計算論的関連性理論における日本語条件文の解釈*

松井 理直

Interpretations of Japanese Conditionals on Computational Relevance Theory

Michinao F. MATSUI

Abstract

Everyday situations require us to reason. For solving complex problems in our real world, it is important not only to get explicit information from the outside world, but also to seek and identify appropriate information by selecting it from a huge amount of knowledge stored in memory. Therefore, reasoning is fundamental to human intelligence. The most important process of reasoning is to select optimal knowledge which is essential to interpretation of current information, and to ignore inappropriate knowledge which is irrelevant. In Relevance Theory, it is claimed that an optimal relevance gives the most appropriate interpretation by means of deductive inference. This paper provides a way of computational formalization of Relevance Theory and proposes a system of an interpretation of some aspects of Japanese conditionals.

1. 序論

推論は人間の思考における最も重要な特性の一つである。推論によって、我々は直接的な経験を経ずして、新たな知識を獲得することができ、現実に適切に対処することができ、過去の出来事の原因を理解し、未来の事態を予測することができる。素晴らしいことに、人間はしばしば適切な推論思考を行い、様々な問題に対してうまく対応することができる。

もちろん、人間が常に正しい思考・推論を行うとは限らない。これにはいくつかの理由がある。まず、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。ま

*本研究は、日本学術振興会科学研究費補助金・基盤研究 (A)「日常的推論の論理と言語形式: 量化表現、条件文、モーダル表現を中心として」(平成 15 年度～平成 18 年度、研究代表者: 郡司 隆男、課題番号 15202009) の援助を受けている。

た、我々の能力そのものにも限界がある。人間という認知主体を取り巻く環境は極めて範囲が広く、かつ常に情報が流動的に変化している世界である。認知主体の持つ知覚・思考・伝達といった情報処理能力は、外界に存在する膨大な情報の一部分しか処理することができない。したがって、完全解を常に求めることができるとは限らない。重要なことは、認知主体がを少しでもよりよい解を得るために、部分情報を手がかりにして可能な限り安定した体制化と推論を行おうとしている点にある。認知の本質は、巨大な認知環境が内包する多様性に適応していく能力といってもよい。

Sperber と Wilson によって提案されている関連性理論 (Sperber & Wilson, 1986) は、認知主体が部分情報からいかに適切に情報の体制化を行うかという問題に対する極めて興味深い理論である。現在、この理論は言語・思考・知覚から社会文化に至る認知活動の幅広い分野に応用されており、人間の知的活動全般を支配する性質を考える上で、大変に重要なモデルを提案している。また、理論の基盤となる前提や原理が明示的に規定されているため、理論を形式的に表現できる可能性を持っている点も魅力である。本稿では、この関連性理論を定量的に形式化する手法の一つを提案し、それに基づいて、言語に関連する推論、特に条件文の解釈過程について考察を行う。

2. 計算論的関連性理論

2.1 関連性理論

まず始めに、関連性理論の枠組みを簡単に概観しておこう。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。

認知主体は、部分情報に一貫性を持たせ、体制化されたものにするため、外界の情報間あるいは自らの想定の間に常に関連性を求める存在である。Sperber & Wilson は、こうした性質を**関連性の認知原理** (*cognitive principle of relevance*) と呼んでいる。

(2) 関連性の認知原理：
人間の認知系は自らにとって関連のある情報に注意を払うようデザインされている。

さらに、情報の受け取りが動的に行われるコミュニケーションでは、**関連性の伝達原理** (communicative principle of relevance) が存在し、情報的意図と伝達的意図が関与する。次節から、この関連性理論の諸定義を計算論の立場から形式化した **計算論的関連性理論**について議論を行う。[1]

2.2 関連性理論の計算論的見解

今、認知主体の作業記憶に、ある意味論 (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 が生じるとする。この時、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。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。

(3) 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。[2] 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。

(3d) のうち、(I) は (3c) の (ii), (iii) より自明である。また、(3c) の (i) も間接的に (3d-I) の効果をもたらすが、これは次節でより詳しく述べる。(3d) 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。[3] この推論の想定確信度については 4 節で議論を行う。

[1]関連性理論の形式化は、本稿で提案する方法以外に、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。

[2]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ができる。しかし、想定確信度は多値の真理値そのものではない。3.1 節を参照のこと。

[3](3d) 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。

2.3 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。

(4) 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。想定確信度と命題の真理値の関係については、3.1 節でより詳しく議論を行う。

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 ((3a) 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	$\mathcal{A}_{xy}$	$\mathcal{A}_{xy_2}$	$\mathcal{A}_{xy_3}$	$\mathcal{A}_{xy_4}$	$\cdots$
	X_2	$\mathcal{A}_{x_2y}$	$\mathcal{A}_{x_2y_2}$	$\mathcal{A}_{x_2y_3}$	$\mathcal{A}_{x_2y_4}$	$\cdots$
	X_3	$\mathcal{A}_{x_3y}$	$\mathcal{A}_{x_3y_2}$	$\mathcal{A}_{x_3y_3}$	$\mathcal{A}_{x_3y_4}$	$\cdots$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

$= \mathcal{A}_{x\bar{y}}$

$\parallel$ $\mathcal{A}_{\bar{x}y}$

$\mathcal{A}_{\bar{x}\bar{y}}$

2.4　想定の変動と関連性の計算

(1), (4) の想定値は相対的なものなので、ある想定値が変化すると、他の想定値にも影響を与える。これは (1c) に示した認知効果の数量的表現となり、このことから (1d) の関連性の計算が可能となる。まず、想定値の変動に伴う効果について見てみよう。今、(5) のような認知環境が構成されている状況下で、想定 $X \wedge Y$ の確信度 $\mathcal{A}_{xy}$ が $+\delta$ だけ変化したとする。この時、新たな認知環境における各想定の確信度は、(6) に示すようにいくつかの可能性が考えられる。

(5)

	$\mathcal{A}_y$	$\mathcal{A}_{\bar{y}}$	計
$\mathcal{A}_x$	a	b	$a+b$
$\mathcal{A}_{\bar{x}}$	c	d	$c+d$
計	$a+c$	$b+d$	1

(6) a.

	$\mathcal{A}_y$	$\mathcal{A}_{\bar{y}}$	計
$\mathcal{A}_x$	$\frac{a+\delta}{1+\delta}$	$\frac{b}{1+\delta}$	$\frac{a+\delta+b}{1+\delta}$
$\mathcal{A}_{\bar{x}}$	$\frac{c}{1+\delta}$	$\frac{d}{1+\delta}$	$\frac{c+d}{1+\delta}$
計	$\frac{a+\delta+c}{1+\delta}$	$\frac{b+d}{1+\delta}$	1

b.

	$\mathcal{A}_y$	$\mathcal{A}_{\bar{y}}$	計
$\mathcal{A}_x$	$a+\delta$	$b-\delta$	$a+b$
$\mathcal{A}_{\bar{x}}$	c	d	$c+d$
計	$a+\delta+c$	$b-\delta+d$	1

c.

	$\mathcal{A}_y$	$\mathcal{A}_{\bar{y}}$	計
$\mathcal{A}_x$	$a+\delta$	b	$a+\delta+b$
$\mathcal{A}_{\bar{x}}$	$c-\delta$	d	$c-\delta+d$
計	$a+c$	$b+d$	1

d. $\cdots$

このことから、ある想定値が 0 である (7-i) と、ある想定が認知環境に組み込まれていない (7-ii) は計算論的な意味が異なることが分かる。

(7) (i)

	$\mathcal{A}_{y_1}$	$\mathcal{A}_{y_2}$	計
$\mathcal{A}_{x_1}$	a	$1-a$	1
$\mathcal{A}_{x_2}$	0	0	0
計	a	$1-a$	1

(ii)

	$\mathcal{A}_{y_1}$	$\mathcal{A}_{y_2}$	計
$\mathcal{A}_{x_1}$	a	$1-a$	1
計	a	$1-a$	1

(7-i) では、想定 X_2 は偽と確信されており、変更が困難である。こうした認知環境の効果は、6 節で見る反事実条件文の解釈などに影響を及ぼす。一方、(7-ii) では、新規の想定 X_2 が得られれば、既知の想定 X_1 の確信度 $\mathcal{A}_{x_1}$ も変化する。これは、(1c), (3c) で見た「新規情報の獲得は関連性の認知効果が高い」ことに対応する計算論的な性質である。

新規情報の獲得と関連性の認知効果に関しては、認知環境の設定そのものにも影響を与える。(6) から分かるように、ある想定の確信度が変化した時、それが認知環境全体にどのような影響を及ぼすかは一意には決定できない。こうした可能性を絞り込むのが、(1d), (3d) に示す関連性の効果である。(1c-i), (3c-i) の「新規想定の獲得が認知効果を高める」という性質は、既に述べたように、新たな想定を認知環境に組み込むことが、いくつかの既存の想定に関する確信度を 1 あるいは 0 に近づける作用を果たすからなのだ。したがって、複数の認知環境の候補があった場合には、確信度が 1 か 0 に近づく既存想定が多いほど、より良い認知環境であるといってよい。これによって、認知環境の候補を一意に絞り込むことが可能となる。

2.5 談話領域と想定確信度の変更可能性

既存想定の確信度を変更する効果を持つ情報は、新規想定のみとは限らない。既に認知環境に組み込まれている想定をより補強するような外界情報、あるいは矛盾するような外界情報も、既存想定の確信度に変更を迫る効果を持つ。また、新規想定を認知環境に組み込んだ時に、想定確信度を容易に変更可能な既存想定と、それが難しい既存想定があると考えられる。これは、(1d), (3d-I) で述べた認知環境の改善に掛かるコストに関わる問題といってよい。

この想定確信度の変更可能性に有意義な示唆を与える枠組みとして、田窪・金水 (1996), 田窪 (1993) によって提案されている談話管理理論 (*Discourse Management Theory*) を取り上げよう。この理論は、話者の談話領域における情報管理のされ方を明確に区別することで、話者と聴き手における知識参照の無限後退という問題をクリアしている理論である。談話管理理論では、話者の談話領域として次の 2 種類を設定する。

(8) a. **D-領域：** 長期記憶内の、既に検証され、同化された直接経験情報や対話現場に存在する直接情報が格納される領域。換言すれば、ある 1 つの談話セッションが始まる以前に獲得されていた情報を格納する領域。

b. **I-領域：** まだ検証され終わっていない情報、推論・伝聞・仮定などで、間接的・仮想的に設定される情報を格納する領域。

想定確信度という観点から談話領域を考えると、D-領域にある単独想定は、その確信度が1か0に近く、また確信度の変更が行われにくいものと見なすことができる。一方、I-領域にある単独想定は、確信度の値はその談話セッションで暫定的に決められ、変更が容易なものと考えてよい。すなわち、I-領域にある想定を変更するほうが、D-領域にある想定を変更するよりもコストが低い。この変更に掛かるコストも認知効果に関わる条件である。本稿では、I-領域に存在する想定の確信度を $(\mathcal{A}_x)$ のように括弧で囲んで表示し、関連性の高い情報を求める際にD-領域とI-領域の情報を区別することで、最適な認知環境の導出を行う。

3. 計算論的関連性理論と論理

3.1 想定確信度・エントロピー・真理値

前節で見たように、想定の確信度とは、ある命題表象に対して与えられる量的な価値であり、情報間の関連性の計算に重要な役割を果たす。本項では、この想定確信度の意味をより明確にするため、命題論理における真理値との関係について議論を行う。

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。

(9) 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 (10) 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。

(10) 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。

(11) 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、$\mathcal{A}_x \geq 0.5$ 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、(11) 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考え方に近づく。何よりも大切なことは、エントロピーと真理値が一定の計算式によって関係付けられるという点、さらにそのエントロピー自体が想定確信度を用いて計算可能であるという点にある。認知環境における想定確信度の状態、すなわち想定間の関連性は、エントロピーの計算や命題の真理値の基礎を成していると言ってもよい。

この性質は、人間の心的情報処理を考える上でも重要なことである。例えば、井上 (2000), Inoue and Den (1999) は、日本語の文理解過程における曖昧性解消に用いられる予測可能性に、エントロピーの計算が極めて重要な役割を果たすと述べている。彼らの研究を計算論的関連性理論の立場から再解釈を行うと、文理解における曖昧性解消に名詞・動詞間における情報の関連性が強く影響するという形で捉え直すことができるだろう。

3.2 様相性

計算論的関連性理論では、想定確信度を用いて、様相論理に近いモダリティの計算も可能である。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 (12) のような対応が存在すると考えられる。

(12) 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、(12) 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。

(13) 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 (13b) の状態とほぼ同一であり、「X ができる」という表現は (13d) の状態とほぼ同じであると考えてよい。このことは、英語において義務的必然性 (deontic necessity) と認識的必然性 (epistemic necessity) がどちらも同じ助動詞 *must* で表現され、また義務的可能性 (deontic possibility) と認識的可能性 (epistemic possibility) がどちらも同一の *can* で表されることに通じる。日本語では、認識的態度と義務的態度は全て語彙的に区別されるが、これは命題の心的態度、特に聴き手に対する認知環境の変更指令のあり方に違いがあり、その違いが言語表現に反映されているものと思われる (この点は本稿では議論を行わない)。

3.3 想定確信度に対する論理演算子の機能

(13) からも分かる通り、関連性理論では、ある発話内容は話者の認知環境を反映したものであると共に、聴き手にある特定の認知環境を形成するよう働きかける指令であると見なすことができる。否定 ($\neg$), 連言 ($\wedge$), 選言 ($\vee$), 排他的選言 ($\oplus$), 含意 ($\rightarrow$) といった論理結合子も、複数の想定を持つ認知環境を構成する指令であるといってよい。

例えば、(14) に示すような認知環境を持っている話者は、「X かつ Y である」という連言を用いた発話を行い、聴き手はこの発話を、自らの認知環境における $X \wedge Y$ の想定確信

度 $\mathcal{A}_{xy}$ を 1 に設定せよという指令として受け取る (聴き手がどのタイプの認知環境を構成するかは、この発話を受け取る直前の状態に依存する)。

(14) a.

想定	Y	$\neg Y$
X	$\mathcal{A}_{xy}=1$	$\mathcal{A}_{x\bar{y}}=0$
$\neg X$	$\mathcal{A}_{\bar{x}y}=0$	$\mathcal{A}_{\bar{x}\bar{y}}=0$

b.

想定	Y
X	$\mathcal{A}_{xy}=1$
$\neg X$	$\mathcal{A}_{\bar{x}y}=0$

c.

想定	Y	$\neg Y$
X	$\mathcal{A}_{xy}=1$	$\mathcal{A}_{x\bar{y}}=0$

d.

想定	Y
X	$\mathcal{A}_{xy}=1$

連言と同様に、否定も比較的単純な認知環境の操作指令である。否定の場合は、それに対応する表現が語彙化可能か否かによって、認知環境におけるあり方が変化するが (表 1, 表 4 の違いを参照のこと)、一般に $1-\mathcal{A}_x$ がその想定確信度となる。

比較的安定した計算である連言・否定とは異なり、「X か Y である」という選言を用いた表現は少し複雑な計算過程となる。選言が聴き手の認知環境を変更する指令として機能する時、(15) のような指令として解釈される。いずれの解釈がなされるかは、連言の場合と同様、直前の聴き手の認知環境の状態に依存する。

(15) a. $1-\mathcal{A}_{\bar{x}\bar{y}}=1$

b. $\mathcal{A}_x+\mathcal{A}_y-(1+n)\cdot\mathcal{A}_{xy}=1$ (ただし $0 \leqq n \leqq 1$)

(15a), (15b) は、どちらも想定 $\neg X \vee \neg Y$ の確信度 $\mathcal{A}_{\bar{x}\bar{y}}$ を 0 に近づける効果があるという共通の計算論的性質を持つ。しかし、その効果が直接的であるか間接的であるかという点が異なる。(15a) は、その計算式からも分かる通り、想定確信度 $\mathcal{A}_{\bar{x}\bar{y}}$ を直接 0 に近づける効果を持つ。すなわち、ダイレクトに変更を受ける想定は $\neg X \wedge \neg Y$ のみであり、このことから、命題論理でいう包含的選言の解釈が導き出される。これに対し、(15b) において想定確信度を直接 0 に近づけるな変更を受けるものは、想定 $X \wedge Y$ の確信度 $\mathcal{A}_{xy}$ である。想定 $\neg X \wedge \neg Y$ の確信度 $\mathcal{A}_{\bar{x}\bar{y}}$ も結果的に 0 に設定されるが、これは 想定 $X, Y, X \wedge Y$ の想定確信度が計算された結果、間接的に 0 になるに過ぎない。こうした計算過程の影響により、(15b) では包含的選言から排他的選言に至るまでの gradual な解釈が成立する。$n=0$ であるなら完全な包含的選言の解釈、$n=1$ の時は完全な排他的選言の解釈である。つまり、n の値によって、包含的選言に近い解釈から排他的選言に近い解釈までの様々な解釈が成り立ち得るわけだ。

4. 認知環境における推論確信度の計算

4.1 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、(3d) 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 (16) 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。

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

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。

4.2 推論確信度の確率論的解釈

式 (16) 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 (16) は

(17)
$$\begin{aligned}\mathcal{A}_{x \to y} &= h_x \cdot \frac{\mathcal{A}_{xy}}{\mathcal{A}_{xy}+\mathcal{A}_{x\bar{y}}} - h_{\bar{x}} \cdot \frac{\mathcal{A}_{\bar{x}y}}{\mathcal{A}_{\bar{x}y}+\mathcal{A}_{\bar{x}\bar{y}}} \\ &= h_x \cdot P(\mathcal{A}_y|\mathcal{A}_x) - h_{\bar{x}} \cdot P(\mathcal{A}_y|\mathcal{A}_{\bar{x}})\end{aligned}$$

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。

4.3 推論確信度の統計学的解釈

式 (16) 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。

回帰係数は変数 X, Y の共分散 $Cov(X,Y)$ を X の分散 $V(X)$ で割ることにより求めることができる。この性質を利用して、表 1 の想定値から回帰係数を計算してみよう。まず、各変数の確率は

$$\begin{aligned}P(X=1) &= \mathcal{A}_{xy}+\mathcal{A}_{x\bar{y}}, \quad & P(X=0) &= \mathcal{A}_{\bar{x}y}+\mathcal{A}_{\bar{x}\bar{y}}, \\ P(Y=1) &= \mathcal{A}_{xy}+\mathcal{A}_{\bar{x}y}, \quad & P(Y=0) &= \mathcal{A}_{x\bar{y}}+\mathcal{A}_{\bar{x}\bar{y}}\end{aligned}$$

となり，ここから変数 X, Y の期待値

$$\begin{aligned}E(X) &= \mathcal{A}_{xy}+\mathcal{A}_{x\bar{y}}, \quad & E(X^2) &= \mathcal{A}_{xy}+\mathcal{A}_{x\bar{y}}, \\ E(Y) &= \mathcal{A}_{xy}+\mathcal{A}_{\bar{x}y}, \quad & E(Y^2) &= \mathcal{A}_{xy}+\mathcal{A}_{\bar{x}y} \qquad E(XY) = \mathcal{A}_{xy}\end{aligned}$$

が得られる。したがって，変数 X の分散 $V(X)$ および X, Y の共分散 $Cov(X, Y)$ は，

$$V(X) = E(X^2) - (E(X))^2 = (\mathcal{A}_{xy}+\mathcal{A}_{x\bar{y}})(\mathcal{A}_{\bar{x}y}+\mathcal{A}_{\bar{x}\bar{y}})$$

$$Cov(X, Y) = E(XY) - E(X)\cdot E(Y) = \mathcal{A}_{xy}\mathcal{A}_{\bar{x}\bar{y}} - \mathcal{A}_{x\bar{y}}\mathcal{A}_{\bar{x}y}$$

となる。これらの値を用いて x から y への回帰係数 β は次のようになり、式 (16) における $h_x, h_{\bar{x}}$ を 1 に設定した式に一致する。

(18)
$$\beta = \frac{Cov(X, Y)}{V(X)} = \frac{\mathcal{A}_{xy}}{\mathcal{A}_{xy}+\mathcal{A}_{x\bar{y}}} - \frac{\mathcal{A}_{\bar{x}y}}{\mathcal{A}_{\bar{x}y}+\mathcal{A}_{\bar{x}\bar{y}}} \quad (= \frac{\mathcal{A}_{xy}}{\mathcal{A}_x} - \frac{\mathcal{A}_{\bar{x}y}}{\mathcal{A}_{\bar{x}}})$$

式 (16) 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。(3d) 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。この点については 5 節および 6 節で議論を行う。

4.4 ゼロを巡る演算と係数

式 (16) における係数 $h_x, h_{\bar{x}}$ は、情報を参照するフレームの大きさと見なすことができる。認知科学において、フレーム問題は長く議論されてきたものであるが、推論想定の確信度を計算する上で代数的にも興味深い性質を持っているので、本項ではこの性質を議論してみたい。

ポイントとなるのは、ゼロの割り算に関る代数的性質である。一般に $a = \frac{b}{c}$ という演算において、c がゼロであることは避けられる。これはゼロの除算がいくつかの不安定な性質を持つためである。まず、$a = \frac{b}{0}$ において、b≠0 の場合を考えてみよう。$a = \frac{b}{c}$ は $a\cdot c = b$ と等価でなければならないが、b≠0, c=0 の時、$b \neq 0$ と $a\cdot 0 = 0$ が同時に成立することになってしまい、両者は矛盾する。したがって、割られる数がゼロでない場合、ゼロによる除算は不能 (演算ができない) となる。一方、$b = 0$ の時は、c=0 であっても、b =0 と $a\cdot 0 = 0$ は同値となり、成立し得る。したがって、0÷0 の演算は、どのような場合でも常に計算可能であり、その答えは無数にある、すなわち演算は不定であるという

ことになる。また、解が不定であっても、それに再びゼロを掛けると、ゼロの積算結果は常にゼロであるため、$0 \cdot \frac{0}{0} = 0$ が成立する。

ここで (16) 式について見てみると、例えば、$\mathcal{A}_{x\to y} = h_x \cdot \frac{\mathcal{A}_{xy}}{\mathcal{A}_{xy}+\mathcal{A}_{x\bar{y}}} - h_{\bar{x}} \cdot \frac{\mathcal{A}_{\bar{x}y}}{\mathcal{A}_{\bar{x}y}+\mathcal{A}_{\bar{x}\bar{y}}}$ において、分母がゼロになる時には、常に分子もゼロになることが分かる。すなわち、(19) 式の演算は、最悪の場合に解が不定になることはあっても、演算が不能になることは決してない。常に演算が実行可能であるという点は、言語の意味解釈を行う上で、一つの重要な性質である。

しかし、意味解釈という場面を考えるなら、演算結果が不定であるという状態は困難を伴う。これはエントロピー最大の状態であり、「非文ではないが、その意味は全く分からない」という、認知環境の改善をもたらさない事態である。もし、認知主体がこういう演算結果に遭遇した時、なんらかの coercion を働かせてでも、最適な意味が導出できるような処理を行わなければならない。この時、(16) 式の係数が重要な役割を果たす。例えば (16) 式の中で、$\frac{\mathcal{A}_{\bar{x}y}}{\mathcal{A}_{\bar{x}y}+\mathcal{A}_{\bar{x}\bar{y}}}$ の部分が $0 \div 0$ となり、解が不定になってしまうことがあったとしよう。これは、$\mathcal{A}_{\bar{x}y}, \mathcal{A}_{\bar{x}\bar{y}}$ のいずれもが 0 になるということに等しいので、想定 $\neg X$ そのものが不成立となっている状態であることが分かる。このままでは、(16) 式全体の解も不定になってしまうわけであるが、ここで係数 $h_{\bar{x}}=0$ を設定すれば、$0 \cdot \frac{0}{0} = 0$ の性質により、解が不定になるという最悪の結果を避けることができる。係数 $h_{\bar{x}}$ を 0 に設定するということは、$\neg X$ を参照フレームから除外するということに他ならない。換言すれば、想定 $\neg X$ そのものが不成立である時、それをフレームから除外することで推論行為を達成可能なのだ。

4.5 日常的推論の一般式

式 (16) 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 (16) 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。[4]

(19) 日常的推論の確信度に関する一般式

$$\begin{aligned}\mathcal{A}_{x\to y} &= \frac{\mathcal{A}_{xy}}{\mathcal{A}_{xy}+\mathcal{A}_{x\bar{y}}} - k_{\bar{x}} \cdot \frac{\mathcal{A}_{\bar{x}y}}{\mathcal{A}_{\bar{x}y}+\mathcal{A}_{\bar{x}\bar{y}}} \\ &= P(\mathcal{A}_y|\mathcal{A}_x) - k_{\bar{x}} \cdot P(\mathcal{A}_y|\mathcal{A}_{\bar{x}})\end{aligned}$$

[4] 言うまでもないが、係数 $k_{\bar{x}}$ の役割についても、4.4 節の議論が成立する。

連言・選言・否定と同様、推論表現 (自然言語ではしばしば if 文で表現される) も、話者の認知環境の反映であり、同時に聴き手に対する認知環境の変更指令として機能する。次節では、式 (19) に基づいて、条件文解釈の心的過程ががどのような性質を持っているのかという点について議論してみよう。

5. 条件文の計算過程

5.1 論理学における真理値について

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。(20) 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。

(20)

x	y	x→y	x↔y
真	真	真	真
真	偽	偽	偽
偽	真	$\boxed{\text{真}}$	偽
偽	偽	$\boxed{\text{真}}$	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 (12) の違いを、式 (19) に反映させてみよう。

まず、(12a) 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 (19) において、$0 \leqq \frac{\mathcal{A}_{xy}}{\mathcal{A}_{xy}+\mathcal{A}_{x\bar{y}}} \leqq 1$, $0 \leqq \frac{\mathcal{A}_{\bar{x}y}}{\mathcal{A}_{\bar{x}y}+\mathcal{A}_{\bar{x}\bar{y}}} \leqq 1$ 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、$\frac{\mathcal{A}_{xy}}{\mathcal{A}_{xy}+\mathcal{A}_{x\bar{y}}} = 1$, $\frac{\mathcal{A}_{xy}}{\mathcal{A}_{\bar{x}y}+\mathcal{A}_{\bar{x}\bar{y}}} = 0$ 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 (20) 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、(12b) に相当する $\mathcal{A}_{x\to y}\geqq 0$ を見てみよう。式 (19) 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、$\frac{\mathcal{A}_{xy}}{\mathcal{A}_{xy}+\mathcal{A}_{x\bar{y}}} = 1$ 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、(20) 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 (19) 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 $\frac{\mathcal{A}_{xy}}{\mathcal{A}_{xy}+\mathcal{A}_{x\bar{y}}}=1$ 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。

5.2 Wason 選択課題

前項で、命題論理における含意と同値が計算論的関連性理論においてどのように導出されるかを見たが、実際に人間が行う推論過程は、含意でも同値でもなような、中途半端な判断が行われることが知られている。中でも、Wason 選択課題 (四枚カード問題) は、人間の推論過程を最も端的に示す心理実験の一つであり、代表的な実験として、例えば次のようなものが挙げられる。

(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：$\boxed{\text{もし表が “A” なら、裏には “7” が印刷されている}}$

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、最低限 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

規則を含意解釈として捉えるなら、$\boxed{\text{A}}$ と $\boxed{2}$ のカードを選択すると正解になり、同値解釈として捉えるなら全てのカードを選択する必要がある。しかし、実際にはこのタイプの Wason 選択課題では、かなりの被験者が $\boxed{\text{A}}$ と $\boxed{7}$ のカードを選択することが知られている (大学生でもほぼ 20% 前後の正答率しか得られないことが多い)。その判断の偏りを生じさせる認知的バイアスとして、確証バイアス (Wason, 1966) やマッチングバイアス (Evans & Lynch, 1973)、主題化効果などが提案されてきた。最近では、Sperber et al. (1995) が関連性理論により、様々なタイプの Wason 選択課題が包括的に説明できると主張している。例えば上記のタイプの実験では、規則に明示された関連性の高い情報に引きずられた結果として判断錯誤が引き起こされるのだと Sperber は考える。ここで、彼の議論を式 (19) により形式的に捉えてみよう。

まず初めに、判断錯誤の特徴をより明確にするため、Wason 課題の正しい解法を見ておく。5.1 節で見たように、規則を含意として捉えるなら、$\mathcal{A}_{x\bar{y}}$ が 0 であることを確認すればよい。すなわち、想定 $X\wedge\neg Y$ が棄却できるかを確認すればよい。したがって、X の想定に関わる事実と $\neg Y$ に関わる事実を「共に」調べる必要があり、そのために、$\boxed{\text{A}}$ と $\boxed{2}$ のカードを選択しなければならない。一方、規則を同値として解釈するのであれ

ば、$\mathcal{A}_{x\bar{y}}$ が 0 であり、かつ $\mathcal{A}_{\bar{x}y}$ が 0 であることを確認することになる。つまり、$X\land\neg Y$ と $\neg X\land Y$ のいずれの想定をも棄却できる事を調べる必要がある。したがって、$X\land\neg Y$ の棄却可能性を調べるため、X と $\neg Y$ の想定に関わる事実を「共に」調べる必要があると同時に、$\neg X\land Y$ の棄却可能性も検証するため、想定 $\neg X$ に関わる事実と想定 Y に関わる事実を「共に」調べる必要に迫られる。この結果、全てのカードが選択される。

しかし、判断錯誤を起こす被験者は、規則から得られる論理に対応した事実を「共に」調べるという演繹推論の作業がを行えない。逆に、彼らは目の前の「事実」から構成可能な認知環境を設定し (すなわち事実から得られる可能な限りの想定を行い)、その認知環境から規則の想定確信度を得るために、式 (19) に則った計算を行って、事実と規則がどの程度の「関連性」を持っているのか調べようとする。しかも、その計算を各事実ごとに「独立に」検証しようと試みる。これこそが Wason 課題が持つ難しさのポイントである。Wason 課題に失敗する原因は、正にこの「独立に」検証するという作業にあるのだ (本当は、前述したように、演繹的に、かつ「共に」検証しなければならない)。その失敗する過程を順に見てみよう。

まず、被験者は規則を読む。この規則から被験者が直接に得る想定は、まず $\boxed{\text{A}}$ に関する想定 X と $\boxed{7}$ に関する想定 Y であり、その確信度各々 1 である。また、この規則は直接経験で確認されているわけではなく、今の時点で初めて得られる知識であるため、これらの想定は I-領域に置かれる。したがって、(22) のような認知環境が設定される。

(22)

	Y : $\boxed{7}$	計
X : $\boxed{\text{A}}$	($\mathcal{A}_{xy}$=1)	($\mathcal{A}_{x}$=1)
計	($\mathcal{A}_{y}$=1)	1

さらに進んで、規則を読み終えた被験者は、想定 X と Y が条件文の関係で結ばれること、すなわち想定 $X\rightarrow Y$ を得て、その想定確信度を設定する。ここで、今までの過程と異なるのは、この条件文の想定確信度が直接外界の事実から得られるものではなく、式 (19) による計算が必要な点である。今、規則に書かれているのは、'if' 条件のみで、'else' 条件は明示されていないため、式 (19) における係数は $k_{\bar{x}}=0$ と設定される。そこで、被験者は想定確信度 $\mathcal{A}_{x\rightarrow y}=\dfrac{\mathcal{A}_{xy}}{\mathcal{A}_{xy}+\mathcal{A}_{x\bar{y}}}$ を計算しようとして、問題に気付く。すなわち、(22) の認知環境には、$\mathcal{A}_{x\rightarrow y}$ の計算に必要な想定 $X\land\neg Y$ の確信度 $\mathcal{A}_{x\bar{y}}$ が存在しないのである。そこで、被験者は間接的にもう一つの想定 $\neg Y$ が必要であることを知り、これを認知環境に組み込む。さらにこの段階では、被験者は $X\rightarrow Y$ を受理しようとするため[5]、確信度 $\mathcal{A}_{x\bar{y}}$=0 (同時に $\mathcal{A}_{\bar{y}}$=0) を暫定的に設定し (規則をそのようや指令として受け取るということだ)、推論の確信度を $\mathcal{A}_{x\rightarrow y}$=1 とする。したがって、認知環境は (23) のように変化する。$\mathcal{A}_{x\bar{y}}$ に括弧が二重についているのは、I-領域にある情報であり、かつそれが間接的に得られたものであるため、想定の変更が極めて容易であることを示す。

[5]これが仮説の確証バイアスの間接的な原因となる。しかし、確証バイアスという方略そのものが存在するのではなく、結果的にそれとほぼ同等の効果が現れるというに過ぎない。

(23)

	Y：$\boxed{7}$	$\neg Y$：$\boxed{7}$以外	計
X：$\boxed{A}$	$(\mathcal{A}_{xy}=1)$	$((\mathcal{A}_{x\bar{y}}=0))$	$(\mathcal{A}_{x}=1)$
計	$(\mathcal{A}_{y}=1)$	$((\mathcal{A}_{\bar{y}}=0))$	1

さて、被験者はいよいよカードを調べに掛かる。選択するべきカードはなるべく関連性の高い顕示的事実、すなわち推論の想定確信度 $\mathcal{A}_{x\to y}$ を 1 か 0 に確定してくれるようなカードだ。そこで、被験者は各カードの関連性がどの程度であるか、各々のタイプごとに「独立に」検証していく (前述したように、この「独立」の検証が落とし穴である)。

まず、被験者は $\boxed{A}$ のカード「だけ」を調べにかかる。このカードを見た被験者は、顕示的情報 (すなわち D-領域に格納される想定) として、$\boxed{A}$ の裏側が未知 (すなわち想定 Y か$\neg Y$ かが不明) であることを知る。そこで、(23) の $\mathcal{A}_{xy}, \mathcal{A}_{x\bar{y}}$ は、D-領域の想定確信度として $\mathcal{A}_{xy}=\mathcal{A}_{x\bar{y}}=0.5$ に書き換えられる。I-領域の確信度を D-領域の確信度として書き換えるため、(24) への変更はコストが掛からず、即時に行われる。この認知環境における推論の確信度は、各想定値を式 (19) に代入して得られる値である $\mathcal{A}_{x\to y}=0.5$ へと一時的に減少してしまう。

(24)

	Y：$\boxed{7}$	$\neg Y$：$\boxed{7}$以外	計
X：$\boxed{A}$	$\mathcal{A}_{xy}=0.5$	$\mathcal{A}_{x\bar{y}}=0.5$	$\mathcal{A}_{x}=1$
計	$\mathcal{A}_{y}=0.5$	$\mathcal{A}_{\bar{y}}=0.5$	1

(24) の推論の確信度は関連性が低いため、被験者は裏側の数字が明確になった場合を仮定し、$\mathcal{A}_{xy}, \mathcal{A}_{x\bar{y}}$ の値を再び変更した時に、認知環境の改善が図れるかを計算する。これは D-領域の想定を I-領域の想定として扱うことになるため、即時の更新は行われず、認知効果が高かった時のみ想定値が変更される。裏の数字は $\boxed{7}$ かそうでないかのいずれかであるから、認知環境としては (25) のような可能性がある。

(25) a.

	Y：$\boxed{7}$	$\neg Y$：$\boxed{7}$以外	計
X：$\boxed{A}$	$(\mathcal{A}_{xy}=1)$	$(\mathcal{A}_{x\bar{y}}=0)$	$(\mathcal{A}_{x}=1)$
計	$(\mathcal{A}_{y}=1)$	$(\mathcal{A}_{\bar{y}}=0)$	1

b.

	Y：$\boxed{7}$	$\neg Y$：$\boxed{7}$以外	計
X：$\boxed{A}$	$(\mathcal{A}_{xy}=0)$	$(\mathcal{A}_{x\bar{y}}=1)$	$(\mathcal{A}_{x}=1)$
計	$(\mathcal{A}_{y}=0)$	$(\mathcal{A}_{\bar{y}}=1)$	1

この (25) の想定を式 (19) に代入する。なお、X の否定想定である $\neg X$ は認知環境に組み込まれていないため、$k_{\bar{x}}=0$ である。その結果、(25a) の想定からは規則の想定確信度として $\mathcal{A}_{x\to y}=1$ が得られ、(25b) の想定からは $\mathcal{A}_{x\to y}=0$ が得られる。すなわち、裏面がどのようなものであれ、$\boxed{A}$ のカードは、最も関連性の高い情報であることが分かる。このことから、被験者は喜んで $\boxed{A}$ のカードを選択する。

次に、被験者は $\boxed{G}$ のカードに意識を集中させる。新しいカードに取り込むので、認知環境は規則から得られる初期状態 (23) にリセットされる。この $\boxed{G}$ のカードから得ら

れる顕示的情報は、$\boxed{\text{A}}$ 以外のカードがあり (想定 $\neg X$ を設定する必要があり)、かつ裏の情報は未知であるというものだ。したがって、D-領域の想定確信度である $\mathcal{A}_{\bar{x}y}=\mathcal{A}_{\bar{x}\bar{y}}=0.5$ を、(23) に組み込む。したがって、認知環境は (26) のように変化する。

(26)

	Y：$\boxed{7}$	$\neg Y$：$\boxed{7}$以外	計
X：$\boxed{\text{A}}$	$(\mathcal{A}_{xy}=0)$	$(\mathcal{A}_{x\bar{y}}=0)$	$(\mathcal{A}_{x}=0)$
$\neg X$：$\boxed{\text{G}}$	$\mathcal{A}_{\bar{x}y}=0.5$	$\mathcal{A}_{\bar{x}\bar{y}}=0.5$	$\mathcal{A}_{\bar{x}}=1$
計	$(\mathcal{A}_{y}=0.5)$	$(\mathcal{A}_{\bar{y}}=0.5)$	1

この認知環境における推論の想定度 $\mathcal{A}_{x\to y}$ は、0÷0 の演算を含むため不定となってしまう。しかも、その不定を消去するための係数が 式 (19) には存在しないため[6]、推論の想定度を決定できない。これは関連性のない状態とは言えないが、非常に不安定な状態であることに代わりはない。そこで、$\boxed{\text{A}}$ のカード同様、裏面が分かった時に、推論の想定値がどのように変更されるのかを計算することになるのだが、裏面の情報が分かったところで、$\mathcal{A}_{xy}=\mathcal{A}_{x\bar{y}}=0$ の部分に影響を与えないため、やはり推論の想定度 $\mathcal{A}_{x\to y}$ の値は不定のままである。そこで、カード $\boxed{\text{G}}$ は認知環境の改善には役立たない情報と見なされ、選択されることはない (実際に Wason 課題で最も選択されにくいカードがこれである)。

次は裏面のカードである。これまでの過程と同様に、被験者は $\boxed{7}$ のカードから、D-領域の想定値 $\mathcal{A}_{xy}=\mathcal{A}_{\bar{x}y}=0.5$ を得て、これを (23) に組み込む。この時の認知環境は (27a) となり、$\mathcal{A}_{\bar{x}\bar{y}}=1-(\mathcal{A}_{xy}+\mathcal{A}_{x\bar{y}}+\mathcal{A}_{\bar{x}y})$ より、間接的に $\mathcal{A}_{\bar{x}\bar{y}}=0$ が得られ、認知環境 (27b) を構成する。この認知環境における推論の想定確信度は、式 (19) より $\mathcal{A}_{x\to y}=0.5$ である。

(27) a.

	Y：$\boxed{7}$	$\neg Y$：$\boxed{7}$以外	計
X：$\boxed{\text{A}}$	$\mathcal{A}_{xy}=0.5$	$(\mathcal{A}_{x\bar{y}}=0)$	$(\mathcal{A}_{x}=0)$
$\neg X$：$\boxed{\text{A}}$以外	$\mathcal{A}_{\bar{x}y}=0.5$		$\mathcal{A}_{\bar{x}}=0.5$
計	$\mathcal{A}_{y}=1$	$(\mathcal{A}_{\bar{y}}=0)$	1

b.

	Y：$\boxed{7}$	$\neg Y$：$\boxed{7}$以外	計
X：$\boxed{\text{A}}$	$\mathcal{A}_{xy}=0.5$	$(\mathcal{A}_{x\bar{y}}=0)$	$(\mathcal{A}_{x}=0)$
$\neg X$：$\boxed{\text{A}}$以外	$\mathcal{A}_{\bar{x}y}=0.5$	$((\mathcal{A}_{\bar{x}\bar{y}}=0))$	$(\mathcal{A}_{\bar{x}}=0.5)$
計	$\mathcal{A}_{y}=1$	$((\mathcal{A}_{\bar{y}}=0))$	1

この推論確信度は関連性が極めて低いため、被験者は「表のアルファベット」が仮に分かった場合の認知環境を設定し、推論の想定値が関連性のあるものに変化するかを検証する。$\boxed{7}$ の表側が $\boxed{\text{A}}$ であった時の認知環境は (28a), $\boxed{\text{A}}$ 以外であった時の認知環境は (28b) のようになる。

[6]式 (16) には h_x が存在するため、X の情報をフレーム外に出せるが、4.5 節で述べたように、日常的推論では顕示的情報 X をフレーム外に出すことは基本的にできない。

(28) a.

	Y：$\boxed{7}$	$\neg Y$：$\boxed{7}$以外	計
X：$\boxed{A}$	$(\mathcal{A}_{xy}=1)$	$(\mathcal{A}_{x\bar{y}}=0)$	$(\mathcal{A}_{x}=1)$
$\neg X$：$\boxed{A}$以外	$(\mathcal{A}_{\bar{x}y}=0)$	$((\mathcal{A}_{\bar{x}\bar{y}}=0))$	$((\mathcal{A}_{\bar{x}}=0))$
計	$(\mathcal{A}_{y}=1)$	$((\mathcal{A}_{\bar{y}}=0))$	1

b.

	Y：$\boxed{7}$	$\neg Y$：$\boxed{7}$以外	計
X：$\boxed{A}$	$(\mathcal{A}_{xy}=0)$	$(\mathcal{A}_{x\bar{y}}=0)$	$(\mathcal{A}_{x}=0)$
$\neg X$：$\boxed{A}$以外	$(\mathcal{A}_{\bar{x}y}=1)$	$((\mathcal{A}_{\bar{x}\bar{y}}=0))$	$(\mathcal{A}_{\bar{x}}=1)$
計	$(\mathcal{A}_{y}=1)$	$((\mathcal{A}_{\bar{y}}=0))$	1

(28) の数値より、やはり 式 (19) を用いて各認知環境における推論の確信度を計算する。$\neg X$ が完全に考慮されているので、$k_{\bar{x}}=1$ とした上で、推論想定値を求めると、(28a) では $\mathcal{A}_{x\to y}=1$ となり、(28b) では値が不定となる。(28a) の状況において、最高の関連性を得ることができるため、この $\boxed{7}$ のカードは選ばれやすい。ただし、$\boxed{A}$ がいずれの認知環境においても最高の関連性が得られていたのに対し、$\boxed{7}$ のカードは一方の認知環境からは最高の関連性が得られるが、もう一方の情報は認知環境の改善をもたらさない情報となってしまうため、$\boxed{7}$ のカードの選択率は、$\boxed{A}$ に比べて減少する (実際にその通りである)。

最後に本来なら正解であるはずの $\boxed{2}$ のカードについて見てみよう。このカードからは、D-領域の想定値 $\mathcal{A}_{\bar{x}y}=\mathcal{A}_{\bar{x}\bar{y}}=0.5$ が得られるため、これを (23) に組み込むと (29a) が得られ、さらに間接的に $\mathcal{A}_{\bar{x}y}=0$ を導出できるため、認知環境は (29b) のように設定される。

(29) a.

	Y：$\boxed{7}$	$\neg Y$：$\boxed{2}$	計
X：$\boxed{A}$	$(\mathcal{A}_{xy}=0)$	$\mathcal{A}_{x\bar{y}}=0.5$	$(\mathcal{A}_{x}=0.5)$
$\neg X$：$\boxed{A}$以外		$\mathcal{A}_{\bar{x}\bar{y}}=0.5$	$\mathcal{A}_{\bar{x}}=0.5$
計	$(\mathcal{A}_{y}=0)$	$\mathcal{A}_{\bar{y}}=1$	1

b.

	Y：$\boxed{7}$	$\neg Y$：$\boxed{2}$	計
X：$\boxed{A}$	$(\mathcal{A}_{xy}=0)$	$\mathcal{A}_{x\bar{y}}=0.5$	$(\mathcal{A}_{x}=0.5)$
$\neg X$：$\boxed{A}$以外	$((\mathcal{A}_{\bar{x}y}=0))$	$\mathcal{A}_{\bar{x}\bar{y}}=0.5$	$(\mathcal{A}_{\bar{x}}=0.5)$
計	$((\mathcal{A}_{y}=0))$	$\mathcal{A}_{\bar{y}}=1$	1

(29b) の推論確信度を計算すると、$\mathcal{A}_{x\to y}=0$ である。事実が確定していない段階において推論の確信度が 0 になっているので、これは認知効果の低い環境といえる ((3d-II) に関係する議論を参照のこと)。そこで、$\boxed{2}$ の表面が分かった時の認知環境を設定し、どの程度の関連性が得られるのかを計算する。表面が $\boxed{A}$ である時と $\boxed{A}$ 以外の時における認知環境は、各々 (30) のようになる。

(30) a.

	Y：$\boxed{7}$	$\neg Y$：$\boxed{2}$	計
X：$\boxed{A}$	$(\mathcal{A}_{xy}=0)$	$(\mathcal{A}_{x\bar{y}}=1)$	$(\mathcal{A}_{x}=1)$
$\neg X$：$\boxed{A}$以外	$((\mathcal{A}_{\bar{x}y}=0))$	$(\mathcal{A}_{\bar{x}\bar{y}}=0)$	$((\mathcal{A}_{\bar{x}}=0))$
計	$((\mathcal{A}_{y}=0))$	$(\mathcal{A}_{\bar{y}}=1)$	1

b.

	Y：$\boxed{7}$	$\neg Y$：$\boxed{2}$	計
X：$\boxed{A}$	$(\mathcal{A}_{xy}=0)$	$(\mathcal{A}_{x\bar{y}}=0)$	$(\mathcal{A}_{x}=0)$
$\neg X$：$\boxed{A}$以外	$((\mathcal{A}_{\bar{x}y}=0))$	$(\mathcal{A}_{\bar{x}\bar{y}}=1)$	$((\mathcal{A}_{\bar{x}}=1))$
計	$((\mathcal{A}_{y}=0))$	$(\mathcal{A}_{\bar{y}}=1)$	1

(30) の値を式 (19) に代入すると、いずれの場合も解は不定となり、推論確信度を決定することができない。この不安定な状態から逃れるため、(30a) に関しては、$\neg X$ に関する情報を認知環境から排除し、式 (19) の係数 $k_{\bar{x}}$ を 0 にすることで、$\mathcal{A}_{x \to y}=0$ を得ることができる。しかし、この推論確信度はかえって関連性の低い情報であることを示してしまう。したがって、本来なら正解である $\boxed{2}$ のカードは残念ながら選択されない。以上が Wason 課題における選択錯誤の心理的過程である。

和田 (2004) は、こうした Wason 選択課題をさらに拡張し、20 枚のカードを用いることで前件・後件の生起確率を変え、それがカード選択にどのような影響を受けるかを実験的に検証している。その結果、主に後件に関わるカードの生起確率がカード選択に影響し、低確率の場合ほどカードの選択率が高くなることが示された。この結果も式 (19) から理解できる。今、カードの生起確率が想定確信度とほぼイコールであるとすると、一枚のカードをめくることによってもたらされる想定確信度の改善 (生起確率の変動) は、もともとの想定確信度が小さいほど、すなわち生起確率が小さなものほど大きくなる。係数 $k_{\bar{x}}$ についても、枚数の少ない条件であるほど、カード一枚当たりにおける係数の設定値は大きなものにできる。したがって、認知環境を改善する関連性の高い情報は、想定確信度の低いもの、すなわち生起確率の小さなものに引きずられやすいことが分かる。

5.3 マッチング選択課題

Wason 選択課題と似て非なる問題として、マッチング課題といわれる以下のような実験課題がある。このマッチング課題では、Wason 課題に比べ、正解率が劇的に上昇することが知られている。しかし、確証バイアスやマッチングバイアスでは、正答率が劇的に上昇することが説明できない。

(31) ある工場では、次の規則に従って、上に文字、下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：もし文字が “A” なら、数字は “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。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> どのカードを調べる必要がありますか。

A 7	A 2	G 7	G 2

計算論的関連性理論を用いると、この実験の心理過程は次のように考えられる。この場合も、実験当初の想定は、$\mathcal{A}_{xy}=\mathcal{A}_{x\bar{y}}=\mathcal{A}_{\bar{x}y}=\mathcal{A}_{\bar{x}\bar{y}}=0.25$ となり、カード製造規則に関する想定の強さは $\mathcal{A}_{x\to y}=0.5-k_{\bar{x}}\cdot 0.5$ である。ここで、もし $\boxed{\text{A7}}$ のカードを「誤り」として除外したとすると、$\mathcal{A}_{xy}=0$, $\mathcal{A}_{x\bar{y}}=\mathcal{A}_{\bar{x}y}=\mathcal{A}_{\bar{x}\bar{y}}=0.33$ (前件否定は考慮されないため $k_{\bar{x}}=0$) より、$\mathcal{A}_{x\to y}=0$ となるため、規則も絶対に誤っているということが分かる。したがって、「除外してはいけない」最も関連性の高いカードと判断できる。次に、$\boxed{\text{A2}}$ のカードを誤りとして除外した場合には、$\mathcal{A}_{x\bar{y}}=0$, $\mathcal{A}_{xy}=\mathcal{A}_{\bar{x}y}=\mathcal{A}_{\bar{x}\bar{y}}=0.33$ (そして $k_{\bar{x}}=0$) より、$\mathcal{A}_{x\to y}=1$ となるため、「除外しなければならない」最も関連性の高いカードと判断される。一方、$\boxed{\text{G7}}$ のカードを誤りとして除外した場合には、$\mathcal{A}_{\bar{x}y}=0$, $\mathcal{A}_{xy}=\mathcal{A}_{x\bar{y}}=\mathcal{A}_{\bar{x}\bar{y}}=0.33$ より、$\mathcal{A}_{x\to y}=0.5$ であるため、当初の想定値と値がほとんど変わらず、関連性の高い情報とはいえないことが分かる。最後の $\boxed{\text{G2}}$ の場合も、$\mathcal{A}_{\bar{x}\bar{y}}=0$, $\mathcal{A}_{xy}=\mathcal{A}_{x\bar{y}}=\mathcal{A}_{\bar{x}y}=0.33$ より、推論想定の強さは $\mathcal{A}_{x\to y}=0.5-k_{\bar{x}}$ で、関連性の高い情報にはならない。この結果、論理的な含意計算の結果と一致する $\boxed{\text{A7}}$ と $\boxed{\text{A2}}$ のカードは必ず選択され、正解率の上昇がもたらされる。

5.4 モダリティと義務／許可スキーマ

Wason 課題に関する面白い現象の 1 つに、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点がある。例えば、次のような問題を考えてみよう。

(32) 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 (33) のような方略で解釈することを言い、許可スキーマとは (34) 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。

(33) 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。

(34) 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。

(35)

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

(35) 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。3.1 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 (19) により $\mathcal{A}_{x\to y}$ の値を求めてみると、$\frac{\mathcal{A}_{xy}}{\mathcal{A}_{xy}+\mathcal{A}_{x\bar{y}}}=1,\quad \frac{\mathcal{A}_{\bar{x}y}}{\mathcal{A}_{\bar{x}y}+\mathcal{A}_{\bar{x}\bar{y}}}\approx 0$ より、$\mathcal{A}_{x\to y}\approx 1$ となることがわかる。(13b) 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、〜ねばらなない」に相当する値である。これが、(33a) 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、(36) のような認知環境を持つ。

(36)

雇用者		平日 休み	平日 労働
休日	労働	$\mathcal{A}_{xy}$	$\mathcal{A}_{x\bar{y}}$
休日	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、(36) の認知環境から、式 (19) 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 (13d) 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 (34a) 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。

5.5 誘導推論

では、他の (33b)～(33d), (34b)～(34d) といったスキーマが出現する理由は何だろうか。坂原 (1985) は、グライス流の会話の原則の立場から、条件文「X ならば Y である」から、「X でないならば Y でない」「Y ならば X である」「Y でないならば X でない」といった誘導推論 (*invited inference*) が起こると述べている。このうち、X→Y と命題論理上等価なものは、対偶条件¬Y→¬X のみであり、¬X→¬Y (古くから前件否定の誤謬として知られてきた) と Y→X は等価な論理とは言えない。しかし、限られた情報量の発話から、最大限の情報を得るのに、この誘導推論は極めて有益な心的操作になるというわけだ。

ここで、想定 $X\to Y$, $\neg X\to\neg Y$, $Y\to X$, $\neg Y\to\neg X$ の想定確信度を求める式を見てみよう。式の構造は (19) と同一である。

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

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

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

この (37) に労働者の認知環境 (35) の想定値を代入すると、$\mathcal{A}_{x\to y}\approx 1$, $\mathcal{A}_{\bar{x}\to\bar{y}}>0$, $\mathcal{A}_{y\to x}>0$, $\mathcal{A}_{\bar{y}\to\bar{x}}\approx 1$ となることが分かる。これは、義務スキーマ (33) のモダリティそのものである。許可スキーマも同様である。雇用者の認知環境 (36) を (37) に代入すると、$\mathcal{A}_{x\to y}>0$, $\mathcal{A}_{\bar{x}\to\bar{y}}\approx 1$, $\mathcal{A}_{y\to x}\approx 1$, $\mathcal{A}_{\bar{y}\to\bar{x}}>0$ となり、許可スキーマ (34) のモダリティに一致する。このように、Wason 課題における各種バイアスやスキーマは、認知環境の性質と推論の想定確信度 (19) の相互作用によってもたらされる性質なのである。

6. 反事実条件文再考

6.1 反事実条件文の性質

可能性と必然性に関わる論理に関係する、もう1つの変わった論理が反事実条件文である。この条件文は、実質条件法と厳格条件法の中間的な強さを持ち、その論理はいずれの論理とも異なっている (Lewis, 1973)。反事実条件文の興味深い性質は、前件か後件

かのいずれかが成立していないことが明確な文であるという点にある。例えば、日本語の反事実条件文でも、(38a) のような文脈から前件が成立しないことが分かるもの、(38b) のように常識などの知識レベルから前件が成立しないことが明確なもの、(38c) のように後件が成立しないことが明確なものが存在し、(39) のように解釈される。

(38) a. 結局、奈緒美は来なかった。彼女が来ていれば、チョコをもらえたのに。
b. 私が鳥だったら、今すぐ奈緒美に会いにいけるのに。
c. あの薬を飲んでいたら、今ごろ死んでいたところだ。

(39) a. 奈緒美は来なかったし、チョコももらえなかった。
b. (鳥ではないし) 今は奈緒美に会いに行けない。
c. 薬は飲まなかったし、現在も生きている。

6.2 反事実条件文における論理上の問題点

反事実条件文の解釈が成立するのに、表出された前件あるいは後件のいずれかを偽であると確信できることが必要であるという性質は、形式論理上の問題点を引き起こす。

まず、後件が偽である (38c) の場合を考えてみよう。真理表 (表 20) からも分かる通り、後件が偽である場合には、前件が真である限り、含意演算にせよ同値演算にせよ、文全体が偽になり、そのままの形で受理することはできない。そこで再解釈として、前件も偽であるとすると、文全体が真となり、認知主体はこの条件文を反事実的解釈として受理可能である。しかし、どのみち再解釈を行うのであれば、なぜ後件を偽のままで置いておく必要があるのだろう。後件に関する信念を変更し、後件を真であると見なせば、前件も真・後件も真、文全体も真という極めて安定した解釈を得ることができるはずである。

もちろん、これに関する回答として、再解釈のコストを挙げることができる。後件が偽と確信している認知主体にとって、その確信を変更するよりも、真か偽か明確ではない前件に関する信念を変更するほうが、より再解釈のコストが低くて済むという説明である。確かにこれは妥当な考え方であるが、それでもいくつかの問題が残る。まず根本的に、反事実条件文の意味理解とはどういうことかという問題がある。確かに、「前件・後件ともに偽」という解釈は現実世界に対応させることが可能なので、安定した理解といえる。では、言語で表現された内容そのものは、どう扱われるべきなのか。そもそも反事実条件文全体を「真」として受理しようとしているのだろうか。近年の比喩理解の心理学的研究で、比喩の理解は再解釈を必要とするような時間のかかるものではなく、直接理解に近い極めて高速な処理であることが明らかにされてきている。反事実条件文に関しても、同様のことがいえる。と同時に、反事実条件文の理解は、明らかにパラフレーズの解釈を引き起こし、その条件文自体を「真」であるとして受理しようとしているのか、疑問が残る。

前件が偽である (38a), (38b) の場合、さらに別の問題が起こる。真理表 (表 20) より、前件が偽の場合、少なくとも含意演算においては、後件が真であっても、文全体を真と

して受理可能である。すなわち、反事実条件文として理解しなくてもよい。同値演算では、後件が偽でない限り、文全体を真として受理できないため、反事実条件文としての再解釈が強制されることがいえるが、ではなぜ、前件が偽である場合、同値演算 (すなわち双条件解釈) を行わなければならないのかという点の説明が必要である。

次節では、計算論的関連性理論における推論演算 (19) を用いて、なぜ反事実条件文の解釈が生起し得るのかを論じる。議論のポイントとして、反事実条件文の解釈において、含意・同値という論理の区別を持ち出さなくても自然に解釈が生じうることを主張する。

6.3 反事実条件文における確信度変更の不可能性

田窪 (1993) は、状態形の条件文が反事実解釈に結びつきやすいことを指摘し、これは状態形が既に決定している現実を指し、それにも関らず、強制的に仮定が行われるために、反事実性が生じるのだと述べている。この「既に決定している」という点は、想定の確信度が 0 か 1 に限りなく近くなっていると捉えなおすことができる。同様に、郡司 (2004) は、(40) の違いを説明するメカニズムとして、アスペクトの縮退という性質を提案し、事態がすでに完了したものという前提の元で出来事を全体として捉える見方が可能であれば、反実仮想解釈が成立すると述べている。

(40) a. あの時、あと 30 分待ってい {たら／れば}、彼に会えただろう。(反実仮想)
  b. 30 分待ってい {たら／れば}、彼に会えた。(進行形)

この「完了したものとして全体として捉える見方」を、想定確信度という点から捉えなおすと、その想定に対応する事態が既に決定済みであり、想定の確信度が確定している (確信度が 1 か 0 に固定されている) 状態と考えられる。一方、反実仮想解釈が進行形解釈の時に生じない理由は、事態が安定していないため、それに対応する想定の確信度が中間的な値を取っているためといってよい。次節以降では、反事実条件文の理解過程を見ながら、この想定確信度の変更不可能性が及ぼす影響について考察する。

6.4 前件の偽が明確である時の直接理解過程

この節では、「健が鳥だったら、今すぐ奈緒美に会えるのに」という反事実条件文の理解過程を見てみよう。なお、「健は鳥である」という想定を X、「健は奈緒美に会える」という想定を Y としておく。

この発話を受理するまで、話者も聴き手も「健が人間である」ことを直接経験として知っている。したがって、D-領域の情報として、

(41)

	計
$\neg X$: *human*(*Ken*)	$\mathcal{A}_{\bar{X}}=1$

という認知環境を持っている。ここで「健が鳥だったら、今すぐ奈緒美に会えるのに」という条件文を受け取った場合、Wason 課題の場合と同様、まず I-領域の想定として、X, Y を認知環境に組み込まなければならないことを知る。さらに、この文が条件文である

ことから、$X \to Y$ という想定を設定し、その確信度を式 (19) から計算しようとする。この時、式の計算を遂行するために、$\neg Y$ の情報も必要であることが間接的にわかる (係数 $k_{\bar{x}}$ に関わらず、$\mathcal{A}_{x\bar{y}}$ の値が決まらない限り、推論想定度を計算できないからである)。しかし、I-領域の想定は、D-領域の想定確信度を変更する力がないため、(42) のような認知環境が構成される。

(42)

	Y : *meet*(*Ken*, *Naomi*)	$\neg Y$	計
X : *bird*(*Ken*)	($\mathcal{A}_{xy}$=0)	(($\mathcal{A}_{x\bar{y}}$=0))	(($\mathcal{A}_x$=0))
$\neg X$: *human*(*Ken*)	$\mathcal{A}_{\bar{x}y}$= 不明	$\mathcal{A}_{\bar{x}\bar{y}}$= 不明	$\mathcal{A}_{\bar{x}}$=1
計	$\mathcal{A}_y$= 不明	$\mathcal{A}_{\bar{y}}$= 不明	1

この時、(19) の前半部分は $0 \leqq \dfrac{\mathcal{A}_{xy}}{\mathcal{A}_{xy}+\mathcal{A}_{x\bar{y}}} \leqq 1$ となり、不定の値である。ここでいかなる場合においても、$\mathcal{A}_{x \to y}$ を想定として受理するためには、常に $\mathcal{A}_{x \to y} \geqq 0$ が成立する必要があるため、$\dfrac{\mathcal{A}_{\bar{x}y}}{\mathcal{A}_{\bar{x}y}+\mathcal{A}_{\bar{x}\bar{y}}}$=0 でなければならない。したがって、$\mathcal{A}_{\bar{x}y}$=0 が成立し、$\mathcal{A}_{xy}$=0, $\mathcal{A}_{x\bar{y}}$=0, $\mathcal{A}_{\bar{x}y}$=0 より $\mathcal{A}_{\bar{x}\bar{y}}$=1 が決まることで、(43) に示す認知環境を得る。すなわち、「健は鳥でない」かつ「健は奈緒美に会えない」ことが理解され、これ以外の解釈は起こりえないことが分かる。さらに、認知環境としてみた場合、各想定が 0 か 1 になっているため、最も関連性の高い認知環境であるといってよい。

(43)

	Y : *meet*(*Ken*, *Naomi*)	$\neg Y$	計
X : *bird*(*Ken*)	($\mathcal{A}_{xy}$=0)	(($\mathcal{A}_{x\bar{y}}$=0))	(($\mathcal{A}_x$=0))
$\neg X$: *human*(*Ken*)	($\mathcal{A}_{\bar{x}y}$=0)	($\mathcal{A}_{\bar{x}\bar{y}}$=1)	($\mathcal{A}_{\bar{x}}$=1)
計	($\mathcal{A}_y$=0)	($\mathcal{A}_{\bar{y}}$=1)	1

なお、この時の推論想定値は $0 \leqq \mathcal{A}_{x \to y} \leqq 1$ の範囲で任意の値となり、何らかの形で関連性のある情報として受理されるが、極めて不安定な状態である。つまり、この条件文そのものの真理値については、真であるとも偽であるとも決定できない。しかし、既に述べたように、反事実条件文を理解する上で、再解釈を行うことなく、正しい意味の直接理解が可能である点は重要な性質と思われる。

6.5 前件の偽が明確である時の誘導推論

しかし、推論の想定確信度が決まらないというのは、関連性の計算にとってよくないことである。そこで、認知主体によっては、誘導推論を開始する。

前件が偽である場合、(37b) の $\mathcal{A}_{y \to x}$ は常に 0 となり、推論想定度自体は決定可能であるが、その値が低いため、関連性のある解釈とはいえない。一方、(37c) では、係数 k_x を 0 に設定しさえすれば、$\mathcal{A}_{\bar{x} \to \bar{y}}$=1 となり、最大の関連性を持つ情報となる。$X$ はもともと誤った想定であることが分かっているため、k_x=0 によって認知環境から排除しても問題は起こらない。したがって、「健は鳥でないので、奈緒美に会えない」という誘導推論が可能になる。(37d) も (37c) と同様、係数 $k_{\bar{y}}$ を 0 にすることで推論の想定確信度を 1 に確定できる。

6.6　後件の偽が明確である時の理解過程

次に、「もしあの薬を飲んでいたら、奈緒美は死んでいたところだ」という後件の偽が明らかな反事実条件文の理解過程を見てみよう。「奈緒美が薬を飲んだ」という想定を X とし、「奈緒美はもう死んだ」という想定を Y とすると、そもそも話者や聴き手は D-領域の想定として、条件文を受理する以前から $\neg Y$ の想定値 $\mathcal{A}_{\bar{y}}=1$ を認知環境に組み込んでいる。この状態で、$X \rightarrow Y$ の条件文を受理した場合、推論の想定値を計算するため、最低限、$\mathcal{A}_{xy}$ と $\mathcal{A}_{x\bar{y}}$ の値を決める必要が出てくる。しかし、いずれも I-領域の想定であり、想定確信度 $\mathcal{A}_{\bar{y}}=1$ を覆す力はない。したがって、I-領域の情報として $\mathcal{A}_{xy}=0$, $\mathcal{A}_{x\bar{y}}=0$ が (44a) のように認知環境に組み込まれる。そうすると、不明になっている想定値は (44b) のように自動的に決定され、$\mathcal{A}_{\bar{y}}=1$ のみが受理される情報であることが分かる。すなわち、「奈緒美は薬を飲んでおらず、かつ死んでもいない」という解釈のみが成立する。この場合も $\mathcal{A}_{x \rightarrow y}$ の値は不定になるが、やはり再解釈なしに目的の解が得られるのだ。

(44)　a.

	Y : *die*(*Naomi*)	$\neg Y$: *alive*(*Naomi*)	計
X : *take*(*Naomi*, *pill*)	($\mathcal{A}_{xy}$=0)	($\mathcal{A}_{x\bar{y}}$=0)	($\mathcal{A}_{x}$=0)
$\neg X$	$\mathcal{A}_{\bar{x}y}$= 不明	$\mathcal{A}_{\bar{x}\bar{y}}$= 不明	($\mathcal{A}_{\bar{x}}$=1)
計	$\mathcal{A}_{y}$=0	$\mathcal{A}_{\bar{y}}$=1	1

b.

	Y : *die*(*Naomi*)	$\neg Y$: *alive*(*Naomi*)	計
X : *take*(*Naomi*, *pill*)	($\mathcal{A}_{xy}$=0)	($\mathcal{A}_{x\bar{y}}$=0)	($\mathcal{A}_{x}$=0)
$\neg X$	((($\mathcal{A}_{\bar{x}y}$=0)))	(($\mathcal{A}_{\bar{x}\bar{y}}$=1))	((($\mathcal{A}_{\bar{x}}$=1)))
計	$\mathcal{A}_{y}$=0	$\mathcal{A}_{\bar{y}}$=1	1

しかし、前件が偽の場合の反事実条件文同様、推論の想定値が不安定であるのは関連性の効果として望ましくないため、認知主体によってはやはり誘導推論を開始する。この過程は、前件が偽の場合と全く同様で、フレームの係数を 0 に置きさえすれば、(37c), (37d) の誘導推論が可能となり、「薬を飲んでいないので、死んでいない」「死んでいないのは薬を飲まなかったからだ」といった解釈が成立する。

6.7　反事実条件文の効果

前述したように、反事実条件文は命題論理や真理条件として考えると、奇妙な性質を持っている。それにも関わらず、この表現が用いられる理由は何か。関連性理論では、これも認知環境の構成に原因があるのだと説明できる。現実をただ単純に述べただけでは、(45a) のような認知環境しか構成されず、ここに別の想定が組み込まれた時には、想定確信度が変化してしまう可能性がある ((4.4 の議論を参照されたい)。しかし、反事実条件文では (45b) のように、1つの表現で豊かな認知環境を生み出すことができ、想定の変更は容易でなくなる。つまり最も安定した解釈が行えるわけだ。これが反事実条件文の使われる理由ではないかと考えられる。さらに、前項まで見てきたように、計算論的関連性理論に基づいて反事実条件文を理解した場合、必ずしも誘導推論のような再解釈は

必要とされず、直接理解を行うことが可能である。この意味でも、反事実条件文は効率の良い表現といってよい。

(45) (a)

	$\mathcal{A}_{y_1}$	$\mathcal{A}_{y_2}$	計
$\mathcal{A}_{x_1}$	0	1	1
計	0	1	1

(b)

	$\mathcal{A}_{y_1}$	$\mathcal{A}_{y_2}$	計
$\mathcal{A}_{x_1}$	0	0	0
$\mathcal{A}_{x_2}$	0	1	1
計	0	1	1

これは、「外に出たら綺麗な景色が見えるよ」といった連言を使った条件表現にも同じことがいえる。もともと、「外に出ていない」という事実が存在し、その想定が D-領域にあるからこそ、この条件文のもたらす「想定 $X \wedge Y$ の確信度 $\mathcal{A}_{xy}$ を 0 から変更せよ」という条件表現と同等の価値を持つようになるのだ。反事実条件文を含め、ある種の条件表現は、直接的な表現よりも認知環境を豊かにしてくれるという点で、効率がよく、高い関連性をもたらしてくれる表現なのである。

7. 総合論議

本稿では、関連性理論の基本的な定義を計算論的に定義する手法を述べ、その応用として各種の条件文がどのように理解されるかを見てきた。議論のポイントは以下の通りである。

(46) a. 想定に量的な想定確信度を与えることで、関連性理論を計算論的に形式化できる。

b. 関連性・エントロピー・真理値・モダリティなどの諸概念は想定確信度から定義・計算可能である。

c. 認知環境は基本的に単独想定か、連言により結びつけられた複合想定によって規定される (すなわち情報の共起関係が基本となる)。

d. 各種の論理演算子は、認知環境を変化させる指令としても機能する。

e. 推論の想定計算は、関連性理論の中でも最も基本的で重要な演算である。

f. 反事実条件文のような表現は、認知環境の設定という点から見ると、極めて妥当性の高い効果的な言明である。

g. 反事実条件文の解釈には、再解釈は必ずしも必要でない。これは即時理解という意味でも重要である。

本稿で述べた関連性理論の形式化は、この理論の持つ特徴の極一部のみを計算論的に表現したに過ぎない。今後は、さらに理論の精緻化を行い、他の条件表現や帰納推論、一般推論なども含めた思考過程の研究を行う予定である。

参考文献

Cheng, Patricia W. & Holyak, Keith J. (1985). Progmatic reasoning schemas. ***Cognitive Psycholoy***, **17**, 391–416.

Evans, Jonathan St. B. T. & Lynch, J. S. (1973). Matching bias in the selection task. ***British Journal of Psychology***, **64**, 391–397.

郡司隆男 (2004). 日本語のアスペクトと反実仮想. *Theoretical and Applied Linguistics at Kobe Shoin*, **7**, 21–34.

井上雅勝 (2000). 『ガーデンパス現象に基づく日本語文理解過程の 実証的研究』. Ph. D. dissertation, 大阪大学.

Inoue, Masakatsu & Den, Yasuharu (1999). Influence of verb-predicability on ambiguity resolution in Japanese. *Proceedings of the 2nd International Conference on Cognitive Science*, **1**, 499–502.

Lewis, David (1973). *Counterfactuals*. Havard University Press: Cambridge, Mass.

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.

坂原茂 (1985). 『日常言語の推論』. 東京大学出版会.

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

Sperber, Dan & Wilson, Deirdre (1986). *Relevance: Communication and Cognition*. Blackwell. 内田聖二ほか 訳 (1993). 『関連性理論—伝達と認知—』. 研究社出版.

田窪行則 (1993). 談話管理理論から見た日本語の反事実条件文. 益岡隆志（編），『日本語の条件表現』, pp. 169–183. くろしお出版, 東京.

田窪行則・金水敏 (1996). 複数の心的領域による談話管理. 『認知科学』, **3** (3), 59–74.

和田一成 (2004). 20 枚版 Wason 選択課題における選択肢の提示確率の効果. 『日本心理学会発表論文集』, **68**, 887.

Wason, Peter. C. (1966). Reasoning. In Foss, B. M. (Ed.), *New Horizons in Psychology*. Harmondsworth: Penguin.

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