

【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19年9月27日(2007.9.27)

【公開番号】特開2005-115922(P2005-115922A)

【公開日】平成17年4月28日(2005.4.28)

【年通号数】公開・登録公報2005-017

【出願番号】特願2004-239993(P2004-239993)

【国際特許分類】

G 06 F 17/27 (2006.01)

【F I】

G 06 F 17/27 M

【手続補正書】

【提出日】平成19年8月9日(2007.8.9)

【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】

テキストから意味構造を識別する方法であって、
前記テキストから意味エンティティを識別すること、

前記意味エンティティを結合して、少なくとも2つの意味構造候補を形成すること、

前記意味構造の尤度に基づいて各意味構造候補の意味スコアを生成すること、

意味構造候補において同じレベルで現れる意味エンティティ対の個々の遷移確率を生成することによって各意味構造候補の遷移スコアを生成することであって、前記遷移確率を生成することは、1組の意味エンティティにおける意味エンティティの種類を用いて前記遷移確率を生成するために使用される値を検索すること、

前記テキスト内の単語の位置及び前記単語から形成される意味エンティティの前記意味構造内の位置に基づいて各意味構造候補の構文スコアを生成すること、

各意味構造の前記構文スコアと前記意味スコアとを結合して、少なくともテキストの一部に対して前記意味構造候補の1つを選択すること

を含むことを特徴とする方法。

【請求項2】

構文スコアを生成することは、意味構造候補内の意味エンティティ間の各親子関係に対して、親意味エンティティと関連する単語のトークン位置を識別することと、子意味エンティティと関連する単語のトークン位置を識別することと、前記トークン位置に基づき確率を判定することを含むことを特徴とする請求項1記載の方法。

【請求項3】

意味エンティティを用いて意味構造候補を形成することは、前記テキストから識別されない前記意味構造候補内に意味エンティティを含めることを含むことを特徴とする請求項1記載の方法。

【請求項4】

意味スコアを生成することは、前記意味エンティティが前記テキストから識別されなかった場合には、ペナルティ要素を用いて意味エンティティの確率を割り引くことを含むことを特徴とする請求項3記載の方法。

【請求項5】

意味エンティティを識別することは、テキストが意味エンティティと関連するスキーマ

を使用することを含むことを特徴とする請求項 1 記載の方法。

【請求項 6】

单一の単語は、前記スキーマ内の複数意味エンティティと関連することを特徴とする請求項 5 記載の方法。

【請求項 7】

意味エンティティを用いて意味構造を形成することは、スキーマ内で定義される意味エンティティ間の関係における前記意味構造の基礎を形成することを含むことを特徴とする請求項 5 記載の方法。

【請求項 8】

意味エンティティ間の前記関係を定義する前記スキーマと、テキストと意味エンティティとを関連付ける前記スキーマとが同じスキーマであることを特徴とする請求項 7 記載の方法。

【請求項 9】

スキーマ内で定義される関係における前記意味構造の基礎を形成することは、

2つの意味エンティティによってインデックスを付けられ、前記スキーマ内に見出される1組の意味構造を含むエントリをもつテーブルを構築することであって、各意味構造は、前記エントリにインデックスを付けるために使用される2つの意味エンティティを含むこと、

前記2つの意味エンティティを使用して前記テーブルを検索し、前記1組の意味構造を見つけること、

前記1組の意味構造内の各意味構造に対して意味構造を形成することを含むことを特徴とする請求項 7 記載の方法。

【請求項 10】

選択された意味構造と、前記テキストから識別される追加的な意味エンティティとを結合して拡張意味構造を形成することを更に含むことを特徴とする請求項 1 記載の方法。

【請求項 11】

前記拡張意味構造の意味スコアと構文スコアを生成することを更に含むことを特徴とする請求項 10 記載の方法。

【請求項 12】

選択された意味構造と追加的な意味エンティティとを結合することは、順序づけられた検索を実行することを含み、前記順序づけられた検索は、

前記追加的な意味エンティティを前記選択された意味構造内の前記エンティティのいずれかの下位に配置できるかどうかを判定すること、

前記選択された意味構造の先端を前記追加的な意味エンティティの下位に配置することができるかどうかを判定すること、

前記追加的な意味エンティティと前記選択された意味構造の前記先端の両方を下位に配置することができる共通上位エンティティがあるかどうかを判定すること

を含むことを特徴とする請求項 10 記載の方法。

【請求項 13】

選択された意味構造がコマンド意味構造でないと決定すること、

コマンド意味構造内に前記選択された意味構造を配置すること、

前記コマンド意味構造を前記テキストの前記意味構造として識別することを更に含むことを特徴とする請求項 1 記載の方法。

【請求項 14】

前記選択された意味構造を複数コマンド意味構造内に配置することを更に含むことを特徴とする請求項 13 記載の方法。

【請求項 15】

各コマンド意味構造にスコアを付けること、

最も高いスコアが付けられたコマンド意味構造を前記テキストの前記意味構造として選択すること

を更に含むことを特徴とする請求項 1 4 記載の方法。

【請求項 1 6】

前記意味構造の前記構文スコアと前記意味スコアに基づき、追加的条件から意味構造候補を枝刈りすることを更に含むことを特徴とする請求項 1 記載の方法。

【請求項 1 7】

コンピュータに、

テキストから意味エンティティを識別するステップと、

前記意味エンティティから 1 組の部分的意味構造を形成するステップと、

前記部分的意味構造の各々にスコアを付け、各部分的意味構造のスコアを提供するステップと、

少なくとも一つの部分的意味構造を枝刈りし、前記部分的意味構造の前記スコアに基づき少なくとも一つの部分的意味構造を保持するステップと、

保持された部分的意味構造から完全な意味構造を形成するステップであって、前記完全な意味構造は、前記テキストから識別される全ての意味エンティティを含むステップと、

前記完全な意味構造がコマンド意味構造でないと決定するステップと、

前記完全な意味構造を複数のコマンド意味構造に配置するステップと、

各コマンド意味構造にスコアを付けるステップと、

前記スコアに基づき前記コマンド意味構造の 1 つを選択するステップ

を実行させるためのコンピュータ実行可能命令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 1 8】

意味エンティティを識別するステップは、意味エンティティがテキストと関連するスキーマを利用することを含むことを特徴とする請求項 1 7 記載のコンピュータ読み取り可能な記録媒体。

【請求項 1 9】

1 組の部分的意味構造を形成するステップは、

テーブルにアクセスして、前記テキストから識別される 2 つの意味エンティティによってインデックス付けされるエントリを見つけることであって、前記エントリは、前記 2 つの意味エンティティの両方を含む複数の意味構造を提供すること、

前記エントリによって提供される前記意味構造の各々の部分的意味構造を形成することを含むことを特徴とする請求項 1 7 記載のコンピュータ読み取り可能な記録媒体。

【請求項 2 0】

意味エンティティ間の関係を記述するスキーマから前記テーブルを構築するステップを更に含むことを特徴とする請求項 1 9 記載のコンピュータ読み取り可能な記録媒体。

【請求項 2 1】

少なくとも一つの部分的意味構造は、前記テキストから識別されない意味エンティティを含むことを特徴とする請求項 1 9 記載のコンピュータ読み取り可能な記録媒体。

【請求項 2 2】

部分的意味構造にスコアを付けるステップは、

前記部分的意味構造内の各エンティティのスコアを計算することを含み、エンティティの前記スコアは、前記エンティティが前記テキストから識別されなかった場合に減らされることを特徴とする請求項 2 1 記載のコンピュータ読み取り可能な記録媒体。

【請求項 2 3】

完全な意味構造を形成するステップは、

前記テキストから識別される意味エンティティと、部分的意味構造とを結合することを含むことを特徴とする請求項 1 7 記載のコンピュータ読み取り可能な記録媒体。

【請求項 2 4】

完全な意味構造を形成するステップは、

前記意味エンティティが前記部分的意味構造内の前記エンティティのいずれかの下位に配置できるかどうかを判定すること、

前記部分的意味構造の先端が意味エンティティの下位に配置できるかどうかを判定すること、

前記部分的意味構造の先端と前記意味エンティティの両方が共通上位エンティティの下位に配置できるかどうかを判定すること

を含むことを特徴とする請求項23記載のコンピュータ読み取り可能な記録媒体。

【請求項25】

部分的意味構造にスコアを付けるステップは、

前記部分的意味構造内の前記意味エンティティと、前記部分的意味構造内の前記意味エンティティ間の関係に基づいて意味スコアを決定することを含むことを特徴とする請求項17記載のコンピュータ読み取り可能な記録媒体。

【請求項26】

部分的意味構造にスコアを付けるステップは、

前記意味エンティティと関連する前記テキスト内の単語の位置に基づいて構文スコアを決定することを更に含むことを特徴とする請求項25記載のコンピュータ読み取り可能な記録媒体。

【請求項27】

部分的意味構造にスコアを付けるステップは、

意味エンティティの順序に基づいて遷移スコアを判定することを更に含み、前記順序は、前記意味エンティティを前記意味エンティティと関連する前記テキストに配置することによって設定されることを特徴とする請求項26記載のコンピュータ読み取り可能な記録媒体。

【請求項28】

遷移確率を生成するために使用される値を検索することは、前記エンティティ対の順序付けがトレーニングデータ内で見られる回数を検索することを含むことを特徴とする請求項1記載の方法。

【請求項29】

遷移確率を生成することは、前記エンティティ対の順序付けが前記トレーニングデータ内で見られる回数を、前記エンティティ対の順序づけが前記トレーニングデータ内で見られる回数の合計で割ることを含むことを特徴とする請求項28記載の方法。