

抽象記号系のモデル：思考の分散表現と自己の中心性

A model of abstract symbolic systems: distributed representation of thought and the centrality of Self

犬童 健良[†]

Kenryo Indo

[†] 関東学園大学

Kanto Gakuen University

kindo@kanto-gakuen.ac.jp

概要

本研究は、表意作用を記号間のマニピュラビリティとして形式化する。表意作用はある対象の集まりの順序の対（プロフィール）の集合をドメイン（定義域）から対象の離散的な勾配、すなわち各順序に対応するエージェント最適反応が恒等写像でない場合として発生する。社会選択理論における Gibbard-Satterthwaite 定理から独裁的な制限ドメインでは非独裁かつ全射的であれば必ずマニピュラブルであるプロフィール対が存在する。本研究では Prolog を用いて意味の流れの場を可視化する実験手法を紹介する。また自己意識、言語的パズル、メタファー、意思決定アノマリーなどへの応用についても議論する。

キーワード：意味構成関数, マニピュラビリティ, 耐戦略性, 自己の分散性と中心性

生として理解しようとした(トム, 1980)。すなわちある勾配ベクトル場（より一般的にファイバー束）におけるトポロジー的性質の変化としての意味作用が可視化される。本研究は、記号の表意作用をマニピュラビリティの概念と関連付けてそのダイナミックな性質を形式化するモデルを提案する。

本研究では制限ドメインの耐戦略的 SCF と独裁および可能ドメインを実験的に生成する犬童(2025)の Prolog プログラムを利用した（ソースコードは Github で公開されている）。例えばボルダールールのプログラムを追加して実験することができる。2 人の投票者で 3 候補者 a,b,c についての選択結果を図 1 に示す。

1. はじめに

表意作用は記号 S、記号が意味する対象 O、および両者を結びつける解釈 I の 3 項関係 $\langle S, O, I \rangle$ として表される (Pierce, 1986; Ogdem & Richards, 1923; 米盛, 1981)。有標 - 無標による二項対立（シニフィアン—シニフィエの関係）による構造化は記号論の基礎であるが、アイコン、インデックス、シンボル、より一般的に 1 次性・2 次性・3 次性に基づくパースの 3 分法との間に微妙な不一致がある (Jakobson, 1984)。また渡辺(1986)は（非フレーゲ的な）パースの含意概念を確率論理の一種（傾向論理）によって表現しようと試みた。しかしパースの言う「生きた論理的帰結」（living logical conclusion）は科学的・実用主義的な推論、その習慣として捉えられるべきものである（パース著作集の 5.491; 米盛(1981)p.215）。

記号論の対象は言語的な素材にとどまらず、任意の対象が何か別のことを意味する現象が前言語的に構造化され、それゆえ構文論・意味論・語用論といった言語学的な区別も予め存在していないことに注意する。パースが目指したのはそのようなダイナミックな意味の認知現象、とくにその解釈項、すなわちマイクロレベルでの自己意識のはたらきである。ルネ・トムは微分位相幾何学を用いて記号過程を力学軌道の構造安定な形態発

図 1 ボルダールールを用いた意味選択関数の例

	1	2	3	4	5	6
1: [a, b, c]	a	a	a	b	a	a
2: [a, c, b]	a	a	a	a	a	c
3: [b, a, c]	a	a	b	b	a	b
4: [b, c, a]	b	a	b	b	c	b
5: [c, a, b]	a	a	a	c	c	c
6: [c, b, a]	a	c	b	b	c	c

行と列の各見出しの数字とカッコ内は 3 候補の好ましさの順序を区別するものであり、行と列が 2 人の投票者の観点に対応する。投票者は記号解釈の観点に読み替えられる。候補者の順序の組（すなわちプロフィール）あるいはその集まりが記号となる可能性があるものである。これらは解釈される記号であると共に対象でもある。数学的にはこれらを適当な被覆の下で脈体（nerve）として構成し直すと抽象的単体ないし単体複体といった代数的な位相空間が得られる。

記号の表意作用（あるいは意味作用）はある記号がそれ自体でなく他の何かを意味する現象である。本研究で提案する抽象的記号系において、表意作用（あるいは

意味作用) はある記号 (の意味) が別の記号 (の意味) にいいかえられること, すなわちマニピュレーションである. 本研究ではマニピュレーション概念の定式化を社会選択理論から借用する. プロフィールの集合は意味構成関数 (SCF) の定義域 (ドメイン) である.

図 1 の SCF のマニピュラビリティを図 2 に示す. * はマニピュラブルなプロフィールであり, その垂直方向あるいは水平方向の別のプロフィールに向かってマニピュラビリティ, つまり意味作用の流れが存在する. 0 で示されるプロフィールはマニピュラブルでない, つまり耐戦略的である. なおこの場合のボルダールールには選択結果のあいまい性があるが, @の部分は何のケースでもつねにマニピュラブルである (この 6 プロフィールは制限ドメインで独裁性に関するッリングの一部である).

図 2 マニピュラビリティと耐戦略性

	1	2	3	4	5	6
1: [a, b, c]	0	0	@	*	0	*
2: [a, c, b]	0	0	0	*	@	*
3: [b, a, c]	@	0	0	0	0	0
4: [b, c, a]	*	*	0	0	*	@
5: [c, a, b]	0	@	0	*	0	0
6: [c, b, a]	*	*	0	@	0	0

プロフィール間のマニピュラビリティはパス流の解釈項 (記号と対象の関係) やソシユール流の二項対立として再解釈される. マニピュラビリティを考慮することのメリットの一つは, 記述力の高い数理的な認知モデリングにおける制約の恣意性 (この弱点は可能世界意味論, メンタルスペース理論, 状況理論, 圏論モデルなどに共通すると思われる) から少なくとも部分的に免れることである.

ポスターではロバ文, 反事実条件文やメタファー, ゲチエ問題, ニューカム問題といった言語・論理・意思決定のアノマリーへの応用についても議論したい.

2. 分散する自己の安定性と中心性

記号過程は意味の場としての順序プロフィールの集合 (意味関数のドメイン) における相互作用であり, 表意作用はプロフィール間のマニピュラビリティとして発生する. 記号系は, ある記号 (の意味) が別の記号 (の

意味) にいいかえられること, すなわちマニピュレーションが起こる場である. また記号過程は意味の場で起こるマニピュレーションの時系列である. この意味作用の場は記憶 (記憶装置) を抽象化したもの, あるいはエージェントが相互作用する環境であるとも解釈でき, そこでは記号を使用することは記号を使用する環境そのものに影響する行為である (なお本研究では立ち入らないが, 行為は語られることで自己を形成し, 意味付けられていくという観点は, アーレント, バルト, クリステヴァらの思想に通じると思われる).

記号過程はマニピュレーションの時系列であり, 思考はマニピュラビリティのネットワーク内に分散表現されつつ, 自己の中心性が保持される (A.O. Rorty の自己欺瞞とアクラシア). この条件は社会選択理論における Gibbard-Satterthwaite 定理 (以下 GS 定理と記す) の制限ドメインへの一般化である.

自己意識が自らを欺かないこと, すなわち思考における自己の安定性はマニピュラビリティの排除 (耐戦略性と呼ばれる) と意味的な多様性 (非独裁性かつ非強制性) をバランスさせる必要がある. 後述するように, いわばセザンヌが絵画に多視点を導入したときのような巧妙な手術が必要である. それらは認知的な阻止 (cognitive blocking) の形式化であり, 位相空間における障害 (obstruction) と解釈できるかもしれない.

また自己の中心性はマニピュラニリティのネットワークが友情グラフ (friendship theorem) の条件を満たす場合に成立する. 記号の 3 項関係は共通友人の条件そのものであるから, 4 サイクルを認知的に阻止することが自己の中心性の十分条件である (Erdős, Rényi, & Sós, 1966).

メンタルモデルに転用されたとき, 耐戦略性は思考することで自分自身が欺かれないという自然な制約である. またそれによって記述力の高いメンタルスペースや圏論などにおけるモデリング上の恣意性が少なくとも部分的に解消される. たとえば反事実的条件の可能世界モデル (Lewis, 1973) は最近接不可能世界の選択あるいは可能世界間の類似性にかんするモデルのあいまい性や恣意性にかんする技術的な問題が知られる (Lakoff (1996) も参照). マニピュラビリティによるアプローチでは極大可能ドメイン, すなわち最小数のプロフィールを認知的阻止することにより, 意思決定の様式 (語り口) が生成するオラクルの系列とその公平指標として明確化される (犬童, 2025).

3. 基本モデル

抽象記号系の具体的なモデルとして、例えば 3 個の (基礎的な) 対象 a, b, c とその順序を考え、またそこから基本的な記号系のモデルを構成する。より一般に観測 (認知) によって確率分布 (対象ごとの確率の予測) が得られると仮定すると、順序はその定性確率に相当する。この順序ないし確率の可能なペアをプロフィールと呼び、プロフィールの集まりを意味が発生する場であり、記号はプロフィールの関数 (それを SCF と呼ぶ) として定義される。暫定的に SCF は意味構成関数、あるいは意味結合関数と読むことにする (名称がふさわしいかどうかは読者の判断に委ねる)。またその定義域を意味作用のドメイン (領域)、あるいはたんにドメインと呼ぶ。このモデルは数理モデルとしては経済学で社会選択関数と呼ばれているものと一致するが、本研究では順序プロフィールの次元数 $n=2$ とする。2 つの確率はそれぞれの予測を報告 (確率を推奨) するエージェントに対応する。

確率はシンプレックスに対応し、順序の変化は隣接する小領域間での基底を変換するオペレーションに相当する。この変換は多様体の接続 (メトリック) の離散的なバージョンである。シンプレックス対内の分割された領域間の境界を辺としてそれに沿って移動することを記号の解釈と考えると、意味作用はそのシーケンス、すなわちマニピュラビリティの発生する辺をたどったパスとして表現される。

このシーケンスはホモトピー法的一种であるラベル法 (指標理論) およびとその擬多様体への一般化として知られる (Shapley, 1974; Lemke & Grotzinger, 1976) に類似する。

4. 独裁定理

意味構成関数 SCF はプロフィールの関数である。特殊な SCF はプロフィールに、偽 (懷疑) か真 (信念)、あるいは 0 か 1 を割り当てる (判別プロセス)。より一般には SCF はプロフィールに、順序を割り当てる (再生のプロセス)、あるいは対象を割り当てる (注目プロセス)。個々のプロフィールを客体、いくつかのプロフィールを集めた集合を述語と考えると、パス = 渡辺風の記号過程として解釈される。社会選択理論の基本性質 (一般不可能性定理, GS 定理) の記号過程としての再解釈は、渡辺 (1981, 1986) の「みにくいアヒルの子定理 (ugly duckling theorem)」のバージョンである。

どの記号も一方のエージェントの観点だけで意味が定まる独裁になり、他方の報告はつねに無視される。これは特徴関数ゲームとして述語の真偽値をモデル化したときに非空コアの存在条件、すなわちバランス集合族が選べること (Bondareva-Shapley theorem) に相当する。

ある記号 (プロフィール) の SCF 値をその記号の「文字通りの意味」と呼ぶことにすると、意味作用は記号間のマニピュラビリティとしてモデル化される。記号が文字通りの意味ではなく、別のプロフィールの SCF 値となること、すなわち「いいかえる」現象としてのマニピュラビリティはゲーム理論的には最適反応に相当しこれを理論的に予測することができる。プロフィールが制限されたドメインにおいて GS 定理のバージョンが成り立つようなプロフィール集合を (γ) 独裁ドメインと呼ぶことにしよう。独裁ドメインにおいて多様性を確保するためにはマニピュレーションを防ぐことができない。いいかえれば意味作用を生じるプロフィールが必ず存在する。社会選択理論の文脈ではマニピュラビリティ概念は、真の選好を偽ること (嘘つき・虚偽申告) であり「悪」である。しかし意味作用のモデルでは「善」になる。暗黙知理論では局所的な記号過程のパーシステントな特性が全体的な空間において意図しない反転を引き起こすことがある。非自明なホロノミーが生じると言い換えられるかもしれない。このモラルインバージョンの現象は現実社会で観察されるが、実はそれが M. ポランニーの暗黙知理論の動機である (Polanyi, 1966)。

とくに極大可能ドメインでは認知的に阻止されたプロフィールペアのいずれか一方を追加することで独裁ドメインに変化するが、認知的阻止を系統的に交換することで独裁と公平の間を結ぶ変形系列が得られる。これは h-コボルディズム定理に類似する。つまり隣接指標の臨界点ペア間のキャンセルによって多様体をより単純なトポロジーに変形することができることに類似する。公平なルールがいかえられて独裁に変わるという観点から、みにくいアヒルの子定理になぞらえて、これを「いいカエルとわるいカエルの補題」と呼ぶことにする。

5. 認知的阻止と極大可能ドメイン

重要なことは自己意識が安定して作動するには、認知的に阻止するプロフィールの部分集合を上手に選ぶ必要があるということである。そのために気づかれな

い、あるいは意識できない部分が生じる。ちなみに定数関数のように結果の多様性を捨てるか、独裁のように観点の多様性を捨てれば自明な耐戦略的ルールを得る。また独裁ドメイン同士も互いに独立ならばそれぞれの多様性を下げ合併すれば可能ドメインに変わる。一方、現実の多様性を非自明な耐戦略性を得るためのミニマルな認知的阻止は基本ケースで 54 通り存在するが、その選び方は超アロー領域と関連しており、適当なプロフィールペアを抜けばよいというわけではない（犬童, 2025）。

繰り返しになるが、独裁ドメイン間を移動する系列（独裁シーケンス）においてはつねに意味作用の発生が保証される。マニピュラビリティの連鎖で生じる記号の言い換えが意味のある思考として成立するためにはこのことが必要である。一方セレンディピティのような現象では思考は一般に散漫であるが、薄々そこに近づく感覚に気づくことある。この性質はホモトピー法に似ている。最終的に完全な解に至るまではつねに何か一つだけ条件が欠けている状態（準完全ラベル）で記号（ラベル）の入れ替えをしていると不意に均衡に到達する。しかしループにはまって出られなくなることもある。

6. まとめ

マニピュラビリティに基づく記号系の数理とその計算的実験は従来の記号論の発見的推論や認知科学的モデルに、少なくとも潜在的に刺激的で新たな視点を投げかけている。記号の意味のインバージョン、意味の勾配ベクトル場におけるホロノミーは、いいカエルとわるいカエルの補題として寓意化される。記号論モデルにおける根本的なパズルであるパース＝ヤコブソンの問題（二重マークドネスの排除）は独裁結果の二者択一性から直接帰結する。これ以外に、反事実条件文、メタファー/アナロジー/アイロニー、潜在的応用として意味論的パズル（ロバ文、再帰代名詞、ゲチエ問題など）、決定基準のあいまいな意思決定問題（ニューカム問題）、あるいはくじ比較（Allais パラドックス）における仮想的結果間の相対的な注目がある。なお Chichilnisky, Heal, Barisinikov らが開拓した位相的社会選択理論にならって、本研究で示した制限ドメイン上での社会選択関数を用いた意味の勾配ベクトル場（より一般的にファイバー束）を脈体（nerve）やその離散モース関数として扱くと、多様体をベースにした軌道の数理のアナロジ

ーをより厳密に扱うことができるようになると思われる。今後の課題としたい。

文献

- Arrow, K. J. (1963). *Social Choice and Individual Values*. Yale University Press.
- 犬童 健良(2025). 論理プログラミングを用いた抽象的社会選択関数における耐戦略性の実験的研究, 関東学園大学経済学紀要, Vol.51, pp.1-64.
https://doi.org/10.20589/kantogakuenecconomics.51.0_1
- Erdős, P., Rényi, A., and Sós, V. T. (1966), *On a problem of graph theory*, Studia Sci. Math. Hungar., 1: 215–235.
- Gibbard, A. (1973). Manipulation of voting schemes: A general result, *Econometrica*, 41, 587-601.
- Jakobson, R. (1984). 記号論の発達略史, ヤコブソン, R. 池上嘉彦・山中 桂一(訳) 言語とメタ言語 勁草書房, 1984, pp.9-41.
- Lakoff, G. (1996). Sorry, I'm not myself today: The metaphor system for conceptualizing the self. In G. Fauconnier & E. Sweetser (Eds.), *Spaces, Words, and Grammar*. Chicago: University of Chicago Press, pp. 91–123.
- Lemke, C.E., Grotzinger, S.J. (1976). On generalizing shapley's index theory to labelled pseudomanifolds. *Mathematical Programming*, 10, 245–262.
- Lewis, D. (1973). *Counterfactuals*. Harvard University Press.
- Ogden, Ch. K., & Richards, I. A. (1923). *The Meaning of Meaning: A Study of the Influence of Language upon Thought and of the Science of Symbolism*. 10th ed., London: Routledge & Kegan Paul.
- Peirce, C. S. (1986). 記号学：パース著作集 内田種臣訳 勁草書房.
- Polanyi, M. (1966). *The Tacit Dimension*. Doubleday & Co. (ボラニー, M. 佐藤 敬三(訳) 暗黙知の次元：言語から非言語へ 紀伊国屋書店)
- Rorty, A. O. (1987). Self-deception, akrasia and irrationality. In Elster, J. (ed.), *The Multiple Self*, Cambridge University Press, pp.115-131.
- Satterthwaite, M. (1975). Strategyproofness and Arrow's conditions: existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory*, 10, 187-217.
- Shapley, L.S. (1974). A note on the Lemke-Howson algorithm. *Mathematical Programming Study*, 1, 175–189.
- トム, R. 彌永昌吉・宇敷重弘(訳)(1980). 構造安定性と形態形成 原書第2版 岩波書店.
- 渡辺 慧 村上陽一郎・丹治 信春(訳)(1981). 知識と推測：科学的認識論 東京図書.
- 渡辺 慧 (1986). 知ること 東京大学出版会.
- 米盛 裕二 (1981). パースの記号学 勁草書房.