

【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27年4月16日(2015.4.16)

【公開番号】特開2012-208924(P2012-208924A)

【公開日】平成24年10月25日(2012.10.25)

【年通号数】公開・登録公報2012-044

【出願番号】特願2012-45250(P2012-45250)

【国際特許分類】

G 06 F 17/30 (2006.01)

【F I】

G 06 F 17/30 3 5 0 C

G 06 F 17/30 3 4 0 Z

G 06 F 17/30 1 7 0 A

【手続補正書】

【提出日】平成27年2月26日(2015.2.26)

【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】

コンピュータ実行可能な文書比較方法であって、

コンピュータが、2つの文書に関する多数の文書間類似値  $s_i(a, b)$  を受信するステップであって、前記文書間類似値は、様々な文書間類似度計算方法  $i$  によって計算される、前記ステップと、

2つの文書  $a$  及び  $b$  のそれぞれの個別の文書間類似度計算方法  $i$  の重み  $(\alpha_i, \beta_i)$  を決定するステップと、

ここで、重み  $\alpha_i$  は文書  $a$  に対応し、重み  $\beta_i$  は文書  $b$  に対応し、1つの文書間類似度計算方法の重みを決定するステップは、

前記重みを初期化するステップと、

前記重みを、ユーザからのフィードバックに基づいて更新するステップと、を含み、各文書の前記文書間類似度計算方法の前記重みを更新するステップは、ユーザからのフィードバックに基づいて、重みに学習アルゴリズムを適用するステップを含み、

前記2つの文書に関する前記個別の文書間類似度計算方法  $i$  の組み合わせた重みを計算するための重み組み合わせ関数  $f(\alpha_i, \beta_i)$  を決定するステップと、

ここで、前記重み組み合わせ関数は、前記文書  $a$  及び  $b$  の間の比較の方向を定め、

コンピュータが、 $k$  個の文書間類似度計算方法によって計算される前記文書間類似値と前記重み組み合わせ関数とに基づいて、

【数1】

$$S(a, b) = \sum_{i=1}^k f(\alpha_i, \beta_i) \cdot s_i(a, b)$$

として、組み合わされた類似値  $S(a, b)$  を生成するステップとを含む、コンピュータ実行可能な文書比較方法。

【請求項 2】

前記文書間類似度計算方法は、テキストに基づいた文書間類似度計算方法、視覚に基づいた文書間類似度計算方法、使用に基づいた文書間類似度計算方法、および、ソーシャルネットワークに基づいた文書間類似度計算方法のうちの1つ以上を含む、請求項1に記載の方法。

【請求項 3】

各文書の前記文書間類似度計算方法の前記重みは、文書タイプ、文書場所、文書構造、文書使用、および、それらの文書の重みのうちの少なくとも1つに基づいて初期化される、請求項1に記載の方法。

【請求項 4】

前記重み組み合わせ関数は、前記重みの、平均値、最小値、または、最大値を計算する実数値関数である、請求項1に記載の方法。

【請求項 5】

文書間類似度レベルを推定するシステムであって、  
文書間類似度レベルを推定するための、少なくとも1つのメモリに結合された少なくとも1つのハードウェア・プロセッサと、

2つの文書に関する多数の文書間類似値  $s_i(a, b)$  を受信するように構成された受信機構であって、前記文書間類似値は、様々な文書間類似度計算方法  $i$  によって計算される、前記受信機構と、

2つの文書  $a$  及び  $b$  のそれぞれの個別の文書間類似度計算方法  $i$  の重み  $(\alpha_i, \beta_i)$  と、

前記2つの文書に関する前記個別の文書間類似度計算方法  $i$  の組み合わされた重みを計算するための重み組み合わせ関数  $f(\alpha_i, \beta_i)$  と、

を決定するように構成された決定機構であって、

ここで、重み  $\alpha_i$  は文書  $a$  に対応し、重み  $\beta_i$  は文書  $b$  に対応し、1つの文書間類似度計算方法の重みの決定は、

前記重みを初期化することと、

前記重みを、ユーザからのフィードバックに基づいて更新することと、を含み、各文書の前記文書間類似度計算方法の前記重みを更新することは、ユーザからのフィードバックに基づいて、重みに学習アルゴリズムを適用することを含み、

前記重み組み合わせ関数は、前記文書  $a$  及び  $b$  の間の比較の方向を定める、決定機構と、

$k$  個の文書間類似度計算方法によって計算される前記文書間類似値と前記重み組み合わせ関数とに基づいて、

【数2】

$$S(a, b) = \sum_{i=1}^k f(\alpha_i, \beta_i) \cdot s_i(a, b)$$

として、組み合わされた類似値  $S(a, b)$  を生成するように構成された、値生成機構とを含む、文書間類似度レベルを推定するシステム。