

(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2008-171181

(P2008-171181A)

(43) 公開日 平成20年7月24日(2008.7.24)

(51) Int.Cl. F I テーマコード (参考)
G06F 17/30 (2006.01) G06F 17/30 140 5B075
 G06F 17/30 330B

審査請求 未請求 請求項の数 6 O L (全 19 頁)

(21) 出願番号 特願2007-3444 (P2007-3444)
 (22) 出願日 平成19年1月11日 (2007.1.11)

(71) 出願人 000003078
 株式会社東芝
 東京都港区芝浦一丁目1番1号
 (71) 出願人 301063496
 東芝ソリューション株式会社
 東京都港区芝浦一丁目1番1号
 (74) 代理人 100058479
 弁理士 鈴江 武彦
 (74) 代理人 100091351
 弁理士 河野 哲
 (74) 代理人 100088683
 弁理士 中村 誠
 (74) 代理人 100108855
 弁理士 蔵田 昌俊

最終頁に続く

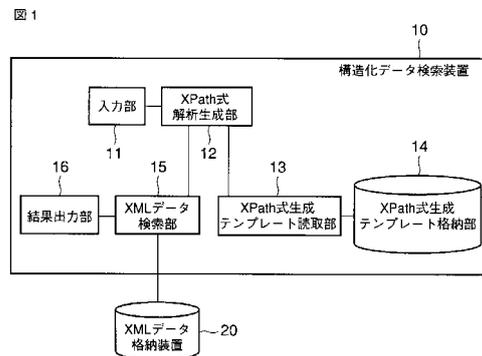
(54) 【発明の名称】 構造化データ検索装置

(57) 【要約】

【課題】ユーザによって指定されたスキーマに準拠したXPath式から、当該指定されたスキーマに準拠しないXMLデータをも検索することを可能とする。

【解決手段】入力部11は、ユーザによって指定されたパス構成要素を含むXPath式(第1の検索式)を入力する。XPath式解析生成部12は、入力されたXPath式を解析し、当該XPath式に含まれるパス構成要素を抽出する。XPath式解析生成部12は、抽出されたパス構成要素に基づいて、XMLデータ格納装置20に格納されている、入力されたXPath式とは異なるスキーマのXMLデータを検索するための定型XPath式(第2の検索式)を生成する。XMLデータ検索部15は、生成された定型XPath式を用いて、当該定型XPath式に合致する構造化データをXMLデータ格納装置20から検索する。結果出力部16は、XMLデータ検索部15の結果を出力する。

【選択図】 図1



【特許請求の範囲】**【請求項 1】**

スキーマが異なる複数の構造化データが格納されている構造化データ格納手段から、データを検索する構造化データ検索装置において、

ユーザによって指定された構成要素を含む構造化データを検索するための第 1 の検索式を入力する入力手段と、

前記第 1 の検索式を解析し、当該第 1 の検索式に含まれる構成要素を抽出する抽出手段と、

前記抽出された構成要素に基づいて、前記構造化データ格納手段に格納されている前記第 1 の検索式とは異なるスキーマの構造化データを検索するための第 2 の検索式を生成する検索式生成手段と、

前記第 2 の検索式を用いて、当該第 2 の検索式に合致する構造化データを前記構造化データ格納手段から検索する検索手段と、

前記検索手段の結果を出力する検索結果出力手段と

を具備することを特徴とする構造化データ検索装置。

【請求項 2】

スキーマが異なる複数の構造化データが格納されている構造化データ格納手段から、データを検索する構造化データ検索装置において、

ユーザによって指定されたスキーマの構造化データを検索するための当該ユーザによって指定された構成要素を含む第 1 の検索式を入力する入力手段と、

前記第 1 の検索式を解析し、当該第 1 の検索式に含まれる構成要素を抽出する抽出手段と、

前記抽出された構成要素に基づいて、前記指定されたスキーマ及び当該指定されたスキーマとは異なるスキーマの構造化データを検索するための第 2 の検索式を生成する検索式生成手段と、

前記第 2 の検索式を用いて、当該第 2 の検索式に合致する構造化データを前記構造化データ格納手段から検索する検索手段と、

前記検索手段の結果を出力する検索結果出力手段と

を具備することを特徴とする構造化データ検索装置。

【請求項 3】

前記構造化データ格納手段に格納されている構造化データを検索するための検索式に含まれる構成要素に変数が埋め込まれた形式で表現される少なくとも 1 つのテンプレートを予め格納するテンプレート格納手段を更に具備し、

前記検索式生成手段は、前記テンプレート格納手段に格納されている少なくとも 1 つのテンプレートに基づいて、当該テンプレートに埋め込まれている変数を前記抽出された構成要素に置換することによって前記第 2 の検索式を生成する

ことを特徴とする請求項 1 または 2 記載の構造化データ検索装置。

【請求項 4】

前記テンプレート格納手段は、前記少なくとも 1 つのテンプレートを含む前記変数の組み合わせがそれぞれ異なる複数のテンプレートを格納し、

前記複数のテンプレートの各々は、検索式の曖昧度を示すレベルに対応付けられており、

前記検索結果出力手段は、前記テンプレートに対応付けられているレベル毎に検索結果として出力する

ことを特徴とする請求項 3 記載の構造化データ検索装置。

【請求項 5】

前記入力手段は、更に曖昧度を示す指定レベル情報を入力し、

前記検索式生成手段は、前記テンプレート格納手段に格納されている複数のテンプレートのうち、前記指定レベル情報によって示される曖昧度より低い曖昧度を示すレベルに対応付けられているテンプレートに基づいて、前記第 2 の検索式を生成する

10

20

30

40

50

ことを特徴とする請求項 4 記載の構造化データ検索装置。

【請求項 6】

前記入力手段は、更に検索件数を示す指定件数情報を入力し、

前記検索式生成手段は、前記テンプレート格納手段に格納されている複数のテンプレートの各々に対応付けられている曖昧度が低い順に、当該テンプレートに基づいて前記第 2 の検索式を生成し、

前記検索手段は、検索された構造化データの数が前記指定件数情報によって示される検索件数を超えるまで、当該第 2 の検索式に合致する構造化データを順次検索する

ことを特徴とする請求項 4 記載の構造化データ検索装置。

【発明の詳細な説明】

10

【技術分野】

【0001】

本発明は、XMLデータベースに蓄積されたスキーマの異なる複数の種類のXMLデータに対して、ユーザによって指定されたXPath式に応じて検索を行う構造化データ検索装置に関する。

【背景技術】

【0002】

一般に、論理構造を持つデータは構造化データと呼ばれる。構造化データにおいて、当該データの論理構造は、当該データ中に記述されたタグによって示される場合がある。このタグを用いて論理構造が表現された構造化データは、計算機で各種目的に合わせて解釈ないし加工して利用する処理に適している。構造化データの代表として、XML (Extensible Markup Language) 形式で記述されたXMLデータが知られている。

20

【0003】

近年、非常に多くのアプリケーションでXMLが用いられるようになり、様々なデータがXML形式で記述されるようになってきている。これにより、XML形式で記述されたXMLデータを検索する技術が、重要となっている。

【0004】

上記したXMLデータでは、複数の構成要素からなる階層化された論理構造を有し、例えば構成要素の要素名等の構造は必ずしも予め定められている必要はない。XMLデータには、例えば異なるスキーマ(データ構造)に準拠するデータが混在している。異なるスキーマとしては、例えばXHTML (Extensible HyperText Markup Language)、NewsML、RSS (Rich Site Summary) またはAtom等が挙げられる。

30

【0005】

このような異なるスキーマに準拠するXMLデータが混在するXMLデータベースに対して、例えばユーザ(検索者)が所望するXMLデータを検索する方法には、例えば以下に示す第1~第4の方法を含む複数の方法がある。

【0006】

まず、第1の方法として、ユーザが例えば異なるスキーマ毎にXPath式(問い合わせ言語)を指定する方法がある。これによれば、指定されたスキーマ毎のXPath式を用いて、検索を実行することにより、異なるスキーマに準拠するXMLデータであっても検索することができる。XPath式は、例えば構成要素を含むパス形式で示される。このXPath式によれば、当該XPath式に含まれる構成要素によりXMLデータの構造を指定することによって、XMLデータを検索することができる。

40

【0007】

第2の方法として、例えばユーザによって指定されたキーワードに基づいて、機械的に全文検索を行う方法がある。これによれば、異なるスキーマに準拠するXMLデータであっても、指定されたキーワードを含むXMLデータを検索することができる。

【0008】

第3の方法として、異なるスキーマ毎のXPath式の対応表を作成しておく方法がある。これによれば、ユーザによって指定されたスキーマに準拠するXPath式が入力された場合

50

であっても、当該指定されたスキーマとは異なるスキーマに準拠するXMLデータを検索することができる。これに関連する技術として、例えばユーザがXMLデータ等の構造を意識することなく検索でき、ユーザが所望するXMLデータをユーザが指定した構造で取得することができる技術が開示されている（例えば、特許文献1を参照）。

【0009】

また、第4の方法として、異なるスキーマの各々に合致するような抽象スキーマを作成する方法がある。これによれば、異なるスキーマの各々の対応をとった中間形式としての抽象スキーマを作成し、当該抽象スキーマとの間でXPath式の変換を行うことで、異なるスキーマに準拠するXMLデータを検索することができる。これに関連する技術として、例えば異なるデータ構造の複数の構造化データの中から、所望の構造化データを検索する際には、そのデータ構造に依存することなく、所望のXMLデータを要素値として持つ構成要素を含む構造化データを容易に検索する技術が開示されている（例えば、特許文献2を参照）。

10

【特許文献1】特開2003-316783号公報

【特許文献2】特開2004-164104号公報

【発明の開示】

【発明が解決しようとする課題】

【0010】

上記したような第1～第4の方法によれば、異なるスキーマに準拠するXMLデータであっても検索することが可能である。

20

【0011】

しかしながら、上記した第1の方法では、ユーザは、異なるスキーマ毎にXPath式を複数指定（入力）する必要があるため手間がかかる。また、ユーザは、異なるスキーマ毎のすべてのXMLデータの構造を把握している必要がある。

【0012】

また、第2の方法では、XMLデータの構造を意識した検索を行うことができないため、多くのノイズデータ（ユーザが必要としていないデータ）が検索結果に含まれる。

【0013】

また、第3の方法では、ユーザにとって意図する検索を行うことは可能であるが、対応表を作成する手間やコストがかかってしまう。また、新規スキーマに準拠するXMLデータ追加された場合には、対応表に新規スキーマのXPath式を追加する必要がある。これにより、更に手間やコストがかかるおそれがある。

30

【0014】

また、第4の方法では、中間形式としての抽象スキーマを作成するコストが大きくなる。また、新規スキーマに準拠するXMLデータが追加された場合には、抽象スキーマに対する変更が発生する可能性がある。これにより、更にコストがかかるおそれがある。

【0015】

本発明の目的は、ユーザによって指定されたスキーマに準拠したXPath式から、当該スキーマに準拠しないXMLデータをも検索することができる構造化データ検索装置を提供することにある。

40

【課題を解決するための手段】

【0016】

本発明の1つの態様によれば、スキーマが異なる複数の構造化データを格納する構造化データ格納手段から、データを検索する構造化データ検索装置が提供される。この構造化データ検索装置は、ユーザによって指定された構成要素を含む構造化データを検索するための第1の検索式を入力する入力手段と、前記第1の検索式を解析し、当該第1の検索式に含まれる構成要素を抽出する抽出手段と、前記抽出された構成要素に基づいて、前記前記構造化データ格納手段に格納されている第1の検索式とは異なるスキーマの構造化データを検索するための第2の検索式を生成する検索式生成手段と、前記第2の検索式を用いて、当該第2の検索式に合致する構造化データを前記構造化データ格納手段から検索する

50

検索手段と、前記検索手段の結果を出力する検索結果出力手段とを具備する。

【発明の効果】

【0017】

本発明によれば、ユーザによって指定されたスキーマに準拠したXPath式から、当該指定されたスキーマに準拠しないXMLデータをも検索することが可能となる。

【発明を実施するための最良の形態】

【0018】

以下、図面を参照して、本発明の実施形態について説明する。

【0019】

図1は、本実施形態に係る構造化データ検索装置の機能構成を示すブロック図である。図1に示すように、構造化データ検索装置10は、入力部11、XPath式解析生成部12、XPath式生成テンプレート読取部13、XPath式生成テンプレート格納部14、XMLデータ検索部15及び結果出力部16を含む。

10

【0020】

また、構造化データ検索装置10の外部には、XMLデータ格納装置20が設けられている。このXMLデータ格納装置20には、例えばスキーマ(データ構造)が異なる複数の構造化データ(以下、XMLデータと表記)が格納されている。XMLデータは、例えば当該XMLデータ中に記述されたタグによって階層化されたデータである。また、異なるスキーマには、例えばXHTML(Extensible HyperText Markup Language)、NewsML、RSS(Rich Site Summary)またはAtom等が含まれる。

20

【0021】

入力部11は、例えばユーザによって指定されたスキーマに準拠するXMLデータを検索するための検索式(第1の検索式)を入力する。この検索式(以下、XPath式と表記)は、パス形式で示される。また、XPath式には、例えばユーザによって指定された構成要素(以下、パス構成要素と表記)が含まれる。XPath式は、当該XPath式に含まれる構成要素の配列により構造化データの構造が表されている。XPath式に含まれるパス構成要素は、例えばXMLデータのタグに該当する要素(または属性)または当該要素に対する制限を示す述部を含む。この述部には、例えば該当する要素の内容に、ある文字列が含まれることを示す「contains」のような関数が含まれる。XPath式は、例えば「/第0要素[第0述部]/第1要素[第1述部]/.../最終要素または属性[最終述部]」のように示される。なお、第n要素(または第n述部)のn(n=0、1、...)は、XPath式に含まれる構成要素の階層を表す。また、[第n述部]及び[最終述部]の部分は、省略することが可能である。

30

【0022】

XPath式解析生成部12は、入力部11に入力されたXPath式(以下、入力XPath式と表記)を解析する。XPath式解析生成部12は、解析結果に応じて、入力XPath式に含まれるパス構成要素を抽出する。また、XPath式解析生成部12は、抽出されたパス構成要素及びXPath式生成テンプレート格納部14に格納されているテンプレートに基づいて、定型XPath式を生成する。XPath式解析生成部12は、ユーザによって指定されたスキーマ及び当該指定されたスキーマとは異なるスキーマに準拠する構造化データを検索するための定型XPath式(第2の検索式)を生成する。

40

【0023】

XPath式生成テンプレート読取部13は、例えばXPath式生成テンプレート格納部14に予め格納されているXPath式生成テンプレート(以下、単にテンプレートと表記)を読み取る。XPath式生成テンプレート読取部13は、読み取られたテンプレートをXPath式解析生成部12に渡す。

【0024】

XPath式生成テンプレート格納部14には、複数のテンプレートが格納されている。この複数のテンプレートは、例えば定型XPath式を生成するために用いられるテンプレートファイルである。複数のテンプレートは、例えばXPath式に含まれるパス構成要素に変数

50

が埋め込まれた形式で表現される。複数のテンプレートの各々は、変数の組み合わせがそれぞれ異なる。テンプレートに埋め込まれている変数は、例えばXPath式に含まれるパス構成要素の該当部分に対応し、機械的にXPath式を生成可能な形式であるものとする。変数は、XPath式に含まれる構成要素の階層を表すように予め定義されている。

【 0 0 2 5 】

また、テンプレートの各々は、例えば当該テンプレートを用いて生成されるXPath式の曖昧度を示すレベル（以下、曖昧レベルと表記）に対応付けてXPath式生成テンプレート格納部 1 4 に格納されている。この曖昧レベルは、例えば 1 または複数の段階に設定されている。

【 0 0 2 6 】

X M L データ検索部 1 5 は、X M L データ格納装置 2 0 に格納されている X M L データのうち、入力XPath式またはXPath式解析生成部 1 2 によって生成された定型XPath式を用いて X M L データを検索する。また、X M L データ検索部 1 5 は、検索された結果を記憶（保持）する記憶部（図示せず）を含む。

【 0 0 2 7 】

結果出力部 1 6 は、X M L データ検索部 1 5 によって検索された X M L データを検索結果として、例えばユーザに対して表示するために出力する。また、結果出力部 1 6 は、例えば入力XPath式を用いて検索された X M L データまたはXPath式解析生成部 1 2 によって生成された定型XPath式を用いて検索された X M L データを区別して出力する。また、結果出力部 1 6 は、XPath式解析生成部 1 2 によって生成された定型XPath式を用いて検索された X M L データを、当該XPath式が生成される際に用いられたテンプレートのレベル毎に検索結果を出力（表示）する。

【 0 0 2 8 】

図 2 は、XPath式解析生成部 1 2 の機能構成を示すブロック図である。XPath式解析生成部 1 2 は、XPath式解析部 1 2 1、テンプレート解析部 1 2 2、変換部 1 2 3 及びXPath式生成部 1 2 4 を含む。

【 0 0 2 9 】

XPath式解析部 1 2 1 は、入力XPath式を解析する。XPath式解析部 1 2 1 は、解析結果に応じて、パス構成要素を抽出する。

【 0 0 3 0 】

テンプレート解析部 1 2 2 は、XPath式生成テンプレート読取部 1 3 によって読み取られたテンプレートを解析する。テンプレート解析部 1 2 2 は、解析結果に応じて、テンプレートの変数（部分）を切り出す。

【 0 0 3 1 】

変換部 1 2 3 は、XPath式解析部 1 2 1 によって抽出されたパス構成要素及びテンプレート解析部 1 2 2 によって切り出された変数から対応付けを行い、対応するパス構成要素を変数に置換する処理を実行する。

【 0 0 3 2 】

XPath式生成部 1 2 4 は、変換部 1 2 3 によって置換処理されたテンプレートから、定型XPath式を生成する。

【 0 0 3 3 】

図 3 は、XPath式生成テンプレート格納部 1 4 のデータ構造の一例を示す。図 3 の例では、XPath式生成テンプレート格納部 1 4 には、テンプレート A ~ F が格納されている。テンプレート A ~ C は曖昧レベル 1 に、テンプレート D 及び E は曖昧レベル 2 に、また、テンプレート F は曖昧レベル 3 に対応付けてXPath式生成テンプレート格納部 1 4 に格納されている。

【 0 0 3 4 】

曖昧レベルは、各テンプレートに基づいて生成された定型XPath式の曖昧度を示す。この曖昧レベルは、例えば曖昧レベル 0 は曖昧性なし、曖昧レベル 3 は全文検索に近い曖昧性のよう規定することができる。なお、本実施形態においては、入力XPathを曖昧レベ

10

20

30

40

50

ル 0 (曖昧性なし)とし、曖昧レベルが高くなる程、曖昧性が高くなるものとする。

【 0 0 3 5 】

以下、XPath式生成テンプレート格納部 1 4 に格納されているテンプレートの具体例について説明する。

曖昧レベル 1 のテンプレート (例えばテンプレート A) の具体例としては、「 /\${Elem.FIRST}//*[\${Predicate.LAST}]」が挙げられる。ここで、「 \${Elem.FIRST}」とは、例えばXPath式に含まれる構成要素の第 0 要素 (階層が 0 である要素) を示す変数であり、入力XPath式に含まれるパス構成要素のうち、最初の要素を示す。また、「 \${Predicate.LAST}」とは、例えば最終述部 (階層が最も低い述部) を示す変数であり、入力XPath式に含まれるパス構成要素のうち、最後の述部を示す。このテンプレートに含まれる「 /」は、当該「 /」の前後の例えば要素同士が親子関係であることを示し、「 //」は、当該「 //」の前後の例えば要素同士が子孫関係 (親子関係を含む) であることを示す。なお、このテンプレートに基づいて生成されるXPath式によれば、例えば入力XPath式に含まれる第 0 要素より以下の階層 (子孫関係) に当該入力XPath式に含まれる最終述部を含む (最終述部によって示される制限を満たす) X M L データが検索される。

10

【 0 0 3 6 】

曖昧レベル 2 のテンプレート (例えばテンプレート D) の具体例としては、「 //\${Elem.LAST}[\${Predicate.LAST}]」が挙げられる。「 \${Elem.LAST}」とは、最終要素を示す変数であり、入力XPath式に含まれるパス構成要素のうち、最後の要素を示す。なお、このテンプレートに基づいて生成されるXPath式によれば、例えば入力XPath式に含まれる最終要素の中 (内容) に当該入力XPath式に含まれる最終述部が含まれている X M L データ検索される。

20

【 0 0 3 7 】

曖昧レベル 3 のテンプレート (例えばテンプレート F) の具体例としては、「 //*[\${Predicate.LAST}]」が挙げられる。なお、このテンプレートに基づいて生成されるXPath式によれば、例えば入力XPath式に含まれる最終述部が含まれる X M L データが検索される。

【 0 0 3 8 】

また、変数には、上記した以外にも、第 n 要素を示す「 \${Elem.n}」、第 0 述部を示す「[\${Predicate.FIRST}]」または第 n 述部を示す「[\${Predicate.n}]」等が含まれる。

30

【 0 0 3 9 】

次に、図 4 のフローチャートを参照して、構造化データ検索装置 1 0 の処理手順について説明する。

まず、入力部 1 1 は、ユーザによって指定されたスキーマに準拠する X M L データを検索するためのXPath式が入力される (ステップ S 1)。

【 0 0 4 0 】

次に、X M L データ検索部 1 5 は、入力部 1 1 に入力されたXPath式 (入力XPath式) を用いて、当該入力XPath式に合致する X M L データを X M L データ格納部 2 0 から検索する (ステップ S 2)。

【 0 0 4 1 】

X M L データ検索部 1 5 によって検索された X M L データは、当該 X M L データ検索部 1 5 に含まれる記憶部に保持される (ステップ S 3)。このとき、この X M L データは、曖昧レベル 0 の検索結果として保持される。

40

【 0 0 4 2 】

XPath式解析生成部 1 2 は、入力XPath式の解析を行う。これにより、XPath式解析生成部 1 2 は、入力XPath式に含まれるパス構成要素の各々を抽出する (ステップ S 4)。

【 0 0 4 3 】

次に、XPath式解析生成部 1 2 は、XPath式生成テンプレート読取部 1 3 に対して、曖昧レベル n のテンプレートをXPath式生成テンプレート格納部 1 4 から読み取る (読み出す) ための要求 (読取要求) を出力する。XPath式生成解析生成部 1 2 は、まず曖昧度が低

50

いテンプレートを読み取るよう読取要求する。

【0044】

XPath式生成テンプレート読取部13は、XPath式解析生成部12からの読取要求に応じて、XPath式生成テンプレート格納部14に格納されているテンプレートを読み取る（ステップS5）。初めに、XPath式生成テンプレート格納部14から、曖昧レベル1のテンプレートが読み取られる。このとき、上記した図3に示すXPath式生成テンプレート格納部14のように、例えば曖昧レベル1に対応付けて複数のテンプレートが格納されている場合には、当該複数のテンプレートの全てを読み取るものとする。XPath式生成テンプレート読取部13は、読み取ったテンプレートをXPath式解析生成部12に出力する（渡す）。

10

【0045】

XPath式解析生成部12は、ステップS4において抽出されたパス構成要素及びXPath式生成テンプレート読取部13によって読み出されたテンプレートに基づいて、定型XPath式を生成する（ステップS6）。このとき、例えばXPath式生成テンプレート読取部13によって複数のテンプレートが読み出された場合には、当該複数のテンプレートの各々について、定型XPath式が生成される。

【0046】

XMLデータ検索部15は、XPath式解析生成部12によって生成された定型XPath式を用いて、当該XPath式に合致するXMLデータをXMLデータ格納装置20から検索する（ステップS7）。XMLデータ検索部15によって検索されたXMLデータは、当該XMLデータ検索部15の記憶部に保持される（ステップS8）。この場合、検索されたXMLデータは、曖昧レベル毎の検索結果として保持される。

20

【0047】

次に、XPath式解析生成部12は、XPath式生成テンプレート格納部14に定型XPath式を生成する処理が実行されていないテンプレート（未処理テンプレート）が存在するかを判定する（ステップS9）。

【0048】

XPath式生成テンプレート格納部14に未処理テンプレートが存在すると判定された場合（ステップS9のYES）、ステップS5に戻って処理が繰り返される。この場合、XPath式解析生成部12は、XPath式生成テンプレート読取部13に対して、曖昧レベル2のテンプレートを読み取るよう読取要求を出力する。これにより、ステップS5では、XPath式生成テンプレート読取部13は、XPath式解析生成部12からの読取要求に応じて、XPath式生成テンプレート格納部14に格納されている曖昧レベル2のテンプレートを読み取る。つまり、ステップS5における曖昧レベルnの値を1つ上げて処理が繰り返される。このように、XPath式生成テンプレート格納部14に格納されているテンプレートは、当該テンプレートの曖昧レベル毎にステップ5～8の処理がされる。

30

【0049】

一方、XPath式生成テンプレート格納部14に未処理テンプレートが存在しないと判定された場合（ステップS9のNO）、結果出力部16は、XMLデータ検索部15の記憶部に保持されている検索結果を例えばユーザに対して表示するために出力する（ステップS10）。結果出力部16は、例えば曖昧レベル毎に検索結果を出力する。

40

【0050】

なお、ここでは、上記したように入力XPath式及びXPath式解析生成部12によって生成された定型XPath式を用いた検索処理（ステップS2及びステップS7）を分けて実行するものとして説明したが、これらの処理は、例えばステップS9及びS10の間で、まとめて実行される構成であってもかまわない。

【0051】

次に、図5のフローチャートを参照して、上記した図4に示すステップS4の処理の詳細な処理手順について説明する。

まず、XPath式解析生成部12のXPath式解析部121は、入力XPath式に含まれる文字

50

列を例えば「/」で分離する（ステップS 1 1）。このとき、XPath式解析部 1 2 1 は、入力XPath式に含まれる文字列を「[]」内に含まれない「/」で分離する。つまり、これによって、1つの要素（または1組の要素及び述部）毎に分離される。

【0052】

次に、XPath式解析部 1 2 1 は、ステップS 1 1において分離された文字列の各々に要素（名）及び述部が含まれる場合には、当該文字列に含まれる要素及び述部を分離する（ステップS 1 2）。ステップS 1 1において分離された文字列に「[]」が含まれる場合に、当該「[]」で囲まれた部分が述部であり、それ以外が要素（または属性）となる。

【0053】

XPath式解析部 1 2 1 は、ステップS 1 1及びS 1 2において分離された要素または述部の各々を、パス構成要素として抽出する（ステップS 1 3）。

10

【0054】

ここで、図6を参照して、上記したパス構成要素を抽出する処理について具体的に説明する。

図6に示すように、入力XPath式は、「/html/head/title[contains(./text(), '耐震偽装')]」であるものとする。以下に、この入力XPath式からパス構成要素を抽出する処理について説明する。

【0055】

この場合、XPath式解析部 1 2 1 は、XPath式に含まれる文字列を「/」で分離する。すなわち、XPath式解析部 1 2 1 は、入力XPath式から「html」、「head」及び「title[contains(./text(), '耐震偽装')]」を分離する。

20

【0056】

次に、分離された「title[contains(./text(), '耐震偽装')]」には、「[]」が含まれるため、要素及び述部が含まれている。このため、XPath式解析部 1 2 1 は、分離された「title[contains(./text(), '耐震偽装')]」を、要素「title」及び述部「[contains(./text(), '耐震偽装')]」を分離する。

【0057】

これにより、XPath式解析部 1 2 1 は、パス構成要素として「html」、「head」、「title」及び「[contains(./text(), '耐震偽装')]」を抽出する。

【0058】

ここで、図6に示すように、XPath式解析部 1 2 1 によって抽出された「html」は入力XPath式の第0要素、「head」は第1要素、「title」は第2要素である。また、「[contains(./text(), '耐震偽装')]」は、入力XPath式の第2述部である。なお、図6に示す「null」は空であることを示し、ここで説明した入力XPath式には、第0及び第1述部は存在しない。

30

【0059】

次に、図7のフローチャートを参照して、上記した図4に示すステップS 6の処理の詳細な処理手順について説明する。

まず、XPath式解析生成部 1 2 のテンプレート解析部 1 2 2 は、XPath式生成テンプレート読取部 1 3 によって読み取られたテンプレートを解析する。これにより、テンプレート解析部 1 2 2 は、解析されたテンプレートから、当該テンプレートに含まれる変数（部分）を切り出す（ステップS 2 1）。

40

【0060】

次に、変換部 1 2 3 は、テンプレート解析部 1 2 2 によって切り出された変数部分を、XPath式解析部 1 2 1 によって抽出されたパス構成要素のうち、当該変数部分に該当するパス構成要素に置換する（ステップS 2 2）。

【0061】

図6に示したXPath式解析部 1 2 1 によって抽出されたパス構成要素の例を用いて具体的に説明する。曖昧レベル1のテンプレートAを「/\${Elem.FIRST}//*[\${Predicate.LAST}]」とする。テンプレートAの変数部分「\${Elem.FIRST}」を、当該抽出されたパス構成

50

要素のうち、第0要素である「html」に置換する。また、変数部分「\${Predicate.LAST}」を、当該抽出されたパス構成要素のうち、第2述部（最終述部）である「contains(./text(), '耐震偽装')」に置換する。

【0062】

次に、XPath式生成部124は、変換部123によって置換されたものを定型XPath式として生成（確定）する（ステップS23）。具体的には、テンプレートの中の変数ではない部分「//」または「[]」と、得られたパス構成要素とを連結する。上記した例では、定型XPath式として、「/html//*[contains(./text(), '耐震偽装')]」が生成される。なお、この定型XPath式によれば、「html」要素より以下の階層に「耐震偽装」の文字列が含まれるXMLデータが検索される。

【0063】

ここで、図8～図12を参照して、XPath式生成テンプレート格納部14に格納されている曖昧レベル毎のテンプレートに応じて、XPath式が生成された際の検索処理について具体的に説明する。

【0064】

図8～図12は、それぞれXMLデータ格納部20に格納されているXMLデータの一例を示す。なお、図8に示すXMLデータ31は、XHTMLに準拠するXMLデータである。図9に示すXMLデータ32は、NewsMLに準拠するXMLデータである。図10に示すXMLデータ33は、RSSに準拠するXMLデータである。図11に示すXMLデータ34は、Atomに準拠するXMLデータである。図12に示すXMLデータ35は、図8のXMLデータ31と同様に、XHTMLに準拠するXMLデータである。

【0065】

ここで、入力XPath式が「/html/head/title[contains(./text(), '耐震偽装')]」である場合を想定する。この入力XPath式は、例えばユーザによって指定されたスキーマ（ここでは、XHTML）に準拠したXPath式である。この場合、例えば図6で説明したように、パス構成要素として入力XPath式の第0要素である「html」、第1要素である「head」、第2要素（最終要素）である「title」及び第2述部（最終述部）である「[contains(./text(), '耐震偽装')]」が抽出されたものとする。

【0066】

また、曖昧レベル毎のテンプレートの例として、曖昧レベル1のテンプレートAは、「/\${Elem.FIRST}//*[\${Predicate.LAST}]」であるものとする。曖昧レベル2のテンプレートDは、「//\${Elem.LAST}[\${Predicate.LAST}]」であるものとする。また、曖昧レベル3のテンプレートFは、「//*[\${Predicate.LAST}]」であるものとする。

【0067】

まず、入力XPath式を用いて検索処理が実行された場合、当該入力XPath式によれば、要素「html」より以下の階層に要素「head」があり、当該要素「head」より以下の階層に要素「title」があり、当該要素「title」の内容に「耐震偽装」の文字列が含まれるXMLデータが検索される。入力XPath式に合致するXMLデータとしては、XMLデータ31～35のうち、XMLデータ31の1件が該当する。つまり、XHTMLに準拠するXMLデータが検索される。

【0068】

次に、曖昧レベル1のテンプレートA及び抽出されたパス構成要素から、定型XPath式として「/html//*[contains(./text(), '耐震偽装')]」が生成される。この定型XPath式を用いて検索処理が実行された場合、当該定型XPath式によれば、要素「html」より以下の階層に「耐震偽装」の文字列が含まれているXMLデータが検索される。この定型XPath式に合致するXMLデータとしては、XMLデータ31～35のうち、XMLデータ31及び35の2件が該当する。つまり、上記した入力XPath式に合致するXMLデータ（XMLデータ31）を含むXHTMLに準拠するXMLデータが検索される。

【0069】

また、曖昧レベル2のテンプレートD及び抽出されたパス構成要素から、定型XPath式

10

20

30

40

50

として「`//*[@title[contains(./text(), '耐震偽装')]]`」が生成される。この定型XPath式を用いて検索処理が実行された場合、当該定型XPath式によれば、要素「title」の内容に「耐震偽装」の文字列が含まれるXMLデータが検索される。この定型XPath式に合致するXMLデータとしては、XMLデータ31～35のうち、XMLデータ31、33及び34の3件が該当する。つまり、ユーザによって指定されたスキーマ(XHTML)とは異なるスキーマ(RSS及びAtom)に準拠するXMLデータ(XMLデータ33及び34)も検索される。

【0070】

また、曖昧レベル3のテンプレートF及び抽出されたパス構成要素から、定型XPath式として「`//*[contains(./text(), '耐震偽装')]`」が生成される。この定型XPath式を用いて検索処理が実行された場合、当該定型XPath式によれば、「耐震偽装」の文字列が含まれるXMLデータが検索される。換言すれば、「耐震偽装」の文字列(キーワード)で全文検索を実行した場合に相当する。この定型XPath式に合致するXMLデータとしては、XMLデータ31～35の全てが該当する。つまり、ユーザによって指定されたスキーマ(XHTML)とは異なるスキーマ(NewsML、RSS及びAtom)に準拠するXMLデータ(XMLデータ32、33及び34)も検索される。

10

【0071】

このように、例えばXHTMLに準拠した入力XPath式のみを用いて検索処理が実行された場合には、XHTMLに準拠したXMLデータ31のみが検索される。しかしながら、XPath式生成テンプレート格納部14に格納されているテンプレート及びXPath式から抽出されたパス構成要素に基づいて、定型XPath式を生成し検索することによって、スキーマが異なる(つまり、XHTMLに準拠していない)XMLデータ32～34であっても検索することができる。

20

【0072】

本実施形態においては、ユーザによって指定された入力XPath式からパス構成要素を抽出し、当該パス構成要素及びテンプレートを用いて定型XPath式が生成される。これにより、ユーザによって指定されたスキーマ以外のスキーマに準拠したXMLデータであっても検索することが可能となる。

【0073】

また、上記実施形態においては、階層構造を意識したテンプレートを作成することで、定型XPath式を用いて検索処理が実行された場合であっても、ノイズデータを少なくすることが可能である。

30

【0074】

また、上記実施形態においては、XPath式生成テンプレート格納部14に格納されているテンプレートは、スキーマに依存しないため、例えば新規のスキーマを持つXMLデータが検索対象となった場合であっても、変更等の処理を行う必要がない。すなわち、スキーマに準拠した入力XPath式によらず、テンプレートからXPath式を作成することにより準拠しないスキーマをもつXMLデータを検索することができる。これにより、新規のスキーマを持つXMLデータが検索対象となった場合であっても、例えばテンプレートを変更の際の手間やコストを削減することが可能となる。

40

【0075】

なお、上記した本実施形態においては、入力XPath式は例えばユーザによって指定されたスキーマに準拠するものとして説明したが、当該入力XPath式は、例えばユーザによって指定されたスキーマに準拠するものでなくとも、XPath式生成テンプレート格納部14に格納されているテンプレートに基づいて、定型XPath式が生成可能なXPath式であればよい。

【0076】

また、入力XPath式と同一の定型XPath式を生成可能なテンプレートが、XPath式生成テンプレート格納部14に存在する場合には、当該入力XPath式を用いて検索処理を実行せず、当該テンプレートを用いて生成された(当該入力XPath式と同一の)定型XPath式のみ

50

を用いて検索処理を行う構成であっても構わない。

【 0 0 7 7 】

[第 1 の 変 形 例]

本実施形態の第 1 の変形例について説明する。本変形例に係る構造化データ検索装置 10 は、入力された指定件数情報によって示される検索件数を超えるまで検索処理を実行する。

【 0 0 7 8 】

図 13 のフローチャートを参照して、構造化データ検索装置 10 の処理手順について説明する。

まず、入力部 11 は、ユーザによって指定されたスキーマに準拠する XML データを検索するための XPath 式及びユーザによって指定された検索件数（検索結果数）を示す指定件数情報が入力される（ステップ S31）。ここでは、ユーザによって指定された検索件数は x （ x は、1 以上の整数）件であるものとして説明する。

【 0 0 7 9 】

次に、XML データ検索部 15 は、入力部 11 に入力された XPath 式（入力 XPath 式）を用いて、当該入力 XPath 式に合致する XML データを XML データ格納部 20 から検索する（ステップ S32）。

【 0 0 8 0 】

XML データ検索部 15 によって検索された XML データは、当該 XML データ検索部 15 に含まれる記憶部に保持される（ステップ S33）。このとき、この XML データは曖昧レベル 0 の検索結果として保持される。

【 0 0 8 1 】

次に、XPath 式解析生成部 12 は、XML データ検索部 15 によって検索された XML データの検索件数が x 件未満であるか否かを判定する（ステップ S34）。検索された XML データの検索件数が x 件未満であると判定された場合（ステップ S34 の YES）、XPath 式解析生成部 12 は、入力 XPath 式の解析を行う。これにより、XPath 式解析生成部 12 は、入力 XPath 式に含まれるパス構成要素の各々を抽出する（ステップ S35）。

【 0 0 8 2 】

次に、XPath 式解析生成部 12 は、XPath 式生成テンプレート読取部 13 に対して、曖昧レベル n のテンプレートを XPath 式生成テンプレート格納部 14 から読み取る（読み出す）ための要求（読取要求）をする。ここでは、XPath 式解析生成部 12 は、まず曖昧度が低いテンプレートから読み取るよう読取要求をする。

【 0 0 8 3 】

XPath 式生成テンプレート読取部 13 は、XPath 式解析生成部 12 からの読取要求に応じて、XPath 式生成テンプレート格納部 14 に格納されているテンプレートを読み取る（ステップ S36）。初めに、XPath 式生成テンプレート格納部 14 から、曖昧レベル 1 のテンプレートが読み取られる。XPath 式生成テンプレート読取部 13 は、読み取ったテンプレートを XPath 式解析生成部 12 に出力する。

【 0 0 8 4 】

XPath 式解析生成部 12 は、ステップ S35 において抽出されたパス構成要素及び XPath 式生成テンプレート読取部 13 によって読み出されたテンプレートに基づいて、定型 XPath 式を生成する（ステップ S37）。

【 0 0 8 5 】

XML データ検索部 15 は、XPath 式解析生成部 12 によって生成された XPath 式を用いて、当該 XPath 式に合致する XML データを XML データ格納装置 20 から検索する（ステップ S38）。XML データ検索部 15 によって検索された XML データは、当該 XML データ検索部 15 の記憶部に保持される（ステップ S39）。この XML データは、曖昧レベル毎の検索結果として保持される。

【 0 0 8 6 】

次に、XPath 式解析生成部 12 は、これまでの総検索結果件数が x 件未満であるか否か

10

20

30

40

50

を判定する（ステップS40）。総検索結果件数は、ここでは、曖昧レベル0及び1の検索結果件数の総数である。

【0087】

総検索結果件数がx件未満であると判定された場合（ステップS40のYES）、ステップS36に戻って処理が繰り返される。この場合、ステップS36において曖昧レベル2のテンプレートについての処理が実行される。つまり、総検索結果件数がx件を越えるまで、曖昧レベルが低い順に処理が繰り返される。

【0088】

一方、総検索結果件数がx件未満でないと判定された場合（ステップS40のNO）、結果出力部16は、XMLデータ検索部15の記憶部に保持されている検索結果を例えばユーザに対して表示するために出力する（ステップS41）。結果出力部16は、例えば曖昧レベル毎に検索結果を出力する。

10

【0089】

一方、ステップS34において、検索されたXMLデータの検索件数がx件未満でないと判定された場合、ステップS41の処理が実行される。

【0090】

上記したように本変形例においては、ユーザによって指定された検索件数に限定されており、また、曖昧レベルが低いテンプレートの順に処理が実行されるため、例えばXPath式生成テンプレート格納部14に格納されているすべてのテンプレートに基づいて生成された定型XPath式を用いて検索処理が実行された場合より、ノイズデータを少なくし、よりユーザにとって所望の検索結果を得ることが可能となる。

20

【0091】

[第2の変形例]

本実施形態の第2の変形例について説明する。本変形例に係る構造化データ検索装置10は、入力部11に入力された指定レベル情報によって示される曖昧レベルに制限して検索処理を実行する。

【0092】

次に、図14のフローチャートを参照して、構造化データ検索装置10の処理手順について説明する。

まず、入力部11は、ユーザによって指定されたスキーマに準拠するXMLデータを検索するためのXPath式及びユーザによって指定された曖昧レベルを示す指定レベル情報が入力される（ステップS51）。ここでは、ユーザによって指定された曖昧レベル（以下、指定曖昧レベルと表記）は、N（Nは、0以上の整数）であるものとして説明する。

30

【0093】

次に、XMLデータ検索部15は、入力部11によって入力されたXPath式（入力XPath式）を用いて、当該入力XPath式に合致するXMLデータをXMLデータ格納部20から検索する（ステップS52）。XMLデータ検索部15によって検索されたXMLデータは、当該XMLデータ検索部15に含まれる記憶部に保持される（ステップS53）。このとき、このXMLデータは曖昧レベル0の検索結果として保持される。

【0094】

次に、XPath式解析生成部12は、入力部11に入力された指定レベル情報によって示される指定曖昧レベル（ここでは、N）が0であるか否かが判定される（ステップS54）。

40

指定曖昧レベルが0でないと判定された場合（ステップS54のNO）、XPath式解析生成部12は、入力XPath式の解析を行う。これにより、XPath式解析生成部12は、入力XPath式に含まれるパス構成要素の各々を抽出する（ステップS55）。

【0095】

次に、XPath式解析生成部12は、XPath式生成テンプレート読取部13に対して、曖昧レベルnのテンプレートをXPath式生成テンプレート格納部14から読み取る（読み出す）ための要求（読取要求）をする。ここでは、XPath式解析生成部12は、まず曖昧度の

50

低いテンプレートから読み取るよう読取要求をする。

【0096】

XPath式生成テンプレート読取部13は、XPath式解析生成部12からの読取要求に応じて、XPath式生成テンプレート格納部14に格納されているテンプレートを読み取る（ステップS56）。初めに、XPath式生成テンプレート格納部14から、曖昧レベル1のテンプレートが読み取られる。XPath式生成テンプレート読取部13は、読み取ったテンプレートをXPath式解析生成部12に出力する。

【0097】

XPath式解析生成部12は、ステップS55において抽出されたパス構成要素及びXPath式生成テンプレート読取部13によって読み出されたテンプレートに基づいて、定型XPath式を生成する（ステップS57）。 10

【0098】

XMLデータ検索部15は、XPath式解析生成部12によって生成されたXPath式を用いて、当該XPath式に合致するXMLデータをXMLデータ格納装置20から検索する（ステップS58）。XMLデータ検索部15によって検索されたXMLデータは、当該XMLデータ検索部15の記憶部に保持される（ステップS59）。このXMLデータは、曖昧レベル毎の検索結果として保持される。

【0099】

次に、XPath式解析生成部12は、指定曖昧レベルがnであるか否かを判定する（ステップS60）。このとき、nの値は1である。つまり、指定曖昧レベルが1であるか否かが判定される。 20

指定レベル情報によって示される曖昧レベルがn（ここでは1）でないと判定された場合（ステップS60のNO）、ステップS56に戻って処理が繰り返される。この場合では、ステップS56において曖昧レベル2のテンプレートについての処理が実行される。つまり、曖昧レベルがNのテンプレートが処理されるまで、曖昧レベルが低い順に処理が繰り返される。換言すると、曖昧レベルがNより低いレベルに対応付けられているテンプレートについてのみ処理が実行される。

一方、指定曖昧レベルがnであると判定された場合（ステップS60のYES）、結果出力部16は、XMLデータ検索部15の記憶部に保持されている検索結果を例えばユーザに対して表示するために出力する（ステップS61）。結果出力部16は、例えば曖昧レベル毎に検索結果を出力する。 30

【0100】

一方、ステップS54において、指定曖昧レベルが0であると判定された場合、ステップS61の処理が実行される。

【0101】

上記したように本変形例においては、ユーザによって指定された曖昧レベルのテンプレートまでを用いて定型XPath式が生成され、検索処理が実行される。これによって、よりユーザにとって所望の検索結果を得ることが可能となる。

【0102】

なお、本願発明は、上記実施形態またはその各変形例そのままに限定されるものではなく、実施段階ではその要旨を逸脱しない範囲で構成要素を変形して具体化できる。また、上記実施形態またはその各変形例に開示されている複数の構成要素の適宜な組合せにより種々の発明を形成できる。例えば、実施形態またはその各変形例に示される全構成要素から幾つかの構成要素を削除してもよい。更に、異なる実施形態またはその各変形例に亘る構成要素を適宜組合せてもよい。 40

【図面の簡単な説明】

【0103】

【図1】本発明の実施形態に係る構造化データ検索装置の機能構成を示すブロック図である。

【図2】図1に示すXPath式解析生成部12の機能構成を示すブロック図。 50

- 【図3】XPath式生成テンプレート格納部14のデータ構造の一例を示す図。
- 【図4】構造化データ検索装置10の処理手順を示すフローチャート。
- 【図5】図4に示すステップS4の処理の詳細な処理手順を示すフローチャート。
- 【図6】パス構成要素を抽出する処理について具体的に説明するための図。
- 【図7】図4に示すステップS6の処理の詳細な処理手順を示すフローチャート。
- 【図8】XMLデータ格納部20に格納されているXMLデータの一例を示す図。
- 【図9】XMLデータ格納部20に格納されているXMLデータの一例を示す図。
- 【図10】XMLデータ格納部20に格納されているXMLデータの一例を示す図。
- 【図11】XMLデータ格納部20に格納されているXMLデータの一例を示す図。
- 【図12】XMLデータ格納部20に格納されているXMLデータの一例を示す図。
- 【図13】本実施形態の第1の変形例に係る構造化データ検索装置10の処理手順を示すフローチャート。
- 【図14】本実施形態の第2の変形例に係る構造化データ検索装置10の処理手順を示すフローチャート。

10

【符号の説明】

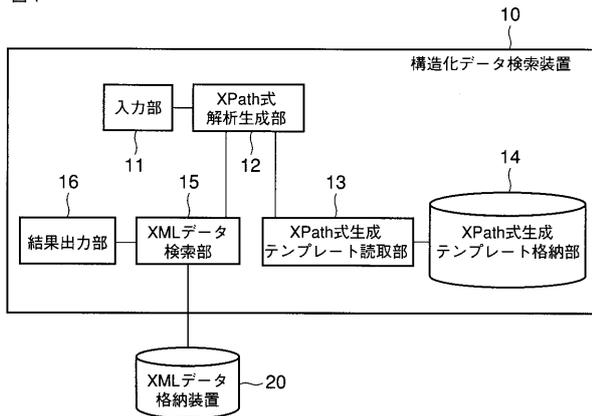
【0104】

10...構造化データ検索装置、11...入力部、12...XPath式解析生成部、13...XPath式生成テンプレート読取部、14...XPath式生成テンプレート格納部、15...XMLデータ検索部、16...結果出力部、20...XMLデータ格納装置、121...XPath式解析部、122...変換部、123...XPath式生成部。

20

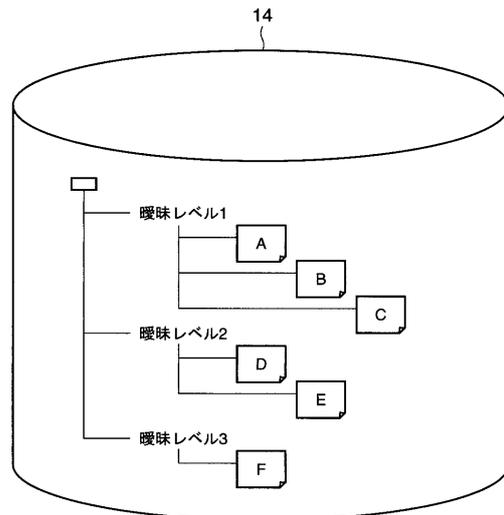
【図1】

図1



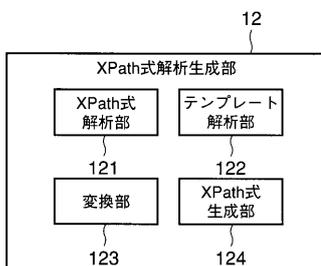
【図3】

図3



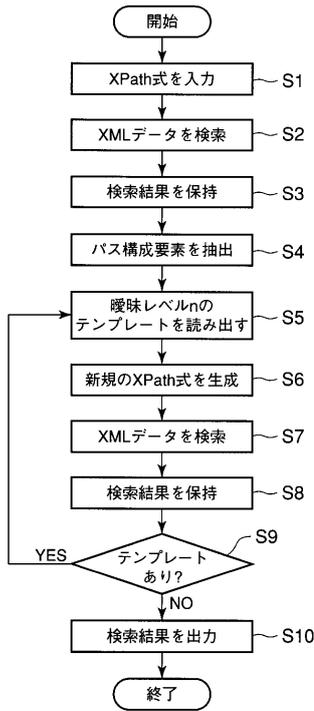
【図2】

図2



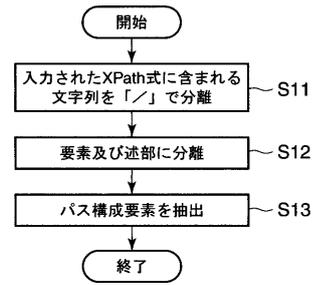
【 図 4 】

図 4



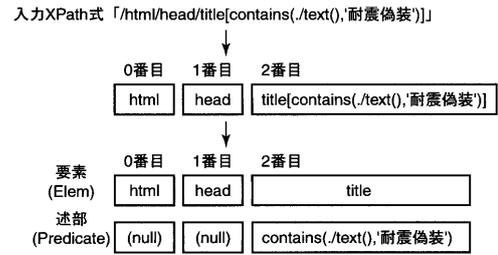
【 図 5 】

図 5



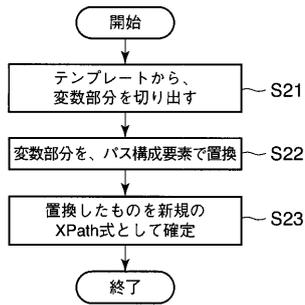
【 図 6 】

図 6



【 図 7 】

図 7



【 図 8 】

図 8

31

```

    <?xml version="1.0" encoding="Shift_JIS"?>
    <html>
    <head>
    <title>[ニュース]東京・府中市のマンションで耐震偽装の疑い</title>
    </head>
    <body>
    <h1>東京・府中市のマンションで耐震偽装の疑い</h1>
    <p>
    (本文省略)
    </p>
    </body>
    </html>
  
```

【 図 9 】

図 9

32

```

<?xml version="1.0"encoding="Shift_JIS"?>
<NewsML>
<!--一部省略-->
<NewsItem>
<!--一部省略-->
<NewsComponent Duid="NC0001"xml:lang="ja">
<NewsLines>
<HeadLine>東京・府中市のマンションで耐震偽装の疑い</HeadLine>
<ByLine Euid="Name">東芝太郎</ByLine>
<ByLine Euid="Company">東芝ソリューション株式会社</ByLine>
<DateLine>20060320T120000+0900</DateLine>
<CopyrightLine>東芝ソリューション株式会社</CopyrightLine>
<KeywordLine>耐震偽装</KeywordLine>
<KeywordLine>マンション</KeywordLine>
</NewsLines>
<!--一部省略-->
<NewsComponent Duid="NC0001-1"xml:lang="ja">
<!--一部省略-->
<DataContent>
(本文省略)
</DataContent>
</NewsComponent>
</NewsItem>
</NewsML>

```

【 図 1 0 】

図 10

33

```

<?xml version="1.0"encoding="Shift_JIS"?>
<rdf:RDF
xmlns=http://purl.org/rss/1.0"
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:content="http://purl.org/rss/1.0/modules/content/"
xmlns:dc="http://purl.org/dc/elements/1.1"
xml:lang="ja">
<channel rdf:about="http://d.hatena.ne.jp/umedamochio/rss">
<title>東芝ニュース</title>
<link>http://example.com/toshiba/news</link>
<description>東芝ニュース</description>
<dc:creator>東芝ソリューション株式会社</dc:creator>
<items>
<rdf:Seq>
<rdf:li rdf:resource="http://example.com/toshiba/news/20060320/01"/>
<rdf:li rdf:resource="http://example.com/toshiba/news/20060320/02"/>
</rdf:Seq>
</channel>
<item rdf:about="http://example.com/toshiba/news/20060320/02">
<title>東京・府中市のマンションで耐震偽装の疑い</title>
<link>http://example.com/toshiba/news/20060320/02</link>
<description>(本文省略)</description>
<content:encoded><![CDATA[(本文省略)]]></content:encoded>
<dc:creator>東芝ソリューション株式会社</dc:creator>
<dc:date>2006-03-20</dc:date>
<dc:subject>社会</dc:subject>
</item>
<item rdf:about="http://example.com/toshiba/news/20060320/01">
<title>日経平均 16000 割れ</title>
<link>http://example.com/toshiba/news/20060320/01</link>
<description>(本文省略)</description>
<content:encoded><![CDATA[(本文省略)]]></content:encoded>
<dc:creator>東芝ソリューション株式会社</dc:creator>
<dc:date>2006-03-20</dc:date>
<dc:subject>経済</dc:subject>
</item>
</rdf:RDF>

```

【 図 1 1 】

図 11

34

```

<?xml version="1.0"encoding="Shift_JIS"?>
<feed
xmlns="http://purl.org/atom/ns#"
version="0.3"
xml:lang="ja">
<!--一部省略-->
<entry xmlns="http://purl.org/atom/ns#">
<!--一部省略-->
<link href="http://example.com/toshiba/news/20060320/02"
title="東京・府中市のマンションで耐震偽装の疑い"/>
<title>東京・府中市のマンションで耐震偽装の疑い</title>
<content>
(本文省略)
</content>
<draft xmlns="http://purl.org/atom-blog/ns#">false</draft>
</entry>
<entry xmlns="http://purl.org/atom/ns#">
<!--一部省略-->
<link href="http://example.com/toshiba/news/20060320/01"
title="日経平均 16000 割れ"/>
<title>日経平均 16000 割れ</title>
<content>
(本文省略)
</content>
<draft xmlns="http://purl.org/atom-blog/ns#">false</draft>
</entry>
</feed>

```

【 図 1 2 】

図 12

35

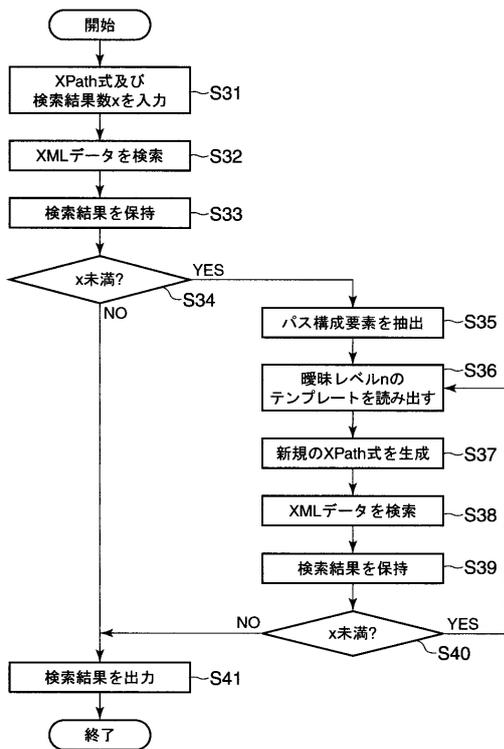
```

<?xml version="1.0"encoding="Shift_JIS"?>
<html>
<head>
<title>[ニュース]日経平均 16000割れ</title>
</head>
<body>
<h1>耐震偽装などが影響、建設業界を中心に日経平均が大幅な下げ</h1>
<p>
(本文省略)
</p>
</body>
</html>

```

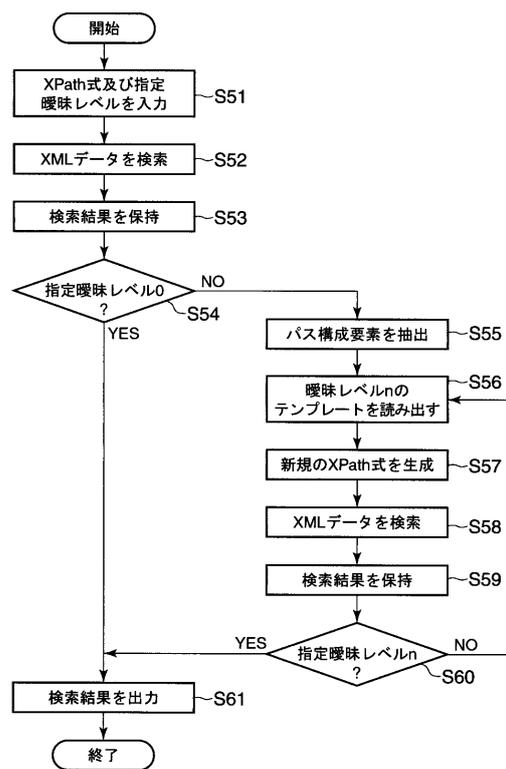
【 図 1 3 】

図 13



【 図 1 4 】

図 14



フロントページの続き

(74)代理人 100075672

弁理士 峰 隆司

(74)代理人 100109830

弁理士 福原 淑弘

(74)代理人 100084618

弁理士 村松 貞男

(74)代理人 100092196

弁理士 橋本 良郎

(72)発明者 矢野 令

東京都港区芝浦一丁目1番1号 東芝ソリューション株式会社内

(72)発明者 川崎 直丸

東京都港区芝浦一丁目1番1号 東芝ソリューション株式会社内

Fターム(参考) 5B075 ND34 PP23 PP26