# Know-How for Motivated BDI Agents

# (Extended Abstract)

Matthias Thimm
Department of Computer Science
Technische Universität Dortmund
Germany
matthias.thimm@udo.edu

Patrick Krümpelmann
Department of Computer Science
Technische Universität Dortmund
Germany
patrick.kruempelmann@udo.edu

## ABSTRACT

The BDI model is well accepted as an architecture for representing and realizing rational agents. The beliefs in this model are focused on the representation of beliefs about the world and other agents and are widely independent from the agents intentions. We argue that also the representation of *know-how*, which captures the beliefs about actions and procedures, has to be taken into account when modeling rational agents. Using the notion of *know-how* as introduced by Singh we formalize and implement a concrete and usable agent architecture that supports and benefits from this representation of procedural beliefs in multiple ways. It also supports the representation of motivations that influence the agent's behavior. We thus enable the agent to reason about its planning capabilities in the same way as it can reason about any other of its beliefs by extending a BDI-based agent architecture to allow the representation of procedural beliefs explicitly as part of the agent's logical beliefs which again influences and enhances the agent's behavior.

## Categories and Subject Descriptors

I.2.11 [**Distributed Artificial Intelligence**]: Intelligent agents, Multiagent systems

## General Terms

Design, Theory, Human Factors

## Keywords

Multiagent System, BDI, Planning, Know-How, Motivation

## 1. INTRODUCTION

Planning and agency are two closely related fields in the research of multiagent systems and artificial intelligence in general. In this extended abstract we take a look at the planning capabilities of an agent from the perspective of knowledge representation by explicitly modeling these capabilities as logical beliefs of the agent in order to enable it to reason about and revise them in the same way as any other logical beliefs. We build upon the commonly known and used BDI model [7] that divides the mental state of an agent into *Beliefs*, *Desires*, and *Intentions*. In [5] Singh introduced the notion of *know-how* in order to relate the components *Beliefs* and *Intentions* more closely. Know-how describes the part of the logical beliefs of an agent that describes structural knowledge to reach certain goals. While in most modern agent architectures [1] beliefs influence the selection and achievement of intentions the other way around is not supported by these systems in general. Singh [5] claims, that "*since rational agency is intimately related to actions and procedures, it is important also to consider the form of knowledge that is about actions and procedures*". We take a first step towards the support and implementation of structural knowledge about procedures, i. e. know-how, in a concrete agent architecture. We support our formal work by an implementation of our know-how formalism in the full fledged BDI multiagent system KiMAS [2, 6] which also supports motivations [4] that may influence goal adoption and the overall behavior of the agent.

We proceed as follows. In Section 2 we give some motivation and continue with a short overview of our formal proposal of know-how in Section 3. In Section 4 we conclude. The technical paper [6] elaborates the ideas introduced here in more detail.

**Cite as:** Know-How for Motivated BDI Agents, (Extended Abstract), Matthias Thimm, Patrick Krümpelmann, *Proc. of 8th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2009)*, Decker, Sichman, Sierra and Castelfranchi (eds.), May, 10–15, 2009, Budapest, Hungary, pp. 1143–1144


## 2. BELIEFS AND KNOW-HOW

The BDI model [7] distinguishes between *Beliefs*, *Desires*, and *Intentions* as the main elements of an agent's mind. Informally speaking, *Beliefs* comprise the agent's beliefs about the world and its current situation, *Desires* represent what it wishes to achieve and hence represent possible goals, whereas *Intentions* model the agent's immediate (sub-)goals and thus focus on the next actions.

Know-how as introduced by Singh [5] is an extension to the traditional BDI model. In [5] it has been argued that on a descriptive layer an explicit distinction between beliefs about the world (know-that) and beliefs about how to achieve certain intentions (know-how) is indispensable in order to model agents and their behavior in a natural way. In modern formalizations of the BDI model or other agent architectures for planning and reasoning [1, 3] planning and beliefs components are mostly kept separate. Although beliefs do (of course) influence intention deliberation and goal generation, the other way round, namely reasoning about the current intentions of the agent and especially reasoning about the capability of how to achieve some state cannot be formalized in a natural way in these systems. Particularly, the pure knowledge of the possibility to achieve a certain intention is crucial for the agent in order to determine if the intention can be pursued or has to be dropped. In general, there are many situations in which it is necessary for the agent to reason about its planning capabilities.

In [5] Singh develops a temporal logic in which know-that and know-how can be specified in order to verify whether an agent is able to achieve certain goals and whether it knows how to achieve it. Here, we take the general idea of know-how one step further on the way from a specification and verification tool towards a programming tool. To our knowledge the concept of know-how has not been developed further since Singh's publications in the late 1990s while planning and intention generation are active fields including some ideas of how to combine state-of-the-art knowledge representation and reasoning techniques with problem solving and planning, e. g. [3]. We introduce know-how as a formalization of an agent's planning capabilities in a declarative manner and show how the agent can use this representation to reason about it.

## 3. KNOW-HOW

The know-how of an agent is structured in *know-how statements* which is the atomic form of a structural piece of information. A *know-how statement* $\sigma$ is a tuple

$$\sigma = (a, (s_1, \ldots, s_n), \{c_1, \ldots, c_m\})$$

with (abstract) goals $a, s_1, \ldots, s_n$ and (abstract) conditions $c_1, \ldots, c_m$. The informal meaning of such a know-how statement is as follows: In order to achieve goal $a$ the agent can try to achieve the subgoals $s_1, \ldots, s_m$, if the conditions $c_1, \ldots, c_m$ are fulfilled with respect to the agent's beliefs. A *know-how base* $\Sigma$ is a set of know-how statements and can be seen as a subset of the agent's logical beliefs.

Given a know-how base $\Sigma$, an agent has the means to achieve a given intention by iteratively querying $\Sigma$ for the needed intentions. We give formal semantics to this intuitive meaning using state transition systems and a sound and complete procedure that implements this semantics in the full paper [6].

One of the main motivations for the explicit representation of know-how and intentions in the agents beliefs is that this enables the agent to reason about these. To be more precise, the information reflected by the know-how is useful for the selection of new goals as the achievement of these is critically dependent on know-how. The structure of know-how also reveals information of the involved conditions and subgoals of the achievement of goals. This information enables the agent to reason about the feasibility and effort as well of the reliability of the achievement of goals. The just described way of reasoning interacts with the agents motivations as these can be dependent on the current level of confidence of the agent in a given situation.

As an example, an agent must be capable of determining if it has the means to achieve a given intention. Let $\Sigma$ be a know-how base. An intention $I$ is *achievable* in $\Sigma$ if

- $I$ is an atomic intention, i. e. $I$ can be fulfilled by an atomic action, or
- there is at least one know-how statement $\sigma \in \Sigma$ with $goal(\sigma) = I$ and every sub-target of $\sigma$ is achievable.

$\Sigma$ is called *sound* if every intention $I$ is achievable.

The definition of sound know-how does not take the conditions of a know-how statement into consideration when determining if an intention is achievable. In order to regard these conditions we also discuss a notion called *reliable know-how*. In our framework reliability of know-how is understood as the robustness of know-how to the incompleteness of information. Reliable know-how is know-how which is known not to fail given the incomplete information available. Let us assume that actions cannot fail, then atomic intentions are *reliably achievable* by definition. For complex intentions *reliability* is recursively defined using the reliability of its subcomponents and a *context* $C$, which is a set of conditions. The intention $I$ is *reliably achievable in* $C$ given a know-how base $\Sigma$ iff

- $I$ is an atomic intention or
- there is at least one know-how statement $\sigma$ with target $I$ and $\sigma$ is reliable in context $C$.

A know-how statement $\sigma$ with sub-targets $I'_1, \ldots, I'_n$ is *reliable in context* $C$ iff

1. each sub-target $I'_i$ $(1 \leq i \leq n)$ of $\sigma$ is reliably achievable in $C \cup \{I_1, \ldots, I_{i-1}\}$ and
2. the conditions of $\sigma$ are fulfilled in $C$.

The treatment of know-how as logical beliefs and the procedures that check the agent's know-how for soundness and reliability can be implemented in logic programming [6].

## 4. CONCLUSION

By explicitly representing an agent's procedural beliefs within its logical beliefs, the agent acquires the capability to reason about its current state of plan deliberation and enables it to treat this kind of beliefs in the same way as its other beliefs. We see this proposal as a first step to the full support of the notion of know-how in a concrete logic-based agent architecture. This constitutes an enhancement of the agents reasoning capabilities as well as it improves the interplay of the agents components. For future work, we plan to exploit and extend the new possibilities opened by our work in terms of reasoning with and about know-how as well as the further integration of agent components.

## 5. REFERENCES

[1] R. H. Bordini, L. Braubach, M. Dastani, A. E. F. Seghrouchni, J. J. Gomez-Sanz, J. Leite, G. O'Hare, A. Pokahr, and A. Ricci. A survey of programming languages and platforms for multiagent systems. *Informatica*, 30:33–44, 2006.

[2] P. Krümpelmann, M. Thimm, M. Ritterskamp, and G. Kern-Isberner. Belief operations for motivated BDI agents. In *Proc. of AAMAS'08*, pages 421–428, 2008.

[3] V. Lifschitz. Answer set programming and plan generation. *Artif. Intel.*, 138(1-2):39–54, June 2002.

[4] M. Luck and M. d'Inverno. Motivated behaviour for goal adoption. In *Proc. of the 4th Australian Workshop on Distributed Artificial Intelligence*, pages 58–73, 1998.

[5] M. P. Singh. Know-how. In A. S. Rao and M. J. Wooldridge, editors, *Foundations of Rational Agency, Applied Logic Series*, pages 105–132. Kluwer, 1999.

[6] M. Thimm and P. Krümpelmann. Know-how for motivated BDI agents (extended version). Technical report, Technische Universität Dortmund, 2009.

[7] G. Weiss, editor. *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*. MIT Press, Cambridge, MA, USA, 1999.