

群れの意思決定能力に関する基礎研究

防衛大学校理工学研究科後期課程

電子情報工学系専攻 情報知能メディア学教育研究分野

フング ニュ ハイ

令和2年3月

目次

第1章 序論	1
1.1 研究背景	1
1.2 研究の目的	3
1.3 本研究の構成	4
第2章 関連研究	5
2.1 生き物がみせる群レベルでの高度な振る舞い	5
2.2 関連研究	8
2.2.1 スwarmロボティックスに関する研究	8
2.2.2 The best-of-n 問題に関する研究.....	10
2.3 まとめ	17
第3章 BRT モデルの提案	18
3.1 はじめに	18
3.2 Iwanaga らのモデル.....	19
3.3 提案手法	22
3.3.1 アルゴリズム	22
3.3.2 バイアス分布	23
3.4 実験	24
3.4.1 BRT モデルの挙動例.....	24
3.4.2 適した選択枝の発見	25
3.4.3 エージェント数と合意時間の関係	26
3.5 まとめ	28

第4章 BRTモデルその①：	
タブーリストをエージェントによる効率的な集団探索	29
4.1 はじめに	29
4.2 提案手法	30
4.3 提案手法の特性の概算	30
4.4 実験	33
4.4.1 UEM エージェント集団の特性	33
4.4.2 UEM エージェントの比率	35
4.5 まとめ	38
第5章 BRTモデルその②：観測範囲が制限されている場合	39
5.1 はじめに	39
5.2 提案手法：BRT-LIS モデル	40
5.3 実験	41
5.3.1 実験設定	41
5.3.2 挙動	43
5.3.3 適した選択枝の探索とその特徴	46
5.4 まとめ	48
第6章 BRTモデルその③：集団行動のプリファレンスの実現	49
6.1 はじめに	49
6.2 提案手法	52
6.2.1 アルゴリズム	52
6.2.2 バイアス値の生成	53
6.3 実験	54
6.3.1 意思決定の挙動	54
6.3.2 集団行動のプリファレンスの実現	59
6.4 まとめ	62

第7章 Q-BRTモデル：2次関数を用いたバイアス分布の生成	63
7.1 はじめに	63
7.2 提案手法：2次関数を用いたバイアス分布	63
7.3 Q-BRTモデルの挙動	65
7.4 数学的な解析とその検証実験	67
7.4.1 任意の選択肢に合意できるか	68
7.4.2 投票フェーズに要する時間の概算	71
7.4.3 適した選択肢の発見に要する平均時間の概算	78
7.4.4 群れサイズのスケーラビリティの検証実験	81
7.4.5 適した選択肢の発見	82
7.5 まとめ	84
第8章 Q-BRTモデルその①： Plural voting からヒントを得た集団探索の高速化	85
8.1 はじめに	85
8.2 提案手法	87
8.2.1 アルゴリズム	87
8.2.2 バイアス分布	88
8.3 意思決定プロセスのダイナミクス	89
8.4 適した選択肢の発見に要する平均時間の概算	91
8.5 実験	95
8.5.1 探索精度の検証	95
8.5.2 探索時間の検証	96
8.6 まとめ	99

第9章 Q-BRT モデルその②：エルファロール・バー問題	100
9.1 はじめに	100
9.2 エルファロール・バー問題	101
9.3 提案手法	102
9.3.1 エルファロール・バー問題の定義	102
9.3.2 役割の導入と選択枝の設定	103
9.3.3 集団の状態の評価法	104
9.3.4 個体の意思決定アルゴリズム	106
9.4 実験	108
9.4.1 意思決定における集団ダイナミクスの例	108
9.4.2 適した選択枝の発見精度	110
9.4.3 マージン q の影響	111
9.5 まとめ	114
第10章 結論と今後の展望	115
謝辞	117
参考文献	118
研究成果	127

第1章 序論

1.1 研究背景

群れでは、個体間のマイクロな相互作用によってマクロな挙動が創発し、そのマクロな挙動によって個体の行動が影響を受け、その結果、全体を支配する秩序が発生する。マイクロの相互作用とマクロな状態との関係は通常非線形であり、適切なマクロ状態を導くマイクロな行動を解析的に求めることは容易ではない。それ故、社会性生物にみられる高度に組織的な振る舞いがボトムアップに導かれる原理や方法は研究者の注目を集めてきた。特に蟻や蜂の群れでの採餌については多くの研究がおこなわれており、個体に採餌と休息の役に割り当てる方法については優れたモデルが提案されている。大まかにいって予め必要な役割の集合(たとえば採餌と休息)とこれらの要素間の関係を個体の行動に埋め込んでおけば、例えば群れのエネルギー状態を一定に維持するように各役割の個体数を調整できる^[7,37,12]。この一方で、役割を調整しきれず群れを目標とするマクロ状態に導くことができないケースもあるはずで、このような局面で如何にして群れが望むマクロ状態に達する新しいマイクロな相互作用を発見し、記憶し、学ぶのかについて、一般的なモデルは提案されておらず、より高度な群れの自己組織化システムの構築のボトルネックとなっている。

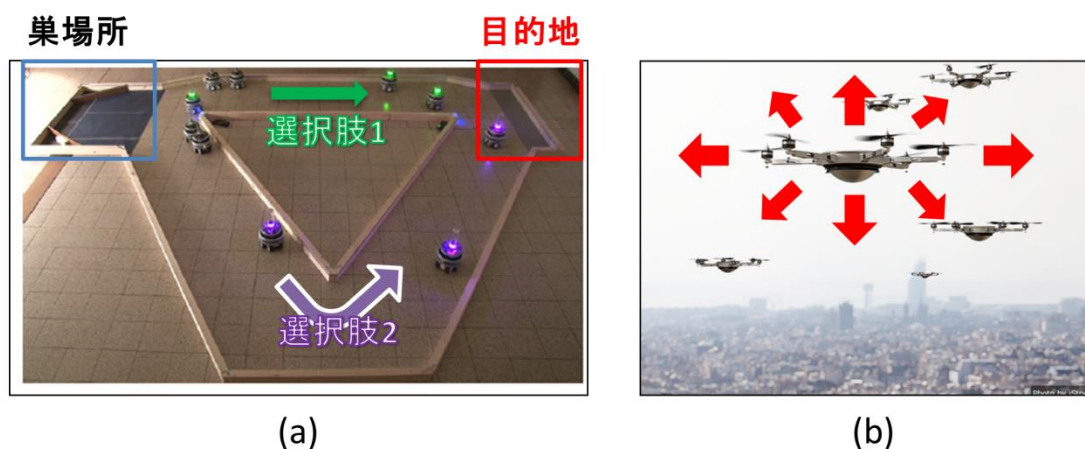


図 1.1 The best-of-n 問題の例：
(a) 2次元の問題^[67], (b) 3次元の問題。

近年、群ロボットの分野では、群れの集団的意思決定の研究として the best-of-n 問題が注目されている。これは、リーダーを使用せず分散エージェントから成る群れが n 個の候補の中から適した集団行動に合意する集団的意思決定問題である。2000 年頃に提案されたものの 2015 年まで長い間、the best-of-n といっても、2 択のケースしか研究されてこなかった。図 1.1 に示すように 2 次元での道路の方向を決めるなら、これらの 2 択の手法で十分だが、例えば、3 次元空間を自由に移動するドローン群の意思決定には新しい理論が必要と考えられている。

このような背景から近年、2 択以上のケースのための開発が盛んになっている。その主な手法は、value-sensitive 及び value-free という概念に則っている。Value-sensitive の概念では、選択肢の支持率が選択肢の良さに大きく依存する。一方、value-free は社会性に基づく方法である。Value-sensitive 概念に基づいた手法は、ミツバチ群が分蜂する際の意思決定メカニズムをモデル化したものである。ミツバチは次の移住先を探す際に、偵察蜂が自身の見積もりを他の蜂に報告し、受け取った蜂はその報告内容が過去の記録よりも高ければその候補に鞍替える。このプロセスを大多数が同じ候補を支持するまで繰り返す^[70]。このようなメカニズムによって多数の選択肢の扱いが可能となったが、value-sensitive の手法には次のような問題が生じる。

- (1) 選択肢の候補を多くすると、候補一つ当りの偵察蜂が少なくなり、候補の真の良さと食い違う誤った見積もりが群れ全体に広まる恐れがある。
- (2) 選択肢の支持率が候補の良さに大きく依存するため突出した候補がない場合などでは支持者数が拮抗したままの状態が続く (デッドロック問題) 原因になる。

この特徴のため、候補の淘汰が起こるような評価以外の変動源が不可欠である。これまでに、ノイズを加える方法、他の候補を抑制する方法など様々な内的ダイナミクスが導入されている。しかし、これらは、候補の本来の良さととのずれが生じさせるため、たとえ候補の数が 2 個の場合でも、7 割程度しか最適な選択肢を選べないケースがある事が報告されている^[63]。

このように、現在の主流の方法で選択肢の数 n を増やすことは難しいことはわかっているが、この 15 年これに変わる新しい考え方は提案されていないのが実情である。

1.2 研究の目的

そこで本研究, the best-of-n 問題を扱い, 個体の意思決定は選択肢の良さに依存しない value-free という概念に基づく手法を提案し, 多数の選択肢 ($n \gg 2$) が扱える群れの意思決定フレームワークの開発を試みる. Value-sensitive 概念と異なり, 社会的な投票モデルを提案する. これにより, 選択肢の良さに関係なくまず群れはいずれ候補の 1 つに合意する (合意というのは全員が同一の選択肢を選択する事である). よい選択肢を見つけるために, 合意した選択肢の良し悪しを各個で評価・判断する. もし, 良い選択肢なら選び続け, 良い選択肢ではないなら, 再び投票を行って, 別の選択肢に合意する. このプロセスを繰り返すことによってよい選択肢を発見できると考える. 以下ではこのモデルを BRT モデルと呼ぶ事にする.

1.3 本研究の構成

本論文の構成は次の通りである。

第1章では、本研究の背景と目的について述べる。

第2章では、従来研究を紹介すると共に本研究の位置付けを明確にする。

第3章では、多数の選択肢が扱える BRT (時間経過に増加する閾値, Bias and Rising Threshold) モデルを提案する。また、計算機実験により、提案手法の特徴を明らかにする。

第4章では、失敗を記憶する UEM (Unsuitable Experience Memory) エージェントを導入して、探索の効率化を試みる。

第5章では、提案した BRT モデルを通信範囲に限りがあるよりリアルな環境での可用性を検討し、通信範囲とパフォーマンスの関係性を明らかにする。

第6章では、群ロボットの応用力を高めるため集団行動のプリファレンスを実現する。

第7章では、解析のしやすい2次関数を用いたバイアス分布を使用する Q-BRT モデルを提案する。また、数学的な分析と計算機実験により、唯一の選択肢に合意する事や、適した選択肢の発見に要する平均時間の概算など Q-BRT モデルの特性を明らかにする。

第8章では、通信帯域に余裕があると仮定して、Plural Voting からヒントを得て、集団探索の高速化を試みる。

第9章では、集団の意思決定の精度を議論する際に分散 AI の分野でよく使われる標準問題の1つであるエルファロール・バー (El Farol Bar) 問題に Q-BRT モデルを用いて Q-BRT モデルの実装方法の例を示す。

第10章では、本論文の要約と今後の展望について述べる。

第2章 関連研究

本章では将来的な目標とするスワームロボティクスが具備すべき特徴を述べ、これに関連する研究について紹介する。

2.1 生き物がみせる群レベルでの高度な振る舞い

群ロボットの研究は、その程度の差はあるにせよ社会的昆虫、魚類、鳥類、細菌、動物といった自然界の生き物の群の解明(特に群れの中での相互作用の様式)をその出発点としている^[6]。群れは、少数の個体からなる小規模な群から、何百万の個体で構成されている複雑に組織化された共同体まで様々なものが知られている。群れは大まかにはローカルコミュニケーションと情報伝達を通して群れを形成していると考えられ、群れを形成することによって単独では解決が難しい課題を解決している。その課題の中には、経路計画^[88]、巣作り^[75]、仕事配分^[5]や巣探し^[3,76]などがある。自然には、異なる個体の能力や情報伝達手段などの内部と異なる生活環境や外敵などの外部によって様々な群れの形態が存在する。そこで、次に群れの例を示す。

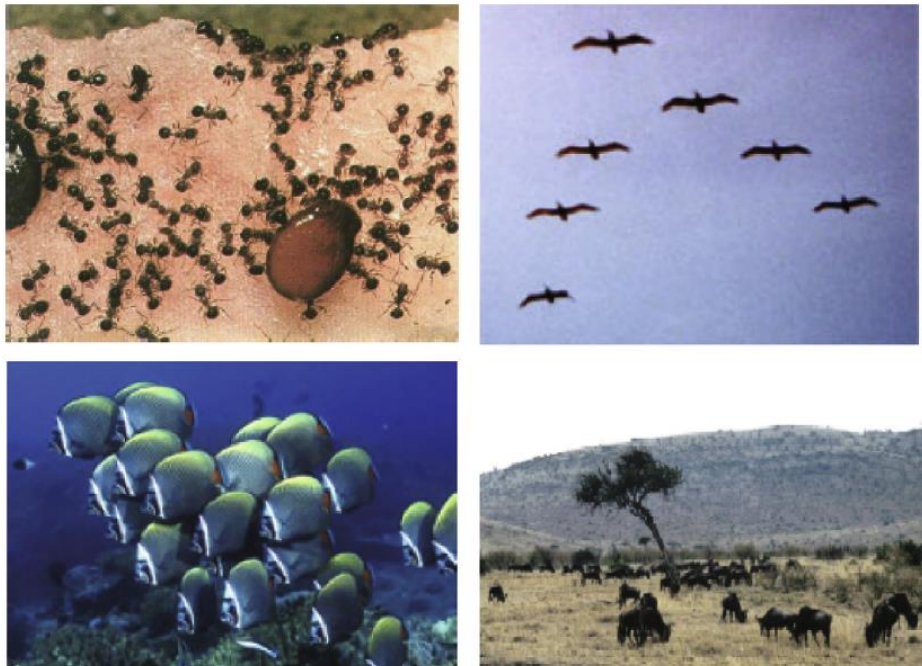


図2.1 自然の生物群^[93]。

- 細菌のコロニー:細菌は、多細胞凝集体を作り、分子レベルの信号を交換して細胞間通信をおこなっている^[73]。細胞分裂によって必要な種類の細胞を生成することによって、長期に渡る生存を可能にしている。異なる細胞タイプへの分化も行える。群を成す事によって例えばバイオフィームの中の細菌の抗菌剤に対する耐性は、同種の個々の細菌の耐性よりも500倍高い^[13]。
- 魚群:魚たちは、目や側線に加え、目と頭の両側に位置する目と肩の「スクールイングマーク」を使って、近隣の魚との関係を感じることができる^[8]。群を構成することによって、魚群は敵から攻撃を減らし^[59]、また捕食を容易にしている^[45]。魚の群れの特徴は耐分裂性と柔軟性だと考えられる。群れの個体数は数千以上に上ることも多く、大規模な群れでは外敵や餌を直接見ることができる個体は限られている。それにも拘わらず群れは高速に移動しても分裂しづらいように思える。
- アリ群とハチ群:アリは、フェロモン、音および物理的な接触を介して互いに通信する^[29]。餌の大小に適した規模のサブグループを形成でき、これによって餌を効率よく巣に運ぶことができる。また、葉きりアリは葉を巣に運ぶアリとこれを警護するアリの二つの役割を創発させて、外敵からの防御と餌の運搬という二つの課題を、群を形成することによって解決している。さらに、うまく餌を取得できたアリは、巣に戻る際にルートをフェロモンでマークする。より優れたルートにはより多くのアリがさらに通行するので、より良いルートが一層強化される。この過程によって最良の経路を群で共有することができる^[22]。またこれらの社会性昆虫では群が複数のカーストに分かれており、役割の分担が行われている。同じ遺伝子をもちながら、大頭アリでは頭が巨大化する働きアリと普通の働きアリに分化し、その形態と仕事の分担を状況に応じて適切に変化させることができる。
- イナゴ:イナゴは環境に応じて行動を大きく変化させることが知られている。これに対してBuhlら^[10]は、モデルを用いてグループ内の個体の密度が増加するにつれて、群が個人の無秩序な動きから高度に整列した集団運動へと変化するこの現象を明らかにしている。
- 人間:タスクや内容などに適した様々なグループを柔軟に作ることが知られている。生活に則したグループでは、家族というグループがよく知られている一方で、寮生活や軍隊などの職と生活が一致しているグループなど、他の生き物と比べて極めて多彩である。

以上から、群を高度だと言わしめている特徴には次のようなものが挙げられる。

- (1) 様々な状況で適切な協調や補完的行動が実現できる。
- (2) より複雑な環境や状況でも適応できる。
- (3) 限られた認知能力とコミュニケーション能力を持つ単純な個で構成された群れから、個体を超えた能力が創発する(高い防御能力, 高い集餌能力(大小サイズが異なる餌の運搬, 地理的に広域に分散する餌の収集)などがある)。
- (4) 環境への適応機能
 - (ア) 役割分担機能:例)餌の大きさに応じた群の形成, 役割の分担比率の調整(大型働きアリと小型働きアリの配分)。
 - (イ) 最適化機能:例)最適なルートへと収斂する機能。
- (5) 群で知識の共有や意志統一が行える。例)蟻の最適なルートに関する情報やミツバチの分蜂活動などが相当する。
- (6) 環境に応じて群全体の動き方を変えることができる(イナゴ)。
- (7) リーダがいなくてもタスク全体を小規模なものに分割してエージェントの作業を関連させて, それらを集団行動に集計するというメカニズム(例: 餌の運搬と警護)。
- (8) 異なる体格や認知能力をもった新個体の発生機能(カースト, 雌雄の比率)。
- (9) 様々なグループを作成することができる。

2.2 関連研究

本節では、本研究に関する関連研究を紹介する。本研究では、スワームロボティックスの分野で注目されている the best-of-n 問題のための手法として BRT モデルを提案する。まず、スワームロボティックスに関する研究を紹介し、群れでしか実現できない行動の重要性を示す。さらに、本研究で扱う the best-of-n 問題に関する関連研究を紹介し、それらの研究と差別化を行う。

2.2.1 スワームロボティックスに関する研究

スワームロボティックスは群知能の考え方を多数のロボット群に適用するものである。主に物理的で移動するエージェントを主体とする研究分野である。たとえばハーバード大の killobot^[42]が知られている。



図2.2 killobot (<http://itvscience.com/robot-swarm/>).

クラスタリングやソーティングは群れでしか成しえない行動の1つでこれまでも多く研究されている^[15,1]. Dario ら^[15]は群れ内の死骸を一か所にまとめる行動はアリのいくつかの種で見られる現象を扱っている。エージェントは近傍の死骸の数に反比例して死骸を運搬する。また近傍の死骸の数に比例して、運搬中の死骸を地面におろす。すなわち死骸のクラスタへ新たな死骸が追加される確率は、付近の死骸の密度によって決まる。実験では、1500 個の死骸が、事前の情報または外部制御なしに徐々に3つの集積所に集まった。Abdelhak ら^[1]はエージェントは、死骸の低密度のスポットから高密度のスポットまで輸送する単純で局所的なルールでの代用を検証した。実験では、群れがコミュニケーションなしで80 個のアイテムを収集するタスクを完了したことを報告している。



図 2.3 McLurkin ら^[39]の被覆問題の実験.

環境で均一に分散させることは被覆問題と呼ばれ、群れでしかできない重要な機能の1つである. McLurkin ら^[39]は, iRobotの群れのためのアルゴリズムを提案した. アルゴリズムは, 交互に実行される2つのステップに分割される. まずはロボットを分散させて, それから境界を検出する. このようにして, 群れは徐々に環境内で拡大することができる.

○ フォーメーション

フォーメーションとはロボットを一定の規則にのっとして配置することである. フォーメーションは自然界でよく見られる現象を理解する為の学術的アプローチである一方で, 様々なタスクをデザインする上での基礎である. 社会集団は, 個体の年齢, 形態, 栄養状態, 個性, リーダーシップの状態の違いにより, 多様な群れを形作る. 通常均一な個体からなる群れが議論されるが, 異質な個体からなるフォーメーションの形成原理についてはあまり研究されていない.

○ フロッキング戦略の基礎

1987年に Reynolds^[66]が提案した「Boids」モデルは, 個体間距離を用いたフロッキングの典型的なモデルである. このモデルは, 宇宙船, UAV, ロボットなどの様々なアプリケーションに広く採用されている. これらのアプリケーションでは, グループ行動をグループレベルで明示的に定義しておらず, 個体のルールとして表現しており, 生物関連の研究者からも高い評価を得ている^[4].

スワームロボティックスのフロッキングにおけるBoidsモデルの最も一般的な応用は, 仮想力の形態である. Hettiarachchi ら^[27]は, 相互作用によって仮想的に生成される物理的な力を使ってロボットの行動を制御するという「Physicomimetics」フレームワークを紹介した.

以上から、スワームロボティクス分野の研究において、自然の生き物の群れから参考してきた特徴を以下に示す。

- ① 扱える選択肢の数：必要なだけ、異なる選択肢を検討できる。
- ② フォールトトレラント：集団として適してない選択肢を選択した時により良いものに切り替える方が望ましい。
- ③ リーダーなし：リーダーがいなくても動作する。
- ④ 動的な環境：刻々と変化する環境下でも有効である。
- ⑤ 最適性：適した選択肢を集団で発見できる。
- ⑥ 高速な探索：適した選択肢の発見に要する時間は短いほうがいい。もし、良いものが見つければ、残りの候補を検討する必要はない。
- ⑦ 群れサイズ恒常性：群を分断することなくその群状態を維持できる。
- ⑧ 群れサイズのスケールビリティ：群れサイズが増加しても問題なく使用できる。

これらの特徴を全て実現することは困難であるが、より優れた群ロボットシステムを構築する為には議論すべき要件である。

2.2.2 The best-of-n 問題に関する研究

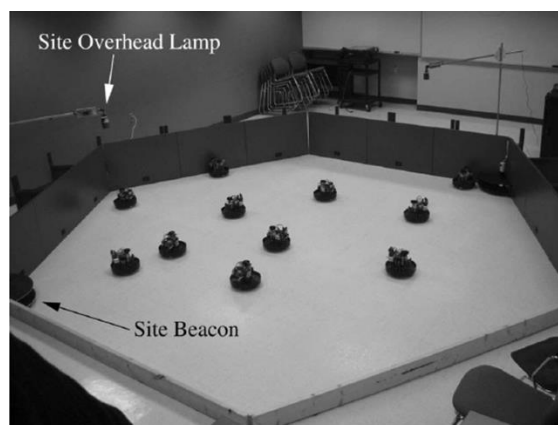
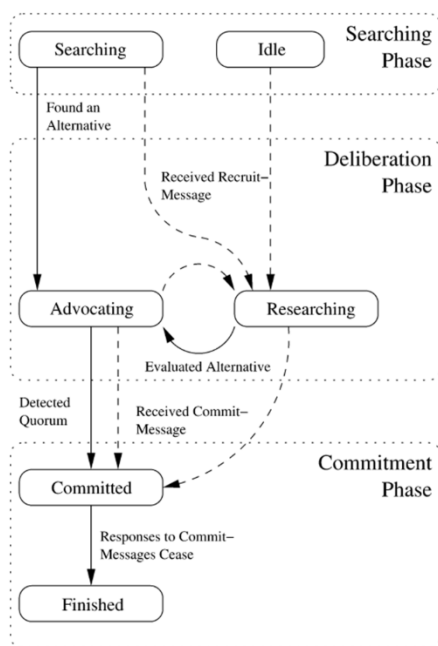
次に、本研究で扱う the best-of-n 問題に関する研究を紹介する。まず、改めてこの問題について説明する。

The best-of-n 問題とは「集合的意思決定問題」の一つのケースに相当する。一般に集合的意思決定問題は、全員が同じ状態に達する「合意問題」と、異なる状態で安定する「タスク割当問題」にわかれる^[84]。合意問題は「連続値」に合意する場合と、「離散値」に合意する場合がある。前者の代表的な問題としては進行方向の合意やランデブー地点の合意などで、アルゴリズムには Boids^[66]やコンセンサスアルゴリズム^[47,48]が知られている。ここで扱うのは後者のケースで、これが一般に the best-of-n 問題と呼ばれている。

現在提案されている the best-of-n 問題の為の手法は主に、value-sensitive 及び value-free という概念に則っている。Value-sensitive 概念では、選択肢の支持率が選択肢の良さに大きく依存する。一方、value-free は社会性に基づく方法である。以下では、この2の概念に基づく手法を紹介する。

- Value-sensitive 概念に基づく手法：

Parker ら^[54]は、2つの巣のうち良いものを選ぶ問題に対してクォーラム決定法を提案している。ロボットは3つのフェーズ、6つの状態をもつ。当初、ロボットは探索フェーズにあり、環境を探索する探索状態か、または休憩状態で待機している。巣が見つかって、その巣の良さを査定すると、ロボットは審議フェーズに移行し、査定した巣の代表状態になる。審議フェーズでは、代表状態のロボットが自身の巣の良さと比例した頻度で募集メッセージを送信する。この募集メッセージを受けたロボットは、再探索状態になりその巣の良さを査定しに行く。時間が経過すると良さの高い巣に属する代表状態のロボットが増加する傾向がある事を示している。各ロボットは受け取った最近の一定のメッセージをみて、同じ意見の個数を調べて、クォーラム値に達したらコミットメント状態になる。一旦、コミットメントフェーズに移行すると審議フェーズに戻れないため、コミットした巣を変更することができない。そのため、審議フェーズでの選択精度が求められ、クォーラムは群れの半数以上に設定する必要があるがそれでも偶然による判断ミスが避けられない。



(a)

(b)

図 2.4 Parker ら^[54]の集団的意思決定アルゴリズム(a)と移動ロボットによる実験(b).

Montes de Oca ら^[44]は 2 択の経路選択問題に対して、Krapivsky ら^[32]の多

数決ルールに不応期間を導入した Lambiotte ら^[35]の合意形成手法の並列性について検討している。初期時にメンバーの意見をランダムに設定する。そして、ランダムに3台のロボットからなるサブグループ(チーム)をk組つくり、多数決ルールを用いてメンバーの意見(経路)を統一する(図 2.5)。意見が揃ったチームは、選んだ経路を往復する。この間はメンバーが意見を変更することができなく、不応期間と呼ぶ。スタート位置に戻ったエージェントは、再び他のものとチームを組み、意思決定を行う。このプロセスを繰り返すと、最も短い経路が選択される傾向がある事を示している。この研究では不応期間の発生確率分布が指数分布と正規分布にした場合の収束性について2択問題を用いて解析的な検討を行っている。シミュレーションにより、Krapivsky ら^[32]の方法よりもこの手法は最適性や群れサイズのスケーラビリティが改善されることが示されているものの収束に要する時間は群れのサイズに応じて増加する。

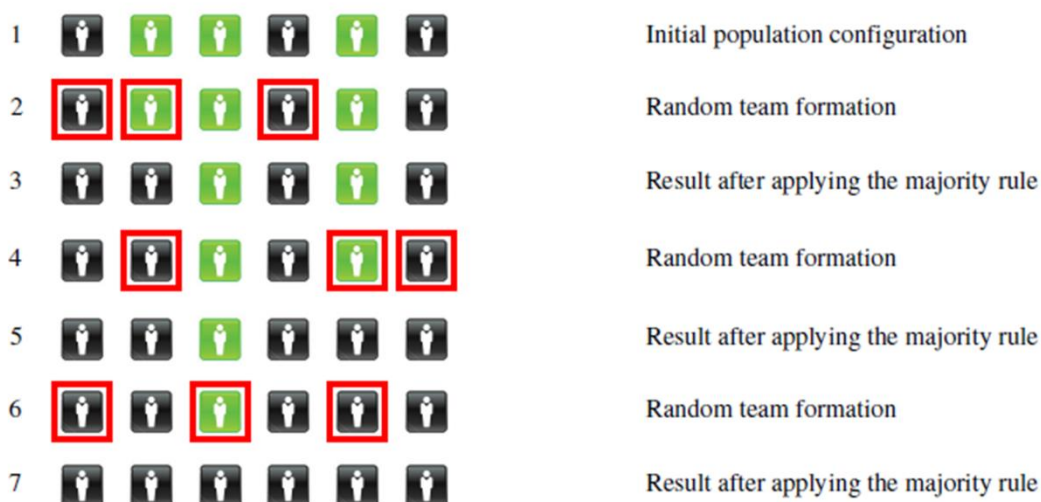


図 2.5 Montes de Oca ら^[44]の集団的意思決定のダイナミクスの例。

Valentini ら^[85]は重み付けされた投票モデルを提案し、群れが2カ所の候補地 A, B から移住先を選ぶ問題を扱っている。各エージェントは候補地ごとに調査状態とリクルート状態があり、移住先を指数分布にしたがうランダム時間調査した結果を元にベストな候補を他のエージェントに通知する。よい候補ほど長く周知され、各エージェントは受信メッセージからランダムに次調査地点を選ぶ。このポジティブフィードバックによって、全体が合意できる場合があることが、マルチエージェントシミュレーションと ODE による分析が示されている。その結果、 $n=2$ のときには、台数が増えると良い候補への

同期確率が向上する一方で、合意に達するまでの時間はたとえ二者の差に大きな差があっても伸びることが示されている。また、候補の評価の差が小さいと精度も悪くなることが示されている。

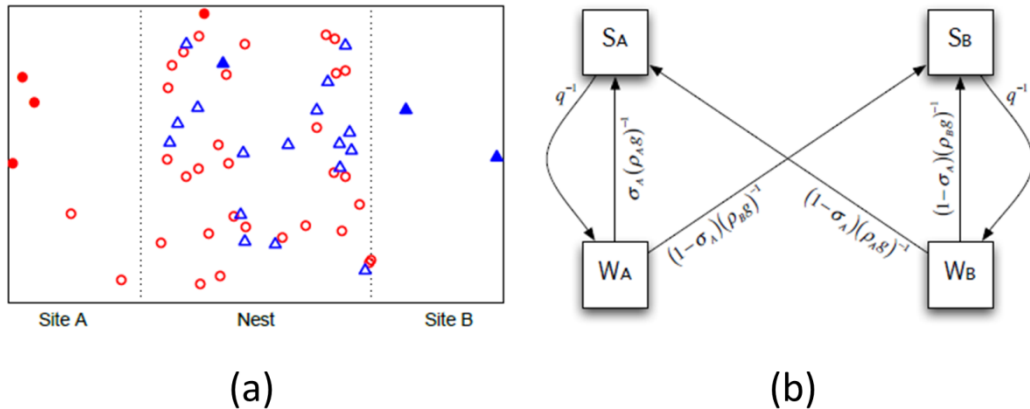


図 2.6 Valentini ら^[85]のエージェントベースのシミュレーションのスクリーンショット(a)と確率的状態遷移(b).

Seeley ら^[70,71]は、ミツバチ群が分蜂する際の行動を観察し、偵察ハチが巣に戻って自身の候補を宣伝する際に他の候補を宣伝する偵察ハチに停止信号を送る事を発見した。この交差抑止のメカニズムをモデル化し、良さが同等の2つの巣を選択する際に起こるデッドロック問題を回避できる事を明らかにしている。Pais ら^[51]も、2つの巣を選択するミツバチの群れの集団的意思決定メカニズムのダイナミクスを分析し、意思決定能力を向上させるための交差抑止の役割を明らかにした。その上で、Seeley ら^[70]の交差抑止にノイズを導入し、群れのパフォーマンスを改善させた。



図 2.7 ミツバチの交差抑止のメカニズム^[51,65,70,71].

Reina ら^[65]は、最短経路を探索する問題($n=2$)に対し、ミツバチ群で観察さ

れた巣の選択行動及び脊椎動物の脳的意思決定行動^[38,70]に基づいた集団的意思決定モデルを提案している。図 2.7 に示すように、ロボットには、どの選択肢も支持しない U 状態、選択肢 A もしくは B を支持するコミット状態という 3 つの状態がある。コミット状態のロボットは、U 状態のものを募集することができる。また、自身と異なる選択肢の支持を抑止し、U 状態に戻す事ができる（交差抑止）。さらに、自身のコミット状態を放棄し、U 状態になることもある。提案手法は、システムの微視的パラメーターを群れの巨視的ダイナミクスにリンクする決定論的および確率論的な数学的モデルの両方によってサポートされている。

Reina ら^[63]は、2 択しか扱わない Pais ら^[51]の意思決定モデルを複数の選択肢が扱えるように拡張を行った。選択肢の良さが同等な対称ケースにおける集団的意思決定のダイナミクスを数学的な解析により明らかにした。また、複数の選択肢の中に 1 つが突出している非対称ケースにおいては、適切にパラメータを設定すれば、選択肢の数 $n \leq 7$ なら、群れは少なくとも約 75% の精度で最も良い巣に収束できる事を明らかにした。

Hasegawa ら^[24]は、複数の選択肢を扱う the best-of- n 問題に対し、簡単な yes/no 決定エージェントによる集団意思決定法を提案している。集団の個体は、ある閾値を持ち、全ての選択肢の良さと比較する。自身の閾値が選択肢の良さより小さいと “yes” と答える。そして、“yes” と答えたエージェント数を観測して、これが一番多い選択肢を選択する。その結果、集団の閾値分布の標準偏差を大きく設定する事により、選択肢の数が 20 でも意思決定の精度が 8 割以上である。

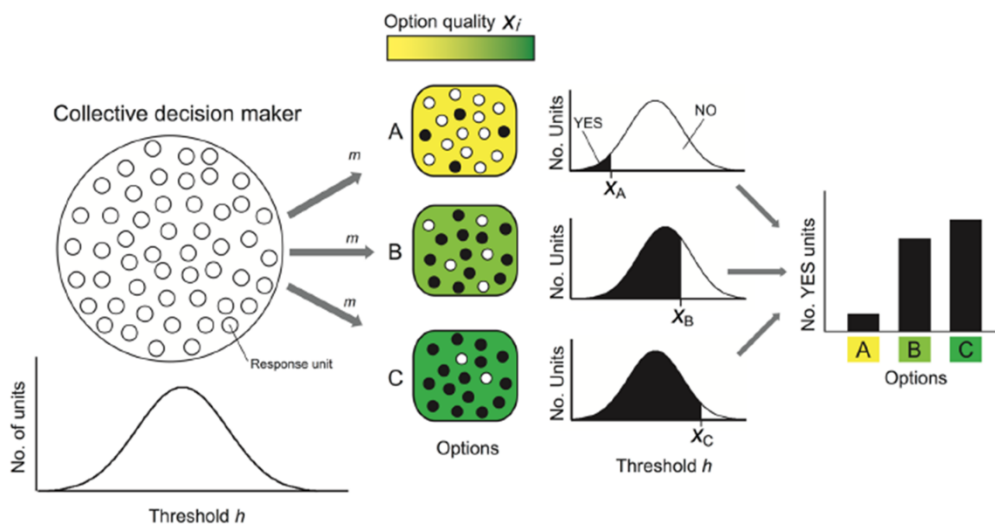


図 2.8 Hasegawa ら^[24]の選択肢の良さを査定するスキーム。

- Value-free 概念に基づく手法：

Iwanaga ら^[28]はある意見に賛成するか反対するかの 2 択問題について、自身が賛成することと、多数派になることと、の 2 つの利得と、個人差を導入して、賛成、反対の両グループのダイナミクスを議論した。その結果、個人差の分布特性が正規分布等の一山型である場合、人数によらない速さで、初期状態に応じて全員が賛成もしくは反対の状態になることを解析的に明らかにしている。ただし、2 択以上の場合については言及していない。この研究については次章でも詳しく説明する。

Wessnitzer ら^[91]は移動する餌を群で追うロボット群のシミュレーションと実機での検証をおこなっている。二匹の移動する餌の内(2 択問題)、一方を全ロボットで追跡する。各ロボットはどちらの餌を追うかを意思表示し、近傍のロボット群で多数決を取ることによって群として追う餌について合意する。全員で一方の餌を捕獲した後、残りの餌に向かう。この手法は節 2.2.1 の要件⑥(群れサイズ恒常性)を満たしているが、2 択のみを扱っていることと、エージェント数に対するスケーラビリティに課題がある。

これらの研究を要件について比較したものを表 2.1 に示す。◎、○と△のマークは、それぞれ節 2.2.1 の要件を満たす具合を表している。この表から、the best-of-n 問題は 2000 年頃に提案されたものの 2015 年まで長い間、the best-of-n といっても、2 択のケースしか研究されてこなかったことがわかる。また、近年、Reina ら^[65]の研究がきっかけで 2 択以上のための開発が盛んになっていることがわかる。

表2.1 関連研究

Requirement	Value-sensitive approaches						Value-free approaches								
	Parker et al (2009)	Montes et al (2011)	Valenti et al. (2014)	Seeley et al. (2012)	Pais et al. (2013)	Reina et al. (2015)	Reina et al. (2017)	Hasegawa et al. (2017)	Iwanaga et al. (2002)	Wessnitzer et al. (2003)	Kubo et al. (2015)	Phung et al. (2017)	Kubo et al. (2017)	Scheider et al. (2016)	The proposed method
1: No. of alternatives (n)	2	2	2	2	2	≥ 2	≥ 2	≥ 2	2	2	2	≥ 2	≥ 2	2	≥ 2
2: Failure recovery	-	-	-	-	-	-	Δ	Δ	-	-	o	\otimes	o	-	\otimes
3: No leader	Δ	o	o	o	o	o	Δ	Δ	o	o	o	o	o	o	o
4: Dynamic environment	-	o	o	-	-	-	-	-	o	o	o	o	o	-	o
5: Optimality	Δ	Δ	o	o	o	o	o	o	\otimes	o	o	\otimes	o	o	\otimes
6: Time limitation	o	o	o	-	o	o	-	-	o	o	o	o	o	o	o
7: Homeostatic	o	o	Δ	Δ	Δ	Δ	\otimes	\otimes	\otimes	\otimes	o	\otimes	o	Δ	\otimes
8: Scalability of N	Δ	Δ	o	o	o	o	o	o	\otimes	o	Δ	\otimes	o	o	\otimes

2.3 まとめ

アリやハチなどは群れをなしている。一匹の能力が限りがあるが、群れとなった大きな集団は、様々な知的な振る舞いを見せている。ロボットにおいても、ロボットが多数集まり群れとして動作する群ロボットシステムは、群れとしての特性と個としての特性を活用して、臨機応変にタスク処理を遂行することが可能となり、人間の管理が難しい環境での活躍が期待される。

現在、群ロボットの分野では、群れの集団行動の意思決定の研究として the best-of-n 問題が注目されている。現在提案されている the best-of-n 問題の為の手法は主に value-sensitive 及び value-free という概念に則っている。Value-sensitive 手法では、選択肢の支持率が選択肢の良さに大きく依存するため、評価のバラツキや突出した候補がない場合などでは合意に到りにくく、結果として多数の選択肢を扱うのが困難である。その一方、value-free 手法では、3 択以上の問題についての理論がなく、群れとしての探索機能が不要な 2 択問題しか扱う事ができない。

第3章 BRTモデルの提案

3.1 はじめに

本章では、多数の選択肢が扱える群れの意思決定フレームワークを構築する。

まず、第2.2.1節で指摘した要件①： $n \geq 2$ でも選択を可能にするため、以下のコンセプトでの手法の開発を検討した。各選択肢の評価が出きった後に最高値の行動を選択する方法では、 n が大きくなるにつれ決定までに要する時間（反応時間）が長くなるので、要件⑥を満たすことが難しい。そこで、選択肢を1つずつ順に評価して、良いとわかれば行動選択を切り上げる方式を採用する。切り上げ方には様々な工夫が施せるが^[78]、ここでは単純にするために、選択肢ごとの善し悪しのみで決めるものとする。

ある選択肢の集団的行動としての善し悪しを評価する際、ミツバチが新しい巣場所の善し悪しを確認する時のように個体ごとに異なる選択肢の評価を試みる方式がある。これは、要件⑦を難しくするばかりか、集団として行動した際の評価とギャップが生じやすくなる。この問題を回避するには実際にその行動を集団で実行して評価することが考えられる。これはWessnitzerら^[91]の方法に似ている。ただし、これを実施するには、集団が評価する選択肢について速やかに合意できることが望ましい。試すべき行動についてすみやかに合意できれば、許される時間内により多くの行動を試すことができるので、実質的に要件①を満たすことができる。逆に合意までに長い時間がかかれば、多くの候補を評価することができない。加えて要件⑧を満たすために、群れのサイズや選択肢の数 n が多くても素早く合意できる必要がある。

以下では、以上を満たすBRTモデル(時間経過とともに増加する閾値, Bias and Rising Threshold)を提案する。BRTモデルでは、すべてのエージェントは各自多数の選択肢の中から1つを選びこれを全員に通知する。この際、エージェントは各選択肢の支持者数を鑑みて自身の選択肢を変更する。全員の選択肢が一致することを合意状態と呼び、これを合意するまで行うものである。これはIwanagaら^[28]の2択の社会モデルを多数の選択肢の場合でも利用できるように拡張すると共に、探索機能を実現する為に時間経過を組み入れたモデルである。

次に、Iwanagaら^[28]のモデルについて詳しく説明する。

3.2 Iwanaga らのモデル

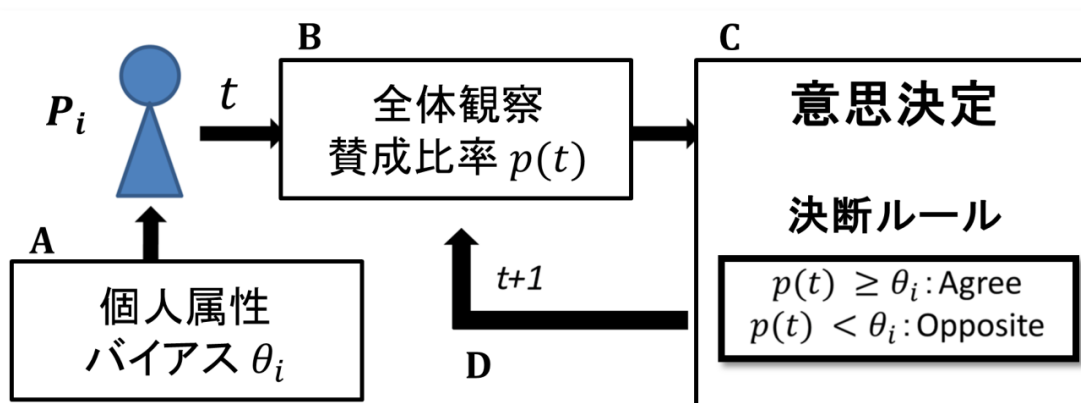


図 3.1 個人の意思決定における社会的皮膚の役割^[28].

Iwanaga ら^[28]は Schelling^[69]のクリティカルマスの研究にヒントを得て、個人の意思決定は自分の好みなどを表す個人属性だけでなく集団全体の雰囲気 ($p(t)$) からも影響を受けると考えた。図 3.1 にその仕組みを示す。ある選択肢に対し、各個人は賛成かもしくは反対の立場をとるものとする。集団全体の雰囲気というのは、自身の選択肢に対する賛成者の比率である。これに対して各個人は個人属性バイアス値 θ_i を持つ。個体は、図 3.1B に示すように、全体を見渡し賛成者の比率 $p(t)$ を観測する。次にこれを用いて個体は、以下のルールに基づいて意思決定を行う (図 3.1C)。

$$\begin{cases} p(t) \geq \theta_i: & \text{Agree} \\ p(t) < \theta_i: & \text{Opposite} \end{cases} \quad (3.1)$$

式(3.1)に示すように、時刻 t でバイアス値 θ_i よりも賛成率 $p(t)$ が高ければ個体 P_i は賛成する。全員意思決定を行った後、時刻 $t+1$ の $p(t+1)$ は変化する (図 3.1D)、再び個人は判断をおこなう (図 3.1C)。

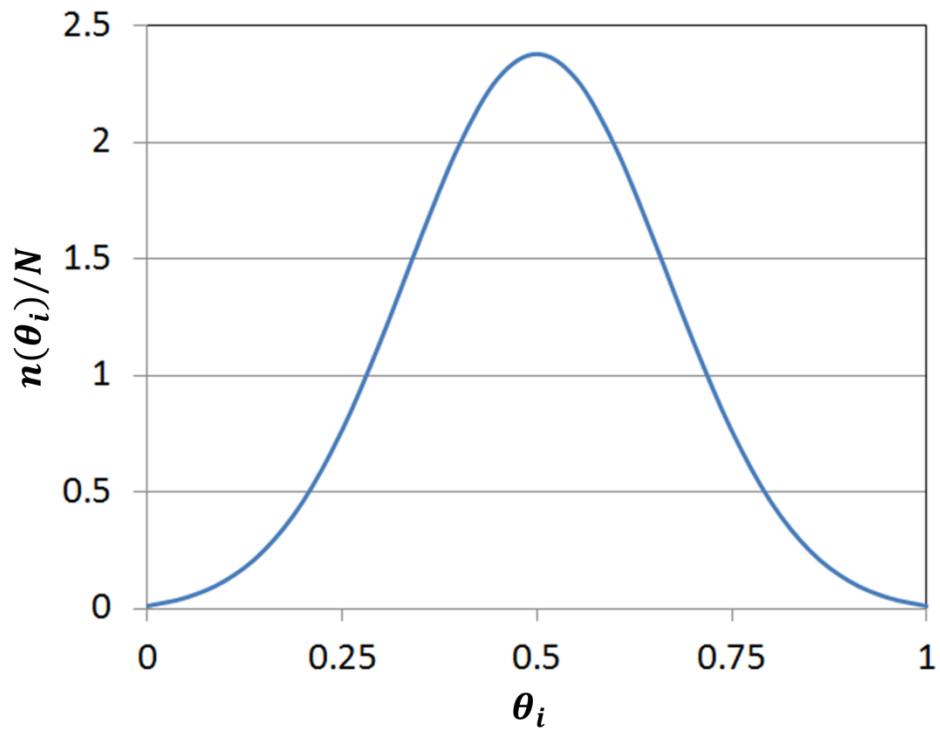


図 3.2 バイアス分布^[28].

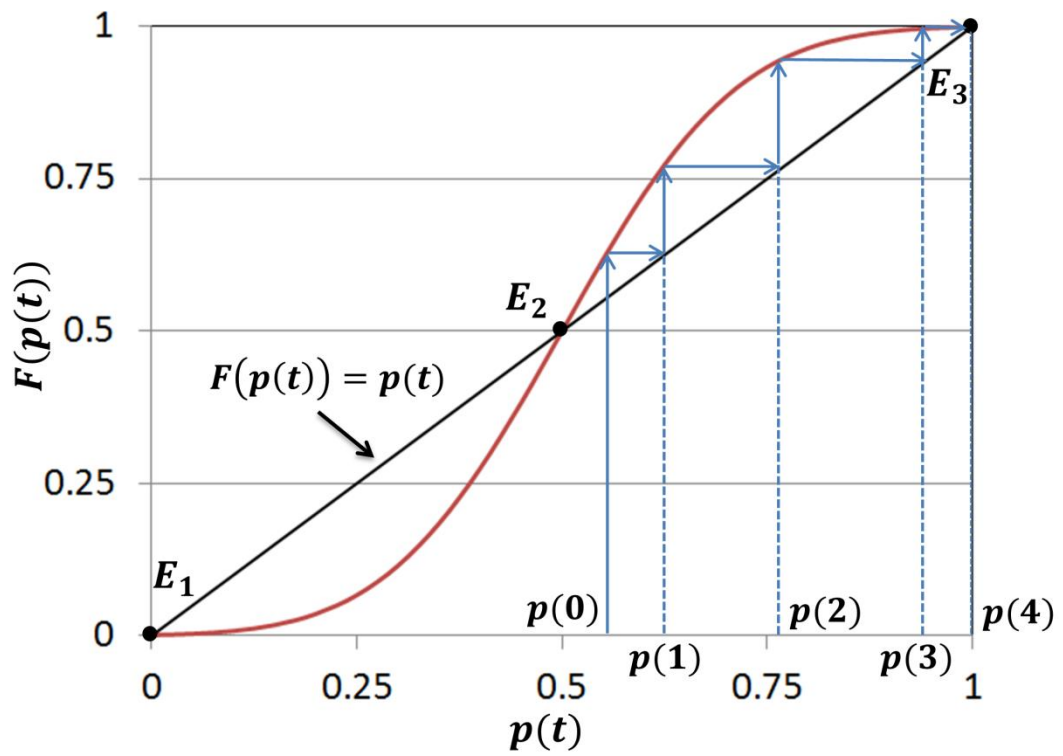


図 3.3 バイアスの累積確率分布^[28].

Iwanaga らは、このように意思決定を定めた時、賛成比率 $p(t)$ のダイナミクスを解析する方法を提案している。図 3.2 のように θ_i の分布を定めると、その累積確率密度分布 $F(\theta)$

$$F(\theta) = \sum_{\theta_i \leq \theta} n(\theta_i) \quad (3.2)$$

は、図 3.3 のように S 字型のカーブを形成する。式(3.1)に従ってエージェントが判断すると、時刻 t での $p(t)$ は次の式に従う。

$$p(t+1) = F(p(t)) \quad (3.3)$$

式(3.3)より

$$p^* = F(p^*) \quad (3.4)$$

となる賛成率 p^* は変化しない事から均衡点と呼ばれる。均衡点には安定点と不安定点がある。図 3.2 と図 3.3 のようにバイアス値 θ_i の分布が釣鐘型の場合、不安定点が 1 つ (E_2) と 2 つの安定点 (E_1, E_3) が発生し、全員が賛成もしくは反対の状態に、滑らかに推移する。すなわち、集団の挙動は以下の 3 つのパターンに分類できる。

- $p(t) = 0.5$ の場合である。図 3.2 では E_2 に相当する。 $F(p(t)) = p(t)$ になるため、変化しない。但し、なんらかの理由によって $p(t)$ が変化すると以下の 2 つのうちのいずれかの状態に移行する。
- $p(t) > 0.5$ の場合である ($E_2 < p(t) \leq 1$)。この区間では $p(t+1) = F(p(t)) > p(t)$ であるため、全員が賛成する状態 ($p(t) = 1$) に速やかに推移する。図 3.3 では $p(0) < p(1) < p(2) < p(3) \simeq 1$ でほぼ均衡状態から約 3 単位時間で全員が賛成する状態に到達している。
- $p(t) < 0.5$ の場合である ($0 \leq p(t) < E_2$)。 $p(t+1) = F(p(t)) < p(t)$ であるため、全員が反対する状態 ($p(t) = 0$) に速やかに推移する。

収束するまでのイテレーション回数は θ_i の分布形状が同じであれば人数には依存しないため、多数でも速やかな収束が期待できる。しかしながら、3 択以上の問題 ($n \geq 3$) についての理論がなく、Iwanaga らのモデルでは群れとしての探索機能が不要な 2 択問題しか扱う事が出来ない。

3.3 提案手法

以上を踏まえて、以下では、Iwanaga らのモデルを多数の選択肢が扱える the best-of- n ($n \gg 2$) 問題に拡張し、新しい群れの意思決定法を提案する。混乱を避けるために、以下では選択肢の数を n ではなく M とし、集団の個体数を N とする。

3.3.1 BRT アルゴリズム

ここで、 N 台のエージェントで構成される集団 $G = \{P_i: i = 1, \dots, N\}$ において、 M 個の選択肢の中から適した選択肢を 1 つ選ぶ意思決定を行うとする。エージェントの行動集合を $A = \{a_j: j = 1, \dots, M\}$ ($M \geq 2$) とし、探すべき適した選択肢を a_{goal} とする ($a_{goal} \in A$)。エージェントは事前に a_{goal} を知らず、全員が合意して初めて合意した選択肢が a_{goal} であるか分かるものとする。各個体は、集団全体での動向を考慮しながら（あるいは周りの圧力を受けながら）自らが選んだ選択肢を選び続ける、あるいはやめて別の選択肢を選ぶかの意思表示をする。すなわち、自らの選好に基づく意見を通そうとする意思と、そのことで周りから孤立することを避けたいとする 2 つの相対峙した心理的な働きによる葛藤を抱えながら、総合的に判断することになる。

また、集団全体の動向が変わったことを察知することで、一度決心したことを後になって変更することもありうる。今エージェント P_i の時刻 t での選択肢を $A_i(t) \in A$ とする。また、エージェント P_i は個人属性バイアス値 θ_i ($0 < \theta_i < 1$) を持つ。 $n(a_j)$ を選択肢 a_j を選択しているエージェントの数とする。全員が同一の選択肢を選択した時、集団がその選択肢に合意したとする。これはたとえばある餌場に集団全体が移動することに相当する。

時刻 t において、エージェント P_i は集団全体で自身と同じ選択肢を選んでいる割合を $n(A_i(t))/N$ 、次の時刻に選ぶ選択肢 $A_i(t+1)$ を次のように決定する。もし

$$\frac{n(A_i(t))}{N} \geq \theta_i + \tau \cdot c_i(t) \cdot (t - t_{i,last}(t)) \quad (3.5)$$

であれば、現在の選択肢 $A_i(t)$ を引き続き選択し、 $A_i(t+1) = A_i(t)$ となる。そうでなければ現在の選択肢 $A_i(t)$ をやめてそれ以外の選択肢を確率的に選ぶ ($A_i(t+1) \in A - A_i(t)$)。

τ は支持者数の増加量を表す定数である。 $t_{i,last}(t)$ はエージェント P_i が最後に選択肢を変更した時刻で、 $(t - t_{i,last}(t))$ で同一選択肢を選び続けている時間に相当し、以下の式で表せる。

$$t_{i,last}(t) = \begin{cases} t & A_i(t) \neq A_i(t-1) \\ t_{i,last}(t-1) & otherwise \end{cases} \quad (3.6)$$

$c(t)$ は選択肢が“適している”時 0 である評価関数である. すなわち

$$c_i(t) = c(t) = \begin{cases} 0 & \forall i, A_i(t) \equiv a_{goal} \\ 1 & otherwise \end{cases} \quad (3.7)$$

提案するアルゴリズムは Iwanaga らモデルに時間経過に増加する閾値(BRT 機構)を導入するものである. 以下では, バイアス値とこの閾値を BRT 機構と呼ぶ. バイアス値 θ_i を釣鐘型に分布すれば, 速やかな合意が実現できるだろう. そして, $c(t)$ は, 自分の状況を見て良いか悪いかを判断する. 適した選択肢ならば $c(t)$ が 0 なので適した合意行動が続く. そうでない場合は, 次第に, BRT 機構が大きくなり, 式(3.5)が満たされなくなって, エージェントは現在の選択肢を諦めて別のものに切り替える. これによって, 速やかに合意するという Iwanaga らモデルの特徴に加えて, 選択肢の良し悪しの判断, 合意行動の切り替えを実現することが期待できる. そして, 合意と合意の切り替えを繰り返すことにより, 適した選択肢を発見することが期待できる.

3.3.2 バイアス分布

ここでは, ガウシアン分布に基づくバイアス値 θ_i の設定方法を提案する.

$$\theta_i = \min(1, \max(0, G(\frac{1}{M}, \sigma_g^2))) \quad (3.8)$$

ここで, $G(\frac{1}{M}, \sigma_g^2)$ は平均 $\frac{1}{M}$, 分散 σ_g^2 の正規分布に基づく乱数で

$$G(\frac{1}{M}, \sigma_g^2) = \frac{1}{\sqrt{2\pi\sigma_g^2}} \exp(-\frac{(x - \frac{1}{M})^2}{2\sigma_g^2}) \quad (3.9)$$

である.

3.4 実験

3.4.1 BRT モデルの挙動例

ここでは、BRT モデルの挙動を図示し、その特性を説明する。

エージェント数 $N = 1000$, $M = 3$, $\sigma_g^2 = \frac{1}{3M}$, $\mu = \frac{1}{M}$ とし、初期時でエージェントはランダムに選択肢を選択しているとした。

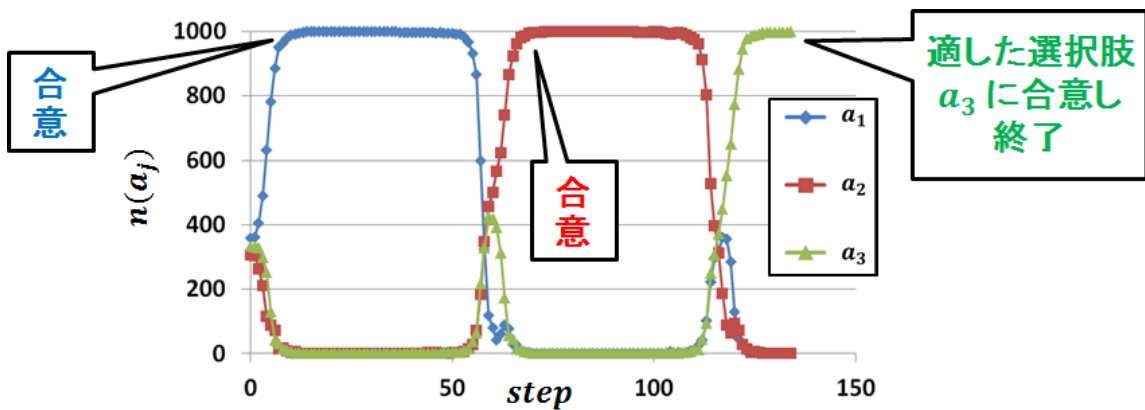


図 3.4 各選択肢の支持者数の遷移例(ランダムスタート).

$$N = 1000, M = 3, \sigma_g^2 = \frac{1}{3M}, \mu = \frac{1}{M}, a_{goal} = a_3, \tau = 0.01.$$

図 3.4 に群れの選択肢ごとの支持者数の遷移を示している。縦軸は各選択肢を選択しているエージェントの割合で、横軸は経過時間である。全エージェントがある同一選択肢を選択するとき、この状態を合意状態と呼ぶ。初期時において、選択肢 a_1 を選択しているエージェントは最も高い支持者数約 36%であった。時間経過により、 a_1 の支持者数がどんどん増加して行く。そして、約 10 ステップ後に、全員で選択肢 a_1 選んで合意状態になっている事がわかる。この時、式(3.5)の左側が 1 になっている。しかし、 a_1 は適した選択肢ではないので式(3.5)の右側の第 2 項は常に増加する。時間が経過すると式(3.5)が満たされなくなり a_1 をやめて別の選択肢を選択するエージェントがどんどん出てくる。これによって、群れの状態は合意状態から、非合意状態に戻り、再び意思決定を行う。このプロセスがもう一回繰り返された。2 回目の合意後約 130 ステップで全員が適した選択肢 a_3 に合意し、終了した。

以上から、3 つの選択肢の中から、群れは速やかに任意の選択肢に合意し、その合意を切り替えて、適した選択肢を発見できる事がわかる。

3.4.2 適した選択肢の発見

ここでは、集団は多数選択肢の中から適した選択肢を発見できるまでに要する時間を計算機実験によって示す。

$N = 100$, $\tau = 0.003$, $\sigma_g^2 = \frac{1}{3M}$, $\mu = \frac{1}{M}$ とした。初期時においてエージェントはランダムに選択肢を選択しているとし、選択肢の数 M を 2, 5, 10, 15, 20 と変えて適した選択肢を発見することができるか検証を行った。予め選択肢候補の 1 つ a_x をランダムに選び、これを適した選択肢とした ($a_{goal} = a_x$)。 a_{goal} はエージェントに伏せておく。これを全員が選択することができたケースをカウントし、それぞれの M に対し 10000 回を実施し、その発見確率を求めた。制限時間を設け、これに達したならば試行を打ち切った。制限時間として 3000 と 30000 ステップの両ケースで実施した。図 3.5 に実験結果を示す。

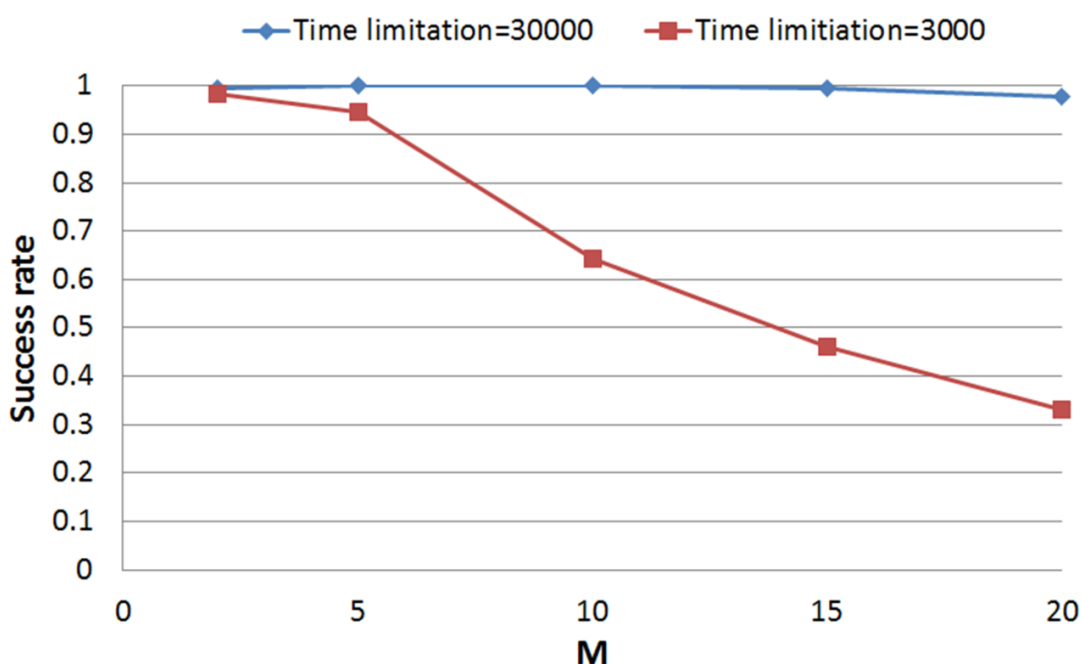


図 3.5 制限時間までの発見率。

$$N = 100, \tau = 0.003, \sigma_g^2 = \frac{1}{3M}, \mu = \frac{1}{M}.$$

縦軸は、制限時間内に適した選択 a_{goal} を発見できた比率で、横軸は選択肢の数 M を表している。□マークと◇マークの折れ線はそれぞれ制限時間 3000 と 30000 ステップのケースを表している。制限時間が 3000 ステップでの発見できた比率は、 M が増加するごとに発見率が低下している事から、 M が増加するに従

って発見に要する時間が延びる事がわかった。一方、制限時間を 30000 ステップにした場合では、どの M においても 100%適した選択枝が発見できた。この事から十分な時間さえあれば提案手法は多数の選択枝の中から適したものを発見することができる。

3.4.3 エージェント数と合意時間の関係

本節では、計算機実験を使用し、エージェント数 N と合意時間の関係を明らかにする。

$M = 6$, $\tau = 0.001$, $\sigma_g^2 = \frac{1}{3M}$, $\mu = \frac{1}{M}$ とした。初期時において集団が合意状態からスタートするとし、エージェント数 N を 10, 40, 160, 640, 2560 と変えて合意状態からスタートして別の状態に合意するまでの時間(合意時間)を計測した。図 3.6 に 10000 試行を実施した結果を示す。

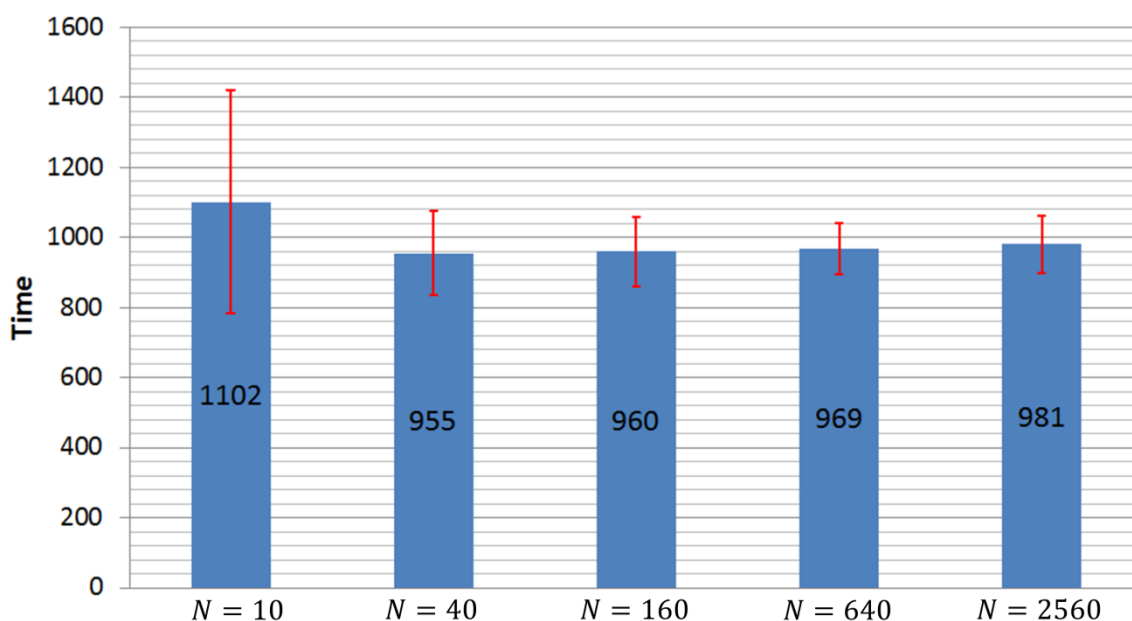


図 3.6 合意状態からスタートして別の状態に合意するまでの平均時間とその

標準偏差. $M = 6$, $\tau = 0.001$, $\sigma_g^2 = \frac{1}{3M}$, $\mu = \frac{1}{M}$.

その結果、エージェント数がとても小さい($N = 10$)ケースを除いて、エージェント数の増加は合意形成に要する時間に影響を与えないことが分かった。標準偏差をみても、エージェント数がとても小さい時($N = 10$)のみ合意に要する時間が

ばらつくが、エージェント数がある程度大きいとこのバラつきが急激に収まることが分かった. 図 3.7 は $N = 10, 160, 2560$ のある試行で用いたバイアス値を累積確率分布で表したものである. この図をみると, $N = 160$ 及び $N = 2560$ のバイアス値の累積確率分布は, ほぼ図 3.3 に示す提案累積確率分布の形になっていることがわかる. 一方, $N = 10$ の場合は, エージェント数が小さいためその累積確率分布は滑らかでない. この荒さが合意に要する時間をばらつかせる原因だと考えられる.

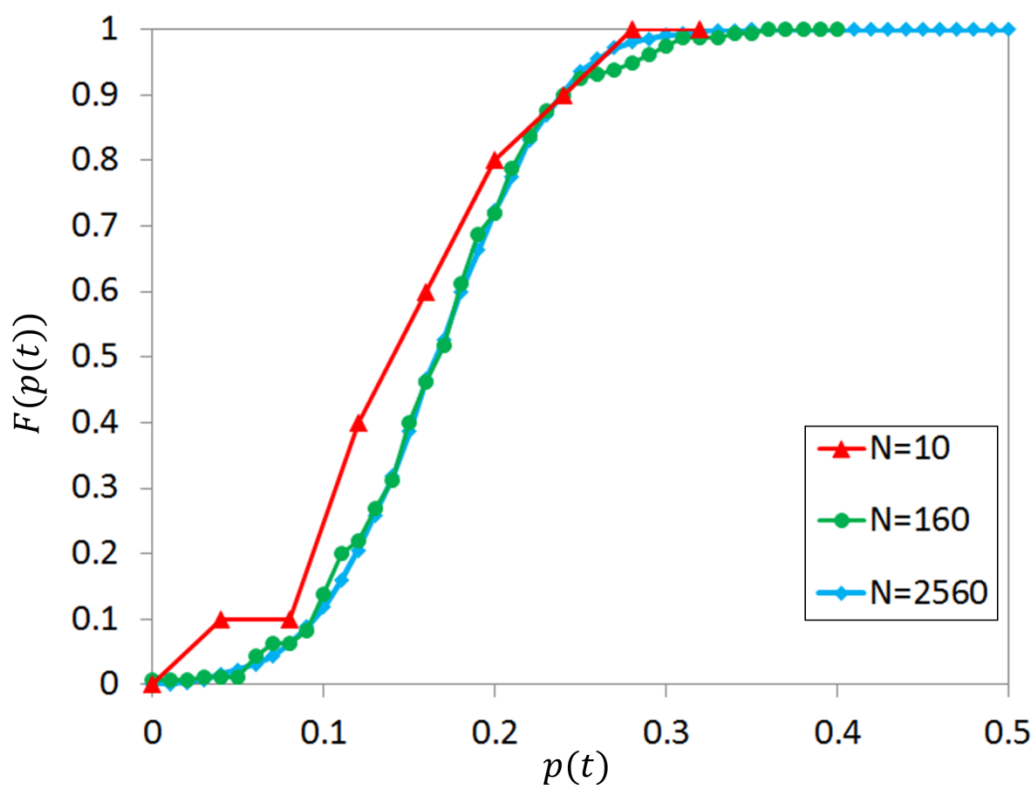


図 3.7 バイアス値の累積確率分布.

$$M = 6, \tau = 0.001, \sigma_g^2 = \frac{1}{3M}, \mu = \frac{1}{M}.$$

以上から, 合意時間はエージェント数 N には依存しない事がわかった. よって, 本モデルは, 巨大数からなる群れロボットや群れのサイズが変化するなどの場合にも利用できる.

3.5 まとめ

本章では, 多数の選択肢が扱える BRT モデルを提案した. Iwanaga らのモデルに時間経過に増加する閾値を導入することにより, 群れは速やかに任意の選択肢に合意し, その合意行動を切り替えて, 適した合意行動を発見できる事を計算機実験により示した. また, BRT モデルでは合意時間はエージェント数 N には依存しないため, 巨大数からなる群れロボットや群れのサイズが変化するなどの場合にも使用できる事を計算機実験によって明らかにした.

第4章 BRTモデルその①：

タブーリストを持つエージェントによる効率的な集団探索

4.1 はじめに

第3章では、多数の選択肢の中から適したものを発見する the best of-n 問題のために、集団レベルで試行錯誤法を使用した BRT モデルを提案した。BRT モデルでは、エージェントは合意したものの適さなかったらその選択肢を記憶して、次時刻ではそれ以外の行動を選択する。この短期記憶は直近の選択肢に限られるため、図 4.1 に示すように少し前に適さなかった選択肢に繰り返し合意してしまうことがある。

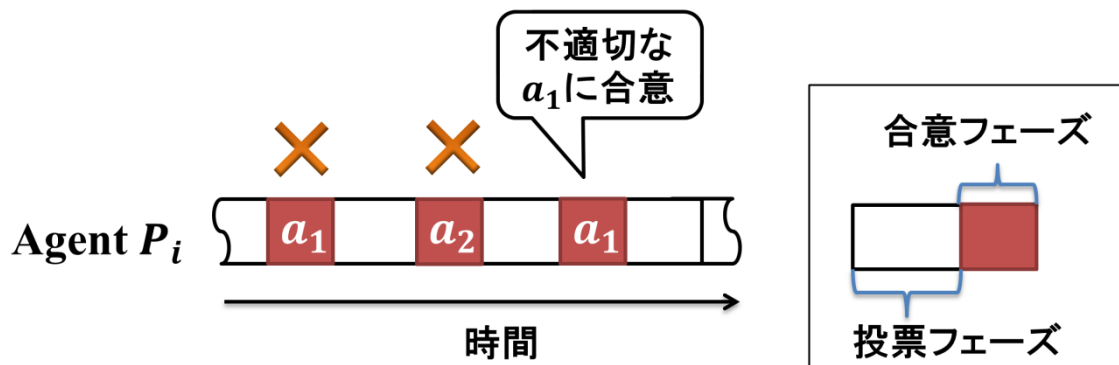


図 4.1 BRT モデルで起こる時間のかかる例.

前回実行時から環境が変化していない静的な環境ではこれは有効的な探索とはいえない。そこで、合意したものの適さなかった選択肢を記憶するエージェントを導入して、探索の効率化を試みる。

4.2 提案手法

エージェントは、合意したものの適さなかった行動を記憶するタブーリスト $L_i(t)$ をもち、合意しながらも不適切と判断出来た時点で確率 p_{in} でこのリストに追加する.

$$L_i(t+1) = \begin{cases} L_i(t) \cup A_i(t) & \text{probability } p_{in} \\ L_i(t) & \text{otherwise} \end{cases} \quad (4.1)$$

$$A_i(t) = \{\forall j, A_i(t) = A_j(t), A_i(t) \neq a_{goal}\} \quad (4.2)$$

新たな選択肢を選択する際には、 $L_i(t)$ に含まない選択肢を確率的に選ぶ. すなわち

$$A_i(t+1) \in A - L_i(t) \quad (4.3)$$

式(4.1)で記憶して、式(4.3)で選択肢の候補を選ぶ BRT エージェントを提案し、以下では UEM (Unsuitable Experience Memory) エージェントと呼ぶ.

4.3 提案手法の特性の概算

表 4.1 発見に要する時間と発見率

回数	候補数	発見率	時間
1 回目	M	$\frac{1}{M}$	$D + C$
2 回目	$M - 1$	$(1 - \frac{1}{M}) \frac{1}{M - 1}$	$2(D + C)$
3 回目	$M - 2$	$(1 - \frac{1}{M})(1 - \frac{1}{M - 1}) \frac{1}{M - 2}$	$3(D + C)$

もし、群れのすべてのエージェントが、 $p_{in} = 1$ で記憶する UEM エージェントの場合、適した選択肢を発見するために要する平均時間は行動候補数 M の線形になると考えられる. 表 4.1 に適した選択肢を発見する過程とそれに要する時間を示す. 今、 M 個の候補の一つが適した行動集合 a_{goal} の唯一の要素とする. 一度も合意していないとき、候補数は M のままである. この中の 1 つに合意したとき、それが a_{goal} である確率は $\frac{1}{M}$ である. C は合意に至るまでに要した時間、 D は合意した選択肢が適切なものか判定するのに要する時間とする. 1 回目に合意した行動が a_{goal} でない場合、2 回目の合意を試みるが、候補となる行動の数は UEM

エージェントではその記憶によって $M - 1$ である。したがって、2 回目に合意した行動が a_{goal} である確率は $(1 - \frac{1}{M}) \frac{1}{M-1}$ である。これより、発見に要する平均時間は

$$\sum_{i=0}^{M-1} (i+1) \frac{1}{M-i} \prod_{j=1}^i (1 - \frac{1}{M+1-j}) = (D+C) \frac{M+1}{2} \quad (4.4)$$

であり、これより UEM エージェントによって探索時間大きく改善できる可能性がある事がわかる。

表 4.2 ある 1 つの未選択肢を支持する確率

回数	BRT エージェント	UEM エージェント
1 回目	$\frac{1}{M}$	$\frac{1}{M}$
2 回目	$\frac{1}{M-1}$	$\frac{1}{M-1}$
3 回目	$\frac{1}{M-1}$	$\frac{1}{M-2}$
x 回目	$\frac{1}{M-1}$	$\frac{1}{M-(x-1)}$

次に、UEM エージェントの比率について検討する。表 4.2 に合意を試みた回数と、任意の未選択肢を選ぶ確率を示す。初めて合意を試みる時には、ある 1 つの選択肢を選ぶ確率は BRT, UEM エージェントと共に $\frac{1}{M}$ である。2 回目では、1 つの選択肢が不適切であったことを両エージェント共に知っているため、共に $\frac{1}{M-1}$ である。しかし、3 回目以降では両エージェントの未選択肢を選ぶ確率は異なってくる。BRT エージェントは $\frac{1}{M-1}$ のままであるが、UEM エージェントではすでに選ばれて不適切だと分かった選択肢は選ばないため選択肢の候補が減り、 $x (\leq M)$ 回目の選択確率は $\frac{1}{M-(x-1)}$ である。探索時間を短縮するにはすでに試した選択肢に合意する回数を減らす必要があるが、これには不適切だと分かった選択肢と未だ試していない選択肢を選ぶエージェント数差を大きくすることが

肝要で、これは両者の期待値の差に相当する。 x 回目の両者の期待値の差は UEM エージェントの個体数を $u (< N)$ とすると

$$\frac{u}{M - (x - 1)} \quad (4.5)$$

である。この式は、UEM エージェントの個体数が多ければ、既に試した選択肢を選ぶことは少なくなる事を示している。一方で、合意を試みる回数が進むにつれ、分母が小さくなるため、すでに選んだ選択肢に合意する事が少なくなる事を表している。従って、UEM エージェントの個体数が少なくとも探索時間の短縮効果が期待できる。

4.4 実験

本節では, UEM エージェントの挙動を示し, 予想された特性を持っている事を確認する. はじめに, UEM エージェントの集団の挙動を示し, M に対して線形の時間で適した選択肢を発見できる事を示す. 次に, UEM エージェントの比率を変えた実験によって, 少ない UEM エージェントでも探索時間の短縮が見込める事を示す.

4.4.1 UEM エージェント集団の特性

ここでは記憶する確率 $p_{in} = 1$ の UEM エージェント 1000 台からなる群れに対し, 選択肢の数 M を 3 から 15 まで変化させた. それぞれについて 10000 試行を実施した. $\tau = 0.005$, $\sigma_g^2 = \frac{1}{6M}$, $\mu = \frac{1}{M}$ とした. 図 4.2 はそのうち $M = 3, 6, 10$ の時の探索時間を累積確率密度分布で表したものである.

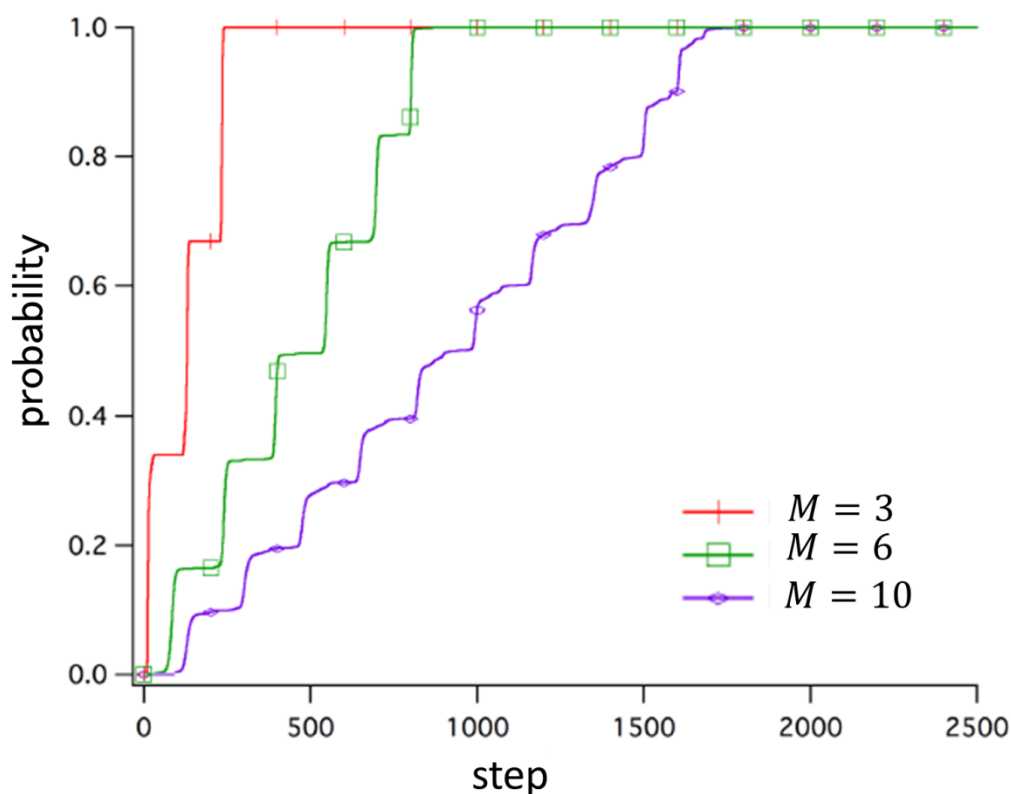


図 4.2 UEM エージェントの集団による探索時間の累積確率密度分布. $\tau = 0.005$,

$$\sigma_g^2 = \frac{1}{6M}, \mu = \frac{1}{M}.$$

横軸に要した時間，縦軸にこの時間までに適した選択枝を発見できた確率を示す．3本の折線いずれも，大きく波打っており，その回数は M 回であった．各凹みは合意・再探索過程に相当している．例えば， $M = 3$ のグラフでは 35 ステップ付近まで，140 ステップ付近まで，245 ステップ付近までの 3 個の凹みがあり，それぞれが合意を試みた回数に相当する．これより 1 回目に合意する比率は 34% (35 ステップ)，2 回目までに合意する確率 68% (140 ステップ) であることがわかった．

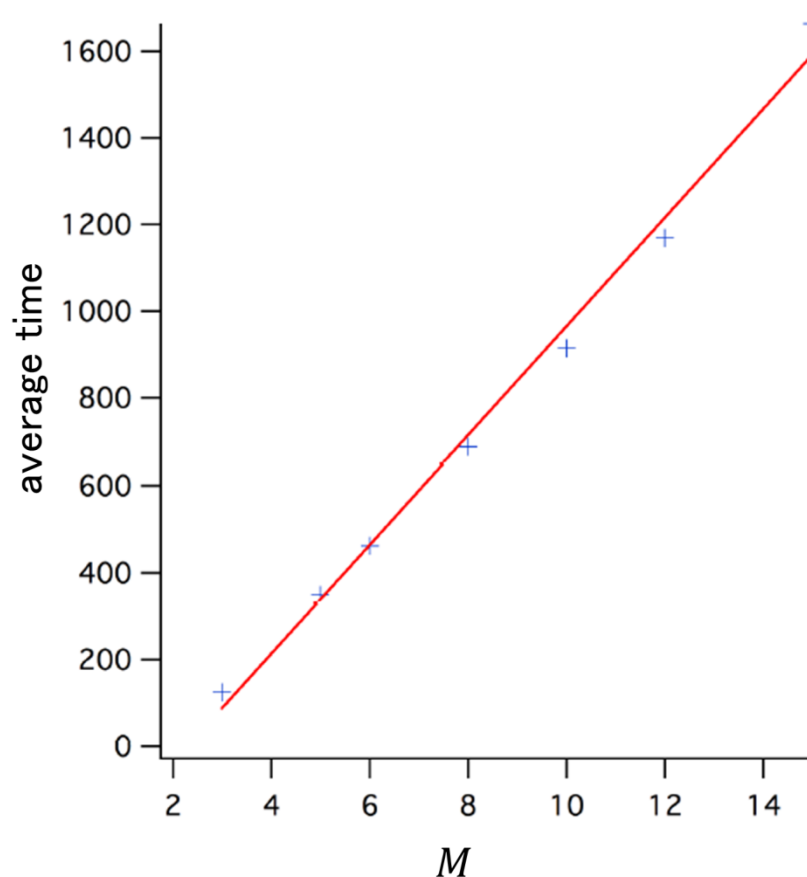


図 4.3 UEM エージェントの集団の平均探索時間と M の関係．

$$\tau = 0.005, \sigma_g^2 = \frac{1}{6M}, \mu = \frac{1}{M}.$$

また， M が大きくなるに連れ，発見に要する時間がかかることがわかる． $M = 3$ で 100%発見するには 610 ステップ必要で， $M = 6$ では 804 ステップ必要で， $M = 10$ では 1720 ステップ必要だった．図 4.3 は平均探索時間と M との関係を +マークで示したものである． M が増加するにしたがって，平均探索時間はほぼ線形に増加しており，式(4.4)での予想と合致している．

4.4.2 UEM エージェントの比率

次に、UEM エージェントの比率を変えて探索時間を計測した。 $N = 1000$ とし、 $p_{in} = 1$ の UEM エージェントの比率 p_{UEM} を 0.0 から 1.0 まで変えて探索時間を計測した。BRT, UEM エージェントのパラメータは $M = 6$, $\tau = 0.005$, $\sigma_g^2 = \frac{1}{6M}$, $\mu = \frac{1}{M}$ とした。

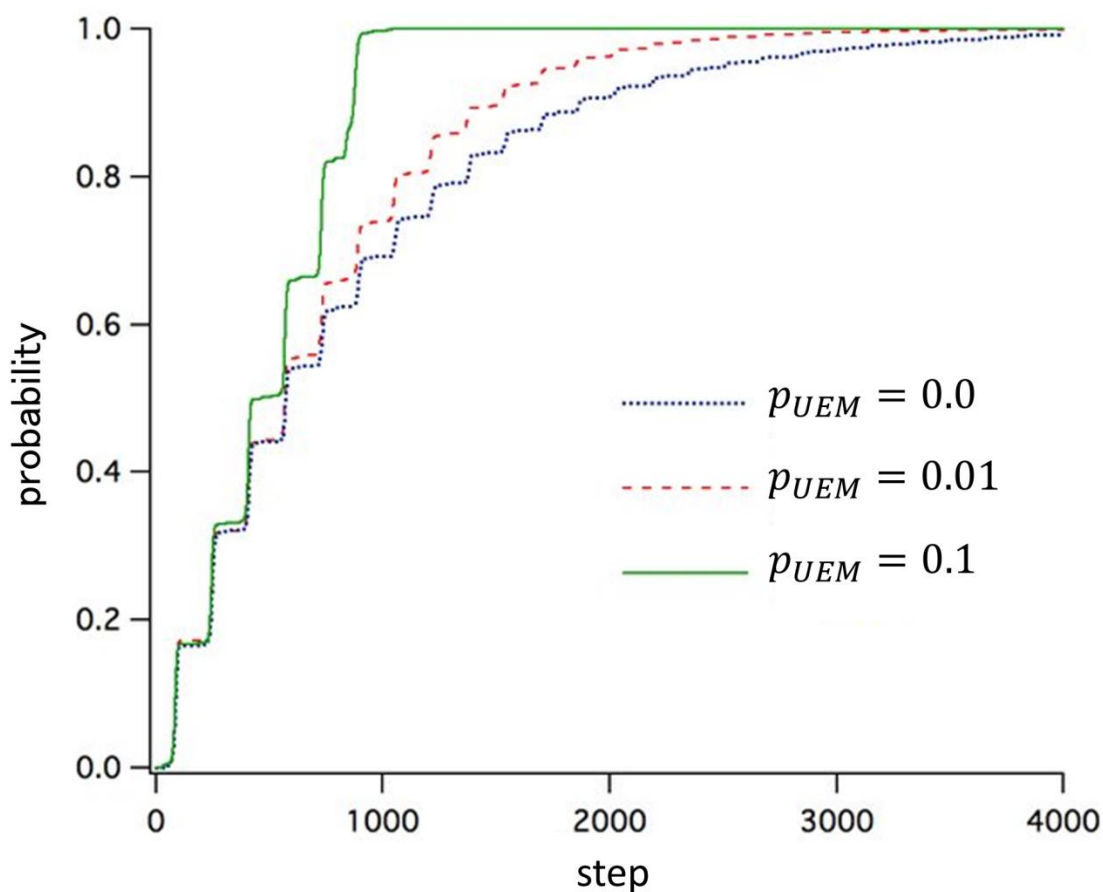


図 4.4 UEM エージェントの集団の探索時間.

$$M = 6, \tau = 0.005, \sigma_g^2 = \frac{1}{6M}, \mu = \frac{1}{M}.$$

図 4.4 は探索時間を累積確率密度分布で表したものである。横軸に要した時間、縦軸にこの時間までに適した選択枝を発見する確率を表している。点線で $p_{UEM} = 0$ の試行を表している。このグラフは頻繁に波打っており、BRT エージェントだけの群れでは、合意に至るまで多数回の合意・変更が必要なのがわか

る. また, 3000 ステップでも 98%に達しておらず, 長時間要する試行も少なくな
い事がわかる. 破線で $p_{UEM} = 0.01$ のケースを示している. このケースでは波打
つ回数は 9 回と $M (= 6)$ よりも多いものの, 3000 ステップまでに 99%以上の
試行で適した選択枝の発見に至っており, 探索時間の改善が見られる. 実線で
 $p_{UEM} = 0.1$ のケースを表している. 大きく波打つ回数は 6 回 ($= M$) で全員が
UEM エージェントの群れの場合と等しい. また探索時間も 1000 ステップまでに
99%の試行で適した選択枝を発見できている. 式(4.5)の概算どおり, 少数の UEM
エージェントでも大きく探索時間が改善できている.

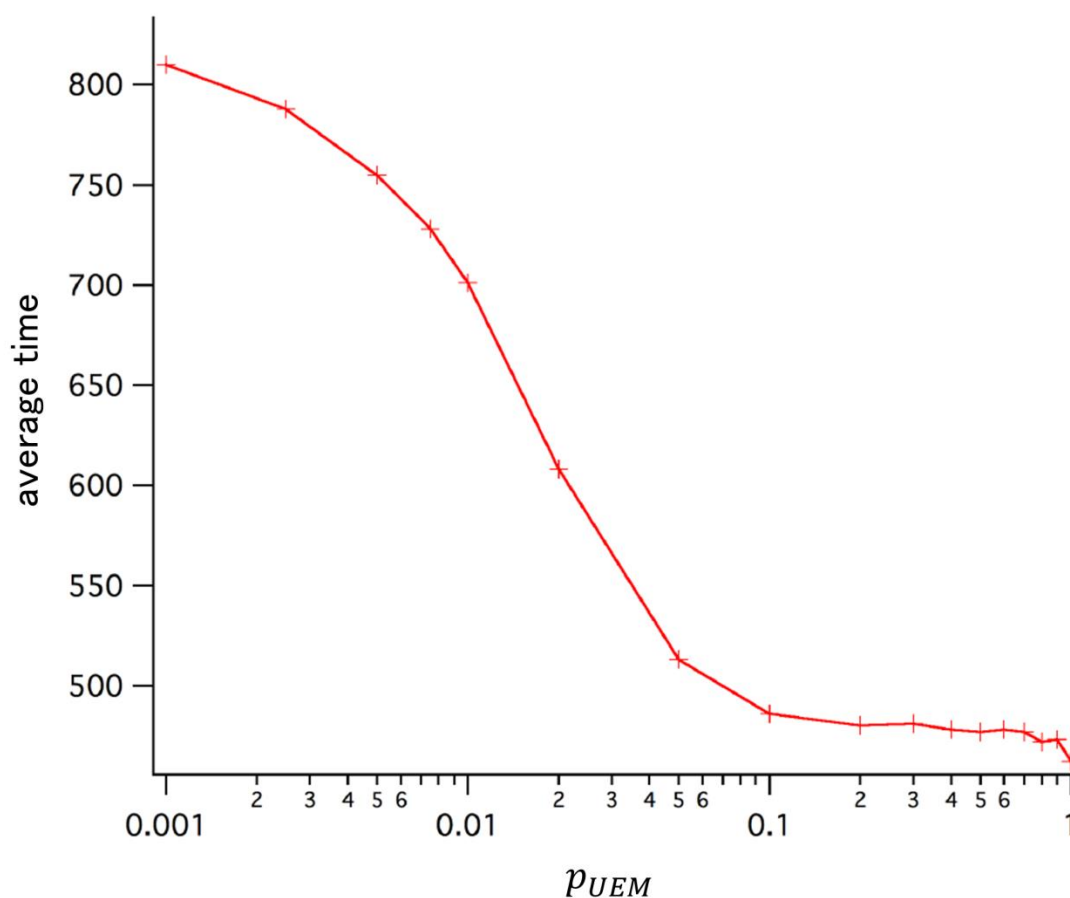


図 4.5 UEM エージェントの比率と探索時間の関係.

$$M = 6, \tau = 0.005, \sigma_g^2 = \frac{1}{6M}, \mu = \frac{1}{M}.$$

図 4.5 に UEM エージェントの比率と平均探索時間の関係を示す. 横軸に UEM
エージェントの比率, 縦軸に平均探索時間を表している. $p_{UEM} = 0$ では平均 834
ステップ要し, $p_{UEM} = 1.0$ では平均 462 ステップ要した. $p_{UEM} = 0.001$ では 810
ステップであり, 微小数 UEM エージェントが混じるだけでも探索時間は改善でき

た. UEM エージェントの比率を高めてゆくと, $p_{UEM} = 0.1$ まで, 劇的な改善がみられた($p_{UEM} = 0.1$ では平均 486 ステップ要した).

この 2 つの実験から, 提案した UEM エージェントは BRT エージェントの探索時間を大きく改善できる事がわかった.

4.5 まとめ

本章では，BRT エージェント群が創発する試行錯誤的探索挙動の効率を改善するために，UEM エージェントと呼ぶ，失敗した経験を記録するタブーリストを持つ BRT エージェントを提案した．群れが UEM エージェントから構成される場合には探索時間が選択肢の数 M に対して線形に増加すること，BRT エージェントと混在する場合には UEM エージェントが少数でも劇的な効果があることを概算し，計算機実験でその傾向を検証した．その結果，期待どおりの挙動を示すことがわかった．

第5章 BRT モデルその②：観測範囲が制限されている場合

5.1 はじめに

BRT モデルでは，群れの個体の意思決定は，個人属性を表すバイアス θ と集団全体の雰囲気・動向から影響を受ける(図 5. 1A)．しかし，現在のロボット群に搭載されているセンサーの性能や実際の作業環境の状況などによって，図 5. 1B に示すように個体が見渡せる範囲が限られている場合がある．従って，ロボット群で利用する為には，このような場合においても BRT モデルが有効である事を示す必要がある．

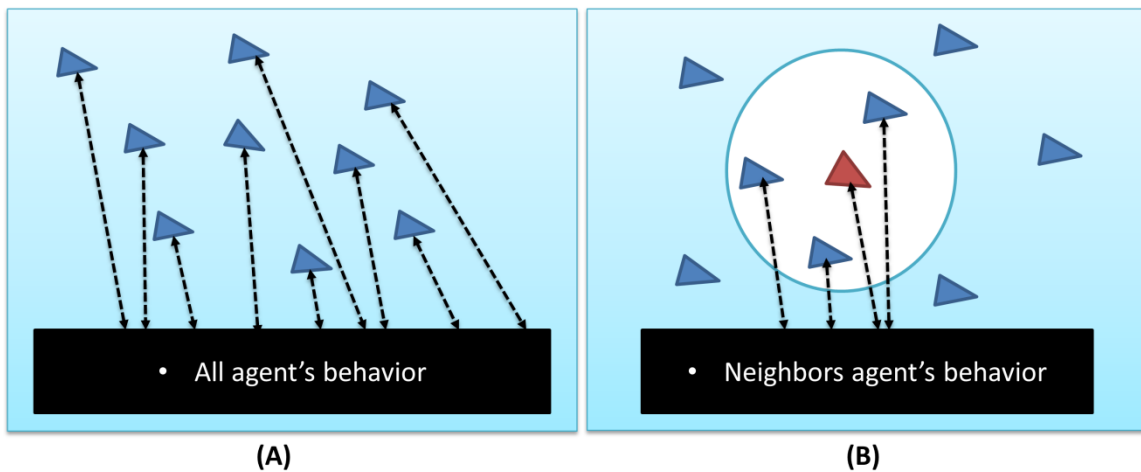


図5.1 (A) 集団全体の情報が共有する場合. (B) ローカル情報共有の場合.

そこで，本章ではローカル情報共有に基づく BRT-LIS (Local Information Sharing)モデルを提案し，エージェントの観測範囲が制限されても適した選択肢が発見できる条件を簡単に調査する．

5.2 提案手法：BRT-LIS モデル

式(3.5)に示すように BRT モデルでは、個体は集団全体を観て自身と同じ選択肢を選んで割合を $n(A_i(t))/N$ を観測して意思決定を行う。以下では、ローカル情報共有に基づく BRT-LIS モデルを考える。

BRT モデルと同様、 N 台のエージェント $P_1, \dots, P_i, \dots, P_N$ ($1 \leq i \leq N$) で構成される集団があるとする。加えて、エージェントの選択肢集合を M 個の $A = \{a_1, \dots, a_j, \dots, a_M\}$ ($M \geq 2$) とする。探すべき適した選択肢を a_{goal} とする ($a_{goal} \in A$)。エージェントは事前に a_{goal} を知らず、全員が合意して初めて合意した選択肢が a_{goal} であるか分かるものとする。エージェント P_i の時刻 t での選択肢を $A_i(t) \in A$ とする。また、エージェント P_i は個人属性バイアス θ_i ($0 < \theta_i < 1$) を持つ。

しかし、BRT モデルと異なり、BRT-LIS モデルでは、個体は半径 r 内にある近傍の情報のみを取得する。時刻 t において $N_i(t)$ を近傍エージェントの数とする。 $n_i(A_i(t))$ を $A_i(t)$ を選択しているエージェント P_i の近傍エージェントの数、全員が同一の選択肢を選択した時、集団がその選択肢に合意したとする。

時刻 t において、エージェント P_i は次の時刻に選ぶ選択肢 $A_i(t+1)$ を次のように決定する。もし

$$\frac{n_i(A_i(t))}{N_i(t)} \geq \theta_i + \tau \cdot c_i(t) \cdot (t - t_{i,last}(t)) \quad (5.1)$$

であれば、現在の選択肢 $A_i(t)$ を引き続き選択し、 $A_i(t+1) = A_i(t)$ となる。そうでなければ、現在の選択肢 $A_i(t)$ をやめてそれ以外の選択肢を確率的に選ぶ ($A_i(t+1) = A - A_i(t)$)。

BRT モデルと同様、 τ は支持者数の増加量を表す定数である。 $t_{i,last}(t)$ はエージェント P_i が最後に選択肢を変更した時刻で、 $(t - t_{i,last}(t))$ で同一選択肢を選び続けている時間に相当し、以下の式で求める。

$$t_{i,last}(t) = \begin{cases} t & A_i(t) \neq A_i(t-1) \\ t_{i,last}(t-1) & otherwise \end{cases} \quad (5.2)$$

$c_i(t)$ は選択肢が“適している”時 0 である関数である。すなわち

$$c_i(t) = \begin{cases} 0 & \forall i, A_i(t) \equiv a_{goal} \\ 1 & otherwise \end{cases} \quad (5.3)$$

である。

5.3 実験

第3章では計算機実験によってBRTモデルではほぼ集団のサイズに依存せず速やかに合意でき、また合意した選択肢が切り替えられる事で適した選択肢が発見できる事が明らかにした。ここでは、簡単なシミュレーションを用いてBRT-LISモデルの特徴と適した選択肢の発見ができるかを明らかにする。

5.3.1 実験設定

まず、ここで扱う実験環境を説明する。図5.2に示すように、周期境界条件の2Dフィールド（空間が繋がっており、フィールドの左・右・上・下端に着くと、反対側から同じ方向で出てくる）の中で移動するエージェント集団にBRT-LISモデルを実装する。黒い○はエージェント P_i を表し、点線の円は円の中心にある P_i の近傍範囲で、灰色の○は P_i の近傍エージェントを表し、青い点は P_i の近傍に属さないエージェントを表す。赤い矢印はエージェントの選択肢 $A = \{a_1, a_2, a_3, a_4\}$ を表し、それぞれ左・右・上・下に移動する行動を意味する。

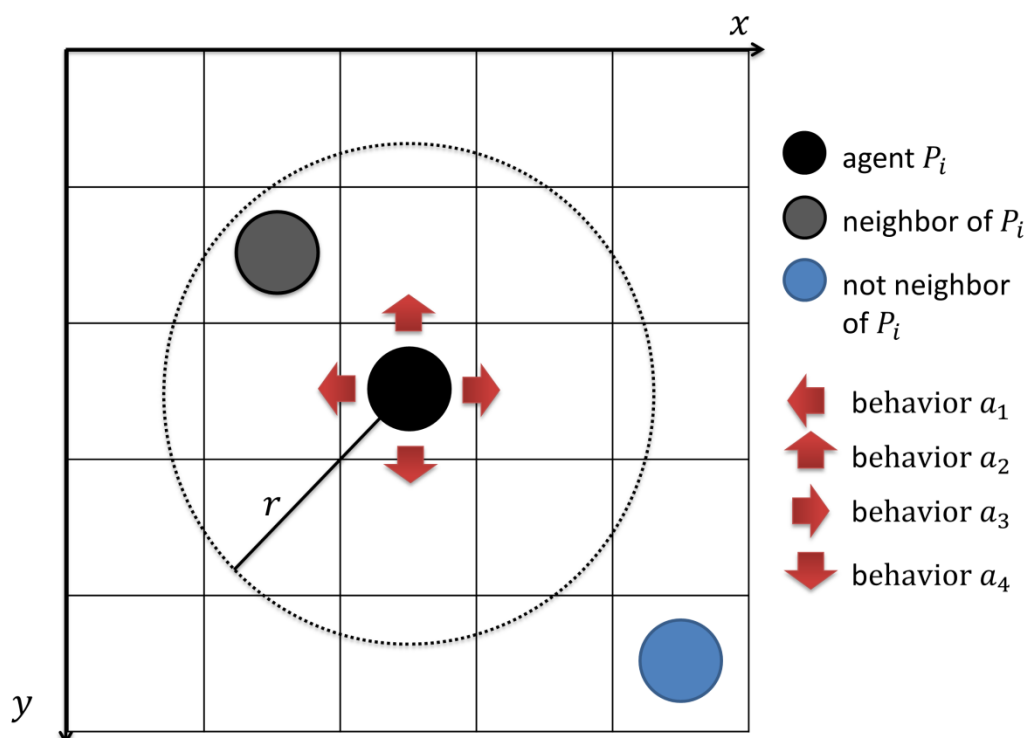


図5.2 実験設定の説明.

表 5.1 実験のパラメータ設定.

記号	定義	設定値
h, w	フィールドのサイズ：縦と横の長さ	600, 600
M	選択枝の数	4
μ	バイアス分布の期待値	$1/M$
σ^2	バイアス分布の分散	$1/(3M)$
τ	支持者数の増加量	0.001

実験のパラメータ設定を表 5.1 に示す. これらのパラメータは, 第 3 章の計算機実験において使用したパラメータと等しい. 図 5.3 に示すように, 初期時に, エージェントをランダムにフィールドに配置し, BRT-LIS アルゴリズムによる意思決定を開始した.

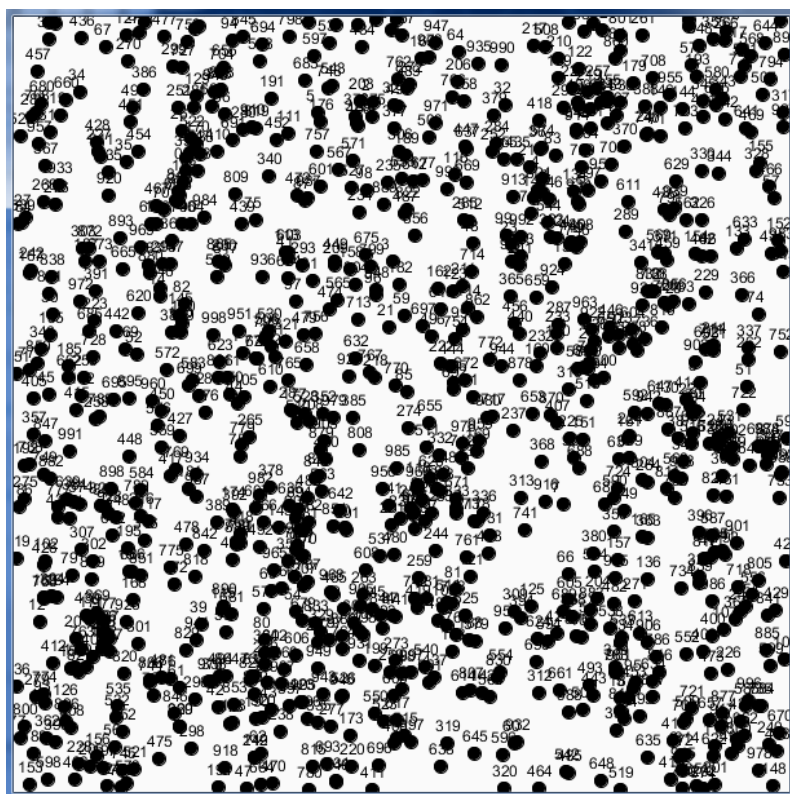


図 5.3 初期時にランダムにエージェントを配置した例.

5.3.2 挙動

まず、BRT-LIS モデルでも合意状態の切り替えについて注目して検討を行う。適した選択肢を設定せず BRT-LIS モデルの挙動を図示し、その特徴を明らかにする。

エージェント数 $N = 1000$ とし、初期時において全エージェントが選択肢 a_1 (左に移動する) を選んでいると設定し、合意状態から意思決定をスタートする。すべてのエージェントの近傍の半径 r を同時に 20, 50, 100, 200 と小さい値から大きい値まで変えて実行した。図 5.4~5.7 にそれぞれ r において各選択肢の支持者数の遷移を示す。

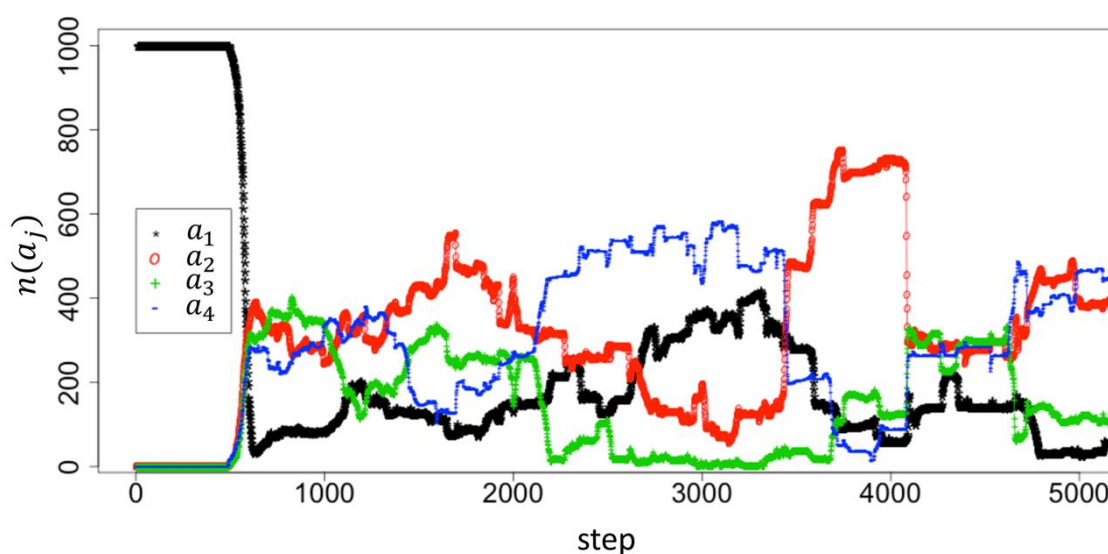


図 5.4 各選択肢の支持者数の遷移. $r = 20$.

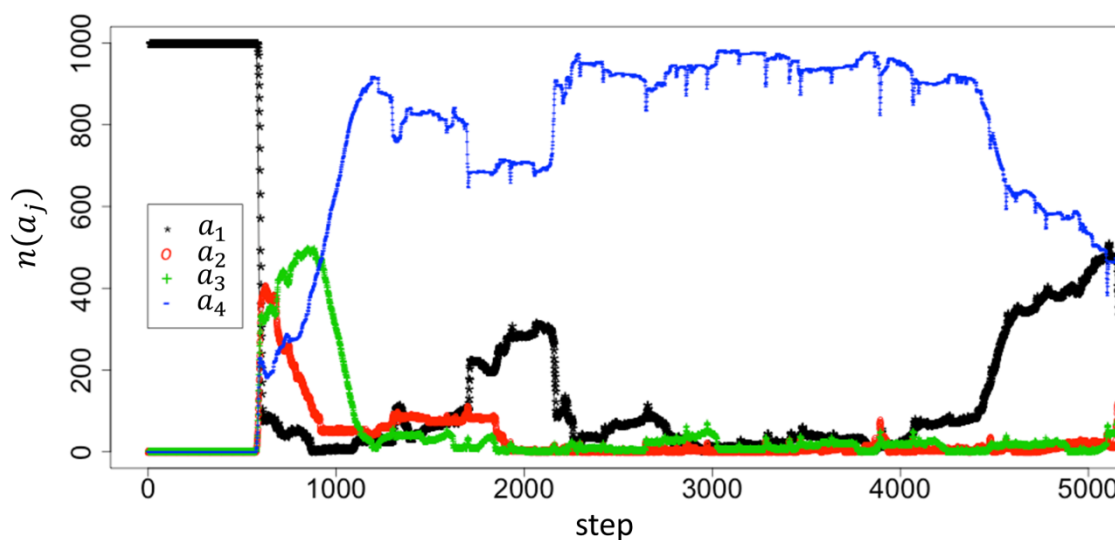


図 5.5 各選択肢の支持者数の遷移. $r = 50$.

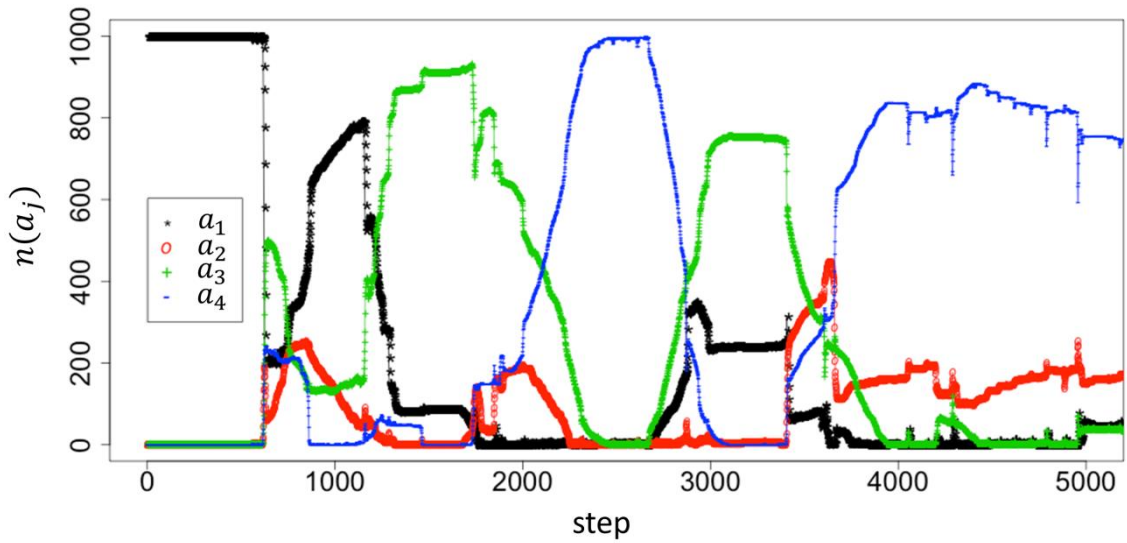


図 5.6 各選択肢の支持者数の遷移. $r = 100$.

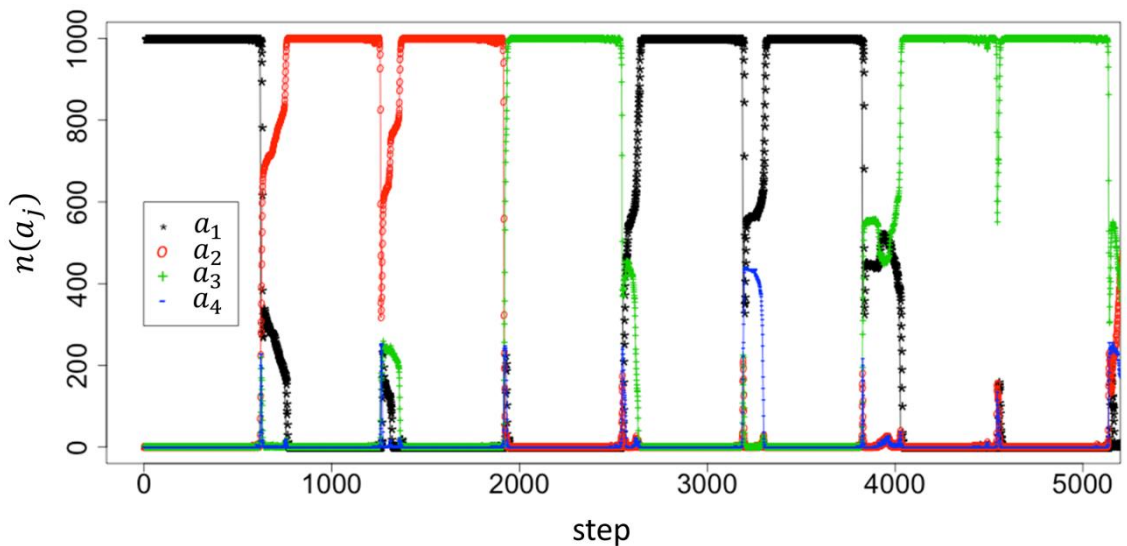


図 5.7 各選択肢の支持者数の遷移. $r = 200$.

これらの図により、エージェントの近傍の半径 r が小さいケース ($r = 20, 50$) ではエージェント全員である 1つの選択肢に合意することなく、フィールドでバラバラに移動し続ける。図 5.8 に個体の観測できる範囲の割合を示すように、半径 $r = 20, 50$ の場合では個体はそれぞれ全体の約 0.3%と 2.2%の領域しか観測できない為、平均的に各個体に約 3, 22 台の近傍のエージェントしかいない為、そのグループ中でさえ合意するのが難しい。従って、半径 r が小さいと全体的に合意状態に至らない事がわかる。

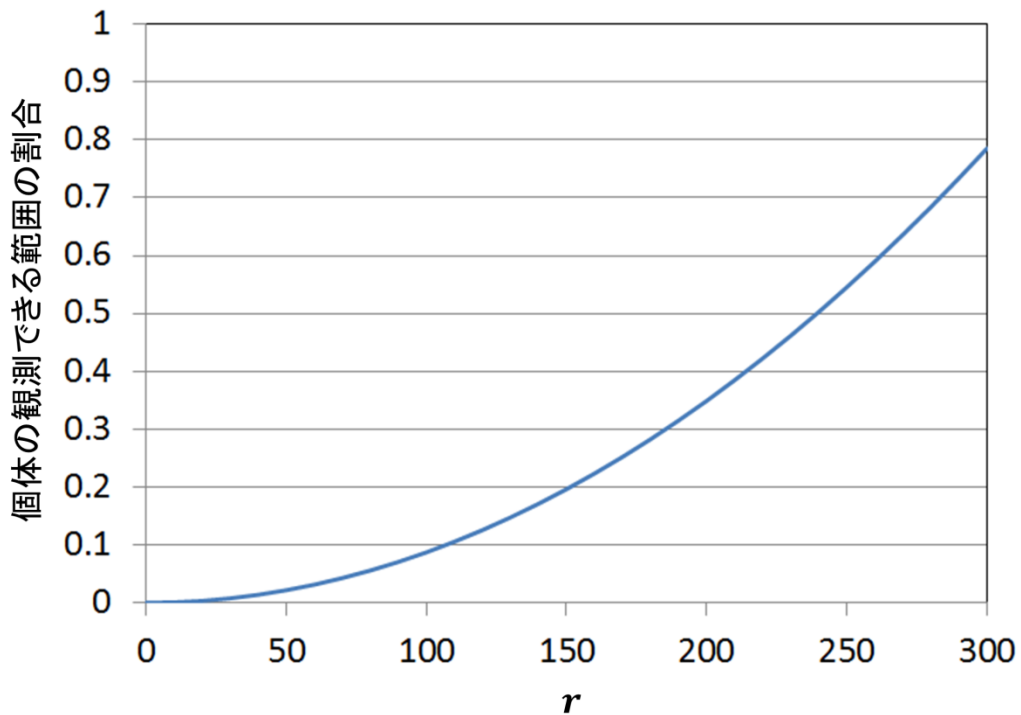


図 5.8 個体の観測できる範囲の割合.

半径 $r = 100$ の場合では、個体は、全体の約 9%しか見えないが、平均的に各個体に約 90 台のエージェントが近傍にいるため、そのグループ中では合意できる。また、他のグループと交差すると拡大したり、もしくは吸収されたりするので、図 5.6 に示すように約 900 ステップ後に選択肢 a_1 の支持者数が急激に増加したが、約 1100 ステップには選択肢 a_1 と入れ替わって選択肢 a_3 支持者数がトップになる。このように、局所的に合意した小さいグループがいくつか発生し、合体と離脱の繰り返しにより大きいグループが形成されるが全体的に合意には至らない事がわかる。

半径 $r = 200$ の場合では、個体の観測できる範囲は全体の約 35%である。図 5.7 に示すように、全体的に合意とその合意の切り替えが速やかに行われている事がわかる。

以上から、BRT-LIS モデルでは、ある程度大きい半径であれば集団は速やかに合意することができ、また合意を切り替えることもできる。従って、適した選択肢が発見できるかどうかについては大いに半径 r に依存することを推測できる。これは動的な環境でのコンセンサスアルゴリズム^[47,48]で広く研究されたトピックであり、本研究ですべてをカバーする事は難しい。そこで、以下では半径 r の影響について簡単な検討を行う。

5.3.3 適した選択枝の探索とその特徴

ここでは、計算機実験により、適した選択枝が発見できる事を示すと共に観測範囲の影響を明らかにする。

まず、適した選択枝の発見には半径 r がどのような影響を与えるかについて詳しく調べる。

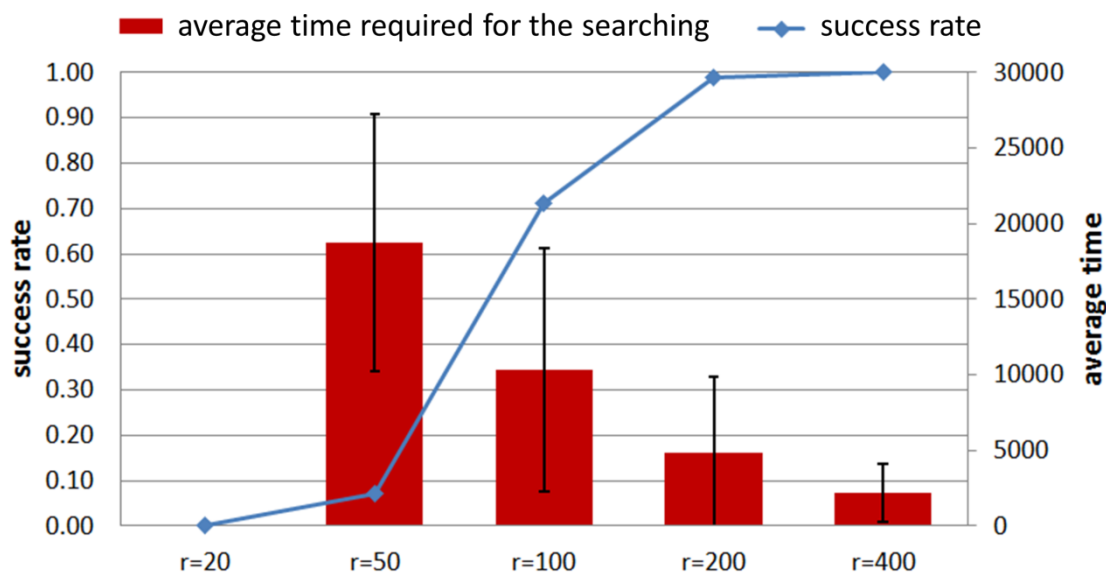


図5.8 適した選択枝の発見に要する平均時間とその発見率。

$$N = 100, A_i(0) = a_1 \forall i \in [1, N], a_{goal} = a_4.$$

実験条件は次のとおりである。エージェント数 $N = 100$, 初期時に全エージェントが選択枝 a_1 を選び、適した選択枝を a_4 とした。半径 r を 20, 50, 100, 200, 400 と変えて実行し、適した選択枝 a_4 が集団で発見できたケースをカウントし、発見までに要した時間を測定した。制限時間を 30000 ステップにし、それぞれについて 1000 試行を実施した。これらの実験結果を図 5.8 および図 5.9 に示す。

図 5.8 に適した選択枝の発見に要する平均時間とその発見率を示す。左側の縦軸と折線は発見率、右側の縦軸と赤い棒は発見に要する平均時間、黒い棒は平均時間の標準偏差を表している。まず、折線を見ると適した選択枝の発見率は、通信範囲の半径 r の増加と共に増加し、半径 r が 100 以上のときは 100% になる事がわかった。これは BRT-LIS モデルが BRT モデルと同じように発見できる事を意味する。また、赤い棒をみると、半径 r が大きければ大きいほど適した選択枝の平均発見時間が、半径 r とほぼ反比例しながら減少している事がわかる。そして、その標準偏差は小さくなっている事もわかる。

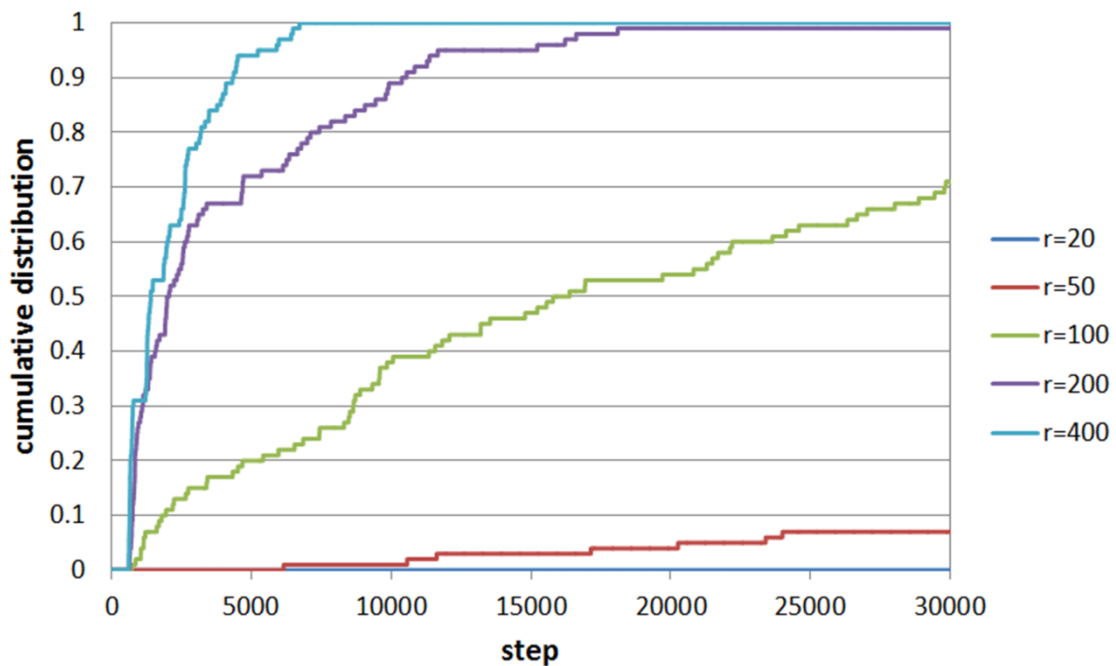


図5.9 探索時間の累積確率密度分布.

$$N = 100, A_i(0) = a_1 \forall i \in [1, N], a_{goal} = a_4.$$

図 5.9 はこの時の探索時間を累積確率密度分布で表したものである。横軸に探索に要した時間, 縦軸にこの時間まで適した選択枝を発見する確率を表している。半径 $r = 20$ の場合では制限時間まで適した選択枝を発見できなかった。半径 $r = 50$ の場合では, 発見率が約 0.8%でしか適した選択枝が発見できなかった。これより半 r が小さすぎると, 適した選択枝が中々発見できない事がわかる。その一方, 半径 $r = 100$ に広げると発見率が約 72%で適した選択枝が発見できるケースが劇的に増加した。半径 r をさらに上げると ($r = 200$) 約 18200 ステップまでにすべての試行において適した選択枝を発見できた事がわかる。

以上から, 少し通信範囲を広げるだけで大幅に精度の改善がみられ, 通信範囲とパフォーマンスには閾值的な関係がある事を推定した。

5.4 まとめ

本章では、提案した BRT モデルを観測範囲に限りがあるよりリアルな環境での可用性を検討した。個体の観測範囲を制限した移動エージェントのシミュレーションにより、パラメータを適切に設定すれば適した選択肢が発見できる事が確認できた。また、少し観測範囲を広げるだけで大幅に精度の改善がみられ、観測範囲とパフォーマンスには閾值的な関係がある事を推定した。これは自然の生き物の群れにおいても起こりうる事で、群れサイズが一定レベルに達さなければ正しい決定を導く事ができない。

第6章 BRT モデルその③：

集団行動のプリファレンスの実現

6.1 はじめに

生物の群れでは集団行動を異なる頻度で行い、生活環境の変化に上手に対応している事がよく知られている。例えば、アフリカスイギュウの群れは、1日の暑い時間帯には、普通、休息のために座り込んでいる。夕暮れが近づくと採食するため群れで移動する。暑い時間帯に移動する群れとはまれである^[46]。すなわち、アフリカスイギュウの集団行動には優先度が存在し、日差しや気温などの厳しい環境の変化に対して集団行動の優先度を適切に調整することにより、安定した群生活を実現している^[61]。このように優先度のある行動を行動のプリファレンスと呼ぶ。BRT モデルにもこのような集団行動のプリファレンスを実現できればロボット集団に高度な応用力を持たせる事が期待できる。



図 6.1 散水穴の付近で休息するアフリカスイギュウの群れ^[46]。

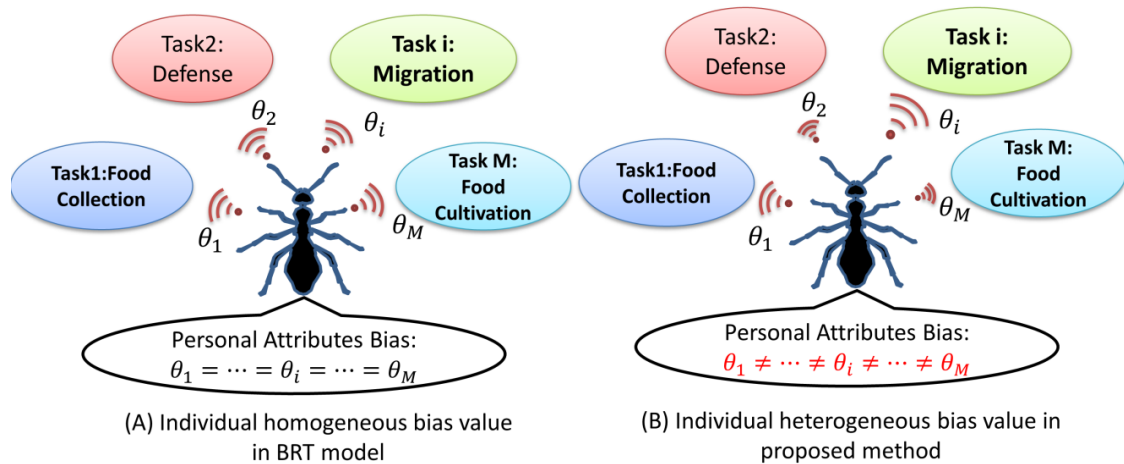


図 6.2 複数の選択肢に対して個体が有する
(A) 同一バイアス値及び(B) 異なる複数バイアス値.

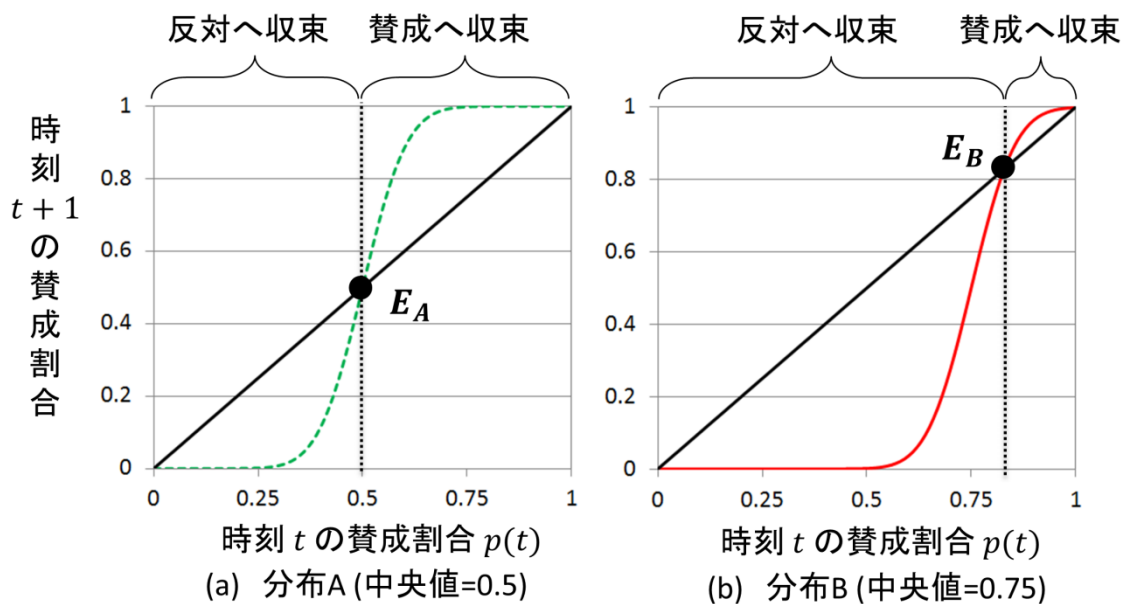


図6.3 バイアス分布の性質がわかる一例.

図 6.2A に示すように、提案した BRT モデルでは、選択肢の候補が M 個があるのに対し、群れの個体は個人属性を表すバイアス値 θ_i を 1 つしか持たない。Iwanaga ら^[28]は、個人属性というのは選択肢に対する選好度と、主流派の一員であることの個人的な望ましき(あるいは孤立することへの恐れ)の強さをバイアス値 θ_i によって表すと定義した。従って、図 6.2B に示すように個体が、異なる選択肢に対して異なるバイアス値を有すれば、BRT モデルによりリアルな社会性を持たせる事が可能かもしれない。異なる選好度や価値観を持つ人々で構成される

集団にはプリファレンスのある集団行動を生成できるかもしれない。

また、第 3.2 節で述べたように、Iwanaga らの研究では、バイアス値の分布を釣鐘型に設定するとバイアス値の累積確率分布には不安定点が発生し、その点を境に、初期状態によって集団は賛成もしくは反対状態に移行する事を明らかにしている。図 6.3 に示すように、中央値の大きい分布 B は分布 A と比較して反対状態を発生させる頻度が賛成状態より高い事がわかる。従って、選択肢の発生頻度が分布の中央値に大きく依存する事がわかる。この性質を利用すれば集団行動のプリファレンスを実現できるかもしれない。

このような背景から、本章では、複数の選択肢に対して異なる中央値を有するバイアス分布を BRT モデルに導入し、集団行動のプリファレンスの実現を試みる。

6.2 提案手法

6.2.1 意思決定のアルゴリズム

ここでは、それぞれの選択肢に対して異なる中央値を有するバイアス分布を用いた HBRT (Heterogeneous Bias and Rising Threshold) モデルの意思決定アルゴリズムを提案する。

BRT モデル同様に、 N 台のエージェントで構成される集団 $G = \{P_i: i = 1, \dots, N\}$ において、 M 個の選択肢の中から適した選択肢を 1 つ選ぶ意思決定を行うとする。エージェントの行動集合を $A = \{a_j: j = 1, \dots, M\} (M \geq 2)$ とし、探すべき適した選択肢を a_{goal} とする ($a_{goal} \in A$)。エージェントは事前に a_{goal} を知らず、全員が合意して初めて合意した選択肢が a_{goal} であるか分かるものとする。エージェント P_i の時刻 t での選択肢を $A_i(t) \in A$ とする。 $n(a_j)$ を選択肢 a_j を選択しているエージェントの数とする。

ここでは、エージェント P_i は、選択肢 a_j に対してバイアス値 θ_{ij} を持つ ($0 < \theta_{ij} < 1$) を持つ。時刻 t において、エージェント P_i は集団全体で自身と同じ選択肢を選んでいる割合を $n(A_i(t))/N$ 、次の時刻に選ぶ選択肢 $A_i(t+1)$ を次のように決定する。もし

$$\frac{n(A_i(t))}{N} \geq \theta_{ij} + \tau \cdot c_i(t) \cdot (t - t_{i,last}(t)) \quad (6.1)$$

であれば、現在の選択肢 $A_i(t)$ を引き続き選択し、 $A_i(t+1) = A_i(t)$ とする。そうでなければ現在の選択肢 $A_i(t)$ をやめてそれ以外の選択肢を確率的に選ぶ ($A_i(t+1) \in A - A_i(t)$)。

BRT モデルと同様、 τ は支持者数の増加量を表す定数である。 $t_{i,last}(t)$ はエージェント P_i が最後に選択肢を変更した時刻で、 $(t - t_{i,last}(t))$ で同一選択肢を選び続けている時間に相当し、以下の式で求める。

$$t_{i,last}(t) = \begin{cases} t & A_i(t) \neq A_i(t-1) \\ t_{i,last}(t-1) & otherwise \end{cases} \quad (6.2)$$

$c_i(t)$ は選択肢が“適している”時 0 である関数である。すなわち

$$c_i(t) = c(t) = \begin{cases} 0 & \forall i, A_i(t) \equiv a_{goal} \\ 1 & otherwise \end{cases} \quad (6.3)$$

である。

6.2.2 バイアス値の生成

提案手法は、BRT モデルと同様に、ガウシアン分布に基づいた鐘型曲線を使用してバイアス値を生成する：

選択肢 a_j に対してエージェント P_i が有するバイアス値

$$\theta_{ij} = \min(1, \max(0, G(\mu_j, \sigma_j^2))) \quad (6.4)$$

である、 $G(\mu_j, \sigma_j^2)$ は平均 μ_j 、分散 σ_j^2 の正規分布に基づく乱数で

$$G(\mu_j, \sigma_j^2) = \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left(-\frac{(x - \mu_j)^2}{2\sigma_j^2}\right) \quad (6.5)$$

である。

6.3 実験

本節では、HBRT モデルの挙動を示し、その特徴を明らかにする。そして、計算機実験により、集団行動のプリファレンスを実現できる事を明らかにする。

6.3.1 意思決定の挙動

はじめに、HBRT モデルの意思決定の挙動を図示し、合意の形成と合意状態の切り替えが発生するかについて検討を行う。そのため、適した選択肢 a_{goal} を設定しない。

実験の設定は以下の通りである。 $N = 1000$, $M = 3$, $\tau = 0.005$ とした。初期時で全員が選択肢 a_1 を選んだ状態からスタートとし、3つの選択肢に対して以下の4つのケースでバイアス値を生成した。異なる中央値のケースに加えて、比較対象として同じ中央値のケースと異なる分散値も用意した。時間経過に各選択肢の支持者数の遷移を図 6.4~図 6.7 に示す。横軸が時間で、縦軸は各選択肢の支持者数 $n(a_j)$, $j = 1, 2, 3$ である。

- ケース 1: 同じ中央値 $\mu_1 = \mu_2 = \mu_3 = \frac{1}{M}$

$$\text{異なる分散値 } \sigma_1^2 = \frac{1.0}{3M}, \sigma_2^2 = \frac{1.0}{6M}, \sigma_3^2 = \frac{1.0}{9M}$$

- ケース 2: 同じ中央値 $\mu_1 = \mu_2 = \mu_3 = \frac{1}{M}$

$$\text{異なる分散値 } \sigma_1^2 = \frac{1.0}{3M}, \sigma_2^2 = \frac{1.0}{6M}, \sigma_3^2 = \frac{1.0}{12M}$$

- ケース 3: 異なる中央値 $\mu_1 = \frac{1}{M}, \mu_2 = \frac{1}{0.5M}, \mu_3 = \frac{1}{1.5M}$

$$\text{同じ分散値 } \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \frac{1.0}{3M}$$

- ケース 4: 異なる中央値 $\mu_1 = \frac{1}{M}, \mu_2 = \frac{1}{1.5M}, \mu_3 = \frac{1}{0.5M}$

$$\text{異なる分散値 } \sigma_1^2 = \frac{1.0}{3M}, \sigma_2^2 = \frac{1.0}{6M}, \sigma_3^2 = \frac{1.0}{9M}$$

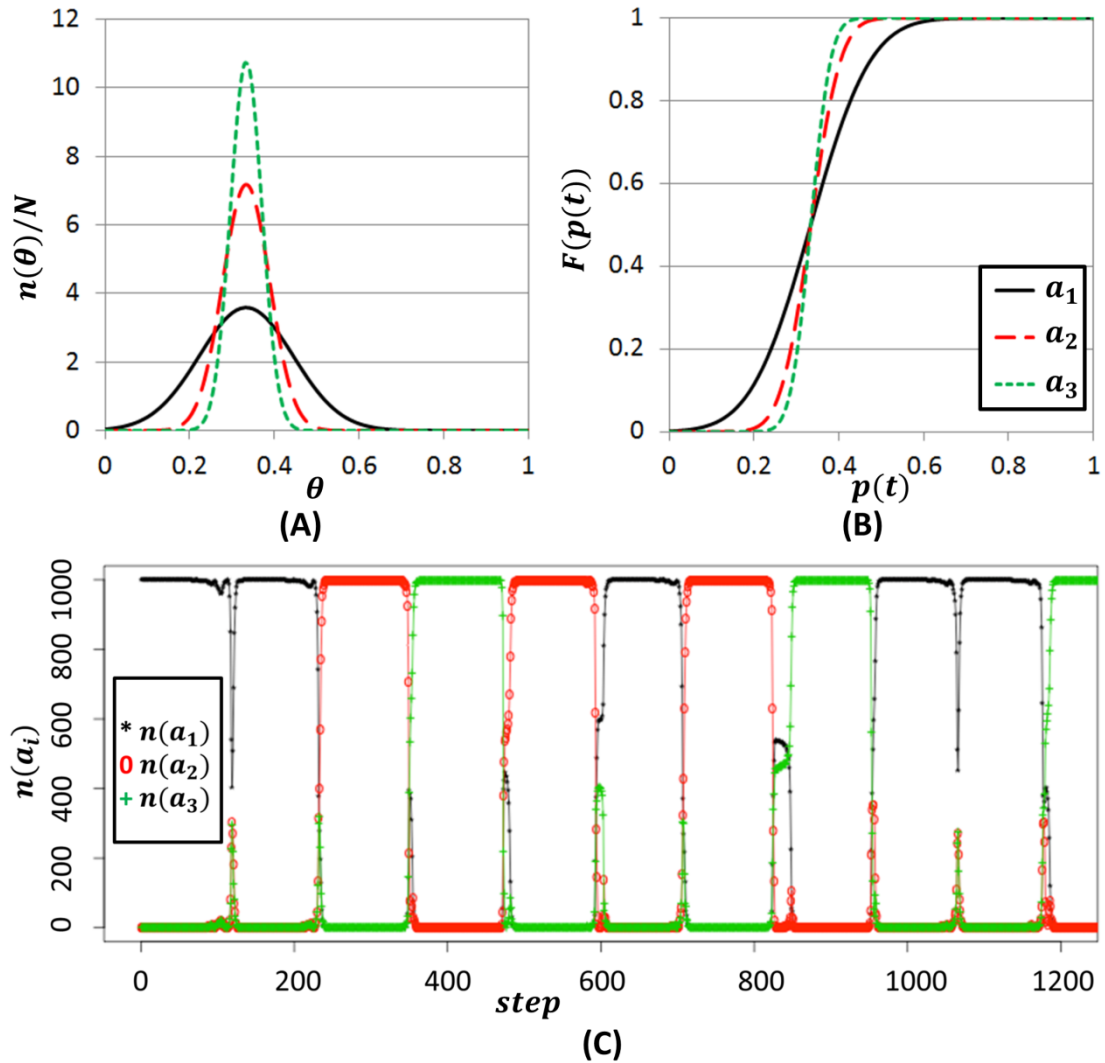


図 6.4 ケース 1 における (A) バイアス分布, (B) 累積確率分布と (C) 時間経過に各選択枝の支持者数の遷移.

ケース 1 では, 3 つの選択枝に対して同じ中央値で異なる分散値の 3 つの分布を使用してバイアス値を生成した. 全体的に合意の形成と合意状態の切り替えができていることが確認できた. また, 分散値が小さいと分布の鐘型が狭くなるためバイアス値 θ_{ij} は平均値あたりに集中し, 少数派の意見が反映されにくく合意状態が安定に長く伸びることがわかった. 図 6.4 から, 比較的に小さい分散値の選択枝 a_2 及び選択枝 a_3 に合意した状態が選択枝 a_1 のときより安定で長く保たれることがわかる. その一方, 選択枝 a_1 に対して分散値が大きいため集団には大きいバイアス値を持つエージェントが比較的に多く存在し, これらのエージェントは速めに式 (6.1) が満たされなくなり, 別の選択枝を選び始める. しかし, 他の選択枝が 2 つあるため支持者の割合が別れ, 自身の大きいバイアス値より小

さくなり、再び別の選択肢を支持しなければならない。これにより、多数派の支持率が減少したり増加したりして不安定な状態となる。この現象は、例えば、約 80 ステップと 185 ステップなどにみられた。

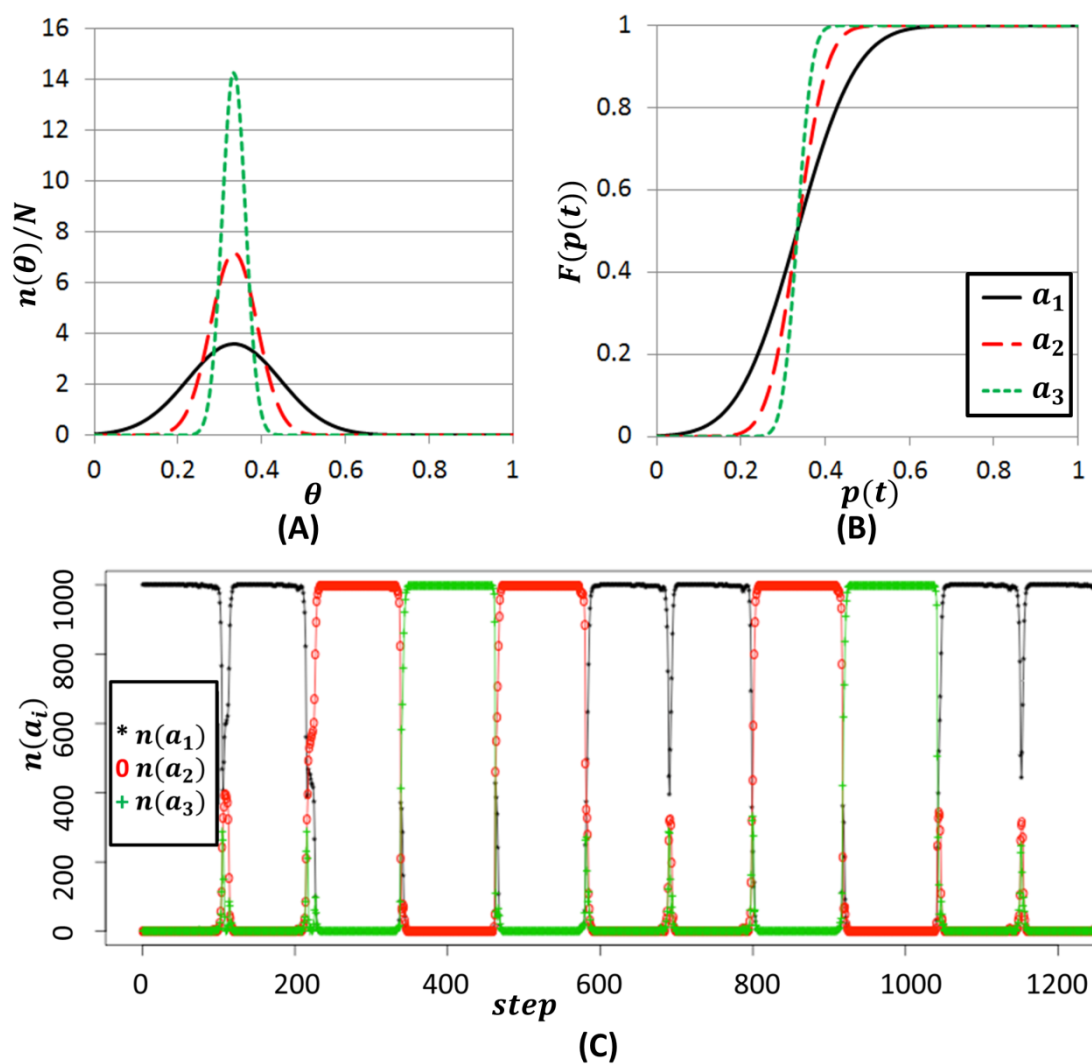


図 6.5 ケース 2 における (A) バイアス分布, (B) 累積確率分布と (C) 時間経過に各選択肢の支持者数の遷移.

ケース 2 では、ケース 1 と同様、全体的に合意の形成と合意状態の切り替えができていたことが確認できた。また、比較的分散値が小さいほど速やかな合意形成と安定で長い合意状態がみられた。

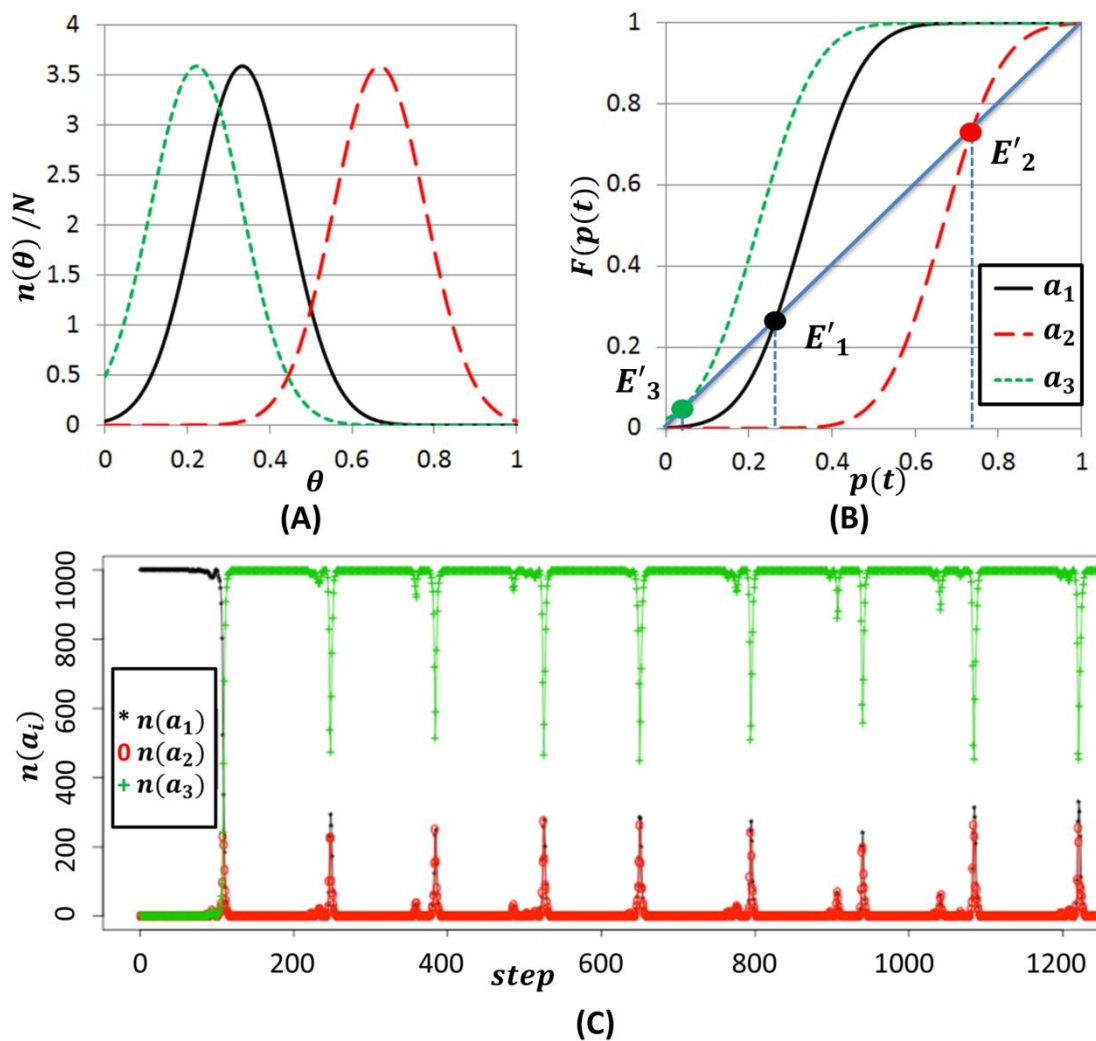


図 6.6 ケース 3 における (A) バイアス分布, (B) 累積確率分布と (C) 時間経過に各選択肢の支持者数の遷移.

ケース 3 では, 3 つの選択肢に対して異なる中央値で同じ分散値の 3 つの分布を使用してバイアス値を生成した. 全体的に合意の形成がみられたが 2 回目の合意状態からずっと選択肢 a_3 が多数派で合意状態になったり合意状態が崩れたりしていた. このような現象は従来の BRT モデルではみられなかった. その原因としては, 約 80 ステップから選択肢 a_1 に合意した状態は崩れて, 約 100 ステップにはそれぞれの支持者の割合がほぼランダムな状態となった (約 $1/3$). 図 6.6(B) に示すように, E'_1 , E'_2 , E'_3 はそれぞれ選択肢 a_1 , a_2 , a_3 に対するバイアス値の累積確率分布の均衡点 (不安定点) を表している. この図から, $E'_3 < E'_1 < \frac{1}{3} < E'_2$ が成り立つ事がわかる. 従って, もし支持者の割合が約 $1/3$ になった場合は, 選択

肢 a_2 の支持者は選択肢 a_1 と a_3 に流れる. しかし, $\mu_1 = \frac{1}{M}$ に集中するバイアス値 θ_{i1} は, $\mu_3 = \frac{1}{1.5M}$ に集中するバイアス値 θ_{i3} より大きいため選択肢 a_2 の支持者はまた選択肢 a_2 と a_3 に流れる. その結果, 選択肢 a_3 の支持者が多数派となり, そして合意状態となった. E'_3 がもの凄く小さく E'_1 と E'_2 と離れているため, 合意状態が崩れてもすぐに元に戻り, 合意状態が再び繰り返される.

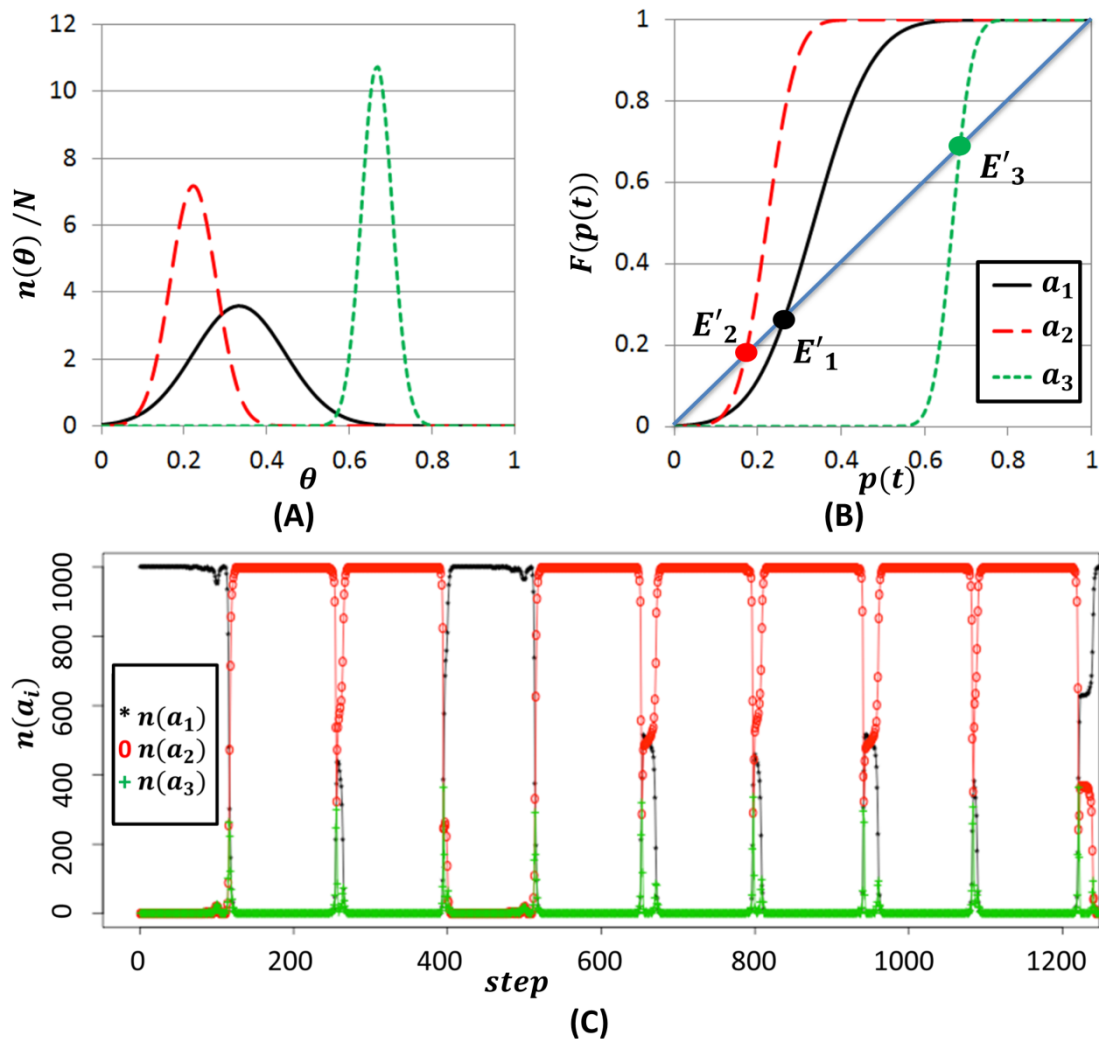


図 6.7 ケース 4 における (A) バイアス分布, (B) 累積確率分布と (C) 時間経過に各選択肢の支持者数の遷移.

ケース 4 では, 3 つの選択肢に対して離れた異なる中央値で異なる分散値の 3 つの分布を使用してバイアス値を生成した. ケース 3 と同様, 最も小さい中央値をもつ選択肢は最も選ばれ合意状態となることがわかった. また, 小さい分散値

の選択肢はより安定に合意状態が保たれていることがもう一度確認できた。

以上から、より小さい中央値の選択肢は集団全体でより高い確率で選ばれることがわかった。また、小さい分散値では、より素早く合意が形成され、安定に合意状態が保たれることもわかった。従って、選択肢のバイアス分布のパラメータである中央値及び分散値を適切に異なる値に設定することにより、集団行動のプリファレンスを実現すること期待できる。

6.3.2 集団行動のプリファレンスの実現

前節では、異なる選択肢に対するバイアス分布の中央値を異なる値に設定することで異なる比率の集団行動が発生することが確認できた。ここでは、この特徴を使用し、集団行動のプリファレンスを実現する。

実験の設定は以下の通りである。 $N = 100$, $M = 3$, $\tau = 0.005$ とした。適した選択肢 a_{goal} を設定せず、初期時で全員が選択肢 a_1 を選んだ状態からスタートとした。分布の設定は前節の4つのケースに加えて以下のケース5を用意する。

- ケース5: 異なる中央値 $\mu_1 = \frac{1}{M}$, $\mu_2 = \frac{1}{1.2M}$, $\mu_3 = \frac{1}{0.8M}$

$$\text{異なる分散値 } \sigma_1^2 = \frac{1.0}{3M}, \sigma_2^2 = \frac{1.0}{6M}, \sigma_3^2 = \frac{1.0}{9M}$$

ケース5は、ケース4と同様に異なる中央値と異なる分散値だが、図6.8に示すように、比較的ケース4より中央値が近いと設定した。こうする事により、ケース4とは異なる集団行動のプリファレンスを実現することが期待できる。

それぞれのケースを10000試行して、制限時間30000ステップ内にそれぞれの選択肢に何回合意したかを測定し、図6.9にその結果を示す。

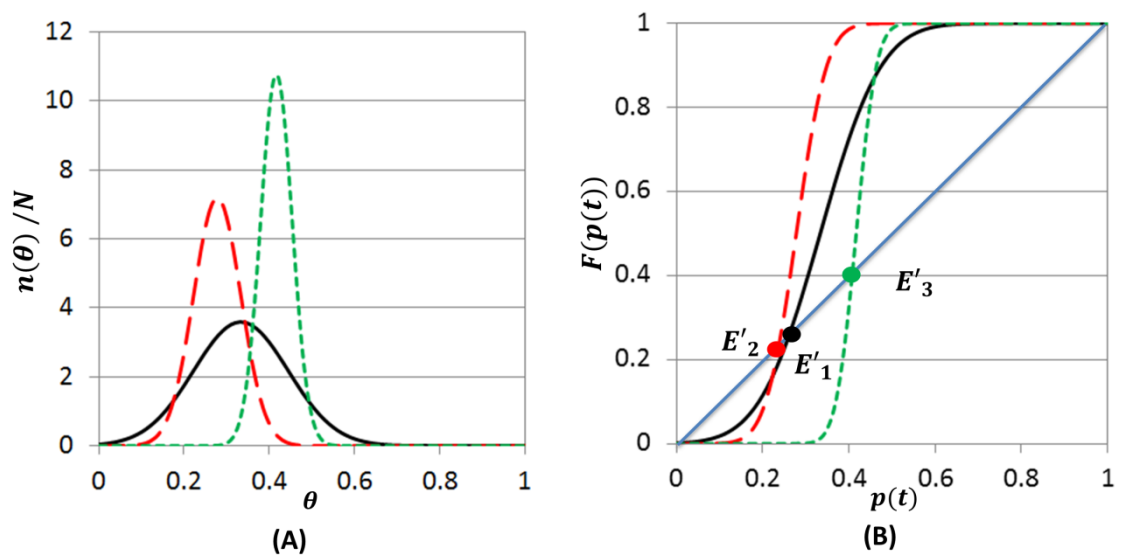


図 6.8 ケース 5 のバイアス分布とその累積分布関数.

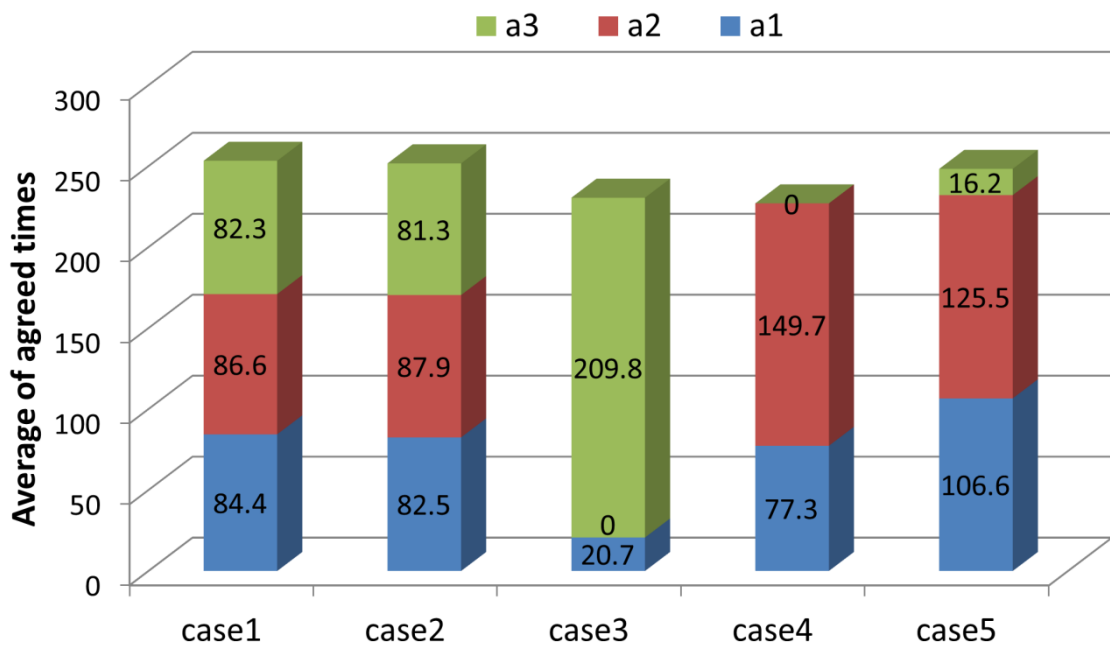


図 6.9 それぞれの選択肢に合意した平均回数.

縦軸が 1 つの試行で合意した平均回数で、青色、赤色、緑色のバーはそれぞれの選択肢 a_1 , a_2 , a_3 の実験結果である. ケース 1 とケース 2 をみると同じ中央値のため、分散値によらずほぼ各選択肢に均等に合意している事がわかる. 各選択肢が選ばれる確率は同じなのでプリファレンスのある集団行動とはいわない. その一方、ケース 3 では、最も中央値が突出している選択肢 a_2 ($\mu_2 = \frac{1}{0.5M} \gg \mu_1 =$

$\frac{1}{M} > \mu_3 = \frac{1}{1.5M}$)は一度も合意されなかった。圧倒的に合意されたのは最も中央値の小さい選択肢 a_3 である。よって、ケース 3 で格差の大きい集団行動のプリファレンスを実現した。ケース 4 は、ケース 3 と同様な中央値を設定したが、両側の分布の分散値を小さくした事により、不安定点 E'_2 と E'_1 (図 6.8(B)) の距離が大きく凝縮されて、選択肢 a_1 が選ばれやすくなった。また、図 6.9(B) に示すように、ケース 5 のバイアス分布はケース 4 のバイアス分布を真ん中に押し寄せられたものである。そのため、3つの不安定点 E'_1, E'_2, E'_3 の距離がさらに凝縮され、選択肢 a_1 と a_3 が選ばれやすくなった。従って、ケース 5 では、格差の少ない集団行動のプリファレンスを実現した。

以上から、選択肢のバイアス分布を適切に調整する事により集団行動のプリファレンスを様々な形で実現できる。

6.4 まとめ

本章では、環境の変化に対して高度な応用力を持つ群ロボットを目指し、集団行動のプリファレンス（優先度）を実現するため、それぞれの選択肢に対して異なる中央値を有するバイアス分布を用いた HBRT モデルを提案した。計算機実験により、それぞれの選択肢に対して個体のバイアス分布の平均値と分散値を適切に変えれば、集団行動のプリファレンスを実現できる事を明らかにした。

第7章 Q-BRT モデル：

2 次関数を用いたバイアス分布の生成

7.1 はじめに

第3章で提案した BRT モデルでは、ガウシアン分布に基づいた鐘型曲線を使用して個人属性を表すバイアス値を生成した。しかし、ガウシアン分布は $\theta_i \leq 0$ と $\theta_i \geq 1$ の領域の処理がヒューリスティックになり、BRT モデルの収束について確率論から離れて議論する事が難しかった。

そこで、本章では、解析のしやすい2次関数を用いたバイアス分布を使用する Q-BRT モデルを提案する。また、数学的な分析と計算機実験を用いて、Q-BRT モデルの収束や適した選択枝の発見に関する特性を明らかにする。

7.2 提案手法：2 次関数を用いたバイアス分布

ここで、より解析しやすい2次関数を用いたバイアス分布を提案する。BRT モデルと同様、 N 台のエージェントで構成される集団 $G = \{P_i: i = 1, \dots, N\}$ は、 M 個の選択枝の中から1つの適した選択枝を探索しているとする。エージェント P_i は個人属性バイアス θ_i ($0 < \theta_i < 1$) を持つ。バイアス θ_i 次のように設定する。

$$n(\theta_i) = \begin{cases} 0 & \theta_i \leq 0 \\ k_1 N \theta_i^2 & 0 < \theta_i \leq \frac{1}{M} \\ k_1 N \left(\theta_i - \frac{2}{M}\right)^2 & \frac{1}{M} < \theta_i < \frac{2}{M} \\ 0 & \frac{2}{M} \leq \theta_i \end{cases} \quad (7.1)$$

ここで、 $n(\theta_i)$ とは、 θ_i をとる人数を表し、 $k_1 = 3M^3/2$ は正規化項である。これにより

$$\frac{\int_0^1 n(\theta_i) d\theta_i}{N} = 1 \quad (7.2)$$

である。

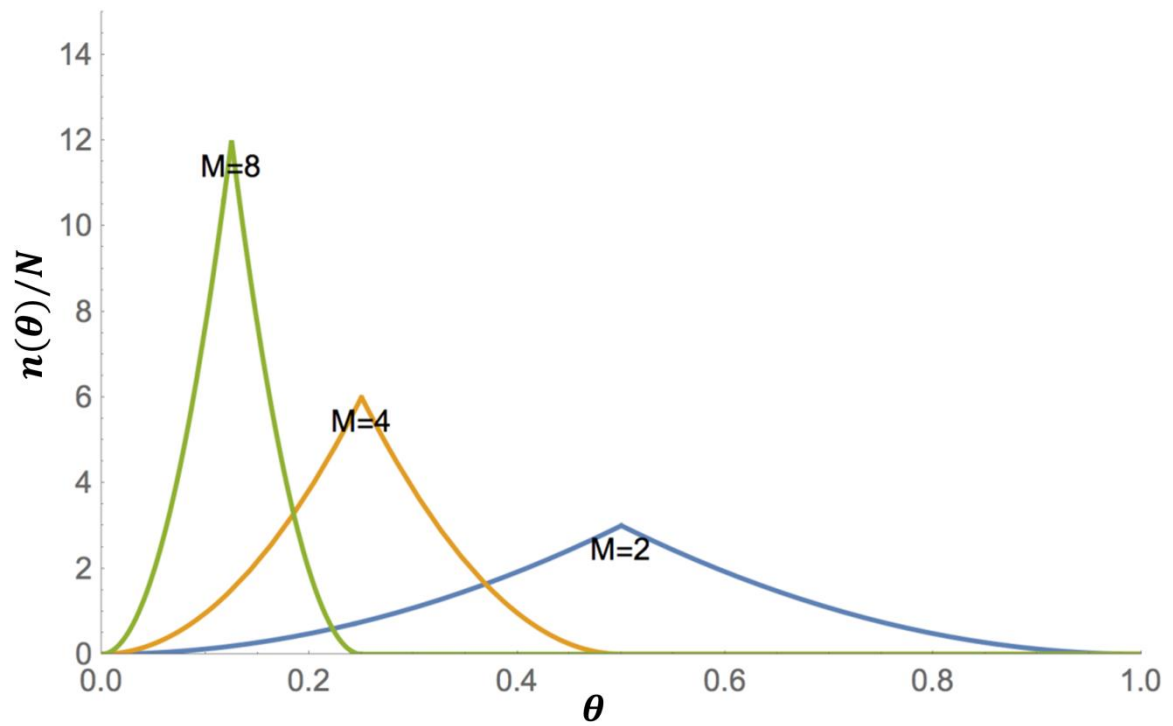


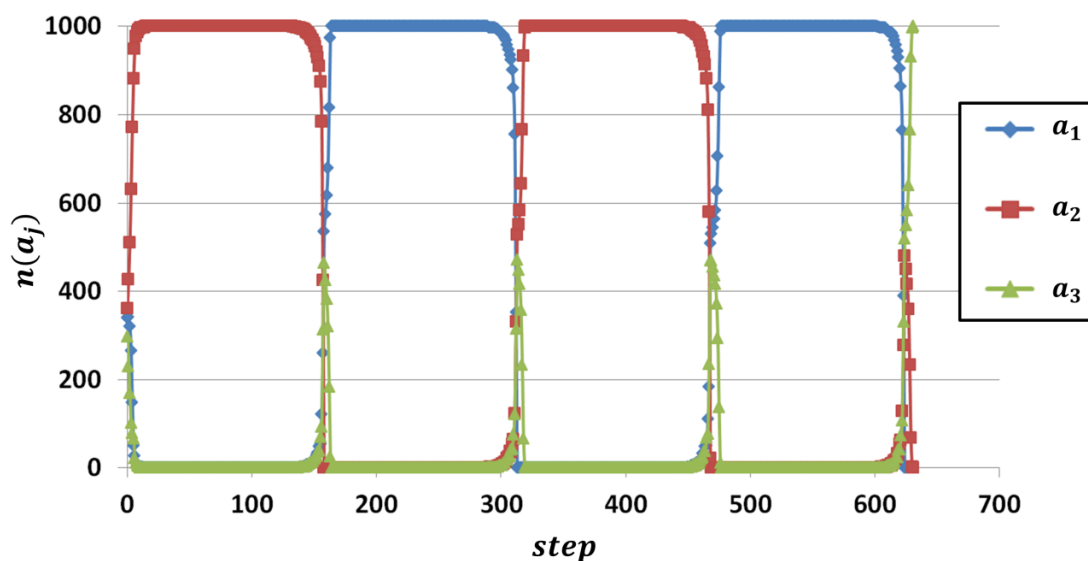
図 7.1 提案手法のバイアス分布の例.

図 7.1 に提案手法のバイアス分布の例を示す. バイアス分布の形状は, 選択肢の数 M にのみ依存し, 中央値 $\mu = 1/M$ が突出している.

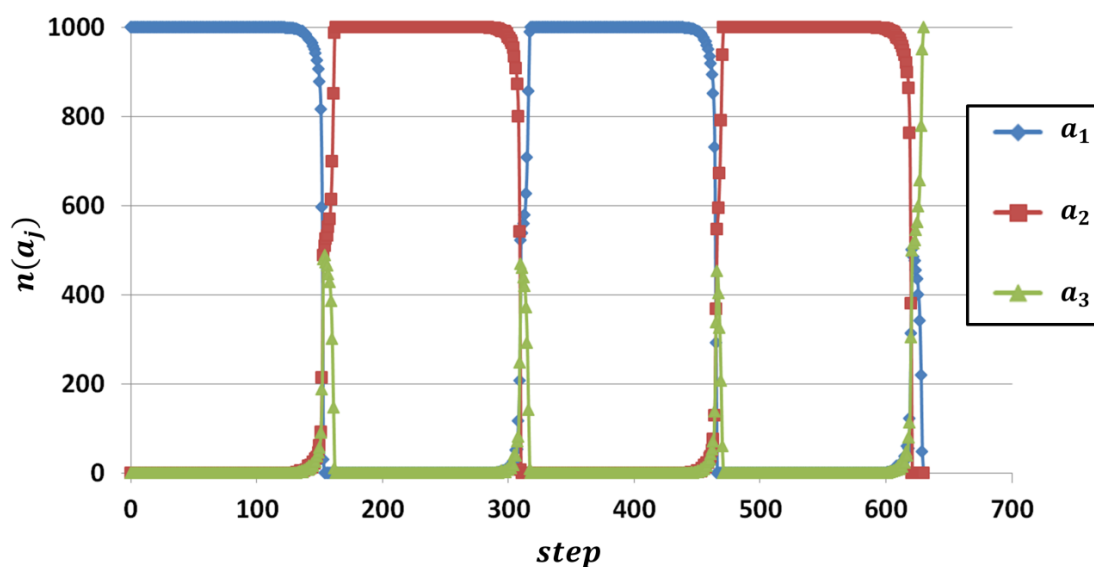
7.3 Q-BRT モデルの挙動

ここでは、以上で提案した Q-BRT モデルの挙動を図示しその特性を明らかにする。

実験の設定は以下の通りである。 $N = 1000$, $M = 3$, $\tau = 0.003$ とし、選択肢の集合 $A = \{a_1, a_2, a_3\}$ の内、適した選択肢 a_{goal} を a_3 と設定した。初期時にエージェントがランダムな選択肢を支持する場合と全員が不適切な合意状態からスタートの 2 通りを実装した。



(a) ランダムにスタート.



(b) 合意状態からにスタート.

図 7.2 各選択肢の支持者数の遷移(初期時にランダムにスタート).

図 7.2(a)はエージェントが初期時にランダムな選択肢を選択している時の各選択肢の支持者数の遷移を表している。横軸が時間で縦軸は各選択肢の支持者数 $n(a_j)$, $i = 1, 2, 3$ を表している。スタートから選択肢 a_2 の支持者数は約 360 で最も多かった為、時間経過により、 a_2 の支持者数がどんどん増加してゆき、約 12 ステップ後に、全員で選択肢 a_2 に合意した事がわかる。しかし、選択肢 a_2 は適した選択肢ではないため、約 140 ステップが経過すると a_2 をやめて別の選択肢を選択するエージェントがどんどん出てくる。これによって、群れの状態は合意状態から、ほぼ初期状態に戻り、再び意思決定を行う。この試行ではこの合意し直しプロセスがもう 3 回繰り返された。約 630 ステップで全員が適した選択肢 a_3 に合意し、終了した。よって、ランダムにスタートしても 3 つの選択肢の中から、群れは合意を切り替えて、適した選択肢を発見できている事がわかる。

図 7.2(b)に不適切な合意状態からの遷移例を示す。初期状態では選択肢 a_1 に合意しているが、約 165 ステップで別の選択肢 (a_2) に全員が合意した。しかし、選択肢 a_2 は適した選択肢ではない為、集団は再び別の選択肢に切り替えて、ついに、約 630 ステップで適した選択肢 a_3 が発見できた。従って、不適切な合意状態からでも適した選択肢を発見できている事がわかる。

以上から、2 次関数を用いたバイアス分布法でも群れは唯一の選択肢に速やかに合意し、適さない合意状態を切り替えて、適した選択肢を発見できる事がわかった。

7.4 数学的な解析とその検証実験

前節では、Q-BRT モデルの挙動を計算機実験により提案モデルの特性を明らかにした。ここでは、その特性に関する数学的な解析を行い、提案モデルの特性を明らかにする。

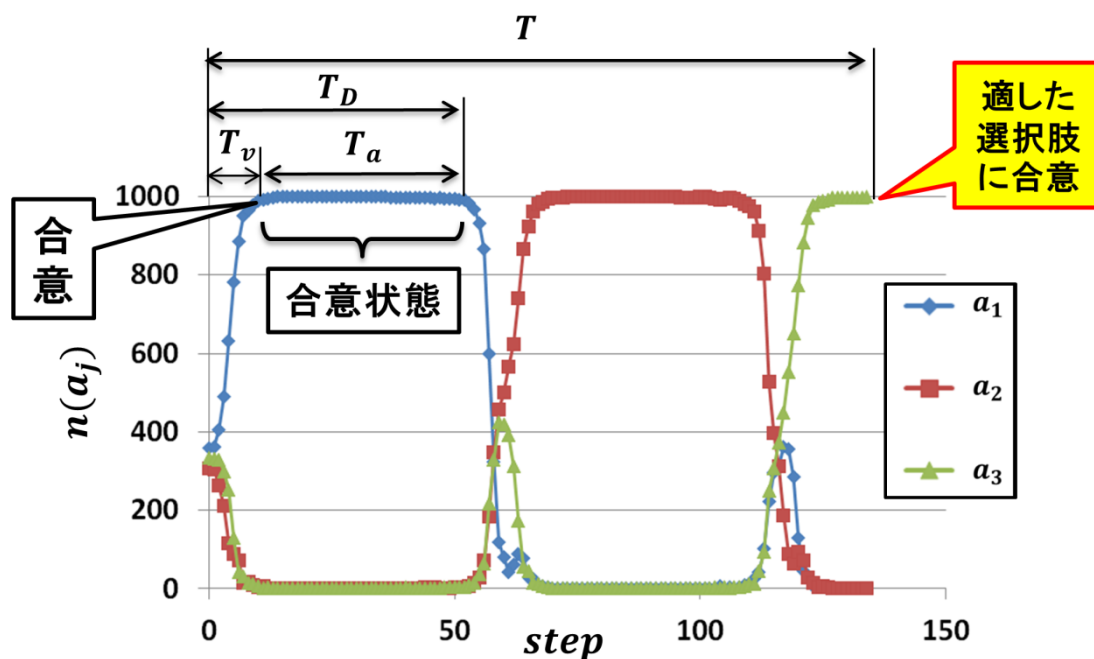


図 7.3 適した選択肢が発見される過程.

表 7.1 用語の定義

記号	定義
T_v	投票フェーズに要する時間
T_a	合意フェーズに要する時間
T_D	1回の試行に要する時間
T	適した選択肢の発見に要する時間
E	適した選択肢の発見に要する試行の回数

まず、初めに、以下で扱う用語を定義する。図 7.3 に示すように、ランダムな状態もしくは合意状態でない状態から開始して合意状態が発生するまでは、“投票フェーズ” (voting phase) と呼ぶ。投票フェーズに要する時間を T_v とする。合意状態が開始してから終了までは“合意フェーズ” (agreement phase) と呼び、要する時間を T_a とする。投票フェーズから合意フェーズが終了までは、“1回の試行” と呼び、それに要する時間を T_D とする。すなわち、

$$T_D = T_v + T_a \quad (7.3)$$

である。また，“適した選択肢の発見に要する時間を T ，要する試行の回数を E とする。表 7.1 にこれらの用語の定義を示す。

7.4.1 任意の選択肢に合意できるか

ここでは、Q-BRT モデルは選択肢の数 M が多数 ($M > 2$) でも、提案したバイアス分布を適切に設定することにより、群れは任意の選択肢に合意し、合意状態に収束する事を数学的に示す。

Iwanaga ら^[28]は、2 択の問題に対して、累積確率密度分布を S 字型のカーブになるようにバイアス分布を設定すると、不安定な均衡点 E_2 が発生し、賛成者の割合 $p(t)$ によって全員が賛成もしくは反対の状態に滑らかに推移する事を明らかにした。すなわち、初期時の賛成者の割合が、均衡点 E_2 より小さい（大きい）と、集団は反対の状態（賛成の状態）に収束する事が知られている。従って、どの M でも、もし、初期時にランダムで選択肢を選ぶ際、不均衡点より大きい支持率を持つ選択肢が必ず 1 つ存在することを証明できれば、群れは、合意することを明らかに出来る。以下では、この証明過程を説明する。

式(7.1)の累積確率密度関数を以下のように定める。

$$f(\theta) = \sum_{\theta_i < \theta} n(\theta_i) \quad (7.4)$$

$$F(\theta) = \frac{f(\theta)}{N} \quad (7.5)$$

図 7.1 に示したバイアス分布の累積確率密度関数 $F(\theta)$ を図 7.4 に示す。 $y = x$ として $F(p(t)) = p(t)$ のグラフも描画している。ここでは、 $p(t)$ とは時刻 t である選択肢を支持するエージェントの比率である。このように $F(\theta)$ のグラフは $F(p(t)) = p(t)$ とそれぞれ 3 点で交差しており、それは図 3.3 に示した $p(t) = 0$ (E_1 に相当) と $p(t) = 1$ (E_3 に相当) の 2 個の安定不動点とその間にある 1 点の不安定不動点 p^* (E_2 に相当) である。 E_2 よりも右側では $p(t + 1) > p(t)$ となるため、 $p(0) > E_2$ の選択肢に集団全体が急速に $p(t) = 1$ に収束する。 p^* の値は

$$F(p^*) = p^* \quad (7.6)$$

の解である。

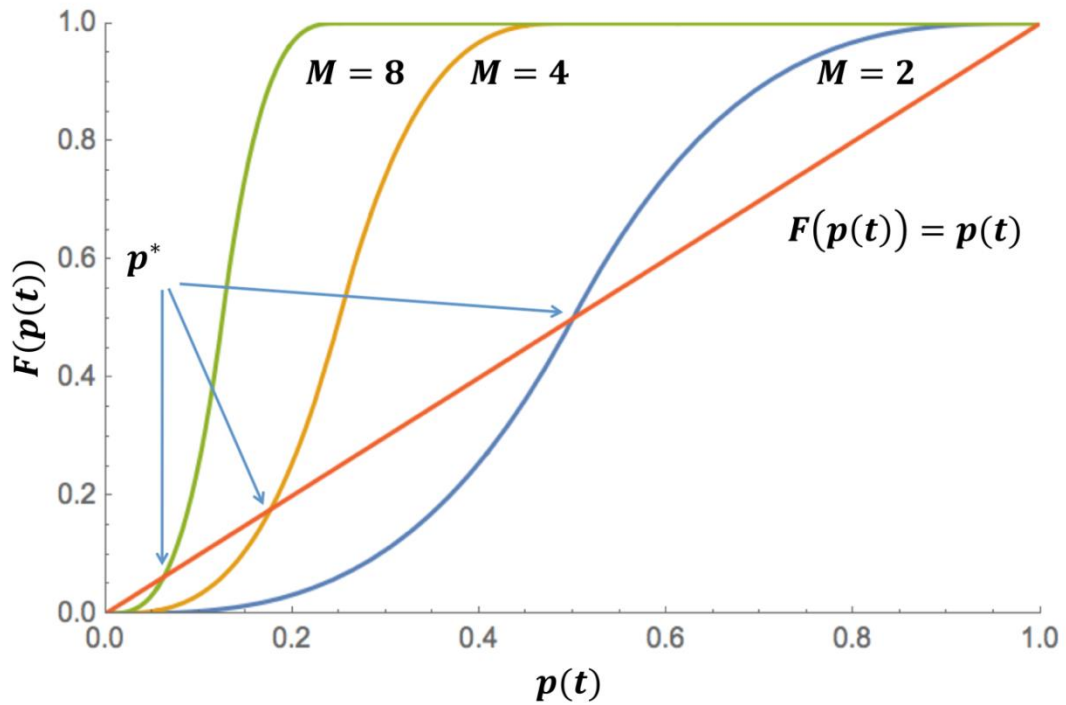


図 7.4 図 7.1 に示したバイアス分布の累積確率密度関数 $F(\theta)$.

式(7.1)について、方程式(7.6)を解くと、その解は $M \geq 2$ において、0 と 1 を除くと

$$p^* = \sqrt{\frac{2}{M^3}} \quad (7.7)$$

である。

次に、初期時に最も多く支持されている選択肢を o_1 とし、 o_1 の支持者の比率と p^* の関係について調べる。もし、その比率が p^* より大きければ、急速に全員がその選択肢を支持する状態に収束する。今、選択肢 o_j の支持者比率を $p(o_j)$ として、選択肢 o_1 が他の選択肢のいずれよりも支持者数が多いならば、次の 2 つの関係が得られる。

$$\sum_{j \in \{1, \dots, M\}} p(o_j) = 1 \quad (7.8)$$

$$p(o_1) > \max(p(o_j)), p(o_j) \geq 0, j \in \{2, \dots, M\} \quad (7.9)$$

ここで、次式が成り立つ。

$$\frac{1}{M} - p^* = \frac{1}{M} - \sqrt{\frac{2}{M^3}} \quad (7.10)$$

$M \geq 2$ なので,

$$\frac{1}{M} \left(\frac{\sqrt{M} - \sqrt{2}}{\sqrt{M}} \right) \geq 0 \quad (7.11)$$

となり,

$$\frac{1}{M} \geq p^* \quad (7.12)$$

が成り立つ。従って,

$$\begin{aligned} p(o_1) - p^* &\geq p(o_1) - \frac{1}{M} \\ &= p(o_1) - \frac{\sum_{j \in \{1, \dots, M\}} p(o_j)}{M} \\ &= p(o_1) \frac{M-1}{M} - \frac{\sum_{j \in \{2, \dots, M\}} p(o_j)}{M} \\ &= \frac{1}{M} \left((p(o_1) - p(o_2)) + \dots + (p(o_1) - p(o_M)) \right) \\ &> 0 \end{aligned} \quad (7.13)$$

以上から、最も支持率の大きい選択肢 o_1 を着目してその支持率が p^* より大きいので必ず群れが o_1 に収束することが証明できた。

7.4.2 投票フェーズに要する時間の概算

次に、投票フェーズに要する時間について数学的に概算を行う。事前の実験によって図 7.5 に示すように、Q-BRT モデルは o_1 の支持率は合意に至るまで、支持者数の変化がある状態と変化がない状態と 2 種類に分類できる。そこで、それぞれを概算し投票フェーズに要する時間を見積もり、その特性を検証する。まず、支持者数の変化がある状態の所要時間を概算する。

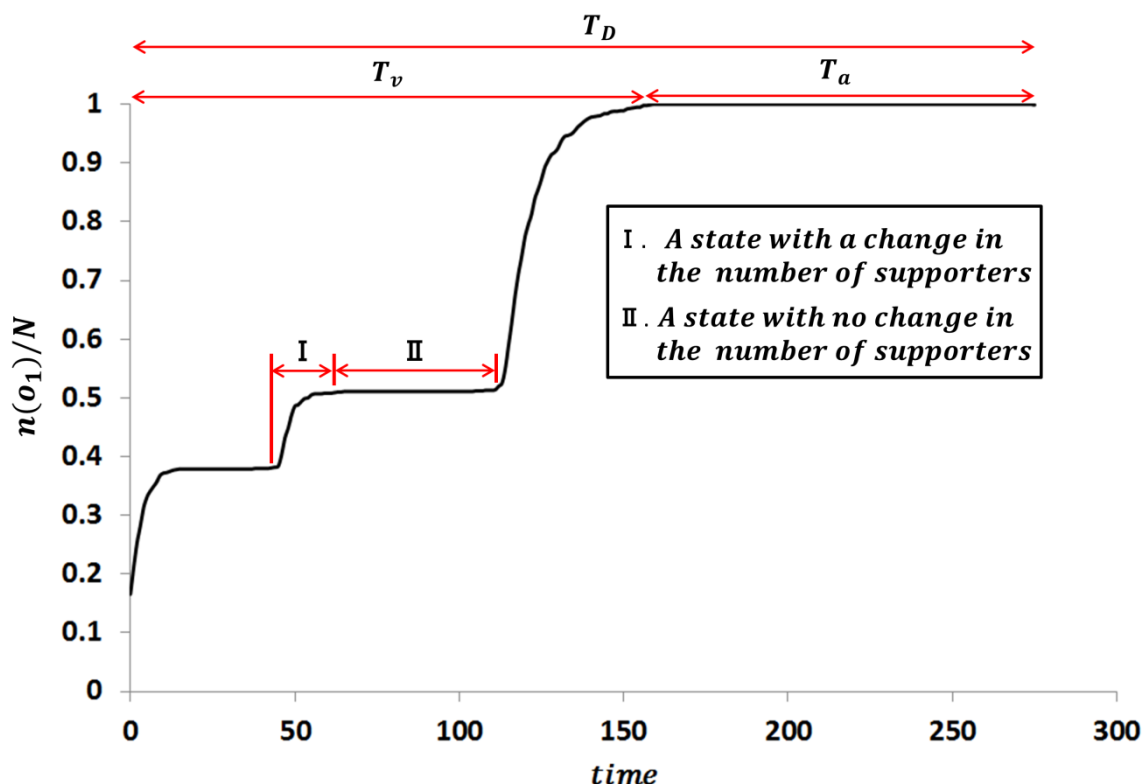


図 7.5 合意状態に至るまで最も支持される選択肢 o_1 の支持率の遷移。支持者数の変化がある状態 I と変化がない状態 II。

選択肢 o_1 が優勢な時、その他の選択肢はほぼ支持されない。そのため、非優勢の選択肢を支持しているエージェントは毎ステップごと支持するものを変更する。 $s[t]$ は時刻 t に選択肢 o_1 の支持者数とする。 o_1 の支持が増える率は $(M - 1)$ 個の選択肢から選択肢 o_1 を選ぶ確率 $1/(M - 1)$ に相当する。ランダムな初期状態から合意するために残りのエージェント全員 $(N - s[t])$ が最終的には選択肢 o_1 を選ぶ必要である。初期状態で選択肢 o_1 の平均支持率は調和級数の最大値 $H[M]/M$ である^[90]。よって、時刻 $t + 1$ ステップ後、選択肢 o_1 の支持者数は

$$s[t+1] = \frac{N-s[t]}{M-1} + s[t] \quad (7.14)$$

ここで,

$$s[0] = \frac{N \cdot H[M]}{M} \quad (7.15)$$

である.

この漸化式を解くと

$$s[t] = \frac{N \cdot (M - M \left(\frac{M-2}{M-1}\right)^t + H[M] \cdot \left(\frac{M-2}{M-1}\right)^t)}{M} \quad (7.16)$$

になる.

選択肢 o_1 を支持しているエージェントの比率が v になる時間 \hat{T}_1 を求めると

$$\hat{T}_1 = \frac{\log \frac{M - M \cdot v}{M - H[M]}}{\log \frac{2 - M}{1 - M}} \quad (7.17)$$

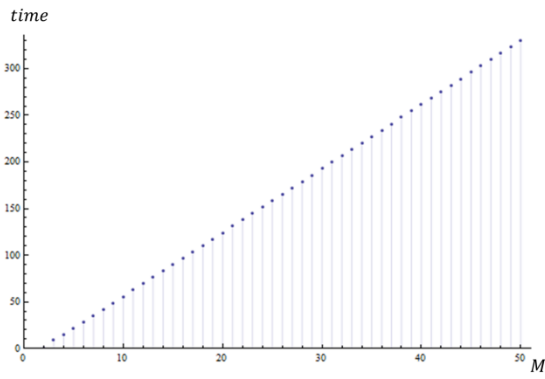
になる.

図 7.6(a)に $v = 0.999$ として選択肢の数 M に対する支持者数の変化がある状態にかかる時間 \hat{T}_1 を図示した. この図に示すように, M の増加に対してほぼ線形で \hat{T}_1 が伸びる事が予測できる.

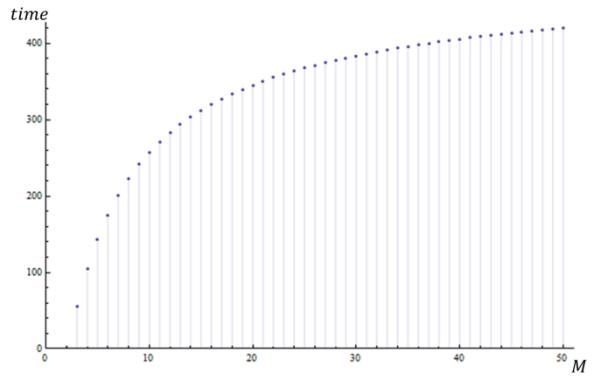
次に, 支持者数の変化がない状態に要する時間を概算する. 前述したようにエージェントが M 個の選択肢からランダムに自身の選択肢を選ぶ時, 各選択肢が占める比率は調和級数に従う. 例えば, 最も少ない選択肢の平均支持比率は $1/M^2$ で, 2番目に少ないものは $\frac{1}{M} \left(\frac{1}{M} + \frac{1}{M-1}\right)$ である. 図 7.7に示すように, 今, o_M を支持するエージェントがいなくなった時, これまで o_M を支持していたエージェントは他の $(M-1)$ 個の選択肢に均等に割当たられると考えて, 他の選択肢の支持率は一様に x_1 増える. よって,

$$x_1 = \frac{1}{M^2(M-1)} \quad (7.18)$$

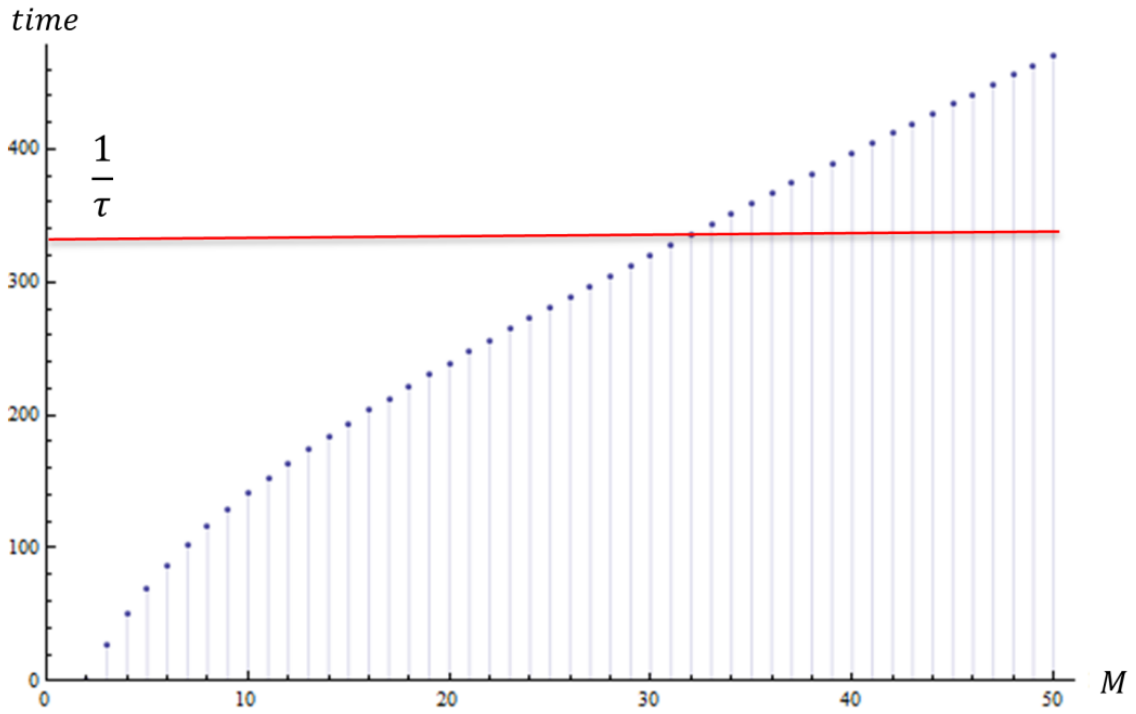
となる. 図 7.7に示すように, x_1 だけ支持者が一律増加するので支持者数の変化が起こらない x_1/τ ステップの時間が発生する. 同様に以後の k 番目についても他の選択肢が増加する率 x_k を求める事ができる. 表 7.2にその結果を示す.



(a) \hat{T}_1



(b) \hat{T}_2



(c) \hat{T}_v

図7.6 選択肢の数 M に対する支持者数の変化がある状態にかかる時間 \hat{T}_1

(a),

支持者数の変化がない状態にかかる時間 \hat{T}_2 (b)

と投票フェーズに要する時間 \hat{T}_v の一例(c). $v = 0.999$, $\tau = 0.001$.

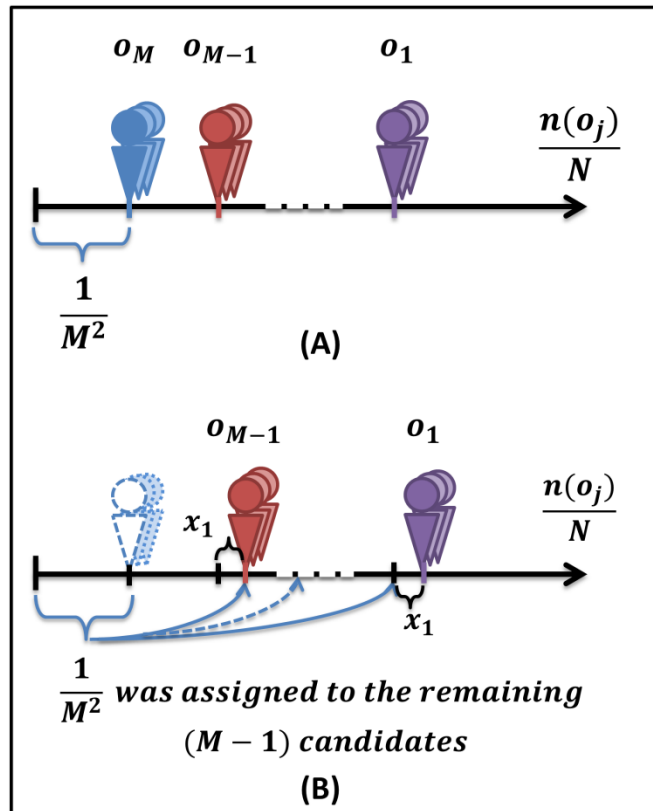


図 7.7 支持者数の変化が起こらない過程.

表 7.2 支持者数の変化が起こらない時間 \hat{T}_2 の概算

昇順	選択肢	支持者数の再分配量	\hat{T}_2
1 st	o_M	$\frac{1}{M^2}$	$\frac{x_1}{\tau} = \frac{1}{M^2(M-1)\tau}$
2 nd	o_{M-1}	$\frac{1}{M} \left(\frac{1}{M} + \frac{1}{M-1} \right)$	$\frac{x_2}{\tau} = \frac{2}{M(M-2)(M-1)\tau}$
3 rd	o_{M-2}	$\frac{1}{M} \left(\frac{1}{M} + \frac{1}{M-1} + \frac{1}{M-2} \right)$	$\frac{x_3}{\tau} = \frac{3}{M(M-3)(M-2)\tau}$
k^{th}	o_{M-k+1}	$\frac{1}{M} \sum_{j=1}^k \frac{1}{M-j+1}$	$\frac{x_k}{\tau} = \frac{k}{M(M-k)(M-k+1)\tau}$

従って、支持者数の変化が起こらない状態 II にかかる時間 \hat{T}_2 は、

$$\hat{T}_2 = \frac{1}{\tau} \sum_{j=1}^{M-2} x_j = \frac{1}{\tau} \sum_{j=1}^{M-2} \frac{j}{M(M-j)(M-j+1)} \quad (7.19)$$

である．図 7.6(b)に $\tau = 0.001$ の時の \hat{T}_2 を示す．

以上から，投票フェーズに要する時間 \hat{T}_v は以下の式で求まる．

$$\begin{aligned} \hat{T}_v &= \hat{T}_1 + \hat{T}_2 \\ &= \frac{\log \frac{M - M \cdot v}{M - H[M]}}{\log \frac{2 - M}{1 - M}} + \frac{1}{\tau} \sum_{j=1}^{M-2} \frac{j}{M(M-j)(M-j+1)} \end{aligned} \quad (7.20)$$

図 7.6(c)に $v = 0.999$, $\tau = 0.001$ の時の \hat{T}_v を示す．

式(7.19)によると，投票フェーズに要する時間は，群れのサイズ N に依存しない事がわかる．また，図 7.6(c)によると，投票フェーズに要する時間 \hat{T}_v は，選択肢の数 M にほぼ線形的に比例している事がわかる．

次に，計算機実験を使用して，以上の概算を検証する．実験の設定は以下の通りである． $N = 1000$, $\tau = 0.001$ とし，選択肢の数 M を変えて投票フェーズに要する時間を測定した．それぞれの M に対して 1000 回の試行を実装した．図 7.8(A)にその結果を示す．横軸は選択肢の数 M で縦軸は投票フェーズに要する時間である．赤い線は $T_v = \frac{1}{\tau}$ で，青い△と点線マークはそれぞれ実験結果の平均値と予測値を表している．予測値はほぼ実験結果と一致している事がわかった．一方，予測値が赤い線の $T_v = \frac{1}{\tau}$ を超えるあたりで実験値が急速に伸びる事がみられた．その原因として，式(3.5)から

$$1 \geq \theta_i + \tau \cdot T_v \quad (7.21)$$

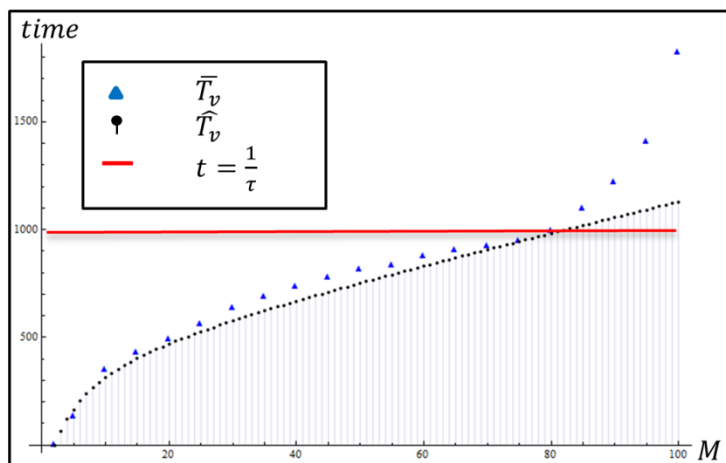
ならば，最も多い支持者数の割合が 1(合意状態)になるまで増加し続ける．すなわち，

$$T_v \leq \frac{1 - \theta_i}{\tau} \quad (7.22)$$

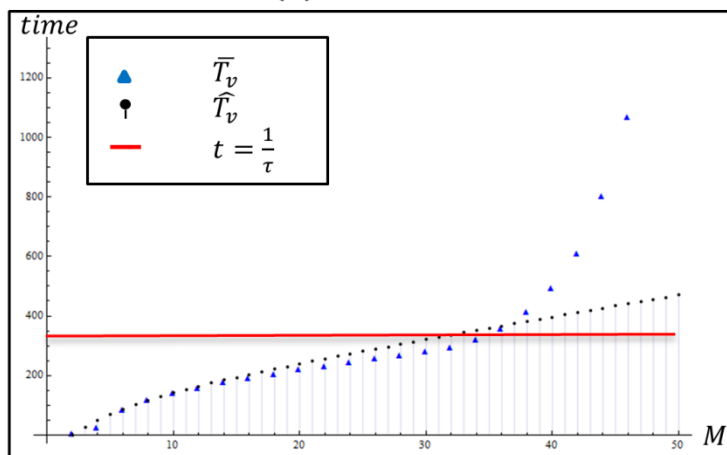
$$T_v < \frac{1}{\tau} \quad (7.23)$$

は，ランダムに選択している状態から集団が合意状態に達する条件である．予測値 \hat{T}_v が $\frac{1}{\tau}$ が超える M と τ を設定すると最も多い支持者数の割合が増加するが，合意状態になることなく途中で減少してしまう．図 7.9 にその現象が起こった $M = 34$, $N = 1000$, $\tau = 0.003$ の場合における各選択肢の支持者数の遷移

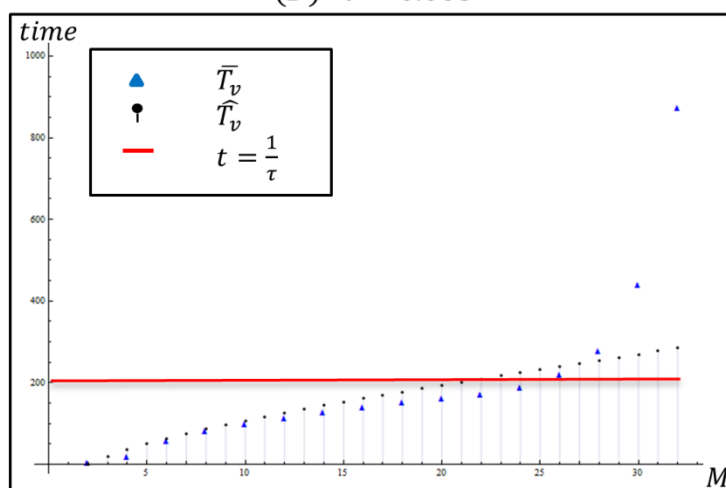
の一例を示す。最も多い支持者数は、時間経過に増加するが、約 320 ステップにその支持者数が合意状態に至らなく急激に減少した。



(A) $\tau = 0.001$



(B) $\tau = 0.003$



(C) $\tau = 0.005$

図7.8 式(7.20)に基づく投票フェイズに要する時間の予測値 \hat{T}_v と

その検証実験の結果(1000回の試行の平均値). $N = 1000$, $\nu = 0.999$.

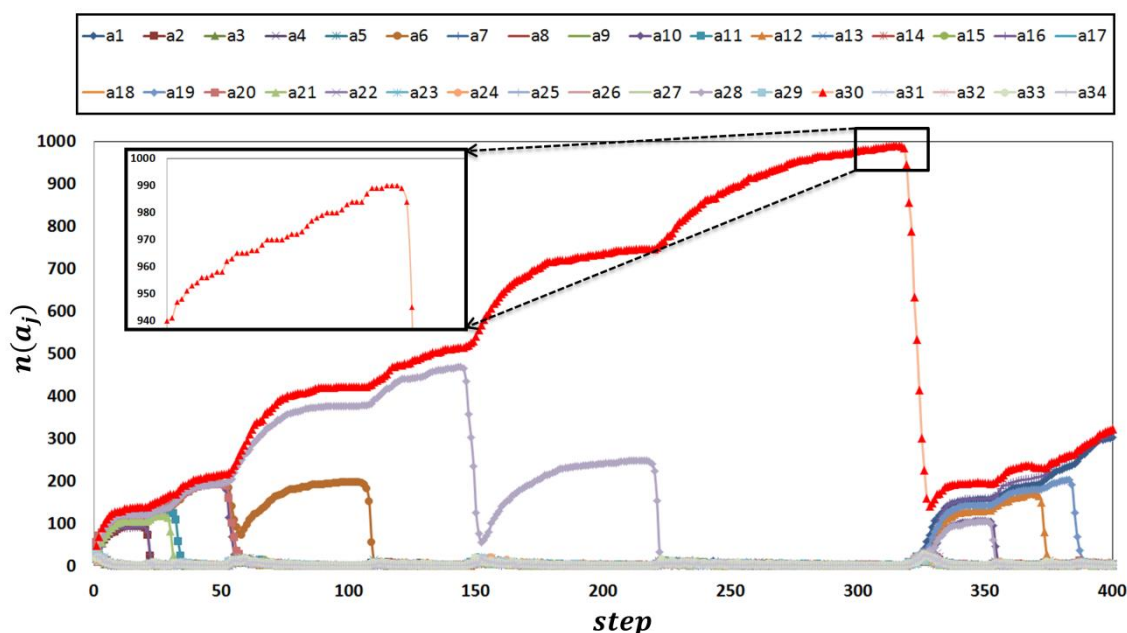


図 7.9 各選択肢の支持者数の遷移の一例. $M = 34$, $N = 1000$, $\tau = 0.003$.

次に, 異なる τ で同じ実験を行う. 以上の実験より大きい $\tau = 0.003$, 0.005 と替えて設定し, それぞれ 1000 試行を実行した. 図 7.8(B)と図 7.8(C)にその結果を示す. 図 7.8(A)と同様, 予測値はほぼ実験結果と一致し, 赤い線の $T_\nu = \frac{1}{\tau}$ 超えるあたりで実験値が急速に伸びる事がみられた. 予測値はほぼ実験結果と一致しているのが $\tau = 0.001$ の実験では M が 85 までで, $\tau = 0.005$ の実験では M が 26 までである. すなわち, 小さい τ の方がより多数の選択肢が扱えるようになる.

以上より, 投票フェーズに要する時間が, 群れのサイズ N に依存せず, 選択肢の数 M に線形的に比例する事がわかった. また, 群れは合意する事ができる条件として, 投票フェーズに要する時間の予測値が $\frac{1}{\tau}$ より小さい値になるような M と τ の組合せを適切に設定することである. そこで, 以下ではこの条件が満たされる M と τ を議論の対象とする.

7.4.3 適した選択肢の発見に要する平均時間の概算

次に、適した選択肢を発見する平均時間 \bar{T}_D について検討する.

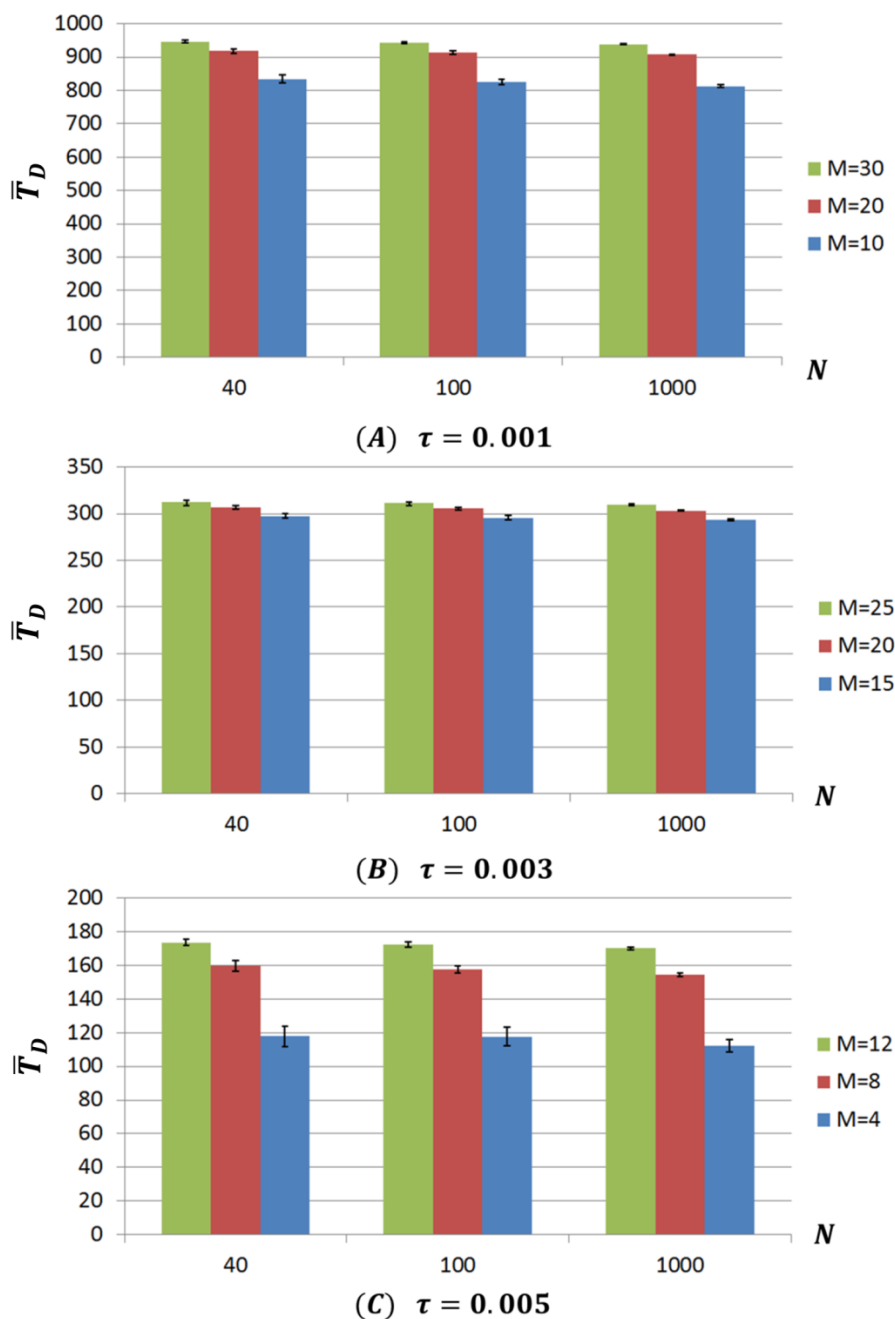


図7.10 1回の試行に要する平均時間 \bar{T}_D とその標準偏差.

前述したように、適した選択肢が見つかるまで、複数の試行(投票と合意)が必要である. よって、1回の試行に要する平均時間 \bar{T}_D は、発見に要する平均時間

を概算するために重要である．そこで，まず，1回の試行に要する平均時間 \bar{T}_D について調べる．選択肢の数 M ，群れのサイズ N と τ を色々変えて1回の試行に要する時間を測定した．それぞれのパラメータの組合せに対して，10000回を実行した．図 7.10 にその結果を示す．縦軸は \bar{T}_D ，横軸が群れのサイズ N で，ひげは標準偏差を表している．これらの結果より，どのパラメータの組合せでも標準偏差がすごく小さいことから，1回の試行に要する時間はほぼ一定な値である事がわかった．従って，以下では，1回の試行に要する時間 T_D がある定数 T'_D であると仮定する．

表 7.3 適した選択肢の発見率と要する時間の推定値

合意	候補の数	発見率	要する時間
1 st	M	$\frac{1}{M}$	T'_D
2 nd	$M - 1$	$(1 - \frac{1}{M}) \frac{1}{M - 1}$	$2T'_D$
3 rd	$M - 1$	$(1 - \frac{1}{M})(1 - \frac{1}{M - 1}) \frac{1}{M - 1}$	$3T'_D$
l^{th}	$M - 1$	$(1 - \frac{1}{M})(1 - \frac{1}{M - 1})^{l-2} \frac{1}{M - 1}$	$l \cdot T'_D$

まず，平均何回合意すれば適した選択肢を発見できるかを概算する． M 個の候補の1つ a_x ($x \in \{1, \dots, M\}$) が適した選択肢 a_{goal} とする ($a_x = a_{goal}$)．一度も合意していないとき，候補数は M のままである．この中1つの試行で合意したとき，それが a_x である確率は $1/M$ である．2回目の合意以降から，個体は1つの選択肢が a_x ではない事を知っている．また，次に合意する選択肢はランダムに決定される．これらの仮定の下では，表 7.3 に適した選択肢の発見率と要する時間の推定値を示す．これより，発見に要する平均時間の推定値 \hat{T} は

$$\begin{aligned} \hat{T} &= T'_D \left(\frac{1}{M} + \sum_{l=2}^{\infty} \left(1 - \frac{1}{M}\right) \left(1 - \frac{1}{M-1}\right)^{l-2} \frac{l}{M-1} \right) \\ &= T'_D \left(\frac{1}{M} + M - 1 \right) \end{aligned} \quad (7.24)$$

である．ここで，

$$\bar{E} = \frac{1}{M} + M - 1 \quad (7.25)$$

が発見に要する平均試行である．従って，適した選択肢の発見に要する平均試行

は、選択肢の数に対してほぼ線形であることを推定できる。

次に、1回の試行に要する時間 T_D について数学的に解析する。 T_D は初期状態から合意状態が終了するまでの時間であるため、式(3.5)が初めて満たされなくなったときの $(t - t_{i,last}(t))$ が \widehat{T}_D に相当する。すなわち、 \widehat{T}_D は、

$$\theta_i + \tau \cdot \widehat{T}_D = 1 \quad (7.26)$$

の方程式の解である。また、最初に式(3.5)が満たされなくなったエージェントは集団のなか最も大きいバイアス値 θ_{max} を有するものである。つまり、 \widehat{T}_D は、

$$\widehat{T}_D = \frac{1 - \theta_{max}}{\tau} \quad (7.27)$$

である。ここで、 $\theta_{max} = \max(\theta_i)$ である。

群れのサイズ N がある程度大きいと

$$\theta_{max} \approx \frac{2}{M} \quad (7.28)$$

のため、

$$\widehat{T}_D \approx \frac{1 - \frac{2}{M}}{\tau} \quad (7.29)$$

となる。式(7.24)に導入すると発見に要する平均時間の推定値 \hat{T} は

$$\hat{T} \approx \frac{1 - \frac{2}{M}}{\tau} \left(\frac{1}{M} + M - 1 \right) \quad (7.30)$$

$$\approx \frac{M + \frac{3}{M} - \frac{2}{M^2} - 3}{\tau}$$

である。

以上から、1回の試行に要する時間がほぼ一定になっており、選択肢の数 M の数が増加すると、発見に要する平均時間は線形的に増加する事がわかった。また、式(7.27)と式(7.28)には群れのサイズ N がない事から1回の試行に要する時間と発見に要する平均時間が、群れのサイズに依存しない事を示した。以下では、これらについての検証実験を行う。

7.4.4 群れサイズのスケラビリティの検証実験

まず、式(7.27)に示した1回の試行に要する時間が、群れのサイズに依存しないことについて検証実験を行い、群れのサイズのスケラビリティの特性を明らかにする。

そこで、群れのサイズ N を色々変えて1回の試行に要する時間 T_D を測定した。ここでは、

$$N = k_2 \cdot M \quad (7.31)$$

とし、条件 $T_D < 1/\tau$ を満たすいくつかの M の値に対して $k_2 = 1, 2, 4, 6, 8, 10, 20, 40, 60, 80, 100$ と色々な値でパラメータを設定した。 $\tau = 0.003$, 制限時間を20000ステップとし、初期時にエージェントはランダムに選択肢を支持するものとする。図7.11に10000回を実施した結果を示す。横軸に k_2 , 縦軸に1回の試行に要する時間 \bar{T}_D を表している。赤点線は、 $T_D = 1/\tau$ を示す。図より $k_2 = 1, 2$ というとても小さい値を除いて、1回の試行に要する平均時間 \bar{T}_D は、 $1/\tau$ より小さい値で、 N の増加に対してほぼ一定である事がわかった。

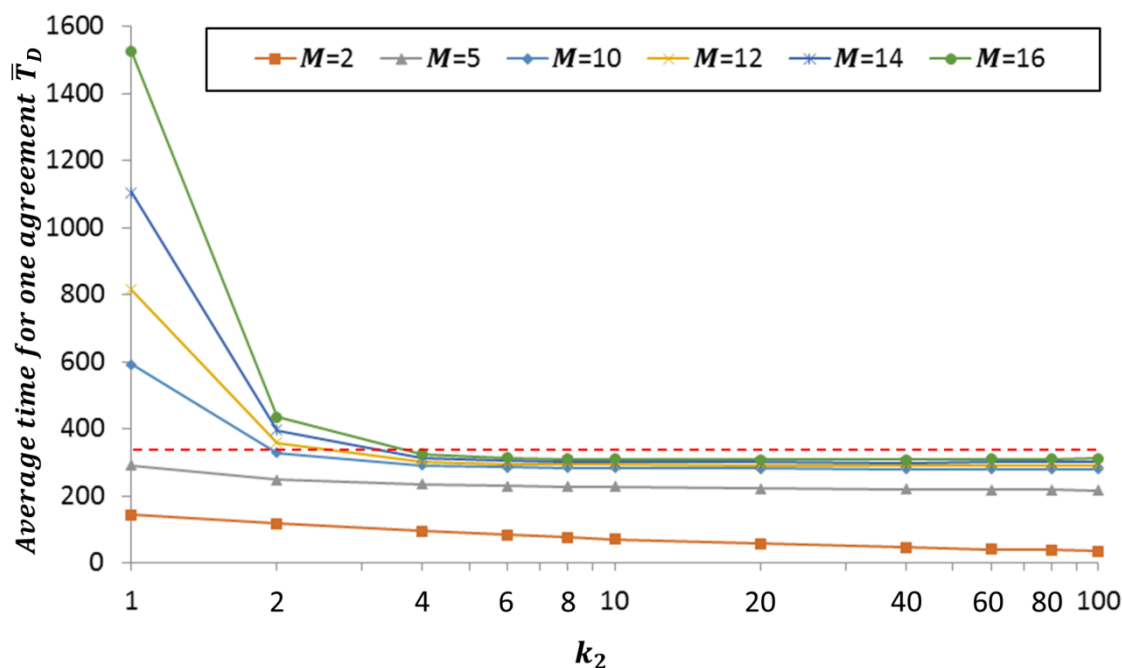


図 7.11 異なる群れサイズにおける1回の試行に要する平均時間 \bar{T}_D .

k_2 が小さいとき ($k_2 = 1, 2$), 1回の試行に要する平均時間がとても長かった。原因としては、 k_2 が小さいと全体的にエージェント数 N が少なく、図3.8に示したようにバイアス値が十分に提案分布の形になってないため、速やかな合意ができなく長い時間がかかったと考えられる。一方、 k_2 がある程度大きい ($k_2 \geq 4$)

とバイアス分布が提案分布と近い形となり、速やかな合意ができ、式(7.27)に示すように1回の試行に要する時間は群れのサイズに依存せず、一定な値となる。

7.4.5 適した選択枝の発見

次に、適した選択枝を発見できる事を計算機実験によって示す。

初期時にはエージェントはランダムに選択枝を選択しているとし、候補の数 M を変えて適した選択枝を発見できるかについて検証している。実験条件は次のとおりである。エージェント数 $N = 100$, $\tau = 0.003$ とする。予め選択枝のなかの1つを選び適した選択枝 a_{goal} と設定する。 a_{goal} を全員が選択することができたケースをカウントし、 M ごとにそれぞれ 10000 試行を実施し、その発見確率と発見に要する時間を求めた。図 7.12 と図 7.13 にその結果を示す。

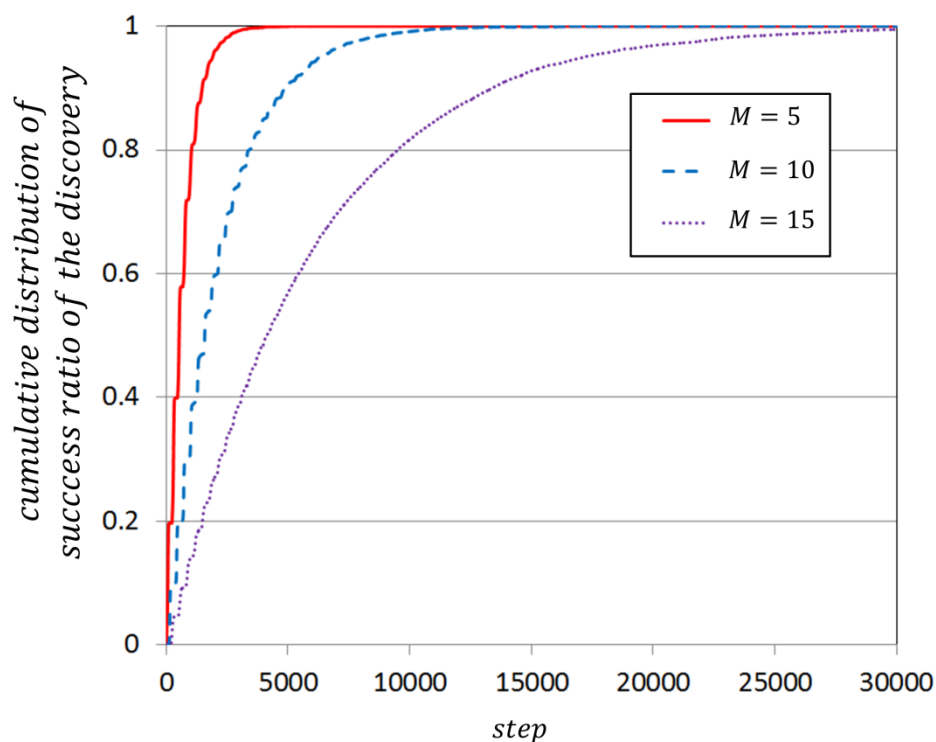


図 7.12 発見率の累積確率密度分布。

図 7.12 は、3つの異なる M における発見率を累積確率密度分布で表したものである。横軸は発見に必要な時間を表し、縦軸は成功率の累積分布を表す。異なる M のすべてのケースで、発見の成功率は非常に高い 99%以上であった。したがって、 M が大きい場合でも、群れは適した選択枝を発見できる事がわかった。

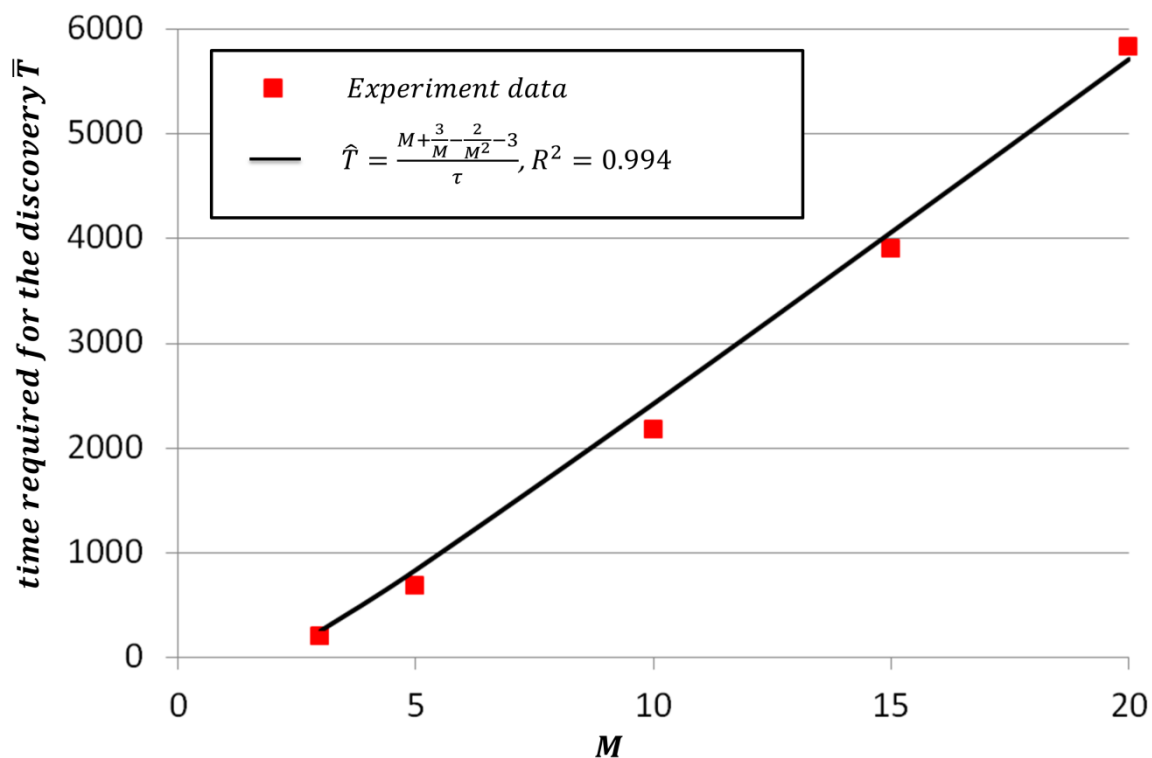


図 7.13 適した選択肢の発見に要する平均時間とその推定値.

図 7.13 に発見に適した選択肢の発見に要する平均時間と式(7.30)の推定値を示す. 横軸は選択肢の数 M を表し, 縦軸は平均時間である. 黒い線は, 式(7.30)を使用した, 検出に必要な平均時間 \hat{T} と M の関係の回帰分析の結果である. この図から, 式(7.30)による推定が適切である事がわかる. また, 決定係数 R^2 が 0.994 (≈ 1) であることから, 残差が非常に小さく, 実験データが推定データとほぼ同じであることを意味する. したがって, 式(7.30)を使用して, 適した選択肢の発見に必要な平均時間を推定できる事がわかった. また, 選択肢の数 M にほぼ比例している事もわかった.

以上より, 提案手法 Q-BRT モデルでは選択肢が沢山存在する時でも本節で示した方法でパラメータを適切に設定すれば群れは高確率で適した選択肢を発見する事が出来る事がわかった. またその場合, 発見までの時間はほぼ線形に増加すると予測できる. 以上から, Q-BRT モデルは高い実用性を有すると共に複雑な問題 ($M \gg 2$) の解決を可能にすると考えられる.

7.5 まとめ

BRT モデルでは、ガウシアン分布に基づいた鐘型曲線を使用して個人属性を表すバイアス値を生成したが、収束について確率論から離れて議論する事が難しかった。そこで、本章では、解析のしやすい2次関数を用いたバイアス分布を使用するQ-BRTモデルを提案した。数学的な分析と計算機実験により、任意の選択肢に合意する事や、適した選択肢の発見に要する平均時間の概算などQ-BRTモデルの特性を明らかにした。

第8章 Q-BRT モデルその①：

Plural voting からヒントを得た集団探索の高速化

8.1 はじめに

前章では、ガウシアン分布の代わりに数学的な解析のしやすい2次関数を用いたバイアス分布を使用する Q-BRT モデルを提案した。この Q-BRT モデルは、多数の選択肢でも非常に高い精度で適した選択肢を発見することができるモデルである。しかし、計算機実験では、適した選択肢を発見するのに長い時間がかかる場合があることを示した。図 8.1 は、Q-BRT モデルにおける適した選択肢の発見に要する時間をヒストグラムのグラフで表したものである。横軸は時間、縦軸がその時間で発見できたケースの数を表している。グラフは右肩下がりしている事からほとんどのケースにおいて 8000 ステップという短い時間で適した選択肢を発見できた事がわかる。その一方、時間が 20000 ステップ以上かかったケースが少なからずみられ、39200 ステップ以上かかったケースもあった。その原因は、エージェントは別の選択肢に切り替えるときに、確率的に選んでいるので、何度選んでも適した選択肢が選ばれない為と考えられる。

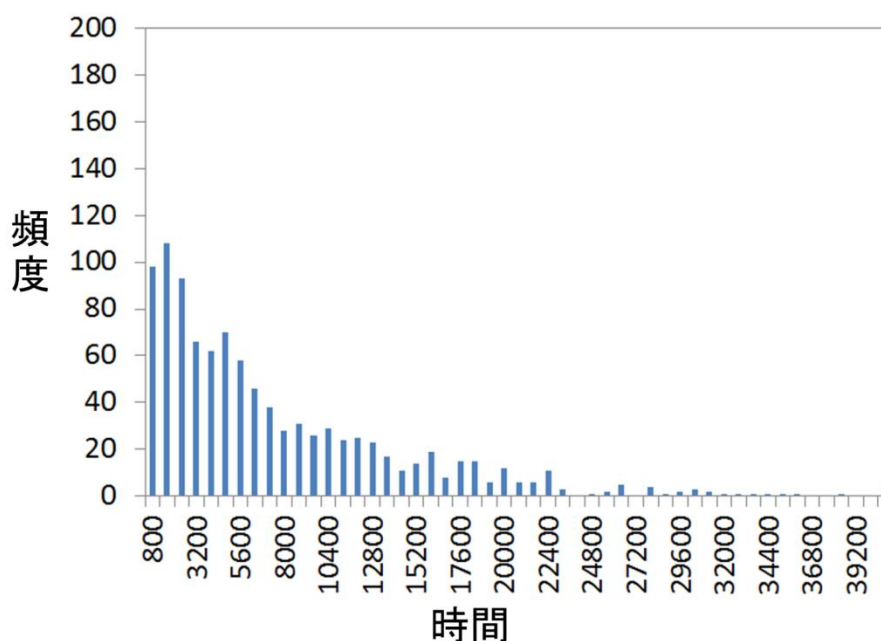


図 8.1 Q-BRT モデルにおける適した選択肢の発見に要する時間のヒストグラム。

第4章では、失敗した経験を記録するタブーリストを持つエージェントを導入して BRT モデルの探索効率を改善した。これは、静的な環境において非常に効果的な手法であるが、適した選択肢が変化する動的な環境では適した選択肢を切り捨てる恐れがある。

そこで、本章では、Plural voting^[19]からヒントを得て、BRT モデルの単一の投票システムを複数の投票システムに拡張し、動的な環境でも対応できる集団探索の高速化を試みる。Plural voting (multiple voting と呼ばれる)とは、1 人の人が一回の選挙で複数回投票できるようにすることである。毎時刻に多数の選択肢の中から 1 つの選択肢にしか投票しない従来の BRT モデルより、複数に投票する方が適した選択肢の発見率を高める事ができる為、発見に要する時間を短縮する事が期待できる。

このような投票システム(選挙システム)からヒントを得た、群ロボットや社会科学などの分野における研究には次のものがある。最も単純なメカニズムの 1 つとしてよく知られているのは、投票者モデル(voter model)である。これは、特に正確に集団的意思決定に導くものの時間はかかる^[36]。他に、投票者モデルより速く合意に導く一般的なモデルとして多数決ルール(majority rule)があるが、投票者モデルより精度がやや劣っている^[32,44]。近年、k-unanimity rule が Scheidler らによって提案され^[67]、そのパフォーマンスは、投票者モデルと多数決ルールモデルの間にある^[82]。そこで、本章では、the best-of-n 問題に対し、群れが高い精度で適した選択肢を迅速に発見できる意思決定アルゴリズムを提供する。

8.2 提案手法

以上を踏まえて、Plural voting からヒントを得て k の複数投票を用いた BRT モデルを提案する。以下では、 k -BRT モデルと呼ぶ。

8.2.1 アルゴリズム

BRT モデルと同様、 N 台のエージェントで構成される集団 $G = \{P_i: i = 1, \dots, N\}$ において、 M 個の選択肢の中から適した選択肢を 1 つ選ぶ意思決定が行われているとする。エージェントの選択肢集合を $A = \{a_j: j = 1, \dots, M\} (M \geq 2)$ とし、探すべき適した選択肢を a_{goal} とする ($a_{goal} \in A$)。エージェントは事前に a_{goal} を知らず、全員が合意して初めて合意した選択肢が a_{goal} であるか分かるものとする。

エージェント P_i は、毎時刻に k 回の投票をすると仮定する ($M > k > 1$)。 k 回の投票において、 M 個の候補の中から k 個の候補に投票する (選択する)。そして、 P_i が M 個の選択肢に対して $B_i = \{\theta_{i1}, \dots, \theta_{ij}, \dots, \theta_{iM}\}$ をもつとする ($0 < \theta_{ij} < 1$)。時刻 t において、 P_i が l 回目の投票で $A_{il}(t) \in A$ の選択肢を支持し ($1 \leq l \leq k$)、その時の意思決定にバイアス値 $\vartheta_{il} \in B_i$ を使用する。

時刻 t において、エージェント P_i は次の時刻の l 回目の投票で選ぶ選択肢 $A_{il}(t+1)$ を次のように決定する。もし

$$\frac{n(A_{il}(t))}{N} \geq \vartheta_{il} + \tau \cdot c_{il}(t) \cdot (t - t_{il,last}(t)) \quad (8.1)$$

であれば、現在の選択肢 $A_{il}(t)$ を引き続き選択し、 $A_{il}(t+1) = A_{il}(t)$ となる。そうでなければ現在の選択肢 $A_{il}(t)$ をやめてそれ以外の選択肢を確率的に選ぶ ($A_{il}(t+1) = A - A_{il}(t)$)。

τ は支持者数の増加量を表す定数である。 $t_{il,last}(t)$ は l 回目の投票でエージェント P_i が最後に選択肢を変更した時刻で、 $(t - t_{il,last}(t))$ で同一選択肢を選び続けている時間に相当し、以下の式で求める。

$$t_{il,last}(t) = \begin{cases} t & A_{il}(t) \neq A_{il}(t-1) \\ t_{il,last}(t-1) & otherwise \end{cases} \quad (8.2)$$

$c_{il}(t)$ は選択肢が“適している”時 0 である関数である。すなわち

$$c_{il}(t) = \begin{cases} 0 & \exists l, \forall i, A_{il}(t) \equiv a_{goal} \\ 1 & otherwise \end{cases} \quad (8.3)$$

k 回の投票のいずれかで、群れ全体が選択肢 a_{goal} に投票すると、 $c_{il}(t) = 0$ となり、意思決定プロセスが成功と見なし終了する。

8.2.2 バイアス分布

本章では、第7章で提案した二次関数を使用し、それぞれの選択肢に対して以下の様にバイアス θ_{ij} を生成する.

$$n(\theta_{ij}) = \begin{cases} 0 & \theta_{ij} \leq 0 \\ k_1 N \theta_{ij}^2 & 0 < \theta_{ij} \leq \frac{1}{M} \\ k_1 N \left(\theta_{ij} - \frac{2}{M}\right)^2 & \frac{1}{M} < \theta_{ij} < \frac{2}{M} \\ 0 & \frac{2}{M} \leq \theta_{ij} \end{cases} \quad (8.4)$$

ここで、 $n(\theta_{ij})$ とは、 θ_{ij} をとる人数を表し、 $k_1 = 3M^3/2$ は正規化項である.

8.3 意思決定プロセスのダイナミクス

ここでは、提案した k-BRT モデルの挙動を図示しその特性を明らかにする。BRT モデルでは、群れが適さない選択肢に合意するのをやめて別の選択肢に合意する事により、複数の選択肢のなかから、適した選択肢を発見する事ができる。このように、k-BRT モデルでも、合意の形成、合意状態の切り替えと適した選択肢の発見が発生するかについて検討を行う。そのため、適した選択肢 a_{goal} がないことを仮定する。

実験の設定は以下の通りである。 $M = 5$, $N = 100$, $\tau = 0.003$, $k = 2$, $a_{goal} \equiv a_5$ と設定し、初期時でエージェントがランダムな選択肢に投票するとした。図 8.2 と図 8.3 に 2 つの投票過程における各選択肢の支持者数の遷移を示す。横軸が時間で縦軸は各選択肢の支持者数を表している。両方の投票過程において、群れは急速に合意状態に達し、合意した選択肢が適したものと判断でない場合、それをやめて別の選択肢に合意し始めることがわかる。2 回目の投票で発見できなかったが、約 1050 ステップに 1 回目の投票では、適した選択肢 a_5 が発見された。

以上から、k-BRT モデルでは、速やかな合意形成と合意状態の切り替え、そして適した選択肢の発見が確認できたことから k-BRT モデルによる集団探索の高速化が期待できる。

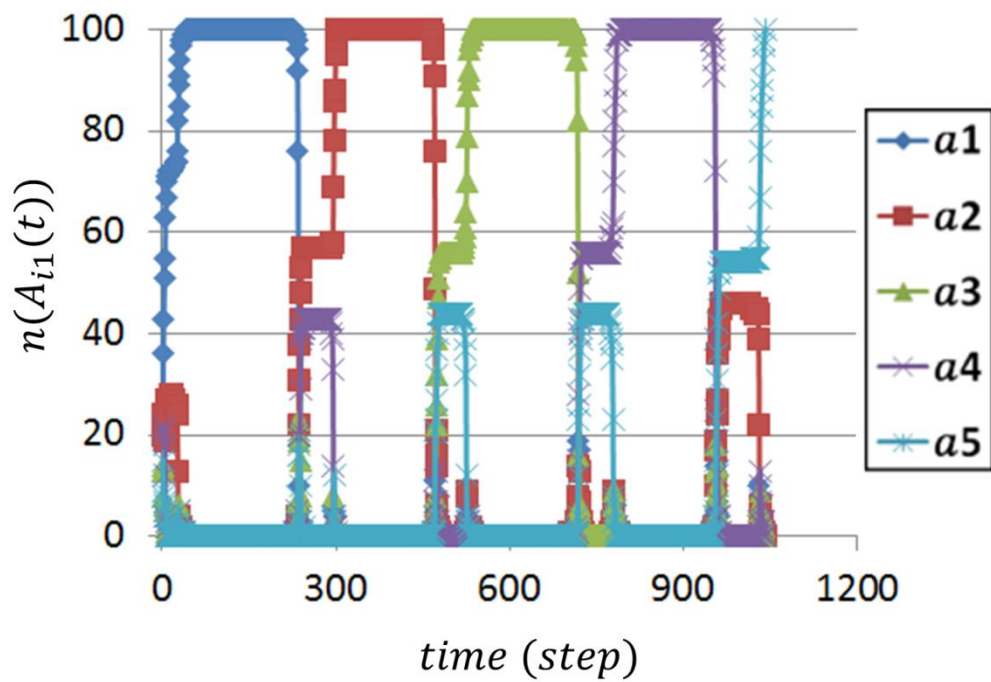


図 8.2 1 回目の投票における各選択肢の支持者数の遷移.
 $M = 5, N = 100, \tau = 0.003, k = 2, a_{goal} \equiv a_5$.

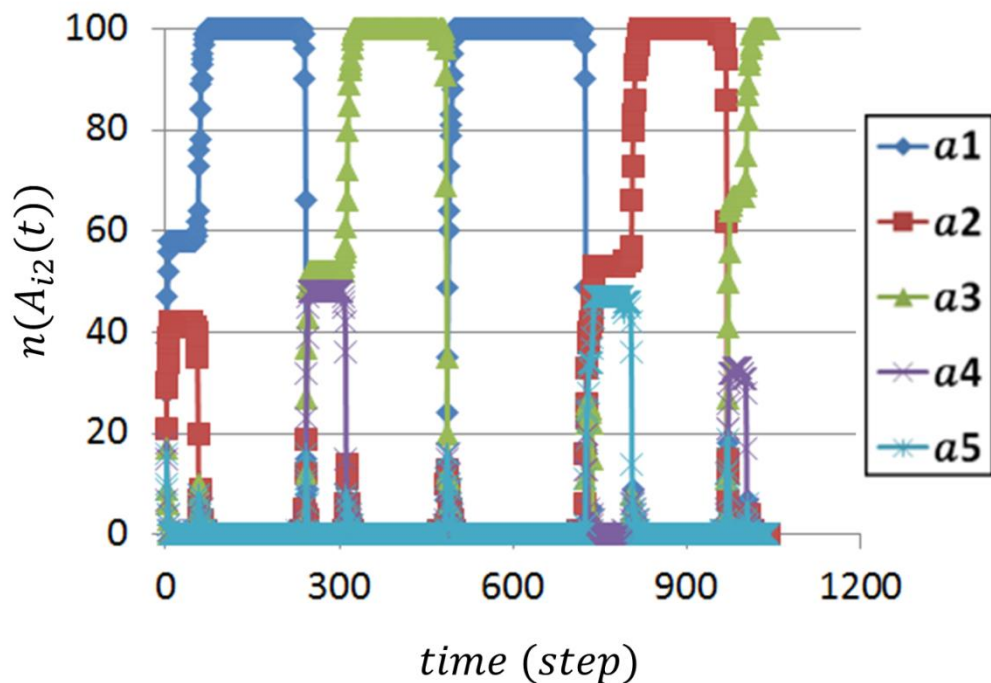


図 8.3 2 回目の投票における各選択肢の支持者数の遷移.
 $k = 2, M = 5, N = 100, \tau = 0.003, a_{goal} \equiv a_5$.

8.4 適した選択肢の発見に要する平均時間の概算

前節では、k-BRT モデルの挙動を図示し、k-BRT モデルによる集団探索の高速化を実現することは大いに期待できる事がわかった。ここでは、数学的な解析を用いて適した選択肢の発見に要する平均時間を概算し、k-BRT による集団探索の特徴を明らかにする。

表8.1 適した選択肢が発見される確率

合意	Q-BRT モデル	k-BRT モデル
1 st	$\frac{1}{M}$	$P(1) = \frac{M^k - (M-1)^k}{M^k}$
2 nd	$\frac{1}{M}(1 - \frac{1}{M})$	$P(1)(1 - P(1))$
3 rd	$\frac{1}{M}(1 - \frac{1}{M})^2$	$P(1)(1 - P(1))^2$
x th	$\frac{1}{M}(1 - \frac{1}{M})^{x-1}$	$P(1)(1 - P(1))^{x-1}$

M 個の候補の1つ a_x ($x \in \{1, \dots, M\}$) が適した選択肢 a_{goal} とする ($a_x = a_{goal}$)。Q-BRTモデルでは、最初に合意するとき、それが a_x である確率は $1/M$ である。それに対して、提案手法では、 $P(1) = (M^k - (M-1)^k)/M^k$ である。表8.1 に両モデルにおける合意時の発見率を示す。従って、 x 回目の合意は、

$$P(x) = P(1)(1 - P(1))^{x-1} \quad (8.5)$$

である。また、

$$\begin{aligned} \ln(P(x)) &= \ln(P(1)(1 - P(1))^{x-1}) \\ &= (x-1)\ln(1 - P(1)) + \ln(P(1)) \\ &= x\ln(1 - P(1)) + \ln(P(1)/(1 - P(1))) \end{aligned} \quad (8.6)$$

と展開することができる。従って、 x 回目の合意は

$$P(x) = \frac{P(1)}{1 - P(1)} e^{x\ln(1 - P(1))} \quad (8.7)$$

である。

図8.4は、合意のときにおける発見率 $P(x)$ の例を示している。発見率 $P(x)$ は x の指数関数のため、一回目の合意では適した選択枝の発見率が最も高く、合意回数を重ねるにつれて、発見率が低下している事がわかる。また、投票回数 k が高いほど、グラフの傾きが大きく、発見率がより高い値からより急激に減少していることから、 k -BRTモデルの方が、少ない合意回数で適した選択枝を発見することが期待できる。

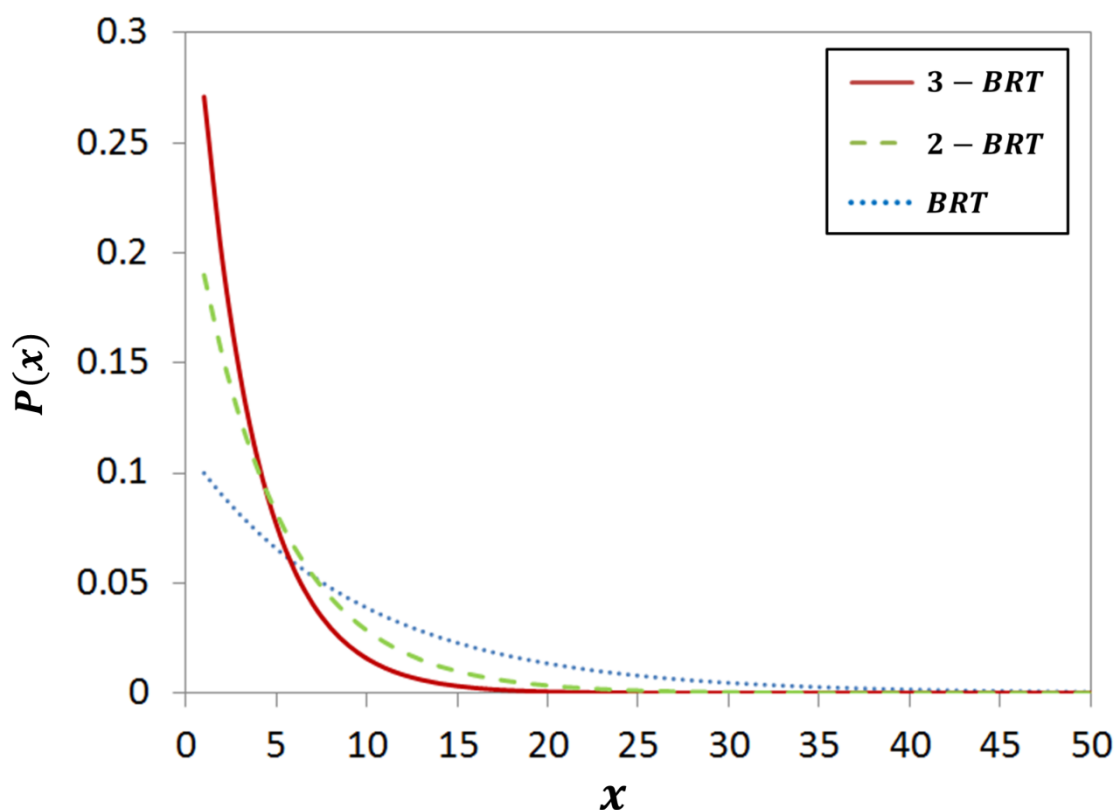


図8.4 合意のときにおける発見率. $M = 10$.

次に、提案手法のパフォーマンスを定量的に評価するために、適した選択枝の発見に要する平均時間 \bar{T} を概算する。第7章で示したように、平均時間 \bar{T} は

$$\bar{T} = \bar{E} \cdot T_D \quad (8.8)$$

である。 T_D は、一回の試行(投票・合意)に要する時間で、

$$T_D = \frac{1 - \frac{2}{M}}{\tau} \quad (8.9)$$

である。

k-BRTモデルにおいて、発見に要する平均試行回数 \bar{E} は、以下のように計算する事ができる。

$$\begin{aligned}\bar{E} &= \sum_{x=1}^{\infty} xP(x) \\ &= \frac{P(1)}{1-P(1)} \sum_{x=1}^{\infty} xe^{x \ln(1-P(1))}\end{aligned}\tag{8.10}$$

また,

$$\sum_{x=1}^{\infty} e^{xb} = \frac{e^b}{1-e^b} = \frac{1}{e^{-b}-1}\tag{8.11}$$

が成り立つため,

$$\frac{d}{db} \sum_{x=1}^{\infty} e^{xb} = \frac{d}{db} (e^{-b}-1)^{-1}\tag{8.12}$$

$$\sum_{x=1}^{\infty} be^{xb} = \frac{e^{-b}}{(e^{-b}-1)^{-2}} = \frac{e^b}{(1-e^b)^2}$$

と展開できる。これを式(8.10)に当てはめると発見に要する平均試行回数 \bar{E} は

$$\begin{aligned}\bar{E} &= \frac{P(1)}{1-P(1)} \cdot \frac{1-P(1)}{P(1)^2} = \frac{1}{P(1)} \\ &= \frac{M^k}{M^k - (M-1)^k}\end{aligned}\tag{8.13}$$

である。

これと式(8.9)を式(8.8)に導入すると適した選択肢の発見に要する平均時間 \bar{T} は

$$\bar{T} = \frac{M^k}{M^k - (M-1)^k} \cdot \frac{1 - \frac{2}{M}}{\tau}\tag{8.14}$$

になる。

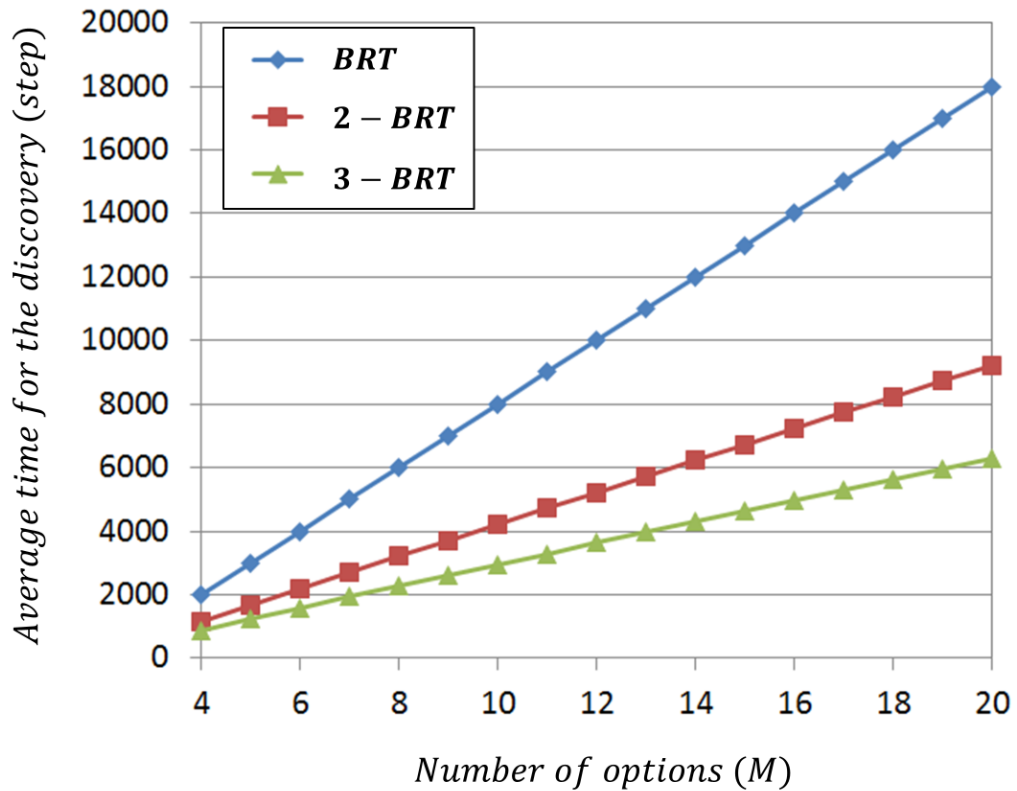


図8.5 式(8.14)に基づいた発見に要する平均時間 \bar{T} の例. $\tau = 0.001$.

図8.5に $\tau = 0.001$ の場合, 式(8.14)に基づいた発見に要する平均時間 \bar{T} を示す. 提案手法は, BRTモデルと同様に, 発見に要する平均時間が選択肢の数に線形的に増加している事がわかる. また, k が大きければ大きいほど, 線形関数の勾配が小さく, 発見に要する時間が劇的に短くなるのがわかる.

以上から, 提案手法は合意回数で適した選択肢の発見率を向上させ, それに要する時間を劇的に短縮する事が期待できる.

8.5 実験

本章では、the best-of- n 問題における探索の高速化を実現するため、Plural voting からヒントを得て k の複数投票を BRT モデルに導入した。前節では、数学的な解析を用いて提案手法による集団探索の高速化の本質や特徴などを明らかにした。以下では、計算機実験を使用し、これらについて検証を行い、集団探索の高速化が実現できることを示す。

8.5.1 探索精度の検証

まず、多数選択枝が存在しているとき、その中から 1 つの適した選択枝を発見できるか、また、その探索精度について検証を行う。

実験の設定は以下の通りである。 $N = 100$, $M = 10$, $\tau = 0.001$, $a_{goal} \equiv a_M$ とし、初期時においてエージェントがランダムに選択枝を選択しているとする。実装する対象として、 $k = 2$, $k = 3$ の提案手法 (2-BRT モデル, 3-BRT モデル) と従来の BRT モデル、と 3 つのモデルを用意する。それぞれのモデルに対し、1000 回を実行し、適した選択枝が発見された時間を記録する。図 8.6 にその結果を示す。

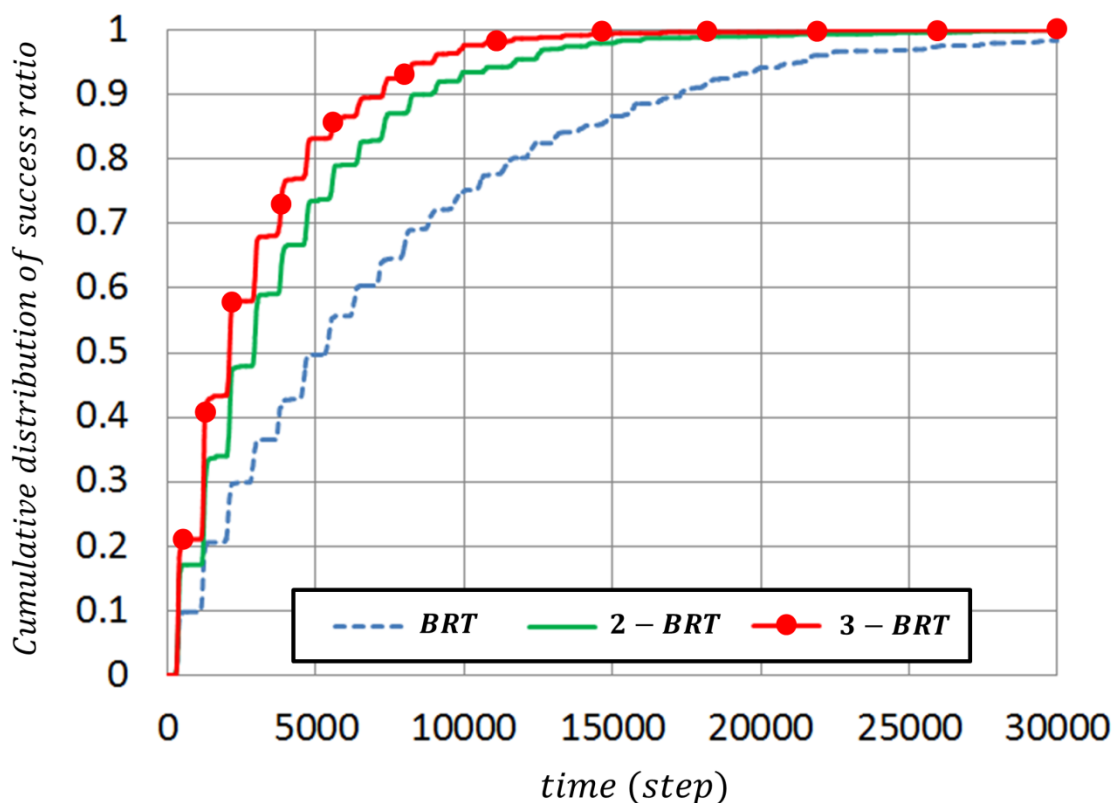


図8.6 探索時間の累積確率密度分布. $N = 100$, $M = 10$, $\tau = 0.001$.

図 8.6 に探索時間を累積確率密度分布で表したものである。横軸に要した時間、縦軸にこの時間までに適した選択肢を発見する確率を表している。全体的に、30000 ステップまでは、すべてのモデルが 99%以上という非常に高い発見率を示している。また、点線で従来の BRT モデルを表していて、十分に長い時間を与えれば発見率が非常に高いが、短い時間が必要な場合ではやや劣っている。それに比べて、提案手法である 2-BRT モデル(平滑線)と 3-BRT モデル(○マーカ付きの線)は、短い時間でも遥かに高い発見率をしめしていることがわかる。例えば、5000 ステップ以内の発見率は、BRT モデルでは約 50%しかないのに対して、2-BRT モデルと 3-BRT モデルでは、それぞれ約 74%と 83%である。

以上から、提案手法は従来の BRT モデルの探索精度を大きく改善できる事がわかった。

8.5.2 探索時間の検証

次に、適した選択肢の発見に要する時間を着目し、提案手法による探索の高速化に関する特徴を明らかにする。

まず、図 8.7 は、前サブセクションの実験で得られた探索時間をヒストグラムで表したものである。縦軸は頻度、横軸が時間を表している。青棒に実験結果、平滑線に式(8.7)を使用した推定値を示す。全てのモデルにおいて、この推定値は、ほぼ実験結果と一致していることから発見頻度が探索時間の指数関数に従っていることがわかる。また、ヒストグラムのグラフが右に歪んでいるが、従来の BRT モデルより提案手法の方が歪度が大きいいため、より少ない時間で適した選択肢を発見できることが確認できた。そして、時間が 20000 ステップ以上かかったケースが少なからずみられた BRT モデルに対して、3-BRT モデルでは 10000 以上かかったケースがほぼない。さらに、提案手法で k を大きくすることにより集団探索がより高速になっていることも確認できた。

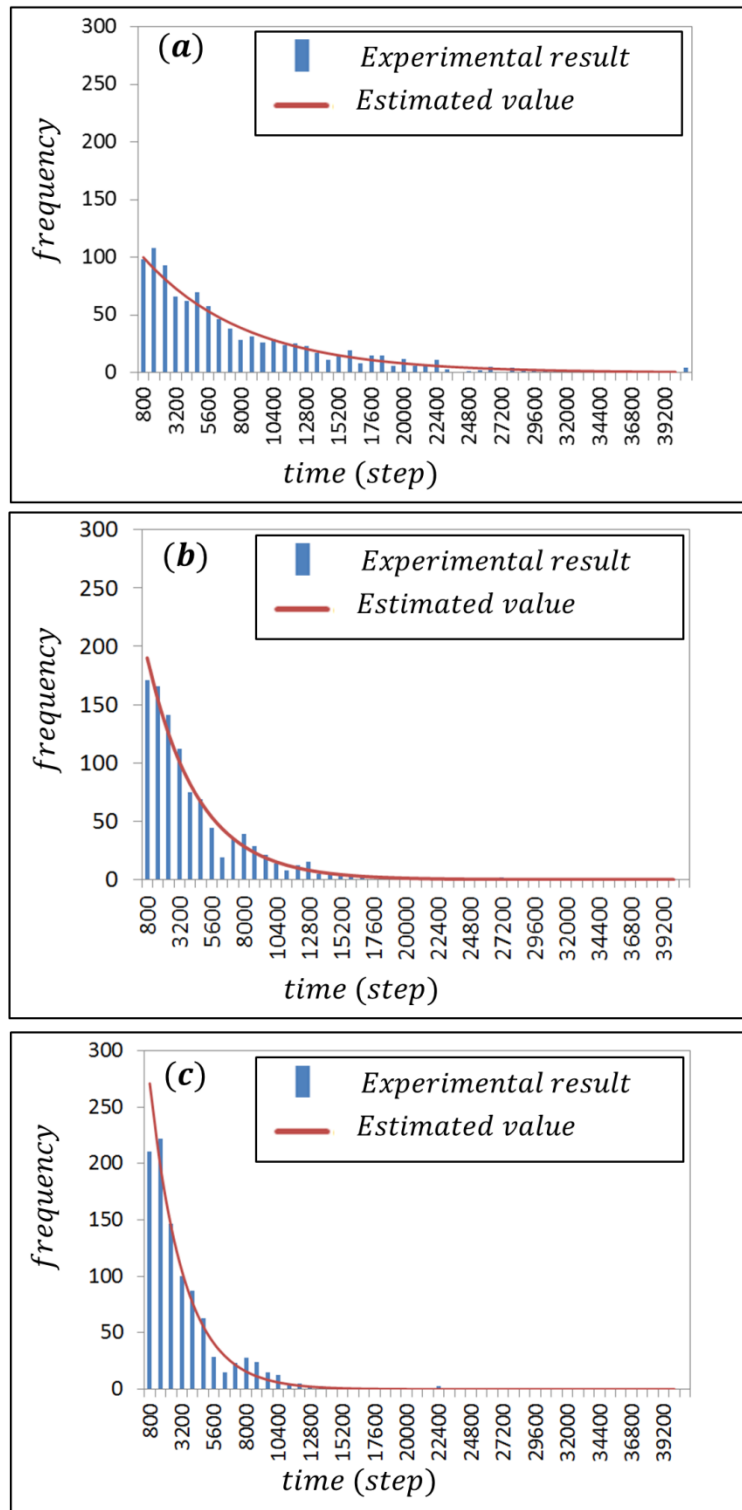


図8.7 探索時間のヒストグラム.

(a) 従来のBRTモデル, (b) 2-BRTモデル, (c) 3-BRTモデル.

$N = 100, M = 10, \tau = 0.001$.

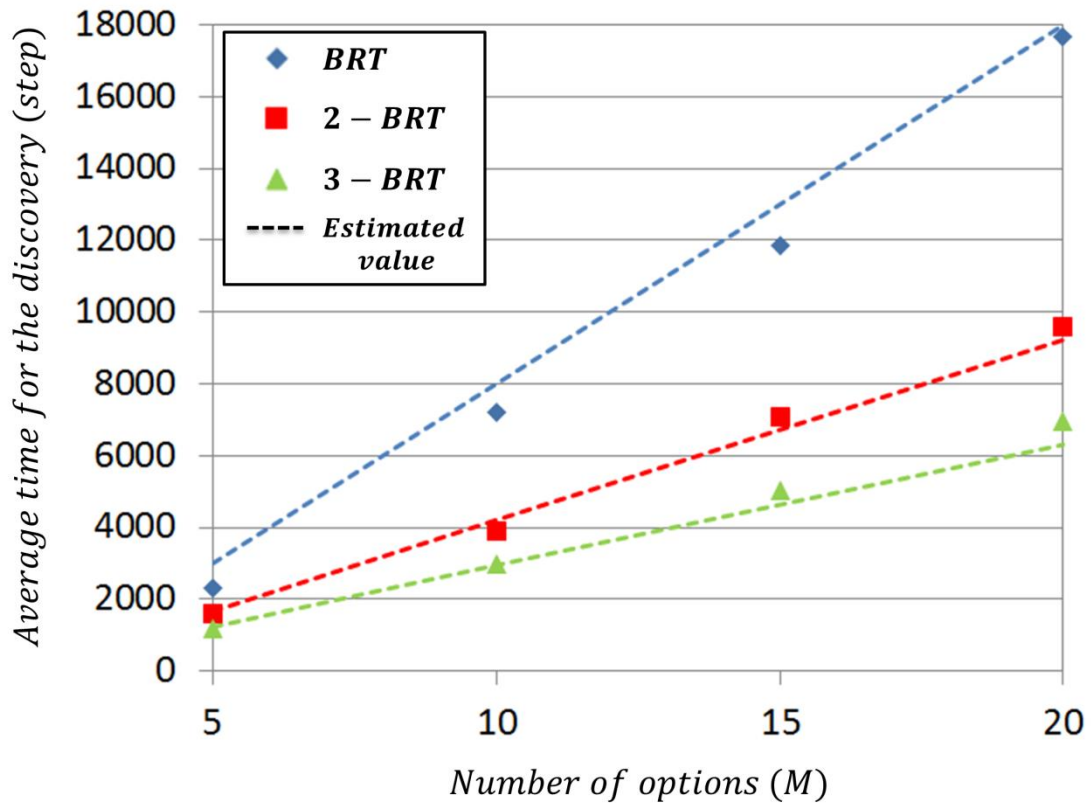


図8.8 適した選択枝の発見に要する平均時間と式(8.14)に基づいた推定値.

$$N = 100, \tau = 0.001, a_{goal} \equiv a_M.$$

次に、発見に要する平均時間と式(8.14)に基づいた推定値との関係性を検証する。実験の設定は前サブセクションと同様に、 $N = 100$, $\tau = 0.001$, $a_{goal} \equiv a_M$ とし、初期時においてエージェントがランダムに選択枝を選択しているとする。選択枝の数 M を変えて、発見に要する時間を測定した。 M ごとにそれぞれ1000試行を実行した。

図8.8に適した選択枝の発見に要する平均時間と式(8.14)に基づいた推定値を示す。横軸は選択枝の数 M を表し、縦軸は平均時間である。◇、□と△のマークはそれぞれBRTモデル、2-BRTモデルと3-BRTモデルの実験結果で、点線は式(8.14)に基づいた推定値を表している。この推定がほぼ実験結果と一致していることから、発見に要する平均時間が選択枝の数 M にほぼ線形的に比例することが確認できた。また、どの M でも、発見に要する平均時間がほぼパラメータ k と反比例していることから、提案手法は発見時間を約 k 倍短縮できる事がわかった。

8.6 まとめ

Q-BRT モデルでは、多数の選択肢からでも非常に高い精度で適した選択肢を発見する事ができるが、多くの計算機実験を通じて適した選択肢を見つけるのに非常に時間がかかるケースがある事がわかってきた。本章では、Plural voting からヒントを得て、集団探索の高速化を試みた。 k 個の投票を導入する事により、適した選択肢の発見時間を約 k 倍短縮できる事がわかった。また、この発見時間を定量的に推定することができるため、時間が限られるタスクや環境などには非常に有効的な手法であろう。

第9章 Q-BRT モデルその②：エルファロール・バー問題

9.1 はじめに

第3章では、the best-of-n 問題に対し、Iwanaga らのモデルに時間経過と共に増加する閾値を導入することにより、従来は困難であった多数の選択肢の扱いを実現した。そして、第7章では、ガウシアン分布の代わりに、解析のしやすい2次関数を用いたバイアス分布を使用する Q-BRT モデルを提案した。数学的な分析と計算機実験により、Q-BRT モデルは、扱える選択肢の数が多く、適した選択肢を素早く発見でき、要する時間の概算ができるという優れた特徴を明らかにした。従って、本提案手法は、可用性が高く、より複雑な問題や様々な分野において適用できると考える。

そこで、本章では、集団の意思決定の精度を議論する際に分散 AI の分野でよく使われる標準問題の1つであるエルファロール・バー (El Farol Bar) 問題に Q-BRT モデルを用いて Q-BRT モデルの実装方法の例を示す。

Arthur^[2]によって提案されたエルファロール・バー問題は、バーに適した嗜好の割合になるよう、エージェントがバーに行くかどうかを決定する問題である。この研究については次節でも詳しく説明するが、例えば Whitehead^[92]は、エージェントが毎ステップ自身の決定を適切に変更すれば、バーの出席者がバーの最適な収容に収束する事を示している。それでもなおローカル意思決定の積み重ねと全体との最適性の両立は依然として群ロボットの分野で重要である。

ここで、エルファロール・バー問題を役割分担問題として扱う。多くの選択肢を扱える特徴を活用することにより従来とは異なるやり方でこの役割分担問題を扱えることを示す。

従来の考え方では、エージェントは、独自にバーに行くあるいは行かないという2つの行動を自由に選択できると仮定するので全体の状態は 2^N の組み合わせになる。従って、扱える選択肢が多い Q-BRT でもこのままでは難しい。一方で個体は社会の動勢に敏感に反応し、このような独立性を持たないのが実態である。そこで、社会的な行動が求められている状況では社会が取り扱う状態は十分に少ないと考えた。それをモードと呼び、選択したモードに合致した行動をエージェントは実施する。つまり、そこで、エージェントには、集団の一員としてバーに行く・行かないという2つの役割がある。多数の選択肢が扱える利点を活かして様々な比率の役割の組み合わせからバーに適したものを探す the best-of-n 問題として実装する。

9.2 エルファロール・バー問題

ここで、エルファロール・バー問題について紹介する。

エルファロール・バー問題はゲーム理論における補完的な問題であり、ニューメキシコ州サンタフェのバーに基づいて、1994年にW. Brian Arthurによって作成された。次のように、定義されている。有限の住民がいるとする。住民みんながエルファロール・バーに行きたいと思っている。しかし、このバーはとても小さく、もし混んでいるなら行っても楽しくない。実際、人々の選好は次のように記述される：

- もし 60%より少ない住民がバーに行けば、彼らはみんな家にいるよりも楽しい時間を過ごすことになる。
- もし 60%より多い住民がバーに行けば、彼らはみんな家にいるよりも悪い時間を過ごすことになる。

全員が同時にバーに行くかどうかを決める必要があるとする。彼らは毎日に彼ら自身がバーに行くかを決める前に、他の人がどれくらいバーに行くのか様子を見ることはできない。このバーは素敵なバーなので誰もが行きたがるが、出席の数が人口の60%以上の場合、集団として最も良い雰囲気にはなれない。したがって、多くの人々が同じ価値観の集団に満足される社会を形成することは困難であり、複雑な社会の研究の基礎として取り上げられている。

9.3 提案手法

9.3.1 エルファロール・バー問題の定義

ここでは、以下のように特別なエルファロール・バー問題を導入する。

N 人の主体で構成される集団 $G=\{P_i:i=1,2,\dots,N\}$ がいるとする。集団はみんな、バーに行きたいと思っている。しかし、このバーはとても小さく、もし混みすぎると行っても楽しくない。 n_{el} ($0 < n_{el} < N$) は、このバーの最適な収容人数である (n_{el} 以上の人数だとバーが混んで楽しめない)。 P_i がバーにいるとき、バーの混む具合によって得した気分もしくは損した気分となる。但し、エージェントはソーシャル・ネットワーキング・サービス (SNS) 等で全員の評価を観測できると仮定して“チーム”として良し悪しを評価する。人々の選好は次の通りである。

- もし、 n_{el} より多くの人々がバーに行くことと決めたならばバーは混んでいてバーに来なければよかったのにと後悔し損した気分となる (一方、バーに行かないと決めた人々はバーに行かなくてよかったと得した気分となる)。そして、人は、この混む具合を SNS 等へ書き込んで、速やかに集団全体へ情報が行き渡る。よって、チームとしては“**良くない集団状態**”だと個体は評価する。
- もし、 n_{el} より少ない人がバーに行くことと決めたならばバーは混まなくてバーに来てよかったと得した気分となるが (一方、バーに行かないと決めた人々はバーに行けばよかったのに損した気分となる)。そして、人は、この混む具合を SNS 等へ書き込んで、速やかに集団全体へ情報が行き渡る。よって、チームとしては“**良い集団状態**”だと個体は評価する。

すなわち、毎時刻においてバーに行くことと決定した人数は n_{el} の値に近ければ近いほど集団はより多くのものが得した気分、損した気分のものが少ない**チーム**として**良い状態**となるものとする。 n_{el} の値とぴったりになれば**最適な状態**となる。以下では、Q-BRT モデルを用いてこの**最適な状態**を見つける方法を提案する。

9.3.2 役割の導入と選択肢の設定

ここでは、役割の導入と選択肢の設定について説明する。

エージェントがバーに行くもしくは行かないという2つの役割の組合せを選択肢として扱うとし、以下のように定義する。

役割：エージェント P_i には $r_i = 1$ (バーに行く)と $r_i = 0$ (バーに行かない)という2つの役割を持つとする。

選択肢の定義：エージェントの選択肢集合を $A = \{a_j : j = 1, \dots, M\} (M \geq 2)$ とし、選択肢 $a_j = \{r_{1j}, \dots, r_{ij}, \dots, r_{Nj}\}$ をエージェントの役割の組合せとする。エージェント P_i の時刻 t での選択肢を $A_i(t) \in A$ とする。選択肢 $A_i(t)$ を選択すると $A_i(t)$ に記録されている自身の役割を果たす(バーに行くもしくは行かない)。例えば、5人から構成される集団がいて、ある時刻に全員が選択肢 $a_1 = \{1, 1, 1, 0, 0\}$ に合意しているとすると、バーに行くのが P_1, P_2 と P_3 で行かないのが P_4 と P_5 である。

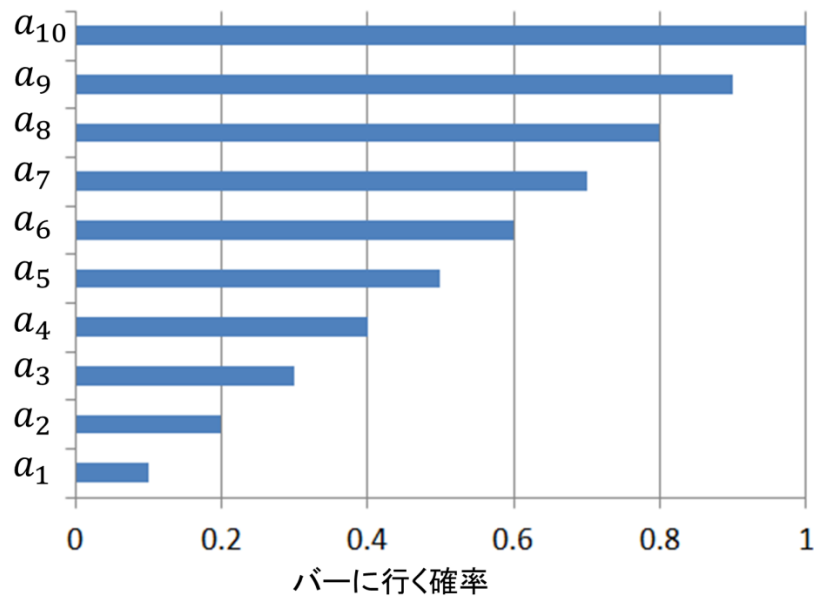


図 9.1 各選択肢におけるバーに行く確率の一例： $M = 10$ 。

予め選択肢の役割を適切に設定できれば、その中から適したものを発見する事が期待できる。ここで、バーの最適な人数 n_{el} がオリジナルの集団の60%に限らず、どの値にも対応できるように、選択肢 a_j に記録されているエージェント P_i の役割 r_{ij} を次のように設定する。

$$r_{ij} = \begin{cases} 1 & \text{random}(0,1) \leq j/M \\ 0 & \text{otherwise} \end{cases} \quad (9.1)$$

すなわち、エージェント P_i が選択肢 a_j を選択するときバーに行く確率は j/M である。例えば、選択肢が 10 ときは、各選択肢におけるバーに行く確率は図 9.1 のようになる。

9.3.3 集団の状態の評価法

次に、集団 G が最適な状態になるよう、Q-BRT モデル集団の評価法について説明する。

最適なものを発見する上で全体の評価と個での評価の関連を適切に定めることが設計上極めて重要である。ここで扱うエルファロール・バー問題の定義に述べたように、バーに行くことと決定した人数は n_{el} の値に近ければ近いほど集団は良い状態になる。集団の状態を評価する方法は色々考えられるがここでは簡単にするため、以下の Loss 関数を使用し、最適な収容人数との差の割合とする。

$$\begin{aligned} L(A_i(t)) &= \left| \frac{n_b(t)}{N} - \frac{n_{el}}{N} \right| \\ &= \left| \frac{n_b(t) - n_{el}}{N} \right| \end{aligned} \quad (9.2)$$

ここで、 $n_b(t)$ は選択肢 $A_i(t)$ を選択するときのバーに行く人数である。

エージェント P_i は評価関数 $c_i(t)$ を持つ。Loss 関数は、

- ある小さな定数 q ($q > 0$) 以下であれば、すなわち、

$$L(A_i(t)) \leq q \quad (9.3)$$

ならば、バーに行く人数は、 n_{el} とほぼ同じ値となり、個体が集団は最適な集団状態に達していると評価する。すなわち、選択肢 $A_i(t)$ が“**適した選択肢**”であることを評価し、 $c_i(t) = 0$ となる。

- q より大きければ、すなわち、

$$L(A_i(t)) > q \quad (9.4)$$

ならば、バーに行く人数は、 n_{el} と離れ、個体が集団は最適な集団状態に達してないと評価する。すなわち、選択肢 $A_i(t)$ が“**適した選択肢でない**”ことを評価し、 $c_i(t) = 1$ となる。

- 評価は全員式(9.2)を用いるものとする。バーに行ったものとそうでないものの差はすぎに集団全体に伝わると想定するものである。

q をマージンと呼び、最適な値(n_{el}/N)との許せる誤差を表している。実の世界では、昆虫や人間社会は常に完璧な状態を求めるわけではなく、**ある程度良けれ**

ばその状態を維持しようとするのである。 q が少し大きい場合、エージェントはバーが少し混んでいても楽しむことができる。一方、 q が 0 に近い場合は、バーに行く人数が n_{el} でない限り、バーを楽しむことはできない。異なる社会には異なるマージン q が存在し、また個人によって q が異なるかもしれない。また、環境の変化によって q の値を適切に調整しなければならない。本章では、簡単のため、集団はみんな同じ小さい q の値をもつとする。

q の値と選択肢の設定によって、 M 個の候補の中に式(9.3)を満たす複数の選択肢が存在することがある（存在しない場合もある）。このような選択肢を“適した選択肢 a_{good} ”と呼ぶ。また、集団を最適な状態にする選択肢 ($L(A_i(t)) = 0$) を“最適な選択肢 a_{best} ”と呼ぶ。従って、

$$L(a_{best}) \leq L(a_{good}) \tag{9.5}$$

が成り立つ。

9.3.4 個体の意思決定アルゴリズム

次に、Q-BRT モデルに基づいた個体の意思決定アルゴリズムについて説明する。

エージェント P_i は個人属性バイアス θ_i ($0 < \theta_i < 1$) を持つ。 $n(a_j)$ を選択肢 a_j を選択しているエージェントの数とする。 図9.2に P_i の意思決定フローチャートを示す。

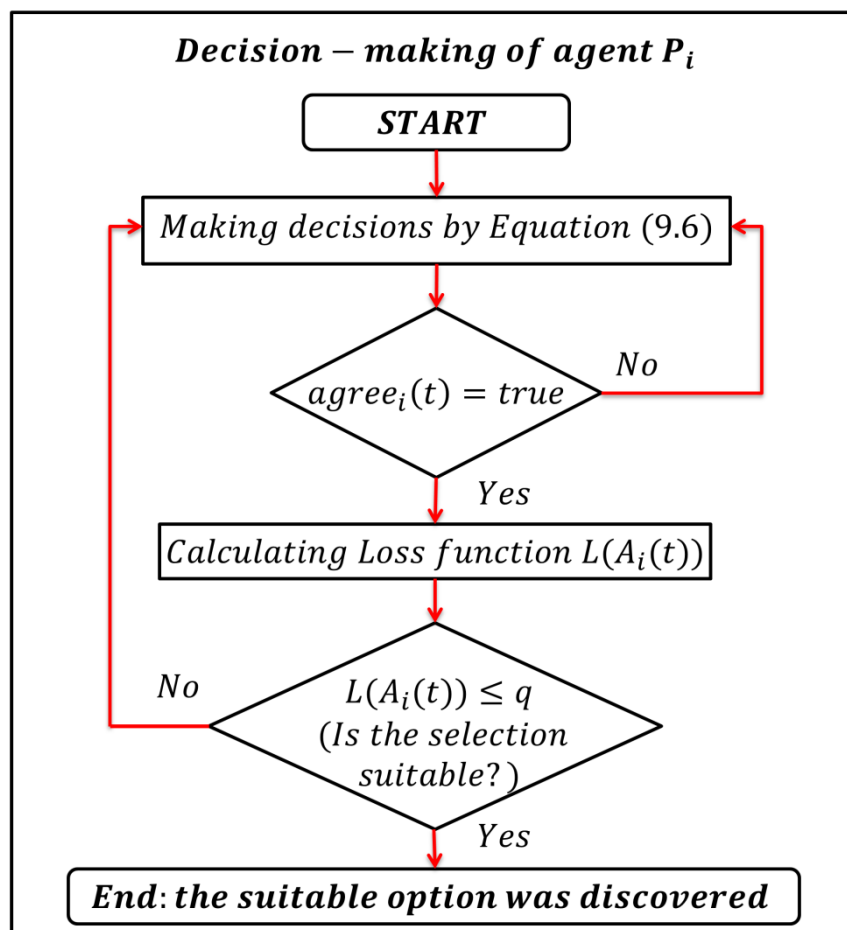


図9.2 Q-BRT アルゴリズムに基づいた
エージェント P_i の意思決定フローチャート。

時刻 t において、エージェント P_i は集団全体で自身と同じ選択肢を選んでいる割合を $n(A_i(t))/N$ とし、次の時刻に選ぶ選択肢 $A_i(t+1)$ を次のように決定する。もし

$$\frac{n(A_i(t))}{N} \geq \theta_i + \tau \cdot c_i(t) \cdot (t - t_{i,last}(t)) \quad (9.6)$$

であれば、現在の選択肢 $A_i(t)$ を引き続き選択し、 $A_i(t+1) = A_i(t)$ となる。 そう

でなければ, 現在の選択肢 $A_i(t)$ をやめてそれ以外の選択肢を確率的に選ぶ ($A_i(t+1) = A - A_i(t)$).

τ は支持者数の増加量を表す定数である. $t_{i,last}(t)$ はエージェント P_i が最後に選択肢を変更した時刻で, $(t - t_{i,last}(t))$ で同一選択肢を選び続けている時間に相当し, 以下の式で求める.

$$t_{i,last}(t) = \begin{cases} t & A_i(t) \neq A_i(t-1) \\ t_{i,last}(t-1) & otherwise \end{cases} \quad (9.7)$$

$c_i(t)$ は選択肢が “適した選択肢” である時 0 となる評価関数である. すなわち

$$c_i(t) = c(t) = \begin{cases} 0 & \forall i, agree_i(t) = true \wedge E(A_i(t)) = true \\ 1 & otherwise \end{cases} \quad (9.8)$$

ここで, $agree_i(t)$ は合意しているかどうかを判断する関数で,

$$agree_i(t) = agree(t) = \begin{cases} true & \exists j, \forall i, A_i(t) = a_j \\ false & otherwise \end{cases} \quad (9.9)$$

である. また, $E(A_i(t))$ は, 選択肢 $A_i(t)$ が式(9.3)を満たせば適した選択肢 a_{good} であることを評価する関数で,

$$E(A_i(t)) = \begin{cases} true & L(A_i(t)) \leq q \\ false & otherwise \end{cases} \quad (9.10)$$

である.

図9.2に示すように, エージェントがQ-BRTモデルに基づいた意思決定を行い, M 個の候補からある任意の選択肢に合意する. その度, 合意した選択肢を評価し, 適した選択肢かどうかを判断する. もし, 適した選択肢ならこれを選び続ける. 適した選択肢ではないなら再び投票を行って, 別の選択肢に合意する. これを繰り返す事によって, 集団にとって適した選択肢を発見することが期待できる.

9.4 実験

エルファロール・バー問題に関する多くの関連研究では、バーの最適な収容人数 n_{el} を集団の 60% として扱っている。本章では、この値に限らず、バーの最適な収容人数 n_{el} がよりデリケートな値（集団の 58%, 59% 等）の場合についても検討を行い、適した選択肢を発見できることを確認し、意思決定に関する特性を明らかにする。

9.4.1 意思決定における集団ダイナミクスの例

まず、最初に集団の意思決定の挙動を図示し、その特徴を明らかにする。

実験の設定は以下の通りである。バーの最適な収容人数 $n_{el} = 0.58 \cdot N$, $N = 100$, $M = 5$, $\tau = 0.003$, $q = 0.03$ とし、初期時でエージェントがランダムに選択肢を選んでいくとした。式(9.1)に基づいて r_{ij} を生成し、その結果、 a_1 から a_5 に記録しているバーに行く人数の割合がそれぞれ 0.16, 0.29, 0.57, 0.82, 1.0 となった。図 9.3 に各選択肢の支持者数の遷移、図 9.4 にその時バーに行く人の割合を示す。

初期時に選択肢 a_1 の支持者数は 28 で最も多かったため、時間経過により、 a_1 の支持者数がどんどん増加してゆき、約 20 ステップ後に、全員で選択肢 a_1 に合意している事がわかる(図 9.3)。図 9.4 の△が付いた線で示すように、 $agree_i(t) = true$ となった合意状態では、エージェントが選択肢 a_1 を評価する。しかし、選択肢 a_1 では 16 人しかバーに行かないため、 $E(a_1) = false$ (式(9.10))となり $c_i(t) = 1$ (式(9.8))となる。よって、式(9.6)の右側が増加し続ける。その結果、200 ステップに合意状態が崩れ、ほぼ初期状態から意思決定が再び行われている。このプロセスがもう 2 回繰り返された。そして、ついに約 790 ステップで全員が選択肢 a_3 に合意し、

$$L(a_3) = |0.57 - 0.58| = 0.01 < q \quad (9.11)$$

のため、選択肢 a_3 が適した選択肢として発見し、終了する。

以上から、提案手法によりバーに適した集団状態を発見できることがわかった。

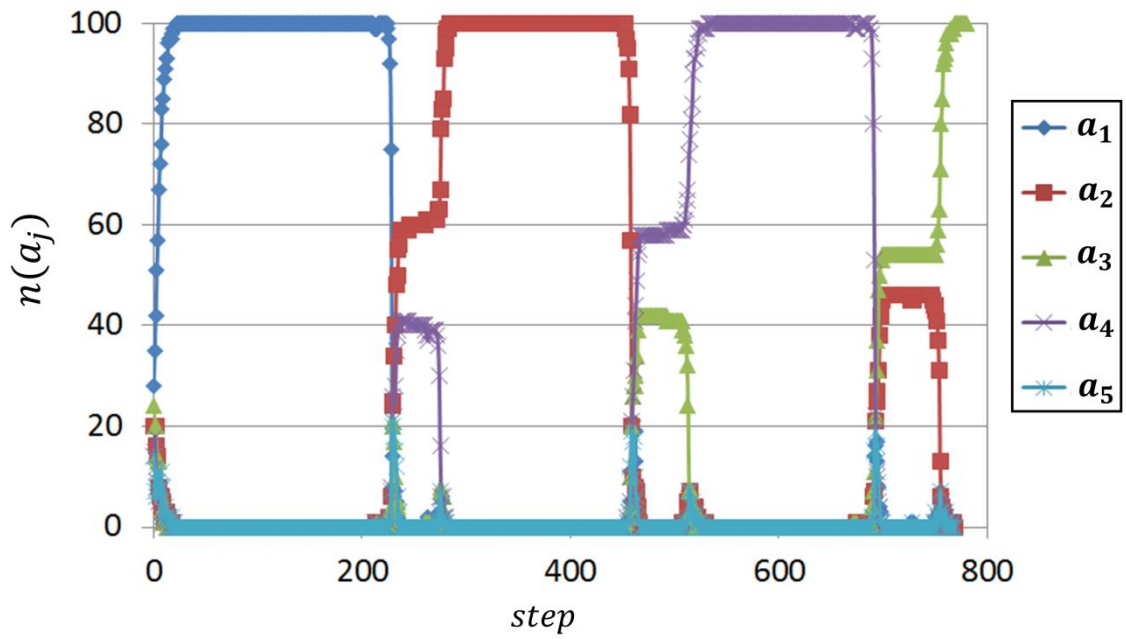


図9.3 各選択肢の支持者数の遷移.

$n_{el} = 0.58 \cdot N$, $N = 100$, $M = 5$, $\tau = 0.003$, $q = 0.03$.

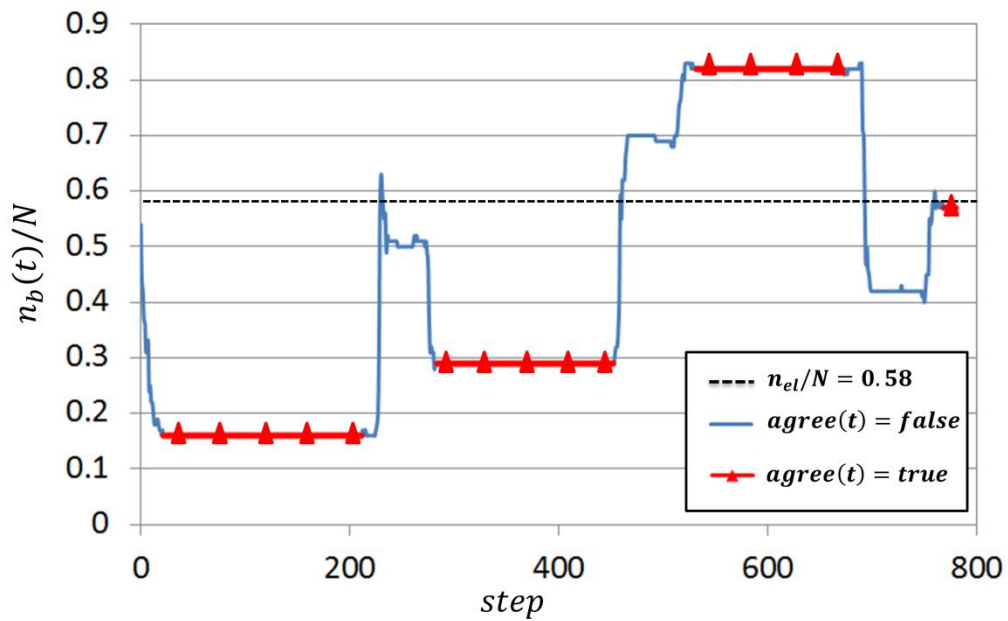


図9.4 バーに行く人の割合の遷移.

$n_{el} = 0.58 \cdot N$, $N = 100$, $M = 5$, $\tau = 0.003$, $q = 0.03$.

9.4.2 適した選択枝の発見精度

前節では，集団の意思決定の一例を挙げて適した集団状態を発見でき事が確認できた．ここでは，バーの最適な人数 n_{el} を 0 から 1 までランダムに設定し，どの n_{el} でも適した集団状態が発見できる事を示す．また，選択枝の数を色々と変えてその発見精度と選択枝の数との関係性を明らかにする．

実験設定は以下の通りである．エージェント数 $N = 100$ ， $\tau = 0.003$ ， $q = 0.03$ ，バーの最適な収容人数を毎回ランダムに設定し ($n_{el} = \text{random}(0,1) \cdot N$)，初期時においてエージェントがランダムに選択枝を選択しているとする．選択枝の数 M を 2, 3, 5, 10, 15, 20 と変え，それぞれの M に対し 1000 回を実行し，制限時間 30000 ステップ以内に適した選択枝が発見された時間を記録する．図 9.5 にその結果を示す．

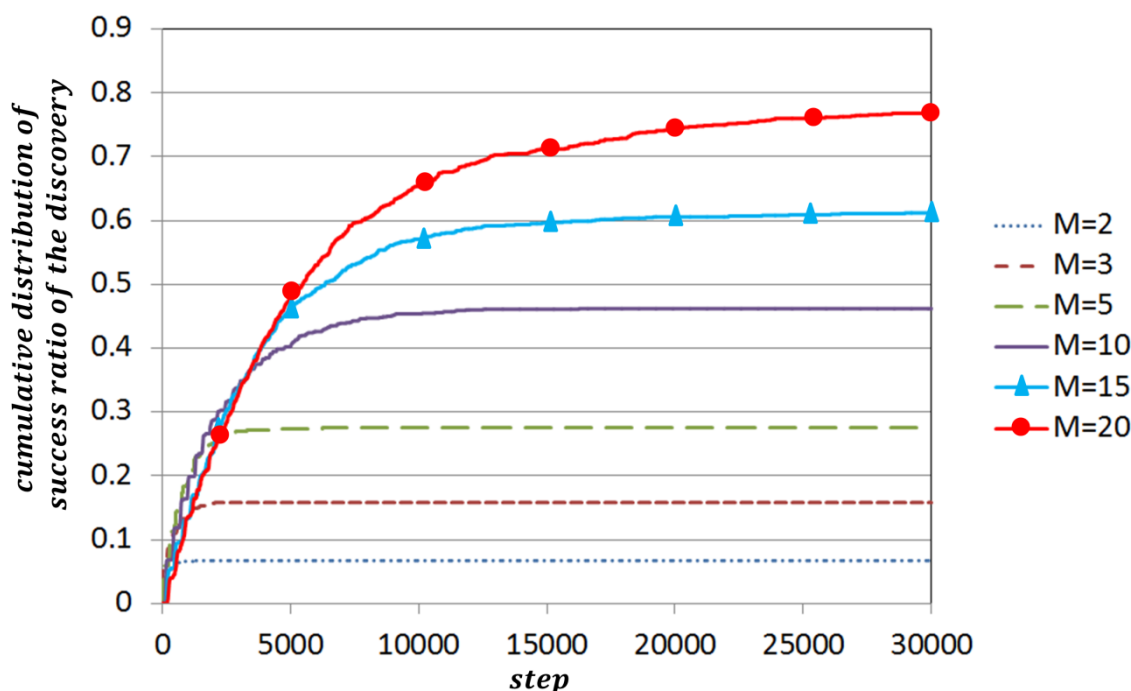


図9.5 異なる M における発見率の累積確率密度分布.
 $N = 100$, $\tau = 0.003$, $q = 0.03$, $n_{el} = \text{random}(0,1) \cdot N$.

図 9.5 は，異なる M における発見率を累積確率密度分布で表したものである．横軸は時間，縦軸はその時間までの発見率を表している．ほぼどの M においても，最初は発見率が急激に上がっているがある時間が経過すると発見率がほぼ一定値に収束している事がわかる．これは， M 候補の中に適した選択枝が存在しないケースと存在するケースがあるからである．適した選択枝が存在するケースに

においては、速やかに発見されるが、存在しないケースにおいては長い時間をかけても見つけれない事がわかった。

その一方、 M が多いほどより高確率かつ短時間で適した選択枝が発見されている事から多数の選択枝が扱える Q-BRT モデルでなければ、 $(0, 1)$ の間に任意に存在する適した集団状態を約 80% ($M = 20$) 近くまで発見できなかった。よって、提案手法は Q-BRT モデルの探索力を活かす事ができた。

9.4.3 マージン q の影響

次に、マージン q は最適な状態の探索にどのような影響を与えるかについて検証を行う。

実験設定は以下の通りである。 $N = 100$, $M = 20$, $\tau = 0.003$ とし、バーの最適な収容人数を毎回ランダムに設定し ($n_{el} = \text{random}(0,1) \cdot N$)、初期時においてエージェントがランダムに選択枝を選択しているとする。マージン $q = 0.01, 0.03, 0.06, 0.09, 0.12$ と色々と変えて、それぞれの q に対し 1000 回を実装し、制限時間 30000 ステップ以内に適した選択枝が発見された時間を測定する。図 9.6 と図 9.7 にその結果を示す。

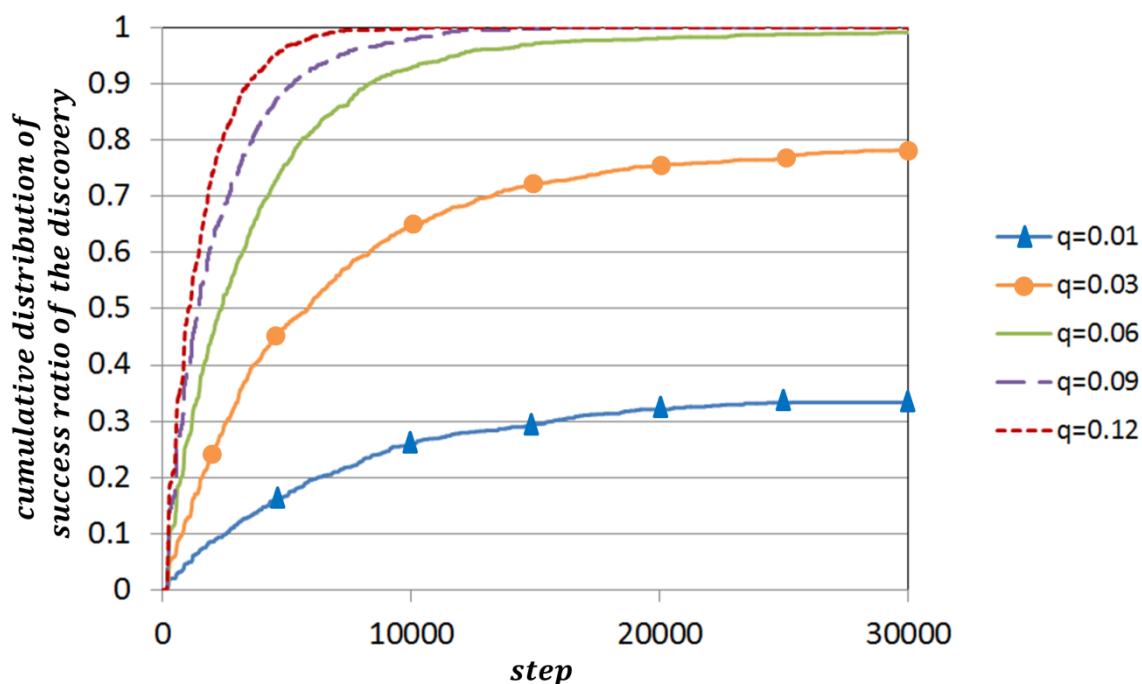


図9.6 異なる q における発見率の累積確率密度分布。
 $N = 100$, $M = 20$, $\tau = 0.003$, $n_{el} = \text{random}(0,1) \cdot N$.

図 9.6 は、異なる q における発見率を累積確率密度分布で表したものである。横軸は時間、縦軸はその時間までの発見率を表している。全体的に、 q が小さい異なっただけで非常に異なる発見率の累積確率密度分布を示している。それに加えて、発見率は、 q の増加に連れて、増加してゆく事がわかる。これは、 q が大きいと、 $(\frac{n_{el}}{N} - q)$ と $(\frac{n_{el}}{N} + q)$ の間に適した集団状態を満たす選択肢がより高確率で存在するからである。しかし、 q が大き過ぎると $(\frac{n_{el}}{N} - q, \frac{n_{el}}{N} + q)$ の間に複数が適した選択肢が存在し、 $\frac{n_{el}}{N}$ の値から離れた選択肢を適した選択肢として発見してしまう恐れがある。その一方、 q が小さすぎると適した選択肢はエージェントが評価できる範囲 $(\frac{n_{el}}{N} - q, \frac{n_{el}}{N} + q)$ に存在せず、結局発見されることはない。この実験設定中では $q = 0.06$ がベストな値で、真の適した選択肢を全ての試行において発見できたものである。式(9.1)によるとエージェントが評価しなければならない一番長い範囲が $(0, \frac{1}{M}]$ であるため、どの $\frac{n_{el}}{N}$ でも必ずその $(\frac{n_{el}}{N} - q)$ と $(\frac{n_{el}}{N} + q)$ の間に少なくとも 1 つの適した選択肢が存在する条件は

$$q > \frac{1}{M} \tag{9.12}$$

である。

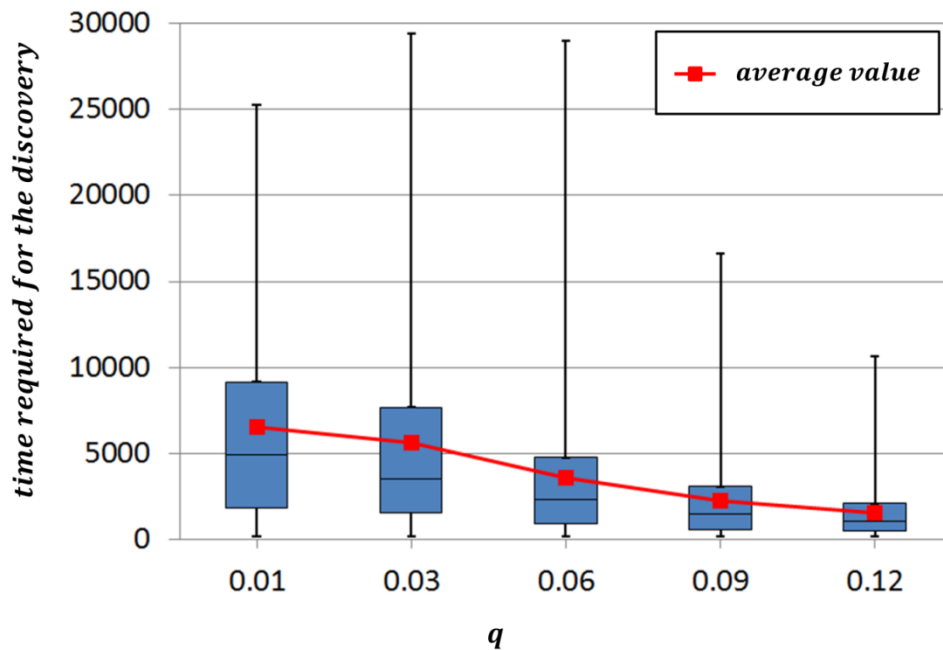


図9.7 q と発見に要する時間の関係.

$$N = 100, M = 20, \tau = 0.003, n_{el} = \text{random}(0,1) \cdot N.$$

図9.7は、 q と発見に要する時間の関係を示す箱ひげ図である。赤い□のマークは、発見に要する平均時間を表している。発見に要する平均時間は、 q とほぼ反比例している事がわかる。また、 q が増加すると、短い時間にデータが集中している事から発見に要する時間のバラツキが減少している事がわかる。

以上から、式(9.12)に示すように、 q が $\frac{1}{M}$ より少し大きな値は高精度かつ短時間で適した集団状態を発見できる事がわかった。

9.5 まとめ

Q-BRT モデルは、多数の選択肢からでも非常に高い精度で適した選択肢を速やかに発見できるアルゴリズムである。本章では、集団の意思決定の精度を議論する際に分散 AI の分野でよく使われる標準問題の 1 つであるエルファロール・バー (El Farol Bar) 問題に Q-BRT モデルを用いて Q-BRT モデルの実装方法の例を示した。個体にバーに行くあるいは行かないという役割を新たに導入し、事前に乱数で生成した個体の役割の組合せを選択肢として使用した。また、バーの最適な人数の割合と許せる誤差を表すマージン q を使用した評価関数に基づいて、個体は合意した選択肢が適したものと判断できるまで各選択肢を取捨選択し続ける。

計算機実験により、適切に事前に個体の役割の組合せを設定すると共に $q > \frac{1}{M}$ である、バーの最適な人数割合がどの値でも、高精度かつ短時間でバーに適した集団状態を発見する事ができた。今後は、事前に個体の役割の組合せを設定しなくても個体が自ら役割を変更するという学習を行い、適した集団状態を発見できる方法について検討する必要がある。

第 10 章 結論と今後の展望

群ロボットの分野では、群れの集団行動の意思決定の研究として the best-of-n 問題という標準問題が知られており、特に注目されている。これは、リーダーを使用せず分散エージェントから成る群れが n 個の候補の中から適した集団行動に合意する集団的意思決定問題である。近年、2 択以上のケースの為の開発が盛んになっている。その主な手法は、value-sensitive という概念に則っている。この概念では選択肢の支持率が選択肢の良さに大きく依存する為、評価のバラツキや突出した候補がない場合などでは合意に到りにくく、結果として多数の選択肢を扱うのが困難である。

そこで、本研究では、この the best-of-n 問題に対して個体の意思決定が選択肢の良さに依存しない value-free という概念に基づく手法を提案した。

第 2 章では、Iwanaga らは、個人の意思決定は自身の好みを表す個人属性だけでなく集団全体の雰囲気からも影響を受けると考え、集団に個人差を導入した社会モデルを提案し、選択肢の良さに依存せず行動が選ばれながらも、速やかに社会が合意状態に達する事を明らかにしている。しかしながら、3 択以上の理論がなく群れとしての探索機能が不要な 2 択問題しか扱う事が出来ない事を述べ、同様の概念に基づくアプローチの研究が不十分である事を指摘した。

第 3 章では、Iwanaga らのモデルに時間経過と共に増加する閾値(BRT, Bias and Rising Threshold)を導入し、集団が適さない選択肢に合意するのをやめて別の選択肢に合意するという試行錯誤ダイナミクスを発生できる BRT モデルを提案した。計算機実験により、群れは任意の選択肢に速やかに合意し、その合意行動を切り替えて、選択肢の数が 20 の場合でも適した合意行動を発見できる事を示した。

第 4 章では、合意したものの適さなかった選択肢を記憶する UEM(Unsuitable Experience Memory)エージェントを導入して、探索の効率化を試みた。計算機実験により、微少数の UEM エージェントを導入するだけでも、発見時間を劇的に改良できる事を明らかにした。

第 5 章では、提案した BRT モデルを観測範囲に限りがあるよりリアルな環境での可用性を検討した。個体の観測範囲を制限した移動エージェントのシミュレーションにより、パラメータを適切に設定すれば適した選択肢が発見できる事が確認できた。また、少し観測範囲を広げるだけで大幅に精度の改善がみられ、観測範囲とパフォーマンスには閾值的な関係がある事を推測した。

第 6 章では、生物がもつ集団行動のプリファレンス（優先度）を実現する為、

各選択肢に対して異なるバイアス値を用いた HBRT モデルを提案した。計算機実験により、集団のバイアス分布の平均値と分散値を適切に変えれば、それぞれの選択肢のプリファレンスを実現できる事を明らかにした。

第7章では、解析のしやすい2次関数を用いたバイアス分布を使用する Q-BRT モデルを提案し、数学的な分析を行った。その結果、選択肢が2つに限らず多数でも群れは任意の選択肢に合意する事を明らかにした。また、投票フェーズに要する時間と適した選択肢の発見に要する平均時間を概算し、これらの時間が群れのサイズに依存せず、選択肢の数に線形的に比例することを明らかにした。また、計算機実験を使用し、扱える選択肢の数は、支持者数の増加量を表す τ に依存し、 τ を小さくするほど扱える選択肢が増える事を明らかにした。

第8章では、Plural voting からヒントを得て、さらなる集団探索の高速化を試みた。数学的な解析と計算機実験により、同時に k 個の投票を実施する事により、適した選択肢の発見時間を約 k 倍短縮できる事を明らかにした。

最後に、本 BRT モデルシステムの利用方法例を示した。題材として集団の意思決定の精度を議論する際に分散 AI の分野でよく使われる標準問題の1つであるエルファロール・バー (El Farol Bar) 問題を用いた。個体に、バーに行くあるいは行かないという役割項を新たに導入し、事前に生成した個体の役割の組合せを選択肢として使用した。また、バーの最適な人数の割合との許せる誤差を表すマージン q を使用した評価関数を基づいて、個体は合意した選択肢が適したものと判断できるまで各選択肢を取捨選択し続ける。計算機実験によって、事前に個体の役割の組合せとマージン q を適切に設定すれば、バーの最適な人数割合がどの値でも、高精度かつ短時間でバーに適した集団状態を発見する事を明らかにした。

以上のように、多数の選択肢の the best-of- n 問題 ($n \gg 2$) でも適用できる群れの意思決定に関するフレームワークを構築できたと考える。

本 BRT モデルは人間の特性を考慮したものの為、ロボットシステムだけでなく人間を含む機械系への適用が可能であるが公平性についての更なる研究が必要である。

謝辞

防衛大学校理工学研究科在学中は、多くの方々にご指導、ご協力頂き、本研究を行うことができました。

特に、研究において多くのご指導、ご鞭撻を賜りました担当指導教官の久保正男准教授に心より御礼申し上げます。久保先生は、いろいろな方向に発散し、一向に収束しなかった私を、様々なアドバイスによって導いてくださいました。また、語学不得意な私に、国内外の学会発表などの機会を与えて下さったことは本当にいい経験になりました。久保先生研究室での6年間で得られた幅広い知識や考え方は、これからの人生においても大いに役に立つと確信しております。

佐藤浩准教授には、研究を進める上で必要な技術やコンピュータ環境の構築等について貴重なご意見、ご助言を頂くとともに、楽しそうに研究されているお姿に感化されました。

最後に、本研究中、有意義な討論をしていただいた研究室の方々をはじめ、多くの示唆を与えて下さった全ての方々に感謝いたします。ありがとうございました。

参考文献

- [1] Abdelhak C., Ilhem K., Philippe G., Adel M. A., Emergent complex behaviors for swarm robotic systems by local rules, In *IEEE workshop on robotic intelligence informationally structured space, IEEE*. pp. 69-76 (2011).
- [2] Arthur W. B., Inductive Reasoning and Bounded Rationality, In *Proceedings of the Hundred and Sixth Annual Meeting of the American Economic Association*, Vol. 84, No. 2, pp. 404-411 (1994).
- [3] Barbaro A., Einarsson B., Birnir B., Sigurðsson S., Valdimarsson H., Paʼlsson O. K., Sveinbjörnsson S., Modelling and simulations of the migration of pelagic fish, *ICES J Mar Sci: Journal du Conseil*, Vol. 66, No. 5, pp. 826-838 (2009).
- [4] Barnes L. E., Fields Mary Anne, Valavanis Kimon P., Swarm formation control utilizing elliptical surfaces and limiting functions, *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, Vol. 39, pp. 1434-1445 (2009).
- [5] Beshers S. N., Fewell J.H., Models of division of labor in social insects, *Annu Rev Entomol*, Vol. 46, pp. 413-440 (2001).
- [6] Bonabeau E., Dorigo M., Theraulaz G., From natural to artificial swarm intelligence, Oxford: Oxford University Press (1999).
- [7] Bonabeau E., Theraulaz G., Deneubourg J. J., Fixed Response Thresholds and the Regulation of Division of Labor in Insect Societies, *Bulletin of Mathematical Biology*. Vol. 60, pp. 753-807 (1998).
- [8] Bone Q., Biology of fishes, England: Taylor & Francis (2008).
- [9] Brambilla M., Ferrante E., Birattari M., Dorigo M., Swarm robotics: a review from the swarm engineering perspective, *Swarm Intell.*, Vol. 7, pp. 1-41 (2013). Doi:10.1007/s11721-012-0075-2.
- [10] Buhl J., Sumpter D. J. T., Couzin I. D, Hale J. J., Despland E., Miller E. R., Simpson S. J., From disorder to order in marching locusts, *Science*, Vol. 312, No. 5778, pp. 1402-1406 (2006).

- [11] Camazine S., Deneubourg J. L., Frank N. R., Sneyd J., Theraulaz G., Bonabeau E., Self-organization in biological systems, Princeton: Princeton University Press (2003).
- [12] Castello E., Yamamoto T., Libera F. D., Liu W., Winfield Alan F. T., Nakamura Y., Ishiguro H., Adaptive Foraging for Simulated and Real Robotic Swarms: The dynamical response threshold approach, *Swarm Intelligence*, Vol. 10, pp. 1-31 (2016). Doi: 10.1007/s11721-015-0117-7.
- [13] Costerton J. W., Lewandowski Z., Caldwell Douglas E., Korber D. R., Lappin-Scott H. M., Microbial biofilms, *Annu Rev Microbiol*, Vol. 49, No. 1, pp. 711-745 (1995).
- [14] Couzin I. D., Ioannou C. C., Demirel G., Gross T., Torney C. J., Hartnett A., et al., Uninformed individuals promote democratic consensus in animal groups. *Science*, Vol. 334, pp. 1578–1580 (2011). Doi: 10.1126/science.1210280.
- [15] Dario F., Claudio M., Bio-inspired artificial intelligence: theories, methods, and technologies, *Massachusetts: The MIT Press* (2008).
- [16] Diego O., John O., In Search of an Understandable Consensus Algorithm. In Proceedings of the 2014 USENIX conference on Annual Technical Conference, pp. 305-320 (2014).
- [17] Garnier S., Combe M., Jost C., Theraulaz G., Do Ants Need to Estimate the Geometrical Properties of Trail Bifurcations to Find an Efficient Route? A Swarm Robotics Test Bed, *PLoS Comput. Biol.* Vol. 9, No.3, (2013). Doi: 10.1371/journal.pcbi.1002903.
- [18] Gauci M., Chen J., Li W., Dodd T. J., Groß R., Self-organized aggregation without computation, *Int. J. Robot. Res.*, Vol. 33, pp. 1145-1161 (2014). Doi: 10.1177/0278364914525244.
- [19] George S. L., Plural Voting (Abolition) Bill (No 42), Parliamentary Debates (Hansard), United Kingdom: House of Commons, (18 May 1892),
- [20] Gordon D. M., Group-level dynamics in harvester ants: young colonies and the role of patrolling, *Anim. Behav.* Vol. 35, pp. 833-834, (1987).

- [21] Gordon D. M., The dynamics of the daily round of the harvester ant colony, *Anim.Behav*, Vol. 34, pp. 1402-1419 (1986).
- [22] Goss S., Aron S., Deneubourg J. L., Pasteels J. M., Self-organized shortcuts in the argentine ant, *Naturwissenschaften*, Vol. 76, No. 12, pp. 579-581 (1989).
- [23] Hamann H., Schmickl T., Wörn H., Crailsheim K., Analysis of emergent symmetry breaking in collective decision making, *Neural Comput. Appl*, Vol. 21, pp. 207-218 (2012).
- [24] Hasegawa E., Mizumoto N., Kobayashi K., Dobata S., Yoshimura J., Watanabe S., Murakami Y., Matsuura K, Nature of collective decision-making by simple yes/no decision units, *Scientific Reports*, Vol. 7, No. 14436 (2017).
- [25] Hatano Y., Mesbahi M., Agreement over random networks, *IEEE Trans. Automat. Contr.*, Vol. 50, pp. 1867-1872 (2005). Doi: 10.1109/TAC.2005.858670.
- [26] Hauert S., Leven S., Varga M., Ruini F., Cangelosi A., Zufferey J. C., et al., Reynolds flocking in reality with fixed-wing robots: communication range vs. maximum turning rate, In *Proceedings of the 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2011 (Piscataway)*, pp. 5015–5020 (2011).
- [27] Hettiarachchi S., Spears William M., Distributed adaptive swarm for obstacle avoidance, *Int J Intell Comput Cybern*, Vol. 2, pp. 644-671 (2009).
- [28] Iwanaga S., Namatame A., The complexity of collective decision, *Nonlinear Dynamics, Psychology, and Life Sciences*, Vol. 6, No.2, pp. 137-158 (2002).
- [29] Jackson D.E., Ratnieks F. L., Communication in ants, *Curr Biol*, Vol. 16, No. 15, pp. 570-577 (2006).
- [30] Johnson L., Choi H. L., How J., The Role of Information Assumptions in Decentralized Task Allocation: A Tutorial, *IEEE Control Systems Magazine*, Vol. 36, No. 4, pp. 45-58 (2016).
- [31] Khatib O., Real-time obstacle avoidance for manipulators and mobile robots, *Int J Robot Res*, Vol. 5, pp. 90-98 (1986).

- [32] Krapivsky P. L., Redner S., Dynamics of majority rule in two-state interacting spin systems, *Physical Review Letters*, Vol. 90, No. 23 (2003).
- [33] Kubo M., Phung N. H., Sato H., Efficient collective search by agents that remember failures, In *Proceedings of Acarob 2018* (2018).
- [34] Kubo M., Phung N. H., Sato H., Namatame A., Direction switch behavior to enclose a pack of targets based on phase transition, In *Proceedings of the 1st International Symposium on Swarm Behavior and Bio-Inspired Robotics*, pp. 295-298 (2015).
- [35] Lambiotte R., Saramäki J., Blondel V. D., Dynamics of latent voters, *Phys. Rev. E Stat. Nonlin. Soft. Matter Phys.*, Vol. 79, 046107 (2009).
- [36] Liggett T., Stochastic Interacting Systems: Contact, Voter and Exclusion Processes, Vol. 324 (S. S. Chern, B. Eckmann, and P. de la), Harpe (eds.), Springer, 1999.
- [37] Liu W, Winfield A., Sa J., Chen J., Dou L., A Strategies for Energy Optimisation in a Swarm of Foraging Robots, *Swarm Robotics*, pp. 14-26, 2006.
- [38] Marshall J. A. R., Bogacz R., Dornhaus A., Planqué R., Kovacs T., Franks N. R., On optimal decision-making in brains and social insect colonies, *J. R. Soc. Interface*, Vol. 6, pp. 1065-1074 (2009). Doi: 10.1098/rsif.2008.0511.
- [39] McLurkin J., Smith J., Distributed algorithms for dispersion in indoor environments using a swarm of autonomous mobile robots, *Distributed autonomous robotic systems*, Vol. 6, Japan: Springer, pp. 399-408 (2007).
- [40] Menzel R., Giurfa M., Cognitive architecture of a mini-brain: the honeybee, *Trends Cogn Sci*, Vol. 5, No. 2, pp. 62-71 (2001).
- [41] Michael R., Christian A., Radhika N., Kilobot: a low cost scalable robot system for collective behaviors. In *IEEE International conference on robotics and automation*, pp. 3293-3298 (2012).
- [42] Michael Rubenstein, Alejandro Cornejo, Radhika Nagpal: Programmable self-assembly in a thousand-robot swarm, *Science*, Vol. 345, No. 6198, pp. 795-799 (2014).

- [43] Mill J. S., *Considerations on Representative Government* (1861).
- [44] Montes de Oca M. A., Ferrante E., Scheidler A., Pinciroli C., Birattari M., Dorigo M., Majority-rule opinion dynamics with differential latency: a mechanism for self-organized collective decision-making, *Swarm Intelligence*, Vol. 5, pp. 305-327 (2011). Doi: 10.1007/s11721-011-0062-z.
- [45] Moyle Peter B., Cech Joseph J., *Fishes: an introduction to ichthyology*, Prentice Hall (1988).
- [46] Nicholas B. Davies, John R. Krebs, Stuart A. West, *An introduction to behavioural ecology*, 4th edition, chapter 6: Living in groups, pp. 147-178, Wiley-Blackwell (2012).
- [47] Olfati-Saber R., Fax J. A., Murray R. M., Consensus and Cooperation in Networked Multi-agent Systems, *In Proceedings of the IEEE*, Vol. 95, No. 1, pp. 215-233 (2007).
- [48] Olfati-Saber R., Murray R. M., Consensus problems in networks of agents with switching topology and time-delays, *IEEE Transactions on automatic control*, Vol. 49, No. 9, pp. 1520-1533 (2004).
- [49] Pacala S. W., Gordon D. M., Godfray H. C. J., Effects of social group size on information transfer and task allocation, *Evol. Ecol*, Vol. 10, pp. 127-165 (1996).
- [50] Page R. E., Mitchell S. D., The genetics of division of labour in honey beecolonies, *Advances in Insect Physiology*, Vol. 23, pp. 117-169 (1991).
- [51] Pais D., Hogan P. M., Schlegel T., Franks N. R., Leonard N. E., Marshall J. A. R., A mechanism for value-sensitive decision-making, *PloS One*, Vol. 8, No. 9 (2013).
- [52] Parker C. A. C., Zhang H., Biologically inspired collective comparisons by robotic swarms, *The International Journal of Robotics Research*, Vol. 30, pp. 524-535 (2011). Doi: 10.1177/0278364910397621.
- [53] Parker C. A. C., Zhang H., Collective unary decision-making by decentralized multiple-robot systems applied to the task-sequencing problem, *Swarm Intelligence*, Vol. 4, pp.199-220 (2010). Doi: 10.1007/s11721-010-0042-8.

- [54] Parker C. A. C., Zhang H., Cooperative decision-making in decentralized multiple-robot systems: the best-of-n problem, *IEEE/ASME Transactions on Mechatronics*, Vol. 14, No. 2, pp. 240-251 (2009).
- [55] Parr L. A., Winslow James T., Hopkins William D., de Waal Frans B. M., Recognizing facial cues: individual discrimination by chimpanzees (pan troglodytes) and rhesus monkeys (macacumulatta), *Journal of Comparative Psychology*, Vol. 114, No. 1, pp. 47-60 (2000).
- [56] Parr L. A., de Waal F. B., Visual kin recognition in chimpanzees, *Nature*, Vol. 399, pp. 647-647 (1999).
- [57] Phung N. H. (In Japanese), A study on extension of enclosing algorithm for mutilple mobile agents, Master's thesis at National Defense Academy (2015).
- [58] Phung N. H., Kubo M., Sato H., Iwanaga S., Agreement Algorithm using the Trial and Error Method at the Macro Level, In *Proceedings of the 2nd International Symposium on Swarm Behavior and Bio-Inspired Robotics*, pp. 82-85 (2017).
- [59] Pitcher T. J., Magurran A. E., Winfield I. J., Fish in larger shoals find food faster, *Behav Ecol Sociobiol*, Vol. 10, No. 2, pp. 149-151 (1982).
- [60] Plural voting (December, 18, 2019). In *Wikipedia: The Free Encyclopedia*. Retrieved from https://en.wikipedia.org/wiki/Plural_voting.
- [61] Prins H. H. T, Ecology and behaviour of the African Buffalo (1996).
- [62] Ravary F., Lecoutey E., Kaminski G., Cha`line N., Jaisson P., Individual experience alone can generate lasting division of labor in ants, *Current Biology*, Vol. 17, No. 15, pp.1308-1312 (2007).
- [63] Reina A., Marshall J. A. R., Trianni V., Bose T., Model of the best-of-N nest-site selection process in honeybees, *Physical Review E*, Vol. 95 (2017). Doi: 10.1103/PhysRevE.95.052411.
- [64] Reina A., Miletitch R., Dorigo M., Trianni V., A quantitative micro-macro link for collective decisions: the shortest path discovery/selection example, *Swarm Intelligence*, Vol. 9, pp. 75-102 (2015). Doi: 10.1007/s11721-015-0105-y.

- [65] Reina A., Valentini G., Fernandez-Oto C., Dorigo M., Trianni V., A design pattern for decentralised decision-making, *PloS One* , Vol. 10, No. 10 (2015). Doi: 10.1371/journal.pone.0140950.
- [66] Reynolds C. W., Flocks, herds and schools: a distributed behavioral model, *ACM SIGGRAPH computer graphics*, Vol. 21, No. 4, pp. 25-34 (1987).
- [67] Scheidler A., Brutschy A., Ferrante E., Dorigo M., The k-unanimity rule for self-organized decision making in swarms of robots, *IEEE Transactions on Cybernetics*, Vol. 46, No. 5, pp. 1175-1188 (2016).
- [68] Scheidler A., Dynamics of majority rule with differential latencies, *Physical Review E*, Vol. 83, No. 031116 (2011). Doi: 10.1103/PhysRevE.83.031116.
- [69] Schelling T. S., *Micromotives and Macrobehavior*, Norton (1978).
- [70] Seeley T. D., Visscher P. K., Schlegel T., Hogan P. M., Franks N. R., Marshall J. A. R., Stop signals provide cross inhibition in collective decision-making by honeybee swarms, *Science*, Vol. 335, No. 6064, pp. 108-111 (2012).
- [71] Seeley T. D., *Honeybee Democracy*, Princeton University Press (2010).
- [72] Shang Y., Bouffanais R., Influence of the number of topologically interacting neighbors on swarm dynamics, *Scientific Repots*, Vol. 4, No. 4184 (2014). Doi: 10.1038/srep04184.
- [73] Shapiro J. A., Thinking about bacterial populations as multicellular organisms, *Annual review of microbiology*, Vol. 52, No. 1, pp. 81-104 (1998).
- [74] Shu-Heng Chen, Umberto Gostoli, Agent-Based Modeling of the El Farol Bar Problem, *Algorithmic Social Science Research Unit. Discussion Papers 1120* (2011).
- [75] Theraulaz G., Gautrais J., Camazine S., Deneubourg J. L., The formation of spatial patterns in social insects: from simple behaviours to complex structures, *Philosophical transactions of the Royal Society of London. Series A, mathematical, Physical and Engineering Science*, Vol. 361, No. 1807, pp. 1263-1282 (2003).

- [76] Thorup K., Alerstam T., Hake M., Kjelle'n N., Bird orientation: compensation for wind drift in migrating raptors is age dependent, In *Proceedings of Biological Sciences*, Vol. 270, Issue Suppl_1 (2003).
- [77] Todd P. M., Borges B., Designing socially intelligent agents for the ultimatum game, *Socially intelligent agents-Papers from the 1997 Fall Symposium*, pp. 134-136 (1997).
- [78] Todd P. M., Searching for the next best mate, simulating social phenomena, *Springer-Verlag*, pp. 419-436 (1997).
- [79] Todd P. M., A Connectionist Approach to Algorithmic Composition, *Computer Music Journal*, Vol. 13, No. 3, pp. 27-43 (1989).
- [80] Tofts C., Franks N.R., Doing the right thing: ants, honeybees and naked mole-rats, *Trends in Ecology & Evolution*, Vol.7, pp. 346-349 (1992).
- [81] Traniello J. F. A., Rosengaus R.B., Ecology, evolution and division of labour in social insects, *Animal Behaviour*, Vol. 53, pp. 209-213 (1997).
- [82] Valentini G., Achieving consensus in robot swarms: design and analysis of strategies for the best-of-n problem, In *Achieving consensus in robot swarms*, Springer International Publishing (2017).
- [83] Valentini G., Brambilla D., Hamann H., Dorigo M., Collective Perception of Environmental Features in a Robot Swarm, In Dorigo M. et al. (eds) *ANTS 2016: Swarm Intelligence*, pp. 65-76, Lecture Notes in Computer Science, Vol. 9882, Springer, Cham.
- [84] Valentini G., Ferrante E., Dorigo M., The best-of-n problem in robot swarms: formalization, state of the art, and novel perspectives, *Frontiers in Robotics and AI*, Vol. 4, Article 9, pp. 1-18 (2017).
- [85] Valentini G., Hamann H., Dorigo M., Self-organized collective decision making: the weighted voter model, In *Proceedings of the 2014 international conference on autonomous agents and multi-agent systems*, pp. 45-52 (2014).

- [86] Vicsek T., Czirók A., Ben-Jacob E., Cohen I., Shochet O., Novel type of phase transition in a system of self-driven particles, *Physical Review Letters*, Vol. 75, pp. 1226-1229 (1995). Doi: 10.1103/PhysRevLett.75.1226.
- [87] Vicsek T., Zafeiris A., Collective motion, *Physics Reports*, Vol. 517, No. 3-4, pp. 71-140 (2012). Doi: 10.1016/j.physrep.2012.03.004.
- [88] Vittori K., Talbot G., Gautrais J., Fourcassie V., Araujo A. F., Theraulaz G., Path efficiency of ant foraging trails in an artificial network, *Journal of Theoretical Biology*, Vol. 239, No. 4, 507-515 (2006).
- [89] Wallraff H. G., de Waal Frans B. M., Avian navigation: pigeon homing as a paradigm, *Germany: Springer* (2005).
- [90] Weisstein E. W., Harmonic Series, MathWorld (accessed 10 October 2018). <http://mathworld.wolfram.com/HarmonicSeries.html>.
- [91] Wessnitzer J., Melhuish C., Collective decision-making and behaviour transitions in distributed ad hoc wireless networks of mobile robots: target-hunting, In: Banzhaf W., Ziegler J., Christaller T., Dittrich P., Kim J.T. (eds) *Advances in Artificial Life. ECAL 2003, Lecture Notes in Computer Science*, Vol. 2801, pp. 893-902.
- [92] Whitehead D., The El Farol Bar Problem Revisited: Reinforcement Learning in a Potential Game, ESE Discussion Papers 186, Edinburgh School of Economics, University of Edinburgh (2008).
- [93] Ying Tan, Zhong-yang Zheng, Research Advance in Swarm Robotics, *Defence technology*, Vol. 9, No. 1, pp. 18-39 (2013).
- [94] Yongming Y., Xihui C. et al., Swarm Robots Task Allocation Based on Local Communication, In *Proceedings of 2010 Int. Conf. on Computer, Mechatronics, Control and Electronic Engineering*, Vol.5, pp. 415-418 (2009).
- [95] 生天目 章, 戦略的意思決定, pp. 119-149, 朝倉書店 (2001).

研究成果

学術論文

- (1) **Phung N. H.**, Kubo M., Sato H., Iwanaga S., Agreement Algorithm using the Trial and Error Method at the Macrolevel, *Artificial Life and Robotics*, Vol. 23, No. 4, pp. 564-570 (2018).
- (2) Kubo M., **Phung N. H.**, Sato H., Efficient Collective Search by Agents that Remember Failures. *Journal of Robotics, Networking and Artificial Life*, Vol. 5, No. 1, pp. 67-70 (2018).
- (3) **Phung N. H.**, Kubo M., Sato H., Efficient Searching by Bias and Raising Threshold Algorithm Using Multiple Voting in the Best-of-n Problem. *International Journal of Computer Theory and Engineering*, Vol. 11, No. 3, pp. 39-45 (2019).
- (4) **Phung N. H.**, Kubo M., Sato H., Agreement Algorithm Based on a Trial and Error Method for the Best of Proportions Problem, *Journal of Robotics and Mechatronics*, Vol. 4, No. 11, pp. 558-565 (2019).

国際学会（査読付）

- (1) **Phung N. H.**, Kubo M., Yamagichi A., Sato H., Iwanaga S., Namatame A., Enclosing of Target Zone, In *Proceedings of the 1st International Symposium on Swarm Behavior and Bio-Inspired Robotics* (2015).
- (2) Kubo M., Sato H., **Phung N. H.**, Namatame A., Direction switch behavior to enclose a pack of targets based on phase transition, In *Proceedings of the 1st International Symposium on Swarm Behavior and Bio-Inspired Robotics* (2015).
- (3) Kubo M., **Phung N. H.**, Sato H., Direction Switch Behavior to Enclose a Target, In *Proceedings of the 2017 International Conference on Artificial Life and Robotics*, pp. 489-493 (2017).
- (4) **Phung N. H.**, Kubo M., Sato H., Iwanaga S., Namatame A., Agreement Algorithm using the Trial and Error Method at the Macro Level, In *Proceedings of the 2nd International Symposium on Swarm Behavior and Bio-Inspired Robotics*, pp. 82-85 (2017).
- (5) **Phung N. H.**, Kubo M., Sato H., Emergence Metastable Swarm Behaviors using Heterogeneous Bias values, In *Proceedings of the 4th Asian Conference on Defence Technology* (2017).
- (6) **Phung N. H.**, Kubo M., Sato H., Improvement of Bias and Rising Threshold Algorithm based on Local Information Sharing, In *Proceedings of the 10th International Conference on Agents and Artificial Intelligence*, Vol. 2, pp. 339-344 (2018).
- (7) **Phung N. H.**, Kubo M., Sato H., El Farol Bar problem by Agreement algorithm based on trial and error behavior at the macro level, In *Proceedings of the 22nd Asia Pacific Symposium on Intelligent and Evolutionary System*, pp. 21-28 (2018).
- (8) **Phung N. H.**, Kubo M., Sato H., Efficient Searching by Bias and Raising Threshold Algorithm Using Multiple Voting in the Best-of-n Problem. In *Proceedings of the 4th International Conference on Robotics and Intelligent Technology* (2019).

口頭発表

- (1) **Phung N. H., Kubo M.**,ロボカップジャパンオープン 2016,於：愛知工業大学 八草キャンパス(ロボット競技会参加).

表彰

- (1) Best student paper award finalists in the first International Symposium on Swarm Behavior and Bio-Inspired Robotics (October 2015).
- (2) Best presentation award in 2nd International Conference on Robotics and Intelligent Technology (May 2019).