

動的環境における自律エージェントの環境観測に関する考察

水野 升裕 佐藤 文明 渡辺 尚 水野 忠則

静岡大学

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。しかしながら、観測するために必要となる観測コストと最終目標を達成するまでのコスト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コストについて考察する。具体的には、典型的な分散協調モデルである「パベルの塔」を対象とし、対象空間の大きさ、エージェント数を変化させ、観測間隔に対する評価を得る。

Environment Observation Cost of Distributed Autonomous Agents in a Dynamic Environment

Masuhiko Mizuno Fumiaki Sato Takashi Watanabe Tadanori Mizuno
Shizuoka University

In a distributed cooperation system, each agent has to observe the environment in order to create an adequate subgoal. Tradeoff exists between cost of observation and achievement of the shared final goal of agents. In this paper, we discuss cost of observation of an environment and achievement of subgoals. Through simulation, we evaluate observation interval against number of agents and moving area size. We found that the model has an optimal observation interval regarding minimizing total cost against number of agents and moving area size, respectively.

1 はじめに

自律エージェントとは、各自の意志決定原理/機構に基づき動作する独立した知的活動体であり、現在、さまざまな目的のために利用され始めている。例えば、仕事の負荷やニュースやメールなどの情報処理の手間を軽減するインテリジェントエージェント [2] やソフトウェアを相互操作できるようにするソフトウェアエージェント [3] などユーザを手助けするエージェントや、複数のエージェントが協調して探索 [4]、プラン認識 [5] などを行って、ある目標を効率的に達成するようなエージェントなどが提案されている [1]。

複数の自律エージェントが協調して目標を達

成するためには、自分がおかれている環境と他のエージェントの意図を把握した上で、効率的な解を得るような副目標を生成することが重要である。環境に関する情報と他エージェントの意図を獲得するための方法としては、

- (1) 通信などにより環境を観測する方法 (明示的方法) [6]
- (2) その時点までに得られた情報を元に環境を予測する方法 (暗示的方法) [7][8]

がある [5]。

マルチエージェントに関する多くの問題の設定は、明示的方法または暗示的方法のいずれの場合であっても、環境観測または予測に要する時間等を見捨てるものと考えている [9]。しか

しながら、現実問題への応用を考えると、環境観測に要する時間、他エージェントとの通信時間、副目標を生成するための思考に要する時間を無視できず、この時間に起因するオーバーヘッドとその効果を考察しなければならない。

著者らは、この問題に対し、まず [15] において自律エージェントの副目標の生成の基礎検討を行った。この中では、典型的なモデルの 1 つである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た。

著者らは [14] において、環境観測による副目標生成の問題に対して、環境観測に必要な時間と副目標の達成に要する時間に対して、観測頻度が最終目標を達成するまでのコストに与える影響を考察した。しかしながら、[14] では、自律エージェントが実施すべき観測頻度が明確化されていなかった。また、初期状態に与えられる観測間隔によっては、環境が非常に変動する場合や初期で与えられた観測間隔が大きすぎる場合があり、適切な副目標を生成できないことがあった。

本研究は、自律エージェントがどの程度の観測頻度で環境を観測すべきかを考察する。まず、エージェント数と格子空間の大きさが最終目標を達成するまでコストにどの程度影響を与えるかをシミュレーションによって評価し考察する。また、エージェントが自律的に観測頻度を設定する適応型環境観測機構を提案する。

2 バベルの塔

2.1 バベルの塔

分散協調問題解決を検討するための典型的なモデルとしては、Tileworld[10]、追跡問題 [11]、MICE[12]、バベルの塔 [13] などがある。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、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 (図 1) は以下のように定義される。

- (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。

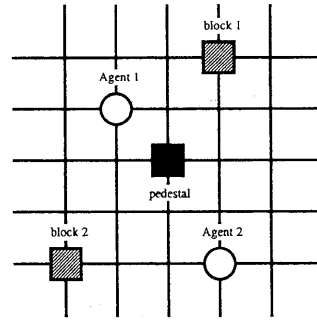


図 1: バベルの塔

- (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乗せて運ぶ。
- (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。

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 環境観測

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に行動をするため、環境が常に変化している。この環境においては、副目標を生成する際にその時点で担当できるブロックを認識し、より効率が良いと思われるブロックを副目標とするために、対象空間

中にあるブロックを観測しなければならない。観測頻度が高いと観測にかかるコスト(以下、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。

また、観測にかかるコストを考慮するため、 A ステップ移動するコストで1つの格子点を観測できるとき、1格子点を観測するコストが 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。エージェントが観測を行ってからまた、次の観測を行うまでの時間を観測間隔(I_0)と呼ぶ。観測コストを含め最終目標を達成するまでにかかるコストをトータルコスト(C_t)と呼ぶ。

さらに、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、通信にかかるコストについては考慮しない。

3 協調プロトコル

我々はこれまでにバベルの塔を効率よく達成するためのエージェントの協調プロトコルを設計してきた。最も効率良く達成できる協調プロトコルとして、LM-DMax(q)を提案した[14]。協調プロトコルは、エージェントが自律的に協調するために必要な協調メカニズムと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通信プロトコルの2つからなる。

3.1 協調メカニズム

本稿では、協調メカニズムとしてLM-DMax(q)アルゴリズム(Limited Movement-Delay Maximizing Algorithm)を用いる。

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。

台座の番号が k 番のとき、 $k + 2$ 番以上のブロックを持ったエージェントは、ブロックを台座に載せられずに待ち状態になる。この場合には自主的に副目標を修正し、その場にブロックにおいて、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。

また、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ふ。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[15]。あるエージェントが台座からある距離にブロックを置くという動作は、他のエージェントにとっては環境が動的に変化するということに他ならない。さらに、エージェント数が大きくなると多くのブロックの配置が変化するため、エージェントにとって環境観測が必要となる。

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、(1) 副目標をたずねるもの(ask-subgoal)と(2) 副目標を答えるもの(reply-subgoal)を用意した。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

動的に変化する環境において、観測間隔がトータルコストに与える影響を調べるため、エージェント数およびブロック数を変化させて、シミュレーションによって評価する。

ブロック数 (n)	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 (l)	50

表 1: シミュレーション仮定

4.1 シミュレーション仮定

本研究では、表 1 に示す仮定のもとで、シミュレーションを行った。また、観測間隔に対する詳細な検討を行うため、格子点における競合はなしとした。ここで、評価に用いる非効率度を定義する。下限 (LB) はすべてのブロックの台座からの往復の距離の和をエージェント数で割ったものであり、下限は完全にすべてのエージェントが並行に動作し、無駄なくすべてのブロックを運んだ場合のステップ数である。また、非効率度とは、各サンプルにおけるステップ数の総和を下限の総和で割ったものであり、非効率度 = 1 は理想値である。

$$LB = \frac{\sum_{j=1}^l d_j}{m}$$

$$I = \frac{\sum_{k=1}^l s_k}{\sum_{k=1}^l LB_k}$$

- m : エージェント数
- d_j : ブロック b_j と台座までの往復の距離
- s_k : サンプル k におけるステップ数
- LB_k : サンプル k における下限
- l : サンプル数

4.2 エージェント数に対する評価

観測間隔 I_0 とエージェント数 m に対する評価を行うため、シミュレーション仮定として、表 1 に加え、格子空間の大きさを 200×200 、1 格子点を観測するコスト $A = 0.0005$ とした。トータルコスト、非効率度に対するシミュレーション結果をそれぞれ図 2、図 3 に示す。図 2 より、各エージェント数に対して、トータルコストが最小となる観測間隔が存在することがわかる。エージェント数それぞれに対して、観測間隔が小さい場合には、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

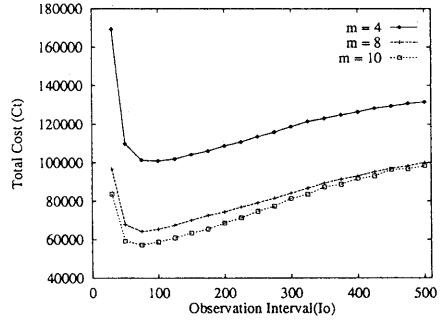


図 2: エージェント数に対する評価 (トータルコスト, $A = 0.0005$)

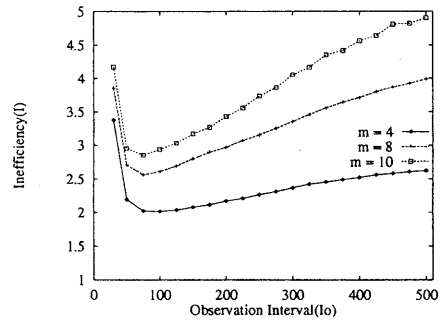


図 3: エージェント数に対する評価 (非効率度, $A = 0.0005$)

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。また、観測間隔が大きい場合には、トータルコストが徐々に増加することがわかる。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る。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、トータルコストが増加することが理由であると考えられる。

また、図 2 からは、エージェント数が多いとトータルコストが低く抑えられるという結果がみられるが、図 3 を併せて考察すると様子が変わってくる。図 3 では、エージェント数に対して非効率度が単調増加の関係になっていることがわかる。エージェント数が小さい場合には、1 エージェントあたりのタスク量が非常に大きく、アイドル状態に陥るエージェントがほとんど存在しないため、エージェントは並行に動作することがで

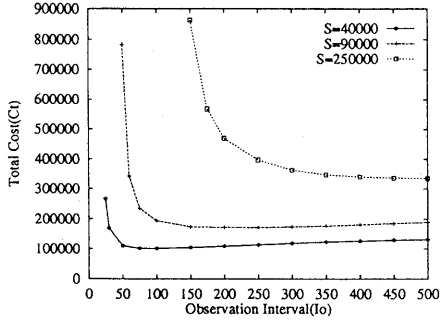


図 4: 格子空間の大きさに対する評価 ($m = 4$, $A = 0.0005$)

きる。一方、エージェント数が大きい場合には、割り当てられるタスクがなくなるエージェントが現れることもある。また、観測がある一定時間をおいてなされるため、あるエージェントにとつて他のエージェントに関する情報にもずれが生じるが、エージェント数が大きい場合には、このずれが大きくなる。そのため、非効率度はエージェント数に対して単調増加となると考えられる。トータルコストを最小限でとどめるためには、格子空間上のエージェント数を考慮し対象空間上における環境変化を予測した上で、観測間隔を設定する必要があるといえる。

4.3 格子空間の大きさに対する評価

観測間隔 I_o と格子空間の大きさに対する評価を行うため、シミュレーション仮定として、表 1 に加え、エージェント数を $m = 4$ 、1 格子点を観測するコスト $A = 0.0005$ とした。シミュレーション結果を図 4 に示す。

図 4 より、格子空間の大きさそれぞれについて、トータルコストを最小にする最適な観測間隔が存在することがわかる。この最適な観測間隔は、格子空間が大きくなるにつれて単調増加する。格子空間が大きい場合に、エージェントは観測する領域が大きくなるため、対象領域を観測するコストは大きくなる。その上、観測間隔が小さい場合には、観測回数も増えるため観測コストが更に大きくなる。この場合には、副目標達成のための行動にかかる時間が少なくなるため、トータルコストが増加する。逆に観測間隔が大きい場合には、各エージェントがもつ情報が現在の状

況と異なりずれが生じるため、トータルコストが増加する。

トータルコストを低く抑えるには、格子空間の大きさに応じて、観測間隔を設定しなければならないといえる。

5 適応型環境観測機構

本章では、環境情報の変化の度合に応じて、エージェントが自律的に観測間隔を設定する適応型環境観測機構を提案する。

5.1 適応型環境観測機構の導入

エージェントに観測間隔を初期状態において与えた場合には、環境が非常に変動する場合や与えられた観測間隔が大きすぎる場合には、適切な副目標を生成することができず、トータルコストが増加するという欠点が存在する。

以上のような状況を打開するために、初期状態の段階でエージェントに観測間隔を一方的に与えるのではなく、エージェントが環境を観測した際に、環境情報の変化の度合に応じて自律的に観測間隔を設定するという適応型環境観測機構の実現が必要となる。

5.2 観測間隔決定アルゴリズム

エージェント a_i が過去に時刻 t_1, t_2 に環境を観測しており ($t_1 < t_2$)、時刻 t_3 に環境を観測した際に ($t_2 < t_3$)、 a_i は以下のように次に観測するまでの時間 I_o を決定する。

$$\begin{aligned} dif1 &:= dif(E(t_2, a_i), E(t_3, a_i), a_i, a_i); \\ dif2 &:= dif(E(t_1, a_i), E(t_2, a_i), a_i, a_i); \\ if (dif2 - dif1) > I_t & \text{ then} \\ & I_o(a_i) := I_o(a_i) \div v; \\ else \\ & I_o(a_i) := I_o(a_i) \times v; \end{aligned}$$

ここで以下の用語を定義する。

$$\begin{aligned} \text{エージェント } i \text{ の観測間隔: } & I_o(a_i) \\ \text{時刻 } t \text{ に } a_i \text{ が観測した環境情報: } & E(t, a_i) \\ E(t_1, a_i) \text{ と } E(t_2, a_j) \text{ の差分: } & dif(t_1, t_2, a_i, a_j) \end{aligned}$$

ここで、 I_t は環境変動に対する閾値、 v は環境観測変動変数 (環境が変動した場合にどの程度観測

間隔を増減させるか)とする($v > 1$)。上記のアルゴリズムでは、過去3回の環境観測結果を用い、環境変動の割合を変換関数 dif によってスカラー値に変換する。この値が環境変動に対する閾値より大きければ、エージェントは環境が大きく変化していると判断し、観測間隔を小さくすることによって、環境変化に敏感に反応しようとする。逆に、閾値より小さければあまり変化していないと認識して、観測間隔を大きくし、観測にかかるコストを減少させようとする。このようにエージェントは環境変化に応じて、観測頻度を自律的に設定し、トータルコストを低く抑えようと努力する。

6 まとめ

本研究は動的に変化する環境において、各エージェントが実施すべき観測頻度について考察した。典型的な分散協調モデルであるバベルの塔を対象とし、エージェント数と格子空間の大きさを変化させてシミュレーションによって評価した。また、エージェントが自律的に環境観測頻度を決定する適応型環境観測機構について述べた。

本稿で述べた適応型環境観測機構においては、エージェントは環境変動の程度を認識し、適切な観測頻度を設定する。しかし、どの程度の環境変動が大きな変化といえるか、また、環境が変化したと認識した場合に、次に観測するまでの時間を決定するための方法はまだ確立されていない。これらの問題は今後の課題であるといえる。

参考文献

- [1] 石田 亨: “エージェントを考える”, 人工知能学会誌, Sep., Vol. 10, No. 5(1995).
- [2] Pattie Maes: “Agents that Reduce Work and Information Overload,” *Comm. of the ACM*, July, Vol. 37, No. 7(1994).
- [3] Michael R. Genesereth and Steven P. Ketchpel: “Software Agents,” *Comm. of the ACM*, July, Vol. 37, No. 7(1994).
- [4] Toru Ishida: “Two is not Always Better than One: Experiences in Real-Time Bidirectional Search,” *International Conference on Multi-agent Systems*, June(1995).
- [5] Marcus J. Huber and Edmund H. Durfee: “Deciding When to Commit To Action During Observation-based Coordination,” *International Conference on Multiagent Systems*, June(1995).
- [6] Smith, R. G.: “The Contract Net Protocol: High-Level Communication and Control in a Distributed Problem Solver”, *IEEE Trans. Comput.*, Vol. 29, No. 12, pp. 1104-1113(1980).
- [7] Genesereth, M.; Ginsberg, M.; and Rosenschein, J.: “Cooperation without communications,” Technical Report 84-36, Stanford Heuristic Programming Project, Computer Science Department, Stanford University, Stanford, California 94305(1984).
- [8] Gmytrasiewicz, P. J.; Durfee, E. H.; and Wehe, D. K.: “A decision-theoretic approach to coordinating multiagent interactions,” *Twelfth International Joint Conference on Artificial Intelligence*(1991).
- [9] Victor Lesser (General chair): *ICMAS-95: First International Conference on Multi-Agent Systems*, The AAAI Press, June(1995).
- [10] M. Pollack and M. Ringuette: “Introducing the Tileworld: Experimentally Evaluating Agent Architectures,” *The Eighth National Conference on Artificial Intelligence*, pp.183-189(1990).
- [11] M. Benda, V. Jagannathan, and R. Dodhiawalla: “On optimal cooperation of knowledge sources,” *The 1988 Workshop on Distributed Artificial Intelligence*, May(1988).
- [12] E. H. Durfee and T. Montgomery: “MICE: A flexible testbed for intelligent coordination experiments,” *The 9th AAAI Distributed Artificial Intelligence Workshop*, pp. 25-40(1989).
- [13] Toru Ishida: “Towards Organizational Problem Solving,” *IEEE International Conference on Robotics and Automation*, pp. 839-845(1993).
- [14] 水野, 渡辺, 水野: “自律エージェントの環境観測に関する考察”, 情報処理学会マルチメディア通信と分散処理ワークショップ論文集, pp.39-46, October(1994).
- [15] T. Watanabe, T. Yamazaki, and M. Mizuno: “Dynamic Subgoal Generation of Autonomous Agents Moving in a Lattice World,” *The Second International Symposium on Autonomous Decentralized Systems*, April(1995).