

q-learning-replicator-dynamics

コーディング規約

- 変数/関数名などの大文字・小文字とprefixは以下の規則に従う.

名称	規則
ローカル変数名	snake_case
グローバル変数名	g_snake_case
クラス	CamelCase
メンバ変数名	m_snake_case
関数名	snake_case()
定数	SNAKE_CASE

ディレクトリ構成

```
root/
├── src/
│   ├── env_game/ # 環境(ゲーム)強化学習とゲームを分離・統合
│   ├── agent/ # エージェント
│   ├── logger/ # ファイル書き込み
│   ├── expriments/ # 強化学習のトレイン, レプリケータによる戦略の分析
│   ├── visualization/ # グラフなどの可視化
│   └── evolutionary_game/ # レプリケータや利得計算
└── outcomes/
    └── yyyymmddThhmmss_hogehoge/
```

プログラム概要

root/main.m

概要

プログラムのメイン関数

機能

- 引数でパラメータを受け取る
- 引数を元に, 環境 (ゲーム) , エージェントを作成
- 強化学習の学習軌跡を保存(入力: 環境, エージェント, 引数) 3.1. 事前に与えた戦略数, 以下の処理を繰り返す. 3.1.1 初期方策(戦略)の設定 3.1.2 指定のエピソード数ゲームを行い, エピソードごとに方策を保存 3.1.3

戦略ごとに方策の推移をファイルに保存

4. ダイナミクスを計算(入力: ダイナミクスのラムダ式, ゲーム, 戦略) 4.1 戦略分布を初期化する(N次元マトリクスの一様分布を作成する) 4.2 各戦略間の利得を求める 4.3 (分布ごとに)各戦略の平均利得を求める 4.4 (分布ごとに)ポピュレーションの平均利得を求める 4.5 式に従い,ダイナミクスを求める 4.6 各分布ごとにダイナミクスを保存する

5. それぞれを描画 5.1 保存されたログファイルからグラフを作成

env_game/matrix_game.m

概要

利得表を持ち, 各戦略間の利得を計算する

機能

1. 各戦略間の利得計算

env_game/base_env.m

概要

強化学習の基底クラス, 抽象メソッドを定義

機能

1. リセット環境のリセット(状態を返す)
2. 行動を元に状態遷移(次状態, 報酬, 終了判定)

env_game/matrix_game_env.m

概要

matrix_gameとbase_envを継承 ワンショットの強化学習環境とマトリクスゲームの両側面をもつ

機能

1. 各戦略間の利得計算 戦略を複数受け取り, それぞれの期待利得を返す
2. リセット環境のリセット(状態を返す) 特殊な処理は不要 状態を返す
3. 行動を元に状態遷移(次状態, 報酬, 終了判定) 各エージェントの行動に対応する利得を報酬とする ワンショットの環境であるため, 常に同じ状態・終了判定を返す

agent/base_agent.m

概要

行動主体の基底クラス 状態を入力して行動を出力する 得た報酬を元に方策を更新

機能

1. 行動決定
2. 方策更新

agent/q_learning_agent.m

概要

base_agentを継承 Q学習エージェント

機能

1. 行動決定 方策から行動を決定
2. 方策更新(入力: 状態, 行動, 次状態, 報酬) Q値を更新 Q値を元に方策を更新(行動選択に従う)
3. Q値を更新 Q値の更新式に従う

agent/action_select.m

概要

行動選択関数をまとめたもの

logger/

省略, fprintfとcsvへの書き込みを行う

experiments/train_agents.m

概要

強化学習の学習を進める

機能

1. 入力(エージェント, 環境, エピソード数, 評価エピソード数, logger)に従い, 強化学習を行う
2. エージェントは与えられたエピソード数, 学習を行う
3. 評価エピソード数ごとに評価を行う
4. loggerを用いて, ログ保存を行う

experiments/run_game.m

進化ダイナミクスに基づきゲームを行う

機能

1. 入力(戦略人口, ゲーム, ダイナミクス, エポック数)に従い, ダイナミクス分析を行う
2. エポック数にダイナミクスに従い, 人口を変異させる

visualization/

省略

evolutionary_game/dynamics.m

概要

各ダイナミクスを保存

機能

戦略人口，戦略間の利得を入力し，戦略人口の変化を出力する

evolutionary_game/utils.m

概要

各戦略の平均利得を求めるなどの一般的な処理を記述