

【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 17 年 9 月 8 日 (2005.9.8)

【公開番号】特開 2005-135039(P2005-135039A)
 【公開日】平成 17 年 5 月 26 日 (2005.5.26)
 【年通号数】公開・登録公報 2005-020
 【出願番号】特願 2003-368241(P2003-368241)
 【国際特許分類第 7 版】
 G 0 6 F 17/28
 【F I】
 G 0 6 F 17/28 Z

【手続補正書】
 【提出日】平成 17 年 5 月 13 日 (2005.5.13)
 【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】

第 1 の言語の入力文を句単位で統語解析し、統語解析の結果を示す木構造を出力するための統語解析手段と、

前記第 1 の言語と第 2 の言語との対訳文からなる用例を複数個格納した用例データベースとを含む機械翻訳装置であって、

前記用例に含まれる前記対訳文は各々句ごとに予め互いに対応付けられており、かつ前記対訳文のうち、前記第 2 の言語の文の各句は、訳語情報と、当該句に付随する、関係情報及びモダリティ情報を含む付随情報とを含む木構造形式で格納されており、前記第 1 の言語の文の各句は、句単位の木構造形式で格納されており、

前記機械翻訳装置はさらに、

前記統語解析手段から前記木構造を受け、当該木構造に含まれる句ごとに、訳語情報を前記用例データベースから抽出するための訳語抽出手段と、

前記統語解析手段から前記木構造を受け、当該木構造の用言パターンと、対訳文のうちの前記第 1 の言語の文の用言パターンが一致する用例を前記用例データベースにおいて検索するための用例検索手段と、

前記訳語抽出手段により抽出された訳語情報と、前記用例検索手段により検索された用例とに基づいて、前記入力文の前記第 2 の言語の訳文を生成するための訳文生成手段とを含む、機械翻訳装置。

【請求項 2】

前記訳文生成手段は、

前記訳語抽出手段により抽出された訳語情報を、前記用例検索手段により検索された用例の木構造と結合し、前記第 2 の言語の木構造を生成するための結合手段と、

前記結合手段により生成された前記第 2 の言語の木構造から、語順決定のための所定の規則にしたがって前記第 2 の言語に含まれる句を直列に配列し、かつ前記第 2 の言語の木構造の前記付随情報を用いて各句の文法的形態を設定するための直列化手段とを含む、請求項 1 に記載の機械翻訳装置。

【請求項 3】

前記結合手段は、前記訳語抽出手段により抽出された各訳語情報で、前記用例検索手段により検索された用例の木構造中の、当該各訳語情報に対応する訳語情報を置換することで

、前記第 2 の言語の木構造を生成するための訳語情報置換手段を含む、請求項 1 又は請求項 2 に記載の機械翻訳装置。

【請求項 4】

前記訳語抽出手段により抽出された訳語情報に対応する句が所定の条件を充足しているときに、当該句の付随情報を抽出するための付随情報抽出手段をさらに含み、

前記結合手段は、前記付随情報が前記付随情報抽出手段により抽出されたときには、前記訳語情報置換手段により生成される前記第 2 の言語の木構造中において当該句に対応する句の付随情報を、前記付随情報抽出手段により抽出された付随情報で置換するための付随情報置換手段をさらに含む、請求項 3 に記載の機械翻訳装置。

【請求項 5】

前記訳語抽出手段は、前記統語解析手段から前記入力文の前記木構造を受け、当該木構造中の句と内容語が一致する前記第 1 の言語の句を含む用例を前記用例データベースにおいて検索し、検索された用例の、当該第 1 の言語の句に対応付けられている訳語情報を抽出するための手段を含む、請求項 4 に記載の機械翻訳装置。

【請求項 6】

前記付随情報抽出手段は、前記抽出するための手段により検索された用例の、前記第 1 の言語の句の係り先の句が、前記木構造中の対応する句の係り先の句と一致するときに、当該第 1 の言語の句の付随情報を抽出するための手段を含む、請求項 5 に記載の機械翻訳装置。

【請求項 7】

前記付随情報抽出手段は、前記抽出するための手段により検索された用例の、前記第 1 の言語の句の主要語の品詞が、前記木構造中の対応する句の品詞と一致するときに、当該第 1 の言語の句の付随情報を抽出するための手段を含む、請求項 5 に記載の機械翻訳装置。

【請求項 8】

コンピュータにより実行されると、当該コンピュータを請求項 1 ～請求項 7 のいずれかに記載の機械翻訳装置として動作させる、機械翻訳プログラム。

【手続補正 2】

【補正対象書類名】明細書

【補正対象項目名】0003

【補正方法】変更

【補正の内容】

【0003】

＜内容語＞

内容語とは、強い意味を持つ語のことをいう。日本語、英語以外の言語でも同様である。本明細書では、品詞によって次の語を内容語とする。すなわち、日本語では形式名詞、助詞、又は助動詞以外の品詞を持つ語のことをいい、英語では、前置詞、冠詞、又は to - 不定詞以外の品詞を持つ語のことを言う。

【手続補正 3】

【補正対象書類名】明細書

【補正対象項目名】0016

【補正方法】変更

【補正の内容】

【0016】

- 従来技術 -

用例ベース翻訳とは、予めある言語（例えば日本語）の文と、その文を別の言語（例えば英語）に翻訳した文とが対になった翻訳用例（以下「用例」と呼ぶ。）を多数格納した用例 DB を準備しておき、入力文が与えられると、この入力文と類似した文を含む用例を用例 DB から探し出し、その用例に含まれる翻訳を組合せて翻訳を行なう機械翻訳手法である。

【手続補正 4】

【補正対象書類名】明細書
【補正対象項目名】0030
【補正方法】変更
【補正の内容】
【0030】

好ましくは、訳文生成手段は、訳語抽出手段により抽出された訳語情報を、用例検索手段により検索された用例の木構造と結合し、第2の言語の木構造を生成するための結合手段と、結合手段により生成された第2の言語の木構造から、語順決定のための所定の規則にしたがって第2の言語に含まれる句を直列に配列し、かつ第2の言語の木構造の付随情報を用いて各句の文法的形態を設定するための直列化手段とを含む。

【手続補正5】
【補正対象書類名】明細書
【補正対象項目名】0035
【補正方法】変更
【補正の内容】
【0035】

さらに好ましくは、付随情報抽出手段は、抽出するための手段により検索された用例の、第1の言語の句の主要語の品詞が、木構造中で対応する句の品詞と一致するときに、当該第1の言語の句の付随情報を抽出するための手段を含む。

【手続補正6】
【補正対象書類名】明細書
【補正対象項目名】0045
【補正方法】変更
【補正の内容】
【0045】

例えば、英語文62の1番目の句「P r e s i d e n t i a l E l e c t i o n」は、関係情報及びモダリティ情報としては何も持たず、訳語情報として「P r e s i d e n t i a l E l e c t i o n」に相当する木構造を格納している。2番目の句「i s s c h e d u l e d」は、関係情報としては何も持たないが、モダリティ情報として「b e + V B N（過去分詞形）」を含み、さらに訳語情報として「s c h e d u l e」を含む。3番目の句は関係情報として「t o」、訳語情報として「h o l d」、モダリティ情報として「b e + V B N（過去分詞形）」を含む。以下同様である。

【手続補正7】
【補正対象書類名】明細書
【補正対象項目名】0048
【補正方法】変更
【補正の内容】
【0048】

図4に、訳語抽出部38の機能を概略的に示す。図4を参照して、入力文が図4の左側に示すような木構造100で表される「検証する会議が」というものであったとする。これに対応する用例として、入力文の「検証する」という句と一致する日本語句「検証する」を持つ用例102を選択する。さらにこの場合、用例102の句「検証する」の係り先である「議論」の品詞が入力文の「会議」と同じであるため、訳語情報だけでなく関係情報及びモダリティ情報も用例102から抽出する。