エージェントの相互信念を扱う拡張BDI logicの演繹体系

新出尚之[†] 高田司郎[‡] 櫟粛之[⋆]

BDI logic は、合理的エージェントの記述に用いられる時相論理体系で、CTL*にエージェントの心的状態を記述する様相オペレータを導入し、述語論理に拡張したものである。我々は、CTL ベースの命題論理に限定した BDI logic に対し、sequent calculus による完全な演繹体系を与えた。しかし、もともとの BDI logic は、単独のエージェントの心的状態に対応する様相オペレータしか持たず、マルチエージェント環境でのエージェントの相互心的環境の記述はできない。本論文では、我々の演繹体系に対し、マルチエージェント環境におけるエージェント毎の心的状態様相オペレータ、および相互心的状態の記述を許すように拡張したものを与え、またその体系によるいくつかの証明の例を示す。

A Deduction System of Extended BDI Logic to Handle Mutual Belief

NIDE Naoyuki[†] Shiro TAKATA[‡] Tadashi ARARAGI[⋆]

BDI logics are extensions of the predicate variant of CTL* to represent the behavior of rational agents by introducing mental state operators. We previously presented deduction systems for CTL-based propositional BDI logics using sequent calculus. However, since the original BDI logics have only operators to represent the mental states of individual agents, they cannot handle mutual mental states for multiple agents. In this paper, we extend our deduction systems so that we can handle individual and mutual mental states for multiple agents in multi-agent environments. We also show some examples of proof in our system.

1 はじめに

BDI logic [3] は、合理的エージェントの仕様記述や、合理的エージェントの実装手法である BDI アーキテクチャの基礎概念として用いられる時相論理体系であり、CTL* [1] を述語論理に拡張して、信念·願望·意図を表す様相オペレータを追加したものである。

その演繹体系としては、CTL ベースの命題論理に限定した BDI logic に対して、これまでに Rao らが与えた Hilbert style のもの [4] や、著者らが与えた sequent calculus によるもの [7] があった。しかし、それらは単一エージェントの心的状態を表すオペレータしか持たず、マルチエージェント環境での相互心的状態 (mutual mental state) などを扱えない。複数のエージェントの心的状態オペレータを扱えるような拡張には [5] などがあるが、演繹体系は扱われていない。

本論文では、我々の演繹体系を拡張して、マルチエージェント環境における複数エージェントのそれぞれの心的状態や、相互心的状態 (mutual mental state) を表すオペレータを導入したものを与え [6]、また、この体系を用いて、エージェント間の相互信念や協調などに関わるいくつかの性質の証明の例を示す。

2 論理式の定義

[†]奈良女子大学, Nara Women's Univ.
[‡]近畿大学, Kinki Univ.
[⋆]NTT CS 研, NTT Communication Science Lab.

エージェントの集合 Gr をあらかじめ決めておく。論理式の定義は、CTL [1] でのそれ (ただし、$\vee$, $\neg$, AX, AU, EU 以外のオペレータは略記として導入する) に、以下の拡張を行ったものである[1]。

- ϕ が論理式、かつ $i \in Gr$, $g \subset Gr$ ならば、$\mathsf{BEL}_i(\phi)$, $\mathsf{DESIRE}_i(\phi)$, $\mathsf{INTEND}_i(\phi)$, $\mathsf{G\text{-}BEL}_g(\phi)$, $\mathsf{G\text{-}DESIRE}_g(\phi)$, $\mathsf{G\text{-}INTEND}_g(\phi)$ もそれぞれ論理式

この他、$\mathsf{E\text{-}BEL}_g(\phi)$ は $\bigwedge_{i\in g} \mathsf{BEL}_i(\phi)$ の、$\mathsf{M\text{-}BEL}_g(\phi)$ は $\mathsf{E\text{-}BEL}_g(\mathsf{G\text{-}BEL}_g(\phi))$ の略記 (DESIRE, INTEND に対しても同様) とし、また、演算子の結合順位は一般的な慣習通りとする。本体系ではエージェントの限量子は導入していない。

直感的には $\mathsf{M\text{-}BEL}_g(\phi)$ が「エージェント群 g が相互信念 ϕ を持つ」を表す。

3 意味論

3.1 $\mathcal{MA}(\mathrm{B}^{KD45}\mathrm{D}^{KD}\mathrm{I}^{KD})_{CTL}$-structure

可能世界の集合 $W(\neq \emptyset)$ を定め、各世界 $w \in W$ に対して、state の集合 $St_w(\neq \emptyset)$ およびその上の serial な関係 $R_w \subset St_w \times St_w$ を定める。直感的には R_w が時間の遷移関係である。また、各世界 $w \in W$ と state $t \in St_w$ の組に対し、命題記号への真偽値割り当て $L(w,t) \subseteq P$ を定める。

[1]推論規則を見通し良くするため、[6] とは少し異なっているが、本質的な変化ではない。なお、述語論理への拡張はまだ行っていない。

さらに、Gr の各要素 i に対し、$W, \bigcup_{w\in W} St_w, W$ 上の serial な 3 項関係 $\mathcal{B}_i \subset W \times \bigcup_{w\in W} St_w \times W$ で、条件「$(w,t,w') \in \mathcal{B}_i$ かつ $t \in St_w$ ならば $t \in St_{w'}$」を満たすものを定め、同様に $\mathcal{D}_i, \mathcal{I}_i$ も定める。これらは (時刻毎に定まる) 世界間の可視関係で、直感的には信念·願望·意図を表す。なお、$\mathcal{B}_i$ については transitive および Euclidean であることも要請する。

以上の組を $\mathcal{MA}(\mathrm{B}^{KD45}\mathrm{D}^{KD}\mathrm{I}^{KD})_{CTL}$-*structure* (Multi-Agent BDI structure) と呼ぶ。

3.2 BDI logic の論理式の解釈と恒真性

$\mathcal{MA}(\mathrm{B}^{KD45}\mathrm{D}^{KD}\mathrm{I}^{KD})_{CTL}$-structure M とその世界 w、state t に対し、論理式 ϕ が真である場合、$(M,w,t) \models \phi$ と書く。論理式の真偽値は以下のように定める。紙面の制約上、古典および CTL のオペレータについては略する。

$(w,t,w') \in \mathcal{B}_i$ なる全ての w' に対し $(M,w',t) \models \phi$ であるとき、$(M,w,t) \models \mathsf{BEL}_i(\phi)$ とする。また、$w = w_0$, $(w_0,t,w_1) \in \mathcal{B}_{i_0}$, $(w_1,t,w_2) \in \mathcal{B}_{i_1}$, $\cdots$(ただし $i_0, i_1, \cdots \in g$) を満たす任意の $w_0, w_1, w_2, \cdots$ と $n \geq 0$ に対し $(M,w_n,t) \models \phi$ であるとき、$(M,w,t) \models \mathsf{G\text{-}BEL}_g(\phi)$ とする。DESIRE, INTEND についても同様に定める。

4 演繹体系

本章では演繹体系 $\mathcal{MA}(\mathrm{B}^{KD45}\mathrm{D}^{KD}\mathrm{I}^{KD})_{CTL}$ を導入する。本体系では、sequent の $\rightarrow$ の両側は論理式の multi set である。以下、ギリシア大文字は論理式の multi set、ギリシア小文字は単一の論理式とし、また、論理式の multi set Γ と単項論理演算子 K に対し、$K(\Gamma)$ は Γ の各論理式に K を付加して得られる multi set を表す。

4.1 推論規則

推論規則は図 1 に挙げられたものである。この体系には cut はない。なお、簡単のため、本論文では [2] で導入した心的状態の整合性公理は扱わない。

名前に「左」「右」がつく規則は、逆向き (結論から前提へ) に見ると、$\rightarrow$ のそれぞれ左か右のどれかの論理式 ϕ のトップレベルのオペレータを除去する役割を持つ。以下これを、その規則の ϕ **に対する適用**と呼ぶ。

4.2 証明可能性の定義

S の推論図を、sequent をノード、S を根とし、推論規則を枝とする木と見て、これを**推論木**と呼ぶ。S の推論木のある葉ノード L が、その祖先 $N(\neq L)$ と同じ sequent であり、かつ、それと同じ sequent が N から L までの経路の途中にはないとき、この経路を**導出のループ**と呼ぶ。[7] と同様、本体系では一定条件下 (後述) で、導出のループを持つような証明を認める。

以下、ある論理式の様相オペレータのうち、他の様相オペレータの scope 内にないものを「最外様相オペレータ」と呼ぶ。

定義 4.1 導出のループ $N_0, N_1, \cdots, N_n$ (N_n が葉) が、ある論理式 $\rho = \mathsf{A}(\phi\ \mathsf{U}\ \psi)$ あるいは $\rho = \mathsf{E}(\phi\ \mathsf{U}\ \psi)$ に対して以下の条件を全て満たすとき、これを、**終局性論理式** ρ **を持つ導出のループ**と呼ぶ。

A1. N_0 から N_n までの過程で 1 回以上、ρ に対する AU 左か EU 左の適用がある。

A2. N_i から N_{i+1} までの過程が、ρ に対する AU 左か EU 左の適用ならば、N_{i+1} はその前提のうち右の sequent である。

A3. N_0 から N_n までの過程のどこかで、Weak によって、最外様相オペレータが時相オペレータ (AX, AU, EU) であるような論理式が消されることはない (ただし、$\rightarrow$ の左あるいは右に、そのような論理式で n 個の同一のものがあるとき、これを m $(n > m > 0)$ 個に減らすことは構わない)。

また、そのループが、ある論理式 $\rho = \mathsf{G\text{-}BEL}_g(\phi)$ に対して以下の条件を全て満たすときにも、これを、終局性論理式 ρ を持つ導出のループと呼ぶ。

B1. N_0 から N_n までの過程で 1 回以上、ρ に対する G-BEL 右の適用がある。

B2. N_i から N_{i+1} までの過程が、ρ に対する G-BEL 右の適用ならば、N_{i+1} はその前提の一番左ではない sequent である。

B3. BEL_i に対する BEL–KD45 の適用の直前に Weak で BEL_j $(i \neq j)$ を消去する場合を除き、N_0 から N_n までの過程のどこかで、Weak によって、最外様相オペレータが信念オペレータ (BEL や G-BEL) であるような論理式が消されることはない (但し書きは A3 と同じ)。

さらに、$\rho = \mathsf{G\text{-}DESIRE}_g(\phi)$ または $\rho = \mathsf{G\text{-}INTEND}_g(\phi)$ に対して同様の条件を満たすときも、同様に呼ぶ。 □

定義 4.2 sequent S が**証明可能**とは、以下の条件を全て満たす S の推論木 T が存在することである。

1. T の全ての葉は、initial sequent をラベルとして持つか、あるいは導出のループの終点である。
2. T の全ての導出のループを合わせたグラフ (導出のループに含まれるノードと枝全てからなるグラフ)

$$\frac{\Gamma \to \phi, \Delta}{\neg\phi, \Gamma \to \Delta}\ (\neg\text{左}) \qquad \frac{\phi, \Gamma \to \Delta \quad \psi, \Gamma \to \Delta}{\phi \lor \psi, \Gamma \to \Delta}\ (\lor\text{左}) \qquad \frac{\phi, \Gamma \to \Delta}{\Gamma \to \neg\phi, \Delta}\ (\neg\text{右}) \qquad \frac{\Gamma \to \phi, \psi, \Delta}{\Gamma \to \phi \lor \psi, \Delta}\ (\lor\text{右}) \qquad \frac{\Gamma \to \Delta}{\Gamma, \Gamma' \to \Delta, \Delta'}\ (\mathrm{Weak})$$

$$\frac{\psi, \Gamma \to \Delta \quad \phi, \mathsf{AX}\,\mathsf{A}(\phi\ \mathsf{U}\ \psi), \Gamma \to \Delta}{\mathsf{A}(\phi\ \mathsf{U}\ \psi), \Gamma \to \Delta}\ (\mathsf{AU}\ \text{左}) \qquad \frac{\Gamma \to \psi, \phi, \Delta \quad \Gamma \to \psi, \mathsf{AX}\,\mathsf{A}(\phi\ \mathsf{U}\ \psi), \Delta}{\Gamma \to \mathsf{A}(\phi\ \mathsf{U}\ \psi), \Delta}\ (\mathsf{AU}\ \text{右})$$

$$\frac{\psi, \Gamma \to \Delta \quad \phi, \mathsf{EX}\,\mathsf{E}(\phi\ \mathsf{U}\ \psi), \Gamma \to \Delta}{\mathsf{E}(\phi\ \mathsf{U}\ \psi), \Gamma \to \Delta}\ (\mathsf{EU}\ \text{左}) \qquad \frac{\Gamma \to \psi, \phi, \Delta \quad \Gamma \to \psi, \mathsf{EX}\,\mathsf{E}(\phi\ \mathsf{U}\ \psi), \Delta}{\Gamma \to \mathsf{E}(\phi\ \mathsf{U}\ \psi), \Delta}\ (\mathsf{EU}\ \text{右})$$

$$\frac{\Gamma \to \Theta}{\mathsf{AX}(\Gamma) \to \mathsf{AX}(\Theta)}\ (\mathsf{AX}\text{–KD}) \qquad \frac{\Gamma \to \Theta}{\mathsf{DESIRE}_i(\Gamma) \to \mathsf{DESIRE}_i(\Theta)}\ (\mathsf{DESIRE}\text{–KD}) \qquad \frac{\Gamma \to \Theta}{\mathsf{INTEND}_i(\Gamma) \to \mathsf{INTEND}_i(\Theta)}\ (\mathsf{INTEND}\text{–KD})$$

$$\frac{\Gamma, \mathsf{BEL}_i(\Gamma) \to \mathsf{BEL}_i(\Delta), \Theta, \mathsf{BEL}_i(\Theta)}{\mathsf{BEL}_i(\Gamma) \to \mathsf{BEL}_i(\Delta), \mathsf{BEL}_i(\Theta)}\ (\mathsf{BEL}\text{–KD45}) \qquad \frac{\phi, \mathsf{BEL}_{i_1}(\mathsf{G\text{-}BEL}_g(\phi)), \cdots, \mathsf{BEL}_{i_n}(\mathsf{G\text{-}BEL}_g(\phi)), \Gamma \to \Delta}{\mathsf{G\text{-}BEL}_g(\phi), \Gamma \to \Delta}\ (\mathsf{G\text{-}BEL}\ \text{左})$$

$$\frac{\Gamma \to \phi, \Delta \quad \Gamma \to \mathsf{BEL}_{i_1}(\mathsf{G\text{-}BEL}_g(\phi)), \Delta \quad \cdots \quad \Gamma \to \mathsf{BEL}_{i_n}(\mathsf{G\text{-}BEL}_g(\phi)), \Delta}{\Gamma \to \mathsf{G\text{-}BEL}_g(\phi), \Delta}\ (\mathsf{G\text{-}BEL}\ \text{右})$$

$$\frac{\phi, \mathsf{DESIRE}_{i_1}(\mathsf{G\text{-}DESIRE}_g(\phi)), \cdots, \mathsf{DESIRE}_{i_n}(\mathsf{G\text{-}DESIRE}_g(\phi)), \Gamma \to \Delta}{\mathsf{G\text{-}DESIRE}_g(\phi), \Gamma \to \Delta}\ (\mathsf{G\text{-}DESIRE}\ \text{左})$$

$$\frac{\Gamma \to \phi, \Delta \quad \Gamma \to \mathsf{DESIRE}_{i_1}(\mathsf{G\text{-}DESIRE}_g(\phi)), \Delta \quad \cdots \quad \Gamma \to \mathsf{DESIRE}_{i_n}(\mathsf{G\text{-}DESIRE}_g(\phi)), \Delta}{\Gamma \to \mathsf{G\text{-}DESIRE}_g(\phi), \Delta}\ (\mathsf{G\text{-}DESIRE}\ \text{右})$$

$$\frac{\phi, \mathsf{INTEND}_{i_1}(\mathsf{G\text{-}INTEND}_g(\phi)), \cdots, \mathsf{INTEND}_{i_n}(\mathsf{G\text{-}INTEND}_g(\phi)), \Gamma \to \Delta}{\mathsf{G\text{-}INTEND}_g(\phi), \Gamma \to \Delta}\ (\mathsf{G\text{-}INTEND}\ \text{左})$$

$$\frac{\Gamma \to \phi, \Delta \quad \Gamma \to \mathsf{INTEND}_{i_1}(\mathsf{G\text{-}INTEND}_g(\phi)), \Delta \quad \cdots \quad \Gamma \to \mathsf{INTEND}_{i_n}(\mathsf{G\text{-}INTEND}_g(\phi)), \Delta}{\Gamma \to \mathsf{G\text{-}INTEND}_g(\phi), \Delta}\ (\mathsf{G\text{-}INTEND}\ \text{右})$$

(ただし、Θ は高々1つの論理式、$i \in Gr$, $g = \{i_1, \cdots, i_n\} \subset Gr$ とする)

図 1: $\mathcal{MA}(\mathrm{B}^{KD45}\mathrm{D}^{KD}\mathrm{I}^{KD})_{CTL}$ の推論規則

の各連結成分 T'(これは T の部分木になる) に対し、T' に含まれる T の導出のループは全て、共通の終局性論理式 ρ を持つ。 □

また、$\to \phi$ という sequent が証明可能であるとき、論理式 ϕ が証明可能であるという。

本体系の健全·完全性を示すには、[7] と本質的に同じ手法を用いる。また、Wang のアルゴリズムを拡張した決定アルゴリズムも存在する。それらについては省略する。

5 例

5.1 協調行為

[5] では、本論文の M-BEL などと同じものを用いて、エージェント間の協調行為に関する様々な概念を記述している。例えば、エージェント群 g が ϕ に対して blind commitment を持つことは

$$\begin{array}{l} \bigwedge_{i \in g}(\neg\,\mathsf{BEL}_i(\phi) \land \\ \quad \mathsf{A}(\mathsf{INTEND}_i(\phi) \land (\mathsf{BEL}_i(\phi) \supset \mathsf{A}(\mathsf{INTEND}_i(\chi_1)\ \mathsf{U}\ \chi_1))\ \mathsf{U}\ \chi_1)) \\ \qquad \text{ただし } \chi_1 = \mathsf{M\text{-}BEL}_g(\phi) \end{array}$$

と書かれる。この式を ψ とし、$i_1, i_2 \in g$ とするとき、$\psi \supset \mathsf{A}(\mathsf{INTEND}_{i_1}(\phi)\ \mathsf{U}\ \mathsf{BEL}_{i_2}(\phi))$ は本体系で証明できる論理式の例である (証明図は省く)。

もう 1 つの例として、joint ability [5] を取り上げる。ここでは [5] の「($\mathsf{Agts}\ \alpha\ g$)」(g は action α を実行する agent group である)、「($\mathsf{Achvs}\ \alpha\ \phi$)」($\alpha$ は ϕ に適したプランである),「($\mathsf{J\text{-}Can}^k\ g\ \phi$)」($k$ 階の joint ability) をそれぞれ $\mathsf{Agts}(\alpha, g)$, $\mathsf{Achvs}(\alpha, \phi)$, $\mathsf{J\text{-}Can}_g^k(\phi)$ と書く。

現在のところ我々の体系は、action expression を用いて action を構成できない。そこで簡単化のため、エージェントが実行可能なプランは $\alpha_1, \cdots, \alpha_\ell$ と列挙でき、かつ各プラン α_i は atomic action $a_1^i, \cdots, a_{n_i}^i$ を順に実行するものに限るとする。また、各 atomic action a を、その action の実行直後に成立する原子命題と同一視し、さらに各 $\mathsf{Agts}(\alpha, g)$ を原子命題として扱う。すると、$\mathsf{Achvs}(\alpha, \phi)$ を次の論理式の略記と捉えることができる。

$$\begin{array}{l} \mathsf{EX}\,\mathsf{EF}(a_1^i \land \mathsf{EX}\,\mathsf{EF}(a_2^i \land \cdots (\cdots \land \mathsf{EX}\,\mathsf{EF}(a_{n_i}^i \land \mathsf{EF}\,\phi) \cdots) \cdots)) \land \\ \mathsf{AX}\,\mathsf{AG}(a_1^i \supset \mathsf{AX}\,\mathsf{AG}(a_2^i \supset \cdots (\cdots \supset \mathsf{AX}\,\mathsf{AG}(a_{n_i}^i \supset \mathsf{AF}\,\phi) \cdots) \cdots)) \end{array}$$

また、$\mathsf{J\text{-}Can}_g^0(\phi)$ (0 階の joint ability; agent group g は何らかの action によって ϕ を直接達成できる) を $\bigvee_{i=1}^{\ell} \mathsf{G\text{-}BEL}_g(\mathsf{Agts}(\alpha_i, g) \land \mathsf{Achvs}(\alpha_i, \phi))$ で定義でき、$k > 0$ に対し $\mathsf{J\text{-}Can}_g^k(\phi)$ を $\mathsf{J\text{-}Can}_g^{k-1}(\mathsf{J\text{-}Can}_g^0(\phi))$ で定義できる。

このとき、例えば $\mathsf{J\text{-}Can}_g^k(\phi) \supset \mathsf{G\text{-}BEL}_g(\mathsf{EF}\,\phi)$ を図 2 のように証明できる (一部略した)。

5.2 相互信念

n 人の子供 ($c_1, \cdots, c_n$ とする) と 1 人の父に関する muddy children 問題を取り上げる。子供は全員赤か青の帽子をかぶっていることと、うち最低 1 人は赤帽であ

$$
\begin{array}{cl}
\vdots & \\ \hline
\bigvee_{j=1}^{4} m_j, \neg m_2, \neg m_3, \neg m_4 \to m_1 & \\ \hline
\phi, \neg m_2, \neg m_3, \neg m_4 \to m_1 & \text{(G-BEL–L)} \\ \hline
\mathsf{BEL}_{c_1}(\phi), \mathsf{BEL}_{c_1}(\neg m_2), \mathsf{BEL}_{c_1}(\neg m_3), \mathsf{BEL}_{c_1}(\neg m_4) \to \mathsf{BEL}_{c_1}(m_1) & \text{(BEL–KD45)} \\ \hline
\mathsf{BEL}_{c_1}(\phi), \xi_{1,2}, \xi_{1,3}, \xi_{1,4}, \neg m_3, \neg m_4, \neg \mathsf{BEL}_{c_1}(m_1) \to m_2 & \text{(Classical)} \\ \hline
\phi, \Gamma, \neg m_3, \neg m_4, \psi_1 \to m_2 & \text{(multiple G-BEL–L)} \\ \hline
\mathsf{BEL}_{c_2}(\phi), \mathsf{BEL}_{c_2}(\Gamma), \mathsf{BEL}_{c_2}(\neg m_3), \mathsf{BEL}_{c_2}(\neg m_4), \mathsf{BEL}_{c_2}(\psi_1) \to \mathsf{BEL}_{c_2}(m_2) & \text{(BEL–KD45)} \\ \hline
\mathsf{BEL}_{c_2}(\phi), \mathsf{BEL}_{c_2}(\Gamma), \xi_{2,3}, \xi_{2,4}, \neg m_4, \mathsf{BEL}_{c_2}(\psi_1), \neg \mathsf{BEL}_{c_2}(m_3) \to m_3 & \text{(Classical)} \\ \hline
\phi, \Gamma, \neg m_4, \psi_1, \psi_2 \to m_3 & \text{(multiple G-BEL–L)} \\ \hline
\mathsf{BEL}_{c_3}(\phi), \mathsf{BEL}_{c_3}(\Gamma), \mathsf{BEL}_{c_3}(\neg m_4), \mathsf{BEL}_{c_3}(\psi_1), \mathsf{BEL}_{c_3}(\psi_2) \to \mathsf{BEL}_{c_3}(m_3) & \text{(BEL–KD45)} \\ \hline
\mathsf{BEL}_{c_3}(\phi), \mathsf{BEL}_{c_3}(\Gamma), \xi_{3,4}, \mathsf{BEL}_{c_3}(\psi_1), \mathsf{BEL}_{c_3}(\psi_2), \neg \mathsf{BEL}_{c_3}(m_3) \to m_4 & \text{(Classical)} \\ \hline
\phi, \Gamma, \psi_1, \psi_2, \psi_3 \to m_4 & \text{(multiple G-BEL–L)} \\ \hline
\mathsf{BEL}_{c_4}(\phi), \mathsf{BEL}_{c_4}(\Gamma), \mathsf{BEL}_{c_4}(\psi_1), \mathsf{BEL}_{c_4}(\psi_2), \mathsf{BEL}_{c_4}(\psi_3) \to \mathsf{BEL}_{c_4}(m_4) & \text{(BEL–KD45)} \\ \hline
\phi, \Gamma, \psi_1, \psi_2, \psi_3 \to \mathsf{BEL}_{c_4}(m_4) & \text{(multiple G-BEL–L)} \\ \hline
\to \phi \wedge \bigwedge \Gamma \wedge \bigwedge_{j=1}^{3} \psi_j \supset \mathsf{BEL}_{c_4}(m_4) & \text{(Classical)}
\end{array}
$$

(ただし $\phi \equiv \mathsf{M\text{-}BEL}_g(\bigvee_{j=1}^{4} m_j)$, $\psi_j \equiv \mathsf{M\text{-}BEL}_g(\neg \mathsf{BEL}_{c_j}(m_j))$, $\xi_{j,j'} \equiv \mathsf{BEL}_{c_j}(\neg m'_j) \Leftrightarrow \neg m'_j$ および $\Gamma \equiv \{\mathsf{M\text{-}BEL}_g(\xi_{j,j'}) \mid 1 \leq j \leq 4, 1 \leq j' \leq 4, j \neq j'\}$)

図 3: muddy children 問題の証明例

$$
\begin{array}{cl}
\vdots & \\ \hline
\mathsf{J\text{-}Can}_g^0(\phi) \supset \mathsf{EF}\,\phi & \\ \hline
\vdots & \\ \hline
\mathsf{J\text{-}Can}_g^{k-2}(\phi) \supset \mathsf{EF}\,\phi & \\ \hline
\vdots & \\ \hline
\mathsf{J\text{-}Can}_g^{k-1}(\phi) \supset \mathsf{EF}\,\phi & \\ \hline
\vdots & \\ \hline
\mathsf{Achvs}(\alpha_i, \mathsf{J\text{-}Can}_g^{k-1}(\phi)) \to \mathsf{EF}\,\phi & \\
\uparrow \qquad \dfrac{\vdots}{\mathsf{BEL}_{i_1}(\mathsf{G\text{-}BEL}_g(\xi_1)) \to \mathsf{BEL}_{i_1}(\psi)} \quad \cdots & \\ \hline
\mathsf{G\text{-}BEL}_g(\xi_1) \to \psi & \text{(G-BEL–L/R, Weak)} \\
\uparrow \qquad \cdots & \\ \hline
\to \mathsf{J\text{-}Can}_g^{k}(\phi) \supset \psi & \text{(Classical)}
\end{array}
$$

(ただし $\psi \equiv \mathsf{G\text{-}BEL}_g(\mathsf{EF}\,\phi)$, $\xi_i \equiv \mathsf{Agts}(\alpha_i, g) \wedge \mathsf{Achvs}(\alpha_i, \mathsf{J\text{-}Can}_g^{k-1}(\phi)))$

図 2: joint ability に関する証明の例

ることは全員が知っており、どの子も他の子の帽子の色は見える。全員十分賢くかつ正直であり、問答は全員に聞こえる。また、子は全員実は赤帽とする。

父は $c_1, c_2, \cdots$ に順に自分の帽子の色を尋ねる。c_{n-1} までは「わからない」と答えるが、それを聞いた c_n は「私は赤帽」と答えることができる。このことは以下の論理式で表され (ここで m_i は「c_i は赤帽」を表し、$g = \{c_1, \cdots, c_n\}$ とする)、例えば $n = 4$ の場合は図 3 のように証明できる (長いため、複数回の規則適用を適宜まとめ、Weak の適用は省いて記した)。

$$
\begin{array}{l}
\mathsf{G\text{-}BEL}_g(\bigvee_{j=1}^{n} m_j) \wedge \\
\bigwedge_{1 \leq j \leq n, 1 \leq j' \leq n, j \neq j'} \mathsf{G\text{-}BEL}_g(\mathsf{BEL}_{c_j}(\neg m_{j'}) \Leftrightarrow \neg m_{j'}) \wedge \\
\bigwedge_{j=1}^{n-1} \mathsf{G\text{-}BEL}_g(\neg \mathsf{BEL}_{c_j}(m_j) \wedge \neg \mathsf{BEL}_{c_j}(\neg m_j)) \supset \mathsf{BEL}_{c_n}(m_n)
\end{array}
$$

6 まとめ

複数エージェントの心的状態や相互心的状態を扱えるように拡張した、BDI logic の演繹体系を示し、またそれを用いて、エージェント間の相互信念などに関するいくつかの性質の証明の例を示した。今後の課題には、ベースとなる BDI logic を CTL* ベースや述語論理に拡張すること、マルチエージェント環境のエージェント記述言語への応用のための実行系の実現、action expression の導入、などが挙げられる。

参考文献

[1] E. Allen Emerson and J. Srinivasan. Branching Time Temporal Logic. In J.W. de Bakker, W.P. de Roever, and G. Rozenberg, editors, *Linear Time, Branching Time and Partial Order in Logics and Models for Concurrency*, pp. 123–172. Springer-Verlag, 1989.

[2] N. Nide, S. Takata, and T. Araragi. Deduction Systems for BDI Logics with Mental State Consistency. In *Proc. of CLIMA '02*, pp. 123–135, 2002.

[3] Anand S. Rao and Michael P. Georgeff. Modeling Rational Agents within a BDI-Architecture. In *Proc. of International Conference on Principles of Knowlegde Representation and Reasoning*, pp. 473–484, 1991.

[4] Anand S. Rao and Michael P. Georgeff. Decision Procedures for BDI Logics. *Journal of Logic and Computation*, Vol. 8, No. 3, pp. 292–343, 1998.

[5] Michael Wooldridge. *Reasoning about Rational Agents*. The MIT Press, 2000.

[6] 新出尚之, 高田司郎, 櫟粛之. 相互信念と協調行為を扱うための演繹体系. 日本ソフトウェア科学会第 19 回大会論文集, 2002. (CD-ROM, file `pdf/6C-2.pdf`).

[7] 新出尚之, 高田司郎, 櫟粛之. BDI Logic の sequent calculus による演繹体系. コンピュータ・ソフトウェア, Vol. 20, No. 1, pp. 66–83, 2003.