

(19) 日本国特許庁 (JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2009-544102

(P2009-544102A)

(43) 公表日 平成21年12月10日 (2009. 12. 10)

(51) Int. Cl.	F I	テーマコード (参考)
G06F 17/21 (2006.01)	G06F 17/21 570L	5B075
G06F 17/30 (2006.01)	G06F 17/30 330B	5B082
G06F 12/00 (2006.01)	G06F 17/21 501T	5B109
	G06F 17/30 140	
	G06F 12/00 547H	
審査請求 未請求 予備審査請求 有 (全 21 頁) 最終頁に続く		

(21) 出願番号 特願2009-520892 (P2009-520892)
 (86) (22) 出願日 平成19年7月9日 (2007. 7. 9)
 (85) 翻訳文提出日 平成21年2月24日 (2009. 2. 24)
 (86) 国際出願番号 PCT/US2007/073087
 (87) 国際公開番号 W02008/011294
 (87) 国際公開日 平成20年1月24日 (2008. 1. 24)
 (31) 優先権主張番号 11/489, 426
 (32) 優先日 平成18年7月18日 (2006. 7. 18)
 (33) 優先権主張国 米国 (US)

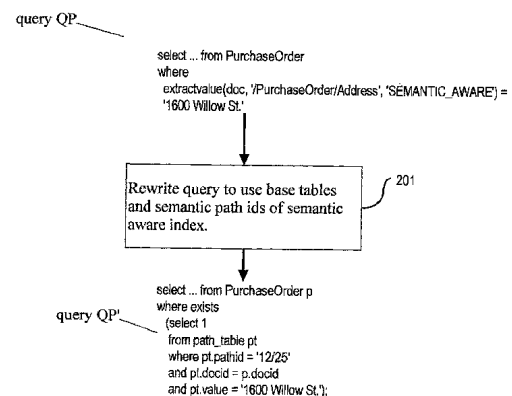
(71) 出願人 502303739
 オラクル・インターナショナル・コーポレーション
 アメリカ合衆国、94065 カリフォルニア州、レッドウッド・ショアーズ、オラクル・パークウェイ、500
 (74) 代理人 100064746
 弁理士 深見 久郎
 (74) 代理人 100085132
 弁理士 森田 俊雄
 (74) 代理人 100083703
 弁理士 仲村 義平
 (74) 代理人 100096781
 弁理士 堀井 豊

最終頁に続く

(54) 【発明の名称】 XML文書の、意味論を意識した処理

(57) 【要約】

XML文書の、意味論を意識した処理は、異なる名前を有するが意味論的に同等である要素を、照会およびスキーマ検証などの要素名に依存した動作を実行する際に、同じ要素として扱う。当該意味論を意識した処理は、意味論的に同等の名前のセットの各要素名を「正規のタグ名」にマップするマッピングに基づいている。



【特許請求の範囲】**【請求項 1】**

コンピュータによって実現されるステップを含む方法であって、前記コンピュータによって実現されるステップは、

正規のタグ名を、第 1 のノードの第 1 の名前と、前記第 1 の名前とは異なる第 2 のノードの第 2 の名前とにマップする意味論的マッピングを格納するステップを含み、XML 文書の集合は前記第 1 のノードおよび前記第 2 のノードを含み、前記コンピュータによって実現されるステップはさらに、

前記意味論的マッピングに基づき、前記第 1 の名前および第 2 の名前を同一の名前として扱うことによりタグ名オペレーションを実行するステップを含む、方法。

10

【請求項 2】

前記タグ名オペレーションは、前記 XML 文書の集合に対して発行されたクエリのコンピュータ処理である、請求項 1 に記載の方法。

【請求項 3】

前記タグ名オペレーションはスキーマ検証を含む、請求項 1 に記載の方法。

【請求項 4】

前記タグ名オペレーションは、前記 XML 文書の集合へのアクセスを管理するレポジトリによって実行される、請求項 1 に記載の方法。

【請求項 5】

前記コンピュータによって実現されるステップは、前記意味論的マッピングを表わすデータを登録する要求を受信するステップと、

20

前記要求に応答して、前記データを前記意味論的マッピングとして登録するステップとをさらに含む、請求項 4 に記載の方法。

【請求項 6】

コンピュータによって実現されるステップを含む方法であって、前記コンピュータによって実現されるステップは、

XML 文書の集合における複数のノードの各ノードについて、

意味論的マッピングに基づいて意味論的経路 ID を生成するステップを含み、

前記複数のノードは第 1 のノードおよび第 2 のノードを含み、

第 1 の名前は前記第 1 のノードまたは前記第 1 のノードの先祖ノードに関連付けられており、

30

第 2 の名前は前記第 2 のノードまたは前記第 2 のノードの先祖に関連付けられており、

前記意味論的マッピングは、正規のタグ名を前記第 1 の名前および前記第 2 の名前にマップし、

前記第 1 のノードおよび前記第 2 のノードのために生成された意味論的経路 ID は同一である、方法。

【請求項 7】

前記各ノードのための意味論的経路 ID は、前記各ノードの経路におけるノードの各々の名前についてのコードを含み、

前記第 1 の名前についてのコードおよび前記第 2 の名前についてのコードは同じである、請求項 6 に記載の方法。

40

【請求項 8】

前記コンピュータによって実現されるステップはさらに、

前記複数のノードのために生成される意味論的経路 ID によって前記複数のノードに索引付けするインデックスを作成するステップを含む、請求項 6 に記載の方法。

【請求項 9】

前記 XML 文書の集合はデータベースサーバによって管理され、前記コンピュータによって実現されるステップはさらに、

XML 文書の集合に対して発行されたクエリを受信するステップを含み、前記クエリは経路を指定し、前記コンピュータによって実現されるステップはさらに、

50

前記経路に基づいて、データベースサーバが前記インデックスにアクセスするよう前記クエリを書換えるステップを含む、請求項 6 に記載の方法。

【請求項 10】

命令の 1 つ以上のシーケンスを搬送するコンピュータ読取可能媒体であって、前記命令の 1 つ以上のシーケンスは、1 つ以上のプロセッサによって実行されると、請求項 1 に記載の方法を前記 1 つ以上のプロセッサに実行させる、コンピュータ読取可能媒体。

【請求項 11】

命令の 1 つ以上のシーケンスを搬送するコンピュータ読取可能媒体であって、前記命令の 1 つ以上のシーケンスは、1 つ以上のプロセッサによって実行されると、請求項 2 に記載の方法を前記 1 つ以上のプロセッサに実行させる、コンピュータ読取可能媒体。

10

【請求項 12】

命令の 1 つ以上のシーケンスを搬送するコンピュータ読取可能媒体であって、前記命令の 1 つ以上のシーケンスは、1 つ以上のプロセッサによって実行されると、請求項 3 に記載の方法を前記 1 つ以上のプロセッサに実行させる、コンピュータ読取可能媒体。

【請求項 13】

命令の 1 つ以上のシーケンスを搬送するコンピュータ読取可能媒体であって、前記命令の 1 つ以上のシーケンスは、1 つ以上のプロセッサによって実行されると、請求項 4 に記載の方法を前記 1 つ以上のプロセッサに実行させる、コンピュータ読取可能媒体。

【請求項 14】

命令の 1 つ以上のシーケンスを搬送するコンピュータ読取可能媒体であって、前記命令の 1 つ以上のシーケンスは、1 つ以上のプロセッサによって実行されると、請求項 5 に記載の方法を前記 1 つ以上のプロセッサに実行させる、コンピュータ読取可能媒体。

20

【請求項 15】

命令の 1 つ以上のシーケンスを搬送するコンピュータ読取可能媒体であって、前記命令の 1 つ以上のシーケンスは、1 つ以上のプロセッサによって実行されると、請求項 6 に記載の方法を前記 1 つ以上のプロセッサに実行させる、コンピュータ読取可能媒体。

【請求項 16】

命令の 1 つ以上のシーケンスを搬送するコンピュータ読取可能媒体であって、前記命令の 1 つ以上のシーケンスは、1 つ以上のプロセッサによって実行されると、請求項 7 に記載の方法を前記 1 つ以上のプロセッサに実行させる、コンピュータ読取可能媒体。

30

【請求項 17】

命令の 1 つ以上のシーケンスを搬送するコンピュータ読取可能媒体であって、前記命令の 1 つ以上のシーケンスは、1 つ以上のプロセッサによって実行されると、請求項 8 に記載の方法を前記 1 つ以上のプロセッサに実行させる、コンピュータ読取可能媒体。

【請求項 18】

命令の 1 つ以上のシーケンスを搬送するコンピュータ読取可能媒体であって、前記命令の 1 つ以上のシーケンスは、1 つ以上のプロセッサによって実行されると、請求項 9 に記載の方法を前記 1 つ以上のプロセッサに実行させる、コンピュータ読取可能媒体。

【請求項 19】

X M L 文書の集合における複数のノードのインデックスを記憶するコンピュータ読取可能媒体であって、

40

前記複数のノードの各ノードは、前記各ノードを含む或る経路に関連付けられており、
前記インデックスの各エントリは、前記複数のノードの特定のノードに対応し、前記特定のノードの或る経路を表わす意味論的経路 I D に前記ノードに関連付け、

前記複数のノードは第 1 のノードおよび第 2 のノードを含み、

第 1 の名前は、前記第 1 のノードまたは前記第 1 のノードの先祖ノードに関連付けられており、

第 2 の名前は、前記第 2 のノードまたは前記第 2 のノードの先祖ノードに関連付けられており、

前記第 1 のノードおよび前記第 2 のノードについてのそれぞれの意味論的経路 I D は同

50

一である、コンピュータ読取可能媒体。

【請求項 20】

前記各ノードについての意味論的経路 ID は、前記各ノードの経路におけるノードの各々の名前についてのコードを含み、

前記第 1 の名前および前記第 2 の名前についてのそれぞれのコードは同じである、請求項 10 に記載のコンピュータ読取可能媒体。

【発明の詳細な説明】

【技術分野】

【0001】

関連出願

10

本願は、2004年7月2日に出願されたシバサンカラン・チャンドラセカラン (Siva sankaran Chandrasekaran) による「XML データにアクセスするためのインデックス (Index For Accessing XML Data)」と題された米国出願連続番号第 10 / 884, 311 号に関し、その内容の全体があらゆる目的で引用によりこの明細書中に援用されている。

【0002】

発明の分野

この発明は XML データの処理に関する。

【背景技術】

【0003】

背景

20

拡張マークアップ言語 (XML) はデータおよび文書についての規格であり、コンピュータ業界で広く受け入れられている。XML は、この明細書中で XML 文書またはそのフラグメントと称されるファイルまたはデータパケットなどのデータの塊に構造を与え、当該構造を記述する。XML 規格により、XML 要素と称される XML エンティティの区域の範囲を定めるタグが提供される。各々の XML 要素は、属性と称される 1 つ以上の名前 - 値の対を含み得る。XML を説明するために以下の XML フラグメント A が与えられる。

【0004】

フラグメント F A

【0005】

【数 1】

30

```
<book>My book
    <publication publisher="Doubleday"
        date="January"></publication>
    <Author>Mark Berry</Author>
    <Author>Jane Murray</Author>
</book>
```

【発明の概要】

40

【発明が解決しようとする課題】

【0006】

XML 要素は、開始タグおよび対応する終了タグによって範囲が定められる。たとえば、セグメント A は、要素の範囲を定めるための開始タグ <Author> および終了タグ </Author> を含む。要素間のデータは要素の内容と称される。この要素の場合、要素の内容はテキストデータ Mark Berry となる。

【0007】

要素は、この明細書中においてはその要素名で言及される。たとえば、開始タグ <publication> および終了タグ </publication> によって範囲が定められる要素は publication と称される。

50

【 0 0 0 8 】

要素内容は、属性および他の要素を含む他のさまざまな種類のデータを含み得る。要素bookは、1つ以上の要素を含む要素の一例である。具体的には、bookは、publicationおよびauthorという2つの要素を含む。別の要素に含まれる要素は、その要素の子孫と称される。このため、要素publicationおよびauthorは要素bookの子孫となる。要素の属性は、その要素が含むものと見なされてもよい。

【 0 0 0 9 】

属性および子孫要素を含む要素を規定することにより、XML文書は、要素とその子孫要素とその属性との間の階層ツリー関係を規定する。このような階層ツリー関係を有する要素の如何なるセットも、この明細書中ではXML文書またはフラグメントと称される。

10

【 0 0 1 0 】

ノードツリーモデル

XMLについての重要な規格は、XQuery1.0およびXPath2.0データモデルである（引用によりこの明細書中に援用されるW3C作業草案（2004年7月9日）を参照されたい）。このモデルの一面は、XML文書が、当該XML文書の階層の性質を反映したノードの階層によって表わされることである。ノードの階層は複数のレベルのノードで構成される。各レベルのノードは各々が、異なるレベルにおける1つ以上のノードにリンクされている。最上のレベルよりも低いレベルにある各ノードは、上方のレベルにおける親ノードのうち1つ以上のノードの子ノードとなる。同じレベルにあるノードは兄弟ノードとなる。ツリー階層またはノードツリーにおいては、各々の子ノードは親ノードを1つしか持たないが、親ノードは子ノードを複数有していてもよい。ツリー階層においては、親ノードがリンクされていないノードは根ノードとなり、子ノードがリンクされていないノードは葉ノードとなる。ツリー階層は単一の根ノードを有する。

20

【 0 0 1 1 】

XML文書を表わすノードツリーにおいては、ノードは、要素に対応し得る。当該ノードの子ノードは、当該要素に含まれる属性または別の要素に対応する。

【 0 0 1 2 】

ノードは名前に関連付けられてもよい。たとえば、要素bookを表わすノードの名前はbookとなる。属性publisherを表わすノードの場合、ノードの名前はpublisherとなる。

30

【 0 0 1 3 】

表現を簡便にするために、XML文書の他の部分および要素は、文書を表わすノードのツリー内のノードと見なされる。このため、名前bookを有するノードの値を「My book」と称することは、ノードbookに関連付けられる要素の値がMy bookであることを表わすのに簡便な方法となる。要素、属性またはノードの名前はまた、この明細書中においてはタグ名と称される。

【 0 0 1 4 】

XML文書におけるノードのための経路は、一連の親 - 子リンクを反映しており、XML文書におけるノードから始まって、階層をさらに下って特定のノードに到達する。たとえば、XML文書の根からノードpublicationまでの経路は、「/book/publication」となる。

40

【 0 0 1 5 】

同じ意味論についてのタグ名の普及

XMLの人気の高い1つの理由として、テキストを含むタグ名が記述的に使用可能であり、このため、要素および属性の意味を伝えるのに使用可能であることが挙げられる。たとえば、要素<address>は、アドレスを表わすデータを格納するのに用いられる。

【 0 0 1 6 】

しかしながら、タグ名は、しばしば、特定のアプリケーションまたはプロジェクトを実現する独立した個体またはグループによって作成される。このため、同じ意味が、最終的に、異なるXML文書内の異なるタグ名で表わされてしまう可能性がある。規格委員会または業界共同事業体からいくつかのXML語彙が出てきているが、これらは、使用されて

50

いるXMLタグ名全体のうちほんのわずかの部分を占めているに過ぎない。タグ名の普及は続いており、同様の意味または同じ意味を示す多くの異なるタグ名が場当たりに作成されている。この問題は、同じ企業内のグループ間や異なる企業間で発生している。

【0017】

たとえば、アドレス値は、1つのXML文書内における要素<Address>によって表わされ得るが、別の文書における異なる要素<Addr>によって表わされてもよい。さらに、これらのタグは異なる名前空間を用いていてもよい。たとえば、企業C1が<c1:Address>を用い得るのに対して、企業C2は<c2:Address>を用いる。XMLの観点から見ると、これらの企業が規定するこれらのタグおよび要素は異なっており、異なる物を意味するものと想定される。

10

【0018】

互いに異なっているが意味論的に同一のものとして扱われる可能性のあるタグ名のセットは、この明細書中においては、意味論的に同等である異種のタグ名と称される。上述の例においては、<Address>、<Addr>、<c1:Address>および<c2:Address>は意味論的に同等である異種のタグ名を有する。

【0019】

レポジトリ内のタグ名の普及

異なる語彙(すなわち、タグ名のセット)に基づいたXML文書が最終的に(XMLデータベースなどの)単一のレポジトリに行き着く場合の多くのシナリオが存在する。これは、データ統合、ウェブサービスおよび内容ルーティングに共通している。このような場合、レポジトリにおけるXML文書の集合にわたってクエリを作成することが非常に困難になる。上述の例においては、文書にわたってアドレスをチェックするためのクエリは、異なる文書における意味論的に同等の要素にアクセスするために異なるタグ名を用いる複雑なクエリを必要とする。

20

【0020】

このようなクエリの実現可能な公式は、以下のとおりである。

【0021】

【数2】

```
select ... from PurchaseOrder
where extractvalue(doc, '/PurchaseOrder/Address') =
'1600 Willow St.'
or extractvalue(doc, '/PurchaseOrder/Addr') = '1600
Willow St.';
```

30

【0022】

明らかに、使用されるXPath式の複雑度が増し、意味論的に同等のタグ名の数XML集合の範囲内で増加すると、上述の方策が実行不可能となる。クエリの複雑さに加えて、このようなクエリは性能が低下してしまう。この欠点は、XPath、XQueryおよびXSLTなどのXMLについての標準的なクエリおよび変換言語すべてが被る。

40

【0023】

非葉ノードについてのタグ名の普及は、タグ名の普及についての問題を悪化させてしまう。先祖ノードの子孫が同じタグ名を有しているが、先祖が意味論的に同等であるが異なる名前を有している場合、異なる経路列は子孫と称される必要がある。たとえば、XML文書のいくつかのセットは、発行者およびそのアドレスを表わす要素を含む。しかしながら、1つのサブセットにおいて要素<publisher>が用いられ、別のサブセットにおいて要素<publishing company>が用いられる。これらはともに子孫要素<address>、<city>および<zip>を含む。これら両方のサブセットについては、意味論的に同等の子孫要素を表わすのに同じタグ名が用いられるが、サブセット間においては、子孫要素を識別す

50

るのに異なるXPath列が用いられなければならない。たとえば、要素<address>を参照するために、1つのサブセットにおいてはXPath列/publisher/address/が用いられ、他方のサブセットにおいてはXPath列/publishing company/address/が用いられる。

【0024】

タグ名の普及に対処する別の方策として、同じ意味に対して同じタグ名を用いるようすべての文書を正規化することが挙げられる。たとえば、XML文書の集合の範囲内で、意味論的に同等のすべてのアドレス要素を<Address>に変更する。すると、XMLの集合を有するアドレス要素にアクセスするクエリは、1つのタグ名を参照するだけで済む。この方策についての主な不利点は、元の文書の忠実性が保全されないことである。

【0025】

上述に基づき、タグ名の普及に対処する方法の改善が必要とされる。

この節に記載される方策は、追求可能な方策であるが、必ずしも以前に想到または追及された方策である必要はない。したがって、特に指定のない限り、この段落に記載される方策はいずれも、単にこの段落における記載によって先行技術と見なされるものと想定されるべきではない。

【0026】

この発明は、同様の参照番号が同様の要素を指す添付の図面において、限定的ではなく例示的に図示される。

【図面の簡単な説明】

【0027】

【図1】この発明の実施例に従った意味論的経路IDに基づいたXMLインデックスを示す図である。

【図2】この発明の実施例に従ったクエリの、意味論を意識した書換えを示す図である。

【図3】この発明の実施例において用いられ得るコンピュータシステムを示す図である。

【発明を実施するための形態】

【0028】

詳細な説明

以下の記載においては、この発明を完全に理解できるようにするために、説明の目的で、多数の具体的な詳細が提供される。しかしながら、これらの具体的な詳細なしにこの発明が実施可能であることが明らかとなるだろう。他の例においては、周知の構造および装置は、この発明を不必要に曖昧にすることを避けるために、ブロック図の形で示される。

【0029】

概要

この明細書中に記載される方策により、「タグ名オペレーション」を実行する際に意味論的に同等である異種のタグを同じタグ名として扱うことが可能となる。タグ名オペレーションは、ノードのタグ名に依存するオペレーションである。タグ名オペレーションの例には、クエリQAなどのXPath列を用いてXMLデータを参照するクエリのコンピュータ処理が含まれる。タグ名オペレーションの別の例としてスキーマ検証があり、XML文書がXMLスキーマに適合するかどうか判断される。

【0030】

当該方策は、意味論的に同等のタグ名のセットの各タグ名を「正規のタグ名」にマップするマッピングに基づいている。意味論的に同等のタグ名は、個々に、正規のタグ名の同義語および互いの同義語と見なされる。タグ名オペレーションは、同義語が、当該同義語がマップされている正規のタグ名と同一であるかのように実行される。この態様でタグ名オペレーションを実行することを、この明細書中では、意味論を意識した処理と称する。

【0031】

たとえば、XML文書の集合は、以下の意味論的に同等のアドレスタグ名のセット、すなわち、Address、Addr、c1:Addressおよびc2:Addressを含む。

【0032】

以下のXMLフラグメントXAは、これらの意味論的に同等のアドレスタグ名を正規の

10

20

30

40

50

タグ名Addressにマップする。

【 0 0 3 3 】

フラグメント X A

【 0 0 3 4 】

【 数 3 】

```
<element name="Address">
  <synonym name="Addr"/>
  <synonym name="c1:Address"/>
  <synonym name="c2:Address"/>
```

10

【 0 0 3 5 】

以下のクエリ Q B をコンピュータ処理すると、

【 0 0 3 6 】

【 数 4 】

```
select ... from PurchaseOrder
where extractvalue(doc, '/PurchaseOrder/Address') =
'500'
```

20

【 0 0 3 7 】

となり、以下の経路によって識別される要素は、以下の経路内にあるものとして扱われる。

【 0 0 3 8 】

クエリ Q B において指定される /PurchaseOrder/Address
/PurchaseOrder/Address、
/PurchaseOrder/Addr、
/PurchaseOrder/c1:Address、および、
/PurchaseOrder/c2:Address。

【 0 0 3 9 】

30

同義語を正規のタグ名にマップするマッピングは、この明細書中では意味論的マッピングと称される。意味論的マッピングを表わす一方法の例として、XML 文書またはフラグメント A などのフラグメントの使用が挙げられる。この発明は、同義語マッピングを表わす特定の如何なる方法にも限定されない。

【 0 0 4 0 】

この発明の一実施例に従うと、タグ名オペレーションの、意味論を意識した処理は、XML レポジトリによって実行される。XML レポジトリは、この用語がこの明細書中で用いられる場合、XML 文書へのアクセスを格納および管理するコンピュータシステムを指す。具体的には、レポジトリは、統合ソフトウェア構成要素とメモリ、ディスクストレージ、コンピュータなどの計算リソースの割当との組合せであり、プロセッサ上で統合ソフトウェア構成要素を実行するためにノード上で処理を行なう。ソフトウェアと計算リソースとの組合せは、XML 文書の保存および XML 文書へのアクセスの管理を専用とする。典型的には、レポジトリは、XML 文書にアクセスするかまたは XML 文書进行处理するようクエリを発行するクライアントの代わりに、XML 文書を保存しかつ XML 文書にアクセスするのに用いられる。レポジトリによって処理されるクエリは、XML クエリ言語（「XQuery」）および XML 経路言語（「XPath」）などの XML 規格に準拠する。XPath は、この明細書中において引用により援用されている XML 経路言語（XML Path Language）（XPath）、バージョン 1.0（W3C 推奨（W3C Recommendation）（1999 年 1 月 16 日））に記載されている。XPath2.0 および XQuery1.0 は、この明細書中において引用により援用されている XQuery1.0 および XPath2.0（W3C 候補推奨（W3C Candidate Rec

40

50

ommendation) (2005年11月3日)に記載されている。

【0041】

経路IDおよびインデックス

この発明の実施例に従うと、XMLレポジトリは意味論的経路IDインデックスを用いる。経路IDは、XML文書内におけるノード間の経路のための識別子である。XML文書におけるノードのための経路は、一連の親-子リンクを反映しており、XML文書におけるノードから始まって、階層をさらに下って特定のノードに到達する。経路は、しばしば、経路におけるノードの名前の連続を表わす列である経路式によって表わされる。たとえば、XML文書D2の根からノードPublicationまでの経路は、経路式「/Book/Publication」によって表わされる。

10

【0042】

ノードの名前は非常に長くてもよい。経路式の長さを短縮し、経路式を記憶するのに必要なストレージの量を減らすために、名前に基づいた経路式の代わりに経路IDを用いてもよい。

【0043】

経路IDは、ノード名の代わりに用いられるノードIDコードからなる。経路IDにおいては、名前に基づいた経路式の対応する各ノード名のためのノードコードが存在する。

【0044】

説明のために、2つのXML文書について考察する。

文書D1

20

【0045】

【数5】

```
<Purchase Order>
```

```
...
```

```
<Addr>10 Main St</Address>
```

```
...
```

```
</Purchase Order>
```

30

【0046】

文書D2

【0047】

【数6】

```
<Purchase Order>
```

```
...
```

```
<Address>500 Oracle Pkwy</Addr>
```

```
...
```

```
</Purchase Order>
```

40

【0048】

ノードIDコード12、23および24は、それぞれノードPurchaseOrder、AddrおよびAddressに割当てられる。このため、経路「/Purchase Order/Addr」についての経路IDは「/12/23」となり、「/Purchase Order/Address」についての経路IDは「/12/24」となる。さらに、経路IDはそれ自体が、経路全体のためにより短い識別子を割当てる、すなわち「/12/23」に対して42、「/12/24」に対して43を割当てる別個のシステム経路IDテーブルに記憶されてもよい。

【0049】

50

経路IDは、経路IDによってXML文書の集合におけるノードに索引付けするインデックスを生成するのに用いられてもよい。経路IDが用いる記憶空間が少ないので、経路IDは、全ノード名に基づいたインデックス経路式のストレージのオーバーヘッドを発生させることなく、それらの経路に基づいて索引付けすべきノードに索引付けする。「XMLデータにアクセスするためのインデックス」には、経路テーブルおよび第2のインデックスを含むインデックスの例が記載される。

【0050】

意味論的経路ID

意味論的経路IDは、経路式の意味論的な等価に基づいて生成される経路IDである。所与の経路式のために、その意味論的に同等の名前ベースの経路式は、その同義語の代わりに正規のタグ名を含む。同義語マッピングを用いて、同義語をどの正規のタグ名にマップするかを決定する。意味論的経路IDは、意味論的に同等の経路式のノードIDコードに基づいており、正規のタグ名のノードIDコードが、正規のタグ名の同義語のノードIDコードの代わりに用いられる。たとえば、正規のタグ名ADDRESSのノードIDコードは25である。このため、経路「/Purchase Order/Addr」についての意味論的経路IDは「/12/25」となり、「/Purchase Order/Address」についても「/12/25」となる。

【0051】

経路IDと同様に、インデックスは、ノードの意味論的経路IDによってXML文書の集合におけるノードに索引付けしてもよい。このようなインデックスは、この明細書中では意味論を意識したインデックスと称される。意味論的に同等の経路を有するノードは同じ意味論的経路IDに索引付けされ、意味論的インデックスの局面は、意味論的に同等である異種の名前をもったノードからのXMLデータのクエリおよび検索を最適化するのに用いることができる。これを如何に実現し得るかについては、XML文書の保存およびXML文書への照会のために構成および/または向上されるオブジェクト/リレーショナルデータベースサーバを含むXMLレポジトリの文脈内において説明される。

【0052】

レポジトリ/データベースサーバ上のXMLストレージ

一実施例に従うと、XMLレポジトリは、XML文書の保存およびXML文書への照会のために構成および/または向上されたオブジェクト/リレーショナルデータベースサーバからなる。このようなデータベースサーバにおいては、XML文書はテーブルの行に保存されてもよく、XML文書のノードは、その行における別個の列に格納される。XML文書全体またはそのフラグメントはまた、列におけるlob(ラージオブジェクト)に保存されてもよい。XML文書はまた、データベースにおけるオブジェクトの階層として保存されてもよく、各オブジェクトは、オブジェクトクラスのインスタンスであり、XML文書の1つ以上の要素を格納する。オブジェクトクラスは、たとえば、要素に対応する構造を規定し、要素の直近の子孫を表わすオブジェクトに対するポインタまたは基準を含む。XML値を保持するデータベースシステムのテーブルおよび/またはオブジェクトは、この明細書中においてはベーステーブルまたはオブジェクトと称される。

【0053】

オブジェクト・リレーショナルデータベースサーバは、少なくとも部分的にXML規格、たとえばXQuery/XPath、および他の規格、たとえばXQL/XML規格(この明細書中において引用により援用されているINCIITS/ISO/IEC 9075-14:2003を参照されたい)に準拠するクエリを実行する。

【0054】

説明のために、この発明の実施例を、XML文書の保存およびXML文書への照会のために構成および/または向上されたオブジェクト/リレーショナルデータベースサーバを含むデータベースサーバの形態であるレポジトリを参照し、そして、XMLデータを格納するためにこのようなデータベースサーバによって用いられるベースデータ構造を参照して説明することとする。しかしながら、この発明の実施例はこのようなレポジトリには限定されない。

10

20

30

40

50

【 0 0 5 5 】

インデックス

一実施例に従うと、データベースサーバは、XML文書の集合に索引付けする「論理インデックス」を維持する。論理インデックスは、1つ以上のXML文書のセットなどの別のデータの塊へのアクセスのために協働的に用いられる複数の構造を含み得る。この発明の実施例に従うと、論理インデックスは、この明細書中においてXMLインデックスと称されており、経路テーブルを含む。当該経路テーブルは、XML文書の集合におけるノードの階層についての情報を含んでおり、ノードの値を含んでいてもよい。論理インデックスは、経路テーブルに索引付けする順序付けされたインデックスを含む他のインデックスを含み得る。順序付けされたインデックスは、インデックスキーに基づいて順序付けされたエントリを含む。

10

【 0 0 5 6 】

図1は、一実施例に従ったXMLインデックスの経路テーブル102を示す。経路テーブルは、XML文書の集合についての階層情報を含む。経路テーブル102を、文書D1およびD2を参照して説明する。

【 0 0 5 7 】

経路テーブル102は、列RID、LOCATOR（ロケータ）、VALUE（値）、ORDERKEY、PATHID（経路ID）およびSEMANTIC PATHID（意味論的経路ID）を含む。経路テーブル102における行は各々、文書D1およびD2を含むXML文書の集合におけるノードに対応する。列RIDは行の行IDを含む。経路テーブル102における特定の行のノードに関して、行IDは、ノードを含むベーステーブルにおける行を識別する。経路テーブル102のエントリの1つのセットが、LOB列における文書D1のノードを保持する行R1を識別する。エントリ103は、文書D1におけるノード/Purchase Order/Addrに対応する。経路テーブル102のエントリの別のセットが、文書D2のノードを含む行R2を識別する。エントリ104は、文書D1におけるノード/Purchase Order/Addressに対応する。

20

【 0 0 5 8 】

列LOCATOR（ロケータ）は、XML文書のデータ表現の範囲内におけるノードの位置を示す値であるノードロケータを含む。たとえば、XML文書を表わすテキストのストリームに関して、ノードロケータは、テキストのストリーム内における、ノードを表わすテキストの開始バイト位置を表わす値であり得る。別の例として、関連するオブジェクトのセットはXML文書のノードを表わし得る。ノードロケータは、ノードを表わすオブジェクトに対する基準であってもよい。

30

【 0 0 5 9 】

列VALUE（値）はノードの値を含む。代替的には、経路テーブルはノードの値を保持する列を省いてもよい。これらの値は、ノードロケータによって識別される位置から検索することによって得ることができる。

【 0 0 6 0 】

列PATHID（経路ID）は経路IDを保持する。エントリおよびそのそれぞれのノードに関して、経路IDの列は、ノードの経路IDを保持する。ノード/Purchase Order/Addrのエントリに関して、経路IDは値「12/23」を保持する。/Purchase Order/Addressのエントリに関して、経路IDは値「/12/24」を保持する。

40

【 0 0 6 1 】

列SEMANTIC PATHID（意味論的経路ID）は意味論的経路IDを保持する。エントリおよびそのそれぞれのノードに関して、SEMANTIC PATHIDはノードの意味論的経路IDを保持する。/Purchase Order/Addrおよび/Purchase Order/Addrが意味論的に同等の経路を有しているので、列SEMANTIC PATHIDにおけるそれぞれの意味論的経路IDは、同一の「12/25」となる。

【 0 0 6 2 】

タグ名オペレーションについての意味論的マッピングの登録

50

ユーザは、XML文書の集合についての意味論的マッピングをレポジトリに登録して、登録された意味論的マッピングに従って、意味論を意識した態様で、XML文書上でレポジトリにタグ名オペレーションを実行させる。一実施例に従うと、登録は、意味論を意識したインデックスを作成するプロセス中に行なわれる。たとえば、意味論的インデックスを作成するために、ユーザは、XML文書の集合についてのXMLインデックスを作成するようデータベースサーバにDDL(「データ定義言語」)コマンドを発行する。当該コマンドは、データベースサーバによって保存され意味論的マッピングを表わすXML文書を参照する。コマンドの受信に応じて、データベースサーバはコマンドを実行して、登録された意味論的マッピングに基づいて意味論を意識したインデックスを作成する。次いで、データベースサーバは、意味論的マッピングに基づいて、かつ意味論的マッピングに従ってタグ名前オペレーションを実行する。意味論を意識したインデックスによって索引付けされたXMLの集合に文書が追加されると、インデックスは意味論的マッピングに従って維持される。

10

【0063】

クエリについての意味論を意識した書換え

図2は、クエリが意味論を意識した態様でコンピュータ処理されるようにデータベースサーバがクエリQPを書換えるクエリ書換え動作を示す。

【0064】

クエリQPは、XML文書D1およびD2を含むXML文書の集合に対してユーザによって発行される。クエリQPは、クエリQPが意味論を意識した方法で評価されるべきであることを指定するパラメータ値「SEMATI_C__AWARE」を有するextractvalue関数を含む。クエリのコンピュータ処理などの意味論を意識した処理は、さまざまな方法で指示可能であり、この発明は、特定の如何なる方法にも限定されない。意味論を意識した処理ではシステムが広範囲に指定され得る。たとえば、ユーザは、XML文書の集合に対して発行されたすべてのクエリが意味論を意識した処理を受けなければならないことを指定し得る。意味論を意識した処理は、たとえば、クライアントがデータベースサーバでセッションを確立することによって、または、クエリQPによって示されるように、明確なクエリパラメータによってセッションレベルで指定され得る。

20

【0065】

フラグメントXAがXMLの集合のための意味論的マッピングとしてデータベースサーバに登録されていたので、意味論を意識した書換えは、この意味論的マッピングおよび経路テーブル102に基づく。

30

【0066】

ステップ200において、クエリQPはクエリQPに書換えられ、これが、extractvalue関数によって与えられる経路の、意味論的に同等の経路IDに適合するエントリを探索する。意味論的に同等の経路IDは「12/25」である。これは、XMLの集合について登録された意味論的マッピングに基づいて生成された。なお、文書内の実際の経路が(/PurchaseOrder/Addressではなく)/PurchaseOrder/Addrであったとしても、文書D2がクエリQPによって選択されることに留意されたい。

40

【0067】

他の実施例

上述のとおり、説明された方策は、さまざまな形態のタグ名オペレーションに適用可能であり、クエリのコンピュータ処理または評価には限定されない。タグ名オペレーションの別の例としてスキーマ検証がある。スキーマ検証によって、XML文書がXMLスキーマに適合するかどうか判断される。

【0068】

XMLスキーマは、特定の種類のXML文書の構造を規定する。たとえば、XMLスキーマは、XML文書に含まれる要素についての名前、XML文書に含まれる要素間の階層関係、およびXML文書に含まれる値の種類を指定し得る。XMLスキーマを規定する規格には、その内容が引用によりこの明細書中に援用されているXMLスキーマ、パート0

50

、パート 1、パート 2 (XML Schema, Part 0, Part 1, Part 2) (W3C 推奨 (2001 年 5 月 2 日)) と、その内容が引用によりこの明細書中に援用されている XML スキーマ、パート 1 (XML Schema, Part 1) : 構造 (第 2 版、W3C 推奨 (2004 年 10 月 28 日)) と、その内容が引用によりこの明細書中に援用されている XML スキーマ、パート 2 (XML Schema, Part 2) : データタイプ (第 2 版、W3C 推奨 (2004 年 10 月 28 日)) とが含まれる。

【0069】

意味論を意識したスキーマ検証の下では、XML スキーマで規定されたノードと意味論的に同等の名前を有する特定のノードは、その特定のノードの真の名前がスキーマ規定のノードとは異なっていたとしても、同じものとして扱われる。意味論的な同等例は、フラグメント X A によって表わされる意味論的マッピングなどの意味論的マッピングに基づいて決定される。

10

【0070】

たとえば、スキーマは、要素 <address> を子 <purchase order> として含むよう XML 文書を規定し得る。意味論を意識した処理がなければ、文書 D 1 は XML スキーマに適合するものとは見なされない。というのも、当該文書は、異なっているが、意味論的に同等の要素 <Addr> を含んでいるからである。意味論を意識した処理の際に、文書 D 1 は適合しているものと見なされる。というのも、意味論的マッピングに基づいて、要素 <Addr> が <Address> と同一のものとして扱われるからである。

20

【0071】

ハードウェア概要

図 3 は、この発明の実施例が実現され得るコンピュータシステム 300 を図示するブロック図である。コンピュータシステム 300 は、情報を通信するためのバス 302 または他の通信メカニズムと、バス 302 に結合され情報を処理するためのプロセッサ 304 とを含む。コンピュータシステム 300 はまた、プロセッサ 304 によって実行されるべき情報および命令を記憶するための、バス 302 に結合されるランダムアクセスメモリ (RAM) または他の動的記憶装置などのメインメモリ 306 を含む。メインメモリ 306 はまた、プロセッサ 304 によって実行されるべき命令の実行中に、一時変数または他の中間情報を記憶するために用いられ得る。コンピュータシステム 300 はさらに、バス 302 に結合される、プロセッサ 304 のための静的情報および命令を記憶するための読出専用メモリ (ROM) 308 または他の静的記憶装置を含む。情報および命令を記憶するための、磁気ディスクまたは光学ディスクなどの記憶装置 310 が設けられ、バス 302 に結合される。

30

【0072】

コンピュータシステム 300 は、コンピュータユーザに情報を表示するための、陰極線管 (CRT) などのディスプレイ 312 にバス 302 を介して結合され得る。プロセッサ 304 に情報および指令選択を通信するための、英数字および他のキーを含む入力装置 314 がバス 302 に結合される。別の種類のユーザ入力装置は、プロセッサ 304 に方向情報および指令選択を通信するため、ならびにディスプレイ 312 上のカーソルの動きを制御するためのマウス、トラックボール、またはカーソル方向キーなどのカーソル制御部 316 である。この入力装置は、典型的には、第 1 の軸 (たとえば x) および第 2 の軸 (たとえば y) の 2 本の軸における 2 自由度を有しており、この装置が平面内の位置を指定することを可能にする。

40

【0073】

この発明は、この明細書中において記載される技術を実現するためのコンピュータシステム 300 の使用に関する。この発明の一実施例に従うと、それらの技術は、プロセッサ 304 がメインメモリ 306 に含まれる 1 つ以上の命令の 1 つ以上のシーケンスを実行することに応答して、コンピュータシステム 300 によって実行される。このような命令は、記憶装置 310 などの別の機械読取可能媒体からメインメモリ 306 に読込まれ得る。メインメモリ 306 に含まれる命令のシーケンスを実行して、プロセッサ 304 に、この

50

明細書中において記載されるプロセスステップを実行させる。代替的な実施例においては、ソフトウェア命令の代わりに、またはソフトウェア命令と組合せて、配線による回路構成を用いてこの発明を実現してもよい。したがって、この発明の実施例は、ハードウェア回路構成およびソフトウェアのいずれの特定の組合せにも限定されない。

【0074】

この明細書中で用いられるような「機械読取可能媒体」という語は、特定の態様で機械に操作させるデータを与えることに関与するいずれかの媒体を指す。コンピュータシステム300を用いて実現される実施例においては、たとえばプロセッサ304に実行するよう命令を与える際に、さまざまな機械読取可能媒体が含まれる。このような媒体は、不揮発性媒体、揮発性媒体および送信媒体を含む多くの形態を取り得るがこれらに限定されない。不揮発性媒体は、たとえば、記憶装置310などの光学または磁気ディスクを含む。揮発性媒体は、メインメモリ306などの動的メモリを含む。送信媒体は、バス302を備える配線を含む、同軸ケーブル、銅線および光ファイバを含む。送信媒体はまた、電波および赤外線データ通信中に発生されるような音波または光波の形態を取り得る。すべてのこのような媒体は、媒体によって搬送される命令が、機械に命令を読み込む物理メカニズムによって検出され得るように具体的でなければならない。

10

【0075】

機械読取可能媒体の一般的な形態は、たとえば、フロッピー（登録商標）ディスク、フレキシブルディスク、ハードディスク、磁気テープ、もしくは他のいずれかの磁気媒体、CD-ROM、他のいずれかの光学媒体、パンチカード、紙テープ、孔のパターンを有する他のいずれかの物理媒体、RAM、PROM、EPROM、FLASH-EPROM、他のいずれかのメモリチップもしくはカートリッジ、後で説明されるような搬送波、またはそこからコンピュータが読込可能な他のいずれかの媒体をも含む。

20

【0076】

実行用の1つ以上の命令の1つ以上のシーケンスをプロセッサ304に搬送する際、さまざまな形態の機械読取可能媒体が含まれ得る。たとえば、命令は初め、リモートコンピュータの磁気ディスクで搬送され得る。リモートコンピュータは、その動的メモリに命令をロードし、モデムを用いて電話線で命令を送ることができる。コンピュータシステム300にローカルなモデムが電話線上のデータを受信し、赤外線送信機を用いてそのデータを赤外線信号に変換することができる。赤外線検出器は赤外線信号で搬送されるデータを受信することができ、適切な回路構成が当該データをバス302に乘せることができる。バス302は、データをメインメモリ306に搬送し、そこからプロセッサ304が命令を取出して実行する。メインメモリ306が受けた命令は、プロセッサ304による実行の前または後に、記憶装置310に任意で記憶され得る。

30

【0077】

コンピュータシステム300はまた、バス302に結合される通信インターフェイス318を含む。通信インターフェイス318は、ローカルネットワーク322に接続されるネットワークリンク320に双方向データ通信結合を与える。たとえば、通信インターフェイス318は、対応する種類の電話線にデータ通信接続を与えるための統合サービスデジタル網（ISDN）カードまたはモデムであり得る。別の例として、通信インターフェイス318は、互換性のあるLANにデータ通信接続を与えるためのローカルエリアネットワーク（LAN）カードであり得る。無線リンクも実現され得る。このようないずれの実現例においても、通信インターフェイス318は、さまざまな種類の情報を表わすデジタルデータストリームを搬送する電気、電磁または光学信号を送受信する。

40

【0078】

ネットワークリンク320は典型的に、1つ以上のネットワークを通じて他のデータ装置にデータ通信を与える。たとえば、ネットワークリンク320は、ローカルネットワーク322を通じてホストコンピュータ324に、またはインターネットサービスプロバイダ（ISP）326によって操作されるデータ機器に接続を提供し得る。ISP326は次いで、現在一般的に「インターネット」328と称される広域パケットデータ通信網を

50

通じてデータ通信サービスを提供する。ローカルネットワーク 322 およびインターネット 328 はともに、デジタルデータストリームを搬送する電気、電磁または光学信号を用いる。さまざまなネットワークを通る信号、ならびにコンピュータシステム 300 との間でデジタルデータを搬送し合う通信インターフェイス 318 を通るネットワークリンク 320 上の信号は、情報を移送する搬送波の例示的な形態である。

【0079】

コンピュータシステム 300 は、ネットワーク、ネットワークリンク 320 および通信インターフェイス 318 を通じて、メッセージを送り、プログラムコードを含むデータを受信することができる。インターネットの例では、サーバ 330 が、インターネット 328、ISP 326、ローカルネットワーク 322 および通信インターフェイス 318 を通じてアプリケーションプログラムについての要求コードを送信し得る。

10

【0080】

受信されたコードは、その受信時にプロセッサ 304 によって実行され得、および／または、後で実行するために記憶装置 310 もしくは他の不揮発性記憶装置に記憶され得る。この態様で、コンピュータシステム 300 は、アプリケーションコードを搬送波の形態で得ることができる。

【0081】

上述の明細書においては、この発明の実施例が、実現例ごとに異なり得る多数の特定のな詳細を参照して説明された。したがって、発明であるもの、および出願人によって発明であると意図されるものを唯一および独占的に示すものは、いずれの後の訂正をも含む、この出願から生じる一組の請求項であり、そのような請求項が生じる特定のな形態におけるものである。そのような請求項に含まれる用語についてこの明細書中において明示的に記載されるいずれの定義も、請求項において用いられるようなそのような用語の意味を決定するものとする。したがって、請求項において明示的に列挙されていない限定、要素、性質、特徴、利点または属性は、そのような請求項の範囲を如何なる意味でも限定するものではない。したがって、明細書および図面は、限定的ではなく例示的に認識されるべきである。

20

【図 1】

XML インデックス 101

経路テーブル 102

RID	ロケータ	値	経路 ID	意味論的 経路 ID
R1	•	•	•	•
R1	•	•	•	•
R1	17	1600 WILLOW	12/23	12/25
R1	•	•	•	•
R2	•	•	•	•
R2	13	1600 WILLOW	12/24	12/25
R2	•	•	•	•
R2	•	•	•	•

文書 D1 文書 D2

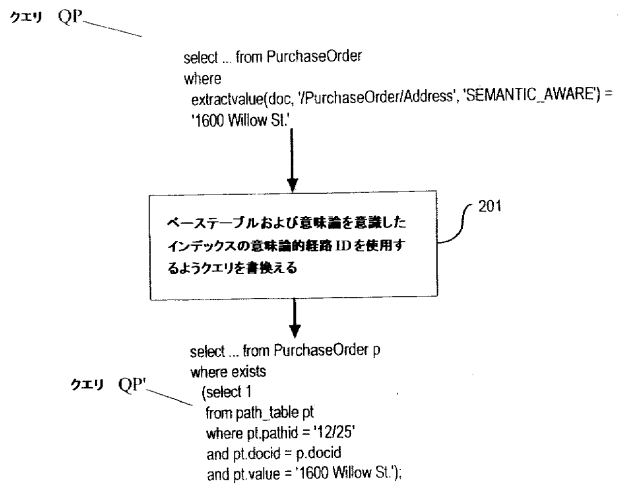
エントリー 103 エントリー 104

PURCHASE ORDER/ADDR PURCHASE ORDER/ADDRESS

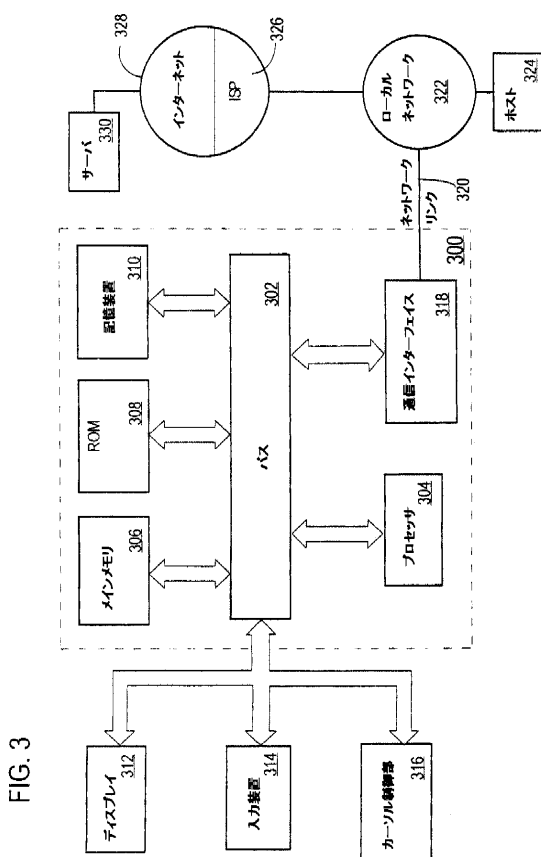
FIG. 1

【図 2】

FIG. 2



【図 3】



【手続補正書】

【提出日】平成20年9月16日(2008.9.16)

【手続補正 1】

【補正対象書類名】特許請求の範囲

【補正対象項目名】全文

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項 1】

コンピュータによって実現されるステップを含む方法であって、前記コンピュータによって実現されるステップは、

正規のタグ名を、第1のノードの第1の名前と、前記第1の名前とは異なる第2のノードの第2の名前とにマップする意味論的マッピングを格納するステップを含み、XML文書の集合は前記第1のノードおよび前記第2のノードを含み、前記コンピュータによって実現されるステップはさらに、

前記意味論的マッピングに基づき、前記第1の名前および第2の名前を同一の名前として扱うことによりタグ名オペレーションを実行するステップを含み、

前記タグ名オペレーションは、前記XML文書の集合に対して発行されたクエリのコンピュータ処理であり、前記クエリはクエリ言語に準拠する、方法。

【請求項 2】

コンピュータによって実現されるステップを含む方法であって、前記コンピュータによって実現されるステップは、

正規のタグ名を、第1のノードの第1の名前と、前記第1の名前とは異なる第2のノードの第2の名前とにマップする意味論的マッピングを格納するステップを含み、XML文書の集合は前記第1のノードおよび前記第2のノードを含み、前記コンピュータによって実現されるステップはさらに、

前記意味論的マッピングに基づき、前記第1の名前および第2の名前を同一の名前として扱うことによりタグ名オペレーションを実行するステップを含み、

前記タグ名オペレーションはXML文書のスキーマ検証である、方法。

【請求項 3】

前記タグ名オペレーションは、前記XML文書の集合へのアクセスを管理するレポジトリによって実行される、請求項 1 または 2 に記載の方法。

【請求項 4】

前記コンピュータによって実現されるステップは、前記意味論的マッピングを表わすデータを登録する要求を受信するステップと、

前記要求に応答して、前記データを前記意味論的マッピングとして登録するステップとをさらに含む、請求項 3 に記載の方法。

【請求項 5】

コンピュータによって実現されるステップを含む方法であって、前記コンピュータによって実現されるステップは、

XML文書の集合における複数のノードの各ノードについて、

意味論的マッピングに基づいて意味論的経路IDを生成するステップを含み、

前記複数のノードは第1のノードおよび第2のノードを含み、

第1の名前は前記第1のノードまたは前記第1のノードの先祖ノードに関連付けられており、

第2の名前は前記第2のノードまたは前記第2のノードの先祖に関連付けられており、

前記意味論的マッピングは、正規のタグ名を前記第1の名前および前記第2の名前にマップし、

前記第1のノードおよび前記第2のノードのために生成された意味論的経路IDは同一である、方法。

【請求項 6】

前記各ノードのための意味論的経路 ID は、前記各ノードの経路におけるノードの各々の名前についてのコードを含み、

前記第 1 の名前についてのコードおよび前記第 2 の名前についてのコードは同じである、請求項 5 に記載の方法。

【請求項 7】

前記コンピュータによって実現されるステップはさらに、

前記複数のノードのために生成される意味論的経路 ID によって前記複数のノードに索引付けするインデックスを作成するステップを含む、請求項 5 に記載の方法。

【請求項 8】

前記 XML 文書の集合はデータベースサーバによって管理され、前記コンピュータによって実現されるステップはさらに、

XML 文書の集合に対して発行されたクエリを受信するステップを含み、前記クエリは経路を指定し、前記コンピュータによって実現されるステップはさらに、

前記経路に基づいて、データベースサーバが前記インデックスにアクセスするよう前記クエリを書換えるステップを含む、請求項 5 に記載の方法。

【請求項 9】

命令の 1 つ以上のシーケンスを搬送するコンピュータ読取可能媒体であって、前記命令の 1 つ以上のシーケンスは、1 つ以上のプロセッサによって実行されると、請求項 1 から 8 のいずれかに記載の方法を前記 1 つ以上のプロセッサに実行させる、コンピュータ読取可能媒体。

【請求項 10】

XML 文書の集合における複数のノードのインデックスを記憶するコンピュータ読取可能媒体であって、

前記複数のノードの各ノードは、前記各ノードを含む或る経路に関連付けられており、

前記インデックスの各エントリは、前記複数のノードの特定のノードに対応し、前記特定のノードの或る経路を表わす意味論的経路 ID に前記ノードに関連付け、

前記複数のノードは第 1 のノードおよび第 2 のノードを含み、

第 1 の名前は、前記第 1 のノードまたは前記第 1 のノードの先祖ノードに関連付けられており、

第 2 の名前は、前記第 2 のノードまたは前記第 2 のノードの先祖ノードに関連付けられており、

前記第 1 のノードおよび前記第 2 のノードについてのそれぞれの意味論的経路 ID は同一である、コンピュータ読取可能媒体。

【請求項 11】

前記各ノードについての意味論的経路 ID は、前記各ノードの経路におけるノードの各々の名前についてのコードを含み、

前記第 1 の名前および前記第 2 の名前についてのそれぞれのコードは同じである、請求項 10 に記載のコンピュータ読取可能媒体。

【国際調査報告】

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2007/073087

A. CLASSIFICATION OF SUBJECT MATTER
INV. G06F17/30

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)
G06F

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

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

EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EUGENE AGICHTEIN Y, C. T. HOWARD HO , VANJA JOSIFOVSKI , JOERG GERHARDT: "Extracting Relations from XML Documents" SPRINGER LECTURE NOTES IN COMPUTER SCIENCE (LNCS) & XSDM 2003, [Online] vol. 2814, 2003, XP002459928 Chicago, USA Retrieved from the Internet: URL: http://www.mathcs.emory.edu/{eugene/papers/xsdm2003.pdf} [retrieved on 2007-11-26] the whole document ----- -/--	1-20

☒ 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 document 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.
- *G* document member of the same patent family

Date of the actual completion of the international search

26 November 2007

Date of mailing of the international search report

07/12/2007

Name and mailing address of the ISA/

European Patent Office, P.B. 6818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Warry, Lawrence

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2007/073087

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	YUH-PYNG SHIEH ET AL: "A reasoning framework for heterogeneous XML" INFORMATION TECHNOLOGY: RESEARCH AND EDUCATION, 2003. PROCEEDINGS. ITRE2003. INTERNATIONAL CONFERENCE ON AUG. 11-13, 2003, PISCATAWAY, NJ, USA, IEEE, 11 August 2003 (2003-08-11), pages 404-411, XP010685481 ISBN: 0-7803-7724-9	1, 6, 19
A	the whole document	2-5, 7-18, 20
A	HUIYONG XIAO, ISABEL F. CRUZ, FEIHONG HSU: "Semantic Mappings for the Integration of XML and RDF Sources" DEPARTMENT OF COMPUTER SCIENCE, UNIVERSITY OF ILLINOIS, [Online] 2004, XP002459929 Chicago, USA Retrieved from the Internet: URL: http://www.cs.uiuc.edu/~advise/publications/dataint/11web04.pdf [retrieved on 2007-11-26] the whole document	1-20
A	FUJIMOTO K ET AL: "A Mapping Scheme of XML Documents into Relational Databases using Schema-based Path Identifiers" WEB INFORMATION RETRIEVAL AND INTEGRATION, 2005. WIRI '05. PROCEEDINGS. INTERNATIONAL WORKSHOP ON CHALLENGES IN TOKYO, JAPAN 08-09 APRIL 2005, PISCATAWAY, NJ, USA, IEEE, 8 April 2005 (2005-04-08), pages 82-90, XP010860562 ISBN: 0-7695-2414-1 the whole document	1-20

フロントページの続き

(51)Int.Cl. F I テーマコード(参考)
G 0 6 F 12/00 5 1 3 D

(81)指定国 AP(BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), EP(AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, SI, SK, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, 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, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW

(74)代理人 100098316

弁理士 野田 久登

(74)代理人 100109162

弁理士 酒井 将行

(74)代理人 100111246

弁理士 荒川 伸夫

(72)発明者 マーシー, ラビ

アメリカ合衆国、 9 4 5 5 5 カリフォルニア州、フレモント、ジェイミー・サークル、 3 3 2 2 7

F ターム(参考) 5B075 NK02 NK43 PP26

5B082 GA08

5B109 NH08 NH20 VA02