

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第3815567号

(P3815567)

(45) 発行日 平成18年8月30日(2006.8.30)

(24) 登録日 平成18年6月16日(2006.6.16)

(51) Int. Cl.

G06F 17/21 (2006.01)

F I

G06F 17/21 501T

G06F 17/21 596Z

請求項の数 4 (全 26 頁)

(21) 出願番号 特願2003-94281 (P2003-94281)
 (22) 出願日 平成15年3月31日 (2003.3.31)
 (65) 公開番号 特開2004-302799 (P2004-302799A)
 (43) 公開日 平成16年10月28日 (2004.10.28)
 審査請求日 平成16年4月28日 (2004.4.28)

(73) 特許権者 000004237
 日本電気株式会社
 東京都港区芝五丁目7番1号
 (74) 代理人 100102864
 弁理士 工藤 実
 (72) 発明者 藤田 悟
 東京都港区芝五丁目7番1号 日本電気株
 式会社内

審査官 和田 財太

(56) 参考文献 特開2002-244894 (JP, A)
)
 国際公開第02/029602 (WO,
 A1)

最終頁に続く

(54) 【発明の名称】 コンピュータシステム、コンピュータプログラム、コンピュータ間の通信方法、構造化文書の符号化方法、符号化された構造化文書の復号方法

(57) 【特許請求の範囲】

【請求項1】

第一のコンピュータと、

第二のコンピュータ

とを備え、

前記第一のコンピュータは、

第一の構造化文書の内部表現データである第一の構造化文書内部表現を生成する第一のアプリケーションプログラムと、

前記第一のアプリケーションプログラムから前記第一の構造化文書内部表現の入力を受けて、前記第一の構造化文書内部表現から前記構造化文書のデータ構造の検証を支援するための情報を取り出し、前記検証を支援するための情報を含めて前記構造化文書を符号化する構造化文書符号化手段と、

前記構造化文書符号化手段の作成する符号化文書を前記第二のコンピュータに対して送信するデータ送信手段

とを有し、

前記第二のコンピュータは、

前記第一のコンピュータから送信された前記符号化文書を受信するデータ受信手段と、

前記符号化文書と文書構造定義ファイルを用いて、前記符号化文書の構造の正しさを検証し、その検証に成功した場合に、前記符号化文書を前記第二のコンピュータ上の第二の構造化文書内部表現として復号化する構造化文書復号化手段

10

20

前記第二の構造化文書内部表現を入力として処理を実行する第二のアプリケーションプログラムを有し、

前記第一のコンピュータの前記構造化文書符号化手段は、

前記第一の構造化文書内部表現を入力として、前記構造化文書に現れる情報タグの中身を解析して、要素名と属性名を抽出して、抽出した要素名が要素名リストの中に既に登録済かどうかを検査し、登録されていない場合にだけ、前記要素名リストに追加登録を行い、抽出した属性名が属性名リストの中に既に登録済かどうかを検査し、登録されていない場合にだけ、前記属性名リストに追加登録を行うことにより、前記要素名リスト及び前記属性名リストを作成する要素名/属性名抽出手段と、

前記構造化文書のタグに対する内容を表すテキストを抽出して、抽出した要素テキストが、要素テキストリストの中に、同じ要素名に対するテキストとして、既に登録済かどうかを検査し、登録されていない場合にだけ、前記要素テキストリストに追加登録を行い、抽出した属性テキストが、属性テキストリストの中に、同じ属性名に対するテキストとして、既に登録済かどうかを検査し、登録されていない場合にだけ、前記属性テキストリストに追加登録を行うことにより、前記要素テキストリストと前記属性テキストリストを作成するテキスト抽出手段と、

前記構造化文書の中の階層的な情報タグ構造及び前記情報タグに保有する属性情報との部分構造を抽出して、抽出した部分構造が、部分構造リストの中に、既に登録済かどうかを検査し、登録されていない場合にだけ、前記部分構造リストに追加登録を行うことにより、前記部分構造リストを作成する部分構造抽出手段と、

前記要素名リストと、前記属性名リストと、前記要素テキストリストと、前記属性テキストリストと、前記部分構造リストを用いて、前記構造化文書の全体構造データを抽出する全体文書構造抽出手段と、

前記要素名リストと、前記属性名リストと、前記要素テキストリストと、前記属性テキストリストと、前記部分構造リストと、前記全体構造データを、前記符号化文書として出力する符号化文書出力手段

とを備え、

前記第二のコンピュータの前記構造化文書復号化手段は、

前記符号化文書を入力として受理して、前記要素名リスト、前記属性名リスト、前記要素テキストリスト、前記属性テキストリスト、前記部分構造リスト、前記全体構造データを抽出する符号化文書入力手段と、

前記文書構造定義ファイルに記載された文書型の定義を利用して、前記要素名リストに対応する前記要素テキストリストの中身の文書型を比較して、前記属性名リストに対応する前記属性テキストリストの中身の文書型を比較して、それぞれのテキストの文書型が前記文書構造定義ファイルに記載された文書型の定義と一致することを検証するテキスト型検証手段と、

前記文書構造定義ファイルに記載された文書型の定義を利用して、前記部分構造リストに登録された部分構造と前記要素名リストと前記属性名リストを比較して、部分文書型が前記文書構造定義ファイルに記載された文書形の定義と一致することを検証する部分構造検証手段と、

前記テキスト型検証手段と前記部分構造検証手段とで検証された文書について、前記要素名リストと、前記属性名リストと、前記要素テキストリストと、前記属性テキストリストと、前記部分構造リストと、前記全体構造データを用いて、前記第二の構造化文書内部表現を再生する構造化文書内部表現出力手段

とを備える

構造化文書の通信システム。

【請求項 2】

請求項 1 記載の構造化文書の通信システムであって、

前記第一のコンピュータの前記検証支援情報抽出手段が、前記第一の構造化文書内部表現に加えて、文書構造定義ファイルを入力として受理し、

10

20

30

40

50

前記要素名/属性名抽出手段と、前記テキスト抽出手段と、前記部分構造抽出手段と、前記全体構造データ抽出手段が、それぞれ前記文書構造定義ファイルの情報に基づいて、前記要素テキストや前記属性テキストの文書型や部分構造型が、前記文書構造定義ファイルに定義される型と一致することを検証する

構造化文書の通信システム。

【請求項3】

請求項1記載の構造化文書の通信システムであって、

前記検証支援情報抽出手段の前記要素名/属性名抽出手段が、抽出した要素名が前記要素名リストの中に既に登録済かどうかを検査し、登録されていない場合にだけ、前記要素名リストに追加登録を行う

構造化文書の通信システム。

【請求項4】

要素名抽出手段と、属性名抽出手段と、テキスト抽出手段と、部分構造抽出手段と、全体文書構造抽出手段と、符号化文書出力手段とを備える第一のコンピュータから、符号化文書入力手段と、テキスト型検証手段と、部分構造検証手段と、構造化文書内部表現出力手段とを備える第二のコンピュータに構造化文書を通信する通信方法であって、

前記第一のコンピュータの前記要素名抽出手段が、構造化文書の内部表現を入力として、前記構造化文書のタグ情報内の要素名を抽出し、前記要素名が要素名リストの中に登録済かどうかを検査し、登録されていない場合は登録を行うステップと、

前記第一のコンピュータの前記属性名抽出手段が、前記タグ情報が属性を持つかどうかを検査し、属性を持つ場合には、属性が複数あればそれを一つずつ順に選択して属性名を抽出し、前記属性名が属性名リストの中に登録済かどうかを検査し、登録されていない場合は登録を行うステップと、

前記第一のコンピュータの前記テキスト抽出手段が、属性値として属性テキストを抽出し、前記属性テキストが属性テキストリストの中に同じ属性名に対して登録済かどうかを検査し、登録されていない場合は登録を行うステップと、

前記第一のコンピュータの前記テキスト抽出手段が、前記タグの内容としてテキストを持つかどうかを検査し、テキストを持つ場合はテキストを抽出し、前記テキストが要素テキストリストの中に同じ要素名に対して登録済かどうかを検査し、登録されていない場合は登録を行うステップと、

前記第一のコンピュータの前記部分構造抽出手段が、対象としている要素を中心に一階層だけの部分構造を抽出し、前記部分構造が部分構造リストの中に登録済かどうかを検査し、登録されていない場合は登録を行うステップと、

前記第一のコンピュータが、前記対象としている要素が下位構造として要素を持つか検査し、下位に要素を持つ場合には、下位の要素が複数の場合は一つずつ順に選択して、前記下位の要素に対して、再帰的に同じ手順で構造化文書を解析するステップと、

前記第一のコンピュータの全体文書構造抽出手段が、前記要素名リストと、前記属性名リストと、前記要素テキストリストと、前記属性テキストリストと、前記部分構造リストを用いて、前記構造化文書の全体文書構造を表す全体構造データを抽出するステップと、

前記第一のコンピュータの符号化文書出力手段が、前記要素名リストと、前記属性名リストと、前記要素テキストリストと、前記属性テキストリストと、前記部分構造リストと、前記全体構造データを、前記符号化文書として出力するステップと、

前記第二のコンピュータの前記符号化文書入力手段が前記符号化文書を受信し、前記符号化文書から前記要素名リストと、前記属性名リストと、前記要素テキストリストと、前記属性テキストリストと、前記部分構造リストと、前記全体文書データを抽出するステップと、

前記第二のコンピュータの前記テキスト型検証手段が、前記要素テキストリストの各要素テキストが前記構造化文書定義ファイルに示される要素の型と適合しているかを検証し、前記属性テキストリストの各属性テキストが前記構造化文書定義ファイルに示される属性の型と適合しているかを検証するステップと、

10

20

30

40

50

前記第二のコンピュータの前記部分構造検証手段が、前記部分構造リストの各部分構造が前記構造化文書定義ファイルに示される文書構造に適合しているかを検証するステップと、

前記第二のコンピュータの前記構造化文書内部表現出力手段が、前記テキスト型検証手段及び前記部分構造検証手段による検証のいずれにも適合した文書について、構造化文書の内部表現を再構成し、前記構造化文書の内部表現をアプリケーションプログラムに受け渡すステップ

とを具備する

構造化文書の通信方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、コンピュータシステムに関し、特に、構造化文書（構造化テキスト）を相互に通信するコンピュータで形成されたコンピュータシステムに関する。

【0002】

【従来の技術】

構造化文書（構造化テキスト）は、文書のデータ構造をあらわす記述（タグ）を文書自体に埋め込んだ文書である。構造化文書は、データ構造について高い柔軟性と拡張性を有しており、このような特長から、構造化文書は、異なるコンピュータ、及び異なるアプリケーション間のデータ交換に広く使用される。XML、及びSGMLは、典型的な構造化文書の規格である。

【0003】

構造化文書の構造を規定するために、文書型定義が使用される。例えば、XML文書の文書型定義としてDTD（Document Type Definition）及びXMLスキーマが使用される。文書型定義は、構造化文書自体に埋め込まれることも、構造化文書とは別に用意されることもある。構造化文書とは別途に用意されて文書型定義を記述したファイルは、文書構造定義ファイルと呼ばれる。

【0004】

構造化文書は、文書型定義に即した形式で正しく記述される必要がある。構造化文書が文書型定義に従わずに記述されていることは、その構造化文書を取り扱うコンピュータが該構造化文書の内容を正しく把握することを妨げる。このため、構造化文書は、そのデータ構造の妥当性が検証されることが望ましい。特許文献1は、構造化文書のデータ構造の検証について開示している。

【0005】

構造化文書の検証における一つの問題は、構造化文書を検証するために必要な演算量が大いことである。構造化文書は、本質的に冗長性が大きく、その大きさが大きい傾向にある。この構造化文書の冗長性は、検証に必要な演算量を増加させ、このため、検証に必要なコストを増大させる。

【0006】

このような問題は、特に、構造化文書を一のコンピュータから他のコンピュータに伝送する場合に顕著になる。構造化文書を伝送する場合、該構造化文書は、送信側コンピュータと受信側コンピュータの両方で検証され得る。送信側コンピュータにおいて構造化文書の検証が行われても、受信した構造化文書のデータの完全性を保証するためには、受信側コンピュータにおいても構造化文書の検証が行われる必要がある。送信側コンピュータと受信側コンピュータの両方で構造化文書の検証が行われることは、検証に必要な演算量が大いという問題を更に深刻にする。

【0007】

【特許文献1】

特開2001-75958号公報

【0008】

10

20

30

40

50

【発明が解決しようとする課題】

本発明の目的は、構造化文書を検証するために必要な演算量を抑制し、これにより検証に必要なコストを減少することにある。

本発明の他の目的は、構造化文書を送信側コンピュータから受信側コンピュータに伝送する場合に、受信側コンピュータが構造化文書を検証するために必要な演算量を抑制し、これにより検証に必要なコストを減少することにある。

【0009】**【課題を解決するための手段】**

以下に、[発明の実施の形態]で使用される番号・符号を用いて、課題を解決するための手段を説明する。これらの番号・符号は、[特許請求の範囲]の記載と[発明の実施の形態]の記載との対応関係を明らかにするために付加されている。但し、付加された番号・符号は、[特許請求の範囲]に記載されている発明の技術的範囲の解釈に用いてはならない。

10

【0010】

本発明によるコンピュータシステムは、構造化文書の内部表現(13)を符号化して符号化文書(4)を作成する符号化手段(12)を含む送信側コンピュータ(1)と、ネットワーク(3)を介して符号化文書(4)を取得し、符号化文書(4)を復号して内部表現(23)を再構成する復号手段(22)を含む受信側コンピュータ(2)とを備えている。符号化手段(12)は、該構造化文書に含まれる記述を重複することなく列挙することによって内部表現(13)から検証支援情報(4a~4e)を生成する検証支援情報抽出手段(15)を含み、且つ、検証支援情報(4a~4e)を含むように符号化文書(4)を生成する。復号手段(22)は、検証支援情報(4a~4e)に基づいて、再構成される内部表現(23)のデータ構造が、所定の文書型宣言に適合しているかを検証する検証手段(25)を含む。当該コンピュータシステムでは、記述の重複が排除された検証支援情報(4a~4e)に基づいて内部表現(23)のデータ構造の妥当性が検証される。このような構成は、検証手順の重複を排除し、従って、構造化文書を検証するために必要な演算量を抑制することができる。

20

【0011】

検証支援情報(4a~4e)は、内部表現(13)を構成する要素のうち、テキストデータを要素内容として含むテキスト含有要素の要素テキストが、重複しないように列挙された要素テキストリスト(4c)を含み、文書型宣言は、テキスト含有要素の型を定義した要素型宣言を含み、検証手段(25)は、該要素テキストが、該型に適合するか否かを検証することが好適である。これにより、同一の要素テキストが内部表現(13)に複数含まれていても、該要素テキストの検証の重複を避けることができる。

30

【0012】

検証支援情報(4a~4e)は、内部表現(13)を構成する要素のうち、属性を有する属性所有要素と、該属性の属性値である属性テキストとが、重複しないように列挙された属性テキストリスト(4d)を含み、文書型宣言は、属性の属性型を定義した属性型宣言を含み、検証手段(25)は、属性テキストが、属性型に適合するか否かを検証することが好適である。これにより、同一の属性テキストが内部表現(13)に複数含まれていても、該属性テキストの検証の重複を避けることができる。

40

【0013】

検証支援情報(4a~4e)は、内部表現(13)に含まれる部分構造が、重複しないように列挙された部分構造リスト(4e)を含み、検証手段(25)は、該部分構造が、文書型宣言に定義された構文に適合するか否かを検証することが好適である。これにより、同一の部分構造が内部表現(13)に複数含まれていても、該部分構造の検証の重複を避けることができる。

【0014】

検証支援情報(4a~4e)が、該要素の要素名を列挙した要素名リスト(4a)と、該属性の属性名を列挙した属性名リスト(4b)と、要素テキストリスト(4c)と、属

50

性テキストリスト(4d)とを含む場合、該部分構造のそれぞれには、構造IDが与えられ、該要素テキストと該属性テキストとのそれぞれにテキストIDが与えられ、該部分構造の部分構造表現は、該要素の要素IDと、該要素が有する属性の属性IDと、該要素の子要素の要素IDと、該親要素が、要素テキストを含むことを示す記号とによって記述され、符号化文書(4)は、該構造IDと該テキストIDとによって内部表現(13)の全体の構造を表現する全体構造データ(4f)を含み、復号手段(22)は、要素名リスト(4a)と属性名リスト(4b)と要素テキストリスト(4c)と属性テキストリスト(4d)と全体構造データ(4f)とに基づいて内部表現(23)を再構成することが好適である。このような構成は、内部表現(13、23)の全体構成を、小さなデータ量で表現することを可能にする。

10

【0015】

本発明によるコンピュータは、構造化文書の内部表現を符号化して符号化文書(4)を作成する符号化手段(12)を備えている。符号化手段(12)は、該構造化文書に含まれる記述を重複することなく列挙することによって内部表現から検証支援情報(4a~4e)を生成する検証支援情報抽出手段(25)を含み、且つ、検証支援情報(4a~4e)を含むように符号化文書(4)を生成する。このようなコンピュータは、上記のコンピュータシステムを好適に実現する。

【0016】

本発明によるコンピュータは、ネットワーク(3)を介して符号化文書(4)を取得し、符号化文書(4)を復号して構造化文書の内部表現を再構成する復号手段(22)を備えている。符号化文書(4)は、該構造化文書に含まれる記述が重複することなく列挙された検証支援情報(4a~4e)を含んでいる。

20

復号手段(22)は、検証支援情報(4a~4e)に基づいて、再構成される内部表現(23)のデータ構造が、所定の文書型宣言に適合しているかを検証する検証手段(25)を含む。このようなコンピュータは、上記のコンピュータシステムを好適に実現する。

【0017】

本発明による構造化文書通信方法は、

(a) 構造化文書の内部表現(13)を符号化して符号化文書(4)を作成するステップと、

(b) ネットワーク(3)を介して符号化文書(4)を取得し、符号化文書(4)を復号して内部表現(23)を再構成するステップとを備えている。

30

該(a)ステップは、

(a1) 前記構造化文書に含まれる記述を重複することなく列挙することによって内部表現(13)から検証支援情報(4a~4e)を生成するステップと、

(a2) 前記検証支援情報(4a~4e)を含むように前記符号化文書(4)を生成するステップ

とを含む。該(b)ステップは、

(b1) 検証支援情報(4a~4e)に基づいて、再構成され内部表現(23)のデータ構造が、所定の文書型宣言に適合しているかを検証するステップ

40

を含む。このような構造化文書通信方法は、記述の重複が排除された検証支援情報(4a~4e)に基づいて内部表現(23)のデータ構造の妥当性を検証する。従って、当該構造化文書通信方法は、検証手順の重複を排除し、従って、構造化文書を検証するために必要な演算量を抑制することができる。

【0018】

本発明によるコンピュータプログラムは、

(a1) 構造化文書に含まれる記述を重複することなく列挙することによって該構造化文書の内部表現(13)から検証支援情報(4a~4e)を生成するステップと、

(a2) 検証支援情報(4a~4e)を含むように符号化文書(4)を生成するステップとをコンピュータ(1)に実行させる。当該コンピュータプログラムは、上述のコンピ

50

ータシステムの実現に好適に使用される。

【0019】

本発明によるコンピュータプログラムは、

(b) ネットワーク(3)を介して符号化文書(4)を取得し、符号化文書(4)を復号して構造化文書の内部表現(23)を再構成するステップ

をコンピュータに実行させるコンピュータプログラムである。符号化文書(4)は、該構造化文書に含まれる記述が重複することなく列挙された検証支援情報(4a~4e)を含む。該(b)ステップは、

(b1) 検証支援情報(4a~4e)に基づいて、再構成される内部表現(23)のデータ構造が、所定の文書型宣言に適合しているかを検証するステップ

を含む。当該コンピュータプログラムは、上述のコンピュータシステムの実現に好適に使用される。

【0020】

【発明の実施の形態】

以下、添付図面を参照しながら、本発明の実施の一形態を説明する。本発明によるコンピュータシステムの実施の一形態では、図1に示されているように、送信側コンピュータ1と受信側コンピュータ2とが設けられる。

【0021】

送信側コンピュータ1には、構造化文書、(例えば、XML文書、及びSGML文書)に対応したアプリケーションプログラム11と構造化文書符号化モジュール12とがインストールされている。アプリケーションプログラム11は、受信側コンピュータ2に送信しようとする構造化文書の内部表現である構造化文書内部表現13を生成する。図10に示されているように、アプリケーションプログラム11は、典型的には、外部から構造化文書13'を受け取り、受け取った構造化文書13'の構造化文書内部表現13を生成するプログラムである。

【0022】

図1に示されているように、送信側コンピュータ1には、送信側コンピュータ2が取り扱う構造化文書の構造を示す文書型宣言を記述した文書構造定義ファイル14が用意されることがある。文書構造定義ファイル14に記述される文書型宣言には、構造化文書内部表現13が表現する構造化文書の文書構造の規則が記述される。より具体的には、文書型宣言は、構造化文書に含まれる要素の型を定義する要素型宣言、要素に与えられる属性の型を定義する属性型宣言を含んでいる。

【0023】

文書構造定義ファイル14が用意される場合、構造化文書内部表現13は、文書構造定義ファイル14に記述されている文書型宣言に従って記述される必要がある。当該コンピュータシステムがXMLを使用する場合、文書型宣言としては、例えば、DTD、XMLスキーマが使用される。

【0024】

構造化文書符号化モジュール12は、アプリケーションプログラム11から渡された構造化文書内部表現13を符号化して符号化文書4を生成するコンピュータプログラムである。生成された符号化文書4は、ネットワーク3を介して受信側コンピュータ2に送信される。

【0025】

受信側コンピュータ2には、アプリケーションプログラム21と構造化文書復号化モジュール22とがインストールされている。構造化文書復号化モジュール22は、受信側コンピュータ2が受信した符号化文書4を復号化して、構造化文書内部表現23を生成するために使用される。アプリケーションプログラム21は、復号化文章復号化モジュール21から構造化文書内部表現23を受け取り、所望の処理を行う。図10に示されているように、アプリケーションプログラム21は、典型的には、構造化文書内部表現23から、それによって表現されている構造化文書24'を再生するプログラムである。送信側コンピ

10

20

30

40

50

ュータ 1 から受信側コンピュータ 2 への通信の間にエラーが発生しなければ、構造化文書内部表現 2 3 は、送信側コンピュータ 1 のアプリケーションプログラム 1 1 が生成する構造化文書内部表現 1 3 に一致する。

【 0 0 2 6 】

図 1 に示されているように、受信側コンピュータ 2 は、受信側コンピュータ 2 が取り扱う構造化文書が従うべき文書型宣言を記述した文書構造定義ファイル 2 4 を保存している。送信側コンピュータ 1 に用意され得る文書構造定義ファイル 1 4 と同様に、文書構造定義ファイル 2 4 に記述されている文書型宣言は、構造化文書内部表現 2 3 に表現された構造化文書の文書構造の規則を記述している。より具体的には、文書構造定義ファイル 2 4 に記述されている文書型宣言は、構造化文書に含まれる要素の型を定義する要素型宣言、要素に与えられる属性の型を定義する属性型宣言を含んでいる。送信側コンピュータ 1 に文書構造定義ファイル 1 4 が用意される場合、受信側コンピュータ 2 に用意される文書構造定義ファイル 2 4 と送信側コンピュータ 1 に用意される文書構造定義ファイル 1 4 とは、一致される必要がある。

10

【 0 0 2 7 】

受信側コンピュータ 2 が構造化文書内部表現 2 3 の内容を正しく把握するためには、構造化文書内部表現 2 3 が文書構造定義ファイル 2 4 に含まれる文書型宣言に従って記述されていること（このような構造化文書内部表現 2 3 の性質は、「構造化文書内部表現 2 3 の妥当性」と呼ばれる。）が保証されなくてはならない。構造化文書内部表現 2 3 が XML 文書の内部表現である場合、これは、構造化文書内部表現 2 3 によって表現される XML 文書が「妥当な XML 文書 (valid XML document) 」であることを保証することに相当する。

20

【 0 0 2 8 】

しかし、既述のように、構造化文書は冗長であるから、構造化文書内部表現 2 3 を直接に検証して構造化文書内部表現 2 3 の妥当性を保証することは、それに必要な演算量が大きく好ましくない。

【 0 0 2 9 】

構造化文書内部表現 2 3 の妥当性を保証するために、符号化文書 4 が下記に述べられるような特別な形式で生成され、その符号化文書 4 に基づいて構造化文書内部表現 2 3 の妥当性が検証される。必要な演算量を抑えて構造化文書内部表現 2 3 の妥当性の保証を容易化するために、符号化文書 4 は、下記の 2 つの特性を備えている。第 1 に、符号化文書 4 は、検証支援情報を含んで構成されている。検証支援情報とは、受信側コンピュータ 2 において構造化文書内部表現 2 3 の妥当性の保証に使用される情報である。第 2 に、符号化文書 4 は、構造化文書内部表現 1 3 の冗長性を排除しながら、構造化文書内部表現 1 3 と同等の内容を記述している。既述のように、構造化文書の冗長性は、その検証に必要な演算量を増加させる。冗長性が少ない符号化文書 4 に基づいて構造化文書内部表現 2 3 の妥当性を保証することにより、妥当性の保証に必要な演算量が抑制される。

30

【 0 0 3 0 】

図 2 は、符号化文書 4 の具体的な構造を示している。符号化文書 4 は、要素名リスト 4 a と、属性名リスト 4 b と、要素テキストリスト 4 c と、属性テキストリスト 4 d と、部分構造リスト 4 e と、全体文書構造データ 4 f とを含んでいる。

40

【 0 0 3 1 】

要素名リスト 4 a は、送信側コンピュータ 1 のアプリケーションプログラム 1 1 によって提供される構造化文書内部表現 1 3 に含まれる要素の要素名を列挙したリストである。図 3 は、要素名リスト 4 a の内容の例を示している。要素名リスト 4 a は、構造化文書内部表現 1 3 に含まれる要素のそれぞれについて、要素に固有に与えられた要素 ID と、その要素の要素名とを記述している。図 3 を参照すると、例えば、要素 ID " E 0 1 " が与えられた要素の要素名は、" 要素名 A " である。

【 0 0 3 2 】

要素名リスト 4 a において、同一の要素名は、重複して列挙されない。ある要素名を有す

50

る要素が構造化文書内部表現 1 3 に複数回現れても、要素名リスト 4 a には、その要素名は一度しか現れない。

【 0 0 3 3 】

属性名リスト 4 b は、構造化文書内部表現 1 3 中で使用されている属性の属性名を列挙したリストである。図 4 は、属性名リスト 4 b の内容の例を示している。属性名リスト 4 a は、構造化文書内部表現 1 3 中で使用されている属性のそれぞれに固有に与えられた属性 ID と、その属性の属性名とを記述している。図 4 を参照すると、例えば、属性 ID " A 0 1 " が与えられている属性の属性名は、" 属性名 A " である。

【 0 0 3 4 】

属性名リスト 4 b において、同一の属性名は、重複して列挙されない。ある属性名を有する属性が構造化文書内部表現 1 3 に複数回現れても、属性名リスト 4 b には、その属性名は一度しか現れない。

10

【 0 0 3 5 】

要素テキストリスト 4 c は、構造化文書内部表現 1 3 に含まれる要素の要素テキストを列挙したリストである。要素テキストとは、テキストデータである要素内容である。図 5 は、要素テキストリスト 4 c の内容の例を示している。要素テキストリスト 4 c は、構造化文書内部表現 1 3 に含まれる全要素のうち、要素テキストを含む要素の要素 ID と、その要素に含まれる要素テキストと、その要素テキストに固有に与えられたテキスト ID とを記述している。構造化文書では、ある要素の要素内容は、子要素のみで構成されることが許されているから、全ての要素が要素テキストリスト 4 c に記載されているとは限らないことに留意されたい。図 5 を参照すると、例えば、要素 ID " E 0 1 " が与えられている要素の要素内容は、テキスト ID " T 0 1 " が与えられている要素テキスト " テキスト A " と、テキスト ID " T 0 2 " が与えられている要素テキスト " テキスト B " と、テキスト ID " T 0 3 " が与えられている要素テキスト " テキスト C " とを含んでいる。

20

【 0 0 3 6 】

要素テキストリスト 4 c において、同一の要素テキストは、同一の要素名に対して、重複して列挙されない。ある要素テキストが構造化文書内部表現 1 3 に複数回現れても、要素テキストリスト 4 c には、その要素テキストは一度しか現れない。

【 0 0 3 7 】

属性テキストリスト 4 d は、構造化文書内部表現 1 3 中で使用されている属性の属性テキストを列挙したリストである。ある属性の属性テキストとは、その属性の属性値を示すテキストデータである。図 6 は、属性テキストリスト 4 d の内容の例を示している。属性テキストリスト 4 d は、構造化文書内部表現 1 3 中で使用されている属性の属性 ID と、その属性の属性テキストと、その属性テキストに固有に与えられたテキスト ID とを記述している。ある属性テキストのテキスト ID は、他の属性テキストのテキスト ID と異なるとともに、上述の要素テキストのテキスト ID と異なるように定められる。

30

【 0 0 3 8 】

図 6 を参照すると、例えば属性 ID " A 0 1 " が与えられている属性の属性テキスト (属性値) は、テキスト ID " T 0 1 " が与えられている属性テキスト " テキスト A " である。

40

【 0 0 3 9 】

属性テキストリスト 4 d において、同一の属性テキストは、同一の属性名に対して、重複して列挙されない。ある属性テキストが構造化文書内部表現 1 3 に複数回現れても、属性テキストリスト 4 c には、その属性テキストは一度しか現れない。

【 0 0 4 0 】

部分構造リスト 4 e は、構造化文書内部表現 1 3 に現れる部分構造を示す部分構造表現を列挙したリストである。部分構造は、一の要素 (親要素) と、その要素の子要素とから構成される。部分構造には、親要素の子要素の子要素は、含まれないことに留意されたい。部分構造表現は、このような部分構造を表現する記載である。ある要素 (親要素) の部分構造表現は、

50

- (1) 親要素の要素 I D
- (2) 親要素に規定されている属性の属性 I D、
- (2) 親要素に含まれる子要素の要素 I D
- (3) 親要素が要素テキストを含む場合、該親要素が要素テキストを含むことを示す記号を含む。

【 0 0 4 1 】

図 7 は、部分構造リスト 4 e の内容の例を示す。部分構造リスト 4 e は、部分構造のそれぞれに固有に与えられた構造 I D とその部分構造の部分構造表現とを記述している。部分構造表現の先頭には、その部分構造を構成する親要素の要素 I D が記述される。図 7 を参照すると、例えば構造 I D " S 0 1 " が与えられている部分構造は、要素 I D " E 0 1 " を有する親要素を含む。その親要素には、属性 I D " A 0 1 " を有する属性が規定され、更に、その親要素は、属性 I D " E 0 2 " を有する子要素を 2 つ含んでいる。また、構造 I D " S 0 4 " が与えられている部分構造は、要素 I D " E 0 4 " を有する親要素を含む。この要素は、ある要素テキストを含む。図 7 において記号 " T " は、要素 I D " E 0 4 " を有する要素が、要素テキストを含むことを意味している。構造 I D " S 0 4 " が与えられている部分構造表現には、親要素以外の要素 I D は含まれておらず、これは、該親要素は子要素を含まないことを示している。

10

【 0 0 4 2 】

部分構造リスト 4 e において、同一の部分構造は、重複して列挙されない。ある部分構造が構造化文書内部表現 1 3 に複数回現れても、部分構造リスト 4 e には、その部分構造は一度しか現れない。

20

【 0 0 4 3 】

全体文書構造データ 4 f は、構造化文書内部表現 1 3 の全体の構造を示すデータであり、構造 I D とテキスト I D とによって記述される。全体文書構造データ 4 f には、構造 I D に続いて、該構造 I D で特定される部分構造の親要素の要素テキスト、又は、その親要素に与えられている属性の属性テキスト(属性値)のテキスト I D が記述される。例えば、構造 I D " S 0 4 " によって特定される部分構造の部分構造表現が、" E 0 4 T " であり、要素 I D " E 0 4 " によって特定される要素名が " D " であり、テキスト I D " T 0 4 " によって特定される要素テキストが、" テキスト D " であるとする。全体文書構造データ 4 f が " S 0 4 T 0 4 " という記載を含む場合、その記載は、要素名が " D " である要素の要素テキストが " テキスト D " であることが表している。構造 I D とテキスト I D とによって記述される全体文書構造データ 4 f は、より小さなデータ量で、構造化文書内部表現 1 3 の全体の構造を記述することができる。

30

【 0 0 4 4 】

上述の要素名リスト 4 a と、属性名リスト 4 b と、要素テキストリスト 4 c と、属性テキストリスト 4 d と、部分構造リスト 4 e と、全体文書構造データ 4 f とは、符号化データ 4 を復号化して構造化文書内部表現 1 3 と同一の構造化文書内部表現 2 3 を復元するために必要な情報を全て含んでいる。

【 0 0 4 5 】

要素名リスト 4 a と、属性名リスト 4 b と、要素テキストリスト 4 c と、属性テキストリスト 4 d と、部分構造リスト 4 e とは、構造化文書内部表現 2 3 を復元するのに必要な情報であるとともに、受信側コンピュータ 2 における構造化文書内部表現 2 3 の妥当性の保証に使用される検証支援情報としても機能する。要素名リスト 4 a と要素テキストリスト 4 c とは、構造化文書内部表現 2 3 に含まれている要素の要素テキストが、その要素について文書構造定義ファイル 2 4 に宣言されている要素型に適合することを検証するために使用される。属性名リスト 4 b と属性テキストリスト 4 d とは、構造化文書内部表現 2 3 で使用されている属性の属性テキストが、その属性について文書構造定義ファイル 2 4 に宣言されている属性型に適合することを検証するために使用される。部分構造リスト 4 e は、構造化文書内部表現 2 3 で使用されている部分構造が、文書構造定義ファイル 2 4 に定義されている文書構造に適合することを検証するために使用される。これらを検証する

40

50

ことにより、符号化文書 4 を復号化して構造化文書内部表現 2 3 を復元する前に、構造化文書内部表現 2 3 の妥当性を保証することが可能になる。

【 0 0 4 6 】

上述のように、要素テキストリスト 4 c、属性テキストリスト 4 d、及び部分構造リスト 4 e からは、重複が排除されている。これは、構造化文書内部表現 2 3 の妥当性を保証するために必要な演算の演算量を有効に減少する。要素テキストリスト 4 c において、同一の要素テキストが重複して列挙されないことは、該要素テキストの妥当性の検証を重複して行う必要をなくす。更に、属性テキストリスト 4 d において、同一の属性テキストが重複して列挙されないことは、該属性テキストの妥当性の検証を重複して行う必要をなくす。加えて、部分構造リスト 4 e において、同一の部分構造が重複して列挙されないことは、該部分構造の妥当性の検証を重複して行う必要をなくす。このように、要素テキストリスト 4 c、属性テキストリスト 4 d、及び部分構造リスト 4 e における重複の排除により、構造化文書内部表現 2 3 の妥当性を保証するために必要な演算の演算量が抑制される。

10

【 0 0 4 7 】

図 8 は、符号化文書 4 を生成する構造化文書符号化モジュール 1 2 の構成を示す。構造化文書符号化モジュール 1 2 は、検証支援情報抽出モジュール 1 5 と、全体文書構造抽出モジュール 1 6 と、符号化文書出力モジュール 1 7 とを備えている。

【 0 0 4 8 】

検証支援情報抽出モジュール 1 5 は、要素名 / 属性名抽出モジュール 1 5 a と、テキスト抽出モジュール 1 5 b と、部分構造抽出モジュール 1 5 c とを含む。要素名 / 属性名抽出モジュール 1 5 a は、構造化文書内部表現 1 3 から、要素名と属性名とを抽出して、要素名リスト 4 a と属性名リスト 4 b とを作成する。テキスト抽出モジュール 1 5 b は、構造化文書内部表現 1 3 から要素テキスト及び属性テキストとを抽出して、要素テキストリスト 4 c と属性テキストリスト 4 d とを作成する。要素テキストリスト 4 c と属性テキストリスト 4 d との作成には、要素名 / 属性名抽出モジュール 1 5 a によって生成された要素名リスト 4 a と属性名リスト 4 b が使用される。部分構造抽出モジュール 1 5 c は、構造化文書内部表現 1 3 から部分構造を抽出して部分構造リスト 4 e を作成する。部分構造抽出モジュール 1 5 c は、抽出した部分構造のそれぞれに構造 ID を与え、更に、部分構造の構造 ID とその部分構造表現とを列挙して部分構造リスト 4 e を作成する。

20

【 0 0 4 9 】

要素名 / 属性名抽出モジュール 1 5 a、テキスト抽出モジュール 1 5 b、及び部分構造抽出モジュール 1 5 c は、送信側コンピュータ 1 が文書構造定義ファイル 1 4 を有している場合に文書構造定義ファイル 1 4 を利用して要素名リスト 4 a、属性名リスト 4 b、要素テキストリスト 4 c、属性テキストリスト 4 d、及び部分構造リスト 4 e を効率的に作成するように構成されていることが好適である。

30

【 0 0 5 0 】

全体文書構造抽出モジュール 1 6 は、構造化文書内部表現 1 3 から全体構造データ 4 f を作成する。全体文書構造抽出モジュール 1 6 は、構造化文書内部表現 1 3 に基づいて構造化文書の全体構造を把握し、その全体構造を構造 ID とテキスト ID とで表現する全体構造データ 4 f を生成する。

40

【 0 0 5 1 】

符号化文書出力モジュール 1 7 は、要素名リスト 4 a と、属性名リスト 4 b と、要素テキストリスト 4 c と、属性テキストリスト 4 d と、部分構造リスト 4 e と、全体構造データ 4 f とを統合して、符号化文書 4 を生成する。

【 0 0 5 2 】

図 9 は、符号化文書 4 を復号して構造化文書内部表現 2 3 を再生する構造化文書復号化モジュール 2 2 の構成を示す。構造化文書復号化モジュール 2 2 は、検証モジュール 2 5 と、符号化文書入力モジュール 2 6 と、構造化文書内部表現出力モジュール 2 7 とを備えている。

【 0 0 5 3 】

50

符号化文書入力モジュール 2 6 は、符号化文書 4 を展開して、要素名リスト 4 a と、属性名リスト 4 b と、要素テキストリスト 4 c と、属性テキストリスト 4 d と、部分構造リスト 4 e と、全体構造データ 4 f とを生成する。

【 0 0 5 4 】

検証モジュール 2 5 は、テキスト型検証モジュール 2 5 a と部分構造検証モジュール 2 5 b とを含む。テキスト型検証モジュール 2 5 a は、要素テキストリスト 4 c に列挙された要素テキストと、属性テキストリスト 4 d に列挙された属性テキストとが、文書構造定義ファイル 2 4 に記述されている文書型宣言に適合しているかを検証する。部分構造検証モジュール 2 5 b は、部分構造リスト 4 e に列挙された部分構造が、文書構造定義ファイル 2 4 に規定されている型に適合しているかを検証する。

10

【 0 0 5 5 】

構造化文書内部表現出力モジュール 2 7 は、要素名リスト 4 a、属性名リスト 4 b、要素テキストリスト 4 c、属性テキストリスト 4 d、部分構造リスト 4 e 及び全体構造データ 4 f から構造化文書内部表現 2 3 を再生する。

【 0 0 5 6 】

続いて、本実施の形態におけるコンピュータシステムの動作が説明される。図 1 1 及び図 1 2 は、構造化文書符号化モジュール 1 2 が構造化文書内部表現 1 3 から符号化文書 4 を生成するために行う処理を示すフローチャートである。構造化文書符号化モジュール 1 2 に構造化文書内部表現 1 3 が与えられ、構造化文書符号化モジュール 1 2 の実行が開始されると（ステップ S 0 1）、構造化文書符号化モジュール 1 2 は、何も列挙されていない要素名リスト 4 a、属性名リスト 4 b、要素テキストリスト 4 c、属性テキストリスト 4 d、及び部分構造リスト 4 e を生成する。更に、構造化文書符号化モジュール 1 2 は、構造化文書内部表現 1 3 に記述されている要素のうち、要素名/属性名抽出モジュール 1 5 a が処理すべき要素を選択する。選択された要素は、以下において、処理対象要素と呼ばれる。構造化文書符号化モジュール 1 2 の実行が開始された直後においては、ルート要素（即ち、最上位の要素）が処理対象要素として選択される。

20

【 0 0 5 7 】

要素名/属性名抽出モジュール 1 5 a は、処理対象要素の要素名を抽出する（ステップ S 0 2）。更に、要素名/属性名抽出モジュール 1 5 a は、処理対象要素の要素名が要素名リスト 4 a に登録されているかを調べる（ステップ S 0 3）。登録されていない場合、要素名/属性名抽出モジュール 1 5 a は、処理対象要素の要素名を、処理対象要素の要素 ID とともに要素名リスト 4 a に登録する（ステップ S 0 4）。ステップ S 0 3、S 0 4 の処理により、要素名は、重複することなく要素名リスト 4 a に登録される。

30

【 0 0 5 8 】

更に、要素名/属性名抽出モジュール 1 5 a は、処理対象要素が属性を持つかを調べる（ステップ S 0 5）。処理対象要素が属性を持たない場合には、構造化文書符号化モジュール 1 2 の処理は、図 1 2 に示されているステップ S 1 4 にジャンプする。

【 0 0 5 9 】

処理対象要素が属性を持つ場合には、要素名/属性名抽出モジュール 1 5 a は、処理対象要素が有する属性の一つを選択する（ステップ S 0 6）。処理対象要素が属性を一つしか有しない場合には、その属性が自動的に選択される。

40

【 0 0 6 0 】

続いて、要素名/属性名抽出モジュール 1 5 a は、選択された属性の属性名を、構造化文書内部表現 1 3 から抽出する（ステップ S 0 7）。更に、要素名/属性名抽出モジュール 1 5 a は、選択された属性の属性名が属性名リスト 4 b に登録されているかを調べる。（ステップ S 0 8）。登録されていない場合、要素名/属性名抽出モジュール 1 5 a は、選択された属性の属性名を、その属性の属性 ID とともに属性名リスト 4 b に登録する（ステップ S 0 9）ステップ S 0 8、S 0 9 の処理により、属性名は、重複することなく属性名リスト 4 b に登録される。

【 0 0 6 1 】

50

更に、要素名/属性名抽出モジュール15 aは、処理対象要素に規定されている全ての属性についてステップS 07からステップS 12の処理、即ち、属性登録処理が行われたかを調べる。属性登録処理が全ての属性について行われていない場合、属性登録処理が行われていない属性について属性登録処理が行われる。

【0062】

処理対象要素に規定されている全ての属性について属性登録処理が行われると、図12に示されているように、テキスト抽出モジュール15 bは、処理対象要素が要素テキストを有するかを調べる(ステップS 14)。有しない場合、構造化文書符号化モジュール12の処理は、ステップS 18にジャンプする。処理対象要素が要素テキストを有する場合、テキスト抽出モジュール15 bは、その要素テキストが要素テキストリスト4 cに登録されているかを調べる(ステップS 11)。登録されていない場合、テキスト抽出モジュール15 bは、その要素テキストにテキストIDを与え、更に、その要素テキストを、処理対象要素の要素IDとテキストIDとともに要素テキストリスト4 cに登録する。ステップS 10及びステップS 11の処理により、要素テキストは、重複することなく要素テキストリスト4 cに登録される。

10

【0063】

続いて、部分構造抽出モジュール15 cは、処理対象要素を親要素として有する部分構造を構造化文書内部表現13から抽出し、その部分構造の部分構造表現を取得する(ステップS 18)。既述の通り、部分構造表現は、要素IDと属性IDで表現されている。部分構造リスト4 eの作成に使用される要素IDは、要素名を検索キーとして要素名リスト4 aを検索することによって取得され、属性IDは、属性名を検索キーとして属性名リスト4 bを検索することによって取得される。

20

【0064】

更に部分構造抽出モジュール15 cは、抽出した部分構造が部分構造リスト4 eに登録されているかを調べる(ステップS 19)。登録されていない場合、部分構造抽出モジュール15 cは、該部分構造に構造IDを与え、更に、該部分構造の部分構造表現を、構造IDとともに部分構造リスト4 eに登録する。

【0065】

続いて、構造化文書符号化モジュール12は、処理対象要素が子要素を有するかを調べる(ステップS 21)。有する場合、構造化文書符号化モジュール12は、該処理対象要素の子要素のうちの一つを選択する(ステップS 22)。更に、構造化文書符号化モジュール12は、選択された子要素を処理対象要素としてステップS 02以下の手順を再帰的に実行する(ステップS 23)。これにより、選択された子要素の要素名と要素IDが要素リスト4 aに登録され、その子要素に与えられている属性の属性名と属性IDとが属性リスト4 bに登録され、その属性の属性テキスト(属性値)が属性テキストリスト4 dに登録され、その子要素の要素テキストが要素テキストリスト4 cに登録され、その子要素を親要素とする部分構造が部分構造リスト4 dに登録される。選択された子要素が、更に子要素を有する場合も同様の処理が行われる。

30

【0066】

全ての要素についてステップS 02からステップS 23の処理が行われると、全体文書構造抽出モジュール16は、構造化文書内部表現13を構造IDとテキストIDとで表現する全体構造データ4 fを生成する。

40

【0067】

更に、符号化文書出力モジュール17が、要素名リスト4 a、属性名リスト4 b、要素テキストリスト4 c、属性テキストリスト4 d、部分構造リスト4 e及び全体構造データ4 fを統合して符号化文書4を生成し、符号化文書4の生成処理が完了する(ステップS 26)。

【0068】

図13は、構造化文書復号化モジュール22が、符号化文書4から生成されるべき構造化文書内部表現23の妥当性を検証し、更に、符号化文書4から構造化文書内部表現23を

50

復元する処理を示すフローチャートである。まず、構造化文書復号化モジュール 2 2 が符号化文書 4 を受け取ると、構造化文書復号化モジュール 2 2 の符号化文書入力モジュール 2 6 は、符号化文書 4 から、要素名リスト 4 a、属性名リスト 4 b、要素テキストリスト 4 c、属性テキストリスト 4 d、部分構造リスト 4 e、及び全体構造データ 4 f を抽出する（ステップ S 3 2 ~ S 3 7）。

【 0 0 6 9 】

更に、構造化文書復号化モジュール 2 2 のテキスト型検証モジュール 2 5 a は、要素テキストリスト 4 c に列挙された要素テキストを検証する（ステップ S 3 8）。要素テキストの検証は、下記の手順で行われる。テキスト型検証モジュール 2 5 a は、要素テキストリスト 4 c に列挙された要素テキストを有する要素の要素名を、要素 ID を検索キーとして要素名リスト 4 a を検索して取得する。更にテキスト型検証モジュール 2 5 a は、取得した要素名を検索キーとして文書構造定義ファイル 2 4 を検索し、その要素名を有する要素の要素型宣言を取得する。更にテキスト型検証モジュール 2 5 a は、要素テキストが、要素型宣言に規定された型に適合しているかを判定する（ステップ S 3 9）。要素テキストが、要素型宣言に規定された型に適合していない場合、即ち、要素テキストに要素型違反がある場合、要素型違反の発生を通知する要素型違反通知を生成する。要素型違反通知は、受信側コンピュータ 2 の表示装置に表示され、更に、送信側コンピュータ 1 に送信されて送信側コンピュータ 1 の表示装置に表示される（ステップ S 4 4）。要素型違反が発生した場合、構造化文書内部表現 2 3 の復元は行われず、構造化文書復号化モジュール 2 2 の処理は終了する（ステップ S 4 7）。

10

20

【 0 0 7 0 】

要素型違反が発生していない場合、テキスト型検証モジュール 2 5 a は、更に、属性テキストリスト 4 d に列挙された属性テキストを検証する（ステップ S 4 0）。属性テキストの検証は下記の手順で行われる。テキスト型検証モジュール 2 5 a は、属性テキストリスト 4 d に列挙された属性テキストを属性値として有する属性の属性名を、属性 ID を検索キーとして属性名リスト 4 b を検索して取得する。更にテキスト型検証モジュール 2 5 a は、取得した属性名を検索キーとして文書構造定義ファイル 2 4 を検索し、その属性名を有する属性の属性型宣言を取得する。更にテキスト型検証モジュール 2 5 a は、属性テキストが、属性型宣言に規定された型に適合しているかを判定する（ステップ S 4 1）。属性テキストが、属性型宣言に規定された型に適合していない場合、即ち、属性テキストに属性型違反がある場合、属性型違反の発生を通知する属性型違反通知を生成する。属性型違反通知は、受信側コンピュータ 2 の表示装置に表示され、更に、送信側コンピュータ 1 に送信されて送信側コンピュータ 1 の表示装置に表示される（ステップ S 4 4）。属性型違反が発生した場合、構造化文書内部表現 2 3 の復元は行われず、構造化文書復号化モジュール 2 2 の処理は終了する（ステップ S 4 7）。

30

【 0 0 7 1 】

属性型違反が発生していない場合、部分構造検証モジュール 2 5 b は、部分構造リスト 4 e に列挙されている部分構造を検証する（ステップ S 4 2）。部分構造の検証は、下記の手順で行われる。部分構造検証モジュール 2 5 b は、部分構造リスト 4 e に列挙された部分構造の部分構造表現に現れている要素 ID を検索キーとして要素名リスト 4 a を検索する。部分構造検証モジュール 2 5 b は、要素名リスト 4 a を検索することにより、その部分構造に関係している要素の要素名を取得する。更に、その部分構造に関係している要素のうち、親要素の要素名を検索キーとして文書構造定義ファイル 2 4 を検索し、その部分構造が取るべき型を取得する。部分構造検証モジュール 2 5 b は、その部分構造が、文書構造定義ファイル 2 4 に規定された型に適合しているかを判定する（ステップ S 4 3）。部分構造が、文書構造定義ファイル 2 4 に規定されている型に適合しない場合、部分構造の型違反の発生を通知する部分構造型違反通知を生成する。部分構造型違反通知は、受信側コンピュータ 2 の表示装置に表示され、更に、送信側コンピュータ 1 に送信されて送信側コンピュータ 1 の表示装置に表示される（ステップ S 4 4）。部分構造型違反が発生した場合、構造化文書内部表現 2 3 の復元は行われず、構造化文書復号化モジュール 2 2 の

40

50

処理は終了する（ステップS 4 7）。

【0072】

要素型違反、属性型違反、及び部分構造型違反のいずれも発生しなかった場合、構造化文書内部表現出力モジュール27は、要素名リスト4a、属性名リスト4b、要素テキストリスト4c、属性テキストリスト4d、部分構造リスト4e及び全体構造データ4fから構造化文書内部表現23を再生する。構造化文書内部表現出力モジュール27は、要素名リスト4a、属性名リスト4b、要素テキストリスト4c、属性テキストリスト4d、及び部分構造リスト4eを参照して全体構造データ4fを解析する。構造化文書内部表現出力モジュール27は、全体構造データ4fを解析することによって構造化文書内部表現23に含まれるべき要素の要素名、その要素の要素内容（即ち、要素テキスト及び子要素の要素名）、その要素に定められた属性の属性名、属性の属性テキスト（属性値）を把握し、更に構造化文書内部表現23を生成する。

10

【0073】

更に構造化文書内部表現出力モジュール27は、型違反が発生しなかったことを通知する正常通知を生成する。正常通知は、受信側コンピュータ2の表示装置に表示され、更に、送信側コンピュータ1に送信されて送信側コンピュータ1の表示装置に表示される（ステップS 46）。構造化文書復号化モジュール22の処理は終了する（ステップS 47）。

【0074】

このように、構造化文書復号化モジュール22は、要素テキストリスト4c、属性テキストリスト4d、及び部分構造リスト4eに基づいて構造化文書内部表現23の妥当性を検証し、その妥当性が保証された後に構造化文書内部表現23を復元する。このような検証方法は、構造化文書内部表現23の妥当性を保証するために必要な演算の演算量を有効に減少する。上述のように、要素テキストリスト4c、属性テキストリスト4d、及び部分構造リスト4eからは、重複が排除されている。要素テキストリスト4cにおいて、同一の要素テキストが重複して列挙されないことは、該要素テキストの妥当性の検証を重複して行う必要をなくす。更に、属性テキストリスト4dにおいて、同一の属性テキストが重複して列挙されないことは、該属性テキストの妥当性の検証を重複して行う必要をなくす。加えて、部分構造リスト4eにおいて、同一の部分構造が重複して列挙されないことは、該部分構造の妥当性の検証を重複して行う必要をなくす。このように、要素テキストリスト4c、属性テキストリスト4d、及び部分構造リスト4eにおける重複の排除により、構造化文書内部表現23の妥当性を保証するために必要な演算の演算量が抑制される。

20

30

【0075】

以下では、構造化文書の処理の実施例が、説明される。

【0076】

【実施例】

本実施例では、図14に示されているXML文書について上記の実施の形態が適用される。該XML文書の構造がアプリケーションプログラム11によって解析され、構造化文書内部表現13が生成される。生成された構造化文書内部表現は、構造化文書符号化モジュール12に引き渡される

【0077】

ステップS 01において要素名リスト4a、属性名リスト4b、要素テキストリスト4c、属性テキストリスト4d、及び部分構造リスト4eが初期化され、更に、ルート要素"Addressbook"が処理対象要素として選択される。ステップS 02においてルート要素の要素名"Addressbook"が抽出される。この要素名は要素名リスト4aに登録されていない。従って、ステップS 04において、図15に示されているように、ルート要素の要素ID"E 01"と要素名"Addressbook"とが要素名リスト4aに登録される。

40

【0078】

ステップS 05において処理対象要素であるルート要素が属性を有するかが調べられる。図14に示されているように、ルート要素は、属性"owner"を有する。ステップS

50

06において属性"owner"が選択され、ステップS07においてその属性の属性名"owner"が抽出される。その属性名は、属性リスト4bに登録されていない(ステップS08)。そのため、ステップS09において、図16に示されているように、該属性の属性ID"A01"と属性名"owner"とが属性名リスト4bに登録される。更に、ステップS10において、該属性の属性テキスト"Suzuki"が抽出される。属性テキスト"Suzuki"は、属性テキストリスト4dに登録されていない(ステップS11)。そのため、ステップS12において、図18に示されているように、該属性の属性ID"A01"と、該属性テキストのテキストID"T10"と、属性テキスト"Suzuki"とが属性テキストリスト4dに登録される。ルート要素は、他の属性を有しないため(ステップS13)、ステップS06からステップS12までの処理は完了する。

10

【0079】

続いて、ステップS14において、処理対象要素であるルート要素が要素テキストを有するかが判断される。ルート要素は、その要素内容として要素テキストを有しないので、処理がステップS18にジャンプする。

【0080】

ステップS18では、ルート要素"Addressbook"を親要素として有する部分構造が抽出される。即ち、ルート要素に与えられた属性と、ルート要素の子要素とが抽出される。該部分構造の部分構造表現は親要素(ルート要素)の要素ID、親要素の属性の属性ID、及び子要素の要素IDで記述される。ルート要素の子要素"Person"の要素IDを"E02"とすると、ルート要素を親要素として有する部分構造の部分構造表現は、例えば、"E01 A01 E02 E02"と定められる。本実施例では、部分構造表現は、親要素の要素ID、親要素が有する属性の属性ID、及び子要素の要素IDが、この順に記載されるという規則に従って作成されている。親要素が属性を複数有する場合には、その属性IDが順次に並べられ、親要素が子要素を複数有する場合には、その要素IDが順次に並べられる。

20

【0081】

更に、ステップS19では抽出された部分構造が部分構造リスト4eに登録されているかが判断される。ルート要素"Addressbook"を親要素として有する部分構造は登録されていないので、図19に示されているように、ステップS20において該部分構造に構造ID"S01"が与えられ、更に、該部分構造の部分構造表現"E01 A01 E02 E02"が、構造ID"S01"とともに部分構造リスト4eに登録される。

30

【0082】

続いて、ステップS21においてルート要素"Addressbook"が子要素を有するかが判断される。図14に示されているように、ルート要素は、子要素"Person"を2つ有している。子要素"Person"が処理対象要素として順次に選択され、ステップS02からS22までの処理が、再帰的に実行される(ステップS22~24)。

【0083】

図14のXML文書を基にして生成される要素名リスト4aが図15に示され、属性名リスト4bが図16に示され、要素テキストリスト4cが図17に示され、属性テキストリスト4dが図18に示され、部分構造リスト4eが図19に示されている。

40

【0084】

更に、全体構造データ4fが生成される。図14に示されているXML文書の全体構成は、図19の部分構造リスト4eに示されている構造IDと、図17、図18の要素テキストリスト4c及び属性テキストリスト4dに列挙されているテキストIDとを用いて、例えば、

```

” S01 T10 S02 S03 S04 T01 S05 T02 S06
  S07 T04 S08 T05 S09 T06 S10 T08 S11
  S03 S04 T01 S05 T03 S12 S07 T04 S08
  T05 S09 T07 S13 T09”

```

と表現される。本実施例では、下記の規則に従って全体構造データ4fが生成されている。最上位の階層の要素を親要素とする部分構造の構造IDが先頭に置かれ、該親要素の属性の属性テキスト及び該親要素の要素テキストのテキストIDが次に並べられる。更に、該親要素の子要素を親要素とする部分構造の構造IDが並べられる。以下再帰的に、該子要素の属性の属性テキスト及び該子要素の要素テキストのテキストIDが並べられ、該子要素の子要素を親要素とする部分構造の構造IDが並べられる。

10

【0085】

上記の全体構造データ4fを理解しやすくするために括弧を挿入すると、全体構造データ4fは、“(S01 T10 (S02 (S03 (S04 T01) (S05 T02))) (S06 (S07 T04) (S08 T05) (S09 T06)) (S10 T08)) (S11 (S03 (S04 T01) (S05 T03))) (S12 (S07 T04) (S08 T05) (S09 T07) (S13 T09)))))”と表現される。

【0086】

先頭の”S01”は、構造ID”S01”を有する部分構造に含まれる親要素がルート要素であることを示している。構造ID”S01”が与えられている部分構造の部分構造表現”E01 A02 E02 E02”は、要素ID”E01”を有する要素(要素”Addressbook”、図15参照)がルート要素であることを示している。ルート要素の要素名は、図15に示されている要素名リスト4aから”Addressbook”であることがわかる。

20

【0087】

更に、構造ID”S01”を有する部分構造の部分構造表現”E01 A02 E02 E02”は、そのルート要素は、属性ID”A01”を有する属性(属性”Owner”、図16参照)を有し、且つ、要素ID”E02”を有する要素(要素”Person”、図15参照)を2つ有していることを示している。全体構造データ4fの”S01”に続く”T10”は、その属性の属性値がテキストID”T10”を有する属性テキスト”Suzuki”(図18参照)であることを示している。

30

【0088】

要素ID”E02”を有する2つの要素”Person”のうちの最初の一方には、表現(S02 ...)が割り当てられ、2番目の方には、表現(S11 ...)が割り当てられる。表現(S02 ...)に含まれる構造ID”S02”を有する部分構造の部分構造表現”E02 E03 E06 E10”から、要素”Person”の最初の一方は、要素ID”E03”、”E06”、及び”E10”をそれぞれに有する要素を子要素として有していることが分かる。要素ID”E03”を有する要素には、表現(S03 ...)が割り当てられ、要素ID”E06”を有する要素には、表現(S06 ...)が割り当てられ、要素ID”E10”を有する要素には、表現(S10 ...)が割り当てられる。以下、同様にして図14のXML文書の全体構造が表現される。但し、このような全体構造を表現するために、他の規則が使用され得ると理解されなくてはならない。

40

【0089】

図15乃至図19に示されている要素名リスト4a、属性名リスト4b、要素テキストリスト4c、属性テキストリスト4d、及び部分構造リスト4eと、上記の全体構造データ4fとが統合されて、符号化文書4が生成される。

【0090】

一方、図14のXML文書に対応する符号化文書4が構造化文書復号化モジュール22に

50

入力されると、ステップ S 3 2 ~ S 3 7 において、その符号化文書 4 から、図 1 5 乃至図 1 9 に示されている要素名リスト 4 a、属性名リスト 4 b、要素テキストリスト 4 c、属性テキストリスト 4 d、及び部分構造リスト 4 e と、上記の全体構造データ 4 f とが取り出される。

【 0 0 9 1 】

続いて、ステップ S 3 8 において、テキスト型検証モジュール 2 5 a により、要素テキストリスト 4 c に挙げられている要素テキストの型が検証される。要素テキストリスト 4 c に挙げられている 9 個の要素テキストのそれぞれについて、その要素テキストを要素内容とする要素の要素 ID が取り出され、取り出された要素 ID を検索キーとして要素名リスト 4 a が検索されて、該要素の要素名が取り出される。その要素名を用いて文書構造定義ファイル 2 4 が検索されて、該要素名を有する要素の要素型宣言が取り出される。更に、要素テキストが該要素型宣言に規定されている型に適合するかが検証される。

10

【 0 0 9 2 】

例えば、図 1 5 を参照して、要素 " L a s t N a m e " について考える。更に、文書構造定義ファイル 2 4 に規定された要素 " L a s t N a m e " の要素型宣言により、要素 " L a s t N a m e " の要素テキストが文字型であると規定されているとする。図 1 7 を参照すると、要素 " L a s t N a m e " は要素テキスト " Y a m a d a " を有しており、要素 " L a s t N a m e " の要素テキストは、正しい型を有していると判断される。

【 0 0 9 3 】

更に例えば、図 1 5 を参照して、要素 " T e l e p h o n e " について考える。文書構造定義ファイル 2 4 に規定された要素 " T e l e p h o n e " の要素型宣言により、要素 " T e l e p h o n e " の要素テキストが、3 桁の整数と、それに続くハイフン " - " と、それに続く 3 桁の整数と、それに続くハイフンと、それに続く 4 桁の整数というフォーマットを有すると規定されているとする。図 1 7 を参照すると、要素 " T e l e p h o n e " は要素テキスト " 0 4 5 - 1 2 3 - 4 5 6 7 " を有しており、要素 " T e l e p h o n e " の要素テキストは、正しい型を有していると判断される。

20

【 0 0 9 4 】

要素テキストが取り得る型は、この他にも、例えば、整数型、浮動小数点型、日付型であり得る。

【 0 0 9 5 】

同様に、全ての要素について要素型違反がないと判断されると、ステップ S 4 0 において、属性テキストリスト 4 d に挙げられている属性テキストの型が検証される。属性テキストリスト 4 d に挙げられている属性テキストのそれぞれについて、その属性テキストを属性値とする属性の属性 ID が取り出され、取り出された属性 ID を検索キーとして属性名リスト 4 b が検索されて、該属性の属性名が取り出される。その属性名を用いて文書構造定義ファイル 2 4 が検索されて、該属性名を有する属性の属性型宣言が取り出される。更に、属性テキストが該属性型宣言に規定されている型に適合するかが検証される。

30

【 0 0 9 6 】

例えば、図 1 6 を参照して、属性 " O w n e r " について考える。更に、文書構造定義ファイル 2 4 に規定された属性 " O w n e r " の属性型宣言により、属性 " O w n e r " の属性テキスト (属性値) が文字型であると規定されているとする。図 1 8 を参照すると、属性 " O w n e r " の要素テキストは、 " S u z u k i " であり、属性 " O w n e r " の要素テキストは、正しい型を有していると判断される。

40

【 0 0 9 7 】

更に、ステップ S 4 2 において部分構造リスト 4 e に列挙されている 1 3 個の部分構造が検証される。例えば、図 1 9 を参照して、構造 ID " S 0 1 " を有する部分構造の部分構造表現は、 " E 0 1 A 0 1 E 0 2 E 0 2 " である。要素名リスト 4 a と属性名リスト 4 b とを参照すると、この部分構造表現は、要素 ID " E 0 1 " を有する要素 " A d d r e s s b o o k " は属性 " O w n e r " を有し、且つ、その子要素が 2 つの要素 " P e r s o n " であることを示している。この部分構造が、文書構造定義ファイル 2 4 に規定

50

されている構文に適合するかが判定される。他の部分構造についても、同様の検証が行われる。

【0098】

以上の工程により、要素テキスト、属性テキスト、及び部分構造の検証が終了する。検証が終了した後、要素名リスト4 a、属性名リスト4 b、要素テキストリスト4 c、属性テキストリスト4 d、及び部分構造リスト4 eと、上記の全体構造データ4 fから、図14のXML文書の構造化文書内部表現23が復元される。

【0099】

このような検証方法は、同一の要素テキスト、属性テキスト、及び部分構造に対して重複して検証することを避けることができる。図14のXML文書を用いてその妥当性を検証しようとするれば、同一の検証対象に対して重複した検証が行われる。例えば、要素"Person"が要素"Name"と要素"Address"を有し、且つ、要素"Name"が要素"LastName"と要素"FirstName"を有するという記述は、図14のXML文書にそれぞれ2回現れる。従って、図14のXML文書を用いて検証が行われた場合、これらの記述は、2回検証されたはずである。更に、要素"LastName"の要素テキスト"Yamada"の型が正しいかという検証も2回繰り返されたはずである。要素"Country"の要素テキスト"Japan"、要素"Prefecture"の要素テキスト"Kanagawa"についても同様である。

10

【0100】

本実施の形態では、冗長性を排除して構造化文書を表示した符号化文書4が送信され、その符号化文書4に基づいて再生されるべき構造化文書内部表現23の妥当性が検証され、妥当性が保証された後に構造化文書内部表現23が再生される。これにより、同一の検証作業の繰り返しが排除され、検証作業のコストを低減することができる。

20

【0101】

図20に示されているように、本実施の形態のコンピュータシステムは、送信側コンピュータ1にデータ圧縮モジュール18を有し、受信側コンピュータ2にデータ解凍モジュール28を有していることが好適である。データ圧縮モジュール18は、符号化文書4を圧縮して圧縮符号化文書4'を生成する。データ解凍モジュール28は、圧縮符号化文書4'を受信して解凍し、符号化文書4を再生する。このような構成は、送信側コンピュータ1から受信側コンピュータ2に通信されるデータの量を抑制する。

30

【0102】

本実施の形態のコンピュータシステムにおいて、要素IDが一連番号を含んで構成される場合、要素名リスト4 aには明示的に要素IDが記述される必要はない。要素IDの代わりに、要素名リスト4 a内において要素名が記述される順番を使用することが可能である。例えば、要素名リスト4 aにおいて第1番目に記述される要素名は、要素ID"E01"を有すると規定され、その要素IDが要素名リスト4 aに記述されないことが可能である。通信コストに敏感な状況において本実施の形態のコンピュータシステムが使用される場合、要素名リスト4 aをより小さなデータ量で表現するメリットは大きい。

【0103】

同様に、属性IDが一連番号を含んで構成される場合、属性名リスト4 bには、明示的に要素IDが記述される必要はない。

40

【0104】

更に同様に、要素テキストリスト4 c及び属性テキストリスト4 dが一のリストに記述され、且つ、テキストIDが一連番号を含んで構成される場合、要素テキストリスト4 c及び属性テキストリスト4 dには、明示的にテキストIDが記述される必要はない。

【0105】

更に同様に、構造IDが一連番号を含んで構成される場合、部分構造リスト4 eには、構造IDが明示的に記述される必要がない。

【0106】**【発明の効果】**

50

本発明により、構造化文書を検証するために必要な演算量を抑制し、これにより、検証に必要なコストを減少することができる。

また、本発明により、構造化文書を送信側コンピュータから受信側コンピュータに伝送する場合に、受信側コンピュータが構造化文書を検証するために必要な演算量を抑制し、これにより、検証に必要なコストを減少することができる。

【図面の簡単な説明】

【図 1】図 1 は、本発明によるコンピュータシステムの実施の一形態を示すブロック図である。

【図 2】図 2 は、符号化文書 4 の内容を示す。

【図 3】図 3 は、要素名リスト 4 a の内容を示す。

【図 4】図 4 は、属性名リスト 4 b の内容を示す。

【図 5】図 5 は、要素テキストリスト 4 c の内容を示す。

【図 6】図 6 は、属性テキストリスト 4 d の内容を示す。

【図 7】図 7 は、部分構造リスト 4 e の内容を示す。

【図 8】図 8 は、構造化文書符号化モジュール 1 2 の構成を示すブロック図である。

【図 9】図 9 は、構造化文書復号化モジュール 2 2 の構成を示すブロック図である。

【図 1 0】図 1 0 は、本実施の形態のコンピュータシステムの変形例を示すブロック図である。

【図 1 1】図 1 1 は、構造化文書内部表現 1 3 から符号化文書 4 を生成する手順を示すフローチャートである。

【図 1 2】図 1 2 は、構造化文書内部表現 1 3 から符号化文書 4 を生成する手順を示すフローチャートである。

【図 1 3】図 1 3 は、符号化文書 4 に基づいて構造化文書内部表現 2 3 の妥当性を検証し、更に、構造化文書内部表現 2 3 を再構成する手順を示すフローチャートである。

【図 1 4】図 1 4 は、XML 文書の例を示す。

【図 1 5】図 1 5 は、図 1 4 の XML 文書の内部表現から生成される要素名リスト 4 a を示す。

【図 1 6】図 1 6 は、図 1 4 の XML 文書の内部表現から生成される属性名リスト 4 b を示す。

【図 1 7】図 1 7 は、図 1 4 の XML 文書の内部表現から生成される要素テキストリスト 4 c を示す。

【図 1 8】図 1 8 は、図 1 4 の XML 文書の内部表現から生成される属性テキストリスト 4 d を示す。

【図 1 9】図 1 9 は、図 1 4 の XML 文書の内部表現から生成される部分構成リスト 4 e を示す。

【図 2 0】図 2 0 は、本実施の形態のコンピュータシステムの他の変形例を示すブロック図である。

【符号の説明】

1 : 送信側コンピュータ

2 : 受信側コンピュータ

3 : ネットワーク

4 : 符号化文書

4 a : 要素名リスト

4 b : 属性名リスト

4 c : 要素テキストリスト

4 d : 属性テキストリスト

4 e : 部分構造リスト

4 f : 全体構造データ

1 1 : アプリケーションプログラム

1 2 : 構造化文書符号化モジュール

10

20

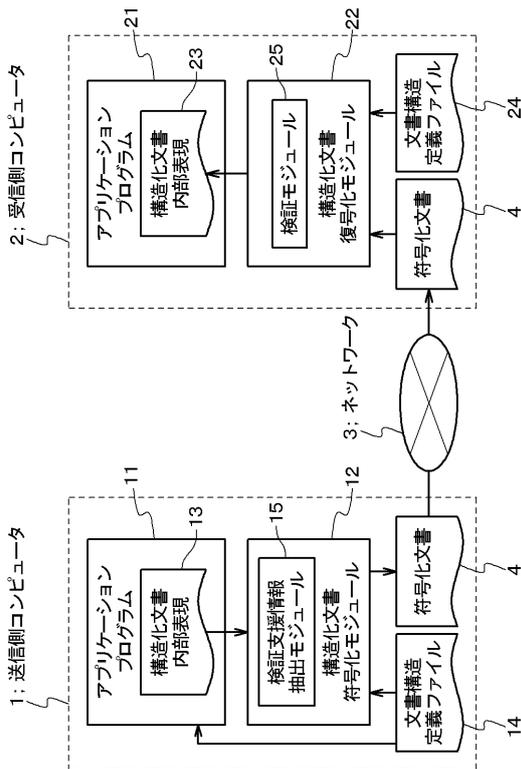
30

40

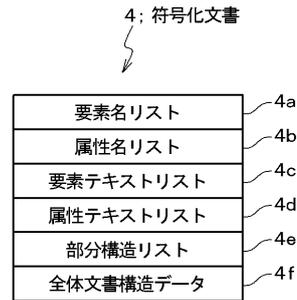
50

- 1 3 : 構造化文書内部表現
- 1 3 ' : 構造化文書
- 1 4 : 文書構造定義ファイル
- 1 5 : 検証支援情報抽出モジュール
- 1 5 a : 要素名 / 属性名抽出モジュール
- 1 5 b : テキスト抽出モジュール
- 1 5 c : 部分構造抽出モジュール
- 1 6 : 全体文書構造抽出モジュール
- 1 7 : 符号化文書出力モジュール
- 2 1 : アプリケーションプログラム
- 2 2 : 構造化文書復号化モジュール
- 2 3 : 構造化文書内部表現
- 2 4 : 文書構造定義ファイル
- 2 5 : 検証モジュール
- 2 5 a : テキスト型検証モジュール
- 2 5 b : 部分構造検証モジュール
- 2 6 : 符号化文書入力モジュール
- 2 7 : 構造化文書内部表現出力モジュール

【 図 1 】



【 図 2 】



【 図 3 】

4a; 要素名リスト

要素 I D	要素名
E01	要素名 A
E02	要素名 B
E03	要素名 C
E04	要素名 D
...	...

【 図 4 】

4b: 属性名リスト

属性ID	属性名
A01	属性名A
A02	属性名B
A03	属性名C
A04	属性名D
...	...

【 図 6 】

4d: 属性テキストリスト

属性ID	テキストID	属性テキスト
A01	T05	テキストA
A01	T06	テキストB
A01	T07	テキストC
A02	T08	テキストD
...

【 図 5 】

4c: 要素テキストリスト

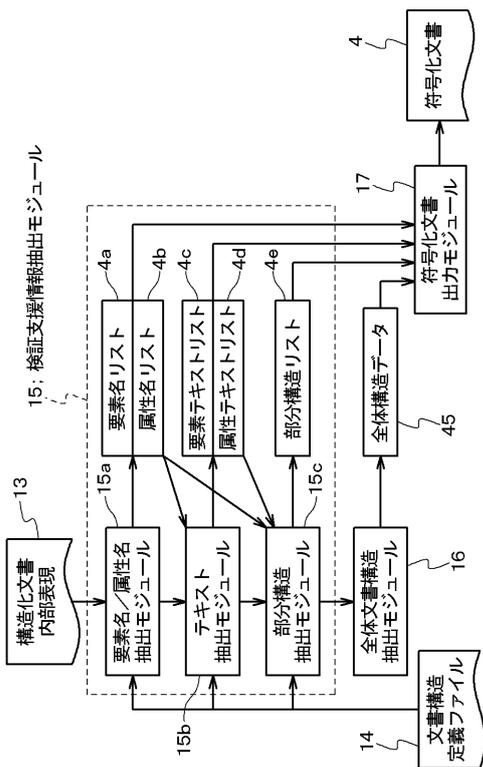
要素ID	テキストID	要素テキスト
E01	T01	テキストA
E01	T02	テキストB
E01	T03	テキストC
E02	T04	テキストD
...

【 図 7 】

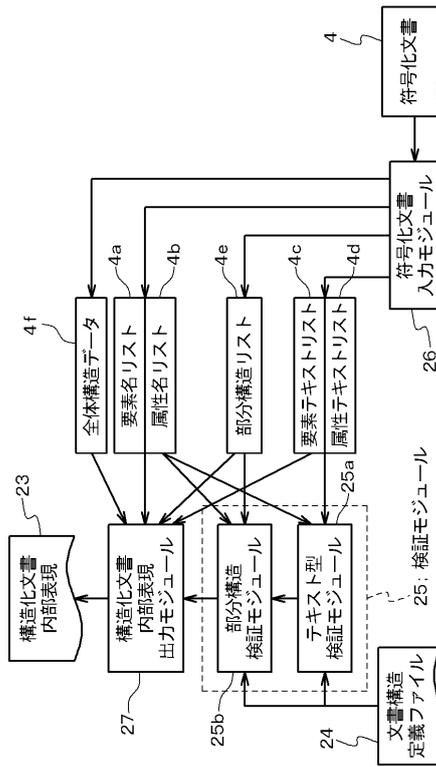
4e: 部分構造リスト

構造ID	部分構造表現
S01	E01 A01 E02 E02
S02	E02 E03 E06 E10
S03	E03 E04 E05
S04	E04 T
...	...

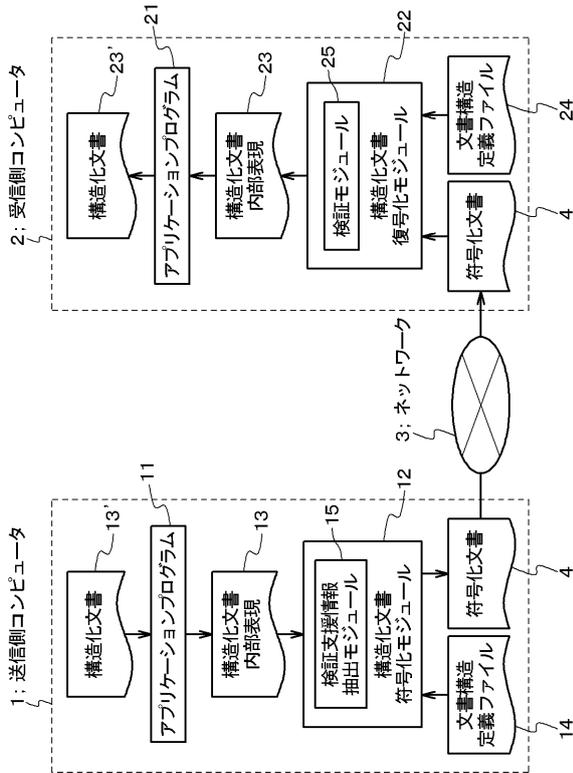
【 図 8 】



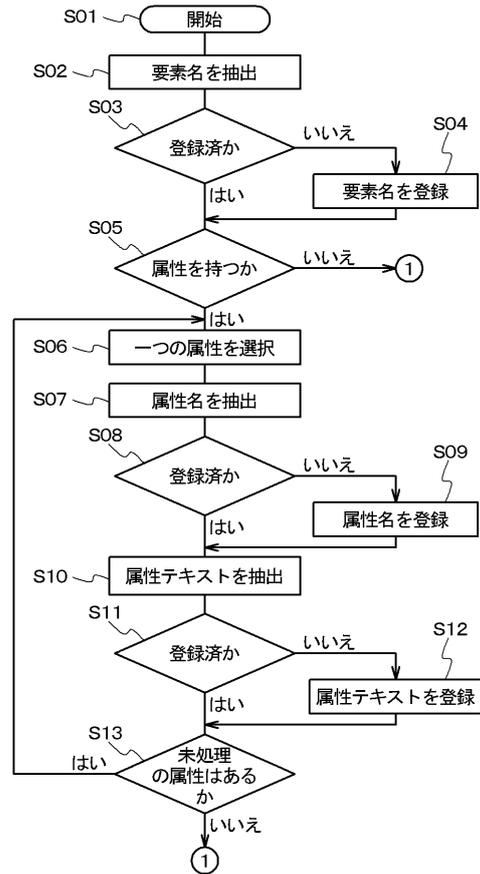
【 図 9 】



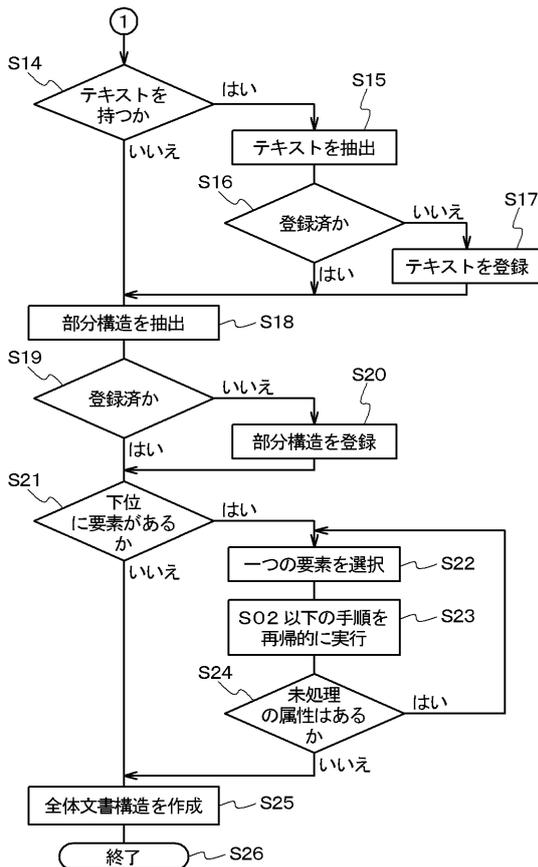
【 図 1 0 】



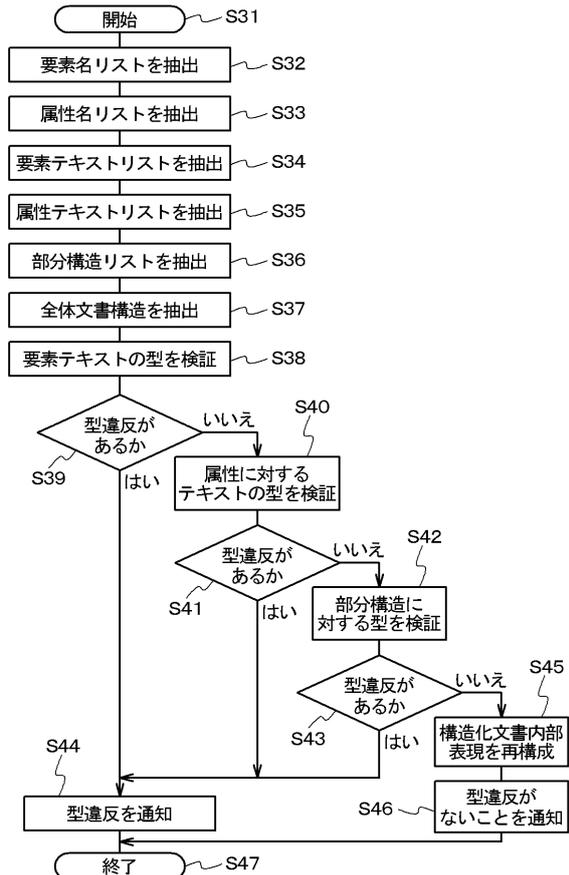
【 図 1 1 】



【 図 1 2 】



【 図 1 3 】



【 図 1 4 】

```

<AddressBook owner = "Suzuki">
  <Person>
    <Name>
      <LastName>Yamada</LastName>
      <FirstName>Taro</FirstName>
    </Name>
    <Address>
      <Country>Japan</Country>
      <Prefecture>Kanagawa</Prefecture>
      <City>Yokohama</City>
    </Address>
    <Telephone>045-123-4567</Telephone>
  </Person>
  <Person>
    <Name>
      <LastName>Yamada</LastName>
      <FirstName>Jiro</FirstName>
    </Name>
    <Address>
      <Country>Japan</Country>
      <Prefecture>Kanagawa</Prefecture>
      <City>Kawasaki</City>
      <Ward>Miyamae</Ward>
    </Address>
  </Person>
</AddressBook>

```

【 図 1 5 】

要素 I D	要素名
E01	AddressBook
E02	Person
E03	Name
E04	LastName
E05	FirstName
E06	Address
E07	Country
E08	Prefecture
E09	City
E10	Telephone
E11	Ward

【 図 1 6 】

属性 I D	属性名
A01	Owner

【 図 1 7 】

要素 I D	テキスト I D	要素テキスト
E04	T01	Yamada
E05	T02	Taro
E05	T03	Jiro
E07	T04	Japan
E08	T05	Kanagawa
E09	T06	Yokohama
E09	T07	Kawasaki
E10	T08	045-123-4567
E11	T09	Miyamae

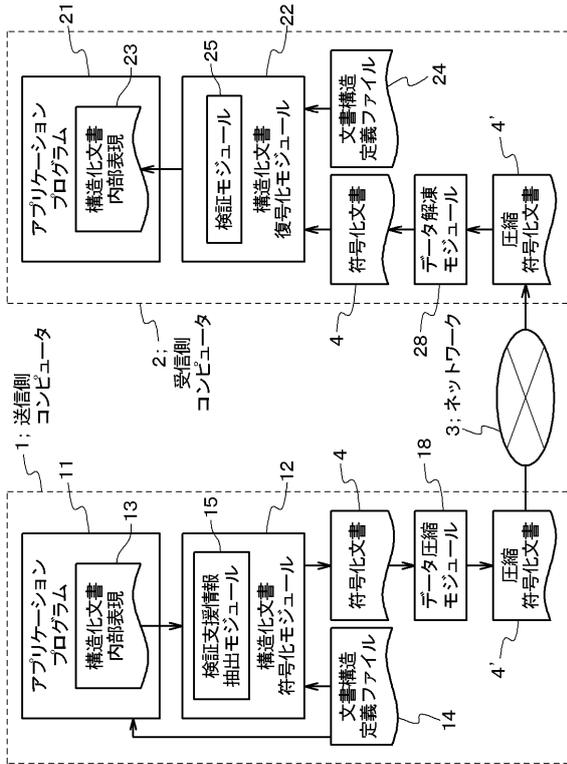
【 図 1 9 】

構造 I D	部分構造表現
S01	E01 A01 E02 E02
S02	E02 E03 E06 E10
S03	E03 E04 E05
S04	E04 T
S05	E05 T
S06	E06 E07 E08 E09
S07	E07 T
S08	E08 T
S09	E09 T
S10	E10 T
S11	E02 E03 E06
S12	E06 E07 E08 E09 E11
S13	E11 T

【 図 1 8 】

属性 I D	テキスト I D	属性テキスト
A01	T10	Suzuki

【 図 20 】



フロントページの続き

(58)調査した分野(Int.Cl. , D B名)

G06F 17/21-17/26