

【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18年8月17日(2006.8.17)

【公開番号】特開2004-62893(P2004-62893A)

【公開日】平成16年2月26日(2004.2.26)

【年通号数】公開・登録公報2004-008

【出願番号】特願2003-188931(P2003-188931)

【国際特許分類】

G 06 F 17/21 (2006.01)

G 06 F 17/30 (2006.01)

【F I】

G 06 F 17/21 5 9 0 E

G 06 F 17/21 5 8 6 A

G 06 F 17/30 1 7 0 A

G 06 F 17/30 3 5 0 C

【手続補正書】

【提出日】平成18年6月30日(2006.6.30)

【手続補正1】

【補正対象書類名】明細書

【補正対象項目名】特許請求の範囲

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項1】文の集合から例文を検索する方法であって、

入力照会文を受け取るステップと、

用語頻度-逆文書頻度アルゴリズムを使用して、前記文の集合から前記入力照会文に対する候補例文を選択するステップと、

前記選択した候補例文と前記入力照会文との間の編集距離に基づいて、前記選択した候補例文を再ランク付けするステップであって、前記選択した候補例文と前記入力照会文との編集距離が、1組の前記選択した候補例文と前記入力照会文とを他の前記入力照会文と前記候補例文との組に変更するのに必要な最小の操作数に基づく再ランク付けするステップと

を含むことを特徴とする方法。

【請求項2】前記選択した候補例文を再ランク付けするステップは、前記選択した候補例文と前記入力照会文との間の重み付き編集距離に基づいて、前記選択した候補例文を再ランク付け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3】重み付き編集距離に基づいて、前記選択した候補例文を再ランク付けするステップは、

各候補例文についての別々の重み付き編集距離を、前記候補例文中の用語の関数として、かつ前記候補例文中の前記用語に対応した重み付きスコアの関数として計算するステップであって、前記重み付きスコアは、前記候補例文中の対応する用語に関連した品詞に基づいて異なる値を有するステップと、

各候補例文について前記計算した別々の重み付き編集距離に基づいて、前記選択した候補例文を再ランク付けするステップとをさらに含むことを特徴とする請求項2に記載の方法。

【請求項4】前記用語頻度-逆文書頻度アルゴリズムを使用して、前記文の集合から前記入力照会文に対する候補例文を選択するステップは、

前記文の集合の文中における対応した用語に関連する品詞をタグ付けするステップと、

前記入力照会文からトップワードを除去するステップと、

前記文の集合における各文について用語頻度 - 逆文書頻度スコアを計算するステップとをさらに含むことを特徴とする請求項 3 に記載の方法。

【請求項 5】 前記用語頻度 - 逆文書頻度アルゴリズムを使用して、前記文の集合から前記入力照会文に対する候補例文を選択するステップは、前記文の集合のうち、しきい値よりも大きい用語頻度 - 逆文書頻度スコアを有する文を前記候補例文として選択するステップをさらに含むことを特徴とする請求項 5 に記載の方法。

【請求項 6】 入力照会文を受け取るステップと、

用語頻度 - 逆文書頻度アルゴリズムを使用して、文の集合から前記入力照会文に対する候補例文を選択するステップと、

前記選択した候補例文と前記入力照会文との間の編集距離に基づいて、前記選択した候補例文を再ランク付けするステップであって、前記選択した候補例文と前記入力照会文との編集距離が、1組の前記選択した候補例文と前記入力照会文とを他の前記入力照会文と前記候補例文との組に変更するのに必要な最小の操作数に基づく再ランク付けするステップとを含むステップを実行する

コンピュータ実行可能命令を有することを特徴とするコンピュータ可読媒体。

【請求項 7】 前記選択した候補例文を再ランク付けするステップは、前記選択した候補例文と前記入力照会文との間の重み付き編集距離に基づいて、前記選択した候補例文を再ランク付けするステップをさらに含むことを特徴とする請求項 6 に記載のコンピュータ可読媒体。

【請求項 8】 重み付き編集距離に基づいて、前記選択した候補例文を再ランク付けするステップは、

各候補例文についての別々の重み付き編集距離を、前記候補例文中の用語の関数として、かつ前記候補例文中の前記用語に対応した重み付きスコアの関数として計算するステップであって、前記重み付きスコアは、前記候補例文中の対応する用語に関連した品詞に基づいて異なる値を有するステップと、

各候補例文についての前記計算した別々の重み付き編集距離に基づいて、前記選択した候補例文を再ランク付けするステップとをさらに含むことを特徴とする請求項 7 に記載のコンピュータ可読媒体。

【請求項 9】 前記用語頻度 - 逆文書頻度アルゴリズムを使用して、前記文の集合から前記入力照会文に対する候補例文を選択するステップは、

前記文の集合の文中における対応した用語に関連する品詞をタグ付けするステップと、前記入力照会文からトップワードを除去するステップと、

前記文の集合における各文について用語頻度 - 逆文書頻度スコアを計算するステップとをさらに含むことを特徴とする請求項 8 に記載のコンピュータ可読媒体。

【請求項 10】 前記用語頻度 - 逆文書頻度アルゴリズムを使用して、前記文の集合から前記入力照会文に対する候補例文を選択するステップは、前記文の集合のうち、しきい値よりも大きい用語頻度 - 逆文書頻度スコアを有する文を前記候補例文として選択するステップをさらに含むことを特徴とする請求項 9 に記載のコンピュータ可読媒体。

【請求項 11】 文の集合から例文を検索するシステムであって、

照会文を受け取る入力と、

用語頻度 - 逆文書頻度アルゴリズムを使用して、前記文の集合から前記照会文に対する候補例文を選択する前記入力に結合された、用語頻度 - 逆文書頻度文検索コンポーネントと、

選択した各候補例文について別々の重み付き編集距離を、前記候補例文における用語の関数として、かつ前記候補例文における前記用語に対応した重み付きスコアの関数として計算する、前記用語頻度 - 逆文書頻度コンポーネントに結合された重み付き編集距離計算コンポーネントであって、前記重み付きスコアは、前記候補例文中の対応した用語に関連する品詞に基づいて異なる値を有する重み付き編集距離計算コンポーネントと、

各候補例文について前記計算した別々の重み付き編集距離に基づいて、前記選択した候

補例文をランク付けする、前記重み付き編集距離計算コンポーネントに結合されたランク付けコンポーネントと
を備えることを特徴とするシステム。