

自律エージェントの環境観測に関する考察

水野 升裕 渡辺 尚 水野 忠則

静岡大学 工学部

分散協調システムにおいては、各自律エージェントは適切な副目標を生成するために環境を詳細に観測する必要がある。しかし、観測に要するコストと最終目標を達成するまでの時間の間にはトレードオフの関係が存在する。本研究では、自律エージェントの環境観測、副目標達成の各動作に費す時間に関して考察する。具体的には、典型的な分散協調モデルである「バベルの塔」を対象とし、各エージェントの担当領域の大きさ、観測に要するコスト、最終目標を達成するまでの時間の関係について検討する。シミュレーションによって全コストを定量的に評価し、環境観測コストに対して全コストが最小となる最適観測間隔が存在することを示す。

Environment Observation Cost of Distributed Autonomous Agents

Masuhiko Mizuno Takashi Watanabe

Tadanori Mizuno

Faculty of engineering, Shizuoka University

3-5-1, Johoku, Hamamatsu, 432 Japan

In a distributed cooperation system, each agent has to observe the environment in order to create an adequate subgoal. This paper discusses tradeoff between cost of observation and achievement of the shared final goal of agents moving in a lattice world. The model referred to as the restricted Tower of Babel is proposed including moving area restriction and observation cost. We have been already discussed an effective policy for the basic Tower of Babel in which each agent regenerates its subgoal independently. Based on the policy, this paper proposes a cooperation mechanism called LM-DMax algorithm in the restricted model. Through simulation, we found that the model has an optimal observation interval regarding minimizing total cost to achieve the final goal.

1 はじめに

分散協調システムとは、単独エージェントでは解決できない問題、または非効率的にしか解決できない問題に対して、複数のエージェントが協力・妥協を通して効率的に解を得るシステムである。分散協調システムにおいては、エージェントがある段階において、自律的に副目標を生成し達成する動作を繰り返すことによって、最終目標を達成する。したがって、各エージェントの副目標の生成が全体の効率を大きく左右する。

分散協調システムにおいて、効率よく最終目標を達成するためには、以下の問題を考察しなければならない。

- (1) 1つの最終目標に対して、複数の副目標が生成される。つまり、エージェントは大規模な問題を分割しなければならない。問題分割を複数のエージェントで行うか、または、単独で行うか、そして、分割をどのように行うかによって、問題解決の効率は、大きく左右される。
- (2) あるエージェントにとっての都合のよい副目標が、他のエージェントにとって不都合となることがある。例えば、エージェントが共有し合っている資源に対する競合が発生し、エージェント群が効率よく動作しない場合がある。また、各エージェントが単独で利得の大きい副目標を決定すると、エージェント間で副目標の重複が生じたり、最終目標達成に不可欠な行動をどのエージェントも副目標としない状態も起こり得る。この場合、エージェントの副目標の修正、再生成等が必要となる。また、エージェントが副目標の修正を行える時点も問題となる。
- (3) 環境の動的変化を考えた場合には、環境の観測、副目標の生成(思考)、副目標の達成(行動)が行われることになる。環境の観測、副目標の生成、副目標の達成に必要な時間が、環境の動的変化に対して無視できない場合には、観測と思考にバランスを考慮しなければならない。例えば、環境を詳細に観測し最適な副目標を生成しようとする、環境の観測と副目標の生成に要する時間、す

なわち、オーバーヘッドが大きくなる。しかも、導出された副目標は既に最適でないかもしれない。これに対し、副目標を決定できる時点が少ない場合には、各エージェントがとっている行動の一貫性が保持されないため、必ずしも最終目標を効率的に達成できるとは限らない。

協調メカニズムの基礎的考察を行うモデルとして提案されているものに、Tileworld[1]、追跡問題[2]、バベルの塔[3]などがある。Tileworldでは、格子状の2次元格子空間上にエージェント、タイル、障害物及び穴が存在し、エージェントがタイルを穴に埋めた場合に、得点が得られる。各エージェントの最終目標は、エージェントが、いかに協調して高得点を上げるかにある。Tileworldにおいてはエージェント数、穴の出現率、障害物の出現率、穴の得点分布、穴の大きさ分布、穴とタイルとの距離分布、穴の時間的得点変化、穴のタイマ切れ、動作と思考の時間配分など、非常に多くのパラメータにより、種々の行動選択を考察することができる。

また、追跡問題では、格子状の2次元空間上に2種類のエージェントが存在し、逃避エージェント数を4つの追跡エージェントが囲むことを目標とする。逃避エージェントの行動を追跡エージェントは予測することはできない。

一方、バベルの塔は、自己組織化のためのモデルとして提案されたものである。このモデルでは、格子状の2次元空間上に1から順に番号のふられたブロックが存在し、エージェントがブロックを中央の台座の上に番号順に載せていく。バベルの塔は、他のモデルに比べて、比較的単純なモデルである。上記(1)の問題については、1つのブロックを台座に運ぶことが分割された問題となる。したがって、(1)について考察する余地はあまりない。次に、(2)に対して著者らは[4]において副目標の生成と達成および、修正に関して検討している。この中で、バベルの塔を効率よく解決するヒューリスティックを数種開発し、それらの効率を比較検討することにより、エージェントの副目標生成、修正について考察している。[4]においては、上記(3)の問題に対して、環境観測に必要な時間、副目標の生成(思考)に

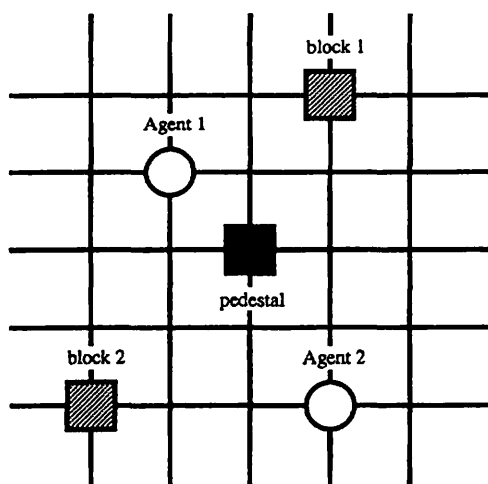


図 1: バベルの塔

必要な時間は、副目標の達成(行動)に要する時間に対して非常に小さいとして無視している。しかしながら、動的に変化する環境や、非常に大きな格子空間やエージェント数を対象とした場合には、この仮定はゆきすぎる。

本研究は、自律エージェントの環境観測、副目標達成の各動作に費す時間に関して考察することを目的とする。すなわち、副目標を生成するために必要となる観測にかかるコストと、ブロックの搬送・移動にかかるコスト、すなわち、不適当な副目標生成によってもたらされるコストのトレードオフを考察するために、エージェントが観測する一定時間間隔を変化させ、エージェントの観測コストを考慮したバベルの塔において、エージェント群の協調動作の考察を行う。

2 バベルの塔

2.1 バベルの塔

バベルの塔(図1)を以下に定義する。

- (1) 格子状の2次元平面上に1から順に番号のふられたブロックが散乱している。
- (2) 平面上に存在する台座の上に1から順にブロックを積み上げることがバベルの塔の最

終目標である。

- (3) エージェントは1つのブロックを自分の上に載せて運ぶ。
- (4) エージェントは1単位時間に隣接する格子点に進むことができる。
- (5) エージェントは問題解決の最中に障害物となるブロックが現れても、それを移動させることはできない。

バベルの塔は、初期状態においてブロックの位置が固定されているため、計画法により最適解を求められる可能性があるが、本研究では以下の理由によりこのアプローチをとらない。

- (1) 非常に大きな格子空間や、ブロック数、エージェント数を対象とした場合、最適解を求めることは非常に困難となる。
- (2) エージェントに観測可能な範囲を限定すると、複雑な制約付きの問題になってしまう。バベルの塔では全空間を瞬時に見渡せると仮定されるが、ロボット制御や交通システムに適應させることを考えると、現実的にはこの仮定は成り立たない。
- (3) バベルの塔の発展形の1つとして、ブロックを動的に発生させる環境を想定することは容易に行え得る。その場合には、事前に最適計画を立てることは不可能となる。

2.2 バベルの塔における副目標

このバベルの塔においては、以下の2種類の副目標が考えられる。この2種類の副目標の決定をどのように行うかが最終目標を達成するまでの効率を大きく左右すると考えられる。

- (1) 副目標とすべき次のブロックの決定
ブロックを運び終えたエージェントが次に運ぶブロックは、エージェントの副目標とみなすことができる。
- (2) 副目標とすべき次の格子点の決定
エージェントが次に移動すべき格子点はエージェントの副目標とみなすことができる。

2.3 エージェントの担当領域

バベルの塔においては、エージェントは対象となる格子空間全体を移動・観測することができる。しかし、パスロケーションシステムや物流システムのように、実際には一部の領域を担当させ、その領域内で作業させるということは日常的に行われている。そこで、エージェントに担当領域を設け、その領域内を移動・観測させるバベルの塔 (以下、移動制約付きバベルの塔と呼ぶ) を考察する。

2.4 観測コストと移動コスト

マルチエージェント環境においては、複数のエージェントが自律的かつ並列的に行動をするため、環境が動的に変化している。この環境においては、副目標を生成する際にその時点に担当できるブロックを認識するために、自分の担当領域の中にあるブロックを観測しなければならない。全観測コストは以下の要素に依存する。

(1) 担当領域の面積

担当領域の面積が大きいと観測コストは大きくなるが、担当できるブロックの数も多くなり、より効率がよくなると考えられるブロックを担当することができ、移動コストが減少する。一方、担当領域が小さいと、場合によってはブロックが1つのエージェントの担当領域に偏り、担当できるブロックがないエージェントが発生するため、バベルの塔を効率的に達成できず移動コストは増加するが、観測コストは小さくなると考えられる。

(2) 観測頻度

より効率がよいと考えられるブロックを副目標とするために、エージェントは自分の担当領域を観測する必要がある。観測頻度が高いと観測コストは大きくなるが、より効率がよいと考えられる副目標を生成することができる。一方、観測頻度が低いと効率がよいと考えられる副目標を生成できないばかりか、他のエージェントによってブロックが移動させられたことも知ることが

できず、非常に効率が悪くなることが考えられる。

(3) 1回の観測にかかる単位観測コスト

1格子点を A ステップ移動するコストで観測できるとき、単位観測コストが A であると定義する。

エージェントが移動しながら観測を行うことも考えられるが、本稿では、簡単のため観測は停止して、かつ、定期的に行うものとする。

また、他のエージェントの副目標を知る手段として通信を利用する。本稿では、通信範囲は格子空間全体とし、通信コストについては考慮しない。

3 協調プロトコル

バベルの塔を効率よく達成するためにエージェントの協調プロトコルを設計した。協調プロトコルは、次の2つからなる。

(1) 協調メカニズム

これは、エージェントが自律的に協調するために必要である。協調メカニズムには環境情報の収集と副目標の生成の2つが含まれる。

(2) 通信プロトコル

エージェント間で情報のやりとりを行う際のとりきめを定義する。

3.1 協調メカニズム

著者らは、バベルの塔を効率よく解決するためのアルゴリズムをすでにいくつか開発している。そのなかで、副目標を達成したとき (台座に到着したとき) に副目標を生成できるアルゴリズムとしては、DMax(q) アルゴリズム (Delay Maximizing Algorithm, 遅れ時間最大化アルゴリズム) が最も効率がよく [4]、さらに、移動制約付きバベルの塔に対しては、LM-DMax(q) アルゴリズム (Limited Movement-Delay Maximizing Algorithm) が有効であることがわかっている [5]。

以下では、 m エージェントと n 個のブロックが存在する環境において、あるエージェントが番号 k のブロック b_k を台座に載せた瞬間とする。

LM-DMax(q) アルゴリズムは、残っているブロックの中で番号が最小のものから q 個のうち、他のエージェントが台座にブロックを置いてから自分がブロックを置くまでの予想の時間 (予想遅れ時間) が最も大きくなるブロックを次の副目標とするアルゴリズムである。ただし、 q 個のうち、自分しか担当できないブロックがある場合には、ただちにそのブロックを担当する。すなわち、エージェント x は以下により副目標 b_j を生成する。

1. b_{k+1} を他のエージェントが担当していないく、 x が b_{k+1} を担当することができれば、 b_{k+1} を副目標とする。
2. 残っているブロックの中で、他のエージェントが担当しているブロックを除く q 個のうち、エージェント x だけが担当することができるブロックがあれば、そのブロックを副目標とする。
3. 残っているブロックの中で、他のエージェントが担当しているブロックを除く q 個のうち、他のエージェントが b_{k-1} を置いてからエージェント x がブロックを置くまでの時間 (遅れ時間) が最も大きいものを副目標とする。

台座の番号が k 番である場合に、 $k+2$ 番以上のブロックを持ったエージェントが台座に接近すると、エージェントはブロックを台座に載せられずに待ち状態になる。このため、エージェントが現在の副目標を達成できないと判断したとき、自主的に副目標を修正し、その場にブロックをおいて、次の副目標を決定する。これを副目標の自主的修正とよぶ。

さらに、格子点競合が存在するときには、他のエージェントやブロックが障害物となり、台座周辺にブロックが集まってしまうと、最終目標を達成できないことが多くなってしまう。そのため、エージェントが自主的修正を行う場合に、台座上のブロックの番号と自分が持っているブロックの番号の差を考慮して、台座からある距離に置く方式を既に提案している。ここで、以下の記号を定義する。

台座の番号：

a

副目標としているブロック番号： b

空き格子点数： c

次の式で与えられる距離 r の格子点上にブロックを置く。

$$r = \lfloor \frac{-1 + \sqrt{1 + 2 \times (b - a - 1) \times c}}{2} \rfloor$$

本稿で用いる副目標生成アルゴリズムは、LM-DMax(q) に副目標の自主的修正と障害物回避を付加したものとする。

3.2 通信プロトコル

LM-DMax(q) アルゴリズムで副目標を生成するために、他のエージェントの副目標を知る必要がある。そのため、通信プロトコルとしては以下のものを設計した。

- (1) 副目標をたずねる (ask-subgoal)
- (2) 副目標を答える (reply-subgoal)

また、通信を行う時点については、観測を行うとき、つまり、定期的に行うものとした。

4 シミュレーションおよび結果

観測間隔および担当領域の大きさがトータルコストに与える影響をシミュレーションによって評価する。ここで、以下の用語を定義する。

重複率 $M(\%)$: 全格子空間に対する、複数のエージェントが担当する領域の占める割合。エージェントの担当領域が大きいと、他のエージェントと重複する領域が大きくなるため、重複率も大きくなる。

4.1 シミュレーション仮定

本研究では、表 1 に示した仮定の下でシミュレーションを行った。

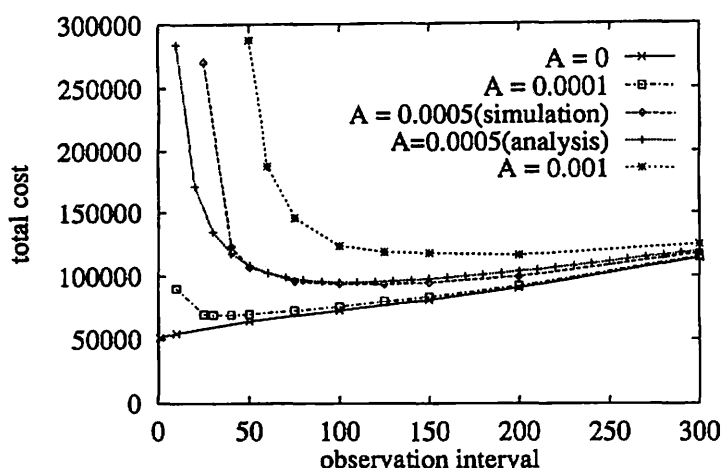


図 2: 観測間隔に対する評価

格子平面の大きさ	200 × 200
エージェント数	4
ブロック数	1000
台座の位置	格子空間中央
エージェントの初期位置	台座
ブロックの位置	ランダム
サンプル数	50
q	10
c	10

表 1: シミュレーション仮定

環境観測の間隔と担当領域の大きさが与える影響について詳細に検討するため、格子点における競合はないものとした。

エージェントは環境を観測するときに、担当領域の面積の大きさ S と単位観測コスト A に比例した時間、すなわち、 $S \times A$ だけ停止する。

4.2 観測間隔に対する評価

観測間隔が効率に与える影響を調べるため、観測間隔 $0 \leq I_o \leq 300$ におけるシミュレーション結果を図 2 に示す。なお、シミュレーション仮定は表 1 に加え、エージェントの担当領域を格子空間全体とした。また、 $A = 0$ は観測に要する

コストがかからないため、これ以上トータルコストが小さくなることはない。

図 2 より、各単位観測コストに対して、トータルコストが最も小さくなる観測間隔が存在することがわかる。

単位観測コスト $A \neq 0$ では、観測間隔が小さくなるとトータルコストが急激に増加していることがわかる。これは、観測間隔が短くなることによって観測に時間を多く必要とし、副目標達成のための行動に時間をかけることができないためであると考えられる。逆に、観測間隔が大きくなるにつれて、徐々にトータルコストが大きくなることもわかる。これは、観測間隔が大きくなるにつれて、実際の環境とエージェントのもつ環境情報にずれが生じ、その不正確な情報をもとにエージェントが副目標を生成するため、適切な副目標を生成することができず、結果的にトータルコストが増加すると考えられる。環境情報のずれが副目標の生成に支障をきたさない程度の観測間隔がトータルコストを最小にすることができると言える。

これらの考察を基に単位観測コストに対する最適な観測間隔を $0 \leq I_o \leq 300$ の範囲で定式化した。トータルコストは移動コスト C_m と観測コスト C_o の和であるから、以下の式が成立す

る[†]。

$$C_t = C_m + C_o \quad (1)$$

単位観測コスト $A = 0$ では、観測コスト $C_o = 0$ であるため、観測間隔の影響だけを受ける。図2から $A = 0$ のときの特性を I_o の線型関数であると近似することができる。すなわち、 $A = 0$ のとき、

$$\begin{aligned} C_t|_{A=0} &= C_m \\ &= C_1 \times I_o + C_2 \end{aligned} \quad (2)$$

とする。(2) 式は、時間をかけて環境観測する間に、始めに観測した情報が古くなり、不適切な副目標を生成することによるコストの増加が観測間隔 I_o の線型関数によって表されることを意味している。

また、観測コストは観測間隔に対しては反比例の関係にあり、単位観測コストと担当領域の面積に対しては比例の関係にある。さらに、式(2)と同様に、単位観測コストに比例して観測時間が増加するため単位観測コストに対して線型関係にあると考えられる。図2より観測コストを求める式は、原点 O を通ることを考慮し、 C_3, C_4 を定数とすると、次の式が成立する。

$$C_o = \frac{(C_3 A + C_4) A S}{I_o} \quad (3)$$

したがって、トータルコスト C_t は以下の式で与えられる。

$$C_t = C_1 I_o + C_2 + \frac{(C_3 A + C_4) A S}{I_o} \quad (4)$$

各エージェントの担当領域の面積は $S = 40000$ である。図2より、 C_1, C_2, C_3, C_4 は、

$$\begin{aligned} C_1 &= 200 \\ C_2 &= 52000 \\ C_3 &= 2 \times 10^7 \\ C_4 &= 1.05 \times 10^5 \end{aligned}$$

[†]複数のエージェントが同時に観測・移動を行うので、(1)式は厳密にはより複雑な式になる。

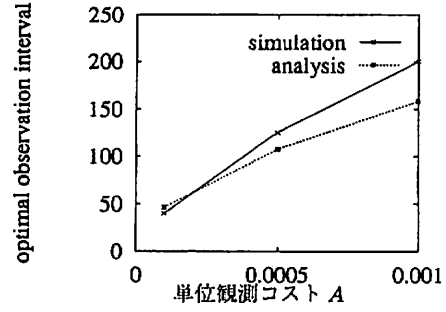


図3: 式(6)で求めた最適観測間隔とシミュレーションから得た最適観測間隔

によって、近似できることがわかった。

ところで、 $\partial C_t / \partial I_o = 0$ となる I_o が最適な観測間隔であるから、(4) 式を I_o で偏微分すると、

$$\frac{\partial C_t}{\partial I_o} = C_1 - \frac{(C_3 A + C_4) A S}{I_o^2} \quad (5)$$

となるから、 $\partial C_t / \partial I_o = 0$ となる I_o は、以下の式で与えられる。

$$I_o = \sqrt{\frac{(C_3 A + C_4) A S}{C_1}} \quad (6)$$

この式によって、単位観測コスト A を与えれば、最適な観測間隔を求めることができる。図2に $A = 0.0005$ の場合の(4) 式の値を示す。

また、(6) 式から求めた最適観測間隔とシミュレーション結果にもとづく最適観測間隔を図3に示す。

4.3 担当領域に対する評価

エージェントの担当領域の大きさがトータルコストに与える影響を調べるため、担当領域の大きさを変化させたときのシミュレーション結果を図4に示す。シミュレーション仮定は表1に加え、観測間隔 $I_o = 100$ 、単位観測コスト $A = 0.0005$ とした。

図4より、重複率 $M = 80\%$ でトータルコストが最小になることがわかる。これは次の理由によるものと考えられる。

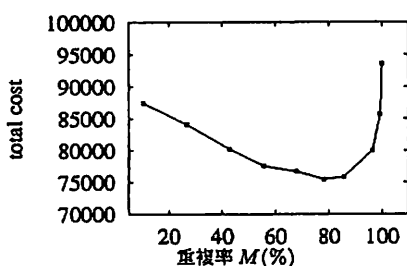


図 4: 担当領域の大きさに対する評価

重複率が 100% に近づくとき、各エージェントの担当領域の面積も大きくなるため、より多くのブロックを副目標とすることができ、効率がよいと思われるブロックを副目標とすることができるため、移動に必要なコストは小さくなる。しかしながら、担当領域の面積の増加によって観測に必要なコストが大きくなり、トータルコストが大きくなる。

一方、重複率が小さくなると、担当領域の面積が小さくなるため、観測に必要なコストは小さくなる。しかし、あるエージェントにブロックが偏った場合には、他のエージェントは副目標とすべきブロックがないという状況が生じ、移動に必要なコストが大きくなるため、トータルコストが大きくなる。

5 まとめ

本研究は自律エージェントの環境観測、副目標達成の各動作に費す時間に関して考察した。典型的な分散協調モデルであるバベルの塔を対象とし、各エージェントの担当領域の大きさ、観測に要するコスト、最終目標を達成するまでの時間の関係について検討した。

本稿では、エージェントに担当領域を設けた制約付きバベルの塔において有効であることがわかっている LM-DMax(q) アルゴリズムを、副目標生成アルゴリズムとして適用した。観測コストを考慮し、環境観測を一定間隔ごとに行った場合に、最小のトータルコストによって最終目標を達成することができる最適観測間隔が存在することをシミュレーションによって示し、さら

に、最適な観測間隔を求める式を定式化した。

また、エージェントに担当領域を設けた場合に最終目標を効率よく達成することができる担当領域の大きさが存在することを示した。

本稿では、エージェント数を固定してきたが、今後の課題として、エージェント数を変化させたときの評価を行うことが挙げられる。さらに、観測間隔を周波数にわりあてて、標準化定理を応用することを今後検討する予定である。

参考文献

- [1] Martha Pollack and Marc Ringuette: "Introducing the Tileworld: Experimentally Evaluating Agent Architectures," In Proceedings of the Eighth National Conference on Artificial Intelligence, p.p.183-189, 1990.
- [2] M.Benda, V.Jagannathan, and R.Dodhiwalla: "On optimal cooperation of knowledge sources," Proceedings of the 1988 Workshop on Distributed Artificial Intelligence, May 1988.
- [3] Toru Ishida: "Towards Organizational Problem Solving," IEEE CONF. ON Robotics and Automation, 1993.
- [4] 山崎 哲哉, 渡辺 尚, 水野 忠則: "分散協調システムにおける副目標の動的生成について", 第 5 回自律分散システム・シンポジウム資料 p.p129-134, Jan 1994
- [5] 水野 升裕, 渡辺 尚, 水野 忠則: "格子状空間を移動するエージェント群の協調動作について", 情報処理学会第 48 回全国大会 4N-5, 1994