明示的認識論理学と動的規範論理学

中　山　康　雄

目　次

明示的認識論理学と動的規範論理学

中　山　康　雄

１．はじめに

本稿では、近年私が展開してきた認識と規範に関する論理体系を新たな視点からまとめるとともに、次の二つのことを目的とする。

１　いくつかの有名なパズルの解法を提案する。

２　動的規範論理学（Dynamic Normative Logic, DNL）を用いて、典型的ゲームがどのように記述できるかを示す。

まず導入として、本稿のテーマに関するこれまでの私の研究過程を記述しておこう。本研究は、2010 年の紀要論文と国際ワークショップ SOCREAL 2010 の論文において、規範体系論理学（Logic for Normative Systems, LNS）を提案したときからはじまっている（中山 2010, Nakayama 2010）。LNS は、法的推論や道徳的推論をも含めた規範的推論を形式的に表現するために構想した体系である。その全貌は、『規範とゲーム』において描かれた（中山 2011）。そこでは、規範体系とゲーム体系という二つの枠組みを基盤にして社会活動を描くことが試みられた。しかしその後も、このアプローチは拡張されている。2012 年の科学基礎論学会での発表「規範体系論理学の特徴づけ」において、典型的なゲームが LNS の枠組みの中で記述できることが示され、SOCREAL 2013 で動的論理学の枠組みが提案された（Nakayama 2013）。そして、LNS や DNL を言語行為論に適用する方法が、中山（2012）や Nakayama (2013, 2014) で提案された。また、このアプローチの総括的な紹介が、Nakayama (*forthcoming*) でなされた。さらに、明示的認識論理学の時間論への適用が中山（2014）で提案されている。

２．明示的認識論理学

明示的認識論理学（Explicit Epistemic Logic, EEL）は、私がさまざまな形で展開してきた論理学体系の基礎となる枠組みであり、規範体系論理学および動的規範論理学から認識に関する部分だけを取り出したものとなっている。この節では、EEL の枠組みとその特性を紹介しておく。

定義１　Tを一階述語論理学（First-order Predicate Logic, FOL）の文集合とする。以下、Tを「信念基盤（belief base）」と呼ぶ。また、「$\Leftrightarrow$」はメタ言語の論理的同値を表すとする。そして、「$T \vdash p$」は「Tからpが演繹的に証明できる」ことを表わすとする。

(1a) $\mathbf{B}_T\, p \Leftrightarrow_{def}$［$T \vdash p$ & T は無矛盾］。

(1b) $\mathbf{M}_T\, p \Leftrightarrow_{def}$［$T \cup \{p\}$ は無矛盾］。

(1c) $Cn(T) =_{def} \{p : T \vdash p\}$。$Cn(T)$ を「信念集合（belief set）」と呼ぶ。

(1d) Xが認識主体のとき、Xが明示的に信じている文集合のこともXと書くことにする。そして、$\mathbf{B}_X\, p$や$\mathbf{M}_X\, p$などの表現を用いることにする。また、Xは個人に限定されていず、集団でもありうるとする。

(1e) $\mathbf{B}_@\, p$は、現実世界でpが真であることを意味するとする。このとき、@は極大無矛盾集合であり、任意の文pについて、「$p \in @$ または $\neg p \in @$」が成り立つ。

(1f)［共通信念基盤］Gは人物の有限集合とする。T_iをiが持っている信念基盤とするとき、$T(G) =_{def} \cap_{i \in G}\, Cn(T_i)$ と定義し、$T(G)$ を「グループ G の共通信念基盤」と呼ぶ。

$\mathbf{B}_T\, p$、$\mathbf{M}_T\, p$、$\mathbf{B}_@\, p$をそれぞれ次のように呼ぶことにする：「信念基盤Tのもとでpが信じられている」、「信念基盤Tのもとでpが可能だと信じられている」、「現実世界でpは真である」。$\mathbf{B}_T\, p$と$\mathbf{M}_T\, p$を日常言語で表現すれば、「（認識状態Tを前提にすれば、）pであるに違いない」、そして、「（認識状態Tを前提にすれば、）pかもしれない」ということになる。

定義１から命題１が成り立つことは、容易に証明できる。

命題１　⇒はメタ言語の「ならば」、& はメタ言語の「かつ」、*not* はメタ言語の否定、そして、*or* はメタ言語の「または」を表すとする

(2a)［演繹的閉包性］　$\mathbf{B}_T\, p \Rightarrow p \in Cn(T)$.

(2b)［Modus Ponens］　$\mathbf{B}_T\, (p \supset q) \Rightarrow (\mathbf{B}_T\, p \Rightarrow \mathbf{B}_T\, q)$.

(2c)［認識的様相性］　$\mathbf{B}_T\, p \Rightarrow \mathbf{M}_T\, p$.

(2d)［単調性］　($T1 \subseteq T2$ & $T2$ は無矛盾) $\Rightarrow (\mathbf{B}_{T1}\, p \Rightarrow \mathbf{B}_{T2}\, p)$.

(2e)［@ の極大無矛盾性］　$\mathbf{B}_@\, p \Leftrightarrow p \in @$.

(2f)［二値性の原理］　$\mathbf{B}_@\, p$ *or* $\mathbf{B}_@\, \neg p$.

(2g)［共通信念］　$\mathbf{B}_{T(G)}\, p \Leftrightarrow \forall i \in G\, \mathbf{B}_{Ti}\, p$. また、$\mathbf{B}_{T(G)}\, p$ と $\mathbf{B}_{Ti}\, p$ を、単純に $\mathbf{B}_G\, p$ および $\mathbf{B}_i\, p$ と書くこともある。

証明　(2a) は、(1a) と (1c) から直ちに帰結する。(2b) は (1a) から、(2c) は (1a) と (1b) から帰結する。 (2d) は (1a) と FOL の単調性から、(2e) と (2f) は (1e) から、(2g) は (1f) から帰結する。　Q.E.D.

本稿で議論する二つのパズルを解くためには、知覚による情報のとり入れを用いる必

要がある。そこで本稿では、次の知覚原理を認めることにする。

(3)［知覚原理］　知覚主体 X が適切な条件のもとで知覚し、それにともなって信念 p を形成したときには、次のことが成り立つ：$\mathbf{B}_{@}p \Rightarrow \mathbf{B}_{x}p.$

次に、明示的認識論理学の観点から二つのパズルを解く方法を提案しておく。

3．二少女の色あてパズル

「二少女の色あてパズル」は、「帽子の色あてパズル」の核の部分を取り出して単純化したパズルであり、八杉・小田（2001）が議論している。それは、次のようなパズルである。

二人の少女アンナとビアンカが一列に並んでいる。エンリコが二人に帽子をかぶせる。ビアンカの帽子の色は白であるが , アンナの色は何でもよい。アンナはビアンカの帽子を見ることができるが自分の帽子は見ることができない。ビアンカはどちらの帽子も見ることができない。エンリコが二人に告げる :「二人の帽子のうち少なくとも一方は白です」、と。

エンリコはまずアンナに「あなたの帽子の色は白ですか ?」と聞く。

(A) アンナは「わかりません」と答える。

(B1) ビアンカはこれを聞き , 自分の知識とする。

エンリコは次にビアンカに「あなたの帽子の色は白ですか ?」と聞く。

(B2) ビアンカは「はい」と答える。

このパズルの場合、信念帰属が重要な役割をはたしている。そしてパズルの解法もこの信念帰属の仕方と深く関わっている。ビアンカは、なぜアンナが「わかりません」と答えたかについて考え、そこからアンナの帽子が白であることをビアンカが見ていたという結論に至り、この推論に基づいて「はい」と答えているのである。このパズルを解くことがむずかしいのは、アンナの知覚を推測してビアンカが見ていないものについて答えるところにある。私はまず、信念帰属と他者からの知覚情報の受け入れに関する次の二原理を認めることから出発する。

(4a) $\mathbf{B}_{[\mathrm{B(PR)} > \mathrm{A}]}\,p \Leftrightarrow$ [B が認める原理 B(PR) と A が信じていると B が思っている文集合から $\mathrm{B_A}\,p$ を導くことができる]。このとき、$\mathbf{B}_{[\mathrm{B(PR)} > \mathrm{A}]}p$ を「原理 B(PR) に基づいて B が信念 p を A に帰属する」と読むことにする。ただし、このパズルに関してビアンカが認める原理 B(PR) は、メタ言語に関する古典論理の推論体系、命題

１、知覚原理 (3) だとする。

(4b) p が A の知覚内容を表す文のときには、次の規定が成り立つ：$\mathbf{B}_{[B(PR) > A]}\, p \Rightarrow \mathbf{B}_{B}\, p$.
つまりこのようなとき、B は A に帰属させた信念を真なものとして受け入れる。

このとき、ビアンカが自分の帽子が白だと推論できることを次のように証明できる。

〔1〕$\mathbf{B}_{[B(PR) > A]}$ (帽子 (B, 白) ∨ 帽子 (A, 白)).　［エンリコの発言を二人とも聴いているから］

〔2〕*not* $\mathbf{B}_{[B(PR) > A]}$ 帽子 (A, 白).　［ビアンカも知っているように、エンリコの質問に「わかりません」とアンナが答えたから］

〔3〕$\mathbf{B}_{@}$ 帽子 (B, Colour) ⇒ $\mathbf{B}_{[B(PR) > A]}$ 帽子 (B, Colour).　［(3)、(4a) より］

〔4〕$\mathbf{B}_{[B(PR) > A]}$ (帽子 (B, Colour) ∧ Colour ≠ 白) ⇒ $\mathbf{B}_{[B(PR) > A]}$ 帽子 (A, 白).［〔1〕より］

〔5〕*not* $\mathbf{B}_{[B(PR) > A]}$ ¬ 帽子 (B, 白).　［〔2〕、〔4〕より］

〔6〕*not* $\mathbf{B}_{@}$ ¬ 帽子 (B, 白).　［〔3〕、〔5〕より］

〔7〕$\mathbf{B}_{@}$ 帽子 (B, 白).　［〔6〕、(2f) より］

〔8〕$\mathbf{B}_{[B(PR) > A]}$ 帽子 (B, 白).　［〔3〕、〔7〕より］

〔9〕$\mathbf{B}_{B}$ 帽子 (B, 白).　［〔8〕、(4b) より］

このように、三つの基本的原理 (3)、(4a)、(4b) を認めれば、ビアンカが最後になぜ「はい」と答えたのかを、EEL の枠組みを用いて説明できる。

４. 泥だらけの子供たちのパズル

「泥だらけの子供たち（Muddy Children）のパズル」（Fagin et al 1995, van Benthem 2011: 12）は、情報更新に関するパズルである。言い換えると、このパズルを解く鍵になるのは、適切な情報更新の処理である。このパズルは、次のように描写できる。

> 三人の子供が外で遊んだ後、三人のうちの二人の子供の額(ひたい)に泥がつく。彼らは他の子供を見ることができるが、自分自身の状態については知ることができない。さて、彼らの父親がやってきて、次のように言う：「お前たちのうちの少なくとも一人がよごれている。」そして彼は問う：「自分がよごれているかどうか、誰か知っているか？」子供たちは、正直に答え、これが繰り返される。問いと回答が繰り返されると、何が起こるだろうか？

ここで、このパズルを記述するためにいくつかの規定を加えることにする。三人の子供を、*a*、*b*、*c* と名付けよう。問いと回答が繰り返されるそれぞれの状態を 0 から始ま

る自然数 n で表そう。また、この三人から成る集合を G と呼ぶ（$G = \{a, b, c\}$）。さらに、状態 n での三人の信念基盤をそれぞれ、$a(n)$、$b(n)$、$c(n)$ で表す。そして、$G(n) = Cn(a(n)) \cap Cn(b(n)) \cap Cn(c(n))$ と定義する。つまり $G(n)$ は、状態 n での三人の子供たちの共通信念基盤を表していることになる。

子供は三人おり、子供の額はきれいか汚いかのどちらかであるので、次のような基礎理論（Elementary Theory）$ET_{mc} = \{(5a), (5b)\}$ が成り立つ。

(5a) $\forall i, j, k\,(i \neq j \wedge j \neq k \wedge k \neq i \supset \forall m\,(m = i \vee m = j \vee m = k))$.
(5b) $\forall i\,(\text{額}\,(i, \text{きれい}) \vee \text{額}\,(i, \text{汚い})) \wedge \neg\exists i\,(\text{額}\,(i, \text{きれい}) \wedge \text{額}\,(i, \text{汚い}))$.

a、b、c は「自分たちの額は、きれいか汚いかのどちらかである」ことを知っている。また、三人は他の子供の額を見ることができる。そこで、次のことが成り立つ。

(5c) $ET_{mc} \subseteq a(0)$ & $ET_{mc} \subseteq b(0)$ & $ET_{mc} \subseteq c(0)$ & $ET_{mc} \subseteq G(0)$ & $ET_{mc} \subseteq @$.
(5d) $\forall I, j \in G\ \forall x\,(i \neq j \Rightarrow (\mathbf{B}_{@}\,\text{額}\,(j, x) \Rightarrow \mathbf{B}_{i(0)}\,\text{額}\,(j, x)))$.

このパズルの状況では、知識が保存されるため、(5e) が成り立つ。そしてそこから、(5f) 〜 (5h) が帰結する。ただしここでは、任意の $i(n)$ が無矛盾だと仮定している。

(5e) $\forall i \in G\ \forall n\,(i(n) \subseteq i(n+1))$.
(5f) $\forall i \in G\ \forall n\,(\mathbf{B}_{i(n)} p \Rightarrow \mathbf{B}_{i(n+1)}\, p)$.
(5g) $\forall n\,(\mathbf{B}_{G(n)} p \Rightarrow \mathbf{B}_{G(n+1)}\, p)$.
(5h) $\forall i \in G\ \forall n\,(\mathbf{B}_{G(n)} p \Rightarrow \mathbf{B}_{i(n)}\, p)$.

このパズルではまず、「お前たちのうちの少なくとも一人がよごれている」と父親が言うことにより、全員の信念が更新される：

(5i) すべての G の構成員 i について、$i(1) = i(0) \cup \{\exists i\,\text{額}\,(i, \text{汚い})\}$. よって、$G(0) \cup \{\exists i\,\text{額}\,(i, \text{汚い})\} \subseteq G(1)$.

このパズルでは、一人の額が汚い場合、二人の額が汚い場合、三人の額が汚い場合という三種類の可能性がある。ここでそれぞれの場合について、三人とも無矛盾な信念を持っているという前提のもとに、推論を記述してみよう。まず、一人だけの額が汚い場合の推論は、次のようになる。

〔1〕 $\mathbf{B}_{@} \exists^{=1} i\,\text{額}\,(i, \text{汚い})$.　［一人だけの額が汚いから］

〔2〕 $\forall i \in G\ (d = i \Leftrightarrow$ 額 $(i,$ 汚い $))$. ［記述を簡略化するために、名前 d を導入］

〔3〕 $\mathbf{B}_{@} \exists^{=2} i$ 額 $(i,$ きれい $)$. ［〔1〕、(5a)、(5b)、(5c) より］

〔4〕 $\forall i \in G\ (i \neq d \Rightarrow \mathbf{B}_{@}$ 額 $(i,$ きれい $))$. ［〔1〕、〔2〕、〔3〕、(5a) より］

〔5〕 $\forall i \in G\ (i \neq d \Rightarrow \mathbf{B}_{d(1)}$ 額 $(i,$ きれい $))$. ［〔4〕、(3)、(5d)、(5f) より］

〔6〕 $\mathbf{B}_{d(1)} \exists^{=2} i\ (i \neq d \wedge$ 額 (i, きれい $))$. ［〔3〕、〔4〕、〔5〕 より］

〔7〕 $\mathbf{B}_{d(1)} \exists i$ 額 $(i,$ 汚い $)$. ［(1a)、(5i) より］

〔8〕 $\mathbf{B}_{d(1)}$ 額 $(d,$ 汚い $)$. ［〔6〕、〔7〕、(1a)、(5a)、(5c)、(5e) より］

〔9〕 $\mathbf{B}_{@} \exists^{=1} i$ 額 $(i,$ 汚い $) \Rightarrow \exists i \in G\ \mathbf{B}_{i(1)}$ 額 $(i,$ 汚い $)$. ［〔1〕、〔8〕 より］

このように、一人だけの額が汚い場合には、その額の汚い一人の子供だけが「自分は知っている。自分の額は汚い」と答えるはずである。

二人の額が汚い場合には、「自分がよごれているかどうか、誰か知っているか？」という最初の問いかけに対しては誰も答えられず、誰も「知らない」と答える。

〔1〕 $\mathbf{B}_{@} \exists^{=2} i$ 額 $(i,$ 汚い $)$. ［二人の額が汚いから］

〔2〕 $\mathbf{B}_{@} \exists^{=1} i$ 額 $(i,$ きれい $)$. ［〔1〕、(5a)、(5b) 、(5i) より］

〔3〕 $\mathbf{B}_{G(2)} \exists^{\geq 2} i$ 額 $(i,$ 汚い $)$. ［最初の問いかけに誰も答えられなかったので、二少女の色合わせパズルと同様にして示すことができる］

〔4〕 額 $(d1,$ 汚い $) \wedge$ 額 $(d2,$ 汚い $) \wedge d1 \neq d2$. ［〔1〕 と 〔2〕 に基づき名前の導入］

〔5〕 $\mathbf{B}_{d1(2)} \exists i\ (i \neq d1 \wedge$ 額 (i, きれい $))$. ［〔2〕、〔4〕、(3) より］

〔6〕 $\mathbf{B}_{d1(2)} \exists^{\geq 2} i$ 額 $(i,$ 汚い $)$. ［〔3〕、(5h) より］

〔7〕 $\mathbf{B}_{d1(2)}$ 額 $(d1,$ 汚い $)$. ［〔5〕、〔6〕、(5a) より］

〔8〕 $\mathbf{B}_{@} \exists^{=2} i$ 額 $(i,$ 汚い $) \Rightarrow \exists i \in G\ \mathbf{B}_{i(2)}$ 額 $(i,$ 汚い $)$. ［〔1〕、〔7〕 より］

このように二人の額が汚い場合には、最初の問いかけに三人とも答えられなかった後に、二回目の問いかけには額が汚い子供は「自分は知っている。自分の額は汚い」と答えるはずである。そして、額が汚い子供は二人いるので、このように答える子供は二人いるはずである。

三人の額が汚い場合には、「自分がよごれているかどうか、誰か知っているか？」という二回の問いかけに対して誰も答えられず、誰もが「知らない」と二回答えることになる。

〔1〕 $\mathbf{B}_{@} \exists^{=3} i$ 額 $(i,$ 汚い $)$. ［三人の額が汚いから］

〔2〕 $\mathbf{B}_{G(2)} \exists^{=3} i$ 額 $(i,$ 汚い $)$. ［二回の問いかけに誰も答えられなかったので、二少女の色合わせパズルと同様にして示すことができる］

〔3〕 $\forall i \in G\ \mathbf{B}_{i(3)}$ 額 $(i,$ 汚い $)$. ［〔2〕、(2g) より］

このように三人の額が汚い場合には、二回の問いかけに三人とも答えられなかった後に、三回目の問いかけには三人の子供全員が「自分の額は汚い」と答えるはずである。

5．規範体系論理学と動的規範論理学

この節では、明示的認識論理学を基盤にして規範体系論理学と動的規範論理学を定義することにする。

(6a) T と OB は FOL の文の集合とする。また､T を｢信念基盤｣､OB を｢義務基盤｣と呼ぶ。さらに、$NS = \langle T, OB \rangle$ を規範体系（*normative system, NS*）と呼ぶ。

(6b) $\mathbf{B}_{NS}\, p \Leftrightarrow_{def} \mathbf{B}_T\, p$,

(6c) $\mathbf{M}_{NS}\, p \Leftrightarrow_{def} \mathbf{M}_T\, \mathrm{p}$.

(6d) $\mathbf{O}_{NS}\, p \Leftrightarrow_{def} [\mathbf{B}_{T \cup OB}\, p$ **&** $not\ \mathbf{B}_T\, \mathrm{p}]$.

(6e) $\mathbf{F}_{NS}\, p \Leftrightarrow_{def} \mathbf{O}_{NS} \neg \mathrm{p}$.

(6f) $\mathbf{P}_{NS}\, p \Leftrightarrow_{def} [\mathbf{M}_{T \cup OB}\, p$ **&** $not\ \mathbf{B}_T\, p]$.

(6g) 規範体系 NS が無矛盾なのは、$T \cup OB$ が無矛盾なとき、かつ、そのときに限る。

$\mathbf{O}_{NS}\, p$、$\mathbf{F}_{NS}\, p$、$\mathbf{P}_{NS}\, p$ をそれぞれ､「NS で p は義務である」､「NS で p は禁止されている」､「NS で p は許されている」と読むことにする。

ここで、規範体系論理学の諸定理のうち代表的なものを記しておく。

命題２　$NS = \langle T, OB \rangle$ は無矛盾な規範体系とする。このとき、次の諸定理が成り立つ。

(7a1) $\mathbf{O}_{NS}\, p \Rightarrow \mathbf{P}_{NS}\, p$.

(7a2) $\mathbf{F}_{NS}\, p \Rightarrow not\ \mathbf{P}_{NS}\, p$.

(7a3) $\mathbf{P}_{NS}\, p \Rightarrow not\ \mathbf{F}_{NS}\, p$.

(7a4) $\mathbf{O}_{NS}\, p \Rightarrow not\ \mathbf{B}_{NS}\, p$.

(7a5) $\mathbf{B}_{NS}\, p \Rightarrow (\mathbf{M}_{NS}\, p\ \&\ not\ \mathbf{O}_{NS}\, p\ \&\ not\ \mathbf{F}_{NS}\, p\ \&\ not\ \mathbf{P}_{NS}\, p)$.

(7b1) $(\mathbf{O}_{NS}\, (p \supset q)\ \&\ \mathbf{B}_{NS}\, p) \Rightarrow \mathbf{O}_{NS}\, q$.

(7b2) $(\mathbf{O}_{NS}\, (p \supset q)\ \&\ \mathbf{O}_{NS}\, p) \Rightarrow \mathbf{O}_{NS}\, q$.

(7b3) $(\mathbf{O}_{NS}\, (p \wedge q)\ \&\ not\ \mathbf{B}_{NS}\, p) \Rightarrow \mathbf{O}_{NS}\, p$.

(7b4) $(\mathbf{O}_{NS}\, (p \wedge q)\ \&\ \mathbf{B}_{NS}\, p) \Rightarrow \mathbf{O}_{NS}\, q$.

(7b5) $(\mathbf{O}_{NS}\, (p \vee q)\ \&\ \mathbf{B}_{NS} \neg p) \Rightarrow \mathbf{O}_{NS}\, q$.

(7b6) $(\mathbf{O}_{NS}\, (p \vee q)\ \&\ \mathbf{F}_{NS}\, p) \Rightarrow \mathbf{O}_{NS}\, q$.

(7b7) $(\mathbf{O}_{NS}\, p\ \&\ \mathbf{B}_{NS} \neg(p \wedge q)) \Rightarrow not\ \mathbf{P}_{NS}\, q$.

(7c1) $\mathbf{O}_{NS}\, (p \supset \neg q) \Leftrightarrow \mathbf{F}_{NS}\, (p \wedge q)$.

(7c2) $(\mathbf{O}_{NS}\, (p \supset \neg q)\ \&\ \mathbf{B}_{NS}\, p) \Rightarrow \mathbf{F}_{NS}\, q$.

(7c3) $(\mathbf{F}_{NS}\,(p \wedge q)\ \&\ \mathbf{B}_{NS}\,p) \Rightarrow \mathbf{F}_{NS}\,q$.

(7c4) $(\mathbf{F}_{NS}\,(p \wedge q)\ \&\ \mathbf{O}_{NS}\,p) \Rightarrow \mathbf{F}_{NS}\,q$.

(7d1) $(\mathbf{P}_{NS}\,(p \supset q)\ \&\ \mathbf{B}_{NS}\,p) \Rightarrow \mathbf{P}_{NS}\,q$.

(7d2) $(\mathbf{P}_{NS}\,(p \supset q)\ \&\ \mathbf{P}_{NS}\,p) \Rightarrow \mathbf{P}_{NS}\,q$.

(7d3) $(\mathbf{P}_{NS}\,(p \vee q)\ \&\ \mathbf{B}_{NS}\neg p) \Rightarrow \mathbf{P}_{NS}\,q$.

(7d4) $(\mathbf{P}_{NS}\,(p \vee q)\ \&\ \mathbf{F}_{NS}\,p) \Rightarrow \mathbf{P}_{NS}\,q$.

(7d5) $(\mathbf{P}_{NS}\,p\ \&\ \mathit{not}\ \mathbf{B}_{NS}\,(p \vee q)) \Rightarrow \mathbf{P}_{NS}\,(p \vee q)$.

(7e1) $(\mathbf{O}_{NS}\ \forall x_1 \ldots \forall x_{\mathrm{n}}\,(P(x_1, \ldots, x_{\mathrm{n}}) \supset Q(x_1, \ldots, x_{\mathrm{n}}))\ \&\ \mathbf{B}_{NS}\,P(a_1, \ldots, a_{\mathrm{n}})\ \&\ \mathit{not}\ \mathbf{B}_{NS}\,Q(a_1, \ldots, a_{\mathrm{n}}))$
$\Rightarrow \mathbf{O}_{NS}\,Q(a_1, \ldots, a_{\mathrm{n}})$.

(7e2) $\mathbf{O}_{NS}\ \forall x_1 \ldots \forall x_{\mathrm{n}}\,(P(x_1, \ldots, x_{\mathrm{n}}) \supset \neg Q(x_1, \ldots, x_{\mathrm{n}})) \Leftrightarrow \mathbf{F}_{NS}\ \exists x_1 \ldots \exists x_{\mathrm{n}}\,(P(x_1, \ldots, x_{\mathrm{n}}) \wedge Q(x_1, \ldots, x_{\mathrm{n}}))$.

(7e3) $(\mathbf{F}_{NS}\ \exists x_1 \ldots \exists x_{\mathrm{n}}\,(P(x_1, \ldots, x_{\mathrm{n}}) \wedge Q(x_1, \ldots, x_{\mathrm{n}}))\ \&\ \mathbf{B}_{NS}\,P(a_1, \ldots, a_{\mathrm{n}})\ \&\ \mathrm{not}\ \mathbf{B}_{NS}\neg Q(a_1, \ldots, a_{\mathrm{n}}))$
$\Rightarrow \mathbf{F}_{NS}\,Q(a_1, \ldots, a_{\mathrm{n}})$.

(7e4) $(\mathbf{P}_{NS}\ \forall x_1 \ldots \forall x_{\mathrm{n}}\,(P(x_1, \ldots, x_{\mathrm{n}}) \supset Q(x_1, \ldots, x_{\mathrm{n}}))\ \&\ \mathbf{B}_{NS}\,P(a_1, \ldots, a_{\mathrm{n}})\ \&\ \mathrm{not}\ \mathbf{B}_{NS}\,\neg Q(a_1, \ldots, a_{\mathrm{n}}))$
$\Rightarrow \mathbf{P}_{NS}\,Q(a_1, \ldots, a_{\mathrm{n}})$.

証明　(7a1) ～ (7a5) は定義 (6b) ～ (6f) から直ちに帰結する。(7b1) を証明するために、$\mathbf{O}_{NS}\,(p \supset q)$ と $\mathbf{B}_{NS}\,p$ を仮定する。すると (6b) と (6d) から、$\mathbf{B}_{T\cup OB}\,(\mathrm{p} \supset \mathrm{q})$ **&** *not* $\mathbf{B}_T\,(p \supset q)$ & $\mathbf{B}_T p$ が成り立つことがわかる。よって、$\mathbf{B}_{T\cup OB}\,q$ が成り立つ。ここで、$\mathbf{B}_T\,q$ と仮定すると、$\mathbf{B}_T\,(p \supset q)$ が帰結するが、これは *not* $\mathbf{B}_T\,(p \supset q)$ と矛盾する。よって、*not* $\mathbf{B}_T\,q$ が成り立つ。すると、(6d) から $\mathbf{O}_{NS}\,q$ が帰結する。よって、(7b1) が成り立つ。他の命題も同様の仕方で証明できる。　Q.E.D.

拙著『規範とゲーム』(2011) においては、(7e1) ～ (7e4) の表現において前件に *not* $\mathbf{B}_{NS}\,Q(a_1, \ldots, a_{\mathrm{n}})$ や *not* $\mathbf{B}_{NS}\,\neg Q(a_1, \ldots, a_{\mathrm{n}})$ の条件が欠けている。ここには、前著の誤りを訂正した定理が記されている。

動的規範論理学では、k を自然数とするとき、規範体系 *NS* の代わりに規範的状態（Normative State）$NS(k)$ を用いる。規範的状態は、信念状態 $T(k)$ と義務状態 $OB(k)$ から構成され、$NS(k) = \langle T(k), OB(k)\rangle$ が成り立つ。ちなみに $NS(k)$ は、k 時点での規範的状態を表わしている。動的規範論理学は、規範的状態の更新（update）を許すような論理体系である。信念状態も義務状態もともに更新できるが、本稿ではおもに信念状態の更新について考察する。

6．動的規範論理学のゲームへの適用

『規範とゲーム』では、一人ゲームは次のように、状態と行為空間と行為選択を規定することによって特徴づけられている（p. 100）。

(8a)［初期状態］初期状態は、ゲームの出発点における状態のことである。
(8b)［行為空間算出関数］行為空間算出関数は、与えられた状態でプレイヤーの行為空間（プレイヤーに許されている行為タイプの集合）を算出する関数である。
(8c)［行為選択］与えられた状態によって対応する行為空間からひとつの行為タイプを選択し、これを実行することにより、行為選択は実行される。
(8d)［状態算出関数］状態算出関数は、先行状態とプレイヤーの先行行為から規定される状態を算出する関数である。
(8e)［終了条件］ゲームが終了する条件を規定する。

この節で示したいのは、典型的なゲームの場合には、このゲーム体系の規定が動的規範論理学の枠組みの中で描けるということである。ここでは、三方陣形成の例を用いてこのことを示そう。三方陣というのは、魔方陣（magic square）のひとつであり、正方形の 3 × 3 のマスに数字を配置し、縦・横・斜めのいずれの列についても、その列の数字の合計が同じになるもののことである。

本稿では、三方陣形成の一人ゲームを記述するのに、次のようなタイプの命題や関数が用いることにする。ただし、k はゲームの進行段階を表わす自然数（これを「ゲーム内時点」と呼ぶ）とする。

行為命題：　書く (数 , 位置 , k).
状態命題：　占領 (k)、未占領 (k)、未使用 (k)、終了 (k)、勝利 (k).

ここで、規範的状態 $NS_{ms}(k) = \langle T_{ms}(k), OB_{ms}\rangle$ & $T_{ms}(k) = ET_{ms} \cup STATE(k)$ を考える。ただし、ET_{ms} は三方陣の基礎理論（elementary theory）を、OB_{ms} は三方陣の義務基盤を、$STATE(k)$ は k 時点での状態を表わすとする。

［三方陣の基礎理論］　ET_{ms} = {(ET1), (ET2), (ET3)}。なお、集合論の（有限集合に限定された）基本的体系は前提されているとする。

(I)［状態遷移に関する規定］　プレイヤーが一桁の数字 x をマス y に置くと、この行為の結果として、マス y がこの数字で占領され、状態が更新される。

(ET1) $\forall x \forall y \forall n$ (書く $(x, y, n) \supset$ (未占領 (n) = 未占領 $(n-1) - \{y\} \wedge$ 占領 (n) = 占領 $(n-1) \cup \{\langle x, y\rangle\} \wedge$ 未使用 (n) = 未使用 $(n-1) - \{x\}$)).

(II)［終了条件に関する規定］　ゲームは、九マスすべてが数字で占領されたときに終了する。このとき、縦列および横列および斜め列の数を足し合わせた合計がすべて同一になったときにプレイヤーの勝利となる。

(ET2) $\forall n$ (終了 $(n) \equiv$ 未占領 $(n) = \varnothing$).

(ET3) $\forall n$ (勝利 $(n) \equiv$ (終了 $(n) \wedge \forall x_a \forall x_b \forall x_c \forall x_d \forall x_e \forall x_f \forall x_g \forall x_h \forall x_i$ (占領 (n) =

$\langle x_a, \mathrm{a}\rangle, \langle x_b, \mathrm{b}\rangle, \langle x_c, \mathrm{c}\rangle, \langle x_d, \mathrm{d}\rangle, \langle x_e, \mathrm{e}\rangle, \langle x_f, \mathrm{f}\rangle, \langle x_g, \mathrm{g}\rangle, \langle x_h, \mathrm{h}\rangle, \langle x_i, \mathrm{i}\rangle\} \supset \exists s\, (x_a+x_b+x_c = s \wedge x_d+x_e+x_f = s \wedge x_g+x_h+x_i = s \wedge x_a+x_d+x_g = s \wedge x_b+x_e+x_h = s \wedge x_c+x_f+x_i = s \wedge x_a+x_e+x_i = s \wedge x_c+x_e+x_g = s))))$.

[三方陣の義務基盤]　OB_{ms} = {(OB1), (OB2), (OB3), (OB4)}。

(OB1) $\forall n\, (\neg$終了$(n-1) \supset \exists^{=1}x\, \exists^{=1}y\,($書く$(x, y, n) \wedge x \in$未使用$(n-1) \wedge y \in$未占領$(n-1))$.
（ゲームが未終了なら、プレイヤーは未使用の数字ひとつをまだ数字が書き込まれていないひとつのマスの中に書かなければならない。）

(OB2) $\forall n\, \neg \exists^{>1}x\, \exists y$ 書く (x, y, n).
（一度に複数の数をマスに書き込んではならない。）

(OB3) $\forall n\, \neg \exists x\, \exists y\,($ 書く $(x, y, n) \wedge \neg \mathrm{x} \in$未使用 $(n-1))$.
（一度使用された数字を書き込んではならない。）

(OB4) $\forall n\, \neg \exists x\, \exists y\,($ 書く $(x, y, n) \wedge \neg \mathrm{y} \in$未占領 $(n-1))$.
（数字がすでに書き込まれたマスに新たに数字を書き込んではならない。）

このとき、$\mathbf{B}_{NSms(n)}\, p \Rightarrow \mathbf{B}_{Tms(n)}\, p$ が証明できる。また、「ゲームが未終了なら、未占領領域に未使用の数を書き入れることが許されている（$\mathbf{P}_{NSms(n)}\, (\neg$ 終了 $(n) \supset \forall x\, \forall y\, (x \in$未使用$(n) \wedge y \in$未占領 $(n) \supset$書く $(x, y, n+1))))$）」ということが、有限モデルの提示によって証明できる。それでは、三方陣のゲームの進行例を動的規範倫理学の枠組みを用いて記述してみよう。

（表１）　初期状態

a	b	c
d	e	f
g	h	i

$STATE(0)$ = { 未使用 (0) = {1,2,3,4,5,6,7,8,9}, 未占領 (0) = {a, b, c, d, e, f, g, h, i}, 占領 (0) =$\varnothing$ }.
$T_{ms}(0) = ET_{ms} \cup STATE(0)$.

（表２）　行為選択「書く (**5**, e, 1)」後の状態

a	b	c
d	**5**	f
g	h	i

$STATE(1) = STATE(0) \cup \{$ 書く $(5, \mathrm{e}, 1)\}$ & $T_{ms}(1) = ET_{ms} \cup \mathrm{STATE}(1)$.
$\mathbf{B}_{Tms(1)}\,($未使用$(1) = \{1,2,3,4,6,7,8,9\} \wedge$ 未占領$(1) = \{\mathrm{a, b, c, d, f, g, h, i}\} \wedge$ 占領 $(1) = \{\langle 5, \mathrm{e}\rangle\} \wedge \neg$ 終了 $(1))$.

第２、３、４、５、６、７段階については、状態描写を省略し、どのような手がそれぞれ実行され、どのような状態更新がなされたのかだけを記述することにする。

$STATE(2) = STATE(1) \cup \{$ 書く $(8, \mathrm{a}, 2)\}$ & $T_{ms}(2) = ET_{ms} \cup STATE(2)$.

$STATE(3) = STATE(2) \cup \{$ 書く $(2, \mathrm{i}, 3)\}$ & $T_{ms}(3) = E_{ms} \cup STATE(3)$.
$STATE(4) = STATE(3) \cup \{$ 書く $(6, \mathrm{c}, 4)\}$ & $T_{ms}(4) = ET_{ms} \cup STATE(4)$.
$STATE(5) = STATE(4) \cup \{$ 書く $(4, \mathrm{g}, 5)\}$ & $T_{ms}(5) = ET_{ms} \cup STATE(5)$.
$STATE(6) = STATE(5) \cup \{$ 書く $(3, \mathrm{d}, 6)\}$ & $T_{ms}(6) = ET_{ms} \cup STATE(6)$.
$STATE(7) = STATE(6) \cup \{$ 書く $(1, \mathrm{f}, 7)\}$ & $T_{ms}(7) = ET_{ms} \cup STATE(7)$.

（表３） 行為選択「書く (7, f, 8)」後の状態

8	1	6
3	5	**7**
4	h	2

$STATE(8) = STATE(7) \cup \{$ 書く $(7, \mathrm{f}, 8)\}$ & $T_{ms}(8) = ET_{ms} \cup STATE(8)$.
$\mathbf{B}_{Tms(8)}$ (未使用 (8) = {9} ∧ 未占領 (8) = {h} ∧ 占領 (8) = {⟨5, e⟩, ⟨8, a⟩,⟨2, i⟩,⟨6, c⟩,⟨4, g⟩,⟨1, b⟩,⟨3, d⟩,⟨7, f⟩} ∧¬ 終了 (8)).

（表４） 行為選択「書く (9, h, 9)」後の状態

8	1	6
3	5	7
4	**9**	2

$STATE(9) = STATE(8) \cup \{$ 書く $(9, \mathrm{h}, 9)\}$ & $T_{ms}(9) = ET_{ms} \cup STATE(9)$.
$\mathbf{B}_{Tms(9)}$ (未使用 (9)=∅ ∧ 未占領 (9) =∅ ∧ 占領 (9) = {⟨5, e⟩,⟨8, a⟩, ⟨2, i⟩,⟨6, c⟩,⟨4, g⟩,⟨1, b⟩,⟨3, d⟩,⟨7, f⟩,⟨9, h⟩} ∧ 終了 (9) ∧ 勝利 (9)).

このように、「書く（数 , 位置 , *k*)」というタイプの行為の遂行によってゲームの状態は進行していき、最終的に終了状態に達することになる。またより複雑な典型的ゲームに関しても、同様な方法で動的規範論理学による記述が可能である [1)]。

7．カードのパズル

カードのパズルというものがある（van Benthem 2011: 8)。本稿では、このパズルを一種のゲームとして記述することにする。

> H、S、N の三人にそれぞれ、赤、白、青のカードが一枚ずつ分けられているとする。各プレイヤーは自分のカードを見ることはできるが、他の人のカードを見ることはできない。実際の分配は、H が赤を持ち、M が白を持ち、S は青を持つというものだったとする。ここで、会話が開始される。
> M は H に問う :「青のカードを持っていますか？」
> H は正直に「いいえ」と答える。

動的規範論理学を用いてこの会話の進行を描くために、規範体系 $NS_X(k)$ の構成要素となる集団 G のための基礎理論と義務基盤を規定する。

[I] $G = \{H, M, S\}$ とする。そして、変項 X は G か H か M か S であるとする。

$NS_X(k) = \langle X(k), OB_G\rangle$ & $X(k) = ET_G \cup STATE_X(k)$.

[II] 集団 G のための基礎理論 ET_G = {(9a), (9b)}.

(9a) $\forall i \exists^{=1} x$ (持つ $(i, x) \wedge (x =$ 赤 $\vee x =$ 白 $\vee x =$ 青 $)) \wedge \neg \exists i \exists j \exists x (i \neq j \wedge$ 持つ $(i, x) \wedge$ 持つ $(j, x))$.

(9b) $\forall n$ ($\forall i \forall x$ (答える $(i, x,$「はい」$, n) \supset$ 持つ $(i, x)) \wedge \forall i \forall x$ (答える $(i, x,$「いいえ」$, n) \supset \neg$ 持つ $(i, x)))$.

[III] 集団 G のための義務基盤：OB_G = {(9c)}

(9c) $\forall n \forall i \forall j \forall x$ (問う $(i, j, x, n) \supset (($ 持つ $(j, x) \supset$ 答える $(j, x,$「はい」$, n+1)) \wedge (\neg$ 持つ $(j, x) \supset$ 答える $(j, x,$「いいえ」$, n+1))))$.

[IV] 初期状態：

(9d) $STATE_G(0) = \emptyset$ & $STATE_H(0) = \{$ 持つ (H, 赤)$\}$ & $STATE_M(0) = \{$ 持つ (M, 白)$\}$ & $STATE_S(0) = \{$ 持つ (S, 青)$\}$.

M は自分が白のカードを持っていることを知っている。そこで M は、他の人のカードの色を知るためには、他の誰かに、「赤を持っているか」、あるいは、「青を持っているか」ということを聞けばよい。そこで M は、「青のカードを持っているか？」と H に聞いたのである。H は「いいえ」と答えたので、H は青ではなく赤のカードを持っていることになる。そこで残る S は、青のカードを持っていることがわかる。この推論を形式的に描くと次のようになる。

〔1〕 $\mathbf{P}_{M(0)}$ 問う (M, H, 青 , 1). ［有限モデルの構成より］

〔2〕 $\mathbf{B}_{@(1)}$ 問う (M, H, 青 , 1). ［M が H に「青か」と問う］

〔3〕 $STATE_G(1) = STATE_G(0) \cup \{$ 問う (M, H, 青 , 1)$\}$. ［〔2〕、(3) より］

〔4〕 $\mathbf{B}_{H(1)} \neg$ 持つ (H, 青). ［(5f)、(9a)、(9d) より］

〔5〕 $\mathbf{O}_{H(1)}$ 答える (H, 青 ,「いいえ」, 2). ［〔3〕〔4〕、(9c) より］

〔6〕 $\mathbf{B}_{@(2)}$ 答える (H, 青 ,「いいえ」, 2). ［H が M に「いいえ」と答える］

〔7〕 $STATE_G(2) = STATE_G(1) \cup \{$ 答える (H, 青 ,「いいえ」, 2)$\}$. ［〔6〕、(3) より］

〔8〕 $\mathbf{B}_{G(2)} \neg$ 持つ (H, 青). ［〔7〕、(9b) より］

〔9〕 $\mathbf{B}_{M(2)} \neg$ 持つ (H, 青). ［〔8〕、(2g) より］

〔10〕 $\mathbf{B}_{M(2)}$ 持つ (M, 白). ［(5f) 、(9d) より］

〔11〕 $\mathbf{B}_{M(2)}$ (持つ (H, 青) $\vee$ 持つ (H, 赤)). ［〔10〕、(9a) より］

〔12〕 $\mathbf{B}_{M(2)}$ 持つ (H, 赤). ［〔9〕、〔11〕 より］

〔13〕 $\mathbf{B}_{M(2)}$ 持つ (S, 青). ［〔10〕、〔12〕、(9a) より］

このように、カードのパズルはコミュニケーションを用いた一種のゲームとして解釈できることがわかる。そして、動的規範論理学は、このパズルの記述に適用できることが明らかになった。

8．まとめ

本稿では、明示的認識論理学、規範体系論理学、動的規範論理学の体系を紹介し、それを用いてどのようにパズルを解くのか、どのように典型的ゲームを表現するのかを示した。これらの論理体系は、命題論理学ではなく一階述語論理学を用いること、および、信念状態や規範状態を明示的に表現できることにその特徴がある。また、ヴィトゲンシュタインの原初的言語ゲームを、動的規範論理学を用いて記述できることが Nakayama (2013a, *forthcoming*) で示唆されている。

注

1）二人ゲームに関しては、誰が手番にあるかを示す状態述語が必要になる。そして、手番にない者にはゲームに関する行為が許されておらず、手番にある者は次の一手を自らの行為空間から選択しなければならない。

文献表

van Benthem, J. (2011) *Logical Dynamics of Information and Interaction*, Cambridge University Press.

Nakayama, Y. (2010), Logical Framework for Normative Systems, *SOCREAL 2010: Proceedings of the 2nd International Workshop On Philosophy and Ethics of Social Reality*, 27–28 March 2010, Hokkaido University, Sapporo, JAPAN, pp. 19-24. http://www.hucc.hokudai.ac.jp/~k15696/home/sr10/sr10-prep.pdf

Nakayama, Y. (2013a), Dynamic Normative Logic and Information Update," T. Yamada (ed.) *SOCREAL 2013: 3rd International Workshop on Philosophy and Ethics of Social Reality, Abstracts*, Hokkaido University, Sapporo, JAPAN, pp. 23-27. http://eprints.lib.hokudai.ac.jp/dspace/handle/2115/55055

Nakayama, Y. (2013b), Speech Acts, Normative Systems, and Local Information Update, *Proceedings of the Tenth International Workshop of Logic and Engineering of Natural Language Semantics 10 (LENLS 10)*, ISBN 978-4-915905-57-5 C3004 (JSAI), pp. 157-170.

Nakayama, Y. (2014), Speech Acts, Normative Systems, and Local Information Update, Y. I. Nakano, *et al*. (eds.) *New Frontiers in Artificial Intelligence (JSAI-isAI 2013 Workshops, Kanagawa, Japan, Selected Papers from LENLS10, JURISIN2013, MiMI2013, AAA2013,* *DDS13)*, Springer, pp. 98-114. (A revised version of (2013b))

Nakayama, Y. (*forthcoming*), Norms and Games as Integrating Components of Social Organizations, H. Ishiguro, *et al*. *Cognitive Neuroscientific Robotics: B Analytic Approaches to Human Understanding*, Springer, 20 頁程度

中山康雄 (2010),「規範体系の分析」『大阪大学大学院人間科学研究科紀要』第 36 巻 , 81-98 頁

中山康雄 (2011),『規範とゲーム　－　社会の哲学入門』勁草書房 .

中山康雄 (2012),「規範体系論理学を基盤にした言語行為の分析」日本認知科学会第 29 回大会発表論文集（CD-ROM 版）, 501-503 頁
http://www.jcss.gr.jp/meetings/JCSS2012/proceedings.html

中山康雄 (2014),「主体と時間と情報更新」『アルケー』関西哲学会年報 No. 22, 53-64 頁

八杉満利子・小田宗兵衛 (2001),「体系からの脱出 : 証明論による解析」『科学基礎論研究』Vol. 28 No.2, 33-38 頁

Explicit Epistemic Logic and Dynamic Normative Logic

Yasuo NAKAYAMA

In this paper, I report and summarize my recent researches on dynamic change of epistemic and normative states. I systematically redefine *Explicit Epistemic Logic* (EEL), *Logic for Normative Systems* (LNS), and *Dynamic Normative Logic* (DNL). There are several logical frameworks for change of epistemic and normative states. Approaches in this paper, have two characteristic features.

1. These frameworks explicitly express information of epistemic and normative states.
2. These frameworks deal with first-order logic.

The paper can be divided into four parts.

In the first part, EEL is defined and its fundamental propositions are introduced and explained. EEL is a framework that represents a belief state based on a belief base T. In EEL, "p is believed under the belief base T" is represented as $\mathbf{B}_T\, p$. In the second part, EEL is applied to two problems, namely to *puzzle of two girls* in Yasugi and Oda (2001) and *puzzle of Muddy Children* in van Benthem (2011). To solve these puzzles, I accept a perception principle and a principle for ascription of beliefs. For Muddy Children, additionally, the belief update has to be taken into consideration.

In the third part of this paper, LNS and DNL are defined. A normative system NS is defined as a pair composed of a belief base T and an obligation base OB ($NS = \langle T, OB \rangle$). Now, we can express not only that p is believed in NS but also that p is obligated (or forbidden or permitted) in NS. They are represented as $\mathbf{B}_{NS}\, p$, $\mathbf{O}_{NS}\, p$, $\mathbf{F}_{NS}\, p$, and $\mathbf{P}_{NS}\, p$. DNL is an extension of LNS equipped with devices for update of the belief base and the obligation base ($NS(k) = \langle T(k), OB(k) \rangle$).

Finally, in the fourth part, DNL is applied to a description of typical games, namely to a play for construction of a magic square and *puzzle of cards* in van Benthem (2011). Furthermore, DNL can be used to describe primitive language games.