Modelling Animal Behaviour Based on Interpretation of Another Animal's Behaviour

Tibor Bosse (tbosse@few.vu.nl) Zulfiqar A. Memon (zamemon@few.vu.nl) Jan Treur (treur@few.vu.nl)

Vrije Universiteit Amsterdam, Department of Artificial Intelligence
De Boelelaan 1081, NL-1081 HV, Amsterdam, The Netherlands

Abstract

For certain animals, the capability to interpret and anticipate on another animal's behaviour may be crucial for survival. To this end, as is often claimed informally, an animal may apply a Theory of Mind to estimate what the other animal has on its mind. This paper uses a formal BDI-based agent model for Theory of Mind to formalise and simulate such a situation. The model uses BDI-concepts to describe a form of metacognition: a cognitive process of an agent about the cognitive process of another agent, which is also based on BDI-concepts. This paper explores whether this formal model is applicable to certain animal species. A specific case study is addressed, which involves the scenario of a prey that manipulates the behaviour of a predator. For this scenario, simulation experiments have been performed, and their results are discussed.

Introduction

For certain animals, to function effectively in interaction with other animals, it is useful if they are able to interpret, estimate and anticipate on potential behaviour of animals around it. It is often assumed that this requires metacognition in some form of Theory of Mind (Baron-Cohen, 1995; Bogdan, 1997; Malle, Moses, and Baldwin, 2001). Such a Theory of Mind can be exploited by an animal in two different manners. The first manner is just to predict the behaviour in advance, in order to be prepared that it will occur. A second manner is to affect the occurrence of behaviour by manipulating the occurrence of circumstances that are likely to lead to it.

One of the ways to model an agent B exploiting a Theory of Mind about an agent A is to use a BDI-model (based on beliefs, desires and intentions) to describe agent A's cognitive processes and actions. To model the agent B's own behaviour a BDI-model can be used as well; in this way within agent B's cognitive processes, at two levels BDI-models play a role. This type of model will be exploited in this paper to model the behaviour of higher animals such as primates and dogs. For example, for agent B the desire is generated that agent A will not perform the action to kill B, and that agent A will in particular not generate the desire or intention to do so. Based on this desire of B, for example, the refined desire of B can be generated that agent A will not believe that agent B is reachable. Based on the latter desire, an intention and action can be generated to present circumstances to agent A that will make A believe that B is not reachable.

The vehicle used to model the two-level BDI-model is the modelling language LEADSTO (Bosse, Jonker, Meij, and Treur, 2007). In this language, direct temporal dependencies between two state properties in successive states are modelled by *executable dynamic properties*. The LEADSTO format is defined as follows. Let α and β be state properties of the form 'conjunction of ground atoms or negations of ground atoms'. In the LEADSTO language the notation $\alpha \twoheadrightarrow_{e, f, g, h} \beta$, means:

If state property α holds for a certain time interval with duration g, then after some delay (between e and f) state property β will hold for a certain time interval of length h.

Here, atomic state properties can have a qualitative, logical format, such as an expression desire(d), expressing that desire d occurs, or a numerical format such as an expression has_value(x, v) which expresses that variable x has value v.

In this paper, first the general BDI-model is explained. This BDI-model is illustrated by a case study about a predator that desires to kill a prey. The next section describes how the simple model can be extended to a two-level BDI-model of an agent that also involves another agent's BDI-model. This two-level BDI-model is illustrated by a case study that elaborates upon the previous example: it addresses the scenario of a prey that has metacognition addressing analysis of the behaviour of a predator, and prevents being attacked. Based on this model, some simulation experiments and their results are discussed.

The BDI-Model

The BDI-model bases the preparation and performing of actions on beliefs, desires and intentions (e.g., Georgeff and Lansky, 1987; Jonker, Treur, and Wijngaards, 2003; Rao and Georgeff, 1991; 1995). This model shows a long tradition in the literature, going back to Aristotle's analysis of how humans (and animals) can come to actions; cf. (Aristotle, 350 BCa; 350BCb). He discusses how the occurrence of certain internal (mental) state properties within the living being entail or cause the occurrence of an action in the external world. Based on this, Aristotle introduced the following pattern to explain action (called practical syllogism):

If	A has a desire D
and	A has the belief that X is a (or: the best) means to achieve D
then	A will do X

The BDI-model incorporates such a pattern to explain behaviour in a refined form. Instead of a process from desire to action in one step, as an intermediate stage first an intention is generated, and from the intention the action is generated. Thus the process is refined into a two-step process. See Figure 1 for the generic structure of the BDI-model in causal-graph-like style, as often used to visualise LEADSTO specifications. Here the box indicates the borders of the agent, the circles denote state properties, and the arrows indicate dynamic properties expressing that one state property leads to (or causes) another state property. In this model, an action is performed when the subject has the intention to do this action and it has the belief that certain circumstances in the world are fulfilled such that the opportunity to do the action is there. Beliefs are created on the basis of observations. The intention to do a specific type of action is created if there is some desire D, and there is the belief that certain circumstances in the world state are there, that make it possible that performing this action will fulfil this desire (this is the kind of rationality criterion discussed above; e.g., what is called means-end analysis is covered by this). Whether or not a given action is adequate to fulfil a given desire depends on the current world state; therefore this belief may depend on other beliefs about the world state. Instantiated relations within the general BDI-model as depicted by arrows in graphical format in Figure 1 can be specified in formal LEADSTO format as follows:

desire(D) ∧ belief(B1) ↠ intention(P)
intention(P) ∧ belief(B2) ↠ performs(P)

with appropriate desire D, action P and beliefs B1, B2. Note that the beliefs used here both depend on observations, as shown in Figure 1. Furthermore, ∧ stands for the conjunction operator (and) between the atomic state properties (in the graphical format denoted by an arc connecting two (or more) arrows). Often, dynamic properties in LEADSTO are presented in *semi-formal* format, as follows:

At any point in time
if desire D is present
and the belief B1 is present
then the intention for action P will occur

At any point in time
if the intention for action P is present
and the belief B2 is present
then the action P will be performed

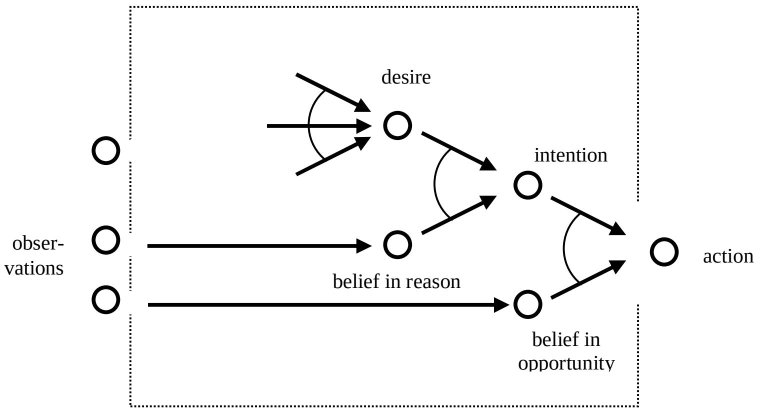

Figure 1: Structure of the general BDI-model.

As a generic template, including a reference to the agent X concerned, this can be expressed by:

For any desire D, world state property Z, and action Y such that has_reason_for(X, D, Z, Y) holds:

desire(X, D) ∧ belief(X, Z) ↠ intention(X, Y)

For any world state property Z and action Y such that is_opportunity_for(X, Z, Y) holds:

intention(X, Y) ∧ belief(X, Z) ↠ performs(X, Y)

Here has_reason_for(X, D, Z, Y) is a relation that can be used to specify which state property Z is considered a reason to choose a certain intention Y for desire D. Similarly is_opportunity_for(X, Z, Y) is a relation that can be used to specify which state property Z is considered an opportunity to actually perform an intended action Y.

Assuming that beliefs are available, what remains to be generated in this model are the desires. For desires, there is no generic way (known) in which they are to be generated in the standard model. Often, in applications, generation of desires depends on domain-specific knowledge.

A BDI-Model for Animal Behaviour

To illustrate the BDI-model described above by a specific example, a specific scenario is addressed, in the domain of a predator that wants to attack a prey. This scenario was inspired by (Bogdan, 1997), who introduces the notion of a *goal setting for interpretation* (i.e., a situation in which an organisms needs to interpret the behaviour of another organism in order to satisfy its private goals), which he illustrates as follows:

> ‘To illustrate, suppose that organism *A* (interpreter) has a private goal (say resting). It interferes with the goal of another organism *S* (subject), which is to eat *A*. Those *A*-type organisms will be selected who manage to form the social or *S*-regarding goal of avoiding the nasty type *S* by countering their inimical behavior, say by threat or deception. The latter goal in turn selects for interpretation, specifically, for interpretation goals such as desire identification and behavior prediction. Those *A*-type organisms are selected who form and reach such interpretation goals. The environment that selected for such accomplishments is a goal setting *of a certain kind*, say of behavior manipulation by behavior prediction and desire identification. There could be as many kinds of goal settings for interpretation as there are interpretation goals and tasks to achieve them, and hence as many skills.’ (Bogdan, 1997), p. 111

Based on this description, a scenario is considered that involves a predator (agent A) and a prey (agent B). Assume that, under certain circumstances, the predator tries to kill the prey, and the prey tries to avoid this by manipulation. First, only the behaviour of the predator is addressed (in which no Theory of Mind is involved). However, in a later section, the cognitive process of the prey involving Theory of Mind is addressed as well. Using the BDI-model as introduced above, the example is made more precise as follows. The *desire* to eat the prey is created after time t by the predator if the following holds at time t:

- the predator has the belief that the prey is alone
  (i.e., not surrounded by other animals)

The *intention* to kill the prey is generated after time t if the following holds at time t:

- the predator has the desire to eat the prey
- the predator has the belief that the prey is weak
  (i.e., that it does not show strong, aggressive behaviour)

The *action* to kill the prey is generated after time t if the following holds at time t:

- the predator has the intention to kill the prey
- the predator has the belief that the prey is slow
  (i.e., that it does not run very fast, so that it can be caught)

Using the generic template discussed, via the relations

has_reason_for(predator, eat_prey,
not(prey_shows_aggressive_behaviour), kill_prey)
is_opportunity_for(predator, not(prey_runs_fast), kill_prey)

the following model for agent predator is obtained:

belief(predator, not(prey_surrounded_by_other_animals)) ↠ desire(predator, eat_prey)
desire(predator, eat_prey) ∧ belief(predator, not(prey_shows_aggressive_behaviour)) ↠ intention(predator, kill_prey)
intention(predator, kill_prey) ∧ belief(predator, not(prey_runs_fast)) ↠ performs(predator, kill_prey)

The Two-Level BDI-Model

According to the intentional stance (Dennett, 1987, 1991), an agent is assumed to decide to act and communicate based on intentional notions such as beliefs about its environment and its desires and intentions. These decisions, and the intentional notions by which they can be explained and predicted, generally depend on circumstances in the environment, and, in particular, on the information on these circumstances just acquired by interaction (i.e., by observation and communication), but also on information acquired by interaction in the past. To be able to analyse the occurrence of intentional notions in the behaviour of an observed agent, the observable behavioural patterns over time form a basis; cf. (Dennett, 1991).

In the model presented in this paper, the instrumentalist perspective is taken as a point of departure for a Theory of Mind. More specifically, the model describes the cognitive process of an agent B that applies the intentional stance to another agent A by attributing beliefs, desires and intentions. Thus, for agent B a Theory of Mind is obtained using concepts for agent A's beliefs, desires and intentions. For example, in case a prey (agent B) fears to be attacked by a predator (agent A), it may analyse in more detail under which circumstances the predator may generate the desire and intention to attack.

As a next step, the model is extended with BDI-concepts for agent B's own beliefs, desires and intentions as well. By doing this, agent B is able to not only *have* a theory about the mind of agent A, but also to *use* it within its own BDI-based cognitive processes to generate its actions. To this end, a number of meta-representations expressed by meta-predicates are introduced, e.g.:

belief(B, desire(A, D))

This expresses that agent B believes that agent A has desire D.

desire(B, not(intention(A, X)))

This expresses that agent B desires that agent A does not intend action X.

belief(B, depends_on(performs(A, X), intention(A, X)))

This expresses that agent B believes that, whether A will perform action X depends on whether A intends to do X. Note that the third meta-statement has a more complex structure than the other two, since it represents a statement about a *dynamic property*, rather than a statement about a *state property*. These dependencies can be read from a graph such as depicted in Figures 1 and 2 (right hand side). For example, it is assumed that agent B knows part of this graph in his Theory of Mind, expressed by beliefs such as:

belief(B, depends_on(performs(A, X), intention(A, X)))
belief(B, depends_on(performs(A, P), belief(A, B2)))
belief(B, depends_on(intention(A, P), desire(A, D)))
belief(B, depends_on(intention(A, P), belief(A, B1)))
belief(B, depends_on(desire(A, D), belief(A, B3)))
belief(B, depends_on(belief(A, X), hears(A, X)))

Desire refinement in the BDI-model for an agent B attributing motivations to an agent A is formulated (in LEADSTO format) by:

desire(B, X) ∧ belief(B, depends_on(X, Y)) ↠ desire(B, Y)
desire(B, X) ∧ belief(B, depends_on(X, not(Y))) ↠ desire(B, not(Y))
desire(B, not(X)) ∧ belief(B, depends_on(X, Y)) ↠ desire(B, not(Y))
desire(B, not(X)) ∧ belief(B, depends_on(X, not(Y))) ↠ desire(B, Y)

Moreover the following schemes for intention and action generation are included in the model. For any desire D, world state property Z, and action Y such that has_reason_for(B, D, Z, Y) holds:

desire(B, D) ∧ belief(B, Z) ↠ intention(B, Y)

For any world state property Z and action Y such that is_opportunity_for(B, Z, Y) holds:

intention(B, Y) ∧ belief(B, Z) ↠ performs(B, Y)

Moreover, some dynamic properties of the world are needed:

performs(B, Y) ∧ has_effect(Y, C) ↠ holds_in_world(C)
holds_in_world(C) ↠ observes(A, C)

For an overview of the complete two-level BDI-model, see Figure 2.

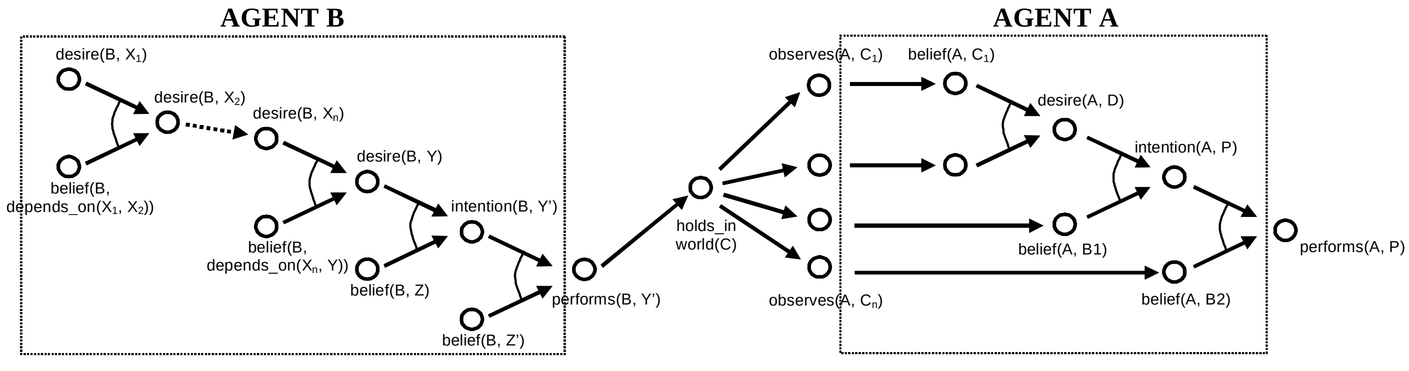

Figure 2: Structure of the two-level BDI-model.

A Two-Level BDI-Model for Animal Behaviour

The above model was used to describe how the prey agent (from the case described earlier) acts in an anticipatory manner to avoid the predator's desire, intention and/or action to occur. The initial desire of the prey is that the predator does not perform the action to kill it:

desire(prey, not(performs(predator, kill(prey))))

Fulfilment of this desire can be obtained in the following three manners:

Avoiding the predator's desire to occur

This can be obtained when the predator observes that the prey is surrounded by other animals. This will make the condition in the predator's desire generation as described earlier fail.

Avoiding the predator's intention to occur (given that the desire occurs)

This can be obtained by refutation of the belief that plays the role of the reason to generate the intention in the predator's intention generation as described earlier, i.e., the belief that the prey is weak (and does not show aggressive behaviour).

Avoiding the predator's action to occur (given that the intention occurs)

This can be obtained by refutation of the belief that plays the role of opportunity in the predator's desire action as described, i.e., the belief that the prey is slow (and does not run fast).

For convenience, the model does not make a selection but addresses all three options to prevent the killing action. This means that the prey generates *desires* for:

- The predator observes that the prey is surrounded by other animals
  observes(predator, prey_surrounded_by_other_animals)
- The predator observes that the prey shows aggressive behaviour
  observes(predator, prey_shows_aggressive_behaviour)
- The predator observes that the prey runs fast
  observes(predator, prey_runs_fast)

To fulfil these desires, *intentions* are to be generated by the prey to actions such as:

- call for help of other animals: call_for_help
- show aggressive behaviour: show_aggressive_behaviour
- run fast: run_fast

Reasons for the prey to choose for these intentions are *beliefs* in, respectively:

- The predator is paying attention to the prey's gaze (so that it will notice it when the prey calls for help of other animals)
  predator_is_noticing_preys_gaze
- The predator is paying attention to the prey's gesture (so that it will notice it when the prey shows aggressive behaviour)
  predator_is_noticing_preys_gesture
- The predator is at a reasonable distance away (so that it is able to run away without being caught)
  predator_is_reasonable_distance_away

Moreover, the intentions of the prey can lead to the corresponding actions when the following *beliefs* of the prey in *opportunities* are there:

- Other animals are around (so that it is possible to call for their help)
  other_animals_around
- The predator is about to attack (so that it is possible to show aggressive behaviour)
  predator_about_to_attack
- No obstacle is blocking the escape route of the prey (so that it is possible to run away)
  no_obstacle

In addition to the generic BDI-model shown before, the following specific relations were used to model the case study:

belief(prey(depends_on(performs(predator, kill(prey)), intention(predator, kill(prey))))
belief(prey(depends_on(performs(predator, kill(prey)), not(belief(predator, prey_runs_fast))))
belief(prey(depends_on(intention(predator, kill(prey)), desire(predator, eat(prey))))
belief(prey(depends_on(intention(predator, kill(prey)), not(belief(predator, prey_shows_aggressive_behaviour))))
belief(prey(depends_on(desire(predator, eat(prey)), not(belief(predator, prey_surrounded_by_other_animals))))
belief(prey(depends_on(belief(predator, prey_surrounded_by_other_animals), observes(predator, prey_surrounded_by_other_animals)))
belief(prey(depends_on(belief(predator, prey_shows_aggressive_behaviour), observes(predator, prey_shows_aggressive_behaviour)))
belief(prey(depends_on(belief(predator, prey_runs_fast), observes(predator, prey_runs_fast)))

has_reason_for(prey, observes(predator, prey_surrounded_by_other_animals), predator_is_noticing_preys_gaze, call_for_help)
has_reason_for(prey, observes(predator, prey_shows_aggressive_behaviour), predator_is_noticing_preys_gesture, show_aggressive_behaviour)
has_reason_for(prey, observes(predator, prey_runs_fast), predator_is_reasonable_distance_away, run_fast)

is_opportunity_for(prey, other_animals_around, call_for_help)
is_opportunity_for(prey, predator_about_to_attack, show_aggressive_behaviour)
is_opportunity_for(prey, no_obstacle, run_fast)

has_effect(call_for_help, prey_surrounded_by_other_animals)
has_effect(show_aggressive_behaviour, prey_shows_aggressive_behaviour)
has_effect(run_fast, prey_runs_fast)

By combining these relations with the generic LEADSTO rules provided in the previous section, a complete executable LEADSTO specification for the two-level BDI-model has been created. This simulation model is shown in the appendix at http://www.cs.vu.nl/~tbosse/tom/ICCM.pdf.

Simulation Experiments

In simulation experiments, the two-level BDI-model has been applied to the case study as described above. To this end, the LEADSTO software environment (Bosse, Jonker, Meij, and Treur, 2007) has been used. In Figure 3 and 4, examples of resulting simulation traces are shown. In these figures, time is on the horizontal axis; the state properties are on the vertical axis. The dark boxes indicate that a state property is true. Note that, due to space limitations, only a selection of the relevant atoms is shown.

Figure 3 is the resulting simulation trace of the situation in which *no* Theory of Mind is involved, i.e., only the behaviour of the predator is addressed, without manipulation by the prey. The trace depicts that the predator initially receives some inputs (e.g., indicated by the state property

observes(predator, not(prey_surrounded_by_other_animals))

at time point 1).

As a result, the predator has made some beliefs (e.g., the state property

belief(predator, not(prey_surrounded_by_other_animals))

at time point 2), which persists for a longer time. Due to this belief, it generates the desire to eat the prey at time point 3

desire(predator, eat(prey))

Based on this desire and the belief

belief(predator, not(prey_shows_aggressive_behaviour))

the predator generates the intention to kill the prey at time point 4:

intention(predator, kill(prey))

Based on this intention and the belief

belief(predator, not(prey_runs_fast))

the predator eventually performs the action of killing the prey at time point 5.

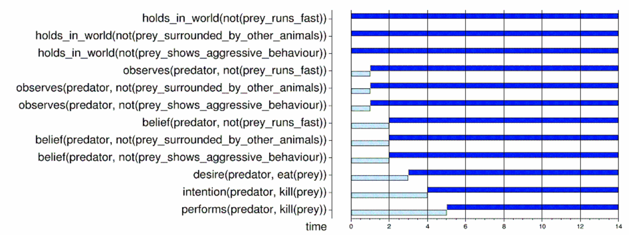

Figure 3: Simulation trace of the predator's behaviour

Figure 4 is the resulting simulation trace of the extended case study, in which the prey agent can act in an anticipatory manner to avoid the predator's desire to eat the prey, and intention and/or action to kill it. Figure 4 shows, among others, that the prey initially desires that the predator does not perform the action to kill it:

desire(prey, not(performs(predator, kill(prey))))

Based on this, the prey eventually generates a number of more detailed desires about what the predator should observe (see, for example, the state property

desire(prey, observes(predator, prey_shows_aggressive_behaviour))

at time point 3). Next, the prey uses these desires to generate some intentions to fulfill these desires (e.g., the state property

intention(prey, show_aggressive_behaviour)

at time point 4). Eventually, when the opportunities are there, these intentions are performed, and the predator observes some new inputs (e.g., the state property

observes(predator, prey_shows_aggressive_behaviour)

at time point 8). As a result, the predator eventually does not generate the action to kill the prey.

Note that in the scenario sketched in Figure 4, the prey takes all possible actions (within the given conceptualization) to fulfill its desires. This is a rather extreme case, since according to the prey's BDI-model, modifying only one of the predator's inputs will be sufficient to make sure that it does not kill the prey. Other traces can be generated in which the prey takes fewer actions to fulfill its desires.

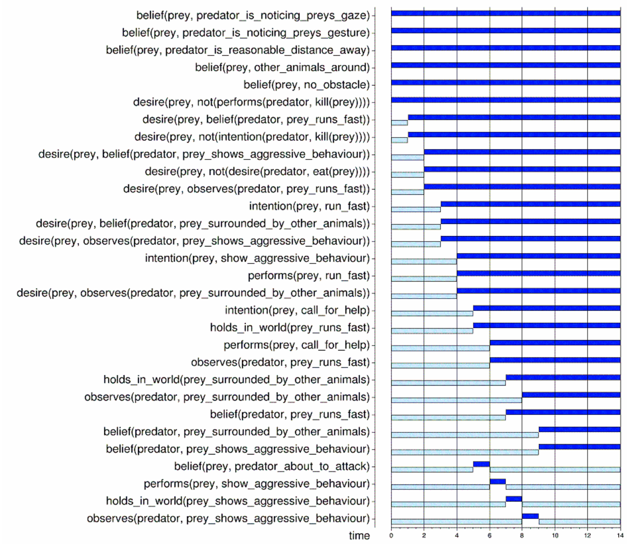

Figure 4: Simulation trace of the prey's manipulation of the predator's behaviour.

Discussion

In order to function well in interaction with other agents, it is very helpful for an agent to have capabilities to predict in which circumstances the agents in its environment will show certain behaviours. To this end, such an agent will have to perform interpretation based on a Theory of Mind (Baron-Cohen, 1995). This type of metacognition is studied in the context of human social interaction, but also in the area of animal behaviour it is addressed; e.g., (Barrett and Henzi, 2005; Bogdan, 1997; Heyes, 1998). In this paper the latter area is addressed. A model for Theory of Mind is applied, which makes use of BDI-concepts at two different levels. First, the model uses BDI-concepts *within* the Theory of Mind (i.e., it makes use of beliefs, desires and intentions to describe the cognitive process of another agent). Second, it uses BDI-concepts for *interpretation of* the Theory of Mind (i.e., it makes use of beliefs, desires and intentions to describe an agent's meta-cognition about the cognitive process of another agent). At this second level, meta-statements are involved, such as 'B believes that A desires d' or 'B desires that A does not intend a'. These meta-statements are about the states occurring within the other agent. In addition, meta-statements are involved about the dynamics occurring within the other agents. An example of such a (more complex) meta-statement is 'B believes that, if A performs a, then earlier he or she intended a'.

The two-level BDI-based model as presented can be exploited both in order to be prepared for the behaviour of another agent, and in order to affect the behaviour of another agent at forehand. The model has been formalised using the modelling language LEADSTO, which describes dynamics in terms of direct temporal dependencies between state properties in successive states. The model not only addresses analysis of the other agent's beliefs, desires and intentions, but also integrates this with the agent's own beliefs, desires and intentions, and actions.

Obviously, empirical validation of the model is a difficult issue. At least, the present paper has indicated that it is possible to apply computational models for Theory of Mind to animal behaviour. Moreover, the model indeed shows the anticipatory behaviour of higher animals as described in literature such as (Bogdan, 1997). In this sense the model has been validated positively. However, notice that this is a relative validation, only with respect to the literature that forms the basis of the model. In cases that the available knowledge about the functioning of such animals is improving, the model can be improved accordingly. In this sense the approach anticipates further development.

Concerning related work, there is a large body of literature on Theory of Mind in non-human primates (e.g., Barrett and Henzi, 2005; Heyes, 1998), in particular in chimpanzees (Matsuzawa, Tomonaga, and Tanaka, 2006) and macaques (Sinha, 2003). This literature illustrates that non-human primates use Theories of Mind about other primates while interacting socially with them in specific types of behaviour like imitation, social relationships, deception, and role-taking. Moreover, recent literature suggests that dogs use a certain kind of Theory of Mind as well (e.g., Horowitz, 2002; Virányi, Topál, Miklósi, and Csányi, 2006). However, none of these papers contains a computational model of Theory of Mind in non-human primates. In contrast, the current paper presents such a model, and illustrates how it can be applied to simulate the behaviour of a prey animal that tries to manipulate the attacking behaviour of a predator. Moreover, a number of other papers propose computational models of Theory of Mind (e.g., Gmytrasiewicz and Durfee, 1995; Marsella, Pynadath, and Read, 2004), but these are not applied explicitly to animal behaviour. For an extensive comparison of our approach to these models, the reader is referred to (Bosse, Memon, and Treur, 2007).

For future research, it is planned to exploit the features of the LEADSTO language for modelling more quantitative, numerical concepts. For example, the possibility to add probabilities to the simulation rules will be explored. In addition, more precise values can be chosen for the timing parameters e, f, g, h mentioned in the introduction. Doing this also makes it possible to make a better comparison between the traces shown in Figure 3 and 4. Currently, the trace in Figure 4 does not contain the first three world states shown in Figure 3. If these were present, the predator would kill the prey before the prey had the chance to manipulate it. By allowing different timing parameters, this problem could be solved. In addition, being able to experiment with the timing parameters would allow the modeller to make the model more realistic.

References

Aristotle (350 BCa). Nicomachean Ethics (translated by W.D. Ross).

Aristotle (350 BCb). De Motu Animalium On the Motion of Animals (translated by A. S. L. Farquharson) .

Baron-Cohen, S. (1995). Mindblindness. MIT Press.

Barrett, L. and Henzi, P. (2005). The social nature of primate cognition. *Proceedings of The Royal Society of Biological Sciences*, vol. 272, pp. 1865-1875.

Bogdan, R.J. (1997). Interpreting Minds. MIT Press.

Bosse, T., Jonker, C.M., Meij, L. van der, and Treur, J. (2007). A Language and Environment for Analysis of Dynamics by Simulation. *Int. Journal of Artificial Intelligence Tools.* To appear, 2007. Earlier version in: Eymann, T. et al. (eds.), Proc. of the 3rd German Conf. on Multi-Agent System Technologies, MATES'05. Springer LNAI, vol. 3550, pp. 165-178.

Bosse, T., Memon, Z.A., and Treur, J. (2007). A Two-Level BDI-Agent Model for Theory of Mind and its Use in Social Manipulation. In: *Proceedings of the AISB 2007 Workshop on Mindful Environments*, pp 335-342.

Dennett, D.C. (1987). The Intentional Stance. MIT Press. Cambridge Mass.

Dennett, D.C. (1991). Real Patterns. The Journal of Philosophy, vol. 88, pp. 27-51.

Georgeff, M. P., and Lansky, A. L. (1987). Reactive Reasoning and Planning. In: Forbus, K. and Shrobe, H. (eds.), Proceedings of the Sixth National Conference on Artificial Intelligence, AAAI'87. Menlo Park, California. American Association for Artificial Intelligence, 1987, pp. 677-682.

Gmytrasiewicz P. J., and Durfee. E. H. (1995). A rigorous, operational formalization of recursive modeling. In: Lesser, V. (ed.), *Proceedings of the First International Conference on Multiagent Systems*, pp. 125-132.

Heyes, C.M. (1998). Theory of mind in nonhuman primates. *Behavioural and Brain Sciences*, vol. 21, pp. 101-134.

Horowitz, A. (2002). The behaviors of theories of mind, and a case study of dogs at play. PhD. Thesis, University of California, 2002.

Jonker, C.M., Treur, J., and Wijngaards, W.C.A., (2003). A Temporal Modelling Environment for Internally Grounded Beliefs, Desires and Intentions. Cognitive Systems Research Journal, vol. 4(3), 2003, pp. 191-210.

Malle, B.F., Moses, L.J., Baldwin, D.A. (2001). Intentions and Intentionality: Foundations of Social Cognition. MIT Press.

Marsella, S.C., Pynadath, D.V., and Read, S.J. (2004). PsychSim: Agent-based modeling of social interaction and influence. In: Lovett, M., Schunn, C.D., Lebiere, C., and Munro, P. (eds.), *Proc. of the Sixth Int. Conference on Cognitive Modeling, ICCM 2004*, pp. 243-248 Pittsburg, Pensylvania, USA.

Matsuzawa, T., Tomonaga, M., and Tanaka, M. (2006). Cognitive Development in Chimpanzees. Springer Verlag, Tokyo, 2006.

Rao, A.S. and Georgeff, M.P. (1995) BDI-agents: from theory to practice. In: Lesser, V. (ed.), Proceedings of the International Conference on Multiagent Systems, pp. 312 – 319.

Rao, A.S. and Georgeff, M.P. (1991). Modelling Rational Agents within a BDI-architecture. In: Allen, J., Fikes, R. and Sandewall, E. (eds.), Proceedings of the Second International Conference on Principles of Knowledge Representation and Reasoning, (KR'91). Morgan Kaufmann, pp. 473-484.

Sinha, A. (2003). A beautiful mind: Attribution and intentionality in wild bonnet macaques. *Current Science*, vol. 85. no. 7, 2003.

Virányi, Zs., Topál, J., Miklósi, Á, and Csányi, V. (2006). A nonverbal test of knowledge attribution: a comparative study of dogs and children. *Animal Cognition*, vol. 9, no. 1, pp. 13-26.