

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号  
特許第4310068号  
(P4310068)

(45) 発行日 平成21年8月5日(2009.8.5)

(24) 登録日 平成21年5月15日(2009.5.15)

(51) Int.Cl.

G 0 6 F 1 7 / 2 4 ( 2 0 0 6 . 0 1 )

F I

G O 6 F 1 7 / 2 4 5 5 4 N

請求項の数 7 (全 10 頁)

|           |                               |           |                     |
|-----------|-------------------------------|-----------|---------------------|
| (21) 出願番号 | 特願2002-114118 (P2002-114118)  | (73) 特許権者 | 000005108           |
| (22) 出願日  | 平成14年4月17日 (2002. 4. 17)      |           | 株式会社日立製作所           |
| (65) 公開番号 | 特開2003-308318 (P2003-308318A) |           | 東京都千代田区丸の内一丁目6番6号   |
| (43) 公開日  | 平成15年10月31日 (2003.10.31)      | (74) 代理人  | 100100310           |
| 審査請求日     | 平成16年12月1日 (2004. 12. 1)      |           | 弁理士 井上 学            |
| 審判番号      | 不服2006-26765 (P2006-26765/J1) | (72) 発明者  | 森本 康嗣               |
| 審判請求日     | 平成18年11月28日 (2006.11.28)      |           | 東京都国分寺市東恋ヶ窪一丁目280番地 |
|           |                               |           | 株式会社日立製作所中央研究所内     |
|           |                               | (72) 発明者  | 須藤 毅                |
|           |                               |           | 東京都江東区新砂一丁目6番27号 株式 |
|           |                               |           | 会社日立製作所公共システム事業部内   |
|           |                               | (72) 発明者  | 小泉 敦子               |
|           |                               |           | 東京都国分寺市東恋ヶ窪一丁目280番地 |
|           |                               |           | 株式会社日立製作所中央研究所内     |

最終頁に続く

(54) 【発明の名称】 文書読解支援方法

(57) 【特許請求の範囲】

【請求項 1】

文書読解支援方法のプログラムを格納するプログラム記憶媒体であって、前記プログラムは、CPUに、  
ファイル装置に記憶された少なくとも2個のテキストを少なくとも、読点又は改行文字の直後で分割することによって、それぞれ構成要素に分割するステップと、  
前記テキストを比較する際に、前記テキストから得られた構成要素同士の比較を、文字を単位とするDPマッチングによって行うステップと、  
前記構成要素同士の比較結果に基づいて作成された構成要素同士の類似行列に基づいて、前記テキスト同士の比較を、構成要素を単位とするDPマッチングにより行うステップと、  
前記構成要素を単位とする比較結果を出力するステップとを実行させることを特徴とするプログラム記憶媒体。

【請求項 2】

請求項 1 に記載のプログラム記憶媒体であって、  
前記比較対象となる少なくとも2個のテキストは、論理的に同一の文書から取り出された部分テキストであることを特徴とするプログラム記憶媒体。

【請求項 3】

請求項 2 に記載のプログラム記憶媒体であって、  
前記論理的に同一の文書が、コンピュータシステム上の同一のファイルであることを特徴

とするプログラム記憶媒体。

【請求項 4】

請求項 2 または 3 に記載のプログラム記憶媒体であって、  
前記プログラムは、CPU にさらに、前記論理的に同一の文書から部分テキストを取り出す際、候補となる部分テキストを抽出して出力するステップと、  
入力手段からの指示に従って前記部分テキストの中から、少なくとも 2 個のテキストを選択するステップとを実行させることを特徴とするプログラム記憶媒体。

【請求項 5】

請求項 4 に記載のプログラム記憶媒体であって、  
前記プログラムは、CPU にさらに、前記候補となる部分テキストを出力する際に、前記候補となる部分テキストの組に関する優先順位を示す情報を出力させるステップを実行させることを特徴とするプログラム記憶媒体。

10

【請求項 6】

請求項 4 に記載のプログラム記憶媒体であって、  
前記プログラムは、CPU にさらに、前記候補となる部分テキストを出力する際に、前記候補となる部分テキストの同士の依存関係を示す情報を出力させるステップを実行させることを特徴とするプログラム記憶媒体。

【請求項 7】

請求項 2 に記載のプログラム記憶媒体において、前記プログラムは、CPU にさらに、  
上記比較対照のテキストそれぞれにのみに含まれる構成要素を比較結果に基づいて選択するユーザ入力を受け付けるステップと、  
前記選択された構成要素中の文字列が上記文書に含まれるか否かを検索するステップと、  
上記文字列の対応箇所を出力するステップとを実行させることを特徴とするプログラム記憶媒体。

20

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、文書の読解支援方法に関し、特に、特許明細書の請求項のような、類似した文が繰り返される箇所を含む文書の読解を支援する方法をコンピュータに実現させるためのプログラムに関する。

30

【0002】

【従来の技術】

ワープロや PC などの普及により、ほとんどの文書が電子的に作成されるようになったことから、計算機を用いて、文書の読解を支援するシステムが多く提案されている。例えば、亀田雅之、日本語文書読解支援系 QJR の検討、情報処理学会自然言語処理研究会資料 NL10-9, pp.57-64 (1996) には、形態素解析や構文解析を施した結果を利用して文書の読解を支援する方法が述べられている。

【0003】

【発明が解決しようとする課題】

しかしながら、現在のところ、構文解析、意味解析、文脈処理などの自然言語処理技術の精度は十分ではないため、誤った解析結果に基づいた支援情報が出力されることが多く、必ずしもユーザの利便性向上には役立っていなかった。一方で、特許や論文のように、一定の約束事に従って作成される文書の場合には、必ずしも、意味解析、文脈処理のような成熟していない技術を利用しなくても、ユーザの利便性向上を図ることが可能である。本発明の目的は、特に、特許明細書の請求項のような、類似した文が繰り返される箇所を含む文書に対し、より簡便な方法によって、確実な効果が得られる文書の読解支援方法を提供することにある。

40

【0004】

【課題を解決するための手段】

本発明の課題は、

50

少なくとも2個のテキストをそれぞれ構成要素に分割し、  
前記テキストを比較する際に、前記テキストから得られた構成要素同士の比較を最初に行い、  
前記構成要素同士の比較結果に基づいて、前記テキスト同士の比較を行い、  
前記比較結果を表示することによって解決できる。

【 0 0 0 5 】

【発明の実施の形態】

以下、本発明の実施例を図面を用いて説明する。

図1に、本発明による特許明細書読解支援システムの構成を示す。本発明による特許明細書読解支援システムは、処理対象となる特許明細書データ12および特許明細書読解支援処理プログラム11が格納されたファイル装置1、各データおよびプログラムが必要に応じてロードされるメモリ2、ファイルおよびメモリ上のプログラムとデータによって処理を実行するCPU3、ユーザへの結果の表示/ユーザからのデータの入力などを行う入出力装置4からなる。特許明細書読解支援処理プログラム11は、請求項認識モジュール111、ユーザ指定認識モジュール112、類似請求項検索モジュール113、請求項依存関係決定モジュール114、請求項比較モジュール115、比較結果表示モジュール116、実施例参照モジュール117からなる。

以下では、本実施例における特許明細書の請求項読解支援処理について図2の処理フローを用いて説明する。

請求項認識モジュール111は、入力明細書中の各請求項を認識する(ステップ11)。請求項の認識には、請求項8のように括弧で示されるタグを利用することができる。そのため、高い精度で各請求項の認識を行うことが可能である。

比較対象となる少なくとも2個の請求項を抽出する(ステップ12)。比較対象となる請求項の抽出を図3に示す処理フローを用いて詳細に説明する。

ユーザ指定認識モジュール112は、ユーザがマウスなどのデバイスを用いて指定した請求項を認識する(ステップ121)。その際、請求項認識モジュールによる認識結果を用いることにより、請求項全体をマウスなどのデバイスで指定することなく、所望の請求項の一部をクリックするなどの操作のみにより、請求項の指定を行うことができる。選択された請求項は、選択されていることが分かるように、選択されていない請求項とは異なる様態、例えば、色を反転するなどして表示する。図4に選択状態の表示例を示す。図4では、請求項6が選択されている。

類似請求項検索モジュール113は、ユーザの指定に従って、ユーザが指定した請求項に類似している請求項を検索する(ステップ122)。類似している請求項の検索は、DPマッチングによって行われる。DPマッチングによる類似請求項検索処理の詳細についてはステップ13の説明において後述する。検索結果は、例えば、図5に示すように表示される。図5では、ユーザが指定した請求項6に類似した請求項が、類似度を示すスコアの降順に表示されている。ユーザは表示された請求項の中から、比較対象としたい請求項を選択すればよい。また、類似度が高い順に予め定められた個数の請求項を選択する、予め定められた閾値より大きいスコアを持つ請求項を選択するなどの方法により、自動的に比較対象となる請求項を選択することも可能である。

請求項依存関係決定モジュール114は、請求項の依存関係を決定する(ステップ123)。請求項の依存関係とは、ある請求項が他の請求項に直接従属している請求項であることを言う。請求項の依存関係の決定は、パターンマッチングによって行われる。ここでパターンとは「X1請求項Nに記載のX2」のようなものである。パターンにおいて、「X1」、「X2」、「N」は任意の文字列とマッチする変数部分である。変数部分には、変数部分とマッチする文字列が満たすべき条件を記述することが出来る。依存関係の認識は、予め設けられたパターンと、各請求項i中の文字列がマッチするかどうかを調べることにより行われる。例えば、請求項iが上記のパターンとマッチし、「N」とマッチした文字列が数値jであれば、請求項iは請求項jに依存しているといったように判定する。このように決定された依存関係は、例えば図6のように表示される。図6では、請求項6がユーザによってマウスなどで選択さ

10

20

30

40

50

れていることが編み掛けによって示されている。そして、請求項6「が」依存している請求項2は下線によって示され、請求項6「に」依存している請求項9が太字の斜体によって示されている。また、別の表示方法として、図7のように依存関係を直接グラフなどの形式で表示しても良い。ユーザは、表示された依存関係を見て、比較対象としたい請求項を選択すれば良い。また、この依存関係に基づいて、比較対象としたい請求項を自動的に選択することもできる。例えば、ユーザがステップ121において、ある請求項を選択し、ステップ123において、請求項の依存関係が自動的に決定されたとする。このとき、依存関係を表すツリー上において、選択されている請求項と予め定められた関係を持つ請求項、例えば、兄弟にあたる請求項、を比較対象として自動的に選択することができる。

請求項比較モジュール115は、ユーザが指定した少なくとも2個以上の請求項を比較し、請求項間で一致する箇所および異なる箇所を認識する(ステップ13)。以下では、本ステップの処理を詳細に説明する。

10

まず本ステップでは、全ての請求項をいくつかの構成要素に分割する。ここで、構成要素とは、請求項を構成する論理的な構成要素などを念頭に置いているが、必ずしも請求項の論理的構造を正しく反映しているように分割する必要はない。例えば、請求項の途中に改行が使われているような場合には、改行を区切りと考えて、請求項を構成要素に分割する。他に、例えば、請求項の途中に読点「、」が使われているような場合には、読点を用いて請求項を分割することもできる。また、「～を含むことを特徴とする～」のような定型的な表現を検出し、この部分を構成要素として分割することも可能である。

次に、全ての請求項から任意の2個を取り出し、比較処理を行う。以下では、図4の2個の請求項、請求項6と請求項7が選択された場合を例に説明する。選択された2個の請求項6、7がそれぞれ、 $N_i$ 個および $N_j$ 個の構成要素からなっているとすると、図4の例を構成要素に分割した結果を図8に示す。図8は、請求項の構成要素を1行毎に示し、その構成要素が $i$ 番目の請求項の $j$ 番目の構成要素である場合に、行頭に、「 $i-j$ 」という形で構成要素のIDを示している。図8の場合、 $N_i$ と $N_j$ はともに5である。次に、それぞれの請求項から取り出した $n_i$ 番目、及び $n_j$ 番目の構成要素同士をDPマッチングによって比較する。DPマッチングについては、例えば、小田中他、情報数学、コロナ社、(2000)などに記載されている。ここでは、簡単にDPマッチングの適用について説明する。

20

図8の2個の請求項のそれぞれ4番目の構成要素(6-4および7-4)同士を比較する場合を考える。このとき、各構成要素の文字数をそれぞれ $M, N$ とする。このとき、構成要素6-4および7-4中の文字に基づいて、図9のような $M \times N$ 行列を構成する。行列の $(m, n)$ 要素には、構成要素6-4の $m$ 番目の文字と構成要素7-4の $n$ 番目の文字が同じであれば1が格納され、異なる場合には0が格納されている。

30

このような行列が与えられたとき、 $(1, 1)$ 要素から $(M, N)$ 要素に至るパスが2個の文のマッチング結果に相当する。図9のパスは、 $(1, 1) \rightarrow (2, 2) \rightarrow \dots \rightarrow (13, 18)$ となっているが、これは、1文字目同士が対応し、2文字目同士が対応し、最終的に構成要素6-4の13番目の文字「段」と構成要素7-4の18番目の文字「段」が対応していることを示している。このモデルでは、「補正された」のように一方にしか含まれていない表現をスキップして照合することなどが容易に行える。パス上の要素の値が0の場合には、対応する要素がない表現であると考えれば良い。DPマッチングは、可能なパスの中から上に述べたような最適なマッチング結果を示すパスを発見する方法の一つである。

40

以下、図10に示す簡単な例を用いて最適なパスの発見方法を説明する。図10では、2個の文字列ABを比較する場合を例に説明している。この場合、1文字目のA同士が対応し、2文字目のB同士が対応するのが最適な結果であるが、これをDPマッチングによって計算する過程を説明する。図10には、3種類の配列が示されている。図10(a)は、図9と同種の配列であり、一方の構成要素の $m$ 文字目と他方の構成要素の $n$ 文字目が一致しているときには $(m, n)$ 要素に1が、異なっているときには0が格納された行列である。この行列の要素を $a(m, n)$ とする。図10(b)は、要素 $(m, n)$ までのスコアの総和 $s(m, n)$ を格納した行列であり、図10(c)は、要素 $(1, 1)$ から $(m, n)$ までのパス長 $l(m, n)$ を格納した行列である。スコアの総和とパス長は、次のように計算される。要素 $(m, n)$ へのパスは、必ず、 $(m-1, n)$ 要素

50

、 $(m-1, n-1)$ 要素、 $(m, n-1)$ 要素のいずれかを通るものとする。このとき、 $s(m-1, n)/(l(m-1, n)+1)$ 、 $s(m-1, n-1)/(l(m-1, n-1)+1)$ 、 $s(m, n-1)/(l(m, n-1)+1)$ の3個の値(パス上の平均スコア)を比較し、最大の値を取る要素から $(m, n)$ 要素へのパスを通るものとする。そして、最大の値を取る要素に対して、 $s(m, n)=s(k, l)+a(m, n)$ 、 $l(m, n)=l(k, l)+1$ と計算する。ただし、 $m=1$ の場合は、 $(m-1, n)$ 要素、 $(m-1, n-1)$ 要素を考慮することができないため、パスは $(m, n-1)$ 要素を通る。図10(b)の場合、 $s(2, 1)=s(1, 2)=1$ 、 $l(2, 1)=l(1, 2)=1$ となるため、 $(2, 2)$ 要素については、 $(1, 1)$   $(2, 1)$   $(2, 2)$ のパスについては、 $s(m, n-1)/(l(m, n-1)+1)$ を計算すると $1/2=0.5$ 、 $(1, 1)$   $(1, 2)$   $(2, 2)$ のパスについては、 $s(m-1, n)/(l(m-1, n)+1)=1/2=0.5$ となり、 $(1, 1)$   $(2, 2)$ のパスについては $s(m-1, n-1)/(l(m-1, n-1)+1)=1/1=1.0$ となるため、 $(1, 1)$   $(2, 2)$ のパスが最も良いことになる。配列 $p$ は $p(m, n)$ 要素に到るパス上の直前のノードを格納する。例えば、 $p(2, 2)$ には $(1, 1)$ を格納する。そして、 $(2, 2)$ の位置から $(1, 1)$ の位置まで、行列 $p$ の示す直前の要素へのパスを辿ることによって最適なパス、すなわち最良のマッチング結果が得られ、 $s(2, 2)/l(2, 2)$ によってこのパスのスコアが得られる。

以上の比較を $N_i \times N_j$ 個(図8の例の場合、 $5 \times 5=25$ 個)の全ての構成要素の組み合わせについて行う。すると、 $N_i \times N_j$ の実数の要素からなる行列を作成することができる。例を図11に示す。この行列は、構成要素を単位とする2個の請求項同士の対応関係を示しており、この行列の $(i, j)$ 要素には、 $i$ 番目の構成要素と $j$ 番目の構成要素との文字単位のマッチング結果によって得られるスコアの値が格納されている。スコアの値は0以上、1以下の実数値である。この行列においてDPマッチングにより最適なパスを求めることによって、構成要素同士の比較の場合と同様に、請求項同士の対応関係とスコアを求めることができる。図11の例では、 $(5-1, 6-1)$   $(5-2, 6-2)$   $(5-3, 6-3)$   $(5-4, 6-4)$   $(5-5, 6-5)$ という構成要素同士の対応関係が得られる。図11の例の場合、スキップされている要素はないが、図9を用いて説明したように、対応する構成要素を持たない、一方の請求項にのみ含まれている構成要素が存在した場合には、最良のパス上であっても、行列の要素の値が小さくなる。ただし、構成要素同士を文字列比較する場合とは異なり、その値は必ずしも0とはならない。それは、「の」のような頻出する文字は、対応しない構成要素であっても共通して含まれていることがあるためである。よって、予め閾値を定めておき、パス上の行列要素の値がこの閾値より小さい場合には、対応する要素ではないものとする。なお、構成要素の対応関係が1対2などの場合もあり得るが、このような場合については、例えば、森本他, 語の対訳知識を用いた対訳テキストの文対応付けアルゴリズム, 言語処理学会第1回年次大会, (1995)に記載されているような方法で、対応関係を検出することができるため、説明は省略する。

以上の処理をステップ12で選択された請求項の任意の組み合わせについて行うことにより、任意の請求項の組に対するスコア、任意の請求項の任意の構成要素と対応する他の請求項の構成要素の一覧とそれぞれに対するスコアなどが、比較結果として得られる。比較結果のデータ構造の説明図を図12に示す。比較結果は、請求項が何個あるかなどの情報を示す請求項リストと各請求項毎に何個の構成要素があるかなどの情報を示す構成要素リスト上に、請求項同士の対応関係を示す請求項対応関係リンクと、構成要素同士の対応関係を示す構成要素対応関係リンクを張ることによって管理する。対応関係リンクは、リンクの値として、対応関係のスコアを保持しておく必要がある。

比較結果表示モジュール116は、請求項比較モジュールの処理結果をユーザに表示する(ステップ14)。表示方法の例の一つを図13に示す。図13では、請求項5、請求項6、請求項7という3個の請求項が同時に表示されている。各請求項は、それぞれ3個、5個、5個の構成要素に分割されている。これらの構成要素は、対応関係に基づき、5種類の様態、すなわち、(1)斜体のフォント、(2)網掛けされた四角で囲まれている、(3)ゴシック体のフォント、(4)下線、(5)透明な四角で囲まれている、によって区別して示されている。例えば、構成要素5-1、すなわち、請求項5の1番目の構成要素は、構成要素6-1、構成要素7-1と斜体のフォントが施されており、対応する構成要素であることが分かる。このような表示は、構成要素の対応関係が抽出されているため、実現が可能となる。もちろん、網掛

10

20

30

40

50

け以外に異なる色を用いるなどの方法でも良い。このように、対応する構成要素は同じ状態での表示が行われるため、請求項間の対応関係を容易に認識することができ、請求項の読解が容易になる。

また表示方法の別の例の一つを図14に示す。図14では、請求項6、請求項7という2個の請求項が同時に表示されている。また、対応する構成要素が、等しい行位置となるように、位置を調整して表示する。これにより、図13の場合と同様に、対応する構成要素を容易に理解でき、請求項の読解が容易になる。また、図14の場合には、対応する構成要素同士について、文字レベルで異なった部分について下線を施している。また、スキップされている構成要素やスコアが低いために対応する構成要素を持たない構成要素を区別して表示することもできる。これにより、請求項間の差異が容易に理解できるようになる。なお、対応する構成要素同士の文字レベルの差異を図13の表示方法について実現するためには若干の工夫が必要となる。なぜならば、図13の表示方法では、3個以上のn個の請求項を同時に比較・表示することが必要となるためである。そのため、差異の表示は例えば、次のようにする。すなわち、全ての請求項に共通である文字列は、下線なし、当該請求項にのみ含まれている文字列には2重下線、それ以外の文字列には1重下線を施す。それ以外にも、ユーザが2個の対応する構成要素の組を選択すると、文字レベルの差異を表示するように設計しても良い。

実施例参照モジュール117は、請求項に対応する実施例中の箇所を発見して表示する(ステップ15)。本ステップの処理を図15の処理フローを用いて詳細に説明する。

ステップ14で表示された結果をユーザが参照し、マウスなどのデバイスを用いてユーザが指定した請求項を認識する(ステップ151)。

ユーザが指定した請求項のみに含まれる構成要素を、この請求項の特徴文字列として取得する(ステップ152)。

取得された特徴文字列を実施例全体の中から検索する(ステップ153)。高速に検索処理を行うためには、例えば、特開平5-174064号公報などに示される方法を用いることができるため説明を省略する。また、特徴文字列となった構成要素の文字列全体を検索してもマッチしない場合が考えられるため、特徴文字列を単語に分割した後に検索を行うことも考えられるが、これらは例えば、長尾編,自然言語処理,岩波講座ソフトウェア科学,岩波書店,(1996)などに記載の技術を用いることが可能であるため、説明を省略する。

なお、以上の実施例では、ステップ12において、比較対象とする請求項を選択した上で、比較を行う方法を説明したが、それ以外に次のようにすることもできる。

ユーザがいずれか1個の請求項をマウスなどのデバイスを選択する。すると、この時点で、選択された請求項とそれ以外の請求項について比較を行う。比較処理は動的に行っても良いし、予め任意の組み合わせについて処理をしておいても良い。このようにして得られた比較結果を表示する。表示は、例えば、図13のように行うことができる。また、別の方法として、比較した結果、最初に選択した請求とのスコアが最も高い請求項を抽出し、図14のような方法で表示することもできる。また、スコア順に比較対象となる請求項を変更することなども容易である。

あるいは、事前に全ての請求項の組み合わせについて処理をしておき、図13のように表示するのも良い。

【0006】

【発明の効果】

本発明によれば、特許の請求項のような、類似した文が繰り返される箇所について、全く同じである箇所と異なる箇所を、異なる様態で表示できるようになるため、文書の読解が容易になる。

【図面の簡単な説明】

【図1】本発明の実施例である文書読解支援システムのブロック図である。

【図2】文書読解支援システムの処理フローである。

【図3】比較対象請求項抽出の処理フローである。

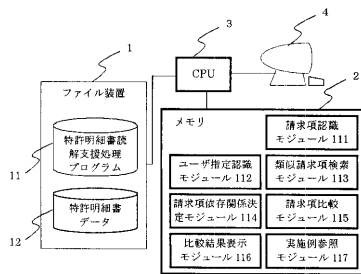
【図4】ユーザ指定によって選択された請求項の表示の例である。

- 【図 5】類似請求項検索結果の表示の例である。  
 【図 6】請求項の依存関係表示の第 1 の例である。  
 【図 7】請求項の依存関係表示の第 2 の例である。  
 【図 8】請求項の構成要素への分割結果の例である。  
 【図 9】構成要素同士の DP マッチングの例である。  
 【図 10】DP マッチングによるパスの計算方法の説明図である。  
 【図 11】請求項同士の DP マッチングの例である。  
 【図 12】比較結果のデータ構造の説明図である。  
 【図 13】請求項の比較結果表示の第 1 の例である。  
 【図 14】請求項の比較結果表示の第 2 の例である。  
 【図 15】実施例参照処理の処理フローである。

10

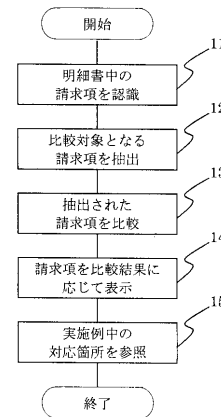
【図 1】

文書読解支援システムのブロック図 図 1



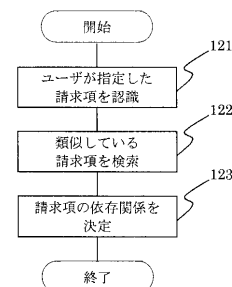
【図 2】

文書読解支援システムの処理フロー 図 2



【図 3】

比較対象請求項抽出の処理フロー 図 3



【 図 4 】

ユーザ指定によって選択された請求項の表示の例 図4

【特許請求の範囲】

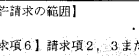
.....

【請求項6】 請求項2、または4記載のX装置であって、  
 文書中に出現する語の候補を入力する手段と、  
 単語の出現頻度に基づいて単語の尤度を決定する手段と、  
 各候補の尤度を出力する手段  
 を含むことを特徴とするX装置。

【請求項7】 請求項2、または4記載のX装置であって、  
 文書中に出現する語の候補を尤度とともに入力する手段と、  
 単語の出現頻度に基づいて単語の尤度を補正する手段と、  
 各候補の補正された尤度を出力する手段  
 を含むことを特徴とするX装置。

【 図 5 】

類似請求項検索結果の表示の例 図5



**【特許請求の範囲】**

【請求項6】請求項2、3または4  
 文書中に出現する語の候補を尤度  
 単語の出現頻度に基づいて単語の  
 各候補の尤度を出力する手段  
 を含むことを特徴とするX装置。

【請求項7】請求項2、3または4  
 文書中に出現する語の候補を尤度  
 単語の出現頻度に基づいて単語の  
 各候補の補正された尤度を出力す  
 るを含むことを特徴とするX装置。

|        |       |
|--------|-------|
| 【請求項7】 | 0.952 |
| 【請求項5】 | 0.805 |
| 【請求項3】 | 0.505 |

【 図 6 】

請求項の依存関係表示の第1の例 図6

☐ ☒

【特許請求の範囲】

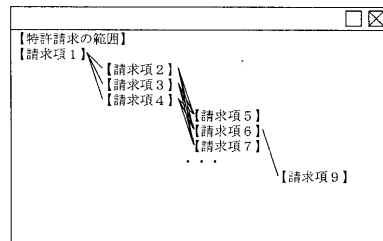
【請求項2】 請求項1記載のX装置であって、  
単語の出現情報に基づいて単語の尤度を決定する手段  
を含むことを特徴とするX装置。

【請求項6】 請求項2、3または4記載のX装置であって、  
文書中に出現する語の候補を入力する手段と、  
単語の出現頻度に基づいて単語の尤度を決定する手段と、  
各候補の尤度を出力する手段  
を含むことを特徴とするX装置。

【請求項9】 請求項6記載のX装置であって、

【圖 7】

請求項の依存関係表示の第2の例 図7



【 図 8 】

請求項の構成要素への分割結果の例 図8

6-1 請求項2、3または4記載のX装置であって、  
6-2 文書中に出現する語の候補を入力する手段と、  
6-3 単語の出現頻度に基づいて単語の尤度を決定する手段と、  
6-4 各候補の尤度を出力する手段  
6-5を含むことを特徴とするX装置。

7-1 請求項2、3または4記載のX装置であって、  
7-2 文書中に出現する語の候補を尤度とともに入力する手段と、  
7-3 単語の出現頻度に基づいて単語の尤度を補正する手段と、  
7-4 各候補の補正された尤度を出力する手段  
7-5を含むことを特徴とするX装置。

【 図 1 0 】

DP マッチングによるパスの計算方法の説明図 図 10

|   | A | B |
|---|---|---|
| A | 1 | 0 |
| B | 0 | 1 |

|   | A   | B   |
|---|-----|-----|
| A | 1,0 | 1,0 |
| B | 1,0 | 1,0 |

|   | A | B |
|---|---|---|
| A | 0 | 1 |
| B | 1 | 1 |

|   | A   | B   |
|---|-----|-----|
| A | -   | 1,1 |
| B | 1,1 | 1,1 |

(a) 配列 a                      (b) 配列 s                      (c) 配列 l                      (d) 配列 p

【 図 1 1 】

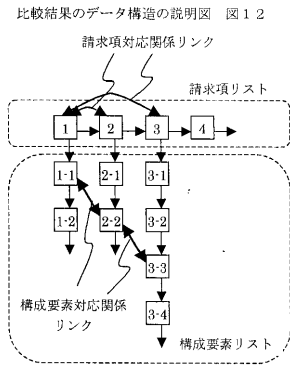
請求項同士の DP マッチングの例 図 1 1

[illegible]

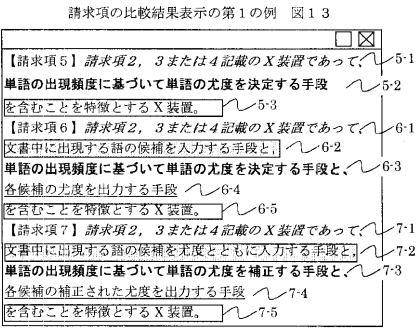
|     | 6-1 | 6-2 | 6-3 | 6-4 | 6-5 |
|-----|-----|-----|-----|-----|-----|
| 5-1 | 1.0 | 0.1 | 0.2 | 0.0 | 0   |
| 5-2 | 0   | 0.9 | 0   | 0   | 0.1 |
| 5-3 | 0   | 0   | 0.8 | 0   | 0   |
| 5-4 | 0.2 | 0   | 0   | 0.7 | 0   |
| 5-5 | 0   | 0.1 | 0   | 0   | 0.8 |



【図 1 2】



【図 1 3】



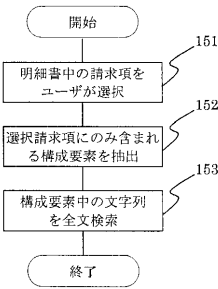
【図 1 4】

請求項の比較結果表示の第 2 の例 図 1 4

|  |   |
|--|---|
| 【請求項 6】<br>請求項 2, 3 または 4 記載の X 装置であって、<br>文書中に出現する語の候補を入力する手段と、<br>単語の出現頻度に基づいて単語の尤度を決定する手段と、<br>各候補の尤度を出力する手段<br>を含むことを特徴とする X 装置。 | 【請求項 7】<br>請求項 2, 3 または 4 記載の X 装置であって、<br>文書中に出現する語の候補を尤度とともに入力する手段と、<br>単語の出現頻度に基づいて単語の尤度を補正する手段と、<br>各候補の補正された尤度を出力する手段<br>を含むことを特徴とする X 装置。 |
|--|---|

【図 1 5】

実施例参照処理の処理フロー 図 1 5



---

フロントページの続き

合議体

審判長 立川 功

審判官 真木 健彦

審判官 長島 孝志

- (56)参考文献 特開 2 0 0 1 - 9 2 8 3 0 ( J P , A )  
特開平 9 - 1 6 7 1 6 4 ( J P , A )  
特開 2 0 0 2 - 6 3 1 9 2 ( J P , A )  
特開平 8 - 1 0 6 4 7 4 ( J P , A )  
特開 2 0 0 0 - 1 1 2 9 4 9 ( J P , A )

- (58)調査した分野(Int.Cl. , D B 名)

G06F17/20-17/26

G06F17/30

G06F17/60-19/00