シミュレータを用いたSCCF性能調査　実験設定

2020/07/08 1AS13 廣瀬翔

1. **実験概要**

2019年度実施したSCCFの特性調査実験では、カテゴリ数が状態行動数の削減のために6個であることや、2回前までの発話内容を格納するなど、状態行動数が少ない状態での学習ができることを確認した。

本実験では、状態行動数を増やし、深層強化学習アルゴリズムをSCCFに適用することで学習が可能であるかについて、性能を調査する。

ただし、様々な設定で実験を実施するため、IDAは実際の利用者ではなく、エージェントを評価するシミュレータを対象として（実験環境を簡単にして）実験を実施する。

○対象アルゴリズム

- REINFORCE

- Deep-Q-Network (DQN)

- Double DQN

- Dueling Network

- PPO(Proximal Policy Optimization Algorithms)など

- その他、時間があれば以下の環境で色々試したいです。

1. エージェントのパラメータ設定

状態: 過去T回までの発話したカテゴリ S

行動: 次に発話するカテゴリ A

エージェントが発話するカテゴリの総数: C

エージェントが持つ、過去の発話履歴の長さ: T (Tステップ前までの発話を保存)

エージェントのもつ状態行動の総数

$$

状態 \cdot 行動 = C^T \cdot C = C^{T+1}

$$

1. 実験設定
2. シミュレータの設定