Appendix: Planning with Multi-agent Belief using Justified Perspectives

Guang Hu, Tim Miller, Nir Lipovetzky

School of Computing and Information Systems The University of Melbourne
Parkville, VIC 3010, AUS
ghu1@student.unimelb.edu.au, tmiller@unimelb.edu.au, nir.lipovetzky@unimelb.edu.au

1 Complete semantics for justified belief

Definition 1.1 (Complete semantics)**.** The complete semantics for justified belief is defined as:

(a)	$(M, \vec{s}) \vDash r(\vec{t})$	iff	$\pi(s_n, r(\vec{t})) = true$
(b)	$(M, \vec{s}) \vDash \phi \wedge \psi$	iff	$(M, \vec{s}) \vDash \phi$ and $(M, s) \vDash \psi$
(c)	$(M, \vec{s}) \vDash \neg\varphi$	iff	$(M, \vec{s}) \nvDash \varphi$
(d)	$(M, \vec{s}) \vDash S_i v$	iff	$v \in$ dom$(O_i(s_n))$
(e)	$(M, \vec{s}) \vDash S_i\varphi$	iff	$\forall \vec{g} \in \vec{S}_G, (M, \vec{g}[\langle O_i(s_n)\rangle]) \vDash \varphi$ or $\forall \vec{g} \in \vec{S}_G, (M, \vec{g}[\langle O_i(s_n)\rangle]) \vDash \neg\varphi$
(f)	$(M, \vec{s}) \vDash K_i\varphi$	iff	$(M, \vec{s}) \vDash \varphi \wedge S_i\varphi$
(g)	$(M, \vec{s}) \vDash B_i\varphi$	iff	$\forall \vec{g} \in \vec{S}_G, (M, \vec{g}[f_i(\vec{s})]) \vDash \varphi$

where: $\vec{S}_G \in \vec{S}$ is the set of all possible global states sequences and $\vec{g}[\vec{s}] = g_1[s_1], \ldots, g_n[s_n]$; and, $g[s]$ means function override: $g[s](v) = s(v)$ when $v \in$ dom(s) and $g(v)$ otherwise; and s_n is the final state in sequence $\vec{s}$; that is, $s_n = \vec{s}(|\vec{s}|)$.

2 Ternary semantics for justified belief

Definition 2.1 (Ternary semantics)**.** The ternary semantics are defined using function T, omitting the model M for readability:

(a)	$T[\vec{s}, r(\vec{t})]$	=	1 if $\pi(s_n, r(\vec{t})) = true$; 0 if $\pi(s_n, r(\vec{t})) = false$; $\frac{1}{2}$ otherwise
(b)	$T[\vec{s}, \phi \wedge \psi]$	=	$\min(T[\vec{s}, \phi], T[\vec{s}, \psi])$
(c)	$T[\vec{s}, \neg\varphi]$	=	$1 - T[\vec{s}, \varphi]$
(d)	$T[\vec{s}, S_i v]$	=	$\frac{1}{2}$ if $i \notin$ dom(s_n) or $v \notin$ dom(s_n) 0 if $v \notin$ dom$(O_i(s_n))$ 1 otherwise
(e)	$T[\vec{s}, S_i\varphi]$	=	$\frac{1}{2}$ if $T[\vec{s}, \varphi] = \frac{1}{2}$ or $i \notin$ dom(s_n); 0 if $T[\langle O_i(s_n)\rangle, \varphi] = \frac{1}{2}$; 1 otherwise
(f)	$T[\vec{s}, K_i\varphi]$	=	$T[\vec{s}, \varphi \wedge S_i\varphi]$
(g)	$T[\vec{s}, B_i\varphi]$	=	$T[f_i(\vec{s}), \varphi]$

where s_n is the final state in sequence $\vec{s}$; that is, $s_n = \vec{s}(|\vec{s}|)$.

3 Proof for KD45$_n$

Now, we give the theorem and proof for KD45$_n$ properties.

Theorem 3.1. The following axioms hold, making this a KD45$_n$ logic:

K (Distribution):	$B_i\varphi \wedge B_i(\varphi \rightarrow \psi) \rightarrow B_i\psi$
D (Consistency):	$B_i\varphi \rightarrow \neg B_i\neg\varphi$
4 (Positive Introspection):	$B_i\varphi \rightarrow B_iB_i\varphi$
5 (Negative Introspection):	$\neg B_i\varphi \rightarrow B_i\neg B_i\varphi$

Proof. Based on the definition of B_i, $M, \vec{s} \vDash B_i\varphi$ is equivalent to $M, f_i(\vec{s}) \vDash \varphi$. From this, axiom K is: $M, f_i(\vec{s}) \vDash \varphi$ and $M, f_i(\vec{s}) \vDash (\varphi \rightarrow \psi)$ imply $M, f_i(\vec{s}) \vDash \psi$, which holds trivially. For the axiom D, when $M, f_i(\vec{s})\varphi$ holds, then it must be that $M, f_i(\vec{s})\neg\varphi$ does not hold.

Axioms 4 and 5 are more involved. The value of a variable from $f_i(\vec{s})$ depends on two values: lt and $\vec{s}$, which are, respectively, the last time the variable was seen by agent i and the input perspectives. The lt depends only on the function O_i. Since $O_i(s) = O_i(O_i(s))$, the lt of $f_i(\vec{s})$ and $f_i(f_i(\vec{s}))$ for each state and each variable are the same.

Now, note that the retrieval function R returns the value $v = e$ if v is in the state s_{lt} (the first line of R), the value of each variable in $f_i(s)$ is the same as its in $f_i(f_i(s))$. Therefore, $f_i(s) = f_i(f_i(s))$.

Given that axiom 4 is equivalent to $M, f_i(\vec{s}) \vDash \varphi$ implies $M, f_i(f_i(\vec{s})) \vDash \varphi$, this holds trivially.

For axiom 5, $M, f_i(s) \nvDash \varphi$ is equivalent to $M, f_i(f_i(s)) \nvDash \varphi$. Based on the definition of B_i, $M, f_i(f_i(s)) \nvDash \varphi$ gives $M, f_i(s) \nvDash B_i\varphi$. Then, based on the definition of $\neg$, we have that $M, f_i(s) \nvDash B_i\varphi$ is equivalent to $M, f_i(s) \vDash \neg B_i\varphi$. Therefore, axiom 5 holds. □

4 Experiment Results for Coin example

	Parameters				Performance					
	$\|Agt\|$	d	$\|\mathcal{G}\|$	$\|P\|$	$\|Gen\|$	$\|Exp\|$	$\|Calls\|$	TIME(s) $calls$	TIME(s) Total	Goal
C01	2	1	1	1	3	2	2	0.0	0.0	$B_a coin = head$
C02	2	1	1	2	7	18	7	0.0	0.0	$B_a coin = tail$
C03	2	1	1	2	5	12	7	0.0	0.0	$B_a coin = head \wedge B_b coin = head$
C04	2	2	1	4	43	126	61	0.0	0.0	$B_a coin = head \wedge B_b coin = tail$
C05	2	3	2	4	46	135	58	0.0	0.0	$B_a coin = head \wedge B_b coin = tail \wedge B_b\ B_a coin = head$
C06	2	2	2	6	414	1239	533	0.5	0.7	$B_a\ B_b coin = tail \wedge B_b B_a coin = head$

Table 1: Experimental results for coin domain