

(19) 日本国特許庁 (JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2020-4001

(P2020-4001A)

(43) 公開日 令和2年1月9日 (2020. 1. 9)

(51) Int.Cl.

G06N 7/00 (2006.01)

F I

G06N 7/00 150

テーマコード (参考)

審査請求 未請求 請求項の数 8 O L (全 15 頁)

(21) 出願番号 特願2018-121859 (P2018-121859)
 (22) 出願日 平成30年6月27日 (2018. 6. 27)

(特許庁注：以下のものは登録商標)

1. J A V A S C R I P T

(71) 出願人 000005108
 株式会社日立製作所
 東京都千代田区丸の内一丁目6番6号
 (74) 代理人 110000198
 特許業務法人湘洋内外特許事務所
 (72) 発明者 今沢 慶
 東京都千代田区丸の内一丁目6番6号 株
 式会社日立製作所内

(54) 【発明の名称】 因果推論モデル構築方法

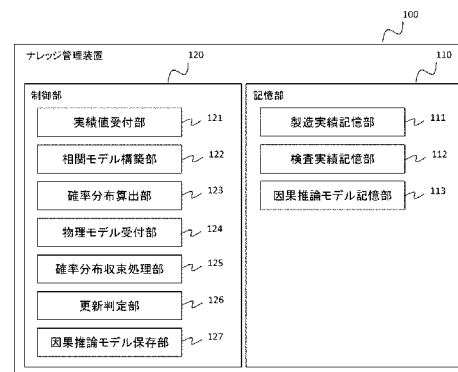
(57) 【要約】 (修正有)

【課題】製造現場において、成立することが確認されている物理モデルなどのナレッジを他の製造現場から相互利用可能な形にモデル化することで、ナレッジの抽出を容易化する。

【解決手段】コンピュータを用いた因果推論モデル構築方法であって、コンピュータは、制御部と、記憶部と、を備える。制御部は、製造工程および検査工程で取得された製造ログデータを読み込む実績値受付ステップと、製造ログデータを用いて相関モデルを構築する相関モデル構築ステップと、物理モデルの入力を受け付ける物理モデル受付ステップと、相関モデルおよび物理モデルの確率分布を算出する確率分布算出ステップと、相関モデルの確率分布を収束計算し物理モデルの確率分布に近似させ、相関モデルを上書きする確率分布収束処理ステップと、相関モデルを因果推論モデルとして記憶部に格納する因果推論モデル保存ステップと、を実施する。

【選択図】図1

図1



【特許請求の範囲】**【請求項 1】**

コンピュータを用いた因果推論モデル構築方法であって、
前記コンピュータは、制御部と、記憶部と、を備え、
前記制御部は、
製造工程および検査工程のいずれかまたは両方で取得されたセンサデータを含む製造ログデータを前記記憶部から読み込む実績値受付ステップと、
前記製造ログデータを用いて相関モデルを構築する相関モデル構築ステップと、
前記製造工程または検査工程に係る物理法則をモデル化した物理モデルの入力を受け付ける物理モデル受付ステップと、
前記相関モデル構築ステップで構築された相関モデルおよび前記物理モデルの確率分布を算出する確率分布算出ステップと、
前記相関モデルの確率分布を収束計算し前記物理モデルの確率分布に近似させ、前記相関モデルを上書きする確率分布収束処理ステップと、
確率分布収束処理ステップの処理結果の相関モデルを因果推論モデルとして前記記憶部に格納する因果推論モデル保存ステップと、を実施する、
ことを特徴とする因果推論モデル構築方法。

10

【請求項 2】

請求項 1 に記載された因果推論モデル構築方法であって、
前記相関モデル構築ステップにおいて、ベイジアンネットワークを用いる、
ことを特徴とする因果推論モデル構築方法。

20

【請求項 3】

請求項 1 に記載された因果推論モデル構築方法であって、
前記物理モデル受付ステップでは、陰関数の形式で記述された物理モデルを受け付ける、
ことを特徴とする因果推論モデル構築方法。

【請求項 4】

請求項 1 に記載された因果推論モデル構築方法であって、
前記確率分布収束処理ステップでは、前記相関モデルの確率分布について摂動処理を行い、前記物理モデルの確率分布とのノルムにより近似を判定する、
ことを特徴とする因果推論モデル構築方法。

30

【請求項 5】

請求項 1 に記載された因果推論モデル構築方法であって、
前記確率分布収束処理ステップでは、前記相関モデルの確率分布について摂動処理を行い、当該摂動処理においては、所定の分布を有するランダムサンプリングに従った誤差値を摂動量として用い、前記物理モデルの確率分布とのノルムにより近似を判定する、
ことを特徴とする因果推論モデル構築方法。

【請求項 6】

請求項 1 に記載された因果推論モデル構築方法であって、
前記確率分布収束処理ステップでは、前記相関モデルの確率分布について摂動処理を繰り返し行い、当該摂動処理においては、所定の分布を有するランダムサンプリングに従った誤差値を徐々に小さな値に収束する摂動量として用い、前記物理モデルの確率分布とのノルムにより近似を判定する、
ことを特徴とする因果推論モデル構築方法。

40

【請求項 7】

請求項 1 に記載された因果推論モデル構築方法であって、
前記確率分布収束処理ステップでは、前記相関モデルの確率分布について摂動処理を繰り返し行い、当該摂動処理においては、所定の分布を有するランダムサンプリングに従った誤差値を徐々に小さな値に収束する摂動量として用い、前記物理モデルの確率分布とのノルムが前記摂動処理前の確率分布と前記物理モデルの確率分布との間のノルムよりも小

50

さくなる場合には、前記摂動処理後の確率分布により前記摂動処理前の確率分布を上書きする、

ことを特徴とする因果推論モデル構築方法。

【請求項 8】

請求項 1 に記載された因果推論モデル構築方法であって、

前記因果推論モデル保存ステップでは、JSON フォーマットにより保存する、

ことを特徴とする因果推論モデル構築方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、因果推論モデル構築方法に関する。

【背景技術】

【0002】

本技術分野の背景技術として、特開 2013-008221 号公報(特許文献 1)がある。この文献には、「計算処理装置は、記号論理における述語論理の推論を実行するための後ろ向き推論の形式のデータフローグラフであって、複数の演算ノードおよび演算ノードの間を各々つなぐ複数のエッジにより述語論理の構造を表わすデータフローグラフに変換するネットワーク作成部と、変換されたデータフローグラフに基づくトークンの伝搬処理により、述語論理に対する導出を実行する計算実行部とを備える。」と記載されている。

【先行技術文献】

【特許文献】

【0003】

【特許文献 1】特開 2013-008221 号公報

【発明の概要】

【発明が解決しようとする課題】

【0004】

上記特許文献 1 に記載された技術では、記号論理による論理推論を実行するための方法が記載されている。しかし、記号論理におけるデータフローだけでは、製造現場において知られている物理モデルのような定量的な因果関係をモデル化することができない。

【0005】

本発明の目的は、製造現場において、成立することが確認されている物理モデルなどのナレッジを他の製造現場から相互利用可能な形にモデル化することで、ナレッジの抽出を容易化する。

【課題を解決するための手段】

【0006】

本願は、上記課題の少なくとも一部を解決する手段を複数含んでいるが、その例を挙げるならば、以下のとおりである。上記課題を解決すべく、本発明の一態様に係る因果推論モデル構築方法は、コンピュータを用いた因果推論モデル構築方法であって、上記コンピュータは、制御部と、記憶部と、を備え、上記制御部は、製造工程または検査工程のいずれかまたは両方で取得されたセンサデータを含む製造ログデータを上記記憶部から読み込む実績値受付ステップと、上記製造ログデータを用いて相関モデルを構築する相関モデル構築ステップと、上記製造工程に係る物理法則をモデル化した物理モデルの入力を受け付ける物理モデル受付ステップと、上記相関モデル構築ステップで構築された相関モデルおよび上記物理モデルの確率分布を算出する確率分布算出ステップと、上記相関モデルの確率分布を収束計算し上記物理モデルの確率分布に近似させ、上記相関モデルを上書きする確率分布収束処理ステップと、確率分布収束処理ステップの処理結果の相関モデルを因果推論モデルとして上記記憶部に格納する因果推論モデル保存ステップと、を実施する、ことを特徴とする。

【発明の効果】

【0007】

10

20

30

40

50

本発明によれば、製造現場において、成立することが確認されている物理モデルなどのナレッジを相互利用可能な形にモデル化することで、ナレッジの抽出を容易化する技術を実現することができる。上記した以外の課題、構成及び効果は、以下の実施形態の説明により明らかにされる。

【図面の簡単な説明】

【0008】

【図1】本発明の第一の実施形態に係るナレッジ管理装置の構成例を示す図である。

【図2】製造実績記憶部のデータ構造例を示す図である。

【図3】検査実績記憶部のデータ構造例を示す図である。

【図4】ナレッジ管理装置のハードウェア構成例を示す図である。

10

【図5】因果推論モデルの図示の例を示す図である。

【図6】因果関係の図示の例を示す図である。

【図7】ナレッジ生成処理のフローの例を示す図である。

【図8】確率分布の収束計算処理のフローの例を示す図である。

【図9】ランダムサンプリングの処理の例を示す図である。

【図10】因果推論モデルのJSONを用いたデータ構造例を示す図である。

【図11】ナレッジ生成処理の出力画面の例を示す図である。

【発明を実施するための形態】

【0009】

以下、本発明に係る実施の形態を図面に基づいて説明する。なお、実施の形態を説明するための全図において、同一の部材には原則として同一の符号を付し、その繰り返しの説明は省略する。また、以下の実施の形態において、その構成要素（要素ステップ等も含む）は、特に明示した場合および原理的に明らかに必須であると考えられる場合等を除き、必ずしも必須のものではないことは言うまでもない。また、「Aからなる」、「Aよりなる」、「Aを有する」、「Aを含む」と言うときは、特にその要素のみである旨明示した場合等を除き、それ以外の要素を排除するものでないことは言うまでもない。同様に、以下の実施の形態において、構成要素等の形状、位置関係等に言及するときは、特に明示した場合および原理的に明らかにそうでないと考えられる場合等を除き、実質的にその形状等に近似または類似するもの等を含むものとする。

20

【0010】

統計的な概念として、相関と因果という、一見類似するが相異なる概念がある。「相関」関係は、一方の変数の値が大きいに他方の変数の値も大きい、というような関係をいう。例えば、体重と身長の関係がこのような関係となる。「因果」関係は、一方の変数の値を変化させたときに、他方の変数の値も変化するというような関係をいい、例えば照明のスイッチと部屋の明るさ、のような関係がこれに相当する。

30

【0011】

ここで、相関関係が因果関係と認められるためには、「相関の強さ（Strength）」、「相関の確からしさ（Consistency）」、「相関の特異性（Specificity）」、「量・反応関係の成立（Biological Gradient）」、「時間的先行性（Temporality）」、「妥当性（Plausibility）」、「先行知見との整合（Coherence）」、「他の知見との整合（Analogy）」、「実験との一致（Experiment）」等の要件を満たす必要があるとの考え方がある（Hillのガイドライン等）。

40

【0012】

そして、情報処理システムを用いて解析したさまざまな測定値の相関関係を利用して、製造工程あるいは検査工程において望む結果を制御する種々の方法が研究されつつある。例えば、ビッグデータ解析等によれば、おおまかな変数（センサー等の取得値）間の相関関係を解析することができるようになりつつある。一方で、相関関係の中には、因果関係を有する関係も含まれるが、因果関係とは異なる関係も含まれるため、解析した相関関係を具体的な製造現場の効率化に適用しようとしても、定量的な因果関係を特定できなけれ

50

ば結果の制御には貢献させづらい。

【0013】

数多ある測定項目の相関関係の中から、因果関係を効率よく抽出することができれば、定量的な因果関係を特定することも容易となる。また、そのような因果関係をナレッジとして蓄積することにより、他事象への展開効率を高めることが可能となると考えられる。そして、製造工程および検査工程にあてはめると、上記の「妥当性 (Plausibility)」は、物理モデルと一致すること (物理法則に反しないこと) と換言し得る。そのため、抽出された相関モデルが、物理モデルと一致するか否かを判定することは、相関関係が因果関係となりうるか否かを判定することと同様の意味がある。

【0014】

図1は、本発明の第一の実施形態に係るナレッジ管理装置の構成例を示す図である。ナレッジ管理装置100は、記憶部110と、制御部120と、を備える。記憶部110には、製造実績記憶部111と、検査実績記憶部112と、因果推論モデル記憶部113と、が含まれる。制御部120には、実績値受付部121と、相関モデル構築部122と、確率分布算出部123と、物理モデル受付部124と、確率分布収束処理部125と、更新判定部126と、因果推論モデル保存部127と、が含まれる。

【0015】

図2は、製造実績記憶部のデータ構造例を示す図である。製造実績記憶部111には、pVar1(111a)と、pVar2(111b)と、pVar3(111c)と、pVar4(111d)と、pVar5(111e)と、pVar6(111f)と、pVar7(111g)と、pVar8(111h)と、pVar9(111j)と、pVar10(111k)と、が含まれる。

【0016】

製造実績記憶部111には、製造装置の操業状態をモニタした製造データが含まれており、その製造データの個別の項目が上記pVar1乃至pVar10の項目名で示されている。例えば、pVar1(111a)は、製品の個体を識別するためのID番号を示す項目名である。2列目以降のpVar2(111b)乃至pVar10(111k)は、それぞれ製造装置の操業状態をセンサ等によりモニタした結果得られたデータである。操業状態をモニタする項目の例としては、処理中の温度、湿度、圧力、電流、電圧、物質質量等がある。

【0017】

これらのデータは、一般に、製品を製造処理している際に定期サンプリングされるデータである。ここで、定期サンプリングとは、例えば、1秒間に100回の頻度等、各種のモニタする項目に応じて所定の頻度で定期的にセンサデータを取得することを指す。

【0018】

一般に、製品の製造処理時間は、定期サンプリングのサンプリング間隔よりも長い。そのため、一つの製品を処理している間に、同じセンサから取得されるデータを複数回取得することになる。したがって、図2のように、1列目に製品の個体を識別するためのID番号を配置する場合、サンプリングデータを複数回分得られるため、同じID番号のデータに重複が生じてしまう。

【0019】

そこで、製品の個体を識別するためのID番号を、各行ベクトルを一意に特定するためのユニークキーとして扱い、各項目には複数回取得したデータの統計値 (平均値や中央値等) を用いることで、製品の個体を識別するためのID番号に重複が生じないようにデータを整形しておくものとする。

【0020】

なお、操業状態をモニタした他の例として、製品の処理に要した処理時間がある。このようなデータは、製品を一つ処理するたびに、1点のデータ (処理時間) が得られる。したがって、このようなデータでは、製品の個体を識別するためのID番号に重複が生じることはないため、統計処理を行わなくても、そのまま用いることができる。

10

20

30

40

50

【 0 0 2 1 】

図 3 は、検査実績記憶部のデータ構造例を示す図である。検査実績記憶部 1 1 2 には、c V a r 1 (1 1 2 a) と、c V a r 2 (1 1 2 b) と、c V a r 3 (1 1 2 c) と、c V a r 4 (1 1 2 d) と、c V a r 5 (1 1 2 e) と、c V a r 6 (1 1 2 f) と、c V a r 7 (1 1 2 g) と、c V a r 8 (1 1 2 h) と、c V a r 9 (1 1 2 j) と、c V a r 1 0 (1 1 2 k) と、が含まれる。

【 0 0 2 2 】

検査実績記憶部 1 1 2 には、検査装置の測定結果である検査データが含まれており、その検査データの個別の項目が上記 c V a r 1 乃至 c V a r 1 0 の項目名で示されている。例えば、c V a r 1 (1 1 2 a) は、製品の個体を識別するための ID 番号を示す項目名である。本データと、製造実績記憶部 1 1 1 の p V a r 1 (1 1 1 a) は対応関係にあり、同じ値は同じ個体であることを示している。したがって、本データを参照することで、製造装置の操業状態をモニタした製造データと検査データの対応関係を取得することができる。

10

【 0 0 2 3 】

2 列目以降の c V a r 2 (1 1 2 b) 乃至 c V a r 1 0 (1 1 2 k) は、それぞれ検査装置がセンサ等により測定した結果得られたデータである。検査データの例としては、製品の特定の箇所の寸法といった物理的な大きさに関連した測定値や、電気的な特性に関する測定値等がある。

【 0 0 2 4 】

20

このような検査データは、数値として測定される。そして、検査データには、規格が設定されていて、規格内に入っているかどうかの判定を行うことがある。上述の c V a r 1 0 (1 1 2 k) が、製品が規格内に入っているか否かの判定結果を表すデータである。本例では、規格に入った場合は、「 O K 」、入らなかった場合は、「 N G 」と格納されている。

【 0 0 2 5 】

このような規格に対する判定結果は、個々の測定値ごとに行い、個々の測定値ごとに対する判定結果を全て記録している場合と、全ての測定値に対する総合判定結果として記録している場合等があるが、図 3 の例においては、判定結果が一つしか記載されておらず、これは、総合判定結果が記載されていることを意味する。

30

【 0 0 2 6 】

実績値受付部 1 2 1 は、製造装置の操業状態をモニタした製造データと、検査データと、のいずれかまたは両方を含む製造ログデータの読み込みを行う。具体的には、実績値受付部 1 2 1 は、記憶部 1 1 0 の製造実績記憶部 1 1 1 と、検査実績記憶部 1 1 2 と、を読み出す。

【 0 0 2 7 】

相関モデル構築部 1 2 2 は、製造ログデータを用いて、相関モデルを構築する。具体的には、相関モデル構築部 1 2 2 は、厳密に因果関係にあるか否かは問わず、K 2 アルゴリズム等の構造学習アルゴリズムを用いて、相関性のある項目を原因と結果に関連付けて、ベイジアンネットワーク (B a y e s i a n N e t w o r k) を構成し、一般に用いられる X M L (e X t e n s i b l e M a r k u p L a n g u a g e) の拡張データ等により保存する。これをグラフィカルデータとして描画すると、図 5 に示されるようなエッジとノードを有する有向モデルとなる。

40

【 0 0 2 8 】

図 5 は、因果推論モデルの図示の例を示す図である。因果推論モデルとは、原因と結果の関係を、頂点、すなわちノードと、矢線、すなわちエッジを用いて示したグラフィカルモデルをいう。本実施形態における相関モデルとは、一般には因果推論モデルと呼ばれるものである。しかし、因果推論モデルという表現は、厳密な因果の要件を求めないものであるために、正確には相関関係を推論したモデルといえる。本実施形態では、因果の要件を求めない相関関係については相関モデルと表現し、物理モデルとの近似を一つの要件と

50

する場合については、相関モデルと区別する意味で、因果推論モデルと表現する。

【0029】

図5では、 $pVar1(401)$ から、 $cVar2(402)$ に向かって矢印、すなわちエッジが延びている。これは、 $pVar1(401)$ と $cVar2(402)$ との間に因果関係が存在していることを示している。また、 $pVar1(401)$ 側が矢印の根元で、 $cVar2(402)$ 側が矢印の先端である。これは、 $pVar1(401)$ が原因で、 $cVar2(402)$ が結果であることを示している。

【0030】

また、因果推論モデルでは、図5に示すようなノードとエッジからなるグラフィカルモデルに対応する確率分布の分解式を持つ。図6に、その一例を示す。

10

【0031】

図6は、因果関係の図示の例を示す図である。図6には、変数 $x1(501)$ 、変数 $x2(502)$ 、変数 $x3(503)$ 、変数 $x4(504)$ からなる因果関係の一例が示されている。これは、確率分布の分解式の左辺、すなわち、分解される側の式は、変数 $x1(501)$ 、変数 $x2(502)$ 、変数 $x3(503)$ 、変数 $x4(504)$ の同時確率分布 $p(x1, x2, x3, x4)$ であることに対応する。また、図6では、変数 $x1(501)$ から変数 $x2(502)$ への矢印、すなわちエッジが延びている。これは、確率分布の分解式の右辺に、条件付き確率 $p(x2 | x1)$ が含まれていることに対応している。以降の例においても、矢印に対応する形で、条件付き確率が右辺に含まれている。

【0032】

20

確率分布算出部123は、相関モデルおよび物理モデルの確率分布を特定する。なお、確率分布算出部123は、相関モデル構築部122により相関モデルが既に確率分布として得られている場合には特段の処理を行わない。物理モデルに関しては、確率分布算出部123は、物理モデルの項目の次元および関係性に応じた配列を生成し、配列の要素の確率を算出する。式(1)は、物理モデルの例である。

【0033】

<式1>

$$y = f(x1, x2) = a_1 \cdot x1 + a_2 \cdot x2 + \varepsilon$$

・ ・ ・ (1)

30

【0034】

式(1)は、目的変数「 y 」と、説明変数「 $x1$ 」、「 $x2$ 」と、の間を線形関係によりモデル化した例である。一般にこのような線形関係でのモデル化の場合、誤差項として「 ε 」が配置されている。この誤差項「 ε 」は、平均が「0」の確率分布として用いられる。例えば、最小二乗法では、正規分布が用いられる。したがって、最小二乗法を用いて構築された線形関係の物理モデルを用いて確率分布の配列を生成する場合、線形式を平均値に持つ多次元正規分布を用いる。より一般には、誤差項の確率分布が未定の場合は、線形関係以外の場合でも、物理モデル式上の確率を「1」、それ以外を「0」とすることで、確率分布を得ることが可能である。

40

【0035】

物理モデル受付部124は、製造工程または検査工程に係る物理法則をモデル化した物理モデルの入力を受け付ける。物理モデルの一例としては、アレニウスの法則や、オームの法則といった数式で定義された関係がある。数式を定義する方法としては、一般的なLaTeX(ラーテック、レイテック)等により定義されている書式を用いることで、任意の数式を定義することが可能である。とくに、物理モデル受付部124は、陰関数(引数のうちの一つの変数の値が残りの変数に関係付けられる)の形式で記載された物理モデルを受け付けるのが望ましいが、これに限られるものではなく、陽関数で記載された物理モデルであってもよい。

【0036】

50

確率分布収束処理部 125 は、相関モデルの確率分布を収束計算し物理モデルの確率分布に近似させる。具体的には、確率分布収束処理部 125 は、目的変数「y」と、説明変数「x1」の値と、の組み合わせおよび当該目的変数「y」における説明変数「x2」の値と、をそれぞれランダムに 1 点選択する。なお、因果推論モデルに条件付き確率（例えば、 $p(x1|y)$ と $p(x2|y)$ 等）が含まれている場合には、確率分布収束処理部 125 は、この条件付き確率に従ったランダムな 1 点を、「x1」と「y」、「x2」と「y」の組み合わせごとに選択する（ランダムサンプリング）。

【0037】

そして、確率分布収束処理部 125 は、選択した点の確率を組み合わせごとに特定し、当該確率をランダムに摂動させる。そして、確率分布収束処理部 125 は、更新判定部 126 に更新可否を判定させ、更新する場合には因果推論モデル保存部 127 に摂動後の状態の因果推論モデルを上書き保存させる。確率分布収束処理部 125 は、このような処理を複数回（例えば、10, 000 回）実施して、相関モデルの確率分布を、物理モデルに近づけるよう収束させる。

10

【0038】

摂動のさせ方としては、確率分布収束処理部 125 は、式（2）および式（3）のようにランダムサンプルした点（それぞれ、式（2）は（x1, y）、式（3）は（x2, y））の確率を変動させる。

【0039】

<式 2>

$$\frac{p(x1|y) + \delta}{\int (p(x1|y) + \delta) dx1}$$

・・・ (2)

20

<式 3>

$$\frac{p(x2|y) + \delta}{\int (p(x2|y) + \delta) dx2}$$

・・・ (3)

30

【0040】

すなわち、上式（2）においては、分子の $p(x1|y)$ は摂動前の確率分布である。上式（2）では、この摂動前の確率分布に、変動値 δ （デルタ）を足し合わせることで、摂動させている。このとき、変動値「 δ 」はランダムに得られた値である。ただし、このランダムな変動値「 δ 」の確率分布には、正規分布や一様分布を用いる。

【0041】

上式（2）の分母は、正規化のためのスカラー倍で、分子において変動値「 δ 」を足し合わせたことにより全確率が「1」ではなくなっていることに対する補正項である。同様に、上式（3）は、 $p(x2|y)$ を摂動する式の例である。このようにして、確率分布収束処理部 125 は、選択した点の確率を摂動させる。

40

【0042】

更新判定部 126 は、摂動により更新した確率分布を採用するか否かの判定を行う。具体的には、更新判定部 126 は、摂動により更新した確率分布が、物理モデルから得られた確率分布に近づいたか否かをカルバック・ライブラー・ダイバージェンスや、L2 距離（L2 ノルム）等の確率分布間の近似を判定可能な擬似的な距離を用いて判定し、近づいたと判断した場合は、摂動により更新した確率分布を採用すると判定する。なお、これに限られるものではなく、例えば、更新判定部 126 は、摂動により更新した確率分布が、物理モデルから得られた確率分布から一定程度までであれば遠ざかった場合であっても摂

50

動により更新した確率分布を採用するようにしてもよい。

【0043】

因果推論モデル保存部127は、確率分布収束処理がなされた結果の相関モデルを、因果推論モデル記憶部113に保存する。具体的には、因果推論モデル保存部127は、所定のJSON(JavaScript Object Notation)と呼ばれるフォーマットで因果推論モデルを保存する。

【0044】

例えば、因果推論モデル保存部127は、相関モデルの各頂点、すなわち、ノードに対して、それぞれ一つずつ、図10に示すJSONのフォーマットのデータを保存する。

【0045】

図10は、因果推論モデルのJSONを用いたデータ構造例を示す図である。「Frame」要素は、確率分布を保存するために、データを離散化する際の階級値を示している。図10の例では、「-0.01から0.01」が階級「1」に属し、「0.01から0.03」が階級「2」に属することを示している。

【0046】

「UniqueName」要素は、コンピュータが変数を一意に認識するためのIDを示している。本例では、「Class」という文字列が、変数を一意に認識するためのIDとして保持されている。

【0047】

「Parents」要素は、原因となる変数を示している。図10の例では、空集合を意味する、[]となっている。これは、そのノードには原因となる変数が存在していないことを表している。

【0048】

「Prob」要素は、条件付き確率を示している。図10の例では、「0.51」と「0.49」が保持されている。これは、対象の変数が、階級「1」、すなわち、値が「-0.01から0.01」を取る確率が、「0.51」であることを示し、階級「2」、すなわち、値が「0.01から0.03」を取る確率が、「0.49」であることを示している。

【0049】

「Qty」要素は、条件付き確率を算出するために用いたデータ点数を保持している。図10の例では、「4297」と「4271」が保持されている。これは、対象の変数が、階級「1」、すなわち、値が「-0.01から0.01」を取ったデータ数が、「4297」点あり、階級「2」、すなわち、値が「0.01から0.03」を取ったデータ数が、「4271」であったことを示している。

【0050】

「Label」要素は、「UniqueName」要素で特定した変数の表示上の名前を保持している。図10の例では、「Yield」という文字列が保持されている。

【0051】

「ParentsFrame」要素は、原因となる変数を離散化する際に用いる階級値を保持している。図10の例では、「Parents」要素に示されたように原因となる変数が存在しないため、空集合を意味する[]となっている。

【0052】

図4は、ナレッジ管理装置のハードウェア構成例を示す図である。ナレッジ管理装置100は、中央処理装置(Central Processing Unit: CPU)101と、メモリ102と、ハードディスク装置(Hard Disk Drive: HDD)などの外部記憶装置103と、キーボードやマウス、バーコードリーダなどの入力装置104と、ディスプレイなどの出力装置105とを備えたコンピュータ、あるいはこのコンピュータを複数備えたコンピュータシステムで実現できる。

【0053】

例えば、制御部120の実績値受付部121と、相関モデル構築部122と、確率分布

10

20

30

40

50

算出部 1 2 3 と、物理モデル受付部 1 2 4 と、確率分布収束処理部 1 2 5 と、更新判定部 1 2 6 と、因果推論モデル保存部 1 2 7 とは、外部記憶装置 1 0 3 に記憶されている所定のプログラムをメモリ 1 0 2 にロードして CPU 1 0 1 で実行することで実現可能であり、記憶部 1 1 0 は、CPU 1 0 1 がメモリ 1 0 2 または外部記憶装置 1 0 3 を利用することにより実現可能である。

【 0 0 5 4 】

なお、これに限られず、ナレッジ管理装置 1 0 0 は、例えば A S I C (A p p l i c a t i o n S p e c i f i c I n t e g r a t e d C i r c u i t) やマイコンで実現されるものであってもよい。

【 0 0 5 5 】

図 7 は、ナレッジ生成処理のフローの例を示す図である。ナレッジ生成処理は、操作者の指示に応じて開始される。

【 0 0 5 6 】

まず、実績値受付部 1 2 1 は、製造ログデータを読み込む (ステップ S 0 0 1)。具体的には、実績値受付部 1 2 1 は、製造装置の操業状態をモニタした製造データと、製造物の検査データと、のいずれかまたは両方を含む製造ログデータの読み込みを行う。例えば、実績値受付部 1 2 1 は、製造実績記憶部 1 1 1 と、検査実績記憶部 1 1 2 と、に格納されている製造データ、検査データのいずれかまたは両方を読み込む。

【 0 0 5 7 】

そして、相関モデル構築部 1 2 2 は、相関モデルを構築する (ステップ S 0 0 2)。具体的には、相関モデル構築部 1 2 2 は、K 2 アルゴリズム等の構造学習アルゴリズムを用いて、相関性のある項目を原因と結果に関連付けて、一般に用いられる X M L の拡張データ等により保存する。

【 0 0 5 8 】

そして、物理モデル受付部 1 2 4 は、物理モデルを読み込む (ステップ S 0 0 3)。具体的には、物理モデル受付部 1 2 4 は、製造工程または検査工程に係る物理法則をモデル化した物理モデルの入力を、操作者の入力により受け付ける。物理モデルの一例としては、アレニウスの法則や、オームの法則といった一般的な電気・物理等の自然法則に関する数式である。

【 0 0 5 9 】

そして、確率分布収束処理部 1 2 5 は、確率分布の収束計処理を行う (ステップ S 0 0 4)。具体的には、確率分布収束処理部 1 2 5 は、図 8 に示す確率分布の収束計算処理を実施する。図 8 に示す確率分布の収束計算処理については、詳細は後述する。

【 0 0 6 0 】

そして、因果推論モデル保存部 1 2 7 は、因果推論モデルを保存する (ステップ S 0 0 5)。具体的には、因果推論モデル保存部 1 2 7 は、確率分布収束処理がなされた結果の相関モデルを、所定の J S O N と呼ばれるフォーマットで因果推論モデル記憶部 1 1 3 に保存する。

【 0 0 6 1 】

以上が、ナレッジ生成処理の流れである。ナレッジ生成処理によれば、製造あるいは検査に係るデータおよび物理モデルを用いて、確度の高い因果推論モデルを構築することができる。

【 0 0 6 2 】

図 8 は、確率分布の収束計算処理のフローの例を示す図である。確率分布の収束計算処理は、ナレッジ生成処理のステップ S 0 0 4 にて開始される。

【 0 0 6 3 】

まず、確率分布算出部 1 2 3 は、物理モデルを確率モデルに変換する (ステップ S 0 0 4 1)。具体的には、確率分布算出部 1 2 3 は、物理モデルに関しては、物理モデルの項目の次元および関係性に応じた配列を生成し、配列の要素ごとに確率を算出する。

【 0 0 6 4 】

10

20

30

40

50

そして、確率分布収束処理部 125 は、目的変数「 y 」をランダムサンプリングし（ステップ S0042）、説明変数「 x 」をランダムサンプリングする（ステップ S0043）。具体的には、確率分布収束処理部 125 は、目的変数と、説明変数の値と、の組み合わせをランダムに 1 点選択する。これは、すべての目的変数、説明変数の組み合わせについて条件に従ったランダムな 1 点をそれぞれ選択することで、取りうる確率をランダムに決定する目的である。

【0065】

図 9 は、ランダムサンプリングの処理の例を示す図である。図 9 は、物理モデルを式（1）に示す線形関係とした場合の例である。式（1）においては、目的変数を「 y 」で表している。図 9 では、横軸 612、622 を目的変数（ y ）、縦軸 611、621 を説明変数（ x_1 または x_2 ）としている。式（1）の例では、説明変数は「 x_1 」と「 x_2 」との二つ存在するため、図 9 においては、左側の分布図 610 が、目的変数「 y 」と説明変数「 x_1 」の確率分布を示し、右側の分布図 620 が、目的変数「 y 」と説明変数「 x_2 」の確率分布を示す。図 9 の分布図 610 および分布図 620 では、色が薄い部分は、確率の値が高く、色が濃い部分は、確率の値が低いことを意味している。

10

【0066】

図 9 において、目的変数をランダムサンプリングすることは、横軸 612、622 のうちの 1 点を選択することに相当する。そして、図 9 においては、横軸上に記載した白抜きの円 613、623 は、サンプリングした「 y 」の値（分布図間で、「 y 」の値は同値）を示している。

20

【0067】

そして、図 9 において、説明変数をランダムサンプリングすることは、縦軸 611、621 のうちの 1 点を選択することに相当する。そして、図 9 においては、塗りつぶしの円 614、624 が、それぞれサンプリングした値に該当する。

【0068】

そして、確率分布収束処理部 125 は、確率分布を誤差値（変動値）「 δ 」により摂動させる（ステップ S0044）。具体的には、確率分布収束処理部 125 は、上式（2）および上式（3）のようにランダムサンプリングした点の確率を変動させる。なお、当該摂動処理においては、所定の分布を有するランダムサンプリングに従った誤差値「 δ 」を、繰り返し処理回数、または前回の摂動による下記の擬似的な距離の変動量に応じて徐々に小さな値に収束する摂動量として用いるようにしてもよい。このようにすると、より精度を高く確率分布を収束しやすくなる。

30

【0069】

そして、更新判定部 126 は、摂動により更新した確率分布を採用するか否かの判定を行う。具体的には、更新判定部 126 は、摂動により更新した確率分布が、物理モデルから得られた確率分布に近づいたか否かをカルバック・ライブラー・ダイバージェンスや、L2 距離（L2 ノルム）等の確率分布間の近似を判定可能な擬似的な距離を用いて判定し、近づいたと判断した場合は、摂動により更新した確率分布を採用すると判定する。なお、近づいた場合だけでなく、一定程度であれば遠ざかった場合であっても採用するようにしてもよいことは、上述のとおりである。

40

【0070】

そして、確率分布収束処理部 125 は、更新した確率分布を採用する場合には、更新処理を行う（ステップ S0045）。具体的には、更新した確率分布を採用する場合には、確率分布を上書きし、採用しない場合には、摂動により更新した確率分布を破棄する。

【0071】

そして、確率分布収束処理部 125 は、終了を判定する（ステップ S0046）。具体的には、確率分布収束処理部 125 は、摂動により更新した確率分布と物理モデルから得られた確率分布に近づいた度合いが所定の閾値よりも小さい場合、すなわちカルバック・ライブラー・ダイバージェンスや、L2 距離（L2 ノルム）等の擬似的な距離の変動量が所定の閾値より小さい場合には、収束したと判定して、確率分布の収束計算処理を終了さ

50

せる。変動量が所定の閾値以上である場合には、継続するため、確率分布収束処理部 125 は、制御をステップ S 0042 に戻す。

【0072】

以上が、確率分布の収束計算処理の流れである。確率分布の収束計算処理によれば、物理モデルを組み込んだ、因果推論モデルを構築可能である。言い換えると、相関モデルの確率分布をランダムに摂動させる処理を繰り返し行うことで、徐々に物理モデルの確率分布に近似させることができるため、因果推論モデルを抽出することができる。

【0073】

図 11 は、ナレッジ生成処理の出力画面の例を示す図である。ナレッジ生成処理の出力画面の例 900 には、因果推論モデル確認ウィンドウ 901 と、物理モデル定義ウィンドウ 903 と、Build ボタン 904 と、Re - Build ボタン 905 と、が含まれる。

10

【0074】

因果推論モデル確認ウィンドウ 901 には、相関モデル構築部 122 が構築した相関モデルがグラフィカルモデルの形で表示される。

【0075】

また、因果推論モデル確認ウィンドウ 901 では、エッジ上の右クリックを受け付けると、Cut メニュー 902 が出現する。Cut メニュー 902 にクリックを受け付けると、不要なエッジが切断され、相関モデル上も関係情報が消去される。

【0076】

20

物理モデル定義ウィンドウ 903 は、物理モデルの数式の入力を受け付ける。具体的には、物理モデル定義ウィンドウ 903 は、L a T e X 等（使用されている自然言語等により地域に応じた数式表記可能な派生を含む）により定義されている書式を用いた数式の入力を受け付ける。図 11 の例では、「 $y=f(x_1, x_2)$ 」に相当する L a T e X 形式の式「 $y=f(x_1, x_2)$ 」の入力例である。なお、物理モデル定義ウィンドウ 903 は、これに限られず、予め定められた数式のスケルトンを選択入力可能とし、スケルトンに対する編集入力を受け付けるようにしてもよい。

【0077】

Build ボタン 904 は、入力を受け付けると、ナレッジ生成処理のステップ S 002 の、相関モデル構築処理を実行する。

30

【0078】

Re - Build ボタン 905 は、因果推論モデル確認ウィンドウ 901 で編集された相関モデルに、物理モデル定義ウィンドウ 903 に入力された物理モデルを反映した因果推論モデルを構築する。すなわち、ナレッジ生成処理を、ステップ S 003 の物理モデル読み込み処理から開始させる。以上が、ナレッジ生成処理の出力画面の例 900 である。

【0079】

以上が、本発明に係る実施の形態である。上記の実施の形態によれば、製造現場において、成立することが確認されている物理モデルなどのナレッジを他の製造現場から相互利用可能な形にモデル化することで、ナレッジの抽出を容易化することができる。

【0080】

40

なお、本発明は上記の実施形態に限定されるものではなく、様々な変形例が含まれる。例えば、上記した実施形態は本発明を分かりやすく説明するために詳細に説明したものであり、必ずしも説明した全ての構成を備えるものに限定されるものではない。

【0081】

また、各実施形態の構成の一部について、他の構成の追加・削除・置換をすることが可能である。

【0082】

また、上記の各構成、機能、処理部等は、それらの一部又は全部を、例えば集積回路で設計する等によりハードウェアで実現してもよい。また、上記の各構成、機能等は、プロセッサがそれぞれの機能を実現するプログラムに応じて演算を実行するソフトウェア制御

50

によって実現してもよい。各機能を実現するプログラム、テーブル、ファイル等の情報は、メモリや、ハードディスク、SSD等の記録装置、または、ICカード、SDカード、DVD等の記録媒体に格納しておき、実行時にRAM(Random Access Memory)等に読み出され、CPU等により実行することができる。

【0083】

また、制御線や情報線は説明上必要と考えられるものを示しており、製品上必ずしも全ての制御線や情報線を示しているとは限らない。実際には殆ど全ての構成が相互に接続されていると考えてもよい。

【0084】

また、上記した各構成、機能、処理部等は、それらの一部又は全部を、例えば別の装置で実行してネットワークを介して統合処理する等により分散システムで実現してもよい。

【0085】

また、上記した実施形態の技術的要素は、単独で適用されてもよいし、プログラム部品とハードウェア部品のような複数の部分に分けられて適用されるようにしてもよい。

【0086】

以上、本発明について、実施形態を中心に説明した。

【符号の説明】

【0087】

100・・・ナレッジ管理装置、110・・・記憶部、111・・・製造実績記憶部、112・・・検査実績記憶部、113・・・因果推論モデル記憶部、120・・・制御部、121・・・実績値受付部、122・・・相関モデル構築部、123・・・確率分布算出部、124・・・物理モデル受付部、125・・・確率分布収束処理部、126・・・更新判定部、127・・・因果推論モデル保存部。

【図1】

【図2】

図1

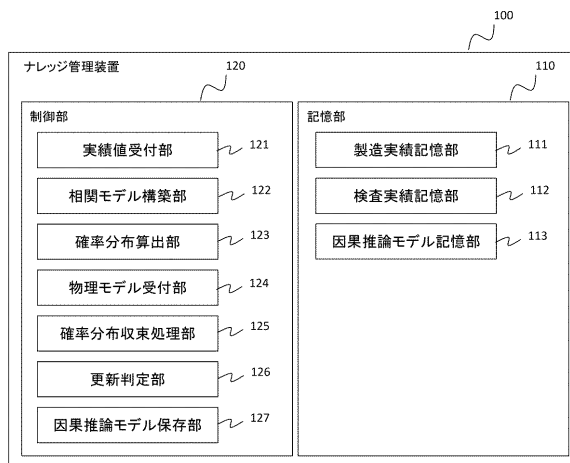


図2

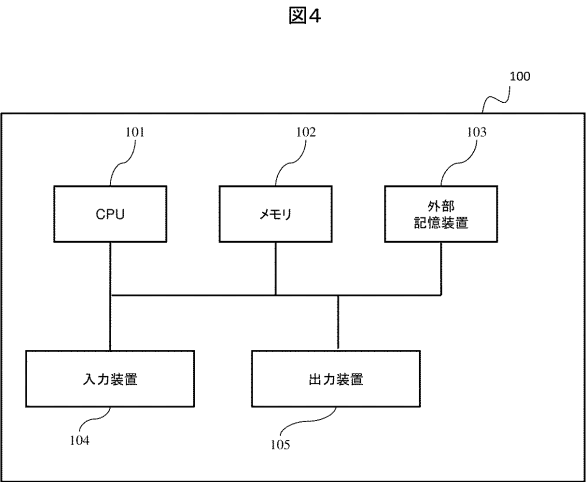
	111a	111b	111c	111d	111e	111f	111g	111h	111i	111j	111k
pVar1	pVar2	pVar3	pVar4	pVar5	pVar6	pVar7	pVar8	pVar9	pVar10		
A01	16.2	0.6	0.6	0	58.2	0.9	5.6	6.8	1		
A02	79.4	0.4	6.8	0	54.1	0.6	8.5	1.3	0.2		
A03	31.1	0.8	0.4	0.1	87	0.9	3.5	7.2	0.3		
A04	52.9	0.5	0.7	0	26.5	0.7	4.5	1.1	0.4		
A05	16.6	0.4	5.2	0.1	31.8	0.6	0.5	1.2	0.1		
A06	60.2	0.9	1	0.1	11.9	0.8	1.8	6.4	0.7		
A07	26.3	0.9	8.2	0.1	94	0.9	6.6	3.3	0.4		
A08	65.4	0.6	8.2	0	64.6	1	3.3	6.5	1		
A09	68.9	0.6	7.2	0	47.9	0	9	7.5	0.4		
A10	74.8	0.6	1.5	0	63.9	0.9	1.2	5.8	0.6		
A11	45.1	0.2	6.6	0.1	54.5	0.6	9.9	7.4	0.2		
A12	8.4	0.3	5.2	0.1	64.7	1	5.4	2.3	0.4		
B01	22.9	0.5	9.7	0	54.4	0.5	7.1	7.3	0.2		
B02	91.3	0.2	6.5	0.1	72.1	0.5	10	9.7	0.8		
B03	15.2	0.8	8	0.1	52.2	0.8	2.9	8.7	0.9		
B04	82.6	0.2	4.5	0.1	99.4	0.2	4.1	0.9	0.4		
B05	53.8	0.2	4.3	0.1	21.9	0.5	4.6	3.7	0.7		

【 図 3 】

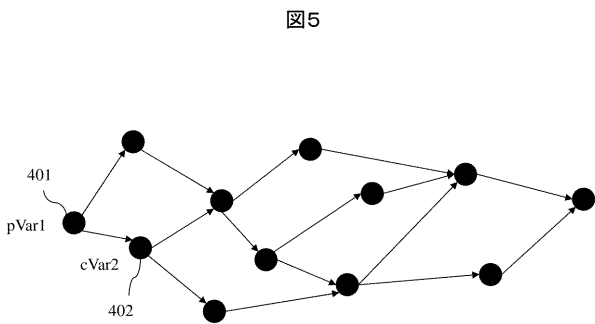
図3

	112a	112b	112c	112d	112e	112f	112g	112h	112j	112k
	cVar1	cVar2	cVar3	cVar4	cVar5	cVar6	cVar7	cVar8	cVar9	cVar10
A01		7.9	50.2	0.9	3.8	0.1	15.1	0.8	3	OK
A02		6.8	92	0.1	1.4	0.1	90.8	0.7	5.9	OK
A03		4.8	0.8	0.9	0.2	0.1	46	0.2	6.3	OK
A04		1.6	20.3	0.1	5.3	0.1	18.1	0.7	9.5	OK
A05		0.9	20.4	0.4	1.4	0.1	93.8	0.6	6.6	OK
A06		8.5	68.3	1	1.9	0.1	84.1	0.6	0.6	OK
A07		6.4	33.5	1	3.7	0.1	23.8	1	7.6	OK
A08		9.2	62.7	0	6.9	0	83.8	0.1	6.3	OK
A09		1.9	18.9	0.7	2.3	0	46	0.9	6.3	NG
A10		7.3	60.9	1	6.1	0	33.6	0.1	1	OK
A11		7.7	3.5	0.9	3.9	0.1	64.7	0.9	7.2	OK
A12		3.2	33.3	0.7	1.6	0.1	75.7	0.2	9.2	OK
B01		1.8	11.4	0.3	6.1	0	60.5	0.7	0.4	OK
B02		1.7	64	0.4	7.4	0	33.3	0.9	9.1	NG
B03		3	66	0	7.7	0	94.7	0.7	0.8	NG
B04		7.5	65.2	0.4	1.8	0	17.8	0.2	1.1	NG
B05		9.7	4.9	0.7	0.7	0.1	71.7	0	7.2	NG

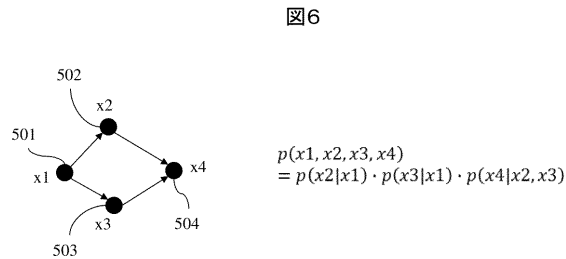
【 図 4 】



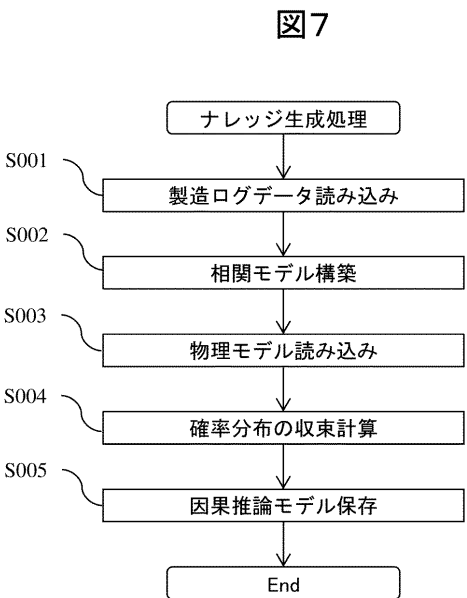
【 図 5 】



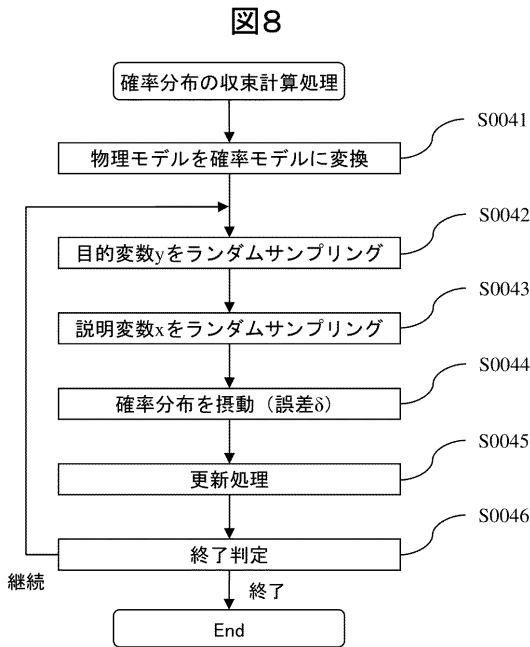
【 図 6 】



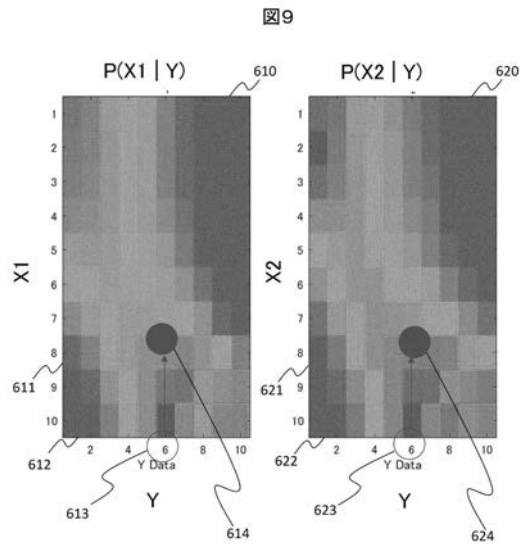
【 図 7 】



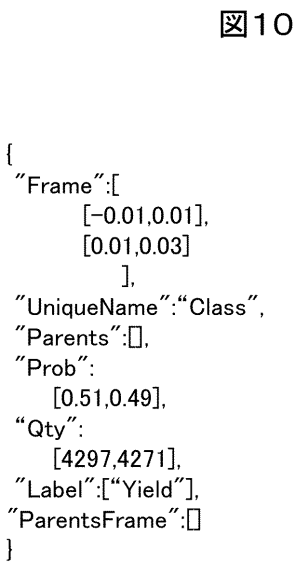
【 図 8 】



【 図 9 】



【 図 1 0 】



【 図 1 1 】

