【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22年12月24日(2010.12.24)

【公開番号】特開2010-257488(P2010-257488A)
【公開日】平成22年11月11日(2010.11.11)
【年通号数】公開・登録公報2010-045
【出願番号】特願2010-176034(P2010-176034)
【国際特許分類】
Ｇ０６Ｆ　17/30　　(2006.01)
【ＦＩ】
Ｇ０６Ｆ　17/30　　３２０Ｂ
Ｇ０６Ｆ　17/30　　３４０Ｚ
Ｇ０６Ｆ　17/30　　１７０Ａ

【手続補正書】
【提出日】平成22年10月13日(2010.10.13)
【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】
受信クエリーを改良するための方法であって、
前記受信クエリーに対応するトップランク付けされた文書の初期グループを作成するように該受信クエリーを処理する段階と、
前記初期グループ内の文書の全部又は一部の各文書を事前計算済みのランク付けされた候補用語のそれぞれのセットに関連付けて、前記ランク付けされた候補用語のそれぞれのセットの各候補用語が、前記各文書内に組み込まれるようにする段階と、
選択関数に従って、前記ランク付けされた候補用語のそれぞれのセットの１つ又はそれ以上内にある候補用語のサブセットを選択する段階と、
前記受信クエリーに応答して、前記トップランク付けされた文書の初期グループと前記候補用語のサブセットとを提示する段階とを含み、
前記トップランク付けされた文書の初期グループ内の文書の全部又は一部について、当該文書と関連する前記ランク付けされた候補用語のそれぞれのセットが、
（Ａ）前記文書の用語を候補用語のマスターリストと比較し、前記用語が前記候補用語のマスターリスト内にある場合に該用語を前記候補用語のセットに加え、
（Ｂ）前記文書の用語と前記候補用語のマスターリストとの比較を繰り返し、
（Ｃ）前記候補用語のセット内の候補用語をランク付けして、これにより前記ランク付けされた候補用語のそれぞれのセットを形成し、
前記比較段階（Ａ）の事例によって候補用語が識別される回数が、前記ランキング（Ｃ）によって使用されて、前記ランク付けされた候補用語のセット内の前記候補用語をランク付けすることを特徴とする方法。
【請求項２】
前記ランキング（Ｃ）が更に、前記それぞれの関連する文書の候補用語の第１位置を用いて前記候補用語をランク付けすることを特徴とする請求項１に記載の方法。
【請求項３】
前記識別が更に、
（Ｃ）第１候補用語が前記候補用語のそれぞれのセット内の第２候補用語のサブセット

である場合に該第１候補用語を廃棄する段階と、
（Ｄ）前記第１候補用語が、前記比較段階（Ａ）の事例によって前記ランク付けされた候補用語のそれぞれのセットに関連付けられた前記文書において識別された回数で前記第２候補用語をクレジットする段階と、
（Ｅ）前記ランク付けされた候補用語のそれぞれのセット内に第２候補用語のサブセットである第１候補用語が存在しなくなるまで前記廃棄段階（Ｃ）及び前記クレジット段階（Ｄ）を繰り返す段階と、
を含む請求項１に記載の方法。
【請求項４】
前記識別が更に、
（Ｃ）第１候補用語が前記ランク付けされた候補用語のそれぞれのセット内の第２候補用語の正字の異形又は屈折異形である場合に、該第１候補用語を廃棄する段階と、
（Ｄ）前記第１候補用語が、前記比較段階（Ａ）の事例によって前記ランク付けされた候補用語のそれぞれのセットに関連付けられた前記文書において識別された回数で第２候補用語をクレジットする段階と、
（Ｅ）前記ランク付けされた候補用語のそれぞれのセット内に第２候補用語の正字の異形又は屈折異形である第１候補用語が存在しなくなるまで、前記廃棄段階（Ｃ）及び前記クレジット段階（Ｄ）を繰り返す段階と、
を含む請求項１に記載の方法。
【請求項５】
前記第２候補用語をクレジットする段階（Ｄ）が更に、
前記第１候補用語及び前記第２候補用語を含む結合用語として前記第２候補用語を上書きする段階を含み、前記比較段階（Ａ）の事例によって最も多い回数識別された前記第１候補用語又は前記第２候補用語のうちの１つが前記結合用語の始めに現れることを特徴とする請求項４に記載の方法。
【請求項６】
前記結合用語の始めに現れている用語だけが前記提示において使用されることを特徴とする請求項５に記載の方法。
【請求項７】
前記トップランク付けされた文書の初期グループ内の文書の全部又は一部について、前記文書と関連するランク付けされた候補用語のそれぞれのセットは、該それぞれのセットの各候補用語について前記文書内の候補用語の第１位置を含むことを特徴とする請求項１に記載の方法。
【請求項８】
前記ランク付けされた候補用語の各それぞれのセットが、前記受信クエリーの処理の前の時点で識別されることを特徴とする請求項１に記載の方法。
【請求項９】
前記識別が、
（Ｃ）固有の用語の閾値数が前記比較段階（Ａ）によって検討されたときに、前記比較段階（Ａ）を終了させる段階と、前記繰り返し段階（Ｂ）を終了させる段階とを更に含む請求項１に記載の方法。
【請求項１０】
前記候補用語のマスターリストが、特定の言語について最適化されることを特徴とする請求項１に記載の方法。
【請求項１１】
前記特定の言語が、英語、スペイン語、フランス語、ドイツ語、ポルトガル語、イタリア語、ロシア語、中国語、又は日本語であることを特徴とする請求項１０に記載の方法。
【請求項１２】
前記トップランク付けされた文書の初期グループ内の文書の全部又は一部が、前記候補用語のマスターリストが最適化される言語と同じであることを特徴とする請求項１０に記

載の方法。
【請求項１３】
前記候補用語のマスターリストの各用語が、語又は句であることを特徴とする請求項１に記載の方法。
【請求項１４】
前記候補用語のマスターリストは、１，０００，０００より多い用語を含むことを特徴とする請求項１に記載の方法。
【請求項１５】
前記方法は更に、前記受信クエリーと前記候補用語のサブセットからの候補用語とを含む変更されたクエリーを使用して、前記処理段階、選択段階、及び提示段階を繰り返すことを含む請求項１に記載の方法。
【請求項１６】
前記選択関数は、
（ｉ）前記トップランク付けされた文書の初期グループ内の文書に関連する前記ランク付けされた候補用語の各それぞれのセット内の各候補用語に重み付け関数を適用し、前記ランク付けされた文書の初期グループ内の各トップランクの文書が閾値ランキングより数値的に小さいランキングを有し、
（ｉｉ）前記候補用語のサブセットについて、最も高い重みを受け取る候補用語を選択する、
ことを含む請求項１に記載の方法。
【請求項１７】
前記重み付け関数によって候補用語に加えられる重みが、（ｉ）前記候補用語を含み、且つ（ｉｉ）トップランクの文書にそれぞれ関連するランク付けされた候補用語のセットの数に応じて決定されることを特徴とする請求項１６に記載の方法。
【請求項１８】
前記重み付け関数によって候補用語に加えられる重みが、（ｉ）前記候補用語を含み、且つ（ｉｉ）トップランクの文書にそれぞれ関連するランク付けされた候補用語のセット内の前記候補用語の平均位置に応じて決定されることを特徴とする請求項１６に記載の方法。
【請求項１９】
前記重み付け関数によって候補用語に加えられる重みが、前記受信クエリーの用語が前記候補用語内に存在するかどうかに応じて決定されることを特徴とする請求項１６に記載の方法。
【請求項２０】
前記重み付け関数によって候補用語に加えられる重みが、前記候補用語の文字数に応じて決定されることを特徴とする請求項１６に記載の方法。
【請求項２１】
前記重み付け関数によって候補用語に加えられる重みが、前記候補用語を含むランク付けされた候補用語のセットに関連付けられた前記トップランクの文書の平均ランクに応じて決定されることを特徴とする請求項１６に記載の方法。
【請求項２２】
前記重み付け関数によって候補用語に加えられる重みが、ＴｅｒｍＣｏｕｎｔ、ＴｅｒｍＰｏｓｉｔｉｏｎ、ＲｅｓｕｌｔＰｏｓｉｔｉｏｎ、ＴｅｒｍＬｅｎｇｔｈ、及びＱｕｅｒｙＩｎｃｌｕｓｉｏｎのいずれかの組合せに応じて決定され、ここで、
ＴｅｒｍＣｏｕｎｔは、（ｉ）前記候補用語を含み、且つ（ｉｉ）トップランクの文書にそれぞれ関連するランク付けされた候補用語のセットの数であり、
ＴｅｒｍＰｏｓｉｔｉｏｎは、（ｉ）前記候補用語を含み、且つ（ｉｉ）トップランクの文書にそれぞれ関連するランク付けされた候補用語のセットにおける候補用語のランク位置の関数であり、
ＲｅｓｕｌｔＰｏｓｉｔｉｏｎは、前記候補用語を含むランク付けされた候補用語のセ

ットに関連付けられるトップランクの文書のランクの関数であり、
ＴｅｒｍＬｅｎｇｔｈは、前記候補用語の文字数（候補用語の複雑性）であり、
ＱｕｅｒｙＩｎｃｌｕｓｉｏｎは、前記受信クエリーの用語が前記候補用語内に存在するかどうかを示す値であることを特徴とする請求項１６に記載の方法。
【請求項２３】
前記重み付け関数によって候補用語に加えられる重みが、次式に応じて決定されることを特徴とする請求項２２に記載の方法。
ＴｅｒｍＣｏｕｎｔ＋ＴｅｒｍＰｏｓｉｔｉｏｎ＋ＲｅｓｕｌｔＰｏｓｉｔｉｏｎ＋ＴｅｒｍＬｅｎｇｔｈ＋ＱｕｅｒｙＩｎｃｌｕｓｉｏｎ
【請求項２４】
ＴｅｒｍＣｏｕｎｔ、ＴｅｒｍＰｏｓｉｔｉｏｎ、ＲｅｓｕｌｔＰｏｓｉｔｉｏｎ、ＴｅｒｍＬｅｎｇｔｈ、及びＱｕｅｒｙＩｎｃｌｕｓｉｏｎは、各々別々に重み付けされることを特徴とする請求項２３に記載の方法。
【請求項２５】
前記方法が更に、前記受信クエリーと前記候補用語のサブセットからの候補用語とを含む変更されたクエリーを使用して、任意選択的に該処理段階、選択段階、及び提示段階を繰り返す段階を含むことを特徴とする請求項２３に記載の方法。
【請求項２６】
前記重み付け関数によって候補用語に加えられる重みが、次式に応じて決定されることを特徴とする請求項２５に記載の方法。
（ＴｅｒｍＣｏｕｎｔ*w_1）＋
（ＴｅｒｍＰｏｓｉｔｉｏｎ*（w_2＋（ＲｅｆｉｎｅｍｅｎｔＤｅｐｔｈ*w_2'））＋
（ＲｅｓｕｌｔＰｏｓｉｔｉｏｎ*w_3）＋
（ＴｅｒｍＬｅｎｇｔｈ*（w_4＋（ＲｅｆｉｎｅｍｅｎｔＤｅｐｔｈ*w_4'））＋
（ＱｕｅｒｙＩｎｃｌｕｓｉｏｎ*（w_5＋（ＲｅｆｉｎｅｍｅｎｔＤｅｐｔｈ*w_5'））
ここで、w_1、w_2、w_3、w_4、w_5、w_2'、w_4'、及びw_5'は別々の重みであり、ＲｅｆｉｎｅｍｅｎｔＤｅｐｔｈは前記受信クエリーに対して前記処理が行われた回数である。
【請求項２７】
コンピュータシステムと共に使用されるコンピュータ読み取り可能な記録媒体であって、受信クエリーを改良するためのクエリー改良提案エンジンに、
前記受信クエリーに対応するトップランク付けされた文書の初期グループを作成するように前記受信クエリーを処理させ、
前記トップランク付けされた文書の初期グループ内の文書の全部又は一部の各文書を、事前計算済みのランク付けされた候補用語のそれぞれのセットに関連付けて、前記ランク付けされた候補用語のそれぞれのセットの各候補用語が前記各文書内に組み込まれるように処理させ、
選択関数に従って、前記候補用語のそれぞれのセットの１つ又はそれ以上内にある候補用語のサブセットを選択させ、
前記受信クエリーに応答して、前記トップランク付けされた文書の初期グループ及び前記候補用語のサブセットを提示させ、
前記トップランク付けされた文書の初期グループ内の文書の全部又は一部について、当該文書と関連するランク付けされた候補用語のそれぞれのセットが、
（Ａ）前記文書の用語を候補用語のマスターリストと比較する処理であって、前記用語が前記候補用語のマスターリスト内にある場合に、前記文書の用語を候補用語として前記文書と関連する前記ランク付けされた候補用語のそれぞれのセットに加える処理と、
（Ｂ）前記文書内の用語の最大数が検討されるまで、前記文書の用語と前記候補用語のマスターリストとの比較を再実行する処理とを行うことによって特定されるようにし、

前記トップランク付けされた文書の初期グループ内の文書の全部又は一部について、候補用語が比較段階（Ａ）のための前記命令の事例によって識別される回数が、前記文書と関連する前記ランク付けされた候補用語のそれぞれのセットに含まれるようにするためのプログラムを記録したコンピュータ読み取り可能な記録媒体。

【請求項２８】

前記候補用語が前記文書の語の第１閾値数内で識別される場合、比較段階（Ａ）のための前記命令の事例によって前記候補用語が識別される前記回数が、重みを増やされることを特徴とする請求項２７に記載のコンピュータ読み取り可能な記録媒体。

【請求項２９】

前記識別が更に、

（Ｃ）第１候補用語が前記ランク付けされた候補用語のそれぞれのセット内の第２候補用語のサブセットである場合に前記第１候補用語を廃棄する命令と、

（Ｄ）前記比較段階（Ａ）の事例によって前記ランク付けされた候補用語のそれぞれのセットに関連する前記文書において前記第１候補用語が識別された回数で前記第２候補用語をクレジットする命令と、

（Ｅ）前記ランク付けされた候補用語のそれぞれのセット内に第２候補用語のサブセットである第１候補用語が存在しなくなるまで前記廃棄（Ｃ）の命令及び前記クレジット（Ｄ）の命令を繰り返す命令と、

を含むことを特徴とする請求項２７に記載のコンピュータ読み取り可能な記録媒体。

【請求項３０】

前記識別が更に、

（Ｃ）第１候補用語が前記ランク付けされた候補用語のそれぞれのセット内の第２候補用語の正字の異形又は屈折異形である場合に、該第１候補用語を廃棄する命令と、

（Ｄ）前記比較段階（Ａ）の事例によって前記ランク付けされた候補用語のそれぞれのセットに関連した前記文書において前記第１候補用語が識別された回数で前記第２候補をクレジットする命令と、

（Ｅ）前記ランク付けされた候補用語のそれぞれのセット内に第２候補用語の正字の異形又は屈折異形である第１候補用語が存在しなくなるまで、前記廃棄の命令（Ｃ）と前記クレジットの命令（Ｄ）とを繰り返す命令と、

を含むことを特徴とする請求項２７に記載のコンピュータ読み取り可能な記録媒体。

【請求項３１】

前記第２候補用語をクレジットする前記命令（Ｄ）が更に、

前記第１候補用語及び前記第２候補用語を含む結合用語として前記第２候補用語を上書きする命令を含み、前記比較する命令（Ａ）の事例によって最も多い回数識別された前記第１候補用語又は該第２候補用語のうちの１つが前記結合用語の始めに現れることを特徴とする請求項３０に記載のコンピュータ読み取り可能な記録媒体。

【請求項３２】

前記結合用語の始めに現れている用語だけが前記提示の命令によって使用されることを特徴とする請求項３１に記載のコンピュータ読み取り可能な記録媒体。

【請求項３３】

前記トップランク付けされた文書の初期グループ内の文書の全部又は一部について、前記文書と関連するランク付けされた候補用語のそれぞれのセットは、該それぞれのセットの各候補用語について前記文書内の候補用語の平均位置を含むことを特徴とする請求項２７に記載のコンピュータ読み取り可能な記録媒体。

【請求項３４】

前記文書内の前記候補用語の平均位置は、前記比較の命令（Ａ）の事例中に識別された前記候補用語の各事例の位置を平均することによって決定されることを特徴とする請求項３３に記載のコンピュータ読み取り可能な記録媒体。

【請求項３５】

前記識別は、

（Ｃ）固有の用語の閾値数が前記比較の命令（Ａ）によって検討されたときに、前記比較の命令（Ａ）を終了させる命令と前記繰り返しの命令（Ｂ）を終了させる命令を更に含むことを特徴とする請求項２７に記載のコンピュータ読み取り可能な記録媒体。
【請求項３６】
前記候補用語のマスターリストが、特定の言語について最適化されることを特徴とする請求項２７に記載のコンピュータ読み取り可能な記録媒体。
【請求項３７】
前記トップランク付けされた文書の初期グループ内の文書の全部又は一部の各文書が、前記候補用語のマスターリストが最適化される言語と同じであることを特徴とする請求項３６に記載のコンピュータ読み取り可能な記録媒体。
【請求項３８】
前記候補用語のマスターリストの各用語が、語又は句であることを特徴とする請求項２７に記載のコンピュータ読み取り可能な記録媒体。
【請求項３９】
前記クエリー改良提案エンジンが更に、前記受信クエリーと前記候補用語のサブセットからの候補用語とを含む変更されたクエリーを使用して、前記処理の命令、前記選択の命令、及び前記提示の命令を繰り返すための命令を含むことを特徴とする請求項２７に記載のコンピュータ読み取り可能な記録媒体。
【請求項４０】
前記選択関数は、
（ｉ）前記トップランク付けされた文書の初期グループ内の文書に関連するランク付けされた候補用語の各それぞれのセット内の各候補用語に重み付け関数を適用する命令であって、前記トップランク付けされた文書の初期グループ内の各文書が閾値ランキングより数値的に小さいランキングを有する命令と、
（ｉｉ）前記候補用語のサブセットについて、最も高い重みを受け取る候補用語を選択する命令と、
を含むことを特徴とする請求項２７に記載のコンピュータ読み取り可能な記録媒体。
【請求項４１】
前記重み付け関数によって候補用語に加えられる重みが、前記候補用語がトップランクの文書の上位部分に現れる回数に応じて決定されることを特徴とする請求項４０に記載のコンピュータ読み取り可能な記録媒体。
【請求項４２】
前記重み付け関数によって候補用語に加えられる重みが、前記候補用語が現れるトップランクの文書内の前記候補用語の位置に応じて決定されることを特徴とする請求項４０に記載のコンピュータ読み取り可能な記録媒体。
【請求項４３】
前記重み付け関数によって候補用語に加えられる重みが、前記受信クエリーの用語が前記候補用語内に存在するかどうかに応じて決定されることを特徴とする請求項４０に記載のコンピュータ読み取り可能な記録媒体。
【請求項４４】
前記重み付け関数によって候補用語に加えられる重みが、前記候補用語の文字数に応じて決定されることを特徴とする請求項４０に記載のコンピュータ読み取り可能な記録媒体。
【請求項４５】
前記重み付け関数によって候補用語に加えられる重みが、前記候補用語を含む前記トップランク付けされた文書の初期グループ内の文書の位置に応じて決定されることを特徴とする請求項４０に記載のコンピュータ読み取り可能な記録媒体。
【請求項４６】
前記重み付け関数によって候補用語に加えられる重みが、ＴｅｒｍＣｏｕｎｔ、ＴｅｒｍＰｏｓｉｔｉｏｎ、ＲｅｓｕｌｔＰｏｓｉｔｉｏｎ、ＴｅｒｍＬｅｎｇｔｈ、及びＱｕ

ｅｒｙＩｎｃｌｕｓｉｏｎのいずれかの組合せに応じて決定され、ここで、
ＴｅｒｍＣｏｕｎｔは、各トップランクの文書の上位部分に前記候補用語が現れる回数であり、
ＴｅｒｍＰｏｓｉｔｉｏｎは、前記候補用語が現れる各トップランクの文書内の前記候補用語の位置の関数であり、
ＲｅｓｕｌｔＰｏｓｉｔｉｏｎは、前記候補用語を含む前記トップランク付けされた文書の初期グループにおける文書の位置の関数であり、
ＴｅｒｍＬｅｎｇｔｈは、前記候補用語の文字数であり、
ＱｕｅｒｙＩｎｃｌｕｓｉｏｎは、前記受信クエリーの用語が前記候補用語内に存在する場合はゼロではなく、ＱｕｅｒｙＩｎｃｌｕｓｉｏｎは、前記受信クエリーの用語が前記候補用語内に存在しない場合はゼロであることを特徴とする請求項４０に記載のコンピュータ読み取り可能な記録媒体。
【請求項４７】
前記重み付け関数によって候補用語に加えられる重みが、次式に応じて決定されることを特徴とする請求項４６に記載のコンピュータ読み取り可能な記録媒体。
ＴｅｒｍＣｏｕｎｔ＋ＴｅｒｍＰｏｓｉｔｉｏｎ＋ＲｅｓｕｌｔＰｏｓｉｔｉｏｎ＋ＴｅｒｍＬｅｎｇｔｈ＋ＱｕｅｒｙＩｎｃｌｕｓｉｏｎ
【請求項４８】
ＴｅｒｍＣｏｕｎｔ、ＴｅｒｍＰｏｓｉｔｉｏｎ、ＲｅｓｕｌｔＰｏｓｉｔｉｏｎ、ＴｅｒｍＬｅｎｇｔｈ、及びＱｕｅｒｙＩｎｃｌｕｓｉｏｎは、各々別々に重み付けされることを特徴とする請求項４７に記載のコンピュータ読み取り可能な記録媒体。
【請求項４９】
受信クエリーを改良するためのコンピュータシステムであって、
中央処理ユニットと、
前記中央処理ユニットに結合された、クエリー改良提案エンジンを記憶するメモリと、を含み、
前記クエリー改良提案エンジンが、
前記受信クエリーに対応するトップランク付けされた文書の初期グループを作成するように前記受信クエリーを処理し、ここで、前記トップランク付けされた文書の初期グループ内の文書の全部又は一部の各文書が、事前計算済みのランク付けされた候補用語のそれぞれのセットに関連付けられ、その結果、前記ランク付けされた候補用語のそれぞれのセットの各候補用語が前記文書内に組み込まれ、さらに、
選択関数に従って、前記候補用語のそれぞれのセットの１つ又はそれ以上にある候補用語のサブセットを選択し、
前記受信クエリーに応答して、前記トップランク付けされた文書の初期グループ及び前記候補用語のサブセットを提示し、
前記トップランク付けされた文書の初期グループ内の文書の全部又は一部について、当該文書と関連するランク付けされた候補用語のそれぞれのセットが、
（Ａ）前記文書の用語を候補用語のマスターリストと比較し、前記用語が前記候補用語のマスターリスト内にある場合に、前記文書の用語を候補用語として前記文書と関連するランク付けされた候補用語のそれぞれのセットに加え、
（Ｂ）前記文書内の用語の最大数が検討されるまで、前記文書の用語と前記候補用語のマスターリストとの比較を再実行することによって特定され、
前記トップランク付けされた文書の初期グループ内の文書の全部又は一部について、該文書と関連する前記ランク付けされた候補用語により、該文書と関連するランク付けされた候補用語のそれぞれのセット内の候補用語の分類が行われ、前記分類が第１分類又は第２分類を含み、さらに、
候補用語が比較段階（Ａ）のための前記命令の事例によって識別される回数が、前記文書と関連する前記ランク付けされた候補用語のそれぞれのセットに含まれることを特徴とするコンピュータシステム。

【請求項５０】
前記識別が更に、
（Ｃ）第１候補用語が前記ランク付けされた候補用語のそれぞれのセット内の第２候補用語のサブセットである場合に前記第１候補用語を廃棄する命令と、
（Ｄ）前記比較段階（Ａ）の事例によって前記ランク付けされた候補用語のそれぞれのセットに関連する前記文書において前記第１候補用語が識別された回数で前記第２候補用語をクレジットする命令と、
（Ｅ）前記ランク付けされた候補用語のそれぞれのセット内に第２候補用語のサブセットである第１候補用語が存在しなくなるまで前記廃棄（Ｃ）の命令及び前記クレジット（Ｄ）の命令を繰り返す命令と、
を含むことを特徴とする請求項４９に記載のコンピュータシステム。
【請求項５１】
前記識別が更に、
（Ｃ）第１候補用語が前記ランク付けされた候補用語のそれぞれのセット内の第２候補用語の正字の異形又は屈折異形である場合に、該第１候補用語を廃棄する命令と、
（Ｄ）前記比較段階（Ａ）の事例によって前記ランク付けされた候補用語のそれぞれのセットに関連した前記文書において前記第１候補用語が識別された回数で前記第２候補をクレジットする命令と、
（Ｅ）前記ランク付けされた候補用語のそれぞれのセット内に第２候補用語の正字の異形又は屈折異形である第１候補用語が存在しなくなるまで、前記廃棄の命令（Ｃ）と前記クレジットの命令（Ｄ）とを繰り返す命令と、
を含むことを特徴とする請求項４９に記載のコンピュータシステム。
【請求項５２】
前記トップランク付けされた文書の初期グループ内の文書の全部又は一部について、前記文書と関連するランク付けされた候補用語のそれぞれのセットは、該それぞれのセットの各候補用語について前記文書内の候補用語の平均位置を含むことを特徴とする請求項４９に記載のコンピュータシステム。
【請求項５３】
前記識別は、
（Ｃ）固有の用語の閾値数が前記比較の命令（Ａ）によって検討されたときに、前記比較の命令（Ａ）を終了させる命令と前記繰り返しの命令（Ｂ）を終了させる命令を更に含むことを特徴とする請求項４９に記載のコンピュータシステム。
【請求項５４】
前記クエリー改良提案エンジンが更に、前記受信クエリーと前記候補用語のサブセットからの候補用語とを含む変更されたクエリーを使用して、前記処理の命令、前記選択の命令、及び前記提示の命令を繰り返すための命令を含むことを特徴とする請求項４９に記載のコンピュータシステム。
【請求項５５】
前記選択関数は、
（ｉ）前記トップランク付けされた文書の初期グループ内の文書に関連するランク付けされた候補用語の各それぞれのセット内の各候補用語に重み付け関数を適用する命令であって、前記トップランク付けされた文書の初期グループ内の各文書が閾値ランキングより数値的に小さいランキングを有する命令と、
（ｉｉ）前記候補用語のサブセットについて、最も高い重みを受け取る候補用語を選択する命令と、
を含むことを特徴とする請求項５０に記載のコンピュータシステム。
【請求項５６】
前記重み付け関数によって候補用語に加えられる重みが、ＴｅｒｍＣｏｕｎｔ、ＴｅｒｍＰｏｓｉｔｉｏｎ、ＲｅｓｕｌｔＰｏｓｉｔｉｏｎ、ＴｅｒｍＬｅｎｇｔｈ、及びＱｕｅｒｙＩｎｃｌｕｓｉｏｎのいずれかの組合せに応じて決定され、ここで、

ＴｅｒｍＣｏｕｎｔは、各トップランクの文書の上位部分に前記候補用語が現れる回数であり、
ＴｅｒｍＰｏｓｉｔｉｏｎは、前記候補用語が現れる各トップランクの文書内の前記候補用語の位置の関数であり、
ＲｅｓｕｌｔＰｏｓｉｔｉｏｎは、前記候補用語を含む前記トップランク付けされた文書の初期グループにおける文書の位置の関数であり、
ＴｅｒｍＬｅｎｇｔｈは、前記候補用語の文字数であり、
ＱｕｅｒｙＩｎｃｌｕｓｉｏｎは、前記受信クエリーの用語が前記候補用語内に存在する場合適用され、ＱｕｅｒｙＩｎｃｌｕｓｉｏｎは、前記受信クエリーの用語が前記候補用語内に存在しない場合に適用されないことを特徴とする請求項５５に記載のコンピュータシステム。
【請求項５７】
前記重み付け関数によって候補用語に加えられる重みが、次式に応じて決定されることを特徴とする請求項５６に記載のコンピュータシステム。
ＴｅｒｍＣｏｕｎｔ＋ＴｅｒｍＰｏｓｉｔｉｏｎ＋ＲｅｓｕｌｔＰｏｓｉｔｉｏｎ＋ＴｅｒｍＬｅｎｇｔｈ＋ＱｕｅｒｙＩｎｃｌｕｓｉｏｎ
【請求項５８】
ＴｅｒｍＣｏｕｎｔ、ＴｅｒｍＰｏｓｉｔｉｏｎ、ＲｅｓｕｌｔＰｏｓｉｔｉｏｎ、ＴｅｒｍＬｅｎｇｔｈ、及びＱｕｅｒｙＩｎｃｌｕｓｉｏｎは、各々別々に重み付けされることを特徴とする請求項５７に記載のコンピュータシステム。
【請求項５９】
前記ランク付けされた候補用語の各それぞれのセットが、前記受信クエリーの処理の前の時点で識別されることを特徴とする請求項２７に記載のコンピュータ読み取り可能な記録媒体。
【請求項６０】
前記ランク付けされた候補用語の各それぞれのセットが、前記受信クエリーの処理の前の時点で識別されることを特徴とする請求項４９に記載のコンピュータシステム。