

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2007-518164

(P2007-518164A)

(43) 公表日 平成19年7月5日(2007.7.5)

(51) Int. Cl.

G06F 17/28 (2006.01)

F I

G06F 17/28

Q

テーマコード (参考)

5B091

審査請求 未請求 予備審査請求 未請求 (全 26 頁)

(21) 出願番号 特願2006-546858 (P2006-546858)
 (86) (22) 出願日 平成17年1月3日 (2005.1.3)
 (85) 翻訳文提出日 平成18年7月4日 (2006.7.4)
 (86) 国際出願番号 PCT/KR2005/000065
 (87) 国際公開番号 W02005/065061
 (87) 国際公開日 平成17年7月21日 (2005.7.21)
 (31) 優先権主張番号 10-2004-0000801
 (32) 優先日 平成16年1月6日 (2004.1.6)
 (33) 優先権主張国 韓国 (KR)

(71) 出願人 506229763
 リ インーソブ
 大韓民国 121-080 ソウル マボ
 ーグ デハーンドン 85-24
 (74) 代理人 100059281
 弁理士 鈴木 正次
 (74) 代理人 100108947
 弁理士 涌井 謙一
 (74) 代理人 100117086
 弁理士 山本 典弘
 (74) 代理人 100124383
 弁理士 鈴木 一永
 (72) 発明者 リ インーソブ
 大韓民国 121-080 ソウル マボ
 ーグ デハーンドン 85-24

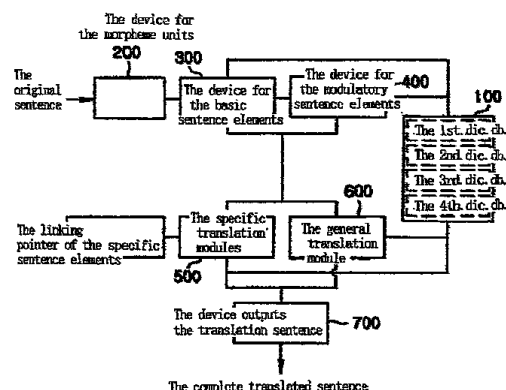
最終頁に続く

(54) 【発明の名称】 自動翻訳装置及びその自動翻訳装置を利用した自動翻訳方法並びにその自動翻訳装置が記録された記録媒体

(57) 【要約】

本発明は、自動翻訳装置及びその自動翻訳装置を利用した自動翻訳方法並びにその自動翻訳装置が記録された記録媒体に関し、より詳しくは、英語を中国語にあるいは中国語を英語に翻訳する場合、目的原語の対訳語文章をより自然かつ正確な翻訳に生成することができる。これは、中国語特有の様々な文法的特性を考慮して作製したアルゴリズムを本発明の自動翻訳装置及び記録媒体に採用することによって実現できた。前記特殊翻訳アルゴリズムは、本発明の基本概念であると同時に主要概念である。

本発明による自動翻訳装置は、辞典DB100、形態素分析手段200、基礎文章成分規定手段300、加工文章要素規定手段400、特殊文章要素翻訳手段500、一般文章要素翻訳手段600、完成翻訳文出力手段700で構成される。特殊文章成分翻訳手段500と連結されている特殊文章要素連結ポインターは、5種類の特殊翻訳モジュール501～505と連係して作動する。前記100～700、及び特殊翻訳モジュール501～505のプロセスを通じて英語と中国語間の翻訳をより



【特許請求の範囲】

【請求項1】

自動翻訳装置は、次のように構成することとを特徴とする。

1. 辞書DB保存手段100: 4つの相互関連の辞書データベースを有しており、4つの各辞書データベースには英語単語の全情報と、対訳中国語単語の情報がそれぞれの言語体系に基づいて記録されている。

2. 形態素分析手段200: 入力された原語文章の単語を形態素単位に分析する。

3. 基礎文章要素規定手段300: 形態素分析手段200で形態素単位に分析された単語の結果を受信し、各品詞で表示される基礎文章要素として規定し、辞書DB保存手段100に基づき入力した英語単語を中国語の対訳語に1次変換する。

4. 加工文章要素規定手段400 (変換翻訳モジュール; modulatory translation module): 基礎文章要素規定手段300で分析した結果を受信し、処理された各品詞を組み合わせて少なくとも2つ以上の品詞が合成語、熟語、あるいは定型化された慣用語句などに変換できるものがあるかどうかをチェックし、変換できるものが導出される場合、少なくとも2つ以上の該当する品詞を結合して、構文的機能を考慮した品詞に変換し、変換された転成品詞を新たに規定する (すなわち品詞が変わる、品詞の再転成が場合によって発生する)。そして、第1、第2、第3の辞書データベースDBに基づいて、基礎文章要素規定手段300により1次処理された、入力された英語単語の対訳中国語がここで再処理される。それは、品詞の再転成がここで発生するためである。

5. 特殊文章要素翻訳手段500: 基礎文章要素規定手段300及び加工文章要素規定手段400の結果、すなわち基礎文章要素あるいは加工文章要素をチェックし、この特殊文章要素翻訳手段500に内蔵されている特殊文章要素連結ポインター (pointer) と連結される識別因子が各単語の文章要素に存在するかどうかを判断する。一つ以上の原言語の文章要素あるいは文章構文から識別因子が検出された場合、一つ以上の原言語の文章要素あるいは文章構文は、500に内蔵された各特殊文章要素翻訳モジュールにより、辞書データベースに定型化された対訳中国語の文章要素あるいは文章構文に再処理され、これによって300及び400で処理された対訳中国語も再処理されながら中国語の語順も再処理される。特殊文章要素翻訳手段500に内蔵された各特殊文章要素翻訳モジュールには5種類のモジュール (501~505) があり、それぞれ独特の文法処理能力を有する連結ポインター (pointer) 及び各文章要素に付与された識別因子によって活性化される。

6. 一般文章要素翻訳手段600: 特殊文章要素翻訳手段500の結果を受信し、特殊文章要素連結ポインターにより再処理された単語を、辞書DB100に記録された中国語の語順規則に従って再構成する必要があるかどうかを判断し、必要に応じて再処理し、残りの原言語の文章要素あるいは文章構文を完成された文章要素に翻訳する一般文章要素翻訳手段600である。さらに、この段階において、中国語文法の独特の規則に従って、(1) 最終の中国語の対訳文章で不要な中国語単語を除去し、(2) 最終の中国語の対訳文章で必ず必要な中国語単語 (対訳語と関係なく) を挿入する過程が含まれる。

7. 完成翻訳文出力手段700: 完成された対訳中国語文章を出力する。

特殊文章要素翻訳手段500に内蔵された5種類の特殊翻訳モジュール (501~505) には、それぞれ互いに異なる処理能力を有する特殊文章要素連結ポインターが含まれている。

第1特殊翻訳モジュール501は、原語文章中にある動詞の中国語対訳動詞に特殊文章要素連結ポインターが存在するかどうか、さらに、▲離▼合▲詞▼識別因子を有するかどうかをチェックする。その二つが同時に存在する場合には、原語文章が過去時制であることを調べ、過去時制であれば、対訳語の文章要素として動詞を翻訳するとき、その翻訳された対訳語の動詞の音節の間に過去時制を表す“了”という文字を挿入して翻訳し、次いで、原語文章中に回数の副詞が存在するかどうかを調べ、それに対応する対訳語を過去表示の“了”の後に位置させて (原語文章が過去時制でない場合には回数の副詞のみ位置する)

10

20

30

40

50

翻訳する。

第2特殊翻訳モジュール502は、原言語の文章要素中に中国語対訳動詞の程度を表す中国語の補語性格の単語が含まれているかどうかを特殊文章要素連結ポインターによってチェックし、その中国語の程度補語が中国語対訳動詞を修飾するという識別因子が存在する場合には、対訳語の文章要素である動詞と中国語の補語性格の単語を翻訳するとき、翻訳された対訳語の動詞とその対訳語の程度補語の間に“得”という文字を挿入して翻訳する。

第3特殊翻訳モジュール503は、原言語の文章要素中に中国語対訳動詞の結果を表す中国語の補語性格の単語が含まれているかどうかを特殊文章要素連結ポインターによってチェックし、その対訳された中国語の結果補語が中国語対訳動詞を修飾するという識別因子が存在する場合には、対訳語の文章要素としてその単語を翻訳するとき、翻訳された対訳語の動詞の後にその中国語の対訳語の結果補語を置いて翻訳する。

第4特殊翻訳モジュール504は、原言語の文章要素中に動詞が「be動詞＋過去分詞」からなる被動文(受動態)であるかどうか、及び原語文章の過去分詞と対応する中国語対訳動詞がヒトや物の動作を支配するかどうか、また、影響を与えるかどうかを特殊文章要素連結ポインター及び識別因子によってチェックする。その結果、全て該当する場合には、被動文の原語文章を対訳語に翻訳するとき、その翻訳された対訳語文章の対訳動詞の前に“被”という文字を挿入するかを決定し、語順の再構成要求に従って翻訳する。

第5特殊翻訳モジュール505は、原言語の文章要素を構成する構文が「主語＋be動詞＋形容詞＋前置詞」からなる形容詞構文であるかどうかを特殊文章要素連結ポインター及び識別因子によってチェックする。その結果、形容詞構文である場合には、対訳語の文章要素を構成する中国語の形容詞構文に翻訳するときに、be動詞に対応する対訳語の動詞を省略し、語順の再構成要求に従って翻訳する。

辞書データベースに内蔵された4つの辞書のすべての英語単語と中国語単語には、それぞれ形態素情報、基礎文章要素情報、加工文章要素情報、特殊文章要素情報が含まれており、特に、特殊文章要素情報には識別因子検索表示が第1、第2、第3の辞書DBに付与されている。識別因子は、各英語単語の特殊文章要素情報をチェックし、5種類の特殊翻訳モジュール(501～505)中の特殊文章要素連結ポインターと連係して機能し、特殊文章要素情報を有する各英語単語を最も適切かつ正確に中国語対訳の単語と対応させて対訳結果を生成するようにする。

5種類の特殊翻訳モジュール(501～505)は、入力された英語文章を当該単語情報の再処理及び再加工過程に誘導し、最も適切かつ正確な最終の中国語対訳文章に変換、生成されるようにする重要な役割をする。

第1辞書DBは一般単語情報を有し、第2、第3辞書DBは合成語、慣用語、熟語などの情報を、第4辞書DBは専門用語情報を有する。

【請求項2】

請求項1に記載の自動翻訳装置は次のように構成することを特徴とする。

辞書DB100は、英語単語及び対訳中国語単語の情報がそれぞれの言語体系に沿って記録されている4つの辞書データベースを有し、第1、第2、第3の辞書DBには、英語と対応する中国単語の合成語、慣用語、熟語などの特殊文章要素情報が記録されている。

前記辞書DB100を参照して、中国語の対訳語及び入力された原英語単語が特殊文章要素翻訳手段500の特殊文章性分連結ポインターによって調査、検索される(500)

前記辞書DB100を参照して、特殊文章要素連結ポインターは、中国語対訳動詞及び入力された原英語の動詞に中国語▲離▼合▲詞▼動詞の識別因子が付与されているかどうかをチェックする。また、原英語文章が過去時制であるかどうか、さらに、原英語文章中に回数の副詞が存在するかどうかをチェックする。これが第1特殊翻訳モジュールの過程である(501)。

前記辞書DB100を参照して、特殊文章要素連結ポインターは、中国語対訳単語に程度補語の識別因子が付与されているかどうかをチェックする。そして、その中国語の程度

10

20

30

40

50

補語単語が中国語対訳動詞を修飾するという識別因子が付与されているかどうかをチェックする。これが第2特殊翻訳モジュールの過程である(502)。

前記辞書DB100を参照して、特殊文章要素連結ポインターは、中国語対訳単語に結果補語の識別因子が付与されているかどうかをチェックする。そして、その中国語の結果補語単語が中国語対訳動詞を修飾するという識別因子が付与されているかどうかをチェックする。これが第3特殊翻訳モジュールの過程である(503)。

前記辞書DB100を参照して、特殊文章要素連結ポインターは、中国語の対訳文及び原英語文に被動文構文を表示する被動文識別因子が付与されているかどうかをチェックする。次いで、原英語文章の過去分詞と対応する中国語対訳動詞がヒトや物の動作を支配するかどうか、また、影響を与えるかどうかを特殊文章要素連結ポインター及び識別因子が判断し、その後、語順の再構成要求の有無及び被動文の英語動詞の後にある英語前置詞に対応する中国語の対訳単語が中国語文法に符合するかどうかを判断する。これが第4特殊翻訳モジュールの過程である(504)。

前記辞書DB100を参照して、特殊文章要素連結ポインターは、中国語対訳構文及び原英語構文に形容詞構文識別因子が付与されているかどうかをチェックする。次いで、形容詞構文である場合には、中国語形容詞構文に翻訳するとき、be動詞に対応する対訳語の動詞を省略し、中国語語順の再構成要求に従って翻訳する。これが第5特殊翻訳モジュールの過程である(505)。

【請求項3】

前記自動翻訳装置及びその自動翻訳生成方法による記録媒体は、次のように構成することを特徴とする。

1. 辞書DB保存手段100：4つの相互関連の辞書データベースを有しており、4つの各辞書データベースには英語単語の全情報と、対訳中国語単語の情報がそれぞれの言語体系に基づいて記録されている。
2. 形態素分析手段200：入力された原語文章の単語を形態素単位に分析する。
3. 基礎文章要素規定手段300：形態素分析手段200で形態素単位に分析された単語の結果を受信し、各品詞で表示される基礎文章要素として規定し、辞書DB保存手段100に基づき入力した英語単語を中国語の対訳語に1次変換する。
4. 加工文章要素規定手段400 (変換翻訳モジュール; *modulatory translation module*)：基礎文章要素規定手段300で分析した結果を受信し、処理された各品詞を組み合わせて少なくとも2つ以上の品詞が合成語、熟語、あるいは定型化された慣用語句などに変換できるものがあるかどうかをチェックし、変換できるものが導出される場合、少なくとも2つ以上の該当する品詞を結合して、構文的機能を考慮した品詞に変換し、変換された転成品詞を新たに規定する(すなわち品詞が変わる、品詞の再転成が場合によって発生する)。そして、第1、第2、第3の辞書データベースDBに基づいて、基礎文章要素規定手段300により1次処理された、入力された英語単語の対訳中国語がここで再処理される。それは、品詞の再転成がここで発生するためである。
5. 特殊文章要素翻訳手段500：基礎文章要素規定手段300及び加工文章要素規定手段400の結果、すなわち基礎文章要素あるいは加工文章要素をチェックし、この特殊文章要素翻訳手段500に内蔵されている特殊文章要素連結ポインター(*pointer*)と連結される識別因子が各単語の文章要素に存在するかどうかを判断する。一つ以上の原言語の文章要素あるいは文章構文から識別因子が検出された場合、一つ以上の原言語の文章要素あるいは文章構文は、500に内蔵された各特殊文章要素翻訳モジュールにより、辞書データベースベースに定型化された対訳中国語の文章要素あるいは文章構文に再処理され、これによって300及び400で処理された対訳中国語も再処理されながら中国語の語順も再処理される。特殊文章要素翻訳手段500に内蔵された各特殊文章要素翻訳モジュールには5種類のモジュール(501~505)があり、それぞれ独特の文法処理能力を有する連結ポインター(*pointer*)及び各文章要素に付与された識別因子によって活性化される。

6. 一般文章要素翻訳手段600:特殊文章要素翻訳手段500の結果を受信し、特殊文章要素連結ポインターにより再処理された単語を、辞書DB100に記録された中国語の語順規則に従って再構成する必要があるかどうかを判断し、必要に応じて再処理し、残りの原言語の文章要素あるいは文章構文を完成された文章要素に翻訳する一般文章要素翻訳手段600である。さらに、この段階において、中国語文法の独特の規則に従って、(1)最終の中国語の対訳文章で不要な中国語単語を除去し、(2)最終の中国語の対訳文章で必ず必要な中国語単語(対訳語と関係なく)を挿入する過程が含まれる。

7. 完成翻訳文出力手段700:完成された対訳中国語文章を出力する。

特殊文章要素翻訳手段500に内蔵された5種類の特殊翻訳モジュール(501~505)には、それぞれ互いに異なる処理能力を有する特殊文章要素連結ポインターが含まれている。

10

第1特殊翻訳モジュール501は、原語文章中にある動詞の中国語対訳動詞に特殊文章要素連結ポインターが存在するかどうか、さらに、▲離▼合▲詞▼識別因子を有するかどうかをチェックする。その二つが同時に存在する場合には、原語文章が過去時制であることを調べ、過去時制であれば対訳語の文章要素として動詞を翻訳するとき、その翻訳された対訳語の動詞の音節の間に過去時制を表す“了”という文字を挿入して翻訳し、次いで、原語文章中に回数の副詞が存在するかどうかを調べ、それに対応する対訳語を過去表示の“了”の後に位置させて(原語文章が過去時制でない場合には回数の副詞のみ位置する)翻訳する。

第2特殊翻訳モジュール502は、原言語の文章要素中に中国語対訳動詞の程度を表す中国語の補語性格の単語が含まれているかどうかを特殊文章要素連結ポインターによってチェックし、その中国語の程度補語が中国語対訳動詞を修飾するという識別因子が存在する場合には、対訳語の文章要素である動詞と中国語の補語性格の単語を翻訳するとき、翻訳された対訳語の動詞とその対訳語の程度補語の間に“得”という文字を挿入して翻訳する。

20

第3特殊翻訳モジュール503は、原言語の文章要素中に中国語対訳動詞の結果を表す中国語の補語性格の単語が含まれているかどうかを特殊文章要素連結ポインターによってチェックし、その対訳された中国語の結果補語が中国語対訳動詞を修飾するという識別因子が存在する場合には、対訳語の文章要素としてその単語を翻訳するとき、翻訳された対訳語の動詞の後にその中国語の対訳語の結果補語を置いて翻訳する。

30

第4特殊翻訳モジュール504は、原言語の文章要素中に動詞が「be動詞+過去分詞」からなる被動文(受動態)であるかどうか、及び原語文章の過去分詞と対応する中国語対訳動詞がヒトや物の動作を支配するかどうか、また、影響を与えるかどうかを特殊文章要素連結ポインター及び識別因子によってチェックする。その結果、全て該当する場合には、被動文の原語文章を対訳語に翻訳するとき、その翻訳された対訳語文章の対訳動詞の前に“被”という文字を挿入するかを決定し、語順の再構成要求の有無に従って翻訳する。

第5特殊翻訳モジュール505は、原言語の文章要素を構成する構文が「主語+be動詞+形容詞+前置詞」からなる形容詞構文であるかどうかを特殊文章要素連結ポインター及び識別因子によってチェックする。その結果、形容詞構文である場合には、対訳語の文章要素を構成する中国語の形容詞構文に翻訳するときに、be動詞に対応する対訳語の動詞を省略し、語順の再構成要求に従って翻訳する。

40

辞書データベースに内蔵された4つの辞書のすべての英語単語と中国語単語には、それぞれ形態素情報、基礎文章要素情報、加工文章要素情報、特殊文章要素情報が含まれており、特に、特殊文章要素情報には識別因子検索表示が第1、第2、第3の辞書DBに付与されている。識別因子は、各英語単語の特殊文章要素情報をチェックし、5種類の特殊翻訳モジュール(501~505)中の特殊文章要素連結ポインターと連係して機能し、特殊文章要素情報を有する各英語単語を最も適切かつ正確に中国語対訳の単語と対応させて対訳結果を生成するようにする。

5種類の特殊翻訳モジュール(501~505)は、入力された英語文章を当該単語情報の再処理及び再加工過程に誘導し、最も適切かつ正確な最終の中国語対訳文章に変換、生

50

成されるようにする重要な役割をする。

第1辞書DBは一般単語情報を有し、第2、第3辞書DBは合成語、慣用語、熟語などの情報を、第4辞書DBは専門用語情報を有する。

【請求項4】

請求項3に記載の自動翻訳装置の記録媒体は次のように構成することを特徴とする。

辞書DB100は、英語単語及び対訳中国語単語の情報がそれぞれの言語体系に沿って記録されている4つの辞書データベースを有し、第1、第2、第3の辞書DBには、英語と対応する中国単語の合成語、慣用語、熟語などの特殊文章要素情報が記録されている。

前記辞書DB100を参照して、中国語の対訳語及び入力された原英語単語が特殊文章要素翻訳手段500の特殊文章性分連結ポインターによって調査、検索される(500) 10

前記辞書DB100を参照して、特殊文章要素連結ポインターは、中国語対訳動詞及び入力された原英語の動詞に中国語▲離▼合▲詞▼動詞の識別因子が付与されているかどうかをチェックする。また、原英語文章が過去時制であるかどうか、さらに、原英語文章中に回数の副詞が存在するかどうかをチェックする。これが第1特殊翻訳モジュールの過程である(501)。

前記辞書DB100を参照して、特殊文章要素連結ポインターは、中国語対訳単語に程度補語の識別因子が付与されているかどうかをチェックする。そして、その中国語の程度補語単語が中国語対訳動詞を修飾するという識別因子が付与されているかどうかをチェックする。これが第2特殊翻訳モジュールの過程である(502)。 20

前記辞書DB100を参照して、特殊文章要素連結ポインターは、中国語対訳単語に結果補語の識別因子が付与されているかどうかをチェックする。そして、その中国語の結果補語単語が中国語対訳動詞を修飾するという識別因子が付与されているかどうかをチェックする。これが第3特殊翻訳モジュールの過程である(503)。

前記辞書DB100を参照して、特殊文章要素連結ポインターは、中国語の対訳文及び原英語文に被動文構文を表示する被動文識別因子が付与されているかどうかをチェックする。次いで、原英語文章の過去分詞と対応する中国語対訳動詞がヒトや物の動作を支配するかどうか、また、影響を与えるかどうかを特殊文章要素連結ポインター及び識別因子が判断し、その後、語順の再構成要求の有無及び被動文の英語動詞の後にある英語前置詞に対応する中国語の対訳単語が中国語文法に符合するかどうかを判断する。これが第4特殊翻訳モジュールの過程である(504)。 30

前記辞書DB100を参照して、特殊文章要素連結ポインターは、中国語対訳構文及び原英語構文に形容詞構文識別因子が付与されているかどうかをチェックする。次いで、形容詞構文である場合には、中国語形容詞構文に翻訳するときに、be動詞に対応する対訳語の動詞を省略し、中国語語順の再構成要求に従って翻訳する。これが第5特殊翻訳モジュールの過程である(505)。

【請求項5】

前記自動翻訳装置による翻訳方法は、次の段階を含むことを特徴とする。

ステップS100：原語文章を入力する段階。

ステップS200：入力された原語文章の単語を形態素単位に分析する段階。 40

ステップS300：形態素単位に分析された単語を各品詞として分析し、基礎文章要素として解釈する段階。

ステップS400：S300で分析された結果を受信し、処理した各品詞を組み合わせ、少なくとも二つ以上の品詞が合成語、熟語、または定型化された慣用語句などに変換できるものがあるかどうかをチェックする。変換できるものが検出される場合には、少なくとも二つ以上の品詞を結合して構文的機能を考慮した品詞に転換し、転換された転成品詞を含む加工文章要素として規定する段階。

ステップS500：基礎文章要素あるいは加工文章要素をチェックし、ステップS500の段階に内蔵されている特殊文章要素の連結ポインター(pointer)と連結される識別因子が各単語の文章要素に存在するかどうかを判断する。一つ以上の原言語の文章要 50

素あるいは文章構文から識別因子が検出される場合には、一つ以上の原言語の文章要素あるいは文章構文は、S 5 0 0 に内蔵された各特殊文章要素の翻訳モジュール 5 0 1 ～ 5 0 5 により辞書データベースに定型化された対訳中国語の文章要素あるいは文章構文に再処理される。これによって、S 3 0 0 及び S 4 0 0 で処理された対訳中国語も再処理されながら中国語の語順が再処理される段階。

ステップ S 6 0 0 : S 5 0 0 の結果を受信し、特殊文章要素連結ポインターによって再処理された単語を、辞書 D B に記録された中国語の語順規則に従って再構成する必要があるかどうかを判断し、必要に応じて再処理し、残りの原言語の文章要素あるいは文章構文を完成された文章要素に翻訳する段階である。さらに、この段階において、中国語文法の独特の規則に従って、(1) 最終の中国語の対訳文章で不要な中国語単語を除去し、(2) 最終の中国語の対訳文章で必ず必要な中国語単語(対訳語と関係なく)を挿入する過程が含まれる。

10

ステップ S 7 0 0 : 完成された対訳文章を出力する段階を含む。

なお、ステップ S 5 0 0 において、5 0 0 に内蔵された(第 1 特殊翻訳モジュール 5 0 1) 特殊文章要素連結ポインター及び識別因子は、原言語の文章要素に対応する中国語対訳動詞が▲離▼合▲詞▼であるかどうかを判断し、▲離▼合▲詞▼である場合には、対訳語の文章要素としてその動詞を翻訳するとき、その翻訳された対訳語の動詞の音節の間に過去時制を表す“了”という文字を挿入(原語文章が過去型である場合)したり、原語文章に回数の副詞が存在する場合には、“了”という文字の後に対訳された中国語の回数の副詞の対訳語を挿入(原語文章が過去型ではない場合には“了”は挿入せず、副詞の対訳語のみ挿入する)して翻訳する段階(S 5 0 1)と、

20

ステップ S 5 0 0 において、対訳された中国語文章の単語のうち、中国語程度補語が存在するかどうか、及びその程度補語が対訳された中国語文章の対訳動詞を修飾するかどうかを 5 0 0 に内蔵された(第 2 特殊翻訳モジュール 5 0 2) 特殊文章要素連結ポインター及び識別因子が判断する。その結果、中国語対訳動詞を修飾する場合には、中国語対訳動詞と対訳された中国語程度補語の間に“得”という文字を挿入し、その後、該当する対訳語の程度補語を配置して翻訳する段階(S 5 0 2)と、

ステップ S 5 0 0 において、対訳された中国語文章の単語のうち、中国語結果補語が存在するかどうか、及びその結果補語が対訳された中国語文章の対訳動詞を修飾するかどうかを 5 0 0 に内蔵された(第 3 特殊翻訳モジュール 5 0 3) 特殊文章要素連結ポインター及び識別因子が判断する。その結果、中国語対訳動詞を修飾する場合には、中国語対訳動詞のすぐ後に該当する対訳語の結果補語を置いて翻訳する段階(S 5 0 3)と、

30

ステップ S 5 0 0 において、第 4 特殊翻訳モジュール 5 0 4 の特殊文章要素連結ポインター及び識別因子が原語文章の文章構造を「b e 動詞＋過去分詞」からなる被動文(受動態)であるかどうかを判断し、被動文であると判断した場合には、原語文章の過去分詞と対応する中国語対訳動詞がヒトや物の動作を支配するかどうか、また、影響を与えるかどうかを判断する。その結果、全て該当する場合には、被動文の原語文章を対訳語に翻訳するとき、その翻訳された対訳語文章の対訳動詞の前に“被”という文字を挿入するかを決定し、その後、語順の再構成要求の有無に従って翻訳する段階(S 5 0 4)と、

ステップ S 5 0 0 において、第 5 特殊翻訳モジュール 5 0 5 の特殊文章要素連結ポインター及び識別因子が原語文章の文章構造を「b e 動詞＋形容詞＋前置詞」からなる形容詞構文であるかどうかを判断し、形容詞構文であると判断した場合には、対訳語の文章要素を構成する中国語の形容詞構文に翻訳するときに、b e 動詞に対応する中国語対訳動詞を省略し、語順の再構成要求に従って翻訳する段階(S 5 0 5)を含む。

40

ステップ S 5 0 0 において、特殊翻訳モジュール 5 0 1 ～ 5 0 5 及び内蔵された特殊文章要素連結ポインター、並びに単語に付与された識別因子が三位一体となって動作する。

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

本発明は、自動翻訳装置及びその自動翻訳装置を利用した自動翻訳方法並びにその自動

50

翻訳装置が記録された記録媒体に関し、より詳しくは、英語を中国語にあるいは中国語を英語に翻訳する場合、既存の種々の英語と中国語間の自動翻訳装置に比してより精度の高くかつ自然な翻訳文を生成することができる装置に関する。

【0002】

本発明は、英語を中国語にあるいは中国語を英語に翻訳する際に、中国語文法特有の様々な文法的特性を“特別なアルゴリズム(algorithms)”で処理した技術を利用することによって良質の翻訳を実現することができる。さらに、本発明は、本発明の出願者が2003年3月試製品(Demo Program;中国語特有の文法的特性を特別なアルゴリズム(algorithms)で処理した技術を採用)を開発し、理論上の問題を実証的に検証、証明した。

10

【背景技術】

【0003】

一般に、コンピュータの記憶装置に一冊の辞書や数冊の文法書などからなる数字、文字、記号などを適切に符号化して記憶させた後、翻訳要請に応じてある国の言語を別の国の言語に翻訳する自動翻訳装置には様々なものが開発されている。

【0004】

一例として、“プロテクターと構文ノードを有するドアフレームを利用した自動翻訳機(Machine Translator utilizing Doorframe of which including Protector and Syntax Node:韓国特許登録10-0329109号)以下、MTDと称する。”による翻訳過程は次のとおりである。

20

【0005】

MTDのユーザが対訳語に翻訳しようとする原始言語を入力すると、自動翻訳装置は、その入力された原始言語の文章を形態素単位に分析し、その分析された形態素単位の各々の単語の品詞を決定する。

【0006】

次に、品詞が決定された単語のうち、一つの単語や句として扱われる熟語、複合名詞などのような固定表現化して使われる単語は、一つの品詞として再決定し、その後、原始言語文章の品詞のうち、重要な役割をする品詞(動詞、接続詞、関係詞、記号)が分析されないようにプロテクターで保護し、その保護されたプロテクター間の部分構文が名詞句であるか、あるいは前置詞句であるかを分析し、適切な構文タグを付与する。

30

【0007】

このように、プロテクターと部分構文からなる原ドアフレームを原ドアフレームのデータベースとマッチングさせて、保護表示されたプロテクター(動詞、接続詞、関係詞、記号)と対応する対訳ドアフレームを選択する。ここで、対訳ドアフレームには前記原文のプロテクターで表示される各品詞が既に翻訳されており、その後、残りの構文タグが与えられた名詞句または前置詞句などを翻訳して、一つの完全な対訳文をMTDが完成される。

【0008】

上述したMTDは、英語と韓国語間の翻訳に適合しており、さらに中国語の場合にも、英語と文章要素の配置順序が一部類似するので英語と中国語間の翻訳にも一部適用可能である。しかしながら、次のような中国語特有の文法的特質のために、MTDによる翻訳は不十分であるという問題点がある。

40

【0009】

第1に、中国語を構成する品詞のうち動詞が、▲離▼合▲詞▼(中国語文法特有の動詞構文の一つであって、一つの単語が2音節以上を有する動詞の前音節と後音節が離れ、その間に原語文章が過去時制である場合には、過去表示の“了”という文字を挿入し、原語文章に回数の副詞が存在する場合には、それに対応する対訳語を過去表示文字の後に順に位置させる)構文動詞として使われる場合、中国語動詞は変換しないので、英語文章が過去時制であることを表す“了”という文字や英語文章の回数の副詞が動詞を構成する音節

50

の間に挿入されなければならない。しかし、過去時制を表す英語文章を中国語文章に翻訳する場合、“了”という文字や回数の副詞の対訳語を挿入せずに翻訳されるという問題点がある。結果的に、MTDは中国語の▲離▼合▲詞▼構文構造を翻訳することができない。

【0010】

第2に、中国語を構成する文章要素として補語があり、その補語が程度補語である場合、中国語文章では独特の構文処理が必要である。文章で程度補語が使われたときに、動詞と補語の間に“得”という文字を挿入して、程度を表していることを表示しなければならない。しかしながら、このような中国語の程度補語が使われた英語文章を中国語文章に翻訳する場合、前記のような中国語特有の文法的な機能を反映しないまま翻訳されるので、その意味が不明確になるという問題点がある。結果的に、MTDは中国語特有の程度補語構文構造を翻訳することができない。

10

【0011】

第3に、中国語を構成する文章要素として補語があり、その補語が結果補語である場合、中国語文章では独特の構文処理が必要である。文章で結果補語が使われたときに、中国語動詞のすぐ後に補語が位置するように語順を調整して表示しなければならない。しかしながら、このような中国語の結果補語が使われた英語文章を中国語文章に翻訳する場合、前記したような中国語特有の文法的な機能が反映されずに翻訳されるので、その意味が不明確になるという問題点がある。結果的に、MTDは中国語特有の結果補語構文構造を翻訳することができない。

20

【0012】

第4に、中国語を構成する文章が被動文である場合、その被動文に使用された中国語対訳文章の動詞が、ヒトや物の動作を支配するか、あるいは影響を与えるかどうかによって、“被”という文字を動詞の前に付けるか否かを決定する。ここで、被動文の英語文章を中国語文章に翻訳する場合、被という文字を付けるかどうか決定する決定手段がないので、翻訳が不明確になるという問題点がある。結果的に、MTDは中国語特有の被動文構文構造を翻訳することができない。

【0013】

第5に、中国語を構成する文章が形容詞構文である場合、その形容詞構文では英語文章のbe動詞が省略されて翻訳されるべきであるが、形容詞構文の英語文章を中国語文章に翻訳する場合、be動詞も一緒に翻訳されて翻訳が不明確になるという問題点がある。このような問題点は、中国語を英語に翻訳をする場合にも同様に発生する。結果的に、MTDは中国語特有の形容詞構文構造を翻訳することができない。

30

【0014】

精度の向上された英語と中国語間の自動翻訳装置を実現するために、上述したような五つの中国語特有の構文構造に対応できる的確な処理が要求される。

【発明の開示】

【発明が解決しようとする課題】

【0015】

本発明は、前記のような従来の問題点を解決するためになされたものであって、その目的は、中国語特有の五つの構文構造に対応することができる翻訳装置を提供することである。

40

【課題を解決するための手段】

【0016】

本発明に係る英語と中国語間の自動翻訳装置は、上述したような中国語特有の五つの構文構造に対応できないという問題点を解決するためのものであり、それを解決するために、英語と中国語間の自動翻訳装置に関連する5種類のアルゴリズム理論を発明し、さらにそのアルゴリズム理論に基づいた英語と中国語間の自動翻訳装置の試製品を開発した。これにより、アルゴリズム理論の整合性及び進歩性が検証された。

【0017】

50

本発明の自動翻訳装置は、次のように構成する。

【0018】

1. 辞書DB保存手段100: 4つの相互関連の辞書データベースを有しており、4つの各辞書データベースには英語単語の全情報と、対訳中国語単語の情報がそれぞれの言語体系に基づいて記録されている。

【0019】

2. 形態素分析手段200: 入力された原語文章の単語を形態素単位に分析する。

【0020】

3. 基礎文章要素規定手段300: 形態素分析手段200で形態素単位に分析された単語の結果を受信し、各品詞で表示される基礎文章要素として規定し、辞書DB保存手段100に基づき入力した英語単語を中国語の対訳語に1次変換する。 10

【0021】

4. 加工文章要素規定手段400 (変換翻訳モジュール; *modulatory translation module*): 基礎文章要素規定手段300で分析した結果を受信し、処理された各品詞を組み合わせて少なくとも2つ以上の品詞が合成語、熟語、あるいは定型化された慣用語句などに変換できるものがあるかどうかをチェックし、変換できるものが導出される場合、少なくとも2つ以上の該当する品詞を結合して、構文的機能を考慮した品詞に変換し、変換された転成品詞を新たに規定する (すなわち品詞が変わる、品詞の再転成が場合によって発生する)。そして、第1、第2、第3の辞書データベースDBに基づいて、基礎文章要素規定手段300により1次処理された、入力された英語単語の対訳中国語がここで再処理される。それは、品詞の再転成がここで発生するためである。 20

【0022】

5. 特殊文章要素翻訳手段500: 基礎文章要素規定手段300及び加工文章要素規定手段400の結果、すなわち基礎文章要素あるいは加工文章要素をチェックし、この特殊文章要素翻訳手段500に内蔵されている特殊文章要素連結ポインター (*pointer*) と連結される識別因子が各単語の文章要素に存在するかどうかを判断する。一つ以上の原言語の文章要素あるいは文章構文から識別因子が検出された場合、一つ以上の原言語の文章要素あるいは文章構文は、500に内蔵された各特殊文章要素翻訳モジュールにより、辞書データベースに定型化された対訳中国語の文章要素あるいは文章構文に再処理され、これによって300及び400で処理された対訳中国語も再処理されながら中国語の語順も再処理される。特殊文章要素翻訳手段500に内蔵された各特殊文章要素翻訳モジュールには5種類のモジュール (501~505) があり、それぞれ独特の文法処理能力を有する連結ポインター (*pointer*) 及び各文章要素に付与された識別因子によって活性化される。 30

【0023】

6. 一般文章要素翻訳手段600: 特殊文章要素翻訳手段500の結果を受信し、特殊文章要素連結ポインターにより再処理された単語を、辞書DB100に記録された中国語の語順規則に従って再構成する必要があるかどうかを判断し、必要に応じて再処理し、残りの原言語の文章要素あるいは文章構文を完成された文章要素に翻訳する一般文章要素翻訳手段600である。さらに、この段階において、中国語文法の独特の規則に従って、(1) 最終の中国語の対訳文章で不要な中国語単語を除去し、(2) 最終の中国語の対訳文章で必ず必要な中国語単語 (対訳語と関係なく) を挿入する過程が含まれる。 40

【0024】

7. 完成翻訳文出力手段700: 完成された対訳中国語文章を出力する。

【0025】

特殊文章要素翻訳手段500に内蔵された5種類の特殊翻訳モジュール (501~505) には、それぞれ互いに異なる処理能力を有する特殊文章要素連結ポインターが含まれている。

【0026】

第1特殊翻訳モジュール501は、原語文章中にある動詞の中国語対訳動詞に特殊文章要素連結ポインター(pointer)が存在するかどうか、さらに、▲離▼合▲詞▼識別因子を有するかどうかをチェックする。その二つが同時に存在する場合には、原語文章が過去時制であるかを調べ、過去時制であれば対訳語の文章要素として動詞を翻訳するとき、その翻訳された対訳語の動詞の音節の間に過去時制を表す“了”という文字を挿入して翻訳し、次いで、原語文章中に回数の副詞が存在するかどうかを調べ、それに対応する対訳語を過去表示の“了”の後に位置させて(原語文章が過去時制でない場合には回数の副詞のみ位置する)翻訳する。

【0027】

第2特殊翻訳モジュール502は、原言語の文章要素中に中国語対訳動詞の程度を表す中国語の補語性格の単語が含まれているかどうかを特殊文章要素連結ポインターによってチェックし、その中国語の程度補語が中国語対訳動詞を修飾するという識別因子が存在する場合には、対訳語の文章要素である動詞と中国語の補語性格の単語を翻訳するとき、翻訳された対訳語の動詞とその対訳語の程度補語の間に“得”という文字を挿入して翻訳する。 10

【0028】

第3特殊翻訳モジュール503は、原言語の文章要素中に中国語対訳動詞の結果を表す中国語の補語性格の単語が含まれているかどうかを特殊文章要素連結ポインターによってチェックし、その対訳された中国語の結果補語が中国語対訳動詞を修飾するという識別因子が存在する場合には、対訳語の文章要素としてその単語を翻訳するとき、翻訳された対訳語の動詞の後にその中国語の対訳語の結果補語を置いて翻訳する。 20

【0029】

第4特殊翻訳モジュール504は、原言語の文章要素中に動詞が「be動詞＋過去分詞」からなる被动文(受動態)であるかどうか、及び原語文章の過去分詞と対応する中国語対訳動詞がヒトや物の動作を支配するかどうか、また、影響を与えるかどうかを特殊文章要素連結ポインター及び識別因子によってチェックする。その結果、全て該当する場合には、被动文の原語文章を対訳語に翻訳するとき、その翻訳された対訳語文章の対訳動詞の前に“被”という文字を挿入するかを決定し、語順の再構成要求の有無に従って翻訳する。

【0030】

第5特殊翻訳モジュール505は、原言語の文章要素を構成する構文が「主語＋be動詞＋形容詞＋前置詞」からなる形容詞構文であるかどうかを特殊文章要素連結ポインター及び識別因子によってチェックする。その結果、形容詞構文である場合には、対訳語の文章要素を構成する中国語の形容詞構文に翻訳するときに、be動詞に対応する対訳語の動詞を省略し、語順の再構成要求に従って翻訳する。 30

【0031】

辞書データベースに内蔵された4つの辞書のすべての英語単語と中国語単語には、それぞれ形態素情報、基礎文章要素情報、加工文章要素情報、特殊文章要素情報が含まれており、特に、特殊文章要素情報には識別因子検索表示が第1、第2、第3の辞書DBに付与されている。識別因子は、各英語単語の特殊文章要素情報をチェックし、5種類の特殊翻訳モジュール(501～505)中の特殊文章要素連結ポインターと連係して機能し、特殊文章要素情報を有する各英語単語を最も適切かつ正確に中国語対訳の単語と対応させて対訳結果を生成するようにする。 40

【0032】

5種類の特殊翻訳モジュール(501～505)は、入力された英語文章を当該単語情報の再処理及び再加工過程に誘導し、最も適切かつ正確な最終の中国語対訳文章に変換、生成されるようにする重要な役割をする。

【0033】

第1辞書DBは一般単語情報を有し、第2、第3辞書DBは合成語、慣用語、熟語などの情報を、第4辞書DBは専門用語情報を有する。

【0034】

次に、入力された英語文章が中国語の対訳文章に翻訳生成される過程を説明する。

【0035】

1. 英語文章の動詞を中国語の動詞に対訳する際に、該動詞が▲離▼合▲詞▼構文の動詞に対訳されるとき、特殊文章要素連結ポインタ及び識別因子は、英語文章が過去時制であるかどうかをチェックし、同時に英語文章に回数の副詞が存在するかどうかをチェックする。

【0036】

2. 中国語対訳文章中に中国語補語(英語文章では主に副詞)が程度補語である場合には、特殊文章要素連結ポインタ及び識別因子は、その中国語補語が対訳された中国語動詞を修飾するかどうかをチェックする。

10

【0037】

3. 中国語対訳文章中に中国語補語(英語文章では主に副詞)が結果補語である場合には、特殊文章要素連結ポインタ及び識別因子は、その中国語補語が対訳された中国語動詞を修飾するかどうかをチェックする。

【0038】

4. 入力された英語文章が被動文(b e + p . p .)である場合には、特殊文章要素連結ポインタ及び識別因子は、対応する中国語対訳動詞がヒトや物の動作を支配するかどうか、また、影響を与えるかどうかを判断する。次いで、英語動詞の後に位置する英語前置詞に対応する中国語の対訳単語が中国語文法に符合するかを判断する。何故ならば、英語前置詞に対応する中国語の対訳単語は、中国語文法上の特別な対訳語条件を満たす必要があるためである。

20

【0039】

5. 入力された英語文章が形容詞構文である場合には、特殊文章要素連結ポインタ及び識別因子がそれをチェックして、入力された英語文章のb e 動詞に対応する中国語の対訳単語を除去し、中国語語順規則に従って語順を再配列する。

【0040】

上述した英語と中国語間の自動翻訳装置及びその自動翻訳生成方法による記録媒体は、次のように構成する。

【0041】

1. 辞書DB保存手段100:4つの相互関連の辞書データベースを有しており、4つの各辞書データベースには英語単語の全情報と、対訳中国語単語の情報がそれぞれの言語体系に基づいて記録されている。

30

【0042】

2. 形態素分析手段200:入力された原語文章の単語を形態素単位に分析する。

【0043】

3. 基礎文章要素規定手段300:形態素分析手段200で形態素単位に分析された単語の結果を受信し、各品詞で表示される基礎文章要素として規定し、辞書DB保存手段100に基づき入力した英語単語を中国語の対訳語に1次変換する。

【0044】

4. 加工文章要素規定手段400(変換翻訳モジュール;modulatory translation module):基礎文章要素規定手段300で分析した結果を受信し、処理された各品詞を組み合わせて少なくとも2つ以上の品詞が合成語、熟語、あるいは定型化された慣用語句などに変換できるものがあるかどうかをチェックし、変換できるものが導出される場合、少なくとも2つ以上の該当する品詞を結合して、構文的機能を考慮した品詞に変換し、変換された転成品詞を新たに規定する(すなわち品詞が変わる、品詞の再転成が場合によって発生する)。そして、第1、第2、第3の辞書データベースDBに基づいて、基礎文章要素規定手段300により1次処理された、入力された英語単語の対訳中国語がここで再処理される。それは、品詞の再転成がここで発生するためである。

40

【0045】

50

5. 特殊文章要素翻訳手段500:基礎文章要素規定手段300及び加工文章要素規定手段400の結果、すなわち基礎文章要素あるいは加工文章要素をチェックし、この特殊文章要素翻訳手段500に内蔵されている特殊文章要素連結ポインター(pointer)と連結される識別因子が各単語の文章要素に存在するかどうかを判断する。一つ以上の原言語の文章要素あるいは文章構文から識別因子が検出された場合、一つ以上の原言語の文章要素あるいは文章構文は、500に内蔵された各特殊文章要素翻訳モジュールにより、辞書データベースベースに定型化された対訳中国語の文章要素あるいは文章構文に再処理され、これによって300及び400で処理された対訳中国語も再処理されながら中国語の語順も再処理される。特殊文章要素翻訳手段500に内蔵された各特殊文章要素翻訳モジュールには5種類のモジュール(501~505)があり、それぞれ独特の文法処理能力を有する連結ポインター(pointer)及び各文章要素に付与された識別因子によって活性化される。 10

【0046】

6. 一般文章要素翻訳手段600:特殊文章要素翻訳手段500の結果を受信し、特殊文章要素連結ポインターにより再処理された単語を、辞書DB100に記録された中国語の語順規則に従って再構成する必要があるかどうかを判断し、必要に応じて再処理し、残りの原言語の文章要素あるいは文章構文を完成された文章要素に翻訳する一般文章要素翻訳手段600である。さらに、この段階において、中国語文法の独特の規則に従って、(1)最終の中国語の対訳文章で不要な中国語単語を除去し、(2)最終の中国語の対訳文章で必ず必要な中国語単語(対訳語と関係なく)を挿入する過程が含まれる。 20

【0047】

7. 完成翻訳文出力手段700:完成された対訳中国語文章を出力する。

【0048】

特殊文章要素翻訳手段500に内蔵された5種類の特殊翻訳モジュール(501~505)には、それぞれ互いに異なる処理能力を有する特殊文章要素連結ポインターが含まれている。

【0049】

第1特殊翻訳モジュール501は、原語文章中にある動詞の中国語対訳動詞に特殊文章要素連結ポインターが存在するかどうか、さらに、▲離▼合▲詞▼識別因子を有するかどうかをチェックする。その二つが同時に存在する場合には、原語文章が過去時制であるかを調べ、過去時制であれば対訳語の文章要素として動詞を翻訳するとき、その翻訳された対訳語の動詞の音節の間に過去時制を表す“了”という文字を挿入して翻訳し、次いで、原語文章中に回数の副詞が存在するかどうかを調べ、それに対応する対訳語を過去表示の“了”の後に位置させて(原語文章が過去時制でない場合には回数の副詞のみ位置する)翻訳する。 30

【0050】

第2特殊翻訳モジュール502は、原言語の文章要素中に中国語対訳動詞の程度を表す中国語の補語性格の単語が含まれているかどうかを特殊文章要素連結ポインターによってチェックし、その中国語の程度補語が中国語対訳動詞を修飾するという識別因子が存在する場合には、対訳語の文章要素である動詞と中国語の補語性格の単語を翻訳するとき、翻訳された対訳語の動詞とその対訳語の程度補語の間に“得”という文字を挿入して翻訳する。 40

【0051】

第3特殊翻訳モジュール503は、原言語の文章要素中に中国語対訳動詞の結果を表す中国語の補語性格の単語が含まれているかどうかを特殊文章要素連結ポインターによってチェックし、その対訳された中国語の結果補語が中国語対訳動詞を修飾するという識別因子が存在する場合には、対訳語の文章要素としてその単語を翻訳するとき、翻訳された対訳語の動詞の後にその中国語の対訳語の結果補語を置いて翻訳する。

【0052】

第4特殊翻訳モジュール504は、原言語の文章要素中に動詞が「be動詞+過去分詞 50

」からなる被動文(受動態)であるかどうか、及び原語文章の過去分詞と対応する中国語対訳動詞がヒトや物の動作を支配するかどうか、また、影響を与えるかどうかを特殊文章要素連結ポインター及び識別因子によってチェックする。その結果、全て該当する場合には、被動文の原語文章を対訳語に翻訳するとき、その翻訳された対訳語文章の対訳動詞の前に“被”という文字を挿入するかを決定し、語順の再構成要求の有無に従って翻訳する。

【0053】

第5特殊翻訳モジュール505は、原言語の文章要素を構成する構文が「主語＋be動詞＋形容詞＋前置詞」からなる形容詞構文であるかどうかを特殊文章要素連結ポインター及び識別因子によってチェックする。その結果、形容詞構文である場合には、対訳語の文章要素を構成する中国語の形容詞構文に翻訳するときに、be動詞に対応する対訳語の動詞を省略し、語順の再構成要求に従って翻訳する。 10

【0054】

辞書データベースに内蔵された4つの辞書のすべての英語単語と中国語単語には、それぞれ形態素情報、基礎文章要素情報、加工文章要素情報、特殊文章要素情報が含まれており、特に、特殊文章要素情報には識別因子検索表示が第1、第2、第3の辞書DBに付与されている。識別因子は、各英語単語の特殊文章要素情報をチェックし、5種類の特殊翻訳モジュール(501～505)中の特殊文章要素連結ポインターと連係して機能し、特殊文章要素情報を有する各英語単語を最も適切かつ正確に中国語対訳の単語と対応させて対訳結果を生成するようにする。

【0055】

5種類の特殊翻訳モジュール(501～505)は、入力された英語文章を当該単語情報の再処理及び再加工過程に誘導し、最も適切かつ正確な最終の中国語対訳文章に変換、生成されるようにする重要な役割をする。 20

【0056】

第1辞書DBは一般単語情報を有し、第2、第3辞書DBは合成語、慣用語、熟語などの情報を、第4辞書DBは専門用語情報を有する。

【0057】

次に、入力された英語文章が中国語の対訳文章に翻訳生成される過程を説明する。

【0058】

1. 英語文章の動詞を中国語の動詞に対訳する際に、該動詞が▲離▼合▲詞▼構文の動詞に対訳されるとき、特殊文章要素連結ポインター及び識別因子は、英語文章が過去時制であるかどうかをチェックし、同時に英語文章に回数の副詞が存在するかどうかをチェックする。 30

【0059】

2. 中国語対訳文章中に中国語補語(英語文章では主に副詞)が程度補語である場合には、特殊文章要素連結ポインター及び識別因子は、その中国語補語が対訳された中国語動詞を修飾するかどうかをチェックする。

【0060】

3. 中国語対訳文章中に中国語補語(英語文章では主に副詞)が結果補語である場合には、特殊文章要素連結ポインター及び識別因子は、その中国語補語が対訳された中国語動詞を修飾するかどうかをチェックする。 40

【0061】

4. 入力された英語文章が被動文(be + p. p.)である場合には、特殊文章要素連結ポインター及び識別因子は、対応する中国語対訳動詞がヒトや物の動作を支配するかどうか、また、影響を与えるかどうかを判断する。次いで、英語動詞の後に位置する英語前置詞に対応する中国語の対訳単語が中国語文法に符合するかを判断する。何故ならば、英語前置詞に対応する中国語の対訳単語は、中国語文法上の特別な対訳語条件を満たす必要があるためである。

【0062】

5. 入力された英語文章が形容詞構文である場合には、特殊文章要素連結ポインター及び 50

識別因子がそれをチェックして、入力された英語文章の be 動詞に対応する中国語の対訳単語を除去し、中国語語順規則に従って語順を再配列する。

【0063】

本発明の自動翻訳方法は、次のように構成する。

【0064】

ステップ S100：原語文章を入力する段階。

【0065】

ステップ S200：入力された原語文章の単語を形態素単位に分析する段階。

【0066】

ステップ S300：形態素単位に分析された単語を各品詞として分析し、基礎文章要素として解釈する段階。 10

【0067】

ステップ S400：S300で分析された結果を受信し、処理した各品詞を組み合わせて、少なくとも二つ以上の品詞が合成語、熟語、または定型化された慣用語句などに変換できるものがあるかどうかをチェックする。変換できるものが検出される場合には、少なくとも二つ以上の品詞を結合して構文的機能を考慮した品詞に転換し、転換された転成品詞を含む加工文章要素として規定する段階。

【0068】

ステップ S500：基礎文章要素あるいは加工文章要素をチェックし、ステップ S500段階に内蔵されている特殊文章要素の連結ポインターと連結される識別因子が各単語の文章要素に存在するかどうかを判断する。一つ以上の原言語の文章要素あるいは文章構文から識別因子が検出される場合には、一つ以上の原言語の文章要素あるいは文章構文は、S500に内蔵された各特殊文章要素の翻訳モジュール501～505により辞書データベースに定型化された対訳中国語の文章要素あるいは文章構文に再処理される。これによって、S300及びS400で処理された対訳中国語も再処理されながら中国語の語順が再処理される段階。 20

【0069】

ステップ S600：S500の結果を受信し、特殊文章要素連結ポインターによって再処理された単語を、辞書DBに記録された中国語の語順規則に従って再構成する必要があるかどうかを判断し、必要に応じて再処理し、残りの原言語の文章要素あるいは文章構文を完成された文章要素に翻訳する段階である。さらに、この段階において、中国語文法の独特の規則に従って、(1)最終の中国語の対訳文章で不要な中国語単語を除去し、(2)最終の中国語の対訳文章で必ず必要な中国語単語(対訳語と関係なく)を挿入する過程が含まれる。 30

【0070】

ステップ S700：完成された対訳文章を出力する段階である。

【0071】

なお、ステップ S500において、500に内蔵された(第1特殊翻訳モジュール501)特殊文章要素連結ポインター及び識別因子は、原言語の文章要素に対応する中国語対訳動詞が▲離▼合▲詞▼であるかどうかを判断し、▲離▼合▲詞▼である場合には、対訳語の文章要素としてその動詞を翻訳するとき、その翻訳された対訳語の動詞の音節の間に過去時制を表す“了”という文字を挿入(原語文章が過去型である場合)したり、原語文章に回数の副詞が存在する場合には、“了”という文字の後に対訳された中国語の回数の副詞の対訳語を挿入(原語文章が過去型ではない場合には“了”は挿入せず、副詞の対訳語のみ挿入する)して翻訳する段階(S501)と、 40

ステップ S500において、対訳された中国語文章の単語のうち、中国語程度補語が存在するかどうか、及びその程度補語が対訳された中国語文章の対訳動詞を修飾するかどうかを500に内蔵された(第2特殊翻訳モジュール502)特殊文章要素連結ポインター及び識別因子が判断する。その結果、中国語対訳動詞を修飾する場合には、中国語対訳動詞と対訳された中国語程度補語の間に“得”という文字を挿入し、その後、該当する対訳 50

語の程度補語を配置して翻訳する段階(S 5 0 2)と、

ステップ S 5 0 0 において、対訳された中国語文章の単語のうち、中国語結果補語が存在するかどうか、及びその結果補語が対訳された中国語文章の対訳動詞を修飾するかどうかを 5 0 0 に内蔵された(第 3 特殊翻訳モジュール 5 0 3)特殊文章要素連結ポインター及び識別因子が判断する。その結果、中国語対訳動詞を修飾する場合には、中国語対訳動詞のすぐ後に該当する対訳語の結果補語を置いて翻訳する段階(S 5 0 3)と、

ステップ S 5 0 0 において、第 4 特殊翻訳モジュール 5 0 4 の特殊文章要素連結ポインター及び識別因子が原語文章の文章構造を「b e 動詞＋過去分詞」からなる被動文(受動態)であるかどうかを判断し、被動文であると判断した場合には、原語文章の過去分詞と対応する中国語対訳動詞がヒトや物の動作を支配するかどうか、また、影響を与えるかどうかを判断する。その結果、全て該当する場合には、被動文の原語文章を対訳語に翻訳するとき、その翻訳された対訳語文章の対訳動詞の前に“被”という文字を挿入するかを決定し、その後、語順の再構成要求の有無に従って翻訳する段階(S 5 0 4)と、

ステップ S 5 0 0 において、第 5 特殊翻訳モジュール 5 0 5 の特殊文章要素連結ポインター及び識別因子が原語文章の文章構造を「b e 動詞＋形容詞＋前置詞」からなる形容詞構文であるかどうかを判断し、形容詞構文であると判断した場合には、対訳語の文章要素を構成する中国語の形容詞構文に翻訳するときに、b e 動詞に対応する中国語対訳動詞を省略し、語順の再構成要求に従って翻訳する段階(S 5 0 5)を含む。

【0 0 7 2】

ステップ S 5 0 0 において、特殊翻訳モジュール 5 0 1 ～ 5 0 5 及び内蔵された特殊文章要素連結ポインター、並びに単語に付与された識別因子が三位一体となって動作する。

【発明を実施するための最良の形態】

【0 0 7 3】

以下、本発明の英語と中国語間の自動翻訳装置の実際翻訳過程について例を挙げて説明する。

【0 0 7 4】

ユーザがインタフェース(例えば、キーボード)を利用して“ I h a v e n o t c h a t t e d w i t h h i m o n c e . ” (図 5 参照)という原語文章を入力すると、入力された原語文章は翻訳要請画面の原語文章入力部に表示される(S 1 0 0)。このように、入力した後、翻訳要請画面に備えられた翻訳要請メニューを選択すれば自動翻訳装置が翻訳を開始する。本発明の自動翻訳装置による翻訳過程について説明する。

【0 0 7 5】

まず、入力された原語文章“ I h a v e n o t c h a t t e d w i t h h i m o n c e . ” は、形態素分析手段 2 0 0 により英語文章を構成する各単語が形態素単位に分析(S 2 0 0)される。形態素単位に分析された英語文章の各単語は、基礎文章要素規定手段 3 0 0 により順次にデータ保存手段 1 0 0 の第 1 辞書 D B に連結され、代名詞、助動詞、否定副詞、動詞(過去時制)、前置詞、名詞、副詞などの各品詞で表示される基礎文章要素として規定される(S 3 0 0)。

【0 0 7 6】

加工文章要素規定手段 4 0 0 は、3 0 0 による単語情報データを検索して、合成語、熟語、あるいは定型化された慣用語句などに変換できるものが存在するかどうかを判断し、変換できるものが検出される場合には、少なくとも二つ以上の該当する品詞を辞書 D B に連結して構文的機能を考慮した品詞に転換し、変換された転成品詞として新たに規定する。同時に、この過程で品詞が再定義される品詞の再転成が必要に応じて発生する。入力された原文章の主語、動詞句(時制を含む)、名詞句、前置詞句、副詞句等の単語関連データが再定義される(S 4 0 0)。

【0 0 7 7】

次に、特殊文章要素翻訳手段 5 0 0 に入った再定義された主語、動詞句(時制を含む)、名詞句、前置詞句、副詞句などの情報は、特殊文章要素連結ポインター及び特定単語の識別因子情報の検索、並びに特定文章要素の語順配置の検索によってそれぞれの 5 種類の該

当する特殊翻訳モジュール（５０１～５０５）に通報される。この過程により、入力された原英語文章が有する特殊文章要素の機能的意味が出現し、これによって原英語文章は最も適切かつ正確な中国語の対訳文章に生成、翻訳される。

【００７８】

入力された原英語文章“ I have not chatted with him once . ”は、辞書DBと２００、３００、４００段階を経て分析された情報が特殊文章要素翻訳手段５００に送信される。

【００７９】

２００、３００、４００段階により原英語文章の“ I ”は中国語“ 我 ”に、“ have ”及び“ not ”は“ 有 ”及び“ ▲没▼ ”に、“ chatted ”は“ ▲間談▼ ”に、“ with ”は“ 跟 ”に、“ him ”は“ 他 ”に、そして“ once ”は“ 一次 ”に対訳される。 10

【００８０】

５００において、特殊文章要素連結ポインターは原英語文章を検索して、動詞“ chatted ”に中国語▲離▼合▲詞▼動詞の識別因子が付与されており、過去時制であるとの情報を取得する。特殊文章要素連結ポインター及び識別因子は、この単語が第１特殊翻訳モジュール５０１で処理されるべきであると判断し、その処理を渡す。第１特殊翻訳モジュール５０１は、辞書DBを検索して、原英語文章中の動詞“ chatted ”に対応する最も正確な中国語の対訳語が“ ▲間談▼ ”ではなく、▲離▼合▲詞▼動詞の“ 聊天儿 ”であると判断する。 20

【００８１】

次いで、特殊文章要素連結ポインター及び識別因子は、“ 聊天儿 ”の音節の間に過去時制を表す中国語の対訳語“ 了 ”を挿入する。その結果、中国語の対訳語の動詞は“ 聊了天儿 ”になる。

【００８２】

その後、特殊文章要素連結ポインター及び識別因子は、原英語文章を検索して、回数の副詞“ once ”を認識、これを第１翻訳モジュール５０１に通知する。通知を受けた第１特殊翻訳モジュール５０１は、辞書DBを参照して“ once ”に対応する中国語“ 一次 ”を抽出し、第１特殊翻訳モジュール５０１、特殊文章要素連結ポインター及び識別因子は、抽出された“ 一次 ”を翻訳された▲離▼合▲詞▼動詞である“ 聊了天儿 ”に挿入して“ 聊了一次天儿 ”を生成する。 30

【００８３】

これのように、入力された原英語文章中の“ chatted ”と“ once ”に対する中国語対訳は正確かつ自然に処理することができた。

【００８４】

本自動翻訳装置の第１特殊翻訳モジュール５０１及び連結ポインター並びに識別因子は、中国語対訳動詞が▲離▼合▲詞▼に属するかどうか、さらに、原英語文章に回数の副詞が存在するかどうかを判断できるようにプログラムされている。

【００８５】

特殊文章要素連結ポインターは、識別因子の付与された品詞が原英語文章にこれ以上存在しないし、また、与えられた語順配置要求も無いと判断する場合、第１特殊翻訳モジュール５０１は特殊文章要素の翻訳を完了する（Ｓ５００）。 40

【００８６】

前記過程により、原英語文章“ I have not chatted with him once . ”は“ 我有▲没▼聊了一次天儿跟他 . ”に対訳される。

【００８７】

その後、一般文章要素翻訳手段６００は、直前の段階であるＳ５００の結果を受信し、特殊文章要素連結ポインターによって再処理された単語を辞書DBに記録された中国語語文法の独特の規則に沿って、１）最終の中国語の対訳文章で不要な中国語単語を除去し、２）最終の中国語の対訳文章で必ず必要な中国語単語（対訳語と関係なく）を挿入する過程 50

を行なう。さらに、辞書DBに記録された中国語の語順規則に従って語順を再構成すべきかどうかを判断し、必要に応じて再処理し、残りの原言語の文章要素あるいは文章構文を完成された文章要素に生成、翻訳する。

【0088】

ステップS600において、入力された英語単語“have”の対訳語“有”が最終の中国語の翻訳文章から除去され、英語単語“with him”の対訳語“跟他”は、中国語の語順規則に従って“▲没▼”の文字のすぐ後に移動配置される。

【0089】

一般文章要素翻訳手段600により、原英語文章“I have not chatted with him once.”は“我▲没▼跟他聊了一次天儿。”に最終完成された翻訳文として生成される(S600)。

10

【0090】

そして、完成翻訳文出力手段700がこれを出し、本自動翻訳装置の画面に“我▲没▼跟他聊了一次天儿。”を表示する。

【0091】

このように、本自動翻訳装置を利用して正確かつ自然な対訳文を得ることができた。中国語の▲離▼合▲詞▼構文をここまで正確かつ自然に翻訳する翻訳技術(Technological Algorithm)はこれまでに存在しなかった。

【0092】

原英語文章の入力から加工文章要素への規定までの過程S100～S400は、上述した過程と同様の流れにより行なわれる。

20

【0093】

次に、特殊文章要素翻訳手段500(特殊翻訳モジュール)が原英語文章の特殊文章要素をどのように扱い処理するか、そして、5種類の特殊翻訳モジュールがどのように作動するかを説明する(図5～図9参照)。

【0094】

原英語文章“He always answers correctly.”を入力する(図6参照)。この文章も辞書DBと200、300、400段階を経ることによって分析、処理、再処理、再定義され、中国語の対訳単語を生成する。

【0095】

入力された原英語文章中の“He”は“他”に、“always”は“▲忠▼是”に、“answers”は“回答”に、“correctly”は“正▲確▼地”に対訳される。

30

【0096】

次に、400の情報が500に伝達されると、特殊文章要素連結ポインターが入力された原文章の英語単語を検索しながら中国語の対訳文章を形成し、“He always answers correctly.”は“他▲忠▼是回答正▲確▼地。”に翻訳生成される。

【0097】

この過程において、すべての英語単語は特殊文章要素連結ポインターにより識別因子が付与されているかどうかチェックされ、“correctly”という単語に識別因子が付与されていると判断する。特殊文章要素連結ポインターは、“correctly”の識別因子と連係して辞書DBに記録された“correctly”関連情報を検索する。

40

【0098】

特殊文章要素連結ポインター及び識別因子は、“correctly”という単語を中国語の程度補語として処理すべきであるという情報を取得する。

【0099】

特殊文章要素連結ポインター及び“correctly”の識別因子は、第2特殊翻訳モジュールで原英語文章が処理されるべきであると判断し、200、300、400過程で導出された中国語の対訳語である“正▲確▼地”が中国語文法に符合しない対訳語であ

50

ると判断する。第2特殊翻訳モジュール502の特殊文章要素連結ポインター及び“correctly”の識別因子は辞書DBを参照して、原英語文章中の“correctly”に対応する最も正確で自然な中国語の対訳語が“正▲確▼地”ではなく“▲対▼”であると判断する。

【0100】

第2特殊翻訳モジュール及び特殊文章要素連結ポインター、“correctly”の識別因子は、2次加工対訳語である“▲対▼”が中国語対訳動詞である“回答”(“answers”の対訳語)を修飾するかどうかを判断する。その結果、修飾していると判断され、中国語動詞を修飾する程度補語であると判明されれば、特殊文章要素連結ポインター及び“correctly”の識別因子は、特殊中国語挿入文字“得”を中国語対訳動詞“回答”の後に置き、2次加工対訳語である“▲対▼”を“得”の後に位置させる。中国語の程度補語の構文規則に従って処理したものである。

【0101】

次に、特殊文章要素翻訳手段500は、入力された原英語文章“He always answers correctly.”に対する中国語の対訳文章が、中国語の語順規則に沿っているかどうかを検査した後、“他▲忠▼是回答得▲対▼.”という文章を提示する。

【0102】

一般文章要素翻訳手段600は、S500の結果を受信してチェックし、その結果に対して再処理すべき内容がないことを判断する。

【0103】

その後、一般文章要素翻訳手段600は、中国語の対訳文章“他▲忠▼是回答得▲対▼.”を提示する。

【0104】

最後に、完成翻訳文出力手段700がそれを出し、本自動翻訳装置の画面に“他▲忠▼是回答得▲対▼.”を表示する。

【0105】

このように、本自動翻訳装置を利用してより正確かつ自然な対訳文を得ることができた。中国語の程度補語構文をここまで正確かつ自然に翻訳、生成した技術はこれまでに存在しなかった。

【0106】

次に、原英語文章“I translated it into English wrongly.”を入力する(図7参照)。この文章も辞書DBと200、300、400過程を経ることにより分析、処理、再処理、再定義され、中国語の対訳単語を生成する。

【0107】

入力された原英語文章において“I”は“我”に、“translated”は“翻▲訳▼”に、“it”は“那个”に、そして“into”は“成”、“English”は“英▲語▼”、“wrongly”は“▲錯誤▼地”に対訳される。

【0108】

次いで、400の情報が500に伝達されれば、特殊文章要素連結ポインターが原文章の英語単語を検索しながら中国語の対訳文章を形成し、“I translated it into English wrongly.”は“我翻▲訳▼那个成英▲語錯誤▼地.”に翻訳生成される。

【0109】

この過程において、すべての英語単語は、特殊文章要素連結ポインターにより識別因子が付与されているかどうかチェックされ、“wrongly”という単語に識別因子が付与されているという情報を得る。特殊文章要素連結ポインターは“wrongly”の識別因子と連係して辞書DBに記録された“wrongly”関連情報を検索する。

【0110】

特殊文章要素連結ポインター及び識別因子は“wrongly”を中国語語法の結果補

10

20

30

40

50

語として処理すべきであると判断する。

【0111】

特殊文章要素連結ポインター及び“w r o n g l y”の識別因子は、第3特殊翻訳モジュールで原英語文章が処理されるべきであると判断し、200、300、400過程で導出された中国語の対訳語である“▲錯誤▼地”が中国語文法に符合しないと判断する。第3特殊翻訳モジュール503の特殊文章要素連結ポインター及び“w r o n g l y”の識別因子は、辞書DBを参照して、原英語文章中の“w r o n g l y”に対応する最も正確で自然な中国語の対訳語が“▲錯誤▼地”ではなく“▲錯▼了”であると判断する。

【0112】

第3特殊翻訳モジュール及び特殊文章要素連結ポインター、“w r o n g l y”の識別因子は、2次加工対訳語である“▲錯▼了”が中国語対訳動詞である“翻▲訳▼”(“t r a n s l a t e d”の対訳語)を修飾するかどうかを判断する。その結果、修飾すると判断され、中国語動詞を修飾する結果補語であると判明されれば、特殊文章要素連結ポインター及び“w r o n g l y”の識別因子は2次加工対訳語である“▲錯▼了”を中国語対訳動詞である“翻▲訳▼”のすぐ後に配置する。中国語の結果補語構文規則と中国語の語順規則に従って処理したものである。

【0113】

その後、一般文章要素翻訳手段600は、直前の段階であるS500の結果を受信し、特殊文章要素連結ポインターにより再処理された単語を、辞書DBに記録された中国語文法の独特の規則に従って、1)最終の中国語の対訳文章で不要な中国語単語を除去し、2)最終の中国語の対訳文章で必ず必要な中国語の単語(対訳語と関係なく)を挿入する過程を行なう。また、辞書DBに記録された中国語の語順規則に従って語順を再構成する必要があるかどうか判断し、必要に応じて再処理し、残りの原言語の文章要素あるいは文章構文を完成された文章要素に生成、翻訳する。

【0114】

S500の結果を受信した一般文章要素翻訳手段600は、中国語の対訳語“那个”(“i t”の中国語の対訳語)のすぐ前に中国語の挿入単語“把”(中国語文法の独特の規則に従って)が位置する必要があると判断する。そして、中国語の対訳語“把那个”は、中国語の語順規則に従って中国語対訳動詞である“翻▲訳▼”のすぐ前に位置する必要があると判断する。このようにして、一般文章要素翻訳手段600は、最終の中国語の対訳文章“我把那个翻▲訳錯▼了成英▲語▼.”を生成する(S600)。

【0115】

最後に、完成翻訳文出力手段700がそれを出し、本自動翻訳装置の画面に“我把那个翻▲訳錯▼了成英▲語▼.”(S600)を表示する。

【0116】

このように、本自動翻訳装置を利用して正確かつ自然な対訳文を得ることができた。中国語の結果補語構文をここまで正確かつ自然に翻訳、生成した翻訳技術(T e c h n o l o g i c a l A l g o r i t h m)はこれまでに存在しなかった。

【0117】

次に、原英語文章、“H e i s b e s i e g e d w i t h v i s i t o r s f r o m a b r o a d.”を入力する(図8参照)。この文章も辞書DBと200、300、400過程を経ることによって分析、処理、再処理、再定義され、中国語の対訳単語を生成する。

【0118】

入力された原英語文章において“H e”は“他”に、“i s”は“是”に、“b e s i e g e d”は“包▲囲▼了”に、“w i t h”は“跟”に、“v i s i t o r s”は“客人”に、“f r o m”は“▲從▼”に、“a b r o a d”は“外国”に対訳される。

【0119】

次いで、400の情報が500に伝達されれば、特殊文章要素連結ポインターが原文の英語単語を検索しながら中国語の対訳文章を形成する。即ち、“H e i s b e s i e

ged with visitors from a broad.”は“他是包▲囲▼了跟客人▲從▼外国.”に翻訳、生成される。

【0120】

この過程で入力された全ての英語単語は、特殊文章要素連結ポインターにより識別因子が付与されているかどうかチェックされ、原英語文章の文章構造が「be動詞＋過去分詞」からなる被動文(受動態)であると判断するように被動文識別因子を有することが判る。特殊文章要素連結ポインター及び識別因子は、辞書DBを参照して原英語文章の文章構造と単語を検索する。

【0121】

特殊文章要素連結ポインター及び文章構造識別因子は、第4特殊翻訳モジュール504で原英語文章が処理されるべきであると判断する。第4特殊翻訳モジュール504及び特殊文章要素連結ポインター並びに文章構造識別因子は、辞書DBを参照して原文中“is besieged”という動詞句を検索すると同時に、中国語の対訳語の動詞“包▲囲▼了”がヒトや物の動作を支配するかどうか、また、影響を与えるかどうかを判断する。その結果、Yesと判断される場合には、翻訳された中国語対訳動詞の前に“被”という文字を挿入することを決定する。第4特殊翻訳モジュール504は、一次的対訳動詞である“包▲囲▼了”を“被包▲囲▼了”に再加工して処理する。 10

【0122】

第4特殊翻訳モジュール504の特殊文章要素連結ポインター及び識別因子はまた、入力された被動文の英語動詞の後に位置する英語前置詞に対応する中国語の対訳単語が中国語文法に符合するかを調べる。被動文の英語前置詞に対応する中国語の対訳単語は、中国語文法において確固に存在する特別な英語前置詞の対訳語条件を充足しなければならないためである。 20

【0123】

以上のような条件を調べるために、特殊文章要素連結ポインター及び識別因子は、被動文の英語前置詞“with”の中国語の対訳語である“跟”が、被動文の英語前置詞に対応する中国語の対訳語条件を満たしているかどうかを調べる。その結果、NOである場合には、辞書DBを再検索し、入力された被動文の英語文章の前置詞に対応する中国語の対訳語条件(中国語文法)に符合する対訳語が“跟”ではなく“与”であるという情報を取得する。 30

【0124】

次いで、第4特殊翻訳モジュール504は辞書DBを参照して、中国語の語順規則に従って原文の前置詞句“with visitors”を結合した中国語の対訳語句“与客人”を、中国語対訳動詞句である“被包▲囲▼了”のすぐ前に配置する。第4特殊翻訳モジュール504は、原文の中国語の対訳文章を“他是与客人被包▲囲▼了▲從▼外国.”に再変換して提示する。

【0125】

その後、一般文章要素翻訳手段600は、直前の段階であるS500の結果を受信し、特殊文章要素連結ポインターにより再処理された単語を辞書DBに記録された中国語文法の独特の規則に従って、最終の中国語対訳文章で不要な中国語単語“是”(動詞beの対訳語)を除去し、辞書DBに記録された中国語の語順規則に従って語順を再構成すべきかどうかを判断する。その結果、前置詞句である“▲從▼外国”(“from abroad”の対訳語)を“与客人”(“with visitors”の対訳語)のすぐ後に位置させる。このようにして、600は最終の中国語の対訳文章として“他与客人▲從▼外国被包▲囲▼了.”を生成する(S600)。 40

【0126】

最後に、完成翻訳文出力手段700がそれを出力し、本自動翻訳装置の画面に“他与客人▲從▼外国被包▲囲▼了.”を表示する。

【0127】

このように、本自動翻訳装置を利用して正確かつ自然な対訳文を得ることができた。中 50

国語の被動文構文をここまで正確かつ自然に翻訳、生成した技術はこれまでに存在しなかった。

【0128】

次に、原英語文章“H e i s s e n s i t i v e t o l i g h t . ”を入力する(図9参照)。この文章も辞書DB及び200、300、400の過程によって分析、処理、再処理、再定義され、中国語の対訳単語を生成する。

【0129】

入力された原英語文章において、“H e ”は“他”に、“i s ”は“是”に、“s e n s i t i v e ”は“敏感”に、“t o ”は“▲対▼”に、“l i g h t ”は“光”に対訳される。

10

【0130】

次いで、400の情報が500に伝達されれば、特殊文章要素連結ポインターが原入力文章の英語単語を検索しながら、中国語の対訳文章を形成し、“H e i s s e n s i t i v e t o l i g h t . ”は“他是敏感▲対▼光.”に翻訳、生成される。

【0131】

この過程において、入力された全ての英語単語は、特殊文章要素連結ポインターにより識別因子が付与されているかが判断され、特殊文章要素連結ポインターは、識別因子の付与された単語を検索されれば、識別因子を活性化させる。

【0132】

特殊文章要素連結ポインターは、原英語文章中“i s s e n s i t i v e ”の文章構造から識別因子を検出し、「b e + 形容詞」からなる形容詞構文であると判断する。そして、特殊文章要素連結ポインター及び識別因子は、原英語文章が第5特殊翻訳モジュール505で処理されるべきであると判断する。

20

【0133】

第5特殊翻訳モジュール505及び特殊文章要素連結ポインター並びに文章構造識別因子は、辞書DBを参照して原文章中の“i s s e n s i t i v e ”という形容詞句に関する情報を検索し、原英語文章中の“i s ”に対応する中国語対訳動詞“是”を除去し、辞書DBに記録された中国語の語順規則に従って中国語の対訳文章を再構成する。

【0134】

このようにして、入力された原英語文章“H e i s s e n s i t i v e t o l i g h t . ”は“他敏感▲対▼光.”に対訳、生成される(S500)。

30

【0135】

その後、一般文章要素翻訳手段600は、直前の段階であるS500の結果を受信し、特殊文章要素連結ポインターにより再処理された単語を、辞書DBに記録された中国語文法の独特の規則に従って、最終の中国語の対訳文章に不要な中国語単語及び必ず必要な中国語単語を調べ、辞書DBに記録された中国語の語順規則に従って語順を再調整する必要があるかどうかを判断する。その結果、前置詞句である“▲対▼光”(“t o l i g h t ”の対訳語)が中国語の形容詞対訳語“敏感”のすぐ前に位置する必要があると判断する。正しい中国語の語順規則に従うためである。これにより、600は、原英語文章の最終の中国語の対訳文章を“他▲対▼光敏感.”のように再構成して生成する(S600)

40

【0136】

その結果、完成翻訳文出力手段700がそれを出し、本自動翻訳装置の画面に“他▲対▼光敏感.”を表示する。

【0137】

このように、本自動翻訳装置により正確かつ自然な対訳文を得ることができた。中国語の被動文構文をここまで正確かつ自然に翻訳、生成した技術(T e c h n o l o g i c a l A l g o r i t h m)はこれまでに存在しなかった。

【0138】

以上のように、当該発明の属する技術分野における通常の知識を有する者であれば、本

50

発明の思想及び範囲内で様々な変形及び均等な他の実施例が可能であり、そのような多様な変形及び改良形態も本発明の権利範囲に属するものである。

【図面の簡単な説明】

【0139】

【図1】 本発明の自動翻訳装置を示すシステム構成図。

【図2】 図1に示した自動翻訳装置の構成要素である特殊文章要素翻訳手段の詳細なブロック図。

【図3】 本発明の自動翻訳装置により翻訳される過程を示したフローチャート。

【図4】 特殊文章要素連結ポインターにより一つ以上の原言語の文章要素あるいは文章構文を検索した場合、その検索された一つ以上の原言語の文章要素あるいは文章構文を各翻訳モジュールによって辞書に定型化された対訳語の文章要素あるいは文章構文に翻訳する段階を詳細に示したフローチャート。

【図5】 図5乃至図9は、本発明の開発者が開発した理論(本発明の特許出願書の内容)に基づいて作製して成功した試製品(Demo Program: 2003.3)を用いて翻訳、生成された実際翻訳文の例であり、図5は、中国語対訳動詞のうち、▲離▼合▲詞▼構文動詞の翻訳過程を示す図。

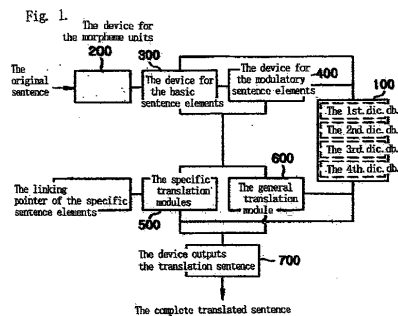
【図6】 本発明の試製品による翻訳例であり、程度補語の翻訳過程を示す図。

【図7】 本発明の試製品による翻訳例であり、結果補語の翻訳過程を示す図。

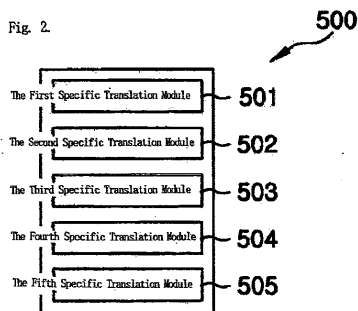
【図8】 本発明の試製品による翻訳例であり、被動文の翻訳過程を示す図。

【図9】 本発明の試製品による翻訳例であり、形容詞構文の翻訳過程を示す図。

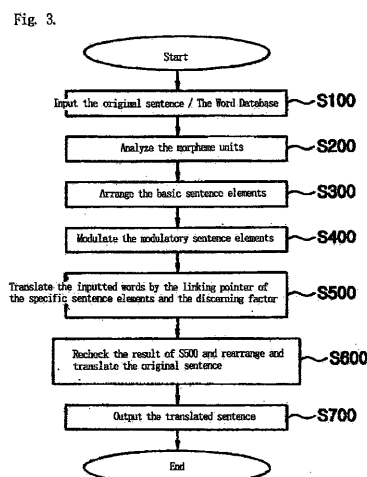
【図1】



【図2】

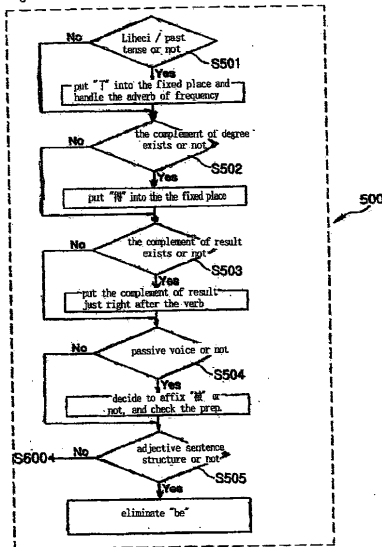


【図3】



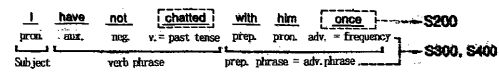
【图 4】

Fig. 4.

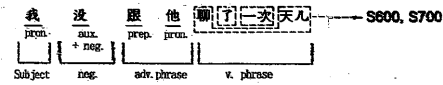


【图 5】

Fig. 5.

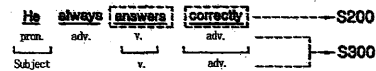


I have not 聊[了]一次天儿 with him → S501



【图 6】

Fig. 6.

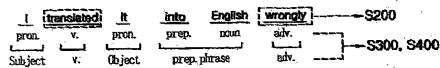


He always 回答[得]对 → S502

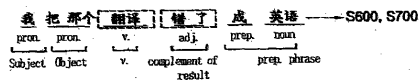


【图 7】

Fig. 7.

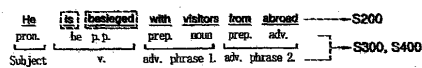


I 翻译[了] it into English → S503
 v. complement of result

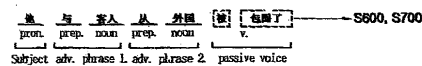


【图 8】

Fig. 8.

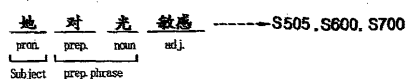
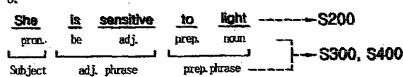


He 被[了] with visitors from abroad → S504



【图 9】

Fig. 9.



【国際調査報告】

INTERNATIONAL SEARCH REPORT

International application No.
PCT/KR2005/000065

A. CLASSIFICATION OF SUBJECT MATTER

IPC7 G06F 17/28

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC7 G06F17/00

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
Korean patents and applications for inventions since 1975

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	KR 2001-57775 A (ETRI INSTITUTE) 5 JUL 2001 SEE THE WHOLE DOCUMENTS	1-5
Y	KR 2002-2527 A (CHO YONG HWAN ETR.) 10 JAN. 2002 SEE THE WHOLE DOCUMENTS	1-5
P.Y.	KR 2004-42231 A (ETRI INSTITUTE) 20 MAY 2004 SEE THE WHOLE DOCUMENTS	1-5

☐ Further documents are listed in the continuation of Box C.☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

13 JUNE 2005 (13.06.2005)

Date of mailing of the international search report

18 JUNE 2005 (18.06.2005)

Name and mailing address of the ISA/KR

Korean Intellectual Property Office
920 Dunsan-dong, Seo-gu, Daejeon 302-701,
Republic of Korea

Facsimile No. 82-42-472-7140

Authorized officer

JEONG, Jae Hoon

Telephone No. 82-42-481-5787



フロントページの続き

(81)指定国 AP(BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), EP(AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW

Fターム(参考) 5B091 AA07 AB11 BA03 BA11 CA24 EA24 EA25

【要約の続き】

自然かつ正確に生成することができる。

本発明は、特許出願者が2003年3月開発した試製品(本発明の核心である特殊翻訳アルゴリズムを適用)により、理論の実際適用が妥当であることが検証された。本発明の実施例(図5、6、7、8、9)は、試製品を実際に作動させて得られた翻訳文の例である。