

(19)日本国特許庁(JP)

## (12)特許公報(B2)

(11)特許番号  
特許第7444248号  
(P7444248)

(45)発行日 令和6年3月6日(2024.3.6)

(24)登録日 令和6年2月27日(2024.2.27)

(51)国際特許分類

G 0 6 F	16/90 (2019.01)	F I	G 0 6 F	16/90
G 0 6 F	16/26 (2019.01)		G 0 6 F	16/26

請求項の数 7 (全26頁)

(21)出願番号	特願2022-524810(P2022-524810)
(86)(22)出願日	令和2年5月21日(2020.5.21)
(86)国際出願番号	PCT/JP2020/020144
(87)国際公開番号	WO2021/234916
(87)国際公開日	令和3年11月25日(2021.11.25)
審査請求日	令和4年8月9日(2022.8.9)

(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
(74)代理人	100103090 弁理士 岩壁 冬樹
(74)代理人	100124501 弁理士 塩川 誠人
(72)発明者	小山田 昌史 東京都港区芝五丁目 7 番 1 号 日本電気 株式会社内
審査官	三橋 竜太郎

最終頁に続く

(54)【発明の名称】 分析装置、分析方法および分析プログラム

## (57)【特許請求の範囲】

## 【請求項 1】

1つ以上のデータの意味を、1つ以上の別のデータの意味に変換する変換ルールを複数個記憶する変換ルール記憶手段と、

与えられた1つ以上のデータの意味と、前記変換ルールとに基づいて、前記意味を持つデータを用いてどのような分析を行えるかを示す情報を導出する分析手段と、

各変換ルールに基づいて、データの意味を表すノードの集合と、変換ルールIDを表すノードの集合とを含む有向2部グラフを生成するグラフ生成手段とを備え、

前記分析手段は、

前記与えられた1つ以上のデータの意味に対応する各ノードを探索開始点と定めた後に、前記有向2部グラフにおいて、前記探索開始点から1つのエッジを介して到達する変換ルールIDに対応するノードを特定し、

特定された前記ノードに対応する変換ルールIDが表す変換ルールにおける変換前のデータの意味に対応する各ノードが全て探索開始点であり、特定された前記ノードが当該探索開始点の全てから到達されている場合に、前記変換ルールにおける変換後のデータの意味を表すノードまでの探索ルートを導出し、

前記変換後のデータの意味を表すノードを探索開始点として定めることを繰り返し、

新たな探索ルートが導出できなくなった時点までに導出された探索ルートを、前記情報として定める

10

20

ことを特徴とする分析装置。

【請求項 2】

1つ以上のデータの意味を、1つ以上の別のデータの意味に変換する変換ルールを複数個記憶する変換ルール記憶手段と、

与えられた1つ以上のデータの意味と、前記変換ルールに基づいて、前記意味を持つデータを用いてどのような分析を行えるかを示す情報を導出する分析手段とを備え、

個々の変換ルールにそれぞれコストが予め定められていて、

各変換ルールに基づいて、データの意味を表すノードの集合と、変換ルール ID を表すノードの集合とを含む有向2部グラフを生成するグラフ生成手段と、

前記有向2部グラフにおける個々のデータの意味に対して、コストの初期値を設定するコスト初期値設定手段とを備え、

前記分析手段は、

前記与えられた1つ以上のデータの意味に対応する各ノードを探索開始点と定めた後に、

前記有向2部グラフにおいて、前記探索開始点から1つのエッジが向かっている変換ルール ID に対応するノードのうち、前記探索開始点に対応するデータの意味のコストと、前記変換ルール ID が表す変換ルールのコストとの和が、所定のコスト上限値以下であるという条件を満たすノードのみを、前記探索開始点から1つのエッジを介して到達されるノードとして特定し、

特定された前記ノードに対応する変換ルール ID が表す変換ルールにおける変換前のデータの意味に対応する各ノードが全て探索開始点であり、特定された前記ノードが当該探索開始点の全てから到達されている場合に、前記変換ルールにおける変換後のデータの意味を表すノードまでの探索ルートを導出し、

前記変換後のデータの意味を表すノードを探索開始点として定めるとともに、所定の条件を満たす場合に、当該ノードに対応するデータの意味のコストを更新すること

を繰り返し、

新たな探索ルートが導出できなくなった時点までに導出された探索ルートを、前記情報をとして定める

ことを特徴とする分析装置。

【請求項 3】

前記コスト初期値設定手段は、

前記与えられた1つ以上のデータの意味のコストをそれぞれ0に設定し、残りのデータの意味のコストを無限大に設定し、

前記分析手段は、

導出された探索ルート上の最後の変換ルール ID が表す変換ルールのコストと、前記変換ルール ID に対応するノードに到達する全ての探索開始点に対応する各データの意味のコストの総和との和が、前記変換ルールにおける変換後のデータの意味のコスト以下である場合に、当該データの意味のコストを、前記和で更新する

請求項2に記載の分析装置。

【請求項 4】

1つ以上のデータの意味を、1つ以上の別のデータの意味に変換する変換ルールを複数個記憶する変換ルール記憶手段と、

与えられた1つ以上のデータの意味と、前記変換ルールに基づいて、前記意味を持つデータを用いてどのような分析を行えるかを示す情報を導出する分析手段とを備え、

前記分析手段は、

変換前のデータの意味が、前記与えられた1つ以上のデータの意味に包含されているという条件を満たす変換ルールを抽出し、

抽出した各変換ルールにおける変換後のデータの意味と、前記与えられた1つ以上のデータの意味との和集合を求め、当該和集合を前記与えられた1つ以上のデータの意味とみなすことを繰り返し、

抽出した変換ルールと、前回抽出した変換ルールとが同一になったならば、前記抽出し

10

20

30

40

50

た変換ルールの集合を、前記情報として定める  
ことを特徴とする分析装置。

【請求項 5】

データが与えられた場合に、当該データの意味を推定する意味推定手段を備える  
請求項 1 から請求項 4 のうちのいずれか 1 項に記載の分析装置。

【請求項 6】

1 つ以上のデータの意味を、1 つ以上の別のデータの意味に変換する変換ルールを複数個記憶する変換ルール記憶手段を備えるコンピュータが、

与えられた 1 つ以上のデータの意味と、前記変換ルールとに基づいて、前記意味を持つデータを用いてどのような分析を行えるかを示す情報を導出し、  
前記情報を導出するときに、

変換前のデータの意味が、前記与えられた 1 つ以上のデータの意味に包含されているという条件を満たす変換ルールを抽出し、

抽出した各変換ルールにおける変換後のデータの意味と、前記与えられた 1 つ以上のデータの意味との和集合を求め、当該和集合を前記与えられた 1 つ以上のデータの意味とみなすことを繰り返し、

抽出した変換ルールと、前回抽出した変換ルールとが同一になったならば、前記抽出した変換ルールの集合を、前記情報として定める

ことを特徴とする分析方法。

【請求項 7】

1 つ以上のデータの意味を、1 つ以上の別のデータの意味に変換する変換ルールを複数個記憶する変換ルール記憶手段を備えるコンピュータに、

与えられた 1 つ以上のデータの意味と、前記変換ルールとに基づいて、前記意味を持つデータを用いてどのような分析を行えるかを示す情報を導出する分析処理を実行させ、  
前記分析処理で、

変換前のデータの意味が、前記与えられた 1 つ以上のデータの意味に包含されているという条件を満たす変換ルールを抽出させ、

抽出した各変換ルールにおける変換後のデータの意味と、前記与えられた 1 つ以上のデータの意味との和集合を求め、当該和集合を前記与えられた 1 つ以上のデータの意味とみなすことを繰り返させ、

抽出した変換ルールと、前回抽出した変換ルールとが同一になったならば、前記抽出した変換ルールの集合を、前記情報として定めさせる

ための分析プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、データの意味が与えられた場合に、その意味を持つデータを用いてどのような分析を行えるのかを分析する分析装置、分析方法および分析プログラムに関する。

【背景技術】

【0002】

特許文献 1 には、シーケンス分析と決定木分析を組み合わせて、精度の高い潜在ターゲットを導出できる装置が記載されている。特許文献 1 に記載された装置は、時系列データを入力し、また、特定属性データ出現集合特有のルールを入力する。そして、その装置は、特定属性データ出現傾向と正の相関があるルールと同一の時系列的推移を持つデータの加工、および、特定属性データ非出現傾向と負の相関があるルールと同一の時系列的推移を持つデータの加工を行う。

【0003】

また、特許文献 2 には、データを格納したテーブルが入力された場合に、そのテーブルのカラム毎に、カラムに格納されたデータの意味を推定する技術が記載されている。ここで、「データの意味」とは、そのデータが表す概念である。各カラムにはカラム名が付与

10

20

30

40

50

されている。しかし、一般に、カラム名は、人間によって決定されるので、カラム名には表記ゆれが生じる。例えば、人物の性別を格納したカラムのカラム名として、「種別」、「男女」等の種々のカラム名が付与され得る。前述のように、「データの意味」とは、そのデータが表す概念であり、カラム名とは区別される。上記の例では、「性別」がデータの意味に該当する。

【先行技術文献】

【特許文献】

【0004】

【文献】特開2005-70913号公報

【文献】国際公開第2018/025706号

10

【発明の概要】

【発明が解決しようとする課題】

【0005】

一般に、多くのデータを持つ者がいる。このような者の例として、例えば、種々の店舗等が挙げられる。ただし、このような者は、店舗に限定されない。

【0006】

このように多くのデータを持つ者は、データを持つ一方で、そのデータを用いてどのような分析を行えるかを分かっていないことが多い。

【0007】

そこで、本発明は、所持されているデータを用いてどのような分析を行えるかを示す情報を導出することができる分析装置、分析方法および分析プログラムを提供することを目的とする。

20

【課題を解決するための手段】

【0008】

本発明による分析装置は、1つ以上のデータの意味を、1つ以上の別のデータの意味に変換する変換ルールを複数個記憶する変換ルール記憶手段と、与えられた1つ以上のデータの意味と、前記変換ルールとに基づいて、前記意味を持つデータを用いてどのような分析を行えるかを示す情報を導出する分析手段と、各変換ルールに基づいて、データの意味を表すノードの集合と、変換ルールIDを表すノードの集合とを含む有向2部グラフを生成するグラフ生成手段とを備え、前記分析手段が、前記与えられた1つ以上のデータの意味に対応する各ノードを探索開始点と定めた後に、前記有向2部グラフにおいて、前記探索開始点から1つのエッジを介して到達する変換ルールIDに対応するノードを特定し、特定された前記ノードに対応する変換ルールIDが表す変換ルールにおける変換前のデータの意味に対応する各ノードが全て探索開始点であり、特定された前記ノードが当該探索開始点の全てから到達されている場合に、前記変換ルールにおける変換後のデータの意味を表すノードまでの探索ルートを導出し、前記変換後のデータの意味を表すノードを探索開始点として定めることを繰り返し、新たな探索ルートが導出できなくなった時点までに導出された探索ルートを、前記情報として定めることを特徴とする。

30

また、本発明による分析装置は、1つ以上のデータの意味を、1つ以上の別のデータの意味に変換する変換ルールを複数個記憶する変換ルール記憶手段と、与えられた1つ以上のデータの意味と、前記変換ルールとに基づいて、前記意味を持つデータを用いてどのような分析を行えるかを示す情報を導出する分析手段とを備え、個々の変換ルールにそれぞれコストが予め定められていて、各変換ルールに基づいて、データの意味を表すノードの集合と、変換ルールIDを表すノードの集合とを含む有向2部グラフを生成するグラフ生成手段と、前記有向2部グラフにおける個々のデータの意味に対して、コストの初期値を設定するコスト初期値設定手段とを備え、前記分析手段が、前記与えられた1つ以上のデータの意味に対応する各ノードを探索開始点と定めた後に、前記有向2部グラフにおいて、前記探索開始点から1つのエッジが向かっている変換ルールIDに対応するノードのうち、前記探索開始点に対応するデータの意味のコストと、前記変換ルールIDが表す変換ルールのコストとの和が、所定のコスト上限値以下であるという条件を満たすノードのみを

40

50

前記探索開始点から 1 つのエッジを介して到達されるノードとして特定し、特定された前記ノードに対応する変換ルール ID が表す変換ルールにおける変換前のデータの意味に対応する各ノードが全て探索開始点であり、特定された前記ノードが当該探索開始点の全てから到達されている場合に、前記変換ルールにおける変換後のデータの意味を表すノードまでの探索ルートを導出し、前記変換後のデータの意味を表すノードを探索開始点として定めるとともに、所定の条件を満たす場合に、当該ノードに対応するデータの意味のコストを更新することを繰り返し、新たな探索ルートが導出できなくなった時点までに導出された探索ルートを、前記情報として定めることを特徴とする。

また、本発明による分析装置は、1 つ以上のデータの意味を、1 つ以上の別のデータの意味に変換する変換ルールを複数個記憶する変換ルール記憶手段と、与えられた 1 つ以上のデータの意味と、前記変換ルールとに基づいて、前記意味を持つデータを用いてどのような分析を行えるかを示す情報を導出する分析手段とを備え、前記分析手段が、変換前のデータの意味が、前記与えられた 1 つ以上のデータの意味に包含されているという条件を満たす変換ルールを抽出し、抽出した各変換ルールにおける変換後のデータの意味と、前記与えられた 1 つ以上のデータの意味との和集合を求め、当該和集合を前記与えられた 1 つ以上のデータの意味とみなすことを繰り返し、抽出した変換ルールと、前回抽出した変換ルールとが同一になったならば、前記抽出した変換ルールの集合を、前記情報として定めることを特徴とする。

#### 【0009】

本発明による分析方法は、1 つ以上のデータの意味を、1 つ以上の別のデータの意味に変換する変換ルールを複数個記憶する変換ルール記憶手段を備えるコンピュータが、与えられた 1 つ以上のデータの意味と、前記変換ルールとに基づいて、前記意味を持つデータを用いてどのような分析を行えるかを示す情報を導出し、前記情報を導出するときに、変換前のデータの意味が、前記与えられた 1 つ以上のデータの意味に包含されているという条件を満たす変換ルールを抽出し、抽出した各変換ルールにおける変換後のデータの意味と、前記与えられた 1 つ以上のデータの意味との和集合を求め、当該和集合を前記与えられた 1 つ以上のデータの意味とみなすことを繰り返し、抽出した変換ルールと、前回抽出した変換ルールとが同一になったならば、前記抽出した変換ルールの集合を、前記情報として定めることを特徴とする。

#### 【0010】

本発明による分析プログラムは、1 つ以上のデータの意味を、1 つ以上の別のデータの意味に変換する変換ルールを複数個記憶する変換ルール記憶手段を備えるコンピュータに、与えられた 1 つ以上のデータの意味と、前記変換ルールとに基づいて、前記意味を持つデータを用いてどのような分析を行えるかを示す情報を導出する分析処理を実行させ、前記分析処理で、変換前のデータの意味が、前記与えられた 1 つ以上のデータの意味に包含されているという条件を満たす変換ルールを抽出させ、抽出した各変換ルールにおける変換後のデータの意味と、前記与えられた 1 つ以上のデータの意味との和集合を求め、当該和集合を前記与えられた 1 つ以上のデータの意味とみなすことを繰り返させ、抽出した変換ルールと、前回抽出した変換ルールとが同一になったならば、前記抽出した変換ルールの集合を、前記情報として定めさせる。

#### 【発明の効果】

#### 【0011】

本発明によれば、所持されているデータを用いてどのような分析を行えるかを示す情報を導出することができる。

#### 【図面の簡単な説明】

#### 【0012】

【図 1】本発明の第 1 の実施形態の分析装置の例を示すブロック図である。

【図 2】変換ルールの例を示す模式図である。

【図 3】第 1 の実施形態の処理経過の例を示すフローチャートである。

【図 4】本発明の第 2 の実施形態の分析装置の例を示すブロック図である。

10

20

30

40

50

【図 5】変換ルールの例を示す模式図である。

【図 6】図 5 に示す各変換ルールに基づいて生成された有向 2 部グラフの例を示す説明図である。

【図 7】第 2 の実施形態の処理経過の例を示すフローチャートである。

【図 8】第 2 の実施形態の処理経過の例を示すフローチャートである。

【図 9】第 2 の実施形態の処理経過の例を示すフローチャートである。

【図 10】ステップ S 1 8 で導出される探索ルートの例を示す模式図である。

【図 11】最終的に得られる探索ルートの例を示す模式図である。

【図 12】本発明の第 3 の実施形態の分析装置の例を示すブロック図である。

【図 13】第 3 の実施形態の処理経過の例を示すフローチャートである。

10

【図 14】第 3 の実施形態の処理経過の例を示すフローチャートである。

【図 15】第 3 の実施形態の処理経過の例を示すフローチャートである。

【図 16】ステップ S 3 1 で設定されたコストを有向 2 部グラフとともに示した模式図である。

【図 17】ステップ S 1 8 で導出される探索ルートの例を示す模式図である。

【図 18】本発明の各実施形態の分析装置に係るコンピュータの構成例を示す概略ブロック図である。

【図 19】本発明の分析装置の概要の例を示すブロック図である。

【発明を実施するための形態】

【0013】

20

以下、本発明の実施形態を図面を参照して説明する。

【0014】

図 1 は、本発明の第 1 の実施形態の分析装置の例を示すブロック図である。本実施形態の分析装置 1 は、取得部 2 と、意味推定部 3 と、意味記憶部 4 と、変換ルール記憶部 5 と、分析部 6 とを備える。

【0015】

取得部 2 は、多くのデータを持つ者によって所持されているデータを取得する。本実施形態では、説明を簡単にするために、取得部 2 が、データを格納したテーブルを取得する場合を例にして説明する。テーブルは、1 つ以上のカラムを含み、各カラムにデータが格納されている。

30

【0016】

なお、データを持つ者によって所持されているデータには、その者が他者から購入したデータが含まれていてもよい。

【0017】

取得部 2 は、例えば、光学ディスク等のデータ記録媒体に記録されたテーブルを読み込むデータ読み込み装置によって実現されてもよいが、取得部 2 は、そのようなデータ読み込み装置に限定されない。例えば、取得部 2 は、通信ネットワークを介して配信されるテーブルを受信する通信インターフェースによって実現されてもよい。

【0018】

意味推定部 3 は、取得部 2 が取得した種々のデータ群毎に、データの意味を推定する。本実施形態では、意味推定部 3 が、取得部 2 が取得したテーブルのカラム毎に、カラムに格納されたデータの意味を推定する場合を例にして説明する。意味推定部 3 によって推定されるデータの意味の数は 1 つ以上である。取得部 2 がデータの意味を推定する方法は、公知の方法でよい。例えば、意味推定部 3 は、特許文献 2 に記載された方法で、カラム毎に、カラムに格納されたデータの意味を推定してもよい。意味推定部 3 によって推定された種々のデータの意味は、データを持つ者から与えられた「データの意味」であると言うことができる。

40

【0019】

意味推定部 3 は、推定によって得た種々の「データの意味」を、意味記憶部 4 に記憶させる。例えば、意味推定部 3 が、推定によって、データの意味として、「身長」、「体重

50

」、「年収」、「年齢」等を得た場合、それらのデータの意味を、意味記憶部4に記憶させる。

【0020】

意味記憶部4は、データの意味を記憶する記憶装置である。

【0021】

変換ルール記憶部5は、複数の変換ルールを記憶する記憶装置である。「変換ルール」とは、1つ以上のデータの意味を、1つ以上の別のデータの意味に変換するルールである。ここで、「1つ以上のデータの意味を、1つ以上の別のデータの意味に変換する」とは、「ある意味を持つデータに基づいて、別の意味を持つデータを計算したり、推定したりすることができる」ということを表している。また、個々の変換ルールにはそれぞれ、変換ルールID(変換ルールの識別情報)が予め定められている。

10

【0022】

図2は、変換ルールの例を示す模式図である。図2では、3つの変換ルールを示している。例えば、図2に示す1番目の変換ルールは、「身長」というデータの意味、および、「体重」というデータの意味を、「BMI(Body Mass Index)」というデータの意味に変換することを表している。このことは、「身長」という意味を持つデータ、および、「体重」という意味を持つデータに基づいて、「BMI」という意味を持つデータを計算できることを表している。また、この変換ルールには、「BMI計算」という変換ルールIDが定められている。

20

【0023】

図2に示す1番目の変換ルールでは、「身長」および「体重」が変換前の「データの意味」に該当し、「BMI」が変換後の「データの意味」に該当する。図2に示すように矢印を用いて模式的に変換ルールを表した場合、図2に示す例では、矢印の左辺に変換前の「データの意味」が示され、矢印の右辺に変換後の「データの意味」が示されている。そのため、便宜的に、変換前のデータの意味を、左辺のデータの意味と称し、変換後のデータの意味を右辺のデータの意味と称する場合がある。

30

【0024】

図2では、各変換ルールにおいて、左辺のデータの意味(変換前のデータの意味)が複数存在する場合を示したが、変換ルールにおいて、左辺のデータの意味の数が1つであってもよい。また、図2では、各変換ルールにおいて、右辺のデータの意味(変換後のデータの意味)の数が1つである場合を示したが、変換ルールにおいて、右辺のデータの意味が複数存在していてもよい。右辺のデータの意味が複数存在するということは、左辺のデータの意味に基づいて、複数種類のデータが得られることを意味している。

30

【0025】

また、以下の説明において、変換ルールIDを“r”と表した場合に、左辺に該当する全てのデータの意味を、“r.input\_semantics”と表す場合がある。同様に、右辺に該当する全てのデータの意味を、“r.output\_semantics”と表す場合がある。例えば、図2に示す1番目の変換ルールでは、“BMI計算.input\_semantics”は、{「身長」、「体重」}であり、“BMI計算.output\_semantics”は、{「BMI」}である。

40

【0026】

各変換ルールは、例えば、分析処理のサービスを提供する者や、データを販売する者等によって提供され、予め変換ルール記憶部5に記憶される。

【0027】

分析部6は、意味推定部3によって得られたデータの意味と、各変換ルールとに基づいて、その意味を持つデータを用いてどのような分析を行えるかを示す情報を導出する。本実施形態では、分析部6は、この情報として、意味推定部3によって得られたデータの意味を起点にして再帰的に抽出した変換ルールの集合を導出する。

40

【0028】

前述のように、意味推定部3によって推定された種々のデータの意味は、データを持つ者から与えられた「データの意味」であると言えることができる。そのため、以下、意味推

50

定部 3 によって推定されたデータの意味を、与えられたデータの意味と記す場合がある。

【 0 0 2 9 】

分析部 6 は、左辺のデータの意味（変換前のデータの意味）が、与えられたデータの意味に包含されているという条件を満たす変換ルールを抽出する。そして、分析部 6 は、抽出した各変換ルールにおける右辺のデータの意味（変換後のデータの意味）と、与えられたデータの意味との和集合を求め、その和集合を、与えられたデータの意味とみなす。分析部 6 は、この動作を繰り返す。

【 0 0 3 0 】

そして、分析部 6 は、新たに抽出した変換ルールと、前回抽出した変換ルールとが同一になつたならば、その新たに抽出した変換ルールの集合を、どのような分析を行えるかを示す情報として定める。

10

【 0 0 3 1 】

意味推定部 3 および分析部 6 は、例えば、分析プログラムに従って動作するコンピュータの C P U ( Central Processing Unit ) によって実現される。例えば、C P U が、コンピュータのプログラム記憶装置等のプログラム記録媒体から分析プログラムを読み込み、その分析プログラムに従って、意味推定部 3 および分析部 6 として動作すればよい。

【 0 0 3 2 】

意味記憶部 4 および変換ルール記憶部 5 は、例えば、コンピュータが備える記憶装置によって実現される。

【 0 0 3 3 】

次に、第 1 の実施形態の処理経過について説明する。図 3 は、第 1 の実施形態の処理経過の例を示すフローチャートである。なお、取得部 2 がデータ（テーブル）を取得しているものとする。また、説明を簡単にするために、変換ルール記憶部 5 が、図 2 に示す 3 つの変換ルールを記憶しているものとする。また、既に説明した事項については、詳細な説明を省略する。

20

【 0 0 3 4 】

まず、意味推定部 3 が、取得部 2 が取得したデータの意味を推定する（ステップ S 1）。意味推定部 3 は、テーブルのカラム毎に、カラムに格納されたデータの意味を推定する。ステップ S 1 では、1 つ以上のデータの意味が得られる。本例では、データの意味として、「身長」、「体重」、「年収」、「年齢」が得られたものとする。以下、ステップ S 1 で得られたデータの意味を、与えられたデータの意味と記す。意味推定部 3 は、与えられたデータの意味（「身長」、「体重」、「年収」、「年齢」）を、意味記憶部 4 に記憶させる。

30

【 0 0 3 5 】

ステップ S 1 の後に、分析部 6 は、左辺のデータの意味（変換前のデータの意味）が、与えられたデータの意味に包含されているという条件を満たす変換ルールを全て抽出する（ステップ S 2）。本例では、与えられたデータの意味は、{「身長」、「体重」、「年収」、「年齢」} である。また、図 2 に示す変換ルール「B M I 計算」の左辺のデータの意味は、{「身長」、「体重」} である。従って、変換ルール「B M I 計算」の左辺のデータの意味は、与えられたデータの意味に包含されている。従って、変換ルール「B M I 計算」は、上記の条件を満たしているので、分析部 6 は、変換ルール「B M I 計算」を抽出する。

40

【 0 0 3 6 】

また、図 2 に示す変換ルール「年齢推定」に関しては、「性別」が、与えられたデータの意味に包含されていない。同様に、変換ルール「癌リスク予測」に関しては、「B M I」が、与えられたデータの意味に包含されていない。よって、変換ルール「年齢推定」および変換ルール「癌リスク予測」は、抽出されない。

【 0 0 3 7 】

ステップ S 2 の次に、分析部 6 は、直近のステップ S 2 で抽出した変換ルールと、前回のステップ S 2 で抽出した変換ルールとが同一であるか否かを判定する（ステップ S 3）

50

。両者が同一でない場合(ステップS3のNo)、ステップS4に移行する。両者が同一である場合(ステップS3のYes)、ステップS5に移行する。最初にステップS3に移行した時点では、ステップS2は1回しか実行されていないため、ステップS4に移行する。従って、ここでは、ステップS4に移行する。

#### 【0038】

ステップS4において、分析部6は、抽出した各変換ルールの右辺のデータの意味(変換後のデータの意味)と、与えられたデータの意味との和集合を求め、その和集合を、与えられたデータの意味とみなす。そして、分析部6は、与えられたデータの意味を、意味記憶部4に記憶させる。

#### 【0039】

本例では、抽出した変換ルール「BMI計算」の右辺のデータの意味は、{「BMI」}である。従って、分析部6は、{「BMI」}と、与えられたデータの意味{「身長」、「体重」、「年収」、「年齢」}との和集合として、{「身長」、「体重」、「年収」、「年齢」、「BMI」}を求める。そして、分析部6は、その和集合{「身長」、「体重」、「年収」、「年齢」、「BMI」}を、与えられたデータの意味とみなす。さらに、分析部6は、{「身長」、「体重」、「年収」、「年齢」、「BMI」}を意味記憶部4に記憶させる。

#### 【0040】

ステップS4の後、分析部6は、ステップS2以降の処理を繰り返す。2回目のステップS2において、与えられたデータの意味は、{「身長」、「体重」、「年収」、「年齢」、「BMI」}となっている。従って、このとき、分析部6は、変換ルール「BMI計算」、および、変換ルール「癌リスク予測」を抽出する(図2参照)。

10

#### 【0041】

次に、分析部6は、再度、ステップS3を実行する。このとき、直近のステップS2で抽出した変換ルールは、変換ルール「BMI計算」、および、変換ルール「癌リスク予測」である。また、前回のステップS2で抽出した変換ルールは、変換ルール「BMI計算」のみである。従って、両者は同一ではないので(ステップS3のNo)、ステップS4に移行する。

#### 【0042】

直近のステップS2で、抽出した2つの変換ルールの右辺の意味は、それぞれ、「BMI」、「癌リスク」である。従って、ステップS4において、分析部6は、{「BMI」、「癌リスク」}と、与えられたデータの意味{「身長」、「体重」、「年収」、「年齢」、「BMI」}の和集合として、{「身長」、「体重」、「年収」、「年齢」、「BMI」、「癌リスク」}を求める。そして、分析部6は、その和集合{「身長」、「体重」、「年収」、「年齢」、「BMI」、「癌リスク」}を、与えられたデータの意味とみなす(ステップS4)。さらに、分析部6は、{「身長」、「体重」、「年収」、「年齢」、「BMI」、「癌リスク」}を意味記憶部4に記憶させる。

30

#### 【0043】

ステップS4の後、分析部6は、3回目のステップS2を実行する。このとき、与えられたデータの意味は、{「身長」、「体重」、「年収」、「年齢」、「BMI」、「癌リスク」}となっている。従って、このとき、分析部6は、変換ルール「BMI計算」、および、変換ルール「癌リスク予測」を抽出する(図2参照)。

40

#### 【0044】

次に、分析部6は、再度、ステップS3を実行する。このとき、直近のステップS2で抽出した変換ルールは、変換ルール「BMI計算」、および、変換ルール「癌リスク予測」である。また、前回のステップS2で抽出した変換ルールも、変換ルール「BMI計算」、および、変換ルール「癌リスク予測」である。従って、両者は同一であるので(ステップS3のYes)、ステップS5に移行する。

#### 【0045】

ステップS5において、分析部6は、直近のステップS2で抽出した変換ルールの集合

50

を、どのような分析を行えるかを示す情報として定め、その情報を出力する（ステップS5）。本例では、分析部6は、変換ルール「BMI計算」、および、変換ルール「癌リスク予測」からなる変換ルールの集合を、どのような分析を行えるかを示す情報として定める。

#### 【0046】

また、ステップS5では、分析部6は、例えば、分析装置1に設けられたディスプレイ装置（図示略）に、情報を表示させてもよい。ただし、表示は、出力態様の一例であり、分析部6は、他の態様で情報を出力してもよい。この点は、後述の他の実施形態でも同様である。

#### 【0047】

第1の実施形態では、分析部6が、ステップS2で、左辺のデータの意味が、与えられたデータの意味に包含されているという条件を満たす変換ルールを全て抽出する。すなわち、ステップS2で抽出される変換ルールは、与えられた意味を持つデータを用いて行える分析を表していることができる。さらに、第1の実施形態では、ステップS4で和集合を求め、その和集合を、与えられたデータの意味とみなすことによって、与えられたデータの意味の数を増加させる。この結果、ステップS2～S4の繰り返し処理において、繰り返し回数が増えるほど、ステップS2で抽出される変換ルールも増加する。ステップS2で抽出される変換ルールが変化しなくなるまで、分析部6は、ステップS2～S4の繰り返し処理を実行するので、与えられた意味を持つデータを用いて行える分析を表す変換ルールをできるだけ多く抽出することができる。従って、本実施形態によれば、ユーザに所持されているデータを用いてどのような分析を行えるかを提示することができる。

10

#### 【0048】

実施形態2.

図4は、本発明の第2の実施形態の分析装置の例を示すブロック図である。本実施形態の分析装置1は、取得部2と、意味推定部3と、意味記憶部4と、変換ルール記憶部5と、グラフ生成部7と、グラフ記憶部8と、分析部26とを備える。

#### 【0049】

第2の実施形態における取得部2、意味推定部3、意味記憶部4および変換ルール記憶部5は、第1の実施形態における取得部2、意味推定部3、意味記憶部4および変換ルール記憶部5と同様であり、説明を省略する。

20

#### 【0050】

グラフ生成部7は、予め変換ルール記憶部5に記憶されている各変換ルールに基づいて、データの意味を表すノードの集合と、変換ルールIDを表すノードの集合とを含む有向2部グラフを生成する。第2の実施形態では、変換ルール記憶部5が、図5に例示する6個の変換ルールを記憶している場合を例にして説明する。また、図6は、図5に示す各変換ルールに基づいて生成された有向2部グラフの例を示す説明図である。

30

#### 【0051】

グラフ生成部7が有向2部グラフを生成する動作の例を以下に示す。ただし、グラフ生成部7は、有向2部グラフを他の方法で生成してもよい。

#### 【0052】

グラフ生成部7は、変換ルール記憶部5に記憶されている各変換ルールのうち、未選択の変換ルールを1つ選択する。グラフ生成部7は、選択した変換ルールの左辺のデータの意味を表すノード、および、右辺のデータの意味を表すノードを生成し、それらのノードを第1のノードの集合に含める。また、グラフ生成部7は、選択した変換ルールの変換ルールIDを表すノードを生成し、第2のノードの集合に含める。第1のノードの集合は、データの意味に対応するノードの集合であり、第2のノードの集合は、変換ルールIDに対応するノードの集合である。そして、グラフ生成部7は、選択した変換ルールの左辺のデータの意味に対応する各ノードそれぞれから、選択した変換ルールの変換ルールIDに対応するノードに向かうエッジを生成する。さらに、グラフ生成部7は、選択した変換ルールの変換ルールIDに対応するノードから、選択した変換ルールの右辺のデータの意味

40

50

に対応する各ノードそれぞれに向かうエッジを生成する。

【0053】

例えば、図5に示す変換ルール「B M I 計算」が選択されたとする。この場合、グラフ生成部7は、「身長」に対応するノード、「体重」に対応するノード、および「B M I」に対応するノードを生成し、それらのノードを第1の集合に含める。また、グラフ生成部7は、「B M I 計算」に対応するノードを生成し、第2の集合に含める。そして、グラフ生成部7は、「身長」に対応するノード、および、「体重」に対応するノードそれから、「B M I 計算」に対応するノードに向かうエッジを生成する。さらに、グラフ生成部7は、「B M I 計算」に対応するノードから、「B M I」に対応するノードに向かうエッジを生成する。

10

【0054】

グラフ生成部7は、未選択の変換ルールが無くなるまで、変換ルールを1つずつ選択し、上記の処理を実行する。ただし、新たに生成しようとするデータの意味に対応するノードが、既に生成されている場合には、グラフ生成部7は、そのノードを重複して生成しなくてよい。例えば、変換ルール「B M I 計算」を選択した後に、変換ルール「癌リスク予測」を選択して上記の処理を実行する場合を考える。変換ルール「癌リスク予測」を選択する場合、グラフ生成部7は、データの意味に対応するノードとして、「B M I」に対応するノード、「年齢」に対応するノード、および、「癌リスク」に対応するノードを生成することになるが、「B M I」に対応するノードは、既に生成済みである。従って、この場合、グラフ生成部7は、「B M I」に対応するノードを重複して生成しなくてよい。

20

【0055】

このようにして生成された有向2部グラフの例を図6に示す。図6では、データの意味に対応するノードから変換ルールIDに対応するノードに向かうエッジを実線で示している。また、変換ルールIDに対応するノードからデータの意味に対応するノードに向かうエッジを破線で示している。換言すれば、第1の集合内のノードから第2の集合内のノードに向かうエッジを実線で示し、第2の集合内のノードから第1の集合内のノードに向かうエッジを破線で示している。

20

【0056】

グラフ生成部7は、生成した有向2部グラフをグラフ記憶部8に記憶させる。グラフ記憶部8は、生成された有向2部グラフを記憶する記憶装置である。

30

【0057】

分析部26は、与えられたデータの意味に対応する各ノードをそれぞれ探索開始点と定める。

【0058】

その後、分析部26は、以下の処理を繰り返す。

【0059】

分析部26は、有向2部グラフにおいて、探索開始点から1つのエッジを介して到達する変換ルールIDに対応するノードを特定する。

【0060】

そして、分析部26は、特定されたノードに対応する変換ルールIDが表す変換ルールの左辺の「データの意味」に対応する各ノードが全て探索開始点であり、特定されたノードがそれらの探索開始点の全てから到達されている場合に、それらの探索開始点から、その変換ルールの右辺のデータの意味を表すノードまでの探索ルートを導出する。

40

【0061】

そして、分析部26は、上述の右辺のデータの意味を表すノードを探索開始点として定める。

【0062】

分析部26は、新たな探索ルートを導出できなくなった時点までに導出された探索ルートを、どのような分析を行えるかを示す情報として定める。

【0063】

50

グラフ生成部 7 および分析部 2 6 は、例えば、例えば、分析プログラムに従って動作するコンピュータの C P U によって実現される。また、グラフ記憶部 8 は、例えば、コンピュータが備える記憶装置によって実現される。

#### 【 0 0 6 4 】

次に、第 2 の実施形態の処理経過について説明する。図 7、図 8 および図 9 は、第 2 の実施形態の処理経過の例を示すフローチャートである。ただし、既に説明した事項については、詳細な説明を省略する。

#### 【 0 0 6 5 】

なお、変換ルール記憶部 5 は、予め図 5 に例示する各変換ルールを記憶しているものとする。また、グラフ生成部 7 は、その各変換ルールに基づいて、図 6 に例示する有向 2 部グラフを既に生成しており、その有向 2 部グラフをグラフ記憶部 8 に記憶させているものとする。また、取得部 2 がデータ（テーブル）を取得しているものとする。

10

#### 【 0 0 6 6 】

まず、意味推定部 3 が、取得部 2 が取得したデータの意味を推定する（ステップ S 1 1）。意味推定部 3 は、テーブルのカラム毎に、カラムに格納されたデータの意味を推定する。ステップ S 1 1 では、1 つ以上のデータの意味が得られる。以下、ステップ S 1 1 で得られたデータの意味を、与えられたデータの意味と記す。本例では、与えられたデータの意味が、{「身長」、「体重」、「年収」、「年齢」}であるものとする。

#### 【 0 0 6 7 】

次に、分析部 2 6 は、予め生成されている有向 2 部グラフにおいて、与えられたデータの意味に対応する各ノードをそれぞれ探索開始点と定める（ステップ S 1 2）。このノードは、有向 2 部グラフの第 1 の集合に属している。また、分析部 2 6 は、探索開始点に対応するデータの意味を、意味記憶部 4 に記憶させる。探索開始点の数は 1 つとは限らない。

20

#### 【 0 0 6 8 】

本例では、「身長」、「体重」、「年収」、「年齢」に対応するそれぞれのノードを、探索開始点とする。

#### 【 0 0 6 9 】

次に、分析部 2 6 は、有向 2 部グラフにおいて、各探索開始点からエッジが向かっている第 2 の集合内のノードを特定する（ステップ S 1 3）。ステップ S 1 3 で特定されるノードの数は、1 つとは限らない。また、ステップ S 1 3 で特定されたノードは、探索開始点から 1 つのエッジを介して到達するノードであると言える。

30

#### 【 0 0 7 0 】

本例では、「身長」、「体重」、「年収」、「年齢」に対応するそれぞれのノードが探索開始点であるので、ステップ S 1 3 では、「B M I 計算」、「年齢推定」、「癌リスク予測」、「年収推定」、「保険金算出」に対応するノードが特定される。

#### 【 0 0 7 1 】

次に、分析部 2 6 は、ステップ S 1 3 で特定されたノードが全てステップ S 1 5 で選択済みであるか否かを判定する（ステップ S 1 4）。

#### 【 0 0 7 2 】

ステップ S 1 3 で特定されたノードのうち、ステップ S 1 5 で選択されていないノードが残っているならば（ステップ S 1 4 の N o ）、ステップ S 1 5 に移行する。

40

#### 【 0 0 7 3 】

ステップ S 1 5 において、分析部 2 6 は、ステップ S 1 3 で特定されたノードのうち、未選択のノードを 1 つ選択する。

#### 【 0 0 7 4 】

次に、分析部 2 6 は、ステップ S 1 5 で選択されたノードに対応する変換ルール I D を持つ変換ルールに、そのノードに到達する各探索開始点それぞれに対応するデータの意味を付加する（ステップ S 1 6）。変換ルール I D を “r” と表した場合に、その変換ルール I D に対応するノードに到達する探索開始点におけるデータの意味を、“r.visited\_semantics” と表すこととする。選択されたノードに対応する変換ルール I D を “r” とすると、ス 50

ステップ S 16 で、分析部 26 は、“r.visited\_semantics”に、具体的なデータの意味を付加する。

【0075】

ステップ S 16 の次に、分析部 26 は、ステップ S 15 で選択されたノードに対応する変換ルール ID を持つ変換ルールの左辺の「データの意味」と、ステップ S 16 でその変換ルールに付加された「データの意味」の集合とが合致するか否かを判定する（ステップ S 17）。選択されたノードに対応する変換ルール ID を “r” とすると、ステップ S 17 で、分析部 26 は、“r.input\_semantics” と、“r.visited\_semantics” に付加された「データの意味」の集合とが合致しているか否かを判定すればよい。

【0076】

ステップ S 17 で合致すると判定されたということは（ステップ S 17 の Yes）、選択されたノードに対応する変換ルール ID を持つ変換ルールにおける変換前の「データの意味」に対応する各ノードが全て探索開始点であり、選択されたノードが、その探索開始点の全てから到達されていることを意味する。

10

【0077】

ステップ S 17 で合致しないと判定された場合（ステップ S 17 の No）、分析部 26 は、ステップ S 14（図 7 参照）以降の処理を繰り返す。

【0078】

ステップ S 17 で合致すると判定された場合（ステップ S 17 の Yes）、分析部 26 は、ステップ S 15 で選択されたノードに対応する変換ルール ID を持つ変換ルールの左辺のデータの意味を表す各探索開始点から、選択されたノードを経由して、その変換ルールの右辺のデータの意味を表すノードに行き着くまでの探索ルートを、新たな探索ルートとして導出する。そして、分析部 26 は、その探索ルートをグラフ記憶部 8 に記憶させる（ステップ S 18）。ただし、既に導出済みの探索ルートと同じ探索ルートについては、分析部 26 は、重複してグラフ記憶部 8 に記憶させなくてよい。

20

【0079】

図 10 は、ステップ S 18 で導出される探索ルートの例を示す模式図である。

【0080】

ステップ S 18 の次に、分析部 26 は、ステップ S 18 で導出された探索ルートの終点に該当するノードを、探索開始点として定める（ステップ S 19）。すなわち、ステップ S 19 において、既存の探索開始点に、新たな探索開始点が追加されることになる。例えば、ステップ S 18 で図 10 に例示する探索ルートを導出した場合、分析部 26 は、その探索ルートの終点に該当する、「B M I」というデータの意味を表すノードを、新たな探索開始点として、既存の探索開始点に追加する。分析部 26 は、ステップ S 19 で定めた探索開始点（換言すれば、新たに追加された探索開始点）に対応するデータの意味を、意味記憶部 4 に記憶させる。

30

【0081】

ステップ S 19 の後、分析部 26 は、ステップ S 14（図 7 参照）以降の処理を繰り返す。

【0082】

ステップ S 14 に移行し、ステップ S 13 で特定されたノードが全てステップ S 15 で選択済みであると判定した場合（ステップ S 14 の Yes）、ステップ 20（図 9 参照）に移行する。

40

【0083】

ステップ S 20 において、分析部 26 は、ステップ S 13 で特定された全てのノードのうちの一部のノードに関して、ステップ S 18 で既存の探索ルートとは異なる新たな探索ルートが導出されたか否かを判定する（ステップ S 20）。

【0084】

ステップ S 13 で特定された全てのノードのうちの一部のノードに関して、ステップ S 18 で新たな探索ルートが導出されていたならば（ステップ S 20 の Yes）、分析部 2

50

6は、ステップS13(図7参照)以降の処理を繰り返す。

【0085】

ステップS13で特定されたいずれのノードに関しても、ステップS18で新たな探索ルートが導出されていないならば(ステップS20のNo)、分析部26は、その時点で得られている探索ルートを、どのような分析を行えるかを示す情報として定め、その情報を出力する(ステップS21)。ステップS21で処理を終了する。

【0086】

図11は、本例で最終的に得られる探索ルートの例を示す模式図である。

【0087】

第2の実施形態によれば、図11に示すように、与えられたデータの意味に対応するノードに基づく探索ルートが得られる。そして、その探索ルート上には、変換ルールIDに対応するノードが含まれている。従って、第1の実施形態と同様に、ユーザに所持されているデータを用いてどのような分析を行えるかを提示することができる。さらに、第2の実施形態では、どのような分析を行えるかを示す情報が、探索ルートの形式で導出されるので、どのような分析を行えるかだけでなく、どのような手順で分析を行うかという分析手順も提示することができる。例えば、図11に例示する探索ルートが提示された場合、ユーザは、BMI計算を行い、その後、癌リスク予測を行うことによって、「癌リスク」が得られるということを理解できる。

10

【0088】

なお、図11では、終点が「癌リスク」のみであるような探索ルートを示しているが、探索ルートが枝分かれていき、終点が複数存在する探索ルートが得られてもよい。そのような探索ルートは、複数種類の分析を行うことができるということを表し、また、それらの分析毎に得られるデータの意味を表している。この点は、後述の第3の実施形態でも同様である。

20

【0089】

実施形態3.

図12は、本発明の第3の実施形態の分析装置の例を示すブロック図である。本実施形態の分析装置1は、取得部2と、意味推定部3と、意味記憶部4と、変換ルール記憶部5と、グラフ生成部7と、グラフ記憶部8と、コスト初期値設定部31と、コスト記憶部32と、分析部36とを備える。

30

【0090】

第3の実施形態では、有向2部グラフの第1の集合に属する個々のノードに対応する個々の「データの意味」にコストが設定される。また、各変換ルールにも予めコストが定められている。また、データを与えるユーザ(データを持つ者)には、コスト上限(以下、max\_costと記す。)が予め定められている。max\_costは、利用可能な変換ルール(換言すれば、利用可能な分析処理)に対する予算の上限を表していると言うことができる。また、max\_costは、所定のコスト上限値と称することもできる。

【0091】

以下、「データの意味」を“s”と表し、そのデータの意味のコストを“s.cost”と表す場合がある。また、ある変換ルールの変換ルールIDを“r”と表し、その変換ルールのコストを“r.cost”と表す場合がある。

40

【0092】

第3の実施形態における取得部2、意味推定部3、意味記憶部4および変換ルール記憶部5は、第1の実施形態や第2の実施形態における取得部2、意味推定部3、意味記憶部4および変換ルール記憶部5と同様であり、説明を省略する。ただし、本実施形態では、変換ルール記憶部5に記憶されている個々の変換ルールには、予め個別にコストが定められている。

【0093】

また、第3の実施形態におけるグラフ生成部7およびグラフ記憶部8は、第2の実施形態におけるグラフ生成部7およびグラフ記憶部8と同様であり、説明を省略する。

50

## 【0094】

コスト初期値設定部31は、有向2部グラフの第1の集合に属する個々のノードに対応する個々のデータの意味に対して、コストの初期値を設定する。このとき、コスト初期値設定部31は、与えられたデータの意味（換言すれば、意味推定部3によって推定されたデータの意味）のコストをそれぞれ“0”に設定し、残りのデータの意味のコストを無限大に設定する。

## 【0095】

コスト記憶部32は、データの意味毎に、データの意味とコストとの組み合わせを記憶する記憶装置である。

## 【0096】

分析部36は、データの意味のコスト、各変換ルールに予め定められたコスト、および、`max_cost`に基づいて、探索開始点からの探索ルートを導出する。

## 【0097】

具体的には、分析部36は、与えられたデータの意味に対応する各ノードをそれぞれ探索開始点と定める。

## 【0098】

その後、分析部36は、以下の処理を繰り返す。

## 【0099】

分析部36は、有向2部グラフにおいて、探索開始点から1つのエッジが向かっている変換ルールIDに対応する各ノードのうち、その探索開始点に対応するデータの意味のコストと、変換ルールIDが表す変換ルールのコストとの和が、`max_cost`以下であるという条件を満たしているノードのみを、その探索開始点から1つのエッジを介して到達されるノードとして特定する。従って、第3の実施形態では、探索開始点に対応するデータの意味のコスト(`si.cost`)と、変換ルールのコスト(`r.cost`)との和が`max_cost`以下であるという条件を満たしている場合にのみ、その変換ルールの変換ルールIDに対応するノードが、その探索開始点から1つのエッジを介して到達されるノードとして特定される。

## 【0100】

そして、分析部36は、特定されたノードに対応する変換ルールIDが表す変換ルールの左辺の「データの意味」に対応する各ノードが全て探索開始点であり、特定されたノードがそれらの探索開始点の全てから到達されている場合に、それらの探索開始点から、その変換ルールの右辺のデータの意味を表すノードまでの探索ルートを導出する。

## 【0101】

そして、分析部36は、上述の右辺のデータの意味を表すノードを探索開始点として定める。このとき、分析部36は、所定の条件が満たされている場合に、その探索開始点に対応するデータの意味のコストを更新する。所定の条件とは、導出された探索ルート上の最後の変換ルールIDが表す変換ルールのコストと、その変換ルールIDに対応するノードに到達する全ての探索開始点に対応する各データの意味のコストの総和との和が、新たに定められた探索開始点に対応するデータの意味のコスト以下であるという条件である。この条件が満たされている場合、分析部36は、新たに定められた探索開始点に対応するデータの意味のコストを、上記の和の値で更新する。また、この条件が満たされていない場合、分析部36は、新たに定められた探索開始点に対応するデータの意味のコストを更新しない。

## 【0102】

分析部36は、新たな探索ルートを導出できなくなった時点までに導出された探索ルートを、どのような分析を行えるかを示す情報として定める。

## 【0103】

コスト初期値設定部31および分析部36は、例えば、分析プログラムに従って動作するコンピュータのCPUによって実現される。また、コスト記憶部32は、例えば、コンピュータが備える記憶装置によって実現される。

## 【0104】

10

20

30

40

50

次に、第3の実施形態の処理経過について説明する。図13、図14および図15は、第3の実施形態の処理経過の例を示すフローチャートである。ただし、既に説明した事項については、詳細な説明を省略する。また、第2の実施形態と同様の処理についても、図7、図8および図9と同一のステップ番号を付し、詳細な説明を省略する。

#### 【0105】

なお、変換ルール記憶部5は、図5に例示する各変換ルールを予め記憶しているものとする。ただし、変換ルール「BMI計算」、「年齢推定」、「癌リスク予測」、「年収推定」、「保険金算出」および「性別判定」にはそれぞれ、コストとして、“5”, “10”, “20”, “20”, “12”および“9”が予め定められているものとする。また、max\_costの値も予め定められているものとする。 10

#### 【0106】

また、グラフ生成部7は、その各変換ルールに基づいて、図6に例示する有向2部グラフを既に生成しており、その有向2部グラフをグラフ記憶部8に記憶させているものとする。また、取得部2がデータ（テーブル）を取得しているものとする。

#### 【0107】

まず、意味推定部3が、取得部2が取得したデータの意味を推定する（ステップS11）。以下、ステップS11で得られたデータの意味を、与えられたデータの意味と記す。本例では、与えられたデータの意味が、{「身長」、「体重」、「年収」、「年齢」}であるものとする。

#### 【0108】

第3の実施形態では、ステップS11の次に、コスト初期値設定部31が、有向2部グラフの第1の集合に属する個々のノードに対応する個々のデータの意味のコストの初期値を設定する（ステップS31）。具体的には、コスト初期値設定部31は、与えられたデータの意味{「身長」、「体重」、「年収」、「年齢」}のコストをそれぞれ0に設定し、残りのデータの意味のコストを無限大に設定する。図16は、ステップS31で設定されたコストを有向2部グラフとともに示した模式図である。図16では、予め定められている変換ルールのコストも図示している。 20

#### 【0109】

なお、ステップS31において、コスト初期値設定部31は、データの意味毎に、データの意味とコストとの組み合わせをコスト記憶部32に記憶させる。 30

#### 【0110】

ステップS31の次に、分析部36は、有向2部グラフにおいて、与えられたデータの意味に対応する各ノードをそれぞれ探索開始点と定める（ステップS12）。本例では、「身長」、「体重」、「年収」、「年齢」に対応するそれぞれのノードを、探索開始点とする。

#### 【0111】

次に、分析部36は、探索開始点から1つのエッジが向かっている変換ルールIDに対応する各ノードのうち、その探索開始点に対応するデータの意味のコストと、変換ルールIDが表す変換ルールのコストとの和が、max\_cost以下であるという条件を満たしているノードのみを、その探索開始点から1つのエッジを介して到達されるノードとして特定する（ステップS32）。 40

#### 【0112】

例えば、「身長」と「BMI計算」とに着目して説明する。「身長」に対応するノードは探索開始点である。また、「身長」のコストは“0”であり、変換ルール「BMI計算」のコストは“5”である（図16参照）。換言すれば、“身長.cost = 0”、“BMI計算.cost = 5”である。従って、「“身長.cost” + “BMI計算.cost”」の値がmax\_cost以下であるならば、「BMI計算」に対応するノードは、「身長」に対応する探索開始点から1つのエッジを介して到達されるノードとして特定される。一方、「“身長.cost” + “BMI計算.cost”」の値がmax\_costよりも大きいならば、「BMI計算」に対応するノードは、「身長」に対応する探索開始点から1つのエッジを介して到達されるノードとして特定されな 50

い。

#### 【0113】

また、ステップS32において、探索開始点から1つのエッジを介して到達されるノードが1つも特定されない場合には、その時点で得られている探索ルートを、どのような分析を行えるかを示す情報として定め、その情報を出力し、処理を終了する。

#### 【0114】

ステップS32の次に、分析部36は、ステップS14以降の処理を行う。ステップS14～S19の動作は、第2の実施形態におけるステップS14～S19の動作と同様であり、説明を省略する。ただし、分析部36は、ステップ14において、ステップS32で特定されたノードが全てステップS15で選択済みであるか否かを判定する。また、分析部36は、ステップS15において、ステップS32で特定されたノードのうち、未選択のノードを1つ選択する。

10

#### 【0115】

ステップS19の後、分析部36は、ステップS18で導出された探索ルート上の最後の変換ルールIDが表す変換ルールのコストと、その変換ルールIDに対応するノードに到達する全ての探索開始点に対応する各データの意味のコストの総和との和が、ステップS19で新たに定められた探索開始点に対応するデータの意味のコスト以下であるという条件を満たすか否かを判定する。その条件が満たされているならば、新たに定められた探索開始点に対応するデータの意味のコストを、上記の和の値で更新する。その条件が満たされていないならば、新たに定められた探索開始点に対応するデータの意味のコストを更新しない(ステップS33)。

20

#### 【0116】

例えば、ステップS18において、図17に示す探索ルートが導出されたとする。この場合、探索ルート上の最後の変換ルールIDは「BMI計算」であり、変換ルール「BMI計算」のコストは“5”である。また、「BMI計算」に対応するノードに到達する全ての探索開始点は、「身長」に対応する探索開始点、および、「体重」に対応する探索開始点である。そして、「身長」のコスト、および、「体重」のコストはいずれも“0”であるので、それらの総和も“0”である。従って、変換ルール「BMI計算」のコスト“5”と、上記の総和“0”との和は“5”である。

30

#### 【0117】

また、図17に示す探索ルートでは、「BMI」に対応する探索開始点が、新たに定められた探索開始点である。そして、「BMI」のコストは無限大である。従って、上記の和“5”は、「BMI」のコスト“無限大”以下であるので、ステップS33における条件は満たされている。従って、分析部36は、「BMI」のコストを“無限大”から、上記の和“5”に更新する。

#### 【0118】

仮に、「BMI」のコストが“4”であるとすると、ステップS33における条件は満たされないことになる。この場合、分析部36は、「BMI」のコストを更新せず、“4”的まととする。

40

#### 【0119】

分析部36は、データの意味のコストを更新する場合、コスト記憶部32に記憶されているデータの意味のコストを更新すればよい。

#### 【0120】

ステップS33の後、分析部36は、ステップS14(図13参照)以降の処理を繰り返す。

#### 【0121】

また、ステップS20, S21は、第2の実施形態におけるステップS20, S21と同様であり、説明を省略する。ただし、ステップS20において、分析部36は、ステップS32で特定された全てのノードのうちの一部のノードに関して、ステップS18で既存の探索ルートとは異なる新たな探索ルートが導出されたか否かを判定する

50

## 【0122】

第3の実施形態においても、第2の実施形態と同様の効果が得られる。さらに、第3の実施形態では、`max_cost`の値に応じて、得られる探索ルートの長さが異なる。すなわち、第3の実施形態では、`max_cost`の値に応じた長さの探索ルートが得られる。このことは、例えば、データを持つ者の予算の範囲内で行える分析を提示できるということを意味している。

## 【0123】

上記の例において、例えば、`max_cost`が“2”であるならば、探索ルートは得られずに処理が終了する。また、上記の例において、例えば、`max_cost`が“15”であるならば、図10に示す探索ルートが導出され、処理が終了する。また、上記の例において、例えば、`max_cost`が“30”であるならば、図11に示す探索ルートが導出され、処理が終了する。このように、第3の実施形態では、`max_cost`の値に応じた長さの探索ルートが得られる。10

## 【0124】

また、第2の実施形態で説明したように、探索ルートが枝分かれしていき、終点が複数存在する探索ルートが得られてもよい。

## 【0125】

また、第2の実施形態や第3の実施形態において、データを持つ者が分析によって得たいと考えているデータがあるならば、そのデータの意味が指定されてもよい。そして、第2の実施形態の分析部26や第3の実施形態の分析部36は、探索ルートが枝分かれしていき終点が複数存在する探索ルートを得た後、指定された「データの位置」を終点とする探索ルートのみを抽出し、その探索ルートを出力してもよい。20

## 【0126】

次に、各実施形態の変形例について説明する。本発明の各実施形態において、分析装置1に、意味推定部3が設けられていなくてもよい。その場合、取得部2は、ユーザに所持されているデータの意味を直接取得してもよい。すなわち、取得部2が、1つ以上の「データの意味」を直接外部から取得してよい。この場合、その「データの意味」を、上記の各実施形態における「与えられたデータの意味」として扱えばよい。

## 【0127】

図18は、本発明の各実施形態の分析装置1に係るコンピュータの構成例を示す概略ブロック図である。例えば、コンピュータ1000は、CPU1001と、主記憶装置1002と、補助記憶装置1003と、インタフェース1004と、データ(テーブル)を読み込むデータ読み込み装置1005とを備える。30

## 【0128】

本発明の各実施形態の分析装置1は、コンピュータ1000によって実現される。分析装置1の動作は、プログラム(分析プログラム)の形式で補助記憶装置1003に記憶されている。CPU1001は、プログラムを補助記憶装置1003から読み出し、そのプログラムを主記憶装置1002に展開し、そのプログラムに従って、上記の各実施形態で説明した処理を実行する。

## 【0129】

補助記憶装置1003は、一時的でない有形の媒体の例である。一時的でない有形の媒体の他の例として、インタフェース1004を介して接続される磁気ディスク、光磁気ディスク、CD-ROM(Compact Disk Read Only Memory)、DVD-ROM(Digital Versatile Disk Read Only Memory)、半導体メモリ等が挙げられる。また、プログラムが通信回線によってコンピュータ1000に配信される場合、配信を受けたコンピュータ1000がそのプログラムを主記憶装置1002に展開し、そのプログラムに従って上記の各実施形態で説明した処理を実行してもよい。40

## 【0130】

また、各構成要素の一部または全部は、汎用または専用の回路(circuitry)、プロセッサ等やこれらの組合せによって実現されてもよい。これらは、単一のチップによって構成されてもよいし、バスを介して接続される複数のチップによって構成されてもよい。各50

構成要素の一部または全部は、上述した回路等とプログラムとの組合せによって実現されてもよい。

【0131】

各構成要素の一部または全部が複数の情報処理装置や回路等により実現される場合には、複数の情報処理装置や回路等は集中配置されてもよいし、分散配置されてもよい。例えば、情報処理装置や回路等は、クライアントアンドサーバシステム、クラウドコンピューティングシステム等、各々が通信ネットワークを介して接続される形態として実現されてもよい。

【0132】

次に、本発明の概要について説明する。図19は、本発明の分析装置の概要の例を示すブロック図である。本発明の分析装置は、変換ルール記憶手段75と、分析手段76とを備える。

10

【0133】

変換ルール記憶手段75（例えば、変換ルール記憶部5）は、1つ以上のデータの意味を、1つ以上の別のデータの意味に変換する変換ルールを複数個記憶する。

【0134】

分析手段76（例えば、分析部6、分析部26または分析部36）は、与えられた1つ以上のデータの意味と、変換ルールとに基づいて、その意味を持つデータを用いてどのような分析を行えるかを示す情報を導出する。

【0135】

そのような構成によって、所持されているデータを用いてどのような分析を行えるかを示す情報を導出することができる。

20

【0136】

また、各変換ルールに基づいて、データの意味を表すノードの集合と、変換ルールIDを表すノードの集合とを含む有向2部グラフを生成するグラフ生成手段（例えば、グラフ生成部7）を備え、

分析手段76（例えば、分析部26）が、

与えられた1つ以上のデータの意味に対応する各ノードを探索開始点と定めた後に、

有向2部グラフにおいて、探索開始点から1つのエッジを介して到達する変換ルールIDに対応するノードを特定し、

30

特定されたノードに対応する変換ルールIDが表す変換ルールにおける変換前のデータの意味に対応する各ノードが全て探索開始点であり、特定されたノードが当該探索開始点の全てから到達されている場合に、その変換ルールにおける変換後のデータの意味を表すノードまでの探索ルートを導出し、

その変換後のデータの意味を表すノードを探索開始点として定めること

を繰り返し、

新たな探索ルートが導出できなくなった時点までに導出された探索ルートを、上記の情報として定める

構成であってもよい。

【0137】

また、個々の変換ルールにそれぞれコストが予め定められていて、

各変換ルールに基づいて、データの意味を表すノードの集合と、変換ルールIDを表すノードの集合とを含む有向2部グラフを生成するグラフ生成手段（例えば、グラフ生成部7）と、

有向2部グラフにおける個々のデータの意味に対して、コストの初期値を設定するコスト初期値設定手段（例えば、コスト初期値設定部31）とを備え、

分析手段76（例えば、分析部36）が、

与えられた1つ以上のデータの意味に対応する各ノードを探索開始点と定めた後に、

有向2部グラフにおいて、探索開始点から1つのエッジが向かっている変換ルールIDに対応するノードのうち、その探索開始点に対応するデータの意味のコストと、変換ル

40

50

ルール ID が表す変換ルールのコストとの和が、所定のコスト上限値以下であるという条件を満たすノードのみを、その探索開始点から 1 つのエッジを介して到達されるノードとして特定し、

特定されたノードに対応する変換ルール ID が表す変換ルールにおける変換前のデータの意味に対応する各ノードが全て探索開始点であり、特定されたノードが当該探索開始点の全てから到達されている場合に、その変換ルールにおける変換後のデータの意味を表すノードまでの探索ルートを導出し、

その変換後のデータの意味を表すノードを探索開始点として定めるとともに、所定の条件を満たす場合に、当該ノードに対応するデータの意味のコストを更新すること

を繰り返し、

新たな探索ルートが導出できなくなった時点までに導出された探索ルートを、上記の情報として定める

構成であってもよい。

#### 【0138】

また、コスト初期値設定手段が、

与えられた 1 つ以上のデータの意味のコストをそれぞれ 0 に設定し、残りのデータの意味のコストを無限大に設定し、

分析手段 76 ( 例えば、分析部 36 ) が、

導出された探索ルート上の最後の変換ルール ID が表す変換ルールのコストと、その変換ルール ID に対応するノードに到達する全ての探索開始点に対応する各データの意味のコストの総和との和が、その変換ルールにおける変換後のデータの意味のコスト以下である場合に、当該データの意味のコストを、上記の和で更新する

構成であってもよい。

#### 【0139】

また、分析手段 76 ( 例えば、分析部 6 ) が、

変換前のデータの意味が、与えられた 1 つ以上のデータの意味に包含されているという条件を満たす変換ルールを抽出し、

抽出した各変換ルールにおける変換後のデータの意味と、与えられた 1 つ以上のデータの意味との和集合を求め、当該和集合を与えられた 1 つ以上のデータの意味とみなすこと

を繰り返し、抽出した変換ルールと、前回抽出した変換ルールとが同一になったならば、その抽出した変換ルールの集合を、上記の情報として定める

構成であってもよい。

#### 【0140】

また、データが与えられた場合に、当該データの意味を推定する意味推定手段 ( 例えば、意味推定部 3 ) を備える構成であってもよい。

#### 【0141】

以上、実施形態を参照して本願発明を説明したが、本願発明は上記の実施形態に限定されるものではない。本願発明の構成や詳細には、本願発明のスコープ内で当業者が理解し得る様々な変更をすることができる。

#### 【産業上の利用の可能性】

#### 【0142】

本発明は、データの意味が与えられた場合に、その意味を持つデータを用いてどのような分析を行えるのかを分析する分析装置に好適に適用可能である。

#### 【符号の説明】

#### 【0143】

1 分析装置

2 取得部

3 意味推定部

4 意味記憶部

10

20

30

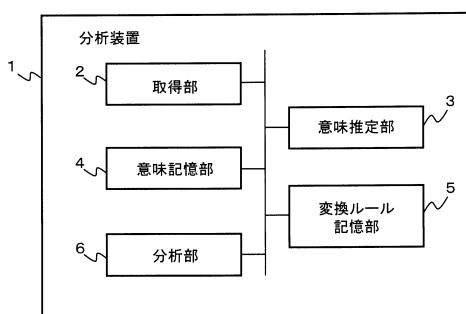
40

50

- 5 変換ルール記憶部  
 6, 26, 36 分析部  
 7 グラフ生成部  
 8 グラフ記憶部  
 31 コスト初期値設定部  
 32 コスト記憶部

【図面】

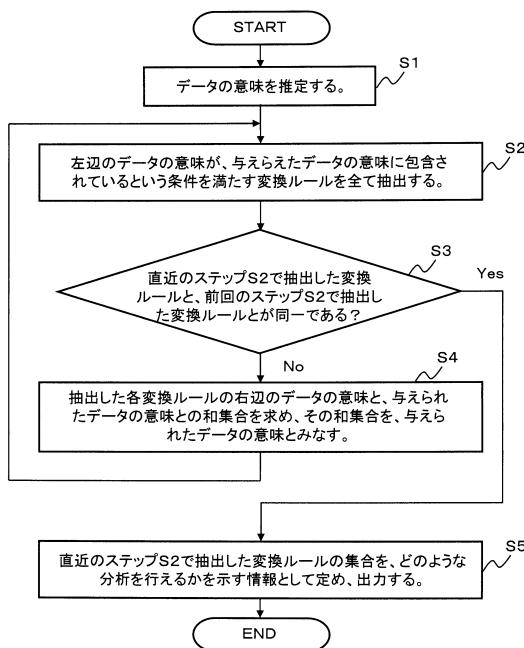
【図 1】



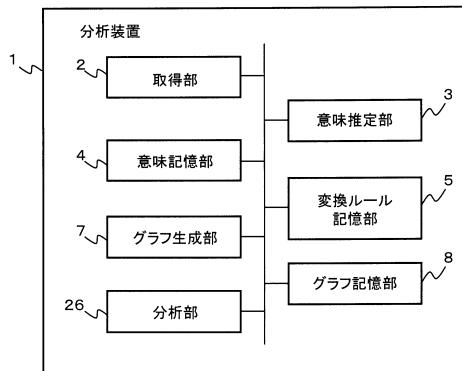
【図 2】



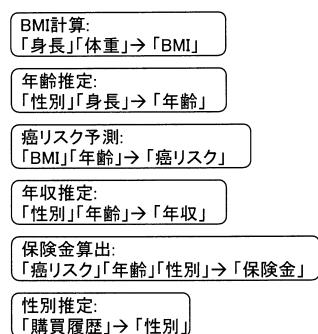
【図 3】



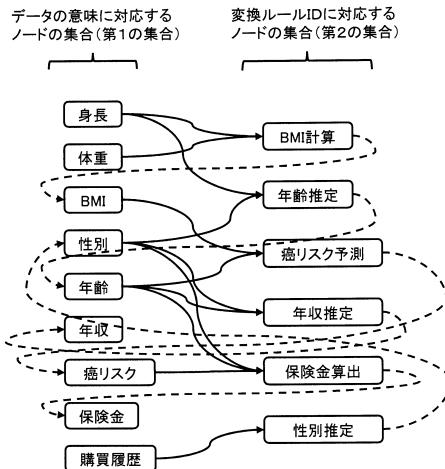
【図 4】



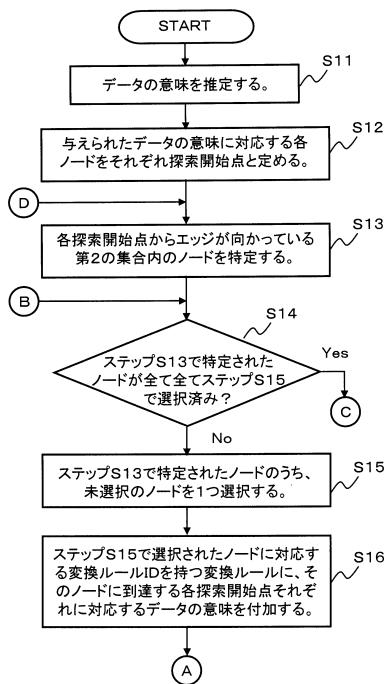
【 図 5 】



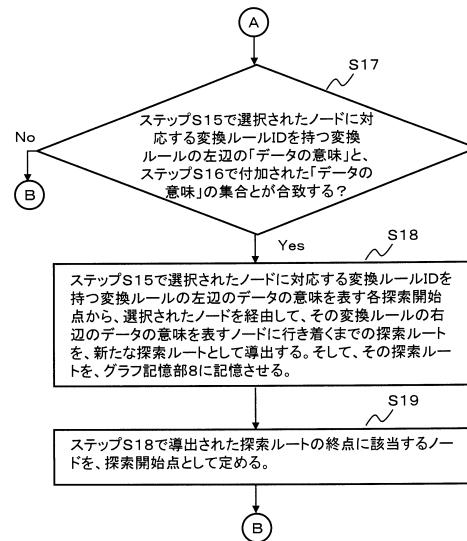
【 図 6 】



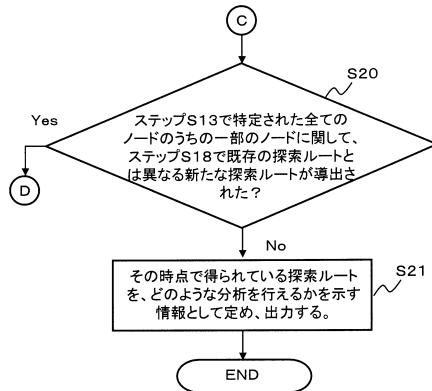
【 図 7 】



【 四 8 】



【図 9】

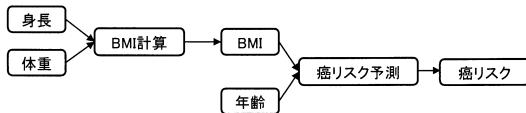


【図 10】

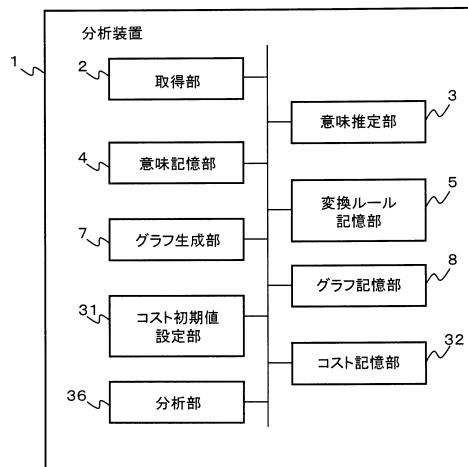


10

【図 11】



【図 12】



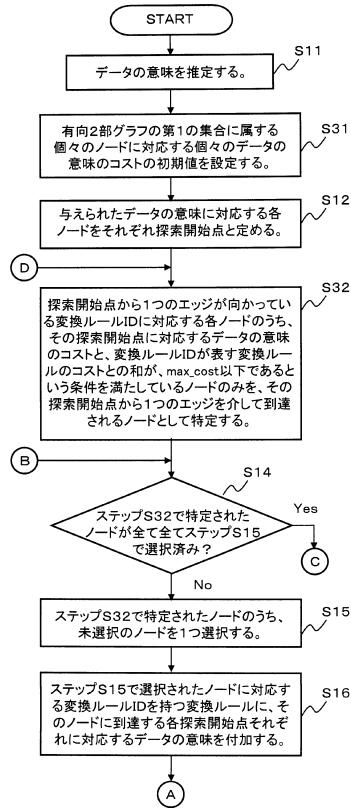
20

30

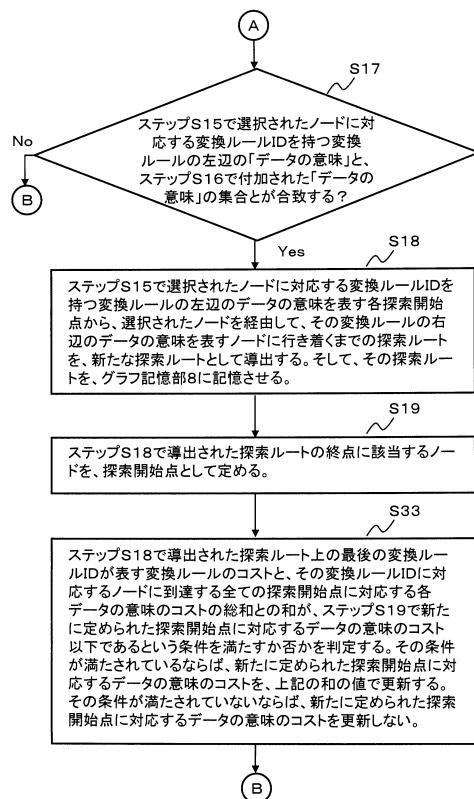
40

50

【図13】



【図14】



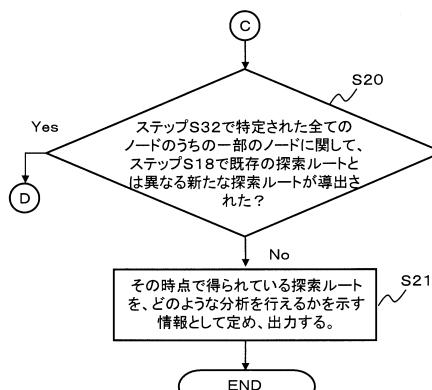
10

20

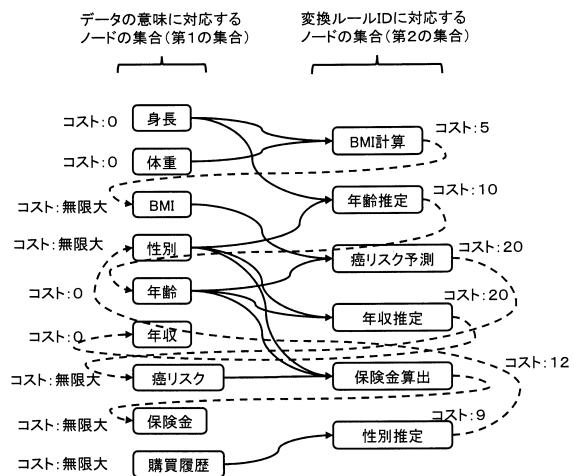
30

40

【図15】

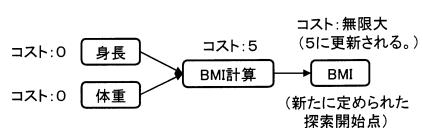


【図16】

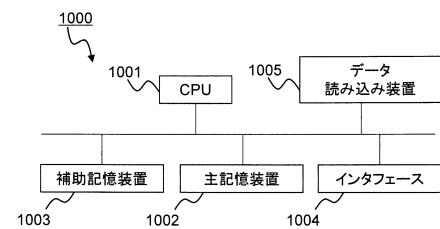


50

【図17】



【図18】



10

【図19】



20

30

40

50

---

フロントページの続き

(56)参考文献      特開2007-241901(JP, A)  
                    特開平08-077010(JP, A)  
                    特開2008-310650(JP, A)

(58)調査した分野 (Int.Cl., DB名)  
G 06 F 16 / 00 - 16 / 958