

分散人工知能における標準的小問題

大沢 英一 沼岡 千里 石田 亨* † ‡ § ¶

分散人工知能におけるエージェント・アーキテクチャや協調問題解決のモデル化に関連したいくつかの標準的小問題 (Tileworld, 追跡問題 (Pursuit game), 囚人のジレンマ (Prisoner's dilemma)) を取り上げる。それにより、当該研究領域における目的と問題点を整理し、また、これまでに提案されてきた解法やアプローチを比較することにより、それらの諸性質と有効性を考察する。

1 はじめに

分散人工知能の目的は、人間の集団や社会などが持つ、さまざまな知的な振舞いを生み出すメカニズムをモデル化し、また、それに基づいてシステムやアルゴリズムなどを構築することである。ところで、分散人工知能研究の歴史を振り返ってみると、それは必ずしも何らかの統一的な基礎理論から出発して発展してきたものではない。むしろ、具体的な応用分野において問題を発見し、それらをモデル化し、そして理論や解法を与えてきたという傾向が顕著にみら

れる。それらの応用分野には、分散センシング、ネットワーク管理、航空管制、協調ロボット、分散スケジューリングなどがある。これらの応用分野において考慮された問題は、比較的大きくて複雑なものであり、また追試を行なうことが困難であった。

よって、これまでに分散人工知能研究が達成してきたことを検討しようとする、まず、(1) 当該研究領域における問題点が何なのかということに関して、十分に一般化された問題領域がなかなか見当たらない。また、(2) そこで開発されてきた解法、方法論の有効性、そして一般性をはかることが困難である。そして、(3) 追試が困難であるために議論がかみあわず、さまざまなモデルが提案されるが評価することが不可能になっている、などの問題に直面する。これらの反省から、誰もが理解でき、類似の実験状況の設定が容易な小問題が必要であるという認識が生まれつつある。

さて、分散人工知能には以下の三つの基本的な研究課題が存在する¹。

エージェント・アーキテクチャ: 複数のエージェントが独立に動作するような環境は、個々のエージェントにとっては動的な環境とみなせる。動的で予測困難な環境で動作するエージェントのアーキテクチャとしてはど

*Canonical Problems in Distributed Artificial Intelligence.

†Ei-Ichi Osawa, Chisato Numaoka, (株) ソニーコンピュータサイエンス研究所, Sony Computer Science Laboratory Inc.

‡Toru Ishida, 日本電信電話 (株) NTT コミュニケーション科学研究所, NTT Communication Science Laboratories.

§コンピュータソフトウェア, Vol.10, No.3(1993), pp.3-20.

¶1993年3月5日受付.

¹分散人工知能の分野にはこれら以外にも、エージェント間の信念の不整合の問題やエージェント間通信における common groundings (共有知識/共有語彙など) の獲得問題など、さまざまな観点から提案された多くの基本的研究課題がある。これらに関する広範なサーベイとしては [3] などを参照されたい。

のようなものが優れているのか。

協調的組織スキーマ：複数のエージェントが共通の目標を達成する過程は協調問題解決と呼ばれる。共通の大域目標を持つ複数エージェントの集団において目標を効率良く達成するためには、各エージェントが個々の局所的な戦略と大域目標達成のための戦略をどのようなスキーマにより整合させるのが好ましいのか。

非協調状況における戦略：エージェント間の利害関係は、協調(上記の協調的組織のように利害が一致する場合)および非協調といった状況に分類される。非協調状況はさらに、競合(利害が完全に対立)と利害が完全には対立しない状況に分類される。これらの非協調状況において、(1) 自己の利益を最大にするためにはどのような戦略に従って意思決定をするべきか、また、(2) さまざまな戦略をとるエージェントの集団は、長期的な観察においてどのような趨勢をたどるのか。

本稿では、上記の三つの研究課題に対して、それぞれ Tileworld [14]、追跡問題 [2, 18]、囚人のジレンマ [6] という標準的小問題²を取り上げる。これら三つの小問題に対しては、複数の研究者が独立の論文においてさまざまな観点からの研究成果を論じている。よってこれらの問題を通して、

- 当該研究領域における目的、中心課題などを整理する
- これまでに提案されてきた解法や方法論を考察し、それらの諸性質と有効性について考察する

²本稿でこれらの問題を標準的と呼ぶ理由は、それらが分散人工知能研究の基本的課題をとらえており、また、それら問題に関して独立の研究者/研究機関による複数の研究が行なわれていることによる。

- 今後の研究の方向性を探る

ことを試みたい。

本論文の構成は、まず第2章でエージェントのアーキテクチャを評価するための実験環境 Tileworld について述べ、第3章と第4章では、それぞれ追跡問題と囚人のジレンマを通して協調の問題に焦点を当てる。2, 3, 4の各章では、まず前述の各問題を紹介し³、それらの主性質と分散人工知能における研究目的について述べ、続いて各問題に関するいくつかの関連研究について述べる。各章の最後では、それらの関連研究の関係をまとめた後に各問題に関する考察を述べる。最後の章では、各章における考察を分散人工知能研究の観点からまとめ、さらに今後の展望について述べる。

2 Tileworld – 動的環境におけるエージェント・アーキテクチャ –

2.1 問題の定義

Tileworld[14]⁴はチェスボードのような2次元格子状上に、エージェント(これまで報告されている範囲では1個)、タイル、障害物および穴(これらは一般に複数個)を配置した、計算機上の仮想的な実験環境である(図1参照)。

エージェントは一つのセル上にちょうど収まる大きさで、上下左右に単位時間当たり1セル動くことができる。タイルも一つのセルにちょ

³Tileworld と追跡問題に関しては、できる限りこれらが提唱された原論文に基づいて定義することとする。しかしながら、これらの問題を提唱している原論文においても、同一実験を厳密に追試するに十分なほど、パラメタなどに関する詳細な指定があるわけではない。また、本稿で紹介しているいくつかの関連研究事例においても、オリジナルの問題設定を多少変更して(または簡略化して)結論を引き出している場合などがある。必要な変更点については各研究事例の紹介の中で述べることとする。

⁴Tileworld はワシントン大学の FTP サイト (cs.washington.edu) から “/pub/new-tileworld.tar.Z” を get することにより入手可能である。

```

# # # # # # # # # # # # # # # # # # # # # # # # # # #
# #   T   T   #   #   T   #   T   #   #   #   #   #   #
# #     #                 2 2   T #
# #     #                 2     #
# #     # # 5          T     #
# #     # # 5 T     #     #
# #     # # 5 T     a     T #
# #     T           #     #
# #     T       T       T   T T #
# #           #   T # T   T     #
# #           # # # # #   T     #
# #           # T   T T     #
# #           #           T     #
# #           #           # # # # #
# #           T       T T     #
# #           # # # # #   T T     #
# #           #           T T     T #
# # # # # # # # # # # # # # # # # # # # # # # # # # # #
a=agent,#=obstacle,T=tile,<digits>=hole

```

図 1: Tileworld の初期状態の例 (Martha Pollack 氏の好意による)

うど取まる大きさで、エージェントに押されることによって、隣接するセルへと滑るよう移動する。障害物はいくつかのセル上に位置し、移動することはない。穴もいくつかのセルから成り、タイルによって埋めることができる。すなわち、穴と同じセル上にタイルが運ばれると、穴とタイルは消滅し空白のセルに戻る。一つの穴を構成する全てのセルが埋められると、エージェントに得点が与えられる。得点は穴ごとに予め決められており、エージェントもそれぞれの穴の価値を知らされている。

穴、タイル、障害物はシミュレーションに先だって設定されたパラメータ (穴の出現率、障害物の出現率、穴の得点のばらつき、穴の大きさとタイルからの距離のばらつき等) に従って現われたり、消えたりする。Tileworld でのエージェントの目標は、穴を埋めることによって可能な限り多くの得点をあげることである。

2.2 問題の背景および目的

人工知能における (狭義の) 問題解決とは、状態空間の中で初期状態から出発し、状態遷移を生じさせる動作の実行を繰り返すことによって、

目標状態に至ることである。古典的なプランニングは、目標状態に至る動作系列を実行前に求めるもので、多大な計算量を要する問題である。したがって、動的な世界に反応しなければならない場合には、古典的プランニングとは異なる、より瞬間的に反応し行動するプランニング手法が考案されなければならない。この種の研究は、実時間プランニングとも即応的 (reactive) プランニングとも呼ばれ活発に討論されてきた ([20, 21])。Tileworld は、まさしくこのような研究の流れを背景に提案されている。

Tileworld で動作するエージェントは変動する環境の中を動き回り、タイルを運び穴を埋めることが要求される。すなわち、限られた時間でプランを構築し、判断を下さなければならない。Tileworld は、このような資源が制約された (resource bounded) エージェントのアーキテクチャを評価する環境として提案され、利用され始めている。

2.3 研究事例: 熟考と反応のトレードオフ

IRMA (the Intelligent Resource-Bounded Machine Architecture) は動的な環境におけるエージェントのアーキテクチャを目標としたもので、Bratman ら [4] により提案され、Pollack ら [14] により Tileworld 上で評価された。IRMA では、環境の変動に反応するために、エージェントが実行中のプランは再考されたり、必要があれば廃棄されたりする。しかし一方、環境の変動への過剰な反応を抑えるために、エージェントのプランは適切なレベルで安定していなければならない。したがって IRMA では、即応 (reactiveness) の対極の概念として熟考 (deliberation) が導入されている。熟考は多くの可能性 (選択候補) の中から何を追求すべきかを判断するプロセスである。

IRMA の大きな特徴は、熟考の対象となる選

択候補を制限するためのフィルタリング機構を備えていることである。フィルタリング機構は、(1) 新たに生成された選択候補とその時点で実行されているプランが両立 (compatible) するかどうか、また、(2) 両立しない場合には、現在実行中のプランを廃棄 (override) すべきか否かを調べる。このフィルタを通過した選択候補だけを熟考の対象とすることによって、妥当なレベルでの安定性を保証しようとする。

実験では、フィルタは以下のように構成された。

両立検査: ごく単純なものが用いられた。新たな選択候補は、“どこどこの穴を今すぐ (あるいは後ほど) 埋める” という形をしており、新しい穴が発見された時に生成される。両立検査では、今すぐ埋めようとする穴が、現在エージェントが埋めつつある穴と異なる場合に限り両立しないとする。すなわち、穴を埋めている最中は、“今すぐ埋める” という形の新たな選択候補は考えない。両立検査をパスした選択候補は熟考の対象となる。

廃棄検査: 両立しない選択候補に対して行なわれる。両立しないと判断された選択候補の穴の得点と、現在埋めつつある穴の得点を比べ、もしその差がフィルタのしきい値を超えていれば、その選択候補を熟考の対象とする。このしきい値は予めパラメータとして与えられる。例えばしきい値をゼロとすると、現在埋めている穴より得点の大きな穴は全て熟考の対象となる。新しい穴に関しては、それが消滅するまでの時間が長い (すなわち、埋める時間的余裕がある) と考えると、しきい値をマイナスに設定してもよい。なお、しきい値を $-\infty$ に設定すると、全ての選択候補が熟考の対象となる。

一方、熟考プロセスは以下の2種が用意された。

- 得点のより高い穴が選択される (本来この程度では熟考とは言えないが、評価のために考えられたものであろう)。
- 労力当たりの得点 (Likely Value: LV) のより高い穴を選択する。

$$LV(h) = \frac{score(h)}{dist(a, h) + \sum_{i=1}^n 2 \times dist(h, t_i)} \quad (1)$$

ここで、 $score(h)$ は穴の得点、 $dist(a, h)$ はその時点のエージェントの位置と穴の距離、 n は穴を埋めるのに必要なタイルの数、 $dist(h, t_i)$ は穴から i 番目に近いタイルまでの距離である。

環境の変動 (穴や障害物、タイルなどの出現/消滅) やエージェントの移動速度に対して、思考速度を変化させて初期的な評価が行なわれた。評価結果として、(1) 思考速度が速い間はフィルタのしきい値は低くてよい (すなわち、多くの選択候補を熟考の対象とする) が、思考速度が遅くなるとフィルタのしきい値を上げていく (すなわち、熟考の対象を絞り込む) 方が効率がよい、(2) しきい値を適切に調節すると、LVを計算する手法の方が単純な手法に比べ効率が良くなる、と報告されている。

2.4 研究事例: コミットメントの効果

コミットメント (commitment) はエージェント間の相互作用を考える上での、大切な概念として研究されている。例えば Cohen らはコミットメントを、達成されるまで持ち続ける目標 (goal) であると定義している [5]。また Kinny らは、エージェントが一度決めたプランをどの程度持ち続けるかを、プランに対するコミットメントの度合 (degree of commitment) と呼んでいる [11]。

Kinny らは、3種のエージェントを導入し評価の対象としている。大胆な (bold) エージェント

は決してプランを変更しない。反対に用心深い (cautious) エージェントは常に (1 動作ごとに) 再プランニングを試みる。その中間的なものとして、ノーマル (normal) なエージェントを導入する。

この実験では障害物やタイルは考慮されていない。穴はエージェントが到達した時点で埋められたと見なされ消滅する。したがって、もともとの Tileworld より、かなり単純化された問題を対象としている。評価の軸は以下の二つである。

環境の変動率: Tileworld が変化するクロック周波数の、エージェントが動作するクロック周波数に対する比率である。変動率が高くなるにつれて、穴が出現消滅する頻度が増大する。これは、問題解決過程で、穴が突然消えることによって目標が消失したり、近くに穴が出現することによってより良い目標が生まれたりすることを意味する。

エージェントの効率: 全部の穴を埋めた場合に与えられる総得点の内、エージェントがどの程度の得点をあげたかを示す比率である。これによって、そのエージェントのアーキテクチャの良否を評価する。

実験はエージェントのプランニングに要する時間をさまざまに変化させながら行なわれた。得られた結果は以下のとおりである。

- 環境の変動がなければエージェントの効率はプランニング時間の長短にかかわらず 1 であるが、変動速度が大きくなるにつれてこの値は減少し、プランニング時間の長短にかかわらず 0 に収束する。
- プランニングに要する時間が長くなるにつれ、大胆なエージェントが用心深いエージェントより、(環境の変動率に関わらず) 効率が良くなる。

- プランニングに要する時間が短くなるにつれて、環境の変動率が大きな領域で変化が起こりはじめる。まずノーマルなエージェントが大胆なエージェントより効率が良くなり、ついで用心深いエージェントが大胆なエージェントを上回るようになる。

2.5 まとめ

人工知能分野においてモデル化の果たす役割は大きい。しかし同時に、提案されたモデルは、いかに人工的に知能を構成するのに寄与するかという尺度で評価されなければならない。Tileworld の提案は、エージェントアーキテクチャの定量的評価に一步踏み出したという点で評価されるべきだろう。これまで得られた結果は比較的自明なものばかりであるが、研究は始まったばかりであり今後期待したい。

Tileworld を理解しにくいものになっているのは、問題設定の複雑さである。穴の出現/消滅は目標の出現/消滅を意味する。障害物の出現/消滅はプランが最適な状態であり続けることがないことを意味する。穴の得点は目標の優先度を表す。穴を複数のタイルによって埋められなければならないことにしているのは、目標を達成するためのコストに変化を持たせるためだろう。この例題を研究に用いる場合には、研究の目的に合わせて、Tileworld から部分問題を切り出すべきであろう。また研究例は報告されていないが、Tileworld を多数のエージェントがいる環境に拡張することは容易であり、分散人工知能のエージェント・アーキテクチャを評価する例題として利用することも可能である。

3 追跡問題 – 協調のための組織的 枠組 –

3.1 問題の定義

無限に広がる2次元格子状に、四つの青いエージェントと一つの赤いエージェントがいる。エージェント達は、2次元格子状を縦横方向に、一度に1セルずつ移動することができる。もちろんしばらく動かないでいることも可能である。1サイクル内に、全てのエージェントが逐次的に動くものとする。青いエージェントは、他の青いエージェントのいるセルには移動できるが、赤いエージェントのいるセルには移動できない。各エージェントの視界は限られており、赤いエージェントを見ることはできるが、他の青いエージェントを見ることはできない。後述する組織構造の制約の範囲内で、通信を行ない他の青いエージェントの位置を知ることができる。ここで四つの青いエージェント達に与えられた使命は、なんとかして赤いエージェントを四方から取り囲むことである。この際、青いエージェント達は協力して赤いエージェントを囲い込むことができる。

3.2 問題の背景および目的

常に動作環境が変化するようなシステムにおいては、問題解決のための全体的な制御を行なうことが非常に困難である。このような環境で独立に動作するエージェントは、能力において限られており、世界に関する部分情報しか得ることができない。このような状況で、エージェントが単独で解決できない問題に直面した時に、協調作業による解決が重要な意味を持つてくる。協調作業に関して考慮されるべき問題の一つは、エージェント間の相互作用の形態である。追跡問題は、このような背景の中で、エージェント間の組織的関係や通信が協調的作業に与える影響

について調査することを目的として提案された。

追跡問題における一つの課題は、青いエージェント間の制御関係を規定する組織構造である。青いエージェントはもちろん、単独で推論し、自分の行動戦略を決定することもできる。この推論の中には、他の青いエージェントが次にどこに移動するかという予測も含まれるであろう。全てのエージェントが全く同様の推論をし、同様な結論を導き出すことができるのならば、このように単独で推論し行動することに何の問題もないであろう。しかし、各エージェントの情報が異なるような状況では、同様な結論が引き出せるとは限らない。他の青いエージェントが、予測に反して行動することもあり得るわけである。そこで、効率の良い問題解決をもたらすために、四つの青いエージェントが、互いの合意を確認しながら行動するための枠組が必要とされる。

追跡問題におけるもう一つの課題は、青いエージェント間の情報交換の形態、さらには情報交換の効率である。青いエージェント達の視界は限られているので、もし単独で問題を解決する場合には、赤いエージェントを探し回らなければならない。しかしながら、もし青いエージェントが通信することによって情報を交換できるならば、無駄に探し回ることなく赤いエージェントの位置を知ることが可能である。しかしながら、通信量が多くなりすぎると、それだけ行動意思決定に時間がかかり、実時間環境では、赤いエージェントに逃げられてしまう恐れもある。

以下では追跡問題を扱った四つの関連研究について、できる限りそれらの論文の記述に忠実に解説を行なう。まずはじめに、Bendaらのオリジナルの研究と Stephens らによってその後行なわれた関連研究を紹介する。これら二つの研究は異なる結論を引き出しており、それらの結論の解釈については十分に検討される必要がある。その後、Gasser らによる研究を紹介する。彼らは組織構造を、問題表現と環境に依存して

決めるべきものととらえている。最後に、Levyらが最近示したアプローチを紹介する。彼らはゲーム理論に基づくエージェントの行動に関心を持っており、この研究もその関心の一部としてとらえることができる。

3.3 研究事例: 組織構造と協調

Bendaらは[2]において、“複数の知的エージェントが協力して一つの解を求める際に、エージェントの間に存在する組織構造が問題解決の効率にどのような影響を与えるか”という問題に取り組んでいる。この問題を説明し評価する目的で、追跡問題を作り上げた⁵。

二つの青いエージェントから構成される基本的組織構造が、青いエージェント間に張られたリンクに対して行なうことのできる操作に応じて定義される。Bendaらが定義している基本的組織構造は、以下の三つである。

1. Communicating-agents(CAs): リンクに沿って、一端のエージェントが、もう一端のエージェントに対しデータを要求することができる。
2. Negotiating-agents(NAs): このリンクの両端にいるエージェントは、データを要求したり、追求するゴールを交渉できる。CAsが、相手からデータを要求する目的以外に通信することができないのに対し、NAsは、データを要求したりする他に、相手と移動に関する交渉を行なうこともできる。
3. Controlling-agents(CLAs): 二つのエージェントが有向リンクによって結合されている。このリンクの始点に位置するマスタエージェントは、終点に位置するスレーブエージェントを完全に制御することができ

⁵後に Stephens と Merx が [18] において、この問題に対し “Pursuit Game” という名称を与えている。

る。原論文には説明がないが、マスタエージェントはスレーブエージェントにデータを要求できるものと考えられる。

二つ以上からなる一般的なエージェントの組織は、以上三つの基本構造の組合せによって決定される。例えば、エージェント A, B, C において、A と B が CLAs であり、A と C が CLAs であれば、A をマスタとし、B と C をスレーブとする階層的組織になる。

実験によって Benda らは、“システム全体として階層的な制御系統を持つ組織の方が、問題解決において優れている”と結論付けている。彼らはその理由として階層的組織では通信量を十分削減できる点をあげている。

Stephens らは、Benda らの定義した三つの基本的組織関係に、通信を行なわないようなエージェントの組織を加えて、6つのシナリオ⁶に関して実験を行なった [18]⁷。この結果として、

- 動的に変化する目標物に対する追従性に関しては、交渉し合う組織の方が階層的に制御する組織よりも優れている、また、
- 完全に目的を遂行するためには、階層的に制御する組織が優れている、

という結論を引き出している。この結論は Benda らのものとは異なっている。ここで、少しこの差異について考えてみる。

まず、Benda らの実験で CLAs が良いとされた理由は、赤いエージェントを捕らえるまでに

⁶シナリオとは、この場合エージェントの初期配置を意味する。

⁷Stephens らの問題定義は、Benda らのものとは異なっている。まず、格子のサイズは 30×30 に制限される。Benda らのものにおいては視界に制限があったが、Stephens らのものにおいては、青いエージェントはいつでも赤いエージェントの位置を知ることができる。赤いエージェントの上下左右の位置は捕獲位置と呼ばれる。青いエージェントがひとたびこの四つの捕獲位置の内一つに到達した場合、以降赤いエージェントの動きに応じて同じ位置を保持するように移動する。

要した通信量が、他に比して少ないことにある。二つの青いエージェントが同じ場所に移動しようとしていると仮定しよう。二つのエージェントが CLAs の関係にあるとすれば、片方が制御命令を送れば問題は解決されるので、1回の送信で済む。一方 NAs の関係にあれば、例えば、自分の考えを相手に知らせるために双方が1回ずつ送信し、次に相手の考えを考慮に入れて意思を決定し、その結果を双方が知らせ合い、さらにその結果に対して合意したことを知らせるとすると、6回の送信がなされることになる。したがって、CLAs におけるエージェント達の移動量が6倍程度にならないと、通信量が NAs と同じ程度にならない。結果として、通信量だけで比較すれば、CLAs の方が優れている。

Stephens らは、これに対して、通信量⁸よりもむしろ、エージェントのヒューリスティック効率を考慮に入れた評価を行なっている。ヒューリスティック効率は、以下の二つの要素によって決定される。一つは初期局面において計算される、四つの青いエージェントから赤いエージェントまでの距離の合計 N_1 である。もう一つは、追跡が終了するもしくは追跡が決して終了しないことがわかるまでに必要とされた局面数を n としたとき、各局面において計算される、四つの青いエージェントから赤いエージェントまでの距離の合計を N_i としたときの、 $T = \sum_{i=1}^n N_i$ である。この時ヒューリスティック効率は、 $E = N_1/T$ で表現される量である。よって、急激に取り囲むことができればこの値は大きなものとなり、なかなか取り囲むことができなければ小さなものとなる。実際、Stephens らの実験においては、

⁸Stephens らの手法によると、各エージェントはまずローカルな黒板に(捕獲位置, コスト)の対をコスト順にソートして書き出し、この黒板上のデータを他の全てのエージェントに送る。後は、こうして集められたデータを元に、特定のアルゴリズムによって各エージェントに対して捕獲位置を割り当てる。このように、Stephens らの設定では、最初に共通知識を作り出すためのみ通信が起こる。

この効率において、NAs は CLAs に勝るという結果が得られた。

CLAs の利点は、一つのエージェントが他の全ての動きを制御できることにある。だが一方で、単一のエージェントの視点のみを反映して、全ての青いエージェントの動きを決定してしまうので、交渉を通じてより良い可能性を見つけるという機会を逃してしまう可能性もある。Stephens らの実験結果は、まさにこうした事情を反映していると考えられる。

3.4 研究事例: 組織的知識の表現と制御構造

Gasser らは、分散知的システムのための一般的協調機構に向けての彼らの研究状況に関する報告の中で、追跡問題を引用している [8]⁹。彼らは、これまでの分散人工知能システムにおける組織構造には、状況変化に適応するような柔軟性がなかったことに着目し、より柔軟性のある組織構造を持つ分散人工知能システムの枠組を提案しようとしている。彼らの視点において、“組織は、エージェント集団の間に存在する構造的関係として、もしくは、エージェントの活動を外から制約するものとして存在するのではなく、(これまでの問題解決を通じて作り出された) エージェント自身の持つ信念、期待、および約束¹⁰の中に概念的に作り出されたもの”である。エージェントはある組織における問題解決活動の中で、関連する他のエージェントが期待通りに行動しているならば、いつも通りに行動することができる。ところが、自分の期待と異

⁹問題設定は [18] に準ずる。

¹⁰期待や約束は信念の一部である。エージェントは信念において、自分が解決すべき部分問題をフレームによって表現している。フレームの中には、エージェントの世界に対する期待を表現している“制約”や、約束ごととして処理されるべき“ルーチン”が記述されている。これらフレーム表現の記述が、組織というものを概念的に作り上げる。

なる動きをするエージェントがいる場合、新しい問題解決行為が発生する。Gasserらは、期待に応じて（概念的）組織が形成されると考えているので、この新しい問題解決の必要性は、新しい（概念的）組織構造の形成、すなわち組織変更を引き起こすことになる。以上のような考えをテストするために、Gasserらは知的協調実験環境 (ICE) と呼ばれるテストベッドを開発した。また、追跡問題における中盤および終盤の問題解決を題材として議論を行なっている。

Gasserらは、まず終盤の問題解決において“四つの青いエージェントが割り当てられた四分円 (quadrant)¹¹ から出ないようにして、赤いエージェントとの距離を狭めることができるならば、かならず赤いエージェントを捕らえることができる”という事実があることに着目する。この事実に着目すると、青いエージェント達がゲームの終盤において解かねばならない問題は、“自分を含む四つのエージェントに対し四分円の割り当てを行なう”ことであり、次にこの“割り当てられた四分円から出ないようにして、その距離を狭める”ことである。この状況を、“青いエージェント達が Lieb configuration にいる”という (図2 参照)。

ここでは上の問題を解くために起こる処理過程を説明する。簡単のため、二つの青いエージェント（四分円 Q_1 にいる B_1 と Q_2 にいる B_2 ）と一つの赤いエージェント（ R_{a1} ）を想定する。ICEにおける問題表現はフレームによって与えられる。脚注でも述べたように、このフレーム表現はエージェントの信念であり、概念的に組織構造を規定している。各四分円には EndGame フレームが割り当てられている。このフレームでは、その四分円においてどの青いエージェン

¹¹まず、赤いエージェントを中心とする円を考え、赤いエージェントに対して前後左右に扇型ができるようにこれを分割する。前後左右はそれぞれ赤いエージェントが移動できる方向である。ここでいう四分円とは、このようにして作られた扇型を指している。

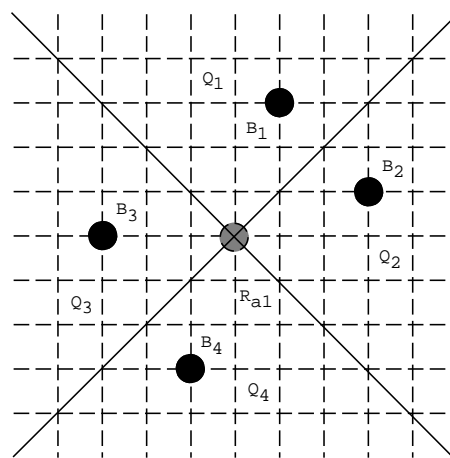


図 2: Lieb configuration の例 (Les Gasser 氏の好意による)

トが赤いエージェントをブロックするべきかという問題を解決する。このフレーム内の制約やルーチンが、この場合に必要とされる組織構造および組織内で期待されるエージェントの役割や行動を規定している。この問題解決に付随して Follow-Lieb-Rules という別の問題解決フレームがある。このフレームでは、割り当てられた四分円を、エージェントの存在すべき領域に関する制約として持っており、現在自分がその四分円にいるかどうか調べ、そうでないならば、制約を保つ状態を回復するような行為を引き起こす。この場合、四分円 Q_1 の Follow-Lieb-Rules フレームにおける制約は $in(B_1, Q_1)$ である。これはエージェント B_1 が四分円 Q_1 にいるべきであるという期待を示している。

例えば、上の初期状況（四分円 Q_1 に B_1 がおり、 Q_2 に B_2 がいる状況）において、EndGame フレームの問題が解決されたとしよう。この結果、エージェント B_1 と B_2 はそれぞれ赤いエージェント R_{a1} の動きを予測し、その結果赤いエージェントを捕獲できるような方向に移動するであろう。この結果、四分円 Q_2 に B_1 と B_2 が同時にいるような状況が生じたとしよう。これは四

分円 Q_1 の Follow-Lieb-Rules フレームにおける制約 $in(B_1, Q_1)$ に違反している。つまり、期待と異なる行動をするエージェントが組織内にいることになる。この結果、エージェント B_1 が Q_1 に戻るといふ新しい問題解決の必要性が生じ、その問題解決に相当する組織を規定するフレームが参照されることになる。

別の例においては、エージェント B_1 と B_2 が共に予測を誤り、互いに存在する四分円を交換している (つまり Q_1 に B_2 がおり、 Q_2 に B_1 がいる) ような状況に対する問題解決が示される。ここでは二通りの解決が可能である。一つは、前の例で見たように、元の四分円に戻るといふ行為を、 B_1 と B_2 が共に行なうことである。そしてもう一つは、二つの四分円は共にただ一つのエージェントによって占有されていけばよいという理由によって、特別な問題解決は何もしないでそのままにしておくというものである。この解法は、組織における役割を変更することに相当する。例えば、そのままの方がコスト的に安価であるならば、この二つめの解法が採用されることになろう。この場合、Follow-Lieb-Rules 中の制約条件の書き換えが必要とされる。

Gasser らの議論は、組織形成、変更の必要性が、問題表現と環境の変化に依存して発生することを示唆している。これは Benda らや Stephens らの実験では全く考慮されていない点である。彼らは、組織構造が静的に与えられ決して変化しないことを仮定している。結果として、彼らの結論は、エージェントが置かれている状況を無視したものとなってしまっている。Gasser らのように、問題表現と環境の変化ということを考えて入れて評価することが、今後望まれるであろう。なお、Gasser らは、これまでに述べたような状況変化に適応する協調組織のモデルを提案しているが、評価結果に関しては述べていない。Gasser らのモデルとは異なるが、大沢は、追跡問題における適応的組織スキーマの問題解

決コストに関する評価結果とその有効性について報告している [19]。

3.5 研究事例: ゲーム理論的アプローチ

Levy らは、同じ追跡問題を題材として、ゲーム理論的戦略に基づくエージェントの組織を考察している [12]¹²。さらに Stephens らの研究で用いられた 6 つのシナリオを利用して、考案された組織を用いた追跡問題の結果を評価している。ここではその解法について紹介する。Levy らの目的は、彼らがこれまで行なってきたゲーム理論的方法論が、できる限り通信を行なわないで、自律的に追跡問題を解決するエージェントによって、どのように利用されるか示すことにある。具体的には、協力ゲームを非協力ゲームと組み合わせた方法が利用される¹³。Levy らの研究では、動的に形成される提携 (coalition)¹⁴ が組織ということになる。

各エージェントは可能な 5 種類の戦略 (上下左右に動く、または動かない) の内から、以下に示す手順で次に実行する戦略を決定する。

1. 全てのエージェントによるあらゆる戦略の組合せに対して、全ての提携に対する利得を計算する。この際、利得計算には、(1) 四つの青いエージェントが全体として、赤いエージェントからどれだけ離れているか、および、(2) 赤いエージェントの脱出方向 (上

¹²問題設定は [18] に準ずる。

¹³協力ゲームとは、コミュニケーションが可能な状況において、各プレーヤがそのとるべき戦略を合意のうえで決定するゲームである。非協力ゲームとは、このような合意による決定ではなく、各自の独立な判断により戦略を決定するゲームのことである。([23] より引用)

¹⁴協力ゲームにおいて、プレイヤーの集合を N としたとき、 N の任意の部分集合を提携と呼ぶ。空集合や、プレイヤー全体からなる集合も提携と考える。一つの提携に属するプレイヤーの間では、十分なコミュニケーションが可能で、提携メンバがどのような行動をとるか十分に協議した上で決定するものとする。各提携は提携としての共同戦略をもって、その中から一つを選択する。([23] より引用)

下左右)をいくつ塞いでいるか、の2点が考慮されている。提携がとる戦略の組に対する利得が計算された後、シャープレイ値(Shapley value)¹⁵と一致するように、その利得を各エージェントの貢献度に応じて分割する。この分割された利得を別払い(side payment)と見ると、ここまですべてが別払いのある協力ゲーム¹⁶に相当する。これらの作業の結果として、四つのエージェントによる可能な戦略の組全てに対して、各エージェントが手にする利得を表現した利得行列が生成される。

2. 1で計算された利得行列を用いて、他の青いエージェントがどの戦略をとるかわからないという前提の下で、非協力ゲームにおける均衡点を求める手法で、次にとるべき戦略を決定する。

この二つのステップによって行なわれていることの意義を説明すると次のようになる。まずより早く赤いエージェントを取り囲むことが目的であるために、利得計算が四つの青いエージェントと赤いエージェントの配置を考慮したものになっている。このために、青いエージェントが各局面でとる戦略に対する利得は、これらの配置によってさまざまに変化する。あらゆる配置の状態をあらかじめ持つておくことができないため、利得行列は各局面で動的に生成される必要がある。この際、各エージェントの持つ利得行列が異なると、ゲームの扱いが難しくなる

¹⁵エージェント i の全ての提携に対する貢献度の加重平均を示したもの。([23]より引用)

¹⁶一般に、利得あるいは効用が何らの制限なく分割可能であり、利得の一部をエージェント間で譲渡でき、かつエージェント A が B に譲渡した場合、 A の失った利得の大きさが、 B の得た利得の大きさに等しい時、そのゲームは譲渡可能効用 (transferable utility) を持つという。そして利得の一部を譲渡すること、およびその譲渡された利得を別払いという ([23]より引用)。別払いのある協力ゲームとは、この譲渡可能効用があるという前提を持つゲームのことをいう。

ため、利得行列の計算においては、協力ゲームの理論を用いて、同様な利得行列が計算されるように配慮されている。これが一つめのステップでなされていることである。次に二つめのステップでは非協力ゲームの理論が用いられているが、これは通信の利用を押えるためである。

すでに何度も述べたように彼らの実験目的の一つは、できる限りエージェント間の通信を減少させることである。実際、青いエージェントが非協力ゲームを利用して動くことによって、ほとんど通信をすることなく動き回ることができるという利点がある。しかし、実験を通じて、通信が全くない状態では、常に追跡を成功させることが必ずしも可能ではないという結論も下している。特に、非協力ゲームにおいて複数の均衡点が存在する場合、通信によってより良い戦略を選ばせることが可能な例が示されている。この結論は、Rosenschein らによる論文 [15]¹⁷におけるものと同様である。

3.6 まとめ

本章では、協調のための組織的枠組を考えるための基本的問題として“追跡問題”を例として取り上げ、いくつかの研究事例について解説を行ってきた。変化し続ける環境において、各エージェントは個々の問題解決活動の中で、それを助けるものとして組織を必要とする場合がある。よって、Gasser らのいうように、エージェントの組織における役割が、エージェントの信念における必要性に応じて決まってくるような問題解決の枠組が重要性を増すことになる。

Benda らの結論では、制御系の整ったシステ

¹⁷Rosenschein と Genesereth は、[15]において、利得関数に基づいて次の手を決定するエージェントが、通信を用いて相手と次の手に関する交渉を行なうモデルを提案している。この研究において、彼らは囚人のジレンマを取り上げ、交渉を行なうことによって、この問題においてジレンマとされている状況を回避することができ、最良の手を選ぶことができることを示した。

ムの方が問題解決能力において優れていることが示された。Stephensらは目的の動的な変化というものを考慮して評価を行なった結果、交渉能力が柔軟な問題解決に役に立つという結論を示している。Levyらの研究では、組織構造よりも、エージェント自身の行動決定能力というものが全面に押し出されるようになる。ゲーム理論的な手法では、組織が提携という形で表現される。提携内では、提携全体の利益、提携内での利得のやりとりといったものを仮定して個々のエージェントが行動を決定することにより、ほとんど通信を行なうことなくほぼ合理的な組織行動をすることができる。しかし、より良い問題解決においては通信能力が必要となる。

追跡問題は問題設定が単純であるため、エージェントの能力を考えることによってさまざまな変形が可能である。このため、組織的活動に関して何を議論したいかによって、その場に応じた設定を行なえばよい。本稿では紹介できなかったが、DurfeeらによるMICE [7]¹⁸は追跡問題の発展形として興味深いものである。MICEでは、赤や青のエージェントの数が増加した場合や、赤や青が食い合ってしまう場合なども考慮されている。

4 囚人のジレンマ – 利害が完全には対立しない状況での協調的行動選択 –

4.1 問題の定義

囚人のジレンマ¹⁹: 2人プレーヤによるゲーム。

¹⁸MICEはミシガン大学のFTPサイト (freebie.engin.umich.edu) から “/pub/MICE/Mice.tar.Z” を get することによって入手可能である。

¹⁹囚人のジレンマは、1950年頃 M. Flood と M. Dresher によって発案され、後に A.W. Tucker によって定式化された。これは元来分散人工知能研究とは異なる文脈において提案されたものである。しかし、その問題は分散人工知能が解決しようとする一つの基本課題をうまく

各プレーヤは**協調**、もしくは**裏切り**という二つの選択肢をとることができる。互いに、相手が次にとる行動を知らないままに、自分の次の行動を選ばなくてはならない。両者が協調した場合の利得は双方に R 、両者とも裏切った場合は双方に P 、一方が協調し他方が裏切った場合は、協調した方に S 、裏切った方に T の利得を与える。ここで各利得 T, R, P, S は $T > R > P > S$ および $R > (T + S)/2$ を満足するようにとる。このゲームを1回だけ、または繰り返し行なう場合、それぞれどのような行動選択の戦略を取るべきか？

4.2 問題の主性質および目的

囚人のジレンマのように、2人のエージェントがいて各エージェントが二つの動作を選択できるゲーム(これを 2×2 ゲームと呼ぶことにする)における双方の利得は 2×2 の行列で表現することができる(例えば図3。この行列の各要素において、\ の左側の値はプレーヤ J の、そして右側はプレーヤ K の利得を意味する。)

$J \setminus K$	協調	裏切り
協調	$R = 3 \setminus R = 3$	$S = 1 \setminus T = 4$
裏切り	$T = 4 \setminus S = 1$	$P = 2 \setminus P = 2$

図3: 囚人のジレンマの例

一般に 2×2 のゲームは、各プレーヤにおいて利得の大きさを相対化する(例えば利得は少ない方からそれぞれ1, 2, 3, 4という相対的な値をとるようにする)と、それは144種類の異なるケースに分類される。囚人のジレンマはその中の一つのケースである。

囚人のジレンマにおける利得に関する最初の制約($T > R > P > S$)は、一方的に裏切った
反映しているために、何人かの分散人工知能研究者により考察対象として取り上げられた。

場合の利得 T が最高で、裏切られた場合の利得 S が最低となることを示している。また、協調しあうときの利得 R は、裏切りあったときの利得 P より大きいことも示している。

2 番目の制約 ($R > (T + S)/2$) は、両者が交互に搾取しあっても (つまり、今回は一方が裏切り、他方が協調する。そして次回は双方その逆の戦略をとるといふように、両者の裏切りと協調の位相が完全にずれた場合)、毎回両者とも協調したときより平均利得が低いというものである。

囚人のジレンマではその利得行列から分かるように、相手がどう出ようと、自分の方は協調するより裏切った方が必ず得である。しかし、両方とも裏切りを選ぶと、両方が協調しあう場合に比べ損をしてしまうところにこのゲームがジレンマと呼ばれるゆえんがある。このゲームはチェスなどのゲームとは異なり、プレーヤどうしの利害は完全には対立しない。

囚人のジレンマを分散人工知能研究の観点からみると、非協調状況の中で、特に、複数エージェントの利害が完全には対立しない状況における、さまざまな行動規範(戦略)に関する考察のための基本的問題として捉えられる。また、囚人のジレンマでは“互いに、相手が次にとる行動を知らないままに、自分の次の行動を選ばなくてはならない”としている。これは、通信が制限されたある状況下で、複数エージェントが意思決定を行なう場合の協調の役割、そして有効な戦略について考察するための問題であるとみなすこともできる。

このゲームは非零和非協力ゲームと呼ばれており、最適戦略(相手がどのような戦略をとろうとも、自己の利得を最大化できる戦略)は存在しないことが知られている。戦略の有効性は相手がとる戦略に依存する。またゲームを1回しか行なわないのか、または繰り返し行なうのかによっても、有効な戦略への考え方が異なる。

1 回きりの囚人のジレンマ・ゲームにおいては、相手の戦略にかかわらず、自分の利益を最大化しようとする、(裏切り、裏切り)という両プレーヤの戦略の組で均衡することがゲーム理論的分析により分かっている(例えば [23])。これは、ケース解析による行動決定の例として知られている。しかし利得行列からすぐ分かるように、この均衡戦略は両者が協調しあったときよりも損になっている。

それでは、両者が協調を選択することが可能となるための条件はなんだろうか？ これを探るのが本節の目的である。以下では、まず、1回の対戦で両者が協調を選択することを可能とするための条件に関する Genesereth らの研究 [9] について述べる。次に、このゲームが繰り返し行なわれる場合に関する Axelrod の一連の研究 [1] を紹介する。

4.3 研究事例：協調における合理性の役割

Genesereth ら [9] は、 2×2 ゲームに関する合理性(最適戦略を選択するための特性)についての研究を行なった。先に述べたように、 2×2 ゲームは 144 種類の異なるケースに分類される。Genesereth らは、それら 144 種類の 2×2 のゲームのうち囚人のジレンマを含む 117 種類をカバーする、合理性に基づいた各種の行動選択方法について報告している。

Genesereth ら以前の分散人工知能の研究では、複数のエージェントが同一設計者の命令の下で相互に協調することを仮定している研究が比較的多かった。つまりこれらの研究では、エージェントは競合することがなく、エージェントが複数いた場合にはそれらのゴールは同一のものか、もしくは矛盾しないものであり、エージェントは自由に相互に助け合うとしていた(このようなエージェントは慈悲深い (benevolent) エージェントと呼ばれる)。従来の分散人工知能の研究は

このような仮定の下で、(1) 行為の同期、(2) 効率の良い通信、(3) 不注意による相互干渉とそれの回避、などの問題に焦点を当てていた。

Genesereth らはこの慈悲深いエージェントの仮定を捨て、自律的で、かつ独立した動機を持つエージェント達が、自己の目標を達成(利得を最大化)するためにどのようにして相互作用すべきかについて考察した。つまり、エージェントどうしは競合したり、無競合であったり、また、利害が完全には対立しないような状況に直面する。彼らは、このようなさまざまな状況においてエージェント間の協調を実現するための合理性の役割とは何かを明らかにしようとした。また、Genesereth らはエージェントどうしが(通信機器の故障などにより)通信をしない状況を仮定している。ちなみに、Rosenschein らの研究 [15] ではこの仮定は外されている。

囚人のジレンマにおいて協調動作(両者が協調を選択する)を可能とするためには、双方のエージェントに**一般合理性**と、**共通振舞い**と呼ばれる動作依存性が要求される。一般合理性とは、ゲームにおいて戦略を選択する手続きに関して、他のどの手続きにも優越されない(より低い利得が得られるようなことがない)ような手続きを選択することである。また、共通振舞いとは、双方の立場を入れ替えれば2人のエージェントが同じ行動をとるという制約である²⁰。一般合理性の利点は、それが共通振舞いと一緒に考慮されると、ある結合動作(仮に xy とする)が他の結合動作(仮に zw とする)に優越されていることがすべてのエージェントにおいて成り立つ場合、結合動作 xy をとり除くことを可能にすることである。

囚人のジレンマでは状況は2人のエージェントに対して対称的、つまり利得行列をみると双方のエージェントにとって対称的となっているので、共通振舞いは2人のエージェントの同一

²⁰ 異質なエージェントが相互作用を持つような環境を考えた場合、この制約は充足困難なものであろう。

動作を保証する。よって、共通振舞いに従う2人のエージェントの結合動作としては(協調, 協調)もしくは(裏切り, 裏切り)が可能であるが、一般合理性に従うと、(協調, 協調)という結合動作が(裏切り, 裏切り)に優越することから(裏切り, 裏切り)を排除することになる。

4.4 研究事例：反復囚人のジレンマ・ゲームにおけるさまざまな戦略の評価

囚人のジレンマ・ゲームが反復される場合であっても、その反復回数を両プレーヤが知っているときには、両プレーヤが4.2節で述べたケース解析に基づき戦略を選択するならば、相変わらず協調関係を引き出すことができない。なぜなら、最終回では両者とも(1回きりのゲームと同じ理屈で)裏切りあうという結論が成り立つので、その前の回でも、最終回に相手が裏切るのを見越しているためにどちらも協調しようとはしない。この論法は繰り返し適用され、どんなに対戦回数が多くてもそれが有限で双方に既知であれば、この論法により前の対戦をたどっていくと最初の回まで戻ってしまうからである。

しかし、両プレーヤが対戦回数を知らないような場合には協調関係が発現しうることが知られている。本節と次節では、これに関連した研究を紹介する。

まず、反復囚人のジレンマ・ゲームにおいては、どのような戦略が有効であり、また、協調関係が成立する条件は何であるか、を考察するためにAxelrod²¹は反復囚人のジレンマ・コンピュータ選手権²²を開催した[1]。この選手権に招待された人々は、非零和ゲームに特有な戦略的可能性について十分知り尽くした14名のゲーム

²¹ Axelrodによる一連の研究は、国際関係論という分散人工知能とは異なる研究分野で発表されたものである。しかしながら彼の一連の研究は、分散人工知能の基礎として十分に評価でき、かつ重要な考察を含んでいる。

²² このコンピュータ選手権に関する考察は、文献[10]の第29章などにも述べられているので参照されたい。

理論の専門家達であった。彼らはこのコンピュータ選手権にプログラムを参加させた。参加者の各プログラムは、自分自身のプログラムや“でたらめ戦略”(協調、裏切りをランダムに選択する)を含めてリーグ戦²³を行なった。各試合における勝ち負けは総得点により決定する²⁴。

このコンピュータ選手権で優勝した Rapoport のプログラムは“しっぺ返し”(TIT-FOR-TAT)と名付けられた戦略をとるものであった。しっぺ返しは、初回は協調し、次からは相手が前回取った行為を選択するものである。この戦略は、相手が裏切らない限り自分から裏切ることはない(これを“上品である”と呼ぶ)が、相手が一度裏切ると、その次の回に必ず報復する(これを“手応えがある”と呼ぶ)。しかし、相手が裏切った後でも改心して協調にもどれば、それ以降は自分も再び協調する(これを“心が広い”と呼ぶ)という特徴を持つ。この選手権における、成績上位の八つのプログラムはどれも上品なものであった。

選手権で10位になった Downing のプログラムは、相手に報復の気配が見えないときには、徹底して裏切って食い逃げを決め込み、逆に報復の気配が見えるときは、こちらも協調するという戦略(収益最大化戦略と呼ばれる)をとった。この収益最大化戦略は、最初相手の協調と裏切りの確率を1/2と仮定して対戦を始める。このように、収益最大化戦略は、最初は相手を手応えのない戦略と仮定しているのだから、対戦の最初の2回はどうしても相手を裏切ることになる。この相手を見積もるための試し見をすることが多くの相手に収益最大化戦略に対して報復するように仕向けてしまい、結果として収益最大化

²³このリーグ戦において、各試合はちょうど200回の対戦からなり、その回数は事前に参加者に知らされていた。しかしながら、その結果は本節の最初で述べたような自明な結果とはならなかった。これに関する考察については後述する。

²⁴勝ち負けの評価基準を総得点ではなく、たとえば勝ち数などにすると、有効な戦略が変わりうる。

戦略は対戦の出だしにおける得点が悪かった。

上品な戦略のうちで、前述の収益最大化原理に基づく戦略と対戦した場合に、状況を悪化させるか、しないかの対応の仕方の違いは戦略の“心の広さ”であった。上品な戦略の中で下位から2番目となった Friedman によるプログラムは最も“心が狭い”戦略をとるものであった。それは相手が1回でも裏切ると、その後ずっと裏切り続けるというものである。

選手権で12位となった Joss のとった戦略はしっぺ返しの焼き直しであるが、これはある確率で相手を裏切って食い逃げしようとする卑劣な戦略である。これはしっぺ返し同様、相手に裏切られると即座に裏切り返すが、相手が協調した後でも9割協調して、1割裏切る。つまり気まぐれに相手の協調を食い逃げしようとする。このような戦略をとると、その気まぐれさの確率に応じてある回数以降に裏切りの応酬が始まることがある。対戦相手が自分自身であったり、しっぺ返しのように報復する場合は、このような卑劣な戦略は割に合わない。

以上、囚人のジレンマ・ゲームにおける各種戦略の諸性質をみてきたが、この選手権から得られた教訓は、互いに泥試合に陥る事態を極力避けることの重要性である。一回の裏切り行為は、それ自体の直接的効果を考えればうまくいくように思える。しかしその性質の真の損失は、それが果てしなき報復戦を呼び起こしてしまうことにある。この選手権で成功したしっぺ返しが成功した要因は、それが相手から搾取する可能性を放棄していることにある。

この節の最初で述べたように、囚人のジレンマ・ゲームが反復される場合であっても、両プレーヤがその反復回数を知っているのなら、両者がケース解析に基づき戦略を選択する場合は、相変わらず協調関係を引き出すことができない。Axelrod の実験でも事前に参加者に対戦回数が知らされていたのだが、それでは、上述のよう

な協調戦略が参加してきたことをどのようにして説明できるのであろうか。

Axelrod の考察によれば、囚人のジレンマ・ゲームにおいては、(1) ゲームが繰り返し行なわれ、しかもその回数が双方に分からない、かつ、(2) 未来係数²⁵が十分に大きい場合には、協調戦略が発現しうる(協調戦略が現われてくることを理論的に裏付けられる)ことが分っている。将来の利得を見積もるための総和をとるときに、未来係数が大きければ大きいほど、将来の対戦が利得の総和に効いてくるので、このような場合には相互協調に向かう芽が現われてきやすいのである。

上で紹介した Axelrod の実験では、反復回数は参加者にとって既知であった。しかしながら、それが 200 回という十分に長い回数であったために、しっぺ返しのような協調戦略をとるプログラムが複数参加してきたと考えられる。また実験結果が示すように、それらの協調関係は実際に有効であった。反復回数が既知であったことによる効果であるが、[1]によれば、200 回の対戦の終盤になると裏切るというような小細工を行なう戦略がいくつか観察されたと報告されている。Axelrod は上述の実験と同様な実験を 2 回行っている。2 回目の実験では、対戦回数は未来係数を導入することにより確率的に決定されることが事前に知らされていた。このよう

²⁵ある対戦が終わった後に、次の対戦を行なうかどうかの確率である。今回の対戦の利得に対する次回の対戦の利得の重要度の度合を指すと解釈することもできる。同一の相手と繰り返しゲームを行なう場合、今回の自分の戦略が今の自分の利得を決めるだけでなく、次回の相手の戦略に影響する可能性がある。よって、未来は現在に影響を及ぼすと考えられるが、しかしながら、未来は現在ほど重要ではない。それは、今回の利得が未来の期待に勝ったり、また、対戦が今回で終わったりすることがあるからである。よって、次回以降の利得は今回の利得よりも必ず軽くみられる。未来係数はこのことを反映している。仮に各対戦の利得が P で未来係数が w であるとすると、今回を含めた将来の期待利得の総和は $P + Pw + Pw^2 + Pw^3 + \dots = P/(1-w)$ となる。よって w が大きいほど未来の期待利得が大きくなる。

な条件下では対戦回数が確定してないので、試合の終盤で小細工を行なう戦略はなくなったと報告されている。

4.5 研究事例：協調関係の繁栄 – 仮想的な生態学的実験からの考察 –

いま仮に、ある集団が“全面裏切り”戦略をとる個体で占められていたとする。この集団の中に、しっぺ返しのような協調戦略をとる個体が現われた²⁶とする。このような協調戦略は、全面裏切り戦略をとる集団の中で定着して繁栄²⁷できるものであろうか？

もしそうでないとするならば、繁栄できるためにはどのような条件が必要なのか？これについて考察するため、Axelrod は第 2 回囚人のジレンマコンピュータ選手権に参加してきた 63 個の戦略を用いて、ある種の仮想的な生態学的模擬実験を行なった [1]。そこでは、1 回の総当たり戦を終えた後に、各戦略の得点に比例して次回(生態学的には次世代と考えてよい)の総当たり戦における各戦略の参加者の比率を変えるというものである。これを何世代も繰り返して行ない、戦略の勢力図がどのように変化するかをみることを行なった。この実験は、Maynard-Smith が提唱している概念である“進化的に安定な戦略”(Evolutionarily Stable Strategy, 略称 ESS)[16, 17]と深く関係する。ESS とは個体群の大部分のメンバーがそれを採用する(つまり、その戦略が繁栄する)と、別の代替戦略にとってかわられることのない戦略と定義される。

さて、第 2 回囚人のジレンマコンピュータ選手権に参加してきた戦略を用いて 1,000 世代に渡る勢力争いのシミュレーションを行なった結

²⁶現われたの意味は、全面裏切り戦略をとる個体で占められている集団に、協調戦略者が外から入り込んできたと考えてもよいし、全面裏切り戦略をとるある特定の個体に何らかの変化(突然変異など)がおきて、協調戦略に変化したと考えてもよい。

²⁷大部分の個体はその戦略を採用するようになること。

果は、1回の総当たり戦の結果と同様にしっぺ返し最優勢となった。また、1回の総当たり戦における上位 1/3 の心の広い戦略は 1,000 世代目にはほぼすべてを占めるようになった。1回の総当たり戦における下位 1/3 の戦略は 50 世代でほぼ消滅した。また、1回の総当たり戦では、それ自身あまり得点を挙げられない戦略と対戦して功を奏するような戦略（“弱いものいじめ”と呼ばれる）は、世代が進むにつれて自分が成功するのに必要な弱者を滅ぼしてしまい、結果的には自滅の道歩んだ。

Axelrod の実験からは上記の結果が報告されているが、実験の初期段階において対戦に参加する戦略をしっぺ返しを含んだいくつかの種類に限定し、ある種の遺伝的プログラム²⁸に基づいてシミュレーションを行ない、より長期的な世代に渡って各戦略の繁栄をみる、という実験の結果が [13] に紹介されている。この実験の結果として、突然変異がある場合は必ずしもしっぺ返しが最良の戦略として進化的に安定な戦略となるのではなく、長期的にはさまざまな異なる戦略が繁栄/衰退を繰り返し、しっぺ返しよりもノイズに強い優れた戦略が発生してくるという進化のダイナミクスが観察されたことが報告されている。そのダイナミクスの内容としては、有限平衡状態、戦略の急激な絶滅、相互依存性の共進化、進化的に安定な戦略などが挙げられている。また、反復囚人のジレンマ・ゲームに関する別実験の結果が [22] に報告されており興味深い。

さて、先に挙げた問い（協調戦略が繁栄するための条件）に対する答であるが、それは次のように述べるができる。仮にある集団がすべて全面裏切り戦略で占められているとする。これは進化的に安定となり、他の戦略をとる単

一の個体がそこに現われたとしても侵入²⁹できる可能性はない。何故なら、たとえそこに協調戦略が一つ現われたとしても、それは必ず全面裏切り戦略より低い利得しか得られないからである。この協調戦略が全面裏切り戦略の集団の中に侵入し繁栄できるためには、その集団に複数（少なくとも二つ以上）の協調戦略者が現われ、それらどうしで対戦する機会が多くなければならない。この“それらどうしで対戦する機会が多くなる”ということは、協調戦略者が他の協調戦略者と対戦する確率 p が次に述べる条件を満たすことを意味している。その条件とは、先に述べた協調しあうときの利得を R 、裏切りあったときの利得を P 、裏切られたときの利得 S 、そして未来係数を w としたときに、 $pR/(1-w) + (1-p)(S + wP/(1-w))$ が $P/(1-w)$ よりも大きくなることである。つまり、全員が裏切る集団における平均利得よりも、大多数が裏切りあう集団において確率 p で協調戦略者どうしが対戦する機会があるときの協調戦略者の平均利得の方が大きい場合である。

このような条件を満たせば、各戦略の総得点に基づき次世代の各戦略の個体数を決定するという環境の下では、しっぺ返しのような協調戦略は最初は少数であっても世代を重ねるごとに次第にその個体数を増し、最終的には全面裏切りにとって代わり進化的に安定な戦略となりうる。

4.6 まとめ

囚人のジレンマは一見単純なゲームであるが、最適な戦略というものは、かなり強い制約を双方のプレーヤにおかない限り存在しない。有効な戦略にしても、ゲームが行なわれる状況（1回か繰り返し行なうか）や相手の戦略など、いろいろ

²⁸ ノイズによる戦略の突然変異、過去の対戦に関する記憶の長さの確率的变化、各戦略の総得点に基づく次世代の戦略の選択、などの機構が導入されている。

²⁹ 集団の全個体がある特定の戦略をとる中に、別の戦略をとる一個体が現われて、それが集団内の他の個体よりも高い得点を上げられれば、この入り込んだ別の戦略は集団に侵入できるといわれる。

るな要因を考慮に入れなければならない。それらの要因に対してさまざまな関連研究が、ゲーム理論における均衡戦略にはじまり、協調における合理性の役割、繰り返し囚人のジレンマ・ゲームによる仮想生態学的な優位性の実験、そして協調の発現と進化、などの広範な観点からの考察を示した。

2人のエージェントが相互作用を持つ状態で、そこで採られる行動の各効用がこの章の最初に与えた問題の定義にある制約を満たすものであれば、それは囚人のジレンマと同じ状況になる。このような明確な効用が得られない状況も多くあるが、それでも例えば、(1) プレーヤどうしの利得は互いに価値を比較することができない、そして、(2) プレーヤの利得は相対的なスケールでしか測ることができない、などの場合においてもそれは囚人のジレンマと同じ状況になることが知られている。

このように条件をゆるめれば、囚人のジレンマに対応する状況はかなり多くなるであろう。エージェントが囚人のジレンマの状況に直面した場合(特にその状況が繰り返される場合)は、ここで述べられたようなさまざまな結果、知見、そして教訓などを考慮に入れて協調という行動選択を行えば、最悪の状況を避けて自己の効用や全体の効用をより高くすることが可能となる。

また、囚人のジレンマ・ゲームの前提条件がそうであるように、Geneserethらの合理的エージェントの研究ではエージェントどうしは戦略を決定する以前に通信を行わないことを前提としている。そのために、(協調, 協調)という結合動作に達するには共通振舞いという強い制約を2人のエージェントに要求することになる。この制約は、異質なエージェントが相互作用を持つような環境では充足困難なものであろう。このような問題に対処する方法を考察するために、本稿では紹介しなかったが、Rosenscheinらの研究[15]では、この通信を行わないという仮

定を外している。この場合、利害が完全には対立しないような状況で、どのような内容の情報をどの程度通信しあえば協調動作に達することができるかが興味ある研究対象となる。このようなエージェント間における情報のやりとりとそれに基づく意思決定は、一般に**取り引き**とか**交渉**などと呼ばれている。分散人工知能における交渉スキーマの研究は、Rosenscheinらの研究以降もさまざまな問題領域で盛んに行なわれており、協調に関するこのようなアプローチの研究も注目されている³⁰。

5 おわりに

分散人工知能において、動的な環境におけるエージェントの行動規範や、協調/非協調状況における相互作用のモデル化などは重要な研究課題である。それらに関する研究を Tileworld, 追跡問題, 囚人のジレンマという三つの小問題を通して考察した。

Tileworld は資源制限を受けたエージェントのアーキテクチャを評価するための動的な実験環境である。ここでは、環境の動的特性に対するエージェントの即応と熟考のトレードオフ、そしてコミットメントの役割などに関して考察した。

追跡問題は、単一の大域目標を持つ複数エージェントの集団において、各エージェントが個々の局所的なプランと大域目標達成のためのプランをどのようなメカニズム/戦略により整合させるかという問題である。この問題では、大域目標を達成するために複数のエージェントが組織を作る場合、(1) どのような組織形態が可能で、(2) 各組織形態はどのような性質を持っていて、また(3) 各組織の問題解決の効率はどうか、などの点から、協調的組織スキーマに関して考察した。

囚人のジレンマは非零和非協力ゲームである。

³⁰交渉に関しては、桑原・石田の解説[24]がある。

それを通して利害が完全には対立しない状況でのさまざまな戦略の相互作用や戦略の進化, そして協調的行動規範を与える合理性について考察した。

本稿では取り扱わなかったが, エージェントの部分性³¹や, 通信におけるエージェント間の不整合の問題なども分散人工知能研究においては基本的な研究課題である (例えば, [3]などを参照されたい)。これらの問題に関しても, 多くの研究を比較し追試を可能とするための標準的な小問題が考案されることが重要であろう。また, それらの問題と同種の問題に対する他の研究分野における考察などを十分に踏まえ, 分散人工知能の基本理論が確立されていくことを望む。最後に, 本稿の内容が, (開放型)分散システムにおける計算実体のパフォーマンスの考察や, 計算実体間の協調モデルを考える上で何らかのお役に立てば幸いである。

謝 辞

このサーベイは, 1991年10月の「第1回マルチエージェントと協調計算研究会」における会場からの要望に応える形で始められたものである。協同作業をサポート頂いた, ソニーコンピュータサイエンス研究所の所真理雄副所長および各研究員, NTTコミュニケーション科学研究所の西川清史所長, 中野良平主幹研究員, 並びに討論頂いた赤埴淳一氏に感謝します。また, 本稿の査読者の方々とソニーコンピュータサイエンス研究所の長尾確氏には, たいへん有益な助言を頂きました。ここに感謝の意を表します。

³¹局所性や資源制限性に起因するエージェントの信念, 知識, そして技能の部分性を指す。

参考文献

- [1] Axelrod, R.: *The Evolution of Cooperation*, Basic Books Inc. (or Penguin Books), 1984. (邦訳: 松田裕之訳, 「つきあい方の科学」, HBJ 出版局, 1987年).
- [2] Benda, M., Jagannathan, V. and Dodiawalla, R.: On Optimal Cooperation of Knowledge Sources, Technical Report BCS-G2010-28, Boeing AI Center, 1985. (Cited in Gasser, L., Rouquette, N., Hill, R.W. and Lieb, J.: Representing and Using Organizational Knowledge in Distributed AI Systems, In Les Gasser and Michael N. Huhns, editors, *Distributed Artificial Intelligence II*, pp. 55-78, 1989).
- [3] Bond, A. and Gasser, L.: An Analysis of Problems and Research in DAI, In Alan Bond and Les Gasser, editors, *Readings in Distributed Artificial Intelligence*, pp. 3-35. Morgan Kaufmann Publishers, Inc., 1988.
- [4] Bratman, M.E., Israel, D.J. and Pollack, M.E.: Plans and Resource Bounded Practical Reasoning, *Computational Intelligence*, Vol. 4, No. 4 (1988), pp. 349-355.
- [5] Cohen, P.R. and Levesque, H.J.: Intention is Choice with Commitment, *Artificial Intelligence*, Vol. 42, No. 2-3 (1990), pp. 213-261.
- [6] Davis, L.: Prisoners, Paradox and Rationality, *American Philosophical Quarterly*, Vol. 14 (1977).
- [7] Durfee, E.H. and Montgomery, T.A.: MICE: A Flexible Testbed for Intelligent

- Coordination Experiments. In *Proceedings of the Ninth Workshop on Distributed Artificial Intelligence*, pp. 25–40, 1989.
- [8] Gasser, L., Rouquette, N, Hill, R.W. and Lieb, J.: Representing and Using Organizational Knowledge in Distributed AI Systems, In Les Gasser and Michael N. Huhns, editors, *Distributed Artificial Intelligence, Volume II*, pp. 55–78. Morgan Kaufmann Publishers, Inc., 1989.
- [9] Genesereth, M.R., Ginsberg, M.L. and Rosenschein, J.S.: Cooperation Without Communication, In *Proceedings of the Fifth National Conference on Artificial Intelligence (AAAI-86)*, pp. 51–57, 1986.
- [10] Hofstadter, D.R.: *METAMAGICAL THEMAS*, Basic Books Inc., 1985. (邦訳: 竹内郁雄他訳,「メタマジック・ゲーム」, 白揚社, 1990年).
- [11] Kinny, D.N. and Georgeff, M.P.: Commitment and Effectiveness of Situated Agents, In *Proceedings of the Twelfth International Joint Conference on Artificial Intelligence (IJCAI-91)*, pp. 82–88, 1991.
- [12] Levy, R. and Rosenschein, J.S.: A Game Theoretic Approach to Distributed Artificial Intelligence, In Eric Werner and Yves Demazeau, editors, *Decentralized A.I. 3*. Elsevier/North Holland, 1992.
- [13] Lindgren, K.: Evolutionary Phenomena in Simple Dynamics, In Christopher G. Langton, Charles Taylor, J. Doynne Farmer, and Steen Rasmussen, editors, *Artificial Life II*, Addison Wesley, 1992, pp. 295–312.
- [14] Pollack, M.E. and Ringuette, M.: Introducing the Tileworld: Experimentally Evaluating Agent Architectures, In *Proceedings of the Eighth National Conference on Artificial Intelligence (AAAI-90)*, pp. 183–189, 1990.
- [15] Rosenschein, J.S. and Genesereth, M.R.: Deals Among Rational Agents, In *Proceedings of the Ninth International Joint Conference on Artificial Intelligence (IJCAI-85)*, pp. 91–99, 1985.
- [16] Maynard Smith, J.: The Theory of Games and the Evolution of Animal Conflict, *Journal of Theoretical Biology*, Vol. 47 (1974), pp. 209–221.
- [17] Maynard Smith, J.: The Evolution of Behavior, *Scientific American*, Vol. 239 (1978), pp. 176–192.
- [18] Stephens, L.M. and Merx, M.: Agent Organization as an Effector of DAI System Performance, In *Proceedings of the Ninth Workshop on Distributed Artificial Intelligence*, pp. 263–292, 1989.
- [19] 大沢英一: 問題空間の変化に適応する協調的組織スキーマ – 追跡ゲームにおける考察 –, 第2回マルチエージェントと協調計算ワークショップ (MACC'92), 1992. (「マルチエージェントと協調計算 II」(近代科学社)に掲載予定).
- [20] 石田亨: 知識表現と動的世界 – 最近のプランニング研究から –, 人工知能学会誌, Vol. 5, No. 2 (1990), pp. 146–153.
- [21] 石田亨: 動的環境でのプランニング – アーキテクチャと技法 –, 計測と制御, Vol. 31, No. 7 (1992), pp. 781–786.

- [22] 竹内郁雄, 山崎憲一: ナノピコ教室解答 (ソフトウェアサンの取引き). *bit*, Vol. 22, No. 2 (1990), pp. 109-113.
- [23] 鈴木光男: ゲーム理論入門, 共立出版, 1981.
- [24] 桑原和宏, 石田亨: 分散人工知能 (2):交渉と均衡化, 人工知能学会誌, Vol. 8, No. 1 (1993), pp. 17-25.