

認知環境の更新に関する妥当な計算手法について*

松井 理直

A Study of an Adequate Algorithm for Updating Cognitive Environments

Michinao F. MATSUI

Abstract

When we acquire the right cognitive environment, we cannot search for every relevant piece of information because our cognitive ability is not limitless. Cognitive agents, however, must have proper cognitive environments from partial information structure. In order to achieve this purpose, we need to continue updating cognitive environments through our experience in real world. This paper proposes a valid algorithm for updating the environments on the basis of the estimated degree of belief, and discusses the qualitative feature of our innate cognitive ability.

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。情報の真偽は真理として決まっており、全知全能の存在者なら真理を知っている。しかし、私たち認知主体は様々な点で限界を持つ存在者であり、その知識体系も基本的に不完全である。認知主体は全ての情報を知ることはできないし、獲得した情報も完全に正しいとは限らない。また、この世界には認知主体にとって例外と感じられる現象が生じることがあり、時に獲得した想定を更新を迫られる。このような不完全で不安定な環境において、認知主体は部分情報を適切に利用し、可能な限り情報に一貫性を持たせ、体制化された想定を持たなければならない。

こうした柔軟性を持った情報処理を行うためには、ある情報を認知環境内に想定として取り込む際、その想定の実偽を明確に決定するのではなく、「どの程度正しそうか」と

*本研究は、日本学術振興会科学研究費補助金・基盤研究 (A)(1)「推論機構の言語的実現とその解釈メカニズムに関する研究」(平成 19-22 年度、研究代表者:田窪行則、課題番号 19102012)の援助を受けている。

いう想定確信度を連続量で持つことが望ましい。またこの確信度の計算は、外界情報と常に相互作用を起こしつつ更新されるシステムであることが望まれる。本稿は、こうした認知環境の更新アルゴリズムと想定確信度の計算方法、およびこうしたアルゴリズム上で動作する認知能力の特性について論じたものである。

2. 想定確信度更新のためのアルゴリズム

2.1 知識と想定

プラトンの『テアイテトス』や『メノン』から現在に至る認識論の歴史において、「認識主体 S が命題 A を知っている」ということは、「 A が真であるということを S が確信しており、 A が実際に真であり、 A が真ということを S が確信する当然の理由がある」ことが基本的条件であると考えられてきた (Ayer, 1981)。簡単に言うと、「知識とは正当化された真なる想定」ということである。誤っていると分かっていることを信じ続けるのは知識ではないし、自分で正しいと思いこんでいても、それが正しくなければ知識とはいえない。また、あることを信じていて、その内容がたまたま正しいことであっても、その正しさが「まぐれ当たり」であるなら、それは知識とはいえない。知識と想定の大きな違いはこの点にある。偽である何かを信じることはできるが、偽である何かを「知る」ことはできない。

問題となるのは、どのような条件を満たせば「正当化された知識」といえるのかという点にある。古くから、正当化の条件とは、十分な証拠に基づいていることだと考えられていた。しかし、強い証拠主義に立つと、ほぼ全ての想定が知識とは言えなくなる。帰納推論は不確実性を持ち、多くの証拠を集めても次の瞬間に反例が見つかるかもしれない。演繹推論でも、命題に十分な証拠を求め続けると、ある命題の証拠となる別の命題を次々と求めなければならず、直接感覚所与以外の情報について無限後退に陥る。直接感覚さえも、ゲティア問題 (Gettier, 1963) のように、常に正当化されるとは限らない。

この証拠主義の問題点を克服するためには、正当化の条件を認識主体の活動のみに限定するのではなく、認識主体の外部に求めることが必要になる。これを認識論の外在主義と言う。外在主義の代表的な考え方として、信頼性主義や整合主義 (Armstrong, 1973) を挙げることができる。信頼性主義の正当化条件は「 S の想定が信頼できるプロセスにより形成された場合」と定義される。整合主義では、「 S の想定が他の想定と矛盾を起こさない」ことが正当化の条件となる。例えば、感覚などから直接得られる基本的情報は、それを歪める阻害要因が明確でない限り、信頼性主義により正当化された想定といってよい。また、ある想定が他の基本想定と矛盾しないのであれば、自分自身でその想定の本拠を説明できない場合でも、整合主義により正当化された想定と見なし得る。

2.2 信頼できる情報獲得過程

認識論の外在主義は懐疑主義を克服する良いアイデアであり、我々の認知能力もこうした特性を備えていると考えられる。ただし、こうした認知能力は例外事象や偶発的な情報といったノイズが存在する現実世界にあつて、そのノイズに耐えうる信頼性を持つ

た判断を行わなければならない。「想定確信度」はこうした柔軟な判断を行うために有利な概念の一つである。想定は知識の基になるものであっても知識そのものではないため、想定確信度は真か偽かという二値に限定されず、連続量であってよい。信号とノイズを切り分ける計算が可能になる。また、真理値についても、想定確信度がある一定の条件を満たせば、その想定肯定形あるいは否定形を知識として獲得すればよい(4節参照)。

問題となるのは、この想定確信度を決定する妥当な方法である。認知主体は、限られた情報と時間の中であっても適切な情報処理を行わなければならない。したがって、大量のデータを必要とする大数の法則に基づくフィッシャー流の確率値は、想定確信度の計算方法として必ずしも妥当ではない。むしろ、各時間ごと個別の情報を獲得し、その度ごとに想定確信度を更新できる演算過程であることが望ましい。次節以降で、こうした想定確信度の計算方法を考察してみよう。

2.3 過去の履歴を伴う想定確信度の更新過程

我々認知主体は、過去・現在・未来という歴史の流れの中に存在している。過去の情報は無知であっても未知ではない。少なくとも過去の各情報は真か偽かのいずれかであることは確信できるので、過去の世界に対して適当な信念確信度を割り振ることができる。また、今現在経験している情報については、それを完全に真なるものとして受理できると考え、懐疑論の立場は採用しない。その代わりに、この瞬間の情報を直接受け入れても、将来的に想定確信度を変更可能な計算過程によって、信頼性判断に破綻を来さないシステムを仮定する。これに対し、未来の情報は既定のものとは未知なるものがあり、後者については過去・現在から予測されるものである。これを概略図で表すと図1のようになる。

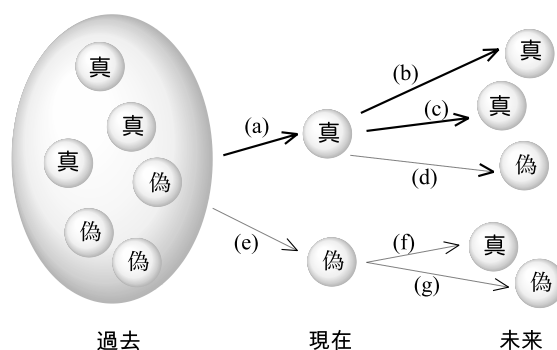


図1: 過去・現在・未来における情報 X の真偽

想定確信度にフィッシャー流の頻度確率を用いるならば、この過去・現在・未来における情報 X の生起頻度を実際に取得しなければならない。しかし、現実問題としてこれは極めて困難である。例えば過去の世界における情報 X は明確にその真理値が決まって

いるが、認知主体の「無知」の故に、過去世界のあり方を完全に再現することはできない。また、未来の世界についても、認知主体の「未知」の故に、明確で安定した1つの世界像を構築することができない。認知主体にとって唯一確実なことは、「今現在、情報 X が生起している」「今現在、否定情報 $\neg X$ が生起している」という現在の事象に関する想定のみである。

このような制限の中で、認知主体に求められる能力は、過去・現在・未来という移りゆく時間の中で、様々な情報経験を通じて、情報 X に関する想定確信度を更新してゆき、なるべく真理に近い想定確信度を獲得するシステムである。例えば、次のような認知機構が考えられる。

- (1) a. 今現在の情報 X に関する想定確信度は、現在実際にその情報が真なるものとして存在している (あるいは否定情報 $\neg X$ が真なるものとして存在している) として受け入れる。
- b. 過去の情報に関しては、暫定的に一つの世界像を構築する。
- c. 未来の情報に関しては、現在の情報から暫定的な予測を行う。
- d. 現在という時が物理的現象 (物理的時間) の変遷と共に変化していくのに従って、暫定的な処理の結果も含んだ想定の確信度を刻々と更新していく。

こうした更新過程は次のようにモデル化できる。今、認知主体がある情報 X を現在初めて直接に体験し、この情報の真偽を決定しようとしている場面を考えてみよう。まず、認知主体は情報 X を認知環境に受理する前に、 $P_0(x)$ という確信度で情報 X が真であるという想定を立て、かつ $P_0(\bar{x})$ という確信度で否定情報 $\neg X$ が真であるという想定を立てる。 $P_0(x), P_0(\bar{x})$ は $0 \leq P_0(x) \leq 1, 0 \leq P_0(\bar{x}) \leq 1, P_0(x) + P_0(\bar{x}) = 1$ を満たす任意の妥当な値であり、図 1 の過去世界における情報 X の暫定的な想定確信度に相当する。本稿で議論するアルゴリズムでは、この $P_0(x), P_0(\bar{x})$ という初期値は適当に決めてもよく、情報 X に関する事象を繰り返し経験することによって、妥当な確信度に収束する (3.1 節参照)。ただし、一般的に過去情報の真偽が無知であるという意味を考慮するなら、 $P_0(x), P_0(\bar{x})$ 共に 0.5 と設定するのが妥当であろう。

これに対し、現在世界の情報に関しては、認知主体はその情報を直接経験できるため、現在世界における情報 X の想定確信度を 1 として設定すればよい。¹ したがって、「過去世界」と「現在世界」を合わせた想定確信度は、肯定情報 X と否定情報 $\neg X$ の想定確信度の比率が $P_0(x) + 1 : P_0(\bar{x})$ を満たす値となる。さらに、この比率に未来世界における想定確信度を追加すると、最も適切な想定確信度が決定できる。

しかし、認知主体にとって未来の事象は未知であるので、その想定確信度も推測するしかない。この推測を行うにあたって最も妥当な方法の 1 つは、今この瞬間における「過去・現在・未来」の関係が、過去においても成立しており、また未来においても同様な関

¹本稿では懐疑論を採らない。4 節で真なる情報の想定確信度が 1 にならないということを議論するが、直接経験に関しては完全な想定確信度を持つとする。なお、直接経験でも不完全な想定確信度しか持てないという懐疑論については、稿を改めて議論する。

係が生じるという再帰的性質を仮定することであろう。この世界は基本的には極端に変化することはないという仮定であるといってもよい。この仮定に従うなら、未来世界において情報 X が真となる想定確信度は、図 1 における “path (a)–(b)” および “path (a)–(c)” の可能性を計算すればよく、これは $P_0(x)^2$ と表せる。単純にいうなら、この数値はいずれ現在となる近接未来において情報 X が成立する期待値に相当する。同様に情報 X が偽となる想定確信度は、図 1 の “path (a)–(d)”, “path (e)–(f)”, “path (e)–(g)” の経路を計算すればよく、これは $P_0(\bar{x})^2 + 2P_0(x)P_0(\bar{x})$ あるいは $1 - P_0(x)^2$ となる。この値は近接未来において否定情報 $\neg X$ が成立する期待値に該当する。

以上のことから、新たに更新すべき情報 X の想定確信度 $P_1(x)$ は、否定情報との比率 $P_1(x) : P_1(\bar{x})$ の比率が $P_0(x) + 1 + P_0(x)^2 : P_0(\bar{x}) + (1 - P_0(x)^2)$ を満たせばよい。ここで、 $1 - P_0(x)^2$ は $P_0(\bar{x})^2 + 2P_0(x)P_0(\bar{x})$ と等しいため、 $P_0(\bar{x}) + (1 - P_0(x)^2)$ は $P_0(\bar{x}) \cdot (2 + P_0(\bar{x}))$ と変形でき、さらに $P_0(x) + 1 + P_0(x)^2$ と $P_0(\bar{x}) + (1 - P_0(x)^2)$ の総計は 3 になることから、情報 X と否定情報 $\neg X$ に対する更新された想定確信度は各々 $P_1(x) = \frac{1 + P_0(x) \cdot (1 + P_0(x))}{3}$, $P_1(\bar{x}) = \frac{P_0(\bar{x}) \cdot (2 + P_0(\bar{x}))}{3}$ として計算できる。

さらに、過去・現在・未来がどのような物理的時間においても基本的に類似した性質を持つという仮定から、上述の性質を一般化することができる。まず、現実世界において、情報 X に関する n 回目の証拠が得られたとき、過去・現在・未来に関する想定確信度は (2) のようになる。

- (2) a. 現実世界における n 回目の証拠が情報 X であった場合、情報 X に関する確信度は以下の通りである。
 1. 過去 (n-1 回目まで) の累積情報の確信度: $P_{n-1}(x)$
 2. 現在の確信度: 1
 3. 近接未来に対する確信度: $P_{n-1}(x)^2$
- b. 同様の場合、情報 $\neg X$ に関する確信度は以下の通りである。
 1. 過去の累積情報に対する確信度: $P_{n-1}(\bar{x}) = 1 - P_{n-1}(x)$
 2. 現在の確信度: 0
 3. 近接未来に対する確信度: $1 - P_{n-1}(x)^2$
- c. 現実世界における n 回目の証拠が情報 $\neg X$ であった場合、情報 X に関する確信度は以下の通りである。
 1. 過去の累積情報に対する確信度: $P_{n-1}(x)$
 2. 現在の確信度: 0
 3. 近接未来に対する確信度: $1 - P_{n-1}(\bar{x})^2$
- d. 同様の場合、情報 $\neg X$ に関する確信度は以下の通りである。
 1. 過去の累積情報に対する確信度: $P_{n-1}(\bar{x}) = 1 - P_{n-1}(x)$
 2. 現在の確信度: 1

3. 近接未来に対する確信度： $P_{n-1}(\bar{x})^2$

このことから、現時点において情報 X に関する n 回目の証拠を得た時の想定確信度は、式 (3) として表現できる。同様に、否定情報 $\neg X$ に関する n 回目の証拠を発見した場合には、否定情報の想定確信度を式 (4) によって計算できる。

(3) 現実の世界において n 回目の証拠が情報 X の成立であった場合：

- a. 情報 X の想定確信度： $P_n(x) = \frac{1 + P_{n-1}(x) \cdot (1 + P_{n-1}(x))}{3}$
- b. 否定情報 $\neg X$ の想定確信度： $P_n(\bar{x}) = \frac{P_{n-1}(\bar{x}) \cdot (2 + P_{n-1}(\bar{x}))}{3}$
- c. $P_0(x), P_0(\bar{x})$ は任意の数値

(4) 現実の世界において n 回目の証拠が否定情報 $\neg X$ の成立であった場合：

- a. 情報 X の想定確信度： $P_n(x) = \frac{P_{n-1}(x) \cdot (2 + P_{n-1}(x))}{3}$
- b. 否定情報 $\neg X$ の想定確信度： $P_n(\bar{x}) = \frac{1 + P_{n-1}(\bar{x}) \cdot (1 + P_{n-1}(\bar{x}))}{3}$
- c. $P_0(x), P_0(\bar{x})$ は任意の数値

2.4 過去世界を考慮しない想定確信度の更新様式

前節の議論では、過去・現在・未来という時の流れの中で、想定確信度を再帰的に更新していく過程を見た。しかし、現実の認知処理においては、過去に一旦受理した情報を取り消さなければならないことがある。推論思考における前提情報を保持できなくなる場合や、反事実条件文のように現実に反する想定を故意に立てて推論を行う場合、あるいは過去の出来事に対する後悔などがこれに当たる。こうした認知処理は、過去情報の累積効果に直接影響されないという点で、過去世界を現在世界に取り込み、同時に処理を行う認知過程と見なすことができる。したがって、このような認知処理では、前節の議論で見た $P_1(x) : P_1(\bar{x})$ の最終的な比率である $P_0(x) + 1 + P_0(x)^2 : P_0(\bar{x}) + (1 - P_0(x)^2)$ という関係式のうち、最初に足しまれている $P_0(x), P_0(\bar{x})$ の数値を含めてはならない。したがって、心理的同時性を満たす場合の想定確信度の更新比率は、 $P_1(x) : P_1(\bar{x})$ について $1 + P_0(x)^2 : 1 - P_0(x)^2$ であればよい。この結果、過去世界を考慮しない想定確信度は、以下の計算式に従って更新すればよいことになる。

(5) 過去世界を考慮しない処理において、n 個目の証拠が情報 X の成立であった場合：

- a. 情報 X の想定確信度： $P_n(x) = \frac{1 + P_{n-1}(x)^2}{2}$
- b. 否定情報 $\neg X$ の想定確信度： $P_n(\bar{x}) = \frac{1 - P_{n-1}(x)^2}{2}$
- c. $P_0(x), P_0(\bar{x})$ は任意の数値

(6) 過去世界を考慮せず、かつ n 個目の証拠が否定情報 $\neg X$ の成立であった場合：

- a. 情報 X の想定確信度： $P_n(x) = \frac{1 - P_{n-1}(\bar{x})^2}{2}$
- b. 否定情報 $\neg X$ の想定確信度： $P_n(\bar{x}) = \frac{1 + P_{n-1}(\bar{x})^2}{2}$
- c. $P_0(x), P_0(\bar{x})$ は任意の数値

過去世界を考慮する確信度更新の式 (3), (4) と本節で見た過去世界を考慮しない確信度更新の式 (5), (6) は、いくつかの点で異なった性質を持つ。次節では、これらの定義式に基づく想定確信度の特徴について議論を行う。

3. 想定確信度の更新様式に関する性質

3.1 真なる想定・知識の獲得

まず初めに、過去の履歴を考慮する信念確信度得の更新式 (3), (4) の特徴から見てみよう。今、情報 X が完全に真であり、世界に否定情報 $\neg X$ が存在しないとする。ただし、認知主体は最初は情報 X が完全に真であることを知らないため、初期想定を $P_0(x) = 0.5, P_0(\bar{x}) = 0.5$ と設定している。ここから、認知主体は現実世界の中で情報 X の存在を何度も経験し、情報 X に関する想定確信度を図 2 のように更新していく。

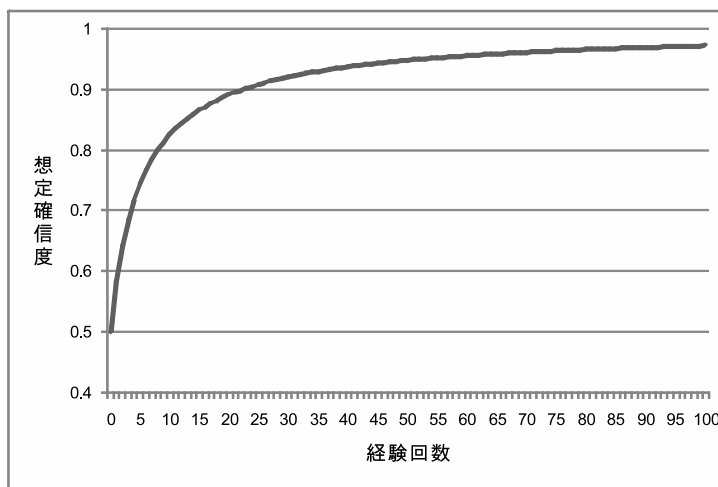


図 2: 初期想定確信度を 0.5 と設定した場合の更新過程

グラフから分かる通り、情報 X に関する想定確信度は経験回数と共に上昇しており、情報 X の存在を 9 回経験した時点で確信度は 0.8 を超え、23 回経験した時点で 0.9 を上回る値となる。また、経験回数が 52 回に達した時、想定確信度は 0.95 を超える。

ここで、事前想定の影響について見てみよう。もし事前に情報 X について真である可能性が高いと考えられる状況であれば、認知主体は初期想定確信度 $P_0(x)$ を高く設定す

ることができる。ここで今、認知主体が $P_0(x) = 0.8, P_0(\bar{x}) = 0.2$ という初期確信度を持っていたとすると、その想定更新は図3のような過程を辿り、経験回数 17 回目で確信度は 0.9 を、経験回数 44 回目で確信度は 0.95 を上回ることになる。

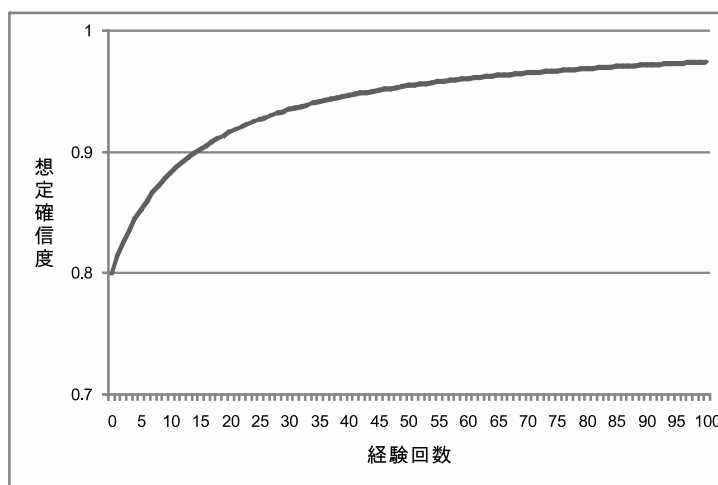


図3: 初期想定確信度を 0.8 と設定した場合の更新過程

逆に事前に情報 X が偽である可能性が高いことと予想される状況であるなら、例えば $P_0(x) = 0.2, P_0(\bar{x}) = 0.8$ といったように $P_0(x)$ が低く設定される。しかし、実際には情報 X が真であり、それ故に情報 X の生起しか経験せず、否定情報 $\neg X$ を経験しないことによって、図4のように想定確信度が更新されていく。初期値が低くても、早い段階で確信度の変更が急速に行われ、情報 X を 10 回経験した時点で既に想定確信度は 0.8 を超え、経験回数 25 回目で 0.9 を上回り、54 回目で 0.95 を超えることが読み取れる。

ここで初期想定確信度が 0.5 であっても 0.2 であっても、想定確信度の更新スピードに大きな違いはない点に注意されたい。この性質は、式 (3), (4) に基づくアルゴリズムにおいて、初期想定が信念の獲得にそれほど強く影響していないことを示している。これは知識獲得において先入観の影響を強く受けないということであり、安定した知識を獲得するためのすぐれた性質である。このアルゴリズムでは、初期状態で否定的な「先入観」を持っていても、現実の情報に肯定的な性質でありさえすれば、初期の想定を急速に変更することができ、初期状態で肯定的な先入観を持っている場合と大きくは異なる情報獲得が可能になっているのである。

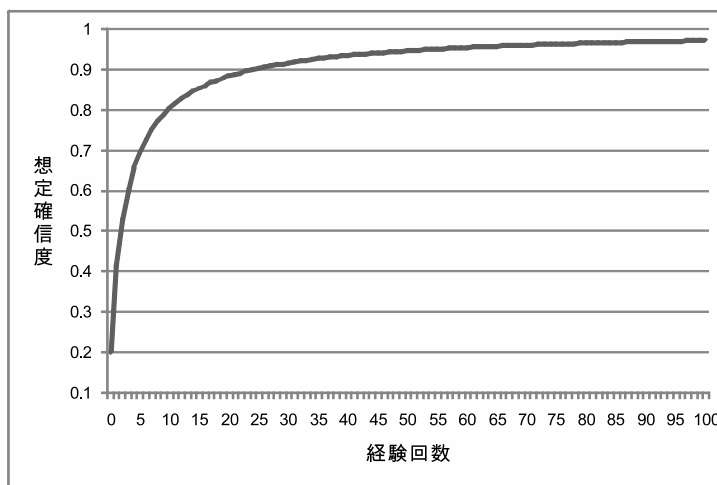


図 4: 初期想定確信度を 0.2 と設定した場合の更新過程

3.2 反例が実在する場合の特徴

前節では、現実世界の事例が情報 X を支持しているのであれば、情報 X に関する正しい想定値を獲得できることを見た。しかし、現実世界に反例が存在した場合は、どのような想定更新になるのだろうか。

今、情報 X の初期想定確信度を $P_0(x) = 0.5$ とし、10 回目に否定情報 $\neg X$ を経験したとする。また、11 回目以降はまた情報 X のみを経験することにしよう。すなわち、10 回目に経験した情報は単なる例外か accidental なデータであったという状況である。この時の想定更新過程を図 5 に示す。

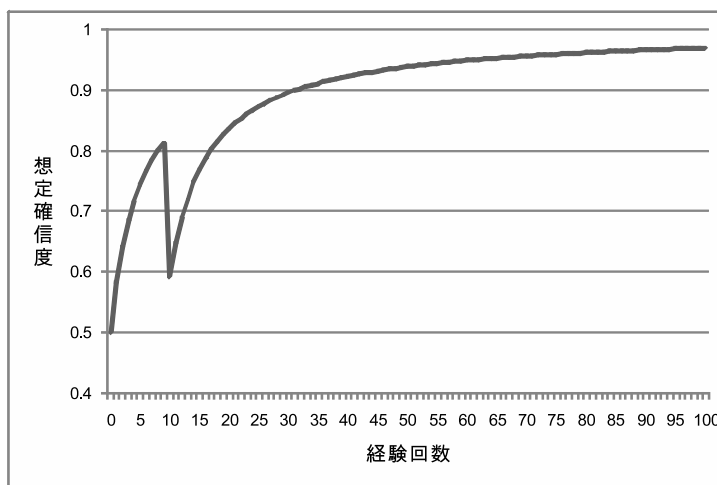


図 5: 10 回目に例外事象を経験した場合

グラフから、9回目に0.79まで上昇していた想定確信度が、例外を経験した時点で急速に0.58まで下がり、その後また確信度が上昇するような変化を示していることが読み取れる。すなわち、0.8近くまで確信していた想定であっても、反例に出会ったならば、一旦真か偽か分からない(多少真である確信のほうが高い)という状態に認知環境を戻し、その後、やはり情報 X は真であるという確信を高めていくような性質を持つのである。

同様に、10回目・11回目に連続して例外事象を経験し、12回目以降は正常事象を経験した更新過程を図6に示す。たった2回しか反例に出会っていないにもかかわらず、想定確信度が初期想定と同じレベル(約0.47)にまで戻っていることが分かる。すなわち、反例となる否定的情報は、仮説を支持する肯定的情報の連続よりもさらに強い効果を持つのである。

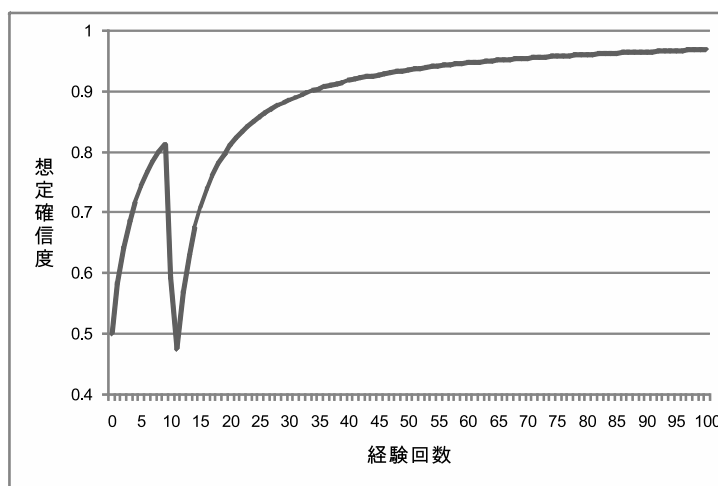


図6: 10回目・11回目に連続して例外事象を経験した場合

反例に出会うタイミングが遅い場合であっても、この性質は大きくは変わらない。図7に60回目で初めて反例を経験した場合を、図6に60回目・61回目に連続して反例を経験した場合の例を示す。59回目の時点で想定確信度は0.954にまで上昇しているが、60回目に例外事象に出会った瞬間、想定確信度は0.65にまで減少していることが分かる。また、60回目・61回目に連続して例外を経験した場合には、想定確信度は初期状態とほぼ同じレベルである0.509にまで下がり、情報 X の真偽が全く決定できないという中立状態に戻る。なお、この傾向は初期確信度が0.5でなくても保たれており、初期確信度とは無関係に成り立つ。

以上の特徴は、本稿で提案した式(3), (4)に基づくアルゴリズムが、初期確信度のみならず、反例や例外を経験するタイミングにも影響されないことを示している。すなわち、本稿で提案する想定確信度の更新様式は、反例に鋭敏な獲得様式を持っており、常に仮説検証を志向する性質を備えているということである。この特徴が正しい「知識」を獲得する上で極めて重要な性質であることは言うまでもない。

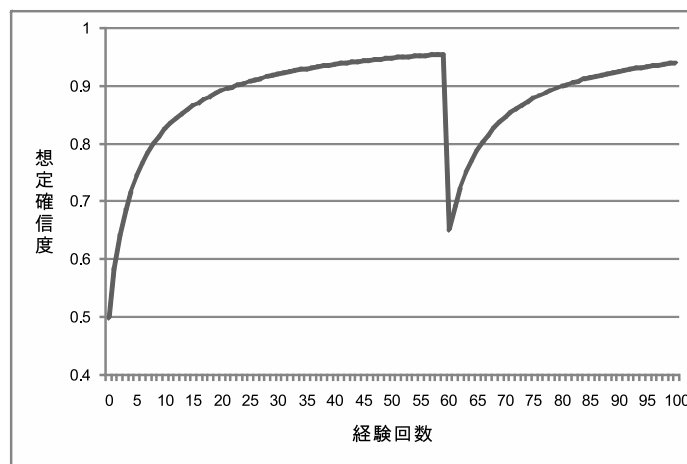


図 7: 60 回目に初めて例外事象を経験した場合

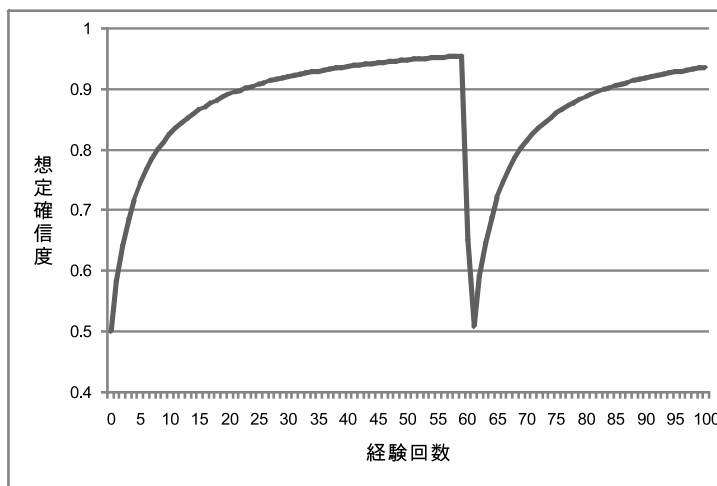


図 8: 60 回目・61 回目に連続して例外事象を経験した場合

3.3 過去世界を考慮しない想定更新過程

前節で述べた過去・現在・未来の状態を推測しつつ想定を更新を行う過程に対し、過去の情報を考慮しない確信度の更新式 (5), (6) は、参照する情報が少なくなる分、確信度の獲得が迅速でありかつ違反例に対しより敏感であるという特徴を持つ。図 9 に、初期想定確信度を $P_0(x) = 0.5, P_0(\bar{x}) = 0.5$ と設定した場合の確信度更新過程を示す。同一の初期想定値で過去情報を取り込む条件である 図 2 と比較して収束速度がより迅速であり、5 回目の情報経験で確信度は 0.8 を越え、15 回目で 0.9 に達している。確信度が 0.95 を越えるのも 34 回目の情報を経験した時点である。

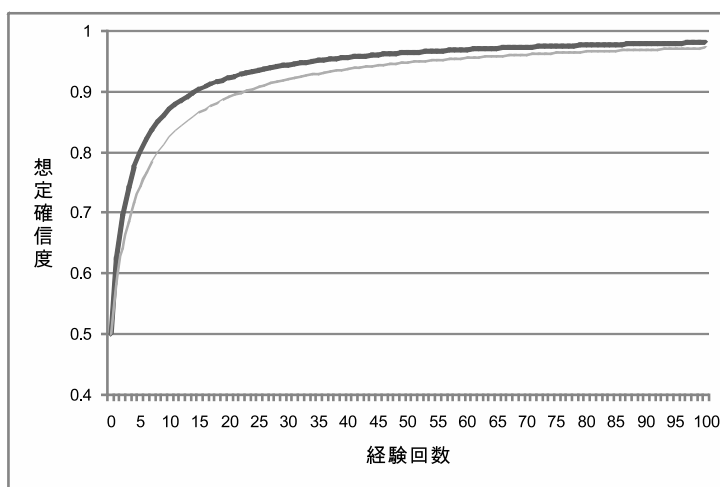


図 9: 初期想定が 0.5 の場合の確信度更新過程

太線：過去世界を考慮しない場合

細線：過去世界を考慮する場合 (図 a1-05t と同一)

一方、反例や例外事例に遭遇した時は、たった 1 つの否定情報を見つけた時点で、即座に真偽を定めない中立的な確信度に戻る。10 回目に例外事例を発見した場合には想定確信度は 0.49 になり、60 回目に初めて例外事例と遭遇した場合でも、その時点で想定確信度は 0.5 にまで落ち込む。すなわち、過去情報を取り込むアルゴリズムに比べ、違反事例や例外事例に対してより鋭敏な性質を持つ。

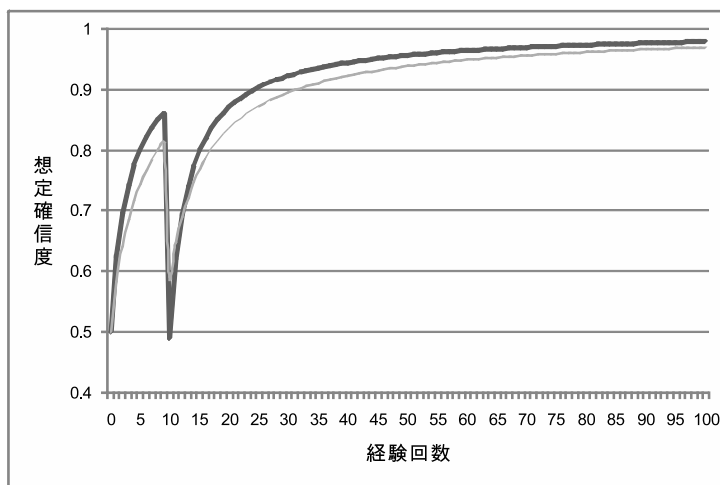


図 10: 10 回目に初めて例外事例を発見した場合の確信度更新過程

太線：過去世界を考慮しない場合

細線：過去世界を考慮する場合 (図 5 と同一)

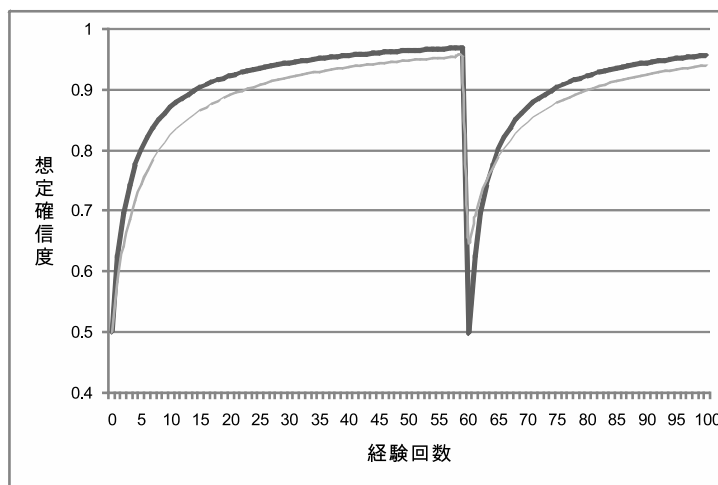


図 11: 60 回目に初めて例外事例を発見した場合の確信度更新過程
 太線：過去世界を考慮しない場合
 細線：過去世界を考慮する場合 (図 7 と同一)

3.4 近接未来に対する期待度

これらのアルゴリズムは、近接未来の予想という点でも興味深い性質を持っている。例えばコイントスをしていて、表が 3 回連続で出ていた時、4 回目も表になる確率を聞かれたとしよう。数学的には何回目であっても表の出る確率は 0.5 であるが、現実こういう賭け事をしている人は、しばしば「そろそろ裏が出るだろう」という期待を持ってしまう。あるいは、表が 20 回連続で出ていたなら、「次こそは裏が出るだろう」と期待するよりも、「このコインは細工がしてある偽物だろう」と考えることだろう。つまり、認知主体の行う予測は時に時に数学的な確率と異なる性質を持つことがある。

こうした認知主体が抱く想定（性質を、確信度更新の点から検討してみよう。(2) で見たように、今現在において情報 X の存在を確認できた時、認知主体は $P_{n-1}(x)^2$ という確信度で、近未来においても情報 X が存在するという期待を抱く。図 12 は、情報 X の真偽について中立的な初期想定 (すなわち $P_0(x) = 0.5$) を持っている時に、情報 X を発見する経験回数と共に近未来に対する期待度がどのように変化しているのかを示したものである。太線は否定情報 $\neg X$ に対する期待度の変化を、細線は肯定情報 X に対する期待度の変化を、点線はその時点での想定確信度の更新過程 (図 2 と同一) を表す。

ここで、図 12 において、最初から数回分の間に起こる否定情報に対する期待値に注目して欲しい。このグラフでは肯定情報 X は完全に真であり、現実場面でも情報 X しか経験しないため、情報 X に対する想定確信度は徐々に高くなっているが、それにも関わらず、最初の数回分の間は、否定情報 $\neg X$ に対する期待度も高くなっている。これは、初期状態が情報の真偽について最も中立的な $P_0(x) = 0.5, P_0(\bar{x}) = 0.5$ となっていることが原因である。最初期の状態では、認知主体は情報 X の真偽が分かっていないため、現実世

界で得たデータが情報 X を支持するものであった時、情報 X に対する想定確信度を上昇させると共に、しかし次は否定情報 $\neg X$ が見つかるだろうという期待も大きくなることを意味している。ただし、この高い期待度は 5 回連続で肯定証拠を得る時くらいまでしか続かず、その後は初期状態の期待値 $P_0(\bar{x})$ を急速に下回っていく。この性質は、現実世界に否定的な想定を確認する情報が無いために、認知主体が情報 X を真理であると理解し始めた行動として解釈できる。

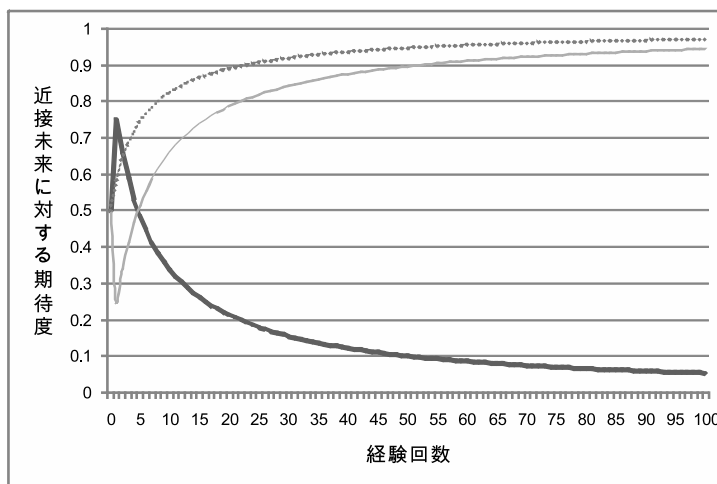


図 12: 情報 X を連続して経験した時の近未来に対する期待度 (初期状態 0.5)

太線：否定情報 $\neg X$ の期待度 細線：肯定情報 X の期待度

点線：想定確信度の更新過程 (図 2 と同一)

この近接未来に対する期待値の変化から、肯定情報に対する確信度と否定情報に対する確信度の和が 1 にならないことがあるという事実に対する示唆も得られる。すなわち、実験場面で否定情報の確信度を計測した時、それが純粹に現在の状況に対する心的状況を反映しているだけでなく、近接未来に起こりうる否定的事象をも予測し、その想定期待度を足し込んでしまうために、総和が 1 を超えてしまうということである。

こうした性質を実験的に制御するためには、実験を開始するに当たり、初期状態を中立的な予測に置いておくのではなく、真か偽かのどちらかに極端に偏るような場面を設定しておけばよいと思われる。近接未来の期待値は、初期状態における想定確信度の影響を受けるからである。例えば、初期状態において否定情報 $\neg X$ に対する想定確信度を高く持っている場合には、近未来に対する期待値も図 12 とは異なった更新過程を辿る。図 13 に初期想定確信度を $P_0(x) = 0.2, P_0(\bar{x}) = 0.8$ とした時の変化を示す。この場合も与えられる情報は肯定証拠 X ばかりであるため、否定情報 $\neg X$ の期待度は一瞬だけ高くなった後は低下し続け、3 回目の情報を経験した時点で既に初期想定確信度 $P_0(\bar{x}) = 0.8$ を下回ってしまう。

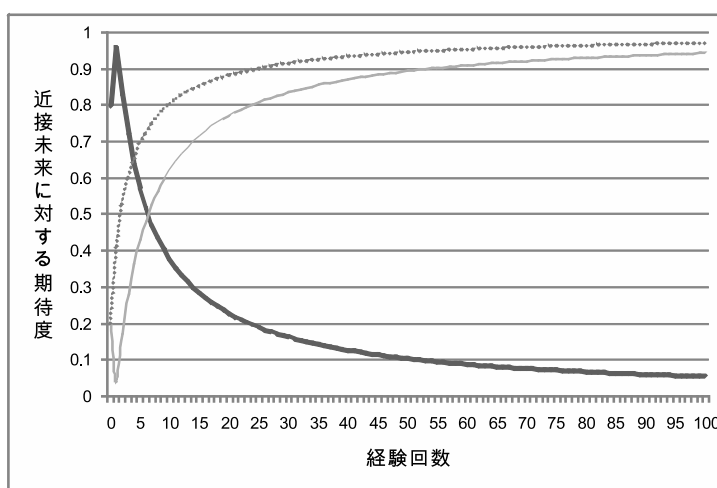


図 13: 情報 X を連続して経験した時の近未来に対する期待度 (初期状態 0.2)

太線：否定情報 $\neg X$ の期待度 細線：肯定情報 X の期待度

点線：想定確信度の更新過程

また、初期状態において肯定的な確信度が非常に高い場合には、現実データに否定的証拠がない限り、近接未来における否定情報の期待値は低下していくだけである (図 14)。

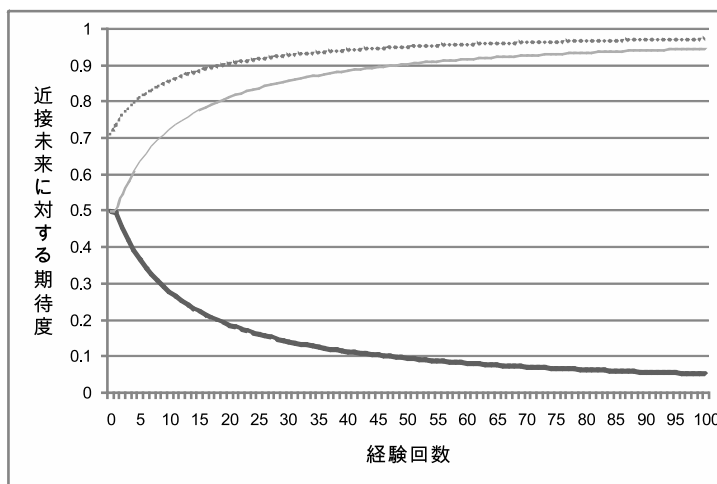


図 14: 情報 X を連続して経験した時の近未来に対する期待度 (初期状態 0.7)

太線：否定情報 $\neg X$ の期待度 細線：肯定情報 X の期待度

点線：想定確信度の更新過程

では、途中で反例事象あるいは例外事象を見つけた場合はどうなるのだろうか。図 15 に 10 回目に初めて反証事例 $\neg X$ を発見した場合、図 16 に 60 回目に初めて反証事例を

発見した場合における想定確信度の変化を示す。

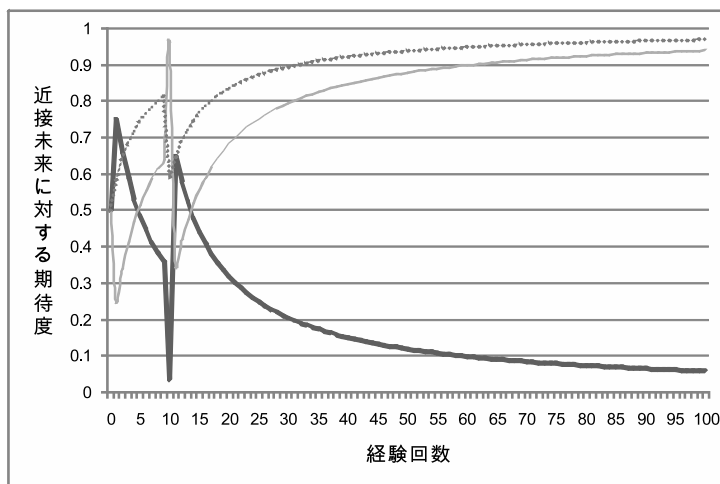


図 15: 10 回目に初めて否定情報を経験した場合 (初期状態 0.5)

太線：否定情報 $\neg X$ の期待度 細線：肯定情報 X の期待度
点線：想定確信度の更新過程

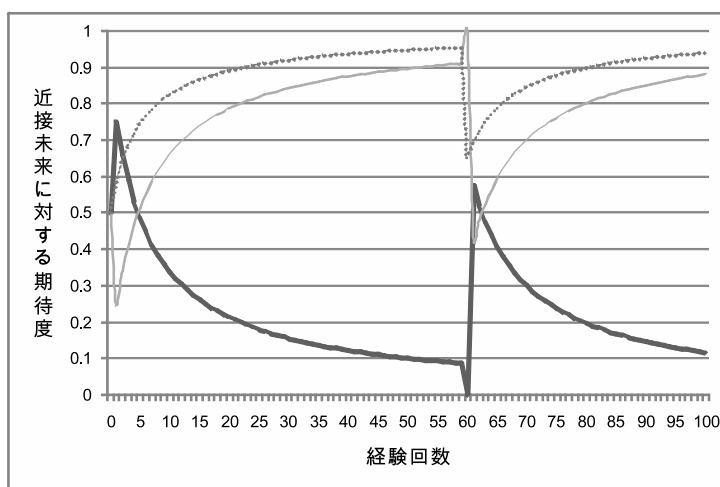


図 16: 60 回目に初めて否定情報を経験した場合 (初期状態 0.5)

太線：否定情報 $\neg X$ の期待度
細線：肯定情報 X の期待度
点線：想定確信度の更新過程

初めて反証事例を経験するのが早い段階であるか遅い段階であるかに関わらず、いずれの場合も、反証事例を発見した瞬間に、近接未来における否定情報 $\neg X$ の期待値が極

端に低くなり、次に肯定情報を見いだしたときに、再び否定情報に対する期待値が高くなっていることが見て取れる。これは現実と未来に対する我々の期待の持ち方のパターンに近い性質とってよいだろう。

3.5 初期状態に拘束される想定の影響

想定確信度を更新していく過程をモデル化するためには、変更を受けにくい初期想定
の存在も考慮する必要がある。このような情報の例として、言語や感覚系における生得
的な知識や、後天的な固定観念などが挙げられる。この初期状態に拘束される想定とは、
現実にとどのような証拠を発見しようと、最初期に持った想定確信度が常に参照されるよ
うな情報と見なしてよい。最初期の確信度を参照するプロセスは複数の方法でモデル化
が可能であるが、最も単純なモデルは、図 1 における path (a) や path (e) において、経験
によって更新された想定値を用いず、あくまで初期設定の想定値を使うことによって表
現できる。この設定の元でシミュレーションを行ったものが、図 17 である。最初の数回
のうちは、実環境で得られるデータの影響を少し受けるが、その後は想定確信度が全く
変化していないことが分かる。

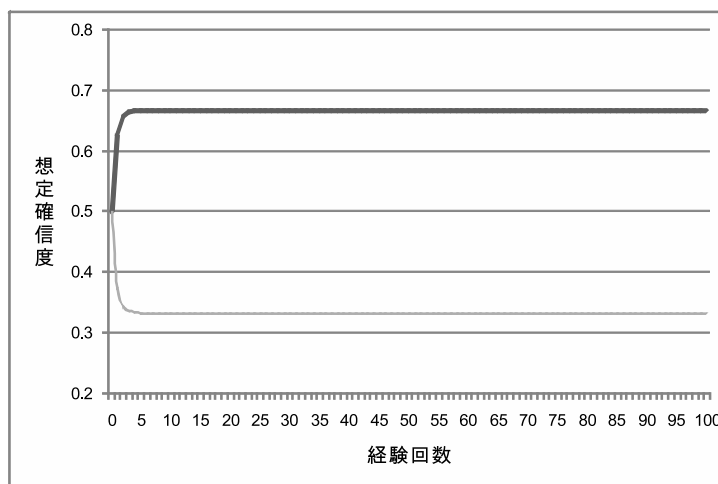


図 17: 初期状態に束縛される想定の高信度更新過程の一例

太線：情報 X の想定確信度

細線：否定情報 $\neg X$ の想定確信度

途中で反例事象を見つけた場合も、想定値の変動がある範囲内で収束しており、初期
状態の想定確信度から大きく変化することがない。ある情報が外界にあった場合でも、
その情報の想定確信度にはほとんど変化が生じない(すなわち認知環境がほとんど変化しな
い)ということは、その情報が認知環境にとって関連性のないデータであることを意味す
る (Sperber & Wilson, 1986)。このことが、初期状態に縛られる想定を持った時に新規の

情報を無視してしまう傾向が生じるさせる。なお、初期状態に依存する想定確信度の変更モデルは、本稿で述べた計算手法とは別の過程も考えることができるが、この点についてはまた別稿に譲りたい。

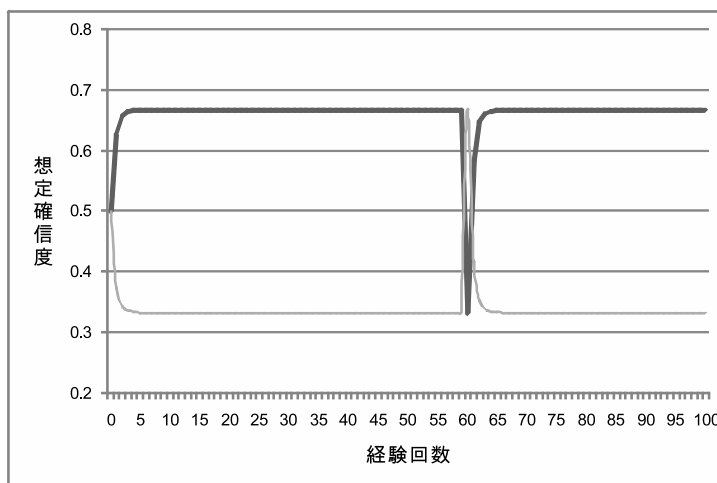


図 18: 初期想定の高く設定した時の確信度更新過程

太線：情報 X の想定確信度

細線：否定情報 $\neg X$ の想定確信度

以上、本節で述べた初期状態に拘束される認知過程を除き、式 (3), (4), (5), (6) の計算過程によって、認知主体はどのような想定状態からスタートしても情報の真理値や情報値とほぼ等価な想定確信度を獲得することができる。しかし、この「ほぼ等価な」という点に問題が残る。図 2?図 16 までのいずれのグラフにおいても、認知主体が持つ想定確信度は 1 に到達していない。真なる情報 X を 100 回連続で経験した場合でも、100% の想定確信度を持っていない。このことを単純に見ると、本アルゴリズムに従う認知主体は、真理に近づけるだけで、真なる情報である「知識」の獲得には至らないように思われる。

「真理とは何か」という哲学的な議論は未だに決着がついていない。また、我々認知主体が常に真理を見出せるとは限らないことも否定できない。しかし、認知主体が何らかの真理を知識として知っていることも又、否定できないことであろう。もし真理という「事象の在り方」を認めることができるとするのなら、認知主体が時に真理を見いだせない原因が存在するはずである。例えば、外的な要因としては、情報が意図的に歪められて伝えられている、情報の伝達経路においてノイズが混在するといった原因が考えられ、内的な要因としては認知の限界や歪みが考えられるであろう。

次節では、真理という事象の性質が明確であるにも関わらず、外界のノイズや我々の認識能力の限界のために、真理を確実に捉えることができているのだという立場で、

真理値と想定確信度の間で成立する数理的な関係について考察する。結論として、ある情報の想定確信度が1に達していなくても、その情報を真として帰納できることを議論し、本アルゴリズムでも知識獲得が行える可能性を示唆したい。

4. 真理判断の指標

4.1 認識過程における真理値と攪乱要因の作用

今、情報 X の真偽を変数 T とする。情報 X が二値の命題である時には、 $T = 1$ の時が真、 $T = -1$ の時が偽となる。三値の真理値を採る命題であるなら、 $T = 0$ の時が未知 (*unknown*) である。情報 X が多値の真理値 (あるいは情報価) を取る場合は、 T を $-1 \leq T \leq 1$ の範囲で真理値 (情報価) に対応させる。

これに対し、真理値の認識を混乱させる外界のノイズや我々の認識能力の限界といった攪乱要因の存在を変数 N で表す。 $N = 1$ は攪乱要因が影響を与えていることを、 $N = 0$ なら攪乱要因が関わらないことを表す。実際には、この N は外的攪乱要因 N_o および認知主体の持つ内的攪乱要因 N_i から成るが、本稿では議論を単純にするために、これら要因を纏めて扱う。

次に、認知主体が情報 X を受理する認知能力を定義する。今、認知主体が真理値を認識する能力の鋭敏さを定数 a 、攪乱要因が認知主体に与える影響の大きさを定数 b で表す。 a, b は共に $0 \leq a, 0 \leq b$ を満たす実数である。以下では、定数 a を認識度指数、定数 b を攪乱度指数と呼ぶ。この時、認知主体の認識量は aT で、認知主体を混乱させる攪乱量は bT で表せる。

この認識度指数 a は、真理値に対する認識能力の敏感さを表す重要な数値である。心理学的にも哲学的にも、認知主体が真理をどのように認識しているかは未だに議論されているところであるが、日本語や英語をはじめとする多くの言語表現を観察する限り、

- (7) a. モダリティのない「 X である」
- b. 認識的必然性を伴う「 X に違いない」
- c. 認識的可能性を伴う「 X かもしれない」

という3段階の区別が少なくとも存在する。こうした区別は様相論理でも扱うことができる。このことから、認識度指数 a は一般的に $a \approx 3$ を満たし、認識度が悪い場合は $a < 3$ 、認識度が鋭敏な場合には $a > 3$ を満たすと考えてよいだろう。

一方、攪乱量 bN は実際には外的攪乱量と内的攪乱量の総和 $bN = b_o N_o + b_i N_i$ である。また、外的攪乱要因も内的攪乱要因も複数の要素が考えられるため、 $b_o N_o$ は $b_o N_o = b_{o1} N_{o1} + b_{o2} N_{o2} + \dots$ として、 $b_i N_i$ も $b_i N_i = b_{i1} N_{i1} + b_{i2} N_{i2} + \dots$ として表されるはずである。しかし、前述したように本稿ではこれらの要因を区別せず、これらの総和である bN だけを扱う。

ここで、認知主体が情報 X の真偽を判断できる可能性の程度を示す指標 V を考える。情報の真理と攪乱要因は独立している見なしてよいので、単純な線形的関係を仮定する。したがって、判断の指標 V は

$$(8) \quad V = aT - bN$$

として表現できる。言うまでもなく、認識度指数 a が小さい時や攪乱要因 $-bN$ が大きければ、認知主体の真理を捉える能力は低下し、認識度指数 a が大きく、かつ攪乱要因 $-bN$ が小さければ、認知主体が真理を認識できる可能性が高くなる。

4.2 真理値と想定確信度の関係

この真理値判断の指標 V は、言うまでもなく想定確信度に影響を与える。この関係は、指標 V と想定確信度のエントロピーによって定義することができる。まず、情報 X に関する判断の容易さは、情報 X と否定情報 $\neg X$ がどの程度大きく違っているかに対応する。この肯定情報と否定情報との区別の明確さは、両者の相対的な比率(オッズ比)の情報量(ロジット)によって表すことができるので、このことから真理判断指標 V は、

$$(9) \quad V = \log_2 \frac{P_n(x)}{P_n(\bar{x})}$$

と表現できる。ここで、 V に関する2つの関係式(8),(9)より、

$$(10) \quad \log_2 \frac{P_n(x)}{P_n(\bar{x})} = aT - bN$$

すなわち

$$(11) \quad \frac{P_n(x)}{P_n(\bar{x})} = \exp(aT - bN)$$

が成立する。この式(11)は、情報 X の確信度と否定情報 $\neg X$ の確信度とのオッズ比 $\frac{P_n(x)}{P_n(\bar{x})}$ が、真理値 T およびノイズ N と指数関数的な関係を持つことを意味している。さらに式(11)に $P_n(\bar{x}) = 1 - P_n(x)$ を代入すると、式(12)を得る。この式(12)は簡潔に式(13)と表現できる。

$$(12) \quad P_n(x) = \frac{\exp(aT - bN)}{1 + (\exp(aT - bN))}$$

$$(13) \quad P_n(x) = \frac{1}{1 + (\exp(-aT + bN))}$$

この式(13)が、情報の真理値が想定確信度に与える影響を表す関係式となる。すなわち、真理値と想定確信度は(攪乱要因を含んだ)ロジスティックな関係を持つ。この式(13)を用いて、次節で「真理」が想定確信度に与える影響について考察を行う。

4.3 真理値の想定確信度からの帰納

今、真理値の認識を攪乱させるノイズが一切ない世界、すなわち $N = 0$ を満たす世界を考える。これにより、真理が想定確信度に与える純粋な影響を見ることが出来る。 $N = 0$ より、真理値と想定確信度との間には、

$$(14) \quad P_n(x) = \frac{1}{1 + (\exp(-aT))}$$

という関係が成立する。ここで、認識度指数 a を標準的な値と考えられる $a = 3$ とすると、真理値と想定確信度は図 19 に示すような関係を持ち、真理値である「真 ($T = 1$)」「未知 ($T = 0$)」「偽 ($T = -1$)」は各々 (15) のような想定確信度に対応する。

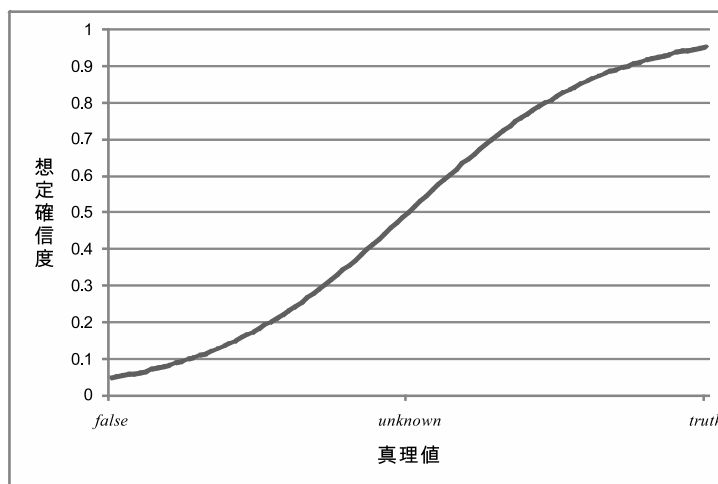


図 19: 真理値と想定確信度との対応関係

$$(15) \quad \begin{aligned} \text{a. 真} &: \frac{1}{1 + (\exp(-3))} = 0.953 \\ \text{b. 未知} &: \frac{1}{1 + (\exp(0))} = 0.5 \\ \text{c. 偽} &: \frac{1}{1 + (\exp(3))} = 0.047 \end{aligned}$$

真理値と想定確信度が一対一対応をするとは限らないため、想定確信度から真理値への推定は帰納推論にならざるを得ないが、(15) より、想定確信度が約 0.953 以上なら「真」、想定確信度が約 0.047 以下なら「偽」と見なしてよいだろう。この帰納推論を認めるのであれば、不完全な認知主体であっても「知識」を獲得できることになる。例えば図 2 でいうなら、情報 X が 53 回連続して出現した時に、想定確信度は 0.953 を超え、この情報 X を「知識」として獲得してよい。過去情報を考慮しない図 9 では、情報 X を 36 回連続して経験した時、想定確信度が 0.953 を超え、この情報 X を「知識」として獲得できることになる。

4.4 認識度指数の影響

前節では標準的な認識度指数である $a = 3$ の場合を考えたが、認知主体の内的な状態や対象となる事象の難易度によって、この認識度指数は変化し得る。図 20 は、攪乱要因

がない状況において、認識度指数が $a = 1, a = 2, a = 3, a = 5$ の場合に、真理値が各々どのような想定確信度に対応するかを示したものである。認識度指数が高くなるにつれて、真なる情報の確信度はより 1 に近づき、偽なる情報の確信度はより 0 に近づいていることが分かる。

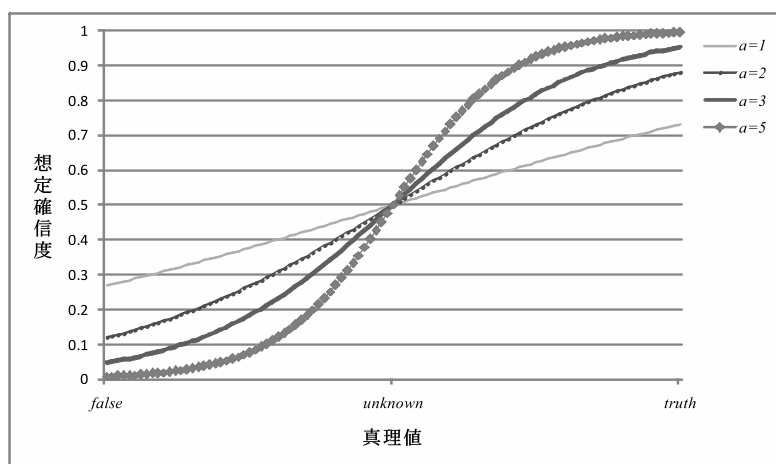


図 20: 認識度指数が想定確信度に与える影響

認識度指数が $a = 1$ の場合は、真理値 (あるいは情報値) と想定確信度が線形的に対応する。また、情報が真である時でも想定確信度は約 0.73、偽なる情報に対する想定確信度は約 0.27 となっており、想定確信度の変動範囲が狭い。このことは、どのような情報であっても想定確信度は比較的近似した値となり、情報の真偽判断が明確にできなくなることを意味している。なお言うまでもないが、認識度指数が $a = 0$ の場合は、いかなる情報に対しても想定確信度は 0.5 となり、真偽の弁別は完全に不可能となる。

一方、 $a = 5$ の場合は、真なる命題に対しては想定確信度が約 0.993 で、偽なる命題に対する想定確信度は約 0.007 となっており、真偽に関してほぼ完全な確信度が得られている。また、真理値が *unknown* (あるいは情報値が 0) に近い情報であるほど、ほんの少しの情報変化が想定確信度を大きく変動させ、真に近い情報あるいは偽に近い情報になるほど、情報変化が想定確信度に与える影響力が減少していることも特徴の一つである。この性質は、認識度指数が上昇するにしたがって、世界をなるべく安定した形で理解する能力も上昇することを示している。認識度指数が極めて大きくなると、 T が 0 以下なら想定確信度は 0 となり、 T が 0 以上なら想定確信度は 1 となる。すなわち、世界をほぼ完全に真か偽かに二分できるような認知環境が構築される。

4.5 攪乱要因の影響

現実の世界では、真理の認識を阻害する攪乱要因が何らかの形で存在すると考えられる。この場合は、以下に述べるように、前節で述べた想定確信度よりも低い値であっても、当該情報を真理と見なせるようになる。

今、攪乱要因が存在する状況 $T = 1$ を設定する。また、この攪乱要因は認識能力と比較して弱い妨害力しか持っていない状況であるとし、攪乱度指数 b を認識度指数 $a = 3$ の半分の比重となる $b = 1.5$ としておこう。この時、式 (13) は $P_n(x) = \frac{1}{1 + (\exp(1.5 - 3T))}$ という形になり、真理値と想定確信度は図 21 に示すような対応関係を持つ。

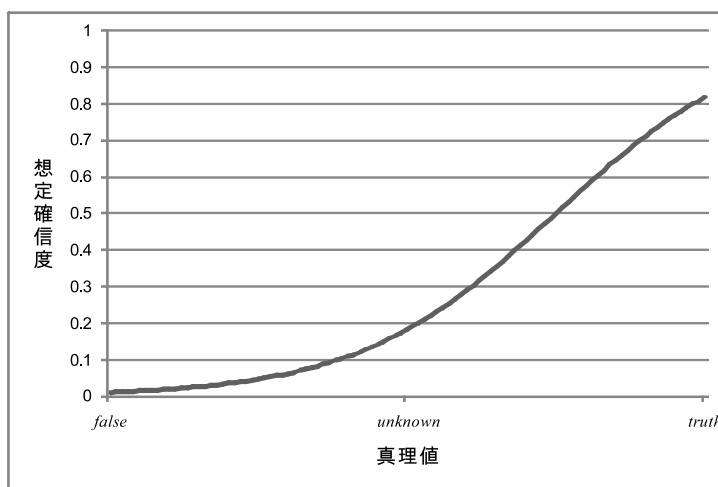


図 21: 攪乱度指数 $b = 1.5$ の状況における真理値と想定確信度の関係

図 19 と比較して明らかなように、攪乱要因が存在すると、真の命題に対しても認識主体はその情報を完全な形で受取することができないため、攪乱要因が無い場合に比べ想定確信度が低くなる。例えば攪乱度指数が 1.5 という状況下では、真なる命題であっても、その想定確信度は約 0.73 程度の値にしかならない。逆に、認知主体の側から見ると、攪乱要因の存在が自明な環境下であれば、想定確信度が多少低くても、帰納的に当該命題を真と判断し得る可能性を持つのである。

4.6 認識の攪乱要因とプラトンの問題

しかし、前節の結論はあくまで可能性に過ぎないという点に注意が必要である。現実世界に何らかの形で攪乱要因が存在するからといって、認識論的に低い想定確信度であっても常に真と判断してよいと言ってしまふことはできない。実際の環境においては、攪乱要因の存在が自明であるとは限らないし、攪乱要因の存在を意識できていたとしても、その影響力の大きさ(攪乱度指数の大きさ)は明確でないことが多い。最も大きな問題は、攪乱度指数が高くなればなるほど、想定確信度はほとんどの情報価に対して極めて低い値を取るようになってしまい、情報の違いを認識することすらできなくなってしまう点にある。この性質を見るために、図 22 に攪乱度指数が想定確信度に与える影響を示す。

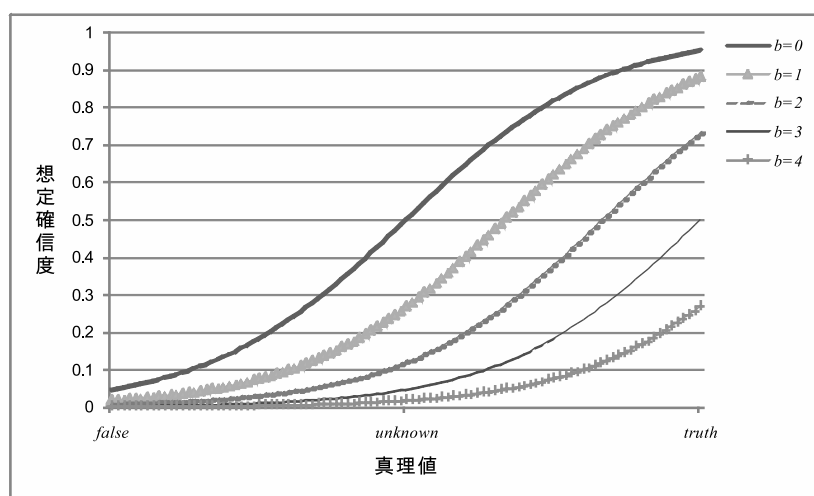


図 22: 攪乱度指数が想定確信度に与える影響 (認識度指数 $a = 3$)

例えば、攪乱度指数が認識度指数と同程度の $b = 3$ であるなら、完全に真なる情報であっても、その情報の想定確信度は 0.5 にしかならない。これは、攪乱要因が存在しない状況における *unknown* な情報に対する想定確信度に等しい。すなわち、攪乱要因が認識能力と等しいくらいの影響力を持っている環境で、認識主体が攪乱要因の存在に気がついていなければ、例え真なる情報であっても、認識主体は *unknown* な情報であると判断してしまうのである。攪乱度指数がさらに大きくなると ($b = 4$ など)、ほとんどの情報量において想定確信度は 0.1 を下回ってしまい、完全に「偽」なる情報と事実上区別がつかなくなってしまう。攪乱度指数がさらに高くなると、真の情報であっても、その想定確信度は偽の情報に対する想定確信度とほとんど違いがなくなってしまう、認知主体は何が正しい情報で何が誤った情報か、全く判断がつかなくなってしまうだろう。

この点を現実の問題に即して考えてみよう。図 22 から分かる通り、真実値が *truth* である情報に対応する想定確信度は、攪乱度指数の値によって大幅に変化する。しかし、*false* の情報に対応する想定確信度は一貫して 0.1 を切るような低い値しか取らない。逆にいうと、ほぼ 0 である想定確信度はどのような状況においても客観的事実として *false* である可能性が高いといえるが、客観的事象を常に *truth* と断定できるような想定確信度は存在しないのである。つまり、想定確信度という観点からすると、認識論的に *truth* よりも *false* のほうが確実なことが言え、誤りを犯しにくい。この認識的性質を明確に意識している興味深い例が、司法判断の大原則である「疑わしきは罰せず」というモットーに見られる。平和な時代の裁判の場では、冤罪だけは避けなければならない。すなわち、有罪判決は確実に罪を犯した人のみ出さなければならない。場合によっては真犯人に対して無罪判決を出してしまうことがあっても、無実の人に有罪判決を下してしまう冤罪だけは避けなければならない。ここから、裁判の現場で「疑わしきは罰せず」という大前提を立てる必然性が出てくる。有罪判決は、この大前提を *false* として棄却した時

のみに出されるものだからであり、想定確信度として確実性を持っているのは *truth* に対応する確信度ではなく、*false* に対応する確信度であるからなのだ。冤罪を避けるために有罪判決だけは確実でなければならない。そのためには仮説を *false* として棄却した時のみ有罪判決にならなければならない。だからこそ裁判の前提は「疑わしきは罰せず(被告は無罪である)」というモットーになっているのである。²

しかし、知識獲得は、真理値という観点から見るとこの裁判制度の話の裏返しになる。真理値が「確実に *truth* となる情報」が知識としての価値を持つのである。知識獲得の問題が本質的に困難である理由がここにある。我々の認識形態すなわち想定確信度の性質は *false* に偏った確実性を持っている。知識獲得に求められる性質とは逆なのだ。我々がほぼ確実に言えることは、「これは真の情報ではなく、したがって知識にはなり得ない」ということであって、「これは真の情報であり、知識である」ということを確実に言うことは認識論的に難しい。この世界がある程度良質で、認識を狂わせる攪乱要因が認識精度よりも低いものであることが保証されていない限り、知識獲得は不可能のように思える。

しかし、現実世界の攪乱要因が認識精度よりも低いという保証はない。そして、人間は「知恵」を持つ存在であり、事物を正しく認識できる能力こそが、人間を人間たらしめている性質である。では、我々はどのようにして知識を獲得しているのだろうか。ここで重要になる概念が、3.5 節で議論した「初期状態に拘束される想定」の存在である。本稿で提案した想定確信度・認知環境の更新アルゴリズムは、3 節で議論した通り、外界に存在する情報に応じて、比較的迅速に妥当な想定確信度を獲得し、また反例や例外情報に敏感な性質を持っている。しかし、3.5 節で見たとおり、初期状態に拘束される想定獲得はその限りではない。この場合の想定確信度の変更過程は、初期状態の想定確信度から大きくは変動せず、例外事象があったとしても、想定確信度の変動はある一定の範囲内に収まる。

既に述べたとおり、初期状態に拘束される想定の例としては、生得的知識や先入観、固定観念などが挙げられる。我々の持っている生得的知識や先入観が極めて悪質なものであるなら、経験を通じた知識獲得は正確には機能しない。しかし、それらが良質のものであるなら、この世界の攪乱要因の影響を逃れ、知識と呼べる想定を獲得できる可能性が残されている。Chomsky (1986) が強調する『プラトンの問題 (Plato's problem)』の重要性はここにある。「刺激の貧困 (poverty of stimulus)」は単に定量的な性質—言語データを大量に集めれば言語の本質が分かる—というものではない。「刺激の貧困」はより本質的な問題—我々の認識を阻害し得る外界のノイズおよび認識主体内部の愚かしさの問題—に関するものである。外界や認知主体内部に存在する攪乱要因が我々の認識精度を上回るものでなければ、データ数が少なくても、真の知識に近づくアルゴリズムを設計し得る。しかし、もし「刺激の貧困」が的を射た指摘であるのならば、Chomsky が一貫

²この議論で鍵になるのは、冤罪だけは犯してはならず、だからこそ有罪判決は「確実」に真犯人に対してのみ出されなければならないというところにある。これは平時にのみ通用する理屈かもしれない。戦争などの国家の緊急時には、「スパイを捕まえること」等の治安対策が重視されるので、「疑わしい奴は捕まえる」という前提を立て、この前提が *false* として棄却される「確実に疑わしくない人」にだけ無罪判決を出すこともあり得る。第二次世界大戦中の日本国内で起こっていたことを考えれば、この話の妥当性が分かるだろう。

して主張している「知識の生得性」が、認識論においても重要な意味を持つことになるのである。この「刺激の貧困」の問題を、認識主体の持つ認識度指数と攪乱要因の強度という点から捉え直すことも、本研究の今後の課題である。

5. 総合論議

Sperber and Wilson (1986) の提案している関連性理論は、認知能力の基本機能を関連性のある情報を見つけ出すことにありと見なしている。これは認知環境の改善を伴う情報のことであり、関連性のある情報が収集できれば、その情報間の推論演算から、ほぼ妥当な帰結を導きだすことができると考えられている。本論文で提案した想定確信度を更新式 (3), (4), (5), (6) は、こうした認知環境を改善していく上で必要となるアルゴリズムである。このアルゴリズムは次のような特徴を持つ。

- (16) a. 初期想定 の 確信度 の 強 さ に は 大 き く 影 響 さ れ な い 。
- b. 現 実 世 界 に 存 在 す る 情 報 の 性 質 に よ っ て 、 確 信 度 は そ の 時 々 で 更 新 さ れ て い く 。
- c. 反 例 事 象 に 比 較 的 鋭 敏 に 反 応 す る 。
- d. 過 去 情 報 を 考 慮 す る 学 習 と 、 過 去 情 報 を 考 慮 し な く て よ い 学 習 を ほ ぼ 同 一 の ア ル ゴ リ ズ ム で 扱 え る (後 者 の ほ う が 簡 略 化 さ れ て い る) 。
- e. 妥 当 な 想 定 確 信 度 の 獲 得 に 、 デ ー タ 数 や 初 期 想 定 は さ ほど 影 響 し な い 。
- f. 想 定 確 信 度 と 真 理 値 の 間 に 、 ロ ジ ス テ ィ ッ ク 回 帰 の 関 係 を 仮 定 す る こ と が で き 、 想 定 の み な ら ず 、 「 知 識 」 の 獲 得 が 可 能 と な っ て い る 。

この性質を、松井 (2009) に組み込めば、関連性の計算を行いながら認知環境を更新するプロセスをモデル化できるであろう。このモデルは松井 (2007)・松井 (2005) で考察した日常推論の性質や条件文の理解といった問題についても、より簡潔な解釈を可能すると思われる。この点については、また稿を改めて論じてみたい。

本研究は「プラトンの問題」を数理的に考察するための準備でもある。ここで提案したアルゴリズムは、既に述べたように比較的収束が早く、大量のデータを必要とせず、また初期想定の影響もほとんど受けない。しかし、それは認識を攪乱する要因の強さが認識精度を超えない限りにおいての話である。認識の攪乱要因、すなわち実世界のノイズや認識主体自身の持つ愚かしさがどの程度のものなのか、認識主体は完全には知ることができない。そうであるなら、本研究から言えることは次のいずれかである。

- (17) a. 認 識 主 体 は 想 定 確 信 度 か ら 事 象 の 真 理 値 を 帰 納 し て お り 、 真 理 を 獲 得 し よ う と す る 。 し か し 、 経 験 か ら の 学 習 だ け で は 外 界 や 認 知 主 体 内 部 に 存 在 す る 攪 乱 要 因 に 阻 害 さ れ 、 真 理 に 到 達 で き な い 。 し た が っ て 、 経 験 以 上 に 真 理 の 獲 得 に 重 要 と な る 要 素 が 存 在 す る 。
- b. 認 識 主 体 は そ も そ も 想 定 確 信 度 か ら 真 理 値 を 帰 納 し よ う と は し て お ら ず 、 真 理 と し て の 「 知 識 」 を 獲 得 し よ う と も し て い な い 。 持 っ て い る も の は 、 そ の 時 々 で 更 新 さ れ て い く 「 想 定 」 の み で あ り 、 こ れ は 経 験 か ら 更 新 さ れ て い く も の で 、 経 験 が 最 も 重 要 な 要 素 で あ る 。

(17) のどちらが正しいかは哲学的にも明らかではない。しかし私は (17a) が正しいと信じる。そして、もし (17a) が正しいのであれば、良質な生得的能力か、あるいは我々の認識を攪乱する要因の程度を遙かに上回る高い認識度指数を持つ認知能力のいずれかが知識獲得にとって必須の能力なのである。

参考文献

Armstrong, David M. (1973). *Belief, Truth and Knowledge*. Cambridge University Press, Chicago.

Ayer, A.J. (1981). 『知識の哲学』. 白水社.

Chomsky, Noam (1986). *Barriers*. The MIT Press, Cambridge.

Gettier, Edmund L. (1963). Is Justified True Belief Knowledge?. *Analysis*, **23**, 121–123.

松井理直 (2005). 計算論的関連性理論における日本語条件文の解釈. *Theoretical and Applied Linguistics at Kobe Shoin*, **8**, 95–122.

松井理直 (2007). 計算論的関連性理論に基づく日常的推論の分析. *Theoretical and Applied Linguistics at Kobe Shoin*, **10**, 45–76.

松井理直 (2009). 認知的関連性の単純かつ妥当な計算方法. *Theoretical and Applied Linguistics at Kobe Shoin*, **12**, 21–36.

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.

Author's E-mail Address: matsui@sils.shoin.ac.jp

Author's web site: <http://sils.shoin.ac.jp/~matsui/>