想定の確信度と真理値*

松井 理直

Degree of Belief and Truth Value

Michinao F. MATSUI

Abstract

We live in environment filled with various information, and 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, to acquire the truth knowledge, and to ignore inappropriate information which is irrelevant. However, it is difficult to define *Truth*, *Knowledge*, and *Correctness* for human being. This paper proposes that human knowledge is justified true belief with *epistemic necessity* and that truth is defined by updating the cognitive environment from the viewpoint of Computational Relevance Theory.

1. はじめに

私たちは様々な情報があふれている環境の中で生きている。それらの情報のうち、いくつかの情報は真であり、いくつかの情報は偽である。こうした情報のうち、その正しさが明確な形で確認されているものは「真理」と呼ばれる。また、私たちはたくさんの考えや思いや想定を持って生きている。こうした想定のうち、ある基準を満たしたものを知識と言う。知識は正しい思考を行う上で最も重要な基盤となる。

全知全能の存在があるならば、その存在にとって「真理」は自明のものである。また、全ての真理を「知識」として持つことができる。しかし、限界のある人間は、いかにし

*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)「計算論的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 11, 25–66, 2008.

である情報が真であることを知るのだろう。どういう信念なら知識といえるものになるのだろう。

本稿は、計算論的関連性理論の観点から、真理値および知識と呼びうる信念の性質について議論したものである。結論として、人間にとって「真」とは必ずしも論理学の計算とは合致しないこと、「知識」の定義において、**認識的必然性**という概念が重要であることを述べる。

2. 計算論的関連性理論

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と思われる。本節では、まず関連性理論を代数的に扱う手法について考察してみよう。

2.1 記憶

人間が認知活動を行う際に利用できる記憶の情報には大きくわけて 2 つの種類がある。一つは固定化され、変更の難しい記憶情報であり、もう一つは比較的自由に変更でき、様々な加工や操作が可能な記憶情報である。心理学の記憶研究と関連づけていえば、前者のタイプは長期記憶 (LTM: long term memory) のデータであり、後者のタイプは作動記憶 (WM: working memory) の情報といってもよい。

長期記憶 (LTM) のデータは減衰や干渉といった何らかの理由で忘却されない限り、ずっと保持される。長期記憶の内容は陳述記憶・非陳述記憶の 2 つに分類される。 陳述記憶とは、簡単にいえば言葉で表現できる記憶であり、宣言的に表現される記憶情報である。陳述記憶はエピソード記憶・意味記憶の 2 つに分類される。非陳述記憶とは、言葉で表現できない非宣言的な記憶であり、手続き記憶とプライミング記憶の 2 つに分類できる。一方、作動記憶 の情報は短期間保持される短期記憶 (STM: short term memory) と共に、中央制御系、音韻ループ、視空間スケッチパッドといった認知的な情報処理を行うための記憶情報を含む。作動記憶の情報は、精緻なリハーサルや検証を経て、長期記憶に送り込まれ、データとして貯蔵されることになる。

こうした概念は言語においても重要なものであり、これらの性質を最も明示的・直接的に取り込んだ理論として、田窪· 金水 (1996b) による**談話管理理論** (discourse management theory) を挙げることができる。この理論では、以下のような性質を持つ D-領域、I-領域という談話領域 (discourse domain) を設定し、各領域における情報処理の指令が言語表現に直接反映されていると考える。

(1) a. **D-領域**

長期記憶内にある情報で、その内容や真偽が既に検証されており、直接経験

情報やエピソード情報、対話の現場情報とリンクされている要素が格納されている。

b. **I-領域**

推論や伝聞など間接的に得られていたり、仮説や仮定など仮に設定されている、まだ検証されていない情報が格納されている。

c. **情報移転制約**

I-領域内の要素は、その要素が設定された談話セッションの間はD-領域に移すことはできない。

こうした概念を用いると、たとえば終助詞の「よ」と「ね」の使い分けも、前者が「I-領域に新規情報を付加し、そこから適切な推論を行わせるためのマーカー」であるのに対し、後者は「情報の内容を計算中であるというマーカー」として分析できる。また、「こそあど言葉」でも、ア系列は「話し手がD-領域の要素を直接指示しているマーカー」として、ソ系列は「話してがI-領域の要素を指示するマーカー」として簡潔に定義できる(金水・田窪, 1990)。

談話管理理論ほど明示的ではないにせよ、安定した情報と一時的に変更・操作可能な情報という区別は、Fauconnier (1994) のメンタル・スペース理論における各種のスペース設定や、Sperber and Wilson (1986) による**関連性理論**における「認知環境」という概念においても重要な役割を担っている。

以下の節では、ベースとなる安定した情報と、一時的な操作可能な情報という区別を、想定の確信度という数的指標によって表現し、それに基づいていかに情報の更新に関する計算を行うか、また論理 (特に真理値) とこうした想定の確信度とがどのように関係するのかという点について議論を行う。理論の枠組みには、関連性理論を採用するが、基本的なアイデアは談話管理理論やメンタル・スペース理論といった他の言語理論にも応用できるものであろう。

2.2 データと想定

認知主体の感覚によって取り込まれたあらゆる外界符号物や、長期記憶に貯蔵されているあらゆる記憶は、認知主体によって常に利用可能になっているとは限らない。こういう種類の情報を特にデータと呼ぼう。これに対して、認知主体が長期記憶にアクセスしたり、外界に注意を払う (意識的注意か無意識かは問わない) ことによって、一部のデータは認知主体にとって利用可能な「**顕在的** (*manifest*) データ」となる。こうした顕在的情報のことを特に**想定** (*assumption*) と呼ぶことにしよう。長期記憶 (あるいはD-領域) の想定内容は固定化されており、作動記憶 (あるいはI-領域) に引き出して、相当な認知的コストをかけない限り変更することはできない。一方、外界から取り込んだばかりの想定 (発話や知覚情報を) や認知主体の内部に生じた仮説 (真偽の検証が行われていない想定) は、まず作動記憶 (I-領域) に取り込まれ、その操作や変更が容易である。作動記憶の情報はその**妥当性について検証**が行われ、相応のリハーサルが行われた後、長期記憶に

貯蔵される。本稿の後半部で、どういうプロセスが妥当な検証といえるのかという点について議論を行う。

2.3 認知環境・認知効果・関連性

データは様々な形で構造化されている。たとえば、認知主体が知っている個々のトークンとしてのコップはデータである。個々のコップに関係づけられた上位カテゴリーであるタイプとしてのコップもデータであり、グラスや皿などと一緒に関係づけられた食器といった上位カテゴリーもデータである。また、「2008 年 1 月 15 日に健が数学の勉強をした」「2008 年 1 月 27 日に奈緒美のガンが発見された」といった個々の体験もデータであり、外界に今現在実際に起こっている種々の事象や、聞いたばかりの発話、得られたばかりの情報などもデータである。こうしたデータが、何らかのきっかけにより認知主体にとって顕在的な情報になり、利用や操作が可能になったものが想定である。顕在的情報である想定の総体を **認知環境** (*cognitive environments*) と言う。認知主体は常に新しい情報を獲得し、また推論によって自主的に新しい想定を生み出す存在であるため、認知環境は常に変化し、更新されている。こうした認知環境の変化の中で、特に認知環境の「改善」をもたらす作用を **認知効果** (*cognitive effects*) という。認知効果は

(2) a. 新しい想定の獲得
b. 不確実な想定の確定、
c. 誤った想定の棄却

のいずれかによってもたらされる。この認知効果という点から、関連性理論の中心概念が以下のように定義される。

(3) **関連性：**
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。

認知主体にとって、この関連性という性質はとても重要なものである。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。しかし、認知主体の持つ知覚・思考・伝達といった情報処理能力は、世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。そうした限界の中で、認知主体は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う。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であり、無限の情報の中で、必要と思われる情報を取捨選択する能力だといってもよい。

効率的に部分情報を処理するためには、当該の情報にとって無関係なものは無視し、関係のあるもののみを選択すればよい。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。[1]

[1] 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*, 情報的意図と伝達的意図) も重要な鍵となるが、本論文では関連性の伝達原理は取り扱わない。

(4) **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。

すなわち、認知主体は関連性の計算ができるが故に、部分情報のみを用いて、世界に一貫性を持たせ、体制化されたものにすることができると考えられるのである。もちろん、「フレーム問題」に代表されるように、関連性の計算が常に成功するとは限らないし、そもそも万能ではない認知主体にとって、「完全性」は最初から放棄せざるを得ない。しかし、妥当な関連性が計算できれば、多くの場合、部分情報だけでも適切な処理を行えるという点が大切なのである。

2.4 想定の確信度

関連性は (意識上か無意識下は別にして) 思考と同じく一種の推論過程である。関連性に基づく推論は、4.3 節で議論する通り、演繹推論 (deduction) と共に、帰納推論 (reduction) と仮説推論 (abduction) を含む。したがって、常に正しい推論であるとは限らない。また、想定そのものも、外界から獲得したばかりの想定や、思考・推論から得られた想定は常に確実な情報であるとは限らない。したがって、各想定がどの程度信頼できるものであるかという指標が必要である。この指標を**想定の確信度**と呼ぼう。この想定確信度は、(2), (3) に対応する形で、以下のような機能を持つ。

(5) 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。

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。$\mathcal{A}(x) < 0$ となる心的情報は関連性を持つ想定と見なされず、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 (あるいは 0) に近づけること、(iii) 想定確信度が 0 になったものは認知環境から棄却し、認知環境をコンパクトに最適化すること、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、効率よく (5b) を引き起こす情報である。すなわち、(i) の新規情報に関しては認知的関連性の確信度がより 1 に近い情報であること、(ii), (iii) の既存情報に関しては確信度の変更がより容易な情報が、関連性の高い情報である。

2.5 知らないということ

「知識」とは何かという問題は古くから哲学で問われてきた疑問であり、戸田山 (2002) で詳しく議論されている通り、未だに解決しているとは言い難い。しかし、何が真なる

知識であるかを明確にできない限り、長期記憶やD-領域に貯蔵すべき情報を選別することができない。本節では、「知識」や「真：という問題を扱う前に、「知らない」とはどういうものであるかについて考察する。一番わかりやすい例として、想定の確信度が0.5という場合を考えてみよう。

(6) a. 今、少々いかがわしいカジノで、「コインの表が出るか裏が出るか」という賭に参加しているとしよう。このカジノでは、時々不正なコインを使って賭をしていることで有名なので、慎重なKさんは賭を始める前にじっくりと600回の賭を観察したところ、ちょうど表も裏もちょうど300回ずつ起こっていた。そこでKさんは「このコインはどうやら正しいコインらしい。だから次に表が出るか裏が出るかは半々の確率だな」と考えた。

b. 同じカジノで、おっちょこちょいのHさんは、いきなり賭に参加することにした。Hさんは、「このコインは不正なものかもしれないし、正しいものかもしれない。正しいコインなら表も裏も同じだろうが、不正なコインなら表が多くでるかもしれないし、裏が多くでるかもしれない。要するに今の段階では何も分からなくて、表が出るかもしれないし裏が出るかもしれないので、どっちに賭けても同じだろう」と考えた。

この例では、Kさんの「表がでる想定確信度」もHさんの「表がでる想定確信度」も共に0.5となる。しかし、その実態は全く異なる。Kさんの場合は、正しいコインであることを「知って」おり、だから「表も裏も同じように出るだろう」と正しく推論した結果の想定確信度である。表も裏も同じように出ることを知っており、**実際**に表が出るのか裏が出るのか「分からない」ことも理解できているので、想定確信度が0.5と設定される。一方、Hさんの場合は何も「知らない」状態であり、0.5という想定確信度は無知であるが故の数値に過ぎない。

想定確信度という観点から言うと、**今この瞬間**の想定確信度からは、「知っているが故に分からない」という想定と、「無知」であるが故の想定とを区別することはできない。[2] しかし、幸運なことに認知主体は時間の流れの中に存在しており、次々に新しい想定を獲得していくことができる。新たな想定を獲得できた時点で、KさんとHさんの想定は大きく変わっていく。上記の状況に続いて、KさんもHさんも、7回のコイン投げを行い、6回表が出て、1回裏が出たとしよう。Kさんにとっては、「このコインは306回表が出て、301回裏が出たので、正しいコインという想定を変更する必要はないな」と考える。Hさんは、「6回も表が出ていて、裏は1回しか出ていない。これは不正なコインの可能性もあるな」と考える。これが「知っている」ことの想定確信度と「知らない」ことの想定確信度の本質的な違いである。つまり、認知環境が常に更新されていることが、「知っているか否か」ということに関してきわめて重要なことなのである。ただし、認知環境が更新されたとしても、「知」としての価値が増えることは保証されていない点に注

[2] 自分自身の心的状態を観察するメタなプロセスを仮定すれば区別することもできるが、本稿では無限後退の可能性を避けるため、こうしたアプローチは取らない。

意されたい。「知っている」ということが保証されるためには、認知環境の更新以外にも様々な手続きが必要である (戸田山, 2002; ノージック, 1997)。しかし、「無知[3]である」ことから徐々に逃れることは、知識獲得においてきわめて重要である。そのために、認知環境の更新が最低限必要なのである。[4]

さらに、もう一つ、「知」と「無知」の違いが想定確信度に与える重大な影響について見ておこう。例として2枚のコインを投げる事象を想像していただきたい。この時、「2枚とも表が出る」という想定の確信度はどのように計算されるだろうか。「無知」ではないKさんは、過去の事象にアクセスできるため、「一つのコインが表、もう一つも表」「一つのコインが表、もう一つのコインが裏」「一つのコインが裏、もう一つのコインが表」「一つのコインが裏、もう一つも裏」という情報にアクセスできる。したがって、「2枚とも表が出る」という想定の確信度は0.25となる。これはいわゆる確率論でいう頻度確率に近い。一方、「無知」であるHさんは「両方が表」「片方が表、もう片方が裏」「両方とも裏」という想定から、「2枚とも表が出る」という想定の確信度を1/3と設定する。これは「構造確率」と呼ばれるものである。ここで、認知主体は十分に証拠がある時には「頻度確率」に近い形で想定確信度を計算し、「無知」である時には「構造確率」によって想定確信度を設定すると仮定しておく。

この仮定は、後述するようにいくつかの心理学的妥当性があると共に、フレーム問題を擬似解決するための一つの方法でもある。単独想定といえども、実際には暗黙の前提(フレーム)や文脈(コンテクスト)の影響を受けており、正確に言うなら当該情報と文脈情報との複合想定である。「想定」は顕在的な情報であるので、暗黙前提や当該文脈以外の状況は計算の必要がない。しかし、逆にいうと、いかに適切な前提を設定するかが問題となる。構造確率の導入はこういう場合に有効である。構造確率の計算はきわめて高速に行えるため、まず構造確率に基づいて不正確ではあるがとりあえず妥当と思われるフレームを設定する。その後、必要があれば、頻度確率に相当する想定確信度および後述する関連性計算の数値に基づいて、より適切にフレームを調整を行えばよい。人間は「正しさ」を最初から求めることはできない。しかし、時間をかけ、認知環境の更新を行うことにより、正しい解に近づいていくことができるというのが本稿の1つの主張である。

2.6 相互作用のある想定を持つ認知環境

認知主体が何かものを考えたり、推論したり、あるいは対話を理解したりする場合、ある一つの情報だけで全てを判断するということはほとんどない。多くの場合、他の情報とどのように関係しているかを考慮した上で、何らかの判断を下すことになる。つまり、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、他の想定と相互作用を起こす複合想定であることが多い。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のような形でとりあえず表現できる。$\mathcal{A}(xy)$ は「情報 X と情報

[3]言うまでもなく無知と未知は異なる。上の例では、Hさんは無知だが、Kさんはそうではない。しかし、608回目の以降の事象はKさんにとってもHさんにとっても未知である。

[4]この議論は、長期記憶あるいはD-領域の情報に関する想定確信度を変更しにくいことの理由でもある。

Y が両立する」という想定に関する確信度、同様に、$\mathcal{A}(x\overline{y})$ は「情報 X と情報 $\neg Y$ の両立」、$\mathcal{A}(\overline{x}y)$ は「情報 $\neg X$ と情報 Y の両立」、$\mathcal{A}(\overline{xy})$ は「情報 $\neg X$ と情報 $\neg Y$ の両立」に関する想定確信度を示し、また C は最大のコンテクストを表す。なお、認知環境における想定は、3.4 節で見るように、「X なら Y」「X は Y」といった条件文のような形で表現されていることがほとんどだと思われるが、ここでは便宜上、情報の両立という形でとりあえず考えておく。

表 1: 認知環境における想定間の可能性

C		情報 Y		
		Y	$\neg Y$	合計
情報X	X	$\mathcal{A}(xy)$	$\mathcal{A}(x\overline{y})$	$\mathcal{A}(x)$
	$\neg X$	$\mathcal{A}(\overline{x}y)$	$\mathcal{A}(\overline{xy})$	$\mathcal{A}(\overline{x})$
	合計	$\mathcal{A}(y)$	$\mathcal{A}(\overline{y})$	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、$\mathcal{A}(xy) + \mathcal{A}(x\overline{y}) + \mathcal{A}(\overline{x}y) + \mathcal{A}(\overline{xy}) = 1$ が成り立つ。また、$\mathcal{A}(x) = \mathcal{A}(xy) + \mathcal{A}(x\overline{y})$, $\mathcal{A}(\overline{x}) = \mathcal{A}(\overline{x}y) + \mathcal{A}(\overline{xy})$, $\mathcal{A}(y) = \mathcal{A}(xy) + \mathcal{A}(\overline{x}y)$, $\mathcal{A}(\overline{y}) = \mathcal{A}(x\overline{y}) + \mathcal{A}(\overline{xy})$ も成立する。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、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、具体的な複数の下位情報に展開してもよい。以下は、情報 X の否定命題「X でない $(\overline{X})$」を明示的な情報 $X_2, X_3, \cdots$ に、情報 X の否定命題 $\overline{Y}$) 明示的な情報 $X_2, X_3, \cdots$ に展開した場合の認知環境である。

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

2.7　想定の両立・制限・独立性

本節で述べる内容は、基本的な確率の性質と同一である。まず、前節で $\mathcal{A}(xy)$ は「情報 X と情報 Y が「両立」するという想定確信度」を表すと述べたが、この「両立」という概念に多少の注意が必要である。[5] 今、ある交差点に立っている信号機 A を考える。も

[5]本節で後述する「条件制限」との混同しないようにされたい。

し、この信号機Aが故障していないという保証があるならば、信号機Aが青でありかつ同時に赤であるということはあり得ない。すなわち、ある特定の時間で、「信号機Aが青である」という事象 X と「信号機Aが赤である」という事象 Y は両立しない。したがって、$\mathcal{A}(xy) = 0$ ということになる。一方、信号機Aとは異なった場所にある信号機Bを考えると、ある特定の時間で「信号機Aが青である」という情報と「信号機Bが赤である」という情報は両立し得る。したがって、$\mathcal{A}(xy) > 0$ である。[6]

次に、事象の「独立」と「従属」という概念を明確にしておこう。今、日本にある信号機Aがアメリカにある信号機Bを考える。この時、「信号機Aが青である」という事象 X と「信号機Aが赤である」という事象 Y は、同時に、しかし互いに無関係に成立しうる。このように、お互いが無関係に両立する想定を、互いに**独立**であると言う。一方、信号機Bが信号機Aと同じ交差点にあり、直角にずれた位置に立っているものだとしよう。この時、ほとんどの場合で、信号機Aが赤になると信号機Bは青になる。[7] すなわち、信号機Bの状態は信号機Aの条件如何によりおおよそ判断がつく。こうした関係を**制限関係**と呼ぶことにしよう。

情報 X の情報 Y に対する制限の強さは、情報 X が生起した時、情報 Y がどの程度「両立して生起し得るか」という観点から計算できる。ここで大切な点は、束縛の強さの計算において、情報 X が生起しなかった場合は全く考慮しなくてよい (情報 X が生起しない時には、情報 X の制限力は発揮されないので、情報 Y は自由に振る舞っていてよい)、というという点にある。したがって、情報 X の情報 Y に対する制限の強さは、いわゆる条件付き確率 $P(Y|X) = \dfrac{P(X\&Y)}{P(X)}$ で表現できる。想定の場合も同じで、想定 x の生起を前提とした時に、想定 y の確信度をどの程度持てるかという計算は、想定 x と想定 y が両立するという確信度を想定 x の確信度で割ったものに等しい。すなわち、

(7) 想定 x を前提とした時の想定 y の確信度：$\mathcal{A}(y|x) = \dfrac{\mathcal{A}(xy)}{\mathcal{A}(x)}$

が成立する。この式は Stalnaker (1970) が、条件文に対する主観的確率の式として定義したものに等しい。また、Lewis (1973) も、条件文の主張可能性を式 (7) によって計算できると主張している。しかし、条件文の主張可能性はむしろ関連性の概念から考えた方がよいと思われる。この点については、3.4 節を参照されたい。

式 (7) は、想定間の「独立度」を考える上でも重要である。想定 x と想定 y が独立であったとしたら、想定 y が生起を前提にしていようといなかろうと、想定 y の確信度に変化は生じないはずである。したがって、$\mathcal{A}(y|x) = \mathcal{A}(y)$ という関係が成立する。この関係式は、(7) より $\dfrac{\mathcal{A}(xy)}{\mathcal{A}(x)} = \mathcal{A}(y)$ となり、式 (8b) に変形できる。すなわち、想定 x と想定

[6]確率論では、両立不可能な関係を「相互に排反」と言う。

[7]信号の変化は、完全に同時に起こるものではなく、信号機Aが赤になったほんのわずかな間、信号機Bも赤になっている。すなわち、「信号機Aが赤である」という状態と「信号機Bが青である」という状態は両立し得るし、「信号機Aが赤である」という状態と「信号機Bが赤である」という状態も両立しうる (どちらも青という状態だけは、故障でない限り起こらない)。ただ、「どちらも赤」という状態より「信号機Aは赤で、信号機Bは青」という状態のほうが起こりやすいというに過ぎない。

y が独立である時、それらが両立するという想定確信度は、各単独想定の確信度の積に等しい。この関係は暫定的な想定確信度を決める際に重要となる。

(8) 想定 x と想定 y が独立である時、以下が成立する：

a. $\mathcal{A}(y|x) = \mathcal{A}(y)$

b. $\mathcal{A}(xy) = \mathcal{A}(x) \cdot \mathcal{A}(y)$

2.8 前提情報と焦点情報

情報の両立という概念と情報の制限という概念の最大の違いは、前者には情報間の対称性が成立するのに対し、後者には情報の非対称性が生じる可能性を持つという点にある。すなわち、$\mathcal{A}(xy) = \mathcal{A}(yx)$ は常に成立し、$\mathcal{A}(xy) \neq \mathcal{A}(yx)$ があり得ないのに対し、$\mathcal{A}(y|x) \neq \mathcal{A}(x|y)$ は成立し得る (偶然 $\mathcal{A}(y|x) = \mathcal{A}(x|y)$ が成立することもある)。ここで、$\mathcal{A}(y|x)$ における想定 x を**前提**と呼び、想定 y を**焦点**と呼ぶことにする。[8] 前提情報・焦点情報は、いわゆる旧情報・新情報とほぼ類似した概念であるが (定延利之・熊谷吉治・苅田修司, 1999)、ここでは I-領域・D-領域の中で既に活性化している情報を前提情報、I-領域の中で初めて活性化を起こした情報 (獲得したばかりの情報も含む) を焦点情報と考える。また、焦点情報・前提情報に関して、以下の制限が掛かると仮定する。

(9) a. 焦点情報は、可能な限り 1 つでなければならない (Chafe, 1994)。

b. 焦点情報は探索しやすいため、頻度確率に近い形で想定確信度を持てるが、前提確率は多くの場合構造確率に近い形で想定確信度が設定される。

なお、残りの $\overline{x}, \overline{y}$ はアクセス可能情報と呼ばれる。表 2.6 における $x_1, x_2, x_3, y_1, y_2, y_3, \cdots$ もアクセス可能情報である。(9) は焦点情報の想定 y に関わる制約であり、この制約から $\overline{y}, y_1, y_2, y_3$ といった想定 y の否定に関わる想定は、想定 x と同レベルまで活性化されることは特殊な状況を除き、基本的にはない。これに対し、$\overline{x}, x_1, x_2, x_3$ といった前提情報の否定に関わる想定は、想定 x とほぼ同レベルまで活性化されることもある。

2.9 関連性の計算

式 (7) は、前提情報 X が焦点情報 Y の生起をどの程度制限しているかという指標としてはよいものである。しかし、情報 X が情報 Y をどの程度「束縛」しているかという指標としては不十分である。例えば、S という操作をしたところ R という反応が 100% の確率であったとする。この時、R という反応を引き出すのに、S という操作が「必要」であるとはいえない。S 以外の操作をした時に、R という反応が起こるかどうかを検証して始めて、S という操作の「必要性」を主張することができる。

この例は、関連性理論における**認知的関連性**をいかに計算すればよいかという点について一つのヒントを与えてくれる。想定 X の生起が想定 Y の生起を引き起こすことが多ければ、まずそれだけで、両者の間にある程度の関連性があると見なせそうである。しか

[8] $\mathcal{A}(x|\overline{y})$ なら y が前提、y が焦点である。

し、これだけでは関連性の計算として十分ではない。さらに、想定 X の生起がなかった時に、想定 Y の生起が起こらないことが確認できれば、十分な関連性があるといえる。[9] 以上のことから、想定 x の想定 y に対する関連性 $\mathcal{A}(x \Rightarrow y)$ に関する一般式として、式 (10) を考えることができる。[10]

(10)
$$\begin{aligned}\mathcal{A}(x \Rightarrow y) &= k_x \cdot \frac{\mathcal{A}(xy)}{\mathcal{A}(x)} - k_{\bar{x}} \cdot \frac{\mathcal{A}(\bar{x}y)}{\mathcal{A}(\bar{x})} \\ &= k_x \cdot \frac{\mathcal{A}(xy)}{\mathcal{A}(xy)+\mathcal{A}(x\bar{y})} - k_{\bar{x}} \cdot \frac{\mathcal{A}(\bar{x}y)}{\mathcal{A}(\bar{x}y)+\mathcal{A}(\overline{xy})}\end{aligned}$$

k_x は 情報 X が成立する状況をどの程度考慮にするか、$k_{\bar{x}}$ は 情報 X が成立しない場合の状況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る。前節との関係でいうと、k_x は 前提想定 x の活性度を、$k_{\bar{x}}$ は前提否定の想定活性度を示す指標といってもよい。いずれの係数も、$0 \leqq k_x \leqq 1$、$0 \leqq k_{\bar{x}} \leqq 1$ 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 の時は当該情報を全く参照しないことを意味する。

一般的に、Y に対する X の関連性を計算する時、想定 x は認知主体にとって顕在的になっていると見なせるため、$k_x = 1$ と置いて差し支えない。一方、バイアス係数 $k_{\bar{x}}$ の数値は情報 X の与えられ方によって異なる。例えば情報 X が「未成年」というものであれば、情報 $\neg X$ である「20 歳以上の成人」は想起されやすいであろうから、バイアス係数 $k_{\bar{x}}$ の数値は 1 に近く設定できる。しかし、情報 X が「25 番目の素数」といったものであれば、情報 $\neg X$ をどの範囲まで探索すべきか明確ではない。「25 番目の素数」以外の数は無限にあるため、全てを探索するのは不可能であるし、「25 番目の素数があった時、それに 3 を足せば 100 になる」といった文の真偽を考えるとき、。「25 番目の素数」以外の数を考える必要もない。こういう場合、バイアス係数 $k_{\bar{x}}$ の数値は 0 に近く設定されるであろう。なお、言うまでもなく、式 (7) は式 (10) の特殊な場合であり、$k_x = 1$, $k_{\bar{x}} = 0$ と置いた場合に等しい。

関連性の確信度 (10) は、$-1 \leqq \mathcal{A}(x \Rightarrow y) \leqq 1$ の範囲を持つ。このうち、認知環境に取り込まれるのは、(5a) の定義から 0 以上のものであるため、関連性について次のような制約を設けることにしよう。

(11) **関連性に基づく認知環境の制約**

a. 関連性の確信度が 0 以上であれば、その情報は認知環境に取り込まれる。複数の解釈が可能であった時は、関連性の高い解釈ほど優先される。

b. 関連性の確信度が 0 未満であれば、その情報は無視されて認知環境に取り込まれないか、あるいは認知環境に取り込み可能な別の表現として理解される。

なお、$a = \frac{b}{c}$ 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。$a = \frac{b}{0}$ において、$b \neq 0$ の場合、こ

[9] 十分な関連性があった時に初めて、我々は両者の間の「因果関係」も疑うことができる。ただし重要なことだが、関連性があっても、因果関係があるとすぐに決めつけることはできない。因果関係と関連性関係は別の概念である。

[10] なお、関連性の定義式は他の方法も考えられるが、式 (10) はその中で最も単純な定義式である。

の演算は不能 (演算ができない) となる。一方、$b = 0$ の時、演算は不定 (解が無数に存在) となる。ここで (10) 式について見てみると、分母がゼロになる時には、常に分子もゼロになり、解は 0〜1 までの任意の実数となる。したがって、ある想定の確信度がゼロであった時、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ない。また、関連性の確信度が不定になってしまったとしても、その情報は一旦認知環境に取り込まれる。ただし、確信度がゼロのままであったなら、認知環境の改善は望めないため、(5c) にしたがって、あるタイミングで認知環境から廃棄される。

2.10 関連性確信度の様々な解釈

ここで、式 (10) の意味を、統計学の観点から解釈しておこう。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 (10) は

(12) $\mathcal{A}(x \Rightarrow y) = h_x \cdot P(\mathcal{A}(y)|\mathcal{A}(x)) - h_{\overline{x}} \cdot P(\mathcal{A}(y)|\mathcal{A}(\overline{x}))$

とも表現できる。同様に、式 (10) 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。

$$P(X = 1) = \mathcal{A}(xy) + \mathcal{A}(x\overline{y}), \quad P(X = 0) = \mathcal{A}(\overline{x}y) + \mathcal{A}(\overline{xy}),$$
$$P(Y = 1) = \mathcal{A}(xy) + \mathcal{A}(\overline{x}y), \quad P(Y = 0) = \mathcal{A}(x\overline{y}) + \mathcal{A}(\overline{xy})$$

$$E(X) = \mathcal{A}(xy) + \mathcal{A}(x\overline{y}), \quad E(X^2) = \mathcal{A}(xy) + \mathcal{A}(x\overline{y}),$$
$$E(Y) = \mathcal{A}(xy) + \mathcal{A}(\overline{x}y), \quad E(Y^2) = \mathcal{A}(xy) + \mathcal{A}(\overline{x}y) \qquad E(XY) = \mathcal{A}(xy)$$

$$\begin{aligned} V(X) &= E(X^2) - (E(X))^2 \\ &= (\mathcal{A}(xy) + \mathcal{A}(x\overline{y}))\,(\mathcal{A}(\overline{x}y) + \mathcal{A}(\overline{xy})) \end{aligned} \qquad \begin{aligned} Cov(X, Y) &= E(XY) - E(X) \cdot E(Y) \\ &= \mathcal{A}(xy)\mathcal{A}(\overline{xy}) - \mathcal{A}(x\overline{y})\mathcal{A}(\overline{x}y) \end{aligned}$$

(13)
$$\begin{aligned} \text{回帰係数}\,\beta &= \frac{Cov(X, Y)}{V(X)} \\ &= \frac{\mathcal{A}(xy)}{\mathcal{A}(xy) + \mathcal{A}(x\overline{y})} - \frac{\mathcal{A}(\overline{x}y)}{\mathcal{A}(\overline{x}y) + \mathcal{A}(\overline{xy})} \qquad \left(= \frac{\mathcal{A}(xy)}{\mathcal{A}(x)} - \frac{\mathcal{A}(\overline{x}y)}{\mathcal{A}(\overline{x})}\right) \end{aligned}$$

したがって、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、係数が小さくなると共に、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。なお、統計学における相関係数は、$r = \sqrt{\mathcal{A}(x \Rightarrow y) \cdot \mathcal{A}(y \rightarrow x)}$ によりを計算可能 (符号は $\mathcal{A}(x \Rightarrow y)$ に一致) であり、またいわゆる決定係数は r^2 に等しい。

2.11 認知環境の更新

(1), (2.6) 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 (2) に示した認知効果の数量的表現にもつながる。簡単に想定値

の変動に伴う効果について見てみよう。今、(14) 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、(15) に示すようにいくつかの可能性が考えられる。

(14)

	$\mathcal{A}(y)$	$\mathcal{A}(\overline{y})$	計
$\mathcal{A}(x)$	a	b	$a+b$
$\mathcal{A}(\overline{x})$	c	d	$c+d$
計	$a+c$	$b+d$	1

(15) a.

	$\mathcal{A}(y)$	$\mathcal{A}(\overline{y})$	計
$\mathcal{A}(x)$	$\frac{a+\delta}{1+\delta}$	$\frac{b}{1+\delta}$	$\frac{a+\delta+b}{1+\delta}$
$\mathcal{A}(\overline{x})$	$\frac{c}{1+\delta}$	$\frac{d}{1+\delta}$	$\frac{c+d}{1+\delta}$
計	$\frac{a+\delta+c}{1+\delta}$	$\frac{b+d}{1+\delta}$	1

b.

	$\mathcal{A}(y)$	$\mathcal{A}(\overline{y})$	計
$\mathcal{A}(x)$	$a+\delta$	$b-\delta$	$a+b$
$\mathcal{A}(\overline{x})$	c	d	$c+d$
計	$a+\delta+c$	$b-\delta+d$	1

c.

	$\mathcal{A}(y)$	$\mathcal{A}(\overline{y})$	計
$\mathcal{A}(x)$	$a+\delta$	b	$a+\delta+b$
$\mathcal{A}(\overline{x})$	$c-\delta$	d	$c-\delta+d$
計	$a+c$	$b+d$	1

d. ...

このことから、ある想定値が 0 である (16-i) と、ある想定が認知環境に組み込まれていない (16-ii) は計算論的な意味が異なることが分かる。

(16) (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

(16-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、(16-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、前述した「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。(15) 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。常に複数の可能性がある。その可能性の中で、最適な認知環境を選び出すのが関連性の計算であり、関連性の確信度がより 1

に近い認知環境ほど、よりよいものであるといってよい。2.5 節で見たとおり、「知っているが分からない」というそもそも「知らない」ということは本質的な違いがある。この違いは、認知環境の更新によってもたらされる。その認知環境の更新は、より関連性の高いものを目指して行われる。そして、各想定がより関連性の高い認知環境を構成しているという条件は、「知識」と呼びうる情報に関する必要条件の一つといってもよい。

3. 想定確信度と真理値

3.1 知識の哲学

A.J. (1981) は、P ということを S が知っていることの条件として、以下の 3 つを挙げている。これは、『メノン』の中でアリストテレスが定義した「知識とは正当化された真なる信念 (knowledge is justified true belief.)」という考え方と基本的に同一である。

(17) a. X が真であるという確信を S が持っている。
　　b. X は真である。
　　c. X が真だということに S が確信を持つのは当然である。

戸田山 (2002) の表現を借りるなら、これらの性質は次のように言い換えることができる。

(18) a. S は X を信じている。
　　b. X を信じている S の信念は真である。
　　c. S の信念は正当化されている。

戸田山 (2002) はこうした知識の定義が適切でないことをゲティア問題などの例を通じて論じ、またゲティア問題をクリアしようとするドレツキの確率論的アプローチなども不十分であると主張している。その上で、戸田山は認識論は次のように作り直されるべきであると指摘する。

(19) a. 自然化された認識論：知識獲得の科学的・心理学的な獲得手段を決定すること。
　　b. 社会化された認識論：個人ではなく集団として何かを知っているという点にまで知識を拡張すること。
　　c. 知識の基盤を信念ではなく、情報に置くこと。
　　d. 命題の真偽を中心概念としないこと。

関連性理論は、こうした戸田山の提案に十分ではないにせよ、ある程度答えられる可能性を持っている。まず、(19a) から見てみよう。妥当な認知環境を生成する関連性計算 (10) を理想的に行うためには、認知主体は $\mathcal{A}(y|x)$ と $\mathcal{A}(y|\bar{x})$ に関する確信度を 1 と 0 に設定できることが要求される。言い換えるなら、「肯定的前提想定 x の元で肯定想定 y を確信している」と同時に、「否定的前提想定 $\bar{x}$ の元では想定 y を完全に否定できる」時に、関連性は最大となる。これを認識論に当てはめるなら、次のように言うことができる。すなわち、「情報 Y について何かを知る」ためには、「情報 X に関する想定 x が満

たされた時に Y に関する想定 y を確信できる」と同時に、「X が成立しないという想定 $\overline{x}$ を持っている時には Y に関する想定 y を完全に否定できる」ような情報Xを見つけ出さなければならない。これが知識獲得の手段の必要条件である (十分条件ではない)。

(19b) の提案は多少やっかいで、**共有知識のパラドックス**を回避する必要がある。まず、「(私は、ではなく) 私たちは K を知っている」という言明が正当化されるには、私たちという集団の中で、誰が K を知っていて、誰が K を知らないかが明確でなければならない。そうでなければ、誤った信念が知識として認定されてしまう。さて、今、K という命題を知っている A さんが、「B さんも K を知っている」ということを知っているとする。これだけでは、集団の知識とはまだ言えない。B さんから見ると、「私 B が K を知っていることを A さんが知っている」ということを B さんが知っていなければならない。しかし、これだけでは正当化は十分ではない。A さんは、「B さんが K を知っていることを私 A が知っているということを、きちんと B さんが知っている」ということも分かっていなければならない。これが無限に続く。同様に、A さんが K を知っていて、B さんが K を知らない場合にも、お互いがそのことを確認して正当化を行おうとした時に、無限後退が発生する。このように集団の知識の「正当化」は、極めて難しい。田窪・金水 (1996b) による談話管理理論の美しい点は、こうした知識共有のパラドックスを回避できるところにある (田窪・金水, 1996a; 三藤, 1999)。談話管理理論では、自分と他者の知識という前提は行われず、あくまで「自分」の持っている知識をどのように管理するかということが問題とされる。既に見たとおり、D-領域は恒常的な知識を管理する場であり、I-領域は検証されていない知識を管理する場である。認知環境は、この D-領域と I-領域にある情報の総体として捉えられる。[11]

(19c) についても、認知環境という点からある程度答えることができるだろう。まず、認知環境に組み込まれている想定 (顕在的情報) は信念ではない。認知環境における想定には、確信度が 0.5 のようなものもあれば、0 に近いものや 1 に近いものなどが全て含まれる。こうした想定のうち、まず確信度が完全に 0 である想定は信念とは呼べない。また、確信度が 0 より大きな想定は、信念になり得る「可能性」があるが、やはり信念ではない。2.5 で述べたように、認知環境の更新が行われていく中で、初めて「信念」と呼べる想定を判断することができ、その中で想定確信度が安定し、1 に漸近するもののみが「知識」と呼んでもよいものなのである。知識が、「正当化された真なる信念」であるのは結果にすぎない。ある瞬間「だけ」の心的状態からは何が知識であるかは保証されない。信念や知識を保証するためには、認知環境の更新履歴、人間の経験累積、時間の流れが必要不可欠なのである。

最後の (19d) は、戸田山 (2002) の提案の中で、最も難しい問題といってよいだろう。真理値は、論理の最も基本的な構成概念であり、論理は意味の基盤である。つまい、真理値が決まらない限り、認知主体は内容の理解ができない。

[11] 一般的な関連性理論では、認知環境の内部を 2 種類に区別していないが、計算論的関連性理論では、この区別は必要であり、I-領域における想定の確信度を変更するほうがコストが低い。

論理学において、真理値はモデルと命題に基づいて決定される。今、ある有限集合を一つ固定し、その元から一定の規則によって生成される集合を言語と呼ぶとき、その一つの言語に属する記号列 (文) が命題である。モデルとは一つの状況・世界のことである。真理値は、命題とモデルの関係を評価したものであり、命題がモデルに当てはまる場合を真といい、そうでない場合を偽と呼ぶ。モデル・命題・真理値のうち、2 つが所与のものであれば、残りの 1 つを決定できる。私たちの言語活動における最も基本的な性質は、これを利用したものである。例えば、我々が意味のある発話を行えるのは、モデルと真理値が明確であるため、何が命題であり、何が命題でないかを決めることができるからである。また、命題と真理値が明確であれば、その命題を成立させるモデルを構成することが可能になり、意味を理解することができる。言語活動だけでなく、思考 (認識・知識) においても、モデル・命題・真理値の関係は重要であろう。命題とモデルが完全に与えられていて初めて、真理に関する最も初歩的な判断が可能になるからである。

一般的な論理学において、世界やモデルは情報の範囲が明確に規定された完全なものとして与えられる。しかし、私たちの日常的な認知活動において、情報の総体は明確に定められるものではない。それは暫定的なものであり、不安定なものであり、せいぜい良くてもある種の近似であったりする。そもそも、何らかの新しい情報を獲得し、認知環境を更新してゆくことは、認知にとって本質的な性質といってもよい。命題もまた常に明確なものとはいえない。我々が利用できる情報の全てが命題として表現されるとは限らないし、たとえ言語により表現された情報であっても、命題の形を成してなしていない情報の欠落した文からものごとを判断しなければならないこともしばしばある。そうなると、モデル・命題・真理値のうち、最も安定して使えなければならないものは真理値・真理判断の能力である。既に見たように、言語産出や言語理解のいずれの場面においても、真理値は明確になっていなければならない。言うまでもなく、否定や連言といった論理子の機能も真理値により定義され、意味の基本となる外延・内包の定義も真理値に依存する。

真理は確かに存在する。ある命題の真理値も決まっている。しかし、現実の認知主体にとって、ある情報が真理か否かが常に明確であるとは限らない。真理値も認知主体には隠されていることが多い。こうした点から、二値の真理値に限定されない多値論理も提案されてきた。次節では、関連性理論の認知環境に基き、想定確信度から認知主体が真理値と 見なしうる 値をどのように計算すればよいかという点について議論を行う。

3.2 想定確信度に基づく真理値の計算

認知主体にとって真理は必ずしも知り得るものではない。しかし、その場合であっても、二値の真理値が必ずしも無意味になるとは限らない。例えば、「今日 (2008 年 2 月 21 日) の午前 11 時は晴れである」という命題は、必ず真か偽かのいずれかである。また、私はその真理値を今は正確には知り得ないが、「雨になるだろう」という判断に従って「傘を持って外出する」という決断をすることができ、実際に傘を持ってでかけるという行動を起こすことができる。すなわち、真理であるかどうかとは別に、認知主体はある

命題について何らかの判断 (区別) を下すことはできる。

このある命題について何らかの判断 (区別) が行えるという性質は、真理値の必要条件の一つである。特に、ある命題とその否定命題との区別が可能であれば、少なくとも最低限の「真偽」の判断を行えるということだ。この性質は、「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。ある命題が未知であれば、その否定命題も未知であるが、ある命題の真偽が明確である場合には、その否定命題の真理値も決定できる。

(20) 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

情報を区別するための指標は様々なものが考えられるであろうが、最も妥当な指標の一つは情報のエントロピー (情報の曖昧度) である。エントロピーが高ければ区別が困難であるし、エントロピーが低ければ区別は容易であろう。そこで、このエントロピーという概念を応用して、真理値の代替物とすることを考える。

まず、想定の確信度を、顕在化された各情報に対して、一定の基準に従って与えられる定量的な価値と見なそう。この時、想定確信度のエントロピーを計算することで、想定の情報量を定義することができる。ここで、情報 X に対する想定 x のエントロピー (情報の曖昧度) は、想定確信度 $\mathcal{A}(x)$ を用いて次式で計算できる。[12]

(21) **想定 x のエントロピー：**$\mathcal{E}(x) = -\mathcal{A}(x)\cdot \log_2\mathcal{A}(x) - (1-\mathcal{A}(x))\cdot \log_2(1-\mathcal{A}(x))$

$\mathcal{E}(x)$ は常に 0〜1 までの値を取る。$\mathcal{E}(x) = 1$ は、情報が曖昧であり、乱雑度が高く、最も不確定な情報であることを示す。$\mathcal{E}(x) = 0$ ならば、確定的な情報であり、曖昧性はない。実際に、$\mathcal{A}(x)$ 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 $\mathcal{E}(x) = 0$ となり、$\mathcal{A}(x)$ が 0.5 の時にはエントロピー最大 ($\mathcal{E}(x) = 1$) となる。[13] 次に、このエントロピーの数値 $\mathcal{E}(x)$ から、多値の判断価値 $\mathcal{V}(x)$ を定義する。

(22) **想定 X の判断価値 $\mathcal{V}(x)$：**

$$\mathcal{V}(x) = \frac{1 + j\cdot(1-\mathcal{E}(x))}{2}$$

ただし、j は ± 記号であり、$\mathcal{A}(x) \geqq 0.5$ の時 $j = 1$, $\mathcal{A}(x) < 0.5$ の時 $j = -1$

この判断価値 $\mathcal{V}(x)$ は 0〜1 までの値を取り、情報のエントロピーの低いければ低いほど 0 か 1 に近い値を、エントロピーが高くなるにつれて 0.5 に近い値を取る。具体的には、

[12] 以下の計算では、単独想定 x を用いているが、両立想定 xy の場合でも同様に計算は可能である。

[13] 前述したように、$\mathcal{A}(x)$ が 0.5 である時、想定 x と $\bar{x}$ を同程度に考慮できることを意味する場合と、は想定 x に関して無知であることを示す場合がある。これらは認識論上は区別すべき事柄であるが、エントロピーという観点からは同一に扱ってよい。

$\mathcal{A}(x) = 1$ の時は $\mathcal{V}(x) = 1$, $\mathcal{A}(x) = 0.5$ ならば $\mathcal{V}(x) = 0.5$, $\mathcal{A}(x) = 0$ では $\mathcal{V}(x) = 0$ といった関係になる。この情報の判断価値を使って、我々はその情報を「採択」するか「棄却」するかを決めることができる。すなわち、$\mathcal{V}(x)$ が判断の閾値 c より大きければ情報を採択、閾値 C 以下であれば情報を棄却すればよい。閾値 C の決め方は様々であるが、例えば情報のエントロピーが最大になる時、$\mathcal{V}(x)$=0.5 となることから、閾値を $C = 0.5$ とし、$\mathcal{V}(x)$ が 0.5 よりも大きければ情報を採択するといったことが考えられる。極端な例でいえば、閾値を $C = 0$ とし、$\mathcal{V}(x) > 0$ なら情報を採択し、$\mathcal{V}(x) = 0$ の場合のみ情報を棄却するといったことも考えられる。逆に、最も厳しい判断基準を考え、$\mathcal{V}(x) = 1$ の時のみ情報を採択し、$\mathcal{V}(x) < 1$ なら情報を全て棄却してしまうということも考えられる。

ここで、ごく標準的な論理の真理値を考える。特に、ここでは包括的選言 (inclusive disjunction) と実質含意 (material implication) を取り上げよう。

(23)

x	y	x∨y	x→y
真	真	真	真
真	偽	真	偽
偽	真	真	真
偽	偽	偽	真

この包括的選言・実質含意の真理表は、偽となる状況は唯一に限定されているのに対し、真となる状況は 3 つの可能性があることを示している。すなわち、偽は絶対的であり、偽でないものは全て真と判断される。

これを、「真理」ではなく、人間の認識に移し替えてみると、次のことがいえる。もし、この論理表現を現実の言語表現と見なすなら、現実に起こった事態は、真となる状況のうちの一つで、残りの二つの状況は実際には起こらなかったことになるのだが、しかし「起こりうる可能性があった」のなら真と見なせる。逆に、偽となる状況は絶対的に偽と判断しうるものである。ここから我々は判断の閾値設定について、一つの値を導き出せる。すなわち、閾値を $C = 0$ とし、判断価値が閾値以上となる $\mathcal{V}(x) > 0$ なら情報を採択し、それを便宜的に「真」と呼ぶ。真となる可能性は多様だが、可能性がある限り、それを真として判断する。一方、判断価値が閾値に一致する $\mathcal{V}(x) = 0$ の場合のみ情報を棄却し、そうした情報を便宜的に「偽」と呼ぶ。偽となる状況は判断価値が 0 以外の場合にはなく、唯一絶対的に限定される。これが標準的な論理に近い真理値をもたらすための、プラグマティックな方法なのである。[14]

3.3 真理値の見込みとしての想定確信度

前節で定義した判断価値 $\mathcal{V}(x)$ と想定確信度 $\mathcal{A}(x)$ の間には、線形ではないが、**単調な関係**が成立する。すなわち、想定確信度が 1 に近づくほど、判断価値は 1 に近づき、確信

[14] 言うまでもなく、判断基準の閾値は任意に設定可能である。また判断価値を二つ以上のクラスに分割してもよい。例えば、基準値を $C_1 = 0$, $C_2 = 1$ と 2 つ設定し、判断価値を $\mathcal{V}(x) = C_1$, $C_1 < \mathcal{V}(x) < C_2$, $\mathcal{V}(x) = C_2$ と 3 段階に分けて、それぞれを「偽」「未知」「真」と呼ぶことにすると、(20b) のような三値論理に対応する真理値を設定できる。

度が 0 に近いほど判断価値も 0 に近くなる。この性質は、想定確信度を真理値の代用品として使い得ることを示唆している。実際、想定確信度の数値を単なる個人的な基準値ではなく、ほぼ全ての人の共通認識における強度と見なすことができるのであれば、情報 X の想定確信度が 0 の時には、即座にその情報の「真理値」は偽であると結論づけてよい。エントロピーの計算を行うまでもない。同様に、情報 X の想定確信度が 0 より大きければ、その情報は「(正しい可能性があるという意味での) 真」と考えることができる。また、真か偽かの基準値 C に関しても、その閾値は状況に合わせて柔軟に設定できることから、想定確信度を真理値と見なすための、適切な閾値 C' の設定を行えばよい。

全知全能ではない認知主体にとっては、実世界で適切な認知・思考を行うために、対象となる想定を暫定的に真あるいは偽であると見なして、推論を進めていかざるを得ない。想定の確信度は、こうした真理値の見込み計算を迅速に行う上で、最もよい指標であると思われる。このような前提に立つと、次節で見るとおり、想定確信度に基づいて、擬似的な命題論理を構成することも可能になる。

3.4 関連性に基づく擬似的な論理計算

想定確信度を真理値の見込みと考えることができるとするならば、命題の「真理値」は二値ではなく、多値のものとなる。多値の真理値を用いても、一貫した論理が構成できることは、これまでに行え割れてきた多くの研究が示す通りである。例えば、代表的な多値論理であるァジィ論理では、各論理演算子のメンバーシップ度を以下のように計算する。なお、Mem(X) は情報 X のメンバーシップ度を表すメンバーシップ関数、min は最小値を求める関数、max は最大値を求める関数である。

(24) a. 否定：$\mathrm{Mem}(\neg X)=1-\mathrm{Mem}(X)$,
b. 連言：$\mathrm{Mem}(X\wedge Y)=\min(\mathrm{Mem}(X),\mathrm{Mem}(Y))$,
c. 選言：$\mathrm{Mem}(X\vee Y)=\max(\mathrm{Mem}(X),\mathrm{Mem}(Y))$ として定義
d. 含意：$\mathrm{Mem}(X\rightarrow Y) = \min(1, 1-\mathrm{Mem}(X)+\mathrm{Mem}(Y))$

こうした論理においても、例えば $\neg(X\wedge Y) = \neg A\vee\neg Y$ といったド・モルガンの法則や、$X\rightarrow Y = \neg X\vee Y$ といったトートロジカルな性質が保たれる。しかし、不自然な点も存在する。例えば、$\mathrm{Mem}(X\rightarrow Y)$ の定義に含まれる $1-\mathrm{Mem}(X)+\mathrm{Mem}(Y)$ という部分は、$X\rightarrow Y = \neg X\vee Y$ というトートロジーを表現したものと理解できるが、同じ定義の中に含まれる $\min(1, \triangle)$ という連言部分の表現は、メンバーシップ度が 1 を超えることを防ぐ以外に、何を「意味」するのかが明確ではない。また、$\mathrm{Mem}(X\rightarrow Y)$ の定義では、前件のメンバーシップ度が小さいほど、後件のメンバーシップ度に関わらず、$X\rightarrow Y$ のメンバーシップ度が高くなることになるが、これは人間の思考過程をうまく捉えてはいない。そこで、計算論的関連性理論に基づき、ファジィ論理とは類似した「擬似論理」の構成を試みて見よう。特に、本節では、命題論理の最も基本的な論理子である否定・含意・連言・選言について簡単に議論を行う。なお、同値と排他的選言に関しては、各々含意と選言を議論する中で触れる。

Sperber and Wilson (1986) は、認知主体にとって「関連性」の計算が認知の中心にあることを主張した。計算論的関連性理論に基づく擬似論理 (以下 RPL: Relevance theoretic Pseudo-Logic と略記) では、こうした考え方に基づき、関連性に関する想定確信度こそが最も基本的な数値であると考える。すなわち新しい認知活動や談話セッションが始まった時、まず妥当な関連性の確信度が設定され (多くの場合、$\mathcal{A}(x \Rightarrow y) = 0.5$ あるいは $\mathcal{A}(x \Rightarrow y) = 1$)、式 (10) に従って、$\mathcal{A}(xy)$ などの確信度が計算されるとする。つまり、認知活動が始まった時点では、式 (10) は右辺から左辺の数値を求めるために使われる。認知活動が継続していくにつれて、式 (10) は左辺から右辺への計算—すなわち、関連性の計算を行うためにも使われることになる。

単独の情報 X に関する想定 x の確信度 $\mathcal{A}(x)$ も、情報 X から単独で決定されるものではない。意識しているか無意識しているかは別にして、あるコンテクスト C における情報 X の関連性確信度 $\mathcal{A}(© \Rightarrow x)$ が単独の想定確信度 $\mathcal{A}(x)$ となる。

(25) 単独の想定確信度：$\mathcal{A}(x) = \mathcal{A}(© \Rightarrow x)$

コンテクスト C は、前提情報の一種であるが、ある認知セッション内で、決して ¬C が考慮されることがないという点で特別な前提情報といえる。記号 © は、C のある認知セッション内で、C の内部情報しか使われないことを意味する。

単独想定の確信度は、コンテクスト C をどのように検索するかによって大きく異なる。もし認知主体がコンテクスト C における様々な情報を十分に走査できるのであれば、想定確信度は頻度確率に近い数値として与えられるであろう。しかし、対象について未知である場合、あるいは単なる仮説である場合には、2.5 節で見たように、構造確率に近い数値を使う。構造確率は、現実を移すには十分な指標ではないが、個別情報を走査する必要がなく、すばやく計算が行えること、またたとえ誤った想定確信度を持っていても、認知環境を更新していく過程で修正すればよいことを考えると、認知主体が構造確率に基づいて想定確信度を決めるのも妥当なことだといえる。この点に関しては、4.5 節も参考されたい。

情報 X の否定については、式 (26) で計算されるものとする。なお、$k_{\bar{x}}$ は式 (10) におけるバイアス係数と同一のもので、否定命題の顕在性を表す。

(26) a. 否定の想定確信度：$\mathcal{A}(\bar{x}) = k_{\bar{x}} \cdot (1 - \mathcal{A}(x))$

否定のフレームを完全に考慮できた時 (すなわち $k_{\bar{x}} = 1$ の時) には、$\mathcal{A}(\bar{x}) = 1 - \mathcal{A}(x)$ となり、ファジィ論理の計算式と一致する。否定命題の顕在性が低い時には、否定の想定確信度は過小評価される。したがって、関連性理論に基づく擬似論理では、$\mathcal{A}(x) + \mathcal{A}(\bar{x}) \leqq 1$ となるため、一般的に、二重否定 ¬¬X の想定確信度が X の想定確信度と一致することが保証されない。二重否定が保証されるのは $k_{\bar{x}} = 1$ の時のみである。また、一般的に排中率 (Law of excluded middle) も保証されていない。ただし、$k_{\bar{x}} = 1$ かつ想定度が 1 か 0 のいずれかの場合に排中率が成立し、一般的な命題論理の否定と一致する。

次に、含意 (implication)・同値 (equivalence) の擬似計算について見てみよう。RPL において、「X ならば Y」という表現は、正に関連性の概念そのものであり、含意や同値に関する擬似的な真理値も、正に関連性の計算式 (10) によって計算される。

ここで、真理値という面から、含意と同値を捉え直してみる。含意における「真」は、前件が真である場合、絶対的な真と判定できるものであるが、前件が偽である時には、真になる可能性を意味する値である (以下の真理表 (27) 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)。一方、同値解釈における「真」は絶対的な完全解釈がなされる。

(27)

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

この可能性・絶対性という違いを、式 (10) に反映させてみよう。今、式 (10) において、$\mathcal{A}(x \Rightarrow y) = 1$ が成立する (すなわち絶対的に真と見なせる) 条件を計算すると、$0 \leqq \dfrac{\mathcal{A}(xy)}{\mathcal{A}(xy) + \mathcal{A}(x\overline{y})} \leqq 1$, $0 \leqq \dfrac{\mathcal{A}(\overline{x}y)}{\mathcal{A}(\overline{x}y) + \mathcal{A}(\overline{xy})} \leqq 1$ より、任意の $k_{\overline{x}}$ に対し、$\dfrac{\mathcal{A}(xy)}{\mathcal{A}(xy) + \mathcal{A}(x\overline{y})} = 1$, $\dfrac{\mathcal{A}(xy)}{\mathcal{A}(\overline{x}y) + \mathcal{A}(\overline{xy})} = 0$ でなければならない。したがって、$\mathcal{A}(x\overline{y}) = 0$, $\mathcal{A}(\overline{x}y) = 0$ となる必要があり、$X \wedge \neg Y$, $\neg X \wedge Y$ が共に偽 (すなわち (27) 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、$\mathcal{A}(x \Rightarrow y) = 1$ を満たす演算は、同値解釈と等しい。

次に、式 (10) において、真の可能性があると見なせる場合、すなわち $\mathcal{A}(x \Rightarrow y) \geqq 0$ となる場合を考えてみよう。任意の $k_{\overline{x}}$ に対し $\mathcal{A}(x \rightarrow y) \geqq 0$ となるには、$\dfrac{\mathcal{A}(xy)}{\mathcal{A}(xy) + \mathcal{A}(x\overline{y})} = 1$ であればよいため、$\mathcal{A}(x\overline{y}) = 0$ のみが成立すればよい。これは (27) の 2 段目に相当する $X \wedge \neg Y$ のみが偽となることに相当する。すなわち、$\mathcal{A}(x \rightarrow y) \geqq 0$ 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 (10) におけるバイアス係数 $k_{\overline{x}}$ を $k_{\overline{x}} = 0$ と設定することに等しい。この時に、完全解釈、すなわち $\mathcal{A}(x \Rightarrow y) = 1$ となる条件を求めると、やはり $\dfrac{\mathcal{A}(xy)}{\mathcal{A}(xy) + \mathcal{A}(x\overline{y})} = 1$ となり、$\mathcal{A}(x\overline{y}) = 0$ のみが成立すればよいことが分かる。

一般的に、関連性の想定確信度 $\mathcal{A}(x \Rightarrow y)$ は 0 から 1 までの様々な値を取る。したがって、$k_{\overline{x}} = 1$ である場合に、関連性の想定確信度が 1 に近づくほど同値に近い解釈が得ら

れる。また、バイアス係数 $k_{\bar{x}}$ も 0 から 1 まで、様々に設定され得る。この時、バイアス係数が大きくなるほど、やはり同値に近い計算が可能になる。なお、バイアス係数 $k_{\bar{x}}$ が $k_{\bar{x}} = 0$ と設定された場合の $\mathcal{A}(x \Rightarrow y)$ は、情報 X の Y に対する制限 $\mathcal{A}(y|x)$ (式 (7) を参照) と等しい点にも注意されたい。

次に、連言の擬似計算について見てみよう。連言 X∧Y の想定確信度 $\mathcal{A}(xy)$ は、以下の条件下で計算される。

(28) a. $k_{\bar{x}} = 0$ かつ
$\mathcal{A}(xy) = \mathcal{A}(x \Rightarrow y) \cdot \mathcal{A}(x)$ あるいは $\mathcal{A}(xy) = \mathcal{A}(y \Rightarrow x) \cdot \mathcal{A}(y)$
b. すなわち、$\mathcal{A}(xy) = \mathcal{A}(y|x) \cdot \mathcal{A}(x)$ あるいは $\mathcal{A}(xy) = \mathcal{A}(x|y) \cdot \mathcal{A}(y)$
c. 情報 X と情報 Y が独立であるなら、$\mathcal{A}(xy) = \mathcal{A}(x) \cdot \mathcal{A}(y)$

(28c) は、(8) で見たとおり、情報 X と情報 Y が独立しているならば、$\mathcal{A}(y) = \mathcal{A}(y|x)$, $\mathcal{A}(x) = \mathcal{A}(x|y)$ が成立するため、(28a, b) から導出される重要な性質である。

最後に選言についてみておこう。これは次式によって計算される。前述した通り、k_x, k_y は、X, Y という肯定情報の顕在性を示すバイアス係数であり、特殊な場合を除き、一般的に 1 と置くことができる。一方、$k_{\bar{x}}$, $k_{\bar{y}}$ は、X, Y の否定情報の顕在性を示すバイアス係数であり、各々 0〜1 までの値を取る。

(29) 選言の想定確信度：$\mathcal{A}(x + y) = \mathcal{A}(x) + \mathcal{A}(y) - (1 + \mathcal{A}(\bar{x}|y) \cdot \mathcal{A}(\bar{y}|x)) \cdot \mathcal{A}(xy)$

式 (29) において、$\mathcal{A}(\bar{x}|y)$, $\mathcal{A}(\bar{y}|x)$ は 0〜1 までの値を取ることから、$1 + \mathcal{A}(\bar{x}|y) \cdot \mathcal{A}(\bar{y}|x)$ は 1〜2 までの値を取り、$\mathcal{A}(\bar{x}|y)$, $\mathcal{A}(\bar{y}|x)$ の数値が大きくなるほど (つまり片方が成立した時に、もう片方が不成立である可能性が高いほど)、$1 + \mathcal{A}(\bar{x}|y) \cdot \mathcal{A}(\bar{y}|x)$ は 2 に近い値を取る。ここで、今、$\mathcal{A}(x) = \mathcal{A}(xy) + \mathcal{A}(x\bar{y})$, $\mathcal{A}(y) = \mathcal{A}(xy) + \mathcal{A}(\bar{x}y)$ より、$\mathcal{A}(x) + \mathcal{A}(y) - 1 \cdot \mathcal{A}(xy) = \mathcal{A}(xy) + \mathcal{A}(x\bar{y}) + \mathcal{A}(\bar{x}y)$, $\mathcal{A}(x) + \mathcal{A}(y) - 2 \cdot \mathcal{A}(xy) = \mathcal{A}(x\bar{y}) + \mathcal{A}(\bar{x}y)$ が成立することから、$\mathcal{A}(\bar{x}|y)$, $\mathcal{A}(\bar{y}|x)$ の数値が小さければ、$\mathcal{A}(x + y)$ は包含的選言に、$\mathcal{A}(\bar{x}|y)$, $\mathcal{A}(\bar{y}|x)$ の数値が大きくなればなるほど、$\mathcal{A}(x + y)$ は排他的選言に近い演算になっていることが分かる。

3.5 モダリティ

関連性に基づく擬似論理では、一種の様相性も扱うことができる。言語表現におけるモダリティは、状況や世界、想定に関して、単にそれが存在すると述べるのではなく、どのように存在するのか、その存在価値はどのようなものなのかを表す意味論的なカテゴリーである。想定確信度は 0〜1 までの連続的な値を取ることから、この値を (かなり荒っぽい形ではあるが) 様相性と関係づけることができる。例えば、話者にとっての想定確信度が 0 から 1 に近づくにつれ、「(前提想定 x において) y でない ($\mathcal{A}(x \Rightarrow y) = 0$)」「$Y$ でありうる ($\mathcal{A}(x \Rightarrow y) \geqq 0$)」、「(きっと) Y であろう ($\mathcal{A}(x \Rightarrow y) > 0$)」「$Y$ にちがいない ($\mathcal{A}(x \Rightarrow y) \approx= 1$)[15]」「$Y$ である ($\mathcal{A}(x \Rightarrow y) = 1$)」といった表現の区分が可能である。こう

[15] ≈ 1 は 1 に極めて近いということを意味する。

した表現は、同時に、聞き手にとってそれ相応の想定確信度を設定せよという指令として解釈される。

3.6 関連性理論から見た義務性と認識性

さらに、日常言語におけるモダリティでは、義務性 (deontic) と認識性 (epistemic) の違いも重要となる。両者の最大の違いは、義務性が基本的に認知主体の外部から与えられる強制力であるのに対し、認識性は自発的な思考・認識過程であるという点に求められる。しあがって、日本語でも英語でも、義務性と認識性は認知環境の計算と設定に関して、その過程が異なる。義務的言明が与えられた時には、その言明を満たす想定確信度になるよう、強制的に認知環境を変更し、設定しなおさなければならない。また、その談話場面が続く限り、認知環境の更新が行わるごとに、義務的言明に違反しない認知環境になっていることを確認しなければならない。これに対し、認識的言明は、ある時点での認知環境から計算される想定確信度に従って、そういう命題が生成されるのであって、認知主体の内部における認知環境の強制的な変更力は持たない。

日本語では、さらに義務性と認識性のモダリティは形態的な違いを持つ。英語では義務性であっても認識性であっても、'*must*' や '*may*' という同一の助動詞を使いうるが、日本語では「〜でなければならない」と「〜にちがいない」という形で、義務的必然性と認識的必然性の区別が行われる。可能性に関しても、義務性と認識性とで、「〜してもよい」と「〜かもしれない」という形態上の区別をつけなければならない。

「(X は)Y でなければならない」という義務的必然性は、「Y でない (Y の否定) ＋れば (条件) ＋なる (X の成立) ＋ない (偽)」という形態素からできているため、論理表現でいうと、¬Y→¬X と解釈される。これは X→Y とトートロジーである。したがって、「未成年 (X) は禁酒 (Y) しなければならない」という言明は、「未成年 (X) なら禁酒 (Y)」と解釈できる。このように義務的必然性は論理のみで十分解釈可能である。このもちろん、関連性理論に従った語用論上の解釈もできる。言語表現から直接得られる論理式は、¬Y→¬X であるため、関連性の計算は、「Y でない」という情報の「X でない」という情報に対する関連性が存在する (0 以上である) という演算に等しい。すなわち、$\mathcal{A}(\bar{y} \Rightarrow \bar{x}) \geqq 0$ を満たすような認知環境の設定が強制される。ここで、$\mathcal{A}(\bar{y} \Rightarrow \bar{x}) = \dfrac{\mathcal{A}(\overline{xy})}{\mathcal{A}(\bar{y})} - k_y \cdot \dfrac{\mathcal{A}(\bar{x}y)}{\mathcal{A}(y)} = \dfrac{\mathcal{A}(\overline{xy})}{\mathcal{A}(x\bar{y}) + \mathcal{A}(\overline{xy})} - k_y \cdot \dfrac{\mathcal{A}(\bar{x}y)}{\mathcal{A}(xy) + \mathcal{A}(\bar{x}y)}$ であり、さらに $0 \leqq \dfrac{\mathcal{A}(\bar{x}y)}{\mathcal{A}(y)} \leqq 1$ より、$\mathcal{A}(\bar{y} \Rightarrow \bar{x}) \geqq 0$ を常に満たすには、$\dfrac{\mathcal{A}(\overline{xy})}{\mathcal{A}(\bar{y})} = \dfrac{\mathcal{A}(\overline{xy})}{\mathcal{A}(x\bar{y}) + \mathcal{A}(\overline{xy})} = 1$ であればよい。したがって、$\mathcal{A}(\overline{xy}) > 0,\ \mathcal{A}(x\bar{y}) = 0$ が導出される。すなわち、「(X は)Y でなければならない」という言明は、「X の否定情報と Y の否定情報は両立できる」「X と Y の否定情報との両立はない」ということになる。例えば、「未成年 (X) は禁酒 (Y) しなければならない」という言明によって、「(未成年の否定情報である) 成人と (禁酒の否定情報である) 飲酒は両立する」「未成年と (禁酒の否定情報である) 飲酒との両立はない」という認知環境の設定が強制される。これは、前述した論理の解釈とほぼ一致する。また、この関連性理論を使った計算過程では、「Y でなければならない」という表現において、「Y でな

い場合」の区別のみが明確に計算されている ($\mathcal{A}(\overline{xy})$ と $\mathcal{A}(x\overline{y})$ の値のみが明示的に計算されている) という点にも注意されたい。例えば、「禁酒しなければならない」という表現では、「禁酒でない場合＝飲酒」の制限事項が計算されており、「禁酒」に関しては未成年であろうと成年であろうと関知しない、ということである。

一方、「Y に違いない」という認識的必然性は、「(前提となる X は) Y に・違い・ない」という形態素から、「(前提となる)X について、Y と異なるという命題が成立しない」と解釈される。まず、この「異なる」という意味が、指示対象の相違ということであれば、これは「¬(X→Y) が偽であることはない」という論理解釈が可能である。すなわち、X→Y が真になる条件を求めればよい。また、もし「異なり」が指示対象の相違という厳密なものではなければ、語用論上の計算を働かせる必要がある。今、「X について Y と (何らかの形で) 異なる」という命題の成立度は、関連性を用いて、$\mathcal{A}(x \Rightarrow y) = 1-d$ と計算できる。d は食い違い度を表す 0 以上 1 以下の数値で、この数値が大きければ大きいほど「違い」が大きい。「違いない」という表現は、この違い d がなければよい (すなわち $d \approx 0$) ということなので、「(X は)Y に違いない」という言明は、前述したように、$\mathcal{A}(x \Rightarrow y) \approx 1$ という計算であるといってよい。ここで、$\mathcal{A}(x \Rightarrow y) = \dfrac{\mathcal{A}(xy)}{\mathcal{A}(x)} - k_{\overline{x}} \cdot \dfrac{\mathcal{A}(\overline{x}y)}{\mathcal{A}(\overline{x})} = \dfrac{\mathcal{A}(xy)}{\mathcal{A}(xy) + \mathcal{A}(x\overline{y})} - k_{\overline{x}} \cdot \dfrac{\mathcal{A}(\overline{x}y)}{\mathcal{A}(\overline{x}y) + \mathcal{A}(\overline{xy})}$ より、$\mathcal{A}(x \Rightarrow y) \leqq 1$ を常に満たすには、$\dfrac{\mathcal{A}(xy)}{\mathcal{A}(x)} = \dfrac{\mathcal{A}(xy)}{\mathcal{A}(xy) + \mathcal{A}(x\overline{y})}$ がほぼ 1 であり、$k_{\overline{y}} \cdot \dfrac{\mathcal{A}(\overline{x}y)}{\mathcal{A}(\overline{x})} = k_{\overline{x}} \cdot \dfrac{\mathcal{A}(\overline{x}y)}{\mathcal{A}(\overline{x}y) + \mathcal{A}(\overline{xy})}$ がほぼ 0 であればよい。前者の条件からは $\mathcal{A}(xy) \neq 0,\ \mathcal{A}(x\overline{y}) \approx 0$ が導出される。すなわち、「X と Y の両立はあり得る」と同時に、「X と Y の否定情報との両立はほぼあり得ない」ということである。また、後者の条件からは、$k_{\overline{x}} \approx 0$ であるか、あるいは $\mathcal{A}(\overline{xy}) \neq 0,\ \mathcal{A}(\overline{x}y) \approx 0$ が導出される。すなわち、前提 Y の否定状況をほぼ全く考えないようにするか、あるいは前提 Y の否定状況を考慮する限り、「¬X と¬Y の両立はあり得る」と同時に、「¬X と Y の両立はほぼあり得ない」ということが満たされなければならない。

この「(X は)Y に違いない」という語用論上の計算で重要な点は、まず安定して計算される想定は、前提情報 X に関する情報—「X と Y の両立はあり得る」と「X と Y の否定情報との両立はほぼあり得ない」という情報—であるというところにある。「Y でなければならない」という表現の関連性計算において、「Y の否定情報」に関する計算が明示的に行われていたこととは極めて対照的である。また、X の否定情報については、考慮の対象にしてもしなくてもよいが、考慮した場合に限り、「¬X と¬Y の両立はあり得る」と同時に、「¬X と Y の両立はほぼあり得ない」という想定まで持つ必要があり、X→Y という含意ではなく、X↔Y という同値に近いものとして、X の性質が認識されているという点にも注意されたい。

以上の議論をまとめてみよう。まず、「(X は)Y でなければならない」という義務的必然性と「(X は)Y に違いない」という認識的必然性の共通点は以下の通りである。

(30) a. 論理的には、どちらも X→Y とトートロジーを成す表現である。

b. したがって、情報 X の否定と情報 Y は両立しない。

しかし、関連性という語用論の計算では、義務的必然性と認識的必然性はその性質が大きく異なる。まず、認知環境の設定に関して、

(31) 義務的必然性は、認知環境の変更に関する外部からの強制力を表すのに対し、認識的必然性は認知環境の自発的な状態計算である。

という違いがある。また、関連性の計算に関しては、

(32) 義務的必然性は $\mathcal{A}(\overline{y} \Rightarrow \overline{x}) \geqq 0$ の関連性計算であり、認識的必然性は $\mathcal{A}(x \Rightarrow y) \approx 1$ という関連性計算である。

という違いがある。また、この関連性の違いから、「安定して計算される情報」について、

(33) a. 義務的必然性では、「Y でない場合」の区別—「X の否定情報と Y の否定情報は両立できる」「X と Y の否定情報との両立はない」—が計算される。

b. 認識的必然性では、「X である場合」の区別—「X と Y の両立はあり得る」「X と Y の否定情報との両立はほぼあり得ない」—が計算される。

という違いも発生する。また、以下のように派生的に計算される内容についての違いも生じ、そのことが論理との食い違いを生む原因にもなる。

(34) a. 義務的必然性では、「Y である場合」の区別を行っても行わなくても、安定して、論理 (実質含意) とほぼ同一の結果を得る。

b. 認識的必然性では、「X でない場合」を考慮に入れない時と入れた時とでは、得られる情報に違いが生じ、前者の場合は論理的な実質含意に、後者の場合は同値計算に近くなる。

日本語のモダリティにおける義務性と認識性の語用論上の違いは、義務的可能性／認識的可能性の計算においても生じる。まず、義務的可能性から見てみよう。「(X は)Y であってよい」という言明は、形態的に「X は Y であるということが *'positive'* である」と解釈できるため、$\mathcal{A}(x \Rightarrow y) > 0$ を満たすような認知環境を構成すればよい。ここで、$\mathcal{A}(x \Rightarrow y) = \dfrac{\mathcal{A}(xy)}{\mathcal{A}(x)} - k_{\overline{x}} \cdot \dfrac{\mathcal{A}(\overline{x}y)}{\mathcal{A}(\overline{x})} = \dfrac{\mathcal{A}(xy)}{\mathcal{A}(xy) + \mathcal{A}(x\overline{y})} - k_{\overline{x}} \cdot \dfrac{\mathcal{A}(\overline{x}y)}{\mathcal{A}(\overline{x}y) + \mathcal{A}(\overline{xy})}$ より、$\mathcal{A}(x \Rightarrow y) \geqq 0$ を常に満たすには、

(35) a. $\mathcal{A}(xy) > 0$ かつ $k_{\overline{x}} = 0$

b. $\mathcal{A}(xy) > 0$ かつ $\mathcal{A}(\overline{x}y) = 0$

のいずれかであればよい。例えば、「成人 (X) は飲酒 (Y) してもよい」という文であれば、以下のいずれかの解釈が行われることになる。X が成立する条件のみを考えた場合には、「許可」の解釈であるのに対し、X が成立しない条件を考慮した場合には、「X でなくかつ Y であることは認められない」という禁止の解釈が生じる点に注意されたい。

(36) a. 成人で飲酒することは認められており ($\mathcal{A}(xy) > 0$)、未成年に関しては何も言及しない ($k_{\bar{x}} = 0$)。

b. 「成人で飲酒すること」は認められており ($\mathcal{A}(xy) > 0$)、「未成年で飲酒すること」は認められない ($\mathcal{A}(\bar{x}y) = 0$)。

これに対し、認識的可能性「(X は)Y かもしれない」という表現は、X の Y に対する関連性を計算してみたところ、その関連性の想定確信度は「知られていない (決められない)」という言明だと考えられる。言い換えるなら、認識的可能性は、情報 X を前提想定として、情報 Y を焦点情報として認知環境に組み込む機能を持つが、想定の確信度に関しては特別な制限を設けないということになる。したがって、認知環境には $x \wedge y$ という想定だけが組み込まれ、$pXY \geqq 0$ という可能性に関する想定確信度が設定される。

なお、義務性・認識性の違いが形態的に表現される日本語と異なり、英語では両者を同一の助動詞によって表現することができる。こうした英語の最も基本的な性質は以下のように捉えうる。[16]

(37) a. 義務的必然性と認識的必然性の論理表現は X→Y で同一である。

b. “(X)mustY” の関連性計算は、$\mathcal{A}(x \Rightarrow y) \approx 1$ によって行われる。

c. 義務的必然性は、バイアス係数 $k_{\bar{x}} = 0$ の元で $\mathcal{A}(x \Rightarrow y) \approx 1$ が計算されればよいのに対し、認識的必然性は $k_{\bar{x}} \leqq 0$ を満たす適切なバイアス係数の元で $\mathcal{A}(x \Rightarrow y) \approx 1$ を満たさなければならない。3.4 節で見たとおり、$\mathcal{A}(x \Rightarrow y)$ の演算はバイアス係数が 0 に近いほど含意計算と同等の性質を持ち、1 に近いほど同値計算と同等の性質を持つ。この結果、(34) と同様の違いが語用論的に生じる。

d. 同様に、義務的可能性は、バイアス係数 $k_{\bar{x}} = 0$ の元で $\mathcal{A}(x \Rightarrow y) > 0$ を満たせばよいのに対し、認識的必然性は $k_{\bar{x}} \leqq 0$ を満たす適切なバイアス係数の元で $\mathcal{A}(x \Rightarrow y) > 0$ を満たさなければならない。

3.7 計算論的関連性理論から見た知識

以上の議論を簡単にまとめると、次のようになる。これを、戸田山 (2002) の提案である「(19d)：命題の真偽を中心概念としないこと」に対する、関連性理論からの一つの答としたい。

(38) a. 関連性に基づく表示を基本とし、想定確信度を真理値の見込みとして利用することにより、一種の多値論理体系 RPL を作ることができる。

b. 各種論理子は、関連性に基づく演算により定義できる。

c. RPL における想定確信度を 0 か 1 のいずれかのみに限定し、かつ、バイアス係数を適切にコントロールすることで、標準的な命題論理と等価な結果が得られる。

[16] もちろん、英語の場合、この性質以外に、動詞の状態性、否定のスコープなどの影響も考慮しなければならない。この点に関しては、また稿を改めて議論を行う。

d. 情報 X の必然性は、義務的必然性にせよ認識的必然性にせよ、関連性の計算の一種として捉えうる。

e. 想定確信度から、エントロピーの計算を含む一定の手続きによって、真理値の必要条件を満たす判断価値を定義することもできるが、想定確信度そのものを真理値と見込んだほうが、より直接的で、かつ整合性のある体系を構築できる。

ここで、再び「知識」とは何かという問題に立ち返ってみたい。伝統的な真理の定義である「正当化された真なる信念」とは (18) のようなものであり、この定義の不備に対する戸田山 (2002) の提案は (19) のようなものであった。

(18) a. S は X を信じている。
 b. X を信じている S の信念は真である。
 c. S の信念は正当化されている。

(19) a. 自然化された認識論：知識獲得の科学的・心理学的な獲得手段を決定すること。
 b. 社会化された認識論：個人ではなく集団として何かを知っているという点にまで知識を拡張すること。
 c. 知識の基盤を信念ではなく、情報に置くこと。
 d. 命題の真偽を中心概念としないこと。

これを関連性理論の立場から再度、定義し直してみよう。まず、(18) における S を一人の認知主体としよう。このままでは (19b) に反するが、これを認知主体の構成する認知環境と考え、その中に認知主体の直接的なアクセスが可能である D-領域と、間接的なアクセス領域である I-領域 という性質をもうけることで、(19b) に一応答えることができる。ただし、一時的な認知環境は、認知主体が何かを忘却してしまったり、あるいは外部の知識にアクセスする手段がなかったりするため、不完全なものになってしまう。したがって、S を「多様な認知セッションにおいて認知主体が構成する認知環境の集合」としよう (以後、単に「認知環境の集合」と呼ぶ)。認知環境の集合は、認知環境の更新履歴と言い換えてもよい。

次に、(19c) を見てみよう。前述した通り、認知環境における各想定とは、認知主体がアクセスできる情報のことであり、何らかの形で想定確信度の付与ができる情報であればよい。極端にいえば、ある神経発火が起こり、その神経発火の量を適切に取り出せるのであれば、それだけも想定となり得る。もちろん、想定とは「認知主体がアクセスできる」情報という制限が付くため、(19c) を完全に満たすものではないが、しかし妥当な制限であるとも思われる。

(19d) については、(38) で見たとおりであり、関連性に基づく表現と想定確信度を中心概念においても、標準的な論理を含む様々な表象の操作が可能となる。そこで、ここまでのところ、(18a) を「認知環境の集合に、想定確信度が 0 以上となる想定 x がある」と

いう形に、(18b) を「x の想定確信度は 1 であり、否定想定 $\overline{x}$ の確信度は 0 である」という形に言い換えておくことにしよう。

残る問題は、(18c) と (19a) である。ここで、前節で議論した認識論的必然性の性質を振り返ってみよう。「(Y は)X に違いない」という認識的必然性が成立するのは、関連性の計算 $\mathcal{A}(x \Rightarrow y) \approx 1$ を満たす必要がある。すなわち、

(39) a. $$\frac{\mathcal{A}(xy)}{\mathcal{A}(x)} = \frac{\mathcal{A}(xy)}{\mathcal{A}(xy) + \mathcal{A}(x\overline{y})} \approx 1$$

b. $$k_{\overline{x}} \cdot \frac{\mathcal{A}(\overline{x}y)}{\mathcal{A}(\overline{x})} = k_{\overline{x}} \cdot \frac{\mathcal{A}(\overline{x}y)}{\mathcal{A}(\overline{x}y) + \mathcal{A}(\overline{xy})} \approx 0$$

が満たされなければならない。(39a) から、$\mathcal{A}(xy) \neq 0,\ \mathcal{A}(x\overline{y}) \approx 0$—すなわち、「X と Y の両立はあり得る」と同時に、「X と Y の否定情報との両立はほぼあり得ない」ということが導出される。情報 X を固定した場合における (18a), (18b) の性質といってもよい。

一方、(39b) からは、$k_{\overline{x}} \approx 0$, あるいは $\mathcal{A}(\overline{xy}) \neq 0,\ \mathcal{A}(\overline{x}y) \approx 0$ が導出される。前述したように、$k_{\overline{x}} \approx 0$ という性質は、前件否定の状況をほぼ全く考慮しないことを意味する。残りの $\mathcal{A}(\overline{xy}) \neq 0,\ \mathcal{A}(\overline{x}y) \approx 0$ という条件は、否定情報 ¬X における (18a), (18b) の性質である。すなわち、前提 X の否定状況を考慮して、「¬X と¬Y の両立はあり得る」と同時に、「¬X と Y の両立はほぼあり得ない」ということが満たされなければならない。

認識的必然性におけるこの 2 つの性質が、(18c) と (19a) に対する一つの答である。まず、「X ならば Y である」が知識となるためには、認知環境の集合 (すなわち認知環境の更新履歴) の中で、認識的必然性を満たさなければならない。またその時、前提情報 Y とともに、前提の否定情報 ¬X も考慮することが望ましい。否定情報を考慮しない場合は、考慮していないことを明示できていることが望ましい。また、もしも前提の否定情報 ¬X まで考慮した場合には、「¬X–Y と ¬X–¬Y が各々両立可能か不可能か」を検証しなければならない。これが「正当化」に関する「知識獲得の科学的・心理学的な獲得手段」であるとしよう。

以上が、計算論的関連性理論の立場からみた「正当化された真なる想定としての知識」の性質である。まとめると、知識とは以下のような性質を持つ想定のことであると考えられる。

(40) ある妥当なコンテクスト C において、以下の条件を満たしたとき、「(X は)Y である」という想定を知識と見なせる。

a. **コンテクスト C の範囲：** I-領域・D-領域を含めた認知環境の集合において、

b. **正当化の手段：** 前提情報 X を十全に考慮すると同時に、その否定情報 ¬X については、考慮していないことを明確にするか、考慮するなら十全に考慮した上で、

c. **真であるという見込み：** X から Y に対する関連性の想定確信度が、認知環境の更新履歴を通じて一貫してほぼ 1 であること。すなわち一貫した認識的必然性を十分に満たすこと。

4. 認知過程における関連性の影響

前節では、「知識」といえるものに、いかに関連性が深く関与しているかを見た。また、認知環境における想定は、例えば$y \Rightarrow x$のような関連性の形でもって表示されているのが基本であることも見た。単独想定や、連言や選言を含んだ想定などは、多くの場合、関連性の想定から計算されるものと考えられる。本節ではこうした関連性の計算過程が、実際の認知活動の中でどのように影響しているのかという点について簡単に触れ、関連性計算の性質について検討してみたい。

4.1 顕在性と構造確率の影響：色選択課題

これまでの様々な心理学的実験で、人間はいくつかの認知的錯誤を起こすことが明らかにされてきた。その認知的錯誤が起こる原因を探ると、逆に人間の認知処理の特性がよく見えてくる。中には、数学的には間違いであるにせよ、そのエラーにはある主の妥当性が認められることも例もある。この節では、こうした認知的錯誤のいくつかを取り上げ、それが関連性理論においてどのように捉えられるものであるのかという点について考えてみたい。まず、最も簡単な例の一つとして、次のような「カード色選択課題」から見てみよう。

> 机の上に3枚のカードがある。一枚は両面赤、一枚は両面青、一枚は片面赤でもう片面が青である。今、一枚取り出したところ、赤いカードだった。さて、このカードの裏面は赤か青か？賭けるとしたらどっちが得だろうか？

論理的には、次のような解になる。今、両面が赤であるカード(これをカードAとしよう)の片面をA1, その裏側をA2とする。同様に、両面青のカード(B)の片面をB1, もう片面をB2, 片面赤のカード(C)の赤側をC1, 青側をC2とする。今、赤い面が出ているので、可能性はA1, A2, C1のいずれかである。したがって、その反対側の面は、A2(赤), A1(赤), C2(青)のいずれかであり、2/3の確率でカードの裏面は赤—つまり赤いほうに賭けるほうが特であると考えられる。

しかし、この問題について、かなりの人が次のように考える。まず、赤いカードが出ているので、このカードはAかCかのいずれかである。Aだとすると裏面は赤、Cだとすると裏面は青である。したがって、赤が出るか青が出るかは1/2の確率で同じである。

この「色選択課題」の誤りの原因は、顕在的な情報が「カードの各面」というカードの持っている一つの性質ではなく、「カードそのもの」に設定され、またそれ故に、頻度確率ではなく、構造確率の計算が行われているところにある。これは数学的にはエラーであるが、しかし認知主体が事物の性質ではなく、事物そのものに注意を向けるのには、それなりの妥当性もあると考えられる。今、「カードの面」という性質に注目すると、A1とB1は共に「赤」という性質を持っており、そこに違いがない。同様に、B2とC2も「青」であり、違いがない。一方、「カードそのもの」に注意を向けると、各カードは互いに完全に違った性質を持つ。このことは、認知主体にとって、どういう情報が顕在的になりやすいかを端的に示している—すなわち、「違いが明確になる」情報ほど顕在性を持

ちやすいということだ。これは認知活動の一般的な性質でもある。知覚でいうなら、感覚が捉えた情報は、まず違いを抽出する微分系で処理される。視覚なら、感覚が受容した表面情報の違いを抽出することで、「エッジ」の検出が行われる。聴覚の場合は、直前の音信号との違いをまず抽出することで、「音イベント」の検出が行われる。言語でも、その本質は「恣意的な分節化」(Saussure, 1916) にあるが、その恣意性はなるべく差異の体系が強くなければならないという制約を受ける。

関連性の表現式が単なる $\mathcal{A}(x|y)$ ではなく、式 (10) を考えた方がよい理由もここにある。$\mathcal{A}(x \Rightarrow y) = k_x \cdot \dfrac{\mathcal{A}(xy)}{\mathcal{A}(x)} - k_{\overline{x}} \cdot \dfrac{\mathcal{A}(\overline{x}y)}{\mathcal{A}(\overline{x})}$ という式は、想定 y の関連性が高くなるためには、その前提となる想定 x と $\overline{x}$ との間でのコントラストが必要となることを意味しており、「違いが明確になる」情報ほど顕在性を持ちやすいという性質を持っている。これは認知システムの基本的な特徴を表現できているといってよいだろう。

4.2 リンダ問題

「リンダ問題」は、人が論理や集合の連言に関する錯誤を起こすこと、特に確率の conjunction rule である $P(X\&Y) \leqq P(X)$ をうまく判断することができないことを示したものである。リンダ問題には様々なバージョンがあるが、よく知られている例は次のようなものである。

(41) リンダは、今年 31 歳になる女性である。彼女は独身で非常に知的で、はっきりものを言う。大学時代は哲学を学び、学生の頃は社会主義と差別問題に関する活動に深く関わっており、核兵器反対のデモにも参加していた。さて、あなたが、この女性の今を推測する場合、次の二つの選択肢のうち、どちらの可能性が高いと考えられるか。

(a) 彼女は銀行員である。
(b) 彼女は銀行員で、女性運動家としても活動している。

数学的には、「銀行員で女性運動家」という集合は、「銀行員」という集合に完全に包含されるため、全女性を母数とした割合を考えると、「銀行員で女性運動家」の確率は「銀行員」の確率よりも必ず低くなるはずである。したがって、論理的な正解は (a) となる。しかし、多くの人が (b) を選択してしまう。

この問題を最初に提出した Kahneman and Tversky (1982) は、このエラーの原因を一種の認知バイアスである代表性ヒューリスティクスの結果であると言っている。このヒューリスティクスのため、被験者は集合関係についての論理的考慮を無視し、リンダについての記述との類似性という観点から選択肢を順序づけてしまうことになる。また、この問題は (a) の選択肢を、(b) との対比から「彼女は銀行員だが、女性運動家ではない」と理解してしまうために起こるエラーであるという指摘もなされている。

関連性理論の立場から見ても、この問題は大変に興味深いものである。人が何かを判断する時に、如何に情報間の関係性が重要であるかを教えてくれる例だからである。リ

ンダ問題における論理的な解は、何の事前情報もなく、単に「リンダという女性がいる。彼女は「銀行員」か「銀行員で女性運動家」かどちらの可能性が高いか」を判断させるのと同一である。すなわち、そもそも「知的ではっきりものが言えて哲学を学んで社会活動をしていた」という情報を考慮する必要がない。

しかし、認知主体は常に「関連性」を求める存在である。「知的ではっきりものが言えて哲学を学んで社会活動をしていた女性」という情報 (これを前提情報 X としよう) が与えられてしまうと、情報 X と「銀行員」との関連性、情報 X と「銀行員かつ女性運動家」の関連性を探し出す。このとき、関連性の計算は次のようになる。なお、$\overline{X}$ は「哲学や社会活動などに興味のない女性」を指す。

(42) a. $\mathcal{A}(X \Rightarrow \text{銀行員}) = \mathcal{A}(\text{銀行員} \mid X) - \mathcal{A}(\text{銀行員} \mid \overline{X})$

b. $\mathcal{A}(X \Rightarrow \text{銀行員} \wedge \text{女性運動家}) =$

$\mathcal{A}(\text{銀行員} \wedge \text{女性運動家} \mid X) - \mathcal{A}(\text{銀行員} \wedge \text{女性運動家} \mid \overline{X})$

(42a) において、「哲学や社会活動に興味を持っている人の中で、銀行員であるような人：$\mathcal{A}(\text{銀行員} \mid X)$」と「哲学や社会活動に興味がない人の中で、銀行員であるような人：$\mathcal{A}(\text{銀行員} \mid \overline{X})$」の数値がほとんど同じであろう。したがって、「銀行員」という情報に対する「哲学や社会活動に興味を持っている人」という情報の関連性は大変に低い。一方、(42b) における「哲学や社会活動に興味を持っている人の中で、銀行員かつ女性運動家である人：$\mathcal{A}(\text{銀行員} \wedge \text{女性運動家} \mid X)$」の確信度は、「哲学や社会活動に興味がない人の中で、銀行員かつ女性運動家である人：$\mathcal{A}(\text{銀行員} \wedge \text{女性運動家} \mid \overline{X})$」の確信度よりもかなり高くなるであろう。したがって、(42b) の関連性確信度はかなり高く計算される。この関連性の違いを生み出しているのは、後件情報である「銀行員」「銀行員かつ女性運動家」という情報なので、「リンダは銀行員かつ女性運動家」であろうという結論を引き出す。

4.3 演繹推論・帰納推論・仮説推論

リンダ問題は日常的によく見かける推論問題の一種である。よく知られているように、推論には演繹推論 (deduction)・帰納推論 (induction)・仮説推論 (abduction) の 3 種があり、それぞれ以下のような性質を持つ。

(43) a. 演繹推論：$X \rightarrow Y$ が真である時、$X \rightarrow Y$ と X から Y を推論する。あるいは $X \rightarrow Y$ と $\neg Y$ から $\neg X$ を推論する。

b. 帰納推論：事例 X と Y から、$X \rightarrow Y$ を推論する。あるいは個々の想定の類似性から一般化された想定を導く。

c. 仮説推論：事例 Y から $X \rightarrow Y$ となる条件をいくつか設定し、その中で最も妥当な条件を使って、X を推論する。

論理学に基づく場合、演繹推論は常に真であることが保証されているのに対し、帰納推論・仮説推論は真になることが保証されていない。そのため、演繹推論の優位性がしばしば強調されてきた。しかし、関連性理論に基づいた擬似的な論理 (3.4 節を参照) では、

想定確信度が連続的な量を取るため、演繹推論の結果も一種の可能性解釈となり、絶対的なものとはいえない。それ故、以下に述べる通り、この3つの推論形式は互いに絡み合い、互いに影響を及ぼす推論となる。

例えば今、後件 Y を説明する条件が、$X_1 \to Y, X_2 \to Y, \cdots, X_n \to Y$ のように、複数個あるとする。一般的な論理学では、$X_1, X_2, \cdots, X_n 1$ のいずれもが真であるならば、結論 Y が真であることも保証され、複数の条件間に優位性の違いは生じない。しかし、計算論的関連性理論のような連続量を持ち込んだ場合には、話が違ってくる。もし、$\mathcal{A}(X_1 \Rightarrow Y)$, $\mathcal{A}(X_2 \Rightarrow Y), \cdots, \mathcal{A}(X_n \Rightarrow Y)$ の関連性確信度が異なっているならば、$X_1, X_2, \cdots, X_n$ の確信度が全て1であったとしても、結論 Y の想定確信度は異なる。仮説推論でも同様に $\mathcal{A}(X_1 \Rightarrow Y)$, $\mathcal{A}(X_2 \Rightarrow Y), \cdots, \mathcal{A}(X_n \Rightarrow Y)$ の関連性確信度が同一であっても、$X_1, X_2, \cdots, X_n$ の確信度が異なれば結論 Y の想定確信度は違ってくる。計算論的関連性理論における推論とは、一つの前提情報 X と、X を前提条件に持つ様々な関連性計算のうち、最も確信度の高い $\mathcal{A}(X \Rightarrow Y)$ を用いて焦点情報 Y を得る過程といってよい。そして、この演繹演算が複数あるという性質が、結果的に帰納推論や仮説推論を引き起こす。すなわち、$\{X_1, Y\}, \{X_2, Y\} \cdots, \{X_n, Y\}$ のペアから考えられる条件 $\mathcal{A}(X_1 \Rightarrow Y), \mathcal{A}(X_2 \Rightarrow Y), \cdots, \mathcal{A}(X_n \Rightarrow Y)$ のうち、最も関連性確信度の高いものを選択する過程が帰納推論である。同様に、$\{X_1, Y\}, \{X_2, Y\} \cdots, \{X_n, Y\}$ のペアから考えられる条件 $\mathcal{A}(X_1 \Rightarrow Y), \mathcal{A}(X_2 \Rightarrow Y), \cdots, \mathcal{A}(X_n \Rightarrow Y)$ のうち、最も関連性確信度の高いもの引き起こす X を選び出すのが仮説推論となるのである。計算論的関連性理論の元では、この3種の推論の中心には常に関連性の計算があり、関連性の高いものに注目するという点で同一の性質を持つ。

4.4 認知的浮動

リンダ問題の難しさは、この問題を解く基準となる関連性の計算を、$\mathcal{A}(X \Rightarrow$ 銀行員$)$ として行うのか、あるいは $\mathcal{A}($銀行員 $\Rightarrow X)$ として行うのかが明確ではない点にも原因がある。この点について、中垣 (2007) は、人間に事象の確率判断を行わせた時、連言確率 $P(X\&Y)$ と条件付き確率 $P(Y|X), P(X|Y)$ の違いがほとんど区別されないことを指摘し、それを**認知的浮動** (*cognitive fluctuation*) と呼んだ。この認知的浮動という現象について、中垣は以下の2点を主張している。

(44) a. 認知的浮動は $P(Y|X) \to P(X\&Y) \to P(X|Y)$ というプロセスを経て引き起こされる。また、このプロセスは可逆性を持つ (すなわち、$P(X|Y) \to P(Y\&X) \to P(Y|X)$ というプロセスも起こる)。

b. 認知的浮動はまた事象Aと事象Bの関連性が高くなければ起こらない。

認知的浮動という現象は、これまでいくつもの実験で主張されてきた人間の確率判断錯誤の現象をよく説明するモデルであり、極めて興味深い。本節では、このモデルを計算論的関連性理論の立場から検討してみよ。まず、(44a) から、人間の確率判断が $P(X\&Y)$ という連言確率の状態を好まず、$P(Y|X)$ あるいは $P(X|Y)$ という条件付き確率を好むという特徴を持つことが伺える。これは心理・行動という観点から見て、興味深く、また

妥当性を持った特徴である。今、探索課題 (行動レベルでも記憶レベルでもよい) の例として、「図書館で、洋書で (A) かつ心理学について書かれている本 (B) を見つけてください」という命令を実行する時を考えてみよう。これは A&B という連言から成る探索課題であるが、図書館の本を片っ端から探して、A&B という性質を満たす本をピックアップする人はまずいない。「心理学のコーナーに行って、その中で洋書を探す」かあるいは「洋書コーナーに行って、心理学の本を探す」という行動を取る。すなわち、効率的に探索を行うには、ある情報を手がかりにして探索範囲をおおまかに絞り込み、その中でさらに詳細な検索を行うことが必要となる。

確率判断を行うときに問題となるのは、この「最初の探索範囲」である。$P(A\&B)$ の値を正確に知るには、「心理学の洋書」の冊数を、「図書館全体の本」の冊数で割り込む必要がある。これは、探索という点からは、極めて効率が悪い。そこで、まず洋書のコーナーに行くわけであるが、ここで、全体の母数を図書館全体の本ではなく、「洋書コーナーに置いてある本 (A)」に設定してしまうと、$P(B|A)$ という条件付き確率を求めることになってしまう。同様に、「心理学のコーナーに置いてある本 (B)」から「洋書 (A)」をを探す時には、ついつい $P(A|B)$ という条件付き確率で考えてしまう。もちろん、$P(A\&B) = P(A|B) \cdot P(B) = P(B|A) \cdot P(A)$ より、条件付き確率に、母数自体の確率 (図書館全体における洋書の割合や心理学の本の割合) を掛け込むと、本当に得たい P(A&B) を知ることができるのだが、「心理学コーナー」に行ったり、「洋書コーナー」に行ったりした時点で、このコーナーだけが「フレーム」となってしまい、「フレーム以外にある本」はアクセス可能ではあるが、活性化されていない状態になる。したがって、この最後のかけ算のプロセスが無視されてしまうのである。このことから分かる通り、(44a) の性質は、「フレーム問題」から引き起こってくるものであり、認知的浮遊が連言で留まらないのは、フレーム問題の計算効率性のためと考えて良い。

また、このプロセスは言語学的に見ても妥当性がある。今、「学生が英語を勉強しています」という文を考えてみよう。これは、中立的な解釈では、「[x が学生という性質を満たしていて、かつ英語を勉強しているという性質も満たす] ような x が存在する」という意味になり、おおよそ $\exists x$[学生 (x)$\wedge$勉強 (x, 英語)] と表現できる。すなわち、連言が関わる解釈であり、確率としては P(学生&英語の勉強) が計算される。ところで、この文は、「英語を勉強しているのは誰ですか」という質問文の答としても機能する。そして、この質問文の答は、「英語を勉強しているのは学生です」というものであってもよい。すなわち、「英語を勉強している」というのは前提情報であり、「学生」というのが焦点情報になるような解釈を受ける。こうした表現の論理形式は、おおよそ、$\forall x$[勉強 (x, 英語)$\rightarrow$学生 (x)] という条件文を含んだものになり、確率としては P(学生 | 英語の勉強) が計算される。[17]。

以上のことから、同じ「X が Y」という文であっても、状況により $P(X\&Y)$ が計算されることもあれば、$P(X|Y)$ が計算されることもあり、両者の混同が起こりやすいのは言語の意味という観点からも必然的であることが分かる。また、$P(X\&Y)$ の計算を生み出

[17] より正確には、$\mathcal{A}$(学生$\Rightarrow$勉強) が計算されていると見るべきであるが、ここでは話を単純化しておく。

す ∃x[学生 (x)∧勉強 (x, 英語)] という論理表現は、学生 (x) と勉強 (x, 勉強) が等価な価値を持っており、焦点情報が 2 つあることになるため、Chafe (1994) の言う新情報の条件、あるいは「焦点情報制約 (9)」に反する。これに対し、P(学生 | 英語の勉強) の元となる ∀x[勉強 (x, 英語)→学生 (x)] という表現は、前提情報が 1 つ、焦点情報が 1 つとなり、Chape の新情報の条件や (9) に一致する。したがって、連言確率よりも条件付き確率のほうが好まれるのは、言語学的に見ても妥当なことなのである。ここで、以上の流れをまとめておこう。[18]

(45) a. $P(Y|X)$ があった時、まず $P(X) = 1$ を仮定する。すなわち、情報 X が全フレームであることを仮定し、情報 $\neg X$ は一切考慮しない。$P(Y)$ は任意の値でよく、フレームの影響は受けない。ここで、$P(Y|X) = P(Y|X) \cdot 1 = P(Y|X) \cdot P(X) = P(X\&Y)$ より、$P(Y|X)$ は $P(X\&Y)$ に浮動する。

b. $P(X\&Y)$ は、焦点情報制約 (9) に反し、また、「A が B」という意味が連言表現と条件表現の曖昧性を持つため、今度は、$P(Y) = 1$, $P(X)$ は任意という条件を満たすフレーム設定を行い、$P(X\&Y) = P(X|Y) \cdot P(Y) = P(X|Y) \cdot 1 = P(X|Y)$ より、$P(X\&Y)$ が $P(Y|X)$ に浮動する。

(45) のポイントは、(45a), (45b) のいずれの過程においても、前提の否定状況を考えない (すなわち、$P(X) = 1$ か $P(Y) = 1$ を仮定する) 点にある。では、しかし、では、この仮定が成立しない場合はどうなるのだろう。ここで、もう認知的浮動のもう一つの条件、すなわち関連性の高さ (44b) という条件について考えてみよう。

今、式 (10) に基づいて、情報 X と情報 Y の関連性を考えてみよう。どちらが前提情報となり、どちらが焦点情報となるかにより、次の 2 通りの計算式が成立する。

(46) a.
$$\mathcal{A}(x \Rightarrow y) = \frac{\mathcal{A}(xy)}{\mathcal{A}(x)} - k_{\bar{x}} \cdot \frac{\mathcal{A}(\bar{x}y)}{\mathcal{A}(\bar{x})} = \frac{\mathcal{A}(xy)}{\mathcal{A}(xy)+\mathcal{A}(x\bar{y})} - k_{\bar{x}} \cdot \frac{\mathcal{A}(\bar{x}y)}{\mathcal{A}(\bar{x}y)+\mathcal{A}(\overline{xy})}$$

b.
$$\mathcal{A}(y \Rightarrow x) = \frac{\mathcal{A}(xy)}{\mathcal{A}(y)} - k_{\bar{y}} \cdot \frac{\mathcal{A}(x\bar{y})}{\mathcal{A}(\bar{y})} = \frac{\mathcal{A}(xy)}{\mathcal{A}(xy)+\mathcal{A}(\bar{x}y)} - k_{\bar{y}} \cdot \frac{\mathcal{A}(x\bar{y})}{\mathcal{A}(x\bar{y})+\mathcal{A}(\overline{xy})}$$

$\mathcal{A}(x \Rightarrow y)$ において、どのようなフレームの大きさであれ (バイアス係数$k_{\bar{x}}$ が任意である時)、関連性が高くなるためには、$\mathcal{A}(x\bar{y})$ と $\mathcal{A}(\bar{x}y)$ の値が極めて小さければよい。すなわち、どのようなフレーム設定が行われているにせよ、「X と$\neg Y$ の両立が見込めず、$\neg X$ と Y の両立も見込めない」という事態は、情報 X から情報 Y への関連性を高くする必要十分条件である。同時に、$\mathcal{A}(x\bar{y})$ と $\mathcal{A}(\bar{x}y)$ の値が極めて小さいときには、$\mathcal{A}(y \Rightarrow x)$ の関連性も常に高く、また $\mathcal{A}(x \Rightarrow y) \approx \mathcal{A}(y \Rightarrow x)$ が見込まれる。ここで認知的浮遊が起こる。

[18] (45) では、$P(Y|X)$ から $P(X|Y)$ への浮動を取り上げるが、$P(X|Y)$ から $P(Y|X)$ への浮動も同様である。

主観確率 $P(Y|X)$ は、$\mathcal{A}(x \Rightarrow y)$ の特別なケースである (バイアス係数 $k_{\bar{x}}$ が極めて小さい)。同様に、主観確率 $P(X|Y)$ は、$\mathcal{A}(y \Rightarrow x)$ の一例である。したがって、関連性の高い状況は、$P(Y|X) \approx P(X|Y)$ を引き起こすことが言える。このプロセスの特徴は、X$ Y という連言表現が明示的に出てこない点にある。すなわち、関連性が高い状況では、$P(Y|X)$ から $P(X|Y)$ への認知的浮動 (あるいは $P(X|Y)$ から $P(Y|X)$ の認知的浮動) は、関連性の計算から直接引き起こされる。

4.5 臨床検査における偽陽性

前述したように、計算論的関連性理論における想定の確信度は、主観確率を元にした条件付き確率計算の一種である。しかし、人間の確率計算能力は、この条件付き確率に対して脆弱であることが知られている。その最も代表的な例として、以下のような「検査薬の偽陽性問題」がある。

> 今、ある非常に深刻な病気 S は発症率が 0.1%であることが知られている。検査薬 T を使うと、この病気 S にかかっている時、99% の確率で検査結果を正しく陽性と診断でき、また病気 S でなかった時は、95%の確率で正しく陰性と判断できる。今、この検査薬 T を使って陽性反応が出た場合、実際に病気 S である確率はどの程度か。

まず、ベイズ則を使って、まず正解を求めておこう。S を病気にかかっている事象、T を陽性反応が出るという事象とした時、求める確率は、以下のように計算できる。

(47)
$$\begin{aligned} P(S|T) &= \frac{P(T|S)P(S)}{P(T|S)P(T) + P(T|\neg S)P(\neg S)} \\ &= \frac{0.99 \times 0.001}{0.99 \times 0.001 + 0.05 \times 0.999} \\ &\approx 0.019 \end{aligned}$$

すなわち、検査薬 T の陽性反応から病気 S と考えられる確率は、わずかに 1.9% ほどである。しかし、多くの人は、90% 以上の確率で病気 S にかかっていると思いこんでしまう。その理由として、しばしば以下の 2 点が指摘されてきた。

(48) a. 事前確率 (この場合は病気 S の発症率が 0.1% であること) を無視してしまう。
  b. 「病気 S なら陽性反応 T」という論理を、論理的にはトートロジーにならない逆の論理「陽性反応 T なら病気 S」と理解されてしまうこと。

(48b) は前節で見た認知的浮動そのものといってよいので、ここでは (48a) について少し考えてみよう。確かに、人は事前確率を適切に取り扱うことができないことがある。しかし、それには一つの理由があるものと考えられる。それは、頻度確率と構造確率の違いに起因するものである。ある認知主体が病気 S の疑いを持った時、この 0.1% という数値はほとんど何の意味も持たないという理由である。0.1% という数値はあくまで頻度確率であって、多くの人間を母数とした場合の統計値に過ぎない。それを「自分」という

個体に当てはめた場合、一生の中で 0.1% 分だけ (80 年の人生とすると、約 55.5 日だけ) 病気 S にかかるということでもない。ほとんどかからない病気ということが数値で示されていても、その一般論が自分に通用するかどうかはまた別問題であり、自分が病気 S にほとんどかからないということを保証するものでもない。すなわち、認知主体が病気 S の疑いを持った時、想定確信度として機能するのは、「病気 S にかかっているのか否か」という**構造確率**なのである。この想定の元では、陽性反応 T から病気 S を疑う確率は、

(49)
$$
\begin{aligned}
P(S|T) &= \frac{P(T|S)P(S)}{P(T|S)P(T)+P(T|\neg S)P(\neg S)} \\
&= \frac{0.99\times 0.5}{0.99\times 0.5+0.05\times 0.5} \\
&\approx 0.952
\end{aligned}
$$

という計算より、95% 以上という数値になり、多くの人が持つ直感的な数値に一致する。また、この偽陽性問題では、病気 S の検査薬 T に対する関連性を簡単に計算でき、その関連性確信度は 0.94 と非常に高いものとなる。

(50) $\mathcal{A}(S \Rightarrow T) = \mathcal{A}(T|S) - \mathcal{A}(T|\overline{S})$ より、
$\mathcal{A}(S \Rightarrow T) = 0.99 - 0.05$ すなわち、病気 S の検査薬 T に対する関連性は 0.94

実際、この関連性の高さは信じるに足る生態学的な妥当性を持っている。何の検査も受けていない人なら、病気 S にかかることはほとんどない (0.1%) と信じることができる。しかし、検査薬 T により陽性反応が出ると、病気 S の確率は数学的に見ても 1.9% 程度に跳ね上がり、その倍率は 19 倍に達する。1.9% という数値は小さく見えるが、しかしそれは全人口を母数とした時の数値である。実際の人数を考えると、0.1% と 1.9% という 20 倍近くの違いは非常に大きなものであり、それだけ重みもある。わたしたちは、1.9% という数値を示された時、その小さな数字の意味を見失いがちである。それよりも、関連性確信度が教えてくれる「危険性」のほうが、生態学的に妥当な意味を持つといってもよい。

認知システムを考える上で難しい問題は、何をもって「正しい」とするかという点にある。本節で見た偽陽性問題において、90% 以上の数値を答とするのは、確かに数学的に正しくない。しかし、その正解の数値である 1.9% という値は、病気の危険性に照らして「正しい」数値とはいえない。前節で見た認知的浮動も同様である。X→Y という推論と、その逆の Y→X という推論は、論理学上ではトートロジーにならない。X→Y と等価な形式は、対偶である¬Y→¬X という推論である。これは確かに正しい。しかし、その「正しさ」が認知主体にとって妥当かどうかは、また別の問題である。次節ではこの点について考察してみよう。

4.6 ヘンペルのカラス

「ヘンペルのカラス」という問題は、「カラスは黒い」ことを対偶によって証明しようとした時に起こる「腑に落ちない感じ」に焦点を当てたものである。まず、「カラスは黒

い」という命題は、その対偶である「黒くないものはカラスでない」とトートロジーになる。したがって、「カラスは黒い」ことを証明するためには、「黒くないものはカラスでない」ことが証明できればよい。こうして世界中の黒くないものを順に調べ、その中にカラスが含まれていない[19]ならば、「カラスは黒い」ということを証明できる。この証明における違和感は、カラスを一羽も調べずに、当該命題である「カラスは黒い」ということが証明されてしまうというところにある。

この証明に違和感が生じる理由の1つとして、現実場面において、「黒くないもの」の数が極めて多く、実際には調べきれないからだと説明されることがある。確かにこれは重要な理由であり、次節で述べる Wason 選択課題においても関与する理由である。しかし、計算論的関連性理論を使うと、この「違和感」に他の理由があることも分かる。今、「カラスが存在する」という情報を X,「色が黒い」という情報を Y とすると、「カラスは黒い」という命題の関連性確信度は、$\mathcal{A}(x \Rightarrow y) = \dfrac{\mathcal{A}(xy)}{\mathcal{A}(xy) + \mathcal{A}(x\overline{y})} - k_{\overline{x}} \cdot \dfrac{\mathcal{A}(\overline{x}y)}{\mathcal{A}(\overline{x}y) + \mathcal{A}(\overline{xy})}$ によって計算される。まず、このことから、「カラス以外の動物で黒いもの」がたくさんあるなら、「カラスは黒い」という命題の関連性確信度は下がってしまう。しかし、ここではこの点については議論せず、「カラス以外の動物」のことを全く考えないことにしよう。この場合は、$\mathcal{A}(x \Rightarrow y) = \dfrac{\mathcal{A}(xy)}{\mathcal{A}(xy) + \mathcal{A}(x\overline{y})}$ のみを計算すればよい。ここで、黒くないもの ($\neg$Y の想定 $\overline{y}$) を全て調べて、もしもその中にカラスがいなければ、$\mathcal{A}(x\overline{y}) = 0$ であることがわかり、$\mathcal{A}(x \Rightarrow y) = \dfrac{\mathcal{A}(xy)}{\mathcal{A}(xy)}$ となる。すなわち、$\mathcal{A}(x \Rightarrow y) = 1$ であり、「カラスは黒い」ということが絶対的に真であることがいえる。しかし、この計算過程には唯一の例外が存在する。それは $\mathcal{A}(xy) = 0$ となる場合である。この時には、$\mathcal{A}(x \Rightarrow y)$ は $\dfrac{0}{0}$ を計算することになり、解は不定となってしまう。ここが、論理と計算論的関連性理論の最大の違いである。

通常の論理[20]では、「カラスが全く存在していない」場合でも、黒くないものを調べて、その中にカラスがいなければ、「カラスは黒い」という命題は真になる。しかし、関連性理論ではそういう結論にならない。一羽でも黒いカラスがいることが保証されているなら (すなわち $\mathcal{A}(x \Rightarrow y) > 0$ が保証されているのなら)、対偶によって命題を証明することが可能である。しかし、いくら対偶命題が成立しても、「黒いカラス」に関する想定が持てないのであれば、真とも偽とも判断できない命題という計算がなされるのである。「ヘンペルのカラス」の証明における違和感の原因はまさにここにあると思われる。たとえ対偶論法を使ったとしても、関連性を求める認知主体にとっては、最低一羽はカラスを調べ、それが黒いことを確認しなければ、「カラスは黒い」という命題は意味をなさないのである。

[19]実際には白いカラスがいるらしい。
[20]直観主義論理学などを除く。

4.7 Wason 選択課題

「ヘンペルのカラス」における論理と関連性の食い違いを最も劇的に示す例が Wason 選択課題である (Wason, 1966)。この課題は、人間が行う推論過程は、論理学でいう含意でも同値でもない中途半端な判断であることを端的に示すものである。例えば次のような例を考えてみよう。

(51) 規則：もし表が “A” なら、裏には “7” が印刷されている

選択肢：$\boxed{\mathrm{A}}$ $\boxed{\mathrm{G}}$ $\boxed{7}$ $\boxed{2}$

規則を含意解釈として捉えるなら、$\boxed{\mathrm{A}}$ 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 $\boxed{\mathrm{A}}$ と $\boxed{7}$ のカードを選択することが知られている。その判断の偏りを生じさせる認知的バイアスとして、確証バイアスやマッチングバイアス、主題化効果などが提案されており、関連性理論に基づく説明も行われている。ここでは、この問題に対する関連性の関与を、式 (10) により形式的に捉えてみよう。

この問題に対して論理推論を使わない被験者は、実際のカードを調べ、ルールとの関連性の高いカードを選択しようとする。まず $\boxed{\mathrm{A}}$ のカードを見ると、このカードはルールの前件を肯定するもので (したがって $k_{\bar{x}} = 0$ と設定できるカードで)、また $\{\mathcal{A}(xy), \mathcal{A}(x\bar{y})\}=\{1,0\}$ あるいは $\{0,1\}$ の可能性があることが分かる。前者の可能性からは規則の想定確信度として $\mathcal{A}(x \Rightarrow y) = 1$ が、また後者の想定からは $\mathcal{A}(x \Rightarrow y) = 0$ が得られ、規則を完全に確定／棄却できる極めて関連性の高いカードであることが分かる。したがって、$\boxed{\mathrm{A}}$ は必ず選択される。

一方、$\boxed{\mathrm{G}}$ のカードは、バイアス係数 $k_{\bar{x}} = 1$ および、$\{\mathcal{A}(\bar{x}y), \mathcal{A}(\overline{xy})\}=\{1,0\}$ あるいは $\{0,1\}$ という可能性を認知主体に提供する情報である。どちらの可能性の場合でも $\mathcal{A}(x \Rightarrow y)$ の値は決定されず、マイナスの数値 (すなわち関連性がなく、認知環境から棄却される) になることもある。すなわち関連性が極めて低く、また認知環境の改善をもたらさないカードである。したがって、このカードは選択されない。

これが $\boxed{7}$ のカードになると、表を見たときに $\boxed{\mathrm{A}}$ であった時には、完全な関連性を持つカードとなる。ただし、表が $\boxed{\mathrm{A}}$ 以外であれば、関連性なしと判断されるカードでもある。したがって、$\boxed{\mathrm{A}}$ のカードほどの選択率ではないにせよ、選ばれやすいカードとなる。そして、最後の $\boxed{2}$ のカードである。論理推論を使うと、このカードは対偶論法に使えるものであり、必ず選ばなければいけないものである。しかし、関連性を使った判断では事情が異なる。このカードでは表が分からないため、バイアス係数は $k_{\bar{x}} > 0$ と設定しなければいけない。ここで、もしこのカードの表面が $\boxed{\mathrm{A}}$ であった時、$\mathcal{A}(xy) = 0$, $\mathcal{A}(x\bar{y}) = 1$ で、その他の想定値は不定であるため、$\mathcal{A}(x \Rightarrow y)$ の数値は $\mathcal{A}(x \Rightarrow y) \leqq 0$ となる。すなわち、単に「誤り」というだけでなく、「誤りか関連性のないカード」という判断になってしまう。また、このカードの表面が $\boxed{\mathrm{A}}$ 以外であったなら、これはルールに対して全く関連性のないカードとなる。したがって、$\boxed{2}$ のカードは基本的に無関係なカードであり、関係があっても「誤り」しか証明しないカードであるため、極めて選択され

にくい。繰り返すが、これは論理推論の立場からは誤った選択である。しかし、「ヘンペルのカラス」で見たように、現実場面ではこうした無関係である可能性のある事象を全て調べ尽くすことは不可能であるし、実際の行動として妥当とはいえないこともある。Wason 選択課題は、人間が不合理な行動をすることを示した実験というよりは、むしろ人間にとって正しく合理的な判断とは何かを示している研究なのである。これは、課題の内容によっては Wason 課題が正しく解かれるということからも分かる。例えば、次のような問題を考えてみよう。

(52) ある工場では、次のような就労規則があります。

- $\boxed{\text{もし休日に労働したなら、平日に休みが取れる。}}$

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 (53) のような方略で解釈することを言い、許可スキーマとは (54) 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。

(53) 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**。

(54) 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。

(55)

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

(55) 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。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。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 (10) により $\mathcal{A}(x \Rightarrow y)$ の値を求めてみると、$\dfrac{\mathcal{A}(xy)}{\mathcal{A}(xy)+\mathcal{A}(x\bar{y})} = 1,\quad \dfrac{\mathcal{A}(\bar{x}y)}{\mathcal{A}(\bar{x}y)+\mathcal{A}(\overline{xy})} \approx 0$ より、$\mathcal{A}(x \Rightarrow y) \approx 1$ となることがわかる。同様に、$\mathcal{A}(\bar{y} \Rightarrow \bar{x}) \geqq 0$ も満たす。3.6 節で述べたとおり、これら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、(53a) 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、(56) のような認知環境を持つ。

(56)

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

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、(56) の認知環境から、式 (**??**) によって推論の想定確信度を求めると、$\mathcal{A}(x \Rightarrow y) > 0$ の条件しか満たさない。これは (**??**d) 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 (54a) のような「〜してもよい」という許可スキーマを生み出すことになる。

以上、いくつかの心理実験を通して、認知主体における関連性計算の役割を見てきた。本節の議論から分かる通り、関連性を用いた擬似的な論理推論は、論理学との比較でい

うと正しいとはいえないことがある。しかし、論理学の正しさが認知主体にとって妥当とは限らない。関連性に基づく計算は、確かに時にエラーを引き起こすこともあるが、多くの場合、適切な答を教えてくれる指標になる。そして何よりも、関連性計算は人間にとって「真」とは何か、どういう想定が「知識」として妥当なものであるかという基準になり得るという点が重要なのである。

5. 総合論議

以上、関連性理論で言われている認知的関連性の原理が、論理判断や知識形成の上でどのような役割を果たしてきているのかを見た。本稿の主張は以下の通りである。

(57) a. 想定の確信度は真理値の見込みとして、推論判断に使いうる。
b. 関連性の計算を基礎にして、擬似的な論理を構築できる。
c. 無限の能力を持たない認知主体にとって、「真理」は認知環境の更新を通じて、初めて知りうるものである。すなわち、ある瞬間の想定だけを取り出しただけでは、何か「真理」であるかを決めることはできない。
d. 「(Y は)X である」という想定が「知識 (正当化された真なる信念)」であるためには、最低限、以下の条件を満たさなければならない。
- ある妥当なコンテクストが設定されており、
- I-領域・D-領域を含めた認知環境の集合において、
- 前提情報 Y を十全に考慮すると同時に、その否定情報 ¬Y については、考慮していないことを明確にするか、考慮するなら十全に考慮した上で、
- 認知環境の更新を通じて、一貫した**認識的必然性**を十分に満たすこと。

e. 論理でいう対偶論法は、認知主体の判断にとって常に絶対的なものとはいえない。

本稿の議論では、ドレツキやノージックによる知識に関する議論や、知識表現の問題 (特に「が」「は」によって表される命題や、条件表現の問題など) については触れることができなかった。これらの論点に関しては、また稿を改めて議論を行いたい。

参考文献

A.J. エイヤー (1981).『知識の哲学』. 白水社.

Chafe, Wallace L. (1994). *Discourse, Consciousness, and Time: the Flow and Displacement of Conscious Experience in Speaking and Writing*. The University of Chicago Press, Chicago.

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

Fauconnier, Gilles (1994). *Mental Spaces: Aspects of Meaning Construction in Natural Language (2nd ed.).* Cambridge University Press. 坂原茂ほか 訳 (1993).『メンタル・スペース理論』. 白水社.

三藤博 (1999). 談話の意味表示. 郡司隆男他（編），『談話と文脈』, pp. 55–91. 岩波書店, 東京.

Kahneman, D. & Tversky, A. (Eds.) (1982). *Judgment under uncertainty: Heuristics and biases.* Cambridge University Press, New York.

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

中垣啓 (2007). 認知的浮動による連言の説明.『日本心理学会第 71 回大会論文集』, p. 859.

ノージックロバート (1997).『考えることを考える (上・下)』. 青土社.

定延利之・熊谷吉治・苅田修司 (1999). 新情報と旧情報. 音声文法研究会（編），『文法と音声 III』, pp. 127–148. くろしお出版, 東京.

金水敏・田窪行則 (1990). 談話管理理論からみた日本語の指示詞.『認知科学の発展』, **3**, 85–115.

Saussure, Ferdinand de. (1916). *Course in General Linguistics.* McGraw, New York. 小林英夫訳 (1972).『一般言語学講義』. 岩波書店.

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.

Stalnaker, Robert (1970). Probabilities and Conditionals. In Harper, W. L., Stalnaker, R., & Pearce, G. (Eds.), *Ifs: Conditionals, Belief, Decision, Chance, and Time*, pp. 129–147. D. Reidel, Dordrecht.

田窪行則・金水敏 (1996a). 対話と共有知識—談話管理理論の立場から.『月刊言語』, **25** (1), 30–39.

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

戸田山和久 (2002).『知識の哲学』. 産業図書.

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

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