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 | ―

Design Framework for Organizations Confronting Uncertain Environments: "Liner Loose Symmetry Society Model" as an Example for Benchmarking Organizational Decision-makings

西川 麻樹*

Asaki Nishikawa

1. はじめに:不確実で変化する環境下における、社会組織による意思決定

人と同様、組織や社会も、常に意思決定を迫られている。しかも多くの場合、選択肢の評価には不確実性が伴う。それは「選択肢Aを選ぶと0.7の確率で・・・」というような形すらしておらず、そもそも、確率が不明であることが多い。ゆえに、各選択肢に対し調査をしなくてはならないが、調査費用は有限である以上、「確率」そのものに到達する遥か以前の段階で、「だいたいの情報での決断」を迫られる。さらに、環境は変化して行く。たとえば、

さらに、環境は変化して行く。たとえば、 1980年代から2010年代まで有効だった「レン タルビデオ」店舗、というビジネスモデルは、 「ネットワーク配信」という選択肢に比べ、い つまで優位性を保ち続けられるのだろうか?組 織は決断しなくてはならない。

しかし、組織や社会の構造によっては、たと え「新しい選択肢の方が良い」と多くのメン バーが判断しても、なぜか古い選択肢が選ばれ 続け、環境と乖離し続けるかもしれない。長い 目で見れば、それは組織や社会にとって危険で あろう。

本稿では、このような不確実かつ変化する環境下で、意思決定を迫られる社会や組織に対し、計算機の助けを得ながらその構造をデザインするフレームワークの提案を行う。

1.1 不確実性下の意思決定とバンディット問題

上述のような、不確実状況下での意思決定を うまく抽象化した、「n本腕バンディット問題」というタスクが、人工知能、強化学習分野 で知られている¹。 「バンディット問題」とは、以下のような設 定の学習課題である。

1) 環境:「当たり」の出る確率が違うn個のス

^{*}東京大学大学院情報学環

キーワード:社会システム論、組織論、社会的意思決定、バンディット問題、生体工学、認知バイアス、基礎情報学、情報社会、 心理学、複雑ネットワーク

本腕の海賊 | という意味) があるとする (一 般に、nは2以上で、問題によっては非常に 多くなるが、本稿では2に固定する)。

- 2) 学習主体(プレイヤー):以下のような制 約を持つ学習主体を想定する。
 - どのバンディットの「当たり」確率に関して も事前情報をも持たない(探索の必要性)
 - ・限られた回数しかバンディットを試すこと ができない(資源の有限性)
- 3) 目的:学習主体は限られた資源を使って、 できるだけ多くの「当たり」を得たい。

学習主体は、「当たり」確率が大きいバン

ロットマシン(=バンディット、つまり「一 ディットがどれか分からないので、とりあえず 引いてみる必要がある(「探索」)。一方、ど のバンディットが一番か分かった場合、他は引 かず、それに集中した方が良い(「利用」)。 しかし、確率は与えられず、探索に使える資源 も有限であるため、どのタイミングで「探索」 と「利用」を切り替える決断をするのかが、難 しい問題になる。

> この問題に対する学習戦略 = アルゴリズム は数多く提案されている²。また「変化する環 境 | を表現するため、正解(「当たり | 確率が 高い腕)が、途中で変化していく問題設定も研 究されている³。

1.2 既存研究:マルチプレイヤーによるバンディッド問題に関するゲーム理論的アプローチ

従来の研究では、「単独の」学習主体(プレ イヤー)がバンディット問題を解くという暗黙 3)選好バイアスに従って、プレイヤーが独立に の前提があった。しかし、何人かの研究者は、 複数の主体がグループとしてバンディット問題 (やそれに類似したタスク)を解くというアプ ローチを試みている4。これは本稿で問題にす る「社会組織が、一つの学習主体として不確実 性下での決断をする | という枠組みと関連し興 味深い。

それらの研究には幾つかの前提がある。

- 1) 比較的少人数のプレイヤー数
- 2) プレイヤーは選択肢に関する選好バイアス

(好み)を持つ

- 選択を行うと、社会全体としてのコストが上 昇する (例えば、正しい腕が選べない)
- 4) 「投票」のように、全体の意見を集約し トップダウンかつ強制的に意思決定を行う のではなく、あくまでボトムアップの意思 決定を行う。

これらの前提は、従来の研究が、基本的に ゲーム理論的なアプローチを採用し、数理的解 析を目標とすることに由来すると思われる。

1.3 本稿でのアプローチ:多様で多数のエージェントからなる組織(社会)を意思決定主体として考える

一方、本稿の研究では、「学習主体」を「異 質で多数の主体からなる、ある構造を持った一 つの社会組織」と読み替える。そして、どのよ

うな「組織構造」なら、不確実な環境への適応 ができるのか?というテーマを研究できるよ う、以下の前提をとる。

- 1) 社会組織は「多数決」のような、トップダ ウンに意思決定結果を強制する制度を持つ
- 2) 社会組織のメンバーは、独立に、バイア スと揺らぎを持って、ボトムアップの意 思決定を行うが、近隣の他者からの影響 も受ける
- 3)組織を構成するメンバーは、異質で、し かも非常に多く存在する

先行研究と異なり、本稿のアプローチはゲー

ム理論とほとんど関係を持たない。関係があると すれば、2) におけるバイアスの存在ぐらいであ る。他の前提については、むしろ先行研究とは対 照的に、多数で異質な主体達が、ボトムアップと トップダウンの意思決定機構を使用して組織全体 の意思決定をする、という構成になる。

この選択は、現実の社会組織が、このような 前提を有することが多いということによる。

1.4 不確実環境化で意思決定する組織デザインの「フレームワーク」という提案

上述のような前提をとることで、ゲーム理論 的アプローチと異なり、数理的な解析は困難に なる。その代わり可能になるのは、様々に異な る社会組織を「意思決定アルゴリズム」と読み 替え、共通の課題に対するパフォーマンスを競 う (ベンチマークを行う)という「組織デザイ ンのフレームワーク」である。なおここで、 「共通の課題」というのは、「変化し不確実な (確率分布が知られていない) 環境で、良い選 択をする」というタスク(本稿では(正解が変 化する)バンディット問題)とする。

なお、ここで一つ目の仮定を置く。

仮定1)「良い」社会組織とは、環境の変化 に適応し、選択を変えることができる社会組織 である

逆に、「悪い」社会組織とは、環境の変化が あっても、それを無視し過去の選択に固執する 社会であろう。これは一見自明に見える仮定だ

が、たとえば「環境へ適応するためメンバーの 厚生を完全に犠牲にする社会組織」というのも 想定可能だから、あくまで仮定である。

ここで、本稿の構成をまとめる。2で、バン ディット問題に対する具体的なアルゴリズムを サンプルとして一つ選び、いくつかの仮定を置 きながら、それを社会組織による意思決定アル ゴリズムに書き換えていく。3では、2のモデ ルを用いて簡単なシミュレーションを行った結 果を報告する。これは上述したベンチマークが 実際に可能であることのデモンストレーション である。なお本稿の目的は、そのような社会組 織デザインのフレームワークが存在可能である こと、同じフレームワーク上で他にも様々なデ ザインが可能であろうことを示すことである。 だから、4は、3の結果を受けての結論とその ような拡張への考察にあてられる。

2. ベンチマーク例としての「線形緩い対称性社会(組織)モデル」

前述したように、バンディット問題へのアルゴリズムは既に多数知られており、中には囲碁

などで華々しい応用を既に持つものもある⁵。

本稿では、その中から「緩い対称性モデル (Loose Symmetry Model 以下LS) しとい う、人間の意志決定と関わりの深いアルゴリズ ムをサンプルとして選ぶ。選択の理由は、LSが 非常にパフォーマンスの良いアルゴリズムであ ることと、後述する解釈と仮定により、社会組 織型アルゴリズムに変換しやすいことである。

2.1 緩い対称性モデル (Loose Symmetry Model)

以下で具体的にその変換を行うが、まずLSア ルゴリズムについて簡単に説明する。LSは篠原 ら⁶が発見したバンディット問題に対するアルゴ リズムで、その後他の研究者との協力によって、 その意味解釈や拡張モデル、人間の因果推定との 深い関係など、意義深い研究が現在進行形で行わ

この改変では、元々単独のプレイヤーが実行 するLSというアルゴリズムを、多様なバイア スを持ち、コミュニケーションする多数のエー ジェントからなる社会組織が実行するよう変換 し、「変化する不確実な環境に適応する社会| に対するモデル兼学習戦略として用いる。変 換後のアルゴリズムを本稿では、線形LS社会 (Linear LS Society 以下LLSS) と呼ぶ。

れている⁷新しいアルゴリズムである。

LSは、バンディット問題に対するパフォーマ ンスがトップクラスであるほか、特にモデルを 変更しないでも、環境が変化するタイプのバン ディット問題に対応可能であり、さらに任意パラ メータを持たないなど、優れた特徴を多数持つ。

$$CP_{A} = \frac{a}{a+b} = \frac{r_{A}}{N_{A}}$$

$$RS_{A} = \frac{a+d}{a+b+c+d} = \frac{r_{A} + \overline{r_{B}}}{N_{All}}$$

$$LS_{A} = \frac{a+(\frac{b}{b+d})d}{a+b+(\frac{a}{a+c})c+(\frac{b}{b+d})d} = \frac{r_{A} + (\frac{\overline{r_{A}}}{N_{fail}})\overline{r_{B}}}{N_{A} + (\frac{\overline{r_{A}}}{N_{suc}})r_{B} + (\frac{\overline{r_{A}}}{N_{fail}})\overline{r_{A}}}$$

$$()->0->CP$$

$$()->1->RS$$

Not rewarded b

d

Reward

С

 r_i : reward of i th arm

N_i: trial number of i th arm

N_{AII}: all trial number

 N_{suc} : trials with rewards

 N_{fail} : trials without rewards

図1 ISの計算式

LSの計算式は、「LSの計算式」という下図 の右側にある数式である。

これだけでは意味が分からないので、以下 で、簡単にその意味を説明する。ただし詳細は 参考文献にあたってほしい。

まず、図の左側にある表は、記号の定義をしている。A,Bは二本のバンディットを区別するラベル、Reward、Non-rewardedはそれぞれ、「当たり」、「はずれ」を意味する。例えば、表中aは「バンディットAを引いて、「当たり」が出た回数」を意味する。以下同様に、たとえばdは「バンディットBを引いて「はずれ」だった回数」などと読む。

右側の数式に移ると、まず一番上の行にCPという計算式が書いてある。CPは「条件付き確率(Conditional Probability)」という意味で、最も素朴なアルゴリズムの一つである。このアルゴリズムでは、腕Aに対する「評価」を決めるために、「Aの当たり回数(=a)÷腕Aを引いた回数(=a+b)」、つまり、「P(当たり | Aを引いた)」という条件付き確率が用いられている。同様にBの評価は、c/(c+d)となることも分かるだろう(なお、数式の一番右は、abcdという記号を使わず、数式部下で定義した記号で意味をとりやすくした式だが、内容は同じである)。

学習主体は、毎回腕を引く前に、それまでのabcdの表(記憶)を用い、AとBの評価を計算し、高い方の腕を選択する。このアルゴリズムには確率は入っていないので、常に評価が高い方が選ばれることと、引かなかった方の腕に関する表はアップデートされないことに注意しよう(なお、このプロセスは全アルゴリズムで

共通である)。

一方、次の行にあるRS(Rigorous Symmetry 完全対称性)というアルゴリズムは、既にかなり複雑な式になっていて分かりにくい。このアルゴリズムの特徴は、今考えている腕(たとえばA)だけではなく、もう一つの腕(たとえばB)のことも念頭にいれ、「Bが「はずれ」だった、ならば、Aが良い」という評価を入れていることと、その分母が、「すべての腕を引いた総数」という形になっていることだ。RSは人間の認知バイアスのうち「対称性バイアス」と「相互排他性バイアス」を満たすように作られており、ある意味、人間の無意識的な推論の偏りを表現したモデルにもなっている8。

最後に本稿で用いるLS(Loose Symmetry 緩い対称性)アルゴリズムについて説明しよう。前述のように、この式は篠原によって発見された後、篠原、高橋、中野らの尽力により、様々な意味付けがなされた。逆に言えば、それらの解釈以前には、パフォーマンスが良いことは分かったが、意味のよくわからない式であった。しかしながら、LSを因果推定のモデルとして採用すると、非常に人間と相関が高い⁹ことが知られている。

LSには様々な解釈が施されているが、本稿では、LSは、CPとRSの「中間」であり、その配合バランスをダイナミックに調整している、という解釈をとる。つまり、LSの()の中にある係数を0にすると、LSはCPになるが、1にすると、RSになる。だから括弧の中をCPとRSを結合するパラメータだとみなし、そこを状況に応じ、ダイナミック(括弧の中はabcdによって決まる変化する値)に変動させているアルゴリズムがLS

であるという解釈だ。なお、この結合がダイナ ミックであることは重要で、固定した場合、どの ような係数であっても人間の因果推定との相関は 低くなってしまうことも知られている。 LSについては他にも興味深い論点が多数存在するが、本稿の趣旨とは関係が無いので、この程度の簡潔な説明にとどめる。

2.2 緩い対称性モデル (LS) の線形化 (Linear LS)

前節で紹介したCP、RS、LSは、あくまで単独の学習主体が、バンディット問題に挑むというセッティングで、選択肢の評価を決めるアルゴリズムであった。

ここでは、それを幾つかの仮定の下に、社会 組織による意思決定アルゴリズムに書き換える。 まず、次のような仮定をする。

仮定2) 社会組織は、精密で判断の遅い少数の 人々(エリート)と、偏見にとらわれやすいが 判断が速い多数の人々(大衆)から構成される

これは非常に強い仮定である。現実の社会がこれほど単純な二分割で捉えられることはまずないだろうし、そもそも「エリート」や「大衆」の定義にも大いに異議があり得るだろう¹⁰。この仮定は、むしろ先の節で指摘した、「LSがCPとRSの二成分のハイブリッドとして考えられること」および「CPおよびRSの変化適応スピードについて経験的に分かっていること(CPは適応が遅いが、精密な結果に達する、RSは速いが不正確な結果になることもある)」を、「エリート」と「大衆」にマッピングしたものと考えた方が良い。ただ、本稿は細分化によるモデルの精緻化ではなく、一つの例を構成することが目的なので、とりあえずこの仮定を採用する。

次に、

仮定3)「エリート」と「大衆」は、それぞれ 意見を投票し、社会はその二つを重みづけて最 終的な意思決定を行う

この仮定は、社会組織の構造を、たとえばエリートの独裁や、乱数による意思決定にはせず、多数決原理を採用し、しかも「エリート」と「大衆」の票が持つ重みが変動しうる形にするという仮定である。

さて、以上の仮定の下で、LSを社会組織で 使えるアルゴリズムにするには、幾つかの修正 が必要になる。

まず、LSの「線形化」が必要になる。次の「LLSおよびLLSSについて」という図を見てほしい。最上段にLinearLS(LLS)という行がある。これは、LSの式を、形式的にCPとRSの足し算の形に書き換えたものだ。「形式的に」というのは、RSにかかっている重みの項wLSの定義を見て分かるように、重みをRSに掛けて計算すれば、重みの中にあるLSの項が出てきて、そのままLSに戻ってしまうことを意味している。つまり、これは単に、LSの式を「グループごとの投票」という構造に当てはめるための下準備にすぎない。

$$LinearLS(LLS) = CP + w_{LS}RS = LS$$

$$w_{LS} = \frac{-CP + LS}{RS}$$

 $LLSSociety(LLSS) = CP'(N) + w_{LS}RS'(M) \neq LS$

CP'(N) = N biased CP agents' local intractions and voting results ratio

a) Biased agents update memory as:

$$r_A(t) + 1.0 \times bias_{n,A}$$

where $bias_{n,A}$ are uniform random numbers from [0,1] fixed value for each agent.

c) Fluctuated choices are implemented by:

$$if(\frac{Q_{n,A}}{Q_{n,A} + Q_{n,B}} > = [0,1]$$
 uniform random number) then A is selected by agent n (else B)

where $Q_{n,i}$ are evaluations of arm i by agent n with algorithm Q.

図2 LLSおよびLLSSについて

2.3 LLSからLLSの社会(組織)モデル(LLSS)へ

先の図の続きをみると、LLSSociety(LLSS)という行がある。ここでは、先のLLSと違うことが行われている。まず、LLSでは単にCPやRSと書かれていた項が、CP'(N)やRS'(M)という項に置き換えられている。そのため、もはやこの左辺はLSとは等しくない。実際CP'(N)のような項は、N体のエージェントからなる「集団」を表現しており、さらに後述する内部相互作用も持つ複雑なものだ。ただし、両者を結合する重み部分についてはLLSと変わらない。

ではこの「集団項」はどのような方針で作られているのだろうか?ここで再び仮定をする。

仮定4) 各エージェントは、独立に意思決定する際に、バイアスと揺らぎにさらされる 仮定5) 各エージェントは、独立に意思決定後、 周囲の影響で意見を変えてしまうことがある これらの仮定は、モデルの作りやすさを向上 する目的ではなく、むしろ単純にアルゴリズム を集団化しても意味が無いので、より社会組織 の持つ「困難さ」を反映する方向に、モデルを シフトするために導入されたものだ。

たとえば「CP集団」は、CPアルゴリズムに従って意思決定するエージェントの集団だが、まず彼らは、非常に大きなバイアス(偏見)を持つ集団としてモデル化されている。図のaという項目をみると分かるように、各々のエージェントは、たとえば選択肢Aを引いて「当たり」が出たときに、それを単純に「プラス1」とはカウントしない。そうではなく、エージェントごとに違う値に(0から1の間にある)一様乱数で初期化された「偏見の値」を、1へ掛けた数でカウントしてしまう。だから、あるエージェントはAが当たることを、Bが当

たることの100倍評価するかもしれないし、別のエージェントは逆かもしれない。これは公平な選択肢の評価という観点からは大きなマイナスだが、現実の人間が持つ価値評価の多様性を、モデルに取り込もうとするささやかな努力である。

また、このバイアスは、「ハイブリッドな集団」をモデル化するという目的で導入されてはいるが、マッシブな集団の意思決定の中に、「ゲーム理論」的な「特定選択肢への選好」を持ち込んでいるとも言える¹¹。

次に、各エージェントは先の単独アルゴリズムと異なり、確率的に意思決定を行う。その方法が図の項目cにある。すなわち、0から1の値をとる乱数を一回使い、その値が「Aの評価÷(Aの評価+Bの評価)」以下だったらAを選び、そうでなかったらBを選ぶ、という揺らぎを持った選択を行う。

さらに、仮定5は、エージェント集団を単なる「集合」にしてしまわない工夫である。彼らは、自分で選択肢を一度決めた後、周囲の意見を聞き多数決で意見を変えてしまう。

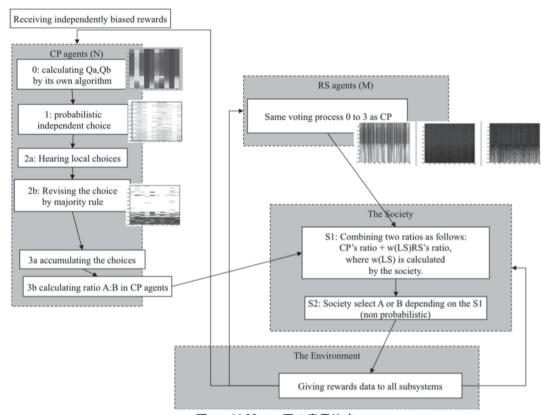


図3 LLSS:一回の意思決定

最後にこれらのプロセスをすべてまとめた「LLSS:一回の意思決定」という図を見てみよう。この図では、LLSSという社会組織が、

「一回の決断をして、その報酬を内部にいる エージェント集団に伝達する」までの過程が描 かれている。 プロセスごとに順序を表す番号がふってあるので、順に見て行く。まずCPエージェント集団に着目する(RS集団でもやることは同様)。この集団は、意思決定にCPを用いるという意味では同質エージェント集団だが、先に述べたように、バイアス、揺らぎ、周囲の環境によって、差異も持った集団でもある。

- の)彼らはまず、過去の記憶(エージェントごとのバイアスで違うabcd表。初期値はランダムに与える)とCPを使って、各腕に対する評価(Qと書いてある)を計算する。なお横にある囲み図では横軸がエージェント、縦軸が時間(上から下に進む)、色で内部状態(色が濃いほどQaの比が高い、など)を表現している。なお、RS集団も同様のグラフを表示しているが、エージェント数が多い。
- 1) それぞれ独立にエージェントが確率的に腕 を選択する。図では選んだ腕に応じて二種 類の色に塗り分けられている。
- 2) a) 周囲の選択をヒアリングし、b) 自分の意見を周囲の多数決で決める(「周囲」の決め方は後述する)。なお二色以外の色

- があるのは、意見同数の場合、とりあえず 「未決」としてグラフを作り、内部状態を 見やすくしているためである。
 - 3) a) 集団内全エージェントの選択を集計 し、b) 集団内部でのAとBの選択比を計 算する。
 - 4) (ここから先は社会全体の意思決定作業になるので、図ではS1、S2と書いてある) S1) 各選択肢の占める比率をLLSSの項に代入する(CP集団の値をCPの項に、RS集団の値をRSの項へ)。ただし、毎回変わる重みであるw(LS)の値は、とりあえず本稿では、「偏見を持たない社会全体の記憶(バイアス=1.0)」を仮定し計算している¹²。 S2) S1の結果を選択肢ごとに作り、値の大きい方を選択する(ここは確率的ではない)
 - 5)環境は、与えられた確率と引かれた腕に応じて、報酬を与える。値は0か1である。
 - 6) 各エージェントは、報酬に自分のバイアス を掛けた値を使って、記憶(abcd表)を アップデートする

以上で、一回の社会的意思決定が終了する。

2.4 三種類のコミュニケーション (集団内相互作用)

なおLLSSでは、エージェントが周囲の意見を聞いて意見を変えるので、「周囲」を決める必要がある。本稿では、とりあえず三種類のつながり方を考えた。次にある図の見方は、先にLLSSの時に説明したのと同様である。

1) WoI (上段二行): 比較用に用意した周囲と 相互作用しない集団。 (上がCP集団、下が RS集団)

- 2) CA (中段二行): 自分の番号 1、+ 1の エージェントの意見を聞く集団 (ただし周期 的境界条件とする)。自分も含めて三体の意 見で多数決をとるので、常に意見が決まる。
- 3) SFG(下段二行): スケールフリーグラフを 作り、その内部にエージェントを置いたも の。よって図の横軸に意味は無い。

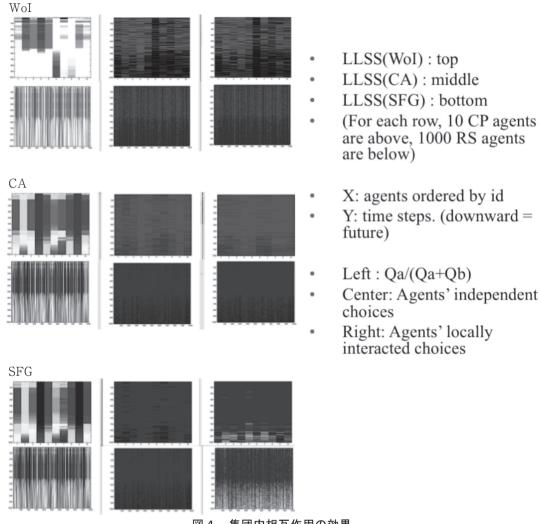


図4 集団内相互作用の効果

以上で、モデルの説明は終わる。

3. シミュレーション結果

3.1 環境と条件

LLSSのパフォーマンスを見るために、様々な難易度のバンディット問題を解かせ、アルゴリズム間比較を行った。また、本稿テーマは「変化する環境への対応」なので、最も単純な

例として、正解 (「当たり」確率の高い腕)の 途中変更を行う。

「問題=環境」としては、各腕の当たり確率 (Pa, Pb) がそれぞれ、 (10%,90%) = 「簡

単な問題」、(30%,70%) = 「普通の問題」、(40%,60%) = 「難しい問題」、の三環境を用意した 13 。この初期設定で500回選択し、その後、Pa, Pbを入れ替えることで「環境の変化」を表現している。変化後さらに500回選択させることで、合計1000回の選択を行うのを1試

行と数える。グラフでは100試行の平均値が表示されている。

エリート (CP集団) のエージェント数は 10、大衆 (RS集団) のエージェント数は1000で ある。また、各エージェントの記憶であるabcd 表は試行ごとにランダムに初期化している。

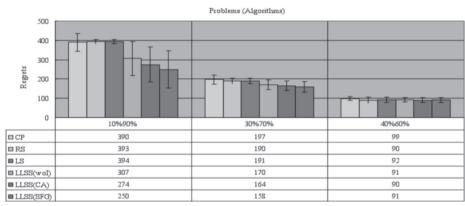
3.2 LS単体とLLSSのパフォーマンス比較

次の図は、上の設定で、各アルゴリズムの「後悔regret」の値を比較した表である。「後悔」は、バンディット問題の評価によく使われる指標の一つで、理想的なプレイヤー(=正解を常に知っている)のパフォーマンスに比べ、どの程度のビハインドで済んでいるのか、という差を表現している。従って「後悔」の値は小さい程、そ

のアルゴリズムの性能は高い。なお、「簡単な問題」ほど、「正解」を見逃すことの損が大きい (=正解の当たり確率が高い)ので、「後悔」の 差が開きやすく、逆に、「難しい問題」では、ど ちらを選んでも確率はあまり変わらないので、差 が開きにくいことに注意してほしい。

図では、まず三つの列が、左から簡単な問





□ RS
□ LS
□ LLSS(wol)
□ LLSS(CA)
□ LLSS(SFG)

□ CP

- Regret: Distance between expected rewards of answer knowing player and record of real rewards
- Regrets with different environments and alogrithms
 - LSCP
 - RS
 - LLSS(woIntraction)
 - LLSS(CA)
- LLSS(SFGraph)
 All values are rounded to intege

図5 アルゴリズムごとの後悔値比較

題、普通の問題、難しい問題に分かれている。 それぞれの列中で横軸はアルゴリズムの違い、 縦軸は100試行の後悔値平均である。また、グラフの下部には、各アルゴリズムの後悔数値が

記載してある。

左列の「簡単な問題」を見てみると、まず 左の三つ(CP、RS、LS)は単独学習主体に よる結果、右の三つ(LLSS(WoI)、LLSS (CA)、LLSS(SFG))は集団内相互作用の 違うLLSSの結果を表現している。

まず左の単独学習主体によるアルゴリズムの差は一見少ないが、CPは標準偏差の値が大きい。これはCPの場合、1試行=1000回の選択回数以内に、正解の入れ替わりに適応できない試行があるからで、この場合に大きく点数が落ちることが値の散らばりの大きさを特徴づける。LSとRSの差はほとんどないが、この二つは元々、問題の状況によっては似たようなパフォーマンスになるので、この結果で問題ない。

次にLLSSの三つを見てみると、「簡単な問題」の平均値に関しては、単独主体アルゴリズムに対し、かなりの改善を示している。これは、各エージェント内部の揺らぎとバイアス、多数決による意思決定の変更という、単独主体アルゴリズムに対する大きな摂動を考えると、興味深い結果である。まず、LLSS(WoI)をみると、集団内相互作用がなくとも、パフォーマンスの改善が起きていることが分かる。次にLLSS(CA)、LLSS(SFG)を比較すると、集団の相互作用構造が複雑なスケールフリーグラフの方で、より大

3.3 割引率の適用とその場合のパフォーマンス

前節の結果だけからは、社会組織版アルゴリズムは単独主体アルゴリズムより、平均値での性能が良いという結論を導くこともできるだろう。しかし、少し拡張したアルゴリズムを考えると、社会組織版のパフォーマンスが単独主体

きなパフォーマンス改善を示している。

ただしこれらの改善は、平均値に関してで、標準偏差は非常に大きい。これは社会全体が片方の選択肢に固執してしまい適応できなくなる事態が、前述のCP以上に多いことが原因である。

また、難易度の低い問題に適応するためには、精度を犠牲にして、短期的な記憶を元に決断する方が優位だが、その場合、当然難しい問題でのパフォーマンスが落ちてしまう。その観点から、右の二列を見てみると、まず「普通の問題」に関しては、差は縮んだものの、単独主体にアルゴリズムに対し優位である点は変わらず、「難しい問題」でもほぼ同等のパフォーマンスになっているので、単純に精度を犠牲にした訳ではないことが分かる。

もちろん、より難しい問題を使えば、精度の差がでる可能性はある。しかしながら、本稿が対象とする社会組織による選択問題では、ほとんど報酬確率が変わらない選択について間違えないことよりも、簡単、もしくは普通の問題について間違えないことの方が重要である。どちらを選んでも対して変わらない選択なら、どちらでも良いからだ。もちろん問題領域によっては非常に小さな差異を見逃さないことが重要でなこともあるが、本稿のテーマには恐らく無関係である。

版に及ばないこともある。次の図はその一例で ある。

本稿のような環境が定常でない問題を扱う場合、アルゴリズムの記憶に何らかの操作をして 適応能力を上げるのが通例である。なかでも代



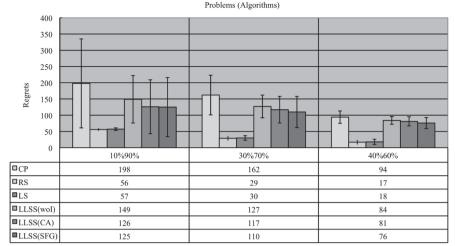


図6 割引率の効果

表的なのは「割引率」という1より僅かに小さな値を、毎回アップデートする際、過去の記憶値に掛けることで、古い記憶を忘れて行く手法である。単純だが効果は絶大で、この図では、LS、RSの成績が劇的に向上している(割引率0.99)。一方、同じ割引率の社会組織版は、二割程度性能向上をしているものの、単独主体ア

ルゴリズムの向上には及ばない。

割引率はアルゴリズムや問題ごとに最適な値が異なるパラメータなので、社会組織版アルゴリズムに対して値が適切ではない可能性はあるが、試行錯誤した限りではそれほど大きな差異はなく、単独主体アルゴリズムに追いつくには、何か抜本的な対策が必要かと思われる。

□СР

□RS □LS

■LLSS(woI)

■LLSS(CA) ■LLSS(SFG)

4. 結論、考察、今後の課題

4.1 結論

本稿では、不確実な環境の変化に適応する組織設計フレームワークを提示するため、人工知能研究でよく知られたバンディット問題をベンチマークに用いた。また組織設計のサンプルとしてLSアルゴリズムの社会組織版をモデル化し、元のアルゴリズムと比較を行った。

社会組織変換の時に加えた揺らぎや大きなバイアスがあるにも関わらず、割引率なしの社会

組織版アルゴリズムは、単独主体版を上回るパフォーマンスを見せた。また、集団内相互作用の効果でパフォーマンスが向上し、しかもそれは集団構造が複雑な方が、高い上昇率を示した。一方、割引率を入れた場合のパフォーマンスについては、単独主体版アルゴリズムを下回る結果になった。

4.2 考察

ここに示したのは、「緩い対称性モデル LS という、バンディッド問題で比較的パ フォーマンスの高いモデルを社会組織型に変形 したアルゴリズムである。前述したように、 LSモデルは、人間の認知プロセスとも関連が 深いモデルだから、本稿の社会組織モデルは、 一種の「生体工学」とも言える。粘菌の動作原 理を基にした社会や、ニューラルネットのよう にSNS+重み付けによって意思決定をする社会 組織など、「生体の意思決定構造にヒントを得 た社会組織デザインとそのベンチマークとして のバンディッド問題」という枠組みは、本稿以 外のサンプルでも構想できるだろう。本稿のモ デルはその一例に過ぎず、生体の仕組みを応用 した社会組織デザインという分野の始まりに過 ぎない。

また、ここで採用したLLSSアルゴリズムは、現在の環境に応じてRS(仮説生成的だが誤ることもある)とCP(最終的には正しいが、遅い)の重み付けを、RS部分に掛けられたウェイトを通じて、動的にバランシングしている。このバランシングは、「組織全体としてどの程度リスクをとるのか」という意味を持つと解釈できるから、リスクと社会という文脈でも興味深い。恐らく、リスク許容のみでもリスク拒否のみでもない、状況によって変化する、別のリスクに対する態度がありうる。

最後に、本稿にネオ・サイバネティクス的

4.3 今後の課題

本稿のフレームワークには、非常に多くの課 題が残されている。まず、本稿ではアルゴリズ 観点¹⁴からのコメントをしておく。本稿の考察は、「社会組織」を一つの「行為者」とみなす視点の取り方を前提としている。さらに、その視点位置は、「社会組織内部で組織構造を設計しうる立場」に置かれている。この組織内組織設計者は、組織のメンバーでありながら、ある意味組織全体を変更できるメタレベルにいる。この二重性が問題になることもある¹⁵が、本稿のモデルでは、「LS的重みによる集計の強制」という形で二重性は回避され、設計者のメタレベル性は安定に確保されている。これは、システムを動かす特殊な論理を、無根拠に選択しているという意味で、「オートポイエーシス」と似た立ち位置ともいえる¹⁶。

もし、この無根拠な選択が問題になるとした ら、恐らく「設計」を「実現」もしくは「運 営」する時だと思われる。設計者が、外側に安 定した立場を確保している限り、それは設計 →実現という一方向的流れになるだろう。そし てそれは、様々な任意パラメータを外側からど う設定するのかという問題として表面化するこ とになるはずだ。その時に必要になるのは、 「(パラメータによらず)不可避的にこのよう な組織になる」というロジックのモデルへの内 包だろう。それがなされないなら、社会の設計 は、「社会主義革命」のような、設計のトップ ダウン強制と現実との乖離という悲劇を繰り返 す可能性がある。

ムのサンプルとしてLSをとりあげたが、他のアルゴリズム群を取り上げない理由は無い。ま

た、環境に関しても、1000回選択500回目で確率交換という非常に単純な枠組みを採用しているが、これも既に、単体版では多様な動的環境の試みが行われている¹⁷。また、本稿の目的から行けば、もっと少ない選択回数でのベンチマークが望ましいかもしれない。組織的意思決定の機会は、同一環境下では、恐らくそれほど多くはないだろうからだ。

また、たとえアルゴリズムをLLSSに絞っても研究課題は多数存在する。まず、通常の多数決との比較をもう少し精密に行う必要がある。 LLSSの特徴は、動的に票の重みが変わる多数決という点にあるが、これは全ての固定した重み付け多数決(もちろん重みが1:1のものも含む)より上のパフォーマンスを示すだろうか?また、集計を担当している社会全体がバイアスを持つ場合の効果、決定論的選択と確率的選択の比較、エリートと大衆の人数比の効果なども重要な課題である。

さらに最終節に示した割引率の効果が低いこ とへの対策を探すことも必要だろう。適切な 割引率の設定は可能なのだろうか?LLSSでは 「記憶する者=学習主体」は多数存在するの で、それに合わせた新しいやり方が必要かもし れない。だが、そもそも、割引率を外から一律 に割り当てるのは、個人の記憶の癖を社会設計 者が勝手に決定していることを意味するから、 モデルとして適切ではないとも言える。

また、非常に大きな課題としては、社会組織版アルゴリズムは個人版アルゴリズムに勝てるのか?という問いがある。計算量としては圧倒的に社会組織版が大きい(ただし並列化は容易である)のに対し、パフォーマンスが向上しないのは、不利になるはずの条件(バイアスや揺らぎ)を入れているという譲歩はあるにせよ、残念な結果である。ただし、本稿でもエージェント間グラフ構造の効果は、スケールフリーグラフの方が、セルオートマトンのような格子より良いという予備的な結果が出ており、この構造をうまく利用した先に、全く想像もしなかったような社会組織のデザインが浮上するかもしれない。

#

- Sutton, Barto, 1998
- 2 バンディット問題を効率よく処理するアルゴリズムは、UCB系のもの(Auer 2010), Softmax (Daw et al. 2006), e-Greedy (Sutton 1998) など多数知られている。
- 3 Auer 2006など
- ⁴ Park 2001, Cesa-Bianchi et al. 2006, Berg 2010, Liu, Zhao 2010, Stone, Kraus 2010, Chang et al 2010, Kalathil et al. 2012 など。
- ⁵ Gelly et al. 2006
- 6 篠原 et al. 2007
- ⁷ Takahashi et al. 2010a、篠原、中野 2007、中野、篠原 2008など参照。なお以後LS関係の式の導出や解釈はこれらの論文による。
- 8 RSは、対称性バイアス、相互排他性バイアスという人間の認知バイアスに加え、排中律をも満たすように作られたアルゴリズムである。
- 9 篠原, 中野 2007。なお高橋 et al 2010bはさらに相関の高いモデルも提出している。
- 10 元々この二分割は東 2011での、熟議するエリートとネットワークをなす大衆という二分法にヒントを得て着想された。
- 11 また、報酬=効用と考えると、ここでは伝統的な効用を巡る議論とは、少し角度をずらした立場が採用されていると解釈もでき

る。伝統的に経済学では、効用について、序数的、基数的という立場が知られている。そして、序数的な方が、他者の効用の不可知性により配慮した立場ともいえる。しかし、本稿では、足し算などの様々な演算を許すという意味では基数的な効用に近い立場でありながら、入ってくる効用単位をどう評価するのかが、非常に大きなノイズにさらされているという意味で、不可知性に配慮した立場でもある。また、西川 2012では、さらに「そもそも序数的ですらない相手の状態」についての議論と、それを使った社会が構成されている。その意味では本稿の方針と西川 2012は、不可知性を巡って相補的な関係にある。

- 12 社会全体が偏見を持つことも勿論可能である。予想と異なり、それほど致命的な欠陥にはならないことが多い。ただ、まだはっきりした結論が出せないので、ここでは最も単純な「偏見の無い社会」を用いる。
- 13 各腕の当たり確率の合計は1である必要はないが、今回はそのような問題を採用している。
- 14 Hansen 2009, Clarke, Hansen 2009, Foerster 2003、日本では西垣 2004など。
- 15 西川 2011、西川 2012など。また、全く別観点から同様の問題について論じた論考としてルーマン1997がある。
- 16 郡司 2004
- ¹⁷ Auer 2006

参考文献

Auer, P., Ortner, R. (2010) UCB revisited: Improved regret bounds for the stochastic multi-armed bandit problem, Periodica Mathematica Hungarica, Volume 61, Numbers 1-2, pp.55-65

Auer, P (2006) Exploration vs. Exploitation Challenge Framework, PASCAL Second Challenges Workshop, Venice

東浩紀 (2011) 『一般意志2.0』 講談社

Berg, S. (2010) Solving Dynamic Bandit Problems and Decentralized Games using the Kalman Bayesian Learning Automaton, Thesis submitted in Partial Fulfillment of the Requirements for the Degree Master of Science in Information and Communication Technology

Cesa-Bianchi, N., Lugosi, G. and Stoltz, G. (2006) Regret Minimization Under Partial Monitoring, *Information Theory Workshop, ITW* '06 Punta del Este. IEEE, pp.72-76

Chang, H. S., Hu, J., Fu, M.C. and Marcus, S.I. (2010), Adaptive Adversarial Multi-Armed Bandit Approach to Two-Person Zero-Sum Markov Games, *Automatic Control*, *IEEE Transactions on*, vol.55, no.2, pp.463-468

Clarke, B., Hansen, M. B. N. (2009) Neocybernetic Emergence: Returning Posthuman, Cybernetics and Human Knowing, vol.16, nos.1-2, pp.83-99

Daw, N., O'Doherty, J., Dayan, P., Seymour, B. & Dolan, R. (2006) Cortical substrates for exploratory decisions in humans, Nature 441, pp.876–879

Foerster, H. v. (2003) Understanding Understanding, Springer

郡司ペギオ-幸夫(2004) 『原生計算と存在論的観測』東京大学出版会

Gelly, S., Wang Y., Munos R. and Teytaud O. (2006) Modification of UCT with patterns in Monte-Carlo Go, Technical Report No.6062, INRIA

Luhmann, N. (1997) Die Gesellschaft der Gesellschaft L&II, Suhrkamp. (=2009,馬場・赤堀・菅 原・高橋訳『社会の社会1・2』 法政大学出版局)

Kalathil, D., Nayyar, N., Jain, R. (2012) Decentralized Learning for Multi-player Multi-armed Bandits, arXiv: 1206.3582v1

Liu, K., Zhao, Q. (2010) Distributed Learning in Multi-Armed Bandit with Multiple Players, *IEEE Transations on Signal Processing*, Vol. 58, No. 11, pp.5667-5681

中野昌宏・篠原修二 (2008) 「対称性バイアスの必然性と可能性――無意識の思考をどうモデル化するか」, 『認知科学』, Vol. 15, No.3, 日本認知科学会, pp.428-441

西川アサキ (2011) 『魂と体、脳』講談社選書メチエ

—— (2012) 「形から逃げ出す生命、ガタリの夢、自身の死を悼むシステム」,『現代思想』,40 巻 10 号, pp.144-170

西垣通(2004) 『基礎情報学』NTT 出版

Park, K. (2001) Essays in Strategic Experimentation, Levine's Working Paper Archive

Stone, P., Kraus, S. (2010) To teach or not to teach? : decision making under uncertainty in ad hoc teams, *Proceeding AAMAS '10 Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems*: volume 1 - Volume 1, pp.117-124 Sutton, R. S., Barto, A. G. (1998) *Reinforcement Learning*: *An Introduction*, MIT Press

篠原 修二, 田口 亮, 桂田 浩一, 新田 恒雄(2007)「因果性に基つく信念形成モテルと N 本腕ハンティット問題への適用」、『人工 知能学会論文誌』、、Vol. 22、No. 1, pp.58–68

篠原修二,中野昌宏 (2007) 「2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, 進化経済学会

Takahashi, T., Nakano, M., and Shinohara, S. (2010a) Cognitive Symmetry: Illogical but Rational Biases, Symmetry: Culture and Science, 21, 1-3, pp.275-294

Takahashi, T., Kohno Y. and Oyo, K. (2010b) Causal Induction Heuristics as Proportion of Assumed-to-be Rare Instances (pARIs). Proceedings of the 7th International Conference on Cognitive Science (ICCS2010), pp.361-362



西川 麻樹 (にしかわ あさき)

1975年2月10日

[専攻領域]哲学、人工知能

[著書・論文] (3本まで、タイトル・発行誌名あるいは発行機関名)

『魂と体、脳 -計算機とドゥルーズで考える心身問題』(講談社, 2011年)

『形から逃げ出す生命、ガタリの夢、自身の死を悼むシステム』(『現代思想』8月、2012年)

"The Emergence and Collapse of the Self Monitoring Center in Multi-agent Systems" Lecture Notes in Artificial Intelligence, 4953, 803-812 (2008)

「所属] 東京大学大学院情報学環助教

[所属学会] 人工知能学会、科学基礎論学会、情報メディア学会など。