Computationele logica

Kamans, Jim 10302905

Roosingh, Sander 11983957

Schenk, Stefan 11881798

November 2017

Possibility: $\sim_a := \leq_a \cup \geq_a$

Information Cell: $w(a) := \{s \in W : w \sim_a s\}$

Knowledge: $\|K_a\varphi\|_M = \{s \in W : s(a) \subseteq \|\varphi\|_M\}$

Conditional Belief: $\|B_a^\varphi\psi\|_M = \{s \in W : best_a(\|\varphi\|_M \cap s(a)) \subseteq \|\psi\|_M\}$

Exercise 1

(a) Show the following equivalence:

$$K_a\varphi \Leftrightarrow B_a^{\neg\varphi} false$$

Take $s \in W$ such that $s \models_M K_a\varphi$.

Therefore for all $t \in s(a)$ it's true that: $t \models_M \varphi$.

Using conditional belief: $\|B_a^{\neg\varphi} false\|_M = \{v \in W : best_a(\|\neg\varphi\|_M \cap v(a)) \subseteq \|false\|_M\}$

Eliminating any world or all worlds either results in a world from s(a), where $K_a\varphi$, or no worlds at all, so $false$.

[ANSWER HERE]

(b) Prove semantically the equivalence claimed on Slide 24 of Lecture Notes 4.2:

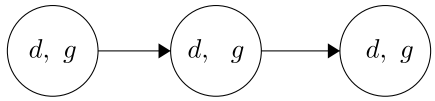
$$B_a\phi \Leftrightarrow \Diamond_a\Box_a\phi$$

where $\Diamond a\phi : \neg\Box a\neg\phi$ is the dual modality to safe belief $\Box a$.

(c) Prove (via a counterexample) that safe belief does NOT imply strong belief; i.e. that

$$\Box_a\varphi \nRightarrow Sb_a\varphi$$

The thought that Albert can be drunk can only be gained by new information, but it's not strongly believed initially:



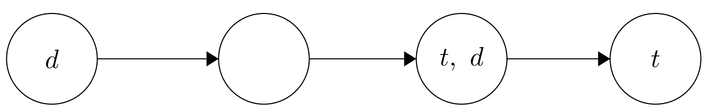
HINT: The positive statements (a) and (b) need general proofs: you need to show that, for every plausibility model and every sentence ϕ, the desired formula is true at all the worlds of that model. But the negative statement (c) must be shown by giving a counterexample: construct some plausibility model and find some sentence φ for which the implication fails to be true at some world of that model (which we can think of as the "real world").

Exercise 2

1. *Write down a logical formula in the language of beliefs, knowledge and conditional beliefs to encode all the above assumptions.*

$$B_a(t \wedge \neg d) \wedge B_a^{(t \wedge d)}(t \wedge d) \wedge B_a^{(\neg t \wedge \neg d)}(\neg t \wedge \neg d) \wedge B_a^{(\neg t \wedge d)}(\neg t \wedge \neg d)$$

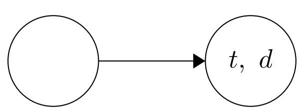
2. *Represent the agent's beliefs (and conditional beliefs), using a plausibility model with four possible worlds. Specify the valuation (which atomic sentences of the two atomic sentences d and t are true at which worlds). Represent the agent's plausibility relation on these worlds, by drawing arrows going from the less plausible worlds to the more plausible ones.*



3. *Suppose somebody who never lies tells our agent "You are close to the target if and only if you believe that you are in a dangerous zone." Write down formally this sentence as a formula φ in doxastic logic (using the atomic sentences).*

   $\varphi := t_a^{(B_a d)} \wedge B_a d^{(t_a)}$

4. *Interpreting the above truthful announcement $!\varphi$ as an update with the sentence φ in the previous part, represent the updated model.*



5. *After the previous announcement, another truthful announcement is made: "You are in a dangerous zone if and only if you don't believe that you are*

in a dangerous zone." Write down formally this sentence as a formula ψ in doxastic logic (using the atomic sentences).

$\psi := d_a^{(B_a \neg d_a)} \wedge B_a \neg d_a^{(d_a)}$

6. *What is the real world? (In other words, answer the question: is the agent in a dangerous zone or not, and is he close to the target or not?) Justify your answer, by interpreting the announcement in the previous part as a new update $\neg\psi$ and representing the updated model.*

   The agent believes that it's most plausible that he's in a dangerous zone, if $\neg\psi$, the agent knows that if he believes this, he's not in a dangerous zone.