Trust and Belief, Interrelation

Besik Dundua[1] and Levan Uridia[2]

[1] DCC-FC & LIACC, University of Porto, Portugal
bdundua@dcc.fc.up.pt
[2] Universidad Rey Juan Carlos de Madrid, Spain
uridia@ia.urjc.es

Abstract. We introduce the modal system $\mathbf{B}_T^2$ which is a multi-modal language designed to talk about trust and belief of two agents. The belief operators are based on modal system **KS**. This gives the main difference with already known system **BA** introduced by Churn-Jung Liau [2] and also carries its own intuitive meaning. As a main result we prove that $\mathbf{B}_T^2$ is sound and complete with respect to the given semantics, which is a mixture of the Kripke and neighborhood semantics.

1 Introduction

In a notion derived from Plato's dialogue Theaetetus, philosophy has traditionally defined *knowledge* as *justified true belief*. The relationship between belief and knowledge is that a belief is a knowledge if the belief is true, and if the believer has a justification for believing it is true. Dropping out the part about justification we can express this connection in a standard modal language by well known translation Sp (splitting translation from **S5** to **KS**), The main clause of which states: $Sp(\mathcal{K}p) = \mathcal{B}p \wedge p$. Mainly because of this connection together with some doxastic properties of the axiom system, the modal logic **KS** could be adopted as a good concurrent to the classical doxastic logic **KD45**. First introduced in the 1980s by Segerberg as a modal logic of *some other time* [8], **KS** was rediscovered as a modal logic of *inequality* [5], [7] and recently it was investigated as a doxastic logic which carries quite adequate properties for modelling beliefs of agents [1]. In this paper we consider only those agents whose doxastic properties are formalized in the modal logic **KS**.

Different types of trust have been proposed and studied in the disciplines like philosophy, economics, computer science, etc. In this paper we focus on the interrelation of trust and belief in the two agents case. On the one hand, we follow the ideas introduced in [2], but on the other hand, we simplify the language in a sense that we leave only two types of modalities: $\square_i$ as a belief operator of agent i and $T_{i,j}$ for trust of agent-i in agent-j. So the logic $\mathbf{B}_T^2$ we introduce is especially designed to talk about belief and trust and their interrelation. As for belief operators, as we already mentioned, they satisfy the axioms of **KS** and this makes the main difference with the logic **BA** discussed in [2].

The paper is organized in the following way: In section 2 we recall the logic **KS**, provide basic definitions and some known facts. In section 3 we introduce

35

the logic $\mathbf{B}_T^2$ and also give its semantics on Kripke's structures. In the same section we prove the main result of the paper which states that the logic $\mathbf{B}_T^2$ is sound and complete with respect to the given semantics.

2   The Modal Logic KS

In 1976 Krister Segerberg [8] explicitly formulated a modal logic **KS** in which the diamond modality $\Diamond$ is interpreted as "somewhere else". In this section we define the system KS and its Kripke semantics.

2.1   Syntax

Definition 2.11 *The normal modal logic* **KS** *is defined in a standard modal language with infinite set of propositional letters $p, q, r, ..$ and connectives $\wedge, \square, \neg,$*

- *The axioms are all classical tautologies plus three axioms containing modal operators. Namely:*

$\square(p \rightarrow q) \rightarrow \square p \rightarrow \square q,$
$\square p \wedge p \rightarrow \square\square p,$
$p \rightarrow \square\Diamond p,$ *where* $\Diamond p \equiv \neg\square\neg p.$

- *The rules of inference are: Modus-ponens, Substitution and Necessitation.*

Observe that doxastic interpretation of the last axiom states that *If p is true then agent believes that it is not the case that he believes the negation of p.*
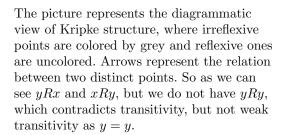
It is an easy task to show that if we add axiom $\square p \rightarrow p$ to the **KS** we will get the classical epistemic system **S5**. Following Smullyan [9] this means that if the **KS**-reasoner is *accurate* (never believes any false proposition) then his beliefs coincide with his knowledge.

2.2   Kripke semantics

Kripke semantics for the modal logic **KS** is provided by weakly-transitive and symmetric Kripke frames. Below we give the definition of weakly-transitive relation.

Definition 2.21 *We will say that a relation $R \subseteq W \times W$ is weakly-transitive if $(\forall x, y, z)(xRy \wedge yRz \wedge x \neq z \Rightarrow xRz).$*

Obviously every transitive relation is weakly-transitive also. Moreover it is immediate to notice that weakly-transitive relations differ from transitive ones just by the occurrence of irreflexive points in clusters. As you can see the frame on the picture is weak transitive, but not transitive.

The picture represents the diagrammatic view of Kripke structure, where irreflexive points are colored by grey and reflexive ones are uncolored. Arrows represent the relation between two distinct points. So as we can see yRx and xRy, but we do not have yRy, which contradicts transitivity, but not weak transitivity as $y = y$.

pic. 1

In the study of modal logic the class of rooted frames plays a central role. Recall that a frame (W, R) is rooted if it contains a point $w \in W$, which can see all other points in W. That is $R(w) \supseteq W - \{w\}$, where $R(w)$ is the set of all successors of w. The class of all rooted, weakly-transitive and symmetric frames can be characterized by the property which we call weak - cluster.

Definition 2.22 *We will say that a relation $R \subseteq W \times W$ is weak-cluster if $(\forall x, y)(x \neq y \Rightarrow xRy)$.*

It is easy to see that every weak-cluster is just a cluster where we allow irreflexive points. We will see the detailed characterization of finite weak-clusters in the next section. The following proposition makes the link between weak-clusters and rooted, weakly-transitive, symmetric frames.

Proposition 2.23 *A frame (W, R) is rooted, weakly-transitive and symmetric iff it is weak-cluster.*

Proof. It is immediate that every weak-cluster is rooted, weakly-transitive and symmetric frame. For the other direction let (W, R) be rooted, weakly-transitive and symmetric frame. Let $w \in W$ be the root. Take arbitrary two distinct points $x, y \in W$. As w is the root, we have: wRx and wRy. Because of symmetry we get xRw. Now as R is weakly-transitive, from $xRy \wedge yRw$ and $x \neq y$ we get xRy. Hence R is a weak-cluster.

For the sake of completeness we will just briefly state the main definitions, like: Kripke model, satisfaction and validity of modal formulas. These definitions are standard and can be found in any modal logic book.

Definition 2.24 *The pair (W, R), with W an arbitrary set (set of possible worlds) and $R \subseteq W \times W$ is called a Kripke frame.*
If we additionally have a third component $V : Prop \times W \to \{0, 1\}$, then we say that we have a Kripke model $M = (W, R, V)$ (Here Prop denotes the set of all propositional letters).

The satisfaction and validity of a modal formula are defined inductively. We just state the base and modal cases here.

Definition 2.25 *For a given Kripke model $M = (W, R, V)$ the satisfaction of a formula at a point $w \in W$ is defined inductively as follows: $w \Vdash p$ iff $V(p, w) = 1$, the boolean cases are standard, $w \Vdash \Box\phi$ iff $(\forall v)(wRv \Rightarrow v \Vdash \phi)$.*

We will say that a formula ϕ is valid in a model (W, R, V) if for every point $w \in W$ we have $w \Vdash \phi$. We will say that a formula ϕ is valid in a frame (W, R) if it is valid in every model (W, R, V) based on a frame (W, R). We will say that a formula ϕ is valid in a class of frames C if ϕ is valid in every frame $(W, R) \in C$.

So far we defined the modal logic **KS** syntactically and we gave the definition of weak-cluster relation. The following theorem links these two notions:

Theorem 2.26 *[4] The modal logic **KS** is sound and complete w.r.t. the class of all **finite, irreflexive** weak-cluster relations.*

Mainly because of the theorem 2.26 the modal logic **KS** is called the modal logic of inequality. As the reader can easily check the interpretation of box in irreflexive weak-clusters boils down to the following: $w \Vdash \Box\phi$ iff $(\forall v)(w \neq v \Rightarrow v \Vdash \phi)$.

3   Trust

In this section we extend the language for multi-agent case and in addition we add modalities for trust. We restrict the language for the case with two agents as far as the other cases (for finite agents) follow as an easy generalizations of two agent case. We take the ideas from [2] and introduce modal logic which has enough expressive power to talk about trust and belief. We do not consider the same language as in [2], but just its fragment, since we are only interested in interrelation between trust and belief. We give the semantics for this logic and as a main result of the paper we prove the completeness of the described logic with respect to the semantics. Main crucial difference from [2] lies in the fact that the doxastic properties of agents follow **KS** axioms not **KD45** axioms as is classically adopted.

3.1   Syntax

The language consists of infinite set of propositional letters $p, q, r, ...$, connectives $\vee, \wedge, \neg, \rightarrow$, and modalities $\Box_1, \Box_2$ $T_{1,2}, T_{2,1}$

Axioms:Each $\Box_i$ satisfies **KS** axioms,

$\vdash T_{1,2}p \leftrightarrow \Box_1 T_{1,2}p,$
$\vdash T_{2,1}p \leftrightarrow \Box_2 T_{2,1}p.$

Rules of inference are: Modus ponens and substitution for each modality, necessitation for $\Box_i$ where $i \in \{1, 2\}$ and the following rule $\frac{\vdash p \leftrightarrow q}{\vdash T_{i,j}p \leftrightarrow T_{i,j}q}$ for each

$T_{i,j}$ with $i, j \in \{1, 2\}$ and $i \neq j$.

The desired interpretation of $T_{i,j}p$ carries the following idea: Agent-i trusts agent-j about the claim p. In these settings the last two axioms have very intuitive meaning, mainly: Agent-i trusts agent-j about the claim p iff Agent-i believes that he trusts agent-j about p. Hence trust does not contradict one's beliefs. These are the only (very natural) restrictions we have on the interrelation between trust and belief.

3.2 Semantics

Kripke semantics is provided by bi-relational Kripke frames together with two neighborhood functions. More formally

Definition 3.21 *A $\mathbf{B}_T^2$-frame $\mathfrak{F}$ is a tuple $(W, R_1, R_2, u_{1,2}, u_{2,1})$, where:*

W is a set of possible worlds,

$R_1, R_2 \subseteq W \times W$ are weakly transitive and symmetric relations,

$u_1 : W \to PP(W), u_2 : W \to PP(W)$ are functions (neighborhood maps), such that the following equalities take place: $u_{i,j}(w) = \bigcap_{v \in R_i(w)} u_{i,j}(v)$, for every $i, j \in \{1, 2\}$ where $i \neq j$.

$\mathbf{B}_T^2$-model is a pair $M = (\mathfrak{F}, V)$, where $\mathfrak{F}$ is a $\mathbf{B}_T^2$-frame and $V : Prop \to P(W)$ is a valuation.

Definition 3.22 *A satisfaction of a formula in a given $\mathbf{B}_T^2$-model $M = (\mathfrak{F}, V)$ and a point $w \in W$ is defined inductively as follows:*
$w \Vdash p$ iff $w \in V(p)$,
$w \Vdash \neg\alpha$ iff $w \nVdash \alpha$,
$w \Vdash \alpha \wedge \beta$ iff $w \Vdash \alpha$ and $w \Vdash \beta$,
$w \Vdash \Box_i\alpha$ iff $(\forall w')(wR_iw' \Rightarrow w' \Vdash \alpha)$,
$w \Vdash T_{i,j}\alpha$ iff $|\alpha| \in u_{i,j}(w)$. Here $|\alpha|$ denotes the set $\{v | v \Vdash \phi\}$.

A formula is valid in in a given $\mathbf{B}_T^2$-model if it is true at every point of the model. A formula is valid in a $\mathbf{B}_T^2$-frame if it is valid in every model based on the frame. A formula is valid in a class of $\mathbf{B}_T^2$-frames if it is valid in every frame in the class.

Theorem 3.23 *The logic $\mathbf{B}_T^2$ is sound and complete with respect to the class of all $\mathbf{B}_T^2$-frames.*

Proof. The soundness easily follows by direct check as for completeness, the proof is standard and therefore we just give a sketch.

Let W be the set of all maximally consistent subsets of formulas in a logic $\mathbf{B}_T^2$. Let us define the relations R_1 and R_2 on W in the following way: For every

$\Gamma, \Gamma' \in W$ we define $\Gamma R_i \Gamma'$ iff $(\forall \alpha)(\Box_i \alpha \in \Gamma \Rightarrow \alpha \in \Gamma')$, where $i \in \{1, 2\}$. The following lemma is proved in [4] when proving completeness of the logic KS. It also directly follows from Sahlqvist theorem and the observation that KS-axioms characterize the class of all weakly-transitive and symmetric frames.

Lemma 1. *[4] Each R_i is weakly-transitive and symmetric.*

So far we defined a set W with two weakly-transitive and symmetric relations R_1, R_2 on it. Now we define functions $u_{1,2}$ and $u_{2,1}$ in the following way:

$$u_{i,j}(\Gamma) = \{\{\Gamma' | \phi \in \Gamma'\} | T_{i,j} \phi \in \Gamma\}.$$

It immediately follows that $u_{i,j}$ are functions defined from W to $PP(W)$. Before we show that $u_{i,j}(\Gamma) = \bigcap_{\Gamma' \in R_i(\Gamma)} u_{i,j}(\Gamma')$, for every $i, j \in \{1, 2\}$, where $i \neq j$ and every $\Gamma \in W$, let us define the valuation and prove the truth lemma.

The valuation V is defined in the following way: $V(\Gamma) = \{p | p \in \Gamma\}$.

Lemma 2 (*Truth*). *For every formula $\alpha \in \mathbf{B}_T^2$ and every point $\Gamma \in W$ of the canonical model, the following equivalence holds: $\Gamma \Vdash \alpha$ iff $\alpha \in \Gamma$.*

Proof. The proof goes by induction on the length of formula. Base case follows immediately from the definition of valuation. Assume for all $\alpha \in \mathbf{B}_T^2$ with length less then k holds: $\Gamma \Vdash \alpha$ iff $\alpha \in \Gamma$.

Let us prove the claim for α with length equal to k. If α is conjunction or negation of two formulas then the result easily follows from the definition of satisfaction relation and the properties of maximal consistent sets, so we skip the proofs. Assume $\alpha = \Box_i \beta$ and assume $\Gamma \Vdash \alpha$. Take a set $B = \{\gamma | \Box_i \gamma \in \Gamma\} \cup \{\neg \beta\}$. The sub claim is that B is inconsistent. Assume not, then there exists $\Gamma' \in W$ such that $\Gamma' \supseteq B$. This by definition of the relation R_i means that $\Gamma R_i \Gamma'$. Now as $\neg \beta \in \Gamma'$, by inductive assumption we get $\Gamma' \Vdash \neg \beta$. Hence we get a contradiction with our assumption that $\Gamma \Vdash \Box_i \beta$. So B is inconsistent. This means that there exists $\gamma_1, \gamma_2, ... \gamma_n \in B$ such that $\vdash \gamma_1 \wedge \gamma_2 \wedge .. \wedge \gamma_n \to \beta$. Applying necessitation rule for $\Box_i$ we get $\vdash \Box_i \gamma_1 \wedge .. \wedge \Box_i \gamma_n \to \Box_i \beta$ so $\Gamma \vdash \Box_i \beta$, hence we get that $\Box_i \beta \in \Gamma$.

We just showed the left-to-right direction of our claim for $\alpha = \Box_i \beta$. For the right-to-left implication assume that $\Box_i \beta \in \Gamma$. By definition of R_i for every Γ' with $\Gamma R_i \Gamma'$ we have $\beta \in \Gamma'$. From this by inductive assumption it follows that $\Gamma' \vdash \beta$. So we imply that $\Gamma \Vdash \Box_i \beta$.

Now assume $\alpha = T_{i,j} \phi$. Assume $\Gamma \Vdash T_{i,j} \phi$. By definition this means that $|\phi| \in u_{i,j}(\Gamma)$. Hence there exists β such that $\{\Gamma'' | \beta \in \Gamma''\} = |\phi|$ with $T_{i,j} \beta \in \Gamma$. This means that we have $\vdash \beta \leftrightarrow \phi$ in $\mathbf{B}_T^2$. Hence by the rule for trust modality we have $\vdash T_{i,j} \beta \leftrightarrow T_{i,j} \phi$. But the last implies that $T_{i,j} \phi \in \Gamma$.

Conversely assume that $T_{i,j} \phi \in \Gamma$ this implies that $\{\Gamma'' | \phi \in \Gamma''\} \in u_{i,j}(\Gamma)$. Now by inductive assumption we know that $\Gamma'' \Vdash \phi$ iff $\phi \in \Gamma''$ hence $|\phi| \in u_{i,j}(\Gamma)$. Hence $\Gamma \Vdash T_{i,j} \phi$.

Now let us show that the model we constructed falls into the class of $\mathbf{B}_T^2$-models. The only thing left to show is the following equality:

$$u_{i,j}(\Gamma) = \bigcap_{\Gamma' \in R_i(\Gamma)} u_{i,j}(\Gamma').$$

Assume $X \in u_{i,j}(\Gamma)$. This means that X is of the form $\{\Gamma'' | \phi \in \Gamma''\}$ for some ϕ with $T_{i,j}\phi \in \Gamma$. Because of the $\mathbf{B}_T^2$ axioms and because Γ is maximally consistent set, we imply that $\Box_i T_{i,j}\phi \in \Gamma$. From this we imply that $T_{i,j}\phi \in \Gamma'$ for every $\Gamma' \in R_i(\Gamma)$. Now by definition of $u_{i,j}$ this means that $\{\Gamma'' | \phi \in \Gamma''\} \in u_{i,j}(\Gamma')$ and as Γ' was arbitrary member of $R_i(\Gamma)$, we get that $X = \{\Gamma'' | \phi \in \Gamma''\} \in \bigcap_{\Gamma' \in R_i(\Gamma)} u_{i,j}(\Gamma')$.

Conversely assume some set $X \subseteq W$ belongs to $\bigcap_{\Gamma' \in R_i(\Gamma)} u_{i,j}(\Gamma')$. By definition this means that there exists a formula ϕ such that $T_{i,j}\phi \in \bigcap_{\Gamma' \in R_i(\Gamma)} \Gamma'$ and $X = \{\Gamma'' | \phi \in \Gamma''\}$. Now as far as $(\forall \Gamma')(\Gamma R_i \Gamma' \Rightarrow T_{i,j}\phi \in \Gamma')$ by truth lemma we get that $(\forall \Gamma')(\Gamma R_i \Gamma' \Rightarrow \Gamma' \Vdash T_{i,j}\phi)$. Hence $\Gamma \Vdash \Box_i T_{i,j}\phi$. Now applying axioms for trust modality we get that $\Gamma \Vdash T_{i,j}\phi$ and hence $X \in u_{i,j}(\Gamma)$. This completes the proof.

4 Conclusions

As a conclusion we mention that the logic described is very much alike to the fragment of BIT defined in [2] and techniques used are also much similar. The only motivation for considering $\mathbf{B}_T^2$ and hence moving from "$KD45$-reasoner" to "KS-reasoner" lies in the future perspective to generalize the semantics of the logic and impose topology instead of neighborhood maps. It is well known that KS has much closer connection with topology then $KD45$. As a future work we use the system PρLog [3] to implement a reasoner for reasoning questions in the $\mathbf{B}_T^2$.

5 Acknowledgments

References

1. P. Blackburn, J.F.A.K. van Benthem, Frank Wolter. *Handbook of Modal Logic.* 2002.
2. Churn-Jung Liau. *Belief, information acquisition, and trust in multi-agent systems - A modal logic formulation.* Journal of Artificial Intelligence, 149, 3160, 2003.

3. J. Coelho, B. Dundua, M. Florido, and T. Kutsia. *A rule-based approach to xml processing and web reasoning*. In P. Hitzler and T. Lukasiewicz, editors, RR, volume 6333 of Lecture Notes in Computer Science, pages 164172. Springer, 2010.
4. L. Esakia. *Weak transitivity-restitution*. Logical Studies 2001, vol 8, 244-255.
5. V. Goranko. *Modal definability in enriched language*. Notre Dame Journal of Formal Logic, 31, 81-105, 1990.
6. P.R. Halmos. *Algebraic Logic*. Chelsea Publishing Company, New York, 1962.
7. M. de Rijke. *The Modal Logic of Inequality*. J. Symbolic Logic, 57, 566-587, 1992.
8. K. Segerberg. ”Somewhere else” and ”some other time”. In: Wright and wrong: *mini-essay in honor of G.H.Wright*, Publ. the group in logic and methodology of Real Finland, 61-64, 1976.
9. R. M. Smullyan. *What Is the Name of This Book?: The Riddle of Dracula and Other Logical Puzzles*.