

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7590984号
(P7590984)

(45)発行日 令和6年11月27日(2024.11.27)

(24)登録日 令和6年11月19日(2024.11.19)

| | | | | |
|-------------------------|---------|-------|-------|--|
| (51)国際特許分類 | F I | | | |
| G 0 6 F 16/28 (2019.01) | G 0 6 F | 16/28 | | |
| G 0 6 N 5/025(2023.01) | G 0 6 N | 5/025 | | |
| G 0 6 N 99/00 (2019.01) | G 0 6 N | 99/00 | 1 8 0 | |

請求項の数 14 (全26頁)

| | | | |
|-------------|----------------------------------|----------|---|
| (21)出願番号 | 特願2021-563019(P2021-563019) | (73)特許権者 | 000004237 日本電気株式会社 東京都港区芝五丁目7番1号 |
| (86)(22)出願日 | 平成31年4月24日(2019.4.24) | (74)代理人 | 100103894 弁理士 家入 健 |
| (65)公表番号 | 特表2022-537009(P2022-537009 A) | (72)発明者 | ウェイ ウェンジュアン 中華人民共和国 1 0 0 6 0 0 ベイジン 、チャオヤン ディストリクト、ドンフ ァンドンルー ナンバー 1 9、リャンマ ーチャオ ディプロマティック オフィス ビルディング、ビルディング ディー 2 、6エフ |
| (43)公表日 | 令和4年8月23日(2022.8.23) | (72)発明者 | リウ チュンチェン 中華人民共和国 1 0 0 6 0 0 ベイジン 、チャオヤン ディストリクト、ドンフ 最終頁に続く |
| (86)国際出願番号 | PCT/CN2019/084049 | | |
| (87)国際公開番号 | WO2020/215237 | | |
| (87)国際公開日 | 令和2年10月29日(2020.10.29) | | |
| 審査請求日 | 令和4年4月22日(2022.4.22) | | |
| 審判番号 | 不服2023-18947(P2023-18947/J 1) | | |
| 審判請求日 | 令和5年11月8日(2023.11.8) | | |

(54)【発明の名称】 データ処理方法、装置及びプログラム

(57)【特許請求の範囲】

【請求項 1】

複数の要素に関する観測サンプル集合を取得することであって、前記観測サンプル集合における1つの観測サンプルは、前記複数の要素の対応する観測値を含むことと、

前記観測サンプル集合に基づいて、前記複数の要素の間に存在する依存関係集合を決定することであって、前記依存関係集合における1つの依存関係は、前記複数の要素における互いに関連する1つの要素ペアを示すことと、

前記依存関係集合に基づいて、前記複数の要素の因果関係シーケンスを決定することであって、前記因果関係シーケンスは、互いに関連する前記要素ペアにおける一方の要素が他方の要素の原因であることを示すことと、を含み、

前記依存関係集合を決定することは、

前記複数の要素における任意の2つの要素について、前記観測サンプル集合における前記2つの要素の対応する観測値に基づいて、前記2つの要素の間の相関係数を推定することと、

前記推定した結果に基づいて、前記依存関係集合を決定するための第1の目標関数を確立することであって、前記第1の目標関数は下式で表され、 S は依存関係を示す精度行列であり、 s_{ij} は予め定義されている係数であり、 S は前記相関係数に基づいて確立される相関係数行列を表し、 $tr(\cdot)$ は行列のトレースを表す、ことと、

【数 1】

10

20

$$tr(S\Omega) - \log|\Omega| + \lambda \sum_{j \neq k} |\Omega_{jk}|$$

0 を満たしつつ前記第 1 の目標関数を最小化することで、前記依存関係集合を決定することと、を含む

10

データ処理装置が実行するデータ処理方法。

【請求項 2】

前記複数の要素は、目標要素を含み、前記方法は、

前記因果関係シーケンスに基づいて、前記複数の要素から前記目標要素の原因となる少なくとも 1 つの要素を決定することと、

前記少なくとも 1 つの要素の観測値を変更することで、前記目標要素の観測値に影響することと、をさらに含む請求項 1 に記載の方法。

【請求項 3】

前記複数の要素に関する変更後の観測サンプル集合を取得することであって、前記変更後の観測サンプル集合における少なくとも 1 つの観測サンプルは、前記少なくとも 1 つの要素の変更後の観測値を含むことと、

20

前記変更後の観測サンプル集合に基づいて、前記因果関係シーケンスを最適化することと、をさらに含む請求項 2 に記載の方法。

【請求項 4】

前記因果関係シーケンスを決定することは、

前記依存関係集合によって示された複数の互いに関連する要素ペアについて、各要素ペアにおける一方の要素から他方の要素への影響を決定することと、

複数の要素ペアについて決定された対応する影響及び前記観測サンプル集合に基づいて、前記因果関係シーケンスを決定することと、を含む請求項 1 に記載の方法。

【請求項 5】

30

各要素ペアにおける一方の要素から他方の要素への影響を決定することは、

所定の分布に基づいて、前記複数の要素ペアについて前記対応する影響を決定するための第 2 の目標関数を確立することであって、前記第 2 の目標関数は下式で表され、行列 B は、複数の要素ペアのそれぞれについて決定された影響を示し、 $X = \{x_1, x_2, \dots, x_D\}$ 、 $R^{N \times D}$ は前記観測サンプル集合を示し、 N は X における観測サンプルの総数を示し、 D は要素の総数を示し、ベクトル $x_i \in R^N$ (ただし、 $1 \leq i \leq D$) は、 i 番目の要素の N 個の観測値を示し、 $x_{i,n}$ (ただし、 $1 \leq i \leq D$ 、かつ $1 \leq n \leq N$) は、 i 番目の要素の n 番目の観測値を示し、ベクトル $\beta_i \in R^{D-1}$ は i 番目の要素についての各要素の影響を示し、 $|B|_0$ は、行列 B における非ゼロ元素の総数を示す、ことと、

【数 2】

40

$$\sum_{i=1}^D \log \left(\sum_{n=1}^N (x_{i,n} - X_{-i,n} \beta_i)^2 \right) + \frac{1}{N} \log(N) |B|_0$$

前記第 2 の目標関数を最小化することで、各要素ペアにおける一方の要素から他方の要素への影響を決定することと、を含む請求項 4 に記載の方法。

【請求項 6】

50

前記因果関係シーケンスを決定することは、

履歴因果関係シーケンス及び前記履歴因果関係シーケンスに対応する履歴因果関係のスコアを取得する操作と、

前記履歴因果関係シーケンス及び前記複数の要素ペアに基づいて、前記因果関係シーケンスに追加する可能性がある1つ又は複数の候補要素を決定する操作と、

前記1つ又は複数の候補要素が存在することに応答し、前記履歴因果関係のスコア、前記複数の要素ペアについて決定された前記対応する影響及び前記観測サンプル集合に基づいて、前記1つ又は複数の候補要素に対応する1つ又は複数の候補因果関係のスコアを決定する操作と、

前記1つ又は複数の候補因果関係のスコアに基づいて、前記1つ又は複数の候補要素から前記因果関係シーケンスに追加しようとする候補要素を選択する操作と、

選択された前記候補要素に基づいて、前記履歴因果関係シーケンス及び前記履歴因果関係のスコアを更新する操作と、を少なくとも1回繰り返して実行することを含む、請求項4に記載の方法。

【請求項7】

前記1つ又は複数の候補要素が存在しないことに応答し、前記履歴因果関係シーケンスを前記因果関係シーケンスとして決定することをさらに含む、請求項6に記載の方法。

【請求項8】

前記履歴因果関係シーケンス及び前記履歴因果関係のスコアを取得することは、

前記履歴因果関係シーケンスを空シーケンスに初期化することと、

前記空シーケンスに対応する初期因果関係のスコアを前記履歴因果関係のスコアとして決定することと、を含む請求項6に記載の方法。

【請求項9】

前記1つ又は複数の候補要素から前記候補要素を選択することは、

前記1つ又は複数の候補因果関係のスコアから最小の候補因果関係のスコアを決定することと、

前記1つ又は複数の候補要素から前記最小の候補因果関係のスコアに関連する前記候補要素を選択することと、を含む請求項6に記載の方法。

【請求項10】

前記1つ又は複数の候補要素から前記候補要素を選択することは、

決定待ちの前記因果関係シーケンスに関連する制約条件を取得することと、

前記候補要素の前記追加が前記制約条件に合致するように、前記1つ又は複数の候補要素から前記因果関係シーケンスに追加しようとする前記候補要素を選択することと、を含む請求項6に記載の方法。

【請求項11】

前記制約条件を取得することは、

前記制約条件を示す情報を取得することと、

前記情報に基づいて前記制約条件を決定することと、を含む請求項10に記載の方法。

【請求項12】

前記制約条件を取得することは、

前記履歴因果関係シーケンス及び前記複数の要素ペアに基づいて、前記制約条件を決定することを含む、請求項10に記載の方法。

【請求項13】

請求項1～12のいずれか1項に記載の方法をデータ処理装置に実行させるためのコンピュータプログラム。

【請求項14】

少なくとも1つのプロセッシングユニットと、

前記少なくとも1つのプロセッシングユニットに結合され、前記少なくとも1つのプロセッシングユニットが実行する、請求項13に記載のコンピュータプログラムを記憶する少なくとも1つのメモリと、を含むデータ処理装置。

10

20

30

40

50

【発明の詳細な説明】**【技術分野】****【0001】**

本開示の実施形態は、機械学習の分野に関し、さらに具体的には、データ処理方法、装置及びコンピュータ可読記憶媒体に関する。

【背景技術】**【0002】**

情報技術の飛躍的な発展に伴い、データの規模が急速に大きくなっている。機械学習は、このような背景及びトレンドでますます広く注目されている。中でも、因果発見は、現実の生活において、例えば、サプライチェーン、医療健康及びリテールなどの分野において幅広く適用されている。ここでは、上記の因果発見とは、複数の要素に関するサンプルデータから複数の要素の間に存在する因果関係を発見することである。例えば、リテールの分野では、因果発見の結果は、各種の販売戦略を立てることを補助するために用いられ、医療健康の分野では、因果発見の結果は、患者についての治療方案などを作成することを補助するために用いられる。

10

【0003】

しかしながら、技術の発展に伴い、単一のシステムにおいて因果関係が存在する可能性のある様々な要素の数は著しく増加している。また、異なるシステム間の連動にも注目されることが多い。これは、因果関係を発見しようとする要素の数が数百から千以上になる可能性のあることを招く。この場合、多くの要素の間に存在する因果関係を迅速かつ正確に発見することは、ますます重要なものとなっている。

20

【発明の概要】**【発明が解決しようとする課題】****【0004】**

本開示の実施形態は、データ処理方法、装置及びコンピュータ可読記憶媒体を提供する。
【課題を解決するための手段】

【0005】

本開示の第1の態様において、データ処理方法が提供される。該方法は、複数の要素に関する観測サンプル集合を取得することであって、該観測サンプル集合における各観測サンプルは、複数の要素の対応する観測値を含むことと、該観測サンプル集合に基づいて、複数の要素の間に存在する依存関係集合を決定することであって、該依存関係集合における1つの依存関係は、複数の要素における互いに関連する1つの要素ペアを指示することと、該依存関係集合に基づいて、複数の要素の因果関係シーケンスを決定することであって、該因果関係シーケンスは、互いに関連する要素ペアにおける一方の要素が他方の要素の原因であることを指示することと、を含む。

30

【0006】

本開示の第2の態様において、データ処理装置が提供される。該装置は、少なくとも1つのプロセッシングユニット及び少なくとも1つのメモリを含む。少なくとも1つのメモリは、少なくとも1つのプロセッシングユニットに結合され、少なくとも1つのプロセッシングユニットが実行する指令を記憶する。指令が少なくとも1つのプロセッシングユニットにより実行される場合、該装置に以下の動作を実行させ、動作は、複数の要素に関する観測サンプル集合を取得することであって、該観測サンプル集合における各観測サンプルは、複数の要素の対応する観測値を含むことと、該観測サンプル集合に基づいて、複数の要素の間に存在する依存関係集合を決定することであって、該依存関係集合における1つの依存関係は、複数の要素における互いに関連する1つの要素ペアを指示することと、該依存関係集合に基づいて、複数の要素の因果関係シーケンスを決定することであって、該因果関係シーケンスは、互いに関連する要素ペアにおける一方の要素が他方の要素の原因であることを指示することと、を含む。

40

【0007】

本開示の第3の態様において、コンピュータ可読記憶媒体が提供され、該コンピュータ

50

可読記憶媒体は、デバイスにより実行される場合、該デバイスに本開示の第1の態様に記載の方法を実行させる機器実行可能な指令を記憶している。

【0008】

発明の概要を提供することは、簡略した形態で一連の概念を紹介するためであり、これらは、以下の具体的な実施形態においてさらに詳細に説明される。発明の概要は、本開示の肝心となる特徴又は必須な特徴を示す意図も、本開示の範囲を限定する意図もない。本開示のその他の特徴は、以下の説明により容易に理解できるはずである。

【図面の簡単な説明】

【0009】

以下の開示内容及び特許請求の範囲から、本発明の目的、利点及び他の特徴はより明らかになる。ここでは、例を示す目的を基にして、図面を参照しながら、好ましい実施形態についての制限的ではない記述を示す。

10

【0010】

【図1A】本開示の実施形態に係る、データ処理用の例示的なシステムを示すブロック図である。

【図1B】本開示の実施形態に係る、データ処理用の例示的なシステムを示すブロック図である。

【0011】

【図2】本開示の実施形態に係る、複数の要素の間の因果関係を決定するための模式図である。

20

【0012】

【図3】本開示の実施形態に係る、例示的な方法を示すフローチャートである。

【0013】

【図4】本開示の実施形態に係る、例示的な方法を示すフローチャートである。

【0014】

【図5】本開示の実施形態に係る、例示的な方法を示すフローチャートである。

【0015】

【図6】本開示の実施形態に係る、例示的な方法を示すフローチャートである。

【0016】

【図7】本開示の実施形態を実施できる例示的な装置を模式的に示すブロック図である。

30

【0017】

各図面において、同一又は対応する符号は同一又は対応する部分を示す。

【発明を実施するための形態】

【0018】

以下、図面を参照しながら、本開示の実施形態をさらに詳細に説明する。本開示は、図面により本開示のいくつかの実施形態を示しているが、様々な形態により実施可能であり、ここで説明する実施形態に限定されるように解釈されるべきでない。逆にこれらの実施形態を提供することは本開示をより徹底的、かつ完全に理解するためであることを理解すべきである。本開示の図面及び実施形態は例示的なものにすぎず、本開示の保護範囲を限定するためのものではないと理解すべきである。

40

【0019】

本開示の実施形態の記述において、用語「含む」及びそれに類似する用語は、開放式であり、即ち「...を含むが、これらに限定されない」と解釈されるべきである。用語「...に基づいて」は、「少なくとも部分的に基づく」と解釈される。用語「1つの実施形態」又は「該実施形態」は、「少なくとも1つの実施形態」と解釈されるべきである。用語「第1」、「第2」などは、異なる又は同一なオブジェクトを指してもよい。以下では、他の明示的及び非明示的な定義も含まれる場合がある。

【0020】

本開示の実施形態において、用語「因果構造」(causal structure)とは、通常、システムにおける各要素の間の因果関係を記述する構造を指し、本明細書にお

50

いて、「因果関係シーケンス」とも呼ばれる。用語「要素」は「変量」とも呼ばれる。用語「観測サンプル」は、直接的に観測可能な複数の要素の1組の観測値を指し、直接的に観測可能な要素は「観測変数」とも呼ばれる。

【0021】

上記のように、実際の生活において、多くの観測変数の間に存在する因果関係を迅速かつ正確に発見することは望まれている。

【0022】

顧客サービスの分野において、どの要素が顧客の通信キャリアについての満足度に影響するかを決定するために、顧客の消費行為データ（例えば、顧客のレベル、毎月を使用したネットの通信量、無料の通信量の比率、毎月を使用したネットの通信量の総額など）、満足度調査データ及びキャリア戦略データを大量に収集してもよい。収集した各タイプのデータは、1種の要素（又は変数）の観測値とも呼ばれる。これらの要素の間に存在する因果関係を発見することで、顧客満足度に影響する1つ又は複数の要素を決定することができる。さらに、当該1つ又は複数の要素の観測値を変更し、或いは当該1つ又は複数の要素について対応する戦略を立てることで、顧客の通信キャリアについての満足度を向上させることができる。

10

【0023】

健康の分野において、患者の血圧に影響する要素を決定するために、例えば、心拍数、心拍出量、アレルギー指数、総末梢血管抵抗、カテコールアミンの放出、血圧など、患者の一連の生理的な指標（即ち、一連の要素の観測値）を大量に収集してもよい。これらの生理的な指標の間に存在する因果関係を発見することで、患者の血圧に影響する生理的な指標（即ち、要素）を決定することができる。さらに、当該生理的な指標に影響し、或いは当該生理的な指標について対応する戦略を立てることで、患者の血圧を安定に保つことができる。

20

【0024】

商品販売の分野において、目標商品（例えば、傘）の売上げに影響する要素を決定するために、外部要素データ（例えば、天気、季節、温度、日付、店舗の大きさなど）、当該商品の販売データ（例えば、当該商品の売上げ、当該商品の価格など）、及び1つ又は複数の関連商品（例えば、アイスクリーム）の販売データなどを収集してもよい。収集した各タイプのデータを1種の要素の観測値とする。これらの要素の間に存在する因果関係を発見することで、目標商品の売上げに影響する1つ又は複数の要素を決定することができる。さらに、当該1つ又は複数の要素の観測値を変更し、或いは当該1つ又は複数の要素について対応する戦略を立てることで、目標商品の売上げを増加させることができる。

30

【0025】

ソフトウェア開発の分野において、故障率及びノ又はソフトウェアの開発周期に影響する要素を決定するために、ソフトウェア開発の各種の要素の情報を収集してもよく、例えば、ソフトウェア開発の全体情報（例えば、開発周期、開発に投入されたリソースなど）及びソフトウェア開発の各々の段階の情報を含むが、これらに限定されない。ソフトウェア開発の各々の段階の情報は、例えばアーキテクチャ段階の情報（例えば、ソフトウェアアーキテクチャ方法、ソフトウェアアーキテクチャ階層の数など）、コーディング段階の情報（例えば、コード長、関数の数、プログラミング言語、モジュールの数など）、テスト段階の情報（例えば、ユニットテストの正確率又は故障率、ブラックボックステストの正確率又は故障率、ホワイトボックステストの正確率又は故障率など）、ソフトウェアの発行後の稼働段階の情報（例えば、稼働段階の正確率又は故障率など）を含んでもよい。収集した各タイプのデータを1種の要素の観測値とする。これらの要素の間に存在する因果関係を発見することで、ソフトウェアの開発周期及びノ又は故障率に影響する1つ又は複数の要素を決定することができる。さらに、当該1つ又は複数の要素の観測値を変更し、或いは当該1つ又は複数の要素について対応する戦略を立てることで、ソフトウェアの開発周期及びノ又は故障率を低下させることができる。

40

【0026】

50

いくつかの従来の解決手段は、主に、少ない要素（例えば、100個以下の要素）を有するシステムについて、制約に基づく方法又はスコアに基づく方法で変数空間全体を検索して可能な因果関係を発見する。例えば、制約に基づく方法では、一般的に、複数の要素に条件付き独立性の検定を行うことで、それらの因果構造を発見する。しかしながら、検索する必要がある要素の数が多い場合に、条件付き独立性の検定の結果が不確実になる。また、因果構造の発見は、常に、何らかの要素を含まないから全ての他の要素を含むまでの条件集合を利用して任意の2つの要素の間の依存関係を検定する必要があるため、膨大な計算オーバーヘッドをもたらす。スコアに基づく方法では、一般的に、因果構造とサンプルデータとの間のマッチング度を判定するスコアを最適化することで因果構造を発見する。しかしながら、検索空間が指数関数以上の速度で増加するため、これらの解決手段は、常に、多くの要素（例えば、数百から千以上になる）についての因果構造（「高次元因果構造」とも呼ばれる）の発見に適用しにくい。

10

【0027】

本開示の実施形態によれば、データ処理のための解決手段を提供する。該解決手段は、高次元因果構造の発見を迅速かつ正確に実現することができるため、上記問題及び/又は他の潜在的な問題を解決することができる。以下、上記例示的なシーンに合わせて本開示の各実施形態を詳細に説明する。これは単に説明することを目的とし、如何なる形態で本発明の範囲を制限するためではないと理解すべきである。

【0028】

図1Aは、本開示の実施形態に係る、データ処理用のシステム100を示す例示的なブロック図である。システム100は、例えば、複数の要素の間の因果関係を発見することができる。図1Aに示すシステム100は、本開示の実施形態における1種の例を実現するものに過ぎず、本開示の範囲を制限するためのものではないと理解すべきである。本開示の実施形態は、他のシステム又はアーキテクチャにも同様に適用される。

20

【0029】

図1Aに示すように、システム100は、因果関係決定装置120を含んでもよい。因果関係決定装置120は、複数の要素に関する観測サンプル集合110を受信するとともに、その中から複数の要素の間の因果関係を指示する因果関係シーケンス130を決定することができる。好ましくは、いくつかの実施形態において、システム100は、さらに複数の要素に関する観測サンプル集合110を採集するための観測サンプル採集装置（図1Aにおいて未図示）を含んでもよい。観測サンプル採集装置は、観測サンプル集合110を取得するように、複数の要素の観測値をリアルタイムに採集し、或いは、定期的に又は不定期的に採集することができる。いくつかの実施形態において、観測サンプル採集装置は、それぞれ異なるタイプの要素の観測値を採集するための1つ又は複数の採集ユニットを含んでもよい。

30

【0030】

観測サンプル集合110は、1つ又は複数の目標要素に関する複数の要素の観測サンプルを含んでもよい。観測サンプル集合110は、例えば、 $X = \{x_1, x_2, \dots, x_D\} \in \mathbb{R}^{N \times D}$ として示されてよく、ただし、Nは観測サンプル集合110における観測サンプルの総数を示し、Dは要素の総数を示し、観測サンプル集合110における各々の観測サンプルはD個の要素を有する1組の観測値を含む。本明細書において、D個の要素を有する集合は、 $V = \{v_1, v_2, \dots, v_D\}$ として示され、かつ v_i （ただし、 $1 \leq i \leq D$ ）は、D個の要素におけるi番目の要素を示す。ベクトル $x_i \in \mathbb{R}^N$ （ただし、 $1 \leq i \leq D$ ）は、i番目の要素のN個の観測値を示す。例えば、 $x_{i,n}$ （ただし、 $1 \leq i \leq D$ 、かつ $1 \leq n \leq N$ ）は、i番目の要素のn番目の観測値を示す。

40

【0031】

上記通信キャリアの顧客満足度に関するシーンを例とする場合に、例えば目標要素は「顧客満足度」であり、要素集合Vは、顧客属性に関連する要素（例えば、顧客レベル、顧客番号など）、顧客行為に関する要素（例えば、毎月を使用したネットの通信量、無料の通信量の比率、毎月を使用したネットの通信量の総額など）、顧客のフィードバックに関

50

連する要素（例えば、クレーム数、顧客満足度）及び顧客について立てた戦略要素（例えば、パケットオーバーの提示回数、タイミングなど）における1種又は複数種などを含んでもよい。要素である「顧客レベル」（即ち、 v_i ）を例とする場合に、ベクトル x_i は、N人の顧客のレベルからなり、かつ $x_{i,n}$ は、N人の顧客におけるn人目の顧客のレベルを示してもよい。要素である「毎月を使用したネットの通信量」（即ち、 v_i ）を例とする場合に、ベクトル x_i は、N人の顧客が毎月を使用したネットの通信量からなり、かつ $x_{i,n}$ は、N人の顧客におけるn人目の顧客が毎月を使用したネットの通信量を示してもよい。因果関係シーケンス130は、例えば、顧客レベル、毎月を使用したネットの通信量、無料の通信量の比率、毎月を使用したネットの通信量の総額、顧客満足度などの要素の間の因果関係を指示し、例えば、目標要素である「顧客満足度」の原因がどの要素であるかを指示することができる。

10

【0032】

上記患者の血圧に関連するシーンを例とする場合に、例えば、目標要素は「血圧」であり、要素集合Vは、心拍数、心拍出量、アレルギー指数、総末梢血管抵抗、カテコールアミンの放出、血圧などを含んでもよい。要素である「心拍数」（即ち、 v_i ）を例とする場合に、ベクトル x_i は、N人の患者の心拍数からなり、かつ $x_{i,n}$ は、N人の患者におけるn人目の患者の心拍数を示してもよい。要素である「心拍出量」（即ち、 v_i ）を例とする場合に、ベクトル x_i は、N人の患者の心拍出量からなり、かつ $x_{i,n}$ は、N人の患者におけるn人目の患者の心拍出量を示してもよい。因果関係シーケンス130は、例えば、心拍数、心拍出量、アレルギー指数、総末梢血管抵抗、カテコールアミンの放出、血圧などの要素の間の因果関係を指示し、例えば、目標要素である「血圧」の原因がどの要素であるかを指示することができる。

20

【0033】

上記商品販売のシーンを例とする場合に、例えば、目標要素は「目標商品の売上げ」であり、要素集合Vは、外部要素（例えば、天気、季節、温度、日付、店舗の大きさなど）、目標商品（例えば、傘）の販売行為に関する要素（例えば、目標商品の売上げ、目標商品の価格など）、1つ又は複数の関連商品（例えば、アイスクリーム）の販売行為に関する要素（例えば、関連商品の売上げ、関連商品の価格など）及び目標商品についての販売戦略要素（例えば、プロモーションの回数、頻度など）における1種又は複数種を含んでもよい。要素である「温度」（即ち、 v_i ）を例とする場合に、ベクトル x_i は、N日の温度からなり、かつ $x_{i,n}$ はn日目の温度を示してもよい。要素である「目標商品の売上げ」（即ち、 v_i ）を例とする場合に、ベクトル x_i は、N日の傘の売上げからなり、かつ $x_{i,n}$ は、n日目の傘の売上げを示してもよい。因果関係シーケンス130は、例えば、天気、季節、温度、日付、店舗の大きさ、目標商品の売上げ、目標商品の価格、関連商品の売上げ、関連商品の価格などの要素の間の因果関係を指示し、例えば、目標要素である「目標商品の売上げ」の原因がどの要素であるかを指示することができる。

30

【0034】

上記ソフトウェア開発のシーンを例とする場合に、例えば、目標要素は、「ソフトウェアの開発周期」又は「ソフトウェアの稼働段階の故障率」であり、要素集合Vは、ソフトウェア開発の全体要素（例えば、開発周期、開発に投入されたりソースなど）及びソフトウェア開発の各々の段階の要素における1種又は複数種を含んでもよい。ソフトウェア開発の各々の段階の要素は、例えばアーキテクチャ段階の要素（例えば、ソフトウェアアーキテクチャ方法、ソフトウェアアーキテクチャ階層の数など）、コーディング段階の要素（例えば、コード長、関数の数、プログラミング言語、モジュールの数など）、テスト段階の要素（例えば、ユニットテストの正確率又は故障率、ブラックボックステストの正確率又は故障率、ホワイトボックステストの正確率又は故障率など）、ソフトウェアの発行後の稼働段階の要素（例えば、稼働段階の正確率、稼働段階の故障率など）を含んでもよい。要素である「開発周期」（即ち、 v_i ）を例とする場合に、ベクトル x_i は、N個のソフトウェア製品の開発周期からなり、かつ $x_{i,n}$ は、n個目のソフトウェア製品の開発周期を示してもよい。要素である「コード長」（即ち、 v_i ）を例とする場合に、ベクト

40

50

ル x_i は、N個のソフトウェア製品のコード長からなり、かつ $x_{i,n}$ は、n個目のソフトウェア製品のコード長を示す。因果関係シーケンス130は、例えば、ソフトウェアの開発周期、開発に投入されたリソース、アーキテクチャ方法、アーキテクチャ階層の数、コード長、関数の数、プログラミング言語、モジュールの数、ユニットテストの正確率又は故障率、ブラックボックステストの正確率又は故障率、ホワイトボックステストの正確率又は故障率、稼働段階の正確率、稼働段階の故障率などの要素の間の因果関係を指示する。例えば、目標要素である「開発周期」の原因がどの要素であるか、目標要素である「稼働段階の故障率」の原因がどの要素であるかを指示することができる。

【0035】

図1Aに示すように、因果関係決定装置120は、例えば、依存関係決定ユニット121及び因果関係決定ユニット122を含んでもよい。いくつかの実施形態において、依存関係決定ユニット121は、観測サンプル集合110に基づいて複数の要素の間に存在する依存関係集合を決定することができる。該依存関係集合における各依存関係は、複数の要素における互いに関連する1つの対応する要素ペアを指示する。いくつかの実施形態において、因果関係決定ユニット122は、依存関係決定ユニット121により決定された依存関係集合に基づいて、要素関係シーケンス130を決定することができる。要素関係シーケンス130は、依存関係が存在する要素ペアの間の因果関係を指示することができる（即ち、一方の要素が他方の要素の原因である）。

10

【0036】

システム100に含まれるこれらの装置及び/又は装置におけるユニットは、例示的なものにすぎず、本開示の範囲を制限するためのものではないと理解すべきである。システム100は、さらに、示されない付加的な装置及び/又はユニットを含んでもよいと理解すべきである。例えば、いくつかの実施形態において、システム100は、さらに、因果関係シーケンス130の態様を表すための因果関係表現装置（未図示）を含んでもよい。

20

【0037】

いくつかの実施形態において、因果関係表現装置は、視覚や聴覚などの異なる形態で因果関係シーケンス130の態様を表現することができる。例えば、因果関係表現装置は、グラフ、マップ、テキストなどの形態で因果関係シーケンス130を表現することができる。いくつかの実施形態において、因果関係表現装置は、因果関係シーケンス130の全ての態様、即ち、全ての要素の間の因果関係を表現することができる。代替的に、いくつかの実施形態において、因果関係表現装置は、因果関係シーケンス130の一部の態様のみを表現し、例えば、1つ又は複数の目標要素に関連する因果関係のみを表現することができる。いくつかの実施形態において、目標要素の原因が複数の要素を含む場合に、因果関係表現装置は、さらに、複数の要素の対応する重要程度を表現し、例えば、異なる色及び/又は異なる重要程度を示す数値などの形態で複数の要素の対応する重要程度を表現することができる。本開示の実施形態は、その態様について制限されない。

30

【0038】

図1Bは、本開示の実施形態に係る、データ処理用のシステム105を示す例示的なブロック図である。システム105は、例えば、図1Aに示すような因果関係シーケンス130を適用し、最適化することができる。図1Bに示すシステム105は、本開示の実施形態における1種の例を実現するものに過ぎず、本開示の範囲を制限するためのものではないと理解すべきである。本開示の実施形態は、他のシステム又はアーキテクチャにも同様に適用される。

40

【0039】

図1Bに示すように、システム105は、観測サンプル影響装置140を含んでもよい。観測サンプル影響装置140は、因果関係シーケンス130に基づいて、複数の要素から目標要素の原因となる少なくとも1つの要素を決定することができる。観測サンプル影響装置140は、少なくとも1つの要素の観測値を変更することで目標要素の観測値に影響することにより、変更後の観測サンプル集合150を取得することができる。変更後の観測サンプル集合150における少なくとも1つの観測サンプルは、少なくとも1つの要

50

素の変更後の観測値を含む。

【0040】

上記通信キャリアの顧客満足度に関するシーンを例とする場合に、目標要素は、例えば、「顧客満足度」であり、因果関係シーケンス130は、例えば、目標要素である「顧客満足度」の原因がどの要素（例えば、パケットを使い切る前の提示、お得なパケットなど）であるかを指示することができる。観測サンプル影響装置140は、例えば、これらの要素の観測値に影響し、変更し、及び/又はこれらの要素について対応する戦略を立てる（例えば、パケットを使い切る前に顧客についてより多くの提示を提供し、顧客についてより多くのお得なパケットを提供する）ことで、顧客の通信キャリアについての満足度を向上させることができる。

10

【0041】

上記患者の血圧に関連するシーンを例とする場合に、目標要素は、例えば、「血圧」であり、因果関係シーケンス130は、例えば、目標要素である「血圧」の原因がどの生理的な指標であるかを指示することができる。観測サンプル影響装置140は、例えば、これらの生理的な指標に影響し、変更し、及び/又はこれらの生理的な指標について対応する戦略を立てることで、患者の血圧を安定に保つことができる。

【0042】

上記商品販売のシーンを例とする場合に、目標要素は、例えば、「傘の売上げ」であり、因果関係シーケンス130は、例えば、目標要素である「傘の売上げ」の原因がどの要素（例えば、天気、販売される傘の数など）であるかを指示することができる。観測サンプル影響装置140は、例えば、これらの要素に影響し、変更し、及び/又はこれらの要素について対応する戦略を立てる（例えば、雨が降るときに販売に供される傘の数を増やす）ことで、目標商品である傘の売上げを増加させることができる。

20

【0043】

上記ソフトウェア開発のシーンを例とする場合に、目標要素は例えば「開発周期」であり、因果関係シーケンス130は、例えば、目標要素である「開発周期」の原因がどの要素（例えば、アーキテクチャ階層の数、プログラミング言語など）であるかを指示することができる。観測サンプル影響装置140は、例えば、これらの要素に影響し、変更し、及び/又はこれらの要素について対応する戦略を立てる（例えば、ソフトウェアアーキテクチャの複雑度を低下し、よりフレンドリーなプログラミング言語を使用するなど）ことで、ソフトウェア開発の周期を短縮することができる。また、例えば、目標要素は「稼働段階のソフトウェアの故障率」であってもよく、因果関係シーケンス130は、例えば、目標要素である「稼働段階のソフトウェアの故障率」の原因がどの要素（例えば、コード長、モジュールの数など）であるかを指示することができる。観測サンプル影響装置140は、例えば、これらの要素に影響し、変更し、及び/又はこれらの要素について対応する戦略を立てる（例えば、コード長を短くし、モジュールの数を減少するなど）ことで、稼働段階のソフトウェアの故障率を低下させることができる。

30

【0044】

図1Bに示すように、システム105は、因果関係最適化装置160を含んでもよい。因果関係最適化装置160は、変更後の観測サンプル集合150に基づいて、因果関係シーケンス130を最適化することができるため、因果関係シーケンス130の正確性を高める。いくつかの実施形態において、因果関係最適化装置160は、変更後の観測サンプル集合150に基づいて、複数の要素の間の因果関係（例えば、因果関係決定装置120の実行過程に類似すること）を改めて発見することができるため、最適化後の因果関係シーケンスを取得する。この形態によれば、本開示の実施形態は、因果発見の正確性及びロバスト性を一層向上させることができる。

40

【0045】

図1Aに示す因果関係決定装置120、図1Bに示す観測サンプル影響装置140及び因果関係最適化装置160は、互いに分離されることが示されているが、これは説明を目的とすることにすぎず、本開示の範囲を限定するためではないと理解すべきである。いく

50

つかの実施形態において、図 1 A に示す因果関係決定装置 1 2 0、図 1 B に示す観測サンプル影響装置 1 4 0 及び因果関係最適化装置 1 6 0 は、同一の物理装置或いは複数の異なる物理装置で実現されることができ、いくつかの実施形態において、図 1 A に示す因果関係決定装置 1 2 0 と図 1 B に示す因果関係最適化装置 1 6 0 は、同一の装置として実現することができる。本開示の実施形態は、その態様について制限されない。

【 0 0 4 6 】

図 2 は、本開示の実施形態に係る、複数の要素の間の因果関係を決定するための模式図である。簡略化のため、かつ説明の便宜上、図 2 において、観測サンプル集合 1 1 0 に係る要素（即ち、観測変数）の数（即ち、 D ）が 5 であると仮定する。図 2 に示すように、観測サンプル集合 1 1 0 は、要素 v_1 、 v_2 、 v_3 、 v_4 及び v_5 についての複数の観測サンプルを含む。依存関係決定ユニット 1 2 1 は、観測サンプル集合 1 1 0 に基づいて、要素 v_1 、 v_2 、 v_3 、 v_4 及び v_5 の間に存在する依存関係集合を決定することができ、スケルトン図 2 1 0 として示される。例えば、依存関係集合 2 1 0 は、要素 v_1 と v_2 が互いに関連し、要素 v_2 と v_4 が互いに関連し、要素 v_3 と v_4 が互いに関連し、かつ要素 v_3 と v_5 が互いに関連することを指示する。因果関係決定ユニット 1 2 2 は、依存関係集合 2 1 0 に基づいて、要素 v_1 、 v_2 、 v_3 、 v_4 及び v_5 の因果関係シーケンスを決定することができ、例えば、有向非巡回グラフ 1 3 0 として示される。例えば、因果関係シーケンス 1 3 0 は、要素 v_2 が要素 v_1 の原因であり（例えば、エッジ $v_2 \rightarrow v_1$ で示される）、要素 v_2 が要素 v_4 の原因であり（例えば、エッジ $v_2 \rightarrow v_4$ で示される）、要素 v_3 が要素 v_4 の原因であり（例えば、エッジ $v_3 \rightarrow v_4$ で示される）、かつ要素 v_3 が要素 v_5 の原因である（例えば、エッジ $v_3 \rightarrow v_5$ で示される）ことを指示する。

【 0 0 4 7 】

図 3 は、本開示の実施形態に係る、複数の要素の間の因果関係を決定するための方法 3 0 0 のフローチャートである。例えば、方法 3 0 0 は、図 1 A に示す因果関係決定装置 1 2 0 により実行されてもよい。方法 3 0 0 は、さらに、示されない付加的な動作を含んでもよく、及び / 又は示されているいくつかの動作を省略してもよいと理解すべきである。本開示の範囲は、その態様について制限されない。

【 0 0 4 8 】

ブロック 3 1 0 において、因果関係決定装置 1 2 0 は、複数の要素に関する観測サンプル集合（例えば、図 1 A 及び図 2 に示す観測サンプル集合 1 1 0）を取得する。観測サンプル集合における 1 つの観測サンプルは、複数の要素の対応する観測値を含む。

【 0 0 4 9 】

ブロック 3 2 0 において、因果関係決定装置 1 2 0（例えば、依存関係決定ユニット 1 2 1）は、観測サンプル集合に基づいて複数の要素の間に存在する依存関係集合を決定する（例えば、図 2 に示すスケルトン図 2 1 0）。依存関係集合における 1 つの依存関係は、複数の要素における互いに関連する 1 つの要素ペアを指示する。

【 0 0 5 0 】

いくつかの実施形態において、依存関係集合を決定するために、因果関係決定装置 1 2 0 は、複数の要素における任意の 2 つの要素の対応する観測値に基づいて、これらの 2 つの要素の間の相関係数を推定することができる。例えば、相関係数は、Spearman 相関係数又は Kendall 相関係数のうちのいずれか 1 つであってもよい。推定された 2 つずつの要素の相関係数に基づいて、因果関係決定装置 1 2 0 は、相関係数行列 S を確立することができる。例えば、要素の総数が D であると仮定すると、 S は、 $D \times D$ の行列である。行列 S における j 行目 k 列目の要素が S_{jk} であると仮定すると、 S_{jk} は、以下のように決定される。

【数 1】

$$S_{jk} = \begin{cases} 2 \sin\left(\frac{\pi}{6} \rho_{jk}\right), & j \neq k \\ 1, & j = k \end{cases} \quad \text{又は} \quad S_{jk} = \begin{cases} \sin\left(\frac{\pi}{2} \tau_{jk}\right), & j \neq k \\ 1, & j = k \end{cases} \quad (1)$$

10

20

30

40

50

ここで、
【数 2】

$$\rho_{jk}$$

は、D個の要素におけるj番目の要素とk番目の要素とのSpearman相関係数を示し、

【数 3】

$$\tau_{jk}$$

は、D個の要素におけるj番目の要素とk番目の要素とのKendallの相関係数を示す。Spearman相関係数及びKendall相関係数の計算は、当業者に知られており、ここでは説明を省略する。また、任意の既知又は将来開発の方法又は手段を利用して2つの要素の間の相関係数を計算することができ、Spearman相関係数及びKendall相関係数に限定されるものではない。Spearman相関係数及びKendall相関係数は、相関係数の例に過ぎず、本開示の範囲を限定することを意図していないと理解すべきである。

【0051】

いくつかの実施形態において、因果関係決定装置120は、推定された相関係数行列Sに基づいて、依存関係集合（即ち、スケルトン図210）を決定するための目標関数（本明細書において、「第1の目標関数」とも呼ばれる）を確立することができる。因果関係決定装置120は、第1の目標関数を最小化することで、依存関係集合を決定することができる。いくつかの実施形態において、例えば、因果関係決定装置120は、グラフィカルLassoアルゴリズムに基づいて、要素の間の対応する依存関係を示す精度行列を学習することができる。例えば、は、以下のように決定される。

【数 4】

$$\Omega = \underset{\Omega \geq 0}{\operatorname{argmin}} \{ \operatorname{tr}(S\Omega) - \log|\Omega| + \lambda \sum_{j \neq k} |\Omega_{jk}| \} \quad (2)$$

ここで、
【数 5】

$$\lambda$$

は、予め定義されている係数である。決定された行列において、j行目k列目の要素

【数 6】

$$\Omega_{jk}$$

が0であれば、D個の要素におけるj番目の要素とk番目の要素とが関連しないことを示

し、要素

【数 7】

$$\Omega_{jk}$$

が 0 でなければ、D 個の要素における j 番目の要素と k 番目の要素とが互いに関連する（必ずしも因果関係を有するとは限らない）ことを示す。行列 M で依存関係集合を示すと仮定すると、行列 M における要素

10

【数 8】

$$M_{jk} = \Omega_{jk} \neq 0$$

である。この形態によれば、因果関係決定装置 1 2 0 は、観測サンプル集合に基づいて、複数の要素の間の依存関係集合を決定することができ、図 2 におけるスケルトン図 2 1 0 に示すとおりである。以下では、「依存関係集合」と「スケルトン図」は、交換可能に使用される。

20

【0 0 5 2】

代替的に、いくつかの実施形態において、因果関係決定装置 1 2 0 は、複数の要素に条件付き独立性の検定を行うことで、依存関係集合 M を決定することができる。条件付き独立性の検定は、所定の条件集合において、2 つの要素が互いに独立しているか否かを決定することができる。従来の条件付き独立性の検定において、2 つの要素が独立しているか否かを判断する場合に、全ての他の要素の任意の組み合わせを上記条件集合とする必要がある。要素の数（即ち、D）が多い場合に、計算オーバーヘッドが非常に大きい。条件集合に多くの他の要素が含まれる場合に、上記 2 つの要素は、互いに独立していると決定されやすい。高次元因果構造の発見を迅速かつ正確に実現するために、いくつかの実施形態において、複数の要素に条件付き独立性の検定を行うことで、依存関係集合を決定する場合に、因果関係決定装置 1 2 0 は、条件集合における他の要素の数を 1 に制限することができる。この形態によれば、条件付き独立性の検定の計算オーバーヘッドを低減するだけでなく、互いに独立していると判定された要素ペアの数を低減することができるため、後続の因果関係シーケンスの発見に役立つ。

30

【0 0 5 3】

この形態によれば、依存関係集合を決定することで、本開示の実施形態は、検索対象の変数空間の大きさを減少させることにより、多くの要素の間の因果関係を迅速に発見することができる。

40

【0 0 5 4】

ブロック 3 3 0 において、因果関係決定装置 1 2 0（例えば、因果関係決定ユニット 1 2 2）は、依存関係集合に基づいて、複数の要素の因果関係シーケンスを決定する。因果関係シーケンスは、互いに関連する要素ペアにおける一方の要素が他方の要素の原因であることを指示することができる。

【0 0 5 5】

いくつかの実施形態において、因果関係シーケンスを決定するために、因果関係決定装置 1 2 0 は、依存関係集合により指示された複数の互いに関連する要素ペアについて、互いに関連する各要素ペアにおける一方の要素から他方の要素への影響を決定することができる。いくつかの実施形態において、各要素ペアにおける一方の要素から他方の要素への

50

影響を決定するために、因果関係決定装置 120 は、所定の分布（例えば、ガウス分布又は他の分布）に基づいて、第 2 の目標関数を確立して、第 2 の目標関数を最小化することで、各要素ペアにおける一方の要素から他方の要素への影響を決定することができる。第 2 の目標関数は、例えば、以下の 2 つの要素に基づいて確立することができる。第 1、発見された因果構造が観測データサンプルについて、良好なデータフィッティング度を持つようにし、第 2、発見された因果構造が依然として疎であるようにする。

【数 9】

いくつかの実施形態において、要素の総数が D であると仮定し、複数の要素ペアのそれぞれについて決定された影響を行列 B で示すと、 B は、以下のように決定される。

$$B = \underset{B=[\beta_1, \dots, \beta_D]}{\operatorname{argmin}} \sum_{i=1}^D \log \left(\sum_{n=1}^N (x_{i,n} - X_{-i,n} \beta_i)^2 \right) + \frac{1}{N} \log(N) |B|_0$$

$$\text{s.t. } G_{\{\beta_1, \dots, \beta_D\}} \in \text{DAG}; G_{\{\beta_1, \dots, \beta_D\}} \subset M \quad (3)$$

ここで、 $X = \{x_1, x_2, \dots, x_D\}$ $\mathbb{R}^{N \times D}$ は、観測サンプル集合を示し、 N は、 X における観測サンプルの総数を示し、 D は、要素の総数を示す。ベクトル x_i \mathbb{R}^N （ただし、 $1 \leq i \leq D$ ）は、 i 番目の要素（即ち、要素 v_i ）の N 個の観測値を示す。 $x_{i,n}$ （ただし、 $1 \leq i \leq D$ 、かつ $1 \leq n \leq N$ ）は、 i 番目の要素（即ち、要素 v_i ）の n 番目の観測値を示す。ベクトル

【数 10】

$$\beta_i \in \mathbb{R}^{D-1} \quad (\text{ただし、} 1 \leq i \leq D)$$

は、要素 v_i についての各要素の対応する影響を示す。例えば、

【数 11】

$$\beta_{ij} \neq 0$$

であれば、要素 v_j が要素 v_i の直接的原因である可能性があることを示し、

【数 12】

$$\beta_{ij} = 0$$

であれば、要素 v_j が要素 v_i に影響しないことを示すため、要素 v_i の直接的原因である可能性がない。

【数 13】

$$|B|_0$$

10

20

30

40

50

は、行列 B における非ゼロ元素の総数を示し、因果構造の分散度についての推定であり、かつ

【数 1 4】

$$|B|_0 = \sum_{i=1}^D |\beta_i|_0$$

である。制約条件

【数 1 5】

$$G_{\{\beta_1, \dots, \beta_D\}} \in \text{DAG}$$

は、決定しようとする因果構造が有向非巡回グラフであることを示し、かつ制約条件

【数 1 6】

$$G_{\{\beta_1, \dots, \beta_D\}} \subset M$$

は、決定しようとする因果構造が以前に決定されたスケルトン図 M（例えば、図 2 に示すスケルトン図 2 1 0）のサブ集合であることを示す。

【0 0 5 7】

いくつかの実施形態において、因果関係決定装置 1 2 0 は、複数の要素ペアについて決定された対応する影響と観測サンプル集合に基づいて、因果関係シーケンス 1 3 0 を決定することができる。例えば、図 4 は、本開示の実施形態に係る、因果関係シーケンスを決定するための方法 4 0 0 のフローチャートである。方法 4 0 0 は、図 1 A に示す因果関係決定装置 1 2 0 によって実行されてもよい。方法 4 0 0 は、さらに、示されない付加的な動作を含んでもよく、及び / 又は示されているいくつかの動作を省略してもよいと理解すべきである。本開示の範囲は、その態様について制限されない。

【0 0 5 8】

ブロック 4 1 0 において、因果関係決定装置 1 2 0 は、履歴因果関係シーケンスと履歴因果関係のスコアを取得することができる。ここでは、履歴因果関係シーケンスを Q_S で示し、履歴因果関係のスコアを $f(Q_S)$ で示すと仮定する。

【0 0 5 9】

いくつかの実施形態において、初期状況で、因果関係決定装置 1 2 0 は、履歴因果関係シーケンスを空シーケンスに初期化することができ、即ち $Q_S = \{\}$ である。因果関係決定装置 1 2 0 は、空シーケンスに対応する初期因果関係のスコアを履歴因果関係のスコアとして決定することができ、即ち、

【数 1 7】

10

20

30

40

50

$$f(Q_s) = \sum_{i=1}^D SBIC(v_i | V \setminus v_i)$$

ただし、 $SBIC(v_i | V \setminus U) = \min_{\beta_i, S(\beta_i) \subseteq (V \setminus U) \cap S(m_i)} BIC(\beta_i, X_i, X_{-i})$ 、かつ

$$BIC(\beta_i, X_i, X_{-i}) = \log(\sum_{n=1}^N (x_{i,n} - X_{-i,n} \beta_i)^2) + \frac{1}{N} \log(N) |\beta_i|_0 \quad (4)$$

上記式(4)において、因果関係シーケンスが有向非巡回グラフ $G = \{V, E\}$ で示されると仮定し、ただし、 V は、グラフ G における全てのノード（例えば、スケルトン図 210 におけるエッジを有する全てのノード）の集合を示し、 E は、グラフ G における全てのエッジの集合を示す。例えば、図 2 に示す例では、 $V = \{v_1, v_2, v_3, v_4, v_5\}$ である。

【数 18】

$$V \setminus U$$

は、 V におけるノード集合 U 以外の全てのノードを示す（即ち、

【数 19】

$$V \setminus v_i$$

は、 V における

【数 20】

$$v_i$$

以外の全てのノードを示す）。

【数 21】

$$S(\beta_i)$$

は、

【数 22】

$$\beta_i$$

のサポート集合を示し、即ち、ノード

【数 23】

10

20

30

40

50

v_i

の親ノード（即ち、要素

【数 2 4】

v_i

の潜在的な原因を示すノード)の集合である。制約条件

【数 2 5】

$$S(\beta_i) \subseteq (V \setminus U \cap S(m_i))$$

は、集合

【数 2 6】

$S(\beta_i)$

が集合

【数 2 7】

$V \setminus U$

と集合

【数 2 8】

$S(m_i)$

との集合のサブ集合を示し、集合

【数 2 9】

$S(m_i)$

は、スケルトン図M（例えば、スケルトン図 2 1 0）においてノード

【数 3 0】

v_i

10

20

30

40

50

とエッジを有するノードの集合を示す。例えば、図 2 に示す例では、 $f(Q_S) = f(\{ \}) = SBIC(v_1 | v_2) + SBIC(v_2 | (v_1, v_4)) + SBIC(v_3 | (v_4, v_5)) + SBIC(v_4 | (v_2, v_3)) + SBIC(v_5 | v_3)$ である。

【0060】

ブロック 420 において、因果関係決定装置 120 は、履歴因果関係シーケンス Q_S と依存関係集合 (例えば、スケルトン図 210) により指示された複数の要素ペアに基づいて、因果関係シーケンスに追加する可能性がある 1 つ又は複数の候補要素を決定する。いくつかの実施形態において、1 つ又は複数の候補要素は、候補ノード集合 $V \setminus Q_S$ に対応する全ての要素を含んでもよく、候補ノード集合 $V \setminus Q_S$ は、ノード集合 V における、 Q_S に含まれるノード以外の全てのノードを示す。

10

【0061】

1 つ又は複数の候補要素が存在しなければ (即ち、集合 $V \setminus Q_S$ が空である)、ブロック 470 において、因果関係決定装置 120 は、履歴因果関係シーケンス Q_S を、決定された因果関係シーケンス 130 として出力することができる。

【0062】

1 つ又は複数の候補要素が存在すれば、ブロック 440 において、因果関係決定装置 120 は、1 つ又は複数の候補要素に対応する 1 つ又は複数の候補因果関係のスコアを決定することができる。例えば、候補ノード集合 $V \setminus Q_S$ における各候補ノード

【数 3 1】

$$v_i$$

20

(即ち、

【数 3 2】

$$v_i \in V \setminus Q_S$$

30

) について、それに対応する候補因果関係シーケンス

【数 3 3】

$$Q_S' = Q_S \cup v_i$$

であり、かつ該候補因果関係シーケンスのスコアが

【数 3 4】

$$f(Q_S') = f(Q_S) + SBIC(v_i | Q_S) - SBIC(v_i | V \setminus v_i)$$

40

である。

【0063】

ブロック 450 において、因果関係決定装置 120 は、決定された 1 つ又は複数の候補因果関係のスコアに基づいて、1 つ又は複数の候補要素から因果関係シーケンスに追加しようとする候補要素を選択することができる。

【0064】

いくつかの実施形態において、因果関係決定装置 120 は、1 つ又は複数の候補因果関係のスコアから最小の候補因果関係のスコアを決定し、最小の候補因果関係のスコアに関

50

連する候補要素を選択することで、因果関係シーケンス 130 に追加することができる。

【0065】

付加的又は代替的に、いくつかの実施形態において、因果関係シーケンスをより迅速に決定するために、因果関係決定装置 120 は、決定待ちの因果関係シーケンスに関連する制約条件を取得することができる。

【0066】

いくつかの実施形態において、因果関係決定装置 120 は、該制約条件を指示するエキスパート情報を取得し、取得されたエキスパート情報に基づいて該制約条件を決定することができる。図 2 に示す例を例とする場合に、エキスパート情報は、例えば、ノード v_3 がノード v_4 の前にあることを指示することができ、即ち、ノード v_3 に対応する要素がノード v_4 に対応する要素の原因である可能性があるが、ノード v_4 に対応する要素がノード v_3 に対応する要素の原因である可能性がない。

10

【0067】

付加的又は代替的に、いくつかの実施形態において、因果関係決定装置 120 は、履歴因果関係シーケンスとスケルトン図 210 に指示された複数の互いに関連する要素ペアに基づいて、制約条件を決定することができる。例えば、図 2 に示す例において、現在の Q_5 は、ノード v_3 に対応する要素がノード v_4 に対応する要素の原因であると仮定する（即ち、因果関係シーケンス 130 にエッジ $v_3 \rightarrow v_4$ が存在する）。また、スケルトン図 M に基づいて、ノード v_1 及び v_2 が互いに関連し、ノード v_2 及び v_4 が互いに関連し、要素 v_3 及び v_4 が互いに関連し、要素 v_3 及び v_5 が互いに関連することを決定することができる。したがって、ノード v_1 、 v_2 及び v_4 は、強連結ノード集合を構成し、ノード v_3 及び v_5 は、強連結ノード集合を構成する。この場合、例えば、因果関係決定装置 120 は、ノード集合 $\{v_3, v_5\}$ がノード集合 $\{v_1, v_2, v_4\}$ の前にあることを決定することができる。即ち、ノード集合 $\{v_3, v_5\}$ におけるあるノードは、ノード集合 $\{v_1, v_2, v_4\}$ におけるあるノードの原因である可能性があるが、ノード集合 $\{v_1, v_2, v_4\}$ における任意のノードは、ノード集合 $\{v_3, v_5\}$ におけるあるノードの原因である可能性がない。

20

【0068】

いくつかの実施形態において、決定待ちの因果関係シーケンスに関連する制約条件を取得することに応答して、因果関係決定装置 120 は、選択された候補要素の追加が取得された制約条件に合致するように、1つ又は複数の候補要素から因果関係シーケンスに追加しようとする候補要素を選択することができる。例えば、最小の候補因果関係のスコアに関連する候補要素の追加が該制約条件に違反する場合に、因果関係決定装置 120 は、他方の候補要素（例えば、次に小さい候補因果関係のスコアに関連する候補要素）を選択することにより、因果関係シーケンス 130 に追加することができる。

30

【0069】

この形態によれば、制約条件を利用することで、因果関係シーケンスの決定過程において、候補要素の数を制限することにより、因果関係シーケンスをより迅速に決定することができる。

【0070】

ブロック 460 において、候補要素が選択されることに応答して、因果関係決定装置 120 は、履歴因果関係シーケンス Q_5 と履歴因果関係のスコア $f(Q_5)$ を更新することができる。例えば、因果関係決定装置 120 は、履歴因果関係シーケンス Q_5 の代わりに、選択された候補要素に対応する候補因果関係シーケンス Q_5' を利用し、履歴因果関係のスコア $f(Q_5)$ の代わりに、候補因果関係シーケンス Q_5' に対応するスコア $f(Q_5')$ を利用することができる。

40

【0071】

いくつかの実施形態において、因果関係決定装置 120 は、全ての可能な候補要素が検索されるまで、方法 400 におけるブロック 410 - 460 を繰り返して実行することができる（即ち、ブロック 470 までに実行する）。

50

【0072】

図5は、本開示の実施形態に係る、目標要素の観測値に影響するための方法500のフローチャートである。例えば、方法500は、図1Bに示す観測サンプル影響装置140により実行されてもよい。いくつかの実施形態において、方法500は、方法300の後に実行されてもよい。方法500は、さらに、示されない付加的な動作を含んでもよく、及び/又は示されているいくつかの動作を省略してもよいと理解すべきである。本開示の範囲は、その態様について制限されない。

【0073】

ブロック510において、観測サンプル影響装置140は、因果関係シーケンスに基づいて、目標要素の原因となる少なくとも1つの要素を複数の要素から決定する。次に、ブロック520において、観測サンプル影響装置140は、少なくとも1つの要素の観測値を変更することで、目標要素の観測値に影響する。いくつかの実施形態において、例えば、観測サンプル影響装置140は、少なくとも1つの要素に影響し、変更し、及び/又は少なくとも1つの要素について対応する戦略を立てることで、目標要素の観測値に影響することができる。

10

【0074】

上記通信キャリアの顧客満足度に関するシーンを例とする場合に、目標要素は、例えば、「顧客満足度」である。観測サンプル影響装置140は、因果関係シーケンス130に基づいて、目標要素である「顧客満足度」の原因がどの要素（例えば、パケットを使い切る前の提示、お得なパケットなど）であるかを決定することができる。観測サンプル影響装置140は、さらに、これらの要素に影響し、変更し、及び/又はこれらの要素について対応する戦略を立てる（例えば、パケットを使い切る前に顧客についてより多くの提示を提供し、顧客についてより多くのお得なパケットを提供する）ことで、顧客の通信キャリアについての満足度を向上させることができる。

20

【0075】

上記患者の血圧に関連するシーンを例とする場合に、目標要素は、例えば、「血圧」である。観測サンプル影響装置140は、因果関係シーケンス130に基づいて、目標要素である「血圧」の原因がどの生理的な指標であるかを決定することができる。観測サンプル影響装置140は、さらに、これらの生理的な指標に影響し、変更し、及び/又はこれらの生理的な指標について対応する戦略を立てることで、患者の血圧を安定に保つことができる。

30

【0076】

上記商品販売のシーンを例とする場合に、目標要素は、例えば、「傘の売上げ」である。観測サンプル影響装置140は、因果関係シーケンス130に基づいて、目標要素である「傘の売上げ」の原因がどの要素（例えば、天気、販売される傘の数など）であるかを決定することができる。観測サンプル影響装置140は、さらに、これらの要素に影響し、変更し、及び/又はこれらの要素について対応する戦略を立てる（例えば、雨が降るときに販売に供される傘の数を増やす）ことで、目標商品である傘の売上げを増加させることができる。

【0077】

上記ソフトウェア開発のシーンを例とする場合に、目標要素は、例えば、「開発周期」である。観測サンプル影響装置140は、因果関係シーケンス130に基づいて、目標要素である「開発周期」の原因がどの要素（例えば、アーキテクチャ階層の数、プログラミング言語など）であるかを決定することができる。観測サンプル影響装置140は、さらに、これらの要素に影響し、変更し、及び/又はこれらの要素について対応する戦略を立てる（例えば、ソフトウェアアーキテクチャの複雑度を低下し、よりフレンドリーなプログラミング言語を使用するなど）ことで、ソフトウェア開発の周期を短縮することができる。また、例えば、目標要素は、「稼働段階のソフトウェアの故障率」であってもよい。観測サンプル影響装置140は、因果関係シーケンス130に基づいて、目標要素である「稼働段階のソフトウェアの故障率」の原因がどの要素（例えば、コード長、モジュール

40

50

の数など)であるかを決定することができる。観測サンプル影響装置140は、さらに、これらの要素に影響し、変更し、及び/又はこれらの要素について対応する戦略を立てる(例えば、コード長を短くし、モジュールの数を減少するなど)ことで、稼働段階のソフトウェアの故障率を低下させることができる。

【0078】

図6は、本開示の実施形態に係る、因果関係を最適化するための方法600のフローチャートである。例えば、方法600は、図1Bに示す因果関係最適化装置160により実行されてもよい。いくつかの実施形態において、方法600は、方法500の後に実行されてもよい。方法600は、さらに、示されない付加的な動作を含んでもよく、及び/又は示されているいくつかの動作を省略してもよいと理解すべきである。本開示の範囲は、その態様について制限されない。

10

【0079】

ブロック610において、因果関係最適化装置160は、複数の要素に関する変更後の観測サンプル集合を取得する。いくつかの実施形態において、変更後の観測サンプル集合における少なくとも1つの観測サンプルは、少なくとも1つの要素(例えば、少なくとも1つの要素が目標要素の原因である)の変更後の観測値を含んでもよい。次に、ブロック620において、因果関係最適化装置160は、変更後の観測サンプル集合に基づいて、因果関係シーケンスを最適化することができる。いくつかの実施形態において、例えば、因果関係最適化装置160は、変更後の観測サンプル集合150に基づいて、複数の要素の間の因果関係(例えば、因果関係決定装置120の実行過程に類似すること)を改めて発見することができるため、最適化後の因果関係シーケンスを取得する。この形態によれば、本開示の実施形態は、因果発見の正確性及びロバスト性を一層向上させることができる。

20

【0080】

図7は、本開示の実施形態を実施するのに適した例示的デバイス700の概略ブロック図である。例えば、図1Aに示す因果関係決定装置120、図1Bに示す観測サンプル影響装置140及び/又は因果関係最適化装置160は、デバイス700により実施されてもよい。図に示すように、デバイス700は、リードオンリーメモリ(ROM)702に記憶されたコンピュータプログラム指令又は記憶ユニット708からランダムアクセスメモリ(RAM)703にロードされたコンピュータプログラム指令に基づいて、様々な適切な動作及び処理を実行することができる中央プロセッシングユニット(CPU)701を含む。RAM703には、デバイス700の操作に必要な様々なプログラム及びデータがさらに記憶されてよい。CPU701、ROM702、及びRAM703は、バス704を介して互いに接続されている。入力/出力(I/O)インタフェース705もバス704に接続されている。

30

【0081】

キーボード、マウスなどの入力ユニット706と、様々なタイプのディスプレイ、スピーカなどの出力ユニット707と、磁気ディスク、光ディスクなどの記憶ユニット708と、ネットワークカード、モデム、無線通信トランシーバなどの通信ユニット709とを含むデバイス700における複数のコンポーネントは、I/Oインタフェース705に接続されている。通信ユニット709は、デバイス700がインターネットなどのコンピュータネットワーク及び/又は様々な電気通信ネットワークを介して他のデバイスと情報/データを交換することを許可する。

40

【0082】

プロセッシングユニット701は、例えば、方法300、400、500及び/又は600である上記各々の過程及び処理を実行するように構成されている。例えば、いくつかの実施形態において、方法300、400、500及び/又は方法600は、記憶ユニット708のような機器読み取り可能な媒体に有形的に含まれるコンピュータソフトウェアプログラムとして実現されてよい。いくつかの実施形態において、コンピュータプログラムの一部又は全ては、ROM702及び/又は通信ユニット709を介してデバイス70

50

0 にロード及び/又はインストールされてよい。コンピュータプログラムが R A M 7 0 3 にロードされて C P U 7 0 1 により実行される場合、前述の方法 3 0 0、4 0 0、5 0 0 及び/又は方法 6 0 0 の 1 つ又は複数のステップを実行することができる。

【 0 0 8 3 】

本開示は、システム、方法及び/又はコンピュータプログラム製品であってもよい。コンピュータプログラム製品は、本開示の様々な態様を実行するためのコンピュータ読み取り可能なプログラム指令が格納されているコンピュータ可読記憶媒体を含んでよい。

【 0 0 8 4 】

コンピュータ可読記憶媒体は、指令実行装置によって使用される指令を保持して記憶することができる物理装置であってもよい。コンピュータ可読記憶媒体は、例えば、電気記憶装置、磁気記憶装置、光記憶装置、電磁記憶装置、半導体記憶装置又は前述の任意の適切な組み合わせであってもよいが、これらに限定されない。コンピュータ可読記憶媒体のより具体的な例（非網羅的リスト）は、ポータブルコンピュータディスク、ハードディスク、ランダムアクセスメモリ（R A M）、読み取り専用メモリ（R O M）、消去可能なプログラマブル読み取り専用メモリ（E P R O M 又はフラッシュメモリ）、スタティックランダムアクセスメモリ（S R A M）、ポータブルコンパクトディスク読み取り専用メモリ（C D - R O M）、デジタル多用途ディスク（D V D）、メモリスティック、フロッピー（登録商標）ディスク、指令が記憶されたパンチカード又は凹溝内の凸構造、及び前述の任意の適切な組み合わせなどの機器的符号化装置を含む。ここで使用されるコンピュータ可読記憶媒体は、無線電波や自由に伝播される他の電磁波、導波管や他の伝播媒体を介して伝播される電磁波（例えば、光ファイバケーブルを介した光パルス）、又はワイヤを介して伝送される電子信号などの、一時的な信号自身として解釈されてはならない。

【 0 0 8 5 】

ここで説明されるコンピュータ読み取り可能なプログラム指令は、コンピュータ可読記憶媒体から様々な計算/処理デバイスにダウンロードするか、又はネットワーク、例えば、インターネット、ローカルエリアネットワーク、ワイドエリアネットワーク及び/又は無線ネットワークを介して外部コンピュータ又は外部記憶装置にダウンロードすることができる。ネットワークは、銅の伝送ケーブル、光伝送ファイバー、無線伝送、ルータ、ファイアウォール、スイッチ、ゲートウェイコンピュータ及び/又はエッジサーバを含んでよい。各計算/処理デバイスにおけるネットワークアダプターカード又はネットワークインタフェースは、ネットワークからコンピュータ読み取り可能なプログラム指令を受信し、各計算/処理デバイスにおけるコンピュータ可読記憶媒体に記憶するために、該コンピュータ読み取り可能なプログラム指令を転送する。

【 0 0 8 6 】

本開示の操作を実行するためのコンピュータプログラム指令は、アセンブリ指令、命令セットアーキテクチャ（I S A）指令、機器指令、機器依存指令、マイクロコード、ファームウェア指令、状態設定データ、或いは 1 種又は複数種のプログラミング言語の任意の組み合わせで書かれたソースコード又はターゲットコードであってもよく、前記プログラミング言語は、S m a l l t a l k、C ++ などのオブジェクト指向プログラミング言語、及び「C」言語又は類似のプログラミング言語などの従来の手続き型プログラミング言語を含む。コンピュータ読み取り可能なプログラム指令は、ユーザのコンピュータで完全に又は部分的に実行されてもよく、1 つの独立したソフトウェアパッケージとして実行されてもよく、部分的にユーザのコンピュータで、部分的にリモートコンピュータで実行されてもよく、完全にリモートコンピュータ又はサーバで実行されてもよい。リモートコンピュータに関する場合、リモートコンピュータは、任意の種類ネットワーク（ローカルエリアネットワーク（L A N）又はワイドエリアネットワーク（W A N）を含む）を通じてユーザのコンピュータに接続されてもよく、（例えば、インターネットサービスプロバイダを利用してインターネットを通じて）外部コンピュータに接続されてもよい。いくつかの実施形態において、例えばプログラマブル論理回路、フィールドプログラマブルゲートアレイ（F P G A）又はプログラマブルロジックアレイ（P L A）を含む電子回路は、

本開示の態様を実行するために、コンピュータ読み取り可能なプログラム指令の状態情報を利用して電子回路をカスタマイズすることにより、コンピュータ読み取り可能なプログラム指令を実行してよい。

【0087】

ここで、本開示の実施形態に係る方法、装置（システム）及びコンピュータプログラム製品のフローチャート及び／又はブロック図を参照して本開示の各態様を説明する。フローチャート及び／又はブロック図の各ブロック及びフローチャート及び／又はブロック図における各ブロックの組み合わせの全ては、コンピュータ読み取り可能なプログラム指令によって実行され得ることを理解されたい。

【0088】

これらのコンピュータ読み取り可能なプログラム指令は、汎用コンピュータ、専用コンピュータ又は他のプログラム可能なデータ処理装置のプロセッシングユニットに提供することができ、それによりマシンを作成して、これらの指令がコンピュータ又は他のプログラム可能なデータ処理装置のプロセッシングユニットによって実行されるときに、フローチャート及び／又はブロック図における1つ又は複数のブロックで規定した機能／動作を実現する装置を作成するようにする。これらのコンピュータ読み取り可能なプログラム指令をコンピュータ可読記憶媒体に記憶してもよく、これらの指令は、コンピュータ、プログラム可能なデータ処理装置及び／又は他のデバイスが決定方式で動作するようにし、それにより、指令が記憶されたコンピュータ読み取り可能な媒体は、フローチャート及び／又はブロック図における1つ又は複数のブロックで規定した機能／動作を実現する各態様の指令を含む、製品を含む。

【0089】

コンピュータ読み取り可能なプログラム指令をコンピュータ、他のプログラム可能なデータ処理装置、又は他のデバイスにローディングして、一連の操作ステップをコンピュータ、他のプログラム可能なデータ処理装置又は他のデバイスで実行することができるようにし、コンピュータによって実現されるプロセスを作成し、それにより、コンピュータ、他のプログラム可能なデータ処理装置、又は他のデバイスで実行される指令がフローチャート及び／又はブロック図における1つ又は複数のブロックで規定した機能／動作を実現することができるようにする。

【0090】

図面におけるフローチャート及びブロック図は、本開示の複数の実施形態に係るシステム、方法及びコンピュータプログラム製品の実現可能な実装アーキテクチャ、機能及び動作を示す。これに関して、フローチャート又はブロック図における各ブロックは、1つのモジュール、プログラムセグメント又は指令の一部を表すことができ、前記モジュール、プログラムセグメント又は指令の一部は、規定された論理機能を実現するための1つ又は複数の実行可能な指令を含む。いくつかの代替的な実現において、ブロックで表示された機能は、図面に表示された順序と異なる順序で発生することができる。例えば、連続して示された2つのブロックは、実際には、基本的に並行して実行される場合や、逆の順序で実行される場合があり、これは、関連する機能によって決定される。ブロック図及び／又はフローチャートにおける各ブロック、及びブロック図及び／又はフローチャートにおけるブロックの組み合わせは、規定された機能又は動作を実行する専用のハードウェアに基づくシステムによって実現されてもよく、専用のハードウェアとコンピュータ指令の組み合わせによって実現されてもよいことにも留意されたい。

【0091】

以上、本開示の各実施形態を説明したが、以上の説明は、網羅的ではなく、例示的なものに過ぎず、開示された各実施形態に限定されない。当業者にとって、説明された各実施形態の範囲及び思想から逸脱することなく、多くの修正及び変更は明らかである。本明細書で使用される用語の選択は、各実施形態の原理、実際の応用又は市場における技術への改善を最もよく説明するか、当業者が本明細書で開示された各実施形態を理解することができるようにすることを意図する。

10

20

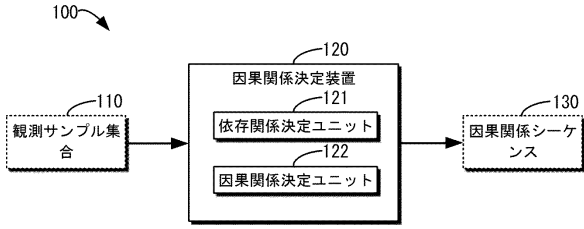
30

40

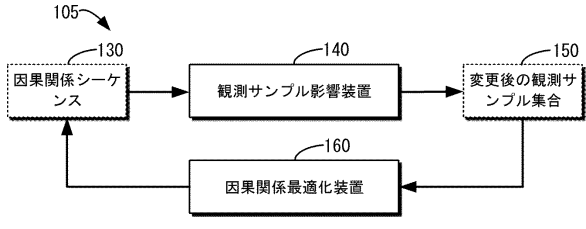
50

【図面】

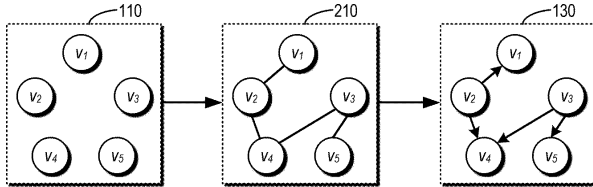
【図 1 A】



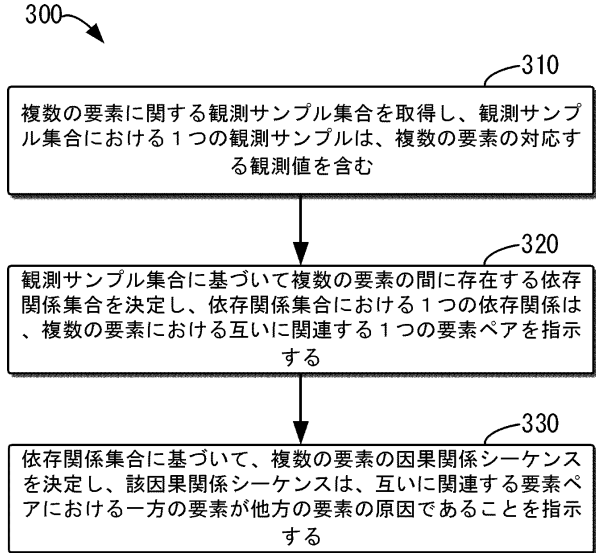
【図 1 B】



【図 2】



【図 3】



10

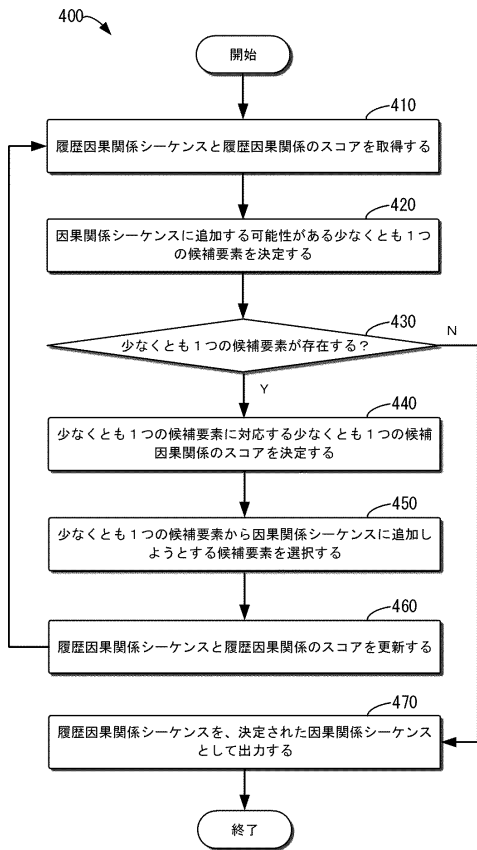
20

30

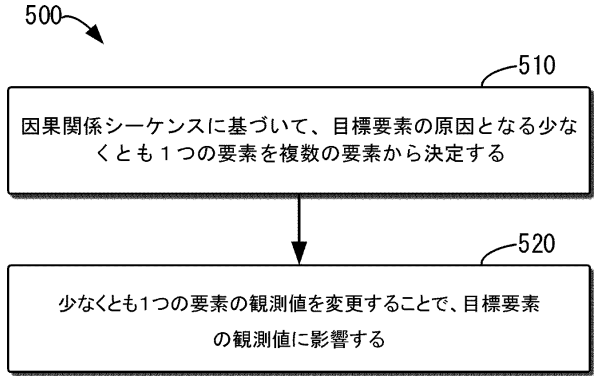
40

50

【図4】



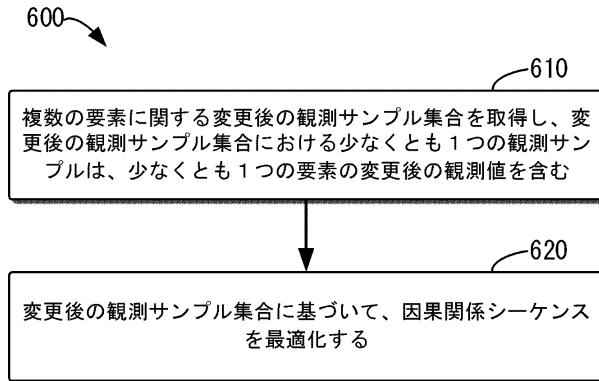
【図5】



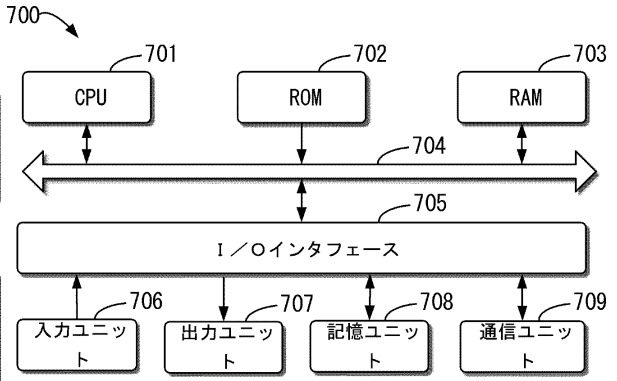
10

20

【図6】



【図7】



30

40

50

フロントページの続き

ァンドンルー ナンバー 19, リャンマーチャオ ディプロマティック オフィス ビルディング,
ビルディング ディー 2, 6 エフ

(72)発明者 ツウイ ルーイエ

中華人民共和国 100600 ベイジン, チャオヤン ディストリクト, ドンファンズンルー ナ
ンバー 19, リャンマーチャオ ディプロマティック オフィス ビルディング, ビルディング デ
ィー 2, 6 エフ

(72)発明者 フェン ルー

中華人民共和国 100600 ベイジン, チャオヤン ディストリクト, ドンファンズンルー ナ
ンバー 19, リャンマーチャオ ディプロマティック オフィス ビルディング, ビルディング デ
ィー 2, 6 エフ

合議体

審判長 吉田 美彦

審判官 林 毅

審判官 大塚 俊範

(56)参考文献 安富祖仁ほか, 遺伝子発現データからの遺伝子間因果関係ネットワーク推定, 情報処理学
会研究報告, 社団法人情報処理学会, 2006年9月15日, Vol. 2006, No.
9 (2006-BIO-6), pp. 9-15

(58)調査した分野 (Int.Cl., DB名)

G06F16/28

G06N5/025

G06N99/00