# 信念と状態遷移を確率的に扱う合理的エージェント向け論理

新出　尚之[†a)]　　高田　司郎[††]

# A Logic for Rational Agents Which Handles Beliefs and State Transitions Probabilistically

NIDE, Naoyuki[†a)] and Shiro TAKATA[††]

**あらまし**　合理的エージェントのモデルである BDI モデルでは、時相論理体系 BDI logic によってエージェントの振る舞いの形式的記述が行える。我々はこれを拡張し、強化学習との結合などを形式化するために必要となる確率的状態遷移の記述を可能にした時相論理体系 $\mathcal{TOMATO}\mathit{es}$ を提案した。しかし $\mathcal{TOMATO}\mathit{es}$ では、状態遷移を表す時相オペレータは確率的に扱えるが、確率的な信念は扱えなかった。一方、実際のエージェントでは確率的な信念を扱うことが必要な応用も多い。そこで、信念と状態遷移をともに確率的に扱える論理モデルが必要となるが、同時にそれは、合理的エージェントに要請される心的状態の整合性を扱えることをも要請される。本論文ではそのような論理モデルとして、$\mathcal{TOMATO}\mathit{es}$ に確率的信念オペレータを導入して拡張した論理体系 $\mathcal{B}$-$\mathcal{TOMATO}\mathit{es}$ を提案し、また、確率的な信念を用いた記述例について述べる。

**キーワード**　合理的エージェント, BDI, 確率, 論理

## 1. はじめに

我々は、BDI モデル [1] をベースに、学習による基本行為の獲得との結合 [2] などによって、動的環境下で目的を達成するために行動できる合理的エージェントの実現を目指し、研究を進めてきた。

BDI モデルでは、時相論理体系 BDI logic によって、エージェントの心的状態 (信念·願望·意図) やその経時変化を論理モデルで記述でき、エージェントの振る舞いについて形式的な議論が行えることが大きな利点の 1 つである。そのため、BDI モデルにそれ以外の意思決定メカニズムを導入して拡張するにあたっても、その拡張を論理モデルの中に取り込めることが望まれる。

我々はこの観点から、我々が BDI モデルに対して行った拡張を単一の論理モデルで扱えるようにするための試みとして、確率的な状態遷移とイベントの選択が記述できるように BDI logic を拡張した論理体系 $\mathcal{TOMATO}\mathit{es}$ [3], [4] を提案し、この体系のもとで BDI モデルと強化学習やエージェント相互間の協調行為との結合に関する記述や推論などが行えることを示した。

しかし、$\mathcal{TOMATO}\mathit{es}$ では確率的な記述が行えるのは時相オペレータ、すなわち状態遷移のみであり、信念オペレータを確率的に記述することができない。一方で、動的環境下、中でも実世界においては、一般に知覚の手段も精度も限られていること、また行動も誤差を免れ得ないことなどから、信念は不確実さを伴う場合が多く、従ってそうした不確実さをもった信念を扱う必要に迫られる応用も多い。代表的な例としては、POMDP [5] の下での強化学習が挙げられる。また、統計的アブダクション [6] も、観察に対する説明を確率分布に基づいて導き出す点で、そのような例の 1 つに挙げられよう。

他に、確率的信念の記述の別な応用として、感情表現の実現が考えられる。近年は、人との対話や交流を行うエージェントに、感情表現を付加する応用が注目されている。BDI モデルの下で感情表現の形式化を試みた例として、Adam ら [7][(注1)] や Steunebrink ら [10] のものがあるが、それらは感情の生起を、信念など心的状態を用いて記述した論理式の真偽としてモデル化しており、基となる心的状態が 0/1 の 2 値のため、感情が生起するか否かの区別しか記述できず、生起された感情の度合い (「すごく嬉しい」「まあまあ嬉しい」など) を表現できていない。これを解決するには、ベースとなる心的状態を確率的に扱える必要がある。

いずれの応用においても、合理的エージェントへの応用として捉えるには、前述の観点から、確率的信念が単独で扱えるだけでは十分でなく、それが BDI モデルと統合的に扱える必要がある。特に、BDI モデルにおいて合理的エージェントに要請される心的状態の整合性は、確率的信念の導入後も引き続き成り立つことが求められる。

そこで本論文では、$\mathcal{TOMATO}\mathit{es}$ に確率的[(注2)]信念の扱いを導入して拡張した論理体系 $\mathcal{B}$-$\mathcal{TOMATO}\mathit{es}$ を提案し、その記述事例を示す。ただし、感情表現の事例については [11] に譲り、本論文では扱わない。

[†] 奈良女子大学 研究院 生活環境科学系, 奈良市
Faculty, Division of Human Life and Environmental Sciences, Nara Women's University, Kita-uoya Nishimachi, Nara-shi, 630-8506 Japan
[††] 近畿大学 理工学部, 東大阪市
Faculty of Science and Engineering, Kinki University, Higashi-Osaka-shi, 577-8502 Japan
a) E-mail: nide@ics.nara-wu.ac.jp

(注1)：これに基づいた、Jason [8] 上での我々による実装の試みもある [9]。
(注2)：「確率」(probability) ではなく「確度」(certainty) と見るべき場合もあるかも知れないが、本論文では両者をあえて区別はしていない。

## 2. 論理体系 $\mathcal{B}$-$\mathscr{TOMATOes}$

$\mathcal{B}$-$\mathscr{TOMATOes}$ は、一階述語論理を拡張した点時刻分岐時相論理に、エージェント毎の心的状態 (信念·願望·意図) オペレータ、およびイベントによる状態 (時刻) 遷移を表現するオペレータを追加したものであり、特に状態遷移オペレータと信念オペレータは確率パラメータを持つ。

### 2.1 構　文

以下、単に「論理式」と言えば $\mathcal{B}$-$\mathscr{TOMATOes}$ の論理式を指す。また、$x, y$ などは一階述語論理の通常の変数記号として用い、$\mathfrak{X}, \mathfrak{Y}$ などは論理式を表す変数記号 (以下「論理式変数記号」と呼ぶ) として用いる。エージェント定数記号およびイベント定数記号の有限集合 $\mathcal{A}, \mathcal{E}$ を選んで固定しておく。$\{p \mid p \in \mathbb{R}, 0 \leqq p \leqq 1\}$ を $[0,1]$ と書く。

論理式を以下の BNF で定義する。ここで、$\phi$ は論理式、$\alpha$ は一階述語論理の原始論理式、$a, e, p, r$ はそれぞれ $\mathcal{A}$, $\mathcal{E}$, $[0,1]$, $\{\geqq, >\}$ の要素を表すメタ変数とする。なお、構文中に「|」を含む論理式があるため、BNF の選択記号は「¦」で表す。また「[ ]*」は反復を表す。

$$\phi ::= \alpha \mathrel{¦} \neg\phi \mathrel{¦} \phi \vee \phi \mathrel{¦} \forall x\phi \mathrel{¦} \mathfrak{X} \mathrel{¦} \mathsf{pos}(e) \mathrel{¦} \mathsf{X}^e({}_{rp}\,\phi[\,|\,{}_{rp}\,\phi]^*) \mathrel{¦}$$
$$\mathsf{BEL}^a({}_{rp}\,\phi[\,|\,{}_{rp}\,\phi]^*) \mathrel{¦} \mathsf{DESIRE}^a\,\phi \mathrel{¦} \mathsf{INTEND}^a\,\phi \mathrel{¦} \mu\mathfrak{X}.\phi$$

ただし、$\mu\mathfrak{X}.\phi$ の $\phi$ の中で、$\mathfrak{X}$ は $\neg$ の奇数段のネストの中に現れてはならない。$\mu$ は 2.2.2 節で述べる最小不動点オペレータである。$\mathscr{TOMATOes}$ との構文上の差は、信念オペレータ BEL が確率パラメータを持つ点だけである(注3)。例えば $\mathsf{BEL}^a({}_{\geqq 0.7}\,\phi \mid {}_{\geqq 0.3}\,\psi)$ は論理式である。

この他、$\wedge, \supset, \Leftrightarrow, \exists$ などを一般的な略記として導入し、オペレータ間に一般的な優先順位 (例えば $\wedge$ が $\vee$ より先に結合するなど) を導入する。また、$\mathsf{BEL}^a({}_{r_1 p_1}\,\phi_1)$ を $\mathsf{BEL}^a_{r_1 p_1}\,\phi_1$ と、さらに $\mathsf{BEL}^a_{\geqq p_1}\,\phi_1 \wedge \neg\,\mathsf{BEL}^a_{> p_1}\,\phi_1$ を $\mathsf{BEL}^a_{= p_1}\,\phi_1$ と略記する。$\mathsf{BEL}^a_{< p_1}\,\phi_1$, $\mathsf{BEL}^a_{\leqq p_1}\,\phi_1$ なども同様な略記として定義できる (以上、$\mathsf{X}^e$ についても同様)。

直感的には、$\mathsf{pos}(e)$ は「イベント $e$ が実行可能」、$\mathsf{X}^e({}_{r_1 p_1}\,\phi_1 \mid \cdots \mid {}_{r_n p_n}\,\phi_n)$ は「イベント $e$ を実行すると、次の状態 (時刻) では確率 $r_1 q_1$ で $\phi_1$ が、$\cdots$、確率 $r_n q_n$ で $\phi_n$ が (排他的に) 成り立つ」、$\mathsf{BEL}^a({}_{r_1 p_1}\,\phi_1 \mid \cdots \mid {}_{r_n p_n}\,\phi_n)$ は「エージェント $a$ は確度 $r_1 q_1$ で $\phi_1$ を、$\cdots$、確度 $r_n q_n$ で $\phi_n$ を (排他的に) 信じている」、$\mathsf{DESIRE}^a\,\phi$, $\mathsf{INTEND}^a\,\phi$ は「エージェント $a$ は $\phi$ を願望している」「(同) 意図している」を表す。ここで「排他的に」とは、例えば $\mathsf{BEL}^a({}_{\geqq 0.5}\,\phi \mid {}_{\geqq 0.5}\,\psi)$ の意味として「確度 0.5 で $\phi$ も $\psi$ も成り立つが、残り 0.5 ではどちらも成り立たない、と信じる」のようなものを排除し、「確度 0.5 ずつで case1 と case2 のどちらかが起き、case1 では $\phi$ が成り立ち ($\psi$ も成り立つかもしれない)、case2 では $\psi$ が成り立つ ($\phi$ も成り立つかもしれない) と信じる」のように、各事象 (ここでは $\phi$ と $\psi$) がそれぞれ別の場合として確度を持つようなもののみを認めることをいう ($\phi$ と $\psi$ が同時に成り立たないという要請を表す「排他」ではない)。

さらに $\mathsf{BEL}^a_{\geqq 1}\,\phi$ を $\mathsf{BEL}^a\,\phi$ と、$\mathsf{X}^e_{\geqq 1}\,\phi$ を $\mathsf{AX}^e\,\phi$ と、$\bigwedge_{e \in \mathcal{E}}(\mathsf{pos}(e) \supset \mathsf{AX}^e\,\phi)$ を $\mathsf{AX}\,\phi$ と略記する。これらは通常の信念様相論理や CTL での BEL や AX に相当する。

なお、心的状態の中では信念にのみ確率を導入している(注4)。BDI モデルでの願望や意図は、エージェントが達成すべき目標とその手段を自らの意思で選択したものであるため、「あることを確率 0.8 で願望/意図する」といったことは考えにくいことがその理由である。

### 2.2 意　味　論

#### 2.2.1 BDI ストラクチャ

以下のものを定めておく。

- 可能世界の集合 $W (\neq \emptyset)$
- $W$ の各要素 $w$ に対し、state の集合 $St_w (\neq \emptyset)$
- $W$ の各要素 $w$ と $St_w$ の各要素 $t \in St_w$ に対し、一階述語論理の項ないし原始論理式を与えてその解釈を返す関数 $i_{w,t}$。ただし解釈に用いる領域 (Domain) $U$、および項の解釈は $w, t$ に関わらず共通とする
- $W$ の各要素 $w$ と $St_w$ の各要素 $t \in St_w$ に対し、$\mathcal{E}$ の空でない部分集合 $pos_{w,t}$
- $W$ の各要素 $w$ と $\mathcal{E}$ の各要素 $e$ に対し、${St_w}^2$ から $[0,1]$ への関数 $R^e_w$。ただし、任意の $t \in St_w$ に対し $\sum_{t' \in St_w} R^e_w(t, t') = 1$ であること
- $\mathcal{A}$ の各要素 $a$ と $\bigcup_{w \in W} St_w$ の各要素 $t$ に対し、集合 $\{w \mid t \in St_w\}$ (以後 $W_t$ と書く) 上の serial な(注5)2 項関係 $\mathcal{D}^t_a, \mathcal{I}^t_a$、および ${W_t}^2$ から $[0,1]$ への関数 $\mathcal{B}^t_a$ であって、以下を満たすもの
  - ⋆ 任意の $w \in W_t$ に対し $\sum_{w' \in W_t} \mathcal{B}^t_a(w, w') = 1$
  - ⋆ $\mathcal{B}^t_a(w, w') > 0$ なる任意の $w, w'$ に対し $w \overset{M}{\equiv}_{t,a} w'$ かつ $w'$ のある subworld $w'' \in W_t$ があって $w\ \mathcal{D}^t_a\ w''$
  - ⋆ $w\ \mathcal{D}^t_a\ w'$ なる任意の $w, w'$ に対し、$w'$ のある subworld $w'' \in W_t$ があって $w\ \mathcal{I}^t_a\ w''$

ただし、ここで $w \overset{M}{\equiv}_{t,a} w'$ とは任意の $w'' \in W_t$ が以下を満たすことと定義する: ①$\mathcal{B}^t_a(w, w'') = \mathcal{B}^t_a(w', w'')$ ②$w\ \mathcal{D}^t_a\ w''$ **iff** $w'\ \mathcal{D}^t_a\ w''$ ③$w\ \mathcal{I}^t_a\ w''$ **iff** $w'\ \mathcal{I}^t_a\ w''$。また、$w'$ が $w$ の subworld であるとは、以下を満たすことと定義し、このとき $w \sqsupseteq w'$ と書く ($w$ 自身も $w$ の subworld の 1 つであることに注意)。

- $St_w = St'_w$　• 任意の $e \in \mathcal{E}$ に対し $R^e_w = R^e_{w'}$
- 任意の $t \in St_w$ に対し $i_{w,t} = i_{w',t}$
- 任意の $t \in St_w$ に対し $pos_{w,t} \supseteqq pos_{w',t}$
- 任意の $t \in St_w, a \in \mathcal{A}$ に対し $w \overset{M}{\equiv}_{t,a} w'$

以上を組にしたものを、ここでは BDI ストラクチャと呼ぶ。大まかには、state は時相論理の「点時刻」で、1 つの可能世界は $\{(t, t') \mid$ ある $e \in pos_{w,t}$ に対し $R^e_w(t, t') > 0\}$ をエッジとする時刻の木であり、エッジは時刻の

(注3): さらに確率パラメータに任意の項を許せば、例えば 3. 節の式 (1) (この式は BEL を略記で含む) を確率パラメータについて全称限量し一般化するなどの応用も可能となる。その代わり、2.2.2 節で触れる演繹体系の構築は難しくなろう。本論文では rough な議論のみ行うため、この点は考慮しない。

(注4): [11] での Desirability などの扱いは本論文とはやや異なっている。

(注5): 集合 $S$ 上の 2 項関係 $R$ が serial であるとは、任意の $s \in S$ に対し、$s\ R\ s'$ となる $s' \in S$ が存在することをいう。

前後関係を表す。$pos_{w,t}$ は state $t$ で実行可能なイベントの集合を表し、$R^e_w(t,t')=p$ は state $t$ でイベント $e$ を実行すると確率 $p$ で次の時刻は state $t'$ になることを表す。$w$ の subworld とは、$w$ と比べて、各 state での実行可能なイベントだけが制限され、それ以外は $w$ と全く同じであるような世界である。

$\mathcal{D}^t_a, \mathcal{I}^t_a$ は時刻 $t$ における可能世界間の可視関係で、エージェント $a$ の願望·意図を表す。$\mathcal{B}^t_a$ は時刻 $t$ におけるエージェント $a$ の確度つき信念を与えるための、可能世界間の確度つき可視関係にあたるものである。

図 1 に BDI ストラクチャの概略を示した (ただしこの図は、簡単のため $\mathcal{I}^t_a$ は表示しておらず、$\mathcal{B}^t_a$ や $\mathcal{D}^t_a$ などについても全て表示してはいない。特に、$\mathcal{D}^t_a$ の serial 性などの性質も反映していない)。state の丸の中に書かれているイベントはその state で可能なイベント、エッジに付いているイベントはそのイベントでの遷移であることを表す。$w'_2$ は $w'_1$ の、$w''_2$ は $w''_1$ の subworld である。

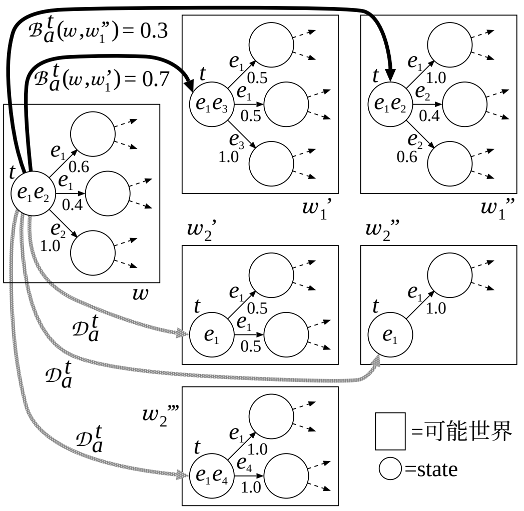


図 1　BDI ストラクチャの概略

### 2.2.2　論理式の解釈

以後、$\{(w,t) \mid w \in W, t \in St_w\}$ を $Swt$ と書く。

BDI ストラクチャ $M$、および論理式変数記号の集合 $\mathcal{V}$ から $2^{Swt}$ への関数 $f_\mathcal{V}$ を 1 つ定めておく。論理式 $\phi$ に対し、その解釈 $[\![\phi]\!]_{\langle M,f_\mathcal{V}\rangle}$ ($\subseteq Swt$) を以下のように定める。

- $\phi$ が原始論理式のとき、$[\![\phi]\!]_{\langle M,f_\mathcal{V}\rangle} = \{(w,t) \mid i_{w,t}(\phi)$ が真$\}$
- $[\![\phi \vee \psi]\!]_{\langle M,f_\mathcal{V}\rangle} = [\![\phi]\!]_{\langle M,f_\mathcal{V}\rangle} \cup [\![\psi]\!]_{\langle M,f_\mathcal{V}\rangle}$
- $[\![\neg\phi]\!]_{\langle M,f_\mathcal{V}\rangle} = Swt \setminus [\![\phi]\!]_{\langle M,f_\mathcal{V}\rangle}$
- $[\![\forall x \phi]\!]_{\langle M,f_\mathcal{V}\rangle} = \bigcap_{u \in U} [\![\phi]\!]_{\langle M[x:=u],f_\mathcal{V}\rangle}$　ただし $M[x:=u]$ は $M$ での $x$ の解釈を $u$ に変更して得られる BDI ストラクチャ
- $[\![\mathsf{pos}(e)]\!]_{\langle M,f_\mathcal{V}\rangle} = \{(w,t) \mid pos_{w,t} \ni e\}$
- $[\![\mathsf{X}^e({}_{r_1 p_1}\phi_1 \mid \cdots \mid {}_{r_n p_n}\phi_n)]\!]_{\langle M,f_\mathcal{V}\rangle} = \{(w,t) \mid$ 各 $i=1,\cdots,n$ について $T_i \subseteqq \{t' \mid (w,t') \in [\![\phi_i]\!]_{\langle M,f_\mathcal{V}\rangle}\}$ かつ $(\sum_{t' \in T_i} R^e_w(t,t'))\ r_i\ p_i$ を満たすような、互いに素な集合 $T_1,\cdots,T_n$ が存在する$\}$ ($r_i$ は $>, \geqq$ のいずれかであることに注意)
- $[\![\mathsf{BEL}^a({}_{r_1 p_1}\phi_1 \mid \cdots \mid {}_{r_n p_n}\phi_n)]\!]_{\langle M,f_\mathcal{V}\rangle} = \{(w,t) \mid$ 各 $i=1,\cdots,n$ について $W_i \subseteqq \{w' \mid (w',t) \in [\![\phi_i]\!]_{\langle M,f_\mathcal{V}\rangle}\}$ かつ $(\sum_{w' \in W_i} \mathcal{B}^t_a(w,w'))\ r_i\ p_i$ を満たすような、互いに素な集合 $W_1,\cdots,W_n$ が存在する$\}$
- $[\![\mathsf{DESIRE}^a\,\phi]\!]_{\langle M,f_\mathcal{V}\rangle} = \{(w,t) \mid w\ \mathcal{D}^t_a\ w'$ なる任意の世界 $w'$ に対して $(w',t) \in [\![\phi]\!]_{\langle M,f_\mathcal{V}\rangle}\}$
- $[\![\mathsf{INTEND}^a\,\phi]\!]_{\langle M,f_\mathcal{V}\rangle}$ については同様
- $\mathfrak{X} \in \mathcal{V}$ のとき、$[\![\mathfrak{X}]\!]_{\langle M,f_\mathcal{V}\rangle} = f_\mathcal{V}(\mathfrak{X})$

また、以上の定義からは、論理式変数記号 $\mathfrak{X}$ の自由な出現を持つ (持たなくても) 論理式 $\phi$ を、$\mathfrak{X}$ の解釈を受け取って $\phi$ の解釈を返す関数 $f_\phi : 2^{Swt} \to 2^{Swt}$ と捉え直すことができる。そこで、

- $[\![\mu\mathfrak{X}.\phi]\!]_{\langle M,f_\mathcal{V}\rangle}$ は、$f_\phi$ の最小不動点である

と定義する。この場合 $f_\phi$ は定義から単調関数なので、最小不動点の存在は保証される。不動点オペレータがあることから、時相論理としては (AX があるため) CTL*より強い表現力を持ち、また相互信念などの概念も記述できる。

$[\![\phi]\!]_{\langle M,f_\mathcal{V}\rangle}$ は、$\phi$ の成り立つ世界と state のペア $(w,t)$ の集合と見ることができる。$[\![\phi]\!]_{\langle M,f_\mathcal{V}\rangle} \ni (w,t)$ であるとき、世界 $w$ の state $t$ で $\phi$ が成り立つという。例えば図 1 において、世界 $w$ の state $t$ で $\mathsf{BEL}^a({}_{\geqq 0.7}\,\mathsf{pos}(e_3) \mid {}_{\geqq 0.3}\,\mathsf{pos}(e_2))$ が成り立つ。論理式 $\phi$ が、任意の $M, f_\mathcal{V}$ に対し $[\![\phi]\!]_{\langle M,f_\mathcal{V}\rangle} = Swt$ を満たすとき、$\phi$ は恒真であるという。

なお、$\mathcal{B}$-TOMATOes では信念の確率分布は離散なものとして扱っている。一方、我々は TOMATOes を状態遷移について連続な確率分布が扱えるように拡張した $\mathcal{C}$-TOMATOes [12] を既に提案している。同じ手法で信念の連続な確率分布を導入することも可能であり、特に実世界への応用を考えればそれは必要でもあると考えられる。しかし本論文では、簡単のためそれは行っていない。

また、TOMATOes と同様の手法で $\mathcal{B}$-TOMATOes に演繹体系を与え、(少なくとも) 健全性を示すことも可能であろう。本論文では演繹体系に関する議論は割愛する。

### 2.3　心的状態の整合性

$\mathcal{B}$-TOMATOes では、在来 BDI logic や TOMATOes での恒真論理式と同様の、次のような論理式が恒真である。

$$\neg\,\mathsf{BEL}^a\,\mathit{false} \qquad \mathsf{BEL}^a_{rp}\,\phi \Leftrightarrow \mathsf{BEL}^a\,\mathsf{BEL}^a_{rp}\,\phi$$
$$\mathsf{DESIRE}^a\,\phi \Leftrightarrow \mathsf{BEL}^a\,\mathsf{DESIRE}^a\,\phi$$
$$\mathsf{INTEND}^a\,\phi \Leftrightarrow \mathsf{BEL}^a\,\mathsf{INTEND}^a\,\phi$$
$$\mathsf{DESIRE}^a\,\theta \supset \mathsf{BEL}^a\,\theta \qquad \mathsf{INTEND}^a\,\theta \supset \mathsf{DESIRE}^a\,\theta$$

ここで $\mathit{false}$ は $\phi \wedge \neg\phi$ の略記とする。上 4 式 (3 行) は、自身の心的状態に関する完全な信念を持つこと、および BEL オペレータに関する KD45 相当の要請である。下 2 式 (1 行) は strong realism [1] と呼ばれ、「エージェントが信念と矛盾する意思決定を行わない」など、合理性のために要請される性質である。ただし下 2 式で $\theta$ は O-formula [1] ($\mathcal{B}$-TOMATOes では BEL, DESIRE, INTEND のスコープ内を除き、pos が $\neg$ の奇数段のネストの中に現れない論理式。直感的には「あることが特定のイベントの選択に

よる未来でのみ成り立つ」を表す) とする。

## 3. 事 例

ここではスペースの都合上、簡単な事例のみ取り上げる。[5] による以下の問題を考える。エージェントが 2 つのドアの前にいる。一方のドアの向こうには虎がおり、もう一方のドアの向こうには宝がある。虎のいるドアを開けると報酬 $-100$、宝のあるドアを開けると報酬 10 を得てそれで 1 エピソードが終わる。また、虎のいるドアを知るために虎の鳴き声を聞くこともできるが、それには報酬 $-1$ がかかる。しかもその知覚は不正確で、鳴き声が聞こえた方のドアに実際に虎がいる確率は 0.85 である。

以下では $\mathsf{BEL}^a({}_{\geq p}\,\mathrm{tiger(left)} \mid {}_{\geq 1-p}\,\mathrm{tiger(right)})$ を $\mathrm{tiger\_left}(p)$ と略記する。エピソードの最初にはエージェント $a$ は虎の位置に関する情報を全く持たないとすると、$a$ の信念では虎の位置は半々の確率で左か右であり、これは $\mathrm{tiger\_left}(0.5)$ と書ける。

ある時刻で $\mathrm{tiger\_left}(p)$ が成り立っており、$a$ が「虎の声を聞く」というアクション (これをイベント hear で表す) を取ったとする。その結果、左から鳴き声が聞こえたので、$a$ は虎が左にいる事後確率を求め、信念を修正する。この過程は

$$\begin{aligned}&\mathrm{tiger\_left}(p) \supset \mathsf{AX}^{\mathrm{hear}}(\mathrm{roar(left)} \wedge \\ &\quad (p' = \tfrac{.85p}{.85p+.15(1-p)}) \supset \mathrm{tiger\_left}(p'))\end{aligned} \tag{1}$$

と書ける。ただし、通常の数学の演算記号を表す $=, +$ などの述語記号や関数記号が一階言語に入っており、解釈も数学でのそれと同じものに固定されているとする (以後も同様)。このように、(ベイズモデルを用いた) エージェントの信念の更新を、確率的信念の経時変化として捉えることができる。

ここまでは現在の状況に関する確率的信念を扱ったが、強化学習の方策を「どの行為を取るか」に関する確率的信念と捉えることもでき、その場合、方策を $\mathsf{BEL}$ オペレータで記述することも可能である。例えば上の問題を、エージェント $a$ が現在の状態 $\mathrm{tiger\_left}(p)$ をもとに行為 hear, open_left, open_right のどれを取るかの戦略を得る強化学習問題と捉えるとする。ある方策を「確率 $q_0$ で hear を、$q_1$ で open_left を、$1-q_0-q_1$ で open_right を実行する」というものとすると、「$p$ がある範囲内 ($p_0 \leqq p \leqq p_1$) に入る場合はこの方策を取る」は $\mathrm{tiger\_left}(p) \wedge (p_0 \leqq p \leqq p_1) \supset \mathsf{BEL}^a({}_{\geqq q_0}\,\mathsf{pos\_only}(\mathrm{hear})|_{\geqq q_1}\,\mathsf{pos\_only}(\mathrm{open\_left})|_{\geqq 1-q_0-q_1}\,\mathsf{pos\_only}(\mathrm{open\_right}))$ と書ける ($\mathsf{pos\_only}(e)$ は $\mathsf{pos}(e) \wedge \bigwedge_{e' \in \mathcal{E}, e' \neq e} \neg\mathsf{pos}(e')$ の略記、すなわち「イベント $e$ のみが選択可能」を表すとする)。

方策により例えばイベント hear の実行が選ばれればエージェントはそれを意図し ($\mathsf{INTEND}^a\,\mathsf{pos}(\mathrm{hear})$)、合理的エージェントの性質「基本行為が選択されれば直ちに実行される」によりイベントが実行される。このように、$\mathcal{B}$-$\mathscr{TOMATO}es$ ではエージェントの行為選択メカニズムなどを、確率的な信念や状態遷移と一体的に扱える (この例はシングルエージェントであったが、マルチエージェントでも同様の扱いが可能である)。

## 4. ま と め

本論文では、BDI logic の我々による拡張 $\mathscr{TOMATO}es$ にさらに確率的信念を導入した体系 $\mathcal{B}$-$\mathscr{TOMATO}es$ を提案し、記述例を示した。

確率的信念、ないしは不確実な信念を扱う論理モデルに関する研究は以前から多数存在する。また、「信念」という形をとらずに確率的な言明を扱う論理の研究も多い。それらと比べ、$\mathcal{B}$-$\mathscr{TOMATO}es$ はエージェントの行為選択メカニズムの形式化を同一の体系の中で可能とし、心的状態の整合性など合理的エージェントに要請される性質も満たせることが利点である。ただし、本論文の目的は形式化を示すことにとどまり、確率的信念を利用した個々の具体的なアルゴリズム、例えば 3. 節で例に挙げた POMDP による確率的信念の更新や強化学習の具体的な手法などには踏み込んでおらず、こうした応用への適用は課題として残る。また、マルチエージェントへの応用を考える場合、例えば [13] などゲーム理論をバックグラウンドとした体系などとの比較も必要であろう。これらを通じて本体系の優位性を示すことが今後の課題である。

## 文 献


[1] A.S. Rao and M.P. Georgeff, “Modeling Rational Agents within a BDI-Architecture,” Readings in Agents, eds. by M.N. Huhns and M.P. Singh, pp.317–328, Morgan Kaufmann, San Francisco, 1997.

[2] 高田司郎，新出尚之，濱砂幸裕，波部 斉，藤田 恵，“アトラクター状態を用いた実世界における基本行為の学習について，” 情報処理学会研究報告 2013-MPS-92，no.24，pp.1–6，2013.

[3] 新出尚之，高田司郎，藤田 恵，“拡張 BDI 論理 $\mathscr{TOMATO}es$ による協調行為のモデル化と応用，” 人工知能学会論文誌，vol.26，no.1，pp.13–24，2011.

[4] 高田司郎，新出尚之，藤田 恵，“拡張 BDI 論理 $\mathscr{TOMATO}es$ を用いた強化学習のモデル化について，” 人工知能学会論文誌，vol.26，no.1，pp.156–165，2011.

[5] L.P. Kaelbling, M.L. Littman, and A.R. Cassandra, “Planning and acting in partially observable stochastic domains,” Artificial Intelligence, vol.101, pp.99–134, 1998.

[6] 佐藤泰介，“統計的アブダクション，” 人工知能学会誌，vol.25，no.3，pp.400–407，2010.

[7] C. Adam, A. Herzig, and D. Longin, “A logical formalization of the OCC theory of emotions,” Synthese, vol.168, no.2, pp.201–248, 2009.

[8] R.H. Bordini, J.F. Hübner, and M. Wooldridge, Programming Multi-Agent Systems in AgentSpeak using Jason, John Wiley & Sons, 2007.

[9] 藤田 恵，清水詩子，新出尚之，“感情表現を行動決定に取り入れた BDI エージェントの実現に向けて，” Proc. of JAWS2011，2011．77fujita.pdf.

[10] B.R. Steunebrink, M. Dastani, and J.-J.C. Meyer, “A formal model of emotion triggers: an approach for BDI agents,” Synthese, vol.185, no.1 Supplement, pp.83–129, 2012.

[11] 池之内彰子，新出尚之，“OCC theory に基づくエージェントの感情表現の論理モデルについて，” Proc. of JAWS2014，pp.91–94，2014.

[12] 新出尚之，高田司郎，藤田 恵，“連続した状態空間での合理的エージェントの行為を扱う論理モデルの試み，” 電子情報通信学会論文誌，vol.J96-D，no.12，pp.2939–2950，2013.

[13] C. Zhou, “Probability logic for Harsanyi type spaces,” http://arxiv.org/abs/1405.6355, 2014.