

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第3999771号
(P3999771)

(45) 発行日 平成19年10月31日(2007.10.31)

(24) 登録日 平成19年8月17日(2007.8.17)

(51) Int. Cl.

G06F 17/28 (2006.01)

F I

G06F 17/28

U

請求項の数 9 (全 19 頁)

(21) 出願番号	特願2004-199606 (P2004-199606)	(73) 特許権者	000003078
(22) 出願日	平成16年7月6日(2004.7.6)		株式会社東芝
(65) 公開番号	特開2006-23844 (P2006-23844A)		東京都港区芝浦一丁目1番1号
(43) 公開日	平成18年1月26日(2006.1.26)	(73) 特許権者	301063496
審査請求日	平成17年3月18日(2005.3.18)		東芝ソリューション株式会社
			東京都港区芝浦一丁目1番1号
		(74) 代理人	100077849
			弁理士 須山 佐一
		(72) 発明者	佐々木 晶
			東京都港区芝浦一丁目1番1号 東芝ソリ
			ューション株式会社内
		(72) 発明者	吉村 裕美子
			東京都港区芝浦一丁目1番1号 東芝ソリ
			ューション株式会社内

最終頁に続く

(54) 【発明の名称】 翻訳支援プログラム、翻訳支援装置、翻訳支援方法

(57) 【特許請求の範囲】

【請求項1】

翻訳辞書、翻訳結果を記憶する翻訳メモリ、翻訳文及び原文に対する編集履歴を記憶する編集履歴記憶部を備えたコンピュータによって、ある言語の原文を他の言語に翻訳処理する翻訳支援プログラムにおいて、

前記コンピュータを、

前記原文を、前記翻訳辞書に基づいて機械翻訳することで翻訳文を生成する翻訳手段と

、
前記翻訳手段により前記原文と前記翻訳文を対応付けて前記翻訳メモリに保存する手段と、

前記翻訳メモリに記憶された原文および前記翻訳文のうち少なくとも一つの文にある文字列が編集された場合、その位置に、前記原文の文字列に対して編集部分を示す特殊記号を付加して特殊記号付きの原文を生成する編集手段と、

前記編集手段により生成された特殊記号付きの原文を、元の原文の編集履歴として前記編集履歴記憶部に記憶する手段

として機能させることを特徴とする翻訳支援プログラム。

【請求項2】

請求項1記載の翻訳支援プログラムにおいて、

前記コンピュータを、

前記特殊記号付きの原文を、特殊記号に従って再度翻訳して新たな翻訳文を生成する再

翻訳手段と、

前記再翻訳手段により生成された新たな翻訳文を翻訳結果の文書として前記特殊記号付きの原文に対応付けて前記翻訳メモリに記憶する手段として機能させることを特徴とする翻訳支援プログラム。

【請求項 3】

請求項 1 記載の翻訳支援プログラムにおいて、

前記特殊記号は、@と@で挟まれた数字で構成されており、前記数字は、この位置に特殊記号を挿入したことを識別するための特殊記号の ID 番号であることを特徴とする翻訳支援プログラム。

【請求項 4】

10

請求項 3 記載の翻訳支援プログラムにおいて、

前記特殊記号付きの原文のうち、同じ ID 番号が付いている原文同士は、編集前は繋がっていたことを示すことを特徴とする翻訳支援プログラム。

【請求項 5】

請求項 3 または 4 記載の翻訳支援プログラムにおいて、

前記数字は、文の文節または文の結合部が検出された際に前記編集手段により連続番号で付与されることを特徴とする翻訳支援プログラム。

【請求項 6】

ある言語の原文を、予め記憶されている翻訳辞書に基づいて機械翻訳することで他の言語の翻訳文を生成する翻訳手段と、

20

前記翻訳手段による翻訳結果を前記原文と前記翻訳文とを対応付けて保存する翻訳メモリと、

前記翻訳手段により翻訳された翻訳文および原文の少なくとも一方を編集した編集履歴が記憶される編集履歴記憶部と、

前記原文および前記翻訳文のうち少なくとも一つの文にある文字列が編集された場合、その位置に、前記原文の文字列に対して編集部分を示す特殊記号を付加して特殊記号付きの原文を生成する編集手段と、

前記編集手段により生成された特殊記号付きの原文を、元の原文の編集履歴として前記編集履歴記憶部に記憶する手段と

を具備したことを特徴とする翻訳支援装置。

30

【請求項 7】

請求項 6 記載の翻訳支援装置において、

前記特殊記号付きの原文を、特殊記号に従って再度翻訳して新たな翻訳文を生成する再翻訳手段と、

前記再翻訳手段により生成された新たな翻訳文を翻訳結果の文書として前記特殊記号付きの原文に対応付けて前記翻訳メモリに記憶する手段と

を具備したことを特徴とする翻訳支援装置。

【請求項 8】

翻訳手段、翻訳辞書、翻訳メモリ、編集手段、編集履歴記憶部を備えたコンピュータによって、ある言語の原文を他の言語に翻訳処理する翻訳支援方法において、

40

前記翻訳手段が、ある言語の原文を、予め記憶されている前期翻訳辞書に基づいて機械翻訳することで他の言語の翻訳文を生成するステップと、

前記翻訳手段が、前記原文と前記翻訳文とを対応付けて翻訳メモリに保存するステップと、

前記原文および前記翻訳文のうち少なくとも一つの文にある文字列が編集された場合、その位置に、前記原文の文字列に対して編集部分を示す特殊記号を前記編集手段が付加して特殊記号付きの原文を生成するステップと、

生成した特殊記号付きの原文を、前記編集手段が元の原文の編集履歴として編集履歴記憶部に記憶するステップと

を有することを特徴とする翻訳支援方法。

50

【請求項 9】

請求項 8 記載の翻訳支援方法において、前記コンピュータは更に再翻訳手段を備え、
前記特殊記号付きの原文を再翻訳手段が特殊記号に従って再度翻訳して新たな翻訳文を生成するステップと、

再度翻訳して生成した新たな翻訳文を前記編集手段が翻訳結果の文書として前記特殊記号付きの原文に対応付けて翻訳メモリに保存するステップと

を有することを特徴とする翻訳支援方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、例えばある言語の文章を他の言語の文章に翻訳する翻訳支援プログラム、翻訳支援装置、翻訳支援方法に関する。

【背景技術】

【0002】

ある言語（原言語）の文書を他の言語（目的言語）の文書に自動的に翻訳する機械翻訳装置がある。

【0003】

従来の機械翻訳装置としては、ある言語で書かれている文（原文）と、その文を他の言語で表した文（訳文）とを対で格納したデータベースである翻訳メモリを参照して、翻訳作業を支援する翻訳支援装置が知られている（例えば特許文献 1 参照）。

【0004】

従来の翻訳支援装置における翻訳処理の流れは以下のようにになっている。

1. 過去に翻訳済みの対訳文を翻訳メモリに予め登録しておき、翻訳対象文と類似の文が翻訳メモリ中にあれば、それを参照して翻訳に利用する。
2. 翻訳対象文と類似の文が翻訳メモリ中にない場合は、機械翻訳を実行させて翻訳文（下訳）を作成し、下訳に人手で適宜修正を加えて訳文を完成する。
3. 翻訳対象文と完成した訳文を翻訳メモリに登録して、新たな翻訳対象文書を翻訳する際に再利用する。

【特許文献 1】特開平 6 - 6 8 1 4 2 号公報

【発明の開示】

【発明が解決しようとする課題】

【0005】

従来の翻訳支援装置では、機械翻訳の翻訳結果を手で修正して翻訳メモリに登録する場合、訳文の修正（後編集）は、簡単に行えても原文の修正（前編集）は簡単には行えないという問題があった。

【0006】

これは、翻訳メモリに登録する原文がオリジナルの原文から変更されていると、別の文書の翻訳時に再利用され難くなるためである。

【0007】

例えば、ある日本語テキスト文書から、同じ内容の h t m l 文書、リッチテキスト文書を作ったものとする。そして、日本語テキスト文書を基にして英語の翻訳文書を作成し、これらを翻訳メモリに登録したものとする。その際に、翻訳メモリに登録した原文はオリジナルの原文から変更されているものとする。

【0008】

この翻訳メモリの内容を使って、h t m l 文書、リッチテキスト文書を翻訳し、英語版の h t m l 文書、リッチテキスト文書を作成するものとする。

【0009】

しかし、翻訳メモリには、オリジナルの原文とは異なる原文が登録されているため、翻訳メモリのデータとの類似率は低くなる。また、機械翻訳装置が翻訳し易いように原文を修正すると、実際には存在しない人工的な文になる場合もあり、このような場合は実際の

10

20

30

40

50

文とマッチしない可能性がさらに高くなる。

【 0 0 1 0 】

そこで、原文の修正をあえて行う場合には、修正前の原文を別途保存しておき、翻訳メモリ登録時に修正後の原文と置き換える必要がある。

【 0 0 1 1 】

しかし、この場合にも、特に原文の分割・結合を行うと、原文との置き換え作業は非常に複雑になるという問題がある。その上、原文の分割・結合が必要になるケースはまれではない。例えば原文である日本語の1文が「が、」でつながる複数の文からなり、文を分割した方が翻訳結果が良くなる場合や、htmlファイルなどの書式情報付き文書においてレイアウトの都合で1文が分割されており、正しく翻訳するには文を結合する必要がある場合など、枚挙にいとまがない。

10

【 0 0 1 2 】

このように、翻訳メモリに登録することを考えると、原文の修正は簡単にはできないため、現実的には機械翻訳結果の修正は、訳文の修正に頼る場合が大半であった。

【 0 0 1 3 】

最初から自分で英文を書き起こせるユーザは、機械翻訳の機能を十分に活用せずに、訳文の上書き編集をして多大な労力を費やし、最初から英文を書き起こすことの難しいユーザは、訳文に問題があることは分かっている、それをどう直せばよいかを推敲できずに機械翻訳された訳文を容認せざるを得ない場合が多くあった。

【 0 0 1 4 】

20

本発明はこのような課題を解決するためになされたもので、原文を修正しても編集前の状態に戻せるような編集処理が行え、機械翻訳の機能および翻訳メモリ中の翻訳資産を有効に活用できる翻訳支援プログラム、翻訳支援装置、翻訳支援方法を提供することを目的としている。

【課題を解決するための手段】

【 0 0 1 5 】

上記した目的を達成するために、本発明の翻訳支援プログラムは、翻訳辞書、翻訳結果を記憶する翻訳メモリ、翻訳文及び原文に対する編集履歴を記憶する編集履歴記憶部を備えたコンピュータによって、ある言語の原文を他の言語に翻訳処理する翻訳支援プログラムにおいて、前記コンピュータを、前記原文を、前記翻訳辞書に基づいて機械翻訳することで翻訳文を生成する翻訳手段と、前記翻訳手段により前記原文と前記翻訳文を対応付けて前記翻訳メモリに保存する手段と、前記翻訳メモリに記憶された原文および前記翻訳文のうち少なくとも一つの文にある文字列が編集された場合、その位置に、前記原文の文字列に対して編集部分を示す特殊記号を付加して特殊記号付きの原文を生成する編集手段と

30

、前記編集手段により生成された特殊記号付きの原文を、元の原文の編集履歴として前記編集履歴記憶部に記憶する手段として機能させることを特徴とする。

上記翻訳支援プログラムにおいて、前記コンピュータを、前記特殊記号付きの原文を、特殊記号に従って再度翻訳して新たな翻訳文を生成する再翻訳手段と、前記再翻訳手段により生成された新たな翻訳文を翻訳結果の文書として前記特殊記号付きの原文に対応付けて前記翻訳メモリに記憶する手段として機能させるよう にしても良い。

40

【 0 0 1 6 】

本発明の翻訳支援装置は、ある言語の原文を、予め記憶されている翻訳辞書に基づいて機械翻訳することで他の言語の翻訳文を生成する翻訳手段と、前記翻訳手段による翻訳結果を前記原文と前記翻訳文とを対応付けて保存する翻訳メモリと、前記翻訳手段により翻訳された翻訳文および原文の少なくとも一方を編集した編集履歴が記憶される編集履歴記憶部と、前記原文および前記翻訳文のうち少なくとも一つの文にある文字列が編集された場合、その位置に、前記原文の文字列に対して編集部分を示す特殊記号を付加して特殊記号付きの原文を生成する編集手段と、前記編集手段により生成された特殊記号付きの原文を、元の原文の編集履歴として前記編集履歴記憶部に記憶する手段とを具備したことを特

50

徴とする。

上記翻訳支援装置において、前記特殊記号付きの原文を、特殊記号に従って再度翻訳して新たな翻訳文を生成する再翻訳手段と、前記再翻訳手段により生成された新たな翻訳文を翻訳結果の文書として前記特殊記号付きの原文に対応付けて前記翻訳メモリに記憶する手段とを備えても良い。

【0017】

本発明の翻訳支援方法は、翻訳手段、翻訳辞書、翻訳メモリ、編集手段、編集履歴記憶部を備えたコンピュータによって、ある言語の原文を他の言語に翻訳処理する翻訳支援方法において、前記翻訳手段が、ある言語の原文を、予め記憶されている前期翻訳辞書に基づいて機械翻訳することで他の言語の翻訳文を生成するステップと、前記翻訳手段が、前記原文と前記翻訳文とを対応付けて翻訳メモリに保存するステップと、前記原文および前記翻訳文のうち少なくとも一つの文にある文字列が編集された場合、その位置に、前記原文の文字列に対して編集部分を示す特殊記号を前記編集手段が付加して特殊記号付きの原文を生成するステップと、生成した特殊記号付きの原文を、前記編集手段が元の原文の編集履歴として編集履歴記憶部に記憶するステップとを有することを特徴とする。 なお、上記翻訳支援方法において、前記特殊記号付きの原文を再翻訳手段が特殊記号に従って再度翻訳して新たな翻訳文を生成するステップと、再度翻訳して生成した新たな翻訳文を前記編集手段が翻訳結果の文書として前記特殊記号付きの原文に対応付けて翻訳メモリに保存するステップとを有してもよい。

10

【0018】

本発明では、選択された文字列の分割位置、結合位置、および一文編集のうちの少なくとも一つを行う編集位置に特殊記号を付加した特殊記号付きの原文を生成し、それを原文の編集履歴として編集履歴記憶部に記憶するので、編集履歴記憶部を参照して特殊記号付きの原文から特殊記号を検出することで、特殊記号付きの原文を編集前の原文の形態に戻すことができる。

20

【0019】

また、特殊記号付きの原文を再度翻訳した新たな翻訳文を翻訳結果の文書として特殊記号付きの原文に対応付けて翻訳メモリに保存するので、高い精度の翻訳結果を再利用できるようになる。

【発明の効果】

30

【0020】

以上説明したように本発明によれば、原文を修正しても編集前の状態に戻せるような編集処理が行え、機械翻訳の機能および翻訳メモリ中の翻訳資産を有効に活用できる。

【発明を実施するための最良の形態】

【0021】

以下、本発明の実施の形態を図面を参照して詳細に説明する。図1は本発明に係る一実施形態の翻訳支援装置全体の構成を示すブロック図である。

この実施形態の翻訳支援装置は、大別して、文分割処理部1、翻訳手段としての翻訳処理部2、編集手段としての原文・訳文編集部3、文書出力部4等の4つの部分からなる。

【0022】

40

文分割処理部1は、翻訳対象の原言語で書かれた原文文書入力部5と、原文文書を所定の文分割規則に従って1文単位に自動的に分割する原文文書自動文分割部6と、自動文分割の際に参照される所定の文分割規則が記憶された文分割規則テーブル7とを有している。原文文書入力部5は、例えばキーボード、マウス、グラフィックユーザインタフェース画面等で構成され、原文および訳文のうち少なくとも一つに対して文字列の編集操作を行う操作手段として機能する。文分割規則テーブル7に記憶されている所定の文分割規則は、例えばメモリにファイルの形態で記憶されていても良く、また、文分割処理部1の処理プログラムの中に予め記述、つまり設定されていても良い。

【0023】

翻訳処理部2は、翻訳処理を制御する翻訳処理制御部8と、翻訳対象のある原言語（例

50

えば日本語等)で書かれた文(翻訳対象文:第1の文書)と、翻訳後の他の言語、つまり目的言語(英語等)で書かれた文(翻訳結果の文書:第2の文書(以下訳文と称す))とが一对(一組)に対応付けられて保存されている翻訳メモリ9と、この翻訳メモリ9の中に、1文単位に分割された原言語からなる原文と類似する文(類似文)があるか否かを検索する翻訳メモリ検索部10と、原文を自動的に目的言語の文に翻訳する機械翻訳処理部11と、この機械翻訳処理部11により参照される翻訳辞書としての翻訳用辞書12を有している。

【0024】

翻訳処理部2は、ある原文を予め記憶されている翻訳用辞書12に基づいて機械翻訳することで目的言語文書である訳語を生成する翻訳手段として機能する。

10

翻訳用辞書12には、言語翻訳用の辞書情報(日→英辞書、英→日辞書等の辞書データ)と、この他、形態素解析、構文解析、意味解析、言語変換用の解析ルール、変換規則等が記憶されている。

【0025】

翻訳メモリ9は、翻訳処理部2による翻訳結果の再利用を目的として翻訳前の文書とこの翻訳前の文書を翻訳した翻訳結果の文書とを対応付けて保存するものである。この翻訳メモリ9は、過去の翻訳実績を再利用するためにデータベースの形態で蓄積し翻訳作業の効率アップを図る機能の一部として用いられる。

【0026】

機械翻訳処理部11は、形態素解析、構文解析、意味解析、言語変換処理部13、訳文生成部14を有している。

20

原文・訳文編集部3は、操作手段により原文に対して行われた文字列の編集操作に応じた位置に特殊記号を付加して第3の文書(図6参照)を生成する手段として機能する。

【0027】

原文・訳文編集部3は、原文、訳文を編集した編集履歴が記憶される編集履歴記憶部15を有している。編集履歴記憶部15は、メモリ、ハードディスク装置に設けられた記憶領域等で実現される。原文・訳文編集部3は、生成した第3の文書を第1の文書の編集履歴として編集履歴記憶部15に記憶する手段として機能する。つまり、編集履歴記憶部15には、原文および訳文のうち、少なくとも一つを編集した編集履歴が記憶される。

原文・訳文編集部3は、操作手段である対訳編集画面によって、原文の中から選択された文字列に対して、編集履歴記憶部15に記憶された特殊記号付きの訳文を基に、分割、結合、および一文編集のうちの少なくとも一つを行う位置を推定し、その位置に特殊記号を付加して特殊記号付きの原文を生成する手段として機能する。原文・訳文編集部3は、生成された特殊記号付きの原文を原文の編集履歴として編集履歴記憶部15に記憶する手段として機能する。

30

【0028】

訳文生成部14は、形態素解析、構文解析、意味解析、言語変換処理部13により形態素解析、構文解析、意味解析、言語変換されて意味をなす語句(文字列)となったものを文の形態に並べる処理を行う。

【0029】

40

文書出力部4は、文の形態の原文・訳文などを文書の形態で出力するものであり、原文、訳文、それぞれの修正文のうち少なくとも一つを出力するプリントドライバ、表示ドライバ等を含むプログラムと、プリンタ、表示装置等のハードウェア等である。

【0030】

この翻訳支援装置のハードウェアは、CPU、メモリ、ハードディスク装置等を備えたコンピュータと、このコンピュータに接続された表示装置および印刷装置等である。ハードディスク装置にはコンピュータシステム全体を動作させるオペレーティングシステム(以下OSと称す)と、機械翻訳を実行する制御プログラム(以下翻訳支援プログラムと称す)がインストールされており、これら翻訳支援プログラム、OS、CPU、メモリ等が協働して、文分割処理部1、翻訳処理部2、原文・訳文編集部3、文書出力部4等の処理

50

動作を実現する。

【0031】

以下、図2を参照してこの実施形態の翻訳支援装置の動作を説明する。図2はこの実施形態の翻訳支援装置の処理全体を示すフローチャートである。

【0032】

この翻訳支援装置では、原文文書入力部5に原文文書が入力されると(図2のステップS201)、原文文書自動文分割部6は、原文文書を1文単位に分割する(ステップS202)。この原文文書自動文分割部6では、日本語の文章の場合、1つの文は、読点を区切りとして分割される。また、英語の文章の場合、1つの文は、ピリオド、コロン、セミコロンなどを区切りとして分割される。この他、例えば括弧、改行記号、htmlのような書式情報付き文書では、1つの文は、改行記号、カッコなどを区切りとして分割される。また、英文の場合、“Dr.”などのようにピリオドが付く語があるため、個別に考慮すべき単語を辞書にまとめ、適宜参照する。上記区切り情報は予めメモリあるいはプログラム上に記憶(設定)されている。

10

【0033】

原文文書自動文分割部6により1文単位に分割された原文は、1文ずつ、最後の文になるまで(ステップS203)、翻訳処理部2へ送られて、翻訳処理部2によって翻訳処理が行われる(ステップS204)。なお、ステップS204の翻訳処理の詳細な内容は後述する。

【0034】

20

翻訳処理が行われた訳文は、原文とともに原文・訳文編集部3へ送られる。原文・訳文編集部3では、訳文および/または原文の編集処理が行われる(ステップS205)。このステップS205の編集処理の詳細な内容については後述する。

【0035】

ステップS205の編集処理が行われた後、編集済みの原文に対する翻訳処理が再度必要な場合、原文・訳文編集部3より翻訳処理部2へ翻訳対象の文が戻される。そして、翻訳処理部2により文書の有無が判定されて(ステップS203)、翻訳処理が行われる(ステップS204)。

なお、翻訳対象の文が複数存在したとしても、ステップS203の判定処理を経ることで、すべての翻訳対象文に対して1文ずつ翻訳処理が実行される。

30

【0036】

翻訳処理が再度必要でない場合(ステップS206のNo)、原文・訳文編集部3は、原文と翻訳結果を文書出力部4へ送り(ステップS207)、原文・訳文編集部3としての処理動作を終了する。

【0037】

< 翻訳処理部2の動作 >

ここで、図3を参照して図2のステップ204で示した翻訳処理について説明する。図3は図2のステップ204で示した翻訳処理を示すフローチャートである。

【0038】

文分割処理部1によって1文単位に分割された翻訳対象文が翻訳処理部2に入力されると(ステップS301)、翻訳処理部2では、翻訳メモリ検索部10が、翻訳処理制御部8からの命令により、翻訳対象文をキーにして翻訳メモリ9を検索することで(ステップS302)、翻訳対象文と類似した文が翻訳メモリ9内に存在するか否かを判定する(ステップS303)。なおステップ302の翻訳メモリ検索部10による翻訳メモリ検索処理の詳細な内容については後述する。

40

【0039】

検索の結果、翻訳メモリ9内に、翻訳対象文と類似した文が存在した場合(ステップS303のYes)、翻訳メモリ検索部10は、その類似文を翻訳処理による訳文と判定して、翻訳結果を表示装置の表示画面へ出力し(ステップS304)、この翻訳対象文に対する翻訳処理を終了する。

50

また、翻訳対象文と類似した文が存在しない場合（ステップ S 3 0 3 の N o ）、翻訳メモリ検索部 1 0 は、翻訳対象文を機械翻訳処理部 1 1 へ送り、機械翻訳処理を実行させる。

機械翻訳処理部 1 1 は、翻訳メモリ検索部 1 0 より受けた取った翻訳対象文に対して形態素解析、構文解析、意味解析、言語変換の各種処理からなる原文解析処理（ステップ S 3 0 5 ）と、訳文生成処理（ステップ S 3 0 6 ）とを行うことで機械翻訳処理を行う。

機械翻訳処理部 1 1 は、機械翻訳処理が終了すると、翻訳結果を原文・訳文編集部 3 へ出力する（ステップ S 3 0 4 ）。

【 0 0 4 0 】

< 原文・訳文編集部 3 の動作 >

続いて、図 4 のフローチャートを参照して、上記図 2 のステップ 2 0 5 で示した原文・訳文編集部 3 の処理の詳細について説明する。

【 0 0 4 1 】

原文・訳文編集部 3 に、翻訳処理部 2 から原文および翻訳処理の結果が送られると、原文・訳文編集部 3 は、原文・訳文の同じ内容を初期値として編集履歴記憶部 1 5 に記憶するとともに、その原文・訳文を表示装置の表示画面に表示する。原文・訳文の編集処理は、表示画面に表示された原文・訳文のうち、ユーザが選択した文に対して実行される。この表示画面は、原文および訳文の対訳編集画面であり、原文および訳文の少なくとも一つに対して編集対象の文字列の選択操作を行う操作手段として機能する。

【 0 0 4 2 】

図 4 には、ユーザが編集したい文を選択してから 1 回の編集作業が終了するまでの流れを示すものであり、すべての編集作業が終了するまで、必要に応じて図 4 の処理が繰り返される。

【 0 0 4 3 】

ユーザにより選択された編集対象の文字列、つまり編集対象文が原文・訳文編集部 3 に入力されると（図 4 のステップ S 4 0 1 ）、原文・訳文編集部 3 は、編集対象文が訳文か原文かを判定する（ステップ S 4 0 2 ）。この判定の仕方としては、入力元（訳文は機械翻訳処理部 1 1 から入力、原文は翻訳処理制御部 8 から入力）がどこであるか、つまりどこから送られてきたかで判定する方法と、ユーザが操作した画面あるいは文自体（原文と訳文を 2 分割画面に別個に表示しているため）で判定する方法がある。

【 0 0 4 4 】

判定の結果、編集対象文が訳文の場合（ステップ S 4 0 2 の Y e s ）、原文・訳文編集部 3 は、訳文編集を行い（ステップ S 4 0 3 ）、編集結果を文書出力部 4 へ出力して（ステップ S 4 1 2 ）、編集処理を終了する。

また、編集対象文が原文の場合（ステップ S 4 0 2 の N o ）、原文・訳文編集部 3 は、編集内容に応じて 1 文の分割を行うか否か（ステップ S 4 0 4 ）、複数文の結合を行うか否か（ステップ S 4 0 7 ）を判定し、この判定結果に応じて処理を行う。

【 0 0 4 5 】

例えば 1 文の分割を行う場合（ステップ S 4 0 4 の Y e s ）、原文・訳文編集部 3 は、文分割処理を実行し（ステップ S 4 0 5 ）、複数文の結合を行う場合には、文結合処理を実行し（ステップ S 4 0 8 ）、これら 2 つ以外の場合、つまり、編集内容が 1 文内での変更のみに留まる場合（ステップ S 4 0 4 の Y e s ）、原文・訳文編集部 3 は、1 文編集処理を実行する（ステップ S 4 0 9 ）。これら文分割、文結合、一文編集等の各処理によって、編集文に特殊記号が付される。各処理の内容については、後で図 5 の例を用いて詳細に説明する。なお、選択入力された文が長い場合、分割と結合を同時に行うことも有り得る。この場合、ステップ S 4 0 5 、S 4 0 8 の処理が同時に実行されることになる。また、他の一文編集処理との組み合わせも考えられる。

【 0 0 4 6 】

これらの処理の後、原文・訳文編集部 3 は、編集履歴記憶部 1 5 に対して原文編集履歴内容の更新処理を行い（ステップ S 4 0 6 ）、編集履歴記憶部 1 5 に、処理内容が時系列

10

20

30

40

50

で記憶される。

【 0 0 4 7 】

原文の編集作業を終了した後、ユーザは再翻訳が必要か否かを判断する。再翻訳が必要と判断したユーザは、表示画面上の翻訳ボタンを操作し、再翻訳が不要と判断したユーザは、表示画面上の翻訳ボタン以外のボタン、あるいはキー操作を行うので、原文・訳文編集部 3 は、原文編集後のユーザの操作に応じて処理内容を変える（ステップ S 4 1 0 ）。

【 0 0 4 8 】

例えばユーザにより表示画面上の翻訳ボタンが操作された場合、原文・訳文編集部 3 は、再翻訳が必要と判定し（ステップ S 4 1 0 の Y e s ）、この場合は、編集された原文を翻訳処理部 2 に渡し、再度の翻訳処理を実行させる（ステップ S 4 1 1 ）。

10

【 0 0 4 9 】

また、ユーザにより他の操作が行われた場合、原文・訳文編集部 3 は、再翻訳を不要と判定し（ステップ S 4 1 0 の N o ）、翻訳処理（ステップ S 4 1 1 ）をスキップし、訳文編集が必要か否かの判定処理を行う（ステップ S 4 1 2 ）。

【 0 0 5 0 】

このステップ S 4 1 2 の判定処理では、ユーザにより次の文（原文あるいは訳文）が編集対象として指定された場合に、原文・訳文編集部 3 は、訳文編集を必要と判定し（ステップ S 4 1 2 の Y e s ）、指定された編集対象文に対して訳文編集を実行し（ステップ S 4 0 3 ）、その後、翻訳結果の文を翻訳メモリ 9 と文書出力部 4 へ出力し（ステップ S 4 1 3 ）、編集処理を終了する。原文・訳文編集部 3 より出力された翻訳結果の文は、特殊記号が付加された編集後の原文に対応付けられて翻訳メモリ 9 に保存（登録）される。

20

【 0 0 5 1 】

また、ユーザにより次の文（原文あるいは訳文）が編集対象として指定されず、訳文編集が不要な場合（ステップ S 4 1 2 の N o ）、原文・訳文編集部 3 は、訳文編集処理（ステップ S 4 0 3 ）をスキップして、編集文を出力し（ステップ S 4 1 3 ）、編集処理を終了する。

【 0 0 5 2 】

< 原文・訳文編集部 3 の動作の実例と翻訳メモリ 9 への登録例 >

図 5 は翻訳処理部 2 から原文・訳文編集部 3 に送られてきた、編集処理前の原文・訳文および編集履歴記憶部 1 5 の内容を示したものである。これらの例を用いて、具体的な原文・訳文編集部 3 の動作と翻訳メモリ 9 への登録内容について説明する。原文・訳文未編集の状態では、編集履歴記憶部 1 5 の内容（原文・訳文）は、それぞれの原文・訳文と全く同じ内容になっている。

30

【 0 0 5 3 】

これらの例では、翻訳処理から出力された翻訳結果は原文の内容を十分に反映したものとは言えず、修正が必要である。

例えば文 5 1 は、ひらがな表記になっているため、「たなか」が人名と認識されていない例である。文 5 2 では、「田中ですが、」の「が、」が、日本語では軽い接続の意味で使われているが、翻訳結果では逆接の意味と解釈され、逆接の接続詞 " although " が出力されている。文 5 3 および文 5 4 は、レイアウトの都合で 1 文が 2 つに分割されたため、正しく翻訳されていない例である。

40

【 0 0 5 4 】

上記の点を考慮して、ユーザが文の選択操作を行い、この選択操作に応じて原文・訳文編集部 3 が文分割処理、文結合処理、1 文編集処理を行い、原文を修正した結果を図 6 に示す。

【 0 0 5 5 】

文 6 1 は、ひらがな表記を漢字表記に修正したものである。漢字表記にすることで「田中」が人名と解釈され、正しい翻訳結果が出力されている。このように意味が一意に決まるように表記や表現を変更することで、正しい翻訳結果が得られる場合が多い。

文 6 2 および文 6 3 は、「が、」で接続された原文を 2 文に分割し、それぞれ文として完

50

結するようにしたものである。これらの文を再度翻訳すると、当然だが、逆接の接続詞althoughは訳文に現れなくなる。このような文の分割処理を行うと、編集履歴記憶部15には、分割した前半文字列の末尾と、後半文字列の先頭に特殊記号" @数字@"が挿入された編集履歴が記憶される。"@"と"@ "の間の数字は、この位置に特殊記号を挿入したことを識別するための特殊記号のID番号であり、同じID番号が付いている原文同士は、編集前は繋がっていたことを示す。数字は、通常、文の文節や文の結合部が検出された際に原文・訳文編集部3により連続番号で付与される。

【0056】

文64は、文53および文54の2文を結合したものである。結合した文の再翻訳結果は、意味の通るものとなっている。このような文の結合処理を行うと、編集履歴記憶部15には文同士の結合部に" @数字@"が挿入された編集履歴が記憶される。

10

【0057】

結合処理の場合は同じID番号を持つ複数の原文は存在しないが、同じIDを持つ原文と訳文は存在する。機械翻訳処理部11は、結合した原文を再翻訳するので、訳文中の結合部は、本来は存在しないはずであるが、編集履歴記憶部15(訳文)に保存されている結合前の翻訳結果と再翻訳結果とを比較することで、原文の結合部に対応する訳文の結合部を推定(特定)し翻訳を行う。例えば文64は、" This processing "の後で切れると推定(特定)される。これは、編集履歴記憶部15に記憶された@前の文字列" This processing "と再翻訳結果の" This processing "が完全一致しているので、訳文の切れ目はprocessingの後であると推定すると、2つの翻訳結果の一致度がもっとも高くなるためである。

20

すなわち、翻訳処理制御部8は、編集履歴記憶部15に記憶された訳文(最翻訳結果)に対して、編集履歴記憶部15に記憶された編集記号付き訳文を基にして、文の切れ目を推定する。

【0058】

図7は原文を修正して機械翻訳処理部11が再翻訳した訳文に、更に修正を加えた例である。

この例は、文74の「処理」に対する訳語「processing」を「process」に変更する修正を加えた例である。このように、原文の修正を行うことで構文的に正しい翻訳結果が得られると、後は翻訳結果の一部を修正するだけで一定水準の訳文を得ることができる。

30

【0059】

したがって、最初から自分で英文を書き起こせるユーザは、訳文を最初から自分で作成する必要がなくなり、翻訳の労力が大幅に軽減できる。

また、最初から英文を自分で作成することが難しいユーザにとっては、わずかな修正で正しい英文が得られるため、機械翻訳の翻訳結果が間違っているとしても、比較的容易に修正が可能となる。

【0060】

図8は上記の翻訳結果から、翻訳メモリ9に登録された内容を示した図である。

翻訳メモリ9に登録された原文は、修正を加えた原文ではなく、編集履歴記憶部15の内容であり、オリジナルの原文の文字列に対して文分割部分や文結合部を示す特殊記号を加えたもの(第3の文書の形態)になっている。

40

【0061】

この実施形態の翻訳支援装置では、このような翻訳メモリ9への登録方法を採用しているが、原文を分割した場合に限り、分割前の原文に復元したものを翻訳メモリ9に登録する、という方法も可能である。この方法によれば、翻訳メモリ検索部10に特別な機能が無くても、翻訳対象の文が編集前のオリジナルの原文と一致する場合、翻訳メモリ9に登録されたオリジナルの原文と対になっている訳文を出力することができる。

【0062】

以下、より汎用的な図8に示した内容で翻訳メモリ9に登録する場合について説明する。

。

50

< 新たな文書翻訳時の翻訳メモリ検索動作 >

図 8 に示した翻訳メモリ 9 を使って、図 9 に示す新たな文書を翻訳する場合の、翻訳メモリ検索の動作について説明する。翻訳メモリ検索は、図 3 のステップ S 3 0 2 に示した処理である。

【 0 0 6 3 】

この場合、翻訳メモリ検索部 1 0 は、翻訳対象の 1 つの文（翻訳文 0 1 と呼ぶとする）に対して、翻訳メモリ 9 内の 1 つのデータ（メモリデータ 0 1 と呼ぶ）と比較処理を行い、比較処理終了後、翻訳メモリ 9 の次のデータとの比較処理を行う。これを繰り返し翻訳文 0 1 と翻訳メモリ 9 のすべてのデータとの比較処理が終了すると、次の翻訳文 0 2 の検索処理を開始する。

10

【 0 0 6 4 】

このような比較方法をとったのは、本実施例の動作を分かりやすく説明するためであり、検査高速化のために、インデックス作成などの他の検索方法をとったとしても、本特許の範囲を逸脱するものではない。なお、比較処理の内容によっては、複数の翻訳文、複数のメモリデータをまとめて比較する場合もある。

【 0 0 6 5 】

図 1 0 および図 1 1 は、翻訳文 0 1 と、翻訳メモリデータ 1 件（メモリデータ 0 1 と呼ぶことにする）との比較処理を示すフローチャートであり、図 1 0 は完全一致検索処理を示し、図 1 1 は類似文検索処理を示す。

【 0 0 6 6 】

翻訳メモリ検索部 1 0 は、まず、図 1 0 の完全一致検索処理を実行した後、完全一致するメモリデータ 0 1 が検出されなかった翻訳文 0 1 に対して、引き続き、図 1 1 の類似文検索処理を実行する。

20

【 0 0 6 7 】

以下では、まず、図 1 0 のフローチャートを用いて検索処理の動作を説明し、図 9 の例文が当てはまるケースに対して具体例を使った説明を加える。

< 翻訳メモリ検索動作 - 完全一致検索 >

翻訳メモリ検索部 1 0 に翻訳対象文（翻訳文 0 1 ）が入力されると（ステップ S 5 0 1 ）、翻訳メモリ検索部 1 0 は、まず、翻訳文 0 1 とメモリデータ 0 1 とを比較して互いが完全一致するか否かを判定する（ステップ S 5 0 2 ）。ここで、完全一致とは、翻訳文 0 1 とメモリデータ 0 1 、つまり翻訳メモリ 9 に記憶されている翻訳文とが一語一句変わらないことを指す。

30

【 0 0 6 8 】

比較の結果、翻訳文 0 1 とメモリデータ 0 1 とが完全一致した場合、翻訳メモリ検索部 1 0 は、翻訳メモリ 9 に一致する文が存在するものと判定し（ステップ S 5 0 2 の Y e s ）、一致した文をメモリデータ 0 1 とする（ステップ S 5 0 3 ）。

【 0 0 6 9 】

このステップ S 5 0 3 のような結果になるケースは、翻訳文 0 1 が図 9 の文 9 0 の場合に相当する。つまり最初に行った翻訳時に、漢字表記になるよう原文を編集していても、翻訳メモリ 9 にはオリジナルのひらがな表記の原文である図 8 の対原文 8 1 「わたしはたなかです。」が登録されていたため、オリジナルと同じひらがな表記の翻訳文 0 1 にメモリデータ 0 1 がマッチして、正しい英文 " I am Tanaka. " が翻訳結果とされる。以上は、1 文編集処理を行って翻訳メモリ 9 のデータがマッチするケースの場合である。

40

【 0 0 7 0 】

ステップ S 5 0 2 の判定処理において、翻訳文 0 1 とメモリデータ 0 1 が完全一致しない場合（ステップ S 5 0 2 の N o ）、翻訳メモリ検索部 1 0 は、翻訳文 0 1 とメモリデータ 0 1 の原文中の " @ " の前にある文字列とを比較して完全一致するか否かを判定する（ステップ S 5 0 4 ）。

【 0 0 7 1 】

この比較判定の結果、翻訳文 0 1 とメモリデータ 0 1 の原文中の " @ " の前にある文字

50

列とが完全一致しない場合（ステップS504のNo）、翻訳メモリ検索部10は、翻訳文01と一致する文は無しと判定し、検索処理を終了する。

【0072】

一方、翻訳文01とメモリデータ01の原文中の“@”の前にある文字列とが完全に一致した場合（ステップS504のYes）、翻訳メモリ検索部10は、“@”がメモリデータ01の文末に存在するか否かを判定する（ステップS505）。

【0073】

この判定の結果、“@”がメモリデータ01の文末に存在した場合（ステップS505のYes）、翻訳メモリ検索部10は、さらに一致しなかった翻訳文中の文字列（この部分を未一致部と呼び、一致している部分を一致部と呼ぶ）が、先頭部分に同じID番号が付いた@記号を持った翻訳メモリ9のデータ（これをメモリデータ02と呼ぶ）と一致するか否かを判定する（ステップS506）。

10

【0074】

この判定の結果、未一致部とメモリデータ02とが完全一致した場合、翻訳メモリ検索部10は、翻訳メモリ9に一致する文が有るものと判定し（ステップS506のYes）、一致した文については、特殊記号“@”を削除すると共にメモリデータ01の後にメモリデータ02を結合し、結合文（メモリデータ01+02）を生成する（S507）。

このS507の結果になるケースは、翻訳文01が図9の文91の場合に相当する。

【0075】

翻訳メモリ検索部10は、翻訳メモリ9を検索した場合、まず「私は田中ですが、」の部分メモリデータの対原文82「私は田中ですが、@1@」の“@1@”の前の部分と一致する（ステップS504）。なお@1@の数字の「1」は新たな文節や結合位置が検出されたときに、機械的に順に付加されている連続番号である。

20

次のステップS506の調査処理において、「彼は中田です。」の部分メモリデータの対原文83「@1@彼は中田です。」の@1@より後ろの文字列と一致する。

この調査結果に基づいて、S507の処理では、“I am Tanaka. He is Nakada.”という、逆接の接続詞の入らない英文の翻訳結果が生成される。

【0076】

このように、翻訳メモリ9の内容を作成するときに、原文を分割して翻訳メモリ9に登録したときに、分割されたデータに、互いに連結していることを示す特殊記号（@1@等）が付加されているので、オリジナルと同じく分割されていない翻訳文にメモリデータがマッチする。以上は、文分割処理を行って登録した翻訳メモリデータがマッチするケースである。

30

【0077】

ステップS506の判定処理で「未一致部」がメモリデータ02と完全一致しない場合（ステップS506のYes）、翻訳メモリ検索部10は、予め設定された数式（関数式）を用いてあいまい一致の一致率を計算し、その計算結果の数値と予め設定されていた基準値とを比較して、計算結果の数値が基準値を超えているか否かを判定する（ステップS508）。

すなわち、類似度がある一定の値以上の場合（ステップS508のYes）、翻訳メモリ検索部10は、翻訳文01の「一致部」とメモリデータ01とがマッチ（一致）したものと見なし、翻訳結果にはメモリデータ01の訳文を出力すると共に、「未一致部」についてはメモリデータとマッチしなかったものと判定する（ステップS509）。

40

類似度がある一定の値に満たない場合（ステップS508のNo）、翻訳メモリ検索部10は、「一致部」、「未一致部」共にメモリデータとマッチしなかったものと判定し、完全一致の検索処理を終了する。

これは、文の一部はマッチしていても、残りの部分の一致度が非常に低い場合は、文全体としての意味を考え直した方が良い場合があるからである。ただしユーザの意志によって、ステップS508の判定処理をスキップする設定に変更し、「一致部」はメモリデータとマッチしたものと判定して、メモリデータの訳文を表示するようなモードを導入する

50

ようにしてもよい。

【 0 0 7 8 】

ステップ S 5 0 9 のような結果になるケースは、翻訳文が図 9 の文 9 2 の場合に相当する。文 9 2 の前半「私は田中ですが、」は、メモリデータの対原文 8 2 の「@」前の部分「私は田中ですが、」と一致するが、後半部の「生まれは千葉です」がメモリデータの対原文 8 3 と一致しているのは助詞「は」と助動詞「です」のみで、一致度は非常に低い（8 語中 3 語）。このため、「一致部」は、「未一致部」とともにメモリデータと不一致と判定される。

ステップ S 5 0 5 の判定処理において、「@」がメモリデータの文末にない場合（ステップ S 5 0 5 の No）、翻訳メモリ検索部 1 0 は、ステップ S 5 0 4 の処理での翻訳メモリ 9 の未一致部が、次の翻訳対象文（翻訳文 0 2 と呼ぶ）と完全一致するか否かを判定する（ステップ S 5 1 0）。

10

未一致部と翻訳文 0 2 が一致した場合（ステップ S 5 1 0 の Yes）、翻訳メモリ検索部 1 0 は、一致文ありと判定し、一致文を、翻訳文 0 1 に対してはメモリデータ 0 1 の「@」前の部分、翻訳文 0 2 に対してはメモリデータ 0 1 の「@」の後の部分とする結合処理を行い（ステップ S 5 1 1）、訳文を生成する。

【 0 0 7 9 】

このステップ S 5 1 1 のような結果になるケースは、翻訳文が図 9 の文 9 3 および文 9 4 の場合に相当する。

【 0 0 8 0 】

20

すなわち、翻訳メモリ検索部 1 0 が翻訳メモリ 9 を検索したところ、まず図 9 の文 9 3 「この処理は」の部分メモリデータの対原文 8 4 の「この処理は@ 2 @」の「@ 2 @」の前の部分と一致する（ステップ S 5 0 4）。

【 0 0 8 1 】

次に、図 9 の文 9 4 「以下のように行います。」の部分「@ 2 @」の後の部分と一致する（ステップ S 5 1 0）。ステップ S 5 1 1 では、この検索結果に基づいて、「This process@ 2 @ is performed as follows.」という訳文（翻訳結果）が生成される。

【 0 0 8 2 】

このように、翻訳メモリ 9 のデータを作成したときに、原文を結合して翻訳メモリ 9 に登録しても、結合部に結合の履歴を示す特殊記号（@ 2 @）が挿入されるので、オリジナルと同じく結合されていない翻訳文にメモリデータがマッチする。

30

【 0 0 8 3 】

また、原文の結合部と対応する結合部が訳文においても特殊記号（@ 2 @）が挿入されているので、結合されていない原文の各部分に対して、対応する訳文を出力することができる。以上は、文結合処理を行って登録した翻訳メモリデータがマッチするケースである。

【 0 0 8 4 】

ステップ S 5 1 0 の処理において、翻訳メモリ 9 の「未一致部」が翻訳文 0 2 と完全一致しない場合（ステップ S 5 1 0 の No）、翻訳メモリ検索部 1 0 は、あいまい一致の一致率を計算し、ステップ S 5 0 8 の判定処理と同様に、類似度がある一定の値を超えているか否かを判定する（ステップ S 5 1 2）。

40

【 0 0 8 5 】

判定の結果、類似度がある一定の値以上の場合（ステップ S 5 1 2 の Yes）、翻訳メモリ検索部 1 0 は、翻訳メモリ 9 の「一致部」を翻訳文 0 1 とマッチしたものと判定し、翻訳結果としてメモリデータの訳文の「@」前の部分を出し、「未一致部」についてはメモリデータとマッチしなかったものと判定する（ステップ S 5 1 3）。

【 0 0 8 6 】

また、類似度がある一定の値に満たない場合（ステップ S 5 1 2 の No）、翻訳メモリ検索部 1 0 は、翻訳文 0 1、0 2 とともにメモリデータとマッチしなかったものと判定し、完全一致の検索処理を終了となる。なお、この他、ステップ S 5 0 8 の基準値との比較判

50

定処理の場合と同様に、ステップS512の比較判定を行わずに、「一致部」のメモリデータの訳文を表示するモードを設定してもよい。

ステップS513のような結果になるケースは、訳文が図9の文95、文96の場合に相当する。

すなわち、文95は、訳メモリデータの対原文84“@”前の文字列と一致するが、文96が“@”後の文字列と一致しているのは動詞「行う」のみである上、活用形が異なり、一致度は非常に低い。このため、訳文01は、訳文02と共にメモリデータと不一致と判定される。

【0087】

< 訳メモリ検索動作 - 類似文一致検索 >

10

図10の完全一致検索で一致文なしと判定された訳対象文に対して、図11の類似文検索を行う。以下では、図11のフローチャートを参照して類似文検索処理について説明する。文分割処理、文結合処理が行われたメモリデータが類似文としてヒットする場合には、図9の具体例を用いて説明する。

訳メモリ検索部10は、図10に示した完全一致検索処理で訳対象文と一致する文が存在しないものと判定すると、類似文検索処理を行う。

【0088】

この場合、図11に示すように、一致文なしの訳対象文が訳メモリ検索部10に入力されると(ステップS601)、訳メモリ検索部10は、まず、訳メモリ9を検索し、訳メモリ9に類似文はあるか否かを判定する(ステップS602)。

20

【0089】

ここでは、訳対象文とメモリデータ01との類似度が一定以上の値か否かで類似文の有無を判定する。

類似文検索での一致度は、日本語の場合、文字列の中で一致した文字数がどのくらいあるかといった文字数割合を計算で求める。また、英語の場合は文字列の中で一致した単語数がどのくらいあるかの単語数割合を計算で求める。

【0090】

活用形が異なる単語は、一定の係数をかけた上で「一致」と判定する。計算の結果として得られた類似度が、ユーザが予め設定しておいた値以上の場合(ステップS602のYes)、訳メモリ検索部10は、訳メモリ9に類似文が存在したものと判定し、類似文をメモリデータ01とする(ステップS603)。

30

【0091】

また、類似度が、ユーザが予め設定しておいた値未満の場合(ステップS602のNo)、訳メモリ検索部10は、訳メモリ9に類似文が存在しないものと判定し、次に訳対象文とメモリデータ01の原文中の“@”の前にある文字列との類似度を判定する(ステップS604)。

訳対象文と“@”の前にある文字列との類似度を判定した結果でも、類似度が一定の値に満たない場合(ステップS604のNo)、訳メモリ検索部10は、類似文なしと判定し、類似文検索処理を終了する。

【0092】

40

一方、類似文ありと判定した場合(ステップS604のYes)、訳メモリ検索部10は、“@”がメモリデータ01の文末にあるか否かを判定する(ステップS605)。

【0093】

この判定の結果、“@”がメモリデータ01の文末にある場合(ステップS605のYes)、訳メモリ検索部10は、さらにステップS604の比較判定処理で一致しなかった訳文中の未一致部と、メモリデータ02との類似度を判定する(ステップS606)。

【0094】

この判定の結果、類似度がある一定以上の値に満たない場合(ステップS606のNo

50

)、翻訳メモリ検索部10は、類似文なしと判定し、処理を終了する。

また、類似度がある一定の値以上の場合(ステップS606のYes)、翻訳メモリ検索部10は、類似文ありと判定し、検索された類似文の特殊記号@数字@を削除し、メモリデータ01の後にメモリデータ02を結合した文を生成する(ステップS607)。

【0095】

ステップS607の結果となるケースは、翻訳文が図9の文97の場合に相当する。

つまり、図9の文97の「わたしは田中ですが、」および翻訳メモリ9の対原文82「私は田中ですが、」と、文98の「かれは中田です。」および翻訳メモリ9の対原文82「彼は中田です。」とは、「私」および「彼」の表記が異なる以外は一致している。

【0096】

ステップS605の比較判定処理において、“@”がメモリデータ01の文末にない場合(ステップS605のNo)、翻訳メモリ検索部10は、さらにステップS604の比較判定処理での翻訳メモリ9の未一致部と、翻訳文02との類似度を判定する(ステップS608)。

【0097】

類似度がある一定以上の値に満たない場合(ステップS608のNo)、翻訳メモリ検索部10は、類似文なしと判定して処理を終了する。

一方、類似度がある一定の値以上の場合(ステップS608のYes)、翻訳メモリ検索部10は、類似文ありと判定し、類似文は、翻訳文01に対してはメモリデータ01の“@”前の部分、翻訳文02に対してはメモリデータ01の“@”の後の部分とした文を生成する(ステップS609)。

ステップS609の結果になるケースは、翻訳文が図9の文98および文99の場合に相当する。

【0098】

つまり図9の文98の「あの処理は」は、メモリデータ対原文84の“@”前部分と指示語が異なるだけであり(「この」と「あの」)、文99はメモリデータの対原文84の“@”後部分と完全一致している。

【0099】

このようにこの実施形態の翻訳支援装置によれば、翻訳メモリ9に原文とその翻訳結果を登録する際、原文に対して編集が行われていた場合、編集前の原文の状態を再現できるよう、原文を編集した文字列の位置に特殊記号(@数字@)を付加した文書(第3の文書)を生成し、それを第1の文書の編集履歴として編集履歴記憶部15に記憶する。

また、生成した第3の文書を再翻訳した翻訳結果の文書(第4の文書)を第3の文書と対応付けて翻訳メモリ9に保存(登録)するので、機械翻訳の精度を向上できると共に、編集履歴記憶部15を参照することで機械翻訳が翻訳し易くなる。

さらに、第3の文書に付加されている特殊記号を付加時の規則で削除することで原文の状態に戻せるので、原文自体に対して編集を気軽に行うことができる。さらに原文に対して文分割、文結合、用語の置換、情報の追加などの編集を気軽に行うことができる。

【0100】

この結果、機械翻訳の能力を十分活用し、翻訳作業の効率を高めることができる。

【図面の簡単な説明】

【0101】

【図1】本発明に係る一実施形態の翻訳支援装置の全体構成を示すブロック図。

【図2】図1の翻訳支援装置の処理全体の流れを表すフローチャート。

【図3】図1の翻訳支援装置の翻訳処理の流れを表すフローチャート。

【図4】図1の翻訳支援装置の編集処理の流れを表すフローチャート。

【図5】編集処理前の原文・訳文および編集履歴記憶部の内容を示す図。

【図6】原文編集後の原文、再翻訳結果および編集履歴記憶部の内容を示す図。

【図7】訳文編集後の原文、再翻訳結果および編集履歴記憶部の内容を示す図。

【図8】翻訳メモリの登録内容を示す図。

10

20

30

40

50

【図 9】新たな文書での翻訳対象文。

【図 10】図 1 の翻訳支援装置において、完全一致検索の際の翻訳メモリ検索処理を示すフローチャート。

【図 11】図 1 の翻訳支援装置において、類似文検索の際の翻訳メモリ検索処理を示すフローチャート。

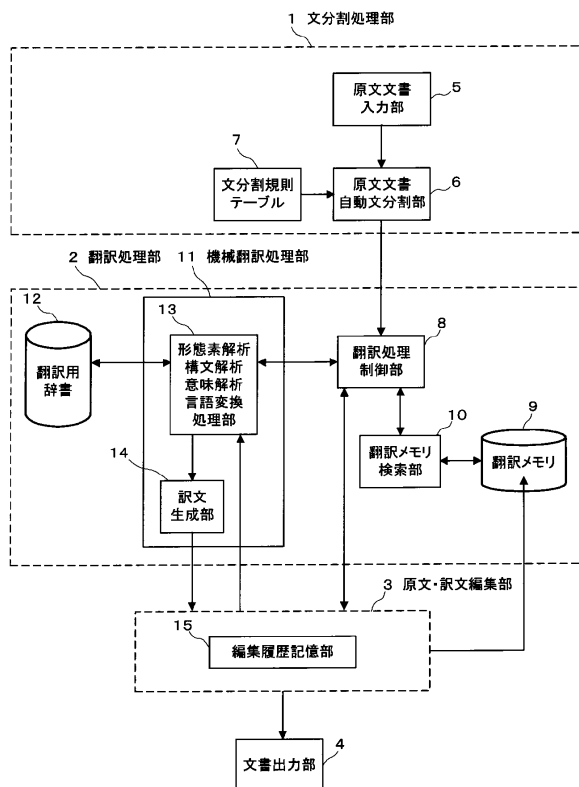
【符号の説明】

【0102】

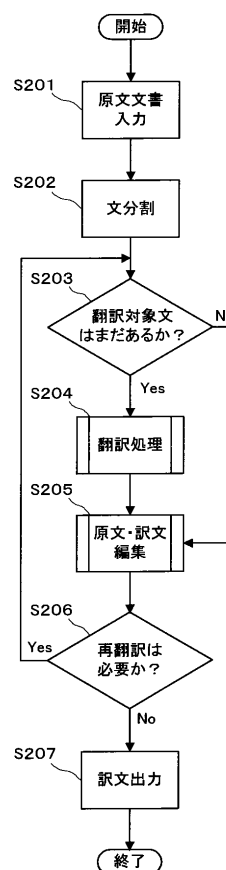
1 ... 文分割処理部、2 ... 翻訳処理部、3 ... 原文・訳文編集部、4 ... 文書出力部、5 ... 原文文書入力部、6 ... 原文文書自動分割部、7 ... 文分割規則テーブル、8 ... 翻訳処理制御部、9 ... 翻訳メモリ、10 ... 翻訳メモリ検索部、11 ... 機械翻訳処理部、12 ... 翻訳用辞書、13 ... 形態素解析、構文解析、意味解析、言語変換処理部、14 ... 訳文生成部、15 ... 編集履歴記憶部。

10

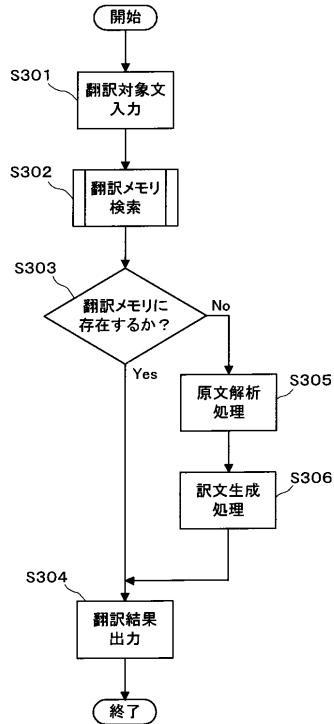
【図 1】



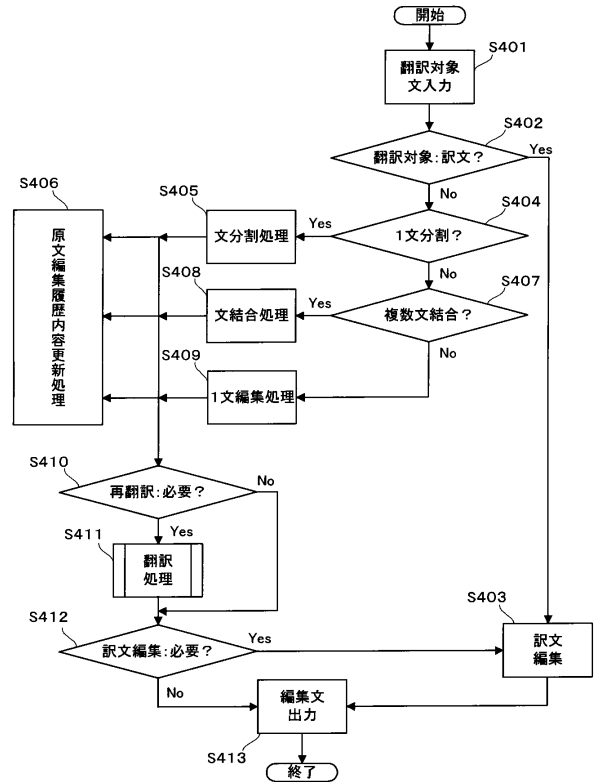
【図 2】



【図 3】



【図 4】



【図 5】

	編集履歴記憶部 (原文)	原文	訳文	編集履歴記憶部 (訳文)
51	わたしは たなかです	わたしは たなかです	I am in a shelf.	I am in a shelf.
52	私は田中ですが、 彼は中田です。	私は田中ですが、 彼は中田です。	He is Nakada although I am Tanaka.	He is Nakada although I am Tanaka.
53	この処理は	この処理は	This processing	This processing
54	以下のように行い ます。	以下のように行い ます。	It carries out as follows.	It carries out as follows.

【図 7】

	編集履歴記憶部 (原文)	原文(編集後)	訳文(再翻訳結果 + 訳文編集)	編集履歴記憶部 (訳文)
71	わたしはたなかです	私は田中です。	I am Tanaka.	I am in a shelf.
72	私は田中ですが、 @1@	私は田中です。	I am Tanaka.	He is Nakada although I am Tanaka.
73	@1@彼は中田です。	彼は中田です。	He is Nakada	
74	この処理は@2@ 以下のように行い ます。	この処理は以下の ように行います。	This processing is performed as follows.	This processing@2@ It carries out as follows.

【図 6】

	編集履歴記憶部 (原文)	原文(編集後)	訳文(再翻訳結果)	編集履歴記憶部 (訳文)
61	わたしは たなかです	わたしは たなかです	I am Tanaka.	I am in a shelf.
62	私は田中ですが、 @1@	私は田中です。	I am Tanaka.	He is Nakada although I am Tanaka.
63	@1@彼は中田で す。	彼は中田です。	He is Nakada	
64	この処理は@2@ 以下のように行い ます。	この処理は以下の ように行います。	This processing is performed as follows.	This processing@2@ It carries out as follows.

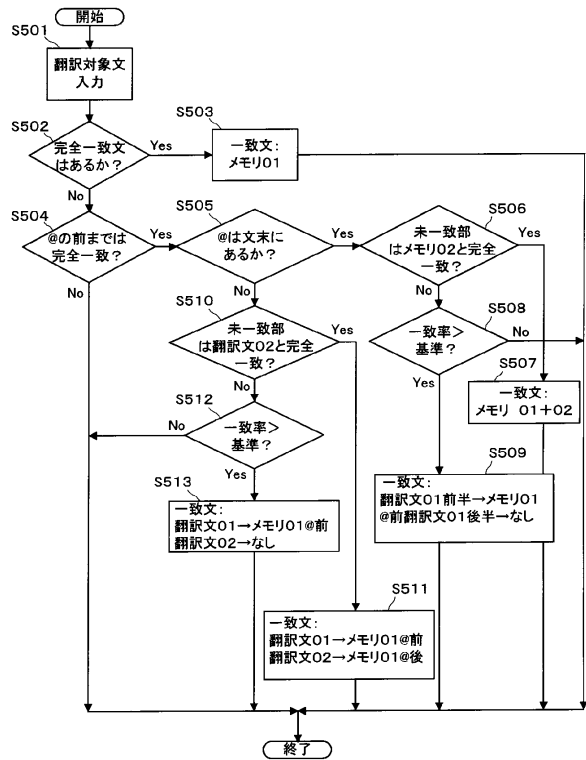
【図 8】

	原文	訳文
81	わたしはたなかです	I am Tanaka.
82	私は田中ですが、 @1@	I am Tanaka.
83	@1@彼は中田で す。	He is Nakada
84	この処理は@2@ 以下のように行い ます。	This processing@2@ is performed as follows.

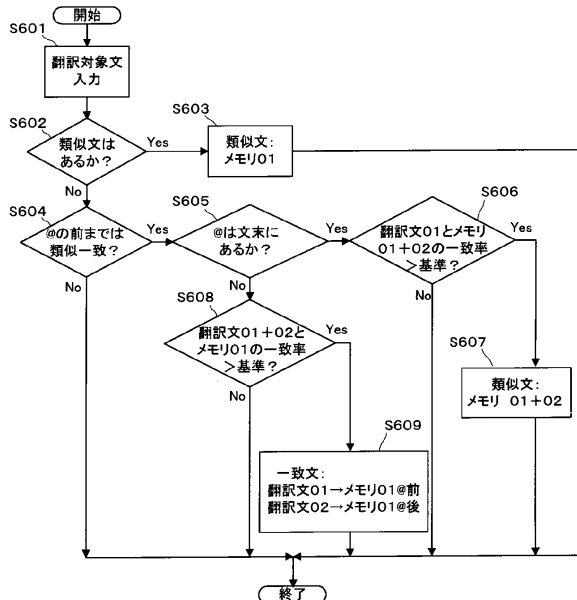
【図 9】

翻訳対象文
90 わたしはたなかです
91 私は田中ですが、彼は中田です。
92 私は田中ですが、生まれは千葉です。
93 この処理は
94 以下のように行います。
95 この処理は
96 明日行う予定です。
97 私は田中ですが、かれは中田です。
98 あの処理は、
99 以下のように行います。

【図 10】



【図 11】



フロントページの続き

審査官 和田 財太

(56)参考文献 特開平06 - 068142 (JP, A)

(58)調査した分野(Int.Cl., DB名)

G06F 17/27 - 17/28