

【公報種別】特許法第17条の2の規定による補正の掲載

【部門区分】第6部門第3区分

【発行日】平成19年8月9日(2007.8.9)

【公開番号】特開2006-39871(P2006-39871A)

【公開日】平成18年2月9日(2006.2.9)

【年通号数】公開・登録公報2006-006

【出願番号】特願2004-217744(P2004-217744)

【国際特許分類】

G 06 F 17/30 (2006.01)

【F I】

G 06 F 17/30 170 A

G 06 F 17/30 320 D

G 06 F 17/30 350 C

【手続補正書】

【提出日】平成19年6月22日(2007.6.22)

【手続補正1】

【補正対象書類名】特許請求の範囲

【補正対象項目名】全文

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項1】

所定の単語に対する類義語を検索する類義語検索装置であって、

前記所定の単語に関する単語情報を取得する単語情報取得手段と、

単語に関する単語情報にその単語に関連してその単語を特徴付ける単語に関する関連単語情報を複数有する関連単語集合情報が1つのデータ構造として関連付けられた単語関連情報が複数格納されるテーブル構造の関連単語テーブルに基づき、前記取得した単語情報の所定の単語に対応する単語関連情報を取得する単語関連情報取得手段と、

前記関連単語テーブルにおける前記取得した単語関連情報の関連単語集合情報に類似する関連単語集合情報を備えた単語関連情報を検索する関連単語検索手段と、

この関連単語検索手段にて検索された前記単語関連情報の単語情報を取得して前記所定の単語に対する類義語として選出する類義語選出手段と、

を具備したことを特徴とした類義語検索装置。

【請求項2】

請求項1に記載の類義語検索装置であって、

前記単語関連情報取得手段は、複数の単語にて構成された文章に出現在する単語情報に、その単語の文章における近傍に位置する他の単語に関する関連単語情報を複数有する関連単語集合情報が1つのデータ構造として関連付けられ、前記文章に出現在する各単語についてそれぞれ構成された複数の単語関連情報が格納されるテーブル構造の関連単語テーブルに基づいて、前記単語関連情報を取得する

ことを特徴とした類義語検索装置。

【請求項3】

請求項2に記載の類義語検索装置であって、

前記関連単語テーブルは、文章に関する文章情報を複数記憶する文章記憶手段に基づいて生成された

ことを特徴とした類義語検索装置。

【請求項4】

請求項2または請求項3に記載の類義語検索装置であって、

前記単語関連情報の単語情報は、形態素解析により前記文章から抽出した単語に関する情報である

ことを特徴とした類義語検索装置。

【請求項 5】

請求項 1 に記載の類義語検索装置であって、

前記単語関連情報取得手段は、単語に関する単語情報に、その単語に関する説明文に出現在する単語に関する関連単語情報を複数有する関連単語集合情報が 1 つのデータ構造として関連付けられた単語関連情報が単語毎に複数格納されるテーブル構造の関連単語テーブルに基づいて、前記単語関連情報を取得する

ことを特徴とした類義語検索装置。

【請求項 6】

請求項 1 に記載の類義語検索装置であって、

前記単語関連情報取得手段は、単語に関する単語情報に、その単語のシソーラス項目となる単語群を構成する各単語に関する関連単語情報を複数有する関連単語集合情報が 1 つのデータ構造として関連付けられ、単語毎に複数格納されるテーブル構造の関連単語テーブルに基づいて、前記単語関連情報を取得する

ことを特徴とした類義語検索装置。

【請求項 7】

請求項 1 ないし請求項 6 のいずれかに記載の類義語検索装置であって、

前記単語情報取得手段にて前記所定の単語に関する単語情報を取得すると、前記関連単語テーブルを作成する関連単語テーブル生成手段を具備した

ことを特徴とした類義語検索装置。

【請求項 8】

請求項 1 ないし請求項 6 のいずれかに記載の類義語検索装置であって、

前記関連単語テーブルを記憶する関連単語記憶手段を具備した

ことを特徴とした類義語検索装置。

【請求項 9】

請求項 1 ないし請求項 8 のいずれかに記載の類義語検索装置であって、

前記関連単語検索手段は、前記所定の単語に対応する単語関連情報の関連単語集合情報に対する前記関連単語テーブルにおける単語関連情報の関連単語集合情報の類似状態に関する類似度情報を演算し、この演算した類似度情報に基づいて前記単語関連情報を検出する

ことを特徴とした類義語検索装置。

【請求項 10】

請求項 9 に記載の類義語検索装置であって、

前記関連単語検索手段は、前記所定の単語に対応する単語関連情報の関連単語集合情報を構成する複数の関連単語情報と、前記関連単語テーブルにおける前記単語関連情報の関連単語集合情報を構成する複数の関連単語情報との同一性に応じた重み付けに関する重み度情報を演算し、この重み度情報に基づいて前記類似度情報を演算する

ことを特徴とした類義語検索装置。

【請求項 11】

請求項 1 ないし請求項 10 のいずれかに記載の類義語検索装置であって、

前記関連単語テーブルの単語関連情報は、前記単語情報の単語の分類に関する単語分類情報を有し、

前記関連単語検索手段は、前記所定の単語に対応する単語関連情報の前記単語分類情報と同一の単語分類情報を有した前記関連単語テーブルにおける単語関連情報から検索することを特徴とした類義語検索装置。

【請求項 12】

請求項 1 ないし請求項 11 のいずれかに記載の類義語検索装置であって、

前記関連単語検索手段は、

検出した前記単語関連情報の単語情報を表示手段に表示させる表示制御手段と、
入力操作可能な操作手段における入力操作に応じて前記表示手段に表示される前記単語
の選択および変更の少なくともいずれか一方の実施により選択された単語および変更され
た単語のうちの少なくともいずれか1つの単語に関する単語関連情報を設定して取得する
単語関連情報設定手段と、を備えた

ことを特徴とした類義語検索装置。

【請求項13】

請求項12に記載の類義語検索装置であって、
前記関連単語検索手段は、単語関連情報設定手段にて取得した前記単語関連情報の関連
単語集合情報に類似する関連単語集合情報を備えた単語関連情報をさらに検索する
ことを特徴とした類義語検索装置。

【請求項14】

請求項1ないし請求項13のいずれかに記載の類義語検索装置であって、
前記類義語選出手段にて類義語として選出した単語情報を、前記所定の単語に対応する
単語情報に関連付けて1つのデータ構造とした類義語情報を生成し、各種情報を記憶する
記憶手段に複数記憶させる類義語情報生成手段を具備した
ことを特徴とした類義語検索装置。

【請求項15】

請求項1ないし請求項14のいずれかに記載の類義語検索装置であって、
前記文章記憶手段は、前記文章を複数含む文献毎の集合とした前記文章情報を複数有す
る文献情報を複数記憶するテーブル構造に構築された
ことを特徴とした類義語検索装置。

【請求項16】

請求項15に記載の類義語検索装置であって、
前記関連単語検索手段は、前記所定の単語が含まれる前記文献の文献情報における前記
文章情報の文章に含まれる前記他の単語に対応する単語関連情報から検索する
ことを特徴とした類義語検索装置。

【請求項17】

請求項15または請求項16に記載の類義語検索装置であって、
前記文献情報は、前記文献の分類に関する文献分類情報を有し、
前記関連単語検索手段は、前記所定の単語が含まれる前記文献の文献情報の文献分類情
報と同一の文献分類情報を有した前記文献情報における前記文章情報の文章に含まれ前記
所定の単語の近傍に位置する前記他の単語に対応する単語関連情報から検索する
ことを特徴とした類義語検索装置。

【請求項18】

請求項15ないし請求項17のいずれかに記載の類義語検索装置であって、
前記文献情報は、前記文献に関する書誌情報を有し、
前記関連単語検索手段は、前記所定の単語が含まれる前記文献の文献情報に関連付けら
れた書誌情報を有した前記文献情報における前記文章情報の文章に含まれ前記所定の単語
の近傍に位置する前記他の単語に対応する単語関連情報から検索する
ことを特徴とした類義語検索装置。

【請求項19】

請求項1ないし請求項18のいずれかに記載の類義語検索装置と、
この類義語検索装置にて類義語として選出した前記単語情報および前記単語情報取得手
段で取得した単語情報のうちの少なくともいずれか1つを含む文章に関する前記文章情報
の検索を実行する検索手段と、
を具備したことを特徴とした情報検索装置。

【請求項20】

請求項15ないし請求項18のいずれかに記載の類義語検索装置と、
この類義語検索装置にて類義語として選出した前記単語情報および前記単語情報取得手

段で取得した単語情報のうちの少なくともいずれか1つを含む文献情報の検索を実行する検索手段と、

を具備したことを特徴とした情報検索装置。

【請求項21】

請求項19または請求項20に記載の情報検索装置であって、

前記検索手段は、前記類義語検索装置にて選出した前記単語情報および前記単語情報取得手段で取得した単語情報を単語を、入力操作可能な操作手段における入力操作に応じて選択および変更のうちの少なくともいずれか一方を実施可能に前記表示手段に表示させ、前記入力操作に応じて前記単語の選択および変更のうちの少なくともいずれか一方が実施されて選択された単語および変更された単語のうちの少なくともいずれか一方の単語に関する単語情報に基づいて検索を実行する

ことを特徴とした情報検索装置。

【請求項22】

演算手段により、複数の単語にて構成された文章に関する文章情報を複数記憶する記憶手段を利用して所定の単語に対する類義語を検索する類義語検索方法であって、

前記演算手段は、

所定の単語に関する単語情報を取得すると、単語に関する単語情報にその単語に関連し、その単語を特徴付ける単語に関する関連単語情報を複数有する関連単語集合情報が1つのデータ構造として関連付けられた単語関連情報を複数記憶するテーブル構造に構築された関連単語記憶領域に基づき、前記取得した所定の単語に対応する前記単語関連情報を取得し、

この取得した単語関連情報を関連単語集合情報に類似する関連単語集合情報を備えた単語関連情報を検索し、

この検索した単語関連情報の単語情報を前記所定の単語に対する類義語として選出する

ことを特徴とする類義語検索方法。

【請求項23】

演算手段を、請求項1ないし請求項18のいずれかに記載の類義語検索装置、または、請求項19ないし請求項21のいずれかに記載の情報検索装置として機能させる

ことを特徴とした検索プログラム。

【請求項24】

請求項22に記載の類義語検索方法を演算手段に実行させる

ことを特徴とした検索プログラム。

【手続補正2】

【補正対象書類名】明細書

【補正対象項目名】0002

【補正方法】変更

【補正の内容】

【0002】

従来、所定の単語に関する関連語を検索する装置として、文書ファイルから抽出した単語を利用する構成が知られている（例えば、特許文献1参照）。この特許文献1に記載のものは、文書ファイルを参照して文書中で前後に共通した単語が連なる2つの異なる単語を関連語の候補対として選定し、二連単語リストを作成する。注目単語に対して、前置単語のリストと、後置単語のリストを生成する。そして、二連単語リストを参照して、各前置単語に対する後置単語の集合と、各後置単語に対する前置単語の集合とを生成する。この語、これらの集合の共通要素から、関連語の候補対を抽出する構成が採られている。

【手続補正3】

【補正対象書類名】明細書

【補正対象項目名】0004

【補正方法】変更

【補正の内容】**【0004】**

しかしながら、上述した特許文献1に記載のような従来の関連語を検索する構成では、二連単語の候補対から前置単語のリストおよび後置単語のリストを生成して、それぞれに対する集合を生成する演算をしているため、演算負荷が比較的大きい問題がある。また、二連単語に基づいて関連語を検索しているため、所定の単語に対する類義語の検索漏れが生じるおそれがある。

【手続補正4】

【補正対象書類名】明細書

【補正対象項目名】0006

【補正方法】変更

【補正の内容】**【0006】**

本発明の類義語検索装置は、所定の単語に対する類義語を検索する類義語検索装置であって、前記所定の単語に関する単語情報を取得する単語情報取得手段と、単語に関する単語情報にその単語に関連してその単語を特徴付ける単語に関する関連単語情報を複数有する関連単語集合情報が1つのデータ構造として関連付けられた単語関連情報が複数格納されるテーブル構造の関連単語テーブルに基づき、前記取得した単語情報の所定の単語に対応する単語関連情報を取得する単語関連情報取得手段と、前記関連単語テーブルにおける前記取得した単語関連情報の関連単語集合情報に類似する関連単語集合情報を備えた単語関連情報を検索する関連単語検索手段と、この関連単語検索手段にて検索された前記単語関連情報の単語情報を取得して前記所定の単語に対する類義語として選出する類義語選出手段と、を具備したことを特徴とする。

【手続補正5】

【補正対象書類名】明細書

【補正対象項目名】0008

【補正方法】変更

【補正の内容】**【0008】**

そして、本発明では、請求項1に記載の類義語検索装置であって、前記単語関連情報取得手段は、複数の単語にて構成された文章に出現する単語情報に、その単語の文章における近傍に位置する他の単語に関する関連単語情報を複数有する関連単語集合情報が1つのデータ構造として関連付けられ、前記文章に出現する各単語についてそれぞれ構成された複数の単語関連情報が格納されるテーブル構造の関連単語テーブルに基づいて、前記単語関連情報を取得する構成とすることが好ましい。この構成では、文章における単語の近傍に位置する他の単語の集合は、その単語の機能や効能または作用あるいは効果、これら機能などを奏すものなどの比較的に関連する単語であるので、他の単語の集合をその単語に関連付けた単語関連情報を複数格納するテーブル構造の関連単語テーブルを利用してことで、所定の単語に対して関連する単語が検索されることとなる。このため、その他の単語の集合に関連付く単語を類義語としてすることで、単語にその単語に関連する単語を関連付けた簡単なテーブル構造を利用して容易に適切な類義語の検索が可能となる。

【手続補正6】

【補正対象書類名】明細書

【補正対象項目名】0029

【補正方法】変更

【補正の内容】**【0029】**

そして、本発明では、請求項19または請求項20に記載の情報検索装置であって、前記検索手段は、前記類義語検索装置にて選出した前記単語情報および前記単語情報取得手段で取得した単語情報の単語を、入力操作可能な操作手段における入力操作に応じて選択

および変更のうちの少なくともいずれか一方を実施可能に前記表示手段に表示させ、前記入力操作に応じて前記単語の選択および変更のうちの少なくともいずれか一方が実施されて選択された単語および変更された単語のうちの少なくともいずれか一方の単語に関する単語情報に基づいて検索を実行する構成とすることが好ましい。この構成では、検索手段により、類義語検索装置にて選出した単語情報および類義語検索装置の単語情報取得手段で取得した単語情報の単語を、入力操作可能な操作手段における入力操作に応じて選択および変更のうちの少なくともいずれか一方を実施可能に表示手段に表示させる。そして、検索手段は、入力操作に応じて選択や変更が実施されて選択された単語および変更された単語のうちの少なくともいずれか一方の単語に関する単語情報に基づいて、文章や文献の検索をする。このことにより、所定の単語や所望の類義語による所望の文章や文献が得られる。

【手続補正7】

【補正対象書類名】明細書

【補正対象項目名】0036

【補正方法】削除

【補正の内容】

【手続補正8】

【補正対象書類名】明細書

【補正対象項目名】0037

【補正方法】削除

【補正の内容】

【手続補正9】

【補正対象書類名】明細書

【補正対象項目名】0049

【補正方法】変更

【補正の内容】

【0049】

検索要求手段252は、キーワード設定手段251で取得した単語情報のキーワードに基づいて、このキーワードを含む文献情報の検索の実行をサーバ装置300に要求する旨の検索要求情報を生成する。すなわち、検索要求情報には、キーワードと、このキーワードを含む文献情報を検索する条件である検索条件情報と、検索を要求する利用者や端末装置200を特定する顧客情報などを有した端末特定情報とが、1つのデータ構造として関連付けられている。そして、検索要求手段252は、生成した検索要求情報を端末通信手段210によりネットワーク110を介してサーバ装置300へ送信する。

【手続補正10】

【補正対象書類名】明細書

【補正対象項目名】0086

【補正方法】変更

【補正の内容】

【0086】

この後、利用者が画面表示に従って端末入力手段220を適宜操作し、設定したキーワードを含む文献の検索を要求する旨の入力操作により、文献検索を実行させる。この文献検索の実行を要求する旨の入力操作としては、例えば表示画面中に設けられ文献検索の実行を促すコマンドボタンの入力操作などが例示できる。この入力操作により、端末処理手段250の検索要求手段252は、設定させたキーワードを含む文献検索の実行をサーバ装置300に要求する旨の検索要求情報を生成し(ステップS102)、端末通信手段210によりネットワーク110を介してサーバ装置300へ送信させる処理をする(ステップS103)。

【手続補正11】

【補正対象書類名】明細書

【補正対象項目名】0100

【補正方法】変更

【補正の内容】

【0100】

また、文章に関する文章情報を複数記憶する文献情報記憶領域341に基づいて単語関連情報343Aを複数格納するテーブル構造の関連単語テーブルを構築している。このため、検索対象の文章の集合となる文献を蓄積するデータベースである文献情報記憶領域341を利用するので、例えば用語辞典などの言語資源を利用する必要がなく、また例えば関連単語記憶領域343としてあらかじめ構築せずに類義語の検索の際に関連単語テーブルを構築することで関連単語テーブルを記憶しておく領域を削減でき、さらなる構成の簡略化が図れる。また、文章を複数含む文献毎の集合として文章情報を複数有する文献情報341Aを複数記憶するテーブル構造に構築された文献情報記憶領域341を利用している。このため、文献毎に文章情報がまとまった文献情報を複数記憶するテーブル構造の文献情報記憶領域341を利用して、文献で利用される積雪な類義語が検索され、文章情報毎に複数格納するテーブル構造に比してデータベースの構築や保守管理などが容易にできるとともに、取得したキーワードや選出した類義語に基づく文献の検索も容易にできる。さらに、文献情報の文章情報の文章における近傍に位置する他の単語を関連単語情報343A21として関連付けている。このため、キーワードを含む文献はそのキーワードに対する類義語の存在頻度が比較的に高いことから、全ての文章情報から検索する必要がなく、一部の文献から効率よく異義語が検索される。

【手続補正12】

【補正対象書類名】明細書

【補正対象項目名】0101

【補正方法】変更

【補正の内容】

【0101】

さらに、単語関連情報343Aの単語情報342A1として、形態素解析などにより文章から抽出した単語の情報としている。このため、既存の形態素解析などにより簡単に単語関連情報343Aを生成でき、簡単な処理で関連単語テーブルが構築され、類義語を検索するための構成や新たな単語関連情報343Aの追加や修正などの保守管理なども容易にできる。

【手続補正13】

【補正対象書類名】明細書

【補正対象項目名】0119

【補正方法】変更

【補正の内容】

【0119】

また、類義語情報を生成して複数記憶するテーブル構造の類義語情報記憶領域344を構築する構成について説明したが、類義語情報を生成しなくてもよい。また、類義語情報を生成する場合でも、アイコン表示をしなくてもよい。なお、このアイコン表示の際、例えばその類義語の利用頻度が高くなるに従って表示形態を変更してもよい。さらには、アイコンの操作により表示される類義語の集合として、類義語の削除や修正、追加などが可能としてもよく、これらの削除や修正、追加の内容を類義語情報記憶領域344に反映させたり、端末装置200のみに記憶させておいたりしてもよい。

【手続補正14】

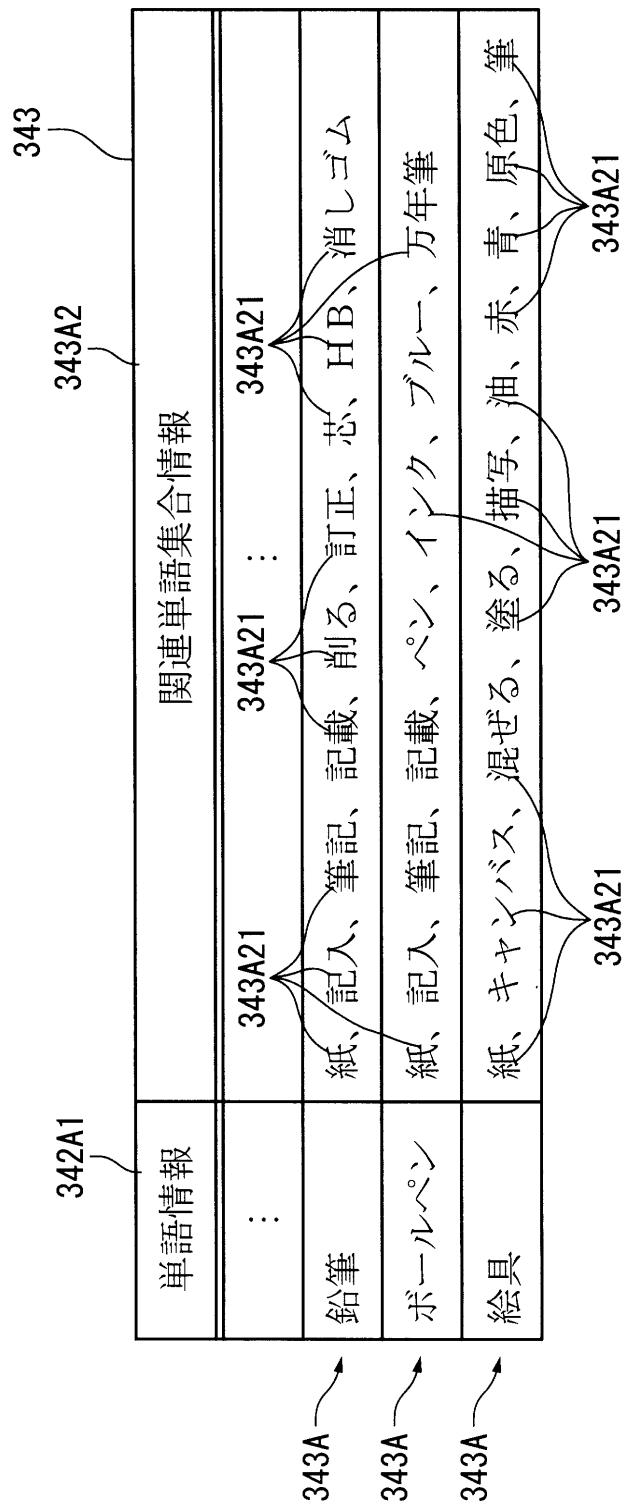
【補正対象書類名】図面

【補正対象項目名】図6

【補正方法】変更

【補正の内容】

【図6】



【手続補正15】

【補正対象書類名】図面

【補正対象項目名】図7

【補正方法】変更

【補正の内容】

【図7】

342A1		344				344A3	
344A2		類義語集合情報				アイコン情報	
:		:				:	
鉛筆		万年筆	ボールペン	筆記具	絵具	…	C02AQB
:							
344A21		344A21				:	
:		:				:	

344A ←