Motivating Agents in Unreliable Environments: A Computational Model

Patrick Krümpelmann, Matthias Thimm, Gabriele Kern-Isberner, and Regina Fritsch

Information Engineering Group, Department of Computer Science,
Technische Universität Dortmund, Germany

Abstract. The development of formal models for rational agents is mainly driven by the well-established BDI approach which divides an agent's mental state into beliefs, desires, and intentions. In this paper, we argue that motivation as well has to be taken into account in order to allow for a flexible and proactive behavior of intelligent agents in unreliable environments. In our approach, *motives* take the role of describing an agent's personality and are the driving force for creating desires and abandoning previously selected goals. We investigate the relationships between motives and their associated desires as well as the impact brought about by the uncertainty and unreliability of the environment.

1 Introduction

Today, formal approaches for representing the mental state of an intelligent agent mostly employ the BDI model, a framework that originated in psychology and describes intelligent behavior such as decision making, deliberation, and means-end reasoning in rational beings. This model divides a mental state into *beliefs*, *desires*, and *intentions* and gives a formal account for their interactions. Beginning with the work [1] many researchers in the field of artificial intelligence and intelligent agents applied this (informal) framework to formalize autonomous intelligent behavior [10].

However, the BDI model is an abstraction of a mental model of an intelligent being and given some sufficiently complex scenario it is not enough to represent proper decision making adequately. In the area of intelligent agents specifically the *desires* of an agent are oversimplified and often assumed to be initially given. A desire represents some state of the world such as "*I want to be rich*" or "*I want to get a professorship*" that the agent wishes to achieve. In contrast to *intentions* which represent a currently pursued course of action the set of desires does not need to be consistent, i. e. completely achievable, as the previously mentioned statements have shown. Furthermore, a single desire might be unachievable in general or unachievable in some specific context, e. g. due to some physical ("*I want to fly*") or mental ("*I want to learn all languages of the world*") limiting factors. Formal models employing the BDI model assume the desires of an agent to be given as agents are typically situated in some constrained environment with limited capabilities and tasks. Looking at the well-known cleaner robot the two desires "*I want to clean the room*" and "*I want to have a high battery level*" are completely sufficient to describe its world. Generally, however, assuming desires to be given inhibits real autonomous behavior. Bringing agents into more complex environments demands for mechanisms that allow an agent to set its desires by itself. Looking at how humans handle their desires, motivation theory [6] bridges the gap between a being's personality and its desires it wishes to satisfy. A *motive* such as "*benevolence*" or "*greed*" is a basic marker of an agent's personality that can be used to create (or abandon) some desire. We illustrate this intuition with a simple example.

Example 1. Given an agent competes with other agents for food, the motive "*greed*" would generate the desire of acquiring as much food as possible, while motive "*benevolence*" would generate the desire of acquiring just as much food as really needed and to ensure that other agents acquire as much food as they need. Both motives might very well be present in the agent's personality but with differing strengths.

In this paper, we develop a formal account for incorporating motivation into the BDI approach. Instead of assuming desires to be given we assume that an agent comprises of some set of basic motives that drives its behavior. Each motive of the agent is equipped with some weight and each motive is coupled with a set of desires that can be positively or negatively influenced by the motive. We give a formal account on the aggregation of the weights of the motives and these couplings in order to determine the desires the agent is motivated to follow. Furthermore, using the notion of *reliability* [4] we investigate how beliefs about actions (*know-how*) and beliefs about the world might influence the deliberation on motivation. More precisely, the motivation to follow some desire is strongly influenced by the uncertainty of the world and the knowledge of the agent that some course of action might not be reliable achievable.

The rest of this paper is organized as follows. In Section 2 we begin by giving some background on BDI agents and a formal account on representing motives and motivation that derives from psychology. In Section 3 we elaborate a simple agent model that integrates handling of motivation. We go on in Section 4 by giving a formal account on the dynamics of motivation and in Section 5 we consider the unreliability of the environment and discuss its influence on the motivational model of the agent. In Section 6 we review related work in Section 7 we conclude.

2 Agents and Motives

The BDI approach is a well-established approach to model rational agents. This model distinguishes between *beliefs*, *desires*, and *intentions* in order to represent human-like reasoning and behavior. In this model, beliefs represent the agent's (subjective) knowledge about itself, the world, and other agents. Desires describe what the agent is longing for in its environment and intentions account for its currently pursued goals and the intended course of action. Given some percept from the environment the agents usually employs some form of belief revision or update in order to incorporate the new information into its own beliefs. Afterwards, by taking the current state of the world into account, the agent considers its desires and selects some desire to be pursued as a *goal*. It appropriately updates its intentions using means-end reasoning and planning techniques [3], selects some course of action, and eventually performs some action in the environment. In general, this process is repeated indefinitely.

While most formalizations of the BDI model [10] assume the agent's desires to be given, a more natural as well as flexible and powerful approach, demands for the possibility to generate desires. The research in psychology describes how desires are created by an agent's motivation [5,6]. In what follows, we formalize the notion of motivation in order to incorporate it into a formal agent architecture. This allows for a more rational representation of an intelligent agent.

The theory of motivation is concerned with the question of how an agent determines its desires. Motivation is driven by *motives* which describe reasons for some specific behavior and are meant to be as basic as possible, e. g. hunger or love. A classification of basic motives is provided by *Maslow's hierarchy of needs* [5] that distinguishes between five levels of motives. From top to bottom motives in the corresponding levels are ordered by their importance and motives in higher levels are only *active* if some appropriate portion of lower level motives are *satisfied*. In its bottom level are the most basic motives, those for *physiological needs* such as health and food followed by *safety needs* (law and order), *love and belonging* (family, friends, love), followed by *esteem* (independence, respect) and finally *self-actualization* such as individuality. Motives of different levels have different susceptibilities to deficiency, in general more basic motives are more prone to deficiency, while top level motives may not be satisfiable. There are critical claims on the applicability of the hierarchy of needs for the human reasoning process due to the simplification of a strict hierarchy [9]. Nonetheless, we choose this conceptual framework to be used within our model as it provides a sufficient abstraction for the relationships of motives.

Let $(\mathcal{L}, \succ)$ be a totally ordered set of *motive levels* such that $L \succ L'$ with $L, L' \in \mathcal{L}$ means that motives on level L are more basic than motives on level L'. In the following we use $\mathcal{L}^M = (\{\mathsf{sa}, \mathsf{es}, \mathsf{lb}, \mathsf{sn}, \mathsf{pn}\}, \succ)$ with $\mathsf{pn} \succ \mathsf{sn} \succ \mathsf{lb} \succ \mathsf{es} \succ \mathsf{sa}$ (pn stands for *physiological needs*, sn for *safety needs*, etc. as listed above), thus using *Maslow's hierarchy of needs* to represent importance of basic motives. Our framework, however, is open to other types or quantities of motive levels. We also adapt the notion of deficiency needs by partitioning the set of motive levels into deficiency and non-deficiency levels. For $\mathcal{L}^M$ we define the deficiency levels as $\mathsf{d}(\mathcal{L}^M) = \{\mathsf{pn}, \mathsf{sn}, \mathsf{lb}\}$. This simplification of the above mentioned susceptibilities of motives might be generalized to a more granular or continuos representation. But it will be sufficient to enable the agent to focus on the deficiency needs in unreliable situations.

For our agents we assume some (finite) set Mot of basic motives. For every motive $m \in \mathsf{Mot}$ let $\mathsf{L}(m) \in \mathcal{L}$ denote the motive level of m.

The importance of individual motives and motive levels for an agent constitutes its personality. To measure importance of motives we employ the unit interval as the general range for weights[1]. In general, a smaller weight indicates a less important motive level.

Definition 1. *A* level weight range function wr *on* $\mathcal{L}$ *maps a motive level* $L \in \mathcal{L}$ *onto a continuous subset of the unit interval, i. e.* $wr(L) = [l_L, u_L]$ *with a lower and upper bound* $l_L, u_L \in [0, 1]$ *and* $l_L \leq u_L$*. We abbreviate* $\delta_L = u_L - l_L$*.*

For a motive level $L \in \mathcal{L}$ the value $wr(L) = [l_L, u_L]$ indicates that each motive belonging to L has at least an importance of l_L and at most an importance of u_L.

[1] Note, that our approach can be generalized to ranges represented by any totally ordered set; we choose to use the unit interval only for reasons of simplicity of presentation.

Example 2. For $\mathcal{L}^M$ an adequate level weight range function wr can be given by

$$\begin{aligned}
wr(\mathsf{pn}) &= [0.75, 1] && (\delta_{\mathsf{pn}} = 0.25)\\
wr(\mathsf{sn}) &= [0.55, 0.85] && (\delta_{\mathsf{sn}} = 0.3)\\
wr(\mathsf{lb}) &= [0.35, 0.65] && (\delta_{\mathsf{lb}} = 0.3)\\
wr(\mathsf{es}) &= [0.15, 0.45] && (\delta_{\mathsf{es}} = 0.3)\\
wr(\mathsf{sa}) &= [0, 0.25] && (\delta_{\mathsf{sa}} = 0.25)
\end{aligned}$$

Notice, that the weight ranges of motive levels might overlap in order to allow for situations where some less basic motive has a stronger influence than some more basic motive. We come back to this issue when taking the reliability of the environment into account.

These ranges describe the general importance of a motive level. The actual importance of a motive level at some point in time, called *level weight*, is given by a function $w : \mathcal{L} \to [0, 1]$ which maps a motive level $L \in \mathcal{L}$ to an element of its level weight range, i. e. $w(L) \in wr(L)$ for all $L \in \mathcal{L}$. Let W denote the set of all such functions w. While weight ranges are assumed to be fixed, level weights are subject to change when an agent acts in some environment and perceives new information about the world.

The weights of motive levels control how desires are created. The links between motives and desires are provided by *motive couplings*.

Definition 2. *A* motive coupling mc *is a tuple* (m, D, cs, ϕ) *with* $m \in \mathsf{Mot}$*, a desire* D*,* $cs \in [-1, 1]$*, and* ϕ *some sentence in the language of the beliefs of the agent. Let* $\mathcal{MC}$ *denote the set of all motive couplings. For a motive coupling* $mc = (m, D, cs, \phi)$ *we abbreviate* $D(mc) = D$*,* $\mathsf{L}(mc) = \mathsf{L}(m)$ *and* $\phi(mc) = \phi$*.*

A motive coupling (m, D, cs, ϕ) denotes some tendency of a motive m to influence the creation of a desire D positively ($cs > 0$) or negatively ($cs < 0$) with *coupling strength* cs if some statement ϕ can be verified in the beliefs of the agent. There, ϕ represents a condition that may trigger the coupling according to the given situation. That is, if for the current beliefs B, we find that $B \models \phi$, then the coupling between motive m and desire D is activated to the degree cs. Let $\mathcal{D}$ denote the set of desires that appear in some motive coupling.

Example 3. Let us consider the motive "*environmental awareness*". This motive is a strong influence on the desire "*save the whales*" with e. g. a coupling strength of 0.9 and a relatively weak influence for the desire "*buy fruits from your own country*" with e. g. a coupling strength of 0.3. Furthermore, it exhibits a negative influence on the desire "*buy a sports car*" with e. g. a coupling strength of -0.9.

Definition 3. *A* motive state M *is a tuple* $M = (\mathcal{M}, MC, wr, w)$ *with* $\mathcal{M} \subseteq \mathsf{Mot}$*,* $MC \subseteq \mathcal{MC}$*,* wr *is a level weight function on* $\mathcal{L}$*, and* $w : \mathcal{L} \to wr(\mathcal{L})$ *denotes the actual weight of each motive level. Let* Ω *denote the set of all motive states.*

We assume the beliefs of an agent to be represented in some logic. Let B be the belief state of an agent at some point in time with $B \in \mathcal{B}$ and $\mathcal{B}$ denotes the set of all possible belief states an agent may have. We assume that the agent maintains some form of structural knowledge on actions or *know-how* within its logical beliefs which is able to assess the reliability of achieving some desire $D \in \mathcal{D}$ in the given situation. We denote by $B \models \mathit{reliable}(D)$ that the current state of beliefs B gives reasonable grounds to believe that the desire D is reliably achievable, cf. [4]. For example, in a situation where the agent is rich the desire "*buy a Ferrari*" is reliably achievable whereas in a situation where the agent is poor it is not.

Definition 4. *Let $D_1, \ldots, D_n \in \mathcal{D}$ and $\mu_1, \ldots, \mu_n \in [-1, 1]$. Then the set of tuples $\gamma = \{(D_1, \mu_1), \ldots, (D_n, \mu_n)\}$ is called a* motivational structure. *Let Γ denote the set of all motivational structures.*

A motivational structure γ describes the motivation to follow the desires of an agent in some specific situation. For a tuple $(D, \mu) \in \gamma$ the value μ is called *motivational value* for D and represents the strength of the motivation. In general, a positive value of μ represents a positive motivation to follow D while a negative value represents motivations to not follow D. A zero value of μ describes a neutral motivation to follow the desire D.

3 An abstract model for motivated agents

As in the standard BDI approach the inner cycle of our formal agent model starts with incorporating newly received percepts into the belief using some belief revision function. As the world might have changed the motive state of an agent might change as well. In order to adapt to a changed world the agent has to reconsider the current weights and weight ranges of its motives.

Example 4. We continue Example 3. There, the motive "*environmental awareness*" lies in level 5 of Maslow's hierarchy of needs: self-actualization. The motive "*food*" lies in the lowest level of physiological needs. If the agent is low on food and low on money there are usually stronger grounds to prefer desires generated by motives in the lower levels. But if the situation changes, e. g. if the agents becomes wealthy, then the weights of the levels might change. Consider some generous level weight ranges, e. g $wr(\mathsf{pn}) = [0.4, 1]$ and $wr(\mathsf{sa}) = [0, 0.6]$. As the situation has become very *reliable* and the agent is not frightened about fulfilling its physiological needs the weights of the higher levels decrease and therefore motives for self-actualization can become even more motivated than motives for physiological needs.

The adjustment of the agent's motive state is performed by some *weight adjustment function* which determines a new level weight for the motives of the agent.

Definition 5. *Let Δ be a function $\Delta : \Omega \times \mathcal{B} \to W$ that determines for the motive state M in a situation described by beliefs B a new level weight function $\Delta(M, B)$.*

After adjusting the motive state of the agent with a new level function the motivational structure might be subject to change as well.

Example 5. We continue Example 4 and consider the desires "*buy cheap food*" and "*buy fruits from your own country*". Usually, in resource-bounded situations the first desire is preferred to the latter one. If the agent becomes wealthy the situation is secure enough to give "*buy fruits from your own country*" a higher motivation.

Definition 6. *Let Λ be a function $\Lambda : \Omega \times \mathcal{B} \rightarrow \Gamma$ that creates a new motivational structure $\Lambda(M, B)$ for some motive state M and beliefs B.*

Having adjusted the motivational structure the agent now has to decide which desires to follow. In general, the agent is best off in selecting its maximally motivated desires as goals, i. e. desires D with $(D, \mu) \in \gamma$ such that μ is maximal in γ. This is true for situations where the agent does not currently pursue any intention or the world has changed drastically and the agent has to reconsider its course of action completely. Usually, the agent is currently pursuing some intentions and switching to a new desire—because it is slightly more motivated as the current ones—does normally not make sense. The agent might start alternating between desires and switching to a new desire likely has further ramifications as it can be incompatible with currently pursued intentions which would need to be dropped. It is crucial to carefully deliberate on this decision as dropping an intention that has been pursued for some time may imply a considerable waste of resources. At this point we do not go into more details on the selection of desires but assume there is some mechanism to either select zero or more desires from the current motivational structure for pursuit. Figure 1 gives a rough overview on the agent model we developed in this section. In the figure, solid lines indicate action flow and dashed lines indicate information flow.

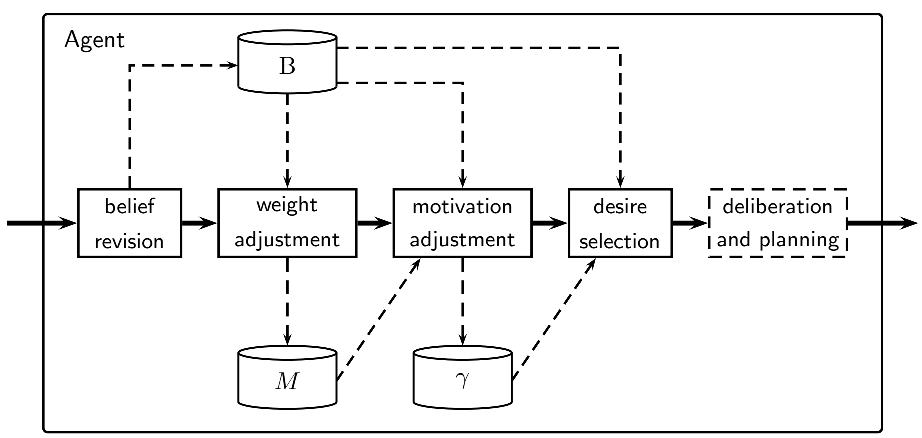

Fig. 1. A simple model of a motivated agent.

4 Adjusting motivation

In this section we elaborate on how the motivation adjustment function Λ of our agent model (see Definition 6) can be realized. In particular we present functions to combine components of the motive state and the belief of the agent to compute the motivational values of desires.

For the computation of the motivational value of a desire we have to combine the coupling strengths with the level weights for each motive coupling. Based on these basic motivations we need to consider the interaction of motivations by combining the basic motivations for the same desire from different motive couplings to determine the resulting motivational value.

For the computation of the basic motivations we introduce a function that combines coupling strengths with the level weights for a motive coupling. For this function from $[0,1] \times [-1,1]$ to $[-1,1]$ we demand the properties associativity, commutativity, monotony and having 1 as its neutral element. Hence, this function should be some kind of a t-norm (appropriately extended to the interval $[-1,1]$) and we picked the product as a sufficient candidate for the following definition.

Definition 7. *Given a motive coupling mc, its coupling strength $cs(mc)$ and level weight $w(mc)$ we define a function $\beta : [0,1] \times [-1,1] \to [-1,1]$ representing a* basic motivation *as:* $\beta(w(mc), cs(mc)) = w(mc) \cdot cs(mc)$.

Assuming positive values of the coupling strength, the influence of the level weight and the coupling strength is symmetric, for a level weight or coupling strength of 0 the resulting basic motivation will be 0. For a level weight of 1 the basic motivation is limited by the coupling strength and vice versa. For values of the level weight and coupling strength other than 0 and 1 the basic motivation is smaller than both, e. g., $0.5 \cdot 0.5 = 0.25$. For negative values of the coupling strength the behavior of the absolute value is the same but the resulting basic motivation is negative, thus acting against the realization of the associated desire.

For each desire there might exist several motive couplings and hence a set of basic motivations results for each desire. In order to determine the motivation value of a desire we combine the basic motivations for it. This combination function has to account for the nature of interaction of different motives for the same desire. In order to realize an adequate combination method, we base our approach for the combination of basic motivations on the parallel combination initially used for certainty factors in the MYCIN expert system [2] and using the commutative and associative aggregation function $f\colon (-1,1) \times (-1,1) \to (-1,1)$ defined as

$$f(x,y) = \begin{cases} x + y - x \cdot y, & \text{if } x, y > 0 \\ x + y + x \cdot y, & \text{if } x, y < 0 \\ \frac{x+y}{1-\min\{|x|,|y|\}}, & \text{else} \end{cases}$$

we define the motivation value $\mu(D)$ of a desire D as follows.

Definition 8. *Let $D \in \mathcal{D}$, let $mc_1, \ldots, mc_l$ be all motive couplings with $D(mc_i) = D$ for $i = 1, \ldots, l$, and let $m_i = \beta(w(mc_i), cs(mc_i))$. Then the* motivation value $\mu(D)$ *for a desire D is defined as*

$$\mu(D) = f(m_1, f(m_2, \ldots, f(m_{l-1}, m_l) \ldots))$$

for $l > 1$ and $\mu(D) = m_1$ otherwise.

The function $\mu(D)$ computes the motivation value for the desire D. For adjusting the motivational structure γ of the agent we compute the motivational value for each desire for which exists at least one active motive coupling.

Definition 9. *For a motive state $MS = (\mathcal{M}, MC, wr, w)$ and a belief state B we set*

$$\begin{aligned} \Lambda(MS, B) = \{(d, \mu(d)) \mid \exists mc \in MC : d = D(mc), \\ B \not\models d, B \models \phi(mc)\} \end{aligned}$$

The motivation structure γ contains the set of desires that are not satisfied and for whose there exists at least one active motive coupling, i. e., the condition of the coupling is satisfied together with their respective current motivational value.

Example 6. We continue Example 3 and instantiate the level weights according to the ranges given in Example 2 assuming a very reliable situation of the agent.

$$\begin{aligned} w(\mathsf{sa}) = 0.25 \in wr(\mathsf{sa}), \\ w(\mathsf{pn}) = 0.75 \in wr(\mathsf{pn}), \\ w(\mathsf{es}) = 0.45 \in wr(\mathsf{es}) \end{aligned}$$

For the motive couplings we formalize those of Example 3 and add the motive "*prestige*" with a positive coupling for "*buy a sports car*".

$$\begin{aligned} &(env\_awareness, save\_whales, 0.9, endangered\_whales), \\ &(env\_awareness, buy\_local\_fruits, 0.3, true), \\ &(evn\_awareness, buy\_sports\_car, -0.9, true), \\ &(prestige, buy\_sports\_car, 1, true) \end{aligned}$$

The beliefs of the agent shall be given by $B = \emptyset$. Given the motive state described above we can compute the corresponding motivational structure as follows. The desire "$save\_whales$" is contained in a motive coupling but as the agent is ignorant it does not believe that whales are endangered and consequently does not have a motivation for saving the whales. The desire "buy_local_fruits" has a motive coupling with level weight 0.25 and coupling strength 0.3. The resulting basic motivation is $\beta(0.25, 0.3) = 0.075$. As there are no other motive couplings for the desire the motivation value is equal to the basic motivation. The desire "buy_sports_car" has two motive couplings from different motives, one positive coupling and one negative. The resulting basic motivations are $\beta(0.25, -0.9) = -0.225$ and $\beta(0.45, 1) = 0.45$. From these we get the motivation value:

$$\mu(buy\_sports\_car) = f(-0.225, 0.45) = \frac{-0.225 + 0.45}{1 - 0.225}$$
$$= \frac{0.225}{0.775} \approx 0.29$$

Thus, the resulting motivation value for "buy_sports_car" is positive such that the motive coupling from the motive "*prestige*" is dominating the coupling from "*environmental awareness*", as expected. The resulting motivational structure of the agent is given by

$$\gamma = \Lambda(MS, B)$$
$$= \{(buy\_local\_fruits, 0.075), (buy\_sports\_car, 0.29)\}.$$

The motivational structure, whose creation is explained in this section, changes if the motive state is adjusted. The following section elaborates the computation of the motive state adjustment.

5 Taking Reliability into account

Agents usually behave differently in reliable vs. unreliable situations. Reliability makes the agent feel safe, whereas unreliability may signalize harassment, even danger. The degree of reliability of the current situation has direct impact on the motivational structure of the agent in our approach. As said before for every desire D we assume that the beliefs B of the agent is able to derive whether the desire is reliably achievable $B \models reliable(D)$. This means that the agent has the means (in form of plans and actions) to achieve the desire D in the current situation, cf. [4]. The belief on the reliability of all desires enables the agent to assess the reliability of the whole situation and gives an idea whether more basic motives, respectively desires, should be pursued instead of more higher motives, respectively desires. For example, if the agent is situated in some resource-bound location, say a desert, then usually a very small portion of the agent's desires is reliably achievable. In unreliable environments the agent should stick to desires that derive from more low-level motives, i. e. from motives that are in deficiency levels.

Definition 10. *Let m be a motive and let $K = \{(m, D_1, cs_1, \phi_1), \ldots, (m, D_l, cs_l, \phi_l)\}$ be the set of motive couplings for m with $cs_i > 0$ for $i = 1, \ldots, l$. Then m is called* satisfiable *wrt. beliefs B, denoted by $B \models sat(m)$, if there exists an $k \in \{1, \ldots, l\}$ desire D with*

$$cs_k \geq \frac{1}{l}\sum_{i=1}^{l} cs_i \quad \textit{and} \quad B \models reliable(D).$$

The above definition says that a motive m is satisfiable if there is a desire D that is coupled to m above average and that is reliable. Given the information on satisfiability of motives in a given situation the agent has to compare its set of motives with the set of motives that are satisfiable. In general, the more motives of an agent are satisfiable the more reliable the current situation is. In this comparison one has to consider the different motive levels as especially the deficiency motives have more weight in assessing the situation as reliable. Therefore, for a motive state $M = (\mathcal{M}, MC, wr, w)$ and beliefs B the reliability of the current situation with respect to the reliability of desires and hence the satisfiability of motives, denoted by rel_sit, is assessed by

$$rel\_sit = Z \cdot \sum_{L \in \mathcal{L}} \frac{l_L + u_L}{2} \frac{|ST_L|}{|S_L|} \tag{1}$$

with (remember that $\mathsf{L}(m) \in \mathcal{L}$ denotes the motive level of m)

$$\begin{aligned} ST_L &= \{m \in \mathcal{M} \mid \mathsf{L}(m) = L \wedge B \models sat(m)\} \\ S_L &= \{m \in \mathcal{M} \mid \mathsf{L}(m) = L\} \\ Z &= \frac{1}{\sum_{L \in \mathcal{L}} \frac{l_L + u_L}{2}} \quad . \end{aligned}$$

In Equation (1) the reliability of the situation is determined as a weighted sum over the proportions of satisfiable motives to the set of all motives in a given level. The weight of the proportions is defined to be the center of the level range which amounts to an average weight for the level. The term rel_sit is normalized by the factor Z and hence we have $0 \leq rel\_sit \leq 1$.

We modeled motivations according to Maslow's hierarchy in Section 2 distinguishing five types of motives of decreasing importance, i. e. the motives on the first level are the most basic ones. In order to implement Maslow's hierarchy, in Definition 1 we defined an interval of weights for each motive level L, assessing a lower bound l_L and an upper bound u_L for the importance of motives of level L, with δ_L denoting the width of the respective weight interval.

As a special feature of our approach, the current weight of the motive levels will depend on the reliability of the current situation: The more reliable the situation is, the more safe will the agent feel, and the less important are the motives of the motive levels that are deficiency needs. Consequently, the non-deficiency motives might obtain more influence. In unreliable situations, the basic motives are more important to sustain the agent's vital functionality. This is realized by adjusting the level weights $w(L)$ in the following way:

$$\begin{aligned} w(L) &= u_L + \delta_L f_{\downarrow}(rel\_sit), \quad L \in \mathsf{d}(\mathcal{L}^M), \\ w(L) &= u_L + \delta_L f_{\uparrow}(rel\_sit), \quad L \notin \mathsf{d}(\mathcal{L}^M), \end{aligned} \tag{2}$$

with a monotonically decreasing function $f_{\downarrow} : [0, 1] \rightarrow [0, 1]$ and a monotonically increasing function $f_{\uparrow} : [0, 1] \rightarrow [0, 1]$. Here, we set $f_{\downarrow}(y) = 1 - y$ and $f_{\uparrow}(y) = y$. Equations (2) realize a level weight adjustment function Δ as specified in Definition 5 and used in Definition 8 for computing the motivational values of desires. The influence of the reliability of the current situation might be modeled by families of more complex functions, even going beyond the dichotomy of deficiency vs. non-deficiency. We demonstrate the adjustment of level weights in the following example.

Example 7. We continue Example 5 and consider the two motives m_1 =“*environmental awareness*” and m_2 = “*food*”. As in Example 4 let the level weight ranges of the motive levels “*physiological needs*” and “*self-actualization*” be given by $wr(\mathsf{pn}) = [0.4, 1]$ and $wr(\mathsf{sa}) = [0, 0.6]$ and let $w(\mathsf{pn}) = 0.8$ and $w(\mathsf{sa}) = 0.4$. Consider the desires d_1 =“*buy fruits from your own country*” and d_2 =“*buy cheap food*” and the motive couplings $(m_1, d_1, 0.3, true)$ and $(m_2, d_2, 0.3, true)$. In this situation the motivational structure amounts to $\gamma = \{(d_1, 0.12), (d_2, 0.24)\}$. Imagine the above motivational structure derives from the situation that the agent is poor and he knows of no reliable way to acquire fruits from its own country. In that situation only the motive “*food*” is satisfied but not the motive “*environmental awareness*”. Consider now the situation that the agent becomes wealthy and does not have to worry the price of local fruits. Therefore, let the current beliefs B of the agent be of a form such that both $B \models reliable(d_1)$ and $B \models reliable(d_2)$. Therefore, both motives “*food*” and “*environmental awareness*” are satisfiable which amounts to

$$rel\_sit = \frac{1}{1}\left(\frac{0.4+1}{2}1 + \frac{0+0.6}{2}1\right) = 1$$

and it follows

$$w(\mathsf{pn}) = u_{\mathsf{pn}} + \delta_{\mathsf{pn}} f_{\downarrow}(rel\_sit) = 0.4 + 0.6(1-1) = 0.4$$
$$w(\mathsf{sa}) = u_{\mathsf{sa}} + \delta_{\mathsf{sa}} f_{\uparrow}(rel\_sit) = 0.0 + 0.6(1-0) = 0.6\,.$$

Computing the new motivational structure for the new level weights yields $\gamma = \{(d_1, 0.18), (d_2, 0.12)\}$. Therefore, in a reliable situation the desire “*buy fruits from your own country*” is more motivated than “*buy cheap food*”.

6 Related Work

There is a large body of established literature on motivation in the field of psychology and philosophy, see e. g. [6], while the literature in artificial intelligence and intelligent agents is rather limited. Nonetheless, on the conceptional level Norman and Long [8] also use motives and motivation. However, our approach goes further than theirs by adapting levels of motives and ideas from Maslow’s hierarchy for the BDI model. We also take the reliability of the environment for the agents motivational state under consideration. The implementation of the motivational model of [8] defines a new agent architecture called *motivated agency* that allows for the generation of motivated goals that are active if a threshold is met. The aim of their model is to limit the number of goals pursued by the agent and stands in parallel to the BDI model whereas our approach is fitted neatly into it [4] and complemented by other extensions such as explicit representation of structural knowledge. We also use a far more structured approach giving a framework of motive levels, their weights and couplings to desires that are used for goal selection.

Other work towards a computational model has been done by Luck and colleagues with in the publication [7]. The approach presented there works with one single structure of motivations. This structure updates the intensity of a given motivation that is aggregated over time until a predefined threshold is met which triggers a goal selection for this motivation and the mitigation of the motivation. The model we present here is far richer, more modular and flexible. We differentiate the concepts of motives, motivation and goal selection, add the levels of motives and consider the reliability of the environment.

7 Summary and Conclusion

The ability to generate desires and goals is a crucial feature of autonomous agents. In this paper, we presented a fully elaborated computational model that allows agents to be driven by motivations which are linked to their needs and desires and the influence of which depends on their current beliefs as well as on their self-evaluation. In summary, our model of a motivated BDI agent is based on the following principal ideas: 1.) we make use of a flexible hierarchy of motive types that roughly follows Maslow's model [5] and distinguishes between deficiency and non-deficiency needs, 2.) desires are linked to motives by couplings that are assigned some degrees of strength (positive or negative) and that can be triggered by conditions found to be true in the agent's belief about the world, and 3.) the interactions between motives on different levels of the hierarchy are processed by use of an aggregation function and give rise to the motivational structure of the agent that guides its current behavior.

In spite of the richness and complexity of our approach, the motives of the agent are clearly seen to be the basic driving component for the agents behavior. This enables the agent to act autonomously and on intrinsic incitement.

References

1. M. E. Bratman. *Intention, Plans, and Practical Reason*. CSLI Publications, 1987.
2. B. G. Buchanan and E. H. Shortliffe. *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*. Addison Wesley, 1984.
3. M. Ghallab, D. Nau, and P. Travers. *Automated Planning: Theory and Practice*. Morgan Kaufmann, 2004.
4. P. Krümpelmann and M. Thimm. A Logic Programming Framework for Reasoning about Know-How. In *Proc. of the 14th Int. Workshop NMR'10*, 2010.
5. A. H. Maslow. *Motivation and personality*. Harper and Row, Cambridge, MA, USA, 1970.
6. A. R. Mele. *Motivation and Agency*. Oxford University Press, 2003.
7. F. Meneguzzi and M. Luck. Motivations as an Abstraction of Meta-level Reasoning. In *Proc. of the 5th Int. Central and Eastern European Conf. on Multi-Agent Systems (CEEMAS'07)*, pages 204–214. Springer, 2007.
8. T. J. Norman and D. Long. Alarms: An implementation of motivated agency. In *Intelligent Agents II - Agent Theories, Architectures, and Languages*, pages 219–234. Springer, 1996.
9. A. Wahba and L. Bridgewell. Maslow reconsidered: A review of research on the need hierarchy theory. *Organizational Behavior and Human Performance*, 15:212–240, 1976.
10. G. Weiss, editor. *Multiagent systems*. MIT Press, 1999.