

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2014年1月16日(16.01.2014)



(10) 国際公開番号

WO 2014/010071 A1

- (51) 国際特許分類:
G06F 19/00 (2011.01) G06Q 10/06 (2012.01)
- (21) 国際出願番号: PCT/JP2012/067887
- (22) 国際出願日: 2012年7月13日(13.07.2012)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (71) 出願人(米国を除く全ての指定国について): 株式会社日立製作所 (HITACHI, LTD.) [JP/JP]; 〒1008280 東京都千代田区丸の内一丁目6番6号 Tokyo (JP).
- (72) 発明者; および
- (75) 発明者/出願人(米国についてのみ): 本林 正裕 (MOTOBAYASHI Masahiro) [JP/JP]; 〒1858601 東京都国分寺市東恋ヶ窪一丁目280番地 株式会社日立製作所 中央研究所内 Tokyo (JP). 中野 定樹 (NAKANO Sadaki) [JP/JP]; 〒1858601 東京都国分寺市東恋ヶ窪一丁目280番地 株式会社日立製作所 中央研究所内 Tokyo (JP). 古川 直広 (FURUKAWA Naohiro) [JP/JP]; 〒1858601 東京都国分寺市東恋ヶ窪一丁目280番地 株式会社日立製作所 中央研究所内 Tokyo (JP).
- (74) 代理人: ポレール特許業務法人 (POLAIRE I.P.C.); 〒1040032 東京都中央区八丁堀二丁目7番1号 Tokyo (JP).
- (81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT,

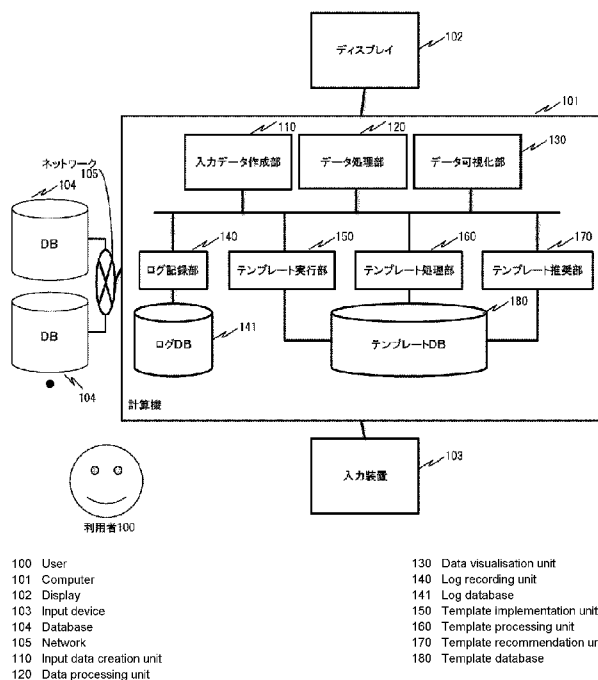
[続葉有]

(54) Title: DATA ANALYSIS ASSISTANCE METHOD AND SYSTEM

(54) 発明の名称: データ分析支援方法及びシステム

【図1】

図1



(57) Abstract: If data to be analysed, or analysis processing or visualisation processing for the data are not fixed, during trial and error the data to be analysed is regenerated and the analysis processing and visualisation processing for the data must be modified. Editing or restructuring the data to be analysed causes sections after the edited or restructured section to become non-replicable, and therefore the present invention provides editing that automatically modifies other processes or presents options for modifying other processes. A computer constituting a system comprises: a template recommendation unit that implements a template recommendation function by which the next task can be recommended to the user; and a template implementation unit that implements a template implementation function by which all types of templates can be implemented.

(57) 要約: 分析対象データやデータに対する分析処理や可視化処理が固定ではない場合には、試行錯誤のたびに、分析対象データを作り変え、分析処理、可視化処理を変更しなければならない。分析対象データの編集や再構築により、その部分より後の部分は再現不可能となるため、編集が他の処理を自動的に変更、あるいは他の処理を変更するための候補を提示する。システムを構成する計算機は、利用者に対して次作業を推奨できるテンプレート推奨機能を実行するテンプレート推奨部、各種のテンプレートを実行できるテンプレート実行機能を実行するテンプレート実行部を備える。



QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

- (84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI,

添付公開書類:

— 国際調査報告 (条約第 21 条(3))

明 細 書

発明の名称： データ分析支援方法及びシステム

技術分野

[0001] 本発明は、データ分析を支援するデータ分析支援方法及びシステムに係り、特に分析対象データやデータに対する処理が固定で無い場合の分析支援技術に関する。

背景技術

[0002] 業務システムなどから蓄積される企業内の膨大なデータを、蓄積、分析、加工して、企業の意思決定に活用するビジネスインテリジェンスという手法がある。この手法は、複数のデータベース（DB）を横断し、多種、大量のデータを高速に分析可能である。この手法は、高速化のために、高度に構造化された分析対象データを分析の目的に応じて予め構築する必要がある、分析内容と必要なデータが決まっている定型分析に強い。

[0003] 一方、分析対象データやデータに対する分析処理や可視化処理が固定ではない場合には、試行錯誤のたびに、分析対象データを作り変えると共に、分析処理、可視化処理を変更しなければならない。このようなデータ分析に関連する背景技術として、特許文献1や特許文献2がある。特許文献1には、コンピュータ側で検知した分析手順を管理サーバのデータベースに記録し、記録した分析手順を第三者が共有・再利用することが記載されている。また、特許文献2には、最適化されたプロセス定義を提供すること、言い換えると、既存プロセスを再利用し、新しいプロセスを作る際に、既存プロセス前後関係を利用し、実行可能な新しいプロセスを提示することが開示されている。

先行技術文献

特許文献

[0004] 特許文献1：特開2008-140278号公報

特許文献2：特開2005-285101号公報

発明の概要

発明が解決しようとする課題

[0005] 上述の背景技術に見るように、分析手順を記録し、編集や再構築することにより再利用する技術では、過去の分析手順の再現は可能である。また過去の分析手順を編集することにより、部分的に再利用し、新たな分析を行うことも可能である。しかしその編集により、編集した部分より後の部分は再現不可能となるため、その編集が他の処理を自動的に変更、あるいは他の処理を変更するための候補を提示するような技術が、求められている。

[0006] 本発明の目的は、上記の課題を解決し、過去の分析手順を編集し、他の処理を自動的に変更、或いは他の処理を変更するための候補を提示することが可能なデータ分析支援方法及びシステムを提供することにある。

課題を解決するための手段

[0007] 本発明においては、上記の目的を達成するため、計算機を用いたデータ分析支援方法であって、計算機は、抽出したデータに入力データ作成処理を施すことにより分析対象データを作成する入力データ作成機能と、データにデータ処理を施し処理結果データを作成するデータ処理機能と、データに可視化処理を施し可視化するデータ可視化機能を実行することによりデータ分析支援の作業を行うデータ分析支援方法を提供する。

[0008] また、本発明においては、上記の目的を達成するため、データ分析支援システムであって、複数のテンプレートを蓄積するテンプレートデータベースを記憶する記憶部と、テンプレートを実行する処理部とを備え、処理部は、データに入力データ作成処理を施すことにより分析対象データを作成する入力データ作成機能と、データにデータ処理を施し処理結果データを作成するデータ処理機能と、データに可視化処理を施し可視化するデータ可視化機能を実行することによりデータ分析支援の作業を行う構成のデータ分析支援システムを提供する。

発明の効果

[0009] 本発明によれば、利用者の分析作業に要する手間を削減することが可能となる。

図面の簡単な説明

[0010] [図1]第一の実施例の分析支援システムのブロック図である。

[図2]第一の実施例に係る、システムの作業1の手順の一例を示すフローチャートである。

[図3A]第一の実施例に係る、テンプレートリストのディスプレイ表示例を示す図である。

[図3B]第一の実施例に係る、テンプレートのデータ構造及びテンプレートの一例、ログのデータ構造及びログの一例を示す図である。

[図3C]第一の実施例に係る、ログのデータ構造及びログの一例を示す図である。

[図3D]第一の実施例に係る、テンプレートログのデータ構造及びテンプレートログの一例を示す図である。

[図3E]第一の実施例に係る、テンプレートリスト及び可視化結果図のディスプレイ表示例を示す図である。

[図3F]第一の実施例に係る、パラメータ変更用のダイアログの一例を示す図である。

[図4A]第一の実施例に係る、作業1の具体例を示す図である。

[図4B]第一の実施例に係る、作業1の具体例を示す図である。

[図5]第一の実施例に係る、図2のS208の処理について詳述する図である。

[図6]第一の実施例に係る、作業2の手順の一例を示すフローチャートである。

[図7A]第一の実施例に係る、作業2の具体例を示す図である。

[図7B]第一の実施例に係る、作業2の具体例を示す図である。

[図8]第一の実施例に係る、図6のS607の処理について詳述する図である。

[図9]第一の実施例に係る、作業3の手順の一例を示すフローチャートである。

[図10A]第一の実施例に係る、作業3の具体例を示す図である。

[図10B]第一の実施例に係る、作業3の具体例を示す図である。

[図11]第一の実施例に係る、図9のS907の内容について詳述する図であり、作業4の手順の一例を示すフローチャートである。

[図12]第一の実施例に係る、図11のS1110の処理について詳述する図である。

[図13]第一の実施例に係る、図11のS1116の処理について詳述する図である。

[図14A]第一の実施例に係る、作業4の具体例を示す図である。

[図14B]第一の実施例に係る、作業4の具体例を示す図である。

[図14C]第一の実施例に係る、作業4の具体例を示す図である。

[図14D]第一の実施例に係る、作業4の具体例を示す図である。

[図14E]第一の実施例に係る、作業4の具体例を示す図である。

[図14F]第一の実施例に係る、作業4の具体例を示す図である。

[図15]第二の実施例に係る、作業2の手順の一例を示すフローチャートである。

[図16]第二の実施例に係る、図15のS1504の処理を詳述する図である。

[図17]第二の実施例に係る、作業1の手順の一例を示すフローチャートである。

[図18]第二の実施例に係る、作業3の手順の一例を示すフローチャートである。

[図19]第二の実施例に係る、図18のS1807の内容について詳述する図であり、作業4の手順の一例を示すフローチャートである。

[図20A]第二の実施例に係る、図18のS1807の内容について詳述する図であり、作業4の手順の一例を示すフローチャートである。

[図20B]第二の実施例に係る、図18のS1807の内容について詳述する図であり、作業4の手順の一例を示すフローチャートである。

[図21A]第二の実施例に係る、作業4の具体例を示す図である。

[図21B]第二の実施例に係る、作業4の具体例を示す図である。

[図21C]第二の実施例に係る、作業4の具体例を示す図である。

[図21D]第二の実施例に係る、作業4の具体例を示す図である。

[図21E]第二の実施例に係る、作業4の具体例を示す図である。

[図21F]第二の実施例に係る、作業4の具体例を示す図である。

[図22]第三の実施例に係る、作業3の手順の一例を示すフローチャートである。

[図23]第三の実施例に係る、図22のS2204の処理を詳述する図である。

[図24]第三の実施例に係る、作業2の手順の一例を示すフローチャートである。

[図25]第三の実施例に係る、作業1の手順の一例を示すフローチャートである。

[図26A]第三の実施例に係る、図25のS2512の内容について詳述する図であり、作業4の手順の一例を示すフローチャートである。

[図26B]第三の実施例に係る、図25のS2512の内容について詳述する図であり、作業4の手順の一例を示すフローチャートである。

[図26C]第三の実施例に係る、図25のS2512の内容について詳述する図であり、作業4の手順の一例を示すフローチャートである。

[図27A]第三の実施例に係る、作業4の具体例を示す図である。

[図27B]第三の実施例に係る、作業4の具体例を示す図である。

[図28A]第四の実施例に係る、分析作業の手順を示す図である。

[図28B]第四の実施例に係る、分析作業の手順を示す図である。

[図29]第五の実施例に係る、分析作業の手順を示す図である。

[図30]第五の実施例に係る、パラメータ編集用のウィンドウの一例を示す図

である。

発明を実施するための形態

[0011] 以下、本発明のデータ分析支援方法、及び装置の各種の実施の形態について、図面を参照しながら説明する。なお、本明細書において、計算機等が実現する各種の機能を、作業或いは部と称する場合がある。例えば、分析対象データを作成するデータ作成機能を、データ作成作業、或いはデータ作成部、分析対象データにデータ処理を施し処理結果データを作成するデータ処理機能を、データ処理作業、或いはデータ処理部、処理結果データに可視化処理を施すデータ可視化機能をデータ可視化作業、或いはデータ可視化部等と呼ぶ場合がある。

実施例 1

[0012] 図1は、本実施例の分析支援システムの一構成例のブロック図である。同図のシステムは、計算機101、表示部として機能するディスプレイ102、入力装置103、及びデータベース(DB)104を備える。計算機101、DB104は、ネットワーク105を介して互いに接続される。

[0013] 計算機101は、利用者100が分析作業を行う際に利用する。さらに、分析作業に関わる種々の機能を提供し、機能実行時のログを収集し、テンプレート状況を合わせて抽出・推奨することにより利用者100の分析作業を支援する。

[0014] 計算機101は、一般的なPC(Personal Computer)が想定されている。計算機101は、図示を省略した中央処理部(Central Processing Unit: CPU)となるプロセッサ、記憶部であるメモリ、及びインターフェース部であるインターフェースを備える。プロセッサは、メモリに記憶されたプログラムを処理することによって、各種処理を実行する。メモリは、処理を実行するためのプログラム、及びデータを記憶する。インターフェースとしては、キーボード、マウスなどの入力装置103に接続するもの、ディスプレイ102に接続するもの、ネットワーク105を介して、DB104などに接続するもの、などを備える。

- [0015] DB 104 は、例えば、企業に関わる情報、各種統計データ、センサ等の時系列データ、Web アクセスログなどの各種データを保持するデータベースである。DB 104 は、図 1 に見るように、図示を省略した外部のストレージ装置に格納され、ネットワーク 105 を介して計算機 101 に接続されるように構成してもよいし、計算機 101 内に含まれるように構成してもよい。
- [0016] 計算機 101 は、入力データ作成部 110、データ処理部 120、データ可視化部 130、ログ記録部 140、ログ DB 141、テンプレート実行部 150、テンプレート処理部 160、テンプレート推奨部 170、テンプレート DB 180 を備える。入力データ作成部 110、データ処理部 120、データ可視化部 130、ログ記録部 140、ログ DB 141、テンプレート実行部 150、テンプレート処理部 160、テンプレート推奨部 170、テンプレート DB 180 は例えば、プログラムであってメモリに記憶され、プロセッサによって実行される機能ブロックである。ログ DB 141、テンプレート DB 180 は、それぞれログの、テンプレートのデータベースを示し、省略された PC 内の記憶部であるメモリに蓄積される。
- [0017] 入力データ作成部 110 は、利用者 100 あるいはテンプレート実行部 150 の指示により DB 104 から必要なデータを抽出し、抽出したデータに入力データ作成処理を施すことにより、分析対象データを作成する。
- [0018] データ処理部 120 は、利用者 100 あるいはテンプレート実行部 150 の指示により、入力データ作成部 110 が作成した分析対象データにデータ処理を施し、処理結果データを作成する。
- [0019] データ可視化部 130 は、利用者 100 あるいはテンプレート実行部 150 の指示により、データ処理部 120 が作成した処理結果データに可視化処理を施し、可視化する。
- [0020] ログ記録部 140 は、入力データ作成部 110、データ処理部 120、データ可視化部 130 の各処理をログ DB 141 に記録する。
- [0021] テンプレート実行部 150 は、利用者 100 の指示により、テンプレート

DB 180からテンプレートを取り出し、テンプレートの内容に従って入力データ作成部110、データ処理部120、データ可視化部130に指示し、分析処理を実行する。

[0022] テンプレート処理部160は、テンプレートからパラメータに関わる情報の抽出、入力データ作成部110が作成した分析対象データやデータ処理部120が作成した処理結果データに対してテンプレートが実行可能かどうかの判定を行う。

[0023] テンプレート推奨部170は、予め定められた基準でテンプレートDB180からテンプレートを抽出し、利用者100に提示する、また、入力データ作成部110が作成した分析対象データやデータ処理部120が作成した処理結果データに対して適用可能なテンプレートをテンプレートDB180から抽出し、利用者100に提示する。

[0024] 次に、第一の実施例の分析支援システムの計算機101での処理について、具体例を参照しながら説明する。本実施例では、利用者100が、作業1「DBから必要なデータを抽出し、抽出したデータに入力データ作成処理を施すことにより分析対象データを作成」、作業2「分析対象データにデータ処理を施し処理結果データを作成」、作業3「処理結果データに可視化処理を施し可視化」、作業4「作業1からやり直す」という順番で分析作業を行う場面を想定する。

[0025] 図2は、第一の実施例の上記作業1の手順の一例を示すフローチャートである。

まず、プロセッサは、テンプレート推奨部170が指定された基準に従い、テンプレートDB180から入力データ作成部110で実行できるテンプレートを抽出し、リストLiに登録する(S201)。ここで基準とは、最近利用者100が使ったテンプレート、全利用者がよく使うテンプレート、利用者100以外の特定の利用者がよく使うテンプレート、部署別によく使われるテンプレート、過去この時期によく使われたテンプレート、などのテンプレートの推奨基準を意味する。

[0026] 次に、抽出したリスト L_i を表示部であるディスプレイ 102 に表示する (S202)。ディスプレイ 102 への表示例を図 3A に示す。利用者 100 がテンプレート T_i を選択した場合、テンプレート実行部 150 が入力データ作成部用のテンプレート T_i を実行する (S203、S204)。図 3B はテンプレートのデータ構造、及びテンプレートの一例を示す図である。ここで示すテンプレートは、入力データ作成部用、データ処理部用、データ可視化部用で同じ構造を想定しているが、各部で実行可能であれば異なる構造であっても構わない。同図の上段は、テンプレートのデータ構造を、同図の下段は、テンプレートの一例を示している。同図の上段に見るように、テンプレートのデータ構造は、その項目名に見るように、パラメータ、処理名、引数からなり、図示する内容のデータが記載される。

[0027] 利用者 100 がテンプレートを選択しなかった場合は、テンプレートによる分析作業の支援は行わず、GUI (Graphical User Interface) 操作などによる通常の分析作業を実施する (S205)。S204 や S205 で指示され、入力データ作成部 110 で実行された入力データ作成処理が正常に終了した場合、その結果作成される分析対象データを D_a とする。またログ記録部 140 が入力データ作成部 110 で実行した処理内容を記録し、ログ DB 141 に登録する (S206、S207)。また、S204 でテンプレートを選択した場合、実行したテンプレートに関するログも併せて記録し、ログ DB 141 に登録する (S207)。

[0028] 図 3C の上段、下段にそれぞれログのデータ構造及びログの一例を示す。同図のログ構造に見るように、項目名としてデータ列名、時刻、処理名、引数がある。また、図 3D の上段、下段にそれぞれテンプレートログのデータ構造及びテンプレートログの一例を示す。テンプレートログ構造の項目としては、図示の通り、時刻、処理種別、テンプレート名、パラメータ値、入力データ列名、出力データ列名、可視化結果図がある。図 3E には、可視化結果図のディスプレイ 102 の表示画面上の表示例を、図 3F にはパラメータ変更用のダイアログの表示画面上の表示例を示したが、これら図 3E、図 3

Fについては後で詳述する。図2の作業1の処理フローチャートに戻り、正常に終了しなかった場合は、ディスプレイ102に問題発生箇所を表示し、終了する(S206、S220)。

[0029] さらに、テンプレートDB180からテンプレート推奨部170が入力データ作成部の処理の結果出力される分析対象データDaに適用可能であり、かつデータ処理部120で実行可能なテンプレートを抽出し、リストLdに登録する(S208)。S208の内容については後で図5を用いて詳述する。

[0030] 図4A、図4Bは、第一の実施例の作業1の具体例を示す図である。具体例としては、選択テンプレートTiの内容が、「指定された二つのテーブルを、指定した列をキー列として統合し、指定した値にキー列名を変更する」の場合に、選択テンプレートTiを実行した結果作成される分析対象データDaの一例を示す。

[0031] 図4Aは、テンプレートTiの例である。この例でのTiのパラメータは5つあり、パラメータ1は、種別がE(その他)で値が「テーブル1」、パラメータ2は種別がE(その他)で値が「テーブル2」、パラメータ3は種別がN(列名)で値が「列A」、パラメータ4は種別がN(列名)で値が「列X」、パラメータ5は、種別がE(その他)で値が「KEY」である。Tiで行う処理はパラメータ1～5を引数として、「指定された2つのテーブル(パラメータ1, 2)を、指定した2つの列(パラメータ3、パラメータ4)をキー列として統合し、指定した値(パラメータ5)にキー列名を変更する」を行う。図4Bは「テーブル統合」のイメージを示す図である。

[0032] 入力データ作成部110で実行可能な処理は図4A、図4Bで示した「テーブル統合」に限ったものではなく、フィルタリングや文字列の置換などSQLやプログラミング言語等で実現可能な処理を複数組み合わせたものを対象とする。

[0033] 図5は、図2のフローチャートで示したS208の処理について詳述するフローチャート図である。

まず、テンプレートログから処理種別が「データ処理用」であるログを全て抽出し、リスト L_2 に登録する。空のリスト L_d を用意する。分析対象データ D_a の列名リストを C_1 とする（S501）。 L_2 の要素数を N_2 、 $K_2=0$ とする（S502）。 L_2 の K_2 番目のログから入力データ列名リスト C_2 を取得し、 $C_2 \subseteq C_1$ である場合、リスト L_d に L_2 の K_2 番目のテンプレートを追加する。この処理を L_2 の要素全てに実行する（S503～S507）。以上が、本実施例における作業1の手順の説明である。

[0034] 図6は、第一の実施例の作業2の手順の一例を示すフローチャートである。

まず、ディスプレイにS208で作成したデータ処理テンプレートリスト L_d を表示する（S601）。ディスプレイ102への表示例は図3Aに示すものと同様である。利用者100がテンプレートを選択した場合、テンプレート実行部150がデータ処理部用のテンプレート T_d を実行する（S602、S603）。図3Bはテンプレートのデータ構造及びテンプレートの一例を示す図である。ここで示すテンプレートは、入力データ作成部用、データ処理部用、データ可視化部用で同じ構造を想定しているが、各部で実行可能であれば異なる構造であっても構わない。

[0035] 利用者100がテンプレートを選択しなかった場合は、テンプレートによる分析作業の支援は行わず、GUI操作などによる通常の分析作業を実施する（S604）。S603やS604で指示され、データ処理部120で実行された分析処理が正常に終了した場合、その結果作成される処理結果データを D_r とする。また、ログ記録部140がデータ処理部120で実行した処理内容を記録し、ログDB141に登録する（S605、S606）。また、S603でテンプレートを選択した場合、実行したテンプレートに関するログも併せて記録し、ログDB141に登録する（S606）。図3Cにログのデータ構造及びログの一例を示す。図3Dにテンプレートログのデータ構造及びテンプレートログの一例を示す。正常に終了しなかった場合、ディスプレイに問題発生箇所を表示し、終了する（S605、S220）。

[0036] さらに、テンプレートDB 180からテンプレート推奨部 170がデータ処理部の処理の結果出力される処理結果データDrに適用可能でありかつデータ可視化部 130で実行可能なテンプレートを抽出し、リストLvに登録する(S607)。S607の内容については図8で詳述する。

[0037] 図7は、第一の実施例の前記作業2の具体例を示す図である。具体例としては、選択テンプレートTdの内容が「指定した列の要素の出現数を数える」の場合に、選択テンプレートTdを実行した結果作成される処理結果データDrの一例を示す。

[0038] 図7Aは、テンプレートTdの例である。この例でのTdのパラメータは1つあり、パラメータ1は、種別がN(列名)で値が「KEY」である。Tdで行う処理はパラメータ1を引数として、「指定した列(パラメータ1)の要素の出現数を数える」を行う。図7Bは「要素の出現数を数える」のイメージを示す図である。

[0039] データ処理部で実行可能な処理は図7で示した「要素の出現数を数える」に限ったものではなく、フィルタリングや文字列の置換、四則演算、統計演算などSQLやプログラミング言語等で実現可能な処理を複数組み合わせたものを対象とする。

[0040] 図8は、図6のフローチャート中のS607の処理について詳述する図である。

まず、テンプレートログから処理種別が「可視化用」であるログを全て抽出し、リストL2に登録する。空のリストLvを用意する。処理結果データDrの列名リストをC1とする(S801)。L2の要素数をN2、K2=0とする(S802)。L2のK2番目のログから入力データ列名リストC2を取得し、C2 $\subseteq$ C1である場合、リストLvにL2のK2番目のテンプレートを追加する。この処理をL2の要素全てに実行する(S803~S807)。以上が、本実施例における作業2の手順の説明である。

[0041] 図9は、第一の実施例の前記作業3の手順の一例を示すフローチャートである。

まず、ディスプレイにS 6 0 7で作成したデータ可視化テンプレートリストLv及びLvに含まれる各テンプレートの可視化例を表示する（S 9 0 1）。ディスプレイ1 0 2への表示例を図3 A、図3 Eに示す。図3 Eの3 1 0はLvの一覧、3 1 1には、3 1 0に表示されているテンプレート名を選択すると過去にこのデータ可視化テンプレートの可視化結果が表示される。

[0042] 利用者1 0 0がテンプレートを選択した場合、テンプレート実行部1 5 0がデータ可視化部用のテンプレートTvを実行する（S 9 0 2、S 9 0 3）。図3 Bはテンプレートのデータ構造及びテンプレートの一例を示す図である。ここで示すテンプレートは、入力データ作成部用、データ処理部用、データ可視化部用で同じ構造を想定しているが、各部で実行可能であれば異なる構造であっても構わない。

[0043] 利用者1 0 0がテンプレートを選択しなかった場合は、テンプレートによる分析作業の支援は行わず、GUI操作などによる通常の分析作業を実施する（S 9 0 4）。S 9 0 3やS 9 0 4で指示され、データ可視化部1 3 0で実行された分析処理が正常に終了した場合、可視化結果をディスプレイ1 0 2に表示する。また、ログ記録部1 4 0がデータ可視化部1 3 0で実行した処理内容を記録し、ログDB 1 4 1に登録する（S 9 0 5、S 9 0 6）。また、S 9 0 3でテンプレートを選択した場合、実行したテンプレートに関するログも併せて記録し、ログDB 1 4 1に登録する（S 9 0 6）。図3 Cにログのデータ構造及びログの一例を示す。図3 Dにテンプレートログのデータ構造及びテンプレートログの一例を示す。正常に終了しなかった場合、ディスプレイに問題発生箇所を表示し、終了する（S 9 0 5、S 2 2 0）。

[0044] さらに、可視化結果が利用者1 0 0の所望のものではなかった場合、利用者1 0 0が前記作業1～3の処理を改変して再実行する（S 9 0 7）。S 9 0 7の内容については図1 1で詳述する。

[0045] 図1 0 A、図1 0 Bは、第一の実施例の前記作業3の具体例を示す図である。具体例としては、選択テンプレートTvの内容が「指定した列を棒グラフで表示し、指定した列を横軸の項目名とする」の場合に、選択テンプレ

トTvを実行した結果作成される可視化の一例を示す。

[0046] 図10Aは、テンプレートTvの例である。この例でのTvのパラメータは2つあり、パラメータ1は、種別がN（列名）で値が「要素の出現数」、パラメータ2は、種別がN（列名）で値が「KEY」である。Tvで行う処理はパラメータ1、2を引数として、「指定した列（パラメータ1）を棒グラフで表示し、指定した列（パラメータ2）を横軸の項目名とする」を行う。図10Bは「棒グラフ」のイメージを示す図である。

[0047] データ可視化部で実行可能な処理は図10で示した「指定した列を棒グラフで表示し、指定した列を横軸の項目名とする」に限ったものではなく、グラフ化、アニメーション、などプログラミング言語等で実現可能な処理を複数組み合わせたものを対象とする。

[0048] 図11は、図9のフローチャートのS907の内容について詳述する図であり、第一の実施例の前記作業4の手順の一例を示すフローチャートである。

まず、入力データ作成部用のテンプレートTiのパラメータを変更するか、入力データ作成部用のテンプレートTiを選択しなおすか、選択する（S1101、S1102）。入力データ作成部用のテンプレートTiを選択しなおす場合、S201に進む。この処理については前述の通りであるため、ここでは割愛する。

[0049] 入力データ作成部用のテンプレートTiのパラメータを変更する場合、テンプレート処理部160がTiからパラメータを抽出し、ディスプレイにパラメータ一覧を表示する（S1103）。利用者100により変更するパラメータが選択された場合、ディスプレイにパラメータ編集用のウィンドウを表示し、利用者100により変更されたパラメータ及びその値を取得し、変更されたパラメータを反映させた入力データ作成部用のテンプレートTiを実行する（S1104～S1107）。

[0050] パラメータ編集用のウィンドウの一例を図3Fに示す。図3Fの301にはパラメータ名、302には値の候補、303には、302で選択された値

が表示され、利用者１００はこの値を採用してもよいし、この値を編集してもよい。３０２に表示される値の候補は、自動的に取得される。例えば、パラメータの種別が「Ｔ」である場合、アクセス可能なテーブル名一覧が３０２に表示される。

[0051] 入力データ作成部１１０で実行された入力データ作成処理が正常に終了した場合、その結果作成される分析対象データをＤａとする。またログ記録部１４０が入力データ作成部１１０で実行した処理内容を記録し、ログＤＢ１４１に登録する（Ｓ１１０８、Ｓ１１０９）。また、実行したテンプレートＴｉに関するログも併せて記録し、ログＤＢ１４１に登録する（Ｓ１１０９）。図３Ｃにログのデータ構造及びログの一例を示す。図３Ｄにテンプレートログのデータ構造及びテンプレートログの一例を示す。正常に終了しなかった場合、ディスプレイに問題発生箇所を表示し、終了する（Ｓ１１０８、Ｓ２２０）。

[0052] さらに、テンプレート処理部１６０が分析対象データＤａに対してＴｄが実行可能かどうか判定し、判定結果からＴｄが実行可能な場合、テンプレート実行部１５０がＴｄを実行する（Ｓ１１１０、Ｓ１１１１、Ｓ１１１３）。Ｓ１１１０の内容については図１２で詳述する。

[0053] データ処理部１２０で実行されたデータ処理が正常に終了した場合、その結果作成される処理結果データをＤｒとする。またログ記録部１４０がデータ処理部１２０で実行した処理内容を記録し、ログＤＢ１４１に登録する（Ｓ１１１４、Ｓ１１１５）。また、実行したテンプレートＴｄに関するログも併せて記録し、ログＤＢ１４１に登録する（Ｓ１１１５）。図３Ｃにログのデータ構造及びログの一例を示す。図３Ｄにテンプレートログのデータ構造及びテンプレートログの一例を示す。

[0054] Ｓ１１１１で実行不可能である場合、またはＳ１１１４で正常終了しなかった場合、テンプレート推奨部１７０が分析対象データＤａに適用可能なデータ処理部用のテンプレートを抽出し、リストＬｄに登録し（Ｓ１１１２）、Ｓ６０１に進む。

[0055] さらに、テンプレート処理部160が処理結果データDrに対してTvが実行可能かどうか判定し、判定結果からTvが実行可能な場合、テンプレート実行部150がTvを実行する(S1116、S1117、S1119)。S1116の内容については図13で詳述する。

[0056] データ可視化部130で実行された可視化処理が正常に終了した場合、可視化結果をディスプレイに表示する。またログ記録部140がデータ可視化部130で実行した処理内容を記録し、ログDB141に登録する(S1120、S1121)。また、実行したテンプレートTvに関するログも併せて記録し、ログDB141に登録する(S1121)。図3Cにログのデータ構造及びログの一例を示す。図3Dにテンプレートログのデータ構造及びテンプレートログの一例を示す。

[0057] S1117で実行不可能である場合、またはS1120で正常終了しなかった場合、テンプレート推奨部170が処理結果データDrに適用可能なデータ可視化部用のテンプレートを抽出し、リストLvに登録し(S1118)、S901に進む。

[0058] 図12は、S1110の処理について詳述する図である。

[0059] まず、テンプレート処理部160が、テンプレートログからデータ処理部のテンプレートTdを実行したログを全て抽出し、リストL2に登録する(S1201)。L2の要素数をN2、 $K2=0$ とする、分析対象データDaの列名リストをC1とする(S1202)。L2のK2番目のログから入力データ列名リストC2を取得し、 $C2 \subseteq C1$ である場合、判定結果を「実行可能」とし、Tdに与えるパラメータをL2のK2番目のログのパラメータ値として、終了する。この処理をL2の要素全てに実行し(S1203～S1207)、判定結果が「実行可能」にならない場合、判定結果を「実行不可能」として終了する(S1203、S1208)。

[0060] 図13は、S1116の処理について詳述する図である。

[0061] まず、テンプレート処理部160が、テンプレートログからデータ可視化部用のテンプレートTvを実行したログを全て抽出し、リストL2に登録す

る (S 1 3 0 1)。L 2 の要素数を N 2、K 2 = 0 とする、処理結果データ D r の列名リストを C 1 とする (S 1 3 0 2)。L 2 の K 2 番目のログから入力データ列名リスト C 2 を取得し、C 2 $\subseteq$ C 1 である場合、判定結果を「実行可能」とし、T v に与えるパラメータを L 2 の K 2 番目のログのパラメータ値として、終了する。この処理を L 2 の要素全てに実行し (S 1 3 0 3 ~ S 1 3 0 7)、判定結果が「実行可能」にならない場合、判定結果を「実行不可能」として終了する (S 1 3 0 3、S 1 3 0 8)。

[0062] 図 1 4 A ~ 図 4 F は、第一の実施例の前記作業 4 の具体例を示す図である。具体例としては、図 4、図 7、図 1 0 の作業 1 ~ 作業 3 の分析処理では、「列 A」及び「列 X」をキー列として入力データ作成処理を行ったが、図 1 4 A ~ 図 4 F の例では S 1 1 0 3 ~ S 1 1 0 5 で入力データ作成部用のテンプレート T i のパラメータ 3 を「列 C」に、パラメータ 4 を「列 Z」にそれぞれ変更し、「列 C」及び「列 Z」をキー列として入力データ作成処理を行った場合に、テンプレート T d、T v を用いてデータ処理、可視化処理を実行する場面を想定する。

[0063] 図 1 4 A は、S 1 1 0 3 ~ S 1 1 0 6 で入力データ作成部用のテンプレート T i のパラメータ 3 を「列 C」に、パラメータ 4 を「列 Z」にそれぞれ変更し「指定された 2 つのテーブル (パラメータ 1, 2) を、指定した 2 つの列 (パラメータ 3、パラメータ 4) をキー列として統合し、指定した値 (パラメータ 5) にキー列名を変更する」を行う。図 1 4 B は「テーブル統合」のイメージを示す図である。

[0064] 図 1 4 C は、選択テンプレート T d の例である。この例では選択テンプレート T d はテンプレート T i の実行の結果の分析対象データ D a に適用可能であるため、パラメータの変更などは行わない。図 1 4 D は「要素の出現数を数える」のイメージを示す図である。

[0065] 図 1 4 E は、テンプレート T v の例である。この例ではテンプレート T v はテンプレート T d の実行の結果の処理結果データ D r に適用可能であるため、パラメータの変更などは行わない。図 1 4 F は「棒グラフ」のイメージ

を示す図である。

- [0066] 以上説明した第一の実施例によれば、分析作業の入力データ作成処理の変更が可能となり、また入力データ作成処理の変更に伴う、データ処理、可視化処理の変更もテンプレートを用いて支援することが可能となるため、利用者100の分析作業の手間を削減することが可能となる。

実施例 2

- [0067] 以上説明した第一の実施例では、具体例としては、利用者100が作業1「DB104から必要なデータを抽出し、抽出したデータに入力データ作成処理を施すことにより分析対象データを作成」、作業2「分析対象データにデータ処理を施し処理結果データを作成」、作業3「処理結果データに可視化処理を施し可視化」、作業4「作業1からやり直す」という順番で分析作業を行う場面を想定したが、第二の実施例では、作業2、作業1、作業3、作業4の順番で分析作業を行う場面を想定したデータ分析支援システムの実施例である。

- [0068] 図15は、第二の実施例の、前記作業2の手順の一例を示すフローチャートである。

まず、テンプレート推奨部170が指定された基準に従い、テンプレートDB180からデータ処理部120で実行できるテンプレートを抽出し、リストLdに登録する(S1501)。ここで指定された基準とは、最近利用者100が使ったテンプレート、全利用者がよく使うテンプレート、利用者100以外の特定の利用者がよく使うテンプレート、部署別によく使われるテンプレート、過去この時期によく使われたテンプレート、などを意味する。

- [0069] 次に、抽出したリストLdをディスプレイ102に表示する(S1502)。ディスプレイ102への表示例を図3Aに示す。利用者100がテンプレートを選択した場合、テンプレート推奨部が選択されたデータ処理テンプレートTdの実行に必要な分析対象データを作成できる入力データ作成テンプレートを抽出し、リストLiに登録する(S1503, S1504)。

[0070] 図16はS1504の処理を詳述する図である。

テンプレートログからTdを実行したログを全て抽出しL1とする。テンプレートログから処理種別が「入力データ作成用」であるログを全て抽出し、L2とする。空のリストをLiとする(S1601)。L1の要素数をN1、L2の要素数をN2、K1=0、K2=0とする(S1602)。L1のK1番目のログから入力データ列名リストC1を取得する(S1603)。L2のK2番目のログから出力データ列名リストC2取得し、 $C2 \subseteq C1$ である場合、リストLiにL2のK2番目のテンプレートを追加する。この処理をL1及びL2の要素全てに実行する(S1603~S1610)。

[0071] 図17は、第二の実施例の前記作業1の手順の一例を示すフローチャートである。

まず、ディスプレイにS1504で作成した入力データ作成テンプレートリストLiを表示する(S1701)。ディスプレイ102への表示例は図3Aに示すものと同様である。利用者100がテンプレートを選択した場合、テンプレート実行部150が入力データ作成部用のテンプレートTiを実行する(S1702、S1703)。図3Bはテンプレートのデータ構造及びテンプレートの一例を示す図である。ここで示すテンプレートは、入力データ作成部用、データ処理部用、データ可視化部用で同じ構造を想定しているが、各部で実行可能であれば異なる構造であっても構わない。

[0072] 利用者100がテンプレートを選択しなかった場合は、テンプレートによる分析作業の支援は行わず、GUI操作などによる通常の実施する(S1704)。S1703やS1704で指示され、入力データ作成部110で実行された入力データ作成処理が正常に終了した場合、その結果作成される分析対象データをDaとする。また、ログ記録部140が入力データ作成部110で実行した処理内容を記録し、ログDB141に登録する(S1705、S1706)。また、S1703でテンプレートを選択した場合、実行したテンプレートに関するログも併せて記録し、ログDB141に登録する(S1706)。図3Cにログのデータ構造及びログの一例を示す

。図3Dにテンプレートログのデータ構造及びテンプレートログの一例を示す。正常に終了しなかった場合、ディスプレイに問題発生箇所を表示し、終了する（S1705、S220）。

[0073] 次に、テンプレート実行部150がデータ処理部用のテンプレートTdを実行する（S1707）。図3Bはテンプレートのデータ構造及びテンプレートの一例を示す図である。ここで示すテンプレートは、入力データ作成部用、データ処理部用、データ可視化部用で同じ構造を想定しているが、各部で実行可能であれば異なる構造であっても構わない。

[0074] データ処理部120で実行された分析処理が正常に終了した場合、その結果作成される処理結果データをDrとする。また、ログ記録部140がデータ処理部120で実行した処理内容を記録し、ログDB141に登録する（S1708、S1709）。また、実行したテンプレートに関するログも併せて記録し、ログDB141に登録する（S1709）。図3Cにログのデータ構造及びログの一例を示す。図3Dにテンプレートログのデータ構造及びテンプレートログの一例を示す。正常に終了しなかった場合、ディスプレイに問題発生箇所を表示し、終了する（S1708、S220）。

[0075] さらに、テンプレートDB180からテンプレート推奨部170がデータ処理部の処理の結果出力される処理結果データDrに適用可能でありかつデータ可視化部130で実行可能なテンプレートを抽出し、リストLvに登録する（S1710）。S1710の内容についてはS607と同様である。

[0076] 図18は、第二の実施例の前記作業3の手順の一例を示すフローチャートである。

[0077] まず、ディスプレイにS1710で作成した可視化テンプレートリストLv及びLvに含まれる各テンプレートの可視化例を表示する（S1801）。ディスプレイ102への表示例は図3A、3Eに示すものと同様である。利用者100がテンプレートを選択した場合、テンプレート実行部150がデータ可視化部用のテンプレートTvを実行する（S1802、S1803）。図3Bはテンプレートのデータ構造及びテンプレートの一例を示す図で

ある。ここで示すテンプレートは、入力データ作成部用、データ処理部用、データ可視化部用で同じ構造を想定しているが、各部で実行可能であれば異なる構造であっても構わない。

[0078] 利用者100がテンプレートを選択しなかった場合は、テンプレートによる分析作業の支援は行わず、GUI操作などによる通常の分析作業を実施する(S1804)。S1803やS1804で指示され、データ可視化部130で実行された分析処理が正常に終了した場合、可視化結果をディスプレイ102に表示する。また、ログ記録部140がデータ可視化部130で実行した処理内容を記録し、ログDB141に登録する(S1805、S1806)。また、S1803でテンプレートを選択した場合、実行したテンプレートに関するログも併せて記録し、ログDB141に登録する(S1806)。図3Cにログのデータ構造及びログの一例を示す。図3Dにテンプレートログのデータ構造及びテンプレートログの一例を示す。正常に終了しなかった場合、ディスプレイに問題発生箇所を表示し、終了する(S1805、S220)。

[0079] さらに、可視化結果が利用者100の所望のものではなかった場合、前記作業1～3の処理を改変して再実行する(S1807)。S1807の内容については図19で詳述する。

[0080] 図19、図20A、図20Bは、S1807の内容について詳述する図であり、第二の実施例の前記作業4の手順の一例を示すフローチャートである。

まず、データ処理部用のテンプレートTdを変更するかどうか、選択する(S1901)。データ処理部用のテンプレートTdを変更する場合、テンプレート推奨部170が指定された基準に従い、テンプレートDB180からデータ処理部120で実行できるテンプレートを抽出し、利用者100に提示し、利用者100の選択を受けて、S1903に進む(S1902)。ここで指定された基準とは、最近利用者100が使ったテンプレート、全利用者がよく使うテンプレート、利用者100以外の特定の利用者がよく使う

テンプレート、部署別によく使われるテンプレート、過去この時期によく使われたテンプレート、などを意味する。

[0081] 次に、テンプレート処理部160がTdからパラメータを抽出し、ディスプレイにパラメータ一覧を表示する(S1903)。利用者100によりパラメータが変更された場合、利用者100により変更されたパラメータ及びその値を取得し、S2001に進む(S1904、S1905)。パラメータ編集用のウィンドウの一例を図3Fに示す。図3Fの302に表示される値の候補は、選択されたパラメータに応じてTiの実行の結果得られるデータDaなどから自動的に入手する。

[0082] 次に、分析対象データDaの列名リストをC0とし、データ処理部用のテンプレートTdから種別が「N」であるパラメータ及び引数の値を取り出し列名リストC2に登録する(S2001、S2002)。C2 $\subseteq$ C0の場合、S1907に進む(S2003)。C2 $\subseteq$ C0ではない場合、Tiの行数をR0、K0=0とする(S2004)。

[0083] Tiの1~K0行目までを実行し、その結果作成されるデータの列名リストをC0'とし、C2 $\subseteq$ C0'である場合、TiのK0行目までをTiとし、S1907に進む。この処理をK0=R0になるまで続け、K0=R0になったらS2010に進む(S2005~S2009)。

[0084] 次に、テンプレートログから処理種別が「入力データ作成用」であるログを全て抽出し、リストL1に登録する(S2010)。L1の要素数をN1、K1=0とする(S2011)。L1のK1番目のログから出力データ列名リストC1取得し、C2 $\subseteq$ C1の場合、L1のK1番目のテンプレートをTiとし、S1907に進む。この処理をK1=R1になるまで続け、K1=R1になったらS2017に進む(S2012~S2016)。

[0085] 次に、L1の要素数をN1、K1=0とする(S2017)。L1のK1番目のテンプレートをTK1とする。TK1の行数をR3、K3=0とする。TK1の1~K3行目までを実行し、その結果作成されるデータの列名リストをC3'とする。C2 $\subseteq$ C3'の場合、TiのK3行目までをTiとし

、S 1 9 0 7に進む。この処理をT K 1の全行とL 1の全要素に適用し、K 1 = N 1となった場合、T iを空にし、S 1 9 0 7に進む（S 2 0 1 8～S 2 0 2 5）。

[0086] 次に、T iが空の場合、テンプレートによる分析作業の支援は行わず、G U I操作などによる通常の分析作業を実施する（S 1 9 0 7、S 1 9 0 8）。T iが空ではない場合、テンプレート実行部1 5 0が入力データ作成部用のテンプレートT iを実行する（S 1 9 0 7、S 1 9 0 9）。S 1 9 0 8やS 1 9 0 9で指示され、入力データ作成部1 1 0で実行された入力データ作成処理が正常に終了した場合、ログ記録部1 4 0が入力データ作成部1 1 0で実行した処理内容を記録し、ログDB 1 4 1に登録する（S 1 9 1 0、S 1 9 1 1）。また、S 1 9 0 9でテンプレートT iを実行した場合、実行したテンプレートに関するログも併せて記録し、ログDB 1 4 1に登録する（S 1 9 1 1）。正常に終了しなかった場合、ディスプレイに問題発生箇所を表示し、終了する（S 1 9 1 0、S 2 2 0）。

[0087] テンプレート実行部1 5 0がT dを実行する（S 1 9 1 2）。データ処理部1 2 0で実行されたデータ処理が正常に終了した場合、その結果作成される処理結果データをD rとする。またログ記録部1 4 0がデータ処理部1 2 0で実行した処理内容を記録し、ログDB 1 4 1に登録する（S 1 9 1 3、S 1 9 1 4）。また、実行したテンプレートT dに関するログも併せて記録し、ログDB 1 4 1に登録する（S 1 9 1 4）。正常に終了しなかった場合、ディスプレイに問題発生箇所を表示し、終了する（S 1 9 1 3、S 2 2 0）。

[0088] さらに、テンプレート処理部1 6 0が処理結果データD rに対してT vが実行可能かどうか判定し、判定結果からT vが実行可能な場合、テンプレート実行部1 5 0がT vを実行する（S 1 9 1 5、S 1 9 1 6、S 1 9 1 7）。S 1 9 1 5の内容については図1 3で既述のためここでは割愛する。

[0089] データ可視化部1 3 0で実行された可視化処理が正常に終了した場合、可視化結果をディスプレイに表示する。またログ記録部1 4 0がデータ可視化

部 1 3 0 で実行した処理内容を記録し、ログ DB 1 4 1 に登録する (S 1 9 1 8、S 1 9 1 9)。また、実行したテンプレート T v に関するログも併せて記録し、ログ DB 1 4 1 に登録する (S 1 9 1 9)。正常に終了しなかった場合、ディスプレイに問題発生箇所を表示し、終了する (S 1 9 1 8、S 2 2 0)。

[0090] S 1 9 1 6 で実行不可能である場合、テンプレート推奨部 1 7 0 が処理結果データ D r に適用可能なデータ可視化部用のテンプレートを抽出し、リスト L v に登録し S 1 8 0 1 に進む (S 1 9 2 0)。

[0091] 図 2 1 A ~ 図 2 1 F は、第二の実施例の前記作業 4 の具体例を示す図である。具体例としては、図 4、図 7、図 1 0 の作業 1 ~ 作業 3 の分析処理では、「要素数を数える」としてデータ処理を行ったが、図 2 1 の例では S 1 9 0 2 ~ S 1 9 0 6 でデータ処理部用のテンプレート T d を「指定した二列のクロス集計を行う」を選択し、パラメータ 1 を「K E Y」、パラメータ 2 を「列 A」とし、テンプレート T i、T v を用いてデータ処理、可視化処理を実行する場面を想定する。

[0092] 図 2 1 A は、テンプレート T d の例である。この例での T d のパラメータは 2 つあり、パラメータ 1 は、種別が N (列名) で値が「K E Y」、パラメータ 2 は、種別が N (列名) で値が「列 A」である。T d で行う処理はパラメータ 1、2 を引数として、「指定した二列 (パラメータ 1、パラメータ 2) のクロス集計を行う」を行う。図 2 1 B は「クロス集計」のイメージを示す図である。

[0093] 図 2 1 C は、テンプレート T i の例である。テンプレート T d の変更では、テンプレート T i に影響が無いいためパラメータの変更などは行わない。図 2 1 D は「テーブル統合」のイメージを示す図である。

[0094] 図 2 1 E は、テンプレート T v の例である。この例では処理結果データ D r の変更に伴い、適用可能なテンプレート T v を抽出する。図 2 1 E は抽出された T v の内容が「指定した 2 列を棒グラフで表示し、指定した列を横軸の項目名とする」の場合を示し、図 2 1 F は「棒グラフ」のイメージを示す

図である。

- [0095] 以上説明した第二の実施例によれば、分析作業のデータ処理の変更が可能となり、またデータ処理の変更に伴う、入力データ作成処理、可視化処理の変更もテンプレートを用いて支援することが可能となるため、利用者100の分析作業の手間を削減することが可能となる。

実施例 3

- [0096] 第一の実施例では、利用者100が作業1「DB104から必要なデータを抽出し、抽出したデータに入力データ作成処理を施すことにより分析対象データを作成」、作業2「分析対象データにデータ処理を施し処理結果データを作成」、作業3「処理結果データに可視化処理を施し可視化」、作業4「作業1からやり直す」を実行する際、作業3、作業2、作業1、作業4の順番で分析作業を行う場面を想定した実施例である。

- [0097] 図22は、第三の実施例の前記作業3の手順の一例を示すフローチャートである。

まず、テンプレート推奨部170が指定された基準に従い、テンプレートDB180から入力データ作成部110で実行できるテンプレートを抽出し、リストLvに登録する(S2201)。ここで指定された基準とは、最近利用者100が使ったテンプレート、全利用者がよく使うテンプレート、利用者100以外の特定の利用者がよく使うテンプレート、部署別によく使われるテンプレート、過去この時期によく使われたテンプレート、などを意味する。

- [0098] 次に、抽出したリストLv及びLvに含まれる各テンプレートの可視化例をディスプレイ102に表示する(S2202)。ディスプレイ102への表示例を図3A、図3Eに示す。利用者100がテンプレートを選択した場合、テンプレート推奨部が選択された可視化テンプレートTvの実行に必要な処理結果データを作成できるデータ処理テンプレートを抽出しリストLdに登録する(S2203, S2204)。

- [0099] 図23はS2204の処理を詳述する図である。

テンプレート処理部がテンプレートログからT_vを実行したログを全て抽出しL₁とする。テンプレートログから処理種別が「データ処理用」であるログを全て抽出し、L₂とする。空のリストをL_dとする（S2301）。L₁の要素数をN₁、L₂の要素数をN₂、K₁=0、K₂=0とする（S2302）。L₁のK₁番目のログから入力データ列名リストC₁を取得する（S2303）。L₂のK₂番目のログから出力データ列名リストC₂取得し、C₂⊆C₁である場合、リストL_dにL₂のK₂番目のテンプレートを追加する。この処理をL₁及びL₂の要素全てに実行する（S2303～S2310）。

[0100] 図24は、第三の実施例の前記作業2の手順の一例を示すフローチャートである。

まず、ディスプレイにS2204で作成したデータ処理テンプレートリストL_dを表示する（S2401）。ディスプレイ102への表示例は図3Aに示すものと同様である。利用者100がテンプレートを選択した場合、テンプレート推奨部がテンプレート推奨部が選択されたデータ処理テンプレートT_dの実行に必要な分析対象データを作成できる入力データ作成テンプレートを抽出し、リストL_iに登録する（S2403）。S2403の内容についてはS1504と同様である。

[0101] 図25は、第三の実施例の前記作業1の手順の一例を示すフローチャートである。

まず、ディスプレイにS2403で作成した入力データ作成テンプレートリストL_iを表示する（S2501）。ディスプレイ102への表示例は図3Aに示すものと同様である。利用者100がテンプレートを選択した場合、テンプレート実行部150が入力データ作成部用のテンプレートT_iを実行する（S2502、S2503）。図3Bはテンプレートのデータ構造及びテンプレートの一例を示す図である。ここで示すテンプレートは、入力データ作成部用、データ処理部用、データ可視化部用で同じ構造を想定しているが、各部で実行可能であれば異なる構造であっても構わない。

[0102] 入力データ作成部 110 で実行された入力データ作成処理が正常に終了した場合、その結果作成される分析対象データを D a とする。また、ログ記録部 140 が入力データ作成部 110 で実行した処理内容を記録し、ログ DB 141 に登録する (S 2504、S 2505)。また、実行したテンプレートに関するログも併せて記録し、ログ DB 141 に登録する (S 2505)。図 3 C にログのデータ構造及びログの一例を示す。図 3 D にテンプレートログのデータ構造及びテンプレートログの一例を示す。正常に終了しなかった場合、ディスプレイに問題発生箇所を表示し、終了する (S 2504、S 220)。

[0103] 次に、テンプレート実行部 150 がデータ処理部用のテンプレート T d を実行する (S 2506)。図 3 B はテンプレートのデータ構造及びテンプレートの一例を示す図である。ここで示すテンプレートは、入力データ作成部用、データ処理部用、データ可視化部用で同じ構造を想定しているが、各部で実行可能であれば異なる構造であっても構わない。

[0104] データ処理部 120 で実行された分析処理が正常に終了した場合、その結果作成される処理結果データを D r とする。また、ログ記録部 140 がデータ処理部 120 で実行した処理内容を記録し、ログ DB 141 に登録する (S 2507、S 2508)。また、実行したテンプレートに関するログも併せて記録し、ログ DB 141 に登録する (S 2508)。図 3 C にログのデータ構造及びログの一例を示す。図 3 D にテンプレートログのデータ構造及びテンプレートログの一例を示す。正常に終了しなかった場合、ディスプレイに問題発生箇所を表示し、終了する (S 2507、S 220)。

[0105] 次に、テンプレート実行部 150 がデータ可視化部用のテンプレート T v を実行する (S 2509)。データ可視化部 130 で実行された分析処理が正常に終了した場合、可視化結果をディスプレイ 102 に表示する。また、ログ記録部 140 がデータ可視化部 130 で実行した処理内容を記録し、ログ DB 141 に登録する (S 2510、S 2511)。また、実行したテンプレートに関するログも併せて記録し、ログ DB 141 に登録する (S 25

11)。正常に終了しなかった場合、ディスプレイに問題発生箇所を表示し、終了する(S2510、S220)。

[0106] さらに、可視化結果が利用者100の所望のものではなかった場合、前記作業1～3の処理を改変して再実行する(S2512)。S2512の内容については図26A～図26Dで詳述する。

[0107] 図26A～図26Dは、S2512の内容について詳述する図であり、第三の実施例の前記作業4の手順の一例を示すフローチャートである。

まず、データ可視化部用のテンプレートTvを変更するかどうか、選択する(S2601)。データ可視化部用のテンプレートTvを変更する場合、テンプレート推奨部170が指定された基準に従い、テンプレートDB180からデータ可視化部130で実行できるテンプレートを抽出し、利用者100に提示し、利用者100の選択を受けて、S2603に進む(S2602)。ここで指定された基準とは、最近利用者100が使ったテンプレート、全利用者がよく使うテンプレート、利用者100以外の特定の利用者がよく使うテンプレート、部署別によく使われるテンプレート、過去この時期によく使われたテンプレート、などを意味する。

[0108] 次に、テンプレート処理部160がTvからパラメータを抽出し、ディスプレイにパラメータ一覧を表示する(S2603)。利用者100によりパラメータが選択された場合、ディスプレイにパラメータ編集用のウィンドウを表示し、利用者100により変更されたパラメータ及びその値を取得し、S2701に進む(S2604、S2605)。パラメータ編集用のウィンドウの一例を図3Fに示す。図3Fの302に表示される値の候補は、選択されたパラメータに応じてTdの実行の結果得られるデータDrなどから自動的に入手する。

[0109] 次に、処理結果データDrの列名リストをC0とし、データ可視化部用のテンプレートTvから種別が「N」であるパラメータ及び引数の値を取り出し列名リストC2に登録する(S2701、S2702)。C2 $\subseteq$ C0の場合、S2607に進む(S2703)。C2 $\subseteq$ C0ではない場合、Tdの行

数を R_0 、 $K_0=0$ とする（S2704）。

[0110] T_d の1～ K_0 行目までを実行し、その結果作成されるデータの列名リストを C_0' とし、 $C_2 \subseteq C_0'$ である場合、 T_d の K_0 行目までを T_d とし、S2607に進む。この処理を $K_0=R_0$ になるまで続け、 $K_0=R_0$ になったらS2710に進む（S2705～S2709）。

[0111] 次に、テンプレートログから処理種別が「データ処理用」であるログを全て抽出し、リスト L_1 に登録する（S2710）。 L_1 の要素数を N_1 、 $K_1=0$ とする（S2711）。 L_1 の K_1 番目のログから出力データ列名リスト C_1 取得し、 $C_2 \subseteq C_1$ の場合、 L_1 の K_1 番目のテンプレートを T_i とし、S2607に進む。この処理を $K_1=R_1$ になるまで続け、 $K_1=R_1$ になったらS2717に進む（S2712～S2716）。

[0112] 次に、 L_1 の要素数を N_1 、 $K_1=0$ とする（S2717）。 L_1 の K_1 番目のテンプレートを T_{K_1} とする。 T_{K_1} の行数を R_3 、 $K_3=0$ とする。 T_{K_1} の1～ K_3 行目までを実行し、その結果作成されるデータの列名リストを C_3' とする。 $C_2 \subseteq C_3'$ の場合、 T_i の K_3 行目までを T_i とし、S2607に進む。この処理を T_{K_1} の全行と L_1 の全要素に適用し、 $K_1=N_1$ となった場合、 T_d を空にし、S2607に進む（S2718～S2725）。

[0113] 次に、 T_d が空の場合、ディスプレイに問題発生箇所を表示し、終了する（S2507、S220）。 T_d が空ではない場合、S2001に進む（S2507）。

分析対象データ D_a の列名リストを C_0 とし、データ処理部用のテンプレート T_d から種別が「N」であるパラメータ及び引数の値を取り出し列名リスト C_2 に登録する（S2001、S2002）。 $C_2 \subseteq C_0$ の場合、S2608に進む（S2003）。 $C_2 \subseteq C_0$ ではない場合、 T_i の行数を R_0 、 $K_0=0$ とする（S2004）。

[0114] T_i の1～ K_0 行目までを実行し、その結果作成されるデータの列名リストを C_0' とし、 $C_2 \subseteq C_0'$ である場合、 T_i の K_0 行目までを T_i とし

、S 1 9 0 7に進む。この処理を $K 0 = R 0$ になるまで続け、 $K 0 = R 0$ になったらS 2 0 1 0に進む（S 2 0 0 5～S 2 0 0 9）。

[0115] 次に、テンプレートログから処理種別が「入力データ作成用」であるログを全て抽出し、リストL 1に登録する（S 2 0 1 0）。L 1の要素数をN 1、 $K 1 = 0$ とする（S 2 0 1 1）。L 1のK 1番目のログから出力データ列名リストC 1取得し、 $C 2 \subseteq C 1$ の場合、L 1のK 1番目のテンプレートを T_i とし、S 2 6 0 8に進む。この処理を $K 1 = R 1$ になるまで続け、 $K 1 = R 1$ になったらS 2 0 1 7に進む（S 2 0 1 2～S 2 0 1 6）。

[0116] 次に、L 1の要素数をN 1、 $K 1 = 0$ とする（S 2 0 1 7）。L 1のK 1番目のテンプレートを $T_{K 1}$ とする。 $T_{K 1}$ の行数をR 3、 $K 3 = 0$ とする。 $T_{K 1}$ の1～K 3行目までを実行し、その結果作成されるデータの列名リストを $C 3'$ とする。 $C 2 \subseteq C 3'$ の場合、 T_i のK 3行目までを T_i とし、S 2 6 0 8に進む。この処理を $T_{K 1}$ の全行とL 1の全要素に適用し、 $K 1 = N 1$ となった場合、 T_i を空にし、S 2 6 0 8に進む（S 2 0 1 8～S 2 0 2 5）。

[0117] 次に、 T_i が空の場合、テンプレートによる分析作業の支援は行わず、G U I操作などによる通常の分析作業を実施する（S 2 6 0 8、S 2 6 0 9）。 T_i が空ではない場合、テンプレート実行部1 5 0が入力データ作成部用のテンプレート T_i を実行する（S 2 6 0 8、S 2 6 1 0）。S 2 6 0 9やS 2 6 1 0で指示され、入力データ作成部1 1 0で実行された入力データ作成処理が正常に終了した場合、ログ記録部1 4 0が入力データ作成部1 1 0で実行した処理内容を記録し、ログDB 1 4 1に登録する（S 2 6 1 1、S 2 6 1 2）。また、S 2 6 1 0でテンプレート T_i を実行した場合、実行したテンプレートに関するログも併せて記録し、ログDB 1 4 1に登録する（S 2 6 1 2）。正常に終了しなかった場合、ディスプレイに問題発生箇所を表示し、終了する（S 2 6 1 1、S 2 2 0）。

[0118] 次に、テンプレート実行部1 5 0が T_d を実行する（S 2 6 1 3）。データ処理部1 2 0で実行されたデータ処理が正常に終了した場合、ログ記録部

140がデータ処理部120で実行した処理内容を記録し、ログDB141に登録する(S2614、S2615)。また、実行したテンプレートTdに関するログも併せて記録し、ログDB141に登録する(S2615)。正常に終了しなかった場合、ディスプレイに問題発生箇所を表示し、終了する(S2614、S220)。

[0119] 次に、テンプレート実行部150がTvを実行する(S2616)。データ可視化部130で実行された可視化処理が正常に終了した場合、可視化結果をディスプレイに表示する。またログ記録部140がデータ可視化部130で実行した処理内容を記録し、ログDB141に登録する(S2617、S2618)。また、実行したテンプレートTvに関するログも併せて記録し、ログDB141に登録する(S2618)。正常に終了しなかった場合、ディスプレイに問題発生箇所を表示し、終了する(S2617、S220)。

[0120] 図27A、図27Bは、第三の実施例の前記作業4の具体例を示す図である。具体例としては、図4、図7、図10の作業1～作業3の分析処理では、「指定した列を棒グラフで表示し、指定した列を横軸の項目名とする」として可視化を行ったが、図27A、図27Bの例ではS2602～S2606でデータ可視化部用のテンプレートTvを「指定した列を円グラフで表示し、指定した列を項目名とする」を選択し、パラメータ1を「要素の出現数」、パラメータ2を「KEY」とし、テンプレートTi、Tdを用いて入力データ作成処理、データ処理を実行する場面を想定する。

[0121] 図27Aは、テンプレートTvの例である。この例でのTvのパラメータは2つあり、パラメータ1は、種別がN(列名)で値が「要素の出現数」、パラメータ2は、種別がN(列名)で値が「KEY」である。Tvで行う処理はパラメータ1、2を引数として、「指定した列(パラメータ1)を円グラフで表示し、指定した列(パラメータ2)を項目名とする」である。図21Bは「円グラフ」のイメージを示す図である。テンプレートTi及びTdによる処理それぞれ図4、図7と同様のため、ここでは割愛する。

[0122] 第三の実施例によれば、分析作業の可視化処理の変更が可能となり、また可視化処理の変更に伴う、入力データ作成処理、データ処理の変更もテンプレートを用いて支援することが可能となるため、利用者100の分析作業の手間を削減することが可能となる。

実施例 4

[0123] 第四の実施例では、利用者100が作業1「DB104から必要なデータを抽出し、抽出したデータに入力データ作成処理を施すことにより分析対象データを作成」、作業2「分析対象データにデータ処理を施し処理結果データを作成」、作業3「処理結果データに可視化処理を施し可視化」、作業4「作業1からやり直す」を実行する際、作業3、作業1、作業2、作業4の順番で分析作業を行う場面を想定した実施例である。

[0124] 図28A、図28Bは第四の実施例の手順の一例を示すフローチャートである。

まず、テンプレート推奨部170が指定された基準に従い、テンプレートDB180からデータ可視化部130で実行できるテンプレートを抽出し、リストLvに登録する(S2801)。ここで指定された基準とは、最近利用者100が使ったテンプレート、全利用者がよく使うテンプレート、利用者100以外の特定の利用者がよく使うテンプレート、部署別によく使われるテンプレート、過去この時期によく使われたテンプレート、などを意味する。

[0125] 次に、抽出したリストLv及びLvに含まれる各テンプレートの可視化例をディスプレイ102に表示する(S2802)。ディスプレイ102への表示例を図3A、図3Eに示す。利用者100がテンプレートを選択した場合、テンプレート推奨部がテンプレートログから選択された可視化テンプレートTvの直前に実行した入力データ作成テンプレートを全て抽出し、リストLiに登録し、ディスプレイにLiを表示する(S2803～S2805)。入力データ作成用テンプレートTiが選択された場合、Tiを実行する(S2806、S2807)。テンプレート推奨部170がテンプレートロ

グからT_iとT_vの間にあるデータ処理テンプレートを全て抽出し、リストL_dに登録し、ディスプレイにL_dを表示する(S2808、S2809)。

[0126] データ処理用テンプレートT_dが選択された場合、T_dのパラメータを変更するかどうか選択する(S2810、S2811)。パラメータを変更する場合、テンプレート処理部170がT_dからパラメータを抽出し、ディスプレイにパラメーター一覧を表示する(S2812)。パラメータが選択された場合、ディスプレイにパラメータ編集用のウィンドウを表示する(S2813、S2814)。パラメータ編集が終了した場合、T_dを実行する(S2815、S2816)。パラメータ変更をしない場合、T_dを実行する(S2816)。パラメータ編集用のウィンドウの例は図3Fに示す。図3Fの302に表示される値の候補は、選択されたパラメータに応じてT_iの実行の結果得られるデータなどから自動的に入手する。

[0127] 次に、T_dのパラメータ変更や可視化のための要素の変更のために、T_vのパラメータを変更する。テンプレート処理部170がT_vからパラメータを抽出し、ディスプレイにパラメーター一覧を表示する(S2817)。パラメータが選択された場合、ディスプレイにパラメータ編集用のウィンドウを表示する(S2818)。パラメータ編集用のウィンドウの例は図3Fに示す。図3Fの302に表示される値の候補は、選択されたパラメータに応じてT_dの実行の結果得られるデータなどから自動的に入手する。

[0128] パラメータ編集が終了した場合T_vを実行し、可視化結果をディスプレイに表示し、ログ記録部140が入力データ作成部110、データ処理部120、データ可視化部130でそれぞれ実行した処理内容を記録し、ログDB141に登録する(S2820～S2822)。

[0129] S2804～S2806では、入力データ作成用テンプレートを選択しているが、図2のS205のように、テンプレートによる分析作業の支援は行わず、GUI操作などによる通常の分析作業を実施しても構わない。

[0130] S2806で入力データ作成用テンプレートT_iを選択した後、T_iのパ

ラメータ変更が可能であるが、その場合、それ以降の処理は実施例 1 と同様となる。

[0131] S 2 8 0 7 ~ S 2 8 0 9 では、データ処理用テンプレートを選択しているが、図 6 の S 6 0 4 のように、テンプレートによる分析作業の支援は行わず、G U I 操作などによる通常の分析作業を実施しても構わない。

[0132] 第四の実施例によれば、分析作業の可視化結果から分析作業の開始が可能となる。つまり可視化結果を実現するために必要な入力データ作成処理、データ処理の構築をテンプレートを利用して支援することが可能となるため、利用者 1 0 0 の分析作業の手間を削減することが可能となる。

実施例 5

[0133] 第五の実施例では、利用者 1 0 0 が作業 1 「D B 1 0 4 から必要なデータを抽出し、抽出したデータに入力データ作成処理を施すことにより分析対象データを作成」、作業 2 「分析対象データにデータ処理を施し処理結果データを作成」、作業 3 「処理結果データに可視化処理を施し可視化」、作業 4 「作業 1 からやり直す」という作業を、作業 1、作業 3、作業 2、作業 4 の順番で分析作業を行う場面を想定した実施例である。

[0134] 図 2 9 は第五の実施例の手順の一例を示すフローチャートである。

まず、テンプレート推奨部 1 7 0 が指定された基準に従い、テンプレート D B 1 8 0 から入力データ作成部 1 1 0 で実行できるテンプレートを抽出し、リスト L i に登録する。また、テンプレートログから L i に含まれる入力データ作成テンプレートの直後に実行されているデータ可視化テンプレートを全て抽出し、リスト L e に登録する (S 2 9 0 1)。ここで指定された基準とは、最近利用者 1 0 0 が使ったテンプレート、全利用者がよく使うテンプレート、利用者 1 0 0 以外の特定の利用者がよく使うテンプレート、部署別によく使われるテンプレート、過去この時期によく使われたテンプレート、などを意味する。

[0135] 次に、抽出したリスト L i をディスプレイ 1 0 2 に表示する。このとき、ユーザの操作により L i からテンプレートが選択された場合、選択されたテ

ンプレートに対応するデータ可視化テンプレートの可視化結果図をL eから抽出し、ディスプレイに表示する（S 2 9 0 2）。ディスプレイ1 0 2への表示例を図3 0に示す。3 0 0 1はL iの一覧、3 0 0 2には、3 0 0 1に表示されているテンプレート名を選択すると過去にこの入力データ作成テンプレートを実行した直後に実行されたデータ可視化テンプレートの可視化結果が表示される。

- [0136] 利用者1 0 0がテンプレートを選択した場合、選択されたテンプレートT iを実行する（S 2 9 0 3, S 2 9 0 4）。テンプレート推奨部がテンプレートログから選択された入力データ作成テンプレートT iの直後に実行したデータ可視化テンプレートを全て抽出し、リストL vに登録し、ディスプレイにL vを表示する（S 2 9 0 5, S 2 9 0 6）。データ可視化テンプレートT vが選択された場合、テンプレート推奨部1 7 0がテンプレートログからT iとT vの間にあるデータ処理テンプレートを全て抽出し、リストL dに登録し、ディスプレイにL dを表示する（S 2 9 0 7、図2 8 BのS 2 8 0 8, S 2 8 0 9）。

これ以降の処理手順は、図2 8 Bで説明済みであるため、割愛する。

- [0137] 第五の実施例によれば、分析作業を開始する際に、分析の対象としたいデータと、そのデータを用いて可視化可能な図を提示することにより、ユーザの分析結果イメージを実現する分析の実施が可能となる。つまり可視化結果を実現するために必要な入力データ作成処理、データ処理の構築をテンプレートを利用して支援することが可能となるため、利用者1 0 0の分析作業の手間を削減することが可能となる。

- [0138] なお、本発明は上記した実施例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、上記した実施例は本発明のより良い理解のために詳細に説明したのであり、必ずしも説明の全ての構成を備えるものに限定されるものではない。

- [0139] また、ある実施例の構成の一部を他の実施例の構成に置き換えることが可能であり、また、ある実施例の構成に他の実施例の構成を加えることが可能

である。また、各実施例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。

[0140] 更に、上述した各構成、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良いし、それらの一部又は全部を実現するプログラムを作成することによりソフトウェアで実現しても良いことは言うまでもない。

符号の説明

- [0141] 1 0 0 利用者
1 0 1 計算機
1 0 2 ディスプレイ
1 0 3 入力装置
1 0 4 データベース (DB)
1 1 0 入力データ作成部
1 2 0 データ処理部
1 3 0 データ可視化部
1 4 0 ログ記録部
1 4 1 ログDB
1 5 0 テンプレート実行部
1 6 0 テンプレート処理部
1 7 0 テンプレート推奨部
1 8 0 テンプレートDB
3 0 1 パラメータ名
3 0 2 値の候補
3 0 3 選択された値
3 1 0、3 0 0 1 テンプレートリスト一覧
3 1 1、3 0 0 2 可視化結果

請求の範囲

- [請求項1] 計算機を用いたデータ分析支援方法であって、
前記計算機は、
データに入力データ作成処理を施すことにより分析対象データを作成する入力データ作成機能と、データにデータ処理を施し処理結果データを作成するデータ処理機能と、データに可視化処理を施し可視化するデータ可視化機能を実行することによりデータ分析支援の作業を行う
ことを特徴とするデータ分析支援方法。
- [請求項2] 請求項1に記載のデータ分析支援方法であって、
前記計算機は、
前記入力データ作成機能、前記データ処理機能、或いは前記データ可視化機能を実行するためのテンプレートを蓄積するテンプレートデータベースを利用し、
利用者に対して次作業についてのテンプレートを推奨するテンプレート推奨機能と、前記テンプレートを実行するテンプレート実行機能を実行する
ことを特徴とするデータ分析支援方法。
- [請求項3] 請求項2に記載のデータ分析支援方法であって、
前記計算機は
利用者が部分的に変更した前記テンプレートに対して、前記テンプレート推奨機能を用いて、前記テンプレートデータベースに蓄積されたテンプレートの中から、適用可能なテンプレートを推奨する
ことを特徴とするデータ分析支援方法。
- [請求項4] 請求項2に記載のデータ分析支援方法であって、
前記計算機は、
前記テンプレート実行機能により、前記入力データ作成機能、前記データ処理機能、前記データ可視化機能を用いて、前記テンプレートを

実行する

ことを特徴とするデータ分析支援方法。

[請求項5]

請求項2に記載のデータ分析支援方法であって、

前記計算機は、

指定された基準に従い、前記テンプレート推奨機能を実行して得られた入力データ作成テンプレートのリストから、選択された前記入力データ作成テンプレートを、前記テンプレート実行機能により実行することを特徴とするデータ分析支援方法。

[請求項6]

請求項2に記載のデータ分析支援方法であって、

前記計算機は、

指定された基準に従い、前記テンプレート推奨機能を実行して得られたデータ処理テンプレートのリストから、選択された前記データ処理テンプレートの実行に必要な前記分析対象データを作成できる、前記入力データ作成テンプレートを抽出することを特徴とするデータ分析支援方法。

[請求項7]

請求項2に記載のデータ分析支援方法であって、

前記計算機は、

指定された基準に従い、前記テンプレート推奨機能を実行して得られたデータ可視化テンプレートリストから、選択された前記データ可視化テンプレートの実行に必要な前記処理結果データを作成できる、前記データ処理テンプレートを抽出することを特徴とするデータ分析支援方法。

[請求項8]

請求項2に記載のデータ分析支援方法であって、

前記計算機は、

指定された基準に従い、前記テンプレート推奨機能を実行して得られたデータ可視化テンプレートリストから、選択された前記データ可視化テンプレートの直前に実行された入力データ作成テンプレートを抽出して得られた入力データ作成テンプレートリストから、選択された

前記入力データ作成テンプレートを、前記テンプレート実行機能により実行する

ことを特徴とするデータ分析支援方法。

[請求項9]

請求項2に記載のデータ分析支援方法であって、

前記計算機は、

指定された基準に従い、前記テンプレート推奨機能を実行して抽出した入力データ作成テンプレートリストから、選択された前記入力データ作成テンプレートの直後に実行されたデータ可視化テンプレートを抽出して得られたデータ可視化テンプレートリストから、選択された前記データ可視化テンプレートと、選択された前記入力データ作成テンプレートとの間で実行されたデータ処理テンプレートを抽出して得られるデータ処理テンプレートリストから、選択された前記データ処理テンプレートを、前記テンプレート実行機能により実行する

ことを特徴とするデータ分析支援方法。

[請求項10]

データ分析支援システムであって、

複数のテンプレートを蓄積するテンプレートデータベースを記憶する記憶部と、

前記テンプレートを実行する処理部とを備え、

前記処理部は、

データに入力データ作成処理を施すことにより分析対象データを作成する入力データ作成機能と、データにデータ処理を施し処理結果データを作成するデータ処理機能と、データに可視化処理を施し可視化するデータ可視化機能を実行することによりデータ分析支援の作業を行う

ことを特徴とするデータ分析支援システム。

[請求項11]

請求項10に記載のデータ分析支援システムであって、

前記処理部は、

利用者に対して次作業についてのテンプレートを推奨するテンプレ-

ト推奨機能を備え、

前記テンプレート推奨機能の実行の結果、選択された前記テンプレートを実行する

ことを特徴とするデータ分析支援システム。

[請求項12]

請求項 1 1 に記載のデータ分析支援システムであって、

前記処理部は

利用者が部分的に変更した前記テンプレートに対して、前記テンプレート推奨機能により、前記テンプレートデータベースに蓄積されたテンプレートの中から、適用可能なテンプレートを推奨する

ことを特徴とするデータ分析支援システム。

[請求項13]

請求項 1 1 に記載のデータ分析支援システムであって、

前記処理部は、

前記入力データ作成機能、前記データ処理機能、前記データ可視化機能を用いて、選択された前記テンプレートを実行する

ことを特徴とするデータ分析支援システム。

[請求項14]

請求項 1 3 に記載のデータ分析支援システムであって、

表示部を更に備え、

前記処理部は、

指定された基準に従い、前記テンプレート推奨機能を実行して得られた前記テンプレートを前記表示部に表示する

ことを特徴とするデータ分析支援システム。

[請求項15]

請求項 1 4 に記載のデータ分析支援システムであって、

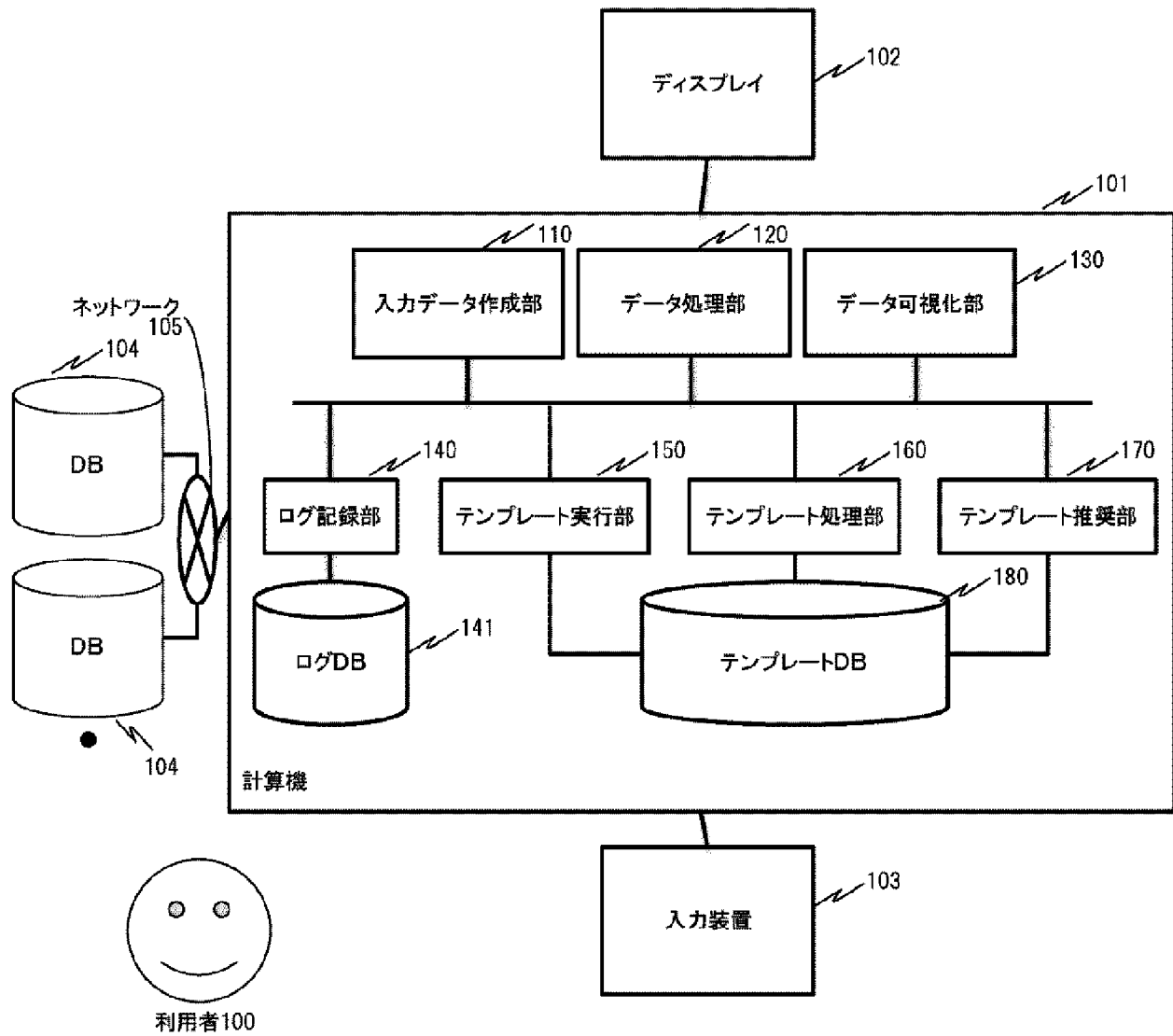
前記処理部は、

前記表示部に表示された前記テンプレートから、選択された前記テンプレートを前記データ可視化機能で実行した得られた図例を前記表示部に同時に表示する

ことを特徴とするデータ分析支援システム。

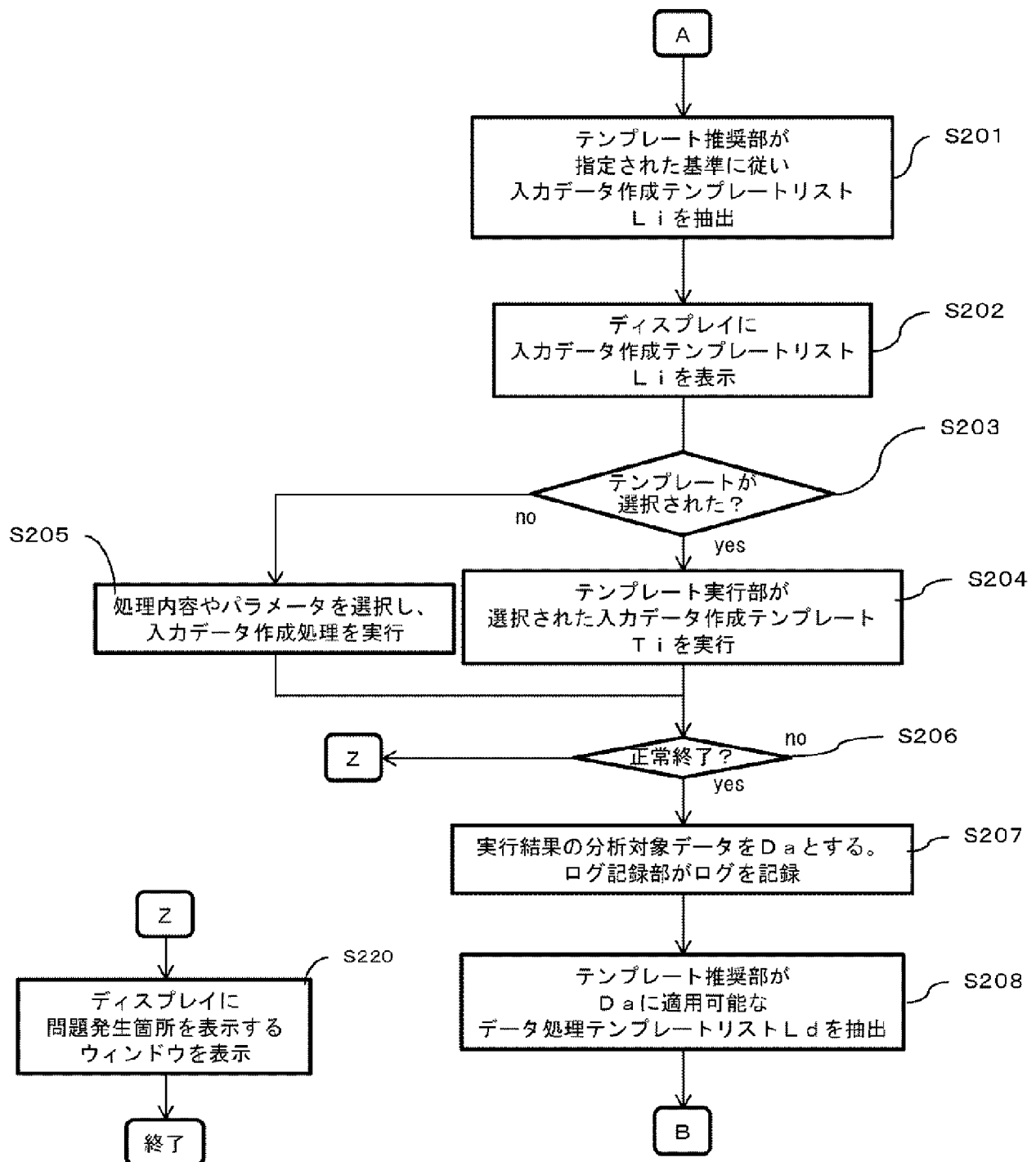
[図1]

図1



[図2]

図2



[図3A]

図 3 A

推奨テンプレートの一例

#	テンプレート名	✓ Li
1	テンプレートA	
2	テンプレートC	
3	テンプレートF	
	...	

[図3B]

図 3 B

テンプレートのデータ構造

項目名	内容
パラメータ	実行時に与える設定値。 実行時に引数の該当するパラメータを設定値に置換する。 パラメータは「種別」と「値」からなる。 種別とは、パラメータ値が何に関連する値なのかを指す値であり、 下記6種類からなる： T：テーブル名、F：ファイル名、N：列名、V：データの要素、 C：事前定義済み定数、E：その他
処理名	実行した処理を特定する名前またはID
引数	処理実行時に与えた引数（複数）



テンプレートの一例

パラメータ1=[種別], [テンプレートでは空白、実行時に別の値に置換する]
 パラメータ2=[種別], [テンプレートでは空白、実行時に別の値に置換する]

処理名1	<u>パラメータ1</u>	[種別], 引数2 ...
処理名2	<u>パラメータ1</u>	[種別], 引数2 ...
処理名3	<u>パラメータ1</u>	[種別], 引数2 ...
処理名4	[種別], 引数B1	[種別], 引数2 ...
処理名5	<u>パラメータ2</u>	[種別], 引数2 ...
処理名6	[種別], 引数D1	<u>パラメータ2</u> ...
...		

[図3C]

図 3 C

ログ構造

項目名	内容
データ列名	処理実行時の分析対象データまたは処理結果データの列名
時刻	処理実行時の時刻
処理名	実行した処理を特定する名前または I D
引数	処理実行時に与えた引数（複数）



ログの一例

データ列名=データ列名1	データ列名2	...
2012/03/01 0:00:00	処理名1	[種別], 引数A1 [種別], 引数A2 ...
2012/03/01 0:01:00	処理名2	[種別], 引数A1 [種別], 引数B2 ...
2012/03/01 0:01:00	処理名3	[種別], 引数A1 [種別], 引数C2 ...
2012/03/01 0:02:00	処理名4	[種別], 引数B1 [種別], 引数D2 ...
2012/03/01 0:02:00	処理名5	[種別], 引数C1 [種別], 引数E2 ...
2012/03/01 0:05:00	処理名6	[種別], 引数D1 [種別], 引数C1 ...
...		

[図3D]

図 3 D

テンプレート構造

項目名	内容
時刻	処理実行時の時刻
処理種別	入力データ作成用/データ処理用/可視化用の各種別
テンプレート名	実行したテンプレート名
パラメータ値	実行時に用いたパラメータ値
入力データ列名	テンプレート実行時に入力する分析対象データまたは処理結果データの列名一覧
出力データ列名	テンプレート実行後の分析対象データまたは処理結果データの列名一覧
可視化結果図	可視化用の場合、可視化結果図を保存したファイルのパス

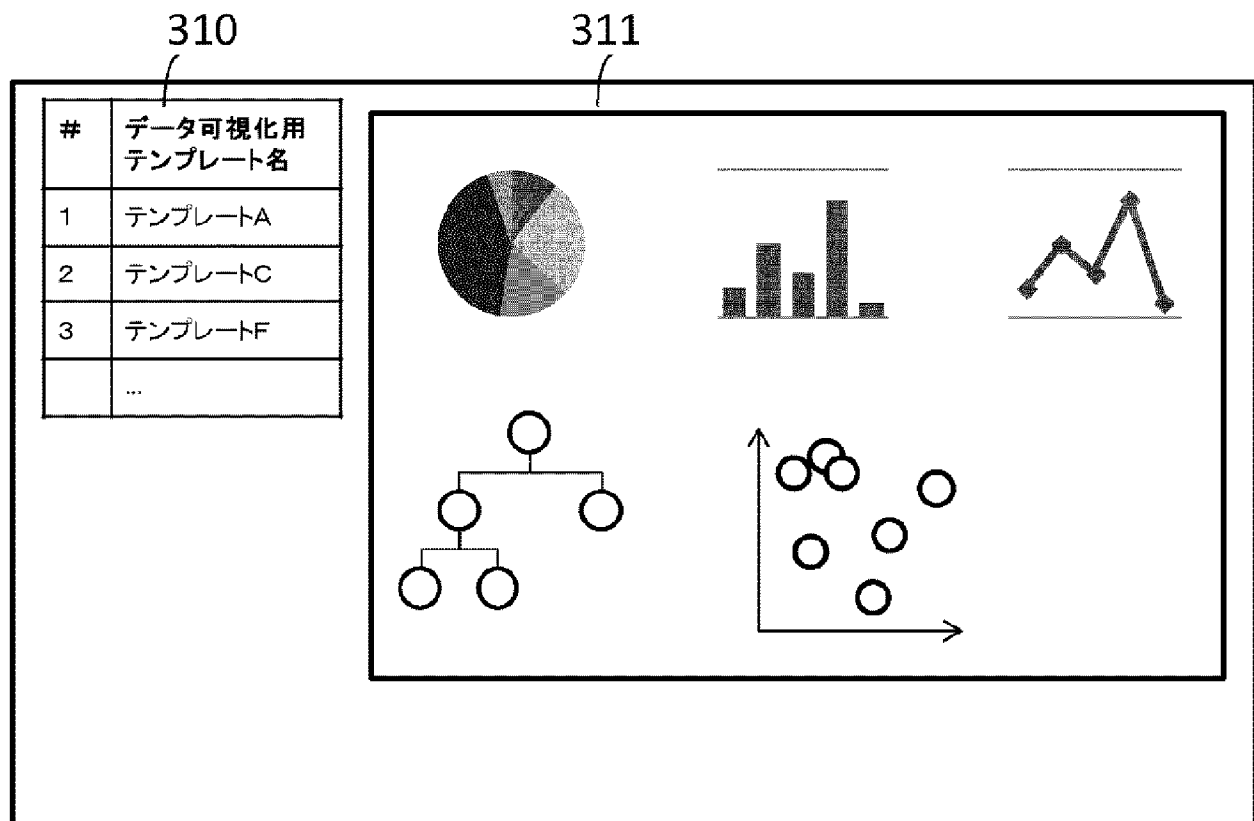


テンプレートの一例

2012/03/01 0:00:00	入力データ作成用	テンプレート名1	パラメータ値1...	入力データ列名1...	出力データ列名1...
2012/03/01 1:01:00	データ処理用	テンプレート名2	パラメータ値2...	入力データ列名2...	出力データ列名2...
...					

[図3E]

図 3 E



[図3F]

図 3 F

301 302 303

パラメータ1

OK Cancel

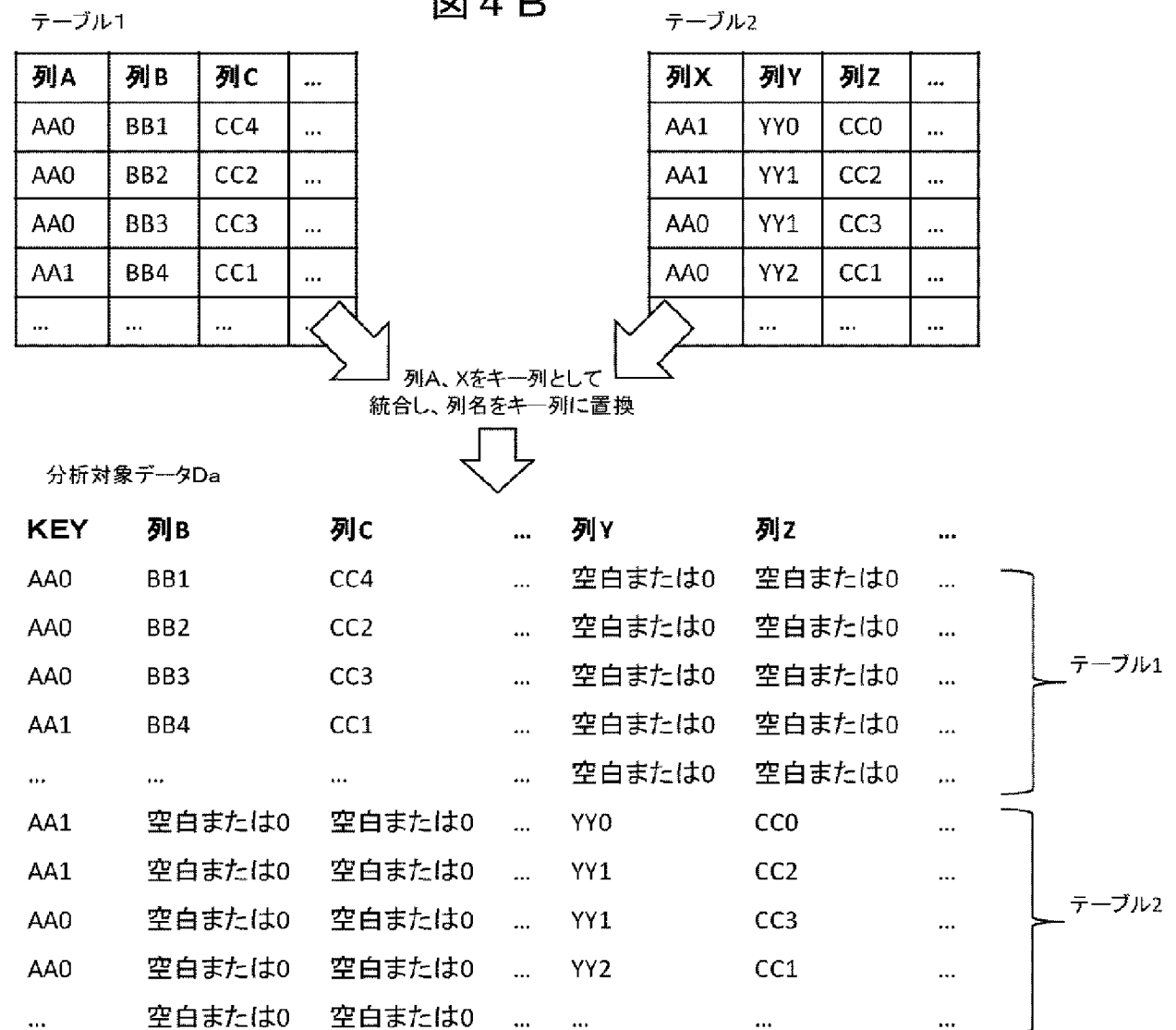
[図4A]

図 4 A

パラメータ 1=E	“テーブル1 “
パラメータ 2=E	“テーブル2”
パラメータ 3=N	“列 A”
パラメータ 4=N	“列 X”
パラメータ 5=N	“KEY”
“テーブル統合” パラメータ 1 パラメータ 2 パラメータ 3 パラメータ 4 パラメータ 5	

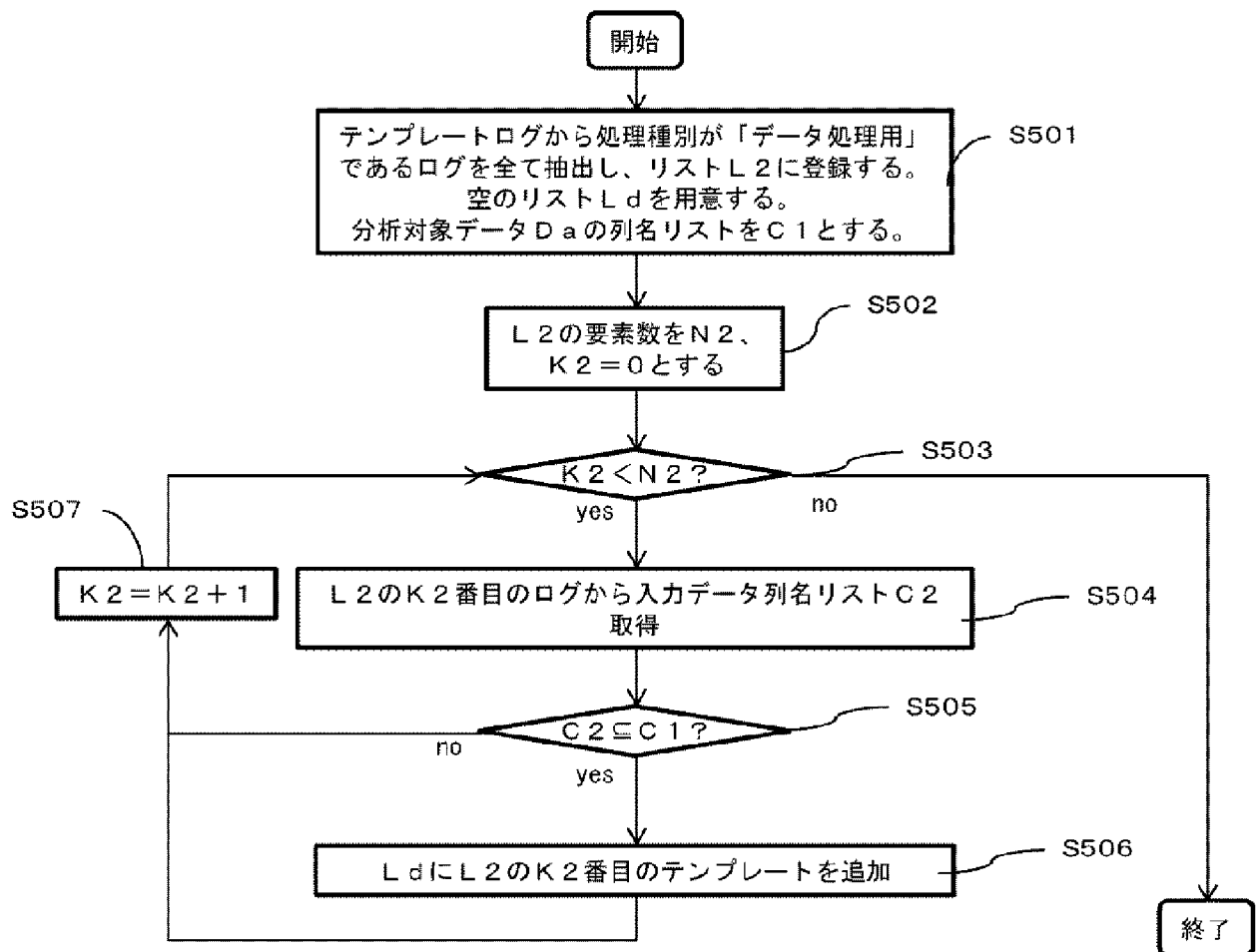
[図4B]

図 4 B



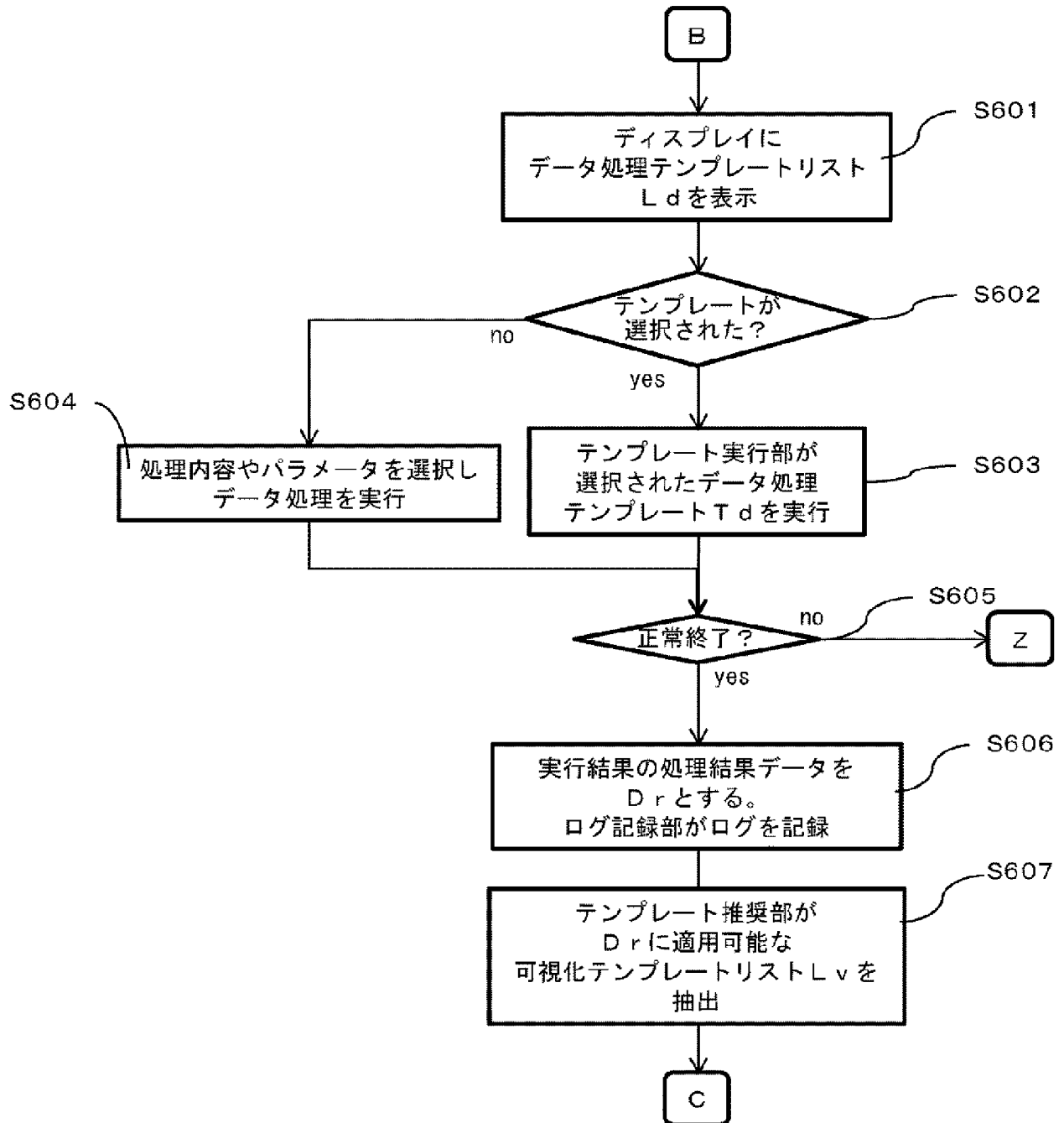
[図5]

図 5



[図6]

図6



[図7A]

図 7 A

パラメータ1=N “KEY”

“要素の出現数を数える”

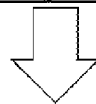
パラメータ1

[図7B]

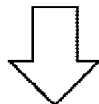
図 7 B

分析対象データ D a

AA0	BB1	CC4	...	空白または0	空白または0	...
AA0	BB2	CC2	...	空白または0	空白または0	...
AA0	BB3	CC3	...	空白または0	空白または0	...
AA1	BB4	CC1	...	空白または0	空白または0	...
...	...	...	...	空白または0	空白または0	...
AA1	空白または0	空白または0	...	YY0	CC0	...
AA1	空白または0	空白または0	...	YY1	CC2	...
AA0	空白または0	空白または0	...	YY1	CC3	...
AA0	空白または0	空白または0	...	YY2	CC1	...
...	空白または0	空白または0	...	...	...	...



列Aの要素の出現数を数える

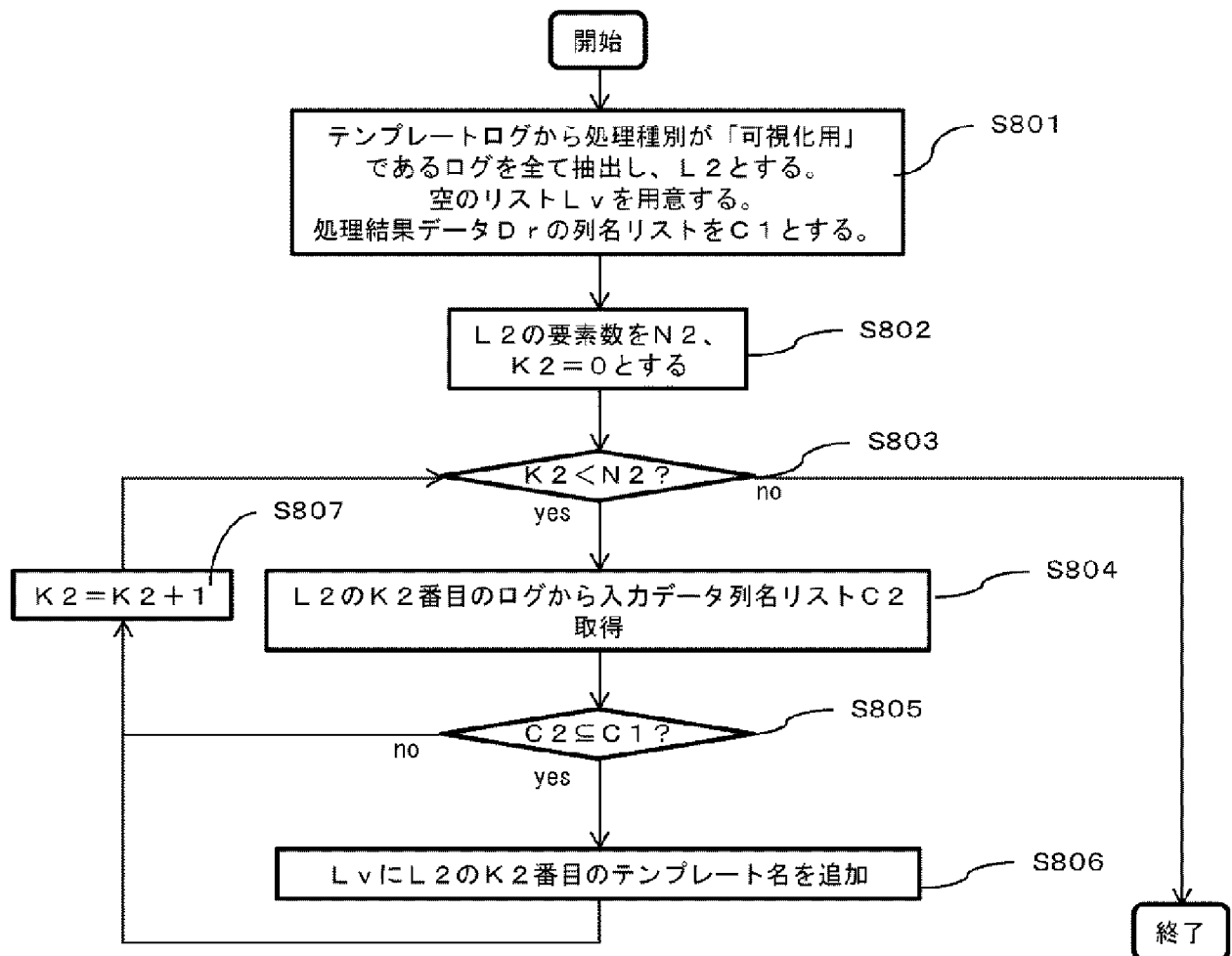


分析対象データ D r

KEY	要素の出現数
AA0	5
AA1	3
AA2	0
AA3	0

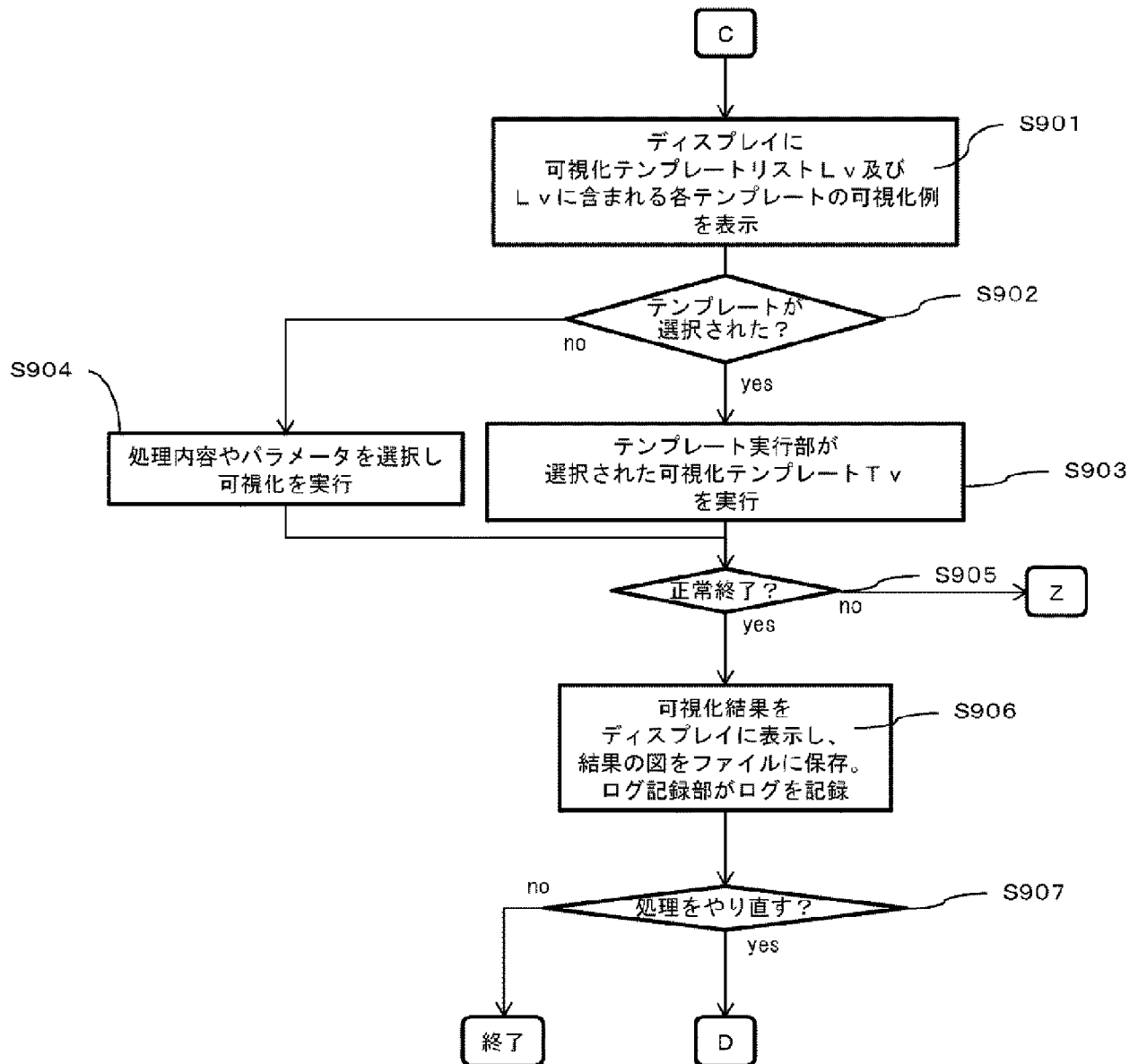
[図8]

図 8



[図9]

図9



[図10A]

図 10 A

パラメータ1=N “要素の出現数”

パラメータ2=N “KEY”

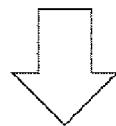
“棒グラフ” パラメータ1 パラメータ2

[図10B]

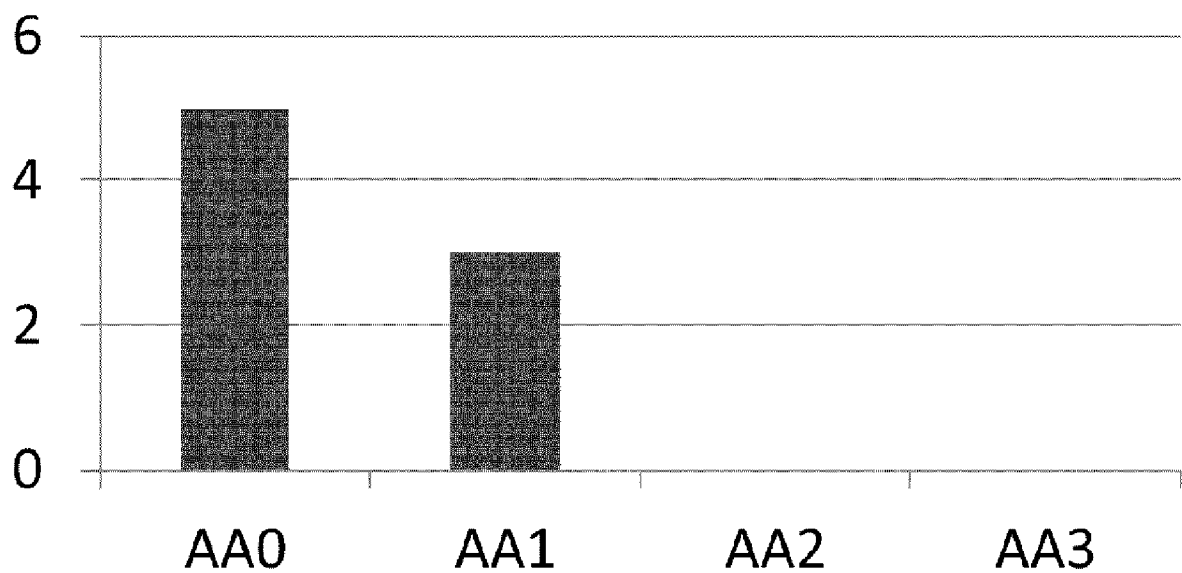
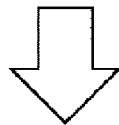
図 1 0 B

分析対象データ D_r

KEY	要素の出現数
AA0	5
AA1	3
AA2	0
AA3	0

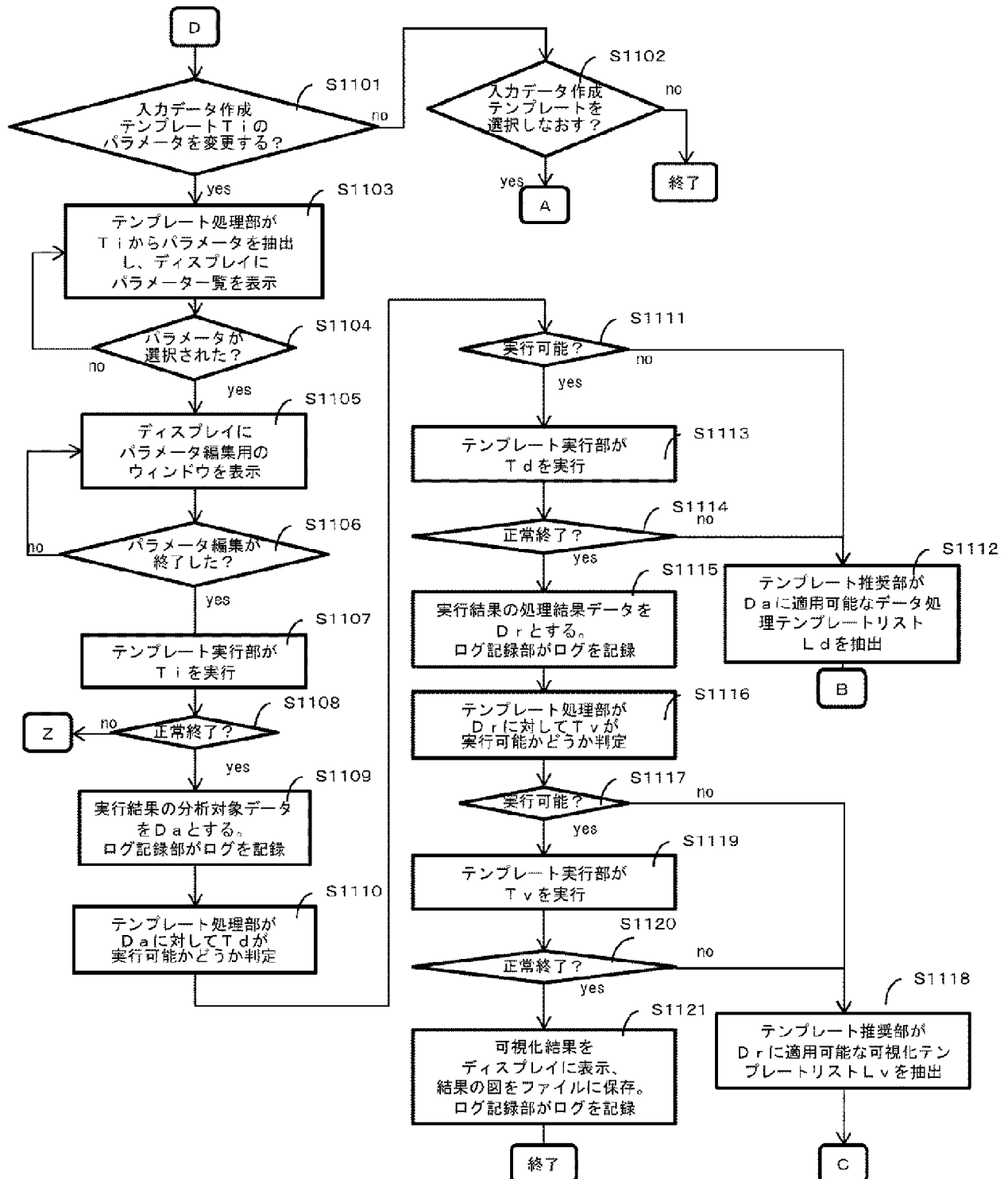


列「要素の出現数」を棒グラフで表示し、列「KEY」を横軸の項目名とする



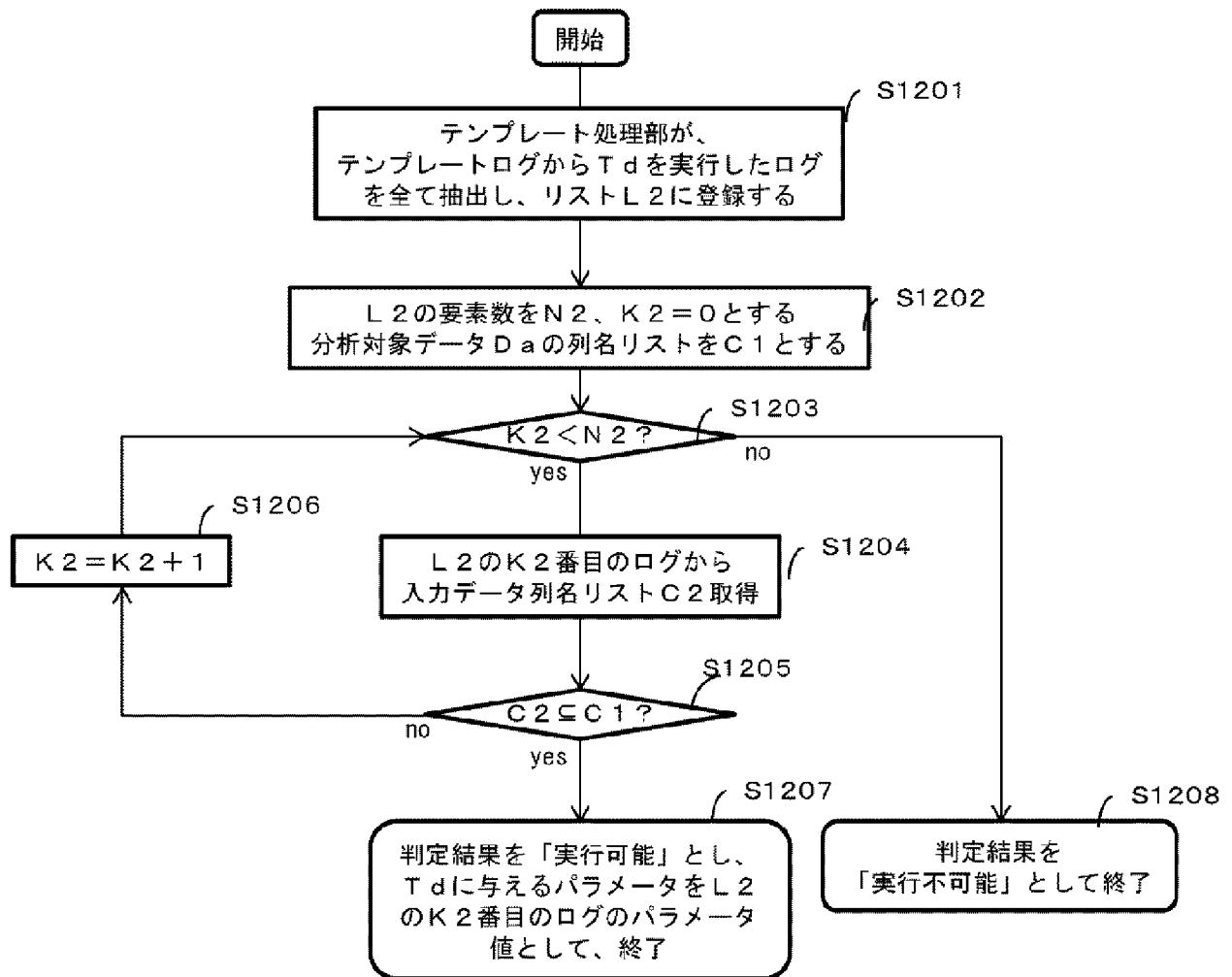
[図11]

図 11



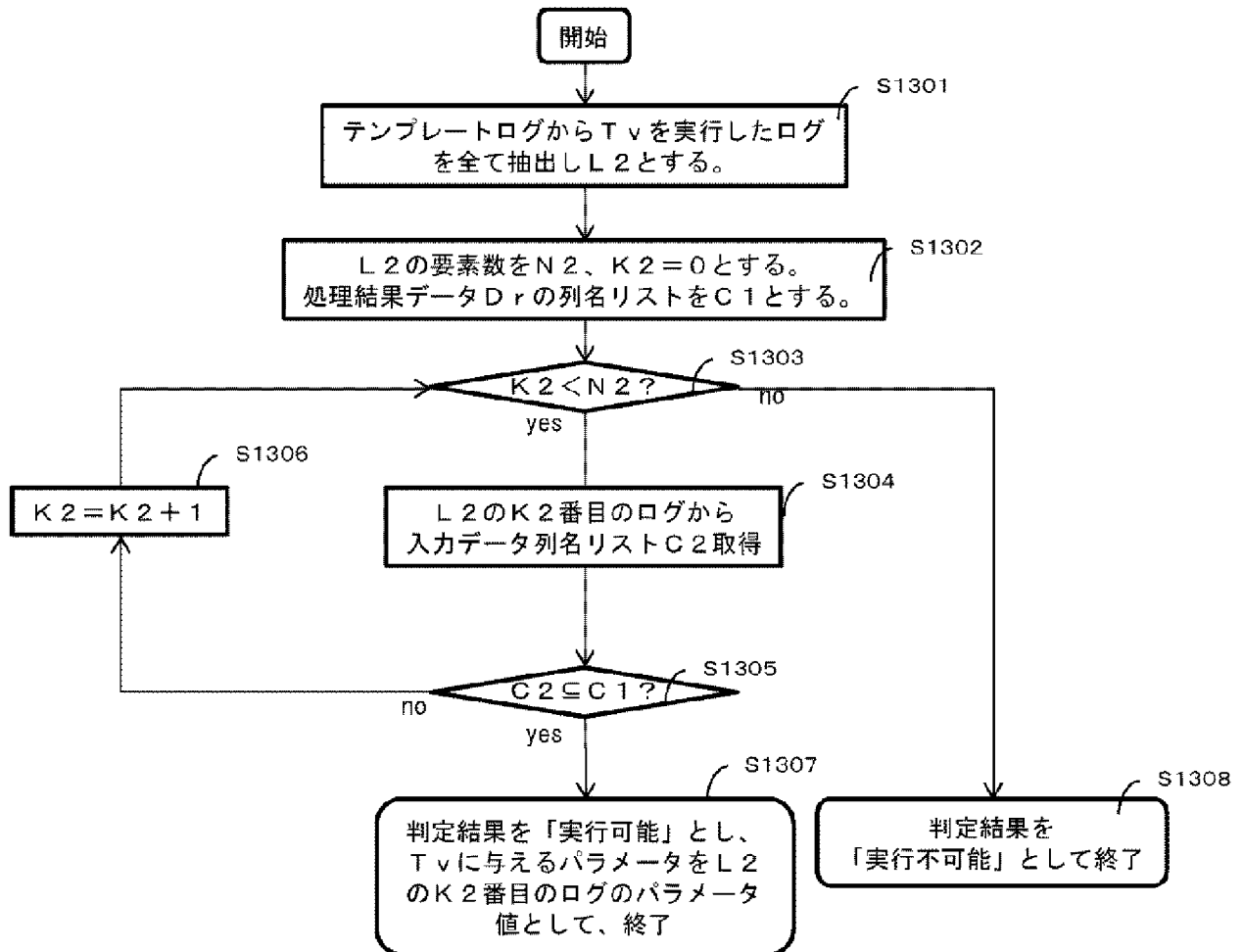
[図12]

図 1 2



[図13]

図 1 3



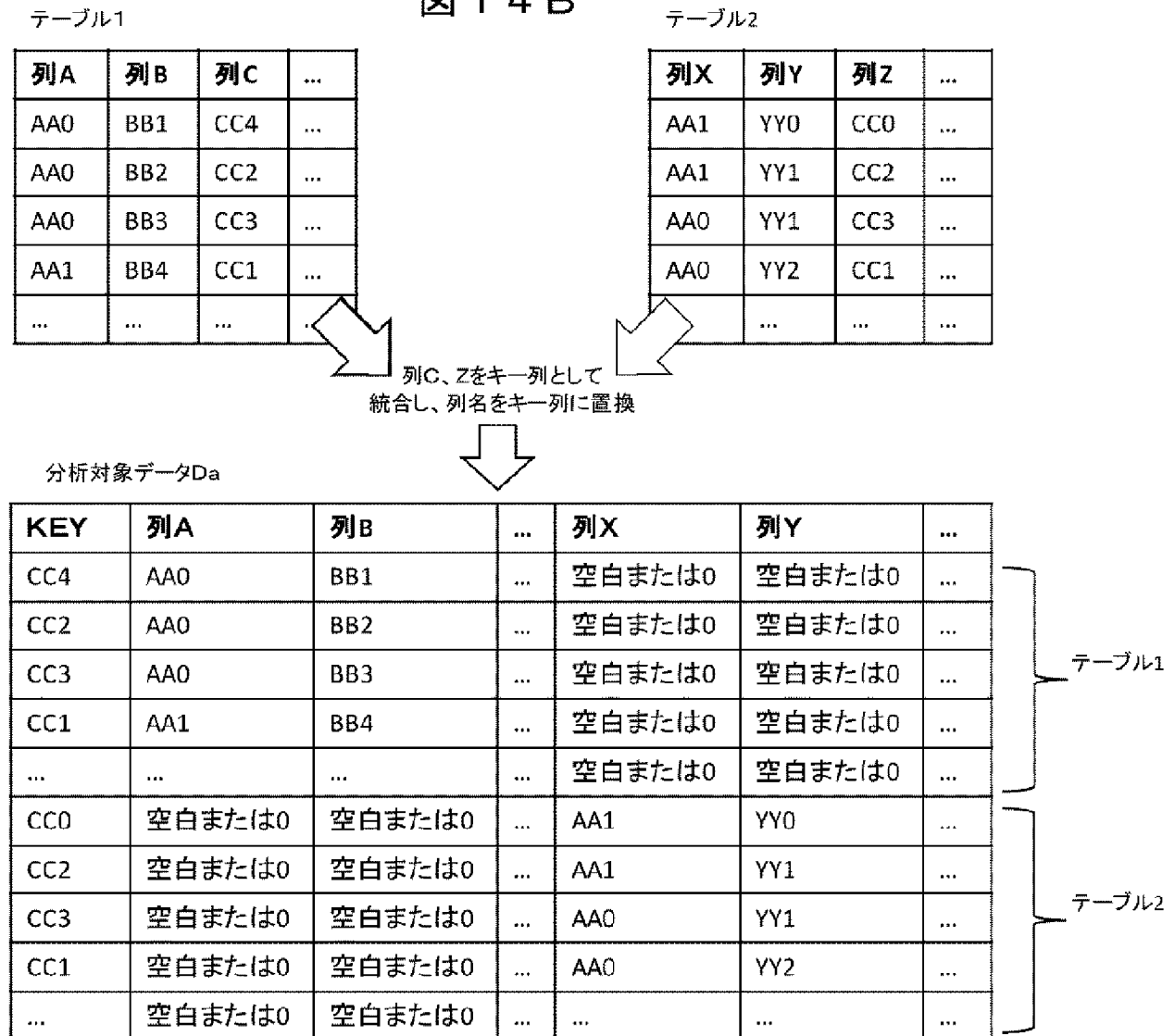
[図14A]

図 1 4 A

パラメータ 1=E	“テーブル1 “
パラメータ 2=E	“テーブル2”
パラメータ 3=N	“列C”
パラメータ 4=N	“列Z”
パラメータ 5=N	“KEY”
“テーブル統合”	
パラメータ 1	パラメータ 2
パラメータ 3	パラメータ 4
パラメータ 5	

[図14B]

図 1 4 B



[図14C]

図 1 4 C

選択テンプレート T d を実行した結果の処理結果データ D r の例
T d の内容が「指定した列の要素の出現数を数える」の場合

パラメータ1=N “KEY”

“要素の出現数を数える”

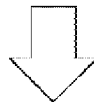
パラメータ1

[図14D]

図 1 4 D

分析対象データ D_a

KEY	列A	列B	...	列X	列Y	...
CC4	AA0	BB1	...	空白または0	空白または0	...
CC2	AA0	BB2	...	空白または0	空白または0	...
CC3	AA0	BB3	...	空白または0	空白または0	...
CC1	AA1	BB4	...	空白または0	空白または0	...
...	...	...	...	空白または0	空白または0	...
CC0	空白または0	空白または0	...	AA1	YY0	...
CC2	空白または0	空白または0	...	AA1	YY1	...
CC3	空白または0	空白または0	...	AA0	YY1	...
CC1	空白または0	空白または0	...	AA0	YY2	...
...	空白または0	空白または0	...	...	...	...



列Aの要素の出現数を数える

分析対象データ D_r

KEY	要素の出現数
CC0	2
CC1	5
CC2	3
CC3	8
CC4	1

[図14E]

図 1 4 E

選択テンプレート T_v を実行した結果の可視化の例T_v の内容が「指定した列を棒グラフで表示し、指定した列を横軸の項目名とする」の場合

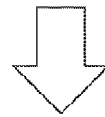
パラメータ1=N “要素の出現数”
 パラメータ2=N “KEY”
 “棒グラフ” パラメータ1 パラメータ2

[図14F]

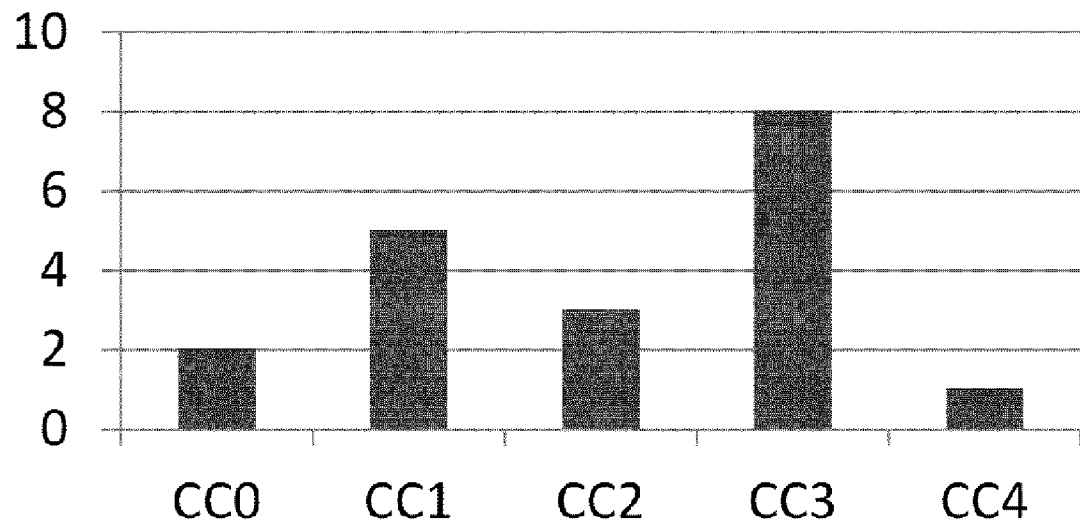
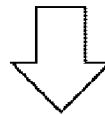
図 1 4 F

分析対象データ D r

KEY	要素の出現数
CC0	2
CC1	5
CC2	3
CC3	8
CC4	1

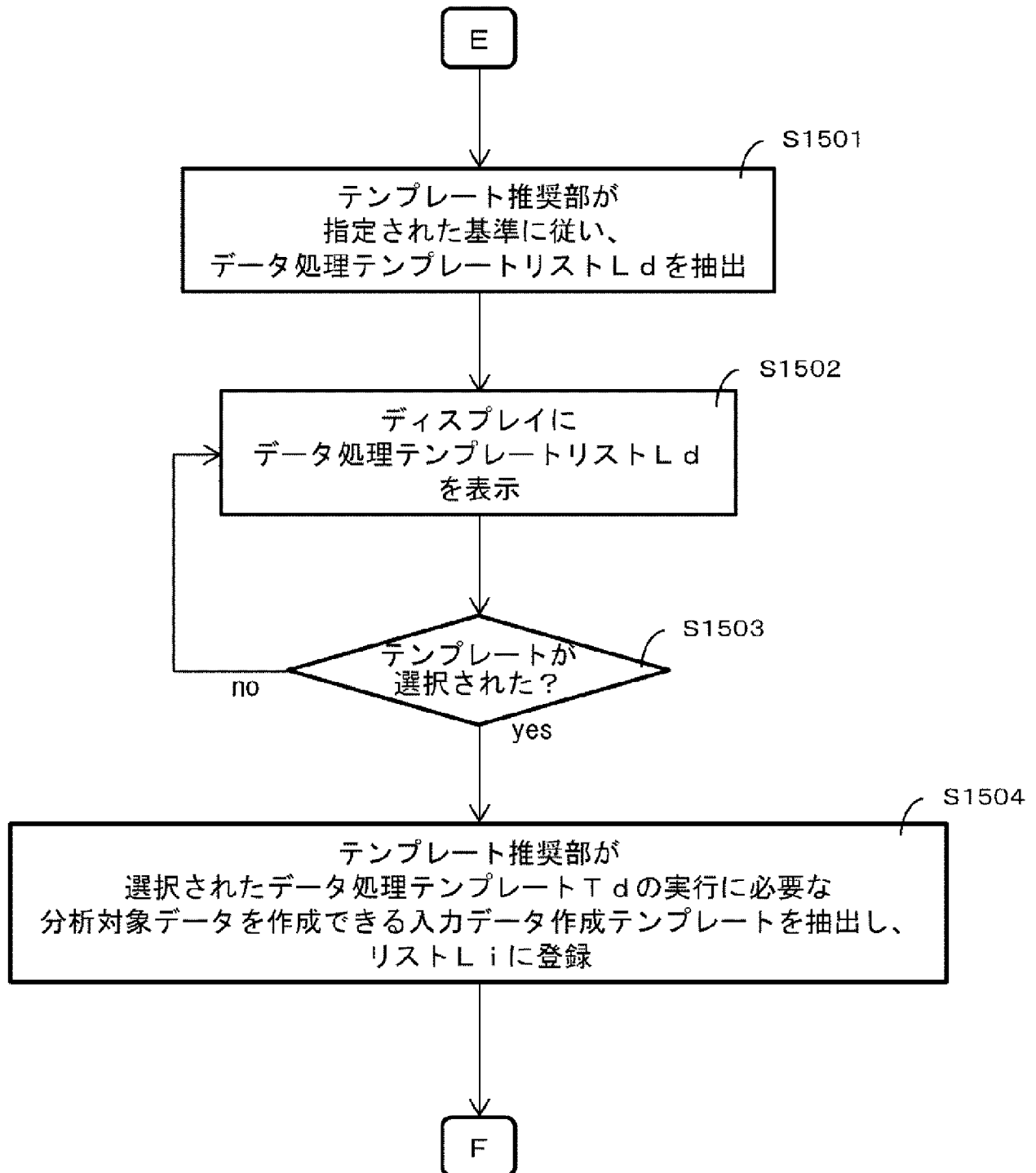


列「要素の出現数」を棒グラフで表示し、列「KEY」を横軸の項目名とする



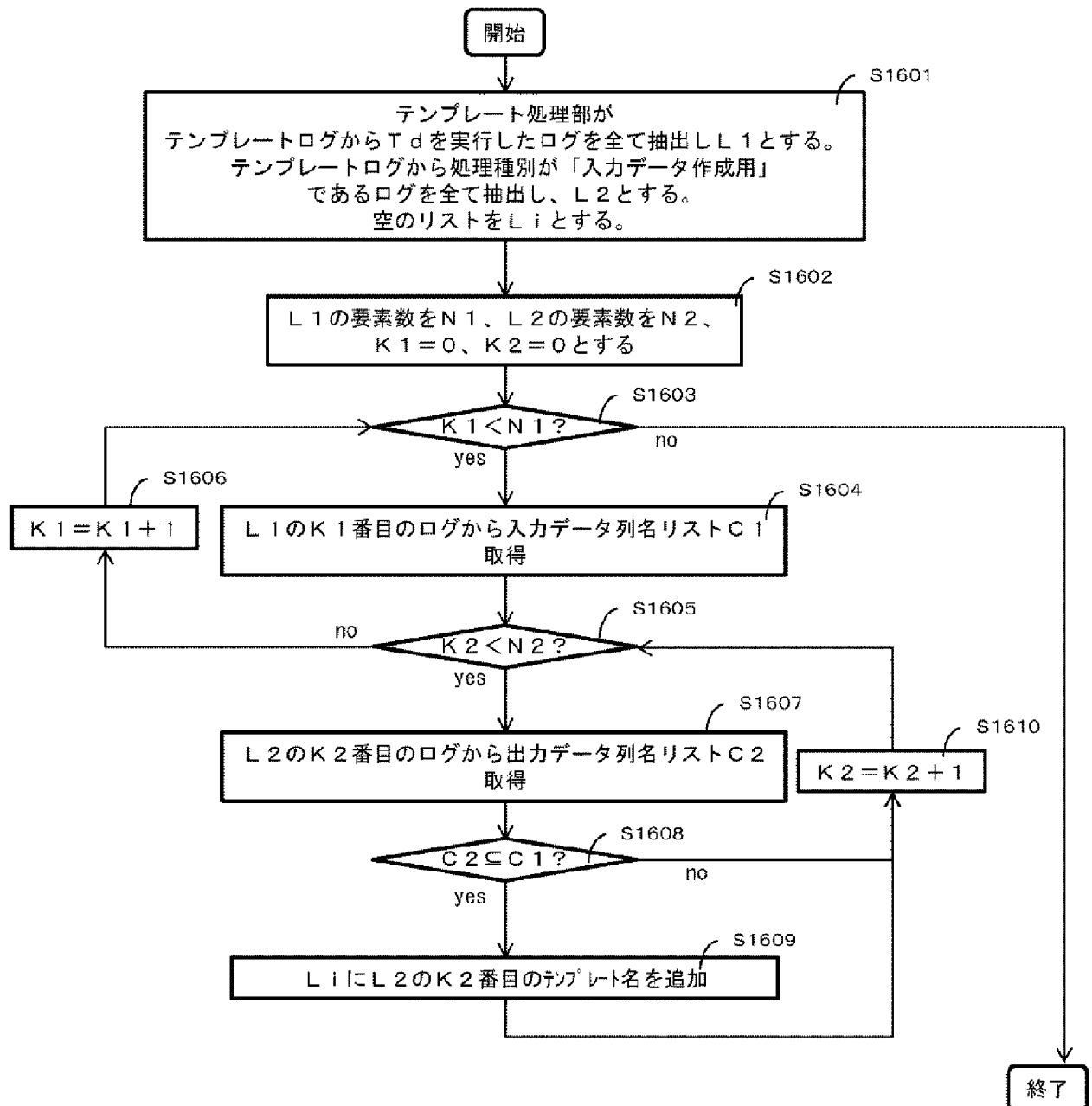
[図15]

図 1 5



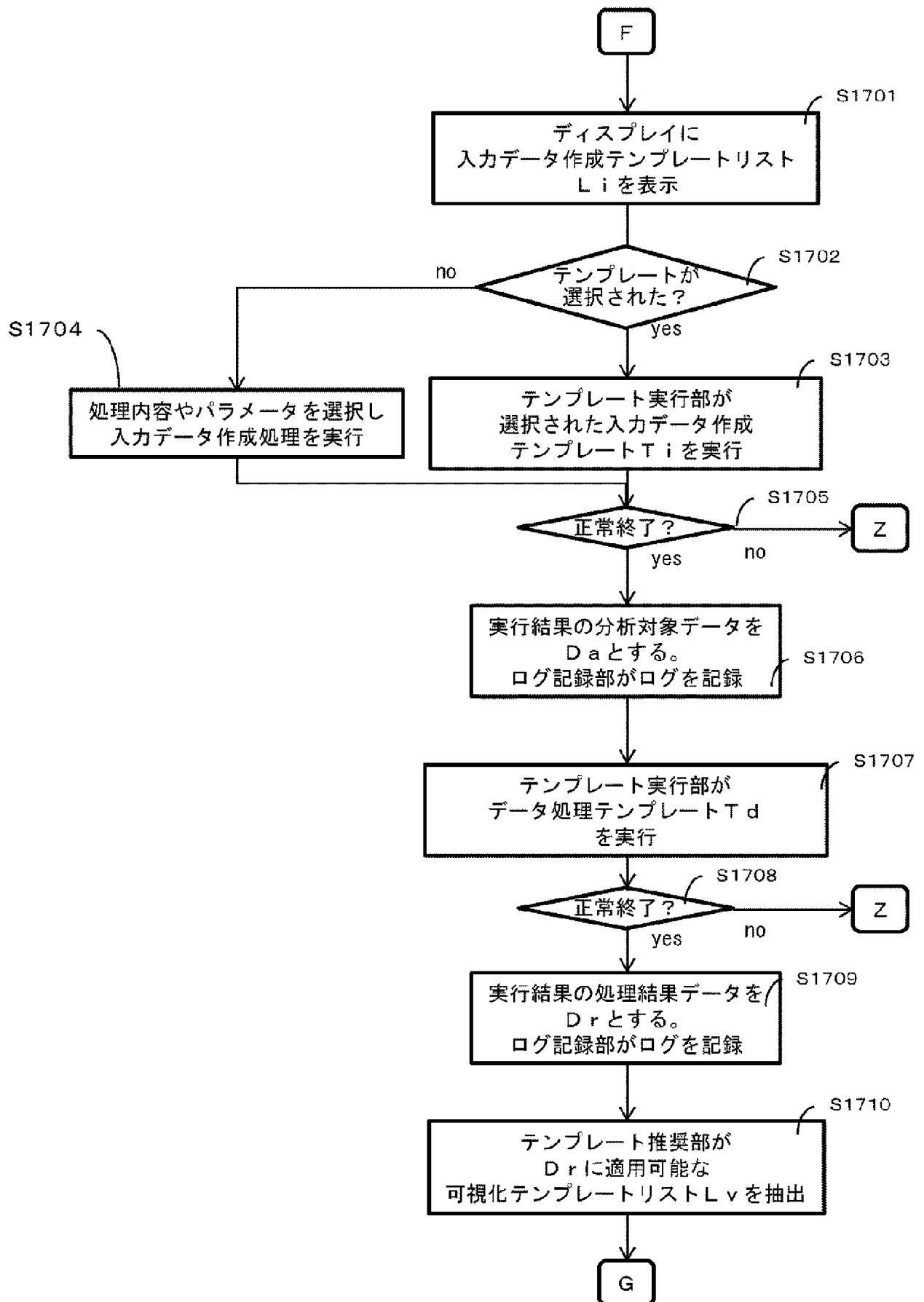
[図16]

図 1 6



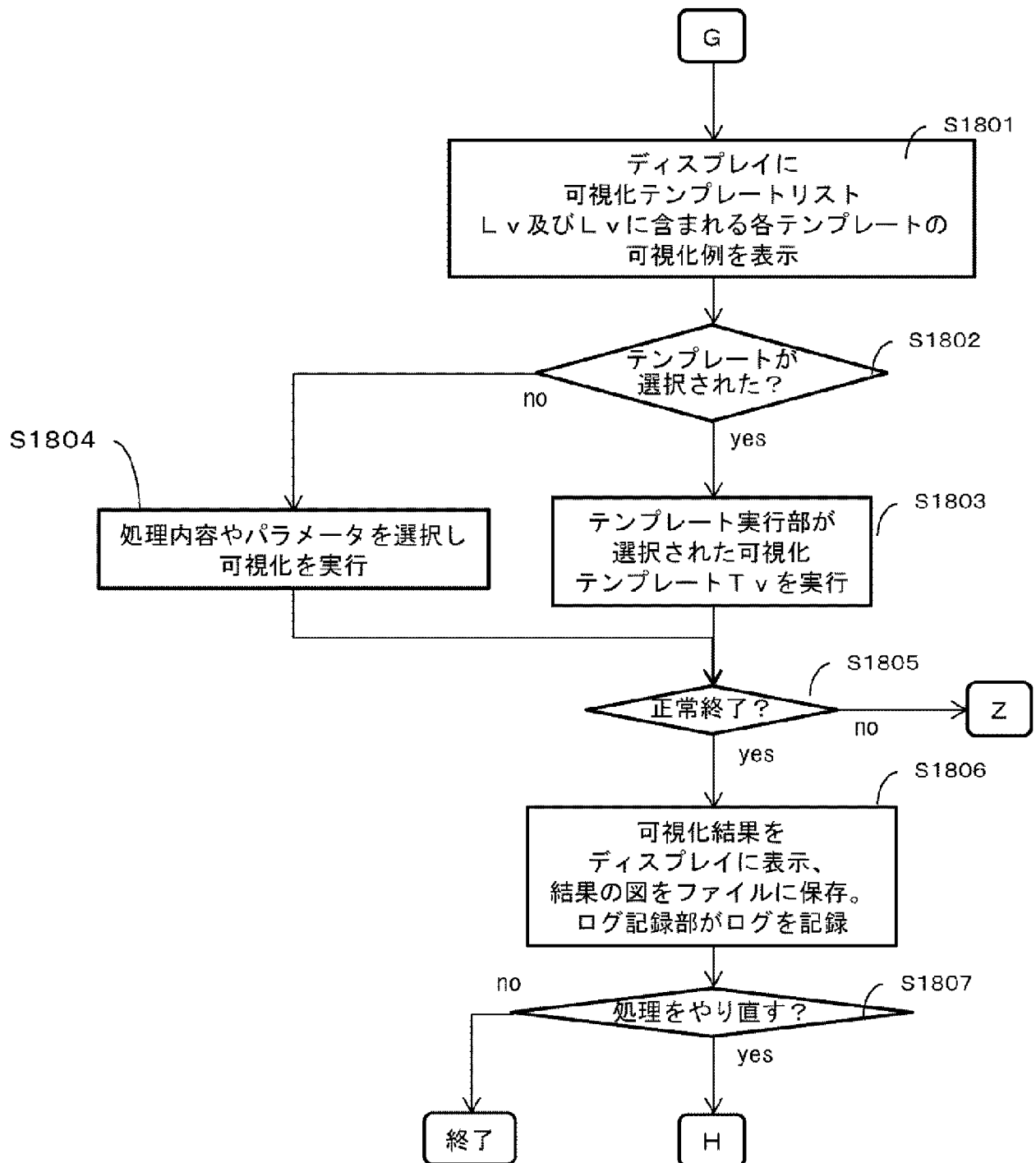
[図17]

図 1 7



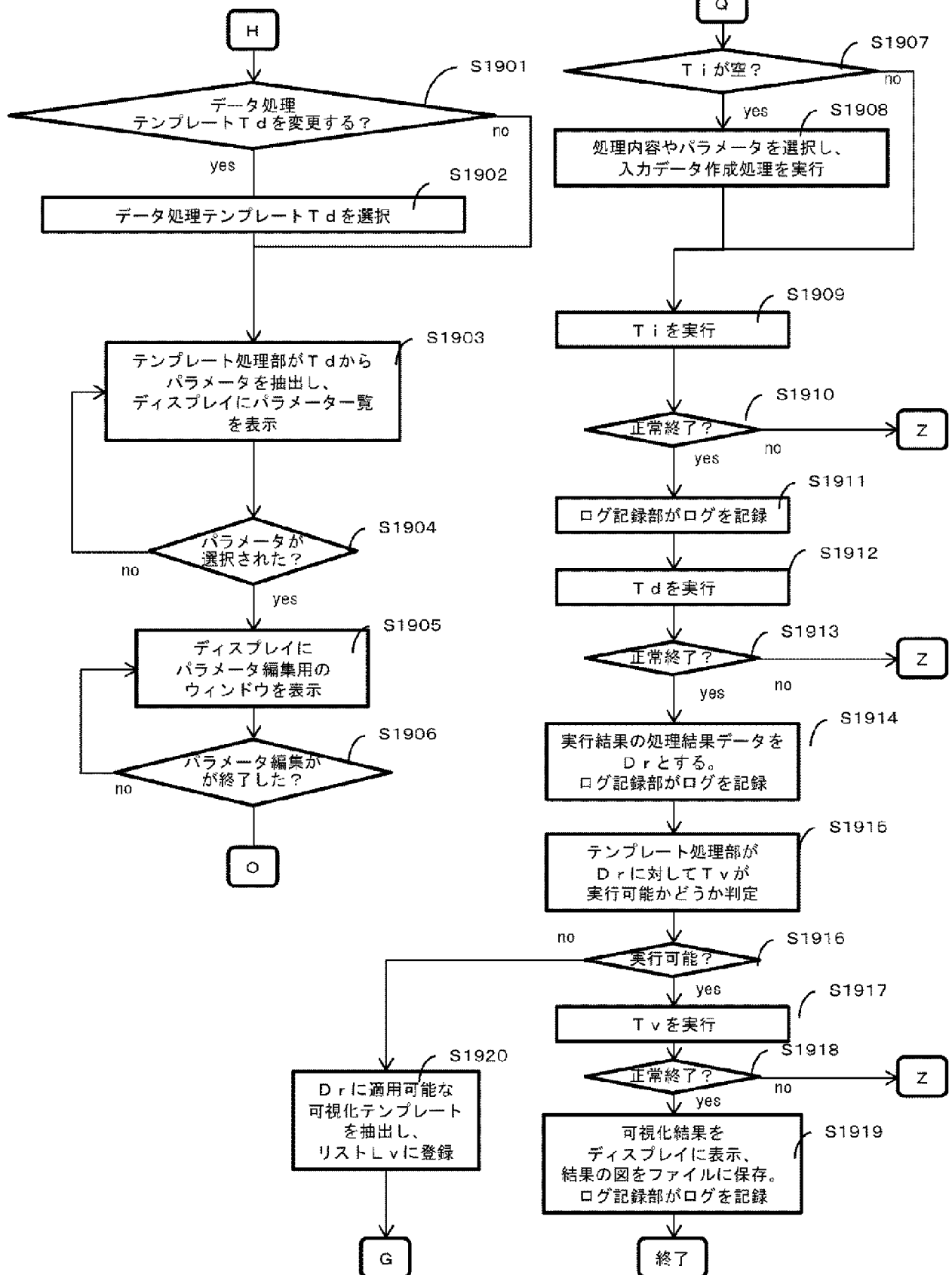
[図18]

図 1 8



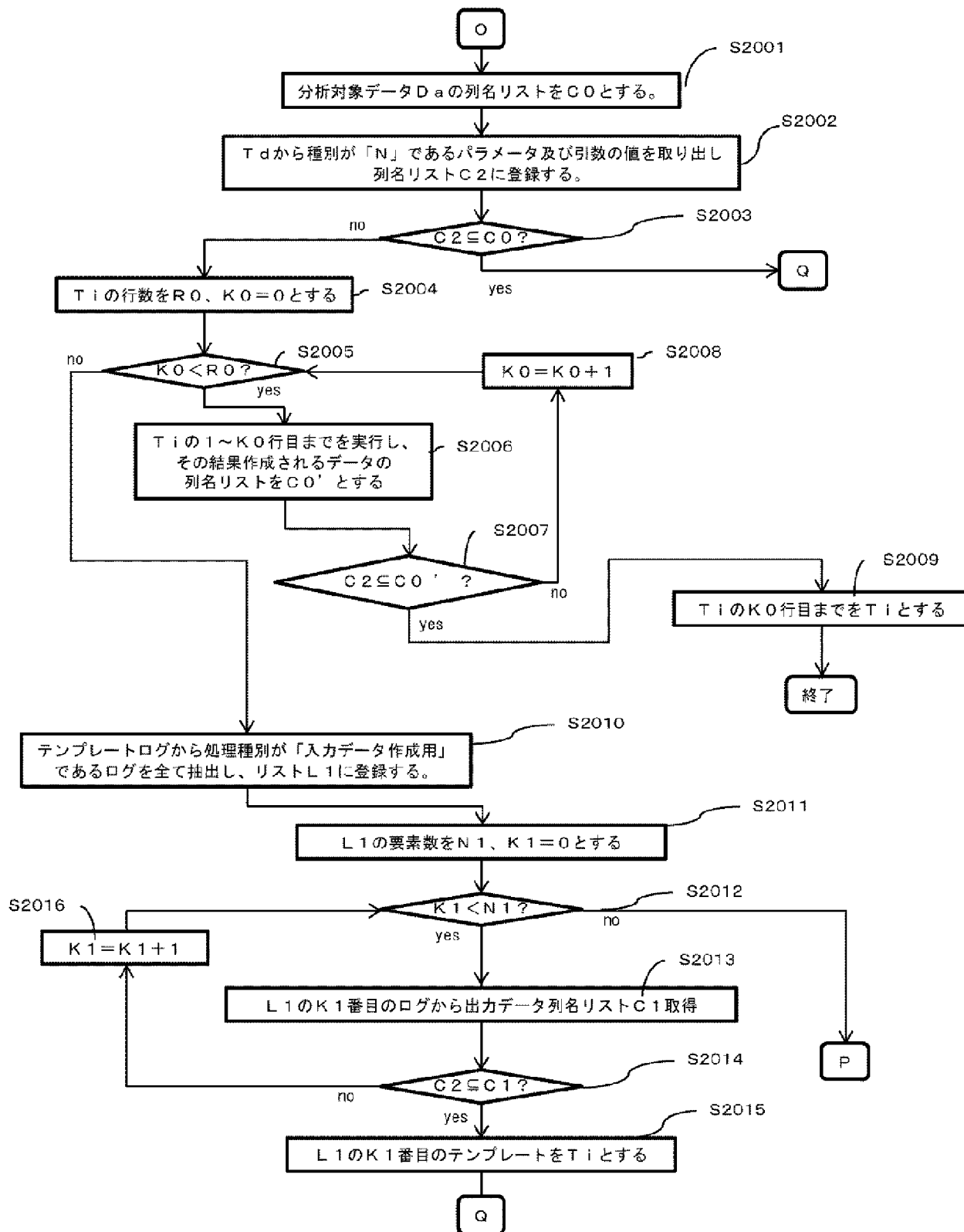
[図19]

図 19



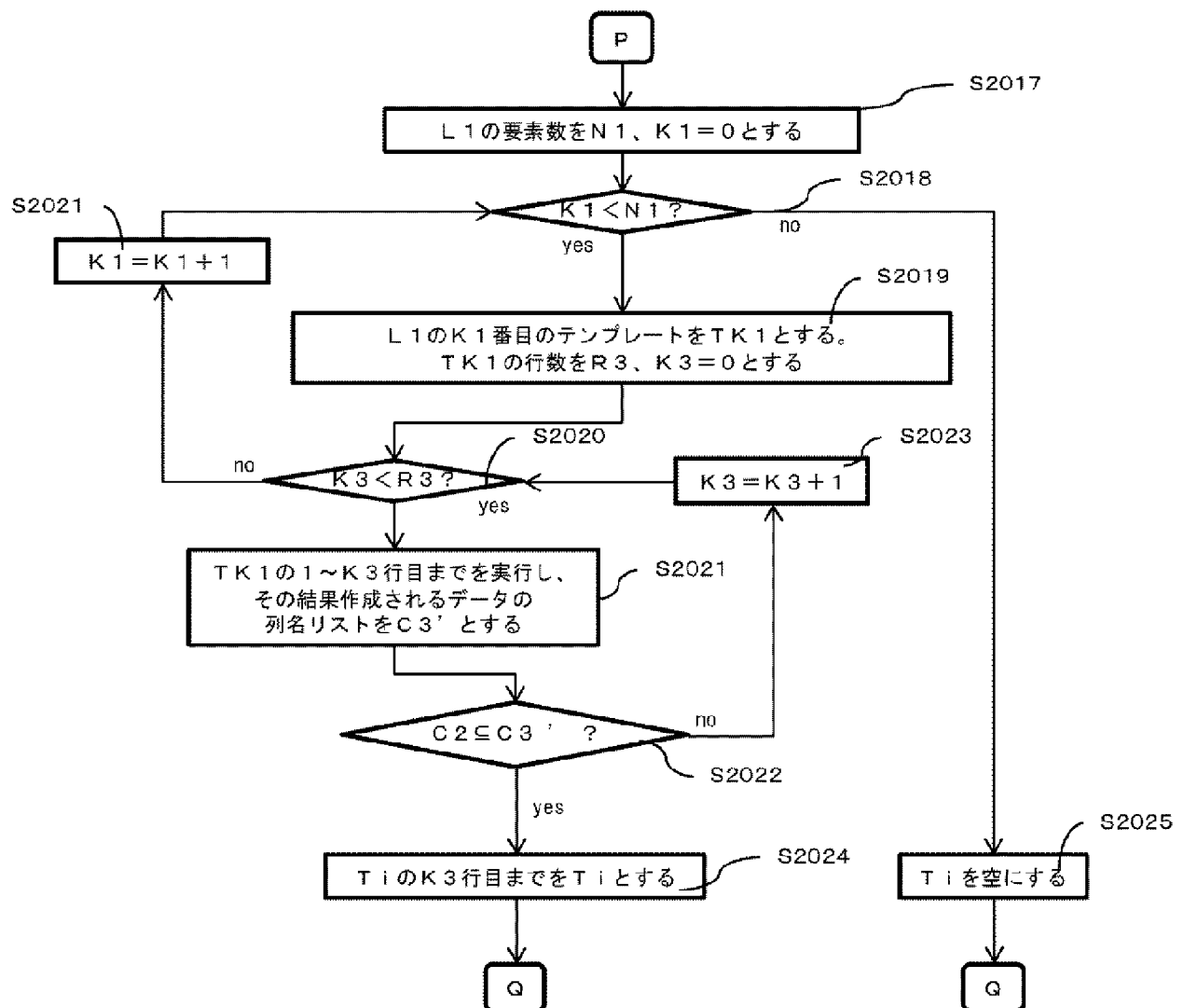
[図20A]

図 20 A



[図20B]

図 20 B



[図21A]

図 2 1 A

選択テンプレートTdを実行した結果の処理結果データDrの例
Tdの内容を「指定した2列のクロス集計を行う」の場合

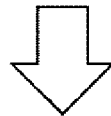
パラメータ1=N	“KEY”
パラメータ2=N	“列A”
“クロス集計” パラメータ1 パラメータ2	

[図21B]

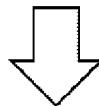
図 2 1 B

分析対象データ D a

CC4	AA0	BB1	...	空白または0	空白または0	...
CC2	AA0	BB2	...	空白または0	空白または0	...
CC3	AA0	BB3	...	空白または0	空白または0	...
CC1	AA1	BB4	...	空白または0	空白または0	...
...	...	...	...	空白または0	空白または0	...
CC0	空白または0	空白または0	...	AA1	YY0	...
CC2	空白または0	空白または0	...	AA1	YY1	...
CC3	空白または0	空白または0	...	AA0	YY1	...
CC1	空白または0	空白または0	...	AA0	YY2	...
...	空白または0	空白または0	...	...	...	...



列Aの要素の出現数を数える



分析対象データ D r

KEY	AA0	AA1
CC0	2	0
CC1	1	4
CC2	3	0
CC3	4	4
CC4	0	1

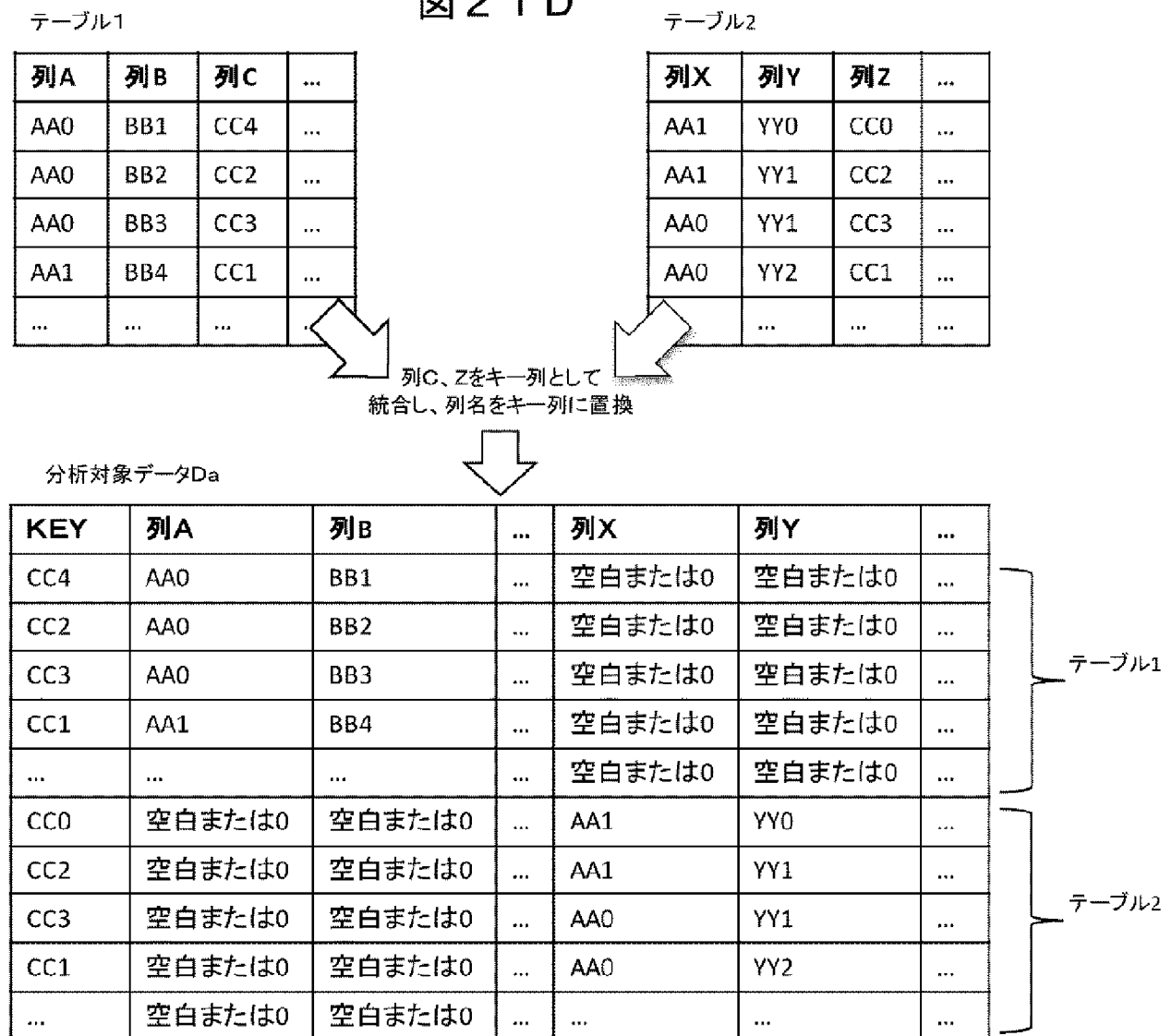
[図21C]

図 2 1 C

パラメータ 1=E	“テーブル1”				
パラメータ 2=E	“テーブル2”				
パラメータ 3=N	“列C”				
パラメータ 4=N	“列Z”				
パラメータ 5=N	“KEY”				
“テーブル統合”	パラメータ 1	パラメータ 2	パラメータ 3	パラメータ 4	パラメータ 5

[図21D]

図 2 1 D



[図21E]

図 2 1 E

選択テンプレートT_vを実行した結果の可視化の例

処理結果データD_rの変更に伴い、適用可能なテンプレートT_vを抽出する。

T_vの内容が「指定した2列を棒グラフで表示し、指定した列を横軸の項目名とする」の場合

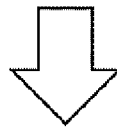
パラメータ1=N	“AA0”
パラメータ2=N	“AA1”
パラメータ2=N	“KEY”
“棒グラフ”	パラメータ1 パラメータ2 パラメータ3

[図21F]

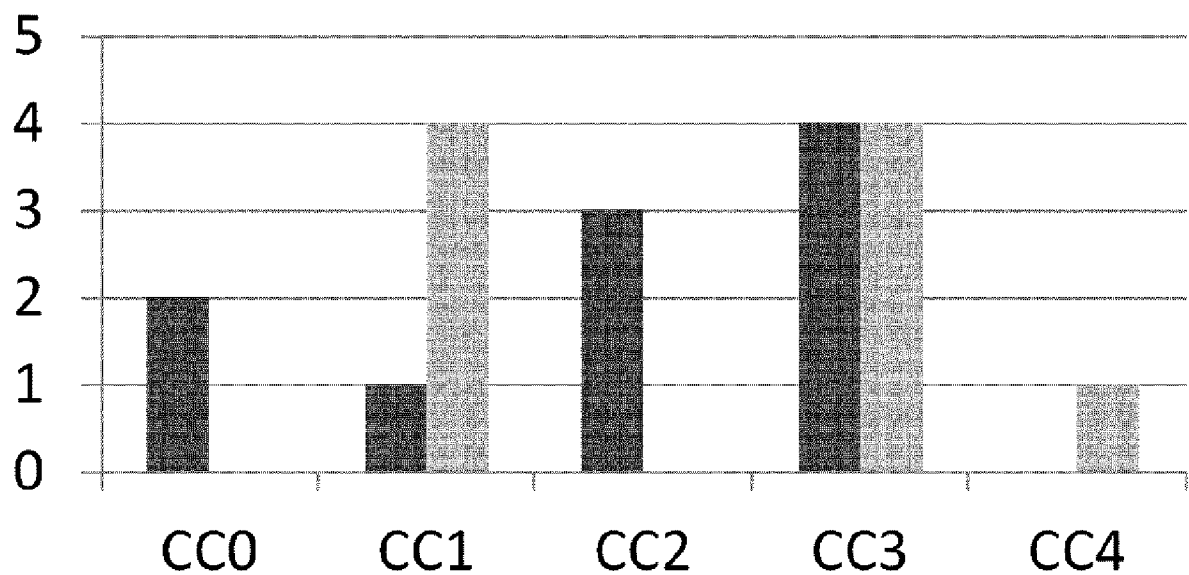
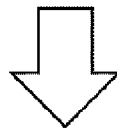
図 2 1 F

処理結果データ D r

KEY	AA0	AA1
CC0	2	0
CC1	1	4
CC2	3	0
CC3	4	4
CC4	0	1

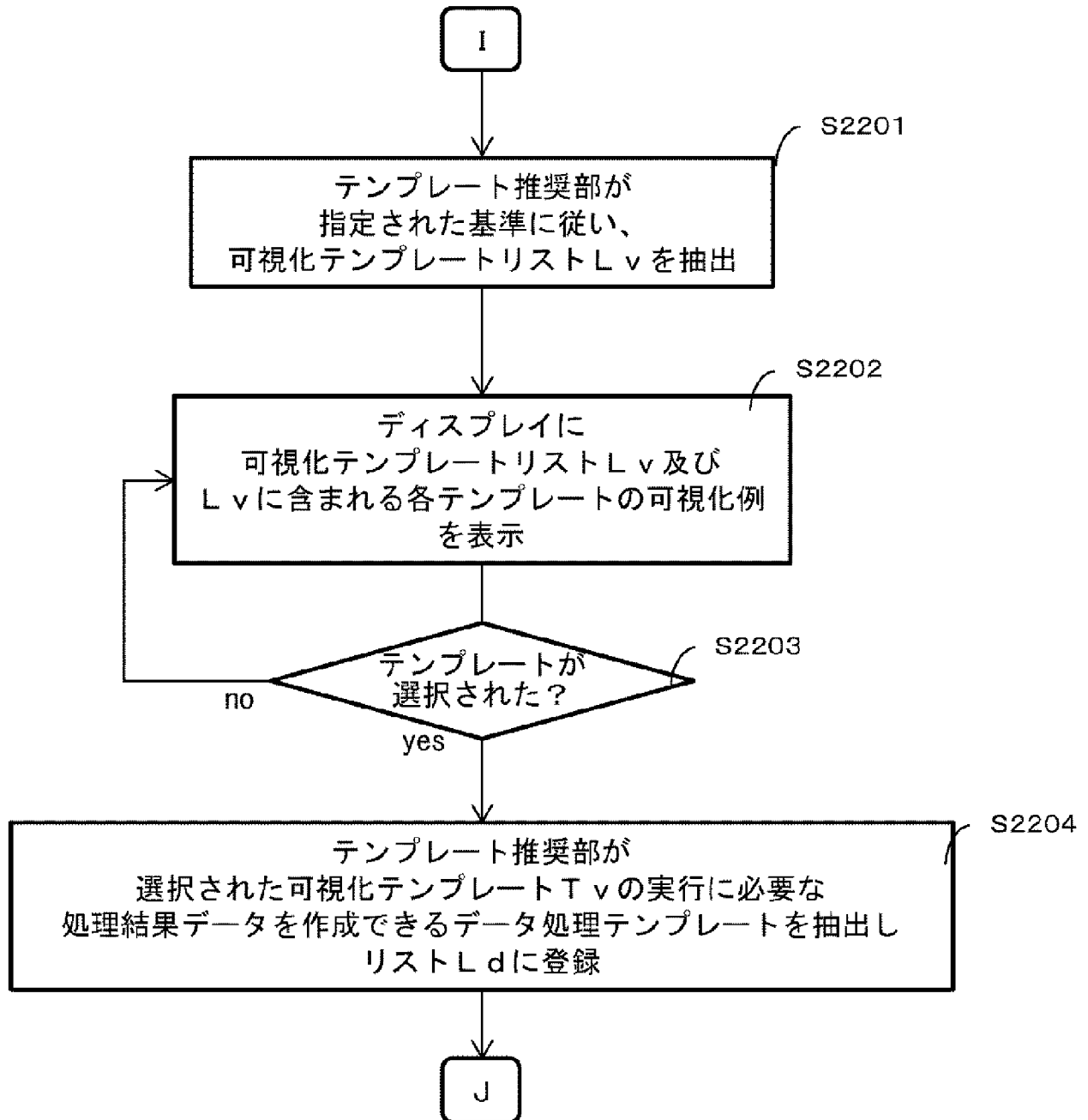


列「AA0、AA1」を棒グラフで表示し、列「KEY」を横軸の項目名とする



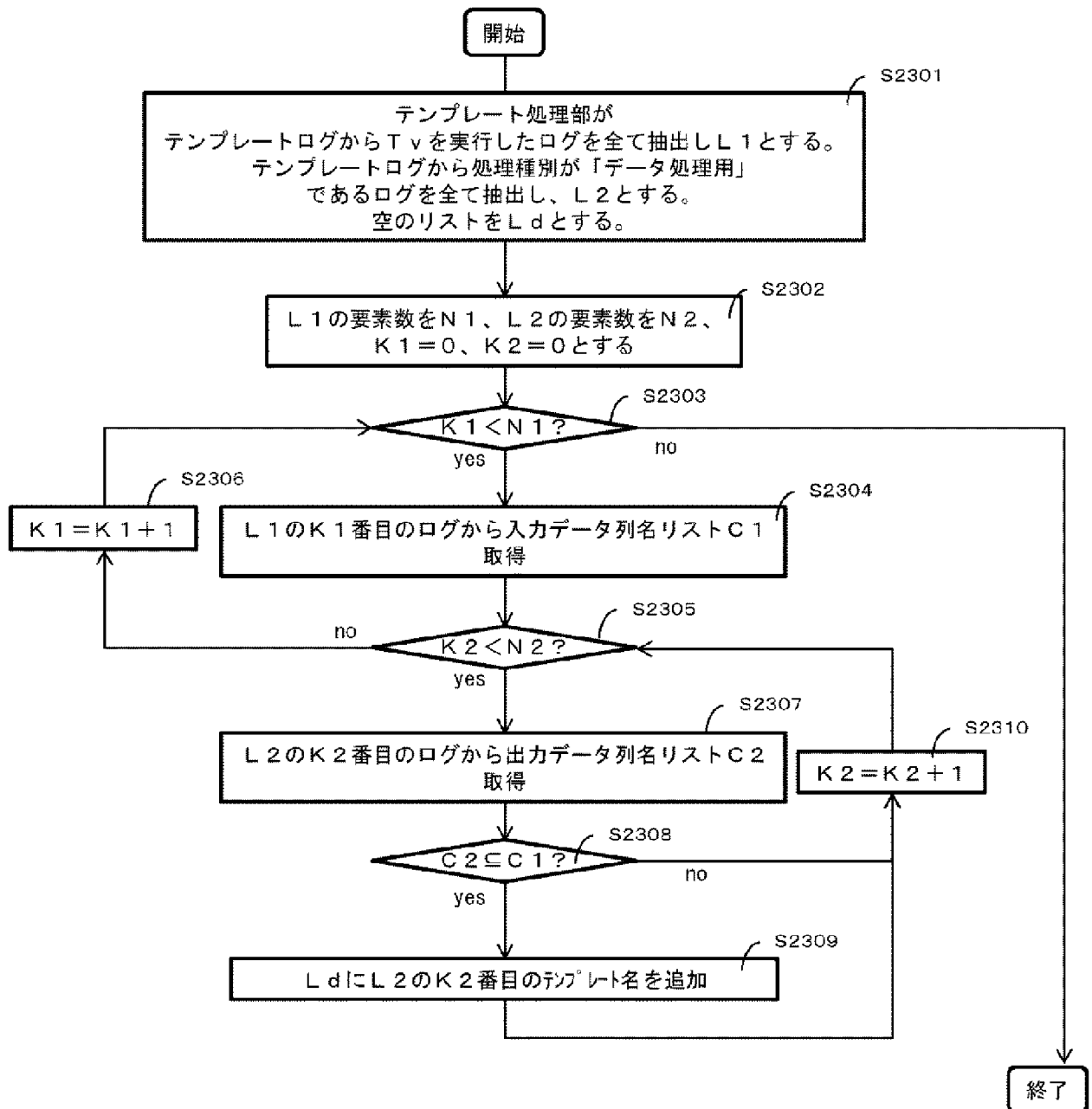
[図22]

図22



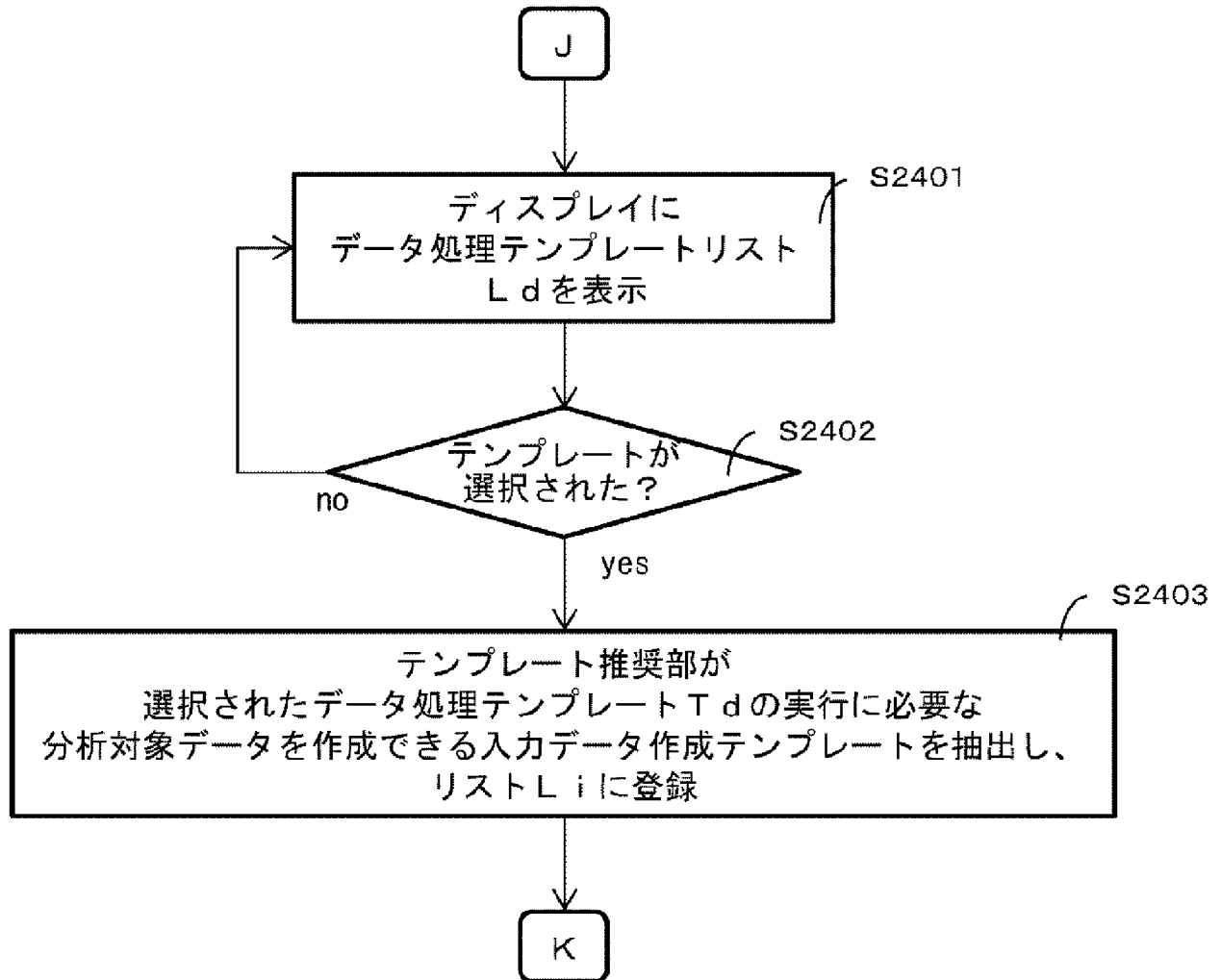
[図23]

図 2 3

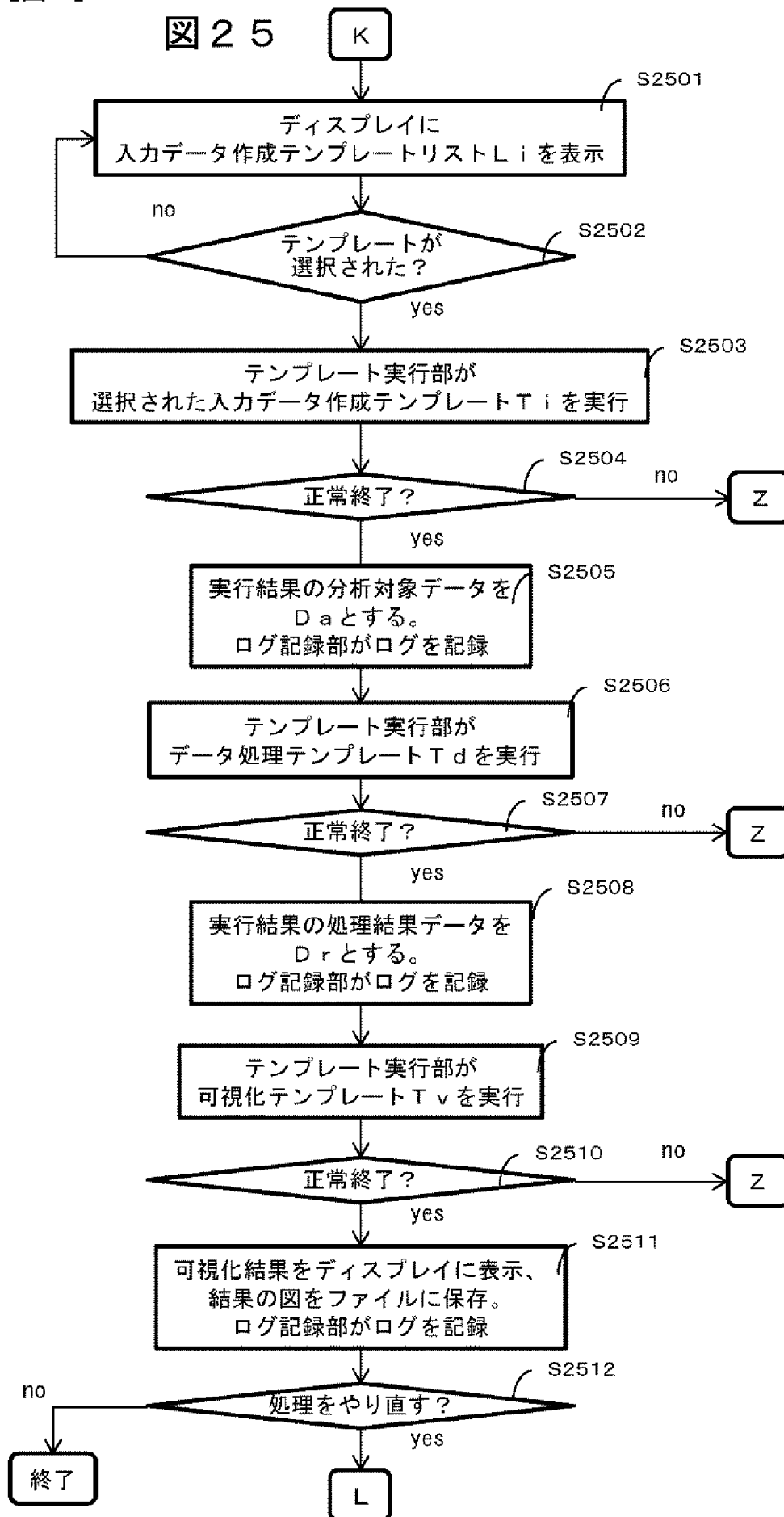


[図24]

図24

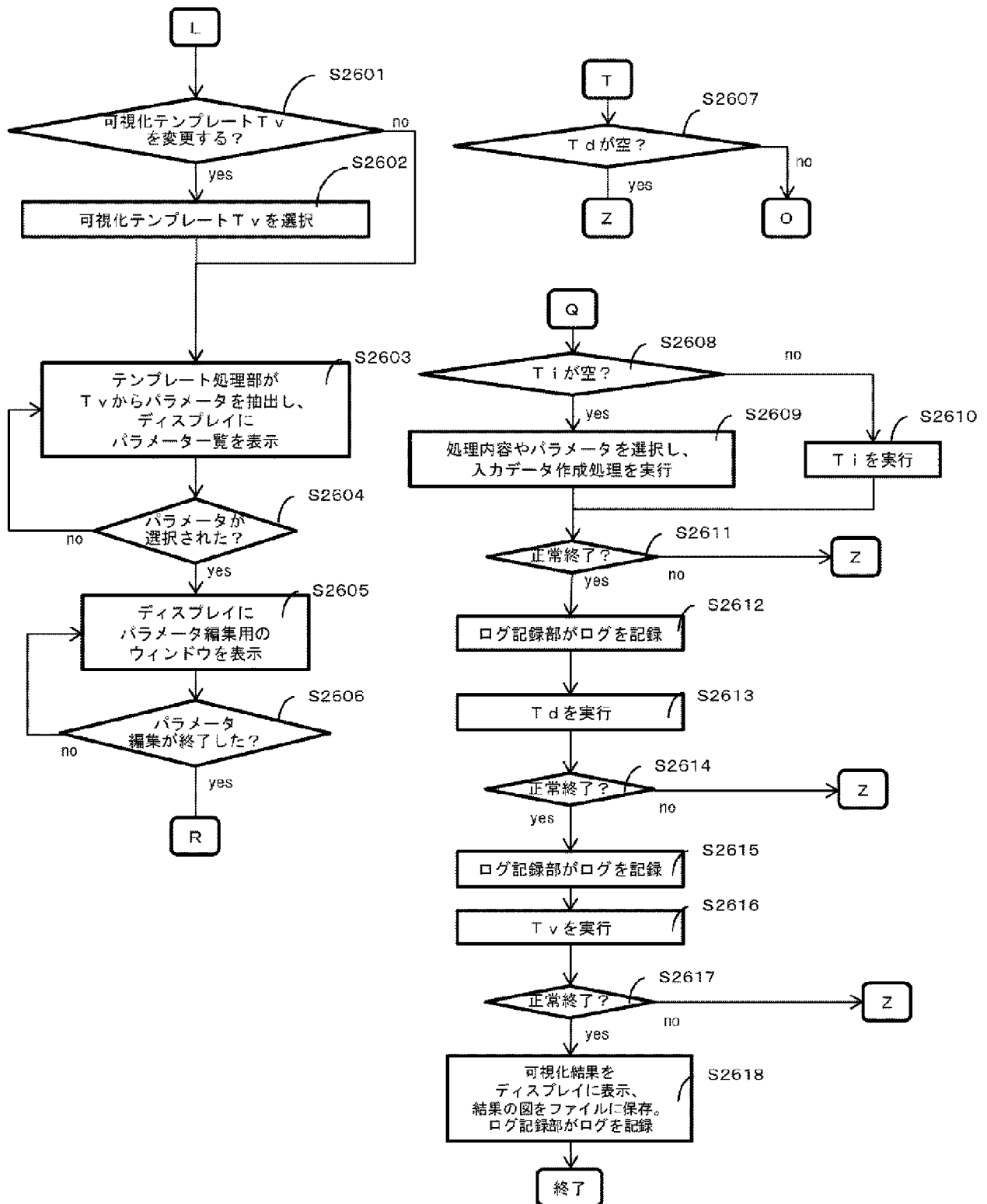


[図25]



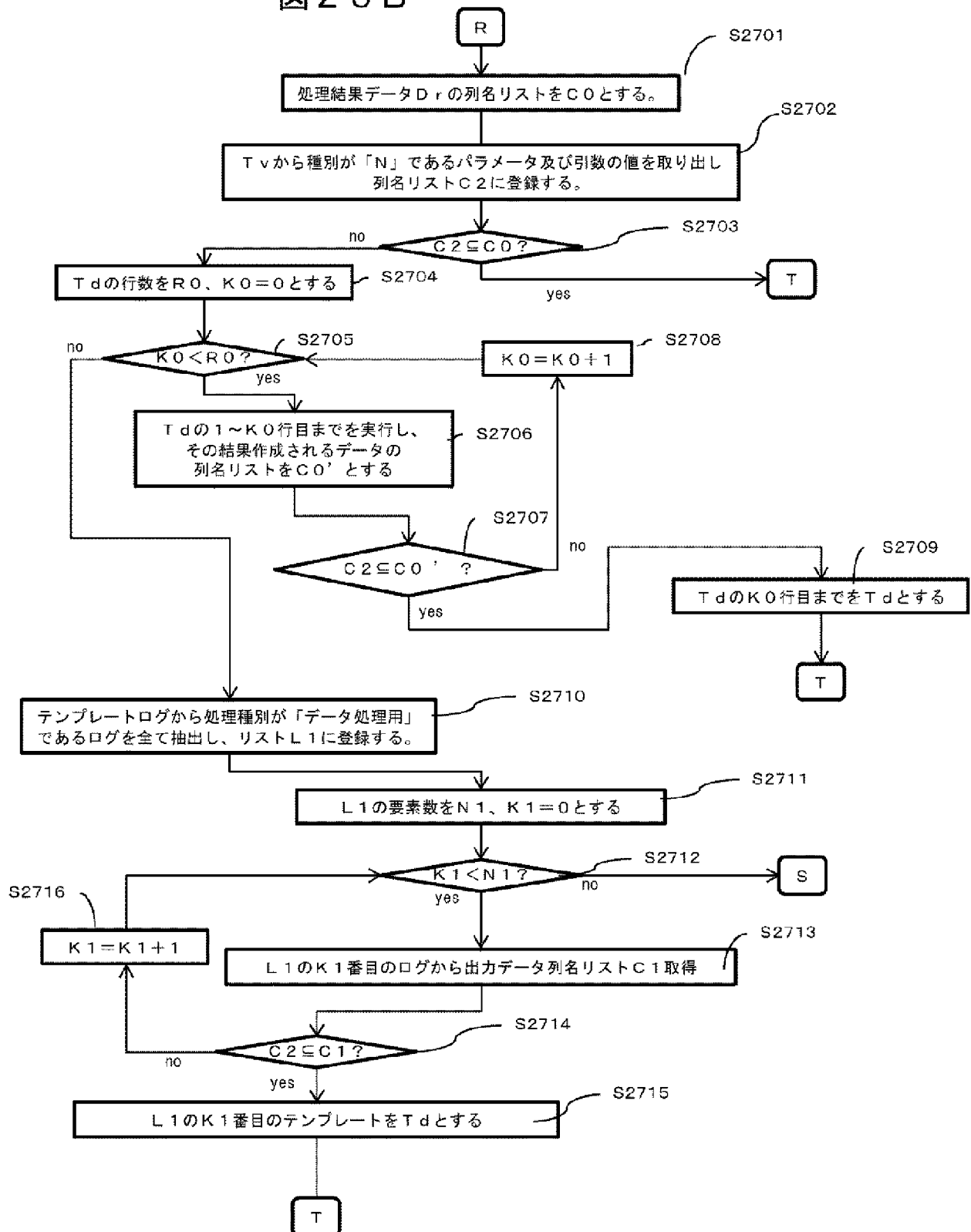
[図26A]

図 2 6 A



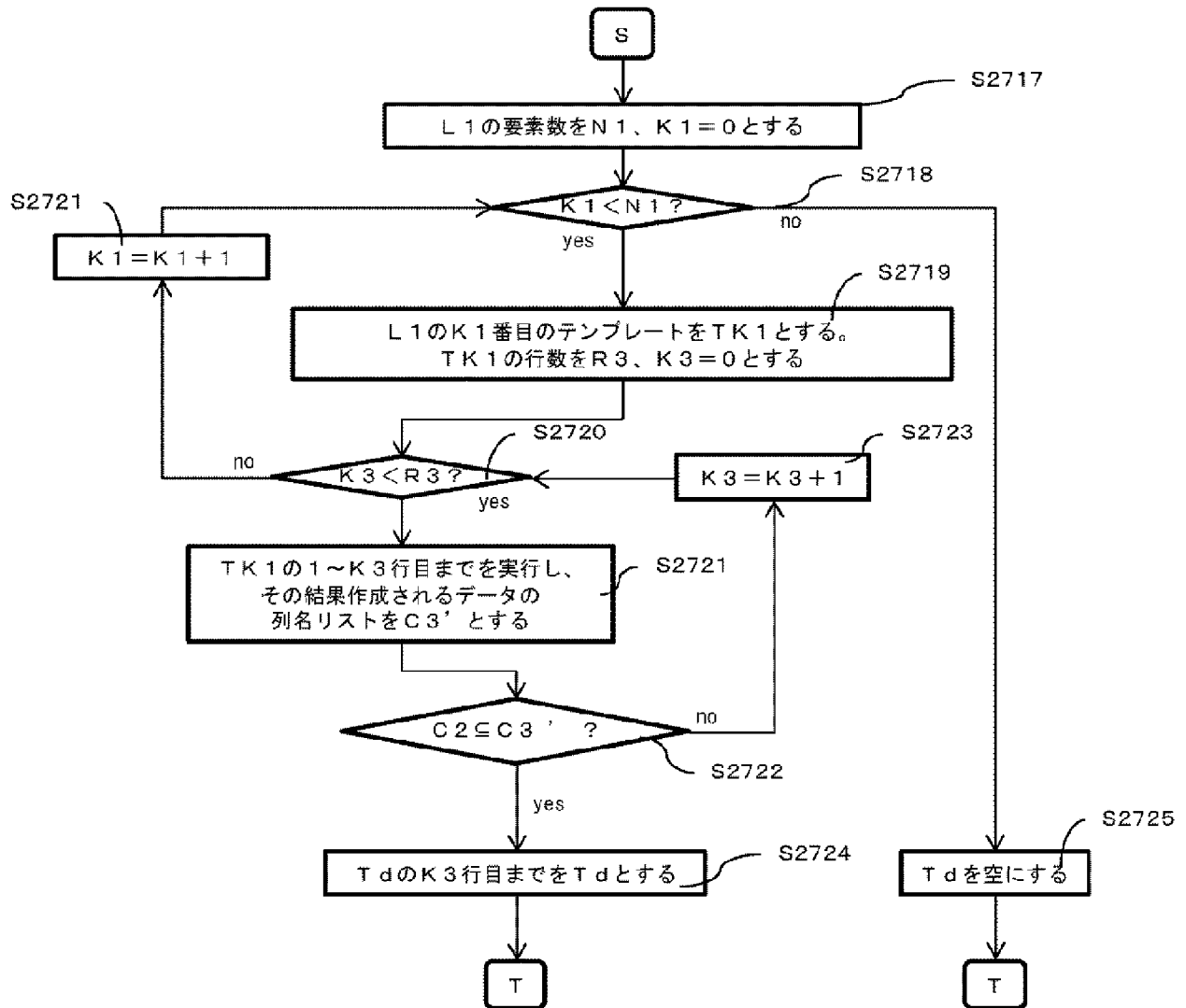
[図26B]

図 2 6 B



[図26C]

図 2 6 C



[図27A]

図 2 7 A

選択テンプレートT_vを実行した結果の可視化の例

T_vの内容が「指定した列を円グラフで表示し、指定した列を項目名とする」の場合

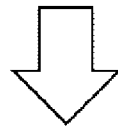
パラメータ1=N	“要素の出現数”
パラメータ2=N	“KEY”
“円グラフ”	パラメータ1 パラメータ2

[図27B]

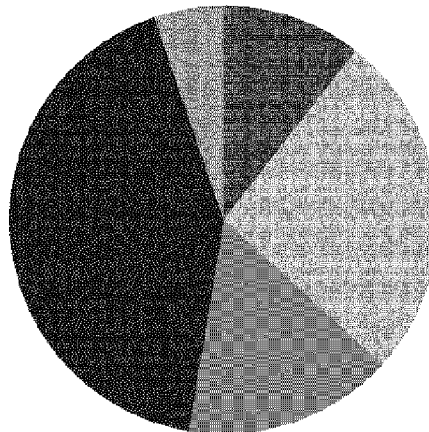
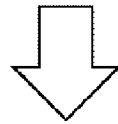
図 2 7 B

分析対象データ D_r

KEY	要素の出現数
CC0	2
CC1	5
CC2	3
CC3	8
CC4	1



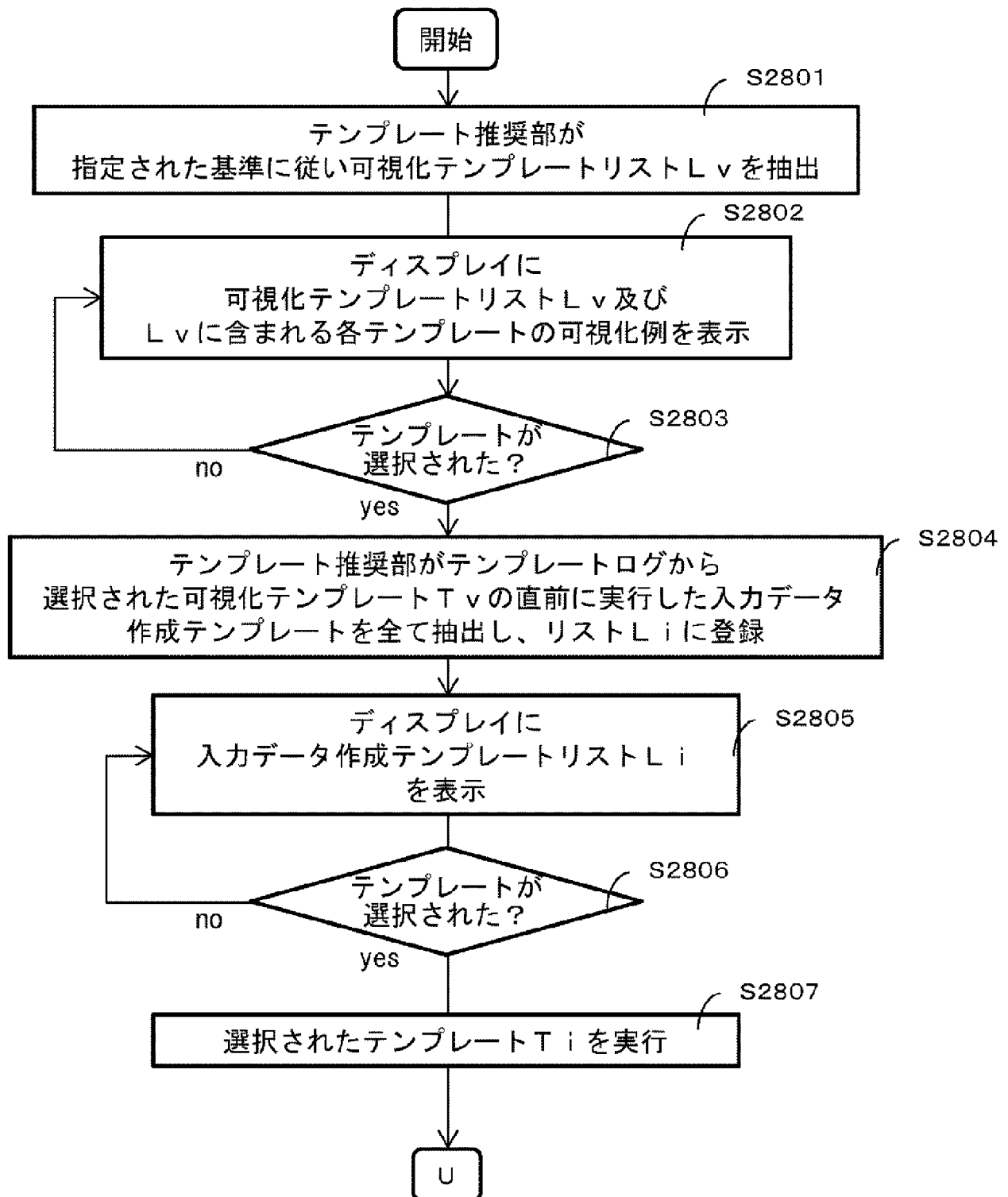
列「要素の出現数」を円グラフで表示し、列「KEY」を横軸の項目名とする



■ CC0 ■ CC1 ■ CC2 ■ CC3 ■ CC4

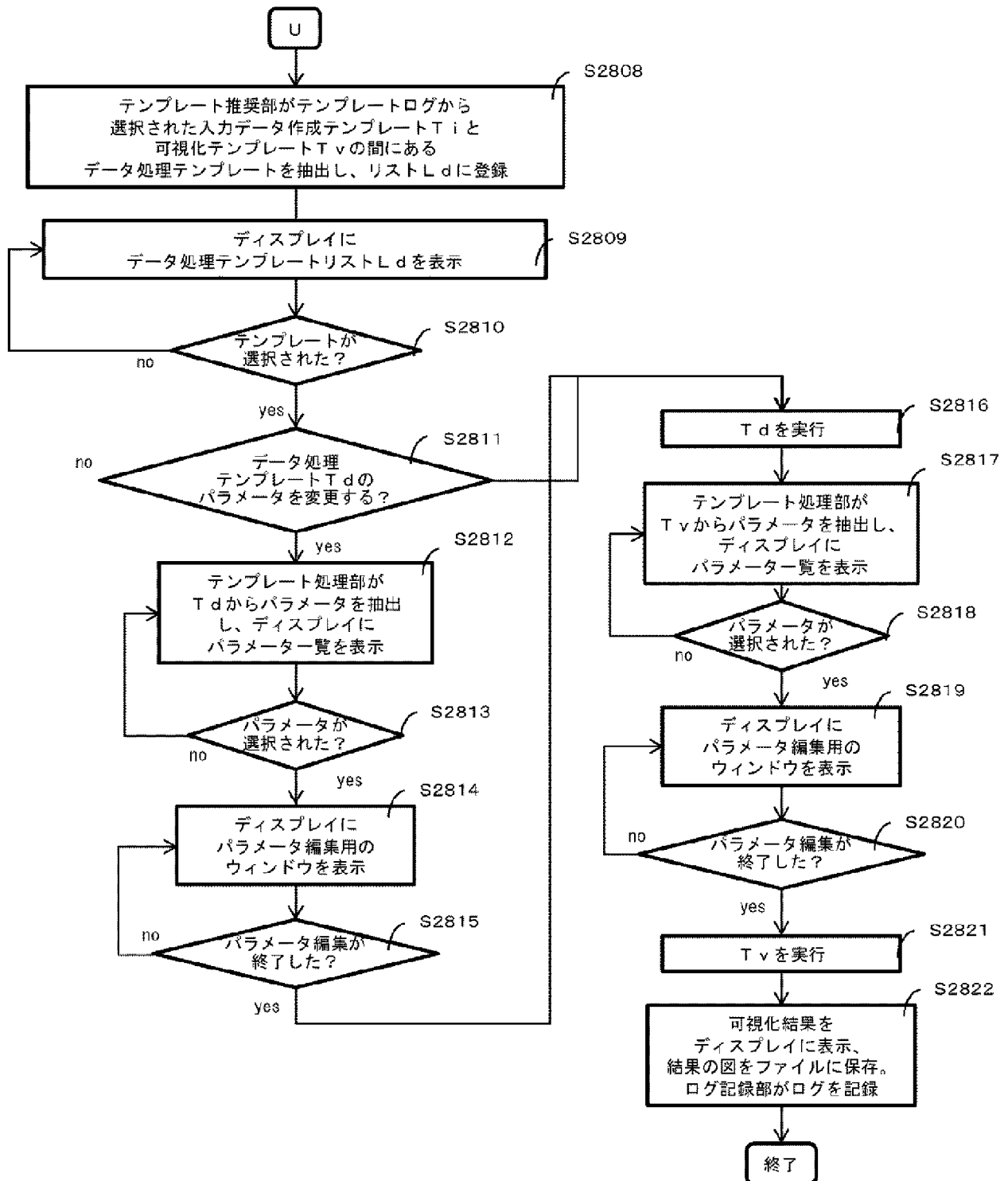
[図28A]

図 2 8 A



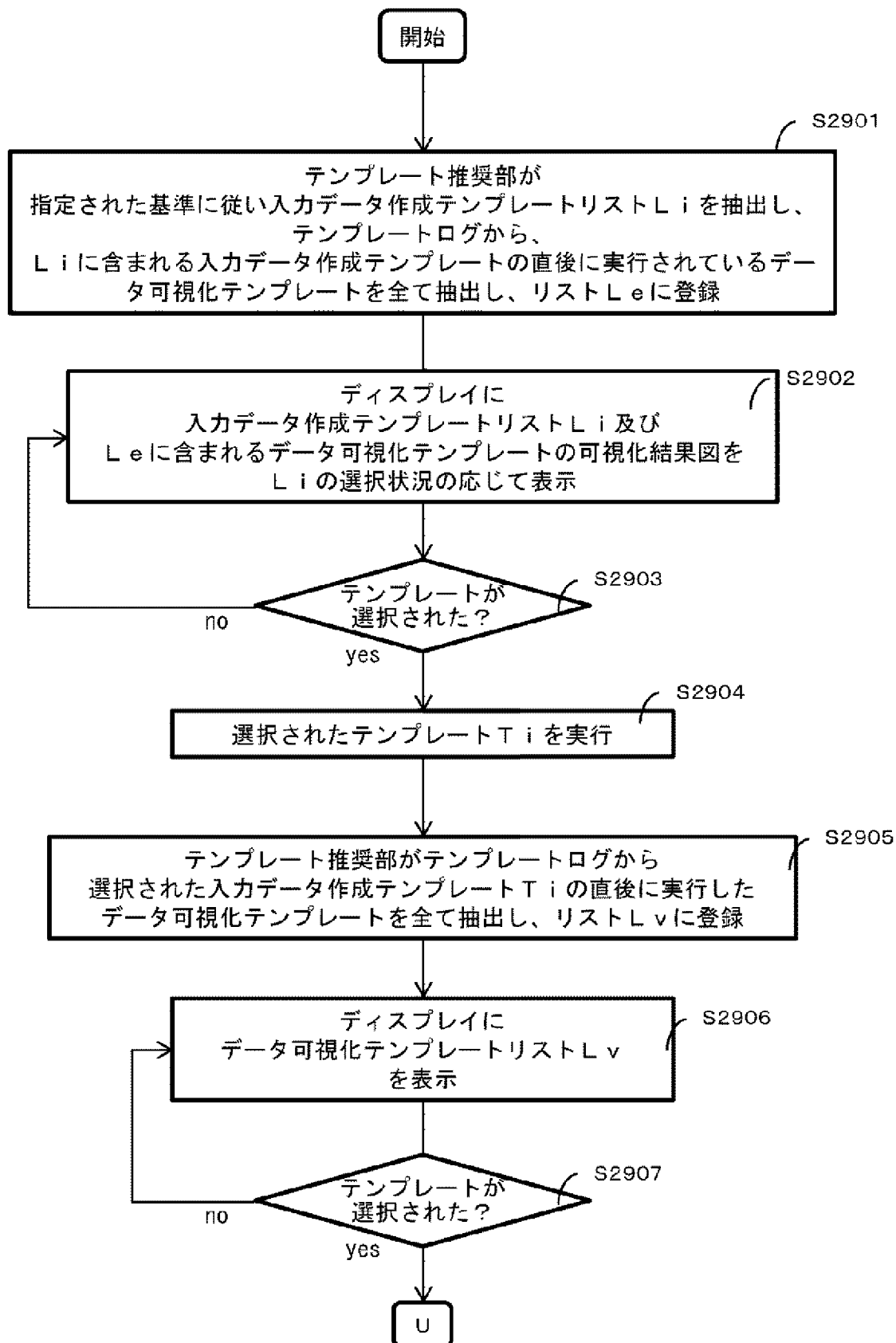
[図28B]

図 28 B



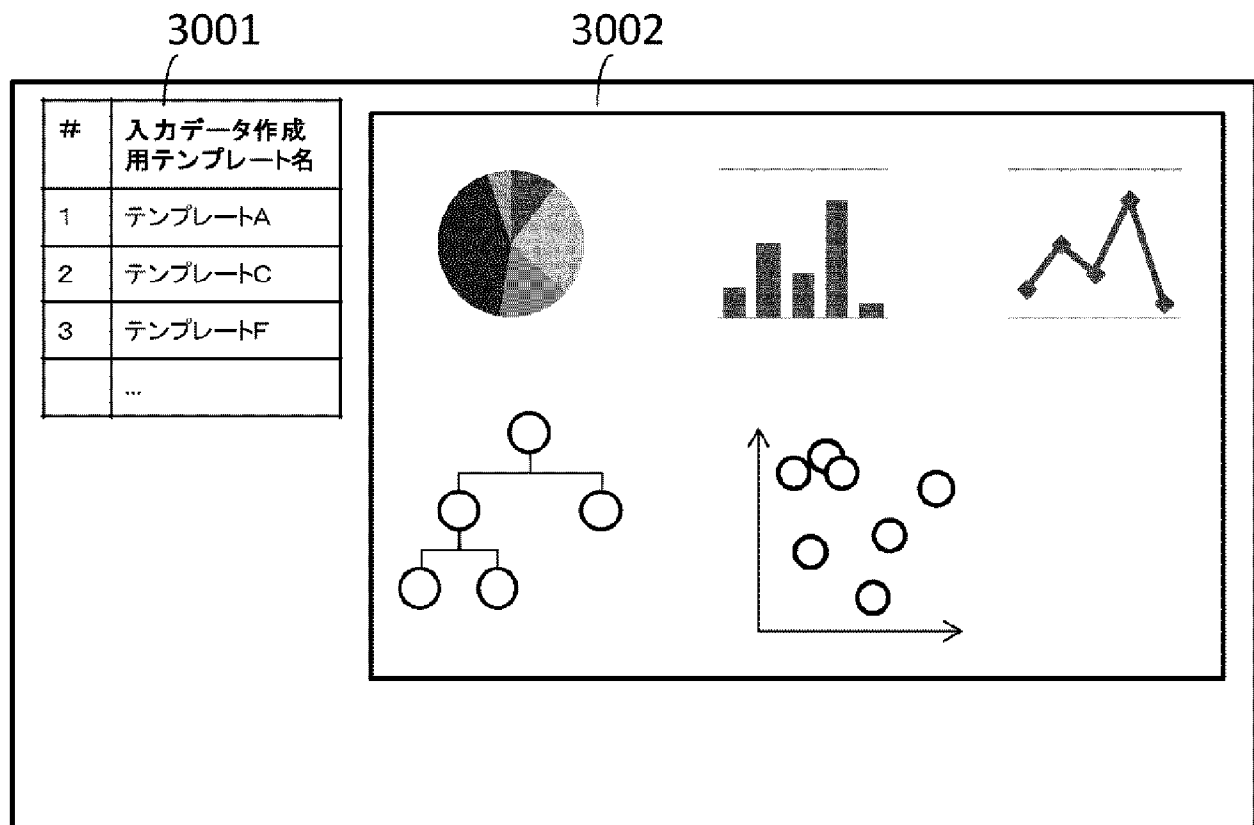
[図29]

図 29



[図30]

図 30



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2012/067887

A. CLASSIFICATION OF SUBJECT MATTER

G06F19/00(2011.01)i, G06Q10/06(2012.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F19/00, G06Q10/06

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho 1922-1996 Jitsuyo Shinan Toroku Koho 1996-2012

Kokai Jitsuyo Shinan Koho 1971-2012 Toroku Jitsuyo Shinan Koho 1994-2012

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y A	JP 2012-73812 A (Hitachi, Ltd.), 12 April 2012 (12.04.2012), paragraphs [0020] to [0059] (Family: none)	1, 10 2-5, 11-15 6-9
Y A	JP 2006-344211 A (Computer Patent Annuities Ltd. Partnership), 21 December 2006 (21.12.2006), particularly, claims 1, 2 & US 2006/0271379 A1 & KR 10-2006-0122756 A	2-5, 11-15 1, 6-10
Y A	JP 2011-258097 A (Hitachi, Ltd.), 22 December 2011 (22.12.2011), particularly, paragraphs [0175] to [0177] (Family: none)	2-5, 11-15 1, 6-10



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T"

later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X"

document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y"

document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&"

document member of the same patent family

Date of the actual completion of the international search

05 October, 2012 (05.10.12)

Date of mailing of the international search report

16 October, 2012 (16.10.12)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2012/067887

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2012-118968 A (International Business Machines Corp.), 21 June 2012 (21.06.2012), entire text; all drawings & US 2012/0143780 A1 & US 2012/0191762 A1 & CN 102542344 A	1-15
A	JP 2003-150650 A (Hitachi, Ltd.), 23 May 2003 (23.05.2003), entire text; all drawings & US 2003/0090526 A1	1-15
A	JP 2003-58697 A (NCR International Inc.), 28 February 2003 (28.02.2003), entire text; all drawings (Family: none)	1-15
A	JP 2003-58562 A (NCR International Inc.), 28 February 2003 (28.02.2003), entire text; all drawings (Family: none)	1-15

A. 発明の属する分野の分類（国際特許分類（I P C））
Int.Cl. G06F19/00(2011.01)i, G06Q10/06(2012.01)i

B. 調査を行った分野

調査を行った最小限資料（国際特許分類（I P C））

Int.Cl. G06F19/00, G06Q10/06

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	1 9 2 2 - 1 9 9 6 年
日本国公開実用新案公報	1 9 7 1 - 2 0 1 2 年
日本国実用新案登録公報	1 9 9 6 - 2 0 1 2 年
日本国登録実用新案公報	1 9 9 4 - 2 0 1 2 年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
X Y A	JP 2012-73812 A（株式会社日立製作所） 2012.04.12, 【0020】～【0059】（ファミリーなし）	1, 10 2-5, 11-15 6-9
Y A	JP 2006-344211 A（コンピューター・パテント・アニュイティーズ・ リミテッド・パートナーシップ） 2006.12.21, 特に【請求項1】、【請求項2】 & US 2006/0271379 A1 & KR 10-2006-0122756 A	2-5, 11-15 1, 6-10

☒ C欄の続きにも文献が列挙されている。

☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの

「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの

「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）

「O」口頭による開示、使用、展示等に言及する文献

「P」国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの

「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの

「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの

「&」同一パテントファミリー文献

国際調査を完了した日

0 5 . 1 0 . 2 0 1 2

国際調査報告の発送日

1 6 . 1 0 . 2 0 1 2

国際調査機関の名称及びあて先

日本国特許庁（I S A / J P）

郵便番号100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官（権限のある職員）

大野 朋也

電話番号 03-3581-1101 内線 3562

5 L

4 5 3 4

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
Y A	JP 2011-258097 A (株式会社日立製作所) 2011.12.22, 特に【0175】～【0177】 (ファミリーなし)	2-5, 11-15 1, 6-10
A	JP 2012-118968 A (インターナショナル・ビジネス・マシーンズ・ コーポレーション) 2012.06.21, 全文、全図 & US 2012/0143780 A1 & US 2012/0191762 A1 & CN 102542344 A	1-15
A	JP 2003-150650 A (株式会社日立製作所) 2003.05.23, 全文、全図 & US 2003/0090526 A1	1-15
A	JP 2003-58697 A (エヌシーアール インターナショナル インコー ポレイテッド) 2003.02.28, 全文、全図 (ファミリーなし)	1-15
A	JP 2003-58562 A (エヌシーアール インターナショナル インコー ポレイテッド) 2003.02.28, 全文、全図 (ファミリーなし)	1-15