(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
2024年8月22日(22.08.2024)

WIPO | PCT

(10) 国際公開番号
WO 2024/171312 A1

(51) 国際特許分類:
G06Q 10/10 (2023.01)

(21) 国際出願番号： PCT/JP2023/005036

(22) 国際出願日： 2023年2月14日(14.02.2023)

(25) 国際出願の言語： 日本語

(26) 国際公開の言語： 日本語

(71) 出願人: 日本電信電話株式会社 (**NIPPON TELEGRAPH AND TELEPHONE CORPORATION**) [JP/JP]; 〒1008116 東京都千代田区大手町一丁目５番１号 Tokyo (JP).

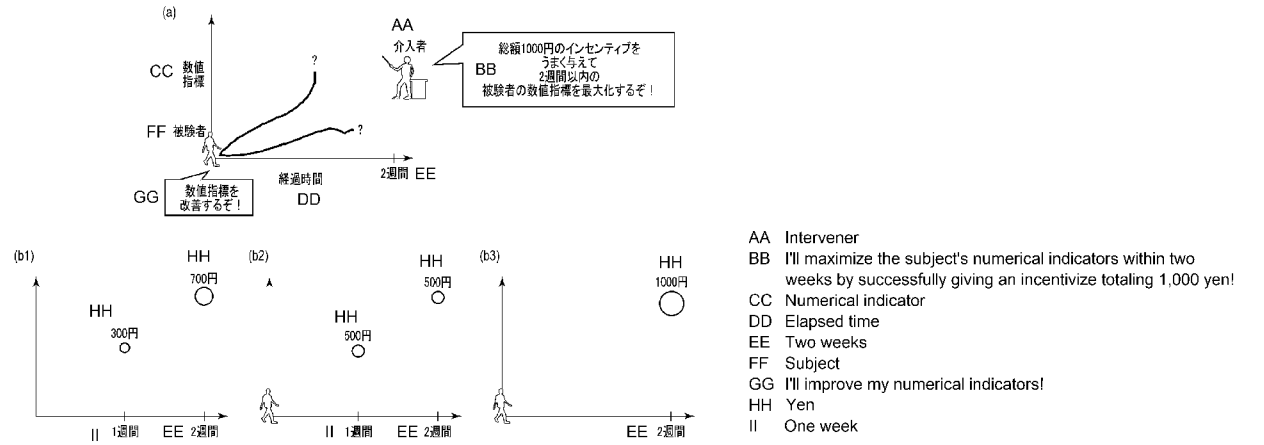
(72) 発明者: 赤木 康紀 (**AKAGI, Yasunori**); 〒1808585 東京都武蔵野市緑町３丁目９－１１ ＮＴＴ 知的財産センタ内 Tokyo (JP). 丸茂 直貴(**MARUMO, Naoki**); 〒1808585 東京都武蔵野市緑町３丁目９－１１ ＮＴＴ 知的財産センタ内 Tokyo (JP). 倉島 健(**KURASHIMA, Takeshi**); 〒1808585 東京都武蔵野市緑町３丁目９－１１ ＮＴＴ 知的財産センタ内 Tokyo (JP).

(74) 代理人: 弁理士法人鈴榮特許綜合事務所 (**SUZUYE & SUZUYE**); 〒1050014 東京都港区芝三丁目２３番１号 セレスティン芝三井ビルディング１１階 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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, 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,

(54) **Title:** INFORMATION PROCESSING DEVICE, INFORMATION PROCESSING METHOD, AND INFORMATION PROCESSING PROGRAM

(54) 発明の名称： 情報処理装置、情報処理方法、および情報処理プログラム



(57) **Abstract:** An information processing device according to one embodiment comprises: an input information acquisition unit that acquires input information including a total budget for remuneration, a number of time steps until a target time, and a present bias strength; a calculation unit that calculates a prescribed threshold value greater than 1/2 on the basis of the total budget for remuneration, the number of time steps, and the present bias strength; an assessment unit that assesses whether the present bias strength is greater than the prescribed threshold value and, if the present bias strength is less than the prescribed threshold value, assesses whether the present bias strength is greater than 1/2; a remuneration schedule calculation unit that calculates a remuneration schedule corresponding to the result of the assessment and including a number of times remuneration is to be given, a time remuneration is to be given, a remuneration to be given, and a required achievement value in terms of time; and an output control unit that performs control so as to output the remuneration schedule.

WO 2024/171312 A1

［続葉有］

SL, 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, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, 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, ME, 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:一実施形態に係る情報処理装置は、報酬の総予算、目標となる時刻までのタイムステップ数、現在のバイアスの強さを含む入力情報を取得する入力情報取得部と、報酬の総予算、タイムステップ数、および現在のバイアスの強さに基づいて１／２よりも大きい所定の閾値を算出する算出部と、現在のバイアスの強さが所定の閾値より大きいかどうかを判定し、現在のバイアスの強さが所定の閾値よりも小さい場合、現在のバイアスの強さが１／２よりも大きいかどうか判定する判定部と、判定の結果に対応する、報酬を与える回数、報酬を与える時刻、与えるべき報酬、および時刻に必要な達成値を含む報酬スケジュールを算出する報酬スケジュール算出部と、報酬スケジュールを出力するように制御する出力制御部と、を備える。

明　細　書

発明の名称 ：

情報処理装置、情報処理方法、および情報処理プログラム

技術分野

[0001] この発明は、情報処理装置、情報処理方法、および情報処理プログラムに関する。

背景技術

[0002] オンライン授業における授業コースの完遂、一定期間における歩数目標の達成、等、所定の目標を達成しようとする人間の行動をモデリングすることは重要である。特に、現在の利益または損失を重く評価し、将来の利益または損失を軽く評価する「現在バイアス」は、目標達成に大きな影響を与えることが知られている。例えば、現在バイアスを考慮したモデリングに基づき最適化された介入を行うことにより、効果的な介入を実現することができる。

[0003] 例えば、非特許文献１では、このようなモデリングの手段の一つであるグラフ理論に基づくモデルが提案されている。このモデルにおいて、人間が取り得る状態が頂点として表現され、各状態において人間が取り得る行動が辺として表現される。また、各辺にはコストが設定されており、これは各行動を取ることの労力を表現している。そして、各頂点には報酬が置かれており、これは、対応する状態に到達することへの報酬を表現している。エージェントは、このグラフ上において、自身が今後取り得る行動の列（グラフ上でのパス）の候補について自身の利得を評価し、利得が最も大きい行動を取ることを選択する。行動の列の利得は、準双曲割引と呼ばれる現在バイアスによって将来のコストを小さく、直近のコストを大きく重み付けして計算される。

[0004] このモデルは、非合理性を含んだ人間の行動をうまく説明することができるものとして注目されており、例えば、非特許文献２および非特許文献３の

様に別のバイアスを含んだモデルなどにも拡張されている。

先行技術文献

非特許文献

[0005] 非特許文献1：Jon Kleinberg and Sigal Oren. Time-inconsistent planning : a computational problem in behavioral economics. In Proceedings of the 15th ACM Conference on Economics and Computation, pages 547-564, 2014.

非特許文献2：Jon Kleinberg, Sigal Oren, and Manish Raghavan. Planning problems for sophisticated agents with present bias. In Proceedings of the 17th ACM Conference on Economics and Computation, pages 343-360, 2016.

非特許文献3：Jon Kleinberg, Sigal Oren, and Manish Raghavan. Planning with multiple biases. In Proceedings of the 18th ACM Conference on Economics and Computation, pages 567-584, 2017.

発明の概要

発明が解決しようとする課題

[0006] しかしながら、上記したモデルにおいて、エージェントの挙動についてシミュレートすることはできるが、全体の予算が決まった報酬をどのように配置するのが最適であるかに関しては答えられないという問題がある。

[0007] この発明は上記事情に着目してなされたもので、その目的とするところは、非特許文献１に開示されたモデルをよりシンプルにしたモデルを作成し、当該モデルにおいて報酬のスケジューリング（全体の予算が決まった報酬をどのタイミングでどの程度の額与えるか）を最適化することにより、被験者（エージェント）である個人のパフォーマンスを最大化す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

[0008] 上記課題を解決するためにこの発明の一態様は、情報処理装置であって、

報酬の総予算、目標となる時刻までのタイムステップ数、現在のバイアスの強さを含む入力情報を取得する入力情報取得部と、前記報酬の総予算、前記タイムステップ数、および前記現在のバイアスの強さに基づいて１／２よりも大きい所定の閾値を算出する算出部と、前記現在のバイアスの強さが前記所定の閾値より大きいかどうかを判定し、前記現在のバイアスの強さが前記所定の閾値よりも小さい場合、前記現在のバイアスの強さが１／２よりも大きいかどうか判定する判定部と、前記判定の結果に対応する、報酬を与える回数、報酬を与える時刻、与えるべき報酬、および前記時刻に必要な達成値を含む報酬スケジュールを算出する報酬スケジュール算出部と、前記報酬スケジュールを出力するように制御する出力制御部と、を備えるようにしたものである。

発明の効果

[0009] この発明の一態様によれば、現在バイアスの影響下における報酬の与え方のスケジュールを最適化することにより、個人のタスクにおけるパフォーマンスを最大化することができる。

図面の簡単な説明

[0010] [図1]図１は、本実施形態のイメージを示した図である。

[図2]図２は、本実施形態に係る情報処理装置のハードウェア構成の一例を示すブロック図である。

[図3]図３は、本実施形態における情報処理装置のソフトウェア構成を、図２に示したハードウェア構成に関連付けて示すブロック図である。

[図4]図４は、情報処理装置が報酬スケジュールを最適化し、最適化した報酬スケジュールを出力するための動作の一例を示すフローチャートである。

[図5]図５は、入力情報の一例を示した図である。

[図6]図６は、最適化スケジュールを出力制御部が表示する際の一例を示している。

[図7]図７は、最適化スケジュールを出力制御部が表示する際の一例を示している。

発明を実施するための形態

[0011] 以下、図面を参照してこの発明に係る実施形態を説明する。なお、以降、説明済みの要素と同一または類似の要素には同一または類似の符号を付し、重複する説明については基本的に省略する。例えば、複数の同一または類似の要素が存在する場合に、各要素を区別せずに説明するために共通の符号を用いることがあるし、各要素を区別して説明するために当該共通の符号に加えて枝番号を用いることもある。

[0012] ［実施形態］

（準備）

最初に本実施形態におけるモデルの詳細を説明する。

頂点集合および辺集合を

$$V=\{(i,\ x)\mid i\in\{0,\ldots,N\},\ x\geqq 0\},$$

$$E=\{((l,\ x),\ (i+1,\ y))\mid i\in\{0,\ldots,N-1\},\ x,\ y\geqq 0\}$$

とする。ここで、ｉは、離散値を取る一方、ｘは、連続値を取る。辺（（ｉ，ｘ），（ｉ＋１，ｙ））のコストは、$(y-x)^2$と記載することができる。頂点（ｉ，ｘ）で得られる報酬をｒ（ｉ，ｘ）と記載する。

[0013] 頂点（ｉ，ｘ）において、ｉは、時刻を、ｘは、タスクの進行度合いを表す数値指標に対応する。例えば、「３０日で１００００００歩（１０万歩）歩く」というタスクを考えた場合、ｉは、現在何日目か、ｘは、現在までに何歩歩いたかを表す。Ｎは、時刻の最大値、すなわち目標となる時刻までのタイムステップ数であり、先の例では「３０日」に対応する。このようにグラフの形を制限することにより、以後の解析的な取り扱いが可能になっている点が非特許文献１のモデルと異なる点である。

[0014] ａ≧０，Ｒ＞０とする、報酬を

$$r(N,\ x)=R1\ [x\geqq a]$$

と定義する。ｉ＜Ｎに対して、

$$r(i,\ x)=0$$

とする。これは、「エージェントが時刻Ｎまでにａ以上のタスクの進行度合いを達成した場合に報酬Ｒを与えること」に対応する。すなわち、ａがエージェント（被験者）である個人の目標であり、Ｒは、報酬に対応する。

[0015] 以下において、ｘ≦ａとする。頂点（ｉ，ｘ）から出発する場合の客観コストｃ（ｉ，ｘ）は、

[0016] [数1]

$$c(i,x) \coloneqq \min\left\{0,\ \min_{y_{i-1},\dots,y_N}\left\{\sum_{j=i+1}^{N}(y_j-y_{j-1})^2 \;\middle|\; y_N \geq a\right\}\right\}$$
$$= \min\left\{0,\ (N-i)\left(\frac{a-x}{N-i}\right)^2 - R\right\}$$

[0017] により定まる。

実際に被験者によって辿られる（主観最適なパス）を（０，z_0），（１，z_1），．．．，（Ｎ，z_N）とすると、（z_i）は、

[0018] [数2]

$$z_0 = 0,$$
$$z_i = \operatorname*{argmin}_{x\in\mathbb{R}}\left\{\frac{1}{\beta}(x-z_{i-1})^2 + c(i,x)\right\},\quad i\in\{1,2,\dots,N\}$$

[0019] により定まる。βは、現在バイアスの強さ、特に準双曲性の強さを表すパラメータであり、０≦β≦１を満たす。βが小さければ、直近のコストに特に大きな重みを置く被験者になり、大きければ未来のコストも評価するような被験者になる。

[0020] これを整理すると、

[0021] [数3]

$$z_i = \begin{cases} a + \dfrac{N-i}{N-i+\beta}(z_{i-1}-a) & \text{if } -\sqrt{R(N-i+\beta)} \leq z_{i-1}-a, \\ z_{i-1} & \text{otherwise} \end{cases} \tag{1}$$

[0022] が分かる。つまり、

[0023] [数4]

$$-\sqrt{R(N-i+\beta)} \leq z_{i-1}-a$$

[0024] の場合、報酬を目指し、

[0025]

[数5]

$$z_{i-1} - a < -\sqrt{R(N-i+\beta)}$$

[0026] の場合、報酬を諦める。

[0027] 式（１）から、一旦報酬を諦めたら再度報酬を目指すことはない。すなわち、ｚ$_{i-1}$＝ｚ$_{i}$ならば、ｚ$_{i-1}$＝ｚ$_{i}$＝…＝ｚ$_{N}$となることが分かる。よって、

[0028] [数6]

$$p_i \coloneqq \frac{N-i}{N-i+\beta}$$

[0029] とおくと、

[0030] [数7]

$$z_i = \begin{cases} a + p_i(z_{i-1} - a) & \text{if } 1 \leq i \leq k, \\ z_{i-1} & \text{if } i > k \end{cases}$$

[0031] となるようなｋ≧０が存在する。このとき、

[0032] [数8]

$$z_N = a\left(1 - \prod_{i=1}^{k} p_i\right)$$

[0033] であり、ｋは、

[0034] [数9]

$$-\sqrt{R(N-j-1+\beta)} > a\left(1 - \prod_{i=1}^{j} p_i\right) - a$$

[0035] を満たす最小の０≦ｊ＜Ｎである。ただし、全ての０≦ｊ＜Ｎで上の条件が成り立たない場合、ｋ＝Ｎとする。以上の条件を整理すると、

[0036] [数10]

$$\frac{1}{N-j-1+\beta}\prod_{i=1}^{j} p_i^2 > \frac{R}{a^2} \tag{2}$$

[0037] となる。

[0038] [数11]

$$q_i \coloneqq \frac{1}{N-i-1+\beta}\prod_{j=1}^{i} p_j^2$$

[0039] とおき、q_1の形状について調べる。最初に、

[0040] [数12]

$$\frac{q_i}{q_{i-1}} = \frac{N-i+\beta}{N-i-1+\beta} p_i^2 = \frac{(N-i)^2}{(N-i+\beta)(N-i-1+\beta)}$$

[0041] である。これにより、$q_i > q_{i-1}$となるiの範囲は、$\beta > 1/2$のとき、

[0042] [数13]

$$N - \frac{\beta(1-\beta)}{2\beta-1} < i$$

[0043] であることが分かる。$\beta \leqq 1/2$のときは、任意のiで$q_i > q_{i-1}$となる。これらのことから以下のことが分かる：

・$\beta > 1/2$且つ

[0044] [数14]

$$\frac{\beta(1-\beta)}{2\beta-1} < 1$$

[0045] の場合（$\beta >$（$\sqrt{5}-1$）／２＝０．６１８．．．）、q_iは、iについて単調減少となる。つまり、βが十分大きい（先見の明がある）場合、報酬を諦めるなら時刻i＝０である。

[0046] ・ある閾値$1/2 < \beta_0 <$（$\sqrt{5}-1$）／２が存在し、

[0047] [数15]

$$\underset{0 \leq i < N}{\operatorname{argmax}}\, q_i = \begin{cases} N-1 & \text{if } \beta < \beta_0, \\ \{0, N-1\} & \text{if } \beta = \beta_0, \\ 0 & \text{if } \beta > \beta_0, \end{cases}$$

[0048] となる。よって、

[0049] [数16]

$$\max_{0 \leq i < N} q_i = \max\{q_0, q_{N-1}\}$$

[0050] である。

[0051] ・$\beta < \beta_0$の場合、a，Rに応じて諦める時刻が０≦ｉ≦Ｎ－２の範囲で変わる。

[0052] ここで、被験者が報酬を必ず獲得できるという条件のもとで、どの程度目

標ａを大きくすることができるか、という問題を考える。Ｒ＞０、現在バイアスパラメータβ、時刻の最大値Ｎは、入力として与えられるとする。つまり最適化問題

[0053] [数17]

$$\max_{a\geq 0} z_N$$

[0054] を解くことを考える。

[0055] （ａ）$\beta \geqq \beta_0$の場合、時刻０で諦めなければ、$z_N = a$となる。そのため、時刻０で諦めないために、

$$q_0 \leqq R / a^2$$

すなわち、

[0056] [数18]

$$a \leq \sqrt{R(N-1+\beta)}$$

[0057] が必要となる。上の条件を満たす範囲でａをなるべく大きく設定するのが良い。すなわち、$\beta \geqq \beta_0$である場合、

[0058] [数19]

$$a = \sqrt{R(N-1+\beta)} \tag{3}$$

[0059] とするのが最適となる。

（ｂ）$\beta < \beta_0$の場合、q_iが単調増加であることから、q_{N-1}が最大である。そのため、時刻Ｎ－１で諦めさせなければ良く、そのために式（２）より、

[0060] [数20]

$$\frac{1}{\beta}\prod_{i=1}^{j} p_i^2 \leq \frac{R}{a^2} \Leftrightarrow a \geq \sqrt{\beta R}\prod_{i=1}^{N-1} p_i = \sqrt{\beta R}\frac{\Gamma(N+\beta)}{\Gamma(N)\Gamma(1+\beta)}$$

[0061] とすればよい。すなわち、$\beta < \beta_0$であるならば、

[0062] [数21]

$$a = \sqrt{\beta R}\frac{\Gamma(N+\beta)}{\Gamma(N)\Gamma(1+\beta)} \tag{4}$$

[0063] とするのが最適である。

[0064] （報酬スケジューリングの最適化）

次に、報酬を与える際にどのようなタイミングでどの程度の報酬を与えることで被験者の到達地点を最大化することができるかという問題を考える。

[0065] 詳しい設定は、以下の通りである。タイムステップ数Ｎの間に総額Ｒの報酬を与える。報酬は、複数回に分割して提示することが可能であり、一度の報酬は、時刻ｎと必要な達成値ａと報酬額ｒの組（ｎ，ａ，ｒ）によって表現される。一つの報酬を提示した場合、被験者は、その報酬を目指して行動し、その報酬を獲得する（被験者が獲得できないような報酬を置くことはできない）。そうなったとき、提示していた報酬は、消滅し、新たな報酬を提示することができるようになる。このとき、総額の報酬Ｒをどのように分割しどのように提示するのが最適なのかを考える。

[0066] 最適解は、βの値によって変わるため、それぞれの場合について場合分けする。

[0067] （１）$\beta \geqq \beta_0$の場合

ｋ回に分けて報酬を与える場合を考える。それぞれの報酬を、（N_1，A_1，R_1），（N_1+N_2，A_1+A_2，R_1+R_2），．．．，（$\Sigma^k_{i=1}N_i$，$\Sigma^k_{i=1}A_i$，$\Sigma^k_{i=1}R_i$）にするとする。このとき、式（３）より、

[0068] [数22]

$$A_i = \sqrt{R_i(N_i - 1 + \beta)}$$

[0069] とするのが最適であり、最終到達地点（時刻がタイムステップ数となった際に必要な達成値）は、

[0070] [数23]

$$\sum_{i=1}^{k} \sqrt{R_i(N_i - 1 + \beta)}$$

[0071] となる。よって、解きたい最適化問題は、

[0072]

[数24]

$$\begin{aligned}\max_{N_i,R_i}\ & \sum_{i=1}^{k}\sqrt{R_i(N_i-1+\beta)},\\ \text{s.t.}\quad & \sum_{i=1}^{k}N_i=N,\\ & \sum_{i=1}^{k}R_i=R,\\ & N_i\in\mathbb{Z}_{>0},\\ & R_i\geq 0.\end{aligned}$$

[0073] となる。この最適化問題の最適値をV_kとすると、

[0074] [数25]

$$\begin{aligned}\sum_{i=1}^{k}\sqrt{R_i(N_i-1+\beta)} &\leq \sqrt{\left(\sum_{i=1}^{k}R_i\right)}\sqrt{\left(\sum_{i=1}^{k}(N_i-1+\beta)\right)}\ (\because \text{Cauchy–Shwartz の不等式})\\ &= \sqrt{R\,(N-k(1-\beta))}\end{aligned}$$

[0075] より、

[0076] [数26]

$$V_k\leq\sqrt{R\,(N-k(1-\beta))}$$

[0077] が成り立つ。また、

[0078] [数27]

$$V_1=\sqrt{R\,(N-1+\beta)}\geq\sqrt{R\,(N-k(1-\beta))}$$

なので、$V_1\geqq V_k$が成り立つ。つまり、ｋ＝１の場合が最適であり、到達地点は、

[0079] [数28]

$$\sqrt{R\,(N-1+\beta)}$$

[0080] である。

[0081] 以上のことから、$\beta\geqq\beta_0$の場合、報酬を分割せず、一括で被験者に報酬を与える方が最終到達地点（タイムステップ数における達成値）が大きくなることが分かる。さらに、分割した場合、損をすることも示唆している。

[0082] （２）$\beta\leqq\beta_0$の場合

ｋ回に分けて報酬を与える場合を考える。それぞれの報酬を（N_1，A_1，

R_1），（N_1+N_2，A_1+A_2，R_1+R_2），．．．．，（$\Sigma^k{}_{i=1}N_i$，$\Sigma^k{}_{i=1}A_i$，$\Sigma^k{}_{i=1}R_i$）にするとする。このとき、式（４）より

[0083] [数29]

$$A_i = \sqrt{\beta R_i}\frac{\Gamma(N_i+\beta)}{\Gamma(1+\beta)\Gamma(N_i)}$$

[0084] とするのが最適であり、最終到達地点は、

[0085] [数30]

$$\sum_{i=1}^{k}\sqrt{\beta R_i}\frac{\Gamma(N_i+\beta)}{\Gamma(1+\beta)\Gamma(N_i)}$$

[0086] となる。よって、解きたい最適化問題は、

[0087] [数31]

$$\begin{aligned}
\max_{N_i,R_i}\ & \sum_{i=1}^{k}\sqrt{R_i}\frac{\Gamma(N_i+\beta)}{\Gamma(N_i)}\\
\text{s.t.}\quad & \sum_{i=1}^{k}N_i = N,\\
& \sum_{i=1}^{k}R_i = R,\\
& N_i \in \mathbb{Z}_{>0},\\
& R_i \geq 0.
\end{aligned}$$

[0088] となる。

[0089] まず、N_iを固定してR_iを動かすことを考える。

[0090] [数32]

$$\begin{aligned}
\sum_{i=1}^{k}\sqrt{R_i}\frac{\Gamma(N_i+\beta)}{\Gamma(N_i)} &\leq \sqrt{\left(\sum_{i=1}^{k}R_i\right)\left(\sum_{i=1}^{k}\left(\frac{\Gamma(N_i+\beta)}{\Gamma(N_i)}\right)^2\right)} \quad (\because \text{Cauchy Shwartz の不等式})\\
&= \sqrt{R}\sqrt{\sum_{i=1}^{k}\left(\frac{\Gamma(N_i+\beta)}{\Gamma(N_i)}\right)^2}
\end{aligned}$$

[0091] より、N_iの最適化問題は、

[0092] [数33]

$$\begin{aligned}
\max_{N_i}\ & \sum_{i=1}^{k}\left(\frac{\Gamma(N_i+\beta)}{\Gamma(N_i)}\right)^2\\
\text{s.t.}\quad & \sum_{i=1}^{k}N_i = N,\\
& N_i \in \mathbb{Z}_{>0}.
\end{aligned}$$

[0093] となる。ここで、等号成立は、

[0094] [数34]

$$Q_i := \frac{\Gamma(N_i+\beta)}{\Gamma(N_i)} \text{ として } R_i := \frac{Q_i^2}{\sum_{j=1}^{k^*} Q_j^2}$$

[0095] のときである。このまま解くのが困難であるため、

[0096] [数35]

$$\frac{\Gamma(N+\beta)}{\Gamma(N)} \approx N^{\beta}$$

[0097] という近似を用いると、

[0098] [数36]

$$\begin{aligned} \max_{N_i} \quad & \sum_{i=1}^{k} N_i^{2\beta} \\ \text{s.t.} \quad & \sum_{i=1}^{k} N_i = N, \\ & N_i \in \mathbb{Z}_{>0}. \end{aligned}$$

[0099] を得る。

[0100] $\beta \geqq 1/2$ の場合、どれかに偏らせるのが最適である。つまり、k＝1とするのが良い。

[0101] $\beta < 1/2$ の場合、$N_i = N/k$ とするのが最適である。このときの最終到達地点は、近似を用いずに計算すると、

[0102] [数37]

$$k\sqrt{\beta\frac{R}{k}}\frac{\Gamma(N/k+\beta)}{\Gamma(1+\beta)\Gamma(N/k)}$$

[0103] と書ける。x：＝N／kとおくと、

[0104] [数38]

$$\sqrt{\frac{\beta N R}{x}}\frac{\Gamma(x+\beta)}{\Gamma(1+\beta)\Gamma(x)} = \frac{\sqrt{\beta N R}}{\Gamma(1+\beta)}\frac{\Gamma(x+\beta)}{\sqrt{x}\Gamma(x)}$$

[0105] となる。よって、

[0106] [数39]

$$\frac{\Gamma(x+\beta)}{\sqrt{x}\Gamma(x)} \tag{5}$$

[0107] が最大になるようにkを計算すれば良い。これは、適当な範囲でkを探索し

、関数値を評価することが可能である。ただしＮ／ｋは、整数になるとは限らないので、その場合、$\Sigma^{k}_{i=1} N_i = N$となるように適当に整数値に丸めることでN_iを決定する。

[0108] 以上をまとめると以下のＡｌｇｏｒｉｔｈｍ１によって報酬スケジュールを決定すれば良い

[0109] [表1]

Algorithm 1: 報酬スケジュールの最適化アルゴリズム

Input: タイムステップ数 N, 現在バイアスの強さ β, 報酬の総予算 R

Output: 報酬の与える回数 k，報酬のスケジューリング $(N_i, A_i, R_i)_{i=1,2,\dots,k}$

β_0 を計算

if $\beta \geq \beta_0$ **then**

  return $k = 1, (N, \sqrt{R(N-1+\beta)}, R)$

else if $\beta \geq \frac{1}{2}$ **then**

  return $k = 1, (N, \sqrt{\beta R}\frac{\Gamma(N+\beta)}{\Gamma(1+\beta)\Gamma(N)}, R)$

else

  k を $\{1, 2, \dots, N\}$ において全探索し，(5) を最大化する k を k^* とする

  N を k^* 等分して $N_1, \dots, N_{k^*}$ とする．ただし，整数にならない場合は適当に $\sum_{i=1}^{k^*} N_i = N$ となるように適当に整数値に丸める

  $Q_i \coloneqq \frac{\Gamma(N_i+\beta)}{\Gamma(N_i)}, R_i \coloneqq \frac{Q_i^2}{\sum_{j=1}^{k^*} Q_j^2}$

  return $k = k^*, (N_i, \sqrt{\beta R}\frac{\Gamma(N_i+\beta)}{\Gamma(1+\beta)\Gamma(N_i)}, R_i)_{i=1,\dots,k^*}$

[0110] なお、Ａｌｇｏｒｉｔｈｍ１を用いた動作の詳細な説明は後述する。

[0111] 図１は、本実施形態のイメージを示した図である。

[0112] 例えば、図１の（ａ）に示すように、被験者が２週間以内に数値指標を上げることを目指すとする。そして、介入者は、総額１０００円のインセンティブ（報酬）を被験者に与えることで被験者が達成する数値指標を最大化したいと望む。

[0113] 図１の（ｂ１）～（ｂ３）に示すように、どのタイミングで被験者にインセンティブ（報酬）を与えるかの例を示す。例えば、（ｂ１）の場合、１週

間後に３００円のインセンティブを被験者に与え、２週間後に７００円のインセンティブを被験者に与える例を示している。同様に、（ｂ２）の場合、１週間後に５００円のインセンティブを被験者に与え、２週間後に５００円のインセンティブを被験者に与える例を示している。（ｂ３）の場合、２週間後に１０００円のインセンティブを与える例を示している。本実施形態では、介入者がどのタイミングでどの程度のインセンティブを被験者に与えれば数値指標を最大化することができるかというスケジュールを提示することが可能となる。

[0114] （構成）

図２は、本実施形態に係る情報処理装置１のハードウェア構成の一例を示すブロック図である。

情報処理装置１は、入力されたデータに基づいて、被験者（エージェント）に報酬を与えるスケジュールを出力することを行うコンピュータである。情報処理装置１は、情報処理装置１を管理する管理者から入力された各種情報に基づいて、被験者が目標を達成するために必要な報酬を与えるスケジュール（タイミング）を作成することが可能である。情報処理装置１は、管理者が持ち運び可能な携帯端末であって良い。しかしながら、情報処理装置１は、携帯端末に限られず、ユーザが作業を行うことが可能な据え置き型のパーソナルコンピュータ等であっても良い。

[0115] 図２を参照すると、情報処理装置１は、制御部１０と、プログラム記憶部２０と、データ記憶部３０と、通信インタフェース４０と、入出力インタフェース５０と、入力装置５１と、出力装置５２と、を備える。制御部１０、プログラム記憶部２０、データ記憶部３０、通信インタフェース４０、および入出力インタフェース５０は、バスを介して互いに通信可能に接続されている。さらに、入出力インタフェース５０は、入力装置２および出力装置３と通信可能に接続される。

[0116] 制御部１０は、情報処理装置１を制御する。制御部１０は、中央処理ユニット（ＣＰＵ：Central Processing Unit）等のハードウェアプロセッサを備

える。例えば、制御部１０は、様々なプログラムを実行することが可能な集積回路であっても良い。

[0117] プログラム記憶部２０は、記憶媒体として、例えば、ＥＰＲＯＭ（Erasable Programmable Read Only Memory）、ＨＤＤ（Hard Disk Drive）、ＳＳＤ（Solid State Drive）等の随時書込みおよび読出しが可能な不揮発性メモリと、ＲＯＭ（Read Only Memory）等の不揮発性メモリとを組み合わせて使用することができる。プログラム記憶部２０は、各種処理を実行するために必要なプログラムを格納している。すなわち、制御部１０は、プログラム記憶部２０に格納されたプログラムを読み出して実行することにより各種制御および動作を実現し得る。

[0118] データ記憶部３０は、記憶媒体として、例えば、ＨＤＤ、メモリカード等の随時書込みおよび読出しが可能な不揮発性メモリと、ＲＡＭ（Random Access Memory）等の揮発性メモリとを組み合わせて使用したストレージである。データ記憶部３０は、制御部１０がプログラムを実行して各種処理を行う過程で取得および生成されたデータを記憶するために用いられる。

[0119] 通信インタフェース４０は、１つ以上の有線または無線の通信モジュールを含む。例えば、通信インタフェース４０は、ネットワークを介して外部装置と無線接続する通信モジュールを含む。また、通信インタフェース４０は、ネットワークを介さずに外部装置と直接接続を可能とする有線の通信モジュールを含んでも良い。さらに、通信インタフェース４０は、近距離無線技術を利用して外部装置と無線接続するための無線通信モジュールを含んでも良い。すなわち、通信インタフェース４０は、制御部１０の制御の下、外部装置との間で通信を行い、各種情報を送受信することができるものであれば一般的な通信インタフェースで良い。

[0120] 入出力インタフェース５０は、入力装置５１および出力装置５２等と接続される。入出力インタフェース５０は、入力装置５１および出力装置５２等の間で情報の送受信を可能にするインタフェースである。入出力インタフェース５０は、通信インタフェース４０と一体であっても良い。例えば、情報

処理装置１と、入力装置２および出力装置３等の少なくとも１つとは、近距離無線技術等を使用して無線接続されており、当該近距離無線技術を用いて情報の送受信を行っても良い。

[0121] 入力装置２は、例えば、情報処理装置１の管理者が各種情報を入力するためのキーボードやポインティングデバイス等を含む。また、入力装置５１は、プログラム記憶部２０またはデータ記憶部３０に格納するべきデータを、ＵＳＢメモリ等のメモリ媒体から読み出すためのリーダや、そのようなデータをディスク媒体から読み出すためのディスク装置を含んでも良い。

[0122] 出力装置３は、情報処理装置１から管理者に提示するべき出力データを表示するディスプレイや、それを印刷するプリンタ等を含む。

[0123] 図３は、本実施形態における情報処理装置１のソフトウェア構成を、図２に示したハードウェア構成に関連付けて示すブロック図である。

制御部１０は、入力情報取得部１０１と、報酬スケジュール最適化部１０２と、出力制御部１０３と、を備える。また、データ記憶部３０は、入力情報記憶部３０１と、報酬スケジュール記憶部３０２と、を備える。

[0124] 入力情報取得部１０１は、入力情報を取得する。入力情報は、タイムステップ数Ｎ、現在のバイアスの強さβ、報酬の総予算Ｒ等を含む。入力情報取得部１０１は、入力情報を入力情報記憶部３０１に記憶させる。現在バイアスの強さβは、介入者が被験者の状態から決定することができる値である。例えば、介入者は、アンケート用紙を用いて被験者の状態を決定し、当該被験者の状態から現在バイアスの強さβを決定する。

[0125] 報酬スケジュール最適化部１０２は、入力情報に基づいて被験者に対してどのようなタイミングで報酬を与えれば数値指標を最大化することができるのかのスケジュールを最適化する最適化部である。例えば、報酬スケジュール最適化部１０２は、上述したＡｌｇｏｒｉｔｈｍ１に従って報酬スケジュールを最適化する。報酬スケジュール最適化部１０２は、β_0算出部１０２１と、β判定部１０２２と、報酬スケジュール算出部１０２３と、を備える。

[0126] β_0算出部１０２１は、入力情報記憶部３０１に記憶された入力情報を読み

出す。そして、β_0算出部１０２１は、少なくともタイムステップ数Ｎおよび現在バイアスの強さβに基づいて、１／２以上である所定の閾値β_0を算出する。なお、所定の閾値β_0の算出方法の詳細は、後述する。

[0127] βが判定部１０２２は、入力情報に含まれる現在バイアスの強さβが算出したβ_0以上であるかどうかを判定する。βがβ_0以上である場合、β判定部１０２２は、入力情報およびβがβ_0以上であることを示す情報を報酬スケジュール算出部１０２３に出力する。さらにβ判定部１０２２は、βがβ_0未満であると判定した場合、βが１／２より大きいかどうかを判定する。

[0128] 報酬スケジュール算出部１０２３は、β比較部の判定した結果に対応する報酬スケジュールを算出する。ここで、報酬スケジュールは、報酬を与える回数、報酬を与える時刻、与えるべき報酬、および当該時刻に必要な達成値を含む。例えば、報酬スケジュール算出部１０２３は、上で説明したＡｌｇｏｒｉｔｈｍ１に従って報酬スケジュールを算出する。なお、報酬スケジュールの算出方法のより詳細な説明は、後述する。

[0129] 出力制御部１０３は、出力装置３のディスプレイ等に報酬スケジュール算出部１０２３により算出された最適化されたスケジュールを表示するように制御する。

[0130] 入力情報記憶部３０１は、入力情報取得部１０１が取得した入力情報を記憶するために用いられる。また、報酬スケジュール記憶部３０２は、報酬スケジュール最適化部１０２により算出された報酬スケジュールを記憶するために用いられる。

[0131] （動作）

図４は、情報処理装置１が報酬スケジュールを最適化し、最適化した報酬スケジュールを出力するための動作の一例を示すフローチャートである。

すなわち、図４は、情報処理装置１が上記したＡｌｇｏｒｉｔｈｍ１を用いて報酬スケジュールを最適化する動作の一例を示している。

[0132] 情報処理装置１の制御部１０がプログラム記憶部２０に記憶されたプログラムを読み出して実行することにより、このフローチャートの動作が実現さ

れる。

[0133] 動作は、情報処理装置１の管理者の指示によって開始する。或いは、制御部１０が通信インタフェース４０を通じて、外部装置から以下で説明する入力情報を受信した際に開始しても良い。

[0134] ステップＳＴ１０１で、入力情報取得部１０１は、入力情報を取得する。入力情報は、タイムステップ数Ｎ、現在のバイアスの強さβ、報酬の総予算Ｒ等を含む。入力情報取得部１０１は、入力情報を入力情報記憶部３０１に記憶させる。

[0135] 図５は、入力情報の一例を示した図である。

図５の例では、タイムステップ数Ｎが２０で、現在のバイアスの強さβが０．６で、報酬の総予算が１００である場合を示している。

[0136] ステップＳＴ１０２で、β_0算出部１０２１は、β_0を算出する。β_0算出部１０２１は、入力情報記憶部３０１に記憶された入力情報を読み出す。ここで、以上で説明した$q_0 = q_{N-1}$となるようなβがβ_0となる。$q_{N-1} - q_0$は、βに関して狭義単調増加である。すなわち、$\beta = 1/2$のとき、$q_{N-1} - q_0 < 0$、$\beta = (\sqrt{5}-1)/2$のとき、$q_{N-1} - q_0 > 0$である。そのため、$q_0 = q_{N-1}$となるようなβ_0がただ１つ存在する。そこで、β_0算出部１０２１は、$1/2 < \beta < (\sqrt{5}-1)/2$の範囲のβを探索し、$q_0 = q_{N-1}$となるβを探索する。或いは、$\beta_0$算出部１０２１は、二分探索を用いて$q_0 = q_{N-1}$となるβを探索しても良い。以上から、$\beta_0$算出部１０２１は、少なくともタイムステップ数Ｎおよび現在バイアスの強さβに基づいて所定の閾値であるβ_0を算出する。β_0算出部１０２１は、算出したβ_0をβ判定部１０２２に出力する。

[0137] ステップＳＴ１０３で、β判定部１０２２は、βがβ_0以上であるかどうかを判定する。そして、β判定部１０２２は、入力情報に含まれるβが算出したβ_0以上であるかどうかを判定する。現在バイアスの強さβがβ_0以上であると判定された場合、β判定部１０２２は、入力情報およびβがβ_0以上であることを示す情報を報酬スケジュール算出部１０２３に出力する。そして、

処理は、ステップＳＴ１０４に進む。一方、βがβ$_0$未満であると判定された場合、処理はステップＳＴ１０５に進む。

[0138] ステップＳＴ１０４で、報酬スケジュール算出部１０２３は、βがβ$_0$より大きいとの判定結果に対応する報酬スケジュールを算出する。例えば、報酬スケジュール算出部１０２３は、上述したように、報酬を分割せず、一括で与える方が良いと判定する。すなわち、報酬スケジュール算出部１０２３は、報酬の回数ｋ＝１の場合が最適であり、報酬を与える時刻がタイムステップ数Ｎであり、当該時刻に必要な達成値が式（３）で示した値であり、与えるべき報酬が報酬の総予算Ｒとなる場合を最適な報酬スケジュールであると算出する。報酬スケジュール算出部１０２３は、算出した報酬スケジュールを報酬スケジュール記憶部３０２に記憶させる。

[0139] ステップＳＴ１０５で、β判定部１０２２は、βが１／２以上であるかどうか判定する。入力情報のβが１／２以上である場合、β判定部１０２２は、βが１／２以上であることを示す情報および入力情報を報酬スケジュール算出部１０２３に出力する。一方、βが１／２未満である場合、β判定部１０２２は、入力情報およびβが１／２未満であることを示す情報および入力情報を報酬スケジュール算出部１０２３に出力する。

[0140] ステップＳＴ１０６で、報酬スケジュール算出部１０２３は、βがβ$_0$より小さいが１／２よりも大きいとの判定結果に対応する報酬スケジュールを算出する。例えば、上述したように、どれかに偏らせるのが最適なスケジュールとなる。そのため、報酬スケジュール算出部１０２３は、報酬の回数ｋ＝１の場合が最適であり、報酬を与える時刻がタイムステップ数Ｎであり、当該時刻に必要な達成値が

[0141] [数40]

$$\sqrt{R(N-1+\beta)}$$

[0142] であり、与えるべき報酬が報酬の総予算Ｒとなる場合を最適な報酬スケジュールであると算出する。報酬スケジュール算出部１０２３は、算出した報酬スケジュールを報酬スケジュール記憶部３０２に記憶させる。

[0143] ステップＳＴ１０７で、報酬スケジュール算出部１０２３は、βが１／２未満である判定結果に対応する報酬スケジュールを算出する。上述したように、βが１／２未満である場合、$N_i = N/k$とするのが最適となる。そこで、報酬スケジュール算出部１０２３は、報酬を与える回数ｋを｛１，２，．．．，Ｎ｝において全探索し、式（５）を最大化するｋをｋ*とする。Ｎをｋ*等分して

[0144] [数41]

$$\sqrt{\beta R}\frac{\Gamma(N+\beta)}{\Gamma(1+\beta)\Gamma(N)}$$

[0145] とする。但し、ｋ*等分の結果、整数にならない場合、

[0146] [数42]

$$\sum_{i=1}^{k^*} N_i = N$$

[0147] となるように整数値を丸めることにより報酬を与える時刻N_iを決定する。さらに上記したように、報酬スケジュール算出部１０２３は、時刻N_i毎に与えるべき報酬が

[0148] [数43]

$$Q_i := \frac{\Gamma(N_i+\beta)}{\Gamma(N_i)}, R_i := \frac{Q_i^2}{\sum_{j-1}^{k^*} Q_j^2}$$

[0149] と算出できる。

[0150] さらに、報酬スケジュール算出部１０２３は、時刻N_iに必要な達成値A_iが

[0151] [数44]

$$\sqrt{\beta R}\frac{\Gamma(N_i+\beta)}{\Gamma(1+\beta)\Gamma(N_i)}$$

[0152] と算出する。すなわち、報酬スケジュール算出部１０２３は、上記のスケジュールを最適化された報酬スケジュールであると算出する。報酬スケジュール算出部１０２３は、算出した報酬スケジュールを報酬スケジュール記憶部３０２に記憶させる。

[0153] ステップＳＴ１０８で、出力制御部１０３は、出力装置３のディスプレイ等に報酬スケジュールを表示するように制御する。出力制御部１０３は、報

酬スケジュール記憶部３０２に記憶された報酬スケジュールを取得し、出力装置３のディスプレイ等に報酬スケジュールを表示させるように制御する。例えば、介入者は、報酬スケジュールに基づいて被験者に対して報酬を与えることにより、被験者のパフォーマンスを最大化することが可能となる。

[0154] 図６および図７は、最適化スケジュールを出力制御部１０３が表示する際の一例を示している。

例えば、図６は、報酬を与える回数ｋを表す図であり、図７は、時刻N_i、時刻N_iに必要な達成値A_i、および与えるべき報酬R_iを表す図である。

[0155] 図６に示すように、報酬を与える回数が３回であり、図７では各回（ｉ＝１，ｉ＝２，ｉ＝３）の時刻N_i、時刻N_iに必要な達成値A_i、および与えるべき報酬R_iの一例を示している。

[0156] 図６および図７に示すように、βが＜１／２の場合、複数回の報酬を与えることになり、報酬を与えるまでの時刻N_i、時刻N_iに必要な達成値A_i、および与えるべき報酬R_iが上記のように決定される。このスケジュールに従って介入者が被験者に対して報酬を与えることにより、被験者は、最終到達地点まで諦めることなく続けることが可能となる。

[0157] （実施形態の作用効果）

本実施形態によれば、情報処理装置１は、現在バイアスの影響下における報酬の与え方のスケジュールを最適化することができる。これにより、対象となる人間（被験者）のタスクにおけるパフォーマンスを最大化することが可能となる。

[0158] ［他の実施形態］

上記の実施形態における情報処理装置１は、各構成要素の動作をプログラムとして構築し、情報処理装置として利用されるコンピュータにインストールして実行させることが可能である。

[0159] また、前記実施形態に記載した手法は、計算機（コンピュータ）に実行させることができるプログラム（ソフトウェア手段）として、例えば磁気ディスク（フロッピー（登録商標）ディスク、ハードディスク等）、光ディスク

（ＣＤ－ＲＯＭ、ＤＶＤ、ＭＯ等）、半導体メモリ（ＲＯＭ、ＲＡＭ、フラッシュメモリ等）等の記憶媒体に格納し、また通信媒体により伝送して頒布することもできる。なお、媒体側に格納されるプログラムには、計算機に実行させるソフトウェア手段（実行プログラムのみならずテーブル、データ構造も含む）を計算機内に構成させる設定プログラムをも含む。本装置を実現する計算機は、記憶媒体に記憶されたプログラムを読み込み、また場合により設定プログラムによりソフトウェア手段を構築し、このソフトウェア手段によって動作が制御されることにより上述した処理を実行する。なお、本明細書で言う記憶媒体は、頒布用に限らず、計算機内部或いはネットワークを介して接続される機器に設けられた磁気ディスク、半導体メモリ等の記憶媒体を含むものである。

[0160] 要するに、この発明は上記実施形態に限定されるものではなく、実施段階ではその要旨を逸脱しない範囲で種々に変形することが可能である。また、各実施形態は可能な限り適宜組み合わせて実施してもよく、その場合組み合わせた効果が得られる。さらに、上記実施形態には種々の段階の発明が含まれており、開示される複数の構成要件における適当な組み合わせにより種々の発明が抽出され得る。

符号の説明

[0161] １…情報処理装置

２…入力装置

３…出力装置

１０…制御部

１０１…入力情報取得部

１０２…報酬スケジュール最適化部

１０２１…β_0算出部

１０２２…β判定部

１０２３…報酬スケジュール算出部

１０３…出力制御部

２０…プログラム記憶部

３０…データ記憶部

３０１…入力情報記憶部

３０２…報酬スケジュール記憶部

４０…通信インタフェース

５０…入出力インタフェース

５１…入力装置

５２…出力装置

請求の範囲

[請求項1] 報酬の総予算、目標となる時刻までのタイムステップ数、現在のバイアスの強さを含む入力情報を取得する入力情報取得部と、

少なくとも前記タイムステップ数および前記現在のバイアスの強さに基づいて１／２よりも大きい所定の閾値を算出する算出部と、

前記現在のバイアスの強さが前記所定の閾値より大きいかどうかを判定し、前記現在のバイアスの強さが前記所定の閾値よりも小さい場合、前記現在のバイアスの強さが１／２よりも大きいかどうか判定する判定部と、

前記判定の結果に対応する、報酬を与える回数、報酬を与える時刻、与えるべき報酬、および前記時刻に必要な達成値を含む報酬スケジュールを算出する報酬スケジュール算出部と、

前記報酬スケジュールを出力するように制御する出力制御部と、

を備える、情報処理装置。

[請求項2] 前記判定の結果が前記現在のバイアスの強さが前記所定の閾値よりも大きいことを示す場合、前記報酬スケジュール算出部は、前記報酬を与える回数が１回であり、前記報酬を与える時刻が前記タイムステップ数であり、前記与えるべき報酬が前記報酬の総予算であり、前記時刻に必要な達成値が

[数1]

$$\sqrt{R(N-1+\beta)}$$

となる前記報酬スケジュールを算出する、ここで、Ｒは、前記報酬の総予算であり、前記Ｎは、前記タイムステップ数であり、βは、前記現在のバイアスの強さである、請求項１に記載の情報処理装置。

[請求項3] 前記判定の結果が前記現在のバイアスの強さが前記所定の閾値よりも小さいが１／２よりも大きいことを示す場合、前記報酬スケジュール算出部は、前記報酬を与える回数が１回であり、前記報酬を与える時刻が前記タイムステップ数であり、前記与えるべき報酬が前記報酬

の総予算であり、前記時刻に必要な達成値が

[数2]

$$\sqrt{\beta R}\frac{\Gamma(N+\beta)}{\Gamma(1+\beta)\Gamma(N)}$$

となる前記報酬スケジュールを算出する、ここで、Ｒは、前記報酬の総予算であり、前記Ｎは、前記タイムステップ数であり、βは、前記現在のバイアスの強さであり、Γは、Γ関数である、請求項１に記載の情報処理装置。

[請求項4] 前記判定の結果が前記現在のバイアスの強さが１／２よりも小さい場合、前記報酬スケジュール算出部は、報酬を与える回数を１から順にタイムステップ数まで全探索することにより報酬を与える回数ｋ＊を決定する、ここでｋは、１以上タイムステップ数未満の整数である、請求項１に記載の情報処理装置。

[請求項5] 前記報酬スケジュール算出部は、前記全探索により、報酬を与える回数をｋ＊であると算出し、前記報酬を与える時刻が前記タイムステップ数を、前記報酬を与える回数ｋ＊で割った値であり、前記与えるべき報酬が

[数3]

$$Q_i \coloneqq \frac{\Gamma(N_i+\beta)}{\Gamma(N_i)}, R_i \coloneqq \frac{Q_i^2}{\sum_j^{k^*}{}_1 Q_j^2}$$

であり、前記時刻に必要な達成値が

[数4]

$$\sqrt{\beta R}\frac{\Gamma(N_i+\beta)}{\Gamma(1+\beta)\Gamma(N_i)}$$

となる前記報酬スケジュールを算出する、ここで、Ｒは、前記報酬の総予算であり、前記Ｎは、前記タイムステップ数であり、βは、前記現在のバイアスの強さであり、Γは、Γ関数である、請求項４に記載の情報処理装置。

[請求項6] 前記報酬スケジュール算出部は、タイムステップ数が前記報酬を与える回数で割り切れない場合、$\Sigma^{k^*}{}_{i=1} N_i = N$となるように丸める

、請求項５に記載の情報処理装置。

[請求項7] 情報処理装置のプロセッサが実行する情報処理方法であって、

報酬の総予算、目標となる時刻までのタイムステップ数、現在のバイアスの強さを含む入力情報を取得することと、

前記報酬の総予算、前記タイムステップ数、および前記現在のバイアスの強さに基づいて１／２よりも大きい所定の閾値を算出することと、

前記現在のバイアスの強さが前記所定の閾値より大きいかどうかを判定することと、

前記現在のバイアスの強さが前記所定の閾値よりも小さい場合、前記現在のバイアスの強さが１／２よりも大きいかどうか判定することと、

前記判定の結果に対応する、報酬を与える回数、報酬を与える時刻、与えるべき報酬、および前記時刻に必要な達成値を含む報酬スケジュールを算出することと、

前記報酬スケジュールを出力するように制御することと、

を備える、情報処理方法。

[請求項8] 情報処理装置のプロセッサによって実行させるための命令を備える情報処理プログラムであって、前記命令は、

報酬の総予算、目標となる時刻までのタイムステップ数、現在のバイアスの強さを含む入力情報を取得することと、

前記報酬の総予算、前記タイムステップ数、および前記現在のバイアスの強さに基づいて１／２よりも大きい所定の閾値を算出することと、

前記現在のバイアスの強さが前記所定の閾値より大きいかどうかを判定することと、

前記現在のバイアスの強さが前記所定の閾値よりも小さい場合、前記現在のバイアスの強さが１／２よりも大きいかどうか判定すること

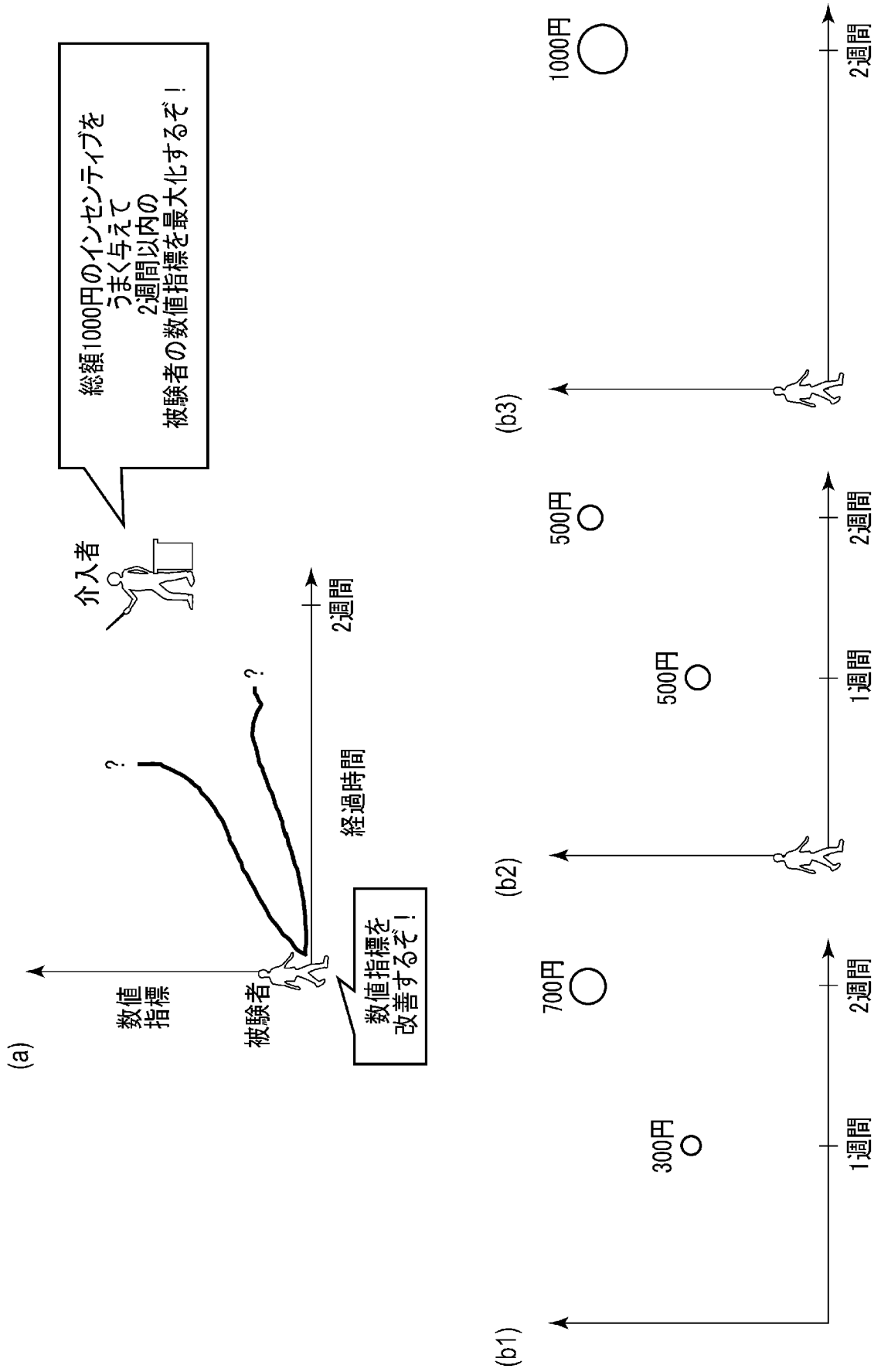
と、

前記判定の結果に対応する、報酬を与える回数、報酬を与える時刻、与えるべき報酬、および前記時刻に必要な達成値を含む報酬スケジュールを算出することと、

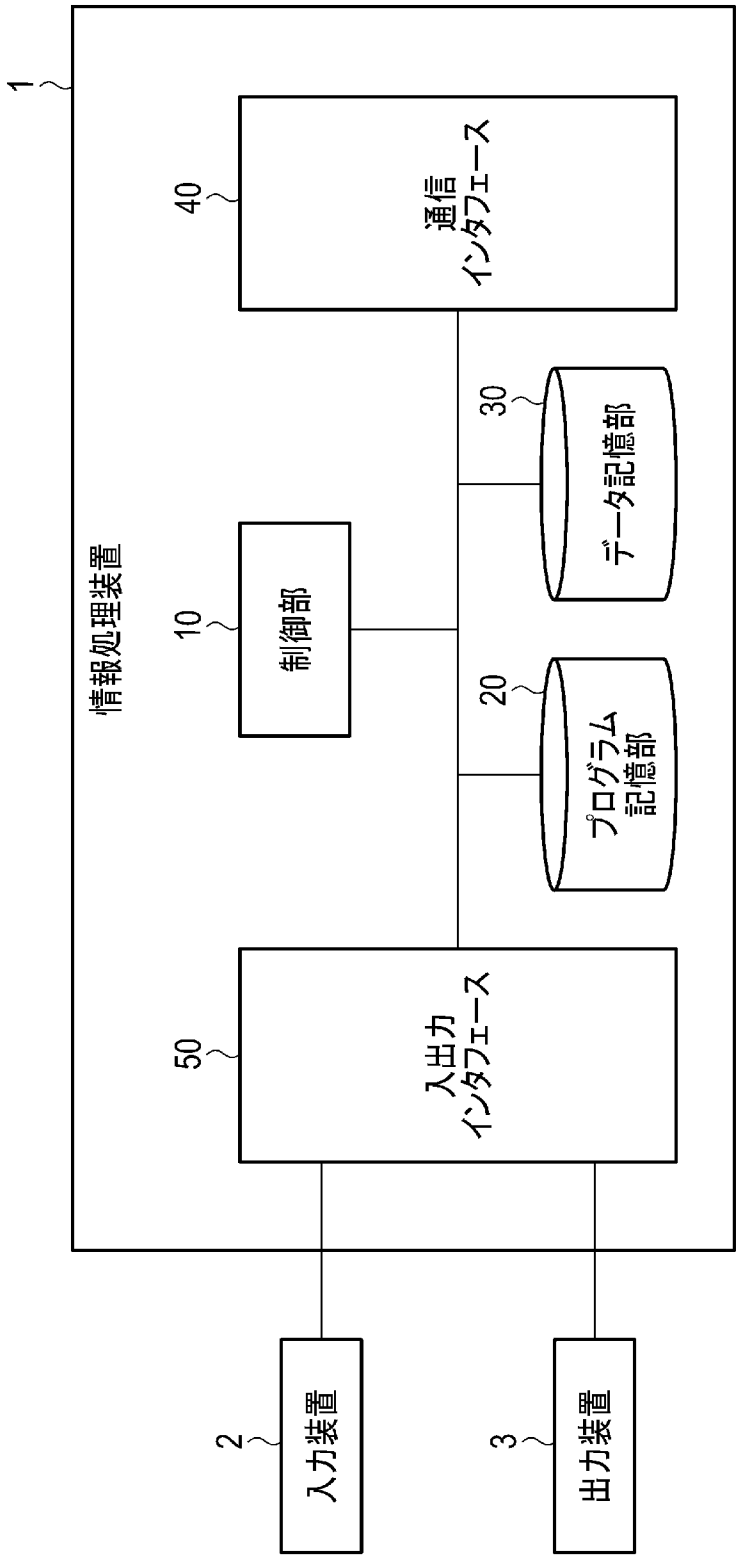
前記報酬スケジュールを出力するように制御することと、

を備える、情報処理プログラム。

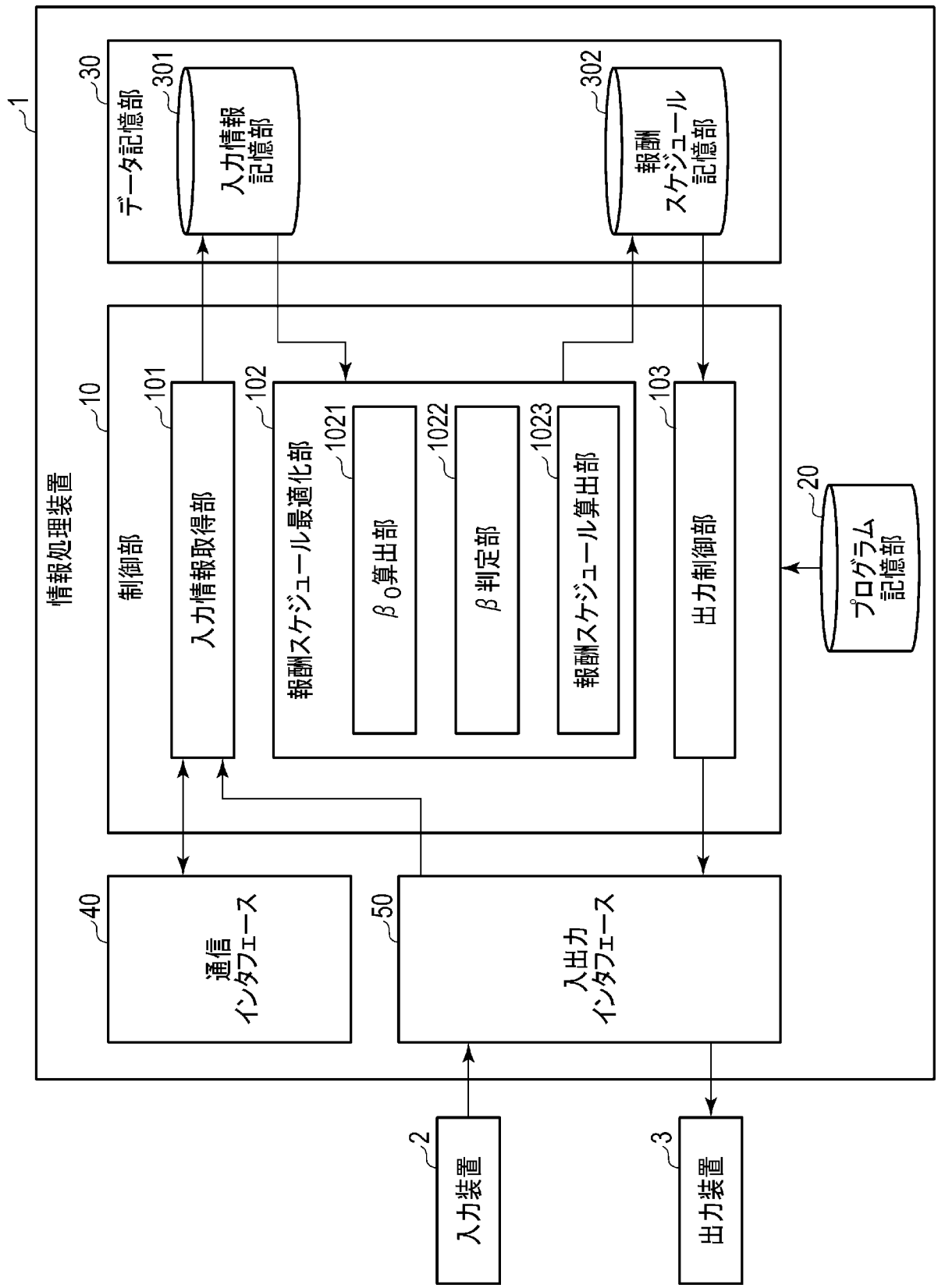
[図1]


(a)
数値指標
被験者
経過時間
2週間
介入者
総額1000円のインセンティブをうまく与えて2週間以内の被験者の数値指標を最大化するぞ！
数値指標を改善するぞ！
?
?
(b1)
300円
700円
1週間
2週間
(b2)
500円
500円
1週間
2週間
(b3)
1000円
2週間

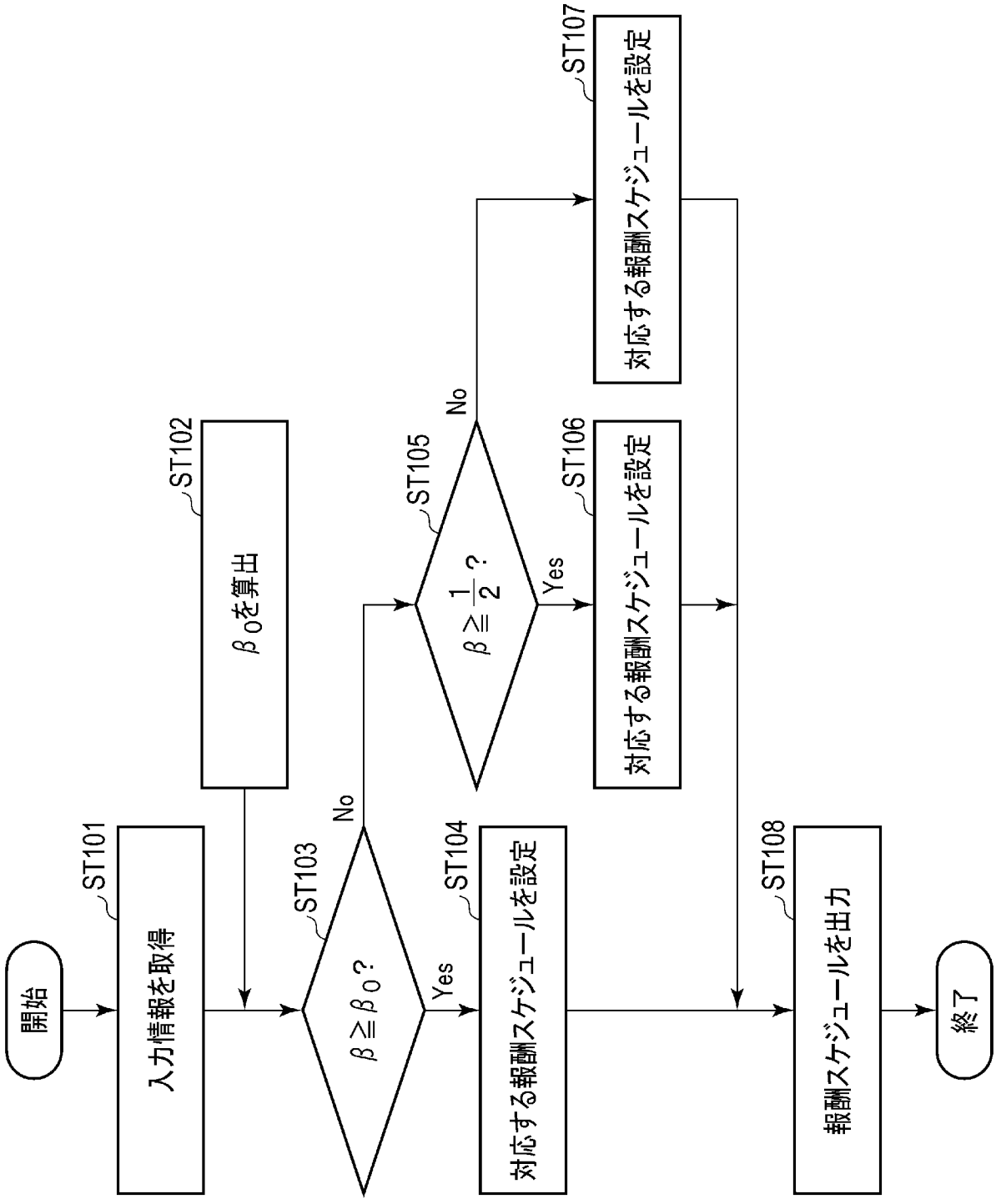
[図2]


1
情報処理装置
40
通信
インタフェース
30
データ記憶部
10
制御部
20
プログラム
記憶部
50
入出力
インタフェース
2
入力装置
3
出力装置

[図3]


1
情報処理装置
30
データ記憶部
301
入力情報記憶部
302
報酬スケジュール記憶部
10
制御部
101
入力情報取得部
102
報酬スケジュール最適化部
1021
β_0算出部
1022
β判定部
1023
報酬スケジュール算出部
103
出力制御部
20
プログラム記憶部
40
通信インタフェース
50
入出力インタフェース
2
入力装置
3
出力装置

[図4]


開始
ST101
入力情報を取得
ST102
β_0を算出
ST103
$\beta \geqq \beta_0$?
No
Yes
ST104
対応する報酬スケジュールを設定
ST105
$\beta \geqq \frac{1}{2}$?
No
Yes
ST106
対応する報酬スケジュールを設定
ST107
対応する報酬スケジュールを設定
ST108
報酬スケジュールを出力
終了

[図5]

	N	β	R
	20	0.6	100

[図6]

	k
	3

[図7]

	i=1	i=2	i=3
N_i	6	7	7
A_i	30	30	30
R_i	33	34	34

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2023/005036

A. CLASSIFICATION OF SUBJECT MATTER

G06Q 10/10(2023.01)i
FI: G06Q10/10

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)

G06Q10/10

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 1971-2023
Registered utility model specifications of Japan 1996-2023
Published registered utility model applications of Japan 1994-2023

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.
A	JP 2013-084175 A (SONY CORPORATION) 09 May 2013 (2013-05-09) entire text	1-8
A	JP 2019-141869 A (FANUC CORPORATION) 29 August 2019 (2019-08-29) entire text	1-8
A	WO 2021/165425 A1 (PHILIP MORRIS PRODUCTS SA) 26 August 2021 (2021-08-26) entire text	1-8
A	JP 2021-192141 A (THE UNIVERSITY OF TOKYO) 16 December 2021 (2021-12-16) entire text	1-8

☐ 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 April 2023	**09 May 2023**
Name and mailing address of the ISA/JP **Japan Patent Office (ISA/JP)** **3-4-3 Kasumigaseki, Chiyoda-ku, Tokyo 100-8915** **Japan**	Authorized officer Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/JP2023/005036

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
JP 2013-084175 A	09 May 2013	US 2013/0246318 A1 entire text US 2015/0278694 A1	
JP 2019-141869 A	29 August 2019	US 2019/0258982 A1 entire text DE 102019001044 A1 CN 110171159 A	
WO 2021/165425 A1	26 August 2021	CN 115104099 A KR 10-2022-0140790 A	
JP 2021-192141 A	16 December 2021	(Family: none)	

国際調査報告

国際出願番号
PCT/JP2023/005036

A. 発明の属する分野の分類（国際特許分類（ＩＰＣ））

G06Q 10/10(2023.01)i
FI: G06Q10/10

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））

G06Q10/10

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	１９７１－２０２３年
日本国実用新案登録公報	１９９６－２０２３年
日本国登録実用新案公報	１９９４－２０２３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
A	JP 2013-084175 A (ソニー株式会社) 09.05.2013 (2013-05-09) 全文	1-8
A	JP 2019-141869 A (ファナック株式会社) 29.08.2019 (2019-08-29) 全文	1-8
A	WO 2021/165425 A1 (PHILIP MORRIS PRODUCTS SA) 26.08.2021 (2021-08-26) 全文	1-8
A	JP 2021-192141 A (国立大学法人 東京大学) 16.12.2021 (2021-12-16) 全文	1-8

☐ 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
25.04.2023	09.05.2023

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）
青柳　光代 5L 4100

電話番号 03-3581-1101 内線 3502

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/JP2023/005036

引用文献	公表日	パテントファミリー文献	公表日
JP 2013-084175 A	09.05.2013	US 2013/0246318 A1 全文 US 2015/0278694 A1	
JP 2019-141869 A	29.08.2019	US 2019/0258982 A1 全文 DE 102019001044 A1 CN 110171159 A	
WO 2021/165425 A1	26.08.2021	CN 115104099 A KR 10-2022-0140790 A	
JP 2021-192141 A	16.12.2021	（ファミリーなし）	