(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2012-173745
(P2012-173745A)
(43)公開日　平成24年9月10日(2012.9.10)

(51)Int.Cl.		F I		テーマコード（参考）
G06F 12/00	(2006.01)	G06F 12/00	546T	5B082

審査請求 未請求　請求項の数 8　ＯＬ　（全 20 頁）

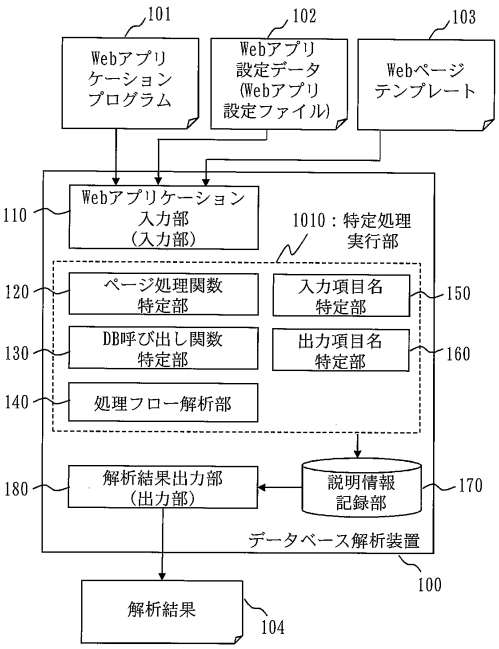
(21)出願番号　特願2011-31645 (P2011-31645)
(22)出願日　平成23年2月17日 (2011.2.17)

(71)出願人　000006013
三菱電機株式会社
東京都千代田区丸の内二丁目７番３号
(74)代理人　100099461
弁理士　溝井　章司
(74)代理人　100122035
弁理士　渡辺　敏雄
(72)発明者　河内　清人
東京都千代田区丸の内二丁目７番３号　三菱電機株式会社内
(72)発明者　桜井　鐘治
東京都千代田区丸の内二丁目７番３号　三菱電機株式会社内

Ｆターム(参考)　5B082 BA09　GA08

(54)【発明の名称】データベース解析装置及びデータベース解析プログラム

(57)【要約】　（修正有）
【課題】各カラムの説明情報を、そのデータベースを利用するＷｅｂアプリケーションプログラムから抽出し、提供する。
【解決手段】データベース解析装置１００は、解析対象のデータベースを使用するＷｅｂアプリケーション１０１、Ｗｅｂアプリケーション１０１の設定に用いるＷｅｂアプリ設定データ１０２、Ｗｅｂページテンプレート１０３を入力するＷｅｂアプリケーション入力部１１０と、Ｗｅｂアプリケーション１０１によって生成されるＷｅｂページに表示される入力フォームの生成に使用される入力フォーム部品であって解析対象データベースの有するテーブルのカラムに対応する入力項目名と対応付けられた入力フォーム部品を、Ｗｅｂアプリケーション１０１と、Ｗｅｂアプリ設定データ１０２と、Ｗｅｂページテンプレート１０３とに基づき特定し、特定した入力フォーム部品に対応する入力項目名を特定する特定処理実行部１０１０と備えた。
【選択図】図１



【特許請求の範囲】
【請求項１】
解析対象のデータベースを使用するＷｅｂアプリケーションプログラムと、前記Ｗｅｂアプリケーションプログラムの設定に使用する所定のＷｅｂアプリ設定情報が記載されたＷｅｂアプリ設定ファイルと、前記ＷｅｂアプリケーションプログラムがＷｅｂページを生成する際の雛型情報を示すＷｅｂページテンプレートとを入力する入力部と、
前記入力部に入力された前記Ｗｅｂアプリケーションプログラムによって生成されるＷｅｂページに表示される入力フォームの生成に使用される入力フォーム部品を、前記Ｗｅｂアプリケーションプログラムと、前記Ｗｅｂアプリ設定ファイルと、前記Ｗｅｂページテンプレートとに基づき特定し、特定した前記入力フォーム部品に対応する前記入力項目名であって前記解析対象のデータベースの有するテーブルのカラムに対応する入力項目名を、前記Ｗｅｂページテンプレートを解析することにより特定する特定処理実行部と、
前記特定処理実行部が特定した前記入力項目名を少なくとも含む情報を、前記Ｗｅｂページに表示される前記入力フォームへ入力される入力データによって更新される前記カラムの説明情報として出力する出力部と
を備えたことを特徴とするデータベース解析装置。
【請求項２】
前記特定処理実行部は、
前記データベースの保有する前記テーブルのカラムから得られるデータが出力される前記Ｗｅｂページにおける出力項目の格納領域を、前記Ｗｅｂアプリケーションプログラムと、前記Ｗｅｂアプリ設定ファイルと、前記Ｗｅｂページテンプレートとに基づき特定し、特定した前記格納領域に対応する出力項目名であって前記データが得られる前記カラムと対応付けられた格納領域の出力項目名を、前記Ｗｅｂページテンプレートを解析することにより特定し、
前記出力部は、
前記特定処理実行部が特定した前記出力項目名を少なくとも含む情報を、前記データの取得元となる前記カラムの説明情報として出力することを特徴とする請求項１記載のデータベース解析装置。
【請求項３】
前記特定処理実行部は、
入力項目が前記テーブルに含まれる場合には、前記テーブルの入力項目に対する見出しに相当する候補である見出し候補を複数抽出し、抽出した複数の前記見出し候補を対象として繰り返し構造が出現するがどうかを判定し、前記繰り返し構造が出現すると判定すると、前記複数の見出し候補の中から前記テーブルの前記入力項目に対する見出しを前記繰り返し構造に基づいて特定し、特定された前記見出しを、前記入力項目名として採用することを特徴とする請求項２記載のデータベース解析装置。
【請求項４】
前記特定処理実行部は、
出力項目が前記テーブルに含まれる場合には、前記テーブルの出力項目に対する見出しに相当する候補である見出し候補を複数抽出し、抽出した複数の前記見出し候補を対象として繰り返し構造が出現するがどうかを判定し、前記繰り返し構造が出現すると判定すると、前記複数の見出し候補の中から前記テーブルの前記出力項目に対する見出しを前記繰り返し構造に基づいて特定し、特定された前記見出しを、前記出力項目名として採用することを特徴とする請求項３記載のデータベース解析装置。
【請求項５】
前記特定処理実行部は、
判定の結果、前記繰り返し構造が出現しないと判定すると、所定の基準に基づいて前記複数の見出し候補の中から少なくとも一つの見出し候補を取り除く取り除き処理を実行し、取り除き処理後の見出し候補を対象として繰り返し構造が出現するがどうかを判定することを特徴とする請求項３または４のいずれかに記載のデータベース解析装置。

【請求項６】
前記特定処理実行部は、
前記Ｗｅｂアプリケーションプログラム内の文字列操作を静的解析で追跡してデータベース呼び出し関数に渡されるＳＱＬ文を特定するＳＱＬ文特定処理を実行すると共に、前記ＳＱＬ文特定処理のときに文字列操作として代入操作と連結操作とのいずれにも該当しない文字列操作がなされたと判定すると、文字列操作に対する正規の生成規則とは異なる仮の生成規則であって前記正規の生成規則に復元可能な仮の生成規則を記録し、前記ＳＱＬ文に基づいてテーブル名、カラム名を特定するに際して前記正規の生成規則が必要な場合にのみ、前記仮の生成規則を前記正規の生成規則に復元することを特徴とする請求項１～５のいずれかに記載のデータベース解析装置。
【請求項７】
前記入力部は、
Ｗｅｂアプリケーションプログラムの格納ディレクトリの情報を示す格納ディレクトリ情報を入力し、
前記データベース解析装置は、さらに、
前記入力部が入力した前記格納ディレクトリ情報を参照することで、前記Ｗｅｂアプリケーションプログラムが使用するフレームワークを特定し、かつ、前記Ｗｅｂアプリケーションプログラムの格納ディレクトリから前記Ｗｅｂアプリ設定情報を抽出するフレームワーク特定部を備えたことを特徴とする請求項１～６のいずれかに記載のデータベース解析装置。
【請求項８】
解析対象のデータベースを使用するＷｅｂアプリケーションプログラムと、前記Ｗｅｂアプリケーションプログラムの設定に使用する所定のＷｅｂアプリ設定情報が記載されたＷｅｂアプリ設定ファイルと、前記Ｗｅｂアプリケーションプログラムが Ｗｅｂページを生成する際の雛型情報を示すＷｅｂページテンプレートとを入力する処理と、
入力された前記Ｗｅｂアプリケーションプログラムによって生成されるＷｅｂページに表示される入力フォームの生成に使用される入力フォーム部品を、前記Ｗｅｂアプリケーションプログラムと、前記Ｗｅｂアプリ設定ファイルと、前記Ｗｅｂページテンプレートとに基づき特定し、特定した前記入力フォーム部品に対応する前記入力項目名であって前記解析対象のデータベースの有するテーブルのカラムに対応する入力項目名を、前記Ｗｅｂページテンプレートを解析することにより特定する処理と、
特定した前記入力項目名を少なくとも含む情報を、前記Ｗｅｂページに表示される前記入力フォームへ入力される入力データによって更新される前記カラムの説明情報として出力する処理と
をコンピュータに実行させるためのデータベース解析プログラム。
【発明の詳細な説明】
【技術分野】
【０００１】
この発明は、データベースを利用するＷｅｂアプリケーションを解析してデータベースのテーブル、カラムの説明情報を抽出するデータベース解析装置及びデータベース解析プログラムに関する。
【背景技術】
【０００２】
一般に製品化されているデータモデリングツールでは、データベース解析機能としてテーブルやカラムの一覧を出力する機能は極めて一般的に提供されている（例えば、非特許文献１）。データベースを解析するだけでは抽出できない情報を、そのデータベースを利用するプログラムを解析することによって、得ようとする技術としては、例えば特許文献１が知られている。特許文献１では、データベースから読み込まれたデータが格納される変数のデータ型や、プログラム内で行われる定数値との比較を調べることで、データベース中の各データ項目のデータ型、上限値、下限値、あるいはループ脱出条件として使用さ

れているかどうか、といった情報を抽出する方法を開示している。
【先行技術文献】
【特許文献】
【０００３】
【特許文献１】国際公開第ＷＯ２００９／０１１０５７号公報
【非特許文献】
【０００４】
【非特許文献１】ＩＢＭ　ＩｎｆｏＳｐｈｅｒｅ　Ｄａｔａ　Ａｒｃｈｉｔｅｃｔ　ｈｔｔｐ：／／ｗｗｗ－０６．ｉｂｍ．ｃｏｍ／ｓｏｆｔｗａｒｅ／ｊｐ／ｄａｔａ／ｏｐｔｉｍ／ｄａｔａ－ａｒｃｈｉｔｅｃｔ／ｆｕｎｃｔｉｏｎ．ｈｔｍｌ
【発明の概要】
【発明が解決しようとする課題】
【０００５】
例えば、資材システムのデータベースに入力された発注金額を、経理システムのデータベース上に反映させる、といった複数のデータベース間のデータ連携システムを開発するには、各データベース間で連携させなければならないテーブル、カラムを特定する作業が必要となる。しかし、従来技術では、データベース内に格納されたテーブルと各テーブル上に存在するカラムの名称（物理名）一覧と、その値域などが抽出されるのみであり、各カラムに格納されるデータの意味（用途）は，カラムやテーブルの名称をもとに仕様書を調査する必要があった。調査には時間がかかる上，仕様書の記載が不十分だったり，内容が古かったりした場合には正しい情報が得られない、といった課題があった。
【０００６】
この発明は、各データベースのカラムの意味を利用者が推測することを容易にするために、各カラムの説明情報をそのデータベースを利用するＷｅｂアプリケーションプログラム（以下、Ｗｅｂアプリケーションという場合もある）から抽出し、提供することを目的とする。
【課題を解決するための手段】
【０００７】
この発明のデータベース解析装置は、
解析対象のデータベースを使用するＷｅｂアプリケーションプログラムと、前記Ｗｅｂアプリケーションプログラムの設定に使用する所定のＷｅｂアプリ設定情報が記載されたＷｅｂアプリ設定ファイルと、前記Ｗｅｂアプリケーションプログラムが Ｗｅｂページを生成する際の雛型情報を示すＷｅｂページテンプレートとを入力する入力部と、
前記入力部に入力された前記Ｗｅｂアプリケーションプログラムによって生成されるＷｅｂページに表示される入力フォームの生成に使用される入力フォーム部品を、前記Ｗｅｂアプリケーションプログラムと、前記Ｗｅｂアプリ設定ファイルと、前記Ｗｅｂページテンプレートとに基づき特定し、特定した前記入力フォーム部品に対応する前記入力項目名であって前記解析対象のデータベースの有するテーブルのカラムに対応する入力項目名を、前記Ｗｅｂページテンプレートを解析することにより特定する特定処理実行部と、
前記特定処理実行部が特定した前記入力項目名を少なくとも含む情報を、前記Ｗｅｂページに表示される前記入力フォームへ入力される入力データによって更新される前記カラムの説明情報として出力する出力部と
を備えたことを特徴とする。
【発明の効果】
【０００８】
この発明により、各カラムの説明情報を、そのデータベースを利用するＷｅｂアプリケーションプログラムから抽出して、利用者に提供することができる。
【図面の簡単な説明】
【０００９】
【図１】実施の形態１のデータベース解析装置１００の構成図。

【図２】実施の形態１のＷｅｂアプリケーションプログラム１０１～Ｗｅｂページテンプレート１０３の概要を示す図。
【図３】実施の形態１のＷｅｂアプリケーション入力部１１０～処理フロー解析部１４０の処理動作を示すフロー。
【図４】実施の形態１の処理フロー解析部１４０の動作を説明する図。
【図５】実施の形態１の入力項目名特定部１５０の動作を示すフロー。
【図６】実施の形態１の入力項目名特定部１５０の動作を説明する図。
【図７】図５のＳ２０４の内容を具体的に説明するフロー。
【図８】図５のＳ２０４の内容を具体的に説明するフロー。
【図９】図５のＳ２０４の内容を具体的に説明するフロー。
【図１０】実施の形態１のデータベース解析装置１００による入力項目名特定の方式が適用される例を示す図。
【図１１】実施の形態１の出力項目名特定部１６０の動作を説明する図。
【図１２】実施の形態１の出力項目名特定部１６０の動作を説明する図。
【図１３】実施の形態１の説明情報記録部１７０によって出力される解析結果を示す図。
【図１４】実施の形態２の処理フロー解析部１４０のフローチャート。
【図１５】実施の形態２の処理フロー解析部１４０の別のフローチャート。
【図１６】実施の形態３のデータベース解析装置３００の構成図。
【図１７】実施の形態４のデータベース解析装置の外観を示す図。
【図１８】実施の形態４のデータベース解析装置のハードウェア構成を示す図。
【発明を実施するための形態】
【００１０】
実施の形態１．
図１は、データベース解析装置１００を示す装置構成図である。図１のように、データベース解析装置１００は、Ｗｅｂアプリケーション入力部１１０（入力部）、ページ処理関数特定部１２０、ＤＢ呼び出し関数特定部１３０、処理フロー解析部１４０、入力項目名特定部１５０、出力項目名特定部１６０、説明情報記録部１７０、および解析結果出力部１８０（出力部）を備える。ページ処理関数特定部１２０～出力項目名特定部１６０は、特定処理実行部１０１０を構成する。なお、以下ではデータベースを「ＤＢ」と表記する場合がある。
【００１１】
Ｗｅｂアプリケーション入力部１１０は、解析対象とするＷｅｂアプリケーションプログラム１０１、Ｗｅｂアプリケーションの設定データであるＷｅｂアプリ設定データ１０２（Ｗｅｂアプリ設定ファイル）、およびＷｅｂページテンプレート１０３を装置内に読み込む。
【００１２】
図２は、Ｗｅｂアプリケーションプログラム１０１～Ｗｅｂページテンプレート１０３の概要を示す図である。
（１）Ｗｅｂアプリケーションプログラム１０１は、ソースコードの形態で入力される。
（２）Ｗｅｂアプリ設定データ１０２には、
（ａ）Ｗｅｂアプリケーションプログラム１０１が受理するＵＲＬの一覧、
（ｂ）ＵＲＬが呼び出されたときに起動される関数名（以下エントリーポイント）、
（ｃ）およびＵＲＬ処理結果として使用すべきＷｅｂページテンプレートの識別子、
が記載されている。なお、本実施の形態１では便宜上、「関数名」と述べているが、必ずしも関数名である必要は無い。例えば、Ｊａｖａ（登録商標）で記述されたＷｅｂアプリケーションでは、ＵＲＬが呼び出されたときに起動されるクラス名さえ判明すれば、呼び出される関数（メソッド）は標準で定められたものが使用される。
（３）Ｗｅｂページテンプレート１０３とは、Ｗｅｂアプリケーションプログラム１０１がＷｅｂページを生成する時の雛形データであり、ＪＳＰ（登録商標）など、一般的なＷｅｂアプリケーションで用いられている技術である。Ｗｅｂアプリケーションプログラム

１０１は、Ｗｅｂページテンプレート１０３を読み込み、その中のパラメータ部分を実行時に生成された値で置換えたデータをブラウザに返す。

【００１３】

（１）ページ処理関数特定部１２０は、図２に示すように、読み込まれたＷｅｂアプリ設定データ１０２から、Ｗｅｂアプリケーションプログラム１０１の各ページを処理するためのエントリーポイントを特定する。

（２）ＤＢ呼び出し関数特定部１３０は、図２に示すように、Ｗｅｂアプリケーションプログラム１０１内に存在するデータベース呼び出し関数を特定する。

（３）処理フロー解析部１４０は、図２に示すように、エントリーポイントからＤＢ呼び出し関数までの、Ｗｅｂアプリケーションプログラム１０１中の処理フローを解析し、エントリーポイントから開始される処理によって更新／参照されるデータベースのテーブル、カラムを特定すると共に、各カラムの入力／出力データに関連するプログラム内の「変数群」を特定する。

【００１４】

（変数群）

ここで、入力データに関連する「変数群」とは、ＤＢ呼び出し関数上でデータベースの更新が発生する場合に、各カラムに対して格納される値に影響を与える変数群を指す。

同様に、出力データに関連する変数群とは、ＤＢ呼び出し関数上でデータベースの検索が発生する場合に、検索結果のうち、各カラムに対応する値が、影響を与える変数群を指す。

【００１５】

入力項目名特定部１５０は、各カラムへの入力データに関連する入力データ関連変数群のうち、解析対象のＷｅｂアプリケーションプログラム１０１上のフォームから入力される変数を特定する。さらに、特定された各変数への入力を与えるフォーム部品が配置されているＷｅｂページテンプレート１０３を解析し、出力されるＷｅｂページのタイトルと、フォーム部品に対応する入力項目名を抽出し、エントリーポイントから開始される一連の処理によって更新されるテーブル上の各カラムに対する説明情報として説明情報記録部１７０に記録する。

【００１６】

同様に、出力項目名特定部１６０は、各カラムからの出力データに関連する変数群のうち、Ｗｅｂページに出力される変数を特定する。さらに、特定された各変数が出力されるＷｅｂページテンプレートを解析し、Ｗｅｂページのタイトルと、出力されるデータを説明する出力項目名を抽出し、エントリーポイントから開始される一連の処理によって参照されるテーブル上の各カラムに対する説明情報として説明情報記録部１７０に記録する。

【００１７】

解析結果出力部１８０は、説明情報記録部１７０によって記録された説明情報を、解析結果１０４として出力する。

【００１８】

図３は、Ｗｅｂアプリケーション入力部１１０～処理フロー解析部１４０の処理動作を示す。次に、図３を参照して、動作を説明する。

（１）まずＷｅｂアプリケーション入力部１１０は、Ｗｅｂアプリケーションプログラム１０１、Ｗｅｂアプリ設定データ１０２、Ｗｅｂページテンプレート１０３を装置内に読み込む（Ｓ１１）。

（２）次に、ページ処理関数特定部１２０は、入力されたＷｅｂアプリ設定データ１０２の情報を用いて、各ページのＵＲＬ、各ページのエントリーポイント、及び関数実行後に表示されるＷｅｂページテンプレートの一覧を作成する（Ｓ１２）。

（３）次に、ＤＢ呼び出し関数特定部１３０は、Ｗｅｂアプリケーションプログラム内を検索し、ＤＢ呼び出し関数の出現位置を特定する（Ｓ１３）。ここで、「ＤＢ呼び出し関数」とは、データベースに対してＳＱＬ文を発行するためにＯＳやフレームワークといったプラットフォームが提供する組み込み関数である。ＤＢ呼び出し関数は、引数としてＳ

ＱＬ文を受け取り（後述の図４）、実行結果（検索の結果得られたデータ等）を返す関数を指す。ＤＢ呼び出し関数の関数名は、あらかじめ定義されているものとする。

【００１９】

処理フロー解析部１４０は、ページ処理関数特定部１２０によって特定された各ページのエントリーポイント、およびＤＢ呼び出し関数特定部１３０によって特定されたプログラム内のＤＢ呼び出し関数一覧を入力し、Ｗｅｂアプリケーションプログラム内の、各ページエントリーポイントからＤＢ呼び出し関数を呼び出す可能性があるかを判定（Ｓ１４－１）する。

図４は、処理フロー解析部１４０の動作を説明する図である。可能性がある場合には、処理フロー解析部１４０は、図４に示すように、エントリーポイント内で行われる文字列操作（代入、連結、置換等）を追跡し、ＤＢ呼び出し関数に渡される引数のうち、ＳＱＬ文を指定する引数に代入される文字列変数が、どのような文字列を格納するかを特定する（Ｓ１４－２）。なお、変数に格納される文字列は、プログラムの分岐や、あるいは外部から入力された値によって不定となる箇所もあるため、一般には、変数が格納し得る文字列全体を表すパターンが特定される。

【００２０】

（コマンド文字列の抽出）

処理フロー解析部１４０は、ＳＱＬ文を指定する引数に代入される文字列のパターンを特定すると、あらかじめ定義されたＳＱＬ文法データと特定した文字列パターンとを比較し、ＳＱＬ文法上、コマンドに相当する文字列を抽出する（Ｓ１４－３）。

【００２１】

（更新命令の場合）

コマンドが、ＩＮＳＥＲＴやＵＰＤＡＴＥといった、データベースの内容を更新する命令であった場合には、処理フロー解析部１４０は、同様にＳＱＬ文法と比較することでテーブル名、カラム名を特定し、さらに各カラムに代入される値に影響を与える変数群を特定する（Ｓ１４－４）。

【００２２】

なお、本実施の形態１では、処理フロー解析部１４０における文字列の特定、およびＳＱＬ文法との比較処理に、下記の＜参考文献＞で開示されている方式を利用することが可能である。詳細については割愛するが、同文献で開示された方式では、エントリーポイントの入口を起点とした制御フローグラフを生成し、制御フロー上に現れる文字列操作を、文字列が格納される変数を非終端記号とした文脈自由文法として表現することで、各文字列変数の値が満たすべき「文法」を特定する。特定された文法を調べることで、任意の文字列変数上で生成される文字列のパターンや、文字列の特定の部分に影響を与える変数群を特定することが可能である。

【００２３】

＜参考文献＞

Ｇ．Ｗａｓｓｅｒｍａｎｎ ａｎｄ Ｚ．Ｓｕ．Ｓｏｕｎｄ ａｎｄ Ｐｒｅｃｉｓｅ Ａｎａｌｙｓｉｓ ｏｆ Ｗｅｂ Ａｐｐｌｉｃａｔｉｏｎｓ ｆｏｒ Ｉｎｊｅｃｔｉｏｎ Ｖｕｌｎｅｒａｂｉｌｉｔｉｅｓ．Ｉｎ Ｃｏｎｆｅｒｅｎｃｅ ｏｎ Ｐｒｏｇｒａｍｍｉｎｇ Ｌａｎｇｕａｇｅ Ｄｅｓｉｇｎ ａｎｄ Ｉｍｐｌｅｍｅｎｔａｔｉｏｎ（ＰＬＤＩ），２００７．

【００２４】

（参照命令の場合）

再び動作の説明に戻る。処理フロー解析部１４０は、コマンドがＳＥＬＥＣＴのように、データベースの内容を参照する命令であった場合には、先ほどと同様、テーブル名を求めた後、ＳＱＬ検索結果を取得する関数にカラム名として渡される文字列を特定し、さらに、同関数の戻り値によって影響される変数群を特定する（Ｓ１４－５）。ＳＱＬ検索結果取得関数の戻り値に影響される変数群を特定するには、プログラム上で、同関数の処理が終了した箇所を起点とし、戻り値が最初に代入された変数を右辺に持つ代入文を追跡す

ればよい。

【００２５】

処理フロー解析部１４０により解析した結果（入出力の変数群）は、次に入力項目名特定部１５０、および出力項目名特定部１６０によって、Ｗｅｂページ上の入力項目名、および出力項目名と対応付けられる。はじめに、図５、図６を参照しながら入力項目名特定部１５０の動作から説明する。

図５、図６は入力項目名特定部１５０の動作を示すフローチャート、概念図である。おもに図６を参照して説明する。

【００２６】

（入力項目名特定部１５０の動作）

入力項目名特定部１５０は、まず、Ｗｅｂページ上のフォームから入力された値が代入される変数を特定する（ステップ２０１）。この特定処理は、各フォーム部品からの入力値を取得する関数をエントリーポイントからの制御グラフ上で探し、戻り値が代入される変数を特定することで容易に実施できる。また、どのフォーム部品からの入力値かも、関数名等から容易に特定可能である。例えば、Ｗｅｂアプリケーション開発フレームワークとして広く使用されているｓｔｒｕｔｓでは、Ｗｅｂページテンプレート上で“ｕｓｅｒ”と名づけられたフォーム部品からの入力値は、ｇｅｔＵｓｅｒ（　）という関数で取得する仕組みとなっており、関数名からｇｅｔを取り除けば、対応するフォーム部品名を取得することができる。

【００２７】

次に、入力項目名特定部１５０は、各フォーム部品からの入力値が代入される変数のうち、データベースの更新に影響を与える変数（更新影響変数ともいう）を特定する（ステップ２０２）。その後、入力項目名特定部１５０は、各変数によって影響を受けるテーブル名およびカラム名を、処理フロー解析部１４０の処理結果（Ｓ１４－４で特定された変数群）と比較することで特定する（ステップ２０３）。これにより、更新に影響を与える変数に値を代入するフォーム部品が特定できるため、このフォーム部品に対応する入力項目名を、Ｗｅｂページテンプレート１０３を解析することで特定し（ステップ２０４）、特定結果を説明情報記録部１７０に蓄積する（ステップ２０５）。

【００２８】

解析対象とするＷｅｂページテンプレート１０３は、現在解析対象としているエントリーポイントに対応するＵＲＬをａｃｔｉｏｎ属性に持つフォームを含んだＷｅｂページテンプレートが選択される。入力項目名特定部１５０がＷｅｂページテンプレートから入力項目名を特定する処理（ステップ２０４）について、図７～図９を参照しながら、詳細に説明する。図７、図８、図９の動作主体は入力項目名特定部１５０である。

【００２９】

はじめに、当該テンプレート内の＜ｔｉｔｌｅ＞～＜／ｔｉｔｌｅ＞に囲まれたテキストを、そのテンプレートによって生成されるＷｅｂページのページ名として記憶する（ステップ３０１）。次に当該フォーム部品がＷｅｂページテンプレートのＨＴＭＬ上、ｔａｂｌｅ要素に含まれているかを確認する（ステップ３０２）。もしｔａｂｌｅ要素に囲まれていないならば、そのフォーム部品を含む行の先頭にある単語を入力項目名とみなす（ステップ３１４）。

【００３０】

ｔａｂｌｅ要素に囲まれていた場合、ｔａｂｌｅ内の各セルに記載された文字列のうち、当該フォーム部品に対応する入力項目名を見つける必要がある。本実施の形態では、次のように入力項目名の特定を実施する。

【００３１】

はじめに、ｔａｂｌｅの各セル（＜ｔｄ＞～＜／ｔｄ＞で囲まれた範囲）のうち、ｒｏｗｓｐａｎ属性が２以上であるセルは、行毎に分解し、各行に同じ値（テキスト、フォーム部品、Ｗｅｂページテンプレート上で出力される際に格納される領域（以下Ｗｅｂアプリケーション出力値格納領域））を含んだセルがｒｏｗｓｐａｎ分だけ並んでいる表へと

変形する（ステップ３０３）。

【００３２】

次に、Ｌ＝１とし（ステップ３０４）、Ｌ＝テーブル行数となるまで以下の処理を繰り返す（ステップ３０５）。

【００３３】

まず、Ｌ行目から、フォーム部品を含まない行を連続して取れる限り取り出す（ステップ３０７）。これらを見出し候補行と呼ぶことにする。もし見出し候補行が存在しなければステップ３５３へ進む。取り出された見出し候補行（行数をＮとする）の下からｓ行目（１≦ｓ≦Ｎ）からｍ行分（１≦ｍ≦Ｎ－ｓ＋１）と同じ構造が、それ以降の行で繰り返し現れる最大回数ｒを求める（ステップ３０９から３１３、３２０から３３３）。同じ構造かどうかは、次の条件をともに満たしている場合とする。

【００３４】

（１）各セルのｃｏｌｓｐａｎ属性値が見出し候補行の対応するセルと同じ値である（ｃｏｌｓｐａｎが無い場合はｃｏｌｓｐａｎ＝１とみなす）（ステップ３２２）。

（２）一回目にマッチしたセルと同じフォーム部品（あるいはテキスト、又はＷｅｂアプリケーション出力値格納領域）が出現する（ステップ３２３，３２４）。なお、テキストは異なっていても構わないものとする。

【００３５】

全てのｓ，ｍに対し、上記処理によってｒを求め、ｍとｒを積算した結果が最大となるｍを見出し行数とする（ステップ３２８，３２９）。もし複数のｍでｍ・ｒが同じ値になった場合には、ｍが最大のものを見出し行数とみなす。対応するｓについては、下からｓ－１行分の見出し候補行は、注釈等、見出しとは無関係なテキストとみなす。もし、全てのｓ，ｍでｒが０になった場合（ステップ３４０）はステップ３５１へ進む。

【００３６】

最後に、繰り返し構造に含まれた各フォーム部品全てに対して、ステップ３４１から３５０に示す方法で入力項目名を与える。ステップ３４７で示される項目名の付与の方法は次の通りである。まず入力項目名Ｗ＝空文字列とする。同一セルの先頭にテキストＷ０が含まれている場合には、Ｗ＝Ｗ０とする。次に、見出し候補行の中から対応するセル（（Ｌ－１）＋ｋ＋Ｎ－（Ｍ＋Ｓ）＋１行、ｃｏｌ列）のテキストＷ１をスラッシュ記号等、適当な区切り文字で付加する。その後、繰り返し構造に現れない見出し候補行の各テキストのうち、同じ列に属するテキストを順に付加していく。ただし、連続して同じテキストが現れる場合は、付加するのは１度だけとし、残りは単に無視されるものとする。

【００３７】

以上の処理を実行後、Ｌ＝Ｌ＋Ｎ＋Ｒ・Ｍとし、ループの先頭に戻る（ステップ３５２）。ステップ３５３～３５８では、表中に見出し候補行が見つけられなかった場合の処理を行う。このような場合、各行は、独立した入力項目を示しているとみなす。ステップ３５６で示される入力項目名の付与の方法は次の通りである。まず入力項目名Ｗ＝空文字列とする。同一セルの先頭にテキストＷ０が含まれている場合には、Ｗ＝Ｗ０とする。次に、セルの左側のセルを順に走査して行き、最初に現れるテキストのみを含んだセルの内容Ｗ１をスラッシュ記号等、適当な区切り文字で付加する。その後、さらに左側のセルを走査していき、次にフォーム部品を含んだセルが現れるか、先頭列に達するまで文字列を順に付加していく。

【００３８】

以上の処理を、再び全てのセルがテキストである行にあたるまで（ステップ３５５）、もしくは表全体の処理を終了するまで（ステップ３５４）繰り返し、その後Ｌ＝ｉに更新後（ステップ３５８）、ループの先頭に戻る。

【００３９】

ループ終了後、Ｗｅｂページテンプレート上の各フォーム部品に対して入力項目名が定まるため、入力項目名を特定する対象であるフォーム部品に対する入力項目名を、当該フォーム部品に対応するデータベースのテーブル名、カラム名、およびテンプレートのペー

ジ名とともに説明情報記録部に保存する（ステップ３１５）。
【００４０】
データベース解析装置１００による本方式を図１０に示す入力フォームに対する入力項目名特定に適用した場合を例とすると、図１０のフォーム部品４０１には、“発注明細／参考ＵＲＬ”という入力項目名が付与される。同様にフォーム部品４０２については、“発注明細／数量”が付与される。フォーム部品４０３では、“納期”が付与され、フォーム部品４０４では“特記事項”が付与される。
【００４１】
（出力項目名特定部１６０の動作）
次に、図１１、図１２を参照して、出力項目名特定部１６０の動作について説明する。
図１１、図１２は出力項目名特定部１６０の動作を示すフローチャート、概念図である。
（１）まず、出力項目名特定部１６０は、Ｗｅｂページ上に出力される変数（出力変数）を特定する（Ｓ２１）。この特定処理は、Ｗｅｂページへの出力値を設定する関数をエントリーポイントからの制御グラフ上で探し、その引数として入力される変数を特定することで容易に実施できる。また、出力されるデータがＷｅｂページテンプレート上でどのような識別名で参照されるかも、関数名等から容易に特定可能である。例えば、ｓｔｒｕｔｓでは、Ｗｅｂページテンプレート上で“ｐｒｉｃｅ”として参照されるデータは、ｓｅｔＰｒｉｃｅ（）という関数で取得する仕組みとなっており、関数名からｓｅｔを取り除けば、対応する出力データ名（変数）を取得することができる。
（２）次に、処理フロー解析部１４０の処理結果と比較することで、各Ｗｅｂページへ出力される出力変数のうち、データベースの検索結果の影響を受ける変数（検索影響変数ともいう）と、影響を受けるカラム名を特定する（Ｓ２２）。これにより、検索結果に影響を受ける検索影響変数に対するＷｅｂアプリケーション出力値格納領域（格納領域）が特定できる（Ｓ２３）。このため、出力項目名特定部１６０は、特定したＷｅｂアプリケーション出力値格納領域に対応する出力項目名を、Ｗｅｂページテンプレートを解析することで特定する（Ｓ２４）。解析対象とするＷｅｂページテンプレートは、現在解析対象としているエントリーポイントの実行結果が出力されるＷｅｂページテンプレートが選択される。すなわち、入力項目名特定部１５０と全く同様な方法（図７～９）で、各出力データに対応する出力項目名を特定する（Ｓ２４）。ただし、入力項目名特定部１５０における前記説明（図７～９）の「フォーム部品」は「Ｗｅｂアプリケーション出力値格納領域」に、「Ｗｅｂアプリケーションからの出力値格納領域」は「フォーム部品」に読み替えた方式となる。
【００４２】
Ｗｅｂアプリケーションプログラム１０１の全てのエントリーポイントに対する解析が終了すると、解析結果出力部１８０が、説明情報記録部１７０に蓄えられた解析結果をユーザに出力する。
図１３は、説明情報記録部１７０によって出力される解析結果を示す。図１３に示すように、テーブル名５０１、カラム名５０２、各カラムに対する入力ページ名５０３、入力項目名５０４、出力ページ名５０５、及び出力項目名５０６である。表中、ｔａｂｌｅ２のカラムｘ２は出力ページ名、出力項目名が記載されていないが、これは同カラムのデータがＷｅｂページ上に現れなかったことを意味するものである。
【００４３】
（１）以上のように、Ｗｅｂアプリケーション内で発生するデータベースアクセスで更新／参照されるカラムに対応するＷｅｂアプリケーションの入出力データに対し、対応する入出力項目名をＷｅｂページテンプレートを解析して特定するようにしているので、データベースの各カラムに対し、そのカラムの意味を利用者が推測するための説明情報を出力することができる。
（２）また、各入出力項目が、Ｗｅｂページテンプレート上でテーブルの中に現れた場合に、テーブル内の見出しに該当するセルを特定することで、表中に入出力項目が現れた場

合にも正しく項目名を識別できるという効果がある。
（３）さらに、見出しの特定において、表の繰り返し構造に着目することで、複数行からなる見出しにも対応することができるという効果がある。
（４）さらに、複数行にわたり連結されているセルを、あらかじめ行毎に分解しておくことで、階層の異なる複数の見出しを含んだ表にも対応できる。
（５）さらに、見出し候補行の下の行からなるべく大きな構造の繰り返しを特定することで、実際の入出力項目を含んだ行には現れない、見出しの見出しを特定することができる。
（６）さらに、上記繰り返し構造を特定する処理において、繰り返し構造が見出せなかった場合、見出し候補行の下から順に取り除いて再び繰り返し構造の特定処理を実施することで、注意書き等、見出しとは無関係な行が挿入されている場合に対応することができる。
（７）さらに、繰り返し構造が見つけられなかった場合には、各フォーム部品に対し、左側のセルを走査していき、最初に見つかる連続したテキストセルを見出しとみなす事で、横方向に見出しを持った表に対応することができる。
（８）さらに、繰り返し構造が途切れた箇所から別途解析を再開することで、複数の構造を持った複雑な表に対しても見出しを特定することができる。
（９）さらに、複数行、又は複数列からなる見出しが特定された場合には、見出しのテキストを区切り文字で連結することで、階層化された入出力項目名とすることができる。

【００４４】

本実施の形態１ではデータベースの参照／更新と関連する入出力項目を特定してから、Ｗｅｂページテンプレート上で、これらの項目に対する項目名を特定することとしたが、データベースへの参照／更新が発生することが判明した時点で、解析対象とすべきＷｅｂページテンプレート上の全ての項目に対し、項目名を計算しておくこともちろん可能である。

【００４５】

なお、ＷｅｂアプリケーションとしてＰＨＰが使用されている場合、各ＵＲＬに対応して起動されるＰＨＰスクリプトの先頭行から順にコードが実行されるため関数の特定は不要である。さらに、ＰＨＰでは、処理プログラムと、処理結果を表示するためのＷｅｂページテンプレートが同一ファイル上で記載される上、ＵＲＬと起動されるＰＨＰスクリプトとの対応はＰＨＰの格納ディレクトリから判別できる。結果としてＰＨＰにおいては、設定データに記載すべき情報は自動で抽出することも可能である。

【００４６】

実施の形態２．

図１４、図１５を参照して実施の形態２を説明する。実施の形態２は、処理フロー解析部１４０に関する。図１４、図１５の動作主体は処理フロー解析部１４０である。以上の実施の形態１では、処理フロー解析部１４０に第１方式（図１４）の静的解析技術を適用したものであるが、次に、第１方式の解析処理を高速化する場合に、ＳＱＬ文の特定における不要な処理を省略する「第２方式」（図１５）を説明する。

【００４７】

まず、図１４を参照しながら第１方式の概要について説明する。第１方式ではまず、Ｗｅｂアプリケーションプログラムを解析し、エントリーポイントを起点とした制御フローグラフを生成する（ステップ６０１）。制御フローグラフとは、プログラム中の命令の実行経路をグラフ化したものである。次に、制御フローグラフ上にある文字列操作（代入、連結、置換等）に対応し、形式言語理論で言うところの生成規則を記録していく（ステップ６０２～６０９）。
（１）文字列操作として代入操作があった場合、代入先変数名を非終端記号とし、代入元の変数、もしくは文字列定数を右辺に持つ生成規則Ａ→Ｘ（Ｘは変数名、もしくは文字列定数）を生成する（ステップ６０３、６０８）。
（２）文字列操作として連結操作があった場合、連結後の文字列が代入される変数名を非

終端記号とし、連結される複数の変数、あるいは文字列定数を右辺に持つ生成規則Ａ→ＸＹ（Ｘ，Ｙは変数名もしくは文字列定数）を生成する（ステップ６０４、６０７）。
（３）その他の場合、つまり第１方式の文字列に変更が行われる操作、例えば文字列置換などは、当該操作によって入力として与えた文字列の生成規則が変換されると考え、当該文字列操作の内容を反映したトランスデューサと呼ばれる有限オートマトンによって、入力変数に対して記録された第１方式の生成規則を変換した上で、生成規則Ａ←Ｘ（Ｘは変換関数への入力変数、あるいは文字列定数）を生成する（ステップ６０５、６０６）。

【００４８】

以上の処理を繰り返すことで、制御フローグラフ上の全ての文字列変数に対する「文法」を定義することができる。最後に、ＤＢ呼び出し関数に引数として渡される変数に対応する非終端記号を開始記号とする文法と、ＳＱＬ文の文法とを比較することで、ＳＱＬ文のコマンドや、カラム名、およびテーブル名に該当する文字列を特定する。さらに、コマンドがｉｎｓｅｒｔやｕｐｄａｔｅといったデータベースを更新するコマンドであった場合には、データベースの各カラムに投入される値に関係している変数群を特定する（ステップ６１０）。

【００４９】

第１方式において、前記（３）に該当する処理、すなわちトランスデューサによる文法変換処理は、計算コストが高い。そこで、本実施の形態２では、（３）の処理を必要最小限に留めることで処理の高速化を実現する。実施の形態２における第２方式の処理について、図１５を参照しながら説明する。

【００５０】

第２方式においても、制御フローグラフを生成し、フローグラフ上の各文字列操作に対応して生成規則を記録していく点は第１方式と同様である（ステップ７０１～７０４、７０７、および７０８）。ただし、文字列操作が（３）に該当する場合だけ第１方式とは異なる処理を行う。第１方式は、入力として与えられた変数に対する生成規則を変換していたが、第２方式では、代わりに特殊な生成規則Ａ←［ダミー］（［ダミー］は便宜上の表記であり、識別可能な記号であれば何でも構わない）を記録する（ステップ７０５）。また、本来行うべき文法変換に必要な情報、すなわち、入力変数名、変換に必要なトランスデューサもあわせて記録しておく（ステップ７０６）。

【００５１】

その後、第１方式と同様に、ＤＢ呼び出し関数に引数として渡される変数に対応する非終端記号を開始記号とする文法と、ＳＱＬ文の文法とを比較する（ステップ７１０）が、この時に、文法上カラム名やテーブル名が指定されるべき箇所に［ダミー］が現れていないか確認する（ステップ７１１）。もしも現れていた場合には、当該［ダミー］に対応する入力変数名と、トランスデューサを取り出し、文法変換（正規の生成規則の復元の一例）を行った上で、再度ＳＱＬ文法との比較を実施する（ステップ７１２）。以上の処理を繰り返し、最終的にカラム名、およびテーブル名に対応する箇所が全て取得できたら処理を終了させる。最後に、コマンドがｉｎｓｅｒｔやｕｐｄａｔｅといったデータベースを更新するコマンドであった場合には、データベースの各カラムに投入される値に関係している変数群を特定するが、これは［ダミー］がＸに依存しているという情報を利用すれば、投入される値に［ダミー］が含まれていたとしても第１方式と同様に特定可能である。

【００５２】

以上のように、文字列変換操作に対して仮の生成規則を記録しておき、カラム名、テーブル名を抽出するために必要な部分に限って文法変換処理を実施することで、処理フロー解析部を高速化することができる。もしも、全てのテーブル名、カラム名に対して文字列変換処理が発生した場合、第２方式は第１方式と同じだけの文法変換処理を行う必要がある。しかし、一般に文字列変換処理が発生するのはデータベースに記録されるデータや、検索文字列部分であり、アクセス対象であるテーブル名、カラム名に対し文字列変換が発生することは稀である。従って、殆どの場合第２方式を適用することで第１方式よりも高速に処理を行うことが可能である。

【００５３】
実施の形態３．
図１６を参照して実施の形態３を説明する。以上の実施の形態１、２では、Ｗｅｂアプリケーション設定データをページ処理関数特定部１２０が解析可能である必要があるが、本実施の形態３では、フレームワーク特定部３１０が、対象Ｗｅｂアプリケーションが開発に使用したフレームワークを特定する。この特定により、ページ処理関数特定部１２０が解析可能な設定データを生成、入力することで、様々なＷｅｂアプリケーション設定データの形式に対応できる。以下にその形態を示す。
【００５４】
図１６は実施の形態３のデータベース解析装置３００の構成図である。データベース解析装置３００は図１に示す形態１のデータベース解析装置１００に対して、さらに、フレームワーク特定部３１０、フレームワーク特定ルール蓄積部３２０、設定データ生成機能蓄積部３３０を付加した。
【００５５】
フレームワーク特定部３１０は、ユーザの入力したＷｅｂアプリケーションのＷｅｂアプリ格納ディレクトリ名１０５－３の中のファイルを調べ、格納されているファイル名などからＷｅｂアプリケーションの開発に使用されたフレームワーク名を特定する。特定にはフレームワーク特定ルール蓄積部３２０に格納されたルールを使用する。フレームワークが特定されると、フレームワーク特定部３１０は、設定データ生成機能蓄積部３３０に蓄えられた設定データ生成機能３３１のうち、特定されたフレームワークに対応した設定データ生成機能３３１を呼び出す。設定データ生成機能３３１は、各フレームワークに固有の方式でＷｅｂアプリケーション格納ディレクトリ中のファイルから設定データに必要な情報を抽出し、ページ処理関数特定部１２０へ入力する。
【００５６】
以上のように、フレームワーク特定部３１０がＷｅｂアプリケーション格納ディレクトリ内を調査してフレームワークを特定する。そして、設定データ生成機能３３１が、フレームワークに固有の方式に従ってＷｅｂアプリケーション格納ディレクトリ内のファイルから情報を抽出し、Ｗｅｂアプリケーション設定データを生成することで、様々なフレームワークに対応することができる、という効果がある。
【００５７】
実施の形態４．
図１７、図１８を参照して実施の形態３を説明する。実施の形態３は、コンピュータであるデータベース解析装置１００、３００のハードウェア構成を説明する。
図１７は、データベース解析装置１００、３００の外観の一例を示す図である。
図１８は、データベース解析装置１００、３００のハードウェア資源の一例を示す図である。
データベース解析装置１００とデータベース解析装置３００とは同様であるので、データベース解析装置１００を例に説明する。
【００５８】
外観を示す図１７において、データベース解析装置１００は、システムユニット８３０、ＣＲＴ（Ｃａｔｈｏｄｅ・Ｒａｙ・Ｔｕｂｅ）やＬＣＤ（液晶）の表示画面を有する表示装置８１３、キーボード８１４（Ｋｅｙ・Ｂｏａｒｄ：Ｋ／Ｂ）、マウス８１５、コンパクトディスク装置８１８（ＣＤＤ：Ｃｏｍｐａｃｔ　Ｄｉｓｋ　Ｄｒｉｖｅ）などのハードウェア資源を備え、これらはケーブルや信号線で接続されている。システムユニット８３０はネットワークに接続している。
【００５９】
またハードウェア資源を示す図１８において、データベース解析装置１００は、プログラムを実行するＣＰＵ８１０（Ｃｅｎｔｒａｌ　Ｐｒｏｃｅｓｓｉｎｇ　Ｕｎｉｔ）を備えている。ＣＰＵ８１０は、バス８２５を介してＲＯＭ（Ｒｅａｄ　Ｏｎｌｙ　Ｍｅｍｏｒｙ）８１１、ＲＡＭ（Ｒａｎｄｏｍ　Ａｃｃｅｓｓ　Ｍｅｍｏｒｙ）８１２、表示装置

８１３、キーボード８１４、マウス８１５、通信ボード８１６、ＣＤＤ８１８、磁気ディスク装置８２０と接続され、これらのハードウェアデバイスを制御する。磁気ディスク装置８２０の代わりに、光ディスク装置、フラッシュメモリなどの記憶装置でもよい。

【００６０】

ＲＡＭ８１２は、揮発性メモリの一例である。ＲＯＭ８１１、ＣＤＤ８１８、磁気ディスク装置８２０等の記憶媒体は、不揮発性メモリの一例である。これらは、「記憶装置」あるいは記憶部、格納部、記録部、バッファの一例である。通信ボード８１６、キーボード８１４などは、入力部、入力装置の一例である。また、通信ボード８１６、表示装置８１３などは、出力部、出力装置の一例である。通信ボード８１６は、ネットワークに接続されている。

【００６１】

磁気ディスク装置８２０には、オペレーティングシステム８２１（ＯＳ）、ウィンドウシステム８２２、プログラム群８２３、ファイル群８２４が記憶されている。プログラム群８２３のプログラムは、ＣＰＵ８１０、オペレーティングシステム８２１、ウィンドウシステム８２２により実行される。

【００６２】

上記プログラム群８２３には、以上の実施の形態の説明において「～部」として説明した機能を実行するプログラムが記憶されている。プログラムは、ＣＰＵ８１０により読み出され実行される。

【００６３】

ファイル群８２４には、以上の実施の形態の説明において、「～の判定結果」、「～の算出結果」、「～の抽出結果」、「～の生成結果」、「～の処理結果」として説明した情報や、データや信号値や変数値やパラメータなどが、「～ファイル」や「～データベース」の各項目として記憶されている。「～ファイル」や「～データベース」は、ディスクやメモリなどの記録媒体に記憶される。ディスクやメモリなどの記憶媒体に記憶された情報やデータや信号値や変数値やパラメータは、読み書き回路を介してＣＰＵ８１０によりメインメモリやキャッシュメモリに読み出され、抽出・検索・参照・比較・演算・計算・処理・出力・印刷・表示などのＣＰＵの動作に用いられる。抽出・検索・参照・比較・演算・計算・処理・出力・印刷・表示のＣＰＵの動作の間、情報やデータや信号値や変数値やパラメータは、メインメモリやキャッシュメモリやバッファメモリに一時的に記憶される。

【００６４】

また、以上に述べた実施の形態の説明において、データや信号値は、ＲＡＭ８１２のメモリ、ＣＤＤ８１８のコンパクトディスク、磁気ディスク装置８２０の磁気ディスク、その他光ディスク、ミニディスク、ＤＶＤ（Ｄｉｇｉｔａｌ・Ｖｅｒｓａｔｉｌｅ・Ｄｉｓｋ）等の記録媒体に記録される。また、データや信号は、バス８２５や信号線やケーブルその他の伝送媒体によりオンライン伝送される。

【００６５】

また、以上の実施の形態の説明において、「～部」として説明したものは、「～手段」、であってもよく、また、「～ステップ」、「～手順」、「～処理」であってもよい。すなわち、「～部」として説明したものは、ソフトウェアのみ、或いは、ソフトウェアとハードウェアとの組み合わせ、さらには、ファームウェアとの組み合わせで実施されても構わない。ファームウェアとソフトウェアは、プログラムとして、磁気ディスク、フレキシブルディスク、光ディスク、コンパクトディスク、ミニディスク、ＤＶＤ等の記録媒体に記憶される。プログラムはＣＰＵ８１０により読み出され、ＣＰＵ８１０により実行される。すなわち、プログラムは、以上に述べた「～部」としてコンピュータを機能させるものである。あるいは、以上に述べた「～部」の手順や方法をコンピュータに実行させるものである。

【００６６】

以上の実施の形態では、データベース解析装置を説明したが、データベース解析装置の

各「～部」の動作を、コンピュータに実行させるためのデータベース解析プログラムとして把握することもできる。あるいは、データベース解析装置の各「～部」の動作を、データベース解析方法として把握することもできる。

【００６７】

以上の実施の形態では以下を特徴とするデータベース解析装置を説明した。

（１）データベース解析装置は、解析対象のデータベースを使用するＷｅｂアプリケーションのソースコードおよび設定ファイル、Ｗｅｂページテンプレートを入力とし、Ｗｅｂアプリケーションの各ページ上の入力フォーム部品を特定し、同フォーム部品に対応する入力項目名を特定して、同入力データによって更新されるデータベースカラムの説明情報として出力する。

（２）データベース解析装置は、さらに、データベースのあるカラムから得られたデータが出力されるＷｅｂページ上の出力項目を特定し、同項目に対応する出力項目名を特定して、当該カラムの説明情報として出力する。

（３）データベース解析装置は、入力／出力項目名の特定を行うにあたり、入力／出力項目がテーブルに含まれていた場合には、テーブルの各項目に対する見出しに相当する部分を、テーブル内の繰り返し構造から特定し、特定された見出しを入力／出力項目名とする。さらにデータベース解析装置は、以下の（ａ）～（ｅ）を特徴とする。

（ａ）データベース解析装置は、複数行にわたり連結されているセルを、あらかじめ行毎に分解しておくことで、階層の異なる複数の見出しを含んだ表にも対応できる。

（ｂ）データベース解析装置は、見出し候補行の下の行からなるべく大きな構造の繰り返しを特定することで、実際の入出力項目を含んだ行には現れない、見出しの見出しを特定することができる。

（ｃ）データベース解析装置は、繰り返し構造を特定する処理において、繰り返し構造が見出せなかった場合、見出し候補行の下から順に取り除いて再び繰り返し構造の特定処理を実施することで、注意書き等、見出しとは無関係な行が挿入されている場合に対応する。

（ｄ）データベース解析装置は、繰り返し構造が現れなかった行に対しては左側のセルを走査していき、最初に見つかる連続したテキストセルを見出しとみなす。

（ｅ）データベース解析装置は、繰り返し構造が途切れた箇所から別途解析を再開することで、複数の構造を持った複雑な表に対しても見出しを特定することができる。

（ｆ）データベース解析装置は、複数行、又は複数列からなる見出しが特定された場合には、見出しのテキストを区切り文字で連結することで、階層化された入出力項目名とする。

（４）データベース解析装置は、Ｗｅｂアプリケーションプログラム内の文字列操作を静的解析で追跡し、ＤＢ呼び出し関数に渡されるＳＱＬ文を特定する際に、解析に時間のかかる文字列変換に対する追跡処理は省略し、テーブル名、カラム名を特定するために最低限必要な箇所のみ上記追跡処理を実施する。

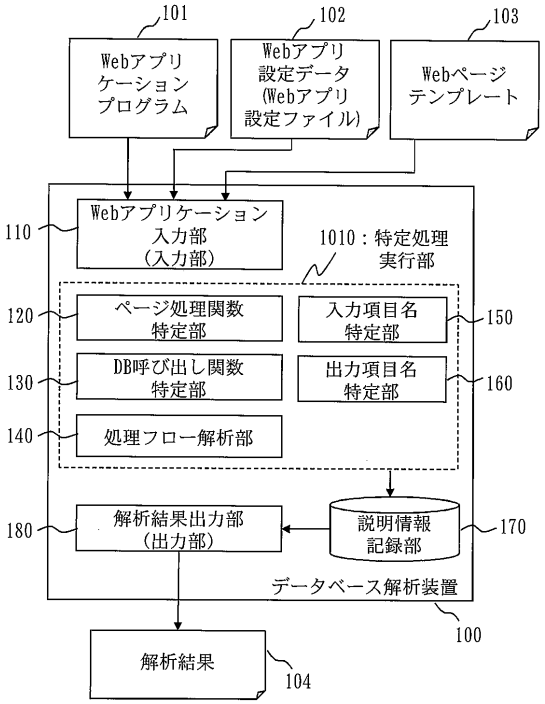
（５）データベース解析装置は、Ｗｅｂアプリケーションが使用している開発フレームワークを特定することで、Ｗｅｂアプリケーション格納ディレクトリから、解析に必要なＷｅｂアプリケーション設定データを抽出する。

【符号の説明】

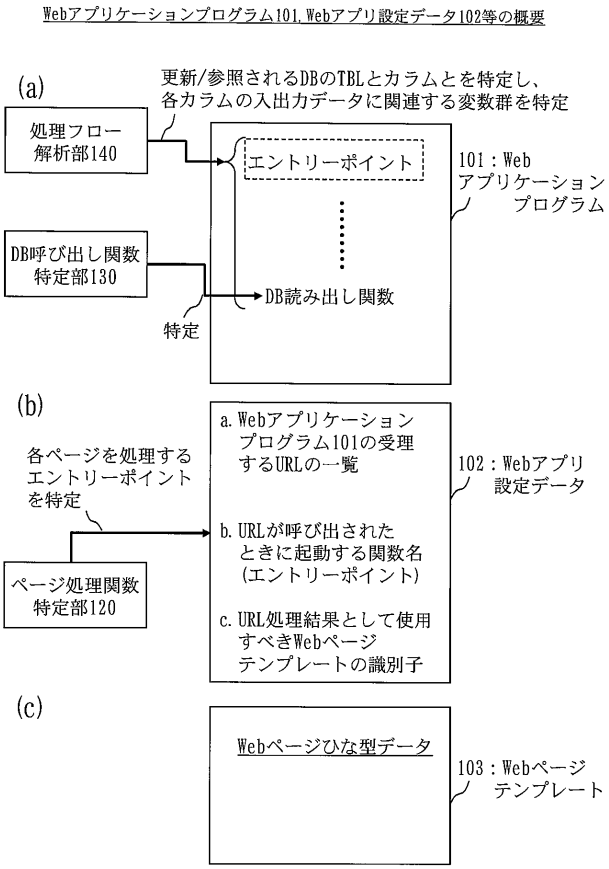
【００６８】

１００，３００　データベース解析装置、１０１　Ｗｅｂアプリケーションプログラム、１０２　Ｗｅｂアプリ設定データ、１０３　Ｗｅｂページテンプレート、１０４　解析結果、１０５－３　Ｗｅｂアプリ格納ディレクトリ名、１１０　Ｗｅｂアプリケーション入力部、１２０　ページ処理関数特定部、１３０　ＤＢ呼び出し関数特定部、１４０　処理フロー解析部、１５０　入力項目名特定部、１６０　出力項目名特定部、１７０　説明情報記録部、１８０　解析結果出力部、１０１０　特定処理実行部、３１０　フレームワーク特定部、３２０　フレームワーク特定ルール蓄積部、３３０　設定データ生成機能蓄積部、３３１　設定データ生成機能。

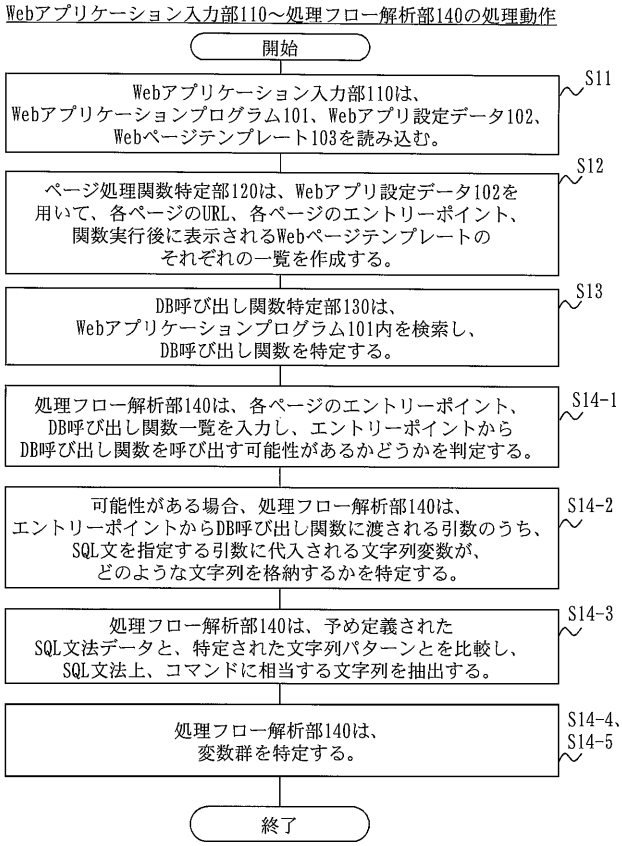
【図１】


101
Webアプリケーションプログラム
102
Webアプリ設定データ(Webアプリ設定ファイル)
103
Webページテンプレート
110
Webアプリケーション入力部(入力部)
1010：特定処理実行部
120
ページ処理関数特定部
150
入力項目名特定部
130
DB呼び出し関数特定部
160
出力項目名特定部
140
処理フロー解析部
180
解析結果出力部(出力部)
説明情報記録部
170
データベース解析装置
100
解析結果
104

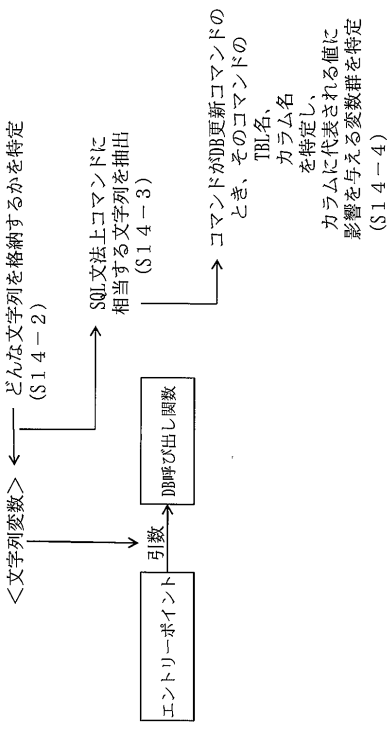
【図２】


Webアプリケーションプログラム101, Webアプリ設定データ102等の概要
(a)
更新/参照されるDBのTBLとカラムとを特定し、各カラムの入出力データに関連する変数群を特定
処理フロー解析部140
エントリーポイント
101：Webアプリケーションプログラム
DB呼び出し関数特定部130
DB読み出し関数
特定
(b)
各ページを処理するエントリーポイントを特定
a. Webアプリケーションプログラム101の受理するURLの一覧
b. URLが呼び出されたときに起動する関数名(エントリーポイント)
c. URL処理結果として使用すべきWebページテンプレートの識別子
102：Webアプリ設定データ
ページ処理関数特定部120
(c)
Webページひな型データ
103：Webページテンプレート

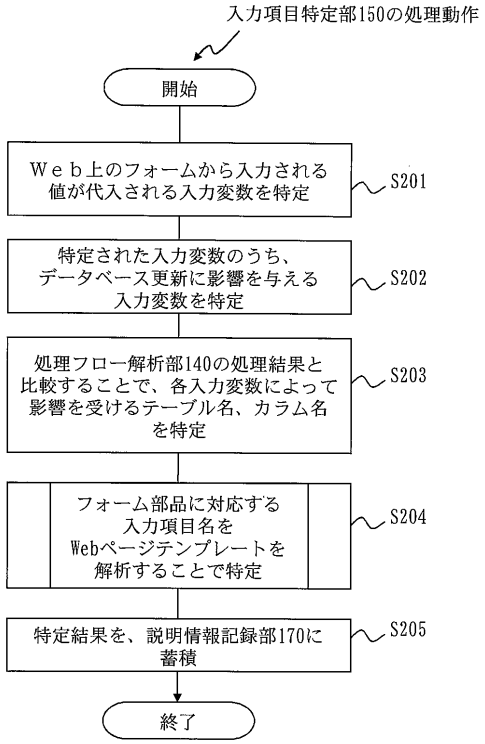
【図３】


Webアプリケーション入力部110～処理フロー解析部140の処理動作
開始
Webアプリケーション入力部110は、Webアプリケーションプログラム101、Webアプリ設定データ102、Webページテンプレート103を読み込む。
S11
ページ処理関数特定部120は、Webアプリ設定データ102を用いて、各ページのURL、各ページのエントリーポイント、関数実行後に表示されるWebページテンプレートのそれぞれの一覧を作成する。
S12
DB呼び出し関数特定部130は、Webアプリケーションプログラム101内を検索し、DB呼び出し関数を特定する。
S13
処理フロー解析部140は、各ページのエントリーポイント、DB呼び出し関数一覧を入力し、エントリーポイントからDB呼び出し関数を呼び出す可能性があるかどうかを判定する。
S14-1
可能性がある場合、処理フロー解析部140は、エントリーポイントからDB呼び出し関数に渡される引数のうち、SQL文を指定する引数に代入される文字列変数が、どのような文字列を格納するかを特定する。
S14-2
処理フロー解析部140は、予め定義されたSQL文法データと、特定された文字列パターンとを比較し、SQL文法上、コマンドに相当する文字列を抽出する。
S14-3
処理フロー解析部140は、変数群を特定する。
S14-4、S14-5
終了

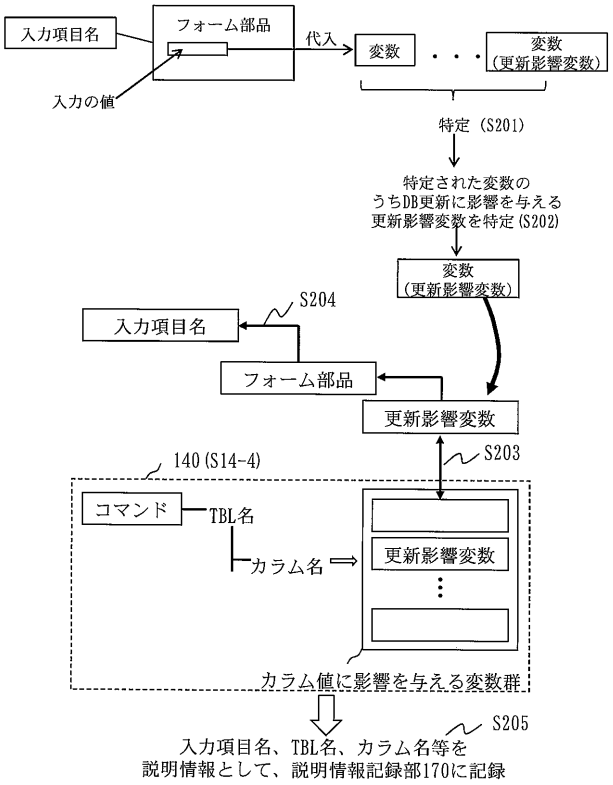
【図４】


どんな文字列を格納するかを特定
(S14-2)
SQL文法上コマンドに相当する文字列を抽出
(S14-3)
コマンドがDB更新コマンドのとき、そのコマンドのTBL名、カラム名を特定し、カラムに代表される値に影響を与える変数群を特定
(S14-4)
<文字列変数>
引数
DB呼び出し関数
エントリーポイント

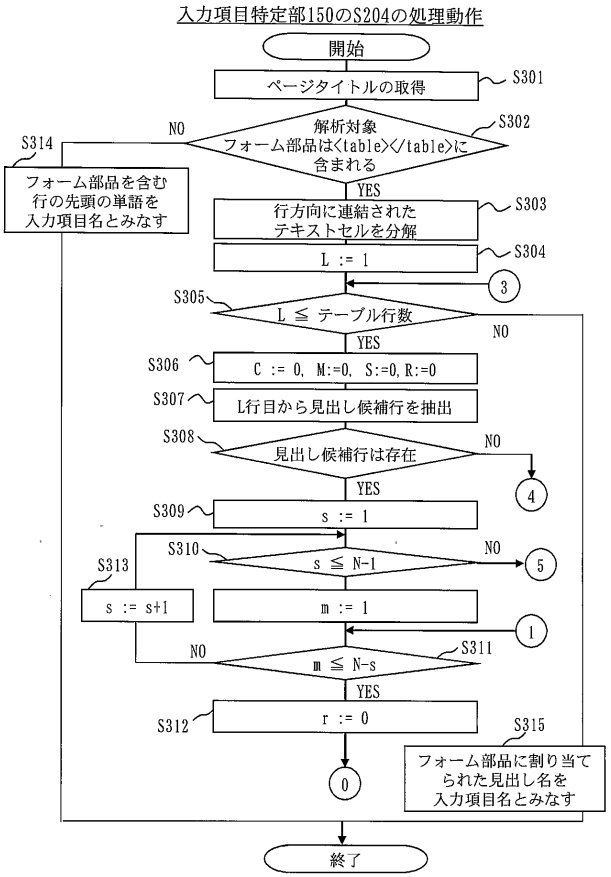
【図５】


入力項目特定部150の処理動作
開始
Ｗｅｂ上のフォームから入力される値が代入される入力変数を特定
S201
特定された入力変数のうち、データベース更新に影響を与える入力変数を特定
S202
処理フロー解析部140の処理結果と比較することで、各入力変数によって影響を受けるテーブル名、カラム名を特定
S203
フォーム部品に対応する入力項目名をWebページテンプレートを解析することで特定
S204
特定結果を、説明情報記録部170に蓄積
S205
終了

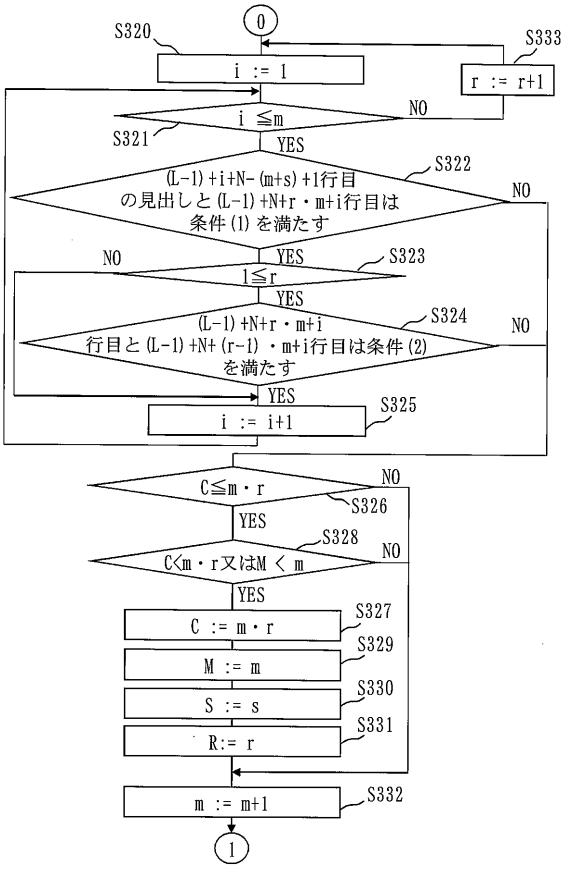
【図６】


入力項目名
フォーム部品
代入
変数
変数（更新影響変数）
入力の値
特定（S201）
特定された変数のうちDB更新に影響を与える更新影響変数を特定（S202）
変数（更新影響変数）
S204
入力項目名
フォーム部品
更新影響変数
S203
140（S14-4）
コマンド
TBL名
カラム名
更新影響変数
カラム値に影響を与える変数群
S205
入力項目名、TBL名、カラム名等を説明情報として、説明情報記録部170に記録

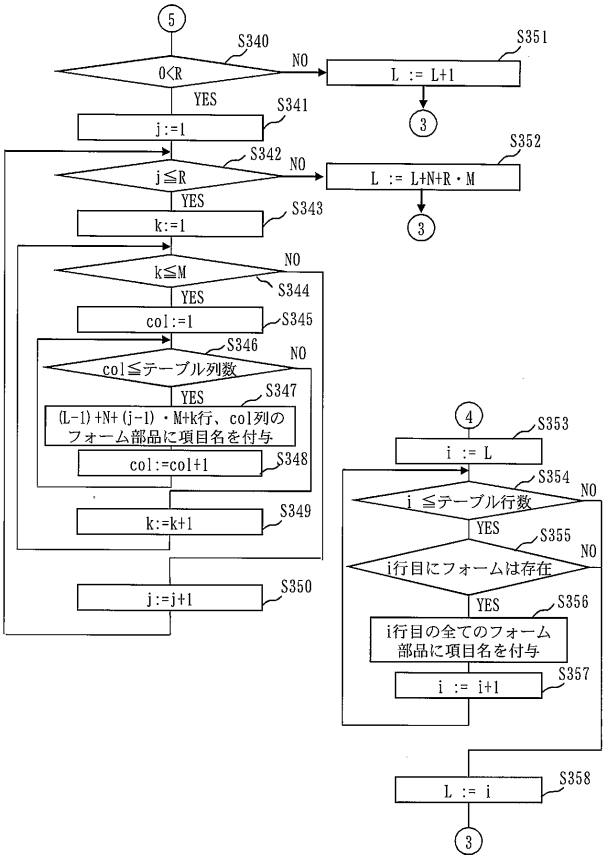
【図７】


入力項目特定部150のS204の処理動作
開始
ページタイトルの取得
S301
解析対象フォーム部品は<table></table>に含まれる
S302
NO
YES
S314
フォーム部品を含む行の先頭の単語を入力項目名とみなす
行方向に連結されたテキストセルを分解
S303
L := 1
S304
3
S305
L ≦ テーブル行数
NO
YES
C := 0, M:=0, S:=0, R:=0
S306
L行目から見出し候補行を抽出
S307
S308
見出し候補行は存在
NO
YES
4
S309
s := 1
S310
s ≦ N-1
NO
5
S313
s := s+1
m := 1
1
S311
m ≦ N-s
NO
YES
S312
r := 0
0
S315
フォーム部品に割り当てられた見出し名を入力項目名とみなす
終了

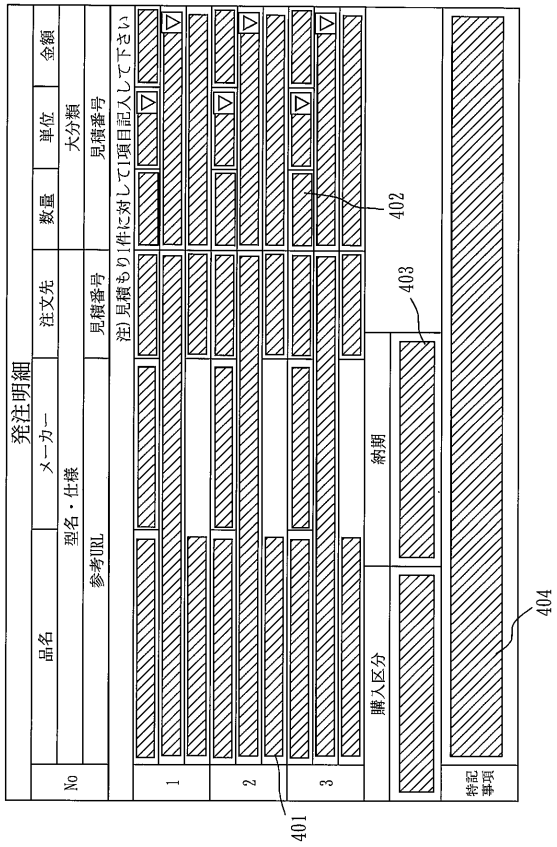
【図８】


0
S320
i := 1
S333
r := r+1
i ≦m
NO
S321
YES
(L-1)+i+N-(m+s)+1行目の見出しと(L-1)+N+r・m+i行目は条件(1)を満たす
S322
NO
YES
NO
S323
1≦r
YES
(L-1)+N+r・m+i行目と(L-1)+N+(r-1)・m+i行目は条件(2)を満たす
S324
NO
YES
i := i+1
S325
C≦m・r
NO
S326
YES
C<m・r又はM<m
S328
NO
YES
C := m・r
S327
M := m
S329
S := s
S330
R:= r
S331
m := m+1
S332
1

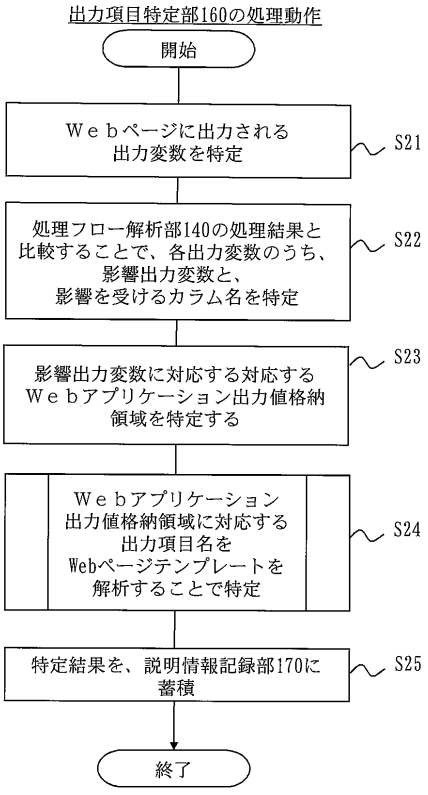
【図９】



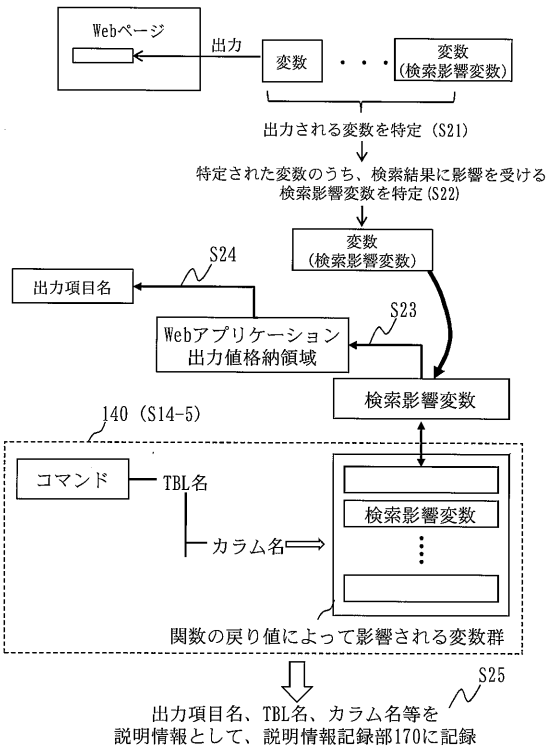
【図１０】



【図１１】



【図１２】

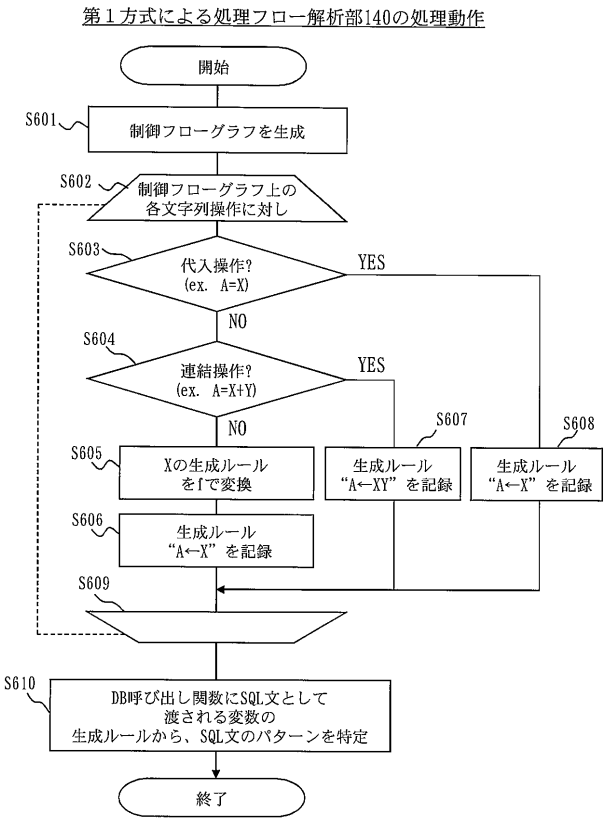


【図１３】

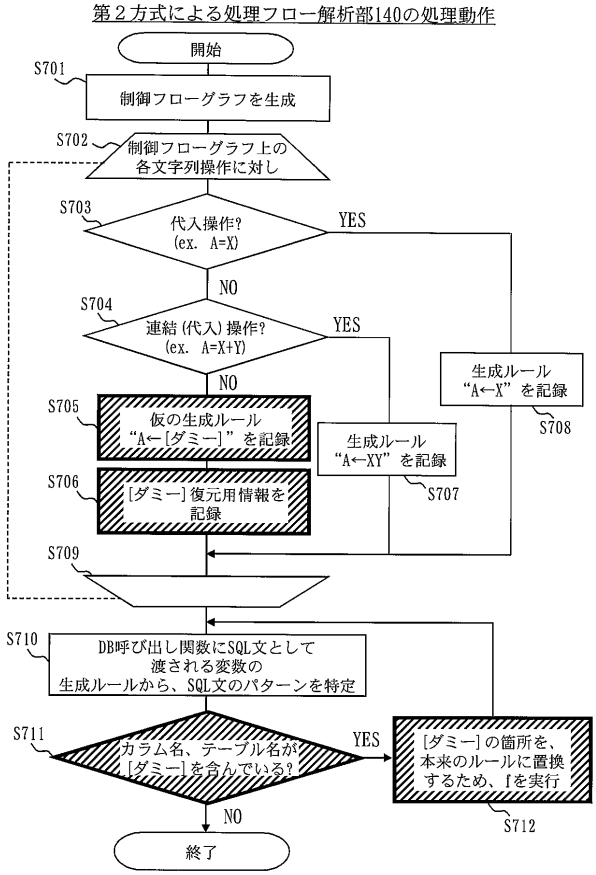
501 テーブル名	502 カラム名	503 入力ページ名	504 入力項目名	505 出力ページ名	506 出力項目名
table1	c1	発注	製品名	発注履歴	製品
	c2	発注	数量	発注履歴	個数
	c3	発注	価格	発注履歴	単価
table2	x1	ユーザ登録	ユーザ名	ユーザ情報	製品
	x2	パスワード変更	パスワード		
	x3	ユーザ登録	住所	ユーザ情報	単価

・・・

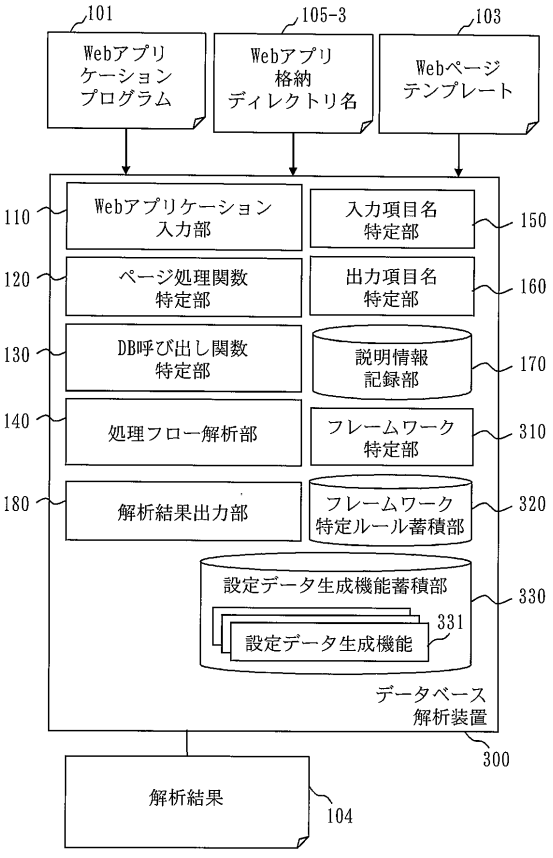
【図１４】



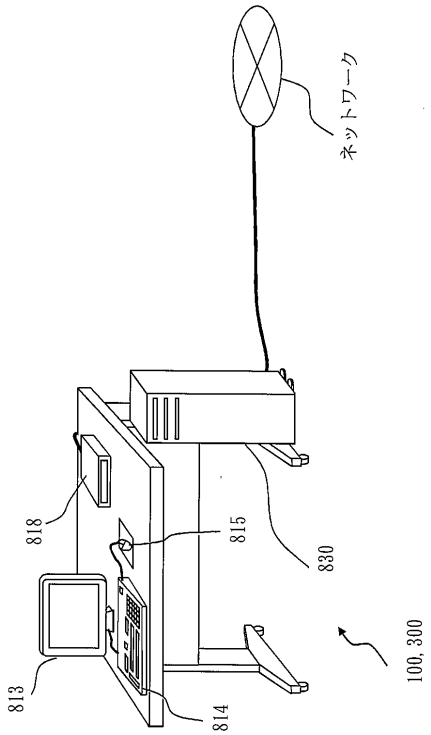
【図１５】



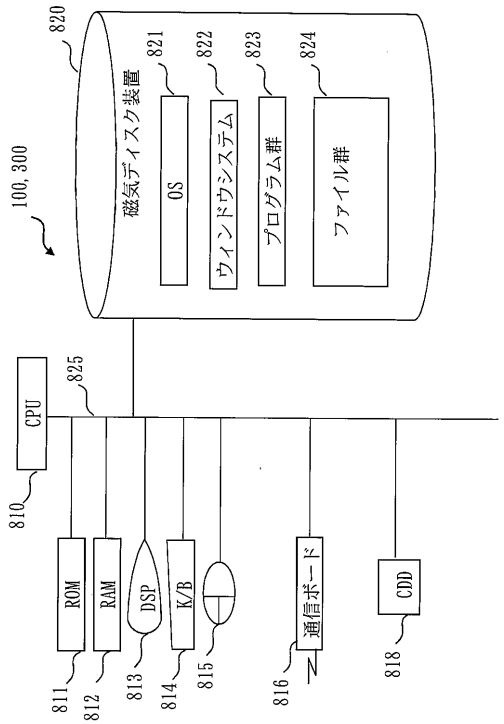
【図１６】



【図１７】


ネットワーク
818
815
830
813
814
100,300

【図１８】


820
磁気ディスク装置
821
OS
822
ウィンドウシステム
823
プログラム群
824
ファイル群
100,300
CPU
825
810
ROM
RAM
DSP
K/B
通信ボード
CDD
811
812
813
814
815
816
818