Multi-agent Justification Logic

Nadezhda Ruhlyadko and Tatiana Yavorskaya

Moscow State University
Department of Mathematics and Mechanics
Moscow, Russia
tatiana_yavorskaya@hotmail.com

Abstract. We introduce the explicit counterparts and prove realization theorem for bimodal logics, describing communication of two agents in case when knowledge (belief) of both agents is described by one of the logics K, T or $\mathsf{K4}$. The agents may interact by either adopting or verifying each others knowledge; interaction can be either one-way or two-ways.

Justification logic introduced by S. Artemov in [1,2] gives formal description of the relation *"the agent knows (believes in) φ because of justification t"*, incorporated in the language of justification logic by atoms of the form $t{:}F$ where t is a justification term and F is a formula. A justification logic can be considered as the explicit variant of a propositional modal logic; the justification counterparts were found for many modal logics such as K, $K4$, D, T, $S4$ ([2,3]).

Multi-agent versions of justification logics were presented and studied in [5] and [4]. In this case interaction between agents is possible; it is expressed via operations which convert justifications of an agent to justifications of the other one. In [5] the logics are formulated for which the agents' knowledge is described by the modal logic $\mathsf{S4}$ and communication between the agents is either *proof adoption* (the agent B can transform evidences of the agent A into his own ones), or *proof verification* (given the evidence t the agent B can construct the evidence for the fact that the agent A accepts t as the evidence for the fact F). If the basic logic of the two agents is as strong as $\mathsf{S4}$ these two types of connection are equivalent. If the reasoning of the agents is described by a weaker system different types of interaction result in different logics, and this is the case we study here.

In all the logics presented here justifications are represented by terms constructed from justification variables and constants by means of operations on justifications: bynary $\cdot$ and $+$ and unary $!$, $\Uparrow_1^2$, $\Uparrow_2^1$, $!_1^2$ and $!_2^1$. The meaning of operations will become clear from the axioms below. Formulas are built from propositional variables by boolean connectives and the justification operators: if t is a justification term and φ is a formula then $t :_i \varphi$ are formulas for $i = 1, 2$ informally interpreted as "the i-the agent accepts t as an evidence for F."

We assume that both agents reason in accordance with the same logic which is either K, T or $\mathsf{K4}$. If we do not assume interaction between the agents this gives us one of the logics K_2, T_2 and $\mathsf{K4}_2$. By J_2 we denote the justification version of the logic K_2, its axioms are formulas $t :_i (F \to G) \to (s :_i F \to [t \cdot s] :_i G)$

and $s:_iF \vee t:_iF \rightarrow [s+t]:_iF$ for $i=1,2$. The logics JT_2 and $\mathsf{JK4}_2$ are the explicit versions of T_2 and $\mathsf{K4}_2$; they are obtained from J_2 by adding the axioms $t:_iF \rightarrow F$ and $t:_iF \rightarrow !t:_it:_iF$ $(i=1,2)$ respectively. The inference rules for justification logics are Modus Ponens and Axiom Necessitation

$$\vdash c_1:_{i_1}c_2:_{i_2}c_3:_{i_3}\ldots c_k:_{i_k}A,$$

where k is a natural number,$i_j \in \{1,2\}$, c_j are justification constants and A is one of the axioms of the corresponding logic.

We consider two types of interaction between agents:

Knowledge adoption. An agent knows everything that the other one knows; operation $\Uparrow_i^j$ converts proofs of the i-th agent into proofs of j-th one.

$(\mathsf{a}\!\uparrow)$	$\Box_1A \rightarrow \Box_2A$	$(\mathsf{a}\!\downarrow)$	$\Box_2A \rightarrow \Box_1A$
$(\mathsf{ja}\!\uparrow)$	$t:_1F \rightarrow \Uparrow_1^2 t:_2F$	$(\mathsf{ja}\!\downarrow)$	$t:_2F \rightarrow \Uparrow_2^1 t:_1F$

Knowledge verification. An agent knows that the other one knows F in case this is true. The corresponding operations are $!_i^j$.

$(\mathsf{v}\!\uparrow)$	$\Box_1A \rightarrow \Box_2\Box_1A$	$(\mathsf{v}\!\downarrow)$	$\Box_2A \rightarrow \Box_1\Box_2A.$
$(\mathsf{jv}\!\uparrow)$	$t:_1F \rightarrow !_1^2t:_2t:_1F$	$(\mathsf{jv}\!\downarrow)$	$t:_2F \rightarrow !_2^1t:_1t:_2F.$

Below is the list of the modal logics together with their justification counterparts: for each modal logic $\mathsf{L} \in \{\mathsf{K}_2, \mathsf{T}_2, \mathsf{K4}_2\}$ we put

- one-way adoption: $\mathsf{L}^{\mathsf{a}} = \mathsf{L} + (\mathsf{a}\!\uparrow)$, $\mathsf{JL}^{\mathsf{a}} = \mathsf{JL} + (\mathsf{ja}\!\uparrow)$,
- one-way verification: $\mathsf{L}^{\mathsf{v}} = \mathsf{L} + (\mathsf{v}\!\uparrow)$, $\mathsf{JL}^{\mathsf{v}} = \mathsf{JL} + (\mathsf{jv}\!\uparrow)$,
- two-way adoption:
  $\mathsf{L}^{\mathsf{a}\uparrow,\mathsf{a}\downarrow} = \mathsf{L} + (\mathsf{a}\!\uparrow) + (\mathsf{a}\!\downarrow)$, $\mathsf{JL}^{\mathsf{a}\uparrow,\mathsf{a}\downarrow} = \mathsf{JL} + (\mathsf{ja}\!\uparrow) + (\mathsf{ja}\!\downarrow)$
- two-way verification:
  $\mathsf{L}^{\mathsf{v}\uparrow,\mathsf{v}\downarrow} = \mathsf{L} + (\mathsf{v}\!\uparrow) + (\mathsf{v}\!\downarrow)$, $\mathsf{JL}^{\mathsf{v}\uparrow,\mathsf{v}\downarrow} = \mathsf{JL} + (\mathsf{jv}\!\uparrow) + (\mathsf{jv}\!\downarrow)$,
- uni-directed adoption and verificational:
  $\mathsf{L}^{\mathsf{a}\uparrow,\mathsf{v}\uparrow} = \mathsf{L} + (\mathsf{a}\!\uparrow) + (\mathsf{v}\!\uparrow)$, $\mathsf{JL}^{\mathsf{a}\uparrow,\mathsf{v}\uparrow} = \mathsf{JL} + (\mathsf{ja}\!\uparrow) + (\mathsf{jv}\!\uparrow)$,
- oppositely directed adoption and verifitation:
  $\mathsf{L}^{\mathsf{a}\uparrow,\mathsf{v}\downarrow} = \mathsf{L} + (\mathsf{a}\!\uparrow) + (\mathsf{v}\!\downarrow)$, $\mathsf{JL}^{\mathsf{a}\uparrow,\mathsf{v}\downarrow} = \mathsf{JL} + (\mathsf{ja}\!\uparrow) + (\mathsf{jv}\!\downarrow)$.

Theorem 1. *For any bi-modal formula F and each logic from the list above except $\mathsf{K}_2^{\mathsf{a}\uparrow,\mathsf{v}\downarrow}$ and $\mathsf{K}_2^{\mathsf{v}\uparrow,\mathsf{v}\downarrow}$ one has $\mathsf{L} \vdash F$ if and only if there exists an assignment F^r of terms to all boxes in F in such a way that $\mathsf{JL} \vdash F^r$.*

References

1. S. Artemov. Explicit provability and constructive semantics. *Bulletin of Symbolic Logic*, 7(1):1–36, 2001.
2. S. Artemov. The Logic of Justification. *The Review of Symbolic Logic*, 1(4):477–513, 2008.
3. Vladimir N. Brezhnev. On explicit counterparts of modal logics. Technical Report CFIS 200005, Cornell University, 2000
4. Antonis Achilleos, Complexity jumps in multiagent justification logic under interacting justifications. E-print 1402.3066, arXiv.org, February 2014.
5. Tatiana Yavorskaya (Sidon), Interacting explicit evidence systems, Theory of Computing Systems, 43(2):272293, August 2008.