

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5701911号  
(P5701911)

(45) 発行日 平成27年4月15日(2015.4.15)

(24) 登録日 平成27年2月27日(2015.2.27)

(51) Int.Cl.

F I

G O 6 F 17/30 (2006.01)

G O 6 F 17/30 3 1 O A

G O 6 F 17/30 3 3 O C

G O 6 F 17/30 3 5 O C

請求項の数 18 (全 27 頁)

(21) 出願番号 特願2012-557043 (P2012-557043)  
 (86) (22) 出願日 平成23年3月9日(2011.3.9)  
 (65) 公表番号 特表2013-531282 (P2013-531282A)  
 (43) 公表日 平成25年8月1日(2013.8.1)  
 (86) 国際出願番号 PCT/US2011/000442  
 (87) 国際公開番号 W02011/112252  
 (87) 国際公開日 平成23年9月15日(2011.9.15)  
 審査請求日 平成25年7月31日(2013.7.31)  
 (31) 優先権主張番号 201010123209.9  
 (32) 優先日 平成22年3月10日(2010.3.10)  
 (33) 優先権主張国 中国(CN)  
 (31) 優先権主張番号 12/932,868  
 (32) 優先日 平成23年3月8日(2011.3.8)  
 (33) 優先権主張国 米国(US)

(73) 特許権者 510330264  
 アリババ・グループ・ホールディング・リ  
 ミテッド  
 ALIBABA GROUP HOLDI  
 NG LIMITED  
 英国領、ケイマン諸島、グランド・ケイマ  
 ン、ジョージ・タウン、ワン・キャピタル  
 ・プレイス、フォース・フロア、ビー・オ  
 ー、ボックス 847  
 (74) 代理人 110000028  
 特許業務法人明成国際特許事務所  
 (74) 代理人 100102989  
 弁理士 井上 佳知

最終頁に続く

(54) 【発明の名称】 クエリモデルに基づく誘導検索

(57) 【特許請求の範囲】

【請求項 1】

コンピュータによって実行される誘導検索の方法であって、新しいクエリを受信し、事前に記憶されたクエリと対応する誘導情報とのマッピング関係を示す複数のモデルを  
取得するために前記新しいクエリを処理し、前記複数のモデルを取得するために前記新し  
いクエリを処理することは、前記新しいクエリの、複数のワードを含む中心フレーズ又は中心ワードを特定するこ  
とを含み、前記モデルは、前記新しいクエリ、前記新しいクエリに基づき変換された情報、また  
はその双方から抽出された情報を含み、前記モデルは、前記新しいクエリを特徴付け、前記新しいクエリに関する前記複数のモデルの対応する複数の類似度を決定し、前記新  
しいクエリに関する前記複数のモデルの対応する複数の類似度を決定することは、前記新しいクエリに関する前記複数のモデルの内の一のモデルの類似度を、前記一の  
モデルにおけるモデルワードの特性、前記一のモデルにおけるスキップされたワードの特  
性、またはそれらの組合せに基づいて算出し、前記スキップされたワードが存在する場合には、前記一のモデルにおける前記スキッ  
プされたワードに基づく前記類似度から値を減じることを含み、前記値を減じることは、前記中心フレーズが特定された場合には、前記新しいクエリにおける前記中心フレ

10

20

ーズに対する前記スキップされたワードの距離に基づいて前記値を決定し、  
前記中心ワードが特定された場合には、前記新しいクエリにおける前記中心ワード  
に対する前記スキップされたワードの距離に基づいて前記値を決定することを含み、  
前記類似度に少なくとも部分的に基づいて前記複数のモデルの少なくとも一つを選択し  
、

事前に記憶されたクエリと対応する誘導情報の複数のマッピング関係を含むデータベ  
ースを検索するために前記選択されたモデルをインデックスとして使用して誘導情報を取得  
し、  
ユーザへの表示のために、前記取得された誘導情報を送信すること、  
を備える方法。

10

【請求項 2】

請求項 1 に記載の方法であって、  
前記複数のモデルを取得するために前記新しいクエリを処理することは、更に、前記中  
心フレーズ又は前記中心ワードを含む指定の長さのフレーズを抽出することを含む、方法  
。

【請求項 3】

請求項 1 に記載の方法であって、  
前記複数のモデルを取得するために、前記新しいクエリを処理することは、前記複数の  
モデルを生成するために前記新しいクエリの中の複数のワードをスキップすることを含む  
、方法。

20

【請求項 4】

請求項 3 に記載の方法であって、  
前記新しいクエリに関する前記複数のモデルの対応する複数の類似度を決定することは  
、前記スキップされたワードの品詞及び前記品詞に関連付けられた事前割り当てスコアに  
少なくとも部分的に基づいて、複数の類似度スコアを決定することを含む、方法。

【請求項 5】

請求項 1 に記載の方法であって、更に、  
前記複数のモデルをそれらの類似度にしたがって順位付けすることを備える方法。

【請求項 6】

請求項 1 に記載の方法であって、更に、  
選択された前記モデルに関連する取得された前記誘導情報の対応する複数の信頼度を決定  
することを備え、前記誘導情報を取得することは、更に、前記信頼度に少なくとも部分的  
に基づいて前記誘導情報の少なくとも幾つかを選択することを含む、方法。

30

【請求項 7】

請求項 6 に記載の方法であって、  
前記複数の信頼度は事前に記憶されたクエリ、ユーザ選択誘導情報およびユーザ選択中  
間情報の過去データに基づいて特定される、方法。

【請求項 8】

請求項 6 に記載の方法であって、  
前記複数の信頼度は事前に記憶されたクエリおよびユーザ選択誘導情報の過去データに  
基づいて特定され、事前に記憶されたクエリと前記ユーザ選択誘導情報が同時に発生する  
発生率に基づく確率値に少なくとも部分的に基づいて決定される、方法。

40

【請求項 9】

請求項 8 に記載の方法であって、  
前記複数の信頼度は、  
前記事前に記憶されたクエリの合計発生回数を与えられた際の前記事前に記憶された  
クエリと前記ユーザ選択誘導情報とが同時に発生する発生率に基づいて第 1 の確率を決定  
すること、

前記事前に記憶されたクエリの合計発生回数を与えられた際の前記事前に記憶された  
クエリとユーザ選択中間情報とが同時に発生する発生率に基づいて第 2 の確率を決定する

50

こと、

前記信頼度を決定するために前記第1の確率と前記第2の確率との加重和を計算すること、

によって決定される、方法。

【請求項10】

請求項1に記載の方法であって、更に、

マシン学習モードで誘導情報を予測することを備える方法。

【請求項11】

誘導検索を提供するためのシステムであって、

1つ又は2つ以上のプロセッサと、

前記1つ又は2つ以上のプロセッサに接続され、前記1つ又は2つ以上のプロセッサに命令を提供するように構成されている1つ又は2つ以上のメモリと、

を備え、

前記1つ又は2つ以上のプロセッサは、

新しいクエリを受信し、

事前に記憶されたクエリと対応する誘導情報とのマッピング関係を示す複数のモデルを取得するために前記新しいクエリを処理し、前記複数のモデルを取得するために前記新しいクエリを処理することは、

前記新しいクエリの、複数のワードを含む中心フレーズ又は中心ワードを特定することを含み、

前記モデルは、前記新しいクエリ、前記新しいクエリに基づき変換された情報、またはその双方から抽出された情報を含み、

前記モデルは、前記新しいクエリを特徴付け、

前記新しいクエリに関する前記複数のモデルの対応する複数の類似度を決定し、前記新しいクエリに関する前記複数のモデルの対応する複数の類似度を決定することは、

前記新しいクエリに関する前記複数のモデルの内の一のモデルの類似度を、前記一のモデルにおけるモデルワードの特性、前記一のモデルにおけるスキップされたワードの特性、またはそれらの組合せに基づいて算出し、

前記スキップされたワードが存在する場合には、前記一のモデルにおける前記スキップされたワードに基づく前記類似度から値を減じることを含み、前記値を減じることは、

前記中心フレーズが特定された場合には、前記新しいクエリにおける前記中心フレーズに対する前記スキップされたワードの距離に基づいて前記値を決定し、

前記中心ワードが特定された場合には、前記新しいクエリにおける前記中心ワードに対する前記スキップされたワードの距離に基づいて前記値を決定することを含み、

前記類似度に少なくとも部分的に基づいて前記複数のモデルの少なくとも一つを選択し、

事前に記憶されたクエリと対応する誘導情報の複数のマッピング関係を含むデータベースを検索するために前記選択されたモデルをインデックスとして使用して誘導情報を取得し、

ユーザへの表示のために前記取得された誘導情報を送信するように、

構成される、システム。

【請求項12】

請求項11に記載のシステムであって、

前記複数のモデルを取得するために前記新しいクエリを処理することは、前記複数のモデルを生成するために前記新しいクエリの中の複数のワードをスキップすることを含む、システム。

【請求項13】

請求項11に記載のシステムであって、

前記誘導情報を取得することは、前記データベースを検索するために前記複数のモデルの少なくとも幾つかのうちの選択された1つのモデルをインデックスとして使用すること

10

20

30

40

50

を含む、システム。

【請求項 1 4】

請求項 1 に記載のシステムであって、更に、

前記 1 つ又は 2 つ以上のプロセッサは、更に、前記複数のモデルに対する取得された前記誘導情報の対応する複数の信頼度を決定するように構成され、前記誘導情報を取得することは、更に、前記信頼度に少なくとも部分的に基づいて、前記誘導情報の少なくとも幾つかを選択することを含む、システム。

【請求項 1 5】

誘導検索を提供するためのコンピュータプログラムであって、

新しいクエリを受信するための機能と、

事前に記憶されたクエリと対応する誘導情報とのマッピング関係を示す複数のモデルを取得するために前記新しいクエリを処理するための機能と、前記複数のモデルを取得するために前記新しいクエリを処理する機能は、

前記新しいクエリの、複数のワードを含む中心フレーズ又は中心ワードを特定することを含み、

前記モデルは、前記新しいクエリ、前記新しいクエリに基づき変換された情報、またはその双方から抽出された情報を含み、

前記モデルは、前記新しいクエリを特徴付け、

前記新しいクエリに関する前記複数のモデルの対応する複数の類似度を決定するための機能と、前記新しいクエリに関する前記複数のモデルの対応する複数の類似度を決定するための機能は、

前記新しいクエリに関する前記複数のモデルの内の一のモデルの類似度を、前記一のモデルにおけるモデルワードの特性、前記一のモデルにおけるスキップされたワードの特性、またはそれらの組合せに基づいて算出し、

前記スキップされたワードが存在する場合には、前記一のモデルにおける前記スキップされたワードに基づく前記類似度から値を減じることを含み、前記値を減じるとは、

前記中心フレーズが特定された場合には、前記新しいクエリにおける前記中心フレーズに対する前記スキップされたワードの距離に基づいて前記値を決定し、

前記中心ワードが特定された場合には、前記新しいクエリにおける前記中心ワードに対する前記スキップされたワードの距離に基づいて前記値を決定することを含み、

前記類似度に少なくとも部分的に基づいて前記複数のモデルの少なくとも一つを選択するための機能と、

事前に記憶されたクエリと対応する誘導情報の複数のマッピング関係を含むデータベースを検索するために前記選択されたモデルをインデックスとして使用して誘導情報を取得するための機能と、

ユーザへの表示のために前記取得された誘導情報を送信するための機能と、

をコンピュータによって実現させるコンピュータプログラム。

【請求項 1 6】

請求項 1 に記載の方法において、前記ワードが存在する場合には、前記一のモデルにおける前記ワードの特性は品詞であり、

前記スキップされたワードが存在する場合には、前記一のモデルにおける前記スキップされたワードの特性は品詞である、方法。

【請求項 1 7】

請求項 1 に記載の方法において、前記モデルを取得することは、前記新しいクエリに対して N - G r a m 技術または S k i p - G r a m 技術を実行することを含む、方法。

【請求項 1 8】

請求項 1 に記載の方法において、前記ワードが存在する場合には、前記一のモデルにおける前記ワードの特性は品詞であり、

前記スキップされたワードが存在する場合には、前記一のモデルにおける前記スキップされたワードの特性は品詞であり、

前記品詞は、副詞、形容詞、数詞、動詞または名詞である、方法。

【発明の詳細な説明】

【技術分野】

【0001】

〔関連出願の相互参照〕

本出願は、あらゆる目的のために、参照によって本明細書に組み込まれる、2010年3月10日に出願された、発明の名称を「METHOD FOR IMPLEMENTING GUIDED SEARCH, GUIDED SEARCH SERVER AND INFORMATION PROCESSING SYSTEM (誘導検索を実行するための方法、誘導検索サーバ、及び情報処理システム)」とする中国特許出願第201010123209.9号に基づく優先権を主張する。

10

【0002】

本発明は、情報処理の技術に関し、特に、誘導検索に関する。

【背景技術】

【0003】

既存の誘導検索システムにおいて、ユーザは、入力テキストに関連する情報を取得するために、クライアントにおいて何らかのテキストをクエリとして入力するのが一般的である。クライアントから提供されるテキストを取得した後、誘導検索サーバは、そのテキストを解析し、ユーザのクエリ意図を特定し、ユーザによって入力されたテキストに関連する情報をクライアントに提供し、ユーザのクエリ範囲を狭め、ユーザが自身の必要とする情報を取り出せるようにする。

20

【0004】

目下のところ、誘導検索サーバは、一般に、クエリ情報と誘導情報との間におけるマッピング関係のマッピングテーブルを事前に記憶している。例えば、クエリ情報は、ユーザがクライアントにおいて入力するテキストであり、誘導情報は、ユーザが入力するテキストに関連する情報である。クエリ情報を取得した後、誘導検索サーバは、クエリ情報に対応する誘導情報をマッピングテーブルで検索し、マッピングテーブルから誘導情報が取り出された場合に、クエリ情報に対応する取り出された誘導情報をクライアントに伝送する。

【0005】

既存の誘導検索システムにおいて、誘導検索サーバは、ユーザが入力するクエリ情報に対応する誘導情報をマッピングテーブルから取り出し、ユーザが入力するクエリ情報がマッピングテーブルの中のクエリ情報と完全に一致する場合にのみ、取り出された誘導情報をクライアントに伝送するのが一般的である。ユーザが入力するクエリ情報が、マッピングテーブルの中のどのクエリ情報とも一致しない場合は、誘導検索サーバは、ユーザが入力するクエリ情報に対応する誘導情報をマッピングテーブルから取り出すことはできないのが一般的である。したがって、既存の誘導検索において、誘導検索サーバは、マッピングテーブルの中に存在しないクエリ情報に対応する誘導情報をクライアントに提供することはできないのが一般的である。

30

【図面の簡単な説明】

【0006】

発明の様々な実施形態が、以下の詳細な説明及び添付の図面で開示される。

40

【0007】

【図1】誘導検索技術を実行するネットワーク環境の一実施形態を示す概略図である。

【0008】

【図2】誘導検索を実施するためのプロセスの一実施形態を示すフローチャートである。

【0009】

【図3】クエリ情報に基づいて1つ又は2つ以上のモデルを導出するためのプロセスの一実施形態を示すフローチャートである。

【0010】

【図4】モデルの類似度スコアを計算するためのプロセスの一実施形態を示すフローチャ

50

ートである。

【 0 0 1 1 】

【図 5】プロセス 2 0 0 を実行するように構成された誘導検索サーバの一実施形態を示すブロック図である。

【 0 0 1 2 】

【図 6】誘導検索を実施するためのプロセスの別の一実施形態を示すフローチャートである。

【 0 0 1 3 】

【図 7】プロセス 6 0 0 を実行するように構成された誘導検索サーバの一実施形態を示すブロック図である。

【発明を実施するための形態】

【 0 0 1 4 】

本発明は、プロセス、装置、システム、合成物、コンピュータによって読み取り可能なストレージ媒体に実装されたコンピュータプログラム製品、並びに / 又は結合先のメモリに記憶された命令及び / 若しくは結合先のメモリによって提供される命令を実行するように構成されたプロセッサなどのプロセッサを含む、数々の形態で実装することができる。本明細書では、これらの実装形態、又は本発明がとりえるその他のあらゆる形態を、技術と称することができる。総じて、開示されたプロセスのステップの順序は、本発明の範囲内で可変である。別途明記されない限り、タスクを実施するように構成されるとして説明されるプロセッサ又はメモリなどのコンポーネントは、所定時にタスクを実施するように一時的に構成される汎用コンポーネントとして、又はタスクを実施するように製造された特殊コンポーネントとして実装することができる。本明細書で使用する「プロセッサ」という用語は、コンピュータプログラム命令などのデータを処理するように構成された 1 つ又は 2 つ以上のデバイス、回路、及び / 又は処理コアを言う。

【 0 0 1 5 】

本発明の原理を例示す添付の図面とともに、以下で、発明の 1 つ又は 2 つ以上の実施形態の詳細な説明が提供される。本発明は、このような実施形態との関連のもとで説明されているが、いかなる実施形態にも限定されない。本発明の範囲は、特許請求の範囲によってのみ限定され、本発明は、数々の代替形態、変更形態、及び均等物を内包している。以下の説明では、本発明の完全な理解を可能にするために、数々の具体的詳細が明記されている。これらの詳細は、例示を目的として提供されるものであり、本発明は、これらの詳細の一部又は全部を伴わずとも、特許請求の範囲にしたがって実施することができる。明瞭さを期するため、本発明に関連する技術分野において知られる技術要素は、本発明が不必要に不明瞭にされないように詳細な説明を省略される。

【 0 0 1 6 】

誘導検索技術が説明される。各種の実施形態において、事前に記憶されたクエリと、対応する誘導情報とのマッピング情報を示す 1 つ又は 2 つ以上のモデルが、新しいクエリに基づいて導出される。誘導情報の取り出しを促すために、作成されたモデルと、新しいクエリとの類似度が決定される。

【 0 0 1 7 】

図 1 は、誘導検索技術を実行するネットワーク環境の一実施形態を示す概略図である。

【 0 0 1 8 】

ユーザは、自身の検索クエリをクライアント 1 0 2 に入力し、クライアント 1 0 2 は、検索を実施するために、インターネットなどのネットワークを通じて誘導検索システム 1 0 1 とやり取りする。各種の実施形態において、クライアント及び誘導検索システムは、パソコン、サーバコンピュータ、スマートフォンなどの手持ち式のすなわち携帯型のデバイス、フラットパネルデバイス、マルチプロセッサシステム、マイクロプロセッサベースのシステム、セットトップボックス、プログラマブル家庭用電子機器、ネットワーク PC、ミニコンピュータ、大規模コンピュータ、特殊用途向けデバイス、上記のシステム若しくはデバイスの任意を含む分散コンピューティング環境、又は 1 つ若しくは 2 つ以上のブ

10

20

30

40

50

ロセッサと、該プロセッサに結合され該プロセッサに命令を提供するように構成されたメモリと、を含むその他のハードウェア/ソフトウェア/ファームウェアの組み合わせなどの、任意の適切なコンピューティングデバイスを使用して実装することができる。

#### 【0019】

クエリ情報は、1つ又は2つ以上の英数字列や記号などを含む要素を含むことができる。例えば、実際の応用では、環境は、電子商取引プラットフォームであり、クエリ情報は、ユーザが興味を持っている製品の名前及び属性を含むことができる。

#### 【0020】

誘導検索システムは、クエリ情報に基づいて誘導情報を検索し、取り出された誘導情報を、ユーザへの表示のためにクライアントに提示する。誘導情報をどのように取得するかの詳細が、以下で説明される。

10

#### 【0021】

本明細書で使用される誘導情報という用語は、ユーザがより迅速に且つより正確に所望の情報を取り出すことを助ける情報を言う。各種の実施形態において、誘導情報は、カテゴリをベースとした誘導情報（例えば、「電子機器」や「健康管理」など）や、属性をベースとした誘導情報（例えば、「赤」や「充電式バッテリーが含まれる」など）などを含む。例示を目的として、以下の実施例では、カテゴリをベースとした誘導情報が広く論じられるが、説明される技術は、属性をベースとした誘導情報又はその他のタイプの誘導情報にも適用することができる。

#### 【0022】

20

一部の実施形態では、（図では別々のデータベースとして示されているが、その他の実装形態では統合可能である）カテゴリ誘導データベース104又は属性誘導データベース106などの誘導情報データベースの中に、クエリと、対応するそれぞれの誘導情報との間における対応関係が記憶されている。対応関係は、過去情報（例えば、これまでのユーザクエリと、ユーザによる選択の結果のカテゴリとに関する記録）、システム構成（例えば、システム管理者によってセットアップされた、クエリ情報と、対応するカテゴリとの間におけるマッピング）、又はその他の任意の適切な技術に基づいて確立することができる。クエリと誘導情報との間におけるマッピング関係は、モデルとして言及される。一部の実施形態では、関係の強さを測定するために、すなわち、誘導情報によって指定される特定のカテゴリや属性などに対して特定のクエリがどの程度の可能性でマッピングされそ

30

#### 【0023】

例えば、或る電子商取引サイトが、製品モデル「N95」のNokiaによる携帯電話及び同じ製品モデルの3Mによるフェイスマスクを販売している。誘導情報データベースは、したがって、クエリ「N95」を、「携帯電話」、「電子機器」、及び「健康&医療」などのカテゴリに対してマッピングする。誘導情報データベースは、また、クエリを、「Nokia」及び「3M」（ブランド名）などの属性に対してもマッピングすることができる。

#### 【0024】

一部の実施形態では、要求したアイテムを検索するために、選択を行うユーザに対して複数の誘導情報が提示される。例えば、「N95」を求めるクエリを入力したユーザには、カテゴリ誘導選択肢「携帯電話」、「電子機器」、及び「健康&医療」を与えられるだろう。ユーザは、検索を遂行するために、自身の意図を最もよく表していると思われるカテゴリを選択する。例えば、検索エンジンは、興味を持たれている（1つ又は2つ以上の）製品を見つけるために、選択されたカテゴリ情報と、元のクエリとを組み合わせることができる。

40

#### 【0025】

図2は、誘導検索を実施するためのプロセスの一実施形態を示すフローチャートである。プロセス200は、図1の100などの誘導検索システムによって実施することができる。

50

## 【 0 0 2 6 】

ステップ 2 0 1 では、ユーザによって入れられたクエリ情報が、クライアントから取得される。各種の実施形態において、クエリ情報は、複数のワード、フレーズ、及び / 又は記号を含むことができる。

## 【 0 0 2 7 】

ステップ 2 0 2 では、1 つ又は 2 つ以上のモデルを導出するために、クエリ情報が処理される。

## 【 0 0 2 8 】

一部の実施形態では、クエリ情報を処理するために及びモデルを導出するために、自然言語処理技術が使用される。クエリ情報を処理してモデルを取得するための技術の例として、構文解析、N - G r a m s (例えば、クエリ情報に基づいて、Nワードのフレーズを導出する)、S k i p - G r a m s (例えば、クエリ情報に基づいて、ワードをスキップされたフレーズを導出する)などが挙げられる。

## 【 0 0 2 9 】

一部の実施形態では、誘導検索サーバは、幾つかのモデルを導出するために、複数の処理技術を使用してクエリ情報を解析する。モデルは、クエリ情報から抽出された及び / 又はクエリ情報に基づいて変換された情報であってクエリ情報を特徴付ける情報を含む。用いられる抽出技術に応じ、モデルは、クエリ情報の一部分及び / 又は翻訳されたクエリ情報を含むことができる。例えば、クエリ情報「F M トランスミッタを伴ったカー M P 3 プレーヤ」は、クエリ文字列の一部分を含む「カー M P 3 プレーヤ」、「カープレーヤ」、及び「F M トランスミッタ」などになるかもしれない。別の例では、同じクエリが、翻訳モデル「カーオーディオ機器」になるかもしれない。

## 【 0 0 3 0 】

ステップ 2 0 3 では、クエリに対するモデルの類似度が決定される。モデルの類似度スコアは、モデルがクエリ情報に類似する程度を表している。類似度は、モデルワードの特性及び / 又はスキップされたワードの特性 (例えば、品詞や単語間距離など) に基づいて計算することができる。一部の実施形態では、モデルの類似度スコアは、モデルの作成にあたってスキップされたワードに少なくとも部分的に基づいて計算される。各スキップされたワードは、ペナルティスコアをもたらす。類似度スコアを計算するために、スキップされたワードの全てのペナルティスコアの合計が計算される。各種の実装形態において、数々の類似度計算技術を用いることができ、その一例が、以下で示される。

## 【 0 0 3 1 】

ステップ 2 0 4 では、誘導検索サーバは、計算された類似度に基づいて、それぞれのモデルを順位付けする。それぞれのモデルは、類似度が高い順又は低い順に順位付けすることができる。一部の実施形態では、このステップは、随意に省略され、類似度に基づいて選択がなされる。

## 【 0 0 3 2 】

クエリ情報「F M トランスミッタを伴ったカー M P 3 プレーヤ」を例にとると、以下のモデル、「プレーヤ」(中心ワード)、「M P 3 プレーヤ」(中心フレーズ)、「カー M P 3 プレーヤ」(指定の長さ 3 を有するフレーズ)、及び「トランスミッタを伴ったカー M P 3 プレーヤ」(ワードをスキップされたフレーズ)が取得されるだろう。各モデルは、類似度スコアに関連付けられる。この例では、スコアが高いほど、モデルが元のクエリに類似している。モデルをスコアが高い順に順位付けすると、「トランスミッタを伴ったカー M P 3 プレーヤ」、「カー M P 3 プレーヤ」、「M S 3 プレーヤ」、及び「プレーヤ」になる。

## 【 0 0 3 3 】

ステップ 2 0 5 では、少なくとも幾つかのモデルに対応する誘導情報が、データベースにおいて探索される。例えば、検索キーワードとしてモデルを使用して、テーブルマッピングクエリ及び検索結果 (例えばカテゴリ) を探索することができる。一部の実施形態では、誘導検索サーバは、それぞれのモデルを類似度の高い順に順位付けする。この場合、

10

20

30

40

50



誘導検索サーバは、順位付けリストの中で1番目のモデルをキーワードとして使用することによって、モデルの中で最も高位のモデルに対応する誘導情報をデータベースにおいて検索することができる。順位付けリストのなかで1番目のモデルに対応する誘導情報が取り出されると、誘導検索サーバは、1番目のモデルに対応する取り出された誘導情報をクライアントに伝送することができ、そうでないならば、誘導検索サーバは、キーワードとしての順位付けリストのなかで2番目のモデルに対応する誘導情報をデータベースで検索することができ、モデルの1つに対応する誘導情報を取り出すまで、又は全てのモデルに対応する誘導情報をデータベースで検索するまで、以下同様に繰り返すことができる。

#### 【0034】

クエリ情報「トランスミッタ911を伴ったカーMP3プレーヤ」を例にとると、誘導検索サーバは、モデル：「プレーヤ」（中心ワード）、「MP3プレーヤ」（中心フレーズ）、「カーMP3プレーヤ」（指定の長さを有するフレーズ）、及び「トランスミッタを伴ったカーMP3プレーヤ」（ワードをスキップされたフレーズ）を抽出するであろう。類似度は、スコアによって表され、誘導検索サーバは、それぞれのモデルのスコアを計算しており、モデルは、スコアの高い順に、それぞれ、「トランスミッタを伴ったカーMP3プレーヤ」、「カーMP3プレーヤ」、「MS3プレーヤ」、及び「プレーヤ」になるだろう。したがって、誘導検索サーバは、モデル：「トランスミッタを伴ったカーMP3プレーヤ」、「カーMP3プレーヤ」、「MS3プレーヤ」、及び「プレーヤ」を、スコアの高い順に順位付けする。その後、誘導検索サーバは、まず、「トランスミッタを伴ったカーMP3プレーヤ」をインデックス/キーワードとして使用し、「トランスミッタを伴ったカーMP3プレーヤ」に対応するカテゴリ（誘導情報）をマッピングテーブルで検索し、対応するカテゴリが取り出されたならば、検索を終了させる。取り出されなかったならば、誘導検索テーブルは、「カーMP3プレーヤ」をキーワードとして使用し、「カーMP3プレーヤ」に対応するカテゴリ（誘導情報）をマッピングテーブルで検索し、対応するカテゴリが取り出されたならば、検索を終了させる。取り出されなかったならば、誘導検索テーブルは、「MP3プレーヤ」をキーワードとして使用し、「MP3プレーヤ」に対応するカテゴリ（誘導情報）をマッピングテーブルで検索し、対応するカテゴリが取り出されたならば、検索を終了させる。取り出されなかったならば、誘導検索テーブルは、「プレーヤ」をキーワードとして使用し、「プレーヤ」に対応するカテゴリ（誘導情報）をマッピングテーブルで検索し、対応するカテゴリが取り出されたならば、検索を終了させる。取り出されなかったならば、マッピングテーブルから、クエリ情報「トランスミッタ911を伴ったカーMP3プレーヤ」に対応するカテゴリは取り出されない。

#### 【0035】

別の例において、誘導検索サーバは、比較的類似度の高い幾つかのモデルに対応する誘導情報を検索することができる、すなわち、誘導検索サーバは、最も類似度の高いモデルに対応する誘導情報がマッピングテーブルから取り出されるかどうかに関係なく、比較的類似度の高い幾つかのモデルに対応する誘導情報をマッピングテーブルから検索することができる。

#### 【0036】

ステップ206では、取得された誘導情報は、ユーザへの表示のために、クライアントに送信される。誘導情報は、検索入力ボックスに関連付けられたプロンプト内に、クライアントのブラウザの表示エリア内に、又は提供される誘導情報に基づいて更なるユーザ選択がなされえるようなその他のフォーマットで表示することができる。

#### 【0037】

各種の実施形態において、誘導検索サーバは、取り出された誘導情報を、様々な条件下で様々なやり方でクライアントに伝送することができる。一部の実施形態では、マッピングテーブル記録は、誘導情報に対するクエリ情報の信頼度レベルを含み、これは、クエリ情報と誘導情報との間における関連性の程度を表している。このような信頼度は、確率又はスコア又は別の形態で表すことができる。信頼度スコアが高いほど、クエリ情報は、より強く誘導情報に関連付けられている。誘導検索サーバが、或るモデルに対応する誘導情

10

20

30

40

50

報を幾つか取り出したならば、サーバは、信頼度の最も高い誘導情報をクライアントに伝送することができる、又は誘導情報を信頼度の順に順位付けし、それぞれの誘導情報を順位付けの順でクライアントに伝送することができる。

【0038】

尚もクエリ情報「トランスミッタを伴ったカーMP3プレーヤ」を例にとり、誘導検索サーバが、モデル「トランスミッタを伴ったカーMP3プレーヤ」をインデックスとして使用し、マッピングテーブルから、2つのカテゴリ「音楽プレーヤ」及び「電子機器製品」、並びに「トランスミッタを伴ったカーMP3プレーヤ」に対するそれらのそれぞれの信頼度90%及び10%を取り出したならば、サーバは、実装形態に応じ、カテゴリ「音楽プレーヤ」のみをクライアントに伝送することができる、又は「プレーヤ」を「電子機器製品」よりも重要であるとして順位付けし、「プレーヤ」及び「電子機器製品」の両方をクライアントに伝送することができる。

10

【0039】

一部の実施形態では、類似度と信頼度との組み合わせにしたがって、クエリ結果がクライアントに返信される。例えば、クエリ情報を取得した後、誘導検索サーバは、クエリ情報を解析して複数のモデルを導出し、それぞれのモデルの類似度を計算し、それらの類似度にしたがってモデルを順位付けする。その後、サーバは、先ず、モデルのなかで1番目のモデルをキーワードとして使用し、その1番目のモデルに対応する誘導情報をマッピングテーブルで検索する。次いで、1番目のモデルに対応する複数の誘導情報、及び1番目のモデルに対するそれらの信頼度が取り出されたならば、サーバは、誘導情報のなかで信頼度の最も高い情報をクライアントに伝送する。

20

【0040】

図3は、クエリ情報に基づいて1つ又は2つ以上のモデルを導出するためのプロセスの一実施形態を示すフローチャートである。プロセス300は、プロセス200の202を実行するために使用することができる。

【0041】

ステップ302では、中心ワード、及び1つ又は2つ以上のワードを含む中心フレーズが抽出される。中心ワード及び中心フレーズは、構文解析及び文法規則に基づいてクエリから抽出される。中心ワード及び中心フレーズは、クエリ情報の基本的な意味を伝える。「トランスミッタを伴ったカーMP3プレーヤ」というクエリを例にとると、構文解析は、中心ワードが「プレーヤ」であること、及び中心フレーズが「MP3プレーヤ」又は「カーMP3プレーヤ」であることを明らかにする。

30

【0042】

ステップ304では、中心ワードに基づいて、クエリに対してN-Gramが実施される。言い換えると、クエリの中の中心ワードが維持され、連なるN個のワードがクエリから抽出される。一部の実施形態では、Nとして2が使用される。例えば、中心ワードが「プレーヤ」であるクエリ「トランスミッタを伴ったカーMP3プレーヤ」は、2-Gram「MP3プレーヤ」及び「を伴ったプレーヤ」になる。

【0043】

ステップ306では、中心フレーズが維持され、(ROUGE評価尺度に基づいて)Skip-Gramが抽出される。中心フレーズに対するスキップ距離がkである1つ又は2つ以上のフレーズが、クエリ情報から抽出される。事前設定が可能であるこのスキップ距離kは、合計でk又はk未満のスキップがN-Gramを構成することを可能にする。クエリ「トランスミッタを伴ったカーMP3プレーヤ」を例にとり、維持される中心フレーズが「カーMP3プレーヤ」であり尚且つスキップ距離が1であるならば、「を伴った」、「FM」、及び「トランスミッタ」をスキップして「カーMP3プレーヤFMトランスミッタ」、「トランスミッタを伴ったカーMP3プレーヤ」、又は「FMを伴ったカーMP3プレーヤ」を取得することが可能である。

40

【0044】

図4は、モデルの類似度スコアを計算するためのプロセスの一実施形態を示すフローチ

50

ャートである。プロセス400は、203を実行するために使用することができる。

【0045】

ステップ402～408では、元のクエリと比べてモデルでスキップされている各ワードに関連付けられたペナルティスコアが決定される。

【0046】

ステップ402では、スキップされたワードの品詞に基づいて、スキップされたワードの第1のペナルティスコアが決定される。元のクエリと比べてモデルでスキップされている各ワードの品詞は、例えば、字句解析技術を使用して決定される。一部の実施形態では、各品詞は、ペナルティスコア  $f_{pos}$  を事前に割り当てられる。一例では、副詞、形容詞、数詞、動詞、及び名詞が、それぞれ0.2、0.4、0.6、0.8、及び1のスコアを事前に割り当てられる。異なる実施形態では、異なる事前割り当て値を使用することができる。スキップされたワードの第1のペナルティスコアは、スキップされたワードの品詞に関連付けられた事前割り当て値に基づいて決定される。

10

【0047】

ステップ404では、元のクエリ情報内の中心ワードに対するスキップされたワードの距離に基づいて、第2のペナルティスコアが決定される。

【0048】

ステップ406では、元のクエリ情報内の中心フレーズに対するスキップされたワードの距離に基づいて、第3のペナルティスコアが決定される。

【0049】

20

ステップ404及びステップ406では、実装形態に応じ、様々な距離式を使用することができる。幾つかの例が、以下で説明される。

【0050】

ステップ408では、スキップされたワードについて、計算されたペナルティスコアの統合値が決定される。一部の実施形態では、スキップされたワードについて、スコアの平均が決定される。一部の実施形態では、スコアの加重平均が使用される。必要に応じ、その他の統合値を使用することもできる。

【0051】

スキップされた全てのワードについて、それらのそれぞれのスコアを取得するために、ステップ402～408を繰り返すことができる。ステップは、例に示されたのとは異なる順序で発生することもある。例では、順次処理が示されているが、ステップは、並列に繰り返すこともできる。

30

【0052】

ステップ410では、スキップされた全てのワードに関連付けられた統合スコアが決定されたときに、その統合ペナルティスコアに基づいて、モデルの類似度スコアが計算される。この例では、クエリ情報を決定するために、ペナルティスコアの合計が使用される。

【0053】

上記のプロセスは、元のクエリが6つのワード「w1 w2 w3 w4 w5 w6」を含む例を使用して説明される。

【0054】

40

上述された技術を使用して、モデル「w3 w5 w6」が取得され、クエリが、中心フレーズ「w4 w5 w6」及び中心ワード「w6」を有するとする。したがって、モデルを取得するために、元のクエリ情報「w1 w2 w3 w4 w5 w6」から「w1」、「w2」、及び「w4」がスキップされている。

【0055】

スキップされたワード「w1」、「w2」、及び「w4」について、ペナルティスコアが決定される。

【0056】

ステップ402では、スキップされたワード「w1」の品詞に基づいて、第1のペナルティスコアが決定される。この例では、w1は、副詞である。上述された、事前割り当て

50

スコアに対する品詞のマッピングの例に基づく、副詞に対応するスコアは、0.2になる。

【0057】

ステップ404では、クエリ情報内におけるスキップされたワードの場所と、クエリ情報内における中心ワードの場所との間の距離に基づく第2のペナルティスコアが、式  $f_{dw} = 1 - (d - 1) \times$  を使用して計算される。ここで、 $d$  は、クエリ内におけるスキップされたワードの場所と、クエリ内における中心ワードの場所との間の距離を表し、 $f_{pos}$ 、 $f_{dw}$ 、及び  $f_{dp}$  のなかでの  $f_{dw}$  の重みを表している。 $w1$  の場合、クエリ情報内における  $w1$  の場所と、クエリ情報内における  $w6$  (中心ワード) の場所との間の距離は、5であり、 $f_{dw}$  は、0.0002であるとされ、したがって、この場合、 $f_{dw} = 1 - (d - 1) \times$  10

【0058】

ステップ406では、クエリ内におけるスキップされたワードの場所と、クエリ情報内における中心フレーズの場所との間の距離に基づく第3のペナルティスコアが、式  $f_{dp} = 1 - d \times$  を使用して計算される。ここで、 $d$  は、クエリ内におけるスキップされたワードの場所と、クエリ情報内における中心フレーズの中央ワードの場所との間の距離を表し (一部の実施形態では、 $d$  は、中心フレーズが偶数のワードを有するならば四捨五入される)、 $f_{pos}$ 、 $f_{dw}$ 、及び  $f_{dp}$  のなかでの  $f_{dp}$  の重みを表している。一部の実施形態では、クエリ情報内におけるスキップされたワードの場所と、クエリ情報内における中心フレーズの場所との間の距離のスコアは、クエリ情報内における中心フレーズの最終ワードの場所と、クエリ情報内における中心フレーズの冒頭ワードの場所との差を計算し、当該差を2で割り算し、当該割り算の結果を整数に四捨五入し、当該整数に中心フレーズ内における冒頭サードの場所を加算し、そこからクエリ情報内におけるスキップされたワードの場所を減算し、最後にその絶対値をとることによって計算される。 $w1$  の場合、クエリ情報内における  $w1$  の場所と、クエリ情報内における  $w5$  (中心フレーズの中央ワード) の場所との間の距離は、4であり、 $f_{dp}$  は、0.0001であるとする。したがって、 $f_{dp} = 1 - d \times$  20

【0059】

この例では、式：スコア =  $(f_{pos} + f_{dw} + f_{dp}) / 3$  を使用して、スキップされたワードの統合スコアが計算される。したがって、ワード「 $w1$ 」をスキップされたクエリのスコアは、スコア1 =  $(0.2 + 0.9992 + 0.9996) / 3 = 0.7329$ である。30

【0060】

ステップ402 ~ 408は、スキップされたワード「 $w2$ 」の統合スコアを計算するために繰り返される。

【0061】

$w2$  は、品詞のスコアが  $f_{pos} = 1$  である名詞であるとする。

【0062】

クエリ情報内における  $w2$  の場所と、クエリ情報内における  $w6$  (中心ワード) の場所との間の距離は、4であり、 $f_{dw}$  は、0.0002であるとされ、この場合、 $f_{dw} = 1 - (d - 1) \times$  40

【0063】

クエリ情報内における  $w2$  の場所と、クエリ情報内における  $w5$  (中心フレーズの中心ワード) の場所との間の距離は、3であり、 $f_{dp}$  は、0.0001であるとされ、したがって、この場合、 $f_{dp} = 1 - d \times$  50

【0064】

スキップされたワード「 $w2$ 」の統合スコアは、したがって、 $(1 + 0.9994 + 0.9997) / 3 = 0.9997$ である。

【0065】

ステップ402 ~ 408は、スキップされたワード「 $w4$ 」の統合スコアを計算するた

めに繰り返される。

【 0 0 6 6 】

w 4 は、品詞に基づくスコアが  $f_{pos} = 0$  である数詞であるとする。

【 0 0 6 7 】

クエリ情報内における w 4 の場所と、クエリ情報内における w 6 (中心ワード) の場所との間の距離は、2 であり、 $d = 0.0002$  であるとされ、したがって、この場合、 $f_{dw} = 1 - (d - 1) \times 0 = 1 - (2 - 1) \times 0.0002 = 0.9998$  である。

【 0 0 6 8 】

クエリ情報内における w 4 の場所と、クエリ情報内における w 5 (中心フレーズの中心ワード) の場所との間の距離は、1 であり、 $d = 0.0001$  であるとされ、したがって、この場合、 $f_{dp} = 1 - d \times 0 = 1 - 1 \times 0.0001 = 0.9999$  である。

【 0 0 6 9 】

スキップされたワード「w 4」の統合スコアは、したがって、 $(0.6 + 0.9998 + 0.9999) / 3 = 0.8666$  である。

【 0 0 7 0 】

4 1 0 では、モデル「w 3 w 4 w 6」の類似度スコアは、したがって、 $0.7329 + 0.9997 + 0.8666 = 2.5992$  である。

【 0 0 7 1 】

プロセス 4 0 0 は、全てのモデルの類似度スコアを取得するために繰り返すことができる。

【 0 0 7 2 】

それぞれのモデルの類似度は、その他のやり方で取得することもできる。例えば、各モデルについてクエリ情報からスキップされたワードの数を計算し、ワードがスキップされるごとに 1 ずつスコアを増分し、したがって、スコアが高いほど類似度が低くなるようにすることができる。それぞれのモデルの類似度を計算するその他のやり方は、ここでは例として挙げられない。

【 0 0 7 3 】

一部の実施形態では、誘導情報は、その信頼度レベルに少なくとも部分的に基づいてユーザに送信される。各種の実施形態において、信頼度レベルは、様々なやり方で計算することができる。例えば、信頼度レベルは、手動で指定することができる。言い換えると、マッピングテーブルの中の信頼度は、手動で設定することができる。別の例では、誘導検索サーバが、クエリ情報と、対応する誘導情報とを記録された過去記録を取り出す。過去記録は、誘導検索サーバとは別のデータベースサーバに記憶することができ、誘導検索サーバは、過去記録が必要とされるときにデータベースサーバから過去記録を取り出すことができる。或いは、過去記録は、誘導検索サーバに記憶することができ、誘導検索サーバは、過去記録が必要とされるときに過去記録を自身の中からローカルに直接取り出すことができる。例えば、過去記録は、クエリ情報と、対応する誘導情報とが検索の中で同時に発生する回数、クエリ情報が単独で発生する回数、対応する誘導情報が単独で発生する回数などを記録している。値の決定は、以下でより詳しく説明される。誘導検索サーバ又は別のサーバが、ユーザによる選択の誘導情報をクライアントから更に取得したならば、その誘導検索サーバ又は別のサーバは、その誘導情報に対するクエリ情報の対応性を決定することができる。過去記録の中のそれぞれのクエリ情報について、誘導検索サーバは、クエリ情報がそれぞれの対応する誘導情報を伴うことの条件付き確率を、対応する誘導情報に対するそのクエリ情報の信頼度として計算する。

【 0 0 7 4 】

一部の実施形態では、過去記録は、クエリ情報と誘導情報との間における直接的対応性、及びクエリ情報と中間情報との間における対応性を記録している。中間情報は、クエリ情報及び誘導情報に対応しているが、クエリ情報と誘導情報とは、互いに対して直接的対応性を有していない。本明細書で使用される中間情報という用語は、或るクエリ情報及び誘導情報の両方から直接的にマッピングされ尚且つクエリ情報と誘導情報との間に非直接

10

20

30

40

50

的な対応関係を形成することができる同じ情報を言う。例えば、クエリ情報が何らかの特定の情報に対応し、尚且つ誘導情報が同じ情報に対応しているとすると、クエリ情報は、したがって、誘導情報に対応している。クエリ情報と誘導情報とが同時的に発生する回数対クエリ情報が発生する合計回数の比率が計算される。この確率は、条件付き確率 A と称される。或いは、誘導検索サーバは、クエリ情報と中間情報とが同時的に発生する回数対クエリ情報が発生する合計回数の比率を計算する。この比率は、条件付き確率 B と称される。その後、誘導検索サーバは、条件付き確率 A と、該条件付き確率 A に対応する重みとの積、及び条件付き確率 B と、該条件付き確率 B に対応する重みとの積を計算し、これら 2 つの積を加算して、クエリ情報が誘導情報を伴うことの条件付き確率とする。条件付き確率 A に対応する重み及び条件付き確率 B に対応する重みは、経験的データから又は線形回帰法によって決定することができる。

10

【 0 0 7 5 】

やはりカテゴリをベースとした誘導情報を例にとり、以下で、誘導情報に対するクエリ情報の信頼度をどのように決定するかの説明が提示される。

【 0 0 7 6 】

ユーザが、カテゴリ誘導データベースの中に対応するカテゴリ集合 { カテゴリ<sub>1</sub>、カテゴリ<sub>2</sub>、……、カテゴリ<sub>j</sub>、…… } を有するクエリ情報：クエリ<sub>i</sub>を入力したならば、そのクエリ<sub>i</sub>がカテゴリ<sub>j</sub>を伴うことの条件付き確率は、 $p(\text{カテゴリ}_j | \text{クエリ}_i)$  として定義される。それぞれのカテゴリについて、条件

【 数 1 】

20

$$\sum_j p(\text{カテゴリ}_j | \text{クエリ}_i) = 1$$

が満たされる。

【 0 0 7 7 】

提示された検索結果に応じて、ユーザは、提示された誘導情報の中のカテゴリを選択することができる、又はカテゴリの代わりにその他の情報を選択することができる。更に、例えばカテゴリに属する製品に関する情報などのその他の被選択情報と、カテゴリとの間にも対応関係があるかもしれないので、その他の被選択情報は、クエリと誘導情報とを非直接的にマッピングする中間情報である。誘導検索サーバ又は別のサーバは、その他の情報とカテゴリとの間におけるマッピング関係を記録されているマッピングテーブルを伴うように、事前に構成することができる。マッピングテーブルは、更に、リアルタイムで更新することができる。したがって、たとえユーザが、クエリ情報を入力した後にカテゴリではなくその他の中間情報を選択したとしても、誘導検索サーバ又は別のサーバは、クエリ情報とカテゴリとの間における対応関係を決定することができる。その他の被選択情報は、中間情報である。したがって、クエリ<sub>i</sub>がカテゴリ<sub>j</sub>を伴うことの条件付き確率が計算されるときは、クエリ情報とカテゴリとの間における直接的対応関係及び非直接的対応関係の両方が考慮されるだろう。

30

【 0 0 7 8 】

一実施形態では、過去データの中でクエリ<sub>i</sub>とカテゴリ<sub>j</sub>とが同時的に発生する回数決定される。ユーザがクエリ<sub>i</sub>を入力し、更にカテゴリ<sub>j</sub>を選択したならば、クエリ<sub>i</sub>とカテゴリ<sub>j</sub>とが同時的に発生する記録回数は、1 つ増分される。誘導検索サーバは、更に、クエリ<sub>i</sub>が発生する合計回数を取り出し、これは、それぞれのユーザが検索システムにクエリ<sub>i</sub>を入力する合計回数を意味する。誘導検索サーバは、クエリ<sub>i</sub>がカテゴリ<sub>j</sub>に直接的に対応することの条件付き確率を、

40

【 数 2 】

$$p'(\text{カテゴリ}_j | \text{クエリ}_i) = \frac{\text{カウント}(\text{カテゴリ}_j, \text{クエリ}_i)}{\text{カウント}(\text{クエリ}_i)}$$

として計算する。ここで、カウント ( カテゴリ<sub>j</sub> , クエリ<sub>i</sub> ) は、クエリ<sub>i</sub>とカテゴリ<sub>j</sub>とが

50

同時的に発生する回数を表し、カウント（クエリ<sub>i</sub>）は、それぞれのユーザがクエリ<sub>i</sub>を入力する合計回数を表している。

【0079】

一部の実施形態では、過去記録を取り出した後に、誘導検索サーバは、中間情報とカテゴリとの間におけるマッピング関係を記録されているマッピングテーブルを取り出す。マッピングテーブルは、データベースサーバ又は誘導検索サーバに記憶することができる。誘導検索サーバは、マッピングテーブルを取り出した後に、カテゴリ<sub>j</sub>のカテゴリに対応するそれぞれの中間情報を取り出す。その後、サーバは、クエリ<sub>i</sub>とそれぞれの中間情報とが同時的に発生する回数を過去記録から取り出す。ユーザがクエリ<sub>i</sub>を入力し、更に、特定の中間情報（例えば、クエリに直接的に対応しているカテゴリではなくその他の製品に関する情報）をクリックしたならば、クエリ<sub>i</sub>と中間情報とが同時的に発生する記録回数は、1つ増分される。更に、誘導検索サーバは、クエリ<sub>i</sub>が発生する合計回数を取り出す。この場合、誘導検索サーバは、クエリ<sub>i</sub>が中間情報に非直接的に対応することの条件付き確率を、

【数3】

$$p''(\text{カテゴリ}_j | \text{クエリ}_i, \text{提供}) = \frac{\text{カウント}(\text{カテゴリ}_j, \text{クエリ}_i | \text{提供})}{\text{カウント}(\text{クエリ}_i | \text{提供})}$$

として計算する。ここで、カウント（カテゴリ<sub>j</sub>，クエリ<sub>i</sub> | 提供）は、クエリ<sub>i</sub>と中間情報とが同時的に発生する回数を表し、中間情報は、カテゴリ<sub>j</sub>に対応する全ての中間情報

【0080】

導出された以上2つの条件付き確率は、カテゴリ<sub>j</sub>に対するクエリ<sub>i</sub>の信頼度を導出するために、線形的に加重される。

【数4】

$$p(\text{カテゴリ}_j | \text{クエリ}_i) = \lambda_1 p'(\text{カテゴリ}_j | \text{クエリ}_i) + \lambda_2 p''(\text{カテゴリ}_j | \text{クエリ}_i, \text{提供})$$

ここで、 $\lambda_1$ 及び $\lambda_2$ は、経験的に又は線形回帰技術を使用して決定することができる重みを表している。

【0081】

なお、プロセス200に示されている方法の実施形態は、発明の一実施形態にすぎず、当業者ならば、その他のヴァリエーションを導き出すことができる。

【0082】

一部の実施形態では、誘導検索サーバは、ステップ201が実施された後に、先ず、取得されたクエリ情報をキーワードとして使用し、該取得されたクエリ情報に対応する誘導情報を、クエリ情報と誘導情報との間におけるマッピング関係を記録されているマッピングテーブルで検索し、誘導情報が存在するならば、その取り出された誘導情報を直接的にクライアントに伝送し、そうでなければ、更にステップ202を実施する。更に、クエリ情報に対応する複数の誘導情報がマッピングテーブルから取り出され、尚且つそれぞれの誘導情報に対するクエリ情報の信頼度がマッピングテーブルに記録されているならば、誘導情報のなかで信頼度の最も高い情報のみがクライアントに伝送される。

【0083】

一部の実施形態では、誘導検索サーバは、マッピングテーブルの中のマッピング関係及び信頼度をリアルタイムで又は定期的に更新し、そうして、誘導情報に対するクエリの正確さを更に向上させる。

【0084】

一部の実施形態では、クエリ情報が、1つの文字、数、記号、又は省略されたフレーズである場合に、誘導検索サーバは、このようなクエリ情報を意味翻訳モードで少なくとも1つのモデルに変換することができる。例えば、誘導検索サーバがクエリ情報「LV」を取得したとすると、誘導検索サーバは、「LV」を「LOUIS VUITTON」に変

10

20

30

40

50

換することができる。実装の便宜を図るために、誘導検索サーバは、クエリ情報と変換後の情報との間におけるマッピング関係を記録されているマッピングテーブルを事前に記憶しておき、クエリ情報を取得した後に、このクエリ情報をキーワードとして使用して、変換後情報をマッピングテーブルで検索することができる。更に、マッピングテーブルは、クエリ情報と変換後情報との間におけるマッピング関係に加えて、変換後情報に対するクエリ情報の信頼度も記録することができる。この場合に、誘導検索サーバが、クエリ情報に対して複数の変換後情報をマッピングテーブルから取り出したならば、誘導検索サーバは、変換後情報のなかで信頼度の最も高い情報をモデルとして選択することができる。

【0085】

一部の実施形態では、誘導検索サーバが、モデルを使用してクエリ情報に対応する誘導情報を導出するならば、誘導検索サーバは、取得されたクエリ情報と導出された誘導情報との間におけるマッピング関係を、クエリ情報と誘導情報との間におけるマッピング関係を記録されているマッピングテーブルに記録することができる。サーバが、導出された誘導情報に対する取得されたクエリ情報の信頼度を計算するならば、サーバは、更に、信頼度をマッピングテーブルに記録することができる。したがって、誘導検索サーバは、同じクエリ情報を取得した後に、クエリ情報と誘導情報との間におけるマッピング関係を記録されているマッピングテーブルから対応する誘導情報を直接的に取り出して、クライアントに伝送される誘導情報を信頼度によって選択することができる。

【0086】

一部の実施形態では、誘導検索サーバが、クエリ情報に対応する複数の誘導情報を、複数のモデルと、それぞれの対応する誘導情報に対するそれぞれのモデルの信頼度とを使用して取り出すならば、誘導検索サーバは、モデルに対応する誘導情報に対する取得されたクエリ情報の、それぞれの信頼度を計算することができる。具体的には、誘導検索サーバが、取り出された特定の誘導情報が1つのモデルのみに対応すると決定したならば、誘導検索サーバは、誘導情報に対するモデルの信頼度を、誘導情報に対する取得されたクエリ情報の信頼度として捉えることができる。誘導検索サーバが、取り出された特定の誘導情報が複数のモデルに対応すると決定したならば、すなわち、同じ誘導情報に各々対応するモデルが複数あるならば、誘導検索サーバは、それぞれのモデルに対する誘導情報の信頼度を加算し、加算結果の平均を、誘導情報に対する取得されたクエリ情報の信頼度とすることができる。

【0087】

一部の実施形態では、誘導検索サーバが、1つのモデルのみを導出するならば、誘導検索サーバは、そのモデルをキーワード/インデックスとして使用して、すなわち、そのモデルを新しいクエリ情報として使用して、モデルに対応する誘導情報をマッピングテーブルで検索する。マッピングテーブルにモデルが記録されていないならば、誘導検索サーバは、モデルに対応する誘導情報を取り出すことができず、モデルに対応する誘導情報がマッピングテーブルに記録されているならば、誘導検索サーバは、モデルに対応する誘導情報を取り出すことができる。

【0088】

一部の実施形態では、誘導検索サーバが、モデルに対応する誘導情報を取り出さないならば、誘導検索サーバは、クエリ情報に対応する誘導情報をマシン学習モードで予測することができる。

【0089】

マシン学習モードでは、取得されたクエリ情報の特徴が特定の誘導情報に属することの確率分布を、過去のクエリ情報の学習及び訓練を通じて計算することができる。例えば、クエリ情報の特徴は、製品のラベル、ブランドのラベル、モデルのラベル、中心ワード、名詞句の集成、中心ワード及びその左隣の名詞句、中心ワード及びその右隣の名詞句などであってよい。

【0090】

具体的には、カテゴリがカテゴリCであるとする、現クエリ情報がカテゴリCに分類

10

20

30

40

50



されることの信頼度は、以下のように表すことができる、すなわち、 $x$ が、例えば製品のラベル、ブランドのラベル、モデルのラベル、中心ワード、名詞句の集成、中心ワード及びその左隣の名詞句、中心ワード及びその右隣の名詞句などの、現クエリ情報から抽出される所定の情報を表すときに、該所定の関連情報 $x$ を与える条件下で現クエリ情報がカテゴリ $C$ に属するとして決定されることの事後確率 $p(c \text{ は正しい} | x)$ として表すことができる。製品のラベルは、現クエリ情報が特定の製品情報であることを示しており、これは、意味特定によって導出される。ブランドのラベルは、現クエリ情報が特定のブランド情報であることを示しており、これは、意味特定によって導出される。モデルのラベルは、現クエリ情報が特定のモデル情報であることを示しており、これは、意味特定によって導出される。

10

【0091】

現クエリ情報から導出される特徴の重みは、従来の確率モデルで訓練することができ、現クエリ情報が特定のカテゴリであると決定されることの条件付き確率は、訓練から導出される対応する特徴値及び重みを使用して、従来の確率モデルで計算することができる。

【0092】

好ましくは、以下の式で表される最大エントロピーモデルが、条件付き確率モデルとして選択される。

【数5】

$$p(y|x) = \frac{1}{Z(x)} \exp\left(\sum_j \lambda_j f_j(x, y)\right)$$

20

ここで、 $y \in \{c \text{ は正しい}, c \text{ は正しくない}\}$ であり、 $x$ は、製品のラベル、ブランドのラベル、モデルのラベル、中心ワード、名詞句の集成、中心ワード及びその左隣の名詞句、中心ワード及びその右隣の名詞句などの、現クエリ情報に対応する所定の情報を表しており、 $f_j$ は、最大エントロピーモデルに対応する特徴値を表しており、例えば、 $x$ がワードレベルの特徴を表しているときに、特徴 $f_j$ は、訓練されたデータのなかで「現クエリ情報のキーワードが「フォン」を含むときは、現クエリ情報はカテゴリ $C$ に属する」の一致を得られるならば、「現クエリ情報のキーワードはフォンを含む」という条件下で値1をとり、そうでないならば、値0をとり、 $\lambda_j$ は、現クエリ情報のそれぞれの特徴に対応する重みを表しており、モデル訓練から導出することができ、 $Z(x)$ は、やはりモデル訓練から導出することができる正規化因子を表している。

30

【0093】

一部の実施形態では、条件付き確率モデルは、例えば線形回帰などのマシン学習モデルから作成される。実装形態によっては、サポートベクターマシンモデルを代わりに採用することができ、これは、クエリ情報とカテゴリとの間における関連性の程度を表す信頼度としてのスコアを計算するために、条件付き確率モデルに代わって使用することができる。

【0094】

一部の実施形態では、カテゴリを複数のレベルに分けることができ、したがって、マルチレベルの分類を採用することができる。例えば、「車両」が第1レベルのカテゴリである、すなわちルートカテゴリと呼ばれ、「車両」に続くレベルのカテゴリが「自動車」、「列車」、「飛行機」、及び「船」を含み、「自動車」に続くレベルのカテゴリが「車」及び「トラック」を含むとする。この場合、誘導検索サーバは、まず、複数の又は全部のルートカテゴリに対するクエリ情報の信頼度をマシン学習モードで計算し、その後、ルートカテゴリのなかで信頼度の最も高いカテゴリに続くレベルのカテゴリに対するクエリ情報の信頼度を計算し、ルートカテゴリのなかで信頼度の最も高いカテゴリに続くレベルのカテゴリに更に続くレベルのカテゴリに対するクエリ情報の信頼度を計算し、以下同様に続く。

40

【0095】

誘導情報（例えばカテゴリ）に対するクエリ情報の信頼度をマシン学習モードで導出し

50

た後、誘導検索サーバは、最も高い信頼度の又は比較的高い信頼度の誘導情報をクライアントに伝送することができる、又は複数の誘導情報を信頼度の順にクライアントに伝送することができる。当業者ならば、誘導検索サーバからクライアントに伝送するための誘導情報の選択を考え出すことが可能であり、ここでは例として挙げられない。

【0096】

もちろん、当業者ならば、図2に例示された方法の実施形態の説明をもとにその他の実施形態を考え出すことが可能であり、ここでは例として挙げられない。

【0097】

図5は、プロセス200を実行するように構成された誘導検索サーバの一実施形態を示すブロック図である。この例では、誘導検索サーバは、クエリ情報と誘導情報との間におけるマッピング関係を記憶されているマッピングテーブルを記憶するように適応されたストレージユニット501と、クライアントから提供されるクエリ情報を取得するように適応された取得ユニット502と、クエリ情報から抽出された又はクエリ情報の意味翻訳を実行することによって得られた情報であってクエリ情報を特徴付ける情報である少なくとも1つのモデルを導出するために、クエリ情報をモデル抽出アルゴリズムで解析するように適応された解析ユニット503と、解析ユニット503がモデルを導出したときに、それらのモデルがクエリ情報に類似する程度を表すそれぞれのモデルの類似度を計算するように適応された計算ユニット504と、それぞれのモデルを類似度の順に順位付けるように適応された順位付けユニット505と、それぞれのモデルを順位付けの順でキーワードとして使用し、それぞれのモデルに対応する誘導情報をマッピングテーブルで検索するための検索ユニット506と、取り出された誘導情報をクライアントに伝送するように適応された伝送ユニット507とを含む。

【0098】

解析ユニット503は、とりわけ、中心ワード及び/又は中心フレーズをクエリ情報から抽出するように適応することができる。

【0099】

解析ユニット503は、とりわけ、中心ワード又は中心フレーズを含む指定の長さのフレーズをクエリ情報から抽出するように適応することができる。

【0100】

モデルについて、計算ユニット504は、とりわけ、クエリ情報と比べてモデルにおいてスキップされたそれぞれのワードのスコアを計算するように及びそれらのスコアの合計をモデルの類似度として計算するように適応することができる。

【0101】

ストレージユニット501は、更に、品詞とスコアとの間におけるマッピング関係を記憶されているマッピングテーブルを記憶することができる。計算ユニット504は、スキップされたワードの品詞を決定するように適応された決定サブユニットと、ワードの品詞に対応するスコアを第1のスコアとしてマッピングテーブルで検索するように適応された検索サブユニットと、クエリ情報内におけるワードの場所と、クエリ情報内における中心ワードの場所との間の距離のスコアを第2のスコアとして計算するように適応された第1の計算サブユニットと、クエリ情報内におけるワードの場所と、クエリ情報内における中心フレーズの場所との間の距離のスコアを第3のスコアとして計算するように適応された第2の計算サブユニットと、クエリ情報と比べてワードをスキップされたモデルのスコアを、既定のスコアアルゴリズムで第1、第2、及び第3のスコアをもとにして計算するように適応された第3の計算サブユニットとを含むことができる。実際の応用では、当業者ならば、実際面での必要に応じて様々なスコアアルゴリズムを考え出すことができる。例えば、第1、第2、及び第3のスコアの平均が計算される、例えば、任意の2つのスコアの平均が計算される、例えば、任意の2つのスコアの平均が計算され、次いで、残りのスコアと加算される又は残りのスコアで乗算されるなどがある。もちろん、その他の形態のスコアアルゴリズムも可能であり、ここでは例として挙げられない。

【0102】

10

20

30

40

50

検索ユニット506は、モデルのなかで1番目のモデルをキーワードとして使用して、その1番目のモデルに対応する誘導情報をマッピングテーブルで検索し、検索ユニット506が、1番目のモデルに対応する誘導情報を取り出したならば、伝送ユニット507は、取り出された誘導情報をクライアントに伝送し、検索ユニット506が、1番目のモデルに対応する誘導情報を取り出さなかったならば、検索ユニット506は、モデルのなかで2番目のモデルをキーワードとして使用して、その2番目のモデルに対応する誘導情報をマッピングテーブルで検索し、検索ユニット506が、モデルの1つに対応する誘導情報を取り出すまで、又はそれぞれのモデルに対応する誘導情報をデータベースで検索するまで、以下同様に繰り返される。

【0103】

10

マッピングテーブルは、更に、クエリ情報と誘導情報との間における関連性の程度を表すものである誘導情報に対するクエリ情報の信頼度を記録することができ、検索ユニット506が、モデルに対応する複数の誘導情報を取り出したならば、伝送ユニット507は、取り出された複数の誘導情報のなかで信頼度の最も高い誘導情報をクライアントに伝送する、又は複数の誘導情報を順位付けしてそれらをクライアントに伝送する。

【0104】

取得ユニット502は、更に、クライアントから提供されるクエリ情報を取得する前に、クエリ情報及び対応する誘導情報を記録されている過去記録を取り出す。この場合、過去記録の中のそれぞれのクエリ情報について、計算ユニット504は、クエリ情報に対応するそれぞれの誘導情報を伴うことの条件付き確率を、対応する誘導情報に対するクエリ情報の信頼度として計算する。計算ユニット504は、信頼度をストレージユニット501に記憶することができる。

20

【0105】

過去記録が、クエリ情報と誘導情報との間、及びクエリ情報と中間情報との間における直接的な対応関係を記録しているならば、計算ユニット504は、クエリ情報と誘導情報とが同時に発生する回数対クエリ情報が発生する合計回数との比を条件付き確率Aとして計算するように適応された第4の計算ユニットと、クエリ情報と中間情報とが同時に発生する回数対クエリ情報が発生する合計回数の比率を条件付き確率Bとして計算するように適応された第5の計算ユニットと、条件付き確率Aと、該条件付き確率Aに対応する重みとの積、及び条件付き確率Bと、該条件付き確率Bに対応する重みとの積を計算し、これら2つの積を加算して、クエリ情報が誘導情報を伴うことの条件付き確率を導出するように適応された、第6の計算ユニットとを含むことができる。

30

【0106】

図5に示される誘導検索サーバは、更に、検索ユニット506がモデルに対応する誘導情報を取り出さなかったときに、クエリ情報に対応する誘導情報をマシン学習モードで予測するために、検索ユニット506によってトリガされるように適応された、マシン学習ユニット508を含むことができる。

【0107】

図5に示される誘導検索サーバの実施形態は、図2に示されたプロセスの実施形態に対応しているので、図5に示された誘導検索サーバの中のそれぞれのユニットの機能及び相互的な協調関係の詳細については、図2に示された方法の実施形態における関連の説明を参照することができ、したがって、説明の繰り返しは、ここでは省略される。

40

【0108】

誘導検索サーバは、情報処理システムに配備することができるので、本発明の一実施形態は、更に、情報処理システムを含む。図1に示されるように、情報処理システムは、クライアント102と、誘導検索サーバ101とを含み、クライアント102は、誘導検索サーバ101と情報のやり取りを実施するように適応され、誘導検索サーバ101は、クライアント102から提供されるクエリ情報を取得し、クエリ情報から抽出された又はクエリ情報の意味翻訳を実行することによって得られた情報であってクエリ情報を特徴付ける情報である少なくとも1つのモデルを導出するためにモデル抽出アルゴリズムでクエリ

50

情報を解析し、複数のモデルが導出されたならば、それらのモデルがクエリ情報に類似する程度を表すためにそれぞれのモデルの類似度を計算し、それぞれのモデルを類似度の順に順位付け、それらのモデルを順位付けの順でキーワードとして使用して、それぞれのモデルに対応する誘導情報を、クエリ情報と誘導情報との間におけるマッピング関係を記憶されているマッピングテーブルで検索し、取り出された誘導情報をクライアント 102 に伝送するように適応される。

#### 【0109】

誘導検索サーバ 101 及びクライアント 102 の具体的な機能、並びにそれらの相互的な協調関係の詳細については、図 2 に示された方法の実施形態及び図 5 に示された方法の実施形態における関連の説明を参照することができ、したがって、説明の繰り返しは、こ

10

#### 【0110】

図 6 は、誘導検索を実施するためのプロセスの別の一実施形態を示すフローチャートである。プロセス 600 は、図 1 の 100 などの誘導検索システムによって実施することができる。

#### 【0111】

ステップ 601 では、クライアントから提供されるクエリ情報が取得される。このステップは、プロセス 200 の 201 と同様である。

#### 【0112】

ステップ 602 において、誘導検索サーバは、クエリ情報に対応する誘導情報をモード

20

#### 【0113】

(1) で検索し、クエリ情報に対応する誘導情報をマシン学習モードで予測する。  
とりわけモード (1) では、誘導検索サーバは、少なくとも 1 つのモデルを導出するためにクエリ情報を解析し、複数のモデルが導出されたならば、誘導検索サーバは、それらのモデルがクエリ情報に類似する程度を表すそれぞれのモデルの類似度を計算し、それぞれのモデルを類似度の順に順位付けし、それらのモデルを順位付けの順でキーワードとして使用して、それぞれのモデルに対応する誘導情報を、クエリ情報と誘導情報との間におけるマッピング関係を記憶されているマッピングテーブルで検索する。

#### 【0114】

モード (1) の詳細については、プロセス 200 のステップ 202 ~ 205 の関連の説明を参照することができ、したがって、説明の繰り返しは、ここでは省略される。

30

#### 【0115】

誘導検索サーバが、クエリ情報に対応する誘導情報をモード (1) で取り出したならば、603 において、誘導検索サーバは、2 つのモードで取り出された及び予測された誘導情報をもとに、クライアントに提供される誘導情報を決定し、クライアントに提供されるものと決定されたその誘導情報をクライアントに伝送する。

#### 【0116】

実際の応用では、誘導検索サーバは、クライアントに提供される誘導情報を様々なやり方で決定することができる。例えば、2 つのモードで取り出された及び予測された同じ誘導情報が、クライアントに伝送される。例えば、それぞれのモードで取り出された及び予測された全ての誘導情報が、クライアントに伝送される。別の例では、最も信頼度の高い誘導情報を、上述のように信頼度モードでクライアントに伝送することができる。更なる例では、複数の誘導情報が順位付けされてクライアントに伝送される。

40

#### 【0117】

図 7 は、プロセス 600 を実行するように構成された誘導検索サーバの一実施形態を示すブロック図である。この例では、誘導検索サーバは、クエリ情報と誘導情報との間におけるマッピング関係を記憶されているマッピングテーブルを記憶するように適応されたストレージユニット 701 と、クライアントから提供されるクエリ情報を取得するように適応された取得ユニット 702 と、クエリ情報に対応する誘導情報をモード (1) で検索するように適応された第 1 の検索ユニット 703 と、クエリ情報に対応する誘導情報をマシ

50

ン学習モードで予測するように適応された第2の検索ユニット704と、第1の検索ユニット703が、クエリ情報に対応する誘導情報をモード(1)で取り出すならば、第1の検索ユニット703によって導出された誘導情報及び第2の検索ユニット704によって予測された誘導情報をもとに、クライアントに提供される誘導情報を決定するように適応された誘導情報決定ユニット705と、クライアントに提供されるものと決定された誘導情報をクライアントに伝送するように適応された伝送ユニット706とを含む。とりわけモード(1)では、誘導検索サーバは、少なくとも1つのモデルを導出するためにクエリ情報を解析し、複数のモデルが導出されたならば、誘導検索サーバは、それらのモデルがクエリ情報に類似する程度を表すためにそれぞれのモデルの類似度を計算し、それぞれのモデルを類似度の順に順位付け、それらのモデルを順位付けの順でキーワードとして使用して、それぞれのモデルに対応する誘導情報を、クエリ情報と誘導情報との間におけるマッピング関係を記憶されているマッピングテーブルで検索する。

10

#### 【0118】

図7に示された誘導検索サーバの実施形態は、図6に示される方法の実施形態に対応しているので、図7に示された誘導検索サーバの中のそれぞれのユニットの機能及び相互的な協調関係の詳細については、図6に示される方法の実施形態における関連の説明を参照することができ、したがって、説明の繰り返しは、ここでは省略される。

#### 【0119】

誘導検索サーバは、情報処理システムに配備することができるので、本発明の実施形態は、更に、情報処理システムを含む。図1に示されるように、情報処理システムは、クライアント102と、誘導検索サーバ101とを含み、クライアント102は、誘導検索サーバ101と情報のやり取りを実施するように適応され、誘導検索サーバ101は、クライアント102から提供されるクエリ情報を取得し、クエリ情報に対応する誘導情報をモード(1)で検索するとともにクエリ情報に対応する誘導情報をマシン学習モードで予測し、クエリ情報に対応する誘導情報がモード(1)で取り出されたならば、2つのモードで取り出された及び予測された誘導情報をもとに、クライアントに提供される誘導情報を決定し、クライアントに提供されるものと決定された誘導情報をクライアントに伝送するように適応される。とりわけモード(1)では、誘導検索サーバは、クエリ情報から抽出された又はクエリ情報の意味翻訳を実行することによって得られた情報であってクエリ情報を特徴付ける情報である少なくとも1つのモデルを導出するためにモデル抽出アルゴリズムでクエリ情報を解析し、複数のモデルが導出されたならば、誘導検索サーバは、それらのモデルがクエリ情報に類似する程度を表すためにそれぞれのモデルの類似度を計算し、それぞれのモデルを類似度の順に順位付け、それらのモデルを順位付けの順でキーワードとして使用して、それぞれのモデルに対応する誘導情報をクエリ情報と誘導情報との間におけるマッピング関係を記憶されているマッピングテーブルで検索する。

20

30

#### 【0120】

誘導検索サーバ101及びクライアント102の具体的な機能、並びにそれらの相互的な協調関係の詳細については、図6に示される方法の実施形態及び図7に示される誘導検索サーバの実施形態における関連の説明を参照することができ、したがって、説明の繰り返しは、ここでは省略される。

40

#### 【0121】

上述のユニットは、1つ又は2つ以上の汎用プロセッサ上で実行されるソフトウェアコンポーネントとして、プログラマブルロジックデバイス及び/若しくは所定の機能を実施するように設計された特殊用途向け集積回路などのハードウェアとして、又はそれらの組み合わせとして実装することができる。一部の実施形態では、ユニットは、本発明の実施形態で説明された方法をコンピュータデバイス(パソコン、サーバ、ネットワーク機器など)に実行させるための幾つかの命令を含み且つ不揮発性のストレージ媒体(光ディスク、フラッシュストレージデバイス、モバイルハードディスクなど)に記憶させることができるソフトウェア製品の形で具現化することができる。ユニットは、1つのデバイスに実装する又は複数のデバイスに分散させることができる。ユニットの機能は、互いに合体さ

50

せる又は複数のサブユニットに更に分けることができる。

【0122】

まとめると、本発明の実施形態では、誘導検索サーバは、クライアントから提供されるクエリ情報を取得した後、少なくとも1つのモデルを導出するためにクエリ情報を解析し、複数のモデルが導出されたならば、誘導検索サーバは、それぞれのモデルの類似度を計算し、類似度の順にそれぞれのモデルを順位付けし、それらのモデルをキーワードとして使用して、それぞれのモデルに対応する誘導情報を、クエリ情報と誘導情報との間のマッピング関係を記録されているマッピングテーブルで検索する。この場合、たとえクライアントから取得されたクエリ情報がマッピングテーブルの中に無いとしても、マッピングテーブルの中に(1つ又は2つ以上の)モデルが存在している可能性があるゆえに、尚も適切な誘導情報を取り出すことが可能であり、ゆえに、マッピングテーブルの中に無いクエリ情報に対応する誘導情報を誘導検索サーバがクライアントに提供できない事態に対処することができる。

10

【0123】

更に、誘導検索サーバが、クエリ情報と誘導情報との間におけるマッピング関係を記録されているマッピングテーブルからモデルに対応する誘導情報を取り出さなかったならば、誘導検索サーバは、取得されたクエリ情報に対応する誘導情報をマシン学習モードで導出することができ、そうして、クライアントから提供されるクエリ情報に対して対応する誘導情報を誘導検索サーバが返せるように更に保証することができる。

20

【0124】

更に、誘導検索サーバは、(1つ又は2つ以上の)モデルを使用して誘導情報を検索することに加えて、取得されたクエリ情報に対応する誘導情報をマシン学習モードで導出し、2つのモードで取り出された及び予測された誘導情報をもとに、クライアントに返される誘導情報を決定することができ、そうして、クライアントに返される誘導情報の理解度及び正確さを更に向上させる。

【0125】

実施形態の以上の説明から明らかなように、当業者ならば、本発明が、基本的な汎用ハードウェアプラットフォームと組み合わせられたソフトウェアの形で実装可能であることを、はっきり理解することができる。このような理解に基づき、先行技術に寄与する本発明による技術的解決策又はその一部は、基本的に、本発明のそれぞれの実施形態にしたがった又は本発明の実施形態の一部にしたがった方法をコンピュータデバイス(パソコン、サーバ、ネットワークデバイスなどでありえる)に実施させるための幾つかの命令を含み尚且つ例えばROM/RAM、磁気ディスク、光ディスクなどのストレージ媒体に記憶可能であるソフトウェア製品の形で具現化することができる。

30

【0126】

本明細書におけるそれぞれの実施形態は、進行的に説明されており、それらの共通の又は同様の部分は、相互に参照可能であり、各実施形態は、(1つ又は2つ以上の)その他の実施形態からのその(1つ又は2つ以上の)違いに重きを置かれている。特に、システムの実施形態の場合は、方法の実施形態と実質的に同様でありそれゆえに簡単に説明されており、その詳細に関しては、方法の実施形態の説明のそれらの部分を参照することができる。

40

【0127】

本発明は、例えば、パソコン、サーバコンピュータ、手持ち式のすなわち携帯型のデバイス、フラットパネルデバイス、マルチプロセッサシステム、マイクロプロセッサベースのシステム、セットトップボックス、プログラマブル家庭用電子機器、ネットワークPC、ミニコンピュータ、大型コンピュータ、上記のシステム若しくはデバイスの任意を含む分散コンピューティング環境などの、汎用又は専用の、数々のコンピュータシステム環境又はコンピュータシステム構成に適用することができる。

【0128】

本発明は、コンピュータによって実行される例えばプログラムモジュールなどのコンピ

50

ユーザ実行可能命令を一般的背景として説明することができる。一般に、プログラムモジュールは、特定のタスクを実行する又は特定の抽象データ型を実装するルーチン、プログラム、オブジェクト、コンポーネント、データ構造などを含む。或いは、本発明は、通信ネットワークを通じて接続されたりリモート処理デバイスによってタスクを実施され尚且ストレージデバイスを含むローカル若しくはリモートのコンピュータストレージ媒体の中にプログラムモジュールを設置可能である分散コンピューティング環境において実行することができる。

【 0 1 2 9 】

本発明は、その実施形態との関連で説明されてきたが、当業者ならば、本発明の趣旨から逸脱することなく数々の変更及びヴァリエーションが可能であることがわかり、本発明は、その特許請求の範囲内にこれらの変更及びヴァリエーションを含むように意図されている。

【 0 1 3 0 】

以上の実施形態は、理解を明瞭にする目的で幾らか詳細に説明されてきたが、本発明は、提供される詳細に限定されない。本発明の実現には、多くの代替的手法がある。開示された実施形態は、例示のためであって、限定的なものではない。

適用例 1：誘導検索の方法であって、新しいクエリを受信し、事前に記憶されたクエリと対応する誘導情報とのマッピング関係を示す複数のモデルを取得するために前記新しいクエリを処理し、前記新しいクエリに関する前記複数のモデルの対応する複数の類似度を決定し、事前に記憶されたクエリと対応する誘導情報との複数のマッピング関係を含むデータベースの中から誘導情報を取得することであって、前記類似度に少なくとも部分的に基づいて前記複数のモデルの少なくとも幾つかを選択し、及び前記データベースを検索するために前記選択されたモデルをインデックスとして使用することを含み、ユーザへの表示のために、前記取得された誘導情報を送信すること、を備える方法。

適用例 2：適用例 1 に記載の方法であって、前記複数のモデルを取得するために前記新しいクエリを処理することは、前記新しいクエリの、複数のワードを含む中心フレーズ又は中心ワードを特定することを含む、方法。

適用例 3：適用例 2 に記載の方法であって、前記複数のモデルを取得するために前記新しいクエリを処理することは、更に、前記中心フレーズ又は前記中心ワードを含む指定の長さのフレーズを抽出することを含む、方法。

適用例 4：適用例 1 に記載の方法であって、前記複数のモデルを取得するために、前記新しいクエリを処理することは、前記複数のモデルを生成するために前記新しいクエリの中の複数のワードをスキップすることを含む、方法。

適用例 5：適用例 4 に記載の方法であって、前記新しいクエリに関する前記複数のモデルの対応する複数の類似度を決定することは、前記スキップされたワードの品詞及び前記品詞に関連付けられた事前割り当てスコアに少なくとも部分的に基づいて、複数の類似度スコアを決定することを含む、方法。

適用例 6：適用例 4 に記載の方法であって、前記新しいクエリに関する前記複数のモデルの対応する複数の類似度を決定することは、前記新しいクエリの中心ワードに対する前記スキップされたワードの距離に少なくとも部分的に基づいて複数の類似度スコアを決定することを含む、方法。

適用例 7：適用例 4 に記載の方法であって、前記新しいクエリに関する前記複数のモデルの対応する複数の類似度を決定することは、前記新しいクエリの中心フレーズに対する前記スキップされたワードの距離に少なくとも部分的に基づいて複数の類似度スコアを決定することを含む、方法。

適用例 8：適用例 1 に記載の方法であって、更に、前記複数のモデルをそれらの類似度にしたがって順位付けすることを備える方法。

適用例 9：適用例 1 に記載の方法であって、前記誘導情報を取得することは、前記データベースを検索するために、前記複数のモデルの少なくとも幾つかのうちの選択された 1 つのモデルをインデックスとして使用することを含む、方法。

適用例 10：適用例 1 に記載の方法であって、更に、前記新しいクエリに関する前記複数のモデルの対応する複数の信頼度を決定することを備え、前記誘導情報を取得することは、更に、前記信頼度に少なくとも部分的に基づいて前記複数のモデルの少なくとも幾つかを選択することを含む、方法。

適用例 11：適用例 10 に記載の方法であって、前記複数のモデルの前記信頼度は過去データに基づいて特定される、方法。

適用例 12：適用例 10 に記載の方法であって、前記複数のモデルのうちの 1 つのモデルの信頼度は過去データに基づいて特定され、前記事前に記憶されたクエリに関連付けられた確率値に少なくとも部分的に基づいて決定される、方法。

適用例 13：適用例 12 に記載の方法であって、前記信頼度は、前記事前に記憶されたクエリの合計発生回数を与えられた際の前記事前に記憶されたクエリと前記誘導情報とが同時に発生する発生率に基づいて第 1 の確率を決定すること、前記事前に記憶されたクエリの合計発生回数を与えられた際の前記事前に記憶されたクエリと中間情報とが同時に発生する発生率に基づいて第 2 の確率を決定すること、前記信頼度を決定するために前記第 1 の確率と前記第 2 の確率との加重和を計算すること、によって決定される、方法。

適用例 14：適用例 1 に記載の方法であって、更に、マシン学習モードで誘導情報を予測することを備える方法。

適用例 15：誘導検索を提供するためのシステムであって、1 つ又は 2 つ以上のプロセッサと、前記 1 つ又は 2 つ以上のプロセッサに接続され、前記 1 つ又は 2 つ以上のプロセッサに命令を提供するように構成されている 1 つ又は 2 つ以上のメモリと、を備え、前記 1 つ又は 2 つ以上のプロセッサは、新しいクエリを受信し、事前に記憶されたクエリと対応する誘導情報とのマッピング関係を示す複数のモデルを取得するために前記新しいクエリを処理し、前記新しいクエリに関する前記複数のモデルの対応する複数の類似度を決定し、事前に記憶されたクエリと対応する誘導情報との複数のマッピング関係を含むデータベースの中から誘導情報を取得することであって、前記類似度に少なくとも部分的に基づいて前記複数のモデルの少なくとも幾つかを選択すること、及び前記データベースを検索するために前記選択されたモデルをインデックスとして使用することを含み、ユーザへの表示のために前記取得された誘導情報を送信するように、構成される、システム。

適用例 16：適用例 15 に記載のシステムであって、前記複数のモデルを取得するために前記新しいクエリを処理することは、前記新しいクエリの、複数のワードを含む中心フレーズ又は中心ワードを特定することを含む、システム。

適用例 17：適用例 15 に記載のシステムであって、前記複数のモデルを取得するために前記新しいクエリを処理することは、前記複数のモデルを生成するために前記新しいクエリの中の複数のワードをスキップすることを含む、システム。

適用例 18：適用例 15 に記載のシステムであって、前記誘導情報を取得することは、前記データベースを検索するために前記複数のモデルの少なくとも幾つかのうちの選択された 1 つのモデルをインデックスとして使用することを含む、システム。

適用例 19：適用例 15 に記載のシステムであって、更に、前記 1 つ又は 2 つ以上のプロセッサは、更に、前記新しいクエリに関する前記複数のモデルの対応する複数の信頼度を決定するように構成され、前記誘導情報を取得することは、更に、前記信頼度に少なくとも部分的に基づいて、前記複数のモデルの少なくとも幾つかを選択することを含む、システム。

適用例 20：誘導検索を提供するためのコンピュータプログラム製品であって、コンピュータによって読み取り可能なストレージ媒体に実装され、新しいクエリを受信するためのコンピュータ命令と、事前に記憶されたクエリと対応する誘導情報とのマッピング関係を示す複数のモデルを取得するために前記新しいクエリを処理するためのコンピュータ命令と、前記新しいクエリに関する前記複数のモデルの対応する複数の類似度を決定するためのコンピュータ命令と、事前に記憶されたクエリと対応する誘導情報との複数のマッピング関係を含むデータベースの中から誘導情報を取得するためのコンピュータ命令であって、前記類似度に少なくとも部分的に基づいて前記複数のモデルの少なくとも幾つかを選

10

20

30

40

50



択すること、及び前記データベースを検索するために前記選択されたモデルをインデックスとして使用することを含む、コンピュータ命令と、ユーザへの表示のために前記取得された誘導情報を送信するためのコンピュータ命令と、を備えるコンピュータプログラム製品。

【図 1】

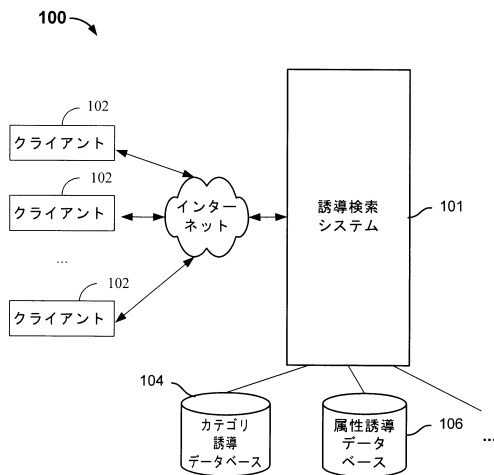


FIG. 1

【図 2】

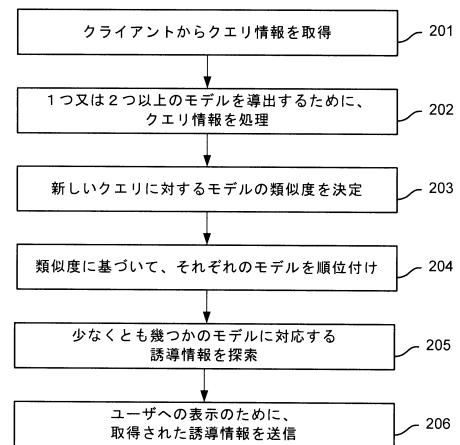


FIG. 2

【図 3】

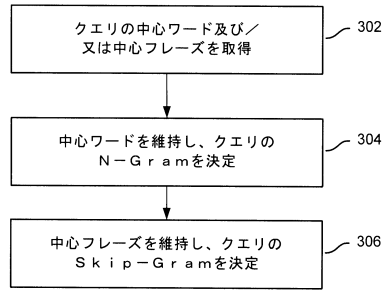


FIG. 3

【図 4】

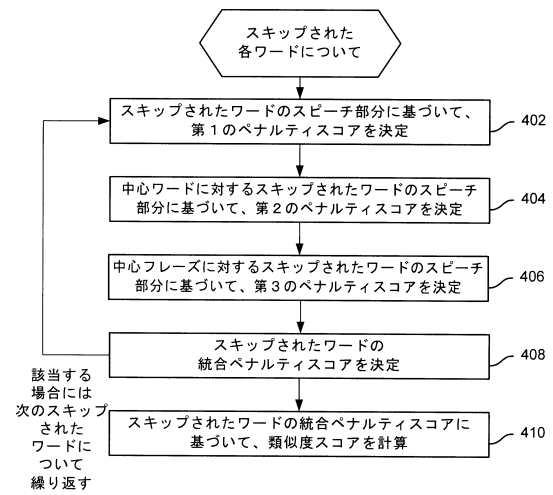


FIG. 4

【図 5】

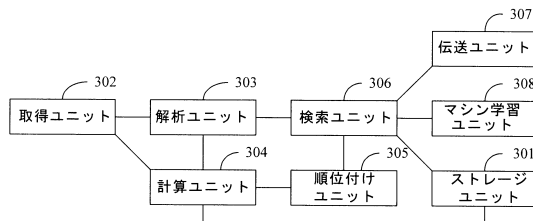


FIG. 5

【図 7】

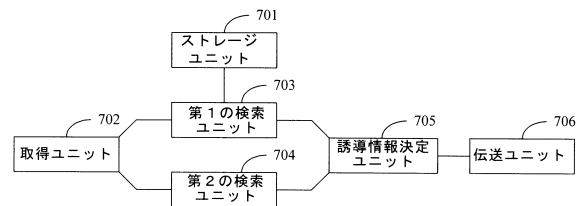


FIG. 7

【図 6】

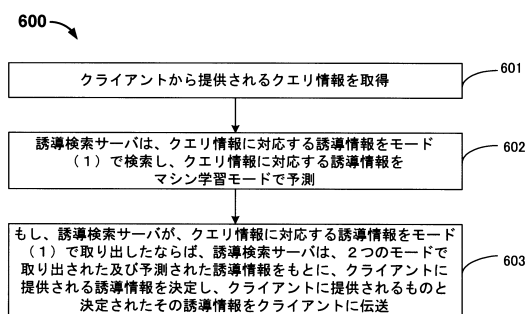


FIG. 6

---

フロントページの続き

- (72)発明者 リアオ・ジエン  
中華人民共和国 ハンチョウ, ワーナー・ロード, 10階, ナンバー391
- (72)発明者 リン・フオン  
中華人民共和国 ハンチョウ, ワーナー・ロード, 10階, ナンバー391
- (72)発明者 チャン・ショウソーン  
中華人民共和国 ハンチョウ, ワーナー・ロード, 10階, ナンバー391
- (72)発明者 チャン・チン  
中華人民共和国 ハンチョウ, ワーナー・ロード, 10階, ナンバー391

審査官 齊藤 貴孝

- (56)参考文献 米国特許出願公開第2009/0248669 (US, A1)  
特開2006-072744 (JP, A)  
特開2009-104602 (JP, A)

- (58)調査した分野(Int.Cl., DB名)  
G06F 17/30