2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology

Levels of Modalities for BDI Logic

Jeff Blee, David Billington, Guido Governatori, and Abdul Sattar
Institute for Integrated and Intelligent Systems, Griffith University, 170 Kessels Road, Nathan, Queensland 4111, Australia.
{J.Blee, D.Billington}@griffith.edu.au, guido@itee.uq.edu.au, A.Sattar@griffith.edu.au

Abstract

The use of rational agents for modelling real world problems has both been heavily investigated and become well accepted, with BDI Logic being a widely used architecture to represent and reason about rational agency. However, in the real world, we often have to deal with different levels of confidence in our beliefs, desires, and intentions. This paper extends our previous framework that integrated qualitative levels into BDI Logic. We describe an expanded set of axioms and properties of the extended logic and also define a detailed non-normal Kripke type semantics.

1. Introduction

Agent Technologies are now well recognized for reasoning about complex real world problems in information and communication [1]. BDI Logic, in particular, is most widely studied for modeling rational agents [2-4] and is used in agent languages such as AgentSpeak and architectures such as JACK and JASON. In BDI, human like reasoning may be modeled by capturing the mentalistic notions of belief, desire, and intention.

In our previous work [5], we argued that in the real world, an agent must have an ability to reason with different levels of mentalistic notions, such as strong belief, moderate belief, weak belief, disbelief, etc. These levels of an agent's attitudes reflect the degree of its confidence about its beliefs, desires, and intentions and thereby allow more versatility in modelling situations. The framework we presented in [5] integrated 5 basic levels of grading of each of the BDI mentalistic notions. In this paper, we refine and extend that framework, and add a detailed semantics.

In Section 2, we present the updated syntax, along with the revised levels of belief, the axioms and properties for belief, and levels of goals (desires and intentions). Section 3 contains the detailed non-normal type Kripke semantics of this logic. The conclusion and future intentions are in Section 4.

2. Framework Syntax

The alphabet of this framework is the union of the following pairwise disjoint sets of symbols: a non-empty countable set $\mathcal{P}$ of atomic propositions; the set $\{\wedge,\vee,\rightarrow,\ \}$ of connectives; the set of brackets {(,[,),]}; and a set of modalities $\{BA, BU_1, BU_2, \ldots, BU_n, BE, BI, BW_n, BW_{n-1}, \ldots, BW_1, BD, DA, DU_1, DU_2, \ldots, DU_n, DE, DI, DW_n, DW_{n-1}, \ldots, DW_1, DD, IA, IU_1, IU_2, \ldots, IU_n, IE, II, IW_n, IW_{n-1}, \ldots, IW_1, ID\}$.

The syntax of the language is as follows:

$$\varphi\ ::-\ p \mid (\Phi\varphi) \mid (\ \ \varphi) \mid (\varphi_1\wedge\varphi_2) \mid (\varphi_1\vee\varphi_2) \mid (\varphi_1\rightarrow\varphi_2)\ ;$$

where $\varphi\in\mathcal{L}$ ($\mathcal{L}$ is the set of all formulae of the alphabet), $p\in\mathcal{P}$, and Φ is an element of the set of modalities.

We write $(\varphi_1 \equiv \varphi_2)$ to abbreviate $(\varphi_1 \rightarrow \varphi_2)\wedge(\varphi_2 \rightarrow \varphi_1)$.

2.1. Levels of Belief

In [5], 5 levels of belief (as well as desire and intention) were presented. BA is absolute belief (equivalent to the normal belief of doxastic logic), BU was 'usual belief' BI doxastic ignorance, BW weak belief or 'usually not' believed and BD signifies disbelief. In this paper, the BU and BW levels are now expanded to *n* levels of each. BI (Doxastic Ignorance) is now altered as follows. *Doxastic Ignorance* denotes no belief is held at all and is similar to the logic presented in [6]. *Doxastic Equivalence* denotes equal evidence for and against a belief. A description of the various levels of belief is now presented.

Description : The belief levels are defined as follows:

BAφ means φ is believed absolutely and is the strongest level (e.g. "the sun will rise tomorrow").

$BU_i\varphi$ means φ is usually believed true (e.g. φ is "the bus will be on time") where $i\in\{1,2,...,n\}$ and

978-0-7695-3496-1/08 $25.00 © 2008 IEEE
DOI 10.1109/WIIAT.2008.231

the number of levels n within BU is domain and application dependent. So, $BU_1\varphi$ means φ is believed less than at BA and more than $BU_2\varphi$ and so on down to BU_n which signifies belief slightly higher than BE and perhaps BI.

$BE\varphi$ means φ is equally believed and disbelieved, that is the agent has equal evidence for and against φ. We label it as *Doxastic Equivalence*.

$BI\varphi$ means φ is neither believed nor disbelieved and is belief absence (labeled *Doxastic Ignorance*).

$BW_i\varphi$ means φ is usually not believed, or only weakly believed, where the number of levels within BW will be the same as within BU. BW_i is the mirror opposite of BU_i so that BW_n is belief slightly less than BE and so on down to BW_1 which is belief slightly higher than total disbelief.

$BD\varphi$ means φ is absolutely disbelieved and is BA's mirror (e.g. φ is "a comet will hit my house today").

Doxastic possibility (P) is the ◊ (diamond) to belief's □ (box) (e.g. $PU_i\varphi \equiv \quad BU_i \quad \varphi$).

The mirroring described above suggests the ability to cut the levels down by approximately half by eliminating BD and the BW levels. However, with only BA, the BU levels, BE, and BI, priority direction between levels could alter depending on a formula's sign. Therefore, all levels are retained to simplify the reasoning using essentially atomic (positive) formulae.

2.2. Belief Axioms and Properties

The belief axioms, and properties that follow from those axioms,. were presented in [5], and we briefly restate them here to include the framework changes. Axiom numbering is prefixed by the letter "A" and properties by the letter "P". Let $\varphi,\psi \in \mathcal{L}$, $n \in \mathbb{Z}^+$, $i \in \{1,2,...,n\}$, $Levels = \{A,U_1,U_2,\ldots,U_n,E,I,W_n,W_{n-1},...,W_1,D\}$, and $\Phi,\Psi,\Omega \in Levels$.

$$BA\varphi \vee BU_1\varphi \vee BU_2\varphi \vee \ldots \vee BU_n\varphi \vee BI\varphi \vee BE\varphi \vee BW_{n-1}\varphi \vee BW_{n-2}\varphi \vee \ldots \vee BW_1\varphi \vee BD\varphi \,. \quad \text{(A1)}$$

$$\text{If } \Phi \neq \Psi, \text{ then } B\Phi\varphi \rightarrow \quad B\Psi\varphi \,. \quad \text{(A2)}$$

$$\text{If } \varphi \equiv \psi, \text{ then } B\Phi\varphi \equiv B\Phi\psi. \quad \text{(A3)}$$

$$BA\varphi \equiv BD \quad \varphi \,. \quad \text{(A4)}$$

$$BD\varphi \equiv BA \quad \varphi \,. \quad \text{(P1)}$$

$$\text{For each } i \in \{1,2,...,n\}, \quad BU_i\varphi \equiv BW_i \quad \varphi \,. \quad \text{(A5)}$$

$$\text{For each } i \in \{1,2,...,n\}, \quad BW_i\varphi \equiv BU_i \quad \varphi \,. \quad \text{(P2)}$$

$$BE\varphi \equiv BE \quad \varphi \,. \quad \text{(A6)}$$

$$BI\varphi \equiv BI \quad \varphi \,. \quad \text{(A7)}$$

A1 and A2 reflect that one, and only one, of the belief levels hold for each φ in $\mathcal{L}$. A3 states that belief does not depend on the syntax of the formula. That the belief modalities BD to BW_n are the mirror opposites of BA to BU_n respectively is reflected in A4, A5, P1, and P2. A6 reflects the intuition that if a formula φ is held in *Doxastic Equivalence*, the agent cannot decide between believing for or against φ (perhaps due to equal evidence/belief for φ and φ), therefore φ is similarly held in doxastic equivalence. If φ is held in *Doxastic Ignorance* the agent has no level of actual belief in φ (perhaps due to no knowledge about φ) and therefore φ is similarly held in doxastic ignorance (reflected in A7). P3 to P5 follow.

$$B\Pi\varphi \rightarrow \quad BI\varphi \wedge \quad BI \quad \varphi \text{ (where } \Pi \in Levels - \{I\}). \quad \text{(P3)}$$

$$BI\varphi \rightarrow \quad BA\varphi \wedge \quad BA \quad \varphi \wedge \quad BU_1\varphi \wedge \quad BU_2\varphi \wedge \ldots \wedge \quad BU_n\varphi \wedge \quad BE\varphi \wedge \quad BU_1 \quad \varphi \wedge \quad BU_2 \quad \varphi \wedge \ldots \wedge \quad BU_n \quad \varphi. \quad \text{(P4)}$$

$$BI\varphi \rightarrow \quad BA\varphi \wedge \quad BD\varphi \wedge \quad BU_1\varphi \wedge \quad BU_2\varphi \wedge \ldots \wedge \quad BU_n\varphi \wedge \quad BE\varphi \wedge \quad BW_n\varphi \wedge \quad BW_{n-1}\varphi \wedge \ldots \wedge \quad BW_1\varphi. \quad \text{(P5)}$$

Definition 1: $P\Phi\varphi = \quad B\Phi \quad \varphi$ (and $B\Phi\varphi = \quad P\Phi \quad \varphi$).

$$BI\varphi \equiv \quad PI\varphi \,. \quad \text{(P6)}$$

P6 follows from A7 and $BI \quad \varphi \equiv \quad PI \quad \varphi \equiv \quad PI\varphi$.

Definition 1.1: An underlined level, $\underline{\Phi}$,is defined as the mirror of the non-underlined level Φ, that is,

$\underline{A} = D$, $\underline{U_i} = W_i$, $\underline{E} = E$, $\underline{I} = I$, $\underline{W_i} = U_i$, and $\underline{D} = A$, where $i \in \{1,2,...,n\}$.

$$P\Phi\varphi \equiv \bigvee[\{B\Psi\varphi : \Psi \in Levels\} - \{B\underline{\Phi}\varphi\}] \quad \text{(P7)}$$

P7 to P11 from [5] are combined to the new P7.

2.2.1. KD45 Axioms: Doxastic Logic and multi-modal BDI Logic includes the axiom system KD45. In [5] it was discussed why these axioms don't hold for our framework. Briefly, it has been shown elsewhere [7], that the K axiom does not hold over modality gradings with K for the BD belief level ($BD(\varphi\rightarrow\psi)\rightarrow(BD\varphi\rightarrow BD\psi)$) being plainly counterintuitive when applied to any example. The axiom D works alright except for the levels BE and BI. For example, using Definition 1 and A6, $BE\varphi \rightarrow PE\varphi$ is converted to $BE\varphi \rightarrow \quad BE\varphi$ which is plainly wrong. BI has the same problem. Axioms 4 and 5 are used for positive and negative introspection respectively, but 4 for BD (i.e. $BD\varphi \rightarrow BD\,BD\varphi$), using P1 and replacing φ with ψ, becomes $BA\psi \rightarrow \quad BA \quad BA\psi$. This is not what we want and the axiom 5 has similar problems for negative introspection. However, alterations, as in A8 to A11, allow us to partially retain K, as well as D, and positive and negative introspection.

$$B\Phi(\varphi \rightarrow \psi) \rightarrow (BA\varphi \rightarrow B\Phi\psi) . \quad \text{(A8)}$$

$$B\Pi\varphi \rightarrow P\Pi\varphi \text{ where } \Pi \in (Levels - \{E, I\}) . \quad \text{(A9)}$$

$$B\Phi\varphi \rightarrow BAB\Phi\varphi . \quad \text{(A10)}$$

$$P\Phi\varphi \rightarrow BAP\Phi\varphi . \quad \text{(A11)}$$

These altered KD45 axioms hold for all belief levels. The only possible problem is regarding the BI and BE levels, in not holding for the D axiom. Due to BI not being a level of actual belief and BE not being definite belief, this problem can be essentially ignored.

2.3. Levels of Goals

Desires can be considered as weak goals and intentions as strong goals (desire + commitment = intention). As with beliefs, it is quite conceivable for an agent to have desires of varying degrees of strength, and allowing this gives our agent more versatility in representing a wider range of situations.

The framework of levels introduced for beliefs is extended to desires and also intentions. Essentially, goals (desires/intentions) are divided into levels in a similar manner to that of beliefs. The difference between goals being that a given desire, among several desires, will only become an intention if it is committed to by the agent. The agent may have many conflicting desires but should rationally have no conflicting intentions. Due to space limitations, the description of desire and intention levels is omitted, but can be intuitively deduced from that for beliefs.

The symbol P_D represents desire possibility and is the diamond (◊) to desire's box (□) ($P_D\Phi\varphi = \ D\Phi\ \ \varphi$). Similarly P_I represents intention possibility ($P_I\Phi\varphi = \ I\Phi\ \ \varphi$). The desire and intention axioms and properties are essentially the same as those of belief, with the letter B being replaced by D or I respectively and P being replaced by P_D or P_I respectively. Therefore we will not take up space here stating them. So, after deliberation, an agent commits to a particular desire, thereby creating an intention of the same level (e.g. $DU_2\varphi$ + *commitment* = $IU_2\varphi$). Naturally there are no direct goal equivalents to the belief introspection axioms, A10 and A11. However, there may be further altered introspection axioms of the form of $D\Phi\varphi \rightarrow BA\, D\Phi\varphi$, so that if φ is desired at a particular level, the agent believes φ desired at that level. Similarly, $P_D\Phi\varphi \rightarrow BA\, P_D\Phi\varphi$ may be applied for negative introspection of desires. Naturally there may be similar extensions for intention.

3. Semantics

The semantics of this logic builds on the semantics of Rao & Georgeff's original logic [3], but using a Minimal / Neighbourhood semantics [8], to allow for the addition of levels of modality. We drop the CTL aspect of the original logic to reduce complexity.

Definition 2: The model is a Kripke structure $\mathcal{M} = \langle W, \beta, \delta, \iota, Th, NB, ND, NI, V\rangle$, where

1) W is a set of worlds.
2) β, δ, and ι are functions for belief, desire and intention respectively which take a world w and a set of worlds X believed (resp. desired, intended) at a level of belief (resp. desire, intention) and returns a number that represents the level of belief (resp. desire, intention) that X is held in at w. Suppose $\gamma \in \{\beta, \delta, \iota\}$ and $X, Y \subseteq W$, then we require:
   a) $\gamma: W \times 2^W \rightarrow (\mathbb{Q} \cap [0,1]) \cup \{u\}$,
      where $\mathbb{Q}$ = rational numbers and u = *undefined*.
   b) $\gamma(w, X) = 1 - \gamma(w, W - X)$,
      and $\gamma(w, X) = u$ iff $\gamma(w, W - X) = u$.
   c) $\gamma(w,X) = 0$ iff $X = \{\}$, and $\gamma(w,X) = 1$ iff $X = W$.
   d) $\forall w' \in Y$, $\beta(w, X) = \beta(w, Y)\bullet\beta(w', X)$
      (For arithmetic working, u is assigned a domain dependent value,(0·5 in the scheduling example)).
3) Th is a threshold function that designates the range covered by each of the levels of belief, desire, or intention. We define Th: $Levels \rightarrow ((\mathbb{Q} \cap [0,1]) \cup \{u\}$ $\times (\mathbb{Q} \cap [0,1]) \cup \{u\}$ (u = *undefined*) such that:
   a) $Th(A) = (1,1)$, $Th(E) = (0{\cdot}5,\ 0{\cdot}5)$, $Th(I) = (u, u)$, and so by 2-2-b, $Th(D) = (0, 0)$.
   b) If $\Phi \neq I$ and $Th(\Phi) = (s,f)$ then $s \neq u$, $f \neq u$, $s \leq f$, and $Th(\underline{\Phi}) = (1-f,\ 1-s)$.
   c) If $i \in \{1,2,...,n\}$ and $Th(U_i) = (s_i, f_i)$, then
      $0{\cdot}5 < s_n < f_n < s_{n-1} < ... < f_2 < s_1 < f_1 < 1$.
4) NB, ND, and NI are 'neighborhood' functions that each map a world and a level to a set of sets of worlds which are belief (resp. desire, intention) accessible to the given world at the given level of belief (resp. desire, intention). If $Z \in \{B,D,I\}$, $B' = \beta$, $D' = \delta$, and $I' = \iota$, then we define NZ: $W \times Levels \rightarrow 2^{2^W}$ as follows. If $\Phi \neq I$ then $NZ(w,\Phi) = \{X \subseteq W : Th(\Phi) = (s, f) \text{ and } s \leq Z'(w,X) \leq f\}$, else $NZ(w, I) = \{X \subseteq W : Z'(w, X) = u\}$. NZ is defined under the following conditions:
   a) If $X \in NZ(w,A)$ and $Y \in NZ(w,\Phi)$ then $X \cap Y \in NZ(w,\Phi)$.
   b) If $X \in NB(w,\Phi)$ then $\{w' : X \in NB(w',\Phi)\} \in NB(w,A)$.
   c) If $X \notin NB(w,\Phi)$ then $\{w' : X \notin NB(w',\Phi)\} \in NB(w,A)$.
   d) $W \in \cup\{NZ(w, \Phi) : \Phi \in Levels\}$ for every $w \in W$.

   Note: c) and d) may be extended to include a goal is believed a goal (change the first two NBs to NZs).

5) V is a function mapping an atomic formula at a world to a truth value. We define V: $\mathcal{P} \times W \rightarrow \{false, true\}$ by $V(p,w) =$ *false* if p is not true at w, or *true* if p is true at w.

Definition 3: The truth of formulae φ in $\mathcal{L}$ is defined:

1) If $\mathcal{M} = \langle W, Th, \beta, \delta, \iota, NB, ND, NI, V \rangle$ is a model, $p \in \mathcal{P}$, and $w \in W$, then we define

$$(\mathcal{M}, w) \vDash p \text{ iff } V(p, w) = true,$$

2) If $\mathcal{M} = \langle W,Th,\beta,\delta,\iota,NB,ND,NI,V \rangle$ is a model, $Z \in \{B,D,I\}$, $\Phi \in Levels$, $w,w' \in W$, and $\varphi \in \mathcal{L}$, the truth of belief, desire, and intention formulae are defined by:

$$(\mathcal{M}, w) \vDash Z\Phi\varphi \text{ iff } ||\varphi|| \in NZ(w, \Phi)$$

$$\text{where } ||\varphi|| = \{w \in W : (\mathcal{M}, w) \vDash \varphi\}.$$

3) If $\mathcal{M} = \langle W,Th,\beta,\delta,\iota,NB,ND,NI,V \rangle$ is a model, $w \in W$, and $\varphi \in \mathcal{L}$, then $(\mathcal{M}, w) \vDash \varphi$ is defined by the standard conditions for the boolean connectives.

Now that $(\mathcal{M}, w) \vDash \varphi$ for an arbitrary formula φ is defined, we have the following definition of a *truth set*.

Definition 4: If φ is any formula then the *truth set* of φ, $||\varphi||$, is defined by

$$||\varphi|| = \{w \in W : (\mathcal{M}, w) \vDash \varphi\}.$$

Reasoning among various mentalistic levels is accomplished through priorities, or a total order, between levels. Note that the place in this order of the level I (with a γ value of u) is domain dependent. It may, most commonly, be placed equal with E at 0·5. For belief, this order of levels would effectively be $BA>BU_1>BU_2>...>BU_n>BE=BI>BW_n>BW_{n-1}>...>BW_1>BD$, and given $\Phi,\Psi \in Levels$, $B\Phi\varphi > B\Psi\psi$ signifies $B\Phi\varphi$ has a higher priority than $B\Psi\psi$. So, $B\Phi\varphi > B\Psi\psi$ means that $||\varphi|| \in NB(w,\Phi)$, $||\psi|| \in NB(w,\Psi)$, and $\beta(w,||\varphi||) > \beta(w,||\psi||)$. This latter may be intuitively written as $\beta(\varphi) > \beta(\psi)$, as $\beta(w, ||\varphi||)$ may be simplified to $\beta(\varphi)$ where the current world w is obvious. Reasoning among desires is carried out in a similar manner. Intentions would be created by the agent committing to the desire (selected from among competing desires) with the highest δ value.

Theorem 1: The logic and the axioms and properties are sound and complete with respect to the semantics (Definitions 1 to 4 inclusive). See http://www.freewebs.com/meaty4ever/XBDI/proofs.pdf for proofs.

An example demonstrating this logic was unable to be included due to space limitations, but can be seen at http://www.freewebs.com/meaty4ever/XBDI/schedule_example.pdf .

4. Conclusion and Future Work

The framework in this paper provides a foundation for a layered BDI architecture. This essentially enables a rational agent to capture commonsense reasoning. We believe that representing and reasoning with levels of mentalistic attitudes significantly enhances an agent's ability to perform human-like practical reasoning in complex domains. The proposed framework is simpler and more intuitive than other BDI frameworks. Intended future work involves implementing this basic framework into an existing BDI platform, most likely the AgentSpeak(L) platform, JASON. Naturally this may necessitate dropping the strictly modal aspects of this logic, but the semantics will be able to be adapted to JASON.

Acknowledgement: We thank the Smart Internet Cooperative Research Centre (SITCRC) for their funding towards this work.

References

[1] M. Luck, P. McBurney, O. Shehory, and S. Willmott, "Agent Technology Roadmap: A Roadmap for Agent Based Computing," AgentLink III, 2005.

[2] M. E. Bratman, *Intention, Plans, and Practical Reason*, Harvard University Press. Cambridge, MA, 1987.

[3] A. S. Rao and M. P. Georgeff, "Decision Procedures for BDI Logics," *Journal of Logic and Computation*, vol. 8, 1998, pp. 293-343.

[4] A. S. Rao, "AgentSpeak(L): BDI Agents speak out in a logical computable language," presented at 7th European Workshop on Modelling Autonomous Agents in a Multi-Agent World, Springer Verlag, Eindhoven, The Netherlands, 1996, pp. 42-55.

[5] J. Blee, D. Billington, and A. Sattar, "Reasoning with Levels of Modalities in BDI Logic," presented at Pacific Rim International Workshop on Multi-Agents (PRIMA07), Springer Verlag, Bangkok, 2007.

[6] W. van der Hoek and A. Lomuscio, "A Logic for Ignorance," in *Electronic Notes in Theoretical Computer Science*, vol. 85, Elsevier, 2004.

[7] W. van der Hoek, "On the Semantics of Graded Semantics," *Journal of Applied Non-Classical Logics*, vol. 2, 1992, pp. 81-123.

[8] B. F. Chellas, *Modal Logic: An Introduction*, Cambridge University Press. Cambridge, MA, 1980.