エージェント間の協調動作に基づく分散的探索手法

北海道大学工学研究科システム情報専攻複雑系工学講座
○ 川村 秀憲　山本 雅人　鈴木 恵二　大内 東

The Distributed Search Algorithm Based on Cooperative Behaviors Between Multi-Agents

Research Group of Complex Systems Eng., Graduate School of Eng., Hokkaido University, JAPAN
○ Hidenori KAWAMURA　Masahito YAMAMOTO　Keiji SUZUKI
and Azuma OHUCHI

Abstract: In this paper, we propose the distributed search algorithm based on cooperative behaviors multi-agents for Vehicle Routing Problems. This algorithm is designed using multi-agent system, in which multi-agents autonomously and independently search the partial solution of problem to decrease their load of tasks usgin pheromone communication. Pheromone communication imitated from living creatures is used as the communication method between multi-agents to emerge the cooperative behavior. Especially, in order to form effective pheromone communication we define the action of agents based on observing other agents. Throughout some computer experiments, we confirm the effectiveness of proposed algorithm from viewpoint of cooperative behavior.

1　はじめに

近年複雑かつ大規模な問題に対して，マルチエージェントシステムによる分散協調的問題解決がその有効性の期待から注目を集めている．しかしマルチエージェントシステムにおいて，複数のエージェント群が有益的に群としての目的を達成する為には，常に他者の状況を鑑み，適応的に協調関係を築きながら組織立って振る舞うことの必要があるが，その様な挙動を示すシステムの設計自体非常に困難な問題であると考えられる．このシステムの実現へ向けて，全体または他者の状況理解と自己の主張を目的としたマルチエージェント間のコミュニケーションが重要な役割を担うであろうことは容易に想像がつく [1]．

そこで本稿では複雑な情報処理や推論の能力，高度なレベルのコミュニケーションなどを行っているとは考え難い昆虫レベルの集団行動に注目する．そこでは例えば社会性昆虫と呼ばれる蟻や蜂などのフェロモンやダンスによるコミュニケーションなどに見られるように，非常に単純な枠組のコミュニケーション手段を用いることにより組織的活動を発現し，群としての目的を達成している．特にフェロモンは群における各個体の環境観察や単純な状況判断から反射的に生成され，それら複数個体のフェロモンの相互作用によって，組織行動を発現する為の有効なコミュニケーション手段として見なすことが出来る [2][3]．それは大規模かつ複雑な問題，特に組合せ最適化問題に対しての分散協調的な最適化手法実現への有力なアプローチになるであろうと期待できる [4][5]．ここで従来の研究の中において，巡回セールスマン問題の解法としてフェロモン情報を利用したアントアルゴリズムが提案されており，その有効性が示されている [6]．アントアルゴリズムにおいて，各エージェントはそれぞれ与えられた問題に対して独自に探索を行い，その時のエージェント間の情報交換の手段としてフェロモン情報を用いている．

しかし本稿では一般に巡回セールスマン問題よりも困難とされる配送経路問題と対象とし [7]，各エージェントは与えられた問題そのものの探索を行うのではなく，それぞれが自律的に与えられた問題から部分問題を生成し，分散的に部分問題を探索する時のコミュニケーションの手段としてフェロモン情報を採用する．そのことによって各エージェントの探索負荷が減少することが期待できる．この際重要となるのはいかに有効なフェロモン情報を生成しそれを利用するかということである．提案手法では実際の生物のように各エージェントが自分と他のエージェントの状況を観察し，簡単なルールに従って反射的にフェロモンを生成する．各エージェントによって生成

されたフェロモン情報から，協調的な振る舞いを発現し，分散適応的に与えられた問題の探索を行うことが期待できる．本稿では以上のマルチエージェントシステムによる分散的な最適化手法を構築する．

2　配送経路問題

配送経路問題は複数の車両がデポ (配送センター) から出発し，荷物の配達を要求する複数の配達先ノードを回って再びデポへ戻ってくる時の総コストの最小化を目的とする n 人巡回セールスマン問題の拡張問題と考えることができ，現実的な制約を加えた種々の応用がある．この問題は各車両に割り当てる配送先ノードとそれぞれの巡回路を同時に決定しなければならず，一般に巡回セールスマン問題より困難とされている．本稿で扱う配送経路問題は与えられた配送先ノードを与えられた台数の車両で均等に分割して複数の巡回路を生成し，そのときの巡回路の総距離の和の最小化を目的とすることとする．

3　分散最適化アルゴリズム

配送経路問題に対する本アルゴリズムは，マルチエージェントシステムによる分散的アルゴリズムで実現される．分散的な探索過程において各エージェントは自律的に与えられた問題から自分の担当するべき部分問題を生成し，その問題に対してのみ探索を行う．このことによって各エージェントは全問題空間を考慮する必要が無くなり，探索の負荷を減らすことが出来ると期待できる．しかし目的は与えられた問題に対する解の目的関数の最小化であるが，全問題空間を意識せず分散的に探索を行う為に各エージェントは部分解の質に関する情報しか得られない．その様な状況において有効な探索を行うには部分問題の決定と探索における自分の部分解の質の追求もさることながら，いかに他のエージェントと協調的な関係を築き上げ全体としての目的を達成するかということが必要である．

そこで本稿では部分問題の生成に対して各エージェント間のノードの奪い合いを許しながら，フェロモン情報を用いることによりエージェント間の協調的な動作と全体としての目的の達成を実現する方法を提案する．このアルゴリズムでは各エージェントはそれぞれ与えられた問題に対する部分解を有し，各自が自分の判断基準に従ってこの部分解の更新を行う．この時，他のエージェントとの解の調整，また与えられた問題に対する各自の部分解の評価のためにフェロモン情報が用いられる．さらに，各エージェントは自分と他者の探索結果を観察しながらそれぞれフェロモン情報を生成し，環境中にフェロモンをおいていく．その中でフェロモン情報が自己組織化され，マルチエージェントによる協調的な探索が実現できる．

3.1　フェロモン

生物の分野においてフェロモンは同種の個体間で交わされる化学信号によるコミュニケーションと定義される．フェロモンは環境の観察や，単純な状況判断からある個体が反射的に環境中に放出する．そこで群行動においては各個体がそれぞれ独自にフェロモンを生成するが，環境中ではそれら複数の個体によって生成されたフェロモンが重ね合わされ，それらの相互作用によって有益的な群行動を誘発する引き金となる．また，いったん放出されたフェロモンは生成した生物から独立した物質となり，その化学的性質から徐々に環境中から蒸発していく．それによって過去の冗長な情報を忘れたり，新しい状況に適応的に対処することが出来る．フェロモンを受け取った個体は，例えば信号の方へ近寄ったり敵を攻撃したりといった特定の行動や，個体内部が生理的に変化するいったことを反射的に誘発する．このようにフェロモンの生成や反応自体は非常に単純な枠組みで扱えるが，例えば社会性昆虫と呼ばれる蟻や蜂などはフェロモンの働きによってその社会的行動が厳密に統制され，はっきりとした分業制のもとで複雑な社会行動を営む．

そこで提案アルゴリズムにおいても同様の枠組みでフェロモンを扱う．配送経路問題の探索において各エージェントは自分と他のエージェントの状態を観察しながら，簡単な状況判断に従って反射的に環境中にフェロモンを置いていく．また，各エージェントはその探索過程において環境中のフェロモンを観察することによって，他のエージェントとの関係を知覚し，それを探索に反映することによって全体としての目的を達成しようとする．ここでは環境中に時間 t における 2 種類のフェロモン$\tau^1_{ki}(t)$ と $\tau^2_{ki}(t)$ を設定する．$\tau^1_{ki}(t)$ はあるエージェント kが与えられた問題における配送先ノード i を用いることの重要性を表し，$\tau^2_{ki}(t)$ はそのノードを用いないことの重要性を表すことを期待する．即ち，これらを用いることによって郡全体の探索過程において，自分の部分解がどの程度有効であるかを評価することが出来る．これらのフェロモンが各エージェントの相互作用によって生成され，適応的な問題分割と探索が出来ると期待できる．

3.2 エージェントの探索アルゴリズム

本アルゴリズムは停止回数 t_{max}まで各エージェントの動作を繰り返す．ここでエージェント kが時間 t において有する部分解を $S_k(t)$ とする．配送経路問題において $S_k(t)$ はデポから出発して自分に割り当てられた各配送先をまわりデポまで戻ってくる巡回路を表すものとする．時間 t において各エージェントは自分の有する部分解 $S_k(t)$ に対する評価関数 (ここでは部分解が表す巡回路の総距離) と，環境中から受け取るフェロモン情報を用いて $S_k(t)$ を更新することが出来る．更新アルゴリズムは以下に従う．

Step1: $S_k^*(t) := S_k(t-1)$

Step2: $S_k^*(t)$ の中から任意のノードを選択し，$S_k^*(t)$ に含まれない任意のノードと入れ替える

Step3: $S_k^*(t)$ に対して反復回数 $t_{localmax}$だけローカルサーチを繰り返す

Step4: 受理確率 $P_k(t)$ を満たした時 $S_k(t) := S_k^*(t)$ とし，満たさない時は $S_k(t) := S_k(t-1)$ とする

ローカルサーチは新しい部分解 $S_k^*(t)$ を素早く改良し，それを構成する配送先ノードから出来うる部分解の評価を安定させる為に行う．ローカルサーチを行わない場合同じノードから出来る部分解のばらつきが大きくなり，有効なフェロモン情報を生成する事が出来なくなる．なお本稿ではローカルサーチを行う際の近傍解の生成法として巡回セールスマン問題で有効であるとされている 2-opt 法を採用した．また，$S_k^*(t)$ の受理確率を以下の式で定義する．

$$P_k(t) = p_k^1(t) \cdot p_k^2(t) \quad (1)$$

$$p_k^1(t) = \begin{cases} 1 & \text{if } \Pi_k^1(t) \geq 0 \\ exp(\Pi_k^1(t)/T_\tau(t)) & \text{otherwise} \end{cases} \quad (2)$$

$$p_k^2(t) = \begin{cases} 1 & \text{if } \Pi_k^2(t) \geq 0 \\ exp(\Pi_k^2(t)/T_E(t)) & \text{otherwise} \end{cases} \quad (3)$$

$$\Pi_k^1(t) = T(S_k^*(t)) - T(S_k(t-1)) \quad (4)$$

$$\Pi_k^2(t) = -\{E(S_k^*(t)) - E(S_k(t-1))\} \quad (5)$$

$$T(S_k(t)) = \sum_{i \in S_k(t)} \tau_{ki}^1(t) + \sum_{i \notin S_k(t)} \tau_{ki}^2(t) \quad (6)$$

ここでフェロモン情報$\tau_{ki}^1(t)$ と$\tau_{ki}^2(t)$ はそれぞれノード i がエージェント kの部分解に属する時の重要度と，属さない時の重要度を表すことが期待される．よって$T(S_k(t))$ はエージェントkに対する部分解$S_k(t)$ のフェロモン量の総和即ちその部分解の重要度を表している．また，$E(S_k(t))$ は部分解 $S_k(t)$ の評価関数 (巡回路の総距離) とする．$p_k^1(t)$ はフェロモン情報に基づき新たに生成した $S_k^*(t)$ が $S_k(t-1)$ よりも重要であるとされた時高確率をとる．同様に $p_k^2(t)$ は $S_k^*(t)$ が $S_k(t-1)$ よりも良い評価を有するとき高確率をとる．つまり各エージェントはより重要で評価の良い解を受理しやすい．この確率において $p_k^1(t)$ と $p_k^2(t)$ をコントロールする温度パラメーターが $T_\rho(t)$ と $T_E(t)$ であり，時間 t において以下の式によって更新される．

$$T_\tau(t) = \rho_\tau \cdot T_\tau(t-1) \quad (7)$$

$$T_E(t) = \rho_E \cdot T_E(t-1) \quad (8)$$

ここで$\rho_\tau (0 \leq \rho_\tau \leq 1)$ と$\rho_E (0 \leq \rho_E \leq 1)$ は温度冷却率であり，探索が進行するにつれて $p_k^1(t)$ と $p_k^2(t)$ の影響を徐々に強めていくことで有効な解へ収束させていく．しかしここで重要なのはこのバランスであり，片方の影響が強すぎたり弱すぎたりすると有効な探索が行えない．従ってこのパラメーターは慎重に設定する必要がある．

3.3 フェロモンの更新

各エージェントはそれぞれの部分解の更新が終了したあと，自分や他のエージェントの部分解を観察しながら環境中にフェロモンを放出する．この時，環境中における各フェロモン情報は以下の式に従って更新される．

$$\tau_{ki}^1(t+1) = \rho \cdot \tau_{ki}^1(t) + \sum_h \Delta\tau_{kih}^1(t) \quad (9)$$

$$\tau_{ki}^2(t+1) = \rho \cdot \tau_{ki}^2(t) + \sum_h \Delta\tau_{kih}^2(t) \quad (10)$$

ここで$\rho (0 \leq \rho \leq 1)$ はフェロモン情報の蒸発に関するパラメーターで $(1-\rho)$ がその蒸発率を表す．フェロモンの蒸発は過去の冗長な情報を忘れ新たに得られた有効な情報を効果的に利用する為に導入される．また$\Delta\tau_{kih}^1(t)$ と $\Delta\tau_{kih}^2(t)$ はそれぞれエージェント h が環境中に放出するフェロモンの変化量を表す．このときエージェント h は自分に関するフェロモンを生成する場合即ち $h = k$の時と，他のエージェントに関するフェロモンを生成する場合即ち $h \neq k$の時に関してそれぞれ以下で説明されるようにフェロモンを環境中に放出する．

まず，$h = k$の時エージェント h はノード i に関して以下の状態のいずれにあるかを判断し，フェロモンを生成する．

状態 1.1: ノード i がエージェント h の部分解に属するか，エージェント h の部分解には属さないが他の任意のエージェントの部分解に属する時

状態 1.2: ノード i がいずれのエージェントの部分解にも属さない時

状態 1.1 の時は，部分解に属するときと属さない時についてそれぞれ$\tau^1_{ki}(t)$ と$\tau^2_{ki}(t)$ にフェロモンを置く．状態 1.2 の時は，そのノードがいずれのエージェントにも使用されていないので，そのノードの使用を促進するために$\tau^1_{ki}(t)$ にフェロモンを置く．

また，$h \neq k$の時はエージェント h はエージェント kとノード i に関して以下の状態のいずれにあるかを判断し，フェロモンを生成する．

状態 2.1: ノード i がエージェント kの部分解に属し，エージェント h の部分解には属さない時

状態 2.2: ノード i がエージェント kの部分解に属さず，他の任意のエージェントの部分解に属す時

状態 2.3: ノード i がエージェント kの部分解に属し，エージェント h の部分解にも属す時

状態 2.4: ノード i がエージェント kの部分解に属さず，他のいずれのエージェントの部分解にも属さない時

状態 2.1，状態 2.2，状態 2.3 にある時はエージェント h はエージェント kの部分解の状態を観察し，ノード i がエージェント kの部分解に属するときと属さない時についてそれぞれ$\tau^1_{ki}(t)$ と$\tau^2_{ki}(t)$ にフェロモンを置く．また，状態 2.4 のときは状態 1.2 の時と同様ノード i がいずれのエージェントにも使用されていないので，そのノードの使用を促進するために$\tau^1_{ki}(t)$ にフェロモンを置く．

エージェント h が環境中に置くフェロモン量は以下の式で計算される．

$$\delta_h(t) = Q/E(S_h(t)) \tag{11}$$

ここで Q は定数である．つまり良い評価の部分解を生成した時そのエージェントのフェロモン量は大きくなり，悪い解の時はフェロモン量が少なくなる．これによってよい評価を有する部分解が促進される．以上をまとめるとエージェント h によるフェロモンの変化量は以下の式で定義できる．

$$\Delta\tau^1_{kih}(t) =$$

$$\begin{cases} \delta_h(t) & \text{if } (k = h \text{ and } i \in S_k(t)) \\ & \text{or } (k \neq h \text{ and } i \in S_k(t) \text{ and } i \notin S_h(t)) \\ & \text{or } (k \neq h \text{ and } i \notin \sum_l S_l(t)) \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

$$\Delta\tau^2_{kih}(t) =$$

$$\begin{cases} \delta_h(t) & \text{if } (i \notin S_k(t) \text{ and } i \in \sum_l S_l(t)) \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

探索の過程においてフェロモン情報が自己組織化され，与えられた問題の完全な分割が達成された時，各エージェントのフェロモン情報には全てのエージェントの部分解の質に関する情報，つまり与えられた問題に対する解の質の情報が反映することになる．

4 計算機実験

本提案アルゴリズムにおいてエージェントの分散協調的な振る舞いを確認する為に 2 種類の計算機実験を行った．ここでは各エージェントが強く自分の部分解の評価を追求しても最適解が得られる問題と，最適解を得る為にはいくつかのエージェントが譲歩し他のエージェントと協調しなければ最適解が得られない問題を設定した．

Q	$T_\tau(0)$	$T_E(0)$	ρ	ρ_τ	ρ_E
800.0	20.0	10.0	0.7	0.9995	0.9995

t_{max}	$t_{localmax}$
2000	100

表 1: 実験 1，2 で用いたパラメータの値

4.1 実験 1

30 の配送先ノードを円状に，その中央にデポを配置した問題に対して実験を行った．この時車両の台数つまりエージェント数は 6 とし，各パラメータは予備実験より表 1の様に決定した．

ここで図 1は探索過程における各エージェントの部分解の評価値の平均値と，全ノードに対する重複したノードの割合を示している．このグラフにおいて平均の評価値は単調に減少しある点から収束しているが，重複の割合はそれ以降にかけても減少し最終的に 0 になっている．各エージェントが探索の早い段階において評価の良い部分解を発見し，それ以降には互いの部分解を調整する為にそれぞれがフェロモン情報に基づいて重複ノードを減らし，協調的に探索を行っている結果であると考えられる．図 2はこの実験において最終的に得られた$\Delta^1_{ki}(t)$ の値であり，ノードの重複の無い有効な問題分割が達成されていることがわかる．図 3はこの実験で得られたこの問題に対する最適解である．

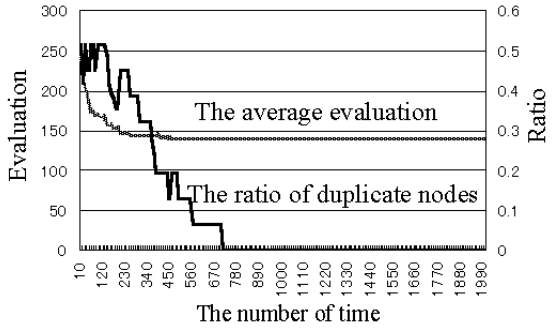

図 1: 実験 1 における各部分解の総評価値と重複ノードの割合の変化

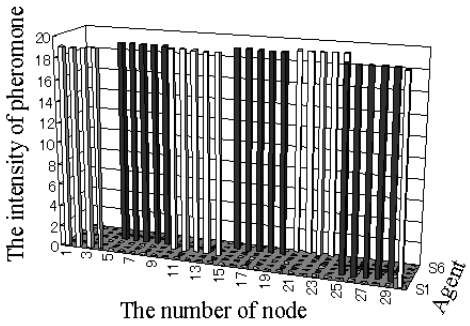

図 2: 実験 1 で得られた$\Delta_{ki}^{1}(t)$ の値

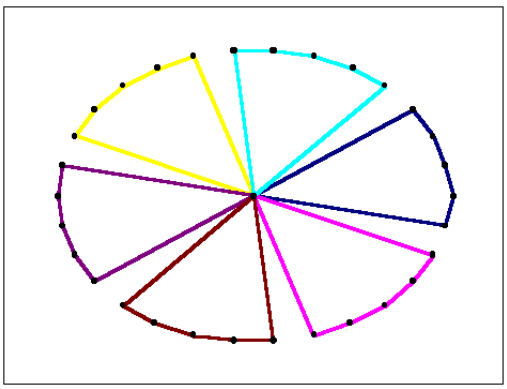

図 3: 実験 1 で得られた最適解

4.2 実験 2

30 の配送先ノードを歯車状に，その中央にデポを配置した問題に対して実験を行った．エージェント数は 6 とし，パラメータは実験 1 と同様表 1の値を用いた．この問題は各エージェントが自分の評価値を優先すると近いノードを用いた部分解に集中してしまうので，有益な解を求める為には各エージェントの協調が必要である．図 4において実験 1 とは違い平均の評価値がなかなか収束しないが，これは分割を変えると生成できる部分解の評価値も変わってしまう為である．この問題の探索においても徐々に重複ノードを減らし協調的な探索を行っているのが確認できる．また図 5は実験 1 と同様問題分割が達成されているが，ここで注目すべき点は各エージェントの部分解に属するノードについて，評価値の悪いエージェントのフェロモン情報と，評価値の良いエージェントのフェロモン情報の値がほぼ同等であることである．これは他のエージェントに譲歩して評価値の悪い部分解を有するエージェントも，良い部分解を有するエージェントと協調状態を結ぶことでフェロモン情報の上で同等の立場にあるということがわかる．これが協調動作を引き起こす理由であると考えられる．図 6がこの実験における各エージェントの部分解の変化である．この図において，有効なフェロモン情報が形成される以前の探索の初期の段階では，各エージェントがそれぞれ近いノードを奪い合っている．しかし探索が進むにつれて有効なフェロモン情報が生成され，徐々に近いノードを譲歩するエージェントが現れ，最終的に最適解に収束しているのが確認できる．

5 終わりに

本稿ではエージェントがフェロモン情報と部分解の評価値に基づいて部分解を更新し，各エージェントが他のエージェントの状態を観察することにより反射的にフェロモンを生成する方法を示した．また，それに基づき自律的に部分問題を生成し探索する中で，他者との協調関係を築き上げ協調分散的に配送経路問題の探索が行えることを示した．本稿で展開した協調のレベルはまだ極めて基礎的なものであり，多様な協調動作の実現を目指して更に様々な角度から検証する必要があると思われる．

参考文献

[1] 鈴木 恵二，吉村 潤，嘉数 侑昇: 疲労度パラメータを導入した行動選択ネットワークによるエージェントの創発的組織化に関する考察，人工知能学会誌，Vol.12，No.1，pp152-159，1997.

[2] 湯嶋 健著: 昆虫のフェロモン, 東京大学出版部, 1976.

[3] William C.Agosta 著, 木村 武二訳: フェロモンの謎 ~生物のコミュニケーション~, 東京化学同人, 1995.

[4] Luca M. Gambardella, Marco Dorigo: Ant-Q: A Reinforcement Learning approach to the traveling

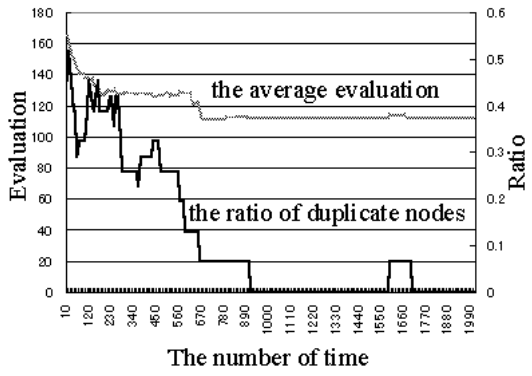

図 4: 実験 2 における各部分解の総評価値と重複ノードの割合の変化

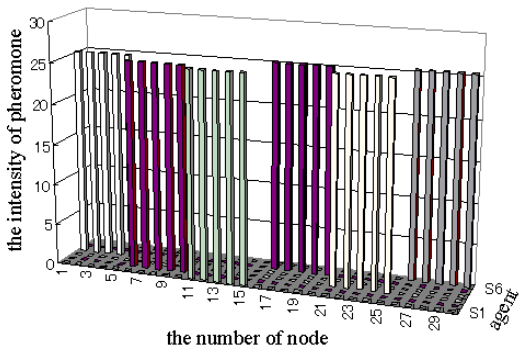

図 5: 実験 2 で得られた$\Delta_{ki}^{1}(t)$ の値

Salesman Problem, Proceedings of ML-95, pp.252-260, 1995.

[5] Devika Subramanian, Peter Druschel and Johnny Chen: Ants and Reinforcement Learning: A Case Study in Routing in Dynamic Networks, the proceedings of IJCAI-97, pp.832-838, 1997.

[6] Marco Dorigo, Vittorio Maniezzo, Alberto Colorni: The Ant Sysytem: Optimization by a colony of cooperating agents, IEEE Transactions on Systems, Man, and Cybernetics-Part B, Vol. 26, No. 1, pp.1-13, 1996.

[7] Ibrahim Hassan Osman: Metastrategy Simulated Annealing and Tabu Search Algorithms for the Vehicle Routing Problem, Annals of Operations Research 41, pp421-451, 1993.

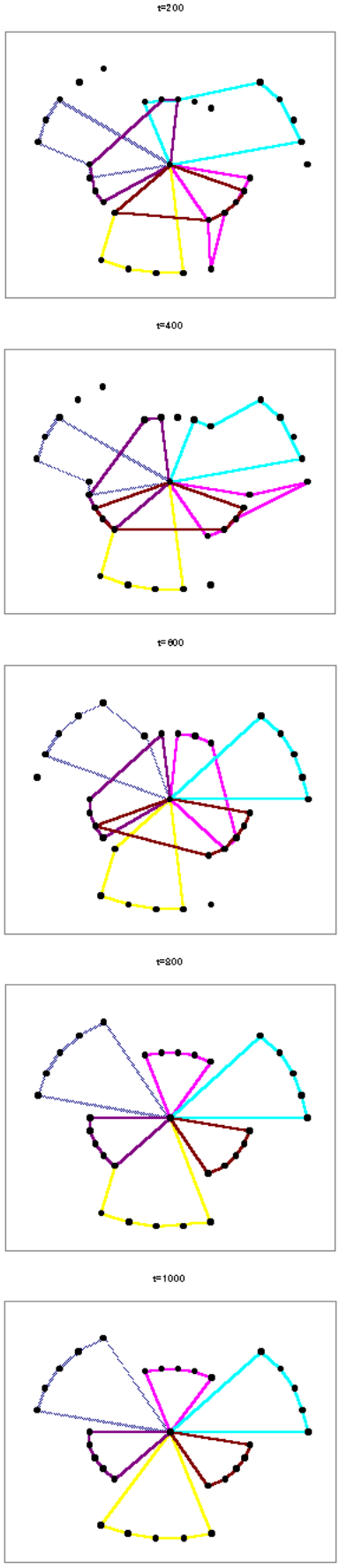

図 6: 実験 2 における各エージェントの部分解の変遷