

【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2年12月17日(2020.12.17)

【公開番号】特開2019-219741(P2019-219741A)

【公開日】令和1年12月26日(2019.12.26)

【年通号数】公開・登録公報2019-052

【出願番号】特願2018-114702(P2018-114702)

【国際特許分類】

G 0 6 N 20/00 (2019.01)

【F I】

G 0 6 N 99/00 1 5 0

【手続補正書】

【提出日】令和2年11月5日(2020.11.5)

【手続補正 1】

【補正対象書類名】特許請求の範囲

【補正対象項目名】全文

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項 1】

対象を制御するための処理の制御内容を決定するためのポリシを学習する計算機システムにおける学習制御方法であって、

前記計算機システムは、

任意のポリシに基づいて前記処理の制御内容を選択するシミュレーションを実現する目標モデルパラメータの一部を変更した遷移モデルパラメータを算出する学習制御部と、

前記学習制御部から入力された前記遷移モデルパラメータ又は前回のシミュレーションの結果に基づいて算出された前記遷移モデルパラメータに基づく前記シミュレーションを複数回実行し、前記シミュレーションの結果に基づいて前記ポリシを更新する学習処理を実行する学習器と、

前記遷移モデルパラメータ及び前記シミュレーションの実行によって更新された前記ポリシに関連する情報を学習履歴として管理する履歴データベースと、を備え、

前記学習制御方法は、

前記学習器が、任意のタイミングで、前記履歴データベースに前記学習履歴を格納する第 1 のステップと、

前記学習器が、任意の回数だけ実行された前記シミュレーションによって更新された前記ポリシの評価値に基づいて、前記学習履歴を利用した前記シミュレーションを実行する必要があるか否かを判定する第 2 のステップと、

前記学習履歴を利用した前記シミュレーションを実行する必要があると判定された場合、前記学習器が、前記履歴データベースから選択された使用学習履歴に基づいて算出された前記遷移モデルパラメータに基づく前記シミュレーションを複数回実行し、前記シミュレーションの結果に基づいて前記ポリシを更新する第 3 のステップと、を含むことを特徴とする学習制御方法。

【請求項 2】

請求項 1 に記載の学習制御方法であって、

前記学習器が、前記評価値に基づいて、前記遷移モデルパラメータを更新するか否かを判定する第 4 のステップと、

前記学習器が、前記遷移モデルパラメータを更新すると判定された場合、前記学習制御部に前記遷移モデルパラメータの更新指示を出力する第 5 のステップと、

前記学習制御部が、前記遷移モデルパラメータの更新指示を受信した場合、現在の前記遷移モデルパラメータを更新し、前記更新された遷移モデルパラメータを前記学習器に出力する第6のステップと、を含むことを特徴とする学習制御方法。

【請求項3】

請求項2に記載の学習制御方法であって、

前記学習制御部は、前記学習履歴の関係を示す履歴関係管理情報を管理し、

前記第1のステップは、

前記学習器が、前記履歴データベースに新規学習履歴を格納する場合、前記新規学習履歴の格納通知を前記学習制御部に通知するステップと、

前記学習制御部が、前記新規学習履歴の生成元となる前記シミュレーションにおいて使用された前記遷移モデルパラメータを算出するために用いられる前記学習履歴と関連付くように前記履歴関係管理情報を更新するステップと、を含むことを特徴とする学習制御方法。

【請求項4】

請求項3に記載の学習制御方法であって、

前記第3のステップは、

前記学習制御部が、前記履歴関係管理情報に基づいて、前記使用学習履歴を選択するステップと、

前記学習制御部が、前記使用学習履歴に基づいて前記遷移モデルパラメータを算出し、前記学習器に出力するステップと、を含むことを特徴とする学習制御方法。

【請求項5】

請求項4に記載の学習制御方法であって、

前記第3のステップは、

前記学習制御部が、前記使用学習履歴に基づいて、前回の前記シミュレーションで使用された前記遷移モデルパラメータの一部を変更することによって新規遷移モデルパラメータを算出し、前記学習器に出力するステップと、

前記学習器が、前記新規遷移モデルパラメータに基づく前記シミュレーションを複数回実行するステップを含むことを特徴とする学習制御方法。

【請求項6】

対象を制御するための処理の制御内容を決定するためのポリシーを学習する計算機システムであって、

任意のポリシーに基づいて前記処理の制御内容を選択するシミュレーションを実現する目標モデルパラメータの一部を変更した遷移モデルパラメータを算出する学習制御部と、

前記学習制御部から入力された前記遷移モデルパラメータ又は前回のシミュレーションの結果に基づいて算出された前記遷移モデルパラメータに基づく前記シミュレーションを複数回実行し、前記シミュレーションの結果に基づいて前記ポリシーを更新する学習処理を実行する学習器と、

前記遷移モデルパラメータ及び前記シミュレーションの実行によって更新された前記ポリシーに関連する情報を学習履歴として管理する履歴データベースと、を備え、

前記学習器は、

任意のタイミングで、前記履歴データベースに前記学習履歴を格納し、

任意の回数だけ実行された前記シミュレーションによって更新された前記ポリシーの評価値に基づいて、前記学習履歴を利用した前記シミュレーションを実行する必要があるかを判定し、

前記学習履歴を利用した前記シミュレーションを実行する必要があると判定された場合、前記学習器が、前記履歴データベースから選択された使用学習履歴に基づいて算出された前記遷移モデルパラメータに基づく前記シミュレーションを複数回実行し、前記シミュレーションの結果に基づいて前記ポリシーを更新することを特徴とする計算機システム。

【請求項7】

請求項6に記載の計算機システムであって、

前記学習器は、

前記評価値に基づいて、前記遷移モデルパラメータを更新するか否かを判定し、

前記遷移モデルパラメータを更新すると判定された場合、前記学習制御部に前記遷移モデルパラメータの更新指示を出力し、

前記学習制御部は、前記遷移モデルパラメータの更新指示を受信した場合、現在の前記遷移モデルパラメータを更新し、前記更新された遷移モデルパラメータを前記学習器に出力することを特徴とする計算機システム。

【請求項 8】

請求項 7 に記載の計算機システムであって、

前記学習制御部は、前記学習履歴の関係を示す履歴関係管理情報を管理し、

前記学習器は、前記履歴データベースに新規学習履歴を格納する場合、前記新規学習履歴の格納通知を前記学習制御部に通知し、

前記学習制御部は、前記新規学習履歴の格納通知を受信した場合、前記新規学習履歴の生成元となる前記シミュレーションにおいて使用された前記遷移モデルパラメータを算出するために用いられる前記学習履歴と関連付くように前記履歴関係管理情報を更新することを特徴とする計算機システム。

【請求項 9】

請求項 8 に記載の計算機システムであって、

前記学習制御部は、

前記学習器によって前記学習履歴を利用した前記シミュレーションを実行する必要があると判定された場合、前記履歴関係管理情報に基づいて、前記使用学習履歴を選択し、

前記使用学習履歴に基づいて前記遷移モデルパラメータを算出し、前記学習器に出力することを特徴とする計算機システム。

【請求項 10】

請求項 9 に記載の計算機システムであって、

前記学習制御部は、前記使用学習履歴に基づいて、前回の前記シミュレーションで使用された前記遷移モデルパラメータの一部を変更することによって新規遷移モデルパラメータを算出し、前記学習器に出力し、

前記学習器は、前記新規遷移モデルパラメータに基づく前記シミュレーションを複数回実行することを特徴とする計算機システム。