

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号  
特許第7159552号  
(P7159552)

(45)発行日 令和4年10月25日(2022.10.25)

(24)登録日 令和4年10月17日(2022.10.17)

(51)国際特許分類	F I
G 0 6 F 16/24 (2019.01)	G 0 6 F 16/24
G 0 6 F 16/2455(2019.01)	G 0 6 F 16/2455
G 0 6 F 16/28 (2019.01)	G 0 6 F 16/28
G 0 6 F 16/903(2019.01)	G 0 6 F 16/903

請求項の数 8 (全24頁)

(21)出願番号	特願2017-242025(P2017-242025)	(73)特許権者	000005223 富士通株式会社 神奈川県川崎市中原区上小田中4丁目1番1号
(22)出願日	平成29年12月18日(2017.12.18)	(74)代理人	100079049 弁理士 中島 淳
(65)公開番号	特開2019-109692(P2019-109692 A)	(74)代理人	100084995 弁理士 加藤 和詳
(43)公開日	令和1年7月4日(2019.7.4)	(74)代理人	100099025 弁理士 福田 浩志
審査請求日	令和2年9月11日(2020.9.11)	(72)発明者	富田 匠 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内
		審査官	三橋 竜太郎

最終頁に続く

(54)【発明の名称】 データ出力プログラム、装置、及び方法

(57)【特許請求の範囲】

【請求項1】

所定の条件を満たすことが既知の既知データを受け付け、  
前記既知データとは異なる複数のデータから、前記既知データと比較した結果の類似度が所定値以上のデータ、及び前記既知データの統計的分析結果に基づき特定されるデータの少なくとも一方を、前記既知データと共に所定の用途に使用するための候補データとして抽出し、  
抽出した候補データを出力する処理をコンピュータに実行させることを特徴とするデータ出力プログラムであって、  
前記候補データとして、前記既知データの統計的分布において、出現頻度が所定の下限値以下の区間に属するデータを抽出するか、又は、  
抽出された前記候補データから、前記既知データの統計的分布において、出現頻度が所定の上限値以上の区間に属するデータの一部を除外する、  
データ出力プログラム。

10

【請求項2】

前記既知データの各々は、1以上の項目を含み、  
前記既知データと、前記項目の値として取り得る組み合わせを含む網羅データとの差分のデータ群から、前記候補データを抽出する  
請求項1に記載のデータ出力プログラム。

【請求項3】

20

前記既知データと比較した結果の類似度が所定値以上のデータ、及び前記既知データの統計的分析結果に基づき特定されるデータの少なくとも一方のデータから、前記既知データとの予め定めた相関ルールの評価値に基づいて、前記候補データを抽出する請求項 1 又は 請求項 2 に記載のデータ出力プログラム。

【請求項 4】

前記既知データは、テスト対象のシステムに入力された実データであり、

前記候補データとして、前記実データと共にテストデータとする候補のデータを抽出する請求項 1 ~ 請求項 3 のいずれか 1 項記載のデータ出力プログラム。

【請求項 5】

前記既知データは、入力データと比較を行う参照データとして記憶部に記憶されたデータであり、

前記記憶部を参照して、受け付けた前記入力データと前記参照データとを比較して、前記参照データと一致しない前記入力データから、前記候補データとして、前記参照データに追加する候補のデータを抽出する

請求項 1 ~ 請求項 3 のいずれか 1 項記載のデータ出力プログラム。

【請求項 6】

出力した前記候補データに対して、ユーザから前記候補データを前記既知データに追加するか否かの指示を受け付け、

受け付けた前記指示に基づいて、前記候補データを前記既知データに追加するか否かを判定する

請求項 1 ~ 請求項 5 のいずれか 1 項記載のデータ出力プログラム。

【請求項 7】

所定の条件を満たすことが既知の既知データを受け付け受付部と、

前記既知データとは異なる複数のデータから、前記既知データと比較した結果の類似度が所定値以上のデータ、及び前記既知データの統計的分析結果に基づき特定されるデータの少なくとも一方を、前記既知データと共に所定の用途に使用するための候補データとして抽出する抽出部と、

抽出した候補データを出力する出力部と、を含むデータ出力装置であって、

前記抽出部が、前記候補データとして、前記既知データの統計的分布において、出現頻度が所定の下限值以下の区間に属するデータを抽出するか、又は、抽出された前記候補データから、前記既知データの統計的分布において、出現頻度が所定の上限值以上の区間に属するデータの一部を除外する、

データ出力装置。

【請求項 8】

所定の条件を満たすことが既知の既知データを受け付け、

前記既知データとは異なる複数のデータから、前記既知データと比較した結果の類似度が所定値以上のデータ、及び前記既知データの統計的分析結果に基づき特定されるデータの少なくとも一方を、前記既知データと共に所定の用途に使用するための候補データとして抽出し、

抽出した候補データを出力する

処理をコンピュータが実行することを特徴とするデータ出力方法であって、

前記候補データとして、前記既知データの統計的分布において、出現頻度が所定の下限值以下の区間に属するデータを抽出するか、又は、

抽出された前記候補データから、前記既知データの統計的分布において、出現頻度が所定の上限值以上の区間に属するデータの一部を除外する、

データ出力方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、データ出力プログラム、データ出力装置、及びデータ出力方法に関する。

10

20

30

40

50

**【背景技術】****【0002】**

従来、データ群の中から、用途に応じた所定の条件を満たすデータを特定し、その用途に用いることが行われている。このようなデータとしては、例えば、システムのテストを行う場合のテストデータや、システム稼動時の異常をアナマリ検知するための参照データなどが挙げられる。

**【0003】**

上記のように、用途に応じたデータを用いる技術に関連して、様々な技術が提案されている。例えば、システムの状況を判断するコンピュータに適用される方法が提案されている。この方法は、システム中の複数の計測対象からそれぞれ計測データを受け取るステップと、各計測対象の複数の属性に対応する複数の分類に従い、各計測データと予め定められた演算アルゴリズムとに基づいて複数の異常値の集合を演算するステップとを含む。また、この方法は、複数の異常値の集合と予め定められた判断アルゴリズムとに基づいて、システムの状況を判断するステップを含む。

10

**【0004】**

また、外部ネットワーク及びLANと、情報処理装置との間に介在し、情報処理装置を不正な侵入から保護する学習型のネットワークセキュリティ装置が提案されている。この装置は、情報処理装置とLANとの間で授受される現行パケットを取り込むネットワークサービスポートの他に、学習ポートを設けている。また、この装置は、現行パケットの監視処理及び学習と並行して、過去にLANから捕捉されてパケット蓄積装置に蓄積された蓄積パケットを、学習ポートを介してネットワークセキュリティ装置に取り込んで学習する。

20

**【先行技術文献】****【特許文献】****【0005】**

【文献】国際公開2012/090718号

特開2007-096735号公報

**【発明の概要】****【発明が解決しようとする課題】****【0006】**

例えば、上記のテストデータとしては、実際にシステムで使用された実データを用いることが考えられる。また、上記のアナマリ検知のための参照データとしては、正常なデータ又は異常なデータとして既知のデータを用いることが考えられる。このような実データや既知のデータでは、テストデータや参照データとして必要なデータが漏れている可能性がある。

30

**【0007】**

一つの側面として、所定の用途に使用するデータを出力する際に、必要なデータの出力漏れを低減することを目的とする。

**【課題を解決するための手段】****【0008】**

一つの態様として、所定の条件を満たすことが既知の既知データを受け付ける。そして、前記既知データとは異なる複数のデータから、前記既知データと比較した結果の類似度が所定値以上のデータ、及び前記既知データの統計的分析結果に基づき特定されるデータの少なくとも一方を、前記既知データと共に所定の用途に使用するための候補データとして抽出し、抽出した候補データを出力する。

40

**【発明の効果】****【0009】**

一つの側面として、所定の用途に使用するデータを出力する際に、必要なデータの出力漏れを低減することができる、という効果を有する。

**【図面の簡単な説明】**

50

【 0 0 1 0 】

【 図 1 】 第 1 実施形態に係るデータ出力装置の機能ブロック図である。

【 図 2 】 実データ群に追加するデータを説明するための図である。

【 図 3 】 網羅データの作成を説明するための図である。

【 図 4 】 候補データの抽出を説明するための図である。

【 図 5 】 候補データの抽出を説明するための図である。

【 図 6 】 追加可否確認画面の一例を示す図である。

【 図 7 】 出力されるテストデータ群の一例を示す図である。

【 図 8 】 出力されるテストデータ群をツリー構造で表した図である。

【 図 9 】 第 1 実施形態に係るデータ出力装置として機能するコンピュータの概略構成を示すブロック図である。 10

【 図 1 0 】 第 1 実施形態におけるデータ出力処理の一例を示すフローチャートである。

【 図 1 1 】 第 2 実施形態に係るデータ出力装置の機能ブロック図である。

【 図 1 2 】 出力するアノマリ検知データを説明するための図である。

【 図 1 3 】 第 2 実施形態に係るデータ出力装置として機能するコンピュータの概略構成を示すブロック図である。

【 図 1 4 】 第 2 実施形態におけるデータ出力処理の一例を示すフローチャートである。

【 図 1 5 】 第 3 実施形態に係るデータ出力装置の機能ブロック図である。

【 図 1 6 】 数値の項目の値を区間毎に区切る統計的手法の一例を説明するための図である。

【 図 1 7 】 実データの統計的分析結果の一例を示す図である。 20

【 図 1 8 】 第 3 実施形態に係るデータ出力装置として機能するコンピュータの概略構成を示すブロック図である。

【 図 1 9 】 第 3 実施形態におけるデータ出力処理の一例を示すフローチャートである。

【 図 2 0 】 第 4 実施形態に係るデータ出力装置の機能ブロック図である。

【 図 2 1 】 相関ルールの評価値の算出を説明するための図である。

【 図 2 2 】 第 4 実施形態に係るデータ出力装置として機能するコンピュータの概略構成を示すブロック図である。

【 図 2 3 】 第 4 実施形態におけるデータ出力処理の一例を示すフローチャートである。

【 発明を実施するための形態 】

【 0 0 1 1 】

以下、図面を参照して、本発明の実施形態の一例を詳細に説明する。 30

【 0 0 1 2 】

< 第 1 実施形態 >

第 1 実施形態では、システム改修後のテストに用いるテストデータ群を出力するデータ出力装置について説明する。

【 0 0 1 3 】

図 1 に示すように、第 1 実施形態に係るデータ出力装置 1 0 には、システム改修前に実際にシステムに入力され実データ群が入力される。また、データ出力装置 1 0 からは、実データ群に、実データ群に含まれるデータ以外のデータが追加されたテストデータ群が出力される。なお、実データ群は、本発明の既知データの一例である。 40

【 0 0 1 4 】

ここで、図 2 を参照して、実データ群に追加するデータの考え方について説明する。

【 0 0 1 5 】

システムのテストを漏れなく行うためには、データの各項目が取り得る値の全ての組み合わせからなる網羅データを用いることが理想である。しかし、網羅データは非常に膨大であるため、網羅データをテストデータとして用いることは、効率面で現実的ではない。

【 0 0 1 6 】

一方、実データをテストデータとすることを考える。この場合、改修後のシステムにおいて、改修前のシステムに入力された実データをテストデータとして用いたテストを行うだけでは、十分でない場合がありえる。 50

## 【 0 0 1 7 】

そこで、本実施形態では、網羅データと実データ群との差分のデータ群から、実データを基準として、テストデータに追加する候補データを抽出し、抽出した候補データからユーザにより選択された候補データを、実データに加えてテストデータ群として出力する。

## 【 0 0 1 8 】

データ出力装置 1 0 は、機能的には、図 1 に示すように、受付部 1 1 と、作成部 1 2 と、抽出部 1 3 と、提示部 1 4 と、出力部 1 5 とを含む。なお、作成部 1 2 及び抽出部 1 3 は、本発明の抽出部の一例である。

## 【 0 0 1 9 】

受付部 1 1 は、データ出力装置 1 0 に入力された実データ群を受け付け、作成部 1 2 及び出力部 1 5 へ受け渡す。

10

## 【 0 0 2 0 】

作成部 1 2 は、受付部 1 1 から受け渡された実データ群に基づいて、網羅データを作成する。例えば、作成部 1 2 は、図 3 の左図に示すような実データ群から、図 3 の右図に示すような網羅データを作成する。なお、図 3 の例では、各データは、「Type」、「ID」、及び「Name」の 3 つの項目を持つ例を示している。また、項目「Type」の値には、「Buy」、「Sell」、及び「Other」の 3 種類が存在し、項目「ID」の値には、「0 1」及び「0 2」の 2 種類が存在し、項目「Name」の値には、「A」及び「B」の 2 種類が存在する。したがって、網羅データには、各項目の取り得る値の全ての組み合わせとして、 $3 \times 2 \times 2 = 12$  個のデータが含まれる。図 3 では、後段の説明のため、網羅データに含まれる各データに番号を付与している。

20

## 【 0 0 2 1 】

抽出部 1 3 は、網羅データと実データ群との差分のデータ群から、実データと類似するデータを候補データとして抽出する。具体的には、抽出部 1 3 は、差分のデータ群に含まれるデータの各々について、そのデータと類似するデータが実データ群に網羅的に存在する場合に、そのデータを候補データとして抽出する。より具体的には、抽出部 1 3 は、データの M 個の項目をワイルドカード (\*) とした探索用データを、ワイルドカードとする項目を異ならせた全てのパターンについて作成する。そして、抽出部 1 3 は、探索用データの全てが、実データ群に含まれる実データのいずれかと一致する場合に、対象のデータと実データとが類似していると判断する。

30

## 【 0 0 2 2 】

例えば、図 3 の右図に示す網羅データのうち、差分のデータ群 ( 4、5、6、9、10、12 番のデータ ) に含まれるデータのうち、4 番の < Buy, 0 2, B > について考える。ここでは、M = 1 の探索用データを作成する場合について説明する。図 4 に示すように、< Buy, 0 2, B > からは、< \*, 0 2, B >、< Buy, \*, B >、及び < Buy, 0 2, \* > が探索用データとして作成される。この場合、いずれの探索用データも、一致する実データが実データ群に存在するため、< Buy, 0 2, B > は、候補データとして抽出される。

## 【 0 0 2 3 】

一方、図 5 に示すように、5 番の < Sell, 0 1, A > からは、< \*, 0 1, A >、< Sell, \*, A >、及び < Sell, 0 1, \* > が探索用データとして作成される。この場合、探索用データ < Sell, 0 1, \* > に一致する実データが実データ群に存在しないため、< Sell, 0 1, A > は、候補データとして抽出されない。

40

## 【 0 0 2 4 】

なお、本実施形態では、あるデータから作成される全ての探索用データが実データと一致する場合に、そのデータを候補データとして抽出する場合について説明するが、これに限定されない。例えば、あるデータから作成された L 個の探索用データのうちの N 個の探索用データが実データと一致する場合に、そのデータを候補データとして抽出してもよい。N は、例えば、 $L \times 0.8$  個や、 $L - 1$  個などとすることができる。

## 【 0 0 2 5 】

50

抽出部 1 3 は、抽出した候補データを提示部 1 4 へ受け渡す。

【 0 0 2 6 】

提示部 1 4 は、抽出部 1 3 から受け渡された候補データを表示装置に表示するなどしてユーザに提示し、その候補データをテストデータに追加するか否かを示す追加可否情報を受け付ける。例えば、提示部 1 4 は、図 6 に示すような追加可否確認画面 2 1 を表示装置に表示する。図 6 の例では、追加可否確認画面 2 1 には、抽出された候補データが表示される候補データ領域 2 2、候補データを選択するためのチェックボックス 2 3、及び追加可否情報を確定する際に選択される確定ボタン 2 4 が含まれる。確定ボタン 2 4 が選択されると、提示部 1 4 は、追加可否情報を受け付ける。追加可否情報は、チェックされたチェックボックス 2 3 に対応する候補データをテストデータに追加し、チェックされていないチェックボックス 2 3 に対応する候補データはテストデータに追加しないことを示す情報である。

10

【 0 0 2 7 】

提示部 1 4 は、追加可否情報に基づいて、追加する候補データを特定し、出力部 1 5 へ受け渡す。

【 0 0 2 8 】

出力部 1 5 は、受付部 1 1 から受け渡された実データ群と、提示部 1 4 から受け渡された候補データとをあわせて、テストデータ群として出力する。図 7 に、出力されるテストデータ群の一例を示す。また、図 8 に、出力されるテストデータ群を、項目毎に階層化し、各項目の値をノードとし、各データの値に対応するノードをエッジで接続したツリー構造で表した図を示す。図 8 では、元の実データ群を実線のノード及びエッジで表し、追加されたデータを点線で表している。また、追加されたデータに類似する元の実データに対応するノードを網掛けのノードで表している。図 7 及び図 8 に示すように、元の実データを基準として、実データに類似するデータが追加されたテストデータ群が出力される。

20

【 0 0 2 9 】

データ出力装置 1 0 は、例えば図 9 に示すコンピュータ 4 0 で実現することができる。コンピュータ 4 0 は、CPU (Central Processing Unit) 4 1 と、一時記憶領域としてのメモリ 4 2 と、不揮発性の記憶部 4 3 とを備える。また、コンピュータ 4 0 は、入力装置、表示装置等の入出力装置 4 4 と、記憶媒体 4 9 に対するデータの読み込み及び書き込みを制御する R / W (Read/Write) 部 4 5 と、インターネット等のネットワークに接続される通信 I / F (Interface) 4 6 とを備える。CPU 4 1、メモリ 4 2、記憶部 4 3、入出力装置 4 4、R / W 部 4 5、及び通信 I / F 4 6 は、バス 4 7 を介して互いに接続される。

30

【 0 0 3 0 】

記憶部 4 3 は、HDD (Hard Disk Drive)、SSD (Solid State Drive)、フラッシュメモリ等によって実現できる。記憶媒体としての記憶部 4 3 には、コンピュータ 4 0 を、データ出力装置 1 0 として機能させるためのデータ出力プログラム 5 0 が記憶される。データ出力プログラム 5 0 は、受付プロセス 5 1 と、作成プロセス 5 2 と、抽出プロセス 5 3 と、提示プロセス 5 4 と、出力プロセス 5 5 とを有する。

【 0 0 3 1 】

CPU 4 1 は、データ出力プログラム 5 0 を記憶部 4 3 から読み出してメモリ 4 2 に展開し、データ出力プログラム 5 0 が有するプロセスを順次実行する。CPU 4 1 は、受付プロセス 5 1 を実行することで、図 1 に示す受付部 1 1 として動作する。また、CPU 4 1 は、作成プロセス 5 2 を実行することで、図 1 に示す作成部 1 2 として動作する。また、CPU 4 1 は、抽出プロセス 5 3 を実行することで、図 1 に示す抽出部 1 3 として動作する。また、CPU 4 1 は、提示プロセス 5 4 を実行することで、図 1 に示す提示部 1 4 として動作する。また、CPU 4 1 は、出力プロセス 5 5 を実行することで、図 1 に示す出力部 1 5 として動作する。これにより、データ出力プログラム 5 0 を実行したコンピュータ 4 0 が、データ出力装置 1 0 として機能することになる。なお、プログラムを実行する CPU 4 1 はハードウェアである。

40

50

## 【0032】

なお、データ出力プログラム50により実現される機能は、例えば半導体集積回路、より詳しくはASIC (Application Specific Integrated Circuit) 等で実現することも可能である。

## 【0033】

次に、第1実施形態に係るデータ出力装置10の作用について説明する。

## 【0034】

データ出力装置10に実データ群が入力され、テストデータ群の出力が指示されると、データ出力装置10において、図10に示すデータ出力処理が実行される。なお、データ出力処理は、本発明のデータ出力方法の一例である。

10

## 【0035】

ステップS11で、受付部11が、データ出力装置10に入力された実データ群を受け付け、作成部12及び出力部15へ受け渡す。

## 【0036】

次に、ステップS12で、作成部12が、受付部11から受け渡された実データ群に基づいて、網羅データを作成する。

## 【0037】

次に、ステップS13で、抽出部13が、網羅データと実データ群との差分のデータ群から、未選択のデータを1つ選択する。

## 【0038】

次に、ステップS14で、抽出部13が、選択したデータに類似する実データが実データ群に存在するか否かを判定する。具体的には、抽出部13は、選択したデータのM個の項目をワイルドカード(\*)とした探索用データを、ワイルドカードとする項目を異ならせた全てのパターンについて作成する。そして、抽出部13は、探索用データの全てが、実データ群に含まれる実データのいずれかと一致する場合に、選択したデータと実データとが類似していると判断する。選択したデータに類似する実データが存在する場合には、処理はステップS15へ移行し、存在しない場合には、処理はステップS16へ移行する。

20

## 【0039】

ステップS15では、抽出部13が、上記ステップS13で選択したデータを候補データとして抽出する。

30

## 【0040】

次に、ステップS16で、抽出部13が、網羅データと実データ群との差分のデータ群に未選択のデータが存在するか否かを判定する。未選択のデータが存在する場合には、処理はステップS13に戻り、全てのデータを選択済みの場合には、処理はステップS17へ移行する。

## 【0041】

ステップS17では、抽出部13が、上記ステップS15で抽出した候補データを提示部14へ受け渡す。そして、提示部14が、抽出部13から受け渡された候補データを表示装置に表示するなどしてユーザに提示する。

## 【0042】

次に、ステップS18で、提示部14が、上記ステップS17で提示した候補データをテストデータに追加するか否かを示す追加可否情報を受け付ける。提示部14は、追加可否情報に基づいて、追加する候補データを特定し、出力部15へ受け渡す。

40

## 【0043】

次に、ステップS19で、出力部15が、上記ステップS11で受付部11から受け渡された実データ群と、上記ステップS18で提示部14から受け渡された候補データとをあわせて、テストデータ群として出力する。そして、データ出力処理は終了する。

## 【0044】

以上説明したように、第1実施形態におけるデータ出力装置10によれば、網羅データと実データ群との差分のデータ群から、実データに類似する候補データを抽出する。そし

50

て、抽出した候補データからユーザにより選択された候補データと、実データとを合わせてテストデータ群として出力する。これにより、システム改修後のテストに用いるテストデータ群を出力する際に、必要なテストデータの出力漏れを低減することができる。

【0045】

<第2実施形態>

次に、第2実施形態について説明する。なお、第2実施形態に係るデータ出力装置において、第1実施形態に係るデータ出力装置10と同様の部分については、同一符号を付して、詳細な説明を省略する。

【0046】

第2実施形態では、システムに入力されるデータの異常を検知するためのアノマリ検知データを出力するデータ出力装置について説明する。

10

【0047】

図11に示すように、第2実施形態に係るデータ出力装置210には、入力データが入力される。また、データ出力装置210は、入力データの異常パターンとして既に判明しているデータがアノマリ検知データとして記憶されたアノマリデータベース(DB)17と接続されている。そして、データ出力装置210からは、アノマリDB17に未登録のアノマリ検知データが出力される。なお、アノマリDB17に記憶されたアノマリ検知データは、本発明の既知データの一例である。

【0048】

ここで、図12を参照して、第2実施形態で出力するアノマリ検知データの考え方について説明する。

20

【0049】

既に判明しているアノマリ検知データと、入力データとを比較することにより、入力データが異常であるか否かを判定することができるが、未知の異常パターンが存在する可能性を考慮すると、既知のアノマリ検知データだけで十分とは言えない。

【0050】

そこで、本実施形態では、既知のアノマリ検知データと一致しない入力データが、既知のアノマリ検知データと類似する場合に、その入力データを候補データとして抽出する。そして、抽出した候補データからユーザにより選択された候補データを、アノマリDB17を更新するための新たなアノマリ検知データとして出力する。

30

【0051】

データ出力装置210は、機能的には、図11に示すように、受付部211と、判定部16と、抽出部213と、提示部14と、出力部215とを含む。なお、判定部16及び抽出部213は、本発明の抽出部の一例である。

【0052】

受付部211は、データ出力装置210に入力された入力データを受け付け、判定部16へ受け渡す。

【0053】

判定部16は、受付部211から受け渡された入力データが、アノマリDB17に記憶されたいずれのかのアノマリ検知データと一致するか否かに基づいて、入力データが異常か否かを判定する。判定部16は、入力データがいずれのかのアノマリ検知データと一致する場合には、入力データが異常であると判定して、表示装置へアラートメッセージを表示するなどして、ユーザへ異常を通知する。また、判定部16は、入力データがいずれのアノマリ検知データにも一致しない場合には、入力データは正常であると判定して、入力データを抽出部213へ受け渡す。

40

【0054】

抽出部213は、既知のアノマリ検知データと比較した結果の類似度が所定値以上の入力データを候補データとして抽出する。入力データが既知のアノマリ検知データに類似するか否かは、第1実施形態における抽出部13と同様に、入力データのM個の項目をワイルドカード(\*)とした探索用データが、アノマリDB17に網羅的に存在するか否かに

50

より判定する。

【0055】

出力部215は、提示部14から受け渡された候補データを、新たなアノマリ検知データとして出力し、アノマリDB17に追加する。

【0056】

データ出力装置210は、例えば図13に示すコンピュータ40で実現することができる。コンピュータ40の記憶部43には、コンピュータ40を、データ出力装置210として機能させるためのデータ出力プログラム250が記憶される。データ出力プログラム250は、受付プロセス251と、判定プロセス56と、抽出プロセス253と、提示プロセス54と、出力プロセス255とを有する。

10

【0057】

CPU41は、データ出力プログラム250を記憶部43から読み出してメモリ42に展開し、データ出力プログラム250が有するプロセスを順次実行する。CPU41は、受付プロセス251を実行することで、図11に示す受付部211として動作する。また、CPU41は、判定プロセス56を実行することで、図11に示す判定部16として動作する。また、CPU41は、抽出プロセス253を実行することで、図11に示す抽出部213として動作する。また、CPU41は、提示プロセス54を実行することで、図11に示す提示部14として動作する。また、CPU41は、出力プロセス255を実行することで、図11に示す出力部215として動作する。これにより、データ出力プログラム250を実行したコンピュータ40が、データ出力装置210として機能することになる。

20

【0058】

なお、データ出力プログラム250により実現される機能は、例えば半導体集積回路、より詳しくはASIC等で実現することも可能である。

【0059】

次に、第2実施形態に係るデータ出力装置210の作用について説明する。

【0060】

システム稼動時に、データ出力装置210に入力データが入力されると、データ出力装置210において、図14に示すデータ出力処理が実行される。

【0061】

ステップS21で、受付部211が、データ出力装置210に入力された入力データを受け付け、判定部16へ受け渡す。

30

【0062】

次に、ステップS22で、判定部16が、受付部211から受け渡された入力データが、アノマリDB17に記憶されたいずれのかのアノマリ検知データと一致するか否かに基づいて、入力データが異常か否かを判定する。入力データがいずれのアノマリ検知データと一致する場合には、処理はステップS23へ移行し、入力データがいずれのアノマリ検知データにも一致しない場合には、処理はステップS24へ移行する。

【0063】

ステップS23では、判定部16が、表示装置へアラートメッセージを表示するなどして、ユーザへ異常を通知する。

40

【0064】

ステップS24では、判定部16が、入力データは正常であると判定して、入力データを抽出部213へ受け渡す。そして、抽出部213が、判定部16から受け渡された入力データが既知のアノマリ検知データに類似するか否かを判定する。類似する場合には、処理はステップS15へ移行し、類似しない場合には、データ出力処理を終了する。

【0065】

ステップS15～S17で、第1実施形態におけるデータ出力処理と同様に、入力データを候補データとして抽出し、抽出した候補データをユーザに提示し、追加可否情報を受け付ける。提示部14は、追加可否情報に基づいて、候補データをアノマリDB17に追

50

加する場合、追加指示と共に候補データを出力部 2 1 5 へ受け渡す。

【 0 0 6 6 】

次に、ステップ S 2 9 で、出力部 2 1 5 が、提示部 1 4 から受け渡された候補データを、新たなアノマリ検知データとして出力し、アノマリ D B 1 7 に追加する。そして、データ出力処理は終了する。

【 0 0 6 7 】

以上説明したように、第 2 実施形態におけるデータ出力装置 2 1 0 によれば、既知のアノマリ検知データと一致しなかった入力データが既知のアノマリ検知データに類似する場合、その入力データを候補データとして抽出する。そして、抽出した候補データからユーザにより選択された候補データを出力し、既知のアノマリ検知データに追加する。これにより、入力データの異常を検知するためのアノマリ検知データを出力する際に、今後入力される入力データに未知の異常パターンが含まれる場合を考慮して、必要なアノマリ検知データの出力漏れを低減することができる。

10

【 0 0 6 8 】

< 第 3 実施形態 >

次に、第 3 実施形態について説明する。なお、第 3 実施形態に係るデータ出力装置において、第 1 実施形態に係るデータ出力装置 1 0 と同様の部分については、同一符号を付して、詳細な説明を省略する。

【 0 0 6 9 】

第 3 実施形態では、第 1 実施形態と同様に、システム改修後のテストに用いるテストデータ群を出力するデータ出力装置について説明する。

20

【 0 0 7 0 】

図 1 5 に示すように、データ出力装置 3 1 0 は、機能的には、受付部 1 1 と、抽出部 3 1 3 と、提示部 1 4 と、出力部 1 5 とを含む。

【 0 0 7 1 】

抽出部 3 1 3 は、実データ群を統計的に分析した結果から、実データ群に加えてテストデータとする候補データを抽出する。具体的には、抽出部 3 1 3 は、実データにおいて数値の項目の値の統計的分布において、出現頻度が低い値のデータを補完するように、候補データを抽出する。

【 0 0 7 2 】

より具体的には、抽出部 3 1 3 は、実データにおける数値の項目の値を、統計的手法により、所定の区間に区切る。統計的手法としては、例えば、等差や等量を用いることができる。等差による区間は、図 1 6 の上図に示すように、数値の項目の値に対して、昇順又は降順に付与したランキングが、1 区間に一定数ずつ含まれるように区切られる。また、等量による区間は、図 1 6 の下図に示すように、数値の項目の最小値から最大値までが、値の大きさが等しくなるように区切られる。

30

【 0 0 7 3 】

例えば、抽出部 3 1 3 は、図 1 7 に示すように、設定ファイルを用いて、実データから統計的分析結果を示す中間ファイルを得る。図 1 7 の例では、設定ファイルにおいて、数値の項目を指定する項目番号（図 1 7 の例では「 2 」）と、等差及び等量の指示（図 1 7 の例では「 r a n k 」及び「 e q u i v a l e n c e 」）とが指示される。また、設定ファイルでは、全体を何区間に区切るかを示すパラメータ（図 1 7 の例では「 3 」及び「 4 」）も指示される。また、中間ファイルには、少なくとも、各区間に属する値の個数の情報が含まれる。抽出部 3 1 3 は、この各区間の値の個数が、所定の下限値以下の区間に属する値を持つデータを候補データとして作成する。該当の区間に属する値は、ランダムに決定してもよいし、予め定めたルール（例えば、その区間の最大値、最小値、中央値等）で決定してもよい。また、作成する候補データの数値以外の項目の値は、実データに存在する値から、ランダムに選択したり、出現頻度が最も高い値を選択したりすればよい。

40

【 0 0 7 4 】

例えば、図 1 7 に示す中間ファイルにおいて、出現頻度が 0 の区間（図 1 7 中の破線部

50

)から候補データを作成する場合、 $\langle 1000, A, 30000, 100 \rangle$ や $\langle 1000, A, 60000, 100 \rangle$ のような候補データを作成することができる。なお、作成する候補データの数は1個に限定されず、他の区間に属する値の個数の平均や最小値などを基準に適宜設定すればよい。

【0075】

データ出力装置310は、例えば図18に示すコンピュータ40で実現することができる。コンピュータ40の記憶部43には、コンピュータ40を、データ出力装置310として機能させるためのデータ出力プログラム350が記憶される。データ出力プログラム350は、受付プロセス51と、抽出プロセス353と、提示プロセス54と、出力プロセス55とを有する。

10

【0076】

CPU41は、データ出力プログラム350を記憶部43から読み出してメモリ42に展開し、データ出力プログラム350が有するプロセスを順次実行する。CPU41は、抽出プロセス353を実行することで、図15に示す抽出部313として動作する。他のプロセスについては、第1実施形態におけるデータ出力プログラム50と同様である。これにより、データ出力プログラム350を実行したコンピュータ40が、データ出力装置310として機能することになる。

【0077】

なお、データ出力プログラム350により実現される機能は、例えば半導体集積回路、より詳しくはASIC等で実現することも可能である。

20

【0078】

次に、第3実施形態に係るデータ出力装置310の作用について説明する。

【0079】

データ出力装置310に実データ群が入力され、テストデータ群の出力が指示されると、データ出力装置310において、図19に示すデータ出力処理が実行される。なお、データ出力処理は、本発明のデータ出力方法の一例である。

【0080】

ステップS11で、受付部11が、実データ群を受け付け、抽出部313へ受け渡す。

【0081】

次に、ステップS35で、抽出部313が、実データにおける数値の項目の値を、統計的手法により、所定の区間に区切る。そして、抽出部313は、各区間に属する値の個数が、所定の下限值以下の区間に属する値を持つデータを候補データとして作成する。

30

【0082】

以下、ステップS17~S19で、第1実施形態におけるデータ出力処理と同様に処理し、データ出力処理は終了する。

【0083】

以上説明したように、第3実施形態におけるデータ出力装置310によれば、実データの統計的分布において、出現頻度が低い値のデータを補完するように、候補データを抽出する。これにより、システム改修後のテストに用いるテストデータ群を出力する際に、必要なテストデータの出力漏れを低減することができる。

40

【0084】

なお、上記第3実施形態では、数値の項目の値の分布を区切る統計的手法として、等差又は等量を用いる場合について説明したが、これに限定されない。例えば、数値の項目の値の分布を正規分布とみなし、偏差値が $\pm 10$ (68.3%)、 $\pm 20$ (95.4%)、 $\pm 30$ (99.73%)、 $\pm 40$ (99.9937%)、及びそれ以上の区間に区切るなどしてもよい。

【0085】

また、上記第3実施形態において、数値の項目の値の統計的分布において、出現頻度が高い値のデータをテストデータから除外するようにしてもよい。これにより、効率的にテストを行うことができるテストデータを出力することができる。

50

【 0 0 8 6 】

< 第 4 実施形態 >

次に、第 4 実施形態について説明する。なお、第 4 実施形態に係るデータ出力装置において、第 1 実施形態に係るデータ出力装置 1 0 と同様の部分については、同一符号を付して、詳細な説明を省略する。

【 0 0 8 7 】

第 4 実施形態では、第 1 実施形態と同様に、システム改修後のテストに用いるテストデータ群を出力するデータ出力装置について説明する。

【 0 0 8 8 】

図 2 0 に示すように、データ出力装置 4 1 0 は、機能的には、受付部 1 1 と、作成部 1 2 と、抽出部 4 1 3 と、評価部 1 8 と、提示部 1 4 と、出力部 1 5 とを含む。なお、作成部 1 2、抽出部 4 1 3、及び評価部 1 8 は、本発明の抽出部の一例である。

10

【 0 0 8 9 】

抽出部 4 1 3 は、第 1 実施形態における抽出部 1 3 と同様に、網羅データと実データ群との差分のデータ群から、実データと類似するデータを候補データとして抽出する。この際、抽出部 4 1 3 は、実データ群を統計的に分析した結果に基づいて、実データを補完した上で、候補データを抽出する。

【 0 0 9 0 】

具体的には、抽出部 4 1 3 は、実データの数値の項目の値の統計的分布において、出現頻度が低い値のデータを実データに追加する。出現頻度が低い値のデータの抽出方法は、第 3 実施形態における抽出部 3 1 3 と同様に、実データにおける数値の項目の値を、統計的手法により、所定の区間に区切り、各区間の値の個数が、所定の下限値以下の区間に属する値を持つデータを抽出すればよい。また、実データに追加する際には、数値以外の項目については、ワイルドカード ( \* ) を設定すればよい。実データが補完されることにより、実データに類似する候補データを抽出する際に、抽出される候補データの範囲を拡大することができる。

20

【 0 0 9 1 】

また、抽出部 4 1 3 は、抽出した候補データから、実データ群を統計的に分析した結果に基づいて、所定の候補データを除外する。具体的には、抽出部 4 1 3 は、実データの数値の項目の値の統計的分布において、出現頻度が高い値の候補データを間引く。より具体的には、抽出部 4 1 3 は、実データにおける数値の項目の値を、統計的手法により、所定の区間に区切り、各区間の値の個数が、所定の上限値以上の区間に属する値を持つ候補データから、所定数の候補データを選択して除外する。所定数は、例えば、除外後のその区間の候補データが上述の上限値となる数とすることができる。また、除外する候補データは、ランダムに選択したり、値の大きさ順に所定個間隔で選択したりすればよい。

30

【 0 0 9 2 】

評価部 1 8 は、抽出部 4 1 3 により抽出された候補データの各々について、実データとの予め定めた相関ルールの評価値を算出する。本実施形態における相関ルールとは、既知の実データがテストデータとして存在する場合に、候補データが高い確率でテストデータとして必要なデータである、というものである。相関ルールの評価値として、本実施形態では、実データ群を、候補データの項目のうち、実データと一致した項目を含むグループ ( L H S ) と、不一致の項目を含むグループ ( R H S ) とした場合の信頼度及びリフト値を用いる。信頼度は、候補データが相関ルールにどの程度適応しているかを表す指標であり、リフト値は、R H S が L H S においてどれだけ特殊であることを示す指標である。いずれの指標も、高いほど、その候補データが相関ルールを満たす候補データである可能性が高いことを示す。以下に、信頼度及びリフト値の一例を示す。

40

【 0 0 9 3 】

信頼度

$$= \text{L H S 及び R H S の両方の条件を満たす実データ数} / \text{L H S を満たす実データ数}$$

$$= ( \text{L H S } \cap \text{R H S } ) / \text{A L L}$$

50

リフト値

$$= \text{LHS 全体における RHS の出現確率} / \text{全実データにおける RHS の出現確率}$$

$$= ( ( \text{LHS} \quad \text{RHS} ) / \text{LHS} ) / ( ( \text{ALL} \quad \text{RHS} ) / \text{ALL} )$$

( ALL は実データの全件数 )

【 0 0 9 4 】

例えば、図 2 1 の上左の図に示す実データ群に対して、上右の図に示す候補データ < Buy , 0 2 , B > の評価値を算出した例を、図 2 1 の下図に示す。図 2 1 は、候補データ < Buy , 0 2 , B > の項目 M 個 ( ここでは、M = 1 ) 違いの実データとして、< Buy , 0 1 , B >、< Buy , 0 2 , A >、及び < Sell , 0 2 , B > の各々について、信頼度及びリフト値を算出した例である。

10

【 0 0 9 5 】

評価部 1 8 は、候補データ毎に算出した評価値が、所定の閾値未満の候補データを、候補データから除外する。例えば、図 2 1 の例で、信頼度の閾値を 0 . 5、リフト値の閾値を 1 . 0 とすると、いずれの信頼度及びリフト値も閾値を超えているため、候補データ < Buy , 0 2 , B > は、除外されることなく、候補データとして採用される。

【 0 0 9 6 】

なお、いずれの信頼度及びリフト値も閾値を超えている場合に、候補データとして採用する場合に限らず、所定割合の信頼度及びリフト値が閾値を越えている場合 ( 例えば、図 2 1 の例では、3 件中 2 件など ) に、候補データとして採用するようにしてもよい。このような、候補データとして採用するための条件や、閾値は、候補データをどの程度の数抽出するかに応じて、適宜設定しておけばよい。

20

【 0 0 9 7 】

データ出力装置 4 1 0 は、例えば図 2 2 に示すコンピュータ 4 0 で実現することができる。コンピュータ 4 0 の記憶部 4 3 には、コンピュータ 4 0 を、データ出力装置 4 1 0 として機能させるためのデータ出力プログラム 4 5 0 が記憶される。データ出力プログラム 4 5 0 は、受付プロセス 5 1 と、作成プロセス 5 2 と、抽出プロセス 4 5 3 と、評価プロセス 5 8 と、提示プロセス 5 4 と、出力プロセス 5 5 とを有する。

【 0 0 9 8 】

CPU 4 1 は、データ出力プログラム 4 5 0 を記憶部 4 3 から読み出してメモリ 4 2 に展開し、データ出力プログラム 4 5 0 が有するプロセスを順次実行する。CPU 4 1 は、抽出プロセス 4 5 3 を実行することで、図 2 0 に示す抽出部 4 1 3 として動作する。また、CPU 4 1 は、評価プロセス 5 8 を実行することで、図 2 0 に示す評価部 1 8 として動作する。他のプロセスについては、第 1 実施形態におけるデータ出力プログラム 5 0 と同様である。これにより、データ出力プログラム 4 5 0 を実行したコンピュータ 4 0 が、データ出力装置 4 1 0 として機能することになる。

30

【 0 0 9 9 】

なお、データ出力プログラム 4 5 0 により実現される機能は、例えば半導体集積回路、より詳しくは ASIC 等で実現することも可能である。

【 0 1 0 0 】

次に、第 4 実施形態に係るデータ出力装置 4 1 0 の作用について説明する。

40

【 0 1 0 1 】

データ出力装置 4 1 0 に実データ群が入力され、テストデータ群の出力が指示されると、データ出力装置 4 1 0 において、図 2 3 に示すデータ出力処理が実行される。なお、データ出力処理は、本発明のデータ出力方法の一例である。

【 0 1 0 2 】

ステップ S 1 1 で、受付部 1 1 が、実データ群を受け付け、作成部 1 2 へ受け渡す。

【 0 1 0 3 】

次に、ステップ S 1 2 で、作成部 1 2 が、実データ群に基づいて、網羅データを作成する。

【 0 1 0 4 】

50

次に、ステップ S 4 1 で、抽出部 4 1 3 が、実データにおける数値の項目の値を、統計的手法により、所定の区間に区切り、各区間の値の個数が、所定の下限値以下の区間に属する値を持つデータで、実データを補完する。

【 0 1 0 5 】

次に、ステップ S 1 3 ~ S 1 6 で、抽出部 4 1 3 が、第 1 実施形態におけるデータ出力処理と同様に、候補データを抽出する。

【 0 1 0 6 】

次に、ステップ S 4 2 で、抽出部 4 1 3 が、実データにおける数値の項目の値を、統計的手法により、所定の区間に区切り、各区間の値の個数が、所定の上限値以上の区間に属する値を持つ候補データから、所定数の候補データを選択して除外する。

10

【 0 1 0 7 】

次に、ステップ S 4 3 で、評価部 1 8 が、抽出部 4 1 3 により抽出された候補データの各々について、実データとの予め定めた相関ルールの評価値を算出する。そして、評価部 1 8 は、算出した評価値が所定の閾値未満の候補データを、候補データから除外する。

【 0 1 0 8 】

以下、ステップ S 1 7 ~ S 1 9 で、第 1 実施形態におけるデータ出力処理と同様に処理し、データ出力処理は終了する。

【 0 1 0 9 】

以上説明したように、第 4 実施形態におけるデータ出力装置 4 1 0 によれば、抽出された候補データのうち、実データとの予め定めた相関ルールの評価値が所定の閾値以上の候補データを採用する。これにより、テストデータとして必要であることの信頼性が高い候補データをテストデータ群に追加することができ、システム改修後のテストに用いるテストデータ群を出力する際に、必要なテストデータの出力漏れを低減することができる。

20

【 0 1 1 0 】

なお、第 4 実施形態では、評価値として、相関ルールの信頼度及びリフト値を用いる場合について説明したが、これに限定されない。候補データが必要なテストデータであることの尤もらしさを評価可能な従来既知の評価手法を用いることができる。

【 0 1 1 1 】

また、上記第 1、第 3、及び第 4 の実施形態では、テストデータ群を出力する場合、第 2 実施形態では、アノマリ検知データを出力する場合について説明したが、これに限定されない。第 1、第 3、及び第 4 の実施形態においても、アノマリ検知データを出力する場合に適用することができる。この場合、実データ群と網羅データとの差分のデータ群から候補データを抽出する処理に変えて、入力データを候補データとして抽出するか否かを判定する処理を行えばよい。また、出力するデータは、テストデータやアノマリ検知のための参照データに限定されない。

30

【 0 1 1 2 】

また、上記各実施形態では、抽出した候補データを一旦ユーザに提示し、ユーザにより選択された候補データを、出力するデータに加える場合について説明したが、これに限定されない。抽出部により抽出された候補データを、そのまま出力するようにしてもよい。

【 0 1 1 3 】

また、上記実施形態では、データ出力プログラム 5 0、2 5 0、3 5 0、4 5 0 が記憶部 4 3 に予め記憶（インストール）されている態様を説明したが、これに限定されない。開示の技術に係るプログラムは、CD-ROM、DVD-ROM、USBメモリ等の記憶媒体に記憶された形態で提供することも可能である。

40

【 0 1 1 4 】

以上の各実施形態に関し、更に以下の付記を開示する。

【 0 1 1 5 】

( 付記 1 )

入力データを受け付け、

予め入力データと比較を行う参照データを記憶する記憶部を参照して、受け付けた前記

50

入力データと参照データとを比較して、前記入力データのうち前記参照データと一致しない入力データがある場合には、前記入力データの属性の入力を受け付けて、受け付けた入力結果に基づいて、前記参照データを更新して前記記憶部に記憶するとともに、更新した参照データをパターン解析して参照データとして追加する候補データを出力する、

処理をコンピュータに実行させることを特徴とするデータ出力プログラム。

【0116】

(付記2)

所定の条件を満たすことが既知の既知データを受け付け、

前記既知データと比較した結果の類似度が所定値以上のデータ、及び前記既知データの統計的分析結果に基づき特定されるデータの少なくとも一方を、前記既知データに追加する候補データとして抽出し、

抽出した候補データを出力する

処理をコンピュータに実行させることを特徴とするデータ出力プログラム。

【0117】

(付記3)

前記既知データの各々は、1以上の項目を含み、

前記既知データと、前記項目の値として取り得る組み合わせを含む網羅データとの差分のデータ群から、前記候補データを抽出する

付記2に記載のデータ出力プログラム。

【0118】

(付記4)

前記候補データとして、前記既知データの統計的分布において、出現頻度が所定の下限値以下の区間に属するデータを抽出する付記2又は付記3に記載のデータ出力プログラム。

【0119】

(付記5)

前記既知データと比較した結果の類似度が所定値以上のデータ、及び前記既知データの統計的分析結果に基づき特定されるデータの少なくとも一方のデータから、前記既知データとの予め定めた相関ルールの評価値に基づいて、前記候補データを抽出する付記2～付記4のいずれか1項に記載のデータ出力プログラム。

【0120】

(付記6)

抽出された前記候補データから、前記既知データの統計的分布において、出現頻度が所定の上限値以上の区間に属するデータの一部を除外する付記2～付記5のいずれか1項に記載のデータ出力プログラム。

【0121】

(付記7)

前記既知データは、テスト対象のシステムに入力された実データであり、

前記候補データとして、前記実データと共にテストデータとする候補のデータを抽出する

付記2～付記6のいずれか1項記載のデータ出力プログラム。

【0122】

(付記8)

前記既知データは、入力データと比較を行う参照データとして記憶部に記憶されたデータであり、

前記記憶部を参照して、受け付けた前記入力データと前記参照データとを比較して、前記参照データと一致しない前記入力データから、前記候補データとして、前記参照データに追加する候補のデータを抽出する

付記2～付記6のいずれか1項記載のデータ出力プログラム。

【0123】

(付記9)

出力した前記候補データに対して、ユーザから前記候補データを前記既知データに追加

10

20

30

40

50

するか否かの指示を受け付け、

受け付けた前記指示に基づいて、前記候補データを前記既知データに追加するか否かを判定する

付記 2 ~ 付記 8 のいずれか 1 項記載のデータ出力プログラム。

【 0 1 2 4 】

( 付記 1 0 )

所定の条件を満たすことが既知の既知データを受け付け受付部と、

前記既知データと比較した結果の類似度が所定値以上のデータ、及び前記既知データの統計的分析結果に基づき特定されるデータの少なくとも一方を、前記既知データに追加する候補データとして抽出する抽出部と、

抽出した候補データを出力する出力部と、

を含むデータ出力装置。

10

【 0 1 2 5 】

( 付記 1 1 )

前記既知データの各々は、1 以上の項目を含み、

前記抽出部は、前記既知データと、前記項目の値として取り得る組み合わせを含む網羅データとの差分のデータ群から、前記候補データを抽出する

付記 1 0 に記載のデータ出力装置。

【 0 1 2 6 】

( 付記 1 2 )

前記抽出部は、前記候補データとして、前記既知データの統計的分布において、出現頻度が所定の下限值以下の区間に属するデータを抽出する付記 1 0 又は付記 1 1 に記載のデータ出力装置。

20

【 0 1 2 7 】

( 付記 1 3 )

前記抽出部は、前記既知データと比較した結果の類似度が所定値以上のデータ、及び前記既知データの統計的分析結果に基づき特定されるデータの少なくとも一方のデータから、前記既知データとの予め定めた相関ルールの評価値に基づいて、前記候補データを抽出する付記 1 0 ~ 付記 1 2 のいずれか 1 項に記載のデータ出力装置。

【 0 1 2 8 】

( 付記 1 4 )

前記抽出部は、抽出された前記候補データから、前記既知データの統計的分布において、出現頻度が所定の上限值以上の区間に属するデータの一部を除外する付記 1 0 ~ 付記 1 3 のいずれか 1 項に記載のデータ出力装置。

30

【 0 1 2 9 】

( 付記 1 5 )

前記既知データは、テスト対象のシステムに入力された実データであり、

前記抽出部は、前記候補データとして、前記実データと共にテストデータとする候補のデータを抽出する

付記 1 0 ~ 付記 1 4 のいずれか 1 項記載のデータ出力装置。

40

【 0 1 3 0 】

( 付記 1 6 )

前記既知データは、入力データと比較を行う参照データとして記憶部に記憶されたデータであり、

前記抽出部は、前記記憶部を参照して、受け付けた前記入力データと前記参照データとを比較して、前記参照データと一致しない前記入力データから、前記候補データとして、前記参照データに追加する候補のデータを抽出する

付記 1 0 ~ 付記 1 4 のいずれか 1 項記載のデータ出力装置。

【 0 1 3 1 】

( 付記 1 7 )

50

出力した前記候補データに対して、ユーザから前記候補データを前記既知データに追加するか否かの指示を受け付ける提示部をさらに含み、

前記出力部は、受け付けた前記指示に基づいて、前記候補データを前記既知データに追加するか否かを判定する

付記 10 ~ 付記 16 のいずれか 1 項記載のデータ出力装置。

【0132】

(付記 18)

所定の条件を満たすことが既知の既知データを受け付け、

前記既知データと比較した結果の類似度が所定値以上のデータ、及び前記既知データの統計的分析結果に基づき特定されるデータの少なくとも一方を、前記既知データに追加する候補データとして抽出し、

抽出した候補データを出力する

処理をコンピュータが実行することを特徴とするデータ出力方法。

【0133】

(付記 19)

入力データを受け付ける受付部と、

予め入力データと比較を行う参照データを記憶する記憶部を参照して、受け付けた前記入力データと参照データとを比較して、前記入力データのうち前記参照データと一致しない入力データがある場合には、前記入力データの属性の入力を受け付けて、受け付けた入力結果に基づいて、前記参照データを更新して前記記憶部に記憶するとともに、更新した参照データをパターン解析して参照データとして追加する候補データを出力する出力部と、を含むデータ出力プログラム。

【0134】

(付記 20)

入力データを受け付け、

予め入力データと比較を行う参照データを記憶する記憶部を参照して、受け付けた前記入力データと参照データとを比較して、前記入力データのうち前記参照データと一致しない入力データがある場合には、前記入力データの属性の入力を受け付けて、受け付けた入力結果に基づいて、前記参照データを更新して前記記憶部に記憶するとともに、更新した参照データをパターン解析して参照データとして追加する候補データを出力する、

処理をコンピュータが実行することを特徴とするデータ出力方法。

【符号の説明】

【0135】

10、210、310、410 データ出力装置

11、211 受付部

12 作成部

13、213、313、413 抽出部

14 提示部

15、215 出力部

16 判定部

17 アノマリデータベース

18 評価部

21 追加可否確認画面

40 コンピュータ

41 CPU

42 メモリ

43 記憶部

49 記憶媒体

50、250、350、450 データ出力プログラム

10

20

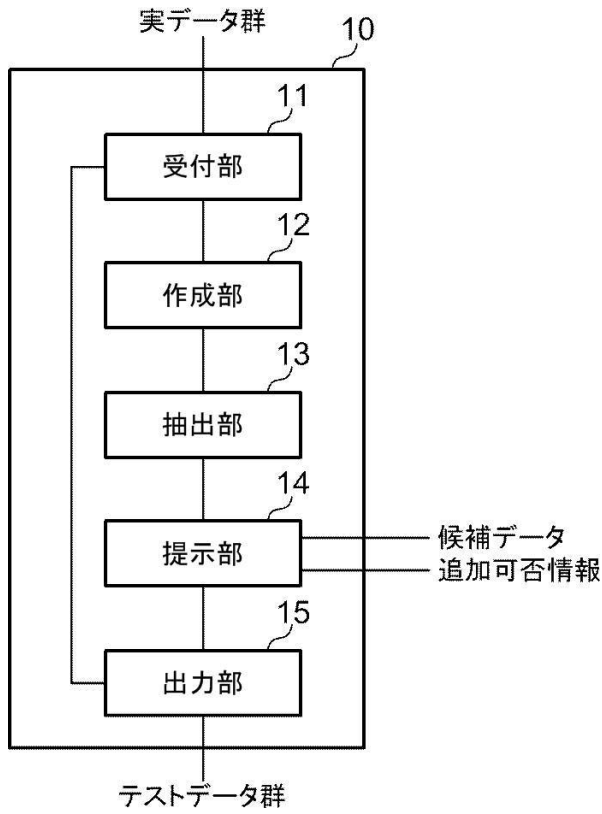
30

40

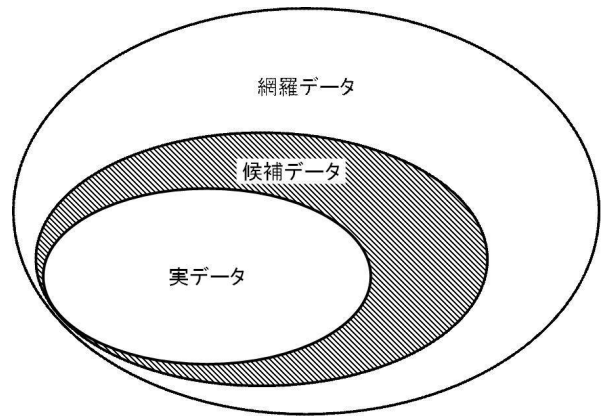
50

【図面】

【図 1】



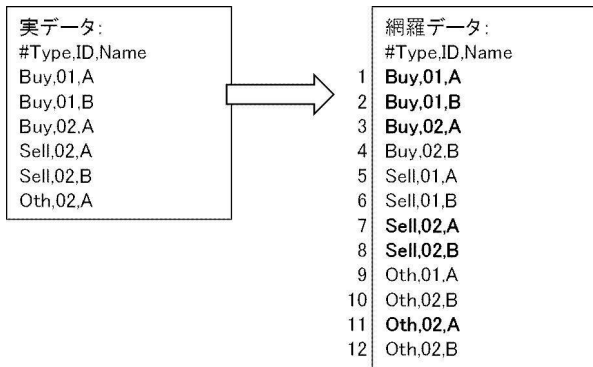
【図 2】



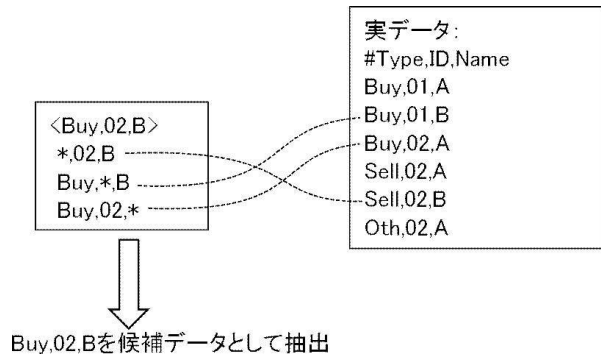
10

20

【図 3】



【図 4】

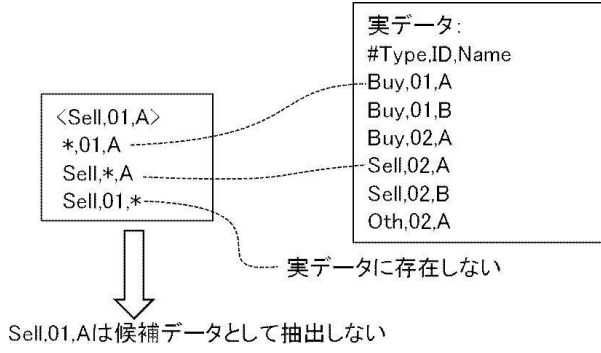


30

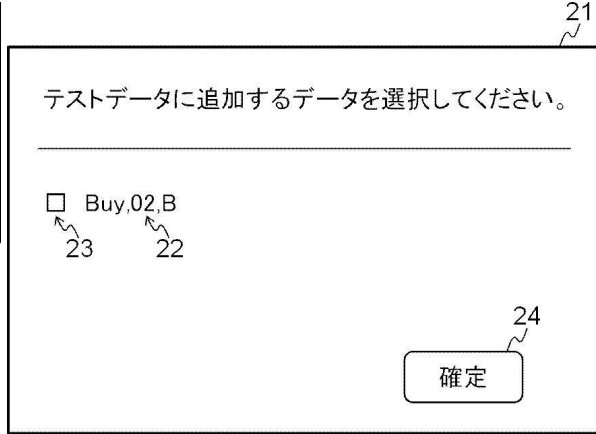
40

50

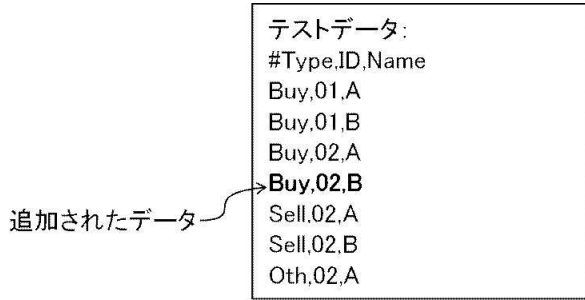
【 図 5 】



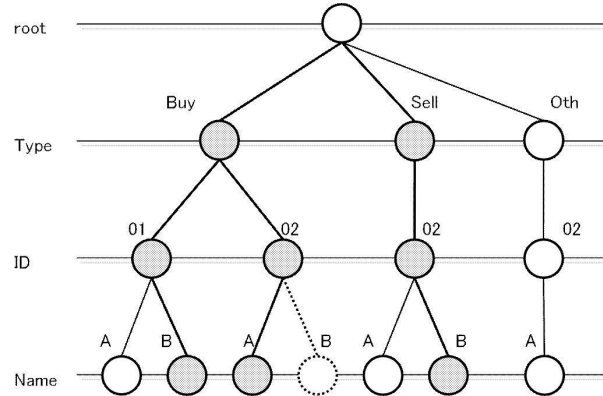
【 図 6 】



【 図 7 】



【 図 8 】



10

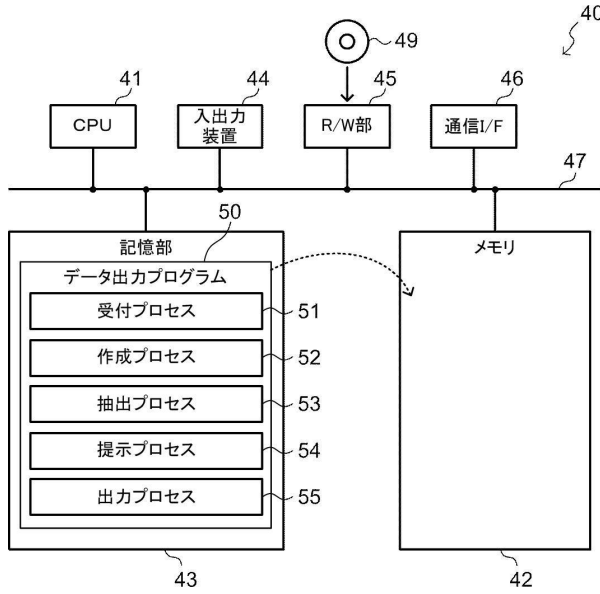
20

30

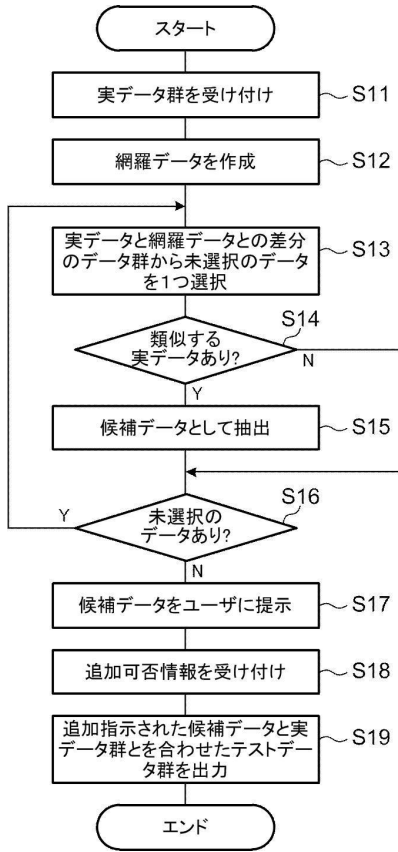
40

50

【図9】



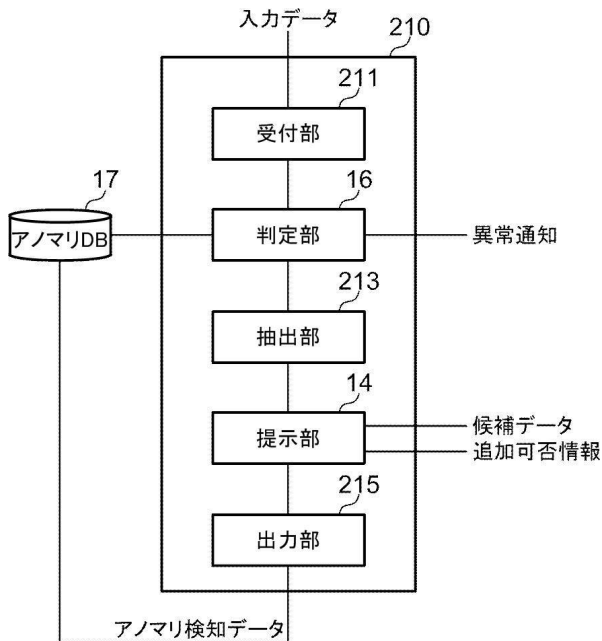
【図10】



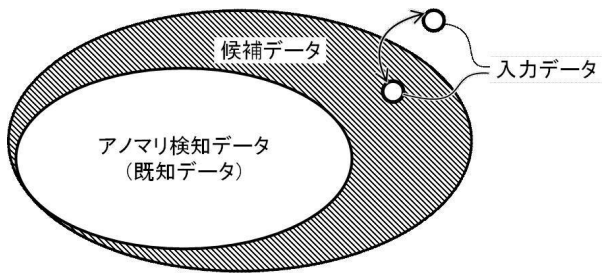
10

20

【図11】



【図12】

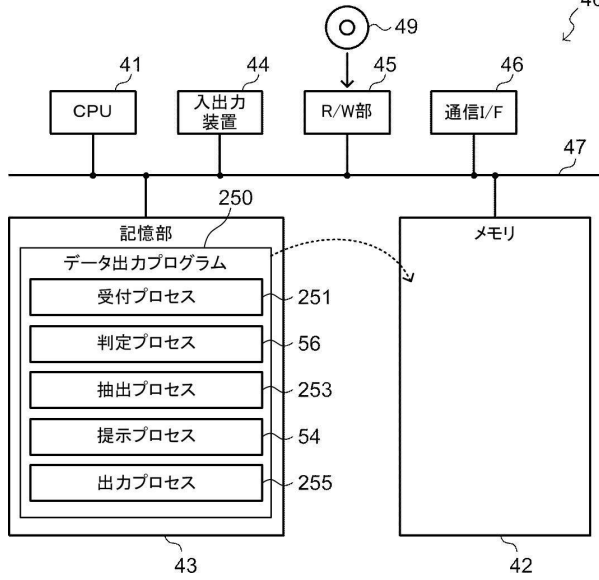


30

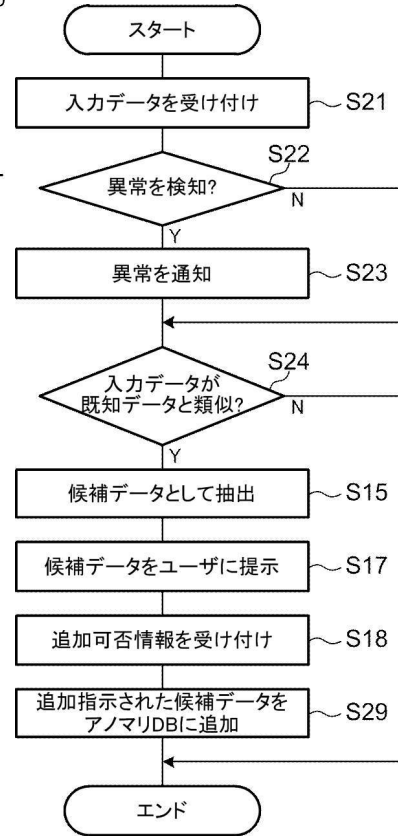
40

50

【図13】



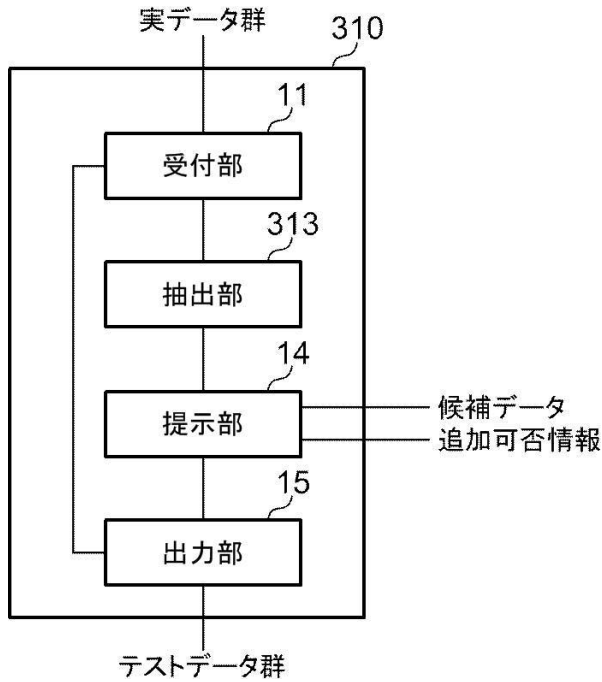
【図14】



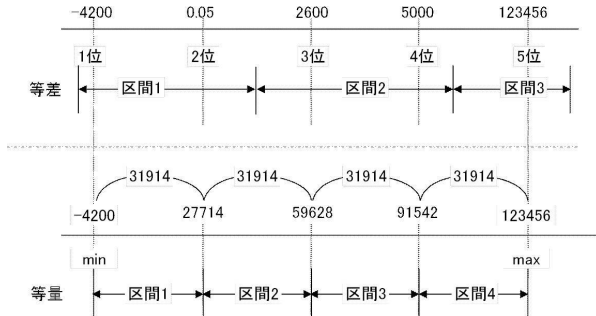
10

20

【図15】



【図16】

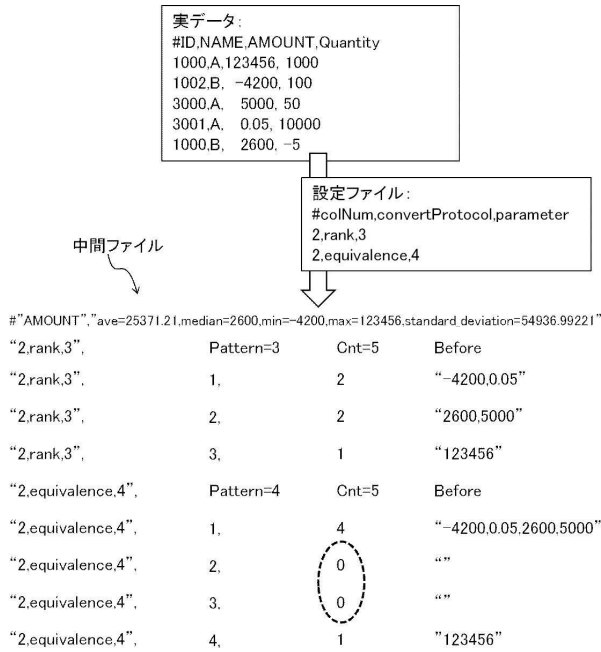


30

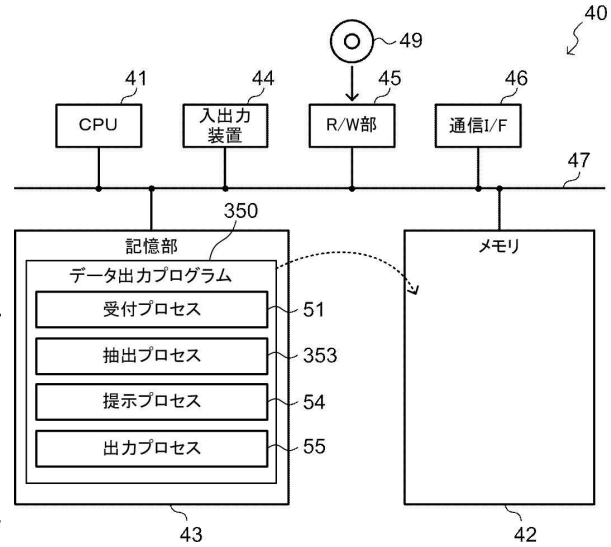
40

50

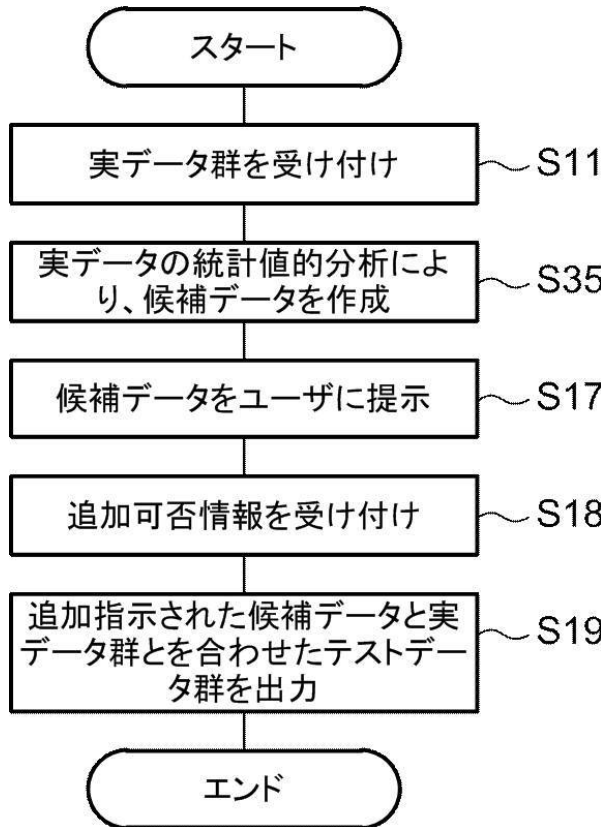
【図17】



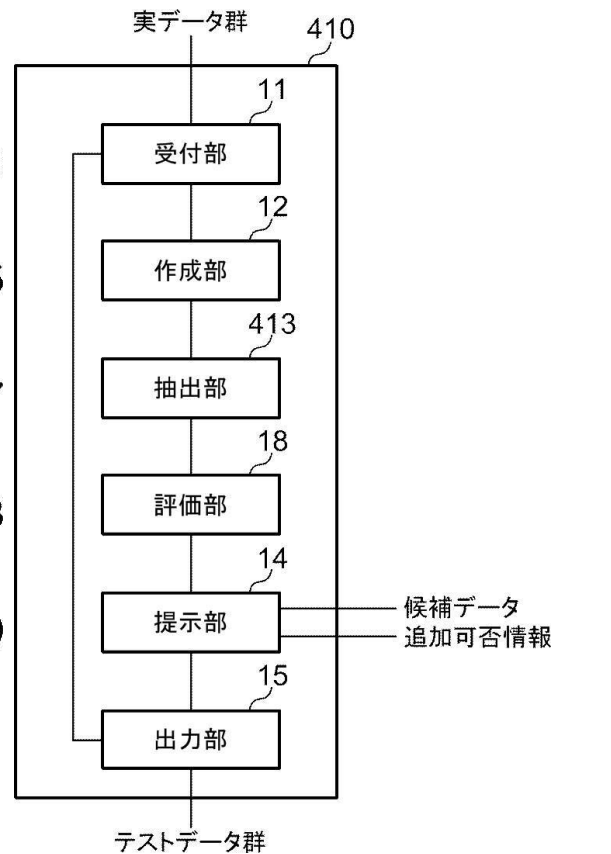
【図18】



【図19】



【図20】



10

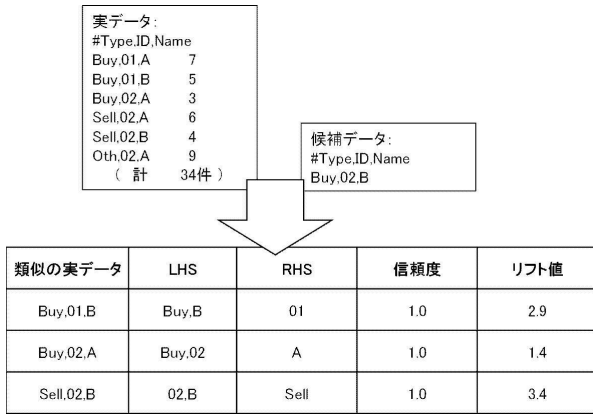
20

30

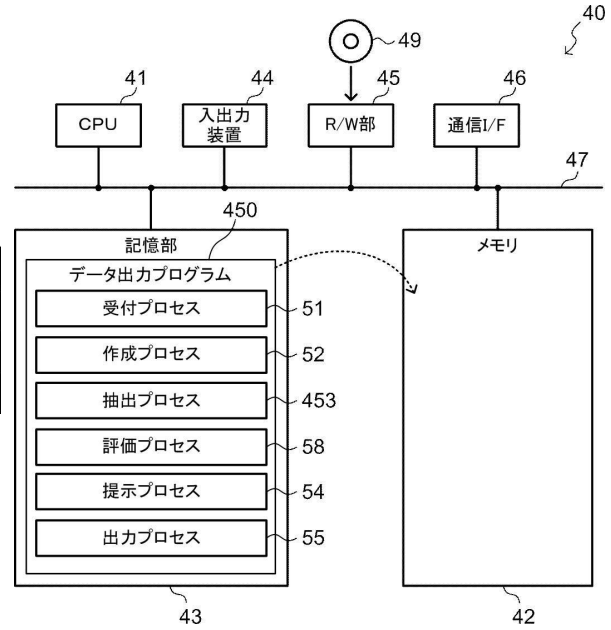
40

50

【図 2 1】



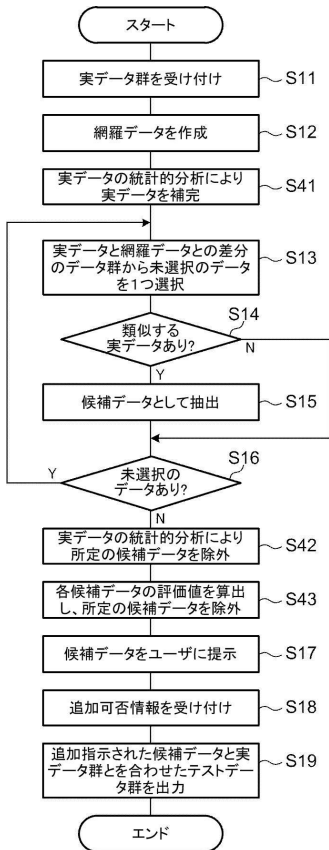
【図 2 2】



10

20

【図 2 3】



30

40

50

---

フロントページの続き

- (56)参考文献 特開 2017 - 021634 (JP, A)  
特開 2014 - 160513 (JP, A)  
国際公開第 2017 / 047296 (WO, A1)
- (58)調査した分野 (Int.Cl., DB名)  
G06F 16/00 - 16/958