(19)日本国特許庁(JP)　(12)公開特許公報(A)　(11)特許出願公開番号
特開2013-109475
(P2013-109475A)
(43)公開日　平成25年6月6日(2013.6.6)

(51)Int.Cl.		F I			テーマコード（参考）
G06F 17/30	(2006.01)	G06F	17/30	220Z	5B091
G06F 17/27	(2006.01)	G06F	17/27	Z	5B109
G06F 17/21	(2006.01)	G06F	17/21	550A	
		G06F	17/30	170A	
		G06F	17/30	210A	

審査請求 未請求　請求項の数 10　ＯＬ　（全 17 頁）

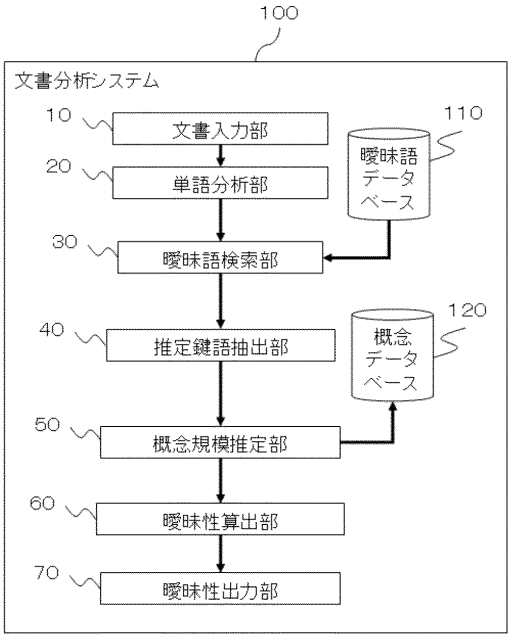
(21)出願番号　特願2011-252657 (P2011-252657)
(22)出願日　平成23年11月18日 (2011.11.18)

(71)出願人　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　100109313
弁理士　机　昌彦
(74)代理人　100124154
弁理士　下坂　直樹
(72)発明者　平尾　英司
東京都港区芝五丁目７番１号
日本電気株式会社内

Ｆターム(参考)　5B091 AA11　CC04　EA02
5B109 QA01

(54)【発明の名称】文書分析システム、文書分析方法およびプログラム

(57)【要約】
【課題】　曖昧語を含む文書について、各使用場面で曖昧語が文書の品質に与える影響の大きさを考慮して、精度よく文書の優先的な修正点や品質を推定する技術を提供する。
【解決手段】　本発明における文書分析システムは、外部から入力された文書を構成する文章に使用されている各単語の単語情報に基づいて、文書中における曖昧語を検索し、曖昧語が有る場合は当該曖昧語を抽出する曖昧語検索部と、推定鍵語を抽出するための推定鍵語抽出ルールに基づいて、文書から各曖昧語に対するそれぞれの推定鍵語を抽出する推定鍵語抽出部と、概念規模推定ルールに基づいて、推定鍵語抽出部で抽出された推定鍵語の概念の規模である概念規模指標を推定する概念規模推定部と、概念規模指標に基づいて、各曖昧語と推定鍵語との組合せ毎に曖昧性情報を算出する曖昧性算出部と、を含む。
【選択図】　図１



【特許請求の範囲】
【請求項１】
外部から入力された文書を構成する文章に使用されている各単語の単語情報に基づいて、前記文書中における、省略された概念を文脈から読み手に推定させる作用を持つ単語である曖昧語を検索し、曖昧語が有る場合は当該曖昧語を抽出する曖昧語検索部と、
前記曖昧語抽出部が抽出した各曖昧語よって省略された概念を読み手が推定する際に手掛かりとなる単語である推定鍵語を抽出するための推定鍵語抽出ルールに基づいて、文書から前記各曖昧語に対するそれぞれの推定鍵語を抽出する推定鍵語抽出部と、
推定鍵語の概念の規模を推定するための概念規模推定ルールに基づいて、前記推定鍵語抽出部で抽出された推定鍵語の概念の規模である概念規模指標を推定する概念規模推定部と、
前記概念規模指標に基づいて、前記各曖昧語と推定鍵語との組合せ毎に曖昧性情報を算出する曖昧性算出部と、
を含む文書分析システム。
【請求項２】
文書の入力を受け付ける文書入力部と、
前記文書から、当該文書を構成する文章に使用されている単語に関する情報である単語情報を抽出する単語分析部と、
省略された概念を文脈から読み手に推定させる作用を持つ単語である曖昧語を格納する曖昧語データベースと、
単語の一般概念情報を蓄積する概念データベースと、
曖昧性情報を出力する曖昧性出力部と、
をさらに含み、
前記曖昧語検索部は、前記単語分析部が抽出した単語情報及び前記曖昧語データベースに基づいて、前記曖昧語を抽出し、
前記概念規模推定部は、前記概念データベース及び所定の概念規模推定ルールに基づいて、前記概念規模指標を推定する、
請求項１に記載の文書分析システム。
【請求項３】
前記推定鍵語抽出部における所定の推定鍵語抽出ルールは、前記曖昧語の直前に存在する単語又は前記曖昧語によって修飾される単語を前記推定鍵語として抽出するルールである、
請求項１又は２に記載の文書分析システム。
【請求項４】
前記単語分析部は、前記文書を構成する文章に対して形態素解析及び構文解析を行うことで、各単語の品詞及び各単語同士の係り受け関係を前記単語情報として抽出し、
前記推定鍵語抽出部における所定の推定鍵語抽出ルールは、曖昧語の直前に存在する単語と係り受けの関係にある単語又は曖昧語を含む文の主語若しくは述語を推定鍵語として抽出するルールである、
請求項１～３のいずれか１項に記載の文書分析システム。
【請求項５】
前記概念データベースは、単語間の同義関係、類義関係、上位関係及び下位関係又は部分関係及び全体関係について、一般概念情報として格納するシソーラスである、
請求項１～４のいずれか１項に記載の文書分析システム。
【請求項６】
前記概念規模推定部における所定の概念規模推定ルールは、各推定鍵語の属する最も高い概念分類上の階層を用いて、階層の高さと単調増加の関係にある値を前記概念規模指標として推定するルールである、
請求項５に記載の文書分析システム。
【請求項７】

前記概念規模推定部における所定の概念規模推定ルールは、各推定鍵語の属する概念分類の下位分類に登録されている単語の総数を概念規模指標とするルールである、
請求項５に記載の文書分析システム。
【請求項８】
前記推定鍵語抽出部は、複数の推定鍵語抽出ルールを持つことで、各曖昧語に対する推定鍵語を複数抽出し、
前記概念規模推定部における所定の概念規模推定ルールは、前記複数抽出した各推定鍵語について算出した概念規模指標のそれぞれについて、単調増加の関係が維持される関数で同一の曖昧語に関する推定鍵語の概念規模指標を集約した値を、最終的な概念規模指標として抽出するルールである、
請求項１～７のいずれか１項に記載の文書分析システム。
【請求項９】
外部から入力された文書を構成する文章に使用されている各単語の単語情報に基づいて、前記文書中における、省略された概念を文脈から読み手に推定させる作用を持つ単語である曖昧語を検索し、曖昧語が有る場合は当該曖昧語を抽出し、
前記抽出した各曖昧語によって省略された概念を読み手が推定する際に手掛かりとなる単語である推定鍵語を抽出するための推定鍵語抽出ルールに基づいて、文書から前記各曖昧後に対するそれぞれの推定鍵語を抽出し、
推定鍵語の概念の規模を推定するための概念規模推定ルールに基づいて、前記抽出した推定鍵語の概念の規模である概念規模指標を推定し、
前記概念規模指標に基づいて、前記各曖昧語と推定鍵語との組合せ毎に曖昧性情報を算出する、
文書分析方法。
【請求項１０】
外部から入力された文書を構成する文章に使用されている各単語の単語情報に基づいて、前記文書中における、省略された概念を文脈から読み手に推定させる作用を持つ単語である曖昧語を検索し、曖昧語が有る場合は当該曖昧語を抽出し、
前記抽出した各曖昧語によって省略された概念を読み手が推定する際に手掛かりとなる単語である推定鍵語を抽出するための推定鍵語抽出ルールに基づいて、文書から前記各曖昧後に対するそれぞれの推定鍵語を抽出し、
推定鍵語の概念の規模を推定するための概念規模推定ルールに基づいて、前記抽出した推定鍵語の概念の規模である概念規模指標を推定し、
前記概念規模指標に基づいて、前記各曖昧語と推定鍵語との組合せ毎に曖昧性情報を算出する、
処理をコンピュータに実行させるプログラム。
【発明の詳細な説明】
【技術分野】
【０００１】
本発明は、文書分析システム、文書分析方法およびプログラムに関し、文書の優先的な修正点や品質を推定する文書分析システム、文書分析方法およびプログラムに関する。
【背景技術】
【０００２】
近年、情報処理装置を用いて、自然言語で書かれた文書を分析して、その文書の優先的な修正点や品質を分析する技術が開発されている。文書分析技術の一例が、特許文献１に「日本語文章修正装置、日本語文章修正方法および日本語文章修正のためのプログラム」として開示されている。
【０００３】
この特許文献１に開示された日本語文章修正装置は、機械翻訳し易い構成に関するルールが登録された文章構成データベースと、機械翻訳し易い用字・用語に関するルールが登録された用字・用語データベースと、機械翻訳し易いスタイルに関するルールが登録され

たスタイルデータベースと、修正箇所抽出手段と、表示手段と、修正手段と、出力手段と、を有する。このような構成を有する類似表現抽出装置は、次のように動作する。

【０００４】

修正箇所抽出手段は、読み込む日本語原文章から、「文を、述語の数だけ作成する」といった上記構成に関するルール、「文は、当て字、誤字、脱字を含まない」といった上記用字・用語に関するルール、または「文は、曖昧な語句を含まない」といった上記スタイルに関するルールに反する構成、用字・用語またはスタイルを抽出する。

【０００５】

表示手段は、上記抽出された構成、用字・用語またはスタイルを表示する。

【０００６】

修正手段は、外部から入力されたデータを用いて上記抽出された構成、用字・用語またはスタイルを訂正、追加または削除して上記日本語原文章を修正する。

【０００７】

出力手段は、修正により得られた翻訳易文章を出力する。

【０００８】

このような構成により、日本語原文章から上記各ルールに反するポイントを抽出・表示し、外部からの訂正、追加、削除といった修正を支援している。

【０００９】

さらに、文書分析技術の他の例が、特許文献２～特許文献５に開示されている。

【００１０】

特許文献２は、データベースに管理されている情報の概念を、各概念の抽象度に応じた配列でオントロジーに登録することで、多種多様なデータベースの追加や削除が頻繁に行われるような場合であっても、データベース選択メニューの維持管理が容易となる技術を開示する。

【００１１】

特許文献３は、検索に際して単語間の概念階層を考慮することにより、適切な検索が可能となる技術を開示する。

【００１２】

特許文献４は、次に入力したい単語の最初の読み情報の所定前に位置する単語からデータベースを参照して入力したい単語を予測表示することで、ユーザがより容易に文章を作成可能な技術を開示する。

【００１３】

特許文献５は、テキスト文書作成時に、作成中のテキスト文書中の任意の文字列に機械的処理に有用な情報（付加情報）を埋め込む等することにより、自動翻訳、自動抄録作成処理等の機械的処理に適したテキスト文書を作成可能な技術を開示する。

【先行技術文献】

【特許文献】

【００１４】

【特許文献１】特開２００７－３１６８３４号公報
【特許文献２】特開２０００－２２２４３６号公報
【特許文献３】特開２００３－０２２２７７号公報
【特許文献４】特開２００５－２３４７４２号公報
【特許文献５】特開平１１－２８２８４４号公報

【発明の概要】

【発明が解決しようとする課題】

【００１５】

特許文献１～５に記載の技術の課題は、「等」、「など」、「とか」又は「以外」といった曖昧語を含む文書の優先的な修正点や品質の推定の精度が低いことである。

【００１６】

その理由は、「等」、「など」、「とか」又は「以外」といった省略された概念を文脈

から読み手に推定させる作用を持つ曖昧語は、文章内での使われ方で文書の品質に与える影響の大きさが異なるためである。そのため、特許文献１に記載の技術のように登録された曖昧語の有無を検出する方法では、各使用場面で曖昧語が文書の品質に与える影響の大きさを考慮できない。

【００１７】

なお、ここでの文書の品質とは、文書の書き手と読み手の間における理解の齟齬の起きにくさや、文書の読み手にとっての理解のしやすさといった、少なくとも非曖昧さを含む特徴であって、文書の情報伝達の性能に関わる特徴を指す。

【００１８】

以上より、本発明の目的は、曖昧語を含む文書について、各使用場面で曖昧語が文書の品質に与える影響の大きさを考慮して、精度よく文書の優先的な修正点や品質を推定することが可能な文書分析システム、文書分析方法およびプログラムを提供することにある。

【課題を解決するための手段】

【００１９】

上記目的を達成するため、本発明に係る文書分析システムは、外部から入力された文書を構成する文章に使用されている各単語の単語情報に基づいて、前記文書中における、省略された概念を文脈から読み手に推定させる作用を持つ単語である曖昧語を検索し、曖昧語が有る場合は当該曖昧語を抽出する曖昧語検索部と、前記曖昧語抽出部が抽出した各曖昧語よって省略された概念を読み手が推定する際に手掛かりとなる単語である推定鍵語を抽出するための推定鍵語抽出ルールに基づいて、文書から前記各曖昧語に対するそれぞれの推定鍵語を抽出する推定鍵語抽出部と、推定鍵語の概念の規模を推定するための概念規模推定ルールに基づいて、前記推定鍵語抽出部で抽出された推定鍵語の概念の規模である概念規模指標を推定する概念規模推定部と、前記概念規模指標に基づいて、前記各曖昧語と推定鍵語との組合せ毎に曖昧性情報を算出する曖昧性算出部と、を含む

上記目的を達成するため、本発明に係る文書分析方法は、外部から入力された文書を構成する文章に使用されている各単語の単語情報に基づいて、前記文書中における、省略された概念を文脈から読み手に推定させる作用を持つ単語である曖昧語を検索し、曖昧語が有る場合は当該曖昧語を抽出し、前記抽出した各曖昧語によって省略された概念を読み手が推定する際に手掛かりとなる単語である推定鍵語を抽出するための推定鍵語抽出ルールに基づいて、文書から前記各曖昧後に対するそれぞれの推定鍵語を抽出し、推定鍵語の概念の規模を推定するための概念規模推定ルールに基づいて、前記抽出した推定鍵語の概念の規模である概念規模指標を推定し、前記概念規模指標に基づいて、前記各曖昧語と推定鍵語との組合せ毎に曖昧性情報を算出する。

【００２０】

上記目的を達成するため、本発明に係るプログラムは、外部から入力された文書を構成する文章に使用されている各単語の単語情報に基づいて、前記文書中における、省略された概念を文脈から読み手に推定させる作用を持つ単語である曖昧語を検索し、曖昧語が有る場合は当該曖昧語を抽出し、前記抽出した各曖昧語によって省略された概念を読み手が推定する際に手掛かりとなる単語である推定鍵語を抽出するための推定鍵語抽出ルールに基づいて、文書から前記各曖昧後に対するそれぞれの推定鍵語を抽出し、推定鍵語の概念の規模を推定するための概念規模推定ルールに基づいて、前記抽出した推定鍵語の概念の規模である概念規模指標を推定し、前記概念規模指標に基づいて、前記各曖昧語と推定鍵語との組合せ毎に曖昧性情報を算出する、処理をコンピュータに実行させる。

【発明の効果】

【００２１】

本発明によれば、曖昧語を含む文書について、各使用場面で曖昧語が文書の品質に与える影響の大きさを考慮して、精度よく文書の優先的な修正点や品質を推定することができる。

【図面の簡単な説明】

【００２２】

【図１】本発明の実施形態１に係る文書分析システム１００の構成を示すブロック図である。
【図２】文書分析システム１００のハードウェア構成を示す図である。
【図３】本発明の実施形態１に係る文書分析システム１００の全体の動作を示すフローチャートである。
【図４】本発明の実施形態２に係る文書分析システム２００の構成を示すブロック図である。
【図５】実施例における文書分析システム１００の構成を示すブロック図である。
【図６】例文から抽出される曖昧語Ｗａｊの一部の例を示す説明図である。
【図７】曖昧語Ｗａｊの直前に存在した単語を推定鍵語Ｗｂｊとする場合の推定鍵語Ｗｂｊの一部の例を示す説明図である。
【図８】曖昧語Ｗａｊの直前に存在する単語と係り受け関係を持っている単語を推定鍵語Ｗｂｊとする場合の推定鍵語Ｗｂｊの一部の例を示す説明図である。
【図９】図７の推定鍵語Ｗｂｊの概念規模指標Ｉｂｊの例を示す説明図である。
【図１０】図８の推定鍵語Ｗｂｊの概念規模指標Ｉｂｊの例を示す説明図である。
【図１１】図９の概念規模指標Ｉｂｊに基づく曖昧性情報Ａｊの例を示す説明図である。
【図１２】図１０の概念規模指標Ｉｂｊに基づく曖昧性情報Ａｊの例を示す説明図である。

【発明を実施するための形態】
【００２３】
［実施形態１］
まず、本発明の実施形態１について、図１～図３を参照して説明する。
【００２４】
図１は、本発明の実施形態１に係る文書分析システム１００の構成を示すブロック図である。
【００２５】
図１を参照すると、本発明の実施形態１に係る文書分析システム１００は、文書入力部１０、単語分析部２０、曖昧語検索部３０、推定鍵語抽出部４０、概念規模推定部５０、曖昧性算出部６０、曖昧性出力部７０、曖昧語データベース１１０及び概念データベース１２０を含む。文書分析システム１００は、上述の各構成要素を、単一の電子機器内又はサーバ及び電子機器並びにこれらを相互に接続するインターネット等の情報通信ネットワークからなるシステム内に、含んでいても良い。
【００２６】
文書分析システム１００は、曖昧語を含む文書について、各使用場面で曖昧語が文書の品質に与える影響の大きさを考慮し、文書の優先的な修正点や品質を推定する、文書分析システムである。ここで、「曖昧語」とは、「等」、「など」、「とか」又は「以外」といった省略された概念を文脈から読み手に推定させる作用を持つ単語をいう。
【００２７】
単一の電子機器で文書分析システム１００を構成する場合、文書分析システム１００は、プログラム制御により動作するコンピュータで実現可能である。
【００２８】
図２は、文書分析システム１００を単一の電子機器で実現する場合のハードウェア構成を示す図である。図２に示すように、文書分析システム１００は、データを入力する入力装置１、データ処理装置２、データ処理装置２での処理結果を出力する出力装置３及び種々のデータベースとして働く記憶装置４を含む。
【００２９】
データ処理装置２は、プログラムを記憶するリードオンリメモリ５（ＲＯＭ５）、データを一時的に記憶するワークエリアとして使用されるランダムアクセスメモリ６（ＲＡＭ６）及びＲＯＭ５に記憶されたプログラムに従って、ＲＡＭ６に記憶されているデータを処理する中央処理装置７（ＣＰＵ７）を含む。

【００３０】
入力装置１は、文書入力部１０の一部を構成する。また、データ処理装置２は、文書入力部１０、単語分析部２０、曖昧語検索部３０、推定鍵語抽出部４０、概念規模推定部５０及び曖昧性算出部６０として動作し、記憶装置４は曖昧語データベース１１０及び概念データベース１２０として動作し、出力装置３は曖昧性出力部７０として動作しても良い。
【００３１】
次に、文書分析システム１００を構成する各構成要素についてさらに詳細に説明する。
【００３２】
文書入力部１０は、文書データを受け付ける。文書入力部１０がデータとして受け付ける文書は、どのような文書であっても良いが、基本的には曖昧語を含む可能性が有り優先的な修正点や品質を分析する対象とする文書である。また、ここで言う文書は、複数の文書の束である文書群でも良い。具体的には、文書入力部１０は、入力装置１とＣＰＵ７とを含む。入力装置１が文書を入力すると、ＣＰＵ７は、その文書データをＲＡＭ６に格納する。
【００３３】
単語分析部２０は、文書入力部１０で入力された文書（文書データ）から、当該文書を構成する文章に使用されている単語に関する情報である単語情報を抽出する。この抽出は、データ処理部のＣＰＵ７によるプログラム制御に基づいて実行される。単語分析部２０は、入力された文書を構成する文章に対して、例えば形態素解析や構文解析を適用することで、各文章に使用されている全単語の単語情報の抽出を行っても良い。
【００３４】
ここで、単語は名詞、動詞、形容詞など単独で意味をなす自立語だけでなく、助詞などを含んでも良い。単語情報は、使用されている単語そのものの情報を含む。単語情報は、単語の存在位置に関する情報を含んでも良い。単語の存在位置に関する情報とは、使用箇所が同定可能な情報であればよく、単語の存在する文の出現順位や頁、目次上の章や節、項などが該当する。
【００３５】
また、単語情報は、必要に応じて単語毎の品詞、格、組み合される助詞、単語間の係り受け関係などを含めても良い。
【００３６】
曖昧語データベース１１０は、曖昧語を格納する。曖昧語データベース１１０は、特定の単語に関する後述する曖昧語検索部３０からの問い合わせに対し、問い合わせ対象の単語が曖昧語として登録されているか検索し応答するデータベースであっても良い。なお、曖昧語データベース１１０は、インターネット上のデータベースであっても良い。
【００３７】
曖昧語検索部３０は、曖昧語データベース１１０を参照して、単語分析部２０で抽出された単語情報、すなわち、各文章に使用されている全単語の単語情報に基づいて、曖昧語を検索する。さらに、検索の結果、曖昧語が曖昧語データベース１１０に有る場合、曖昧語検索部３０は、当該曖昧語を抽出する。また、曖昧語検索部３０は、曖昧語と合わせて、曖昧語を含む文の出現順位や頁、目次上の章や節、項などに関する情報である、曖昧語の存在位置に関する情報を抽出しても良い。
【００３８】
推定鍵語抽出部４０は、曖昧語検索部３０が抽出した各曖昧語について、推定鍵語抽出ルールに基づいて、文書から推定鍵語を抽出する。推定鍵語とは、曖昧語によって省略された概念を読み手が推定する際に手掛かりとなる単語である
ここで、推定鍵語抽出ルールは、曖昧語によって省略された概念を読み手が推定する際に手掛かりとなる単語を抽出するルールであればどの様なルールでも良い。推定鍵語抽出ルールは、例えば、曖昧語の直前に存在する単語又は曖昧語によって修飾される単語（被修飾語）を推定鍵語とするルールでも良い。推定鍵語抽出ルールは、曖昧語の直前に存在

する単語と係り受けの関係にある単語又は曖昧語を含む文の主語や述語などのいずれかに該当する単語を推定鍵語とするルールでも良い。又は、推定鍵語抽出ルールは、上述の推定鍵語とされる単語の組合せを推定鍵語とするルールでも良い。

【００３９】

概念データベース１２０は、収集された単語の概念分類および一般的な同義語、類義語、用法などの一般概念情報を格納する。概念データベース１２０は、特定の単語に関する問い合わせに対し、単語の一般概念情報を検索し応答するデータベースであっても良い。概念データベース１２０は、単語の上位関係及び下位関係、部分関係及び全体関係、同義関係又は類義関係などによって単語を分類し、体系づけたシソーラスなどでも良い。なお、概念データベース１２０は、インターネット上のデータベースであっても良い。

【００４０】

概念規模推定部５０は、推定鍵語抽出部４０で抽出された推定鍵語を受信し、概念データベース１２０を参照することにより得られた、推定鍵語に関する一般概念情報、及び所定の概念規模推定ルールに基づいて、推定鍵語の概念の規模である概念規模指標を推定する。

【００４１】

ここで、概念規模指標の推定を行う「概念規模推定ルール」は、推定鍵語の概念の相対的な大きさを推定できるルールであれば良い。例えば、概念規模推定ルールは、同義語の個数または類義語の個数を概念規模指標とするルールでも良い。

【００４２】

又は、概念データベース１２０として複数の階層での概念が一般概念情報として登録されたシソーラスを用いて概念ベクトルを作成した場合、概念規模推定ルールは、階層の高さと単調増加の関係にある値を概念規模指標とするルールでも良い。具体的には、概念規模推定ルールは、各推定鍵語の属する最も高い概念分類上の階層を用いて、最も高い概念分類が大分類の場合は５、中分類の場合は４、小分類の場合は３、語句の場合は２、登録がない場合は１に設定する。このように、概念規模推定ルールは、階層の高さと単調増加の関係にある値を概念規模指標とするルールでも良い。又は、概念規模推定ルールは、各推定鍵語の属する概念分類の下位分類に登録されている単語の総数などを概念規模指標とするルールでも良い。

【００４３】

曖昧性算出部６０は、概念規模推定部５０で推定した概念規模指標に基づいて、各推定鍵語と曖昧語の組合せ毎に曖昧性情報を算出する。

【００４４】

ここで、曖昧性情報は、各推定鍵語と曖昧語の組合せ毎の、曖昧性に関する指標である曖昧性指標でも良い。たとえば、曖昧性指標は、概念規模指標と単調増加の関係となる関数で導かれる曖昧性に関する指標である。曖昧性情報は、曖昧性指標に加えて、入力した文書における推定鍵語及び曖昧語の位置などに関する情報を含んでも良い。又は、曖昧性情報は、各推定鍵語と曖昧語の組合せ毎の曖昧性指標を、文書全体または任意の範囲で集計した指標であって、かつ、文書の品質を表す指標でも良い。

【００４５】

曖昧性出力部７０は、曖昧性算出部６０で算出した曖昧性情報を出力する。出力形態は、所要の形態で出力すれば良い。例えば、曖昧性出力部７０は、文書内における各推定鍵語と曖昧語の組合せについて、色分け、太字による強調又は文字の拡大などによる明示といった加工を施して、当該加工が施された文書全体を出力しても良い。

【００４６】

又は、曖昧性出力部７０は、各推定鍵語と曖昧語の組合せを抽出した表を出力しても良い。曖昧性出力部７０は、曖昧性指標が任意に設定された閾値より大きい各推定鍵語と曖昧語の組合せのみに限定して出力しても良い。曖昧性出力部７０は、曖昧性指標に基づく色分けや、太字による強調、又は単語の文字の大きさへの強弱の付与などを行っても良い。

【００４７】
各推定鍵語と曖昧語の組合せ毎の曖昧性指標を文書全体または任意の範囲で集計した指標であって、文書の品質を表す指標を曖昧性情報として出力する場合、曖昧性出力部７０は、曖昧性情報を表形式で出力しても良い。
【００４８】
または、曖昧性出力部７０は、ユーザによる各出力形態の選択を受け付けて、ベースとなる表示形態から必要に応じて表に移行できるように構成されても良い。また、曖昧性出力部７０は、必要に応じて動詞や名詞などを選択的に出力するようにしても良い。
【００４９】
曖昧性出力部７０は、曖昧性情報を出力するとともに、ユーザによる文章の修正を受け付けても良い。
【００５０】
次に、図３のフローチャートを参照して、本発明の実施形態１に係る文書分析システム１００の全体の動作について詳細に説明する。なお、図３に示すフローチャート及び以下の説明は処理の一例であり、適宜求める処理に応じて処理順等を入れ替えたり処理を戻したり繰り返したりすることを行っても良い。
【００５１】
文書入力部１０は、分析の対象とする文書の入力を受け付ける（ステップＡ１）。
【００５２】
単語分析部２０は、文書を構成する各文章に形態素解析や構文解析を適用することで、各文章に使用されている全単語の単語情報の抽出を行う（ステップＡ２）。
【００５３】
曖昧語検索部３０は、単語分析部２０で抽出された各文章に使用されている全単語の単語情報に基づいて、文書中の曖昧語の有無について曖昧語データベース１１０を参照し、曖昧語が有る場合は当該曖昧語を抽出する（ステップＡ３）。
【００５４】
推定鍵語抽出部４０は、各曖昧語について推定鍵語抽出ルールに基づき、曖昧語によって省略された概念を読み手が推定する際に手掛かりとなる単語を推定鍵語として抽出する（ステップＡ４）。
【００５５】
概念規模推定部５０は、推定鍵語の一般概念情報について概念データベース１２０を参照し、得られた一般概念情報及び所定の概念規模推定ルールに基づいて、推定鍵語の概念の規模である概念規模指標を推定する（ステップＡ５）。
【００５６】
曖昧性算出部６０は、概念規模推定部５０が推定した概念規模指標に基づいて、各推定鍵語と曖昧語の組合せ毎に曖昧性情報を算出する（ステップＡ６）。
【００５７】
曖昧性出力部７０は、曖昧性算出部６０が算出した曖昧性情報を出力する（ステップＡ７）。
【００５８】
以上説明したように、実施形態１に係る文書分析システム１００によれば、曖昧語を含む文書について、各使用場面で曖昧語が文書の品質に与える影響の大きさを考慮し、精度よく文書の優先的な修正点や品質を推定することが可能になる。、
その理由は、推定鍵語抽出部４０が文書において曖昧語によって省略された概念を読み手が推定する際に手掛かりとなる単語を推定鍵語として抽出し、概念規模推定部５０が推定鍵語の概念の規模である概念規模指標を推定するからである。各推定鍵語と曖昧語の組合せ毎に曖昧性指標を算出することで、文書分析システム１００は、文書の品質に影響を与える文章内での曖昧語の使われ方まで考慮した分析が可能になる。
【００５９】
［実施形態２］

次に、本発明の実施形態２について、図４を参照して説明する。

【００６０】

図４は、本発明の実施形態２に係る文書分析システム２００の構成を示すブロック図である。

【００６１】

図４を参照すると、本発明の実施形態２に係る文書分析システム２００は、曖昧語検索部３５、推定鍵語抽出部４５、概念規模推定部５５及び曖昧性算出部６５を含む。

【００６２】

曖昧語検索部３５は、外部から入力された文書を構成する文章に使用されている各単語の単語情報に基づいて、文書中における、省略された概念を文脈から読み手に推定させる作用を持つ単語である曖昧語を検索し、曖昧語が有る場合は当該曖昧語を抽出する。

【００６３】

推定鍵語抽出部４５は、曖昧語抽出部３５が抽出した各曖昧語によって省略された概念を読み手が推定する際に手掛かりとなる単語である推定鍵語を抽出するための推定鍵語抽出ルールに基づいて、文書から各曖昧語に対するそれぞれの推定鍵語を抽出する。

【００６４】

概念規模推定部５５は、推定鍵語の概念の規模を推定するための概念規模推定ルールに基づいて、推定鍵語抽出部４５で抽出された推定鍵語の概念の規模である概念規模指標を推定する。

【００６５】

曖昧性算出部６５は、概念規模推定部５５で推定された概念規模指標に基づいて、各推定鍵語と曖昧語の組合せ毎に曖昧性情報を算出する。

【００６６】

以上説明したように、文書分析システム２００によれば、曖昧語を含む文書について、各使用場面で曖昧語が文書の品質に与える影響の大きさを考慮して、精度よく文書の優先的な修正点や品質を推定することができる。

【００６７】

なお、本発明に係る文書分析方法は、上述した文書分析システム１００又は文書分析システム２００が実行する動作方法である。また、本発明に係るプログラムは、実施形態１又は実施形態２において説明した各動作を、コンピュータに実行させるプログラムであれば良い。

【００６８】

［実施例］

次に、図５乃至図１２を参照して、具体的な実施例を用いて、本発明の実施形態１に係る文書分析システム１００の動作について説明する。

【００６９】

本実施例では、次のことを目的としている。

【００７０】

先ず、文書分析システム１００は、情報システム構築に関する提案書や仕様書といった、曖昧な箇所を排除すべき文書Ｄ内で、曖昧語Ｗａが文書Ｄの品質に与える曖昧性である曖昧性指標ＩＡを各曖昧語Ｗａの使用場面毎に算出する。曖昧語Ｗａは、「等」、「など」、「とか」又は「以外」といった省略された概念を文脈から読み手に推定させる作用を持つ単語である。

【００７１】

文書分析システム１００は、算出した曖昧性指標ＩＡに基づき、各曖昧語の文書中の位置や曖昧性の程度に関する曖昧性情報Ａを推定する。文書分析システム１００は、推定された曖昧性情報Ａを出力することで、文書Ｄにおける優先的に修正すべき曖昧箇所の把握や、複数の文書間の比較として品質の低い文書の明確化などを容易にし、文書の改善を効率化する。

【００７２】

また、本実施例では、文書分析システム１００は、図５に示されるように、文書解析システム３００と、インターネット・サーバ４００とを含むものとする。

【００７３】

文書解析システム３００は、分析実施者Ｂの持つＰＣ端末５００上で動作し、入力部６００及び出力部７００を介して、分析実施者Ｂが曖昧性指標Ａを推定したい文書群を構成する文書の入力と、曖昧性情報Ａの提示を実現する。

【００７４】

インターネット・サーバ４００は、通信ネットワークを介して文書解析システム３００を実装した分析実施者Ｂの持つＰＣ端末５００と接続されている。インターネット・サーバ４００は、文書解析システム３００からの単語の問い合わせに対し、曖昧語Ｗａとしての登録の有無の検索を可能にする。また、インターネット・サーバ４００は、単語の意味などの概念情報の問い合わせに対し、単語の概念分類や一般的な曖昧語や類義語、用法に関連する一般概念情報Ｃｇの検索を可能にする。

【００７５】

次に、図５と図１との対応関係について説明する。

【００７６】

文書入力部１０、単語分析部２０、曖昧語検索部３０、推定鍵語抽出部４０、概念規模推定部５０及び曖昧性算出部６０は、文書解析システム３００内に含まれる。曖昧性出力部７０は、出力部７００として動作する。曖昧語データベース１１０及び概念データベース１２０は、インターネット・サーバ４００内に含まれる。

【００７７】

この様な構成要素を含む文書解析システム３００及びインターネット・サーバ４００は、以下のように動作する。

【００７８】

文書解析システム３００は、入力部６００を介して、情報システム構築に関する提案書や仕様書といった、分析実施者Ｂが曖昧な箇所を排除するための曖昧性情報Ａを得たい文書Ｄの入力を受け付ける。文書解析システム３００は、文書Ｄを構成する文章毎に形態素解析及び構文解析を適用し、文書を構成する単語に分解し、その係り受け関係や格を解析することで、単語Ｗとその格、係り受け構造等を単語情報として抽出する。なお、ここでの「単語」は、「に関連する」のような複数の単語からなる表現などを含んでも良い。さらに、文書解析システム３００は、文書の文の出現順に通し番号として文番号を付け、各単語を含む文の文番号を単語情報に加える。

【００７９】

インターネット・サーバ４００は、省略された概念を文脈から読み手に推定させる作用を持つ曖昧語や曖昧表現に関する情報である曖昧語情報Ｃａを格納する。また、インターネット・サーバ４００は、任意の単語や表現の情報を抽出する検索エンジンなどの機能も有することで、文書解析システム３００からの問い合わせに応じて、問い合わせ対象の単語が曖昧語情報Ｃａに存在するかどうかを判定し、提示しても良い。

【００８０】

なお、抽出する曖昧語は省略された概念を文脈から読み手に推定させる作用を持つ曖昧語であれば何でも良い。曖昧語は、例えば、「等」、「など」、「とか」、「以外」、「といった」、「みたいな」などの単語が該当する。また、曖昧語は、「のような」、「に関する」、「に関連する」、「に関わる」などの助詞と組み合された複数の単語からなる表現などを含んでも良い。

【００８１】

文書解析システム３００は、文書Ｄに含まれる全ての単語Ｗｉ（ｉ＝１、２、・・・、ｎ）についてインターネット・サーバ４００に曖昧語情報Ｃａの曖昧語に該当する単語であるかどうかを問い合わせる。文書解析システム３００は、曖昧語に該当するという判定結果となった単語Ｗｊ（ｊ＝１、２、・・・、ｍ）を曖昧語Ｗａｊ（ｊ＝１、２、・・・、ｍ）とし、曖昧語の存在する文の文番号と共に抽出する。なお、文書Ｄ内に同一の曖昧

語が複数回使用されていた場合も使用場面はそれぞれ異なるため、文書解析システム３００は、複数の同一の曖昧語を別々に抽出する。

【００８２】

例えば、文書Ｄの一部が「本購買システムは国内拠点と海外支社間の物流業務の効率化を目的としている。開発担当者は本システムの性能を関連する文書等の要件を満たすようにすること。また要件に変更が有れば、都度協議する。協議結果、変更が生じた場合は議事録等に記載し、開発物はこの記載に沿うように開発を行う。開発システムの入力、出力等は旧システムと同じとする。新たな処理として新法に対応するため輸出規定に該当するかどうかの判定などが加わる。他にも簿価の計算式の係数などを旧版値から換算する必要がある。これらの処理の他には機能面に大きな変化などは無い。」という文面からなる場合を考える。この文書に関して、文書解析システム３００は、図６に示すように、一文毎に文番号を付与し、曖昧語が存在する文に関しては、各文の文番号に対応して曖昧語Ｗａｊを記載した表を作成することで、曖昧語を抽出しても良い。

【００８３】

文書解析システム３００は、曖昧語によって省略された概念を読み手が推定する際に手掛かりとなる推定鍵語Ｗｂｊとして、曖昧語Ｗａｊの直前に存在する単語を抽出する。例えば、図６の事例で有れば、文書解析システム３００は、推定鍵語Ｗｂｊとして図７の一番右の列に示すような単語を抽出する。

【００８４】

曖昧語Ｗａｊの直前に存在する単語を推定鍵語Ｗｂｊとすることが有効な理由は、「文書等」という表記が有った時、読み手は「文書」に類するその他のものを推定することになるためである。

【００８５】

又は、文書解析システム３００は、曖昧語Ｗａｊの直前に存在する単語と係り受け関係を持っている単語を推定鍵語Ｗｂｊとして抽出しても良い。例えば、図６の事例で有れば、文書解析システム３００は、推定鍵語Ｗｂｊとして図８の一番右の列に示すような単語を抽出する。

【００８６】

曖昧語Ｗａｊの直前に存在する単語と係り受け関係を持っている単語を推定鍵語Ｗｂｊとすることが有効な理由は、「係数などを旧版値から換算する」という表記が有った時、読み手は「換算する」という動作がそぐうものを推定することになるためである。

【００８７】

なお、推定鍵語Ｗｂｊは曖昧語Ｗａｊに対して少なくとも一つ抽出されれば良いが、文書解析システム３００は、一つの曖昧語に対して複数の推定鍵語Ｗｂｊを抽出しても良い。

【００８８】

インターネット・サーバ４００は、単語の一般的な上位関係及び下位関係、部分関係及び全体関係、同義関係又は類義関係などによって単語を分類し、体系づけたシソーラスの一般概念情報Ｃｇを格納する。また、インターネット・サーバ４００は、任意の単語の情報を抽出する検索エンジンなどの機能も有することで、文書解析システム３００からの問い合わせに応じて、一般概念情報Ｃｇを抽出し、提示しても良い。具体的には、インターネット・サーバ４００は、問い合わせ対象の単語の一般的な概念分類として「大分類」、「中分類」、「小分類」、「語句」といった分類及び当該分類に帰属する単語の数を一般概念情報Ｃｇとして抽出し、提示しても良い。

【００８９】

文書解析システム３００は、推定鍵語Ｗｂｊのそれぞれの一般概念情報Ｃｇについてインターネット・サーバ４００に問い合わせを行う。文書解析システム３００は、インターネット・サーバ４００内に保存されたシソーラスの一般概念情報Ｃｇの分類体系から、各推定鍵語Ｗｂｊが属する最も高い概念分類を抽出する。文書解析システム３００は、その概念分類の下位分類に登録されている単語の個数を概念規模指標Ｉｂｊの値として抽出す

る。
【００９０】
例えば、「文書」という推定鍵語Ｗｂｊが一般概念情報Ｃｇの分類体系では「小分類」に分類されている場合、その下位分類である「語句」とその同義語として登録されている単語を抽出する。「文書」という推定鍵語Ｗｂｊの下位分類である語句とその同義語として、例えば「書状」、「書類」、「ドキュメント」、「公文書」、「私文書」、「私書」、「契約書」、「覚書」、・・・・といった単語が、計１５６個存在するとする。本実施例においては、文書解析システム３００は、１５６という値を概念規模指標Ｉｂｊの値とする。
【００９１】
また、「議事録」のように推定鍵語Ｗｂｊが一般概念情報Ｃｇの分類体系で「語句」に分類されている場合、下位に帰属する単語は無いので、文書解析システム３００は概念規模指標を１としても良い。又は、文書解析システム３００は、「議事録」の同義語や類義語として登録されている語を一般概念情報Ｃｇから抽出し、抽出できた単語の数を概念規模指標としても良い。例えば、「議事録」の同義語又は類義語として「メモ」、「記録」及び「会議録」の３語が得られたとすると、文書解析システム３００は、概念規模指標Ｉｂｊの値を３としても良い。一般概念情報Ｃｇの分類体系に存在しない単語の場合は帰属する単語は無いので、文書解析システム３００は概念規模指標を１としても良い。
【００９２】
図７の事例で有れば、概念規模指標Ｉｂｊは図９の一番右の列に示すように、図８の事例で有れば、概念規模指標Ｉｂｊは図１０の一番右の列に示すように算出される。
【００９３】
または、文書解析システム３００は、曖昧語Ｗａｊの直前に存在した単語と、曖昧語Ｗａｊの直前に存在する単語と係り受け関係を持っている単語の両方を推定鍵語Ｗｂｊとして抽出しても良い。この場合は、文書解析システム３００は、複数の概念規模指標Ｉｂｊと単調増加の関係が維持される相加平均や相乗平均のような関数で、同一の曖昧語Ｗａｊに関するそれぞれの推定鍵語Ｗｂｊに対応する概念規模指標Ｉｂｊを集約した値を、改めて最終的な概念規模指標Ｉｂｊとしても良い。
【００９４】
次に文書解析システム３００は、推定した各曖昧語Ｗａｊに対応する概念規模指標Ｉｂｊを文書内での概念規模指標Ｉｂの最大値Ｉｂｍａｘで除して正規化した値を、各推定鍵語Ｗｂｊと曖昧語Ｗａｊの組合せ毎の曖昧性指標Ｉａｊとして算出する。文書解析システム３００は、曖昧性指標Ｉａｊ及び曖昧語Ｗａｊの文番号を曖昧語Ｗａｊの曖昧性情報Ａｊとする。また、概念規模指標Ｉｂｊを文書全体で集計した値を文書Ｄの品質を表す曖昧性情報Ａｄとして算出する。
【００９５】
例えば、図９の事例で有れば、曖昧性情報Ａｊは図１１のように算出される。図１１は、「文書等」の部分が、「文書」という一般的な語では「等」の概念の推定範囲が広く曖昧性が高いことを示す。また、図１０の事例で有れば、曖昧性情報Ａｊは図１２のように算出される。図１２は、「出力等は旧システムと同じ」の部分が、「同じ」という汎用性の高い語では「等」の概念を限定しにくく曖昧性が高いことを示す。
【００９６】
文書解析システム３００は、曖昧性情報Ａｊに基づき、「文書等」や「出力等は旧システムと同じ」など、文書Ｄ内における各推定鍵語Ｗｂｊと曖昧語Ｗａｊの組合せ部分を着色し明示することで、修正すべき曖昧な文の箇所を分かりやすくし表示する。また、文書解析システム３００は、曖昧性情報Ａｄに基づき、推定鍵語Ｗｂｊと曖昧語Ｗａｊの組合せの曖昧性情報Ａｄを文書Ｄ全体および、目次の章単位などで集計して表もしくはグラフなどの形式で出力しても良い。当該出力により、文書解析システム３００は、文書Ｄの品質を表すメトリクス（評価尺度）及び修正すべき曖昧な章を判断する情報を提供する。
【００９７】

以上、各実施形態及び実施例を参照して本発明を説明したが、本発明は以上の実施形態及び実施例に限定されるものではない。本発明の構成や詳細には、本発明のスコープ内で同業者が理解し得る様々な変更をすることができる。

【産業上の利用可能性】

【００９８】

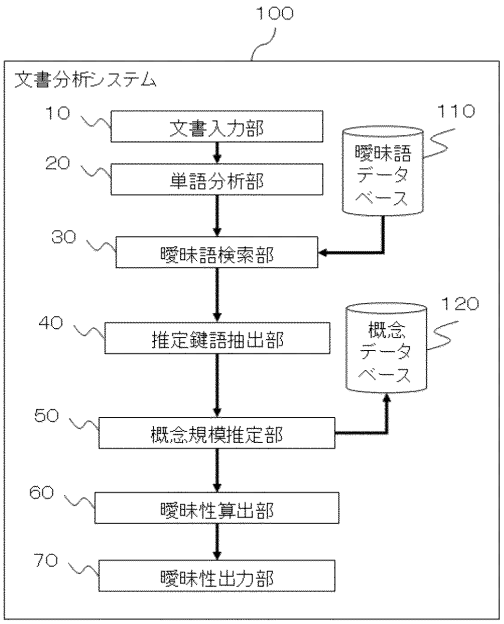
本発明によれば、ソフトウェアやシステムの開発における要件定義などの作業においてやり取りされる各種文書に関して、文書の曖昧さを特に問題の大きい場所から優先的に修正することが可能になる。これにより文書作成や文書レビューの効率化に繋がる。また、複数の読み手の間に異なる解釈が起きる状況などが減少し、手戻りの減少や顧客満足の向上などシステム開発の効率化に関する用途に適用できる。

【符号の説明】

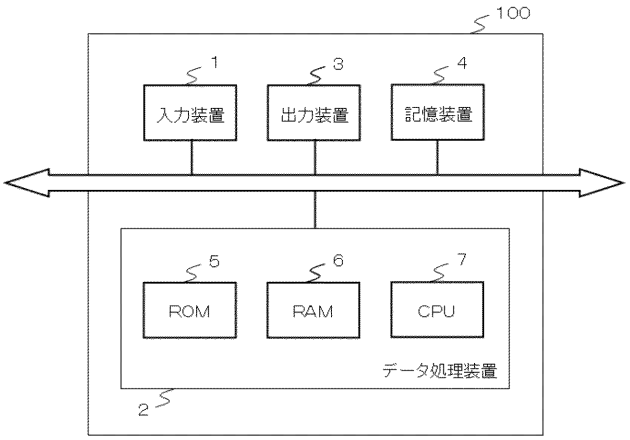
【００９９】

１　入力装置
２　データ処理装置
３　出力装置
４　記憶装置
５　ＲＯＭ
６　ＲＡＭ
７　ＣＰＵ
１０　文書入力部
２０　単語分析部
３０、３５　曖昧語検索部
４０、４５　推定鍵語抽出部
５０、５５　概念規模推定部
６０、６５　曖昧性算出部
７０　曖昧性出力部
１００、２００　文書分析システム
１１０　曖昧語データベース
１２０　概念データベース
３００　文書解析システム
４００　インターネット・サーバ
５００　ＰＣ端末
６００　入力部
７００　出力部

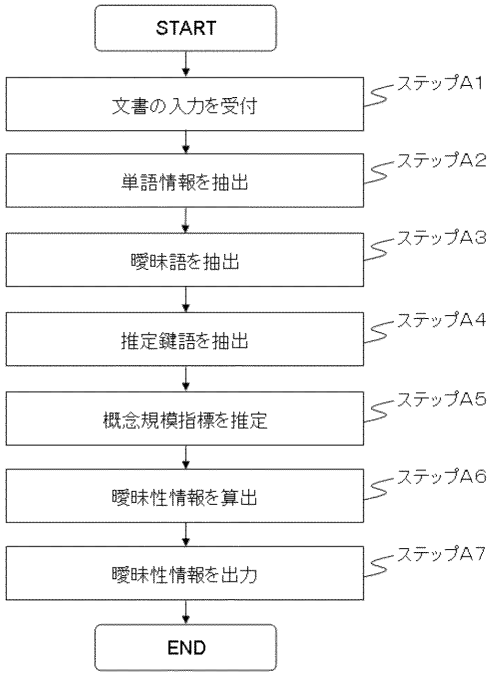
【図１】


100
文書分析システム
10
文書入力部
20
単語分析部
30
曖昧語検索部
40
推定鍵語抽出部
50
概念規模推定部
60
曖昧性算出部
70
曖昧性出力部
110
曖昧語データベース
120
概念データベース

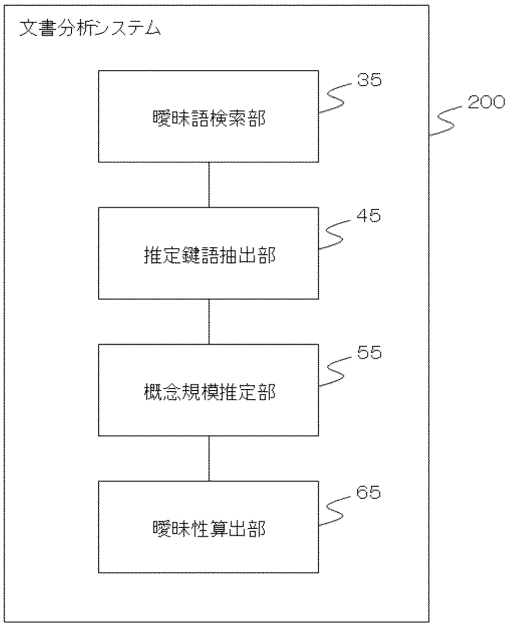
【図２】


100
1
入力装置
3
出力装置
4
記憶装置
5
ROM
6
RAM
7
CPU
データ処理装置
2

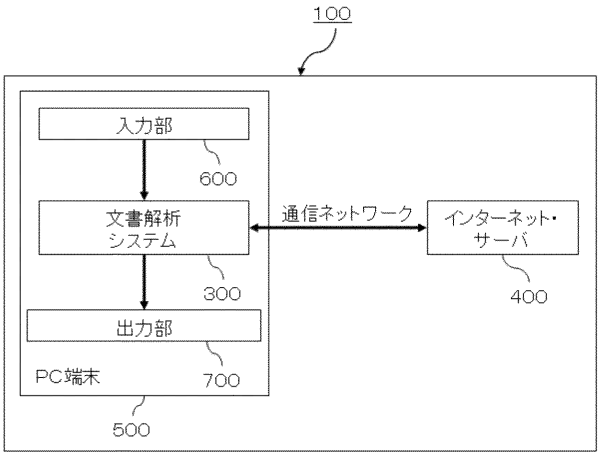
【図３】


START
文書の入力を受付
ステップA1
単語情報を抽出
ステップA2
曖昧語を抽出
ステップA3
推定鍵語を抽出
ステップA4
概念規模指標を推定
ステップA5
曖昧性情報を算出
ステップA6
曖昧性情報を出力
ステップA7
END

【図４】


文書分析システム
200
曖昧語検索部
35
推定鍵語抽出部
45
概念規模推定部
55
曖昧性算出部
65

【図５】



【図６】

文番号	文面	曖昧語Wa
1	本購買システムは国内拠点と海外支社間の物流業務の効率化を目的としている。	-
2	開発担当者は本システムの性能を関連する文書等の要件を満足するようにすること。	等
3	また要件に変更が有れば、都度協議する。	-
4	協議結果、変更が生じた場合は議事録等に記載し、開発物は二の記載に沿うように開発を行う。	等
5	開発システムの入力、出力等は旧システムと同じとする。	等
6	新たな処理として新法に対応するための輸出規定に該当するかどうかの判定などが加わる。	など
7	他にも諸価の計算式の係数などを旧版値から換算する必要がある。	など
8	これらの処理の他には機能面に大きな変化などは無い。	など
・	・	・
・	・	・
・	・	・

【図７】

文番号	文面	曖昧語Wa	推定課語Wb（直前語）
2	開発担当者は本システムの性能を関連する文書等の要件を満足するようにすること。	等	文書
4	協議結果、変更が生じた場合は議事録等に記載し、開発物は二の記載に沿うように開発を行う。	等	議事録
5	データ開発システムの入力、出力等は旧システムと同じとする。	等	出力
6	新たな処理として新法に対応するための輸出規定に該当するかどうかの判定などが加わる。	など	判定
7	他にも諸価の計算式の係数などを旧版値から換算する必要がある。	など	係数
8	これらの処理の他には機能面に大きな変化などは無い。	など	変化
・	・	・	・
・	・	・	・
・	・	・	・

【図８】

文番号	文面	曖昧語Wa	推定課語Wb（係り受け）
2	開発担当者は本システムの性能を関連する文書等の要件を満足するようにすること。	等	満足する
4	協議結果、変更が生じた場合は議事録等に記載し、開発物は二の記載に沿うように開発を行う。	等	記載し
5	データ開発システムの入力、出力等は旧システムと同じとする。	等	同じ
6	新たな処理として新法に対応するための輸出規定に該当するかどうかの判定などが加わる。	など	加わる
7	他にも諸価の計算式の係数などを旧版値から換算する必要がある。	など	換算する
8	これらの処理の他には機能面に大きな変化などは無い。	など	無い
・	・	・	・
・	・	・	・
・	・	・	・

【図９】

文番号	文面	曖昧語Wa	推定鍵語Wb〈直前語〉	最大概念分類	概念規模指標Bj
2	開発担当者は本システムの性能を関連する文書等の要件を満足するようにすること。	等	文書	小分類	156
4	協議結果、変更が生じた場合は議事録等に記載し、開発物はこの記載に沿うように開発を行う。	等	議事録	語句	3
5	データ開発システムの入力、出力等は旧システムと同じとする。	等	出力	語句	7
6	新たな処理として新法に対応するため輸出規定に該当するかどうかの判定などが加わる。	など	判定	語句	4
7	他にも算価の計算式の係数などを旧版価から換算する必要がある。	など	係数	語句	2
8	これらの処理の他には機能面に大きな変化などは無い。	など	変化	中分類	32
・ ・ ・	・ ・ ・	・ ・ ・	・ ・ ・	・ ・ ・	・ ・ ・

【図１０】

文番号	文面	曖昧語Wa	推定鍵語Wb〈係り受け〉	最大概念分類	概念規模指標Bj
2	開発担当者は本システムの性能を関連する文書等の要件を満足するようにすること。	等	満足する	中分類	21
4	協議結果、変更が生じた場合は議事録等に記載し、開発物はこの記載に沿うように開発を行う。	等	記載	語句	5
5	データ開発システムの入力、出力等は旧システムと同じとする。	等	同じ	大分類	122
6	新たな処理として新法に対応するため輸出規定に該当するかどうかの判定などが加わる。	など	加わる	語句	6
7	他にも算価の計算式の係数などを旧版価から換算する必要がある。	など	換算	語句	4
8	これらの処理の他には機能面に大きな変化などは無い。	など	無い	中分類	16
・ ・ ・	・ ・ ・	・ ・ ・	・ ・ ・	・ ・ ・	・ ・ ・

【図１１】

文番号	文面	曖昧語Wa	推定鍵語Wb〈直前語〉	曖昧性情報Aj
2	開発担当者は本システムの性能を関連する文書等の要件を満足するようにすること。	等	文書	1.000
4	協議結果、変更が生じた場合は議事録等に記載し、開発物はこの記載に沿うように開発を行う。	等	議事録	0.019
5	データ開発システムの入力、出力等は旧システムと同じとする。	等	出力	0.045
6	新たな処理として新法に対応するため輸出規定に該当するかどうかの判定などが加わる。	など	判定	0.026
7	他にも算価の計算式の係数などを旧版価から換算する必要がある。	など	係数	0.013
8	これらの処理の他には機能面に大きな変化などは無い。	など	変化	0.205
・ ・ ・	・ ・ ・	・ ・ ・	・ ・ ・	・ ・ ・

【図１２】

文番号	文面	曖昧語Wa	推定鍵語Wb〈係り受け〉	曖昧性情報Aj
2	開発担当者は本システムの性能を関連する文書等の要件を満足するようにすること。	等	満足する	0.172
4	協議結果、変更が生じた場合は議事録等に記載し、開発物はこの記載に沿うように開発を行う。	等	記載	0.041
5	データ開発システムの入力、出力等は旧システムと同じとする。	等	同じ	1.000
6	新たな処理として新法に対応するため輸出規定に該当するかどうかの判定などが加わる。	など	加わる	0.049
7	他にも算価の計算式の係数などを旧版価から換算する必要がある。	など	換算	0.003
8	これらの処理の他には機能面に大きな変化などは無い。	など	無い	0.131
・ ・ ・	・ ・ ・	・ ・ ・	・ ・ ・	・ ・ ・