情報処理学会論文誌　数理モデル化と応用　Vol. 4　No. 3　59–72 (July 2011)

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への外部の行為選択機構の導入に関わるさまざまな性質の記述や推論が行える．本論文では，$\mathcal{TOMATO}$の演繹体系を用いて，確率的状態遷移をともなうエージェントのモデル化，およびその性質に関する証明を扱う例を示す．特に，確率的状態遷移の導入が本質的に必要となる，BDIエージェントへの強化学習の取り込みに主眼をおいて述べる．

Modeling Probabilistic State Transitions Using $\mathcal{TOMATO}$ and Its Application

Naoyuki Nide,[†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 introductions of the external action decision mechanisms into them. In this paper, using the deduction system of $\mathcal{TOMATO}$, we give a detailed explanation for modeling of agents with probabilistic state transitions, and some examples of proofs about them. In particular, we pay special attention to the modeling of the introduction of reinforcement learning into BDI agents.

†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

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．BDIモデルに基づいて実現されるエージェントは，BDIエージェントと呼ばれる．

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[16]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」などがあげられる．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[11),21)]．

$\mathcal{TOMATO}$の利点としては，4章で述べるように，従来のBDI logicでは扱えなかった，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）[12)]における状態遷移を確率遷移オペレータで記述できることや，従来のBDI logicを拡張した論理体系$\mathcal{LORA}$[18)]では無限の論理和/積を用いて表現するしかなかった相互信念や共同意図などを，不動点オペレータを用いて有限の論理式で記述できることなどがあげられる．また，それらの性質についてシーケント計算での演繹が可能なことも利点である．

本論文では，確率的状態遷移をともなうBDIエージェントを$\mathcal{TOMATO}$でモデル化し，さらにその演繹体系を用いてエージェントの性質に関する証明を行う例について述べる．特に，上述したBDIモデルへの拡張のうち強化学習との結合に着目し，これを$\mathcal{TOMATO}$の確率遷移オペレータを用いてBDIの論理モデルに取り込む事例，さらにシーケント計算による演繹を学習過程の表現へ応用する例を示す．

そのためにまず，2 章で確率的状態遷移の導入の動機を述べ，次いで 3 章で $\mathcal{TOMATO}$ の導入を行い，特に，$\mathcal{TOMATO}$ で用いている Kripke 構造である BDI ストラクチャにおける，確率的状態遷移に関するモデル化の概念と推論規則について詳細に述べる．続いて，4 章で確率的状態遷移に関する事例，および 5 章では特に強化学習との結合への応用を念頭に置いた事例について，$\mathcal{TOMATO}$ を用いた記述や証明の例を与える．6 章･7 章でまとめる．なお，協調行為などの記述への応用については別論文[21] に譲る．

2. 確率的状態遷移への動機

我々は，BDI エージェントと強化学習を統合して，熟考的行為と状況に応じて瞬時に実行される反射的行為を兼ね備えた合理的エージェントを提案している[20]．これは，たとえば通い慣れた道を通る際には，実践的推論をそのつど行うことなく体が周囲の状況に反応して経路選択したり，あるいはサッカー競技において，厳しい練習の積み重ねで獲得した技能に従って瞬時に適切な行為をしたりするなど，学習により獲得された反射的行為を，BDI エージェントに取り入れ，より人間に近い振舞いが行えるとする考えである．

我々は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）での状態遷移を BDI モデルの中で記述する試みを行った[22]．しかし，MDP は基本的に確率的遷移であり，確率的な遷移オペレータを持たない BDI logic では「可能な遷移先の中のどこかへ動く」ことしか書けない．たとえば「状態 s_1 にいて行動 e_1 をとると，確率 0.7 で状態 s_2 へ動いて即時報酬 3 を受け取り，確率 0.3 で状態 s_3 へ動いて即時報酬 5 を受け取る」という状況は，従来の BDI logic では，確率を捨象して「どちらかになる」というレベルでしか書けず，確率的遷移の表現が必要となる．

確率的遷移を扱えるように CTL を拡張した論理体系としては PCTL[5] が知られるが，これは 6.2.1 項で述べるように，ある状態から延びるパス（時刻の列）単位に確率を記述するため，BDI モデルが対象とする各行動（イベント）による遷移を区別して個々に遷移確率を与えることは難しいと考えられる．

3. 論理体系 $\mathcal{TOMATO}$

本章では様相論理体系 $\mathcal{TOMATO}$ 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 章で述べたような，BDI モデルが元来持たない行動選択機構の外部からの導入を扱いやすくすることを重視している．特に，確率的状態遷移オペレータの導入は，2 章に述べた動機に基づく．

3.1 論　理　式

3.1.1 構　　文

まず $\mathcal{TOMATO}$ の論理式の定義を与える．以下では，単に「論理式」といえば $\mathcal{TOMATO}$ 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．

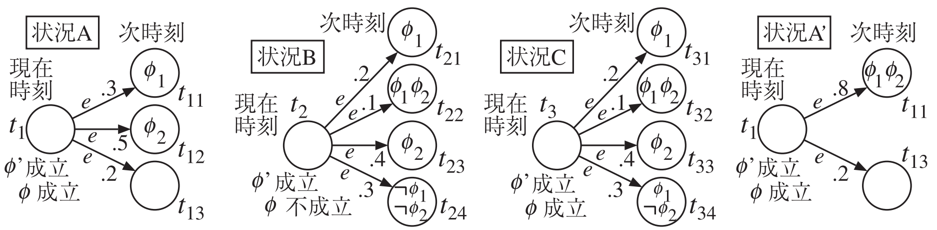
一階言語 $\mathcal{L}$，論理式変数記号の集合 $\mathcal{V}$，イベント定数記号の集合 $\mathcal{E}$，および，エージェント定数記号の集合 $\mathcal{A}$ を各 1 つずつ適当に選んで与えておく．ただし $\mathcal{E}$, $\mathcal{A}$ は有限集合とする．以下，$\{p \mid p \in \mathbb{R}, 0 \leq p \leq 1\}$ を $[0,1]$ と書く．

- 一階述語論理の原始論理式は（$\mathcal{TOMATO}$ 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，$r_i \in \{\geq, >\}$，$p_i \in [0,1]$ かつ $\sum_{i=1}^{n} p_i \leq 1$ 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 $\mathcal{TOMATO}$ の構文の定義である．このほか，$\wedge$ や $\supset$ や $\Leftrightarrow$ や $\exists$ を一般的な略記として導入し，加えて *false* を $p \wedge \neg p$ の略記（p は適当に 1 つ選んで固定した論理式），*true* を $\neg$*false* の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･$\wedge$･$\vee$･$\supset$ の順に先に結合するものとする．さらに，$\wedge$･$\vee$ 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[8]，ν は最大不動点オペレータである．

また，$\mathsf{X}^e_{<p}\,\phi$, $\mathsf{X}^e_{\leq p}\,\phi$, $\mathsf{X}^e_{=p}\,\phi$, $\mathsf{X}^e_{\neq p}\,\phi$ は各々，$\neg\mathsf{X}^e_{\geq p}\,\phi$, $\neg\mathsf{X}^e_{>p}\,\phi$, $\mathsf{X}^e_{\geq p}\,\phi \wedge \mathsf{X}^e_{\leq p}\,\phi$, $\neg\mathsf{X}^e_{=p}\,\phi$ 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，$\mathsf{AX}\,\phi$ は $\bigwedge_{e \in \mathcal{E}} \mathsf{AX}^e\,\phi$ の略記とする．



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

図 1　X^e オペレータの解釈の直感的説明
Fig. 1　Informal explanation of the interpretation of X^e operator.

3.1.2　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．例として $\phi = \mathsf{X}^e({}_{\geq .3}\,\phi_1 \mid {}_{\geq .5}\,\phi_2)$, $\phi' = \mathsf{X}^e_{\geq .3}\,\phi_1 \wedge \mathsf{X}^e_{\geq .5}\,\phi_2$ とすると，ϕ 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，一方，ϕ' 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．たとえば**図 1** の状況 A では現在時刻（t_1）で ϕ, ϕ' とも成り立つ一方，状況 B では現在時刻（t_2）で ϕ' は成り立つが ϕ は成り立たない．t_2 では，ϕ_1 が成り立つ 1 時刻後の状態と，ϕ_2 が成り立つ 1 時刻後の状態とを重ならないように（かつ，遷移確率がそれぞれ 0.3, 0.5 以上となるように）とれないからである．なお，ϕ は ϕ_1 と ϕ_2 が背反であることを要求するものではないので，状況 C の現在時刻（t_3）では ϕ は（ϕ' も）成り立つ．ϕ_1 が成り立つ 1 時刻後の状態として t_{34}，ϕ_2 が成り立つ 1 時刻後の状態として t_{32} と t_{33} をとれるためである．

ただし，状態遷移を記述する立場からは，複数の同等な状態を単一の状態と同一視してよいことがある．たとえば状況 A において，もし t_{11} と t_{12} で成り立つ論理式の集合が（様相オペレータを持つものも含め）まったく同一な（そして ϕ_1 も ϕ_2 も成り立つ）場合は，t_{11} と t_{12} を 1 つの状態と見なし，状況 A を状況 A$'$ と同一視してかまわない．このため，状況 A$'$ でも上記 ϕ は成り立つと定める．

AX オペレータと不動点オペレータを持つ分岐時相論理は，CTL*より高い記述力を持つことが知られており[4),10)]，このため $\mathcal{TOMATO}$ は従来の BDI logic を包含している．たとえば $\mu\mathfrak{X}.(\psi \vee \phi \wedge \mathsf{AX}\,\mathfrak{X})$ を $\mathsf{A}(\phi\,\mathsf{U}\,\psi)$ と略記すれば，これは CTL での同じ式（until オペレータ）と同等の表現になる．

$\mathsf{BEL}^a\,\phi$, $\mathsf{DESIRE}^a\,\phi$, $\mathsf{INTEND}^a\,\phi$ 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．たとえば $\mathsf{BEL}^a\,\mathsf{INTEND}^b\,\phi$ で「エージェント b が ϕ を意図していることをエージェント a が信じる」を表現できる．また，$\mathsf{AF}\,\phi$ を $\mathsf{A}(\mathit{true}\,\mathsf{U}\,\phi)$ の略記とする（「ϕ がいつかは成り立つ」を表す）と，$\mathsf{INTEND}^a\,\mathsf{AF}\,\phi \supset \mathsf{A}(\mathsf{INTEND}^a\,\mathsf{AF}\,\phi\,\mathsf{U}\,\mathsf{BEL}^a\,\phi)$ で blind commitment と呼ばれるコミットメント戦略[14)]（ϕ をいつか達成することを意図すれば，その達成を信じるまではその意図を保持する）を表現できる．なお，現時点の $\mathcal{TOMATO}$ は，心的状態オペレータには確率を導入していない．

さらに，不動点オペレータと心的状態オペレータの併用によって，たとえば「エージェント群 $\mathcal{G}$（$\subseteq \mathcal{A}$）が ϕ を相互信念として持つ」を $\nu\mathfrak{X}.\bigwedge_{a\in\mathcal{G}} \mathsf{BEL}^a(\phi \wedge \mathfrak{X})$ で表現するようなこともでき，これは協調行為などのモデル化に有用である．ただし，本論文では確率的状態遷移のモデル化に主眼を置くため，以後これについては触れない．

3.2　意　味　論

3.2.1　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 はすべての可能世界のすべての 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[⋆1] 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 $R^e_w \subseteq St_w \times St_w$ と，R^e_w から $[0, 1]$ への関数 $\mathcal{P}^e_w$．ただし，任意の $t \in St_w$ および任意の $e \in \mathcal{E}$ に対し，$\sum_{t' \in \{t' \mid t\,R^e_w\,t'\}} \mathcal{P}^e_w(t, t') = 1$ 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^e_w は可能世界 w 内の時刻の前後関係で，$t\,R^e_w\,t'$ および $\mathcal{P}^e_w(t, t') = p$ は state t でイベント e を実行すると確率 p で次

⋆1 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．また，R が Euclidean であるとは，任意の s, s', $s'' \in S$ に対し，$s\,R\,s'$ かつ $s\,R\,s''$ ならば $s'\,R\,s''$ となることをいう．

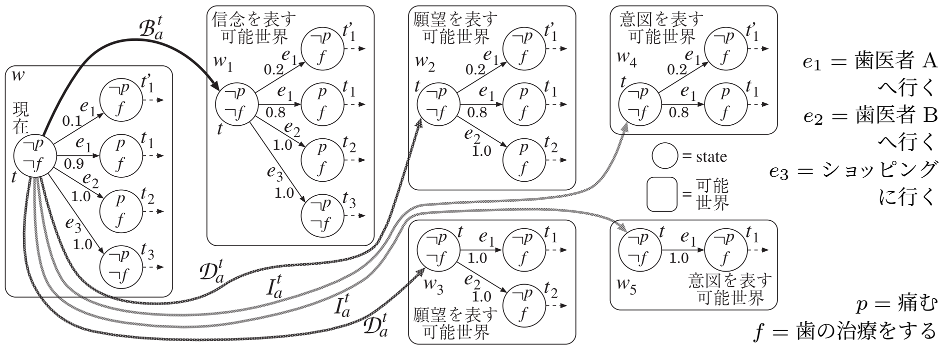


図 2　BDI ストラクチャの概略

Fig. 2　Overview of BDI structure.

の時刻は state t' になることを表す．$\mathcal{B}^t_a$, $\mathcal{D}^t_a$, $\mathcal{I}^t_a$ 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 の信念･願望･意図を表す（**図 2** に概略を示した．ただしこの図は $\mathcal{B}^t_a$ などのすべてを表示したものではなく，$\mathcal{B}^t_a$ の serial 性などの性質も反映していない）．

この図は Rao ら[14] による合理的エージェントの行動の決定過程のモデル化に用いられているもの（ただし少し改変し，確率的遷移を追加した）でもある．世界 w の時刻 t でエージェント a は「歯医者に行けば歯は治せるが，治療で痛む確率が高い」「ショッピングに行けば治療で痛むことはないが歯は治せない」という信念を持ち（この例ではたまたま a にとっての「現実世界」w も信念の世界 w_1 とよく似ている），可能な未来のうち歯を治す未来を願望し，歯医者 A に行く（e_1 を実行する）選択をしてそれを意図とする*1．

なお，R^e_w は serial としたため，形式的にはどの 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．たとえば「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．図 2 では死状態への遷移は省いた．

*1 ただし現実のエージェントでは，e_1 を実行する選択は主体的に行えるが，その結果図 2 の t_1，t'_1 のいずれに行くかは選択できないであろう．この選択は環境から与えられるものである．

3.2.2　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（$[\![\phi]\!]_{\langle M, f_{\mathcal{V}}\rangle} \subseteq \mathit{Swt}$ 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
- $[\![\mathsf{X}^e({}_{r_1 p_1}\, \phi_1 \mid \cdots \mid {}_{r_n p_n}\, \phi_n)]\!]_{\langle M, f_{\mathcal{V}}\rangle} = \{(w,t) \mid T = \{t' \mid t\ R^e_w\ t'\}$ とするとき，各 $t' \in T$，および $i = 1, \cdots, n$ のそれぞれについて，$\mathfrak{p}^{t'}_i \in [0,1]$ が存在して以下を満たす：①各 $t' \in T$ に対し $\sum_{i=1}^{n} \mathfrak{p}^{t'}_i \leq \mathcal{P}^e_w(t,t')$　②$(w,t') \notin [\![\phi_i]\!]_{\langle M, f_{\mathcal{V}}\rangle}$ ならば $\mathfrak{p}^{t'}_i = 0$　③$1 \leq i \leq n$ に対し $\sum_{t' \in T} \mathfrak{p}^{t'}_i\ r_i\ p_i\}$（$r_i$ は $>, \geq$ のいずれかであることに注意）
- $[\![\mathsf{BEL}^a\, \phi]\!]_{\langle M, f_{\mathcal{V}}\rangle} = \{(w,t) \mid w\, \mathcal{B}^t_a\, w'$ なる任意の世界 w' に対して $(w',t) \in [\![\phi]\!]_{\langle M, f_{\mathcal{V}}\rangle}\}$
- $[\![\mathsf{DESIRE}^a\, \phi]\!]_{\langle M, f_{\mathcal{V}}\rangle}$, $[\![\mathsf{INTEND}^a\, \phi]\!]_{\langle M, f_{\mathcal{V}}\rangle}$ 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 $f_\phi : \mathit{Swt} \to \mathit{Swt}$ ととらえ直せる．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．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[17]．

$[\![\phi]\!]_{\langle M, f_{\mathcal{V}}\rangle}$ 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．

たとえば，図 1 の状況 A$'$ がある世界 w の時刻の木となっているとき，(w,t_1) で $\mathsf{X}^e({}_{\geq .3}\, \phi_1 \mid {}_{\geq .5}\, \phi_2)$ が成り立つ．$\mathfrak{p}^{t_{11}}_1 = 0.3$，$\mathfrak{p}^{t_{11}}_2 = 0.5$，$\mathfrak{p}^{t_{13}}_1 = \mathfrak{p}^{t_{13}}_2 = 0$ とすればよいからである．また，図 2 の (w,t) では $\mathsf{DESIRE}^a\, \mathsf{AX}^{e_1}\, f$ が成り立つ．

3.3　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，Γ, Δ などギリシャ文字の大文字（「$'$」付きの Γ' 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

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

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

$$\frac{\Gamma, \mathsf{BEL}^a\,\Gamma \to \mathsf{BEL}^a\,\Delta, \Theta, \mathsf{BEL}^a\,\Theta}{\mathsf{BEL}^a\,\Gamma \to \mathsf{BEL}^a\,\Delta, \mathsf{BEL}^a\,\Theta}\,\text{BEL-KD45} \quad \frac{\Gamma \to \Theta}{\mathsf{DESIRE}^a\,\Gamma \to \mathsf{DESIRE}^a\,\Theta}\,\mathsf{DESIRE}\text{-KD} \quad \frac{\Gamma \to \Theta}{\mathsf{INTEND}^a\,\Gamma \to \mathsf{INTEND}^a\,\Theta}\,\mathsf{INTEND}\text{-KD} \quad \frac{\cdots \;\; \bigwedge_{\mathfrak{S} \subseteq \{1,\cdots,n\}, \mathfrak{S} \neq \emptyset} \mathsf{X}^e_{r_{\mathfrak{S}}\,p_{\mathfrak{S}}}(\bigvee_{i \in \mathfrak{S}} \phi_i) \;\; \cdots}{\cdots \;\; \mathsf{X}^e({}_{r_1 p_1}\,\phi_1 \mid \cdots \mid {}_{r_n p_n}\,\phi_n) \;\; \cdots}\,\mathsf{X}_{\text{excl}}$$

図 3　$\mathscr{TOMATO}$ の推論規則（規則 X-KD を除く．また，規則 ∀左・∀右・X_{excl} については 3.3.1 項に注記あり）
Fig. 3　Inference rules of $\mathscr{TOMATO}$ (excluding X-KD rule).

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「→」の左右は論理式の multi set とする（したがって Exchange の規則がない）．

シーケント ‘$\Gamma \to \Delta$’ の解釈は，論理式 $\bigwedge \Gamma \supset \bigvee \Delta$ の解釈と定義する（ただし $\bigwedge \emptyset = \mathit{true}, \bigvee \emptyset = \mathit{false}$ とする）．

3.3.1　推 論 規 則

推論規則は**図 3** にあげたものである．ただし，このほかに「→」の左の X^e オペレータに関する推論規則 X-KD があり，これについては別途 3.3.2 項で述べる．

また，∀ 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 X_{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 $i = 1, \cdots, n$ に対し規則中の r_i は $\geq$ か $>$ のいずれかとし，$p_{\mathfrak{S}} = \sum_{i \in \mathfrak{S}} p_i$，また $r_{\mathfrak{S}}$ は，もし $\{r_i \mid i \in \mathfrak{S}\} = \{\geq\}$ ならば $\geq$，さもなければ $>$ とする．この規則は直感的には「ある $\mathfrak{S} \subseteq \{1, \cdots, n\}$ について，$\bigvee_{i \in \mathfrak{S}} \phi_i$ が確率 $\sum_{i \in \mathfrak{S}} p_i$ 以上（またはそれより大）で成り立たなければ，各 ϕ_i（$i \in \mathfrak{S}$）の成り立つ状態を確率 r_i 以上（またはそれより大）で，しかも重ならないようにとることはできない（したがって $\mathsf{X}^e({}_{r_1 p_1}\,\phi_1 \mid \cdots \mid {}_{r_n p_n}\,\phi_n)$ は成り立たない）」ととらえられる．たとえば，以下は X_{excl} 規則の例である．

$$\frac{\mathsf{X}^e_{\geq .7}\,\phi_1 \wedge \mathsf{X}^e_{\geq .3}\,\phi_2 \wedge \mathsf{AX}^e(\phi_1 \vee \phi_2) \to \psi}{\mathsf{X}^e({}_{\geq .7}\,\phi_1 \mid {}_{\geq .3}\,\phi_2) \to \psi}$$

3.3.2　X^e オペレータに関する推論規則

図 3 のほかに，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 $\sum_{Q \subseteq \mathfrak{Q}} v(Q) = 1$ を満たし，かつ，$1 \leq i \leq n$ なるすべての整数 i について $(\sum_{Q \in \{T \mid T \subseteq \mathfrak{Q}, \psi_i \in T\}} v(Q))\, r_i\, p_i$ 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 $Q \subseteq \mathfrak{Q}$ 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，$\{Q \subseteq \mathfrak{Q} \mid v(Q) > 0\}$ 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 $Z, Z' \subseteq 2^{\mathfrak{Q}}$ に対し，ある $Q \in Z$ と $Z'' \subseteq 2^{Q}$ 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，したがって 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すべての列挙を $Z_1 = \{Q_{1,1}, \cdots, Q_{1,m_1}\}, \cdots, Z_k = \{Q_{k,1}, \cdots, Q_{k,m_k}\}$ 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 $\mathscr{TOMATO}$ の推論規則である[⋆1]．

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

以降，この規則について例による直感的説明を行う．

⋆1 X-KD の定義には，充足可能性という（意味論上の）概念は使われていないので注意されたい．

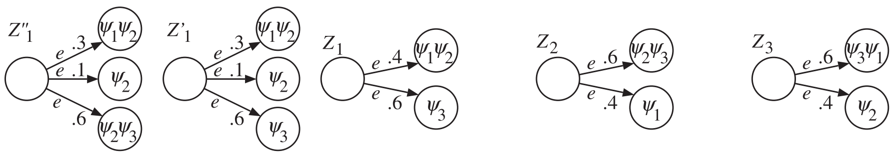


図 4　Σ の要充足集合の事例
Fig. 4　Example of the satisfaction request set of Σ.

たとえば $\Sigma = \{\mathsf{X}^e_{\geq 0.3}\,\psi_1, \mathsf{X}^e_{\geq 0.4}\,\psi_2, \mathsf{X}^e_{\geq 0.6}\,\psi_3\}$，$\mathfrak{Q} = \{\psi_1, \psi_2, \psi_3\}$ とすると，$v(\{\psi_1,\psi_2\}) = 0.4$，$v(\{\psi_3\}) = 0.6$，その他の $Q \subseteq \mathfrak{Q}$ に対しては $v(Q) = 0$，で与えられる v は Σ の確率分配関数の 1 つであり，$\mathrm{req}_v(\Sigma) = \{\{\psi_1,\psi_2\},\{\psi_3\}\} = Z_1$ は Σ 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，Σ 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 $Q \subseteq \mathfrak{Q}$ に対しては $v''(Q) = 0$，で与えられる v'' による $\mathrm{req}_{v''}(\Sigma) = Z''_1$ や，$v'(\{\psi_1,\psi_2\}) = 0.3$，$v'(\{\psi_2\}) = 0.1$，$v'(\{\psi_3\}) = 0.6$，その他の $Q \subseteq \mathfrak{Q}$ に対しては $v'(Q) = 0$，で与えられる v' による $\mathrm{req}_{v'}(\Sigma) = Z'_1$ もともに Σ 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．

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にらんで設けられている．タブロー法でシーケント ‘$\Gamma \rightarrow$’ 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．したがって，Γ の本質的な要充足集合のどれにも，‘$Q \rightarrow$’ が証明可能であるような要素 Q があることを示せば，‘$\Gamma \rightarrow$’ 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．

先の例の Σ の場合，Σ の本質的な要充足集合すべてを列挙すると，先の Z_1 のほかに $Z_2 = \{\{\psi_2,\psi_3\},\{\psi_1\}\}$，$Z_3 = \{\{\psi_3,\psi_1\},\{\psi_2\}\}$ がある．よって，推論規則としては

$$\frac{\psi_1,\psi_2 \rightarrow \quad \psi_2,\psi_3 \rightarrow \quad \psi_2 \rightarrow}{\Sigma \rightarrow} \qquad (\Sigma = \{\mathsf{X}^e_{\geq 0.3}\,\psi_1, \mathsf{X}^e_{\geq 0.4}\,\psi_2, \mathsf{X}^e_{\geq 0.6}\,\psi_3\})$$

など 8 つが作れる．しかし，上に例示した規則の前提のうち ‘$\psi_2 \rightarrow$’ に注目すると，その左の 2 つの前提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

$$\frac{\psi_1,\psi_2 \rightarrow \quad \psi_2,\psi_3 \rightarrow \quad \psi_3,\psi_1 \rightarrow}{\Sigma \rightarrow} \qquad \frac{\psi_1 \rightarrow}{\Sigma \rightarrow} \qquad \frac{\psi_2 \rightarrow}{\Sigma \rightarrow} \qquad \frac{\psi_3 \rightarrow}{\Sigma \rightarrow}$$

という 4 つの規則のみが残り，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（なお，このうち右の 3 つは $\frac{\phi \rightarrow}{\mathsf{X}^e_{\geq p}\,\phi \rightarrow}$（ここで $0 < p \leq 1$）という 1 つの形の規則に集約可）．

なお，タブロー法で規則 X-KD を（逆向きに）適用するには，$\mathsf{X}^e({}_{r_1 p_1}\,\phi_1 \mid \cdots \mid {}_{r_n p_n}\,\phi_n)$ の形の論理式を，すべて $n = 1$ の形に分解することと，X^e オペレータをすべて「→」の左に移すことが必要である．規則 $\mathsf{X}_{\mathrm{excl}}$, $\mathsf{X}_{\geq}$ 右，$\mathsf{X}_{>}$ 右はそのために設けられた規則である．

3.3.3　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 $\frac{S_1,\cdots,S_n}{S}$（$n \geq 0$）が存在し，$S_1,\cdots,S_n$ がすべて証明可能または 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）　$S =$ ‘$\Gamma, \mu\mathfrak{X}.\phi \rightarrow \Delta$’（ただし $\mathfrak{X}$ は Γ, 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，‘$\Gamma, \phi^n(\mathfrak{X}) \rightarrow \Delta$’ が $\{$‘$\Gamma, \mathfrak{X} \rightarrow \Delta$’$\}$ から導出可能

シーケント ‘$\rightarrow \phi$’ 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[11]．ただし本論文では，$\mathcal{TOMATO}$ による確率的状態遷移のモデル化，およびそれらの上での証明の例に的をしぼって議論するため，健全性や完全性の証明については割愛する．

⋆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）．

4. 確率的状態遷移のモデル化および証明の例

4.1 確率的状態遷移に関する証明例

元来の BDI logic では「どちらかになる」としか記述できなかった，2 章の事例は，$\mathcal{TOMATO}$ の確率的遷移オペレータを用いると，$at(s_1) \supset \mathsf{X}^{e_1}({}_{\geq .7}\, at(s_2) \wedge reward(3) \mid {}_{\geq .3}\, at(s_3) \wedge reward(5))$ と書ける（ここで $at(s)$ は「現在の状態が s である」を，$reward(r)$ は「即時報酬が r である」を表す原始論理式とする）．この論理式を ϕ とし，$\psi = \exists x(reward(x) \wedge x \geq 3)$ 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，$\phi \wedge at(s_1) \supset \mathsf{AX}^{e_1}\, \psi$ 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 以上の即時報酬を受ける」を示せる．この証明は**図 5** に示す．図中 X-KD 規則の適用において，$\{\mathsf{X}^{e_1}_{\geq 1}\, \xi_1, \mathsf{X}^{e_1}_{>0}\, \xi_2\}$ の本質的な要充足集合は $\{\{\xi_1, \xi_2\}\}$ のみである（ただし ξ_1, ξ_2 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 Peano の公理を用いるなどして行えるものとする[⋆1]．

一方，$\phi' = at(s_1) \supset \mathsf{X}^{e_1}_{\geq .7}(at(s_2) \wedge reward(3)) \wedge \mathsf{X}^{e_1}_{\geq .3}(at(s_3) \wedge reward(5))$ とすると，先と同じ ψ に対し $\phi' \wedge at(s_1) \supset \mathsf{AX}^{e_1}\, \psi$ は証明できない．次の時刻で，$at(s_2)$ が成り立つ状態と $at(s_3)$ が成り立つ状態を重ならないようにとれる保証がないためである．しかし，「$at(s_2)$ と $at(s_3)$ が排他的にしか成り立たない」がつねに成り立つことを条件に入れた式は証明可能となる．いま，$\mathsf{AG}\,\xi$ を $\nu\mathfrak{X}.(\xi \wedge \mathsf{AX}\,\mathfrak{X})$ の略記とする（CTL の $\mathsf{AG}\,\xi$ と同等で，「ξ が現在から未来永劫成り立つ」の意味になる）．すると，$\mathsf{AG}\,\neg(at(s_2) \wedge at(s_3)) \wedge \phi' \wedge at(s_1) \supset \mathsf{AX}^{e_1}\, \psi$ は**図 6** のように証明できる．図中の X-KD 規則では，$\{\mathsf{X}^{e_1}_{\geq 1}\, \xi_1, \mathsf{X}^{e_1}_{\geq .7}\, \xi_2, \mathsf{X}^{e_1}_{\geq .3}\, \xi_3, \mathsf{X}^{e_1}_{>0}\, \xi_4\}$（$\xi_1$〜$\xi_4$ は任意の論理式）の本質的な要充足集合が $\{\{\xi_1, \xi_2, \xi_3\}, \{\xi_1, \xi_4\}\}$, $\{\{\xi_1, \xi_3, \xi_4\}, \{\xi_1, \xi_2\}\}$ および $\{\{\xi_1, \xi_4, \xi_2\}, \{\xi_1, \xi_3\}\}$ であることを使っている．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，「I+W」は Initial と Weak をまとめて適用したことを表す（以後の図でも同様）．また，図中では次に示す 2 つの派生規則

$$\dfrac{\Gamma, \phi \to \Delta}{\Gamma, \mathsf{AG}\,\phi \to \Delta}\ \mathsf{AG}\,\text{左} \qquad \dfrac{\Gamma, \mathsf{AX}^{e}\,\mathsf{AG}\,\phi \to \Delta}{\Gamma, \mathsf{AG}\,\phi \to \Delta}\ \mathsf{AG}\,\text{左2}$$

を使った（AG 左2 は**図 7** から，また AG 左もほぼ同様にして得られる）．

⋆1 厳密には，このためには形式的体系に等号を導入することが必要になる．ただし，5.1 節に述べるように，数値比較の部分は axiom として扱う方式も考えられる．

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{q_1 \to q_1}\ \text{Initial} \quad \dfrac{\dfrac{\vdots}{\to 3 \geq 3}}{q_1 \to 3 \geq 3}\ \text{Weak}}{q_1 \to q_1 \wedge 3 \geq 3}\ \wedge\text{右}}{p_1, q_1 \to q_1 \wedge 3 \geq 3}\ \text{Weak}}{\zeta_1 \to q_1 \wedge 3 \geq 3}\ \wedge\text{左}}{\zeta_1 \to \psi}\ \exists\text{右}}{①\ \zeta_1, \neg\psi \to}\ \neg\text{左}$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{q_2 \to q_2}\ \text{Initial} \quad \dfrac{\dfrac{\vdots}{\to 5 \geq 3}}{q_2 \to 5 \geq 3}\ \text{Weak}}{q_2 \to q_2 \wedge 5 \geq 3}\ \wedge\text{右}}{p_2, q_2 \to q_2 \wedge 5 \geq 3}\ \text{Weak}}{\zeta_2 \to q_2 \wedge 5 \geq 3}\ \wedge\text{左}}{\zeta_2 \to \psi}\ \exists\text{右}}{②\ \zeta_2, \neg\psi \to}\ \neg\text{左}$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{①\ \zeta_1, \neg\psi \to \quad ②\ \zeta_2, \neg\psi \to}{\zeta_1 \vee \zeta_2, \neg\psi \to}\ \vee\text{左}}{\mathsf{AX}^{e_1}(\zeta_1 \vee \zeta_2), \mathsf{X}^{e_1}_{>0}\, \neg\psi \to}\ \text{X-KD}}{\mathsf{X}^{e_1}_{\geq .7}\, \zeta_1, \mathsf{X}^{e_1}_{\geq .3}\, \zeta_2, \mathsf{AX}^{e_1}(\zeta_1 \vee \zeta_2), \mathsf{X}^{e_1}_{>0}\, \neg\psi, at(s_1) \to}\ \text{Weak}}{\mathsf{X}^{e_1}_{\geq .7}\, \zeta_1, \mathsf{X}^{e_1}_{\geq .3}\, \zeta_2, \mathsf{AX}^{e_1}(\zeta_1 \vee \zeta_2), at(s_1) \to \mathsf{AX}^{e_1}\, \psi}\ \mathsf{X}_{\geq}\text{右}}{\mathsf{X}^{e_1}_{\geq .7}\, \zeta_1 \wedge \mathsf{X}^{e_1}_{\geq .3}\, \zeta_2 \wedge \mathsf{AX}^{e_1}(\zeta_1 \vee \zeta_2), at(s_1) \to \mathsf{AX}^{e_1}\, \psi}\ \wedge\text{左 2 回}$$

$$\dfrac{\mathsf{X}^{e_1}_{\geq .7}\, \zeta_1 \wedge \mathsf{X}^{e_1}_{\geq .3}\, \zeta_2 \wedge \mathsf{AX}^{e_1}(\zeta_1 \vee \zeta_2), at(s_1) \to \mathsf{AX}^{e_1}\, \psi}{\eta, at(s_1) \to \mathsf{AX}^{e_1}\, \psi}\ \mathsf{X}_{\text{excl}}$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{at(s_1) \to at(s_1)}\ \text{Initial}}{at(s_1) \to at(s_1), \mathsf{AX}^{e_1}\, \psi}\ \text{Weak} \quad \eta, at(s_1) \to \mathsf{AX}^{e_1}\, \psi}{\phi, at(s_1) \to \mathsf{AX}^{e_1}\, \psi}\ \supset\text{左}}{\phi \wedge at(s_1) \to \mathsf{AX}^{e_1}\, \psi}\ \wedge\text{左}}{\to \phi \wedge at(s_1) \supset \mathsf{AX}^{e_1}\, \psi}\ \supset\text{右}$$

ここで

$$\phi = at(s_1) \supset \eta,\ \ \eta = \mathsf{X}^{e_1}({}_{\geq .7}\, \zeta_1 \mid {}_{\geq .3}\, \zeta_2),$$
$$\zeta_1 = p_1 \wedge q_1,\ \ \zeta_2 = p_2 \wedge q_2,$$
$$p_1 = at(s_2),\ \ q_1 = reward(3),$$
$$p_2 = at(s_3),\ \ q_2 = reward(5),$$
$$\psi = \exists x(reward(x) \wedge x \geq 3).$$

また，以下の派生規則を用いた：

$$\dfrac{\Gamma, \phi, \psi \to \Delta}{\Gamma, \phi \wedge \psi \to \Delta}\ \wedge\text{左} \qquad \dfrac{\Gamma \to \Delta, \phi \quad \Gamma \to \Delta, \psi}{\Gamma \to \Delta, \phi \wedge \psi}\ \wedge\text{右}$$

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

図 5　証明図の例 1（①, ②の部分は後の図でも使用）
Fig. 5　Example of proof figure (1).

4.2 イベント選択に関する証明例

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，$e_1$ 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（3.2.1 項参照）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 $\mathsf{AX}^{e_1}\, \mathsf{X}^{e_3}_{\geq .4}\, \phi \wedge \mathsf{AX}^{e_2}\, \mathsf{X}^{e_3}_{\geq .2}\, \phi \wedge \mathsf{AX}^{e_3}\, dead \supset \mathsf{AX}(\neg dead \supset \mathsf{X}^{e_3}_{\geq .2}\, \phi)$ の証

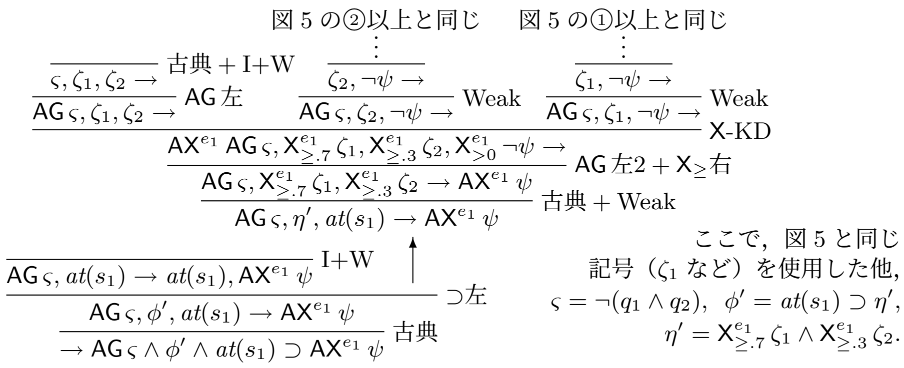


図 6　証明図の例 2
Fig. 6　Example of proof figure (2).

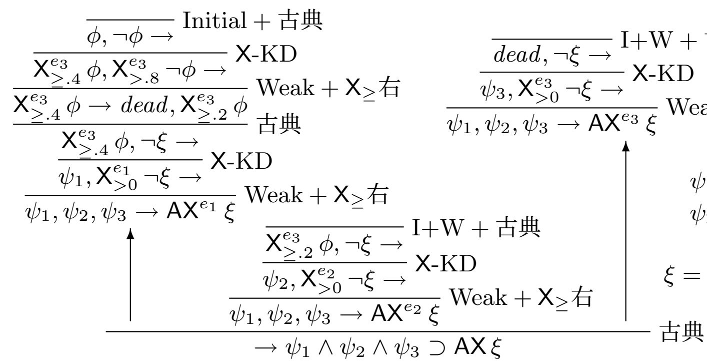


図 9　証明図の例 3
Fig. 9　Example of proof figure (3).

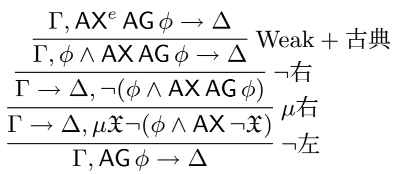


図 7　AG 左2 規則の派生
Fig. 7　Deriving AG-L2 rule.

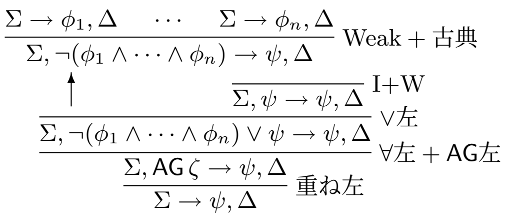


図 8　AG-BC 規則の派生
Fig. 8　Deriving AG-BC rule.

明により示せる．その概要を**図 9** に示す．

5.　強化学習との結合のモデル化

本章ではもう少し進めて，2 章で述べたような強化学習と 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．具体的には，強化学習におけるグリーディ方策や学習の過程などを，$\mathcal{TOMATO}$ による証明過程として記述する．

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，5.1 節に述べるように，ここに Prolog の組み込み述語のような仕組み（ただちに導けてほしい論理式をそのまま axiom として用意しておく）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装方式で，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．現在のエージェント開発システムは，AgentSpeak [13] の実装である Jason [2] など，BDI ベースのものといえど，プランの実行前提条件のチェックなどは論理式で宣言的に書けるものの，その他の部分，特に信念の加除などエージェントの振舞いを規定する部分は手続き的に書かれることが多く，論理式の宣言的意味論と一致するようにはなっていない．

5.1　準備— AG-BC 規則の導入

まず，我々は**図 8** より，以下の派生規則を得る（ここで $\Sigma \ni \mathsf{AG}\,\zeta$，$\zeta = \forall x_1 \cdots \forall x_m \xi$ かつある $t_1, \cdots, t_m$ があって $\xi[x_1 := t_1]\cdots[x_m := t_m] = (\phi_1 \wedge \cdots \wedge \phi_n \supset \psi)$ とする．ただし $m \geq 0$, $n \geq 0$）．

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

以後，$\Sigma \ni \mathsf{AG}\,\zeta$ のとき，Σ が axiom ζ を持つという（あるいは，Σ をエージェントのプログラムと見なして，エージェントが axiom ζ を持つともいう）．するとこの規則は，「あるエージェントが 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，$\phi_1, \phi_2, \cdots, \phi_n$ が成り立つことを結論できればよい」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 AG-Backchain（AG-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 $\mathsf{BEL}^a\,\phi_1, \cdots, \mathsf{BEL}^a\,\phi_n$ から信念 $\mathsf{BEL}^a\,\psi$ を得ること

ができる．ここで，$\mathsf{BEL}^a(\phi_1 \land \cdots \land \phi_n \supset \psi)$（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もっぱら BEL を使うことになろう）．

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とらえることができる．特に，数値の演算や比較などについては，Σ が演算結果に関する無限個の axiom（あるいは，そのうち証明中で必要とする有限個）を直接持つとし，必要な式はただちに Σ から導けると考えることで，（算術に関する公理のような）論理的複雑さを避けつつ扱うことができる．たとえば，あるエージェント a が $90 > 80$ と $\forall x(\mathit{temperature}(x) \land x > 80 \supset \mathsf{INTEND}^a\ \mathit{alarm})$ を axiom として持ち，かつ現時点で $\mathit{temperature}(90)$ をすでに得ているなら，

$$\dfrac{\Sigma \to \mathit{temperature}(90) \qquad \overline{\Sigma \to 90 > 80}\ \textsf{AG-BC}}{\Sigma \to \mathsf{INTEND}^a\ \mathit{alarm}}\ \textsf{AG-BC}$$

より $\mathsf{INTEND}^a\ \mathit{alarm}$ 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．このとき，数値比較の結果である $90 > 80$ は直接 axiom として持たせていることに注意．あるいは，実装を考えるうえでは Σ に $90 > 80$ を明示的には持たせず，代わりに '$\Sigma \to 90 > 80$' をただちに証明可能なシーケントとして扱うこともできる（いい換えれば，Prolog の組み込み述語 is と同様の実現を考える．論理プログラミング的実装を検討するなら，このような仕組みで十分であろう）．

5.2　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（4.1 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 の行動価値が r であることを $\mathit{act\_val}(e, r)$ で表す（他の状態における行動価値，たとえば，状態 s_i での行動 e の行動価値が r であることは，$\mathsf{AG}(at(s_i) \supset \mathit{act\_val}(e, r))$ と表される）．ただしここでは，イベント定数記号 e を（若干 abuse して）一階言語の定数記号としても用いるものとする．

すると，エージェントがグリーディ方策[16] をとることは，（イベント定数記号の集合 $\mathcal{E}$ が $\{e_1, \cdots, e_m\}$ だとして）

$$\dfrac{\dfrac{\Sigma \to \mathit{act\_val}(e_1,4) \qquad \overline{\Sigma \to 4 \le 7}\ \textsf{AG-BC}}{\Sigma \to \mathit{act\_val}(e_1,4) \land 4 \le 7 \lor \mathsf{AX}^{e_1}\ \mathit{dead}}\ \text{Weak} + 古典 \qquad \dfrac{\Sigma \to \mathit{act\_val}(e_2,7) \qquad \overline{\Sigma \to 7 \le 7}\ \textsf{AG-BC}}{\Sigma \to \mathit{act\_val}(e_2,7) \land 7 \le 7 \lor \mathsf{AX}^{e_2}\ \mathit{dead}}\ \text{Weak} + 古典 \qquad \dfrac{\Sigma \to \mathsf{AX}^{e_3}\ \mathit{dead}}{\Sigma \to \mathit{act\_val}(e_3,x) \land x \le 7 \lor \mathsf{AX}^{e_3}\ \mathit{dead}}\ \text{Weak} + 古典}{\Sigma \to \mathsf{INTEND}^a\ \mathit{does}(e_2)}\ \textsf{AG-BC}$$

図 10　推論による動作記述へのグリーディ手法の取り込み
Fig. 10　Importing greedy method info the action description by deduction.

$$\forall r_1 \cdots \forall r_m \left(\bigwedge_{1 \le i \le m} (\mathit{act\_val}(e_i, r_i) \land r_i \le r_j \lor \mathsf{AX}^{e_i}\ \mathit{dead}) \supset \mathsf{INTEND}^a\ \mathit{does}(e_j) \right) \tag{1}$$

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 つの異なる行動の行動価値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すべての行動のうち，行動価値が最大の行動を e_j とし，その行動をとることを意図する」を意味する．また，$\mathit{does}(e)$ は「基本行為 e を実行する」を表す論理式であり，合理的エージェントへの要請として，これを意図すれば（$\mathsf{INTEND}^a\ \mathit{does}(e)$ が成り立てば）行動 e が実際に外界に向かって起こされる[14] ものとする（あるいは，$\mathsf{INTEND}^a\ \mathit{does}(e_j)$ を $\mathsf{INTEND}^a(\neg\,\mathsf{AX}^{e_j}\ \mathit{dead} \land \bigwedge_{e \in \mathcal{E}, e \ne e_j} \mathsf{AX}^e\ \mathit{dead})$ の略記のごとくとらえてもよい．すなわち，この意図によって e_j を実行する未来のみが選択される）．

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 $\mathit{act\_val}(e_1, 4)$, $\mathit{act\_val}(e_2, 7)$, $\mathsf{AX}^{e_3}\ \mathit{dead}$ がその axiom 群からの結論としてすでに得られているとする．このとき，AG-BC 規則を用いた**図 10** の証明図（ここで Σ は少なくとも，$m = 3$, $j = 2$ の場合の式 (1) と，$4 \le 7$, $7 \le 7$ を axiom として持つとする．それ以外の axiom も持ちうる）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とる）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（以降，axiom の

最も外の全称束縛は適宜省いて書く)，$(random(x) \wedge x \geq 0.1 \supset$ グリーディ手法をとる$)$ $\wedge(random(x) \wedge x < 0.1 \supset$ ランダムに行動を選択$)$ などと書けばよい．

また，強化学習の課題においては，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，その遷移の際の即時報酬の推定値 r_s，および遷移後の各状態の状態価値の推定値 v_s を既知として，その行動の行動価値を式 $\sum_s p_s(r_s + \gamma v_s)$ （γ はあらかじめ決めた割引率）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）を持っていれば同様に可能である（現在の状態の状態価値が v であることを $st\_val(v)$ で表す）．

$$\mathsf{X}^e(_{\geq p_1} at(s_1) \wedge reward(r_1) \wedge st\_val(v_1) \mid \cdots \mid _{\geq p_n} at(s_n) \wedge reward(r_n) \wedge st\_val(v_n)) \wedge \sum_{1\leq i\leq n} p_i = 1 \wedge \sum_{1\leq i\leq n} p_i(r_i + \gamma v_i) = r \supset act\_val(e, r)$$

5.3　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[16] 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行動価値の推定値をもとに（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，元の状態での今とった行動の行動価値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 をとることの行動価値の推定値を $Q(s, e)$ 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

$$Q(s, e) := Q(s, e) + \text{差分}$$

として行われる．ここで「差分」は，行動 e をとった結果の即時報酬や次状態などによって決まり，また学習アルゴリズムによっても異なる．

この場合，行動価値の推定値は可変となるので，本節では，5.2 節と違って行動価値に関する情報は信念として（たとえば $\mathsf{BEL}^a(at(s) \supset act\_val(e, r))$ 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 5.2 節の式 (1) とは違い，たとえば $\mathsf{BEL}^a\, at(s) \wedge \bigwedge_{1\leq i\leq m}(\mathsf{BEL}^a(at(s) \supset act\_val(e_i, r_i)) \wedge r_i \leq r_j \vee \mathsf{AX}^{e_i}\, dead) \supset \mathsf{INTEND}^a\, does(e_j)$ のように，行動価値に関しては信念を用いることになる[⋆1]．

⋆1 ここでは，現在の状態に関する情報も信念とした．また，その他の部分，たとえば「行動 e が可能か」といった情報も信念とすることも考えられる．

そして図 10 と同様に行動を決定し，その行動によって行動価値の信念を更新するのだが，差分は観測された次状態や即時報酬といった，行動の結果として次の時刻にならないと分からない情報を用いて計算するので，現時刻では行動価値の更新はできない．そこで，axiom として次のもの

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

を用意しておく．$calc\_delta(d)$ は「計算された差分の値が d である」を表す．この axiom と $\mathsf{BEL}^a\, at(s)$, $\mathsf{BEL}^a(at(s) \supset act\_val(e, r))$, および前節と同様の過程で得られる行動の決定 $\mathsf{INTEND}^a\, does(e)$ 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としては，axiom の集合 Σ に，($\mathsf{AG}\xi$ でなく) ξ を次の時刻でのみ臨時に加える（ように実装する）．すると，次の時刻で差分が計算されて $calc\_delta(d)$ および $r + d = r'$ が得られれば

$$\frac{\Sigma \rightarrow calc\_delta(d) \quad \Sigma \rightarrow r + d = r'}{\Sigma \rightarrow \mathsf{BEL}^a(at(s) \supset act\_val(e, r'))}$$

より $\mathsf{BEL}^a(at(s) \supset act\_val(e, r'))$ を新たな信念とすることができる（この場合 $\Sigma \ni \mathsf{AG}\xi$ でなく $\Sigma \ni \xi$ だが，AG-BC とまったく同じ形の規則が使えることは容易に得られる．Σ が ζ を永続的に成り立つ axiom として持つことが $\Sigma \ni \mathsf{AG}\xi$，その時刻のみの axiom として持つことが $\Sigma \ni \xi$ だととらえることができる）．

このほかに axiom として次のものを用意しておけば，

$$\neg(\mathsf{BEL}^a\, at(s) \wedge \mathsf{INTEND}^a\, does(e)) \wedge \mathsf{BEL}^a(at(s) \supset act\_val(e, r)) \supset \mathsf{AX}^e\, \mathsf{BEL}^a(at(s) \supset act\_val(e, r)) \tag{2}$$

現在の状態と今実行した行動の対以外に関する行動価値の信念は更新されずに保たれる．

5.4　学習結果の BDI エージェントによる利用

我々の目的は，2 章で述べたような，熟考的行為と反射的行為の双方の能力を取り込んだエージェントのモデル化と実現にある．そのためには，前節までのように単に強化学習のメカニズムを確率的遷移による時相論理で表現できるだけではなく，これと BDI モデルによる行動選択の過程を結合して，1 つの論理モデルの中で書けることが求められる．

一例として，我々がテストベッドに用いているカヌーレーシング[20] の事例において，エージェント a がゴールに到達したいという願望を持ち，その達成のためには流れが速いか遅

いかなどでプランを使い分けるとする．たとえば，「中間地点 s_i, s_j を経てゴールへ着く」というプラン $planX$ があり，流れが速くかつエージェント b が同じルートを通ろうとしていないと考えられるならばこのプランを用いるとしよう．そして，地点間の移動には強化学習で得た技能を利用するものとする[⋆1]．なお以後，3.1.2 項で導入した until や AF オペレータを用い，さらに $\mathsf{A}(\phi\,\mathsf{N}\,\psi)$ は $\mathsf{A}(\neg\psi\,\mathsf{U}\,(\phi\wedge\psi))$ の略記（「次に ψ が成り立つとき，ϕ も成り立つ」を表す．atnext オペレータと呼ばれるもの）とする．また，$at(s)$ や $flow(fast)$ などの原始論理式は，記号の名前から想定できる意味を持つものとする．

$planX$ は（簡単のため，意図が達成不能になったという信念を得た場合の意図の放棄については記述を省くと）おおむね

$$\mathsf{INTEND}^a\,\mathsf{AF}\,achieve\_planX \supset stage1$$

$$stage1 \supset \mathsf{INTEND}^a\,\mathsf{AF}\,at(s_i) \qquad stage1 \supset \mathsf{A}(stage2\,\mathsf{N}\,\mathsf{BEL}^a\,at(s_i))$$
$$stage2 \supset \mathsf{INTEND}^a\,\mathsf{AF}\,at(s_j) \qquad stage2 \supset \mathsf{A}(stage3\,\mathsf{N}\,\mathsf{BEL}^a\,at(s_j))$$
$$stage3 \supset \mathsf{INTEND}^a\,\mathsf{AF}\,at(goal) \qquad stage3 \supset \mathsf{A}(\mathsf{BEL}^a\,achieve\_planX\,\mathsf{N}\,\mathsf{BEL}^a\,at(goal))$$

のように表現でき，またこのプランの利用は

$$\mathsf{DESIRE}^a\,\mathsf{AF}\,at(goal) \wedge \mathsf{BEL}^a\,flow(fast) \wedge \mathsf{BEL}^a\,\neg\,\mathsf{INTEND}^b\,achieve\_planX \supset \mathsf{INTEND}^a\,\mathsf{AF}\,achieve\_planX$$

のように書ける．これらを axiom として持つエージェントは，$\mathsf{DESIRE}^a\,\mathsf{AF}\,at(goal)$, $\mathsf{BEL}^a\,flow(fast)$,$\mathsf{BEL}^a\,\neg\,\mathsf{INTEND}^b\,achieve\_planX$ を成り立たせれば，

$$\dfrac{\dfrac{\dfrac{\Sigma \to \mathsf{DESIRE}^a\,\mathsf{AF}\,at(goal) \quad \Sigma \to \mathsf{BEL}^a\,flow(fast) \quad \Sigma \to \mathsf{BEL}^a\,\neg\,\mathsf{INTEND}^b\,achieve\_planX}{\Sigma \to \mathsf{INTEND}^a\,\mathsf{AF}\,achieve\_planX}}{\Sigma \to stage1}}{\Sigma \to \mathsf{INTEND}^a\,\mathsf{AF}\,at(s_i)}$$

より，まずは $\mathsf{INTEND}^a\,\mathsf{AF}\,at(s_i)$ を導く．ここまでは BDI モデルのみによる意思決定である．

ここでエージェントは，強化学習で獲得済みの，カヌーを漕いで目的地に向かうための技能を使って s_i を目指すのだが，そのような技能を状況に応じていくつか（別々の学習で）獲得しており，流れが速いときに s_i に向かう場合は技能$_1$ を使うとする．このためには，$\mathsf{INTEND}^a\,\mathsf{AF}\,at(s_i) \wedge \mathsf{BEL}^a\,flow(fast) \supset \mathsf{A}(use\_skill(1)\,\mathsf{U}\,\mathsf{BEL}^a\,at(s_i))$ という axiom を与えておく．すると，上記に続いて $\mathsf{A}(use\_skill(1)\,\mathsf{U}\,\mathsf{BEL}^a\,at(s_i))$ が導かれる．このとき，実装としては $\mathsf{BEL}^a\,at(s_i)$ が成り立つ時刻まで $use\_skill(1)$ を axiom に臨時に追加する（until の扱いについては，厳密には Temporal Prolog[19] のような実行形態が必要となろう）．

学習結果は，$\mathsf{AG}(use\_skill(n) \supset (at(s) \supset act\_val(e,r)))$ という形の axiom を各 n，s，e について必要なだけ持っておくことで表す．すると，あとは現在の位置に関する情報 $at(s)$ があれば，現時刻の各行動 e の行動価値を知ることができ，5.2 節式 (1) のようなグリーディ行動の axiom を用いることで，図 10 と同様に s_1 に達するまでの行動を決定することができる．

以上のように，強化学習と BDI モデルの結合を $\mathcal{TOMATO}$ を用いて 1 つの論理モデルの中で形式化することができる．

6. 考　　察

前章まで，$\mathcal{TOMATO}$ で用いている Kripke 構造である BDI ストラクチャにおける，確率的状態遷移に関するモデル化の概念と推論規則について詳細に述べた．また，確率的状態遷移に関する事例について，$\mathcal{TOMATO}$ を用いた証明を与えた．しかし，本論文では扱わなかったが今後検討を要すると考えられる課題も多い．本章では，まず $\mathcal{TOMATO}$ の適用範囲を述べ，次いで課題のいくつかについて述べる．

6.1 $\mathcal{TOMATO}$ の適用範囲

確率論における Kolmogorov の「確率の公理」[7] は以下の 3 点である．

（1） 任意の事象 E について $0 \leq P(E) \leq 1$

（2） 全事象 Ω について $P(\Omega) = 1$

（3） $\{E_i\}_{i\in\mathbb{N}}$ が可算無限個の排反な事象ならば $P(\bigcup_{i\in\mathbb{N}} E_i) = \sum_{i\in\mathbb{N}} P(E_i)$

3.3.2 項の $Q \subseteq \Omega$ は，上記の任意の事象 E に対応しており，$\sum_{Q\subseteq\Omega} v(Q) = 1$ を満たす確率分配関数 v は，有限事象を対象とした上記の確率の公理を適用して作成されている．この v で分配された Q を要素に持つ本質的な要充足集合は，上記の確率の公理を満足する次時刻に生起する事象の集合を表現したものであり，X^e オペレータは，確率の公理を満足する確率的状態遷移に適用することができる．よって，$\mathcal{TOMATO}$ は有限事象を対象とした上記の確率の公理を満足する確率的状態遷移を適用範囲とする．つまり，確率的な状態遷移を持つ有限オートマトンを対象としている．

⋆1 この例では地点間の移動には学習の結果を用いるとしているが，もちろん応用によっては，学習を用いず基本行為の部分まで熟考で得たプランによって決めることもありうる．

また，$\mathcal{TOMATO}$ の推論規則は，BDI ストラクチャや論理式の解釈に依存していないため，タブロー法を用いて推論規則を逆向きに適用することで，前提に初期シーケントや証明済みの論理式を持ち，結論に証明したい論理式を持つ証明図を作成できれば，その論理式が恒真，つまり，定理であることが証明できる．

しかし，現時点の $\mathcal{TOMATO}$ では，たとえば $\mathsf{X}^{e_1}_{\geq .6}\,\mathsf{X}^{e_2}_{\geq .4}\,\phi \supset \mathsf{X}^{e_1 e_2}_{\geq .24}\,\phi$ のような記述ができない．すなわち，逐次に実行されるイベント $e_1 e_2$ に関する確率を扱うことができない．このような論理式を扱うためには，逐次イベントの構文，意味論，および，同時確率に関する「乗法の定理」を満足する推論規則を導入する必要がある．

以上をまとめると，$\mathcal{TOMATO}$ は，Kolmogorov の確率の公理を満足する有限事象を対象とした確率的状態遷移における（逐次でない）単独のイベントに関する定理を証明できる．

6.2 課　　題

6.2.1 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，3.3.2 項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あげられる．

PCTL [5] 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．たとえば，「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，2 章でも述べたが，PCTL のような方式では，$\mathcal{TOMATO}$ のようにイベントごとに確率を記述することは難しい．PCTL や，同様に確率を扱えてしかも連続時間論理である CSL [1] を記述言語とし，MDP や連続時間 Markov 連鎖を対象とするモデル検査ツール PRISM [9] のような例もあるが，これも記述言語が PCTL や CSL である限り，モデルの方に複数のイベントが入っていても，記述言語側ではその区別を扱えないことになる．我々の $\mathcal{TOMATO}$ は，強化学習を含め，エージェントの行為の選択を扱うことを目的とするため，本質的にイベントの区別の記述が必要になる．

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[⋆1]．証明系の構築と記述力のバランスは課題の 1 つである．

⋆1 確率の記述を 0 と 1 に制限した qualitative PCTL でさえ，文献 3) の時点で演繹体系は（タブロー法によるもの以外も含め）まだ知られていない．

6.2.2 本質的な要充足集合の自動構築

4 章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

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

に解があるかどうかで分かるので，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すべての要充足集合を，したがってすべ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．たとえば p_i がすべ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．

6.2.3 心的状態の維持の扱い

BDI モデルのように，心的状態を明示的概念として持つエージェントモデルにおいては，心的状態は，デフォルトでは時間の経過にともなって保持されるべきだと考えられる．しかし，BDI logic にとっての心的状態は，単に時刻によって異なる可視世界をともなう様相オペレータであり，ある時刻とその次の時刻の信念には論理的にまったく関係はない．したがって，信念が維持されるという自然な性質を表現するためには，わざわざ nonlogical axiom（たとえば 5.3 節式 (2) のような）を導入しなければならず，論理体系としては不自然さを否めない．

一方で，多くの BDI エージェントの実装では，信念など心的状態は，`add-belief` や `del-belief` などのイベントにより手続き的にデータベースに加除されるものであり，保持は容易だが，論理的な扱いとの整合は保証されない．AgentSpeak(L) [13] のように，手続き的な心的状態の加除に対し，それが満たすべき性質に関する証明論を提供することで両者のギャップを埋めているものもあるが，動作が論理式の宣言的意味論で与えられるようには必ずしもなっていないという点は解消されない．また，心的状態が一階述語論理式に制限される欠点もある．

信念がデフォルトで維持されるような論理の試みには，たとえば文献 15) がある．ただこれは非分岐時相論理であり，記述力の不足が懸念される．

また，文献 6) のように心的状態の更新を時刻の変化ではなくモデル自体の更新として扱うものもある．本論文 5.3 節では，Kripke 構造での信念世界を時間の経過に従って順次決めていく形で信念の更新を表現しているが，今後はこうした手法との比較なども必要であろう．

6.2.4　エージェントの実装への応用

5.1 節以降に述べた方式によれば，𝒯𝒪ℳ𝒜𝒯𝒪 を実行可能なエージェント記述言語として応用する可能性も考えられる．これも今後の検討課題の 1 つである．

7. 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 𝒯𝒪ℳ𝒜𝒯𝒪 を用いて，確率的状態遷移をともな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．また，今後の課題を 6 章で述べた．提案した 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) Aziz, A., Sanwal, K., Singhal, V. and Brayton, R.: Verifying Continuous Time Markov Chains, *Proc. Conference on Computer-Aided Verification*, pp.269–276 (1996).
2) Bordini, R.H., Hübner, J.F. and Wooldridge, M.: *Programming Multi-Agent Systems in AgentSpeak using Jason*, John Wiley & Sons (2007).
3) Brázdil, T., Forejt, V., Křetínský, J. and Kučera, A.: The Satisfiability Problem for Probabilistic CTL, *Proc. 23rd Annual IEEE Symposium on Logic in Computer Science*, pp.391–402 (2008).
4) Dam, M.: Translating CTL into the modal μ-calculus, Technical Report ECS-LFCS-90-123, Laboratory for Foundations of Computer Science, Department of Computer Science, University of Edinburgh (1990).
5) 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).
6) Kobayashi, M. and Tojo, S.: Agent Communicability in Belief Update Logic, *Proc. DALT 2008*, pp.206–221 (2008).
7) Kolmogorov, A.N.: *Foundations of the Theory of Probability*, 2nd edition, Chelsea Publishing (1960).
8) Kozen, D.: Results on the propositional μ-calculus, *Theoretical Computer Science*, Vol.27, pp.333–354 (1983).
9) Kwiatkowska, M., Norman, G. and Parker, D.: Probabilistic Symbolic Model Checking with PRISM: A Hybrid Approach, *Proc. 8th International Conference on Tools and Algorithms for the Construction and Analysis of Systems*, pp.52–66 (2002).
10) Manolios, P.: Mu-Calculus Model-Checking, *Computer-Aided reasoning: ACL2 case studies*, pp.93–111, Kluwer Academic Publishers (2000).
11) NIDE, N., Takata, S. and Fujita, M.: BDI logic with probabilistic transition and fixed-point operator, *Proc. CLIMA '09*, pp.71–86 (2009).
12) Puterman, M.L.: *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, John Wiley & Sons (1994).
13) Rao, A.S.: AgentSpeak(L): BDI agents speak out in a logical computable language, *Proc. MAAMAW-96*, LNAI, Vol.1038, pp.42–55, Springer-Verlag (1996).
14) Rao, A.S. and Georgeff, M.P.: Modeling Rational Agents within a BDI-Architecture, *Reading in Agents*, Huhns, M.N. and Singh, M.P. (Eds.), pp.317–328, Morgan Kaufmann, San Francisco (1997).
15) Su, K., Sattar, A., Wang, K., Luo, X., Governatori, G. and Padmanabhan, V.: Observation-based Model for BDI-Agents, *Proc. AAAI 2005*, pp.190–195 (2005).
16) 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).
17) Tarski, A.: A Lattice-Theoretical Fixpoint Theorem and its Application, *Pacific Journal of Mathematics*, Vol.5, pp.285–309 (1955).
18) Wooldridge, M.: *Reasoning about Rational Agents*, The MIT Press (2000).
19) 桜川貴司：Temporal Prolog: A Programming Language Based on Temporal Logic, コンピュータ・ソフトウェア，Vol.4, No.3, pp.199–211 (1987).
20) 高田司郎，新出尚之，山川　宏，宮崎和光，太田正幸：強化学習で獲得したスキルを実践的推論する BDI の実現方式について，エージェント合同シンポジウム（JAWS2004）講演論文集，pp.517–524 (2004).
21) 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).
22) 新出尚之，高田司郎，山川　宏，宮崎和光，太田正幸：BDI と強化学習の世界モデルの対応付けについて，エージェント合同シンポジウム（JAWS2004）講演論文集，pp.378–385 (2004).

(平成 22 年 11 月 18 日受付)
(平成 23 年 1 月 25 日再受付)
(平成 23 年 3 月 25 日採録)

新出　尚之

1986 年京都大学理学部卒業．1988 年同大学院理学研究科数理解析専攻修士課程修了．同年同専攻博士課程中退．同年京都大学情報処理教育センター助手．1992 年奈良女子大学理学部情報科学科講師．2008 年同准教授，現在に至る．博士（情報科学）．興味を持つ分野は，時相論理による証明およびプログラミングシステム，特に自律エージェントシステムの構築．日本ソフトウェア科学会，電子情報通信学会，日本ロボット学会各会員．

高田　司郎（正会員）

1979 年大阪大学基礎工学部情報工学科卒業．1993 年奈良先端科学技術大学院大学情報科学研究科博士前期課程入学．1999 年同科博士後期課程修了．1979 年 CSK 入社．1993 年けいはんな入社．1999 年 ATR 知能映像通信研究所入所．2001 年 ATR メディア情報科学研究所客員研究員．2002 年福岡工業大学情報工学部管理情報工学科助教授．2003 年より近畿大学理工学部情報学科助教授（現，准教授）．博士（工学）．知能ロボット，コミュニケーション，形式的仕様記述，合理的エージェントに興味を持つ．日本ソフトウェア科学会，人工知能学会，日本ロボット学会，言語処理学会各会員．

藤田　　恵

2006 年奈良女子大学理学部卒業．2009 年同大学院人間文化研究科博士前期課程情報科学専攻修了．同年同大学院人間文化研究科博士後期課程複合現象科学専攻入学，現在に至る．興味を持つ分野は，知的エージェントの構築，知的エージェントによるロボット制御．日本ソフトウェア科学会，人工知能学会，日本ロボット学会各学生会員．