

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

An Attempt to Construct Logical Model of Rational Agents Acting in Continuous State Space

新出 尚之
NIDE, Naoyuki

奈良女子大学大学院自然科学系
Division of Natural Sciences, Faculty, Nara Women's University
nide@ics.nara-wu.ac.jp

高田 司郎
Shiro TAKATA

近畿大学理工学部
Faculty of Science and Engineering, Kinki University
shiro@info.kindai.ac.jp

藤田 恵
Megumi FUJITA

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
saboten@ics.nara-wu.ac.jp

keywords: Agent framework, Rational agents, BDI logic, Real world

Summary

A major different between the real world and the virtual world, for rational agents act in the former in response to the changes of their environments, is that the environment is continuous rather than discrete. Thus, for modeling of such agents, it is natural and desirable to use a model in which the probability distribution of the reachable states as a result of some act is continuous. In this paper, we propose a logical system $C\text{-}\mathcal{ITCMA}\mathcal{ITCes}$, a further extension of the extended BDI logic $\mathcal{ITCMA}\mathcal{ITCes}$ which we had proposed, to make an attempt to handle the continuous state space. In $C\text{-}\mathcal{ITCMA}\mathcal{ITCes}$, we can describe probability density functions of the states reachable from the current state. We present some examples of descriptions about probabilistic move on the plane of coordinates.

1. はじめに

我々は、実世界における合理的エージェントの行為選択に関する研究を進めている。

実世界で身体性を持って機能するエージェントは、環境の多様な変化に対応して、目的を整合的に達成していく必要がある。また、①情報獲得に時間がかかる ②得られる情報が完全でない ③センサのノイズが不可避 ④実世界を離散的には記述できない ⑤複数のタスクを並行して扱わねばならない ⑥実世界は固有のダイナミズムで変化し続ける ⑦実世界は予測不能 — などの問題に直面することが指摘されている [Pfeifer 06b]。

これまでに我々は、それらの問題への部分的解答として、BDI モデルを実装した BDI ロボットの有効性を示した [藤田 12a, 藤田 12b]。BDI モデルは、熟考やコミットメントなどの概念を持ち、これによって、実世界のエージェントないしロボットが、単一の目標達成に特化するのではなく、環境の多様性に対応して合理的かつ柔軟に振る舞う、より一般的な行為選択のモデルを提供する。また、上記の①②⑤への対応が可能なることも示した。

一方、BDI モデルのみでは実現しにくい反射的行動に関する技能の学習を、強化学習で補う手法について、我々はカヌーレーシングをテストベッドとして取り組み、強化学習の獲得方策から生成したサブプランを BDI での意思決定に用いる手法を示した [高田 10, 新出 10]。

これらの研究において、エージェントにおける環境は離散な状態空間 (グリッド) を用いた仮想世界上のものとして扱ったが、実世界を扱うには、離散的な記述ではエージェントが実世界における環境を正しく認識することが十分にできない。この点、すなわち前記④への対処は、実世界におけるエージェントがよりよく機能するために残された課題の 1 つである。そこで我々は次のステップとして、連続な状態空間を持つ仮想世界においてエージェントの持つ意思決定を実世界での行為選択と統合して扱う手法を提案している [高田 12]。

さて、BDI モデルには、その 1 構成要素として様相論理体系 BDI logic があり、エージェントの性質や振る舞いなどに関する形式的記述を行うことができる。これが BDI モデルのもう 1 つの利点でもある。しかし、BDI モデルと強化学習など他の行為決定手段との結合において、

後者が形式的記述から外れてしまうことは、この利点をそぐため望ましくない。このため我々はこれまでに、BDI logic を拡張して、イベントによる状態遷移に確率を記述できるようにした論理体系 $\mathcal{TOMATOCes}$ [新出 11a] を導入し、また、これによって強化学習と BDI の結合による意思決定がモデル化できることを示した [高田 11]。しかし、連続的な状態空間を扱うには、行為の結果到達する状態が連続的な確率分布をするような論理モデルを用いることが自然かつ望ましいが、それはまだできていなかった。

そこで本論文では、 $\mathcal{TOMATOCes}$ をさらに拡張した論理体系 $\mathcal{C-TOMATOCes}$ (Continuous $\mathcal{TOMATOCes}$) を導入し、連続的な状態空間と、行為の結果到達する状態が連続的な確率分布をするようなモデルを扱うことを試みる。さらに、いくつかの記述例を示すとともに、現状での問題点についても述べる。

2. 論理体系 $\mathcal{C-TOMATOCes}$

本節では様相論理体系 $\mathcal{C-TOMATOCes}$ を導入する。この論理は、点時刻分岐時相論理に、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 (信念・願望・意図) オペレータ、およびあるイベントによる状態遷移先の確率密度を表現するためのオペレータを導入したものである。

2.1 確率密度の記述の導入

厳密な構文や意味論については 2.2 節以降に述べ、ここでは $\mathcal{C-TOMATOCes}$ で導入した状態遷移先の確率密度に関するオペレータについて説明する。

$\mathcal{TOMATOCes}$ では、論理式 $X^e(\geq_{.3} \phi_1 | \geq_{.5} \phi_2)$ で、「イベント e によって 1 時刻経過すると、0.3 以上の確率で ϕ_1 が成り立つ状態へ、0.5 以上の確率で ϕ_2 が成り立つ状態へ遷移する (しかも、両者の状態を重ならないように取ることができる。図 1 の右ではなく左のように)」を表現していた。 $\mathcal{C-TOMATOCes}$ ではこの部分を確率密度

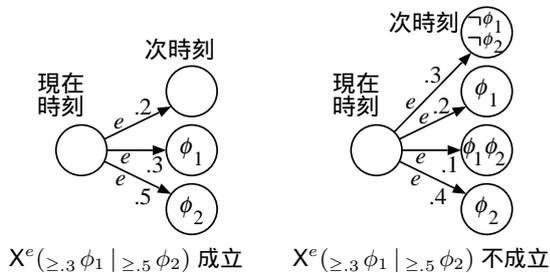


図 1 $\mathcal{TOMATOCes}$ の X^e オペレータ

関数を書けるように拡張し、自由変数 x が現れる項 τ と論理式 ϕ を用いて $X_{@_{\tau}}^e \phi$ と書けば「イベント e で 1 時刻後へ遷移したとき、 ϕ を満たす x の値の確率密度は τ である」を表すものとする。 $\mathcal{TOMATOCes}$ との構文的な差

はこれだけである。

例えば*1、平均 μ で分散 σ^2 の正規分布の確率密度関数 $\mathcal{N}(x | \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$ に対し、いま、3 引数の関数記号 norm があって、項 $\text{norm}(x, \mu, \sigma^2)$ は $\mathcal{N}(x | \mu, \sigma^2)$ と*2解釈されるよう固定しておくものとする。すると、 $\mathcal{C-TOMATOCes}$ では $X_{@_{\text{norm}(x, \mu, \sigma^2)}}^e \text{loc}(x)$ と書くことで「イベント e で 1 時刻遷移した後では、 $\text{loc}(x)$ を満たす x の値は平均 μ で分散 σ^2 の正規分布をする」(図 2) が記述できる ($\text{loc}(x)$ を満たす x の値は 1 状態あたりでは 1 つしかないものとする)。 $\text{loc}(x)$ が何らかの位置情報を表す論理式だとすれば、これによって、連続的な確率分布を行うような位置移動が表現できることになる。

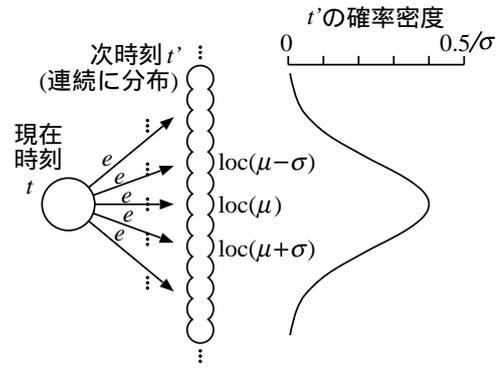


図 2 $\mathcal{C-TOMATOCes}$ の確率分布遷移オペレータ

先に述べた $\mathcal{TOMATOCes}$ の X^e オペレータの一般形、 $X^e(\geq_{p_1} \phi_1 | \dots | \geq_{p_n} \phi_n)$ (ただし p_1, \dots, p_n は定数) は、確率密度関数の記述に Dirac の Delta 関数 [Weisstein] を用いれば、 $\mathcal{C-TOMATOCes}$ の $X_{@}^e$ でも記述できる。すなわち、2 引数関数記号 deltasum の解釈を

$$\text{deltasum}(x, [p_1, \dots, p_n]) = \sum_{i=1}^n p_i \delta(x - i) + (1 - \sum_{i=1}^n p_i) \delta(x)$$

となるよう*3固定しておくことで、(ϕ に自由に出現しない変数記号 x を 1 つ選んで)

$$X_{@_{\text{deltasum}(x, [p_1, \dots, p_n])}}^e \bigwedge_{i=1}^n (x = i \supset \phi_i) \quad (1)$$

と書ける (なお、ここを含めてこれ以後、論理式に $+$, $-$, \leq , $<$ など通常の数学で用いられる記号が述語記号や関数記号として現れる場合、解釈も通常の数学でのそれらに固定しておくものとする)。図 3 に式 $X^e(\geq_{.3} \phi_1 | \geq_{.5} \phi_2)$ の場合を示した。

以上から、 $X_{@}^e$ は従来の $\mathcal{TOMATOCes}$ の確率的遷移オペレータを一般化したものと考えられることができる。以後 $\mathcal{C-TOMATOCes}$ では、 $X^e(\geq_{p_1} \phi_1 | \dots | \geq_{p_n} \phi_n)$ は式 (1) の

*1 以下では「 σ^2 」を、文法上は単一の定数記号と見なすものとする。

*2 正しくは、解釈を与える記号 $[[]]$ (2.3.2 節) を用いて $\mathcal{N}([x] | [\mu], [\sigma^2])$ と書くべきであるが、ここでは粗い議論をする。

*3 ここでの $[]$ は Prolog のリストにあたるものである。あるいは単に n -tuple の記号と思ってよい。

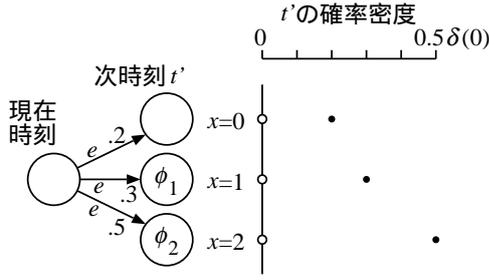


図3 C-TMSTLの X^e オペレータをエミュレート

略記とする。特に、 $n = 1$ の場合はさらに $X_{\geq p_1}^e \phi_1$ と略記する（「イベント e で 1 時刻遷移した後、 ϕ が成り立つ確率は p 以上」を表す）。

また、 $X_{=p}^e \phi$ は $X_{\geq p}^e \phi \wedge X_{\geq 1-p}^e \neg \phi$ の略記とする。これは「イベント e で 1 時刻遷移した後、 ϕ が成り立つ確率は p 」を表す。

2.2 構文

本節では $C\text{-TMSTL}$ の論理式の定義を与える。以下では、単に「論理式」と言えば $C\text{-TMSTL}$ 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（論理式変数記号に対する限量の構文はない）。

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{L} には 2 引数の述語記号として等号 $=$ を含むものとし、また \mathcal{E} , \mathcal{A} は有限集合、 \mathcal{V} は無限集合とする。

以下が $C\text{-TMSTL}$ の構文の定義である。

- 一階述語論理の原始論理式は ($C\text{-TMSTL}$ 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 (\mathcal{L} の) 変数記号ならば $\forall x \phi$ も論理式
 - $e \in \mathcal{E}$ ならば $\text{pos}(e)$ は論理式
 - $e \in \mathcal{E}$ 、 τ が 1 つ以上の自由変数を持つ \mathcal{L} の項、 ϕ が論理式ならば $X_{@ \tau}^e \phi$ も論理式。なおこの場合、 τ の自由変数は、 $X_{@ \tau}^e \phi$ の中では構文上束縛されているものと見なす
 - ϕ が論理式、 $a \in \mathcal{A}$ ならば $\text{BEL}^a \phi, \text{DESIRE}^a \phi$ および $\text{INTEND}^a \phi$ 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、 \mathfrak{x} が ϕ の中で \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。 μ は最小不動点オペレータ [Kozen 83] と呼ばれるもの
- $\text{pos}(e)$ は、直感的には「現時刻でイベント e が実行可能である」を表す。また、 $\text{BEL}^a \phi, \text{DESIRE}^a \phi, \text{INTEND}^a$

ϕ は、各々「エージェント a が ϕ という信念/願望/意図を持つ」を表す。 $X_{@ \tau}^e \phi$ の τ は確率密度関数を表す項である ($X_{@ \tau}^e \phi$ の直感的な意味については 2.1 節に述べた)。

この他、 \wedge や \supset や \Leftrightarrow や \exists を一般的な略記として導入する*4。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 (加えて、 $\wedge \cdot \vee$ は左結合、 \supset は右結合とする)。

不動点オペレータは、CTL*[Emerson 89] 相当の時相オペレータや、相互信念などを記述できるように導入されているものである。本論文では相互信念などの概念は使わないので、CTL* の時相オペレータについてのみ簡単に述べておく。いま、 $AX^e \phi$ を $X_{\geq 1}^e \phi$ の略記とすると、これは「イベント e で 1 時刻後に遷移した先では常に (確率 1 で) ϕ が成り立つ」を表す。さらに $AX \phi$ を $\bigwedge_{e \in \mathcal{E}} (\text{pos}(e) \supset AX^e \phi)$ の略記とする (\mathcal{E} を有限集合としているため、これは論理式として有限に記述できることに注意) と、これは「(可能なイベントによって遷移した) 次の時刻には必ず ϕ が成り立つ」を表し、CTL* での $AX \phi$ に相当するものとなる。あとは、 AX と不動点オペレータを用いることにより、時相論理として CTL* より強い表現力を得ることができる [Manolios 00, Dam 90]。例えば $\mu \mathfrak{x} . (\psi \vee \phi \wedge AX \mathfrak{x})$ を $A(\phi U \psi)$ と略記することにすれば、これで CTL* の until オペレータ（「現在以後 ψ が成り立つまで ϕ が成り立つ」）と同じことが書ける。また、CTL* の $AF \phi$ （「現在以後いつかは ϕ が成り立つ」）も同様の略記として書ける。

2.3 意味論

§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。また、 $=$ の解釈は等号の性質を満たすこと (i.e. 任意の項 T_1, T_2 に対し、 $i_{w,t}$ で $T_1 = T_2$ が真 iff $i_{w,t}$ での T_1 と T_2 の解釈が同じ)。
- 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 $\text{pos}_{w,t}$
- W の各要素 w と \mathcal{E} の各要素 e に対して、 St_w 上の serial な*5 2 項関係 $R_w^e \subseteq St_w \times St_w$ 、および R_w^e か

*4 \supset が略記のため、例えば $\mathfrak{x} \supset \psi$ の中で \mathfrak{x} は \neg の 1 段のネストの中に出現することに注意。

*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。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。

ら $[0, \infty]$ への関数 \mathcal{F}_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、 $\int_{s \in \{t' | t R_w^e t'\}} \mathcal{F}_w^e(t, s) ds$ が定義でき*6、かつその値が1であること

- A の各要素 a と $\bigcup_{w \in W} St_w$ の各要素 t に対し、集合 $\{w | t \in St_w\}$ 上の serial な 2 項関係 B_a^t, D_a^t, I_a^t であって、以下を満たすもの:

[i] $w B_a^t w'$ を満たす任意の $w, w', w'' \in W$ に対し、 $R \in \{B_a^t, D_a^t, I_a^t\}$ のそれぞれについて $w R w''$ iff $w' R w''$ 。特に、 B_a^t は transitive かつ Euclidean であることになる

[ii] $w'' B_a^t w$ を満たす任意の $w'', w \in W$ に対し、 w のある subworld $w' \in W$ があって $w'' D_a^t w'$ 。また、 D_a^t と I_a^t のペアについても同様の性質が成り立つこと

ただし、ここで 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 $t \in St_w$ に対し $i_{w,t} = i_{w',t}$
- 任意の $t \in St_w$ に対し $pos_{w,t} \supseteq pos_{w',t}$
- 任意の $t \in St_w, w'' \in W$ に対し $w B_a^t w''$ iff $w' B_a^t w''$ 。 D_a^t, I_a^t についても同様
- 任意の $e \in \mathcal{E}$ について $R_w^e = R_{w'}^e, \mathcal{F}_w^e = \mathcal{F}_{w'}^e$

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、 $\mathcal{F}_w^e(t, t')$ は確率密度であり、 $t R_w^e t'$ および $\mathcal{F}_w^e(t, t') = f$ は、state t でイベント e を実行すると次の時刻が state t' になる確率密度が f であることを表す。また、 $pos_{w,t}$ は state t で実行可能なイベントの集合を表す。

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。この概念は、BDI モデルにおけるエージェントの合理性 (心的状態の整合性) を扱うために導入されたものである。本論文では詳細は略す。

B_a^t, D_a^t, 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 4 に BDI ストラクチャの概略を示した (概略図のため、 B_a^t などを全て示したのではなく、 R_w^e や B_a^t の serial 性なども反映していない)。各 state、および各イベント e 毎に、その state からの R_w^e による時間経過 (状態遷移) が、図 2 と同様の確率分布をしている。state の円内に書かれているイベントはその state で可能なイベント、状態遷移の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 (その state で可能なイベントでの遷移のみ図示した)。 w_3 は w_2 の、 w_2 は w_1 の、 w_5 は w_4 の subworld である (w_1 の subworld でない w_4

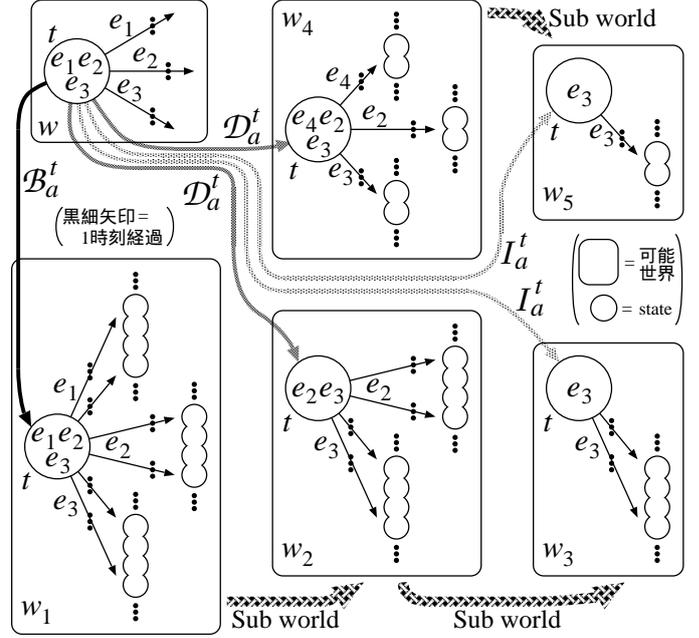


図 4 BDI ストラクチャの概略

が存在しても構わない)。

§2 論理式の解釈

以後、 $\{(w, t) | w \in W, t \in St_w\}$ 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 ($[\phi]_{\langle M, f_{\mathcal{V}} \rangle} \subseteq Swt$ である)。

- ϕ が原始論理式の時、 $[\phi]_{\langle M, f_{\mathcal{V}} \rangle} = \{(w, t) | M$ での $i_{w,t}$ で ϕ が真}
- $[\phi \vee \psi]_{\langle M, f_{\mathcal{V}} \rangle} = [\phi]_{\langle M, f_{\mathcal{V}} \rangle} \cup [\psi]_{\langle M, f_{\mathcal{V}} \rangle}$
- $[\neg \phi]_{\langle M, f_{\mathcal{V}} \rangle} = Swt \setminus [\phi]_{\langle M, f_{\mathcal{V}} \rangle}$
- $[\forall x \phi]_{\langle M, f_{\mathcal{V}} \rangle} = \bigcap_{u \in U} [\phi]_{\langle M^u, f_{\mathcal{V}} \rangle}$ ここで、 M^u は M の全ての state 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
- $[pos(e)]_{\langle M, f_{\mathcal{V}} \rangle} = \{(w, t) | pos_{w,t} \ni e\}$
- $[X_{@ \tau}^e \phi]_{\langle M, f_{\mathcal{V}} \rangle} \ni (w, t)$ iff τ に出現する自由変数 x_1, \dots, x_n の解釈の n -tuple $\langle v_1, \dots, v_n \rangle$ の集合を \mathcal{V} 、 $\{t' | t R_w^e t'\} = \mathcal{T}$ とするとき、 \mathcal{T} から \mathcal{V} へのある写像 \mathcal{M} があって、任意の $v = \langle v_1, \dots, v_n \rangle \in \mathcal{V}$ に対し、以下の 2 条件を満たす: ① $\mathcal{M}(t') = v$ なる任意の $t' \in \mathcal{T}$ に対し $[\phi]_{\langle M^v, f_{\mathcal{V}} \rangle} \ni (w, t')$ ② $\sum_{t' \in \mathcal{T}, \mathcal{M}(t')=v} \mathcal{F}_w^e(t, t') = [\tau]_{\langle M^v, f_{\mathcal{V}} \rangle}^{(w, t)}$ 。ただし M^v は M の全ての state での x_1, \dots, x_n の解釈を v_1, \dots, v_n に変更して得られる BDI ストラクチャとし、 $[\tau]_{\langle M^v, f_{\mathcal{V}} \rangle}^{(w, t)}$ とは τ の $\langle M^v, f_{\mathcal{V}} \rangle$ および (w, t) での解釈とする
- $[BEL^a \phi]_{\langle M, f_{\mathcal{V}} \rangle} = \{(w, t) | w B_a^t w' \text{ なる任意の世界 } w' \text{ に対して } (w', t) \in [\phi]_{\langle M, f_{\mathcal{V}} \rangle}\}$
- $[DESIRE^a \phi]_{\langle M, f_{\mathcal{V}} \rangle}, [INTEND^a \phi]_{\langle M, f_{\mathcal{V}} \rangle}$ については同様

*6 この積分の定義は、状態空間に適切に位相や測度を入れておき、そのもとで行うものとする。

• $x \in \mathcal{V}$ のとき、 $[[x]]_{\langle M, f_V \rangle} = f_V(x)$

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

• $[[\mu x. \phi]]_{\langle M, f_V \rangle}$ 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 [Tarski 55]。以上で論理式の解釈の定義を終わる。

$X_{@t}^e \phi$ の解釈は、2.1 節の $X_{@norm(x, \mu, \sigma^2)}^e \text{loc}(x)$ の例で言えば、 \mathcal{M} が実数 x に $\text{loc}(x)$ を満たす状態 t' を対応させる写像、条件①は状態 $\mathcal{M}(x)$ で実際に $\text{loc}(x)$ が成り立つこと、条件②は t から $\mathcal{M}(x)$ へ遷移する確率密度が $\text{norm}(x, \mu, \sigma^2)$ で与えられることを要請している。

$[[\phi]]_{\langle M, f_V \rangle}$ 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。 $[[\phi]]_{\langle M, f_V \rangle} \ni (w, t)$ であるとき、世界 w の state t で ϕ が成り立つという (紛れがなければ、 w, t を明示せず単に「 ϕ が成り立つ」ということもある)。

3. 記述例

本節では *C-TCMASAGes* によるエージェントの振る舞いの記述例として、我々のテストベッドであるカヌーレーシングを意識して、平面上での確率的な位置移動を考える。

3.1 行為の効果

座標平面上の位置 $(0, 0)$ にいるエージェントが、行為 e を実行 (これをイベントと捉える) すると、 $(1, 1)$ に動こうとするが、実際に到達する位置の分布は、平均 $(1, 1)$ 、共分散行列 $0.1I$ (ただし I は 2 次元単位行列) の 2 次元正規分布に従うとする (図 5)。この正規分布の確率密度関

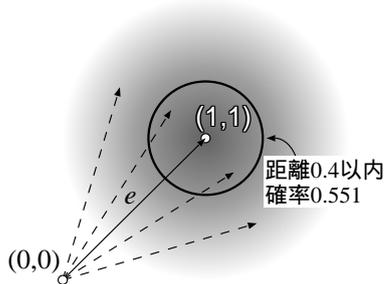


図 5 連続的な確率分布を持つ移動の例

数は $\frac{1}{2\pi \cdot 0.1} \exp\left(-\frac{1}{2 \cdot 0.1}((x-1)^2 + (y-1)^2)\right)$ なので、項 $\text{norm2d}(x, y)$ がこの式に解釈されるよう固定しておけば、

$$\text{loc}(0, 0) \supset X_{@norm2d(x, y)}^e \text{loc}(x, y) \quad (2)$$

でこの振る舞いを記述することができる (エージェントの位置を 2 引数述語記号 loc で表現している)。また、エージェントがこの式に従って振る舞う場合、例えば

$$\text{loc}(0, 0) \supset X_{=1-\exp(-0.8)}^e (\text{loc}(x, y) \supset (x-1)^2 + (y-1)^2 \leq 0.4^2) \quad (3)$$

が成り立つこと、すなわち次の時刻に一定の確率 $1 - e^{-0.8}$ (≈ 0.551) で点 $(1, 1)$ から距離 0.4 以内に着くことも (意味論的議論により) 示せる。

3.2 行為の決定

次に、行為の決定に関する例を示す。正規分布では数値計算が煩雑になるので、 (x_1, y_1) に動こうとすると図 6 のような確率分布で動く場合を考える。

$y_1 + 2$	5%	5%	5%	5%
$y_1 + 1$	5%	10%	10%	5%
y_1	5%	10%	10%	5%
$y_1 - 1$	5%	5%	5%	5%
$y_1 - 2$				
	$x_1 - 2$	$x_1 - 1$	x_1	$x_1 + 1$

(各マス目の中では一様分布)

図 6 平面上での移動先の簡略化された確率分布

いま、 x 軸正方向に流れている図 7 のような川*7 (斜線部は岸、黒塗り部は岩) をエージェント a がボートで下っており、現在位置は $(0, 0)$ とする。エージェントの行為は「左に漕ぐ」(e_1)、「漕がない」(e_2)、「右に漕ぐ」(e_3) の 3 つが選択可能で、それぞれ $(4, 2)$, $(4, 0)$, $(4, -2)$ 地点に行こうとするが、実際にはその地点を中心とする図 6 のような確率分布で移動する。例えば行為 e_3 を取ると、図 7 の網掛け部内に図 6 のような確率分布で到達する。

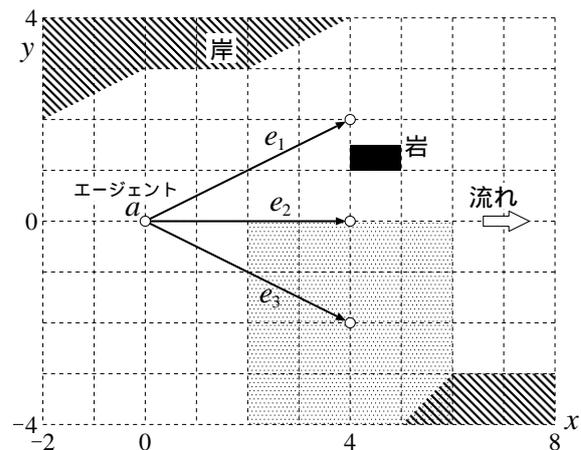


図 7 平面上を移動するエージェント

図 6 の分布の確率密度は

*7 わかりやすさのために格子は書いてあるが、グリッド状の離散な状態に区切られているわけではなく、エージェントの座標は実数値をとる。

$$P(x, y) = \begin{cases} 0.1 & \text{if } x_1 - 1 \leq x \leq x_1 + 1, \\ & y_1 - 1 \leq y \leq y_1 + 1 \\ 0.05 & \text{if 上記以外で } x_1 - 2 \leq x \leq x_1 + 2, \\ & y_1 - 2 \leq y \leq y_1 + 2 \\ 0 & \text{otherwise} \end{cases}$$

と書けるので、項 $P(x, y, x_1, y_1)$ の解釈を上式の値に固定しておけば、エージェント a の行為とその結果は

$$\text{loc}(0, 0) \supset X_{@P(x, y, 4, 2)}^{e_1} \text{loc}(x, y) \wedge X_{@P(x, y, 4, 0)}^{e_2} \text{loc}(x, y) \wedge X_{@P(x, y, 4, -2)}^{e_3} \text{loc}(x, y)$$

と書ける。本節では以後この式を ϕ と書き、 a はこれを信念として持っている ($BEL^a \phi$) もとする。

ここで、さらに a が川の(岸の)形や岩の位置・形を知って(信念として持って)いれば、 a は3つの各行為で岸または岩にぶつかる確率 (e_1 で 0.1、 e_2 で 0.025、 e_3 で 0.025) を知ることは可能である。すなわち、「岸が岩にぶつかる」を命題 *danger* で表すと) 以下が成り立つ。

$$BEL^a(\phi \wedge \text{loc}(0, 0)) \supset BEL^a(X_{=0.1}^{e_1} \text{danger} \wedge X_{=0.025}^{e_2} \text{danger} \wedge X_{=0.025}^{e_3} \text{danger})$$

a が、*danger* の確率を最も低くするよう振る舞うならば、このとき a は行為として e_2 と e_3 のどちらかをとることを願望し、いずれか1つを選んで(ここでは e_3 とする)それを意図として実行することになる。これは、

$$\begin{aligned} BEL^a(\phi \wedge \text{loc}(0, 0)) &\supset \\ &DESIRE^a(\text{pos}(e_2) \wedge \text{pos}(e_3)) \\ BEL^a(\phi \wedge \text{loc}(0, 0)) &\supset INTEND^a(\text{pos}(e_3)) \end{aligned}$$

の2式で書き表せる。実は図4はこの状況を示したものである。 w_1 は可能な行為に関する信念、その subworld である w_2 は *danger* の確率を最低にして次の地点に進もうという願望、さらにその subworld w_3 は具体的にどれかの行為を選んで生成した意図を表す。このように *Contextual Models* では、BDI モデルの枠内で、効果が連続的な確率分布をするような行為について議論できる。

3.3 プランの切り替え

本節では、連続的な確率分布を持つ遷移に関する知識を用いたプランの切り替えの例を述べる。話をさらに簡単にするため、行為の結果の移動先の分布は次元とする。

図8上を移動するエージェント a がある。 a は最初、図の y 軸上 ($x=0$) のどこかにいて、最終的には地点 G を目指しており、そのためのプランとしては当初「地点 A の十分近くを経由して G へ行く」というものを持っているとする(「十分近く」とはここでは、図中の A を通る太線の範囲内、すなわち $(5, -1)$ と $(5, 1)$ の間を指すものとする)。すると、そのため当初 a は「いつかは A の十分近くに達する」という意図を持ち、そしてその意図を持っている限り、 A に向けて x 軸方向に進む行為 (e_A

で表す)を取ろうとする(以後この記号を、 $does(e_A)$ の形で「その行為を取る」を表す論理式としても用いる)。このことは、「 A の十分近くに達している」を $near_A$ で表すと、2.2 節で導入した AF オペレータを用いて

$$INTEND^a AF near_A \supset does(e_A)$$

と書ける。

a が (x, y) にいるとき、行為 e_A を取ると実際には地点 $(x+1, y')$ へ動くものとする。ただし y' は、以下のように決まる y_1 を中心に、正規分布 $\mathcal{N}(y' | y_1, 1)$ で分布する: $y < -1$ なら $y_1 = y + 1$ 、 $y > 1$ なら $y_1 = y - 1$ 、いずれでもなければ $y_1 = 0$ 。すなわち a は、 y 座標を、 A の y 座標である 0 へ近づける方向へ最大 1 ほど動かそうとするが、実際の移動先は分散 1 の 1 次元正規分布に従う。また、 x 軸方向には必ず 1 だけ動く。例えば a が $(0, 7)$ にいる場合、 e_A を行うと図中のような移動をする。

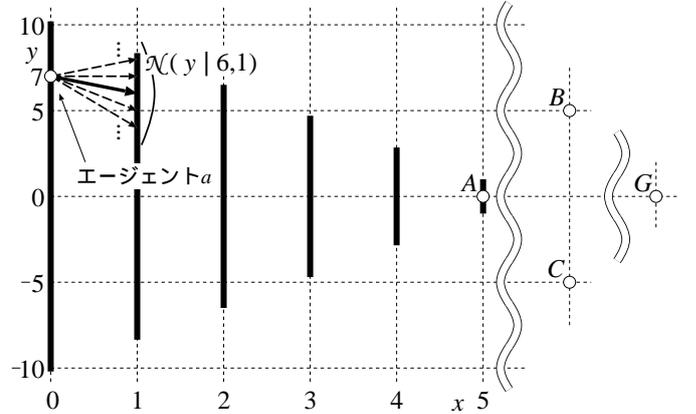


図8 プランを持って平面上を移動するエージェント

以上をエージェント a が知っている場合、そのことは、

$$\begin{aligned} y < -1 &\supset \text{move_target}(y, y + 1) \\ y > 1 &\supset \text{move_target}(y, y - 1) \\ -1 \leq y \wedge y \leq 1 &\supset \text{move_target}(y, 0) \end{aligned}$$

で定義される中間述語 move_target を使って

$$\begin{aligned} BEL^a(\text{loc}(x, y) \wedge \text{move_target}(y, y_1)) &\supset \\ BEL^a X_{@norm(y', y_1, 1)}^{e_A} \text{loc}(x+1, y') &\quad (4) \end{aligned}$$

と書ける(この式の記述には2.1節の norm を使った)。

しかし、 $x=0, \dots, 4$ の各段階では、それぞれ図の太線(おおむね $(0, -10.20) \sim (0, 10.20)$, $(1, -8.36) \sim (1, 8.36)$, $(2, -6.52) \sim (2, 6.52)$, $(3, -4.68) \sim (3, 4.68)$, $(4, -2.83) \sim (4, 2.83)$ の範囲内)にいないと、 e_A によって1つ右の太線の範囲に行ける確率が20%に満たない。特に、 $x=4$ の段階で太線内にいないと、 $near_A$ を達成できる確率が20%に満たない。このことは、式(4)から

$$\begin{aligned} BEL^a(\text{loc}(4, y) \wedge \neg(-2.83 \leq y \wedge y \leq 2.83)) &\supset \\ X_{\geq 8}^{e_A}(\text{loc}(5, y) \wedge \neg(-1 \leq y \wedge y \leq 1)) &\quad (5) \end{aligned}$$

(および他の x 座標の値と y 座標の範囲に関する同様の式) を (意味論的に) 導く形で示せる。

そこで、行動中に太線から外れたら、near_A の達成は難しいと見てこれをあきらめることを考えよう。これは、まず「太線を外れた」を表す runoff_thickline を

$$\begin{aligned} & \text{loc}(5, y) \wedge \neg \text{near_A} \supset \text{runoff_thickline} \\ & \text{loc}(x, y) \wedge x < 5 \wedge X_{\geq 8}^{eA} \text{runoff_thickline} \supset \\ & \text{runoff_thickline} \end{aligned}$$

のように定義し、これを用いて until オペレータにより

$$\begin{aligned} & \text{INTEND}^a \text{AF near_A} \supset \text{A}(\text{INTEND}^a \text{AF near_A} \cup \\ & (\text{BEL}^a \text{near_A} \vee \text{BEL}^a \text{runoff_thickline})) \end{aligned}$$

と書ける。これは「A に近づこうという意図は、実際に A に近づいたと信じるか、あるいは太線から外れた (ため達成困難だ) と信じるまで持続する」を意味し、一種の single-minded コミットメント戦略 [Rao 97] を実現していることになる。

a が意図 $\text{INTEND}^a \text{AF near_A}$ を失った後、まだ G に到達する目標を保持していれば、通常の BDI モデルの枠組の下、適宜他の地点 B や C を経由して G に行く新たなプランを立てて行動を続けることになる。

3.4 強化学習との結合

我々は [新出 11b] にて、 \mathcal{TCMAS} ($\mathcal{TCMASes}$ の前身) のもとで、数値演算の部分を nonlogical axiom として与えることによって、BDI と強化学習の結合をモデル化し、獲得方策による行為の決定や学習過程、プラン内での学習結果の利用などに関する推論を行う事例を与えた。これを実世界へ拡張するには、行為による遷移先が連続分布する場合を扱える必要がある。その一例を示す。

上記事例の中では、強化学習の課題において、ある行動 e によって状態 s_i ($1 \leq i \leq n$ 、ここで n は既知で固定とする) へ遷移する確率 p_i と、その遷移の際の即時報酬の推定値 r_i 、および遷移後の各状態の状態価値の推定値 v_i から、 e の行動価値 r を式 $\sum_{i=1}^n p_i (r_i + \gamma v_i)$ (ここで γ は割引率) によって求める過程を、下記のような axiom で表現した (現在位置・即時報酬・状態価値・行動価値を表す述語をそれぞれ at, reward, st_val, act_val とする)。

$$\begin{aligned} & X_{\geq p_1}^e \text{at}(s_1) \wedge \text{reward}(r_1) \wedge \text{st_val}(v_1) \mid \cdots \mid \\ & \quad \geq p_n \text{at}(s_n) \wedge \text{reward}(r_n) \wedge \text{st_val}(v_n) \wedge \\ & \quad \sum_{1 \leq i \leq n} p_i = 1 \wedge \sum_{1 \leq i \leq n} p_i (r_i + \gamma v_i) = r \\ & \supset \text{act_val}(e, r) \end{aligned}$$

\mathcal{TCMAS} では遷移先の状態の連続分布は扱えなかったが、 $\mathcal{TCMASes}$ では、状態 x へ遷移する確率密度が $p(x)$ だとすると、上記に相当することを

$$\begin{aligned} & X_{\geq p(x)}^e \text{at}(x) \wedge \int_x p(x) (\text{reward}(x) + \gamma \cdot \text{st_val}(x)) dx = r \\ & \supset \text{act_val}(e, r) \end{aligned}$$

と表現できる (なおここでは、reward, st_val は状態を引数に取って即時報酬・状態価値を返す関数を表す関数記号とした。また、 \int も通常の数学のそれを解釈に持つ関数記号として扱う)。

ただし、 $\mathcal{TCMASes}$ はあくまで形式的な表現の枠組みを提供するものであり、実際に強化学習を行うとして、状態が無限にある場合に、各状態での行動と行動価値の対を、計算機の有限な記憶の上でどう扱うかなどの問題への対処は、 $\mathcal{TCMASes}$ で表現しただけでは行えず、別途検討する必要がある。

4. 考察

本節では、 $\mathcal{TCMASes}$ の現状での問題点などについて論じる。

4.1 有用性や記述力などの検討

現時点では、連続な空間でのエージェントのあるべき振る舞い自体に関する検討がまだ十分行えてはいないため、 $\mathcal{TCMASes}$ の有用性も十分示せているとは言えない。 $\mathcal{TCMASes}$ の記述力に関する検討なども含め、今後の課題として残されている。

特に、現状の $\mathcal{TCMASes}$ では、連続な分布が扱えるのは状態空間に限られており、行為については連続な分布を扱えない (2.2 節でイベント記号の集合 \mathcal{E} を有限集合としているため、行為は有限通りとなる)。例えば 3.2 節の例では e_1, e_2, e_3 という離散な 3 つの行為を扱ったが、実際には「 r ほど左に漕ぐ」のような連続分布をとる行為を扱いたいという要求は自然なものであろう。ただ、「左に漕ぐ」という行為と漕ぐ量 r を分離し、例えば「 r ほど左に漕ぐと次の時刻の位置は平均 $f(r)$ 、分散 σ^2 の正規分布をする」を

$$\text{row_amount}(r) \supset X_{\text{norm}(x, f(r), \sigma^2)}^{\text{row_left}} \text{loc}(x)$$

と表現する (ここで row.amount は漕ぐ量、row.left は左に漕ぐという行為を表すとする) ことは可能であるので、これはそう重大な制約ではないかもしれない。

また、時間についても、BDI logic のベースとなる時相論理 CTL* が離散時間のため、現状では離散として扱っているが、これを連続にしようとする論理モデルとしても根本的な変化になる他、応用面でも、強化学習と BDI の結合に用いているマルコフモデルを連続時間に変更する必要があるなど影響が大きく、今のところ難しいと考えられる。

4.2 推論システム

他の課題として、演繹体系や推論エンジンのようなものを作るかというものがある。 $\mathcal{TCMASes}$ には我々がシーケント計算による演繹体系を提案したが [新出 11a]、 $\mathcal{TCMASes}$ については確率分布に任意の関数が書け

るため、演繹体系の構築は困難と考えられる。一方、3・4節で述べた $\mathcal{TCMAS}\mathcal{TCes}$ の事例では、Prolog を拡張した論理プログラミングシステムのような形で、BDI と強化学習の結合に関する推論を行う例を示しているが、 $C\text{-}\mathcal{TCMAS}\mathcal{TCes}$ のもとでも、ある程度不完全でもよければ、数値演算などに関する十分な nonlogical axiom を与えることによって、同様の仕組みを実現し、例えば 3・1 節の式 (2) から (3) を、また 3・3 節の式 (4) から (5) 相当の式を、意味論的議論によらずに導くようなことは可能かもしれない。また、そのようなものができるならば、論理型エージェント記述言語としての応用も期待する。

4.3 他の論理体系との比較

$C\text{-}\mathcal{TCMAS}\mathcal{TCes}$ は、状態遷移の連続な確率分布を扱う論理であるため、可能世界の集合に位相や濃度を導入している。同様に、可能世界の集合に位相などの幾何学的概念を持ち込んでいる研究としては、Awodey ら [Awodey 08] のように、可能世界がなす位相空間上の開核演算を必然様相オペレータと捉えるものがあるが、こちらは世界の構造と様相オペレータの性質の関連を主な研究対象としたもので、 $C\text{-}\mathcal{TCMAS}\mathcal{TCes}$ とは目的が異なり、確率分布の記述などは扱っていない。

また、連続的な確率分布を扱う論理の試みとしては Pfeifer らの SPL [Pfeifer 06a] があるが、これは真偽値に確率分布を導入したものであり、様相オペレータを持たないため、可能世界や状態遷移のような概念は扱えない。 $C\text{-}\mathcal{TCMAS}\mathcal{TCes}$ と同様に分岐時相論理に確率分布を入れたものとしては、Shakarian らの APT [Shakarian 12] や Kwiatkowska らの PTCTL [Kwiatkowska 02] など、主にモデル検査を念頭に置いた一連の研究があるが、これらはいずれも「 ϕ が成り立てば 5 時刻以内に 50 ~ 60% の確率で ψ が成り立つ」(APT での事例) のように、時間の経過 (パス) 上で特定の事象が起こる確率を記述するものであり、 $C\text{-}\mathcal{TCMAS}\mathcal{TCes}$ のように状態遷移先が確率分布するような事例を扱うものではない。一方これらは、時間のパス上での確率記述を扱えたり、また連続時間マルコフ過程を扱えるものがある [Aziz 96] など、現状の $C\text{-}\mathcal{TCMAS}\mathcal{TCes}$ にはない能力を持っており、それらとの比較も今後の課題の 1 つである。ただし、3・4 節で述べた強化学習との結合の事例などでは、連続的な状態分布に対する価値推定のようなことを行う必要があり、そうした応用には $C\text{-}\mathcal{TCMAS}\mathcal{TCes}$ の方が向いていると考えられる。

5. ま と め

本論文では、状態遷移の確率密度を記述できるように BDI logic を拡張した論理体系 $C\text{-}\mathcal{TCMAS}\mathcal{TCes}$ を提案し、またこれを用いて、連続的な確率分布を持つ状態空間を遷移するエージェントの振る舞いを、BDI の枠組みの下で形式化した例を示した。今後は、4 章で述べた諸問題

に関する検討を順次進めていくことが課題である。

◇ 参 考 文 献 ◇

- [Awodey 08] Awodey, S. and Kishida, K.: Topology and Modality: The Topological Interpretation of First-Order Modal Logic, *Review of Symbolic Logic*, Vol. 1, No. 02 (Special issue), pp. 146–166 (2008)
- [Aziz 96] Aziz, A., Sanwal, K., Singhal, V., and Brayton, R.: Verifying Continuous Time Markov Chains, in *Proc. of Conference on Computer-Aided Verification*, pp. 269–276 (1996)
- [Dam 90] 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)
- [Emerson 89] Emerson, E. A. and Srinivasan, J.: Branching Time Temporal Logic, in de Bakker, J. W., de Roeper, W. P., and Rozenberg, G. eds., *Linear Time, Branching Time and Partial Order in Logics and Models for Concurrency*, pp. 123–172, Springer-Verlag (1989)
- [Kozen 83] Kozen, D.: Results on the propositional μ -calculus, *Theoretical Computer Science*, Vol. 27, pp. 333–354 (1983)
- [Kwiatkowska 02] Kwiatkowska, M., Norman, G., Segala, R., and Sproston, J.: Automatic Verification of Real-time Systems with Discrete Probability Distributions, *Theoretical Computer Science*, Vol. 282, No. 1, pp. 101–150 (2002)
- [Manolios 00] Manolios, P.: Mu-Calculus Model-Checking, in *Computer-Aided Reasoning: ACL2 case studies*, pp. 93–111, Kluwer Academic Publishers (2000)
- [Pfeifer 06a] Pfeifer, N. and Kleiter, G. D.: Towards a probability logic based on statistical reasoning, in *Proceedings of the 11th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems*, pp. 2308–2315 (2006)
- [Pfeifer 06b] Pfeifer, R. and Bongard, J. C.: *How the Body Shapes the Way We Think: A New View of Intelligence*, The MIT Press (2006), (細田耕, 石黒章夫 (訳). 知能の原理 — 身体性に基づく構成論的アプローチ —, 共立出版, 2010)
- [Rao 97] Rao, A. S. and Georgeff, M. P.: Modeling Rational Agents within a BDI-Architecture, in Huhns, M. N. and Singh, M. P. eds., *Readings in Agents*, pp. 317–328, Morgan Kaufmann, San Francisco (1997)
- [Shakarian 12] Shakarian, P., Simari, G. I., and Subrahmanian, V. S.: Annotated Probabilistic Temporal Logic: Approximate Fixpoint Implementation, *ACM Transactions on Computational Logic*, Vol. 13, No. 2, pp. 13:1–13:33 (2012)
- [Tarski 55] Tarski, A.: A Lattice-Theoretical Fixpoint Theorem and its Application, *Pacific Journal of Mathematics*, Vol. 5, pp. 285–309 (1955)
- [Weisstein] Weisstein, E. W.: Delta Function, <http://mathworld.wolfram.com/DeltaFunction.html>: A Wolfram Web Resource
- [高田 10] 高田 司郎, 新出 尚之: 方策から生成したプランライブラリを用いる BDI エージェント, in *Proc. of JAWS2010* (2010), T-3.pdf
- [高田 11] 高田 司郎, 新出 尚之, 藤田 恵: 拡張 BDI 論理 $\mathcal{TCMAS}\mathcal{TCes}$ 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)
- [高田 12] 高田 司郎, 新出 尚之: 行為のアトラクター状態を考慮した知能ロボットについて, in *Proc. of JAWS2012* (2012), 69.pdf
- [新出 10] 新出 尚之, 高田 司郎, 片山 寛子: 方策から獲得したサブプランの汎用化について, in *Proc. of JAWS2010* (2010), T-4.pdf
- [新出 11a] 新出 尚之, 高田 司郎, 藤田 恵: 拡張 BDI 論理 $\mathcal{TCMAS}\mathcal{TCes}$ 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)
- [新出 11b] 新出 尚之, 高田 司郎, 藤田 恵: 拡張 BDI 論理 $\mathcal{TCMAS}\mathcal{TCes}$ を用いた確率的状態遷移のモデル化とその応用, *情報処理学会論文誌 数理モデル化と応用*, Vol. 4, No. 3, pp. 59–72 (2011)
- [藤田 12a] 藤田 恵: BDI モデルに基づく自律エージェントの行動決定アルゴリズムの拡張および応用, 博士論文, 奈良女子大学大学院人間文化研究科 (2012)
- [藤田 12b] 藤田 恵, 片山 寛子, 新出 尚之, 高田 司郎: 実世界の多様性に適応した BDI ロボットについて, *情報処理学会論文誌 数理モデル化と応用*, Vol. 5, No. 1, pp. 50–64 (2012)