

部分情報とその情報構造

Partial Information and its Information Structure

鈴木 志 元
Yukimoto Suzuki

Résumé

Almost all information systems have their own logic. A lot of logical systems have been created to capture various phenomenon. Classical logic such as Propositional Logic is the representative example of logical systems. And there is the one which was created to complement the defect of classical logic. It is called Non-Classical Logic. The main example of non-classical logic is called Intuitionistic Logic which was created to treat partial information. In this paper, we explain intuitionistic logic briefly, then discuss the information structure based on intuitionistic logic. And finally, we introduce intuitionistic-information retrieval system.

This paper is the preparation for developing the general theory of partial information and the theory of information retrieval systems which can treat partial information.

- I. はじめに
- II. 部分情報と情報検索
 - A. 情報検索における不確定性
 - B. 翻訳の不確定性とフレーム問題
 - C. 論理的アプローチ
- III. 情報の論理構造
 - A. 直観主義論理
 - B. 直観主義論理の代数的定式化
 - C. Kripke 意味論
 - D. 部分情報と曖昧性
 - E. 直観主義論理と位相構造
- IV. 情報検索の直観主義的定式化
- V. おわりに

鈴木志元：近畿大学，大阪府東大阪市小若江 3-4-1

Yukimoto Suzuki: Kinki University, Kowakae 3-4-1, Higasiosaka, Osaka

1995年11月9日受付

I. はじめに

情報検索システムの背後には、何らかの論理がある。あるいは、すべてのシステムの背後には何らかの論理があるといってもいいだろう（ただし、ここでいっている論理とは、いわゆる数理論理学において明確に定式化されているものだけをさしているのではない）。例えば、ブールの情報検索システムは命題論理という論理体系に基づいている。ブールの情報検索システムについては様々な議論がなされてきたが、それらの議論で指摘されてきた長所と短所は、概ね、命題論理の長所と短所によるものである。このように論理は検索システムを基本的に支配しているといえるだろう。

ブールの検索システムの短所を補うために、様々な情報検索システムが提案されてきた。例えば、それらの中でも論理を意識したものとして、ファジー情報検索モデルがある。これは情報検索における情報の部分性ないし不確定性を扱おうとするものである。情報の部分性を扱える論理体系はファジー論理だけではない。ファジー論理が登場するはるか以前から多くの試みがなされ、多くの論理体系が提案されてきた。古典論理の不備を補い拡張しようとして提案された論理体系は非古典論理と呼ばれているが、このような非古典論理の代表と言えるものが直観主義論理である。

本稿では、第II章で部分的情報についての論理的議論の必要性を論じ、第III章では直観主義論理にもとづく情報のとらえ方や情報構造について説明する。さらに第IV章では直観主義論理にもとづく情報検索システムを紹介する。

II. 部分情報と情報検索

A. 情報検索における不確定性

Popperによれば、知識は推測からはじまる。推測を仮説といってもいいし、理論といってもよい。理論をテストし、より厳しいテストに耐える理論におきかえることによって科学的理論は成長する。いくつかのテストをパスした理論でも絶対的な真とはみなされない。かくれたエラーがあるかもしれないからである。このように理論はいつ

までも真と確定されることはなく、推測的であり続ける。科学の必要条件は、客観性でもなければ真実でもない。それは、体系的に自己を批判し続けるという態度である。

またPopperは3つの世界（世界1：物理的世界、世界2：主観的知識の世界、世界3：客観的知識の世界）を想定した。情報検索において、適合情報がもれていないことを確認する唯一の方法は、記録された情報をすべて直接に調べること、すなわち世界3をすべて調べることである。このような作業に終わりはない。調べている間にも新しい情報が作られているからである。どんなに注意深く検索を行ったとしても、常に新しく情報が生まれてくるために、情報検索は決して完全性に到達することはなく、必然的に不確定で、いつまでも閉じることはない。すなわち、情報検索は本質的に推測なのである。

以上はSwanson¹⁾が述べた情報検索における本質的不確定性の議論の概要であるが、この他にも情報検索における不確定性の議論は数多くある。一般的に言えば、人間に関わる限り不確定性は存在すると言えるのかも知れない。例えば、図書館を考えてみれば、情報伝達行為に関わる人間として索引者と仲介者（図書館員）と利用者があり、それぞれが不確定性の発生源になると考えられる。索引者の非一貫性がそうであり、利用者のASK（変則的知識状態）やラベル効果などがそうである²⁾。

B. 翻訳の不確定性とフレーム問題

前節では情報検索に関わる不確定性について述べたが、理論や用語をめぐって根本的な不確定性のあることが哲学や人工知能で問題となっている。

第一に取り上げなければならないのはQuine³⁾による「翻訳の不確定性」であろう。これは、「異なる言語の間の翻訳には、必然的に不確定性が伴い、翻訳は原則的に一義には定まらない」というものである。未開民族の言語を調査しようとする研究者が根底的翻訳を行うにあたって利用できるデータは、観察可能な刺激と観察可能な行為の2

つのみであって、これらから翻訳の手引きを一通りに決定することはできないというのである。つまりウサギが飛び出したとき原住民が「ギャバガイ！」と叫んだとしても、それがウサギを意味した言葉なのかどうかはわからない。ウサギを意味するのであろうという推測をテストする事はできるが、いつまでも確定はできないというのである。

異なる言語(ないし理論)の間での翻訳(理解)が不可能であるなら、そこで用いられている用語について統一的な理解が得られるはずもない。つまり、用言の統制や翻訳はあるパラダイムのなかでのみ可能なものであって、統一シソーラスなどというものは原理的に不可能である、ということになる。ということは、情報検索ないし索引作業においてどれほど細心に言葉を選んでみても、情報生成者から情報受容者への伝達には本質的に不確定性がつきまわっており、完全な情報検索はありえないということであろう。

さらに話を広げてみよう。人工知能ないし哲学において問題となっているものにフレーム問題がある。Nilsson⁴⁾によれば、「フレーム問題とは、行動により変化しない背景を変化する前景から区別することである」という。積み木の世界のような小さな閉ざされた世界でなら、状態を表現する命題を明示的に書き並べてコンピュータに推論させることも可能であろう。しかし、現実世界における行動を問題にすると、これは非常に難しい。例えば、私がコーヒーを飲むにしても、これによってコーヒーの量やカップの移動、私の胃袋の調子や頭の働き、などの変化が引き起こされる。一方、部屋の壁の色や私の名前、などといった諸々のことは変化しない。このような数多くの事実の中から重要な部分を取り出し、重要でない部分を切り捨てると言うことは、コンピュータにとって非常に難しいことなのである。

大澤⁵⁾はフレーム問題を「ある行為に相関して、その行為にレリヴァントな(関係ある)ものとイレリヴァントなものとを、どうやって効果的に弁別することができるかという問題」であるという。人間は、その限られた認識能力によって限定

された入力情報の中から、経験の教えに従って、結構、効率的にレリヴァントな情報を選択することができるが、それが正しい選択であるという保証はどこにもない。またコンピュータの場合、膨大な入力情報を処理することで選択を行っている。しかし、行為の決定に関わる適合情報のみを限定して収集することはできないし、たとえ適合情報をその中に含むような膨大な情報を入手したと仮定しても、それらの情報を完全に(効果的な時間内に)処理できる能力はない。このように人間やコンピュータの活動には、情報の部分性(知識の部分性と情報処理能力の部分性)と情報の部分性にもとづく不確定性が必然的につきまわるのである。

C. 論理的アプローチ

図書館情報学において、このような部分性は情報要求論にからんで問題となってきた。レファレンス・ライブラリアンが情報要求を特定しようとしても、そこには様々な要素が関与しており、情報の部分性に起因する不確定性がどうしても残る。これについて論じているのが Dervin であり、あくまでもレリヴァントな状況を探そうとしているのが Saracevic であろうか。いわゆる認知的アプローチと呼ばれてきたもののほとんどはこの問題をめぐる議論であったと考えていいだろう。

一方、システムそのものに何らかの部分性を扱わせようという方向の研究もあった。ファジー情報検索システムや確率的情報検索システムなどがその例である。ファジー論理は「あいまいさ」を扱うとされており、応用分野においてはかなり成功しているようであるが、本稿ではあえてファジー論理以外の論理に注目した。というのも、ファジー論理の概念的基盤がそれほど確固としたものではなく、強い応用指向のみが根底にあるように思えるからである。例えば、ファジー論理は、論理体系としての基本的要請である統語論や意味論、あるいは論理体系の無矛盾性や完全性などを備えていない。どのような「あいまいさ」であっても対応できる融通無碍さがファジー論理の長所なのであるが、それは同時に、一步踏み込んで

考えようとするときの短所（制約）にもなっている。

我々は情報検索の背後には情報の部分性があるということを出発点とする。そうするとその部分性のありようとそれへの対処の仕方の解明が情報検索システムの問題として浮かび上がってくる。そして部分性という困難な問題を考える枠組みとして、あえて基礎の確立された数理論理という固いものを採用しようと考えた。というのも、よほど確固とした試金石がなければ、「あいまいさ」や「不確定性」といった言葉の力に流されてしまう可能性が大きいからである。

そこで本稿では、そのような論理体系の代表として直観主義論理を取り上げ、この論理に基づく情報のとらえ方、情報構造の定式化の仕方を展開し、その1つの応用として直観主義に基づく情報検索システムを紹介する。

III. 情報の論理構造

真偽が部分的であるような命題の論理を展開するのに最適な論理が直観主義論理である。本章では、部分的情報を扱う枠組みとしての直観主義論理とそれに基づく情報構造を展開する。

A. 直観主義論理⁶⁾

オランダの数学者 Brouwer は 1908 年に提出した博士論文においてはじめて数学における直観主義という主張をおこなった。彼は人間の認識の不完全性に基ついた論理を構築しようとして直観主義論理を構想したのである。直観主義論理が古典論理と最も明確に異なっている点は、真偽の 2 値性に対する扱いである。古典論理においては、すべての命題は真か偽かのどちらかであるが、Brouwer はこれが神の立場にたった論理であって人間の論理ではないと否定する。彼は論理も人間の認識の不完全性に立脚したものでなければならぬと考えた。つまり、現時点において（あるいは神ならぬ人間の本質的な制約のもとでは）真偽の判定のつかない命題が存在することは当然であって、すべての命題が真か偽かのいずれかであるという古典論理の前提には無理があると考えた

のである。例えば、円周率 $\pi=3.14159\dots$ の小数展開の n 桁目は偶数である、という命題は n をいくらかでも（無限に）大きくすることができるという意味で本質的に（人間にとっては）真偽を明確に示さない命題である。

直観主義論理はある命題が成り立つことを確認する方法があるときに限ってその命題が真であると認める。したがって古来よく用いられてきた背理法という証明方法は直観主義論理では認められない。というのも、背理法はある命題の否定が成り立たないことを述べているだけであって、その命題を明示的に確認しているわけではないからである。命題には真とも偽ともいえないものがあることを認めた（すなわち、真理値の 2 値性を否定した）以上、ある命題の否定が成り立たないからといって、その命題が真であるとは主張できないことになる。

このように、人間の部分的な認識能力に立脚し、部分的な情報のもとにおける推論を形式的に体系化したものが直観主義論理である。情報システムにおける最も根底的な問題のうちの 1 つは、そこで扱われる情報の部分性の取扱いであるが、直観主義論理はまさにこの問題に最適な論理である。

B. 直観主義論理の代数的定式化

ここでは、前節で述べた直観主義論理が擬ブール代数 (pseudo-Boolean algebra) という構造をもつことを示そう。情報検索の論理が直観主義論理とみなせる、というだけでは単なるお話であって、これをシステムとして定式化しようとするなら何らかの形式化が必要だからである。構造を定式化する 1 つの方法が代数的構造としての定式化である。

まず補元の概念を拡張することから始めよう。通常の集合論において、集合 s の補元 $\neg s$ とは s と交わらない最大の集合である。言い替えると、 $s \cap x = 0$ (空集合) となる最大の x である。この定義を束 (lattice)⁷⁾ における議論でも用いることにしよう。このとき 0 は束のゼロ元であり、最大の元 x は擬補元 (pseudo-complement) と呼ばれ

る。

ここで用いたゼロ元 0 を束の任意の元 t に置き換えることによって擬補元の概念をさらに拡張しよう。すなわち、 $s \cap x \leq t$ を満たす最大の元 x を t に関する s の擬補元 (pseudo-complement of s relative to t) といい、これを $s \rightarrow t$ で表す。“ \rightarrow ” は通常の「ならば」であると考えてもいい。それは次の2点が言えるからである。

- (1) (p) かつ $(p \text{ ならば } q)$ から q が導かれる。
- (2) $(p \text{ かつ } r)$ から q が導かれるとき、 r から $(p \rightarrow q)$ が導かれる ($p \rightarrow q$ は最大)。

例えば、 p = 「東大阪市民である」、 q = 「日本国民である」としよう。 r として、「東大阪市民であるなら大阪府民である」とか「東大阪市民であるなら近畿の人間である」とかささまざまな推論が考えられるけれど、それらのうち最大なものは「東大阪市民であるなら日本国民である」であろう。これを相対擬補元というのである。この名称は特別の場合として $q=0$ であるときを考えれば納得がいく。このとき、 $p \rightarrow 0$ は $p \cap x = 0$ となる x であって、これは通常の意味での p の補元である。つまり、相対擬補元は通常の意味での補元という概念を拡張したものになっているのである。

束の任意の元 a, b に対して、相対擬補元が存在するとき、この束を相対擬補束 (relatively pseudo-complemented lattice) という。これは3つの2項演算子をもつ代数 $\langle S, \cup, \cap, \rightarrow \rangle$ とも考えられる。

相対擬補束は最大元 1 をもち (これは $a \rightarrow a$ である)、分配束であることがいえる。相対擬補束が最小元 0 を持ち、任意の元 a に対して、 $-a = a \rightarrow 0$ によって擬補元 (pseudo-complement) $-a$ が定義されているとき、この相対擬補束を擬ブール代数 (ないし Heyting 代数) という。擬ブール代数は3つの2項演算子と1つの1項演算子をもつ代数 $\langle S, \cup, \cap, \rightarrow, - \rangle$ とも考えられる。

この定義によると、 $-a \leq a$ は必ずしも成り立たない。なぜなら、 $-a$ とは $a \cap x \leq 0$ と満たす最大の x であった。すなわち、 $a \cap -a \leq 0$ である。 \cap は対称であるから、 $-a \leq a \rightarrow 0$ であると同時に $a \leq -a \rightarrow 0$ でもある。ところが、 $-a \rightarrow 0 = -a$ で

あるから、 $a \leq -a$ が成り立ってしまうのである。したがって、命題論理に対応する代数構造がブール代数であったように、排中律を認めない直観主義論理に対応する代数構造が擬ブール代数となる。

C. Kripke 意味論

1965年、Kripke は様相論理の意味論の系として、擬ブール代数に基づく直観主義論理の意味論的解釈を論じた。ここでは公理系にはふれずに、非形式的な形でこの意味論をスケッチしよう。

Kripke 意味論の中心的考えは、命題の真偽を知識の (現在の) 状態に対して相対化することである。よって命題は単に真なのではなく、現在の知識状態に関して真となる。この知識状態は可能世界と呼ばれている。Kripke 意味論を展開する前に様相論理における可能世界意味論を簡単に復習しておこう。

可能世界意味論の基本にある考えは、我々が世界についての完全な知識をもっていないなら、いくつかの可能性を想定するだろう、ということである。これは Shannon 流の情報理論とも相通じるところがあり、奇異な考え方ではない。つまり、知識があればそれだけ候補の数を減らすことができるということである。

ある可能世界 s における情報をもとにして、主体 i が可能世界 s' も可能であると考えたとき、可能世界 s と可能世界 s' との間には到達可能性関係 (accessibility relationship) があるという。

様相論理を公理に基づいて体系化しようとするとき、到達可能性関係のとらえ方によって様々な様相論理体系ができる。例えば、 $\Box P \rightarrow P$ という関係を公理として認めたとしよう。この関係は \Box を「必然性」と考えれば当然成り立つものであろう。しかし、 \Box を「信じている」と考えれば必ずしも成り立たない。信じているものが事実であるとは限らないからである。また時間的に解釈して「いつかは必ず真になる」と解釈しても、この関係は成り立たない。私はいつか必ず死ぬであろうが、今はまだ死んではないからである。これを到達可能性で言い替えれば、私が死んでいる世界

は可能であるが、その世界から現在のこの世界へは到達不可能である。これは上の関係が成り立たない例である。つまり、関係 $\Box P \rightarrow P$ は可能世界の間の到達可能性が反射的 (reflexive: 可能世界 A から可能世界 B へ到達可能であるならば、可能世界 B から可能世界 A へも到達可能である) であることを表していると考えられるのである。言い替えるなら、様相というものを可能世界の集合の間に成り立つ関係として幾何的にとらえることができるのである。そしてどのような到達可能性を公理として採用するかは、どのような様相 (時間なのか信念なのか等) を問題にしているのかに依存する。ちなみに上の関係式 $\Box P \rightarrow P$ は公理 T といわれ、この公理を満たす様相命題論理体系は様相論理 T と名付けられている。

直観主義論理においては可能世界の間の到達可能関係を単純な半順序であると想定している。ある可能世界において真となる文は後の状態においても常に真であると考えるのは自然であろう。従って、ある可能世界においてある命題が真であるならば、その可能世界から到達可能なすべての可能世界においてもその命題の真理は継続される。これを順序概念で表現すれば半順序 (partial order)⁹⁾ ということになるのである。これは時間についての順序であると解釈していいだろう。すなわち、我々が情報や知識を獲得していくさまを半順序で表現しているのである (知識を獲得する方法の可能性は多様であるので、全順序 (total order) とはならず一般に半順序となる)。すると可能世界の集合は過去の状態を含むばかりでなく、今現在の知識状態から我々が到達しうるすべての未来の状態も含むことになる。このような概念の集合体は Kripke 枠 (Kripke frame) $K = \langle S, \leq \rangle$ と呼ばれている。

命題論理において命題にそれを真とする要素の集合を対応させたが、この考えは直観主義論理においても利用することができる。命題をその命題が真となる可能世界の部分集合に対応させるのである。真理の継続性はこのような部分集合が K のフィルター⁹⁾ となることと同等、あるいは時間的順序のもとで継承的に閉じていることと同等で

ある。K 上のフィルターの集合を F とかこう。Kripke 付値とは命題の集合からフィルター集合への関数 V, すなわち、各命題 p にフィルター $V(p)$ を対応させるものである。このフィルターは命題 p が真となる可能世界の集合である。

可能世界のモデルを用いれば、直観主義論理が排中律を満たさないことは容易に示すことができる。2つの可能世界 $s \leq s'$ を考えよう。命題 p が s' で成り立っているとき (これを $s' \models p$ とかく), $s \models (p \vee \neg p)$ とはならないのである。なぜなら、もし $s \models \neg p$ なら $s \leq s'$ だから s' でも $s' \models \neg p$ となるはずであるが、 $s' \models p$ なのだからこれは不可能である。また s ではまだ十分な情報がなく命題 p が成り立つとはいえない。したがって $s \models (p \vee \neg p)$ はいえないことになる。

D. 部分情報と曖昧性

直観主義論理ではある可能世界において命題 p が成り立たないということを「その可能世界においては情報が足りなくて命題 p が成り立つのかどうか分からない」と解釈する。これは決して「 $\neg p$ である」ということではない。直観主義論理において否定は否定的命題が確認されたという具合に強く定義されている。すなわち $S \models \neg p$ とは S から到達可能なすべての可能世界において $\neg p$ となるときである、と解釈されている。一般に論理の意味論においては命題の意味としてその真偽値を考えるが、可能世界と命題との組に対して真偽値を割り当てる解釈関数を i とすると、 $i(p, s) = 1$ とは可能世界 s において p であることを知っているという意味であり、 $i(p, s) = 0$ とは $\neg p$ であることを知っているという意味である。すると成り立つかどうか分からない命題も当然あるので、解釈関数は全域関数 (total function: 領域のすべての変数に対して値が定義されている) とはならず、必然的に部分関数 (partial function) となる。すなわち*を未定義と考えれば、

$i(p, s) = 1$ 可能世界 s において p であることを知っている

$i(p, s) = 0$ 可能世界 s において $\neg p$ であることを知っている

$i(p, s) = *$ 可能世界 s において p が成り立つかどうか知らない

と表すことができる。

この意味論は曖昧さを表現するのに便利である。例えば、「背が高い」という述語を明確に満たす人も明確に満たさない人もいるだろうが、いわゆる境界線上にいる人も多いであろう。明確に満たす人の集合を肯定的外延 (positive extension), 明確に満たさない人の集合を否定的外延 (negative extension) と呼ぶとき、これらの間のギャップが境界事例である。

一般にこのような外延の決定は判断基準に依存する。そして判断基準という情報の量に応じて可能世界間の関係を定式化することができる。つまりこれらの可能世界には判定基準の正確さによる順序がつけられていると考えるのである。

Kripke 枠に対応するものとして曖昧枠 $F = \langle S, \leq, D \rangle$ を定義しよう¹⁰⁾。ここで、 S は可能世界の集合、 \leq は半順序、 D は個体の領域である。ただしこれらの可能世界の中には最大元 T が存在するものとする。それは全く曖昧さのない情報の完全な世界のことである。

可能世界 s においてはその世界の判断基準に従って肯定的外延と否定的外延が決定される。可能世界 T においてはこれらの決定が完全になされ、すべての元が肯定的ないし否定的に分類される。 T 以外の可能世界においては、どちらとも判断しかねる元が存在する。ある可能世界における情報を拡大してより正確な判定基準をもった可能世界に到達する方法は幾つも考えられる。

i^+ と i^- は命題 p と可能世界 S の組にその肯定的外延および否定的外延に対応させる解釈関数とする。すると $s \leq s'$ であるとき、 $i^+(p, s) \subseteq i^+(p, s')$ 、 $i^-(p, s) \subseteq i^-(p, s')$ が成り立つ。すなわち、これらの解釈関数は単調であって、判断基準が正確になるほど肯定的ないし否定的外延と決定されるものの数は増えていくのである。この極限に世界 T がある。

先ほど命題論理に部分集合の集合を対応させて命題論理の体系を説明したのと同じように、命題に肯定的外延と否定的外延の組を対応させて論理

体系を定式化する事ができる。例えば否定は、 $\neg \langle X, Y \rangle = \langle Y, X \rangle$, and には $\langle X, Y \rangle \vee \langle Z, V \rangle = \langle X \cap Z, Y \cup V \rangle$ などと定めるのである。このようにして作られた構造を de Morgan 束といい、命題論理のブール代数に対応するものである。これが真理値のギャップを表現していることに注目してほしい。

可能世界 s における真偽はその世界のみに限られた局所的なものであって、命題の真偽といえるものではない。新しい真偽の概念が必要である。可能世界 s の拡張となるすべての可能世界 s' で $s' \models p$ となると p は超真 (supertrue)、可能世界 s の拡張となるすべての可能世界 s' で $s' \models \neg p$ となると p は超偽 (superfalse) であるといわれている。

E. 直観主義論理と位相構造¹¹⁾

我々は単独の情報ではなく情報の集合 (情報システム) を考えている。情報システムは論理をもち、論理は様々な形に構造化される。これまでには主に代数構造として定式化される様子を見てきたが、情報構造の一般論を考えるとき、さらに位相構造も取り入れると幾何的なイメージが大きく広がる。ここでは直観主義論理を位相的に定式化してみよう。

ブール代数の最も一般的な例は通常の部分集合の族 (部分集合の集合) であるが、擬ブール代数の最も一般的な例は位相空間における開集合の族 (開集合の集合) なのである。すなわち、命題論理において各命題にその命題を満たす要素の集合を対応させたように、直観主義論理においては命題 ρ に位相空間の開部分集合 $\langle \rho \rangle$ を対応させるのである。前節の曖昧さの説明の中で境界線上にある事例ということを述べたが、位相空間において境界線を含まない集合が開集合、境界線を含む集合が閉集合である。従って、命題に開集合を対応させることは、境界線上にある曖昧な事例を切り捨て、明確に立証される命題だけを認めていこうという直観主義論理の考えに沿ったものであることがわかる。

ちなみに位相空間とは距離空間を抽象したもの

であって、近さを開集合に含まれるか否かで表現しようとしたものである。また位相空間の基底 (basis) とは、位相空間の任意の点を十分詳細に表現できる小さな部分集合を含む部分集合の族 B である。点 x を含む開集合を点 x の近傍 (neighbourhood) というが、この近傍が基底の元である時、これは基本近傍と呼ばれる。正確に言えば、基本近傍系 $N(x)$ とは、 x の任意の近傍 V に対して、 $U \subseteq V$, $U \in N(x)$ となる開集合の族 $N(x)$ のことである。

位相空間における開集合の族を考えよう。開集合の和と積はまた開集合であるし、相対擬補元を、 $s \rightarrow t = \text{Int}((X-s) \cup t)$ (Int は開集合をとる演算子) とすればこれも開集合であるから、開集合の族は各演算子に関して構造的に閉じており (すなわち、演算子を作用させた結果もまた開集合になる)、ある種の代数構造であると考えられる。さらに相対擬補元が定義されているのだから相対擬補束である。この相対擬補束の稠密¹²⁾で開な部分集合の全体は、補集合をとるという演算に関して閉じているので、この演算を補元演算子 c とし、これは擬ブール代数となる。

二重否定したものがもとの命題に等しくならない ($A^{\circ\circ}$ が A とならない) 例としても位相空間における開集合が最適である。一般に開集合の補集合は閉集合となるが、ここでは開集合のみを考えているのだから、補元をとる演算子として、補集合をとりさらに開集合をとる演算子を考える。すなわち、補集合をとる演算子を c 、開集合をとる演算子を int としたとき、補元演算子を $\text{compl} = \text{int} \circ c$ で定義するのである。実数全体を全体集合とし、いま $A = (-\infty, 0) \cup (0, \infty)$ としよう。これは実数空間において稠密であって、 A の補集合の開集合は空集合である。しかし空集合の補集合の開集合は実数空間であって、もとの集合 A とは一致しない。

命題をある位相空間の開集合をみなそうというのだが、この位相空間の基本近傍系の元 U があって $U \subseteq \langle \rho \rangle$ であるとき「 U は ρ を導く (証明する)」ということにしよう (順序としての論理を思い出していただきたい)。ここで「証明」という

言葉をだしたが、直観主義論理においては明確に証明されたもののみが命題として認められるからである。証明関係を開集合間の包含関係に対応させたのが直観主義論理の位相的表現なのである。

このように定義すると命題間の通常の結合子が次のように定義できる。 $U_i (i=1, 2)$ が ρ_i を導くとき、 $U \subseteq U_1 \cap U_2$ となる U があって、 U が ρ_1 と ρ_2 を導く。この条件を満たす U の中で最大のものは $\langle \rho_1 \rangle^c \cup \langle \rho_2 \rangle$ であって、これを $\langle \rho_1 \rightarrow \rho_2 \rangle$ とかく。結合子 \cap についても同様である。

つぎに含意を定義しよう。 $\langle \rho_1 \rangle^c \cup \langle \rho_2 \rangle$ における基本近傍系 U をとると、 U と交わっている任意の開集合 U_1 に対して ρ_2 を導く開集合 U_2 を細かく制限することができる。この性質を持つ U は $\langle \rho_1 \rangle^c \cup \langle \rho_2 \rangle$ という制約をみたく最大の開集合である。したがって、 $\text{Int}(\langle \rho_1 \rangle^c \cup \langle \rho_2 \rangle)$ によって $\langle \rho_1 \rightarrow \rho_2 \rangle$ を定義することができる。

また ρ が位相空間 X において真であるとは、 $\langle \text{Cl}(\rho) \rangle = X$ が成り立つときであると定義しよう。

位相空間を用いた直観主義論理の解釈は、古典論理の解釈もその特例として含むという意味で一般的なものである。位相空間の特別の場合として1つの点のみからなる空間 $X = \{a\}$ を考えよう。これは開集合として全空間 X と空集合の2つだけをもつ。すると従来通り、真の命題には全空間 X が対応し、偽の命題には空集合が対応して、真偽の2値のみからなる古典論理が表現されたことになるのである。

IV. 情報検索の新しい定式化

直観主義論理に基づけば部分情報が表現できたのだから、曖昧な (部分的にしか決定されていない) 情報をも扱う情報検索を定式化するのに、基本論理として直観主義論理を採用することも可能であろう。本章では Jaegermann や Lipski などによって展開された直観主義論理にもとづく情報検索理論を紹介する¹³⁾。

文献の主題に関する情報が部分的であって、文献を索引するための情報が不足しているときにも使える情報検索システムを考えてみよう。例えば、ある属性の値 (例えば、あるブロックの色と

いう属性の値)が一義に(青であるとは)定まらず、いくつかの値の集まり(青か緑)になってしまうような場合である。つまり実際の値が明確に(一義に)は知られていない状態である。ただ真の値がその集合の中にあるということは仮定しておく。

おもちゃのブロックの検索を考えよう。属性として色に注目する。色は{青, 緑, 赤}の3色であって、各ブロックにこれらのうちの1つが割り当てられているとする。いつでもその色が認識できるとは限らない(十分な照明がないとか)。これが部分的な情報の状態である。つまり、そのブロックが青色をしているのか緑色をしているのか明確にはいいがたい場合である。

情報が部分的であって「青か緑かどちらかである」という不確定な記述をも含めたシステムをつくらうとすれば、どのような定式化を行わなければならないであろうか。色が青であるとわかっているブロックの集合を $d(\text{青})$ 、色が青か緑であるとわかっている(どちらであるかはわからない)ブロックの集合を $d(\text{青}, \text{緑})$ と表すことにする。ブロックの集合の包含関係を見ると、 $d(\text{青})$ と $d(\text{緑})$ は $d(\text{青}, \text{緑})$ に含まれており、 $d(\text{青}, \text{緑})$ は $d(\text{青}) \cup d(\text{緑})$ よりも大きい。なぜなら、 $d(\text{青}, \text{緑})$ には青とも緑とも確定していないブロックが含まれているからである。

論理は推論の体系であるから、ある種の制約を満たす順序として表現することができる。したがって、情報検索システムの論理も属性値の間の順序関係によって表現することができる。ここで属性値間の順序関係をその属性値を満たす(すなわち、その属性値を持つと知られている)対象の集合 $d(*)$ の間の順序関係によって表すことにする。たとえば、「青」という属性値も「緑」という属性値も「青か緑」という属性値に含まれる、という順序関係が成り立っており、このような順序関係によってシステムの論理は決定されるのである。

このシステムにおいて、 $d(\text{青}) + d(\text{not 青})$ はブロック全体とはならない。青であるブロックをすべて明確に認識しているわけではないからであ

る。これは検索の論理が排中律を満たさないということであって、これからも情報検索の論理として直観主義論理の方が自然であることがうかがえる。

情報検索システム(あるいは、情報システム一般)は人間が情報を加工する以上、現実の不完全なモデルであって、システムは現実の断片を提供するにすぎない。システムの利用者はシステムから現実の完全な情報を獲得できると期待しているのだろうか、あるいはシステムで知り得る限りと割り切っているのだろうか。

「知られている」ということ、つまり人間の認識をもとにすれば情報検索システムの機能は、「**について述べている」と知られている」文献だけを提供するものから、「**について述べているかも知れない」文献をも提供するものまで、さまざまな情報検索を考えることができるだろう。

不完全なシステムに新たな情報がつけ加わり(たとえば、明るさが増して色が明確に見えるようになって)、それまで不確定だったものが確定されると、新しい情報システムが構成されることになる。このような新しい情報の追加によって、属性値の集合という不確定な記述が唯一の属性値という確定記述に変更されていった究極にあるのが完全情報システムである。これは情報検索システムがその時点における知識の状態に対応している、ということの意味するものであり、情報検索の意味論がKripke 枠によって解釈可能であることを示唆している。これは可能世界の順序に関して極大な可能世界 T に対応するものであって、完全システムにおいては排中律も成立し、システムの論理は命題論理となる。

情報検索システムはディスクリプター集合のつくる代数構造から文献集合の部分集合への写像として記述することができる。例えば、ブールの情報検索においてはディスクリプターはブール演算子で結合されて検索式となり、検索式の全体はブール代数となった。またそれぞれのディスクリプターに対応する文献集合の間でブール演算子に対応する集合演算を行い、結果となる文献集合を生成した。これは全文献集合の部分集合である。

すなわち、ブールの情報検索においては、ディスクリプター集合の作るブール代数から文献集合の部分集合への対応は集合演算で規定されていた。

ブール情報検索の場合、ディスクリプターから構成された検索式はブール代数となったが、部分情報を扱うという条件のもとでは、ディスクリプターから構成される代数は擬ブール代数となることが示せるのである。このとき、そのような情報検索システムの内部論理は直観主義論理となる。

V. おわりに

状況意味論をはじめとして部分的情報をめぐる議論が盛んになってきた。直観主義論理や Kleene の 3 値論理などはファジー論理が生まれる以前からあるのだが、最近になって論理的意味論やプログラミング意味論の発展と呼応するように、部分情報を扱うための枠となる論理体系の整備が行われはじめています。もちろんそれらの展開の基礎には直観主義論理がある。

本稿では部分情報をふまえた情報構造の一般論への足がかりとして直観主義論理を中心とした解説を行った。情報構造は代数的に定式化されるのが普通であるが、今後の展開の可能性を考え、あえて（情報検索システムの定式化にはまだ取り入れられていない）位相的構造の説明も行った。情報構造というテーマそのものは図書館情報学のテーマとしてそれほどズレたものとは思えないし、位相的などらえ方は幾何的なイメージの豊かさをもった魅力的なものだからである。

今回のテーマの中心は情報構造であって情報検索の定式化はその情報構造を用いた一つの例であった。さらにより具体的な場面における情報構造の定式化が今後の課題である。

注・引用文献

- 1) Swanson, D. R. "Undiscovered Public Knowledge". *The Library Quarterly*. vol. 56, no. 2, April, p. 103-118 (1986)
- 2) Ingwersen, P. 情報検索研究: 認知的アプローチ. トップラン. 1995.
- 3) クワイン, W.V.O. ことばと対象. 勁草書房. 1984.
- 4) Genesereth, M. R. & Nilsson, N. J. 人工知能基礎

論. オーム社. 1993, p. 324.

- 5) 大澤真幸, 松原仁, 黒崎政男. "一般フレーム問題とは何か". *現代思想*. vol. 18, no. 7, p. 168-191 (1990)
- 6) 直観主義論理については
小野寛晰. 情報科学における論理. 日本評論社. 1994. などを参考にした。
なお, 円周率の例は次の文献による。
竹内外史. 数学的世界観. 紀伊國屋書店. 1982.
- 7) 束とは半順序集合であって, 任意の 2 元の上限と下限が存在するものをいう。上限とはこの 2 元より大きなものの中で最小なもの, 下限とはこの 2 元より小さなものの中で最大のものである。順序が入っている多くの構造 (例えば, 論理構造) は束という構造になっている。
- 8) 半順序とは通常の順序関係を抽象したものである。つまり, aRa, aRb かつ bRa なら $a=b$, aRb かつ bRc なら aRc という 3 つの条件を満たす関係を半順序関係という。任意の 2 元をとったとき, それらの間に必ずしも半順序関係があるとは限らない。任意の 2 元が比較可能であるというこの条件をも満たすものが全順序関係である。全順序とはすべての元が一直線上にあるのに対し, 半順序は枝分かれしており, 枝でつながっていないものは比較できないのだと考えればいだろう。
- 9) フィルターとは次の条件を満たす部分集合 F である。
(i) $1 \in F$
(ii) $a \in F, b \in F$ ならば, $a \cap b \in F$
(iii) $a \in F, b \in X$ ならば, $a \cup b \in F$
フィルター F は部分的真をも含む真理集合の構造を表していると考えられる。
- 10) Landman, F. *Structure for Semantics*. Kluwer Academic. 1991, ch. 2. を参考にした。
- 11) 位相空間については
竹之内脩. トポロジー. 廣川書店. 1972. 竹之内脩. 集合・位相. 筑摩書房. 1970. 等, 多くの図書が出版されている。また, 直観主義論理の位相空間による定式化は次の文献を参考にした。
van Dalen, Dirk "Intuitionistic Logic". Gabbay, D. et. al. eds. *Handbook of Philosophical Logic*. Vol. 3, p. 225-339 (1986)
- 12) 部分集合が稠密 (dense) であるとは, ほぼ全体集合に等しいくらい大きいということである。補集合をとれば空集合になる集合として定義してもよい。例えば, 有理数の集合は実数の集合において稠密である。有理数の数と実数の数を比べると, 圧倒的に実数の数の方が多いのだけれど, 実数全体の集合における有理数全体の補集合の開集合をとれば, 空集合になる。言い替えると, 勝手に実数を選んだとき, それにいくらでも近い有理数を選べるくらい, 有理数は実数の集合の中にピッタリと密に (dense) 詰め込まれているので

ある。

- 13) Jaegermann, M. "Information storage and retrieval systems with incomplete information I". *Fundamenta Informaticae*. vol. 2, p. 17-41 (1978)

Lipski, W. "On semantic issues connected with incomplete information databases". *ACM Trans. Database Syst.* vol. 4, no. 3, p. 262-296 (1979)