Observation Expectation Reasoning in Agent Systems*

Bình Vũ Trần and **James Harland** and **Margaret Hamilton**
School of Computer Science and Information Technology
RMIT University
Melbourne, Australia
{tvubinh, jah, mh}@cs.rmit.edu.au

Abstract

The computational grounding problem – the gap between the mental models of an agent and its computational model – is a well known problem within the agent research community. For years, it has been believed that obscure ontological status is the principal cause. This acute problem hampers the speed of agent oriented development. In this work, we propose an alternative way for modelling intelligent agents through the concepts of observation and expectation to avoid this problem.

1 Introduction

The view of agents as intentional systems by Dennett (1987) has been dominant for many years. Mental attitudes such as knowledge, belief, desires, hopes, fears, etc. have been formally analysed to predict the intelligent behaviour of an agent. Following Hintikka (1962), this study is usually carried out using modal logics with possible-worlds semantics. In such theories, modal operators are added to represent an agent's mental attitudes. However, the three questions raised by Wooldridge in his thesis (Wooldridge 1992) about the ontological status of possible worlds remain improperly answered: "Do they really exist? If so, where are they? How do they map onto an agent's physical architecture?" This is usually referred as **the computational grounding problem** of agent theories.

Wooldridge (2000) summarised an approach taken in (Fagin *et al.* 1995; Rosenschein 1985; Wooldridge 1992; Wooldridge & Lomuscio 2001) towards this problem: to have a direct interpretation of a modal formula in terms of program computations. This approach is based on the assumption that "in general, there is *no relationship* between models $mod(\mathcal{L})$ for (a modal logical language) $\mathcal{L}$ (representing a theory of agency) and computations $\mathcal{C}$." The best known example given for this approach is epistemic logic, the modal logic of knowledge (Fagin *et al.* 1995). Consequently, this approach classifies a number of other useful theories of agency such as Cohen-Levesque's theory of intention (Cohen & Levesque 1990) and Rao-Georgeff's Belief-Desire-Intention (BDI) logics (Rao & Georgeff 1991b; 1998) as **ungrounded**.

In this paper, we defy this assumption and point out two basic anomalies which affect theories of agency from the language of specification and the semantic point of view. First, we trace back to what is known as the ***asymmetry problem of modal logic*** (Blackburn 2000b): Although possible worlds are crucial to Kripke semantics, nothing in modal syntax is able to represent them. This results in inadequacy and difficulty of using modal logic as representation formalism and reasoning systems. By adopting Blackburn et al's hybrid logics (Areces, Blackburn, & Marx 2001; Blackburn 2000a; Blackburn & Tzakova 1999), we show that it is possible to construct a corresponding computational model for any agent theory specified by a hybrid language.

However, it is also commonly argued that the introduction of a mechanism to represent possible worlds in modal syntax may answer the Wooldridge's third question but does not resolve his two first questions about possible worlds. The approach and theory of knowledge summarised by Wooldridge (2000) are essentially based on *sensory observations*. A crucial assumption in this approach is that "*there is no uncertainty about the result of performing an action in some state*". Hence if there exists a difference between a belief and a sensory information, the belief finds no ground in this framework though the sensory information can be incorrect or imperfect. Although Wooldridge also claims "dropping this assumption is not problematic," there is no pivotal work showing how useful the introduction of uncertainty would be providing the grounds for mental attitudes such as beliefs, desires.

So, what should be the grounds for mental attitudes such as beliefs, desires? Jakob Fries conceived the existence of non-intuitive grounds of knowledge and Leonard Nelson advocated this idea in (Nelson 1949). Close to this approach, Karl Popper (1969, p. 47) proposed a more specific notion: the notion of *expectations*. Each agent is born with expectations, the *psychologically or genetically a priori*, i.e. prior to all observational experience. The crucial point of this approach is: once an expectation is disappointed by observations, it creates a problem. The process of error elimination using critique continuously generates new expectations and also new problems. *The growth of knowledge proceeds from old problems to new problems, by means of conjectures and*

*We would like to thank Rajeev Goré and three anonymous reviewers for their useful comments and suggestions.

refutations (Popper 1972, pp 258, 259).

The goal of this paper is to introduce a new formal reasoning system based on expectations which does not suffer the fate of the computational grounding problem. In section §2, we investigate the expressiveness problem of modal language to outline a general framework establishing the relationship between a theory of agency and its computational model. Section §3 introduces a possible-worlds model where every possible world has a corresponding grounds which can be translated to a computational model. In section §4, we introduce the observation refutation method as a fundamental tool in knowledge evolution. Section §5 discusses further an approach to integrate criticism into the process. We then discuss some significant issues in our work and others in section §6.

2 Closing The Theory and Practice Gap

As discussed in his thesis (Wooldridge 1992) and in a recent work (Wooldridge 2000), Wooldridge claims: there is *no relationship* between a set of models $mod(\mathcal{L})$ for a modal logical language $\mathcal{L}$ representing a theory of agency and a set of computations $\mathcal{C}$ which simulates the specified properties. Hence, Wooldridge opts for the approach in (Fagin *et al.* 1995; Rosenschein 1985; Wooldridge 1992; Wooldridge & Lomuscio 2001) to directly use a computational model for deriving formulae of the specification language. Opposed to this claim, we argue that the relationships between models (including computational models) preserving all properties of modal languages have been very well studied under the notion of *bisimulation* (Blackburn, de Rijke, & Venema 2001) (cf. p-relations (van Benthem 1983, Definition 3.7)). Bisimulations are *many-to-many* relations. Hence, the problem is not because of the non-existing relationship. It is a selection problem: which corresponding model is the most appropriate one? The reason revealed by the study of bisimulation is that modal languages are not expressive enough to define the various properties of a possible worlds frame.

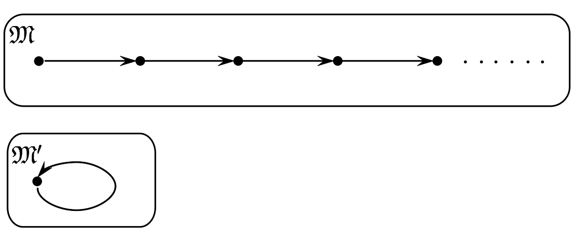

Figure 1: Invariance and modal languages

For example, let $\mathfrak{M}$ be a model which has the natural numbers in their usual order as its frame ($W = \mathbb{N}$) and every propositional symbol is $\mathtt{true}$ at every world, and $\mathfrak{M}'$ be another model which has a **single** reflexive world as its frame and all propositional symbols are $\mathtt{true}$ at this world (*see* **Figure. 1**). Apparently, one is infinite and irreflexive whilst the other is finite and reflexive. However, they are both indistinguishable under the same modal language.

Recognising *"the inability to show a* **one-to-one correspondence** *between the model theory, proof theory, and the abstract interpreter"*, Rao (1996) attempted to increase the expressive power for BDI languages. However, the resulting language AgentSpeak(L) is no longer a modal language.

Blackburn *et al*, following Prior's (1967), have thoroughly studied hybrid logics (Areces, Blackburn, & Marx 2001; Blackburn 2000a; Blackburn & Tzakova 1999). By uniquely tagging a label to each world, a bisimulation between two hybrid models ensures that not only points named by the same label are linked to each other and but most importantly **only** to each other. Hybrid bisimulation therefore becomes *isomorphic* (Areces, Blackburn, & Marx 2001). Crucially, the increase of expressive power comes with **no cost**: The satisfiability problem remains decidable in *PSPACE-complete* (Areces, Blackburn, & Marx 2001). This approach is exactly what we, human beings, do when we get lost in a jungle or in a maze by using landmarks or observing path-turning angles respectively. This assists us in envisaging the world structure in our mind as precisely as the real world that we observe.

3 The Grounds of Knowledge

3.1 Expectation and Observation

In answering the questions about the ontological status of possible worlds: "Do they really exist? If so, where are they?" Wooldridge (1992; 2000), following (Fagin, Halpern, & Vardi 1992), takes computational states as the grounds of knowledge. The grounds correspond to an agent's sensory experience of the real world — intuitive immediate knowledge. This approach offers a powerful analytical tool in various problems such as knowledge-based protocols. Unfortunately, intuitive immediate knowledge cannot be grounds for mental states such as beliefs, which can hold false statements according to sensory observations. For example, a belief statement "David Copperfield cannot fly" may well become false when one sees his body floating in the air. Grounds for such statements are discovered, in the Kant-Friesian school of philosophical theory, as *non-intuitive immediate knowledge*. According to Nelson (Nelson 1949), proof, demonstration and deduction are three possible ways to ground a proposition: 1) *proof* provides justification using logical derivation; 2) *demonstration* verifies judgements by pointing out the intuition on which they are grounded. The task is done through *sensory observations*; 3) Finally, *deduction* provides justification on a non-intuitive grounds – the *immediate knowledge of pure reason*. This task remains within the limits of *mental self-observation*.

Through this analysis, the realm of observation emerges as an important means for judgements. It represents the relation between the material world $\mathbb{G}$, on which intuitive grounds lies, and the mind of each agent a_i, on which non-intuitive grounds lies. Herein, we call the mental state corresponding to the occurrence of each observation *expectation*. Formally,

Definition 1. *An observation relation* $\mathbf{O} \subseteq \mathbb{G} \times \mathcal{E}_i$ *is a relation between the real world* $\mathbb{G}$ *and a subset of mental* *states called expectation set $\mathcal{E}_i$, where $i \in \mathcal{I}$ is the identity of an agent a_i.*

Sensory observations are obtained via sensors (e.g. the eagle's eyes ς) where mental self-observations are expressed through effectors (e.g. wings ε). An important note here is that sensory observations are full in the sense that they fully associate the states of the material world and the states of mind. However, since mental observations are *non-intuitive* it does not contain the association with $\mathbb{G}$. Hence, to complete a mental (effective) observation, it must be associated with at least a sensory observation. Let $\mathbb{S} = \bigcup_{i\in\mathcal{I}} \mathbb{S}_i$ and $\mathbb{E} = \bigcup_{i\in\mathcal{I}} \mathbb{E}_i$ be respectively the sensor and effector sets of an observation system, where $\mathbb{S}_i$ and $\mathbb{E}_i$ are respectively the sets of sensors and effectors of an individual agent a_i.

The formation of these primitive observations is called an *observation method*. An observation method which contains only one primary sensor or effector is called primitive observation method $\mathbb{M}_0$ (e.g. $\epsilon, \varsigma \in \mathbb{M}_0$). A more complicated set of observation methods $\mathbb{M}_k$ would arrange the k expectations of other observation methods in some order to generate new expectations about the world. These expectations are also associated with global states to form more complex observations.

Definition 2. ***(Observation methods)*** *An observation method family is a set of observation method sets $\mathcal{M} = \{\mathbb{M}_k\}_{k\in\mathbb{N}}$ where $\mathbb{M}_k$ is a set of observation methods of arity k for every $k \in \mathbb{N}^+$. $\mathbb{M}_0 = \mathbb{S} \cup \mathbb{E}$ is called primitive observation method set. $\mathbb{M}_k$ is inductively defined as follows:*

- *$e \in \mathcal{E}$ for all $e \in \mu_0, \forall \mu_0 \in \mathbb{M}_0$*
- *$\mu_k(e_1, \dots, e_k) \subseteq \wp(\mathcal{E})$ for all $\mu_k \in \mathbb{M}_k$ and $e_1, \dots, e_k \in \mathcal{E}$*

3.2 Expectation Model

Before defining a modal logical language which can also describe possible worlds in its syntax, in this section, we first describe the set of possible worlds $\mathbb{G}_i$ of an agent a_i. Each possible world g carries the agent a_i's information about its environment. There are two closely related sources of information: from the agent's set of sensors $\mathbb{S}_i$ and from the results of the agent's effectors $\mathbb{E}_i$. By organising the sources of information, an agent will obtain certain information about its environment. For example, when an eagle turns its head, its eyes obtain information differently; or, when the eagle flaps its wings, in its mind, the resulting mental state could be it is already 5 metres away from where it was, but this can only be verified by its eyes. We call each possible world (each possible way of organising information sources) an *observation*. On the one hand, similar to epistemic logic, the grounds of sensory observations are information from sensors. On the other, following philosophers Fries, Nelson and Popper, we ground effective observations on non-intuitive grounds, *inborn expectations*. Inborn expectations are genetically given to an agent at its birth. We call the valid set of formulae at an observation its *expectations*.

The expectation language $\mathcal{L}$ is similar to the language of propositional logic augmented by the modal operator $\mathcal{E}_i$ and the observation operators $@_s$, where s is an observation label i.e. an atomic proposition which is $\mathtt{true}$ *at exactly one possible world in any model.*

Definition 3. (Expectation language) *Let Φ be a set of atomic expectation propositions. Let Ξ be a nonempty set of observation labels disjoint from Φ. An expectation language $\mathcal{L}$ over Φ and Ξ where $p \in \Phi$ and $s \in \Xi$ is defined as follows:*

$$\varphi := s \mid p \mid \neg\varphi \mid \varphi\wedge\varphi \mid \varphi\vee\varphi \mid \varphi \to \varphi \mid \langle\mathcal{E}_i\rangle\varphi \mid [\mathcal{E}_i]\varphi \mid @_s\varphi.$$

Two observations are said to be related if one can be obtained by changing (adding/removing) the information sources (sensors/effectors) of the other. In other words, it is possible to reach the other observation by changing the information sources from the current one. Let $\sim_e^i \subseteq \mathbb{G}_i \times \mathbb{G}_i$ be the set of such related observations. The pair $\mathcal{F} = \langle\mathbb{G}_i, \sim_e^i\rangle$ is called an *observation frame*. The interpretation of an agent a_i's expectations is defined by the function $\pi : \Phi \cup \Xi \to \wp(\mathbb{G}_i)$. The crucial difference from orthodox modal logic in this definition is that for every observation label $s \in \Xi$, π returns a **singleton**. In other words, s is $\mathtt{true}$ at a unique observation, and therefore tags this observation (Blackburn 2000a). The triple $\mathfrak{M} = \langle\mathbb{G}_i, \sim_e^i, \pi\rangle$ is called an expectation model.

Definition 4. *The semantics of expectation logic $\mathcal{L}$ are defined via the satisfaction relation $\models$ as follows*

1. *$\langle\mathfrak{M}, g\rangle \models p$ iff $g \in \pi(p)$ (for all $p \in \Phi$)*
2. *$\langle\mathfrak{M}, g\rangle \models \neg\varphi$ iff $\langle\mathfrak{M}, g\rangle \not\models \varphi$*
3. *$\langle\mathfrak{M}, g\rangle \models \varphi \vee \psi$ iff $\langle\mathfrak{M}, g\rangle \models \varphi$ or $\langle\mathfrak{M}, g\rangle \models \psi$*
4. *$\langle\mathfrak{M}, g\rangle \models \varphi \wedge \psi$ iff $\langle\mathfrak{M}, g\rangle \models \varphi$ and $\langle\mathfrak{M}, g\rangle \models \psi$*
5. *$\langle\mathfrak{M}, g\rangle \models \varphi \to \psi$ iff $\langle\mathfrak{M}, g\rangle \not\models \varphi$ or $\langle\mathfrak{M}, g\rangle \models \psi$*
6. *$\langle\mathfrak{M}, g\rangle \models \langle\mathcal{E}_i\rangle\varphi$ iff $\langle\mathfrak{M}, g'\rangle \models \varphi$ for some g' such that $g \sim_e^i g'$*
7. *$\langle\mathfrak{M}, g\rangle \models [\mathcal{E}_i]\varphi$ iff $\langle\mathfrak{M}, g'\rangle \models \varphi$ for all g' such that $g \sim_e^i g'$*
8. *$\langle\mathfrak{M}, g\rangle \models s$ iff $\pi(s) = \{g\}$, for all $s \in \Xi$, g is called the denotation of s*
9. *$\langle\mathfrak{M}, g\rangle \models @_s\varphi$ iff $\langle\mathfrak{M}, g_s\rangle \models \varphi$ where g_s is the denotation of s.*

where 1 – 7 are standard in modal logics with two additions of hybrid logics in 8 and 9.

Thus, if $[\mathcal{E}_i]\varphi$ is true in some state $g \in \mathbb{G}_i$, then if the agent a_i takes any further possible observation g' from g, φ will be its expectation at g'. For example, let φ be "the sun is shining". If the eagle is making observation g: taking visual images, then if it makes a further observation g', either to use its eyes again or flap its wings to move forward, it will still expect the sun to be shining.

An observation statement $@_s\varphi$ says, whilst taking the observation named s, the agent a_i has an expectation φ. The definition of the observation operator (item 9.) allows the agent to retrieve information from another observation.

4 Reasoning about Unexpectedness – Observation calculus

We now come to the heart of the paper. To illustrate why expectation reasoning is useful, imagine a hungry eagle is chasing an agile sparrow in a maze which neither of them knows how to get through in advance. Suddenly, the sparrow darts behind a wall and disappears from the eagle's field of view. Naturally the eagle would very rarely stop the chase and rest. It continues by making conjectures, *its expectations*, whereabouts the sparrow is, and acts accordingly. However, the eagle's expectations are often violated. The sparrow may not reappear at the end of the wall as expected or it may be caught in a dead-end. By reasoning about such unexpectedness critically, the eagle would be able to eliminate erroneous expectations and also generate new ones finding a way to catch the sparrow and get out of the maze. Throughout all of his well-known books about scientific knowledge, Popper (1959; 1969; 1972) summarised the above reasoning process by the following schema

$$P_1 \rightarrow TT \rightarrow EE \rightarrow P_2$$

Here P stands for problem, TT stands for tentative theory, and EE stands for error-elimination. *The first problems* are created when an agent's *inborn expectations* are disappointed by some observation. The ensuing growth of knowledge may then be described as the process of correcting and modifying previous knowledge through new observations using refutation methods.

Analytic semantic tableaux methods are refutation methods that have received much attention recently in automated theorem proving. A technique by Fitting (1996), which defers the choice of free-variables until more information is available, has been used to reduce search space and the non-determinism inherent in automated proof search. This technique resembles the ability to use expectations as assumptions to delay a current obstructed observation until justified as well as to use expectations as conjectures to find a path through a maze in the above example. Among different approaches using free variables in the labels of semantic modal tableaux, Beckert and Goré's string matching technique (Beckert & Goré 1997) can be used to describe the connection between sensory observations and effective observations. Before introducing our refutation methods, we now elaborate on a construction for the observation frame in §3.2 as follows.

Firstly, an observation (a possible world) is described in detail by a sequence σ of sensors and effectors. For example, the eagle's first observation, looking forward using its eyes $\varsigma \in \mathbb{S}_i$, is represented by the string $\sigma = \varsigma$. Its next observation by subsequently using its wings ε is represented by $\sigma = \varsigma.(\varsigma)$. The brackets denote that the resulting effect of ε is non-intuitive and this effect can be verified using the eagle's eyes ς. We use $[\varsigma]$ to denote a generic observation (either sensory or effective observation). An assumption can be made by substituting a free variable x of the variable set Ψ at some position in the sequence. Each of these sequences can be named using the set observation labels Ξ and by the naming function N.

Definition 5. ***(Linear observation method syntax)*** *Let $\Psi = \{x, y, \ldots\}$ be a set of assumptions which are originally not bound to any sensors or effectors. Let Γ be a set of observation sequences. A linear observation method can be expressed by a string $\sigma \in \Gamma$ defined inductively as follows:*

i. *ς is an observation sequence for all $\varsigma \in \mathbb{S}_i$;*

ii. *If σ is an observation sequence, then so are $\sigma.\eta$ (if $\eta \in \mathbb{S}$) and $\sigma.(\eta)$ (if $\eta \in \mathbb{E}$);*

iii. *If σ is an observation sequence, then for all $x \in \Psi$ $\sigma.(x)$ is also an observation sequence but not $\sigma.x$. $\sigma.(x)$ stands for all possible successors of the observation sequence σ.*

iv. *$prefix(\sigma) = \{\tau \mid \sigma = \tau.\theta\}$ is a function which returns a set of all prefixes of an observation sequence σ.*

v. *The function $N : \Xi \rightarrow \Gamma$ assigns each label in Ξ to an observation sequence.*

Example 1. *Let ε be the wings of the eagle in the above example and ς be its eyes. Since the eagle's eyes can verify the eagle's position after flapping its wings, we can denote this effect by (ς). Whilst chasing the sparrow in the cave system, the eagle can simply use an observation sequence such as $\varsigma.(\varsigma).\varsigma.(\varsigma)\ldots$ to firstly search for the sparrow by its eyes and then flap its wings flying towards the sparrow. If the eagle flies into a dark area, its eyes no longer see the sparrow, and hence it needs to make an assumption x which appears in the observation sequence as $\varsigma.(\varsigma).(x).(\varsigma)\ldots$ By naming the two sequences as $s = \varsigma.(\varsigma)$ and $t = \varsigma.(\varsigma).\varsigma$ respectively, we can describe the eagle's expectation after making an observation using s as $@_s q$ where q means "energy burnt out". The agent's expectation about "the sparrow is caught" p by taking a further observation to s (creating observation sequence t, $@_s\langle\mathcal{E}_i\rangle t$) can be expressed using $@_s\langle\mathcal{E}_i\rangle p$.*

By this novel representation, effective observations and assumptions have not only found their non-intuitive grounds but have also been intertwined closely with sensory observations. This approach hence offers a more powerful analytical tool to wider range of applications, for example, optimising knowledge-based protocols in unstable environment, cooperative problem solving etc. It also promises the ability to build grounded semantics for the theory of intention by Cohen and Levesque (1990) or BDI logics by Rao and Georgeff (1991b).

We identify the following properties as essential for an observation frame. Firstly, once an agent a_i is created, its set of sensors and effectors should be considered as fixed (e.g. eyes, wings, nose, skin…) However, the agent can extend its observation further by incorporating sensors and effectors from other agents in its observation (through mirror, tools…). But any extended observations should be ultimately rooted to the agent's innate set of sensors $\mathbb{S}_i$ (Definition 6(i)). Secondly, the interpretation should allow the justification of an expected sensing effect generated by an effector when it is possible to place the corresponding sensor for that effect (Definition 6(ii)). Thirdly, if any preceding observation sequence cannot be interpreted (explained), then there will be no interpretation for any subsequent observation based on the observation sequence (Definition 6(iii)). Fourthly, if the agent takes another sensory observation following an interpreted observation sequence then the new observation sequence must also be interpreted (Definition

6(iv)). Finally, by taking another observation based on the current observation sequence, the agent also associates its expectations with these observations (Definition 6(v)). Formally, these properties are stated as follows

Definition 6. *An observation interpretation is a pair $\langle \mathcal{F}, I \rangle$ where $\mathcal{F} = \langle \mathbb{G}_i, \sim_e^i \rangle$ is an observation method frame and I is a function $I : \mathcal{L}_{\mathbb{S}\cup\mathbb{E}} \rightarrow \mathbb{G}_i \cup \{\bot\}$ which tells how the real world $\mathbb{G}$ is reflected into an agent's mind through the sequences of all available sensors and effectors in this observation frame. $\mathbb{G}_i$ is the reflected part of the global world $\mathbb{G}$ in the agent's a_i's mind through in this observation frame. An observation method interpretation function must satisfy the following properties:*

i. ***(Individuality)*** *$I(\eta) \in \mathbb{G}_i$ for all $\eta \in \mathbb{S}_i$;*
ii. ***(Justification)*** *$I(\sigma.(\eta)) = I(\sigma.\eta)$ for all $\sigma.(\eta)$ and $\sigma.\eta$ in $\mathcal{L}_{\mathbb{S}\cup\mathbb{E}}$;*
iii. ***(Entirety)*** *for all $\sigma \in \mathcal{L}_{\mathbb{S}\cup\mathbb{E}}$, if $I(\tau) = \bot$ for some $\tau \in prefix(\sigma)$ then $I(\sigma) = \bot$;*
iv. ***(Constructability)*** *for all labels $\tau.\eta \in prefix(\sigma) \cup \{\sigma\}$, if $I(\tau) \in \mathbb{G}_i$ then $I(\tau.\eta \in \mathbb{G}_i)$*
v. ***(Expectability)*** *$\tau = \sigma.[\eta], I(\sigma) \in \mathbb{G}_i, I(\tau) \in \mathbb{G}_i$ iff $I(\sigma) \sim_e^i I(\tau)$.*

Example 2. *If the observation sequence $\sigma = \varsigma_1.\varsigma_2.\varsigma_3$ (where ς_1 senses hunger, ς_2 brings visual images, ς_3 brings touching sense) is satisfied then $I(\varsigma_1), I(\varsigma_1.\varsigma_2)$, $I(\varsigma_1.\varsigma_2.\varsigma_3)$ must be defined. If we have an observation sequence $\sigma = \varsigma_1.(\varsigma_2).\varsigma_3$, where (ς_2) is the effect of flapping wings then $I(\varsigma_1.(\varsigma_2))$ need not be defined (i.e. $I(\varsigma_1.(\varsigma_2)) = \bot$) since that observation may not be captured or interpreted. However if it is, especially when σ exists, then by justification $I(\varsigma_1.(\varsigma_2).\varsigma_3)$ must be interpreted.*

With the refinement of observations and observation methods defined above, we can now describe our *refutation method*. An expectation $@_s\varphi$ can be refuted by constructing a tableaux proof with $@_s\varphi$ at its root, where s is a name for the observation from a built-in sensor ς and φ is the unexpected information perceived by ς.

Definition 7. ***(Tableau observation proofs)*** *A sequence $\mathcal{T}_0, \ldots, \mathcal{T}_r$ of tableaux is an observation proof for the unsatisfiability of a formula φ if:*

i. *$\mathcal{T}_0$ consists of the single node $@_s\varphi$ where s is a name for some built-in sensor $\varsigma \in \mathbb{S}_i$.*
ii. *for $1 \leq m \leq r$, the tableau $\mathcal{T}_m$ is constructed from $\mathcal{T}_{m-1}$ by applying an expansion rule in Table 1, the testimony rule, or the discard rule; and*
iii. *all branches in $\mathcal{T}_r$ are marked as atomically closed.*

These tableaux are constructed by taking subsequent observations using the tableau expansion rules given in **Table. 1**. Though all rules look standard in any **KE** system (D'Agostino & Mondadori 1994), the tableau modality expansion rules have some distinctive significance to be discussed here. Firstly, it is important to note that the only branching rule in this rule set is **PB** (principle of bivalence), which considerably reduces the search space. This rule further insists that the choice to pursue a path is dependent on the truth value of the expectations themselves not on the connectives that link together. This also impacts how the nature of an assumption changes. Normally, a universal assumption x is introduced into an observation sequence t by the reduction of *universal* rules ($\neg\Diamond$- or $\Box$-rule). For example, the eagle may have the expectation "All sparrows are agile". Hence, the eagle can assume that in any further observation it would see agile sparrows. The assumption is no longer universal (or being *freed*), when the eagle is able to observe some sluggish sparrow in one of its further observation. A system with **PB** rule can easily show this change. In classic methods such as Beckert and Goré's free-variable tableaux method (Beckert & Goré 1997), it would not be possible to represent it since the branching rule is based on disjunction rule.

Secondly, the *existential* rules ($\Diamond-$ and $\neg\Box-$rules) require specifically that when they are applied to $@_s\langle\mathcal{E}_i\rangle\varphi$ or $\neg@_s[\mathcal{E}_i]\varphi$ respectively, some subsequent observation must be able to perceive φ. The label a is defined by $N(a) = \sigma.\lceil\varphi\rceil$ where $N(s) = \sigma$ and $\lceil\varphi\rceil$ is a bijection from the set of formulae to ground observation sequences which can sense φ. Here, $\lceil\varphi\rceil$ is a Gödelisation of φ itself. In Popper's language, this means an agent must be able to observe an instance of φ using some specialised tool $\lceil\varphi\rceil$.

In a strong observation proof system, all assumptions must be justified by testimony and discard rules. These rules are defined as follows

Definition 8. ***(Testimony rule)*** *Given a tableau $\mathcal{T}$, a new tableau $\mathcal{T}' = \mathcal{T}\theta$ may be constructed from $\mathcal{T}$ by applying a substitution θ to $\mathcal{T}$ that instantiates free assumptions in $\mathcal{T}$ with other free assumptions or with sensory observations.*

Definition 9. ***(Discard rule)*** *Given a tableau $\mathcal{T}$ and a substitution $\delta : \mathcal{U}(\mathcal{T}) \rightarrow \mathbb{S}$ that instantiates universal assumptions in $\mathcal{T}$ with sensory observations, the* ***discard rule*** *constructs a new tableau $\mathcal{T}'$ from $\mathcal{T}$ by marking $\mathcal{B}$ in $\mathcal{T}$ as closed provided that:*

i. *the branch $\mathcal{B}\delta$ of $\mathcal{T}\delta$ contains a pair $@_s\varphi$ and $\neg@_s\varphi$*
ii. *the observation sequence named by s is ground and justified.*

The main difference between the testimony rule and the discard rule is that, the testimony rule allows us to replace any **free** assumption at any time in a tableau. The main advantage of this is that it allows us to close a branch of the observation tableau without obtaining real-world observations.

The tableau construction by applying the tableau rules above gives us all the formulae in a tableau of the form of observation satisfaction statements $@_s\varphi$ or $\neg@_s\varphi$. Each time an expansion rule is applied, one or two formulae yielded by this rule are added to the set Σ of observation satisfaction statements. We define the satisfaction by ground observation for this set as follows

Definition 10. ***(Ground observation satisfaction)*** *Let Σ be a set of observation satisfaction statements. We say that Σ is satisfied by an observation in $\mathfrak{M}$, where $N(s)$ is a ground observation sequence, if and only if for all formulae in Σ:*

i. *If $@_s\varphi \in \Sigma$ then $\mathfrak{M}, g_s \models \varphi, g_s \in \mathbb{G}_i$ and*
ii. *If $\neg@_s\varphi \in \Sigma$ then $\mathfrak{M}, g_s \not\models \varphi, g_s \in \mathbb{G}_i$*

Disjunction rules

$$\frac{\begin{array}{c}@_s(\varphi \vee \psi)\\ \neg @_s\varphi\end{array}}{@_s\psi}(\vee 1) \qquad \frac{\begin{array}{c}@_s(\varphi \vee \psi)\\ \neg @_s\psi\end{array}}{@_s\varphi}(\vee 2) \qquad \frac{\neg @_s(\varphi \vee \psi)}{\begin{array}{c}\neg @_s\varphi\\ \neg @_s\psi\end{array}}(\neg\vee)$$

Conjunction rules

$$\frac{\begin{array}{c}\neg @_s(\varphi \wedge \psi)\\ @_s\varphi\end{array}}{\neg @_s\psi}(\neg\wedge 1) \qquad \frac{\begin{array}{c}\neg @_s(\varphi \wedge \psi)\\ @_s\psi\end{array}}{\neg @_s\varphi}(\neg\wedge 2) \qquad \frac{@_s(\varphi \wedge \psi)}{\begin{array}{c}@_s\varphi\\ @_s\psi\end{array}}(\wedge)$$

Implication rules

$$\frac{\begin{array}{c}@_s(\varphi \to \psi)\\ @_s\varphi\end{array}}{@_s\psi}(\to 1) \qquad \frac{\begin{array}{c}@_s(\varphi \to \psi)\\ \neg @_s\psi\end{array}}{\neg @_s\varphi}(\to 2) \qquad \frac{\neg @_s(\varphi \to \psi)}{\begin{array}{c}@_s\varphi\\ \neg @_s\psi\end{array}}(\neg\to)$$

Negation rules

$$\frac{@_s\neg\varphi}{\neg @_s\varphi}(\neg) \qquad \frac{\neg @_s\neg\varphi}{@_s\varphi}(\neg\neg)$$

Satisfaction rules

$$\frac{@_s @_t\varphi}{@_t\varphi}(@) \qquad \frac{\neg @_s @_t\varphi}{\neg @_t\varphi}(\neg @)$$

Naming rules

$$\frac{[s\ on\ branch]}{@_s s}(Ref) \qquad \frac{@_t s}{@_s t}(Sym) \qquad \frac{@_t s\ @_t\varphi}{@_s\varphi}(Nom) \qquad \frac{@_s\langle\mathcal{E}_i\rangle t\ @_t\varphi}{@_s\langle\mathcal{E}_i\rangle\varphi}(Bridge)$$

Modality rules

$$\frac{@_s\langle\mathcal{E}_i\rangle\varphi}{\begin{array}{c}@_s\langle\mathcal{E}_i\rangle a\\ @_a\varphi\end{array}}(\Diamond) \qquad \frac{\neg @_s\langle\mathcal{E}_i\rangle\varphi}{\begin{array}{c}@_s\langle\mathcal{E}_i\rangle t\\ \neg @_t\varphi\end{array}}(\neg\Diamond) \qquad \frac{@_s[\mathcal{E}_i]\varphi}{\begin{array}{c}@_s\langle\mathcal{E}_i\rangle t\\ @_t\varphi\end{array}}(\Box) \qquad \frac{\neg @_s[\mathcal{E}_i]\varphi}{\begin{array}{c}@_s\langle\mathcal{E}_i\rangle a\\ \neg @_a\varphi\end{array}}(\neg\Box)$$

Principle of bivalence

$$\frac{}{@_s\varphi \mid \neg @_s\varphi}(PB)$$

Principle of non-contradiction

$$\frac{\begin{array}{c}@_s\varphi\\ \neg @_s\varphi\end{array}}{\times}(PNC)$$

Table 1: Inference rules for observation system

Here g_s is the denotation of s under π. We say that Σ is satisfiable by observation if and only if there is a standard expectation model in which it is satisfied by observation.

By the integration of hybrid logics (Areces, Blackburn, & Marx 2001; Blackburn 2000a; Blackburn & Tzakova 1999), free-variable modal tableaux (Beckert & Goré 1997), and the **KE** system (D'Agostino & Mondadori 1994) we still obtain the following important results:

Theorem 1. ***(Soundness)*** *Suppose there is a closed tableau $\mathcal{T}$ for $\neg @_s\varphi$. Then φ is satisfiable.*

Proof. Using Smullyan's unifying notation to reduce the number of cases and the naming function N to map each observation sequence to an observation label, we can easily prove that any extension $\mathcal{T}'$ of a satisfiable tableau $\mathcal{T}$ by applying the above construction rules is also satisfiable. Hence, by contradiction suppose φ is invalid. Then there is some model $\mathfrak{M}$ and an observation (using a built-in sensor) about the world ς, such that $\mathfrak{M}, w \models \neg @_\varsigma\varphi$. This means that the tableau $\mathcal{T}_0$ consisting of a single node carrying $\neg @_\varsigma\varphi$ is satisfiable. Hence, any extension of $\mathcal{T}_0$ including $\mathcal{T}$ is also satisfiable. This contradicts with the hypothesis that $\mathcal{T}$ is closed and hence not satisfiable. Thus φ must be satisfiable. □

To show completeness of this system we define Hintikka sets like hybrid logics (Blackburn 2000a). It is easy to prove the following lemmas from the definitions:

Definition 11. *Let Ξ be a set of observation sequences that occur in a set of observation satisfaction statements H and $\sim_o$: $\Xi \times \Xi$ where $s \sim_o t$ iff $@_s t \in H$. Clearly, from Hintikka set definition we can see $\sim_o$ is an equivalence relation. Let $|s|$ be the equivalence class of $s \in \Xi$ under $\sim_o$. $\sim_o$ is called an observation equivalence relation.*

Lemma 1. *Two observation sequences are equivalent if and only if all expectations derived from these observations are exactly similar. Formally,*

$$s \sim_o t \Leftrightarrow (\forall\varphi, @_s\varphi \in H \Leftrightarrow @_t\varphi \in H)$$

Definition 12. *Given a Hintikka set H, let $\mathfrak{M}^H = \langle W^H, \sim^H, \pi^H\rangle$ be any triple that satisfies the following condition:*

1. $W^H = \{|s| \mid s \in \Xi\}$.
2. $|s| \sim^H |t|$ *iff* $@_s\langle\mathcal{E}_i\rangle t \in H$

3. $\pi^H(\varphi) = \{|s| \mid @_s\varphi \in H\}$ *for all atoms that occur in* H

Lemma 2. *Any triple* $\mathfrak{M}^H = \langle W^H, \sim^H, \pi^H \rangle$ *defined as above is a standard model.*

Lemma 3. *If* H *is an Hintikka set, and* $\mathfrak{M}^H = \langle W^H, \sim^H, \pi^H \rangle$ *is a model induced by* H *then*

i. If $@_s\varphi \in H$*, then* $\mathfrak{M}^H, |s| \models \varphi$.

ii. If $\neg @_s\varphi \in H$*, then* $\mathfrak{M}^H, |s| \not\models \varphi$.

Lemma 4. *If* $\mathcal{B}$ *is an open branch of the tableau* $\mathcal{T}$ *then the set* H *of all formulae on* $\mathcal{B}$ *is a Hintikka set.*

Theorem 2. ***(Strong completeness for grounded observation system)*** *Any consistent set of expectations in a countable language is satisfiable in a countable standard model.*

Proof. Suppose $\mathcal{E} = \{@_s e_1, \ldots, @_s e_n\}$ is a consistent set of expectations. Enumerate through each expectation in this set and apply inference rules when necessary to construct a tableau $\mathcal{T}$. Assume that the tableau $\mathcal{T}$ is closed. That means, if an agent observes the world by following any branch $\mathcal{B}$ in $\mathcal{T}$ it is always expecting contradicting observations. Obviously, if the branch $\mathcal{B}$ is closed, it must be closed after a finite number of steps m. As soon as the agent discovers a contradiction on a path, it immediately stops observing in that direction. Let $\min(B)$ be the set of observations on a branch $\mathcal{B}$ when the contradiction occurs. Let $\min(\mathcal{T})$ be a tree incorporating all of those branches. By König's lemma, it is easy to show that $\min(\mathcal{T})$ is finite.

However, by applying the $\wedge$-rule $n-1$ times for the formula $@_s(e_1 \wedge \ldots \wedge e_n)$, we also obtain n nodes $@_s e_1, \ldots, @_s e_n$ on a single branch as we can do with $\mathcal{E}$. The remaining construction hence is exactly similar to the construction from the set $\mathcal{E}$. Therefore, we also obtain a closed tableau $\mathcal{T}'$. That means $@_s(e_1, \ldots e_n)$ or equivalently $\neg @_s \neg(e_1, \ldots e_n)$ (by applying $\neg\neg$-rule) is unsatisfiable. This means $\neg(e_1, \ldots e_n)$ is provable. From definition, this also means $\mathcal{E}$ is inconsistent. From this contradiction, we can conclude that our assumption that the tableau $\mathcal{T}$ is closed was wrong. $\mathcal{T}$ must contain at least an open branch $\mathcal{B}$. By lemma 4, all the formulae on this $\mathcal{B}$ form a Hintikka set H which then can be used to construct a standard expectation model $\mathfrak{M}^H$ at observation $|s|$. ☐

5 Reasoning about unexpectedness: Criticism

Besides refutation method which provides a useful way to determine an appropriate theory, criticism plays a significant role in knowledge evolution. Originating from Kant (1893), criticism searches for contradictions and their elimination.

We now present a case study to illustrate how the observation-expectation reasoning system works in a concrete problem: *the bit transmission problem*. The bit transmission problem was first introduced by Halpern and Zuck (1992) as communications between two processes over a faulty communication medium. There is no guarantee that any messages sent by either agent are correctly received. Hence, when sending out a bit, each agent needs to know whether the bit they sent is correctly received by the other.

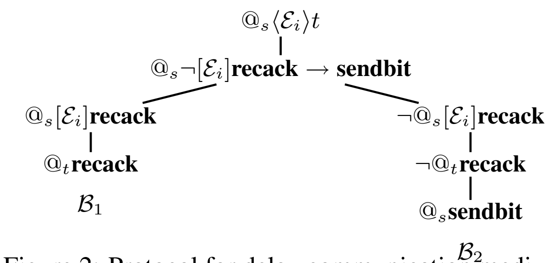

Figure 2: Protocol for delay communication medium

Otherwise, they will resend the bit until it is correctly received. Halpern and Zuck's insight was sending acknowledgement messages carrying the knowledge state of the sender of the message.

However, when the bit takes more than one cycle to be delivered, this approach can be inefficient since the sender has to wait until the bit is received. The problem becomes more interesting when the channel is working properly at certain times, while failing at others. Worse, the failures at different times may be totally different. The ability of agents to ignore insignificant errors (since they can correct them from what they know) and to construct new optimised protocols in an unstable environment thus becomes more desirable in such kind of environment.

A solution to the delay communication medium is to continuously send other bits during the delay time. Let s be the current observation. If the sender expects the delayed receipt of acknowledgement would be at observation t denoted by $@_s\langle\mathcal{E}_i\rangle t$, it would move on to the next observation and send another bit. Otherwise, it keeps sending the current bit. We denote this by $@_s\neg[\mathcal{E}_i]$**recack** $\rightarrow$ **sendbit**. The crucial difference between this and the knowledge-based protocol is the communication delay. In the knowledge based reasoning, this time must be known exactly. In expectation reasoning, t can contain an assumption which can flexibly be justified. For example, we have the assumption that our $@_t$**recack** will be justified by a, $@_t a$ where $@_a$**recack**. Unfortunately, an unexpected error occurs. Though the bit has been sent but at the observation a we cannot receive the acknowledgement as expected. From critical examination, we can derive that $\neg @_t a$.

This creates another problem: what is an appropriate justification? In other words, how to make further conjectures? We are currently investigating the possibility of using the triadic structure: Thesis-Antithesis-Synthesis by Hegel in his *Science of Logic* (Hegel 1929). Unlike classical logic, where a double negation ("A is not $\neg A$") would simply reinstate the original thesis, Hegel suggested synthesis as the third emerging element for a higher rationality. Hence, the contradiction $@_t$**recack** and $\neg @_a$**recack** when the justification $@_t a$ is made, can be analysed in a synthetic observation c where $@_c @_t$**recack** and $@_c \neg @_a$**recack**. This suggests a way to represent the justification strategy.

6 Discussion

A common approach when using modal logic in formal analysis of mental attitudes is to give a modal operator for each attitude. Then, the relationship of these attitudes are studied through interaction axioms. Belief-Desire-Intention (BDI) by Rao and Georgeff (1991b) is one of the most well-known studies using this approach. Roughly speaking, beliefs represent the agent's current information about the world; desires represent the state of the world which the agent is trying to achieve; and intentions are chosen means to achieve the agent's desires. Following the philosopher Bratman (1987), Rao and Georgeff formalised the constraints (Asymmetry thesis, non-transference and side-effect free principles) between these attitudes in (Rao & Georgeff 1991a). The major drawback of this approach is to find a grounds for such interactions. For example, it is difficult to give an analysis of why a very hungry eagle is still chasing its prey though it believes that the chase would take all of its remaining energy. It is even more difficult to explain when an information should be considered belief or knowledge. The term *expectation* has also been used as a modal operator (David, Brezner, & Francez 1994) with a weaker meaning (*most*) than necessity operator (*all*). This is achieved via the notion of *majority* defined by semi-filter. Even though this approach is very interesting, similar to the above approach, it is unable to provide the grounds for possible worlds.

A more recent work by Brunet (2002) also takes a closer look at the modal logic approach of observation-based logic. By defining an ordering of informational content captured by partial and imperfect observations, Brunet introduces *representation structures* as a basis for a knowledge representation formalism. Each representation is a set of descriptions from a given point of view. However, since observations can be partial and imperfect, there could exist several points of view for an observation. The set of transformation functions plays an important role to define such algebraic structures between different representations (points of views). Although these concepts (observation, point of view, transformation functions) appear similar to our concepts (observation, expectation, observation methods) in this paper, there are some differences between the two works. One of them is that an agent in our framework can use the expectation language to reason about its observation method structures. This is done with the adopted feature from hybrid logic.

In agent system, an agent often has incomplete access to its environment. Hence, it is significant not only to represent such situations but also to describe the reasoning process an agent takes when dealing with them. $\mathcal{VSK}$ logic (Wooldridge & Lomuscio 2000) is one of the efforts to formalise what information is true in the environment, what an agent can perceive and then know about its environment. A state of environment is captured into the agent's mind using *visibility function* (a.k.a *observation function* by van der Meyden (van der Meyden 1998)). Further, this work assumes "there is no uncertainty about performing an action in some state." However, we are unaware of any further work attempt to drop this assumption. Here, we introduce effective observations which can also be interpreted in many ways as sensory observations in (Wooldridge & Lomuscio 2000) using an observation interpretation function. The justification property of this observation interpretation function is the tool to remove uncertainty from effective observations, and hence opens a new approach to drop the assumption.

7 Concluding remarks

In this paper, we have elaborated the computational grounding problem by providing an additional ontological grounds: the *non-intuitive immediate knowledge*. This is used as the grounds for effective observations and assumptions. The observation and expectation concepts can be further used for describing other concepts such as belief, desire and intention (Bratman 1987; Cohen & Levesque 1990; Rao & Georgeff 1991b). One of such attempts was presented in (Trần, Harland, & Hamilton 2003). The work expands further this representation to formally describe part of Popper's logic of scientific discovery (Popper 1959). This includes the observation refutation method, and the process of error elimination using critique.

Apart from elaborating in more detail the error-elimination process, another chief issue that we would like to address in our future work is an efficient implementation of this novel theoretical foundation. As previously discussed, the integration of hybrid logics, free-variable modal tableaux and **KE** system is very promising. Though **KE** system is space efficient, time efficiency remains a challenge (Endriss 1999). On the other hand, free-variable modal tableaux is known for being time efficient. Hence, some possibilities to improve the complexity and efficiency are currently being explored.

References

Areces, C.; Blackburn, P.; and Marx, M. 2001. Hybrid logics: Characterization, interpolation and complexity. *Journal of Symbolic Logic* 66(3):977 – 1010.

Beckert, B., and Goré, R. 1997. Free variable tableaux for propositional modal logics. In *Proceedings of International Conference on Theorem Proving with Analytic Tableaux and Related Methods*.

Blackburn, P., and Tzakova, M. 1999. Hybrid languages and temporal logic. *Logic Journal of the IGPL* 7:27–54.

Blackburn, P.; de Rijke, M.; and Venema, Y. 2001. *Modal logic*. Cambridge University Press.

Blackburn, P. 2000a. Internalizing labelled deduction. *Journal of Logic and Computation* 10:137–168.

Blackburn, P. 2000b. Representation, reasoning, and relational structures: a hybrid logic manifesto. *Logic Journal of the IGPL* 8(3):339–625.

Bratman, M. E. 1987. *Intention, Plans, and Practical Reason*. Harvard University Press.

Brunet, O. 2002. A modal logic for observation-based knowledge representation. In Gore, R.; Mendler, M.; and de Paiva, V., eds., *Proceedings of IMLA'02 Workshop,*, 69–81.

Cohen, P. R., and Levesque, H. J. 1990. Intention is choice with commitment. *Artificial Intelligence* 42:213–261.

D'Agostino, M., and Mondadori, M. 1994. The Taming of the Cut. Classical Refutations with Analytic Cut. *Journal of Logic and Computation* 4:285–319.

David, S. B.; Brezner, O.; and Francez, N. 1994. An expectation modal logic. Technical report, Computer Science, Technion.

Dennett, D. C. 1987. *The Intentional Stance*. Cambridge, MA: The MIT Press.

Endriss, U. 1999. A Time Efficient KE Based Theorem Prover. In Murray, N. V., ed., *Automated Reasoning with Analytic Tableaux and Related Methods, International Conference, Tableaux-1999, Proceedings*, volume 1617 of *LNAI*, 313–318. Springer-Verlag.

Fagin, R.; Halpern, J. Y.; Moses, Y.; and Vardi, M. Y. 1995. *Reasoning about Knowledge*. Cambridge, Massachusetts: The MIT Press.

Fagin, R.; Halpern, J. Y.; and Vardi, M. Y. 1992. What can machines know? on the properties of knowledge in distributed systems. *Journal of the ACM* 39(2):328–376.

Fitting, M. 1996. *First-Order Logic and Automated Theorem Proving*. Springer Verlag, 2nd edition edition.

Halpern, J. Y., and Zuck, L. D. 1992. A little knowledge goes a long way: knowledge-based derivations and correctness proofs for a family of protocols. *Journal of the ACM* 39(3):449–478.

Hegel, G. W. F. 1929. *Hegel's science of logic*. Allen & Unwin.

Hintikka, J. 1962. *Knowledge and Belief: An Introduction to the Logic of The Two Notions*. Ithaca, New York: Cornell University Press.

Kant, I. 1893. *Critique of pure reason*. London: Bell.

Nelson, L. 1949. *Socratic Method and Critical Philosophy*. Yale University Press.

Popper, K. R. 1959. *The Logic of Scientific Discovery*. Cambridge University Press.

Popper, K. R. 1969. *Conjectures and Refutations: The Growth of Scientific Knowledge*. Routledge and Kegan Paul, 3rd edition.

Popper, K. R. 1972. *Objective Knowledge: An evolutionary approach*. Oxford University Press.

Prior, A. 1967. *Past, Present and Future*. Oxford: Oxford University Press.

Rao, A., and Georgeff, M. 1991a. Asymmetry thesis and side-effect problems in linear-time and branching-time intention logics. In Myopoulos, J., and Reiter, R., eds., *Proceedings of the 12th International Joint Conference on Artificial Intelligence (IJCAI-91)*, 498–505. Sydney, Australia: Morgan Kaufmann publishers Inc.: San Mateo, CA, USA.

Rao, A., and Georgeff, M. 1991b. Modelling rational agents within a BDI-architecture. In Fikes, R., and Sandewall, E., eds., *Proceedings of the Second International Conference on Principles of Knowledge Representation and Reasoning*, 473–484.

Rao, A., and Georgeff, M. 1998. Decision procedures for bdi logics. *Journal of Logic and Computation* 8(3):293344.

Rao, A. 1996. AgentSpeak(L): BDI Agents speak out in a logical computable language. In Van de Velde, W., and Perram, J., eds., *Proceedings of the Seventh Workshop on Modelling Autonomous Agents in Multi-Agent World (MAAMAW'96)*, volume 1038 of *Lecture Notes in Artificial Intelligence*, 42–55. Eindhoven, The Netherlands: Springer-Verlag.

Rosenschein, S. J. 1985. Formal theories of knowledge in ai and robotics. *New Generation Computing* 345–357.

Trần, B.; Harland, J.; and Hamilton, M. 2003. A Combined Logic of Expectation and Observation (A generalization of BDI logics). In Leite, J. A.; Omicini, A.; Sterling, L.; and Torroni, P., eds., *Declarative Agent Languages and Technologies, First International Workshop, DALT 2003. Melbourne, Victoria, July 15th, 2003. Workshop Notes*, 162–177.

van Benthem, J. 1983. *Modal Logic and Classical Logic*. Naples: Bibliopolis.

van der Meyden, R. 1998. Common knowledge and update in finite environments. *Information and Computation* 140(2):115–157.

Wooldridge, M., and Lomuscio, A. 2000. Reasoning about visibility, perception, and knowledge. In Jennings, N., and Lespérance, Y., eds., *Intelligent Agents VI*, volume Lecture Notes in AI Volume. Springer-Verlag.

Wooldridge, M., and Lomuscio, A. 2001. A computationally grounded logic of visibility, perception, and knowledge. *Logic Journal of the IGPL* 9(1):257–272.

Wooldridge, M. 1992. *The Logical Modelling of Computational Multi-Agent Systems*. PhD thesis, University of Manchester.

Wooldridge, M. 2000. Computationally Grounded Theories of Agency. In Durfee, E. H., ed., *Proceedings of the Fourth International Conference on Multi-Agent Systems (ICMAS 2000)*, volume 9. IEEE Press.