

【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23年1月27日(2011.1.27)

【公表番号】特表2010-511966(P2010-511966A)

【公表日】平成22年4月15日(2010.4.15)

【年通号数】公開・登録公報2010-015

【出願番号】特願2009-540463(P2009-540463)

【国際特許分類】

G 06 F 17/21 (2006.01)

【F I】

G 06 F 17/21 550 J

【手続補正書】

【提出日】平成22年12月3日(2010.12.3)

【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】

コンピュータシステムにより実行する方法であって、

テキストサンプルからの1つまたは複数のコロケーションをコーパスと比較するステップであって、前記1つまたは複数のコロケーションのそれぞれを含む1つまたは複数のクエリー語を使用して、1つまたは複数のウェブ検索を実行するステップを含み、検索が実行される前記1つまたは複数のコロケーションのそれぞれに対して、前記クエリー語の1つが、前記コロケーションに一致するよう事前に選択された閾値を満たす検索結果を与えるまで、または事前に選択された閾値を満たすことなく、前記コロケーションを含む前記クエリー語の全てが使用されるまで、前記コロケーションを含む前記1つまたは複数のクエリー語のそれぞれに対して検索が実行されるステップと、

前記コロケーションが、前記コーパス内で好ましくないかどうかを特定するステップと、

前記コロケーションが好ましくないかどうかの表示を、出力装置を介して提供するステップと、

1つの前記好ましくないコロケーション内の1つの単語を置換するワイルドカードを用いて、1つまたは複数のクエリー語を作成するステップと、

前記クエリー語に対して単語コロケーションリファレンスを検索するステップと、

前記ワイルドカードを置換する候補の単語の割合が比較的高い検索結果を特定するステップと、

前記候補の単語を有する前記検索結果を、前記出力装置を介して潜在的に適切な単語コロケーションとして与えるステップと
を含むことを特徴とする方法。

【請求項2】

前記コーパスが、ウェブ上のコンテンツを含むことを特徴とする請求項1に記載の方法と、

【請求項3】

前記コロケーションが前記コーパス内で見出されない場合、または、前記コロケーションが、前記コーパス内に前記コロケーションのファジーマッチが有意に存在することを示す、事前に選択された閾値一致スコアより上のスコアを取れない場合、前記コロケーショ

ンは好ましくないことを特徴とする請求項 1 に記載の方法。

【請求項 4】

前記コーパスが、ネットワーク上でアクセス可能なコンテンツを含むことを特徴とする請求項 1 に記載の方法。

【請求項 5】

検索が実行されるコロケーションのそれぞれに対して、前記コロケーションを含むクエリー語が、前記コーパス内に前記クエリー語のファジーマッチが有意に存在することを示す、事前に選択された閾値一致スコアより上のスコアを取れない場合、前記コロケーションは好ましくないことを特徴とする請求項 1 に記載の方法。

【請求項 6】

少なくとも 1 つの前記クエリー語が、前記コロケーションを含む文を含むことを特徴とする請求項 1 に記載の方法。

【請求項 7】

少なくとも 1 つの前記クエリー語が、前記コロケーションを含む要約された文を含み、前記要約された文が、前記コロケーションを含む文から付属部分を取り除くことによって形成されることを特徴とする請求項 1 に記載の方法。

【請求項 8】

少なくとも 1 つのクエリー語が、前記コロケーションを含むチャンクの対を含むことを特徴とする請求項 1 に記載の方法。

【請求項 9】

少なくとも 1 つのクエリー語が、前記コロケーションを含む単語の対を含むことを特徴とする請求項 1 に記載の方法。

【請求項 10】

前記テキストサンプルに品詞タグを付けるステップをさらに含み、前記コーパスと比較される前記コロケーションは、動詞 - 名詞コロケーション、前置詞 - 名詞コロケーション、形容詞 - 名詞コロケーション、および動詞 - 副詞コロケーションの少なくとも 1 つを含むものが前記テキストサンプルから選択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】

前記テキストサンプルは、対象の言語であり、前記方法は、前記コーパスをスクリーニングして前記コロケーションを前記対象の言語で表示されるコンテンツのみと比較するステップをさらに含むことを特徴とする請求項 1 に記載の方法。

【請求項 12】

前記テキストサンプルをスクリーニングして非ネイティブの使用を示すべく表示されているかを判定するステップと、前記テキストサンプルが非ネイティブ使用を示すべく表示されている場合に、前記テキストサンプルからの前記コロケーションを前記コーパスと比較する前記ステップを自動的に開始するステップとをさらに含むことを特徴とする請求項 1 に記載の方法。

【請求項 13】

前記クエリー語が、1 つまたは複数の文テンプレート、チャンク・テンプレート、および単語テンプレートを含み、候補の単語を含む検索結果の割合に、文テンプレートに対しては最も高く、チャンク・テンプレートに対してはそれより低く、および単語テンプレートに対しては最も低い、クエリー・テンプレートの重み付けを乗ずることにより、前記ワイルドカードを置換する候補の単語の割合が比較的高い検索の結果を評価することを特徴とする請求項 1 に記載の方法。

【請求項 14】

前記クエリー語が、1 つまたは複数の動詞 - 名詞コロケーション、前置詞 - 名詞コロケーション、形容詞 - 名詞コロケーション、および動詞 - 副詞コロケーションを含み、前記ワイルドカードが、動詞 - 名詞コロケーション中の動詞、動詞 - 名詞コロケーション中の名詞、前置詞 - 名詞コロケーション中の前置詞、形容詞 - 名詞コロケーション中の形容詞、または動詞 - 副詞コロケーション中の副詞として選択されることを特徴とする請求項 1

に記載の方法。

【請求項 1 5】

ユーザが、潜在的に適切な単語コロケーションの1つを選択して、これが対応する好ましくない単語コロケーションを置換することを可能にするステップをさらに含むことを特徴とする請求項1に記載の方法。

【請求項 1 6】

コンピュータシステムによって実行可能な命令を記憶したコンピュータ読み取り可能記憶媒体であって、前記命令が、前記コンピュータシステムを、

テキスト内の単語コロケーションの表示を受け取り、

前記表示された単語コロケーションに関連する1つまたは複数のクエリー・テンプレートのそれぞれに対してウェブ検索を実行し、前記クエリー・テンプレートの1つが、前記単語コロケーションが見出された文を含み、前記クエリー・テンプレートの1つが、前記単語コロケーションが見出された文に基づく要約された文を含み、前記クエリー・テンプレートの1つが、前記単語コロケーションを含むチャunkの対を含み、前記クエリー・テンプレートの1つが、前記単語コロケーションを含む個々の単語の対を含み、

前記文を含む前記クエリー・テンプレートが正確に一致すること、または前記要約された文を含む前記クエリー・テンプレートが正確に一致すること、または前記チャunkの対に対する事前に選択された閾値より大きい前記チャunkの対を含む前記クエリー・テンプレートの一一致スコア、または前記個々の単語の対に対する事前に選択された閾値より大きい前記個々の単語の対を服務前記クエリー・テンプレートの一一致スコアによって示される、前記単語コロケーションが通常の使用に対応しているか、または好ましくないコロケーションであることを、前記1つまたは複数のクエリー・テンプレートのそれぞれに対してウェブ検索した結果が示すかどうかを評価し、

ユーザが認識可能な出力装置を介して、前記単語コロケーションが通常の使用に対応しているかどうか、または前記単語コロケーションが好ましくないかどうかを示し、

1つの前記好ましくないコロケーション内の1つの単語を置換するワイルドカードを用いて、1つまたは複数のクエリー語を作成し、

前記クエリー語に対して単語コロケーションリファレンスを検索し、

前記ワイルドカードを置換する候補の単語の割合が比較的高い検索結果を特定し、

前記候補の単語を有する前記検索結果を、前記出力装置を介して潜在的に適切な単語コロケーションとして与える

ように設定することを特徴とするコンピュータ読み取り可能記憶媒体。

【請求項 1 7】

コンピュータシステムであって、前記コンピュータシステムに記憶されたコンピュータ実行可能命令が前記コンピュータシステムの処理装置によって実行されると、

テキスト内の単語コロケーションを特定し、

1つまたは複数の前記単語コロケーションに基づく一組のクエリー・テンプレートに対してウェブを検索し、

前記検索の結果が、前記単語コロケーションが前記ウェブ上には比較的少ないということを示すかどうかを、ユーザ出力装置を介して示し、

1つの前記比較的少ない単語コロケーション内の1つの単語を置換するワイルドカードを用いて、1つまたは複数のクエリー語を作成し、

前記クエリー語に対して前記ウェブを検索し、

前記ワイルドカードを置換する候補の単語の割合が比較的高い検索結果を特定し、

前記候補の単語を有する前記検索結果を、前記出力装置を介して前記比較的少ない単語コロケーションに対する可能な置換として与える

ように構成されたことを特徴とするコンピュータシステム。