

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4102897号  
(P4102897)

(45) 発行日 平成20年6月18日(2008.6.18)

(24) 登録日 平成20年4月4日(2008.4.4)

(51) Int. Cl. F I  
**G 0 6 F 17/28 (2006.01)** G O 6 F 17/28 U  
**G 0 6 F 17/21 (2006.01)** G O 6 F 17/21 5 5 0 A

請求項の数 16 (全 23 頁)

(21) 出願番号	特願2003-186186 (P2003-186186)	(73) 特許権者	302029118
(22) 出願日	平成15年6月30日(2003.6.30)		渋谷 正敏
(65) 公開番号	特開2004-133896 (P2004-133896A)		東京都大田区南蒲田1丁目6番3号
(43) 公開日	平成16年4月30日(2004.4.30)	(73) 特許権者	393009909
審査請求日	平成17年5月10日(2005.5.10)		佐藤 謙治
(31) 優先権主張番号	特願2002-236372 (P2002-236372)		東京都中野区中野2-13-21-303
(32) 優先日	平成14年8月14日(2002.8.14)	(72) 発明者	渋谷 正敏
(33) 優先権主張国	日本国(JP)		東京都大田区南蒲田1丁目6番3号松原荘
特許権者において、権利譲渡・実施許諾の用意がある。		(72) 発明者	佐藤 謙治
			東京都中野区中野3-11-10-205
		審査官	和田 財太

最終頁に続く

(54) 【発明の名称】 特許明細書デバッグツールおよび特許明細書デバッグツールプログラム

(57) 【特許請求の範囲】

【請求項1】

データ入出力部を有する特許明細書デバッグツールであって、  
 前記データ入出力部に入力された日本語の特許明細書のクレーム文を解析する辞書ファイルと、前記辞書ファイルを用いて解析された結果として得られた前記クレーム文の格助詞「と」に先行して記載されている名詞または名詞と同じ働きを持つ名詞句をクレーム文の技術要素として特定する第1の技術要素特定手段と、前記辞書ファイルを用いて解析された結果として得られた前記クレーム文のものを指し示す修飾語である指示語に続けて記載されている名詞または名詞と同じ働きを持つ名詞句をクレーム文の技術要素として特定する第2の技術要素特定手段と、前記辞書ファイルを用いて解析して得られた前記クレーム文の末尾に記載されている名詞または名詞と同じ働きを持つ名詞句を発明の主題として特定する主題特定手段と、前記第1および第2の技術要素特定手段により該特定された技術要素の中から、前記主題特定手段により特定された前記主題に直接関係する該技術要素を1次の技術要素として特定する1次の技術要素特定手段と、この1次の技術要素特定手段により特定された該1次の技術要素と前記主題とを図面の上に配置し、該1次の技術要素と前記主題との関係を図式として表示するクレーム文第一の図式化手段とを備えたことを特徴とする特許明細書デバッグツール。

【請求項2】

前記辞書ファイルには日本文の自動解析に用いる形態素辞書にクレーム文の技術要素として用いられた用語を名詞相当の用語として追加登録した形態素解析用辞書を含むことを特

徴とする請求項 1 に記載の特許明細書デバッグツール。

【請求項 3】

前記主題特定手段には、前記クレーム文の中から前提部と特徴部とに分離する要素を検出し、前記クレーム文を前提部と特徴部とに分離する機能と前記前提部の主題を特定する機能と、前記クレーム文を前提部と特徴部に分離する要素を検出できない場合は、前記クレーム文には前提部がなく、前記クレーム文を全て特徴部として処理する機能とを含むことを特徴とする請求項 1 または請求項 2 に記載の特許明細書デバッグツール。

【請求項 4】

前記クレーム文第一の図式化手段には、前記前提部の主題と前記特徴部の主題ごとに前記 1 次の技術要素との関係を図式として表示する機能を有することを特徴とする請求項 1 乃至 3 のいずれかに記載の特許明細書デバッグツール。

10

【請求項 5】

前記第 1 および第 2 の技術要素特定手段のうち、前記指示語により特定された技術要素に関して前記クレーム文の中の参照元の技術要素を特定する参照元技術要素特定手段と、この参照元技術要素特定手段により特定された該参照元の技術要素と該指示語により特定された技術要素とを前記クレーム文第一の図式化手段により表示された前記図式の上で線で結び、なおかつ、該指示語を該図式から削除した該図式を表示するクレーム文第二の図式化手段とを備えたことを特徴とする請求項 1 乃至 4 のいずれかに記載の特許明細書デバッグツール。

【請求項 6】

20

前記第 1 および第 2 の技術要素特定手段の中に、前記指示語の後に動詞または動詞と同じ働きを持つ動詞句または用言が続き、その後に名詞または名詞と同じ働きをもつ名詞句が続く場合は、該動詞または該動詞句または該用言を除いた該名詞または名詞と同じ働きをもつ名詞句を技術要素として特定する機能を含むことを特徴とする請求項 1 乃至 5 のいずれかに記載の特許明細書デバッグツール。

【請求項 7】

前記第 1 および第 2 の技術要素特定手段の中に、前記クレーム文の中の並立の意味を有する接続詞を格助詞「と」に変換する機能を含むことを特徴とする請求項 1 乃至 6 のいずれかに記載の特許明細書デバッグツール。

【請求項 8】

30

前記第 1 および第 2 の技術要素特定手段の中で、格助詞「と」に前置された語句として特定された技術要素を「と」による技術要素とし、この格助詞「と」の後に続く形態素の品詞と読点または符号とにより前記「と」による技術要素の属性を定め、該属性が前記クレーム文の中で先行した「と」による技術要素の属性から変化した状態を遷移状態として、該遷移状態により、前記「と」による技術要素どうしの従属関係を演算式またはこの演算式に準じた方式で決定する方法を前記 1 次の技術要素特定手段に含むことを特徴とする請求項 1 乃至 7 のいずれかに記載の特許明細書デバッグツール。

【請求項 9】

前記 1 次の技術要素特定手段のうち、前記「と」による技術要素どうしの従属関係を演算式で決定する方法の中に、前記「と」による技術要素に対し位置づけの値を定義する演算式を含み、該位置づけの値を定義する演算式は、前記遷移状態に基づいて定義され、また、特別に定義された技術要素については位置づけの値を補正できる機能を含み、前記位置づけの値の大小の比較により前記 1 次の技術要素を特定する機能を含むことを特徴とする請求項 1 乃至 8 のいずれかに記載の特許明細書デバッグツール。

40

【請求項 10】

前記「と」による技術要素を特定する方法の中に、前記格助詞「と」が並立の意味を持たない場合は、該格助詞「と」に前置される語句を「と」による技術要素として抽出しない機能を含むことを特徴とする請求項 1 乃至 9 のいずれかに記載の特許明細書デバッグツール。

【請求項 11】

50

前記第1および第2の技術要素特定手段により特定された技術要素の最後尾の形態素が助詞の場合は、該助詞を除いたものを技術要素として特定する機能を有することを特徴とする請求項1乃至10のいずれかに記載の特許明細書デバッグツール。

【請求項12】

前記前提部に属する前記技術要素のうち、他で参照されない技術要素があれば、第一の警告を出力することを特徴とする請求項1乃至11のいずれかに記載の特許明細書デバッグツール。

【請求項13】

前記「と」による技術要素の属性の並びから判断して、格助詞「と」の使い方として、文法に反する場合は、第二の警告を出力することを特徴とする請求項1乃至12のいずれかに記載の特許明細書デバッグツール。

10

【請求項14】

前記主題特定手段により特定された前記前提部の主題と前記の特徴部の主題の比較を行い、それぞれの主題が異なる場合は、「クレーム文の前提部と特徴部との主題が異なる」ことを意味する文章または図形を出力することを特徴とする請求項1乃至13のいずれかに記載の特許明細書デバッグツール。

【請求項15】

前記クレーム文の中の前記技術要素の前記上位概念の用語を前記辞書ファイルの中の上位概念用語辞典から検出して出力する技術要素上位概念検出手段を備えたことを特徴とする請求項1乃至14のいずれかに記載の特許明細書デバッグツール。

20

【請求項16】

前記特徴部の中の複数の技術要素のいずれかと前記前提部の中の複数の技術要素の前記上位概念の用語のいずれかとを比較して同じ語句があれば、第三の警告を出力することを特徴とする請求項1乃至15のいずれかに記載の特許明細書デバッグツール。

【発明の詳細な説明】

【0001】

【発明が属する技術分野】

この発明は、日本語で作成された特許明細書のクレーム文を解析して図式化表示してクレーム文としての問題点を自動検出し出力する特許明細書デバッグツールに関するものである。

30

【0002】

【従来の技術】

特許明細書の作成を支援するツールとしては、特開2001-306754号や特開2000-48013号に開示されているように、特許明細書の作成の手順や作成の要点やチェックリストを表示するツールや、特開2002-207720号に開示されているように、学术论文を特許明細書に変換するツールがある。

【0003】

また、特開平9-293075号に開示されているように、特許明細書のクレーム文をパターン照合した上で階層データを作り第二の言語に翻訳するツールがある。

【0004】

40

その他、本発明の関連技術として重要な日本文を自動解析する技術としては、機械翻訳関係の特開平11-25087号、特開平7-295985等をはじめ多数の文献がある。

【0005】

【発明が解決しようとする課題】

日本語の特許明細書の不慣れな者にとって、特許明細書を作成するのは、たいへん困難な作業であり、特許技術者に依頼する場合も小規模の企業にとっては、高額な負担となり、特許出願を困難なものとしている。特に、クレーム文は長文となる場合があり、さらに種々の指示語の存在が、クレーム文の解読を難解なものにするため、日本語の特許明細書の作成に不慣れな者にとって、自分が作成した特許明細書のクレーム文が自分の意図したように作成されているのか判断するのが困難だけでなく、出願人が専門の特許技術者に特

50

許明細書の作成を依頼した場合でも、作成された特許明細書のクレーム文が出願人に属する発明者が自分が意図したとおりに書かれているか明確に判断できない場合が多く、誤った出願をしてしまうという問題もある。

【0006】

日本語の特許明細書のクレーム文を明確に解読できるようにする方法として、クレーム文を技術要素に分解して図式化表示する方法は有効であるが、従来の技術である特開平9-293075号のように、パターン照合する方法では、日本語の特許明細書のクレーム文のように多種多様な形態をとるものにとって、実用的なレベルになるまでクレーム文のパターンを準備することは事実上困難であり、より日本語の特許明細書のクレーム文の実情にあった方法でクレーム文を解析し、図式化表示する方法が必要になる。

10

【0007】

また、熟練した特許技術者でも、扱う特許明細書の技術分野に不慣れな場合には、不必要な技術要素を過大評価してクレーム文の前提部に記載してしまい請求の範囲を狭めたり、技術要素としてより上位概念の語句を用いてクレーム文を作成できる場合でも、その技術分野に不慣れな場合には、上位概念の語句が思い浮かばなかったりするために、技術要素としては発明者より提示された下位概念の語句をそのまま用いて出願してしまい請求の範囲を狭めてしまう問題点もある。

【0008】

また、本発明の関連技術として重要な日本文を自動解析する技術としては、「原文」を「形態素解析」し、「構文解析」し、「意味解析」していくトランスファー方式が通常用いられているが、通常の記事の場合には、自然言語のあいまい性のために「構文解析」の段階で、まだ十分に精度の良い解析ができないのが現状であるが、特許明細書のクレーム文ではあいまい性をともなう表現が少なく、また常用される特殊な表現方法に着目して、現状の日本文の自動解析技術でクレーム文を解析できるようにする必要がある。

20

【0009】

本発明は上記の問題点を解決するためになされたものであり、多種多様な文構造を持つ日本語の特許明細書のクレーム文であっても、そのクレームを構成する技術要素を格助詞「と」の前に前置される特性と、クレーム文内の技術要素の相互の関係を示すために用いられる「前記」または「該」などの修飾語（以下、ものを指し示す修飾語という意味で「指示語」と表記する）の後に技術要素が配置される特性とに着目し、

30

また、一般の日本文の自動解析に用いる形態素辞書にクレーム文の技術要素として用いられた用語であれば、たとえば、「エンジン電子制御装置」のように本来4つの形態素「エンジン」、「電子」、「制御」、「装置」に分解される用語であっても、クレーム文の技術要素として使用された実績があれば1つの名詞相当の用語として追加登録できる「クレーム文解析用形態素辞書」を用いてクレーム文を形態素解析し、

その形態素解析の結果から格助詞「と」に前置された語句からクレーム文の技術要素を抽出し、その技術要素間の従属関係を論理的に解析して、クレーム文である請求項の末尾に記載される発明を特定する要素（以下「主題」と記載する）に直結する技術要素を1次の技術要素として抽出し、その主題と1次の技術要素ごとに枝分けした木構造で図式化表示し、さらに、クレーム文に含まれる指示語で特定できる技術要素を抽出して、指示語の種類に基づき参照元の技術要素と参照先の技術要素の関係を定めて、上記の木構造の上で技術要素間の引用関係を線で結ぶことで明確に表示し、クレーム文の内容を図式化表示することにより、出願人の発明者や特許明細書作成者がクレーム文が意図したように作成できたかを明確に判断できるようにする。

40

【0010】

また、前提部が存在するクレーム文では、クレーム文を前提部と特徴部に分離し、前提部と特徴部の主題を抽出し、前提部と特徴部ごとに上記の木構造の表示と図式化表示を行うほか、

前提部に他で参照されていない技術要素を自動的に検出し出力するようにして、請求の範

50

困が狭くなることを防止したり、技術要素の上位概念の語句を検出出力し、また、前提部に特徴部より下位概念の語句があれば自動的に検出して、出願人の発明者や特許明細書作成者がクレーム文の技術要素の再検討が容易にできるようにする。

【0011】

【課題を解決するための手段】

この目的を達成するために、本発明の請求項1に記載の特許明細書デバッグツールは、データ入出力部を有する特許明細書デバッグツールであって、データ入出力部を有する特許明細書デバッグツールであって、

前記データ入出力部に入力された日本語の特許明細書のクレーム文を解析する辞書ファイルと、前記辞書ファイルを用いて解析された結果として得られた前記クレーム文の格助詞「と」に先行して記載されている名詞または名詞と同じ働きを持つ名詞句をクレーム文の技術要素として特定する第1の技術要素特定手段と、前記辞書ファイルを用いて解析された結果として得られた前記クレーム文のものを指し示す修飾語である指示語に続けて記載されている名詞または名詞と同じ働きを持つ名詞句をクレーム文の技術要素として特定する第2の技術要素特定手段と、前記辞書ファイルを用いて解析して得られた前記クレーム文の末尾に記載されている名詞または名詞と同じ働きを持つ名詞句を発明の主題として特定する主題特定手段と、前記第1および第2の技術要素特定手段により該特定された技術要素の中から、前記主題特定手段により特定された前記主題に直接関係する該技術要素を1次の技術要素として特定する1次の技術要素特定手段と、この1次の技術要素特定手段により特定された該1次の技術要素と前記主題とを図面の上に配置し、該1次の技術要素と前記主題との関係を図式として表示するクレーム文第一の図式化手段とを備えたことを特徴とするものである。従って、この発明によれば、クレーム文第一の図式化手段により、クレーム文を分解して、その従属関係を図式化して表示されるようになり、クレーム文の構成が明確に理解できるようになる。

【0012】

また、請求項2に記載の特許明細書デバッグツールは、前記辞書ファイルには日本文の自動解析に用いる形態素辞書にクレーム文の技術要素として用いられた用語を名詞相当の用語として追加登録した形態素解析用辞書を含むことを特徴とするものであり、従って、特殊な専門用語や複雑な形態素から構成される技術用語がクレーム文内にあっても容易に技術要素を抽出できるようになる。

【0013】

また、請求項3に記載の特許明細書デバッグツールは、前記主題特定手段には、前記クレーム文の中から前提部と特徴部とに分離する要素を検出し、前記クレーム文を前提部と特徴部とに分離する機能と前記前提部の主題を特定する機能と、前記クレーム文を前提部と特徴部に分離する要素を検出できない場合は、前記クレーム文には前提部がなく、前記クレーム文を全て特徴部として処理する機能とを含むことを特徴とするものであり、従って、前提部付きのクレーム文と、構成要素列挙型のクレーム文とに識別して処理できるようになる。

【0014】

また、請求項4に記載の特許明細書デバッグツールは、前記クレーム文第一の図式化手段には、前記前提部の主題と前記特徴部の主題ごとに前記1次の技術要素との関係を図式として表示する機能を有することを特徴とするものであり、従って、前提部付きのクレーム文であれば、前提部と特徴部に分けて図式化表示できるようになり、構成要素列挙型のクレーム文であれば、クレーム文をすべて特徴部として扱い、クレーム文を図式化して表示できるようになり、クレーム文の構成が明確に理解できるようになる。

【0015】

また、請求項5に記載の特許明細書デバッグツールは、前記第1および第2の技術要素特定手段のうち、前記指示語により特定された技術要素に関して前記クレーム文の中の参照元の技術要素を特定する参照元技術要素特定手段と、この参照元技術要素特定手段により特定された該参照元の技術要素と該指示語により特定された技術要素とを前記クレーム文

10

20

30

40

50

第一の図式化手段により表示された前記図式の上で線で結び、なおかつ、該指示語を該図式から削除した該図式を表示するクレーム文第二の図式化手段とを備えたことを特徴とするものである。従って、クレーム文第二の図式化手段により、参照元技術要素特定手段により定義された技術要素どうしの関係が明確に理解できるようになる。

【 0 0 1 6 】

また、請求項 6 に記載の特許明細書デバッグツールは、前記第 1 および第 2 の技術要素特定手段の中に、前記指示語の後に動詞または動詞と同じ働きを持つ動詞句または用言が続き、その後に名詞または名詞と同じ働きをもつ名詞句が続く場合は、該動詞または該動詞句または該用言を除いた該名詞または名詞と同じ働きをもつ名詞句を技術要素として特定する機能を含むことを特徴とするものであり、従って、指示語の後の動詞句を除かないと技術要素を抽出できない場合でも技術要素を抽出できるようになる。

10

【 0 0 1 7 】

また、請求項 7 に記載の特許明細書デバッグツールは、前記第 1 および第 2 の技術要素特定手段の中に、前記クレーム文の中の並立の意味を有する接続詞を格助詞「と」に変換する機能を含むことを特徴とするものであり、従って、この方法により接続詞「または」等により並立された語句を技術要素として抽出できるようになる。

【 0 0 1 8 】

また、請求項 8 に記載の特許明細書デバッグツールは、前記第 1 および第 2 の技術要素特定手段の中で、格助詞「と」に前置された語句として特定された技術要素を「と」による技術要素とし、この格助詞「と」の後に続く形態素の品詞と読点または符号とにより前記「と」による技術要素の属性を定め、該属性が前記クレーム文の中で先行した「と」による技術要素の属性から変化した状態を遷移状態として、該遷移状態により、前記「と」による技術要素どうしの従属関係を演算式またはこの演算式に準じた方式で決定する方法を前記 1 次の技術要素特定手段に含むことを特徴とするものである。従って、前記「と」による技術要素どうしの従属関係が演算式のみで決定できるようになるので、プログラムの処理の負担が軽くなるばかりでなく、多種多様な形態をとる日本語のクレーム文に対して柔軟な処理が可能になる。

20

【 0 0 1 9 】

また、請求項 9 に記載の特許明細書デバッグツールは、前記 1 次の技術要素特定手段のうち、前記「と」による技術要素どうしの従属関係を演算式で決定する方法の中に、前記「と」による技術要素に対し位置づけの値を定義する演算式を含み、該位置づけの値を定義する演算式は、前記遷移状態に基づいて定義され、また、特別に定義された技術要素については位置づけの値を補正できる機能を含み、前記位置づけの値の大小の比較により前記 1 次の技術要素を特定する機能を含むことを特徴とするものであり、従って、前記 1 次の技術要素を演算式のみで決定できるようになるので、プログラムの処理の負担が軽くなるばかりでなく、多種多様な形態をとる日本語のクレーム文に対して柔軟な処理が可能になる。

30

【 0 0 2 0 】

また、請求項 10 に記載の特許明細書デバッグツールは、前記「と」による技術要素を特定する方法の中に、前記格助詞「と」が並立の意味を持たない場合は、該格助詞「と」に前置される語句を「と」による技術要素として抽出しない機能を含むことを特徴とするものであり、従ってこの方法により並立を意味しない格助詞「と」に前置される語句を技術要素として抽出されることを防止できるようになる。

40

【 0 0 2 1 】

また、請求項 11 に記載の特許明細書デバッグツールは、前記第 1 および第 2 の技術要素特定手段により特定された技術要素の最後尾の形態素が助詞の場合は、該助詞を除いたものを技術要素として特定する機能を有することを特徴とするものであり、従ってこの方法により最後尾に助詞となる技術要素を抽出しなくなる。

【 0 0 2 2 】

また、請求項 12 に記載の特許明細書デバッグツールは、前記前提部に属する前記技術要

50

素のうち、他で参照されない技術要素があれば、第一の警告を出力することを特徴とするものである。従って、前提部に他で引用されない技術要素があるのに気づき、クレーム文の再検討が容易にできるようになる。

【 0 0 2 3 】

また、請求項 1 3 に記載の特許明細書デバッグツールは、前記「と」による技術要素の属性の並びから判断して、格助詞「と」の使い方として、文法に反する場合は、第二の警告を出力することを特徴とするものである。従って、格助詞「と」の誤った使い方に気づき、クレーム文の再検討が容易にできるようになる。

【 0 0 2 4 】

また、請求項 1 4 に記載の特許明細書デバッグツールは、前記主題特定手段により特定された前記前提部の主題と前記の特徴部の主題の比較を行い、それぞれの主題が異なる場合は、「クレーム文の前提部と特徴部との主題が異なる」ことを意味する文章または図形を出力することを特徴とするものであり、従って、前提部と特徴部との主題を同じものにすべきかの検討が容易にできるようになる。

【 0 0 2 5 】

また、請求項 1 5 に記載の特許明細書デバッグツールは、前記クレーム文の中の前記技術要素の前記上位概念の用語を前記辞書ファイルの中の上位概念用語辞典から検出して出力する技術要素上位概念検出手段を備えたことを特徴とするものである。従って、技術要素を上位概念に置き換えるべきか等のクレーム文の再検討が容易にできるようになる。

【 0 0 2 6 】

また、請求項 1 6 に記載の特許明細書デバッグツールは、前記特徴部の中の複数の技術要素のいずれかと前記前提部の中の複数の技術要素の前記上位概念の用語のいずれかとを比較して同じ語句があれば、第三の警告を出力することを特徴とするものである。従って、前提部に、特徴部より下位概念の語句が使われているのに気づき、クレーム文の再検討が容易にできるようになる。

【 0 0 2 7 】

さらに、請求項 1 7 に記載のコンピュータプログラムを汎用のパーソナルコンピュータ等にインストールすることにより、汎用のパーソナルコンピュータを用いて、容易に特許明細書のクレーム文を図式化表示できるようになるため、クレーム文の構成やクレーム文の中の技術要素の関連を明確に理解できるようになり、また、クレーム文の再検討が容易にできるようになる。

【 0 0 2 8 】

【発明の実施の形態】

以下、この発明の特許明細書デバッグツールをコンピュータプログラムとして具体化した実施の形態について説明する。

【 0 0 2 9 】

図 1 は汎用のパーソナルコンピュータのブロック図であり、この汎用のパーソナルコンピュータにコンピュータプログラムとしての特許明細書デバッグツールをインストールして、本発明の機能を具体化させるものである。

【 0 0 3 0 】

中央演算装置である CPU 1 と、RAM または ROM で構成される内部メモリ 2 と、入出力インターフェース 3 とはバスライン 1 1 で接続され、入出力インターフェース 3 には、データ通信端末 4 とデータ出力端末 5 とデータ入力端末 6 と外部メモリドライブ端末 7 とが接続されている。

【 0 0 3 1 】

データ通信端末 4 としては、外部サーバと接続するための TA (ターミナルアダプタ) や通信モデム等があり、データ出力端末 5 としては、モニター画面やプリンター等があり、データ入力端末 6 としては、キーボードやマウス等があり、外部メモリドライブ端末 7 としては、記憶媒体の HD (ハードディスク) 1 2 や CD (コンパクトディスク) や FD (フロッピディスク) 等をドライブする HD ドライバーや CD ドライバーや FD ドライバー

10

20

30

40

50

等がある。

【0032】

OS（オペレーティングシステム）8とOS8の上で動く特許明細書デバッグツール9は外部メモリ（CPU1にバスライン11で直結されていないメモリ）ドライブ端末7の一つであるHDドライバーによりアクセスされるHD（ハードディスク）12の中にインストールされていて、OS8が内部メモリに書き込まれた後、特許明細書デバッグツール9の一部または全部が内部メモリ2に書き込み、図14や図5のフローチャート等で示すプログラムを実行する形態をとる。

【0033】

特許明細書デバッグツール9で解析対象とする図3に示すようなクレーム文は、データ入力端末6のキーボードを介して入力されたり、記憶媒体に記憶されたものを外部メモリドライブ端末7を介して入力されたり、データ通信端末4を介して外部サーバから入力され、特許明細書データファイル10に格納される。

10

【0034】

また、データ入力端末6のキーボードやマウス等は、特許明細書デバッグツール9を操作するために使われ、データの加工結果や図4に示すような技術要素相関図等は、データ出力端末5のモニター画面やプリンタ等に出力される。

【0035】

図2は、特許明細書デバッグツール9と特許明細書データファイル10との構成図である。特許明細書デバッグツール9はデバッグを実行する実行ファイル20とクレーム文解析用辞書ファイル29から構成され、実行ファイル20は、デバッグ対象のクレーム文を形態素解析（クレーム文をクレーム文解析用辞書ファイル29に内蔵された形態素解析用辞書30で定義された文字列に分解し、品詞等を解析）するための形態素解析部21と、形態素解析の結果とクレーム文解析用辞書ファイル29に内蔵された前提部判定ファイル32とにより、クレーム文の特徴部と前提部とを抽出し、特徴部と前提部との主題を抽出する主題抽出部22と、格助詞「と」の前に置かれた名詞句をクレーム文の技術要素として抽出する「と」による技術要素抽出部23と、主題抽出部22で抽出した主題と「と」による技術要素抽出部23で抽出した技術要素の関係を木構造等で出力する木構造出力部24と、「上記」、「前記」、「この」、「これらの」、「該」、「同」などの指示語により特定できる技術要素を抽出する指示語による技術要素抽出部25と、技術要素相互の関連を示す図を出力する技術要素相関図出力部26と、技術要素の上位概念の検出を行う上位概念検索部27と、クレーム文のエラー等を出力するエラーメッセージ出力部28とに分かれていて、それぞれコンピュータプログラムとして機能する。

20

30

【0036】

クレーム文解析用辞書ファイル29は、形態素解析用辞書30と上位概念辞書31と前提部判定用語ファイル32と前処理ファイル33と指示語ファイル34とポジション特別処理ファイル35とから構成される。

【0037】

形態素解析用辞書30は、日本語の単語の辞書で単語の品詞や活用変化や語尾変化や単語と単語との関係から品詞を確定させるための文法等が格納されていて、形態素解析部21でクレーム文を形態素解析するために用いられる。この形態素解析用辞書30は、一般の日本文の自動解析に用いられる形態素辞書であっても良いが、この形態素辞書にクレーム文の技術要素として用いられた用語であれば、たとえば、「エンジン電子制御装置」のように本来4つの形態素「エンジン」、「電子」、「制御」、「装置」に分解される用語であっても、「エンジン電子制御装置」を1つの名詞相当の用語として追加登録できる「クレーム文解析用辞書としての形態素解析用辞書30」であっても良い。

40

【0038】

上位概念辞書31は、語句に対する上位概念の語句が格納されていて、例えば、本発明でクレーム文の技術要素として「ばね」や「ゴム」といった語句が抽出された場合、それらの語句の上位概念として「弾性体」という語句が上位概念辞書31に格納されていて、そ

50



の上位概念の語句が自動的に検索できるようにするために用いられる。

【0039】

前提部判定用語ファイル32は、このファイルに登録されている語句、例えば「において、」や「であって、」等の語句で区切られたクレーム文のうち、文頭からこれらの語句までを前提部とし、これに続き文末までを特徴部と判定するために用いられる。

【0040】

前処理用ファイル33は、技術要素を抽出するためにクレーム文の中に、このファイルに登録されている語句を指定された語句に変換するもので、例えば、「及び」という接続詞は格助詞「と」に変換するように定められていれば、クレーム文の中の「及び」を「と」に変換するように用いられる。

10

【0041】

指示語ファイル34は、このファイルに登録されている語句で、クレーム文内の技術要素の相互の関係を示すために用いる指示語、例えば、図15に示すような「上記」、「前記」、「この」、「同」、「該」等などのものを指し示す修飾語で、これらの指示語に続き、直後の動詞句のような用言を除いた語句と、他の部分で一致する語句を検出し、その語句をあらたな技術要素と決める時に使用し、同ファイルで、それぞれの登録されている語句に定義されている「参照距離」の数値である「+2」、「+1」、「-1」、「-2」により、同ファイルに登録されている語句により決められた技術要素の参照対象とするもう一方の技術要素を決めるもので、「-2」であれば、先に出た技術要素のうち、一番先に出た技術要素を参照対象とし、「-1」であれば、先に出た技術要素のうち、一番最後に出た技術要素を参照対象とし、「+1」であれば、後から出る技術要素のうち、一番先に出る技術要素を参照対象とし、「+2」であれば、後から出る技術要素のうち、一番最後に出る技術要素を参照対象とするように用いられる。

20

【0042】

ポジション特別処理ファイル35は、このファイルに登録されている語句、例えば、「特徴」等が技術要素として検出された場合は、後述する位置づけPSの値を定められた方法に従い操作するために用いられる。

【0043】

図2の特許明細書データファイル10の構成部品として、クレームファイル50には、デバッグの対象とする特許明細書の請求項の原文がクレーム文として格納されていて、アドレス付き技術要素ファイル51は、図7に示すように、デバッグ対象のクレーム文が、形態素と技術要素に分解された表形式のファイルとして格納され、形態素と技術要素を「クレーム要素」の列に記載し、その左列を「アドレス」の項目としてアドレスを記し、「クレーム要素」の右の列には、「語義」、「属性TX」、「位置づけPS」、「参照技術要素のアドレス」、「上位概念用語」を項目とする列が設けられ、「語義」の列には各要素の品詞や技術要素や主題や特徴部、前提部のそれぞれの文頭、文末等を表す語句を記し、「属性TX」の列には各要素の後述する属性TXが記され、「位置づけPS」の列には技術要素の後述する位置づけPSの値が記され、「参照技術要素のアドレス」の列には、指示語により特定された技術要素の参照対象とする技術要素のアドレスが記され、「上位概念用語」の列には、各技術要素の上位概念用語が記され、技術要素相関図ファイル52には、図4のような木構造と木構造上で技術要素の相互の関係を図示するためのデータが格納されている。

30

40

【0044】

図5は、デバッグ対象のクレーム文を入力した後、図4のような技術要素相関図を出力するまでの処理手順を示したフローチャートである。

【0045】

具体的には、図3に示すクレーム文「AとBとに保持されたCと、前記Cを支持するDと、EがHに変位した時、前記Dを駆動するIとを備えたJにおいて、KがLした時に、前記IがM以下となるNと、前記KがOした時に、前記Nが、PとQとRとを牽引してSにするTとを備えたことを特徴とするJ。」を例にして説明する。なおA～Tの部分には、

50

任意の普通名詞等で構成される技術的な用語が入るものとする。

【 0 0 4 6 】

図5の処理S1では、デバッグ対象のクレーム文を形態素解析部21において、形態素解析用辞書30を参照して形態素解析(文を形態素解析用辞書30に登録された文字列に分解し、品詞等を解析する)の処理が行われる。また、形態素解析用辞書30に登録されていない未知語は、すべて名詞句として処理する。

【 0 0 4 7 】

次に、処理S2において、形態素解析の解析の結果、前処理ファイル33に登録されている語句がデバッグ対象のクレーム文の中にあれば、前処理ファイル33で指定された語句に変換する。例えば、接続詞の「および」または「及び」が前処理ファイル33に登録されていて、接続詞の「および」または「及び」を格助詞「と」に変換するように指定されていれば、そのように変換する。これは、文法上、本来は「東京と大阪とを結ぶ新幹線」と記すべきところ、クレーム文で「東京及び大阪とを結ぶ新幹線」という表し方をすることがあり、この表し方のため後述する方法で技術要素を抽出できないことがあるので、処理S2のような特別な処置を設定する必要がある。また、「東京と大阪を結ぶ新幹線」と表す場合もあるが、これは並立する語句には、すべて格助詞「と」を付けるという本来の形に反しているため、この場合には後述する方法で大阪を技術要素として抽出できなくても良いものとする。

【 0 0 4 8 】

次に、処理S3において、クレーム文の全文にアドレスを付ける。アドレスの値はクレーム文頭からの文字数で良い。

【 0 0 4 9 】

図6に処理S3までの形態素解析の結果の一部を示す。左の列から「アドレス」、「形態素相当」、「品詞相当」である。「品詞相当」の中で使われる「名詞句」は、全体として一つの名詞と同じはたらきをするものや、形態素解析用辞書30に登録されている単語ごとに「名詞句」定義されている場合も「名詞句」として表す。例えば、「エンジン電子制御装置」は本来の形態素としては、「エンジン(普通名詞)」、「電子(普通名詞)」、「制御(普通名詞)」、「装置(普通名詞)」に分解され、この4つの語句が重なってきた語句であるが、形態素解析用辞書30に「品詞相当」が「名詞句」として「エンジン電子制御装置」が形態素相当として登録されていれば、「最長一致法」(特開平11-25089号参照)を用いて形態素解析をした場合、「エンジン電子制御装置」が「名詞句」として解析される。「品詞相当」の中で使われる「動詞句」は、全体として一つの動詞と同じはたらきをするものや、形態素解析用辞書30に登録されている単語ごとに「動詞句」と定義されている場合も「動詞句」として表す。ここで、図6の列の項目名を「形態素相当」としているのは、本来の形態素と異なるものが含まれ、「品詞相当」としているのは、本来の文法の品詞と異なるものが含まれるためである。

【 0 0 5 0 】

次に、処理S4において、前提部または特徴部の主題抽出を行う。この処理のより詳細な処理手順を図8に示す。

【 0 0 5 1 】

図5の処理3で、クレーム文の全文にクレーム文頭からの文字数でアドレスを付け、図6に示すように形態素相当の各語句の先頭の文字のアドレスを各形態素相当の語句のアドレスとして表した場合、図8の処理S401で、アドレスの最小値(図6の例では1)がクレーム文頭のアドレス値、アドレス最大値(図7の例では121)がクレーム文末のアドレス値であり、なおかつ特徴部文末のアドレスとすることができる。

【 0 0 5 2 】

次に、処理S402で、検索範囲をクレーム全文に設定し、次に、判定処理S403で、クレーム全文の中に、前提部判定用語ファイル32に登録されている語句(例えば、「において」、「であって、」のようにクレーム文で前提部の区切りとして常用される語句)が有るかどうかを判定する。

## 【 0 0 5 3 】

判定処理 S 4 0 3 の判定結果が “ 有る ” の場合、判定処理 S 4 0 4 で、前提部判定用語ファイル 3 2 に登録されている語句が 1 項のクレーム文中に 2 つ以上有るかどうかを判定し、2 つ以上あれば、処理 S 4 0 5 において、“前提部判定用語が 2 回以上使われているエラー”として、エラー処理をして、処理 S 4 0 6 において、全処理を終了し、判定処理 S 4 0 4 の判定結果が “ 2 つ以上無い ” 場合には、クレーム文中に前提部が存在するので、次に、処理 S 4 0 7 で、クレーム文の文頭のアドレス（図 6、図 7 の例では 1）を前提部文頭のアドレスとし、次に、処理 S 4 0 8 において、処理 S 4 0 3 で検出した語句の先頭のアドレス（図 7 の例では 5 0）を前提部の文末のアドレスとし、次に、処理 S 4 0 9 において、前提部文末の次の形態素相当のアドレス（図 7 の例では、“K” が次の形態素相当になり、そのアドレスは 5 5）を特徴部の文頭のアドレスとし、次に、処理 S 4 1 0 において、前提部文末（図 7 の例では、アドレス 5 0 の “ において、 ”）直前の動詞句（図 7 の例では、アドレス 4 6 の “ 備えた ”）を含まない独立した最長の名詞句（図 7 の例では、アドレス 4 9 の “ J ”）を「前提部主題」とし、「前提部主題」の先頭の文字のアドレスを前提部主題のアドレスとし、前提部主題の属性 TX を 1 として記録し、処理 S 4 1 3 に進む。

10

## 【 0 0 5 4 】

ここで、属性 TX とは、図 5 の処理 S 4 以下で抽出される「主題」、「技術要素」等のクレーム要素について、その属性を判別するための数値で、図 9 に示す表に従い定義され、属性 TX = 0 は、後述する特徴部主題であり、属性 TX = 1 は、前提部主題であり、属性 TX = 3 は、後述する技術要素であって、後ろに読点が続いている格助詞「と」の前の動詞句を含まない独立した最長の名詞句であり、属性 TX = 4 は、後述する技術要素であって、後ろに格助詞が続いている格助詞「と」の前の動詞句を含まない独立した最長の名詞句であり、属性 TX = 5 は、後述する技術要素であり、後ろに読点や格助詞を伴わない格助詞「と」の前の動詞句を含まない独立した最長の名詞句であり、属性 TX = 7 は、後述する技術要素であって、後述する指示語により指定された最長の語句として、それぞれ定義される。

20

## 【 0 0 5 5 】

また、上記の “ 独立した最長の名詞句 ” とは、用言を含まない “ 「名詞句 1」 + 格助詞 + 「名詞句 2」 ” という語句な場合に、格助詞が所有・限定を意味する「の」が使われている場合は、「名詞句 1」が所有者で、「名詞句 2」は「名詞句 1」の従属者の立場になり独立していないため、「名詞句 2」だけでは「前提部主題」とならないことを意味し、格助詞が前の語句に主語格を与える「が」や目的格を与える「を」などである場合には、「名詞句 2」は「名詞句 1」に対して独立した立場であるため、「名詞句 2」だけで「前提部主題」になることを意味する。例えば「エンジンの冷却装置」という語句では、所有・限定を意味する格助詞「の」があるため独立していない名詞句「冷却装置」だけでは「前提部主題」にならないことを意味している。また、“ 「名詞句 1」 + 格助詞 + 「名詞句 2」 ” という語句の格助詞として、並立を意味する「と」が使われている場合も、「名詞句 2」と「名詞句 1」は 2 つで 1 組と考えられ、「名詞句 2」が独立した最長の名詞句とは考えられない。

30

40

## 【 0 0 5 6 】

図 8 の判定処理 S 4 0 3 の結果が “ （前提部判定用語が）無い ” の場合は、処理 S 4 1 1 において、クレーム文頭を特徴部文頭のアドレスとし、次に、処理 S 4 1 2 において、前提部に該当するアドレスが無いという意味で、前提部文頭のアドレスおよび前提部文末のアドレスを 0 に設定し、処理 S 4 1 3 に進む。構成要素列挙型のクレーム文のように前提部が無いクレーム文は、この処理で全文が特徴部として以下処理される。

## 【 0 0 5 7 】

処理 S 4 1 3 において、特徴部文末（図 7 の例ではアドレス 1 2 1 の “ 。 ”）直前の動詞句（図 7 の例ではアドレス 1 1 8 の “ する ”）を含まない独立した最長の名詞句（図 7 の例ではアドレス 1 2 0 の “ J ”）を「特徴部主題」とし、「特徴部主題」の先頭の文字の

50

アドレスを「特徴部主題」のアドレス（図7の例では120）とし、特徴部の属性TXを図9の定義に従い0として、判定処理S414に進む。

【0058】

判定処理S414において、特徴部主題と前提部主題の比較を行い、両者が等しければ処理S4を終了し、等しくなければ、処理S415で、“特徴部主題と前提部主題が等しくない”というメッセージを出力するための処理を行ってから処理S4を終了する。

【0059】

次に、図5の処理S5において、格助詞「と」で特定できる技術要素の検出を行う。この処理のより詳細な処理手順を図10に示す。

【0060】

処理S501で検索範囲をクレーム全文に設定し、その検索範囲において、処理S502で、格助詞「と」の手前の動詞句を含まない独立した最長の名詞句を*i*番目の技術要素N<sub>*i*</sub>として抽出し、技術要素N<sub>*i*</sub>の先頭の文字のアドレスを技術要素N<sub>*i*</sub>のアドレスA<sub>*i*</sub>として、図12のように記録する。

【0061】

図12は、図3に示すクレーム文を解析して抽出した格助詞「と」の前の技術要素である。ここで*i* = 6の“M以下”は、「IがM以下となる」において、格助詞「と」に先行する語句は“IがM以下”となるが、“I”は主語格を示す格助詞「が」により分離されるため、“独立した最長の名詞句”として“M以下”が抽出されたことになり、*i* = 12の“特徴”は、「ことを特徴とする」において、格助詞「と」に先行する語句は“ことを特徴”になるが、目的格を示す格助詞「を」により、“ことを”と“特徴”とが分離されるため、独立した最長の名詞句として“特徴”が抽出されたことになる。“M以下”と“特徴”は、結果の意味を示す格助詞「と」についていたもので、その他の並立の意味を示す格助詞「と」についたものと性格が異なるが、処理S502を実行すると技術要素として抽出される。

【0062】

“M以下”や“特徴”がクレーム文の技術要素として抽出されるのが不都合であれば、処理S1の形態素解析の他に、格助詞「と」の意味を解析する処理を追加し、“結果の意味を示す格助詞「と」に前置される語句は技術要素としない”、または、“並立の意味を示す格助詞「と」に前置される語句を技術要素とする”というルールを追加設定するようにしてもよい。

【0063】

次に、処理S503において、技術要素N<sub>*i*</sub>の属性TX<sub>*i*</sub>を図9の属性TXを定義する表に従い記録する。図12に図3のクレーム文の技術要素N<sub>*i*</sub>の属性TX<sub>*i*</sub>を示す。

*i* = 1のアドレスA<sub>1</sub> = 1の技術要素N<sub>1</sub>の“ A ”は、格助詞「と」のみに前置されているので、TX<sub>1</sub> = 5となり、

*i* = 2のアドレスA<sub>2</sub> = 3の技術要素N<sub>2</sub>の“ B ”は、“格助詞「と」 + 格助詞「に」”に前置されているので、TX<sub>2</sub> = 4となり、

*i* = 3のアドレスA<sub>3</sub> = 11の技術要素N<sub>3</sub>の“ C ”は、“格助詞「と」 + 読点”に前置されているので、TX<sub>3</sub> = 3となる。

以下同様の方法で、

技術要素N<sub>4</sub>の“ D ”は、TX<sub>4</sub> = 3となり、

技術要素N<sub>5</sub>の“ I ”は、TX<sub>5</sub> = 4となり、

技術要素N<sub>6</sub>の“ M以下 ”は、TX<sub>6</sub> = 5となり、

技術要素N<sub>7</sub>の“ N ”は、TX<sub>7</sub> = 3となり、

技術要素N<sub>8</sub>の“ P ”は、TX<sub>8</sub> = 5となり、

技術要素N<sub>9</sub>の“ Q ”は、TX<sub>9</sub> = 5となり、

技術要素N<sub>10</sub>の“ R ”は、TX<sub>10</sub> = 4となり、

技術要素N<sub>11</sub>の“ T ”は、TX<sub>11</sub> = 4となり、

技術要素N<sub>12</sub>の“ 特徴 ”は、TX<sub>12</sub> = 5となる。

10

20

30

40

50

## 【 0 0 6 4 】

次に、処理 S 5 0 4 において、位置づけ P S の初期値 (  $i = 1$  の技術要素 N 1 の位置づけ P S の値 ) である P S 1 を 0 に設定する。ここで、位置づけ P S とは、技術要素 N  $i$  の次元の高低を判定する数値で、より数値が大きければ、より主要な技術要素となり、より数値が小さければ、より枝葉な技術要素となる。前提部または特徴部において、最も大きな数値の位置づけ P S の値をもつ技術要素が 1 次の技術要素となり、その P S の値より 1 小さい P S の値を持つ技術要素が 2 次の技術要素となる。

## 【 0 0 6 5 】

次に、処理 S 5 0 5 において、クレーム文の文頭から順次抽出された順に技術要素 N  $i$  の属性 T X  $i$  の値を 1 つ前の属性 T X  $i - 1$  からの遷移の状態により、図 1 1 に示す表の定義に基づき技術要素 N  $i$  の位置づけ P S  $i$  の値を設定していく処理を行う。

10

## 【 0 0 6 6 】

図 3 のクレーム文の場合、図 1 2 に示すように  $i = 1$  の技術要素 N 1 の “ A ” の位置づけ P S 1 は初期設定された 0 のままであるが、 $i = 2$  の技術要素 N 2 の “ B ” の位置づけ P S 2 は、属性 T X が T X 1 = 5 から T X 2 = 4 に遷移したため、図 1 1 の表の定義の “ P S = P S ” に従い、P S 1 と同じ 0 となる。

## 【 0 0 6 7 】

図 1 1 において、位置づけ P S の値が 1 増えるのは、T X が 「 4 から 3 に 」、 「 4 から 4 に 」、 「 4 から 5 に 」、 「 5 から 3 に 」 遷移した場合であり、位置づけ P S の値が 1 減るのは、T X が 「 3 から 5 に 」 遷移した場合のみであり、位置づけ P S の値が増減しないのは、T X が 「 3 から 3 に 」、 「 3 から 4 に 」、 「 5 から 4 に 」、 「 5 から 5 に 」 遷移した場合である。なお、図 1 1 に示す定義は、日本語の文例の検証をもとに定めたものであるが、より煩雑な文章に応用するためにはより高度な定義を必要とする。

20

## 【 0 0 6 8 】

図 1 2 に示すように、 $i = 3$  の技術要素 N 3 の “ C ” は属性 T X 3 が 「 T X 2 の 4 から 3 に 」 遷移したため、位置づけ P S は 1 つ増え、P S 3 は 1 になる。すなわち、位置づけ P S が 1 となった技術要素 N 3 の “ C ” は、位置づけ P S が 0 の技術要素 N 1 の “ A ” や技術要素 N 2 の “ B ” より主要な技術要素の位置づけになったと考えられる。

## 【 0 0 6 9 】

以下同様にして、 $i = 7$  の技術要素 N 7 の “ N ” は属性 T X 7 が 「 4 から 3 に 」 遷移したため、位置づけ P S は 1 増え P S 7 は 3 になり、 $i = 8$  の技術要素 N 8 の “ P ” は属性 T X 8 が 「 3 から 5 に 」 遷移したため、位置づけ P S は 1 減り P S 8 は 2 になる。すなわち、技術要素 N 8 の “ P ” は、技術要素 N 7 の “ N ” より、位置づけ P S が 1 つ小さくなったので、1 レベル主要な技術要素では無くなったことになる。

30

## 【 0 0 7 0 】

また、 $i = 1 1$  の技術要素 N 1 1 の “ T ” は属性 T X が 「 4 から 4 に 」 遷移したため、位置づけ P S は 1 増え P S 1 1 は 3 になり、 $i = 1 2$  の技術要素 N 1 2 の “ 特徴 ” は、属性 T X が 「 4 から 5 に 」 遷移するため、位置づけ P S は 1 増え、この段階では、P S 1 2 は 4 になり、特徴部の中で最も大きな位置づけ P S となる。

## 【 0 0 7 1 】

次に、処理 S 5 0 6 で、全技術要素 N  $i$  を検索対象とし、判定処理 S 5 0 7 において、各技術要素 N  $i$  が 「 ポジション特別処理ファイル 」 3 5 に登録されている語句と一致するかどうかを判定し、NO、すなわち、一致するものが無ければ処理 S 5 0 9 に進むが、YES、すなわち、一致するものがあれば処理 S 5 0 8 で、その一致する技術要素 N  $i$  の位置づけ P S  $i$  の値から 1 0 0 減じたものを P S  $i$  に設定しなおしてから、処理 S 5 0 9 へ進む。「ポジション特別処理ファイル」に “ 特徴 ” が登録されている場合は、図 1 2 の例のように、 $i = 1 2$  の技術要素 N 1 2 は、“特徴”であるため、処理 S 5 0 8 において、位置づけ P S 1 2 の値は 4 であったが、1 0 0 減じられ - 9 6 となる。

40

## 【 0 0 7 2 】

次に、処理 S 5 0 9 において、アドレスが特徴部に属し、位置づけ P S  $i$  が最大となる技

50

術要素 N 1 を特徴部主題に関する “ 1 次の技術要素 ” として抽出し、処理 S 5 1 0 において、アドレスが前提部に属し、位置づけ P S i が最大となる技術要素 N 1 を前提部主題に関する “ 1 次の技術要素 ” として抽出する。

【 0 0 7 3 】

図 1 2 の例では、特徴部主題に関する 1 次の技術要素 N i は、位置づけ P S i が特徴部で最大の 3 となる i = 7 の “ N ” と i = 1 1 の “ T ” となり、前提部主題に関する 1 次の技術要素 N i は、位置づけ P S i が前提部で最大の 1 となる i = 3 の “ C ” と、i = 4 の “ D ” と、i = 5 の “ I ” となり、図 4 に示すように、それぞれ、特徴部主題の “ J ” と前提部主題の “ J ” とに直結する位置づけの技術要素となる。

【 0 0 7 4 】

ここで、位置づけ P S のもうひとつの特性として、それぞれの特徴部と前提部において、最大の P S 値にならない技術要素は、次に P S 値が 1 大きくなる技術要素に従属する特性がある。図 1 2 の前提部では、i = 1 の “ A ” と i = 2 の “ B ” との位置づけ P S は 0 であるが、次に P S が 1 大きくなる i = 3 の “ C ” に “ A ” と “ B ” とが同列で従属することになり、特徴部では i = 6 の “ M 以下 ” が i = 7 の “ N ” に従属することになり、i = 8 乃至 1 0 の “ P ”、“ Q ”、“ R ” は次に位置づけ P S が 1 大きくなる i = 1 1 の “ T ” に従属しているように、図 1 2 の位置づけ P S i の並び方を解析するだけで、どの技術要素がどの技術要素に従属するか容易に判断できるので、各助詞「と」で特定した技術要素を用いて木構造を作成する場合、図 1 2 の位置づけ P S の並びを解析して、より細かい木構造を作成することができるようになる。

【 0 0 7 5 】

図 1 0 の判定処理 S 5 0 7 と処理 S 5 0 8 で、“特徴”の位置づけ P S の値を特別処理として、1 0 0 減じたのは、“特徴”が特徴部主題に関する 1 次の技術要素として適切でないと考えたためである。

【 0 0 7 6 】

次に、処理 S 5 1 1 において、処理 S 5 0 1 から処理 S 5 1 0 までの処理の間で、文法上、格助詞「と」が正しく使われていないと考えられる場合は、警告を出力するためのエラー処理を行う。例えば、属性 T X が 3、4、5 となる格助詞「と」が 1 個だけ存在する場合や、属性 T X が 3 となる格助詞「と」の後に、属性 T X が 4 となる格助詞「と」が無い場合等は警告を出力する。

【 0 0 7 7 】

次に、図 5 の処理 S 6 において、前提部または特徴部の主題の説明部を抽出する。

【 0 0 7 8 】

処理 S 5 1 0 で前提部の 1 次の技術要素を抽出できた場合、前提部で最後の 1 次の技術要素の次の語句から、前提部主題の直前の語句までを前提部主題の説明部として抽出し、処理 S 5 1 0 で前提部の 1 次の技術要素を抽出できなかった場合は、前提部の文頭から前提部主題の直前の語句までを前提部主題の説明部として抽出する。図 3 のクレーム文の例では、図 7、図 1 2 に示すように、アドレス 4 3 の技術要素 “ I ” が最後の 1 次の技術要素となるため、次のアドレス 4 4 の「と」から前提部主題であるアドレス 4 9 の “ J ” の直前の語句「備えた」までの「とを備えた」を前提部主題の説明部として抽出する。

【 0 0 7 9 】

処理 S 5 0 9 で特徴部の 1 次の技術要素を抽出できた場合、特徴部で最後の 1 次の技術要素の次の語句から、特徴部主題の直前の語句までを特徴部主題の説明部として抽出し、処理 S 5 0 9 で特徴部の 1 次の技術要素を抽出できなかった場合は、特徴部の文頭から特徴部主題の直前の語句までを特徴部主題の説明部として抽出する。図 3 のクレーム文の例では、図 7、図 1 2 に示すように、アドレス 1 0 6 の技術要素 “ T ” が最後の 1 次の技術要素となるため、次のアドレス 1 0 7 の「と」から特徴部主題であるアドレス 1 2 0 の “ J ” の直前の語句「する」までの「とを備えたことを特徴とする」を特徴部主題の説明部として抽出する。

【 0 0 8 0 】

10

20

30

40

50

次に、図5の処理S7において、処理S5で抽出した格助詞「と」で特定できる技術要素を説明する技術要素説明部を抽出する。

【0081】

処理S7において、1次の技術要素 $N_i$ がある場合には、

(1) その1次の技術要素 $N_i$ の手前にある別の1次の技術要素 $N_i$ の直後の格助詞「と」と別の助詞または読点を除いた次の語句と、

(2) 前提部の文頭と、

(3) 特徴部の文頭とのいずれかのアドレスの値から、1次の技術要素 $N_i$ のアドレスの値を引いた値が、負の値で最も0に近くなる語句から、その1次の技術要素 $N_i$ の直前の語句までをその1次の技術要素説明部として抽出する。

10

【0082】

図3のクレーム文の例では、図12に示すように前提部の第1番目の1次の技術要素は、 $i = 3$ の“C”で、技術要素説明部は、前提部文頭から、“C”の直前までの「AとBとに保持された」になり、前提部の第2番目の1次の技術要素は、 $i = 4$ の“D”で、“D”には手前に別の1次の技術要素“C”があり、“C”の直後の格助詞「と」と読点を除いた語句から、“D”の直前の語句までの「前記Cを支持する」が技術要素説明部になる。同様にして、前提部第3番目の1次の技術要素“I”の技術要素説明部は「EがHに変位した時、前記Dを駆動する」であり、特徴部第1番目の1次の技術要素“N”の技術要素説明部は「KがLした時に、前記IがM以下となる」であり、特徴部第2番目の1次の技術要素“T”の技術要素説明部は「前記KがOした時に、前記Nが、PとQとRとを牽引してSにする」である。

20

【0083】

ここまでの説明で、クレーム文を特徴部と前提部の各主題と、各主題の説明部と、各主題に関する1次の技術要素と、各1次の技術要素の技術要素説明部に分解することができたので、図5の処理S8で木構造の出力処理を行う。

【0084】

図4は、図3のクレーム文を木構造で図示したものであるが、処理S8の段階では、まだ指示語の「前記」等は削除されていない。図4の例のように、前提部の主題と特徴部の主題とは分離して配置し、各主題の近くに、主題に関する1次の技術要素を分離して他と干渉しないように配置し、主題と1次の技術要素とをそれぞれ枠で囲み、主題の枠とその主題に関する1次の技術要素の枠とを線で結ぶ。次に、各主題の主題説明部をそれぞれの主題の近くに、1次の技術要素や枠と枠とを結ぶ線と干渉しないように配置する。次に、各1次の技術要素の技術要素説明部を関係する1次の技術要素の近くに他と干渉しないように配置し、その技術要素説明部を枠で囲み、その枠と、その技術要素説明部と関係する1次の技術要素を囲む枠とを線で結ぶ。また、1次の技術要素以外の技術要素も枠で囲み、クレーム文の木構造として出力する。

30

【0085】

次に、図5の処理9の指示語で特定できる2箇所の技術要素の抽出を行う。ここで、指示語とは、図15に示すように指示語ファイル34に登録されている語句のことで、例えば「上記」、「前記」、「この」、「同」、「該」などの語句があり、これらの語句に続き直後の動詞句を除いた語句と、他の部分で最も長い文で一致する語句を技術要素として特定するために用いられる語句である。また、指示語ファイル34では、登録されている語句ごとに「参照距離」の数値が「+2」、「+1」、「-1」、「-2」のいずれかの値で設定されていて、同ファイルに登録された指示語で特定された技術要素と、他の場所の同じ語句の技術要素のうち、参照対象とする技術要素をこの「参照距離」の数値により決めることとし、“「参照距離」の数値により参照技術要素を決めるルール”として、「参照距離」が「-2」であれば、一番先に出た同じ語句の技術要素を参照技術要素とし、

40

「参照距離」が「-1」であれば、直前の同じ語句の技術要素を参照技術要素とし、

「参照距離」が「+1」であれば、直後の同じ語句の技術要素を参照技術要素とし、

50

「参照距離」が「+2」であれば、一番後に出る同じ語句の技術要素を参照技術要素とする。例えば、指示語ファイル34には、「前記」が登録されていて、その「参照距離」が「-2」と設定されているので、「AはBで、前記AはCになるが、前記AはDにならない。」という文では、「Aは」は、指示語「前記」で特定された技術要素になるが、上記のルールに従うと、2番目と3番目の「Aは」の参照技術要素は、いずれも1番最初に出てくる「Aは」になり、3番目の「Aは」の参照技術要素が2番目の「Aは」にはならないことになる。

【0086】

また、他の部分で最も長い文で一致する語句を探す時に、指示語ファイル34に登録されている語句に続いた“直後の動詞句を除く”理由は、「電波の干渉量を測定するステップと、この測定した干渉量に基づき」の例のように、指示語「この」の後の動詞句「測定した」を除かないと、「干渉量」を技術要素として、抽出できないためである。

【0087】

図13に処理S9のより詳細な処理手順を示す。

処理手順の主要な点として、処理S903で、指示語ファイル34の中から「指示語」を選択し、判定処理S904で、クレーム文の中に該当する「指示語」があるかを検索し、有れば処理S910で「指示語」の後の直後の動詞句を除いた語句のうち、最初の句読点までの語句を取り出し「検索バッファ」にセットする。この時、この語句の先頭のアドレスASiと、指示語Sjiとを記録しておく。次に、判定処理S911において、クレーム文の中の他の場所に、「検索バッファ」に登録された語句と同じ語句が有るか検索し、無ければ、処理S914で「検索バッファ」の語尾1文字を減らし、再度、判定処理S911で「検索バッファ」と同じ語句がクレーム文の中に有るかの判定処理を行う。

【0088】

処理S914の後で、判定処理S915で「検索バッファ」の文字数が0かどうか判定し、0でなければ、判定処理S911を行うが、0であれば参照する技術要素が無いのに指示語を用いているので、処理S916でエラーとして記録して、判定処理S907に進む。

【0089】

判定処理S911において、他の場所に「検索バッファ」と同じ語句があるとなった場合には、処理S912で、この時の「検索バッファ」内の語句を「技術要素」として登録し、この「指示語」と新たに検出された「技術要素」の「アドレス」を記録し、次に処理S913において、他の場所の同じ語句の「技術要素」を全て検索して、「指示語」とその検索された「技術要素」の「アドレス」を記録し、判定処理S907に進む。

【0090】

判定処理S907で、指示語ファイル34に別の「指示語」が無いとなった場合には、処理S917に進み、処理S910で記録したアドレスASiごとの指示語Sjiに対する「参照距離」を指示語ファイル34より求め、“「参照距離」の数値により参照技術要素を決めるルール”に従い、アドレスASiの「技術要素」に対する「参照技術要素」を特定し、相互の参照技術要素のアドレス(複数参照されていれば複数のアドレス)を“アドレス付き技術要素ファイル51”に記録する。このアドレスは、図7の例であれば、「参照技術要素のアドレス」の列に記録し、参照元の技術要素であれば参照先の技術要素のアドレスが、参照先の技術要素であれば参照元の技術要素のアドレスが、それぞれ記録されていて、当該する技術要素が他のどの技術要素と関連しているか容易にわかるようにする。

【0091】

図3のクレーム文の例では、「前記」が指示語として図15の指示語ファイル34に参照距離が「-2」として登録されているので、「C」、「D」、「I」、「Kが」、「N」が指示語で特定される技術要素として抽出される。また、参照技術要素は、それぞれの技術要素に1つずつ存在している。技術要素として、「Kが」のような語句が抽出されて不都合であれば、“技術要素のうち最後の語句が助詞であればその助詞を削除する”等のル

10

20

30

40

50



ールを追加設定してもよい。

【0092】

次に、図5の処理S10において、指示語ファイル34に登録された指示語があれば、処理S8で出力した木構造の中から、その指示語を削除する。ここで指示語を削除するのは、図式を見やすくするためである。

【0093】

次に、図5の処理S11において、処理10までの木構造の中で、処理S9で抽出した技術要素を枠で囲み、参照元の技術要素の枠と参照先の技術要素の枠とを線で結び、両方の技術要素が相互に関係していることを明確に示す図式を出力する。図3のクレーム文の例では、図4のような技術要素相関図を出力できる。

10

【0094】

以上が、デバッグ対象のクレーム文を入力した後、技術要素相関図を出力するまでの具体的な手順である。

【0095】

次に、技術要素相関図を出力した後の処理として、図14の処理手順に示すように、処理21で、検討対象をアドレス付き技術要素ファイル51の前提部とし、次に判定処理S22において、処理S21で定めた検索対象の中の技術要素について、参照技術要素のアドレスが空欄のものがあるかを調べ、無ければ処理S24に進むが、有れば、処理S23で“それらの技術要素が参照されていない”ことを警告の意味で出力し、処理S24に進む。例えば、図4、図7の例では、前提部の「A」、「B」は技術要素として抽出されたが、他のどこからも参照されていないので、メッセージの出力対象となる。ここで、前提部の「A」と「B」とが、どこからも参照されていないことを出力する理由は、「A」と「B」との記載が請求の範囲を狭めていないか、あるいは、「A」と「B」という技術要素が将来使われなくなった場合に不利な記載にならないか、「A」と「B」とに関する記載を削除して請求項として成立するかどうか等を出願者が検討しやすくするためである。

20

【0096】

次に、S24で検討対象をクレーム全文の技術要素に設定し、処理25で上位概念辞書31を用いて、技術要素の上位概念の用語を2次の上位概念(上位概念のその上の上位概念)まで検索し、アドレス付き技術要素ファイル51に記録し、次に、判定処理S26で、前提部にある技術要素の上位概念用語と特徴部の技術要素とに同じ語句があるかを調べ、無ければ終了するが、有れば処理S27で「前提部に特徴部より下位概念の用語が有る」ことを出力し、全処理を終了する。処理S25で上位概念用語を図7のようにアドレス付き技術要素ファイル51に出力するのは、出願者が、より上位概念の用語に置き換えられて出願できるかを検討しやすくするためであり、処理S27のような出力するのは、特徴部より前提部の方が狭まった書き方になっていないか出願者が検討しやすくするためである。

30

【0097】

上記の実施の形態は、請求項1項分のクレーム文を図式化表示する方法について説明したが、請求項が2項以上のクレーム文を入力した場合も、句点または“請求項”等のキーワードを基準にして、文の切れ目を見つけ、同様の手順で請求項2項以上の技術要素相関図等を出力できる。

40

【0098】

上記の実施の形態では、格助詞「と」に前置される技術要素の属性TXを図9にも示すように3種類のみを設定して説明したが、この属性TXを4種類以上設定し、図11に示すような属性TXの遷移状態に対する位置づけPSを定義する方法もそれに対応するように変更してもよいし、図2に示すクレーム文解析用辞書ファイル29に6種類の辞書ファイルのみについて説明したが、その他の辞書ファイルを追加設定し、あらたに追加設定した辞書ファイルを用いて図式化表示の方法を変更してもよい。

【0099】

上記の実施の形態では、格助詞の「と」と指示語による技術要素を抽出する方法のみを説

50

明したが、「または」や「かつ」などの接続詞により技術要素を抽出してクレーム文の一部を集合理論の図で図式化表示をしたり、構文解析や意味解析の手法を取り入れて、クレーム文の一部を簡易なフローチャート図やシステムブロック図で図式化表示することもできる。

#### 【0100】

上記の実施の形態では、クレーム文解析用辞書ファイル29の形態素解析用辞書30や上位概念辞書31は、外部メモリのHDに格納して使用するよう説明したが、形態素解析用辞書30は、逐次公開される特許願公開文献の請求項に記載されている技術要素を抽出し追加登録したものをダウンロード等により使用できるようにしてもよいし、上位概念辞書31はインターネット上のWebサーバーに格納して使用するようにして、これを利用する特許技術者が辞書の内容を随時更新できるように構成してもよい。

10

#### 【0101】

上記の実施の形態は、汎用のパーソナルコンピュータで動かすコンピュータプログラムとして、実施の形態を説明したが、コンピュータであれば、汎用パーソナルコンピュータに限らず、CPUとプログラムを搭載した専用機として特許明細書デバッグツールを構成してもよい。

#### 【0102】

##### 【発明の効果】

以上説明したように、本発明によれば、多種多様な文構造を持つ日本語の特許明細書のクレーム文であっても、「において、」や「であって、」のようなクレーム文の前提部を示す常用の語句により、クレーム文の前提部と特徴部との分離と主題の抽出とができ、並立の意味を有する格助詞「と」に前置された語句により技術要素を抽出でき、格助詞「と」に続く助詞または読点の状態により技術要素の属性を数値化し、その1変数の属性の数値の遷移の状態から、技術要素の位置づけや従属関係を論理的に解析できるようになり、また、主題の説明部や技術要素の説明部を分離することにより、パターン照合法を用いずに、論理的手法で、多種多様な形態をとる日本語の特許明細書のクレーム文を木構造のような図式化表示ができるようになる。

20

#### 【0103】

また、「前記」や「上記」などの指示語に続く語句からクレーム文の技術要素を抽出し、指示語の種類により、参照元、参照先の関係を定義し、上記の木構造の上で技術要素相互の関係性を明確に図式化表示できるようになる。

30

#### 【0104】

クレーム文を木構造で表し、技術要素の関係を明確に図式化表示することにより、出願人の発明者や特許明細書作成者が意図したようにクレーム文が作成されているか、容易に確認できるようになるだけでなく、格助詞「と」に関する文法上の誤りを自動的に検出できる他、前提部に他で参照されない技術要素があれば、それを自動的に検出できるため、出願人の発明者や特許明細書作成者がクレーム文の見直しをすることが容易になる。

#### 【0105】

上記で検出したクレーム文の技術要素に関して、上位概念の語句を自動的に検出して出力し、前提部に特徴部より下位概念の語句があれば自動的に検出されるため、出願人の発明者や特許明細書作成者がクレーム文の技術要素の見直しを容易にできるようになる。

40

#### 【0106】

日本文の自動解析に用いる形態素辞書にクレーム文の技術要素として用いられた用語を名詞相当の用語として追加登録した形態素解析用辞書を用いることにより、特殊な専門用語や複雑な形態素から構成される技術用語がクレーム文内にあっても容易に技術要素を抽出できるようになる。

##### 【図面の簡単な説明】

##### 【図1】ブロック図

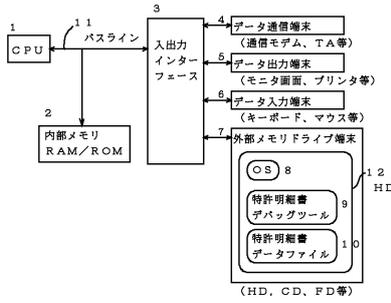
##### 【図2】特許明細書デバッグツールの構成図

##### 【図3】デバッグ対象クレーム文

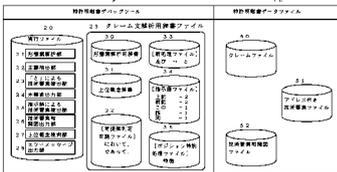
50

【図 4】技術要素相関図	
【図 5】フローチャート	
【図 6】S 3 の段階での形態素解析結果	
【図 7】アドレス付き技術要素ファイル 5 1 の例	
【図 8】前提部または特徴部の主題抽出のフローチャート	
【図 9】属性 T X 表	
【図 10】格助詞「と」で特定できる技術要素の抽出フローチャート	
【図 11】属性 T X の遷移状態に対する位置づけ P S を決める式	
【図 12】格助詞「と」の前の技術要素	
【図 13】指示語で特定できる技術要素の抽出フローチャート	10
【図 14】ゼネラルフローチャート	
【図 15】指示語ファイルの表	
【符号の説明】	
1 CPU	
2 内部メモリ	
3 入出力インターフェース	
4 データ通信端末	
5 データ出力端末	
6 データ入力端末	
7 外部メモリドライブ端末	20
8 OS	
9 特許明細書デバッグツール	
10 特許明細書データファイル	
11 バスライン	
12 HD	
20 実行ファイル	
21 形態素解析部	
22 主題部抽出部	
23 「と」による技術要素抽出部	
24 木構造出力部	30
25 指示語による技術要素抽出部	
26 技術要素相関図出力部	
27 上位概念検索部	
28 エラーメッセージ出力部	
29 クレーム文解析用辞書ファイル	
30 形態素解析用辞書	
31 上位概念辞書	
32 前提部判定用語ファイル	
33 前処理ファイル	
34 指示語ファイル	40
35 ポジション特別処理ファイル	
50 クレームファイル	
51 アドレス付き技術要素ファイル	
52 技術要素相関図ファイル	
図中、同一符号は同一または相当部分を示す。	

【図1】



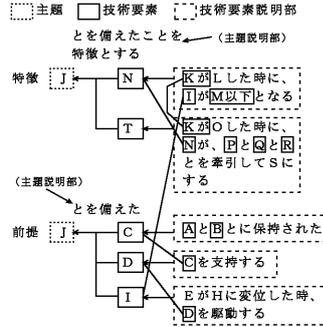
【図2】



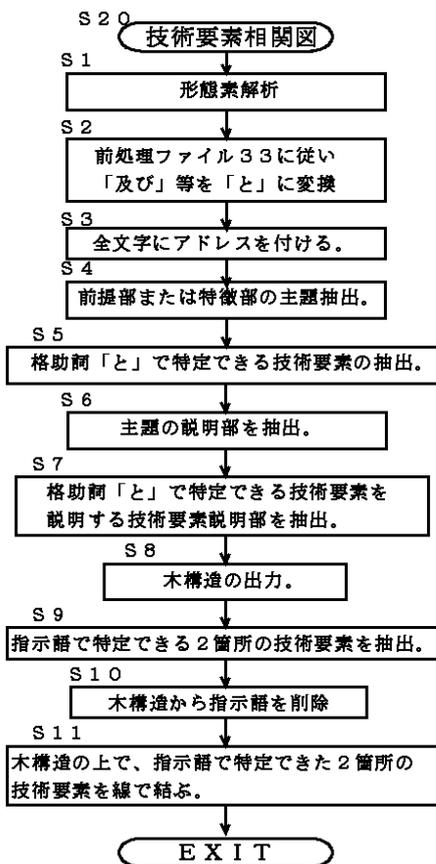
【図3】

【請求項1】  
AとBとに保持されたCと、前記Cを支持するDと、EがHに変位した時、前記Dを駆動するIとを備えたJにおいて、KがLした時に、前記IがM以下となるNと、前記KがOした時に、前記Nが、PとQとRとを牽引してSにするTとを備えたことを特徴とするJ。

【図4】



【図5】



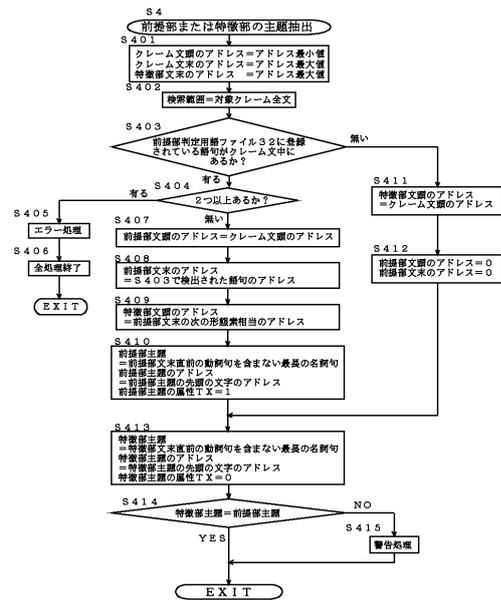
【図6】

アドレス	形態素相当	品詞相当
1	A	名詞句
2	と	格助詞
3	B	名詞句
4	と	格助詞
5	に	格助詞
6	保持された	動詞句
1 1	C	名詞句
1 2	と	格助詞
1 3	,	読点
1 4	前記	指示語
1 6	C	名詞句
1 7	を	格助詞
1 8	支持する	動詞句
2 2	D	名詞句
2 3	と	格助詞
2 4	,	読点

【図 7】

アドレス	クレーム要素	語彙	属性 TX	位置づけ PS	参照技術要素のアドレス	上位概念用語
1	A	技術要素	5	0	-	AA
		クレーム文頭				
		前読部文頭				
2	と	格助詞				
3	B	技術要素	4	0	-	bbb
4	と	格助詞				
5	に	格助詞				
-----						
43	I	1次の技術要素	4	1	55	ii
44	と	格助詞				
45	を	格助詞				
46	備えた	動詞句				
49	J	前読部主題	1			
50	において、	前読部文末				
55	K	技術要素	7		88	kkk
		特徴部文頭				
56	が	格助詞				
57	しした	動詞句				
60	時	普通名詞				
61	に	格助詞				
62	、	読点				
63	前記	指示語				
65	I	技術要素	7		43	ii
66	が	格助詞				
-----						
106	T	1次の技術要素	4	3	-	tTt
107	と	格助詞				
108	を	格助詞				
109	備えた	動詞句				
112	こと	名詞句				
114	を	格助詞				
115	特徴	技術要素	5	-96	-	-
117	と	格助詞				
118	する	動詞				
120	J	特徴部主題	0			
121	、	句点				
		特徴部文末				
		クレーム文末				

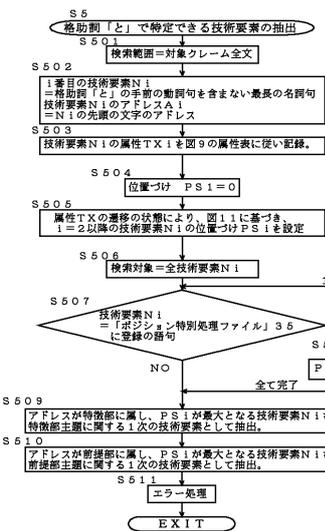
【図 8】



【図 9】

属性 TX	クレーム要素
0	特徴部主題
1	前読部主題
2	-
3	技術要素 (格助詞「と」+読点に前置される動詞句を含まない独立した最長の名詞句)
4	技術要素 (格助詞「と」+格助詞に前置される動詞句を含まない独立した最長の名詞句)
5	技術要素 (属性 TX = 3 または 4 以外の格助詞「と」に前置される動詞句を含まない独立した最長の名詞句)
6	-
7	技術要素 (指示語により指定された最長の語句)

【図 10】



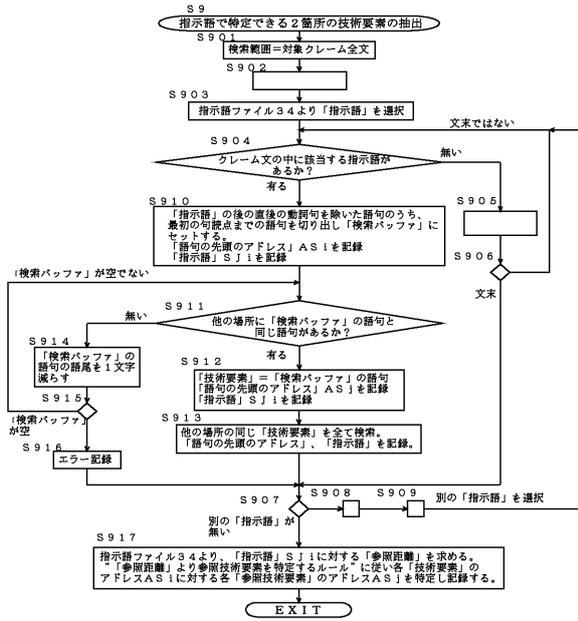
【図 12】

順番 i	技術要素 Ni のアドレス Ai	技術要素 Ni	属性 TX i	位置づけ PS i	注釈
	(1)				
1	1	A	5	0	前読部文頭
2	3	B	4	0	
3	1 1	C	3	1	1 次 の 技 術 要 素
4	2 2	D	3	1	1 次 の 技 術 要 素
5	4 3	I	4	1	1 次 の 技 術 要 素
	(5 5)				特 徴 部 文 頭
6	6 7	M 以下	5	2	
7	7 3	N	3	3	1 次 の 技 術 要 素
8	9 1	P	5	2	
9	9 3	Q	5	2	
10	9 5	R	5	2	
11	1 0 6	T	4	3	1 次 の 技 術 要 素
12	1 1 5	特徴	5	(4 →) - 96	

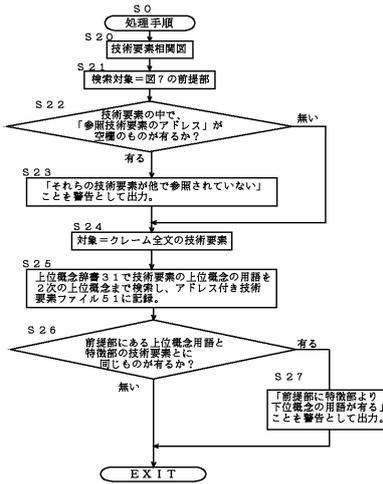
【図 11】

属性 TX 遷移状態	位置づけ PS を決める式
3 → 3	PS = PS
3 → 4	PS = PS
3 → 5	PS = PS - 1
4 → 3	PS = PS + 1
4 → 4	PS = PS + 1
4 → 5	PS = PS + 1
5 → 3	PS = PS + 1
5 → 4	PS = PS
5 → 5	PS = PS

【図13】



【図14】



【図15】

指示語	参照距離
上記	-2
前記	-2
この	-1
該	-1
同	-1
-----	-----

---

フロントページの続き

- (56)参考文献 新森昭宏他, 手がかり句を用いた特許請求項の修辭構造解析, 情報処理学会研究報告2002 - SLP - 41, 日本, 社団法人情報処理学会, 2002年 5月24日, Vol. 2002, No. 50, p. 1 - p. 8  
新森昭宏他, 手がかり句を用いた特許請求項の修辭構造解析, 情報処理学会研究報告2002 - NL - 149, 日本, 社団法人情報処理学会, 2002年 5月24日, Vol. 2002, No. 44, p. 65 - p. 72

- (58)調査した分野(Int.Cl., DB名)  
G06F 17/21-17/28