

多目的が非同期に発生する問題のための マルチエージェントリアクティブプランニング法の提案

Real-time Reactive Planning for Multi-Goals in Dynamic Environment

巖康平*1 森山甲一*2 沼尾正行*2 栗原聡*2
Kohei Iwao Koichi Moriyama Masayuki Numao Satoshi Kurihara

*1大阪大学大学院 情報科学研究科 情報数理学専攻

Department of Information and Physical Sciences, Graduate School of Information Science and Technology, Osaka University

*2大阪大学産業科学研究所

The Institute of Scientific and Industrial Research, Osaka University

The multiagent plan coordination problem arises whenever multiple agents plan to achieve their individual goals independently. Especially in a dynamic environment, the coordination process is required to respond reactively to environmental changes. In this paper, we introduce a novel multiagent planning coordination using shared memory architecture. In this architecture, each agent can plan asynchronously by indirect communication between each other via the shared memory.

1. はじめに

1.1 研究背景

我々が生活する実空間は、構造が複雑であり、しかも常に変化し続ける動的な環境である。これらの環境で動作する自律行動主体には、非同期に発生する複数のゴールに対応し、プランニングを行わなければならない。複数のゴールへのプランニングを行う場合、ゴールごとにその実現を担当する複数の Agent から成るマルチエージェントシステム (MAS) を用いることが多い。そして MAS では、各 Agent 間のプランニングに矛盾や干渉が起きないように何らかの協調機構が必要となる。MAS における協調をテーマとした従来研究 [1, 2] は、Agent 間の結びつきが強く、協調動作が完了して初めてプランを生成できるものが多い。しかし、実環境のように動的な環境ではこの方法では即応性が低下し正しく動作しない可能性が高いため、新たな協調手法を考える必要がある。

1.2 研究目的

本田らは複数の Agent が共有メモリを介して間接的に協調を行うことで Agent 間の独立性を高め、動的環境下でも効果的に動作するプランニング法が提案している [4]。本研究では、この本田らの提案するプランニング法について詳細な検証を行い、判明した Agent 間の不平等性の解決策を提案する。

2. Action Value 更新による時間制約下での間接的協調

本章では、本田らの提案するプランニング法について説明する。

2.1 Action Value の導入

0 から 1 の間で変化する Action Value という値を新たに考え、この Action Value の値の大小を比較することで、プラン

協調時にどちらのアクションを実行すべきか判断基準を得ることができるように定義する。

Action Value は状態空間グラフ (図 1) 上のアクションの重要性であり、その値はそのアクションを選択し中間目標状態まで遷移する場合の推定コストと代替アクションを選択し中間目標状態まで遷移する場合の推定コストを比較して以下のように定義する。*1

$$\begin{aligned} \text{ActionValue}(i, j) &= \begin{cases} 1 & (n = 1) \\ \frac{\sum_{l=1}^n \text{TotalCost}(i, l) - \text{TotalCost}(i, j)}{(n-1)\sum_{l=1}^n \text{TotalCost}(i, l)} & (n > 1) \end{cases} \end{aligned}$$

ここで

$$\begin{aligned} \text{TotalCost}(i, j) &= \text{COST}(S_i^{(k)}, S_{i+1}^{(j)}) + E\text{-COST}(S_{i+1}^{(j)}, \text{subgoal}) \end{aligned}$$

この Action Value により、Action Value が 1 のアクションは代替案が存在しないためプラン実行に必須のアクション、高 Action Value のアクションは代替案が非効率であるため競合する他のアクションよりも優先すべきアクションといった判断が可能となる。

2.2 システム構成と動作プロセス

各ゴールを担当する複数の Personal Agent は他の Personal Agent と直接情報を交換するのではなく、Shared Memory を介して間接的に情報の共有を図る (図 2)。Personal Agent はプランを生成する Planner とプランを実行する Executor から成り、Washington R. によるプランニングと実行の統合モデル [3] に従って動作する。

Planner は RTA*等のアルゴリズムを用いて中間目標状態 S_{subgoal} までのプラン及び初期 Action Value を求め Shared

*1 本論文の 2.4 節までは同氏の人工知能学会 (2008) への投稿論文より引用したものである。

連絡先: 巖康平, 大阪大学 産業科学研究所 沼尾研究室, 〒567-0047 大阪府茨木市美穂ヶ丘 8-1, Tel:06-6879-8426, Fax:06-6879-8428, E-mail:iwao@ai.sanken.osaka-u.ac.jp

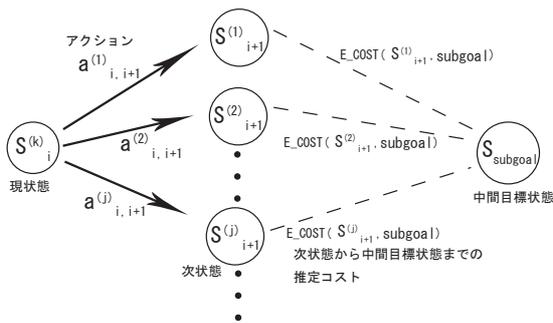


図 1: 状態空間と Action Value

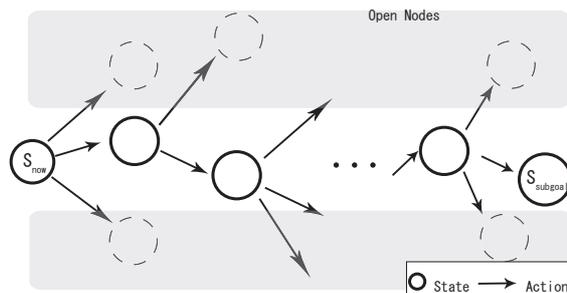


図 3: プランニング過程

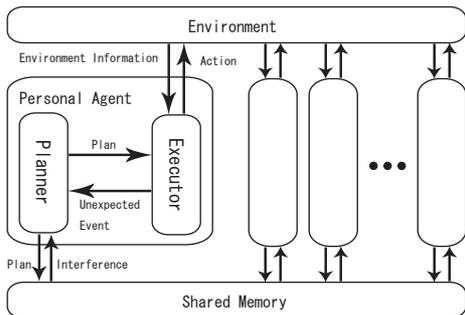


図 2: システム構成

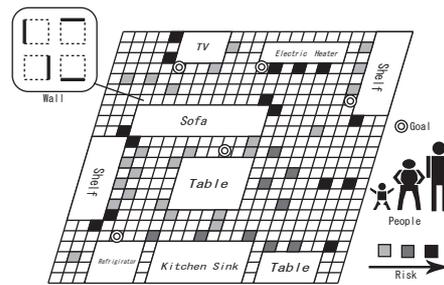


図 4: シミュレータ

Memory に渡す。そして Executor からのプラン実行失敗通知や Personal Agent からの干渉発生通知を受けない限り、未展開ノード (Open Nodes) 中の各状態から $S_{subgoal}$ までの推定コストの計算を繰り返し行う (図 3)。

Personal Agent は Shared Memory に置いたプランと他の Personal Agent のプランを常にモニタしている。干渉が発生した場合、つまり自身のアクションが他者のアクションの効果 (または前提条件) を否定するような場合、2.1 節で述べたように Action Value を用いて判断を行う。この際実行すべきでないプランを渡した Personal Agent が再プランを行うが、 $S_{subgoal}$ までの最善の代替案を準備しているためその影響は最小限に留まる。

2.3 検証シミュレータ

当手法の検証としてシミュレータ実験を行った。本シミュレータは、人々 (People)、目的地 (Goal)、リスク (Risk)、家具 (Table 等)、壁 (Wall) から構成される (図 4)。このシミュレータ上で、人々をリスクから回避させつつ目的地まで誘導するために必要な壁操作のプランニングを行う。各人をサポートする Personal Agent の壁操作は場合によっては他の Personal Agent と干渉する*2 こともあるため適切に協調を行う必要がある。

2.4 検証実験結果

リスク数と壁生成に必要なリードタイムの二つのパラメータを変化させ、100 人が移動を行った際の、目的地へ到達できた人の割合 (以下、目的地到達率と呼ぶ) を測定したところ、提案手法から協調機構を除いた $Planning_A$ と比較した場合、干渉が少ない状況では両者に差がないものの、干渉が多発する状況では提案手法側で協調による目的地到達率の向上が確認で

きた。

3. 問題点

3.1 プラン同士の干渉による各 Personal Agent への影響検証

本田らの行った 2.4 章の実験によって、Action Value 更新による間接的協調が到達率の向上に効果があると推測できるものの、干渉が各 Personal Agent にどのような影響を与えているのかは不明瞭である。

そこで、先と同じシミュレータを用いて Personal Agent 個々について更なる検証を行った。今回はリスク数は 40、壁生成に必要なリードタイムは 3 と一定とし、シミュレータ上の人数 n を 100, 150, 200, 250 とし、移動を行った際の、人を目的地に到達させることの出来た Personal Agent について、生成後初めて構築したプランでの目的地までの歩数 (以下 Plansteps) と、実際に目的地までに要した歩数 (以下 Goalsteps) とを比較した。

3.2 検証実験の結果と考察

人を目的地に到達させることの出来た各 Personal Agent を、 $Goalsteps/Plansteps$ の値の小さなものから順に並べたものを図 5 に示す。

目的地に到達できた人の大半は、初めに予定していた Plansteps 通りの歩数で到達できているものの、一方で図の右端に見られるように、初めの予定の 2 倍以上の歩数を要して目的地に到達している人が存在していることが確認できる。

これは、先の提案手法が、シミュレータ上の全体のコストを低くするようなアルゴリズムとなっているため、全体の合理性を優先した結果、一部の Personal Agent が何度も自らのプランの変更を余儀なくされているためだと考えられる。

*2 生成した壁により他者が通路を移動できなくなる、撤去したため他者が必要なタイミングに壁を生成できなくなる等の例が挙げられる。

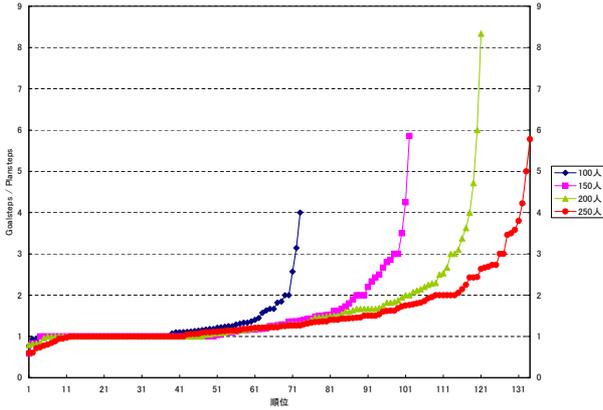


図 5: 各 Personal Agent の Goalsteps/Plansteps

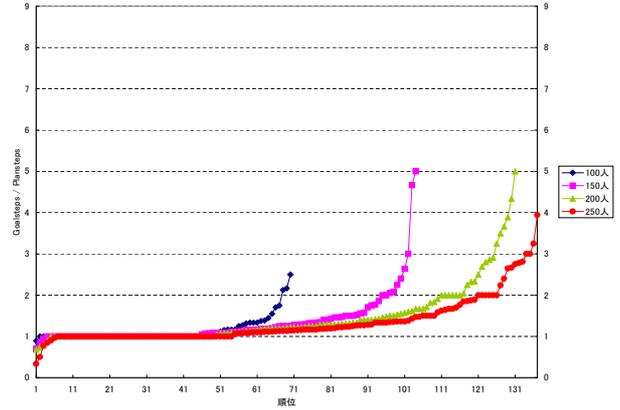


図 6: 各 Personal Agent の Goalsteps/Plansteps

4. Action Value 補正による Personal Agent 間の不平等性の緩和

4.1 Action Value の補正 1

2.1 節で提案している Action Value に、何度もプランの変更を強いられた Personal Agent については優先度を上げるような補正を行うことで、目的地到達率を出来るだけ維持しながら、3.2 節で確認された Personal Agent 間の不平等性の緩和を試みる。

補正を行った Action Value を New Action Value とし、新しく以下のように定義する。

$$NewActionValue = \frac{ActionValue + \alpha}{2}$$

$$\alpha = \frac{n^k - 1}{n^{k_{max}} - 1}$$

ここで、 k はプラン変更時 1 増加し、プラン変更回避時に 1 減少する変数とし、 k についての最大値 k_{max} を設ける。 k_{max}, n は共に任意の自然数とする。

右辺第 2 項によって、New Action Value の値はプランの変更回数が多くなれば全体の合理性にかかわらず大きくなり、それによってプランの変更回数の一部の Personal Agent への偏りを防ぐことが出来ると考えられる。

4.2 Action Value の補正 2

当節では、4.1 節とは異なる補正を用いて、Personal Agent 間の不平等性の緩和を試みる。補正を行った Action Value を New Action Value とし、新しく以下のように定義する。

$$NewActionValue = \frac{ActionValue + \alpha}{2}$$

$$\alpha = \frac{\log_n(k+1)}{\log_n(k_{max}+1)}$$

ここで k は先と同様、プラン変更時 1 増加し、プラン変更回避時に 1 減少する変数とし、 k についての最大値 k_{max} を設ける。 k_{max}, n は共に任意の自然数とする。

4.1 節で提案した α は k の指数関数であるため、プランの変更回数が多ければ多いほど α が大きくなるものの、 α を規格化してしまっているため、 k が小さいとき α は非常に小さな値になってしまう。そこで、 α を本節のように k の対数関数に依存させることで、プランの変更回数が多いときの補正よりも、変更回数が少しでもあるときの補正を優先することで、Personal Agent 間の不平等性の緩和を試みる。

5. 補正検証実験

5.1 補正 1 検証実験

3.1 節と同様に、リスク数は 40、壁生成に必要なリードタイムは 3 と一定とし、シミュレータ上の人数 n を 100, 150, 200, 250 とし、移動を行った際の、人を目的地に到達させることの出来た Personal Agent について、生成後初めて構築したプランでの目的地までの歩数 (以下 Plansteps) と、実際に目的地までに要した歩数 (以下 Goalsteps) とを比較した。なお、 k_{max}, n については、 $k_{max} = 5, n = 3$ とした。

5.2 補正 1 検証実験の結果と考察

人を目的地に到達させることの出来た各 Personal Agent を、Goalsteps/Plansteps の値の小さなものから順に並べたものを図 6 に示す。

図の右端に、初めの予定の 2 倍以上の歩数を要している人は変わらず確認されたものの、3.2 と比較して、その傾向は緩和されていることが確認できる。

特に 200 人、250 人のとき効果がよく現れているのは、4.1 節で提案した補正方法が変更回数が多いときにより効果を発揮するため、比較的干渉の起こりにくい 100 人、150 人のとき効果が現れず、反対に 200 人、250 人のとき効果が顕著になりやすいためと考えられる。

また、表 1 に補正 1 を用いる前と後の目的地到達率を示す。これら目的地到達率を比較すると、補正 1 を用いても目的地到達率は維持できていると考えられ、New Action Value によって目的地到達率の維持と、Personal Agent 間の不平等性の緩和を実現できていると言える。

5.3 補正 2 検証実験

3.1 節、5.1 節と同様に、リスク数は 40、壁生成に必要なリードタイムは 3 と一定とし、シミュレータ上の人数 n を 100, 150, 200, 250 とし、移動を行った際の、人を目的地に到達させることの出来た Personal Agent について、生成後初めて構築した

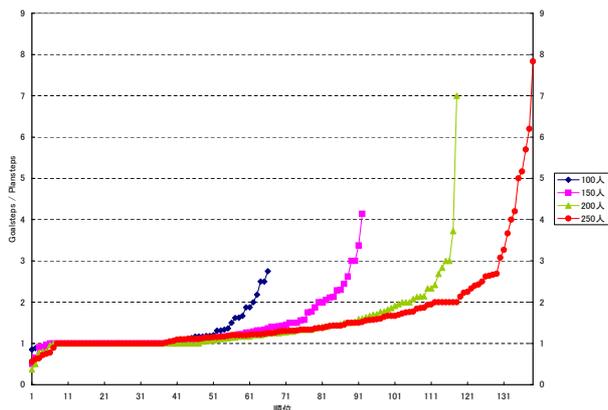


図 7: 各 Personal Agent の Goalsteps/Plansteps

プランでの目的地までの歩数 (以下 Plansteps) と、実際に目的地までに要した歩数 (以下 Goalsteps) とを比較した。なお、 $k_{max,n}$ については、 $k_{max} = 5, n = e$ とした。

5.4 補正 2 検証実験の結果と考察

人を目的地に到達させることの出来た各 Personal Agent を、Goalsteps/Plansteps の値の小さなものから順に並べたものを図 7 に示す。

200 人、250 人のときは補正を行わないときと比べても効果は無く、一方 100 人、150 人のとき補正によって Goalsteps/Plansteps の値の偏りを緩和できていることが確認できる。

これは、4.2 節で提案した補正方法が変更回数の少ないときに効果を発揮するため、比較的干渉の起こりやすい 200 人、250 人のとき効果が現れず、反対に 100 人、150 人のとき効果が顕著になったと考えられる。

また、先の表 1 に補正 2 を用いたときの目的地到達率を加えて示す。これら補正 2 を用いた場合と、補正を行わなかった場合の目的地到達率を比較すると、200 人、250 人のときの目的地到達率は維持できているものの、100 人、150 人のとき目的地到達率がやや減少している。これは、上記のように補正 2 の影響が各 Personal Agent のプラン変更回数の少ないとき、すなわちプランの干渉の起こりにくい人数の少ないときに現れやすいため、全体の合理性を担う Action Value の項の影響が相対的に小さくなってしまい、目的地到達率が減少したものと考えられる。つまり、図 7 に現れている偏りの緩和は、人を目的地に到達させることの出来た Personal Agent の減少によるものであり、補正による効果ではないとも考えられ、更なる検証が必要である。

目的地到達率 [%] \ 人数	100 人	150 人	200 人	250 人
補正なし	73.0	68.0	60.5	53.6
補正 1 使用	70.0	69.3	65.5	54.8
補正 2 使用	66.0	61.3	59.0	55.6

表 1: 補正 1, 2 検証実験:目的地到達率 [%]

6. まとめ

本研究では、すでに提案されている Action Value 更新による間接的協調を用いた即応性のある協調機構について更なる検

証を行い、明らかとなった不平等性の緩和を試みた。緩和方法について異なる手法が必要であるほか、今後ユビキタスホームのように動的環境下で、複数のタスクに対処する様々なシステムへの適用が考えられる。

参考文献

- [1] Jeffrey S. Cox and Edmund H. Durfee. An efficient algorithm for multiagent plan coordination. In *Fifth International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-06)*, 2006.
- [2] Yannis Dimopoulos and Pavlos Moraitis. Multi-agent coordination and cooperation through classical planning. In *IEEE/WIC/ACM International Conference on Intelligent Agent Technology (IAT'06)*, 2006.
- [3] Washington R. Incremental planning for truly integrated planning and reaction. In *Fifth Scandinavian Conference on Artificial Intelligence (SCAI 1995)*, 1995.
- [4] 本田誠一. 多ゴールが非同期に発生する環境下における実時間リアクティブプランニング法の研究. Master's thesis, 大阪大学 大学院 情報科学研究科, 2008.