

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2007-241910

(P2007-241910A)

(43) 公開日 平成19年9月20日(2007.9.20)

(51) Int. Cl. F I テーマコード(参考)  
**G06F 17/28 (2006.01)** G O 6 F 17/28 U 5 B O 9 1

審査請求 未請求 請求項の数 10 O L (全 16 頁)

(21) 出願番号	特願2006-66836 (P2006-66836)	(71) 出願人	301022471 独立行政法人情報通信研究機構 東京都小金井市貫井北町4-2-1
(22) 出願日	平成18年3月13日(2006.3.13)	(74) 代理人	100130111 弁理士 新保 斉
		(72) 発明者	内元 清貴 東京都小金井市貫井北町4-2-1 独立 行政法人情報通信研究機構内
		(72) 発明者	井佐原 均 東京都小金井市貫井北町4-2-1 独立 行政法人情報通信研究機構内
		Fターム(参考)	5B091 AA03 BA03 CD11 EA25

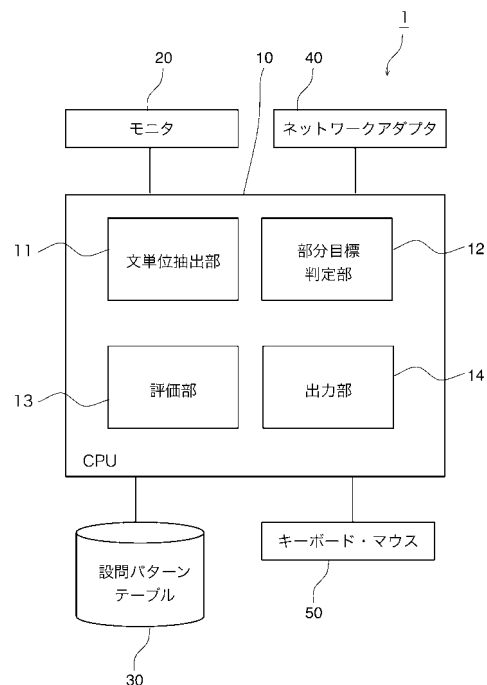
(54) 【発明の名称】 機械翻訳評価装置及び方法

(57) 【要約】

【課題】 機械翻訳の翻訳結果を自動評価する際に、単文のみを対象としても誤りの生じやすい翻訳結果を重点的に確認することができ、高速かつ簡便な方法で自動評価を行うこと。

【解決手段】 機械翻訳結果を自動的に評価する機械翻訳評価装置であって、記憶手段30に格納した、評価事項である部分目標について機械翻訳結果が適合するか否かを判定するための設問パターンを用い、機械翻訳結果である翻訳先テキストに対して、少なくとも該設問パターンに適合するか否かの結果を含む条件から部分目標に係る判定を行う部分目標判定手段12と、部分目標判定において適合した設問パターンと適合しなかった設問パターンに関する情報の少なくともいずれかに基づく所定の計算式を用いて機械翻訳結果の評価値を算出する評価手段13と、評価値を出力する出力手段14とを備える。

【選択図】 図1



**【特許請求の範囲】****【請求項 1】**

機械翻訳結果を自動的に評価する機械翻訳評価装置であって、  
記憶手段に格納した、評価事項である部分目標について機械翻訳結果が適合するか否かを判定するための設問パターンを用い、

機械翻訳結果である翻訳先テキストに対して、少なくとも該設問パターンに適合するか否かの結果を含む条件から部分目標に係る判定を行う部分目標判定手段と、

該部分目標判定において適合した設問パターンと適合しなかった設問パターンに関する情報の少なくともいずれかに基づく所定の計算式を用いて機械翻訳結果の評価値を算出する評価手段と、

評価値を出力する出力手段と

を備えたことを特徴とする機械翻訳評価装置。

10

**【請求項 2】**

前記所定の計算式において、設問パターンに応じて重み付けを行う

請求項 1 に記載の機械翻訳評価装置。

**【請求項 3】**

前記機械翻訳評価装置において、

機械翻訳結果である翻訳先テキストから文を抽出する文単位抽出手段を備え、

抽出された文に対して部分目標判定手段が処理する

請求項 1 又は 2 に記載の機械翻訳評価装置。

20

**【請求項 4】**

記憶手段に予め格納される、又は情報取得手段から取得される、翻訳元言語の翻訳元テキストと、機械翻訳結果である翻訳先テキストと、正解の翻訳結果である参照テキストとを用い、

該翻訳元テキストと、該参照テキストとを比較して予め定義されて記憶手段に格納された設問パターン書式に該当する語句を挿入することで設問パターンを定義する設問パターン定義手段を備えた

請求項 1 ないし 3 に記載の機械翻訳評価装置。

**【請求項 5】**

記憶手段に予め格納される、又は情報取得手段から取得される、翻訳元言語の翻訳元テキストと、機械翻訳結果である翻訳先テキストと、正解の翻訳結果である参照テキストとを用い、

該翻訳元テキストと、該参照テキストとを比較して予め定義された設問の中から使用する設問を選択する設問設定手段と、

選択された設問を参照して予め定義されて記憶手段に格納された設問パターン書式に該当する語句を挿入することで設問パターンを定義する設問パターン定義手段と

を備えた

請求項 1 ないし 3 に記載の機械翻訳評価装置。

30

**【請求項 6】**

機械翻訳結果を自動的に評価する機械翻訳評価方法であって、

評価事項である部分目標について機械翻訳結果が適合するか否かを判定するための設問パターンを用い、

部分目標判定手段が、機械翻訳結果である翻訳先テキストに対して、少なくとも該設問パターンに適合するか否かの結果を含む条件から部分目標に係る判定を行う部分目標判定ステップ、

評価手段が、該部分目標判定において適合した設問パターンと適合しなかった設問パターンに関する情報の少なくともいずれかに基づく所定の計算式を用いて機械翻訳結果の評価値を算出する評価ステップ、

出力手段が、評価値を出力する出力ステップ

を含むことを特徴とする機械翻訳評価方法。

40

50

**【請求項 7】**

前記所定の計算式において、設問パターンに応じて重み付けを行う  
請求項 6 に記載の機械翻訳評価方法。

**【請求項 8】**

前記機械翻訳評価方法において、  
文単位抽出手段が、機械翻訳結果である翻訳先テキストから文を抽出する文単位抽出ステップを含み、  
抽出された文に対して部分目標判定ステップ以下を処理する  
請求項 6 又は 7 に記載の機械翻訳評価方法。

**【請求項 9】**

翻訳元言語の翻訳元テキストと、機械翻訳結果である翻訳先テキストと、正解の翻訳結果である参照テキストとを用い、  
設問パターン定義手段が、該翻訳元テキストと、該参照テキストとを比較して予め定義された設問パターン書式に該当する語句を挿入することで設問パターンを定義する設問パターン定義ステップを含み、  
該定義された設問パターンを用いる  
請求項 6 ないし 8 に記載の機械翻訳評価方法。

10

**【請求項 10】**

翻訳元言語の翻訳元テキストと、機械翻訳結果である翻訳先テキストと、正解の翻訳結果である参照テキストとを用い、  
設問設定手段が、該翻訳元テキストと、該参照テキストとを比較して予め定義された設問の中から使用する設問を選択する設問設定ステップ、  
設問パターン定義手段が、  
選択された設問を参照して予め定義された設問パターン書式に該当する語句を挿入することで設問パターンを定義する設問パターン定義ステップ  
を含み、  
該定義された設問パターンを用いる  
請求項 6 ないし 8 に記載の機械翻訳評価方法。

20

**【発明の詳細な説明】****【技術分野】**

30

**【0001】**

本発明はコンピュータを用いて機械翻訳結果を自動的に評価する装置及び方法に係り、より詳しくは評価事項としての部分目標への適合の有無により評価を行う機械翻訳評価装置及び方法に関する。

**【背景技術】****【0002】**

機械翻訳技術において、機械翻訳の品質評価は重要な課題であると認識されてきた。近年、その品質評価を自動化しその性能を向上させようという試みが数多くなされている。(非特許文献 1 ないし 5 参照)。

自動評価の性能が向上することにより、機械翻訳システムの利用や改良が効率良くできるようになることが期待されるからである。例えば、非特許文献 4 によれば、自動評価の指標をシステムパラメータのチューニングに利用することにより、翻訳性能が向上したという報告がある。この報告は、自動評価の性能が向上すれば翻訳性能が自動的に向上することを示している。

40

**【0003】**

しかし、これまでに提案されている自動評価手法では、300 文程度のまとまったデータがあれば翻訳システムの優劣を判別することができるが、個々の文について各システムの翻訳の優劣を判別するのは難しい。

**【0004】**

例えば、従来技術として、特許文献 1 の技術では、機械翻訳文を解析する第 1 翻訳文

50

解析ステップ、望ましい翻訳文を解析する第2翻訳文解析ステップ、二つの翻訳文の間の語句の対応関係を解析する語句対応づけステップ、対応づけられた語句の割合と対応する語句の一致度に基づいて二つの翻訳文の一致度を算出する翻訳文一致度算出ステップによって機械翻訳文を評価することが提案されている。

【0005】

このように、望ましい翻訳文と、機械翻訳文との一致度(類似度)を利用する方法は一般的であるが、1つの文章間だけでは翻訳結果の適切な評価は行えない。

【0006】

また、特許文献2に開示される技術では、第1の言語で表わされた原文を取得し、順方向機械翻訳部によって、原文を第2の言語の順方向翻訳文に機械翻訳し、3種類の逆方向機械翻訳部によって、順方向翻訳文を第1の言語に機械翻訳して、3個の逆方向翻訳文を生成し、3個の逆方向翻訳文と原文とに基づいて、順方向翻訳文の良否を評価する。

10

【0007】

この方法は逆方向翻訳の結果を用いて、順方向翻訳の性能を測るものであるが、複数の逆方向機械翻訳部を用いなければならない上に、各逆方向機械翻訳部の精度が相当に高くなければ、結局正確な評価を行うことは難しい。

【0008】

さらに、特許文献3の技術では、機械的に翻訳された翻訳文の善し悪しを自動的に評価する機械翻訳文の自動評価方法として、学習ステップと評価ステップとからなるものを提案している。

20

学習ステップは、予め人手によってランク付けが行なわれているとともに、予め人手によって正解翻訳文が用意されている複数の学習用機械翻訳文それぞれを、それに対する正解翻訳文との間の複数種類の編集距離を用いて、編集距離の種類数に応じた要素からなるベクトルに符号化するステップ、および得られたベクトルを訓練事例としてランク付け規則を学習するステップを備えている。

【0009】

この方法も、複数の機械翻訳文と正解翻訳文とを比較することで評価を行うものであるから、特許文献1と同様の問題を有している。

【0010】

【非特許文献1】Sonja Niessen, Franz Josef Och, Gregor Leusch, and Hermann Ney. An Evaluation Tool for Machine Translation: Fast Evaluation for MT Research. In Proceedings of the LREC 2000, pp. 39-45, 2000年

30

【非特許文献2】Kishore Papineni, Salim Roukos, Todd Ward, and Weiing. Zhu. BLEU: a Method for Automatic Evaluation of Machine Translation. In Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL), pp. 311-318, 2002年

【非特許文献3】NIST. Automatic Evaluation of Machine Translation Quality Using N-gram Co-Occurrence Statistics. Technical report, NIST, 2002年

【非特許文献4】Franz Josef Och. Minimum Error Training in Statistical Machine Translation. In Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics (ACL), pp. 160-167, 2003年

40

【非特許文献5】Joseph P. Turian, Luke Shen, and I. Dan Melamed. Evaluation of Machine Translation and its Evaluation. In Proceedings of the MT Summit IX, pp. 386-393, 2003年

【特許文献1】特開平05-298360号公報

【特許文献2】特開2005-078318号公報

【特許文献3】特開2003-085172号公報

【発明の開示】

【発明が解決しようとする課題】

【0011】

50

本発明は、上記従来技術の有する問題点に鑑みて創出されたものであり、機械翻訳の翻訳結果を自動評価する際に、単文のみを対象としても誤りの生じやすい翻訳結果を重点的に確認することができ、高速かつ簡便な方法で自動評価を行うための技術を提供する。

【課題を解決するための手段】

【0012】

本発明は次のような手段により上記の課題の解決を図る。

一般に、ある文を翻訳する際には、英日翻訳で言えば前置詞や不定詞の訳し分けのように翻訳品質を良好に保つために満たすべき条件がひとつ以上存在する。

それらの条件を設問の形で各テスト文に付与したテストセットと、個々の設問に対する回答を自動推定するシステムを作成することにより、従来手法に比べ個々の翻訳文の品質をより適切に自動評価することが可能となる。

10

【0013】

この具体的な方法として、請求項1に記載の機械翻訳評価装置は、記憶手段に格納した、評価事項である部分目標について機械翻訳結果が適合するか否かを判定するための設問パターンを用い、機械翻訳結果である翻訳先テキストに対して、少なくとも該設問パターンに適合するか否かの結果を含む条件から部分目標に係る判定を行う部分目標判定手段と、該部分目標判定において適合した設問パターンと適合しなかった設問パターンに関する情報の少なくともいずれかに基づく所定の計算式を用いて機械翻訳結果の評価値を算出する評価手段と、評価値を出力する出力手段とを備えたことを特徴とする。

【0014】

20

請求項2に記載の機械翻訳評価装置は、前記所定の計算式において、設問パターンに応じて重み付けを行うことを特徴とする。

【0015】

請求項3に記載の機械翻訳評価装置は、機械翻訳結果である翻訳先テキストから文を抽出する文単位抽出手段を備え、抽出された文に対して部分目標判定手段が処理することを特徴とする。

【0016】

請求項4に記載の機械翻訳評価装置は、記憶手段に予め格納される、又は情報取得手段から取得される、翻訳元言語の翻訳元テキストと、機械翻訳結果である翻訳先テキストと、正解の翻訳結果である参照テキストとを用い、該翻訳元テキストと、該参照テキストとを比較して予め定義されて記憶手段に格納された設問パターン書式に該当する語句を挿入することで設問パターンを定義する設問パターン定義手段を備えたことを特徴とする。

30

【0017】

請求項5に記載の機械翻訳評価装置は、記憶手段に予め格納される、又は情報取得手段から取得される、翻訳元言語の翻訳元テキストと、機械翻訳結果である翻訳先テキストと、正解の翻訳結果である参照テキストとを用い、該翻訳元テキストと、該参照テキストとを比較して予め定義された設問の中から使用する設問を選択する設問設定手段と、選択された設問を参照して予め定義されて記憶手段に格納された設問パターン書式に該当する語句を挿入することで設問パターンを定義する設問パターン定義手段とを備えたことを特徴とする。

40

【0018】

また、本発明は次のような機械翻訳評価方法を提供することもできる。

すなわち、請求項6に記載の機械翻訳評価方法は、評価事項である部分目標について機械翻訳結果が適合するか否かを判定するための設問パターンを用い、部分目標判定手段が、機械翻訳結果である翻訳先テキストに対して、少なくとも該設問パターンに適合するか否かの結果を含む条件から部分目標に係る判定を行う部分目標判定ステップ、評価手段が、該部分目標判定において適合した設問パターンと適合しなかった設問パターンに関する情報の少なくともいずれかに基づく所定の計算式を用いて機械翻訳結果の評価値を算出する評価ステップ、出力手段が、評価値を出力する出力ステップの各ステップを含む。

【0019】

50

請求項 7 に記載の機械翻訳評価方法は、前記所定の計算式において、設問パターンに応じて重み付けを行うことを特徴とする。

【0020】

請求項 8 に記載の機械翻訳評価方法は、上記に加えて文単位抽出手段が、機械翻訳結果である翻訳先テキストから文を抽出する文単位抽出ステップを含み、抽出された文に対して部分目標判定ステップ以下を処理することを特徴とする。

【0021】

請求項 9 に記載の機械翻訳評価方法は、翻訳元言語の翻訳元テキストと、機械翻訳結果である翻訳先テキストと、正解の翻訳結果である参照テキストとを用い、設問パターン定義手段が、該翻訳元テキストと、該参照テキストとを比較して予め定義された設問パターン書式に該当する語句を挿入することで設問パターンを定義する設問パターン定義ステップを含み、該定義された設問パターンを用いることを特徴とする。

10

【0022】

請求項 10 に記載の機械翻訳評価方法は、翻訳元言語の翻訳元テキストと、機械翻訳結果である翻訳先テキストと、正解の翻訳結果である参照テキストとを用い、設問設定手段が、該翻訳元テキストと、該参照テキストとを比較して予め定義された設問の中から使用する設問を選択する設問設定ステップ、設問パターン定義手段が、選択された設問を参照して予め定義された設問パターン書式に該当する語句を挿入することで設問パターンを定義する設問パターン定義ステップを含み、該定義された設問パターンを用いることを特徴とする。

20

【発明の効果】

【0023】

本発明は、上記構成を備えることにより次のような効果を奏する。

請求項 1 又は 6 に記載の発明によると、部分目標として設問パターンを用い、適合した設問パターンと適合しなかった設問パターンに関する情報というコンピュータ上の処理が容易な指標により機械翻訳結果を評価することができる。

特に、機械翻訳で生じる典型的な誤りについて重点的に評価することが可能であり、1 つの文章についても適切な評価が行える。

【0024】

請求項 2 又は 7 に記載の発明によると、設問パターンに応じて重み付けを行うことができるので、より高精度な機械翻訳評価を行うことができる。

30

【0025】

請求項 3 又は 8 に記載の発明によると、複数の文を含むテキストについても、文単位抽出手段が各文を抽出して、部分目標判定手段及び評価手段による評価を行うことができる。

【0026】

請求項 4 又は 9 に記載の発明によると、正解の翻訳結果である参照テキストを用いることで、設問パターン書式から設問パターンを自動的に定義することが可能であり、設問パターンの充実に寄与する。

【0027】

請求項 5 又は 10 に記載の発明によると、人手によって予め定義され、人間による評価を行いやすい設問をある程度用意しておくことで、翻訳元テキストと参照テキストとから適当な設問を設定することができる。さらに、設問をコンピュータによる処理が行いやすい設問パターンとして定義することができるので、人間によっても確認しやすく、かつ本発明の自動評価が可能な方法を提供することができる。

40

【発明を実施するための最良の形態】

【0028】

以下、本発明の実施形態を、図面に示す実施例を基に説明する。なお、実施形態は下記に限定されるものではない。

本発明に係る設問パターンを予め用意して処理する実施例 1、設問パターンを自動的に

50

定義する実施例 2、さらに設問を選択して設問パターンを定義する実施例 3 をそれぞれ説  
述する。

【実施例 1】

【0029】

本発明による機械翻訳評価装置（以下、本装置と呼ぶ。）を図 1 に示す。本装置（1）  
は、公知のパーソナルコンピュータによって構成することが便宜である。パーソナルコン  
ピュータには、周知のように演算・情報処理を司る CPU（10）と、表示装置であるモ  
ニタ（20）、ハードディスクなどの外部記憶装置（30）、インターネット等と接続し  
外部サーバ・クライアントからの情報の取得、出力を可能にするネットワークアダプタ（  
40）、ユーザが直接データを入力するキーボード・マウス（50）などを備えている。 10

【0030】

CPU（10）では、複数の文から各文を抽出する文単位抽出部（11）と、外部記憶  
装置（30）に格納された設問パターンテーブルに該当するか否かを判定する部分目標判  
定部（12）、部分目標の判定結果から機械翻訳結果の評価値を所定の計算式に基づいて  
算出する評価部（13）、該評価結果を出力する出力部（14）を備える。

【0031】

図 2 には機械翻訳評価方法に係るフローチャートを示す。

まず機械翻訳結果である翻訳先テキスト（100）を外部記憶装置からの読み込み又は  
ネットワークアダプタからの取得により CPU（10）に取得すると、文単位抽出部（1  
1）が文単位に分割して抽出する。（S1） 20

翻訳先テキストの言語によって公知の抽出方法を適宜選択することができるが、例えば  
日本語であれば読点や段落、改行コードなどをテキストから検索して、その位置で文が終  
了すると判定することができる。

【0032】

なお、本発明では予め文単位に分割された翻訳先テキストを入力してもよいので、必ず  
しも文単位抽出部（11）を備えていなくてもよい。

【0033】

次に、部分目標判定部（12）の処理（S2）に進む。

本発明では、機械翻訳評価を行うために、設問パターンを用いる。設問パターンは外部  
記憶装置（30）にデータテーブルの形式で格納されており、例えば次の表 1 のように定  
義される。 30

【0034】

【表 1】

翻訳元テキスト	設問パターン
I am taking him to the shop this afternoon.	○つれていく(つもり) ○つれていこう
The company now faces additional trouble.	○ています ている ×(いま げんざい)の(かいしゃ きぎ ょう)
It is important for him to come early.	○(かれが)はやくくる(こと の)は ×かれ(にとって のために)
He could swim.	○(およげるだろう およぐ(でしょう  であろう))

40

【0035】

上記表において、翻訳元テキストに対して、各設問パターンが定義されており、設問パ  
ターンの「～が含まれる」、×は「～が含まれない」を表す。

まず日本語を対象とする場合には、あらかじめ漢字を句読点を含まない仮名文に変換し 50

た後に処理を行う。第1文を「私は今日の午後、彼を店に連れて行く」と訳した場合には、「わたしはきょうのごごかれをみせにつれていく」と変換する。このような変換は辞書データベース(図示しない)を用いて仮名に置換することで容易に行える。さらに周知の形態素解析処理手段を用いて高精度な変換を行うこともできる。一例としてJUMAN(非特許文献6参照)を用いて処理することができる。

【0036】

【非特許文献6】<http://www.kc.t.u-tokyo.ac.jp/nl-resource/juman.html>

【0037】

ここで、第3文「It is important for him to come early.」においてfor himを「彼にとって」と訳するのは最適ではなく、本文では「彼が早く来ることは重要である」「早く来ることは、彼にとって重要です。」「彼が早く来ることは重要です。」などと訳するのがよい。

10

そこで、設問パターンには、「(かれが)はやくくる(こと|の)は」が含まれる、と定義しておき、翻訳先テキストに「かれが」が前置されても、されなくてもよく、続いて「はやくくることは」又は「はやくくるのは」が含まれるか否か、2択で判定できるようにしてある。

【0038】

同時に設問パターンに「かれ(にとって|のために)」が含まれない、と定義してあるので、翻訳先テキストに「かれにとって」又は「かれのために」が含まれないか否か、2択で判定できる。

20

このような設問パターンは多数用意しておき、翻訳元テキストも例えば第1文の主語について「(I am|I was|You are|You were|She is)」というように複数の変形を許容するように定義するか、あるいは全ての組み合わせを定義しておくのが望ましい。

【0039】

部分目標判定部(12)では、このような設問パターンテーブル(30)を参照して、翻訳先テキスト(100)が該当するか否か判定する。該判定は、上述した通りまず翻訳先テキスト(100)を仮名で表記された文章に変換処理した後、設問パターンに定義された文字列が含まれる、又は含まれないか否か比較処理することで判定することができる。

【0040】

なお、最も簡単な形態としては予め入力される翻訳先テキスト(100)に対して部分目標判定する設問パターンが決まっている場合が考えられるが、通常は翻訳先テキストと共に、翻訳先テキストへの対訳付けがされた翻訳元テキストを同時に入力し、部分目標判定部(12)において、翻訳元テキスト中に設問パターンテーブルの「翻訳元テキスト」に該当する語句・文が含まれている場合には、部分目標判定の処理を行うようにするのが望ましい。

30

これにより、任意の翻訳元テキスト・翻訳先テキストを入力することで、自動的に設問パターンを選択して部分目標を判定することができる。

【0041】

部分目標が判定された後、評価部(13)における評価(S3)を行う。

40

本発明における評価としては、入力された翻訳先テキスト(100)に対して部分目標判定で設問パターンと合致すると判定された個数と、合致しないと判定された個数とを用いて予め定めた計算式で算出する。

【0042】

その1つの方法は数1によるものであり、

(数1)

評価値 = (合致した個数) / (合致した個数 + 合致しなかった個数)

により評価値を得る。この場合、評価値が1に近いほど正確な機械翻訳結果であると評価できる。しきい値を用いて、一定の評価値以上の場合に合格、それ未満の場合に不合格などと判定してもよい。

50



## 【 0 0 4 3 】

さらに、従来の技術でも用いられている類似度を考慮した評価を行うこともできる。この場合、翻訳先テキスト(100)の他に、予め翻訳元テキストについて正解である訳出例を外部記憶装置(30)に格納しておき、その訳出例を参照テキストとして用いる。

部分目標の達成度を $Q$ 、翻訳先テキストと参照テキストとの類似度を $S$ とし、ある翻訳先テキストに対する評価値 $A$ を次の式で定義する。

## 【 0 0 4 4 】

## 【数 2】

$$A = S + \lambda \times Q$$

10

## 【数 3】

$$Q = \begin{cases} 1 : \sum_{i=1}^n f_i \times Q_i > 0 \text{ のとき} \\ 0 : \sum_{i=1}^n f_i \times Q_i = 0 \text{ のとき} \\ -1 : \sum_{i=1}^n f_i \times Q_i < 0 \text{ のとき} \end{cases}$$

20

## 【 0 0 4 5 】

ここで、 $Q$ は $n$ 個の部分目標が選択された場合の部分目標達成度を表わす。 $f_i$ は部分目標達成度の重みである。 $Q$ の値は、部分目標の合否による多数決により決まり、達成されている部分目標の方が多ければ1、少なければ-1、同数であれば0となる。 $Q_i$ ( $0 \leq i < n$ )は $i$ 番目のカテゴリを持つ部分目標に対する達成度を表わす。

この $Q_i$ は部分目標が達成されていれば1、達成されていなければ-1の値を返す。 $f_i$ は、評価対象の翻訳先テキストに依存して設問テーブル中の項目で判定する場合には1、しない場合には0をとる。(部分目標判定部での選択を基に決定する)

## 【 0 0 4 6 】

$S$ は翻訳先テキストと参照テキストとの類似度である。類似度としては任意の方法で算出することができる。例えば、機械翻訳自動評価手法としてよく用いられるBLEU(非特許文献7を参照)を拡張した指標を用いる。BLEU scoreの計算式は次の式により表わされる。

30

## 【 0 0 4 7 】

【非特許文献7】Linguistic Data Annotation Specification: Assessment of Fluency and Adequacy in Arabic-English and Chinese-English Translations. <http://www ldc.upenn.edu/Projects/TIDES/Translation/TransAssess02.pdf>, 2002年

## 【 0 0 4 8 】

## 【数 4】

$$\log \text{BLEU} = \min\left(1 - \frac{r}{c}, 0\right) + \sum_{n=1}^N \frac{1}{N} \log p_n$$

40

## 【 0 0 4 9 】

数4において、 $r$ は参照文の単語長、 $c$ はテスト文の単語長、 $N$ は考慮する単語 $n$ -gramの最大の $n$ の値を表わす。 $p_n$ は次の式で表わされる。

## 【 0 0 5 0 】

## 【数 5】

$$p_n = \frac{\sum_{C \in \text{参照文集合}} \sum_{n\text{-gram} \in C} \text{Count}_{clip}(n\text{-gram})}{\sum_{C' \in \text{参照文集合}} \sum_{n\text{-gram}' \in C'} \text{Count}(n\text{-gram}' )}$$

## 【0051】

ここで、 $\text{Count}(n\text{-gram})$  はテスト文における単語  $n\text{-gram}$  の出現頻度を表わす。 $\text{Count}_{clip}(n\text{-gram})$  は、次の式で表わされる。 10

## 【0052】

## 【数 6】

$$\text{Count}_{clip}(n\text{-gram}) = \min(\text{Count}, \text{Max\_Ref\_Count})$$

## 【0053】

ここで、 $\text{Max Ref Count}$  はテスト文における単語  $n\text{-gram}$  の出現頻度を表わす。BLEUからの拡張は、単語  $n\text{-gram}$  を依存構造木から抽出する点、テスト文と参照文を入れ替えて計算した BLEU score ともとの BLEU score の調和平均を採用した点、シソーラスを用いて汎化した点の3点である。本技術については、非特許文献 8 に記載している。 20

## 【0054】

【非特許文献 8】 Kiyotaka Uchimoto, Naoko Hayashida, Toru Ishida, and Hitoshi Isahara. Automatic Rating of Machine Translatability. In Proceedings of the MT Summit X, pp. 235-242, 2005年

## 【0055】

以上の数 2 及び 3 による評価値 A は、翻訳先テキストと参照テキストとの類似度に加えて、設問パターンに合致した部分目標の多数決とその部分目標に対する重み から算出するものである。 30

## 【0056】

重み は、設問パターンテーブルに予め定義する。特に重要な設問パターンに重みをつけるようにすればよいが、次のように高精度な の設定を行うこともできる。

すなわち、予め複数の機械翻訳結果について fluency (流暢さ) と adequacy (妥当性) を人手によって数値評価し、 を例えば 0 から 0.1 ずつ増加させていって、両者の数値評価と評価部で計算される A との相関が最大となるときの を求める。これにより、評価部 (13) における評価の高精度化が図られる。

## 【0057】

ここで、重み について Q についてのみ設定したが、類似度 S についても重み S を用いて類似度をどの程度重視するか、重み付けしてもよい。 40

## 【0058】

以上のような評価部 (13) での評価値 (101) は出力部 (14) によって出力処理 (S4) される。出力方法としては、外部記憶装置への書き出しや、ネットワークアダプタを介してのデータ送信、モニタでの表示などが可能である。

## 【実施例 2】

## 【0059】

本発明の第 2 の実施例として、上記した設問パターンを、表 1 のように予め定義しておくのではなく、設問パターン書式に従って自動的に定義することも可能である。

本実施例では図 3 のように設問パターン定義部 (15) を CPU (10) に備え、外部 50

記憶装置（30）には設問パターン書式を予め定義しておく。

設問パターン書式としては表2に示すようなものを用いる

【0060】

【表2】

設問パターン書式
(語句)を含んでいるか?
(語句)を含んでいないか
(品詞)を用いているか?

10

【0061】

上記表1で「～を含んでいる」等と具体的に定義していたのに対して、表2では(語句)の部分は代入すべき空欄として定義されている。また、特定の品詞として用いられていなければならない場合も「(品詞)を用いているか?」という書式で定義されている。

【0062】

このような書式を用いて、設問パターン定義部(15)は(語句)や(品詞)に該当する言葉を自動的に定義する。図4は実施例2に係るフローチャートである。なお、設問パターン定義ステップ(S10)は部分目標判定(S2)よりも前であれば文単位抽出処理(S1)との先後は問わない。

20

【0063】

まず翻訳元テキスト(102)と、参照テキスト(103)とを設問パターン定義部(15)が読み出し、翻訳先テキストに含まれていなければならない語句や品詞を抽出する。その際、全ての訳語を抽出するのではなく、特徴的な訳語についてのみ設問パターンとして定義する。

【0064】

この方法としては、図示しない辞書データベースにおいて、複数の意味(例えばa、b、cの3種類の意味)を持つ単語Wが翻訳元テキスト(102)に含まれているか検索して抽出し、その単語が参照先テキスト(103)においてどの意味で用いられているかを比較し調べる。その結果がaの意味であったときに、設問パターン定義部(15)は設問パターン書式にあてはめて当該翻訳元の文に対して「aを含んでいるか?」という設問パターンを定義する。

30

【0065】

同時に、「bを含んでいないか」「cを含んでいないか」という設問パターンも定義する。

このとき、図示しない同義語データベースを用い、aの同義語a1,a2,a3についても「(a|a1|a2|a3)を含んでいるか?」という設問パターンとして定義してもよい。「含んでいないか」についても同様である。

【0066】

このような訳語選択の場合の他、例えば「They speak English in New Zealand」の翻訳元テキストに対して「ニュージーランドでは英語を話す」という参照テキストがあるとき、単語Theyについて辞書データベースから得られる「彼らは」「それらは」といった訳語が参照テキストには含まれていないことが両者の比較から検出できる。

40

このとき、設問パターン定義部(15)は「(かれらは|それらは)を含んでいないか」という設問パターンとして定義することができる。

【0067】

同様に「She lived there by herself」(翻訳元テキスト)に対して参照テキスト「彼女は独りで暮らしている」を比較すると、byの訳語である「よって」を含んでいないことが設問パターンとして定義できる。

50

同時に、両者の文の対訳関係が予め定義されているとき、又は周知の対訳関係の抽出処理によって判明するときには、「by herself」が「独りで」と訳されていることから、「ひとりでを含んでいるか？」という設問パターンとなる。

同義語データベースを用いれば、「ひとり(だけ|きり)で|たんどくで|たんしんで含んでいるか？」という設問パターンを定義できる。

【0068】

「The percentage of stomach cancer among the workers appears to be the highest for any asbestos workers」の翻訳元テキストに対して参照テキスト「労働者中の胃ガンの割合はアスベスト労働者で最高となるようだ」があるとき、対訳関係から appear to が「ようだ」であることがわかる。

10

「ようだ」を辞書データベースで検索すると助動詞であることもわかるから、設問パターン定義部(15)は appear to について「助動詞を含んでいるか？」という設問パターンを定義することができる。

このように、訳語の選択だけでなく、正しい品詞で用いられているかについて部分目標の評価を行うこともできる。

【0069】

翻訳元テキスト、翻訳先テキスト、参照テキストの間の語句間の対応付けを自動的に抽出する方法は公知であるが、例えば、非特許文献9に開示されるシステムでは多数の対訳文を入力することにより、自動的に精度良く単語間のアラインメントを行うことができる。

20

【0070】

【非特許文献9】<http://www.fjoch.com/GIZA++.html>

【0071】

そして、以上のように翻訳元テキストと、参照テキストの組み合わせから設問パターンを自動定義することができ、これを用いて実施例1の機械翻訳評価を行うことが可能である。このとき、翻訳元テキストや翻訳先テキストは、実施例2で設問パターンを定義したときと異なるテキストであってよい。

【実施例3】

【0072】

設問パターンは上記したようにコンピュータによって判定が容易に高速に判定が行えるものであるが、翻訳評価を行うための設問が人為的に付与されている場合に、その設問は必ずしも設問パターンと同様の処理が可能とは言えない。

30

例えば、非特許文献10で示されたテストセットが知られている。

【0073】

【非特許文献10】井佐原均，内野一，荻野紫穂，奥西稔幸，木下聡，柴田昇吾，杉尾俊之，高山泰博，土井伸一，永野正，成田真澄，野村浩郷．開発者の視点からの機械翻訳システムの技術的評価- テストセットを用いた品質評価法-．自然言語処理，Vol. 3，No. 3，pp. 83-102,1996年

【0074】

このテストセットに付与されている特徴的な情報としては、翻訳結果を評価するための yes/no 設問が挙げられる。この設問は、例えば、「for が「～で」のように原因・理由を表すように訳されていますか？」といったもので、この設問に対し人間が yes/no で回答することによって、翻訳結果を客観的に評価することができるようになっている。

40

【0075】

例えば、上記の例文「The percentage of stomach cancer among the workers appears to be the highest for any asbestos workers.」についての設問は「appear to が「ようだ」のように助動詞として訳されていますか？」と定義されている。

設問は主として文法的な観点からカテゴリ分けされており、上の設問は連鎖動詞に関するものである。

【0076】

50

本発明の第3の実施例は、このような複数の設問が定義されているテストセットを用いて設問パターンを定義するものである。図5は本実施例の機械翻訳評価装置、図6は本実施例のフローチャートである。

表3は、外部記憶装置(30)に格納される設問データベースの一例である。

【0077】

【表3】

翻訳元テキスト	設問
She lived there by herself	"by herself"が「独りで」のように訳されていますか？
The percentage of stomach cancer among the workers appears to be the highest for any asbestos workers	"for"が「一で」のように原因・理由を表すように訳されていますか？

10

【0078】

このような設問を多数用意した設問データベースを用い、設問設定部では、評価すべき翻訳先テキストについてどの設問を用いるのがよいか、選択する。この選択方法としては、参照テキストの中に設問データベースの翻訳元テキスト欄で定義されたのと同じ又は類似の文があるときに、当該設問を選択する。

翻訳元テキストでは表3に示したような具体的な文の他「\* live \* by \*self」(\*は任意、動詞は原形に変換して登録する)などのような形で定義してもよい。

20

【0079】

また、類似の文については、公知の方法によって類似度を算出し、設問データベース記載の翻訳元テキストと類似度の高いものに設問を選択するほか、重要な語句を翻訳元テキストに指定しておき、該単語を含むものについて設問を選択するようにすることができる。例えば「She [lived] there [by herself]」として重要な語句を[]で指定する。そして、[lived]と[by herself]が含まれる文が参照テキストから抽出された場合に、その文に対応する翻訳先テキストについて"by herself"が「独りでのように訳されていますか？」という設問を選択する。

30

【0080】

このように設問設定部(16)で設問を選択した後、設問パターン定義部(17)において、設問パターン書式に基づいて設問パターンを定義する。この処理は、人間にとってわかりやすい設問を、コンピュータ処理が容易な設問パターンに変換する処理とすることができる。

【0081】

例えば、「"by herself"が「独りで」のように訳されていますか？」という設問に対して表2の設問パターン書式では「(語句)を含んでいるか？」が妥当し、「独りを含んでいるか？」という設問パターンに変換する。

具体的には、表4に示すような設問・設問パターン変換テーブルを外部記憶装置(30)に備えておくのが簡便である。

40

【0082】

【表4】

設問	設問パターン
〇〇が△△のように訳されていますか？	△△を含んでいるか？
〇〇が△△のように□□を表すように訳されていますか？	(□□を表す語句)を含んでいるか？

【0083】

50

上段の例では、 の部分は不要であるから除去して、「 を含んでいるか？」という単純な設問パターンに変換する。 の同義語を加えてもよい。

下段の例では「 のように」は単なる例示であり、 を表すようにというさらに有益な情報が含まれているため、「 を表す」語句を辞書データベースから抽出して代入する。

【 0 0 8 4 】

このような変換テーブルを利用することによって、設問パターン定義部 ( 1 7 ) がコンピュータが処理しやすい設問パターンに変換する処理を行うことができる。

【 0 0 8 5 】

以上説述したとおり、本発明は機械翻訳結果を自動評価する方法であり、客観的かつ高速な評価を実現する。また、このようにして得られた評価値を機械翻訳装置にフィードバックすることによって、高い翻訳精度を実現する機械翻訳装置に寄与させることができる。

【 図面の簡単な説明 】

【 0 0 8 6 】

【 図 1 】 本発明の機械翻訳評価装置 ( 第 1 実施例 ) の構成図である。

【 図 2 】 本発明の機械翻訳評価方法 ( 第 1 実施例 ) のフローチャートである。

【 図 3 】 本発明の機械翻訳評価装置 ( 第 2 実施例 ) の構成図である。

【 図 4 】 本発明の機械翻訳評価方法 ( 第 2 実施例 ) のフローチャートである。

【 図 5 】 本発明の機械翻訳評価装置 ( 第 3 実施例 ) の構成図である。

【 図 6 】 本発明の機械翻訳評価方法 ( 第 3 実施例 ) のフローチャートである。

【 符号の説明 】

【 0 0 8 7 】

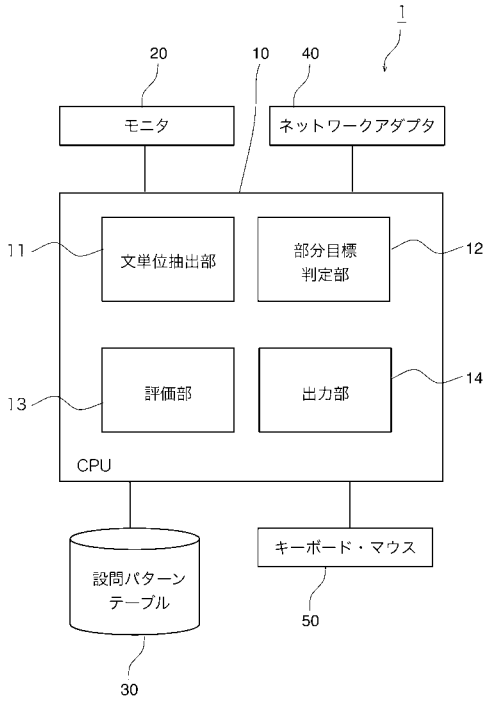
- 1 機械翻訳評価装置
- 1 0 CPU
- 1 1 文単位抽出部
- 1 2 部分目標判定部
- 1 3 評価部
- 1 4 出力部
- 2 0 モニタ
- 3 0 外部記憶装置
- 4 0 ネットワークアダプタ
- 5 0 キーボード・マウス

10

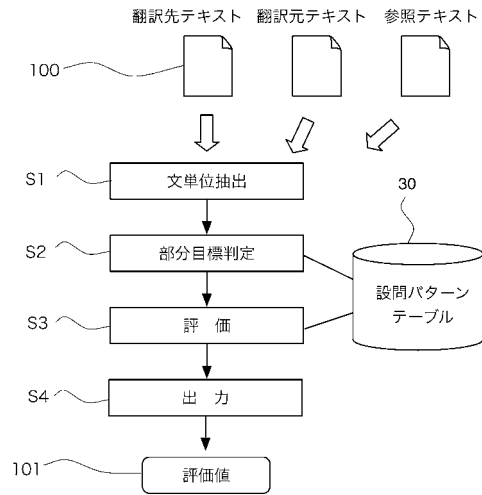
20

30

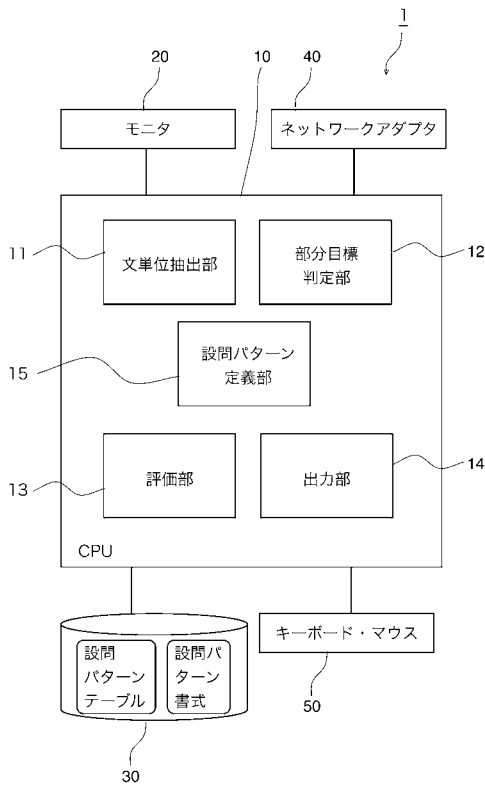
【 図 1 】



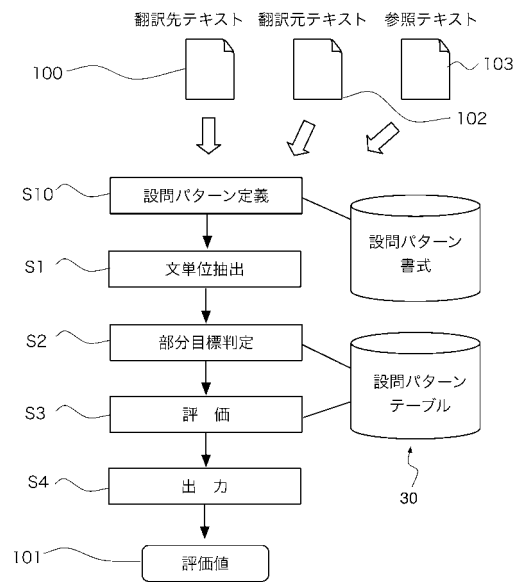
【 図 2 】



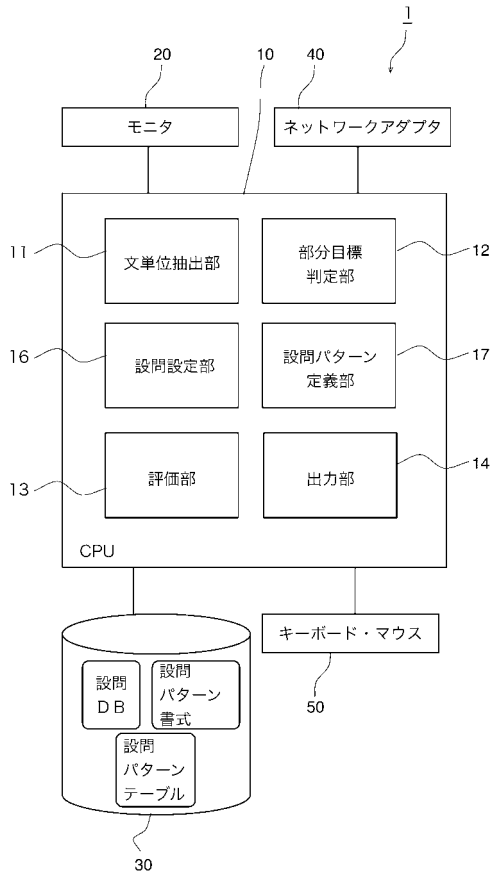
【 図 3 】



【 図 4 】



【 図 5 】



【 図 6 】

