, respectively, the proof for the above proposition is easy to see.

The general approximation quality of both $\mathcal{I}_0^{\leq}$ and $\mathcal{I}_0^{\geq}$ is hard to assess and depends heavily on the actual knowledge base. Appendix C on page 253ff. contains some examples that illustrate the approximation quality of $\mathcal{I}^{\leq}$ and $\mathcal{I}^{\geq}$. As one can see, the function $\mathcal{I}^{\leq}$ almost always coincides with the function Inc^* while $\mathcal{I}^{\geq}$ deviates to Inc^* by a factor of at most 2.2.

3.4.3 *Extensions*

In the following, we consider two extensions of the underlying probabilistic framework of probabilistic conditional logic and discuss the applicability of the MinDev inconsistency measures in these extensions. We investigate inconsistency measurement in the framework of bounded probabilistic conditionals (see Section 2.3) and within the framework of linear probabilistic knowledge bases, cf. (Paris, 1994; Daniel, 2009). We restrain our attention to the inconsistency measure Inc^* but the discussion and the technical results carry over to the approximations $\mathcal{I}^{\leq}$ and $\mathcal{I}^{\geq}$ in a similar fashion.

Bounded Probabilistic Conditionals

In Section 2.3 we presented *bounded probabilistic conditionals*, i. e., conditionals of the form $(\psi\,|\,\phi)[l,u]$ with $\phi,\psi \in \mathcal{L}(\mathsf{At})$, $l,u \in [0,1]$ and $l \leq u$ denoting lower and upper bounds to the probability of the conditional. Given a knowledge base $\mathcal{R}$ consisting of bounded probabilistic conditionals a probability function P satisfies $\mathcal{R}$, denoted by $P \models^{pr} \mathcal{R}$, if and only if for every conditional $(\psi\,|\,\phi)[l,u] \in \mathcal{R}$ it holds that $P(\phi) = 0$ or $P(\psi\,|\,\phi) \in [l,u]$. Using this definition we can consider the problem of inconsistency measurement in the same way as above. Furthermore, we can easily extend the approach of inconsistency measurement using distance minimization to bounded probabilistic conditionals as follows.

We extend the notion of the characteristic function to include bounded knowledge bases as well. In the following, let $\mathcal{R}$ be a knowledge base of bounded probabilistic conditionals with

$$\langle\mathcal{R}\rangle = ((\psi_1\,|\,\phi_1)[l_1,u_1],\ldots,(\psi_n\,|\,\phi_n)[l_n,u_n]) \quad .$$

The characteristic function $\Lambda_{\mathcal{R}}$ for $\mathcal{R}$ is the function

$$\Lambda_{\mathcal{R}} : ([0,1]\times[0,1])^{|\mathcal{R}|} \to \mathfrak{P}((\mathcal{L}(\mathsf{At})\,|\,\mathcal{L}(\mathsf{At}))^{pr,pr})$$

with

$$\Lambda_{\mathcal{R}}((x_1,y_1),\ldots,(x_n,y_n)) =_{def} \{(\psi_1\,|\,\phi_1)[x_1,y_1],\ldots,(\psi_n\,|\,\phi_n)[x_n,y_n]\} \quad .$$

Then we can extend the notion of the MinDev inconsistency measure for bounded conditionals as follows.

Definition 3.12 (MinDev inconsistency measure for bounded conditionals)**.** Let $\mathcal{R}$ be a knowledge base of bounded conditionals with $\langle\mathcal{R}\rangle = ((\psi_1\,|\,\phi_1)[l_1,u_1],\ldots,(\psi_n\,|\,\phi_n)[l_n,u_n])$. Let $\mathsf{Inc}^*_b : \mathfrak{P}((\mathcal{L}(\mathsf{At})\,|\,\mathcal{L}(\mathsf{At}))^{pr,pr}) \to [0,\infty)$ be defined as

$$\begin{aligned}\mathsf{Inc}^*_b(\mathcal{R}) =_{def} \min\{&|l_1-x_1|+|u_1-y_1|+\ldots+|l_n-x_n|+|u_n-y_n| \mid\\ &\Lambda_{\mathcal{R}}((x_1,y_1),\ldots,(x_n,y_n)) \text{ is consistent}\} \quad .\end{aligned}$$

Inc_b^* is called the MinDev *inconsistency measure for bounded conditionals.*

Proposition 3.27. *Let $\mathcal{R}$ be a knowledge base of bounded conditionals with $\langle\mathcal{R}\rangle = ((\psi_1 \,|\, \phi_1)[l_1, u_1], \ldots, (\psi_n \,|\, \phi_n)[l_n, u_n])$ and let $(x_1^*, y_1^*), \ldots, (x_n^*, y_n^*) \in [0,1]^2$ such that*

$$\mathsf{Inc}_b^*(\mathcal{R}) = |l_1 - x_1^*| + |u_1 - y_1^*| + \ldots + |l_n - x_n^*| + |u_n - y_n^*| \tag{3.20}$$

and $\Lambda_{\mathcal{R}}((x_1^*, y_1^*), \ldots, (x_n^*, y_n^*))$ is consistent. Then for every $i = 1, \ldots, n$ it cannot be the case that both $x_i^* \neq l_i$ and $y_i^* \neq u_i$.

Proof. Without loss of generality assume that $x_1^* \neq l_1$ and $y_1^* \neq u_1$. Let P be a probability function with $P \models^{pr} \Lambda_{\mathcal{R}}((x_1^*, y_1^*), \ldots, (x_n^*, y_n^*))$ and consider $p = P(\psi_1 \,|\, \phi_1)$. It follows that $p \in [x_1^*, y_1^*]$ and $p \notin [l_1, u_1]$, otherwise $\Lambda_{\mathcal{R}}((0,0), (x_2^*, y_2^*), \ldots, (x_n^*, y_n^*))$ would be consistent, contradicting the minimality in Equation (3.20). If $p < l_i$ then $p \in [x_1^*, u_1]$ and $\Lambda_{\mathcal{R}}((x_1^*, 0), (x_2^*, y_2^*), \ldots, (x_n^*, y_n^*))$ would be consistent, contradicting the minimality in Equation (3.20). It follows that $p > u_i$, so $p \in [l_1, y_1^*]$ and $\Lambda_{\mathcal{R}}((0, y_1^*), (x_2^*, y_2^*), \ldots, (x_n^*, y_n^*))$ is consistent, contradicting again the minimality in Equation (3.20). It follows that at either $x_1^* = l_1$ or $y_1^* = u_1$. $\square$

Proposition 3.28. *For $\mathcal{R} = \{(\psi_1 \,|\, \phi_1)[l_1, u_1], \ldots, (\psi_n \,|\, \phi_n)[l_n, u_n]\}$ with $l_i = u_i$ for all $i = 1, \ldots, n$ it follows that $\mathsf{Inc}_b^*(\mathcal{R}) = \mathsf{Inc}^*(\mathcal{R}')$ with $\mathcal{R}' =_{def} \{(\psi_1 \,|\, \phi_1)[l_1], \ldots, (\psi_n \,|\, \phi_n)[l_n]\}$.*

Proof. Let $\mathcal{R}$ be a knowledge base of bounded conditionals with $\langle\mathcal{R}\rangle = ((\psi_1 \,|\, \phi_1)[l_1, u_1], \ldots, (\psi_n \,|\, \phi_n)[l_n, u_n])$ and let $\mathcal{R}'$ be a knowledge base with $\langle\mathcal{R}'\rangle = (\psi_1 \,|\, \phi_1)[l_1], \ldots, (\psi_n \,|\, \phi_n)[l_n]$. The proof of the claim follows directly from the fact that $P \models^{pr} (\psi \,|\, \phi)[d]$ if and only if $P \models^{pr} (\psi \,|\, \phi)[d, d]$ for every probability function P. Therefore, for $x_1^*, \ldots, x_n^*$ with $\mathsf{Inc}^*(\mathcal{R}') = |l_1 - x_1^*| + \ldots + |l_n - x_n^*|$ and $\Lambda_{\mathcal{R}'}(x_1^*, \ldots, x_n^*)$ is consistent it follows that

$$\Lambda_{\mathcal{R}}((\min\{x_1^*, l_1\}, \max\{x_1^*, u_1\}), \ldots, (\min\{x_n^*, l_n\}, \max\{x_n^*, u_n\}))$$

is consistent as well and (due to $l_i = u_i$ for $i = 1, \ldots, n$)

$$\begin{aligned}
\mathsf{Inc}_b^*(\mathcal{R}) &\geq |l_1 - \min\{x_1^*, l_1\}| + |u_1 - \max\{x_1^*, u_1\}| + \ldots + \\
&\qquad |l_n - \min\{x_n^*, l_n\}| + |u_n - \max\{x_n^*, u_n\}| \\
&= |l_1 - x_1| + \ldots + |l_n - x_n| = \mathsf{Inc}^*(\mathcal{R}')
\end{aligned}$$

Similarly, one obtains $\mathsf{Inc}_b^*(\mathcal{R}) \leq \mathsf{Inc}^*(\mathcal{R}')$ by starting with $\mathsf{Inc}_b^*(\mathcal{R}') = |l_1 - x_1^*| + |u_1 - y_1^*| + \ldots + |l_n - x_n^*| + |u_i - y_n^*|$ such that $\Lambda_{\mathcal{R}}((x_1^*, y_1^*), \ldots, (x_n^*, y_n^*))$ is consistent and then applying Proposition 3.27. $\square$

As before, Inc_b^* can be phrased as the solution of an optimization problem. For every $i = 1, \ldots, n$ we introduce variables $\eta_i^l, \eta_i^u \in [0,1]$ that measure the minimal deviation of the value of the probabilistic conditional $(\psi_i \,|\, \phi_i)[l_i, u_i]$

to the lower and upper side, respectively. Let α_ω denote the probability of some possible world $\omega \in \Omega(\mathsf{At})$. Now, for every conditional $(\psi_i \mid \phi_i)[l_i, u_i]$, $i = 1, \ldots, n$, consider the constraints

$$\sum_{\omega \in \mathsf{Mod}^{\mathsf{P}}(\psi_i \wedge \phi_i)} \alpha_\omega \geq (l_i - \eta_i^l) \cdot \sum_{\omega \in \mathsf{Mod}^{\mathsf{P}}(\phi_i)} \alpha_\omega \tag{3.21}$$

$$\sum_{\omega \in \mathsf{Mod}^{\mathsf{P}}(\psi_i \wedge \phi_i)} \alpha_\omega \leq (u_i + \eta_i^u) \cdot \sum_{\omega \in \mathsf{Mod}^{\mathsf{P}}(\phi_i)} \alpha_\omega \tag{3.22}$$

to comprehend for the fact that the modified conditional $(\psi_i \mid \phi_i)[l_i - \eta_i^l, u_i + \eta_i^u]$ is satisfiable. Further integrity constraints are given via

$$\eta_1^l \leq l_1, \eta_1^u \leq 1 - u_1, \ldots, \eta_n^l \leq l_n, \eta_n^u \leq 1 - u_n \tag{3.23}$$

$$\sum_{\omega \in \Omega(\mathsf{At})} \alpha_\omega = 1 \tag{3.24}$$

$$\alpha_\omega \geq 0 \qquad \text{for all } \omega \in \Omega(\mathsf{At}) \quad . \tag{3.25}$$

We denote by $\mathsf{DevConsLin}_b(\mathcal{R})$ the set of constraints (3.21), (3.22), (3.23), (3.24), and (3.25) for a knowledge base $\mathcal{R}$ of bounded probabilistic conditionals. Consider minimizing the function

$$f_b(\eta_1^l, \eta_1^u, \ldots, \eta_n^l, \eta_n^u) =_{def} |\eta_1^l| + |\eta_1^u| + \cdots + |\eta_n^l| + |\eta_n^u|$$

with respect to $\mathsf{DevConsLin}_b(\mathcal{R})$. By construction it follows that $\mathsf{Inc}_b^*(\mathcal{R}) = f^b(\eta_1^l, \eta_1^u, \ldots, \eta_n^l, \eta_n^u)$ if $\eta_1^l, \eta_1^u, \ldots, \eta_n^l, \eta_n^u$ minimize f_b with respect to the constraints in $\mathsf{DevConsLin}_b(\mathcal{R})$.

Linear Probabilistic Knowledge Bases

Probabilistic conditionals are a natural means to represent probabilistic beliefs but fail to represent certain more complex pieces of information. Consider the statement *"The probability of Anna winning some competition is at least twice as large as the probability of Bob winning some competition"*. Clearly, there is no way of expressing these statements using solely probabilistic conditionals. In the following, we consider *linear probabilistic knowledge bases* to adhere for this lack of expressivity and show how the previous discussion on inconsistency measurement can be applied to this framework.

A *linear probabilistic constraint* c is a linear constraint of the form

$$c \;:\quad b \geq h_1\langle\phi_1\rangle_P + \ldots + h_m\langle\phi_m\rangle_P \tag{3.26}$$

with formulas $\phi_1, \ldots, \phi_m \in \mathcal{L}(\mathsf{At})$, and real values $b, h_1, \ldots, h_m \in \mathbb{R}$. Let K denote the term on the right-hand side of Equation (3.26). As a notational convenience we sometimes write linear constraints of the form $b \leq K$ (equivalent to $-b \geq -K$) and $b = K$ (equivalent to the two constraints $b \geq K$ and $b \leq K$).

A finite set of probabilistic constraints $\mathcal{R} = \{c_1, \ldots, c_n\}$ is called a *linear probabilistic knowledge base*, cf. (Daniel, 2009). A probability function $P : \Omega(\mathsf{At}) \to [0,1]$ *satisfies* a probabilistic constraint c of the form (3.26), denoted by $P \models^{pr} c$, if and only if

$$b \geq h_1 P(\phi_1) + \ldots + h_m P(\phi_m) \quad . \tag{3.27}$$

A probability function P satisfies a knowledge base $\mathcal{R}$, denoted by $P \models^{pr} \mathcal{R}$, if and only if P satisfies every constraint in $\mathcal{R}$, i. e., for every $c \in \mathcal{R}$ it holds that $P \models^{pr} c$.

Example 3.17. Let a denote the proposition *"Anna wins some competition"* and b denote the *"Bob wins some competition"*. Then the statement *"The probability of Anna winning some competition is at least twice as large as the probability of Bob winning some competition"* can be expressed as a linear probabilistic constraint $0 \leq \langle a \rangle_P - 2\langle b \rangle_P$.

Example 3.18. The constraint imposed by a probabilistic conditional r with $r = (\psi \,|\, \phi)[d]$ can be represented as a linear probabilistic constraint via $0 = \langle \phi \wedge \psi \rangle_P - d\langle \phi \rangle_P$. Furthermore, the constraint imposed by a bounded probabilistic conditional $(\psi \,|\, \phi)[l, u]$ can be represented as the two linear probabilistic constraints $0 \leq \langle \phi \wedge \psi \rangle_P - l\langle \phi \rangle_P$ and $0 \geq \langle \phi \wedge \psi \rangle_P - u\langle \phi \rangle_P$.

As the previous examples show linear probabilistic knowledge bases are a strict generalization of both probabilistic conditional knowledge bases and probabilistic conditional knowledge bases with bounded conditionals.

A probabilistic constraint c is said to be in *canonical form* if it has the form

$$0 \geq \sum_{\omega \in \Omega(\mathsf{At})} h_\omega \langle \omega \rangle_P$$

for some $h_\omega \in [-1, 1]$. Note that every probabilistic constraint can be rewritten to be in canonical form (Paris, 1994; Daniel, 2009). As before, we assume that every probabilistic constraint c is *self-consistent*, i. e., there is always a P with $P \models^{pr} c$. Using linear probabilistic knowledge bases we can consider the problem of inconsistency measurement in the same way as before and develop an approach of inconsistency measurement using distance minimization for this framework as follows.

Let $\mathcal{R}$ be a linear probabilistic knowledge base with $\langle \mathcal{R} \rangle = (c_1, \ldots, c_n)$. Without loss of generality we assume that every c_i is in canonical form

$$c_i \;:\quad 0 \geq \sum_{\omega \in \Omega(\mathsf{At})} h^i_\omega \langle \omega \rangle_P$$

with $h^i_\omega \in [-1, 1]$ for every $\omega \in \Omega(\mathsf{At})$ and $i = 1, \ldots, n$. For every $i = 1, \ldots, n$ and $\omega \in \Omega(\mathsf{At})$ we introduce a variable $\tau^i_\omega \in \mathbb{R}$ that measures the minimal deviation for the corresponding h^i_ω. Let α_ω denote the probability of a possible world $\omega \in \Omega(\mathsf{At})$. Now, for every constraint c_i, $i = 1, \ldots, n$, consider the constraint

$$0 \geq \sum_{\omega \in \Omega(\mathsf{At})} (h^i_\omega + \tau^i_\omega)\alpha_\omega \tag{3.28}$$

to comprehend for the fact that the linear probabilistic knowledge base $\mathcal{R}'$ with $\langle \mathcal{R}' \rangle = (c'_1, \ldots, c'_n)$ of modified probabilistic constraints

$$c'_i \;:\quad 0 \geq \sum_{\omega \in \Omega(\mathsf{At})} (h^i_\omega + \tau^i_\omega)\langle \omega \rangle_p \quad (i = 1, \ldots, n)$$

is satisfiable. As before, further integrity constraints are given via

$$\sum_{\omega \in \Omega(\mathsf{At})} \alpha_\omega = 1 \quad \text{and} \tag{3.29}$$

$$\alpha_\omega \geq 0 \quad \text{for all } \omega \in \Omega(\mathsf{At}) \quad . \tag{3.30}$$

We denote with $\mathsf{DevConsLin}_{lp}(\mathcal{R})$ the set of constraints (3.28), (3.29), and (3.30) for a linear probabilistic knowledge base $\mathcal{R}$. Consider minimizing the function

$$f_{lp}((\tau^i_\omega)^{i=1,\ldots,n}_{\omega \in \Omega(\mathsf{At})}) =_{def} \sum_{i=1}^{n} \sum_{\omega \in \Omega(\mathsf{At})} |\tau^i_\omega|$$

with respect to $\mathsf{DevConsLin}_{lp}(\mathcal{R})$. Let $\mathsf{Inc}^*_{lp}(\mathcal{R})$ denote the solution for f_{lp} in this optimization problem.

Proposition 3.29. *Let $\mathcal{R}$ be a linear probabilistic knowledge base with $\langle \mathcal{R} \rangle = (c_1, \ldots, c_n)$. If $c_i \in \mathcal{R}$ is of the form*

$$c_i \;:\quad 0 = \langle \phi_i \psi_i \rangle_p - d_i \langle \phi_i \rangle_p$$

for all $i = 1, \ldots, n$ then

$$\mathsf{Inc}^*_{lp}(\mathcal{R}) = \mathsf{Inc}^*(\mathcal{R}')$$

with $\mathcal{R}' = \{(\psi_1 \,|\, \phi_1)[d_1], \ldots, (\psi_n \,|\, \phi_n)[d_n]\}$.

Proof. The proof directly follows from the fact that $P \models^{pr} (\psi \,|\, \phi)[d]$ if and only if $P \models^{pr} (0 = \langle \phi\psi \rangle_p - d_i \langle \phi \rangle_p)$ for every probability function P (see also the proof of Proposition 3.28 on page 74). □

3.5 RELATED WORK

The problem of measuring inconsistency in probabilistic knowledge bases is relatively novel and has—to our knowledge—only been addressed before in (Rödder and Xu, 2001), (Daniel, 2009), and (Thimm, 2009a). While our previous discussion on inconsistency measurement has been based on the work (Thimm, 2009a) an evaluation and comparison with the works (Rödder and Xu, 2001) and (Daniel, 2009) is given below. We also have a closer look on the approach of η-consistency (Knight, 2001, 2002) which uses probability theory to measure inconsistencies in classical theories. Further related work is also concerned with measuring inconsistency in classical theories, see e. g. the works by Hunter et. al. (Hunter, 2002, 2003; Konieczny *et al.*, 2003; Hunter and Konieczny, 2004; Grant and Hunter, 2006; Hunter and Konieczny, 2006, 2008; Grant and Hunter, 2008; Hunter and Konieczny, 2010). While (Hunter, 2002, 2003; Konieczny *et al.*, 2003; Hunter and Konieczny, 2004, 2006, 2008, 2010) deal with measuring inconsistency in propositional logic, the works (Grant and Hunter, 2006, 2008) consider first-order logic. Those works also take a principled approach to measuring inconsistency and many of our properties have been adapted from e. g. (Hunter and Konieczny, 2006, 2008). Furthermore, the inconsistency measures presented in Section 3.3 are straightforward translations of inconsistency measures from those works. However, Hunter et. al. are working with classical theories and as such do not have to deal with probabilities as a means for knowledge representation. In order to adhere for the presence of probabilities we introduced the MinDev inconsistency measure which has no correspondent in the classical setting.

We go on by taking a closer look on the works by Knight (Knight, 2001, 2002), Rödder and Xu (Rödder and Xu, 2001), and Daniel (Daniel, 2009).

3.5.1 *η-Consistency*

The works (Knight, 2001, 2002) employ probability theory to measure inconsistency in classical theories. Let At be a propositional signature. Then a set of propositional sentences $\Phi \subseteq \mathcal{L}(\mathsf{At})$ is *η-consistent* if there is a probability function P on $\mathcal{L}(\mathsf{At})$ in the sense of Section 2.3 such that $P(\phi) \geq \eta$ for all $\phi \in \Phi$. The rationale behind this definition is that the probability function P measures the *compatibility* of the interpretations of $\mathcal{L}(\mathsf{At})$ from the point of view of P. The higher the probability $P(\phi)$ of some $\phi \in \Phi$ the better ϕ can be explained by the interpretations of $\mathcal{L}(\mathsf{At})$ with respect to P. A set $\Phi \subseteq \mathcal{L}(\mathsf{At})$ is *maximally η-consistent* if Φ is η-consistent and for every $\eta' > \eta$ it is not η'-consistent. Consider $\mathsf{At} = \{a\}$ and $\Phi = \{a, \neg a\}$. Then Φ is maximally $1/2$-consistent as P with $P(a) = 1/2$ and $P(\neg a) = 1/2$ is a valid probability function and there can be no probability function P' such that $P'(a) > 1/2$ and $P'(\neg a) > 1/2$. It can be shown, see (Knight, 2002), that for every $\Phi \subseteq \mathcal{L}(\mathsf{At})$ there is always some η with $0 \leq \eta \leq 1$ such that Φ is maximally η-consistent. In particular, for consistent Φ it holds that Φ is maximally 1-consistent.

From a categorical point of view the concept of η-consistency is hardly comparable to our notion of inconsistency measures as (Knight, 2001, 2002) use a classical knowledge representation formalism and we employ a probabilistic framework. However, a translation of η-consistency to probabilistic conditional logic seems to be feasible as we already translated several other inconsistency measures for classical logic to probabilistic conditional logic. The main idea of η-consistency lies in finding a probability function on the interpretations of the knowledge base that maximizes the probability of the sentences in the knowledge base. When switching to our probabilistic framework the interpretations for knowledge bases are probability functions themselves. Then, a straightforward application of the paradigm of η-consistency to our framework may be defined as follows. Let $\hat{P} : \mathcal{P}^{\mathrm{P}}(\mathsf{At}) \to [0,1]$ be a probability function on $\mathcal{P}^{\mathrm{P}}(\mathsf{At})$ such that $\hat{P}(P) > 0$ only for finitely many $P \in \mathcal{P}^{\mathrm{P}}(\mathsf{At})$. Define

$$\hat{P}(r) =_{def} \sum_{P \in \mathcal{P}^{\mathrm{P}}(\mathsf{At}), P \models^{pr} r} \hat{P}(P)$$

for a probabilistic conditional r. This means, that the probability (in terms of $\hat{P}$) of a probabilistic conditional is the sum of the probabilities of probability functions that satisfy r. By defining $\hat{P}(r)$ as above we strictly follow the idea of (Knight, 2001, 2002). While in (Knight, 2001, 2002) formulas of the object level are propositional formulas, here formulas of the object level are probabilistic conditionals. Accordingly, we define the probability of a formula of the object level as the sum of the probabilities of all interpretations satisfying this formula, i. e., the sum of the probabilities $\hat{P}(P)$ of probability functions P that satisfy the probabilistic conditional. Then, a knowledge base $\mathcal{R}$ is η-consistent if there is a probability function $\hat{P}$ on $\mathcal{P}^{\mathrm{P}}(\mathsf{At})$ such that $\hat{P}(r) \geq \eta$ for every $r \in \mathcal{R}$. Note that this definition as well is a straightforward translation from the concept of probability functions on classical logic to the concept of probability functions on probability functions on classical logic. In particular, given a probability function $\hat{P} : \mathcal{P}^{\mathrm{P}}(\mathsf{At}) \to [0,1]$, the higher the probability $\hat{P}(r)$ for some probabilistic conditional r the better r can be explained by the probability functions in $\mathcal{P}^{\mathrm{P}}(\mathsf{At})$ from the point of view of $\hat{P}$.

At a first glance, this definition also preserves many properties from η-consistency for classical theories. In particular, we can make the following observations.

Proposition 3.30. *If $\mathcal{R}$ is consistent then $\mathcal{R}$ is maximally 1-consistent.*

Proof. As $\mathcal{R}$ is consistent let P be a probability function such that $P \models^{pr} \mathcal{R}$. Define $\hat{P}_P$ via $\hat{P}_P(P) = 1$ and $\hat{P}_P(P') = 0$ for all $P' \in \mathcal{P}^{\mathrm{P}}(\mathsf{At})$ with $P' \neq P$. It follows that $\mathcal{R}$ is 1-consistent with respect to $\hat{P}_P$ and as there can be no $\eta' > 1$ such that $\mathcal{R}$ is η'-consistent it follows that $\mathcal{R}$ is maximally 1-consistent. $\square$

The following proposition is a direct extension of Theorem 2.12 in (Knight, 2002).

Proposition 3.31. *If $\mathsf{MI}(\mathcal{R}) = \{\mathcal{R}\}$ then $\mathcal{R}$ is maximally $(1 - 1/|\mathcal{R}|)$-consistent.*

Proof. Let $\mathcal{R} = \{r_1, \ldots, r_n\}$ and let $\mathcal{R}_1, \ldots, \mathcal{R}_n$ be defined via $\mathcal{R}_i =_{def} \mathcal{R} \setminus \{r_i\}$ for $i = 1, \ldots, n$. Each $\mathcal{R}_i$ for $i = 1, \ldots, n$ is consistent as $\mathcal{R}$ is minimally inconsistent. Therefore, let $P_1, \ldots, P_n$ be probability functions with $P_i \models^{pr} \mathcal{R}_i$ for $i = 1, \ldots, n$. Define $\hat{P}$ through $\hat{P}(P_i) =_{def} 1/n$ and $\hat{P} = 0$ for all $P \in \mathcal{P}^{\mathrm{P}}(\mathsf{At})$ with $P \notin \{P_1, \ldots, P_n\}$. Note that every r_i is contained in every $\mathcal{R}_j$ with $j \neq i$. Therefore, all probability functions P_j with $j \neq i$ satisfy r_i and it follows

$$
\begin{aligned}
\hat{P}(r_i) &= \hat{P}(P_1) + \ldots + \hat{P}(P_{i-1}) + \hat{P}(P_{i+1}) + \ldots + \hat{P}(P_n) \\
&= \frac{n-1}{n} = 1 - \frac{1}{n}
\end{aligned}
$$

It follows that $\hat{P}(r_i) = 1 - 1/n$ for every $i = 1, \ldots, n$ and, hence, $\mathcal{R}$ is $(1 - 1/n)$-consistent. It is also easy to see that there can be no $\hat{P}'$ with $\hat{P}'(r_i) > 1 - 1/n$ for all $i = 1, \ldots, n$, see (Knight, 2002) for details. $\square$

Although it seems that η-consistency is apt for measuring inconsistency in probabilistic conditional logic it exhibits some problems as the following example shows.

Example 3.19. Let $\mathcal{R} =_{def} \{r_1, r_2\}$ and $\mathcal{R}' =_{def} \{r_1', r_2'\}$ be knowledge bases given via

$$
\begin{aligned}
r_1 &=_{def} (a)[0.51] & \qquad r_2 &=_{def} (a)[0.49] \\
r_1' &=_{def} (a)[0.9] & \qquad r_2' &=_{def} (a)[0.1]
\end{aligned}
$$

Note that $\mathcal{R}'$ is far more inconsistent than $\mathcal{R}$ with respect to the minimal distance to a consistent knowledge base, in particular it holds that $\mathsf{Inc}^*(\mathcal{R}) = 0.02$ and $\mathsf{Inc}^*(\mathcal{R}') = 0.8$. However, the notion of η-consistency cannot differentiate between those knowledge base and it is easy to see that both $\mathcal{R}$ and $\mathcal{R}'$ are maximally $1/2$-consistent (this is an implication of Proposition 3.31).

The above observation suggests that η-consistency suffers from the same disadvantages as the inconsistency measures presented in Section 3.3. However, we leave it for future work to investigate the notion of η-consistency further for probabilistic conditional logic.

3.5.2 *An Inconsistency Measure based on Generalized Divergence*

The work discussed above comes from the field of measuring inconsistency in classical theories. We now turn to an inconsistency measure that has been explicitly developed for the use in probabilistic conditional logic. In (Rödder and Xu, 2001) an inconsistency measure is presented that bases on the notion of *generalized divergence* which generalizes cross-entropy. Given vectors $\vec{y}, \vec{z} \in (0,1]^n$ with $\vec{y} = (y_1, \ldots, y_n)$ and $\vec{z} = (z_1, \ldots, z_n)$, the generalized divergence $D(\vec{y}, \vec{z})$ from $\vec{y}$ to $\vec{z}$ is defined

$$D(\vec{y}, \vec{z}) =_{def} \sum_{i=1}^{n} y_i \mathsf{ld} \frac{y_i}{z_i} - y_i + z_i \quad .$$

We abbreviate further

$$D^2(\vec{y}, \vec{z}) =_{def} D(\vec{y}, \vec{z}) + D(\vec{z}, \vec{y}) = \sum_{i=1}^{n} y_i \mathsf{ld} \frac{y_i}{z_i} + z_i \mathsf{ld} \frac{z_i}{y_i} \quad .$$

Let $\mathcal{R}$ be a knowledge base with $\langle \mathcal{R} \rangle = (r_1, \ldots, r_n)$ and $r_i = (\psi_i \,|\, \phi_i)[d_i]$ for $i = 1, \ldots, n$. Then the inconsistency measure Inc_{gd} is defined via

$$\begin{aligned} &\mathsf{Inc}_{gd}(\mathcal{R}) =_{def} \min\{D^2(\vec{y}, \vec{z}) \mid P \in \mathcal{P}^{\mathrm{P}}(\mathsf{At}) \text{ and} \\ &\quad y_i = (1 - d_i)P(\psi_i \phi_i) \text{ and } z_i = d_i P(\overline{\psi_i} \phi_i) \text{ for } i = 1, \ldots, n\} \quad . \end{aligned} \tag{3.31}$$

Let $\vec{y}^*, \vec{z}^*, P^*$ be some parameters such that $D^2(\vec{y}^*, \vec{z}^*)$ is minimal and $y_i^* = (1 - d_i)P^*(\psi_i \phi_i)$ and $z_i^* = d_i P^*(\overline{\psi_i} \phi_i)$ are satisfied for $i = 1, \ldots, n$. Then it follows that

$$y_i^* - z_i^* = P^*(\psi_i \phi_i) - d_i P^*(\phi_i) \quad . \tag{3.32}$$

for $i = 1, \ldots, n$. Minimizing $D^2(\vec{y}, \vec{z})$ amounts to finding a probability function P^* such that $\vec{y}^*$ and $\vec{z}^*$ are as close as possible to each other with respect to D^2. In particular, if there is a P^* such that $y^* = z^*$ it follows that $P^*(\psi_i \phi_i) - d_i P^*(\phi_i) = 0$ and therefore $P^* \models^{pr} (\psi_i \,|\, \phi_i)[d_i]$ (for $i = 1, \ldots, n$), i.e, $\mathcal{R}$ is consistent. Furthermore, the more y_i^* differs from z_i^* the more $P^*(\psi_i \,|\, \phi_i)$ differs from d_i (for $i = 1, \ldots, n$). The measures Inc_{gd} and Inc^* are similar in spirit as they both minimize the distance of a knowledge base to a consistent one. However, the implementation of those measures is different as they use different distance measures. The MinDev inconsistency measure uses the 1-norm distance and the measure Inc_{gd} uses generalized divergence.

In (Rödder and Xu, 2001), no justification is given why Inc_{gd} is a reasonable inconsistency measure for probabilistic conditional knowledge bases. In particular, no motivation is given for the use of generalized divergence as a distance measure and its commonsense interpretation. However, due to the similarity of Inc_{gd} and Inc^* (in terms of the employed paradigm) it follows that Inc_{gd} satisfies the same properties as Inc^*.

Theorem 3.2. *The function* Inc_{gd} *satisfies (Consistency), (Monotonicity), (Super-Additivity), (Weak Independence), (Independence), and (Continuity).*

The proof of Theorem 3.2 can be found in Appendix A on page 236. As for (MinInc Separability) we can also only conjecture its satisfaction for Inc_{gd}.

Conjecture 3.2. *The function* Inc_{gd} *satisfies (*MinInc *Separability).*

The function Inc_{gd} fails to satisfy (Irrelevance of Syntax) for the same reasons as Inc^* fails to satisfy (Irrelevance of syntax), i. e., Inc_{gd} exhibits the same problem as Inc^* in Example 3.13 on page 65. This also applies to the properties (Penalty) and (Normalization).

The difference between Inc^* and Inc_{gd} lies mainly in the scaling of the inconsistency values. Consider the knowledge bases $\mathcal{R}_1$, $\mathcal{R}_2$, and $\mathcal{R}_3$ given via

$$\begin{aligned}
\mathcal{R}_1 &=_{def} \{(b \mid a)[0.9], (a)[0.9], (b)[0.1]\} \\
\mathcal{R}_2 &=_{def} \{(b \mid a)[0.9], (a)[0.9], (b)[0.01]\} \\
\mathcal{R}_3 &=_{def} \{(b \mid a)[0.9], (a)[0.9], (b)[0.001]\}
\end{aligned}$$

As one can see, the difference between the above knowledge bases lie only in the value of the last probabilistic conditional $(b)[x]$ which has been modified by the value -0.09 from the first to the second knowledge base and by -0.009 from the second to the third knowledge base. The inconsistency values of $\mathcal{R}_1$, $\mathcal{R}_2$, and $\mathcal{R}_3$ with respect to the inconsistency measures Inc^* and Inc_{gd} amount to[2]

$$\begin{aligned}
\mathsf{Inc}_{gd}(\mathcal{R}_1) &\approx 2.536 & \qquad \mathsf{Inc}^*(\mathcal{R}_1) &= 0.71 \\
\mathsf{Inc}_{gd}(\mathcal{R}_2) &\approx 4.361 & \qquad \mathsf{Inc}^*(\mathcal{R}_2) &= 0.8 \\
\mathsf{Inc}_{gd}(\mathcal{R}_3) &\approx 5.904 & \qquad \mathsf{Inc}^*(\mathcal{R}_3) &= 0.809 \quad .
\end{aligned}$$

The differences of the inconsistency values with respect to Inc^* perfectly resemble the modifications in the original knowledge bases, i. e., it holds that

$$\begin{aligned}
\mathsf{Inc}^*(\mathcal{R}_2) - \mathsf{Inc}^*(\mathcal{R}_1) &= 0.09 \\
\mathsf{Inc}^*(\mathcal{R}_3) - \mathsf{Inc}^*(\mathcal{R}_2) &= 0.009 \quad .
\end{aligned}$$

This is not the case for the measure Inc_{gd} as the differences of the inconsistency values amount to

$$\begin{aligned}
\mathsf{Inc}_{gd}(\mathcal{R}_2) - \mathsf{Inc}_{gd}(\mathcal{R}_1) &= 1.825 \\
\mathsf{Inc}_{gd}(\mathcal{R}_3) - \mathsf{Inc}_{gd}(\mathcal{R}_2) &= 1.543 \quad .
\end{aligned}$$

Note also that the ratio of the change in the inconsistency value does not match the modifications of the knowledge bases. More specifically, the change of the probabilities from $\mathcal{R}_1$ to $\mathcal{R}_2$—the probability of (b) has been changed by the value 0.09 from 0.1 to 0.01—is ten times more drastic than the change of the probabilities from $\mathcal{R}_2$ to $\mathcal{R}_3$—by the value 0.009from 0.01

2 The inconsistency values with respect to Inc_{gd} have been computed using the SPIRIT expert system shell version 3.7.3.2.

to 0.001. However, the ratio in the change of the inconsistency values with respect to Inc_{gd} amounts to

$$\frac{1.825}{1.543} \approx 1.183$$

while for the measure Inc^* we get $0.09/0.009 = 10$.

We argue that this *linear* behavior of Inc^* with respect to modifications is more intuitive as it helps the knowledge engineer to assess the severity of the inconsistency in a better way. An inconsistency value $x = \mathsf{Inc}^*(\mathcal{R})$ can be *roughly* interpreted as the need to change the probabilities of the conditionals by a total of x. By also taking the number of conditionals in the knowledge base into account the knowledge engineer gets a good idea of the effort needed to manually repair the knowledge base. As for Inc_{gd}, there is no linear connection between the inconsistency value and the effort needed to restore consistency. However, one might also argue that other criteria are more appropriate for comparing inconsistency values than those linear measures, e. g. comparing the values on a logarithmic scale. In particular, as conditional probabilities are defined as the ratio of probabilities differences in conditional probabilities may have different effects depending on the actual conditional probability. For example, varying conditional probabilities between values 0.4 and 0.5 has by far less effect (in terms of the underlying probabilistic model) than varying the probability between 0 and 0.1, as the latter contains the transition between strict and uncertain belief[3]. We leave it for future work to investigate this topic.

3.5.3 *Candidacy Degrees of Best Candidates*

The motivation of the work (Daniel, 2009) is very similar to ours, namely, probabilistic reasoning with inconsistent information. Among others, one contribution of (Daniel, 2009) is an inconsistency measure on knowledge bases of probabilistic constraints. In particular, the work (Daniel, 2009) focuses on linear probabilistic knowledge bases as discussed in Section 3.4.3 but also considers generalizations such as polynomial probabilistic knowledge bases. However, in order to compare it to our work we simplify several notations and present the inconsistency measure Inc^h_μ of (Daniel, 2009) only for probabilistic conditional logic.

The central notion of (Daniel, 2009) is the *candidacy function*. A candidacy function is similar to a fuzzy set (Gerla, 2001) as it assigns a degree of membership of a probability function belonging to the models of a knowledge base. For the rest of this section let At be some propositional signature. Then a candidacy function $\mathfrak{C}$ is a function $\mathfrak{C} : \mathcal{P}^{\mathrm{P}}(\mathsf{At}) \to [0,1]$. A uniquely determined candidacy function $\mathfrak{C}_{\mathcal{R}}$ can be assigned to a (consistent or inconsistent) knowledge base $\mathcal{R}$ as follows. For a probability function $P \in \mathcal{P}^{\mathrm{P}}(\mathsf{At})$ and a set of probability functions $S \subseteq \mathcal{P}^{\mathrm{P}}(\mathsf{At})$ let $d(P,S)$ denote

3 I thank Gabriele Kern-Isberner for pointing this out to me.

the distance of P to S with respect to the Euclidean norm, i. e., $d(P,S)$ is defined via

$$d(P,S) =_{def} \inf \left\{ \sqrt{\sum_{\omega \in \Omega(\mathsf{At})} (P(\omega) - P'(\omega))^2} \mid P' \in S \right\} \quad .$$

Let $h : \mathbb{R}^+ \to (0,1]$ be a strictly decreasing, positive, and continuous log-concave function with $h(0) = 1$. Then the candidacy function $\mathfrak{C}^h_{\mathcal{R}}$ for a knowledge base $\mathcal{R}$ is defined as

$$\mathfrak{C}^h_{\mathcal{R}}(P) =_{def} \prod_{r \in \mathcal{R}} h\left(\sqrt{2^{|\mathsf{At}|}} d(P, \mathsf{Mod}^{\mathrm{Pr}}(\{r\}))\right)$$

for every $P \in \mathcal{P}^{\mathrm{P}}(\mathsf{At})$. Note that the definition of the candidacy function $\mathfrak{C}^h_{\mathcal{R}}$ depends on the size of the signature At. The intuition behind this definition is that a probability function P that is near to the models of each probabilistic conditional in $\mathcal{R}$ gets a high candidacy degree in $\mathfrak{C}^h_{\mathcal{R}}(P)$. It is easy to see that it holds that $\mathfrak{C}^h_{\mathcal{R}}(P) = 1$ if and only if $P \models^{pr} \mathcal{R}$. Using the candidacy function $\mathfrak{C}^h_{\mathcal{R}}$ the inconsistency measure Inc^h_{μ} can be defined via

$$\mathsf{Inc}^h_{\mu}(\mathcal{R}) =_{def} 1 - \max_{P \in \mathcal{P}^{\mathrm{P}}(\mathsf{At})} \mathfrak{C}^h_{\mathcal{R}}(P)$$

for a knowledge base $\mathcal{R}$. In (Daniel, 2009) it is shown that Inc^h_{μ} satisfies (among others) the following properties.

Proposition 3.32. Inc^h_{μ} *satisfies (Consistency), (Monotonicity), (Continuity), and (Normalization).*

The function Inc^h_{μ} does not satisfy (Irrelevance of Syntax) as we show in the following example.

Example 3.20. Let $\mathcal{R} =_{def} \{r_1, r_2\}$ and $\mathcal{R}' =_{def} \{r_1, r_2, r_3, r_4\}$ be knowledge bases with

$$r_1 =_{def} (a)[0.7] \qquad\qquad r_2 =_{def} (a)[0.3]$$
$$r_3 =_{def} (\neg a)[0.3] \qquad\qquad r_4 =_{def} (\neg a)[0.7]$$

Note that $\mathcal{R}_1 \equiv^{\mathrm{cond}} \mathcal{R}_2$ as $r_1 \equiv^{\mathrm{pr}} r_3$ and $r_2 \equiv^{\mathrm{pr}} r_4$. Let $P \in \mathcal{P}^{\mathrm{P}}(\mathsf{At})$ be some probability function, $h : \mathbb{R}^+ \to (0,1]$ a strictly decreasing, positive, and continuous log-concave function with $h(0) = 1$ and define

$$x_i =_{def} h\left(\sqrt{2^{|\mathsf{At}|}} d(P, \mathsf{Mod}^{\mathrm{Pr}}(\{r_i\}))\right)$$

for $i = 1, \ldots, 4$. Note that $x_1 = x_3$ as $\mathsf{Mod}^{\mathrm{Pr}}(\{r_1\}) = \mathsf{Mod}^{\mathrm{Pr}}(\{r_3\})$ and $x_2 = x_4$ as $\mathsf{Mod}^{\mathrm{Pr}}(\{r_2\}) = \mathsf{Mod}^{\mathrm{Pr}}(\{r_4\})$. As $\mathcal{R}$ is inconsistent it follows that it cannot be the case that both $x_1 = x_3 = 1$ and $x_2 = x_4 = 1$ because P cannot be a model of both r_1 and r_2. It follows that $x_1 x_2 < 1$ and $x_3 x_4 < 1$ and hence

$$\mathfrak{C}^h_{\mathcal{R}}(P) = x_1 x_2 > x_1 x_2 x_3 x_4 = C^h_{\mathcal{R}'}(P) \quad .$$

Therefore, for every $P \in \mathcal{P}^{\mathrm{P}}(\mathsf{At})$ it holds that $\mathfrak{C}^h_{\mathcal{R}}(P) > \mathfrak{C}^h_{\mathcal{R}'}(P)$ and consequently $\mathsf{Inc}^h_\mu(\mathcal{R}) < \mathsf{Inc}^h_\mu(\mathcal{R}')$ violating (Irrelevance of Syntax).

In Example 3.2 on page 51 we talked about the issue of an inconsistency measure satisfying all three of (Consistency), (Super-Additivity), and (Normalization). We showed that an inconsistency measure that does not take the cardinality of the signature into account cannot satisfy all these properties at once. As one can see above, the function Inc^h_μ takes the cardinality of the signature into account and it may be possible that Inc^h_μ satisfies (Super-Additivity). However, this is not the case as the following example shows.

Example 3.21. Let $\mathsf{At} = \{a_1, a_2\}$ be a propositional signature and let $\mathcal{R}_1 =_{def} \{r_1, r_2\}$ and $\mathcal{R}_2 =_{def} \{r_3, r_4\}$ be knowledge bases with

$$r_1 =_{def} (a_1)[1] \qquad\qquad r_2 =_{def} (a_1)[0]$$
$$r_3 =_{def} (a_2)[1] \qquad\qquad r_4 =_{def} (a_2)[0]$$

and let $\mathcal{R} =_{def} \mathcal{R}_1 \cup \mathcal{R}_2$. Note that both $\mathcal{R}_1$ and $\mathcal{R}_2$ are inconsistent and $\mathcal{R}_1 \cap \mathcal{R}_2 = \emptyset$. As Inc^h_μ is defined on the semantic level and does not take the names of propositions into account it follows that $\mathsf{Inc}^h_\mu(\mathcal{R}_1) = \mathsf{Inc}^h_\mu(\mathcal{R}_2)$. As the situations in $\mathcal{R}_1$ and $\mathcal{R}_2$ are symmetric and $\mathcal{R}_i$ is symmetric with respect to r_1 and r_2 and with respect to r_3 and r_4 there are probability functions P_i with $\mathsf{Inc}^h_\mu(\mathcal{R}_i) = 1 - C^h_{\mathcal{R}_i}(P_i)$ for $i = 1, 2$ and

$$\begin{aligned} d(P_1, \mathsf{Mod}^{\mathrm{Pr}}(\{r_1\})) &= d(P_1, \mathsf{Mod}^{\mathrm{Pr}}(\{r_2\})) \\ &= d(P_2, \mathsf{Mod}^{\mathrm{Pr}}(\{r_3\})) \\ &= d(P_2, \mathsf{Mod}^{\mathrm{Pr}}(\{r_4\})) \quad . \end{aligned}$$

Let $x =_{def} d(P_1, \mathsf{Mod}^{\mathrm{Pr}}(\{r_1\}))$ and let $h^* : \mathbb{R}^+ \to (0, 1]$ be a strictly decreasing, positive, and continuous log-concave function with $h^*(0) = 1$ and $h^*\left(\sqrt{2^{|\mathsf{At}|}}x\right) = 0.5$. Then it follows $C^{h^*}_{\mathcal{R}_1}(P_1) = 0.25$ and $\mathsf{Inc}^{h^*}_\mu(\mathcal{R}_1) = 0.75$. In order to satisfy (Super-Additivity) $\mathsf{Inc}^{h^*}_\mu$ must satisfy

$$\mathsf{Inc}^{h^*}_\mu(\mathcal{R}) \geq \mathsf{Inc}^{h^*}_\mu(\mathcal{R}_1) + \mathsf{Inc}^{h^*}_\mu(\mathcal{R}_2) = 1.5$$

which is a contradiction since $\mathsf{Inc}^{h^*}_\mu$ satisfies (Normalization).

In (Daniel, 2009) it is shown that Inc^h_μ satisfies several other properties that cannot be related directly to our properties of Section 3.2. For example, (Daniel, 2009) investigates the property *consequence invariance* which is similar to (Independence) but not equivalent. The property *consequence invariance* is defined as

$$\text{If } \mathcal{R} \models^* r \text{ then } \mathsf{Inc}(\mathcal{R}) = \mathsf{Inc}(\mathcal{R} \cup \{r\})$$

with $\models^*$ is one of $\models_{\text{ic}}, \models_{\text{ff}}$ as defined in (Daniel, 2009). It is shown that Inc^h_μ satisfies *consequence invariance* for both entailment relations. However, there are no results on the behavior of Inc^h_μ by addition of unrelated conditionals such as free conditionals in the sense of Definition 3.2 on page 48. However, it can be easily seen that Inc^h_μ also fails to satisfy (Penalty) for similar reasons as Inc^* fails to satisfy (Penalty). For the knowledge base $\mathcal{R} =_{def} \{(b \,|\, a)[1], (a)[1], (b)[0]\}$ let P' be such that

$$\max_{P \in \mathcal{P}^{\mathrm{P}}(\mathsf{At})} \mathfrak{C}^h_{\mathcal{R}}(P) = \mathfrak{C}^h_{\mathcal{R}}(P') \quad . \tag{3.33}$$

In other words, P' is a probability function that has the maximal candidacy degree with respect to $\mathcal{R}$. As $\mathcal{R}$ is inconsistent, it follows that P' fails to satisfy at least one of the probabilistic conditionals of $\mathcal{R}$. Assume that it holds that $P' \not\models^{pr} (b \,|\, a)[1]$ which implies $P'(a) > 0$. Consider the knowledge base $\mathcal{R}' =_{def} \mathcal{R} \cup \{r'\}$ with $r' =_{def} (b \,|\, a)[P'(b \,|\, a)]$. As Inc^h_μ satisfies (Monotonicity) it follows $\mathsf{Inc}^h_\mu(\mathcal{R}') \geq \mathsf{Inc}^h_\mu(\mathcal{R})$ and due to $h\left(\sqrt{2^{|\mathsf{At}|}} d(P', \mathsf{Mod}^{\mathrm{Pr}}(\{r'\}))\right) = 1$ (as $d(P', \mathsf{Mod}^{\mathrm{Pr}}(\{r'\})) = 0$) it follows that P' also satisfies

$$\max_{P \in \mathcal{P}^{\mathrm{P}}(\mathsf{At})} \mathfrak{C}^h_{\mathcal{R}'}(P) = \mathfrak{C}^h_{\mathcal{R}'}(P') \quad .$$

Therefore, P' has also maximal candidacy degree with respect to $\mathcal{R}'$ which is clear as we only added information consistent with P' (otherwise P' would have violated (3.33). It follows $\mathsf{Inc}^h_\mu(\mathcal{R}') \leq \mathsf{Inc}^h_\mu(\mathcal{R})$ and as $\{(b \,|\, a)[1], (a)[1], r'\}$ is a minimal inconsistent subset of $\mathcal{R}'$ this contradicts (Penalty). Similar observations can be made when $P' \not\models^{pr} (a)[1]$ or $P' \not\models^{pr} (b)[0]$. It is outside the scope of this thesis to check whether Inc^h_μ satisfies any of the properties (Weak Independence), (Independence), and (MinInc Separability). Hence, we leave a deeper discussion of the inconsistency measure Inc^h_μ for future work.

3.6 SUMMARY AND DISCUSSION

In this chapter we discussed the problem of analyzing and measuring inconsistencies in probabilistic conditional logic. We developed a series of rationality postulates for inconsistency measures and extended inconsistency measures for classical theories to the probabilistic setting. Furthermore, we proposed a novel inconsistency measure for probabilistic conditional logic that bases on a specific notion of distance to consistent knowledge bases. We showed that the MinDev inconsistency measure is more apt for the probabilistic setting than the traditional approaches. We continued with the development of approximations to the MinDev inconsistency measure and extended it to two more general frameworks for probabilistic reasoning. Finally, we reviewed related work regarding inconsistency measurement in probabilistic reasoning.

Property	Inc^d	$\mathsf{Inc}^{\mathsf{MI}}$	$\mathsf{Inc}_0^{\mathsf{MI}}$	$\mathsf{Inc}_C^{\mathsf{MI}}$	$\mathsf{Inc}_{C,0}^{\mathsf{MI}}$	Inc^*	Inc_0^*	Inc_{gd}	Inc_μ^h
(Consistency)	X	X	X	X	X	X	X	X	X
(Irr. of Syntax)	X								
(Monotonicity)	X	X		X		X		X	X
(Super-Additivity)		X		X		X		X	
(Weak Indep.)	X	X		X		X		X	?
(Independence)	X	X		X		X		X	?
(MinInc Sep.)		X		X		?		?	?
(Penalty)		X		X					
(Continuity)						X	X	X	X
(Normalization)	X		X		X		X		X

Table 3: Comparison of inconsistency measures

Table 3 shows a comparison of the inconsistency measures discussed in this chapter with respect to their properties (a question mark indicates that it has not been proven whether this property holds or not). As one can see, the measures that satisfy a maximal number of properties are $\mathsf{Inc}^{\mathsf{MI}}$, $\mathsf{Inc}_C^{\mathsf{MI}}$, Inc^*, and Inc_{gd} (under the assumption that Conjectures 3.1 and 3.2 are true). These measures differ only in their satisfaction of the properties (Penalty) and (Continuity). In the presence of quantified knowledge representation satisfaction of (Continuity) seems more important than satisfaction of (Penalty), see the discussion at the end of Section 3.3. Due to the lack of formal results regarding several properties of Inc_μ^h a classification of Inc_μ^h is not possible at this time. Appendix C on page 253ff. lists inconsistency values for the inconsistency measures Inc^d, $\mathsf{Inc}^{\mathsf{MI}}$, $\mathsf{Inc}_0^{\mathsf{MI}}$, $\mathsf{Inc}_C^{\mathsf{MI}}$, $\mathsf{Inc}_{C,0}^{\mathsf{MI}}$, Inc^*, Inc_0^*, and the approximations $\mathcal{I}^{\leq}$, $\mathcal{I}^{\geq}$, $\mathcal{I}_0^{\leq}$, and $\mathcal{I}_0^{\geq}$ for Inc^* and Inc_0^*, respectively, on several benchmark examples. Some of these examples have already been discussed throughout this chapter. The implementations of the inconsistency measures—which were used for computing the values shown in Appendix C—can be found in the Tweety library for artificial intelligence[4].

4 `http://sourceforge.net/projects/tweety/`

Analyzing inconsistencies is of major concern in the area of knowledge representation as consistency is a necessary prerequisite for many knowledge representation formalisms. In particular, the task of inference bases mostly on the consistency of the underlying information. In this chapter we discussed this problem for a specific approach to uncertain knowledge representation, i. e. for probabilistic conditional logic. However, analyzing inconsistencies is only the first step towards *handling inconsistencies*. From our current point there are mainly two possible directions for approaching this goal: 1.) one can acknowledge inconsistency and find a way to perform reasoning based on inconsistent information, or 2.) one can remove the inconsistency and perform reasoning based on consistent information. For (propositional) probabilistic conditional logic an approach that pursues the first direction has previously been proposed under the notion of *paraconsistent probabilistic reasoning* in (Daniel, 2009). In the next chapter we pursue the second direction by exploiting the ideas developed in this chapter for solving the conflicts in a knowledge base. We also consider pursuing the first direction but for relational probabilistic conditional logic in Chapter 6.

4 SOLVING CONFLICTS USING INCONSISTENCY MEASURES

The problem of resolving inconsistencies is *the* major issue in many subfields of artificial intelligence. As noted at the end of the previous chapter there are basically two approaches for dealing with inconsistencies in general. On the one hand, one can try to cope with inconsistencies in the knowledge representation formalism and resolve contradictory information when information is to be inferred from the knowledge base. This approach is pursued in many fields such as default reasoning (Reiter, 1980; Antoniou, 1999), answer set programming (Gelfond and Lifschitz, 1991; Gelfond and Leone, 2002), or argumentation (Bench-Capon and Dunne, 2007; Rahwan and Simari, 2009). Frameworks from these fields usually allow knowledge bases to contain contradictory information but provide methods that guarantee conflict-free inference. For example, in formal frameworks for argumentation such as defeasible logic programming (Garcia and Simari, 2004) the inference procedure includes an analysis phase where proofs for and proofs against some proposition are compared with each other in order to come up with a coherent view on the information. On the other hand, one can insist on a knowledge representation formalism that allows only for consistent knowledge bases and try to maintain consistency whenever contradictory information has to be added. This approach is pursued in many fields as well, such as belief revision (Alchourrón *et al.*, 1985; Hansson, 1999) and information fusion (Bloch and Hunter, 2001). The motivation for achieving a consistent state of the knowledge base derives from application scenarios such as updating or merging beliefs of experts. For instance, given a set of experts on some field such as medicine, joining the beliefs of the experts usually results in an inconsistent knowledge base. The goal of information fusion is to come up with a consistent representation of the merged knowledge of these experts. As our motivation for dealing with inconsistent information in probabilistic conditional logic is the same as for information fusion, i. e. joining beliefs from different sources, we are going to require consistency for knowledge bases as well and focus on the problem of how to restore consistency. However, we are not considering the sources within our discussion but only consider an inconsistent knowledge base as given.

As for frameworks based on classical logic the issue of handling inconsistency has been addressed in several works within the information fusion and belief merging community, see e. g. (Bloch and Hunter, 2001) for a survey and (Everaere *et al.*, 2008; Lang and Marquis, 2010) for some recent developments. Most of these approaches rely on removing specific pieces of information in order to restore consistency. While this method is also possible for probabilistic conditional logic it seems rather extreme in the presence of a much less invasive approach to alter the information represented by a knowledge base: modifying probabilities. However, in (Finthammer *et al.*, 2007) a method is presented that—based on a set of heuristics—restores consistency in probabilistic conditional knowledge bases essentially by removing conditionals. We review the methods of (Finthammer *et al.*, 2007) in Section 4.5.3. In contrast to (Finthammer *et al.*, 2007), we focus our discussion on solving conflicts via modifying probabilities and keeping the qualitative structure of a knowledge base intact. Of course, we do so by obeying the *minimal change paradigm*. But before coming to the actual problem of restoring consistency we have a look at *culpability measures* (Daniel, 2009) first. A culpability measure is a generalization of an inconsistency measure that does not assign a degree of inconsistency to a knowledge base but to each conditional of a knowledge base separately. Therefore, a culpability measure allows not only to get an idea of the severity of the inconsistency but also to obtain information on which parts of the knowledge base are to be blamed for producing an inconsistency. We propose and discuss two different culpability measures. One of them—the Shapley culpability measure—is inspired by previous work by Hunter and Konieczny (Hunter and Konieczny, 2006, 2008, 2010) who use the Shapley value (Shapley, 1953) to define a culpability measure for classical theories.

Using culpability measures we investigate the problem of restoring consistency in a probabilistic conditional knowledge base in a principled way. We implement the general requirement of *minimal change* by developing a series of rationality postulates for consistency restorers that are influenced partially by the role of culpabilities of conditionals. The central idea of restoring consistency in knowledge bases relies on the requirement to keep the structure of a knowledge base intact and to modify the probability of each conditional proportionally to the culpability of the conditional in a minimal way such that the resulting knowledge base is consistent. We take some first steps into the development of constructive approaches for restoring consistency and propose two different families of consistency restorers. Our first approach relies on the notion of *creeping functions* which are functions $\Xi_{\mathcal{R}}$ on $[0,1]$ for a knowledge base $\mathcal{R}$ such that $\Xi_{\mathcal{R}}(0) = \mathcal{R}$ and $\Xi_{\mathcal{R}}(1)$ is consistent. We give three different implementations of such creeping functions based on culpability measures. The task of restoring consistency then reduces to finding a minimal δ^* such that $\Xi_{\mathcal{R}}(\delta^*)$ is consistent. Our second approach is more declarative in nature and is inspired by the characterization of Inc^* using the optimization problem $\mathsf{DevCons}(\mathcal{R})$. As discussed in Chapter 3 the cardinality of the set $\mathcal{MD}(\mathcal{R})$ (see page 63) of vectors that minimize the 1-norm distance to a consistent knowledge base is usually greater than one. Therefore, just using one of these arguments to define the outcome of restoring consistency is not appropriate as it violates the functional requirement of a consistency restorer, i. e., for an inconsistent knowledge base $\mathcal{R}$ the outcome $\mathcal{R}'$ of the consistency restoring process should be uniquely determined. Our second approach relies on restricting $\mathsf{DevCons}(\mathcal{R})$ by introducing requirements based on culpabilities such that the outcome is uniquely determined.

This chapter is organized as follows. In the next section we discuss the general concept of culpability measures and propose two specific measures, the Shapley culpability measure and the mean distance culpability measure. We continue in Section 4.2 with discussing the problem of restoring consistency in knowledge bases in a principled fashion by developing a series of rationality postulates for methods that restore consistency. Afterwards, we propose two different families of methods for restoring consistency. First, in Section 4.3 we discuss the method of creeping functions which are a constructive approach to restore consistency. Afterwards, in Section 4.4 we discuss a declarative approach that bases on the optimization problem to determine the minimal distance to consistent knowledge bases developed in the previous chapter. We go on with a review of related work in Section 4.5 and conclude with a summary and some remarks in Section 4.6.

4.1 CULPABILITY MEASURES

In the following, we adopt the notion of *culpability measure* (Daniel, 2009) in order to give a more fine-grained idea of how inconsistency is distributed over the pieces of information in a knowledge base.[1] Note, that the actual definition of a culpability measure given here has been modified with respect to the definition in (Daniel, 2009) in order to fit our framework.

Definition 4.1 (Culpability measure)**.** A *culpability measure* $C^{\mathcal{R}}$ for a knowledge base $\mathcal{R} = \{r_1, \ldots, r_n\}$ is a function mapping each $r_i \in \mathcal{R}$ to a non-negative real value $C^{\mathcal{R}}(r_i) \in \mathbb{R}_0^+$, i.e., it holds that $C^{\mathcal{R}} : \mathcal{R} \to \mathbb{R}_0^+$.

A culpability measure $C^{\mathcal{R}}$ for a knowledge base $\mathcal{R}$ assigns to each conditional $r_i \in \mathcal{R}$ a quantitative value of the "blame" to be given for creating an inconsistency. We call $C^{\mathcal{R}}(r)$ for $r \in \mathcal{R}$ the *culpability value* of r in $\mathcal{R}$ with respect to $C^{\mathcal{R}}$. When talking about culpability values we omit the reference to the culpability measure when this is clear from context. A larger value of $C^{\mathcal{R}}(r_i)$ refers to a larger culpability, so if $C^{\mathcal{R}}(r_i) > C^{\mathcal{R}}(r_j)$ the conditional r_i is more responsible for creating an inconsistency than r_j.

Some properties expected from culpability measures can be found in (Hunter and Konieczny, 2008) and are rephrased here to fit our framework. Let $\mathcal{R} = \{r_1, \ldots, r_n\}$ be a knowledge base, Inc be an inconsistency measure, and $C^{\mathcal{R}}$ be a culpability measure for $\mathcal{R}$.

(Inc-Distribution) $\sum_{r \in \mathcal{R}} C^{\mathcal{R}}(r) = \mathsf{Inc}(\mathcal{R})$.

(Inc-Symmetry) For all $r, r' \in \mathcal{R}$, if $\mathsf{Inc}(\mathcal{R}' \cup \{r\}) = \mathsf{Inc}(\mathcal{R}' \cup \{r'\})$ for all $\mathcal{R}' \subseteq \mathcal{R} \setminus \{r, r'\}$ then $C^{\mathcal{R}}(r) = C^{\mathcal{R}}(r')$.

1 Culpability measures are referred to as *inconsistency values* in the works by Hunter and Konieczny, see e.g. (Hunter and Konieczny, 2008, 2010)

(Minimality) If r is a free conditional in $\mathcal{R}$ then $C^{\mathcal{R}}(r) = 0$.

(Decomposability) If $\mathsf{MI}(\mathcal{R}_1 \cup \mathcal{R}_2) = \mathsf{MI}(\mathcal{R}_1) \cup \mathsf{MI}(\mathcal{R}_2)$ and $\mathsf{MI}(\mathcal{R}_1) \cap \mathsf{MI}(\mathcal{R}_2) = \emptyset$ then for $r \in \mathcal{R}_1 \cap \mathcal{R}_2$ it holds that $C^{\mathcal{R}_1 \cup \mathcal{R}_2}(r) = C^{\mathcal{R}_1}(r) + C^{\mathcal{R}_2}(r)$.

The property (Inc-Distribution) states that the culpability measure $C^{\mathcal{R}}$ distributes the full inconsistency value $\mathsf{Inc}(\mathcal{R})$ between the individual conditionals. Property (Inc-Symmetry) states that two conditionals r and r' that play symmetric roles in producing inconsistencies in $\mathcal{R}$ have identical culpability values. As free conditionals do not take part in producing inconsistencies the property (Minimality) requires them to have culpability values of zero. The property (Decomposability) is a straightforward extension of the property (MinInc Separability) to culpability measures, cf. Section 3.2. In (Hunter and Konieczny, 2008) another property called (Dominance) is mentioned that demands for a piece of information α to have no smaller culpability value than some piece β if β semantically follows from α. In our framework semantical entailment is trivial. By defining $r \models^{pr} r'$ for conditionals r and r' if and only if for every probability function P with $P \models^{pr} r$ it follows $P \models^{pr} r'$, we get $r \models^{pr} r'$ if and only if $r \equiv^{\mathrm{pr}} r'$. By remembering the geometrical interpretation of probabilistic conditionals this is easy to see. Two hyperplanes H_1, H_2 of dimension m are related to each other in only one of three possible ways: either H_1 and H_2 intersect in a hyperplane of dimension $m-1$, or H_1 and H_2 are parallel, or H_1 and H_2 are identical. For two conditionals r and r' this means that either r and r' are consistent but there are P and P' with $P \models^{pr} r$ and $P \not\models^{pr} r'$ and $P' \models^{pr} r'$ and $P' \not\models^{pr} r$, or r and r' are inconsistent, or $r \equiv^{\mathrm{pr}} r'$. It follows that the property (Dominance) is tautological in our framework and thus is neglected in this thesis.

As for inconsistency measures we also consider the property of continuity for culpability measures.

Definition 4.2 (Characteristic culpability function)**.** Let $C^{\mathcal{R}}$ be a culpability measure and let $\mathcal{R}$ be a knowledge base with $\langle \mathcal{R} \rangle = (r_1, \ldots, r_n)$. The function

$$\theta^{\dagger}_{C^{\mathcal{R}},\mathcal{R}} : [0,1]^n \to [0,\infty)^n$$

with

$$\theta^{\dagger}_{C^{\mathcal{R}},\mathcal{R}}(x_1, \ldots, x_n) =_{def} (C^{\Lambda_{\mathcal{R}}(x_1,\ldots,x_n)}(r_1), \ldots, C^{\Lambda_{\mathcal{R}}(x_1,\ldots,x_n)}(r_n))$$

is called the *characteristic culpability function* of $C^{\mathcal{R}}$ and $\mathcal{R}$.

Using the characteristic culpability function we specify our requirement of continuity as follows.

(Continuity) The characteristic culpability function $\theta^{\dagger}_{\mathcal{C}^{\mathcal{R}},\mathcal{R}}$ is continuous on $[0,1]^{|\mathcal{R}|}$ (with respect to the standard topology on $\mathbb{R}^n$).

Culpability measures can be used to guide the knowledge engineer to restore consistency in an inconsistent knowledge base as they can be used as an indicator for how to change the probabilities of the conditionals appropriately. Before investigating these possibilities we first develop two culpability measures, one of them applying the *Shapley value* and the other one using distance minimization.

4.1.1 *Shapley Culpability Measure*

In the following, we use the *Shapley value* (Shapley, 1953) as a means to measure the culpability of each conditional for creating inconsistencies. The Shapley value is a well-known solution for coalition games in game theory and before continuing with defining the culpability measure based on the Shapley value we give a brief overview on coalition game theory as motivation.

Coalition Game Theory

Coalition game theory is concerned with games where players can form coalitions in order to maximize their own payoff of the game.

Definition 4.3 (Coalition game)**.** A *coalition game* (N, v) is composed of a set of players $N \subseteq \mathbb{N}$ and a function $v : \mathfrak{P}(N) \to \mathbb{R}$ with $v(\emptyset) = 0$ and $v(S \cup T) \geq v(S) + v(T)$ for $S, T \subseteq N$ with $S \cap T = \emptyset$.

For every possible coalition $C \subseteq N$ of players in the game, the value $v(C)$ determines the payoff this coalition gets. As this payoff must be distributed on the members of C, every player has to evaluate himself, which coalition to form in order to maximize his or her own expected payoff. Not every player has to expect the same payoff, as players may be more or less important for the forming of coalitions. Consider the following example taken from (Hunter and Konieczny, 2006).

Example 4.1. Let the set of players N be defined as $N =_{def} \{1,2,3\}$ and the function $v : \mathfrak{P}(N) \to \mathbb{R}$ be defined as

$$\begin{aligned} v(\{1\}) &=_{def} 1 & v(\{2\}) &=_{def} 0 & v(\{3\}) &=_{def} 1 \\ v(\{1,2\}) &=_{def} 10 & v(\{2,3\}) &=_{def} 11 & v(\{1,3\}) &=_{def} 4 \\ v(\{1,2,3\}) &=_{def} 12 \end{aligned}$$

In this game, not every player should expect the same payoff, as for instance it is more advantageous for player 1 to form a coalition with player 2 rather than with player 3 alone.

A solution to a coalition game (N, v) consists of an assignment $S_i(v)$ of payoffs to each player $i \in N$, that is *fair* in the sense that every player gets as much payoff as his or her contribution in the grand coalition N weighs. Some formal desirable properties of a solution are as follows.

(Efficiency) $\sum_{i \in N} S_i(v) = v(N)$

(Symmetry) For all $i, j \in N$, if $v(C \cup \{i\}) = v(C \cup \{j\})$ for all $C \subseteq N \setminus \{i, j\}$ then $S_i(v) = S_j(v)$

(Dummy) If for $i \in N$ it holds that $v(C) = v(C \cup \{i\})$ for all $C \subseteq N$ then it holds that $S_i(v) = 0$

(Additivity) $S_i(v + w) = S_i(v) + S_i(w)$ for all $i \in N$

A solution should comprehend for the fact, that the value to be distributed among the players is the maximal value that can be achieved, cf. (Efficiency). If two players are indistinguishable by their contributions to the coalitions they deserve the same payoff, cf. (Symmetry); if a player does not contribute to any coalition at all his or her payoff should be zero, cf. (Dummy). The property (Additivity) describes the desired behavior of a solution if two coalition games are combined.

It can be shown (Shapley, 1953), that the *Shapley value* defined as follows is the only solution for a coalition game that satisfies (Efficiency), (Symmetry), (Dummy), and (Additivity).

Definition 4.4 (Shapley value)**.** Let (N, v) be a coalition game. The *Shapley Value* $S_v(i)$ for a player $i \in N$ is defined as

$$S_v(i) =_{def} \sum_{C \subseteq N} \frac{(|C| - 1)!(|N| - |C|)!}{|N|!}(v(C) - v(C \setminus \{i\})$$

Consider the following example taken from (Hunter and Konieczny, 2006).

Example 4.2. The Shapley values for the players $1, 2, 3$ from Example 4.1 are

$$S_v(i) \approx 2.83 \qquad S_v(i) \approx 5.83 \qquad S_v(i) \approx 3.33 \quad .$$

Measuring Culpabilities using the Shapley Value

The approach discussed in this section is inspired by the works of Hunter and Konieczny (Hunter and Konieczny, 2006, 2008, 2010) where the Shapley value is used to investigate the causes of inconsistency in classical propositional knowledge bases.

Definition 4.5 (Probabilistic Shapley culpability measure)**.** Let Inc be an inconsistency measure and $\mathcal{R}$ be a knowledge base. We define the *probabilistic Shapley culpability measure* $S^{\mathcal{R}}_{\mathsf{Inc}}$ for $\mathcal{R}$ and with respect to Inc as

$$S^{\mathcal{R}}_{\mathsf{Inc}}(r) =_{def} \sum_{C\subseteq\mathcal{R}} \frac{(|C|-1)!(n-|C|)!}{n!}(\mathsf{Inc}(C) - \mathsf{Inc}(C \setminus \{r\}))$$

for all $r \in \mathcal{R}$.

Using the probabilistic Shapley culpability measure we can obtain more specific information about how the inconsistency is distributed among the probabilistic conditionals of a knowledge base. In the following we use the MinDev inconsistency measure Inc^* for the application of the probabilistic Shapley culpability measure.

Example 4.3. Consider again the knowledge base $\mathcal{R}_1 = \{r_1, r_2, r_3, r_4\}$ from Example 3.11 on page 64 given via

$$r_1 = (b\,|\,\overline{a})[0.8] \qquad r_2 = (b\,|\,a)[0.6]$$
$$r_3 = (a)[0.5] \qquad r_4 = (b)[0.2]$$

with $\mathsf{Inc}^*(\mathcal{R}_1) = 0.5$. There, we have

$$S^{\mathcal{R}_1}_{\mathsf{Inc}^*}(r_1) \approx 0.15 \qquad S^{\mathcal{R}_1}_{\mathsf{Inc}^*}(r_2) \approx 0.117$$
$$S^{\mathcal{R}_1}_{\mathsf{Inc}^*}(r_3) \approx 0.05 \qquad S^{\mathcal{R}_1}_{\mathsf{Inc}^*}(r_4) \approx 0.183 \quad .$$

The distribution of the probabilistic Shapley culpability values indicates that the conditional $r_4 = (b)[0.2]$ is more responsible for the inconsistency in $\mathcal{R}_1$ and $r_3 = (a)[0.5]$ is less responsible. This can be justified as both rules r_1 and r_2 describe an influence of a on b and—assuming that the knowledge base describes causal rather than diagnostic information—thus state that a is more entrenched or more basic than b. Thus, rule r_4 that gives a probability of b not conditioned on anything else, is most dangerous for consistency. As a more formal justification for this distribution take into account that the set of minimal inconsistent subsets of $\mathcal{R}$ is given via $\mathsf{MI}(\mathcal{R}_1) = \{\{r_1, r_2, r_4\}, \{r_1, r_3, r_4\}, \{r_2, r_3, r_4\}\}$. As one can see, r_4 is the only conditional that is in every minimal inconsistent subset and therefore seems to be most problematic.

Example 4.4. Consider again the knowledge base $\mathcal{R} = \{r_1, r_2, r_3\}$ from Example 3.10 on page 63 given via

$$r_1 = (b\,|\,a)[1] \qquad r_2 = (a)[1] \qquad r_3 = (b)[0]$$

with $\mathsf{Inc}^*(\mathcal{R}) = 1$. There, we have

$$S^{\mathcal{R}}_{\mathsf{Inc}^*}(r_1) \approx 0.33 \qquad S^{\mathcal{R}}_{\mathsf{Inc}^*}(r_2) \approx 0.33 \qquad S^{\mathcal{R}}_{\mathsf{Inc}^*}(r_3) \approx 0.33 \quad .$$

Here it is clear, that all three probabilistic conditionals are equally responsible for the inconsistency in $\mathcal{R}$ as $\mathsf{MI}(\mathcal{R}) = \{\{r_1, r_2, r_3\}\}$.

Example 4.5. Consider again the knowledge base $\mathcal{R}_3 = \{r_1, r_2, r_3, r_4, r_5\}$ from Example 3.12 on page 64 given via

$$r_1 = (a \mid c)[0.7] \qquad r_2 = (b \mid c)[0.8] \qquad r_3 = (a)[0.2]$$
$$r_4 = (b)[0.3] \qquad r_5 = (c)[0.5]$$

with $\mathsf{Inc}^*(\mathcal{R}_3) = 0.25$. There we have

$$S^{\mathcal{R}_3}_{\mathsf{Inc}^*}(r_1) \approx 0.062 \qquad S^{\mathcal{R}_3}_{\mathsf{Inc}^*}(r_2) \approx 0.045 \qquad S^{\mathcal{R}_3}_{\mathsf{Inc}^*}(r_3) \approx 0.062$$
$$S^{\mathcal{R}_3}_{\mathsf{Inc}^*}(r_4) \approx 0.045 \qquad S^{\mathcal{R}_3}_{\mathsf{Inc}^*}(r_5) \approx 0.036 \quad .$$

The probabilistic Shapley culpability measure satisfies the same properties as the Shapley value.

Proposition 4.1. *If* Inc *is an inconsistency measure that satisfies (Consistency) and (Super-Additivity), then the probabilistic Shapley culpability measure* S_{Inc} *satisfies (*Inc*-Distribution), (*Inc*-Symmetry), and (Minimality).*

Proof. It suffices to show that Inc satisfies the preconditions on functions in coalition games, cf. Definition 4.3 on page 93. Then $S^{\mathcal{R}}_{\mathsf{Inc}}$ is a valid definition of a Shapley value according to Definition 4.4 on page 94 and thus satisfies the restated properties due to (Shapley, 1953). But this is clear, as Inc satisfies $\mathsf{Inc}(\emptyset) = 0$ due to (Consistency) and $\mathsf{Inc}(S \cup T) \geq \mathsf{Inc}(S) + \mathsf{Inc}(T)$ with $S \cap T = \emptyset$ due to (Super-Additivity). ☐

The probabilistic Shapley culpability measure also satisfies the property (Additivity) as pointed out in Section 4.1.1. However, we do not consider a restated version of this property for culpability measures as combination of inconsistency measures seems to be of no practical relevance, see (Hunter and Konieczny, 2008) for a discussion.

From Proposition 4.1 and Propositions 3.11 on page 54, 3.13 on page 57, and Theorem 3.1 on page 65 we obtain the following statement.

Corollary 4.1. *Let* $\mathsf{Inc} \in \{\mathsf{Inc}^{\mathsf{MI}}, \mathsf{Inc}^{\mathsf{MI}}_C, \mathsf{Inc}^*\}$. *The probabilistic Shapley inconsistency measure* S_{Inc} *satisfies (*Inc*-Distribution), (*Inc*-Symmetry), and (Minimality).*

Satisfaction of (Decomposability) is only achieved for a special type of probabilistic Shapley culpability measure, cf. (Hunter and Konieczny, 2008).

Proposition 4.2. *Let* $\mathcal{R}$ *be a knowledge base and* Inc *be an inconsistency measure with* $\mathsf{Inc}(\mathcal{M}) = 1$ *for each* $\mathcal{M} \in \mathsf{MI}(\mathcal{R})$. *A culpability measure* $C^{\mathcal{R}}$ *satisfies (*Inc*-Distribution), (*Inc*-Symmetry), (Minimality), and (Decomposability) if and only if* $\mathsf{Inc} = \mathsf{Inc}^{\mathsf{MI}}$ *and* $C^{\mathcal{R}} = S^{\mathcal{R}}_{\mathsf{Inc}^{\mathsf{MI}}}$.

The above proposition states that $C^{\mathcal{R}} = S^{\mathcal{R}}_{\mathsf{Inc}^{\mathsf{MI}}}$ is uniquely identified by the properties (Inc-Distribution), (Inc-Symmetry), (Minimality), (Decomposability), and the additional demand that it holds that $\mathsf{Inc}(\mathcal{M}) = 1$ for a minimal inconsistent set $\mathcal{M}$. The proof of Proposition 4.2 is analogous to the proof of Proposition 5 in (Hunter and Konieczny, 2008).

As a further side note, the probabilistic Shapley culpability measure $S^{\mathcal{R}}_{\mathsf{Inc}^{\mathsf{MI}}}$ has a nice characterization as the following theorem shows which extends a result from (Hunter and Konieczny, 2008).

Theorem 4.1. *Let $\mathcal{R}$ be a knowledge base and $r \in \mathcal{R}$. Then*

$$S^{\mathcal{R}}_{\mathsf{Inc}^{\mathsf{MI}}}(r) = \sum_{\mathcal{M} \in \mathsf{MI}(\mathcal{R}), r \in \mathcal{M}} \frac{1}{|\mathcal{M}|}$$

The proof of the above theorem is straightforward and follows the proof of Proposition 4 in (Hunter and Konieczny, 2008). Note that the above characterization of the probabilistic Shapley culpability measure resembles the notion of the MI^{C} inconsistency measure, cf. Section 3.3.

The probabilistic Shapley culpability measure $S^{\mathcal{R}}_{\mathsf{Inc}}$ does not satisfy (Continuity) for arbitrary inconsistency measure Inc, as the following example shows.

Example 4.6. Consider the knowledge base $\mathcal{R}_{x,y}$ with $\langle \mathcal{R}_{x,y} \rangle = (r_1, r_2)$ and

$$r_1 =_{def} (a)[x] \qquad\qquad r_2 =_{def} (a)[y] \quad .$$

For $x = y$ it is $\mathsf{Inc}_d(\mathcal{R}_{x,x}) = 0$ and therefore $S^{\mathcal{R}_{x,x}}_{\mathsf{Inc}_d}(r_1) = S^{\mathcal{R}_{x,x}}_{\mathsf{Inc}_d}(r_2) = 0$. However, for every x, y with $x \neq y$ it is $\mathsf{Inc}_d(\mathcal{R}_{x,y}) = 1$ and $S^{\mathcal{R}_{x,y}}_{\mathsf{Inc}_d}(r_1) = S^{\mathcal{R}_{x,y}}_{\mathsf{Inc}_d}(r_2) = 0.5$. In particular, for fixed y and $x \neq y$ it holds that

$$\lim_{x \to y} \theta^{\dagger}_{S^{\mathcal{R}_{1,1}}_{\mathsf{Inc}_d}, \mathcal{R}_{1,1}}(x, y) = (0.5, 0.5)$$

but $\theta^{\dagger}_{S^{\mathcal{R}_{1,1}}_{\mathsf{Inc}_d}, \mathcal{R}_{1,1}}(y, y) = (0, 0)$.

However, for continuous inconsistency measures the probabilistic Shapley culpability measure is continuous as well.

Proposition 4.3. *If* Inc *satisfies (Continuity) then $S^{\mathcal{R}}_{\mathsf{Inc}}$ satisfies (Continuity) for every knowledge base $\mathcal{R}$.*

Proof. We only have to show that $\theta^{\dagger}_{S^{\mathcal{R}}_{\mathsf{Inc}}, \mathcal{R}}$ is continuous in each dimension. But this is clear due to the continuity of Inc and the continuity of the operations summation and multiplication in Definition 4.5. $\square$

4.1.2 *Mean Distance Culpability Measure*

When considering again the MinDev inconsistency measure and the optimization problem of minimizing f_{Inc^*} with respect to $\mathsf{DevCons}(\mathcal{R})$ from Section 3.4, one might get the intuition that a solution of this optimization problem already defines a culpability measure. In Example 3.11 from page 64, a solution of $\mathsf{DevCons}(\mathcal{R})$ is $\eta_1^* = \eta_2^* = \eta_3^* = 0$ and $\eta_4^* = 0.5$, therefore assigning each of the conditionals r_1, r_2, r_3 a "culpability" of zero and r_4 the value 0.5. This is true for this solution, but remember that—in general—there is no unique $(x_1, \ldots, x_n)$ that minimizes f_{Inc^*} with respect to $\mathsf{DevCons}(\mathcal{R})$. Other (minimal) solutions may assign a value of zero to r_4 thus marking r_4 as harmless. But still, distance minimization is a rational choice for measuring inconsistency and can be made fit to be applied to the problem of culpability measurement when considering not a single but the set $\mathcal{MD}(\mathcal{R})$ of all solutions to minimizing f_{Inc^*} with respect to $\mathsf{DevCons}(\mathcal{R})$. We abbreviate

$$\mathcal{MD}_i(\mathcal{R}) =_{def} \{x_i \mid (x_1, \ldots, x_n) \in \mathcal{MD}(\mathcal{R})\}$$

for $i = 1, \ldots,$ and define the *mean distance culpability measure* as follows.

Definition 4.6 (Mean distance culpability measure)**.** Let $\mathcal{R}$ be a knowledge base with $\langle \mathcal{R} \rangle = ((\psi_1 \mid \phi_1)[d_1], \ldots, (\psi_n \mid \phi_n)[d_n])$. Then the *mean distance culpability measure* $A^{\mathcal{R}}$ is defined as

$$A^{\mathcal{R}}((\psi_i \mid \phi_i)[d_i]) =_{def} \left| d_i - \frac{\sup \mathcal{MD}_i(\mathcal{R}) + \inf \mathcal{MD}_i(\mathcal{R})}{2} \right|$$

for $i = 1, \ldots, n$.

We further define the *sign of culpability* $\mathsf{SignCulp}^{\mathcal{R}}((\psi_i \mid \phi_i)[d_i]) \in \{-1, 0, 1\}$ of $(\psi_i \mid \phi_i)[d_i]$ in $\mathcal{R}$ via

$$\mathsf{SignCulp}^{\mathcal{R}}((\psi_i \mid \phi_i)[d_i]) =_{def} \mathsf{sgn}\left(d_i - \frac{\sup \mathcal{MD}_i(\mathcal{R}) + \inf \mathcal{MD}_i(\mathcal{R})}{2}\right)$$

where sgn is the *signum function* and $i = 1, \ldots, n$.

Example 4.7. We continue Example 3.10 from page 63 and consider the knowledge base $\mathcal{R}$ with $\langle \mathcal{R} \rangle = (r_1, r_2, r_3)$ and

$$r_1 = (b \mid a)[1] \qquad r_2 = (a)[1] \qquad r_3 = (b)[0] \quad .$$

It holds that $\mathsf{Inc}^*(\mathcal{R}) = 1$ and there are three solutions satisfying the corresponding optimization problem that maximizes or minimizes, respectively, one η-value at a time:

$$\eta_1^1 = -1 \qquad \eta_2^1 = 0 \qquad \eta_3^1 = 0$$

$$\eta_1^2 = 0 \qquad \eta_2^2 = -1 \qquad \eta_3^2 = 0$$
$$\eta_1^3 = 0 \qquad \eta_2^3 = 0 \qquad \eta_3^3 = 1 \quad .$$

Therefore we have $A^{\mathcal{R}}(r_1) = A^{\mathcal{R}}(r_2) = A^{\mathcal{R}}(r_3) = 1/2$ and

$$\mathsf{SignCulp}^{\mathcal{R}}(r_1) = -1 \quad \mathsf{SignCulp}^{\mathcal{R}}(r_2) = -1 \quad \mathsf{SignCulp}^{\mathcal{R}}(r_3) = 1$$

We now turn to the formal properties of $A^{\mathcal{R}}$.

Proposition 4.4. *$A^{\mathcal{R}}$ satisfies (Minimality) and (Continuity).*

Proof. Let $\mathcal{R}$ be a knowledge base with $\langle\mathcal{R}\rangle = (r_1, \ldots, r_n)$ and $r_i = (\psi_i \,|\, \phi_i)[d_i]$ for $i = 1, \ldots, n$.

(Minimality) If $r_i \in \mathcal{R}$ is free in $\mathcal{R}$ then for $\mathcal{R}' = \Lambda_{\mathcal{R}}(d_1 + \eta_1, \ldots, d_n + \eta_n)$ with $\mathsf{Inc}^*(\mathcal{R}) = x = |\eta_1| + \ldots + |\eta_n|$ it follows $\eta_i = 0$. Otherwise the knowledge base $\mathcal{R}' \setminus \{(\psi_i \,|\, \phi_i)[d_i + \eta_i]\}$ would be consistent as well and therefore $\mathsf{Inc}^*(\mathcal{R} \setminus \{r_i\}) = x - \eta_j < x$ which contradicts that r_i is free in $\mathcal{R}$. It follows that $\inf \mathcal{MD}_i(\mathcal{R}) = \sup \mathcal{MD}_i(\mathcal{R}) = 0$ and $A^{\mathcal{R}}(r_i) = 0$.

(Continuity) As $\theta_{\mathsf{Inc},\mathcal{R}}$ is continuous so are $\sup \mathcal{MD}_i(\mathcal{R})$ and $\inf \mathcal{MD}_i(\mathcal{R})$ for every $i = 1, \ldots, n$ and consequently $\theta^{\dagger}_{A^{\mathcal{R}},\mathcal{R}}$ as well. □

As Example 4.7 has shown $A^{\mathcal{R}}$ does not satisfy (Inc^*-Distribution) in general. This also holds for every other inconsistency measure Inc discussed before and the property (Inc-Distribution). The culpability measure $A^{\mathcal{R}}$ also fails to satisfy (Decomposition) as the following example shows.

Example 4.8. Consider the knowledge bases $\mathcal{R}_1 =_{def} \{r_1, r_2\}$ and $\mathcal{R}_2 =_{def} \{r_1, r_3\}$ given via

$$r_1 =_{def} (a)[0.7] \qquad r_2 =_{def} (a)[0.3] \qquad r_3 =_{def} (\neg a)[0.7] \quad .$$

It follows that for $\mathcal{R} =_{def} \mathcal{R}_1 \cup \mathcal{R}_2$ it holds that

$$\mathsf{MI}(\mathcal{R}) = \{\{r_1, r_2\}, \{r_1, r_3\}\} = \mathsf{MI}(\mathcal{R}_1) \cup \mathsf{MI}(\mathcal{R}_2)$$

and also $\mathsf{MI}(\mathcal{R}_1) \cap \mathsf{MI}(\mathcal{R}_2) = \emptyset$. However, for $r_1 \in \mathcal{R}_1 \cap \mathcal{R}_2$ it holds that

$$A^{\mathcal{R}_1}(r_1) = 0.2 \quad \text{and} \quad A^{\mathcal{R}_2}(r_1) = 0.2$$

but $A^{\mathcal{R}}(r_1) = 0 \neq A^{\mathcal{R}_1}(r_1) + A^{\mathcal{R}_2}(r_1)$.

Furthermore, $A^{\mathcal{R}}$ also fails to satisfy (Inc^*-Symmetry). Consider the following counterexample.

Example 4.9. Let $\mathcal{R} =_{def} \{r_1, r_2, r_3\}$ given via

$$r_1 = (b \mid a)[0.9] \qquad r_2 = (a)[0.9] \qquad r_3 = (b)[0.1] \quad .$$

Note that $\mathsf{Inc}^*(\mathcal{R}) = 0.81$ and modifying r_3 to 0.81 is the *only* possible minimal modification. However, as every subset of $\mathcal{R}$ is consistent, i. e. for every two $\mathcal{R}', \mathcal{R}'' \subsetneq \mathcal{R}$ it holds that $\mathsf{Inc}^*(\mathcal{R}') = \mathsf{Inc}^*(\mathcal{R}'') = 0$. Therefore, satisfaction of (Inc^*-Symmetry) demands that $A^{\mathcal{R}}(r_2) = A^{\mathcal{R}}(r_3)$. As $A^{\mathcal{R}}(r_2) = 0$ and $A^{\mathcal{R}}(r_3) = 0.71$ the culpability measure $A^{\mathcal{R}}$ fails to satisfy (Inc^*-Symmetry). This also holds for every other inconsistency measure Inc discussed before and the property (Inc-Symmetry).

The above discussion raises the question whether the mean distance culpability measure is a meaningful measure for probabilistic conditional logic. With the probabilistic Shapley culpability measure we already have a measure that satisfies all our desired properties and we have shown that $A^{\mathcal{R}}$ violates three of those properties. However, due to its definition based on the MinDev inconsistency measure, $A^{\mathcal{R}}$ allows to make some distinct observations. Consider the following example.

Example 4.10. Consider the knowledge base $\mathcal{R} =_{def} \{r_1, r_2, r_3\}$ given via

$$r_1 =_{def} (a)[0.3] \qquad r_2 =_{def} (b)[0.4] \qquad r_3 =_{def} (a \wedge b)[0.6] \quad .$$

Note that for a consistent modification of $\mathcal{R}$ the probabilities of both r_1 and r_2 must be at least as large as the probability of r_3. It holds that $\mathsf{Inc}^*(\mathcal{R}) = 0.3$ and assigning probability 0.4 to each conditional yields a consistent knowledge base. In particular, the probability of r_2 has not been changed by this assignment. Assume there is also a consistent knowledge base that changes the probability of r_2 to $0.4 + \alpha$ for some α. It follows that the probability of r_3 has to be modified by at least $0.6 - 0.4 + \alpha$. Now, only r_1 has a too small probability to achieve consistency and the only possible solution is either to change the probability of r_1 to at least $0.4 + \alpha$ as well, or to change the probability of r_3 down to 0.3 or less, or something in between. In any case, this yields a total modification of $0.3 + \alpha$ which is not minimal if $\alpha \neq 0$. It follows that every consistent knowledge base with minimal distance to $\mathcal{R}$ does not change the probability of r_2 and it follows

$$A^{\mathcal{R}}(r_1) = 0.05 \qquad A^{\mathcal{R}}(r_2) = 0 \qquad A^{\mathcal{R}}(r_3) = 0.25$$

Although r_2 is not free in $\mathcal{R}$ it is somewhat "neutral" in the light of the inconsistency produced by r_1 and r_3. Considering e. g. the culpability measure $S^{\mathcal{R}}_{\mathsf{Inc}^{\mathsf{MI}}_C}$ it follows that

$$S^{\mathcal{R}}_{\mathsf{Inc}^{\mathsf{MI}}_C}(r_1) = 0.25 \qquad S^{\mathcal{R}}_{\mathsf{Inc}^{\mathsf{MI}}_C}(r_2) = 0.25 \qquad S^{\mathcal{R}}_{\mathsf{Inc}^{\mathsf{MI}}_C}(r_3) = 0.5$$

thus rendering r_2 as culpable as r_1 for producing the inconsistency. This is due to the fact that $\mathsf{Inc}^{\mathsf{MI}}$ and $S^{\mathcal{R}}_{\mathsf{Inc}^{\mathsf{MI}}_C}$ only considers the membership of conditionals to minimal inconsistent subsets and not their role in these sets. However, consider the culpability measure $S^{\mathcal{R}}_{\mathsf{Inc}^*}$ with the culpability values

$$S^{\mathcal{R}}_{\mathsf{Inc}^*}(r_1) \approx 0.083 \qquad S^{\mathcal{R}}_{\mathsf{Inc}^*}(r_2) \approx 0.033 \qquad S^{\mathcal{R}}_{\mathsf{Inc}^*}(r_3) \approx 0.183 \quad .$$

As one can see, $S^{\mathcal{R}}_{\mathsf{Inc}^*}$ also takes those roles into account and assigns r_2 a smaller culpability value than r_1.

The assignment of zero culpability of $A^{\mathcal{R}}$ to r_2 is very drastic, considering that r_2 is indeed part of a minimal inconsistent subset. This behavior of $A^{\mathcal{R}}$ is even more apparent in the following example.

Example 4.11. Consider the knowledge base $\mathcal{R} =_{def} \{r_1, r_2, r_3, r_4\}$ given via

$$r_1 =_{def} (b \mid a)[0.8] \qquad r_2 =_{def} (b \mid \neg a)[0.6]$$
$$r_3 =_{def} (a)[0.5] \qquad r_4 =_{def} (b)[0.2]$$

For $\mathcal{R}$ it holds that

$$A^{\mathcal{R}}(r_1) = A^{\mathcal{R}}(r_2) = A^{\mathcal{R}}(r_3) = 0 \qquad A^{\mathcal{R}}(r_4) = 0.5$$

and

$$S^{\mathcal{R}}_{\mathsf{Inc}^{\mathsf{MI}}_C}(r_1) = S^{\mathcal{R}}_{\mathsf{Inc}^{\mathsf{MI}}_C}(r_2) = S^{\mathcal{R}}_{\mathsf{Inc}^{\mathsf{MI}}_C}(r_3) \approx 0.222 \qquad S^{\mathcal{R}}_{\mathsf{Inc}^{\mathsf{MI}}_C}(r_4) \approx 0.333 \quad .$$

Note that the large culpability value $A^{\mathcal{R}}(r_4)$ derives from the fact that r_4 is contained in all minimal inconsistent subsets of $\mathcal{R}$ (it holds that $\mathsf{MI}(\mathcal{R}) = \{\{r_1, r_2, r_4\}, \{r_1, r_3, r_4\}, \{r_2, r_3, r_4\}\}$).

The above examples show that $S^{\mathcal{R}}$ is more "balanced" when assigning culpability values of members of minimal inconsistent subsets. In fact, due to the structure of $S^{\mathcal{R}}$ it follows that every probabilistic conditional that is a member of at least one minimal inconsistent subset gets a culpability value greater than zero. As the above examples showed, this is not the case for $A^{\mathcal{R}}$.

4.2 Principled Consistency Restoring

While inconsistency and culpability measures help to analyze and understand the inconsistencies in a knowledge base there is still the open problem of *repairing* the knowledge base to restore consistency. This is a crucial task as repairing an inconsistent knowledge base requires modifying it and any modification alters the beliefs that were initially intended to be represented.

With the use of inconsistency and culpability measures a knowledge engineer has already a powerful set of tools to guide him to manually repair the knowledge base. However, in the following we consider *automatic repair* of inconsistent knowledge bases that adheres to some rationality postulates. We develop several approaches that use inconsistency and culpability measures to restore consistency. Note that these approaches work on the syntactic level only and semantical aspects of the represented beliefs are not adhered for. In practice it is up to the knowledge engineer to verify the results of an automatic repair of an inconsistent knowledge base.

Before discussing rationality postulates for automatic repair we first discuss a naive approach inspired by the solutions to the optimization problem $\mathsf{DevCons}(\mathcal{R})$. By definition, these already describe minimal adjustments to be made in order to restore consistency. By considering the culpability measure $C_{\mathcal{R}}$ for the conditionals in $\mathcal{R}$ one could select the most appropriate solution and modify the knowledge base accordingly. But as pointed out in the previous section, the solutions $\mathcal{MD}(\mathcal{R})$ given by the optimization problem $\mathsf{DevCons}(\mathcal{R})$ are not always the "rational" ones to choose, despite the fact that they describe minimal adjustments. Consider again the knowledge base $\mathcal{R}$ from Example 3.10 on page 63 with $\langle\mathcal{R}\rangle = (r_1, r_2, r_3)$ and

$$r_1 = (b \,|\, a)[1] \qquad\qquad r_2 = (a)[1] \qquad\qquad r_3 = (b)[0] \quad .$$

Some minimal solutions of the optimization problem $\mathsf{DevCons}(\mathcal{R})$ are $\eta_1^* = -1$ or $\eta_2^* = -1$ or $\eta_3^* = 1$ (all other values being zero) and thus there are three "extreme" adjustments $\mathcal{R}^1, \mathcal{R}^2, \mathcal{R}^3$ of $\mathcal{R}$ according to these values[2]:

$$\begin{aligned}
\mathcal{R}^1 &=_{def} \{\, (b \,|\, a)[0], (a)[1], (b)[0] \,\} \\
\mathcal{R}^2 &=_{def} \{\, (b \,|\, a)[1], (a)[0], (b)[0] \,\} \\
\mathcal{R}^3 &=_{def} \{\, (b \,|\, a)[1], (a)[1], (b)[1] \,\}
\end{aligned}$$

But clearly, none of these solutions seem to be an "appropriate" adjustment of $\mathcal{R}$ due to their large discrepancies in the modeled beliefs.

We go on by discussing some rationality postulates for restoring consistency. Let undef denote a not defined function value. Let At be a propositional signature and Υ be a function

$$\Upsilon : \mathfrak{P}((\mathcal{L}(\mathsf{At}) \,|\, \mathcal{L}(\mathsf{At}))^{pr}) \to \mathfrak{P}((\mathcal{L}(\mathsf{At}) \,|\, \mathcal{L}(\mathsf{At}))^{pr}) \cup \{\mathsf{undef}\} \quad .$$

In the following Υ is called a *consistency restorer* and is intended to map a possibly inconsistent knowledge base $\mathcal{R}$ to a consistent knowledge base $\Upsilon(\mathcal{R})$. For the rest of this section let $\mathcal{R}$ with

$$\langle\mathcal{R}\rangle = ((\psi_1 \,|\, \phi_1)[d_1], \ldots, (\psi_n \,|\, \phi_n)[d_n])$$

2 Note that there are infinitely many consistent and minimal solutions.

be a possibly inconsistent knowledge base. Following (Finthammer *et al.*, 2007), there are three basic approaches (and combinations thereof) for restoring consistency:

1. *removing* conditionals such that $\Upsilon(\mathcal{R}) \subseteq \mathcal{R}$,
2. *qualitative modification* of conditionals, i. e., a conditional $(\psi \mid \phi)[d]$ is modified to $(\psi' \mid \phi')[d]$, and
3. *quantitative modification* of conditionals, i. e., a conditional $(\psi \mid \phi)[d]$ is modified to $(\psi \mid \phi)[d']$.

The second approach has already been investigated in (Rödder and Xu, 2001) and in a heuristic fashion in (Ludolph, 2009). A combination of the first and third approach has been discussed (also heuristically) in (Finthammer *et al.*, 2007). Another possibility of restoring consistency—also followed in (Rödder and Xu, 2001; Finthammer *et al.*, 2007)—is by considering *bounded conditionals* and widen the probability intervals for conditionals appropriately, cf. Section 2.3. However, this last approach changes the formalism for knowledge representation which is not a reasonable thing to do when restoring consistency in probabilistic conditional logic. In this thesis we focus on the third option of quantitative modifications and proceed with a *principled* instead of heuristic approach. We give a comparison with the previously mentioned approaches at the end of this chapter.

Restoring consistency in a knowledge base $\mathcal{R}$ is similar to the problem of *information fusion* (Bloch and Hunter, 2001) or *belief merging* (Konieczny and Pino-Pérez, 1998, 2005; Grégoire and Konieczny, 2006). Belief merging considers the problem of joining a set of knowledge bases (possibly coming from different sources) into a single knowledge base that describes the original distributed knowledge in the best way. This problem is also similar to the problems discussed in *social choice theory* (Arrow, 1950; Kelly, 1988). In social choice theory the goal is to aggregate interests or votes in order to come up with a decision that is favorable in the light of the individual interests. A well-known application for social choice theory is the problem of *voting*, i. e. of constructing a voting mechanism that is, in some sense, fair. Arrow's famous impossibility theorem (Arrow, 1950) states that there is no such thing as a fair voting mechanism or "every voting mechanism is flawed". Nonetheless, in the following we borrow many properties that were originally stated for social choice theory and belief merging and apply them to our framework of restoring consistency.

Let Υ be a consistency restorer. The basic demands for such a function are summarized in the following five properties.

(Existence) For every knowledge base $\mathcal{R}$ it holds that $\Upsilon(\mathcal{R}) \neq \mathsf{undef}$.

(Uniqueness) For every knowledge base $\mathcal{R}$ the value of $\Upsilon(\mathcal{R})$ is uniquely determined.

(Structural Preservation) For every knowledge base $\mathcal{R}$ it holds that either $\Upsilon(\mathcal{R}) = \{(\psi_1 \,|\, \phi_1)[d'_1], \ldots, (\psi_n \,|\, \phi_n)[d'_n]\}$ or $\Upsilon(\mathcal{R}) = \mathsf{undef}$.

(Success) For every knowledge base $\mathcal{R}$ it holds that if $\Upsilon(\mathcal{R}) \neq \mathsf{undef}$ then $\Upsilon(\mathcal{R})$ is consistent.

(Irrelevance of Syntax) For knowledge bases $\mathcal{R}_1$ and $\mathcal{R}_2$ it holds that if $\mathcal{R}_1 \equiv^{\text{cond}} \mathcal{R}_2$ then $\Upsilon(\mathcal{R}_1) \equiv^{\text{cond}} \Upsilon(\mathcal{R}_2)$.

The property (Existence) demands that the result of restoring consistency in $\mathcal{R}$ is well-defined, i. e., for every knowledge base $\mathcal{R}$ we get a knowledge base $\Upsilon(\mathcal{R})$. The property (Uniqueness) demands that the result of restoring consistency is uniquely determined. Note that (Uniqueness) is trivially satisfied by every Υ due to its functional property. As pointed out above we focus on quantitative modifications of conditionals which is described by the property (Structural Preservation). The property (Success) describes the very basic demand of a consistency restorer to restore consistency. The property (Irrelevance of Syntax) demands that restoring consistency in cond-equivalent knowledge bases yields again cond-equivalent knowledge bases. One may argue that a similar property can be stated using the notion of kb-equivalence. But remember that kb-equivalence does not distinguish between inconsistent knowledge bases, see Proposition 2.4 on page 29. As for inconsistent $\mathcal{R}_1, \mathcal{R}_2$ it is always the case that $\mathcal{R}_1 \equiv^{\text{kb}} \mathcal{R}_2$ defining (Irrelevance of Syntax) based on kb-equivalence amounts to requiring that the outcome of restoring consistency is the same, regardless of the structure of $\mathcal{R}_1$ and $\mathcal{R}_2$, respectively. Of course, this is not desirable as e. g. restoring consistency in both $\mathcal{R}_1 =_{def} \{(a)[0.7], (a)[0.3]\}$ and $\mathcal{R}_2 =_{def} \{(b)[0.7], (b)[0.3]\}$ should not yield equivalent consistent knowledge bases.

For all following properties we assume that Υ satisfies (Structural Preservation). If also $\Upsilon(\mathcal{R}) \neq \mathsf{undef}$ for a knowledge base $\mathcal{R}$ with $\langle \mathcal{R} \rangle = ((\psi_1 \,|\, \phi_1)[d_1], \ldots, (\psi_n \,|\, \phi_n)[d_n])$ we assume

$$\langle \Upsilon(\mathcal{R}) \rangle = ((\psi_1 \,|\, \phi_1)[d_1^*], \ldots, (\psi_n \,|\, \phi_n)[d_n^*]) \quad .$$

Further, we abbreviate $r_i =_{def} (\psi_i \,|\, \phi_i)[d_i]$ and $r_i^* =_{def} (\psi_i \,|\, \phi_i)[d_i^*]$ for $i = 1, \ldots, n$.

Due to the *minimal-change* paradigm a consistent $\mathcal{R}$ has an obvious solution.

(Consistency) For consistent $\mathcal{R}$ it holds that $\Upsilon(\mathcal{R}) = \mathcal{R}$.

As discussed above, the problem of restoring consistency bears much resemblance to social choice theory. Consequently, many of the principles of social choice theory can be adapted for consistency restoring such as the following ones, cf. (Arrow, 1950; Kelly, 1988).

(Pareto-Efficiency) If $\Upsilon(\mathcal{R}) \neq \mathsf{undef}$ then for every $i = 1, \ldots, n$ and every $d \in [0,1]$ such that $|d - d_i| < |d_i^* - d_i|$ it follows that the knowledge base $\{(\psi_1 \,|\, \phi_1)[d_1^*], \ldots, (\psi_i \,|\, \phi_i)[d], \ldots, (\psi_n \,|\, \phi_n)[d_n^*]\}$ is inconsistent.

In voting theory, a candidate a is called *Pareto-efficient* (or *Pareto-optimal*) if there is no other candidate b that is at least equally preferred by all voters and there is no voter that prefers b to a. This means that all other outcomes for the voting procedure disadvantage at least one voter. Here, the property (Pareto-Efficiency) demands that the result of restoring consistency should be as close as possible to the original knowledge base and no modified conditional should deviate more as needed from its original one. In particular, every other consistent knowledge base deviates to a larger extent from the original knowledge base in at least one probabilistic conditional.

(Non-Dictatorship) For every non-tautological $(\psi_i \,|\, \phi_i)[d_i]$ there is a knowledge base $\mathcal{R}$ with $(\psi_i \,|\, \phi_i)[d_i] \in \mathcal{R}$ such that $\Upsilon(\mathcal{R}) \neq \mathsf{undef}$ and $d_i^* \neq d_i$.

The property (Non-Dictatorship) demands that no non-tautological conditional takes the role of a *dictator* and is never modified in any knowledge base. In voting theory, this property states that there is no single voter that can dictate the outcome.

(Non-Imposition) Υ is surjective.

In voting theory, the property (Non-Imposition) states that every possible outcome of a voting, e.g. the election of any candidate, should be achievable. In our setting, this means that every knowledge base can be the result of consistency restoration. Note that (Non-Imposition) cannot (and should not) be satisfied by any meaningful consistency restorer Υ as $\mathfrak{P}((\mathcal{L}(\mathsf{At}) \,|\, \mathcal{L}(\mathsf{At}))^{pr})$ also contains inconsistent knowledge bases which should not be contained in the image of Υ. Hence, every function Υ that satisfies (Success) cannot satisfy (Non-Imposition). Therefore, let $\mathfrak{P}((\mathcal{L}(\mathsf{At}) \,|\, \mathcal{L}(\mathsf{At}))^{pr})^* \subseteq \mathfrak{P}((\mathcal{L}(\mathsf{At}) \,|\, \mathcal{L}(\mathsf{At}))^{pr})$ be the set of consistent knowledge bases and consider the following restriction of (Non-Imposition).

(Rational Non-Imposition) It holds that $\mathsf{Im}\ \Upsilon = \mathfrak{P}((\mathcal{L}(\mathsf{At}) \,|\, \mathcal{L}(\mathsf{At}))^{pr})^*$.

The property (Rational Non-Imposition) means that every consistent knowledge base can be a function value.

Proposition 4.5. *If* Υ *satisfies (Existence), (Success) and (Consistency) then* Υ *satisfies (Rational Non-Imposition).*

Proof. Let Υ satisfy (Existence), (Success) and (Consistency). Due to (Existence) and (Success) it holds that $\mathsf{Im}\ \Upsilon \subseteq \mathfrak{P}((\mathcal{L}(\mathsf{At}) \,|\, \mathcal{L}(\mathsf{At}))^{pr})^*$. For every knowledge base $\mathcal{R} \in \mathfrak{P}((\mathcal{L}(\mathsf{At}) \,|\, \mathcal{L}(\mathsf{At}))^{pr})^*$ it holds that $\Upsilon(\mathcal{R}) = \mathcal{R}$ due to (Consistency) and it follows $\mathsf{Im}\ \Upsilon = \mathfrak{P}((\mathcal{L}(\mathsf{At}) \,|\, \mathcal{L}(\mathsf{At}))^{pr})^*$. $\square$

Another property discussed in social choice theory is (Independence of Irrelevant Alternatives). Basically, in the scenario of voting theory this property states that given two elections A and B with participants $1, \ldots, m$, if i prefers candidate a to candidate b in A whenever i prefers candidate a to candidate b in B for $i = 1, \ldots, m$, then candidate a is preferred to candidate b in the outcome of the election A if and only if candidate a is preferred to candidate b in the outcome of the election B. This means that preferences regarding other combinations of candidates are irrelevant for the decision regarding just a and b. There is no direct and meaningful translation of this property into the framework of restoring consistency. In the previous translations of properties from social choice theory the notion of "preference" has been translated using the notion of consistency. There, a conditional $r \in \mathcal{R}$ "prefers" a conditional r' to r'' if r and r' are consistent and r and r'' are inconsistent. A straightforward translation of (Independence of Irrelevant Alternatives) into our framework could be phrased as follows.

> Let $\mathcal{R}_1$ and $\mathcal{R}_2$ be two knowledge bases with $|\mathcal{R}_1| = |\mathcal{R}_2| = n$. If there is a bijective function $\sigma : \mathcal{R}_1 \to \mathcal{R}_2$ such that for every $r \in \mathcal{R}_1$, if $\{r, r'\}$ is consistent for some r' whenever $\{\sigma(r), r'\}$ is consistent then $r \in \Upsilon(\mathcal{R}_1)$ whenever $\sigma(r) \in \Upsilon(\mathcal{R}_2)$.

For two arbitrary knowledge bases $\mathcal{R}_1$ and $\mathcal{R}_2$ the prerequisite of the above statement is hard to fulfill except in trivial cases. Also, there is no clear motivation why to demand this property of consistency restorers. Hence, we neglect considering this property in the following.

So far we have not taken culpability measures into account when restoring consistency. A culpability measure is a powerful tool that leads to the culprits of creating inconsistencies and a rational consistency restorer should adhere to a culpability measure in order to modify conditionals with respect to their culpabilities. In the following, we discuss properties relating culpability measures and consistency restorers. The general idea is that, if a culpability measure assigns a large degree of culpability to a conditional r and a much smaller degree of culpability to a conditional r' then r should befall a larger modification than r'.

Let $C^{\mathcal{R}}$ be a culpability measure and remember that for a knowledge base $\mathcal{R}$ with $\langle \mathcal{R} \rangle = ((\psi_1 \,|\, \phi_1)[d_1], \ldots, (\psi_n \,|\, \phi_n)[d_n])$ we assume

$$\langle \Upsilon(\mathcal{R}) \rangle = ((\psi_1 \,|\, \phi_1)[d_1^*], \ldots, (\psi_n \,|\, \phi_n)[d_n^*]) \quad .$$

($C^{\mathcal{R}}$-Conformity) If $\Upsilon(\mathcal{R}) \neq \mathsf{undef}$ then

$$C^{\mathcal{R}}((\psi_i \,|\, \phi_i)[d_i]) \geq C^{\mathcal{R}}((\psi_j \,|\, \phi_j)[d_j]) \text{ implies } |d_i^* - d_i| \geq |d_j^* - d_j|$$

for every $i, j = 1, \ldots, n$.

Satisfying ($C^{\mathcal{R}}$-Conformity) means that a conditional that is less responsible for an inconsistency is not harder penalized than a conditional that is more responsible.

(Inverse $C^{\mathcal{R}}$-Conformity) If $\Upsilon(\mathcal{R}) \neq \mathsf{undef}$ then

$$|d_i^* - d_i| \geq |d_j^* - d_j| \quad \text{implies} \quad C^{\mathcal{R}}((\psi_i \,|\, \phi_i)[d_i]) \geq C^{\mathcal{R}}((\psi_j \,|\, \phi_j)[d_j])$$

for every $i, j = 1, \ldots, n$.

The property (Inverse $C^{\mathcal{R}}$-conformity) is the counterpart of the property ($C^{\mathcal{R}}$-conformity), i.e. a conditional that is harder penalized than another one is more responsible for an inconsistency with respect to the given culpability measure.

As for inconsistency measures we demand some form of continuity as well for consistency restorers. Remember that $\mathfrak{P}^{\text{ord}}(X)$ denotes the set of all vectors of elements of X, cf. page 27. Let $\mathcal{R}$ be a knowledge base. Let $\zeta_{\mathcal{R}}$ be the function $\zeta_{\mathcal{R}} : \mathfrak{P}^{\text{ord}}((\mathcal{L}(\mathsf{At}) \,|\, \mathcal{L}(\mathsf{At}))^{pr}) \to [0,1]^{|\mathcal{R}|}$ defined via

$$\zeta_{\mathcal{R}}(((\psi_1 \,|\, \phi_1)[d_1], \ldots, (\psi_n \,|\, \phi_n)[d_n])) =_{def} (d_1, \ldots, d_n) \quad .$$

for $((\psi_1 \,|\, \phi_1)[d_1], \ldots, (\psi_n \,|\, \phi_n)[d_n]) \in \mathfrak{P}^{\text{ord}}((\mathcal{L}(\mathsf{At}) \,|\, \mathcal{L}(\mathsf{At}))^{pr})$ and let $\varsigma_{\mathcal{R}}^{\Upsilon} : [0,1]^{|\mathcal{R}|} \to [0,1]^{|\mathcal{R}|}$ be defined via $\varsigma_{\mathcal{R}}^{\Upsilon}(x_1, \ldots, x_{|\mathcal{R}|}) =_{def} \zeta_{\mathcal{R}}(\langle \Upsilon(\Lambda_{\mathcal{R}}(x_1, \ldots, x_{|\mathcal{R}|})) \rangle)$ for all $x_1, \ldots, x_{|\mathcal{R}|} \in [0,1]^{|\mathcal{R}|}$. Then we demand satisfaction of the following property.

(Continuity) For every $\mathcal{R}$ the function $\varsigma_{\mathcal{R}}^{\Upsilon}$ is continuous (with respect to the standard topology on $\mathbb{R}^n$).

Table 4 gives an overview on the properties discussed above. There, only those properties are listed that are regarded as desirable for consistency restorers. Consequently, Table 4 does not list the property (Non-Imposition).

We go on by taking some first steps in developing concrete approaches for consistency restorers. In the following, we develop two families of consistency restorers, one defined using *creeping functions* and one by distance minimization.

4.3 SOLVING CONFLICTS BY PENALIZING CULPABILITIES

Let $\mathcal{R}$ be a knowledge base. Our first approach to restore consistency is a procedural approach based on a *creeping function* $\Xi_{\mathcal{R}}$ given via

$$\Xi_{\mathcal{R}} : [0,1] \to \mathfrak{P}((\mathcal{L}(\mathsf{At}) \,|\, \mathcal{L}(\mathsf{At}))^{pr}) \quad .$$

A creeping function $\Xi_{\mathcal{R}}$ is defined as a function that "creeps" from the original (inconsistent) knowledge base towards a consistent one. The creeping is controlled by a parameter $\delta \in [0,1]$ such that $\Xi_{\mathcal{R}}(0) = \mathcal{R}$ and for larger δ the inconsistency in $\Xi_{\mathcal{R}}(\delta)$ vanishes. Given such a function we look for the minimal parameter δ^* such that $\Xi_{\mathcal{R}}(\delta^*)$ is consistent. In the following, we discuss three different creeping functions, each more sophisticated than the previous one.

Property	Description
(Existence)	It holds that $\Upsilon(\mathcal{R}) \neq \mathsf{undef}$
(Uniqueness)	The value of $\Upsilon(\mathcal{R})$ is unique
(Structural Preservation)	$\Upsilon(\mathcal{R}) = \mathsf{undef}$ or $\Upsilon(\mathcal{R})$ has the same structure as $\mathcal{R}$
(Success)	If $\Upsilon(\mathcal{R}) \neq \mathsf{undef}$ then $\Upsilon(\mathcal{R})$ is consistent
(Irrelevance of Syntax)	If $\mathcal{R}_1 \equiv^{\text{cond}} \mathcal{R}_2$ then $\Upsilon(\mathcal{R}_1) \equiv^{\text{cond}} \Upsilon(\mathcal{R}_2)$
(Consistency)	For consistent $\mathcal{R}$ it holds that $\Upsilon(\mathcal{R}) = \mathcal{R}$
(Pareto-Efficiency)	The deviation from $\Upsilon(\mathcal{R})$ to $\mathcal{R}$ is minimal
(Non-Dictatorship)	For every non-tautological r there is an $\mathcal{R}$ such that r has been modified in $\Upsilon(\mathcal{R})$
(Rational Non-Imposition)	$\mathsf{Im}\ \Upsilon = \mathfrak{P}((\mathcal{L}(\mathsf{At}) \mid \mathcal{L}(\mathsf{At}))^{pr})^*$
($C^{\mathcal{R}}$-Conformity)	If $C^{\mathcal{R}}(r) \geq C^{\mathcal{R}}(r')$ then r is modified to no less extent than r' in $\Upsilon(\mathcal{R})$
(Inverse $C^{\mathcal{R}}$-Conformity)	If r is modified to no less extent than r' in $\Upsilon(\mathcal{R})$ then $C^{\mathcal{R}}(r) \geq C^{\mathcal{R}}(r')$
(Continuity)	$\varsigma^{\Upsilon}_{\mathcal{R}}$ is continuous

Table 4: Properties of consistency restorers

In order to ensure that we end up with a consistent knowledge base we need the following notation.

Definition 4.7 (Uniform conditional probability)**.** Let At be propositional signature, let $(\psi \mid \phi)$ be a conditional and P_0 the uniform probability function on $\Omega(\mathsf{At})$. Then the *uniform conditional probability* of $(\psi \mid \phi)$ is defined as $\mathsf{ucp}(\psi \mid \phi) =_{def} P_0(\psi \mid \phi)$ if $P_0(\phi) > 0$ and $\mathsf{ucp}(\psi \mid \phi) =_{def} 0$ otherwise.

Note, that for every set $\mathcal{R} = \{(\psi_1 \mid \phi_1), \ldots, (\psi_n \mid \phi_n)\}$ of qualitative conditionals the set $\{(\psi_1 \mid \phi_1)[\mathsf{ucp}(\psi_1 \mid \phi_1)], \ldots, (\psi_n \mid \phi_n)[\mathsf{ucp}(\psi_n \mid \phi_n)]\}$ is consistent, cf. Proposition 3.17 on page 60.

4.3.1 *Unbiased Creeping*

The first creeping function we consider is a very simple one that actually does not consider any culpability measure at all. We give this definition to have a starting point for the development of more sophisticated creeping functions in the following.

Definition 4.8 (Unbiased creeping function)**.** The *unbiased creeping function* $\Xi^U_\mathcal{R}$ of $\mathcal{R}$ with $\langle\mathcal{R}\rangle = ((\psi_1 \,|\, \phi_1)[d_1], \ldots, (\psi_n \,|\, \phi_n)[d_n])$ is the function $\Xi^U_\mathcal{R}$: $[0,1] \to \mathfrak{P}((\mathcal{L}(\mathsf{At}) \,|\, \mathcal{L}(\mathsf{At}))^{pr})$ defined as

$$\Xi^U_\mathcal{R}(\delta) =_{def} \Lambda_\mathcal{R}((1-\delta)d_1 + \delta\mathsf{ucp}(\psi_1 \,|\, \phi_1), \ldots, (1-\delta)d_n + \delta\mathsf{ucp}(\psi_n \,|\, \phi_n))$$

for $\delta \in [0,1]$.

Note, that $\Xi^U_\mathcal{R}(0) = \mathcal{R}$ and that $\Xi^U_\mathcal{R}(1)$ is satisfiable by the uniform probability function. By increasing the parameter δ from zero to one we perform a continues "creep" from the (inconsistent) starting point to the necessarily consistent goal $\Xi^U_\mathcal{R}(1)$. Therefore, at some point along the way the inconsistency in the knowledge base has to vanish (as the creeping is continuous), see remark below. Note also, that "inconsistency vanishing" is continuous as we have shown that there exists a continuous inconsistency measure (Inc^*).

Definition 4.9 (Unbiased creeping consistency restorer)**.** Let Υ^U be the function $\Upsilon^U : \mathfrak{P}((\mathcal{L}(\mathsf{At}) \,|\, \mathcal{L}(\mathsf{At}))^{pr}) \to \mathfrak{P}((\mathcal{L}(\mathsf{At}) \,|\, \mathcal{L}(\mathsf{At}))^{pr})$ defined as

$$\Upsilon^U(\mathcal{R}) =_{def} \Xi^U_\mathcal{R}(\delta^*)$$

with $\delta^* =_{def} \min\{\delta \in [0,1] \mid \Xi^U_\mathcal{R}(\delta) \text{ is consistent}\}$.

Remark 4.1. We show now that the set

$$\begin{aligned} C \;=\; \{ & ((1-\delta)d_1 + \delta\mathsf{ucp}(\psi_1 \,|\, \phi_1), \ldots, (1-\delta)d_n + \delta\mathsf{ucp}(\psi_n \,|\, \phi_n)) \in [0,1]^n \mid \\ & \delta \in [0,1] \text{ and} \\ & \Lambda_\mathcal{R}((1-\delta)d_1 + \delta\mathsf{ucp}(\psi_1 \,|\, \phi_1), \ldots, (1-\delta)d_n + \delta\mathsf{ucp}(\psi_n \,|\, \phi_n)) \\ & \text{ is consistent}\} \end{aligned}$$

is closed. Note that $C \subseteq \rho(\mathcal{P}_\mathcal{R})$ with $\rho(\mathcal{P}_\mathcal{R})$ from the proof of Proposition 3.18 on page 61. Assume that

$$(d_1, \ldots, d_n) \neq (\mathsf{ucp}(\psi_1 \,|\, \phi_1), \ldots, \mathsf{ucp}(\psi_n \,|\, \phi_n)) \quad . \tag{4.1}$$

Otherwise $\mathcal{R} = \Xi^U_\mathcal{R}(0)$ is already consistent. Consider a sequence $(x^i_1, \ldots, x^i_n) \in C$ for $i \in \mathbb{N}$ such that

$$\lim_{i\to\infty}(x^i_1, \ldots, x^i_n) = (y_1, \ldots, y_n) \quad .$$

For each $(x^i_1, \ldots, x^i_n)$ let $\delta_i \in [0,1]$ be such that

$$(x^i_1, \ldots, x^i_n) = ((1-\delta_i)d_1 + \delta_i\mathsf{ucp}(\psi_1 \,|\, \phi_1), \ldots, (1-\delta_i)d_n + \delta_i\mathsf{ucp}(\psi_n \,|\, \phi_n)) \quad .$$

Note that due to (4.1) each δ_i is uniquely determined. As the mapping $\delta_i \mapsto (x^i_1, \ldots, x^i_n)$ is continuous it follows that $\lim_{i\to\infty}\delta_i = \delta^*$ and as $[0,1]$

is closed it follows that $\delta^* \in [0,1]$. As $\rho(\mathcal{P}_\mathcal{R})$ is closed it follows that $(y_1,\ldots,y_n) \in \rho(\mathcal{P}_\mathcal{R})$ because $(x_1^i,\ldots,x_n^i) \in C \subseteq \rho(\mathcal{P}_\mathcal{R})$ for each $i \in \mathbb{N}$. Therefore, $\Lambda_\mathcal{R}(y_1,\ldots,y_n) = \Xi^U_\mathcal{R}(\delta^*)$ is consistent and $(y_1,\ldots,y_n) \in C$. As C is closed and there is a unique and continuous mapping between each $(x_1,\ldots,x_n) \in C$ and $\delta \in [0,1]$, assuming that (4.1) holds, it follows that the set $\{\delta \in [0,1] \mid \Xi^U_\mathcal{R}(\delta) \text{ is consistent}\}$ is closed as well. As the minimum of a closed set of reals is well-defined so is the unbiased creeping consistency restorer.

Before investigating the properties of Υ^U we first have a look at some examples.

Example 4.12. We continue Example 3.10 from page 63 and consider the knowledge base $\mathcal{R} = \{r_1, r_2, r_3\}$ given via

$$r_1 = (a \,|\, b)[1] \qquad r_2 = (b)[1] \qquad r_3 = (a)[0] \quad .$$

Note that $\mathsf{ucp}(a) = 0.5$, $\mathsf{ucp}(b) = 0.5$ and $\mathsf{ucp}(a \,|\, b) = 0.5$. The smallest value of δ such that $\Xi^U_\mathcal{R}(\delta)$ is consistent is $\delta^* \approx 0.763898$. The resulting knowledge base is

$$\begin{aligned}\Upsilon^U(\mathcal{R}) = \{ \quad & (a \,|\, b)[0.61805],\\ & (b)[0.61805],\\ & (a)[0.38195] \quad \} \quad .\end{aligned}$$

Example 4.13. Consider the knowledge base $\mathcal{R}'_2 =_{def} \{r_1, r_2, r_3, r_4\}$ that extends the knowledge base from Example 4.12 and is given via

$$\begin{aligned} r_1 &=_{def} (a \,|\, b)[1] & r_2 &=_{def} (b)[1]\\ r_3 &=_{def} (a)[1] & r_4 &=_{def} (c)[0.3] \quad . \end{aligned}$$

Note that r_4 is a free conditional in $\mathcal{R}'_2$ and therefore does not participate in the inconsistency in $\mathcal{R}'_2$, i. e., it holds that e. g. $S^{\mathcal{R}'_2}_{\mathsf{Inc}^*}(r_4) = A^{\mathcal{R}'_2}(r_4) = 0$. Still, the probability of r_4 is modified when applying unbiased creeping:

$$\begin{aligned}\Upsilon^U(\mathcal{R}'_2) = \quad \{ \quad & (a \,|\, b)[0.61805],\\ & (b)[0.61805],\\ & (a)[0.38195],\\ & (c)[0.45278] \quad \} \quad .\end{aligned}$$

Example 4.14. Consider the knowledge base $\mathcal{R}''_2 =_{def} \{r_1, r_2, r_3\}$ that is a variant of the knowledge base from Example 4.12 and given via

$$r_1 =_{def} (a \,|\, b)[1] \qquad r_2 =_{def} (b)[1] \qquad r_3 =_{def} (a)[0.75] \quad .$$

Applying $\Upsilon^{\mathcal{U}}$ to $\mathcal{R}_2''$ yields $\Upsilon^{\mathcal{U}}(\mathcal{R}_2'') = \{r_1', r_2', r_3'\}$ given via

$$r_1' = (a \mid b)[0.80905] \qquad r_2' = (b)[0.80905] \qquad r_3' \quad = (a)[0.65453] \quad .$$

Note that although conditional r_3 in $\mathcal{R}_2''$ has a probability that is too low to be consistent with the other two conditionals, the probability of r_3 gets decreased in $\Upsilon^{\mathcal{U}}(\mathcal{R}_2'')$ due to the uniform probability of 0.5.

We now turn to the formal properties of $\Upsilon^{\mathcal{U}}$.

Theorem 4.2. $\Upsilon^{\mathcal{U}}$ *satisfies (Existence), (Uniqueness), (Structural Preservation), (Success), (Irrelevance of Syntax), (Consistency), (Rational Non-Imposition), and (Continuity).*

The proof of Theorem 4.2 can be found in Appendix A on page 237. In general, the function $\Upsilon^{\mathcal{U}}$ does not satisfy (Non-Dictatorship). Consider the non-tautological conditional $r = (a)[0.5]$ and any knowledge base $\mathcal{R}$ with $r \in \mathcal{R}$. As $\mathsf{ucp}(a) = 0.5$ it follows $(1 - \delta)0.5 + \delta 0.5 = 0.5$ for any $\delta \in [0.1]$ and therefore $r \in \Upsilon^{\mathcal{U}}(\mathcal{R})$. As a matter of fact, $\Upsilon^{\mathcal{U}}$ is only dictatorial for conditionals whose probability is exactly the uniform conditional probability of the conditional.

Note that, in general, $\Upsilon^{\mathcal{U}}$ is neither $A^{\mathcal{R}}$- nor $S^{\mathcal{R}}$-conform (and neither inverse $A^{\mathcal{R}}$- nor $S^{\mathcal{R}}$-conform) and not Pareto-efficient, see counter examples above.

4.3.2 *Penalized Creeping*

The unbiased creeping function satisfies many of the desired properties but treats all conditionals in the knowledge base the same. In order to incorporate the role the individual conditionals play in creating inconsistencies, we consider now the culpability measure $A^{\mathcal{R}}$ to guide the creeping function. Let $\alpha = (\alpha_1, \ldots, \alpha_n)$ with

$$\alpha_i =_{def} \mathsf{SignCulp}^{\mathcal{R}}(r_i) A^{\mathcal{R}}(r_i)$$

be the *culpability vector* of $\mathcal{R}$. Note that instead of $A^{\mathcal{R}}$ we could also use another culpability measure but we stick to $A^{\mathcal{R}}$ for matters of presentation. Furthermore, let

$$\hat{\alpha} = (\hat{\alpha}_1, \ldots, \hat{\alpha}_n) =_{def} \left(\frac{\alpha_1}{\alpha_{\min}}, \ldots, \frac{\alpha_n}{\alpha_{\min}} \right)$$

with $\alpha_{\min} =_{def} \min\{|\alpha_i| \mid \alpha_i \neq 0, i = 1, \ldots, n\}$ be the *normalized culpability vector* (if $\alpha_1 = \ldots = \alpha_n = 0$ the value $\alpha_{\min}$ is defined to be 1).

Example 4.15. We continue Example 4.7 from page 98. Here we have $\alpha = (-1/2, -1/2, 1/2)$ and $\hat{\alpha} = (-1, -1, 1)$.

In the following the vector $\hat{\alpha}$ is used as a *weighted search direction* for finding the next consistent knowledge base. In order to stay in the space of probabilistic knowledge bases let $u : \mathbb{R} \to [0,1]$ be the function

$$u(x) =_{def} \begin{cases} x & x \in [0,1] \\ 0 & x < 0 \\ 1 & x > 1 \end{cases}$$

In the following definition, the function u is used to ensure that the values of conditionals describe probabilities.

Definition 4.10 (Penalizing creeping function)**.** Let $\mathcal{R}$ be a knowledge base with $\langle \mathcal{R} \rangle = (r_1, \ldots, r_n)$ and $r_i = (\psi_i \mid \phi_i)[d_i]$ for $i = 1, \ldots, n$ and let α be the culpability vector of $\mathcal{R}$. Then the *penalizing creeping function* $\Xi^P_{\mathcal{R}}$ of $\mathcal{R}$ is the function $\Xi^P_{\mathcal{R}} : [0,1] \to \mathfrak{P}((\mathcal{L}(\mathsf{At}) \mid \mathcal{L}(\mathsf{At}))^{pr})$ defined via

$$\Xi^P_{\mathcal{R}}(\delta) =_{def} \Lambda_{\mathcal{R}}(u(d_1 + \delta\hat{\alpha}_1), \ldots, u(d_n + \delta\hat{\alpha}_n))$$

for $\delta \in [0,1]$.

Note that $\Xi^P_{\mathcal{R}}(0) = \mathcal{R}$ but it is not easy to see that $\Xi^P_{\mathcal{R}}(\delta)$ is consistent for some $\delta \in [0,1]$. As a matter of fact, neither a proof nor a counterexample has been found that shows whether $\Xi^P_{\mathcal{R}}(\delta)$ is consistent for some $\delta \in [0,1]$.

Conjecture 4.1. *For every $\mathcal{R}$ the knowledge base $\Xi^P_{\mathcal{R}}(\delta)$ is consistent for some $\delta \in [0,1]$.*

Consider the following example.

Example 4.16. We continue Example 4.15 from page 111 and consider the knowledge base $\mathcal{R}_2$ with $\langle \mathcal{R} \rangle = (r_1, r_2, r_3)$ and

$$r_1 = (a \mid b)[1] \qquad r_2 = (b)[1] \qquad r_3 = (a)[0] \quad .$$

For $\mathcal{R}_2$ it holds that $\hat{\alpha} = (-1, -1, 1)$. For $\delta \approx 0.38915$ we get

$$\Xi^P_{\mathcal{R}_2}(\delta) = \quad \{ \quad (a \mid b)[0.61805], (b)[0.61805], (a)[0.38195] \quad \} \quad .$$

which is consistent.

By also considering the examples in Appendix C on page 253—which also show that for the given knowledge bases there is some δ such that $\Xi^P_{\mathcal{R}}(\delta)$ is consistent—we are justified in believing Conjecture 4.1 to be true. However, due to the lack of a formal proof we have to consider the situation that the set

$$\{\delta \in [0,1] \mid \Xi^P_{\mathcal{R}}(\delta) \text{ is consistent}\}$$

may be empty. Therefore, we can define the penalizing creeping consistency restorer only for the case where this set is not empty.

Definition 4.11 (Penalizing creeping consistency restorer)**.** Let Υ^P be a function $\Upsilon^P : \mathfrak{P}((\mathcal{L}(\mathsf{At}) \mid \mathcal{L}(\mathsf{At}))^{pr}) \to \mathfrak{P}((\mathcal{L}(\mathsf{At}) \mid \mathcal{L}(\mathsf{At}))^{pr})$ defined as

$$\Upsilon^P(\mathcal{R}) =_{def} \Xi^P_{\mathcal{R}}(\delta^*)$$

with $\delta^* =_{def} \min\{\delta \in [0,1] \mid \Xi^P_{\mathcal{R}}(\delta)$ is consistent$\}$ if $\{\delta \in [0,1] \mid \Xi^P_{\mathcal{R}}(\delta)$ is consistent$\} \neq \emptyset$ and $\Upsilon^P(\mathcal{R}) =_{def} \mathsf{undef}$ otherwise.

Remark 4.1 on page 109 applies to the penalizing creeping consistency restorer as well. We now have a look at some examples.

Example 4.17. We continue Example 4.13 from page 110 with the knowledge base $\mathcal{R}'_2 = \{r_1, r_2, r_3, r_4\}$ given via

$$\begin{aligned} r_1 &= (a \mid b)[1] & r_2 &= (b)[1] \\ r_3 &= (a)[1] & r_4 &= (c)[0.3] \quad . \end{aligned}$$

Applying Υ^P to $\mathcal{R}'_2$ yields

$$\begin{aligned} \Upsilon^P(\mathcal{R}'_2) = \quad \{ \quad & (a \mid b)[0.61805], (b)[0.61805], \\ & (a)[0.38195], (c)[0.3] \quad \} \quad . \end{aligned}$$

As one can see, Υ^P conforms to $A^{\mathcal{R}'_2}$ by not penalizing r_4.

Example 4.18. We continue Example 4.14 from page 110 with the knowledge base $\mathcal{R}''_2 = \{r_1, r_2, r_3\}$ given via

$$r_1 = (a \mid b)[1] \qquad r_2 = (b)[1] \qquad r_3 = (a)[0.75] \quad .$$

Applying Υ^P on $\mathcal{R}''_2$ yields

$$\Upsilon^P(\mathcal{R}''_2) = \quad \{ \quad (a \mid b)[0.9142], (b)[0.9142], (a)[0.8358] \quad \} \quad .$$

Note that the probability of r_3 has been modified using the correct sign of its culpability.

As we have to adhere for the possibility of $\Xi^P_{\mathcal{R}}(\delta)$ being inconsistent for every $\delta \in [0,1]$ it may be the case that $\Upsilon^P(\mathcal{R}) = \mathsf{undef}$ for some $\mathcal{R}$. It follows that Υ^P does not satisfy (Existence) in general and therefore neither (Continuity) nor (Rational Non-Imposition).

Theorem 4.3. *The function Υ^P satisfies (Uniqueness), (Structural Preservation), (Success), (Consistency), and (Non-Dictatorship).*

The proof of Theorem 4.3 can be found in Appendix A on page 238.

Sadly, whether Υ^P satisfies ($A^{\mathcal{R}}$-Conformity) or (Inverse $A^{\mathcal{R}}$-Conformity) can only be conjectured but seems to be justified by considering the above examples.

Conjecture 4.2. *The function* Υ^P *satisfies (*$A^{\mathcal{R}}$*-Conformity) and (Inverse* $A^{\mathcal{R}}$*-Conformity).*

The problem in providing a formal proof for the above conjecture lies in the role of the function u in Definition 4.10. Consider some knowledge base $\mathcal{R}$ with $r, r' \in \mathcal{R}$ such that $r = (\psi \mid \phi)[0.1]$ and $r' = (\psi' \mid \phi')[0.3]$. Assume that the normalized culpability values of r and r' are 2 and 4, respectively. It follows the probabilities of r and r' in $\Xi^P_{\mathcal{R}}(\delta)$ are given by the functions $h_1(\delta) = u(0.1 + 2\delta)$ and $h_2(\delta) = u(0.3 + 4\delta)$, respectively. Now assume that the minimal δ^* such that $\Xi^P_{\mathcal{R}}(\delta^*)$ is consistent is given by $\delta^* = 0.5$. Then it follows that $(\psi \mid \phi)[1], (\psi' \mid \phi')[1] \in \Xi^P_{\mathcal{R}}(\delta^*)$. Hence, the probability of r has been modified by 0.9 and the probability of r' has been modified by 0.7. It follows that Ξ^P violates ($A^{\mathcal{R}}$-Conformity). The one key assumption in this scenario is that the minimal δ^* with $\Xi^P_{\mathcal{R}}(\delta^*)$ being consistent is large enough such that the probabilities of both r and r' reach the border of the interval $[0, 1]$. Remember that the mean distance culpability measure assigns the mean deviation of a probabilistic conditional's probability in the set $\mathcal{MD}(\mathcal{R})$. Consider again the conditional r' which has a normalized culpability value 4. As the probability of r' is 0.3, the maximal value for a deviation of r' is 0.7. Assuming that this is also the mean deviation this amounts to $0.7/4 = 0.175$ being the factor used for normalization and therefore $2 \cdot 0.175 = 0.35$ is the mean deviation of r. This amounts to r maximally deviating by a value of $2 \cdot 0.35 = 0.7$ in $\mathcal{MD}(\mathcal{R})$ (when the infimum of the deviation is 0 and the supremum is 0.7). However, in $\Xi^P_{\mathcal{R}}(\delta^*)$ the probabilistic conditional r deviates by a value of 0.9 which seems to be too drastic to be imaginable.

Under the assumption that both Conjecture 4.1 and Conjecture 4.2 are true we can strengthen Theorem 4.3 as follows.

Theorem 4.4. *If Conjectures 4.1 and 4.2 are true then* Υ^P *satisfies (Existence), (Uniqueness), (Structural Preservation), (Success), (Consistency), (Rational Non-Imposition), (Continuity), (Non-Dictatorship), (*$A^{\mathcal{R}}$*-Conformity), and (Inverse* $A^{\mathcal{R}}$*-Conformity).*

The proof of Theorem 4.4 can be found in Appendix A on page 239. In general, Υ^P does not satisfy (Irrelevance of Syntax) and (Pareto-Efficiency).

Example 4.19. Consider the probabilistic conditionals

$$r_1 =_{def} (a)[0.7] \qquad r_2 =_{def} (a)[0.4] \qquad r_3 =_{def} (\neg a)[0.6]$$

and the knowledge bases $\mathcal{R}_1 =_{def} \{r_1, r_2\}$ and $\mathcal{R}_2 =_{def} \{r_1, r_2, r_3\}$. As $r_2 \equiv^{\text{pr}} r_3$ it follows $\mathcal{R}_1 \equiv^{\text{cond}} \mathcal{R}_2$. The culpability values of the conditionals with respect to to their knowledge bases and the culpability measure $A^{\mathcal{R}}$ are given via

$$A^{\mathcal{R}_1}(r_1) = 0.15 \qquad A^{\mathcal{R}_1}(r_2) = 0.15$$
$$A^{\mathcal{R}_2}(r_1) = 0.3 \qquad A^{\mathcal{R}_2}(r_2) = 0 \qquad A^{\mathcal{R}_2}(r_3) = 0 \quad .$$

Note that $A^{\mathcal{R}_2}(r_2) = A^{\mathcal{R}_2}(r_3) = 0$ as modifying both r_2 and r_3 yields a more drastic change than just r_1. It follows that

$$\Upsilon^P(\mathcal{R}_1) = \{(a)[0.55]\} \qquad \Upsilon^P(\mathcal{R}_2) = \{r_2, r_3\}$$

with $\Upsilon^P(\mathcal{R}_1) \not\equiv^{\text{cond}} \Upsilon^P(\mathcal{R}_2)$, thus violating (Irrelevance of Syntax).

Example 4.20. Consider the knowledge base $\mathcal{R} =_{def} \{r_1, r_2, r_3, r_4, r_5\}$ given via

$$r_1 =_{def} (b|a)[1] \qquad r_2 =_{def} (a)[1] \qquad r_3 =_{def} (b)[0]$$
$$r_4 =_{def} (c)[0.3] \qquad r_5 =_{def} (c)[0.7]$$

and the culpability measure $A^{\mathcal{R}}$ with culpability values given as

$$A^{\mathcal{R}}(r_1) = 0.5 \qquad A^{\mathcal{R}}(r_2) = 0.5 \qquad A^{\mathcal{R}}(r_3) = 0.5$$
$$A^{\mathcal{R}}(r_4) = 0.2 \qquad A^{\mathcal{R}}(r_5) = 0.2 \quad .$$

It follows that it holds that $\Upsilon^P(\mathcal{R}) = \{r'_1, r'_2, r'_3, r'_4\}$ with

$$r'_1 = (b|a)[0.5] \qquad r'_2 = (a)[0.5] \qquad r'_3 = (b)[0.5] \qquad r'_4 = (c)[0.5] \quad .$$

However, also the knowledge base $\mathcal{R}' =_{def} \{r''_1, r''_2, r''_3, r''_4\}$ with

$$r''_1 =_{def} (b|a)[0.5] \quad r''_2 =_{def} (a)[0.6] \quad r''_3 =_{def} (b)[0.5] \quad r''_4 =_{def} (c)[0.5]$$

is consistent and $|1 - 0.6| < |1 - 0.5|$, i. e., the probabilistic conditional r_2 has been modified less in $\mathcal{R}'$ than in $\Upsilon^P(\mathcal{R})$. It follows that Υ^P does not satisfy (Pareto-Efficiency) in general.

4.3.3 *Smoothed Penalized Creeping*

The lack of a formal proof for Conjecture 4.1 and thus the possible dissatisfaction of (Existence) is a major drawback of the penalizing creeping consistency restorer. We now look at a hybrid approach that combines the unbiased and the penalizing consistency restorer in order to regain the satisfaction of (Existence). Our idea bases on the approach of the unbiased creeping function. We want to model a creeping function that starts with the original knowledge base for $\delta = 0$ and creeps towards a knowledge base that is satisfiable by the uniform probability function for $\delta = 1$.

In contrast to the unbiased creeping function, for each probabilistic conditional the gradient of this creeping is not a simple convex combination but obeys the culpability of the conditional. For that, let $\gamma, \beta_1, \beta_2, \theta \in \mathbb{R}$ and $v^{\theta}_{\gamma,\beta_1,\beta_2} : [0,1] \to [0,1]$ be the function

$$v^{\theta}_{\gamma,\beta_1,\beta_2}(x) =_{def} \begin{cases} \beta_2 + \gamma\theta x & \beta_2 < \beta_1 \text{ and } \beta_2 + \gamma\theta x < \beta_1 \\ & \quad \text{and } \gamma \neq 0 \\ \beta_2 - \gamma\theta x & \beta_2 > \beta_1 \text{ and } \beta_2 - \gamma\theta x > \beta_1 \\ & \quad \text{and } \gamma \neq 0 \\ \beta_2 & \gamma = 0 \text{ and } x \in [0,1) \\ \beta_1 & \text{otherwise} \end{cases} \tag{4.2}$$

The function $v^{\theta}_{\gamma,\beta_1,\beta_2}$ describes a weighted linear approach from β_2 to β_1 with gradient γ. Note that for $l, r \in \mathbb{R}$ with $l < r$ it holds that $v^{1}_{r/l,r,l}(x) = xl + (1-x)r$. The parameter θ is a scaling parameter. Figure 5 shows the general function $v^{\theta}_{\gamma,\beta_1,\beta_2}(x)$ and Figure 6 shows the specific function $v^{1}_{0.5,0.5,0.8}(x)$. Note also that $v^{\theta}_{\gamma,\beta_1,\beta_2}(x) = \beta_2$ for $\gamma = 0$ and $x \neq 1$ and $v^{\theta}_{\gamma,\beta_1,\beta_2}(x) = \beta_1$ for $\gamma = 0$ and $x = 1$. Hence, $v^{\theta}_{\gamma,\beta_1,\beta_2}$ is not continuous for $\gamma = 0$.

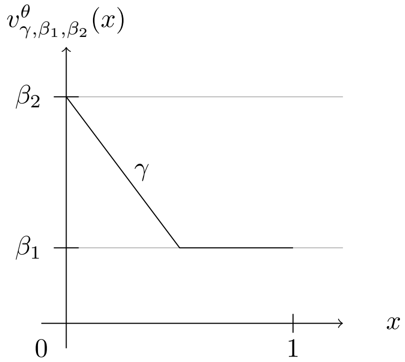

Figure 5: The function $v^{\theta}_{\gamma,\beta_1,\beta_2}(x)$

Our third approach of a creeping function uses the function v for creeping towards a knowledge base that can be satisfied by the uniform probability function while at the same time adhering to a culpability measure $C^{\mathcal{R}}$. In order to ensure that we end with a knowledge base that is satisfiable with the uniform probability function we have to set the scaling parameter θ for v appropriately. In the *smoothed penalizing creeping function* that is to be defined we set the parameter β_1 of v to the uniform probability of the conditional. For $\gamma \neq 0$ we want v to be continuous so we have to ensure that v approaches β_1 when going with δ to 1, i. e., we have to ensure that

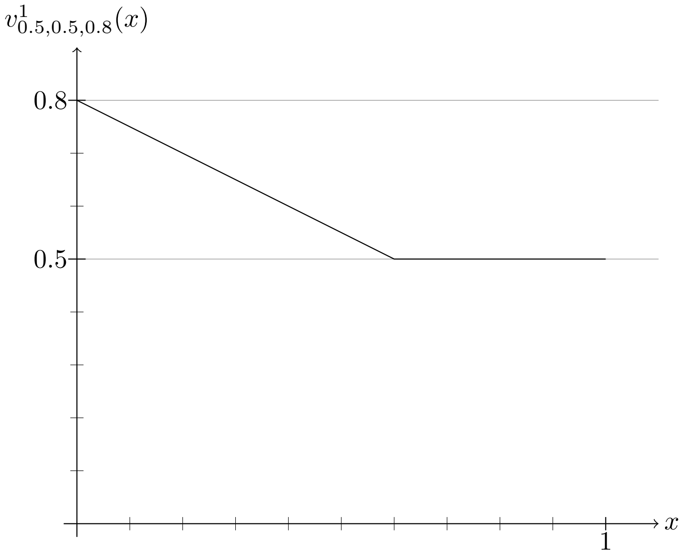

Figure 6: The function $v^1_{0.5,0.5,0.8}(x)$

the absolute value of the gradient of v is not too small, cf. Figure 5. For a culpability measure $C^{\mathcal{R}}$ let

$$\theta^* =_{def} \inf\{\theta \in \mathbb{R}^+_0 \mid \forall i : \mathsf{ucp}(\psi_i \mid \phi_i) \in \mathsf{Im}\ v^{\theta}_{C^{\mathcal{R}}(r_i),\mathsf{ucp}(\psi_i \mid \phi_i),d_i}\} \tag{4.3}$$

denote the *smoothed scaling factor* of $C^{\mathcal{R}}$ and $\mathcal{R}$ with $\langle\mathcal{R}\rangle = (r_1, \ldots, r_n)$ and $r_i = (\psi_i \mid \phi_i)[d_i]$ for $i = 1, \ldots, n$. The smoothed scaling factor θ^* is the smallest scaling factor such that for every probabilistic conditional in the knowledge base it is ensured that $v^{\theta}_{C^{\mathcal{R}}(r_i),\mathsf{ucp}(\psi_i \mid \phi_i),d_i}(1) = \mathsf{ucp}(\psi_i \mid \phi_i)$ (this follows from the monotonicity of v) and that $v^{\theta}_{C^{\mathcal{R}}(r_i),\mathsf{ucp}(\psi_i \mid \phi_i),d_i}$ is continuous for $C^{\mathcal{R}}(r_i) \neq 0$. The factor θ^* can also be characterized via

$$\theta^* = \min\left\{\frac{|d_i - \mathsf{ucp}(\psi_i \mid \phi_i)|}{C^{\mathcal{R}}(r_i)} \mid i = 1, \ldots, n\right\} \quad . \tag{4.4}$$

This is easy to see, as due to the monotonicity of v it suffices to require

$$v^{\theta}_{C^{\mathcal{R}}(r_i),\mathsf{ucp}(\psi_i \mid \phi_i),d_i}(1) = \mathsf{ucp}(\psi_i \mid \phi_i) \tag{4.5}$$

for each $i = 1, \ldots, n$. Assume that $d_i < \mathsf{ucp}(\psi_i \mid \phi_i)$. By considering the first case of (4.2) we want θ to satisfy

$$d_i + C^{\mathcal{R}}(r_i) \cdot \theta \cdot 1 = \mathsf{ucp}(\psi_i \mid \phi_i)$$

which is equivalent to

$$\theta = \frac{\mathsf{ucp}(\psi_i \,|\, \phi_i) - d_i}{C^{\mathcal{R}}(r_i)} \quad .$$

Similarly, for the case $d_i > \mathsf{ucp}(\psi_i \,|\, \phi_i)$ we get

$$\theta = \frac{d_i - \mathsf{ucp}(\psi_i \,|\, \phi_i)}{C^{\mathcal{R}}(r_i)}$$

and together

$$\theta = \frac{|d_i - \mathsf{ucp}(\psi_i \,|\, \phi_i)|}{C^{\mathcal{R}}(r_i)} \quad .$$

As we consider a finite set of conditionals the infimum in Equation (4.3) becomes a minimum in Equation (4.4). We are now able to define the *smoothed penalizing creeping function* as follows.

Definition 4.12 (Smoothed penalizing creeping function)**.** The *smoothed penalizing creeping function* $\Xi^S_{\mathcal{R},C}$ of $\mathcal{R}$ with $\langle \mathcal{R} \rangle = (r_1, \ldots, r_n)$ and $r_i = (\psi_i \,|\, \phi_i)[d_i]$ for $i = 1, \ldots, n$ and C is the function $\Xi^S_{\mathcal{R},C} : [0,1] \to \mathfrak{P}((\mathcal{L}(\mathsf{At}) \,|\, \mathcal{L}(\mathsf{At}))^{pr})$ defined via

$$\Xi^S_{\mathcal{R},C}(\delta) =_{def} \Lambda_{\mathcal{R}}(v^{\theta^*}_{C^{\mathcal{R}}(r_1),\mathsf{ucp}(\psi_1 \,|\, \phi_1),d_1}(\delta), \ldots, v^{\theta^*}_{C^{\mathcal{R}}(r_n),\mathsf{ucp}(\psi_n \,|\, \phi_n),d_n}(\delta))$$

for $\delta \in [0,1]$ and θ^* being the smoothed scaling factor of C and $\mathcal{R}$.

Note that for every knowledge base $\mathcal{R}$ and culpability measure C it holds that $\Xi^S_{\mathcal{R},C}(0) = \mathcal{R}$ and $\Xi^S_{\mathcal{R},C}(1)$ is satisfiable by the uniform probability function. In particular, it holds that

$$\Xi^S_{\mathcal{R},C}(1) =_{def} \Lambda_{\mathcal{R}}(\mathsf{ucp}(\psi_1 \,|\, \phi_1), \ldots, \mathsf{ucp}(\psi_n \,|\, \phi_n))$$

as ensured by the definition of $v^{\theta}_{\gamma,\beta_1,\beta_2}$ above.

Definition 4.13 (Smoothed penalizing creeping consistency restorer)**.** Let Υ^S_C be a function $\Upsilon^S_C : \mathfrak{P}((\mathcal{L}(\mathsf{At}) \,|\, \mathcal{L}(\mathsf{At}))^{pr}) \to \mathfrak{P}((\mathcal{L}(\mathsf{At}) \,|\, \mathcal{L}(\mathsf{At}))^{pr})$ defined as

$$\Upsilon^S_C(\mathcal{R}) =_{def} \Xi^U_{\mathcal{R},C}(\delta^*)$$

with $\delta^* =_{def} \min\{\delta \in [0,1] \mid \Xi^P_{\mathcal{R}}(\delta) \text{ is consistent}\}$.

Theorem 4.5. Υ^S_C *satisfies (Existence), (Uniqueness), (Structural preservation), (Success), (Consistency), and (Rational Non-Imposition).*

The proof of Theorem 4.5 can be found in Appendix A on page 239. The function Υ^S_C does not satisfy (Irrelevance of Syntax) which can be shown by using the same example as for the penalizing consistency restorer, cf. Example 4.19 on page 114. It also does not satisfy (Non-Dictatorship) using the same argumentation as for the unbiased creeping function. The function Υ^S_C does not satisfy (Continuity) as $v^{\theta}_{\gamma,\beta_1,\beta_2}$ is discontinuous for $\gamma = 0$. In general Υ^S_C does not satisfy (C-Conformity) nor (Inverse C-Conformity) as well. The following example illustrates this issue.

Example 4.21. Consider the knowledge base $\mathcal{R} =_{def} \{r_1, r_2, r_3, r_4\}$ given via

$$r_1 =_{def} (a)[0.6] \qquad r_2 =_{def} (a)[0.7]$$
$$r_3 =_{def} (b)[0.8] \qquad r_4 =_{def} (b)[0.85] \quad .$$

It holds that $A^{\mathcal{R}}(r_1) = A^{\mathcal{R}}(r_2) = 0.05$ and $A^{\mathcal{R}}(r_3) = A^{\mathcal{R}}(r_4) = 0.025$ and $\Upsilon^S_{A^{\mathcal{R}}}(\mathcal{R}) = \{(a)[0.5], (b)[0.5]\}$ as $\mathsf{ucp}(a) = \mathsf{ucp}(b) = 0.5$. Hence, $\Upsilon^S_{A^{\mathcal{R}}}$ violates ($A^{\mathcal{R}}$-Conformity) as the culpability value of r_1 is at least as large as the culpability value of r_2 but r_2 has been modified more drastically, i. e., $A^{\mathcal{R}}(r_1) \geq A^{\mathcal{R}}(r_2)$ but $|0.6 - 0.5| < |0.7 - 0.5|$. Furthermore, $\Upsilon^S_{A^{\mathcal{R}}}$ violates (Inverse $A^{\mathcal{R}}$-Conformity) as r_3 has been modified more drastically than r_1 but the culpability value of r_1 is larger than the culpability value of r_3, i. e., $|0.8 - 0.5| \geq |0.6 - 0.5|$ but $A^{\mathcal{R}}(r_1) > A^{\mathcal{R}}(r_3)$.

Similar observations like in the above example can be made for other culpability measures. The function Υ^S_C also fails to satisfy (Pareto-Efficiency), in particular, $\Upsilon^S_{A^{\mathcal{R}}}$ violates (Pareto-Efficiency) in the same example used for proving this statement for the penalizing consistency restorer, cf. Example 4.20 on page 115 (see also Table 27 in Appendix C on page 263).

4.4 SOLVING CONFLICTS BY BALANCED DISTANCE MINIMIZATION

In the following, we describe a declarative approach to restore consistency that is inspired by the computation of the MinDev inconsistency measure.

Let $\mathcal{R}$ be a knowledge base with $\langle\mathcal{R}\rangle = (r_1, \ldots, r_n)$ and $r_i = (\psi_i \mid \phi_i)[d_i]$ for $i = 1, \ldots, n$, $C^{\mathcal{R}}$ be some culpability measure, and let $\eta_1, \ldots, \eta_n$ be some variables. Then consider again the set of constraints $\mathsf{DevCons}(\mathcal{R})$ (see page 64).

$$\sum_{\omega \in \mathsf{Mod}^{\mathsf{P}}(\psi_i \wedge \phi_i)} \alpha_\omega = (d_i + \eta_i) \cdot \sum_{\omega \in \mathsf{Mod}^{\mathsf{P}}(\phi_i))} \alpha_\omega \tag{4.6}$$

$$0 \leq d_1 + \eta_1 \leq 1, \ldots, 0 \leq d_n + \eta_n \leq 1 \tag{4.7}$$

$$\sum_{\omega \in \Omega(\mathsf{At})} \alpha_\omega = 1 \tag{4.8}$$

$$\alpha_\omega \geq 0 \quad \text{for all } \omega \in \Omega(\mathsf{At}) \quad . \tag{4.9}$$

Let C^R be a culpability measure. In order to explicitly impose (C^R-Conformity) and (Inverse C^R-Conformity) to be satisfied consider

$$|\eta_i| \geq |\eta_j| \quad \text{for } i, j = 1, \ldots, n \text{ with } i \neq j \text{ and } C^R(r_i) \geq C^R(r_j) \tag{4.10}$$

Let $\mathsf{CRDevCon}(C, \mathcal{R})$ denote the set of constraints (4.6), (4.7), (4.8), (4.9), and (4.10) for some culpability measure C. Now consider minimizing

$$f_C^B(\eta_1, \ldots, \eta_n) =_{def} |\eta_1| + \cdots + |\eta_n| \tag{4.11}$$

with respect to $\mathsf{CRDevCon}(C, \mathcal{R})$.

Definition 4.14 (Balanced consistency restorer)**.** Let C be some culpability measure and $\Upsilon_C^B : \mathfrak{P}((\mathcal{L}(\mathsf{At}) \mid \mathcal{L}(\mathsf{At}))^{pr}) \rightarrow \mathfrak{P}((\mathcal{L}(\mathsf{At}) \mid \mathcal{L}(\mathsf{At}))^{pr})$ be defined as

$$\Upsilon_C^B(\mathcal{R}) =_{def} \Lambda_{\mathcal{R}}(d_1 + \eta_1^*, \ldots, d_n + \eta_n^*)$$

where $\eta_1^*, \ldots, \eta_n^*$ are uniquely determined variable assignments that minimize f_C^B with respect to $\mathsf{CRDevCon}(C, \mathcal{R})$. If there is no unique optimal solution then $\Upsilon_C^B(\mathcal{R}) =_{def} \mathsf{undef}$.

As the following example shows there are indeed cases where no unique optimal solution for $\mathsf{CRDevCon}(C, \mathcal{R})$ can be obtained.

Example 4.22. Let $\mathcal{R}$ be the knowledge base given via $\langle\rangle\rangle = (r_1, r_2, r_3)$ and

$$r_1 =_{def} (a)[0.3] \qquad r_2 =_{def} (b)[0.4] \qquad r_3 =_{def} (a \wedge b)[0.6] \quad .$$

Consider the culpability measure $A^{\mathcal{R}}$ with the culpability values

$$A^{\mathcal{R}}(r_1) = 0.05 \qquad A^{\mathcal{R}}(r_2) = 0 \qquad A^{\mathcal{R}}(r_3) = 0.25 \quad .$$

Let $\eta_1^*, \eta_2^*, \eta_3^*$ be such that $\Lambda_{\mathcal{R}}(0.3 + \eta_1^*, 0.4 + \eta_2^*, 0.6 + \eta_3^*)$ is consistent and $|\eta_1^*| + |\eta_2^*| + |\eta_3^*|$ is minimal with obeying (4.10) which amounts to (among others) $\eta_1^* \leq \eta_3^*$. It follows that for each $x \in [0, 1]$ the values

$$\eta_1^* = 0.3x \qquad \eta_2^* = 0 \qquad \eta_3^* = -0.3(1 - x)$$

yield consistent $\Lambda_{\mathcal{R}}(0.3 + \eta_1^*, 0.4 + \eta_2^*, 0.6 + \eta_3^*)$ and $|\eta_1^*| + |\eta_2^*| + |\eta_3^*| = 0.3$ minimal.

Nevertheless, Υ^B_C still satisfies (Uniqueness) as for $\mathcal{R}$ from the above example it follows $\Upsilon^B_C(\mathcal{R}) = \mathsf{undef}$ and this is uniquely determined. However, from the above discussion it also follows that Υ^B_C does not satisfy (Existence), (Continuity) and (Rational Non-Imposition) in general.

Theorem 4.6. *Let C satisfy (Inc^*-symmetry). Then Υ^B_C satisfies (Uniqueness), (Structural Preservation), (Success), (Consistency), (C-Conformity), (Inverse C-Conformity), and (Non-Dictatorship).*

The proof of Theorem 4.6 can be found in Appendix A on page 240. The function Υ^B_C does not satisfy (Irrelevance of Syntax) which can be shown by using the same example as for the penalizing consistency restorer, cf. Example 4.19 on page 114. Furthermore, the function Υ^B_C does not satisfy (Pareto-Efficiency) in general as the following example shows.

Example 4.23. Let $\mathcal{R} =_{def} \{r_1, r_2, r_3, r_4, r_5, r_6\}$ be a knowledge base given via

$$\begin{array}{lll} r_1 =_{def} (b \mid a)[1] & r_2 =_{def} (a)[1] & r_3 =_{def} (b)[0] \\ r_4 =_{def} (d \mid c)[0.9] & r_5 =_{def} (c)[0.9] & r_6 =_{def} (d)[0.1] \end{array} .$$

It follows

$$\begin{aligned} S^{\mathcal{R}}_{\mathsf{Inc}^{\mathsf{MI}}}(r_1) &= S^{\mathcal{R}}_{\mathsf{Inc}^{\mathsf{MI}}}(r_2) = S^{\mathcal{R}}_{\mathsf{Inc}^{\mathsf{MI}}}(r_3) = S^{\mathcal{R}}_{\mathsf{Inc}^{\mathsf{MI}}}(r_4) = S^{\mathcal{R}}_{\mathsf{Inc}^{\mathsf{MI}}}(r_5) \\ &= S^{\mathcal{R}}_{\mathsf{Inc}^{\mathsf{MI}}}(r_6) \approx 0.111 \end{aligned}$$

and it holds that $\Upsilon^B_{S^{\mathcal{R}}_{\mathsf{Inc}^{\mathsf{MI}}}}(\mathcal{R}) = \{r'_1, r'_2, r'_3, r'_4, r'_5, r'_6\}$ with

$$\begin{array}{lll} r'_1 = (b \mid a)[0.618] & r'_2 = (a)[0.618] & r'_3 = (b)[0.382] \\ r'_4 = (d \mid c)[0.518] & r'_5 = (c)[0.518] & r'_6 = (d)[0.482] \end{array} .$$

Note, however, that $\Upsilon^B_{S^{\mathcal{R}}_{\mathsf{Inc}^{\mathsf{MI}}}}(\mathcal{R})$ is not Pareto-efficient as $\mathcal{R}' = \{r''_1, r''_2, r''_3, r''_4, r''_5, r''_6\}$ with

$$\begin{array}{lll} r''_1 = (b \mid a)[0.618] & r''_2 = (a)[0.618] & r''_3 = (b)[0.382] \\ r''_4 = (d \mid c)[0.618] & r''_5 = (c)[0.618] & r''_6 = (d)[0.382] \end{array}$$

is consistent as well.

The previous example showed that there are cases where satisfaction of (C-Conformity) disallows satisfaction of (Pareto-Efficiency). There, all probabilistic conditionals are assigned the same culpability value by $S^{\mathcal{R}}_{\mathsf{Inc}^{\mathsf{MI}}}$ and there are two minimal inconsistent subsets of $\mathcal{R}$—$\{r_1, r_2, r_3\}$ and $\{r_4, r_5, r_6\}$—that both model a similar inconsistency but differ slightly in their probabilities. However, the culpability measure $S^{\mathcal{R}}_{\mathsf{Inc}^{\mathsf{MI}}}$ is indifferent about this and as $\Upsilon^{B}_{S^{\mathcal{R}}_{\mathsf{Inc}^{\mathrm{MI}}}}$ satisfies ($S^{\mathcal{R}}_{\mathsf{Inc}^{\mathrm{MI}}}$-Conformity) each probabilistic conditional is modified by the same value. Looking at the set $\{r_1, r_2, r_3\}$, one can see that each probabilistic conditional has to be modified by at least a value of (approximately) 0.383. For $\{r_4, r_5, r_6\}$ a modification by 0.283 suffices, but as $\Upsilon^{B}_{S^{\mathcal{R}}_{\mathsf{Inc}^{\mathrm{MI}}}}$ satisfies ($S^{\mathcal{R}}_{\mathsf{Inc}^{\mathrm{MI}}}$-Conformity) all probabilistic conditionals are modified by 0.383. This yields a solution that is not Pareto-efficient. Note also that the observation in Example 4.23 is not restricted to culpability measures that do not satisfy (Continuity). The same reasoning applies for e.g. the knowledge base $\mathcal{R} =_{def} \{r_1, r_2, r_3, r_4, r_5\}$ given via

$$r_1 =_{def} (b \mid a)[1] \qquad r_2 =_{def} (a)[1] \qquad r_3 =_{def} (b)[0]$$
$$r_4 =_{def} (c)[1] \qquad r_5 =_{def} (c)[0]$$

and the culpability measure $A^{\mathcal{R}}$ as

$$A^{\mathcal{R}}(r_1) = A^{\mathcal{R}}(r_2) = A^{\mathcal{R}}(r_3) = A^{\mathcal{R}}(r_4) = A^{\mathcal{R}}(r_5) = 0.5 \quad .$$

4.5 RELATED WORK

The notion of culpability measures used in this chapter has been adopted from the work (Daniel, 2009) and that work also proposes a specific implementation for a culpability measure. Moreover, the work (Daniel, 2009) also considers the problem of probabilistic reasoning from inconsistent knowledge bases. In particular, (Daniel, 2009) extends reasoning based on the principle of maximum entropy to inconsistent knowledge bases and therefore takes a different approach for handling inconsistency. We discuss the approach of (Daniel, 2009) in more depth below. Culpability measures have also been used under the notion of *inconsistency values* in the works of Hunter et. al., cf. e.g. (Hunter and Konieczny, 2008, 2010). However, those works are concerned with measuring culpabilities in classical logic. In this thesis we adopted several ideas and approaches of (Hunter and Konieczny, 2006, 2008, 2010) for the framework of probabilistic conditional logic. In particular, we extended the Shapley culpability measure and were able to transfer several results to this more expressive framework.

Other related work for our approach to restoring consistency in probabilistic knowledge base consists mainly of the works (Rödder and Xu, 2001) and (Finthammer *et al.*, 2007) which are also reviewed below.

4.5.1 *Culpabilities and Candidacy degrees*

Recall that the inconsistency measure Inc^{h}_{μ} is defined as

$$\mathsf{Inc}^{h}_{\mu}(\mathcal{R}) =_{def} 1 - \max_{P \in \mathcal{P}^{\mathrm{P}}(\mathsf{At})} \mathfrak{C}^{h}_{\mathcal{R}}(P) \quad ,$$

i. e., $\mathsf{Inc}^h_\mu(\mathcal{R})$ is one minus the maximal candidacy degree of a probability function, cf. Section 3.5.3. For the candidacy function $\mathfrak{C}^h_\mathcal{R}$ of a knowledge base $\mathcal{R}$ let

$$\hat{\Omega}^h_\mathcal{R} =_{def} \arg \max_{P \in \mathcal{P}^\mathsf{P}(\mathsf{At})} \mathfrak{C}^h_\mathcal{R}(P)$$

be the set of probability functions that are assigned a maximal candidacy degree. Then a culpability measure $C^\mathcal{R}_h$ can be defined via

$$C^\mathcal{R}_h(r) =_{def} 1 - \mathfrak{C}^h_{\{r\}}(P)$$

for $r \in \mathcal{R}$ and some $P \in \hat{\Omega}^h_\mathcal{R}$. In (Daniel, 2009) Daniel admits that $C^\mathcal{R}_h(r)$ is not well-defined if there are $P, P' \in \hat{\Omega}^h_\mathcal{R}$ with $\mathfrak{C}^h_{\{r\}}(P) \neq \mathfrak{C}^h_{\{r\}}(P')$ for some $r \in \mathcal{R}$. However, it is conjectured that $C^\mathcal{R}_h$ is well-defined if $h = h_{\mathrm{HG}}$ is defined via

$$h_{\mathrm{HG}}(x) =_{def} 1 + \frac{2}{\sqrt{\pi}} \int_0^{-\frac{x}{\sqrt{2}}} e^{-t^2} dt \quad .$$

The function h_{HG} is called the *half-Gaussian blur* (with reliability 0.5). Under the assumption that it holds that $\mathfrak{C}^{h_{\mathrm{HG}}}_{\{r\}}(P) = \mathfrak{C}^{h_{\mathrm{HG}}}_{\{r\}}(P')$ for all $P, P' \in \hat{\Omega}^{h_{\mathrm{HG}}}_\mathcal{R}$ and all $r \in \mathcal{R}$, Daniel proves that $C^\mathcal{R}_{h_{\mathrm{HG}}}$ satisfies (among others) an alternative notion of ($\mathsf{Inc}^{h_{\mathrm{HG}}}_\mu$-Distribution) defined on multiplication rather than summation. More precisely, it is shown that it holds that

$$\mathsf{Inc}^{h_{\mathrm{HG}}}_\mu(\mathcal{R}) = 1 - \prod_{r \in \mathcal{R}} (1 - C^\mathcal{R}_{h_{\mathrm{HG}}}(r))$$

which derives directly from the definition. As a consequence, $C^\mathcal{R}_{h_{\mathrm{HG}}}$ does not satisfy ($\mathsf{Inc}^{h_{\mathrm{HG}}}_\mu$-Distribution) in general which is defined as

($\mathsf{Inc}^{h_{\mathrm{HG}}}_\mu$-Distribution) $\sum_{r \in \mathcal{R}} C^\mathcal{R}(r) = \mathsf{Inc}^{h_{\mathrm{HG}}}_\mu(\mathcal{R})$

for some culpability measure $C^\mathcal{R}$. In other words, while the general property (Inc-Distribution) says that the culpability measure and the inconsistency measure Inc have an additive relationship, the inconsistency measure $\mathsf{Inc}^{h_{\mathrm{HG}}}_\mu$ is the *noisy-or* of the culpability measure $C^\mathcal{R}_{h_{\mathrm{HG}}}$, cf. (Pearl, 1998). It is also shown that $C^\mathcal{R}_{h_{\mathrm{HG}}}$ satisfies a similar property like (Minimality). In particular, it holds that

$$\text{If } \mathcal{R} \models^* r \text{ then } C^\mathcal{R}_{h_{\mathrm{HG}}}(r) = 0$$

with $\models^*$ is one of $\models_{\mathrm{ic}}$, $\models_{\mathrm{ff}}$ as defined in (Daniel, 2009). As for the properties of Inc^h_μ no statements with respect to free conditionals are given in (Daniel, 2009).

Using candidacy functions Daniel extends reasoning based on the principle of maximum entropy (see also Section 2.3) as follows. Recall that for consistent $\mathcal{R}$ the maximum entropy model P^* of $\mathcal{R}$ is uniquely determined by

$$P^* = \arg \max_{P \models^{pr} \mathcal{R}} H(P) \quad , \tag{4.12}$$

see also Definition 2.27 on page 32. This definition is extended in (Daniel, 2009) by defining P^* via

$$P^* =_{def} \arg \max_{P \in \hat{\Omega}^h_{\mathcal{R}}} H(P) \quad . \tag{4.13}$$

This means that the maximum entropy model of a (possibly inconsistent) knowledge base $\mathcal{R}$ is selected among the probability functions that have the maximal candidacy degree with respect to $\mathcal{R}$. Note that (4.13) is equivalent to (4.12) if $\mathcal{R}$ is consistent as for all P it holds that $P \models^{pr} \mathcal{R}$ if and only if $\mathfrak{C}^h_{\mathcal{R}}(P) = 1$. For consistent $\mathcal{R}$ the latter is equivalent to $P \in \hat{\Omega}^h_{\mathcal{R}}$. Hence, the inference process defined in (Daniel, 2009) clearly extends reasoning based on the principle of maximum entropy to inconsistent knowledge bases and it is also shown that P^* as defined in (4.13) satisfies several rationality postulates like *uniqueness* and *irrelevant information*. As discussed before, this approach of handling inconsistencies is orthogonal to our approach. Both approaches are justifiable and it depends on the intended application which paradigm is more suitable.

4.5.2 *Qualitative Modification and Generalized Divergence*

In (Rödder and Xu, 2001) three approaches are proposed for restoring consistency in a knowledge base $\mathcal{R}$. The first two approaches are very similar and follow the paradigm of *qualitative modifications* of conditionals, see also (Rödder and Xu, 1999) and (Kern-Isberner and Rödder, 2003). In those approaches each probabilistic conditional $r_i = (\psi_i \,|\, \phi_i)[d_i] \in \mathcal{R}$ is extended to $(\psi_i \,|\, \phi_i \wedge w_i)[d_i]$ with a new proposition w_i for $i = 1, \ldots, n$. By doing so, inconsistencies in the former knowledge base are resolved as the actual probabilities of the new conditionals heavily depend on the probabilities of the w_i ($i = 1, \ldots, n$) which are unrestricted in general. The third approach of (Rödder and Xu, 2001) is to define the new probability of a probabilistic conditional $(\psi \,|\, \phi)[d] \in \mathcal{R}$ to be $P^*(\psi \,|\, \phi)$ where P^* is the unique probability function for the solution of $\mathsf{Inc}_{gd}(\mathcal{R})$, cf. Section 3.5.2 and particularly Equation (3.31) on page 81. Note that there is no motivation and no evaluation of those approaches given in (Rödder and Xu, 2001). The qualitative approach is hard to compare to our notion of consistency restorers as the former does not satisfy (Structural Preservation) which is a necessary requirement for our approaches to restoring consistency. The approach of using the solution P^* that minimizes the inconsistency measure Inc_{gd} to define the new probabilities of the conditionals falls into the same category as the consistency restorers developed in this chapter. As the inconsistency measure Inc_{gd} is very similar to Inc^* the outcomes of the third approach of (Rödder and Xu, 2001) can be expected to be similar to the balanced consistency restorer Υ^B_C when ignoring the culpabilities of the probabilistic conditionals. However, as there are no results on the quality of the approaches of (Rödder and Xu, 2001) further discussion has to be postponed for future work.

4.5.3 Heureka

The work (Finthammer *et al.*, 2007)—see also (Finthammer, 2008)—share the same motivation as our work in this chapter, namely, restoring consistency in probabilistic conditional knowledge bases. However, while we take a principled approach and base our notion of consistency restoring on a theoretical foundation the works (Finthammer *et al.*, 2007) handle the problem of restoring consistency in a more pragmatic way. More precisely, the system Heureka presented in (Finthammer *et al.*, 2007) restores consistency mainly by 1.) removing probabilistic conditionals and 2.) treating probabilistic conditionals as bounded probabilistic conditionals and widening their interval appropriately. By assigning priorities to the probabilistic conditionals and specifying optimization criteria like "minimize number of removed/modified conditionals" or "minimize sum of priorities of removed/modified conditionals" the user can control the result of the restoration to some extent. Restoring consistency by removing rules is a rather drastic approach and results in a lot of information getting lost. The main issue with widening the intervals of probabilistic conditionals is a categorical one as consistency is restored by switching to a more expressive framework. If one is dependent on using the framework of probabilistic conditional logic this approach is not applicable. However, the idea of widening the intervals of probabilistic conditionals is similar in spirit to the approach taken by creeping functions. By a stepwise widening of the intervals a creeping function can be simulated. Nonetheless, in (Finthammer *et al.*, 2007) no specific strategy is given of how to widen the intervals appropriately, except a uniform widening of the intervals of all conditionals which is a similar to the unbiased creeping function.

In contrast to the work of (Finthammer *et al.*, 2007) our approach is principled and we have shown that the consistency restorers investigated in this section satisfy several quality criteria. Another issue with the approach of (Finthammer *et al.*, 2007) is the need to specify the priorities of conditionals used by the heuristics to restore consistency. In (Finthammer *et al.*, 2007) no hint is given of how to obtain the priorities other than in a user-specified manner. However, one possible approach to obtain priorities is to use (inverse) culpability measures. A large culpability value can be translated to a small priority with the intended meaning that this rule may be likely removed or changed. A more throughout discussion of this idea may be part of future work.

Nonetheless, the aim of the system Heureka lies in effective and simple consistency restoration which makes it suitable for real-world applications if the results can be verified by a knowledge engineer.

4.6 SUMMARY AND DISCUSSION

In this chapter we investigated approaches for restoring consistency in probabilistic conditional knowledge bases. In particular, we extended the notion of inconsistency measures to culpability measures which allow to assign a degree of culpability to each probabilistic conditional in a knowledge base. We gave two implementations of culpability measures: the mean distance culpability measure and the Shapley culpability measure. Using these culpability measures we approached the problem of consistency restoration in a principled fashion. We focused our attention to methods for consistency restoration that keep the structure of the knowledge base intact and are only allowed to modify probabilities of conditionals. We devised a series of rationality postulates and adopted many other from the related field of belief merging. Several of these postulates are influenced by the culpability values of the probabilistic conditionals and follow the idea that probabilistic conditionals that are more culpable are modified to more extent. In the following, we developed two families of approaches of consistency restorers. The first one based on the notion of a creeping function and we gave three instantiations for specific creeping functions. Our second approach extends the optimization problem used to determine the value of the Min-Dev inconsistency measure from the previous chapter by incorporating a culpability measure that guides the search for finding a reasonable solution. Finally, we reviewed related work with respect to culpability measures and consistency restoration.

Table 5 gives an overview on the properties that are satisfied by the consistency restorers investigated in this chapter. The notation "(X)" means that it is only conjectured that the corresponding consistency restorer satisfies the property and the notation "A" means satisfaction of the conformity properties with respect to the mean distance culpability measure A. As one can see, none of the approaches satisfies (Pareto-Efficiency) in general. The discussion regarding Example 4.23 on page 121 suggests that (Pareto-Efficiency) and (C-Conformity) are conflicting demands, at least for those culpability measures discussed in this chapter. However, further work is mandatory and we only took a first step in the field of principled consistency restoration in probabilistic knowledge bases. Thus, we are currently not in the situation to state an analogon of Arrow's famous impossibility theorem for principled consistency restoration, cf. (Arrow, 1950).

Appendix C on page 253 ff. lists the outcomes of applying the consistency restorers Υ^{U}, Υ^{P}, Υ^{S}_{C}, and Υ^{B}_{C} for different choices of the C on several benchmark examples, some of them have already been discussed in this chapter. Prototypical implementations of the consistency restorers—which

Property	Υ^{U}	Υ^{P}	Υ^{S}_{C}	Υ^{B}_{C}
(Existence)	X	(X)	X	
(Uniqueness)	X	X	X	X
(Structural Preservation)	X	X	X	X
(Success)	X	X	X	X
(Irrelevance of Syntax)	X			
(Consistency)	X	X	X	X
(Pareto-Efficiency)				
(Non-Dictatorship)		X		X
(Rational Non-Imposition)	X	(X)	X	
($C^{\mathcal{R}}$-Conformity)		(A)		X
(Inverse $C^{\mathcal{R}}$-Conformity)		(A)		X
(Continuity)	X	(X)		

Table 5: Comparison of consistency restorers

were used for computing the values shown in Appendix C—can be found in the Tweety library for artificial intelligence[3].

In this chapter and the previous one we investigated the issue of inconsistencies in (propositional) probabilistic conditional logic in great depth. It remains to investigate whether our approaches are applicable if we switch to probabilistic conditionals based on first-order logic. As we will see in the following chapter there are more fundamental problems in the relational context that have to be solved first.

3 http://sourceforge.net/projects/tweety/

5 RELATIONAL PROBABILISTIC CONDITIONAL LOGIC

Hitherto we considered probabilistic knowledge representation employing propositional models of belief. For the rest of this thesis we switch to a more expressive framework, namely, probabilistic knowledge representation on (restricted) first-order logic. As has been pointed out in Section 2.4, probabilistic reasoning on relational domains is a rather novel research area within artificial intelligence. However, most of the existing approaches are primarily concerned with machine learning problems, and do not care about logical or formal properties of relational probabilistic reasoning. The following example, inspired by (Delgrande, 1998), illustrates that defining a proper semantics for first-order probabilistic knowledge bases is not straightforward. Let $\mathit{elephant}(\mathsf{X})$ denote that X is an elephant, $\mathit{keeper}(\mathsf{X})$ means that X is a keeper, and $\mathit{likes}(\mathsf{X},\mathsf{Y})$ denotes that X likes Y. Consider the following set of relational conditionals.

$$r_1 =_{def} (\mathit{likes}(\mathsf{X},\mathsf{Y}) \mid \mathit{elephant}(\mathsf{X}) \wedge \mathit{keeper}(\mathsf{Y}))[0.6]$$
$$r_2 =_{def} (\mathit{likes}(\mathsf{X},\mathsf{fred}) \mid \mathit{elephant}(\mathsf{X}) \wedge \mathit{keeper}(\mathsf{fred}))[0.4]$$
$$r_3 =_{def} (\mathit{likes}(\mathsf{clyde},\mathsf{fred}) \mid \mathit{elephant}(\mathsf{clyde}) \wedge \mathit{keeper}(\mathsf{fred}))[0.7]$$

An informal interpretation of the above conditionals can be given as follows. With a probability of 0.6 elephants like keepers (r_1), with a probability of 0.4 elephants like keeper Fred (r_2), and with probability 0.7 elephant Clyde likes keeper Fred (r_3). From the point of view of commonsense reasoning this knowledge base makes perfect sense: conditional r_1 expresses that in some given population, choosing randomly an elephant-keeper-pair, we would expect that the elephant likes the keeper with probability 0.6. However, keeper Fred and elephant Clyde are exceptional—mostly, elephants do not like Fred, but Clyde likes (even) Fred.

The example is ambiguous and its formal interpretation via conditional probabilities is hard to grasp. The probabilistic conditional r_1 expresses a belief about the whole population and r_3 clearly expresses individual belief on the specific individuals Clyde and Fred. The probabilistic conditional r_2 seems to represent something in between as it takes both the whole population and the specific individual Fred into account. In many approaches such as BLPs and MLNs relational rules are grounded, and the probability is attached to each instance. For r_1, this means:

$$\mathit{elephant}(\mathsf{a}) \wedge \mathit{keeper}(\mathsf{b}) \quad \rightarrow \quad \mathit{likes}(\mathsf{a},\mathsf{b}) \quad [0.6] \qquad \text{for all } \mathsf{a},\mathsf{b} \in U.$$

Here U is a properly (or arbitrarily) chosen universe. Besides the question, *how* U should be chosen, there are two other problems. First, grounding turns the relational statement r_1 into a collection of statements of the same type as r_3, i. e. statements about individual beliefs. The population aspect gets lost, more precisely: r_1 is no longer a statement describing a generic behavior in a population but is understood as a template that is applicable in precisely the same manner to all individuals. As a consequence, each individual is treated the same way. Secondly, naive grounding techniques make the knowledge base inconsistent, as setting $\mathsf{a} = \mathsf{clyde}$ and $\mathsf{b} = \mathsf{fred}$ renders the above instance contradictory to the beliefs represented in r_3. So, grounding has to take further constraints into account, in order to return a consistent knowledge base, see e. g. (Fisseler, 2010; Loh *et al.*, 2010).

In contrast to the inconsistencies investigated in the previous two chapters the inconsistency expressed in, e. g., the conditionals r_1 and r_2 above is neither accidental nor derives from false or uncertain information. Here, the problem lies not in the dissatisfiability of the conditionals with respect to the semantics outlined above but in the semantics itself. As discussed above, by simply treating a conditional like r_1 as a schema for its instances, information gets lost. In this chapter, we discuss the problem of how to define reasonable semantics for relational conditionals that allows conditionals like r_1, r_2, and r_3 to be consistent. By doing so we take another approach as the one that we pursued in the previous two chapters. There, we accepted the semantics for (propositional) probabilistic conditional logic and enabled reasoning based on this semantics by restoring consistency. Here, we acknowledge inconsistency with respect to naive grounding and develop new semantics that allow for consistent reasoning. In particular, we propose two approaches for giving formal semantics to relational probabilistic knowledge bases that aim at catching properly the commonsense intuition and resolving ambiguities.

This chapter is organized as follows. In Section 5.1 we look a little more deeply at some problems in existing approaches that combine first-order logic and probabilistic reasoning, mainly with respect to non-monotonic reasoning. Afterwards in Section 5.2 we present *relational probabilistic conditional logic* (RPCL), a very simple probabilistic logic that extends propositional probabilistic conditional logic. In Section 5.3 we discuss semantical issues with this logic and propose two different approaches. In Section 5.4 we analyze this logic and investigate properties for the proposed semantics. In Section 5.5 we review related work and in Section 5.6 we conclude this chapter.

The work reported in this chapter is partially joint work with Gabriele Kern-Isberner, cf. *"Publications and Disclaimer"* on page v.

5.1 RELATIONAL PROBABILISTIC MODELS AND NMR

The development of relational probabilistic models has been driven mainly by motivation coming from the field of relational databases and specifically by learning tasks within relational settings. Therefore most approaches allow for easy learning of knowledge bases and provide fast methods for inferences. But when it comes to default reasoning or *non-monotonic reasoning* (NMR) in general, most approaches fail to be applicable.

Consider the omnipresent penguins example, see e. g. (Finthammer and Thimm, 2011). For a relational representation of this example we consider a set of birds and want to be able to state rules about their ability to fly. In particular, we want to state that birds typically fly (with e. g. probability 0.9), that penguins typically do not fly (they fly only with e. g. probability 0.01), and that every penguin is a bird (with a probability of 1). This example shows how well a formalism deals with conflicting information on exceptional individuals. Given a particular bird Tweety for which we have no belief of being a penguin we expect the formalism to derive with a high probability that Tweety does actually fly. Adding the information that Tweety is a penguin the formalism should derive no longer that Tweety flies. Although Tweety is still a bird the more specific information that penguins do not fly shall override the general rule of birds flying. We represent this example using *Bayesian logic programs* as follows, cf. Section 2.4.1.

Example 5.1. Let *bird*/1, *penguin*/1, and *flies*/1 be unary predicates. Then the above rules can be stated as the set $\{c_1, c_2, c_3\}$ of Bayesian clauses with

$$\begin{aligned}
c_1 &=_{def} (\mathit{bird}(\mathsf{X}) \mid \mathit{penguin}(\mathsf{X}))\\
c_2 &=_{def} (\mathit{flies}(\mathsf{X}) \mid \mathit{bird}(\mathsf{X}))\\
c_3 &=_{def} (\mathit{flies}(\mathsf{X}) \mid \mathit{penguin}(\mathsf{X}))
\end{aligned}$$

and their conditional probability distributions $\{\mathsf{cpd}_{c_1}, \mathsf{cpd}_{c_2}, \mathsf{cpd}_{c_3}\}$ by

$$\begin{aligned}
\mathsf{cpd}_{c_1}(\mathsf{true},\mathsf{true}) &=_{def} 1 & \mathsf{cpd}_{c_1}(\mathsf{false},\mathsf{true}) &=_{def} 0\\
\mathsf{cpd}_{c_1}(\mathsf{true},\mathsf{false}) &=_{def} 0.5 & \mathsf{cpd}_{c_1}(\mathsf{false},\mathsf{false}) &=_{def} 0.5\\
\mathsf{cpd}_{c_2}(\mathsf{true},\mathsf{true}) &=_{def} 0.9 & \mathsf{cpd}_{c_2}(\mathsf{false},\mathsf{true}) &=_{def} 0.1\\
\mathsf{cpd}_{c_2}(\mathsf{true},\mathsf{false}) &=_{def} 0.2 & \mathsf{cpd}_{c_2}(\mathsf{false},\mathsf{false}) &=_{def} 0.8\\
\mathsf{cpd}_{c_3}(\mathsf{true},\mathsf{true}) &=_{def} 0.01 & \mathsf{cpd}_{c_3}(\mathsf{false},\mathsf{true}) &=_{def} 0.99\\
\mathsf{cpd}_{c_3}(\mathsf{true},\mathsf{false}) &=_{def} 0.3 & \mathsf{cpd}_{c_3}(\mathsf{false},\mathsf{false}) &=_{def} 0.7 \quad .
\end{aligned}$$

Note that some of the probabilities defined for each conditional probability distribution are somewhat arbitrary. The problem is that defining a probability for a rule, given that its premise is not fulfilled, is a hard task. Consider clause c_2 saying that birds usually fly. But what is the probability of a non-bird flying? It is a serious drawback of Bayesian logic programs (and Bayes nets in general) that they demand a full specification of a conditional probability distribution when complete information is not available.

The specification of the BLP in the previous example lacks one missing piece of information: the combining rules. These play an essential role in defining the inferential semantics for a BLP and different combining rules may yield substantially different results. Consider the query

$$Q_1 =_{def} \quad \mathit{flies}(\mathsf{tweety}) \mid \mathit{bird}(\mathsf{tweety}) \quad ,$$

e. g., the query for the probability of Tweety flying given that Tweety is a bird. Imagine we define *noisy-or* to be the combining function for *flies*. Then the probability of the above query amounts to 0.9 as only clause c_2 is applicable. Now consider the query

$$Q_2 =_{def} \quad \mathit{flies}(\mathsf{tweety}) \mid \mathit{penguin}(\mathsf{tweety}) \quad .$$

Assuming again that *noisy-or* is the combining function for *flies* the probability of this query amounts to 0.901. This result derives from the fact that both clauses c_2 and c_3 are used when determining the probability of Q_2 and combined using *noisy-or*. Obviously, this inference is not intended and one might ask if the choice of the combining rule was inadequate. As a matter of fact using *noisy-and*[1] gives a more appropriate probability of 0.083 in the example above. But now imagine there are penguins that due to some super-hero powers have a high probability of flying. Using a new predicate *superpower*/1 we can state this property via

$$c_4 =_{def} \quad (\mathit{flies}(\mathsf{X}) \mid \mathit{penguin}(\mathsf{X}), \mathit{superpower}(\mathsf{X}))$$

and an adequate conditional probability distribution cpd_{c_4} for c_4 with $\mathsf{cpd}_{c_4}(\mathsf{true}, \mathsf{true}, \mathsf{true}) = 0.8$. Considering the query

$$Q_3 =_{def} \quad \mathit{flies}(\mathsf{tweety}) \mid \mathit{penguin}(\mathsf{tweety}), \mathit{superpower}(\mathsf{tweety})$$

for the BLP containing $\{c_1, \ldots, c_4\}$ this yields a probability of approximately 0.26 (using the combining rule *noisy-and*) that is far too small if we have a super-penguin. Note that the probabilities of the queries Q_1, Q_2 remain the same in the new BLP. In general, the choice of the combining rule depends on whether the combination should strengthen or weaken some probability. But this choice itself is not a matter of numbers but a matter of semantics of the applied clauses and, hence, cannot be determined by the BLP alone. The above discussion applies to other approaches for statistical relational learning that employ similar concepts as combining rules such as *probabilistic relational models* (Getoor *et al.*, 2007), *relational Bayesian networks* (Jaeger, 2002), and *logical Bayesian networks* (Fierens *et al.*, 2004).

Besides the inability to model rules with exceptions adequately, another drawback of BLPs (and similar approaches) has already been mentioned in Example 5.1. In order to obtain a unique model for a BLP the full specification of the conditional probability distribution of each clause is mandatory.

1 For two probabilities p_1, p_2 the noisy-and of p_1 and p_2 is defined to be $p_1 \cdot p_2$.

There exist refinements of BLPs that try to overcome some of their drawbacks. Logical Bayesian networks (Fierens *et al.*, 2004) introduce purely logical statements that are not interpreted in a probabilistic sense. This allows for omitting the specification of unknown probabilistic dependencies such as—in the above example—the probability of an individual flying if the individual is not a bird. Furthermore, the Balios engine (Kersting and Dick, 2004) for Bayesian logic programs allows the specification of logical predicates as well and supports the inclusion of a purely logical specification of background knowledge with Prolog. Still, for both these refinements of BLPs, if one wants to define a probabilistic relationship between two probabilistic variables the conditional probability distribution has to be fully specified.

The discussion above motivates the need for a flexible semantics of relational rules that allows for exceptions and enables default reasoning. In the following we propose relational probabilistic conditional logic as a framework for specifying such semantics.

5.2 syntax of rpcl

We already presented most of the syntactical concepts of relational probabilistic conditional logic (RPCL) in Chapter 2. In particular, in Section 2.1.2 we introduced the language $\mathcal{L}^{\not\forall\not\exists}(\Sigma, V)$, the quantifier-free fragment of first-order logic. Further, we restrict the underlying signature as follows.

Definition 5.1 (Simple relational signature)**.** A first-order signature $\Sigma = (U, Pred, Func)$ is a *simple relational signature* if and only if *Pred* is finite and $Func = \emptyset$.

If Σ is a simple relational signature then for a set of variables V the resulting language $\mathcal{L}^{\not\forall\not\exists}(\Sigma, V)$ is called *simple relational language*. In the following, let $\Sigma = (U, Pred, \emptyset)$ be a simple relational signature with some infinite pool of constants U and let V be a set of variables. Then $(\mathcal{L}^{\not\forall\not\exists}(\Sigma, V) \mid \mathcal{L}^{\not\forall\not\exists}(\Sigma, V))^{pr}$ is the corresponding probabilistic conditional language, cf. Definition 2.24 on page 26. In the following we call $(\mathcal{L}^{\not\forall\not\exists}(\Sigma, V) \mid \mathcal{L}^{\not\forall\not\exists}(\Sigma, V))^{pr}$ a *relational probabilistic conditional language*. A *knowledge base* $\mathcal{R}$ is a subset of $(\mathcal{L}^{\not\forall\not\exists}(\Sigma, V) \mid \mathcal{L}^{\not\forall\not\exists}(\Sigma, V))^{pr}$ and as before probabilistic conditionals of the form $(\psi \mid \top)[d]$ are abbreviated by $(\psi)[d]$.

For a probabilistic conditional $(\psi \mid \phi)[d] \in (\mathcal{L}^{\not\forall\not\exists}(\Sigma, V) \mid \mathcal{L}^{\not\forall\not\exists}(\Sigma, V))^{pr}$ and a knowledge base $\mathcal{R} \in \mathfrak{P}((\mathcal{L}^{\not\forall\not\exists}(\Sigma, V) \mid \mathcal{L}^{\not\forall\not\exists}(\Sigma, V))^{pr})$ we define

$$\begin{aligned}\mathsf{Const}((\psi \mid \phi)[d]) &=_{def} \mathsf{Const}(\psi) \cup \mathsf{Const}(\phi)\\ \mathsf{Const}(\mathcal{R}) &=_{def} \bigcup_{r \in \mathcal{R}} \mathsf{Const}(r) \quad .\end{aligned}$$

Let $D \subseteq U$ be a finite set of constants. We extend the function $\mathsf{gnd}_D(\cdot)$ in a straightforward fashion to conditionals $r = (\psi \,|\, \phi)$ and probabilistic conditionals $r' = (\psi \,|\, \phi)[d]$ via

$$\mathsf{gnd}_D(r) =_{def} \{(\theta(\psi) \,|\, \theta(\phi)) \mid \theta \in \Gamma^{\mathsf{gnd}}(\Sigma) \wedge \mathsf{Im}\,\theta \subseteq D\} \quad \text{and}$$
$$\mathsf{gnd}_D(r') =_{def} \{(\theta(\psi) \,|\, \theta(\phi))[d] \mid \theta \in \Gamma^{\mathsf{gnd}}(\Sigma) \wedge \mathsf{Im}\,\theta \subseteq D\}$$

and to knowledge bases $\mathcal{R}$ via

$$\mathsf{gnd}_D(\mathcal{R}) =_{def} \bigcup_{r \in \mathcal{R}} \mathsf{gnd}_D(r) \quad .$$

If $\{\mathsf{X}_1, \ldots, \mathsf{X}_n\}$ is the set of free (and not bound) variables in ϕ or ψ we also write $(\psi \,|\, \phi)[d]$ as $(\psi(\vec{\mathsf{X}}) \,|\, \phi(\vec{\mathsf{X}}))[d]$ with $\vec{\mathsf{X}} = (\mathsf{X}_1, \ldots, \mathsf{X}_n)$. If $\vec{\mathsf{a}} = (\mathsf{a}_1, \ldots, \mathsf{a}_n)$ is a vector of constants of the same length we denote by $(\psi(\vec{\mathsf{a}}) \,|\, \phi(\vec{\mathsf{a}}))[d]$ the conditional $(\psi' \,|\, \phi')[d]$ that is the same as $(\psi \,|\, \phi)[d]$ but every occurrence of X_i is replaced by a_i for $i = 1, \ldots, n$. If a probabilistic conditional $(\psi \,|\, \phi)[d]$ contains at least one free variable we say that $(\psi \,|\, \phi)[d]$ is an *open* probabilistic conditional.

5.3 SEMANTICS OF RPCL

Introducing relational aspects in probabilistic statements raises some ambiguity on the understanding of these statements. Consider again the example from the beginning of this chapter.

Example 5.2. Let $\mathcal{R}_{zoo} =_{def} \{r_1, r_2, r_3\}$ be the knowledge base given via

$$r_1 =_{def} (\mathit{likes}(\mathsf{X}, \mathsf{Y}) \,|\, \mathit{elephant}(\mathsf{X}) \wedge \mathit{keeper}(\mathsf{Y}))[0.6]$$
$$r_2 =_{def} (\mathit{likes}(\mathsf{X}, \mathsf{fred}) \,|\, \mathit{elephant}(\mathsf{X}) \wedge \mathit{keeper}(\mathsf{fred}))[0.4]$$
$$r_3 =_{def} (\mathit{likes}(\mathsf{clyde}, \mathsf{fred}) \,|\, \mathit{elephant}(\mathsf{clyde}) \wedge \mathit{keeper}(\mathsf{fred}))[0.7] \quad .$$

The knowledge base $\mathcal{R}_{zoo}$ describes the relationships between keepers and elephants in a zoo. The intuitive meaning behind conditional r_1 is that generally some elephant likes a keeper with a probability 0.6. Conditional r_2 says that keeper Fred is exceptional and an elephant likes him only with a probability 0.4 and conditional r_3 states that elephant Clyde likes Fred with a probability of 0.7.

Clearly, the three probabilistic conditionals from Example 5.2 form a hierarchy of specificity from the most general probabilistic conditional r_1 to the most specific probabilistic conditional r_3. As pointed out in Section 5.1 the use of *combining functions* is not adequate for representing non-monotonic reasoning behavior which is exactly the kind of reasoning expected for the knowledge base $\mathcal{R}_{zoo}$. Looking closer at the probabilistic conditionals r_1, r_2, r_3 and their intended meaning it becomes clear that the type of information represented by the conditionals differs. Conditional r_3 contains no variables but gives a probabilistic statement on the relationship of the individuals Clyde and Fred. This means that the *true* state of the relationship is known only with some uncertainty of 0.7 and thus r_3 describes a *degree of belief* on the truth of $\mathit{likes}(\mathsf{clyde}, \mathsf{fred})$. Both conditionals r_1 and r_2 contain free variables and thus describe knowledge ranging over the whole population. When representing a conditional of the form r_2 a statistical interpretation like "*40% of all elephants like Fred*" is applicable that does not allow for universal instantiation if information on specific individuals is present.

In the following we investigate two approaches that give formal semantics to knowledge bases like $\mathcal{R}_{zoo}$. As in Section 2.3 semantics are given to relational probabilistic conditionals by means of probability functions $P : \Omega(\Sigma) \rightarrow [0, 1]$ with $\Omega(\Sigma)$ being the set of Herbrand interpretations of Σ, cf. Definition 2.14 on page 17. Besides the restrictions made to probability functions in Definition 2.15 on page 19 we also demand

1. if $\omega \in \Omega(\Sigma)$ is infinite then $P(\omega) = 0$, and
2. it holds that $P(\omega) \neq 0$ only for finitely many $\omega \in \Omega(\Sigma)$.

We require probability functions to satisfy these two properties in order to avoid technical difficulties in handling infinite sums. As we only consider finite knowledge bases these demands are of no concern regarding the expressivity of our logic.[2] In the following let $\mathcal{P}^{\mathrm{F}}(\Sigma)$ denote the set of all these probability functions. A probability function $P \in \mathcal{P}^{\mathrm{F}}(\Sigma)$ is extended to sentences ϕ (ground formulas) of $\mathcal{L}^{\not\forall\not\exists}(\Sigma, \emptyset)$ via

$$P(\phi) =_{def} \sum_{\omega \in \Omega(\Sigma),\ \omega \models^{\mathrm{F}} \phi} P(\omega) \quad . \tag{5.1}$$

A function P *satisfies* a ground probabilistic conditional $(\psi \,|\, \phi)[d]$, denoted by $P \models^{pr} (\psi \,|\, \phi)[d]$, if and only if

$$P(\psi \mid \phi) = d \quad \text{or} \quad P(\phi) = 0 \quad . \tag{5.2}$$

Note that this relation is identical to the definition of satisfaction for propositional probabilistic conditional logic—see Equation (2.6) on page 28—but does not suffice to give full semantics to relational probabilistic knowledge bases as these may contain open probabilistic conditionals. It remains to define $P \models^{pr} (\psi \,|\, \phi)[d]$ for non-ground conditionals $(\psi \,|\, \phi)[d]$. We address this issue in the following two sections which present the *averaging* and *aggregating semantics*, respectively.

For the upcoming discussion we need some further notation.

2 Remember that every probabilistic conditional is finite as well, cf. Definition 2.24 on page 26.

Definition 5.2 (Relevance set)**.** Let $D \subseteq U$ be finite. Then the *set of relevant Herbrand interpretations* $\Omega(\Sigma, D)$ of $\Omega(\Sigma)$ with respect to D is the set $\Omega(\Sigma, D) \subseteq \Omega(\Sigma)$ defined as $\Omega(\Sigma, D) =_{def} \{\omega \in \Omega(\Sigma) \mid \mathsf{Const}(\omega) \subseteq D\}$.

Clearly, it holds that $\Omega(\Sigma, D) = \Omega((D, Pred, \emptyset))$ if $\Sigma = (U, Pred, \emptyset)$. Note, that the set of relevant Herbrand interpretations is the semantical counterpart to the grounding function $\mathsf{gnd}_D(\cdot)$.

5.3.1 *Averaging Semantics*

Example 5.2 on page 134 showed that, in general, universal instantiation of an open probabilistic conditional $(\psi \mid \phi)[d]$ does not yield an equivalent representation of the intended meaning of $(\psi \mid \phi)[d]$. More precisely, it demands that every instantiation *inherits* the probability d which is not adequate in the context of exceptional individuals. Moreover, having more specific information on specific instantiations should not render the knowledge base inconsistent as other instantiations might balance out exceptions. Consider the following example from statistics.

Example 5.3. Imagine a table with ten balls on it and every ball is covered by a drapery. It is certain knowledge that nine of these balls are blue and one is red. Taking a look under the first drapery we discover a blue ball. What is the probability of a blue ball being under the second drapery? Let $D =_{def} \{\mathsf{c}_1, \ldots, \mathsf{c}_{10}\}$ denote the objects discovered under the draperies when lifting the draperies in an arbitrary order. We represent the scenario after we discovered that the ball under the first drapery is blue. Let $\mathcal{R}_{dr} =_{def} \{r_{1,1}, \ldots, r_{1,10}, r_2, r_3\}$ be given via

$$r_{1,1} =_{def} (ball(\mathsf{c}_1))[1] \quad \ldots \quad r_{1,10} =_{def} (ball(\mathsf{c}_{10}))[1]$$
$$r_2 =_{def} (blue(\mathsf{X}) \mid ball(\mathsf{X}))[0.9]$$
$$r_3 =_{def} (blue(\mathsf{c}_1))[1]$$

The conditionals $r_{1,1}, \ldots, r_{1,10}$ represent the facts that there is a ball under each drapery. Conditional r_2 says that a ball is blue with probability 0.9 and conditional r_3 states that the object discovered when lifting the first drapery is blue. Clearly, the probability of discovering a blue object under the second drapery is $8/9$, given that the object is a ball, and any reasonable semantics for relational conditionals should allow both $\mathcal{R}_{dr}$ and $\mathcal{R}_{dr} \cup \{(blue(\mathsf{c}_2))[8/9]\}$ to be at least satisfiable.

In the previous example conditional r_2 defined an *expected value* for the probability of discovering a blue ball. The additional information that a blue ball has already been discovered changes the expected value of discovering another blue ball correspondingly. Therefore mutual influences of different conditionals have to be taken into account when defining meaning to a knowledge base.

The approach of *averaging semantics* (Thimm, 2009b; Kern-Isberner and Thimm, 2010) generalizes the above intuition by interpreting open probabilistic conditionals of the form $(\psi(\vec{\mathsf{X}}) \mid \phi(\vec{\mathsf{X}}))[d]$ to describe an expected value on the probability of $(\psi(\vec{\mathsf{a}}) \mid \phi(\vec{\mathsf{a}}))$ for some randomly chosen $\vec{\mathsf{a}}$ in some given adequately large but finite domain $D \subseteq U$. Thus, given the actual probabilities of $(\psi(\vec{\mathsf{a}}) \mid \phi(\vec{\mathsf{a}}))$ for each possible instantiation $\vec{\mathsf{a}}$ we expect the *average* of these probabilities should match d. For a probability function $P \in \mathcal{P}^{\mathrm{F}}(\Sigma)$ and a conditional $(\psi(\vec{\mathsf{X}}) \mid \phi(\vec{\mathsf{X}}))$ we abbreviate with

$$\mathsf{gnd}_D^P(\psi(\vec{\mathsf{X}}) \mid \phi(\vec{\mathsf{X}})) =_{def} \{(\psi' \mid \phi') \in \mathsf{gnd}_D(\psi(\vec{\mathsf{X}}) \mid \phi(\vec{\mathsf{X}}))) \mid P(\phi') > 0\}$$

the set of ground instances of $(\psi(\vec{\mathsf{X}}) \mid \phi(\vec{\mathsf{X}}))$ for which the premise has a non-zero probability. Then P $\varnothing$-satisfies a probabilistic conditional $r = (\psi(\vec{\mathsf{X}}) \mid \phi(\vec{\mathsf{X}}))[d]$ with respect to finite D with $\mathsf{Const}(r) \subseteq D \subseteq U$ and $D \neq \emptyset$, denoted by $P, D \models^{pr}_{\varnothing} (\psi(\vec{\mathsf{X}}) \mid \phi(\vec{\mathsf{X}}))[d]$, if and only if the following two conditions are satisfied:

1. $P(\omega) = 0$ for every $\omega \notin \Omega(\Sigma, D)$, and
2. it holds that either $|\mathsf{gnd}_D^P(\psi(\vec{\mathsf{X}}) \mid \phi(\vec{\mathsf{X}}))| = 0$ or

$$\frac{\sum\limits_{(\psi(\vec{\mathsf{a}}) \mid \phi(\vec{\mathsf{a}})) \in \mathsf{gnd}_D^P((\psi(\vec{\mathsf{X}}) \mid \phi(\vec{\mathsf{X}})))} P(\psi(\vec{\mathsf{a}}) \mid \phi(\vec{\mathsf{a}}))}{|\mathsf{gnd}_D^P(\psi(\vec{\mathsf{X}}) \mid \phi(\vec{\mathsf{X}}))|} = d \quad . \tag{5.3}$$

The first condition is merely a technical amenity. By requiring 1.) we can focus on probability functions that only take those Herbrand interpretations into account that do not mention elements outside the domain under discourse. We come back to this issue in Chapter 6. The interpretation behind requirement 2.) is that a probability function P $\varnothing$-satisfies a probabilistic conditional $(\psi(\vec{\mathsf{X}}) \mid \phi(\vec{\mathsf{X}}))[d]$ if the average of probabilities of the individual instances of $(\psi(\vec{\mathsf{X}}) \mid \phi(\vec{\mathsf{X}}))[d]$ is d. In Equation (5.3), the numerator of the fraction on the left-hand side sums up the conditional probabilities of the different instances of $(\psi(\vec{\mathsf{X}}) \mid \phi(\vec{\mathsf{X}}))[d]$ for which the premise has non-zero probability. Note, that we use $\mathsf{gnd}_D^P(\cdot)$ instead of $\mathsf{gnd}_D(\cdot)$ as $P(\psi(\vec{\mathsf{a}}) \mid \phi(\vec{\mathsf{a}}))$ is not defined for $(\psi(\vec{\mathsf{a}}) \mid \phi(\vec{\mathsf{a}})) \in \mathsf{gnd}_D((\psi(\vec{\mathsf{X}}) \mid \phi(\vec{\mathsf{X}}))) \setminus \mathsf{gnd}_D^P((\psi(\vec{\mathsf{X}}) \mid \phi(\vec{\mathsf{X}})))$ due to $P(\phi(\vec{\mathsf{a}})) = 0$. Therefore, by considering only ground instances in $\mathsf{gnd}_D^P((\psi(\vec{\mathsf{X}}) \mid \phi(\vec{\mathsf{X}})))$ we average only the probabilities of ground instances that are *relevant* for the open probabilistic conditional.

Remark 5.1. Let $r = (\psi \mid \phi)[d]$ be a ground conditional and P be a probability function. Then (5.3) becomes

$$\frac{P(\psi \mid \phi)}{|\mathsf{gnd}_D^P(\psi \mid \phi)|} = d \quad . \tag{5.4}$$

It follows that P $\varnothing$-satisfies r if either $|\mathsf{gnd}_D^P(\psi \mid \phi)| = 0$ or $P(\psi \mid \phi) = d$. Notice, that the first statement is equivalent to $P(\phi) = 0$. It follows that

ω	$P(\omega)$	ω	$P(\omega)$
$\emptyset$	0	$\{a(\mathsf{a}_2), b(\mathsf{a}_1)\}$	0
$\{a(\mathsf{a}_1)\}$	0	$\{a(\mathsf{a}_2), b(\mathsf{a}_2)\}$	0
$\{a(\mathsf{a}_2)\}$	0	$\{b(\mathsf{a}_1), b(\mathsf{a}_2)\}$	0
$\{b(\mathsf{a}_1)\}$	0	$\{a(\mathsf{a}_1), a(\mathsf{a}_2), b(\mathsf{a}_1)\}$	0.1
$\{b(\mathsf{a}_2)\}$	0	$\{a(\mathsf{a}_1), a(\mathsf{a}_2), b(\mathsf{a}_2)\}$	0.3
$\{a(\mathsf{a}_1), a(\mathsf{a}_2)\}$	0.1	$\{a(\mathsf{a}_1), b(\mathsf{a}_2), b(\mathsf{a}_2)\}$	0
$\{a(\mathsf{a}_1), b(\mathsf{a}_2)\}$	0	$\{a(\mathsf{a}_2), b(\mathsf{a}_1), b(\mathsf{a}_2)\}$	0
$\{a(\mathsf{a}_1), b(\mathsf{a}_2)\}$	0	$\{a(\mathsf{a}_1), a(\mathsf{a}_2), b(\mathsf{a}_1), b(\mathsf{a}_2)\}$	0.5

Table 6: The probability function P for the knowledge base $\mathcal{R}$ in Example 5.4 (all Herbrand interpretations omitted have probability 0)

$\models^{pr}_{\emptyset}$ is a clear generalization of the standard probabilistic semantics for propositional probabilistic conditional logic, cf. Equation (5.2) on page 135.

As before, a probability function P $\emptyset$-satisfies a knowledge base $\mathcal{R}$ with respect to finite D with $\mathsf{Const}(\mathcal{R}) \subseteq D \subseteq U$ and $D \neq \emptyset$, denoted $P, D \models^{pr}_{\emptyset} \mathcal{R}$, if P $\emptyset$-satisfies every probabilistic conditional $r \in \mathcal{R}$ with respect to D. We say that $\mathcal{R}$ is $\emptyset$-consistent with respect to finite $D \subseteq U$ if and only if there is at least one P with $P, D \models^{pr}_{\emptyset} \mathcal{R}$, otherwise $\mathcal{R}$ is $\emptyset$-inconsistent with respect to D. Two knowledge bases $\mathcal{R}_1$ and $\mathcal{R}_2$ are $\emptyset$-equivalent, denoted by $\mathcal{R}_1 \equiv^{\emptyset} \mathcal{R}_3$, if and only if for every P and D it holds that $P, D \models^{pr}_{\emptyset} \mathcal{R}_1$ whenever $P, D \models^{pr}_{\emptyset} \mathcal{R}_2$.

Example 5.4. Consider the knowledge base $\mathcal{R} =_{def} \{r_1, r_2, r_3\}$ with

$$r_1 =_{def} (b(\mathsf{X}) \mid a(\mathsf{X}))[0.7] \qquad r_2 =_{def} (a(\mathsf{X}))[1] \qquad r_3 =_{def} (b(\mathsf{a}_1))[0.6]$$

and let $D =_{def} \{\mathsf{a}_1, \mathsf{a}_2\}$. Consider the probability function P given in Table 6. As one can see, it holds that $P, D \models^{pr}_{\emptyset} \mathcal{R}$:

- it holds that $P, D \models^{pr}_{\emptyset} r_1$ as $P(b(\mathsf{a}_1) \mid a(\mathsf{a}_1)) = {}^{P(a(\mathsf{a}_1)b(\mathsf{a}_1))}/_{P(a(\mathsf{a}_1)} = {}^{0.6}/_{1} = 0.6$ and $P(b(\mathsf{a}_2) \mid a(\mathsf{a}_2)) = {}^{P(a(\mathsf{a}_2)b(\mathsf{a}_2))}/_{P(a(\mathsf{a}_2)} = {}^{0.8}/_{1} = 0.8$ and hence ${}^{(0.8+0.6)}/_{2} = 0.7$,
- it holds that $P, D \models^{pr}_{\emptyset} r_2$ as ${}^{(P(a(\mathsf{a}_1)) + P(a(\mathsf{a}_2)))}/_{2} = 1$, and
- it holds that $P, D \models^{pr}_{\emptyset} r_3$ as $P(b(\mathsf{a}_1)) = 0.6$.

Example 5.5. We continue Example 5.3. Let P be some probability function that satisfies

$$P(ball(\mathsf{c}_1)) = \ldots = P(ball(\mathsf{c}_{10})) = 1$$
$$P(blue(\mathsf{c}_1)) = 1$$
$$P(blue(\mathsf{c}_2)) = \ldots = P(blue(\mathsf{c}_{10})) = {}^{8}/_{9}$$

and also $P(\omega) = 0$ for every $\omega \notin \Omega(\Sigma, D)$ (it is easy to see that such a probability function exists). The function P satisfies $r_{1,1}, \ldots, r_{1,10}, r_3$ with respect to D by definition. Furthermore, P satisfies r_2 with respect to D due to

$$P(\mathit{blue}(\mathsf{c}_1) \,|\, \mathit{ball}(\mathsf{c}_1)) = 1$$
$$P(\mathit{blue}(\mathsf{c}_2) \,|\, \mathit{ball}(\mathsf{c}_2)) = \ldots = P(\mathit{blue}(\mathsf{c}_{10}) \,|\, \mathit{ball}(\mathsf{c}_{10})) = 8/9$$

and

$$\frac{1 + 9\frac{8}{9}}{10} = 0.9 \quad .$$

This implies $P, D \models^{pr}_{\varnothing} \mathcal{R}_{dr}$. Furthermore, $P, D \models^{pr}_{\varnothing} \mathcal{R}_{dr} \cup \{(\mathit{blue}(\mathsf{c}_2))[8/9]\}$ as well by definition, and therefore $\mathcal{R}_{dr} \cup \{(\mathit{blue}(\mathsf{c}_2))[8/9]\}$ is satisfiable with respect to averaging semantics.

Example 5.6. We continue Example 5.2 from page 134. Let $D =_{def} \{\mathsf{fred}, \mathsf{clyde}, \mathsf{dumbo}, \mathsf{dave}\}$ and let P be a probability function that satisfies

$$P(\mathit{elephant}(\mathsf{dumbo})) = P(\mathit{elephant}(\mathsf{clyde})) = 1$$
$$P(\mathit{elephant}(\mathsf{fred})) = P(\mathit{elephant}(\mathsf{dave})) = 0$$
$$P(\mathit{keeper}(\mathsf{dumbo})) = P(\mathit{keeper}(\mathsf{clyde})) = 0$$
$$P(\mathit{keeper}(\mathsf{fred})) = P(\mathit{keeper}(\mathsf{dave})) = 1$$
$$P(\mathit{likes}(\mathsf{clyde}, \mathsf{fred})) = 0.7$$
$$P(\mathit{likes}(\mathsf{dumbo}, \mathsf{fred})) = 0.1$$
$$P(\mathit{likes}(\mathsf{clyde}, \mathsf{dave})) = P(\mathit{likes}(\mathsf{dumbo}, \mathsf{dave})) = 0.8 \quad .$$

It follows that $P, D \models^{pr}_{\varnothing} r_3$ by definition. For r_2 it holds that

$$\begin{aligned}
&\mathsf{gnd}_D((\mathit{likes}(\mathsf{X}, \mathsf{fred}) \,|\, \mathit{elephant}(\mathsf{X}) \wedge \mathit{keeper}(\mathsf{fred}))) = \{ \\
&\quad r_2^1 =_{def} \quad (\mathit{likes}(\mathsf{clyde}, \mathsf{fred}) \,|\, \mathit{elephant}(\mathsf{clyde}) \wedge \mathit{keeper}(\mathsf{fred})), \\
&\quad r_2^2 =_{def} \quad (\mathit{likes}(\mathsf{dumbo}, \mathsf{fred}) \,|\, \mathit{elephant}(\mathsf{dumbo}) \wedge \mathit{keeper}(\mathsf{fred})), \\
&\quad r_2^3 =_{def} \quad (\mathit{likes}(\mathsf{dave}, \mathsf{fred}) \,|\, \mathit{elephant}(\mathsf{dave}) \wedge \mathit{keeper}(\mathsf{fred})), \\
&\quad r_2^4 =_{def} \quad (\mathit{likes}(\mathsf{fred}, \mathsf{fred}) \,|\, \mathit{elephant}(\mathsf{fred}) \wedge \mathit{keeper}(\mathsf{fred})) \quad \}
\end{aligned}$$

and $\mathsf{gnd}^P_D((\mathit{likes}(\mathsf{X}, \mathsf{fred}) \,|\, \mathit{elephant}(\mathsf{X}) \wedge \mathit{keeper}(\mathsf{fred}))) = \{r_2^1, r_2^2\}$ as for r_2^3 and r_2^4 it holds that $\mathit{elephant}(\mathsf{dave}) \wedge \mathit{keeper}(\mathsf{fred})$ and $\mathit{elephant}(\mathsf{fred}) \wedge \mathit{keeper}(\mathsf{fred})$ have probability zero, respectively. It follows $P, D \models^{pr}_{\varnothing} r_2$ due to

$$\frac{P(r_2^1) + P(r_2^2)}{|\mathsf{gnd}^P_D((\mathit{likes}(\mathsf{X}, \mathsf{fred}) \,|\, \mathit{elephant}(\mathsf{X}) \wedge \mathit{keeper}(\mathsf{fred})))|} = \frac{0.7 + 0.1}{2} = 0.4 \, .$$

Furthermore, for r_1 it holds that

$$\begin{aligned}
&\mathsf{gnd}^P_D((likes(\mathsf{X},\mathsf{Y}) \,|\, elephant(\mathsf{X}) \wedge keeper(\mathsf{Y}))) = \{ \\
&\quad r_1^1 =_{def} \quad (likes(\mathsf{clyde},\mathsf{fred}) \,|\, elephant(\mathsf{clyde}) \wedge keeper(\mathsf{fred})), \\
&\quad r_1^2 =_{def} \quad (likes(\mathsf{dumbo},\mathsf{fred}) \,|\, elephant(\mathsf{dumbo}) \wedge keeper(\mathsf{fred})), \\
&\quad r_1^3 =_{def} \quad (likes(\mathsf{clyde},\mathsf{dave}) \,|\, elephant(\mathsf{clyde}) \wedge keeper(\mathsf{dave})), \\
&\quad r_1^4 =_{def} \quad (likes(\mathsf{dumbo},\mathsf{dave}) \,|\, elephant(\mathsf{dumbo}) \wedge keeper(\mathsf{dave})) \quad \} .
\end{aligned}$$

Similarly it follows $P, D \models^{pr}_{\varnothing} r_1$ as

$$\begin{aligned}
&\frac{P(r_1^1) + P(r_1^2) + P(r_1^3) + P(r_1^4)}{|\mathsf{gnd}^P_D((likes(\mathsf{X},\mathsf{Y}) \,|\, elephant(\mathsf{X}) \wedge keeper(\mathsf{Y})))|} \\
=&\frac{0.7 + 0.1 + 0.8 + 0.8}{4} = 0.6 \quad .
\end{aligned}$$

5.3.2 *Aggregating Semantics*

Averaging semantics preserves the interpretation of a conditional probability as subjective belief in the conclusion given the premise holds. Therefore conditional probabilities are only defined for ground conditionals and the probability of an open conditional $(\psi(\vec{\mathsf{X}}) \,|\, \phi(\vec{\mathsf{X}}))$ is defined in terms of conditional probabilities of its instances. When considering a relational language, one might argue whether a conditional should be interpreted in this manner or whether conditional probability should be redefined on a higher level, incorporating the relational structure of the language. In the following we give a novel approach for defining conditional probabilities in a relational setting.[3]

We first consider unconditioned formulas $(\phi(\vec{\mathsf{X}}))[d]$ with free (and not bound) variables $\vec{\mathsf{X}}$. Let ω be some Herbrand interpretation and D with $\mathsf{Const}(\phi(\vec{\mathsf{X}})) \subseteq D \subseteq U$ and $D \neq \emptyset$ the finite set of constants under consideration. Treating ω as a statistical sample we can count the number of true instances of $\phi(\vec{\mathsf{X}})$ in ω under D and determine the average number of true instances via

$$f^D_\omega(\phi(\vec{\mathsf{X}})) =_{def} \frac{|\{\phi(\vec{\mathsf{a}}) \,|\, \phi(\vec{\mathsf{a}}) \in \mathsf{gnd}_D(\phi(\vec{\mathsf{X}})) \mid \omega \models^{\mathrm{F}} \phi(\vec{\mathsf{a}})\}|}{|\mathsf{gnd}_D(\phi(\vec{\mathsf{X}}))|} \quad . \tag{5.5}$$

Note that $|\mathsf{gnd}_D(\phi(\vec{\mathsf{X}}))| > 0$ for every $\phi(\vec{\mathsf{X}})$ as $D \neq \emptyset$. For ground $\phi(\vec{\mathsf{a}})$ we get

$$f^D_\omega(\phi(\vec{\mathsf{a}})) = \begin{cases} 1 & \text{if } \omega \models^{\mathrm{F}} \phi(\vec{\mathsf{a}}) \\ 0 & \text{if } \omega \not\models^{\mathrm{F}} \phi(\vec{\mathsf{a}}) \end{cases}$$

3 The following elaboration (until right after Equation 5.8) is due to Gabriele Kern-Isberner and stems from personal communication.

and then (5.5) amounts to

$$f_\omega^D(\phi(\vec{\mathsf{X}})) = \frac{\sum_{\phi(\vec{\mathsf{a}})\in \mathsf{gnd}_D(\phi(\vec{\mathsf{X}}))} f_\omega^D(\phi(\vec{\mathsf{a}}))}{|\mathsf{gnd}_D(\phi(\vec{\mathsf{X}}))|} \quad .$$

Considering some probability function $P \in \mathcal{P}^{\mathrm{F}}(\Sigma)$ (thus describing either a series of samples or subjective beliefs in each interpretation being the actual world) we can appropriately define

$$P(\phi(\vec{\mathsf{X}}); D) =_{def} \sum_{\omega\in\Omega(\Sigma)} f_\omega^D(\phi(\vec{\mathsf{X}}))P(\omega) \tag{5.6}$$

to be the weighted sum of the average frequencies. Rearranging (5.6) yields

$$\begin{aligned}
P(\phi(\vec{\mathsf{X}}); D) &= \sum_{\omega\in\Omega(\Sigma)} f_\omega^D(\phi(\vec{\mathsf{X}}))P(\omega) \\
&= \sum_{\omega\in\Omega(\Sigma)} \frac{\sum_{\phi(\vec{\mathsf{a}})\in \mathsf{gnd}_D(\phi(\vec{\mathsf{X}}))} f_\omega^D(\phi(\vec{\mathsf{a}}))}{|\mathsf{gnd}_D(\phi(\vec{\mathsf{X}}))|} P(\omega) \\
&= \frac{\sum_{\omega\in\Omega(\Sigma)} \sum_{\phi(\vec{\mathsf{a}})\in \mathsf{gnd}_D(\phi(\vec{\mathsf{X}}))} f_\omega^D(\phi(\vec{\mathsf{a}}))P(\omega)}{|\mathsf{gnd}_D(\phi(\vec{\mathsf{X}}))|} \\
&= \frac{\sum_{\phi(\vec{\mathsf{a}})\in \mathsf{gnd}_D(\phi(\vec{\mathsf{X}}))} \sum_{\omega\in\Omega(\Sigma)} f_\omega^D(\phi(\vec{\mathsf{a}}))P(\omega)}{|\mathsf{gnd}_D(\phi(\vec{\mathsf{X}}))|} \\
&= \frac{\sum_{\phi(\vec{\mathsf{a}})\in \mathsf{gnd}_D(\phi(\vec{\mathsf{X}}))} P(\phi(\vec{\mathsf{a}}))}{|\mathsf{gnd}_D(\phi(\vec{\mathsf{X}}))|}
\end{aligned} \tag{5.7}$$

and thus also a statistical justification for (5.3) on page 137 for the case of unconditioned formulas. But instead of applying (5.7) in the same way to conditionals we give a new definition of the conditional probability via

$$P(\psi(\vec{\mathsf{X}}) \,|\, \phi(\vec{\mathsf{X}}); D) = \frac{P(\psi(\vec{\mathsf{X}})\phi(\vec{\mathsf{X}}); D)}{P(\phi(\vec{\mathsf{X}}); D)} \quad \text{if} \quad P(\phi(\vec{\mathsf{X}}); D) > 0 \tag{5.8}$$

thus carrying over the definition of conditional probability to a relational setting. By applying (5.7) and assuming $P(\phi(\vec{\mathsf{X}}); D) > 0$ this yields

$$P(\psi(\vec{\mathsf{X}}) \,|\, \phi(\vec{\mathsf{X}}); D) = \\
\frac{|\mathsf{gnd}_D(\phi(\vec{\mathsf{X}}))| \sum_{(\psi(\vec{\mathsf{a}})\phi(\vec{\mathsf{a}}))\in \mathsf{gnd}_D(\phi(\vec{\mathsf{X}})\psi(\vec{\mathsf{X}}))} P(\psi(\vec{\mathsf{a}})\phi(\vec{\mathsf{a}}))}{|\mathsf{gnd}_D(\phi(\vec{\mathsf{X}})\psi(\vec{\mathsf{X}}))| \sum_{\phi(\vec{\mathsf{a}})\in \mathsf{gnd}_D(\phi(\vec{\mathsf{X}}))} P(\phi(\vec{\mathsf{a}}))} \quad .$$

Note that the equation $|\mathsf{gnd}_D(\phi(\vec{\mathsf{X}})\psi(\vec{\mathsf{X}}))| = |\mathsf{gnd}_D(\phi(\vec{\mathsf{X}}))|$ does not hold in general. For example, consider the conditional $r = (p(\mathsf{X},\mathsf{Y}) \,|\, q(\mathsf{X}))$ with

$$\begin{aligned}\mathsf{gnd}_{\{\mathsf{c}_1,\mathsf{c}_2\}}(p(\mathsf{X},\mathsf{Y})q(\mathsf{X})) &= \{p(\mathsf{c}_1,\mathsf{c}_1)q(\mathsf{c}_1), p(\mathsf{c}_1,\mathsf{c}_2)q(\mathsf{c}_1),\\ &\qquad p(\mathsf{c}_2,\mathsf{c}_1)q(\mathsf{c}_2), p(\mathsf{c}_2,\mathsf{c}_2)q(\mathsf{c}_2)\}\\ \mathsf{gnd}_{\{\mathsf{c}_1,\mathsf{c}_2\}}(q(\mathsf{X})) &= \{q(\mathsf{c}_1), q(\mathsf{c}_2)\}\end{aligned}$$

However, if every variable that occurs in $\psi(\vec{\mathsf{X}})$ also appears in $\phi(\vec{\mathsf{X}})$ then we get $|\mathsf{gnd}_D(\phi(\vec{\mathsf{X}})\psi(\vec{\mathsf{X}}))| = |\mathsf{gnd}_D(\phi(\vec{\mathsf{X}}))|$ and accordingly

$$P(\psi(\vec{\mathsf{X}}) \,|\, \phi(\vec{\mathsf{X}}); D) = \frac{\sum_{(\psi(\vec{\mathsf{a}})\phi(\vec{\mathsf{a}}))\in\mathsf{gnd}_D(\phi(\vec{\mathsf{X}})\psi(\vec{V\mathsf{X}}))} P(\psi(\vec{\mathsf{a}})\phi(\vec{\mathsf{a}}))}{\sum_{\phi(\vec{\mathsf{a}})\in\mathsf{gnd}_D(\phi(\vec{\mathsf{X}}))} P(\phi(\vec{\mathsf{a}}))} \quad . \tag{5.9}$$

A conditional that satisfies the above property is also called *safe* or *range-restricted*, cf. e. g. (Kersting and Raedt, 2007). In the following, we are only dealing with range-restricted conditionals and, consequently, we use the simple form (5.9) for discussion.

We are now able to define the *aggregating semantics* as follows. A probability function $P \in \mathcal{P}^{\mathrm{F}}(\Sigma)$ $\odot$-satisfies $r = (\psi(\vec{\mathsf{X}}) \,|\, \phi(\vec{\mathsf{X}}))[d]$ with respect to finite D with $\mathsf{Const}(r) \subseteq D \subseteq U$ and $D \neq \emptyset$, denoted by $P, D \models^{pr}_{\odot} (\psi(\vec{\mathsf{X}}) \,|\, \phi(\vec{\mathsf{X}}))[d]$ if and only if the following two conditions are satisfied:

1. $P(\omega) = 0$ for every $\omega \notin \Omega(\Sigma, D)$, and
2. it holds that

$$P(\phi(\vec{\mathsf{X}}); D) = 0 \quad \text{or} \quad P(\psi(\vec{\mathsf{X}}) \,|\, \phi(\vec{\mathsf{X}}); D) = d \quad . \tag{5.10}$$

As for the averaging semantics, condition 1.) is a technical amenity. Note that this definition nicely resembles the satisfaction relation in the propositional case, cf. Equation (5.2) on page 135.

Remark 5.2. As for $\models^{pr}_{\varnothing}$, the relation $\models^{pr}_{\odot}$ coincides with the standard probabilistic semantics for ground probabilistic conditionals. Let $r = (\psi \,|\, \phi)[d]$ be a ground probabilistic conditional and P be a probability function. Then (5.8) becomes

$$P(\psi \,|\, \phi; D) = \frac{P(\psi\phi; D)}{P(\phi; D)} = \frac{\frac{P(\psi\phi)}{1}}{\frac{P(\phi)}{1}} = \frac{P(\psi\phi)}{P(\phi)} \quad .$$

It follows that P $\odot$-satisfies r if either $P(\phi; D) = 0$ or $P(\psi \,|\, \phi) = d$. Notice, that the first statement is equivalent to $P(\phi) = 0$. It follows that $\models^{pr}_{\odot}$ is a clear generalization of the standard probabilistic semantics for propositional probabilistic conditional logic, cf. Equation (5.2) on page 135.

As before, a probability function P $\odot$-satisfies a knowledge base $\mathcal{R}$ with respect to finite D with $\mathsf{Const}(\mathcal{R}) \subseteq D \subseteq U$ and $D \neq \emptyset$, denoted by

$P, D \models^{pr}_{\odot} \mathcal{R}$, if P $\odot$-satisfies every probabilistic conditional $r \in \mathcal{R}$ with respect to D. We say that $\mathcal{R}$ is $\odot$-consistent with respect to D if and only if there is at least one P with $P, D \models^{pr}_{\odot} \mathcal{R}$, otherwise $\mathcal{R}$ is $\odot$-inconsistent with respect to D. Two knowledge bases $\mathcal{R}_1$ and $\mathcal{R}_2$ are $\odot$-equivalent, denoted by $\mathcal{R}_1 \equiv^{\odot} \mathcal{R}_2$, if and only if for every P and D it holds that $P, D \models^{pr}_{\odot} \mathcal{R}_1$ whenever $P, D \models^{pr}_{\odot} \mathcal{R}_2$.

Example 5.7. Consider the knowledge base $\mathcal{R} = \{r_1, r_2, r_3\}$ from Example 5.4 on page 138 with

$$r_1 = (b(\mathsf{X}) \,|\, a(\mathsf{X}))[0.7] \qquad r_2 = (a(\mathsf{X}))[1] \qquad r_3 = (b(\mathsf{a}_1))[0.6]$$

and $D = \{\mathsf{a}_1, \mathsf{a}_1\}$. As in Example 5.4 consider the probability function P given in Table 6 (see page 138). As one can see, it holds that $P, D \models^{pr}_{\odot} \mathcal{R}$ as well:

- it holds that $P \models^{pr}_{\odot} r_1$ as $P(b(\mathsf{a}_1)a(\mathsf{a}_1)) = 0.6$ and $P(b(\mathsf{a}_2)a(\mathsf{a}_2))) = 0.8$, and as well $P(a(\mathsf{a}_1)) = P(a(\mathsf{a}_2)) = 1$; hence $(0.8 + 0.6)/2 = 0.7$;
- it holds that $P \models^{pr}_{\odot} r_2$ as $(P(a(\mathsf{a}_1)) + P(a(\mathsf{a}_2)))/2 = 1$, and
- it holds that $P \models^{pr}_{\odot} r_3$ as $P(b(\mathsf{a}_1)) = 0.6$.

Examples 5.4 and 5.7 show that both proposed semantics coincide on the given simple knowledge base. We investigate the similarities and differences of the both semantics further in the next section.

5.4 PROPERTIES AND ANALYSIS

Due to Remark 5.1 on page 137 and Remark 5.2 on page 142 both semantics agree on ground conditionals. Furthermore, it is straightforward to show that $\models^{pr}_{\varnothing}$ and $\models^{pr}_{\odot}$ also agree on probabilistic facts (that may contain variables).

Proposition 5.1. *Let $P \in \mathcal{P}^F(\Sigma)$ be a probability function, $D \subseteq U$ finite, and $(\psi)[d]$ a probabilistic fact. Then it holds that $P, D \models^{pr}_{\varnothing} (\psi)[d]$ if and only if $P, D \models^{pr}_{\odot} (\psi)[d]$.*

Proof. It holds that $P, D \models^{pr}_{\varnothing} (\psi)[d]$ if and only if

$$\frac{\sum_{\psi' \in \mathsf{gnd}_D(\psi)} P(\psi')}{|\mathsf{gnd}_D(\psi)|} = d$$

as the premise of a probabilistic fact, i. e. a tautology, has probability greater zero. In particular, due to Equation 5.7 (see page 141) it holds that

$$P, D \models^{pr}_{\odot} (\psi)[d] \quad \textit{iff} \quad \frac{\sum_{\psi' \in \mathsf{gnd}_D(\psi)} P(\psi')}{|\mathsf{gnd}_D(\psi)|} = d \qquad \square$$

In a straightforward way one can also accept the following statement.

Proposition 5.2. *Let* $r = (\psi \mid \phi)$ *be a conditional and* $\mathsf{gnd}_D(r) = \{(\psi_1 \mid \phi_1), \ldots, (\psi_n \mid \phi_n)\}$ *be the set of its ground instances with respect to a finite set* D *with* $\mathsf{Const}(r) \subseteq D \subseteq U$. *Let* P *be a probability function. If* $P(\psi_1 \mid \phi_1) = \ldots = P(\psi_n \mid \phi_n) = d$ *then both* $P, D \models^{pr}_{\odot} (\psi \mid \phi)[d]$ *and* $P, D \models^{pr}_{\varnothing} (\psi \mid \phi)[d]$.

Proof. Let it hold that $P, D \models^{pr}_{\odot} (\psi \mid \phi)[d']$. As $P(\psi_i \phi_i) = d \cdot P(\phi_i)$ for every $i = 1, \ldots, n$ and by considering Equation 5.6 (see page 141) the value d' amounts to

$$\begin{aligned} d' &= \frac{P(\psi_1 \phi_1) + \ldots + P(\psi_n \phi_n)}{P(\phi_1) + \ldots + P(\phi_n)} \\ &= \frac{d \cdot P(\phi_1) + \ldots + d \cdot P(\phi_n)}{P(\phi_1) + \ldots + P(\phi_n)} = d \quad . \end{aligned}$$

By considering Equation 5.3 (see page 137) and $P, D \models^{pr}_{\varnothing} (\psi \mid \phi)[d']$ we also get

$$\begin{aligned} d' &= \frac{P(\psi_1 \mid \phi_1) + \ldots + P(\psi_n \mid \phi_n)}{n} \\ &= \frac{nd}{n} = d \quad . \end{aligned}$$

Note that $P(\phi_i) > 0$ as $P(\psi_i \mid \phi_i) = d$ is (implicitly) well-defined for $i = 1, \ldots, n$. ☐

For general probabilistic conditionals, however, the averaging and aggregating semantics turn out to be different, as the following example shows.

Example 5.8. Let $a/1$ and $b/1$ be two predicates, let $D =_{def} \{\mathsf{a}_1, \ldots, \mathsf{a}_5\}$, and consider the knowledge base $\mathcal{R}$ given via

$$\begin{aligned} \mathcal{R} =_{def} \{ \quad & (a(\mathsf{a}_1))[0.5], && (a(\mathsf{a}_2))[0.1], \\ & (a(\mathsf{a}_3))[0.9], && (a(\mathsf{a}_4))[0.6], \\ & (a(\mathsf{a}_5))[0.4], && (b(\mathsf{a}_1)a(\mathsf{a}_1))[0.5], \\ & (b(\mathsf{a}_2)a(\mathsf{a}_2))[0.1], && (b(\mathsf{a}_3)a(\mathsf{a}_3))[0.9], \\ & (b(\mathsf{a}_4)a(\mathsf{a}_4))[0.4], && (b(\mathsf{a}_5)a(\mathsf{a}_5))[0.1] \quad \} \quad . \end{aligned}$$

In addition, consider the conditional $r = (b(\mathsf{X}) \mid a(\mathsf{X}))[0.8]$. On the one hand, any probability function P with $P, D \models^{pr}_{\odot} \mathcal{R}$ also obeys $P, D \models^{pr}_{\odot} r$ as

$$\begin{aligned} & \frac{P(b(\mathsf{a}_1)a(\mathsf{a}_1)) + \cdots + P(b(\mathsf{a}_5)a(\mathsf{a}_5))}{P(a(\mathsf{a}_1)) + \cdots + P(a(\mathsf{a}_5))} \\ &= \frac{0.5 + 0.1 + 0.9 + 0.4 + 0.1}{0.5 + 0.1 + 0.9 + 0.6 + 0.4} \\ &= 0.8 \quad . \end{aligned}$$

On the other hand, no probability function P with $P, D \models^{pr}_{\varnothing} \mathcal{R}$ obeys $P, D \models^{pr}_{\varnothing} r$ due to

$$\begin{aligned}
&\frac{1}{5}(P(b(\mathsf{a}_1) \mid a(\mathsf{a}_1)) + \cdots + P(b(\mathsf{a}_5) \mid a(\mathsf{a}_5))) \\
&\qquad = \frac{1}{5}\left(\frac{0.5}{0.5} + \frac{0.1}{0.1} + \frac{0.9}{0.9} + \frac{0.4}{0.6} + \frac{0.1}{0.4}\right) \\
&\qquad = 0.78\overline{3} \neq 0.8 \quad .
\end{aligned}$$

As $P, D \models^{pr}_{\odot} \mathcal{R}$ is equivalent to $P, D \models^{pr}_{\varnothing} \mathcal{R}$ due to Proposition 5.1 the different semantics may lead to different inferences. Furthermore, the two semantics feature a different notion of consistency as $\mathcal{R} \cup \{r\}$ is $\varnothing$-inconsistent with respect to D but $\odot$-consistent with respect to D.

Although the previous example suggests that the difference of the two proposed semantics is marginal, in the following, we show that the difference can be made arbitrarily large.

Lemma 5.1. *Let $n > 1$ be some integer and let $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n \in (0, 1]$ with $\alpha_i \leq \beta_i$ for all $i = 1, \ldots, n$. Then*

$$\left| \frac{\frac{\alpha_1}{\beta_1} + \ldots + \frac{\alpha_n}{\beta_n}}{n} - \frac{\alpha_1 + \ldots + \alpha_n}{\beta_1 + \ldots + \beta_n} \right| < \frac{n-1}{n} \tag{5.11}$$

The proof of Lemma 5.1 can be found in Appendix A on page 240. The bound of $(n-1)/n$ is also the least upper bound as the following example shows.

Example 5.9. Let $n > 1$ be some integer and let $x \geq 2$ be some positive real value. Define $\alpha_1 = \ldots = \alpha_n = \beta_1 = \ldots = \beta_{n-1} = 1/x$ and $\beta_n = 1 - 1/x$. Observe, that for $x \geq 2$ it follows $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n \in (0, 1]$ and $\alpha_i \leq \beta_i$ for every $i = 1, \ldots, n$. Then it holds that

$$\frac{\frac{\alpha_1}{\beta_1} + \ldots + \frac{\alpha_n}{\beta_n}}{n} = \frac{n - 1 + \frac{\frac{1}{x}}{1-\frac{1}{x}}}{n} \stackrel{x\to\infty}{\to} \frac{n-1}{n}$$

and

$$\frac{\alpha_1 + \ldots + \alpha_n}{\beta_1 + \ldots + \beta_n} = \frac{\frac{n}{x}}{\frac{n-1}{x} + 1 - \frac{1}{x}} \stackrel{x\to\infty}{\to} 0 \quad .$$

Corollary 5.1. *Let $P \in \mathcal{P}^F(\Sigma)$ be some probability function, $D \subseteq U$ finite, and $(\psi(\vec{\mathsf{X}}) \mid \phi(\vec{\mathsf{X}}))$ be some probabilistic conditional with $|\mathsf{gnd}^P_D((\psi(\vec{\mathsf{X}}) \mid \phi(\vec{\mathsf{X}})))| > 1$. If $P, D \models^{pr}_{\varnothing} (\psi(\vec{\mathsf{X}}) \mid \phi(\vec{\mathsf{X}}))[d_1]$ and $P, D \models^{pr}_{\odot} (\psi(\vec{\mathsf{X}}) \mid \phi(\vec{\mathsf{X}}))[d_2]$ then*

$$|d_1 - d_2| < \frac{|\mathsf{gnd}^P_D((\psi(\vec{\mathsf{X}}) \mid \phi(\vec{\mathsf{X}})))| - 1}{|\mathsf{gnd}^P_D((\psi(\vec{\mathsf{X}}) \mid \phi(\vec{\mathsf{X}})))|}$$

Proof. This follows directly from Lemma 5.1 and the fact that P both $\varnothing$- and $\odot$-satisfies $(\psi(\vec{X}) \mid \phi(\vec{X}))[d]$ with respect to D for some d (therefore all appearing probabilities of premises are non-zero). Note that Equation (5.8) on page 141 considers all groundings of a conditional and not only those groundings with non-zero probability, cf. Equation (5.7) on page 141. However, if the premise of grounding $(\psi(\vec{a}) \mid \phi(\vec{a}))$ has probability zero so has the conjunction $\psi(\vec{a}) \wedge \phi(\vec{a})$. Therefore, both the numerator and denominator of (5.8) can be simplified by only summing over groundings in $\mathsf{gnd}^P_D((\psi(\vec{X}) \mid \phi(\vec{X})))$ which justifies the application of Lemma 5.1. $\square$

Basically, the above corollary says that the more constants we consider the more likely it is that the semantics differ significantly.

5.5 RELATED WORK

The most related works to the framework developed in this chapter are (Fisseler, 2010) and (Loh *et al.*, 2010)[4]. Both works also aim at extending probabilistic conditional logic and reasoning with the principle of maximum entropy to the relational case. Furthermore, the works (Jaeger, 1995) and (Halpern, 1990; Grove *et al.*, 1996a,b) also consider relational probabilistic reasoning as well as the works on statistical relational learning, cf. Section 2.4. All these works also share similarities with the inferential approach developed in the next chapter. In particular, most approaches for statistical relational learning such as BLPs (see Section 2.4.1) and MLNs (see Section 2.4.2) rely on inferential semantics and thus are reviewed at the end of the next chapter. In this section, however, we restrain the comparison of our framework with the works (Halpern, 1990; Jaeger, 1995; Grove *et al.*, 1996a,b; Fisseler, 2010; Loh *et al.*, 2010) to a syntactical and semantical comparison and continue with a comparison of inference later in Section 6.4.

5.5.1 *Grounding Semantics for RPCL*

The works (Fisseler, 2010) and (Loh *et al.*, 2010) share the same motivation as our work, namely, to extend probabilistic conditional logic to relational settings. In contrast to the present work those works base their semantics mainly on the assumption that open conditionals are to be treated as schemas for their instances. As for the syntax, both frameworks use basically the same formalism as RPCL but with one extension. Both frameworks allow the specification of *grounding constraints* which can be attached to probabilistic conditionals in order to avoid unwanted ground instances of the conditionals to hold. For instance, Example 5.2 (see page 134) can be rephrased with grounding constraints yielding $\mathcal{R}^*_{zoo} =_{def} \{r_1, r_2, r_3\}$ with

4 The last work is based on the diploma thesis (Loh, 2009) which was written under the author's supervision, see also (Thimm *et al.*, 2010).

$$r_1 =_{def} (likes(\mathsf{X},\mathsf{Y}) \,|\, elephant(\mathsf{X}) \wedge keeper(\mathsf{Y}))[0.6][\mathsf{Y} \neq \mathsf{fred}]$$
$$r_2 =_{def} (likes(\mathsf{X},\mathsf{fred}) \,|\, elephant(\mathsf{X}) \wedge keeper(\mathsf{fred}))[0.4][\mathsf{X} \neq \mathsf{clyde}]$$
$$r_3 =_{def} (likes(\mathsf{clyde},\mathsf{fred}) \,|\, elephant(\mathsf{clyde}) \wedge keeper(\mathsf{fred}))[0.7]$$

The informal interpretation for the above conditionals is straightforward. For example, the probabilistic conditional r_1 says that every elephant likes every keeper except Fred with probability 0.6. By employing grounding constraints one can get rid of the unnecessary inconsistencies in the plain grounding of $\mathcal{R}^*_{zoo}$. This approach is pursued in depth in (Fisseler, 2010) but (Loh *et al.*, 2010) also allows for *strategic grounding* when grounding constraints may not be specified. In the following we restrain our attention to the approach of (Loh *et al.*, 2010) as—syntactically and semantically—this approach subsumes the approach of relational probabilistic conditional reasoning in (Fisseler, 2010).

The basic notion of the approach of (Loh *et al.*, 2010) is that of a *grounding operator* which is defined as follows.

Definition 5.3 (Grounding operator)**.** A *grounding operator* $\mathcal{G}$ is a function $\mathcal{G} : \mathfrak{P}(\mathcal{L}^{\forall\!\!\!/\exists\!\!\!/}(\Sigma, V)) \times \mathfrak{P}(U) \to \mathfrak{P}(\mathcal{L}^{\forall\!\!\!/\exists\!\!\!/}(\Sigma, \emptyset))$.

In general, a grounding operator $\mathcal{G}$ *grounds* a relational knowledge base $\mathcal{R}$ to a propositional knowledge base (for which interpretation is straightforward) using a finite set of constants D with $\mathsf{Const}(\mathcal{R}) \subseteq D \subseteq U$.
The semantics of the approach of (Loh *et al.*, 2010) is defined as follows. Let $P \in \mathcal{P}^{\mathrm{F}}(\Sigma)$ be a probability function. Then P *$\mathcal{G}$-satisfies* a knowledge base $\mathcal{R}$ under grounding operator $\mathcal{G}$ with respect to finite D with $\mathsf{Const}(\mathcal{R}) \subseteq D \subseteq U$, denoted by $P, D \models^{pr}_{\mathcal{G}} \mathcal{R}$, if and only if $P \models^{pr} r$ for all $r \in \mathcal{G}(\mathcal{R}, D)$, cf. Equation (5.2) on page 135. We say that $\mathcal{R}$ is $\mathcal{G}$-consistent with respect to finite D if and only if there is at least one P with $P, D \models^{pr}_{\mathcal{G}} \mathcal{R}$, otherwise $\mathcal{R}$ is $\mathcal{G}$-inconsistent with respect to D.

The semantics of this approach relies heavily on the actual definition of a grounding operator. The simplest approach to ground a knowledge base is *universal instantiation*.

Definition 5.4 (Naive grounding operator)**.** The *naive grounding operator* $\mathcal{G}_U$ is defined as $\mathcal{G}_U(\mathcal{R}, D) =_{def} \bigcup_{r \in \mathcal{R}} \mathsf{gnd}_D(r)$ if $\mathsf{Const}(\mathcal{R}) \subseteq D$ and D is finite.

As discussed before, without taking grounding constraints into consideration, the naive grounding of a knowledge base like in Example 5.2 (see page 134) yields a $\mathcal{G}_U$-inconsistent ground knowledge base with respect to the given D. In order to avoid this situation, a series of different grounding operators are proposed in (Loh *et al.*, 2010) that allow for consistent grounding. The most sophisticated of these is the *specificity grounding operator* $\mathcal{G}_{sp}$. This operator removes from the plain grounding of a knowledge base all conflicting probabilistic conditionals and leaves only those that are most specific. We illustrate the specificity grounding operator on Example 5.2.

Example 5.10. The ground knowledge base $\mathcal{G}_{sp}(\mathcal{R}_{\text{zoo}}, D)$ obtained when applying $\mathcal{G}_{sp}$ to $\mathcal{R}_{\text{zoo}}$ with $D =_{def} \{\mathsf{clyde}, \mathsf{dumbo}, \mathsf{dave}, \mathsf{fred}\}$ is given as

$$\begin{aligned}
\mathcal{G}_{sp}(\mathcal{R}_{\text{zoo}}, D) = \{ \\
&(\mathit{likes}(\mathsf{clyde}, \mathsf{dave}) \mid \mathit{elephant}(\mathsf{clyde}) \wedge \mathit{keeper}(\mathsf{dave}))[0.6], \\
&(\mathit{likes}(\mathsf{dumbo}, \mathsf{dave}) \mid \mathit{elephant}(\mathsf{dumbo}) \wedge \mathit{keeper}(\mathsf{dave}))[0.6], \\
&(\mathit{likes}(\mathsf{dumbo}, \mathsf{fred}) \mid \mathit{elephant}(\mathsf{dumbo}) \wedge \mathit{keeper}(\mathsf{fred}))[0.4], \\
&(\mathit{likes}(\mathsf{clyde}, \mathsf{fred}) \mid \mathit{elephant}(\mathsf{clyde}) \wedge \mathit{keeper}(\mathsf{fred}))[0.7], \ldots\}
\end{aligned}$$

Note that we omitted listing ground conditionals that mention either $\mathit{elephant}(\mathsf{dave})$, $\mathit{elephant}(\mathsf{fred})$, $\mathit{keeper}(\mathsf{clyde})$, $\mathit{keeper}(\mathsf{dumbo})$ in the premise due to simplicity of presentation. As one can see, the specificity grounding operator removed the ground instance

$$(\mathit{likes}(\mathsf{clyde}, \mathsf{fred}) \mid \mathit{elephant}(\mathsf{clyde}) \wedge \mathit{keeper}(\mathsf{fred}))[0.4]$$

as r_3 is more specific than r_2 which produced the above ground instance. Similarly, the ground instances

$$\begin{aligned}
&(\mathit{likes}(\mathsf{clyde}, \mathsf{fred}) \mid \mathit{elephant}(\mathsf{clyde}) \wedge \mathit{keeper}(\mathsf{fred}))[0.7] \quad \text{and} \\
&(\mathit{likes}(\mathsf{dumbo}, \mathsf{fred}) \mid \mathit{elephant}(\mathsf{dumbo}) \wedge \mathit{keeper}(\mathsf{fred}))[0.6]
\end{aligned}$$

are removed due to the availability of more specific conditionals. Considering again the knowledge base $\mathcal{R}^*_{\text{zoo}}$ (see page 146) one can see that it holds that $\mathcal{G}_{sp}(\mathcal{R}_{\text{zoo}}, D) = \mathcal{G}_U(\mathcal{R}^*_{\text{zoo}}, D)$ as the explicitly modeled grounding constraints in $\mathcal{R}^*_{\text{zoo}}$ match the intuition behind specificity.

The formal definition of the specificity grounding operator can be found in (Loh *et al.*, 2010).

A similar remark like Remarks 5.1 on page 137 and 5.2 on page 142 can also be phrased for grounding semantics.

Remark 5.3. Let $r = (\psi \mid \phi)[d]$ be a ground conditional, P be a probability function, and D finite with $\mathsf{Const}(r) \subseteq D \subseteq U$ and $D \neq \emptyset$. Then for any grounding operator $\mathcal{G}$ it holds that $P, D \models^{pr}_{\mathcal{G}} r$ if and only if $P \models^{pr} r$. Therefore, grounding semantics is also a clear generalization of the standard probabilistic semantics for propositional probabilistic conditional logic, cf. Equation (5.2 on page 135).

If one restricts attention to the syntactical framework of RPCL (thus forbidding grounding constraints) the most obvious problem with the naive grounding operator is that it often produces a $\mathcal{G}_U$-inconsistent ground knowledge base with respect to the given D. Consider the following example.

Example 5.11. Let $\mathcal{R} =_{def} \{r_1, r_2\}$ be a knowledge base given via

$$r_1 =_{def} (a(\mathsf{X}))[0.7] \qquad\qquad r_2 =_{def} (a(\mathsf{c}_1))[0.4]$$

and $D =_{def} \{\mathsf{c}_1, \mathsf{c}_2, \mathsf{c}_3\}$. Grounding $\mathcal{R}$ with $\mathcal{G}_U$ results in a ground knowledge base $\mathcal{G}_U(\mathcal{R}, D)$ with $(a(\mathsf{c}_1))[0.7], (a(\mathsf{c}_1))[0.4] \in \mathcal{G}_U(\mathcal{R})$. It follows that $\mathcal{G}_U(\mathcal{R}, D)$ is $\mathcal{G}_U$-inconsistent with respect to D as there can be no P with both $P(a(\mathsf{c}_1)) = 0.7$ and $P(a(\mathsf{c}_1)) = 0.4$. However, $\mathcal{R}$ is both $\varnothing$-consistent with respect to D and $\odot$-consistent with respect to D. For example, for every probability function P with $P(a(\mathsf{c}_1)) = 0.4$ and $P(a(\mathsf{c}_2)) = P(a(\mathsf{c}_3)) = 0.85$ it holds that $P, D \models^{pr}_{\varnothing} \mathcal{R}$ and $P, D \models^{pr}_{\odot} \mathcal{R}$.

While the above example shows that there are cases where naive grounding semantics fail but averaging and aggregating do not, there is no example for the opposite direction. In fact, it holds the following statement.

Proposition 5.3. *Let $\mathcal{R}$ be a $\mathcal{G}_U$-consistent knowledge base with respect to $D \subseteq U$. Then $\mathcal{R}$ is also $\varnothing$- and $\odot$-consistent with respect to D and for every probability function $P \in \mathcal{P}^F(\Sigma)$ with $P, D \models^{pr}_{\mathcal{G}_U} \mathcal{R}$ it follows $P, D \models^{pr}_{\varnothing} \mathcal{R}$ and $P, D \models^{pr}_{\odot} \mathcal{R}$.*

Proof. Let $\mathcal{R}$ be $\mathcal{G}_U$-consistent with respect to D and let P be a probability function with $P, D \models^{pr}_{\mathcal{G}_U} \mathcal{R}$. Let $r = (\psi \,|\, \phi)[d] \in \mathcal{R}$ and let

$$S = \{(\psi_1 \,|\, \phi_1)[d], \ldots, (\psi_n \,|\, \phi_n)[d]\} = \mathsf{gnd}_r(D)$$

be the ground instances of r. Then it follows that $S \subseteq \mathcal{G}_U(\mathcal{R}, D)$ as well and therefore $P(\psi_1 \,|\, \phi_1) = \ldots = P(\psi_n \,|\, \phi_n) = d$. It follows that P also satisfies (5.3) (see page 137) and therefore $P, D \models^{pr}_{\varnothing} \mathcal{R}$. By Proposition 5.2 (see page 144) it also follows $P, D \models^{pr}_{\odot} \mathcal{R}$. $\square$

It follows that both averaging and aggregating semantics are clearly more expressive than naive grounding semantics on RPCL. However, the matter changes when we allow grounding constraints.

Example 5.12. Let $\mathcal{R} =_{def} \{r_1, r_2\}$ be a knowledge base given via

$$r_1 =_{def} (a(\mathsf{X}))[0][\mathsf{X} \neq \mathsf{c}_1] \qquad\qquad r_2 =_{def} (a(\mathsf{c}_1))[d]$$

with $d \in (0, 1]$. Then $\mathcal{R}$ is $\mathcal{G}_U$-consistent with respect to every finite D with $\mathsf{c}_1 \in D$. By neglecting the grounding constraint of r_1 and interpreting $\mathcal{R}$ in RPCL it follows that $\mathcal{R}$ is both $\varnothing$- and $\odot$-inconsistent with respect to to every finite D with $\mathsf{c}_1 \in D$. This is clear, as r_1 basically demands that $P(a(\mathsf{c})) = 0$ for *every* $\mathsf{c} \in D$ in order for $P, D \models^{pr}_{\varnothing} r_1$ or $P, D \models^{pr}_{\odot} r_1$ to be satisfied. For example, Equation 5.3 (see page 137) yields $P(a(\mathsf{c}_1)) + \ldots + P(a(\mathsf{c}_m)) = 0$ if $D = \{\mathsf{c}_1, \ldots, \mathsf{c}_m\}$. This can only be achieved if

$P(a(\mathsf{c}_1)) = \ldots = P(a(\mathsf{c}_m)) = 0$. As a consequence, it cannot hold that $P, D \models^{pr}_{\varnothing} r_2$.

In the above example both averaging and aggregating semantics fail as $\mathcal{R}$ represents a scenario with a debatable strong exception and we neglected the grounding constraint of r_1 for interpretation. In this thesis we refrain from introducing grounding constraints in RPCL as a matter of clarity. A grounding constraint puts the burden of identifying exceptions for a probabilistic conditional onto the knowledge engineer. Furthermore, identifying exceptions may be hard or even impossible in practice. Nonetheless, RPCL might be as well augmented with grounding constraints and both averaging and aggregating can be modified in a straightforward way to consider only ground instances of conditionals that obey the grounding constraints. Then the knowledge base $\mathcal{R}$ in Example 5.12 would be both $\varnothing$- and $\odot$-consistent with respect to D.

Example 5.13. Consider again the knowledge base $\mathcal{R} = \{r_1, r_2\}$ but without the grounding constraint of r_1, i. e.

$$r_1 =_{def} (a(\mathsf{X}))[0] \qquad\qquad r_2 =_{def} (a(\mathsf{c}_1))[d] \quad .$$

Obviously, $\mathcal{R}$ is $\mathcal{G}_U$-inconsistent with respect to every finite D with $\mathsf{c}_1 \in D$. However, $\mathcal{R}$ is $\mathcal{G}_{sp}$-consistent with respect to every finite D with $\mathsf{c}_1 \in D$ as $(a(\mathsf{c}_1))[0] \notin \mathcal{G}_{sp}(\mathcal{R})$ as r_2 is more specific than r_1 with respect to c_1.

Although the previous example suggests that specificity grounding semantics avoids the drawbacks of naive grounding semantics completely the notion of specificity is not always applicable in the presence of conflicting information.

Example 5.14. Consider the knowledge base $\mathcal{R} =_{def} \{r_1, r_2, r_3\}$ given via

$$r_1 =_{def} (a(\mathsf{X}))[0.7] \quad r_2 =_{def} (a(\mathsf{X}) \,|\, b(\mathsf{X}))[1] \quad r_3 =_{def} (b(\mathsf{c}_1))[1] \quad .$$

If e. g. $D = \{\mathsf{c}_1, \mathsf{c}_2\}$ then $\mathcal{R}$ is both $\varnothing$- and $\odot$-consistent with respect to D. For example, let P be such that

$$P(b(\mathsf{c}_1)) = P(a(\mathsf{c}_1)) = 1 \qquad P(b(\mathsf{c}_2)) = 0 \qquad P(a(\mathsf{c}_2)) = 0.4$$

Then $P, D \models^{pr}_{\varnothing} r_3$ by definition, $P, D \models^{pr}_{\varnothing} r_2$ as only the ground instance of r_2 with c_1 is considered in Equation (5.3) on page 137 and $P(b(\mathsf{c}_1)) = P(a(\mathsf{c}_1)) = 1$, and finally $P, D \models^{pr}_{\varnothing} r_1$ as the average of $P(a(\mathsf{c}_1)) = 1$ and $P(a(\mathsf{c}_2)) = 0.4$ is exactly 0.7. It follows $P, D \models^{pr}_{\varnothing} \mathcal{R}$ and similarly $P, D \models^{pr}_{\odot} \mathcal{R}$ as well. However, $\mathcal{R}$ is $\mathcal{G}_{sp}$-inconsistent with respect to every finite D with $\mathsf{c}_1 \in D$ because of the following reason. The notion of specificity employed in (Loh *et al.*, 2010) relies mainly on syntactical comparison of conditionals. More specifically, a ground instance of a probabilistic conditional is removed if there is another conditional that is "structurally identical" to the first conditional but more specific, cf. (Loh *et al.*, 2010). In $\mathcal{R}$ no two conditionals are comparable using this notion and therefore no ground instance is removed at all yielding $\mathcal{G}_{sp}(\mathcal{R}, D) = \mathcal{G}_U(\mathcal{R}, D)$. In particular, it holds that $(a(\mathsf{c}_1))[0.7]$, $(b(\mathsf{c}_1))[1]$, $(a(\mathsf{c}_1) \mid b(\mathsf{c}_1))[1] \in \mathcal{G}_{sp}(\mathcal{R}, D)$ which amounts to $\mathcal{G}_{sp}(\mathcal{R}, D)$ being $\mathcal{G}_{sp}$-inconsistent with respect to D.

Another similar approach to the one discussed above is probabilistic logic programming under maximum entropy (Kern-Isberner and Lukasiewicz, 2004). That work considers bounded probabilistic conditionals as discussed in Section 2.3 and bases on a simple relational language without disjunctions. The semantics of the relational probabilistic framework of (Kern-Isberner and Lukasiewicz, 2004) is based on plain grounding and is similar to the approaches of (Fisseler, 2010) and (Loh *et al.*, 2010). More specifically, a probability function P satisfies a bound relational probabilistic conditional $(\psi \mid \phi)[l, u]$ if and only if for every ground instance $(\psi' \mid \phi')[l, u]$ of $(\psi \mid \phi)[l, u]$ it holds that $P(\psi \mid \phi) \in [l, u]$. Consequently, everything that has been said regarding the relation of our approach to the naive grounding operator $\mathcal{G}_U$ of (Loh *et al.*, 2010) also applies to (Kern-Isberner and Lukasiewicz, 2004).

5.5.2 *First-order Probabilistic Logic*

There are several works by Bacchus, Halpern and colleagues that deal with probabilistic reasoning on first-order logic, see e. g. (Bacchus, 1990; Halpern, 1990; Grove *et al.*, 1994; Bacchus *et al.*, 1996). From a conceptual point of view, the basic difference of those works with respect to the present work is that we aim at extending a propositional probabilistic framework—namely, probabilistic conditional logic—with first-order concepts while in those works the aim lies in introducing probabilistic concepts into first-order logic. In the following, we have a closer look on syntax and semantics of some of the logics proposed in (Bacchus, 1990; Halpern, 1990; Grove *et al.*, 1994; Bacchus *et al.*, 1996).[5]

In (Halpern, 1990)—which is inspired by and extends work of (Bacchus, 1990)—two logics and a combination thereof are proposed that introduce probabilities in first-order formulas. The first language $\mathcal{L}_1$ augments a first-order language $\mathcal{L}(\Sigma, V)$ with statements of the form $w_{\mathsf{X}_1,\ldots,\mathsf{X}_m}(\phi) \geq \alpha$ with $\phi \in \mathcal{L}(\Sigma, V)$, $\mathsf{X}_1, \ldots, \mathsf{X}_m$ are free and not bound variables in ϕ, and $\alpha \in [0, 1]$. The intuition behind the statement $w_{\mathsf{X}_1,\ldots,\mathsf{X}_m}(\phi) \geq \alpha$ is that the probability of choosing some random $\mathsf{c}_1, \ldots, \mathsf{c}_m$ that satisfy $\phi[\mathsf{X}_1/\mathsf{c}_1, \ldots, \mathsf{X}_m/\mathsf{c}_m]$ is greater or equal to α. Hence, the probability function P under consideration is actually a probability function on the domain. More specifically, an interpretation for $\mathcal{L}_1$ is a tuple (I, P) with a first-order interpretation $I = (U_I, f_I^U, Pred_I, Func_I) \in \mathsf{Int}(\Sigma)$ (see Definition 2.9 on page 15) and a

5 Some of the syntax and semantics of the logics presented in this section are simplified due to matters of presentation.

probability function $P : U_I \to [0,1]$. The function P is extended to sets of elements $G \subseteq U_I$ via $P(G) = \prod_{\mathsf{c} \in G} P(\mathsf{c})$. Then an interpretation (I, P) satisfies a statement $w_{\mathsf{X}_1,\ldots,\mathsf{X}_m}(\phi) \geq \alpha$, denoted by $(I, P) \models_1^{pr} w_{\mathsf{X}_1,\ldots,\mathsf{X}_m}(\phi) \geq \alpha$, if and only if

$$\sum_{I \models^{\mathrm{F}} \phi[\mathsf{X}_1/\mathsf{c}_1,\ldots,\mathsf{X}_m/\mathsf{c}_m]} P(\{\mathsf{c}_1, \ldots, \mathsf{c}_m\}) \geq \alpha \quad . \tag{5.12}$$

The relation $\models_1^{pr}$ extends in a straightforward fashion to statements of the form $w_{\mathsf{X}_1,\ldots,\mathsf{X}_m}(\phi) \leq \alpha$ and $w_{\mathsf{X}_1,\ldots,\mathsf{X}_m}(\phi) = \alpha$. For other first-order formulas the relation $\models_1^{pr}$ is the same as $\models^{\mathrm{F}}$, i. e., it holds that $(I, P) \models_1^{pr} \phi$ if and only if $I \models^{\mathrm{F}} \phi$ for $\phi \in \mathcal{L}(\Sigma, V)$, see also page 15ff. Consider the following example taken from (Halpern, 1990). Let $\Sigma = (\{\mathsf{c}_1, \mathsf{c}_2, \mathsf{c}_3\}, \{son\}, \emptyset)$ be a first-order signature and $((\{\mathsf{c}_1, \mathsf{c}_2, \mathsf{c}_3\}, \mathsf{id}, \{son(\mathsf{c}_1, \mathsf{c}_2)\}, \emptyset), P)$ be an interpretation[6] with $P(\mathsf{c}_1) = 1/3$, $P(\mathsf{c}_2) = 1/2$, and $P(\mathsf{c}_3) = 1/6$. Then we have

$$(I, P) \models_1^{pr} w_{\mathsf{X}}(son(\mathsf{X}, \mathsf{c}_3)) = 0 \tag{5.13}$$

$$(I, P) \models_1^{pr} w_{\mathsf{X}}(son(\mathsf{c}_1, \mathsf{X})) = \frac{1}{2} \tag{5.14}$$

$$(I, P) \models_1^{pr} w_{\mathsf{X},\mathsf{Y}}(son(\mathsf{X}, \mathsf{Y})) = \frac{1}{6} \quad . \tag{5.15}$$

In particular, (5.15) holds as there is exactly one instantiation of $son(\mathsf{X}, \mathsf{Y})$ that is satisfied in I, namely $son(\mathsf{c}_1, \mathsf{c}_2)$, and the probability of "drawing" both c_1 and c_2 is $P(\mathsf{c}_1)P(\mathsf{c}_2) = 1/6$. As a consequence, the informal interpretation of $\mathcal{L}_1$ is a statistical one. Formulas in $\mathcal{L}_1$ are statistical expressions that rely on the probability of selecting some elements from the underlying domain. While $\mathcal{L}_1$ is apt for representing those statistical expressions it fails to model degrees of belief such as "*the probability that Tweety flies is* 0.7". As the formula $flies(\mathsf{tweety})$ is ground there is no element from the domain to be chosen and it always holds that either $I, P \models_1^{pr} w_{\mathsf{X}}(flies(\mathsf{tweety})) = 0$ or $I, P \models_1^{pr} w_{\mathsf{X}}(flies(\mathsf{tweety})) = 1$ for every P and I. In (Halpern, 1990) the logic $\mathcal{L}_2$ is proposed to complement $\mathcal{L}_1$ in this matter. The language $\mathcal{L}_2$ augments a first-order language $\mathcal{L}(\Sigma, V)$ with statements of the form $w(\phi) \geq \alpha$ with a sentence $\phi \in \mathcal{L}(\Sigma, V)$ and $\alpha \in [0, 1]$. The informal interpretation of a statement $w(\phi) \geq \alpha$ is that the probability of the formula ϕ being true is at least α. The semantics for $\mathcal{L}_2$ are, essentially, the same as for probabilistic conditional logic. For a first-order sentence $\phi \in \mathcal{L}(\Sigma, V)$ and a probability function $P : \mathsf{Int}(\Sigma) \to [0, 1]$ the probability of ϕ can be determined via

$$P(\phi) = \sum_{I \in \mathsf{Int}(\Sigma),\ I \models^{\mathrm{F}} \phi} P(I) \quad .$$

A first-order interpretation I and a probability function $P : \mathsf{Int}(\Sigma) \to [0, 1]$ satisfy a statement $w(\phi) \geq \alpha$, denoted by $(I, P) \models_2^{pr} w(\phi) \geq \alpha$, if and

6 The function id is the *identity*, i. e., it holds that $\mathsf{id}(x) = x$ for all $x \in \mathrm{Dom}\ \mathsf{id}$.

only if $P(\phi) \geq \alpha$. The relation $\models^{pr}_{2}$ extends in a straightforward fashion to statements of the form $w(\phi) \leq \alpha$ and $w(\phi) = \alpha$. Note that this relation is actually independent of I but for other formulas $\models^{pr}_{2}$ is the same as $\models^{\mathrm{F}}$ (ignoring P).

The logics $\mathcal{L}_1$ and $\mathcal{L}_2$ are combined in the logic $\mathcal{L}_3$ which allows for both, statements of the form $w_{\mathsf{X}_1,\ldots,\mathsf{X}_m}(\phi) \geq \alpha$ and statements of the form $w(\phi) \geq \alpha$. Semantics are given to $\mathcal{L}_3$ with the use of two different probability functions, the first one $P_1 : U_I \to [0,1]$ represents a probability distribution on the elements of the domain, and the second one $P_2 : \mathsf{Int}(\Sigma) \to [0,1]$ represents a probability distribution on the first-order interpretations. Hence, an interpretation for $\mathcal{L}_3$ is a tuple (I, P_1, P_2) with a first-order interpretation I and probability functions P_1 and P_2. A formula of the form $w_{\mathsf{X}_1,\ldots,\mathsf{X}_m}(\phi) \geq \alpha$ is satisfied by (I, P_1, P_2), denoted by $(I, P_1, P_2) \models^{pr}_{3} w_{\mathsf{X}_1,\ldots,\mathsf{X}_m}(\phi) \geq \alpha$, if and only if $(I, P_1) \models^{pr}_{1} w_{\mathsf{X}_1,\ldots,\mathsf{X}_m}(\phi) \geq \alpha$. A formula of the form $w(\phi) \geq \alpha$ is satisfied by (I, P_1, P_2), denoted by $(I, P_1, P_2) \models^{pr}_{3} w(\phi) \geq \alpha$, if and only if $(I, P_2) \models^{pr}_{2} w(\phi) \geq \alpha$. Note that the language $\mathcal{L}_3$ also allows the nesting of expressions of both types which complicates the definition of the semantics. However, the interpretation of such nested statements should be clear from the presented basic definitions. The language $\mathcal{L}_3$ also allows for conditionals to appear in a probability statement. As for probabilistic conditional logic, the probability of a conditional is the conditional probability of the conditional's conclusion given its premise. For example, a valid sentence in $\mathcal{L}_3$ is

$$\begin{aligned}\phi =_{def} \; & (w(w_{\mathsf{X}}(\mathit{flies}(\mathsf{X}) \,|\, \mathit{bird}(\mathsf{X})) > 0.99) < 0.2) \wedge \\ & \qquad (w(w_{\mathsf{X}}(\mathit{flies}(\mathsf{X}) \,|\, \mathit{bird}(\mathsf{X})) > 0.9) > 0.95) \quad .\end{aligned}$$

The above statement basically states, that the subjective probability of the statistical statement that at least 99 % of all birds fly is less than 0.2 and that the subjective probability of the statistical statement that at least 90 % of all birds fly is at least 0.95. Note that ϕ is also satisfiable in $\mathcal{L}_3$.

From the point of view of knowledge representation the logic $\mathcal{L}_3$ is clearly more expressive than RPCL, partially due to the availability of quantifiers. However, our semantical notions differ from both probabilistic concepts employed in $\mathcal{L}_3$. Consider the probabilistic fact

$$r =_{def} (\mathit{flies}(\mathsf{X}))[0.9] \tag{5.16}$$

and averaging semantics $\models^{pr}_{\varnothing}$. Then for a probability function P and some finite $D \subseteq U$ with $D = \{\mathsf{c}_1, \ldots, \mathsf{c}_m\}$ it holds that $P, D \models^{pr}_{\varnothing} r$ if $P(\mathit{flies}(\mathsf{c}_1)) + \ldots + P(\mathit{flies}(\mathsf{c}_m)) = m \cdot 0.9$. If the agent believes r then it may be unsure about the actual probabilities of the flying capabilities of the actual individuals but still believes that, in average, every individual flies with probability 0.9. Now consider the statement

$$r' =_{def} w_{\mathsf{X}}(\mathit{flies}(\mathsf{X})) = 0.9$$

in $\mathcal{L}_3$. In order to interpret r' we do not only need a probability function P_1 but the current state of the world, i. e. an interpretation I, as well. Then r' is satisfied with respect to P_1 and I if there are constants $\mathsf{c}_1, \ldots, \mathsf{c}_n$ such that the interpretation of each c_i flies in I (for $i = 1, \ldots, n$) and $P_1(\{\mathsf{c}_1\}) + \ldots + P_1(\{\mathsf{c}_n\}) = 0.9$. Note that, in general, it may be the case that $\{\mathsf{c}_1, \ldots, \mathsf{c}_n\}$ do not represent 90 % of the domain. If the agent believes r' and it has no idea of the underlying probability distribution on the domain elements then r' represents much less information than r. It could be the case that in I there is only a single domain element that flies but the probability that this element is chosen is 0.9 with respect to P_1. Clearly, this interpretation deviates substantially from the intuition of r. In many scenarios, P_1 might be assumed to be a uniform probability function, cf. (Halpern, 1990). Then an interpretation I satisfies r' if for 90 % of the domain elements $\mathit{flies}(\mathsf{X})$ is true in I. If an agent believes in both r' and P_1 being uniform then the agent is unsure which of the present domain elements fly but is sure that 90 % of them do fly. This interpretation is similar to the one for r but more restraining. In particular, if the agent believes every situation to be (equally) likely where 90 % of the elements do fly then it also believes that every particular domain element flies with a probability of 0.9, i. e., from $(I, P_1, P_2) \models_3^{pr} r'$ with P_1 and P_2 uniform it follows $(I, P_1, P_2) \models_3^{pr} w(\mathit{flies}(\mathsf{c})) = 0.9$ for every $\mathsf{c} \in U$. Imagine there are ten elements in the domain, then there are ten possible situations where nine of the ten elements fly and one does not. Therefore, every element flies in nine of ten possible situations and—assuming that P_2 is uniform—it follows the claim. This implication is also known as Miller's principle (Miller, 1966; Skyrms, 1980) and *direct inference* in (Bacchus *et al.*, 1996). But also assuming a uniform P_2 reduces probabilistic reasoning to simple counting.

Observe that P_1 and P_2 are basically independent when it comes to interpretation. Let $U = \{\mathsf{c}_1, \mathsf{c}_2, \mathsf{c}_3\}$ be the domain under discourse and consider the following formula

$$(w_{\mathsf{X}}(\mathit{flies}(\mathsf{X})) = 0.9) \wedge w(\mathit{flies}(\mathsf{c}_1)) = 0.5 \wedge w(\mathit{flies}(\mathsf{c}_2)) = 0.5 \\ \wedge w(\mathit{flies}(\mathsf{c}_2)) = 0.5 \quad . \tag{5.17}$$

Note that the above formula is satisfiable in $\mathcal{L}_3$ as the first sub-formula is interpreted with respect to P_1 while the latter three sub-formulas are interpreted with respect to P_2. A similar example has also been used in (Jaeger, 1995) to show several inadequacies of $\mathcal{L}_3$ (see also below). Representing the above formula in RPCL yields

$$r_1 =_{def} (\mathit{flies}(\mathsf{X}))[0.9] \qquad r_2 =_{def} (\mathit{flies}(\mathsf{c}_1))[0.5]$$
$$r_3 =_{def} (\mathit{flies}(\mathsf{c}_2))[0.5] \qquad r_4 =_{def} (\mathit{flies}(\mathsf{c}_3))[0.5]$$

and obviously $\{r_1, r_2, r_3, r_4\}$ is both $\varnothing$- and $\odot$-inconsistent with respect to $\{\mathsf{c}_1, \mathsf{c}_2, \mathsf{c}_3\}$. Concerning formal interpretation, our approach makes no distinction between statistical probabilities and degrees of beliefs. In particular, in our approach all conditionals are interpreted using the same probability function. Moreover, $\mathcal{L}_3$ allows for representing both statistical probabilities involving undetermined domain elements and degrees of belief on determined domain elements but is not able to represent degrees of belief on undetermined domain elements. More specifically, statements like r_1 above that represent average probabilities cannot be represented in $\mathcal{L}_3$ without listing the involved constants explicitly.

In (Jaeger, 1995) Jaeger develops a probabilistic first-order logic $\mathcal{L}^\beta$ that fixes the problem of $\mathcal{L}_3$ regarding satisfiability of formulas like (5.17). Syntactically, the logic $\mathcal{L}^\beta$ is very similar to $\mathcal{L}_3$. On the one hand, $\mathcal{L}^\beta$ extends $\mathcal{L}_3$ by introducing a special set of *event symbols* which are treated differently from ordinary constants, cf. (Jaeger, 1995). On the other hand, $\mathcal{L}^\beta$ restricts the syntax by disallowing nesting of statements with w.[7] We do not go into the details of the semantics of $\mathcal{L}^\beta$ but only state that those are similar to the ones of $\mathcal{L}_3$ with only subtle differences on the technical level that require a huge notational introduction. We rather discuss some examples concerning satisfiability. In particular, Jaeger shows in (Jaeger, 1995) that the formula

$$(\neg\exists \mathsf{X} : p(\mathsf{X})) \wedge (w(p(\mathsf{c})) = 1)$$

is satisfiable in $\mathcal{L}_3$ but not in $\mathcal{L}^\beta$. This formula states that 1.) there is no element a in the domain such that $p(\mathsf{a})$ is true and 2.) our degree of belief that $p(\mathsf{c})$ is true is one. Obviously, the above formula should be inconsistent but $\mathcal{L}_3$ allows satisfiability. Note that this example is similar in spirit to the formula (5.17) discussed above. The straightforward translation of the above formula to RPCL can be given via $\mathcal{R} =_{def} \{(p(\mathsf{X}))[0], (p(\mathsf{c}))[1]\}$ which is both $\varnothing$- and $\odot$-inconsistent with respect to every D with $\mathsf{c} \in D$. Therefore, both RPCL and $\mathcal{L}^\beta$ meet our intuition in interpreting the above formula while $\mathcal{L}_3$ does not. Furthermore, consider the following example that has also been taken from (Jaeger, 1995).

> Let f be a movie with starring actors *actorA* and *actorB*. We further know that 80 % of the movies starring *actorA* are American productions and only 20 % of the movies starring *actorB* are American productions. We also know that f is the only movie where *actorA* and *actorB* star together. We also think that f is an American production with probability 0.5.

Intuitively, the above scenario makes sense from a commonsensical perspective. We formalize this scenario using the syntax $\mathcal{L}_3$ as follows.

$$\begin{aligned}&\exists^{=1}\mathsf{X} : \mathit{starsActorA}(\mathsf{X}) \wedge \mathit{starsActorB}(\mathsf{X}) \wedge \\ &\quad w_{\mathsf{X}}(\mathit{american}(\mathsf{X}) \mid \mathit{starsActorA}(\mathsf{X})) = 0.8 \wedge \\ &\quad w_{\mathsf{X}}(\mathit{american}(\mathsf{X}) \mid \mathit{starsActorB}(\mathsf{X})) = 0.2 \wedge\end{aligned}$$

7 Furthermore, $\mathcal{L}^\beta$ restricts the syntax by disallowing free variables inside w statements but this restriction has already been applied in the presentation of $\mathcal{L}_3$ above.

$$w(starsActorA(\mathsf{f}) \wedge starsActorB(\mathsf{f})) = 1 \wedge$$
$$w(american(\mathsf{f})) = 0.5 \quad .$$

In the above formula the symbol $\exists^{=1}$ is meant to abbreviate "there exists exactly one". In (Jaeger, 1995) it is shown that the above formula is satisfiable in $\mathcal{L}_3$ but not in $\mathcal{L}^\beta$. We can introduce the quantifier $\exists^{=1}$ into the definition of the otherwise quantifier-free language $\mathcal{L}^{\not\forall\not\exists}(\Sigma, V)$ via $I, VA \models^{\mathrm{F}} \exists^{=1}X : \phi$ if and only if for some variable assignment VA' that is the same as VA but possibly $VA'(X) \neq VA(X)$ it holds that $I, VA' \models^{\mathrm{F}} \phi$ and for all VA'' with $VA''(X) \neq VA'(X)$ it holds that $I, VA'' \not\models^{\mathrm{F}} \phi$. Then, a translation of the above formula in RPCL can be given via a knowledge base $\mathcal{R} = \{r_1, \ldots, r_5\}$ with

$$r_1 =_{def} (\exists^{=1}\mathsf{X} : starsActorA(\mathsf{X}) \wedge starsActorB(\mathsf{X}))[1]$$
$$r_2 =_{def} (american(\mathsf{X}) \mid starsActorA(\mathsf{X}))[0.8]$$
$$r_3 =_{def} (american(\mathsf{X}) \mid starsActorB(\mathsf{X}))[0.2]$$
$$r_4 =_{def} (starsActorA(\mathsf{f}) \wedge starsActorB(\mathsf{f}))[1]$$
$$r_5 =_{def} (american(\mathsf{f}))[0.5] \quad .$$

Then $\mathcal{R}$ is both $\varnothing$- and $\odot$-consistent with respect to D with $\mathsf{f} \in D$ and $|D| > 4$. More precisely, for $D = \{\mathsf{f}, \mathsf{g}_1, \ldots, \mathsf{g}_4\}$ a probability function P that satisfies

$$\begin{aligned}
P(american(\mathsf{f})) &= 0.5\\
P(starsActorA(\mathsf{f})) &= P(starsActorB(\mathsf{f})) = 1\\
P(american(\mathsf{g}_1)) &= P(american(\mathsf{g}_2)) = 0.95\\
P(starsActorA(\mathsf{g}_1)) &= P(starsActorA(\mathsf{g}_2)) = 1\\
P(starsActorB(\mathsf{g}_1)) &= P(starsActorB(\mathsf{g}_2)) = 0\\
P(american(\mathsf{g}_3)) &= P(american(\mathsf{g}_4)) = 0.05\\
P(starsActorA(\mathsf{g}_3)) &= P(starsActorA(\mathsf{g}_4)) = 0\\
P(starsActorB(\mathsf{g}_3)) &= P(starsActorB(\mathsf{g}_4)) = 1
\end{aligned}$$

both $\varnothing$- and $\odot$-satisfies $\mathcal{R}$ with respect to D. Note that P satisfies r_1 with respect to D as f is the only movie where both *actorA* and *actorB* star (in g_1 and g_2 only *actorA* stars and in g_3 and g_4 only *actorB* stars). Also, r_2 is satisfied as $0.95 + 0.95 + 0.5 = 3 \cdot 0.8$ and similarly r_3 is satisfied as $0.05 + 0.05 + 0.5 = 3 \cdot 0.2$. The conditionals r_4 and r_5 are satisfied by definition. Therefore, both RPCL and $\mathcal{L}_3$ meet our intuition in interpreting the above formula while $\mathcal{L}^\beta$ does not.

The semantical approach of (Halpern, 1990) occurs with minor variations in (Grove *et al.*, 1994; Bacchus *et al.*, 1996). Syntactically, the logics employed in (Grove *et al.*, 1994; Bacchus *et al.*, 1996) are the same as $\mathcal{L}_1$ with the difference that instead of the strict relations $\leq$ and $=$, that are used to represent statistical statements, approximate relations $\preceq$ and $\approx$ are introduced. An expression of the form $w_X(\phi) \approx \alpha$ then is informally interpreted by saying that "approximately a proportion of α elements of the domain satisfy ϕ". However, in (Grove *et al.*, 1994) a statement like $w_X(\phi) \approx \alpha$ is rewritten into a statement of the form $\alpha - \epsilon \leq w_X(\phi) \leq \alpha + \epsilon$ with some "small" $\epsilon > 0$ and interpreted using the semantics of $\mathcal{L}_1$. Furthermore, the works (Grove *et al.*, 1994; Bacchus *et al.*, 1996) simplify the setup for their logic a little bit by assuming a uniform probability function on the domain elements for interpretation. It follows that the above discussion also applies in the same way to (Grove *et al.*, 1994; Bacchus *et al.*, 1996).

5.6 summary and discussion

In this chapter we introduced relational probabilistic conditional logic as an extension to probabilistic conditional logic that bases on a simple relational signature for knowledge representation. While traditional semantics for first-order probabilistic logics treat universal quantification in the strict sense of first-order logic, we proposed two novel semantics that give alternative meanings to open probabilistic conditionals. In particular, we presented *averaging* and *aggregating* semantics which both allow for exceptions to the probabilities represented by open probabilistic conditionals. While averaging semantics demands that the average of the conditional probabilities of instances of an open probabilistic conditional matches the conditional's probability, aggregating semantics employs a generalized definition of conditional probabilities that involves the instances of a probabilistic conditional together and not separately. We compared both semantics and showed that they coincide on restricted knowledge bases but also may differ significantly in the general case. We also compared our semantics with similar approaches from the literature, in particular with the grounding semantics of (Fisseler, 2010; Loh *et al.*, 2010; Kern-Isberner and Lukasiewicz, 2004) and the works (Halpern, 1990; Grove *et al.*, 1994; Bacchus *et al.*, 1996; Jaeger, 1995).

The ability to model relations among individuals is crucial in knowledge representation and reasoning. In this chapter we presented a novel approach for incorporating relational aspects into a framework for probabilistic reasoning. Our semantical notions differ significantly from previous approaches as we do not extend the classical interpretation of universal quantification. In Section 5.5 we investigated the relationships of our approaches with others from the literature. We showed that our semantical notions seem to outperform existing semantics in terms of common sense understanding. In particular, we have shown that the examples of (Jaeger, 1995)—that were used to illustrate shortcomings of the logics $\mathcal{L}_3$ and $\mathcal{L}^\beta$—can be represented in RPCL and interpreted according to commonsense. However, it is also clear that RPCL is less expressive in several aspects than e. g. $\mathcal{L}_3$ as the latter uses full first-order logic and allows for nesting of probability statements. For example, the statement "my degree of belief that 90 % of all birds fly is 0.7" is expressible in $\mathcal{L}_3$ but not in RPCL. From the perspective of commonsense reasoning it is arguable whether such expressivity is needed. Most people would probably be indifferent in deciding whether the formalization $w(w_{\mathsf{X}}(\mathit{flies}(\mathsf{X}) \mid \mathit{bird}(\mathsf{X})) = 0.9) = 0.7$ is necessary or a probabilistic conditional like $(\mathit{flies}(\mathsf{X}) \mid \mathit{bird}(\mathsf{X}))[0.8]$ describes an adequately similar state of affairs (given our semantical notions).

While in Chapters 3 and 4 we adopted the standard semantics for probabilistic conditional logic in order to analyze and resolve inconsistencies, in this chapter we neglected standard semantics that rendered knowledge bases like $\{(\mathit{flies}(\mathsf{X}))[0.9], (\mathit{flies}(\mathsf{tweety}))[0.1]\}$ inconsistent in order to come up with more adequate notions of consistency and inconsistency. In particular, we laid the foundations for further research in relational probabilistic conditional logic. One possible direction for further research is now to investigate inconsistency measures for relational probabilistic conditional logic. However, our motivation for the investigation of inconsistency measures in propositional probabilistic conditional logic was the inability to use model-based inductive reasoning processes—such as reasoning based on the principle of maximum entropy—in the presence of inconsistent information. In this chapter we just defined syntax and semantics for relational probabilistic conditional logic and we have not yet touched the issue of inductive reasoning. As a consequence, we postpone the issue for inconsistency measurement in relational probabilistic conditional logic for future work and continue with an investigation of the issue of reasoning itself in RPCL based on the semantics developed in this chapter.

6 REASONING AT MAXIMUM ENTROPY IN RPCL

In the previous chapter we gave a semantical account for treating relational probabilistic conditionals. In Section 2.3 we pointed out a simple way to reason with (propositional) probabilistic conditionals that bases on determining upper and lower bounds for queries, i. e., by determining the set of all models of a knowledge base and compute the range of probabilities of a single conditional within this set. This approach is, of course, also applicable for relational probabilistic conditionals but suffers from the same disadvantages as RPCL subsumes PCL for both averaging and aggregating semantics. However, for PCL we already discussed reasoning based on the principle of maximum entropy which turned out to be a satisfactory reasoning mechanism, cf. Definition 2.27 on page 32. Consequently, in this chapter we discuss extending reasoning based on the principle of maximum entropy to RPCL. Before doing so we first have a look on relational probabilistic reasoning in RPCL in general and illustrate possible problems with means of several benchmark examples. As mentioned earlier, relational probabilistic reasoning has been discussed before in the field of statistical relational learning—see Section 2.4 on page 32—but not in the sense of *default reasoning* or *reasoning with exceptions*. However, there are some works that discuss these topics as well for relational probabilistic frameworks, see e. g. (Jaeger, 1995; Bacchus *et al.*, 1996). We give a comparison of those works with ours later in this chapter and borrow several benchmark examples in order to evaluate our approach. Besides benchmark examples another important evaluation criterium is the satisfaction of properties. There are many works that deal with properties for non-monotonic in general, see e. g. (Makinson, 1989). We also take a principled approach to the problem of relational probabilistic reasoning and adopt several properties for non-monotonic reasoning from (Makinson, 1989) and also develop a series of desirable properties tailored specifically for relational probabilistic inference. We go on by applying the principle of maximum entropy to define inference operators based on the two semantics developed in the previous chapter and compare these operators with existing approaches.

This chapter is organized as follows. In Section 6.1 we start by investigating the problem of relational probabilistic reasoning using benchmark examples and develop a series of rationality postulates. We go on in Section 6.2 with proposing specific inference operators that base on the principle of maximum entropy and employ the semantical notions developed in the previous chapter. In Section 6.3 we analyze and compare the behavior of our inference operators and give hints to related work in Section 6.4. In Section 6.5 we conclude this chapter with some final remarks.

6.1 PROBABILISTIC REASONING AND DESIRABLE PROPERTIES

In this thesis we focus on inductive inference for RPCL, i. e., on the problem of finding a "good" probability function $P_{\mathcal{R}}$ that satisfies all probabilistic conditionals of a knowledge base $\mathcal{R}$, given one of the two proposed semantics. More specifically, we are interested in an operator $\mathcal{I}(\mathcal{R}, D)$ that takes a knowledge base $\mathcal{R}$ and a finite set of constants D with $\mathsf{Const}(\mathcal{R}) \subseteq D \subseteq U$ with $D \neq \emptyset$ as input and returns a probability function $P = \mathcal{I}(\mathcal{R}, D) \in \mathcal{P}^{\mathrm{F}}(\Sigma)$ as output such that P describes $\mathcal{R}$ in a commonsensical manner. In particular, the resulting function should be a model of $\mathcal{R}$ with respect to D and therefore $\mathcal{I}$ should implement a model-based inductive reasoning process in the spirit of (Paris, 1994). For the rest of this section let $\models_{\circ} \in \{\models^{pr}_{\varnothing}, \models^{pr}_{\odot}\}$ be one of the semantical entailment relations proposed before. In the following, we discuss some benchmark examples and state some properties that a reasonable model-based inference operator $\mathcal{I}$ should observe.

In contrast to approaches for statistical relational learning our aim for defining relational probabilistic reasoning lies in flexibility with respect to reasoning with exceptions. This demand has already been illustrated in Example 5.2 on page 5.2, see also (Delgrande, 1998). To clarify the expected behavior of $\mathcal{I}$ with respect to this example we restate it again.

Example 6.1. Consider again the knowledge base $\mathcal{R}_{zoo} = \{r_1, r_2, r_3\}$ from Example 5.2 with

$$r_1 = (likes(\mathsf{X},\mathsf{Y}) \mid elephant(\mathsf{X}) \wedge keeper(\mathsf{Y}))[0.6] \tag{6.1}$$
$$r_2 = (likes(\mathsf{X},\mathsf{fred}) \mid elephant(\mathsf{X}) \wedge keeper(\mathsf{fred}))[0.4] \tag{6.2}$$
$$r_3 = (likes(\mathsf{clyde},\mathsf{fred}) \mid elephant(\mathsf{clyde}) \wedge keeper(\mathsf{fred}))[0.7] \tag{6.3}$$

As a minimal requirement we demand $\mathcal{I}$ to agree at least with the following demands. Clearly, $\mathcal{I}$ should obey that the probability of Clyde liking Fred is 0.7 as this demand has been specified explicitly in $\mathcal{R}_{zoo}$. Furthermore, as r_2 basically represents that Fred is liked less by elephants than the average keeper there should be at least one elephant that likes Fred with approximately the represented probability. As only Fred seems to be an exceptional keeper, consequently, there should be another keeper that is liked by some elephant with roughly the probability represented by r_1.

Next, we consider a relational version of the well-known Tweety-example that has also already been discussed before.

Example 6.2. Let $\mathcal{R}_{\mathrm{birds}}$ be the knowledge base $\mathcal{R}_{\mathrm{birds}} =_{def} \{r_1, \ldots, r_6\}$ given via

$$r_1 =_{def} (bird(\mathsf{tweety}))[1] \tag{6.4}$$
$$r_2 =_{def} (bird(\mathsf{brian}))[1] \tag{6.5}$$

$$r_3 =_{def} (penguin(\mathsf{opus}))[0.9] \tag{6.6}$$
$$r_4 =_{def} (flies(\mathsf{X}) \mid bird(\mathsf{X}))[0.6] \tag{6.7}$$
$$r_5 =_{def} (flies(\mathsf{X}) \mid penguin(\mathsf{X}))[0.01] \tag{6.8}$$
$$r_6 =_{def} (bird(\mathsf{X}) \mid penguin(\mathsf{X}))[1.0] \tag{6.9}$$

The above knowledge base represents a situation where we have three individuals Tweety, Brian, and Opus and we know that both Tweety and Brian are birds (r_1 and r_2) and Opus is most likely a penguin (r_3). Furthermore, we know that birds are likely to fly (r_4), penguins are unlikely to fly (r_5), and every penguin is a bird (r_6). Note, that in difference to most formalizations of this example we are not completely sure whether Opus is a penguin. However, we expect the probability of Opus flying to be much more biased on the probability of r_5 than r_4. In particular, we expect that the probability of Opus flying should be around 0.01 and the probability of Tweety flying should be around 0.6. Furthermore, as we have represented the same knowledge for both Tweety and Brian we expect inferences on those to be the same.

The next example is inspired by (Fisseler, 2010).

Example 6.3. Let $\mathcal{R}_{\text{flu}} =_{def} \{r_1, r_2, r_3\}$ be the knowledge base given via

$$r_1 =_{def} (flu(\mathsf{X}))[0.2] \tag{6.10}$$
$$r_2 =_{def} (flu(\mathsf{X}) \mid susceptible(\mathsf{X}))[0.3] \tag{6.11}$$
$$r_3 =_{def} (flu(\mathsf{X}) \mid contact(\mathsf{X},\mathsf{Y}) \wedge flu(\mathsf{Y}))[0.4] \tag{6.12}$$

The above knowledge base models contagiosity of flu within some population. The probabilistic conditional r_1 states that in general someone catches a flu with probability 0.2, while the probabilistic conditional r_2 gives a higher probability of 0.3 to someone who is susceptible. Finally, the probabilistic conditional r_3 models a situation where someone can get infected by someone who is already infected. In particular, we expect that the probability of a specific individual Anna catching a flu is heavily influenced by the probabilities of other individuals of the domain catching a flu. More precisely, we expect that the probability of Anna catching a flu increases if she has contact with more individuals who have a high probability of catching a flu.

The next example is adapted from (Bacchus *et al.*, 1996).

Example 6.4. Consider the knowledge base $\mathcal{R}_{chirps} =_{def} \{r_1, r_2, r_3, r_4\}$ with

$$r_1 =_{def} (chirps(\mathsf{X}) \mid bird(\mathsf{X}))[0.9]$$

$$r_2 =_{def} (chirps(\mathsf{X}) \,|\, magpie(\mathsf{X}), moody(\mathsf{X}))[0.2]$$
$$r_3 =_{def} (bird(\mathsf{X}) \,|\, magpie(\mathsf{X}))[1]$$
$$r_4 =_{def} (magpie(\mathsf{tweety}))[1]$$
$$r_5 =_{def} (bird(\mathsf{huey}))[1]$$
$$r_6 =_{def} (bird(\mathsf{dewey}))[1]$$

The beliefs represented in $\mathcal{R}_{chirps}$ describe the default probabilities that a bird chirps (r_1) and that a moody magpie chirps (r_2). Knowing that every magpie is a bird (r_3) and given an actual magpie Tweety (r_4) and birds Huey and Dewey (r_5 and r_6) the question at hand is to which probability Tweety chirps. Having no belief whether Tweety is moody or not the choice of the correct "reference class" is inconclusive. However, r_2 seems to model an exception to the general rule r_1 in case of a moody magpie. As a consequence, one would expect that the "normal" state of mind of a magpie is not being moody. As it is not explicitly modeled that Tweety is moody one would expect that the probability of Tweety chirping should be biased on the probability of r_1. The minimal requirement for $\mathcal{I}$ should be to give a probability of Tweety chirping in between 0.9 and 0.2.

The above examples illustrate the expected behavior for $\mathcal{I}$ in specific situations. We now continue with a more principled approach to capture desirable properties for relational probabilistic reasoning.

One of the most basic demands for default reasoning is satisfaction of the system P properties (Makinson, 1989). These properties have been stated for abstract inference relations and serve as a "lower" bound for approaches to default reasoning. When departing from classical deduction—which is not suitable for default reasoning, cf. Section 2.1.3—the first property one has to abandon is *monotonicity*. Mainly, the system P properties aim at giving other reasonable properties that take the place of monotonicity. In our framework these properties can be stated as follows.

(Reflexivity) For all $(\psi \,|\, \phi)[d] \in \mathcal{R}$ it holds that $\mathcal{I}(\mathcal{R}, D), D \models_\circ (\psi \,|\, \phi)[d]$.

(Left Logical Equivalence) If $\mathcal{R}_1 \equiv^\circ \mathcal{R}_2$ then $\mathcal{I}(\mathcal{R}_1, D) = \mathcal{I}(\mathcal{R}_2, D)$.

(Right Weakening) If both $\mathcal{I}(\mathcal{R}, D), D \models_\circ (\psi \,|\, \phi)[d]$ and $(\psi \,|\, \phi)[d] \models^{pr} (\psi' \,|\, \phi')[d']$ then $\mathcal{I}(\mathcal{R}, D), D \models_\circ (\psi' \,|\, \phi')[d']$.

(Cumulativity) If $\mathcal{I}(\mathcal{R}, D), D \models_\circ (\psi \,|\, \phi)[d]$ then $\mathcal{I}(\mathcal{R}, D), D \models_\circ (\psi' \,|\, \phi')[d']$ if and only if $\mathcal{I}(\mathcal{R} \cup \{(\psi \,|\, \phi)[d]\}, D), D \models_\circ (\psi' \,|\, \phi')[d']$.

Note that (Cumulativity) subsumes both *cautious monotony* and *cut* (Makinson, 1989).

Our next demand for an operator $\mathcal{I}$ is a more technical one. As an inconsistent knowledge base $\mathcal{R}$ has no models and therefore an operator $\mathcal{I}$

cannot determine any model of $\mathcal{R}$ for inference, let undef be a new symbol for this case. Then we require the following property.

(Well-Definedness) It holds that $\mathcal{I}(\mathcal{R}, D) \in \mathcal{P}^{\mathrm{F}}(\Sigma)$ and $\mathcal{I}(\mathcal{R}, D), D \models_{\circ} \mathcal{R}$ if and only if D with $\mathsf{Const}(\mathcal{R}) \subseteq D \subseteq U$ is finite and $\mathcal{R}$ is $\circ$-consistent with respect to D. Otherwise it holds that $\mathcal{I}(\mathcal{R}, D) = \mathsf{undef}$.

Basically, the property (Well-Definedness) demands that if a knowledge base is consistent then model-based reasoning should be possible.

When considering knowledge bases based on a relational language the beliefs one obtains for specific individuals is of special interest. An important demand to be made is that for indistinguishable individuals, the same information should be obtained. Here, indistinguishability is defined with respect to the information expressed by $\mathcal{R}$. More specifically, if the explicit information encoded in $\mathcal{R}$ for two different constants $\mathsf{c}_1, \mathsf{c}_2 \in D$ is the same, the probability function $P = \mathcal{I}(\mathcal{R}, D)$ should treat them as indistinguishable. We formalize this indistinguishability by introducing an equivalence relation on constants. Remember that $[\cdot]$ denotes the replacement operator, cf. pages 16 and 27.

Definition 6.1 ($\mathcal{R}$-equivalence)**.** Let $\mathcal{R}$ be a knowledge base. The constants $\mathsf{c}_1, \mathsf{c}_2 \in U$ are *$\mathcal{R}$-equivalent*, denoted by $\mathsf{c}_1 \equiv_{\mathcal{R}} \mathsf{c}_2$, if and only if $\mathcal{R} = \mathcal{R}[\mathsf{c}_1 \leftrightarrow \mathsf{c}_2]$.

Observe that $\equiv_{\mathcal{R}}$ is indeed an equivalence relation, i. e., it is reflexive, transitive, and symmetric. Two $\mathcal{R}$-equivalent constants c_1 and c_2 are indistinguishable with respect to knowledge base $\mathcal{R}$. That is, $\mathcal{R}$ models exactly the same knowledge on both c_1 and c_2 and, in particular, if $\mathcal{R}$ contains a probabilistic conditional of the form $(\psi(\mathsf{c}_1) \,|\, \phi(\mathsf{c}_1))[d]$ then $\mathcal{R}$ also contains the probabilistic conditional $(\psi(\mathsf{c}_2) \,|\, \phi(\mathsf{c}_2))[d]$. Note also that every two $\mathsf{c}_1, \mathsf{c}_2 \in D$ with $\mathsf{c}_1, \mathsf{c}_2 \notin \mathsf{Const}(\mathcal{R})$ are $\mathcal{R}$-equivalent.

Definition 6.2 ($\mathcal{R}$-equivalence class)**.** A set $S \subseteq U$ is called *$\mathcal{R}$-equivalence class* if and only if $S = \{\mathsf{c}' \mid \mathsf{c}' \equiv_{\mathcal{R}} \mathsf{c}\}$ for some $\mathsf{c} \in U$. Let $\mathfrak{S}(\mathcal{R})$ denote the set of all $\mathcal{R}$-equivalence classes.

Note that the notion of $\mathcal{R}$-equivalence bears a resemblance with the notion of *reference classes* (Bacchus *et al.*, 1996) but on a purely syntactical level.

Using $\mathcal{R}$-equivalence we can state our demand for equal treatment of indistinguishable individuals as follows.

(Prototypical Indifference) Let $\mathcal{R}$ be a knowledge base, D a finite set with $\mathsf{Const}(\mathcal{R}) \subseteq D \subseteq U$ and $D \neq \emptyset$, and ψ a ground sentence. If $\mathcal{I}(\mathcal{R}, D) \neq \mathsf{undef}$ then for any $\mathsf{c}_1, \mathsf{c}_2 \in D$ with $\mathsf{c}_1 \equiv_{\mathcal{R}} \mathsf{c}_2$ it follows $\mathcal{I}(\mathcal{R}, D)(\psi) = \mathcal{I}(\mathcal{R}, D)(\psi[\mathsf{c}_1 \leftrightarrow \mathsf{c}_2])$.

The above property states that given two $\mathcal{R}$-equivalent constants $\mathsf{c}_1, \mathsf{c}_2$, i. e. $\mathsf{c}_1 \equiv_{\mathcal{R}} \mathsf{c}_2$, a sentence ϕ should have the same inferred probability as the sentence $\phi[\mathsf{c}_1 \leftrightarrow \mathsf{c}_2]$ which results in replacing c_1 with c_2 and vice versa. For example, we expect $b(\mathsf{c}_1, \mathsf{c}_2)$ to have the same probability as $b(\mathsf{c}_2, \mathsf{c}_1)$ but also $c(\mathsf{c}_1)$ to have the same probability as $c(\mathsf{c}_2)$ (if $\mathsf{c}_1 \equiv_{\mathcal{R}} \mathsf{c}_2$).

Example 6.5. Consider again Example 6.2 from page 160. There, we expect e. g. $\mathcal{I}(\mathcal{R}_{\text{birds}}, D)(\mathit{flies}(\mathsf{tweety})) = \mathcal{I}(\mathcal{R}_{\text{birds}}, D)(\mathit{flies}(\mathsf{brian}))$ as $\mathsf{tweety} \equiv_{\mathcal{R}_{\text{birds}}} \mathsf{brian}$.

Demanding satisfaction of the property (Prototypical Indifference) seems reasonable due to the *principle of symmetry*, cf. e. g. (Paris, 2000). In its most general form this principle states that "*Similar problems should have similar solutions*". For the case of $\mathcal{R}$-equivalent constants the problems of determining the probability of sentences that differ only in these constants can be regarded not just as "similar" but, in fact, as "equivalent". Following the principle of symmetry it is self-evident that the solutions to both of these problems should be similar or, in fact, identical. This means that the probabilities of both sentences should be the same.

An even more basic demand than (Prototypical Indifference) is that renaming a constant should have no impact on the information that can be derived for it.

(Name Irrelevance) Let $\mathcal{R}$ be a knowledge base, D finite with $\mathsf{Const}(\mathcal{R}) \subseteq D \subseteq U$ and $D \neq \emptyset$, $\mathsf{d} \in U \setminus D$ a constant not appearing in D, and ψ a ground sentence. If $\mathcal{I}(\mathcal{R}, D) \neq \mathsf{undef}$ then for every $\mathsf{c} \in D$, it holds that $\mathcal{I}(\mathcal{R}[\mathsf{d}/\mathsf{c}], (D \cup \{\mathsf{d}\}) \setminus \{\mathsf{c}\}) \neq \mathsf{undef}$ and $\mathcal{I}(\mathcal{R}, D)(\psi) = \mathcal{I}(\mathcal{R}[\mathsf{d}/\mathsf{c}], (D \cup \{\mathsf{d}\}) \setminus \{\mathsf{c}\})(\psi[\mathsf{d}/\mathsf{c}])$.

This property simply states that renaming a constant c in $\mathcal{R}$ to d—thus removing c from the underlying domain D but adding d—yields the same inferences. The justification for demanding this property is obvious. An inference operator should not depend on the syntactical identifier of an individual but on the role an individual plays in the knowledge base. Although (Name Irrelevance) seems to be the weaker demand, surprisingly, every function $\mathcal{I}$ satisfying (Name irrelevance) also satisfies (Prototypical Indifference).

Proposition 6.1. *If $\mathcal{I}$ satisfies (Name Irrelevance) then $\mathcal{I}$ satisfies (Prototypical Indifference).*

Proof. Let $\mathcal{R}$ be a knowledge base, D finite with $\mathsf{Const}(\mathcal{R}) \subseteq D \subseteq U$, and $\mathsf{d}_1, \mathsf{d}_2 \notin D$ with $\mathsf{d}_1 \neq \mathsf{d}_2$ (as U is infinite such d_1, d_2 exist). Let furthermore $\mathsf{c}_1, \mathsf{c}_2 \in D$ with $\mathsf{c}_1 \equiv_{\mathcal{R}} \mathsf{c}_2$ and $\mathsf{c}_1 \neq \mathsf{c}_2$. Then it holds for ground ϕ:

$$\begin{aligned}
&\mathcal{I}(\mathcal{R}, D)(\phi) \\
&\qquad = \mathcal{I}(\mathcal{R}[\mathsf{d}_1/\mathsf{c}_1], (D \cup \{\mathsf{d}_1\}) \setminus \{\mathsf{c}_1\})(\phi[\mathsf{d}_1/\mathsf{c}_1]) \\
&\qquad = \mathcal{I}(\mathcal{R}[\mathsf{d}_1/\mathsf{c}_1, \mathsf{d}_2/\mathsf{c}_2], (D \cup \{\mathsf{d}_1, \mathsf{d}_2\}) \setminus \{\mathsf{c}_1, \mathsf{c}_2\})(\phi[\mathsf{d}_1/\mathsf{c}_1, \mathsf{d}_2/\mathsf{c}_2])
\end{aligned}$$

As $\mathsf{c}_1, \mathsf{c}_2 \notin (D \cup \{\mathsf{d}_1, \mathsf{d}_2\}) \setminus \{\mathsf{c}_1, \mathsf{c}_2\}$ it holds that

$$\begin{aligned}
&\mathcal{I}(\mathcal{R}[\mathsf{d}_1/\mathsf{c}_1, \mathsf{d}_2/\mathsf{c}_2], (D \cup \{\mathsf{d}_1, \mathsf{d}_2\}) \setminus \{\mathsf{c}_1, \mathsf{c}_2\})(\phi[\mathsf{d}_1/\mathsf{c}_1, \mathsf{d}_2/\mathsf{c}_2]) \\
&\quad = \mathcal{I}(\mathcal{R}[\mathsf{d}_1/\mathsf{c}_1, \mathsf{d}_2/\mathsf{c}_2][\mathsf{c}_2/\mathsf{d}_1, \mathsf{c}_1/\mathsf{d}_2], \\
&\qquad\quad (((D \cup \{\mathsf{d}_1, \mathsf{d}_2\}) \setminus \{\mathsf{c}_1, \mathsf{c}_2\}) \cup \{\mathsf{c}_1, \mathsf{c}_2\}) \setminus \{\mathsf{d}_1, \mathsf{d}_2\}) \\
&\qquad\quad (\phi[\mathsf{d}_1/\mathsf{c}_1, \mathsf{d}_2/\mathsf{c}_2][\mathsf{c}_2/\mathsf{d}_1, \mathsf{c}_1/\mathsf{d}_2]) \quad .
\end{aligned}$$

Due to

$$\begin{aligned}
&\mathcal{R}[\mathsf{d}_1/\mathsf{c}_1, \mathsf{d}_2/\mathsf{c}_2][\mathsf{c}_2/\mathsf{d}_1, \mathsf{c}_1/\mathsf{d}_2] = \mathcal{R}[\mathsf{c}_2/\mathsf{c}_1, \mathsf{c}_1/\mathsf{c}_2] = \mathcal{R} \\
&(((D \cup \{\mathsf{d}_1, \mathsf{d}_2\}) \setminus \{\mathsf{c}_1, \mathsf{c}_2\}) \cup \{\mathsf{c}_1, \mathsf{c}_2\}) \setminus \{\mathsf{d}_1, \mathsf{d}_2\} = D
\end{aligned}$$

and

$$\phi[\mathsf{d}_1/\mathsf{c}_1, \mathsf{d}_2/\mathsf{c}_2][\mathsf{c}_2/\mathsf{d}_1, \mathsf{c}_1/\mathsf{d}_2] = \phi[\mathsf{c}_1/\mathsf{c}_2, \mathsf{c}_2/\mathsf{c}_1]$$

this yields $\mathcal{I}(\mathcal{R}, D)(\phi) = \mathcal{I}(\mathcal{R}, D)(\phi[\mathsf{c}_1/\mathsf{c}_2, \mathsf{c}_2/\mathsf{c}_1])$. □

However, (Prototypical Indifference) and (Name Irrelevance) are, in general, not equivalent as the following (artificial) example shows.

Example 6.6. Let $\mathcal{R} =_{def} \{r_1, r_2\}$ be given via

$$r_1 =_{def} (a(\mathsf{c}_1))[0.3] \qquad\qquad r_2 =_{def} (a(\mathsf{c}_2))[0.3]$$

and let $P = \mathcal{I}(\mathcal{R}, D)$ with $D = \{c_1, c_2\}$ be such that[1]

$$\begin{aligned}
P(\{a(\mathsf{c})\}) &= 0.3 \quad \text{for } \mathsf{c} \in \{\mathsf{c}_1, \mathsf{c}_2\} \\
P(\{a(\mathsf{c})\}) &= 0.0 \quad \text{for every } \mathsf{c} \notin \{\mathsf{c}_1, \mathsf{c}_2\} \\
P(\emptyset) &= 0.4 \\
P(\omega) &= 0.0 \quad \text{for every } \omega \in \Omega(\Sigma) \setminus (\{\{a(\mathsf{c})\} \mid \mathsf{c} \in U\} \cup \{\emptyset\})
\end{aligned}$$

and let $\mathcal{I}(\mathcal{R}', D') = \mathsf{undef}$ for any $\mathcal{R}' \neq \mathcal{R}$ and $D' \neq D$. In particular, $\mathcal{I}$ does not satisfy (Well-Definedness). However, note that $P = \mathcal{I}(\mathcal{R}, D)$ is in fact a probability function as $P(\{a(\mathsf{c}_1)\}) + P(\{a(\mathsf{c}_2)\}) + P(\emptyset) = 1$. Then $\mathcal{I}$ still satisfies (Prototypical Indifference) as $\mathsf{c}_1 \equiv_{\mathcal{R}} \mathsf{c}_2$ and P is indifferent with respect to c_1 and c_2 by definition, and $\mathsf{c} \equiv_{\mathcal{R}} \mathsf{c}'$ for every $\mathsf{c}, \mathsf{c}' \in U \setminus D$ and P is indifferent with respect to c and c' by definition as well.

1 $\emptyset$ is the empty Herbrand interpretation, i. e. the interpretation where each ground atom is *false*.

However, $\mathcal{I}$ does not satisfy (Name Irrelevance) as for $c \in U \setminus D$ it holds that $\mathcal{I}(\mathcal{R}[\mathsf{c}_1/\mathsf{c}], (D \cup \{\mathsf{c}\}) \setminus \{\mathsf{c}_1\}) = \mathsf{undef}$.

Although (Prototypical Indifference) is a weaker property than (Name Irrelevance) there are several generalizations that follow from (Prototypical Indifference).

Proposition 6.2. *Let $\mathcal{I}$ satisfy (Prototypical Indifference). Let $\mathcal{R}$ be a knowledge base and D finite with $\mathsf{Const}(\mathcal{R}) \subseteq D \subseteq U$ and $D \neq \emptyset$.*

1. *Let ϕ, ψ be two ground sentences. For $\mathsf{c}_1, \mathsf{c}_2 \in D$ with $\mathsf{c}_1 \equiv_{\mathcal{R}} \mathsf{c}_2$ it holds that $\mathcal{I}(\mathcal{R}, D)(\psi \mid \phi) = \mathcal{I}(\mathcal{R}, D)(\psi[\mathsf{c}_1 \leftrightarrow \mathsf{c}_2] \mid \phi[\mathsf{c}_1 \leftrightarrow \mathsf{c}_2])$.*
2. *Let $S \in \mathfrak{S}(\mathcal{R})$, $\mathsf{c}_1, \ldots, \mathsf{c}_n \in S$, and $\sigma : S \to S$ be a permutation on S, i. e. a bijective function on S. Then it holds that*

$$\mathcal{I}(\mathcal{R}, D)(\phi) = \mathcal{I}(\mathcal{R}, D)(\phi[\sigma(\mathsf{c}_1)/\mathsf{c}_1, \ldots, \sigma(\mathsf{c}_n)/\mathsf{c}_n]) \quad .$$

The proof of Proposition 6.2 can be found in Appendix A on page 243.

The following postulate focuses on the implications that a probabilistic conditional $r = (\psi(\vec{\mathsf{X}}) \mid \phi(\vec{\mathsf{X}}))[d]$ should have for the probability of a proper instantiation $P(\psi(\vec{\mathsf{a}}) \mid \phi(\vec{\mathsf{a}}))$. Our intention about r is that, in general, the conditional probability of $\psi(\vec{\mathsf{a}})$ given $\phi(\vec{\mathsf{a}})$ "should" be (around) d. But surely, we cannot guarantee that every possible instantiation r' of r conforms to a strict interpretation of this demand. This follows mainly from the fact, that using ground conditionals we should be able to give exceptions to this rule, cf. Example 6.1 on page 160. What we really want to describe when representing a population-based statement r is that *given an adequate large domain*, the respective conditional probability for constant tuples that may serve as prototypes converges towards d. This behavior resembles the intuition behind the "Law of Large Numbers" (Bernoulli, 1713).

(Convergence) Let $D_1 \subsetneq D_2 \subsetneq \ldots$ be a sequence of finite sets satisfying $\mathsf{Const}(\mathcal{R}) \subseteq D_i \subseteq U$ for all $i \in \mathbb{N}$ and $D_1 \neq \emptyset$. For a probabilistic conditional $r \in \mathcal{R}$ with $r = (\psi(\vec{\mathsf{X}}) \mid \phi(\vec{\mathsf{X}}))[d]$ let $(\psi(\vec{\mathsf{a}}) \mid \phi(\vec{\mathsf{a}}))[d]$ be a proper instantiation of r with constants $\vec{\mathsf{a}}$ that do not appear in $\mathcal{R}$. If $\mathcal{I}(\mathcal{R}, D_i) \neq \mathsf{undef}$ for $i \in \mathbb{N}$ then it holds that

$$\lim_{i \to \infty} \mathcal{I}(\mathcal{R}, D_i)(\psi(\vec{\mathsf{a}}) \mid \phi(\vec{\mathsf{a}})) = \alpha \quad .$$

The important aspect of population-based statements is their capability of expressing a general behavior within a population while allowing for exceptions. So, population-based statements are to reflect some kind of *expected value* over the set of individual instantiations that aggregates individual behaviors. As such, if the probability of one instantiation of a population-based statements lies below the probability assigned to the statement there has to be another instantiation with a probability higher than this probability value in order to compensate for the other exception.

(Compensation) Let $\mathcal{I}(\mathcal{R}, D) \neq \mathsf{undef}$ and $r \in \mathcal{R}$ a non-ground conditional with $r = (\psi(\vec{\mathsf{X}}) \,|\, \phi(\vec{\mathsf{X}}))[d]$ and $0 < d < 1$. If $\vec{\mathsf{a}_1}$ is a vector of constants such that $\mathcal{I}(\mathcal{R}, D)(\psi(\vec{\mathsf{a}_1}) \,|\, \phi(\vec{\mathsf{a}_1})) < d$ then there is another vector of constants $\vec{\mathsf{a}_2}$ with $\mathcal{I}(\mathcal{R}, D)(\psi(\vec{\mathsf{a}_2}) \,|\, \phi(\vec{\mathsf{a}_2})) > d$.

Furthermore, when considering non-ground conditionals $(\psi(\vec{\mathsf{X}}) \,|\, \phi(\vec{\mathsf{X}}))[d]$ with $d \in \{0, 1\}$ no compensation for exceptions is possible, thus requiring *direct inference* (Bacchus *et al.*, 1996) for this particular case.

(Strict Inference) Let $\mathcal{R}$ be a knowledge base and let $(\psi(\vec{\mathsf{X}}) \,|\, \phi(\vec{\mathsf{X}}))[d] \in \mathcal{R}$ be a non-ground conditional with $d \in \{0, 1\}$. If $\mathcal{I}(\mathcal{R}, D) \neq \mathsf{undef}$ then for every $(\psi(\vec{\mathsf{a}}) \,|\, \phi(\vec{\mathsf{a}})) \in \mathsf{gnd}_D((\psi(\vec{\mathsf{X}}) \,|\, \phi(\vec{\mathsf{X}})))$ it follows

$$\mathcal{I}(\mathcal{R}, D)(\psi(\vec{\mathsf{a}}) \,|\, \phi(\vec{\mathsf{a}})) = d \quad .$$

Table 7 gives an overview on the properties for inductive inference operators discussed so far.

Property	Description
(Reflexivity)	$\mathcal{I}(\mathcal{R}, D), D \models_\circ r$ for $r \in \mathcal{R}$
(Left Logical Equivalence)	If $\mathcal{R}_1 \equiv^\circ \mathcal{R}_2$ then $\mathcal{I}(\mathcal{R}_1, D) = \mathcal{I}(\mathcal{R}_2, D)$
(Right Weakening)	$\mathcal{I}(\mathcal{R}, D), D \models_\circ r$ and $r \models^{pr} r'$ implies $\mathcal{I}(\mathcal{R}, D), D \models_\circ r'$.
(Cumulativity)	$\mathcal{I}(\mathcal{R}, D), D \models_\circ r$ implies $\mathcal{I}(\mathcal{R}, D), D \models_\circ r'$ whenever $\mathcal{I}(\mathcal{R} \cup \{r\}, D), D \models_\circ r'$
(Well-Definedness)	$\mathcal{I}(\mathcal{R}, D)$ is well-defined
(Prototypical Indifference)	$\mathsf{c}_1 \equiv_{\mathcal{R}} \mathsf{c}_2$ implies $\mathcal{I}(\mathcal{R}, D)(\psi) = \mathcal{I}(\mathcal{R}, D)(\psi[\mathsf{c}_1 \leftrightarrow \mathsf{c}_2])$
(Name Irrelevance)	$\mathcal{I}(\mathcal{R}, D)(\psi) = \mathcal{I}(\mathcal{R}[\mathsf{d}/\mathsf{c}], (D \cup \{\mathsf{d}\}) \setminus \{\mathsf{c}\})(\psi[\mathsf{d}/\mathsf{c}])$ for $\mathsf{d} \in \mathcal{U} \setminus D$
(Convergence)	$\lim_{i \to \infty} \mathcal{I}(\mathcal{R}, D_i)(\psi(\vec{\mathsf{a}}) \,|\, \phi(\vec{\mathsf{a}})) = \alpha$ for $(\psi(\vec{\mathsf{X}}) \,|\, \phi(\vec{\mathsf{X}}))[d] \in \mathcal{R}$ and some $\vec{\mathsf{a}}$
(Compensation)	If $(\psi(\vec{\mathsf{X}}) \,|\, \phi(\vec{\mathsf{X}}))[d] \in \mathcal{R}$ then $\mathcal{I}(\mathcal{R}, D)(\psi(\vec{\mathsf{a}_1}) \,|\, \phi(\vec{\mathsf{a}_1})) < d$ implies $\vec{\mathsf{a}_2}$ with $\mathcal{I}(\mathcal{R}, D)(\psi(\vec{\mathsf{a}_2}) \,|\, \phi(\vec{\mathsf{a}_2})) > d$ for some $\vec{\mathsf{a}_1}, \vec{\mathsf{a}_2}$.
(Strict Inference)	$\mathcal{I}(\mathcal{R}, D)(\psi(\vec{\mathsf{a}}) \,|\, \phi(\vec{\mathsf{a}})) = d$ for all $\vec{\mathsf{a}}$ if $(\psi(\vec{\mathsf{X}}) \,|\, \phi(\vec{\mathsf{X}}))[d] \in \mathcal{R}$ with $d \in \{0, 1\}$

Table 7: Properties of inductive inference operators

6.2 PROBABILISTIC INFERENCE BY MAXIMIZING ENTROPY

In the propositional case, reasoning based on the principle of maximum entropy has proven to be a suitable approach for commonsense reasoning as it features several nice properties, cf. Section 2.3. To recall, for a propositional signature At the entropy $H(P)$ of a probability function $P : \Omega(\mathsf{At}) \to [0,1]$ is defined via

$$H(P) =_{def} - \sum_{\omega \in \Omega(\mathsf{At})} P(\omega) \operatorname{ld} P(\omega)$$

and measures the amount of indeterminateness inherent in P, cf. Definition 2.26 on page 31. By selecting the unique probability function P^* among all probabilistic models of a propositional knowledge base $\mathcal{R}$ that has maximal entropy, i. e. by computing the solution to the optimization problem

$$P^* =_{def} \mathsf{ME}(S) = \arg \max_{P \models^{pr} \mathcal{R}} H(P),$$

we get the one probability function that satisfies $\mathcal{R}$ and adds as little information as necessary. For the relational case we parametrize the entropy $H_D(P)$ of a probability function $P \in \mathcal{P}^{\mathrm{F}}(\Sigma)$ with the set $D \subseteq U$ of the constants under consideration via

$$H_D(P) =_{def} - \sum_{\omega \in \Omega(\Sigma, D)} P(\omega) \operatorname{ld} P(\omega) \quad . \tag{6.13}$$

As both our semantics require $P(\omega) = 0$ for $\omega \notin \Omega(\Sigma, D)$ in order for $P, D \models_{\circ} \mathcal{R}$ to hold, the above definition only neglects terms that are zero anyway.

As we are interested in generalizing the propositional ME-operator to the first-order case, we postulate a proper form of compatibility to the propositional ME-inference, in addition to the postulates stated for general inference operators in the previous section. Let $\mathcal{R}$ be a ground knowledge base, i. e., each $r \in \mathcal{R}$ is ground. Let At be the set of ground atoms that appear in $\mathcal{R}$. Then $\mathcal{R}$ can also be considered as a propositional knowledge base with respect to $(\mathcal{L}(\mathsf{At}) \,|\, \mathcal{L}(\mathsf{At}))^{pr}$ and the expression $\mathsf{ME}(\mathcal{R})$ is well-defined, cf. Definition 2.27 on page 32. For ground knowledge bases the operation $\mathcal{I}$ should coincide with the ME-operator.

(ME-Compatibility) Let $\mathcal{R}$ be a ground knowledge base. If ϕ is ground then $\mathsf{ME}(\mathcal{R})(\phi) = \mathcal{I}(\mathcal{R}, \mathsf{Const}(\mathcal{R}))(\phi)$.

After having introduced the averaging and the aggregating semantics for relational probabilistic knowledge bases, now we apply the maximum entropy principle to the respective model sets to single out “best” models.

6.2.1 *Averaging Inference*

In the following we define our first variant of an ME-inference $\mathcal{I}_{\varnothing}$ which bases on the averaging semantics as proposed in Section 5.3.1. A preliminary discussion of this operator can also be found in (Thimm, 2009b). By applying the principle of maximum entropy in a straightforward way we would like to define $\mathcal{I}_{\varnothing}$ via

$$\mathcal{I}_{\varnothing}(\mathcal{R}, D) = \arg \max_{P,D \models^{pr}_{\varnothing} \mathcal{R}} H_D(P) \tag{6.14}$$

with a knowledge base $\mathcal{R}$ and finite D with $\mathsf{Const}(\mathcal{R}) \subseteq D \subseteq U$ and $D \neq \emptyset$. However, this definition would presuppose that the probability function P that satisfies $\mathcal{R}$ with respect to D and maximizes $H_D(P)$ is uniquely determined. While this is the case for standard probabilistic semantics for propositional probabilistic conditional logic it is not necessarily true for averaging semantics in relational probabilistic conditional logic. In the propositional case the set of models of knowledge base is convex, i. e., for a propositional probabilistic conditional $(\psi \mid \phi)[d]$ and two probability functions P_1, P_2 with $P_1 \models^{pr} (\psi \mid \phi)[d]$ and $P_2 \models^{pr} (\psi \mid \phi)[d]$ it follows that $P_\delta \models^{pr} (\psi \mid \phi)[d]$ with $P_\delta(\omega) =_{def} \delta P_1(\omega) + (1-\delta) P_2(\omega)$ for every $\omega \in \Omega(\mathsf{At})$ and $\delta \in [0,1]$, cf. Equation (2.6) on page 28. It also holds that maximizing entropy (a strictly concave function) over a convex set has a unique solution (Boyd and Vandenberghe, 2004). In the relational case under the averaging semantics, the set of models of a relational probabilistic knowledge base $\mathcal{R}$ is not convex in general as the following example shows.

Example 6.7. Consider the knowledge base $\mathcal{R} =_{def} \{(b(\mathsf{X}) \mid a(\mathsf{X}))[0.3]\}$ and let $D = \{\mathsf{c}_1, \mathsf{c}_2\}$. Then let P_1 be a probability function that satisfies

$$P_1(b(\mathsf{c}_1)a(\mathsf{c}_1)) = 0.05 \qquad P_1(a(\mathsf{c}_1)) = 0.5$$
$$P_1(b(\mathsf{c}_2)a(\mathsf{c}_2)) = 0.01 \qquad P_1(a(\mathsf{c}_2)) = 0.02$$

It follows that $P_1(b(\mathsf{c}_1) \mid a(\mathsf{c}_1)) = 0.1$ and $P_1(b(\mathsf{c}_2) \mid a(\mathsf{c}_2)) = 0.5$ and therefore $P_1, D \models^{pr}_{\varnothing} \mathcal{R}$. Furthermore, let P_2 be a probability function that satisfies

$$P_2(b(\mathsf{c}_1)a(\mathsf{c}_1)) = 0.2 \qquad P_2(a(\mathsf{c}_1)) = 1$$
$$P_2(b(\mathsf{c}_2)a(\mathsf{c}_2)) = 0.1 \qquad P_2(a(\mathsf{c}_2)) = 0.25$$

It follows that $P_2(b(\mathsf{c}_1) \mid a(\mathsf{c}_1)) = 0.2$ and $P_2(b(\mathsf{c}_2) \mid a(\mathsf{c}_2)) = 0.4$ and therefore $P_2, D \models^{pr}_{\varnothing} \mathcal{R}$ as well. Now consider the convex combination Q of P_1 and P_2 defined via $Q(\omega) =_{def} 1/3 P_1(\omega) + 2/3 P_2(\omega)$ for all $\omega \in \Omega(\Sigma)$. Then it follows

$$Q(b(\mathsf{c}_1)a(\mathsf{c}_1)) = \frac{1}{3} P_1(b(\mathsf{c}_1)a(\mathsf{c}_1)) + \frac{2}{3} P_2(b(\mathsf{c}_1)a(\mathsf{c}_1))$$

$$= \frac{1}{3}0.05 + \frac{2}{3}0.2 = \frac{3}{20}$$

and similarly

$$Q(a(\mathsf{c}_1)) = \frac{13}{15} \qquad Q(b(\mathsf{c}_2)a(\mathsf{c}_2)) = \frac{7}{100} \qquad Q(a(\mathsf{c}_2)) = \frac{13}{75} \quad .$$

It follows that

$$Q(b(\mathsf{c}_1) \mid a(\mathsf{c}_1)) = 0.18 \qquad Q(b(\mathsf{c}_2) \mid a(\mathsf{c}_2)) = \frac{21}{52} \approx 0.4038$$

and therefore $Q, D \not\models^{pr}_{\varnothing} \mathcal{R}$.

Although the set of models of a knowledge base $\mathcal{R}$ is not convex in general, the optimization problem (6.15) may still be uniquely solvable as maximizing a strictly concave function over a convex set is only a sufficient but no necessary condition for a unique solution. As for the knowledge base $\mathcal{R}$ in the above example, $\mathcal{I}_{\varnothing}(\mathcal{R}, D)$ is indeed well-defined, cf. Proposition 6.6 on page 177. However, up until now no formal proof for and against the unique solvability of (6.14) has been found. As a consequence we take a cautious approach and define $\mathcal{I}_{\varnothing}(\mathcal{R}, D)$ via

$$\mathcal{I}_{\varnothing}(\mathcal{R}, D) =_{def} \begin{cases} \arg \max\limits_{P,D \models^{pr}_{\varnothing} \mathcal{R}} H_D(P) & \text{if } \mathcal{R} \text{ } \varnothing\text{-consistent with respect to } D \\ & \text{ and } \mathsf{Const}(\mathcal{R}) \subseteq D \text{ and } D \text{ is finite} \\ & \text{ and solution is unique} \\ \mathsf{undef} & \text{otherwise} \end{cases} \tag{6.15}$$

The second case in the above definition catches knowledge bases $\mathcal{R}$ that are $\varnothing$-inconsistent with respect to D or where the optimization problem of the first case is not uniquely solvable. Obviously, $\mathcal{I}_{\varnothing}$ is a model-based inductive inference operator using semantics $\models^{pr}_{\varnothing}$.

In the following we give some theoretical results that the proposed operator $\mathcal{I}_{\varnothing}$ fulfills several of the desired properties discussed in in the previous section. However, due to the discussion above no formal proof for the satisfaction or dissatisfaction of (Well-definedness) has been found yet.

Conjecture 6.1. *$\mathcal{I}_{\varnothing}$ satisfies (Well-definedness).*

Theorem 6.1. *$\mathcal{I}_{\varnothing}$ satisfies (Reflexivity), (Left Logical Equivalence), (Right Weakening), (Cumulativity), (Name Irrelevance), (Prototypical Indifference), (ME-Compatibility), (Compensation), and (Strict Inference). If Conjecture 6.1 is true then $\mathcal{I}_{\varnothing}$ also satisfies (Convergence).*

The proof of Theorem 6.1 can be found in Appendix A on page 244.

We continue by investigating the behavior of $\mathcal{I}_\varnothing$ on the benchmark examples from Section 6.1.

Example 6.8. We continue Example 6.1 (see page 160) and consider $D =_{def} \{\mathsf{clyde}, \mathsf{dumbo}, \mathsf{giddy}, \mathsf{fred}, \mathsf{dave}\}$. Let $\mathcal{R}_{zoo2} = \mathcal{R}_{zoo} \cup \{r_4, r_5, r_6, r_7\}$ with

$$r_4 = (elephant(\mathsf{clyde}))[1] \tag{6.16}$$
$$r_5 = (elephant(\mathsf{giddy}))[1] \tag{6.17}$$
$$r_6 = (keeper(\mathsf{fred}))[1] \tag{6.18}$$
$$r_7 = (keeper(\mathsf{dave}))[1] \tag{6.19}$$

Note that we have no belief of Dumbo being an elephant. In the following we give the probabilities of several instantiations of *likes* in $\mathcal{I}_\varnothing(\mathcal{R}_{zoo2}, D)$.

$$\mathcal{I}_\varnothing(\mathcal{R}_{zoo2}, D)(likes(\mathsf{clyde}, \mathsf{dave})) \approx 0.723 \tag{6.20}$$
$$\mathcal{I}_\varnothing(\mathcal{R}_{zoo2}, D)(likes(\mathsf{dumbo}, \mathsf{dave})) \approx 0.642 \tag{6.21}$$
$$\mathcal{I}_\varnothing(\mathcal{R}_{zoo2}, D)(likes(\mathsf{giddy}, \mathsf{dave})) \approx 0.723 \tag{6.22}$$
$$\mathcal{I}_\varnothing(\mathcal{R}_{zoo2}, D)(likes(\mathsf{clyde}, \mathsf{fred})) = 0.7 \tag{6.23}$$
$$\mathcal{I}_\varnothing(\mathcal{R}_{zoo2}, D)(likes(\mathsf{dumbo}, \mathsf{fred})) \approx 0.387 \tag{6.24}$$
$$\mathcal{I}_\varnothing(\mathcal{R}_{zoo2}, D)(likes(\mathsf{giddy}, \mathsf{fred})) \approx 0.36 \tag{6.25}$$
$$\mathcal{I}_\varnothing(\mathcal{R}_{zoo2}, D)(elephant(\mathsf{dumbo})) \approx 0.312 \tag{6.26}$$

Note how the deviations brought about by the exceptional individuals Clyde and Fred have to be balanced out by the other individuals. For example, the probabilities of the individual elephants liking Dave are greater than the probabilistic conditional r_1 specified them to be. This is because the probabilities of the elephants liking Fred is considerably smaller as demanded by the probabilistic conditional r_2. Nonetheless, the average of the conditional probabilities do indeed satisfy the conditionals in $\mathcal{R}_{zoo}$. Notice furthermore, that the probability of Dumbo being an elephant is very small—see (6.26)—considering that maximum entropy is achieved by deviating only as little as possible from the uniform probability function. But due to the interaction of the conditionals in $\mathcal{R}_{zoo}$, a smaller probability of Dumbo being an elephant is necessary in order to achieve the correct average conditional probabilities defined in the knowledge base. Thus, the belief of Dumbo being an elephant alleviates due to the premise of believing in the defined conditionals. However, the resulting probabilities (approximately) match our expectations from Example 6.1.

Example 6.9. Consider again the knowledge base $\mathcal{R}_{\text{birds}}$ from Example 6.2 (see page 160) and $D =_{def} \{\mathsf{tweety}, \mathsf{brian}, \mathsf{opus}\}$. Applying $\mathcal{I}_\varnothing$ on $\mathcal{R}_{\text{birds}}$ yields the following results on several queries:

$$\mathcal{I}_\varnothing(\mathcal{R}_{birds}, D)(flies(\mathsf{tweety})) \approx 0.84 \tag{6.27}$$
$$\mathcal{I}_\varnothing(\mathcal{R}_{birds}, D)(flies(\mathsf{brian})) \approx 0.84 \tag{6.28}$$

$$\mathcal{I}_{\varnothing}(\mathcal{R}_{birds}, D)(\mathit{flies}(\mathsf{opus})) \approx 0.12 \tag{6.29}$$
$$\mathcal{I}_{\varnothing}(\mathcal{R}_{birds}, D)(\mathit{penguin}(\mathsf{tweety})) \approx 0.079 \tag{6.30}$$
$$\mathcal{I}_{\varnothing}(\mathcal{R}_{birds}, D)(\mathit{penguin}(\mathsf{brian})) \approx 0.079 \tag{6.31}$$

Due to (Prototypical Indifference) both birds Tweety and Brian fly with a probability of 0.84, see (6.27) and (6.28). As both Tweety and Brian are birds—see (6.4) and (6.5)—this probability is slightly higher than expected, cf. (6.7). This is due to the fact that the major deviation caused by Opus has to be compensated for. Opus flies only with a probability of 0.12—see (6.29)—as it is highly believed that Opus is a penguin and penguins fly with a very small probability, cf. (6.8) and (6.9). Furthermore, both Tweety and Brian are believed to be penguins with a very small probability of 0.079, cf. (6.30) and (6.31). As our domain consists of only three birds and (6.7) demands that the average probability of a bird flying is 0.6 the possibility of Tweety and Brian being penguins diminishes.

Example 6.10. We continue Example 6.3 (see page 161). Applying $\mathcal{I}_{\varnothing}$ on $\mathcal{R}_{flu}$ yields the following results on several queries:

$$\mathcal{I}_{\varnothing}(\mathcal{R}_{flu}, D)(\mathit{flu}(\mathsf{anna})) \approx 0.2 \tag{6.32}$$
$$\mathcal{I}_{\varnothing}(\mathcal{R}_{flu}, D)(\mathit{flu}(\mathsf{anna}) \mid \mathit{contact}(\mathsf{anna}, \mathsf{bob}) \wedge \mathit{flu}(\mathsf{bob})) \approx 0.4 \tag{6.33}$$
$$\mathcal{I}_{\varnothing}(\mathcal{R}_{flu}, D)(\mathit{flu}(\mathsf{anna}) \mid \mathit{contact}(\mathsf{anna}, \mathsf{bob}) \wedge \mathit{flu}(\mathsf{bob}) \wedge \mathit{contact}(\mathsf{anna}, \mathsf{carl}) \wedge \mathit{flu}(\mathsf{carl})) \approx 0.6 \tag{6.34}$$
$$\mathcal{I}_{\varnothing}(\mathcal{R}_{flu}, D)(\mathit{contact}(\mathsf{bob}, \mathsf{carl})) \approx 0.49 \tag{6.35}$$
$$\mathcal{I}_{\varnothing}(\mathcal{R}_{flu}, D)(\mathit{contact}(\mathsf{bob}, \mathsf{carl}) \mid \mathit{flu}(\mathsf{bob}), \mathit{flu}(\mathsf{carl}))) \approx 0.657 \tag{6.36}$$

Observe that we also stated some conditional queries involving actually present evidence. Formulating queries in this form—for example considering the second query that models "what is the probability of Anna having a flu given that Anna had contact with Bob and Bob had the flu"—yields in general different inferences than adding the evidence to the knowledge base—in this case $(\mathit{contact}(\mathsf{anna}, \mathsf{bob}))[1.0]$ and $(\mathit{flu}(\mathsf{bob})[1.0]$—and querying the new knowledge base just for $\mathit{flu}(\mathsf{anna})$, cf. (Pearl, 1998) for a discussion on this topic.

The inferences drawn from $\mathcal{R}_{flu}$ using $\mathcal{I}_{\varnothing}$ resemble quite nicely the intuition behind the modeled beliefs. The probability of Anna having a flu (6.32) exactly models the expected probability when including conditional (6.10). The same is true for the probability of Anna having a flu given that Anna had contact with Bob and Bob had the flu (6.33). Furthermore, if a person had contact with multiple persons who have the flu the probability of having a flu increases (6.34). Applying the principle of maximum entropy to completely unspecified beliefs usually yields a probability function that is as close to the uniform probability function as possible. As one can see from the probability of Bob having contact with Carl (6.35) this might decrease if the corresponding formula appears in the premise of another conditional in the knowledge base (see 6.12 on page 161), see also (Paris, 1994) for a discussion. But knowing that two persons have a flu increases the probability of these two persons having contact (6.36).

Example 6.11. We continue Example 6.4 (see page 161) with $D =_{def} \{\mathsf{tweety}, \mathsf{huey}, \mathsf{dewey}\}$. Applying $\mathcal{I}_{\varnothing}$ on $\mathcal{R}_{\text{chirps}}$ yields the following results on several queries:

$$\mathcal{I}_{\varnothing}(\mathcal{R}_{\text{chirps}}, D)(\mathit{chirps}(\mathsf{tweety})) \approx 0.894 \tag{6.37}$$

$$\mathcal{I}_{\varnothing}(\mathcal{R}_{\text{chirps}}, D)(\mathit{moody}(\mathsf{tweety})) \approx 0 \tag{6.38}$$

$$\mathcal{I}_{\varnothing}(\mathcal{R}_{\text{chirps}}, D)(\mathit{magpie}(\mathsf{huey})) \approx 0.612 \tag{6.39}$$

$$\mathcal{I}_{\varnothing}(\mathcal{R}_{\text{chirps}}, D)(\mathit{chirps}(\mathsf{huey})) \approx 0.913 \tag{6.40}$$

$$\mathcal{I}_{\varnothing}(\mathcal{R}_{\text{chirps}}, D)(\mathit{moody}(\mathsf{huey})) \approx 0.432 \tag{6.41}$$

Using averaging inference the probability of Tweety being moody is approximately zero and he chirps with nearly a probability of 0.9. It follows that an implicit closed world assumption has been made on the probability of Tweety being moody. This can be explained by the large differences between the probabilities of moody magpies chirping (0.2) and the probability of ordinary bird chirping (0.9). As we have two birds that are not specified to be magpies their probability of chirping is mainly influenced by the probabilistic conditional r_1. In order to get an average probability of 0.9 for Tweety, Huey, and Dewey chirping, the probability of Tweety chirping cannot be less than 0.7. It follows that it is unlikely for Tweety being moody as moody magpies chirp only with small probability.

6.2.2 *Aggregating Inference*

In a similar manner like above, we define the ME-inference operator $\mathcal{I}_{\odot}$ that is based upon the semantics $\models^{pr}_{\odot}$. Let

$$\mathcal{I}_{\odot}(\mathcal{R}, D) =_{def} \begin{cases} \arg\max\limits_{P, D \models^{pr}_{\odot} \mathcal{R}} H_D(P) & \text{if } \mathcal{R}\ \odot\text{-consistent with respect to } D \\ & \text{ and } \mathsf{Const}(\mathcal{R}) \subseteq D \text{ and } D \text{ is finite} \\ \mathsf{undef} & \text{otherwise} \end{cases} \tag{6.42}$$

Obviously, $\mathcal{I}_{\odot}$ is a model-based inductive inference operator using semantics $\models^{pr}_{\odot}$. In this semantical context, the probabilistic conditionals from $\mathcal{R}$ induce linear constraints on the probabilities of the possible worlds so that the set of probability functions satisfying $\mathcal{R}$ forms a convex set. This makes the solution to the optimization problem (6.42) unique (if a solution exists).

Lemma 6.1. *Let $r = (\psi(\vec{\mathsf{X}}) \mid \phi(\vec{\mathsf{X}}))[d]$ be a probabilistic conditional, D finite with* $\mathsf{Const}(r) \subseteq D \subseteq U$ *and $D \neq \emptyset$, and Sol_r^D the set of probability functions that satisfy r, i. e., it holds that $Sol_r^D = \{P \mid P, D \models_{\odot}^{pr} (\psi(\vec{\mathsf{X}}) \mid \phi(\vec{\mathsf{X}}))[d]\}$. Then Sol_r is convex.*

The proof of Lemma 6.1 can be found in Appendix A on page 247.

Proposition 6.3. *Let $\mathcal{R}$ be a $\odot$-consistent knowledge base with respect to D. Then the value of $\mathcal{I}_{\odot}(\mathcal{R}, D)$ is uniquely determined.*

Proof. For any knowledge base $\mathcal{R}$ the set of probability functions that satisfy $\mathcal{R}$ is a convex set due to Lemma 6.1 and the fact that the intersection of two convex sets is again a convex set. The entropy is a strict concave function and maximization of a strict concave function over a convex set has a unique solution (Boyd and Vandenberghe, 2004). □

Due to the above Proposition $\mathcal{I}_{\odot}$ complies with all our desired properties.

Proposition 6.4. *$\mathcal{I}_{\odot}$ satisfies (Reflexivity), (Left Logical Equivalence), (Right Weakening), (Cumulativity), (Well-Definedness), (Name Irrelevance), (Prototypical Indifference), (*ME*-Compatibility), (Convergence), (Strict Inference), and (Compensation).*

The proof of Proposition 6.4 can be found in (Thimm and Kern-Isberner, 2011) and (Thimm *et al.*, 2011b). Note that, due to Lemma 6.1 and Proposition 6.3, proving (Well-Definedness) for $\mathcal{I}_{\odot}$ is easy.

In the following we apply $\mathcal{I}_{\odot}$ to the very same examples used in the previous section.

Example 6.12. We apply $\mathcal{I}_{\odot}$ onto the knowledge base $\mathcal{R}_{zoo2}$ from Example 6.8 (see page 171). This yields the following inferences:

$$\mathcal{I}_{\odot}(\mathcal{R}_{zoo2}, D)(likes(\mathsf{clyde}, \mathsf{dave})) \approx 0.8 \tag{6.43}$$
$$\mathcal{I}_{\odot}(\mathcal{R}_{zoo2}, D)(likes(\mathsf{dumbo}, \mathsf{dave})) \approx 0.64 \tag{6.44}$$
$$\mathcal{I}_{\odot}(\mathcal{R}_{zoo2}, D)(likes(\mathsf{giddy}, \mathsf{dave})) \approx 0.8 \tag{6.45}$$
$$\mathcal{I}_{\odot}(\mathcal{R}_{zoo2}, D)(likes(\mathsf{clyde}, \mathsf{fred})) = 0.7 \tag{6.46}$$
$$\mathcal{I}_{\odot}(\mathcal{R}_{zoo2}, D)(likes(\mathsf{dumbo}, \mathsf{fred})) \approx 0.356 \tag{6.47}$$
$$\mathcal{I}_{\odot}(\mathcal{R}_{zoo2}, D)(likes(\mathsf{giddy}, \mathsf{fred})) \approx 0.196 \tag{6.48}$$
$$\mathcal{I}_{\odot}(\mathcal{R}_{zoo2}, D)(elephant(\mathsf{dumbo})) \approx 0.475 \tag{6.49}$$

The results are similar to those computed by using $\mathcal{I}_{\varnothing}$ in the example above. In particular, with regard to liking Dave, both approaches calculate very similar probabilities for all individuals mentioned in the queries. Here, Dumbo—the individual not known to be an elephant—likes Dave with a lower probability than the elephants Clyde and Giddy, cf. (6.43), (6.44), and (6.45). More substantial differences can be noticed with respect to the elephants' liking the moody keeper Fred. For Giddy liking Fred, $\mathcal{I}_{\odot}$ returns a considerably lower probability than $\mathcal{I}_{\varnothing}$, see (6.48). On the other hand, $\mathcal{I}_{\odot}$ is more cautious when processing information on Dumbo, its probability of being an elephant is nearly 0.5 (6.49), while $\mathcal{I}_{\varnothing}$ suggests that Dumbo is not an elephant.

Example 6.13. We apply $\mathcal{I}_{\odot}$ onto the knowledge base $\mathcal{R}_{birds}$ from Example 6.9 (see page 171). This yields the following inferences

$$\mathcal{I}_{\odot}(\mathcal{R}_{birds}, D)(\mathit{flies}(\mathsf{tweety})) \approx 0.85 \tag{6.50}$$
$$\mathcal{I}_{\odot}(\mathcal{R}_{birds}, D)(\mathit{flies}(\mathsf{brian})) \approx 0.85 \tag{6.51}$$
$$\mathcal{I}_{\odot}(\mathcal{R}_{birds}, D)(\mathit{flies}(\mathsf{opus})) \approx 0.10 \tag{6.52}$$
$$\mathcal{I}_{\odot}(\mathcal{R}_{birds}, D)(\mathit{penguin}(\mathsf{tweety})) \approx 0.079 \tag{6.53}$$
$$\mathcal{I}_{\odot}(\mathcal{R}_{birds}, D)(\mathit{penguin}(\mathsf{brian})) \approx 0.079 \tag{6.54}$$

As in the previous example, the inferences drawn using $\mathcal{I}_{\odot}$ are very similar to the ones using $\mathcal{I}_{\varnothing}$. The probabilities of Tweety and Brian being penguins (0.079) are exactly the same as in Example 6.9. There are only minor differences in the probabilities of the instantiations of *flies*. While using $\mathcal{I}_{\varnothing}$ the probability of Tweety and Opus flying is, respectively, 0.84 and 0.12, here we have 0.85 and 0.10.

As for Example 6.3 (see page 161), applying the operator $\mathcal{I}_{\odot}$ on $\mathcal{R}_{\mathit{flu}}$ yields the exact same inferences as $\mathcal{I}_{\varnothing}$. This is due to the fact that both operators fulfill (Prototypical indifference). Consider the probabilistic conditional $(\mathit{flu}(\mathsf{X}) \,|\, \mathit{susceptible}(\mathsf{X}))[0.3] \in \mathcal{R}_{\mathit{flu}}$. As no constant is mentioned in $\mathcal{R}_{\mathit{flu}}$ all of $\mathsf{anna}, \mathsf{bob}, \mathsf{carl}$ belong to the same $\mathcal{R}_{\mathit{flu}}$-equivalence class and therefore it follows

$$\begin{aligned}
&\mathcal{I}_{\circ}(\mathcal{R}_{\mathit{flu}}, D)(\mathit{flu}(\mathsf{anna}) \,|\, \mathit{susceptible}(\mathsf{anna}))\\
&\quad = \mathcal{I}_{\circ}(\mathcal{R}_{\mathit{flu}}, D)(\mathit{flu}(\mathsf{bob}) \,|\, \mathit{susceptible}(\mathsf{bob}))\\
&\quad = \mathcal{I}_{\circ}(\mathcal{R}_{\mathit{flu}}, D)(\mathit{flu}(\mathsf{carl}) \,|\, \mathit{susceptible}(\mathsf{carl}))
\end{aligned}$$

for any $\circ \in \{\varnothing, \odot\}$ due to Proposition 6.2 on page 166. This directly yields

$$\mathcal{I}_{\circ}(\mathcal{R}_{\mathit{flu}}, D)(\mathit{flu}(\mathsf{anna}) \,|\, \mathit{susceptible}(\mathsf{anna})) = 0.3$$

for $\circ \in \{\varnothing, \odot\}$ as can also be seen in Example 6.10 on page 172. If we add probabilistic facts like $(\mathit{contact}(\mathsf{anna}, \mathsf{bob}))[1]$ or $(\mathit{flu}(\mathsf{bob}))[1]$ to $\mathcal{R}_{\mathit{flu}}$ the situation changes and now different inferences can be drawn from the different semantics. One thing to notice about this particular special case of a knowledge base—a knowledge base that mentions no constants—is that there is a direct method to reasoning. As has been discussed above due to (Prototypical Indifference) all inferences drawn from different instantiations are identical. As a result, replacing $(\mathit{flu}(\mathsf{X}) \,|\, \mathit{susceptible}(\mathsf{X}))[0.3]$ by its universal instantiations

$$(\mathit{flu}(\mathsf{anna}) \mid \mathit{susceptible}(\mathsf{anna}))[0.3]$$
$$(\mathit{flu}(\mathsf{bob}) \mid \mathit{susceptible}(\mathsf{bob}))[0.3]$$
$$(\mathit{flu}(\mathsf{carl}) \mid \mathit{susceptible}(\mathsf{carl}))[0.3]$$

amounts to the very same ME-function. We come back to this issue in the next section.

Example 6.14. We continue Example 6.4 (see page 161) with $D =_{def} \{\mathsf{tweety}, \mathsf{huey}, \mathsf{dewey}\}$. Applying $\mathcal{I}_{\odot}$ on $\mathcal{R}_{\text{chirps}}$ yields the following results on several queries:

$$\mathcal{I}_{\odot}(\mathcal{R}_{\text{chirps}}, D)(\mathit{chirps}(\mathsf{tweety})) \approx 0.863 \tag{6.55}$$
$$\mathcal{I}_{\odot}(\mathcal{R}_{\text{chirps}}, D)(\mathit{moody}(\mathsf{tweety})) \approx 0.118 \tag{6.56}$$
$$\mathcal{I}_{\odot}(\mathcal{R}_{\text{chirps}}, D)(\mathit{magpie}(\mathsf{huey})) \approx 0.361 \tag{6.57}$$
$$\mathcal{I}_{\odot}(\mathcal{R}_{\text{chirps}}, D)(\mathit{chirps}(\mathsf{huey})) \approx 0.916 \tag{6.58}$$
$$\mathcal{I}_{\odot}(\mathcal{R}_{\text{chirps}}, D)(\mathit{moody}(\mathsf{huey})) \approx 0.366 \tag{6.59}$$

The above inferences are also very similar to the ones with averaging inference, cf. Example 6.11 on page 173. The main difference is that the probability of Tweety being moody is considerably higher (0.118 compared to 0). This specific case illustrates one drawback of averaging inference. For the knowledge base $\mathcal{R}_{\text{chirps}}$ the averaging inference operator determines such a low probability of Tweety being moody because a higher probability would influence the average probability of $\mathit{chirps}(\mathsf{X})$ considerably. The higher the probability of $\mathit{moody}(\mathsf{tweety})$ the lower the probability of $\mathit{chirps}(\mathsf{tweety})$ as Tweety is the only (known) magpie in the domain and r_2 (see page 161) demands 0.2 to be the probability of moody magpies. As r_1 demands an average probability of 0.9 for $\mathit{chirps}(\mathsf{X})$ the upper bound of 1 for probabilities influences the lower bound for exceptions to r_1. As a consequence, the relatively small domain size imposes Tweety being not moody. This problem does not occur with aggregating semantics to such a large extent as the probabilities of both the conclusion and the premise together with the conclusion are aggregated first, and the conditional probability is determined afterwards. This allows for Tweety being moody with a larger probability.

6.3 ANALYSIS AND COMPARISON

One can notice that the inferences drawn from both operators are very similar. This is not surprising as, basically, both operators satisfy the desired properties which heavily restrict the choice for rational inference operators. However, this observation is quite interesting from a computational point of view as solving the optimization problems (6.15) (see page 170) and (6.42) (see page 173) require different approaches: while (6.15) is a non-convex optimization problem Equation (6.42) describes a convex optimization problem. For the latter efficient solvers are available (Boyd and Vandenberghe, 2004).

By following the observations made in Section 5.4 concerning the relationships of averaging and aggregating semantics we can also state some relationships between $\mathcal{I}_{\varnothing}$ and $\mathcal{I}_{\odot}$.

Proposition 6.5. *Let $\mathcal{R}$ be a knowledge base that consists only of probabilistic facts and ground conditionals. Then for every finite D with $\mathsf{Const}(\mathcal{R}) \subseteq D \subseteq U$ it follows $\mathcal{I}_{\varnothing}(\mathcal{R}, D) = \mathcal{I}_{\odot}(\mathcal{R}, D)$.*

Proof. By Proposition 5.1 on page 143 it follows that $\models^{pr}_{\varnothing}$ and $\models^{pr}_{\odot}$ agree on probabilistic facts, that is, for every P it holds that $P, D \models^{pr}_{\varnothing} (\psi)[d]$ whenever $P, D \models^{pr}_{\odot} (\psi)[d]$. Due to remarks 5.1 on page 137 and 5.2 on page 142 both semantics also agree on ground conditionals, i. e., it holds for every P that $P, D \models^{pr}_{\varnothing} (\psi \,|\, \phi)[d]$ whenever $P, D \models^{pr}_{\odot} (\psi \,|\, \phi)[d]$ if $(\psi \,|\, \phi)[d]$ is ground. It follows that for every P it holds that $P, D \models^{pr}_{\varnothing} \mathcal{R}$ whenever $P, D \models^{pr}_{\odot} \mathcal{R}$. Hence, the optimization problems (6.15) on page 170 and (6.42) on page 173 are defined on the same set of probability functions and it follows $\mathcal{I}_{\varnothing}(\mathcal{R}, D) = \mathcal{I}_{\odot}(\mathcal{R}, D)$. $\square$

As has already been hinted in the discussion of Example 6.3 for aggregating inference on page 175, there is also another class where averaging and aggregating inference coincides.

Proposition 6.6. *Let $\mathcal{R}$ be a knowledge base with $\mathsf{Const}(\mathcal{R}) = \emptyset$. Then for every finite $D \subseteq U$ it holds that*

$$\mathcal{I}_{\varnothing}(\mathcal{R}, D) = \mathcal{I}_{\odot}(\mathcal{R}, D) \quad .$$

Proof. Let $\mathcal{R}$ be $\odot$-consistent with respect to D and let $P_1^* = \mathcal{I}_{\odot}(\mathcal{R}, D)$. As $\mathcal{I}_{\odot}$ satisfies (Prototypical indifference) it follows that $P_1^*(\psi) = P_1^*(\psi[\mathsf{c}_1 \leftrightarrow \mathsf{c}_2])$ for every sentence ψ and every constant $\mathsf{c}_1, \mathsf{c}_2 \in D$ as all constants are in the same $\mathcal{R}$-equivalence class. In particular, for $(\psi \,|\, \phi)[d] \in \mathcal{R}$ with $\{(\psi_1 \,|\, \phi_1), \ldots, (\psi_m \,|\, \phi_m)\} = \mathsf{gnd}_D((\psi \,|\, \phi))$ it follows that $P_1^*(\psi_1\phi_1) = \ldots = P_1^*(\psi_m\phi_m)$ and $P_1^*(\phi_1) = \ldots = P_1^*(\phi_m)$. In order for $P_1^*, D \models^{pr}_{\odot} (\psi \,|\, \phi)[d]$ to hold it follows that $P_1^*(\psi_1 \,|\, \phi_1) = \ldots = P_1^*(\psi_m \,|\, \phi_m) = d$ and therefore $P_1^*, D \models^{pr}_{\varnothing} (\psi \,|\, \phi)[d]$ as well. Hence, it follows $P_1^*, D \models^{pr}_{\varnothing} \mathcal{R}$ and $\mathcal{R}$ is also $\varnothing$-consistent with respect to D. Assume $\mathcal{I}_{\varnothing}(\mathcal{R}, D) = P_2^*$ and $P_2^* \neq P_1^*$. As $\mathcal{I}_{\varnothing}$ satisfies (Prototypical Indifference) P_2^* satisfies $P_2^*(\psi) = P_2^*(\psi[\mathsf{c}_1 \leftrightarrow \mathsf{c}_2])$ for every ψ and $\mathsf{c}_1, \mathsf{c}_2 \in D$. It follows that $P_2^*, D \models^{pr}_{\odot} \mathcal{R}$ as well. Assume that $H_D(P_2^*) > H_D(P_1^*)$, then it can not be the case that $P_1^* = \mathcal{I}_{\odot}(\mathcal{R}, D)$ as P_2^* has higher entropy and $\odot$-satisfies $\mathcal{R}$. So assume $H_D(P_2^*) < H_D(P_1^*)$. But as discussed above P_1^* also $\varnothing$-satisfies $\mathcal{R}$ and therefore P_2^* cannot have maximal entropy. It follows $\mathcal{I}_{\varnothing}(\mathcal{R}, D) = \mathcal{I}_{\odot}(\mathcal{R}, D)$. $\square$

Note that the above proposition does not imply that $\varnothing$-satisfaction and $\odot$-satisfaction coincides for a knowledge base $\mathcal{R}$ with $\mathsf{Const}(\mathcal{R}) = \emptyset$. Consider again Example 5.8 on page 144 and the knowledge base $\mathcal{R}' = \{(b(\mathsf{X}) \mid a(\mathsf{X}))[0.8]\}$ which mentions no constants. In Example 5.8 it is shown that there exists a probability function P with $P, D \models^{pr}_{\odot} \mathcal{R}'$ and $P, D \not\models^{pr}_{\varnothing} \mathcal{R}'$ for some D. Therefore, $\varnothing$-satisfaction and $\odot$-satisfaction is not equivalent for this case of knowledge bases.

For the general case, $\mathcal{I}_{\varnothing}$ and $\mathcal{I}_{\odot}$ may differ significantly. The following result is a direct application of Corollary 5.1 on page 145.

Corollary 6.1. *Let $\mathcal{R}$ be some knowledge base, let D be finite with* $\mathsf{Const}(\mathcal{R}) \subseteq D \subseteq U$*, and let $(\psi \mid \phi)$ be some conditional. Then the following statements hold:*

1. *If $\mathcal{I}_{\varnothing}(\mathcal{R}, D), D \models^{pr}_{\varnothing} (\psi \mid \phi)[d_1]$ and $\mathcal{I}_{\varnothing}(\mathcal{R}, D), D \models^{pr}_{\odot} (\psi \mid \phi)[d_2]$ then*

$$|d_1 - d_2| < \frac{|\mathsf{gnd}_D^{\mathcal{I}_{\varnothing}(\mathcal{R},D)}((\psi \mid \phi))| - 1}{|\mathsf{gnd}_D^{\mathcal{I}_{\varnothing}(\mathcal{R},D)}((\psi \mid \phi))|} \quad .$$

2. *If $\mathcal{I}_{\odot}(\mathcal{R}, D), D \models^{pr}_{\varnothing} (\psi \mid \phi)[d_1']$ and $\mathcal{I}_{\odot}(\mathcal{R}, D), D \models^{pr}_{\odot} (\psi \mid \phi)[d_2']$ then*

$$|d_1' - d_2'| < \frac{|\mathsf{gnd}_D^{\mathcal{I}_{\odot}(\mathcal{R},D)}((\psi \mid \phi))| - 1}{|\mathsf{gnd}_D^{\mathcal{I}_{\odot}(\mathcal{R},D)}((\psi \mid \phi))|} \quad .$$

The above corollary shows that, in general, the inference operators $\mathcal{I}_{\varnothing}$ and $\mathcal{I}_{\odot}$ may come up with rather different probability functions for the same knowledge base. We illustrate this situation with the following example.

Example 6.15. Consider the scenario of a bird sanctuary. We know that there are exactly 1000 birds in this sanctuary, divided into two species: the striped sea eagle and the rare snoring ostrich[2]. Statistically seen, 999 of these birds are striped sea eagles and 1 of them is a snoring ostrich and no bird can be both at the same time. It is common knowledge that all striped sea eagles do fly and that snoring ostriches do not fly. Furthermore, only a few striped sea eagles are pink but every snoring ostrich is pink. This scenario can be represented as the knowledge base $\mathcal{R}_1 =_{def} \{r_1, \ldots, r_7\}$ given via

$$\begin{array}{ll} r_1 =_{def} (sse(\mathsf{X}))[0.999] & r_2 =_{def} (so(\mathsf{X}))[0.001], \\ r_3 =_{def} (sse(\mathsf{X}) \wedge so(\mathsf{X}))[0] & r_4 =_{def} (flies(\mathsf{X}) \mid sse(\mathsf{X}))[1], \\ r_5 =_{def} (flies(\mathsf{X}) \mid so(\mathsf{X}))[0] & r_6 =_{def} (pink(\mathsf{X}) \mid sse(\mathsf{X}))[0.001] \\ r_7 =_{def} (pink(\mathsf{X}) \mid so(\mathsf{X}))[1] & \end{array}$$

where $sse(\mathsf{X})$ means that X is a striped sea eagle, $so(\mathsf{X})$ means that X is a snoring ostrich, $flies(\mathsf{X})$ means that X flies, and $pink(\mathsf{X})$ means that X is

2 These species are just made up.

pink. Note that $\mathcal{R}_1$ does not mention any constants. Due to Proposition 6.6 we obtain

$$P_1 =_{def} \mathcal{I}_{\varnothing}(\mathcal{R}_1, D) = \mathcal{I}_{\odot}(\mathcal{R}_1, D)$$

for every D with $D \neq \emptyset$. The question we want to address is "*What is the probability of a pink bird flying?*", i. e., we want to assess the probability of the conditional $(\mathit{flies}(\mathsf{X}) \,|\, \mathit{pink}(\mathsf{X}))$. For $\mathcal{R}_1$ we get

$$\begin{aligned}
&P_1, D \models^{pr}_{\varnothing} (\mathit{flies}(\mathsf{X}) \,|\, \mathit{pink}(\mathsf{X}))[0.499] \\
&P_1, D \models^{pr}_{\odot} (\mathit{flies}(\mathsf{X}) \,|\, \mathit{pink}(\mathsf{X}))[0.499]
\end{aligned}$$

because $P_1(\mathit{flies}(\mathsf{a}) \,|\, \mathit{pink}(\mathsf{a})) = 0.499$ for every $\mathsf{a} \in D$. Consider now a slightly different scenario where $D =_{def} \{\mathsf{b}_1, \ldots, \mathsf{b}_{1000}\}$ is the actual set of birds in the sanctuary and let $\mathsf{b}_1, \ldots, \mathsf{b}_{999}$ be striped sea eagles and let there be a single snoring ostrich b_{1000}. This can be represented as the knowledge base $\mathcal{R}_2 =_{def} \{r'_{1,1}, \ldots, r'_{1,999}, r'_2, \ldots, r'_7\}$ given via

$$\begin{aligned}
r'_{1,i} &=_{def} (\mathit{sse}(\mathsf{b}_i))[1] \qquad \text{for } i = 1, \ldots, 999 \\
r'_2 &=_{def} (\mathit{so}(\mathsf{b}_{1000}))[1] \\
r'_3 &=_{def} (\mathit{sse}(\mathsf{X}) \wedge \mathit{so}(\mathsf{X}))[0] \\
r'_4 &=_{def} (\mathit{flies}(\mathsf{X}) \,|\, \mathit{sse}(\mathsf{X}))[1] \\
r'_5 &=_{def} (\mathit{flies}(\mathsf{X}) \,|\, \mathit{so}(\mathsf{X}))[0], \\
r'_6 &=_{def} (\mathit{pink}(\mathsf{X}) \,|\, \mathit{sse}(\mathsf{X}))[0.001] \\
r'_7 &=_{def} (\mathit{pink}(\mathsf{X}) \,|\, \mathit{so}(\mathsf{X}))[1] \quad .
\end{aligned}$$

For $\mathcal{R}_2$ we obtain

$$\begin{aligned}
&\mathcal{I}_{\varnothing}(\mathcal{R}_2, D), D \models^{pr}_{\varnothing} (\mathit{flies}(\mathsf{X}) \,|\, \mathit{pink}(\mathsf{X}))[0.999] \\
&\mathcal{I}_{\odot}(\mathcal{R}_2, D), D \models^{pr}_{\odot} (\mathit{flies}(\mathsf{X}) \,|\, \mathit{pink}(\mathsf{X}))[0.499] \quad .
\end{aligned}$$

As one can see $\mathcal{I}_{\odot}$ makes no distinction between the knowledge bases $\mathcal{R}_1$ and $\mathcal{R}_2$ with respect to the probabilistic conditional $r = (\mathit{flies}(\mathsf{X}) \,|\, \mathit{pink}(\mathsf{X}))$ and assigns the probability 0.499 in both cases. The operator $\mathcal{I}_{\varnothing}$, however, assigns a probability 0.499 to r in $\mathcal{R}_1$ and 0.999 in $\mathcal{R}_2$. On the one hand, representing the open probabilistic fact $(\mathit{sse}(\mathsf{X})[0.999]$ as the set of ground facts $(\mathit{sse}(\mathsf{b}_1))[1], \ldots, (\mathit{sse}(\mathsf{b}_{999}))[1]$ seems to be equivalent when fixing the domain D. As a consequence, an inference operator should make no distinction between $\mathcal{R}_1$ and $\mathcal{R}_2$. On the other hand, note that $\mathcal{R}_1$ and $\mathcal{R}_2$ are neither $\varnothing$- nor $\odot$-equivalent with respect to D. The knowledge base $\mathcal{R}_1$ gives no information on the actual distribution of $\mathsf{b}_1, \ldots, \mathsf{b}_{1000}$ to the different species. The operator $\mathcal{I}_{\varnothing}$ is able to recognize this difference. However, whether it is justified to assign the probability 0.999 to r in $\mathcal{R}_2$ depends on the interpretation of r from the point of view of commonsense reasoning. As for aggregating semantics the probability of r is interpreted by taking the probabilities of the premises into account as well. On the one hand, the probability of r is influenced by the probabilities of the instances $(\mathit{flies}(\mathsf{b}_1) \,|\, \mathit{pink}(\mathsf{b}_1)), \ldots, (\mathit{flies}(\mathsf{b}_{999}) \,|\, \mathit{pink}(\mathsf{b}_{999}))$ only to a small extent as the probability of the premises $\mathit{pink}(\mathsf{b}_1), \ldots, \mathit{pink}(\mathsf{b}_{999})$ is rather low (0.001 to be precise). On the other hand, the probability of r is heavily influenced by the probability of the instance $(\mathit{flies}(\mathsf{b}_{1000}) \,|\, \mathit{pink}(\mathsf{b}_{1000}))$ as the premise $\mathit{pink}(\mathsf{b}_{1000})$ has probability one. As $\mathit{pink}(\mathsf{b}_{1000})$ has such a high probability aggregating semantics classifies b_{1000} as a good "reference" for the applicability of r. Averaging semantics on the other side is not influenced by the actual probabilities of the premise. The ground conditionals $(\mathit{flies}(\mathsf{b}_1) \,|\, \mathit{pink}(\mathsf{b}_1)), \ldots, (\mathit{flies}(\mathsf{b}_{999}) \,|\, \mathit{pink}(\mathsf{b}_{999}))$ all hold with probability 1 as $\mathsf{b}_1, \ldots, \mathsf{b}_{999}$ fly independently of their color. The ground conditional $(\mathit{flies}(\mathsf{b}_{1000}) \,|\, \mathit{pink}(\mathsf{b}_{1000}))$ has probability 0 as b_{1000} does not fly independently of the color. Therefore, interpreting r as "usually, pink objects fly" on the given domain is ambiguous. Aggregating semantics acknowledges this indifference by assigning a probability of approximately 0.5 to r which is justifiable as flying objects are rarely pink and non-flying objects are always pink. However, the probability of $(\mathit{flies}(\mathsf{c}) \,|\, \mathit{pink}(\mathsf{c}))$ is one for 99.9 % of the population (for both $\mathcal{I}_{\varnothing}(\mathcal{R}_2, D)$ and $\mathcal{I}_{\odot}(\mathcal{R}_2, D)$) which also justifies assigning probability 0.999 to r.

There seems to be no definite answer to the question which of the both semantics is more appropriate for interpreting relational conditionals. Both meanings are justifiable by considering a specific perspective on their meaning. This perspective might be influenced by the actual knowledge base and the intended meaning of the probabilistic conditionals. It follows that there are knowledge bases where the averaging semantics might be more suitable than the aggregating semantics and vice versa. In particular, in the above example there are two different views which justify application of one specific semantics.

We go on by comparing both our approaches to related work in the literature.

6.4 RELATED WORK

The related work for our approaches to inference is basically the same as for the previous chapter. In particular, the most related works are (Fisseler, 2010) and (Loh *et al.*, 2010) which both use a relational extension of probabilistic conditional logic and employ the principle of maximum entropy for reasoning. Also, the work (Kern-Isberner and Lukasiewicz, 2004) falls into this category. Further related work comprises the work on first-order probabilistic logics such as (Bacchus, 1990; Halpern, 1990; Grove *et al.*, 1994; Bacchus *et al.*, 1996) and (Jaeger, 1995) and, of course, the field of statistical relational learning, cf. Section 2.4. In the following, we have a closer look on each of those works.

6.4.1 *Grounding Semantics and Maximum Entropy*

In Section 5.5.1 we already established several relationships between averaging and aggregating semantics with grounding semantics employed in the works (Fisseler, 2010; Loh *et al.*, 2010; Kern-Isberner and Lukasiewicz, 2004). As before, we focus discussion on the approach of (Loh *et al.*, 2010) as—syntactically and semantically—it subsumes the approach of relational probabilistic conditional reasoning in (Fisseler, 2010) and is similar to the one of (Kern-Isberner and Lukasiewicz, 2004). The way inference is defined in (Kern-Isberner and Lukasiewicz, 2004) differs slightly from (Loh *et al.*, 2010) as it employes an explicit notion of a *closed world assumption* but see (Loh, 2009) for a comparison of the approach of (Loh *et al.*, 2010) with (Kern-Isberner and Lukasiewicz, 2004). The work (Fisseler, 2010) also investigates the properties of reasoning based on the principle of maximum entropy in more depth. In particular, (Fisseler, 2010) presents among other things several criteria that, if satisfied by a knowledge base, allow for computing the ME-function in a simplified manner. However, we do not go into details of those features and concentrate on the semantical and inferential approach of (Loh *et al.*, 2010) which is the same as used in (Fisseler, 2010).

In the following, we restrict the approach of (Loh *et al.*, 2010) to the syntax of RPCL. In particular, we ignore the possibility to specify grounding constraints for probabilistic conditionals, cf. Section 5.5.1. The work (Loh *et al.*, 2010) also employs the principle of maximum entropy for reasoning, just like the present work. More precisely, if $\mathcal{R}$ is a knowledge base, D is finite with $\mathsf{Const}(\mathcal{R}) \subseteq D \subseteq U$ and $D \neq \emptyset$, and $\mathcal{G}$ is some grounding operator (see Definition 5.3 on page 147), then the *grounding inference operator* $\mathcal{I}_\mathcal{G}$ with respect to $\mathcal{G}$ is defined via

$$\mathcal{I}_\mathcal{G}(\mathcal{R}, D) =_{def} \arg \max_{P,D \models^{pr}_{\mathcal{G}} \mathcal{R}} H_D(P) \tag{6.60}$$

if $\mathcal{R}$ is $\mathcal{G}$-consistent with respect to D and $\mathcal{I}_\mathcal{G}(\mathcal{R}, D) =_{def} \mathsf{undef}$ otherwise. It is easy to see that $\mathcal{I}_\mathcal{G}$ satisfies several of our desired properties from Section 6.1.

Theorem 6.2. *Let $\mathcal{G}$ be some grounding operator. Then the inference operator $\mathcal{I}_\mathcal{G}$ satisfies (Left Logical Equivalence), (Right Weakening), (Cumulativity), (Well-Definedness), (Name Irrelevance), and (Prototypical Indifference).*

The proof of Theorem 6.2 can be found in Appendix A on page 248. The operator $\mathcal{I}_\mathcal{G}$ fails to satisfy (Reflexivity) in general as the following example shows.

Example 6.16. We continue Example 5.10 from page 148. Let $\mathcal{R}_{\text{zoo}} = \{r_1, r_2, r_3\}$ be given via

$$r_1 = (\mathit{likes}(\mathsf{X}, \mathsf{Y}) \,|\, \mathit{elephant}(\mathsf{X}) \wedge \mathit{keeper}(\mathsf{Y}))[0.6]$$
$$r_2 = (\mathit{likes}(\mathsf{X}, \mathsf{fred}) \,|\, \mathit{elephant}(\mathsf{X}) \wedge \mathit{keeper}(\mathsf{fred}))[0.4]$$
$$r_3 = (\mathit{likes}(\mathsf{clyde}, \mathsf{fred}) \,|\, \mathit{elephant}(\mathsf{clyde}) \wedge \mathit{keeper}(\mathsf{fred}))[0.7]$$

and let $D =_{def} \{\mathsf{clyde}, \mathsf{dumbo}, \mathsf{fred}, \mathsf{dave}\}$. Consider the specificity grounding operator $\mathcal{G}_{sp}$, cf. Example 5.10. As pointed out in Example 5.10 it holds that

$$(\mathit{likes}(\mathsf{clyde}, \mathsf{fred}) \,|\, \mathit{elephant}(\mathsf{clyde}) \wedge \mathit{keeper}(\mathsf{fred}))[0.4] \notin \mathcal{G}_{sp}(\mathcal{R}, D)$$

but also

$$(\mathit{likes}(\mathsf{clyde}, \mathsf{fred}) \,|\, \mathit{elephant}(\mathsf{clyde}) \wedge \mathit{keeper}(\mathsf{fred}))[0.7] \in \mathcal{G}_{sp}(\mathcal{R}, D) \quad .$$

It follows that $\mathcal{I}_{\mathcal{G}_{sp}}(\mathit{likes}(\mathsf{clyde}, \mathsf{fred}) \,|\, \mathit{elephant}(\mathsf{clyde}) \wedge \mathit{keeper}(\mathsf{fred})) = 0.7$ and therefore

$$\mathcal{I}_{\mathcal{G}_{sp}}(\mathcal{R}, D), D \not\models^{pr}_{\mathcal{G}_{sp}} (\mathit{likes}(\mathsf{X}, \mathsf{fred}) \,|\, \mathit{elephant}(\mathsf{X}) \wedge \mathit{keeper}(\mathsf{fred}))[0.4]$$

violating (Reflexivity).

The operator $\mathcal{I}_{\mathcal{G}}$ fails to satisfy (Reflexivity) as in determining $\mathcal{I}_{\mathcal{G}}(\mathcal{R}, D)$ the interactions of the conditionals are adhered for and specific instances of $\mathcal{G}_{sp}(\mathcal{R}, D)$ may be removed. When evaluating whether $P, D \models^{pr}_{\mathcal{G}_{sp}} r$ holds there is no such interaction and the universal instantiation of r has to be evaluated. However, this is not the case for the naive grounding operator.

Proposition 6.7. *The inference operator $\mathcal{I}_{\mathcal{G}_U}$ satisfies (Reflexivity).*

Proof. Let $\mathcal{R}$ be a knowledge base that is $\mathcal{G}_U$-consistent with respect to D. Then $\mathcal{I}_{\mathcal{G}_U}(\mathcal{R}, D), D \models^{pr}_{\mathcal{G}_U} \mathcal{G}_U(\mathcal{R}, D)$. In particular, for every $r \in \mathcal{R}$ the probability function $\mathcal{I}_{\mathcal{G}_U}(\mathcal{R}, D)$ satisfies every instance of r and hence $\mathcal{I}_{\mathcal{G}_U}(\mathcal{R}, D), D \models^{pr}_{\mathcal{G}_U} r$. □

Note that Theorem 6.2 made no assumptions on the grounding operator whatsoever. By requiring the reasonable property of $\mathcal{G}(\mathcal{R}, D) = \mathcal{R}$ for ground $\mathcal{R}$ and every D we obtain another property.

Proposition 6.8. *If $\mathcal{G}(\mathcal{R}, D) = \mathcal{R}$ for ground $\mathcal{R}$ and every D then $\mathcal{I}_{\mathcal{G}}$ satisfies (ME-compatibility).*

Proof. If $\mathcal{G}(\mathcal{R}, D) = \mathcal{R}$ then Equation (6.60) is semantically equivalent to (2.8) (see page 32) and it follows $\mathsf{ME}(\mathcal{R})(\phi) = \mathcal{I}_{\mathcal{G}}(\phi)$ for every sentence ϕ. □

In particular, the above proposition is applicable for all grounding operators considered in (Loh *et al.*, 2010) and especially for the naive and the specificity grounding operator.

Proposition 6.9. *If $\mathcal{G}$ satisfies*

$$(\psi(\vec{\mathsf{a}}) \mid \phi(\vec{\mathsf{a}}))[d] \in \mathcal{G}(\mathcal{R}, D) \quad \textit{whenever} \quad (\psi(\vec{\mathsf{X}}) \mid \phi(\vec{\mathsf{X}}))[d] \in \mathcal{R}$$

and $\vec{\mathsf{a}}$ does not appear in $\mathcal{R}$ then $\mathcal{I}_{\mathcal{G}}$ satisfies (Convergence).

Proof. As $(\psi(\vec{\mathsf{a}}) \mid \phi(\vec{\mathsf{a}}))[d] \in \mathcal{G}(\mathcal{R}, D)$ it follows $\mathcal{I}_{\mathcal{G}}(\mathcal{R}, D)(\psi(\vec{\mathsf{a}}) \mid \phi(\vec{\mathsf{a}})) = d$. When extending D this probability stays the same as $\vec{\mathsf{a}}$ does not appear in $\mathcal{R}$. ☐

Proposition 6.9 is also applicable for all grounding operators considered in (Loh *et al.*, 2010) and especially for the naive and the specificity grounding operator.

In general, $\mathcal{I}_{\mathcal{G}}$ fails to satisfy (Compensation) for every grounding operator $\mathcal{G}$ considered in (Loh *et al.*, 2010) and especially for the naive and the specificity grounding operator.

Example 6.17. Consider the knowledge base $\mathcal{R} =_{def} \{r_1, r_2\}$ given via

$$r_1 =_{def} (p(\mathsf{X}))[0.7] \qquad\qquad r_2 =_{def} (p(\mathsf{c}_1))[0.4]$$

Let $D =_{def} \{\mathsf{c}_1, \mathsf{c}_2\}$ then it holds that $\mathcal{G}_{sp}(\mathcal{R}, D) = \{(p(\mathsf{c}_1))[0.4], (p(\mathsf{c}_2))[0.7]\}$ and it follows that $\mathcal{I}_{\mathcal{G}_{sp}}(\mathcal{R}, D)(p(\mathsf{c}_1)) = 0.4 < 0.7$ but there is no c such that $\mathcal{I}_{\mathcal{G}_{sp}}(p(\mathsf{c})) > 0.7$. By defining $\mathcal{R}' = \{r'_1, r'_2\}$ via

$$r'_1 =_{def} (p(\mathsf{X}))[1] \qquad\qquad r'_2 =_{def} (p(\mathsf{c}_1))[0]$$

one can also see that $\mathcal{I}_{\mathcal{G}_{sp}}$ fails to satisfy (Strict inference) because it holds that $\mathcal{I}_{\mathcal{G}_{sp}}(\mathcal{R}, D)(p(\mathsf{c}_1)) = 0$. Note also that $\mathcal{R}'$ is $\mathcal{G}_{sp}$-consistent with respect to D.

In the previous example it has also been shown that $\mathcal{I}_{\mathcal{G}_{sp}}$ violates (Strict Inference). This is not the case for $\mathcal{I}_{\mathcal{G}_U}$.

Proposition 6.10. *The inference operator $\mathcal{I}_{\mathcal{G}_U}$ satisfies (Strict Inference).*

Proof. Let $(\psi \mid \phi)[d] \in \mathcal{R}$ with $d \in \{0, 1\}$. Then $(\psi' \mid \phi')[d] \in \mathcal{G}_U(\mathcal{R})$ for every instance $(\psi' \mid \phi')[d]$ of $(\psi \mid \phi)[d]$. it follows $\mathcal{I}_{\mathcal{G}_U}(\mathcal{R}, D)(\psi' \mid \phi') = d$. ☐

In terms of satisfaction of properties one can see that both $\mathcal{I}_{\varnothing}$ and $\mathcal{I}_{\odot}$ outperform $\mathcal{I}_{\mathcal{G}}$ in general. But by exploiting Propositions 5.3 (see page 149) and 6.6 (see page 177) we can find a class of knowledge bases where all inference operators agree.

Proposition 6.11. *Let $\mathcal{R}$ be a knowledge base with $\mathsf{Const}(\mathcal{R}) = \emptyset$. Then it holds that*

$$\mathcal{I}_{\mathcal{G}_U}(\mathcal{R}, D) = \mathcal{I}_{\varnothing}(\mathcal{R}, D) = \mathcal{I}_{\odot}(\mathcal{R}, D) \quad .$$

Proof. By Proposition 6.6 it already follows that $\mathcal{I}_{\varnothing}(\mathcal{R}, D) = \mathcal{I}_{\odot}(\mathcal{R}, D)$. Furthermore, as shown in the proof of Proposition 6.6 it follows that for $(\psi \,|\, \phi)[d] \in \mathcal{R}$ with $\{(\psi_1 \,|\, \phi_1), \ldots, (\psi_m \,|\, \phi_m)\} = \mathsf{gnd}_D((\psi \,|\, \phi))$ it holds that $\mathcal{I}_{\odot}(\mathcal{R}, D)(\psi_1 \,|\, \phi_1) = \ldots = \mathcal{I}_{\odot}(\mathcal{R}, D)(\psi_m \,|\, \phi_m) = d$. Hence, it holds that $\mathcal{I}_{\odot}(\mathcal{R}, D) \models^{pr}_{\mathcal{G}_U} \mathcal{R}$ as well. Assume $\mathcal{I}_{\odot}(\mathcal{R}, D) \neq \mathcal{I}_{\mathcal{G}_U}(\mathcal{R}, D)$. Then there is a P with $P, D \models^{pr}_{\mathcal{G}_U} \mathcal{R}$ with $H_D(P) > H_D(\mathcal{I}_{\odot}(\mathcal{R}, D))$. By Proposition 5.3 it follows that $P, D \models^{pr}_{\odot} \mathcal{R}$ as well, contradicting $\mathcal{I}_{\odot}(\mathcal{R}, D)$ to have maximal entropy. Hence, it follows $\mathcal{I}_{\odot}(\mathcal{R}, D) = \mathcal{I}_{\mathcal{G}_U}(\mathcal{R}, D) = \mathcal{I}_{\varnothing}(\mathcal{R}, D)$. In particular, if $\mathcal{R}$ is e. g. $\varnothing$-inconsistent with respect to D then $\mathcal{R}$ is also $\odot$-inconsistent and $\mathcal{G}_U$-inconsistent with respect to D (and vice versa). In this case it follows $\mathcal{I}_{\mathcal{G}_U}(\mathcal{R}, D) = \mathcal{I}_{\varnothing}(\mathcal{R}, D) = \mathcal{I}_{\odot}(\mathcal{R}, D) = \mathsf{undef}$. □

Proposition 6.11 allows for relying on the naive grounding inference operator for inference when a knowledge base mentions no constants. Employing $\mathcal{I}_{\mathcal{G}_U}$ rather than $\mathcal{I}_{\varnothing}$ or $\mathcal{I}_{\odot}$ gives computational advantages as propositional ME-reasoner such as SPIRIT (Rödder and Meyer, 1996) can be facilitated.

6.4.2 *First-order Probabilistic Logic and Random Worlds*

As already discussed in Section 5.5.2 the work (Halpern, 1990) introduces three logics that aim at augmenting first-order logic with concepts of statistical probabilities and degrees of belief. The work of (Halpern, 1990) is continued in (Grove *et al.*, 1994; Bacchus *et al.*, 1996) with the proposal of a specific account for reasoning based on the *random-worlds method*. The logical framework $\mathcal{L}^{\approx}$ of (Grove *et al.*, 1994; Bacchus *et al.*, 1996) bases on the logic $\mathcal{L}_1$ from (Halpern, 1990) but allows for approximate relations in probability terms, e. g., the formula $w_{\mathsf{X}}(\mathit{flies}(\mathsf{X}) \,|\, \mathit{bird}(\mathsf{X})) \approx 0.9$ is interpreted as "approximately 90 % of the birds in the domain do fly". Grove et. al. introduce these approximate relations in order to circumvent numerical anomalies. For example, the statement $w_{\mathsf{X}}(\mathit{flies}(\mathsf{X}) \,|\, \mathit{bird}(\mathsf{X})) = 0.9$ implies that the number of elements in the domain is a multiple of 10. The logic $\mathcal{L}^{\approx}$ can be embedded in $\mathcal{L}_1$ by translating a formula $w_{\mathsf{X}}(\mathit{flies}(\mathsf{X}) \,|\, \mathit{bird}(\mathsf{X})) \approx 0.9$ into $0.9 - \epsilon \leq w_{\mathsf{X}}(\mathit{flies}(\mathsf{X}) \,|\, \mathit{bird}(\mathsf{X})) \leq 0.9 + \epsilon$ with some "small" $\epsilon > 0$. Then the semantics[3] of $\mathcal{L}^{\approx}$ is the same as for $\mathcal{L}_1$. However, (Grove *et al.*, 1994; Bacchus *et al.*, 1996) assume that the probability function underlying the domain U_I for some interpretation $I = (U_I, f_I^U, \mathit{Pred}_I, \mathit{Func}_I) \in \mathsf{Int}(\Sigma)$ is a uniform probability function. Consequently, the notion of satisfaction can be simplified as follows. Let $\Sigma =_{def} (U, \mathit{Pred}, \mathit{Func})$ be some first-order signature with finite U. Then a formula of the form $w_{\mathsf{X}_1,\ldots,\mathsf{X}_k}(\phi) \approx \alpha$ is satisfied with

3 We simplify semantics of $\mathcal{L}^{\approx}$ a bit for ease of presentation.

respect to $\epsilon > 0$ and a first-order interpretation $I = (U_I, f_I^U, Pred_I, Func_I) \in \mathsf{Int}(\Sigma)$, denoted by $I \models_{\approx}^{\epsilon} w_{\mathsf{X}_1,\ldots,\mathsf{X}_k}(\phi) \approx \alpha$, if and only if

$$\frac{1}{|U|^k}|\{(\mathsf{c}_1,\ldots,\mathsf{c}_k) \in U^k \mid I \models_{\approx}^{\epsilon} \phi[\mathsf{X}_1/\mathsf{c}_1,\ldots,\mathsf{X}_k/\mathsf{c}_k]\}| \in [\alpha-\epsilon, \alpha+\epsilon] \quad .$$

This means that an interpretation I satisfies a formula $w_{\mathsf{X}_1,\ldots,\mathsf{X}_k}(\phi) \approx \alpha$ if "approximately" a number of $\alpha|U|^k$ different selections of constants from U of length k satisfies ϕ. For formulas in $\mathcal{L}(\Sigma, V)$ the relation $\models_{\approx}^{\epsilon}$ is the same as $\models^{\mathrm{F}}$. As an example for a formula in $\mathcal{L}^{\approx}$ consider

$$\begin{aligned}\phi_{\text{birds}} =_{def}\ & w_{\mathsf{X}}(\mathit{flies}(\mathsf{X}) \mid \mathit{bird}(\mathsf{X})) \approx 0.9 \wedge w_{\mathsf{X}}(\mathit{flies}(\mathsf{X}) \mid \mathit{penguin}(\mathsf{X})) \approx 0 \\ & \wedge (\forall \mathsf{X} : \mathit{penguin}(\mathsf{X}) \Rightarrow \mathit{bird}(\mathsf{X}))\end{aligned}$$

which represents the situation that most birds fly, nearly all penguins do not fly, and all penguins are birds.

Inference in (Grove *et al.*, 1994; Bacchus *et al.*, 1996) is defined in terms of degrees of belief with respect to a knowledge base like ϕ_{birds}. The *random-worlds* method assumes all interpretations of a first-order signature to be equally likely and computes the degree of belief of some sentence ψ given ϕ to be the ratio of the number interpretations that satisfy $\psi \wedge \phi$ and the number of interpretations that satisfy just ϕ. More precisely, let $\mathsf{Int}(\Sigma, n) \subseteq \mathsf{Int}(\Sigma)$ denote the set of first-order interpretations $I = (U_I, f_I^U, Pred_I, Func_I) \in \mathsf{Int}(\Sigma)$ such that $|U_I| = n$ and define

$$\#\mathsf{worlds}_n^{\epsilon}(\phi) =_{def} |\{I \in \mathsf{Int}(\Sigma, n) \mid I \models_{\approx}^{\epsilon} \phi\}| \tag{6.61}$$

to be the number of interpretations with domain size n that satisfy the sentence ϕ with respect to a a parameter $\epsilon > 0$. Then the *degree of belief* of a sentence ψ with respect to a sentence ϕ and domain size n, denoted by $Pr_n^{\epsilon}(\psi \mid \phi)$, is defined as

$$Pr_n^{\epsilon}(\psi \mid \phi) =_{def} \frac{\#\mathsf{worlds}_n^{\epsilon}(\psi \wedge \phi)}{\#\mathsf{worlds}_n^{\epsilon}(\phi)} \quad . \tag{6.62}$$

which is the ratio of the number worlds that satisfy $\psi \wedge \phi$ and the number of worlds that satisfy just ϕ. In contrast to our definition of aggregating semantics—see Equation (5.10) on page 142—defining the degree of belief as in Equation (6.62) assumes a uniform distribution of the interpretations as no particular probability function on the worlds is taken into account in (6.62).

In (Grove *et al.*, 1994; Bacchus *et al.*, 1996), the limit when ϵ goes to zero and n goes to infinity of this degree of belief is considered for practical purposes. Provided that some convergence criteria are satisfied this can be defined via

$$Pr_{\infty}(\psi \mid \phi) =_{def} \lim_{\epsilon \to 0} \lim_{n \to \infty} Pr_n^{\epsilon}(\psi \mid \phi) \quad ,$$

see (Grove *et al.*, 1994; Bacchus *et al.*, 1996) for technical details. For example, considering again the formula ϕ_{birds} we get

$$Pr_{\infty}(\mathit{flies}(\mathsf{tweety}) \wedge \mathit{penguin}(\mathsf{tweety}) \mid \phi_{\text{birds}}) = 0 \quad . \tag{6.63}$$

Let $\mathcal{L}_1^{\approx}$ denote the fragment of $\mathcal{L}^{\approx}$ that contains only unary predicates. In (Grove *et al.*, 1994; Bacchus *et al.*, 1996) a method based on the principle of maximum entropy is proposed that eases computing (6.63) for $\mathcal{L}_1^{\approx}$. However, the application of the entropy function differs significantly compared to our understanding. More precisely, in our work we employ entropy as a function on probability functions that gives an idea on the amount of information represented by a probability function. Bacchus et. al. consider the entropy as a function of first-order interpretations in order to classify interpretations with respect to their statistical information. Let $\Sigma =_{def} (U, Pred, Func)$ be a first-order signature such that $Pred = \{p_1, \ldots, p_m\}$ contains only unary predicates. Then an *atom* of Σ is defined to be an expression $\dot{p}_1 \ldots \dot{p}_m$ with $\dot{p}_i \in \{p_1, \neg p_1\}$. For an atom $A =_{def} \dot{p}_1 \ldots \dot{p}_m$ and a first-order interpretation $I =_{def} (U_I, f_I^U, Pred_I, Func_I) \in \mathsf{Int}(\Sigma)$ let

$$A(I) =_{def} |\{f_I^U(\mathsf{c}) \in U_I \mid I \models^{\mathrm{F}} \dot{p}_1(\mathsf{c}) \wedge \ldots \wedge \dot{p}_m(\mathsf{c})\}|$$

be the number of elements of U_I that satisfy A in I. Let $\{A_1, \ldots, A_l\}$ be the set of all atoms of Σ. Then the entropy $H(I)$ of an interpretation I is defined as

$$H(I) =_{def} -\sum_{i=1}^{l} \frac{A_i(I)}{|U_I|} \mathsf{ld} \frac{A_i(I)}{|U_I|} \quad .$$

In (Grove *et al.*, 1994; Bacchus *et al.*, 1996) it is shown that in order to determine (6.62) it suffices to consider only a subset of interpretations in (6.61), that is, the one that have "near" maximum entropy. In (Grove *et al.*, 1994; Bacchus *et al.*, 1996) it is argued that this approach fails when the signature under consideration contains non-unary predicates.

We already compared syntax and semantics of RPCL with the logic $\mathcal{L}_1$ in Section 5.5.2. The most significant drawback of $\mathcal{L}_1$ (from the point of view of our motivation) is the lack to express degrees of belief on individuals. For example, it is not possible to state a statement like "I believe with a degree of 0.7 that Tweety flies" in $\mathcal{L}_1$. The same applies to $\mathcal{L}^{\approx}$ as $\mathcal{L}^{\approx}$ can be embedded in $\mathcal{L}_1$ as discussed above. The use of approximate relations brings some benefits for knowledge representation issues but observe that RPCL can be equipped easily with a similar notion by introducing approximate relations into (5.3) (see page 137) and (5.10) (see page 142), respectively. However, $\mathcal{L}_1$ allows the representation of statistical statements such as "(approximately) 90 % of birds fly" which is also possible in RPCL but interpreted a bit softer than in the statistical setting of $\mathcal{L}^{\approx}$. Therefore, our framework allows for combining both statistical (or rather *population-based*) statements and statements on degrees of belief. If we omit representing information on particular individuals completely we obtain the following observation.

Proposition 6.12. *Let* $\mathcal{R} = \{r_1, \ldots, r_n\}$ *with* $r_i = (\psi_i(\vec{\mathsf{X}}_i) \,|\, \phi_i(\vec{\mathsf{X}}_i))[d_i]$ *for* $i = 1, \ldots, n$ *be a knowledge base and* $\mathsf{Const}(\mathcal{R}) = \emptyset$*. Let* D *be finite with* $\emptyset \neq D \subseteq U$*, then for every* $i = 1, \ldots, n$ *and every instance* $(\psi_i(\vec{\mathsf{c}}) \,|\, \phi_i(\vec{\mathsf{c}}))[d_i]$ *of* r_i *it holds that*

$$\begin{aligned}\mathcal{I}_{\emptyset}(\mathcal{R}, D)(\psi_i(\vec{\mathsf{c}}) \,|\, \phi_i(\vec{\mathsf{c}})) &= \mathcal{I}_{\odot}(\mathcal{R}, D)(\psi_i(\vec{\mathsf{c}}) \,|\, \phi_i(\vec{\mathsf{c}})) \\ &= \mathcal{I}_{\mathcal{G}_U}(\mathcal{R}, D)(\psi_i(\vec{\mathsf{c}}) \,|\, \phi_i(\vec{\mathsf{c}})) \\ &= Pr_{\infty}(\psi_i(\vec{\mathsf{c}}) \,|\, \phi_{\mathcal{R}} \wedge \phi_i(\vec{\mathsf{c}})) \\ &= d_i\end{aligned}$$

with

$$\phi_{\mathcal{R}} =_{def} w_{\vec{\mathsf{X}}_1}(\psi_1(\vec{\mathsf{X}}_1) \,|\, \phi_1(\vec{\mathsf{X}}_1)) = d_1 \wedge \ldots \wedge w_{\vec{\mathsf{X}}_n}(\psi_n(\vec{\mathsf{X}}_n) \,|\, \phi_n(\vec{\mathsf{X}}_n)) = d_n \quad .$$

Proof. The first two equalities have already been stated in Proposition 6.11 on page 184. The final one is a direct application of Theorem 4.1 in (Grove *et al.*, 1994). □

Nonetheless, reasoning bases on averaging or aggregating inference is not the same as reasoning with Pr in general. Consider the knowledge base $\mathcal{R} =_{def} \{(\mathit{flies}(\mathsf{X}))[0.8]\}$ with $D =_{def} \{\mathsf{tweety}, \mathsf{huey}, \mathsf{louie}, \mathsf{dewey}, \mathsf{opus}\}$ and

$$\psi =_{def} \mathit{flies}(\mathsf{tweety}) \wedge \mathit{flies}(\mathsf{huey}) \wedge \mathit{flies}(\mathsf{louie}) \wedge \mathit{flies}(\mathsf{dewey}) \wedge \mathit{flies}(\mathsf{opus}) \quad .$$

Then $Pr_5^{\epsilon}(\psi \,|\, \phi_{\mathcal{R}}) = 0$ (for $\epsilon < 0.2$) as for every I with $I \models_{\approx}^{\epsilon} \phi_{\mathcal{R}}$ it holds that $I \not\models^{\mathrm{F}} \phi$. However, it holds that $\mathcal{I}_{\emptyset}(\mathcal{R}, D)(\psi) = \mathcal{I}_{\odot}(\mathcal{R}, D)(\psi) \approx 0.3278$. As one can see, the random-worlds method completely neglects the possibility that all five domain elements may fly at the same time. From the perspective of the random-worlds approach this is justified as it takes a statistical perspective on knowledge bases. When statistics dictate that four of five elements fly then it is impossible that five elements may fly at the same time. This example also shows that both averaging and aggregating semantics do not have a clear statistical interpretation but mix population-based statements and statements on degrees of belief into a homogeneous view on the modeled knowledge.

In (Grove *et al.*, 1994; Bacchus *et al.*, 1996) it is also claimed that reasoning with maximum entropy fails when the signature contains non-unary predicates. This is true for the logic of (Grove *et al.*, 1994; Bacchus *et al.*, 1996) but not for our approach as the entropy is used in a completely different context which is orthogonal to the context of (Grove *et al.*, 1994; Bacchus *et al.*, 1996). The problem in extending the principle of maximum entropy in (Grove *et al.*, 1994; Bacchus *et al.*, 1996) to non-unary signatures lies in the lack of an adequate definition of the entropy of a first-order interpretation that bases on such a signature. In (Grove *et al.*, 1994; Bacchus *et al.*, 1996) the entropy of an interpretation is defined based on the distribution of the domain elements on the atoms of the signature. More precisely, the elements are partitioned into sets of constants such that each set satisfies the same atoms and the entropy is defined based on the size of those sets. Note that this concept is similar to our concept of $\mathcal{R}$-equivalence classes. However, imagine there is a binary predicate $p/2$ in the signature. Then it is hard to extend the definition of the entropy of a first-order interpretation in a meaningful manner as it is not clear how the concept of an *atom* in the meaning of (Grove *et al.*, 1994; Bacchus *et al.*, 1996) can be generalized. In particular, it is not clear whether sets of pairs of constants have to be considered or some nested classification of constants is more adequate. Bacchus et. al. conclude that the principle of maximum entropy cannot be employed for computing degrees of belief based on the random-worlds method in arbitrary signatures. While this conclusion seems to be completely justified in the context of (Grove *et al.*, 1994; Bacchus *et al.*, 1996) it should be noted that this statement does not apply to our application of the entropy function. The entropy of a probability function is well-defined even if the probability function is based on a non-unary language.

6.4.3 *Statistical Relational Learning*

The areas of statistical relational learning and probabilistic inductive logic programming are concerned with the development of frameworks that combine probabilistic reasoning and knowledge representation using first-order logic. Many of these approaches focus on learning models from data and not on knowledge representation and reasoning. In Section 2.4 we presented two representatives for approaches to statistical relational learning, namely, *Bayesian logic programs* (BLPs) and *Markov logic networks* (MLNs). In the following, we focus on comparing our framework with those two approaches.

As has already been pointed out in Section 5.1, from the perspective of knowledge representation BLPs suffer from the drawback that a complete specification of the conditional probability distribution for each clause is mandatory. This demand may become problematic when crucial information has to be represented that may be unknown. For example, consider some symptom s and some disease d. There may be probabilistic information available that links the symptom to the disease such as "if symptom s is present then the patient has disease d with probability 0.8". However, when specifying this rule as a Bayesian clause, also the probability for the statement "if symptom s is not present then patient has disease d" has to be specified. Usually, such a probability is hard to assess or may even be impossible. Our framework of RPCL does not exhibit this kind of drawback but allows the specification of both uncertain and incomplete information. Other drawbacks of BLPs include the inability of the representation of rules with exceptions and the need to specify combining rules, see also Section 5.1. Another issue that can be raised is an implicit demand on the structure of a BLP. Consider again Example 6.3 from page 161 with $\mathcal{R}_{\text{flu}} = \{r_1, r_2, r_3\}$ given via

$$r_1 = (\mathit{flu}(\mathsf{X}))[0.2]$$
$$r_2 = (\mathit{flu}(\mathsf{X}) \mid \mathit{susceptible}(\mathsf{X}))[0.3]$$
$$r_3 = (\mathit{flu}(\mathsf{X}) \mid \mathit{contact}(\mathsf{X}, \mathsf{Y}) \wedge \mathit{flu}(\mathsf{Y}))[0.4]$$

The above example cannot be represented as a BLP, even if one can provide for the missing probabilities. The problem is, that the ground Bayesian network created when grounding a BLP for some query has to be acyclic. In the example above the predicate *contact* is assumed to describe a symmetric relation: if person A had contact with person B it follows that person B had contact with person A. Given a set of evidences of the form $\{\mathit{contact}(\mathsf{anna}, \mathsf{bob}), \mathit{contact}(\mathsf{anna}, \mathsf{carl}), \mathit{susceptible}(\mathsf{anna}), \ldots\}$ it should be required that these evidences obey the symmetry of *contact*. It follows that the ground Bayesian network for some query like $\mathit{flu}(\mathsf{anna})$ is cyclic, and therefore renders inference impossible. Nonetheless, by demanding that 1.) the evidence for the corresponding BLP is not symmetric with respect to *contact* and 2.) there are no cycles in the transitive closure of *contact* instances, consistent reasoning becomes possible. But note that such restrictions may result in unexpected inferences. However, from the perspective of computational complexity both averaging and aggregating inference are far more demanding than inference with BLPs. For BLPs, existing algorithms for constructing derivation trees as in Prolog (Covington *et al.*, 1996) and both exact and approximate inference algorithms for Bayesian networks can be used. For RPCL, solving the optimization problems (6.15) (see page 170) and (6.42) (see page 173) directly is, in general, infeasible as the number of variables—i. e. the number of Herbrand interpretations—is exponential in the size of the input. Although we have a closer look on this problem in the next chapter, probabilistic inference with BLPs remains computationally more attractive.

In contrast to most approaches discussed before MLNs do not suffer from conflicts that arise in a plain grounding of the knowledge base as the weights of an MLN have no probabilistic interpretation, see (Fisseler, 2008; Thimm *et al.*, 2011a) for some discussion. Every formula in an MLN is interpreted as an influence on the probabilities of certain interpretations and the connection between weight and probability depends heavily on the weights of other formulas. As a consequence, a weighted formula may have different effects in different MLNs and (nearly) every MLN is consistent with respect to the inferential semantics given by (2.11) (see page 40). From the point of view of knowledge representation this property is problematic and unintuitive. In contrast to MLNs our approach for relational probabilistic conditionals have a clear probabilistic interpretations and consistency is well-defined.

6.5 SUMMARY AND DISCUSSION

In this chapter we discussed inductive inference in relational probabilistic conditional logic. We took a principled approach in investigating the space of possible inductive inference operators by proposing a series of rationality postulates and discussing several benchmark examples. We used the semantical notions developed in the previous chapter to develop inference operators that are based on the principle of maximum entropy. We showed that (in principle) both operators satisfy our desired properties and we analyzed and compared these operators in more depth afterwards. In the previous section, we reviewed related work and compared other approaches with our framework.

Probabilistic inference based on the principle of maximum entropy has proven to be a powerful reasoning method in propositional frameworks for knowledge representation and reasoning (Paris, 1994; Kern-Isberner, 2001). Indeed, this principle has been characterized as an optimal inference method in various frameworks and there are several axiomatic derivations, cf. e. g. (Shore and Johnson, 1980; Paris and Vencovská, 1997; Kern-Isberner, 2001). In the past ten years a lot of work has been done on lifting propositional models for probabilistic reasoning to the relational case so it seems natural to investigate the possibilities of applying the principle of maximum entropy on relational settings as we have done in this chapter. Early attempts for combining first-order logic and probabilistic reasoning had been made during the 90s by Bacchus, Halpern and colleagues, cf. Section 6.4.2. Although many of those logics are more expressive than RPCL they lack several features that are desirable from a commonsensical point of view. For example, the language $\mathcal{L}_3$ of (Halpern, 1990) allows for the specification of both statistical statements and statements on degrees of belief. However, the semantics for these two types of information is completely decoupled which results in such surprising observations like that the formula

$$(\neg\exists \mathsf{X} : p(\mathsf{X})) \wedge (w(p(\mathsf{c})) = 1)$$

is satisfiable in $\mathcal{L}_3$, cf. Section 5.5.2. Both our notions of averaging and aggregating semantics make no difference in the interpretation of population-based statements and statements on degree of belief. In RPCL only one probability function is used for interpretation and probabilities stemming from population-based statements and probabilities stemming from statements on degrees of belief influence each other yielding a complete and homogeneous picture on the domain of interest. By employing the principle of maximum entropy for reasoning we obtain an inductive inference mechanism that is both flexible and simple in usage. Furthermore, in contrast to other works on applying ME-inference in relational settings such as (Fisseler, 2010; Loh *et al.*, 2010; Kern-Isberner and Lukasiewicz, 2004) our approach does not feature a direct propositional correspondent and thus cannot be modeled with existing propositional frameworks in a concise way. On the one hand this is a drawback as we cannot employ existing reasoners for propositional ME-inference like SPIRIT (Rödder and Meyer, 1996). On the other hand this shows the advantage of our approach. Although employing a rather restricted first-order language our semantical proposals clearly extend the expressive power of other approaches and allow for the representation of complex interrelationships between different pieces of information, as shown in Section 6.4.1. Furthermore, approaches to statistical relational learning are not apt for allowing a non-monotonic reasoning behavior. The focus of this field clearly lies in machine learning tasks on structured domains that contain no exceptional individuals. From the perspective of knowledge representation both BLPs and MLNs are too restricting and their semantical notions is hard to comprehend, see also (Finthammer and Thimm, 2011) and (Thimm *et al.*, 2011a).

To summarize, in contrast to the approaches discussed in the previous section the work presented here does not treat conditionals with free variables as schemas for their instances. If $(\psi' \mid \phi')$ is an instance of $(\psi \mid \phi)$, then the actual probability in the ME-model of a ground probabilistic conditional $(\psi' \mid \phi')$ may differ significantly from the probability of some $(\psi \mid \phi)[d]$ represented in the knowledge base. Given that the underlying language contains some minimum number of constants, exceptions to a probabilistic conditional can be compensated. This allows for a great flexibility when representing relational probabilistic beliefs and is also inherently important for a non-monotonic reasoning behavior. Our approaches aim at reflecting an overall behavior within a population to which each individual contributes, while at the same time allowing individuals to deviate drastically from that behavior. In this way, both class beliefs and individual beliefs can be represented and processed within one framework.

Both operators fulfill (in principle) the catalogue of desired properties so the question remains which semantics and therefore which inference operator is the more favorable choice? And are there other reasonable possibilities for semantics and inference operators that should be investigated? Clearly, the second question cannot be answered with a "no" as none of the proposed inference operators is characterized by our desired principles for reasoning. To do so other principles have to be found that may fully characterize ME-inference in relational settings like e. g. (Shore and Johnson, 1980; Paris and Vencovská, 1997; Kern-Isberner, 2001) for propositional frameworks. As for the first question, we already discussed this question in Example 6.15 (see page 178) and showed that there may be no definite answer. From a computational point of view the operator $\mathcal{I}_{\odot}$ and thus the semantics $\models^{pr}_{\odot}$ seems to be the favorable choice for reasoning in first-order conditional logic. While Equation (6.15) (see page 170) describes a non-convex optimization problem that is hard to solve in practice, Equation (6.42) (see page 173) induces a convex optimization problem for which efficient algorithms are available (Boyd and Vandenberghe, 2004). Still, a straightforward implementation of both problems yields an exponential transformation due to the exponential number of Herbrand interpretations. Implementations of both inference operators are available within the Tweety library for artificial intelligence[4]. Performing inference using these prototypical implementations took from hours up to days for all but the smallest examples. Accordingly, there is need for a more usable algorithm to compute an ME-function based on our semantics. In the next chapter, we have a look at a more effective method for reasoning in RPCL.

4 http://sourceforge.net/projects/tweety/

7 LIFTED INFERENCE IN RPCL

In the previous chapter we discussed reasoning in RPCL based on the principle of maximum entropy. One important property of the developed approaches—and any reasonable approach for probabilistic reasoning in relational settings—is (Prototypical Indifference). Basically, this property states that if a knowledge base $\mathcal{R}$ contains exactly the same information for two different constants c_1 and c_2 then reasoning with $\mathcal{R}$ is indifferent with respect to c_1 and c_2, i. e., the probability of every statement, either a probabilistic fact or a probabilistic conditional, is the same as the probability of the same statement when c_1 and c_2 are exchanged. It follows that a probability function P^* that is obtained by an inductive reasoning mechanism that obeys (Prototypical Indifference)—such as averaging and aggregating inference—carries a lot of redundant information. For one, P^* is defined on the whole set of Herbrand interpretations $\Omega(\Sigma)$ which is infinite due to U being infinite. However, due to the demand that $P^*(\omega) \neq 0$ only for finitely many $\omega \in \Omega(\Sigma)$ only a finite number of Herbrand interpretations are needed to specify P^*. Moreover, as we defined probabilistic satisfaction for averaging and aggregating semantics via the condition that $P(\omega) = 0$ for every $\omega \in \Omega(\Sigma)$ with $\mathsf{Const}(\omega) \nsubseteq D$—see pages 137 and 142—we can restrict our discussion to probability functions defined on all $\omega \in \Omega(\Sigma)$ with $\mathsf{Const}(\omega) \subseteq D$ only. Nonetheless, as pointed out above P^* still carries a lot of redundant information due to its indifference on constants that are in the same $\mathcal{R}$-equivalence class. In this chapter we introduce *condensed probability functions* as a compact way to represent a probability function that is the result of the application of an inductive inference mechanism that obeys (Prototypical Indifference). Condensed probability functions are defined on *reference worlds* which subsume a whole set of Herbrand interpretations that, basically, model the same situation modulo exchanging constants from the same $\mathcal{R}$-equivalence class. Using reference worlds and condensed probability functions we are able to rephrase the optimization problems (6.15) and (6.42) (see pages 170 and 173, respectively) in a computationally feasible way. It turns out, that if we restrict the language under consideration to unary predicates we avoid the exponential blow-up deriving from considering all Herbrand interpretations for a finite $D \subseteq U$. Of course, considering languages that contain only unary predicates is a huge restriction but unfortunately the approach we develop does not significantly lower the computational complexity for more general languages.

The problem of encompassing the computational complexity of relational probabilistic reasoning is a relatively novel area within the field of statistical relational learning and similar research areas. Most efficient inference algorithms for current approaches to statistical relational learning rely on approximation algorithms, such as the standard inference algorithm for Markov logic networks—see Section 2.4.2—employed in the Alchemy system[1] which bases on Markov chain Monte Carlo methods (Richardson and Domingos, 2006). There are only very few approaches that discuss *exact* inference such as (Poole, 2003; de Salvo Braz *et al.*, 2005; Milch *et al.*, 2008). The approach developed in this chapter is exact as well, provided that the optimization algorithm used for maximizing entropy is exact.

This chapter is organized as follows. In Section 7.1 we discuss lifted inference in RPCL by introducing condensed probability functions and developing an approach to reason with condensed probability functions instead of ordinary probability functions. Afterwards in Section 7.2 we analyze the computational advantages we gain by employing lifted inference. In Section 7.3 we briefly discuss extending lifted inference to non-unary languages which turns out to be infeasible. We discuss related topics to lifted inference in Section 7.4 and conclude with some final remarks in Section 7.5.

7.1 LIFTED INFERENCE

In this section we develop an approach for lifted inference in RPCL. This approach bases on a compact representation of probability functions that we introduce in the subsequent section. Due to combinatorial issues the approach we develop is limited to languages that contain only predicates of arity one. As a consequence, let $\Sigma =_{def} (U, Pred, \emptyset)$ be a simple relational signature that contains only predicates of arity one. For the rest of this section, let Σ be fixed. Before giving an overview of our approach to lifted inference we discuss the problem of the standard representation of probability functions that derive from the application of an inductive inference operator that satisfies (Prototypical Indifference) by means of an example.

Example 7.1. Let $D =_{def} \{\mathsf{tweety}, \mathsf{huey}, \mathsf{dewey}, \mathsf{louie}\}$ be the domain under discourse and let $\mathcal{R}_{\text{birds2}} =_{def} \{r_1, r_2\}$ be the knowledge base given via

$$r_1 =_{def} (\mathit{flies}(\mathsf{X}))[0.8] \qquad r_2 =_{def} (\mathit{flies}(\mathsf{tweety}))[0.3] \quad .$$

Let $\mathcal{I}$ be an inference operator that satisfies (Prototypical Indifference) and let $P^* =_{def} \mathcal{I}(\mathcal{R}, D)$ with $P^* \neq \mathsf{undef}$. In general, P^* is defined on the whole set of Herbrand interpretations $\Omega(\Sigma)$ which is infinite as U is infinite. However, we are only interested in statements concerning the constants in D and therefore only Herbrand interpretations $\omega_0, \ldots, \omega_{15}$ given via

$$\omega_0 =_{def} \emptyset$$
$$\omega_1 =_{def} \{\mathit{flies}(\mathsf{tweety})\}$$
$$\omega_2 =_{def} \{\mathit{flies}(\mathsf{huey})\}$$

1 alchemy.cs.washington.edu

$$\begin{aligned}
\omega_3 &=_{def} \{\mathit{flies}(\mathsf{dewey})\}\\
\omega_4 &=_{def} \{\mathit{flies}(\mathsf{louie})\}\\
\omega_5 &=_{def} \{\mathit{flies}(\mathsf{tweety}), \mathit{flies}(\mathsf{huey})\}\\
\omega_6 &=_{def} \{\mathit{flies}(\mathsf{tweety}), \mathit{flies}(\mathsf{dewey})\}\\
\omega_7 &=_{def} \{\mathit{flies}(\mathsf{tweety}), \mathit{flies}(\mathsf{louie})\}\\
\omega_8 &=_{def} \{\mathit{flies}(\mathsf{huey}), \mathit{flies}(\mathsf{dewey})\}\\
\omega_9 &=_{def} \{\mathit{flies}(\mathsf{huey}), \mathit{flies}(\mathsf{louie})\}\\
\omega_{10} &=_{def} \{\mathit{flies}(\mathsf{dewey}), \mathit{flies}(\mathsf{louie})\}\\
\omega_{11} &=_{def} \{\mathit{flies}(\mathsf{tweety}), \mathit{flies}(\mathsf{huey}), \mathit{flies}(\mathsf{dewey})\}\\
\omega_{12} &=_{def} \{\mathit{flies}(\mathsf{tweety}), \mathit{flies}(\mathsf{huey}), \mathit{flies}(\mathsf{louie})\}\\
\omega_{13} &=_{def} \{\mathit{flies}(\mathsf{tweety}), \mathit{flies}(\mathsf{dewey}), \mathit{flies}(\mathsf{louie})\}\\
\omega_{14} &=_{def} \{\mathit{flies}(\mathsf{huey}), \mathit{flies}(\mathsf{dewey}), \mathit{flies}(\mathsf{louie})\}\\
\omega_{15} &=_{def} \{\mathit{flies}(\mathsf{tweety}), \mathit{flies}(\mathsf{huey}), \mathit{flies}(\mathsf{dewey}), \mathit{flies}(\mathsf{louie})\}
\end{aligned}$$

are of interest to us. It is easy to see, that we can rewrite P^* into a probability function Q^* that is defined on $\omega_0, \ldots, \omega_{15}$ and that satisfies $Q^*(\psi) = P^*(\psi)$ for any ground sentence ψ that does not mention any constant in $U \setminus D$ (we give the detailed definitions below). However, even Q^* needs exponential space as the number of interpretations to be considered is exponential in the number of constants $|D|$. The $\mathcal{R}$-equivalence classes $\mathfrak{S}(\mathcal{R}) = \{S_1, S_2\}$ of $\mathcal{R}$ are given via

$$S_1 = \{\mathsf{tweety}\} \qquad\qquad S_2 = \{\mathsf{huey}, \mathsf{dewey}, \mathsf{louie}\} \cup U \setminus D \quad .$$

and as P^* and Q^* are the results of an inference operator that satisfies (Prototypical Indifference) it follows that e. g. $Q^*(\psi) = Q^*(\psi[\mathsf{huey} \leftrightarrow \mathsf{dewey}])$ for every ground sentence ψ. In particular, as Herbrand interpretations can be understood as ground conjunctions we obtain

$$\begin{aligned}
Q^*(\omega_2) &= Q^*(\omega_3) = Q^*(\omega_4)\\
Q^*(\omega_5) &= Q^*(\omega_6) = Q^*(\omega_7)\\
Q^*(\omega_8) &= Q^*(\omega_9) = Q^*(\omega_{10})\\
Q^*(\omega_{11}) &= Q^*(\omega_{12}) = Q^*(\omega_{13}) \quad .
\end{aligned}$$

Therefore, it suffices to represent Q^* by only eight Herbrand interpretations—i. e., only by ω_0, ω_1, ω_2, ω_5, ω_8, ω_{11}, ω_{14}, and ω_{15}—as the other eight contain only redundant information.

In the following we elaborate on the idea suggested in the above example. First, we clarify the idea of rewriting a probability function to fit into exponential space when $D \subseteq U$ is fixed. Remember that $\Omega(\Sigma, D) = \Omega((D, Pred, \emptyset))$ is the set of *set of relevant Herbrand interpretations*, cf. Definition 5.2 on page 136.

Definition 7.1 (Expansion set)**.** Let $D \subseteq U$ be finite. Then the *expansion set* $\varrho(\omega)$ of $\omega \in \Omega(\Sigma, D)$ is defined via

$$\varrho(\omega) =_{def} \{\omega' \in \Omega(\Sigma) \mid \omega \subseteq \omega' \text{ and } \mathsf{Const}(\omega' \setminus \omega) \cap D = \emptyset\} \quad .$$

The expansion set $\varrho(\omega)$ of a Herbrand interpretation of $\omega \in \Omega(\Sigma, D)$ contains all Herbrand interpretations that agree with ω on the part relevant with respect to D. Note that it is always the case that $\omega \in \varrho(\omega)$.

Example 7.2. We continue Example 7.1. Consider $U =_{def} D \cup \{\mathsf{c}_1, \mathsf{c}_2, \ldots\}$. Then clearly $\Omega(\Sigma, D) = \{\omega_0, \ldots, \omega_{15}\}$ and e. g.

$$\begin{aligned} \varrho(\omega_1) = \{ \quad & \omega_1, \\ & \{\mathit{flies}(\mathsf{tweety}), \mathit{flies}(\mathsf{c}_1)\}, \\ & \{\mathit{flies}(\mathsf{tweety}), \mathit{flies}(\mathsf{c}_2)\}, \\ & \{\mathit{flies}(\mathsf{tweety}), \mathit{flies}(\mathsf{c}_1), \mathit{flies}(\mathsf{c}_2)\}, \ldots \quad \} \quad . \end{aligned}$$

Definition 7.2 (Focused probability function)**.** Let $P : \Omega(\Sigma) \to [0, 1]$ be a probability function and $D \subseteq U$ finite. Then the *focused probability function* $P|_D$ of P is the probability function $P|_D : \Omega(\Sigma, D) \to [0, 1]$ with

$$P|_D(\omega) =_{def} \sum_{\omega' \in \varrho(\omega)} P(\omega') \quad \text{for all } \omega \in \Omega(\Sigma, D) \quad .$$

Note that a focused probability function is a similar notion than a marginal probability function, cf. (Jaynes, 2003).

Proposition 7.1. *Let P be a probability function, $D \subseteq U$ finite, and ψ a ground sentence with $\mathsf{Const}(\psi) \subseteq D$. Then it holds that $P(\psi) = P|_D(\psi)$.*

Proof. Note that if for $\omega \in \Omega(\Sigma, D)$ it holds that $\omega \models^{\mathrm{F}} \psi$ then $\omega' \models^{\mathrm{F}} \psi$ for every $\omega' \in \varrho(\omega)$. Furthermore, if for $\omega' \in \Omega(\Sigma)$ it holds that $\omega' \models^{\mathrm{F}} \psi$ then $\omega' \in \varrho(\omega)$ for some $\omega \in \Omega(\Sigma, D)$ with $\omega \models^{\mathrm{F}} \psi$ by definition of $\varrho(\cdot)$ and the fact that $\mathsf{Const}(\psi) \subseteq D$. It follows that

$$\begin{aligned} P|_D(\psi) &= \sum_{\omega \in \Omega(\Sigma, D), \omega \models^{\mathrm{F}} \psi} P|_D(\omega) \\ &= \sum_{\omega \in \Omega(\Sigma, D), \omega \models^{\mathrm{F}} \psi} \sum_{\omega' \in \varrho(\omega)} P(\omega') \end{aligned}$$

$$= \sum_{\omega' \in \Omega(\Sigma), \omega' \models^{\mathrm{F}} \psi} P(\omega)$$
$$= P(\psi) \quad . \qquad \square$$

The above proposition states that we can reason with $\mathcal{I}(\mathcal{R}, D)|_D$ instead of $\mathcal{I}(\mathcal{R}, D)$ as $\mathcal{I}(\mathcal{R}, D)(\omega) = 0$ for every $\omega \in \Omega(\Sigma)$ with $\mathsf{Const}(\omega) \not\subseteq D$ anyway. However, $\mathcal{I}(\mathcal{R}, D)|_D$ still needs exponential space for storage.

Proposition 7.2. *It holds that* $|\Omega(\Sigma, D)| = 2^{|D||Pred|}$.

Proof. The size of the Herbrand base $\mathsf{At}(\Sigma')$ of $\Sigma' = (D, Pred, \emptyset)$ is $|D||Pred|$ as *Pred* consists only of predicates of arity one. Consequently, the number of subsets of $\mathsf{At}(\Sigma')$ is $2^{|D||Pred|}$. $\square$

In the following we introduce *condensed probability functions* that allow for an even more compact way to represent a probability function that is the result of the application of an inductive inference operator $\mathcal{I}$ that satisfies (Prototypical Indifference). For the rest of this section let $\mathcal{R}$ be a fixed knowledge base, $\mathcal{I}$ an inductive inference operator that satisfies (Prototypical Indifference), such as $\mathcal{I}_\varnothing$ or $\mathcal{I}_\odot$, D a fixed and finite set with $\mathsf{Const}(\mathcal{R}) \subseteq D \subseteq U$, and $P^* = \mathcal{I}(\mathcal{R}, D)|_D$. Let furthermore $\mathfrak{S}(\mathcal{R}) = \{S_1, \ldots, S_n\}$ be the set of $\mathcal{R}$-equivalence classes of $\mathcal{R}$, see Definition 6.2 on page 163. As we are only interested in the constants in D let $\mathfrak{S}|_D(\mathcal{R}) =_{def} \{\underline{S}_1, \ldots, \underline{S}_n\}$ be the projection of $\mathfrak{S}(\mathcal{R})$ onto D, i. e., it holds that $\underline{S}_i = S_i \cap D$ for every $i = 1, \ldots, n$.

7.1.1 *Condensed Probability Functions*

In Section 6.1 the notion of $\mathcal{R}$-equivalence has been introduced as a relation among constants, cf. Definition 6.1 on page 163. We can generalize this relation to be applicable on Herbrand interpretations as follows.

Definition 7.3 ($\mathcal{R}$-equivalence II)**.** Let $\omega_1, \omega_2 \in \Omega(\Sigma, D)$. We say that ω_1 and ω_2 are *$\mathcal{R}$-equivalent*, denoted by $\omega_1 \equiv_{\mathcal{R}} \omega_2$, if there is a set $T = \{(\mathsf{c}_1^1, \mathsf{c}_2^1), \ldots, (\mathsf{c}_1^G, \mathsf{c}_2^G)\} \subseteq \underline{S}_1 \times \underline{S}_1 \cup \ldots \cup \underline{S}_n \times \underline{S}_n$ such that

$$\omega_1 = \omega_2[\mathsf{c}_1^1 \leftrightarrow \mathsf{c}_2^1] \ldots [\mathsf{c}_1^G \leftrightarrow \mathsf{c}_2^G] \quad .$$

Basically, two Herbrand interpretations ω_1 and ω_2 are $\mathcal{R}$-equivalent if we can permute elements within each $\mathcal{R}$-equivalence class such that ω_2 becomes ω_1.

Example 7.3. We continue Example 7.1. Here we have

$$\omega_2 \equiv_{\mathcal{R}_{\text{birds2}}} \omega_3 \equiv_{\mathcal{R}_{\text{birds2}}} \omega_4$$

$$\omega_5 \equiv_{\mathcal{R}_{\text{birds2}}} \omega_6 \equiv_{\mathcal{R}_{\text{birds2}}} \omega_7$$
$$\omega_8 \equiv_{\mathcal{R}_{\text{birds2}}} \omega_9 \equiv_{\mathcal{R}_{\text{birds2}}} \omega_{10}$$
$$\omega_{11} \equiv_{\mathcal{R}_{\text{birds2}}} \omega_{12} \equiv_{\mathcal{R}_{\text{birds2}}} \omega_{13}$$

Proposition 7.3. *$\equiv_{\mathcal{R}}$ is an equivalence relation.*

The proof of Proposition 7.3 can be found in Appendix A on page 249. As $\equiv_{\mathcal{R}}$ is an equivalence relation on elements of $\Omega(\Sigma, D)$ both the $\mathcal{R}$-equivalence class $[\omega]$ of $\omega \in \Omega(\Sigma, D)$ given via

$$[\omega] =_{def} \{\omega' \in \Omega(\Sigma, D) \mid \omega \equiv_{\mathcal{R}} \omega'\}$$

and the quotient set $\Omega(\Sigma, D)/_{\equiv_{\mathcal{R}}}$ of $\Omega(\Sigma, D)$ given via

$$\Omega(\Sigma, D)/_{\equiv_{\mathcal{R}}} =_{def} \{[\omega] \mid \omega \in \Omega(\Sigma, D)\}$$

are well-defined.

Proposition 7.4. *For every $\omega_1, \omega_2 \in \Omega(\Sigma, D)$ with $\omega_1 \equiv_{\mathcal{R}} \omega_2$ it holds that $P^*(\omega_1) = P^*(\omega_2)$.*

Proof. Let A_{ω_1} and A_{ω_2} be the *complete conjunctions* characterizing ω_1 and ω_2, respectively, i. e., A_{ω_1} and A_{ω_2} are defined via (remember that $\mathsf{At}(\Sigma)$ denotes the Herbrand base with respect to the signature Σ)

$$A_{\omega_1} =_{def} \bigwedge \omega_1 \wedge \neg \bigvee \mathsf{At}((D, Pred, \emptyset)) \setminus \omega_1$$
$$A_{\omega_2} =_{def} \bigwedge \omega_2 \wedge \neg \bigvee \mathsf{At}((D, Pred, \emptyset)) \setminus \omega_2 \quad .$$

Note that both A_{ω_1} and A_{ω_2} can be written as conjunctions by applying De Morgan's Law. It follows that ω_1 is the only model that satisfies A_{ω_1} and ω_2 is the only model that satisfies A_{ω_2}. Hence, it holds that $P^*(\omega_1) = P^*(A_{\omega_1})$ and $P^*(\omega_2) = P^*(A_{\omega_2})$. For $\omega_1 \equiv_{\mathcal{R}} \omega_2$ there is a set $T = \{(\mathsf{c}_1^1, \mathsf{c}_2^1), \ldots, (\mathsf{c}_1^G, \mathsf{c}_2^G)\} \subseteq \underline{S}_1 \times \underline{S}_1 \cup \ldots \cup \underline{S}_n \times \underline{S}_n$ with

$$\omega_1 = \omega_2[\mathsf{c}_1^1 \leftrightarrow \mathsf{c}_2^1] \ldots [\mathsf{c}_1^G \leftrightarrow \mathsf{c}_2^G] \quad .$$

It follows that

$$A_{\omega_1} = A_{\omega_2}[\mathsf{c}_1^1 \leftrightarrow \mathsf{c}_2^1] \ldots [\mathsf{c}_1^G \leftrightarrow \mathsf{c}_2^G]$$

holds as well (note that applying the above replacement on the not-negated part of A_{ω_2} results in the not-negated part of A_{ω_1}; as both A_{ω_1} and A_{ω_2} list *every* ground atom and a replacement maintains the structure of a formula the negated part of A_{ω_2} has to become the negated part of A_{ω_1}; otherwise the not-negated part has to change as well). Via iterative application of 2.) in Proposition 6.2 (see page 166) it follows that $P^*(A_{\omega_1}) = P^*(A_{\omega_2})$

and therefore $P^*(\omega_1) = P^*(\omega_2)$ (note that every replacement above only substitutes constants with constants of the same $\mathcal{R}$-equivalence class). $\square$

The above proposition shows that the probability function P^* carries a lot of redundant information stemming from the $\mathcal{R}$-equivalence of certain Herbrand interpretations. In the following, we exploit this observation by using $\Omega(\Sigma, D)/_{\equiv_\mathcal{R}}$ instead of $\Omega(\Sigma, D)$ for redefining P^*. To do so, we go on by developing a method that enumerates the elements of $\Omega(\Sigma, D)/_{\equiv_\mathcal{R}}$ in an efficient way.

For the rest of this section let $Pred =_{def} \{p_1, \ldots, p_P\}$ be the set of (unary) predicates of Σ.

Definition 7.4 (Truth configuration)**.** A *truth configuration* t for $p_1, \ldots, p_P$ is an expression $t =_{def} \dot{p}_1 \ldots \dot{p}_P$ with $\dot{p}_i \in \{p_i, \overline{p}_i\}$ for $i = 1, \ldots, P$. Let Θ denote the set of all truth configurations.

A truth configuration describes a specific situation that can be assigned to a constant c. Thus, a truth configuration that is assigned to c characterizes the state of c in some interpretation as it enumerates which attributes (unary predicates) apply for c and which do not. The concept of truth configurations has also been employed under the notion of *atom* in (Grove *et al.*, 1994; Bacchus *et al.*, 1996), see also Section 6.4.2. For a constant c and a truth configuration $t = \dot{p}_1 \ldots \dot{p}_P$ define

$$\begin{aligned} t(\mathsf{c}) &=_{def} \{\dot{p}_1(\mathsf{c}), \ldots, \dot{p}_P(\mathsf{c})\} \\ t^\wedge(\mathsf{c}) &=_{def} \dot{p}_1(\mathsf{c}) \wedge \ldots \wedge \dot{p}_P(\mathsf{c}) \\ t^+(\mathsf{c}) &=_{def} t(\mathsf{c}) \cap \mathsf{At}((D, Pred, \emptyset)) \\ t^-(\mathsf{c}) &=_{def} t(\mathsf{c}) \setminus t^+(\mathsf{c}) \quad . \end{aligned}$$

Furthermore, for a ground sentence ϕ and constants $\mathsf{c}_1, \ldots, \mathsf{c}_n$ let

$$\begin{aligned} \Theta(\phi, \mathsf{c}_1) &=_{def} \{t \in \Theta \mid t^\wedge(\mathsf{c}_1) \wedge \phi \not\models^{\mathrm{F}} \bot\} \\ \Theta(\phi, \mathsf{c}_1, \ldots, \mathsf{c}_n) &=_{def} \Theta(\phi, \mathsf{c}_1) \times \ldots \times \Theta(\phi, \mathsf{c}_n) \quad . \end{aligned}$$

The set $\Theta(\phi, \mathsf{c}_1)$ contains all those truth configurations t for a constant c_1 that are compatible with some sentence ϕ. The set $\Theta(\phi, \mathsf{c}_1, \ldots, \mathsf{c}_n)$ extends this notion to tuples of constants.

Example 7.4. Let $Pred =_{def} \{p_1/1, p_2/1\}$ and let $\psi =_{def} p_1(\mathsf{c}) \wedge (p_2(\mathsf{c}) \vee p_2(\mathsf{d}))$. Then it holds that $\Theta(\phi, \mathsf{c}) = \{p_1 p_2, p_1 \overline{p}_2\}$.

Definition 7.5 (Instance assignment)**.** An *instance assignment* I is a function $I : \mathfrak{S}|_D(\mathcal{R}) \to \mathbb{N}_0$ that satisfies $I(\underline{S}_i) \leq |\underline{S}_i|$ for all $i = 1, \ldots, n$. Let $\mathfrak{I}$ denote the set of all instance assignments.

Definition 7.6 (Reference world)**.** A *reference world* $\hat{\omega}$ is a function $\hat{\omega} : \Theta \to \mathfrak{I}$ that satisfies

$$\sum_{t \in \Theta} \hat{\omega}(t)(S_i) = |S_i| \quad \text{(for all } i = 1, \ldots, n\text{)} \quad . \tag{7.1}$$

Let $\hat{\Omega}$ be the set of all reference worlds.

Basically, a reference world is a function that maps a truth configuration to the number of constants of each $\mathcal{R}$-equivalence class that satisfy this truth configuration. As we show later, a reference world is a compact representation of $[\omega]$ for some $\omega \in \Omega(\Sigma, D)$. The normalization constraint (7.1) ensures that each constant in D is assigned exactly one truth configuration.

Example 7.5. We continue Example 7.1. The set of truth configurations $\Theta = \{t_1, t_2\}$ with respect to D and $\mathcal{R}$ is given via

$$t_1 = \mathit{flies} \qquad\qquad t_2 = \overline{\mathit{flies}} \quad .$$

Furthermore, the set of instance assignments $\mathfrak{I} = \{I_1, \ldots, I_8\}$ with respect to D and $\mathcal{R}$ is given via

$$\begin{array}{llll}
I_1(\underline{S}_1) = 0 & I_1(\underline{S}_2) = 0 & I_2(\underline{S}_1) = 0 & I_2(\underline{S}_2) = 1 \\
I_3(\underline{S}_1) = 0 & I_3(\underline{S}_2) = 2 & I_4(\underline{S}_1) = 0 & I_4(\underline{S}_2) = 3 \\
I_5(\underline{S}_1) = 1 & I_5(\underline{S}_2) = 0 & I_6(\underline{S}_1) = 1 & I_6(\underline{S}_2) = 1 \\
I_7(\underline{S}_1) = 1 & I_7(\underline{S}_2) = 2 & I_8(\underline{S}_1) = 1 & I_8(\underline{S}_2) = 3 \quad .
\end{array}$$

Finally, the set $\hat{\Omega} = \{\hat{\omega}_1, \ldots, \hat{\omega}_8\}$ of reference worlds with respect to D and $\mathcal{R}$ is given via

$$\begin{array}{llll}
\hat{\omega}_1(t_1) = I_1 & \hat{\omega}_1(t_2) = I_8 & \hat{\omega}_2(t_1) = I_2 & \hat{\omega}_2(t_2) = I_7 \\
\hat{\omega}_3(t_1) = I_3 & \hat{\omega}_3(t_2) = I_6 & \hat{\omega}_4(t_1) = I_4 & \hat{\omega}_4(t_2) = I_5 \\
\hat{\omega}_5(t_1) = I_5 & \hat{\omega}_5(t_2) = I_4 & \hat{\omega}_6(t_1) = I_6 & \hat{\omega}_6(t_2) = I_3 \\
\hat{\omega}_7(t_1) = I_7 & \hat{\omega}_7(t_2) = I_2 & \hat{\omega}_8(t_1) = I_8 & \hat{\omega}_8(t_2) = I_1 \quad .
\end{array}$$

For example, the intuitive description of the reference world $\hat{\omega}_3$ is that $\hat{\omega}_3$ represents a state where the one element of $\underline{S}_1$ does not fly and two elements of $\underline{S}_2$ do fly. Therefore $\hat{\omega}_3$ subsumes the Herbrand interpretations ω_8, ω_9, and ω_{10} from Example 7.1.

In the following we show that $\hat{\Omega}$ is indeed a characterization of the quotient set $\Omega(\Sigma, D)/_{\equiv_{\mathcal{R}}}$. For that, consider the following definition.

Definition 7.7 (Equivalence mapping)**.** The *equivalence mapping* κ is the function $\kappa : \Omega(\Sigma, D) \to \hat{\Omega}$ defined as $\kappa(\omega) =_{def} \hat{\omega}$ with

$$\hat{\omega}(\dot{p}_1 \dots \dot{p}_P)(\underline{S}_i) =_{def} |\{\mathsf{c} \in \underline{S}_i \mid \omega \models^{\mathrm{F}} \dot{p}_1(\mathsf{c}) \wedge \dots \wedge \dot{p}_P(\mathsf{c})\}|$$

for every $\dot{p}_1 \dots \dot{p}_P \in \Theta$ and $i = 1, \dots, n$.

The function κ maps a Herbrand interpretation $\omega \in \Omega(\Sigma, D)$ onto a reference world $\hat{\omega} \in \hat{\Omega}$ with the intended meaning that $\kappa(\omega)$ is the (unique) reference world that represents ω. It holds that $\kappa(\omega) = \hat{\omega}$ whenever $\hat{\omega}$ assigns the same number of elements of an $\mathcal{R}$-equivalence class $\underline{S}_i$ to some truth configuration t as ω contains specific instances of this truth configuration for elements in $\underline{S}_i$.

Example 7.6. We continue Examples 7.1 and 7.5 and consider the Herbrand interpretation

$$\omega_2 = \{\mathit{flies}(\mathsf{huey})\} \quad .$$

For the $\mathcal{R}$-equivalence class $\underline{S}_1 = \{\mathsf{tweety}\}$ and the single element tweety $\in \underline{S}_1$ it holds that ω_2 satisfies $\omega_2 \models^{\mathrm{F}} t_2^{\wedge}(\mathsf{tweety})$ with $t_2 = \overline{\mathit{flies}}$. Therefore, it has to hold that $\kappa(\omega_2)(t_2)(\underline{S}_1) = 1$. For the $\mathcal{R}$-equivalence class $\underline{S}_2 = \{\mathsf{huey}, \mathsf{dewey}, \mathsf{louie}\}$ there is the one element huey $\in \underline{S}_2$ for which it holds that ω_2 satisfies $\omega_2 \models^{\mathrm{F}} t_1^{\wedge}(\mathsf{huey})$ with $t_1 = \mathit{flies}$. Furthermore, there are two elements dewey, louie $\in \underline{S}_2$ with $\omega_2 \models^{\mathrm{F}} t_2^{\wedge}(\mathsf{dewey})$ and $\omega_2 \models^{\mathrm{F}} t_2^{\wedge}(\mathsf{louie})$, respectively. Therefore, it has to hold that both $\kappa(\omega_2)(t_2)(\underline{S}_2) = 2$ and $\kappa(\omega_2)(t_1)(\underline{S}_2) = 1$. It follows that $\kappa(\omega_2) = \hat{\omega}_2$.

Proposition 7.5. *The function κ is surjective.*

Proof. Let $\hat{\omega} \in \hat{\Omega}$ be a reference world, let $\Theta = \{t_1, \dots, t_T\}$, and let $\underline{S}_i = \{\mathsf{c}_1^i, \dots, \mathsf{c}_{n_i}^i\}$ for every $\underline{S}_i \in \mathfrak{S}|_D(\mathcal{R})$. Let $\omega \in \Omega(\Sigma, D)$ be the Herbrand interpretation defined via

$$\omega =_{def} \bigcup_{i=1}^{T} \bigcup_{j=1}^{n} \{p \in t^+(\mathsf{c}_k^j) \mid \sum_{l=1}^{i-1} \hat{\omega}(t_l)(\underline{S}_j) < k \leq \sum_{l=1}^{i} \hat{\omega}(t_l)(\underline{S}_j)\} \qquad (7.2)$$

It is clear that ω is indeed a Herbrand interpretation, i. e., it holds that $\omega \subseteq \mathsf{At}((D, \mathit{Pred}, \emptyset))$, as ω is constructed from elements from $t^+(\mathsf{c}_1), \dots, t^+(\mathsf{c}_m)$ for constants $\mathsf{c}_1, \dots, \mathsf{c}_m$ and each $t^+(\mathsf{c}_1)$ is a subset of $\mathsf{At}(\Sigma)$. Furthermore it holds that $\kappa(\omega) = \hat{\omega}$ as ω contains at least

$$\sum_{l=1}^{i} \hat{\omega}(t_l)(\underline{S}_j) - \sum_{l=1}^{i-1} \hat{\omega}(t_l)(\underline{S}_j) = \hat{\omega}(t_i)(\underline{S}_j)$$

instances satisfying t_i for $\underline{S}_i$ and no constant $\mathsf{c} \in \underline{S}_i$ is used to satisfy two different truth configurations by the construction in Equation (7.2). Notice also, that each c_k^j in Equation (7.2) is well-defined as $k > 0$ and

$$k \leq \sum_{l=1}^{n} \hat{\omega}(t_l)(\underline{S}_j) = |\underline{S}_j| \quad . \qquad \square$$

Now, let the *span number* $\rho_{\hat{\omega}}$ of a reference world $\hat{\omega} \in \hat{\Omega}$ be defined as

$$\rho_{\hat{\omega}} =_{def} \prod_{i=1}^{n} \binom{|\underline{S}_i|}{\hat{\omega}(t_1)(\underline{S}_i), \ldots, \hat{\omega}(t_T)(\underline{S}_i)}$$

with $\Theta = \{t_1, \ldots, t_T\}$ and

$$\binom{n}{k_1, \ldots, k_r} =_{def} \frac{n!}{k_1! \cdots k_r!}$$

being the *multinomial coefficient* indexed by n and $k_1, \ldots, k_r$ with $n = k_1 + \ldots + k_r$. We define

$$\binom{n}{k_1, \ldots, k_r} =_{def} 0$$

if any $k_i < 0$ for $i = 1, \ldots, n$. Note that $\rho_{\hat{\omega}}$ is well-defined as $\hat{\omega}(t_1)(\underline{S}_i) + \ldots + \hat{\omega}(t_T)(\underline{S}_i) = |\underline{S}_i|$ for every reference world $\hat{\omega}$. The span number of a reference world $\hat{\omega}$ is exactly the number of Herbrand interpretations that are subsumed by $\hat{\omega}$.

Proposition 7.6. *It holds that* $|\kappa^{-1}(\hat{\omega})| = \rho_{\hat{\omega}}$ *for every* $\hat{\omega} \in \hat{\Omega}$.

Proof. Let $\hat{\omega} \in \hat{\Omega}$ be a reference world, let $\Theta = \{t_1, \ldots, t_T\}$, and let $\underline{S}_i$ be some $\mathcal{R}$-equivalence class. The value $\hat{\omega}(t, \underline{S}_i) = H$ represents the fact that in $\hat{\omega}$ there are H different constants from the $\mathcal{R}$-equivalence class $\underline{S}_i$ that satisfy $t \in \Theta$. Note that every two truth configurations $t, t' \in \Theta$ describe mutual exclusive scenarios for some constant c and the set Θ is exhaustive, i. e., for every constant c there applies exactly one truth configuration in a Herbrand interpretation ω. It follows that the set of constants $\underline{S}_i$ has to be distributed among the truth configurations. This is equivalent to the combinatorial interpretation of multinomial coefficients. The term

$$\binom{|\underline{S}_i|}{\hat{\omega}(t_1)(\underline{S}_i), \ldots, \hat{\omega}(t_T)(\underline{S}_i)}$$

denotes the number of combinations of partitioning the set $\underline{S}_i$ into T subsets of sizes $\hat{\omega}(t_1)(\underline{S}_i), \ldots, \hat{\omega}(t_T)(\underline{S}_i)$, respectively. Each such combination can be combined with a similar selection for the other $\mathcal{R}$-equivalence classes.

To each such *joint* combination an actual Herbrand interpretation can be constructed that is mapped to $\hat{\omega}$ by κ. Hence it follows that

$$|\kappa^{-1}(\hat{\omega})| = \prod_{i=1}^{n} \binom{|\underline{S}_i|}{\hat{\omega}(t_1)(\underline{S}_i), \ldots, \hat{\omega}(t_T)(\underline{S}_i)} = \rho_{\hat{\omega}} \quad . \qquad \square$$

The following proposition states that $\hat{\Omega}$ indeed characterizes $\Omega^D/_{\equiv_\mathcal{R}}$.

Proposition 7.7. *For every $\omega_1, \omega_2 \in \Omega(\Sigma, D)$ it holds that $\omega_1 \equiv_\mathcal{R} \omega_2$ if and only if $\kappa(\omega_1) = \kappa(\omega_2)$.*

Proof. We have to show both directions. First, let it hold that $\omega_1 \equiv_\mathcal{R} \omega_2$. Then there is a set $T = \{(\mathsf{c}_1^1, \mathsf{c}_2^1), \ldots, (\mathsf{c}_1^G, \mathsf{c}_2^G)\} \subseteq \underline{S}_1 \times \underline{S}_1 \cup \ldots \cup \underline{S}_n \times \underline{S}_n$ with

$$\omega_1 = \omega_2[\mathsf{c}_1^1 \leftrightarrow \mathsf{c}_2^1] \ldots [\mathsf{c}_1^G \leftrightarrow \mathsf{c}_2^G] \quad .$$

As $\mathsf{c}_1^1, \mathsf{c}_2^1 \in \underline{S}_i$ for some i it follows that $\kappa(\omega_2)(t)(\underline{S}_i) = \kappa(\omega_2[\mathsf{c}_1^1 \leftrightarrow \mathsf{c}_2^1])(t)(\underline{S}_i)$ for every $t \in \Theta$, i.e., the number of constants of $\underline{S}_i$ that satisfy t keeps the same when exchanging c_1^1 with c_2^1. Furthermore, for $k \neq i$ it holds that $\kappa(\omega_2)(t)(\underline{S}_k) = \kappa(\omega_2[\mathsf{c}_1^1 \leftrightarrow \mathsf{c}_2^1])(t)(\underline{S}_k)$ for every $j = 1, \ldots, P$ as well as no constant of $\underline{S}_k$ that satisfies t is affected by the replacement. This condition is maintained for multiple exchanges of constants from the same $\mathcal{R}$-equivalence class. It follows that $\kappa(\omega_2)(t)(\underline{S}_i) = \kappa(\omega_1)(t)(\underline{S}_i)$ for every $i = 1, \ldots, n$ and every $t \in \Theta$, i.e. $\kappa(\omega_1) = \kappa(\omega_2)$. For the other direction, let $\kappa(\omega_1) = \kappa(\omega_2) = \hat{\omega}$. For every $t \in \Theta$ let $L(t, \omega_1)$ denote the set of constants that satisfy t in ω_1 and let $L(t, \omega_2)$ denote the set of constants that satisfy t in ω_2. As $\kappa(\omega_1) = \kappa(\omega_2)$ it follows

$$|L(t, \omega_1) \cap \underline{S}_i| = \hat{\omega}(t)(\underline{S}_i) = |L(t, \omega_2) \cap \underline{S}_i|$$

for every $i = 1, \ldots, n$. Furthermore, it holds that $L(t, \omega) \cap L(t', \omega) = \emptyset$ for different $t, t' \in \Theta$ and $\omega \in \{\omega_1, \omega_2\}$ as truth configurations are mutual exclusive. As truth configurations are also exhaustive we can define bijections $\sigma_{t,i} : L(t, \omega_1) \cap \underline{S}_i \to L(t, \omega_2) \cap \underline{S}_i$ for every $t \in \Theta$ and $i = 1, \ldots, n$. In other words, applying $\sigma_{t,i}$ as a replacement onto ω_1 amounts to

$$\sigma_{t,i} \left(\bigcup_{\mathsf{c} \in \underline{S}_i} \{p \in t^+(\mathsf{c})\} \cap \omega_1 \right) = \bigcup_{\mathsf{c} \in \underline{S}_i} \{p \in t^+(\mathsf{c})\} \cap \omega_2 \quad .$$

As

$$\bigcup_{t \in \Theta} \mathsf{Dom}\, \sigma_{t,i} = \underline{S}_i \quad \text{and} \quad \bigcup_{t \in \Theta} \mathsf{Im}\, \sigma_{t,i} = \underline{S}_i$$

the function $\sigma_i : \underline{S}_i \to \underline{S}_i$ with $\sigma_i(\mathsf{c}) = \sigma_{t,i}(\mathsf{c})$ for $\mathsf{c} \in \mathsf{Dom}\, \sigma_{t,i}$ is a permutation on $\underline{S}_i$. As $\underline{S}_i \cap \underline{S}_j = \emptyset$ for $i \neq j$ and $\underline{S}_1 \cup \ldots \cup \underline{S}_n = D$ the function

$\sigma : D \to D$ with $\sigma(\mathsf{c}) = \sigma_i(\mathsf{c})$ for $\mathsf{c} \in \mathsf{Dom}\ \sigma_i$ is a permutation on D. By construction it holds that $\omega_2 = \sigma(\omega_1)$. By the fact that every permutation can be represented as a product of transpositions (Beachy and Blair, 2005) it follows that $\omega_1 \equiv_{\mathcal{R}} \omega_2$ □

Rephrasing the above proposition we obtain the following concise statement.

Corollary 7.1. *The function* $\iota : \Omega(\Sigma, D)/_{\equiv_{\mathcal{R}}} \to \hat{\Omega}$ *with* $\iota([\omega]) = \kappa(\omega)$ *is a bijection.*

Example 7.7. We continue Example 7.5. There it holds that

$$\begin{array}{lll} [\omega_0] = \{\omega_0\} & [\omega_1] = \{\omega_1\} & [\omega_2] = \{\omega_2, \omega_3, \omega_4\} \\ [\omega_5] = \{\omega_5, \omega_6, \omega_7\} & [\omega_8] = \{\omega_8, \omega_9, \omega_{10}\} & [\omega_{11}] = \{\omega_{11}, \omega_{12}\omega_{13}\} \\ [\omega_{14}] = \{\omega_{14}\} & [\omega_{15}] = \{\omega_{15}\} & \end{array}$$

and

$$\begin{array}{llll} \iota([\omega_0] = \hat{\omega}_0 & \iota([\omega_1] = \hat{\omega}_5 & \iota([\omega_2] = \hat{\omega}_2 & \iota([\omega_5] = \hat{\omega}_6 \\ \iota([\omega_8] = \hat{\omega}_3 & \iota([\omega_{11}] = \hat{\omega}_7 & \iota([\omega_{14}] = \hat{\omega}_4 & \iota([\omega_{15}] = \hat{\omega}_8 \end{array} .$$

After having established the equivalence of $\Omega(\Sigma, D)/_{\equiv_{\mathcal{R}}}$ and $\hat{\Omega}$ we now turn to the issue of representing P^* on the basis of $\hat{\Omega}$. For that we need the following notation.

Definition 7.8 (Prototypical uniformity)**.** A probability function $P : \Omega(\Sigma, D) \to [0, 1]$ is called *prototypically uniform* with respect to D if for $\omega_1, \omega_2 \in \Omega(\Sigma, D)$ with $\omega_1 \equiv_{\mathcal{R}} \omega_2$ it follows that $P(\omega_1) = P(\omega_2)$.

Corollary 7.2. P^* *is prototypically uniform with respect to D.*

Proof. This follows directly from Proposition 7.4. □

Prototypically uniform probability functions can be be concisely represented using *condensed probability functions*.

Definition 7.9 (Condensed probability function)**.** Let P be a probability function $P : \Omega(\Sigma, D) \to [0, 1]$ that is prototypically uniform with respect to D. Then the *condensed probability function* $\hat{P}$ for P is the probability function $\hat{P} : \hat{\Omega} \to [0, 1]$ defined via

$$\hat{P}(\hat{\omega}) =_{def} P(\omega) \quad \text{for some } \omega \text{ with } \kappa(\omega) = \hat{\omega} \text{ and for all } \hat{\omega} \in \hat{\Omega} \quad .$$

Let $\hat{\mathcal{P}}_D$ denote the set of all condensed probability functions.

As $P(\omega_1) = P(\omega_2)$ follows from $\kappa(\omega_1) = \kappa(\omega_2)$ for prototypically uniform P, the probability function $\hat{P}$ is well-defined. It also holds that the mapping between prototypically uniform probability functions and condensed probability functions is bijective.

Proposition 7.8. *Let P_1, P_2 be prototypically uniform probability functions with respect to D. If $P_1 \neq P_2$ then $\hat{P}_1 \neq \hat{P}_2$.*

Proof. If $P_1 \neq P_2$ then there is a Herbrand interpretation $\omega \in \Omega(\Sigma, D)$ with $P_1(\omega) \neq P_2(\omega)$. It follows directly that $\hat{P}_1(\kappa(\omega)) = P_1(\omega) \neq P_2(\omega) = \hat{P}_2(\kappa(\omega))$ and therefore $\hat{P}_1 \neq \hat{P}_2$. $\square$

For the prototypically uniform probability function P^*, its condensed probability function $\hat{P}^*$, and a ground sentence ψ it follows directly by definition that

$$\hat{P}^*(\psi) =_{def} P^*(\psi) = \sum_{\omega \in \Omega(\Sigma, D),\ \omega \models^{\mathrm{F}} \psi} \hat{P}^*(\kappa(\omega)) \quad . \tag{7.3}$$

As one can see, one can determine the probability of any ground sentence using $\hat{P}^*$ instead of P^*. However, the sum in the above equation still considers every Herbrand interpretation in $\Omega(\Sigma, D)$ and there may be a $\hat{\omega} \in \hat{\Omega}$ such that $\hat{P}^*(\hat{\omega})$ appears more than once as $\omega, \omega' \in \Omega(\Sigma, D)$ may yield $\kappa(\omega) = \kappa(\omega') = \hat{\omega}$. In the next section we consider the question of how to determine the probability of ψ *without* considering $\Omega(\Sigma, D)$ but only $\hat{\Omega}$ instead, i. e. we want to consider each $\hat{\omega} \in \hat{\Omega}$ only once. Afterwards, we address the issue of determining $\hat{P}^*$ directly from a knowledge base $\mathcal{R}$ without determining P^* first.

7.1.2 *Lifted Inference and Maximum Entropy*

Looking closer at Equation (7.3) one can see that the probability of a single reference world $\hat{\omega}$ may occur more than once within the sum as for different $\omega, \omega' \in \Omega(\Sigma, D)$ with $\omega \models^{\mathrm{F}} \psi$ and $\omega' \models^{\mathrm{F}} \psi$ it may hold that $\kappa(\omega) = \kappa(\omega')$. Therefore, (7.3) can be rewritten to

$$\hat{P}(\psi) = \sum_{\hat{\omega} \in \hat{\Omega}} \Lambda(\hat{\omega}, \psi) \hat{P}(\hat{\omega}) \quad . \tag{7.4}$$

for some $\Lambda(\hat{\omega}, \psi) \in \mathbb{N}_0$ that represents the number of occurrences of $\hat{P}(\hat{\omega})$ in (7.3). More specifically, it holds that

$$\Lambda(\hat{\omega}, \psi) = |\{\omega \in \Omega(\Sigma, D) \mid \kappa(\omega) = \hat{\omega} \wedge \omega \models^{\mathrm{F}} \psi\}|,$$

i. e., $\Lambda(\hat{\omega}, \psi)$ is the number of Herbrand interpretations in (7.3) that satisfy ψ and are mapped by κ to $\hat{\omega}$.

Example 7.8. We continue Example 7.7. There we have that

$$\begin{aligned} P^*(\mathit{flies}(\mathsf{tweety})) &= 1 \cdot \hat{P^*}(\hat{\omega}_5) + 3 \cdot \hat{P^*}(\hat{\omega}_6) + 3 \cdot \hat{P^*}(\hat{\omega}_7) + 1 \cdot \hat{P^*}(\hat{\omega}_8) \\ P^*(\mathit{flies}(\mathsf{huey})) &= 1 \cdot \hat{P^*}(\hat{\omega}_2) + 2 \cdot \hat{P^*}(\hat{\omega}_3) + 1 \cdot \hat{P^*}(\hat{\omega}_4) + 1 \cdot \hat{P^*}(\hat{\omega}_6) + \\ &\qquad 2 \cdot \hat{P^*}(\hat{\omega}_7) + 1 \cdot \hat{P^*}(\hat{\omega}_8) \\ &= \hat{P^*}(\mathit{flies}(\mathsf{dewey})) = \hat{P^*}(\mathit{flies}(\mathsf{louie})) \quad . \end{aligned}$$

In particular, we have $\Lambda(\hat{\omega}_6, \mathit{flies}(\mathsf{tweety})) = 3$ as $\iota^{-1}(\hat{\omega}_6) = \{\omega_5, \omega_6, \omega_7\}$ with

$$\begin{aligned} \omega_5 &= \{\mathit{flies}(\mathsf{tweety}), \mathit{flies}(\mathsf{huey})\} \\ \omega_6 &= \{\mathit{flies}(\mathsf{tweety}), \mathit{flies}(\mathsf{dewey})\} \\ \omega_7 &= \{\mathit{flies}(\mathsf{tweety}), \mathit{flies}(\mathsf{louie})\} \end{aligned}$$

and it holds that $\omega_i \models^{\mathrm{F}} \mathit{flies}(\mathsf{tweety})$ for $i = 5, 6, 7$.

Note, however, that determining $\Lambda(\hat{\omega}, \psi)$ by its definition above still requires considering all Herbrand interpretations in $\Omega(\Sigma, D)$. By exploiting combinatorial patterns within the structure of $\Omega(\Sigma, D)$ we can avoid considering $\Omega(\Sigma, D)$ as a whole and characterize $\Lambda(\hat{\omega}, \psi)$ as follows.

Proposition 7.9. *Let ψ be a conjunction of ground literals, let $\mathsf{Const}(\psi) = \{\mathsf{c}_1, \ldots, \mathsf{c}_m\}$, and let $\Theta = \{t_1, \ldots, t_T\}$. Then*

$$\Lambda(\hat{\omega}, \psi) = \sum_{(t'_1, \ldots, t'_m) \in \Theta(\psi, \mathsf{c}_1, \ldots, \mathsf{c}_m)} \prod_{i=1}^{n} \binom{|\underline{S}_i \setminus \mathsf{Const}(\phi)|}{\alpha_i^{t_1}(t'_1, \ldots, t'_m), \ldots, \alpha_i^{t_T}(t'_1, \ldots, t'_m)}$$

with $\alpha_i^t(t'_1, \ldots, t'_m) =_{def} \hat{\omega}(t)(\underline{S}_i) - |\{k \mid t'_k = t \wedge \mathsf{c}_k \in \underline{S}_i\}|$.

The proof of Proposition 7.9 can be found in Appendix A on page 249. Note that there is no more reference to $\Omega(\Sigma, D)$ in the above characterization of $\Lambda(\hat{\omega}, \psi)$.

In order to determine $\hat{P^*}(\psi)$ for an arbitrary ground sentence ψ remember that ψ can be rewritten to be in disjunctive normal form, see e.g. (Russell and Norvig, 2009). Assume ψ to be in disjunctive normal form and let $c(\psi)$ denote the set of conjuncts of ψ. By iterative application of 2.) of Proposition 2.1 (see page 20) it follows for every probability function P that

$$P(\psi) = \sum_{\psi' \in c(\psi)} P(\psi') - \sum_{(\psi', \psi'') \in c(\psi)^2,\ \psi' \neq \psi''} P(\psi' \wedge \psi'') \quad .$$

As for $\hat{P^*}$, for every $\psi', \psi'' \in c(\psi)$ the terms $\hat{P^*}(\psi')$ and $\hat{P^*}(\psi' \wedge \psi'')$ are well-defined by Equation (7.4) and Proposition 7.9. If we require ψ to be in a disjunctive form such that every two disjuncts are mutually exclusive we get the even simpler form

$$P(\psi) = \sum_{\psi' \in c(\psi)} P(\psi')$$

which derives directly from above as $P(\psi' \wedge \psi'') = 0$ for every $\psi', \psi'' \in c(\psi)^2$ with $\psi' \neq \psi''$.

Hitherto we have shown that $\hat{P^*}$ compactly represents P^* and that $\hat{P^*}$ can be used for reasoning just as P^*. Nonetheless, in order to determine $\hat{P^*}$ one needs to compute P^* first using e. g. the optimization problems stated in Equation (6.15) or Equation (6.42), see pages 170 and 173, respectively. These optimization problems still need exponential space as P^* needs exponential space. In the following, we show that we can modify both (6.15) and (6.42) in a straightforward fashion to determine $\hat{P^*}$ directly without exponential overhead. Although the approach of condensed probability distributions is applicable to any inductive inference mechanism that obeys (Prototypical Indifference) we restrain our attention to $\mathcal{I}_\varnothing$ and $\mathcal{I}_\odot$ as developed in the previous chapter. As a consequence, from now one we assume that P^* has been determined either by (6.15) or (6.42).

For a condensed probability function $\hat{P}$ we define the *entropy* $H_D(\hat{P})$ of $\hat{P}$ to be the entropy of P, i. e. $H_D(\hat{P}) =_{def} H_D(P)$, which is equivalent to

$$\begin{aligned}
H_D(\hat{P}) &= H_D(P) \\
&= - \sum_{\omega \in \Omega(\Sigma),\ \mathsf{Const}(\omega) \subseteq D} P(\omega) \operatorname{ld} P(\omega) \\
&= - \sum_{\omega \in \Omega(\Sigma, D)} P(\omega) \operatorname{ld} P(\omega) \\
&= - \sum_{\omega \in \Omega(\Sigma, D)} \hat{P}(\kappa(\omega)) \operatorname{ld} \hat{P}(\kappa(\omega)) \\
&= - \sum_{\hat{\omega} \in \hat{\Omega}} \sum_{\omega \in \kappa^{-1}(\hat{\omega})} \hat{P}(\kappa(\omega)) \operatorname{ld} \hat{P}(\kappa(\omega)) \\
&= - \sum_{\hat{\omega} \in \hat{\Omega}} \rho_{\hat{\omega}} \hat{P}(\hat{\omega}) \operatorname{ld} \hat{P}(\hat{\omega})
\end{aligned}$$

and thus can be determined by just considering $\hat{\Omega}$.

Proposition 7.10. *Let $\mathcal{S}$ be a set of prototypically uniform probability functions with respect to D and*

$$\hat{\mathcal{S}} =_{def} \{\hat{P} \mid P \in \mathcal{S}\} \quad .$$

If the probability function $P^* = \arg\max_{P \in \mathcal{S}} H_D(P)$ is uniquely determined so is $\hat{Q}^* = \arg\max_{\hat{P} \in \hat{\mathcal{S}}} H_D(\hat{P})$, and it holds that $\hat{Q}^* = \hat{P}^*$.

Proof. Let $P^* = \arg\max_{P \in \mathcal{S}} H_D(P)$ be uniquely determined and assume that there are $\hat{Q}_1^*$ and $\hat{Q}_2^*$ with $\hat{Q}_1^* \neq \hat{Q}_2^*$ and $H_D(\hat{Q}_1^*) = H_D(\hat{Q}_2^*) =$

$\max_{\hat{P}\in\hat{\mathcal{S}}} H_D(\hat{P})$. As $H_D(P) = H_D(\hat{P})$ for every $P \in \mathcal{S}$ it follows that $H_D(Q_1^*) = H_D(Q_2^*)$ as well and also $\max_{P\in\mathcal{S}} H_D(P) = \max_{\hat{P}\in\hat{\mathcal{S}}} H_D(\hat{P})$. By Proposition 7.8 on page 205 it also holds that $Q_1^* \neq Q_2^*$ and it follows that $H_D(Q_1^*) = H_D(Q_2^*) = H_D(P^*) = \max_{P\in\mathcal{S}} H_D(P)$ and therefore the function P^* would not be uniquely determined, contradicting the assumptions. It follows that $Q_1^* = Q_2^* = P^*$ and therefore $\hat{Q}_1^* = \hat{Q}_2^* = \hat{P}^*$ proving the second part of the claim. □

Proposition 7.11. *Let* $\models_\circ \in \{\models_\varnothing^{pr}, \models_\odot^{pr}\}$. *For a knowledge base* $\mathcal{R}$ *and a finite* D *with* $\mathsf{Const}(\mathcal{R}) \subseteq D \subseteq U$ *let*

$$\mathcal{S} =_{def} \{P \mid P, D \models_\circ \mathcal{R}\} \quad ,$$

and let $\mathcal{S}' \subseteq \mathcal{S}$ *be its subset of prototypically uniform probability functions with respect to* D. *If* $\arg\max_{P\in\mathcal{S}} H_D(P)$ *is uniquely determined then it holds that*

$$\arg\max_{P\in\mathcal{S}'} H_D(P) = \arg\max_{P\in\mathcal{S}} H_D(P) \quad .$$

Proof. Corollary 7.2 on page 204 already stated that $\mathcal{I}_\circ(\mathcal{R}, D)$ is prototypically uniform with respect to D if $\mathcal{I}_\circ$ satisfies (Prototypical Indifference). When determining the solution to $\arg\max_{P\in\mathcal{S}} H_D(P)$ we can therefore restrain attention to the set of prototypically uniform probability functions $\mathcal{S}'$. □

The implications of the above two propositions are as follows. Instead of determining first $P^* = \mathcal{I}(\mathcal{R}, D)|_D$ via (6.15) or (6.42) and then determining $\hat{P}^*$ we can directly determine $\hat{P}^*$ by rewriting (6.15) or (6.42). For (6.15) this is

$$\widehat{\mathcal{I}_\varnothing(\mathcal{R}, D)}|_D = \begin{cases} \underset{\hat{P}\in\hat{\mathcal{P}}_D \text{ and } \hat{P},D\models_\varnothing^{pr}\mathcal{R}}{\arg\max} H_D(P) & \text{if } \mathcal{R}\ \varnothing\text{-consistent wrt. } D \\ & \quad \text{and } \mathsf{Const}(\mathcal{R}) \subseteq D \\ & \quad \text{and } D \text{ is finite} \\ & \quad \text{and solution is unique} \\ \mathsf{undef} & \text{otherwise} \end{cases} \tag{7.5}$$

and similarly for (6.42) this is

$$\widehat{\mathcal{I}_\odot(\mathcal{R}, D)}|_D = \begin{cases} \underset{\hat{P}\in\hat{\mathcal{P}}_D \text{ and } \hat{P},D\models_\odot^{pr}\mathcal{R}}{\arg\max} H_D(P) & \text{if } \mathcal{R}\ \odot\text{-consistent wrt. } D \\ & \quad \text{and } \mathsf{Const}(\mathcal{R}) \subseteq D \\ & \quad \text{and } D \text{ is finite} \\ \mathsf{undef} & \text{otherwise} \end{cases} \tag{7.6}$$

Note that both $\hat{P}, D \models_\varnothing^{pr} \mathcal{R}$ and $\hat{P}, D \models_\odot^{pr} \mathcal{R}$ can be checked directly for $\hat{P}$ by employing Equation (7.4) on page 205 and Proposition 7.9 on page 206.

7.2 ANALYSIS

In the following we analyze the computational benefits of using $\hat{P}^*$ instead of P^*. In particular, we are interested in the question how the cardinality of $\hat{\Omega}$ compares to the cardinality of $\Omega(\Sigma, D)$ with respect to the number of constants $|D|$ considered. Proposition 7.2 on page 197 already established that $|\Omega(\Sigma, D)| = 2^{|D||Pred|}$ and therefore the space needed to represent P^* is exponential in both $|D|$ and $|Pred|$. We do not expect to avoid an exponential blow-up in the number of predicates in *Pred* but we show that $|\hat{\Omega}|$ is not exponential in $|D|$ any more.

Remember that each $\hat{\omega} \in \hat{\Omega}$ satisfies

$$\sum_{t \in \Theta} \hat{\omega}(t)(\underline{S}_i) = |\underline{S}_i| \quad (\text{for all } i = 1, \ldots, n),$$

see Definition 7.6 on page 199. This means, that for each $\hat{\omega}$ the constants of each $\underline{S}_i$ are distributed among the truth configurations in Θ. Note that the number of truth configurations is $2^{|Pred|}$, i. e. $|\Theta| = 2^{|Pred|}$. A distribution of constants of $\underline{S}_i$ among Θ can be combined with any distribution of constants of $\underline{S}_j$ for every $i \neq j$, yielding a single reference world $\hat{\omega}$. In order to count the number of reference worlds we need to multiply the number of combinations one can distribute the constants of $\underline{S}_i$ onto the truth configurations in Θ with the number of combinations for every other $\underline{S}_j$ ($i \neq j$). Remember that $\mathfrak{S}|_D(\mathcal{R}) = \{\underline{S}_1, \ldots, \underline{S}_n\}$. Then the previous considerations amount to

$$|\hat{\Omega}| = \prod_{i=1}^{n} |\{(l_1, \ldots, l_{2^{|Pred|}}) \in \mathbb{N}_0^{2^{|Pred|}} \mid l_1 + \ldots + l_{2^{|Pred|}} = |\underline{S}_i|\}| \quad . \tag{7.7}$$

Each factor in the product of the above equation represents the number of combinations the constants of a single $\mathcal{R}$-equivalence class can be distributed among the possible truth configurations in Θ. The condition $l_1 + \ldots + l_{2^{|Pred|}} = |\underline{S}_i|$ ensures that each constant is exactly assigned one truth configuration in every combination. By multiplying the number of the combinations for each $\mathcal{R}$-equivalence class we get the number of combinations the $\mathcal{R}$-equivalence classes can be distributed among the truth configurations in Θ which is exactly the number of reference worlds.

Equation (7.7) gives no direct hint on the space needed to represent $\hat{\Omega}$ in terms of $|D|$ and $|Pred|$. But it is possible to rewrite (7.7) as follows.

Definition 7.10 (Cardinality generator)**.** The *cardinality generator* g_c is the function $g_c : \mathbb{N}_0^2 \to \mathbb{N}_0$ defined via

$$g_c(n_1, n_2) =_{def} \begin{cases} \sum_{i=0}^{n_2} g_c(n_1 - 1, n_2 - i) & \text{if } n_2 > 0 \text{ and } n_1 > 0 \\ 1 & \text{if } n_2 = 0 \\ 0 & \text{otherwise} \end{cases} .$$

Basically, the cardinality generator g_c is a recursive function that maps two integers n_1 and n_2 to the sum $g_c(n_1 - 1, 0) + \ldots + g_c(n_1 - 1, n_2)$. The intuition behind using g_c to enumerate the number of reference worlds is as follows. The first argument of g_c is meant to represent the number of truth configurations and the second the number of constants in a $\mathcal{R}$-equivalence class. By defining $g_c(n_1, n_2) = g_c(n_1 - 1, 0) + \ldots + g_c(n_1 - 1, n_2)$ we say that the number of combinations to distribute n_2 constants on n_1 truth configuration is equal to the number of combinations to distribute zero constants on $n_1 - 1$ truth configurations plus the number of combinations to distribute one constant on $n_1 - 1$ truth configurations, and so on. The first case describes a setting where we assign all n_2 constants to the n_1th truth configuration and as there are no remaining constants left this amounts to the number of $g_c(n_1 - 1, 0)$ remaining combinations. The second case describes a setting where we assign $n_2 - 1$ constants to the n_1th truth configuration and the remaining single constant to the remaining $n_1 - 1$ truth configurations. The final case describes the setting of assigning no constant the n_1th truth configuration and the remaining n_2 constants to the remaining $n_1 - 1$ truth configurations. The second case of the definition of g_c is the basic case for the setting that we have to distribute zero constants on n_1 truth configurations. Obviously, there is only one possible assignment, namely, assigning to each truth configuration zero constants. The third case of the definition of g_c ensures that there are no combinations if the number of truth configurations is zero. Consider $g_c(1, 3)$ as the number of combinations to distribute three constants on one truth configuration. Applying the first case of the definition of g_c yields

$$g_c(1, 3) = g_c(0, 0) + g_c(0, 1) + g_c(0, 2) + g_c(0, 3)$$

and therefore the number of combinations to distribute 3 constants on 1 truth configuration is to assign all three constants to the one truth configuration, or to assign zero, one, or two to it. Obviously, the latter cases are not valid and the only valid assignment is that three constants are assigned to the one truth configuration. Due to the third case in the definition of g_c the terms $g_c(0, 1)$, $g_c(0, 2)$, and $g_c(0, 3)$ are set to zero.

Proposition 7.12. *It holds that*

$$|\hat{\Omega}| = \prod_{i=1}^{n} g_c(2^{|Pred|}, |\underline{S}_i|) \quad . \tag{7.8}$$

Proof. We already established that $g_c(2^{|Pred|}, |\underline{S}_i|)$ is the number of combinations of distributing $|\underline{S}_i|$ constants to $2^{|Pred|}$ truth configurations. It follows that

$$g_c(2^{|Pred|}, |\underline{S}_i|) = |\{(l_1, \ldots, l_{2^{|Pred|}}) \in \mathbb{N}_0^{2^{|Pred|}} \mid l_1 + \ldots + l_{2^{|Pred|}} = |\underline{S}_i|\}| \quad .$$

By applying the above to (7.7) we obtain the claim. □

Still, Equation (7.8) does not allow to get an idea of the size of $|\hat{\Omega}|$. However, the function g_c can be bounded from above as follows.

Lemma 7.1. *For $n_1, n_2, n_2' \in \mathbb{N}_0$ with $n_1, n_2, n_2' > 0$ it holds that $n_2 < n_2'$ implies $g_c(n_1, n_2) \leq g_c(n_1, n_2')$.*

Proof. For $n_2 < n_2'$ it holds that

$$\begin{aligned} g_c(n_1, n_2) &= \sum_{i=0}^{n_2} g_c(n_1 - 1, n_2 - i) \\ &\leq \sum_{i=0}^{n_2} g_c(n_1 - 1, n_2 - i) + \sum_{i=n_2+1}^{n_2'} g_c(n_1 - 1, n_2' - i) \\ &= \sum_{i=0}^{n_2'} g_c(n_1 - 1, n_2 - i) = g_c(n_1, n_2') \end{aligned}$$ □

Lemma 7.2. *It holds that $g_c(n_1, n_2) \leq (n_2 + 1)^{n_1}$ for every $n_1, n_2 \in \mathbb{N}_0$.*

Proof. We prove the above statement by induction on the structure of g_c. First, consider $n_2 = 0$ and $n_1 \in \mathbb{N}_0$. Then

$$g_c(n_1, n_2) = 1 = 1^{n_1} = (n_2 + 1)^{n_1} \quad .$$

Now consider $n_1 = 0$ and $n_2 \in \mathbb{N}_0$ with $n_2 > 0$. Then

$$g_c(n_1, n_2) = 0 \leq 1 = (n_2 + 1)^{n_1} \quad .$$

Now assume $g_c(n_1', n_2') \leq (n_2' + 1)^{n_1'}$ for every $n_1' < n_1$ and $n_2' \leq n_2$. It remains to show $g_c(n_1, n_2) \leq (n_2 + 1)^{n_1}$ for $n_1 > 0$ and $n_2 > 0$. Remember that for this case g_c is monotonously increasing in the second argument, cf. Lemma 7.1.

$$\begin{aligned} g_c(n_1, n_2) &= \sum_{i=0}^{n_2} g_c(n_1 - 1, n_2 - i) \leq (n_2 + 1) g_c(n_1 - 1, n_2) \\ &\leq (n_2 + 1)(n_2 + 1)^{n_1 - 1} = (n_2 + 1)^{n_1} \end{aligned}$$ □

As a direct consequence from the above lemma we can give an upper bound on $|\hat{\Omega}|$ as follows.

Theorem 7.1. *Let $\mathcal{R}$ be a knowledge base. Then*

$$|\hat{\Omega}| \leq (|\mathsf{Const}(\mathcal{R})| + 1)(|D| + 1)^{2^{|Pred|}} \quad .$$

Proof. By Proposition 7.12 it holds that

$$|\hat{\Omega}| = \prod_{i=1}^{n} g_c(2^{|Pred|}, |\underline{S}_i|) \quad .$$

As $\mathcal{R}$ mentions exactly $|\mathsf{Const}(\mathcal{R})|$ different constants it follows

$$|\mathfrak{S}(\mathcal{R})| \geq |\mathsf{Const}(\mathcal{R})| + 1$$

as all constants not appearing in $\mathcal{R}$ belong to the same $\mathcal{R}$-equivalent class. Furthermore, each $\underline{S} \in \mathfrak{S}|_D(\mathcal{R})$ contains at most $|D|$ constants. It follows that

$$|\hat{\Omega}| \leq (|\mathsf{Const}(\mathcal{R})| + 1) g_c(2^{|Pred|}, |D|)$$

and by Lemma 7.2 it follows

$$|\hat{\Omega}| \leq (|\mathsf{Const}(\mathcal{R})| + 1)(|D| + 1)^{2^{|Pred|}} \quad . \qquad \square$$

The obvious observation to be made when comparing $|\Omega(\Sigma, D)|$ to the upper bound of $|\hat{\Omega}|$ is that the latter is not exponential in the number of constants $|D|$. But note that the complexity increased with respect to $|Pred|$. While $|\Omega(\Sigma, D)|$ is exponential in $|Pred|$, the above bound for $|\hat{\Omega}|$ is exponential in $2^{|Pred|}$. However, we believe that this is due to the very coarse estimation in Lemma 7.2. Experiments suggest that g_c can be much better estimated.

Conjecture 7.1. *It holds that* $g_c(n_1, n_2) \leq (n_2 + 1)^{2 \mathsf{ld}\, n_1}$ *for every* $n_1, n_2 \in \mathbb{N}_0$ *(with* $\mathsf{ld}\, 0 = 0$*).*

Confirmation of the above conjecture would result in an upper bound of $(|\mathsf{Const}(\mathcal{R})| + 1)(|D| + 1)^{2|Pred|}$ which is far more beneficial than the result of Theorem 7.1. However, until now no formal proof for the above conjecture has been found.

Table 8 shows some exemplary cardinalities of $\Omega(\Sigma, D)$ and $\hat{\Omega}$ for different value of $|D|$ and $|Pred|$. The knowledge base $\mathcal{R}$ used to determine the $\mathcal{R}$-equivalences classes in $\hat{\Omega}$ mentions a single constant yielding $\mathfrak{S}(\mathcal{R}) = \{\{\mathsf{c}\}, U \setminus \{\mathsf{c}\}\}$ for $\mathsf{Const}(\mathcal{R}) = \{\mathsf{c}\}$. Table 8 shows that especially for this kind of scenarios employing $\hat{\Omega}$ rather than $\Omega(\Sigma, D)$ is computationally beneficial. The numbers in Table 8 also justify the belief in Conjecture 7.1.

7.3 GENERALIZING LIFTED INFERENCE

In the following we discuss the problem of generalizing lifted inference to first-order signatures containing predicates of arity greater one. Let $\Omega(\Sigma)$ be the set of interpretations for some simple relational signature $\Sigma =_{def} (U, Pred, \emptyset)$ that contains at least one non-unary predicate and let $\mathcal{R}$ be some knowledge base on Σ. In contrast to the case without non-unary predicates there is no simple and compact representation of $\Omega(\Sigma, D)/_{\equiv_\mathcal{R}}$ and, in particular, no compact way to enumerate the elements of $\Omega(\Sigma, D)/_{\equiv_\mathcal{R}}$. In contrast to a strictly unary signature there is no similar

$\lvert Pred\rvert$	$\lvert D\rvert$	$\lvert\Omega(\Sigma,D)\rvert$	$\lvert\hat{\Omega}\rvert$
1	1	2	2
1	2	4	4
1	3	8	6
1	4	16	8
1	8	256	16
1	16	65536	32
1	32	4294967296	64
2	1	4	4
2	2	16	16
2	3	64	36
2	4	256	64
2	8	65536	256

Table 8: Comparison of $|\Omega(\Sigma, D)|$ and $|\hat{\Omega}|$ with respect to a simple relational signature $\Sigma = (U, Pred, \emptyset)$ and a knowledge base $\mathcal{R}$ with $|\mathsf{Const}(\mathcal{R})| = 1$

concept of *truth configuration* if the signature contains a non-unary predicate. Consider a predicate $p/2$ and $\mathcal{R}$-equivalence classes S_1 and S_2. Then there are six different instantiations of p that have to be considered as essentially different with respect to $\mathcal{R}$-equivalence. For constants $\mathsf{c}_1 \in S_1$ and $\mathsf{c}_2 \in S_2$ we have the variants $p(\mathsf{c}_1, \mathsf{c}_2)$ and $p(\mathsf{c}_2, \mathsf{c}_1)$; for $\mathsf{c}_1 \in S_i$ we have $p(\mathsf{c}_1, \mathsf{c}_1)$ for $i = 1, 2$; for $\mathsf{c}_1, \mathsf{c}_2 \in S_i$ with $\mathsf{c}_1 \neq \mathsf{c}_2$ we have $p(\mathsf{c}_1, \mathsf{c}_2)$ for $i = 1, 2$. An extended notion of truth configuration must adhere to this combinatorial observation and also take the relations into account that arise by transitivity. Consider the following example.

Example 7.9. Let $\mathcal{R} =_{def} \{r_1, r_2, r_3\}$ be the knowledge base given via

$$r_1 =_{def} (d(\mathsf{a}, \mathsf{b}))[1]$$
$$r_2 =_{def} (d(\mathsf{b}, \mathsf{a}))[1]$$
$$r_3 =_{def} (p(\mathsf{X}, \mathsf{X}) \,|\, d(\mathsf{X}, \mathsf{Y}), d(\mathsf{Y}, \mathsf{X}))[1] \quad .$$

Let $D =_{def} \{\mathsf{a}, \mathsf{b}, \mathsf{c}_1, \ldots, \mathsf{c}_n\}$ be the set of constants, then the $\mathcal{R}$-equivalence classes are $S_1 = \{\mathsf{a}, \mathsf{b}\}$ and $S_2 = \{\mathsf{c}_1, \ldots, \mathsf{c}_n\} \cup U \setminus D$. Consider the Herbrand interpretations ω_1 and ω_2 given via

$$\omega_1 =_{def} \{p(\mathsf{a}, \mathsf{b}), p(\mathsf{a}, \mathsf{a})\} \qquad\qquad \omega_2 =_{def} \{p(\mathsf{a}, \mathsf{b}), p(\mathsf{b}, \mathsf{b})\} \quad .$$

Both interpretations mention the same constants from the same $\mathcal{R}$-equivalence class and both interpretations instantiate p once with the same constant in both arguments and once with different constants in the arguments. However, it holds that $\omega_1 \not\equiv_{\mathcal{R}} \omega_2$ as by switching a and b in ω_1 yields $\{p(\mathsf{b}, \mathsf{a}), p(\mathsf{b}, \mathsf{b})\}$ which might not have the same probability of ω_2 in $\mathcal{I}(\mathcal{R}, D)$.

In the unary case, we used truth configurations to be able to enumerate the elements of $\Omega(\Sigma, D)/_{\equiv_\mathcal{R}}$ in an effective way without considering $\Omega(\Sigma, D)$ itself. In the non-unary case there seems to be no simple way to extend the concept of truth configuration. This observation has also been made by Grove et. al. in (Grove *et al.*, 1994) when they attempted to generalize the notion of entropy of an interpretation to non-unary languages, cf. Section 6.4.2. In particular, the argumentation of Grove et. al. is as follows.[2]

> "In a unary language, [truth configurations] are useful because they are simple descriptions that summarize everything that might be known about a domain element in a model. But consider a language with a single binary predicate [$p(\mathsf{X}, \mathsf{Y})$]. Worlds over this language include all finite graphs (where we think of [$p(\mathsf{X}, \mathsf{Y})$] as holding if there is an edge from [X] to [Y]). In this language, there are infinitely many properties that may be true or false about a domain element. For example, the assertions 'the node [X] has m neighbors' are expressible in the language for each m. Thus, in order to partition the domain elements according to the properties they satisfy, we would need to define infinitely many partitions. Furthermore, it can be shown that 'typically' (i.e., in almost all graphs of sufficiently great size) each node satisfies a different set of first-order properties. Thus, in most graphs, all the nodes are 'different', so a partition of domain elements into a finite number of '[truth configurations]' makes little sense."

Although the framework of (Grove *et al.*, 1994) makes more use of first-order elements than we do, the above argumentation is also applicable for our situation.

However, the approach of lifted inference developed in this chapter can be applied for non-unary languages by determining first $\Omega(\Sigma, D)$ and afterwards (by pair-wise comparisons) merge $\mathcal{R}$-equivalent interpretations to reference worlds (yielding the quotient set $\Omega(\Sigma, D)/_{\equiv_\mathcal{R}}$). Note that we lose the computational advantage of avoiding to consider the full set $\Omega(\Sigma, D)$ in this approach. It is also questionable whether using $\Omega(\Sigma, D)/_{\equiv_\mathcal{R}}$ instead of $\Omega(\Sigma, D)$ for inference is beneficial. In Example 7.9 we illustrated that many Herbrand interpretations turn out to be not $\mathcal{R}$-equivalent. Table 9 shows the cardinalities of both $\Omega(\Sigma, D)$ and $\Omega(\Sigma, D)/_{\equiv_\mathcal{R}}$, depending on the size of D and with respect to a signature containing a single binary predicate and a knowledge base $\mathcal{R}$ with $\mathsf{Const}(\mathcal{R}) = \emptyset$. As $\mathcal{R}$ mentions no constants there is only a single $\mathcal{R}$-equivalence class which makes this scenario the simplest imaginable. Nonetheless, the cardinality of $\Omega(\Sigma, D)/_{\equiv_\mathcal{R}}$—although being significantly smaller than the cardinality of $\Omega(\Sigma, D)$—still seems to grow exponentially in the number of constants considered. Until now, no formal proofs for lower or upper bounds on the growing behavior of $|\Omega(\Sigma, D)/_{\equiv_\mathcal{R}}|$

2 The excerpt is taken from (Grove *et al.*, 1994), p. 67

$\|D\|$	$\|\Omega(\Sigma, D)\|$	$\|\Omega(\Sigma, D)/_{\equiv_{\mathcal{R}}}\|$
1	2	2
2	16	10
3	512	244
4	65536	12235

Table 9: Comparison of $|\Omega(\Sigma, D)|$ and $|\Omega(\Sigma, D)/_{\equiv_{\mathcal{R}}}|$ with respect to a simple relational signature that contains a single binary predicate and a knowledge base $\mathcal{R}$ with $\mathsf{Const}(\mathcal{R}) = \emptyset$

have been found. However, Table 9 gives reason to believe that there is no polynomial upper bound for $|\Omega(\Sigma, D)/_{\equiv_{\mathcal{R}}}|$ in $|D|$. As a consequence, lifted inference can be doubted to be beneficial at all for non-unary languages (in our setting).

7.4 RELATED WORK

As our semantical notions and the use of reasoning based on the principle of maximum entropy is novel for treating relational probabilistic conditionals there are only few works related to the approach developed in this chapter. The observation that a probability function stemming from the application of entropy maximization carries redundant information with respect to indistinguishable constants has also been made in (Fisseler, 2010) for the case of grounding semantics. There, the term *involution invariance* is used to catch the equivalent behavior of ground conditionals with respect to their Lagrange multipliers in (6.60) on page 181. However, (Fisseler, 2010) does not consider exploiting this observation in order to represent the ME-function explicitly by reference worlds. The notion of *lifted inference* used in this chapter has been adopted from the works (Poole, 2003; de Salvo Braz *et al.*, 2005; Milch *et al.*, 2008) which also use this notion to describe effective reasoning procedures for relational probabilistic knowledge. Although the knowledge representation formalisms of those works differ from our work, the motivation and ideas of those works are similar to ours. We have a closer look on those works in the following.

The work (de Salvo Braz *et al.*, 2005)—which extends work begun in (Poole, 2003)—develops an algorithm for lifted probabilistic inference in parametrized belief networks (Horsch and Poole, 1990) which are similar to Bayesian logic programs, cf. Section 2.4.1. Consequently, we use BLPs to illustrate the ideas of (de Salvo Braz *et al.*, 2005) and (Poole, 2003). The basic idea of (Poole, 2003; de Salvo Braz *et al.*, 2005) is the observation that in order to determine the probability of some query, the information used to infer the probability can be partitioned with respect to the information we have for specific individuals. This approach uses the technique of *variable elimination* of (Zhang and Poole, 1996) to simplify equations like (2.10) on page 37 with respect to equivalencies of undistinguishable constants. We do not give a formal description of the algorithms developed in (Poole, 2003; de Salvo Braz *et al.*, 2005) but rather give an idea of the approach by means of an example. Consider the Bayesian clause c given via

$$c =_{def} (p(\mathsf{X}) \,|\, q(\mathsf{X},\mathsf{Y}), r(\mathsf{Y}))$$

and some arbitrary conditional probability distribution cpd_c for c. Let furthermore *noisy-or* be the designated combining rule for c and let

$$\begin{aligned} E_{n,m} =_{def} \{&q(\mathsf{c},\mathsf{d}_1),\ldots,q(\mathsf{c},\mathsf{d}_{n+m})\}\cup \\ &\{r(\mathsf{d}_1),\ldots,r(\mathsf{d}_n),\neg r(\mathsf{d}_{n+1}),\ldots,\neg r(\mathsf{d}_{n+m})\} \end{aligned}$$

be some evidence with $n, m \in \mathbb{N}$. In order to compute the probability of the query $Q = (p(\mathsf{c}) \,|\, E)$ one has to instantiate a ground Bayesian network for the node $p(\mathsf{c})$ with parents $q(\mathsf{d}_1),\ldots,q(\mathsf{d}_{n+m}), r(\mathsf{c},\mathsf{d}_1),\ldots,r(\mathsf{c},\mathsf{d}_{n+m})$. This amounts to

$$\begin{aligned} P(Q) = 1 - (1 - P(p(\mathsf{c}) \,|\, q(\mathsf{c},\mathsf{d}_1), r(\mathsf{d}_1)))\cdot\ldots\cdot \\ (1 - P(p(\mathsf{c}) \,|\, q(\mathsf{c},\mathsf{d}_{n+m}), r(\mathsf{d}_{n+m}))) \quad . \end{aligned}$$

Note that we have the same information for the constants $\mathsf{d}_1,\ldots,\mathsf{d}_n$ and $\mathsf{d}_{n+1},\ldots,\mathsf{d}_m$, respectively. It follows that it holds that

$$\begin{aligned} P(p(\mathsf{c}) \,|\, q(\mathsf{c},\mathsf{d}_1), r(\mathsf{d}_1)) &= \ldots = P(p(\mathsf{c}) \,|\, q(\mathsf{c},\mathsf{d}_n), r(\mathsf{d}_n)) \\ &= \mathsf{cpd}_c(\mathsf{true},\mathsf{true},\mathsf{true}) \\ P(p(\mathsf{c}) \,|\, q(\mathsf{c},\mathsf{d}_{n+1}), r(\mathsf{d}_{n+1})) &= \ldots = P(p(\mathsf{c}) \,|\, q(\mathsf{c},\mathsf{d}_{n+m}), r(\mathsf{d}_{n+m})) \\ &= \mathsf{cpd}_c(\mathsf{true},\mathsf{true},\mathsf{false}) \end{aligned}$$

and therefore

$$P(Q) = 1 - (1 - \mathsf{cpd}_c(\mathsf{true},\mathsf{true},\mathsf{true}))^n (1 - \mathsf{cpd}_c(\mathsf{true},\mathsf{true},\mathsf{false}))^m \; .$$

As one can see, we can avoid grounding the full Bayesian logic program by just considering *prototypical* groundings for c. In (Poole, 2003; de Salvo Braz *et al.*, 2005) this idea is elaborated and a series of algorithms is developed that apply this approach to general parametrized belief networks (or Bayesian logic programs). Obviously, the ideas of (Poole, 2003; de Salvo Braz *et al.*, 2005) are very similar to ours and differences lie mainly in the framework used for knowledge representation and the technical implementation. The work (Poole, 2003) uses parametrized belief networks and inference bases on Bayesian networks and (de Salvo Braz *et al.*, 2005) uses a framework similar to Markov logic networks (see Section 2.4.2). Note that both formalisms are first-order extensions of probabilistic networks as discussed in Section 2.2. We use relational probabilistic conditional logic and inference based on the principle of maximum entropy. Furthermore, we developed an explicit computational model for representing prototypically uniform probability functions and showed that the use of this model is beneficial in terms of computational complexity. In (Poole, 2003) no hints on the computational advantages of applying first-order variable elimination are given but (de Salvo Braz *et al.*, 2005) gives an experimental evaluation that resembles our observations from Conjecture 7.1 on page 212.

The work (Milch *et al.*, 2008) extends the approaches of (de Salvo Braz *et al.*, 2005) and (Poole, 2003) by considering *counting formulas*. A counting formula is a probabilistic rule like e. g.

$$c' =_{def} (expert(\mathsf{X}) \,|\, \#_{\mathsf{Y}}[publication(\mathsf{X}, \mathsf{Y})])$$

which states that the probability of X being an expert (in some field) depends on his or her number of publications (in the very same field). The notion of conditional probability distributions can be extended to adhere for counting formulas by allowing natural number to occur as parameters. For example, a suitable conditional probability distribution $\mathsf{cpd}_{c'}$ for c' is a function $\mathsf{cpd}_{c'} : \mathbb{B} \times \mathbb{N} \to [0,1]$ with e. g. $\mathsf{cpd}_{c'}(\mathsf{true}, 5) = 0.7$. So, $\mathsf{cpd}_{c'}$ assigns the probability 0.7 to X being an expert if X has 5 publications. Counting formulas do not take the actual instances, i. e. the actual publications, but only the number of them into account. In (Milch *et al.*, 2008) it is argued that there are many probabilistic relationships that can be modeled more appropriately with counting formulas than with ordinary clauses. Having enriched the knowledge representation formalism with counting formulas the work (Milch *et al.*, 2008) shows that the approach of (de Salvo Braz *et al.*, 2005; Poole, 2003) can be generalized to allow for a *lifted* way to deal with these formulas as well. More recent work extends lifted inference to other *aggregation functions* such as maximum, minimum, or average (Kisyński and Poole, 2009).

7.5 SUMMARY AND DISCUSSION

In this chapter, we developed a computational account for effective probabilistic inference with relational probabilistic conditionals. In particular, we introduced the notions of reference worlds and condensed probability functions which allow for a compact representation of probability functions that arise from the application of inference operators satisfying (Prototypical Indifference). Due to the equivalence of interpretations that only differ in constants from the same $\mathcal{R}$-equivalence class, those interpretations can be pooled together and used as a single entity. Condensed probability functions are defined on the set of reference worlds and exhibit the same reasoning behavior as the original probability functions, given that those are indifferent with respect to constants from the same $\mathcal{R}$-equivalence class. We also developed an inference procedure that determines the probability of some sentence without considering all interpretations but only reference worlds. Furthermore, we showed that the inference operators developed in the previous chapter can be modified in order to compute the condensed ME-function in a single step without considering the Herbrand interpretations at all. We analyzed the computational benefits of our approach and concluded that we avoid the exponential blow-up in the number of constants that have to be considered. Our approach is—using the given formalization—only applicable for unary languages and we briefly discussed the issues that arise when considering non-unary languages. Finally, we compared our approach to the existing literature on lifted inference in first-order probabilistic models.

The approach developed in this chapter gives directions for efficient implementation of reasoning based on the principle of maximum entropy. However, the work reported is only a first step towards this goal and suffers from two major discrepancies. Firstly, we restricted lifted inference to the case of unary languages which, in practice, is a demand that cannot be easily fulfilled. One of the main advantages of first-order extensions of probabilistic reasoning is the capability to reason over relations. By abandoning non-unary predicates we lose much of the expressive power of relational probabilistic conditional logic. However, note that even by restricting attention to unary languages we do not get the equivalence to propositional probabilistic models due to our semantical notions. For example, the knowledge base $\mathcal{R} =_{def} \{r_1, r_2\}$ with

$$r_1 =_{def} (\mathit{flies}(\mathsf{X}))[0.9]$$
$$r_2 =_{def} (\mathit{flies}(\mathsf{tweety}))[0.3]$$

cannot be represented with a propositional probabilistic model that exhibits the same inference behavior. Secondly, in order to determine the (condensed) ME-function of a knowledge base $\mathcal{R}$ we have to solve a complex optimization problem. The amount of time needed to obtain an optimal solution is hard to predict in general as it depends on both the structure of the knowledge base and the optimization algorithm used for solving the problem, cf. (Boyd and Vandenberghe, 2004). However, there are approaches to avoid solving problems like (6.15) and (6.42) (see pages 170 and 173, respectively) for the propositional case. For example, in (Meyer, 1997) an approximate algorithm for computing the ME-function for propositional probabilistic conditional logic is developed. This algorithm bases on an iterative computation of the ME-function using a compact representation that employs *clique trees*. In this approach the ME-function is not represented using interpretations but sub-interpretations that stem from certain relationships and independences of propositions in the knowledge base. Furthermore, instead of solving the optimization problem (2.8) (see page 32) in a direct fashion, a series of probability functions is determined that converges to the ME-function. The algorithm of (Meyer, 1997) benefits from several characteristic properties of the ME-function in the propositional case and it is to investigate if these properties (or similar ones) can be found for our semantical approaches.

8 SUMMARY AND FINAL REMARKS

In this chapter we conclude this thesis by summarizing its content, giving hints to further and future work, and by making some final remarks.

8.1 SUMMARY

In this thesis we discussed the problem of probabilistic reasoning with incomplete and inconsistent information. We addressed this problem using both propositional and relational conditional logic.

In Chapters 3 and 4 we studied inconsistencies in propositional probabilistic conditional logic. Inconsistencies arise easily when experts share their beliefs to build up a common knowledge base. When relying on model-based inference methods such as inference via the principle of maximum entropy then inconsistencies render a knowledge base useless as there is no model to select from. Consequently, we investigated the problem of inconsistent knowledge bases from both an analytical and practical perspective. In Chapter 3 we introduced inconsistency measures for the framework of probabilistic conditional logic. Inconsistency measures are analytical tools to assess the severity of inconsistencies in a knowledge base and have been used previously in classical frameworks only. To our knowledge, our investigation of inconsistency measures on probabilistic conditional logic is the first one that covered this matter in such a breadth. We developed a list of rationality postulates for measuring inconsistency in a probabilistic framework and developed an approach that satisfies most of those. This measure bases on the distance of an inconsistent knowledge base to the nearest consistent knowledge base by interpreting probabilities of conditionals as coordinates. Though computationally this novel measure is hard to handle we developed approximations that are fast in practice. We continued with a more practical perspective on inconsistencies in probabilistic conditional logic in Chapter 4. First, we extended the concept of inconsistency measures towards culpability measures. A culpability measures assigns no degree of inconsistency to the whole knowledge base but to each probabilistic conditional of a knowledge base. That way one can determine the conditionals to be blamed for the inconsistency. We developed two culpability measures—one employing the Shapley value from coalition game theory and one extending the inconsistency measure of Chapter 3—and showed that these measures satisfy several desirable properties. We continued by employing inconsistency and culpability measures for the task of solving conflicts. We presented a series of rationality postulates for restoring consistency in inconsistent knowledge bases and developed two families of consistency restoring methods. The first one bases on the notion of a creeping function and is constructive in nature. The second one is declarative and employs distance minimization in a balanced fashion. Both approaches satisfy several rationality postulates and we illustrated their advantages and disadvantages in several examples.

Probabilistic conditional logic has been well-studied in the literature before but we also addressed the problem of relational probabilistic reasoning in probabilistic conditional logic. Chapters 5, 6, and 7 covered the issue of representing and reasoning with relational probabilistic conditionals. As this is quite a novel research area we started in Chapter 5 with laying the syntactical and semantical foundations for knowledge representation with relational probabilistic conditional logic. We introduced novel semantical approaches that incorporate information on the population under consideration into the formal interpretation of conditionals. We continued in Chapter 6 with a discussion on the problem of reasoning in this novel framework and developed a series of postulates for model-based inductive inference. As the principle of maximum entropy has proven to be a powerful approach for reasoning in propositional probabilistic conditional logic we extended this principle to the relational case and developed inference methods based on maximum entropy for the new semantics. These approaches satisfy (most of) the rationality postulates and we illustrated their use on several examples. Though, the computational task of determining a probability function with maximum entropy in relational probabilistic conditional logic is quite demanding. That is why we investigated lifted inference with relational probabilistic conditionals in Chapter 7. Due to the satisfaction of (Prototypical Indifference) of the ME-inference methods, a probability function with maximum entropy can be compactly represented by a condensed probability function. We developed constructive approaches for computing a condensed probability function with maximum entropy and determining the probability of an arbitrary sentence. We showed that by using lifted inference in unary languages we gain substantial advantages regarding computational complexity.

8.2 FURTHER AND FUTURE WORK

The work reported in this thesis is not the only one that investigates probabilistic reasoning with incomplete and inconsistent information. Although we have given pointers to relevant literature throughout this thesis we want to stress two specific works that are highly related to our works. The problem of measuring and dealing with inconsistency in probabilistic frameworks has also been investigated in (Daniel, 2009) which has already been discussed in Section 3.5 and in Section 4.5. The main difference of the course of action followed in (Daniel, 2009) in comparison to the one here, is that in (Daniel, 2009) reasoning based on the principle of maximum entropy is extended to be applicable on inconsistent knowledge bases. There is no modification of the existing knowledge base which is the approach pursued in this thesis. But this distinction is merely of relevance as each approach can be transformed into the other without problems. Having a method for restoring consistency and an inference mechanism that only works on consistent knowledge bases, by combining these two into one process one gets an inference mechanism on inconsistent knowledge bases. Further, having an inference mechanism that works on inconsistent knowledge bases allows for consistent modification of the knowledge base as new probabilities for the conditionals can be determined by the inference mechanism. More differences arise on a technical level which were already discussed in Section 3.5 and in Section 4.5. The second work we like to highlight is (Fisseler, 2010) which discusses learning and modeling with probabilistic conditional logic and also considers reasoning based on the principle of maximum entropy in a relational context. However, (Fisseler, 2010) focuses an grounding semantics and encompasses the problem of inconsistent knowledge bases by introducing constraint formulas on variables. This approach requires that the knowledge engineer is capable of identifying exceptions for open probabilistic conditionals himself and we already argued that this requirement hampers a modular usage of knowledge bases, cf. Section 5.5.1. Nonetheless, (Fisseler, 2010) investigates the properties of the ME-function for a knowledge base in great detail and identifies several criteria that allow for an effective computation of this function. The work (Loh *et al.*, 2010) extends the grounding semantics of (Fisseler, 2010) by considering different grounding strategies and an in-depth comparison of our work with (Loh *et al.*, 2010) has been already given in Sections 5.5.1 and 6.4.1.

Current and future work on the areas covered in this thesis include extending the approaches to inconsistency measurement and solving conflicts on more expressive probabilistic languages. We already addressed some of these issues in Section 3.4.3 where we extended the distance minimization inconsistency measure to bounded conditionals and linear knowledge bases. These extensions have to be investigated in more depth and other language extensions such as *polynomial probabilistic knowledge bases* (Daniel, 2009) should be looked into as well. The issue of consistency restoration of probabilistic conditional knowledge bases is quite novel to the scientific community. To our knowledge, only the works (Finthammer *et al.*, 2007; Rödder and Xu, 2001) have discussed this problem before, see Sections 4.5.2 and 4.5.3. Consequently, we only took first steps and more research and (empirical) evaluation is required. This also applies to some extent to the work on relational probabilistic conditional logic. However, the main issue for future work there lies in efficient inference algorithms. Even by employing lifted inference the bound on the size of the application is too low for real-world applications. One particular direction for future research here is investigating whether the approach of (Meyer, 1997) is applicable in the relational domain, cf. Section 7.5. Furthermore, the issue of inconsistency is present in the relational context as well and another direction for future research lies in investigating to what extent the discussion on inconsistency measures in this thesis also applies to the relational case.

Most of the methods developed in this thesis are already prototypically implemented and available within the Tweety library for artificial intelligence[1]. Future work also includes further development of these implementations.

8.3 CONCLUSION

In this thesis we discussed semantical problems of probabilistic conditional logic. On the one hand, we discussed inconsistencies in propositional probabilistic conditional logic and hence investigated semantical deficiencies on contradictory information. On the other hand, we lay semantical foundations for relational probabilistic conditional logic that allow for a consistent treatment of "contradictory" information from the beginning. For the latter, "contradictory" means differing information on population and individuals. We addressed and answered the research questions, that were posed in the introduction (see Section 1.2), as follows. In particular, we discussed the question

> *How to analyze inconsistencies in probabilistic conditional logic and how to measure their severities?*

in great detail by translating inconsistency measures for classical theories into the probabilistic setting and developing a novel inconsistency measure that is apt for probabilistic conditional logic. We showed that this measure can be approximated with effective methods and extended it to more expressive probabilistic settings. These tools help the knowledge engineer to fulfill his or her task of building a consistent knowledge base that can be used for reasoning. For settings where a manual repair of an inconsistent knowledge base is infeasible we discussed the research question

> *How to restore consistency in inconsistent probabilistic knowledge?*

by investigating several approaches for automatic restoration of consistency that are based on culpability measures. It turned out to be quite hard to show the satisfaction of several desirable properties for consistency restorers but, nevertheless, we gave a set of different consistency restorers the knowledge engineer can choose from. Both inconsistency measures and consistency restorers give many insights into to semantical notion of consistency in probabilistic conditional logic. However, when turning to relational extensions for probabilistic reasoning we discovered that the notion of consistency is hard to grasp due to the inadequacy of the traditional semantics in a relational setting. Consequently, we discussed the question

> *How to express relational knowledge in probabilistic conditional logic and what is a meaningful interpretation of relational conditionals?*

1 http://sourceforge.net/projects/tweety/

and introduced two novel semantics for relational probabilistic conditional logic. These semantics extend classical semantics by allowing the treatment of both population-based statements and statements on the degree of belief. We used these semantical notions to obtain an answer to the question

How can one reason with relational probabilistic conditionals?

and extended the reasoning based on the principle of maximum entropy to the relational setting. Our inference operators turned out to be powerful with respect to many commonsense properties and we illustrated their behavior on several benchmark examples. Although reasoning in RPCL in a straightforward fashion is hardly feasible in general we proposed a mechanism that allows for a polynomial representation in the number of constants.

To summarize, in this thesis we discussed the general research question

How to infer knowledge from incomplete, uncertain, and possibly inconsistent information?

in both propositional and relational probabilistic conditional logic. We allowed for incomplete information in knowledge bases as reasoning based on the principle of maximum entropy allows completing this information in an unbiased and sound way. Uncertainty can be expressed using probabilities and we differentiated between uncertainty of subjective beliefs (for propositional probabilistic conditional logic and for statements on the degree of belief in relational probabilistic conditional logic) and uncertainty of population-based statements (in relational probabilistic conditional logic). We dealt with inconsistency in propositional probabilistic conditional logic by investigating inconsistency measures, culpability measures, and consistency restorers and we dealt with inconsistency in relational probabilistic conditional logic by investigating novel semantical notions that circumvent inconsistencies when the information represented is to be treated as consistent from a commonsensical perspective.

Reasoning with incomplete, uncertain, and possibly inconsistent information is of crucial concern in every real-world expert system as complete and correct information is rarely available. In this thesis we dealt with several problems in this area, with a focus on inconsistent information.

A PROOFS OF TECHNICAL RESULTS

Proposition 2.1 (page 20). *Let $X_1, X_2 \subseteq \mathcal{X}$. Then it holds that*

1. $P(X_1) = \sum_{x \in X_1} P(x)$,
2. $P(X_1 \cup X_2) = P(X_1) + P(X_2) - P(X_1 \cap X_2)$, *and*
3. $P(\mathcal{X} \setminus X_1) = 1 - P(X_1)$.

Proof.

1. We prove the statement by induction on the cardinality of X_1. For $X_1 = \{x_1\}$ of cardinality one it holds that $P(X_1) = P(\{x_1\}) = \sum_{x \in X_1} P(x)$ by definition. Let X_1 be of cardinality n, i.e., $X_1 = X_1' \cup \{x_n\}$ for some $X_1' \subseteq \mathcal{X}$ of cardinality $n-1$ and some $x_n \in \mathcal{X}$. By assumption $P(X_1') = \sum_{x \in X_1'} P(x)$ and due to the second Kolmogorov axiom of probability (see Definition 2.15 on page 19) it holds that $P(X_1) = P(X_1') + P(x_n)$ and it follows $P(X_1) = \sum_{x \in X_1'} P(x) + P(x_n) = \sum_{x \in X_1} P(x)$. If the cardinality of $X_1 = \{x_1, x_2, \ldots\}$ is infinite then the observation holds for every finite subset of X_1 and as $P(\mathcal{X})$ is finite due to the first Kolmogorov axiom of probability (see Definition 2.15 on page 19) every sum on probabilities of elements of $\mathcal{X}$ is finite as well yielding
$$\begin{aligned} P(X_1) &= \lim_{n\to\infty} P\left(\bigcup_{i=1}^{n} \{x_i\}\right) = \lim_{n\to\infty} \sum_{i=1}^{n} P(x_i) \\ &= \sum_{i=1}^{\infty} P(x_i) \quad . \end{aligned}$$
2. From the second Kolmogorov axiom (see Definition 2.15 on page 19) and $(S \setminus S') \cap (S \cap S') = \emptyset$ for sets S, S' it follows that
$$\begin{aligned} &P(X_1) + P(X_2) - P(X_1 \cap X_2) \\ &\quad = P(X_1 \setminus X_2) + P(X_1 \cap X_2) + \\ &\qquad\quad P(X_2 \setminus X_1) + P(X_1 \cap X_2) - P(X_1 \cap X_2) \\ &\quad = P(X_1 \setminus X_2) + P(X_1 \cap X_2) + P(X_2 \setminus X_1) \\ &\quad = P((X_1 \setminus X_2) \cup (X_1 \cap X_2) \cup (X_2 \setminus X_1)) \\ &\quad = P(X_1 \cup X_2) \end{aligned}$$
3. As $(\mathcal{X} \setminus X_1) \cap X_1 = \emptyset$, it holds, due to the second Kolmogorov axiom of probability (see Definition 2.15 on page 19), that $P(\mathcal{X} \setminus X_1) +$

$P(X_1) = P(\mathcal{X} \setminus X_1 \cup X_1) = P(\mathcal{X})$ and, due to the first Kolmogorov axiom of probability, $P(\mathcal{X}) = 1$. $\square$

Proposition 2.2 (page 21). *Let P be a probability function on $\mathcal{L}(\mathsf{At})$ and $\phi, \psi \in \mathcal{L}(\mathsf{At})$.*

1. *If $\phi \models^P \perp$ then $P(\phi) = 0$.*
2. *If $\top \models^P \phi$ then $P(\phi) = 1$.*
3. *If $\phi \equiv^P \psi$ then $P(\phi) = P(\psi)$.*
4. *If $\phi \wedge \psi \models^P \perp$ then $P(\phi \vee \psi) = P(\phi) + P(\psi)$.*
5. *It holds that $P(\neg\phi) = 1 - P(\psi)$.*
6. *If $\phi \models^P \psi$ then $P(\phi) \leq P(\psi)$.*

Proof.

1. If $\phi \models^P \perp$ then $\mathsf{Mod}^P(\phi) \subseteq \mathsf{Mod}^P(\perp) = \emptyset$, i.e. $\mathsf{Mod}^P(\phi) = \emptyset$. It follows that there is no $\omega \in \Omega(\mathsf{At})$ with $\omega \models^P \phi$ and therefore $P(\phi) = \sum_{\omega \models^P \phi,\ \omega \in \Omega(\mathsf{At})} P(\omega) = 0$.

2. If $\top \models^P \phi$ then $\mathsf{Mod}^P(\top) \subseteq \mathsf{Mod}^P(\phi)$, i.e. $\mathsf{Mod}^P(\phi) = \mathsf{Int}(\mathsf{At})$. It follows that for every $\omega \in \Omega(\mathsf{At})$ it holds that $\omega \models^P \phi$ and therefore $P(\phi) = \sum_{\omega \models^P \phi,\ \omega \in \Omega(\mathsf{At})} P(\omega) = \sum_{\omega \in \Omega(\mathsf{At})} P(\omega) = 1$.

3. If $\phi \equiv^P \psi$ then $\mathsf{Mod}^P(\phi) = \mathsf{Mod}^P(\psi)$ and therefore $\omega \models^P \phi$ whenever $\omega \models^P \psi$ for every $\omega \in \Omega(\mathsf{At})$. It follows that
$$P(\phi) = \sum_{\omega \models^P \phi,\ \omega \in \Omega(\mathsf{At})} P(\omega) = \sum_{\omega \models^P \psi,\ \omega \in \Omega(\mathsf{At})} P(\omega) = P(\psi) \quad .$$

4. This follows directly from the second Kolmogorov axiom (see Definition 2.15 on page 19) and due to $\mathsf{Mod}^P(\phi \wedge \psi) = \mathsf{Mod}^P(\phi) \cap \mathsf{Mod}^P(\psi)$ and $\mathsf{Mod}^P(\phi \vee \psi) = \mathsf{Mod}^P(\phi) \cup \mathsf{Mod}^P(\psi)$.

5. This follows directly from property 3.) of Proposition 2.1 on page 20.

6. If $\phi \models^P \psi$ then $\mathsf{Mod}^P(\phi) \subseteq \mathsf{Mod}^P(\psi)$ and therefore $\{\omega \in \Omega(\mathsf{At}) \mid \omega \models^P \phi\} \subseteq \{\omega \in \Omega(\mathsf{At}) \mid \omega \models^P \psi\}$. It follows that
$$P(\phi) = \sum_{\omega \models^P \phi,\ \omega \in \Omega(\mathsf{At})} P(\omega) \leq \sum_{\omega \models^P \psi,\ \omega \in \Omega(\mathsf{At})} P(\omega) = P(\psi)$$
as $P(\omega) \geq 0$ for every $\omega \in \Omega(\mathsf{At})$. $\square$

Proposition 2.3 (page 29). *Let $(\psi \mid \phi)[d]$ and $(\psi' \mid \phi')[d']$ be some probabilistic conditionals. It holds that $(\psi \mid \phi)[d] \equiv^{pr} (\psi' \mid \phi')[d']$ if and only if either*

1. *$\phi \equiv^P \phi'$ and $\psi \wedge \phi \equiv^P \psi' \wedge \phi'$ and $d = d'$ or*

2. *$\phi \equiv^P \phi'$ and $\psi \wedge \phi \equiv^P \overline{\psi'} \wedge \phi'$ and $d = 1 - d'$ or*

3. *both $(\psi \,|\, \phi)[d]$ and $(\psi' \,|\, \phi')[d']$ are not self-consistent or*

4. *both $(\psi \,|\, \phi)[d]$ and $(\psi' \,|\, \phi')[d']$ are tautological.*

Proof. We have to show both directions. First, let it hold that $\phi \equiv^{\mathrm{P}} \phi'$, $\psi \wedge \phi \equiv^{\mathrm{P}} \psi' \wedge \phi'$, and $d = d'$. Then it holds that

$$\begin{aligned} P \models^{pr} (\psi \,|\, \phi)[d] \quad &\textit{iff} \quad P(\psi \wedge \phi) = d \cdot P(\phi) \\ &\textit{iff} \quad P(\psi' \wedge \phi') = d' \cdot P(\phi') \\ &\textit{iff} \quad P \models^{pr} (\psi' \,|\, \phi')[d'] \end{aligned}$$

for every probability function P. It follows $(\psi \,|\, \phi)[d] \equiv^{\mathrm{pr}} (\psi' \,|\, \phi')[d']$. A similar reasoning applies for the case of $\phi \equiv^{\mathrm{P}} \phi'$ and $\psi \equiv^{\mathrm{P}} \overline{\psi'}$ and $d = 1 - d'$. If both $(\psi \,|\, \phi)[d]$ and $(\psi' \,|\, \phi')[d']$ are not self-consistent or are tautological then the condition of equivalence is trivially satisfied. For the other direction, let it be the case that $(\psi \,|\, \phi)[d] \equiv^{\mathrm{pr}} (\psi' \,|\, \phi')[d']$ and assume that 1.), 2.), 3.), and 4.) are false. We abbreviate

$$\begin{aligned} M_{12} &=_{def} \mathsf{Mod}^{\mathrm{P}}(\phi \wedge \psi) \qquad & M_1 &=_{def} \mathsf{Mod}^{\mathrm{P}}(\phi) \\ M'_{12} &=_{def} \mathsf{Mod}^{\mathrm{P}}(\phi' \wedge \psi') \qquad & M'_1 &=_{def} \mathsf{Mod}^{\mathrm{P}}(\phi') \end{aligned}$$

Note that it holds that $M_{12} \subseteq M_1$ and $M'_{12} \subseteq M'_1$. As 4.) is assumed to be false, not both $(\psi \,|\, \phi)[d]$ and $(\psi' \,|\, \phi')[d']$ are tautological. Without loss of generality, assume that $(\psi \,|\, \phi)[d]$ is tautological but $(\psi' \,|\, \phi')[d']$ is not. In this case there is a P with $P \models^{pr} (\psi \,|\, \phi)[d]$ but $P \not\models^{pr} (\psi' \,|\, \phi')[d']$, so it cannot be the case that $(\psi \,|\, \phi)[d] \equiv^{\mathrm{pr}} (\psi' \,|\, \phi')[d']$ contradicting the premise. It follows that both $(\psi \,|\, \phi)[d]$ and $(\psi' \,|\, \phi')[d']$ are not tautological. With a similar argumentation it follows that both $(\psi \,|\, \phi)[d]$ and $(\psi' \,|\, \phi')[d']$ are self-consistent. It follows that $M_{12} \neq M_1$ and $M'_{12} \neq M'_1$ as e. g. for $(\psi \,|\, \phi)[d]$ with $M_{12} = M_1$ it follows that $\phi \equiv^{\mathrm{P}} \phi \wedge \psi$ and consecutively $\phi \models^{\mathrm{P}} \psi$ which means that $(\psi \,|\, \phi)[d]$ is tautological for $d = 1$ and not self-consistent for $d \in (0, 1]$ contradicting the previous observation. Consider the following case differentiation:

1. $M_1 \cap M'_1 = \emptyset$:
   Define $P_1 : \Omega(\mathsf{At}) \to [0, 1]$ as follows

$$\begin{aligned} P_1(\omega_I) &=_{def} \frac{d}{3} \cdot \frac{1}{|M_{12}|} && \text{for } I \in M_{12} \\ P_1(\omega_I) &=_{def} \frac{1-d}{3} \frac{1}{|M_1 \setminus M_{12}|} && \text{for } I \in M_1 \setminus M_{12} \\ P_1(\omega_I) &=_{def} 0 && \text{for } I \in M'_1 \setminus M'_{12} \\ P_1(\omega_I) &=_{def} \frac{2}{3} \frac{1}{|M'_{12}|} && \text{for } I \in M'_{12} \\ P_1(\omega_I) &=_{def} 0 && \text{otherwise} \end{aligned}$$

Furthermore, define $P_2 : \Omega(\mathsf{At}) \to [0,1]$ as follows

$$\begin{aligned}
P_2(\omega_I) &=_{def} P_1(\omega_I) && \text{for } I \in M_1 \\
P_2(\omega_I) &=_{def} \frac{2}{3}\frac{1}{|M'_1|} && \text{for } I \in M'_1 \\
P_2(\omega_I) &=_{def} 0 && \text{otherwise}
\end{aligned}$$

Then P_1 is a probability function as

$$\begin{aligned}
\sum_{\omega \in \Omega(\mathsf{At})} P_1(\omega) = |M_{12}| \cdot \frac{d}{3} \cdot \frac{1}{|M_{12}|} + \\
|M_1 \setminus M_{12}| \cdot \frac{1-d}{3} \frac{1}{|M_1 \setminus M_{12}|} + \\
|M'_{12}| \cdot \frac{2}{3} \frac{1}{|M'_{12}|} \\
= 1
\end{aligned}$$

and it holds that

$$\begin{aligned}
P_1(\psi \,|\, \phi) &= \frac{P_1(\phi \wedge \psi)}{P_1(\phi)} \\
&= \frac{|M_{12}| \cdot \frac{d}{3} \cdot \frac{1}{|M_{12}|}}{|M_{12}| \cdot \frac{d}{3} \cdot \frac{1}{|M_{12}|} + |M_1 \setminus M_{12}| \cdot \frac{1-d}{3} \frac{1}{|M_1 \setminus M_{12}|}} \\
&= d
\end{aligned}$$

and

$$P_1(\psi' \,|\, \phi') = \frac{P_1(\phi' \wedge \psi')}{P_1(\phi')} = 1 \quad .$$

Furthermore, P_2 is a probability function as

$$\begin{aligned}
\sum_{\omega \in \Omega(\mathsf{At})} P_2(\omega) = |M_{12}| \cdot \frac{d}{3} \cdot \frac{1}{|M_{12}|} + \\
|M_1 \setminus M_{12}| \cdot \frac{1-d}{3} \frac{1}{|M_1 \setminus M_{12}|} + \\
|M'_1| \cdot \frac{2}{3} \frac{1}{|M'_1|} \\
= 1
\end{aligned}$$

and it holds that $P_2(\psi \,|\, \phi) = d$ and

$$P_2(\psi' \,|\, \phi') = \frac{P_2(\phi' \wedge \psi')}{P_2(\phi'} = \frac{\sum_{I \in M_{12}} P(\omega_I)}{\sum_{I \in M_1} P(\omega_I)} = \frac{|M_{12}| \cdot \frac{1}{|M'_1|}}{|M_1| \cdot \frac{1}{|M'_1|}} < 1$$

as $|M'_1| > |M'_{12}|$ due to $M'_1 \neq M'_{12}$ and $M'_{12} \subseteq M'_1$. It follows that both P_1 and P_2 satisfy $(\psi \,|\, \phi)[d]$ but at least one of P_1 and P_2 cannot satisfy $(\psi' \,|\, \phi')[d']$. Therefore $(\psi \,|\, \phi)[d] \not\equiv^{\text{pr}} (\psi' \,|\, \phi')[d']$ contradicting the premise.

2. $M_1 \cap M'_1 \neq \emptyset$, $M_1 \not\subseteq M'_1$, $M'_1 \not\subseteq M_1$, $M_{12} \not\subseteq M'_1$, and $M'_{12} \not\subseteq M_1$:
We abbreviate $L =_{def} M_1 \cap M'_1$. Define $P_1 : \Omega(\mathsf{At}) \to [0,1]$ as follows

$$\begin{aligned}
P_1(\omega_I) &=_{def} \frac{d}{3} \cdot \frac{1}{|M_{12} \setminus L|} && \text{for } I \in M_{12} \setminus L \\
P_1(\omega_I) &=_{def} \frac{1-d}{3} \cdot \frac{1}{|(M_1 \setminus M_{12}) \setminus L|} && \text{for } I \in (M_1 \setminus M_{12}) \setminus L \\
P_1(\omega_I) &=_{def} \frac{2}{3} \cdot \frac{1}{|M'_{12} \setminus L|} && \text{for } I \in M'_{12} \setminus L \\
P_1(\omega_I) &=_{def} 0 && \text{otherwise}
\end{aligned}$$

Furthermore, define $P_2 : \Omega(\mathsf{At}) \to [0,1]$ as follows

$$\begin{aligned}
P_2(\omega_I) &=_{def} \frac{d}{3} \cdot \frac{1}{|M_{12} \setminus L|} && \text{for } I \in M_{12} \setminus L \\
P_2(\omega_I) &=_{def} \frac{1-d}{3} \cdot \frac{1}{|(M_1 \setminus M_{12}) \setminus L|} && \text{for } I \in (M_1 \setminus M_{12}) \setminus L \\
P_2(\omega_I) &=_{def} \frac{2}{3} \cdot \frac{1}{|(M'_1 \setminus M'_{12}) \setminus L|} && \text{for } I \in (M'_1 \setminus M'_{12}) \setminus L \\
P_2(\omega_I) &=_{def} 0 && \text{otherwise}
\end{aligned}$$

Note that all sets $M_{12} \setminus L$, $(M_1 \setminus M_{12}) \setminus L$, $M'_{12} \setminus L$, $(M'_1 \setminus M'_{12}) \setminus L$ are non-empty due to the above assumptions. Then P_1 is a probability function as

$$\begin{aligned}
&\sum_{\omega \in \Omega(\mathsf{At})} P_1(\omega) \\
&\qquad = |M_{12} \setminus L| \cdot \frac{d}{3} \cdot \frac{1}{|M_{12} \setminus L|} + \\
&\qquad\qquad |(M_1 \setminus M_{12}) \setminus L| \cdot \frac{1-d}{3} \cdot \frac{1}{|(M_1 \setminus M_{12}) \setminus L|} + \\
&\qquad\qquad |(M'_1 \setminus M'_{12}) \setminus L| \cdot \frac{2}{3} \cdot \frac{1}{|(M'_1 \setminus M'_{12}) \setminus L|} \\
&\qquad = 1
\end{aligned}$$

and as in the previous case it holds that $P_1(\psi \,|\, \phi) = d$ and $P_1(\psi' \,|\, \phi') = 1$. Furthermore, with a similar argumentation P_2 is a probability function and it holds that $P_2(\psi \,|\, \phi) = d$ and $P_2(\psi' \,|\, \phi') < 1$. It follows that both P_1 and P_2 satisfy $(\psi \,|\, \phi)[d]$ but at least one of P_1 and P_2 cannot satisfy $(\psi' \,|\, \phi')[d']$. Therefore $(\psi \,|\, \phi)[d] \not\equiv^{\text{pr}} (\psi' \,|\, \phi')[d']$ contradicting the premise.

For the remaining cases with $M_1 \cap M_1' \neq \emptyset$ similar constructions for P_1 and P_2 can be given which all result in a contradiction to the premises. This means that either 1.), 2.), 3.), or 4.) must be true. $\square$

Lemma 3.1 (page 49). *Let $\mathcal{R}$ be a knowledge base and let r be a probabilistic conditional with $r \notin \mathcal{R}$. If $\mathsf{At}(\{r\}) \cap \mathsf{At}(\mathcal{R}) = \emptyset$ then r is a free conditional in $\mathcal{R} \cup \{r\}$.*

Proof. Assume that r is not a free conditional in $\mathcal{R} \cup \{r\}$. Then there is a set $\mathcal{M} \in \mathsf{MI}(\mathcal{R})$ with $r \in \mathcal{M}$. As $\mathcal{M} \setminus \{r\}$ is consistent and $\mathsf{At}(\mathcal{M} \setminus \{r\}) \cap \mathsf{At}(\{r\}) = \emptyset$ let P_1 be a probability function on $\mathcal{L}(\mathsf{At} \setminus \mathsf{At}(\{r\}))$ with $P_1 \models^{pr} \mathcal{M} \setminus \{r\}$. As r is self-consistent let P_2 be a probability function on $\mathcal{L}(\mathsf{At}(\{r\}))$ with $P_2 \models^{pr} r$. Let $\omega \in \Omega(\mathsf{At})$ and define $\omega_{\mathcal{A}}$ with $\mathcal{A} \subseteq \mathsf{At}$ to be the projection of ω on $\mathcal{A}$, i. e. $\omega_{\mathcal{A}} =_{def} \bigwedge\{l \in \mathsf{Lit}(\mathcal{A}) \mid \omega \models^{\mathrm{P}} l\}$. Define a probability function P on $\mathcal{L}(\mathsf{At})$ via

$$P(\omega) =_{def} P_1\left(\omega_{\mathsf{At}\setminus\mathsf{At}(\{r\})}\right) \cdot P_2\left(\omega_{\mathsf{At}(\{r\})}\right) \quad .$$

Note that $f : \Omega(\mathsf{At}) \to \Omega(\mathsf{At} \setminus \mathsf{At}(\{r\})) \times \Omega(\mathsf{At}(\{r\}))$ with

$$f(\omega) =_{def} \left(\omega_{\mathsf{At}\setminus\mathsf{At}(\{r\})}, \omega_{\mathsf{At}(\{r\})}\right)$$

is a bijection. It follows that P is indeed a probability function as

$$\begin{aligned}
\sum_{\omega\in\Omega(\mathsf{At})} P(\omega) &= \sum_{\omega\in\Omega(\mathsf{At})} P_1\left(\omega_{\mathsf{At}\setminus\mathsf{At}(\{r\})}\right) \cdot P_2\left(\omega_{\mathsf{At}(\{r\})}\right) \\
&= \sum_{(\omega_1,\omega_2)\in\Omega(\mathsf{At}\setminus\mathsf{At}(\{r\}))\times\Omega(\mathsf{At}(\{r\}))} P_1(\omega_1)P_2(\omega_2) \\
&= \sum_{\omega_1\in\Omega(\mathsf{At}\setminus\mathsf{At}(\{r\}))} \sum_{\omega_2\in\Omega(\mathsf{At}(\{r\}))} P_1(\omega_1)P_2(\omega_2) \\
&= \sum_{\omega_1\in\Omega(\mathsf{At}\setminus\mathsf{At}(\{r\}))} \left(P_1(\omega_1) \cdot \sum_{\omega_2\in\Omega(\mathsf{At}(\{r\}))} P_2(\omega_2) \right) \\
&= 1
\end{aligned}$$

Furthermore, for $\omega \in \Omega(\mathsf{At} \setminus \mathsf{At}(\{r\}))$ it holds that

$$\begin{aligned}
P(\omega) &= \sum_{\omega'\in\Omega(\mathsf{At}(\{r\}))} P(\omega \wedge \omega') \\
&= \sum_{\omega'\in\Omega(\mathsf{At}(\{r\}))} P_1(\omega)P_2(\omega') \\
&= P_1(\omega) \sum_{\omega'\in\Omega(\mathsf{At}(\{r\}))} P_2(\omega') \\
&= P_1(\omega)
\end{aligned}$$

and similarly $P(\omega') = P_2(\omega')$. It follows that $P \models^{pr} \mathcal{M} \setminus \{r\}$ and $P \models^{pr} r$ contradicting the assumption that $\mathcal{M}$ is a minimal inconsistent subset. □

Proposition 3.9 (page 53)**.** *The function* Inc^d *satisfies (Consistency), (Irrelevance of Syntax), (Monotonicity), (Weak Independence), (Independence), and (Normalization).*

Proof. We only show that Inc^d satisfies (Consistency), (Irrelevance of Syntax), (Monotonicity), (Independence), and (Normalization) as (Weak Independence) follows from (Independence) due to Proposition 3.6 on page 49.

(Consistency) A knowledge base $\mathcal{R}$ is consistent if and only if $\mathsf{Inc}^d(\mathcal{R}) = 0$ by definition.

(Irrelevance of Syntax) From $\mathcal{R}_1 \equiv^{\text{cond}} \mathcal{R}_2$ follows $\mathcal{R}_1 \equiv^{\text{kb}} \mathcal{R}_2$ by Proposition 2.5 on page 30. Therefore, $\mathcal{R}_1$ is inconsistent if and only if $\mathcal{R}_2$ is inconsistent. It follows $\mathsf{Inc}^d(\mathcal{R}_1) = \mathsf{Inc}^d(\mathcal{R}_2)$.

(Monotonicity) If $\mathcal{R}$ is inconsistent so is any superset of $\mathcal{R}$. It follows $\mathsf{Inc}^d(\mathcal{R}) = 1 = \mathsf{Inc}^d(\mathcal{R} \cup \{r\})$. If $\mathcal{R}$ is consistent then $\mathsf{Inc}^d(\mathcal{R} \cup \{r\}) \geq 0$ by definition.

(Independence) If $\mathcal{R}$ is consistent and r is a free conditional in $\mathcal{R} \cup \{r\}$ then $\mathcal{R} \cup \{r\}$ is consistent due to Proposition 3.3 on page 48 and $\mathsf{Inc}^d(\mathcal{R} \cup \{r\}) = 0 = \mathsf{Inc}^d(\mathcal{R})$. If $\mathcal{R}$ is inconsistent so is any superset of $\mathcal{R}$ and hence $\mathsf{Inc}^d(\mathcal{R} \cup \{r\}) = 1 = \mathsf{Inc}^d(\mathcal{R})$.

(Normalization) For every $\mathcal{R}$ it holds that either $\mathsf{Inc}^d(\mathcal{R}) = 0$ or $\mathsf{Inc}^d(\mathcal{R}) = 1$ and therefore $\mathsf{Inc}^d(\mathcal{R}) \in [0, 1]$. □

Proposition 3.11 (page 54)**.** *The function* $\mathsf{Inc}^{\mathsf{MI}}$ *satisfies (Consistency), (Monotonicity), (Super-Additivity), (Weak Independence), (Independence), (*MinInc *separability), and (Penalty).*

Proof. We only show that $\mathsf{Inc}^{\mathsf{MI}}$ satisfies (Consistency), (Super-Additivity), (MinInc Separability), and (Penalty), as (Monotonicity) follows from (Super-Additivity) due to Proposition 3.2 on page 47, (Weak Independence) follows from (Independence) due to Proposition 3.6 on page 49, and (Independence) follows from (MinInc Separability) due to Proposition 3.7 on page 50.

(Consistency) If $\mathcal{R}$ is consistent it follows that $\mathsf{MI}(\mathcal{R}) = \emptyset$ and therefore $\mathsf{Inc}^{\mathsf{MI}}(\mathcal{R}) = 0$. If $\mathcal{R}$ is inconsistent then $\mathsf{MI}(\mathcal{R}) \neq \emptyset$ and $\mathsf{Inc}^{\mathsf{MI}}(\mathcal{R}) > 0$.

(Super-Additivity) Let $\mathcal{R} \cap \mathcal{R}' = \emptyset$. Due to Proposition 3.5 on page 49 it holds that $\mathsf{MI}(\mathcal{R}) \subseteq \mathsf{MI}(\mathcal{R} \cup \mathcal{R}')$ and $\mathsf{MI}(\mathcal{R}') \subseteq \mathsf{MI}(\mathcal{R} \cup \mathcal{R}')$. Due to $\mathcal{R} \cap \mathcal{R}' = \emptyset$ it follows that $\mathsf{MI}(\mathcal{R}) \cap \mathsf{MI}(\mathcal{R}') = \emptyset$ and therefore $\mathsf{Inc}^{\mathsf{MI}}(\mathcal{R} \cup \mathcal{R}') = |\mathsf{MI}(\mathcal{R} \cup \mathcal{R}')| \geq |\mathsf{MI}(\mathcal{R}) \cup \mathsf{MI}(\mathcal{R}')| = |\mathsf{MI}(\mathcal{R})| + |\mathsf{MI}(\mathcal{R}')| = \mathsf{Inc}^{\mathsf{MI}}(\mathcal{R}') + \mathsf{Inc}^{\mathsf{MI}}(\mathcal{R}')$.

(MinInc Separability) Let $\mathcal{R}_1, \mathcal{R}_2$ be knowledge bases with $\mathsf{MI}(\mathcal{R}_1 \cup \mathcal{R}_2) = \mathsf{MI}(\mathcal{R}_1) \cup \mathsf{MI}(\mathcal{R}_2)$ and $\mathsf{MI}(\mathcal{R}_1) \cap \mathsf{MI}(\mathcal{R}_2) = \emptyset$. It follows directly that $\mathsf{Inc}^{\mathsf{MI}}(\mathcal{R}_1 \cup \mathcal{R}_2) = |\mathsf{MI}(\mathcal{R}_1 \cup \mathcal{R}_2)| = |\mathsf{MI}(\mathcal{R}_1)| + |\mathsf{MI}(\mathcal{R}_2)| = \mathsf{Inc}^{\mathsf{MI}}(\mathcal{R}_1) + \mathsf{Inc}^{\mathsf{MI}}(\mathcal{R}_2)$.

(Penalty) Let $r \notin \mathcal{R}$ be a conditional that is not free in $\mathcal{R} \cup \{r\}$. By the facts that $\mathsf{MI}(\mathcal{R}) \subseteq \mathsf{MI}(\mathcal{R} \cup \{r\})$ and that there is a $\mathcal{M} \in \mathsf{MI}(\mathcal{R} \cup \{r\})$ with $r \in \mathcal{M}$ it follows that $|\mathsf{MI}(\mathcal{R})| < |\mathsf{MI}(\mathcal{R} \cup \{r\})|$ and therefore $\mathsf{Inc}^{\mathsf{MI}}(\mathcal{R}) < \mathsf{Inc}^{\mathsf{MI}}(\mathcal{R} \cup \{r\})$. □

Proposition 3.12 (page 55). *The function* $\mathsf{Inc}^{\mathsf{MI}}_0$ *satisfies (Consistency) and (Normalization).*

Proof.

(Consistency) It holds that $\mathsf{Inc}^{\mathsf{MI}}_0(\mathcal{R}) = 0$ for consistent $\mathcal{R}$ by Proposition 3.11 on page 54. It holds that $\mathsf{Inc}^{\mathsf{MI}}_0(\mathcal{R}) > 0$ for inconsistent $\mathcal{R}$ by Proposition 3.11 on page 54 and the fact that $\gamma_{\mathcal{R}} \neq 0$ because $\mathcal{R}$ is inconsistent and therefore non-empty.

(Normalization) For empty $\mathcal{R}$ it follows $\mathsf{Inc}^{\mathsf{MI}}_0(\mathcal{R}) = 0$. By Proposition 3.1 on page 55 it follows for non-empty $\mathcal{R}$ that

$$\mathsf{Inc}^{\mathsf{MI}}_0(\mathcal{R}) = \frac{\mathsf{Inc}^{\mathsf{MI}}(\mathcal{R})}{\gamma_{\mathcal{R}}} \leq \frac{\gamma_{\mathcal{R}}}{\gamma_{\mathcal{R}}} = 1 \quad .$$ □

Proposition 3.13 (page 57). *The function* $\mathsf{Inc}^{\mathsf{MI}}_C$ *satisfies (Consistency), (Monotonicity), (Super-Additivity), (Weak Independence), (Independence), (*MinInc *Separability), and (Penalty).*

Proof. We only show that $\mathsf{Inc}^{\mathsf{MI}}_C$ satisfies (Consistency), (Super-Additivity), (MinInc Separability), and (Penalty) as (Monotonicity) follows from (Super-Additivity) due to Proposition 3.2 on page 47, (Weak Independence) follows from (Independence) due to Proposition 3.6 on page 49, and (Independence) follows from (MinInc Separability) due to Proposition 3.7 on page 50.

(Consistency) If $\mathcal{R}$ is consistent it follows that $\mathsf{MI}(\mathcal{R}) = \emptyset$ and therefore $\mathsf{Inc}^{\mathsf{MI}}_C(\mathcal{R}) = 0$ (the empty sum). If $\mathcal{R}$ is inconsistent then $\mathsf{MI}(\mathcal{R}) \neq \emptyset$ with $\mathcal{M} \in \mathsf{MI}(\mathcal{R})$ and $|\mathcal{M}| > 0$. It follows that $\mathsf{Inc}^{\mathsf{MI}}_C(\mathcal{R}) > 0$.

(Super-Additivity) Let $\mathcal{R} \cap \mathcal{R}' = \emptyset$. Due to Proposition 3.5 on page 49 it holds that $\mathsf{MI}(\mathcal{R}) \subseteq \mathsf{MI}(\mathcal{R} \cup \mathcal{R}')$ and $\mathsf{MI}(\mathcal{R}') \subseteq \mathsf{MI}(\mathcal{R} \cup \mathcal{R}')$. Due to $\mathcal{R} \cap \mathcal{R}' = \emptyset$ it follows that $\mathsf{MI}(\mathcal{R}) \cap \mathsf{MI}(\mathcal{R}') = \emptyset$ and therefore

$$\mathsf{Inc}^{\mathsf{MI}}_C(\mathcal{R} \cup \mathcal{R}') = \sum_{\mathcal{M} \in \mathsf{MI}(\mathcal{R} \cup \mathcal{R}')} \frac{1}{|\mathcal{M}|}$$

$$\geq \sum_{\mathcal{M}\in\mathsf{MI}(\mathcal{R})} \frac{1}{|\mathcal{M}|} + \sum_{\mathcal{M}\in\mathsf{MI}(\mathcal{R}')} \frac{1}{|\mathcal{M}|}$$
$$= \mathsf{Inc}^{\mathsf{MI}}_{C}(\mathcal{R}) + \mathsf{Inc}^{\mathsf{MI}}_{C}(\mathcal{R}') \quad .$$

(MinInc Separability) Let $\mathcal{R}_1, \mathcal{R}_2$ be knowledge bases with $\mathsf{MI}(\mathcal{R}_1 \cup \mathcal{R}_2) = \mathsf{MI}(\mathcal{R}_1) \cup \mathsf{MI}(\mathcal{R}_2)$ and $\mathsf{MI}(\mathcal{R}_1) \cap \mathsf{MI}(\mathcal{R}_2) = \emptyset$. It follows directly that

$$\begin{aligned} \mathsf{Inc}^{\mathsf{MI}}_{C}(\mathcal{R}_1 \cup \mathcal{R}_2) &= \sum_{\mathcal{M}\in\mathsf{MI}(\mathcal{R}_1\cup\mathcal{R}_2)} \frac{1}{|\mathcal{M}|} \\ &= \sum_{\mathcal{M}\in\mathsf{MI}(\mathcal{R}_1)} \frac{1}{|\mathcal{M}|} + \sum_{\mathcal{M}\in\mathsf{MI}(\mathcal{R}_2)} \frac{1}{|\mathcal{M}|} \\ &= \mathsf{Inc}^{\mathsf{MI}}_{C}(\mathcal{R}_1) + \mathsf{Inc}^{\mathsf{MI}}_{C}(\mathcal{R}_2) \quad . \end{aligned}$$

(Penalty) Let $r \notin \mathcal{R}$ be a conditional that is not free in $\mathcal{R} \cup \{r\}$. By the facts that $\mathsf{MI}(\mathcal{R}) \subseteq \mathsf{MI}(\mathcal{R} \cup \{r\})$ and that there is a $\mathcal{M} \in \mathsf{MI}(\mathcal{R} \cup \{r\})$ with $r \in \mathcal{M}$ it follows that $|\mathsf{MI}(\mathcal{R})| < |\mathsf{MI}(\mathcal{R} \cup \{r\})|$ and therefore $\mathsf{Inc}^{\mathsf{MI}}_{C}(\mathcal{R}) < \mathsf{Inc}^{\mathsf{MI}}_{C}(\mathcal{R} \cup \{r\})$. □

Proposition 3.15 (page 58)**.** *The function* $\mathsf{Inc}^{\mathsf{MI}}_{C,0}$ *satisfies (Consistency) and (Normalization).*

Proof.

(Consistency) It holds that $\mathsf{Inc}^{\mathsf{MI}}_{C,0}(\mathcal{R}) = 0$ for consistent $\mathcal{R}$ by Proposition 3.13 on page 57. It holds that $\mathsf{Inc}^{\mathsf{MI}}_{C,0}(\mathcal{R}) > 0$ for inconsistent $\mathcal{R}$ by Proposition 3.13 on page 57 and the fact that $\gamma_{\mathcal{R}} \neq 0$ as $\mathcal{R}$ is inconsistent and therefore non-empty.

(Normalization) For empty $\mathcal{R}$ it follows $\mathsf{Inc}^{\mathsf{MI}}_{C,0}(\mathcal{R}) = 0$. By Proposition 3.14 on page 58 it follows for non-empty $\mathcal{R}$ that

$$\mathsf{Inc}^{\mathsf{MI}}_{C,0}(\mathcal{R}) = 2\frac{\mathsf{Inc}^{\mathsf{MI}}_{C,0}(\mathcal{R})}{\gamma_{\mathcal{R}}} \leq 2\frac{\gamma_{\mathcal{R}}}{2\gamma_{\mathcal{R}}} = 1 \quad .$$ □

Theorem 3.1 (page 65)**.** Inc^* *satisfies (Consistency), (Monotonicity), (Super-Additivity), (Weak Independence), (Independence), and (Continuity).*

Proof. We only show that Inc^* satisfies (Consistency), (Super-Additivity), (Independence), and (Continuity) as (Monotonicity) follows from (Super-Additivity) due to Proposition 3.2 on page 47 and (Weak Independence) follows from (Independence) due to Proposition 3.6 on page 49.

(Consistency) If $\mathcal{R}$ with $\langle\mathcal{R}\rangle = ((\psi_1 \,|\, \phi_1)[d_1], \ldots, (\psi_n \,|\, \phi_n)[d_n])$ is consistent so is $\Lambda_{\mathcal{R}}(d_1, \ldots, d_n)$ and therefore $\mathsf{Inc}^*(\mathcal{R}) = 0$.

(Super-Additivity) Let $\mathcal{R}_1$ with $\langle\mathcal{R}_1\rangle = (r_1,\ldots,r_n)$ and $\mathcal{R}_2$ with $\langle\mathcal{R}_2\rangle = (r_{n+1},\ldots,r_m)$ be two knowledge bases with $r_i =_{def} (\psi_i \mid \phi_i)[d_i]$ for $i = 1,\ldots,m$, $\mathcal{R}_1 \cap \mathcal{R}_2 = \emptyset$ and let $\mathcal{R} =_{def} \mathcal{R}_1 \cup \mathcal{R}_2$ with $\langle\mathcal{R}\rangle = (r_1,\ldots,r_m)$. Let $x_1,\ldots,x_m \in [0,1]$ such that $\Lambda_{\mathcal{R}}(x_1,\ldots, x_m)$ is consistent and $\mathsf{Inc}^*(\mathcal{R}) = |d_1 - x_1| + \ldots + |d_m - x_m|$. As the knowledge base $\Lambda_{\mathcal{R}}(x_1,\ldots,x_m)$ is consistent so is any subset of $\Lambda_{\mathcal{R}}(x_1,\ldots, x_m)$, in particular $\Lambda_{\mathcal{R}_1}(x_1, \ldots,x_n)$ and $\Lambda_{\mathcal{R}_2}(x_{n+1},\ldots,x_m)$ are consistent. It follows that $\mathsf{Inc}^*(\mathcal{R}_1) \leq |d_1 - x_1| + \ldots + |d_n - x_n|$ and $\mathsf{Inc}^*(\mathcal{R}_2) \leq |d_{n+1} - x_{n+1}| + \ldots + |d_m - x_m|$ and therefore $\mathsf{Inc}^*(\mathcal{R}) \geq \mathsf{Inc}^*(\mathcal{R}_1) + \mathsf{Inc}^*(\mathcal{R}_2)$.

(Independence) Before proving (Independence) we first show that from both $\mathcal{R} \cup \{(\psi \mid \phi)[d_1]\}$ and $\mathcal{R} \cup \{(\psi \mid \phi)[d_2]\}$ being consistent for some knowledge base $\mathcal{R}$ and $d_1 \leq d_2$ it follows that $\mathcal{R} \cup \{(\psi \mid \phi)[y]\}$ is consistent for every $y \in [0,1]$ that satisfies $d_1 \leq y \leq d_2$. So let $\mathcal{R}$ be some consistent knowledge base and $(\psi \mid \phi)[d_1]$ be a probabilistic conditional such that both $\mathcal{R} \cup \{(\psi \mid \phi)[d_1]\}$ and $\mathcal{R} \cup \{(\psi \mid \phi)[d_2]\}$ are consistent with $d_1 \leq d_2$. Let $P_1 \models_\circ \mathcal{R} \cup \{(\psi \mid \phi)[d_1]\}$ and let $P_2 \models_\circ \mathcal{R} \cup \{(\psi \mid \phi)[d_2]\}$. If $P_1(\phi) = 0$ then clearly $P_1 \models_\circ \mathcal{R} \cup \{(\psi \mid \phi)[y]\}$ for every $y \in [0,1]$ due to our definition of probabilistic satisfaction. If $P_2(\phi) = 0$ then $P_2 \models_\circ \mathcal{R} \cup \{(\psi \mid \phi)[y]\}$ for every $y \in [0,1]$ accordingly. So assume $P_1(\phi) > 0$ and $P_2(\phi) > 0$. Let $\delta \in [0,1]$ and consider the probability function P_δ defined via $P_\delta(\omega) = \delta P_1(\omega) + (1-\delta)P_2(\omega)$ for all $\omega \in \Omega(\mathsf{At})$. Then $P_\delta \models_\circ \mathcal{R}$ for all $\delta \in [0,1]$ as the set of models of a knowledge base is a convex set, cf. (Paris, 1994). Furthermore, note that $P_\delta(\phi) > 0$ for every $\delta \in [0,1]$ as both $P_1(\phi) > 0$ and $P_2(\phi) > 0$. Then $P_\delta(\psi \mid \phi)$ is continuous in δ and for every $y \in [d_1,d_2]$ there is a $\delta_y \in [0,1]$ such that $P_{\delta_y}(\psi \mid \phi) = y$. It follows that $P_{\delta_y} \models_\circ \mathcal{R} \cup \{(\psi \mid \phi)[y]\}$ for every $y \in [d_1,d_2]$ and therefore $\mathcal{R} \cup \{(\psi \mid \phi)[y]\}$ is consistent for every $y \in [d_1,d_2]$.

Let now $\mathcal{R}$ with $\langle\mathcal{R}\rangle = (r_1,\ldots,r_n)$ and $r_i = (\psi_i \mid \phi_i)[d_i]$ for $i = 1,\ldots,n$ be a knowledge base and let $r = (\psi \mid \phi)[d]$ be free in $\mathcal{R} \cup \{r\}$. Assume that $\mathcal{R}$ is also a minimal inconsistent set, i. e. $\mathsf{MI}(\mathcal{R}) = \{\mathcal{R}\}$. Let $\mathsf{Inc}^*(\mathcal{R}) = x$ and let $(x_1,\ldots,x_n) \in [0,1]^n$ be such that $\Lambda_{\mathcal{R}}(x_1,\ldots,x_n)$ is consistent and $|d_1 - x_1| + \ldots + |d_n - x_n| = x$. Consider now $\mathcal{R}'$ with $\langle\mathcal{R}\rangle = ((\psi_1 \mid \phi_1)[d_1],\ldots,(\psi_n \mid \phi_n)[d_n],(\psi \mid \phi)[d])$. It suffices to show that $\Lambda_{\mathcal{R}'}(x_1,\ldots,x_n,d)$ is consistent. Define $C_j = \mathcal{R} \setminus \{(\psi_j \mid \phi_j)[d_j]\}$ for every $j = 1,\ldots,n$. Then both C_j and $C_j \cup \{r\}$ are consistent. Let d_j be such that there is a P with $P \models_\circ C_j \cup \{r\}$, $P \models_\circ (\psi_j \mid \phi_j)[d'_j]$ and $|d_j - d'_j|$ is minimal. It follows that $|d_j - d'_j| \geq x$ (otherwise this would contradict $\mathsf{Inc}^*(\mathcal{R}) = x$). Assume w.l.o.g. $d'_j > d_j$. As $\{r, r_j\}$ is consistent as well (as r is free) it follows that $\{r, (\psi_j \mid \phi_j)[y]$ is consistent for every $y \in [d_j, d'_j]$ due to our elaboration above. As $|d_j - x_j| \leq x$ it follows $x_j \in [d_j, d'_j]$ as well (or $x_j \in [d'_j, d_j]$ if $d_j > d'_j$). Hence, $\{r, (\psi_j \mid \phi_j)[x_j]\}$ is consistent for every $j = 1,\ldots,n$. As $\Lambda_{\mathcal{R}}(x_1,\ldots,x_n)$

is consistent and r is consistent with every combination of conditionals in $\Lambda_{\mathcal{R}}(x_1,\ldots,x_n)$ it follows that $\Lambda_{\mathcal{R}'}(x_1,\ldots,x_n,d)$ is consistent. The above can be generalized if $\mathcal{R}$ contains multiple minimal inconsistent subsets by iteratively considering each minimal inconsistent subset of $\mathcal{R}$.

(Continuity) Let $r = (\psi\,|\,\phi)[d]$ be a self-consistent and non-tautological probabilistic conditional and consider $\pi_r : [0,1] \times \mathbb{R}^{|\Omega(\mathsf{At})|} \to [0,1]$ that maps a value $x \in [0,1]$ and a point $p \in \mathbb{R}^{|\Omega(\mathsf{At})|}$ to the Euclidean distance, i. e. the 2-norm distance, from p to $H_{(\psi\,|\,\phi)[x]}$. Hence, π_r is defined via $\pi_r(x,p) = d(p, H_{(\psi\,|\,\phi)[x]})$ if $d(\cdot,\cdot)$ is the 2-norm distance. We show now that π_r is continuous with respect to the standard topology. Let $\{\omega_1,\ldots,\omega_m\} = \Omega(\mathsf{At})$ be the set of possible worlds. Then $H_{(\psi\,|\,\phi)[x]}$ is the set of points $(\alpha_1,\ldots,\alpha_m)$ that satisfy

$$\sum_{i=1,\ \omega_i\in\mathsf{Mod}^{\mathrm{P}}(\psi\phi)}^{m} (1-x)\alpha_i - \sum_{i=1,\ \omega_i\in\mathsf{Mod}^{\mathrm{P}}(\overline{\psi_i}\phi_i)}^{m} x\alpha_i = 0 \quad , \tag{3.1}$$

see also Equation (3.2) on page 45. By reordering the elements of $\Omega(\mathsf{At})$ we can write Equation (3.1) as

$$(1-x)\alpha_1 + \ldots + (1-x)\alpha_{k_1} - x\alpha_{k_1+1} - \ldots - x\alpha_{k_2} = 0$$

for some $k_1, k_2 \in \mathbb{N}$ with $k_1 \leq k_2 \leq m$. Then the Euclidean distance $d(p, H_{(\psi\,|\,\phi)[x]})$ from a point $p = (p_1,\ldots,p_m)$ to the hyperplane $H_{(\psi\,|\,\phi)[x]}$ can be computed via

$$d(p, H_{(\psi\,|\,\phi)[x]}) = \frac{(1-x)p_1 + \ldots + (1-x)p_{k_1} - xp_{k_1+1} - \ldots - xp_{k_2}}{\sqrt{k_1(1-x)^2 + (k_2-k_1)x^2}} \tag{3.2}$$

where $\sqrt{k_1(1-x)^2 + (k_2-k_1)x^2}$ is the length of the normal vector

$$(\underbrace{(1-x),\ldots,(1-x)}_{k_1\text{ times}}, \underbrace{x,\ldots,x}_{k_2-k_1\text{ times}}, \underbrace{0,\ldots,0}_{m-k_2-k_1\text{ times}})$$

of $H_{(\psi\,|\,\phi)[x]}$. Note that the denominator of (3.2) is always greater than zero as r is both self-consistent and non-tautological which amounts to $k_1 > 0$ and $k_2 - k_1 > 0$. As (3.2) is continuous with respect to $p_1,\ldots,p_m$ and x it follows that $\pi_r(x,p)$ is continuous in both x and p. Note also that the minimum function is continuous if its arguments are continuous. Now consider $\theta_{\mathsf{Inc}^*,\mathcal{R}}$ which can be characterized as

$$\theta_{\mathsf{Inc}^*,\mathcal{R}}(x_1,\ldots,x_n) = \min\{\pi_r(x_1,p) + \ldots + \pi_r(x_n,p) \mid p \in \mathbb{R}^{|\Omega(\mathsf{At})|} \cap H_0\} \quad .$$

It follows that $\theta_{\mathsf{Inc}^*,\mathcal{R}}$ is continuous. $\square$

Proposition 3.21 (page 66). *The function* Inc_0^* *satisfies (Consistency), (Continuity), and (Normalization).*

Proof.

(Consistency) If $\mathcal{R} = \emptyset$ then $\mathcal{R}$ is trivially consistent and $\mathsf{Inc}_0^*(\mathcal{R}) = 0$. Furthermore, it holds that $\mathcal{R} \neq \emptyset$ is consistent if and only if $\mathsf{Inc}_0^*(\mathcal{R}) = 0$ as $\mathsf{Inc}^*(\mathcal{R})$ satisfies (Consistency).

(Continuity) For $\mathcal{R} \neq \emptyset$ the function $\theta_{\mathsf{Inc}_0^*,\mathcal{R}}$ is continuous as $\theta_{\mathsf{Inc}^*,\mathcal{R}}$ is continuous. For $\mathcal{R} = \emptyset$ the constant function 0 is trivially continuous.

(Normalization) This follows directly from Proposition 3.20 on page 66. $\square$

Theorem 3.2 (page 81). *The function* Inc_{gd} *satisfies (Consistency), (Monotonicity), (Super-Additivity), (Weak Independence), (Independence), and (Continuity).*

Proof. We only give the proofs for (Consistency) and (Super-Additivity) to show the similarity to the corresponding proofs of Inc^*. The remaining proofs are also similar to the ones of Inc^*.

(Consistency) Let $\mathcal{R}$ with $\langle \mathcal{R} \rangle = (r_1, \ldots, r_n)$ and $r_i =_{def} (\psi_i \,|\, \phi_i)[d_i]$ for $i = 1, \ldots, n$. For consistent $\mathcal{R}$ there is a probability function P with $P \models^{pr} \mathcal{R}$. It follows that for each $i = 1, \ldots, n$ it holds that $(1 - d_i)P(\psi_i\phi_i) = d_i P(\overline{\psi_i}\phi_i)$. Therefore setting, $y_i = z_i = (1 - d_i)P(\psi_i\phi_i)$ fulfills the constraints $y_i = (1 - d_i)P(\psi_i\phi_i)$ and $z_i = d_i P(\overline{\psi_i}\phi_i)$ for $i = 1, \ldots, n$, and it also holds that $D^2(\vec{y}, \vec{z}) = 0$. This amounts to $\mathsf{Inc}_{gd}(\mathcal{R}) = 0$. If $\mathcal{R}$ is inconsistent then for all $P \in \mathcal{P}^{\mathrm{P}}(\mathsf{At})$ there is a $j \in \{1, \ldots, n\}$ such that $(1 - d_j)P(\psi_i\phi_j) \neq d_j P(\overline{\psi_j}\phi_j)$ which amounts to $y_j \neq z_j$. As $D^2(\vec{y}, \vec{z}) > 0$ if $y_j \neq z_j$ for some j, it follows $\mathsf{Inc}_{gd}(\mathcal{R}) > 0$.

(Super-Additivity) Let $\mathcal{R}_1$ with $\langle \mathcal{R}_1 \rangle = (r_1, \ldots, r_n)$ and $\mathcal{R}_2$ with $\langle \mathcal{R}_2 \rangle = (r_{n+1}, \ldots, r_m)$ be two knowledge bases with $r_i = (\psi_i \,|\, \phi_i)[d_i]$ for $i = 1, \ldots, m$, $\mathcal{R}_1 \cap \mathcal{R}_2 = \emptyset$ and let $\mathcal{R} =_{def} \mathcal{R}_1 \cup \mathcal{R}_2$ with $\langle \mathcal{R} \rangle = (r_1, \ldots, r_m)$. Let $\vec{y}, \vec{z} \in (0, 1]^m$ be such that

$$\mathsf{Inc}_{gd}(\mathcal{R}) = D^2(\vec{y}, \vec{z}) \tag{3.3}$$

and there is a P^* that such $y_i = (1 - d_i)P^*(\psi_i\phi_i)$ and $z_i = d_i P^*(\overline{\psi_i}\phi_i)$. Let $\vec{y}_{k,l}$ denote the restriction of the vector $\vec{y}$ to the coordinates k to l with $k \leq l$. Observe that it holds that

$$D^2(\vec{y}, \vec{z}) = D^2(\vec{y}_{1,n}, \vec{z}_{1,n}) + D^2(\vec{y}_{n+1,m}, \vec{y}_{n+1,m}) \tag{3.4}$$

and $D^2(\vec{y}_{1,n}, \vec{z}_{1,n})$ is a valid element of the set in (3.31) on page 81 with respect to P and $\mathcal{R}_1$, and same applies for $D^2(\vec{y}_{n+1,m}, \vec{y}_{n+1,m})$ with respect to P and $\mathcal{R}_2$. It follows

$$D^2(\vec{y}_{1,n}, \vec{z}_{1,n}) \geq \mathsf{Inc}_{gd}(\mathcal{R}_1)$$
$$D^2(\vec{y}_{n+1,m}, \vec{y}_{n+1,m}) \geq \mathsf{Inc}_{gd}(\mathcal{R}_2)$$

and together with (3.3) and (3.4) it follows $\mathsf{Inc}_{gd}(\mathcal{R}) \geq \mathsf{Inc}_{gd}(\mathcal{R}_1) + \mathsf{Inc}_{gd}(\mathcal{R}_2)$. □

Theorem 4.2 (page 111)**.** Υ^U *satisfies (Existence), (Uniqueness), (Structural Preservation), (Success), (Irrelevance of Syntax), (Consistency), (Rational Non-Imposition), and (Continuity).*

Proof.

(Existence) This is clear as $\mathsf{undef} \notin \mathsf{Dom}\ \Upsilon^U$ and $\Xi^U_{\mathcal{R}}$ is well-defined as $\mathsf{ucp}(\psi \,|\, \phi)$ is well-defined.

(Uniqueness) This is clear due to the functional definition of Υ^U and Remark 4.1 on page 109.

(Structural Preservation) This follows directly from Definition 4.8 on page 109.

(Success) This follows from the fact that $\Xi^U_{\mathcal{R}}(1)$ is guaranteed to be consistent, see Proposition 3.17 on page 60.

(Irrelevance of Syntax) Let $\mathcal{R}_1, \mathcal{R}_2$ be knowledge bases with $\mathcal{R}_1 \equiv^{\text{cond}} \mathcal{R}_2$. First, we show that

$$\Xi^U_{\mathcal{R}_1}(\delta) \equiv^{\text{cond}} \Xi^U_{\mathcal{R}_2}(\delta)$$

for every $\delta \in [0,1]$. Let $r_1 \in \mathcal{R}_1$ with $r_1 = (\psi \,|\, \phi)[d]$ and $r_2 \in \mathcal{R}_2$ with $r_2 = (\psi' \,|\, \phi')[d']$ such that $r_1 \equiv^{\text{pr}} r_2$. Let $u_1 =_{def} \mathsf{ucp}(\psi \,|\, \phi)$ and $u_2 =_{def} \mathsf{ucp}(\psi' \,|\, \phi')$. Consider the following case differentiation.

1. If r_1 is tautological so is r_2 by Proposition 2.3 on page 29). It follows $u_1 = d$ and $u_2 = d'$ and therefore $\Xi^U_{\{r_1\}}(\delta) \equiv^{\text{cond}} \Xi^U_{\{r_2\}}(\delta)$ for every $\delta \in [0,1]$ as both r_1 and r_2 are not modified during creeping.
2. Assume that r_1 is not tautological. As r_1 is also self-consistent it follows that r_2 has the form 1.) or 2.) in Proposition 2.3 on page 29. Without loss of generality assume it holds that $\phi \equiv^{\text{P}} \phi'$ and $\psi \wedge \phi \equiv^{\text{P}} \psi' \wedge \phi'$ and $d = d'$. It also follows $u_1 = u_2$ as $(\psi \,|\, \phi)$ has the same probability in P_0 as the equivalent $(\psi' \,|\, \phi')$. It follows that $\Xi^U_{\{r_1\}}(\delta) \equiv^{\text{cond}} \Xi^U_{\{r_2\}}(\delta)$ for every $\delta \in [0,1]$ as $r'_1 \in \Xi^U_{\{r_1\}}(\delta)$ has the form $r'_1 = (\psi \,|\, \phi)[d'']$ and $r'_2 \in \Xi^U_{\{r_2\}}(\delta)$ has

the form $r_1' = (\psi' \,|\, \phi')[d''']$ with $d'' = d'''$ and $r_1' \equiv^{\mathrm{pr}} r_2'$ (note that due to $d = d'$ and $u_1 = u_2$ the creeping keeps the identical probabilities intact)[1].

By applying the above observation to all probabilistic conditionals in $\mathcal{R}_1$ and $\mathcal{R}_2$ it follows that $\Xi^{\mathcal{U}}_{\mathcal{R}_1}(\delta) \equiv^{\mathrm{cond}} \Xi^{\mathcal{U}}_{\mathcal{R}_2}(\delta)$ for every $\delta \in [0,1]$.

Assume that $\Upsilon^{\mathcal{U}}(\mathcal{R}_1) \not\equiv^{\mathrm{cond}} \Upsilon^{\mathcal{U}}(\mathcal{R}_2)$. Let δ_1^* and δ_2^* be minimal such that $\Xi^{\mathcal{U}}_{\mathcal{R}_1}(\delta_1^*)$ and $\Xi^{\mathcal{U}}_{\mathcal{R}_2}(\delta_2^*)$ is consistent, respectively. It follows that $\delta_1^* \neq \delta_2^*$ and, without loss of generality, $\delta_1 < \delta_2$ and $\Xi^{\mathcal{U}}_{\mathcal{R}_2}(\delta_1^*)$ is inconsistent. But it also holds that $\Xi^{\mathcal{U}}_{\mathcal{R}_1}(\delta_1^*) \equiv^{\mathrm{cond}} \Xi^{\mathcal{U}}_{\mathcal{R}_2}(\delta_1^*)$ and by Proposition 2.5 on page 30 also $\Xi^{\mathcal{U}}_{\mathcal{R}_1}(\delta_1^*) \equiv^{\mathrm{kb}} \Xi^{\mathcal{U}}_{\mathcal{R}_2}(\delta_1^*)$. Therefore, every model of $\Xi^{\mathcal{U}}_{\mathcal{R}_1}(\delta_1^*)$ is also a model of $\Xi^{\mathcal{U}}_{\mathcal{R}_2}(\delta_1^*)$ and as $\Xi^{\mathcal{U}}_{\mathcal{R}_1}(\delta_1^*)$ is consistent so is $\Xi^{\mathcal{U}}_{\mathcal{R}_2}(\delta_1^*)$ contradicting $\Upsilon^{\mathcal{U}}(\mathcal{R}_1) \not\equiv^{\mathrm{cond}} \Upsilon^{\mathcal{U}}(\mathcal{R}_2)$.

(Consistency) Let $\mathcal{R}$ be consistent. Then $\Xi^{\mathcal{U}}_{\mathcal{R}}(0) = \mathcal{R}$ is consistent and $\Upsilon^{\mathcal{U}}(\mathcal{R}) = \mathcal{R}$.

(Rational Non-Imposition) This follows from Proposition 4.5 on page 105 and the fact that $\Upsilon^{\mathcal{U}}$ satisfies (Existence), (Success), and (Consistency).

(Continuity) Let $\mathcal{R}$ with $\langle\mathcal{R}\rangle = (r_1,\ldots,r_n)$ and $r_i = (\psi_i \,|\, \phi_i)[d_i]$ for $i = 1,\ldots,n$. We abbreviate

$$\begin{aligned}\Lambda^{\dagger}_{\mathcal{R}}(d_1,\ldots,d_n,\delta) =_{def} \Lambda_{\mathcal{R}}((1-\delta)d_1 + \delta\mathsf{ucp}(\psi_1 \,|\, \phi_1),\ldots,\\ (1-\delta)d_n + \delta\mathsf{ucp}(\psi_n \,|\, \phi_n)) \quad .\end{aligned}$$

As $(1-\delta)x + \delta t$ is continuous for fixed t and variables x and δ, so is $\mathsf{Inc}^* \circ \Lambda^{\dagger}_{\mathcal{R}}(d_1,\ldots,d_n,\delta)$ with respect to $d_1,\ldots,d_n,\delta$. As $\mathsf{Inc}^*(\mathcal{R}) = 0$ if and only if $\mathcal{R}$ is consistent it follows that

$$\delta(d_1,\ldots,d_n) =_{def} \min\{\delta \in [0,1] \mid \Lambda^{\dagger}_{\mathcal{R}}(d_1,\ldots,d_n,\delta) \text{ is consistent}\}$$

is continuous as well and as it holds that

$$\Upsilon^{\mathcal{U}}(\mathcal{R}) = \Lambda^{\dagger}_{\mathcal{R}}(d_1,\ldots,d_n,\delta(d_1,\ldots,d_n))$$

so is $\varsigma^{\Upsilon^{\mathcal{U}}}_{\mathcal{R}}$ continuous. □

Theorem 4.3 (page 113). *The function Υ^{P} satisfies (Uniqueness), (Structural Preservation), (Success), (Consistency), and (Non-Dictatorship).*

Proof.

(Uniqueness) This is clear due to the functional definition of Υ^{P}.

1 For the case that $\phi \equiv^{\mathrm{P}} \phi'$ and $\psi \wedge \phi \equiv^{\mathrm{P}} \overline{\psi'} \wedge \phi'$ and $d = 1 - d'$ it follows $u_1 = 1 - u_2$ and accordingly $d'' = 1 - d'''$

(Structural Preservation) This follows directly from the definition of Υ^P, cf. Definition 4.10 on page 112.

(Success) If $\Upsilon^P(\mathcal{R}) \neq \mathsf{undef}$ then $\{\delta \in [0,1] \mid \Xi^P_{\mathcal{R}}(\delta)$ is consistent$\} \neq \emptyset$ and $\Upsilon^P(\mathcal{R})$ is consistent.

(Consistency) Let $\mathcal{R}$ be consistent. Then $\Xi^P_{\mathcal{R}}(0) = \Lambda_{\mathcal{R}}(d_1, \ldots, d_n) = \mathcal{R}$ is consistent and $\Upsilon^P(\mathcal{R}) = \mathcal{R}$.

(Non-Dictatorship) Let $(\psi \mid \phi)[d]$ be a non-tautological probabilistic conditional and consider $\mathcal{R} =_{def} \{(\psi \mid \phi)[d], (\psi \mid \phi)[d']\}$ with some $d' \in [0,1]$ and $d \neq d'$. Then $\mathcal{R}$ is obviously inconsistent and $\alpha = ((d'-d)/2, (d-d')/2)$ and $\hat{\alpha} = (1,-1)$ or $\hat{\alpha} = (-1,1)$. It follows that $\Upsilon^P(\mathcal{R}) = \{(\psi \mid \phi)[(d+d')/2]\}$ and $(\psi \mid \phi)[d]$ is not dictatorial in $\mathcal{R}$. $\square$

Theorem 4.4 (page 114). *If Conjectures 4.1 and 4.2 are true then* Υ^P *satisfies (Existence), (Uniqueness), (Structural Preservation), (Success), (Consistency), (Rational Non-Imposition), (Continuity), (Non-Dictatorship), ($A^{\mathcal{R}}$-Conformity), and (Inverse $A^{\mathcal{R}}$-Conformity).*

Proof. Due to Theorem 4.3 on page 113 the function Υ^P satisfies (Uniqueness), (Structural Preservation), (Success),(Consistency), and (Non-Dictatorship). By Conjectures 4.1 on page 112 and 4.2 on page 114 the function Υ^P satisfies (Existence), ($A^{\mathcal{R}}$-conformity), and (Inverse $A^{\mathcal{R}}$-Conformity). Then Υ^P also satisfies (Rational Non-Imposition) by Proposition 4.5 on page 105. As $\mathsf{undef} \notin \mathsf{Im}\ \Upsilon^P_{\mathcal{R}}$ due to (Existence) it follows that Υ^P also satisfies (Continuity) with a similar argumentation as in the proof of Theorem 4.2 on page 111 and the fact that $A^{\mathcal{R}}$ satisfies (Continuity) due to Proposition 4.4 on page 99. $\square$

Theorem 4.5 (page 118). Υ^S_C *satisfies (Existence), (Uniqueness), (Structural Preservation), (Success), (Consistency), and (Rational Non-Imposition).*

Proof.

(Existence) This is clear as $\mathsf{undef} \notin \mathsf{Dom}\ \Upsilon^S_C$ and $\Xi^S_{\mathcal{R},C}$ is well-defined as $\mathsf{ucp}(\psi \mid \phi)$ is well-defined.

(Uniqueness) This is clear due to the functional definition of Υ^S_C.

(Structural Preservation) This follows directly from Definition 4.12 on page 118.

(Success) This follows from the fact that $\Xi^S_{\mathcal{R},C}(1)$ is guaranteed to be consistent, see Proposition 3.17 on page 60.

(Consistency) Let $\mathcal{R}$ be consistent. Then $\Xi^S_{\mathcal{R},C}(0) = \mathcal{R}$ is consistent and $\Upsilon^P_C(\mathcal{R}) = \mathcal{R}$.

(Rational Non-Imposition) This follows from Proposition 4.5 on page 105 and the fact that Υ_C^S satisfies (Existence), (Success), and (Consistency). □

Theorem 4.6 (page 121). *Let C satisfy (Inc^*-symmetry). Then Υ_C^B satisfies (Uniqueness), (Structural Preservation), (Success), (Consistency), (C-Conformity), (Inverse C-Conformity), and (Non-Dictatorship).*

Proof.

(Uniqueness) This is clear due to the functional definition of Υ_C^B, cf. Definition 4.14 on page 120. If a unique solution for minimizing f_C^B with respect to OPT_C^B exists then uniqueness of $\Upsilon_C^B(\mathcal{R})$ is clear. Otherwise $\Upsilon_C^B(\mathcal{R}) = \mathsf{undef}$ is also uniquely determined.

(Structural Preservation) This follows directly from the definition of Υ_C^B, cf. Definition 4.14 on page 120.

(Success) If $\Upsilon_C^B(\mathcal{R}) \neq \mathsf{undef}$ then $\Upsilon_C^B(\mathcal{R})$ is consistent by definition.

(Consistency) Let $\mathcal{R}$ be consistent. Then $C^{\mathcal{R}}(r) = 0$ for every $r \in \mathcal{R}$ and $\Upsilon_C^B(\mathcal{R}) = \mathcal{R}$ is the unique single adjustment.

(Non-Dictatorship) Let $(\psi \mid \phi)[d]$ be a non-tautological probabilistic conditional and consider the knowledge base $\mathcal{R} = \{(\psi \mid \phi)[d], (\psi \mid \phi)[d']\}$ with some $d' \in [0,1]$ and $d \neq d'$. Then $\mathcal{R}$ is obviously inconsistent and $\mathsf{Inc}^*(\mathcal{R}) = |d - d'|$ and $C^{\mathcal{R}}((\psi \mid \phi)[d]) = C^{\mathcal{R}}((\psi \mid \phi)[d'])$ as C satisfies (Inc^*-symmetry). It follows that $\Upsilon_C^B(\mathcal{R}) = \{(\psi \mid \phi)[(d+d')/2]\}$ and $(\psi \mid \phi)[d]$ is not dictatorial in $\mathcal{R}$.

(C-Conformity) This follows directly from the constraint (4.10) on page 120 which is included in $\mathsf{CRDevCon}(C, \mathcal{R})$.

(Inverse C-Conformity) This follows directly from the constraint (4.10) on page 120 which is included in $\mathsf{CRDevCon}(C, \mathcal{R})$. □

Lemma 5.1 (page 145). *Let $n > 1$ be some positive integer and let $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n \in (0,1]$ with $\alpha_i \leq \beta_i$ for all $i = 1, \ldots, n$. Then*

$$\left| \frac{\frac{\alpha_1}{\beta_1} + \ldots + \frac{\alpha_n}{\beta_n}}{n} - \frac{\alpha_1 + \ldots + \alpha_n}{\beta_1 + \ldots + \beta_n} \right| < \frac{n-1}{n} \tag{5.5}$$

Proof. In order to comprehend the course of the proof we first show the case $n = 2$. We have to show that

$$-\frac{1}{2} < \frac{\frac{\alpha_1}{\beta_1} + \frac{\alpha_2}{\beta_2}}{2} - \frac{\alpha_1 + \alpha_2}{\beta_1 + \beta_2} < \frac{1}{2}$$

Consider first

$$
\begin{aligned}
& & \frac{\frac{\alpha_1}{\beta_1}+\frac{\alpha_2}{\beta_2}}{2} - \frac{\alpha_1+\alpha_2}{\beta_1+\beta_2} < \frac{1}{2} \\
\textit{iff} & & \frac{\alpha_1}{\beta_1}+\frac{\alpha_2}{\beta_2} - \frac{2\alpha_1+2\alpha_2}{\beta_1+\beta_2} < 1 \\
\textit{iff} & & \alpha_1\beta_1\beta_2+\alpha_1\beta_2^2+\alpha_2\beta_1^2+\alpha_2\beta_1\beta_2 \\
& & -2\alpha_1\beta_1\beta_2 - 2\alpha_2\beta_1\beta_2 < \beta_1^2\beta_2+\beta_1\beta_2^2 \\
\textit{iff} & & \alpha_1\beta_1\beta_2+\alpha_2\beta_1\beta_2+\beta_1^2\beta_2+\beta_1\beta_2^2-\alpha_1\beta_2^2-\alpha_2\beta_1^2 > 0 \\
\textit{iff} & & \alpha_1\beta_1\beta_2+\alpha_2\beta_1\beta_2+\beta_1^2\underbrace{(\beta_2-\alpha_2)}_{x_1}+\beta_2^2\underbrace{(\beta_1-\alpha_1)}_{x_2} > 0
\end{aligned}
$$

Due to $\alpha_1 \leq \beta_1$ and $\alpha_2 \leq \beta_2$ it follows $x_1, x_2 \geq 0$. Due to the strict positivity of all α_i, β_i $(i = 1, \ldots, n)$ the above inequality is satisfied. For the other direction we assume the contrary. Consider

$$
\begin{aligned}
& & \frac{\frac{\alpha_1}{\beta_1}+\frac{\alpha_2}{\beta_2}}{2} - \frac{\alpha_1+\alpha_2}{\beta_1+\beta_2} \leq -\frac{1}{2} \\
\textit{iff} & & \alpha_1\beta_2^2+\alpha_2\beta_1^2+\beta_1^2\beta_2+\beta_1\beta_2^2-\alpha_1\beta_1\beta_2-\alpha_2\beta_1\beta_2 \leq 0
\end{aligned}
$$

Due to $\alpha_1 \leq \beta_1$ and $\alpha_2 \leq \beta_2$ it follows

$$
\begin{aligned}
& & \alpha_1\beta_2^2+\alpha_2\beta_1^2+\beta_1^2\beta_2+\beta_1\beta_2^2-\beta_1^2\beta_2-\beta_1\beta_2^2 \leq 0 \\
\textit{iff} & & \alpha_1\beta_2^2+\alpha_2\beta_1^2 \leq 0
\end{aligned}
$$

which is a contradiction since $\alpha_1, \alpha_2, \beta_1, \beta_2 > 0$.

We continue with the general case $n > 1$. We have to show that

$$
-\frac{n-1}{n} < \frac{1}{n}\sum_{j=1}^{n}\frac{\alpha_j}{\beta_j} - \frac{\sum_{j=1}^{n}\alpha_j}{\sum_{i=1}^{n}\beta_i} < \frac{n-1}{n} \quad .
$$

Consider first

$$
\frac{n-1}{n} - \frac{1}{n}\sum_{j=1}^{n}\frac{\alpha_j}{\beta_j} + \frac{\sum_{j=1}^{n}\alpha_j}{\sum_{i=1}^{n}\beta_i} > 0 \tag{5.6}
$$

$$
\begin{aligned}
\textit{iff} \quad & (n-1) - \sum_{j=1}^{n}\frac{\alpha_j}{\beta_j} + \frac{\sum_{j=1}^{n}n\alpha_j}{\sum_{i=1}^{n}\beta_i} > 0 \\
\textit{iff} \quad & (n-1)\prod_{k=1}^{n}\beta_k\sum_{i=1}^{n}\beta_i - \sum_{j=1}^{n}\alpha_j\prod_{k=1,k\neq j}^{n}\beta_k\sum_{i=1}^{n}\beta_i + \sum_{j=1}^{n}n\alpha_j\prod_{k=1}^{n}\beta_k > 0 \\
\textit{iff} \quad & (n-1)\prod_{k=1}^{n}\beta_k\sum_{i=1}^{n}\beta_i - \sum_{j=1}^{n}\left[\alpha_j\prod_{k=1,k\neq j}^{n}\beta_k\left(\sum_{i=1,i\neq j}^{n}\beta_i+\beta_j\right)\right] +
\end{aligned}
$$

$$
\begin{aligned}
& \sum_{j=1} n\alpha_j \prod_{k=1}^{n} \beta_k > 0 \\
\textit{iff}\quad & (n-1)\prod_{k=1}^{n}\beta_k \sum_{i=1}^{n}\beta_i - \sum_{j=1}^{n}\left[\alpha_j \prod_{k=1, i\neq j}^{n}\beta_k \sum_{i=1, i\neq j}^{n}\beta_i + \alpha_j \prod_{k=1}^{n}\beta_k\right] + \\
& \sum_{j=1}^{n} n\alpha_j \prod_{k=1}^{n}\beta_k > 0 \\
\textit{iff}\quad & (n-1)\prod_{k=1}^{n}\beta_k \sum_{i=1}^{n}\beta_i - \sum_{j=1}^{n}\alpha_j \prod_{k=1, i\neq j}^{n}\beta_k \sum_{i=1, i\neq j}^{n}\beta_i - \sum_{j=1}^{n}\alpha_j \prod_{k=1}^{n}\beta_k + \\
& \sum_{j=1}^{n} n\alpha_j \prod_{k=1}^{n}\beta_k > 0 \\
\textit{iff}\quad & (n-1)\prod_{k=1}^{n}\beta_k \sum_{i=1}^{n}\beta_i - \sum_{j=1}^{n}\alpha_j \prod_{k=1, i\neq j}^{n}\beta_k \sum_{i=1, i\neq j}^{n}\beta_i + \\
& \underbrace{\sum_{j=1}^{n}(n-1)\alpha_j \prod_{k=1}^{n}\beta_k}_{C} > 0
\end{aligned}
$$

Observe that $C > 0$ as $\alpha_i, \beta_i > 0$ for all $i = 1, \ldots, n$, so it suffices to show

$$
\begin{aligned}
& (n-1)\prod_{k=1}^{n}\beta_k \sum_{i=1}^{n}\beta_i - \sum_{j=1}^{n}\alpha_j \prod_{k=1, i\neq j}^{n}\beta_k \sum_{i=1, i\neq j}^{n}\beta_i \geq 0 \\
\textit{iff}\quad & \sum_{i=1}^{n}(n-1)\beta_i^2 \prod_{k=1, k\neq i}^{n}\beta_k - \sum_{j=1}^{n}\alpha_j \prod_{k=1, k\neq j}^{n}\beta_k \sum_{i=1, i\neq j}^{n}\beta_i \geq 0 \\
\textit{iff}\quad & \sum_{i=1}^{n}(n-1)\beta_i^2 \prod_{k=1, k\neq i}^{n}\beta_k - \sum_{j=1}^{n}\sum_{i=1, i\neq j}^{n}\alpha_j\beta_i^2 \prod_{k=1, k\neq j, k\neq i}^{n}\beta_k \geq 0 \\
\textit{iff}\quad & \sum_{j=1}^{n}(n-1)\beta_j^2 \prod_{k=1, k\neq j}^{n}\beta_k - \sum_{j=1}^{n}\sum_{i=1, i\neq j}^{n}\alpha_j\beta_i^2 \prod_{k=1, k\neq j, k\neq i}^{n}\beta_k \geq 0 \\
\textit{iff}\quad & \sum_{j=1}^{n}\sum_{i=1, i\neq j}^{n}\beta_j^2 \prod_{k=1, k\neq j}^{n}\beta_k - \sum_{j=1}^{n}\sum_{i=1, i\neq j}^{n}\alpha_j\beta_i^2 \prod_{k=1, k\neq j, k\neq i}^{n}\beta_k \geq 0 \\
\textit{iff}\quad & \sum_{j=1}^{n}\sum_{i=1, i\neq j}^{n}\left[\beta_j^2 \prod_{k=1, k\neq j}^{n}\beta_k - \alpha_j\beta_i^2 \prod_{k=1, k\neq i, k\neq j}^{n}\beta_k\right] \geq 0 \\
\textit{iff}\quad & \sum_{j=1}^{n}\sum_{i=1, i\neq j}^{n}\beta_j^2 \prod_{k=1, k\neq i, k\neq j}^{n}\beta_k \underbrace{\left[\beta_j - \alpha_j\right]}_{D} \geq 0
\end{aligned}
$$

As $\beta_j \geq \alpha_j$ for every $j = 1, \ldots, n$ it follows $D \geq 0$ and we have proven (5.6). For the other direction we assume the contrary. Consider

$$
\begin{aligned}
&\frac{n-1}{n} + \frac{1}{n}\sum_{j=1}^{n}\frac{\alpha_j}{\beta_j} - \frac{\sum_{j=1}^{n}\alpha_j}{\sum_{i=1}^{n}\beta_i} \leq 0 \\
\textit{iff}\quad & (n-1)\prod_{k=1}^{n}\beta_k\sum_{i=1}^{n}\beta_i + \sum_{j=1}^{n}\alpha_j\prod_{k=1,k\neq j}^{n}\beta_k\sum_{i=1}^{n}\beta_i - \sum_{j=1}^{n}n\alpha_j\prod_{k=1}^{n}\beta_k \leq 0 \\
\textit{iff}\quad & (n-1)\prod_{k=1}^{n}\beta_k\sum_{i=1}^{n}\beta_i + \sum_{j=1}^{n}\alpha_j\prod_{k=1,k\neq j}^{n}\beta_k\sum_{i=1,i\neq j}^{n}\beta_i + \sum_{j=1}^{n}\alpha_j\prod_{k=1}^{n}\beta_k \\
&\quad - \sum_{j=1}^{n}n\alpha_j\prod_{k=1}^{n}\beta_k \leq 0 \\
\textit{iff}\quad & (n-1)\prod_{k=1}^{n}\beta_k\sum_{i=1}^{n}\beta_i + \sum_{j=1}^{n}\alpha_j\prod_{k=1,k\neq j}^{n}\beta_k\sum_{i=1,i\neq j}^{n}\beta_i \\
&\quad - (n-1)\sum_{j=1}^{n}\alpha_j\prod_{k=1}^{n}\beta_k \leq 0
\end{aligned}
$$

Since $\beta_j \geq \alpha_j$ for every $j = 1, \ldots, n$ it follows

$$
\begin{aligned}
& (n-1)\prod_{k=1}^{n}\beta_k\sum_{i=1}^{n}\beta_i + \sum_{j=1}^{n}\alpha_j\prod_{k=1,k\neq j}^{n}\beta_k\sum_{i=1,i\neq j}^{n}\beta_i \\
&\quad - (n-1)\sum_{j=1}^{n}\beta_j\prod_{k=1}^{n}\beta_k \leq 0 \\
\textit{iff}\quad & \sum_{j=1}^{n}\alpha_j\prod_{k=1,k\neq j}^{n}\beta_k\sum_{i=1,i\neq j}^{n}\beta_i \leq 0
\end{aligned}
$$

which is a contradiction since $\alpha_i, \beta_i > 0$ for $i = 1, \ldots, n$. □

Proposition 6.2 (page 166). *Let $\mathcal{I}$ satisfy (Prototypical Indifference). Let $\mathcal{R}$ be a knowledge base and D finite with $\mathsf{Const}(\mathcal{R}) \subseteq D \subseteq U$ and $D \neq \emptyset$.*

1. *Let ϕ, ψ be two ground sentences. For $\mathsf{c}_1, \mathsf{c}_2 \in D$ with $\mathsf{c}_1 \equiv_{\mathcal{R}} \mathsf{c}_2$ it holds that $\mathcal{I}(\mathcal{R}, D)(\psi \,|\, \phi) = \mathcal{I}(\mathcal{R}, D)(\psi[\mathsf{c}_1 \leftrightarrow \mathsf{c}_2] \,|\, \phi[\mathsf{c}_1 \leftrightarrow \mathsf{c}_2])$.*
2. *Let $S \in \mathfrak{S}(\mathcal{R})$, $\mathsf{c}_1, \ldots, \mathsf{c}_n \in S$, and $\sigma : S \to S$ be a permutation on S, i. e. a bijective function on S. Then it holds that*

$$\mathcal{I}(\mathcal{R}, D)(\phi) = \mathcal{I}(\mathcal{R}, D)(\phi[\sigma(\mathsf{c}_1)/\mathsf{c}_1, \ldots, \sigma(\mathsf{c}_n)/\mathsf{c}_n]) \quad .$$

Proof.

1. Because of (Prototypical Indifference) it holds that

$$\mathcal{I}(\mathcal{R}, D)(\phi) = \mathcal{I}(\mathcal{R}, D)(\phi[\mathsf{c}_1 \leftrightarrow \mathsf{c}_2]) \quad \text{and}$$

$$\mathcal{I}(\mathcal{R},D)(\psi \wedge \phi) = \mathcal{I}(\mathcal{R},D)((\psi \wedge \phi)[\mathsf{c}_1 \leftrightarrow \mathsf{c}_2])$$

and hence

$$\begin{aligned}\mathcal{I}(\mathcal{R},D)(\psi \mid \phi) &= \frac{\mathcal{I}(\mathcal{R},D)(\psi \wedge \phi)}{\mathcal{I}(\mathcal{R},D)(\phi)}\\ &= \frac{\mathcal{I}(\mathcal{R},D)((\psi \wedge \phi)[\mathsf{c}_1 \leftrightarrow \mathsf{c}_2])}{\mathcal{I}(\mathcal{R},D)(\phi[\mathsf{c}_1 \leftrightarrow \mathsf{c}_2])}\\ &= \mathcal{I}(\mathcal{R},D)(\psi[\mathsf{c}_1 \leftrightarrow \mathsf{c}_2] \mid \phi[\mathsf{c}_1 \leftrightarrow \mathsf{c}_2])\end{aligned}$$

due to $(\psi \wedge \phi)[\mathsf{x}_i/\mathsf{y}_i]_{i=1,\ldots,n} = \psi[\mathsf{x}_i/\mathsf{y}_i]_{i=1,\ldots,n} \wedge \phi[\mathsf{x}_i/\mathsf{y}_i]_{i=1,\ldots,n}$.

2. This follows from the fact that every permutation can be represented as a product of transpositions (Beachy and Blair, 2005), i. e. permutations that exactly transpose two elements. Let $\sigma_1,\ldots,\sigma_m$ be such transpositions of σ, i. e., it holds that $\sigma = \sigma_m \circ \ldots \circ \sigma_1$. Due to (Prototypical Indifference) it holds that

$$\mathcal{I}(\mathcal{R},D)(\phi) = \mathcal{I}(\mathcal{R},D)(\phi[\sigma_1(\mathsf{c}_1)/\mathsf{c}_1,\ldots,\sigma_1(\mathsf{c}_\mathsf{n})/\mathsf{c}_\mathsf{n}])$$

as σ_1 transposes only two elements, i. e., it holds that there are distinct $i,j \in \{1,\ldots,n\}$ with $\sigma_1(\mathsf{c}_i) = \mathsf{c}_j$ and $\sigma_1(\mathsf{c}_j) = \mathsf{c}_i$ and $\sigma_1(\mathsf{c}_k) = \mathsf{c}_k$ for every k with $k \neq i$ and $k \neq j$. Similarly, for each $i = 2,\ldots,m$ it holds that

$$\begin{aligned}&\mathcal{I}(\mathcal{R},D)(\phi[\sigma_1(\mathsf{c}_1)/\mathsf{c}_1,\ldots,\sigma_1(\mathsf{c}_\mathsf{n})/\mathsf{c}_\mathsf{n}]\ldots\\ &\qquad\qquad [\sigma_{i-1}(\mathsf{c}_1)/\mathsf{c}_1,\ldots,\sigma_{i-1}(\mathsf{c}_\mathsf{n})/\mathsf{c}_\mathsf{n}])\\ &\quad = \mathcal{I}(\mathcal{R},D)(\phi[\sigma_1(\mathsf{c}_1)/\mathsf{c}_1,\ldots,\sigma_1(\mathsf{c}_\mathsf{n})/\mathsf{c}_\mathsf{n}]\ldots\\ &\qquad\qquad [\sigma_i(\mathsf{c}_1)/\mathsf{c}_1,\ldots,\sigma_i(\mathsf{c}_\mathsf{n})/\mathsf{c}_\mathsf{n}]) \quad .\end{aligned}$$

Via transitivity it follows

$$\begin{aligned}\mathcal{I}(\mathcal{R},D)(\phi) = \mathcal{I}(\mathcal{R},D)(&\phi[\sigma_1(\mathsf{c}_1)/\mathsf{c}_1,\ldots,\sigma_1(\mathsf{c}_\mathsf{n})/\mathsf{c}_\mathsf{n}]\ldots\\ &[\sigma_m(\mathsf{c}_1)/\mathsf{c}_1,\ldots,\sigma_m(\mathsf{c}_\mathsf{n})/\mathsf{c}_\mathsf{n}]) \quad .\end{aligned}$$

which is equivalent to the claim due to $\sigma = \sigma_m \circ \ldots \circ \sigma_1$. □

Theorem 6.1 (page 170). *$\mathcal{I}_\varnothing$ satisfies (Reflexivity), (Left Logical Equivalence), (Right Weakening), (Cumulativity), (Name Irrelevance), (Prototypical Indifference), (*ME*-Compatibility), (Compensation), and (Strict Inference). If Conjecture 6.1 is true then $\mathcal{I}_\varnothing$ also satisfies (Convergence).*

Proof.

(Reflexivity) If $\mathcal{I}_\varnothing(\mathcal{R},D) \neq \mathsf{undef}$ it follows that $\mathcal{I}_\varnothing(\mathcal{R},D), D \models^{pr}_\varnothing \mathcal{R}$ by definition and therefore $\mathcal{I}_\varnothing(\mathcal{R},D), D \models^{pr}_\varnothing r$ for every $r \in \mathcal{R}$.

(Left Logical Equivalence) Let $\mathcal{R}_1$ and $\mathcal{R}_2$ be knowledge bases with $\mathcal{R}_1 \equiv^{\varnothing} \mathcal{R}_2$. This means for all P and D that $P, D \models^{pr}_{\varnothing} \mathcal{R}_1$ whenever $P, D \models^{pr}_{\varnothing} \mathcal{R}_2$. Then $\mathcal{I}_{\varnothing}(\mathcal{R}_1, D)$ and $\mathcal{I}_{\varnothing}(\mathcal{R}_2, D)$ are defined on the same set of probability functions in (6.15) on page 170 and therefore $\mathcal{I}_{\varnothing}(\mathcal{R}_1, D) = \mathcal{I}_{\varnothing}(\mathcal{R}_2, D)$.

(Right Weakening) Let $\mathcal{R}$ and D such that $\mathcal{I}_{\varnothing}(\mathcal{R}, D), D \models^{pr}_{\varnothing} (\psi \mid \phi)[d]$ and let $(\psi \mid \phi)[d] \models^{pr} (\psi' \mid \phi')[d']$. As $\mathcal{I}_{\varnothing}(\mathcal{R}, D)$ satisfies $(\psi \mid \phi)[d]$ with respect to D it also satisfies$(\psi' \mid \phi')[d']$ with respect to D. Hence, it follows $\mathcal{I}_{\varnothing}(\mathcal{R}, D), D \models^{pr}_{\varnothing} (\psi' \mid \phi')[d']$.

(Cumulativity) For every probability function P and set D it holds that from $P, D \models^{pr}_{\varnothing} \mathcal{R} \cup \{(\psi \mid \phi)[d]\}$ it follows that $P, D \models^{pr}_{\varnothing} \mathcal{R}$ as $\mathcal{R} \subseteq \mathcal{R} \cup \{(\psi \mid \phi)[d]\}$. It also holds that $\mathcal{I}_{\varnothing}(\mathcal{R}, D), D \models^{pr}_{\varnothing} \mathcal{R} \cup \{(\psi \mid \phi)[d]\}$ as $\mathcal{I}_{\varnothing}(\mathcal{R}, D), D \models^{pr}_{\varnothing} (\psi \mid \phi)[d]$ by assumption. Suppose it holds that $\mathcal{I}_{\varnothing}(\mathcal{R} \cup \{(\psi \mid \phi)[d]\}, D) \neq \mathcal{I}_{\varnothing}(\mathcal{R}, D)$ then

$$H_D(\mathcal{I}_{\varnothing}(\mathcal{R} \cup \{(\psi \mid \phi)[d]\}, D)) > H_D(\mathcal{I}_{\varnothing}(\mathcal{R}, D)) \quad .$$

It follows that $\mathcal{I}_{\varnothing}(\mathcal{R} \cup \{(\psi \mid \phi)[d]\}, D)$ should be the ME-model of $\mathcal{R}$ as well because $\mathcal{I}_{\varnothing}(\mathcal{R} \cup \{(\psi \mid \phi)[d]\}, D), D \models^{pr}_{\varnothing} \mathcal{R}$. Hence, $\mathcal{I}_{\varnothing}(\mathcal{R} \cup \{(\psi \mid \phi)[d]\}, D) = \mathcal{I}_{\varnothing}(\mathcal{R}, D)$ and it follows that $\mathcal{I}_{\varnothing}(\mathcal{R}, D), D \models^{pr}_{\varnothing} (\psi' \mid \phi')[d']$ whenever $\mathcal{I}_{\varnothing}(\mathcal{R} \cup \{(\psi \mid \phi)[d]\}, D), D \models^{pr}_{\varnothing} (\psi' \mid \phi')[d']$ for every $(\psi' \mid \phi')[d']$.

(ME-Compatibility) Let $\mathcal{R}$ be a ground knowledge base. Due to Remark 5.1 on page 137 the operator $\models^{pr}_{\varnothing}$ is equivalent to $\models^{pr}$ in the propositional case. Then Equation (6.15) on page 170 also becomes equivalent to the propositional case. Hence, holds that $\mathsf{ME}(\mathcal{R}')(\phi) = \mathcal{I}_{\varnothing}(\mathcal{R}, D)(\phi)$ for every ground sentence ϕ.

(Name Irrelevance Shore and Johnson showed in (Shore and Johnson, 1980) that the principle of entropy is indifferent to syntactical variants and coordinate transformations. As renaming a constant changes the represented knowledge only on a syntactical way we obtain the function $\mathcal{I}_{\varnothing}(\mathcal{R}[\mathsf{d}/\mathsf{c}], (D \cup \{\mathsf{d}\}) \setminus \{\mathsf{c}\})$ via

$$\mathcal{I}_{\varnothing}(\mathcal{R}[\mathsf{d}/\mathsf{c}], (D \cup \{\mathsf{d}\}) \setminus \{\mathsf{c}\})(\omega) =_{def} \mathcal{I}_{\varnothing}(\mathcal{R}, D)(\omega[\mathsf{d} \leftrightarrow \mathsf{c}])$$

for all $\omega \in \Omega(\Sigma)$.

(Prototypical Indifference) This follows by Proposition 6.1 on page 164.

(Compensation) Let $\mathcal{R}$ be a knowledge base and $(\psi(\vec{\mathsf{X}}) \mid \phi(\vec{\mathsf{X}}))[d] \in \mathcal{R}$ a non-ground probabilistic conditional with $d \in (0, 1)$. Suppose

$$\mathcal{I}_{\varnothing}(\mathcal{R}, D)(\psi(\vec{\mathsf{c}}) \mid \phi(\vec{\mathsf{c}})) < d$$

for all $(\psi(\vec{\mathsf{c}}) \mid \phi(\vec{\mathsf{c}}))[d] \in \mathsf{gnd}_D^{\mathcal{I}_{\varnothing}(\mathcal{R}, D)}((\psi(\vec{\mathsf{c}}) \mid \phi(\vec{\mathsf{c}})))$. Then (for finite D) it holds that

$$\begin{aligned} & \frac{\sum_{(\psi(\vec{c})\,|\,\phi(\vec{c}))\in \mathsf{gnd}_D^{\mathcal{I}_\varnothing(\mathcal{R},D)}((\psi(\vec{X})\,|\,\phi(\vec{X})))} P(\psi(\vec{c})\,|\,\phi(\vec{c}))}{|\mathsf{gnd}_D^{\mathcal{I}_\varnothing(\mathcal{R},D)}((\psi(\vec{X})\,|\,\phi(\vec{X})))|} \\ < \quad & \frac{d\cdot|\mathsf{gnd}_D^{\mathcal{I}_\varnothing(\mathcal{R},D)}((\psi(\vec{X})\,|\,\phi(\vec{X})))|}{|\mathsf{gnd}_D^{\mathcal{I}_\varnothing(\mathcal{R},D)}((\psi(\vec{X})\,|\,\phi(\vec{X})))|} = d \end{aligned}$$

contradicting $\mathcal{I}_\varnothing(\mathcal{R},D), D \models^{pr}_\varnothing \mathcal{R}$.

(Strict Inference) Let $\mathcal{R}$ be a knowledge base and $(\psi(\vec{X})\,|\,\phi(\vec{X}))[d] \in \mathcal{R}$ a non-ground probabilistic conditional with $d = 1$ (the case of $d = 0$ can be shown analogously). Suppose

$$\mathcal{I}_\varnothing(\mathcal{R},D)(\psi(\vec{c})\,|\,\phi(\vec{c})) < 1$$

for some $(\psi(\vec{c})\,|\,\phi(\vec{c}))[d] \in \mathsf{gnd}_D^{\mathcal{I}_\varnothing(\mathcal{R},D)}(\mathcal{I}_\varnothing(\mathcal{R},D))(\psi(\vec{c})\,|\,\phi(\vec{c}))$. Then (for finite D) it holds that

$$\begin{aligned} & \frac{\sum_{(\psi(\vec{c})\,|\,\phi(\vec{c}))\in \mathsf{gnd}_D^{\mathcal{I}_\varnothing(\mathcal{R},D)}((\psi(\vec{X})\,|\,\phi(\vec{X})))} P(\psi(\vec{c})\,|\,\phi(\vec{c}))}{|\mathsf{gnd}_D^{\mathcal{I}_\varnothing(\mathcal{R},D)}((\psi(\vec{X})\,|\,\phi(\vec{X})))|} \\ < \quad & \frac{|\mathsf{gnd}_D^{\mathcal{I}_\varnothing(\mathcal{R},D)}((\psi(\vec{X})\,|\,\phi(\vec{X})))|}{|\mathsf{gnd}_D^{\mathcal{I}_\varnothing(\mathcal{R},D)}((\psi(\vec{X})\,|\,\phi(\vec{X})))|} = 1 \end{aligned}$$

contradicting $\mathcal{I}_\varnothing(\mathcal{R},D), D \models^{pr}_\varnothing \mathcal{R}$.

(Convergence) Assume that $\mathcal{I}_\varnothing$ satisfies (Well-definedness). Let $\mathcal{R}$ be a knowledge base and D finite with $\mathsf{Const}(\mathcal{R}) \subseteq D \subseteq U$ with $D \neq \emptyset$ and $P^* =_{def} \mathcal{I}_\varnothing(\mathcal{R},D) \neq \mathsf{undef}$. Let $r = (\psi(\vec{X})\,|\,\phi(\vec{X}))[d] \in \mathcal{R}$ be a conditional with $\vec{X} = (X_1,\ldots,X_h)$ and $c_1,\ldots,c_n$ the constants that appear in $\mathcal{R}$. Let furthermore

$$\{d_1,\ldots,d_m\} = D \setminus \{c_1,\ldots,c_n\} \quad ,$$

so it follows $|D| = n+m$. Let $\vec{d}_1,\ldots,\vec{d}_k$ be all vectors of constants in $\{d_1,\ldots,d_m\}$ of length h such that for every $\vec{d}_i = (d'_1,\ldots,d'_h)$ it holds that $d'_e \neq d'_j$ for $e \neq j$ (for $i = 1,\ldots,k$ and $e,j = 1,\ldots,h$). Let $\vec{c}_1,\ldots,\vec{c}_l$ be all remaining vectors of constants in D. It follows that $(l+k) = (|D|)^h = (n+m)^h$ and

$$k = m^{\underline{h}} =_{def} m(m-1)\ldots(m-h+1) \quad \text{(the falling factorial)}$$

and thus $l = (n+m)^h - m^{\underline{h}}$. Let $P^*_{\vec{c}}$ denote $P^*(\psi(\vec{c})\,|\,\phi(\vec{c}))$ for a vector $\vec{c}$. In order to have $P^*, D \models^{pr}_\varnothing r$ it must hold that $P^*_{\vec{c}_1} + \ldots + P^*_{\vec{c}_l} + P^*_{\vec{d}_1} + \ldots + P^*_{\vec{d}_k} = d\cdot(k+l)$. From (Prototypical Indifference) and 1.) and 2.)

in Proposition 6.2 on page 166 it follows that $P^*_{\vec{\mathsf{d}}_1} = \ldots = P^*_{\vec{\mathsf{d}}_k}$. Define $P^*_k =_{def} P^*_{\vec{\mathsf{d}}_1}$, so it holds that $P^*_{\vec{\mathsf{d}}_1} + \ldots + P^*_{\vec{\mathsf{d}}_k} = kP^*_k$. It follows

$$\begin{aligned} P^*_k &= \frac{d \cdot (k+l) - P^*_{\vec{\mathsf{c}}_1} - \ldots - P^*_{\vec{\mathsf{c}}_l}}{k} \leq \frac{d \cdot (k+l)}{k} \\ &= d \underbrace{\frac{(n+m)^h}{m^{\underline{h}}}}_{\overset{m\to\infty}{\to} 1} \overset{m\to\infty}{\to} d \end{aligned}$$

Similarly it holds that

$$\begin{aligned} P^*_k &= \frac{d \cdot (k+l) - P^*_{\vec{\mathsf{c}}_1} - \ldots - P^*_{\vec{\mathsf{c}}_l}}{k} \geq \frac{d \cdot (k+l) - l}{k} \\ &= d \underbrace{\frac{(n+m)^h}{m^{\underline{h}}}}_{\overset{m\to\infty}{\to} 1} - \underbrace{\frac{(n+m)^h - m^{\underline{h}}}{m^{\underline{h}}}}_{\overset{m\to\infty}{\to} 0} \overset{m\to\infty}{\to} d \quad . \end{aligned}$$

Due to (Well-definedness) all (implicitly) appearing probability functions are well-defined and it follows $P^*_k \to d$ for $m \to \infty$. □

Lemma 6.1 (page 174). *Let $r = (\psi(\vec{\mathsf{X}}) \,|\, \phi(\vec{\mathsf{X}}))[d]$ be a probabilistic conditional, D finite with $\mathsf{Const}(r) \subseteq D \subseteq U$ and $D \neq \emptyset$, and Sol^D_r the set of probability functions that satisfy r, i.e., it holds that $Sol^D_r = \{P \mid P, D \models^{pr}_{\odot} (\psi(\vec{\mathsf{X}}) \,|\, \phi(\vec{\mathsf{X}}))[d]\}$. Then Sol_r is convex.*

Proof. Let P_1 and P_2 be some probability functions with $P_1, D \models^{pr}_{\odot} r$ and $P_2, D \models^{pr}_{\odot} r$. We have to show that any convex combination of P_1 and P_2 satisfies r as well. Let Q be a convex combination of P_1 and P_2, i.e., let $\delta \in (0,1)$ be fixed and define $Q(\omega) =_{def} \delta P_1(\omega) + (1-\delta)P_2(\omega)$ for any $\omega \in \Omega(\Sigma)$. Then it holds that $Q(\psi') = \delta P_1(\psi') + (1-\delta)P_2(\psi')$ for any ground formula ψ' as well. Let $\{(\psi_1 \,|\, \phi_1), \ldots, (\psi_n \,|\, \phi_n)\} = \mathsf{gnd}_D((\psi(\vec{\mathsf{X}}) \,|\, \phi(\vec{\mathsf{X}})))$. Then it holds that

$$\frac{\sum_{i=1}^n P_j(\psi_i \phi_i)}{\sum_{i=1}^n P_j(\phi_i)} = d \tag{6.7}$$

for $j = 1, 2$ and we have to show that

$$\frac{\sum_{i=1}^n Q(\psi_i \phi_i)}{\sum_{i=1}^n Q(\phi_i)} = d$$

which is equivalent to

$$\frac{\delta \sum_{i=1}^n P_1(\psi_i \phi_i) + (1-\delta) \sum_{i=1}^n P_2(\psi_i \phi_i)}{\delta \sum_{i=1}^n P_1(\phi_i) + (1-\delta) \sum_{i=1}^n P_2(\phi_i)} = d \tag{6.8}$$

If $d = 0$ then $P_j(\psi_i \wedge \phi_i) = 0$ for all $i = 1, \ldots, n$ and $j = 1, 2$ due to (6.7). Then also $Q(\psi_i \wedge \phi_i) = 0$ for all $i = 1 \ldots, n$ and it follows $Q, D \models^{pr}_{\odot} r$. We continue with $d > 0$. Then (6.8) is equivalent to (all appearing sums are meant to range over $i = 1, \ldots, n$)

$$\begin{aligned}
&\delta \sum P_1(\psi_i\phi_i) + (1-\delta) \sum P_2(\psi_i\phi_i) = d\delta \sum P_1(\phi_i) + d(1-\delta) \sum P_2(\phi_i) \\
&\textit{iff} \quad \frac{\delta \sum P_1(\psi_i\phi_i)}{d\delta \sum P_1(\phi_i)} + \frac{(1-\delta) \sum P_2(\psi_i\phi_i)}{d\delta \sum P_1(\phi_i)} = 1 + \frac{d(1-\delta) \sum P_2(\phi_i)}{d\delta \sum P_1(\phi_i)} \\
&\textit{iff} \quad 1 + \frac{(1-\delta) \sum P_2(\psi_i\phi_i)}{d\delta \sum P_1(\phi_i)} = 1 + \frac{(1-\delta) \sum P_2(\phi_i)}{\delta \sum P_1(\phi_i)} \\
&\textit{iff} \quad \frac{1}{d}(1-\delta) \sum P_2(\psi_i \wedge \phi_i) = (1-\delta) \sum P_2(\phi_i) \\
&\textit{iff} \quad 1 = 1
\end{aligned}$$

and it follows $Q, D \models^{pr}_{\odot} r$. □

Theorem 6.2 (page 181)**.** *Let $\mathcal{G}$ be some grounding operator. Then the inference operator $\mathcal{I}_{\mathcal{G}}$ satisfies (Left Logical Equivalence), (Right Weakening), (Cumulativity), (Well-Definedness), (Name Irrelevance), and (Prototypical Indifference).*

Proof.

(Left Logical Equivalence) For grounding semantics, equivalence of knowledge bases is straightforwardly defined as follows. Let $\mathcal{G}$ be a grounding operator. Then $\mathcal{R}_1 \equiv_{\mathcal{G}} \mathcal{R}_2$ if and only if $\mathcal{G}(\mathcal{R}_1, D) \equiv^{kb} \mathcal{G}(\mathcal{R}_2, D)$ in the classical sense. By assuming $\mathcal{R}_1 \equiv_{\mathcal{G}} \mathcal{R}_2$ it follows $P, D \models_{\mathcal{G}} \mathcal{R}_1$ whenever $P, D \models_{\mathcal{G}} \mathcal{R}_2$. It follows that (6.60) on page 181 is defined on the same set of probability functions for both $\mathcal{R}_1$ and $\mathcal{R}_2$ and therefore $\mathcal{I}_{\mathcal{G}}(\mathcal{R}_1, D) = \mathcal{I}_{\mathcal{G}}(\mathcal{R}_2, D)$.

(Right Weakening) Let $\mathcal{R}$ and D such that $\mathcal{I}_{\mathcal{G}}(\mathcal{R}, D), D \models_{\mathcal{G}} (\psi \,|\, \phi)[d]$ and let $(\psi \,|\, \phi)[d] \models^{pr} (\psi' \,|\, \phi')[d']$. As $\mathcal{I}_{\mathcal{G}}(\mathcal{R}, D)$ satisfies $(\psi \,|\, \phi)[d]$ with respect to D it also satisfies $(\psi' \,|\, \phi')[d']$ with respect to D. Hence, it follows $\mathcal{I}_{\mathcal{G}}(\mathcal{R}, D), D \models_{\mathcal{G}} (\psi' \,|\, \phi')[d']$.

(Cumulativity) For every P and D it holds that $P, D \models_{\mathcal{G}} \mathcal{R}$ follows from $P, D \models_{\mathcal{G}} \mathcal{R} \cup \{(\psi \,|\, \phi)[d]\}$ as $\mathcal{R} \subseteq \mathcal{R} \cup \{(\psi \,|\, \phi)[d]\}$. It also holds that $\mathcal{I}_{\mathcal{G}}(\mathcal{R}, D), D \models_{\mathcal{G}} \mathcal{R} \cup \{(\psi \,|\, \phi)[d]\}$ as $\mathcal{I}_{\mathcal{G}}(\mathcal{R}, D), D \models_{\mathcal{G}} (\psi \,|\, \phi)[d]$ by assumption. Suppose that $\mathcal{I}_{\mathcal{G}}(\mathcal{R} \cup \{(\psi \,|\, \phi)[d]\}, D) \neq \mathcal{I}_{\mathcal{G}}(\mathcal{R}, D)$ then $H_D(\mathcal{I}_{\mathcal{G}}(\mathcal{R} \cup \{(\psi \,|\, \phi)[d]\}, D)) > H_D(\mathcal{I}_{\mathcal{G}}(\mathcal{R}, D))$. It follows that $\mathcal{I}_{\mathcal{G}}(\mathcal{R} \cup \{(\psi \,|\, \phi)[d]\}, D)$ should be the ME-model of $\mathcal{R}$ as well because $\mathcal{I}_{\mathcal{G}}(\mathcal{R} \cup \{(\psi \,|\, \phi)[d]\}, D), D \models_{\mathcal{G}} \mathcal{R}$. Hence, $\mathcal{I}_{\mathcal{G}}(\mathcal{R} \cup \{(\psi \,|\, \phi)[d]\}, D) = \mathcal{I}_{\mathcal{G}}(\mathcal{R}, D)$ and therefore it holds that $\mathcal{I}_{\mathcal{G}}(\mathcal{R}, D), D \models_{\mathcal{G}} (\psi' \,|\, \phi')[d']$ whenever $\mathcal{I}_{\mathcal{G}}(\mathcal{R} \cup \{(\psi \,|\, \phi)[d]\}, D), D \models_{\mathcal{G}} (\psi' \,|\, \phi')[d']$ for every probabilistic conditional $(\psi' \,|\, \phi')[d']$.

(Well-Definedness) If $\mathcal{R}$ is $\mathcal{G}$-consistent with respect to D then (6.60) on page 181 is semantically equivalent to (2.8) on page 32 which is well-defined.

(Name Irrelevance) This is obvious as the principle of maximum entropy is unbiased to renaming of constants, cf. Proposition 6.1 on page 170.

(Prototypical Indifference) This follows from Proposition 6.1 on page 164. $\square$

Proposition 7.3 (page 198). *$\equiv_{\mathcal{R}}$ is an equivalence relation.*

Proof. We have to show that $\equiv_{\mathcal{R}}$ is reflexive, symmetric, and transitive.

(Reflexivity) For $\omega \in \Omega(\Sigma, D)$ it holds that $\omega = \omega$ (using the identity replacement). It follows that $\omega \equiv_{\mathcal{R}} \omega$.

(Symmetry) Let $\omega_1, \omega_2 \in \Omega(\Sigma, D)$ with $\omega_1 \equiv_{\mathcal{R}} \omega_2$. Then there is a set $T = \{(\mathsf{c}_1^1, \mathsf{c}_2^1), \ldots, (\mathsf{c}_1^G, \mathsf{c}_2^G)\} \subseteq \underline{S}_1 \times \underline{S}_1 \cup \ldots \cup \underline{S}_n \times \underline{S}_n$ with

$$\omega_1 = \omega_2[\mathsf{c}_1^1 \leftrightarrow \mathsf{c}_2^1] \ldots [\mathsf{c}_1^G \leftrightarrow \mathsf{c}_2^G] \quad .$$

As substituting c_i^1 with c_i^2 and vice versa is a symmetric operation for all $i = 1, \ldots, G$ it also holds that $\omega_2 = \omega_1[\mathsf{c}_1^1 \leftrightarrow \mathsf{c}_2^1] \ldots [\mathsf{c}_1^G \leftrightarrow \mathsf{c}_2^G]$. It follows that $\omega_1 \equiv_{\mathcal{R}} \omega_2$.

(Transitivity) Let $\omega_1, \omega_2 \in \Omega(\Sigma, D)$ with $\omega_1 \equiv_{\mathcal{R}} \omega_2$. Then there is a set $T = \{(\mathsf{c}_1^1, \mathsf{c}_2^1), \ldots, (\mathsf{c}_1^G, \mathsf{c}_2^G)\} \subseteq \underline{S}_1 \times \underline{S}_1 \cup \ldots \cup \underline{S}_n \times \underline{S}_n$ with

$$\omega_1 = \omega_2[\mathsf{c}_1^1 \leftrightarrow \mathsf{c}_2^1] \ldots [\mathsf{c}_1^G \leftrightarrow \mathsf{c}_2^G] \quad .$$

Furthermore, let $\omega_3 \in \Omega(\Sigma, D)$ with $\omega_2 \equiv_{\mathcal{R}} \omega_3$. Then there is a set $T = \{(\mathsf{d}_1^1, \mathsf{d}_2^1), \ldots, (\mathsf{d}_1^{G'}, \mathsf{d}_2^{G'})\} \subseteq \underline{S}_1 \times \underline{S}_1 \cup \ldots \cup \underline{S}_n \times \underline{S}_n$ with

$$\omega_2 = \omega_3[\mathsf{d}_1^1 \leftrightarrow \mathsf{d}_2^1] \ldots [\mathsf{d}_1^{G'} \leftrightarrow \mathsf{d}_2^{G'}] \quad .$$

Then it holds that

$$\begin{aligned}\omega_1 &= \omega_2[\mathsf{c}_1^1 \leftrightarrow \mathsf{c}_2^1] \ldots [\mathsf{c}_1^G \leftrightarrow \mathsf{c}_2^G] \\ &= \omega_3[\mathsf{d}_1^1 \leftrightarrow \mathsf{d}_2^1] \ldots [\mathsf{d}_1^{G'} \leftrightarrow \mathsf{d}_2^{G'}][\mathsf{c}_1^1 \leftrightarrow \mathsf{c}_2^1] \ldots [\mathsf{c}_1^G \leftrightarrow \mathsf{c}_2^G] \quad .\end{aligned}$$

It follows that $\omega_1 \equiv_{\mathcal{R}} \omega_3$. $\square$

Proposition 7.9 (page 206). *Let ψ be a conjunction of ground literals, let $\mathsf{Const}(\psi) = \{\mathsf{c}_1, \ldots, \mathsf{c}_m\}$, and let $\Theta = \{t_1, \ldots, t_T\}$. Then*

$$\Lambda(\hat{\omega}, \psi) = \sum_{(t_1', \ldots, t_m') \in \Theta(\psi, \mathsf{c}_1, \ldots, \mathsf{c}_m)} \prod_{i=1}^{n} \binom{|\underline{S}_i \setminus \mathsf{Const}(\phi)|}{\alpha_i^{t_1}(t_1', \ldots, t_m'), \ldots, \alpha_i^{t_T}(t_1', \ldots, t_m')}$$

with $\alpha_i^t(t_1', \ldots, t_m') =_{def} \hat{\omega}(t)(\underline{S}_i) - |\{k \mid t_k' = t \wedge \mathsf{c}_k \in \underline{S}_i\}|$.

Proof. Let us first consider the case of $\Lambda(\hat{\omega}, \psi) > 0$. In order for $\Lambda(\hat{\omega}, \psi) > 0$ to be true, for every constant $\mathsf{c} \in \mathsf{Const}(\psi)$ there must be a truth configuration $t \in \Theta$ such that $t^{\wedge}(\mathsf{c}) \wedge \psi$ is satisfiable and $\hat{\omega}(t)(\underline{S_i}) > 0$ for $\mathsf{c} \in \underline{S_i}$. This means that there must be at least one constant from $\underline{S_i}$ assigned to t such that c can fulfill the role of this constant. More precisely, as there may be multiple constants in ψ compatible with t and belonging to $\underline{S_i}$ the value $\hat{\omega}(t)(\underline{S_i})$ must be at least as large as the number of those constants. In order to determine $\Lambda(\hat{\omega}, \psi)$ we have to consider every compatible truth configuration for every constant in ψ, that is every $(t'_1, \ldots, t'_m) \in \Theta(\psi, \mathsf{c}_1, \ldots, \mathsf{c}_m)$. Let such a $(t'_1, \ldots, t'_m)$ be denoted by *truth profile*. A truth profile $(t'_1, \ldots, t'_m)$ picks one truth configuration t'_i for every constant c_i appearing in ψ such that $(t'_i)^{\wedge}(\mathsf{c}_i) \wedge \psi$ is satisfiable. It follows that every Herbrand interpretation $\omega \in \kappa^{-1}(\hat{\omega})$ with $\omega \models^{\mathrm{F}} (t'_1)^{\wedge}(\mathsf{c}_1) \wedge \ldots \wedge (t'_m)^{\wedge}(\mathsf{c}_m)$ also satisfies $\omega \models^{\mathrm{F}} \psi$. By considering all truth profiles compatible with ψ and counting the number of Herbrand interpretations that fulfill the previous relationship for each such truth profile we counted the number of all Herbrand interpretations in $\kappa^{-1}(\hat{\omega})$ that satisfy ψ. Let $G(t'_1, \ldots, t'_m)$ denote the number of Herbrand interpretations $\omega \in \kappa^{-1}(\hat{\omega})$ with $\omega \models^{\mathrm{F}} (t'_1)^{\wedge}(\mathsf{c}_1) \wedge \ldots \wedge (t'_m)^{\wedge}(\mathsf{c}_m)$. Then we can write

$$\Lambda(\hat{\omega}, \psi) = \sum_{(t'_1,\ldots,t'_m) \in \Theta(\psi, \mathsf{c}_1,\ldots,\mathsf{c}_m)} G(t'_1, \ldots, t'_m) \quad .$$

As shown in Proposition 7.6 on page 202 the number of Herbrand interpretations represented by a reference world $\hat{\omega}$ is given via

$$\rho_{\hat{\omega}} = \prod_{i=1}^{n} \binom{|\underline{S_i}|}{\hat{\omega}(t_1)(\underline{S_i}), \ldots, \hat{\omega}(t_T)(\underline{S_i})} \tag{7.9}$$

It may not be the case that every Herbrand interpretation ω represented by $\hat{\omega}$ actually satisfies ψ. Given some truth profile $(t'_1, \ldots, t'_m)$ for the constants appearing in ψ several characteristics of ω are already predetermined by $(t'_1, \ldots, t'_m)$. Consider an $\mathcal{R}$-equivalence class $\underline{S_i}$ and $\mathsf{c}_j \in \mathsf{Const}(\mathcal{R})$ that is already assigned to the truth configuration t'_j. Then this assignment has to be taken into account when determining the number of Herbrand interpretations ω represented by $\hat{\omega}$ and also obeying $\omega \models^{\mathrm{F}} (t'_j)^{\wedge}(\mathsf{c}_j)$. In particular, the term $|\{k \mid t'_k = t \wedge \mathsf{c}_k \in \underline{S_i}\}|$ represents the number of constants in $\mathsf{Const}(\mathcal{R}) \cap \underline{S_i}$ that are already assigned to the truth configuration t. Consequently, the term

$$\alpha_i^t(t'_1, \ldots, t'_m) = \hat{\omega}(t)(\underline{S_i}) - |\{k \mid t'_k = t \wedge \mathsf{c}_k \in \underline{S_i}\}| \quad .$$

is the number of remaining constants in $\underline{S_i}$ that have not been assigned a truth configuration yet by the truth profile $(t'_1, \ldots, t'_m)$. So, in order the determine the number of Herbrand interpretations that are represented by $\hat{\omega}$ and are compatible with $(t'_1, \ldots, t'_m)$ one needs to distribute the remaining constants of each $\mathcal{R}$-equivalence class onto the the remaining places assigned for each truth configuration by $\hat{\omega}$. The term

$$\binom{|\underline{S}_i \setminus \mathsf{Const}(\phi)|}{\alpha_i^{t_1}(t'_1, \ldots, t'_m), \ldots, \alpha_i^{t_T}(t'_1, \ldots, t'_m)} \tag{7.10}$$

is exactly the number of those combinations. Note also, that if the current truth profile is incompatible with $\hat{\omega}$—i. e. there are already more constants from one $\mathcal{R}$-equivalence class $\underline{S}_i$ assigned to a truth configuration t than expected by $\hat{\omega}(t)(\underline{S}_i)$—then $\alpha_i^{t_1}(t'_1, \ldots, t'_m) < 0$ and by definition the term (7.10) is zero. This is also the intended meaning as in this case there is no Herbrand interpretation represented by $\hat{\omega}$ that is compatible with the current truth profile. Finally, we have to multiply the term (7.10) for each $\mathcal{R}$-equivalence class and obtain

$$G(t'_1, \ldots, t'_m) = \prod_{i=1}^{n} \binom{|\underline{S}_i \setminus \mathsf{Const}(\phi)|}{\alpha_i^{t_1}(t'_1, \ldots, t'_m), \ldots, \alpha_i^{t_T}(t'_1, \ldots, t'_m)}$$

which proves the claim. □

C EXAMPLES – PCL

In the following, we list further examples employing the inconsistency measures, culpability measures, and consistency restorers from Chapters 3 and 4. The functions have been applied on the following knowledge bases.

$$\mathcal{R}_0 =_{def} \{\ r_1 = (a)[0.3],\ r_2 = (a)[0.7]\ \}$$

$$\mathcal{R}_1 =_{def} \{\ r_1 = (b|a)[1]\ \ r_2 = (a)[1]\ \ r_3 = (b)[0]\ \}$$

$$\mathcal{R}_2 =_{def} \{\ r_1 = (a)[0.3]\ \ r_2 = (a)[0.7]\ \ r_3 = (b)[0.8]\ \}$$

$$\mathcal{R}_3 =_{def} \{\ r_1 = (a)[0.3]\ \ r_2 = (b)[0.4]\ \ r_3 = (a \wedge b)[0.6]\ \}$$

$$\mathcal{R}_4 =_{def} \{\ r_1 = (b|a)[0.8]\ \ r_2 = (b|\overline{a})[0.6]\ \ r_3 = (a)[0.5]\ \ r_4 = (b)[0.2]\ \}$$

$$\mathcal{R}_5 =_{def} \{\ r_1 = (b|a)[0.7]\ \ r_2 = (c|b)[0.6]\ \ r_3 = (a)[0.9]\ \ r_4 = (c)[0.1]\ \}$$

$$\mathcal{R}_6 =_{def} \{\ r_1 = (b|a)[0.7]\ \ r_2 = (c|b)[0.6]\ \ r_3 = (a)[0.9]\ \ r_4 = (c)[0.1]\ \ r_5 = (b)[0.8]\ \}$$

$$\mathcal{R}_7 =_{def} \{\ r_1 = (a \wedge b)[0.7]\ \ r_2 = (b|a)[0.9]\ \ r_3 = (c|b)[0.1]\ \ r_4 = (a|c)[0.1]\ \ r_5 = (a|\overline{c})[0.2]\ \}$$

$$\mathcal{R}_8 =_{def} \{\ r_1 = (b|a)[1]\ \ r_2 = (a)[1]\ \ r_3 = (b)[0]\ \ r_4 = (c)[0.3]\ \ r_5 = (c)[0.7]\ \}$$

$$\mathcal{R}_9 =_{def} \{\ r_1 = (b|a)[0.8]\ \ r_2 = (c|b)[0.9]\ \ r_3 = (d|c)[0.7]\ \ r_4 = (e|d)[0.9]\ \ r_5 = (a|e)[0.1]\ \ r_6 = (a)[0.9]\ \}$$

Tables 10 to 28 with even numbers show inconsistency values with respect to the inconsistency measures Inc, $\mathsf{Inc}^{\mathsf{MI}}$, $\mathsf{Inc}^{\mathsf{MI}}_0$, $\mathsf{Inc}^{\mathsf{MI}}_{C,0}$, Inc^*, Inc^*_0, and the approximations $\mathcal{I}^{\leq}$, $\mathcal{I}^{\leq}_0$, $\mathcal{I}^{\geq}_0$, and $\mathcal{I}^{\geq}$ on each of the above knowledge bases as well as the culpability values of each probabilistic conditional with respect to the culpability measures A and S. Tables 11 to 29 with odd numbers show the values of the modified probabilistic conditionals after application of the consistency restorer Υ^{U}, Υ^{P}, Υ^{S}_{C}, and Υ^{B}_{C} for the various culpability measures. In each of the Tables 11 to 29 with odd numbers the modified probability of a probabilistic conditional r_i is shown under column r'_i. Cells containing a "—" indicate that the value is not defined due to the non-existence of a unique function value of Υ^{B}_{C}. Those tables also show the distance to the original knowledge base with respect to the following distance concepts. Let $d_1, \ldots, d_n$ the original values of the probabilistic conditionals in $\mathcal{R}$ and $d'_1, \ldots, d'_n$ the modified values in $\Upsilon(\mathcal{R})$. Then the distances are defined via

$$
\begin{aligned}
|\cdot|_1 &=_{def} |d_1 - d'_1| + \ldots + |d_n - d'_n| && \text{(the 1-norm distance)}\\
|\cdot|_2 &=_{def} \sqrt{(d_1 - d'_1)^2 + \ldots + (d_n - d'_n)^2} && \text{(the 2-norm distance)}\\
\min &=_{def} \min\{|d_1 - d'_1|, \ldots, |d_n - d'_n|\} && \text{(minimal deviation)}\\
\max &=_{def} \max\{|d_1 - d'_1|, \ldots, |d_n - d'_n|\} && \text{(maximal deviation)}\\
avg &=_{def} \frac{|d_1 - d'_1| + \ldots + |d_n - d'_n|}{n} && \text{(average deviation)}\\
\# &=_{def} \text{number of modified values}
\end{aligned}
$$

All appearing numbers are rounded off to three decimal places.

Inc	$\mathsf{Inc}(\mathcal{R}_0)$
Inc^d	1
$\mathsf{Inc}^{\mathsf{MI}}$	1
$\mathsf{Inc}^{\mathsf{MI}}_0$	0.5
$\mathsf{Inc}^{\mathsf{MI}}_C$	0.5
$\mathsf{Inc}^{\mathsf{MI}}_{C,0}$	0.5
Inc^*	0.4
Inc^*_0	0.2
$\mathcal{I}^{\leq}$	0.4
$\mathcal{I}^{\leq}_0$	0.2
$\mathcal{I}^{\geq}$	0.4
$\mathcal{I}^{\geq}_0$	0.2

$C^{\mathcal{R}_0}$	$C(r_1)$	$C(r_2)$
$\mathsf{SignCulp}^{\mathcal{R}_0}$	$+1$	-1
$A^{\mathcal{R}_0}$	0.2	0.2
$S^{\mathcal{R}_0}_{\mathsf{Inc}^d}$	0.5	0.5
$S^{\mathcal{R}_0}_{\mathsf{Inc}^{\mathsf{MI}}}$	0.5	0.5
$S^{\mathcal{R}_0}_{\mathsf{Inc}^{\mathsf{MI}}_0}$	0.25	0.25
$S^{\mathcal{R}_0}_{\mathsf{Inc}^{\mathsf{MI}}_C}$	0.25	0.25
$S^{\mathcal{R}_0}_{\mathsf{Inc}^{\mathsf{MI}}_{C,0}}$	0.25	0.25
$S^{\mathcal{R}_0}_{\mathsf{Inc}^*}$	0.2	0.2
$S^{\mathcal{R}_0}_{\mathsf{Inc}^*_0}$	0.1	0.1
$S^{\mathcal{R}_0}_{\mathcal{I}^{\leq}}$	0.2	0.2
$S^{\mathcal{R}_0}_{\mathcal{I}^{\leq}_0}$	0.1	0.1
$S^{\mathcal{R}_0}_{\mathcal{I}^{\geq}}$	0.2	0.2
$S^{\mathcal{R}_0}_{\mathcal{I}^{\geq}_0}$	0.1	0.1

Table 10: Inconsistency and culpability values for $\mathcal{R}_0$

Y	r_1'	r_2'	$\lvert\cdot\rvert_1$	$\lvert\cdot\rvert_2$	min	max	*avg*	#
Υ^U	0.5	0.5	0.4	0.283	0.2	0.2	0.2	2
Υ^P	0.5	0.5	0.4	0.283	0.2	0.2	0.2	2
$\Upsilon^S_{S^{\mathcal{R}_0}_{\mathsf{Inc}_d}}$	0.5	0.5	0.4	0.283	0.2	0.2	0.2	2
$\Upsilon^S_{S^{\mathcal{R}_0}_{\mathsf{Inc}^{\mathsf{MI}}}}$	0.5	0.5	0.4	0.283	0.2	0.2	0.2	2
$\Upsilon^S_{S^{\mathcal{R}_0}_{\mathsf{Inc}^{\mathsf{MI}}_C}}$	0.5	0.5	0.4	0.283	0.2	0.2	0.2	2
$\Upsilon^S_{S^{\mathcal{R}_0}_{\mathsf{Inc}^*}}$	0.5	0.5	0.4	0.283	0.2	0.2	0.2	2
$\Upsilon^S_{S^{\mathcal{R}_0}_{\mathcal{I}^{\leq}}}$	0.5	0.5	0.4	0.283	0.2	0.2	0.2	2
$\Upsilon^S_{S^{\mathcal{R}_0}_{\mathcal{I}^{\geq}}}$	0.5	0.5	0.4	0.283	0.2	0.2	0.2	2
$\Upsilon^S_{A^{\mathcal{R}_0}}$	0.5	0.5	0.4	0.283	0.2	0.2	0.2	2
$\Upsilon^B_{S^{\mathcal{R}_0}_{\mathsf{Inc}_d}}$	0.5	0.5	0.4	0.283	0.2	0.2	0.2	2
$\Upsilon^B_{S^{\mathcal{R}_0}_{\mathsf{Inc}^{\mathsf{MI}}}}$	0.5	0.5	0.4	0.283	0.2	0.2	0.2	2
$\Upsilon^B_{S^{\mathcal{R}_0}_{\mathsf{Inc}^{\mathsf{MI}}_C}}$	0.5	0.5	0.4	0.283	0.2	0.2	0.2	2
$\Upsilon^B_{S^{\mathcal{R}_0}_{\mathsf{Inc}^*}}$	0.5	0.5	0.4	0.283	0.2	0.2	0.2	2
$\Upsilon^B_{S^{\mathcal{R}_0}_{\mathcal{I}^{\leq}}}$	0.5	0.5	0.4	0.283	0.2	0.2	0.2	2
$\Upsilon^B_{S^{\mathcal{R}_0}_{\mathcal{I}^{\geq}}}$	0.5	0.5	0.4	0.283	0.2	0.2	0.2	2
$\Upsilon^B_{A^{\mathcal{R}_0}}$	0.5	0.5	0.4	0.283	0.2	0.2	0.2	2

Table 11: Repaired knowledge bases for $\mathcal{R}_0$

Inc	$\mathsf{Inc}(\mathcal{R}_1)$
Inc^d	1
$\mathsf{Inc}^{\mathsf{MI}}$	1
$\mathsf{Inc}_0^{\mathsf{MI}}$	0.333
$\mathsf{Inc}_C^{\mathsf{MI}}$	0.333
$\mathsf{Inc}_{C,0}^{\mathsf{MI}}$	0.222
Inc^*	1
Inc_0^*	0.333
$\mathcal{I}^{\leq}$	1
$\mathcal{I}_0^{\leq}$	0.333
$\mathcal{I}^{\geq}$	1.168
$\mathcal{I}_0^{\geq}$	0.389

$C^{\mathcal{R}_1}$	$C^{(}r_1)$	$C(r_2)$	$C(r_3)$
$\mathsf{SignCulp}^{\mathcal{R}}$	-1	-1	$+1$
A^R	0.5	0.5	0.5
$S^R_{\mathsf{Inc}^d}$	0.333	0.333	0.333
$S^R_{\mathsf{Inc}^{\mathsf{MI}}}$	0.333	0.333	0.333
$S^R_{\mathsf{Inc}_0^{\mathsf{MI}}}$	0.111	0.111	0.111
$S^R_{\mathsf{Inc}_C^{\mathsf{MI}}}$	0.111	0.111	0.111
$S^R_{\mathsf{Inc}_{C,0}^{\mathsf{MI}}}$	0.074	0.074	0.074
$S^R_{\mathsf{Inc}^*}$	0.333	0.333	0.333
$S^R_{\mathsf{Inc}_0^*}$	0.111	0.111	0.111
$S^R_{\mathcal{I}^{\leq}}$	0.333	0.333	0.333
$S^R_{\mathcal{I}_0^{\leq}}$	0.111	0.111	0.111
$S^R_{\mathcal{I}^{\geq}}$	0.389	0.389	0.389
$S^R_{\mathcal{I}_0^{\geq}}$	0.13	0.13	0.13

Table 12: Inconsistency and culpability values for $\mathcal{R}_1$

Y	r_1'	r_2'	r_3'	$\lvert\cdot\rvert_1$	$\lvert\cdot\rvert_2$	min	max	*avg*	#
Y^U	0.618	0.618	0.382	1.146	0.662	0.382	0.382	0.382	3
Y^P	0.618	0.618	0.382	1.146	0.662	0.382	0.382	0.382	3
$Y^S_{S^{\mathcal{R}_1}_{\mathsf{Inc}_d}}$	0.618	0.618	0.382	1.146	0.662	0.382	0.382	0.382	3
$Y^S_{S^{\mathcal{R}_1}_{\mathsf{Inc}^{\mathsf{MI}}}}$	0.618	0.618	0.382	1.146	0.662	0.382	0.382	0.382	3
$Y^S_{S^{\mathcal{R}_1}_{\mathsf{Inc}_C^{\mathsf{MI}}}}$	0.618	0.618	0.382	1.146	0.662	0.382	0.382	0.382	3
$Y^S_{S^{\mathcal{R}_1}_{\mathsf{Inc}^*}}$	0.618	0.618	0.382	1.146	0.662	0.382	0.382	0.382	3
$Y^S_{S^{\mathcal{R}_1}_{\mathcal{I}^{\leq}}}$	0.618	0.618	0.382	1.146	0.662	0.382	0.382	0.382	3
$Y^S_{S^{\mathcal{R}_1}_{\mathcal{I}^{\geq}}}$	0.618	0.618	0.382	1.146	0.662	0.382	0.382	0.382	3
$Y^S_{A^{\mathcal{R}_1}}$	0.618	0.618	0.382	1.146	0.662	0.382	0.382	0.382	3
$Y^B_{S^{\mathcal{R}_1}_{\mathsf{Inc}_d}}$	0.618	0.618	0.382	1.146	0.662	0.382	0.382	0.382	3
$Y^B_{S^{\mathcal{R}_1}_{\mathsf{Inc}^{\mathsf{MI}}}}$	0.618	0.618	0.382	1.146	0.662	0.382	0.382	0.382	3
$Y^B_{S^{\mathcal{R}_1}_{\mathsf{Inc}_C^{\mathsf{MI}}}}$	0.618	0.618	0.382	1.146	0.662	0.382	0.382	0.382	3
$Y^B_{S^{\mathcal{R}_1}_{\mathsf{Inc}^*}}$	0.618	0.618	0.382	1.146	0.662	0.382	0.382	0.382	3
$Y^B_{S^{\mathcal{R}_1}_{\mathcal{I}^{\leq}}}$	0.618	0.618	0.382	1.146	0.662	0.382	0.382	0.382	3
$Y^B_{S^{\mathcal{R}_1}_{\mathcal{I}^{\geq}}}$	0.618	0.618	0.382	1.146	0.662	0.382	0.382	0.382	3
$Y^B_{A^{\mathcal{R}_1}}$	0.618	0.618	0.382	1.146	0.662	0.382	0.382	0.382	3

Table 13: Repaired knowledge bases for $\mathcal{R}_1$

Inc	$\mathsf{Inc}(\mathcal{R}_2)$
Inc^d	1
$\mathsf{Inc}^{\mathsf{MI}}$	1
$\mathsf{Inc}_0^{\mathsf{MI}}$	0.333
$\mathsf{Inc}_C^{\mathsf{MI}}$	0.5
$\mathsf{Inc}_{C,0}^{\mathsf{MI}}$	0.333
Inc^*	0.4
Inc_0^*	0.133
$\mathcal{I}^{\leq}$	0.4
$\mathcal{I}_0^{\leq}$	0.133
$\mathcal{I}^{\geq}$	0.4
$\mathcal{I}_0^{\geq}$	0.133

$C^{\mathcal{R}_2}$	$C(r_1)$	$C(r_2)$	$C(r_3)$
$\mathsf{SignCulp}^{\mathcal{R}}$	+1	−1	0
A^R	0.2	0.2	0
$S^R_{\mathsf{Inc}^d}$	0.5	0.5	0
$S^R_{\mathsf{Inc}^{\mathsf{MI}}}$	0.5	0.5	0
$S^R_{\mathsf{Inc}_0^{\mathsf{MI}}}$	0.2	0.2	0
$S^R_{\mathsf{Inc}_C^{\mathsf{MI}}}$	0.25	0.25	0
$S^R_{\mathsf{Inc}_{C,0}^{\mathsf{MI}}}$	0.2	0.2	0
$S^R_{\mathsf{Inc}^*}$	0.2	0.2	0
$S^R_{\mathsf{Inc}_0^*}$	0.067	0.067	0
$S^R_{\mathcal{I}^{\leq}}$	0.2	0.2	0
$S^R_{\mathcal{I}_0^{\leq}}$	0.067	0.067	0
$S^R_{\mathcal{I}^{\geq}}$	0.2	0.2	0
$S^R_{\mathcal{I}_0^{\geq}}$	0.067	0.067	0

Table 14: Inconsistency and culpability values for $\mathcal{R}_2$

Y	r_1'	r_2'	r_3'	$\lvert\cdot\rvert_1$	$\lvert\cdot\rvert_2$	min	max	*avg*	#
Y^U	0.5	0.5	0.5	0.7	0.327	0.2	0.3	0.233	3
Y^P	0.5	0.5	0.8	0.4	0.283	0	0.2	0.133	2
$Y^S_{S^{\mathcal{R}_2}_{\mathsf{Inc}_d}}$	0.5	0.5	0.8	0.4	0.283	0	0.2	0.133	2
$Y^S_{S^{\mathcal{R}_2}_{\mathsf{Inc}^{\mathsf{MI}}}}$	0.5	0.5	0.8	0.4	0.283	0	0.2	0.133	2
$Y^S_{S^{\mathcal{R}_2}_{\mathsf{Inc}_C^{\mathsf{MI}}}}$	0.5	0.5	0.8	0.4	0.283	0	0.2	0.133	2
$Y^S_{S^{\mathcal{R}_2}_{\mathsf{Inc}^*}}$	0.5	0.5	0.8	0.4	0.283	0	0.2	0.133	2
$Y^S_{S^{\mathcal{R}_2}_{\mathcal{I}^{\leq}}}$	0.5	0.5	0.8	0.4	0.283	0	0.2	0.133	2
$Y^S_{S^{\mathcal{R}_2}_{\mathcal{I}^{\geq}}}$	0.5	0.5	0.8	0.4	0.283	0	0.2	0.133	2
$Y^S_{A^{\mathcal{R}_2}}$	0.5	0.5	0.8	0.4	0.283	0	0.2	0.133	2
$Y^B_{S^{\mathcal{R}_2}_{\mathsf{Inc}_d}}$	0.5	0.5	0.8	0.4	0.283	0	0.2	0.133	2
$Y^B_{S^{\mathcal{R}_2}_{\mathsf{Inc}^{\mathsf{MI}}}}$	0.5	0.5	0.8	0.4	0.283	0	0.2	0.133	2
$Y^B_{S^{\mathcal{R}_2}_{\mathsf{Inc}_C^{\mathsf{MI}}}}$	0.5	0.5	0.8	0.4	0.283	0	0.2	0.133	2
$Y^B_{S^{\mathcal{R}_2}_{\mathsf{Inc}^*}}$	0.5	0.5	0.8	0.4	0.283	0	0.2	0.133	2
$Y^B_{S^{\mathcal{R}_2}_{\mathcal{I}^{\leq}}}$	0.5	0.5	0.8	0.4	0.283	0	0.2	0.133	2
$Y^B_{S^{\mathcal{R}_2}_{\mathcal{I}^{\geq}}}$	0.5	0.5	0.8	0.4	0.283	0	0.2	0.133	2
$Y^B_{A^{\mathcal{R}_2}}$	0.5	0.5	0.8	0.4	0.283	0	0.2	0.133	2

Table 15: Repaired knowledge bases for $\mathcal{R}_2$

Inc	$\mathsf{Inc}(\mathcal{R}_3)$
Inc^d	1
$\mathsf{Inc}^{\mathsf{MI}}$	2
$\mathsf{Inc}_0^{\mathsf{MI}}$	0.667
$\mathsf{Inc}_C^{\mathsf{MI}}$	1
$\mathsf{Inc}_{C,0}^{\mathsf{MI}}$	0.667
Inc^*	0.3
Inc_0^*	0.1
$\mathcal{I}^{\leq}$	0.3
$\mathcal{I}_0^{\leq}$	0.133
$\mathcal{I}^{\geq}$	0.3
$\mathcal{I}_0^{\geq}$	0.133

$C^{\mathcal{R}_3}$	$C(r_1)$	$C(r_2)$	$C(r_3)$
$\mathsf{SignCulp}^{\mathcal{R}}$	$+1$	0	-1
A^R	0.05	0	0.25
$S^R_{\mathsf{Inc}^d}$	0.167	0.167	0.667
$S^R_{\mathsf{Inc}^{\mathsf{MI}}}$	0.5	0.5	1
$S^R_{\mathsf{Inc}_0^{\mathsf{MI}}}$	0.139	0.139	0.389
$S^R_{\mathsf{Inc}_C^{\mathsf{MI}}}$	0.25	0.25	0.5
$S^R_{\mathsf{Inc}_{C,0}^{\mathsf{MI}}}$	0.139	0.139	0.389
$S^R_{\mathsf{Inc}^*}$	0.083	0.033	0.183
$S^R_{\mathsf{Inc}_0^*}$	0.028	0.011	0.061
$S^R_{\mathcal{I}^{\leq}}$	0.083	0.033	0.183
$S^R_{\mathcal{I}_0^{\leq}}$	0.028	0.011	0.061
$S^R_{\mathcal{I}^{\geq}}$	0.084	0.034	0.184
$S^R_{\mathcal{I}_0^{\geq}}$	0.028	0.011	0.061

Table 16: Inconsistency and culpability values for $\mathcal{R}_3$

Y	r_1'	r_2'	r_3'	$\lvert\cdot\rvert_1$	$\lvert\cdot\rvert_2$	min	max	avg	#
Y^U	0.41	0.455	0.41	0.355	0.226	0.055	0.19	0.118	3
Y^P	0.35	0.4	0.35	0.3	0.254	0	0.25	0.134	2
$Y^S_{S^{\mathcal{R}_3}_{\mathsf{Inc}_d}}$	0.36	0.46	0.36	0.36	0.255	0.06	0.24	0.12	3
$Y^S_{S^{\mathcal{R}_3}_{\mathsf{Inc}^{\mathsf{MI}}}}$	0.4	0.5	0.4	0.4	0.245	0.1	0.2	0.133	3
$Y^S_{S^{\mathcal{R}_3}_{\mathsf{Inc}_C^{\mathsf{MI}}}}$	0.4	0.5	0.4	0.4	0.245	0.1	0.2	0.133	3
$Y^S_{S^{\mathcal{R}_3}_{\mathsf{Inc}^*}}$	0.394	0.437	0.394	0.337	0.229	0.037	0.206	0.112	3
$Y^S_{S^{\mathcal{R}_3}_{\mathcal{I}^{\leq}}}$	0.394	0.437	0.394	0.337	0.229	0.037	0.206	0.112	3
$Y^S_{S^{\mathcal{R}_3}_{\mathcal{I}^{\geq}}}$	0.394	0.437	0.394	0.337	0.229	0.037	0.206	0.112	3
$Y^S_{A^{\mathcal{R}_3}}$	0.35	0.4	0.35	0.3	0.254	0	0.249	0.134	2
$Y^B_{S^{\mathcal{R}_3}_{\mathsf{Inc}_d}}$	0.3	0.4	0.3	0.3	0.3	0	0.3	0.1	1
$Y^B_{S^{\mathcal{R}_3}_{\mathsf{Inc}^{\mathsf{MI}}}}$	0.3	0.4	0.3	0.3	0.3	0	0.3	0.1	1
$Y^B_{S^{\mathcal{R}_3}_{\mathsf{Inc}_C^{\mathsf{MI}}}}$	0.3	0.4	0.3	0.3	0.3	0	0.3	0.1	1
$Y^B_{S^{\mathcal{R}_3}_{\mathsf{Inc}^*}}$	–	–	–	–	–	–	–	–	
$Y^B_{S^{\mathcal{R}_3}_{\mathcal{I}^{\leq}}}$	–	–	–	–	–	–	–	–	
$Y^B_{S^{\mathcal{R}_3}_{\mathcal{I}^{\geq}}}$	–	–	–	–	–	–	–	–	
$Y^B_{A^{\mathcal{R}_3}}$	–	–	–	–	–	–	–	–	

Table 17: Repaired knowledge bases for $\mathcal{R}_3$

Inc	$\mathsf{Inc}(\mathcal{R}_4)$
Inc^d	1
$\mathsf{Inc}^{\mathsf{MI}}$	3
$\mathsf{Inc}^{\mathsf{MI}}_0$	0.5
$\mathsf{Inc}^{\mathsf{MI}}_C$	1
$\mathsf{Inc}^{\mathsf{MI}}_{C,0}$	0.333
Inc^*	0.5
Inc^*_0	0.125
$\mathcal{I}^{\leq}$	0.5
$\mathcal{I}^{\leq}_0$	0.125
$\mathcal{I}^{\geq}$	0.8
$\mathcal{I}^{\geq}_0$	0.2

$C^{\mathcal{R}_4}$	$C(r_1)$	$C(r_2)$	$C(r_3)$	$C(r_4)$
$\mathsf{SignCulp}^{\mathcal{R}}$	0	0	0	+1
A^R	0	0	0	0.5
$S^R_{\mathsf{Inc}^d}$	0.167	0.167	0.167	0.5
$S^R_{\mathsf{Inc}^{\mathsf{MI}}}$	0.667	0.667	0.667	1.0
$S^R_{\mathsf{Inc}^{\mathsf{MI}}_0}$	0.097	0.097	0.097	0.208
$S^R_{\mathsf{Inc}^{\mathsf{MI}}_C}$	0.222	0.222	0.222	0.333
$S^R_{\mathsf{Inc}^{\mathsf{MI}}_{C,0}}$	0.065	0.065	0.065	0.14
$S^R_{\mathsf{Inc}^*}$	0.149	0.116	0.05	0.183
$S^R_{\mathsf{Inc}^*_0}$	0.04	0.029	0.013	0.046
$S^R_{\mathcal{I}^{\leq}}$	0.149	0.116	0.05	0.183
$S^R_{\mathcal{I}^{\leq}_0}$	0.04	0.029	0.013	0.046
$S^R_{\mathcal{I}^{\geq}}$	0.238	0.186	0.08	0.293
$S^R_{\mathcal{I}^{\geq}_0}$	0.06	0.047	0.02	0.073

Table 18: Inconsistency and culpability values for $\mathcal{R}_4$

Y	r_1'	r_2'	r_3'	r_4'	$\lvert\cdot\rvert_1$	$\lvert\cdot\rvert_2$	min	max	*avg*	#
Y^U	0.5	0.5	0.5	0.5	0.7	0.436	0	0.3	0.175	3
Y^P	0.8	0.6	0.5	0.7	0.5	0.5	0	0.5	0.125	1
$\mathrm{Y}^S_{S^{\mathcal{R}_4}_{\mathsf{Inc}_d}}$	0.5	0.5	0.5	0.5	0.7	0.436	0	0.3	0.175	3
$\mathrm{Y}^S_{S^{\mathcal{R}_4}_{\mathsf{Inc}^{\mathsf{MI}}}}$	0.5	0.5	0.5	0.5	0.7	0.436	0	0.3	0.175	3
$\mathrm{Y}^S_{S^{\mathcal{R}_4}_{\mathsf{Inc}^{\mathsf{MI}}_C}}$	0.5	0.5	0.5	0.5	0.7	0.436	0	0.3	0.175	3
$\mathrm{Y}^S_{S^{\mathcal{R}_4}_{\mathsf{Inc}^*}}$	0.5	0.5	0.5	0.5	0.7	0.436	0	0.3	0.175	3
$\mathrm{Y}^S_{S^{\mathcal{R}_4}_{\mathcal{I}^{\leq}}}$	0.5	0.5	0.5	0.5	0.7	0.436	0	0.3	0.175	3
$\mathrm{Y}^S_{S^{\mathcal{R}_4}_{\mathcal{I}^{\geq}}}$	0.5	0.5	0.5	0.5	0.7	0.436	0	0.3	0.175	3
$\mathrm{Y}^S_{A^{\mathcal{R}_4}}$	0.5	0.5	0.5	0.5	0.7	0.436	0	0.3	0.175	3
$\mathrm{Y}^B_{S^{\mathcal{R}_4}_{\mathsf{Inc}_d}}$	0.8	0.6	0.5	0.7	0.5	0.5	0	0.5	0.125	1
$\mathrm{Y}^B_{S^{\mathcal{R}_4}_{\mathsf{Inc}^{\mathsf{MI}}}}$	0.8	0.6	0.5	0.7	0.5	0.5	0	0.5	0.125	1
$\mathrm{Y}^B_{S^{\mathcal{R}_4}_{\mathsf{Inc}^{\mathsf{MI}}_C}}$	0.8	0.6	0.5	0.7	0.5	0.5	0	0.5	0.125	1
$\mathrm{Y}^B_{S^{\mathcal{R}_4}_{\mathsf{Inc}^*}}$	0.8	0.6	0.5	0.7	0.5	0.5	0	0.5	0.125	1
$\mathrm{Y}^B_{S^{\mathcal{R}_4}_{\mathcal{I}^{\leq}}}$	0.8	0.6	0.5	0.7	0.5	0.5	0	0.5	0.125	1
$\mathrm{Y}^B_{S^{\mathcal{R}_4}_{\mathcal{I}^{\geq}}}$	0.8	0.6	0.5	0.7	0.5	0.5	0	0.5	0.125	1
$\mathrm{Y}^B_{A^{\mathcal{R}_4}}$	0.8	0.6	0.5	0.7	0.5	0.5	0	0.5	0.125	1

Table 19: Repaired knowledge bases for $\mathcal{R}_4$

Inc	$\mathsf{Inc}(\mathcal{R}_5)$
Inc^d	1
$\mathsf{Inc}^{\mathsf{MI}}$	1
$\mathsf{Inc}_0^{\mathsf{MI}}$	0.167
$\mathsf{Inc}_C^{\mathsf{MI}}$	0.25
$\mathsf{Inc}_{C,0}^{\mathsf{MI}}$	0.083
Inc^*	0.278
Inc_0^*	0.069
$\mathcal{I}^{\leq}$	0.278
$\mathcal{I}_0^{\leq}$	0.069
$\mathcal{I}^{\geq}$	0.355
$\mathcal{I}_0^{\geq}$	0.089

$C^{\mathcal{R}_5}$	$C(r_1)$	$C(r_2)$	$C(r_3)$	$C(r_4)$
$\mathsf{SignCulp}^{\mathcal{R}}$	0	0	0	+1
A^R	0	0	0	0.278
$S^R_{\mathsf{Inc}^d}$	0.25	0.25	0.25	0.25
$S^R_{\mathsf{Inc}^{\mathsf{MI}}}$	0.25	0.25	0.25	0.25
$S^R_{\mathsf{Inc}_0^{\mathsf{MI}}}$	0.042	0.042	0.042	0.042
$S^R_{\mathsf{Inc}_C^{\mathsf{MI}}}$	0.063	0.063	0.063	0.063
$S^R_{\mathsf{Inc}_{C,0}^{\mathsf{MI}}}$	0.021	0.021	0.021	0.021
$S^R_{\mathsf{Inc}^*}$	0.07	0.07	0.07	0.07
$S^R_{\mathsf{Inc}_0^*}$	0.018	0.018	0.018	0.018
$S^R_{\mathcal{I}^{\leq}}$	0.07	0.07	0.07	0.07
$S^R_{\mathcal{I}_0^{\leq}}$	0.018	0.018	0.018	0.018
$S^R_{\mathcal{I}^{\geq}}$	0.089	0.089	0.089	0.089
$S^R_{\mathcal{I}_0^{\geq}}$	0.022	0.022	0.022	0.022

Table 20: Inconsistency and culpability values for $\mathcal{R}_5$

Y	r_1'	r_2'	r_3'	r_4'	$\lvert\cdot\rvert_1$	$\lvert\cdot\rvert_2$	min	max	*avg*	#
Y^U	0.621	0.56	0.742	0.258	0.435	0.24	0.04	0.158	0.109	4
Y^P	0.7	0.6	0.9	0.378	0.278	0.077	0	0.278	0.07	1
$\mathrm{Y}^S_{S^{\mathcal{R}_5}_{\mathsf{Inc}_d}}$	0.576	0.5	0.776	0.224	0.472	0.237	0.1	0.124	0.118	4
$\mathrm{Y}^S_{S^{\mathcal{R}_5}_{\mathsf{Inc}^{\mathsf{MI}}}}$	0.576	0.5	0.776	0.224	0.472	0.237	0.1	0.124	0.118	4
$\mathrm{Y}^S_{S^{\mathcal{R}_5}_{\mathsf{Inc}_C^{\mathsf{MI}}}}$	0.576	0.5	0.776	0.224	0.472	0.237	0.1	0.124	0.118	4
$\mathrm{Y}^S_{S^{\mathcal{R}_5}_{\mathsf{Inc}^*}}$	0.576	0.5	0.776	0.224	0.472	0.237	0.1	0.124	0.118	4
$\mathrm{Y}^S_{S^{\mathcal{R}_5}_{\mathcal{I}^{\leq}}}$	0.576	0.5	0.776	0.224	0.472	0.237	0.1	0.124	0.118	4
$\mathrm{Y}^S_{S^{\mathcal{R}_5}_{\mathcal{I}^{\geq}}}$	0.576	0.5	0.776	0.224	0.472	0.237	0.1	0.124	0.118	4
$\mathrm{Y}^S_{A^{\mathcal{R}_5}}$	0.7	0.6	0.9	0.378	0.278	0.077	0	0.278	0.07	1
$\mathrm{Y}^B_{S^{\mathcal{R}_5}_{\mathsf{Inc}_d}}$	0.581	0.481	0.781	0.218	0.714	0.238	0.119	0.119	0.119	4
$\mathrm{Y}^B_{S^{\mathcal{R}_5}_{\mathsf{Inc}^{\mathsf{MI}}}}$	0.581	0.481	0.781	0.218	0.714	0.238	0.119	0.119	0.119	4
$\mathrm{Y}^B_{S^{\mathcal{R}_5}_{\mathsf{Inc}_C^{\mathsf{MI}}}}$	0.581	0.481	0.781	0.218	0.714	0.238	0.119	0.119	0.119	4
$\mathrm{Y}^B_{S^{\mathcal{R}_5}_{\mathsf{Inc}^*}}$	0.581	0.481	0.781	0.218	0.714	0.238	0.119	0.119	0.119	4
$\mathrm{Y}^B_{S^{\mathcal{R}_5}_{\mathcal{I}^{\leq}}}$	0.581	0.481	0.781	0.218	0.714	0.238	0.119	0.119	0.119	4
$\mathrm{Y}^B_{S^{\mathcal{R}_5}_{\mathcal{I}^{\geq}}}$	0.581	0.481	0.781	0.218	0.714	0.238	0.119	0.119	0.119	4
$\mathrm{Y}^B_{A^{\mathcal{R}_5}}$	0.7	0.6	0.9	0.378	0.278	0.077	0	0.278	0.07	1

Table 21: Repaired knowledge bases for $\mathcal{R}_5$

Inc	$\mathsf{Inc}(\mathcal{R}_6)$
Inc^d	1
$\mathsf{Inc}^{\mathsf{MI}}$	3
$\mathsf{Inc}^{\mathsf{MI}}_0$	0.3
$\mathsf{Inc}^{\mathsf{MI}}_C$	0.917
$\mathsf{Inc}^{\mathsf{MI}}_{C,0}$	0.183
Inc^*	0.407
Inc^*_0	0.082
$\mathcal{I}^{\leq}$	0.407
$\mathcal{I}^{\leq}_0$	0.082
$\mathcal{I}^{\geq}$	0.463
$\mathcal{I}^{\geq}_0$	0.093

$C^{\mathcal{R}_6}$	$C(r_1)$	$C(r_2)$	$C(r_3)$	$C(r_4)$	$C(r_5)$
$\mathsf{SignCulp}^{\mathcal{R}}$	0	0	0	$+1$	-1
A^R	0	0	0	0.335	0.072
$S^R_{\mathsf{Inc}^d}$	0.183	0.183	0.183	0.183	0.267
$S^R_{\mathsf{Inc}^{\mathsf{MI}}}$	0.583	0.583	0.583	0.583	0.667
$S^R_{\mathsf{Inc}^{\mathsf{MI}}_0}$	0.054	0.054	0.054	0.054	0.082
$S^R_{\mathsf{Inc}^{\mathsf{MI}}_C}$	0.174	0.174	0.174	0.174	0.222
$S^R_{\mathsf{Inc}^{\mathsf{MI}}_{C,0}}$	0.032	0.032	0.032	0.032	0.057
$S^R_{\mathsf{Inc}^*}$	0.028	0.132	0.029	0.132	0.086
$S^R_{\mathsf{Inc}^*_0}$	0.006	0.026	0.006	0.026	0.017
$S^R_{\mathcal{I}^{\leq}}$	0.029	0.132	0.029	0.132	0.086
$S^R_{\mathcal{I}^{\leq}_0}$	0.006	0.026	0.006	0.026	0.017
$S^R_{\mathcal{I}^{\geq}}$	0.035	0.153	0.035	0.153	0.088
$S^R_{\mathcal{I}^{\geq}_0}$	0.007	0.031	0.007	0.031	0.018

Table 22: Inconsistency and culpability values for $\mathcal{R}_6$

Y	r'_1	r'_2	r'_3	r'_4	r'_5	$\lvert\cdot\rvert_1$	$\lvert\cdot\rvert_2$	min	max	*avg*	#
Y^U	0.582	0.541	0.663	0.337	0.623	0.828	0.401	0.059	0.237	0.166	5
Y^P	0.7	0.6	0.9	0.435	0.728	0.407	0.343	0	0.335	0.081	2
$Y^S_{S^{\mathcal{R}_6}_{\mathsf{Inc}_d}}$	0.526	0.5	0.726	0.274	0.547	0.875	0.406	0.1	0.253	0.175	5
$Y^S_{S^{\mathcal{R}_6}_{\mathsf{Inc}^{\mathsf{MI}}}}$	0.567	0.5	0.767	0.3	0.6	0.766	0.354	0.1	0.2	0.153	5
$Y^S_{S^{\mathcal{R}_6}_{\mathsf{Inc}^{\mathsf{MI}}_C}}$	0.517	0.5	0.717	0.283	0.566	0.883	0.406	0.1	0.234	0.177	5
$Y^S_{S^{\mathcal{R}_6}_{\mathsf{Inc}^*}}$	0.651	0.5	0.851	0.326	0.653	0.571	0.296	0.049	0.226	0.114	5
$Y^S_{S^{\mathcal{R}_6}_{\mathcal{I}^{\leq}}}$	0.651	0.5	0.851	0.326	0.653	0.571	0.296	0.049	0.226	0.114	5
$Y^S_{S^{\mathcal{R}_6}_{\mathcal{I}^{\geq}}}$	0.651	0.5	0.851	0.326	0.653	0.571	0.296	0.049	0.226	0.114	5
$Y^S_{A^{\mathcal{R}_6}}$	0.7	0.6	0.9	0.435	0.728	0.407	0.343	0	0.335	0.081	2
$Y^B_{S^{\mathcal{R}_6}_{\mathsf{Inc}_d}}$	0.651	0.45	0.851	0.25	0.554	0.644	0.332	0.049	0.246	0.129	5
$Y^B_{S^{\mathcal{R}_6}_{\mathsf{Inc}^{\mathsf{MI}}}}$	0.651	0.45	0.851	0.25	0.554	0.644	0.332	0.049	0.246	0.129	5
$Y^B_{S^{\mathcal{R}_6}_{\mathsf{Inc}^{\mathsf{MI}}_C}}$	0.651	0.45	0.851	0.25	0.554	0.644	0.332	0.049	0.246	0.129	5
$Y^B_{S^{\mathcal{R}_6}_{\mathsf{Inc}^*}}$	0.7	0.406	0.9	0.296	0.73	0.46	0.285	0	0.196	0.092	3
$Y^B_{S^{\mathcal{R}_6}_{\mathcal{I}^{\leq}}}$	0.7	0.406	0.9	0.296	0.73	0.46	0.285	0	0.196	0.092	3
$Y^B_{S^{\mathcal{R}_6}_{\mathcal{I}^{\geq}}}$	0.7	0.406	0.9	0.296	0.73	0.46	0.285	0	0.196	0.092	3
$Y^B_{A^{\mathcal{R}_6}}$	0.7	0.6	0.9	0.435	0.728	0.408	0.343	0	0.335	0.081	2

Table 23: Repaired knowledge bases for $\mathcal{R}_6$

Inc	$\mathsf{Inc}(\mathcal{R}_7)$
Inc^d	1
$\mathsf{Inc}^{\mathsf{MI}}$	2
$\mathsf{Inc}_0^{\mathsf{MI}}$	0.2
$\mathsf{Inc}_C^{\mathsf{MI}}$	0.667
$\mathsf{Inc}_{C,0}^{\mathsf{MI}}$	0.133
Inc^*	0.522
Inc_0^*	0.104
$\mathcal{I}^{\leq}$	0.522
$\mathcal{I}_0^{\leq}$	0.104
$\mathcal{I}^{\geq}$	1.147
$\mathcal{I}_0^{\geq}$	0.295

$C^{\mathcal{R}_7}$	$C(r_1)$	$C(r_2)$	$C(r_3)$	$C(r_4)$	$C(r_5)$
$\mathsf{SignCulp}^{\mathcal{R}}$	-1	0	0	0	$+1$
A^R	0.5	0	0	0	0.022
$S^R_{\mathsf{Inc}^d}$	0.417	0	0.083	0.083	0.417
$S^R_{\mathsf{Inc}^{\mathsf{MI}}}$	0.667	0	0.333	0.333	0.667
$S^R_{\mathsf{Inc}_0^{\mathsf{MI}}}$	0.096	0	0.026	0.026	0.096
$S^R_{\mathsf{Inc}_C^{\mathsf{MI}}}$	0.222	0	0.111	0.111	0.222
$S^R_{\mathsf{Inc}_{C,0}^{\mathsf{MI}}}$	0.064	0	0.017	0.017	0.064
$S^R_{\mathsf{Inc}^*}$	0.208	0	0.037	0.066	0.208
$S^R_{\mathsf{Inc}_0^*}$	0.042	0	0.007	0.013	0.042
$S^R_{\mathcal{I}^{\leq}}$	0.208	0	0.037	0.066	0.208
$S^R_{\mathcal{I}_0^{\leq}}$	0.045	0	0.007	0.013	0.042
$S^R_{\mathcal{I}^{\geq}}$	0.504	0.08	0.032	0.352	0.504
$S^R_{\mathcal{I}_0^{\geq}}$	0.1	0.016	0.006	0.07	0.1

Table 24: Inconsistency and culpability values for $\mathcal{R}_7$

Y	r_1'	r_2'	r_3'	r_4'	r_5'	$\lvert\cdot\rvert_1$	$\lvert\cdot\rvert_2$	min	max	avg	#
Y^U	0.25	0.5	0.5	0.5	0.5	1.95	0.879	0.3	0.45	0.39	5
Y^P	0.2	0.9	0.1	0.1	0.222	0.522	0.505	0	0.5	0.104	2
$Y^S_{S^{\mathcal{R}_7}_{\mathsf{Inc}_d}}$	0.25	0.9	0.5	0.5	0.5	1.55	0.783	0	0.45	0.31	4
$Y^S_{S^{\mathcal{R}_7}_{\mathsf{Inc}^{\mathsf{MI}}}}$	0.25	0.9	0.5	0.5	0.5	1.55	0.783	0	0.45	0.31	4
$Y^S_{S^{\mathcal{R}_7}_{\mathsf{Inc}_C^{\mathsf{MI}}}}$	0.25	0.9	0.5	0.5	0.5	1.55	0.783	0	0.45	0.31	4
$Y^S_{S^{\mathcal{R}_7}_{\mathsf{Inc}^*}}$	0.25	0.9	0.5	0.5	0.5	1.55	0.783	0	0.45	0.31	4
$Y^S_{S^{\mathcal{R}_7}_{\mathcal{I}^{\leq}}}$	0.25	0.9	0.5	0.5	0.5	1.55	0.783	0	0.45	0.31	4
$Y^S_{S^{\mathcal{R}_7}_{\mathcal{I}^{\geq}}}$	0.25	0.9	0.5	0.5	0.5	1.55	0.783	0	0.45	0.31	4
$Y^S_{A^{\mathcal{R}_7}}$	0.25	0.9	0.1	0.1	0.5	0.75	0.541	0	0.45	0.15	2
$Y^B_{S^{\mathcal{R}_7}_{\mathsf{Inc}_d}}$	0.402	0.9	0.1	0.1	0.462	0.56	0.397	0	0.298	0.112	2
$Y^B_{S^{\mathcal{R}_7}_{\mathsf{Inc}^{\mathsf{MI}}}}$	0.402	0.9	0.1	0.1	0.462	0.56	0.397	0	0.298	0.112	2
$Y^B_{S^{\mathcal{R}_7}_{\mathsf{Inc}_C^{\mathsf{MI}}}}$	0.402	0.9	0.1	0.1	0.462	0.56	0.397	0	0.298	0.112	2
$Y^B_{S^{\mathcal{R}_7}_{\mathsf{Inc}^*}}$	0.402	0.9	0.1	0.1	0.462	0.56	0.397	0	0.298	0.112	2
$Y^B_{S^{\mathcal{R}_7}_{\mathcal{I}^{\leq}}}$	0.402	0.9	0.1	0.1	0.462	0.56	0.397	0	0.298	0.112	2
$Y^B_{S^{\mathcal{R}_7}_{\mathcal{I}^{\geq}}}$	0.402	0.9	0.1	0.1	0.462	0.56	0.397	0	0.298	0.112	2
$Y^B_{A^{\mathcal{R}_7}}$	—	—	—	—	—	—	—	—	—	—	—

Table 25: Repaired knowledge bases for $\mathcal{R}_7$

Inc	$\mathsf{Inc}(\mathcal{R}_8)$
Inc^d	1
$\mathsf{Inc}^{\mathsf{MI}}$	2
$\mathsf{Inc}^{\mathsf{MI}}_0$	0.2
$\mathsf{Inc}^{\mathsf{MI}}_C$	0.833
$\mathsf{Inc}^{\mathsf{MI}}_{C,0}$	0.167
Inc^*	1.4
Inc^*_0	0.28
$\mathcal{I}^{\leq}$	1.4
$\mathcal{I}^{\leq}_0$	0.28
$\mathcal{I}^{\geq}$	1.567
$\mathcal{I}^{\geq}_0$	0.313

$C^{\mathcal{R}_8}$	$C(r_1)$	$C(r_2)$	$C(r_3)$	$C(r_4)$	$C(r_5)$
$\mathsf{SignCulp}^{\mathcal{R}}$	-1	-1	$+1$	$+1$	-1
A^R	0.5	0.5	0.5	0.2	0.2
$S^R_{\mathsf{Inc}^d}$	0.133	0.133	0.133	0.3	0.3
$S^R_{\mathsf{Inc}^{\mathsf{MI}}}$	0.333	0.333	0.333	0.5	0.5
$S^R_{\mathsf{Inc}^{\mathsf{MI}}_0}$	0.012	0.012	0.012	0.082	0.082
$S^R_{\mathsf{Inc}^{\mathsf{MI}}_C}$	0.111	0.111	0.111	0.25	0.25
$S^R_{\mathsf{Inc}^{\mathsf{MI}}_{C,0}}$	0.022	0.022	0.022	0.05	0.05
$S^R_{\mathsf{Inc}^*}$	0.333	0.333	0.333	0.2	0.2
$S^R_{\mathsf{Inc}^*_0}$	0.067	0.067	0.067	0.04	0.04
$S^R_{\mathcal{I}^{\leq}}$	0.333	0.333	0.333	0.2	0.2
$S^R_{\mathcal{I}^{\leq}_0}$	0.067	0.067	0.067	0.04	0.04
$S^R_{\mathcal{I}^{\geq}}$	0.373	0.373	0.373	0.045	0.045
$S^R_{\mathcal{I}^{\geq}_0}$	0.075	0.075	0.075	0.009	0.009

Table 26: Inconsistency and culpability values for $\mathcal{R}_8$

Y	r_1'	r_2'	r_3'	r_4'	r_5'	$\lvert\cdot\rvert_1$	$\lvert\cdot\rvert_2$	min	max	*avg*	#
Y^U	0.5	0.5	0.5	0.5	0.5	1.9	0.911	0.2	0.5	0.38	5
Y^P	0.5	0.5	0.5	0.5	0.5	1.9	0.911	0.2	0.5	0.38	5
$Y^S_{S^{\mathcal{R}_8}_{\mathsf{Inc}_d}}$	0.618	0.618	0.382	0.5	0.5	1.546	0.72	0.2	0.382	0.309	5
$Y^S_{S^{\mathcal{R}_8}_{\mathsf{Inc}^{\mathsf{MI}}}}$	0.618	0.618	0.382	0.5	0.5	1.546	0.72	0.2	0.382	0.309	5
$Y^S_{S^{\mathcal{R}_8}_{\mathsf{Inc}^{\mathsf{MI}}_C}}$	0.618	0.618	0.382	0.5	0.5	1.546	0.72	0.2	0.382	0.309	5
$Y^S_{S^{\mathcal{R}_8}_{\mathsf{Inc}^*}}$	0.618	0.618	0.382	0.5	0.5	1.546	0.72	0.2	0.382	0.309	5
$Y^S_{S^{\mathcal{R}_8}_{\mathcal{I}^{\leq}}}$	0.618	0.618	0.382	0.5	0.5	1.546	0.72	0.2	0.382	0.309	5
$Y^S_{S^{\mathcal{R}_8}_{\mathcal{I}^{\geq}}}$	0.618	0.618	0.382	0.5	0.5	1.546	0.72	0.2	0.382	0.309	5
$Y^S_{A^{\mathcal{R}_8}}$	0.5	0.5	0.5	0.5	0.5	1.9	0.911	0.2	0.5	0.38	5
$Y^B_{S^{\mathcal{R}_8}_{\mathsf{Inc}_d}}$	0.618	0.618	0.382	0.5	0.5	1.546	0.72	0.2	0.382	0.309	5
$Y^B_{S^{\mathcal{R}_8}_{\mathsf{Inc}^{\mathsf{MI}}}}$	0.618	0.618	0.382	0.5	0.5	1.546	0.72	0.2	0.382	0.309	5
$Y^B_{S^{\mathcal{R}_8}_{\mathsf{Inc}^{\mathsf{MI}}_C}}$	0.618	0.618	0.382	0.5	0.5	1.546	0.72	0.2	0.382	0.309	5
$Y^B_{S^{\mathcal{R}_8}_{\mathsf{Inc}^*}}$	0.618	0.618	0.382	0.5	0.5	1.546	0.72	0.2	0.382	0.309	5
$Y^B_{S^{\mathcal{R}_8}_{\mathcal{I}^{\leq}}}$	0.618	0.618	0.382	0.5	0.5	1.546	0.72	0.2	0.382	0.309	5
$Y^B_{S^{\mathcal{R}_8}_{\mathcal{I}^{\geq}}}$	0.618	0.618	0.382	0.5	0.5	1.546	0.72	0.2	0.382	0.309	5
$Y^B_{A^{\mathcal{R}_8}}$	0.618	0.618	0.382	0.5	0.5	1.546	0.72	0.2	0.382	0.309	5

Table 27: Repaired knowledge bases for $\mathcal{R}_8$

Inc	$\mathsf{Inc}(\mathcal{R}_9)$
Inc^d	1
$\mathsf{Inc}^{\mathsf{MI}}$	1
$\mathsf{Inc}^{\mathsf{MI}}_0$	0.05
$\mathsf{Inc}^{\mathsf{MI}}_C$	0.167
$\mathsf{Inc}^{\mathsf{MI}}_{C,0}$	0.017
Inc^*	0.355
Inc^*_0	0.059
$\mathcal{I}^{\leq}$	0.318
$\mathcal{I}^{\leq}_0$	0.053
$\mathcal{I}^{\geq}$	0.695
$\mathcal{I}^{\geq}_0$	0.116

$C^{\mathcal{R}_9}$	$C(r_1)$	$C(r_2)$	$C(r_3)$	$C(r_4)$	$C(r_5)$	$C(r_6)$
$\mathsf{SignCulp}^{\mathcal{R}}$	0	0	0	0	0	-1
A^R	0	0	0	0	0	0.355
$S^R_{\mathsf{Inc}^d}$	0.167	0.167	0.167	0.167	0.167	0.167
$S^R_{\mathsf{Inc}^{\mathsf{MI}}}$	0.167	0.167	0.167	0.167	0.167	0.167
$S^R_{\mathsf{Inc}^{\mathsf{MI}}_0}$	0.008	0.008	0.008	0.008	0.008	0.008
$S^R_{\mathsf{Inc}^{\mathsf{MI}}_C}$	0.028	0.028	0.028	0.028	0.028	0.028
$S^R_{\mathsf{Inc}^{\mathsf{MI}}_{C,0}}$	0.003	0.003	0.003	0.003	0.003	0.003
$S^R_{\mathsf{Inc}^*}$	0.059	0.059	0.059	0.059	0.059	0.059
$S^R_{\mathsf{Inc}^*_0}$	0.01	0.01	0.01	0.01	0.01	0.01
$S^R_{\mathcal{I}^{\leq}}$	0.053	0.053	0.053	0.053	0.053	0.053
$S^R_{\mathcal{I}^{\leq}_0}$	0.009	0.009	0.009	0.009	0.009	0.009
$S^R_{\mathcal{I}^{\geq}}$	0.116	0.116	0.116	0.116	0.116	0.116
$S^R_{\mathcal{I}^{\geq}_0}$	0.019	0.019	0.019	0.019	0.019	0.019

Table 28: Inconsistency and culpability values for $\mathcal{R}_9$

Y	r_1'	r_2'	r_3'	r_4'	r_5'	r_6'	$\lvert\cdot\rvert_1$	$\lvert\cdot\rvert_2$	min	max	avg	#
Y^U	0.684	0.784	0.584	0.784	0.216	0.784	0.678	0.277	0.113	0.113	0.113	6
Y^P	0.8	0.9	0.7	0.9	0.1	0.545	0.355	0.126	0	0.355	0.059	1
$\mathrm{Y}^S_{S^{\mathcal{R}_9}_{\mathsf{Inc}_d}}$	0.711	0.811	0.611	0.811	0.189	0.811	0.531	0.217	0.089	0.089	0.089	6
$\mathrm{Y}^S_{S^{\mathcal{R}_9}_{\mathsf{Inc}^{\mathsf{MI}}}}$	0.711	0.811	0.611	0.811	0.189	0.811	0.531	0.217	0.089	0.089	0.089	6
$\mathrm{Y}^S_{S^{\mathcal{R}_9}_{\mathsf{Inc}^{\mathsf{MI}}_C}}$	0.711	0.811	0.611	0.811	0.189	0.811	0.531	0.217	0.089	0.089	0.089	6
$\mathrm{Y}^S_{S^{\mathcal{R}_9}_{\mathsf{Inc}^*}}$	0.711	0.811	0.611	0.811	0.189	0.811	0.531	0.217	0.089	0.089	0.089	6
$\mathrm{Y}^S_{S^{\mathcal{R}_9}_{\mathcal{I}^{\leq}}}$	0.711	0.811	0.611	0.811	0.189	0.811	0.531	0.217	0.089	0.089	0.089	6
$\mathrm{Y}^S_{S^{\mathcal{R}_9}_{\mathcal{I}^{\geq}}}$	0.711	0.811	0.611	0.811	0.189	0.811	0.531	0.217	0.089	0.089	0.089	6
$\mathrm{Y}^S_{A^{\mathcal{R}_9}}$	0.8	0.9	0.7	0.9	0.1	0.545	0.355	0.126	0	0.355	0.059	1
$\mathrm{Y}^B_{S^{\mathcal{R}_9}_{\mathsf{Inc}_d}}$	0.711	0.811	0.611	0.811	0.189	0.811	0.531	0.217	0.089	0.089	0.089	6
$\mathrm{Y}^B_{S^{\mathcal{R}_9}_{\mathsf{Inc}^{\mathsf{MI}}}}$	0.711	0.811	0.611	0.811	0.189	0.811	0.531	0.217	0.089	0.089	0.089	6
$\mathrm{Y}^B_{S^{\mathcal{R}_9}_{\mathsf{Inc}^{\mathsf{MI}}_C}}$	0.711	0.811	0.611	0.811	0.189	0.811	0.531	0.217	0.089	0.089	0.089	6
$\mathrm{Y}^B_{S^{\mathcal{R}_9}_{\mathsf{Inc}^*}}$	0.711	0.811	0.611	0.811	0.189	0.811	0.531	0.217	0.089	0.089	0.089	6
$\mathrm{Y}^B_{S^{\mathcal{R}_9}_{\mathcal{I}^{\leq}}}$	0.711	0.811	0.611	0.811	0.189	0.811	0.531	0.217	0.089	0.089	0.089	6
$\mathrm{Y}^B_{S^{\mathcal{R}_9}_{\mathcal{I}^{\geq}}}$	0.711	0.811	0.611	0.811	0.189	0.811	0.531	0.217	0.089	0.089	0.089	6
$\mathrm{Y}^B_{A^{\mathcal{R}_9}}$	0.8	0.9	0.7	0.9	0.1	0.545	0.355	0.126	0	0.355	0.059	1

Table 29: Repaired knowledge bases for $\mathcal{R}_9$

LIST OF FIGURES

LIST OF TABLES

BIBLIOGRAPHY

André Ahuja and Wilhelm Rödder. Project Risk Management by a Probabilistic Expert System. In *Proceedings of the International Conference on Operations Research 2002*, pages 329–334, 2002. (Cited on page 2.)

Carlos E. Alchourrón, Peter Gärdenfors, and David Makinson. On the Logic of Theory Change: Partial Meet Contraction and Revision Functions. *Journal of Symbolic Logic*, 50(2):510–530, 1985. (Cited on pages 10 and 89.)

Kim Allan Andersen and Daniele Pretolani. Easy Cases of Probabilistic Satisfiability. *Annals of Mathematics and Artificial Intelligence*, 33(1):69–91, 2001. (Cited on pages 43 and 45.)

Grigoris Antoniou. A Tutorial on Default Logics. *ACM Computing Surveys*, 31(4):337–359, 1999. (Cited on pages 18, 43, and 89.)

Kenneth J. Arrow. A Difficulty in the Concept of Social Welfare. *Journal of Political Economy*, 58(4):328–346, 1950. (Cited on pages 103, 104, and 126.)

Franz Baader, Diego Calvanese, Deborah McGuinness, Daniele Nardi, and Peter Patel-Schneider, editors. *The Description Logic Handbook - Theory, Implementation and Applications*. Cambridge University Press, 2003. (Cited on pages 2 and 12.)

Franz Baader, Ian Horrocks, and Ulrike Sattler. Description Logics as Ontology Languages for the Semantic Web. In *Mechanizing Mathematical Reasoning*, volume 2605 of *Lecture Notes in Computer Science*, pages 228–248. Springer-Verlag, 2005. (Cited on page 2.)

Fahiem Bacchus, Adam J. Grove, Joseph Y. Halpern, and Daphne Koller. From Statistical Knowledge Bases to Degrees of Belief. *Artificial Intelligence*, 87(1–2):75–143, 1996. (Cited on pages 151, 154, 156, 157, 159, 161, 163, 167, 180, 184, 185, 186, 187, 188, 199, 286, and 290.)

Fahiem Bacchus. *Representing and Reasoning with Probabilistic Knowledge*. The MIT Press, 1990. (Cited on pages 151 and 180.)

John A. Beachy and William D. Blair. *Abstract Algebra*. Waveland Press, Inc., third edition, 2005. (Cited on pages 204 and 244.)

Christoph Beierle and Gabriele Kern-Isberner. *Methoden wissensbasierter Systeme: Grundlagen, Algorithmen, Anwendungen*. Vieweg+Teubner, fourth edition, 2008. (Cited on pages 1, 9, 21, 23, 25, 26, 27, and 31.)

Christoph Beierle, Marc Finthammer, Gabriele Kern-Isberner, and Matthias Thimm. Automated reasoning for relational probabilistic knowledge representation. In Jüergen Giesl and Rainer Hähnle, editors, *Proceedings of the Fifth International Joint Conference on Automated Reasoning (IJCAR'10)*, volume 6173 of *Lecture Notes in Artificial Intelligence*, pages 218–224. Springer-Verlag, 2010. (Cited on page v.)

Christoph Beierle, Bernhard Freund, Gabriele Kern-Isberner, and Matthias Thimm. Can Bobby Demand Delivery? Towards a Knowledge-based System for Private Law. In *Proceedings of the 24th Workshop on Constraint Logic Programming (WLP'10)*, Cairo, Egypt, 2010. (Cited on page vii.)

Christoph Beierle, Bernhard Freund, Gabriele Kern-Isberner, and Matthias Thimm. Using Defeasible Logic Programming for Argumentation-Based Decision Support in Private Law. In *Proceedings of the Third International Conference on Computational Models of Argument (COMMA'10)*, 2010. (Cited on page vii.)

Nuel D. Belnap. How Computer should Think. In Gilbert Ryle, editor, *Contemporary Aspects of Philosophy*, pages 30–56. Oriel Press, 1976. (Cited on pages 18 and 43.)

Nuel D. Belnap. A Useful Four-Valued Logic. In J. Michael Dunn and George Epstein, editors, *Modern Uses of Multiple-Valued Logic*, pages 7–37. Reidel Publishing Company, 1977. (Cited on pages 18 and 43.)

Trevor J. M. Bench-Capon and Paul E. Dunne. Argumentation in Artificial Intelligence. *Artificial Intelligence*, 171(10–15):619–641, 2007. (Cited on pages 2, 18, 43, and 89.)

Salem Benferhat, Didier Dubois, and Henri Prade. Possibilistic and Standard Probabilistic Semantics of Conditional Knowledge Bases. *Journal of Logic and Computation*, 9(6):873–895, 1999. (Cited on pages 3 and 26.)

Jakob Bernoulli. Usum & Applicationem Praecedentis Doctrinae in Civilibus, Moralibus & Oeconomicis. In *Ars Conjectandi*, chapter 4. Thurneysen Brothers, 1713. (Cited on page 166.)

Jean-Yves Béziau, Walter Carnielli, and Dov Gabbay, editors. *Handbook of Paraconsistency*. Studies in Logic. College Publications, 2007. (Cited on pages 18 and 43.)

Joachim Biskup, Gabriele Kern-Isberner, and Matthias Thimm. Towards Enforcement of Confidentiality in Agent Interactions. In Maurice Pagnucco and Michael Thielscher, editors, *Proceedings of the Twelfth International Workshop on Non-Monotonic Reasoning (NMR'08)*, pages 104–112. University of New South Wales, Technical Report No. UNSW-CSE-TR-0819, 2008. (Cited on page viii.)

Isabelle Bloch and Anthony Hunter. Fusion: General Concepts and Characteristics. *International Journal of Intelligent Systems*, 16(10):1107–1134, 2001. (Cited on pages 4, 89, and 103.)

Innocentius M. Bocheński. *A History of Formal Logic*. University of Notre Dame Press, 1961. Translated and edited by Ivo Thomas. (Cited on page 10.)

George Boole. *An Investigation of the Laws of Thought on Which are Founded the Mathematical Theories of Logic and Probabilities*. Cambridge University Press, 2009. Reprinted with corrections. (Cited on page 10.)

Nicolas Bourbaki. *Topological Vector Spaces*. Elements of Mathematics. Springer-Verlag, 1987. (Cited on page 62.)

Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004. (Cited on pages 68, 72, 169, 174, 177, 191, and 218.)

Ronald J. Brachman and Hector J. Levesque. *Knowledge Representation and Reasoning*. The Morgan Kaufmann Series in Artificial Intelligence. Morgan Kaufmann Publishers, 2004. (Cited on pages 1, 10, and 12.)

Alonzo Church. An Unsolvable Problem of Elementary Number Theory. *American Journal of Mathematics*, 58:345–363, 1936. (Cited on page 12.)

A. Cocura, K. Kersting, C. Plageman, W. Burgard, and L. De Raedt. Learning Relational Navigation Policies. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS-06)*, 2006. (Cited on page 33.)

Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. John Wiley and Sons, second edition, 2006. (Cited on page 47.)

Michael A. Covington, Donald Nute, and Andre Vellino. *Prolog Programming in Depth*. Prentice Hall, 1996. (Cited on pages 14 and 189.)

Lionel Daniel. *Paraconsistent Probabilistic Reasoning*. PhD thesis, L'École Nationale Supérieure des Mines de Paris, 2009. (Cited on pages 3, 4, 5, 43, 46, 73, 76, 77, 78, 83, 84, 85, 86, 88, 90, 91, 122, 123, 124, 220, and 221.)

John Davies, Dieter Fensel, and Frank Van Harmelen, editors. *Towards the Semantic Web: Ontology-driven Knowledge Management*. John Wiley and Sons, 2003. (Cited on page 1.)

Martin Davis, George Logemann, and Donald Loveland. A Machine Program for Theorem Proving. *Communications of the ACM*, 5(7):394–397, 1962. (Cited on page 12.)

Luc De Raedt, Paolo Frasconi, Kristian Kersting, and Stephen H. Muggleton, editors. *Probabilistic Inductive Logic Programming: Theory and Applications*, volume 4911 of *Lecture Notes in Computer Science*. Springer-Verlag, 2008. (Cited on pages 3 and 33.)

Rodrigo de Salvo Braz, Eyal Amir, and Dan Roth. Lifted First-Order Probabilistic Inference. In *In Proceedings of the 19th International Joint Conference on Artificial Intelligence (IJCAI'05)*, pages 1319–1325, 2005. (Cited on pages 194, 215, 216, and 217.)

James P. Delgrande. On First-Order Conditional Logics. *Artificial Intelligence*, 105(1–2):105–137, 1998. (Cited on pages 129 and 160.)

Pedro Domingos and Matthew Richardson. Markov Logic: A Unifying Framework for Statistical Relational Learning. In Lise Getoor and Ben Taskar, editors, *Introduction to Statistical Relational Learning, see (Getoor and Taskar, 2007)*, chapter 12. The MIT Press, 2007. (Cited on pages 33 and 38.)

Didier Dubois, Jérôme Lang, and Henri Prade. Inconsistency in Possibilistic Knowledge Bases: To live with it or not live with it. In Lotfi A. Zadeh and Janusz Kacprzyk, editors, *Fuzzy Logic for the Management of Uncertainty*, pages 335–351. Jon Wiley and Sons, New York, 1992. (Cited on page 43.)

Patricia Everaere, Sébastien Konieczny, and Pierre Marquis. Conflict-Based Merging Operators. In Gerhard Brewka and Jérôme Lang, editors, *Principles of Knowledge Representation and Reasoning: Proceedings of the Eleventh International Conference (KR 2008)*, pages 348–357. AAAI Press, 2008. (Cited on page 89.)

Daan Fierens, Hendrik Blockeel, Jan Ramon, and Maurice Bruynooghe. Logical bayesian networks. In Sašo Džeroski and Hendrik Blockeel, editors, *Proceedings of the 3nd international workshop on multi-relational data mining*, pages 19–30, 2004. (Cited on pages 132 and 133.)

Marc Finthammer and Matthias Thimm. An Integrated Development Environment for Probabilistic Relational Reasoning. *Logic Journal of the IGPL*, 2011. To appear. (Cited on pages v, 41, 131, and 191.)

Marc Finthammer, Gabriele Kern-Isberner, and Manuela Ritterskamp. Resolving Inconsistencies in Probabilistic Knowledge Bases. In *KI 2007: Advances in Artificial Intelligence – 30th Annual German Conference on Artifical Intelligence*, volume 4667 of *Lecture Notes in Computer Science*, pages 114–128. Springer-Verlag, 2007. (Cited on pages 3, 4, 43, 90, 103, 122, 125, and 221.)

Marc Finthammer, Sebastian Loh, and Matthias Thimm. Towards a Toolbox for Relational Probabilistic Knowledge Representation, Reasoning, and Learning. In *Workshop on Relational Approaches to Knowledge Representation and Learning, Proceedings*, pages 34–48, 2009. (Cited on page v.)

Marc Finthammer. *Behandlung von Inkonsistenzen in probabilistischen Wissensbasen*. VDM Verlag Dr. Müller, 2008. (Cited on pages 43 and 125.)

Jens Fisseler. Toward Markov Logic with Conditional Probabilities. In David C. Wilson and H. Chad Lane, editors, *Proceedings of the Twenty-First International FLAIRS Conference*, pages 643–648. AAAI Press, 2008. (Cited on pages 39, 41, and 189.)

Jens Fisseler. *Learning and Modeling with Probabilistic Conditional Logic*, volume 328 of *Dissertations in Artificial Intelligence*. IOS Press, 2010. (Cited on pages 3, 4, 5, 27, 130, 146, 147, 151, 157, 161, 180, 181, 190, 215, and 221.)

Frederic Brenton Fitch. *Symbolic Logic: An Introduction*. The Ronald Press Company, 1952. (Cited on page 9.)

Melvin Fitting. *First-Order Logic and Automated Theorem Proving*. Texts in Computer Science. Springer-Verlag, second edition, 1996. (Cited on page 12.)

Nir Friedman, Lise Getoor, Daphne Koller, and Avi Pfeffer. Learning Probabilistic Relational Models. In Thomas Dean, editor, *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence (IJCAI 99)*, pages 1300–1309. Morgan Kaufmann Publishers, 1999. (Cited on page 32.)

Dov M. Gabbay. Classical vs. Non-Classical Logic. In Dov M. Gabbay, Christoper J. Hogger, and John A. Robinson, editors, *Handbook of Logic in Artificial Intelligence and Logic Programming*, volume 2, chapter 2.6. Oxford University Press, 1994. (Cited on page 9.)

A. Garcia and Guillermo R. Simari. Defeasible Logic Programming: An Argumentative Approach. *Theory and Practice of Logic Programming*, 4(1–2):95–138, 2004. (Cited on page 89.)

Diego R. García, Sebastián Gottifredi, Patrick Krümpelmann, Matthias Thimm, Gabriele Kern-Isberner, Marcelo A. Falappa, and Alejandro J. García. On Influence and Contractions in Defeasible Logic Programming. In *Proceedings of the Eleventh International Conference on Logic Programming and Nonmonotonic Reasoning*, 2011. (Cited on page viii.)

Michael R. Garey and David S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. Series of Books in the Mathematical Sciences. W. H. Freeman, 1990. (Cited on page 12.)

Michael Gelfond and Nicola Leone. Logic Programming and Knowledge Representation – The A-Prolog Perspective. *Artificial Intelligence*, 138(1–2):3–38, 2002. (Cited on pages 15, 18, 33, and 89.)

Michael Gelfond and Vladimir Lifschitz. Classical Negation in Logic Programs and Disjunctive Databases. *New Generation Computing*, 9:365–385, 1991. (Cited on pages 15, 18, 33, and 89.)

Giangiacomo Gerla. *Fuzzy Logic: Mathematical Tools for Approximate Reasoning*. Trends in Logic. Springer-Verlag, 2001. (Cited on pages 19 and 83.)

Vladimir Geroimenko. *Dictionary of XML Technologies and the Semantic Web*. Springer Professional Computing Series. Springer-Verlag, 2004. (Cited on page 2.)

Lise Getoor and Ben Taskar, editors. *Introduction to Statistical Relational Learning*. The MIT Press, 2007. (Cited on pages 3, 33, 272, 274, and 277.)

Lise Getoor, Nir Friedman, Daphne Koller, Avi Pfeffer, and Ben Taskar. Probabilistic Relational Models. In Lise Getoor and Ben Taskar, editors, *Introduction to Statistical Relational Learning, see (Getoor and Taskar, 2007)*, chapter 5. The MIT Press, 2007. (Cited on page 132.)

Moises Goldszmidt, Paul Morris, and Judea Pearl. A Maximum Entropy Approach to Nonmonotonic Reasoning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(3):220–232, 1993. (Cited on page 32.)

John Grant and Anthony Hunter. Measuring Inconsistency in Knowledgebases. *Journal of Intelligent Information Systems*, 27:159–184, 2006. (Cited on pages 43, 46, and 78.)

John Grant and Anthony Hunter. Analysing Inconsistent First-Order Knowledgebases. *Artificial Intelligence*, 172(8–9):1064–1093, 2008. (Cited on page 78.)

Eric Grégoire and Sébastien Konieczny. Logic-based Approaches to Information Fusion. *Information Fusion*, 7:4–18, 2006. (Cited on page 103.)

Adam J. Grove, Joseph Y. Halpern, and Daphne Koller. Random Worlds and Maximum Entropy. *Journal of Artificial Intelligence Research*, 2:33–88, 1994. (Cited on pages 151, 156, 157, 180, 184, 185, 186, 187, 188, 199, 214, 286, and 290.)

Adam J. Grove, Joseph Y. Halpern, and Daphne Koller. Asymptotic Conditional Probabilities: The Non-Unary Case. *The Journal of Symbolic Logic*, 61(1):250–276, 1996. (Cited on pages 33 and 146.)

Adam J. Grove, Joseph Y. Halpern, and Daphne Koller. Asymptotic Conditional Probabilities: The Unary Case. *SIAM Journal on Computing*, 25(1):1–51, 1996. (Cited on page 146.)

Jan Gullberg. *Mathematics: From the Birth of Numbers*. W. W. Norton & Company, 1997. (Cited on page 19.)

Joseph Y. Halpern. An Analysis of First-Order Logics of Probability. *Artificial Intelligence*, 46:311–350, 1990. (Cited on pages 146, 151, 152, 154, 156, 157, 180, 184, 190, and 286.)

Pierre Hansen and Brigitte Jaumard. Probabilistic Satisfiability. Technical report, Les Cahiers du GERAD, École Polytechique de Montréal, 1996. (Cited on pages 43 and 45.)

Sven Ove Hansson. *A Textbook of Belief Dynamics: Theory Change and Database Updating*. Applied Logic Series. Springer-Verlag, 1999. (Cited on pages 4, 10, 46, and 89.)

David Hilbert and Wilhelm Ackermann. *Grundzüge der theoretischen Logik (Principles of Mathematical Logic)*. Springer-Verlag, 1928. (Cited on page 12.)

Daniel Hölzgen, Thomas Vengels, Patrick Krümpelmann, Matthias Thimm, and Gabriele Kern-Isberner. Argonauts – A Working System for Motivated Cooperative Agents. *Annals of Mathematics and Artificial Intelligence*, 2011. To appear. (Cited on page viii.)

Alfred Horn. On Sentences which are True of Direct Unions of Algebras. *Journal of Symbolic Logic*, 16:14–21, 1951. (Cited on page 14.)

Michael C. Horsch and David Poole. A Dynamic Approach to Probabilistic Inference using Bayesian Networks. In Piero P. Bonissone, Max Henrion, Laveen N. Kanal, and John F. Lemmer, editors, *Proceedings of the Sixth Conference on Uncertainty in Artifical Intelligence*, pages 155–161. Elsevier, 1990. (Cited on page 215.)

Anthony Hunter and Sébastien Konieczny. Approaches to Measuring Inconsistent Information. In Leopoldo E. Bertossi, Anthony Hunter, and Torsten Schaub, editors, *Inconsistency Tolerance*, volume 3300 of *Lecture Notes in Computer Science*, pages 189–234. Springer-Verlag, 2004. (Cited on page 78.)

Anthony Hunter and Sébastien Konieczny. Shapley Inconsistency Values. In *Proceedings of the Tenth International Conference on Knowledge Representation (KR'06)*, pages 249–259. AAAI Press, 2006. (Cited on pages 47, 78, 90, 93, 94, and 122.)

Anthony Hunter and Sébastien Konieczny. Measuring Inconsistency through Minimal Inconsistent Sets. In Gerhard Brewka and Jérôme Lang, editors, *Proceedings of the Eleventh International Conference on Principles of Knowledge Representation and Reasoning (KR'2008)*, pages 358–366, Sydney, Australia, 2008. AAAI Press. (Cited on pages 47, 49, 56, 57, 78, 90, 91, 92, 94, 96, 97, and 122.)

Anthony Hunter and Sébastien Konieczny. On the Measure of Conflicts: Shapley Inconsistency Values. *Artificial Intelligence*, 174(14):1007–1026, 2010. (Cited on pages 43, 46, 53, 54, 78, 90, 91, 94, and 122.)

Anthony Hunter. Measuring Inconsistency in Knowledge via Quasi-classical Models. In *Proceedings of the 18th American National Conference on Artificial Intelligence (AAAI'2002)*, pages 68–73. The MIT Press, 2002. (Cited on page 78.)

Anthony Hunter. Evaluating the Significance of Inconsistencies. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI'03)*, pages 468–473. Morgan Kaufmann Publishers, 2003. (Cited on page 78.)

Peter Jackson. *Introduction to Expert Systems*. Addison Wesley, third edition, 1998. (Cited on page 1.)

Manfred Jaeger. *Default Reasoning about Probabilities*. PhD thesis, Universität des Saarlandes, 1995. (Cited on pages 146, 154, 155, 156, 157, 159, 180, and 286.)

Manfred Jaeger. Relational Bayesian Networks: a Survey. *Electronic Transactions in Artificial Intelligence*, 6, 2002. (Cited on page 132.)

Edwin T. Jaynes. *Probability Theory: The Logic of Science*. Cambridge University Press, 2003. (Cited on pages 19 and 196.)

Phillip Judson. *Knowledge-based Expert Systems in Chemistry*. RSC Theoretical and Computational Chemistry Series. Springer-Verlag, 2009. (Cited on page 1.)

Jerry Kelly. *Social Choice Theory: An Introduction*. Springer-Verlag, 1988. (Cited on pages 103 and 104.)

Gabriele Kern-Isberner and Thomas Lukasiewicz. Combining Probabilistic Logic Programming with the Power of Maximum Entropy. *Artificial Intelligence, Special Issue on Nonmonotonic Reasoning*, 157(1–2):139–202, 2004. (Cited on pages 3, 4, 5, 27, 151, 157, 180, 181, and 190.)

Gabriele Kern-Isberner and Wilhelm Rödder. Belief Revision and Information Fusion in a Probabilistic Environment. In *Proceedings of the 16th International FLAIRS Conference, (FLAIRS'03)*, pages 506–510. AAAI Press, 2003. (Cited on page 124.)

Gabriele Kern-Isberner and Matthias Thimm. Novel semantical approaches to relational probabilistic conditionals. In Fangzhen Lin, Ulrike Sattler, and Mirosław Truszczyński, editors, *Proceedings of the Twelfth International Conference on the Principles of Knowledge Representation and Reasoning (KR'10)*. AAAI Press, 2010. (Cited on pages v, 5, and 137.)

Gabriele Kern-Isberner, Matthias Thimm, and Marc Finthammer. Qualitative Knowledge Discovery. In Klaus-Dieter Schewe and Bernhard Thalheim, editors, *Proceedings of the Thirrd International Workshop on Semantics in Data and Knowledge Bases (SDKB)*, volume 4925 of *Lecture Notes in Computer Science*, pages 88–113. Springer-Verlag, 2008. (Cited on page viii.)

Gabriele Kern-Isberner, Matthias Thimm, Marc Finthammer, and Jens Fisseler. Mining default rules from statistical data. In H. Chad Lane and Hans W. Guesgen, editors, *Proceedings of the 22th International FLAIRS Conference (FLAIRS'09)*. AAAI Press, 2009. (Cited on page viii.)

Gabriele Kern-Isberner. *Conditionals in Nonmonotonic Reasoning and Belief Revision*. Number 2087 in Lecture Notes in Computer Science. Springer-Verlag, 2001. (Cited on pages 2, 3, 26, 27, 28, 32, 190, and 191.)

Kristian Kersting and Luc De Raedt. Bayesian Logic Programs. In Celine Rouveirol and Michele Sebag, editors, *Inductive Logic Programming: Proceedings of the Tenth International Conference (ILP 2000)*. CEUR-WS.org, 2000. (Cited on page 34.)

Kristian Kersting and Uwe Dick. Balios – The Engine for Bayesian Logic Programs. In *Proceedings of the Eighth European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD-2004)*, pages 549–551, 2004. (Cited on page 133.)

Kristian Kersting and Luc De Raedt. Bayesian Logic Programming: Theory and Tool. In Lise Getoor and Ben Taskar, editors, *Introduction to Statistical Relational Learning, see (Getoor and Taskar, 2007)*, chapter 10. The MIT Press, 2007. (Cited on pages 33, 36, 38, and 142.)

Jacek Kisyński and David Poole. Lifted Aggregation in Directed First-order Probabilistic Models. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI'09)*, pages 1922–1929, 2009. (Cited on page 217.)

Kevin M. Knight. Measuring Inconsistency. *Journal of Philosophical Logic*, 31:77–98, 2001. (Cited on pages 43, 46, 59, 78, and 79.)

Kevin M. Knight. *A Theory of Inconsistency*. PhD thesis, University Of Manchester, July 2002. (Cited on pages 59, 78, 79, and 80.)

Stanley Kok, Parag Singla, Matthew Richardson, Pedro Domingos, Marc Sumner, Hoifung Poon, Daniel Lowd, and Jue Wang. *The Alchemy System for Statistical Relational AI: User Manual*. Department of Computer Science and Engineering, University of Washington, 2008. (Cited on page 41.)

Andrey Nikolaevich Kolmogorov. *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Springer-Verlag, 1933. (Cited on page 19.)

Sébastian Konieczny and Ramón Pino-Pérez. On the Logic of Merging. In Anthony G. Cohn, Lenhard K. Schubert, and Stuart C. Shapiro, editors, *Proceedings of the Sixth International Conference on Principles of Knowledge Representationa and Reasoning (KR'98)*. Morgan Kaufmann Publisher, 1998. (Cited on pages 46 and 103.)

Sébastien Konieczny and Ramón Pino-Pérez. Propositional Belief Base Merging or how to Merge Beliefs/Goals Coming from Several Sources and some Links with Social Choice Theory. *European Journal of Operational Research*, 160(3):785–802, 2005. (Cited on page 103.)

Sébastien Konieczny, Jérôme Lang, and Pierre Marquis. Quantifying Information and Contradiction in Propositional Logic Through Test Actions. In *Proceedings of the 18th International Joint Conference on Artificial Intelligence (IJCAI'03)*, pages 106–111. Morgan Kaufmann Publishers, 2003. (Cited on page 78.)

Sébastien Konieczny, Jérôme Lang, and Pierre Marquis. Reasoning under Inconsistency: The Forgotten Connective. In *Proceedings of the 19th International Joint Conference on Artificial Intelligence (IJCAI'05)*, pages 484–489. Morgan Kaufmann Publishers, 2005. (Cited on page 43.)

Patrick Krümpelmann and Matthias Thimm. A Logic Programming Framework for Reasoning about Know-How. In *Proceedings of the 14th International Workshop on Non-Monotonic Reasoning (NMR'10)*, 2010. (Cited on page viii.)

Patrick Krümpelmann, Matthias Thimm, Manuela Ritterskamp, and Gabriele Kern-Isberner. Belief Operations for Motivated BDI Agents. In Lin Padgham, David C. Parkes, Joerg P. Müller, and Simon Parsons, editors, *Proceedings of the Seventh International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2008)*, pages 421–428, 2008. (Cited on page viii.)

Henry Ely Kyburg, Ronald P. Loui, and Greg N. Carlson, editors. *Knowledge Representation and Defeasible Reasoning*. Studies in Cognitive Systems. Springer-Verlag, 1990. (Cited on page 2.)

Jérôme Lang and Pierre Marquis. Reasoning under Inconsistency: A Forgetting-based Approach. *Artificial Intelligence*, 174(12–13):799–823, 2010. (Cited on page 89.)

Steffen Lauritzen and David J. Spiegelhalter. Local Computations with Probabilities in Graphical Structures and their Applications to Expert Systems. *Journal of the Royal Statistical Society, Series B*, 50(2):415–448, 1988. (Cited on pages 22 and 23.)

Huma Lodhi and Stephen H. Muggleton. Modelling metabolic pathways using stochastic logic programs-based ensemble methods. In *Proceedings of the Second International Conference on Computational Methods in System Biology*. Springer, 2004. (Cited on page 33.)

Sebastian Loh, Matthias Thimm, and Gabriele Kern-Isberner. On the Problem of Grounding a Relational Probabilistic Conditional Knowledge Base. In *Proceedings of the 14th International Workshop on Non-Monotonic Reasoning (NMR'10)*, 2010. (Cited on pages v, 3, 4, 5, 27, 130, 146, 147, 148, 150, 151, 157, 180, 181, 183, 190, and 221.)

Sebastian Loh. Grundierungstechniken für Relationale Probabilistische Inferenz, 2009. Diploma thesis, Technische Universität Dortmund, Germany. (Cited on pages 146 and 181.)

Jana Ludolph. Qualitative Methoden zur Sicherung der Konsistenz bei probabilistischen Wissensbasen, 2009. Diploma thesis, Technische Universität Dortmund, Germany. (Cited on page 103.)

Thomas Lukasiewicz. Probabilistic Deduction with Conditional Constraints over Basic Events. *Journal of Artificial Intelligence Research*, 10(199–241), 1999. (Cited on page 30.)

David Makinson. General Theory of Cumulative Inference. In Michael Reinfrank, Johan de Kleer, Matthew Ginsberg, and Erik Sandewall, editors, *Non-Monotonic Reasoning: Proceedings of the Second International Workshop*, volume 346 of *Lecture Notes in Artificial Intelligence*, pages 1–18. Springer-Verlag, 1989. (Cited on pages 32, 159, and 162.)

Elliott Mendelson. *Introduction to Mathematical Logic*. Chapman & Hall/CRC, fourth edition, 1997. (Cited on page 9.)

Carl-Heinz Meyer. *Korrektes Schließen bei unvollständiger Information*. PhD thesis, FernUniversität in Hagen, Germany, 1997. (Cited on pages 218 and 221.)

Brian Milch, Luke S. Zettlemoyer, Kristian Kersting, Michael Haimes, and Leslie Pack Kaelbling. Lifted Probabilistic Inference with Counting Formulas. In *Procedings of the 23rd AAAI Conference on Artificial Intelligence*, pages 1062–1608, 2008. (Cited on pages 194, 215, and 217.)

David Miller. A Paradox of Information. *British Journal for the Philosophy of Science*, 17, 1966. (Cited on page 154.)

James Munkres. *Topology*. Prentice Hall, second edition, 1999. (Cited on page 62.)

Fabian Neuhaus, Stig Andur Pedersen, Uwe Scheffler, Heinrich Wansing, and Vincent Hendricks, editors. *First-Order Logic Revisited – Proceedings of the conference "75 Years of First-Order Logic" at Humboldt University, Berlin (Germany), September 18-21, 2003*. Logos-Verlag, 2004. (Cited on page 12.)

Donald Nute and Charles Cross. Conditional Logic. In Dov Gabbay and Franz Guenther, editors, *Handbook of Philosophical Logic*, volume 4, pages 1–98. Kluwer Academic Publishers, 2002. (Cited on page 26.)

Donald Nute. Defeasible Logic. In *Handbook of Logic in Artificial Intelligence and Logic Programming*, volume 3 of *Nonmonotonic Reasoning and Uncertain Reasoning*, pages 353–395. Oxford University Press, 1994. (Cited on pages 18 and 43.)

Jeff B. Paris and Alena Vencovská. In Defence of the Maximum Entropy Inference Process. *International Journal of Approximate Reasoning*, 17(1):77–103, 1997. (Cited on pages 190 and 191.)

Jeff B. Paris. *The Uncertain Reasoner's Companion: A Mathematical Perspective*. Cambridge Tracts in Theoretical Computer Science. Cambridge University Press, 1994. (Cited on pages 2, 4, 19, 20, 28, 31, 32, 46, 73, 76, 160, 173, 190, and 234.)

Jeff B. Paris. Common Sense and Maximum Entropy. *Synthese*, 117(1):75–93, 2000. (Cited on page 164.)

Giovanni Parmigiani. *Modeling in Medical Decision Making: A Bayesian Approach*. Statistics in Practice. John Wiley and Sons, 2002. (Cited on page 2.)

Judea Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann Publishers, 1998. (Cited on pages 2, 19, 20, 22, 23, 24, 26, 32, 34, 36, 37, 123, and 172.)

Charles S. Peirce. *Collected Papers of Charles Sanders Peirce: Pragmatism and Pragmaticism*. Number 5–6 in Paul Weiss and Charles Hartshorne, editors, Scientific Metaphysics. Harvard University Press, 1972. (Cited on page 9.)

David Picado-Muiño. Measuring and Repairing Inconsistency in Probabilistic Knowledge Bases. *International Journal of Approximate Reasoning*, 2011. (Cited on pages 62 and 67.)

David Poole. First-Order Probabilistic Inference. In Georg Gottlob and Toby Walsh, editors, *Proceedings of the 18th International Joint Conference on Artificial Intelligence (IJCAI-03)*, pages 985–991. Morgan Kaufmann Publishers, 2003. (Cited on pages 194, 215, 216, and 217.)

James Popple. *A Pragmatic Legal Expert System*. Applied Legal Philosophy Series. Dartmouth Publishing Group, 1996. (Cited on page 1.)

Iyad Rahwan and Guillermo R. Simari, editors. *Argumentation in Artificial Intelligence*. Springer-Verlag, 2009. (Cited on pages 2, 18, 43, and 89.)

Raymond Reiter. A Logic for Default Reasoning. *Artificial Intelligence*, 13(1–2):81–132, 1980. (Cited on pages 18, 43, and 89.)

Nicholas Rescher and Ruth Manor. On Inference from Inconsistent Premises. *Theory and Decision*, 1:179–219, 1970. (Cited on page 43.)

Matthew Richardson and Pedro Domingos. Markov Logic Networks. *Machine Learning*, 62(1–2):107–136, 2006. (Cited on pages 33, 38, 39, 40, and 194.)

Wilhelm Rödder and Carl-Heinz Meyer. Coherent Knowledge Processing at Maximum Entropy by SPIRIT. In *Proceedings of the Twelfth Conference on Uncertainty in Artificial Intelligence (UAI'96)*, pages 470–476, 1996. (Cited on pages 2, 184, and 191.)

Wilhelm Rödder and Longgui Xu. Entropy-driven Inference and Inconsistency. In David Heckerman and Joe Whittaker, editors, *Proceedings of The Seventh International Workshop on Artificial Intelligence and Statistics*, pages 272–277. Morgan Kaufmann Publishers, 1999. (Cited on page 124.)

Wilhelm Rödder and Longgui Xu. Elimination of inconsistent knowledge in the probabilistic expertsystem-shell spirit (in german). In B. Fleischmann, R. Lasch, K. Derigs, W. Domschke, and K. Riedler, editors, *Operations Research Proceedings: Selected Papers of the Symposium on Operations Research 2000*, pages 260–265. Springer-Verlag, 2001. (Cited on pages 3, 4, 43, 46, 77, 78, 80, 81, 103, 122, 124, 125, and 221.)

Wilhelm Rödder, Ivan Ricardo Gartner, and Sandra Rudolph. Entropy-driven Portfolio Selection - A Downside and Upside Risk Framework. Technical report, Diskussionsbeitrag der Fakultät für Wirtschaftswissenschaft, FernUniversität in Hagen, 2009. (Cited on page 2.)

Wilhelm Rödder. Conditional Logic and the Principle of Entropy. *Artificial Intelligence*, 117(1):83–106, 2000. (Cited on pages 3 and 26.)

Stuart Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach*. Prentice Hall, second edition, 2009. (Cited on pages 1, 16, and 206.)

Glenn Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, 1976. (Cited on page 19.)

Lloyd Stowell Shapley. A Value for n-Person Games. In Harold William Kuhn and Albert William Tucker, editors, *Contributions to the Theory of Games II*, volume 28 of *Annals of Mathematics Studies*, pages 307–317. Princeton University Press, 1953. (Cited on pages 90, 93, 94, and 96.)

John E. Shore and Rodney W. Johnson. Axiomatic Derivation of the Principle of Maximum Entropy and the Principle of Minimum Cross-Entropy. *IEEE Transactions on Information Theory*, 26:26–37, 1980. (Cited on pages 32, 190, 191, and 245.)

Edward .H. Shortliffe and Bruce G. Buchanan. A Model of Inexact Reasoning in Medicine. *Mathematical Biosciences*, 23(3–4):351–379, 1975. (Cited on page 1.)

William Siler and James J. Buckley. *Fuzzy Expert Systems and Fuzzy Reasoning*. John Wiley and Sons, 2005. (Cited on page 2.)

Brian Skyrms. Higher Order Degrees of Belief. In David Hugh Mellor, editor, *Prospects of Pragmatism: Essays in Honor of F.P. Ramsey*. Cambridge University Press, 1980. (Cited on page 154.)

Ofer Strichmann and Stefan Szeider, editors. *Theory and Applications of Satisfiability Testing - SAT 2010, 13th International Conference, Edinburgh, UK, July 11–14, 2010*, volume 6175 of *Lecture Notes in Computer Science*. Springer-Verlag, 2010. (Cited on page 12.)

Matthias Thimm and Alejandro J. García. Classification and Strategical Issues of Argumentation Games on Structured Argumentation Frameworks. In Wiebe van der Hoek, Yves Lespérance Gal A. Kaminka, Michael Luck, and Sandip Sen, editors, *Proceedings of the Ninth International Joint Conference on Autonomous Agents and Multi-Agent Systems 2010 (AAMAS'10)*, 2010. (Cited on page vii.)

Matthias Thimm and Gabriele Kern-Isberner. A Distributed Argumentation Framework using Defeasible Logic Programming. In Philippe Besnard, Sylvie Doutre, and Anthony Hunter, editors, *Proceedings of the Second International Conference on Computational Models of Argument (COMMA'08)*, Frontiers in Artificial Intelligence and Applications, pages 381–392. IOS Press, 2008. (Cited on page vii.)

Matthias Thimm and Gabriele Kern-Isberner. A Distributed Argumentation Framework using Defeasible Logic Programming (Extended Version). Technical report, Technische Universität Dortmund, 2008. (Cited on page vii.)

Matthias Thimm and Gabriele Kern-Isberner. On the Relationship of Defeasible Argumentation and Answer Set Programming. In Philippe Besnard, Sylvie Doutre, and Anthony Hunter, editors, *Proceedings of the Second International Conference on Computational Models of Argument (COMMA'08)*, Frontiers in Artificial Intelligence and Applications, pages 393–404. IOS Press, 2008. (Cited on page vii.)

Matthias Thimm and Gabriele Kern-Isberner. On the Relationship of Defeasible Argumentation and Answer Set Programming (Extended Version). Technical report, Technische Universität Dortmund, 2008. (Cited on page vii.)

Matthias Thimm and Gabriele Kern-Isberner. On Probabilistic Inference in Relational Conditional Logics, 2011. Submitted. (Cited on pages v and 174.)

Matthias Thimm and Patrick Krümpelmann. Know-How for Motivated BDI Agents (Extended Abstract). In Keith S. Decker, Jaime Simão Sichman, Carles Sierra, and Cristiano Castelfranchi, editors, *Proceedings of the Eighth International Conference on Autonomous Agents and Multiagent Systems (AAMAS'09)*, pages 1143–1144, 2009. (Cited on page viii.)

Matthias Thimm and Patrick Krümpelmann. Know-How for Motivated BDI Agents (Extended Version). Technical report, Technische Universität Dortmund, Department of Computer Science, 2009. (Cited on page viii.)

Matthias Thimm, Alejandro J. Garcia, Gabriele Kern-Isberner, and Guillermo R. Simari. Using Collaborations for Distributed Argumentation with Defeasible Logic Programming. In Maurice Pagnucco and Michael Thielscher, editors, *Proceedings of the 12th International Workshop* *on Non-Monotonic Reasoning (NMR'08)*, pages 179–188. University of New South Wales, Technical Report No. UNSW-CSE-TR-0819, 2008. (Cited on page vii.)

Matthias Thimm, Marc Finthammer, Sebastian Loh, Gabriele Kern-Isberner, and Christoph Beierle. A System for Relational Probabilistic Reasoning on Maximum Entropy. In Hans W. Guesgen and R. Charles Murray, editors, *Proceedings of the 23rd International FLAIRS Conference (FLAIRS'10)*. AAAI Press, 2010. (Cited on pages v and 146.)

Matthias Thimm, Marc Finthammer, Gabriele Kern-Isberner, and Christoph Beierle. Comparing Approaches to Relational Probabilistic Reasoning: Theory and Implementation, 2011. Submitted. (Cited on pages v, 189, and 191.)

Matthias Thimm, Gabriele Kern-Isberner, and Jens Fisseler. Relational Probabilistic Conditional Reasoning at Maximum Entropy. In *Proceedings of the The Eleventh European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU'11)*, 2011. (Cited on pages v and 174.)

Matthias Thimm. *Verteilte logikbasierte Argumentation: Konzeption, Implementierung und Anwendung im Rechtswesen (Diploma Thesis)*. VDM Verlag Dr. Müller, 2008. (Cited on page vii.)

Matthias Thimm. Measuring inconsistency in probabilistic knowledge bases. In Jeff Bilmes and Andrew Ng, editors, *Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence (UAI'09)*, 2009. (Cited on pages v, 5, 65, 77, and 78.)

Matthias Thimm. Representing Statistical Information and Degrees of Belief in First-Order Probabilistic Conditional Logic. In *Workshop on Relational Approaches to Knowledge Representation and Learning, Proceedings*, pages 49–63, 2009. (Cited on pages v, 5, 137, and 169.)

Matthias Thimm. Realizing Argumentation in Multi-Agent Systems using Defeasible Logic Programming. In Peter McBurney, Iyad Rahwan, Simon Parsons, and Nicolas Maudet, editors, *Argumentation in Multi-Agent Systems, Post-Proceedings of the Sixth International Workshop (ArgMAS), Budapest, Hungary, May 2009 (revised, selected and invited papers)*, volume 6057 of *Lecture Notes in Artificial Intelligence*. Springer-Verlag, 2010. (Cited on page vii.)

Alan Turing. On Computable Numbers, with an Application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society, Series 2*, 42:230–265, 1937. (Cited on page 12.)

Miklos A. Vasarhelyi, Enrique Bonson, and Rani Hoitash, editors. *Artificial Intelligence in Accounting & Auditing: International Perspectives*. Rutgers

Series in Accounting Information Systems. Markus Wiener Publishers, 2005. (Cited on page 1.)

Nevin L. Zhang and David Poole. Exploiting Causal Independence in Bayesian Network Inference. *Journal of Artificial Intelligence Research*, 5:301–328, 1996. (Cited on page 215.)

$\mathbb{N}_0$	Natural numbers with zero
$\mathbb{R}$	Real numbers
$\mathbb{R}_0^+$	Non-negative real numbers
$\mathbb{Q}$	Rational numbers
$\mathbb{Q}_0^+$	Non-negative rational numbers
$\mathfrak{P}(X)$	Power set of X, i. e. the set of subsets of X
$\mathfrak{P}^{\mathrm{ord}}(X)$	Set of all vectors of elements of X
$\mathsf{Dom}\, f$	Domain of a function f, i. e., if $f : X \to Y$ then $\mathsf{Dom}\, f = X$
$\mathsf{Im}\, f$	Image of a function f, i. e. $\mathsf{Im}\, f = \{f(x) \mid x \in \mathsf{Dom}\, f\}$
$\binom{n}{k}$	Binomial coefficient indexed by n and k
$\binom{n}{k_1,\ldots,k_r}$	Multinomial coefficient indexed by n and $k_1, \ldots, k_r$
$\vec{x}$	Vector $x = (x_1, \ldots, x_n)$
$\vert S \vert$	Cardinality of a set S
$\mathsf{ld}\, x$	Binary logarithm of x
$\mathsf{ln}\, x$	Natural logarithm of x
$\mathsf{exp}(x)$	Exponential function with base e
$D(\vec{y}, \vec{z})$	Generalized divergence of $\vec{y}$ to $\vec{z}$, cf. page 81
$\mathsf{sgn}(x)$	Sign of x
$\mathsf{cl}(\mathcal{G})$	Set of cliques of graph $\mathcal{G}$
$\mathsf{nd}(v)$	Set of non-descendants of a node v in a directed graph $\mathcal{G}$
$\mathsf{pa}(v)$	Set of parents of a node v in a directed graph $\mathcal{G}$
$\mathbb{B}$	Set of Boolean truth values true and false
$\top$	Tautology, cf. Definition 2.2 on page 10 and Definition 2.8 on page 13
$\bot$	Contradiction, cf. Definition 2.2 on page 10 and Definition 2.8 on page 13
At	Propositional signature (a set of atoms), cf. Definition 2.1 on page 10
$\mathsf{At}(S)$	Set of atoms appearing in a set S of formulas or conditionals
$\mathsf{At}(\Sigma)$	Herbrand base of Σ, cf. Definition 2.13 on page 17

Σ	First-order signature, cf. Definition 2.5 on page 13
$\mathsf{Lit}(\mathsf{At})$	Set of propositional literals with respect to At, cf. page 10
$\mathsf{Lit}(\Sigma, V)$	Set of first-order literals with respect to Σ and V, cf. page 14
$\mathsf{Terms}(\Sigma, V)$	Set of first-order terms with respect to Σ and V, cf. Definition 2.6 on page 13
$\mathcal{L}(\mathsf{At})$	Propositional language with respect to At, cf. Definition 2.2 on page 10
$\mathcal{L}(\Sigma, V)$	First-order language with respect to Σ and V, cf. Definition 2.8 on page 13
$\mathcal{L}^{\not\forall\not\exists}(\Sigma, V)$	Quantifier-free fragment of $\mathcal{L}(\Sigma, V)$, cf. page 14
$\mathcal{L}$	Logical language, i. e. either $\mathcal{L}(\mathsf{At})$ or $\mathcal{L}(\Sigma, V)$, cf. page 18
$(\mathcal{L} \mid \mathcal{L})$	Conditional language on $\mathcal{L}$, cf. Definition 2.23 on page 26
$(\mathcal{L} \mid \mathcal{L})^{pr}$	Probabilistic conditional language on $\mathcal{L}$, cf. Definition 2.24 on page 26
$(\mathcal{L} \mid \mathcal{L})^{pr,pr}$	Probabilistic bounded conditional language on $\mathcal{L}$, cf. page 30
$\mathcal{L}_1$	Probabilistic language of (Halpern, 1990), cf. Section 5.5.2 on page 151ff.
$\mathcal{L}_2$	Probabilistic language of (Halpern, 1990), cf. Section 5.5.2 on page 151ff.
$\mathcal{L}_3$	Probabilistic language of (Halpern, 1990), cf. Section 5.5.2 on page 151ff.
$\mathcal{L}^{\approx}$	Probabilistic language of (Grove *et al.*, 1994; Bacchus *et al.*, 1996), cf. Section 6.4.2 on page 184ff.
$\mathcal{L}^{\beta}$	Probabilistic language of (Jaeger, 1995), cf. Section 5.5.2 on page 151ff.
$\mathsf{Int}(\mathsf{At})$	Set of propositional interpretations $I : \mathsf{At} \to \{\mathsf{true}, \mathsf{false}\}$, cf. Definition 2.3 on page 11
$\mathsf{Int}(\Sigma)$	Set of first-order interpretations $I = (U_I, f_I^U, Pred_I, Func_I)$ with respect to Σ, cf. Definition 2.9 on page 15
$\mathcal{P}^{\mathrm{P}}(\mathsf{At})$	Set of all probability functions on $\Omega(\mathsf{At})$, cf. page 20
$\mathcal{P}^{\mathrm{F}}(\Sigma)$	Set of all probability functions on $\Omega(\Sigma)$, cf. page 135
$\hat{\mathcal{P}}_D$	Set of condensed probability functions wrt. D, cf. Definition 7.9 on page 204
$\Omega(\mathsf{At})$	Set of propositional possible worlds, cf. page 11
$\Omega(\Sigma)$	Set of Herbrand interpretations with respect to Σ, cf. Definition 5.2 on page 136

$\Omega(\Sigma, D)$	Set of relevant Herbrand interpretations of $\Omega(\Sigma)$ wrt. D, cf. Definition 2.14 on page 17
$\hat{\Omega}$	Set of reference worlds, cf. Definition 7.6 on page 199
$\models^{\mathrm{P}}$	Satisfaction relation for propositional logic, cf. page 11
$\models^{\mathrm{F}}$	Satisfaction relation for first-order logic, cf. page 15
$\models^{pr}$	Satisfaction relation for probabilistic conditional logic, cf. page 28 and page 135
$\models^{pr}_{int}$	Inference relation for bounded probabilistic conditionals, cf. page 30
$\models^{pr}_{\varnothing}$	Satisfaction relation for averaging semantics, cf. Section 5.3.1 on page 136ff.
$\models^{pr}_{\odot}$	Satisfaction relation for aggregating semantics, cf. Section 5.3.2 on page 140ff.
$\models_{\mathcal{G}}$	Satisfaction relation for grounding semantics, cf. Section 5.5.1 on page 146ff. and Section 6.4.1 on page 181ff.
$\models^{pr}_{1}$	Satisfaction relation for $\mathcal{L}_1$, cf. Section 5.5.2 on page 151ff.
$\models^{pr}_{2}$	Satisfaction relation for $\mathcal{L}_2$, cf. Section 5.5.2 on page 151ff.
$\models^{pr}_{3}$	Satisfaction relation for $\mathcal{L}_3$, cf. Section 5.5.2 on page 151ff.
$\models^{\epsilon}_{\approx}$	Satisfaction relation for $\mathcal{L}^{\approx}$, cf. Section 6.4.2 on page 184ff.
$\mathsf{Mod}^{\mathrm{P}}(\Phi)$	Set of propositional models ($\subseteq \mathsf{Int}(\mathsf{At})$) of a set Φ of propositional formulas, cf. Definition 2.4 on page 11
$\mathsf{Mod}^{\mathrm{F}}(\Phi)$	Set of first-order models ($\subseteq \mathsf{Int}(\Sigma)$) of a set Φ of first-order formulas, cf. Definition 2.4 on page 11
$\mathsf{Mod}^{\mathrm{Pr}}(X)$	Set of probability functions ($\subseteq \mathcal{P}^{\mathrm{P}}(\mathsf{At})$) of a probabilistic conditional or a knowledge base X, cf. page 28
$\equiv^{\mathrm{P}}$	Propositional equivalence, cf. page 12
$\equiv^{\mathrm{F}}$	First-order equivalence, cf. page 16
$\equiv^{\mathrm{pr}}$	Probabilistic equivalence, cf. page 29
$\equiv^{\mathrm{kb}}$	Probabilistic kb-equivalence, cf. page 29
$\equiv^{\mathrm{cond}}$	Probabilistic cond-equivalence, cf. page 29
$\equiv^{\varnothing}$	Equivalence wrt. averaging semantics, cf. page 138
$\equiv^{\odot}$	Equivalence wrt. aggregating semantics, cf. page 143
$\equiv_{\mathcal{R}}$	$\mathcal{R}$-equivalence, cf. Definition 6.1 on page 163 and Definition 7.3 on page 197
$\mathfrak{S}(\mathcal{R})$	Set of $\mathcal{R}$-equivalences classes, cf. Definition 6.2 on page 163
$\mathfrak{S}\|_{D}(\mathcal{R})$	Set of focused $\mathcal{R}$-equivalences classes wrt. D, cf. page 197

$\Gamma(\Sigma, V)$	Set of replacements with respect to Σ and V, cf. Definition 2.12 on page 16
$\Gamma^{\text{gnd}}(\Sigma, V)$	Set of ground replacements with respect to Σ and V, cf. page 16
$\mathsf{Const}(X)$	Set of constants appearing in a first-order formula, a set of first-order formulas, a relational probabilistic conditional, or a set of relational probabilistic conditionals X, cf. pages 17 and 133
$\mathsf{gnd}_D(X)$	Set of ground instances of a first-order formula, a set of first-order formulas, a relational probabilistic conditional, or a set of relational probabilistic conditionals X with respect to D, cf. pages 17 and 134
$\mathsf{gnd}_D^P((\psi \mid \phi))$	Set of ground instances $(\psi' \mid \phi')$ of a relational probabilistic conditional $(\psi \mid \phi)$ with respect to D that satisfy $(P(\phi') > 0$, cf. page 137
$H(P)$	Entropy of a probability function P, cf. Definition 2.26 on page 31
$H_D(P)$	Entropy of a probability function P wrt. D, cf. page 168 and page 207
P_0	Uniform probability function, cf. page 19
$\mathsf{ucp}(\psi \mid \phi)$	Probability of $(\psi \mid \phi)$ in the uniform probability function, cf. Definition 4.7 on page 108
$\underline{P}$	Focused probability function of P, cf. Definition 7.2 on page 196
$\mathsf{ME}(\mathcal{R})$	ME model of a knowledge base $\mathcal{R}$ of propositional probabilistic conditionals, cf. page 32
$\perp\!\!\!\perp_{\mathrm{P}}$	Probabilistic resp. conditional independence, cf. Definition 2.16 on page 21 and Definition 2.18 on page 22
$\perp\!\!\!\perp_{\mathcal{G}}$	Graph separation, cf. Definition 2.20 on page 24
CPD_p	Set of conditional probability distributions for a predicate p, cf. Definition 2.29 on page 34
$\Lambda_{\mathcal{R}}$	Characteristic function of $\mathcal{R}$, cf. Definition 3.3 on page 51
$\theta_{\mathsf{Inc},\mathcal{R}}$	Characteristic inconsistency function of Inc and $\mathcal{R}$, cf. Definition 3.4 on page 51
$\theta^{\dagger}_{C,\mathcal{R}}$	Characteristic culpability function of C and $\mathcal{R}$, cf. Definition 4.2 on page 92
Inc^d	Drastic inconsistency measure, cf. Definition 3.5 on page 53
$\mathsf{Inc}^{\mathsf{MI}}$	MI inconsistency measure, cf. Definition 3.6 on page 54

$\mathsf{Inc}_0^{\mathsf{MI}}$	Normalized MI inconsistency measure, cf. Definition 3.7 on page 55
$\mathsf{Inc}_C^{\mathsf{MI}}$	MI^C inconsistency measure, cf. Definition 3.8 on page 57
$\mathsf{Inc}_{C,0}^{\mathsf{MI}}$	Normalized MI^C inconsistency measure, cf. Definition 3.9 on page 58
Inc^*	MinDev inconsistency measure, cf. Definition 3.10 on page 60
Inc_0^*	Normalized MinDev inconsistency measure, cf. Definition 3.11 on page 66
$\mathcal{I}^{\geq}$	Upper approximation of Inc^*, cf. page 72
$\mathcal{I}^{\leq}$	Lower approximation of Inc^*, cf. page 71
Inc_b^*	Bounded MinDev inconsistency measure, cf. page 73
Inc_{LP}^*	MinDev inconsistency measure on linear constraints, cf. page 77
Inc_μ^h	Inconsistency measure based on candidacy degrees, cf. page 84
Inc_{gd}	Inconsistency measure based on generalized divergence, cf. page 81
$\mathsf{MI}(\mathcal{R})$	Set of minimal inconsistent subsets of $\mathcal{R}$, cf. page 48
$\mathcal{MD}(\mathcal{R})$	Set of probability values for knowledge bases with minimal (1-norm) distance to $\mathcal{R}$, cf. page 63
$\mathcal{MD}_i(\mathcal{R})$	Set of minimal distances for the probabilistic conditional r_i, cf. page 98
$\mathcal{PMD}(\mathcal{R})$	Set of probability functions for knowledge bases with minimal (1-norm) distance to $\mathcal{R}$, cf. page 63
$\mathsf{Cons}(\mathcal{R})$	Constrained satisfaction problem to determine consistency of $\mathcal{R}$, cf. page 45
$\mathsf{DevCons}(\mathcal{R})$	Set of constraints used for the optimization problem to compute $\mathsf{Inc}^*(\mathcal{R})$, cf. page 60
$\mathsf{DevConsLin}(\mathcal{R})$	Set of constraints used for the optimization problems to compute $\mathcal{I}^{\geq}(\mathcal{R})$ and $\mathcal{I}^{\leq}(\mathcal{R})$, cf. page 72ff.
$\mathsf{DevConsLin}_b(\mathcal{R})$	Set of constraints used for the optimization problem to compute $\mathsf{Inc}_b^*(\mathcal{R})$, cf. page 73
$\mathsf{DevConsLin}_{lp}(\mathcal{R})$	Set of constraints used for the optimization problem to compute $\mathsf{Inc}_{LP}^*(\mathcal{R})$, cf. page 77
$\mathsf{CRDevCon}(C,\mathcal{R})$	Set of constraints used for computing $\Upsilon_C^B(\mathcal{R})$, cf. Definition 4.14 on page 120
$S_{\mathsf{Inc}}^{\mathcal{R}}$	Shapley culpability measure, cf. Definition 4.5 on page 95

$A^{\mathcal{R}}$	Mean distance culpability measure, cf. Definition 4.6 on page 98
$\mathsf{SignCulp}^{\mathcal{R}}(r)$	Sign of culpability of r in $\mathcal{R}$, cf. page 98
Υ^{U}	Unbiased consistency restorer, cf. Definition 4.9 on page 109
Υ^{P}	Penalizing consistency restorer, cf. Definition 4.11 on page 113
Υ^{S}_{C}	Smoothed penalizing consistency restorer, cf. Definition 4.13 on page 118
Υ^{B}_{C}	Balanced consistency restorer, cf. Definition 4.14 on page 120
Ξ	Creeping function, cf. page 107
Ξ^{U}	Unbiased creeping function, cf. Definition 4.8 on page 109
Ξ^{P}	Penalizing creeping function, cf. Definition 4.10 on page 112
Ξ^{S}	Smoothed penalized creeping function, cf. Definition 4.12 on page 118
$\mathcal{I}_{\varnothing}$	Averaging inference operator, cf. Section 6.2.1 on page 169ff.
$\mathcal{I}_{\odot}$	Aggregating inference operator, cf. Section 6.2.2 on page 173ff.
$\mathcal{I}_{\mathcal{G}}$	Grounding inference operator, cf. Section 6.4.1 on page 181ff.
Pr^{ϵ}_{n}	Degree of belief for domain size n in the framework of (Grove *et al.*, 1994; Bacchus *et al.*, 1996), cf. Section 6.4.2 on page 184ff.
Pr_{∞}	Degree of belief for infinite domain size in the framework of (Grove *et al.*, 1994; Bacchus *et al.*, 1996), cf. Section 6.4.2 on page 184ff.
$\varrho(\omega)$	Expansion set of ω, cf. Definition 7.1 on page 196
Θ	Set of truth configurations, cf. Definition 7.4 on page 199
$\mathfrak{I}$	Set of instance assignments, cf. Definition 7.5 on page 199
κ	Equivalence mapping, cf. Definition 7.7 on page 200
$\rho_{\hat{\omega}}$	Span number of $\hat{\omega}$, cf. page 202
g_{c}	Cardinality generator, cf. Definition 7.10 on page 209

INDEX