

(19)日本国特許庁(JP)

(12)公開特許公報(A)

(11)公開番号

特開2023-164155

(P2023-164155A)

(43)公開日 令和5年11月10日(2023.11.10)

(51)国際特許分類

G 0 6 N 20/00 (2019.01)

F I

G 0 6 N 20/00

審査請求 未請求 請求項の数 15 O L (全21頁)

(21)出願番号	特願2022-75509(P2022-75509)	(71)出願人	000005108 株式会社日立製作所 東京都千代田区丸の内一丁目6番6号
(22)出願日	令和4年4月28日(2022.4.28)	(74)代理人	110001689 青稜弁理士法人
		(72)発明者	土屋 祐太 東京都千代田区丸の内一丁目6番6号 株式会社日立製作所内
		(72)発明者	森 靖英 東京都千代田区丸の内一丁目6番6号 株式会社日立製作所内
		(72)発明者	恵木 正史 東京都千代田区丸の内一丁目6番6号 株式会社日立製作所内

(54)【発明の名称】 情報処理装置、機械学習方法、及び情報処理方法

(57)【要約】

【課題】

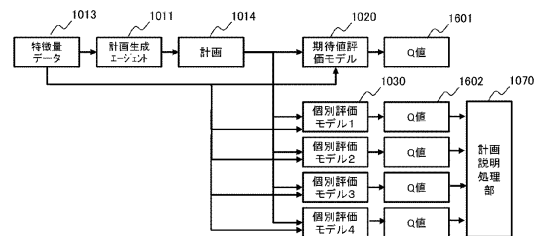
AIがどのような将来のシナリオを想定して出力を行っているのかを、ユーザが容易に判断し得る技術を提供することにある。

【解決手段】

本発明の好ましい一側面は、確率的な状態遷移を伴う環境から観測された状態に基づいて応答を出力するエージェントと、前記確率的な状態遷移の一部が起こるものとして前記応答を評価する個別評価モデルと、前記応答に基づく情報に対応付けて前記評価に基づく情報を出力する計画説明処理部と、を有する情報処理装置である。

【選択図】 図16

図16



**【特許請求の範囲】****【請求項 1】**

確率的な状態遷移を伴う環境から観測された状態に基づいて応答を出力するエージェントと、  
前記確率的な状態遷移の一部が起こるものとして前記応答を評価する個別評価モデルと、  
前記応答に基づく情報に対応付けて前記評価に基づく情報を入力する計画説明処理部と、  
を有する情報処理装置。

**【請求項 2】**

上記請求項 1 において、  
前記エージェントおよび前記個別評価モデルは機械学習モデルであり、  
前記状態は前記環境から得られる特徴量であり、  
前記個別評価モデルは前記特徴量と前記応答を入力として前記応答を評価する、  
情報処理装置。

**【請求項 3】**

上記請求項 2 において、  
前記個別評価モデルはQ値を用いて前記応答を評価する、  
情報処理装置。

**【請求項 4】**

上記請求項 2 において、  
前記個別評価モデルを複数種類有し、  
前記計画説明処理部は、ユーザからの質問を入力とし、前記質問に基づいて複数の個別評価モデルから所定の個別評価モデルを選択する質問処理部を有する、  
情報処理装置。

**【請求項 5】**

上記請求項 2 において、  
前記個別評価モデルを複数種類有し、  
前記計画説明処理部は、前記評価が所定の条件を満たす個別評価モデルに関する説明を出力する説明生成部を有する、  
情報処理装置。

**【請求項 6】**

上記請求項 2 において、  
前記個別評価モデルを複数種類有し、  
前記計画説明処理部は、複数種類の前記個別評価モデルの評価を同時に表示する説明生成部を有する、  
情報処理装置。

**【請求項 7】**

上記請求項 3 において、  
前記計画説明処理部は、前記Q値を他の数値に変換して出力する説明生成部を有する、  
情報処理装置。

**【請求項 8】**

上記請求項 2 において、  
さらに前記応答と同じ形式のデータからなるユーザ計画を処理するユーザ計画処理部を備え、  
前記ユーザ計画処理部は、前記個別評価モデルに前記特徴量と前記ユーザ計画を入力として前記ユーザ計画を評価させ、  
前記計画説明処理部は、さらに前記ユーザ計画の評価に基づく情報を入力する、  
情報処理装置。

**【請求項 9】**

10

20

30

40

50

請求項 2 のエージェントを機械学習する機械学習方法であって、  
前記エージェント、前記個別評価モデル、および確率的な状態遷移の全体を見て期待值的に Q 値を評価する期待値評価モデルの学習を行う際に、  
学習用データを用いて、前記エージェントと前記期待値評価モデルの学習を行い、  
前記学習用データの一部のみを用いて、前記個別評価モデルの学習を行う、  
機械学習方法。

【請求項 10】

請求項 9 の機械学習方法であって、  
前記エージェントと前記期待値評価モデルの学習を、アクタークリティック手法を用いた強化学習で実行する、  
機械学習方法。

10

【請求項 11】

請求項 10 の機械学習方法であって、  
前記個別評価モデルを学習する際に、前記期待値評価モデルの出力を利用する、  
機械学習方法。

【請求項 12】

請求項 11 の機械学習方法であって、  
前記個別評価モデルの出力と前記期待値評価モデルの出力との誤差を計算することで、  
出力の一貫性を担保しながら前記個別評価モデルを学習する、  
機械学習方法。

20

【請求項 13】

確率的な状態遷移を伴う環境に基づく特徴量を入力し応答を出力する第 1 の学習モデルと、前記確率的な状態遷移の一部が確定したものとして前記応答を評価する第 2 の学習モデルと、を備える情報処理装置が実行する情報処理方法であって、  
前記第 1 の学習モデルに、前記特徴量を入力し応答を出力させる第 1 のステップ、  
前記第 2 の学習モデルに、前記特徴量と前記応答を入力して前記応答の評価値を得る第 2 のステップ、  
前記応答に対応付けて前記評価値に基づく情報を入力する第 3 のステップ、  
を実行する情報処理方法。

30

【請求項 14】

請求項 13 の情報処理方法であって、  
前記第 2 の学習モデルを複数準備し、  
ユーザが指定した条件に当てはまる前記第 2 の学習モデルの評価値に基づく情報を入力すること、および、ユーザが指定した条件に当てはまる評価値を入力した前記第 2 の学習モデルに基づく情報を入力すること、の少なくとも一つを実行する、  
情報処理方法。

【請求項 15】

請求項 13 の情報処理方法であって、  
前記第 1 の学習モデルを、アクタークリティック手法を用いた強化学習で学習し、  
前記第 2 の学習モデルを、前記第 1 の学習モデルの学習に用いた学習データのうち、前記確定した状態遷移を持つデータのみを用いて学習する、  
情報処理方法。

40

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、機械学習システムが出力する計画や行動を評価し、説明を提示する技術に関する。

【背景技術】

【0002】

機械学習システムの 1 つである強化学習は、行動により報酬が与えられる環境(タスク)

50

において、適切な報酬につながる行動が出力されるように、機械学習モデル（人工知能：AI）のパラメータを学習していく仕組みである。その高い性能から、社会インフラや医療現場などの業務にまで適用範囲を広げている。例えば、予想される自然災害などによる被害を最小限に抑えるために、あらかじめ人員などのリソースを適切に配置する事前対策計画を立案することができる。しかし、このようなミッションクリティカルな業務において機械学習システムを活用するためには、高い有用性に加え、透明性、公平性、解釈性など様々な性質についての要件を満たすことが求められる。そこで、機械学習システムの判断根拠を説明する技術であるXAI（eXplainable AI）の研究が急速に進展している。

【0003】

強化学習を対象としたXAI技術として、非特許文献1では、AIモデルに入力した画像の中でAIが重要視した部分を、ヒートマップで可視化している。特に教師あり学習の枠組みでは、このような入力データに対する説明技術の開発が盛んに行われている。一方で、強化学習におけるAIの行動は、将来得られる報酬や事象を見越して学習されるものであるため、入力データを用いた「過去向きの説明」ではなく、AIが意図した将来の事象に対する「未来向きの説明」への注目が集まっている。

10

【0004】

例えば、非特許文献2では、説明対象の行動以降に起こるであろう将来の事象（状態遷移）の系列（以下、シナリオ）について、最も起こる確率の高いシナリオを説明に用いる方法を提案している。

【0005】

また、非特許文献3は、将来起こりうる全ての状態遷移と行動に対するテーブルを出力する教師あり学習のAIモデルによって、強化学習AIの行動の意図を可視化する手法を提案している。

20

【0006】

さらに、特許文献1では、行動の良さを表すQ値と呼ばれる値を評価するAIを、目的関数ごとに分ける手法を提案している。これにより、複数の目的を同時に満たす行動を学習しやすくなり、各目的関数の重み調整への示唆も与えられる。

【先行技術文献】

【特許文献】

【0007】

【特許文献1】特開2019-159888号公報

【非特許文献】

【0008】

【非特許文献1】S. Greydanus, A. Koul, J. Dodge, and Alan Fern, "Visualizing and Understanding Atari Agents," Proceedings of the 35th International Conference on Machine Learning, Stockholm, Sweden, PMLR 80, 2018.

【非特許文献2】J. V. D. Waa, J.V. Diggelen, K.V.D. Bosch, and M. Neerinx, "Contrastive Explanations for Reinforcement Learning in terms of Expected Consequences," ArXiv, 1807.08706, 2018.

【非特許文献3】H. Yau, C. Russell, and S. Hadfield, "What Did You Think Would Happen? Explaining Agent Behaviour through Intended Outcomes," Workshop on Extending Explainable AI Beyond Deep Models and Classifiers, Vienna, Austria, PMLR 119, 2020.

30

40

【発明の概要】

【発明が解決しようとする課題】

【0009】

非特許文献2に記載の技術では、AIの意図を解釈するには不十分である。強化学習は様々なシナリオを想定し、期待値的に有効な行動を選んだのであり、例えば、確率が低くてもAIの行動が非常に効果的なシナリオや、依然低い報酬となるリスクシナリオを含んでいる。よって、最も起こる確率の高いシナリオのみでは、AIの意図を説明するために

50

十分な情報を得ることができない。断定的に1つのシナリオを選ぶのではなく、ユーザの関心に合わせてシナリオを選択する機能が求められる。

【0010】

非特許文献3に記載の技術では、AIの意図した状態を網羅的に比較することができる反面、現実では非常に多くの状態遷移や行動が考えられるため、現場での適用は難しい。

【0011】

特許文献1に記載の技術では、XAIは考慮されていないが、仮に複数の目的関数を利用してAIの意図を説明しようとした場合でも、複数の目的関数からAIの重視したものを抽出することはできるが、AIが想定した具体的な将来のシナリオを判断することは難しい。

【0012】

そこで本発明の課題は、AIがどのような将来のシナリオを想定して出力を行っているのかを、ユーザが容易に判断し得る技術を提供することにある。

【課題を解決するための手段】

【0013】

本発明の好ましい一側面は、確率的な状態遷移を伴う環境から観測された状態に基づいて応答を出力するエージェントと、前記確率的な状態遷移の一部が起こるものとして前記応答を評価する個別評価モデルと、前記応答に基づく情報に対応付けて前記評価に基づく情報を出力する計画説明処理部と、を有する情報処理装置である。

【0014】

より具体的には、上記において、前記エージェントおよび前記個別評価モデルは機械学習モデルであり、前記状態は前記環境から得られる特徴量であり、前記個別評価モデルは前記特徴量と前記応答を入力として前記応答を評価する。

【0015】

さらに具体的には、前記エージェント、前記個別評価モデル、および確率的な状態遷移の全体を見て期待値的にQ値を評価する期待値評価モデルの学習を行う際に、学習用データを用いて、前記エージェントと前記期待値評価モデルの学習を行い、前記学習用データの一部のみを用いて、前記個別評価モデルの学習を行う。

【0016】

本発明の好ましい他の一側面は、確率的な状態遷移を伴う環境に基づく特徴量を入力し応答を出力する第1の学習モデルと、前記確率的な状態遷移の一部が確定したものとして前記応答を評価する第2の学習モデルと、を備える情報処理装置が実行する情報処理方法であって、前記第1の学習モデルに、前記特徴量を入力し応答を出力させる第1のステップ、前記第2の学習モデルに、前記特徴量と前記応答を入力して前記応答の評価値を得る第2のステップ、前記応答に対応付けて前記評価値に基づく情報を出力する第3のステップ、を実行する情報処理方法である。

【発明の効果】

【0017】

AIがどのような将来のシナリオを想定して出力を行っているのかを、ユーザが容易に判断し得る技術を提供することができる。上記した以外の課題、構成及び効果は、以下の実施形態の説明により明らかにされる。

【図面の簡単な説明】

【0018】

【図1】機械学習システム評価装置のシステム構成（ハードウェア）の一例を示すブロック図である。

【図2】環境データのデータ構造の一例を示す表図である。

【図3】特徴量データのデータ構造の一例を示す表図である。

【図4】計画のデータ構造の一例を示す表図である。

【図5】環境遷移条件のデータ構造の一例を示す表図である。

【図6】計画生成エージェント、期待値および個別評価モデルの学習を行う作業フローの一例を示す流れ図である。

10

20

30

40

50

【図7】個別評価条件のデータ構造の一例を示す表図である。

【図8】機械学習モデルの学習段階における画面出力の一例を示すイメージ図である。

【図9】機械学習モデルの誤差計算・モデル更新を行う作業フローの一例を示す流れ図である。

【図10】機械学習システムが出力した行動や計画における意図の説明を行う作業フローの一例を示す流れ図である。

【図11】シナリオ選択条件のデータ構造の一例を示す表図である。

【図12】機械学習システム評価結果の画面出力の一例を示すイメージ図である。

【図13】ユーザ計画との比較を用いた機械学習システム評価装置のシステム構成（ハードウェア）の一例を示すブロック図である。

【図14】ユーザ計画との比較を用いた機械学習システムの説明を行う作業フローの一例を示す流れ図である。

【図15】ユーザ計画との比較を用いた機械学習システムの説明を行う画面出力の一例を示すイメージ図である。

【図16】実施例の概略構成を説明するブロック図である。

【発明を実施するための形態】

【0019】

実施の形態について、図面を用いて詳細に説明する。ただし、本発明は以下に示す実施の形態の記載内容に限定して解釈されるものではない。本発明の思想ないし趣旨から逸脱しない範囲で、その具体的構成を変更し得ることは当業者であれば容易に理解される。

【0020】

以下に説明する実施例の構成において、同一部分又は同様な機能を有する部分には同一の符号を異なる図面間で共通して用い、重複する説明は省略することがある。

【0021】

同一あるいは同様な機能を有する要素が複数ある場合には、同一の符号に異なる添字を付して説明する場合がある。ただし、複数の要素を区別する必要がない場合には、添字を省略して説明する場合がある。

【0022】

本明細書等における「第1」、「第2」、「第3」などの表記は、構成要素を識別するために付するものであり、必ずしも、数、順序、もしくはその内容を限定するものではない。また、構成要素の識別のための番号は文脈毎に用いられ、一つの文脈で用いた番号が、他の文脈で必ずしも同一の構成を示すとは限らない。また、ある番号で識別された構成要素が、他の番号で識別された構成要素の機能を兼ねることを妨げるものではない。

【0023】

図面等において示す各構成の位置、大きさ、形状、範囲などは、発明の理解を容易にするため、実際の位置、大きさ、形状、範囲などを表していない場合がある。このため、本発明は、必ずしも、図面等を開示された位置、大きさ、形状、範囲などに限定されない。

【0024】

本明細書で引用した刊行物、特許および特許出願は、そのまま本明細書の説明の一部を構成する。

【0025】

本明細書において単数形で表される構成要素は、特段文脈で明らかに示されない限り、複数形を含むものとする。

【0026】

なお、以下では専ら、予想される自然災害などによる被害を最小限に抑えるために、あらかじめ人員などのリソースを適切に配置する事前対策計画を立案する強化学習システムに関して述べるが、各方法は、ロボットやゲームAIの行動選択、列車や自動車の運転制御や、従業員のシフトスケジュールなど、環境から観測された状態に合わせて行動や計画（計画は予定された行動なので、両者を合わせて単に行動ということがある）を出力する一般的な強化学習の対象となる問題において、広く適用できるものである。

10

20

30

40

50

## 【0027】

実施例の提案する情報処理装置、機械学習方法、及び情報処理方法は、確率等の条件によって状態遷移を伴う環境から観測された状態に合わせて行動や計画を出力するエージェント部分と、分割して評価する状態遷移の条件をユーザが指定する部分と、指定した条件で分割された将来の状態遷移ごとに行動や計画の価値を推定する部分と、ユーザからの質問を処理する部分と、質問処理結果に対応する状態遷移を選択して将来の状態と報酬を計算する部分と、得られた情報を用いて行動や計画の意図の説明を生成する部分を有する。

## 【0028】

かかる構成によれば、AIが出力した行動や計画に対して、非常に多くの状態遷移が存在する問題設定においても、ユーザ指定の条件で分割された状態遷移ごとに価値を評価することで、ユーザの関心に基づいてAIの想定した具体的な将来のシナリオが提示され、AIの行動の意図を解釈するために有用な情報を得ることができる。

10

## 【0029】

以下、図面を参照して、本発明の幾つかの実施例を説明する。但し、それらの実施例は本発明を実現するための一例に過ぎず、本発明の技術的範囲を限定するものではない。本発明の思想ないし趣旨から逸脱しない範囲で、その具体的構成を変更し得ることは当業者であれば容易に理解される。

## 【0030】

以下に説明する発明の構成において、同一又は類似する構成又は機能には同一の符号を付し、重複する説明は省略する。

20

## 【0031】

図面等において示す各構成の位置、大きさ、形状、及び範囲等は、発明の理解を容易にするため、実際の位置、大きさ、形状、及び範囲等を表していない場合がある。したがって、本発明では、図面等に開示された位置、大きさ、形状、及び範囲等に限定されない。以下、本発明の実施の形態について、図面を参照して説明する。

## 【実施例1】

## 【0032】

図1は、本発明の実施例を実現する機械学習システム評価装置の構成例である。機械学習システム評価装置は、記憶装置1001、処理装置1002、入力装置1003、出力装置1004から構成される。

30

## 【0033】

記憶装置1001は、HDD(Hard Disk Drive)及びSSD(Solid State Drive)等、データを永続的に格納する汎用の装置であり、計画情報1010、AIが出力した計画に対して複数の状態遷移における期待値的な良さを評価する機械学習モデルである期待値評価モデル1020、AIが出力した計画をユーザ指定の条件をもとに状態遷移ごとに分割して評価する機械学習モデルである個別評価モデル1030、および計画説明情報1040を有する。なお、記憶装置1001は、その他の装置と同様の端末上ではなくクラウドや外部サーバー上に存在し、ネットワーク経由でデータが参照できる構成であっても問題ない。

## 【0034】

計画情報1010は、環境から観測された状態に合わせて計画を出力する計画生成エージェント1011、環境についての情報を格納した環境データ1012(図2を参照)、エージェントの入力データとなる特徴量データ1013(図3を参照)、エージェントから出力された計画1014(図4を参照)、環境の状態遷移条件を指定する環境遷移条件1015(図5を参照)、各機械学習モデルを学習するための入力データとなるモデル学習用データ1016、評価モデルによる計画の評価結果1017を含む。

40

## 【0035】

計画説明情報1040は、AIの計画を状態遷移ごとに分割評価する条件である個別評価条件1041、AIの計画に対するユーザからの質問データ1042、質問に基づいて指定された状態遷移の条件を格納するシナリオ選択条件1043、質問に対する回答とな

50

る回答データ 1044 を含む。

【0036】

処理装置 1002 は、汎用の計算機であり、その内部に機械学習モデル処理部 1050、環境処理部 1060、計画説明処理部 1070、画面出力部 1080、データ入力部 1090 を、ソフトウェアプログラムとしてメモリに格納されるなどの形態で有する。

【0037】

計画説明処理部 1070 は、個別評価モデル 1030 の処理を行う個別評価処理部 1071、ユーザからの質問データ 1042 やシナリオ選択条件 1043 の処理を行う質問処理部 1072、ユーザへの回答データ 1044 を生成する説明生成部 1073 を含む。

【0038】

画面出力部 1080 は、計画 1014 や回答データ 1044 をディスプレイ出力可能な形式に変換する際に使用される。

【0039】

データ入力部 1090 は、ユーザからのパラメータや質問を設定する際に使用される。

【0040】

入力装置 1003 は計算機用の汎用的な入力装置、例えばマウスやキーボード、タッチパネル等である。

【0041】

出力装置 1004 は、ディスプレイ等の装置であり、画面出力部 1080 を通してユーザとインタラクションするための情報が表示される。ただし、機械学習システムの評価結果を人間が確認する必要が無い場合(例えば、評価結果を別システムに直接渡す場合)は、出力装置は設けなくても良い。

【0042】

以上の構成は、単体の装置で構成してもよいし、あるいは、装置の任意の部分が、ネットワークで接続された他のコンピュータで構成されてもよい。本実施例中、ソフトウェアで構成した機能と同等の機能は、FPGA (Field Programmable Gate Array)、ASIC (Application Specific Integrated Circuit) などのハードウェアでも実現できる。

【0043】

図 2 は、環境データ 1012 の一例を示す図である。環境データ 1012 は、時間変化しないデータ 1012C、時間変化を伴うデータ 1012V、機械学習パラメータ 1012P から構成される。ここでは、エリア 1 ~ 3 において送配電を行うためのリソースを適切に配置する計画を立案する例を考える。

【0044】

時間変化しないデータ 1012C は、各データ項目のカテゴリ情報を表すカテゴリ 21、その値 22 から構成されるデータベースである。一例として、エリアごとの発電所等の施設数や、エリア間の距離が記録される。

【0045】

時間変化を伴うデータ 1012V は、時間断面を示すステップ数 23、データ項目のカテゴリ 24、値 25 から構成されるデータベースである。一例として、時間変動するエリアごとの電力需要や温度が記録される。

【0046】

機械学習パラメータ 1012P は、機械学習時に用いる各パラメータのカテゴリ 26、その値 27 から構成されるデータベースである。

【0047】

環境データ 1012 は、例えば、ユーザにより入力され、又は所定の情報処理装置から取得される情報である。また、データの形式はテーブルデータに限らず、例えば画像情報や計算式でもよい。

【0048】

図 3 は、特徴量データ 1013 の一例を示す図である。特徴量データ 1013 は、各特

10

20

30

40

50

微量のカテゴリ 3 1、その値 3 2 から構成されるデータベースである。特徴量データ 1 0 1 3 は、環境データ 1 0 1 2 をもとにして環境処理部 1 0 6 0 によって生成され、値 3 2 は計画生成エージェント 1 0 1 1 に入力可能なデータ形式（主に数値またはカテゴリデータ）となる。なお、本明細書における「状態」とは、各時間ステップにおける特徴量データ 1 0 1 3 を指す。

【 0 0 4 9 】

図 4 は、計画 1 0 1 4 の一例を示す図である。計画 1 0 1 4 は、対象の時間ステップを表すステップ数 4 1、計画項目のカテゴリ 4 2、その値 4 3 で構成されるデータベースである。計画 1 0 1 4 は、計画生成エージェント 1 0 1 1 によって時間ステップごとに出力される。この例の計画 1 0 1 4 は、あるエリアにおいて送配電を行うためのリソース（人員）を、適切に配置する計画である。

10

【 0 0 5 0 】

図 5 は、環境遷移条件 1 0 1 5 の一例を示す図である。環境遷移条件 1 0 1 5 は、対象の時間ステップを表すステップ数 5 1、遷移条件項目のカテゴリ 5 2、その値 5 3 で構成されるデータベースである。環境遷移条件 1 0 1 5 は、確率や条件式によって定義され、環境処理部 1 0 6 0 によって次の時間ステップの特徴量データ 1 0 1 3 に反映される。なお、本明細書における「事象の発生」は、環境遷移が起こることを指す。この例の環境遷移条件は、ステップ毎のエリア毎の停電確率を示す。

【 0 0 5 1 】

以下、機械学習システム評価装置の動作プロセスを説明する。本実施例は、大きく学習段階と説明段階に分けられる。

20

【 0 0 5 2 】

図 6 は、計画生成エージェント 1 0 1 1、期待値評価モデル 1 0 2 0、個別評価モデル 1 0 3 0 の学習プロセスの一例を説明するフロー図である。本実施例では、時間ステップごとに環境から観測された状態に合わせてエージェントが行動や計画を出力するエピソードを複数回繰り返すことで学習用データを蓄積し（s 6 0 3 ~ s 6 1 0）、逐次誤差関数の計算と機械学習モデルの更新を行い（s 6 1 2）、精度の高いモデルの学習を行う。

【 0 0 5 3 】

フロー図に基づく動作は以下の通りである。

【 0 0 5 4 】

ステップ s 6 0 1：ユーザが個別評価条件 1 0 4 1 を指定する。GUI上でインタラクティブに設定する、ファイルなどの形式で他の情報処理装置からデータを送る、などの方法で行う。クラスタリングなどのアルゴリズムによる自動分類でもよい。詳細は図 7 および図 8 で説明する。

30

【 0 0 5 5 】

ステップ s 6 0 2：個別評価条件 1 0 4 1 に基づいて、個別評価処理部 1 0 7 1 が個別評価モデル 1 0 3 0 を生成する。モデル数は、個別評価条件 1 0 4 1 にて決められた条件によって決定される。個別評価条件 1 0 4 1 とそれによって生成されたモデル数の例については、図 7 と図 8 に述べる。個別評価モデル 1 0 3 0 は、ニューラルネットワークなどの機械学習モデルを想定している。個別評価モデル 1 0 3 0 は、期待値評価モデル 1 0 2 0 と同様の特徴量を入力データとして扱うことができ、出力も期待値評価モデル 1 0 2 0 と同様に、評価結果 1 0 1 7 へ格納される Q 値というスカラー値を基本的な構成としている。Q 値の詳細については、図 9 にて述べる。

40

【 0 0 5 6 】

ステップ s 6 0 3：学習用データの蓄積およびモデル更新を行うエピソードループを開始する。

【 0 0 5 7 】

ステップ s 6 0 4：環境データ 1 0 1 2 のうち、時間変化しないデータ 1 0 1 2 C および時間変化を伴うデータ 1 0 1 2 V から最初の時間ステップのデータを入力として、環境処理部 1 0 6 0 が特徴量データ 1 0 1 3（図 3）を出力する。

50

## 【 0 0 5 8 】

ステップs 6 0 5 : 1つのエピソードにおける各時間ステップの処理を行うループを開始する。エピソードは、複数の時間ステップから構成される。時間ステップの数は、環境データの時間変化を伴うデータ1 0 1 2 Vや機械学習パラメータ1 0 1 2 Pによって指定される。例えば、台風の上陸から過ぎ去るまでがエピソードで、時間ステップは1 3時、1 4時、1 5時...のようなものである。時間ステップ変化の際に、環境がどのように変わるのか(どこで停電が起こるのか)を、環境遷移条件1 0 1 5が決めている。

## 【 0 0 5 9 】

ステップs 6 0 6 : 特徴量データ1 0 1 3を入力として、計画生成エージェント1 0 1 1が計画1 0 1 4を出力する。エージェントは、一般的なニューラルネットワークなどの機械学習モデルである。 10

## 【 0 0 6 0 】

ステップs 6 0 7 : 環境データ1 0 1 2のうち時間変化を伴うデータ1 0 1 2 Vから次の時間ステップのデータ項目、ステップs 6 0 6で出力された計画1 0 1 4、環境遷移条件1 0 1 5を入力として、環境処理部1 0 6 0が次の時間ステップの特徴量データ1 0 1 3を生成する。

## 【 0 0 6 1 】

ステップs 6 0 8 : 現在の時間ステップおよび次の時間ステップの特徴量データ1 0 1 3、計画1 0 1 4を入力として、環境処理部1 0 6 0が報酬を計算する。報酬は、状態遷移前後において、エージェントの出力した計画によって得られる利益や罰則を表す値であり、一般的にスカラー値となる。本実施例では、人員などリソースを配置することのコスト、適切な配置によって削減できた被害量などが当てはまる。ステップs 6 0 6 ~ s 6 0 8の処理は、アクタークリティック(Actor-Critic)として知られる強化学習を適用している。 20

## 【 0 0 6 2 】

ステップs 6 0 9 : 環境処理部1 0 6 0が、現在の時間ステップおよび次の時間ステップの特徴量データ1 0 1 3、計画1 0 1 4、ステップs 6 0 8で生成された報酬値、個別評価条件1 0 4 1に対応したラベル(図7を参照)を1つのタプル等にまとめる。ラベルの条件判定は、ステップs 6 0 7で処理した状態遷移に基づく。作成した学習用データは、環境処理部1 0 6 0によって、モデル学習用データ1 0 1 6として蓄積される。 30

## 【 0 0 6 3 】

ステップs 6 1 0 : 環境データの時間変化を伴うデータ1 0 1 2 Vや機械学習パラメータ1 0 1 2 Pにて指定されたステップ数だけ繰り返す。なお、ステップ数は条件式等で指定される場合もある。環境処理部1 0 6 0にて終了・継続の判定を行う。

## 【 0 0 6 4 】

ステップs 6 1 1 : 機械学習パラメータ1 0 1 2 Pにて指定されたモデル更新頻度の条件に合うかどうかを環境処理部1 0 6 0にて判定する。合う場合はステップs 6 1 2、合わない場合はステップs 6 1 3へ進む。

## 【 0 0 6 5 】

ステップs 6 1 2 : 蓄積されたデータを用いて機械学習モデルを学習・更新する。詳細プロセスは図9にて述べる。 40

## 【 0 0 6 6 】

ステップs 6 1 3 : 機械学習パラメータ1 0 1 2 Pにて指定されたエピソード数だけ繰り返す。環境処理部1 0 6 0にて終了・継続の判定を行う。

## 【 0 0 6 7 】

図7は、個別評価条件1 0 4 1の一例を示す図である。個別評価条件1 0 4 1は、各条件のラベル7 1と、その条件内容である条件7 2から構成されるデータベースである。

## 【 0 0 6 8 】

ラベル7 1は、図6のステップs 6 0 9で学習データに格納され、どの個別評価の条件に対応するかどうかの目印として用いる。 50

## 【0069】

条件72は、環境遷移条件1015に対応し、どの事象が起きたかどうか、その環境遷移によって引き起こされる影響の大きさなどが格納される。また、1つの条件に対して複数の環境遷移が該当してもよい。

## 【0070】

条件72は、例えば、先に設定した環境遷移条件1015に基づいてユーザが指定する。条件72は、時間ステップに依存しない条件(どの時間ステップでも適用できるもの)、もしくは、時間ステップごとの条件を記載することができる。条件72は、具体的なプログラムの変数名や値に紐づいた条件(Aという変数がXという値以上になったら...)、あるいは、環境遷移条件1015と対応した条件(環境遷移条件に記載される「エリア1の停電」が発生したら...)を記載することができる。

10

## 【0071】

図7の例では、時間ステップに依存しない条件が書かれている。よって、環境遷移条件1015によってどこかの時間ステップで「エリア1の停電」が発生したら、個別評価条件1041に従って、ステップs609で学習用データに「ラベル1」が付される。もし、個別評価条件1041が「時間ステップ1でエリア1の停電」「時間ステップ2でエリア1の停電」のように、時間ステップごとに定義されていれば、その該当時間ステップでの環境遷移条件の発生に合わせてラベルが付される。

## 【0072】

図8は、ユーザが個別評価条件1041を入力するインターフェースおよびモデル学習結果のファイル出力の一例を示す図である。個別評価条件1041のファイル入力部801、モデル学習をスタートするボタン802、学習された出力モデルのファイル803から構成される。図8では、図7の個別評価条件1041によって5つのラベルが指定されていたため、個別評価モデルが5つ作成されている。

20

## 【0073】

図9は、図6のステップs612で行われる機械学習モデルの誤差計算・モデル更新プロセスの一例を説明するフロー図である。学習用データからサンプリングされたデータを用いて誤差関数を計算し、各モデルパラメータを更新する。なお、用いられる機械学習モデルはニューラルネットワークを例として示すが、強化学習に用いることができれば細かな指定はない。

30

## 【0074】

ステップs901：機械学習モデル処理部1050が、モデル学習用データ1016から任意のデータをサンプリングする。総数や条件は環境データ1012で指定しても良い。

## 【0075】

ステップs902：サンプリングされたそれぞれの学習データに対して、状態遷移前の特徴量データ1013、計画1014を入力として、期待値評価モデル1020がQ値を出力する。Q値はその状態における計画の良さを表す強化学習における一般的なスカラ値であるが、Q値以外でも計画の良さを表す値であればよい。評価した学習データは、Q値を対応付けて評価結果1017へと格納する。期待値評価モデル1020は公知の手法で、例えば環境処理部によって生成されるものとする。

40

## 【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 $\alpha$ 、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。

## 【0077】

## 【数1】

50

$$\left( R_{t+1} + \gamma \max_{a_{t+1}} Q_{EX\_target}(s_{t+1}, a_{t+1}) - Q_{EX}(s_t, a_t) \right)^2$$

… (1)

【0078】

10

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<sub>t+1</sub>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<sub>t+1</sub>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。

【0079】

ステップs904：機械学習モデル処理部1050が、計画生成エージェント1011の学習を行う。一般的なQ学習では、ステップs902で評価結果1017に格納されたデータのQ値平均を-1倍したものを、誤差関数として学習させる。計画生成エージェントは、より大きなQ値となる計画を立てられるように、学習を進める。

20

【0080】

ステップs905：各個別評価モデル1030についてモデル更新処理を行う。

【0081】

ステップs906：機械学習モデル処理部1050が、ステップs901でサンプリングされたモデル学習用データ1016について、対応する個別評価のラベルを持つデータを抽出する。

【0082】

ステップs907：サンプリングされたそれぞれの学習データに対して、状態遷移前の特徴量データ1013、計画1014を入力として、個別評価モデル1030がQ値を出力する。評価した学習データは、Q値を対応付けて評価結果1017へと格納する。

30

【0083】

ステップs908：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、個別評価モデル1030を更新する。一般的には、各個別評価モデルに対して、ステップs903の処理を行えばよい。

【0084】

ただし、ターゲットネットワークとして更新前の個別評価モデルを用いてしまうと、期待値評価モデル1020と異なる方向で誤差を最小化するために学習してしまう。そこで、ターゲットネットワークとして期待値評価モデル1020を用い、状態遷移ごとの価値を推定する部分と、期待値的な価値との誤差を計算することで、期待値Q値と個別のQ値の一貫性を保った値を用いて、本実施例の目的であるユーザの関心に合わせた粒度でのQ値分解が可能となる。また、各個別評価モデルは独立しているため、並列処理によって学習の高速化が可能である。

40

【0085】

ステップs909：全ての個別評価モデル1030についてモデル更新処理を行ったら終了する。なお、ステップs906にて1つもデータがサンプリングできなかったモデルについては、更新処理は行わなくてもよい。

【0086】

図10は、学習した個別評価モデル1030を活用した機械学習システムの説明プロセスの一例を示す図である。説明段階では、ユーザからの質問を処理し、状態遷移ごとに個

50

別評価したQ値ベクトルから対応する状態遷移をシミュレートし、結果を表示する。このプロセスによって、AIがどのような将来のシナリオを期待して計画を行ったかどうかを解釈することが可能となる。

【0087】

ステップs101：環境処理部1060が、説明の対象となる時間ステップにおける特徴量データ1013を環境データ1012から生成する。対象の時間ステップや条件は、データ入力部1090を用いてユーザが指定するか、他の情報処理装置から指定する。

【0088】

ステップs102：特徴量データ1013を入力として、計画生成エージェント1011が計画1014を出力する。

【0089】

ステップs103：特徴量データ1013と計画1014を入力として、期待値評価モデル1020と個別評価モデル1030がQ値を出力する。なお、用いる個別評価モデルは、環境処理部1060が環境データ1012を参照し、現在の時間ステップで起こりうる状態遷移に対応するもののみを用いる。

【0090】

ステップs104：入力装置1003により、ユーザが質問データ1042を入力する。質問入力には、データ入力部1090を用いてGUI上にてファイルをアップロードする、自然言語にて入力する、などの方法を用いる。

【0091】

ステップs105：質問処理部1072が、ユーザからの質問データ1042とシナリオ選択条件1043（図10を参照）を入力として、ステップs103にて個別評価モデル1030から出力された個別評価のQ値ベクトルから適切な状態遷移を選択する。

【0092】

ステップs106：選択された状態遷移をシミュレーションするために、環境データ1012と計画1014を用いて、環境処理部1060が次の時間ステップの特徴量データ1013を生成する。

【0093】

ステップs107：現在の時間ステップおよび次の時間ステップの特徴量データ1013、計画1014を入力として、環境処理部1060が報酬を計算する。

【0094】

ステップs108：説明生成部1073がユーザへの回答データ1044を生成する。

【0095】

ステップs109：画面出力部1080が、回答データ1044などをGUI形式に変換し、出力装置1004で表示する（図12を参照）。

【0096】

図11は、シナリオ選択条件1043の一例を示す図である。シナリオ選択条件1043は、ユーザからの質問111と、対応する状態遷移112から構成されるデータベースである。質問処理部1072が、ユーザからの質問データ1042を質問111へ対応する形式へ変換することで、シナリオ選択条件1043から適切な状態遷移112を選択することができる。例えば、「最も期待した状態遷移は？」という質問に対して、Q値最大値を示す状態遷移を表示することで、計画1014がもっとも効果の発揮する具体的な事象を知ることができる。なお、シナリオ選択条件1043はテーブルデータの形式に限らず、条件式などでも問題ない。また、最大あるいは最小に限らず、所定の条件を満たすQ値を状態遷移112としてもよい。

【0097】

図12は、説明生成部1073が生成した回答データ1044などの画面出力の一例を示す図である。出力された計画1014の例1201、計画1014をグラフィカルに可視化した例1202、シナリオ選択条件1043のファイル入力部1203、ユーザからの質問データ1042のファイル入力部1204、ファイルアップロードと説明を開始す

10

20

30

40

50

るボタン 1205 を有する。

【0098】

また、ユーザからの質問文の表示例 1206、回答文 1207、複数の個別評価モデル 1030 の Q 値を併記して選ばれた状態遷移をハイライトした Q 値ベクトル 1208、選ばれた状態遷移後の環境と報酬をグラフィカルに可視化した例 1209 から構成される。

【0099】

まず、計画生成エージェント 1011 から出力された計画 1014 をもとに表示される例 1201、例 1202 に対して、ユーザがシナリオ選択条件 1043 および質問データ 1042 を、ファイル入力部 1203 およびファイル入力部 1204 を用いてアップロードする。

10

【0100】

次に、質問データ 1042 をシナリオ選択条件 1043 と照らし合わせながら、質問処理部 1072 が状態遷移 112 を決定し、回答文 1207、Q 値ベクトル 1208、表示例 1209 を回答データ 1044 として示している。

【0101】

この例では、最も期待したシナリオをユーザが聞いているため、最も大きい Q 値を示す状態遷移が選択され、Q 値ベクトル 1208 でもハイライト表示されている。なお、計画情報と回答情報は同時に画面に出さず、2 つの画面の切り替えなどの方法で提示してもよい。

【0102】

なお、本実施例では Q 値をユーザに提示したが、この値は抽象的なため、説明としてふさわしくない場合がある。その際は、図 10 のステップ s103 で Q 値を出力するに加えて、環境処理部 1060 を用いてより Q 値をユーザの解釈しやすい値に変換してもよい。例えばこの実施例では、停電復旧時間や人員などリソース稼働率が該当する。その他にも、状態遷移の確率値を分けて表示、各状態遷移の不確実性を求めて AI の計画に対する確信度を見積もることもできる。不確実性については、公知のアンサンブル学習等を活用した推定方法を用いる。

20

【0103】

また、本実施例では 1 つの時間ステップにおける状態遷移と Q 値を示したが、複数の時間ステップの系列を提示することで、さらに解釈性が高まることもある。その場合は、図 10 の説明プロセスを任意の回数繰り返すか、環境データ 1012 やシナリオ選択条件 1043 によって条件を指定する方法が挙げられる。

30

【0104】

さらに、本実施例では主に各個別評価モデルに対して 1 つの状態遷移が指定された場合を想定しているが、複数の状態遷移を対応させてもよい。説明段階では、シナリオ選択条件 1043 によって、どの状態遷移を用いるかを指定する。1 つだけではなく、複数の状態遷移を表示してもよい。

【0105】

加えて、得られた Q 値ベクトルは説明のみならず、計画生成エージェント 1011 の性能向上を目的とした追加学習のための方針を決めるヒントとしても活用できる。例えば、熟練者から見て重要と思われる将来の事象に対して Q 値が小さい場合、その状態遷移を回答データとして表示することで、ユーザはその事象が発生するエピソードを追加学習するように方針を決めることができる。

40

【実施例 2】

【0106】

図 13 は、第 2 の実施例の機械学習システム評価装置を示すブロック図である。本実施例では、ユーザの想定する計画との比較により、AI が出力した計画 1014 の解釈性を向上させる方法を述べる。

【0107】

第 2 の実施例を実施するための一例として、図 1 を拡張した図 13 に示す装置を用いる

50

。図1の装置図からの追加点として、記憶装置1001の計画説明情報1040におけるユーザ計画1345、処理装置1002の計画説明処理部1070におけるユーザ計画処理部1374が存在する。これらの具体的な活用方法は、後段の説明にて述べる。

【0108】

図14は、ユーザ計画との比較を用いた説明プロセスの一例を示す図である。多くのプロセスは図10と類似しているため、異なる点についてのみ詳細を述べる。

【0109】

ステップs1401～s1403：図10のステップs101～s103と同様である。

【0110】

ステップs1404：ユーザは、質問データ1042に加えて、ユーザの想定するユーザ計画1345を入力する。データの形式は、AIによる計画1014と同様である。

【0111】

ステップs1405：ユーザ計画1345に対して、期待値評価モデル1020と個別評価モデル1030がQ値を出力する。

【0112】

ステップs1406：質問処理部1072が、ユーザからの質問データ1042とシナリオ選択条件1043（図10を参照）を入力として、AIの計画1014とユーザ計画1345の個別評価のQ値ベクトルをそれぞれ比較し、適切な状態遷移を選択する。例えば、「なぜユーザ計画よりもAIの計画の方が良いのか？」という質問の場合は、AIの計画1014においてQ値が大きく、ユーザ計画1345では低いQ値を示す状態遷移を選択することで、各計画の意図する将来の事象において大きな差となる項目を示すことができる。

【0113】

ステップs1407～s1410：図10のステップs106～s109と同様である。ただし、ユーザ計画1345についても、ユーザ計画処理部1374にて、AIによる計画と同様の処理を行い、回答データに処理結果を加える。

【0114】

図15は、ユーザ計画1345との比較を用いた説明の画面出力の一例を示す図である。AIによる計画1014の例1501、計画1014をグラフィカルに可視化した例1502、シナリオ選択条件1043のファイル入力部1503、ユーザからの質問データ1042のファイル入力部1504、ユーザの想定するユーザ計画1345のファイル入力部1505、ファイルアップロードと説明を開始するボタン1506、ユーザからの質問文の表示例1507、回答文1508、選ばれた状態遷移をハイライトしたAI計画のQ値ベクトル1509、選ばれた状態遷移後の環境とAI計画の報酬をグラフィカルに可視化した例1510、選ばれた状態遷移をハイライトしたユーザ計画1345のQ値ベクトル1511、選ばれた状態遷移後の環境とユーザ計画の報酬をグラフィカルに可視化した例1512から構成される。まず、出力された計画1014に対して、ユーザがシナリオ選択条件1043および質問データ1042をアップロードする。次に、シナリオ選択条件1043と照らし合わせながら、質問処理部1072が状態遷移を決定し、回答データ1044として回答文1508から可視化例1512までを示している。この例では、各計画の意図する将来の事象において差が最も大きい状態遷移を選ぶことで、「なぜユーザ計画よりもAIの計画の方が良いのか？」という質問に対して、AIの意図を解釈するための情報を提示している。なお、各項目は別々の画面に表示されていてもよい。

【0115】

図16は図10で説明した実施例の構成を概念的に示すブロック図である。実施例の構成はアクタークリティック(Actor-Critic)として知られる強化学習を適用している。アクタークリティックは、環境から観測された状態を基に行動(Action)を選択して実行するアクター(Actor)と、アクターが選択した行動を評価するクリティック(Critic)で構成される強化学習のフレームワークである。アクターは評価をもとに計画(行動)の最適化を行う。

10

20

30

40

50

## 【0116】

実施例においてアクターには、計画生成エージェント1011が対応する。計画生成エージェント1011は、環境データ1012から作られる特徴量データ1013を入力とし、計画1014を生成する。計画1014と環境データの時間変化を伴うデータ1012Vと環境遷移条件1015から、環境処理部1060が次の時間ステップの特徴量データ1013を生成する(状態遷移が発生する)。

## 【0117】

クリティック1603には、期待値評価モデル1020が対応する。すでに述べた実施例では、期待値評価モデル1020は、特徴量データ1013と計画1014を入力として、その状態における計画(行動)の良さを表すQ値1601を出力する。ここで、期待値評価モデル1020が出力するQ値1601は、全ての状態遷移関数に関する期待値を示す。

10

## 【0118】

実施例では、既述のように1または複数の個別評価モデル1030を備えており、XAIの機能を実現する。個別評価モデル1030は、計画生成エージェント1011が出力した計画1014を任意の条件をもとに状態遷移ごとに分割して評価する機械学習モデルである。言い換えれば、個別評価モデルは、期待値評価モデルが評価の基にする確率的な状態遷移の一部を固定したものを評価するモデルである。

## 【0119】

期待値評価モデル1020が、確率的な状態遷移の全体を見て、期待値的にQ値を評価するのに対して、個別評価モデル1030は、確率的な状態遷移の一部が起こるものとして固定し、その際のQ値を評価する。個別評価モデル1030が出力するQ値1602に基づいて、計画説明処理部1070は計画生成エージェント1011の計画1014についての説明情報を生成する。

20

## 【0120】

個別評価モデル1030はそれぞれ異なるシナリオに基づいた評価を行うため、計画生成エージェント1011が出力した計画1014が、どのようなシナリオに対して意味があるかを、個別評価モデル1030の出力するQ値1602に基づいて知ることができる。

## 【0121】

以上の実施例によると、確率等の条件によって状態遷移を伴う環境から観測された状態に合わせて行動や計画を出力するエージェント部分と、ユーザの関心に基づいて計画の個別評価条件を指定する部分と、将来の状態遷移ごとの価値を推定する個別評価モデル部分と、ユーザからの質問を処理する部分と、処理した結果に対応する状態遷移をもつ個別評価モデルを選択して将来の状態と報酬を計算する部分と、得られた情報を用いて行動や計画の意図の説明を生成する部分を用いることで、強化学習に基づく機械学習システムが出力した行動や計画の意図を解釈するために、AIが想定した具体的な将来のシナリオを、ユーザの関心に合わせて提示することができる。

30

## 【0122】

上記実施例によれば、機械学習モデルの出力をシナリオごとに容易に解釈可能となるため、効率の良い計画の立案が可能となり、消費エネルギーが少なく、炭素排出量を減らし、地球温暖化を防止、持続可能な社会の実現に寄与することができる。

40

## 【符号の説明】

## 【0123】

1001...記憶装置、1002...処理装置、1011...計画生成エージェント、1012...環境データ、1013...特徴量データ、1014...計画、1015...環境遷移条件、1016...モデル学習用データ、1017...評価結果、1020...期待値評価モデル、1030...個別評価モデル、1041...個別評価条件、1042...質問データ、1043...シナリオ選択条件、1044...回答データ、1050...機械学習モデル処理部、1060...環境処理部、1071...個別評価処理部、1072...質問処理部、1073...説明生成

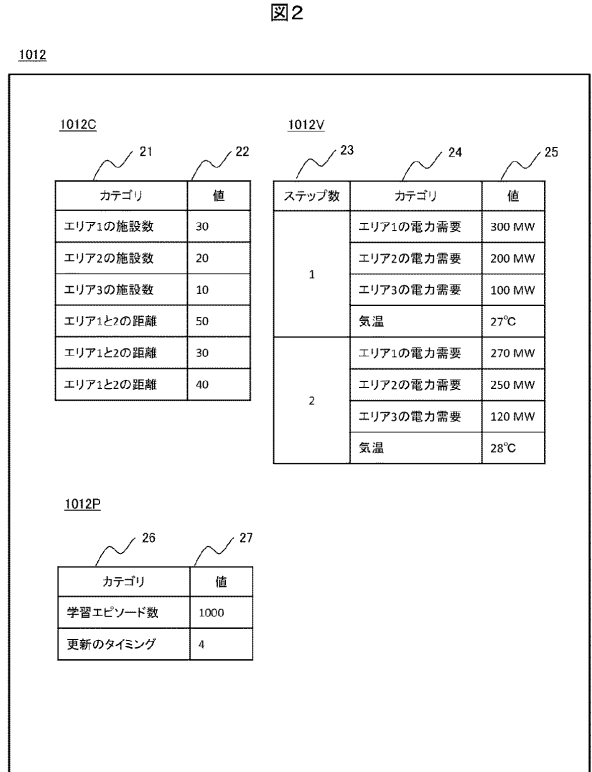
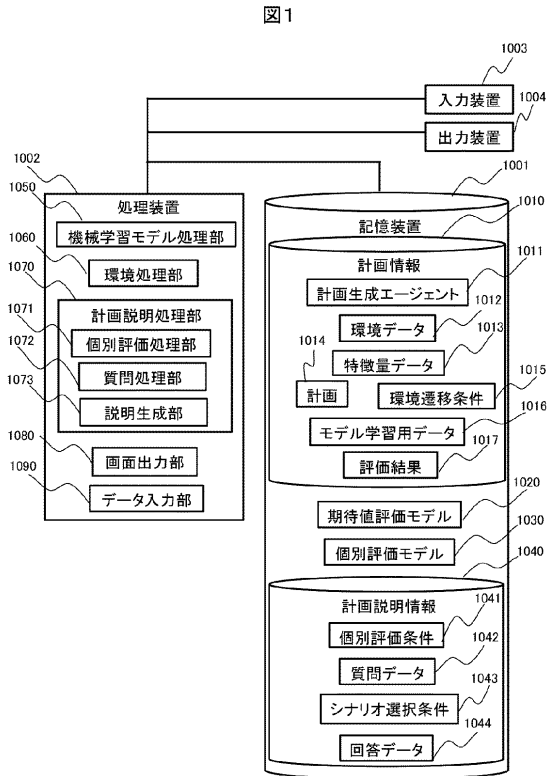
50

部、1080...画面出力部、1090...データ入力部、1208...Q値ベクトル、1345...ユーザ計画、1374...ユーザ計画処理部

【図面】

【図1】

【図2】



10

20

30

40

50

【 図 3 】

図3

1013

31

32

特徴量名	値
エリア1の施設数	30
エリア2の施設数	20
エリア3の施設数	10
エリア1と2の距離	50
エリア1と2の距離	30
エリア1と2の距離	40
エリア1の電力需要	300
エリア2の電力需要	200
エリア3の電力需要	100
気温	27
エリア1の停電	1
エリア2の停電	0
エリア3の停電	0

【 図 4 】

図4

1014

41

42

43

ステップ数	カテゴリ	値
1	エリア1の人員数	15
	エリア2の人員数	20
	エリア3の人員数	10

10

20

【 図 5 】

図5

1015

51

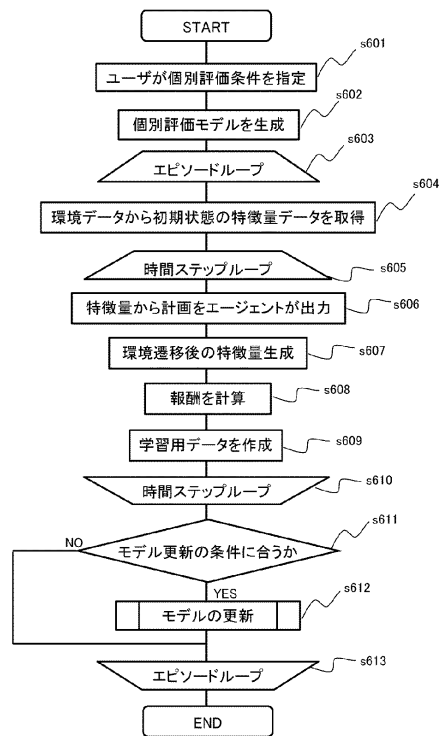
52

53

ステップ数	カテゴリ	値
1	エリア1の停電確率	0.1
	エリア2の停電確率	0.2
	エリア3の停電確率	0.1
2	エリア1の停電確率	0.15
	エリア2の停電確率	0.1
	エリア3の停電確率	0.1

【 図 6 】

図6



30

40

50

【 図 7 】

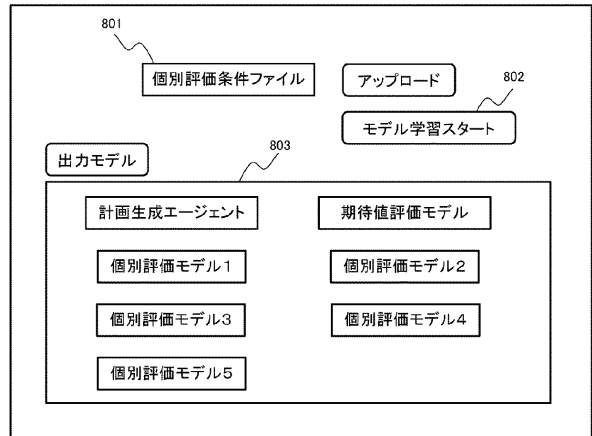
図7

1041

ラベル	条件
1	エリア1の停電
2	エリア2の停電
3	エリア3の停電
4	被害量が300MW以上
5	停電なし

【 図 8 】

図8

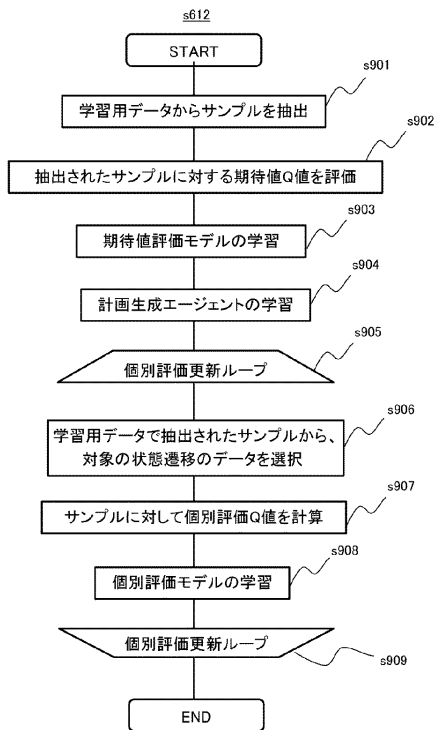


10

20

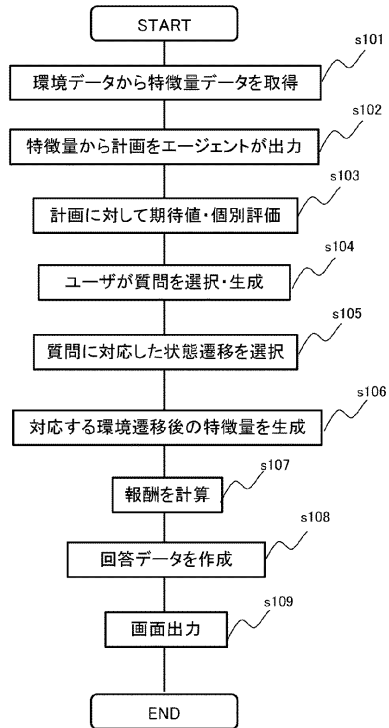
【 図 9 】

図9



【 図 10 】

図10



30

40

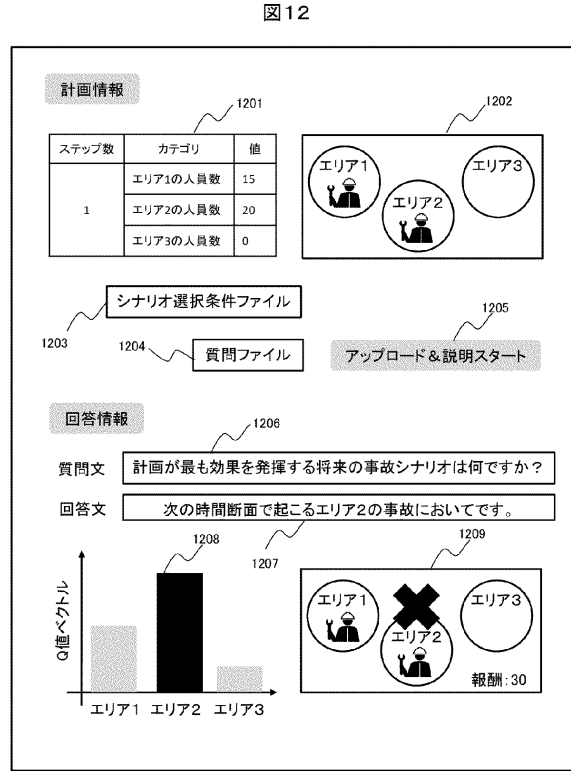
50

【 図 1 1 】

図 11

1043	111	112
質問	状態遷移	
最も期待した状態遷移は？	Q値ベクトル最大値	
リスクとなる状態遷移は？	Q値ベクトル最小値	

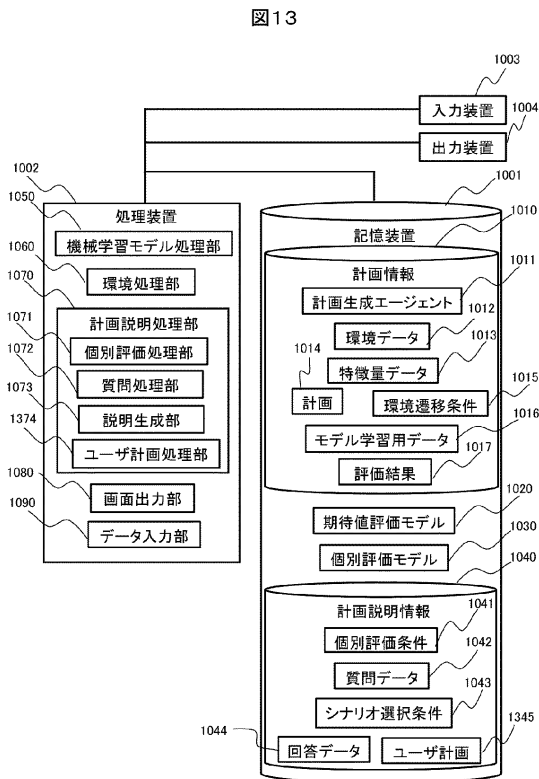
【 図 1 2 】



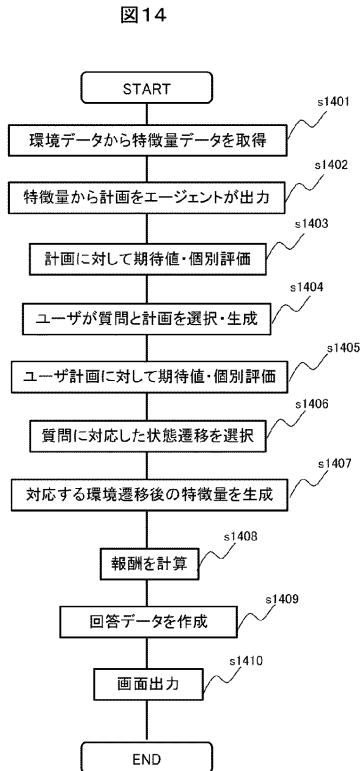
10

20

【 図 1 3 】



【 図 1 4 】



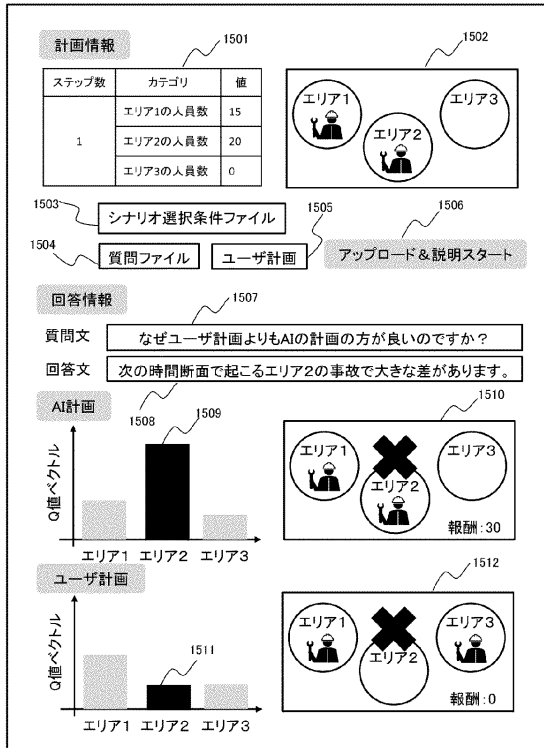
30

40

50

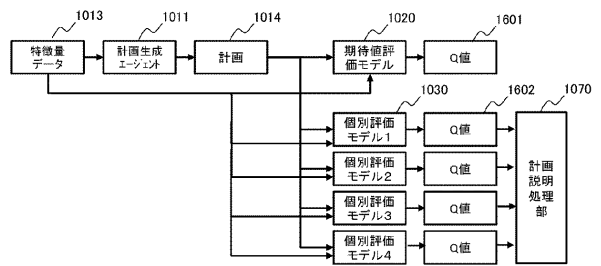
【 図 1 5 】

図15



【 図 1 6 】

図16



10

20

30

40

50