

大阪体育大学紀要  
第20巻 (1989) pp.171-184

# フレーム問題の諸相

齊藤了文

元年3月31日受付

Some Aspects of Frame Problem

BULLETIN  
OF  
OSAKA COLLEGE OF PHYSICAL EDUCATION

大阪体育大学紀要  
第20巻(1989) pp.171-184

## フレーム問題の諸相

齊藤了文

元年3月31日受付

### Some Aspects of Frame Problem

Norifumi SAITO

In this paper, I want to elucidate some aspects of frame problem. First, I look over the background, where frame problem appears. In Artificial Intelligence (AI), frame problem appears in the context of knowledge representation. And I pick up some presuppositions involved in this background. Secondly, "the solutions" of frame problem are presented. One of these "solutions" is related to nonmonotonic logic. This logic is the extension of classic logic and is so called common-sense reasoning. And I want develop the significance of frame problem. According to Hashida, the clue is the partiality of information. But this idea is crucial, only when view of knowledge changes. Traditionally, justification is the merkmal of knowledge. But AI presents the applicational view of knowledge. This view and the partiality of information will show us new looks of human and artificial intelligence.

フレーム問題は、差し当たり「行為が行われる時、どの性質が持続し、どの性質が変化するかを、計算論的に穏当な仕方で in a computationally reasonable manner 記述する問題」<sup>1)</sup>だと言うことができる。

「計算論的に穏当」という言葉さえそれほど厳密に考えなければ、なぜこんなことが問題になるか、あまりよくは分からないであろう。しかし実は、このフレーム問題は、人間の心について考えようとする哲学においても、また機械によって知能を実現しようとしている人工知能 (AI) においても、大きな関心がよせられている問題なのである。

なぜこのようなことが、問題にされるようになったのか(第一部)。また、この問題をもとにして、AIの内部でどのような方向に研究が進んでいるのか。また、どのような哲学的な考察が必要になるのか(第二部)。この論文では、このような事

柄について概観してみたいと思う。

#### I

フレーム問題が生じた背景を理解するために、人工知能研究の方法及び歴史の一部を概観する。まず、コンピュータを計算する機械と考える通常のイメージから、記号処理をする機械とみなすようになることが重要である。

「通常の考え方では、コンピュータは高速の数計算器である。この見解は部分的にしか正しくない。事実、デジタル型コンピュータは、情報の処理および変換のための、明確に定義された well-defined 過程を実行することができる、一般的な記号-処理装置 symbol-processing device である。」<sup>2)</sup>

ここでは二つの語句に注目する。まず、コンピュータは「明確に定義された」事柄を扱うということである。少なくとも、明確に定義されてい

ない事柄をコンピュータで処理することは非常に困難である。そして、次にコンピュータは「一般的な記号処理装置」とみなせるということである。コンピュータは合計を計算したり、科学技術計算をするためにももちろん用いられるが、それだけでなく、例えばファミコン・ゲームをするためにも用いられている。そして、ゲームにおいては、コンピュータの内部で用いられている信号は、結局は絵の移動や操作を表す記号として解釈されている。

### I. 1

まず、チェスをするプログラムを手掛かりとして人間の知的な行動を真似ようとした、人工知能の初期の試みを概観しよう。

「チェスは、非常に知的なゲームである。競争を曖昧にするような偶然性なしに、2つの知性を、互いにはなはだ複雑な形勢の中で対抗させるゲームであるから、2人とも、自分の形勢を、完全に理解することは期待できないが、十分分析はできるゲームであるから、互いに相手をより深く読もうとすることは可能である。・・・中略・・・もし、うまくできたチェスの機械を作ることができたら、人間の知的な努力の核心を見抜いてしまったように感じられるであろう。」<sup>3)</sup>

これが、初期における、人工知能に対する基本的な姿勢であった。それでは、このチェスをするという「知的」な課題は、どのようにして機械的に実現されるのであろうか。「シャノンが認めているように、チェスは有限なゲームである。チェスには、有限個の状態しかなく、おのおのの状態には有限個の駒運びがあるにすぎない。チェスのルールは、どの試合も終わりがあるということを保証している。すなわち、どのような状態も結局は勝ちか負けか引き分けかに至るのである。このように考えればチェスは、枝分かれした“木”で

完全に記述される。この“木”では、分岐点はある状態に対応し、枝はおのおのの状態からの選択しうる駒運びに対応する。この“木”の全体を眺め、そしておのおの選択の最終的な結果のすべてを見ることができるとしては、チェスは単純なゲームであるということは、直観的に明らかであり容易に証明できる。勝敗を決める最終の位置から出発して、ゲームをする人は、おのおの分岐点でどちらの枝が自分や相手にとって一番よいかを決めながら、自分の次の手を選択しようというところまで逆に進んでいくことができる。」<sup>4)</sup>

もちろん、この枝分かれした“木”の全体をたどることができれば、機械的にチェスの必勝法が見つかるはずである。これは、「理論的には」可能である。しかしシャノンは、「一世紀の間にこの“木”を調べるために利用できる時間は、 $10^{16}$ マイクロ秒より少ないのに、ほぼ $10^{120}$ ほど、しらべなければならぬやり方がある」<sup>5)</sup>という概算を行っている。従って、力づくで解法をもとめることは、組み合わせ的な爆発に陥り、「實際上」不可能なのである<sup>6)</sup>。

このように、問題の定義が明確であり、しかも解決のための手段も明確に定義されているにもかかわらず、目標とされる解決に至るためには、何らかの探索 search が必要となる、このような問題を機械を使って、効率的に解決しようというのが人工知能の方法であり目的であった。

### I. 2

さて、この方法をもう少し形式的に記述することにしよう。

チェスの場合を考えると、「初期状態 initial state」は、試合を始める時の盤面の状態であり、「目標状態 goal state」は、相手のキングが詰んだ状態である。初期状態から始めて、目標状態を

めざすのであるが、もちろん一足飛びに目標状態に至るわけではない。それで、駒を動かすことによって、目標状態をめざすのである。この場合、駒を動かすことは、或る状態を別の状態へと遷移させる「オペレータ operator」である。この3つの言葉を使って述べると、問題を解決するということは、初期状態を目標状態へ変換するオペレータの系列を求めることであると言える。別の例を使ってもうすこし詳しくみてみよう。

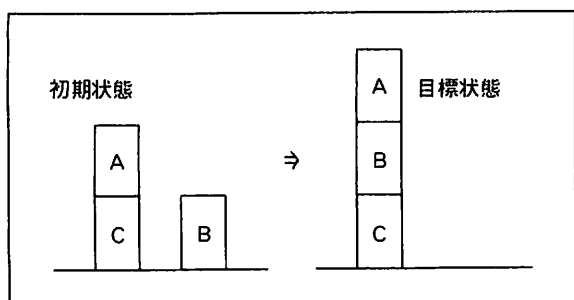


図 1

さて、上のような、テーブルの上にある積み木を用いて塔を作る問題を考えてみよう<sup>7)</sup>。物体 u が物体 v の上に載っているという関係を ON(u,v) と書くことにし、テーブルを TABLE と表記すると、図 1 の「状態記述 state description」は次のようになる。

初期 状態	ON(A, C)	ON(A, B)	目標 状態
	ON(B, TABLE) ⇔	ON(B, C)	
	ON(C, TABLE)	ON(C, TABLE)	

さて、オペレータの入力となる状態は、一定の条件を満足しなければならない。この条件のことを「前提条件 precondition」という。また、オペ

レータの適用によって状態がどのように変化するかを表すためには、適用によって成立しなくなる記述と、新しく成立するようになる記述を示せばよい。前者を「削除リスト delete list」、後者を「追加リスト add list」とよぶ。さて、図 1 でロボットが積み木を動かすと考えると、ロボットの動作として例えば次の 4 種類のオペレータを考える。

1. テーブルの上の積み木 x を持ち上げる。
2. 積み木 y の上の積み木 x を持ち上げる。
3. 積み木 x をテーブルの上に置く。
4. 積み木 x を積み木 y の上に置く。

ロボットが積み木 x を持ち上げているという関係を HOLD(x) と書き、「x の上には何もない」という関係を CLEAR(x) と書くことにする。すると図 1 の状態記述は、次のように書き直される。

初期 状態	ON(A, C)	ON(A, B)	目標 状態
	ON(B, TABLE)	ON(B, C)	
	ON(C, TABLE) ⇔	ON(C, TABLE)	
	CLEAR(A)	CLEAR(A)	
	CLEAR(B)		

そして前提条件とオペレータの効果は表 1 で表される。

また、「拘束条件 constraint」とは、目標を達成するさいに守らねばならない条件である。例えば、「x を持ち上げる」というオペレータを考えると、この前提条件として“x ≠ TABLE”を含めることができる。しかし、これを前提条件から取り除き、代わりに「テーブルを持ち上げることはできない」という拘束条件をつけることができる。「そうすれ

表 1

オペレータ	前提条件	削除リスト	追加リスト
1	CLEAR(x)	CLEAR(x), ON(x, TABLE)	HOLD(x)
2	CLEAR(x)	CLEAR(x), ON(x, y)	HOLD(x), CLEAR(x)
3	HOLD(x)	HOLD(x)	ON(x, TABLE), CLEAR(x)
4	HOLD(x) ∧ CLEAR(y)	HOLD(x)	ON(x, y), CLEAR(x)

ば、オペレータを適用する前に前提条件を調べる代わりに、適用した後の状態 HOLD(TABLE)がないことを確かめればよい。』<sup>8)</sup>

このようにして、「問題解決は、初期状態から出発して、適用可能なオペレータをつぎつぎに作用させて、拘束条件を守りながら目標状態に変換することである。問題の状態はその間に多くの状態を経由して変換される。これは、状態が状態空間の中を移動してゆくとみなすことができる。したがって、問題解決は状態空間の探索といえることができる。』<sup>9)</sup>

このような探索は、コンピュータを知的にしているといえる。つまり、普通のプログラムは、計算したり、集計したりしても、それは人間が予め考えておいた通りに（高速に）計算するにすぎなかった。それに対して、探索を行った上で、行動を行うとすると、最適な解が見つかった場合には、全く合理的で適切な行動をすることになる。これは、頭の中で計画を練ったあと、行動を行うというイメージに合っている。しかも、この行動そのものは、人間が直接に命じて行わせたものではないのである<sup>10)</sup>。

### 1. 3

そしてこの探索の方法は、ゲームという問題に限らず一般に人工知能の研究にとって基本的な方法となっている。例えば、定理の証明、文の解析、シーン解析などにおいても用いられている<sup>11)</sup>。

まず、定理の証明についてコメントを行う。この場合、前提から結論が導出されるかどうかを決定するための基本的な方法は、「諸前提」と「結論の否定」とをつけ加えた論理式の集合が全体として矛盾していることを示すことである。導出原理 resolution principle を使ってこれを証明するとき、意味の木 semantic tree を用いるなら、これは或る意味で探索の問題を含むことになる。そして

この導出原理を使って上に述べた積み木の例やその他の問題解決を行うことが可能である<sup>12)</sup>。

またシーン解析や文の解析においては、ゲームや問題解決とは違って、目標状態が何かが最初からははっきりしていないことが特徴である。

まず、シーン解析についてコメントをしておく。立体的な積み木を、線だけで描いた図がある場合、その線画の各線をどう解釈するかを決めるのが、シーンの解析である。例えば、それは境界線であるとか、凸を示す線であるとか、凹を示す線であるとかを弁別しようというのである。このことによつて、どのような積み木がどのように積まれているかが分かるようになるのである。ここで、シーン解析のウォルツのプログラムが成功した理由は、自然の拘束を利用したことにある。つまり、線画に現れる線分のつながり方のうち、論理的に可能なつながり方をもとに考えたのではなく、物理的に可能なつながり方を考えることによつて、線画の解釈を進めようとしたのである。組み合わせ可能性でなく、実現可能性を考えることによつて、探索の数を減らしたことが、彼のプログラムの成功につながったのである。

また、文の解析についても次のように言われる。

「一般に意味的な解析を無視し統語解析のみに限定すれば、数十語の単語からなる文を統語解析すると数百から数万の統語解析結果が得られることがある。・・中略・・このようにして、文が長くなるにつれて曖昧性の数の組み合わせが増え、莫大な数の統語解析結果が得られることになる。しかもそれらのほとんどが意味的に異常な解釈を強制するものであることも問題である。計算言語学の研究者の悪戦苦闘の歴史は、解析結果に含まれる曖昧性をどのように解消するかにあったと言つてよい。』<sup>13)</sup>

文の解析においては、例えばどの単語が動詞であり、どれがそれにかかる副詞であるかといった

ことを決定しようとする。そして、この場合も、組み合わせ的な可能性をただ求めることにはあまり意味はない。そのうちのどれが求める解析結果であるかを判定することが問題なのである。また意味や文脈なども使って、組み合わせ的可能性のうちの一部だけを（与えられた文脈の中でどの統語解釈が妥当かが問題である<sup>14)</sup>）を考えようとするのが問題となっていたのである。

探索が問題になる場合でも、すべての可能性を求めるのではなく、現実的な可能性を考えることによって、求める解答が、適当な有限時間内に得られることになるのである。

#### I. 4

さて、このように人工知能の基本的な研究方法を概観したので、これからフレーム問題がどのようにして登場するかを見ることにしよう。

まず注意しておくべきことは、論理的な推論や探索が必要とされているということは、ある問題について、問題となっているすべての領域の「事実」についてすべて知っており、その事実を必要に応じて適当に取り出せるという知識のモデルをとっていないということである。つまり、全知のモデルではない。更に個々の事実の全てを、一般規則を抜きにして個々別々に知っているとはできない。しかも、予測や計画においては「事実」になっていない、また「事実」になることのできない多くの可能性を取り扱わねばならない。従って少なくとも、「基本的な」いくつかの事実や規則を知った上で、後の事実はそこから「演繹」することで対応しようとするのである<sup>15)</sup>。

さて、現実の世界で行為するということを考えると、この点をもうすこし詳しく考えなければならない。「どの状況でも平均して6個の異なる動作があるのが普通で、そして望まれる目標を達成するためには4つの連続した動作が典型的であると

いうのであれば、 $64 = 1296$ の考慮すべき中間および終端状況があることになります。数千個以上の状況の1つひとつを定義する何千という事実を記憶することは、何百万を超える事実を記憶することを意味します。そしてそれは、少なくともすべての事実をすぐにアクセスできる記憶に持っておこうというのであれば、普通に使えるコンピュータではまだ不可能です。』<sup>16)</sup>

つまり、あらわれる状況を客観的に、文脈をなれて記述しようとするれば、それは実際上不可能であることが分かるのである。従って、「始めようとしている何千という動作の1つひとつがせいぜい3つの事実の変更を引き起こすのであれば、問題解決木における各節点の変更情報だけを記憶することはできそうです。』<sup>17)</sup> 従って、状況名をつけることによって状況を覚えておくことにすれば、あとはどのように状況が変更したかを書くだけで済むことになるように思われよう。

このとき、一つの動作によって、大抵の状況は変化しないと考えられている。従って例外だけを、変化したものだけを記憶することで十分だと考えられている。いわば私たちが変化を理解する場合には、全世界の或る瞬間でのすべての事態を次の瞬間のすべての事態と比較しているのではなく、他のすべては同じだとしておいて、例えば目の前のイヌだけが動いていると考えているのである。そしてこのような見方を表すような、知識表現を行おうとしているのである。

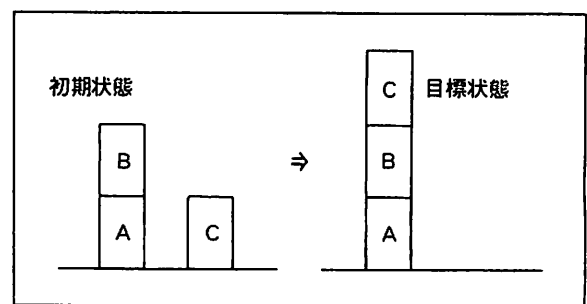


図 2

このことを少し形式化してみていくために、前に述べた積み木の例を少し変形して（図2参照）扱おう<sup>18)</sup>。行動する前の世界の状態と行動した後の相互関係を記述するために、述語論理式の中に状態を示す引数を入れ、初期状態を So とする。すると状態記述は次のようになる。

ON(A, TABLE, So)

ON(B, A, So)

ON(C, TABLE, So)

CLEAR(B, So)

CLEAR(C, So)

このとき、積み木の移動を行う操作 trans に関する知識は、次のように表現される。

$(\forall x) (\forall y) (\forall s) [\text{Clear}(x, s) \wedge \text{Clear}(y, s) \rightarrow \text{On}(x, y, \text{do}(\text{trans}(x, y), s))]$

ここで、関数  $\text{do}(u, v)$  の値は、『状態  $v$  において、操作  $u$  を行った後の状態』と考える。このとき、上の論理式は、『ある状態  $s$  において  $\text{Clear}(x)$  と  $\text{Clear}(y)$  が成立しているならば、その状態  $s$  において積み木  $x$  を積み木  $y$  の上に移動する操作  $\text{trans}(x, y)$  を実行した後の状態  $\text{do}(\text{trans}(x, y), s)$  においては、 $\text{On}(x, y)$  が成立する。』ということを表す。

この場合、目標状態  $(\exists s) [\text{On}(A, B, s) \wedge \text{On}(B, C, s) \wedge \text{On}(C, \text{TABLE}, s)]$  を証明すればよくなる。

しかし、ここに問題が生ずる。導出法を使って空節を導こうとしても、目標状態の否定およびその他の式からは矛盾が導出できないのである。つまり、

ON(B, A,  $\text{do}(\text{trans}(C, B), \text{So})$ )

ON(A, TABLE,  $\text{do}(\text{trans}(C, B), \text{So})$ )

がそれぞれ、ON(B, A, So) と ON(A, TABLE, So) と同じであることがわからないため、空節が導けず証明ができないのである。

この記述の意味を考えてみると、積み木 C を積み木 B の上に移動しても、「積み木 B が積み木 A

の上にある」「積み木 A がテーブルの上にある」という事実は変化しないということがわかる。すると、人間にとっては自明な事実の表現がいわば理解されていなかったことが問題にされるであろう。

このような失敗を防ぐために、例えば「ある積み木を移動させても、それとは異なる積み木どうしの上下関係は、移動後の状態でも変わらない」といったことを形式的に表現して導入する必要がある。

しかし、問題になるのは、このようなことが、様々な場面で生じているということである。つまり、「物を移動しても、サイズは変化しない」「物を移動しても、色は変化しない」「物を移動しても、その重さは変化しない」といった、当面の行為とは全く無関係な事実を表現しなければならないのである。こうした事実も表現しておかないと、求める結論が演繹されない場合があるのである。

或る行為によっては、大抵の状況は変化せず、枠（フレーム）に固定されていて、行為の効果はそれによって直接影響を受ける事実だけに関係し、その他のものは行為後の世界においてもそのまま成立すると考えると、この困難は避けることができるであろう<sup>19)</sup>。しかし、実は行為の副作用が起きることによって、簡単にはいかないのである。このため、この困難はフレーム問題と呼ばれるのである。

## II

フレーム問題がこのようにして生じてくるとすると、これに対していくつかの方向での対処が存在する。これはフレーム問題が生じた前提を考察することによって理解される。その前提のうち幾つかあげてみよう。

1. 結局はコンピュータが、明確に定義されたことしか扱えず、しかも記号処理を行うことによっ

てこのような問題が生じたのだ。

2. 計算量のある程度にしばることが、この問題の本質をなしている。
3. 明示的に述べていない大抵の状況は変化しないということを、うまく表現する手段がなかったために、このような困難が生じたのだ。
4. 基本的には時間的変化の表現がうまくいっていないことが問題である。

これらの問題点のうち後2者は、基本的に知識表現の問題である。つまり差し当たり「大抵の状況は変化しない」ということを表現しなければならない。しかもそのとき、副作用もうまく処理できるものでなければならぬのである。従って或る意味では技術的な問題に関係する。まずこの問題を見てみよう。

## II. 1

例えば、図1に関して述べたように（STRIPSで使用された）削除リストと追加リストを用いる方法がある。しかし、それでも現実の世界を考えると、この方法ではうまくいかない。記述の量が非常に増えてしまうのである。

「明示的に記述してあるものだけが変化し、それ以外は何も変化しない」ということを表現するということが、そしてそれを論理的演繹と調和させることがまた目指された<sup>20)</sup>。これをめざそうとしたのが、マクダーモットとドイルの「非単調論理」、レイターの「デフォルト論理」マッカーシーの「サーカムスクリプション（極小限定）」である。（これらは、その性質の共通性のため、一括して非単調論理と言われることがある。）例えばサーカムスクリプションとは、何が変化し、何が変化しないかを、個別的に明示的に述べるのを避けるために、「結論にジャンプする」<sup>21)</sup>推論規則である。すなわち「対象Oが、事実Aから性質Pを持つと推論されたら、性質Pを持つ対象物はOに限る」と

いうメタな推論規則である<sup>22)</sup>。

これは或る意味で常識的推論をモデル化している。従って、例外だけを明示的に書いておき、普通は一般的な規則に従ってものを考えるということモデル化しているのである。図2について述べたように、このような考えこそ、通常の論理体系では証明できないことなのである。古典論理を用いて結論を得ようとするならば、すべての状態と行動に関して、変化するかどうかの完全な情報を与えねばならない。これに対して、「常識はある種の情報の状態に関係し、ある情報を手にして別の情報はないときに目的を達する能力である」<sup>23)</sup>。このように考えられた限りでの「常識」を形式的に表現することが問題になっているのである。これを形式化しようという試みが、前提が増すに従って結論が増す「単調論理」（古典論理はすべてこの性質をもつ）に対して、新たな情報が得られたために今まで得ていた結論を廃棄することを可能にする「非単調論理」なのである。

しかし、このサーカムスクリプションにおいては、意図しないモデルが出てくることがある。この実例がイェール射撃問題と言われるものである。そしてこれをどう扱うかということについてさまざまな考察がなされている<sup>24)</sup>。しかし少なくともまだ、フレーム問題の解決として一般に認められているものはない。

また、イェール射撃問題は、「デフォルト推論を時間の推論に形式的に適用すると直観に反する事実を推論する」<sup>25)</sup>ことであるとも言われる。ここで問題になるのは、「時間の推論」つまり、時間に伴う世界の推移の表現と推論を行うことである。古典的な「時制論理 temporal logic だけでは、変化についての推論の方法を与えない」<sup>26)</sup>のである。つまり、理想化された世界ではなく、現実の複雑な環境を考えると、未来を予測する際の、効率と信頼性をどう調和させて推論するかということ考



えにいれなければならないのである<sup>27)</sup>。

このようにフレーム問題は、論理的、技術的方面では、「常識的推論」をどう定式化するかという問題や時間的变化についての推論をどう定式化するかという問題へと展開しているのである。

## II. 2

次に、単なる技術的な問題を越えた一般的な問題を考察しよう。これは、前に述べたように、記号処理をしているということと、情報処理の量的限界の問題である。哲学的議論は多数行われているが、ここではその一例を部分的に指摘するにとどめる。

フレーム問題の生じた理由を、記号処理自体、つまり表象主義と計算主義に見るみかたがある（例えばドレイファス）。そして、人間とAIの相違にのみ着目して議論を進めている。つまり、人間はフレーム問題を解決しているにもかかわらず、AIはそれができていない。しかも、すべての知識を形式化する限り、AIは、現に解決できていないだけでなく、原理的にできないのだと主張するのである<sup>28)</sup>。このように人間は機械ではないということを主張するために、多くの哲学的考察は行われてきたのである。

## II. 3

さて、この小論では、別の側面から問題を見ていくことにする。もちろん、AIは人間の知性のモデルたろうとしている。しかし、モデルである限りは、人間の心との相違は常に出てくる。従って、人間の心の問題を現在のAIがすべて解決したと考えると、どのような欠点が見えてくるかという問いをたてるよりも、少なくとも現在のAIで説明している部分についてそれと人間の知性との共通の問題状況を考察する方がより有益ではないかと思われる。つまりAIが人間の意識を実現でき

るかできないかという、人間の「心の本性」に係わる問題はさておいて、人間にとって使いやすいコンピュータを作るという面でも、フレーム問題は考察する価値があると思えるのである。

このときまず考慮したい主張は、情報の部分性という主張である。「フレーム問題の本質は、巨大な全体のうちのごく部分的な情報しか記述しないし処理できない」<sup>29)</sup>ということにあるという主張である。これは、上で、「組み合わせ的な爆発」とか「實際上不可能」とか言ったことに対応する。実際に動くシステムを作り上げようとするときに、情報の部分性は常に顧慮しなければならない条件である。

「情報の部分性に対処することは、世界に関する莫大な量の知識を部分的に取り込み、さらにそれを部分的に処理するだけで、大部分の場合にうまくいくようにする、ということである。その場合、情報を部分的にしか参照していないから、間違いが起こり得ることを覚悟しなければならない。そして、その誤りによる被害を最小に抑えるということが、即ちフレーム問題を「解く」ということなのである。この意味において、フレーム問題とは程度と効率の問題であり、形式論理の問題ではない。また、この問題の深遠さは、人間の知能の柔軟性が、誤りを犯し得るという代償を払うことによって初めて可能となっている、という点にある。よって、フレーム問題は確かにAIの一般問題とならざるをえない。」<sup>30)</sup>ここでは、前節で述べた技術的問題、知識表現の問題を越えて、フレーム問題を一般的に解釈すべきことが述べられている。

フレーム問題をどのような範囲の問題と認めるかという事は、もちろん多くの議論があるところである<sup>31)</sup>。しかし、その定義の価値は、そこからどのような洞察が得られるかに依存すると考えても、発展途上の学問にとっては許されるのではな

いかと思う<sup>32)</sup>。この意味で、情報の部分性に着目することは重要だと考える。

## II. 4

さて、人間における情報の部分性、知識の有限性については、哲学や宗教の分野でしばしば論じられていることと同じように思えるかもしれない。なぜ、フレーム問題に結びついて出てくる「有限性」「情報の部分性」が、哲学的に陳腐な主張に留まらず、新たな観点を提供するかを考えなければならぬ。

そのためにまず、「フレーム問題」が初めて登場したマッカーシーとヘイズの論文を見てみよう。そこにおいては形而上学的に十分な表現と対比して、認識論的に十分な表現を求めることが主題となっていたのである。そして、そのような問題設定の下で、「状況」と状況を定義域とする関数「フルーエント」を導入したのである。そして結局フレーム問題が生じたのである。

さて、形而上学的に十分な表現とは、例えば「一組の粒子間にそれぞれ働く力を通じて相互作用をしている、粒子の集まりとしての世界の表現」とか「離散的オートマトンが相互作用するシステムとしての表現」である。これは世界について一般的に、原理的には表現できるということを意味している。認識論的に十分な表現には、「世界の局面に対して現実に持っている諸事実を表現するためにそれを実際的に用いることができる」<sup>33)</sup>ことが必要である。例えば、「ジョンは家にいる」とか「犬はネコをおいかける」といった事実を表現することができなければならない。

彼らはこのような区別を行って、認識論的に十分な表現の必要性を述べるのである。それでは、この区別はどのようなことを含意しているのだろうか。

古典力学におけるラプラスの魔にとっては、形

而上学的に十分な表現で満足するであろう。そしてそのとき、副作用や例外を特定できないという意味での「フレーム問題」は生じない。というのは、世界のすべての状態を微分方程式と初期値で表現しようというのがラプラスの魔で代表される立場であり、この場合、世界のどこで何が起きているかということは、粒子レベルですべて分かっていることになるからである。このときは、例外を考える必要はない。形而上学的に十分な表現で満足すれば、変化しないものの記述にはあまり患わされないで、少なくとも近似値という仕方で世界のすべてを理解していると主張できるように思われる。そして、このような方向に進むことは、世界をある意味で物体からなる一様な世界とみなすという、物理学の基本的考えを述べたものとなる。

このように考えると、(副作用や例外と関わる)フレーム問題には、記憶や探索量の限界という知識の有限性以外に、常識的表現を導入する必要があるという「認識論的に十分な表現」という条件が関与しているように思える。そしてこの場合には、設計者が理解できないようなプログラムはコントロールできないため危険である<sup>34)</sup>という基本的な立場に立って、知識表現が考えられているのである。つまり、機械の暴走を許さず、機械を支配しようという意図にもとずく、知識表現である。

認識論的に十分な表現は、私たちの意図や行動の目的といったものが自然に表現できるとか、コンピュータを道具と考える限りその「ふるまいや思考」をある程度予測し制御できるという利点をもっている。しかしこの表現はそれだけを取り出すと欠点ももっているのである。そしてそれを端的に言い表すと「内的処理」<sup>35)</sup>を忘れていくということになるであろう。

火星は太陽の回りをまわっている。しかし、火星はその運動方程式を計算しながら運行している

わけではない。人間の知能を模擬しようとしている AI システムにおいても同じことが言える。「ここには、模擬する主体 (AI 研究者、認知科学者) が解釈を与える記号と模擬対象 (人間、エキスパート) が実際に持つ記号という、2 種類の記号が現れる。外側の記号、内側の記号といってもよい。」<sup>36)</sup>

私たちの意識的思考においては、この 2 種類の記号は非常に近い。しかし、このような区別に対して、内側の記号も結局はコミュニケーションできて初めて意味を持つことから、内側の記号や内的処理に関与する必要はないと言われるかもしれない。しかし例えば「関連知識の想起」における計算といった問題を考える場合には、ある事象が別の事象によって想起される「され易さの度合」とか、ある特定の外的状況に置かれた場合の特定の知識の想起の「され易さの度合」といったような、ある意味で内的な計算の状態を表示する枠組みが不可欠になる。<sup>37)</sup> 意識的、自覚的思考は関連知識の想起が生じた後で問題になるにすぎない。その意味で内的処理を特に取り出して考察することには意義があるだろう。

しかしそれに留まることなく、「内的処理に関与しない論理派の「機械の知」は、関連知識の想起 (フレーム問題)、定式化の切り換えといった典型的な問題に遭遇して、決定的に破綻する。これらの問題は、いずれも、ブライテンベルグのモデルが示唆するように、内省できない「必然的な情報処理」が関与している。」<sup>38)</sup> このように辻井は述べるのである。この主張は恐らく正しい方向を示しているように思えるが、少なくとも「認識論的に十分な表現」だけではうまくいかないということは、認められるように思える。しかも、論理的定式化だけでなく、その「処理」を特に考えようという主張は重要だと思われる。

## II. 5

もうひとつ別の側面から、情報の部分性を考えてみる。さて、「AI とは、人間が扱えるような部分情報問題を扱おうとするものである」<sup>39)</sup> と言えるとすれば、この「人間が扱える」という部分を説明する必要がある。ここでは「有限的条件を持っているからこそ知的能力の発揮ということに意味がある」<sup>40)</sup> ということに注目すべきであろう。これは人間の知性、人間が知的であるということ、どのように理解するかということに関係している。

さて、哲学の伝統においては、知識とは基礎づけをもった真なる命題というように考えられていた。基礎づけがない、ただの主張は、とても知識に値しないと考えられていた。科学に関しても、経験に基づく正当化をもつかどうかを重視されていた。どのようにしても疑えない真理を知識だと理解してきた。このような思考法は、多数の民族の間の折り合いをつけるために、契約や法律を正当化の根拠とする社会においては、また基礎づけの究極の根拠として絶対の一神を認めようとする社会においては、なかなか有益であったであろう。

さて、正当化という観点で求められるものは、結局は普遍的な法則や規則である。自然科学においても、このことはいえる。しかし、その法則が与えられた時に、現実の事例においてその法則をどう適用するかという問題が残っている。この適用が自明でないこと、それを考察するのが AI の研究ではないだろうか。(法則がみつからず、曖昧な状況を研究するということは、AI だけでなく、あらゆる発展途上の科学が行っていることである。) 自然科学はそれ自身の法則を求めて研究する。それに対して AI は、その法則を現実の様々な状況のもとで、どのようなときにどう使えばいいかを特に研究する。AI (特にフレーム問題) が教

えてくれることは、このような、知性のいままであまり省みられなかった側面である。そして、この観点に立って初めて、有限性、情報の部分性は問題として重要な意味をもつことになるのである。部分的な情報をもっているにもかかわらず、私たちは、この複雑な世界でうまく生きている。動物にもできていることとして、知性を理解する際に軽視されていたことが、あらためて注目され、考察、研究を要することとされるのである。このような知性の姿を探求する AI は、科学の成果を実際に使えるようにすることを目指すという意味で、純粋応用科学という特徴づけが可能かもしれない。

ともかく、正当化という観点でのみ知識を理解することは、著書の集合として、データベースとして図書館にある本そのものが知識であるということにもなりかねない。少なくとも、私たちにとっては、知識のある人というのは、いつでも適当なときに、適切なことを教えてくれるものだという思いがある。つまり、以上のような知識を「使える」<sup>41)</sup>かどうか、知的であるかどうかの分かれ目になっているようである。そしてこの「適用可能性」「応用可能性」が自明<sup>42)</sup>であれば、知識を使えるということをそれほど問題にする必要はないであろう。そして、私たち人間にとっては、ある程度の目標や一般的な規則といった知識が与えられれば、その実現の「原理的可能性」は明らかだと思われる。しかし、この適用可能性は、実は自明ではないのである。まず注目すべきことは、人間の有限性であり、情報の部分性である。そのため規則の適用の仕方の考察が必要になったのである。ヒューリスティックを使うということも、規則のすべてを予め試してみることはできないからなのである。そして更に、認識論的に十分な表現を使うということは、その表現の適用が自明だということを含んでいるわけではないのである。そ

こにおいて、内的処理や計算過程の研究が必要なのである<sup>43)</sup>。そしてフレーム問題はこの点までも示唆しているように思える。

適用の自明性に関して、もう一つ付け加えておく。たとえば演繹においては、推論の必然性が言われる。これを見ると、規則の適用はなぜ自明だといえないかと思われるかもしれない。しかし、推論の各段階において、どの推論規則と公理または定理を利用するかは、証明を試みている場合には予め分かっている。前提の真理性が結論の真理性を導くという必然性は、規則の適用の非自明性とは無関係なのである。

## II. 6

このように、フレーム問題から展開される諸考察を概観してきたが、最後に規則の適用に関して一つ確認をした上で、以上の論点をまとめることにする。

規則の適用という言い方は、トップ・ダウンのものの見方になっている。つまり、主体が規則のものに付与するというイメージがある。それに対して、目の前に広がるシーンに対しては、規則の適用とは、そこから、或る与えられた概念にあてはまるものを取り出すこと、それが何であるかを理解することになる。ボトム・アップの言い方をすると、見えているものを類別することになる。類別と適用とは、ある場合には、同じ事柄の裏表になるのである。「規則の適用」という言い方だけを取り出すと、主体の能動性が強調されたり、或る場合には論理的演繹をイメージするかもしれないので、この点を付言しておく。

さて、以上の議論を要約すると、フレーム問題は、その本質的な論点として、人間の有限性、情報の部分性という主張を含んでいる。しかし、この主張が特に意味をもってくるのは、「認識論的に十分な表現」を求めようとした枠組みにおいてで

あった。そして、この枠組みは、内的処理を無視しているという欠点があるとも言われていた。少なくとも、この内的処理を問題にするということは、知識をどのように使っていくかということに関係している。同じ結論は、知識をどう捉えるかという問題にも関係している。正当化に基づく知識観においては、フレーム問題はロボットの作成にのみ係わる問題だと思われやすい。図書館そのものでなく、それを使えるというところに知識の姿を見る場合に、しかもこのとき、人間の有限性、情報の部分性という条件が関与することによって、初めてフレーム問題が、私たちにとっても重要な問題だとみとめられるようになると思う。そして、使えることと、内的処理との結びつきは、我々の意識的判断の必須の構成要素が、無意識的なもの（もちろんここでは、フロイトの意味で抑圧されたものではない）であることを示唆しているように思える。

このように、これらの解釈は、結局は「規則の適用」ということの解明が必要だということを示している。そしてこれらは、人間の技能、カントのいう判断力とも関係する。そしてこれらは、人間にとっては熟練によって達成されると言われている<sup>44)</sup>。AIの立場からも、また人間の知の解明という点からも、規則の適用という知の有り方を、何からのモデルや仮説を建てることによって研究していくことが、私たちの今後の課題となるであろう。

#### 注

- 1) Frank M. Brown : Introduction p. v in "THE FRAME PROBLEM IN ARTIFICIAL INTELLIGENCE Proceedings of the 1987 Workshop"
- 2) Edward A. Feigenbaum & Julian Feldman eds. : Introduction p. 1 "Computers and Thought"阿部, 横山監修『コンピュータと思考』p.3
- 3) Allen Newell, J. C. Shaw, H. A. Simon : Chess-Playing Programs and the Problem of Complexity p. 39 in "Computers and Thought"阿部, 横山監修『コンピュータと思考』p.56
- 4) Allen Newell, J. C. Shaw, H. A. Simon : Chess-Playing Programs and the Problem of Complexity p. 42 in "Computers and Thought"阿部, 横山監修『コンピュータと思考』p.60
- 5) Allen Newell, J. C. Shaw, H. A. Simon : Chess-Playing Programs and the Problem of Complexity p. 43 in "Computers and Thought"阿部, 横山監修『コンピュータと思考』p.61
- 6) Marvin Minsky : Steps toward Artificial Intelligence p. 408 in "Computers and Thought"阿部, 横山監修『コンピュータと思考』p.531f.
- 7) ここでの説明は、『岩波講座 情報科学 22 人工知能』白井良明・辻井潤一による。
- 8) 白井良明・辻井潤一：『岩波講座 情報科学 22 人工知能』p.17
- 9) 白井良明・辻井潤一：『岩波講座 情報科学 22 人工知能』p.18
- 10) 辻井潤一：『知識の表現と利用』p.4f.参照。
- 11) 例えば, Patrick Henry Winston : Artificial Intelligence second edition p. 87 を参照。
- 12) 白井良明・辻井潤一：『岩波講座 情報科学 22 人工知能』第5章3節参照。
- 13) 田中穂積：計算言語学のすすめ p.9 in『数理科学 1989/3』
- 14) 辻井潤一：計算言語学と理論言語学——統語論とパーサ研究の交流——p.287 in『人工知能学会誌 1988/5』

- 15) J. McCarthy and P. J. Hayes : Some Philosophical Problems from the Standpoint of Artificial Intelligence p. 33 in “Readings in Nonmonotonic Reasoning”三浦訳『人工知能(AI)の観点から見た哲学的諸問題』in『哲学2』p.216
- 16) Bertram Raphael 溝口など訳：『考えるコンピュータ』p.163
- 17) Bertram Raphael 溝口など訳：『考えるコンピュータ』p.163
- 18) 以下の記述は、辻井潤一：『知能の表現と利用』p.192ff.に依存する。
- 19) J. McCarthy and P. J. Hayes : Some Philosophical Problems from the Standpoint of Artificial Intelligence p. 38 in “Readings in Nonmonotonic Reasoning”三浦訳『人工知能(AI)の観点から見た哲学的諸問題』in『哲学4』p.136
- 20) 古典論理の不十分性と非単調性の導入の必要性については、Matthew L. Ginsberg : Introduction p. 5f. in “Readings in Nonmonotonic Reasoning”を参照。
- 21) J. McCarthy :Circumscription ——A Form of Non-Monotonic Reasoning p. 466 in “Readings in Artificial Intelligence”
- 22) 中川裕志：論理+サーカムスクリプション=常識推論 p.14f. in『人工知能学会誌 1987/3』
- 23) J. McCarthy：コンピュータに常識を与える p. 684 in『科学 1987/11』
- 24) 例えば、J. McCarthy：コンピュータに常識を与える in『科学 1987/11』、佐藤健：Yale Shooting問題とその解決へのアプローチ in『人工知能学会誌 1988/2』、松原仁・山本和彦：フレーム問題、非単調論理、イェール射撃問題の間の関係についての一考察 in『人工知能学会誌 1989/1』などを参照。
- 25) 新田克己：時間を含む知識の表現と推論 p.567 in『人工知能学会誌 1988/9』
- 26) Yoav Shoham : “Reasoning About Change” p. 95
- 27) Yoav Shoham : “Reasoning About Change” p. 16ff.
- 28) Hubert L. Dreyfus & Stuart E. Dreyfus : Mind over Machine 邦訳『純粹人工知能批判』など参照。
- 29) 橋田浩一：AIとは何でないか——情報の部分性について p.907 in『bit 1988/8』
- 30) 松原仁・橋田浩一：人間におけるフレーム問題の解決不能性について in『人工知能基礎論研究会(第5回)発表資料 1989/2』
- 31) Zenon W. Pylyshyn ed. : The Robot’s Dilemma, Frank M. Brown : The Frame Problem in Artificial Intelligence また、松原仁・山本和彦：フレーム問題について in『人工知能学会誌 1987/9』などを参照。
- 32) 松原仁：AI研究者は今日も夢みる p.229 in『現代思想 1989/3』
- 33) J. McCarthy and P. J. Hayes : Some Philosophical Problems from the Standpoint of Artificial Intelligence p. 29 in “Reading in Nonmonotonic Reasoning”三浦訳『人工知能(AI)の観点から見た哲学的諸問題』in『哲学1』p.258f
- 34) J. McCarthy and P. J. Hayes : Some Philosophical Problems from the Standpoint of Artificial Intelligence p. 28 in “Reading in Nonmonotonic Reasoning”三浦訳『人工知能(AI)の観点から見た哲学的諸問題』in『哲学1』p.255
- 35) 以下の議論は、辻井潤一：機械の知 in『機械の知 人間の知』を基にしている。
- 36) 辻井潤一：機械の知 in『機械の知 人間の知』p.42
- 37) 辻井潤一：機械の知 in『機械の知 人間の知』p.50
- 38) 辻井潤一：機械の知 in『機械の知 人間の知』p.50
- 39) 橋田浩一：AIとは何でないか——情報の部分性について p.898 in『bit 1988/8』
- 40) 拙論：人工知能研究と人間の有限性 p.215 in『大阪体育大学紀要 第18巻 1987』

- 41) ヴォルフの学問観と対比してカントの学問観を見てみると、カントは証明を暗記することでは尽くせない学問観を提示している。これは、正当化とは違った、現実の複雑な状況に対処していこうとする理性を考えていることを意味する。拙論：カントの学問観 in『関西哲学会紀要第20冊1986/3』参照。
- 42) この論点については、フレーゲが心理主義批判によって、論理学の独立性（自己正当化）を論じたときにも、残存する問題点であると指摘しておいた。拙論：認識と心理主義批判 in『大阪体育大学紀要第16巻1985』参照。
- 43) これこそ、辻井潤一：機械の知 in『機械の知人間の知』の基本的主張である。
- 44) I. Kant : Kritik der reinen Vernunft B. 172