マルチエージェント実時間探索における組織化とその評価

北村 泰彦　　寺西 憲一　　辰巳 昭治

年　月　日

まえがき

試行錯誤を伴うような非決定的問題解決のための基本的な解決手法として，これまでに様々な探索手法が研究されてきた．深さ優先探索や幅優先探索等の力ずく探索手法は計算量の組合せ的爆発を招き，それに対処するために発見的知識を用いて枝刈りを行う や反復深化 等のヒューリスティック探索手法が提案されるようになった．しかしながら一般的には，その計算量は問題のサイズ（状態数）に対して指数的に増加する． や は探索と移動が分離されたオフライン探索と見なすことができるが， は一定時間内の探索結果を基に移動を行なわねばならないというゲーム探索の概念を導入することにより，先読み探索と移動を交互に行うオンライン探索手法として実時間 を提案した． では最適解を保証することはできないが，先読み探索の範囲は有限であることから計算量は改善され，解の長さに対して線形となる．

は準最適解を求める探索アルゴリズムと見なせるが，その解の質を向上させるために には二つの手法が提案されている．一つは先読み探索の範囲を広げることであるが，その計算量は先読み探

人工知能学会誌　　採録決定

索の深さに対して指数的となる．もう一つはアルゴリズムに学習機能を付加した であるが，解の質を向上させるためには同一の問題を何度も繰り返し探索しなければならない．

さて， は探索に関わるエージェントの数を増加させることにより解の質を向上させるマルチエージェント実時間探索 を提案した ． では同一の問題に対して複数のエージェントはそれぞれ自律的，並行的に を実行する．エージェントが探索過程で分岐に遭遇し，そのそれぞれの経路が同等に望ましいときには，エージェントはランダムに経路を選択する．それゆえ，エージェント数を増加させれば，異なる経路がより多く探索されることになり，確率的により良い解経路が得られることになる． では，計算量はエージェント数に対して線形であり， は パズル等の問題に対しては，先読み探索の範囲を広げた（熟考する）少数のエージェントよりも，先読み範囲を最小限にした（即応的な）多数のエージェントの方が有効であることを示している．さらに アルゴリズムは並列システムに実装することも容易であり，その有効性はさらに大きいといえる．

われわれが において興味あるのは，より望ましい経路をより速く発見するためにエージェントはいかに協力すべきかという点である． の では，エージェントは独立して経路選択を行うだけで，エージェント間で経路選択に関する調整は行われていない．したがって冗長な経路探索が行われる可能性が大きく，必ずしも効率よい探索が群として行われているとはいえなかった．そこで本論文では，エージェントが相互に協調して組織的に探索を行う手法を研究の対象とする．すなわちエージェントが互いの位置関係を把握することにより，分散と密集に基づく二つのエージェント組織化手法を提案する．対象となる問題に応じて適切に組織化されたエージェント群は，ランダムな経路選択を行なうエージェント群と比べて，探索効率の向上をもたらすことが予想され，その評価のために迷路探索問題 と パズル問題 を用い，シミュレーション実験を通してその有効性を明らかにする．

マルチエージェント実時間探索

グラフに関する基礎概念

まず数学的な準備としてグラフに関する基本的な用語を定義する．

グラフ は節点 の集合 と（有向）辺 の集合 の組 で表される．ここでは， とする．あるグラフ において， であれば， は の子 ， は の親 と呼ぶ．節点の列 は の時に から への経路 と呼び， を経路の長さ と呼ぶ．特に， のとき，その経路を閉路 と呼ぶ．辺にコストが 但し は正の実数の集合 として与えられる場合には，経路 のコストは で与えられる．

木 は親をもたない根 （節点）を除く全ての節点がたった つの親をもつグラフである．子を持たない木の節点は葉 と呼ばれる．根からある節点 までの経路の長さを の深さ と呼ぶ．

問題の定式化

人工知能で扱うような非決定的な問題解決は以下のような状態空間表現により定式化することができる ．

問題 は，組 により与えられる．ここで， は空でない状態 の集合， は状態遷移を生じさせるオペレータ の集合， は初期状態 ， は目標状態 の集合である．この時，組 は状態を節点，オペレータを辺とみなすことによりグラフとなり，状態空間グラフ と呼ぶ．

求めるべき問題の解とは初期状態から目標状態に至る任意の経路ただし，である．オペレータにコストが与えられているとき，解のコストはその経路のコストとして与えられる．あるコストの解が存在し，それよりも小さいコストの解が存在しなければ，その解は最適解であるという．

ヒューリスティック探索

状態空間グラフ探索の最も基本的な操作は，ある状態にオペレータを適用して得られる子状態を求める状態生成である．特に全ての子を生成することを状態展開と呼び，このとき親状態は展開されたという．一般的なグラフ探索は初期状態から始めて，目標状態が発見されるまで，次々と生成された状態を展開することにより実行される．通常，探索過程を記録しておくために，生成された子にはその親に戻るポインタが付けられ，生成された状態とポインタから初期状態を根とした探索木が形成される．

展開する状態の順序によって深さ優先や幅優先など，さまざまな探索手法が利用可能であるが，このような力ずくの探索手法は状態空間が大きい問題に対しては計算量の組合せ的爆発により現実には実行不可能である．そこで，それぞれの状態に評価値を与え，評価値の高い状態の展開を優先することにより，探索効率の向上を図っている．最適解を求める探索では一般に，状態の評価値は以下に示す評価関数により与えている．

ここでは初期状態から状態までの最適経路のコスト，は状態から目標状態までの最適経路のコストである．しかしながら，このとは探索の途中では正しい値を得ることができないので，などのヒューリスティック探索では以下に示す推定値を評価値として用いる．

ここで，は既に得られている探索木の中で，初期状態から状態までの最適経路のコストとする．また，はから目標状態までの最適経路のコストの推定値であり，ヒューリスティック関数として発見的に与えられるものである．アルゴリズムでは解の最適性を保証する適格性の条件として，全ての状態においてが満たされなければならない．

実時間探索

アルゴリズムは経路探索（求解）と移動（解の適用）が分離されたオフライン探索と見なすことができる．アルゴリズムは最適解を保証するが，一般的には問題のサイズ（状態数）に応じて指数的な計算量を必要とするために大きな探索空間を持つ問題に適用することは困難であった．は一定の先読み探索と移動を交互に行うことにより，最適解を保証することはできないけれども計算量を削減することのできるオンライン探索手法として以下に示すアルゴリズムを提案した．

〔初期化〕とする．

〔展開〕を展開して，その子状態の集合をとする．

〔終了判定〕を満たす目標状態が存在するなら，として，終了．

〔先読み探索〕全てのについて，を経由して目標状態に至る場合のの評価値を計算する．ここで，はから深さまでの先

読み探索結果をもとに次のように計算する．

ただし，　　　　は　を根とする深さ　の先読み探索木の葉状態の集合とし，
はその時点で知られている　から　までの最適経路のコストとする．

〔移動候補選択〕　　　　　　　　　　　　となる隣接状態　を求める．複数存在するときはランダムに選択する．

〔推定コスト更新〕　　の値を二番目に小さい
に更新する．もし存在しなければ，　　　　とする．

〔移動〕　　　　とする．

へ．

では移動しながら探索を行うので，状態の評価値には　のような初期状態からのコストを表すは含まれず，目標状態までのコスト　のみを用いている．したがって探索の経過が評価値には現れないので，同じ状態を何度も訪問する無限ループに陥る可能性がある．しかし　　では　　においてが単調増加するように更新されるので，無限ループに陥らず，解が存在すれば必ず発見されるというアルゴリズムの完全性　　　　　　が保証されている．

また，　　として更新される二番目に小さい評価値　　　の値は，状態　から移動の対象となる以外の状態を経由して目標状態まで至る経路の最小推定コストを示している．　　において　を二番目に小さい評価値に更新することは，以前に訪問した状態への再訪問を抑制して，探索効率を向上させることになる．なぜならば，二番目に小さい値に更新することは最小値に更新する場合と比べて，その状態の推定コストを大きくし，移動先としてその状態を選択する可能性を低くするからである．その反面，この値は状態の評価値としては過大評価を与えている場合もある．

さて　　は　　において解の質を改善する二つの方法を示している．一つは先読み探索において先読み深さ　を増大させることであるが，計算量は先読みの深さに対して指数的に増加するという性質がある．もう一つは各状態から目標状態までの推定コスト　を学習する方法である．この場合，先に述べたように　　における　　の更新は過大評価となる場合があるので，最小評価値　　　に更新するアルゴリズムを用いることになる．これによって　の値は徐々に真の値に収束することが保証されるが，同一の状態に繰り返し訪問する回数は　　より多くなり，解を求める探索速度は低下する．また学習する場合には同じ問題を連続的に繰り返し探索する必要があり，計算量も増加する．

マルチエージェント実時間探索

において解の質を改善するもう一つの方法として，　　　　　　はマルチエージェント実時間探索　　　　を提案した．　　　　では，複数のエージェントはそれぞれ自律的，並行的に先読み深さで　　を実行する．　　において最小評価値を持つ移動候補が複数存在するときには，エージェントはランダムに移動先を選択する．

の有効性は次の　点にまとめられる．

〔発見効果〕エージェント数を増加させれば，それだけ異なる経路がより多く探索されることになり，解の質が改善される．また一部のエージェントが探索に行き詰まっても残りのエージェントが探索を続けることができる．

〔学習効果〕エージェント数を増加させれば，推定コストの更新が活発になり，解の質が改善される．

また，　　　における探索時間は，先読み深さの増大に対しては指数的に増大するが，エージェント数の増加に対しては線形にしか増大しないので計算量の面から　　　　　は有利であるといえる．さらに　　　　は並列システムへの実装が容易であり，探索の高速化が期待できるという利点もある．

ただし　　　の提案した　　　　　には二つの問題点がある．第一の問題点は　　の更新手法として二番目に小さい評価値に更新する　　　方式を用いている点である．複数のエージェントが　　を共有する　　　　　では，　　を二番目に小さい評価値に更新することは，状態の評価値としては過大評価になり，他のエージェントの探索を妨げる可能性がある．一方で，　　　　方式の推定コスト更新は，推定距離の過大評価を行なわないが，エージェントの再訪問を抑制しないため探索時間を増加させるという欠点がある．

そこで本研究では，　　　方式と　　　方式を組み合わせた推定コスト更新手法を提案する．すなわち，エージェントが互いに共有する大局的推定コスト　　　と個別に管理する局所的推定コスト　　　を用い，　　　は　　　方式で更新することにより推定コストの過大評価を防ぎ，同時に　　　を　　　方式で更新することにより再訪問を抑制するようにしている．　　　のアルゴリズムの　　　と　　　は以下のように修正される．

〔先読み探索〕全ての　　　　　について，　を経由して目標状態に至る場合の　の評価値　　　　　　　　　　　　　を計算する．ここで，

は未訪問

は訪問済

である．

〔推定コスト更新〕　　　　　　　　　とする．また　　　の値を二番目に小さい　　　　　　　　　　に更新する．もし存在しなければ，

とする．

方式と　　　方式を組合わせた推定コスト更新手法を用いた　　　　　は，　　　　の提案した　　　方式のみのものに比べて，解の質に関して改善が見られた．また探索速度に関してはエージェント密度が高くなるにつれ効果が現れたが，低い場合は逆に探索速度は低下した 寺西　．これはエージェントが密集して探索する場合に　　　方式のみでは，エージェントは互いに探索を妨げあっていることを示している．

第二の問題点はエージェント群の組織化に関してである．　　　　の　　　　　において，同等な評価値をもつ複数の子状態に遭遇した場合，移動候補選択はランダムに行われていたが，これではエージェントの探索経路が重複することを避けることができない．このような探索経路の重複は，その探索の冗長性からエージェントの群としての探索効率を低下させることになる．そこで次節ではエージェントが互いの位置関係を把握することにより，分散と密集に基づいて組織的に探索を行う手法を提案する．

分散と密集に基づくエージェントの組織化

の　　　　　では移動候補選択においてエージェント間で明示的な調整は行われていなかった．そこで本研究では，　　　　　においてエージェント間で調整を行う組織化手法を提案する．対象となる問題に応じて適切に組織化されたエージェント群は，ランダムな経路選択を行なうエージェント群と比べて，探索効率の向上をもたらすことが予想される．組織化手法として，分散　　　　　と密集　　　　　に基づく二つの組織化手法を提案する．分散に基づく組織化は，エージェントが互いに反発しながら移動することにより，より広い探索空間に位置することが促進され，　　　　　の発見効果が強化される．また密集に基づく組織化は，エージェントが互いに誘引しながら移動することにより，探索空間の狭い範囲に密集して，　　　　　の学習効果を強化する．

分散に基づく組織化

分散に基づく組織化手法は，エージェントは互いに他のエージェントとの位置関係を把握し，エージェントが同等に望ましい経路に遭遇した場合には他のエージェントから離れる経路を選択するように調整する．

ここでエージェント間の位置関係を表現するものとして，エージェントの隣接度 を定義する．エージェント が状態 に位置する場合の隣接度 は以下のように計算される．

ここで， は に関わるエージェントの集合， はエージェント が位置する状態， は状態 から状態 までの経路の推定コストとする．すなわち，エージェントの隣接度は経路コストを尺度として最も近くに位置するエージェントとの距離を表している．

次にエージェントが互いに離れるように移動するかどうかを決定する基準として反発領域 を定義する．エージェント の現在の状態 における反発領域 は以下のように計算する．

ここで， は初期状態である．この反発領域はエージェントが初期状態に位置するとき，最も大きくなり，目標状態に近づくにつれて小さくなるように工夫している．この結果，探索初期においては各エージェントは互いに離れる経路を選択する．しかし目標状態へ近づくにつれて反発領域は小さくなり，目標状態に収束する．また，パラメータ を変化させることによっても，反発領域の大きさを調整することができる．

以上のような隣接度 と反発領域 を用いて， アルゴリズムの移動候補選択ステップは以下のように変更される．

〔移動候補選択〕エージェント は移動候補集合 から次のように移動先 を決定する．

となる を選択.

となる を選択.

移動候補が複数存在するとき，ランダムに選択する．

すなわち全ての移動候補が反発領域内にあれば，その中でもっとも隣接度が大きくなるものを選択する．反発領域の外側に候補が存在すれば，その中からランダムに選択する．

密集に基づく組織化

密集に基づく組織化手法では，エージェントは互いに他のエージェントとの位置関係を把握し，エージェントが同等に望ましい経路に遭遇した場合には他のエージェントに近づく経路を選択するように調整する．

ここではエージェント間の分離度 を定義する．エージェント の状態 に位置する場合における分離度 は隣接度と同様に以下のように計算する．

密集による組織化では，エージェントが誘引領域 で示される範囲内にに位置するように経路選択を調整する．すなわち， の値が大きいとき各エージェントはまばらになり， の値が小さいとき各エージェントは互いに密集して探索を進めることになる．

密集に基づく組織化では アルゴリズムの移動候補選択ステップは以下のようになる．

〔移動候補選択〕エージェント　は移動候補集合
から次のように移動先　を決定する．

となる　を選択．

となる　を選択．

移動候補が複数存在するとき，ランダムに選択する．

すなわち全ての移動候補が誘引領域外にあれば，その中でもっとも遠隔度の小さいものを選択する．誘引領域の内側に候補が存在すれば，その中からランダムに選択する．

また，分散に基づく組織化と密集に基づく組織化を併用することも容易である．この場合は　と　で定義されるドーナツ状の領域にエージェントが位置するように経路選択が行われる．以上のような組織化を行う場合の付加計算量はエージェント間の隣接度と分離度の計算に伴うものであり，それは　　である．ここで　は探索に関わるエージェントの集合である．

評価実験と考察

評価問題

における組織化の効果に関する評価を，以下に示す迷路探索問題と　パズル問題を用いて行なう．

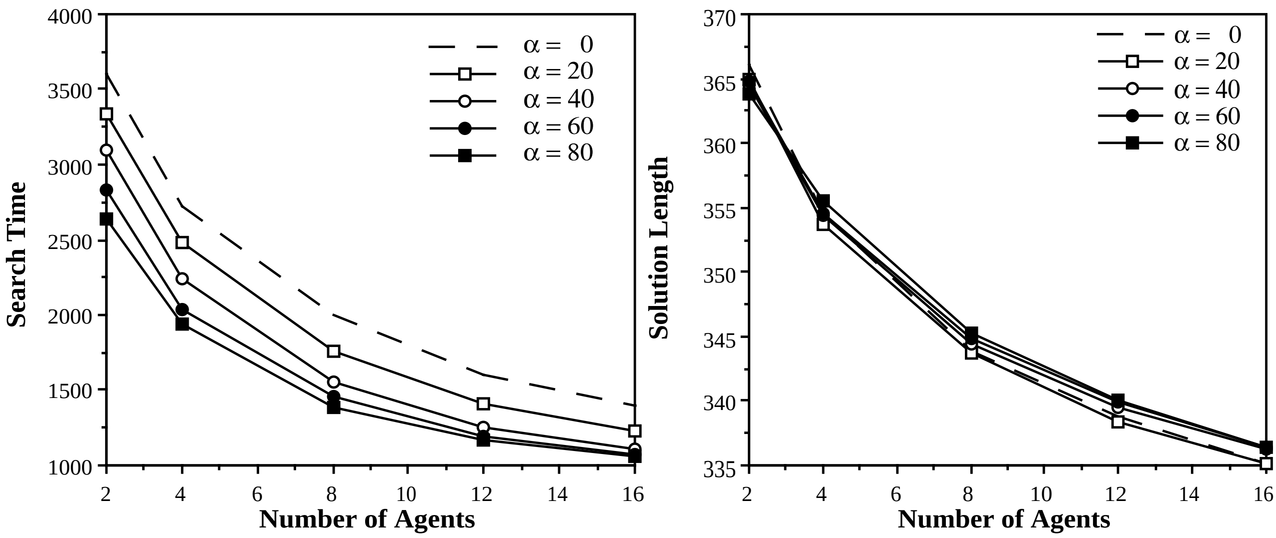
迷路探索問題　格子状グラフ上の迷路において入口から出口までの経路を求める問題である　．エージェントには上下左右のみの移動を許し，斜めには移動できないとする．エージェントの　回の移動につきコスト　かかるとする．また移動不可能なマス目をランダムに発生し，障害物としている．問題には，ランダムに発生させた障害物の比率を　とし，入口　，出口　とする　の格子状グラフを用いている．推定コストの初期値には，目標状態までのユークリッド距離 対角線距離 を用いている．また，隣接度，遠隔度 の計算にもユークリッド距離を用いている．また，ランダムな選択が探索結果に影響を与えるため，題を　回実行してその平均値を取った．実験結果は　個の問題 を解いた平均値である．

パズル問題　パズル問題は，　個の数が付けられた正方形のタイルとブランク タイルの存在しない位置 を含んだ，　の正方形の枠からなっている．ブランクを移動する毎にコスト　かかるとする．タイルの配置を与えられた目標状態まで移動させることが問題となる．推定コストの初期値，隣接度，遠隔度には，タイルの現在位置と目標位置との間のマンハッタン距離の総和を用いている．迷路探索問題と同じく　題を　回実行してその平均値を取った．実験結果は　で示されている　題の内，題の問題 を解いた平均値である．

評価実験

迷路探索問題，　パズル問題ともに，エージェントは並列に動作可能で，回の移動毎に　単位時間かかると仮定している．またエージェントは大局的推定コスト　を共有している．ただし先読み探索や組

解の存在するもののみを選択している．
問題の難度を配慮して上級　問，中級　問，下級　問を選択している．



(a) 探索時間　　(b) 解の質

図　分散に基づく組織化の迷路探索問題における効果

織化などの計算コストは含めておらず，組織化の手法に応じた移動先決定の優劣のみを比較している．実験では探索時間　と解の質　を測定するが，複数のエージェントの中で最初に目標状態に到達したエージェントの所要時間が探索時間となり，その解経路の中で閉路を除いたものが解の質を表している．

分散に基づく組織化の実験では誘引領域を無制限　とし，反発領域のパラメータ　を変化させた．　の時は　の手法と等価になる．迷路探索問題の実験結果を図　，　パズル問題の実験結果を図　に示す．いずれも横軸はエージェント数　である．実験の結果から分散に基づく組織化は探索時間においては迷路探索問題で効果を示したが，　パズル問題では探索時間，解の質ともに性能が低下している．

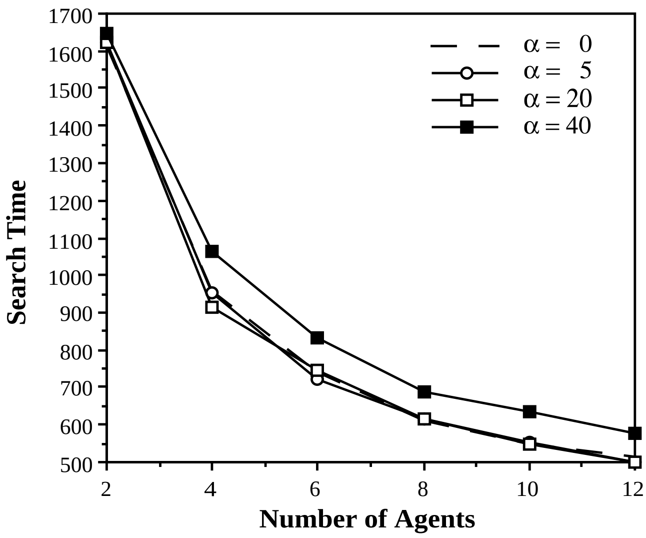
次に密集に基づく組織化の実験では反発領域を　とし，誘引領域　を変化させた．　の時は　の手法と等価になる．迷路探索問題の実験結果を図　，　パズル問題の実験結果を図　に示す．この結果から密集に基づく組織化は迷路探索問題では性能が悪化し，　パズル問題では性能が向上するという，分散に基づく組織化とは対照的な結果が得られた．

考察

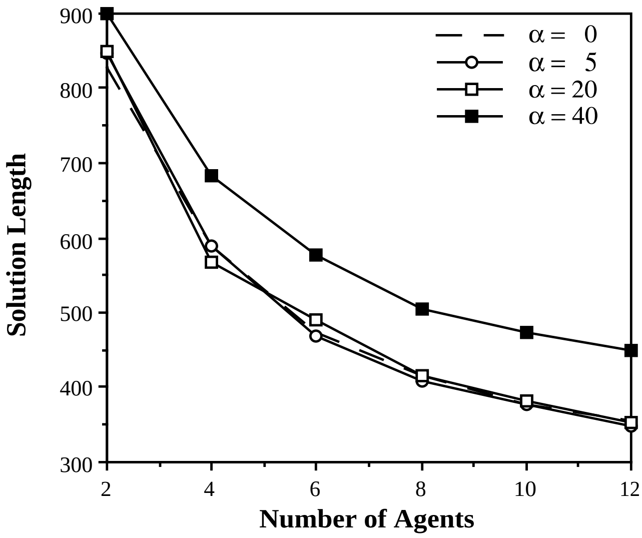
　におけるエージェント組織化手法の効果は推定凹部　石田　を用いて説明できる．推定凹部とは直観的にいえば，目標状態への推定コストの局所的な凹みである．実時間探索を行なうエージェントは単純に最小な評価値を持つ状態を選択し，移動を続けるので，容易に推定凹部の底へ移動してしまう．推定凹部に陥ったエージェントは推定凹部の各状態の推定コスト　の更新を繰り返し，推定凹部を埋めなければそこから抜け出すことはできない．

ここで迷路探索問題と　パズル問題がどのような推定凹部を持っているかどうかを，エージェントが目標状態に到達するまでに通過した各状態の最適コストと推定コストの初期値の差を以下のように測定することで確認してみる．

エージェント数　の　において，エージェントが目標状態に到達するまでに通過した各状態

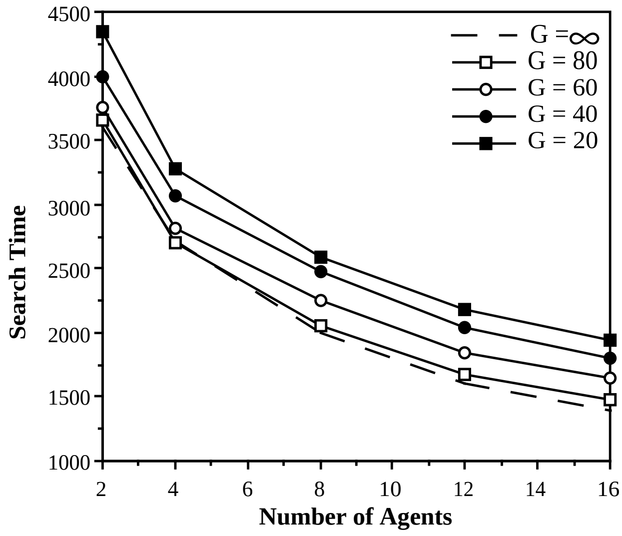


(a) 探索時間

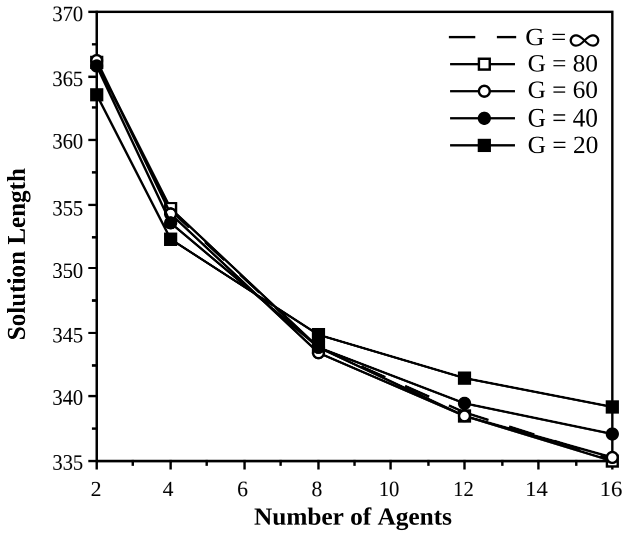


(b) 解の質

図　　分散に基づく組織化の　　パズル問題における効果

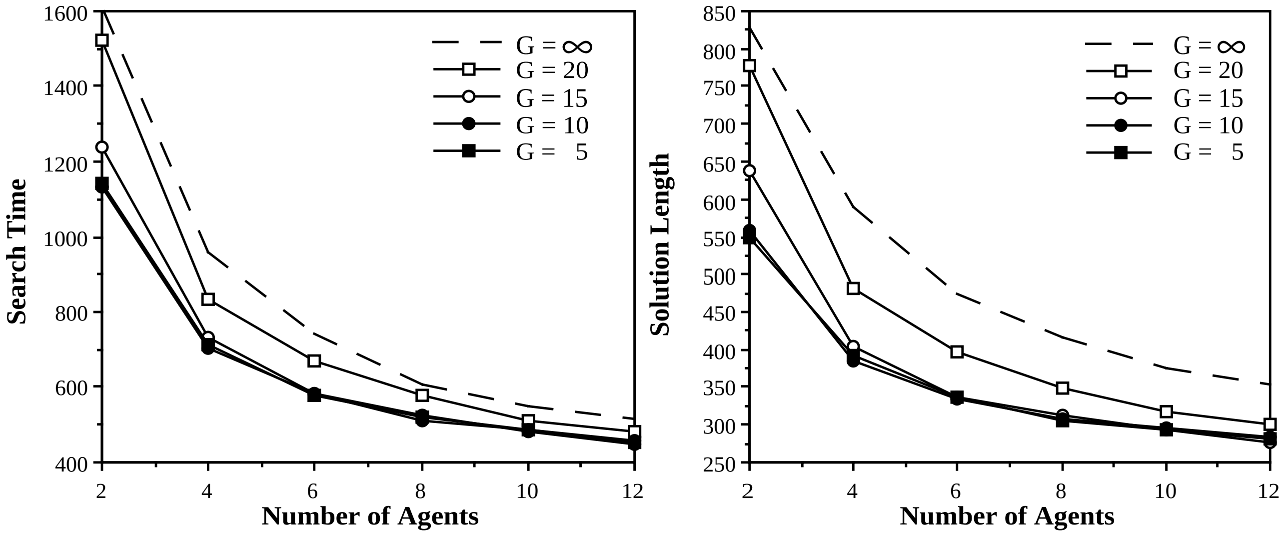


(a) 探索時間



(b) 解の質

図　　密集に基づく組織化の迷路探索問題における効果



(a) 探索時間　　(b) 解の質

図　密集に基づく組織化の　パズル問題における効果

の推定コストの初期値を記憶する．

目標状態から幅優先探索を行なうことにより，目標状態からエージェントが通過した全ての状態までの最適コストを計算し，記憶する．

各状態における最適コストと初期推定コストの差，すなわち推定凹部の深さを計算する．

ただし，　パズル問題では探索空間が非常に大きくなるので，その代わりに　パズル問題と同様の推定凹部を持つと考えられる　パズル問題を用いて測定を行なった．

推定凹部に関して迷路探索問題の測定結果を図　に，　パズル問題のの測定結果を図　に示す．横軸は推定凹部の深さ　，縦軸はその値を持つ状態の割合　である．迷路探索問題では，推定凹部の深さは　から　までの広い範囲に分布し，同じ値である状態の割合は全て　未満である．逆に，　パズル問題においては推定凹部の深さは　から　までの狭い範囲に集中している．これらを比較すると迷路探索問題は　パズル問題に比べて深い推定凹部が多数分布していることが確認できる．

すなわち，迷路探索問題は，障害物による壁の存在がその向こう側にある目標状態に対して推定凹部を作るため，深い推定凹部が状態空間に点在するが，　パズルや　パズル問題では迷路探索問題のような深い推定凹部がほとんど現れず，浅い推定凹部が広い範囲に分布しているといえる．

したがって密集に基づく組織化を迷路探索問題に適用した場合，多くのエージェントが集団として推定凹部の底へ吸い込まれてしまう．一方で，エージェントを互いに反発させる分散に基づく組織化手法では，エージェントはそれぞれ間隔をあけながら移動するので，深い推定凹部に陥らないような経路を選択するエージェントの割合は多くなり，発見効果が有効に働いている．したがって，迷路探索問題には反発領域を大きくし，エージェント群の範囲を広げる分散に基づく組織化手法は有効であるといえる．

しかし　パズル問題では深い推定凹部は少なく，多数の浅い推定凹部が存在する．その際にはエージェントはばらばらで移動するよりも，協力して推定凹部を埋めて行くほうが効果的であるといえる．すなわち，あるエージェントが行なった推定距離の更新を他のエージェントが利用できる密集に基づく組織が適している．誘引領域を狭めれば学習効果が強化され，よりコストの小さい経路を見つける可能性が高くな

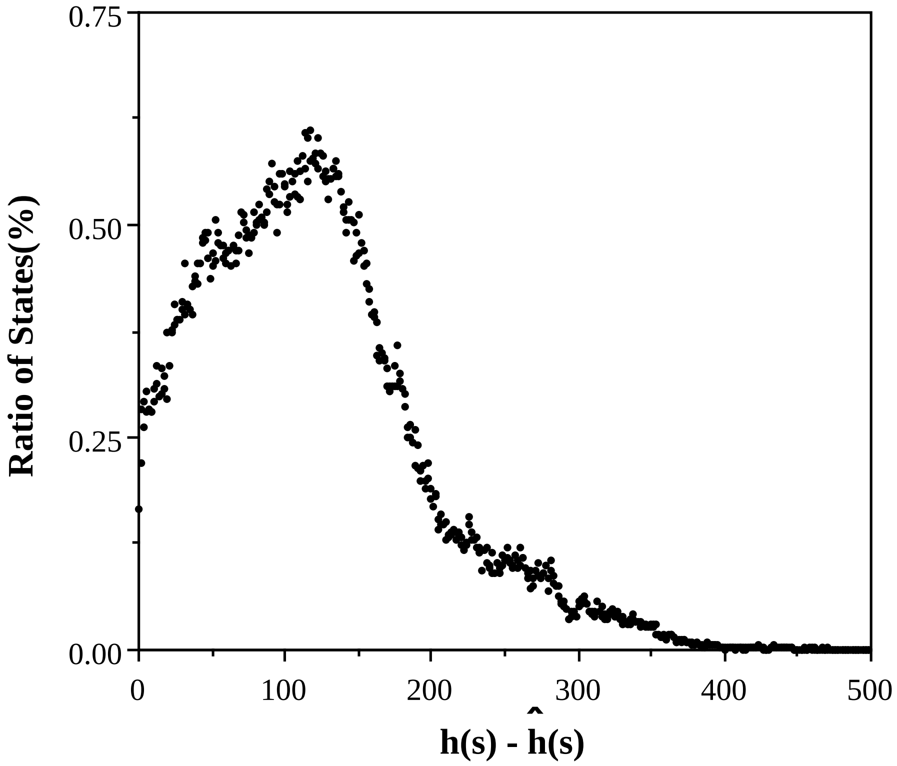


図　　迷路探索問題の推定凹部

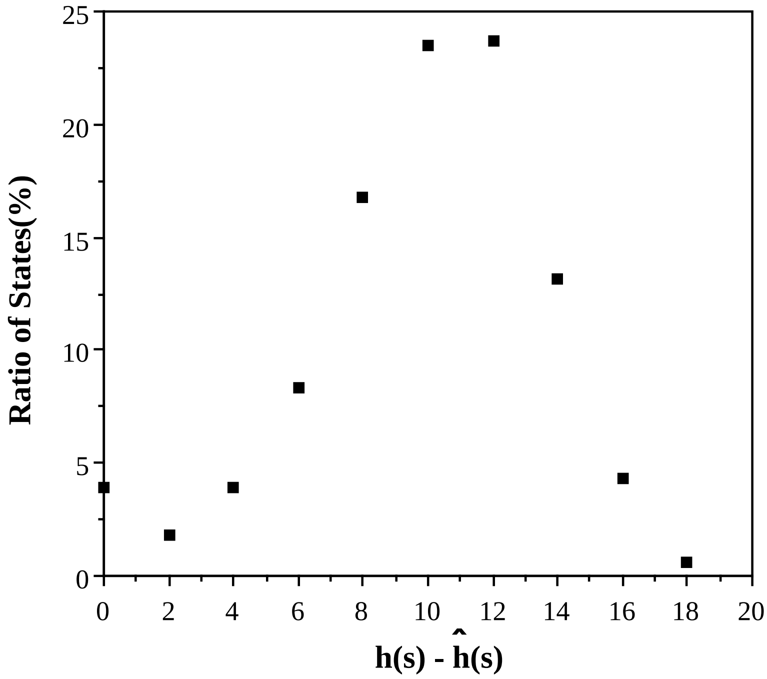


図　　　パズル問題の推定凹部

る．したがって　パズル問題には密集に基づく組織化手法が有効であるといえる．

むすび

マルチエージェント実時間探索　　　　　において，分散と密集に基づく二つの組織化手法の提案とその評価を行った．実時間探索に関わるエージェントはそれぞれ相互の位置関係を把握し，分散の場合はより離れる方向に，密集の場合はより近づく方向に移動しながら探索を行う．　　　　　には発見効果と学習効果の二つの利点があり，分散と密集に基づく組織化はそれぞれを強化する性質がある．評価のために，推定凹部の視点からそれぞれ対照的な特徴を持つ迷路探索問題と　パズル問題を用いた．深い推定凹部が点在する迷路探索問題では発見効果を強化した分散に基づく組織化が，浅い推定凹部が偏在する　パズル問題では学習効果を強化した密集に基づく組織化が有効であることが示された．以上のことから対象となる問題の性質に応じて適切な組織化手法を用いれば，従来の組織化を行わない手法に比べて，探索効率を向上させることができる．推定凹部と有効な組織化の関係は明らかになったといえるが，さらに詳細な反発領域や誘引領域の決定は今後の課題である．また，探索の過程でエージェントを動的に組織化することも興味ある課題の一つといえる．

参考文献

人工知能　共立出版　東京

寺西　寺西憲一　マルチエージェント実時間探索における組織化手法に関する研究　大阪市立大学工学研究科電気工学専攻

石田　石田亨　実時間両方向探索　マルチエージェントと協調計算　近代科学社