

【公報種別】特許法第17条の2の規定による補正の掲載
 【部門区分】第6部門第3区分
 【発行日】令和7年2月25日(2025.2.25)

【公開番号】特開2023-164155(P2023-164155A)
 【公開日】令和5年11月10日(2023.11.10)
 【年通号数】公開公報(特許)2023-212
 【出願番号】特願2022-75509(P2022-75509)
 【国際特許分類】
 G 0 6 N 2 0 / 0 0 (2 0 1 9 . 0 1)
 【 F I 】
 G 0 6 N 2 0 / 0 0

10

【手続補正書】
 【提出日】令和7年2月14日(2025.2.14)
 【手続補正1】
 【補正対象書類名】明細書
 【補正対象項目名】0076
 【補正方法】変更
 【補正の内容】

20

【0076】

ステップs903：機械学習モデル処理部1050が、評価結果1017を用いて誤差関数を計算し、モデルを更新する。例えば、一般的なQ学習の枠組みでは、遷移前時間ステップを t 、遷移後を $t+1$ 、報酬を R_{t+1} 、学習率を α 、遷移前状態を s_t 、遷移後状態を s_{t+1} 、計画を a_t 、次の時間ステップの計画を a_{t+1} 、Q値を Q とすると、誤差関数は以下のように表される。

【手続補正2】
 【補正対象書類名】明細書
 【補正対象項目名】0078
 【補正方法】変更
 【補正の内容】
 【0078】

30

ここで、 Q_{EX} はステップs902で計算したQ値であり、 Q_{EX_target} は、次の時間ステップにおける状態 s_{t+1} と、それを入力として計画生成エージェント1011が出力する計画 a_{t+1} について評価されたQ値である。なお、一般的なQ学習では学習安定化を目的として、 Q_{EX_target} の評価は、ステップs902で用いたモデルが更新される1つ前の期待値評価モデル1020であるターゲットネットワークと呼ばれるものを用いる。学習率は、環境データ1012に含めて指定する機械学習用のパラメータである。

40

50