On Calculi and Kripke Semantics for Multi-Agent Systems within the KARO Framework

Renate A. Schmidt[1] and Dmitry Tishkovsky[2]

[1] Department of Computer Science, The University of Manchester, UK, rschmidt@cs.man.ac.uk
[2] Department of Computer Science, The University of Manchester, UK, dmitry@cs.man.ac.uk

A multi-agent system is a collection of semi-autonomous processes known as 'agents' which live in dynamic and unpredictable environments. Agents continuously interact which each other and the environment and need to reason about their own and other agent's attitudes. A number of theoretical frameworks specifying agent-based systems have been proposed, for example, the BDI model [1, 2], the powerful KARO framework [3, 4], temporal logics of knowledge and belief [5–7], and, more recently, a modal logic framework of belief and change [8], and an epistemic dynamic logic [9].

The long term goal of our work is to provide practical proof systems and automated reasoning tools for the development of multi-agent systems. In previous papers various inference systems were proposed for a core fragment of the KARO framework [10, 11]. Currently we are looking into different ways of enriching the core KARO logic that would accommodate more of the features of the KARO framework. In particular, here, we propose various extensions with features that allow for the specification of the abilities and commitments of agents.

Formally, the language of the KARO framework is defined over the following primitive types: a countable set of propositional variables, a countable set $\mathbf{AtAc}$ of atomic action variables, and a non-empty recursive enumerable set $\mathbf{Ag}$ of agents. The sets of formulae ($\mathbf{For}$) and actions ($\mathbf{Ac}$) in the language are the smallest sets that satisfy the following conditions.

1. $\mathbf{AtAc} \subseteq \mathbf{Ac}$ and $\mathbf{Var} \cup \{\top, \bot\} \subseteq \mathbf{For}$
2. if $\phi \in \mathbf{For}$ then $\phi? \in \mathbf{Ac}$ and $\neg\phi, \mathbf{K}_i\phi, \mathbf{W}_i\phi \in \mathbf{For}$ $(i \in \mathbf{Ag})$
3. if $\alpha \in \mathbf{Ac}$ then $\alpha^*, \mathbf{c}\alpha, \mathbf{u}\beta \in \mathbf{Ac}$ and $\mathbf{A}_i\alpha, \mathbf{Comm}_i\alpha \in \mathbf{For}$ $(i \in \mathbf{Ag})$
4. if $\alpha, \beta \in \mathbf{Ac}$ then $\alpha \cup \beta$, $\alpha;\beta \in \mathbf{Ac}$
5. if $\phi, \psi \in \mathbf{For}$ then $\phi \vee \psi \in \mathbf{For}$
6. if $\alpha \in \mathbf{Ac}$ and $\phi \in \mathbf{For}$ then $[\alpha]_i\phi \in \mathbf{For}$ $(i \in \mathbf{Ag})$

As usual we use the following abbreviations: $\phi \to \psi$ for $\neg\phi \vee \psi$, $\phi \wedge \psi$ for $\neg(\neg\phi \vee \neg\psi)$, $\langle\alpha\rangle_i\phi$ for $\neg[\alpha]_i\neg\phi$ and $\phi \leftrightarrow \psi$ for $(\phi \to \psi) \wedge (\psi \to \phi)$.

The basic KARO-logic, called $\mathbf{KARO}$, is combination of PDL for $[\alpha]_i$, $\mathbf{S5}$ for $\mathbf{K}_i$, $\mathbf{K}$ for $\mathbf{W}_i$ with the following extra axioms $(i \in \mathbf{Ag})$.

1. $\mathbf{A}_i\phi? \leftrightarrow \phi$
2. $\mathbf{A}_i\alpha \cup \beta \leftrightarrow \mathbf{A}_i\alpha \vee \mathbf{A}_i\beta$
3. $\mathbf{A}_i\alpha;\beta \leftrightarrow \mathbf{A}_i\alpha \wedge [\alpha]_i\mathbf{A}_i\beta$
4. $\mathbf{A}_i\alpha^* \to \mathbf{A}_i\alpha \wedge [\alpha]_i\mathbf{A}_i\alpha^*$
5. $\mathbf{A}_i\alpha \wedge [\alpha^*]_i(\mathbf{A}_i\alpha \to [\alpha]_i\mathbf{A}_i\alpha) \to \mathbf{A}_i\alpha^*$
6. $[\mathbf{c}\alpha]_i\mathbf{Comm}_i\alpha \wedge [\mathbf{u}\alpha]_i\neg\mathbf{Comm}_i\alpha$
7. $\mathbf{Comm}_i\alpha \cup \beta \leftrightarrow \mathbf{Comm}_i\alpha \vee \mathbf{Comm}_i\beta$
8. $\mathbf{Comm}_i\alpha;\beta \leftrightarrow \mathbf{Comm}_i\alpha \wedge [\alpha]_i\mathbf{Comm}_i\beta$
9. $\mathbf{Comm}_i\alpha^* \to \mathbf{Comm}_i\alpha \wedge [\alpha]_i\mathbf{Comm}_i\alpha^*$
10. $\mathbf{Comm}_i\alpha \wedge [\alpha^*]_i(\mathbf{Comm}_i\alpha \to [\alpha]_i\mathbf{Comm}_i\alpha) \to \mathbf{Comm}_i\alpha^*$

We define the following extensions of $\mathbf{KARO}$.

$$\mathbf{KARO}_0 := \mathbf{KARO} + \{\mathbf{Comm}_i\alpha \to \mathbf{K}_i\mathbf{Comm}_i\alpha\}_{i\in\mathbf{Ag}}$$
$$\mathbf{KARO}_1 := \mathbf{KARO} + \{[\alpha]_i\mathbf{K}_i\phi \to \mathbf{K}_i[\alpha]_i\phi\}_{i\in\mathbf{Ag}}$$
$$\mathbf{KARO}_2 := \mathbf{KARO}_0 + \{[\alpha]_i\mathbf{K}_i\phi \to \mathbf{K}_i[\alpha]_i\phi\}_{i\in\mathbf{Ag}}$$

A $\mathbf{KARO}$-*frame* is any tuple

$$F = \langle S, Q_i^a, R_i^{\mathbf{K}}, R_i^{\mathbf{W}}, A_i^a, Q_i^{\mathbf{c}a}, Q^{\mathbf{c}\phi?}, S_i^{\cup}, S_i^{\circ}, S_i^{*}, S_i^{\mathbf{u}}\rangle_{\phi\in\mathbf{For}, a\in\mathbf{AtAc}, i\in\mathbf{Ag}},$$

where S is any non-empty set of states; Q_i^a, $Q_i^{\mathbf{c}a}$, $Q_i^{\mathbf{c}\phi?}$ and $R_i^{\mathbf{W}}$ are binary relations on S; $R_i^{\mathbf{K}}$ is an equivalence relation on S; A_i^a is a subset of S; and $S_i^{\cup}(t,\alpha,\beta)$, $S_i^{\circ}(t,\alpha,\beta)$, $S_i^{\mathbf{u}}(t,\alpha)$ and $S_i^{*}(t,\alpha)$ are non-empty subsets of S, for any t in S and any actions α and β.

We say that a subset S_0 of S is K_i*-closed* if for any s from S_0 and any s' from S we have that if sK_is' then $s' \in S_0$.

Define for a binary relation Q on S the following conditions ($i \in \mathbf{Ag}$):

(F1) if $sR_i^{\mathbf{K}}s'$ and sQt, for some states s, s', t from S, then there exists a state t' in S such that $tR_i^{\mathbf{K}}t'$ and $s'Qt'$,

(F2) if $tR_i^{\mathbf{K}}t'$, for some states t, t' from S, and there exists a state s in S such that sQt, then there exists a state s' in S such that $s'Qt'$.

By definition, F is a $\mathbf{KARO}_1$*-frame* if for each i in $\mathbf{Ag}$, all Q_i^a, $Q_i^{\mathbf{c}a}$ and $Q_i^{\mathbf{c}\phi?}$ satisfy property (F1) and for all actions α and β, the sets $S_i^{\cup}(t,\alpha,\beta)$, $S_i^{\circ}(t,\alpha,\beta)$, $S_i^{\mathbf{u}}(t,\alpha)$ and $S_i^{*}(t,\alpha)$ are K_i-closed. F is a $\mathbf{KARO}_2$*-frame* if F is $\mathbf{KARO}_1$-frame and all $Q_i^{\mathbf{c}a}$ and $Q_i^{\mathbf{c}\phi?}$ satisfy property (F2).

A *model* M is any tuple $\langle F, V\rangle$, where F is a frame and V is a function, mapping each propositional variable p to some subset $V(p)$ of the set of states of F.

Using parallel induction on the length of formulae and on the length of actions it is possible to define a binary relation Q_i^α and a set A_i^α, for any action α and agent i, and a meta-relation $M, s \models \phi$, for any state s in model M.

From the definition, each model M can be viewed as a tuple $\langle S, Q_i, R_i^{\mathbf{K}}, R_i^{\mathbf{W}}, A_i, \models\rangle_{i\in\mathbf{Ag}}$, where S is a non-empty set of states, $\models$ is a truth relation on $S \times \mathbf{For}$, $R_i^{\mathbf{K}}$ is an equivalence relation on S, $R_i^{\mathbf{W}}$ and $Q_i(\alpha)$ are binary relations on S, and $A_i(\alpha) \subseteq S$ for each action α. In addition the following hold.

1. $s \models \mathbf{K}_i\phi$ iff $\forall t(sR_i^K t \Rightarrow t \models \phi)$;
2. $s \models \mathbf{W}_i\phi$ iff $\forall t(sR_i^W t \Rightarrow t \models \phi)$;
3. $s \models [\alpha]_i\phi$ iff $\forall t(sQ_i(\alpha)t \Rightarrow t \models \phi)$;
4. $s \models \mathbf{A}_i\alpha$ iff $s \in A_i(\alpha)$;
5. $s \models \mathbf{Comm}_i\alpha$ iff $\exists w \in S(wQ_i(\mathbf{c}\alpha)s)$;
6. $sQ_i(\mathbf{u}\alpha)t \Rightarrow \forall v\neg(vQ_i(\mathbf{c}\alpha)t)$.

We have proved the soundness for $\mathbf{KARO}$, $\mathbf{KARO}_0$, $\mathbf{KARO}_1$, and $\mathbf{KARO}_2$.

Theorem 1. *(i)* $\mathbf{KARO}$ *is sound w.r.t. all* $\mathbf{KARO}$*-frames, (ii)* $\mathbf{KARO}_0$ *and* $\mathbf{KARO}_2$ *are sound w.r.t. all* $\mathbf{KARO}_2$*-frames, (iii)* $\mathbf{KARO}_1$ *is sound w.r.t. all* $\mathbf{KARO}_1$*-frames.*

We conjecture: (i) $\mathbf{KARO}$, $\mathbf{KARO}_1$ and $\mathbf{KARO}_2$ are complete w.r.t. all $\mathbf{KARO}$-, $\mathbf{KARO}_1$- and $\mathbf{KARO}_2$-frames, respectively. (ii) $\mathbf{KARO}_0$ is not complete w.r.t. any class of $\mathbf{KARO}$-frames. (iii) $\mathbf{KARO}$ is decidable.

References

1. A. S. Rao. Decision procedures for propositional linear-time belief-desire-intention logics. In *Proc. ATAL'95*, volume 1037 of *LNAI*, pages 102–118. Springer, 1996.
2. A. S. Rao and M. P. Georgeff. Modeling rational agents within a BDI-architecture. In *Proc. KR'91*, pages 473–484. Morgan Kaufmann, 1991.
3. B. van Linder, W. van der Hoek, and J.-J. C. Meyer. Formalizing abilities and opportunities of agents. *Fundamenta Informaticae*, 34(1, 2):53–101, 1998.
4. B. van Linder, J.-J. C. Meyer, and W. van der Hoek. Formalising motivational attidutes of agents using the KARO framework. Rep. UU-CS-1997-03, Utrecht Univ., 1997.
5. R. Fagin, J. Y. Halpern, Y. Moses, and M. Y. Vardi. *Reasoning About Knowledge*. MIT Press, 1995.
6. J. Y. Halpern and Y. Moses. A guide to completeness and complexity for modal logics of knowledge and belief. *Artif. Intell.*, 54:319–379, 1992.
7. J. Y. Halpern and M. Y. Vardi. The complexity of reasoning about knowledge and time. I lower bounds. *J. Computer and System Sci.*, 38:195–237, 1989.
8. A. Herzig, J. Lang, D. Longin, and T. Polacsek. A logic for planning under partial observability. In *Proc. AAAI'2000*, August 2000.
9. A. Herzig and D. Longin. Belief dynamics in cooperative dialogues. *J. Semantics*, 17(2), 2001. To appear.
10. U. Hustadt, C. Dixon, R. A. Schmidt, and M. Fisher. Normal forms and proofs in combined modal and temporal logics. In *Proc. FroCoS'2000*, volume 1794 of *LNAI*, pages 73–87. Springer, 2000.
11. U. Hustadt, C. Dixon, R.A. Schmidt, M. Fisher, Meyer J.-J., and W. van der Hoek. Verification within the KARO agent theory. In *Proc. FAABS*. LNCS, Springer, 2000.