

AI 特許紹介(53)

AI 特許を学ぶ！究める！

～深層強化学習によるトカマクプラズマ磁場制御特許～

2023年6月9日

河野特許事務所

所長弁理士 河野英仁

「AI 特許紹介」シリーズは、注目すべき AI 特許のポイントを紹介します。熾烈な競争となっている第4次産業革命下では AI 技術がキーとなり、この AI 技術・ソリューションを特許として適切に権利化しておくことが重要であることは言うまでもありません。

AI 技術は Google, Microsoft, Amazon を始めとした IT プラットフォーマ、研究機関及び大学から毎週のように新たな手法が提案されており、また AI 技術を活用した新たなソリューションも次々とリリースされています。

本稿では米国先進 IT 企業を中心に、これらの企業から出願された AI 特許に記載された AI テクノロジー・ソリューションのポイントをわかりやすく解説致します。

1.概要

特許出願人 Deepmind

出願日 2020年4月28日

公開日 2022年11月1日

公開番号 WO2023281048

発明の名称 ニューラルネットワークを用いた磁気閉じ込め装置の磁場制御

048 特許は、深層強化学習を用いることにより、トカマクプラズマの磁場制御を行う技術に関し、特に深層強化学習により学習されたプラズマ閉じ込めニューラルネットワークを使用して、磁気閉じ込めデバイスの磁場を制御するための制御信号を選択するための制御ポリシーを実装するものである。

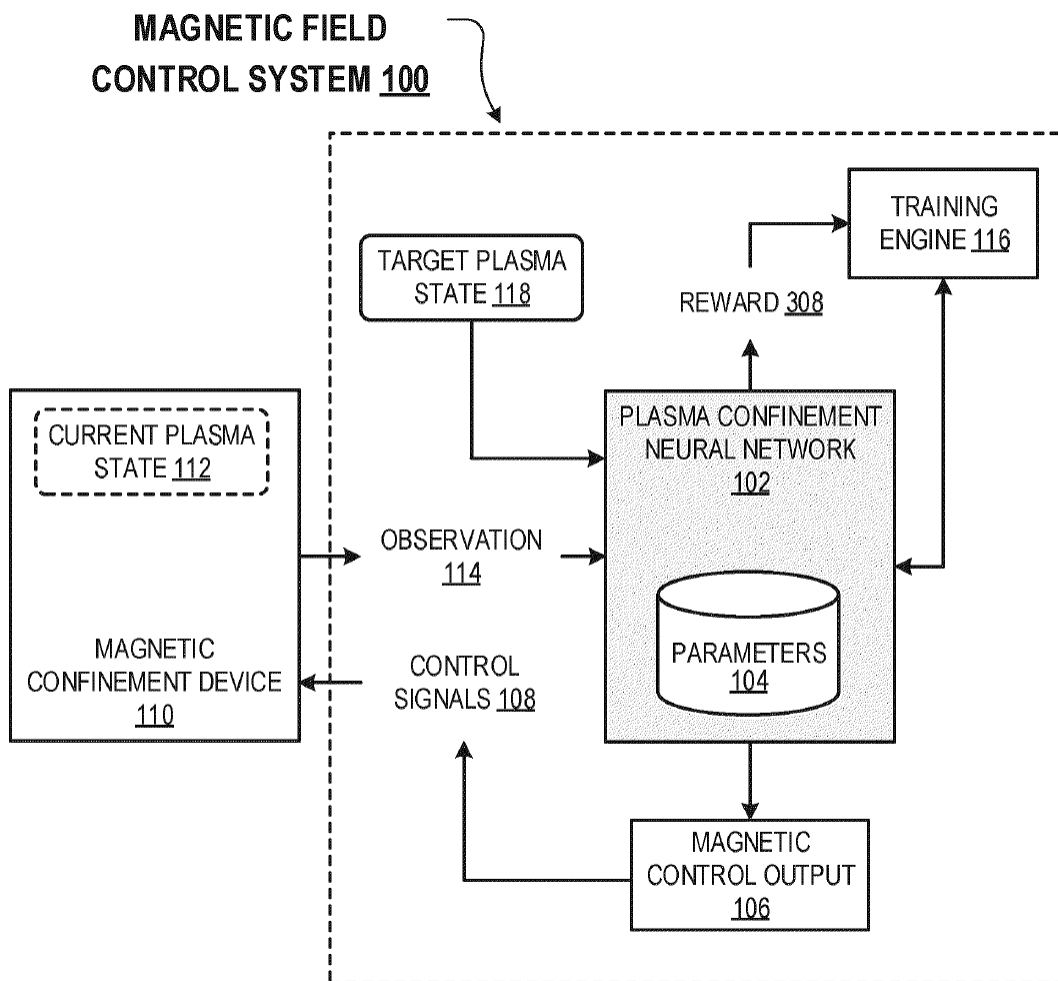
2.特許内容の説明

核融合炉は、高温プラズマで発生する核融合反応から生成される熱を使用して、放射性廃棄物をほとんど使用せずに電力を生成できる。しかしながら、制御された核融合を達成する上で最も困難な問題の1つは、チャンバ内に高温、高圧のプラズマを適切に閉

じ込めることである。極端な温度（摂氏数千万度から数億度など）のため、プラズマはチャンバのどの表面とも直接接触することができず、チャンバ内を真空に保つ必要がある。これは、プラズマ固有の不安定性によってさらに複雑になる。

048 特許のシステムは、プラズマ閉じ込めニューラルネットワークを使用して、磁気閉じ込めデバイスの磁場を制御するための制御信号を選択するための制御ポリシーを実装する。

図1は、磁場制御システム100を示す。



プラズマ閉じ込めニューラルネットワーク102は、ニューラルネットワーク102がどのようにデータを処理するかを指示するネットワークパラメータ104のセットを含む。プラズマ閉じ込めは、初期のプラズマ形成段階、その後のプラズマ平衡への安定化、最終的なプラズマ分解段階など、複数の過渡期を伴う可能性のある高度な時間的手順で

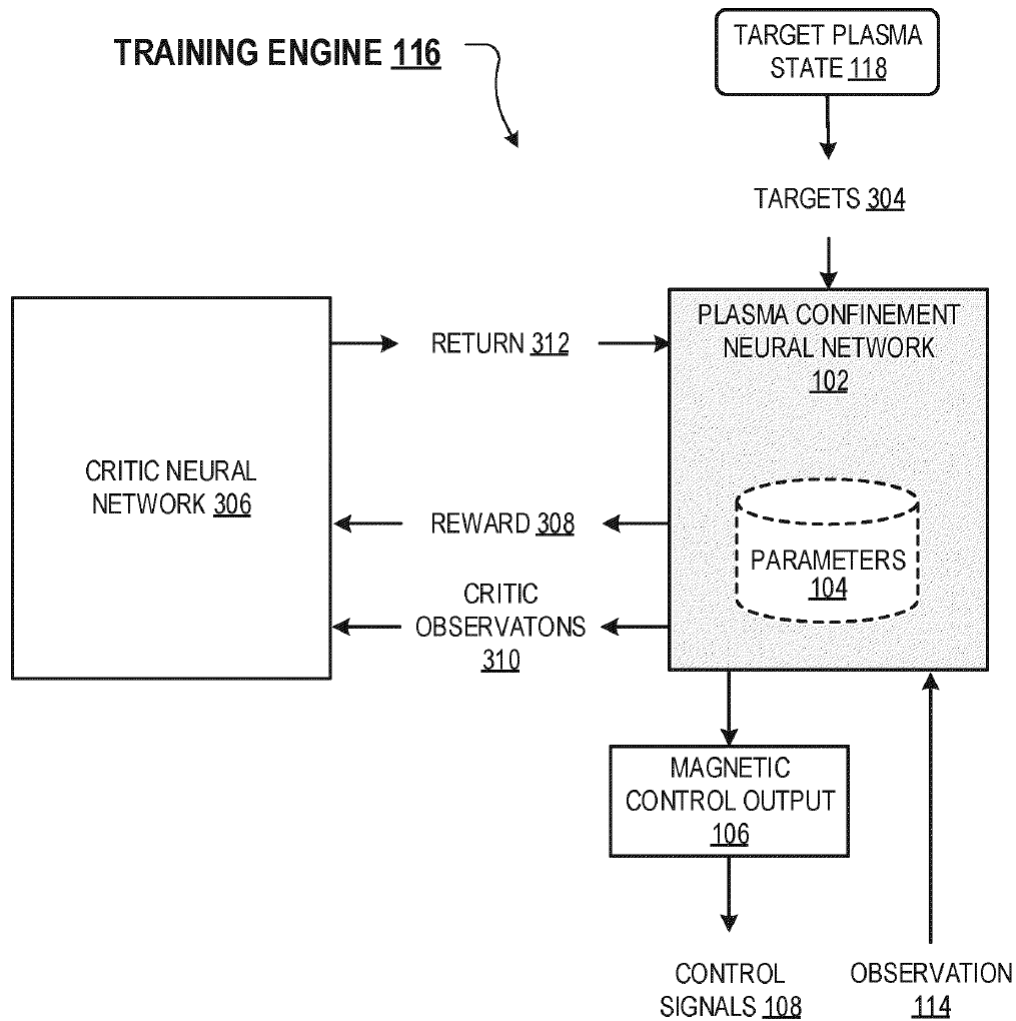
ある。プラズマ固有の不安定性のため、ニューラルネットワーク 102 は、これらの不安定性を修正するために短い時間スケールで応答する必要がある。

各時間ステップで、制御システム 100 は制御ループを実行する。ニューラルネットワーク 102 は、磁気閉じ込め装置 110 のチャンバ内のプラズマ 112 の現在の状態を特徴付ける観察 114 を受信する。報酬 308 は、現在のプラズマ状態 112 に基づいて時間ステップについて決定する。一般に、制御システム 100 は、時間ステップ間で変化し得るプラズマ 118 の目標状態に対して現在のプラズマ状態 112 を評価することによって報酬 308 を決定する。

次に、観察結果 114 は、ネットワークパラメータ 104 に従ってニューラルネットワーク 102 によって処理され、磁気制御出力 106 を生成する。磁気制御出力 106 は、磁気閉じ込めデバイス 110 の磁場を調整するための制御信号 108 を特徴付ける。その結果、現在のプラズマ状態 112 の進展に直接影響する時間ステップでの観察 114 に応答して、制御信号 108 によって磁場を制御することができる。

次に、制御システム 100 は、次の時間ステップのために制御ループを繰り返す。時間ステップに対する報酬は、トレーニングエンジン 116 によって利用され、たとえば強化学習技術を使用して、ニューラルネットワーク 102 のネットワークパラメータ 104 をトレーニングする。

下記図は、プラズマ閉じ込めニューラルネットワーク 102 およびクリティックニューラルネットワーク 306 を共同でトレーニングするために Actor-Critic(アクタークリティック)強化学習技術を使用するトレーニングエンジン 116 の例を示す。



トレーニングエンジン 116 は、プラズマ閉じ込めニューラルネットワーク 102 をトレーニングして、「リターン」312 を増加させる制御信号 108 を生成する。リターン 312 は、プラズマ閉じ込めニューラルネットワーク 102 のクリティック観察 310 を処理することによって、クリティックニューラルネットワーク 306 によって生成される。

クリティック観察 310 は、報酬 308 に基づいて観察 114 に応答して生成される制御信号 108 を特徴付ける。この場合、リターン 312 は、報酬の累積尺度、例えば、時間割引された報酬の合計など、報酬の割引された予想される将来の尺度を指す。

アクタークリティック強化学習技術は、クリティックニューラルネットワーク 306 の出力、つまりリターン 312 を直接的または間接的に使用して、プラズマ閉じ込めニューラルネットワーク 102 をトレーニングする。

クリティック観察 310 は、時間ステップの磁場制御システム 100 の制御ループに関与するすべてのデータ、すなわち、観察 114、ターゲット 304、および制御信号 108 を含む。クリティック 306 は、クリティック観察 310 を、時間ステップについて決定された報酬 308 とともに処理して、リターン 312 を生成する。リターン 312 は、特定の時間ステップにおける制御システム 100 の累積的な将来の報酬を予測する。

軌道を完了した後、訓練エンジン 116 は、各時間ステップでのリターン 312 を実際の累積将来報酬と比較する。訓練エンジン 116 は、クリティックニューラルネットワーク 306 を訓練することができ、すなわち、ネットワークパラメータを更新することによって、累積された将来の報酬を正確に予測するリターン 312 を生成することができる。

逆に、トレーニングエンジン 116 は、プラズマ閉じ込めニューラルネットワーク 102 をトレーニングして、クリティック 306 から生成されるリターン 312 を最大化する制御信号 108 を生成する。

3.クレーム

048 特許のクレーム 1 は以下の通りである。

1.磁気閉じ込め装置のチャンバ内にプラズマを閉じ込めるための磁場を制御するための制御信号を生成するために、1つまたは複数のデータ処理装置によって実行される方法において、

複数の時間ステップのそれぞれで、

磁気閉じ込め装置のチャンバ内のプラズマの現在の状態を特徴付ける観測を取得し、

プラズマ閉じ込めニューラルネットワークを使用して、磁気閉じ込め装置のチャンバ内のプラズマの現在の状態を特徴付ける観測を含む入力を処理し、プラズマ閉じ込めニューラルネットワークは、複数のネットワークパラメータを有し、ネットワークパラメータに従って観測を含む入力を処理して、磁気閉じ込め装置の磁場を制御するための制御信号を特徴付ける磁気制御出力を生成するように構成され、

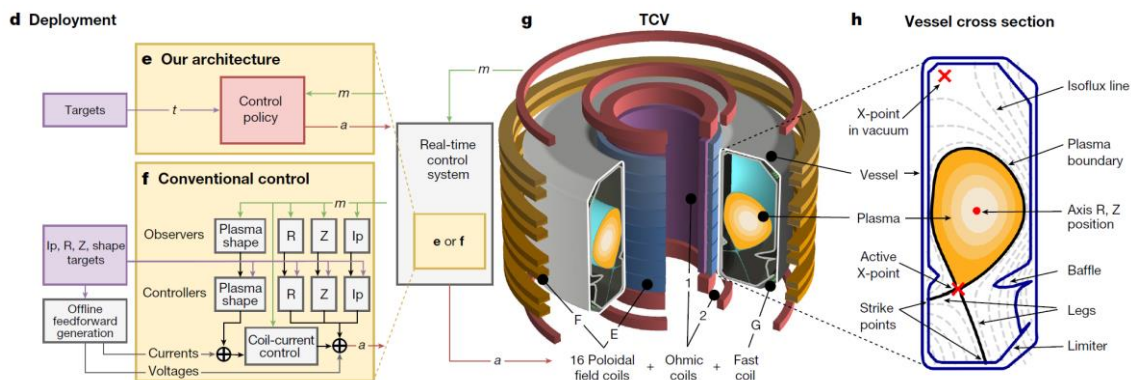
磁気制御出力に基づいて、磁気閉じ込め装置の磁場を制御するための制御信号を生成する。

4. 本特許に関連する論文

本特許に関する論文 “Magnetic control of tokamak plasmas through deep

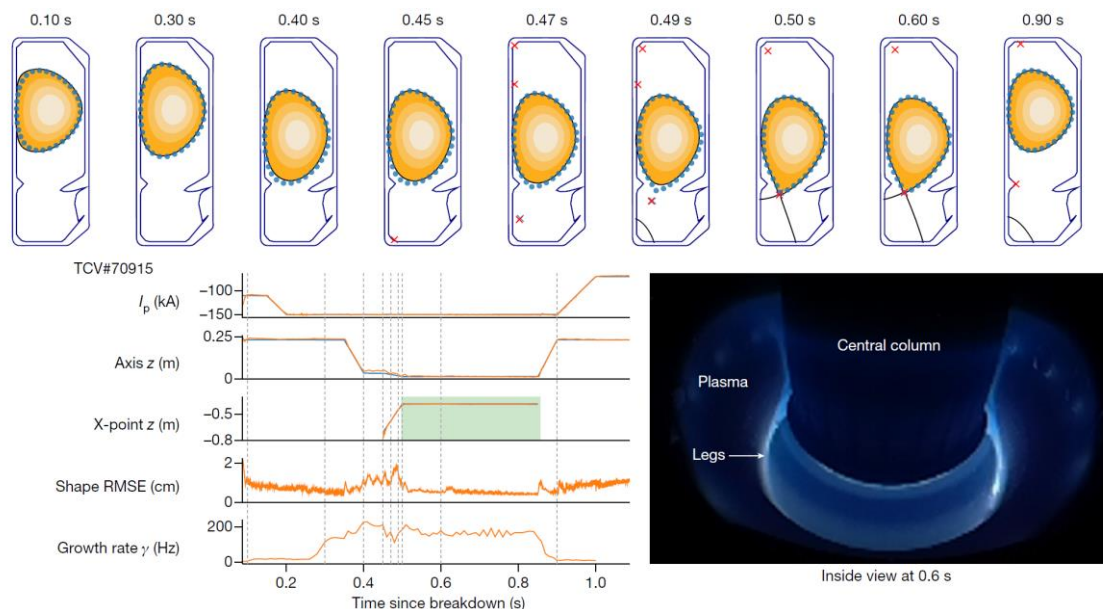
reinforcement learning”¹が、サイエンス誌上で Deepmind の Jonas Degraive 氏らにより公表されている。

下記図は実装例を示す説明図である。



d-h は、TCV(Tokamak à Configuration Variable)とリアルタイムデプロイ制御システムの相互作用を示しており、19 個のコイルすべてを直接制御するために、多くのサブコンポーネントで構成される従来のコントローラー f か、単一のディープニューラルネットワークを使用するアーキテクチャ e のいずれかを使用して実装される。

g は、TCV 及び 19 の作動コイルを示す図である。容器の高さは 1.5 m、短半径は 0.88 m、容器の半幅は 0.26 m である。J は重要な側面がラベル付けされた、容器とプラズマの断面図である。



¹ Jonas Degraive et al. “Magnetic control of tokamak plasmas through deep reinforcement learning” Nature | Vol 602 | 17 February 2022

上記図は、磁場制御システムを使用した複数のプラズマ機能の制御を示すTCV#70915の実験データである。上側の図は、実験後の平衡再構築（実線）と比較して、半径2 cmの標的形狀点（点）を示す。

下側右の図は0.6秒段階における容器内部の画像を示す。脚部で分流されたプラズマが示されている。

下側左の図は、再構成された観察結果と比較された目標時間トレースを示しており、方向転換されたプラズマの窓がマークされている（緑色長方形）。初期制限フェーズ（0.1秒から0.45秒）では、 I_p 二乗平均平方根誤差（RMSE）は0.71 kA（ターゲットの0.59%）であり、形状RMSEは0.78 cm（容器の半幅の3%）である。分流フェーズ（0.55秒から0.8秒）では、 I_p と形状のRMSEはそれぞれ0.28 kAと0.53 cm（0.2%と2.1%）であり、0.62 kAおよび0.75 cm（0.47%および2.9%）のウィンドウ全体（0.1秒から1.0秒）にわたるRMSEが得られる。

このように、本強化学習アーキテクチャが、放電実験の関連するすべての段階で正確なプラズマ制御が可能であることを示している。

以上

著者紹介

河野英仁

河野特許事務所、所長弁理士。立命館大学情報システム学博士前期課程修了、米国フランクリンピアースローセンター知的財産権法修士修了、中国清華大学法学院知的財産夏季セミナー修了、MIT(マサチューセッツ工科大学)コンピュータ科学・AI研究所 AIコース修了。

[AI特許コンサルティング](#)、[医療AI特許コンサルティング](#)の他、米国・中国特許の権利化・侵害訴訟を専門としている。著書に「世界のソフトウェア特許(共著)」、「FinTech特許入門」、「[AI/IoT特許入門3](#)」、「[ブロックチェーン3.0](#)(共著)」がある。