(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2021年11月4日(04.11.2021)

WIPO | PCT

(10) 国際公開番号
WO 2021/220556 A1

(51) 国際特許分類:
G06N 20/00 (2019.01)

(21) 国際出願番号： PCT/JP2021/001234

(22) 国際出願日： 2021年1月15日(15.01.2021)

(25) 国際出願の言語： 日本語

(26) 国際公開の言語： 日本語

(30) 優先権データ：
特願 2020-078883 2020年4月28日(28.04.2020) JP
特願 2020-116497 2020年7月6日(06.07.2020) JP

(71) 出願人:ソニーグループ株式会社(**SONY GROUP CORPORATION**) [JP/JP]; 〒1080075 東京都港区港南１丁目７番１号 Tokyo (JP).

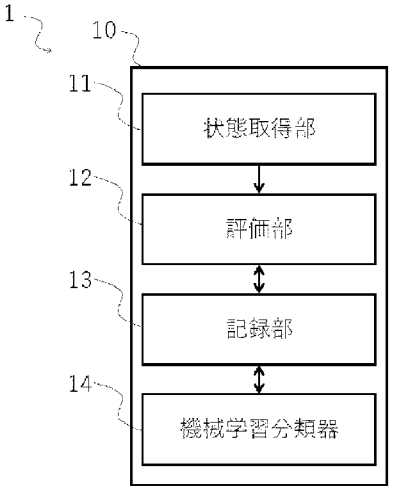
(72) 発明者:滝 裕子(**TAKI Yuko**); 〒1080075 東京都港区港南１丁目７番１号 ソニー株式会社内 Tokyo (JP). 高塚 進(**TAKATSUKA Susumu**); 〒1080075 東京都港区港南１丁目７番１号 ソニー株式会社内 Tokyo (JP). 鉄川 弘樹(**TETSUKAWA Hiroki**); 〒1080075 東京都港区港南１丁目７番１号 ソニー株式会社内 Tokyo (JP).

(74) 代理人: 渡邊 薫(**WATANABE Kaoru**); 〒1080014 東京都港区芝四丁目１０番５号 ヒューリック田町ビル６階 薫風国際特許事務所 Tokyo (JP).

(81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,

(54) **Title:** MACHINE LEARNING SYSTEM, MACHINE LEARNING DEVICE, AND MACHINE LEARNING METHOD

(54) 発明の名称： 機械学習システム、機械学習装置、及び機械学習方法



11... State acquisition unit
12... Evaluation unit
13... Recording unit
14... Machine learning classifier

(57) **Abstract:** Provided are a machine learning system, a machine learning device, and a machine learning method for performing machine learning of a correlation between behavior of a person and the environment surrounding the person so as to promote shift from behavior of the person to a target behavior. The present technology provides a machine learning system provided with at least: a state acquisition unit that at least acquires state information concerning behavior of a person; an evaluation unit that evaluates environment information concerning the environment surrounding the person at the time of acquisition of the state information and the state information so as to obtain a value function; and a machine learning classifier that, in order to encourage shift of behavior to a target behavior, performs reinforcement learning of the value function and selects the environment information when the value function become the highest.

WO 2021/220556 A1

［続葉有］

ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類:

一 国際調査報告(条約第21条(3))

(57) 要約:人の行動とその人の周囲の環境との相関関係を機械学習することにより、その人の行動をターゲット行動に促す機械学習システム、機械学習装置、及び機械学習方法を提供する。本技術は、人の行動に関する状態情報を少なくとも取得する状態取得部と、前記状態情報を取得したときの前記人の周囲の環境に関する環境情報と、前記状態情報と、を評価して価値関数を得る評価部と、前記行動をターゲット行動に促すために、前記価値関数を強化学習して、該価値関数が最も高くなるときの前記環境情報を選定する機械学習分類器と、を少なくとも備えている、機械学習システムを提供する。

明　細　書

発明の名称 ：

機械学習システム、機械学習装置、及び機械学習方法

技術分野

[0001] 本技術は、機械学習システム、機械学習装置、及び機械学習方法に関する。

背景技術

[0002] 従来、人の行動に関する情報をコンピュータが機械学習することにより、前記行動をターゲット行動に促す技術が利用されている。

例えば、特許文献１では、「コンピュータネットワークを利用して消費を誘発して販売を促進させる販売促進情報を消費者に提供する販売促進システム」について開示されている。この特許文献１では、前記販売促進情報が提供された後の消費者の行動に基づいて機械学習を行うことについて説明されている。

先行技術文献

特許文献

[0003] 特許文献1：特開２０１９－０２８８９９号公報

発明の概要

発明が解決しようとする課題

[0004] 従来、人の行動をターゲット行動に促すために、例えば特許文献１において説明されている販売促進情報の提供のように、直接的な表現で情報を人に提供する技術が利用されている。

[0005] しかし、人の行動変容の中には、その人の周囲の環境の変化に応じて、無自覚になされる行動変容がありうる。

[0006] そこで本技術では、人の行動とその人の周囲の環境との相関関係を機械学習することにより、その人の行動をターゲット行動に促す機械学習システム、機械学習装置、及び機械学習方法を提供することを主目的とする。

課題を解決するための手段

[0007] 本技術は、人の行動に関する状態情報を少なくとも取得する状態取得部と、前記状態情報を取得したときの前記人の周囲の環境に関する環境情報と、前記状態情報と、を評価して価値関数を得る評価部と、前記行動をターゲット行動に促すために、前記価値関数を強化学習して、該価値関数が最も高くなるときの前記環境情報を選定する機械学習分類器と、を少なくとも備えている、機械学習システムを提供する。

前記評価部が、前記ターゲット行動に関するターゲット状態情報と、前記状態情報と、の差分に基づいて報酬を算出し、前記報酬と、前記環境情報と、前記状態情報と、に基づいて、前記価値関数を算出してよい。

前記機械学習システムが、複数のターゲット行動情報を含むターゲット状態関連情報を保有してよい。

前記ターゲット状態関連情報には、時間別ターゲット状態情報及び／又は段階別ターゲット状態情報が含まれていてよい。

前記環境情報には、香り、照明、温度、湿度、映像、又は音声に関する情報が含まれていてよい。

香り制御部をさらに備えており、前記香り制御部が、前記機械学習分類器が選定した前記環境情報に基づいて、発生する香りを制御してよい。

香り付着部をさらに備えており、前記香り付着部が、前記機械学習分類器が選定した前記環境情報に基づいて、香りを物に付着し、前記機械学習分類器が、前記環境情報に基づいて、前記香り制御部及び前記香り付着部のいずれが香りを発生させるかを判定してよい。

照明制御部をさらに備えており、前記照明制御部が、前記機械学習分類器が選定した前記環境情報に基づいて、照射する光を制御してよい。

空気調節部をさらに備えており、前記空気調節部が、前記機械学習分類器が選定した前記環境情報に基づいて、温度及び／又は湿度を制御してよい。

映像制御部をさらに備えており、前記映像制御部が、前記機械学習分類器が選定した前記環境情報に基づいて、表示する映像を制御してよい。

音声制御部をさらに備えており、前記音声制御部が、前記機械学習分類器が選定した前記環境情報に基づいて、再生する音声を制御してよい。

前記価値関数が複数の価値グループに区分けされており、前記機械学習分類器が、前記複数の価値グループのそれぞれが保有する前記価値関数を用いることができる。

複数の状態取得部と、達成難易度算出部と、をさらに備えており、前記達成難易度算出部が、複数の状態取得部のそれぞれが取得した状態情報に基づいて、前記ターゲット行動に関する達成難易度を算出してよい。

前記達成難易度には、前記ターゲット行動に促した割合を示す達成率が含まれてよい。

前記達成難易度には、前記ターゲット行動に促すまでの標準の時間を示す標準達成時間が含まれてよい。

前記達成難易度には、前記ターゲット行動に促したときの前記環境情報の項目の数の平均を示す平均環境項目数が含まれてよい。

また、本技術は、人の行動に関する状態情報を少なくとも取得する状態取得部と、前記状態情報を取得したときの前記人の周囲の環境に関する環境情報と、前記状態情報と、を評価して価値関数を得る評価部と、前記行動をターゲット行動に促すために、前記価値関数を強化学習して、該価値関数が最も高くなるときの前記環境情報を選定する機械学習分類器と、を少なくとも備えている、機械学習装置を提供する。

また、本技術は、人の行動に関する状態情報を少なくとも取得することと、前記状態情報を取得したときの前記人の周囲の環境に関する環境情報と、前記状態情報と、を評価して価値関数を得ることと、前記行動をターゲット行動に促すために、前記価値関数を強化学習して、該価値関数が最も高くなるときの前記環境情報を選定することと、を少なくとも含んでいる、機械学習方法を提供する。

図面の簡単な説明

[0008] [図1]本技術の一実施形態に係る機械学習システム１の構成を示すブロック図

である。

[図2]本技術の一実施形態に係る行動変容の一例を示す概念図である。

[図3]本技術の一実施形態に係る行動変容の一例を示す概念図である。

[図4]本技術の一実施形態に係る行動変容の一例を示す概念図である。

[図5]本技術の一実施形態に係る行動変容の一例を示す概念図である。

[図6]本技術の一実施形態に係る行動変容の一例を示す概念図である。

[図7]本技術の一実施形態に係る行動変容の一例を示す概念図である。

[図8]本技術の一実施形態に係る行動変容の一例を示す概念図である。

[図9]本技術の一実施形態に係るターゲット状態関連情報の一例を示すデータベースである。

[図10]本技術の一実施形態に係る機械学習装置１０のハードウェア構成を示すブロック図である。

[図11]本技術の一実施形態に係る機械学習システム１の構成を示すブロック図である。

[図12]本技術の一実施形態に係る香り制御部２３の構成を示すブロック図である。

[図13]本技術の一実施形態に係る機械学習システム１の構成を示すブロック図である。

[図14]本技術の一実施形態に係る機械学習システム１の手順を示すフローチャートである。

[図15]本技術の一実施形態に係る照明制御部２４の構成を示すブロック図である。

[図16]本技術の一実施形態に係る空気調節部２５の構成を示すブロック図である。

[図17]本技術の一実施形態に係る映像制御部２６の構成を示すブロック図である。

[図18]本技術の一実施形態に係る音声制御部２７の構成を示すブロック図である。

[図19]本技術の一実施形態に係る機械学習装置１０が用いるデータベースの一例である。

[図20]本技術の一実施形態に係る機械学習装置１０の手順の一例を示すフローチャートである。

[図21]本技術の一実施形態に係る機械学習システム１の構成を示すブロック図である。

[図22]本技術の一実施形態に係る達成難易度算出部５４が算出する達成難易度について説明する図である。

[図23]本技術の一実施形態に係る達成難易度算出部５４が算出する達成難易度について説明する図である。

[図24]本技術の一実施形態に係る機械学習方法の手順を示すフローチャートである。

発明を実施するための形態

[0009] 以下、本技術を実施するための好適な形態について説明する。以下に説明する実施形態は、本技術の代表的な実施形態の一例を示したものであり、これにより本技術の範囲が狭く解釈されることはない。また、各図は模式図であり、必ずしも厳密に図示されたものではない。

[0010] 本技術の説明は以下の順序で行う。

１．本技術の第１の実施形態（機械学習システムの例１）

（１）概要

（２）評価部

（３）機械学習分類器

（４）行動変容のフロー

（５）複数のターゲット状態情報

（６）ハードウェア構成

２．本技術の第２の実施形態（機械学習システムの例２）

（１）概要

（２）香り制御部

（３）香り付着部
（４）照明制御部
（５）空気調節部
（６）映像制御部
（７）音声制御部
３．本技術の第３の実施形態（機械学習システムの例３）
４．本技術の第４の実施形態（機械学習システムの例４）
５．本技術の第５の実施形態（機械学習システムの例５）
（１）概要
（２）達成難易度
６．本技術の第６の実施形態（機械学習方法）

[0011] ［１．本技術の第１の実施形態（機械学習システムの例１）］

［（１）概要］

本技術の一実施形態に係る機械学習システムは、人の行動に関する情報と、その人の周囲の環境に関する情報と、を評価して機械学習することにより、行動と環境との相関関係を得ることができる。これにより、環境を制御することにより人の行動をターゲット行動に促すことができるようになる。

[0012] 本技術の一実施形態に係る機械学習システムの構成について図１を参照しつつ説明する。図１は、本技術の一実施形態に係る機械学習システム１の構成を示すブロック図である。

[0013] 図１に示されるとおり、機械学習システム１は、例えば機械学習装置１０を含むことができる。機械学習装置１０は、例えば、状態取得部１１、評価部１２、記録部１３、及び機械学習分類器１４などを備えることができる。

[0014] 状態取得部１１は、人の行動に関する状態情報を少なくとも取得する。これにより、機械学習システム１は、環境の変化に応じた人の行動変容を把握できる。

[0015] 状態情報には、例えば、Ｗｅｂサイトの閲覧などに用いられるＣｏｏｋｉｅ、ＥＣ（Electronic Commerce）サイトなどにおける購入履歴、ＧＰＳ（Gl

obal Positioning System）などを用いて得られる位置情報、チャットなどの会話履歴、その他センシング技術を用いて得られる情報などが含まれる。

[0016] さらに状態情報には、人が存在する地域の天気又は気温に関する情報が含まれていてもよい。これにより、機械学習システム１は、人が存在する地域の天気又は気温などに特有の行動変容を把握できる。

[0017] 評価部１２は、状態情報を取得したときの人の周囲の環境に関する環境情報と、前記状態情報と、を評価して価値関数を得る。これにより、環境情報と状態情報との相関関係が得られる。具体的な評価手段については後述する。

[0018] 環境情報には、例えば、香り、照明、温度、湿度、映像、又は音声に関する情報などが含まれている。環境情報の具体例については後述する。

[0019] 記録部１３は、状態情報や環境情報などを記録する。なお、機械学習システム１は、状態情報は取得するが、環境情報は取得せずに記録しているものを用いる。

[0020] 機械学習分類器１４は、人の行動をターゲット行動に促すために、価値関数を強化学習して、該価値関数が最も高くなるときの前記環境情報を選定する。これにより、行動と環境との相関関係が得られる。

[0021] 機械学習の手法は特に限定されないが、例えば強化学習が用いられることができる。強化学習とは、ソフトウェアが現在の状態（本技術における状態情報）を観測して取るべき行動（本技術における環境情報の変更）を決定する問題を扱う機械学習の一種である。エージェント（本技術における機械学習分類器１４）は、行動を試行錯誤して強化学習することにより、価値が最も高くなるときの行動を決定できる。

[0022] 従来、強化学習を実現するための手法として、例えばモンテカルロ法、動的計画法、ＳＡＲＳＡ（state-action-reward-state-action）、あるいはＱ学習（Q-learning）などが用いられている。本技術については、強化学習の一例であるＱ学習を用いて説明する。なお、Ｑ学習のほかの強化学習手法が本技術に用いられてもよい。

[0023] なお、図示を省略するが、機械学習装置１０は、それぞれの構成要素を制御する制御部や、通信ネットワークを介して通信を行う通信インタフェースなどを有していてもよい。

[0024] ［（２）評価部］

上述したように、評価部１２は、状態情報を取得したときの人の周囲の環境に関する環境情報と、前記状態情報と、を評価して価値関数を得る。

[0025] これを実現するための手段は特に限定されないが、例えば評価部１２は、報酬算出部（図示省略）及び価値算出部（図示省略）を有することができる。

[0026] 環境情報の変更に応じて、人の行動に関する状態情報が変化しうる。報酬算出部は、機械学習システム１が環境情報を変更したときの、ターゲット行動に関するターゲット状態情報と、状態情報との差分に基づいて報酬Ｒを算出する。報酬Ｒの値が大きいほど、ターゲット状態情報と状態情報との差分が小さいことを示す。つまり、報酬Ｒの値が大きいほど、人の行動がターゲット行動に近いことを示す。

[0027] 報酬Ｒは、ターゲット状態情報に応じた得点Ｐｔと、状態情報に応じた得点Ｐｍと、を用いて、例えば下記の式（１）で表現できる。

[0028] [数1]

$$R = 1 - \frac{|Pt - Pm|}{Pt} \quad \cdots (1)$$

[0029] 具体例を用いて説明する。「ＥＣサイトを利用して商品Ａを購入する」という行動をターゲット行動に設定する。そして、環境情報の変更によって人がターゲット行動をしたとき、得点として５点が設定されるとする。

[0030] また、環境情報の変更によって、人がターゲット行動に近い行動である「商品Ａが掲載されているＷｅｂサイトを閲覧する」という行動をしたとき、得点として２点が設定されるとする。

[0031] 上記の式（１）に当てはめると、ターゲット状態情報に応じた得点Ｐｔは「５」である。環境情報の変更によって人がターゲット行動をしたとき、状

態情報に応じた得点Ｐｍも「５」である。このとき、報酬Ｒは「１」となる。

[0032] 環境情報の変更によって人がターゲット行動に近い行動をしたとき、状態情報に応じた得点Ｐｍは「２」である。このとき、報酬Ｒは「０．４」となる。

[0033] 環境情報の変更によって人がこの２つの行動ではない行動をしたとき、状態情報に応じた得点Ｐｍは「０」とする。このとき、報酬Ｒも「０」となる。

[0034] このことから、環境情報の変更による人の行動がターゲット行動に近いほど、報酬Ｒの値が高くなることがわかる。報酬算出部は、報酬Ｒが最も高くなるときの環境情報の変更を行ったときの報酬Ｒを算出する。

[0035] 価値算出部は、報酬Ｒと、環境情報と、状態情報と、に基づいて、価値関数Ｑを算出する。価値算出部は、報酬Ｒが最も高くなるときの環境情報の変更を行ったときの状態情報に基づいて、価値関数Ｑを算出する。例えば時刻ｔにおける状態情報s_tに対して環境情報の変更a_tが行われたときの価値関数はＱ（s_t，a_t）と表現される。

[0036] この価値関数Ｑは、例えば記録部１３が記録してもよい。より具体的には、記録部１３は、状態情報又は環境情報ごとに価値関数Ｑをテーブル上に記録してもよい。

[0037] ［（３）機械学習分類器］

上述したように、機械学習分類器１４は、価値関数Ｑを強化学習して、価値関数Ｑが最も高くなるときの環境情報を選定する。

[0038] この強化学習について説明する。機械学習分類器１４は、価値関数Ｑが最も高くなるように試行錯誤と学習を自動的に行う。価値関数Ｑが高いほど、人の行動がターゲット行動に近いことになる。価値関数Ｑが最も高くなるように強化学習することによって、機械学習分類器１４は、人の行動をターゲット行動に促すことができる。

[0039] 価値関数Ｑが最も高くなるときの環境情報が決定したとき、機械学習分類

器１４は、価値関数Ｑを更新する。例えば時刻ｔにおける状態情報$ｓ_{t}$に対して、環境情報の変更$ａ_{t}$が行われ、時刻ｔ＋１における状態情報$ｓ_{t+1}$に遷移したとき、価値関数Ｑ（$ｓ_{t}$，$ａ_{t}$）は、下記の数式（２）を用いて更新される。

[0040] [数2]

$$Q(s_t,a_t) \leftarrow Q(s_t,a_t) + \alpha\left[R_{t+1} + \gamma \max_{a} Q(s_{t+1},a) - Q(s_t,a_t)\right] \quad \cdots (2)$$

[0041] ここで、αは学習係数である。学習係数αは０より大きく１以下の値が設定されるが、０．１程度の値が用いられることが多い。

[0042] $Ｒ_{t+1}$は、この状態情報の遷移によって得られた報酬である。

[0043] γは割引率である。割引率γは０より大きく１以下の値が設定されるが、０．９～０．９９程度の値が用いられることが多い。

[0044] ｍａｘＱ（$ｓ_{t+1}$，ａ）は、将来の理想の価値関数である。ｍａｘＱ（$ｓ_{t+1}$，ａ）は、時刻ｔ＋１における状態$ｓ_{t+1}$において、価値関数Ｑが最も高くなるときの行動ａを選択したときの価値関数である。この価値関数ｍａｘＱ（$ｓ_{t+1}$，ａ）に、割引率γが乗算される。

[0045] 機械学習分類器１４は、上記の数式（２）を用いて価値関数Ｑを更新し続け、価値関数Ｑが最も高くなるときの環境情報を選定する。これにより、機械学習分類器１４は、ターゲット行動に促すことができる環境情報を選定できる。

[0046] ［（４）行動変容のフロー］

ターゲット行動に至るまでに、複数の行動変容が行われることが想定される。このことについて図２を参照しつつ説明する。図２は、本実施形態に係る行動変容の一例を示す概念図である。図２に示されるとおり、ターゲット行動に「動画共有サイトで動画を閲覧する」という行動が設定されている。そして、このターゲット行動に促すための複数の行動変容のフローが設定されている。機械学習分類器１４は、強化学習を繰り返すことにより、この行動変容のフローを構成することができる。機械学習分類器１４は、このフローをたどることで、人の行動をターゲット行動に促すことができる。

[0047] また、ターゲット行動に対する近さに応じて、人の行動が複数のレベルに区分けされている。例えば第１のレベルに設定される行動は、ターゲット行動に最も近い行動であってよい。第２のレベルに設定される行動は、ターゲット行動にその次に近い行動であってよい。第１のレベルの行動に係る価値関数Ｑは、第２のレベルの行動に係る価値関数Ｑより高い。

[0048] この例では、第１のレベルに「トイレに行ってきた」及び「ソファに座った」という行動が設定されている。第２のレベルに「子供が寝た」「帰宅した」「テーブルから離れた」「飲酒した」という行動が設定されている。そして、複数の行動のそれぞれが接続されることにより、行動変容のフローが構成されている。例えば、この個人の行動変容の特性においては、「子供が寝た」という行動が行われたときは、「トイレに行く」という行動が行われる傾向にあることが示されている。

[0049] この行動変容のフローの他の例を図３～７に示す。図３～７は、本実施形態に係る行動変容の一例を示す概念図である。図３では、「ＥＣサイトを利用する」というターゲット行動に促すための行動変容のフローの一例が示されている。図４では、「ＳＮＳ（Social Networking Service）を閲覧する」というターゲット行動に促すための行動変容のフローの一例が示されている。図５では、「高額の買い物をする」というターゲット行動に促すための行動変容のフローの一例が示されている。図６では、「ビールを飲む」というターゲット行動に促すための行動変容のフローの一例が示されている。図７では、「就寝する」というターゲット行動に促すための行動変容のフローの一例が示されている。このように、ターゲット行動に応じて様々な行動変容のフローが構成される。

[0050] また、ターゲット行動が同じであっても、ターゲット行動に促すための行動変容のフローは、個人によって異なっていてよい。このことについて図８を参照しつつ説明する。図８は、本実施形態に係る行動変容の一例を示す概念図である。図８Ａは、ある個人の行動変容の一例を示している。図８Ａに示されるとおり、ターゲット行動として「衣類を購入する」という行動が設

定されている。

[0051] 一方で、図８Ｂは、図８Ａに示される個人ではない個人の行動変容の一例を示している。図８Ｂに示されるとおり、ターゲット行動は図８Ａと同じであるが、ターゲット行動に促すための行動変容のフローが異なっている。

[0052] [（５）複数のターゲット状態情報]

本技術の一実施形態に係る機械学習装置１０は、１つのターゲット行動に関するターゲット状態情報を保有していてよいが、複数のターゲット行動に関する複数のターゲット状態情報を保有していてもよい。この複数のターゲット行動の一部又は全部は、例えば時間別及び／又は段階別に設定されることができる。

[0053] 複数のターゲット行動の一部又は全部は、例えば時間別に設定されることができる。具体例を用いて説明すると、複数のターゲット行動の一部又は全部は、１日のうちの第１の時間帯（例えば午前０～６時台）におけるターゲット行動、第２の時間帯（例えば午前７～午後７時台）におけるターゲット行動、及び第３の時間帯（例えば午後８～１１時台）におけるターゲット行動などに設定されることができる。

[0054] 第１の時間帯（例えば午前０～６時台）におけるターゲット行動は、例えば「睡眠をとる」であってよい。第２の時間帯（例えば午前７～午後７時台）におけるターゲット行動は、例えば「食品Ｓを食べる」であってよい。第３の時間帯（例えば午後８～１１時台）におけるターゲット行動は、例えば「飲料Ｔを飲む」であってよい。

[0055] 複数のターゲット行動の一部又は全部が時間別に設定されることにより、例えば時間帯に応じてターゲット行動が柔軟に設定されることができる。例えば、機械学習装置１０は、「午後３時に食品Ｓを食べる」というターゲット行動に促すことができる。

[0056] あるいは、複数のターゲット行動の一部又は全部は、例えば段階別に設定されることができる。具体例を用いて説明すると、複数のターゲット行動の一部又は全部は、第１の段階におけるターゲット行動及び第２の段階におけ

るターゲット行動に設定されることができる。第１の段階におけるターゲット行動は、例えば「店舗Ｕに行く」であってよい。第２の段階におけるターゲット行動は、例えば「食品Ｓを食べる」であってよい。

[0057] 複数のターゲット行動の一部又は全部が段階別に設定されることにより、例えば一連の流れを有する複数のターゲット行動が設定されることができる。例えば、機械学習装置１０は、「店舗Ｕに行ったときは食品Ｓを食べる」というターゲット行動に促すことができる。

[0058] あるいは、複数のターゲット行動の一部又は全部は、例えば時間別及び段階別に設定されることができる。具体例を用いて説明すると、前記第３の時間帯において、前記第１及び第２の段階におけるターゲット行動に設定されることができる。

[0059] 複数のターゲット行動の一部又は全部が時間別及び段階別に設定されることにより、例えば一連の流れを有する複数のターゲット行動が、時間帯に応じて柔軟に設定されることができる。例えば、機械学習装置１０は、「朝、店舗Ｕに行ったときは食品Ｓを食べる」というターゲット行動に促すことができる。

[0060] これを実現するために、本技術の一実施形態に係る機械学習装置１０は、複数のターゲット状態情報を含むターゲット状態関連情報を保有してよい。

[0061] ターゲット状態関連情報について図９を参照しつつ説明する。図９は、本実施形態に係るターゲット状態関連情報の一例を示すデータベースである。図９では、時間帯とターゲット行動とが関連付けあれている。第１の時間帯（午前０～６時台）におけるターゲット行動は「就寝する」である。第２の時間帯（午前７～午後７時台）におけるターゲット行動は「食品Ｓを食べる」である。第３の時間帯（午後８～１１時台）におけるターゲット行動は、例えば「飲料Ｔを飲む」である。なお、段階別に設定されているターゲット状態関連情報についても同様である。

[0062] ターゲット状態関連情報は、例えば機械学習装置１０が備える記録部１３などに記録されることができる。なお、ターゲット状態関連情報は、機械学

習装置１０のほかのコンピュータ装置が保有してよい。例えば、ターゲット状態関連情報は、クラウド上のサーバに保有されていてよい。このとき、機械学習装置１０は、情報通信ネットワークを介して前記サーバからターゲット状態関連情報を受信してよい。

[0063] ［（６）ハードウェア構成］

機械学習装置１０のハードウェア構成について図１０を参照しつつ説明する。図１０は、本実施形態に係る機械学習装置１０のハードウェア構成を示すブロック図である。図１０に示されるとおり、機械学習装置１０は、構成要素として、ＣＰＵ１０１、ストレージ１０２、ＲＡＭ（Random Access Memory）１０３、及び通信インタフェース１０４を備えることができる。それぞれの構成要素は、例えばデータの伝送路としてのバスで接続されている。

[0064] ＣＰＵ１０１は、例えばマイクロコンピュータにより実現され、機械学習装置１０のそれぞれの構成要素を制御する。ＣＰＵ１０１は、例えば、評価部１２や機械学習分類器１４として機能しうる。この機械学習分類器１４は、例えばプログラムにより実現されうる。このプログラムをＣＰＵ１０１が読み込むことによって機能しうる。

[0065] ストレージ１０２は、ＣＰＵ１０１が使用するプログラムや演算パラメータ等の制御用データ等を記憶する。ストレージ１０２は、例えばＨＤＤ（Hard Disk Drive）又はＳＳＤ（Solid State Drive）等を利用することにより実現されうる。ストレージ１０２は、例えば記録部１３として機能しうる。

[0066] ＲＡＭ１０３は、例えば、ＣＰＵ１０１により実行されるプログラム等を一時的に記憶する。

[0067] 通信インタフェース１０４は、例えばＷｉ－Ｆｉ、Ｂｌｕｅｔｏｏｔｈ（登録商標）、ＬＴＥ（Long Term Evolution）等の通信技術を利用して、情報通信ネットワークを介して通信する機能を有する。

[0068] 機械学習分類器１４などを実現するプログラムは、機械学習システム１のほかのコンピュータ装置又はコンピュータシステムに格納されてもよい。この場合、機械学習システム１は、このプログラムが有する機能を提供するク

ラウドサービスを利用することができる。このクラウドサービスとして、例えばＳａａＳ（Software as a Service）、ＩａａＳ（Infrastructure as a Service）、ＰａａＳ（Platform as a Service）等が挙げられる。

[0069] さらにこのプログラムは、様々なタイプの非一時的なコンピュータ可読媒体（non-transitory computer readable medium）を用いて格納され、コンピュータに供給することができる。非一時的なコンピュータ可読媒体は、様々なタイプの実体のある記録媒体（tangible storage medium）を含む。非一時的なコンピュータ可読媒体の例は、磁気記録媒体（例えばフレキシブルディスク、磁気テープ、ハードディスクドライブ）、光磁気記録媒体（例えば光磁気ディスク）、Compact Disc Read Only Memory（CD-ROM）、CD-R、CD-R/W、半導体メモリ（例えば、マスクROM、Programmable ROM（PROM）、Erasable PROM（EPROM）、フラッシュROM、Random Access Memory（RAM））を含む。また、上記プログラムは、様々なタイプの一時的なコンピュータ可読媒体（transitory computer readable medium）によってコンピュータに供給されてもよい。一時的なコンピュータ可読媒体の例は、電気信号、光信号、及び電磁波を含む。一時的なコンピュータ可読媒体は、電線及び光ファイバ等の有線通信路、又は無線通信路を介して、上記プログラムをコンピュータに供給できる。

[0070] ［２．本技術の第２の実施形態（機械学習システムの例２）］

［（１）概要］

本技術の一実施形態に係る機械学習システム１は、人の周囲の環境を制御するための環境制御装置を備えていてよい。環境制御装置は、機械学習分類器１４が選定した環境情報に基づいて、人の周囲の環境を制御する。これにより、機械学習システム１は、人の行動をターゲット行動に促すことができる。

[0071] 機械学習システム１は、人の行動をターゲット行動に促すことにより、例えば販売を促進できる。機械学習システム１は、顧客に商品を購入させるために店舗の内部又は外部の環境を制御できる。なお、機械学習システム１が

利用されるのは実店舗に限られず、例えばＥＣサイトであってもよい。あるいは、機械学習システム１は、ダウンロード又はストリーミングされるコンテンツや、Ｗｅｂサイトにも利用され、コンテンツやＷｅｂサイトの閲覧を促進できる。

[0072] あるいは、機械学習システム１は、人の行動をターゲット行動に促すことにより、例えば人の生活習慣を改善できる。具体的には、機械学習システム１は、禁煙や禁酒などを促すために環境を制御できる。さらに、例えば買い物依存、睡眠不足、運動不足などの改善に機械学習システム１が利用されてもよい。

[0073] あるいは、機械学習システム１は、人の行動をターゲット行動に促すことにより、例えば選挙において投票に行かせることができる。

[0074] あるいは、機械学習システム１は、人の行動をターゲット行動に促すことにより、例えば公衆衛生やモラルなどを向上させることができる。具体的には、ごみの分別、ごみのポイ捨てや駆け込み乗車などの抑制、咳エチケットや手洗いなどの順守、整列、暴力の抑制などに機械学習システム１が利用されてもよい。

[0075] あるいは、機械学習システム１は、人の行動をターゲット行動に促すことにより、例えば仕事の効率を向上させることができる。具体的には、集中力の向上、よそ見の抑制、オンオフの切り替えなどに機械学習システム１が利用されてもよい。

[0076] 本実施形態に係る機械学習システム１の構成について図１１を参照しつつ説明する。図１１は、本実施形態に係る機械学習システム１の構成を示すブロック図である。なお、第１の実施形態と同様の構成要素については、同じ符号を付し、詳細な説明は省略する。

[0077] 図１１に示されるとおり、本実施形態に係る機械学習システム１は、環境制御装置２０をさらに備えていてよい。環境制御装置２０及び機械学習装置１０は、情報通信ネットワーク４０を介して接続されている。環境制御装置２０は、機械学習分類器１４が選定した環境情報に基づいて、例えば香り、

照明、温度、湿度、映像、又は音声などの環境を制御する。

[0078] 環境制御装置２０は、例えば通信制御部２１、メモリ２２、香り制御部２３、照明制御部２４、空気調節部２５、映像制御部２６、及び音声制御部２７などを有することができる。

[0079] なお、香り制御部２３、照明制御部２４、空気調節部２５、映像制御部２６、及び音声制御部２７については、環境制御装置２０は、これら全ての構成要素を有していなくてもよく、少なくともいずれか一つの構成要素を有していればよい。

[0080] また、機械学習システム１は、複数の環境制御装置２０を備えていてよい。例えば機械学習システム１が２つの環境制御装置２０を備えているとき、一方の環境制御装置２０が香り制御部２３を有しており、他方の環境制御装置２０が照明制御部２４を有していてもよい。

[0081] 通信制御部２１は、情報通信ネットワーク４０を介して機械学習装置１０との情報を通信できる。また、通信制御部２１は、それぞれの構成要素を制御してもよい。

[0082] メモリ２２は、環境制御装置２０が用いる情報、例えば環境情報など記録できる。

[0083] なお、機械学習装置１０が備える機械学習分類器１４などは、例えば環境制御装置２０に備えられていてもよいし、その他のコンピュータ装置に備えられていてもよい。

[0084] [（２）香り制御部]

人の周囲の香りを制御するために、環境制御装置２０は、例えば香り制御部２３を備えることができる。香り制御部２３は、機械学習分類器１４が選定した環境情報に基づいて、発生する香りを制御する。香り制御部２３を備える環境制御装置２０は、例えばアロマディフューザーなどが用いられることにより実現できる。

[0085] なお、香りには、人が香りとして認知できるもののほか、人が香りとして認知できないが吸引されることにより人に対して何らかの作用を及ぼすもの

も含まれる。例えば、吸引される医療用鎮静剤、あるいは、酸素や二酸化炭素などに代表されるような、無臭であるが吸引されることにより人の体調に作用する気体なども香りに含まれる。

[0086] 人は、香り制御部２３が制御するその人に最適化された香りを吸引することにより、無自覚にターゲット行動に促される。

[0087] 香り制御部２３の構成について図１２を参照しつつ説明する。図１２は、本実施形態に係る香り制御部２３の構成を示すブロック図である。図１２に示されるとおり、香り制御部２３は、例えば添加成分カートリッジ２３１、香り制御部２３２、及び香り出力部２３３などを有することができる。

[0088] 添加成分カートリッジ２３１は、香りの添加成分を収容する部品である。添加成分カートリッジ２３１は、交換可能であってよい。添加成分カートリッジ２３１は、例えば添加成分を収容したボンベ、瓶、缶などの容器、添加成分を吸着した紙、不織布、又は石などの素材、あるいは添加成分を混合したワックスや石鹸などの固形物などであってよい。

[0089] 添加成分は、例えば粉末やゲルなどを含む固体、液体、気体であってよいし、これらの混合物であってよい。添加成分は、例えば天然から抽出される天然香料、化学合成から得られる合成香料、又はこれらが調合される調合香料などであってよい。あるいは、添加成分は、香料を含んでいなくてもよい。

[0090] 香り制御部２３２は、環境情報に基づいて、香りを生成するための添加成分を制御する。香り制御部２３２は、例えば添加成分が調合されるときの、それぞれの添加成分の比率などを決定できる。あるいは、香り制御部２３２は、希釈倍率を決定してもよい。前記比率又は前記希釈倍率は、機械学習分類器１４が選定した環境情報に応じて決定される。

[0091] あるいは、香り制御部２３３は、香りの出力に関する例えば噴射圧や噴射回数などを制御してもよい。前記噴射圧又は前記噴射回数は、機械学習分類器１４が選定した環境情報に応じて決定される。

[0092] 香り出力部２３３は、香り制御部２３２により決定された情報に基づいて

、香りを出力する。

[0093] 香り制御部２３を備える機械学習システム１は、特定の香りを人に吸引させることにより、人の行動をターゲット行動に促すことができる。例えば、生理的に特定の商品を購入したくなる香りを人に吸引させることにより、機械学習システム１は、その人に通信販売で特定の商品を購入させたり、又は店舗に向かわせたりすることができる。

[0094] あるいは、香り及びコンテンツが関連付けられていてもよい。これにより、機械学習システム１は、強化学習する前に、香り及びコンテンツの関連付けを人に無意識に学習させることができる。

[0095] 例えば、機械学習システム１は、特定の香りを人に吸引させるとともに、特定の映像コンテンツを人に視聴させることができる。この映像コンテンツには、例えば特定の商品に関する広告が含まれている。これにより、機械学習システム１は、特定の香り及び特定の商品の関連付けを人に無意識に学習させることができる。この特定の香りを人に吸引させることにより、機械学習システム１は、その人に通信販売で特定の商品を購入させたり、又は店舗に向かわせたりすることができる。

[0096] あるいは、香り及び環境が関連付けられていてもよい。この環境は、人が現実に体験する場所又は物などに関する環境である。この場所には、例えば、店舗、公共交通機関、映画館、劇場、テーマパークなどが含まれる。これにより、機械学習システム１は、強化学習する前に、香り及び環境の関連付けを人に無意識に学習させることができる。

[0097] 例えば、機械学習システム１は、特定の店舗に訪れた人に、特定の香りを吸引させることができる。これにより、機械学習システム１は、特定の香り及び特定の店舗の関連付けを人に無意識に学習させることができる。この特定の香りを店舗とは異なる場所で人に吸引させることにより、機械学習システム１は、その人を店舗に向かわせたり、店舗に陳列されている商品を通信販売で購入させたりすることができる。

[0098] 例えば、機械学習システム１は、例えばコーヒーを飲むなどの体験によっ

て、特定の商品から生じる香りを吸引させることができる。これにより、機械学習システム１は、特定の香り及び特定の商品の関連付けを人に無意識に学習させることができる。香り制御部２３がこの特定の香りを発生することにより、機械学習システム１は、この特定の商品を通信販売で購入させたり、この特定の商品が陳列されている店舗に向かわせたりすることができる。

[0099] ［（３）香り付着部］

香りは、物に付着されてもよい。物とは、例えばターゲット行動を促す人に配達される衣類、本、雑貨、販売促進物、梱包材などであってよい。人は、物に付着されその人に最適化された香りを吸引することにより、無自覚にターゲット行動に促される。

[0100] これを実現するために、機械学習システム１は、香り付着部を備えることができる。このことについて図１３を参照しつつ説明する。図１３は、本実施形態に係る機械学習システム１の構成を示すブロック図である。なお、第１の実施形態と同様の構成要素については、同じ符号を付し、詳細な説明は省略する。図１３に示されるとおり、機械学習システム１は、香り付着部３０をさらに備えている。香り付着部３０及び機械学習装置１０は、情報通信ネットワーク４０を介して接続されている。なお、香り付着部３０は、香り制御部２３と同様の構成であってよい。

[0101] 上述した香り制御部２３は、ターゲット行動に促される人の周囲に配される。一方香り付着部３０で、香り付着部３０は、例えば物を出荷する工場などに配される。香り付着部３０は、機械学習分類器１４が選定した環境情報に基づいて、香りを物に付着する。

[0102] 機械学習分類器１４は、前記環境情報に基づいて、香り制御部２３及び香り付着部３０のいずれが香りを発生させるかを判定する。

[0103] このときの機械学習システム１の手順について図１４を参照しつつ説明する。図１４は、本実施形態に係る機械学習システム１の手順を示すフローチャートである。

[0104] 図１４に示されるとおり、機械学習の初期段階においては、機械学習分類

器１４が香りと行動との相関関係を強化学習し（ステップＳ１１）、機械学習分類器１４が環境情報を選定し（ステップＳ１２）、香り制御部２３が香りを発生させることを機械学習分類器１４が判定し（ステップＳ１３）、香り制御部２３が人の周囲の香りを制御する（ステップＳ１４）。香りと行動との相関関係が十分に学習できるまで、ステップＳ１１～Ｓ１４が繰り返される（ステップＳ１５：Ｎｏ）。

[0105] 香りと行動との相関関係が十分に学習できた段階（ステップＳ１５：Ｙｅｓ）で、香り付着部３０が香りを発生させることを機械学習分類器１４が判定し（ステップＳ１６）、香り付着部３０が香りを物に付着する（ステップＳ１７）。

[0106] これにより、機械学習システム１は、人の周囲の香りをより柔軟に制御できる。例えば、機械学習の初期段階（例えば学習開始から１～３ヶ月程度）においては、人の周囲に配される香り制御部２３が短期間で香りを変更しつつ、機械学習分類器１４が香りと行動との相関関係を高効率で強化学習する。機械学習分類器１４は、ターゲット行動に最適な香りを判定する。

[0107] その後、ターゲット行動に最適な香りが付着された物が配達などされることにより、継続してターゲット行動に促すことができる。香り付着部３０が長期間で香りを変更しつつ、機械学習分類器１４が香りと行動との相関関係を継続して強化学習する。

[0108] ［（４）照明制御部］

図１１の説明に戻る。人の周囲の照明を制御するために、環境制御装置２０は、例えば照明制御部２４を備えることができる。照明制御部２４は、機械学習分類器１４が選定した環境情報に基づいて、照射する光を制御する。照明制御部２４を備える環境制御装置２０は、例えば電球（いわゆるスマート電球を含む）などが用いられることにより実現できる。

[0109] 人は、照明制御部２４が照射するその人に最適化された光を視認することにより、無自覚にターゲット行動に促される。

[0110] 照明制御部２４の構成について図１５を参照しつつ説明する。図１５は、

本実施形態に係る照明制御部２４の構成を示すブロック図である。図１５に示されるとおり、照明制御部２４は、例えば光制御部２４１及び光出力部２４２などを有することができる。

[0111] 光制御部２４１は、出力する光の表現を制御する。具体的に説明すると、光制御部２４１は、例えば光の色温度や輝度などを決定できる。前記色温度又は前記輝度は、機械学習分類器１４が選定した環境情報に応じて決定される。例えば色温度は３５００～３９００Ｋなどと決定されてよいし、輝度は３０００～４０００ｌｍなどと決定されてよい。なお、このように決定された値に範囲がある場合は、光制御部２４１は、この範囲の中からランダムに値を決定してよい。機械学習装置１０は、強化学習を積み重ねながら、この範囲を狭めることができる。なお、以下に説明する他の構成要素についても同様である。

[0112] 光出力部２４２は、光制御部２４１により決定された情報に基づいて、光を出力する。

[0113] [（５）空気調節部]

人の周囲の空気を調節するために、環境制御装置２０は、例えば空気調節部２５を備えることができる。空気調節部２５は、機械学習分類器１４が選定した環境情報に基づいて、温度及び／又は湿度を制御する。空気調節部２５を備える環境制御装置２０は、例えばエアコンなどが用いられることにより実現できる。

[0114] 人は、空気調節部２５により制御されるその人に最適化された温度及び／又は湿度により、無自覚にターゲット行動に促される。

[0115] 空気調節部２５の構成について図１６を参照しつつ説明する。図１６は、本実施形態に係る空気調節部２５の構成を示すブロック図である。図１６に示されるとおり、空気調節部２５は、例えば空気制御部２５１及び空気出力部２５２などを有することができる。

[0116] 空気制御部２５１は、空気の温度及び／又は湿度を決定できる。前記温度及び／又は前記湿度は、機械学習分類器１４が選定した環境情報に応じて決

定される。例えば温度は２５．５～２７．５度などと決定されてよいし、湿度は４５～５０％などと決定されてよい。

[0117] 空気出力部２５２は、空気制御部２５１により決定された情報に基づいて、空気を出力する。

[0118] [（６）映像制御部]

人に対して表示される映像を制御するために、環境制御装置２０は、例えば映像制御部２６を備えることができる。映像制御部２６は、機械学習分類器１４が選定した環境情報に基づいて、表示する映像を制御する。映像制御部２６を備える環境制御装置２０は、例えばテレビジョン、携帯ゲーム機、ＰＣ、タブレット、スマートフォン、ＨＭＤ（Head Mounted Display）、ウェアラブル装置、カーナビなどが用いられることにより実現できる。

[0119] なお、映像には動画のほか静止画も含まれる。また、映像には音声が含まれていてもよい。

[0120] 人は、映像制御部２６が表示するその人に最適化された映像を視認することにより、無自覚にターゲット行動に促される。

[0121] 映像制御部２６の構成について図１７を参照しつつ説明する。図１７は、本実施形態に係る映像制御部２６の構成を示すブロック図である。図１７に示されるとおり、映像制御部２６は、例えば映像選択部２６１及び映像表示部２６２などを有することができる。

[0122] 映像選択部２６１は、出力する映像を選択する。この選択する手段は特に限定されないが、例えば、映像選択部２６１は、映像ファイルが記録されているアドレスや、広告バナーのコードなどを用いて決定できる。前記アドレスやコードなどは、機械学習分類器１４が選定した環境情報に応じて決定される。また、映像選択部２６１は、複数の映像ファイルを合成したり編集したりしてもよい。さらには、映像選択部２６１は、映像の色温度や輝度などを調整してもよい。

[0123] なお、映像ファイルは映像制御部２６の内部に記録されていてもよいし、映像制御部２６の外部に記録されていてもよい。

[0124] 映像表示部２６２は、映像選択部２６１により決定された情報に基づいて、映像を出力する。

[0125] [（７）音声制御部]

人に対して再生される音声を制御するために、環境制御装置２０は、例えば音声制御部２７を備えることができる。音声制御部２７は、機械学習分類器１４が選定した環境情報に基づいて、再生する音声を制御する。音声制御部２７を備える環境制御装置２０は、例えばスピーカー（いわゆるスマートスピーカーやストリーミング機能付きスピーカーなどを含む）、タブレット、スマートフォン、ヘッドフォン、ウェアラブル装置、カーステレオなどが用いられることにより実現できる。

[0126] 人は、音声制御部２７が再生するその人に最適化された音声を視聴することにより、無自覚にターゲット行動に促される。

[0127] 音声制御部２７の構成について図１８を参照しつつ説明する。図１８は、本実施形態に係る音声制御部２７の構成を示すブロック図である。図１８に示されるとおり、音声制御部２７は、例えば音声選択部２７１及び音声出力部２７２などを有することができる。

[0128] 音声選択部２７１は、再生する音声を選択する。この選択する手段は特に限定されないが、例えば、音声選択部２７１は、音声ファイルが記録されているアドレスや、広告バナーのコードなどを用いて決定できる。前記アドレスやコードなどは、機械学習分類器１４が選定した環境情報に応じて決定される。また、音声選択部２７１は、複数の音声ファイルを合成したり編集したりしてもよい。さらには、音声選択部２７１は、音高や音量などを調整してもよい。

[0129] なお、音声ファイルは音声制御部２７の内部に記録されていてもよいし、音声制御部２７の外部に記録されていてもよい。

[0130] 音声出力部２７２は、音声選択部２７１により決定された情報に基づいて、音声を出力する。

[0131] [３．本技術の第３の実施形態（機械学習システムの例３）]

本技術の一実施形態に係る機械学習装置１０は、ターゲット行動ごとに価値関数Ｑ、状態情報ｓ、及び環境情報の変更ａを記録できる。そして、機械学習装置１０は、人の行動とその人の周囲の環境との相関関係を強化学習することにより、ターゲット行動に促すことができる環境情報を選定できる。

[0132] このとき、行動と環境との相関関係が類似する複数の人は、一つの価値グループに設定されることができる。例えば、香りと温度変化を感じたときに特定のターゲット行動に促されやすい複数の人は、一つのグループとして設定されることができる。

[0133] このことについて図１９を参照しつつ説明する。図１９は、本実施形態に係る機械学習装置１０が用いるデータベースの一例である。図１９に示されるとおり、このデータベースは、ターゲット行動ごとの価値関数Ｑ、状態情報ｓ、及び環境情報の変更ａを保有している。この価値関数Ｑは、状態情報ｓ及び環境情報の変更ａの相関関係の類似の度合いに従って、複数の価値グループＱ１～８に区分けされている。複数の価値グループＱ１～８のそれぞれに対して、その価値グループに所属する人ｈ０１～３２が関連付けられている。それぞれの人に対して、その人の特徴を示す属性情報Ａ～Ｃが関連付けられていてもよい。

[0134] 強化学習の対象者が所属する価値グループがわかるとき、機械学習装置１０は、この価値グループに係る価値関数Ｑなどの情報を用いることができる。既に強化学習済みの価値関数Ｑなどの情報を用いることにより、機械学習装置１０は、例えば強化学習の過程の一部を省略でき、より短期間で強化学習できる。

[0135] 具体例を用いて説明する。ある人に対するターゲット行動が「テレビゲームをする」と設定されているとする。強化学習により、この人は、香りと温度変化に影響を受けたときにテレビゲームをする傾向にあることがわかっているとする。

[0136] 次に、ターゲット行動が「テレビゲームをする」から「ビールを飲む」に変更になったとする。しかし、これまでの強化学習により、この人は香りと

温度変化に影響を受けやすいことがわかっている。このことから、香りと温度変化に影響を受けやすい価値グループの価値関数Ｑ及び過去に高い報酬が得られた環境情報の変更ａなどの情報が、新たなターゲット行動に促すための強化学習の初期値に用いられることができる。機械学習装置１０は、既に強化学習された情報を初期値にして、強化学習を開始することができる。

[0137] また、既に強化学習済みの価値関数Ｑなどの情報は、同じ価値グループに所属する別の人の行動に関する強化学習に用いられてもよい。図１９を参照しつつ説明すると、例えば価値グループＱ４に所属する人ｈ１３の行動に関して強化学習済みの価値関数Ｑなどの情報は、同じ価値グループＱ４に所属する人ｈ１４の行動に関する強化学習に用いられることができる。

[0138] 本実施形態に係る機械学習装置１０の手順について図２０を参照しつつ説明する。図２０は、本実施形態に係る機械学習装置１０の手順の一例を示すフローチャートである。

図２０に示されるとおり、機械学習装置１０が備える状態取得部１１が、状態情報を取得する（ステップＳ２１）。

[0139] 次に、機械学習装置１０が備える評価部１２が、状態情報に基づいて報酬及び価値関数を算出する（ステップＳ２２）。

[0140] 次に、機械学習装置１０が備える機械学習分類器１４が、価値関数を更新する（ステップＳ２３）。

[0141] 次に、さらなる行動変容を学習するために、機械学習分類器１４が、環境情報を選定する（ステップＳ２４）。

[0142] 次に、機械学習分類器１４が、所定の条件を満たすか否かを判定する（ステップＳ２５）。この判定条件は特に限定されないが、例えば価値関数の更新回数が所定の閾値を超過したか否かにより判定されてもよい。

[0143] 所定の条件を満たすとき（ステップＳ２５：Ｙｅｓ）、機械学習分類器１４が、前記データベースを参照して、類似するグループの価値関数Ｑや過去に高い報酬が得られた環境情報の変更ａなどの情報を取得する（ステップＳ２６）。このデータベースは、機械学習装置１０が備えていてもよいし、機

械学習装置１０のほかのコンピュータ装置が備えていてもよい。これにより、機械学習装置１０は、既に強化学習された情報を用いて、強化学習をすることができる。

[0144] 一方で、所定の条件を満たさないとき（ステップＳ２５：Ｎｏ）、類似するグループの価値関数は取得されない。

[0145] 次に、機械学習分類器１４が、強化学習を終了するか否かを判定する（ステップＳ２７）。この判定条件は特に限定されないが、例えば価値関数が所定の閾値より高いか否かにより判定されてもよい。

[0146] 機械学習を終了しないと判定されたとき（ステップＳ２７：Ｎｏ）、ステップＳ２１～２６の手順が繰り返される。

[0147] 機械学習を終了すると判定されたとき（ステップＳ２７：Ｙｅｓ）、機械学習分類器１４が、環境情報を選定する（ステップＳ２８）。

[0148] ［４．本技術の第４の実施形態（機械学習システムの例４）］

ターゲット行動はランダムに設定されてもよい。特定のターゲット行動に固定されず、さまざまなターゲット行動に促すことにより、機械学習分類器１４は、行動と環境との相関関係を強化学習できる。機械学習分類器１４は、この強化学習により、例えば行動との関係性が低いと思われる環境の変化であっても、行動の兆しや連続性などの規則性を発見できる。

[0149] 図１９を再び参照しつつ説明する。このデータベースに登録されている複数の人のそれぞれに対して、その人の特徴を示す属性情報が関連付けられている。例えば属性Ａは「照明の色温度が３６５０Ｋであり、照明の輝度が３０００ｌｍであり、温度が２６．５度であるときに、ビールを飲む傾向がある」という特徴を有する人であってよい。例えば属性Ｃは「温度が２５度であり、湿度が４８％であり、香りに添加成分Ｔが含まれているときに、ＥＣサイトを閲覧する傾向がある」という特徴を有する人であってよい。

[0150] これを実現するために、図２０に示されるフローチャートにおいて、例えば類似するグループの価値関数を取得する（ステップＳ２６）のではなく、ランダムに選定したグループの価値関数を取得してもよい。

[0151] ［５．本技術の第５の実施形態（機械学習システムの例５）］

［（１）概要］

本技術の一実施形態に係る機械学習システム１は、複数の機械学習装置を含んでいてよい。このことについて図２１を参照しつつ説明する。図２１は、本技術の一実施形態に係る機械学習システム１の構成を示すブロック図である。

[0152] 図２１に示されるとおり、機械学習システム１は、例えば複数の機械学習装置１０ａ～１０ｄを備えることができる。複数の機械学習装置１０ａ～１０ｄのそれぞれは、例えば、状態取得部１１ａ～１１ｄ、評価部１２ａ～１２ｄ、記録部１３ａ～１３ｄ、及び機械学習分類器１４ａ～１４ｄなどを備えることができる。また、複数の機械学習装置１０ａ～１０ｄのそれぞれには、環境制御装置（図示省略）が接続されていてよい。なお、機械学習装置の数は特に限定されない。

[0153] さらに、機械学習システム１は、達成難易度算出装置５０を含むことができる。達成難易度算出装置５０は、図１０に示されるようなハードウェア構成でありうる。達成難易度算出装置５０は、例えば情報通信ネットワーク４０などを介して複数の機械学習装置１０ａ～１０ｄのそれぞれと接続されており、複数の機械学習装置１０ａ～１０ｄのそれぞれから得られた情報を集計し、状態情報と環境情報との相関関係の傾向を得ることができる。より具体的に説明すると、達成難易度算出装置５０は、ターゲット行動に関する達成難易度を算出することができる。ターゲット行動に促したことを達成したと定義するとき、達成難易度は、この達成の困難さを示す。この達成難易度についての詳細な説明は後述する。

[0154] 達成難易度算出装置５０は、例えば情報取得部５１と、被験者情報記録部５２と、行動情報記録部５３と、達成難易度算出部５４と、を備えることができる。

[0155] 情報取得部５１は、複数の機械学習装置１０ａ～１０ｄのそれぞれが得た状態情報を取得する。情報取得部５１は、例えば通信インタフェース１０４

などが用いられることにより実現できる。

[0156] 複数の機械学習装置１０ａ～１０ｄのそれぞれが対象とする被験者は異なっていてよい。被験者情報記録部５２は、複数の機械学習装置１０ａ～１０ｄのそれぞれが対象とする被験者に関する情報を保有する。この情報には、例えば被験者の識別番号、性別、又は年齢などが含まれる。被験者情報記録部５２は、例えばストレージ１０２などが用いられることにより実現できる。

[0157] 行動情報記録部５３は、複数の機械学習装置１０ａ～１０ｄのそれぞれが対象とする行動に関する情報を保有する。この情報には、例えばターゲット行動に関する情報、状態情報、又は状態情報に関する履歴情報などが含まれる。行動情報記録部５３は、例えばストレージ１０２などが用いられることにより実現できる。

[0158] 達成難易度算出部５４は、複数の状態取得部１１ａ～１１ｄのそれぞれが取得した状態情報に基づいて、ターゲット行動に関する達成難易度を算出することができる。達成難易度算出部５４は、例えばＣＰＵ１０１及びプログラムなどが用いられることにより実現できる。

[0159] なお、達成難易度算出部５４は、達成難易度算出装置５０が備えていてもよいし、複数の機械学習装置１０ａ～１０ｄのそれぞれが備えていてもよいし、複数の環境制御装置（図示省略）のそれぞれが備えていてもよい。

[0160] また、図示を省略するが、機械学習システム１は、複数の達成難易度算出装置を備えることができる。複数の達成難易度算出装置の中には、複数の機械学習装置のうち特定の機械学習装置から得られた情報を集計する中継用の達成難易度算出装置があってもよい。

[0161] ［（２）達成難易度］

上述したように、達成難易度は、ターゲット行動に促すときの困難さを示す。達成難易度を算出することにより、機械学習システム１は、例えばターゲット行動に促されやすい又は促されにくい被験者を導出したり、ターゲット行動に促されやすい又は促されにくい環境情報を導出したりすることがで

きる。

[0162] ターゲット行動に促されやすい被験者の集団をアダプティブグループと定義し、ターゲット行動に促されにくい被験者の集団をチャレンジグループと定義する。機械学習システム１は、このアダプティブグループを導出することにより、例えば商品の開発や広告宣伝活動などにおいて、商品のターゲット層を導出できる。このターゲット層には、例えば年齢や性別などが含まれる。商品のターゲット層を導出することにより、商品の開発や広告宣伝活動などが効率化される。

[0163] 例えばコンビニエンスストアなどに設けられているＰＯＳ（Point of sale）システムと、地域のイベント（例えば運動会や花火大会など）とが関連付けられることができる。これにより、機械学習システム１は、このイベントの開催時に購入されやすい商品や、この商品のターゲット層などが導出できる。

[0164] あるいは、商品には、爆発的にヒットする商品と、ロングセラー商品とがある。前者の開発や広告宣伝活動などについてはアダプティブグループが活用され、後者の開発や広告宣伝活動などについてはチャレンジグループが活用されることができる。

[0165] さらには、ターゲット行動は、基本となる基本ターゲット行動と、この基本ターゲット行動に関連付けられる応用ターゲット行動と、に分類されることができる。基本ターゲット行動は、行動の種類を大まかに分類したものであり、例えば、外出、飲食、又は購入などである。応用ターゲット行動は、基本ターゲット行動をより具体的に示したものであり、例えば、セールの日に特定の店舗に行く、特定の場所に旅行に行く、地域のイベント活動に参加する、などである。

[0166] 機械学習システム１は、まず、応用ターゲット行動に促されやすいアダプティブグループを導出する。機械学習システム１は、複数の応用ターゲット行動のそれぞれに関するアダプティブグループを導出し、アダプティブグループに関する情報を訴求することにより、基本ターゲット行動に関するアダ

プティブグループを導出することができる。つまり、機械学習システム１は、複数のアダプティブグループに共通する傾向を得ることができる。これにより、例えば商品開発にあたり、今まで気づかなかった新たなターゲット層が導出されることができる。

[0167] なお、このアダプティブグループの導出は、第２の実施形態において説明したような生活習慣の改善などにも活用されてもよい。

[0168] 達成難易度には、例えばターゲット行動に促した割合を示す達成率ｒが含まれていてよい。達成率ｒが高い被験者は、アダプティブグループに分類される。

[0169] 達成率ｒは、ターゲット行動に促した状態情報の件数ｎと、ターゲット行動に促されなかった状態情報を含むすべての状態情報の件数n_{all}と、を用いて、例えば下記の式（３）で表現できる。

[0170] [数3]

$$r = \frac{100n}{n_{all}} \quad \cdots (3)$$

[0171] または、達成難易度には、例えばターゲット行動に促すまでの標準の時間を示す標準達成時間ｓが含まれていてよい。標準達成時間ｓが短い被験者は、アダプティブグループに分類される。

[0172] 標準達成時間ｓは、ターゲット行動に促すまでの時間を示す達成時間ｘと、ターゲット行動に促すまでの平均の時間を示す平均達成時間ｐと、を用いて、例えば下記の式（４）で表現できる。平均達成時間ｐは、達成時間ｘの合計を、すべての状態情報の件数n_{all}
で除算することにより算出できる。

[0173] [数4]

$$s = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - p)^2} \quad \cdots (4)$$

[0174] なお、極端に達成時間が長い被験者の影響を受けないように、標準偏差を

用いて標準達成時間ｓを算出しているが、標準偏差ではなく平均を用いる平均達成時間ｐが達成難易度に含まれてもよい。

[0175] または、達成難易度には、例えばターゲット行動に促したときの環境情報の項目の数の平均を示す平均環境項目数ｑが含まれていてよい。この環境情報の項目には、例えば香り、照明、温度、湿度、映像、又は音声などが含まれる。平均環境項目数ｑが少ない被験者は、アダプティブグループに分類される。例えば、香りにのみ影響される被験者は、香りと温度の両方でなければ影響されない被験者に比べて、ターゲット行動に促されやすいといえる。

[0176] 平均環境項目数ｑは、ターゲット行動に促した状態情報の件数ｎと、ターゲット行動に促したときの環境情報の項目の数ｅと、を用いて、例えば下記の式（５）で表現できる。なお、達成難易度算出部５４は、平均ではなく式（４）のように標準偏差を算出してもよい。

[0177] [数5]

$$q=\frac{1}{n}\sum_{i=1}^{n}e_i \quad \cdots (5)$$

[0178] 平均環境項目数ｑの算出とあわせて、環境情報の項目の名称が記録されてもよい。例えば行動情報記録部５３が、この環境情報の項目の名称を記録することができる。これにより、機械学習システム１は、特定の環境情報に行動変容を促されやすいアダプティブグループを導出することができる。例えば、機械学習システム１は、香りに行動変容を促されやすいアダプティブグループを導出することができる。

[0179] なお、達成難易度には、達成率ｒ、標準達成時間ｓ、及び平均環境項目数ｑのうち少なくとも１つが含まれていればよい。ただし、例えば達成率ｒのみが達成難易度に含まれている場合よりも、達成率ｒ及び標準達成時間ｓの両者が達成難易度に含まれている場合の方が、アダプティブグループの導出がより容易になる。

[0180] 達成難易度について図２２を参照しつつ説明する。図２２は、本実施形態に係る達成難易度算出部５４が算出する達成難易度について説明する図であ

る。図２２に示されるとおり、達成難易度に含まれる達成率（ＡＲ：Achievement rate）、標準達成時間（ＳＡＴ：Standard achievement time）、及び平均環境項目数（ＮＫＶ：Number of key variables）が示されている。

[0181] 図２２Ａは、日常的に運動をしていない被験者に対して、「運動をする」というターゲット行動を促したときの達成難易度を示している。大分類に示されるとおり、このターゲット行動に促される被験者は、年齢が２０～５９歳の男女である。この被験者全体の達成率は３０％、標準達成時間は５４時間、平均環境項目数は２となっている。

[0182] 大分類を細分化したものが中分類である。ここでは例として性別に基づいて細分化されている。男性の被験者の達成率は３１％、標準達成時間は５５時間、平均環境項目数は１となっている。一方で、女性の被験者の達成率は２９％、標準達成時間は５３時間、平均環境項目数は３となっている。このことから、女性よりも男性の被験者の方が、達成率が高いことがわかる。つまり、達成率に着目すると、男性がアダプティブグループに該当するといえる。

[0183] 中分類を細分化したものが小分類である。ここでは例として年齢に基づいて細分化されている。一例を挙げると、年齢が２０～３９歳である男性の被験者グループの達成率は３４％、標準達成時間は３８時間、平均環境項目数は１となっている。この小分類に示されている４つのグループのうち、このグループが、達成率が最も高く、標準達成時間が最も短く、平均環境項目数が最も少ない。つまり、このグループがアダプティブグループであるといえる。機械学習システム１は、このようにアダプティブグループを導出することができる。このアダプティブグループに対して、「運動をする」というターゲット行動に関連する商品やサービスなどに関する販売促進活動が行われることができる。

[0184] 図２２Ｂは、過去５年間のうち一度も選挙に行ったことがない被験者に対して、「選挙に行く」というターゲット行動を促したときの達成難易度を示している。小分類に示されている４つのグループのうち、年齢が５０～７９

歳である女性の被験者グループが、達成率が最も高く、標準達成時間が最も短く、平均環境項目数が最も少ない。つまり、このグループがアダプティブグループであるといえる。

[0185] このように、機械学習システム１は、達成率、標準達成時間、又は平均環境項目数を算出することにより、アダプティブグループを導出することができる。例えば、機械学習システム１は、「ビールを購入する」というターゲット行動に関して、達成率が８０％以上、標準達成時間が３時間以下のアダプティブグループを導出することができる。ビール会社は、このアダプティブグループに対して、新商品の販売開始時に、この新商品の広告宣伝活動を行うことができる。

[0186] または、例えば、機械学習システム１は、「テレビジョン放送や動画配信サービスなどの動画を視聴する」というターゲット行動に関して、達成率が９０％以上、平均環境項目数が２以下のアダプティブグループを導出することができる。動画サブスクリプションサービスの事業者は、このアダプティブグループに対して、動画サブスクリプションサービスの加入に関する広告宣伝活動を行うことができる。さらには、動画サブスクリプションサービスの加入後も、事業者は、このアダプティブグループに対して、契約の継続を促す広告宣伝活動を行うことができる。

[0187] この達成難易度は、ターゲット行動に促すことを繰り返すうちに、その値が変化することがある。このことについて図２３を参照しつつ説明する。図２３は、本実施形態に係る達成難易度算出部５４が算出する達成難易度について説明する図である。

[0188] 図２３Ａは、「商品Ｓを２回以上購入する」というターゲット行動に促すときの達成難易度を示している。なお、このときの達成率については、商品Ｓを２回購入すれば達成率は１００％、商品Ｓを１回購入すれば達成率は５０％としている。

[0189] 図２３Ａでは、達成難易度に基づいて被験者が複数のグループに分類されている。例えば第１のグループＧ１については、被験者の人数が３９６人、

達成率が８６％、標準達成時間が６７時間、平均環境項目数が２となっている。また、第２のグループＧ２については、被験者の人数が２８３人、達成率が６２％、標準達成時間が１２０時間、平均環境項目数が３となっている。この４つのグループのうち、この第１のグループＧ１及び第２のグループＧ２が、比較的に達成率が高く、標準達成時間が短く、平均環境項目数が少ない。つまり、この第１のグループＧ１及び第２のグループＧ２がアダプティブグループであるといえる。商品Ｓの販売会社は、このアダプティブグループに対して、商品Ｓに関するターゲット行動に促すことができる。

[0190] 続いて、機械学習システム１は、この第１のグループＧ１及び第２のグループＧ２に対しては、商品Ｓに関するターゲット行動に促しつつ、比較的にチャレンジグループである第３のグループＧ３及び第４のグループＧ４に対して、別の商品である商品Ｔに関するターゲット行動に促す。

[0191] 図２３Ｂは、「商品Ｔを２回以上購入する」というターゲット行動に促すときの達成難易度を示している。図２３Ｂに示されるとおり、第３のグループＧ３については、達成率が６８％、標準達成時間が２５８時間、平均環境項目数が３となっている。この４つのグループのうち、この第３のグループＧ３が、比較的に達成率が高く、標準達成時間が短く、平均環境項目数が少ない。つまり、この第３のグループＧ３がアダプティブグループであるといえる。

[0192] つまり、商品Ｓのターゲット層を第１のグループＧ１及び第２のグループＧ２とし、商品Ｔのターゲット層を第３のグループＧ３とすることにより、例えば商品の売上げの向上、又は販売促進活動の効率化などが行われることができる。

[0193] ［６．本技術の第６の実施形態（機械学習方法）］

本技術の一実施形態に係る機械学習方法は、コンピュータ装置を利用して人の行動とその人の周囲の環境との相関関係を機械学習する方法である。本実施形態に係る機械学習方法について図２４を参照しつつ説明する。図２４は、本実施形態に係る機械学習方法の手順を示すフローチャートである。図

２４に示されるとおり、本実施形態に係る機械学習方法は、人の行動に関する状態情報を少なくとも取得すること（ステップＳ１）と、前記状態情報を取得したときの前記人の周囲の環境に関する環境情報と、前記状態情報と、を評価して価値関数を得ること（ステップＳ２）と、前記行動をターゲット行動に促すために、前記価値関数を強化学習して、該価値関数が最も高くなるときの前記環境情報を選定すること（ステップＳ３）と、を少なくとも含んでいる。

[0194] 本実施形態に係る機械学習方法は、第１～４の実施形態に係る技術を利用してもよい。よって、この技術についての再度の説明を省略する。

[0195] なお、本明細書中に記載した効果はあくまで例示であって限定されるものではなく、また他の効果があってもよい。

[0196] なお、本技術は、以下のような構成をとることもできる。

［１］

人の行動に関する状態情報を少なくとも取得する状態取得部と、

前記状態情報を取得したときの前記人の周囲の環境に関する環境情報と、前記状態情報と、を評価して価値関数を得る評価部と、

前記行動をターゲット行動に促すために、前記価値関数を強化学習して、該価値関数が最も高くなるときの前記環境情報を選定する機械学習分類器と、を少なくとも備えている、機械学習システム。

［２］

前記評価部が、

前記ターゲット行動に関するターゲット状態情報と、前記状態情報と、の差分に基づいて報酬を算出し、

前記報酬と、前記環境情報と、前記状態情報と、に基づいて、前記価値関数を算出する、

［１］に記載の機械学習システム。

［３］

複数のターゲット行動情報を含むターゲット状態関連情報を保有する、

［１］又は［２］に記載の機械学習システム。

［４］

前記ターゲット状態関連情報には、時間別ターゲット状態情報及び／又は段階別ターゲット状態情報が含まれている、

［３］に記載の機械学習システム。

［５］

前記環境情報には、香り、照明、温度、湿度、映像、又は音声に関する情報が含まれている、

［１］～［４］のいずれか一つに記載の機械学習システム。

［６］

香り制御部をさらに備えており、

前記香り制御部が、前記機械学習分類器が選定した前記環境情報に基づいて、発生する香りを制御する、

［１］～［５］のいずれか一つに記載の機械学習システム。

［７］

香り付着部をさらに備えており、

前記香り付着部が、前記機械学習分類器が選定した前記環境情報に基づいて、香りを物に付着し、

前記機械学習分類器が、前記環境情報に基づいて、前記香り制御部及び前記香り付着部のいずれが香りを発生させるかを判定する、

［６］に記載の機械学習システム。

［８］

照明制御部をさらに備えており、

前記照明制御部が、前記機械学習分類器が選定した前記環境情報に基づいて、照射する光を制御する、

［１］～［７］のいずれか一つに記載の機械学習システム。

［９］

空気調節部をさらに備えており、

前記空気調節部が、前記機械学習分類器が選定した前記環境情報に基づいて、温度及び／又は湿度を制御する、

［１］～［８］のいずれか一つに記載の機械学習システム。

［１０］

映像制御部をさらに備えており、

前記映像制御部が、前記機械学習分類器が選定した前記環境情報に基づいて、表示する映像を制御する、

［１］～［９］のいずれか一つに記載の機械学習システム。

［１１］

音声制御部をさらに備えており、

前記音声制御部が、前記機械学習分類器が選定した前記環境情報に基づいて、再生する音声を制御する、

［１］～［１０］のいずれか一つに記載の機械学習システム。

［１２］

前記価値関数が複数の価値グループに区分けされており、

前記機械学習分類器が、前記複数の価値グループのそれぞれが保有する価値関数を用いる、

［１］～［１１］のいずれか一つに記載の機械学習システム。

［１３］

複数の状態取得部と、

達成難易度算出部と、をさらに備えており、

前記達成難易度算出部が、複数の状態取得部のそれぞれが取得した前記状態情報に基づいて、前記ターゲット行動に関する達成難易度を算出する、

［１］～［１２］のいずれか一つに記載の機械学習システム。

［１４］

前記達成難易度には、前記ターゲット行動に促した割合を示す達成率が含まれる、

［１３］に記載の機械学習システム。

［１５］

前記達成難易度には、前記ターゲット行動に促すまでの標準の時間を示す標準達成時間が含まれる、

［１３］又は［１４］に記載の機械学習システム。

［１６］

前記達成難易度には、前記ターゲット行動に促したときの前記環境情報の項目の数の平均を示す平均環境項目数が含まれる、

［１３］～［１５］のいずれか一つに記載の機械学習システム。

［１７］

人の行動に関する状態情報を少なくとも取得する状態取得部と、

前記状態情報を取得したときの前記人の周囲の環境に関する環境情報と、前記状態情報と、を評価して価値関数を得る評価部と、

前記行動をターゲット行動に促すために、前記価値関数を強化学習して、該価値関数が最も高くなるときの前記環境情報を選定する機械学習分類器と、を少なくとも備えている、機械学習装置。

［１８］

人の行動に関する状態情報を少なくとも取得することと、

前記状態情報を取得したときの前記人の周囲の環境に関する環境情報と、前記状態情報と、を評価して価値関数を得ることと、

前記行動をターゲット行動に促すために、前記価値関数を強化学習して、該価値関数が最も高くなるときの前記環境情報を選定することと、を少なくとも含んでいる、機械学習方法。

符号の説明

[0197] １：機械学習システム

１０：機械学習装置

１１：状態取得部

１２：評価部

１３：記録部

１４：機械学習分類器

２０：環境制御装置

２３：香り制御部

２４：照明制御部

２５：空気調節部

２６：映像制御部

２７：音声制御部

３０：香り付着部

４０：情報通信ネットワーク

５０：達成難易度算出装置

５１：情報取得部

５２：被験者情報記録部

５３：行動情報記録部

５４：達成難易度算出部

Ｓ１：状態情報を少なくとも取得すること

Ｓ２：環境情報と、前記状態情報と、を評価して価値関数を得ること

Ｓ３：環境情報を選定すること

請求の範囲

[請求項1] 人の行動に関する状態情報を少なくとも取得する状態取得部と、

前記状態情報を取得したときの前記人の周囲の環境に関する環境情報と、前記状態情報と、を評価して価値関数を得る評価部と、

前記行動をターゲット行動に促すために、前記価値関数を強化学習して、該価値関数が最も高くなるときの前記環境情報を選定する機械学習分類器と、を少なくとも備えている、機械学習システム。

[請求項2] 前記評価部が、

前記ターゲット行動に関するターゲット状態情報と、前記状態情報と、の差分に基づいて報酬を算出し、

前記報酬と、前記環境情報と、前記状態情報と、に基づいて、前記価値関数を算出する、

請求項１に記載の機械学習システム。

[請求項3] 複数のターゲット行動情報を含むターゲット状態関連情報を保有する、

請求項１に記載の機械学習システム。

[請求項4] 前記ターゲット状態関連情報には、時間別ターゲット状態情報及び／又は段階別ターゲット状態情報が含まれている、

請求項３に記載の機械学習システム。

[請求項5] 前記環境情報には、香り、照明、温度、湿度、映像、又は音声に関する情報が含まれている、

請求項１に記載の機械学習システム。

[請求項6] 香り制御部をさらに備えており、

前記香り制御部が、前記機械学習分類器が選定した前記環境情報に基づいて、発生する香りを制御する、

請求項１に記載の機械学習システム。

[請求項7] 香り付着部をさらに備えており、

前記香り付着部が、前記機械学習分類器が選定した前記環境情報に

基づいて、香りを物に付着し、

前記機械学習分類器が、前記環境情報に基づいて、前記香り制御部及び前記香り付着部のいずれが香りを発生させるかを判定する、

請求項６に記載の機械学習システム。

[請求項8] 照明制御部をさらに備えており、

前記照明制御部が、前記機械学習分類器が選定した前記環境情報に基づいて、照射する光を制御する、

請求項１に記載の機械学習システム。

[請求項9] 空気調節部をさらに備えており、

前記空気調節部が、前記機械学習分類器が選定した前記環境情報に基づいて、温度及び／又は湿度を制御する、

請求項１に記載の機械学習システム。

[請求項10] 映像制御部をさらに備えており、

前記映像制御部が、前記機械学習分類器が選定した前記環境情報に基づいて、表示する映像を制御する、

請求項１に記載の機械学習システム。

[請求項11] 音声制御部をさらに備えており、

前記音声制御部が、前記機械学習分類器が選定した前記環境情報に基づいて、再生する音声を制御する、

請求項１に記載の機械学習システム。

[請求項12] 前記価値関数が複数の価値グループに区分けされており、

前記機械学習分類器が、前記複数の価値グループのそれぞれが保有する前記価値関数を用いる、

請求項１に記載の機械学習システム。

[請求項13] 複数の状態取得部と、

達成難易度算出部と、をさらに備えており、

前記達成難易度算出部が、複数の状態取得部のそれぞれが取得した前記状態情報に基づいて、前記ターゲット行動に関する達成難易度を

算出する、

請求項１に記載の機械学習システム。

[請求項14] 前記達成難易度には、前記ターゲット行動に促した割合を示す達成率が含まれる、

請求項１３に記載の機械学習システム。

[請求項15] 前記達成難易度には、前記ターゲット行動に促すまでの標準の時間を示す標準達成時間が含まれる、

請求項１３に記載の機械学習システム。

[請求項16] 前記達成難易度には、前記ターゲット行動に促したときの前記環境情報の項目の数の平均を示す平均環境項目数が含まれる、

請求項１３に記載の機械学習システム。

[請求項17] 人の行動に関する状態情報を少なくとも取得する状態取得部と、

前記状態情報を取得したときの前記人の周囲の環境に関する環境情報と、前記状態情報と、を評価して価値関数を得る評価部と、

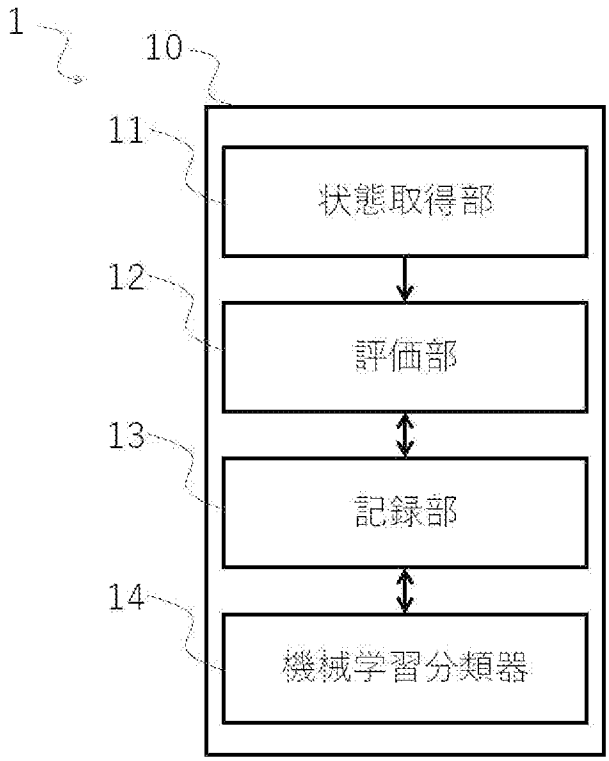
前記行動をターゲット行動に促すために、前記価値関数を強化学習して、該価値関数が最も高くなるときの前記環境情報を選定する機械学習分類器と、を少なくとも備えている、機械学習装置。

[請求項18] 人の行動に関する状態情報を少なくとも取得することと、

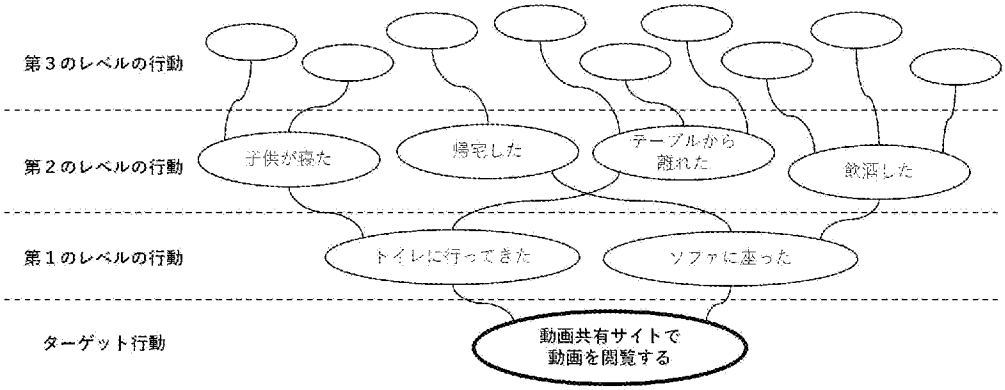
前記状態情報を取得したときの前記人の周囲の環境に関する環境情報と、前記状態情報と、を評価して価値関数を得ることと、

前記行動をターゲット行動に促すために、前記価値関数を強化学習して、該価値関数が最も高くなるときの前記環境情報を選定することと、を少なくとも含んでいる、機械学習方法。

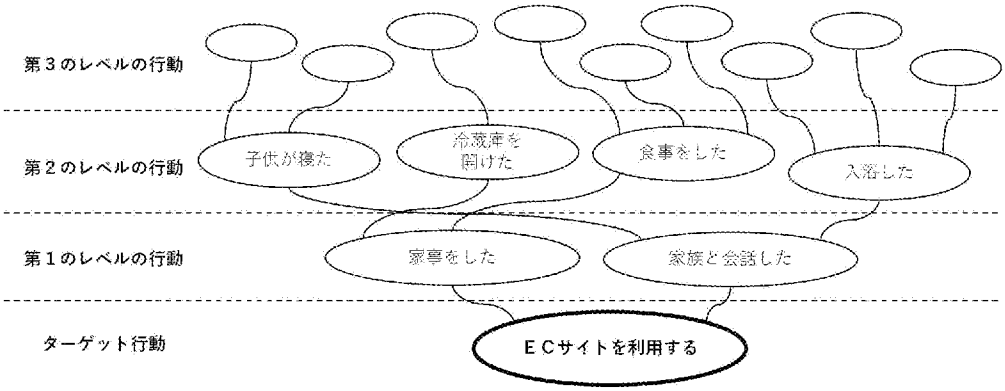
[図1]


1
10
11
状態取得部
12
評価部
13
記録部
14
機械学習分類器

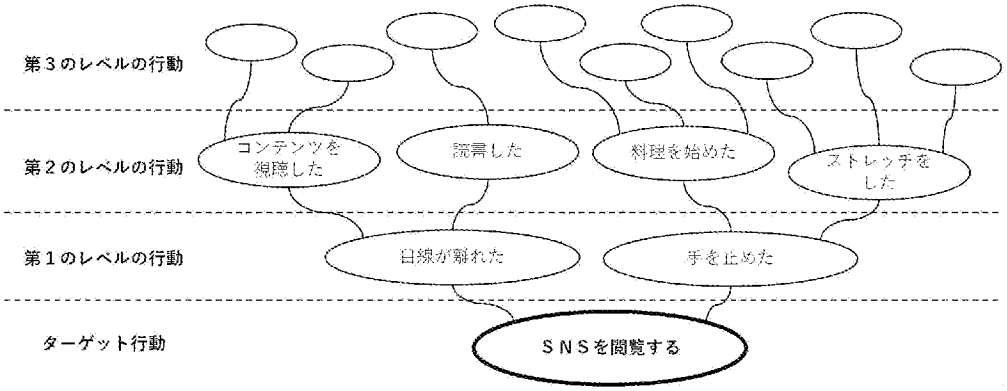
[図2]


第３のレベルの行動
第２のレベルの行動
子供が寝た
帰宅した
テーブルから
離れた
飲酒した
第１のレベルの行動
トイレに行ってきた
ソファに座った
ターゲット行動
動画共有サイトで
動画を閲覧する

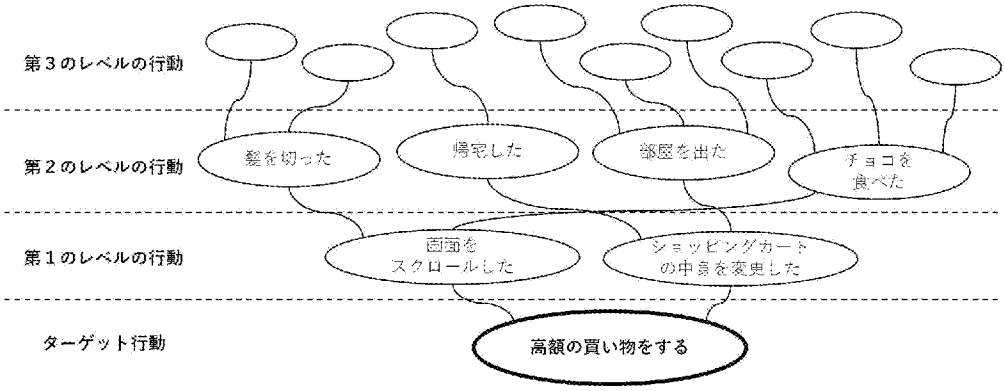
[図3]


第３のレベルの行動
第２のレベルの行動
子供が寝た
冷蔵庫を
開けた
食事をした
入浴した
第１のレベルの行動
家事をした
家族と会話した
ターゲット行動
ＥＣサイトを利用する

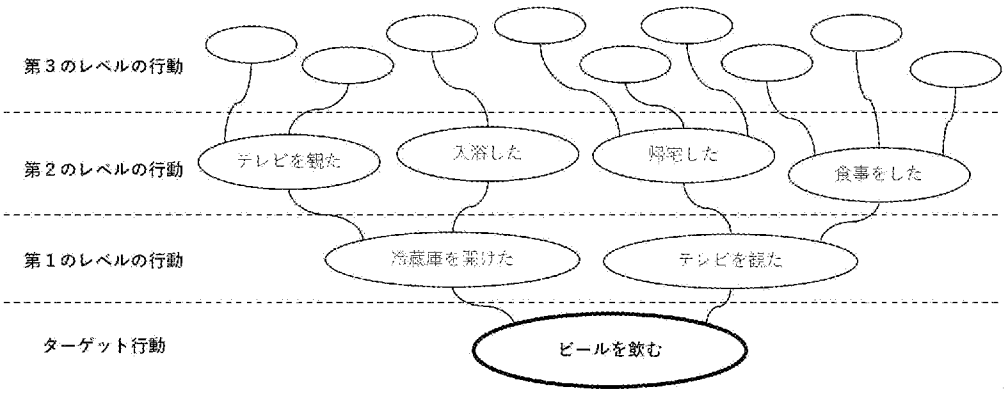
[図4]


第３のレベルの行動
第２のレベルの行動
コンテンツを視聴した
読書した
料理を始めた
ストレッチをした
第１のレベルの行動
目線が離れた
手を止めた
ターゲット行動
ＳＮＳを閲覧する

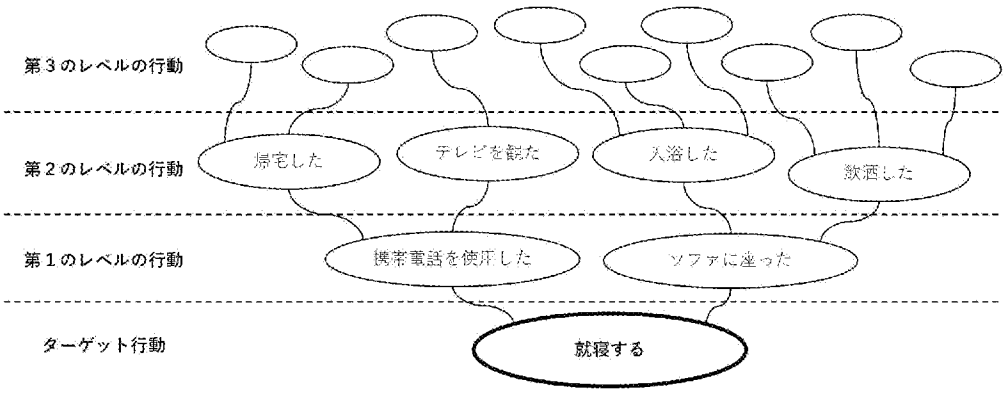
[図5]


第３のレベルの行動
第２のレベルの行動
髪を切った
帰宅した
部屋を出た
チョコを食べた
第１のレベルの行動
画面をスクロールした
ショッピングカートの中身を変更した
ターゲット行動
高額の買い物をする

[図6]

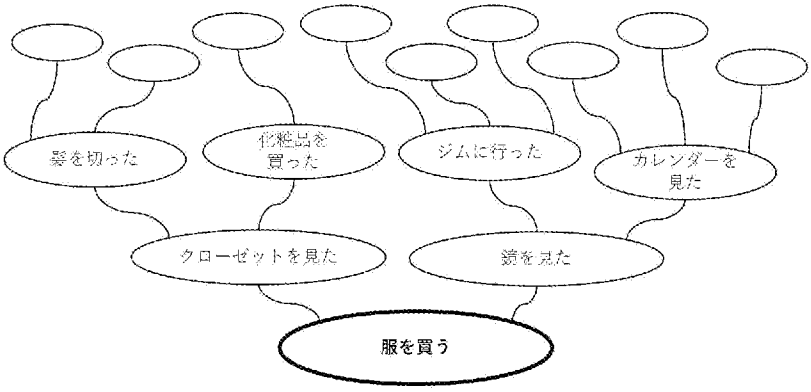

第３のレベルの行動
第２のレベルの行動
テレビを観た
入浴した
帰宅した
食事をした
第１のレベルの行動
冷蔵庫を開けた
テレビを観た
ターゲット行動
ビールを飲む

[図7]

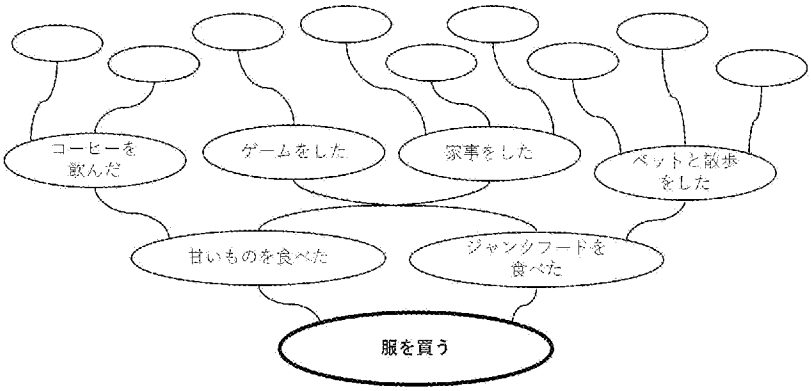

第３のレベルの行動
第２のレベルの行動
帰宅した
テレビを観た
入浴した
飲酒した
第１のレベルの行動
携帯電話を使用した
ソファに座った
ターゲット行動
就寝する

[図8]

A



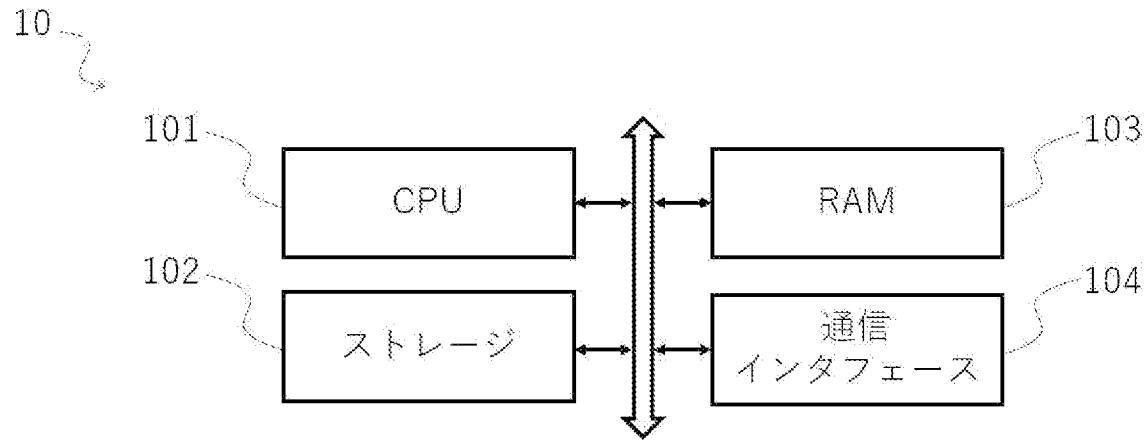
B



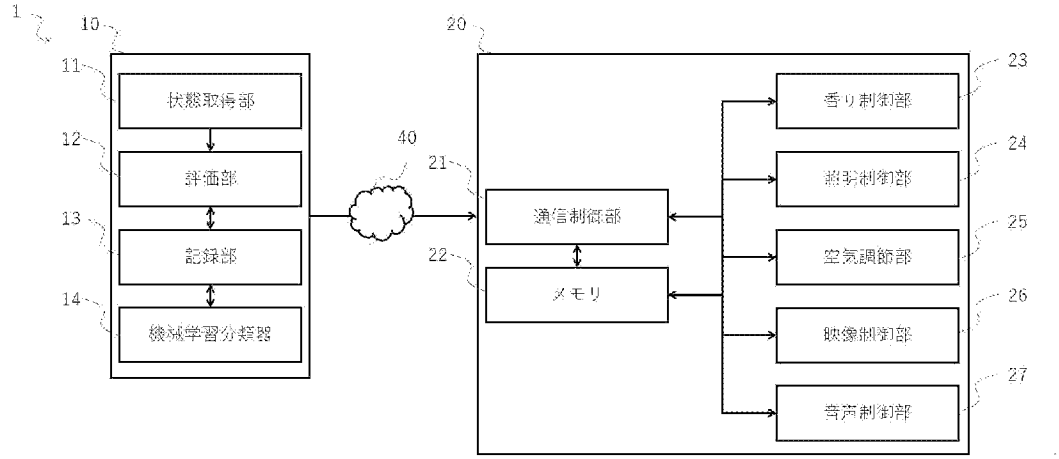
[図9]

時間帯	ターゲット行動
第１の時間帯（午前０～６時台）	就寝する
第２の時間帯（午前７～午後７時台）	食品Ｓを食べる
第３の時間帯（午後８～１１時台）	飲料Ｔを飲む

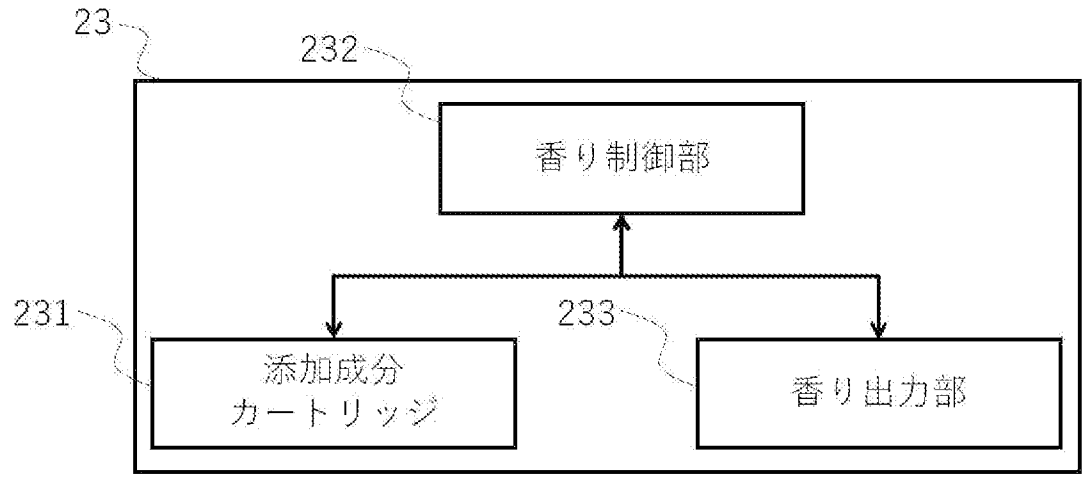
[図10]



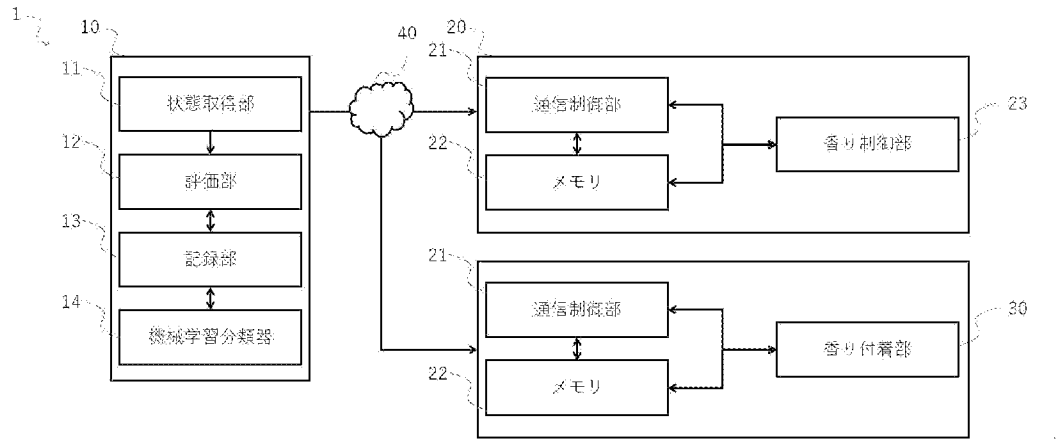
[図11]


1
10
11
状態取得部
12
評価部
13
記録部
14
機械学習分類器
40
20
21
通信制御部
22
メモリ
23
香り制御部
24
照明制御部
25
空気調節部
26
映像制御部
27
音声制御部

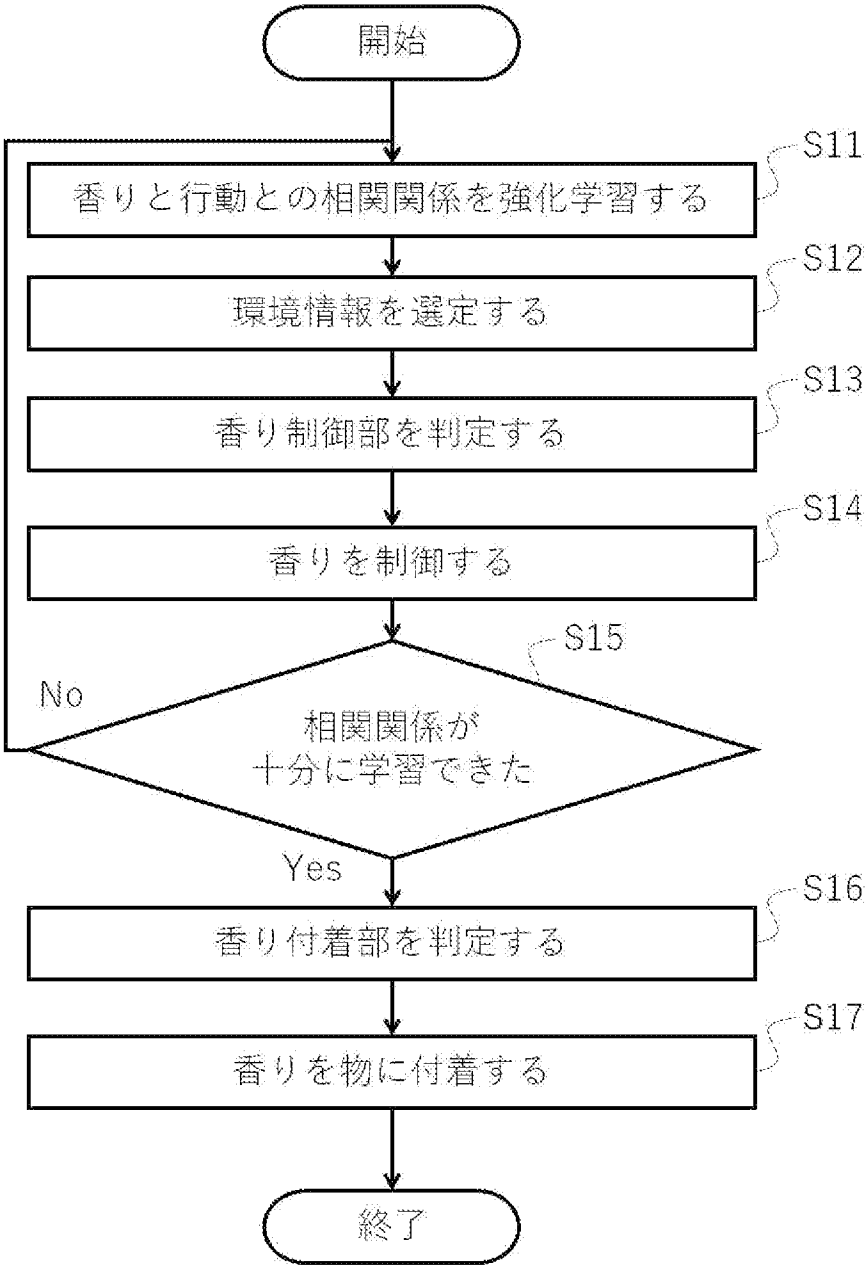
[図12]


23
232
香り制御部
231
添加成分
カートリッジ
233
香り出力部

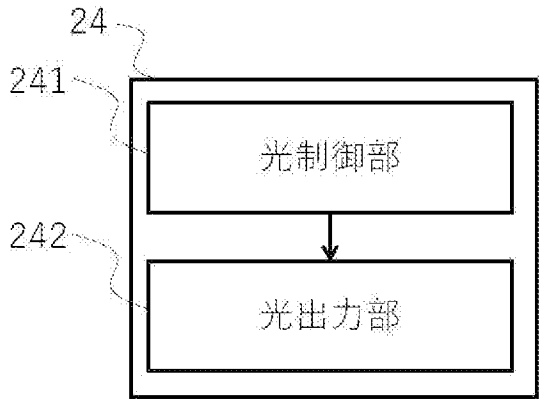
[図13]


1
10
11
状態取得部
12
評価部
13
記録部
14
機械学習分類器
40
20
21
通信制御部
22
メモリ
23
香り制御部
21
通信制御部
22
メモリ
30
香り付着部

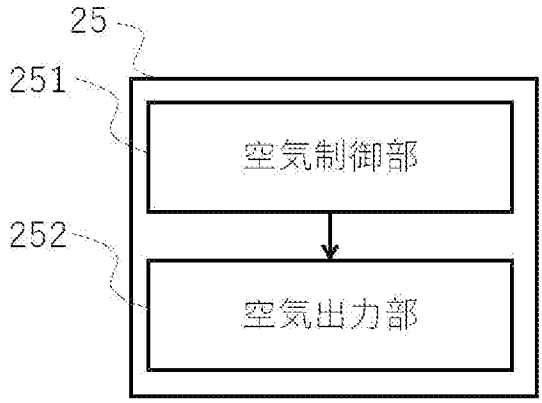
[図14]


開始
S11
香りと行動との相関関係を強化学習する
S12
環境情報を選定する
S13
香り制御部を判定する
S14
香りを制御する
S15
No
相関関係が
十分に学習できた
Yes
S16
香り付着部を判定する
S17
香りを物に付着する
終了

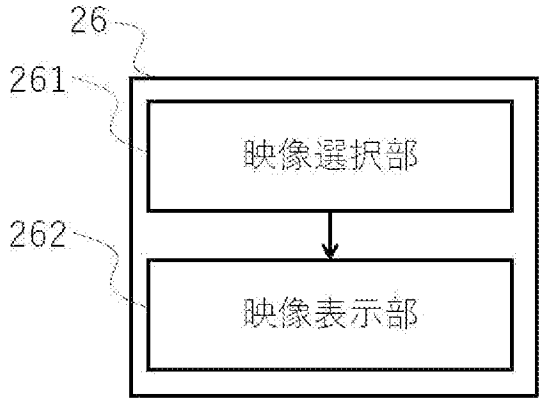
[図15]


24
241
光制御部
242
光出力部

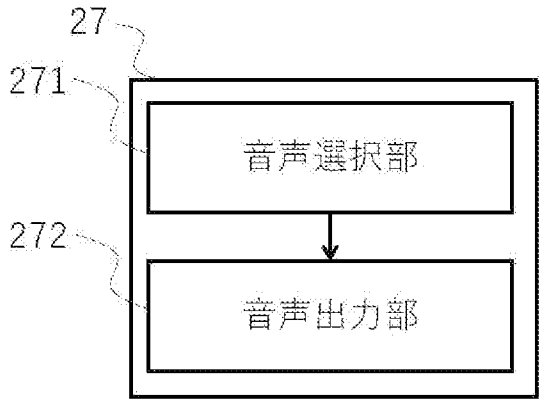
[図16]



[図17]



[図18]

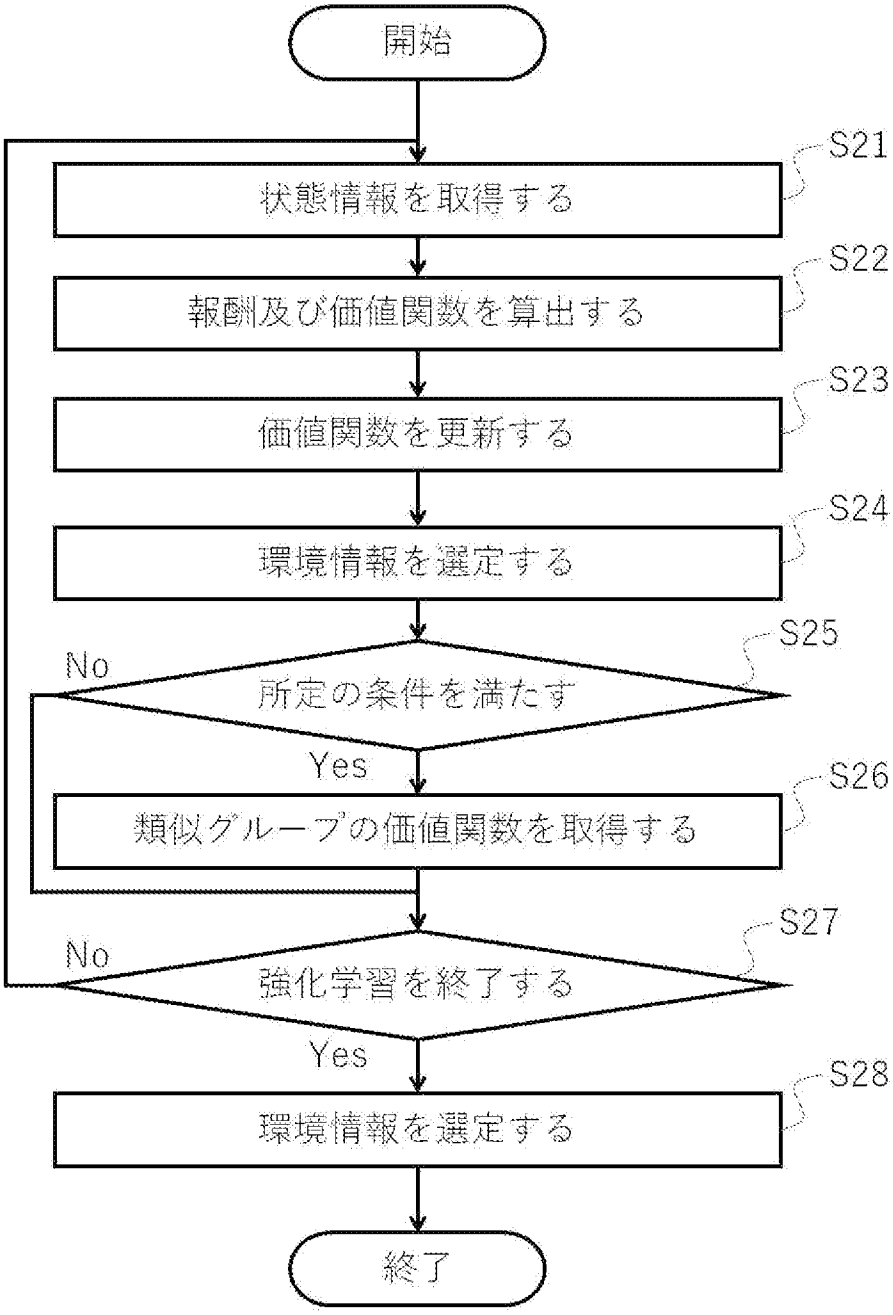


[図19]

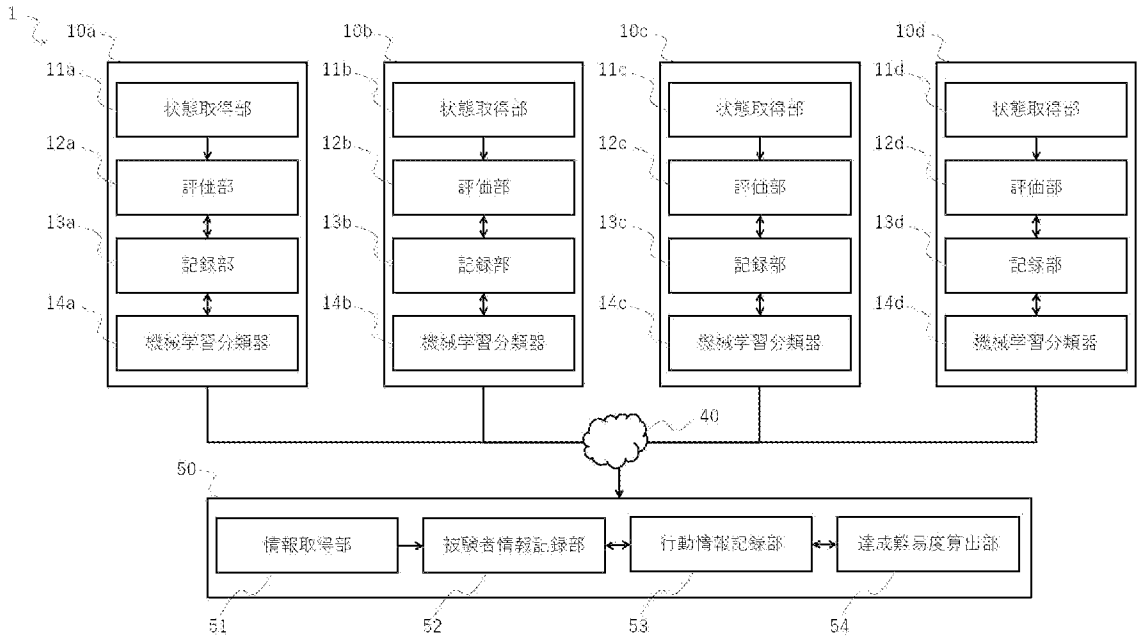
ターゲット行動	第１のターゲット行動															
価値グループ	Q1				Q2				Q3				Q4			
状態情報	s1				s2				s3				s4			
環境所法	a1				a2				a3				a4			
属性	A	B	A	A	C	B	C	A	C	B	C	A	C	B	C	A
人	h01	h02	h03	h04	h05	h06	h07	h08	h09	h10	h11	h12	h13	h14	h15	h16

ターゲット行動	第２のターゲット行動															
価値グループ	Q5				Q6				Q7				Q8			
状態情報	s5				s6				s7				s8			
環境所法	a5				a6				a7				a8			
属性	A	B	A	A	C	B	C	A	C	B	C	A	C	B	C	A
人	h17	h18	h19	h20	h21	h22	h23	h24	h25	h26	h27	h28	h29	h30	h31	h32

[図20]


開始
S21
状態情報を取得する
S22
報酬及び価値関数を算出する
S23
価値関数を更新する
S24
環境情報を選定する
S25
No
所定の条件を満たす
Yes
S26
類似グループの価値関数を取得する
S27
No
強化学習を終了する
Yes
S28
環境情報を選定する
終了

[図21]


1
10a
10b
10c
10d
11a
11b
11c
11d
状態取得部
12a
12b
12c
12d
評価部
13a
13b
13c
13d
記録部
14a
14b
14c
14d
機械学習分類器
40
50
情報取得部
被験者情報記録部
行動情報記録部
達成難易度算出部
51
52
53
54

[図22]

A

大分類	中分類	小分類
性別：男性及び女性 年齢：20～59歳 AR　：30% SAT　：54時間 NKV　：2	性別：男性 年齢：20～59歳 AR　：31% SAT　：55時間 NKV　：1	性別：男性 年齢：20～39歳 AR　：34% SAT　：38時間 NKV　：1
		性別：男性 年齢：40～59歳 AR　：28% SAT　：72時間 NKV　：1
	性別：女性 年齢：20～59歳 AR　：29% SAT　：53時間 NKV　：3	性別：女性 年齢：20～39歳 AR　：31% SAT　：45時間 NKV　：2
		性別：女性 年齢：40～59歳 AR　：27% SAT　：61時間 NKV　：4

B

大分類	中分類	小分類
性別：男性及び女性 年齢：20～79歳 AR　：17% SAT　：275.75時間 NKV　：1.5	性別：男性 年齢：20～79歳 AR　：9.5% SAT　：297.5時間 NKV　：1.5	性別：男性 年齢：20～49歳 AR　：7% SAT　：336時間 NKV　：2
		性別：男性 年齢：50～79歳 AR　：12% SAT　：259時間 NKV　：1
	性別：女性 年齢：20～79歳 AR　：24.5% SAT　：254時間 NKV　：1.5	性別：女性 年齢：20～49歳 AR　：10% SAT　：321時間 NKV　：2
		性別：女性 年齢：50～79歳 AR　：39% SAT　：187時間 NKV　：1

[図23]

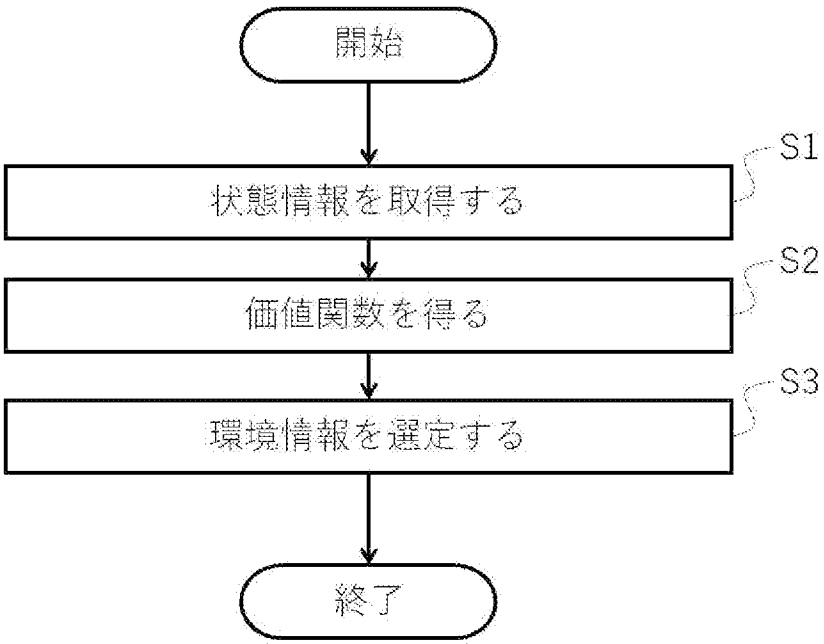
A

[グループG1] 人数：396人 AR　：86% SAT　：67時間 NKV　：2
[グループG2] 人数：283人 AR　：62% SAT　：120時間 NKV　：3
[グループG3] 人数：186人 AR　：38% SAT　：389時間 NKV　：4
[グループG4] 人数：135人 AR　：12% SAT　：512時間 NKV　：5

B

[グループG1] 人数：396人 AR　：36% SAT　：467時間 NKV　：5
[グループG2] 人数：283人 AR　：45% SAT　：240時間 NKV　：3
[グループG3] 人数：186人 AR　：68% SAT　：258時間 NKV　：3
[グループG4] 人数：135人 AR　：14% SAT　：508時間 NKV　：5

[図24]



INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2021/001234

A. CLASSIFICATION OF SUBJECT MATTER
Int. Cl. G06N20/00(2019.01)i
FI: G06N20/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
Int. Cl. G06N3/00-99/00

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
Published examined utility model applications of Japan 1922-1996
Published unexamined utility model applications of Japan 1971-2021
Registered utility model specifications of Japan 1996-2021
Published registered utility model applications of Japan 1994-2021

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y A	WO 2020/027174 A1 (DAINIPPON PRINTING CO., LTD.) 06 February 2020, paragraphs [0032]-[0057], [0093]-[0102], [0112], [0134]-[0157], fig. 1, 2, 10, 15, 16	1-3, 5-6, 8-18 4, 7
Y A	JP 2019-28899 A (ASCON CO., LTD.) 21 February 2019, paragraphs [0018]-[0026], fig. 1	1-3, 5-6, 8-18 4, 7
Y A	WO 2019/087854 A1 (PANASONIC INTELLECTUAL PROPERTY MANAGEMENT CO., LTD.) 09 May 2019, paragraphs [0037]-[0042], [0100], [0101], fig. 3	1-3, 5-6, 8-18 4, 7

☒ Further documents are listed in the continuation of Box C.
☒ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
25.03.2021	06.04.2021

Name and mailing address of the ISA/	Authorized officer
Japan Patent Office 3-4-3, Kasumigaseki, Chiyoda-ku, Tokyo 100-8915, Japan	Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2021/001234

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	山崎惇広,外１名．逐次的に分割された問題空間における複素強化学習．電子情報通信学会論文誌．01 May 2011, vol. J94-D, no. 5, pp. 872-880, in particular, "5. Simulation Experiment", (YAMAZAKI, Atsuhiro et al. Complex-valued reinforcement learning in sequentially-partitioned problem space. IEICE Transactions on Information and Systems.)	13-16

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/JP2021/001234

Patent Documents referred to in the Report	Publication Date	Patent Family	Publication Date
WO 2020/027174 A1	06.02.2020	(Family: none)	
JP 2019-28899 A	21.02.2019	(Family: none)	
WO 2019/087854 A1	09.05.2019	CN 111316313 A	

国際調査報告

国際出願番号
PCT/JP2021/001234

A. 発明の属する分野の分類（国際特許分類（ＩＰＣ））

G06N 20/00(2019.01)i
FI: G06N20/00

B. 調査を行った分野

調査を行った最小限資料（国際特許分類（ＩＰＣ））

G06N3/00-99/00

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	１９２２－１９９６年
日本国公開実用新案公報	１９７１－２０２１年
日本国実用新案登録公報	１９９６－２０２１年
日本国登録実用新案公報	１９９４－２０２１年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

C. 関連すると認められる文献

引用文献のカテゴリー*	引用文献名　及び一部の箇所が関連するときは、その関連する箇所の表示	関連する請求項の番号
Y	WO 2020/027174 A1 (大日本印刷株式会社) 06.02.2020 (2020-02-06) [0032]-[0057], [0093]-[0102], [0112], [0134]-[0157], 図1-2, 10, 15-16	1-3,5-6,8-18
A		4,7
Y	JP 2019-28899 A (株式会社アスコン) 21.02.2019 (2019-02-21) ［００１８］－［００２６］，図１	1-3,5-6,8-18
A		4,7
Y	WO 2019/087854 A1 (パナソニックＩＰマネジメント株式会社) 09.05.2019 (2019-05-09) [0037]-[0042], [0100]-[0101], 図3	1-3,5-6,8-18
A		4,7
Y	山崎惇広、外１名，逐次的に分割された問題空間における複素強化学習，電子情報通信学会論文誌，2011.05.01，Vol.J94-D No.5，pp.872-880 特に”5. シミュレーション実験”	13-16

☐ Ｃ欄の続きにも文献が列挙されている。

☑ パテントファミリーに関する別紙を参照。

＊ 引用文献のカテゴリー
“A” 特に関連のある文献ではなく、一般的技術水準を示すもの
“E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
“L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）
“O” 口頭による開示、使用、展示等に言及する文献
“P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献
“T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの
“X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
“Y” 特に関連のある文献であって、当該文献と他の１以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
“&” 同一パテントファミリー文献

国際調査を完了した日	国際調査報告の発送日
25.03.2021	06.04.2021

名称及びあて先

日本国特許庁(ISA/JP)
〒100-8915
日本国
東京都千代田区霞が関三丁目４番３号

権限のある職員（特許庁審査官）

杉浦　孝光 5B 5287

電話番号 03-3581-1101 内線 3545

国際調査報告
パテントファミリーに関する情報

国際出願番号
PCT/JP2021/001234

引用文献	公表日	パテントファミリー文献	公表日
WO 2020/027174 A1	06.02.2020	（ファミリーなし）	
JP 2019-28899 A	21.02.2019	（ファミリーなし）	
WO 2019/087854 A1	09.05.2019	CN 111316313 A	