

強化学習の実践

参加者 安田 尚平, 小林 陸人, 柏村 峻, 濱田 崇志, 谷水 周平

企画概要

この企画では、自律移動型ロボットの競技会であるロボカップ@ホームリーグ シミュレーション部門への出場を目指し、そしてその競技会への参加を通してロボットの制御方法から強化学習を学びました。

強化学習とは

環境に対する最適な行動を学習する。

● 学習方法 (α : 学習率, γ : 割引率)

■ SARSA

$$Q(s_t, a_t) = Q(s_t, a_t) + \alpha(r_{t+1} + \gamma Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t))$$

■ Q-Learning

$$Q(s_t, a_t) = Q(s_t, a_t) + \alpha(r_{t+1} + \gamma \max_{a'} Q(s_{t+1}, a') - Q(s_t, a_t))$$

● 行動選択方法

■ ϵ -greedy法

● 確率 ϵ ($0 < \epsilon \leq 1$) でランダムな行動選択

● 確率 $1 - \epsilon$ で最大の価値関数になる行動を選択

Soft max法

● 価値関数に重みを付け

確率的に行動選択



ロボカップとは

- 人が操作するのではなくロボット自身がその環境を感知し考えて動く自律移動型ロボットによる競技会
- 今回私たちはその中のロボカップ@ホームリーグのシミュレーションオープンプラットフォームリーグに参加しました

ロボカップ@ホームリーグとは

- 日常生活でのロボットの利用を想定した部門
- 「家庭環境においていかに人間との暮らしに役立つ作業を遂行できるか」を競技内で競う



シミュレーション部門とは

- 仮想空間上で競技を実行
- Handyman
- Interactive Cleanup
- Human Navigation
- 以上の3つのタスクを実践して競い合う



INTERACTIVE CLEANUP

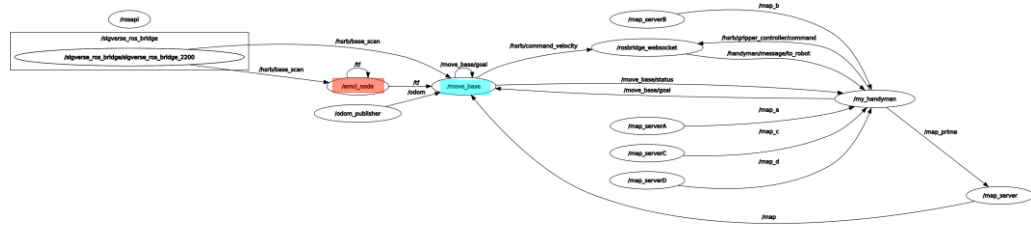
- ユーザの指さしジェスチャによって指示された物体を運搬するタスク

HUMAN NAVIGATION

- ユーザが行おうとしている行動を自然言語によってうまくサポートしユーザーが迷わないようにガイドするタスク

HANDYMAN

- ユーザから一部誤った指示を貰い、ロボット自身がその指示の間違い部分を判断してフィールド内の目的の物場所を探し出すタスク
- 環境地図を与えて、emcl(自己位置推定用パッケージ)とmove base(移動のためのパッケージ)を用いて目的の部屋まで自律移動する。



大会の結果とまとめ

- 初出場にもかかわらず得点でき、5位タイの結果となった(全6チーム)
- Handymanタスクにて得点できた
 - 他2つのタスクは棄権し、0点だった
- 本番ではレイアウト名が予想と異なっていたためプログラムが対応できなかった(ロボットが地図情報を手に入れられず自己位置推定が機能しなかった)
- 結果 適切な走行ができなかった。

プロジェクトのまとめ

- 想定以上にプログラム作成に苦戦してしまった
- 強化学習の基礎とロボットのプログラムの基礎を学べた
- 今後は、ロボットに強化学習を実装したい
- 大会に参加でき、得点もできた
- このPRで得た経験と知識を今後の研究に活かしていきたい