



Knowing How to Play: Uniform Choices in Logics of Agency

Nicolas Troquard
(IRIT, LOA)

In the last years there has been increasing interest in logics enabling reasoning about strategies of agents and coalitions of agents, and the agents' knowledge about such strategies. Such logics combine two kinds of modal logics:

- logics of knowledge such as S5, and multiagent versions thereof; such logics have modal operators K_a , where $K_a \phi$ is read 'agent a knows that ϕ ' ;
- logics of agency, including in particular Coalition Logic (CL) and Alternating-time Temporal Logic (ATL) [1]; such logics have constructions such as CL's $[A] \phi$ or ATL's $\langle\langle A \rangle\rangle X \phi$, both (roughly) reading 'group of agents A has an action to ensure that ϕ holds (whatever the other agents choose to do)'.

While each of these logics is by now well-established, the interaction between knowledge and agency is less consensual. A straightforward combination of for example ATL and epistemic logic (called ATEL) was proposed in [8]. In ATEL one can express things such as 'agent a has an action to ensure that ϕ ', but ignores that ψ . It turned out that ATEL is not sufficient for modeling sentences like 'agent a knows how to ensure ϕ '. The problem can be highlighted by the following example.

Example 1 There is a switch, a lamp, and a blind agent a_1 , which ignores whether the light is on or off. a_1 can toggle the switch (and it knows that), and a_1 can remain passive. Clearly, $\langle\langle a_1 \rangle\rangle X \text{light}$ holds here, i.e., a_1 can ensure that the light is on (viz. by toggling the switch if the light is off, and by doing nothing if the light is already on). We should also be able to conclude that a_1 does not know which action to perform in order to do this.

ATEL makes us conclude here that $K_{a_1} \langle\langle a_1 \rangle\rangle X \text{light}$, i.e. the blind agent a_1 knows that it has an action to ensure the light is on. The problem is that this strategy is what has been called non-uniform: it makes a_1 choose different actions in possible worlds that are indistinguishable for him. Multiagent variants of our example can also be devised.

Several authors have proposed modified versions of ATEL, trying to accommodate in one way or another the notion of uniform strategy [6, 9, 7]. It seems to be fair to say that all these attempts resulted in rather complex formalisms with heavy notations, and that there is no consensus up to now what the appropriate logic of knowledge and strategies is.

We here take as our starting point a slightly different logic of agency that has been developed in philosophical logic. Just as ATL, the logic of 'Seeing To It That' (STIT) [5] is a

modal logic enabling us to speak about time and agents' choices of actions. In STIT, CL's and ATL's \forall forall- \exists exists-quantification (there is a strategy of group A such that for all actions of the other agents) is split up into two different modal operators:

- an operator of historical possibility ;
- an operator of “seeing to it that” $St\hat{t}$.

In previous work [3] we have shown that STIT is at least as expressive as ATL. We have proved this by translating ATL into STIT. The main clauses of the translation map ATL's $\langle\langle A \rangle\rangle\phi$ (group A has a strategy to ensure that ϕ) into STIT's $\diamond St\hat{t}_A \phi$ (it is possible that group A sees to it that ϕ).¹

In this presentation we argue that the STIT framework can easily account for uniform strategies. To support our claim, we first present a straightforward solution in STIT logic augmented by a modal operator of knowledge. Then we offer a simplification, by introducing a modal logic of knowledge-based uniform agency, for choices, alias one-step strategies. Originally presented at AAMAS'06 [4], we shall push the presentation towards the recent perspective of [2].

References

- [1] R. Alur, T. Henzinger, and O. Kupferman. Alternating-time temporal logic. In Proceedings of the 38th IEEE Symposium on Foundations of Computer Science, Florida, October 1997.
- [2] J. Broersen, A. Herzig, and N. Troquard. A STIT-extension of ATL. In Tenth European Conference on Logics in Artificial Intelligence (JELIA'06), Liverpool, England, UK, volume 4160 of Lecture Notes in Artificial Intelligence, pages 69–81. Springer, 2006.
- [3] J. Broersen, A. Herzig, and N. Troquard. Embedding Alternating-time Temporal Logic in Strategic STIT Logic of Agency, 2006. To appear in Journal of Logic and Computation.
- [4] A. Herzig and N. Troquard. Knowing How to Play: Uniform Choices in Logics of Agency . In G. Weiss and P. Stone, editors, 5th International Joint Conference on Autonomous Agents & Multi Agent Systems (AAMAS-06) , Hakodate, Japan, pages 209–216. ACM Press, 2006.
- [5] J. F. Horty. Agency and Deontic Logic. Oxford University Press, Oxford, 2001.
- [6] W. Jamroga and W. van der Hoek. Agents that know how to play. *Fundamenta Informaticae*, 2004.
- [7] W. Jamroga and T. Ågotnes. What agents can achieve under incomplete information. In Proceedings of AAMAS'06, 2006.
- [8] W. van der Hoek and M. Wooldridge. Tractable multiagent planning for epistemic goals. In AAMAS '02: Proceedings of the first international joint conference on Autonomous agents and multiagent systems, pages 1167–1174, New York, NY, USA, 2002. ACM Press.
- [9] S. van Otterloo and G. Jonker. On epistemic temporal strategic logic. *Electronic Notes in Theoretical Computer Science*, 85(2), 2004. In Proceedings of the 2nd Int. Workshop on Logic and Communication in Multiagent Systems (LCMAS 2004).

¹ The STIT operator used here is the strategic STIT.