Inference over knowledge and beliefs in a modal framework

Fernando R. Velázquez-Quesada

F.R.VelazquezQuesada@uva.nl

1 Introduction

In recent years there has been a *dynamic turn* in logic, bringing into focus the study of the processes that transform information. Several works turned the attention to the effect public announcements have over the knowledge of a set of agents, and the ideas have been extended with the study of more complex actions ([1]) and the effect of them over a more wide propositional attitudes ([5, 2]).

Some other works have focused not in the effect of external interaction, but in the internal dynamics (inferences) that lead our reasoning process. Based on a implicit/explicit information representation, there are temporal approaches, considering how information changes through time ([3]), and there are dynamic ones, naming the actions that change it ([4]). There are also works combining internal and external dynamics, like [6] and [7].

The works on internal dynamics have focused on the effect of deduction over knowledge, but in general we have information that does not have to be true, and it is usually modified via inferences that are not truth-preserving. Our proposal extends previous works by allowing us to describe not only the way deduction affects knowledge, but also to describe the way inference in general affects beliefs. For easiness of reading and writing, here we consider the single-agent case, but the generalization to a multi-agent setting is straightforward.

2 The proposal

Let $\mathtt{P}$ be a set of atomic propositions, and denote the propositional language by $\mathcal{I}$. Our basic semantic model is a Kripke model in which each world, besides having an atomic valuation, assigns three sets to the agent. The first and the second are sets of formulas in $\mathcal{I}$, representing her explicit knowledge ($Y_{\mathcal{K}}$) and her explicit beliefs ($Y_{\mathcal{B}}$) in that world; the third is a set of *rules*[1] based on $\mathcal{I}$ (Z), interpreted as the rules the agent can apply. More formally, an *information model* is a tuple $M = \langle W, R, V, Y_{\mathcal{K}}, Y_{\mathcal{B}}, Z\rangle$, with W the set of worlds, $R \subseteq W \times W$ the accessibility relation, $V : W \to \wp(\mathtt{P})$ the valuation, and $Y_{\mathcal{K}}, Y_{\mathcal{B}} : W \to \wp(\mathcal{I})$, $Z : W \to \wp(\mathtt{R})$ as described. As usual, we assume that knowledge is truth information (if $\gamma \in Y_{\mathcal{K}}(w)$, then γ is true at w) and that the agent believes whatever she knows, but not the other way around ($Y_{\mathcal{K}} \subseteq Y_{\mathcal{B}}$).

We consider a language that extends that of epistemic logic by adding operators to talk about the sets $Y_{\mathcal{K}}$, $Y_{\mathcal{B}}$ and Z:

$$\varphi ::= \top \mid p \mid I_{\mathcal{K}}\,\gamma \mid I_{\mathcal{B}}\,\gamma \mid L\,\rho \mid \neg\varphi \mid \varphi \vee \psi \mid \Diamond\varphi$$

with $p \in \mathtt{P}, \gamma \in \mathcal{I}$ and ρ a rule. Formulas $I_{\mathcal{K}}\,\gamma$, $I_{\mathcal{B}}\,\gamma$ and $L\,\rho$ are read as *"the agent explicitly knows γ"*, *"the agent explicitly believes γ"* and *"the agent can apply rule ρ"*, respectively. The satisfiability relation $\Vdash$ between the pointed model (M, w) and formulas of the language is defined as usual, with the cases of $I_{\mathcal{K}}\,\gamma$, $I_{\mathcal{B}}\,\gamma$ and $L\,\rho$ given by just looking at the sets $Y_{\mathcal{K}}$, $Y_{\mathcal{B}}$ and Z, respectively.

Inference can now be defined as a model-operation that modifies the explicit information sets ($Y_{\mathcal{K}}$ or $Y_{\mathcal{B}}$). For example, deduction over knowledge is defined as an operation that adds the

[1] A rule ρ is given as its finite set of premises $\text{prem}(\rho)$ and its conclusion $\text{conc}(\rho)$; the set of rules is denoted by $\mathtt{R}$.

conclusion of the rule to the explicit knowledge set of those worlds where we have all the premises and the rule. Given a model $M = \langle W, R, V, Y_{\mathcal{K}}, Y_{\mathcal{B}}, Z \rangle$, the one that results from the deductive application of the rule ρ is given by $M_{\mathcal{K},\rho} = \langle W, R, V, Y'_{\mathcal{K}}, Y_{\mathcal{B}}, Z \rangle$, where for every $w \in W$, we have $Y'_{\mathcal{K}}(w) := Y_{\mathcal{K}}(w) \cup \{\text{conc}(\rho)\}$ if $\text{prem}(\rho) \subseteq Y_{\mathcal{K}}(w)$ and $\rho \in Z(w)$, and $Y'_{\mathcal{K}}(w) := Y_{\mathcal{K}}(w)$ otherwise. Our two model assumptions are preserved, as can be easily checked, and the language can be extended by closing it under the operator $\langle \mathrm{D}_\rho \rangle$. Formulas of the form $\langle \mathrm{D}_\rho \rangle\, \varphi$ are read as *"there is a way to apply ρ deductively after which φ is the case"*, and their semantics are given in terms of the operation: $(M, w) \Vdash \langle \mathrm{D}_\rho \rangle\, \varphi$ if and only if $(M, w) \Vdash I_{\mathcal{K}}\ \text{prem}(\rho) \wedge L\, \rho$ and $(M_{\mathcal{K},\rho}, w) \Vdash \varphi$.

But now we can define also deduction over beliefs. Maybe we do not know if γ is really the case, but we can still use deduction to see which new beliefs we can get from it (*"I believe that it is raining, so I believe the streets are wet"*). The correspondent model operation is similar to the previous one, but this time affecting the explicit belief set $Y_{\mathcal{B}}$. Just as in the knowledge-based deduction case, this is a step-by-step approach to reach the deductive closure of a set of formulas. In this case, the original set of formulas can be considered as the *belief base*, a concept commonly used in *belief revision*.

Moreover, now we can define inferences different from deduction. For example, a version of default inference can be defined as an operation that does not need all the premises of the rule in order to be applied, and that adds not only the conclusion but also all the assumed premises. Here it is not that relevant in which set (knowledge or beliefs) we look for the premises because the inference is not truth-preserving, and therefore we cannot add the conclusion of the rule to the explicit knowledge set. Nevertheless, we can add it to the explicit beliefs, providing a way to generate *beliefs* from *knowledge*. Other non-truth-preserving inferences, like abduction or belief revision, can also be explored in this framework.

3 Work to do

The work is in initial stages. For each inference we want to express, we need to define formally the model operation, taking care that all the required properties are preserved. Then we need to verify that the preconditions of the operations can be expressed in the language, in order to be able to present sound and complete axiomatizations of the system. We want to provide formulas describing not only the way explicit/implicit knowledge/beliefs are affected after the different inferences, but also they way the different inferences interact with each other.

References

[1] A. Baltag and L. S. Moss. Logics for epistemic programs. *Synthese*, 139(2):165–224, 2004.

[2] A. Baltag and S. Smets. A qualitative theory of dynamic interactive belief revision. In G. Bonanno, W. van der Hoek, and M. Wooldridge, editors, *Logic and the Foundations of Game and Decision Theory*, volume 3 of *Texts in Logic and Games*, pages 13–60. A.U.P., 2008.

[3] H. N. Duc. Reasoning about rational, but not logically omniscient, agents. *Journal of Logic and Computation*, 7(5):633–648, 1997.

[4] M. Jago. Rule-based and resource-bounded. In T. Agotnes et. al., editors, *Proceedings of the Workshop on Logics for Resource-Bounded Agents, ESSLLI 2006*, pages 63–77, 2006.

[5] J. van Benthem. Dynamic logic for belief revision. *Journal of Applied Non-Classical Logics*, 17(2):129–155, 2007.

[6] J. van Benthem. Merging observational and access dynamics in logic. *Journal of Logic Studies*, 1(1), 2008.

[7] F. R. Velazquez-Quesada. Inference and update. In J. van Benthem and E. Pacuit, editors, *Proceedings of the Workshop on Logic and Intelligent Interaction, ESSLLI 2008*, 2008.