

複数の目的地・エージェントを考慮した Value Iteration Networks による経路探索

蔡金雨 李子龍 長名優子

東京工科大学 コンピュータサイエンス学部

1 はじめに

経路探索問題に関する研究の歴史は古く、様々な手法が研究されている。特に、近年では、深層強化学習を利用した経路探索に対する研究も盛んに行われている。そのような手法の1つとして、Deep Q-Network[1]を用いた経路探索が提案されている。Deep Q-NetworkはDeep Learningの代表的な手法である畳み込みニューラルネットワーク[2]を用いて強化学習の代表的な手法であるQ Learning[3]を実現する手法である。この手法では、ある状態のときにどの行動をとるかを試行錯誤によって学習することができる。しかし、計画をたてて行動することはできないため、環境の変化に弱く、別の環境で学習した情報を別の環境で利用することは難しいという問題がある。

それに対し、Value Iteration Networks[4]という手法が提案されている。この手法では、プランニング自体を畳み込みニューラルネットワークを用いて学習することで、他の環境で学習を行うことで未知の環境でもゴールにたどり着くことが可能となっている。しかしながら、Value Iteration Networksを用いた経路探索では、ゴールが複数存在するような環境に対応できない、複数のエージェントが存在するような状況に対応できないという問題がある。

本研究では複数の目的地・エージェントを考慮した Value Iteration Networks による経路探索を提案する。

2 Value Iteration Networks

図1に Value Iteration Networks の処理の流れを示す。Value Iteration Networks の処理は、(1) 畳み込みニューラルネットワークによる報酬と状態遷移の推定、(2) VI (Value Iteration) モジュールによる報酬マップからの価値マップの生成、(3) 注意、(4) 行動の選択の4つに大きく分けることができる。

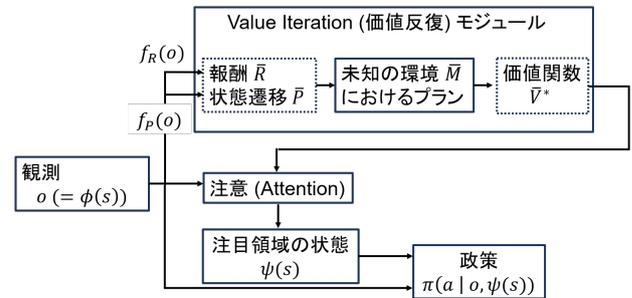


図1: Value Iteration Networks の処理の流れ

3 複数の目的地・エージェントを考慮した Value Iteration Networks による経路探索

ここでは、提案する複数の目的地・エージェントを考慮した Value Iteration Networks による経路探索について説明する。提案手法では

- (1) 複数の目的地を考慮した経路探索
- (2) 複数のエージェントを考慮した経路探索
- (3) 複数の目的地・エージェントを考慮した経路探索

を実現する。

3.1 対象とする課題

エージェントが動き回るフィールドはグリッド状に区切られたものを想定し、1つのマスには1つのエージェントのみが存在できるものとする。エージェントはフィールド全体を観測できるものとし、フィールドの画像とエージェントの位置を観測として使用する。エージェントは8近傍のいずれかのマスに移動するような行動をとる。

3.2 複数の目的地を考慮した経路探索

複数の目的地を考慮した経路探索は出口が複数あり、いずれかの出口に到達できればよいことを目的とした経路探索である。複数のドアがある部屋にいて、どこからでもいいのでできるだけ短時間で部屋の外に出たいような場合に相当する。複数の出口が存在する

Route Search by Value Iteration Networks considering Multiple Destinations and Agents
Cai Jinyu, Li Zilong and Yuko Osana (Tokyo University of Technology, osana@stf.teu.ac.jp)

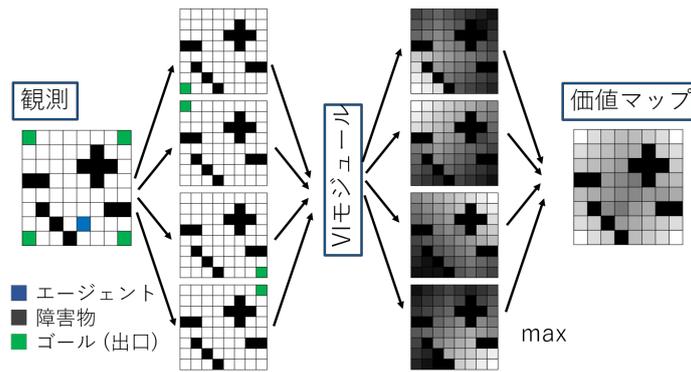


図 2: 複数の目的地を考慮した経路探索

場合には、それぞれの出口に対する価値マップを VI モジュールを用いて生成する。すべての価値マップに対して MAX をとった価値マップを生成し、その価値マップに基づいて行動を決定することになる。図 2 の例では出口が 4 つあるのでそれぞれの出口に対応する 4 つの価値マップが VI モジュールで生成され、それを統合して最終的な価値マップが作られることになる。

3.3 複数のエージェントを考慮した経路探索

エージェントが複数存在する場合には、VI モジュールで得られた価値マップだけでなく、混雑状況も考慮した価値に基づいて行動を決定することになる。エージェントは障害物がなく、他のエージェントがいないマスにのみ移動することができる。エージェントがいるマスには移動できないため、エージェントをどのような順番で移動させるかが問題になってくるが、提案手法では出口に近いエージェントから順に移動の処理を行うものとする。

3.4 複数の目的地・エージェントを考慮した経路探索

複数の目的地・エージェントを考慮した経路探索は 3.2, 3.3 の 2 つの手法を合わせた形になる。まず、それぞれの出口に対する価値マップを VI モジュールを用いて生成し、すべての価値マップを合わせた価値マップを生成する。これはすべてのエージェントに対して共通したものになる。次に着目するエージェントが 8 近傍に移動した場合のそれ以降の経路を仮に取得し、その経路上に存在する混雑領域を考慮した 8 近傍の価値を算出する。この処理はそれぞれのエージェントに対して行うことになる。算出された価値に基づいて行

動を選択し、移動を行う。この処理は出口に近いエージェントから順に実行される。なお、このとき他のエージェントがいるマスに移動するような行動は選択することができない。

4 計算機実験

提案手法を用いて複数の環境において実験を行い、複数の目的地・エージェントを考慮した経路探索が行えることを確認した。

参考文献

- [1] V. Mnih, K. Kavukcuoglu, D. Silver, A. Grave, I. Antonoglou, D. Wierstra and M. Riedmiller : “Playing Atari with deep reinforcement learning,” NIPS Deep Learning Workshop, 2013.
- [2] V.Mnih et al. : “Human-level control through deep reinforcement learning,” Nature, No.518, pp.529–533, 2015.
- [3] C. J. C. H. Watkins and P. Dayan : “Technical Note: Q-Learning,” Machine Learning, Vol.8, pp.55–68, 1992.
- [4] A. Tamar Y. Wu, G. Thomas, S. Levine and P. Abbeel : “Value iteration networks,” Annual Conference on Neural Information Processing Systems, 2016.