

内部観測と、組織内コミュニケーション、組織による意思決定： 生体的組織デザインフレームワークの構築に向けて

Communication and Decision-makings in/by Organizations :
Towards Construction of Designing Framework for Organizations Inspired by Biology

西川 麻樹*¹

Asaki NISHIKAWA

*¹ 東京大学大学院情報学環

The University of Tokyo Interfaculty Initiative

組織の行為は、二重の内部観測的プロセスを含む。組織構成要素となるエージェント達による「内部観測 = 相互観測 = 組織内コミュニケーション」、組織全体を単一エージェントとした場合の「選択 = 環境の観測 = 意思決定」である。本研究では、生物や人間の持つ内部観測 = 認知機構を利用した組織モデルに、人工知能に与えるタスクを学習させることで、組織デザインに内部観測的プロセスを応用するフレームワークを模索する。

1. はじめに

Boids と呼ばれる鳥の群れに対するモデル [Reynolds 12] が提出されて以来、膨大な量の研究が行われている。比較的最近になり、「群れ」全体を単位とみなし意志決定をさせる研究が現れ始めた。本稿では、「群れ」を柔らかい構造を持った「組織」の抽象モデルと考え、それを組織デザインに応用する方法について考える端緒としたい。

1.1 「抽象的組織」としての自己駆動粒子系

「自己駆動粒子系 Self-propelled Particles」とは、多くの変数を含む Boids を単純化（例えば、速度を一定にするなど）し、研究しやすくしたモデル [Vicsek 95] の名称である。自己駆動粒子系以外にも、Boids を起源とする群れのモデルは膨大だが（最近のレビューとして [Romanczuk 12] など）、本稿ではそれらを緩く総称する意味で、「自己駆動粒子系」と呼んでおく*¹。

自己駆動粒子系は、大まかにいって、次の三つの規則で結合した個体 = 粒子（鳥など）の集団（= 群れ）である。各個体は、自分の周囲にいる個体群（近傍）に関して以下のようなこと計算し自分の動きを決める。まず、個体間には、衝突を回避する斥力が働き、近傍にいる他個体と一定以下に接近すると、それらを選ける向きに方向を変える。次に、自分の近傍にいる個体群と方向を揃え、さらに遠ざかり過ぎることのないよう、近傍の中心に向かう引力も働く。最後にこれらの力で決まる方向がノイズで揺らぐ*²。ノイズはモデルに入りきらない様々な要因（個体の意志や風などによる乱れなど）を表現している。これらの力の結合定数、ノイズの大きさ、近傍の決め方など様々な要因が絡み、様々な定式化がありうる。

粒子を「個人」、形成される群れを「組織」とすれば、それは自生的組織の抽象的モデルと考えることも出来る。「自生的」

というのは、人間の組織には、固定された命令系統のような、ある意味外から与えられた構造があるが、自己駆動粒子系が成す群れの場合、方向が揃ったり、階層構造ができるとしても、相互作用ルールから自己組織化するものだからだ。しかし、人間の組織にしても、原始的には、このような集団の自生があったともいえるし、命令系統については後から付加もできる。その意味で、自己駆動粒子系は、「組織」の非常に抽象的なモデル、もしくは組織内で非公式的・自生的に起きるコミュニケーション構造の理論と考えてもよいだろう。

1.2 自己駆動粒子系による意志決定とフレームワークの必要性

しかし、自己駆動粒子系が、鳥の群れのように全体として揃った振る舞いを見せても、それだけでは人間の組織と比較できない。人間の成す組織は、ほとんどの場合、なんらかの意志決定を下すが、単に自己駆動粒子系の物性を研究しても、意志決定に関しては知見が得られないからだ。

しかし、既に群れによる意志決定に関する研究は始まっている。たとえば 5% から 10% という比較的小数の個体が情報を与えられただけで、群れ全体が方向を決定し、しかもその小集団内での非常に僅かな意見の差が、全体の方向を変更するという研究があり [Couzin 05] 注目された（一種の続編として [Couzin 11]）。また、擬似的な敵のいる流路を用いた選択 [Ward 11]、光勾配の認識と回避運動 [Berdahl 13]、近傍への信頼度を含めたベイズ的決断モデル [Pérez-Escudero 11]、敵からの逃亡 [Ioannou 12]、選択肢の数による効果 [Nicolis 11] [Miller 13] などもある。

一方、著者は、人間の認知バイアスと二院制にヒントを得た組織的意志決定モデルを既に提出している [西川 12] が、これは自己駆動粒子系とは無関係である。しかし、これらの研究は両方とも組織の意志決定を研究し、そこから何らかの組織デザインのヒントを得ようとしているとも解釈できる。が、両者は今のところ比較可能な形式を持たない。そこで両者を抽象的に比較するフレームワークとして、強化学習における単純なタスクである「*n* 本腕バンディット問題 [Sutton 98]」を用いたい。

n 本腕バンディット問題とは、簡単に言えば、*n* 個の当たり確率が異なるスロットマシン（腕 = 選択肢）があり、プレイヤーは限られた回数で、できるだけ当たりの回数を増やすという問題である。プレイヤーは確率を知らないの、一応多くの腕を引いてみる必要がある（探索）が、どこかで決断を下し一番良いはずの腕に集中した選択をした方がよい（利用）。探

連絡先: 113-0033

東京都文京区本郷 7-3-1

東京大学大学院情報学環 西川麻樹

e-mail: asaki@iii.u-tokyo.ac.jp

*¹ 例えば本稿で主に参照する [Berdahl 13] のモデルは、[Vicsek 95] と Boids の中間に位置するような定式化であり、斥力、方向、吸引の結合定数を無くすかわりに、斥力を最優先する、などの変更を行っている。

*² 便宜的にこれらの規則を順番に適用するように書いているが、多くのモデルでは三つの力は適当な割合で混ぜられる。

索と利用は基本的にはトレードオフ関係にあるので、それをうまくバランスさせるアルゴリズムが求められ、既に盛んに研究されている（幾つか挙げると、UCB[Auer 10]、SoftMax、-Greedy[Vermorel 05]、LS[篠原 2][Takahashi 11] など。）本稿では自己駆動粒子系にバンディット問題を解かせ、その意志決定プロセスに関し予備的な研究を行いたい。

2. モデル

2.1 本稿での自己駆動粒子系の定義

前述したように自己駆動粒子系には Boids 以来様々な形式がある。ここでは、既に似たタイプの意志決定研究に使われているという理由から、なるべく [Berdahl 13] に近づけた定式化を行う。

まず系は 2 次元とし、 N 体の個体が存在する。各個体 i は、 $(x_{it}, v_{it}, m_{it}, \theta_q)$ という変数の組である。ここで x_{it} は時刻 t での位置、 v_{it} は速度である。 x_{it}, v_{it} への自己駆動粒子系としてのアルゴリズムは [Berdahl 13] と同じものを採用するのでそちらを参照されたい。ただし、バンディット問題に対する報酬記憶である m_{it}, θ_q については、該当するものがないので、以下に説明する。

2.2 問題の定式化

群れの「意志決定 = 腕を引く」を、本稿ではあえて抽象化し、「群れが二つの道のどちらに移動するか？」というような具体的な定式化をしない。その理由は、1) 結果が具体的な環境（「道」の作り方など）に依存する可能性が高いため、抽象的な「意志決定」の方が望ましい、2) 抽象的な定式化の方が、具体的なアルゴリズムを利用する幅は広がる、などである。

そこで本稿では、「意志決定する = 群れの平均速度 V_t が、個体の最高速度 v_{max} の一定割合以上になる $|V_t| > v_{max}c$ 」とする（ここで $|V_t|$ は群れの平均速度ベクトルの大きさ、 c は適当なパラメータ）。これを「意志決定」とする理由は、群れを成したが、個体達の方向性が全くランダムな場合、群れは一方所に留まり、平均速度はゼロに近くなる一方、ある方向に揃って移動する群れでは、平均速度が最高速度に近くなっていくことによる。

本稿では、選択肢は 2 本とする。そして例えば、「90 度と 270 度」のような選択肢集合 O を、「角度 a の単位ベクトル v_a の集合」として考える。意志決定が行われた時は、 O の要素と、意志決定時の平均方向ベクトル V_t との内積 $V_t v_a$ をとり、大きい方を群れが選択したことにする。

意志決定が行われた場合、あらかじめ決めておいた確率 $P_O = (p_{a1}, p_{a2})$ に従って、「報酬 r_t 」が与えられる。報酬の分配は均等とし、 $r_t = 0$ または 1 である。この問題の「正解」は、確率が高い方の腕を引き続けることであるが、「正解」を引いても報酬が貰えるとは限らないことに注意されたい（これは、「光強度」や「敵の姿」が常に存在し、多少揺らぐタイプの意志決定課題との違いとなる）。

2.3 記憶形式とフィードバック

群れによる意志決定が行われたとして、個体はその結果を学習する方式には非常に大きな自由度がある。ここでは、以下の三つを問題にしよう。

1) 意志決定後に、「結果 (= 報酬の有無)」との関連で、個体が自分の「状況」をどのように認識するのか？

2) ある状況とその結果の組が得られた後、それを記憶するアルゴリズムをどうするか？

3) 記憶をどのように行動にフィードバックするのか？

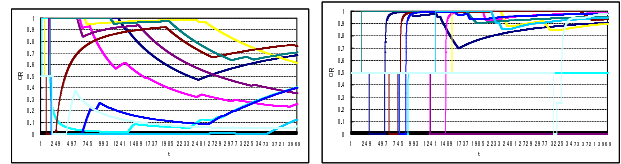


図 1: 正解率の時系列 (右は学習有り、左は無し)

1) に関しては、たとえば既に、周囲の個体に対する信頼度と、自分の得た情報への信頼度をバランスさせるようなアルゴリズムが知られている [Pérez-Escudero 11]。しかし、本稿では可能な限りモデルを単純にするため、「状況」を、意志決定が行われた時の自分の角度とする。角度は 0-360 度の間で q 分割された区間で粗視化して表現される。 $\theta_{i,q}$ をエージェント i の q 番目の区間としよう。 q の最低は 2 で、上限は一応 360 としておく。ここで注意して欲しいのは、群れの平均方向 $\theta_{t,N}$ やその選択 a_t と、その状況での、ある特定個体の方向 $\theta_{i,t}$ には、かなりの差がありうることだ。だから、群れの平均方向から外れた向きの個体は、報酬の有無に関し、間違った記憶（例えば全く逆方向と相関させて憶えるなど）を持つこともある。

また、2) に対し、本稿では、分割された角度に対する条件付き確率 $m_{it}, \theta_q = \frac{r_t(\theta_{i,q})}{n_t(\theta_{i,q})}$ を用いる。 $r_t(\theta_{i,q})$ は、時刻 t までの区間 $\theta_{i,q}$ の報酬総計、 $n_t(\theta_{i,q})$ は該当区間の意志決定総計である。エージェントは意志決定が行われた時刻で、「(全体ではなく)自分の」方向 $\theta_{i,t}$ が、どの区間 $\theta_{i,q}$ に含まれているかをまず決定する。その後、その区間に対する総計 $n_t(\theta_{i,q})$ を一つ増やし、 $r_t(\theta_{i,q})$ に r_t (0 または 1) を加える。意志決定が無い時刻では、 m_{it}, θ_q は変化しない。記憶の方法 (= バンディット問題への戦略) には、前述したように他にも多くの選択肢がありうる。が、本稿では個体の能力ではなく、組織のそれに関心があるので性能よりも単純さが優先される。

また、3) に関しては、群れによる光勾配学習モデル [Berdahl 13] に、試行回数が増えるにつれ漸近的に近づくように、 $|v_{i,t+1}| = v_{min} + m_{it}, \theta_q (v_{max} - v_{min})$ という形の式を使う。これは、[Berdahl 13] でのフィードバック式 $|v'_{i,t+1}| = v_{min} + l_x (v_{max} - v_{min})$ の、 l_x (ある位置 x での光強度) を、 m_{it}, θ_q に置換したものである。つまり、記憶によって、報酬が与えられる頻度の高い方向に対しては、速度が大きくなるようなフィードバックをかける。[Berdahl 13] で「光強度」は、例えば値 0.2 のように一気に与えられ、ノイズを除きほぼ一定だが、バンディット問題では、一回一回の選択は 0 または 1 の報酬を与え、それが多数回繰り返された後、理想的には、その状況（この場合は方向）の報酬頻度が判明することになる。ここで「理想的には」と断るのは、前述したように、個体は誤った学習をする可能性があるからだ。いわば本稿の問題設定は、[Berdahl 13] で一瞬で計算される光強度を、群れの構造を考慮して「微分」したような位置づけになるともいえる。

3. 結果

3.1 学習可能性

図 1 は、 $P_O = (0.7, 0.3)$ の選択肢集合に関する「正解率」の時系列である。正解率とは、「(実際に報酬が与えられたかどうかとは関係なく) 最も高い確率の腕を選んでいる回数/全試行回数」 $CR_t = \frac{n_{t,a}^*}{n_t}$ である（ここで $n_{t,a}^*$ は最も高い確率で報酬を与える選択肢の選択総計、 n_t は、全選択総計）。

初期置 (x_{i0}, v_{i0}) を一様乱数で変えた 10 回の群れ時系列を、

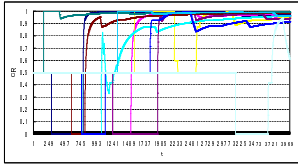
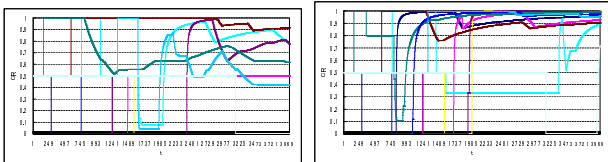
図 2: 記憶の粒度を $q = 10$ とした場合の正解率

図 3: 問題の構造を変えた場合の成功率

全てプロットしてある^{*3}。

左が学習のない場合、右が学習のある場合のグラフで、左では正解率が 1 に収束する傾向があることから、群れが正解方向方向を学習している事が分かるが、その収束開始タイミングはバラバラである。

その理由として初期段階の水平のラインがみられる。これは、一度正解し、そのまま変わらない場合と、(意志決定は、ある程度方向が揃わないとなされないため) 同じ正解率のままずっと意志決定せず、正解率が固定されている状況があるためだ。また、グラフが垂直に変化し、その後水平に保たれるのは、例えば、1 回正解し、その後、ずっと正解し続ける(この場合、後で下がっても下げ幅は小さい)か、意志決定をしない(分母 = 意志決定総計が小さいので、急激に下がる)ような事例を意味する。この場合、群れは内部に大きな揺らぎをためながら、ふらふらと動き回っていることが多い。学習のある右のグラフでも、ときどき正解率が下がり、再度上昇している。これは、群れが内的な揺らぎで全体の方向を変えてしまっても、またすぐに正しい方向を選択し直す能力を持っていることを意味する。もちろん、このような方向転換は偶然にも起こりうる。が、左右のグラフを見ると、両者の傾向が異なる事が分かる。

3.2 粗視化の効果

先の結果は $q = 2$ という極めて粗い構造の記憶を持つ個体群によるものだった。同じ問題に対し、記憶の粒度を変えてみたのが次の図 2 である。 $q = 10$ では、学習の開始が多少遅れ、かわりに収束の精度が多少良くなるが、その傾向はほぼ $q = 2$ と似たようなものになる。

この結果は、「選択枝の分割方向」と「個体の記憶分割方向」が揃っていることに原因がある。通常バンディット問題では、選択枝集合 = 可能性に関しては、個体が正確な知識を持つことを仮定する。だから、全体と個体の問題 = 選択構造は同じである方がむしろ自然ともいえる。だが、ここで、選択枝の方向を個体と問題で 90 度変えてみよう。すると、結果は図 3 のようになる。左が $q = 2$ 、右が $q = 10$ で、選択枝の角度以外は前の図と同様のパラメータである。選択枝の回転に対し、 $q = 10$ がほとんど結果を変えないのに対し、 $q = 2$ は問題と記憶の構造が全く違うため、なかなか正解できなくなってしまう。

*3 パラメータは $N = 100, L = 480$ (L は空間の大きさ)、 $q = 2, c = 0.4, CR$ の初期値は 50 %, m_{it, θ_q} は、全ての区間に対し $\frac{1}{2}$ で初期化した。以下特に断らない限り同じ設定。

4. 考察

4.1 他の自己駆動粒子系による意志決定研究との関係

自己駆動粒子系の意志決定研究で、[Ward 11] は似た枠組みであるが、擬似的な敵や流路というモデルに依存している。それに対し、本稿は群れの平均方向 = バンディット問題における腕の選択という抽象化をすることで、他の意志決定、強化学習モデルとの比較を可能にしている。

一方、光勾配の学習 [Berdahl 13] では、ローカルな光強度のみで群れが光の複雑な明暗を学習する。これに対し、本稿は、光強度 = 選択無限大における挙動という方針でモデル化した。これにより、本稿の枠組みは、学習のプロセス研究を可能にする。それは本稿の予備的研究段階でも、図に見られる学習タイミングのずれという形で見えているが、より詳細には、以下に述べる群れの内部構造という視点が不可欠だと思われる。

4.2 組織の構造と群れの形状効果

本稿では、既に意志決定研究に使われているという理由から、Couzin らの自己駆動粒子系 [Berdahl 13] に近いモデルを採用した。しかし、このモデルで形成される群れの形状はパラメータ減少の副作用か、実際の群れとはかなり異なる。一方、実際の群れの形状特性(進行方向に伸びすぎない、急激なターン、方向、速度の相関係数のべき分布など)を模倣できるモデルも既にある [Hemelrijk 12][Niizato 12]。

このようなモデルに、本稿で与えた記憶モデルを与え、結果を比較することは興味深い。シミュレーション画面を観察すると、群れが間違った解にはまり抜け出せなくなる原因の一つが、群れの形が伸びきってしまい、実際の群れでみられる急ターンができないことにあるのが分かる。急激なターンを実現するには内部構造の複雑な秩序が有効である [Hemelrijk 12] ことが示唆されているから、本稿のフレームワーク内でその効果を確認したい。それは、単純化された群れモデルによる自己駆動粒子系意志決定研究とは異なる方向性を与える可能性がある。

4.3 記憶の形式と「問題」に対する知識表現

ところで $q = 2$ の成績が良いのは、全体の選択枝構造と、個体の記憶構造に矛盾が無いためであった。つまり「正しい選択枝」がどれかは分からないが、「選択枝の正しい集合」については、個体と問題が共通の土俵を持っていた。しかし、現実の問題に照らした場合、むしろ「選択枝の正しい集合」についての知識を、個体が知っているとは仮定する方が難しい。

全体の選択枝の方向と、個体の記憶方向がずれていることを、むしろ積極的に利用し、そもそも選択枝の空間がどのようになっているのか知らないことの有効性についてモデル化・研究することも可能であり、それは「人間」のような複雑な個体が成す組織、あるいは、個体の可能性と全体の可能性が、相互にキャンセルしあうような新しい問題のあり方について有効な視座を与えるかもしれない。本稿のモデルには内部観測的方法論を入れる余裕が無かったが、群れの構造に関する [Niizato 12] の知見と、全体と個体の可能性のズレは、組織的意志決定という問題に、内部観測の視座を入れる二つのポイントになりうる。

4.4 組織論に対し群れ研究が持つ意義

本稿の定式化では、群れは常に意志決定をしているわけではない。実際、群れの形状を動的に変更できれば、都合の悪い方向の時や、未定の場合には意志決定頻度を下げて先送りし、既に確信が高いような角度にいる場合は、逆に群れの形状を尖らせて連続した選択を行う、といった振る舞いがみられる。

これは「群れ」という「アルゴリズム」特有の「迷い」を表現しているともいえる。このような迷い方が有効なものかどうかは、現段階では分からない。が、それを判定する枠組みとして、他の明示的空間的要素を持たないアルゴリズムと比較することは自然だろう。

筆者は既に全く別の組織的意志決定のモデルを提出している [西川 12]。それは、組織による意志決定という以外、ほとんどどの共通点も持たないアルゴリズムであるが、抽象的なフレームワークゆえに比較が可能になる。

また、群れが多くの計算資源を用いた並列計算である以上、単独個体より優れた意志決定モデルが出現しても不思議はない。事実、限られた条件ではあるが、[西川 12] では、単独でのアルゴリズムより良い成績を示すことがある (なお、[西川 12] では、本稿と対照的に、記憶アルゴリズムを取り替えることで、アルゴリズムによらない結果を取り出そうとしている)。

組織論の観点からは、群れに実際には存在しない「紐」のような構造を付加することも興味深い。これは、人間の組織が、通例短期では変動しにくい構造を持つことを、「紐」で表現し、それが群れの意志決定にどのような影響を与えるのかをみることになる。自生した組織に、例えばツリー状の命令経路を強制した場合、その効果や副作用は何か? 命令系統を役職や責任という概念で固定することが、有効な状況とそうでない状況とは何だろうか? 我々は、自身に適した組織デザインの仕方を、未だ知らないかもしれない。

参考文献

- [Auer 10] Auer, P. and Ortner, R.: UCB revisited: Improved regret bounds for the stochastic multi-armed bandit problem, *Periodica Mathematica Hungarica*, Vol. 61, No. 1-2, pp. 55–65 (2010)
- [Berdahl 13] Berdahl, A., Torney, C. J., Ioannou, C. C., Faria, J. J., and Couzin, I. D.: Emergent Sensing of Complex Environments by Mobile Animal Groups, *Science*, Vol. 339, No. 6119, pp. 574–576 (2013)
- [Couzin 05] Couzin, I. D., Krause, J., Franks, N. R., and Levin, S. A.: Effective leadership and decision-making in animal groups on the move, *Nature*, Vol. 433, No. 7025, pp. 513–516 (2005)
- [Couzin 11] Couzin, I. D., Ioannou, C. C., Demirel, G., Gross, T., Torney, C. J., Hartnett, A., Conradt, L., Levin, S. A., and Leonard, N. E.: Uninformed individuals promote democratic consensus in animal groups, *science*, Vol. 334, No. 6062, pp. 1578–1580 (2011)
- [Hemelrijk 12] Hemelrijk, C. K. and Hildenbrandt, H.: Schools of fish and flocks of birds: their shape and internal structure by self-organization, *Interface Focus*, Vol. 2, No. 6, pp. 726–737 (2012)
- [Ioannou 12] Ioannou, C., Guttal, V., and Couzin, I.: Predatory fish select for coordinated collective motion in virtual prey, *Science*, Vol. 337, No. 6099, pp. 1212–1215 (2012)
- [Miller 13] Miller, N., Garnier, S., Hartnett, A. T., and Couzin, I. D.: Both information and social cohesion determine collective decisions in animal groups, *Proceedings of the National Academy of Sciences* (2013)
- [Nicolis 11] Nicolis, S. C., Zabzina, N., Latty, T., and Sumpter, D. J.: Collective irrationality and positive feedback, *PLoS One*, Vol. 6, No. 4, p. e18901 (2011)
- [Niizato 12] Niizato, T. and Gunji, Y.-P.: Fluctuation-Driven Flocking Movement in Three Dimensions and Scale-Free Correlation, *PloS one*, Vol. 7, No. 5, p. e35615 (2012)
- [Pérez-Escudero 11] Pérez-Escudero, A. and Polavieja, de G. G.: Collective animal behavior from Bayesian estimation and probability matching, *PLoS computational biology*, Vol. 7, No. 11, p. e1002282 (2011)
- [Reynolds 12] Reynolds, C.: Boids (Flocks, herds, and schools: a distributed behavioral model), *URL http://www.red3d.com/cwr/boids/*. Last checked on November (2012)
- [Romanczuk 12] Romanczuk, P., Bär, M., Ebeling, W., Lindner, B., and Schimansky-Geier, L.: Active Brownian particles, *The European Physical Journal Special Topics*, Vol. 202, No. 1, pp. 1–162 (2012)
- [Sutton 98] Sutton, R. S. and Barto, A. G.: *Reinforcement learning: An introduction*, Vol. 1, Cambridge Univ Press (1998)
- [Takahashi 11] Takahashi, T., Oyo, K., and Shinohara, S.: A loosely symmetric model of cognition, in *Advances in Artificial Life. Darwin Meets von Neumann*, pp. 238–245, Springer (2011)
- [Vermorel 05] Vermorel, J. and Mohri, M.: Multi-armed bandit algorithms and empirical evaluation, in *Machine Learning: ECML 2005*, pp. 437–448, Springer (2005)
- [Vicsek 95] Vicsek, T., Czirók, A., Ben-Jacob, E., Cohen, I., and Shochet, O.: Novel type of phase transition in a system of self-driven particles, *Physical Review Letters*, Vol. 75, No. 6, pp. 1226–1229 (1995)
- [Ward 11] Ward, A. J., Herbert-Read, J. E., Sumpter, D. J., and Krause, J.: Fast and accurate decisions through collective vigilance in fish shoals, *Proceedings of the National Academy of Sciences*, Vol. 108, No. 6, pp. 2312–2315 (2011)
- [篠原 2] 篠原修二, 中野昌宏: 本腕バンディット問題に対する「緩い対称性モデル」の有効性: 因果推論における対称性バイアスと相互排他性バイアス, *進化経済学論集 第 11 集* (2)
- [西川 12] 西川麻樹: 不確実な環境で決断する組織のデザインフレームワーク: そのベンチマーク例としての「緩い対称性型組織モデル」, *情報学研究: 東京大学大学院情報学環紀要*, No. 83, pp. 51–68 (2012)