Theorem Proving for Synchronous Systems with No Learning

Cláudia Nalon
Department of Computer Science, University of Liverpool
Liverpool, L69 7ZF, United Kingdom
C.Nalon@csc.liv.ac.uk

February 26, 2001

Combinations of non-classical logics is an area of increasing interest as such logics have been shown to be useful for reasoning about complex situations, such as, for instance, the specification and verification of distributed [3] and multi-agent systems [7]. For instance, modal logics of belief and knowledge have been used to represent the informational component of a system; and dynamic or temporal logics have been used to describe the system's dynamic behaviour. In multi-agent systems, the motivational component, i.e. the intentions or goals of the agents, is often represented by the modal logic KD [7].

Here, we examine a *Temporal Logic of Knowledge*, abbreviated as $\mathsf{KL}_{(n)}$, in which the dynamic and informational components of distributed and multi-agent systems can be represented [3]. Given a logical characterisation of a system, it is then necessary to verify if this specification is correct. In other words, we need to *prove* that the specification matches the desired requirements and properties for the system.

When time and knowledge do not interact, proofs can be obtained essentially by combining existing theorem proving algorithms for temporal logics and logics of knowledge; information between the two dimensions may be passed through the propositional fragment which is shared by both logics. However, for particular applications, it is necessary to restrict the class of systems we might wish to consider. This is the case, for instance, when we are interested in how the knowledge of an agent evolves over time. By adding *interaction* axioms, i.e., axioms that include operators of both logics, some interesting properties can be captured. For instance, by adding the axiom $\mathrm{K}_i \bigcirc \phi \Rightarrow \bigcirc \mathrm{K}_i \phi$ (where K_i can be read as 'agent i knows' and $\bigcirc$ as 'the next moment in time'), synchronous systems with perfect recall (where agents do not forget their history) can be described. Nevertheless, there are few works on proof methods for interacting logics, maybe because of the complexity of the validity problem for such classes of logics. It has been shown [5] that when knowledge and time do not interact, the complexity of the validity problem for the combined logic is the same as the complexity for the logics considered alone, i.e. PSPACE for the single agent case. However, when time and knowledge interact, the complexity of the validity problem increases dramatically. For instance, if the synchrony and perfect recall axiom is added, the complexity for the single agent case is double-exponential time.

Interactions between knowledge and time, their properties, and issues of complexity have been discussed in [3, 8, 4]. We restrict attention to the single agent case of synchronous systems with no learning. A system is said to be *synchronous* if the agent has access to a common external clock. No learning can be seen as the dual of perfect recall: an agent does not learn if, at certain moment in time, she is not able to distinguish between two situations, so she will never able to distinguish between them. Formally, we capture this property by adding the axiom $\vdash \bigcirc \mathrm{K}\, \varphi \Rightarrow \mathrm{K} \bigcirc \varphi$ to $\mathsf{KL}_{(1)}$ [4]. The resulting logic is abbreviated as $\mathsf{KL}^{sync,nl}$, whose syntax and semantics are the same as in $\mathsf{KL}_{(1)}$.

Our proposed proof method for $\mathsf{KL}^{sync,nl}$ is essentially the same as the method for $\mathsf{KL}_{(1)}$ [2]: the approach is clausal; the clausal form separates out the different dimensions for reasoning. Also, as in [1], after translation, further clauses are added: here, for each step clause, that is, a clause in the form $X \Rightarrow \bigcirc Y$, we add the corresponding SNL-clause $\neg\mathrm{K}\neg X \Rightarrow \bigcirc\neg\mathrm{K}\neg Y$. Some additional renaming is necessary in order to retain the normal form. These new clauses ensure that the constraints expressed by the no learning and synchrony axiom are made explicit before applying subsequent resolution rules.

The set of inference rules are the same as in the method for $\mathsf{KL}_{(1)}$, that is, no new inference rules need to be introduced, and so existing theorem provers for temporal logics of knowledge can be easily adapted to deal with this particular interaction. The method has been shown to be sound, complete, and termination is guaranteed [6].

Extending this approach to the multi-agent case is ongoing work. Also, other interactions, such as perfect recall and no learning in *non synchronous* systems, will be investigated as well as corresponding interactions between belief and time. Adding interaction axioms increases (sometimes dramatically) the complexity of the logic. Consequently, the development of strategies to guide the search for a proof is crucial in order to make these techniques useful in practical contexts and this is also part of our future work.

References

[1] C. Dixon and M. Fisher. Clausal Resolution for Logics of Time and Knowledge with Synchrony and Perfect Recall. In *Proceedings of ICTL 2000*, Leipzig, Germany, 2000.

[2] C. Dixon and M. Fisher. Resolution-Based Proof for Multi-Modal Temporal Logics of Knowledge. In S. Goodwin and A. Trudel, editors, *Proceedings of the Seventh International Workshop on Temporal Representation and reasoning (TIME'00)*, pages 69–78, Cape Breton, Nova Scotia, Canada, July 2000. IEEE Computer Society Press.

[3] R. Fagin, J. Y. Halpern, Y. Moses, and M. Y. Vardi. *Reasoning About Knowledge.* MIT Press, 1995.

[4] J. Y. Halpern, R. van der Meyden, and M. Vardi. Complete Axiomatizations for Reasoning about Knowledge and Time. Submitted for publication, 1999.

[5] J. Y. Halpern and M. Y. Vardi. The Complexity of Reasoning about Knowledge and Time: Extended Abstract. In *Proceedings of the Eighteenth Annual ACM Symposium on Theory of Computing*, pages 304–315, Berkeley, California, May 1986.

[6] C. Nalon, C. Dixon, and M. Fisher. Theorem-Proving for Temporal Logics of Knowledge with No Learning and Synchrony. Submitted for publication, February 2001.

[7] A. S. Rao and M. P. Georgeff. Modeling Rational Agents within a BDI-Architecture. In R. Fikes and E. Sandewall, editors, *Proceedings of Knowledge Representation and Reasoning (KR&R-91)*, pages 473–484. Morgan-Kaufmann, April 1991.

[8] R. van der Meyden. Axioms for Knowledge and Time in Distributed Systems with Perfect Recall. In *Logic in Computer Science*, pages 448–457, 1994.