計算論的関連性理論と命題論理*

松井 理直

Computational Relevance Theory and Propositional Logic

Michinao F. MATSUI

Abstract

Reasoning is fundamental to human intelligence, but the result of reasoning is different from the result of mathematical logic. In this paper the main point is that the essence of reasoning is the computation based on *Relevance*, and that human's logical act is possible by the simplest parameter settings of the computation of *Relevance*.

推論は人間の知性における最も基本的な特性である。しかし、人間はいわゆる数学的な論理と常に一致する推論を行うわけではない。本稿では、人間の推論は**関連性** に基づく計算であること、またその計算において最も単純なパラメータ設定を行うと、論理と同じ行動が可能になることを主張する。

1. はじめに

人間の思考は非常に精密な特性を持っているが、その推論・思考の形態は数学・論理学の観点から見て、常に正しいとは限らない。これには、正しい推論に必要となる正確な証拠を常に入手できるわけではないという理由と共に、我々の推論能力の特性が反映されているからと考えられる。本論文は、Sperber と Wilson によって提案されている**関連性理論** が主張する推論形式を人間の思考の特性と捉え、どのような状況で論理学的に正しい推論が行われるのか、どのような場合に数学的な結果と人間の思考が食い違うのかという点について議論を行ったものである。

*本研究は、日本学術振興会科学研究費補助金・基盤研究 (C)(1)「計算論関連性理論に基づく条件文理解過程の理論的・実証的研究」(平成 17 年度〜平成 20 年度、研究代表者: 松井 理直、課題番号 17500176) の援助を受けている。

Theoretical and Applied Linguistics at Kobe Shoin 9, 57–71, 2006.

2. 計算論的関連性理論

2.1 関連性の形式化

本節では、まず関連性理論の枠組みを簡単に述べ、それを数学的に形式化する手法について外観する。まず、この理論の中心をなす**関連性**の概念は (1) のように定義される。

(1) a. ある事実や刺激を持つ状況が認知主体に表象され、その表象を「真実あるいは真実であろう」として受理可能である時、その状況を **顕在的** (*manifest*) であるという。本稿では、顕在的事実のことを**想定** (*assumption*) と呼ぶ。

b. ある認知主体における想定の総体を **認知環境**(*cognitive environments*) と呼ぶ。

c. 認知環境の改善をもたらす作用を **認知効果** (*cognitive effects*) という。認知効果は (i) 新しい想定の獲得、(ii) 不確実な想定の確定、(iii) 誤った想定の棄却、によってもたらされる。

d. 不必要なコストを払うことなしに認知効果をもたらす情報のことを、**関連性** (*relevance*) を持つ情報という。

この定義を形式化する手法には、Marin (1999), Marin (2003) による意志決定理論に基づいたモデルなど、いくつかのアプローチが可能であるが、本稿では以下のアプローチを採用する。今、意味論 (言語の意味部門) において、顕示的情報の表象である想定 X が生じた場合、語用論の計算として、想定 X の確信度 (想定の強さ) $\mathcal{A}_x$ が設定されるとする。この想定確信度 $\mathcal{A}_x$ は以下のような性質を持ち、これが (1) の計算論的な見方となる。

(2) a. 心的情報は -1 から 1 までの実数によってその価値が示され、このうち 0〜1 までの値を持つ心的情報が想定 (顕在的事実) としての価値を持つ。この値を**想定確信度**と呼ぶ。すなわち、想定 X の確信度 $\mathcal{A}_x$ は 0〜1 までの実数値で表され、確信度が 1 に近いほど強く確信されている想定であり、0 に近づく程、信念の弱い想定である。なお、情報価値が 0 未満の心的情報は想定としての価値がない。

b. $0 \leqq \mathcal{A}_x \leqq 1$ ならば、想定 X として認知環境の中に組み込まれ、保持される。$\mathcal{A}_x < 0$ となる心的情報は想定と見なされず、認知環境に組み込まれない。

c. 認知効果は、(i) 0 以上の想定確信度を持つ新規想定を認知環境に組み混むこと、(ii) 既存の想定確信度をより 1 に近づけること、(iii) 既存の想定確信度をより 0 に近づけること、によってもたらされる。

d. 認知効果をもたらす関連性の高い想定とは例えば以下のようなものである。
(I) より簡単に想定確信度を 1 か 0 に近づけることができる想定。
(II) 推論の想定確信度が 1 に近い認知環境。

(2a) で定義された想定確信度は、命題の真理値と密接な関係を持ち、想定確信度が 1 に近いほど「真」と見なされており、0 に近いほど「偽」と見なされている命題と見なすこともできる。しかし、想定確信度は多値の真理値そのものではない。この点について

は、3.2 節で議論を行う。(2b) の定義では、o 想定確信度が $\mathcal{A}_x < 0$ となる情報は認知環境に組み込まれないという点が特に重要である。(2c) の定義は特に問題ないであろう。

(2d) のうち、(I) は (2c) の (ii), (iii) より自明である。また、(2c) の (i) も間接的に (2d-I) の効果をもたらす。(2d) の (II) は、人間の思考そのものの性質ともいえる。認知主体は、外界から得られた情報を蓄えるだけの存在ではない。得られた情報から積極的に新しい知識を生成し、未知の可能性を追求できる存在である。こうした自発的に作られる知識を生成する際、推論思考は最も重要な心的能力であり、できる限り正しいと思われる推論を行うことが必要不可欠である。そのため、推論の確信度がより高くなるような認知環境を積極的に選び取っていく。ここで、(2d) において、(II) の条件が (I) の条件と異なり、確信度が 0 の場合に言及されていない点に特に注意されたい。これは誤った推論は無意味であることに由来する。

2.2 想定確信度と認知環境の設定

前節で見たように、関連性の計算には、認知環境における想定確信度の変動が極めて重要になる。本節では、この認知環境の性質について考察する。まず、認知環境の議論に入る前に、2 つのタイプの想定を定義しておく。

(3) a. 単独想定：「X である」という単独の命題に対応する想定。ただし、「X でない」という否定命題に関しては、それが語彙化可能である場合 (Sperber, Cara, & Girotto, 1995) に限り、単独想定となる。

b. 複合想定：単独想定が複数組み合わさった想定。計算論的関連性理論では、連言で結びつけられた複合想定が最も基本的なものであると仮定する。選言を含む複合想定や含意想定、語彙化されていない否定想定などは、各種の連言想定から一定の計算を経て求められるものとする。

さて、一般に認知環境内における想定は単独想定のみであることは少なく、他の想定や文脈情報・背景情報などと相互作用を起こすことが多い。認知環境におけるこうした想定の相互作用や共起関係の可能性は、表 1 のように示すような連言で単独想定を繋いだ各種複合想定の確信度によって表現できる。

表 1: 認知環境における想定間の可能性

		情報 Y		
		Y	$\neg Y$	合計
情報 X	X	$\mathcal{A}_{xy}$	$\mathcal{A}_{x\bar{y}}$	$\mathcal{A}_x$
	$\neg X$	$\mathcal{A}_{\bar{x}y}$	$\mathcal{A}_{\bar{x}\bar{y}}$	$\mathcal{A}_{\bar{x}}$
	合計	$\mathcal{A}_y$	$\mathcal{A}_{\bar{y}}$	1

表 1 において、各想定確信度の値は相対的な関係を表している点に注意されたい。したがって、$\mathcal{A}_{xy} + \mathcal{A}_{x\bar{y}} + \mathcal{A}_{\bar{x}y} + \mathcal{A}_{\bar{x}\bar{y}} = 1$ が成り立つ。また、$\mathcal{A}_x = \mathcal{A}_{xy} + \mathcal{A}_{x\bar{y}}$, $\mathcal{A}_{\bar{x}} = \mathcal{A}_{\bar{x}y} + \mathcal{A}_{\bar{x}\bar{y}}$,

$\mathcal{A}_y = \mathcal{A}_{xy} + \mathcal{A}_{\bar{x}y}$, $\mathcal{A}_{\bar{y}} = \mathcal{A}_{x\bar{y}} + \mathcal{A}_{\bar{x}\bar{y}}$ も成立する。なお、絶対的な関係として言えることは、$\mathcal{A}_{xy} = 0$ ならば想定 $X \wedge Y$ は偽であり、$\mathcal{A}_{xy} \neq 0$ ならば想定 $X \wedge Y$ は真の可能性があるということだけである。

認知環境に組み込まれている限り、各想定値は 0〜1 までの範囲を取り ((2a) を参照)、また全想定値の合計が 1 になることから、表 1 における各想定確信度はその命題の主観的生起確率と見なすことができる。したがって、表 1 における否定情報は、具体的な下位情報に展開してもよい。この場合、認知環境は次のように設定される。なお、この時、$\mathcal{A}_{x_2y}+\mathcal{A}_{x_3y}+\cdots+\mathcal{A}_{x_ny}=\mathcal{A}_{\bar{x}y}$, $\mathcal{A}_{xy_2}+\mathcal{A}_{xy_3}+\cdots+\mathcal{A}_{xy_n}=\mathcal{A}_{x\bar{y}}$, $\mathcal{A}_{x_2y_2}+\mathcal{A}_{x_2y_3}+\cdots+\mathcal{A}_{x_ny_n}=\mathcal{A}_{\bar{x}\bar{y}}$ が成立する。

		情報 Y					
		Y	Y_2	Y_3	Y_4	$\cdots$	
情報X	X	$\mathcal{A}_{xy}$	$\mathcal{A}_{xy_2}$	$\mathcal{A}_{xy_3}$	$\mathcal{A}_{xy_4}$	$\cdots$	$=\mathcal{A}_{x\bar{y}}$
	X_2	$\mathcal{A}_{x_2y}$	$\mathcal{A}_{x_2y_2}$	$\mathcal{A}_{x_2y_3}$	$\mathcal{A}_{x_2y_4}$	$\cdots$	
	X_3	$\mathcal{A}_{x_3y}$	$\mathcal{A}_{x_3y_2}$	$\mathcal{A}_{x_3y_3}$	$\mathcal{A}_{x_3y_4}$	$\cdots$	
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	
		$\Vert$ $\mathcal{A}_{\bar{x}y}$					$\Vert$ $\mathcal{A}_{\bar{x}\bar{y}}$

2.3 推論の確信度に関する一般式

関連性理論では推論思考が中心的な役割を果たす。推論は直接経験を経ずに知識を増やす有効な手段であり、(2d) で見たように、認知効果を高める関連性情報を探り当てる上でも極めて重要な心的操作なのである。したがって、推論の確信度に関して、想定間の含意関係・依存関係を直接計算できるシステムがあったほうが望ましい。計算論的関連性理論では、表 1 における想定 $\mathcal{A}_x$ の想定 $\mathcal{A}_y$ に対する関連性の強度を、式 (4) により定義する。これは X → Y という推論の確信度に関する一般形といってもよい。

(4) $$\mathcal{A}_{x\to y} = h_x \cdot \frac{\mathcal{A}_{xy}}{\mathcal{A}_{xy} + \mathcal{A}_{x\bar{y}}} - h_{\bar{x}} \cdot \frac{\mathcal{A}_{\bar{x}y}}{\mathcal{A}_{\bar{x}y} + \mathcal{A}_{\bar{x}\bar{y}}}$$

h_x は 情報 X が成立する場合の顕示的情報をどの程度考慮にするか、$h_{\bar{x}}$ は 情報 X が成立しない場合の顕示的情報をどの程度考慮するかというバイアスを意味する係数であり、フレーム問題におけるフレームの大きさを決定するものとも見なせる。いずれの係数も、$0 \leqq h_x \leqq 1$, $0 \leqq h_{\bar{x}} \leqq 1$ を満たし、値が 1 の時は関連情報を完全に考慮することを、値が 0 の時は情報を参照しないことを意味する。

式 (4) は、確率論や統計学の観点からその意味を解釈することができる。前述したように、想定確信度は一種の主観的確率と見なせる。今、a という条件の元で b が起こる条件付き確率を $P(b|a)$ と表すとすると、式 (4) は

$$(5)\quad \mathcal{A}_{x\to y} = h_x \cdot \frac{\mathcal{A}_{xy}}{\mathcal{A}_{xy}+\mathcal{A}_{x\bar{y}}} - h_{\bar{x}} \cdot \frac{\mathcal{A}_{\bar{x}y}}{\mathcal{A}_{\bar{x}y}+\mathcal{A}_{\bar{x}\bar{y}}} = h_x \cdot P(\mathcal{A}_y|\mathcal{A}_{\bar{x}}) - h_{\bar{x}} \cdot P(\mathcal{A}_y|\mathcal{A}_{\bar{x}})$$

と表現できる。すなわち係数を無視するなら、前件が真になることを前提とした時に後件が真になる確率から、前件が偽である条件下で後件が成立する確率を引いたものということになり、後件の生起に関して前件の性質がどのような制限を掛けているのかを定量的に計算したものだと考えてよい。

同様に、式 (4) は統計学で用いられる回帰直線の考え方からも解釈することができる。回帰直線とは、データのペアに対し、なるべく誤差の少ない形で $y = \alpha + \beta x$ という関係式に当てはめた時の直線のことをいい、この式の β を回帰係数と呼ぶ。回帰係数は回帰直線の傾きを表す係数であり、x が y にどの程度影響を及ぼしているかを示す係数でもある。式 (4) を回帰直線の傾きと見なすと、この式における推論確信度の効果がよりよく分かる。すなわち、$\mathcal{A}_{x\to y} = 1$ の時は想定 X と Y の間に強い関係があり、したがって何らかの関連性が存在する可能性が高い (絶対に関連性があるとは言えない)。一方、$\mathcal{A}_{x\to y} = 0$ の時は想定 X と Y の間に何の関連性も存在しない。(2d) の (II) において、推論の想定確信度が 1 に近づく認知環境がよいとされているのは、こうした理由による。

なお、回帰係数は要因 X と Y の相互作用の強さを表す相関係数とも密接な関わりを持つ。$r = \sqrt{\mathcal{A}_{x\to y} \cdot \mathcal{A}_{y\to x}}$ により相関係数を計算可能 (符号は $\mathcal{A}_{x\to y}$ に一致) であり、またいわゆる決定係数は r^2 に等しい。r を論理演算子として見るなら、同値計算に相当する。ただし、「論理演算子として」という但し書きに注意されたい。認知主体が行う自然言語の解釈では、この式による計算は行われていないと思われる。例えば、言語における ‘if 文’ の解釈は、含意・同値・その中間的なものという様々な論理の形態に対応するが、これらは全て同一の式から計算されていると考えられる。

2.4 日常的推論の一般式

式 (4) における係数は、統計学上は X と Y により構成される空間に影響し、係数が 1 の時は規直交基底の空間となり、係数が小さくなるに従って、空間は小さくゆがんでいくものと解釈することができる。確率論的には、前述したように、前件の成立・不成立をどの程度考慮するかというフレームの大きさに対応する。ここで、自然言語における「もし X なら Y」という表現を考えた場合、前件を全く考慮しないという状況は語用論的にいって不自然である。しがたって、$h_x \neq 0$ が成立する。このため、式 (4) における係数は 2 種類設ける必要がなくなり、$h_x, h_{\bar{x}}$ の相対的な強さを表す係数 $k_{\bar{x}}$ により表現することができる。

$$(6)\quad \mathcal{A}_{x\to y} = \frac{\mathcal{A}_{xy}}{\mathcal{A}_{xy}+\mathcal{A}_{x\bar{y}}} - k_{\bar{x}} \cdot \frac{\mathcal{A}_{\bar{x}y}}{\mathcal{A}_{\bar{x}y}+\mathcal{A}_{\bar{x}\bar{y}}} = P(\mathcal{A}_y|\mathcal{A}_{\bar{x}}) - k_{\bar{x}} \cdot P(\mathcal{A}_y|\mathcal{A}_{\bar{x}})$$

式 (6) における係数 $k_{\bar{x}}$ は状況によって異なるが、例えば前件の否定状況が起こりやすいほど否定状況の顕在性も高いと仮定するならば、便宜上 $k_{\bar{x}} = \mathcal{A}_{\bar{x}}$, すなわち $k_{\bar{x}} = \mathcal{A}_{\bar{x}y} + \mathcal{A}_{\bar{x}\bar{y}}$ と置くことも可能であろう。この時、式 (6) は以下のように簡略化できる。

$$(7) \quad \mathcal{A}_{x \to y} = \frac{\mathcal{A}_{xy}}{\mathcal{A}_{xy} + \mathcal{A}_{x\bar{y}}} - \mathcal{A}_{\bar{x}y}$$

3. 関連性理論と論理との関係

3.1 様相性

前節で見た想定確信度は、様相論理のオペレータに近いものとして捉えることが可能である。例えば、必然性 (necessity)・可能性 (possibility) と想定確信度の間には、最も基本的な関係として (8) のような対応が存在すると考えられる。

(8) a. 想定 X の必然性：$\mathcal{A}_x = 1$ (近似解釈として $\mathcal{A}_x \approx 1$)
b. 想定 X の可能性：$\mathcal{A}_x \geq 0$ (近似解釈として $\mathcal{A}_x > 0$)

発話においては、(8) をさらに発展させた拡張解釈が行われる。その拡張の原因となるのは、「顕示的刺激である発話は、話者の側からすれば自らの持つ想定の中で最大の関連性を持つような情報の反映であり、聞き手にとっては、処理に見合う程度の関連性があることを保証するような情報である」という関連性の要請である。このことから、発話の認識的 (epistemic) な様相性はおおまかに次のように捉えることができる。

(9) a. X である： 話者にとっては認知環境に確信度が $\mathcal{A}_x$=1 であるような想定 X が存在することの反映であり、聴き手に取っては想定 X の確信度を 1 に設定せよという指令である。なお、$\mathcal{A}_{\bar{x}}$ については $\mathcal{A}_{\bar{x}} = 0$ であってもよいし (すなわち想定 $\neg X$ が認知環境に組み込まれていてもよいし)、想定 $\neg X$ が認知環境に存在していなくてもよい。

b. X に違いない： 話者の認知環境に確信度が $\mathcal{A}_x \approx 1$ であるような想定 X が存在した時の反映であり、聴き手に取っては想定 X の確信度を 1 にできる限り近づけよという指令である。想定 $\neg X$ については $\mathcal{A}_x$=1 の時であるなら、認知環境に組み込まれていても組み込まれていなくてもよいが、$\mathcal{A}_x \approx 1$ である場合には、確信度 $\mathcal{A}_{\bar{x}} = 1 - \mathcal{A}_x$ となるような形で認知環境に組み込まれる。

c. X であろう： 話者の認知環境における想定 X の確信度が、認知環境に組み込まれている他のいずれの想定の確信度よりも高い状態の反映であり、聴き手にとっては、想定 X の確信度を認知環境における他の想定確信度よりも大きくせよという指令である。

d. X でありうる： 話者の認知環境において、想定 X の確信度が $\mathcal{A}_x > 0$ であることの反映であり、聴き手にとっては想定 X を認知環境内に組み込めという指令である。「X かもしれない」も同様の条件であるが、上述

の表現との排反性から、想定確信度 $\mathcal{A}_x$ の値はそれほど大きくない状態に設定されていると考えられる。このことは、X 以外の想定が認知環境内に存在することを示唆する (想定確信度の和が 1 になることから)。

想定確信度という観点からは、認識的態度と義務的 (deontic) 態度には大きな違いは生じないものと思われる。すなわち、「X でなければならない」という表現は (9b) の状態とほぼ同一であり、「X ができる」という表現は (9d) の状態とほぼ同じであると考えてよい。

3.2 エントロピーと真理値

同様に、想定確信度はある種の数学的な操作を通して、一般的な命題論理における真理値に変換することができる。真理値の最も基本的な特徴は、ある命題について何らかの判断 (区別) が行えるという事にある。この判断を binary に行った場合、区別された一方を「真」、もう一方を「偽」と見なすことができる。言い換えるなら、ある命題とその否定命題との区別が可能であれば、少なくとも「真偽」の判断を行えるということだ。この真理値の性質は、「未知」という基準を取り込んだ三値論理の真理値を考えるとより明確になる。

(10) a. 二値論理

X	$\neg X$
真	偽
偽	真

b. 三値論理

X	$\neg X$
真	偽
未知	未知
偽	真

この「ある命題とそれ以外の命題の区別が可能か否か」という性質は、情報論でいうエントロピーの概念で表現できる。そこで、エントロピーの計算を通して、想定確信度と命題の真理値を結びつけてみよう。今、想定 X の確信度を $\mathcal{A}_x$ とした時、X 以外の全て想定について、その確信度の和を取ると $1-\mathcal{A}_x$ となる。したがって、想定 X のエントロピーは式 (11) で与えられる。なお、X は単一命題に対応する想定であっても、複合命題に対応する想定であってもよい。

(11) **想定 X のエントロピー**: $E_x = -\mathcal{A}_x \cdot \log_2 \mathcal{A}_x - (1-\mathcal{A}_x) \cdot \log_2(1-\mathcal{A}_x)$

エントロピーは常に 0〜1 までの値を取る。例えば、$\mathcal{A}_x$ が 1 (想定 X に絶対な確信を抱いており、$\neg X$ の可能性はない) か 0 である ($\neg X$ しか想定されていない) 時、エントロピーは最小の E_x=0 となり、$\mathcal{A}_x$ が 0.5 (想定 X と $\neg X$ を同程度に考慮) の時にはエントロピー最大 (E_x=1) となる。このエントロピー E_x から、多値の判断価値 V_x を定義する。

(12) **想定 X の判断価値**: $V_x = j \cdot (1 - E_x)$

ただし j は ± の符号であり、$\mathcal{A}_x \geqq 0.5$ の時 $j = 1$, $\mathcal{A}_x < 0.5$ の時 $j = -1$

この判断価値 V_x は −1〜1 までの値を取り、$\mathcal{A}_x = 1$ の時 $V_x = 1$, $\mathcal{A}_x = 0$ の時 $V_x = -1$, $\mathcal{A}_x = 0.5$ の時 $V_x = 0$ となる。この連続量 V_x をカテゴリカルに区切れば、命題論理の真

理値と見なすことができる。例えば一般的な二値論理であるなら、$V_x \neq -1$ なら「真」、$V_x = -1$ なら「偽」と見なせばよい。あるいは、$V_x = 1$ の時に「真」、$V_x \neq 1$ の時に「偽」とするような二値論理を構成することも理論上は可能である。こうした二値論理では、包含的選言や含意の代わりに、排他的選言や同値が基本論理となる。三値論理であるなら、$V_x = 1$ なら「真」、$-1 < V_x < 1$ なら「未知 (X か $\neg X$ かの絶対判断ができない)」、$V_x = -1$ の時に「偽」とするような論理を構成する方法が考えられる。もちろん二値論理と同様、これとは別の方法でカテゴリカルな割り当てを行った異なるタイプの三値論理を構成することもできる。例えば、$V_x = 0$ の時に「未知」、V_x が正の実数であれば「真」、負であれば「偽」となるような三値論理も理論上は可能である。

なお、(12) はあくまでエントロピーと真理値の関係を示す 1 つに手段に過ぎず、他の計算法を設定することもできる。例えば、$(1 + j \cdot (1 - E_x))/2$ のような式を定義すれば、値が 0 の時に「偽」、0 以外であるなら「真」と見なせるような論理を構成でき、二値論理におけるブール代数の計算結果と近い答を得ることができる。どのような数学的な変換を考えるにせよ、最も大切なポイントは、いわゆる真理値といわれるものが、連続的な値である想定確信度から計算ができること、その計算過程は情報のエントロピーの計算とほぼ等価なものであるという点にある。換言するならば、関連性理論の立場から論理を見た場合、真理値とは情報を区別するために最低限必要な 2 つの範疇を最も単純な形で実現したものと見なすことができる。そしてその範疇境界をどのように設定するかによって、例えば含意と同値の違い、あるいは exclusive disjunction と inclusive disjunction の違いが生起するのである。

3.3 推論思考と論理的含意

人間の行う推論が、時に論理学の含意計算とは異なる性質を持つことはよく知られた事実である。これにはいくつかの理由があるが、まず根本的な問題として、二値論理における「真」という判断の価値を考えなければならない。二値論理においては、一般的に「偽」は絶対的なものであるが、「真」は可能性を表すものとして解釈される。しかし、論理学における含意 (implication) と同値 (equivalence) を比較した場合、含意における「真」は (前件が偽である時には) 可能性を意味するものであるが、同値解釈における「真」は絶対的な完全解釈がなされる。(13) において□で囲まれている部分が、完全解釈ではなく、可能性解釈を受けている真理値である。

(13)

x	y	x→y	x↔y
真	真	真	真
真	偽	偽	偽
偽	真	$\boxed{真}$	偽
偽	偽	$\boxed{真}$	真

論理学においては、可能性解釈と完全解釈の違いは問題を起こさないが、心的態度としての命題 (想定) を考えた場合には、3.2 節で見たような必然性／可能性の違いが生じる。この (8) の違いを、式 (6) に反映させてみよう。

まず、(8a) に相当する $\mathcal{A}_{x\to y}=1$ の場合を見る。これは、$X\to Y$ の必然性、すなわち完全解釈を要求することになるので、命題論理の観点からいうと同値解釈に相当する。今、式 (6) において、$0 \leqq \dfrac{\mathcal{A}_{xy}}{\mathcal{A}_{xy}+\mathcal{A}_{x\bar{y}}} \leqq 1,\ 0 \leqq \dfrac{\mathcal{A}_{\bar{x}y}}{\mathcal{A}_{\bar{x}y}+\mathcal{A}_{\bar{x}\bar{y}}} \leqq 1$ が成立するため、任意の $k_{\bar{x}}$ に対し、$\mathcal{A}_{x\to y}=0$ が成立するためには、$\dfrac{\mathcal{A}_{xy}}{\mathcal{A}_{xy}+\mathcal{A}_{x\bar{y}}}=1,\ \dfrac{\mathcal{A}_{xy}}{\mathcal{A}_{\bar{x}y}+\mathcal{A}_{\bar{x}\bar{y}}}=0$ でなければならない。したがって、$\mathcal{A}_{x\bar{y}}=0, \mathcal{A}_{\bar{x}y}=0$ となる必要があり、$X\wedge\neg Y, \neg X\wedge Y$ が共に偽 (すなわち (13) における 2 段目と 3 段目が共に偽) の解釈を受けなければならないことが分かる。すなわち、$\mathcal{A}_{x\to y}=1$ を満たす演算は、同値解釈と等しい。

次に、(8b) に相当する $\mathcal{A}_{x\to y} \geqq 0$ を見てみよう。式 (6) において、任意の $k_{\bar{x}}$ に対し $\mathcal{A}_{x\to y} \geqq 0$ となるには、$\dfrac{\mathcal{A}_{xy}}{\mathcal{A}_{xy}+\mathcal{A}_{x\bar{y}}}=1$ であればよい。したがって、$\mathcal{A}_{x\bar{y}}=0$ のみが成立すればよく、(13) の 2 段目に相当する $X\wedge\neg Y$ のみが偽となることが分かる。すなわち、$\mathcal{A}_{x\to y} \geqq 0$ という可能性解釈は、命題論理における含意解釈と等価である。

含意解釈に関しては、次のように考えることもできる。含意において、絶対的な真理値の判断が成されているのは、前件が真である場合のみである。言い換えるなら、含意解釈というのは、前件が真であるようなフレームを設定し、このフレーム内においてのみ完全解釈を行うということでもある。こうしたフレーム設定は、式 (6) における係数 $k_{\bar{x}}$ を $k_{\bar{x}}=0$ と設定することに等しい。この時に、完全解釈、すなわち $\mathcal{A}_{x\to y}=1$ となる条件を求めると、やはり $\dfrac{\mathcal{A}_{xy}}{\mathcal{A}_{xy}+\mathcal{A}_{x\bar{y}}}=1$ となり、$\mathcal{A}_{x\bar{y}}=0$ のみが成立すればよいことが分かる。

このように、関連性の想定確信度の計算におけるパラメータ設定によって、論理と同等の結果を導出することができる。換言するなら、論理とは関連性計算におけるパラメータ設定を究極にまで簡略化したもの、関連性の計算式を最もシンプルにしたものなのだ。つまり、ある人間が論理と同等の行動ができるのは、「論理そのもの」を知っているというより、最も単純で適切な関連性の計算を行った結果、それが数学的な論理と一致するのだということである。また、この立場では、含意と同値の違いは可能性解釈か絶対的解釈かという、モダリティの近い観点から再分析することが可能となる。

4. 様相性を巡るパラドックス

4.1 グライスのパラドックス

前節の結論は論理における対偶表現のパラドックスからも見て取れる。一般に、ある含意表現とその対偶となる含意表現は、必ず論理学的に同一の真理値となる。しかし、その含意表現にモダリティ (あるいは事態の量化) が含まれている場合、「グライスの逆理 (Grice's paradox)」が生じることがある。Grice's paradox とは、ある含意表現における元々の量化の値と、その対偶表現の量化の値との間にズレが生じ、ある含意表現とその対偶が論理学的に同一でなくなることを言う。Kratzer (1986) が挙げている Grice's paradox

の例は以下のようなものである。

> ヨグとゾグがチェスの対戦を行う。ただし、10 回に 9 回はヨグが白である。これまでに 100 回の対戦を行ったところ、ヨグが白の時は、90 回中 80 回ヨグが勝った。一方ヨグが黒の時は、10 回中 10 回ともヨグは負けている。この状況で、次のような発話は適切だろうと思われる。
>
> (14) a. もし、ヨグが白ならば、彼が勝った可能性は 8/9 である。
> b. もし、ヨグが負けたなら、彼が黒であった可能性は 1/2 である。
>
> 一般量化子の理論からも分かる通り、この発話は、次のような論理式で表すことができる。なお、最初の XX-probably は [ ] 内の命題の起こりうる確率を示している。
>
> (15) a. 8/9-probably [if Yog had white, Yog won]
> b. 1/2-probably [if Yog lost, Yog had black]
>
> ここで Grice's paradox が生じる。すなわち、(15) における [ ] 内の命題は、対偶の関係になっている。それにもかかわらず、その確率は異なってしまっている。すなわち、論理的に等価なものとはいえない。このパラドックスをどのように解消すべきか。

確かに、Grice's paradox は論理的には興味深い問題である。しかし、語用論の立場からの回答は、とても単純なものである。これは paradox ではない、すなわち語用論的に、「A ならば B」とその対偶表現「B でないなら A でない」は必ずしも等価でなくてよい、というのが Grice's paradox に対する回答となる。計算論的関連性理論の枠組みを用いると、この回答を形式化して説明することができる。今、式 (6) から、$\mathcal{A}_{x\to y}$ と $\mathcal{A}_{\bar{y}\to\bar{x}}$ の想定確信度は

(16) a. $$\mathcal{A}_{x\to y} = \frac{\mathcal{A}_{xy}}{\mathcal{A}_{xy} + \mathcal{A}_{x\bar{y}}} - k_{\bar{x}} \cdot \frac{\mathcal{A}_{\bar{x}y}}{\mathcal{A}_{\bar{x}y} + \mathcal{A}_{\bar{x}\bar{y}}}$$

b. $$\mathcal{A}_{\bar{y}\to\bar{x}} = \frac{\mathcal{A}_{\bar{x}\bar{y}}}{\mathcal{A}_{\bar{x}\bar{y}} + \mathcal{A}_{\bar{x}y}} - k_{y} \cdot \frac{\mathcal{A}_{\bar{x}y}}{\mathcal{A}_{\bar{x}y} + \mathcal{A}_{xy}}$$

として計算でき、両者の確信度は一致しない。しかし、$\mathcal{A}_{xy} + \mathcal{A}_{x\bar{y}} + \mathcal{A}_{\bar{x}y} + \mathcal{A}_{\bar{x}\bar{y}} = 1$ が成り立つことから、他の推論の想定確信度よりは両者が近い値を持つことも分かる。特徴的なのは、式 (16) における係数を消去した簡易バージョンの式 (7) を用いた場合で、この時には式 (16) は

(17) a. $$\mathcal{A}_{x\to y} = \frac{\mathcal{A}_{xy}}{\mathcal{A}_{xy} + \mathcal{A}_{x\bar{y}}} - \mathcal{A}_{\bar{x}y}$$

b. $$\mathcal{A}_{\bar{y}\to\bar{x}} = \frac{\mathcal{A}_{\bar{x}\bar{y}}}{\mathcal{A}_{\bar{x}\bar{y}} + \mathcal{A}_{\bar{x}y}} - \mathcal{A}_{\bar{x}y}$$

となり、$\mathcal{A}_{x\bar{y}} = 0$ (すなわち X が真かつ Y が偽の時は偽) が成立する場合に限って、ある含意表現とその対偶表現の値は一致することが分かる (その想定確信度は 1 ではなく、$1 - \mathcal{A}_{\bar{x}y}$ となる)。

さらに、計算論的関連性理論では、Grice's paradox の問題を「ベイズの錯誤」と同等のものとして捉えることができる。Kratzer の挙げている例に即して考えてみよう。まずこの例から構成される認知環境の想定確信度は以下のようになる。

(18)

	勝	負	計
白	0.8	0.1	0.9
黒	0	0.1	0.1

ここから「もしヨグ白なら彼は勝つ」という推論の想定確信度を求めてみると、$\frac{0.8}{0.9} - k_{\bar{x}} \cdot \frac{0}{0.1}$, すなわち約 88.9% の確信度となる。なお、この例では、ヨグが黒の時に勝つ確率が 0 であるため、「たまたま」(15a) における確率と一致する。

次に、この認知環境を持っている人が「もしヨグが負けたなら、彼は黒であった」という命題の確信度をどのように計算するかを見てみよう。以下の議論における最大のポイントは、確率のとらえ方にある。Gillies (2000) で詳しく論じられている通り、確率には**客観的確率** (objective probability) と**認識論的確率** (epistemic/subjective probability) を個別に定義できる。例えば、コインを 2 枚投げた場合、一枚が表で他の一枚が裏になる確率は、客観的確率に基づくと 1/2 になる。客観的確率では 2 枚のコインが token として区別されるため、{ コイン 1, コイン 2}={ 表, 表 }, { 表, 裏 }, { 裏, 表 }, { 裏, 裏 } という事象が考えられるからである。一方、認識論的確率では、2 枚のコインは type として同一であるため、2 枚のコインは { 表, 表 }, { 表, 裏 }, { 裏, 裏 } のいずれかの状態にあると見なされ、その結果、一枚が表で他の一枚が裏になる認識論的確率は 1/3 になる。この客観的確率と認識論的確率の違いは、物理学でも量子力学などで重要な役割を果たす。

計算論的関連性理論では、顕示的な情報に関しては、個々の token が明確に意識されるため、客観的確率が計算可能であるが、顕示的でない情報については認識論的確率の計算が行われると考える。ここで、いわゆる条件付き確率における「前提確率の無視」というベイズの**問題**が発生する。具体的に見てみよう。

今、この人が「もしヨグが負けたなら」という思考を開始した場合、顕在化される情報は「勝ち負け」に関するものであり、「白黒」の情報ではない。したがって、(18) の認知環境は次のように再構成される。

(19)

	勝	負	計
白	???	???	0.5
黒	???	???	0.5

ここで問題となるのは ??? の数値であるが、「勝ち負け」に関する情報は顕在的であるから、勝負の割合は (18) における比率が維持される。その結果、認知環境全体の想定値は、

(20)

	勝	負	計
白	0.4/0.9	0.05/0.9	0.5
黒	0	0.5	0.5

この状況で、「もしヨグが負けたなら、彼は黒であった」という命題の確信度を計算すると、求めてみると、$\dfrac{0.5}{0.5+(0.05/0.9)} - k_{\bar{y}} \cdot \dfrac{0}{0+(0.4/0.9)}$ となり、想定確信度はちょうど 90% となる。これは、「もしヨグが勝つなら彼は白である」の想定確信度 88.9% と完全には一致しないが、極めて近い値になる。少なくとも、(15) における確率値よりも、そのギャップが小さくなり、この程度の想定確信度の違いは言語表現に反映されにくいため、語用論的に paradox が解消されるといってもよい。なお、「ヨグが黒の時に勝つ確率が 0」でない場合には、元の推論と対偶表現における想定確信度の違いはより大きなものになるが、それでも Kratzer の挙げている論理的な (すなわち客観的確率に基づく) 量化の値よりも違いは必ず小さくなる。したがって、Grice's paradox における計算論的関連性理論の見方は、以下のようになる。

- ある推論表現とその対偶表現は語用論的に一致するとは限らない
- ただし、その食い違いは「論理的な計算」から導き出される。ものよりも常に小さい値となり、特にある条件を満たす場合 (前件が真でありかつ後件が偽である想定確信度が 0 に近い場合、あるいは前件が偽でありかつ後件が真である場合の想定確信度が 0 に近い場合) には、Grice's paradox はほぼ解消される。

4.2 スキーマのパラドックス

人間の行う推論が、論理学の含意計算とは異なる性質を持つことを示す代表的な例として、Wason による「四枚カード選択問題」がよく知られている。Wason 選択課題の最も一般的な形は以下のようなものだ。

(21) ある工場では、次の規則に従って、表に文字、裏に数字を印刷したラベルを製造しています。

- ラベル製造規則： もし表が "A" なら、裏には "7" が印刷されている

今、この工場で作られた次のような 4 枚のラベルがあります。ラベル 1, 2 は表が見えており、ラベル 3, 4 は裏が見えています。この 4 枚のラベルについて、上の規則が守られているかどうか確かめる時、<u>最低限</u> どのカードをひっくり返して調べる必要がありますか。

(表)	(表)	(裏)	(裏)
A	G	7	2

論理学の含意計算に基づけば、$\boxed{\text{A}}$ と $\boxed{2}$ のカードを選択するのが正解ということになるが、多くの人は $\boxed{2}$ のカードを選択せず、$\boxed{\text{y}}$ のカードを選択してしまう。同値計算に基づくと、4 枚のカード全てを選択しなければならないため、含意計算であれ同値計算であれ、いずれにせよ論理的な答と一致しない選択が行われるのである。

しかし、Wason 課題では、課題の内容によっては時に劇的に正答率が上がるという興味深い性質がある。例えば、次のような問題を考えてみよう。

(22) ある工場では、次のような就労規則があります。

- もし休日に労働したなら、平日に休みが取れる。

今、この工場で働いている人について、就労規則が守られているかどうかを調べたいとします。あなたが**雇用者側**で、労働者が就労規則を守っているか調べるとするなら、どういうタイプの人を調べればよいでしょうか。また、あなたが**労働者側**で、雇用者が就労規則を守っているか調べるとするなら、どういうタイプの人を調べますか。(a) 休日に労働した人、(b) 休日をきちんと取れた人、(c) 平日に休んだ人、(d) 平日に労働した人

このWason課題の興味深いポイントは、雇用者側の観点から考えた場合と、労働者側の観点から考えた場合とで、反応パターンが異なる点にある。雇用者側の立場では、前件否定である (b) と後件肯定である (c) が選ばれやすい。すなわち、論理通りの演繹推論とは全く異なる反応パターンとなる (5.2節で見た一般的なWason課題とも違い、前件否定が選ばれやすくなる点に注意されたい)。しかし、労働者側の立場に立つと、一転して、正解である前件肯定である (a) と後件否定である (d) の選ばれる確率が上昇する。

Cheng and Holyak (1985) は、こうした現象を義務スキーマ・許可スキーマという方略の影響によるものであると説明している。義務スキーマとは、上記の就業規則のような法的言明を (23) のような方略で解釈することを言い、許可スキーマとは (24) に示す方略で解釈することを指す。ゴシックで示したモダリティの点で、両者はちょうど鏡像関係となる。

(23) a. もし前件を満たすなら、後件を満たしていな**ければならない**。
b. もし前件を満たさないなら、後件を満たしていな**くてもよい**。
c. もし後件を満たすなら、前件を満たしてい**てもよい**。
d. もし後件を満たさないなら、前件を満たしてい**てはいけない**。

(24) a. もし前件を満たすなら、後件を満たしてい**てもよい**。
b. もし前件を満たさないなら、後件を満たしてい**てはいけない**。
c. もし後件を満たすなら、前件を満たしていな**ければならない**。
d. もし後件を満たさないなら、前件を満たしていな**くてもよい**。

ここで、従業員の立場に立つと、就労規則は「義務スキーマ」として解釈されるため、必然性の高い前件肯定と後件否定が選ばれやすくなり、結果的に演繹推論と同一の結果が得られる。一方、雇用者側は就労規則を「許可スキーマ」として見なしやすいため、蓋然性の高い前件否定と後件肯定が選択される。

計算論的関連性理論に基づくと、こうしたスキーマ自体が認知環境における想定確信度から自然に計算される推論であるとことが分かる。まず労働者側の立場からこの問題を見てみよう。働く側からすれば、休日に労働し、かつ平日にも労働するという事は考えられない、という心理を自然に持つ。したがって、次のような認知環境が構成される。

(25)

従業者		平日	
		休み	労働
休日	労働	$\mathcal{A}_{xy}$	$\mathcal{A}_{x\bar{y}} = 0$
	休み	$\mathcal{A}_{\bar{x}y}$	$\mathcal{A}_{\bar{x}\bar{y}}$

(25) において、$X \wedge \neg y$ の想定値 $\mathcal{A}_{x\bar{y}}$ が 0 になっている点に注目されたい。3.2 節で見た通り、これは正に命題論理における含意と等価な認知環境である。したがって、労働者側に立った場合、規則を「正しく」理解できるようになるのは自然なことなのだ。

さらに、平日が休みで休日も休みになるという想定の確信度 $\mathcal{A}_{\bar{x}y}$ も、0 とは言わないまでも、0 に近い $\mathcal{A}_{\bar{x}y} \approx 0$ になると考えられる。この時、式 (6) により $\mathcal{A}_{x \to y}$ の値を求めてみると、$\frac{\mathcal{A}_{xy}}{\mathcal{A}_{xy} + \mathcal{A}_{x\bar{y}}} = 1$, $\frac{\mathcal{A}_{\bar{x}y}}{\mathcal{A}_{\bar{x}y} + \mathcal{A}_{\bar{x}\bar{y}}} \approx 0$ より、$\mathcal{A}_{x \to y} \approx 1$ となることがわかる。(9b) で見たように、この $\mathcal{A}_{x \to y}$ の確信度はモダリティにおける必然性「〜に違いない、〜ねばならない」に相当する値である。これが、(23a) のスキーマが生じる理由である。

もう一方の雇用者側から規則を見てみよう。雇用者にとって許し難いのは、平日も休日も休んで、仕事をしない労働者である。したがって、(26) のような認知環境を持つ。

(26)

雇用者		平日	
		休み	労働
休日	労働	$\mathcal{A}_{xy}$	$\mathcal{A}_{x\bar{y}}$
	休み	$\mathcal{A}_{\bar{x}y} = 0$	$\mathcal{A}_{\bar{x}\bar{y}}$

これは、命題論理における含意とも同値とも等価にならない認知環境である。したがって、雇用者側に立つと、規則を論理に合った形で理解することが困難になる。また、雇用者にとっては、$\mathcal{A}_{\bar{x}y}$ の値さえ 0 であれば、他の $\mathcal{A}_{xy}, \mathcal{A}_{x\bar{y}}$ の想定値は同じような大きさであってよい。したがって、(26) の認知環境から、式 (6) によって推論の想定確信度を求めると、$\mathcal{A}_{x \to y} > 0$ の条件しか満たさない。これは (9d) の可能性性に関するモダリティに相当する値であり、したがって (24a) のような「〜であり得る」という許可スキーマを生み出すことになる。このように、計算論的関連性理論に基づくと、どのような状況において人間が論理的な結果と同じ行動を行い、どのような状況では食い違った行動を行うのかを正確に予測することができる。

5. 総合論議

本稿では、まず関連性理論を計算論的に形式化する手法を概観し、それが論理とどのような関係にあるのかをいくつかの例を通して見た。本稿で主張したい点は、以下の通りである。

(27) a. 人間の思考・推論は関連性に基づいて行われる。

b. 関連性の計算における各パラメータの調整方法を「学習」した人は、あたかも「論理的な思考」が可能であるかのように振る舞える。すなわち、人は「論

理」を知ってはいないが、数学的な論理を実行するための「関連性の計算方法」は学習可能である。

c. 同様に、計算論的関連性理論における想定確信度は、論理における真理値に変換可能である。

d. 想定確信度は、様相論理における様相オペレータと等価な計算を行える。

e. 様相性を持った論理においては、対偶表現に paradox が生じることがあるが、計算論的関連性理論に基づくと、どのような場合に paradox が解消され、どのような場合に paradox が残るのかを予測可能である。

f. また、対偶表現の paradox はそもそも語用論的に paradox ではない。最も適切な表現を選択する語用論的計算においては、ある含意表現とその対偶表現は必ずしも一致する必要はない。

本稿で述べた議論は、さらに一般的な推論や帰納推論にも拡張可能である。これらの議論については、また稿を改めて議論を行いたい。

参考文献

Cheng, P. W. & Holyak, K. J. (1985). Progmatic reasoning schemas. *Cognition*, **17**, 391–416.

Gillies, Donald (2000). *Philosophical Theories of Probability*. Routledge.

Kratzer, Angelika (1986). Conditionals. *Chicago Linguistic Society*, **22**, 1–15.

Marin, Arthur (1999). Information, relevance, and social decisinmaking: some principles and results of decision-theoretic semantics. In Moss, L.S., Ginzburg, J., & de Rijke, M. (Eds.), *Logic, Language, and Computation. Vol2.*, pp. 179–221. Stanford CA: CSLI Publications.

Marin, Arthur (2003). *Relevance and Decision-Theoretic Semantics*. Handout. 15th European Summer School in Logic, Language and Information. August 18–29, 2003, Technical University of Vienna.

Sperber, Dan, Cara, F., & Girotto, V. (1995). Relevance theory explains the selection task. *Cognition*, **57**, 31–95.

Author's E-mail Address: `matsui@sils.shoin.ac.jp`
Author's web site: `http://sils.shoin.ac.jp/~matsui/`