情報処理学会研究報告
IPSJ SIG Technical Report

Vol.2010-MPS-81 No.23
Vol.2010-BIO-23 No.23
2010/12/17

拡張 BDI 論理 $\mathcal{TOMATO}$ を用いた 確率的状態遷移のモデル化とその応用

新出尚之[†1]　高田司郎[†2]　藤田恵[†3]

BDI エージェントのモデル化に用いられる様相論理体系 BDI logic に、確率的状態遷移と不動点オペレータを導入して拡張した論理体系が $\mathcal{TOMATO}$ である。これらの拡張を用いて、$\mathcal{TOMATO}$ では BDI エージェントおよびそれと外部の行為選択機構との結合にかかわるさまざまな性質の記述や推論が行える。本論文では、$\mathcal{TOMATO}$ の意味論で用いられるクリプケ構造である BDI ストラクチャを用いて、確率的状態遷移のモデルを扱う方法、および、それに関する証明の例について示す。特に、確率的状態遷移の導入が本質的に必要となる、強化学習との結合に主眼をおいて述べる。

Modeling probabilistic state transitions using $\mathcal{TOMATO}$ and its application

NIDE, NAOYUKI,[†1] SHIRO TAKATA[†2]
and MEGUMI FUJITA[†3]

$\mathcal{TOMATO}$ is an extension of BDI logic, which introduced probabilistic state transitions and fix-point operators. In $\mathcal{TOMATO}$, using those extended notions, we can strictly describe and infer various properties of BDI agents and combinations of them with the external action decision mechanisms. In this paper, we give a detailed explanation of modeling of probabilistic state transitions with the Kripke structure used in $\mathcal{TOMATO}$, called BDI structure, and the inference rules of $\mathcal{TOMATO}$. In addition, we give some proof examples using $\mathcal{TOMATO}$. Especially, we pay particular attention to the modeling of the combination with reinforcement learning.

1. はじめに

BDI モデルでは、合理的エージェントは、明示的に信念・願望・意図の 3 種類の心的状態を持ち、これら心的状態を保持・更新することで意思決定を行い、目的を達成するように振る舞う。BDI モデルの特徴の 1 つは、これら 3 種類の心的状態やその時間的変化を明示的に記述できる BDI logic という様相論理体系を持ち、これによって合理的エージェントの心的状態や振る舞いに関する形式的な議論や証明が行えることが保証される点である。この点が合理的エージェントの設計上大きな利点と考えられ、BDI モデルが受け入れられる一要因となってきた。

しかし、合理的エージェントの研究の進展においては、元来の BDI モデルにはなかったさまざまな拡張が試みられてきている。それらと元来の BDI logic で扱える概念との間にずれがあると、BDI logic による十分な形式化が行えず、合理的エージェントに関する厳密な議論を行えるという BDI モデルの利点の一端が損なわれることになる。その例としては、強化学習[5] における「確率的状態遷移」や、マルチエージェント環境における相互信念や共同意図などを必要とする「協調行為」などが挙げられる。特にこれらの概念は、実世界の合理的エージェントの実現においては今後重要になると考えられる。そこで我々は、これらの概念をそれぞれ形式的に扱えるように、従来の BDI logic を拡張して確率的遷移と不動点オペレータを導入した論理体系 $\mathcal{TOMATO}$ (Theory about Observations of Multi-Agents with Tense and Odds) を提案している[4),7)]。

本論文では、上述した BDI モデルへの拡張のうち特に強化学習との結合に着目し、これを $\mathcal{TOMATO}$ での確率的状態遷移の記述を用いて BDI の論理モデルに取り込む事例、特にエージェントの性質に関する証明の例について述べる。

2. 論理体系 $\mathcal{TOMATO}$

本節では様相論理体系 $\mathcal{TOMATO}$ を導入する。この論理は、点時刻分岐時相論理に、不動点オペレータと確率的状態遷移オペレータ、およびマルチエージェント環境における個々の

†1 奈良女子大学理学部
Facluty of Science, Nara Women's University
†2 近畿大学理工学部
School of Science and Engineering, Kinki University
†3 奈良女子大学大学院人間文化研究科
Graduate School of Humanities and Sciences, Nara Women's University

Vol.2010-MPS-81 No.23
Vol.2010-BIO-23 No.23
2010/12/17

エージェントに関する心的状態オペレータを導入したものであり、1 節で述べたような BDI モデルへの外部の行動選択機構の導入を扱いやすくすることを重視している。

2.1 論 理 式

2.1.1 構 文

まず $\mathscr{TCMATC}$ の論理式の定義を与える。以下では、単に「論理式」と言えば $\mathscr{TCMATC}$ の論理式を指すものとする。変数記号として、x や y などは一階述語論理での通常の変数記号として用い、$\mathfrak{X}$ や $\mathfrak{Y}$ などは論理式を表現する変数記号として用いる。後者は以降「論理式変数記号」と呼称し、主として不動点オペレータ使用時の変数記号として用いる。

一階言語 $\mathcal{L}$、論理式変数記号の集合 $\mathcal{V}$、イベント定数記号の集合 $\mathcal{E}$、および、エージェント定数記号の集合 $\mathcal{A}$ を各 1 つずつ適当に選んで与えておく。ただし $\mathcal{E}, \mathcal{A}$ は有限集合、$\mathcal{V}$ は無限集合とする。以下、$\{p \mid p \in \mathbb{R}, 0 \leq p \leq 1\}$ を $[0,1]$ と書く。

- 一階述語論理の原始論理式は ($\mathscr{TCMATC}$ の) 論理式
- ϕ, ψ が論理式ならば $\phi \vee \psi, \neg\phi$ も論理式
- ϕ が論理式、x が変数記号ならば $\forall x \phi$ も論理式
- $e \in \mathcal{E}$、n が正整数、個々の $i = 1, 2, \cdots, n$ に関して ϕ_i が論理式、$p_i \in [0,1]$、$r_i \in \{\geq, >\}$ ならば $\mathsf{X}^e({}_{r_1 p_1}\,\phi_1 \mid \cdots \mid {}_{r_n p_n}\,\phi_n)$ も論理式。特に $n = 1$ の場合は $\mathsf{X}^e({}_{r_1 p_1}\,\phi_1)$ の略記として $\mathsf{X}^e_{r_1 p_1}\,\phi_1$ を用いる。
- ϕ が論理式、$a \in \mathcal{A}$ ならば $\mathsf{BEL}^a\,\phi, \mathsf{DESIRE}^a\,\phi, \mathsf{INTEND}^a\,\phi$ も論理式
- $\mathfrak{X} \in \mathcal{V}$ ならば $\mathfrak{X}$ は論理式
- ϕ が論理式、$\mathfrak{X} \in \mathcal{V}$ で、ϕ の中で $\mathfrak{X}$ が $\neg$ の奇数段のネストの中に出現しないならば $\mu\mathfrak{X}.\phi$ は論理式。

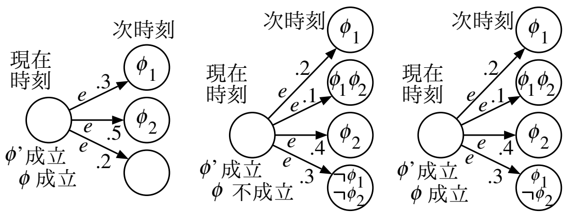
以上が $\mathscr{TCMATC}$ の構文の定義である。この他、$\wedge$ や $\supset$ や $\Leftrightarrow$ や $\exists$ を一般的な略記として導入する。また、必要に応じて括弧で曖昧さを除去し、括弧がない場合の論理オペレータの結合順序は、単項オペレータ $\cdot \wedge \cdot \vee \cdot \supset$ の順に先に結合するものとする。さらに、$\wedge \cdot \vee$ は左結合、$\supset$ は右結合とする。

$\nu\mathfrak{X}.\phi$ は $\neg\mu\mathfrak{X}.(\neg\phi[\mathfrak{X} := \neg\mathfrak{X}])$ の略記とする。μ は最小不動点オペレータと呼ばれるものであり[3)]、ν は最大不動点オペレータである。

また、$\mathsf{X}^e_{<p}\,\phi, \mathsf{X}^e_{\leq p}\,\phi, \mathsf{X}^e_{=p}\,\phi, \mathsf{X}^e_{\neq p}\,\phi$ はそれぞれ、$\neg\,\mathsf{X}^e_{\geq p}\,\phi, \neg\,\mathsf{X}^e_{>p}\,\phi, \mathsf{X}^e_{\geq p}\,\phi \wedge \neg\,\mathsf{X}^e_{>p}\,\phi, \mathsf{X}^e_{>p}\,\phi \vee \neg\,\mathsf{X}^e_{\geq p}\,\phi$ の略記として導入する。さらに、$\mathsf{AX}^e\,\phi$ は $\mathsf{X}^e_{\geq 1}\,\phi$ の略記、$\mathsf{AX}\,\psi$ は $\bigwedge_{e \in \mathcal{E}} \mathsf{AX}^e\,\psi$ の略記とする。

2.1.2 直感的な解釈

X^e オペレータは、CTL の next time オペレータ AX をイベント e に関する確率的状態遷移に拡張したものである。たとえば、$\mathsf{X}^e({}_{\geq .3}\,\phi_1 \mid {}_{\geq .5}\,\phi_2)$ は直感的には「イベント e を実行すると、1 時刻後に 0.3 以上の確率で ϕ_1 が成り立ち、この ϕ_1 が成立する状態とは別な 1 時刻後の状態において 0.5 以上の確率で ϕ_2 が成り立つ」と解釈される。一方、$\mathsf{X}^e_{\geq .3}\,\phi_1 \wedge \mathsf{X}^e_{\geq .5}\,\phi_2$ は直感的には「イベント e を実行すると、1 時刻後に 0.3 以上の確率で ϕ_1 が成り立ち、かつ、0.5 以上の確率で ϕ_2 が成立する」と解釈される。図 1 のように、前者は、論理式 ϕ_1 が成り立つ 1 時刻後の状態と、論理式 ϕ_2 が成り立つ 1 時刻後の状態とをそれぞれ別々に (かつ、遷移確率がそれぞれ 0.3, 0.5 以上となるように) 取れることを要求しているが、後者はそのような制約はない。



$$\phi = \mathsf{X}^e({}_{\geq .3}\,\phi_1 \mid {}_{\geq .5}\,\phi_2), \quad \phi' = \mathsf{X}^e_{\geq .3}\,\phi_1 \wedge \mathsf{X}^e_{\geq .5}\,\phi_2$$

図 1 X^e オペレータの直感的解釈

また、$\mathsf{BEL}^a\,\phi, \mathsf{DESIRE}^a\,\phi, \mathsf{INTEND}^a\,\phi$ は、それぞれ「エージェント a が ϕ という信念/願望/意図を持つ」を表す。

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、一階言語の解釈 (変数割り当てを含む) $i_{w,t}$。すなわち、領域 U、および定数記号・述語記号・関数記号・変数記号の解釈の組。ただし、領域 U は全ての可能世界の全ての state に対して共通であること。
- $\mathcal{A}$ の各要素 a と $\bigcup_{w \in W} St_w$ の各要素 t に対し、集合 $\{w \mid t \in St_w\}$ 上の serial, transitive かつ Euclidean な 2 項関係 $\mathcal{B}^t_a$、および同集合上の serial な 2 項関係 $\mathcal{D}^t_a, \mathcal{I}^t_a$

- W の各要素 w と $\mathcal{E}$ の各要素 e に対して、St_w 上の serial な 2 項関係 $R_w^e \subset St_w \times St_w$ と、R_w^e から $[0,1]$ への関数 $\mathcal{P}_w^e$。ただし、任意の $t \in St_w$ および任意の $e \in \mathcal{E}$ に対し、$\sum_{t' \in \{t' \mid t\, R_w^e\, t'\}} \mathcal{P}_w^e(t,t') = 1$ であること

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」に相当し、1 つの可能世界は時刻の木である。R_w^e は可能世界 w 内の時刻の前後関係で、$t\ R_w^e\ t'$ および $\mathcal{P}_w^e(t,t') = p$ 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。$\mathcal{B}_a^t, \mathcal{D}_a^t, \mathcal{I}_a^t$ 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 の信念･願望･意図を表す (図 2 に概略を示した)。

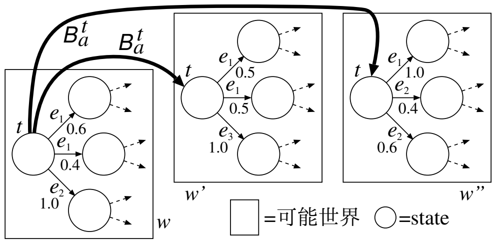


図 2　BDI ストラクチャの概略

R_w^e は serial としたため、どの state でもどのイベントも実行可能なことになる。しかし、実際には特定の state で実行可能なイベントの集合は決まっているのが普通である。この性質は、いわゆる「死状態」を設け、実行可能なイベント以外ではその状態にのみ遷移し、その状態では特別な原始論理式 *dead* が成り立つ、という扱いにすれば表現はできる。例えば「イベント e が実行できれば、e の実行直後は ϕ である」は $\neg \mathsf{AX}^e\, dead \supset \mathsf{AX}^e\, \phi$ で表現できる。

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 $\mathcal{V}$ から 2^{Swt} への関数 $f_{\mathcal{V}}$ を 1 つ定めておく。

論理式 ϕ に対し、その解釈 $[\![\phi]\!]_{\langle M, f_{\mathcal{V}}\rangle}$ を以下のように定める ($[\![\phi]\!]_{\langle M, f_{\mathcal{V}}\rangle} \subset \mathit{Swt}$ である)。

- ϕ が原始論理式のとき、$[\![\phi]\!]_{\langle M, f_{\mathcal{V}}\rangle} = \{(w,t) \mid M$ での $i_{w,t}$ で ϕ が真$\}$
- $[\![\phi \vee \psi]\!]_{\langle M, f_{\mathcal{V}}\rangle} = [\![\phi]\!]_{\langle M, f_{\mathcal{V}}\rangle} \cup [\![\psi]\!]_{\langle M, f_{\mathcal{V}}\rangle}$
- $[\![\neg\phi]\!]_{\langle M, f_{\mathcal{V}}\rangle} = \mathit{Swt} \setminus [\![\phi]\!]_{\langle M, f_{\mathcal{V}}\rangle}$
- $[\![\forall x \phi]\!]_{\langle M, f_{\mathcal{V}}\rangle} = \bigcap_{u \in U} [\![\phi]\!]_{\langle M^u, f_{\mathcal{V}}\rangle}$　ここで、M^u は M での x の解釈を u に変更して得られる BDI ストラクチャとする
- $[\![\mathsf{X}^e({}_{r_1 p_1}\,\phi_1 \mid \cdots \mid {}_{r_n p_n}\,\phi_n)]\!]_{\langle M, f_{\mathcal{V}}\rangle} = \{(w,t) \mid i = 1, \cdots, n$ のそれぞれについて $T_i \subset \{t' \mid (w,t') \in [\![\phi_i]\!]_{\langle M, f_{\mathcal{V}}\rangle}\}$ かつ $(\sum_{t' \in T_i} \mathcal{P}_w^e(t,t'))\ r_i\ p_i$ を満たすような、$\{t' \mid t\ R_w^e\ t'\}$ の互いに素な部分集合 $T_1, \cdots, T_n$ が存在する$\}$ (r_i は $>, \geq$ のいずれかであることに注意)
- $[\![\mathsf{BEL}^a\, \phi]\!]_{\langle M, f_{\mathcal{V}}\rangle} = \{(w,t) \mid w\ \mathcal{B}_a^t\ w'$ なる任意の世界 w' に対して $(w',t) \in [\![\phi]\!]_{\langle M, f_{\mathcal{V}}\rangle}\}$
- $[\![\mathsf{DESIRE}^a\, \phi]\!]_{\langle M, f_{\mathcal{V}}\rangle}, [\![\mathsf{INTEND}^a\, \phi]\!]_{\langle M, f_{\mathcal{V}}\rangle}$ 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 ϕ を、$\mathfrak{X}$ の解釈を受け取って ϕ の解釈を返す関数 $f_\phi : \mathit{Swt} \to \mathit{Swt}$ 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_ϕ の最小不動点である

と定義する。この場合、定義から f_ϕ は単調関数となるので、最小不動点の存在は保証される[6]。

$[\![\phi]\!]_{\langle M, f_{\mathcal{V}}\rangle}$ は、ϕ 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 ϕ が成り立つという。論理式 ϕ が、任意の $M, \mathcal{V}$ に対し $[\![\phi]\!]_{\langle M, f_{\mathcal{V}}\rangle} = \mathit{Swt}$ を満たすとき、ϕ は恒真であるという。

2.3 演 繹 体 系

本節では $\mathcal{TOMATO}$ のシーケント計算による演繹体系を与える。なお、文中ではシーケントの範囲を明示するために、シーケント全体を [ ] でくくることがある。

以降、Σ, Δ などギリシャ文字の大文字 (「$'$」付きの Σ' なども含む) は論理式 0 個以上の並びとする。ただし Θ のみ、論理式 0 個または 1 個を表すとする。また、論理式の並び Γ と単項論理オペレータ K に対し、$K\,\Gamma$ は Γ の各論理式に K を前置して得られる論理式の並びを表す。

本体系では、α 同値な論理式を同一視する。また、シーケントの「$\to$」の左右は論理式の multi set とする (従って Exchange の規則がない)。

シーケント $[\Sigma \to \Delta]$ の解釈は、論理式 $\bigwedge \Sigma \supset \bigvee \Delta$ の解釈と定義する (ただし $\bigwedge \emptyset = \mathit{true}, \bigvee \emptyset = \mathit{false}$ とする。ここで true は適当なトートロジーの、false は $\neg \mathit{true}$ の略記)。

2.3.1 推 論 規 則

推論規則は図 3 に挙げたものである。ただし「$\to$」の左の X^e オペレータに関する推論規則は図 3 には示していない。これについては別途 2.3.2 節で述べる。また、$\forall$ 左の t は任意

の項、∀ 右の y は結論に現れない新しい変数記号とする。これらの規則は、命題論理に制限する場合は使われない。

推論規則 $\mathsf{X}_{\mathrm{excl}}$ は、「前提に書かれている形の部分論理式がシーケント中に現れれば、その部分を結論に書かれている式に置き換えてよい」ことを表す。ただし $n \geq 2$ で、各 i に対し ψ_i は $\neg\mathfrak{X}_1 \wedge \cdots \wedge \neg\mathfrak{X}_{i-1} \wedge \mathfrak{X}_i \wedge \neg\mathfrak{X}_{i+1} \wedge \cdots \wedge \neg\mathfrak{X}_n$ の形 (ここで $\mathfrak{X}_1, \cdots, \mathfrak{X}_n$ は規則の結論に自由に出現しない論理式変数記号) である。

2.3.2 X^e オペレータに関する推論規則

図 3 の他に、X^e に関する以下の推論規則がある。

$\Gamma = \{\mathsf{X}^e_{r_1\,p_1}\,\psi_1, \cdots, \mathsf{X}^e_{r_n\,p_n}\,\psi_n\}$ とし (ただし各 $r_1, \cdots, r_n$ は $\geq$ または $>$)、$\mathfrak{Q} = \{\psi_1, \cdots, \psi_n\}$ とする。

関数 $v : 2^{\mathfrak{Q}} \to [0,1]$ が性質 $\sum_{Q \subset \mathfrak{Q}} v(Q) = 1$ を満たし、かつ、$1 \leq i \leq n$ なる全ての整数 i について $(\sum_{Q \in \{T | T \subset \mathfrak{Q}, \psi_i \in T\}} v(Q))\, r_i\, p_i$ を満たす (r_i は $\geq$ または $>$ であることに注意) とき、v を Γ の確率分配関数と呼ぶ。大まかには、確率分配関数とは、$\mathfrak{Q}$ の各要素 ψ_i が次の時刻に成り立つ確率が $r_i\, p_i$ を満たすことを保証するように、各 $Q \subset \mathfrak{Q}$ の成り立つ state への遷移確率を決めるものである。Γ の確率分配関数 v に対し、$\{Q \subset \mathfrak{Q} \mid v(Q) > 0\}$ を v による Γ の要充足集合と呼び、$\mathrm{req}_v(\Gamma)$ と書く。Z が (ある v による) Γ の要充足集合であって、Z のどの要素も充足可能である場合、Z は充足可能であるという。Γ が充足可能であることと、Γ の充足可能な要充足集合が存在することは同値である。

要充足集合 $Z, Z' \subset 2^{\mathfrak{Q}}$ に対し、ある $Q \in Z$ と $Z'' \subset 2^{Q}$ があって $Z' = (Z \cup Z'') \setminus \{Q\}$ を満たす場合、$Z \succ Z'$ と書く。Z が充足可能なら Z' も充足可能である (Q も、従って Z'' のどの要素も充足可能だからである)。

Z が Γ の要充足集合であり、$Z \succ Z'$ を満たす Γ の要充足集合 Z' が存在しない場合、Z は本質的であるという。

Γ の本質的な要充足集合全ての列挙を $Z_1 = \{Q_{1,1}, \cdots, Q_{1,m_1}\}, \cdots, Z_k = \{Q_{k,1}, \cdots, Q_{k,m_k}\}$ とする。このとき、任意の整数列 $j_1, \cdots, j_k$ (ただし $1 \leq j_1 \leq m_1, \cdots, 1 \leq j_k \leq m_k$) について、以下は $\mathcal{TOMATO}$ の推論規則である。

$$\frac{Q_{1,j_1} \to \quad \cdots \quad Q_{k,j_k} \to}{\Gamma \to}\ \mathsf{X\text{-}KD}$$

以降、この規則について例による直感的説明を行う。

例えば $\Gamma = \{\mathsf{X}^e_{\geq 0.3}\,\psi_1, \mathsf{X}^e_{\geq 0.4}\,\psi_2, \mathsf{X}^e_{\geq 0.6}\,\psi_3\}, \mathfrak{Q} = \{\psi_1, \psi_2, \psi_3\}$ のとき、$v(\{\psi_1, \psi_2\}) = 0.4, v(\{\psi_3\}) = 0.6,$ その他の $Q \subset \mathfrak{Q}$ に対しては $v(Q) = 0,$ で与えられる v は Γ の確率分配関数の 1 つであり、$\mathrm{req}_v(\Gamma) = \{\{\psi_1, \psi_2\}, \{\psi_3\}\} = Z_1$ は Γ の要充足集合の 1 つである。このとき、Z_1 が充足可能であるならば、Γ も充足可能であることに注意 (図 4 の Z_1 参照)。

また、$v''(\{\psi_1, \psi_2\}) = 0.3, v''(\{\psi_2\}) = 0.1, v''(\{\psi_2, \psi_3\}) = 0.6,$ その他の $Q \subset \mathfrak{Q}$ に対しては $v''(Q) = 0,$ で与えられる v'' による $\mathrm{req}_{v''}(\Gamma) = Z''_1$ や、$v'(\{\psi_1, \psi_2\}) = 0.3, v'(\{\psi_2\}) = 0.1, v'(\{\psi_3\}) = 0.6,$ その他の $Q \subset \mathfrak{Q}$ に対しては $v'(Q) = 0,$ で与えられる v' による $\mathrm{req}_{v'}(\Gamma) = Z'_1$ もともに Γ の要充足集合であるが、いずれも本質的ではない。$Z''_1 \succ Z'_1 \succ Z_1$ であるためである。これに対し Z_1 は本質的である。

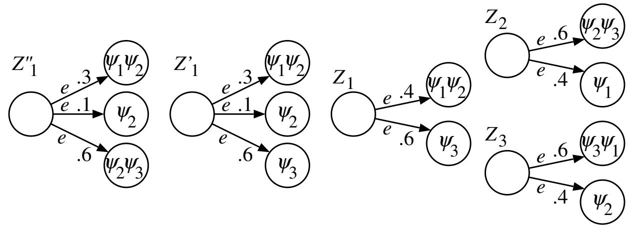


図 4 Γ の要充足集合の事例

さて、規則 X-KD は、タブロー法によるシーケントの証明の構築を睨んで設けられている。タブロー法でシーケント $[\Gamma \to]$ が証明できることを示すには、Γ が充足可能でないことを示すことになる。よって、Γ の充足可能な要充足集合が存在しないことを示せばよい。

ここで、Γ の充足可能な要充足集合が存在するならば、本質的で充足可能な要充足集合が存在することに注意。なぜなら、$Z_1 \succ Z_2 \succ \cdots$ を満たすどのような $Z_1, Z_2, \cdots$ も無限列とはならない[*1]ため、充足可能な要充足集合が存在するならば、そこから出発して、$\succ$ の連鎖で必ず、本質的で充足可能な要充足集合に到達できるからである。

以上の対偶から、「Γ の本質的で充足可能な要充足集合が存在しない」ことを示せばよい。さらにこの条件は「Γ の本質的な要充足集合のどれにも、充足可能でない要素が 1 つ以上存在する」と同値である。従って、Γ の本質的な要充足集合のどれにも、$[Q \to]$ が証明可能であるような要素 Q があることを示せば、$[\Gamma \to]$ が証明できたとしてよい。規則 X-KD はこのように作られている。

⋆1 $f(Z) = \sum_{Q \in Z} 2^{|Q| \cdot 2^{|\mathfrak{Q}|}}$ と定義すると、$Z \succ Z'$ ならば $f(Z) > f(Z')$ であることによる ($\mathfrak{Q}$ が有限であることを使っている)。

$$\frac{}{\phi \to \phi}\,\text{Initial} \quad \frac{\Sigma \to \Delta}{\Sigma, \Sigma' \to \Delta, \Delta'}\,\text{Weak} \quad \frac{\Sigma, \phi, \phi \to \Delta}{\Sigma, \phi \to \Delta}\,\text{重ね左} \quad \frac{\Sigma \to \Delta, \phi, \phi}{\Sigma \to \Delta, \phi}\,\text{重ね右} \quad \frac{\Sigma \to \Delta, \phi}{\Sigma, \neg\phi \to \Delta}\,\neg\text{左} \quad \frac{\Sigma, \phi \to \Delta}{\Sigma \to \Delta, \neg\phi}\,\neg\text{右} \quad \frac{\Sigma, \phi \to \Delta \quad \Sigma, \psi \to \Delta}{\Sigma, \phi \vee \psi \to \Delta}\,\vee\text{左} \quad \frac{\Sigma \to \Delta, \phi, \psi}{\Sigma \to \Delta, \phi \vee \psi}\,\vee\text{右}$$

$$\frac{\Sigma, \phi[x := t] \to \Delta}{\Sigma, \forall x\phi \to \Delta}\,\forall\text{左} \quad \frac{\Sigma \to \Delta, \phi[x := y]}{\Sigma \to \Delta, \forall x\phi}\,\forall\text{右} \quad \frac{\Gamma, \phi[\mathfrak{X} := \mu\mathfrak{X}.\phi] \to \Delta}{\Gamma, \mu\mathfrak{X}.\phi \to \Delta}\,\mu\text{左} \quad \frac{\Gamma \to \Delta, \phi[\mathfrak{X} := \mu\mathfrak{X}.\phi]}{\Gamma \to \Delta, \mu\mathfrak{X}.\phi}\,\mu\text{右} \quad \frac{\Gamma, \mathsf{X}^e_{>1-p}\,\neg\phi \to \Delta}{\Gamma \to \Delta, \mathsf{X}^e_{\geq p}\,\phi}\,\mathsf{X}_{\geq}\text{右} \quad \frac{\Gamma, \mathsf{X}^e_{\geq 1-p}\,\neg\phi \to \Delta}{\Gamma \to \Delta, \mathsf{X}^e_{>p}\,\phi}\,\mathsf{X}_{>}\text{右}$$

$$\frac{\Gamma, \mathsf{BEL}^a\,\Gamma \to \mathsf{BEL}^a\,\Delta, \Theta, \mathsf{BEL}^a\,\Theta}{\mathsf{BEL}^a\,\Gamma \to \mathsf{BEL}^a\,\Delta, \mathsf{BEL}^a\,\Theta}\,\text{BEL-KD45} \quad \frac{\Gamma \to \Theta}{\mathsf{DESIRE}^a\,\Gamma \to \mathsf{DESIRE}^a\,\Theta}\,\text{DESIRE-KD} \quad \frac{\Gamma \to \Theta}{\mathsf{INTEND}^a\,\Gamma \to \mathsf{INTEND}^a\,\Theta}\,\text{INTEND-KD} \quad \frac{\cdots\ \mathsf{X}^e_{r_1\,p_1}(\phi_1 \wedge \psi_1) \wedge \cdots \wedge \mathsf{X}^e_{r_n\,p_n}(\phi_n \wedge \psi_n)\ \cdots}{\cdots\ \mathsf{X}^e({}_{r_1 p_1}\,\phi_1 \mid \cdots \mid {}_{r_n p_n}\,\phi_n)\ \cdots}\,\mathsf{X}_{\text{excl}}$$

図 3　$\mathcal{TOMATO}$ の推論規則 (「→」の左の X^e に関する規則を除く)

先の例の場合、本質的な要充足集合全てを列挙すると、先の Z_1 の他に $Z_2 = \{\{\psi_2, \psi_3\}, \{\psi_1\}\}, Z_3 = \{\{\psi_3, \psi_1\}, \{\psi_2\}\}$ がある。よって、推論規則としては

$$\frac{\psi_1, \psi_2 \to \quad \psi_2, \psi_3 \to \quad \psi_2 \to}{\mathsf{X}^e_{\geq .3}\,\psi_1, \mathsf{X}^e_{\geq .4}\,\psi_2, \mathsf{X}^e_{\geq .6}\,\psi_3 \to}$$

など 8 つが作れる。しかし、上に例示した規則の前提中 ψ_2 に注目すると、その左 2 つは冗長で取り除ける。同様に他の規則からも冗長性を取り除くと、結局図 5 の 4 規則のみ残り、他の 4 規則はそれらより前提が多いため冗長で不要となる。

$$\frac{\psi_1, \psi_2 \to \quad \psi_2, \psi_3 \to \quad \psi_3, \psi_1 \to}{\mathsf{X}^e_{\geq .3}\,\psi_1, \mathsf{X}^e_{\geq .4}\,\psi_2, \mathsf{X}^e_{\geq .6}\,\psi_3 \to} \qquad \frac{\psi_1 \to}{\mathsf{X}^e_{\geq .3}\,\psi_1, \mathsf{X}^e_{\geq .4}\,\psi_2, \mathsf{X}^e_{\geq .6}\,\psi_3 \to}$$

$$\frac{\psi_2 \to}{\mathsf{X}^e_{\geq .3}\,\psi_1, \mathsf{X}^e_{\geq .4}\,\psi_2, \mathsf{X}^e_{\geq .6}\,\psi_3 \to} \qquad \frac{\psi_3 \to}{\mathsf{X}^e_{\geq .3}\,\psi_1, \mathsf{X}^e_{\geq .4}\,\psi_2, \mathsf{X}^e_{\geq .6}\,\psi_3 \to}$$

図 5　X^e オペレータに関する推論規則の例

なお、タブロー法で規則 X-KD を (逆向きに) 適用するには、X^e オペレータを全て「→」の左に移す必要がある。規則 $\mathsf{X}_{\geq}$ 右と $\mathsf{X}_{>}$ 右はそのために設けられた規則である。また、$\mathsf{X}^e({}_{r_1 p_1}\,\phi_1 \mid \cdots \mid {}_{r_n p_n}\,\phi_n)$ の形の論理式を、全て $n = 1$ の形に分解する必要もあり、X_{excl} はそのための規則である。

2.3.3　証明可能性の定義

以下のいずれかが成り立つとき、シーケント S はシーケントの集合 L から導出可能であるという。

(1)　$S \in L$

(2)　推論規則 $\dfrac{S_1, \cdots, S_n}{S}(n \geq 0)$ が存在し、$S_1, \cdots, S_n$ が全て証明可能または L から導出可能

また、以下のいずれかが成り立つとき、シーケント S は証明可能であるという。ただしここで $\phi^n(\mathfrak{X})$ は以下のように定義される: $\phi^0(\mathfrak{X}) = \mathfrak{X}, \phi^n(\mathfrak{X}) = \phi[\mathfrak{X} := \phi^{n-1}(\mathfrak{X})]$.

(1)　S は $\emptyset$ から導出可能

(2)　$S = [\Sigma, \mu\mathfrak{X}.\phi \to \Delta]$ (ただし $\mathfrak{X}$ は Σ, Δ に自由に出現しない) であり、かつある正整数 n が存在して、$[\Sigma, \phi^n(\mathfrak{X}) \to \Delta]$ が $\{[\Sigma, \mathfrak{X} \to \Delta]\}$ から導出可能

シーケント $[\to \phi]$ が証明可能であるとき、論理式 ϕ は証明可能であるという。

この演繹体系の健全性と、(少なくとも命題論理に制限しての) 完全性を示すことができる[4)]。ただし本論文では、それらについては割愛する。

3.　確率的状態遷移のモデル化および証明の例

3.1　確率的状態遷移に関する基本的な証明例

強化学習タスクのモデル化の基本である MDP (マルコフ決定過程) の典型的な事例として「状態 s_1 にいて行動 e_1 をとると、確率 0.7 で状態 s_2 へ動いて報酬 3 を受け取り、確率 0.3 で状態 s_3 へ動いて報酬 5 を受け取る」という状況を考えよう。これは、元来の BDI logic では単に「どちらかになる」というレベルでしか記述できなかった。$\mathcal{TOMATO}$ の確率的遷移オペレータを用いると、$\mathit{at}(s_1) \supset \mathsf{X}^{e_1}({}_{\geq .7}\,\mathit{at}(s_2) \wedge \mathit{reward}(3) \mid {}_{\geq .3}\,\mathit{at}(s_3) \wedge \mathit{reward}(5))$ と書ける。この論理式を ϕ とする。いま、$3 \geq 3$ および $5 \geq 3$ が何らかの方法で証明可能であると仮定しよう。すると我々は、$\phi \wedge \mathit{at}(s_1) \supset \mathsf{AX}^{e_1}\,\exists x(\mathit{reward}(x) \wedge x \geq 3)$ を証明することによって、「ϕ が成り立つならば、状態 s_1 においてイベント e_1 を実行すると 3 以上の報酬を受けることができる」を示せる。この証明は図 6 に示す。この中の X-KD ルールの適用において、$\{\mathsf{X}^{e_1}_{\geq .7}\,\xi_1, \mathsf{X}^{e_1}_{\geq .3}\,\xi_2, \mathsf{X}^{e_1}_{>0}\,\xi_3\}$ の本質的な要充足集合は $\{\{\xi_1, \xi_2\}, \{\xi_3\}\}$、$\{\{\xi_2, \xi_3\}, \{\xi_1\}\}$、および、$\{\{\xi_3, \xi_1\}, \{\xi_2\}\}$ である (ただし ξ_1, ξ_2, ξ_3 は任意の論理式) ことを使っている。また、$3 \geq 3$ および $5 \geq 3$ の実際の証明は、自然数に関するペアノの公理を用いるなどして可能であるものとする。

別な例として、いま、イベントは e_1, e_2, e_3 の 3 種類だけであり ($\mathcal{E} = \{e_1, e_2, e_3\}$)、現在、

Vol.2010-MPS-81 No.23
Vol.2010-BIO-23 No.23
2010/12/17

$$
\begin{prooftree}
\AxiomC{}
\RightLabel{Initial}
\UnaryInfC{$p_1 \wedge q_1 \wedge \mathfrak{X}_1, p_2 \wedge q_2 \wedge \neg\mathfrak{X}_1, \mathfrak{X}_2 \to \mathfrak{X}_2$}
\RightLabel{$\neg$左}
\UnaryInfC{$p_1 \wedge q_1 \wedge \mathfrak{X}_1, \neg\mathfrak{X}_2, p_2 \wedge q_2 \wedge \neg\mathfrak{X}_1, \mathfrak{X}_2 \to$}
\RightLabel{$\wedge$左}
\UnaryInfC{$p_1 \wedge q_1 \wedge \mathfrak{X}_1, \neg\mathfrak{X}_2, \zeta_1 \to$}
\RightLabel{$\wedge$左}
\UnaryInfC{$\zeta_0, \zeta_1 \to$}
\AxiomC{}
\RightLabel{Initial}
\UnaryInfC{$p_2, q_2, \neg\mathfrak{X}_1, \mathfrak{X}_2 \to q_2$}
\RightLabel{$\wedge$左}
\UnaryInfC{$p_2 \wedge q_2, \neg\mathfrak{X}_1, \mathfrak{X}_2 \to q_2$}
\RightLabel{$\wedge$左}
\UnaryInfC{$p_2 \wedge q_2 \wedge \neg\mathfrak{X}_1, \mathfrak{X}_2 \to q_2$}
\RightLabel{$\wedge$左}
\UnaryInfC{$\zeta_1 \to q_2$}
\AxiomC{$\vdots$}
\noLine
\UnaryInfC{$\to 5 \geq 3$}
\RightLabel{Weak}
\UnaryInfC{$\zeta_1 \to 5 \geq 3$}
\RightLabel{$\wedge$右}
\BinaryInfC{$\zeta_1 \to q_2 \wedge 5 \geq 3$}
\RightLabel{$\exists$右}
\UnaryInfC{$\zeta_1 \to \exists x\psi$}
\RightLabel{$\neg$左}
\UnaryInfC{$\zeta_1, \neg\exists x\psi \to$}
\AxiomC{}
\RightLabel{Initial}
\UnaryInfC{$p_1, q_1, \mathfrak{X}_1, \neg\mathfrak{X}_2 \to q_1$}
\RightLabel{$\wedge$左}
\UnaryInfC{$p_1 \wedge q_1, \mathfrak{X}_1, \neg\mathfrak{X}_2 \to q_1$}
\RightLabel{$\wedge$左}
\UnaryInfC{$p_1 \wedge q_1 \wedge \mathfrak{X}_1, \neg\mathfrak{X}_2 \to q_1$}
\RightLabel{$\wedge$左}
\UnaryInfC{$\zeta_0 \to q_1$}
\AxiomC{$\vdots$}
\noLine
\UnaryInfC{$\to 3 \geq 3$}
\RightLabel{Weak}
\UnaryInfC{$\zeta_0 \to 3 \geq 3$}
\RightLabel{$\wedge$右}
\BinaryInfC{$\zeta_0 \to q_1 \wedge 3 \geq 3$}
\RightLabel{$\exists$右}
\UnaryInfC{$\zeta_0 \to \exists x\psi$}
\RightLabel{$\neg$左}
\UnaryInfC{$\zeta_0, \neg\exists x\psi \to$}
\RightLabel{X-KD}
\TrinaryInfC{$\mathsf{X}^e_{\geq .7}\,\zeta_0, \mathsf{X}^e_{\geq .3}\,\zeta_1, \mathsf{X}^e_{>0}\,\neg\exists x\psi \to$}
\RightLabel{Weak}
\UnaryInfC{$\mathsf{X}^e_{\geq .7}\,\zeta_0, \mathsf{X}^e_{\geq .3}\,\zeta_1, \mathsf{X}^e_{>0}\,\neg\exists x\psi, \mathit{at}(s_1) \to$}
\RightLabel{$\mathsf{X}_{\geq}$右}
\UnaryInfC{$\mathsf{X}^e_{\geq .7}\,\zeta_0, \mathsf{X}^e_{\geq .3}\,\zeta_1, \mathit{at}(s_1) \to \mathsf{AX}^{e_1}\,\exists x\psi$}
\RightLabel{$\wedge$左}
\UnaryInfC{$\mathsf{X}^e_{\geq .7}\,\zeta_0 \wedge \mathsf{X}^e_{\geq .3}\,\zeta_1, \mathit{at}(s_1) \to \mathsf{AX}^{e_1}\,\exists x\psi$}
\RightLabel{$\mathsf{X}_{\mathrm{excl}}$}
\UnaryInfC{$\phi', \mathit{at}(s_1) \to \mathsf{AX}^{e_1}\,\exists x\psi$}
\AxiomC{}
\RightLabel{Initial}
\UnaryInfC{$\mathit{at}(s_1) \to \mathit{at}(s_1), \mathsf{AX}^{e_1}\,\exists x\psi$}
\RightLabel{$\supset$左}
\BinaryInfC{$\phi, \mathit{at}(s_1) \to \mathsf{AX}^{e_1}\,\exists x\psi$}
\RightLabel{$\wedge$左}
\UnaryInfC{$\phi \wedge \mathit{at}(s_1) \to \mathsf{AX}^{e_1}\,\exists x\psi$}
\RightLabel{$\supset$右}
\UnaryInfC{$\to \phi \wedge \mathit{at}(s_1) \supset \mathsf{AX}^{e_1}\,\exists x\psi$}
\end{prooftree}
$$

ここで $\phi = \mathit{at}(s_1) \supset \phi', \phi' = \mathsf{X}^{e_1}(_{\geq .7}\, p_1 \wedge q_1 \mid _{\geq .3}\, p_2 \wedge q_2), p_1 = \mathit{at}(s_2), p_2 = \mathit{at}(s_3), q_1 = \mathit{reward}(3), q_2 = \mathit{reward}(5)$,
$\psi = \mathit{reward}(x) \wedge x \geq 3, \zeta_0 = p_1 \wedge q_1 \wedge \mathfrak{X}_1 \wedge \neg\mathfrak{X}_2, \zeta_1 = p_2 \wedge q_2 \wedge \neg\mathfrak{X}_1 \wedge \mathfrak{X}_2$. また、以下の派生規則を用いた:

$$
\frac{\Sigma, \phi, \psi \to \Delta}{\Sigma, \phi \wedge \psi \to \Delta}\ \wedge\text{左} \qquad
\frac{\Sigma \to \Delta, \phi \quad \Sigma \to \Delta, \psi}{\Sigma \to \Delta, \phi \wedge \psi}\ \wedge\text{右} \qquad
\frac{\Sigma \to \Delta, \phi \quad \Sigma, \psi \to \Delta}{\Sigma, \phi \supset \psi \to \Delta}\ \supset\text{左} \qquad
\frac{\Sigma, \phi \to \Delta, \psi}{\Sigma \to \Delta, \phi \supset \psi}\ \supset\text{右} \qquad
\frac{\Sigma \to \Delta, \phi[x := t]}{\Sigma \to \Delta, \exists x\phi}\ \exists\text{右}
$$

図 **6** 証明図の例 1

e_1 を実行すると、次時刻に e_3 を実行したとき ϕ が成り立つ確率が 0.4 以上、また、e_2 を実行すると、次時刻に e_3 を実行したとき ϕ が成り立つ確率が 0.2 以上であり、さらに e_3 は現在実行不可能 (実行すると死状態 (2.2.1 節参照) へ行く) とする。この場合、可能なイベントのどれを実行しても、次時刻に e_3 を実行したとき ϕ が成り立つ確率が 0.2 以上であることは、論理式 $\mathsf{AX}^{e_1}\,\mathsf{X}^{e_3}_{\geq .4}\,\phi \wedge \mathsf{AX}^{e_2}\,\mathsf{X}^{e_3}_{\geq .2}\,\phi \wedge \mathsf{AX}^{e_3}\,\mathit{dead} \supset \mathsf{AX}(\neg\mathit{dead} \supset \mathsf{X}^{e_3}_{\geq .2}\,\phi)$ の証明により示せる。その概要を図 7 に示す。なお、図中の「古典」は古典論理オペレータに対する規則をまとめて適用したことを表す。

4. 強化学習との結合のモデル化

本節ではもう少し進めて、強化学習と BDI エージェントの結合において、強化学習の動作部分を、BDI モデルにおける推論による動作記述の中に取り込む例を示す。

ここでは、あらかじめ与えた nonlogical axiom からの推論という形で、エージェントの動作を記述する。nonlogical axiom の部分は、エージェントに与えたプログラムとも解釈できるし、ここに Prolog の組み込み述語のような仕組みを想定することも可能である。この方式によれば、論理型言語のような実行方式で、論理式で書いた仕様をそのままプログラムとして、その宣言的意味と一致するようにエージェントを動作させることに道が開ける。

4.1 準備—BC 規則の導入

いま、$\mathsf{AG}\,\phi$ を $\nu\mathfrak{X}.(\phi \wedge \mathsf{AX}\,\mathfrak{X})$ の略記とする (「ϕ が現在から未来永劫成り立つ」の意味になる)。このとき図 8 より、以下の派生規則が得られる (ここで $\zeta = \phi_1 \wedge \cdots \wedge \phi_n \supset \psi$、ただし $n \geq 0$ とする)。

$$
\frac{\Sigma, \mathsf{AG}\,\zeta \to \phi_1, \Delta \quad \cdots \quad \Sigma, \mathsf{AG}\,\zeta \to \phi_n, \Delta}{\Sigma, \mathsf{AG}\,\zeta \to \psi, \Delta}\ \text{Backchain (BC)}
$$

$$
\dfrac{
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{\phi, \neg\phi \to}\ \text{古典}+\text{Initial}}{\mathsf{X}^{e_3}_{\geq .4}\,\phi, \mathsf{X}^{e_3}_{>.8}\,\neg\phi \to}\ \text{X-KD}}{\mathsf{X}^{e_3}_{\geq .4}\,\phi \to dead, \mathsf{X}^{e_3}_{\geq .2}\,\phi}\ \text{Weak}+\mathsf{X}_{\geq}\text{右}}{\mathsf{X}^{e_3}_{\geq .4}\,\phi, \neg\xi \to}\ \text{古典}}{\dfrac{\psi_1, \mathsf{X}^{e_1}_{>0}\,\neg\xi \to}{\psi_1, \psi_2, \psi_3 \to \mathsf{AX}^{e_1}\,\xi}\ \text{Weak}+\mathsf{X}_{\geq}\text{右}}\ \text{X-KD}
\qquad
\dfrac{\dfrac{\dfrac{}{\mathsf{X}^{e_3}_{\geq .2}\,\phi, \neg\xi \to}\ \text{古典}+\text{Initial}}{\psi_2, \mathsf{X}^{e_2}_{>0}\,\neg\xi \to}\ \text{X-KD}}{\psi_1, \psi_2, \psi_3 \to \mathsf{AX}^{e_2}\,\xi}\ \text{Weak}+\mathsf{X}_{\geq}\text{右}
\qquad
\dfrac{\dfrac{\dfrac{}{dead, \neg\xi \to}\ \text{古典}+\text{Initial}}{\psi_3, \mathsf{X}^{e_3}_{>0}\,\neg\xi \to}\ \text{X-KD}}{\psi_1, \psi_2, \psi_3 \to \mathsf{AX}^{e_3}\,\xi}\ \text{Weak}+\mathsf{X}_{\geq}\text{右}
}{\to \psi_1 \wedge \psi_2 \wedge \psi_3 \supset \mathsf{AX}\,\xi}\ \text{古典}
$$

ここで $\psi_1 = \mathsf{AX}^{e_1}\,\mathsf{X}^{e_3}_{\geq .4}\,\phi, \psi_2 = \mathsf{AX}^{e_2}\,\mathsf{X}^{e_3}_{\geq .2}\,\phi, \psi_3 = \mathsf{AX}^{e_3}\,dead, \xi = \neg dead \supset \mathsf{X}^{e_3}_{\geq .2}\,\phi$.

図 7　証明図の例 2

これは、あるエージェントが常に成り立つ nonlogical axiom ζ を持っている場合、現在 ψ が成り立つことを結論するためには、$\phi_1 \sim \phi_n$ が成り立つことを結論できればよいものと読むことができ、論理プログラムにおける導出原理 (後向き推論) に相当する使い方ができる (そのためここではこの規則を Backchain (BC) と呼ぶ)。特に、ζ が $\mathsf{BEL}^a\,\phi_1 \wedge \cdots \wedge \mathsf{BEL}^a\,\phi_n \supset \mathsf{BEL}^a\,\psi$ の形をしている場合、エージェント a は信念 $\mathsf{BEL}^a\,\phi_1 \sim \mathsf{BEL}^a\,\phi_n$ から信念 $\mathsf{BEL}^a\,\psi$ を得ることができる。ここで、$\mathsf{BEL}^a(\phi_1 \wedge \cdots \wedge \phi_n \supset \psi)$ (この式を ζ' とする) は ζ より強い ($\zeta' \supset \zeta$ が証明できる) ため、ζ' を axiom として持っている場合も同様に議論できる。(ただし本論文では簡単のため、以下では一部を除き BEL を付けずに議論する。現実のエージェントでは、推論の材料としては専ら BEL を使うことになろう。)

$$
\dfrac{\dfrac{\dfrac{}{\Sigma, \mathsf{AG}\,\zeta, \psi, \mathsf{AX}\,\mathsf{AG}\,\zeta \to \psi, \Delta}\ \text{Initial} \qquad \dfrac{\Sigma, \mathsf{AG}\,\zeta \to \phi_1, \Delta \quad \cdots \quad \Sigma, \mathsf{AG}\,\zeta \to \phi_n, \Delta}{\Sigma, \mathsf{AG}\,\zeta, \neg(\phi_1 \wedge \cdots \wedge \phi_n), \mathsf{AX}\,\mathsf{AG}\,\zeta \to \psi, \Delta}\ \text{古典}+\text{Weak}}{\Sigma, \mathsf{AG}\,\zeta, \neg(\phi_1 \wedge \cdots \wedge \phi_n) \vee \psi, \mathsf{AX}\,\mathsf{AG}\,\zeta \to \psi, \Delta}\ \vee\text{左}}{\Sigma, \mathsf{AG}\,\zeta \to \psi, \Delta}\ \text{重ね左}+\text{古典}+\mu\text{右}
$$

ここで $\zeta = \phi_1 \wedge \cdots \wedge \phi_n \supset \psi$ とする

図 8　Backchain ルールの派生

すると、エージェントを論理式で表現されたプログラムに従って動作する論理プログラミングシステムのように捉えることができる。特に、数値演算など、Prolog での組み込み述語に相当するような仕組みも、あらかじめ十分なだけ axiom として与えてあるとして論じることができる。例えばエージェント a の内部で、x の現在の値が 10 であることが既に得られており、かつ $10 > 5$ と $x > 5 \supset \mathsf{INTEND}^a\,alarm$ が axiom にあれば、$\mathsf{INTEND}^a\,alarm$ を結論することができ、エージェントは警報を発する意図を持つことができる。

4.2　グリーディ方策のモデル化

いま、状態の有限集合 $s_1, \cdots, s_k$ があって、エージェント a は常にこのいずれかの状態にあるものとし、状態 s_i にあることを (3.1 節と同様) 原始論理式 $at(s_i)$ で表すとする。また、現在の状態での行動 (イベント) e の期待報酬が r であることを $exp\_rew(e, r)$ で表す (他の状態における期待報酬、例えば、状態 s_i での行動 e の期待報酬の現在の値が r であることは、$\mathsf{AG}(at(s_i) \supset exp\_rew(e, r))$ と表される)。

すると、エージェントがグリーディ方策[5] を取ることは、(イベント定数記号の集合 $\mathcal{E}$ が $\{e_1, \cdots, e_m\}$ だとして)

$$\bigwedge_{1 \leq i \leq m} (exp\_rew(e_i, r_i) \wedge r_i \leq r_j \vee \mathsf{AX}^{e_i}\,dead) \supset \mathsf{INTEND}^a\,does(e_j)$$

という axiom (が各 $j = 1, \cdots, m$ について、計 m 個ある) として書ける (簡単のため、2 つの異なる行動の期待報酬が一致することは考えないものとする)。これは「可能でないものを除いた全ての行動のうち、期待報酬最大の行動を e_j とし、その行動を取ることを意図する」を意味する。また、$does(e_j)$ を意図すれば、その結果行動 e_j が外界に向かって起こされるものとする。

いま、この axiom (と数値比較に関する十分な axiom) を含む axiom 群があって、かつ $\mathcal{E} = \{e_1, e_2, e_3\}$ とし、現在 $exp\_rew(e_1, 0.3), exp\_rew(e_2, 0.5), \mathsf{AX}^{e_3}\,dead$ がその axiom 群からの結論として既に得られているとする。このとき、BC 規則を用いた図 9 の証明図 (ここで $\Sigma \supset \{\mathsf{AG}\,0.3 \leq 0.5, \mathsf{AG}\,0.5 \leq 0.5, \mathsf{AG}\,\zeta\}$、ただし ζ は上記の axiom で $m = 3, j = 2, r_1 = 0.3, r_2 = 0.5$ の場合の式とする。それ以外の axiom も Σ に含まれうる。なお図では一部、複数の規則の適用をまとめて示している箇所がある) から、「この axiom 群をプログラムとして動くエージェントは行動 e_2 を選ぶ」が示されたことになる。

なお、ε グリーディ方策 (i.e. 一定の少ない確率でランダムに行動を選び、それ以外の場合はグリーディ手法を取る) は、各時刻ごとに (0 以上 1 未満の) 異なる引数で成り立つ 1 引数述語 *random* があるものとすれば記述できる。axiom として $(random(x) \wedge x \geq 0.1 \supset$ グリーディ手法を取る$) \wedge (random(x) \wedge x < 0.1 \supset$ ランダムに行動を選択$)$ などと書けばよい。

また、確率的遷移においては、ある行動 e について、その行動によって各状態 s へ遷移する確率 p_s と、その遷移の際の期待即時報酬 r_s を既知として、その行動の期待報酬を式

$\sum_s p_s r_s$ によって求めたいことがある。これは、下記の axiom (と、やはり数値計算に関する十分な axiom) があれば同様に可能である。

$$\mathsf{X}^e(_{\geq p_1}\, \mathit{at}(s_1) \wedge \mathit{reward}(r_1) \mid \cdots \mid {}_{\geq p_n}\, \mathit{at}(s_n) \wedge \mathit{reward}(r_n)) \wedge \sum_{1\leq i\leq n} p_i = 1 \wedge \sum_{1\leq i\leq n} p_i r_i = r \supset \mathit{exp\_rew}(e, r)$$

$$\dfrac{\dfrac{\dfrac{\Sigma \to \mathit{exp\_rew}(e_1, 0.3)}{\Sigma \to \mathit{exp\_rew}(e_1, 0.3) \wedge 0.3 \leq 0.5}}{\Sigma \to \mathit{exp\_rew}(e_1, 0.3) \wedge 0.3 \leq 0.5 \vee \mathsf{AX}^{e_1}\, \mathit{dead}} \quad \begin{array}{c}\dfrac{\dfrac{\Sigma \to \mathit{exp\_rew}(e_2, 0.5)}{\Sigma \to \mathit{exp\_rew}(e_2, 0.5) \wedge 0.5 \leq 0.5}}{\Sigma \to \mathit{exp\_rew}(e_2, 0.5) \wedge 0.5 \leq 0.5 \vee \mathsf{AX}^{e_2}\, \mathit{dead}} \\ \uparrow \end{array} \quad \dfrac{\Sigma \to \mathsf{AX}^{e_3}\, \mathit{dead}}{\Sigma \to \mathit{exp\_rew}(e_3, x) \wedge x \leq 0.5 \vee \mathsf{AX}^{e_3}\, \mathit{dead}}}{\Sigma \to \mathsf{INTEND}^a\, \mathit{does}(e_2)}\ \mathrm{BC}$$

図 9 推論による動作記述へのグリーディ手法の取り込み

4.3 学習過程のモデル化

強化学習でよく知られる Sarsa や Q などの学習アルゴリズムは、TD 学習[5] と呼ばれる手法に属する。この方法は、現時点での各行動の期待報酬の推定値をもとに (ε グリーディ方策などで) 行動を選択し、観測された次状態と即時報酬に基づいて、元の状態での今取った行動の期待報酬の推定値を更新するというものである。すなわち、ある状態 s で行動 e を取ることの期待報酬の推定値を $Q(s, e)$ と表すと (これは論理式ではなく、通常の数式に現れる関数表記である)、その更新は $Q(s, e) := Q(s, e) + 差分$ として行われる。ここで「差分」は、行動 e を取った結果の即時報酬や次状態などによって決まり、また学習方式によっても異なる。

この場合、期待報酬の推定値は可変となるので、本節では、4.2 節と違って期待報酬に関する情報は信念として (例えば $\mathsf{BEL}^a(\mathit{at}(s) \supset \mathit{exp\_rew}(e, r))$ のような形で) 保持してこれを用い、必要に応じて更新するものとする。

そこで、行動を決定する axiom は 4.2 節のものと違い、例えば $\mathsf{BEL}^a\, \mathit{at}(s) \wedge \bigwedge_{1\leq i\leq m}(\mathsf{BEL}^a(\mathit{at}(s) \supset \mathit{exp\_rew}(e_i, r_i)) \wedge r_i \leq r_j \vee \mathsf{AX}^{e_i}\, \mathit{dead}) \supset \mathsf{INTEND}^a\, \mathit{does}(e_j)$ のように、期待報酬に関しては信念を用いることになる[⋆1]。

そして行動した結果、期待報酬を更新する。その表現のために、次の axiom

$$\mathsf{BEL}^a\, \mathit{at}(s) \wedge \mathsf{BEL}^a(\mathit{at}(s) \supset \mathit{exp\_rew}(e, r)) \wedge \mathsf{INTEND}^a\, \mathit{does}(e) \supset \mathsf{AX}^e\, \xi$$
$$ただし\ \xi = \mathit{calc\_delta}(d) \wedge (r + d = r') \supset \mathsf{BEL}^a(\mathit{at}(s) \supset \mathit{exp\_rew}(e, r'))$$

を用意しておく。$\mathit{calc\_delta}(d)$ は「計算された差分の値が d である」を表す。この axiom と $\mathsf{BEL}^a\, \mathit{at}(s)$, $\mathsf{BEL}^a(\mathit{at}(s) \supset \mathit{exp\_rew}(e, r))$, および前節と同様の過程で得られる行動の決定 $\mathsf{INTEND}^a\, \mathit{does}(e)$ があれば、$\mathsf{AX}^e\, \xi$ を導くことができる。このときエージェントの実装は、次の時刻で ξ を成り立たせる (ように実装する)。そして、次の時刻で差分が計算され $\mathit{calc\_delta}(d) \wedge (r + d = r')$ が成り立てば、$\mathsf{BEL}^a(\mathit{at}(s) \supset \mathit{exp\_rew}(e, r'))$ を新たな信念とすることができる。(その他の信念は保持する必要があるが、ここでは略する)

以上の例のように、$\mathcal{TOMATO}$ によって、確率的遷移を含め、強化学習のメカニズムを BDI の論理モデル内に取り込むことが可能となる。

5. 考　　察

5.1 PCTL との比較

$\mathcal{TOMATO}$ での確率つき時相オペレータは、CTL での next time オペレータに確率を導入したものである。これは、2.3.2 節で述べたようにタブロー法による証明系の構築を意識しているからであるが、難点として、確率つきの記述が次の時刻との間の遷移に限られる点が挙げられる。

PCTL[2] では、確率は path formula に対して導入される。すなわち、時刻の列 (パス) 上での確率を記述できる。例えば、「確率 0.9 以上で将来 ϕ を達成できる」を論理式 $[\mathit{true}\, \mathcal{U}\, \phi]_{\geq 0.9}$ で記述可能である。現在の $\mathcal{TOMATO}$ ではこのような記述ができない。

ただし、PCTL のような方式では、$\mathcal{TOMATO}$ のようにイベント毎に確率を記述することは難しい。また PCTL では確率表現の自由度が高すぎ、タブロー法による証明系は作りにくいのではないかと考えられる[⋆2]。証明系の構築と記述力のバランスは課題の 1 つである。

5.2 本質的な要充足集合の自動構築

3 節の事例を証明する際は、証明図の展開で出現する $\Gamma = \{\mathsf{X}^e_{r_1\, p_1}\, \psi_1, \cdots, \mathsf{X}^e_{r_n\, p_n}\, \psi_n\}$ に対する本質的な要充足集合を手動で求め、X-KD の推論規則を作成して、証明図を構築した。しかし、自動証明への応用を考える場合、要充足集合を自動的に求めることが必要である。

集合 $\{Q_1, \cdots, Q_m\}$(ここで各 $Q_j = \{\psi_{j,1}, \cdots, \psi_{j,\ell_j}\}$ とする) が要充足集合かどうかは、m 変数連立線形不等式

⋆1 ここでは、現在の状態に関する情報も信念とした。

⋆2 確率の記述を 0 と 1 に制限した qualitative PCTL でさえ、文献[1] の時点で演繹体系はまだ知られていない。

Vol.2010-MPS-81 No.23
Vol.2010-BIO-23 No.23
2010/12/17

$$\begin{cases} \sum_{1\leq j\leq m} x_j = 1 & \\ x_j > 0 & (\text{for } 1 \leq j \leq m) \\ (\sum_{1\leq j\leq m, \psi_i \in Q_j} x_j)\; r_i \; p_i & (\text{for } 1 \leq i \leq n) \end{cases}$$

に解があるかどうかでわかるので、この計算が有限時間で具体的に可能ならば、候補の全数検査によって全ての要充足集合を、従って全ての本質的な要充足集合を求めるアルゴリズムがあることになる。例えば p_i が全て有理数の場合がそうである。一方、p_i に計算不能な実数が含まれている場合などは、要充足集合を求めることも不可能であろう。また、計算が可能であっても、応用上はより効率的な要充足集合の求め方が必要と考えられる。それらも含め、今後の課題である。

6. ま と め

本論文では、BDI logic に確率的遷移と不動点オペレータを導入して我々が提案した論理体系 $\mathcal{TOMATO}$ を用いて、確率的状態遷移を伴う応用の記述や証明の例を示した。特に、BDI モデルに望まれている拡張の 1 つである強化学習との結合について、その動作を BDI の論理モデルでの推論による動作記述へ取り込む事例を述べ、それらに対する形式的な扱いが可能であることを示した。提案した $\mathcal{TOMATO}$ が合理的エージェントのモデル化および実現のための有意義な道具として資することを期待する。

参 考 文 献

1) Brázdil, T., Forejt, V., Křetínský, J. and Kučera, A.: The Satisfiability Problem for Probabilistic CTL, *Proc. of 23rd Annual IEEE Symposium on Logic in Computer Science*, pp. 391–402 (2008).
2) Hansson, H. and Jonsson, B.: A Logic for Reasoning about Time and Reliability, *Formal Aspects of Computing*, Vol.6, No.5, pp.512–535 (1994).
3) Kozen, D.: Results on the propositional μ-calculus, *Theoretical Computer Science*, Vol.27, pp.333–354 (1983).
4) NIDE, N., Takata, S. and Fujita, M.: BDI logic with probabilistic transition and fixed-point operator, *Proc. of CLIMA '09*, pp.71–86 (2009).
5) Sutton, R.S. and Barto, A.G.: *Reinforcement Learning: An Introduction*, The MIT Press (1998). (三上貞芳, 皆川雅章 (共訳). 強化学習, 森北出版, 2002).
6) Tarski, A.: A Lattice-Theoretical Fixpoint Theorem and its Application, *Pacific Journal of Mathematics*, Vol.5, pp.285–309 (1955).
7) 新出尚之：確率的遷移と不動点オペレータを持つ BDI logic について，合同エージェントワークショップ&シンポジウム (JAWS2008) 講演論文集 (2008).