Proof in the Core of the KARO Agent Theory

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

The use of *agents* is now seen as an essential tool in representing, understanding and implementing complex software systems. In particular, the characterisation of complex components as *intelligent* or *rational* agents allows the system designer to analyse applications at a much higher level of abstraction. The KARO (knowledge, actions, results, opportunities) framework [6] combines actions (via propositional dynamic logic -$\mathcal{PDL}$) with knowledge (the modal logic S5) and wishes (the modal logic KD). For example the KARO formula $[\mathsf{do}_i(\alpha)]\varphi$ means that φ holds in all worlds which result from agent i performing action α and $\langle\mathsf{do}_i(\alpha)\rangle\varphi$ means that there is a world resulting from agent i performing action α such that φ holds. The ability ($\mathbf{A}_i\alpha$) and opportunity ($\mathbf{O}_i\varphi$) operators mean, informally, that agent i has the ability or opportunity to perform action α. The implementability operator $\Diamond\varphi$ means that following a sequence of actions φ holds.

At Liverpool we have been examining approaches to proof in the KARO framework [5] following two main themes:-

- translation of the core of KARO into a fragment of first-order classical logic and first-order classical resolution;
- the representation of the core of KARO using the fusion of CTL and S5 and clausal resolution directly on this combined logic;

where the core of KARO is a restriction of the full KARO framework that is outlined below. Here we give an overview to the latter approach.

To obtain the core of KARO we simplify KARO as follows. The ability ($\mathbf{A}_i\alpha$) and opportunity ($\mathbf{O}_i\varphi$) operators are defined as $\langle\mathsf{do}_i(\alpha)\rangle\top$. We disallow the unbounded repetition operator α^* on actions and we exclude wishes from the language. There is no interaction between the dynamic and modal components. Finally we use a limited version of the implementability operator.

We use the simple observation that the use of $\mathcal{PDL}$ in the KARO framework is very similar to the use of branching-time temporal logic. Thus, we attempt to use a simple CTL branching-time temporal logic to represent the dynamic component of the core KARO logic. CTL was first described in [1] and can be distinguished from the variety of branching time temporal logics proposed in the literature, as every temporal operator, for example '$\bigcirc$' (in the next moment) must be preceded immediately by a path operator, for example $\mathbf{A}$ (on all paths) or $\mathbf{E}$ (on some path). Thus an example of a CTL formula is $\mathbf{A}\bigcirc\varphi$, meaning in all possible futures, in the next moment φ holds.

First, dynamic operators involving non-atomic actions are normalised so that they apply to atomic actions using transformations such as the following:-

$$\langle\mathsf{do}_i(\alpha\vee\beta)\rangle\psi \rightarrow \langle\mathsf{do}_i(\alpha)\rangle\psi \vee \langle\mathsf{do}_i(\beta)\rangle\psi$$

Thus, in the absence of the unbounded repetition operator, this provides an equivalent formula to the original where the dynamic operators only apply to atomic actions. Next dynamic operators and implementability are replaced by CTL formulae by the following rules for example,

$$\begin{array}{rll} \langle \mathtt{do}_i(a)\rangle\varphi & \text{is replaced by} & \mathbf{E}\bigcirc(done_i^a \wedge \varphi), \\ [\mathtt{do}_i(a)]\varphi & \text{is replaced by} & \mathbf{A}\bigcirc(done_i^a \Rightarrow \varphi) \text{ and} \\ \Diamond\varphi & \text{is replaced by} & \mathbf{E}\Diamond\varphi. \end{array}$$

where $done_i^a$ is a propositional variable uniquely associated with agent i and atomic action a and $\mathbf{E}\Diamond\varphi$ is a CTL formula meaning informally φ holds now or on somepath at some future moment. Initially we work in the logic CTL, with multi-modal S5, the semantics of which are given in, for example [4], with no interactions. A clausal resolution system for the fusion of CTL and S5 can be defined similar to that given in [2] for the fusion of CTL and KD45 where the modal component is replaced by the relevant modal resolution rules for S5 (see for example [3]).

We are currently considering extensions to the core. The addition of wishes for example requires the addition of resolution rules for the KD modal operator for wishes. Allowing the unbounded repetition operator on atomic actions may be tackled by a direct translation into CTL using the temporal logic *until* operator as follows

$$\langle \mathtt{do}_i(a^*)\rangle\varphi \text{ is replaced by } \varphi \vee \mathbf{E}\bigcirc(done_i^a\,\mathcal{U}\,\varphi).$$

Similarly a closer definition of implementability to that intended by the KARO developers may be achieved by utilising the *until* operator

$$\Diamond_i\varphi \text{ is replaced by } \varphi \vee \mathbf{E}\bigcirc(actions_i\,\mathcal{U}\,\varphi)$$

where $actions_i \Rightarrow \bigvee_{a\in Act} done_i^a$. Finally, the work we have done considering interactions between propositional linear time temporal logic and the modal logic S5 may be useful in allowing interaction between $\mathcal{PDL}$ and S5 in KARO.

Acknowledgements This is joint work with Ullrich Hustadt, Michael Fisher, Renate A. Schmidt, John-Jules Meyer, and Wiebe van der Hoek (Liverpool, Manchester, Utrecht).

References

[1] E. M. Clarke and E. A. Emerson. Design and Synthesis of Synchronisation Skeletons Using Branching Time Temporal Logic. In D. Kozen, editor, *Proceedings of the Workshop on the Logic of Programs*, volume 131 of *Lecture Notes in Computer Science*, pages 52–71. Springer-Verlag, 1981.

[2] C. Dixon, M. Fisher, and A. Bolotov. Resolution in a Logic of Rational Agency. In *14th European Conference on Artificial Intelligence (ECAI 2000)*, pages 358–362, Berlin, Germany, August 2000. IOS Press.

[3] C. Dixon, M. Fisher, and M. Wooldridge. Resolution for Temporal Logics of Knowledge. *Journal of Logic and Computation*, 8(3):345–372, 1998.

[4] J. Y. Halpern and M. Y. Vardi. The Complexity of Reasoning about Knowledge and Time. I Lower Bounds. *Journal of Computer and System Sciences*, 38:195–237, 1989.

[5] U. Hustadt, C. Dixon, R. A. Schmidt, M. Fisher, J. J. Meyer, and W. van der Hoek. Verification within the KARO Agent Theory. In *Proceedings of the first Goddard Workshop on Formal Approaches to Agent-Based Systems*, Lecture Notes in Computer Science, Goddard Space, Flight Center, Greenbelt, Maryland, USA, April 2000. Springer.

[6] B. van Linder, W. van der Hoek, and J. J. Ch. Meyer. Formalising motivational attitudes of agents: On preferences, goals and commitments. In M. Wooldridge, J. P. Müller, and M. Tambe, editors, *Intelligent Agents II (LNAI 1037)*, pages 17–32. Springer-Verlag, 1996.