

(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2014-56492

(P2014-56492A)

(43) 公開日 平成26年3月27日(2014.3.27)

(51) Int.Cl.	F I	テーマコード (参考)
G06F 17/28 (2006.01)	G06F 17/28 Z	5B091
G06F 17/30 (2006.01)	G06F 17/30 170A	
	G06F 17/30 350C	
	G06F 17/30 340B	

審査請求 未請求 請求項の数 10 O L (全 14 頁)

(21) 出願番号 特願2012-201747 (P2012-201747)
 (22) 出願日 平成24年9月13日 (2012.9.13)

(71) 出願人 000004237
 日本電気株式会社
 東京都港区芝五丁目7番1号
 (74) 代理人 100109313
 弁理士 机 昌彦
 (74) 代理人 100124154
 弁理士 下坂 直樹
 (72) 発明者 三浦 貢
 東京都港区芝五丁目7番1号
 日本電気株式会社内
 Fターム(参考) 5B091 BA00 CC00 DA00 EA17

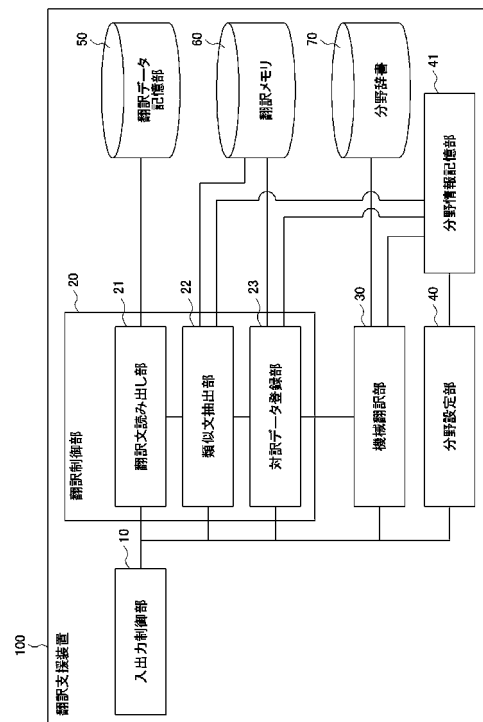
(54) 【発明の名称】 翻訳支援装置、翻訳支援方法および翻訳支援プログラム

(57) 【要約】

【課題】 収集された対訳データから利用する対訳データを選択する際の精度を向上することができる翻訳支援装置、翻訳支援方法および翻訳支援プログラムを提供する。

【解決手段】 翻訳支援装置は、原文と当該原文の翻訳文との組み合わせである対訳データを格納する翻訳メモリと、翻訳対象文の翻訳に利用される候補となる前記対訳データを前記翻訳メモリから抽出する類似文抽出部と、前記類似文抽出部により抽出された対訳データのうち、前記類似文抽出部により抽出された対訳データのうち、前記翻訳対象文の翻訳に利用された対訳データに対して、予め設定された分野を表す情報を関連付けると共に、当該分野を表す情報を関連付けた対訳データを前記翻訳メモリに登録する対訳データ登録部とを備え、前記類似文抽出部は、前記抽出した対訳データのうち、当該対訳データに関連付けられた分野を表す情報が、前記予め設定された分野と一致する対訳データを、前記翻訳対象文の翻訳に利用することに適した対訳データとして順位付ける。

【選択図】 図1



【特許請求の範囲】

【請求項 1】

原文と当該原文の翻訳文との組み合わせである対訳データを格納する翻訳メモリと、
翻訳対象文の翻訳に利用される候補となる前記対訳データを前記翻訳メモリから抽出する類似文抽出部と、

前記類似文抽出部により抽出された対訳データのうち、前記翻訳対象文の翻訳に利用された対訳データに対して、予め設定された分野を表す情報を関連付けると共に、当該分野を表す情報を関連付けた対訳データを前記翻訳メモリに登録する対訳データ登録部とを備え、

前記類似文抽出部は、前記抽出した対訳データのうち、当該対訳データに関連付けられた分野を表す情報が、前記予め設定された分野と一致する対訳データを、前記翻訳対象文の翻訳に利用することに適した対訳データとして順位付ける

翻訳支援装置。

【請求項 2】

前記対訳データ登録部は、前記翻訳対象文の翻訳に利用された対訳データに対して、前記予め設定された分野の翻訳に利用された回数を、前記分野を表す情報として関連付ける請求項 1 記載の翻訳支援装置。

【請求項 3】

前記類似文抽出部により抽出された対訳データのいずれも翻訳に利用しない前記翻訳対象文の機械翻訳を行う機械翻訳部をさらに備え、

前記対訳データ登録部は、前記翻訳対象文と、前記機械翻訳部により取得された当該翻訳対象文の翻訳文との組み合わせに、前記予め設定された分野を表す情報を関連付けて、前記翻訳メモリに登録する

請求項 1 または請求項 2 記載の翻訳支援装置。

【請求項 4】

前記対訳データ登録部は、前記翻訳対象文の翻訳に利用された対訳データに対して、前記予め設定された分野の翻訳に利用された回数を、前記分野を表す情報として関連付け、

前記類似文抽出部は、前記抽出した対訳データのうち、前記予め設定された分野の翻訳に利用された回数が最も多い対訳データを、前記翻訳対象文の翻訳に利用することに適した対訳データとして順位付ける

請求項 3 記載の翻訳支援装置。

【請求項 5】

原文と当該原文の翻訳文との組み合わせである対訳データを格納する翻訳メモリから、翻訳対象文の翻訳に利用される候補となる前記対訳データを抽出し、

前記抽出された対訳データのうち、前記翻訳対象文の翻訳に利用された対訳データに対して、予め設定された分野を表す情報を関連付けると共に、当該分野を表す情報を関連付けた対訳データを前記翻訳メモリに登録し、

前記対訳データを抽出するに際し、前記抽出した対訳データのうち、当該対訳データに関連付けられた分野を表す情報が、前記予め設定された分野と一致する対訳データを、前記翻訳対象文の翻訳に利用することに適した対訳データとして順位付ける

翻訳支援方法。

【請求項 6】

前記対訳データを前記翻訳メモリに登録するに際し、前記翻訳対象文の翻訳に利用された対訳データに対して、前記予め設定された分野の翻訳に利用された回数を、前記分野を表す情報として関連付ける

請求項 5 記載の翻訳支援方法。

【請求項 7】

前記抽出された対訳データのいずれも翻訳に利用しない前記翻訳対象文の機械翻訳を行い、

前記対訳データを前記翻訳メモリに登録するに際し、前記翻訳対象文と、前記機械翻訳

10

20

30

40

50

により取得された当該翻訳対象文の翻訳文との組み合わせに、前記予め設定された分野を表す情報を関連付けて、前記翻訳メモリに登録する

請求項 5 または請求項 6 記載の翻訳支援方法。

【請求項 8】

前記対訳データを前記翻訳メモリに登録するに際し、前記翻訳対象文の翻訳に利用された対訳データに対して、前記予め設定された分野の翻訳に利用された回数を、前記分野を表す情報として関連付け、

前記対訳データを抽出するに際し、前記抽出した対訳データのうち、前記予め設定された分野の翻訳に利用された回数が最も多い対訳データを、前記翻訳対象文の翻訳に利用することに適した対訳データとして順位付ける

10

請求項 7 記載の翻訳支援方法。

【請求項 9】

原文と当該原文の翻訳文との組み合わせである対訳データを格納する翻訳メモリから、翻訳対象文の翻訳に利用される候補となる前記対訳データを抽出する処理と、

前記抽出された対訳データのうち、前記翻訳対象文の翻訳に利用された対訳データに対して、予め設定された分野を表す情報を関連付けると共に、当該分野を表す情報を関連付けた対訳データを前記翻訳メモリに登録する処理と、

前記対訳データを抽出するに際し、前記抽出した対訳データのうち、当該対訳データに関連付けられた分野を表す情報が、前記予め設定された分野と一致する対訳データを、前記翻訳対象文の翻訳に利用することに適した対訳データとして順位付ける処理とを、

20

コンピュータに実行させる翻訳支援プログラム。

【請求項 10】

前記対訳データを前記翻訳メモリに登録するに際し、前記翻訳対象文の翻訳に利用された対訳データに対して、前記予め設定された分野の翻訳に利用された回数を、前記分野を表す情報として関連付ける処理を、

コンピュータに実行させる請求項 9 記載の翻訳支援プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、翻訳支援装置、翻訳支援方法および翻訳支援プログラムに関する。

30

【背景技術】

【0002】

過去に翻訳者が翻訳した結果を再利用することにより、翻訳作業を効率化する翻訳支援装置がある。翻訳支援装置は、翻訳者が翻訳した結果を利用するので、機械翻訳と比較して、品質の高い翻訳を実施することが可能である。一方、翻訳者が翻訳した文が事前に翻訳支援装置に登録されていない場合は、翻訳ができないという問題がある。

【0003】

上記問題に対処するために、翻訳メモリ、すなわち原文とその翻訳文とを一对とする対訳データを広範囲に収集し、その対訳データの集まりを翻訳支援装置において利用することが考えられる。しかしながら、当該対処法では、対訳データが広範囲に収集されるものであるため、その対訳データに含まれる翻訳文が、原文の翻訳文として適切であるか否かの判断が困難になるという問題がある。適切な判断なしに収集された対訳データを利用することは、翻訳の質の低下につながるという問題がある。

40

【0004】

さらに上記問題に対処する技術として、例えば、特許文献 1 は、機械翻訳装置を開示する。この機械翻訳装置は、各種の分野の文章が混在する原文を翻訳する際に、原文を指定した範囲に分け、その範囲の分野を推定することにより、分野に適した精度の高い翻訳を実施している。

【0005】

また、特許文献 2 は、複数の辞書データベースを用いる機械翻訳装置を開示する。この

50

機械翻訳装置は、対訳が記述され、対訳の内容に応じた分類コードが付与された対訳文書を用いて訳語対を分類し、その訳語対を分類ごとに辞書データベースに格納する。この構成により、該機械翻訳装置は、分類に沿った訳語を利用することで、翻訳の精度を向上している。

【先行技術文献】

【特許文献】

【0006】

【特許文献1】特開平5 - 257966号公報

【特許文献2】特開2008 - 083849号公報

【発明の概要】

【発明が解決しようとする課題】

【0007】

特許文献1に開示される機械翻訳装置では、使用する辞書に含まれる訳語のうち、翻訳対象文に関わる分野に適した訳語を使用することにより翻訳の精度を向上することはできる。しかしながら、該機械翻訳装置は、収集された対訳データから利用する対訳データを選択する際の精度を向上することはできないという課題がある。

【0008】

また、特許文献2に開示される機械翻訳装置では、対訳文書に基づいて辞書を分類することにより、使用する辞書に含まれる訳語の精度を向上することはできる。しかしながら、該機械翻訳装置は、上記と同様に、収集された対訳データから利用する対訳データを選択する際の精度を向上することはできないという課題がある。

【0009】

本願発明は、上記課題を鑑みてなされたものであり、収集された対訳データから利用する対訳データを選択する際の精度を向上することができる翻訳支援装置、翻訳支援方法および翻訳支援プログラムを提供することを主要な目的とする。

【課題を解決するための手段】

【0010】

本発明に係る第1の翻訳支援装置は、原文と当該原文の翻訳文との組み合わせである対訳データを格納する翻訳メモリと、翻訳対象文の翻訳に利用される候補となる前記対訳データを前記翻訳メモリから抽出する類似文抽出部と、前記類似文抽出部により抽出された対訳データのうち、前記翻訳対象文の翻訳に利用された対訳データに対して、予め設定された分野を表す情報を関連付けると共に、当該分野を表す情報を関連付けた対訳データを前記翻訳メモリに登録する対訳データ登録部とを備え、前記類似文抽出部は、前記抽出した対訳データのうち、当該対訳データに関連付けられた分野を表す情報が、前記予め設定された分野と一致する対訳データを、前記翻訳対象文の翻訳に利用することに適した対訳データとして順位付ける。

【0011】

本発明に係る第1の翻訳支援方法は、原文と当該原文の翻訳文との組み合わせである対訳データを格納する翻訳メモリから、翻訳対象文の翻訳に利用される候補となる前記対訳データを抽出し、前記抽出された対訳データのうち、前記翻訳対象文の翻訳に利用された対訳データに対して、予め設定された分野を表す情報を関連付けると共に、当該分野を表す情報を関連付けた対訳データを前記翻訳メモリに登録し、前記対訳データを抽出するに際し、前記抽出した対訳データのうち、当該対訳データに関連付けられた分野を表す情報が、前記予め設定された分野と一致する対訳データを、前記翻訳対象文の翻訳に利用することに適した対訳データとして順位付ける。

本発明に係る第1の翻訳支援プログラムは、原文と当該原文の翻訳文との組み合わせである対訳データを格納する翻訳メモリから、翻訳対象文の翻訳に利用される候補となる前記対訳データを抽出する処理と、前記抽出された対訳データのうち、前記翻訳対象文の翻訳に利用された対訳データに対して、予め設定された分野を表す情報を関連付けると共に、当該分野を表す情報を関連付けた対訳データを前記翻訳メモリに登録する処理と、前記対

10

20

30

40

50

訳データを抽出するに際し、前記抽出した対訳データのうち、当該対訳データに関連付けられた分野を表す情報が、前記予め設定された分野と一致する対訳データを、前記翻訳対象文の翻訳に利用することに適した対訳データとして順位付ける処理とを、コンピュータに実行させる。

【0012】

なお同目的は、上記の各構成を有する翻訳支援装置および翻訳支援方法を、コンピュータによって実現するコンピュータ・プログラム、およびそのコンピュータ・プログラムが格納されている、コンピュータ読み取り可能な記憶媒体によっても達成される。

【発明の効果】

【0013】

本願発明によれば、収集された対訳データから利用する対訳データを選択する際の精度を向上することができる効果が得られる。

【図面の簡単な説明】

【0014】

【図1】本発明の第1の実施形態に係る翻訳支援装置の構成を示すブロック図である。

【図2】本発明の第1の実施形態に係る翻訳支援装置のハードウェア構成を示す図である。

【図3】本発明の第1の実施形態に係る翻訳支援装置の分野設定部が利用者に提示する分野設定の画面の一例を示す図である。

【図4】本発明の第1の実施形態に係る翻訳支援装置の分野辞書に格納される辞書データの一例を示す図である。

【図5】本発明の第1の実施形態に係る翻訳支援装置の翻訳メモリに格納される対訳データの一例を示す図である。

【図6】本発明の第1の実施形態に係る翻訳支援装置の翻訳メモリに格納される対訳データの他の例を示す図である。

【図7】本発明の第1の実施形態に係る翻訳支援装置の翻訳制御部の動作を説明するフローチャートである。

【図8】本発明の第2の実施形態に係る翻訳支援装置の構成を示すブロック図である。

【発明を実施するための形態】

【0015】

第1の実施形態

本発明の実施形態について図面を参照して詳細に説明する。

【0016】

図1は、本発明の第1の実施形態に係る翻訳支援装置100の構成を示すブロック図である。図1に示すように、翻訳支援装置100は、入出力制御部10、翻訳制御部20、機械翻訳部30、分野設定部40、分野情報記憶部41、翻訳データ記憶部50、翻訳メモリ60および分野辞書70を備える。翻訳制御部20は、翻訳文読み出し部21、類似文抽出部22、対訳データ登録部23を備える。

【0017】

各構成要素の概略について説明する（詳細は後述する）。

【0018】

入出力制御部10は、利用者からデータの入力を受け付けたり、利用者に対してデータを提示したりする。翻訳制御部20は、翻訳文読み出し部21、類似文抽出部22および対訳データ登録部23において、過去に翻訳された結果を用いながら、翻訳データ記憶部50に格納される翻訳対象文書の翻訳を実施する。

【0019】

翻訳データ記憶部50は、翻訳の対象となる翻訳対象文書を格納する。翻訳メモリ60は、過去に翻訳された結果である対訳データを多数格納する。分野辞書70は、見出し語（単語）と、その語の分野ごとに適した訳語とを含む辞書データを格納する。

【0020】

10

20

30

40

50

機械翻訳部 30 は、利用者からの指示に回答して、翻訳対象文の機械翻訳を実施する。分野設定部 40 は、利用者から翻訳文の分野の設定を受け付けると共に、受け付けた分野を表す情報を分野情報記憶部 41 に格納する。翻訳支援装置 100 の利用者は、機械翻訳を実施する際にその精度を上げるために、翻訳文の分野を分野設定部 40 を介して予め設定しておく。

【0021】

図 1 の翻訳支援装置 100 の入出力制御部 10、翻訳制御部 20、機械翻訳部 30、分野設定部 40 は、コンピュータにより実現した場合、図 2 に例示するハードウェア構成を有する。図 2 に示す構成は、CPU (Central Processing Unit) 200、メモリ等の記憶媒体 201 および記憶媒体 201 に含まれるプログラム 202 を備える。CPU 200 は、各種ソフトウェア・プログラム (コンピュータ・プログラム) を実行することにより、翻訳支援装置 100 の全体的な動作を司る。本実施形態および以下に示す他の実施形態において、CPU 200 が、メモリ等の記憶媒体 201 を適宜参照しながら、以下に示す翻訳支援装置 100 が備える各機能 (各部) のソフトウェア・プログラムを実行する。

10

【0022】

より具体的には、CPU 200 は、メモリ等の記憶媒体 201 を適宜参照しながら、翻訳支援装置 100 が備える機能を実施するソフトウェア・プログラムを実行することにより、入出力制御部 10、翻訳制御部 20、機械翻訳部 30 および分野設定部 40 等のソフトウェア・プログラムを実行する。

20

【0023】

次に、翻訳の手順について具体的に説明する。

【0024】

翻訳支援装置 100 の利用者は、まず、入出力制御部 10 を介して分野設定部 40 に対し、翻訳文の分野を設定する。ここで分野を設定するのは、機械翻訳部 30 が機械翻訳を実施する際に使用する分野辞書 70 の分野を設定するためである。図 3 は、分野設定部 40 が入出力制御部 10 を介して利用者に提示する分野設定の画面の一例を示す図である。図 3 に示すように、各種分野の一覧を表示するとともにいずれかの分野の選択を受け付けるようにしてもよい。利用者は、入出力制御部 10 を介して分野を選択すると、分野設定部 40 は選択された分野を分野情報記憶部 41 に格納しておく。

30

【0025】

図 4 は、分野辞書 70 に格納される辞書データの一例を示す図である。図 4 に示すように、辞書データは、見出し語 (単語) と、その語の分野ごとに適した訳語とを有する。例えば、「operation」という単語は、医学分野の翻訳で使用される場合は「手術」という訳語が用いられ、経済分野の翻訳で使用される場合は「操作」という訳語が用いられる。

【0026】

機械翻訳部 30 は、翻訳対象文の機械翻訳を実施する指示を受けると、分野情報記憶部 41 に格納された分野に関連付けられた分野辞書 70 の訳語を用いて、機械翻訳を実施する。このように、機械翻訳部 30 は機械翻訳の際に分野辞書 70 を用いることにより、機械翻訳の精度を向上する。

40

【0027】

図 5 は、翻訳メモリ 60 に格納される対訳データの一例を示す図である。図 5 に示すように、対訳データは、原文と、その原文を過去に翻訳者が翻訳したり、機械翻訳の結果を翻訳者が修正等したりした結果である翻訳文との、一对の組み合わせである。ここでは、日本語の原文と、それを英語に翻訳した翻訳文との組み合わせである対訳データを格納することを示す。翻訳メモリ 60 は、広範囲に収集された対訳データを多数格納している。

【0028】

図 6 は、翻訳メモリ 60 に格納される対訳データの他の例を示す図である。図 6 に示す対訳データには、原文とその翻訳文に加えて、該対訳データに含まれる翻訳文が過去の翻訳において利用された分野を表す情報が関連付けられている。ここでは、分野を表す情報として、例えば、各対訳データが過去の翻訳において利用された分野の回数を関連付ける

50

ことを示す。例えば、対訳データ 1 は、過去に、医学分野で 1 回、経済分野で 5 回、化学分野で 2 回利用されたことを示す。対訳データに分野を表す情報を関連付けることについて、詳細は後述する。

【0029】

図 7 は、翻訳制御部 20 による翻訳の動作を説明するフローチャートである。図 7 を参照して、翻訳制御部 20 による翻訳の動作について説明する。

【0030】

翻訳支援装置 100 の利用者は、上述のように分野を設定した後、入出力制御部 10 を介して翻訳を実施する指示を出す。入出力制御部 10 は、上記指示に応答して、翻訳文読み出し部 21 に翻訳開始を指示する。翻訳文読み出し部 21 は、上記指示に応答して、翻訳データ記憶部 50 から翻訳の対象となる翻訳文書を読み出すとともに、そのうちの 1 文を翻訳対象文として類似文抽出部 22 に通知する (ST ステップ 101)。

【0031】

類似文抽出部 22 は、受け取った翻訳対象文の翻訳に利用される候補となる対訳データを、翻訳メモリ 60 から抽出する (ST ステップ 102)。すなわち、類似文抽出部 22 は、翻訳対象文と一致または類似する原文を含む 1 または複数の対訳データを、候補として翻訳メモリ 60 から抽出すると共に、類似する順に順位付ける。翻訳対象文と一致または類似する原文を含む対訳データを抽出することは、例えば、特開平 7 - 160720 号公報に開示されるため、その詳細な説明は省略する。

【0032】

類似文抽出部 22 は、抽出した各対訳データに、分野情報が関連付けられているかどうかを判断する (ST ステップ 103)。いずれの対訳データにも分野情報が関連付けられていない場合、類似文抽出部 22 は、翻訳対象文と最も類似すると順位付けた原文の翻訳文を、入出力制御部 10 を介して利用者に表示する (ST ステップ 104)。

【0033】

一方、類似文抽出部 22 は、抽出した対訳データのいずれかに分野情報が関連付けられている場合は、対訳データの順位付けを変更する。すなわち、抽出した対訳データに関連付けられた分野を表す情報のうち、分野情報記憶部 41 に記憶される分野を表す情報と一致するものがあれば、その対訳データを最も類似する対訳データとして順位付ける。例えば図 6 に示すように、過去の翻訳に利用された分野の回数が分野情報として各対訳データに関連付けられている場合、類似文抽出部 22 は、分野情報記憶部 41 に記憶される分野の回数が最も多い対訳データを最も類似する対訳データとして順位付けてもよい。また、抽出した対訳データに関連付けられた分野を表す情報のうち、分野情報記憶部 41 に格納される分野を表す情報と一致するものがない場合、類似文抽出部 22 は、ST ステップ 104 と同様に最も類似すると判断した原文の翻訳文を表示してもよい。

【0034】

類似文抽出部 22 は、最も類似すると判断した対訳データに含まれる翻訳文を入出力制御部 10 を介して利用者に表示する (ST ステップ 105)。利用者は、表示された翻訳文を見て、それが対訳として利用できるか否かを判断し (ST ステップ 106)、その判断の結果を入出力制御部 10 に指示する。入出力制御部 10 は、利用者からの指示が、対訳として利用できない旨の指示であった場合、翻訳対象文を機械翻訳部 30 に通知する。機械翻訳部 30 は、受け取った翻訳対象文を、分野辞書 70 に格納される辞書データのうち分野情報記憶部 41 に記憶される分野の訳語を参照しながら、翻訳する (ST ステップ 107)。機械翻訳部 30 は、機械翻訳の結果を対訳データ登録部 23 に通知する。

【0035】

一方、ST ステップ 106 における利用者からの指示が、対訳として利用できる旨の指示であった場合、入出力制御部 10 は、対訳データ登録部 23 に対して当該対訳データの登録を指示する。

【0036】

対訳データ登録部 23 は、ステップ ST 107 における機械翻訳の結果を受け取ると、

10

20

30

40

50

その結果を翻訳メモリ60に登録する。このとき、対訳データ登録部23は、原文と、機械翻訳により取得された翻訳文とを含む対訳データと、分野情報記憶部41に記憶される分野の利用回数「1」とを、関連付けて、翻訳メモリ60に登録する(STステップ108)。

【0037】

一方、対訳データ登録部23は、ステップST106における対訳データを対訳として利用できる旨の指示を受けたときは、当該対訳データに、分野情報記憶部41に記憶される分野の利用回数を現在の値に「1」加算したものを翻訳メモリ60に格納する(STステップ108)。このように、対訳データ登録部23は、対訳データが当該分野において過去に利用された回数を対訳データに関連付けて保存し、類似文抽出部22は、過去に当該分野で使用された回数の多い対訳データを最も類似する翻訳文として選択するので、精度の高い翻訳文を利用者に提示することができる。

10

【0038】

対訳データ登録部23は、翻訳メモリ60に対して上記対訳データの登録が終了すると、その旨を翻訳文読み出し部21に通知する。翻訳文読み出し部21は、次の翻訳対象文が存在するか否かを判断する(STステップ109)。次の翻訳対象文が存在する場合、翻訳文読み出し部21は、ステップST101に処理を戻す。翻訳制御部20は、以降同様に、翻訳対象文がすべて翻訳されるまで、上記動作を繰り返す。一方、翻訳対象文が存在しない場合、翻訳文読み出し部21は、翻訳の終了を入出力制御部10を介して利用者に通知する。

20

【0039】

なお、利用者は、入出力制御部10を介して提示された翻訳文を適宜修正し、修正後の翻訳文と原文とを対訳データとして、対訳データ登録部23が翻訳メモリ60に登録してもよい。また、対訳データ登録部23は、上述のように翻訳が実施されたタイミングで一文ずつ翻訳メモリ60に登録してもよいし、全文の翻訳が終了したタイミングで一括して翻訳メモリ60に登録してもよい。

【0040】

以上のように、第1の実施形態1では、分野設定部40が利用者から設定された翻訳文の分野を表す情報を分野情報記憶部41に格納し、翻訳文読み出し部21は、翻訳データ記憶部50から翻訳対象文を読み出す。類似文抽出部22は、読み出された翻訳対象文の翻訳に利用される候補となる対訳データを翻訳メモリ60から抽出すると共に、翻訳対象文と原文とが類似する順に順位付ける。類似文抽出部22は、抽出した対訳データのうち、分野情報が関連付けられ、かつその分野が分野情報記憶部41に記憶される分野と一致するものを、最も類似する対訳データとして順位づける。対訳データ登録部23は、最も類似する対訳データを翻訳に利用すると判断された場合、当該対訳データに、分野情報記憶部41に記憶される分野を表す情報を関連付けて、翻訳メモリ60に格納する。この構成により、第1の実施形態によれば、翻訳に利用された分野を表す情報が対訳データに関連付けられると共に、今回の翻訳と分野が一致する対訳データが候補として提示されるので、収集された対訳データから利用する対訳データを選択する際の精度を向上することができる効果が得られる。

30

40

【0041】

また、抽出されたいずれの対訳データも翻訳に利用しない翻訳対象文の機械翻訳を、機械翻訳部30が実施する。そして、対訳データ登録部23は、機械翻訳部30による翻訳文と原文との組み合わせに、分野情報記憶部41に記憶される分野を関連付けた対訳データを翻訳メモリ60に登録する。この構成により、新たな対訳データに分野を関連付けて登録することができるので、次回以降の対訳データを選択する際の精度をさらに向上することができる効果が得られる。

【0042】

また、対訳データ登録部23は、分野を表す情報として、当該対訳データが過去の翻訳において利用された分野の回数を関連付け、類似文抽出部22は、分野情報記憶部41に

50

記憶される分野で過去に利用された回数が最も多い対訳データを、最も類似する対訳データとして順位付けるので、対訳データを選択する際の精度をさらに向上することができる効果が得られる。

第2の実施形態

図8は、本発明の第2の実施形態に係る翻訳支援装置300の構成を示すブロック図である。図8に示すように、翻訳支援装置300は、翻訳メモリ301、類似文抽出部302および対訳データ登録部303を備える。

【0043】

翻訳メモリ301は、原文と当該原文の翻訳文との組み合わせである対訳データを格納する。類似文抽出部302は、翻訳対象文の翻訳に利用される候補となる対訳データを翻訳メモリ301から抽出する。対訳データ登録部303は、類似文抽出部302により抽出された対訳データのうち、翻訳対象文の翻訳に利用された対訳データに対して、予め設定された分野を表す情報を関連付けると共に、当該分野を表す情報を関連付けた対訳データを翻訳メモリ301に登録する。また、類似文抽出部302は、抽出した対訳データのうち、当該対訳データに関連付けられた分野を表す情報が、予め設定された分野と一致する対訳データを、翻訳対象文の翻訳に利用することに適した対訳データとして順位付ける。

10

【0044】

翻訳メモリ301は、上記実施形態の翻訳メモリ60に相当する。類似文抽出部302および対訳データ登録部303は、それぞれ上記実施形態の類似文抽出部22および対訳データ登録部23に相当する。

20

【0045】

以上のように、第2の実施形態によれば、翻訳支援装置300は、上記構成を有するので、収集された対訳データから利用する対訳データを選択する際の精度を向上することができる効果が得られる。

【0046】

なお、上述した各実施形態では、図1に示す翻訳支援装置100の入出力制御部10、翻訳制御部20、機械翻訳部30、分野設定部40を、図2に示すCPU200が実行する一例として、ソフトウェア・プログラムによって実現する場合について説明した。しかしながら、図1に示す各ブロックに示す機能は、一部または全部を、ハードウェアとして

30

【0047】

また、上述した各実施形態を例に説明した本発明は、上述した翻訳支援装置に対して、その説明において参照したフローチャート(図7)の機能を実現可能なコンピュータ・プログラムを供給した後、そのコンピュータ・プログラムを、CPU200に読み出して実行することによって達成される。

【0048】

また、係る供給されたコンピュータ・プログラムは、読み書き可能なメモリ(一時記憶媒体)またはハードディスク装置等のコンピュータ読み取り可能な記憶デバイスに格納すればよい。そして、このような場合において、本発明は、係るコンピュータ・プログラムを表すコード或いは係るコンピュータ・プログラムを格納した記憶媒体によって構成されると捉えることができる。

40

【産業上の利用可能性】

【0049】

本発明は、例えば翻訳メモリを備えた翻訳支援装置に適用できる。

【0050】

上記の実施形態の一部または全部は、以下の付記のようにも記載されうるが、以下には限られない。

(付記1)

原文と当該原文の翻訳文との組み合わせである対訳データを格納する翻訳メモリと、

50

翻訳対象文の翻訳に利用される候補となる前記対訳データを前記翻訳メモリから抽出する類似文抽出部と、

前記類似文抽出部により抽出された対訳データのうち、前記翻訳対象文の翻訳に利用された対訳データに対して、予め設定された分野を表す情報を関連付けると共に、当該分野を表す情報を関連付けた対訳データを前記翻訳メモリに登録する対訳データ登録部とを備え、

前記類似文抽出部は、前記抽出した対訳データのうち、当該対訳データに関連付けられた分野を表す情報が、前記予め設定された分野と一致する対訳データを、前記翻訳対象文の翻訳に利用することに適した対訳データとして順位付ける

翻訳支援装置。

10

(付記 2)

前記対訳データ登録部は、前記翻訳対象文の翻訳に利用された対訳データに対して、前記予め設定された分野の翻訳に利用された回数を、前記分野を表す情報として関連付ける付記 1 記載の翻訳支援装置。

(付記 3)

前記類似文抽出部により抽出された対訳データのいずれも翻訳に利用しない前記翻訳対象文の機械翻訳を行う機械翻訳部をさらに備え、

前記対訳データ登録部は、前記翻訳対象文と、前記機械翻訳部により取得された当該翻訳対象文の翻訳文との組み合わせに、前記予め設定された分野を表す情報を関連付けて、前記翻訳メモリに登録する

20

付記 1 または付記 2 記載の翻訳支援装置。

(付記 4)

前記対訳データ登録部は、前記翻訳対象文の翻訳に利用された対訳データに対して、前記予め設定された分野の翻訳に利用された回数を、前記分野を表す情報として関連付け、

前記類似文抽出部は、前記抽出した対訳データのうち、前記予め設定された分野の翻訳に利用された回数が最も多い対訳データを、前記翻訳対象文の翻訳に利用することに適した対訳データとして順位付ける

付記 3 記載の翻訳支援装置。

(付記 5)

原文と当該原文の翻訳文との組み合わせである対訳データを格納する翻訳メモリから、翻訳対象文の翻訳に利用される候補となる前記対訳データを抽出し、

30

前記抽出された対訳データのうち、前記翻訳対象文の翻訳に利用された対訳データに対して、予め設定された分野を表す情報を関連付けると共に、当該分野を表す情報を関連付けた対訳データを前記翻訳メモリに登録し、

前記対訳データを抽出するに際し、前記抽出した対訳データのうち、当該対訳データに関連付けられた分野を表す情報が、前記予め設定された分野と一致する対訳データを、前記翻訳対象文の翻訳に利用することに適した対訳データとして順位付ける

翻訳支援方法。

(付記 6)

前記対訳データを前記翻訳メモリに登録するに際し、前記翻訳対象文の翻訳に利用された対訳データに対して、前記予め設定された分野の翻訳に利用された回数を、前記分野を表す情報として関連付ける

40

付記 5 記載の翻訳支援方法。

(付記 7)

前記抽出された対訳データのいずれも翻訳に利用しない前記翻訳対象文の機械翻訳を行い、

前記対訳データを前記翻訳メモリに登録するに際し、前記翻訳対象文と、前記機械翻訳により取得された当該翻訳対象文の翻訳文との組み合わせに、前記予め設定された分野を表す情報を関連付けて、前記翻訳メモリに登録する

付記 5 または付記 6 記載の翻訳支援方法。

50

(付記 8)

前記対訳データを前記翻訳メモリに登録するに際し、前記翻訳対象文の翻訳に利用された対訳データに対して、前記予め設定された分野の翻訳に利用された回数を、前記分野を表す情報として関連付け、

前記対訳データを抽出するに際し、前記抽出した対訳データのうち、前記予め設定された分野の翻訳に利用された回数が最も多い対訳データを、前記翻訳対象文の翻訳に利用することに適した対訳データとして順位付ける

付記 7 記載の翻訳支援方法。

(付記 9)

原文と当該原文の翻訳文との組み合わせである対訳データを格納する翻訳メモリから、
10 翻訳対象文の翻訳に利用される候補となる前記対訳データを抽出する処理と、

前記抽出された対訳データのうち、前記翻訳対象文の翻訳に利用された対訳データに対して、予め設定された分野を表す情報を関連付けると共に、当該分野を表す情報を関連付けた対訳データを前記翻訳メモリに登録する処理と、

前記対訳データを抽出するに際し、前記抽出した対訳データのうち、当該対訳データに関連付けられた分野が、前記予め設定された分野と一致する対訳データを、前記翻訳対象文の翻訳に利用することに適した対訳データとして順位付ける処理とを、

コンピュータに実行させる翻訳支援プログラム。

(付記 10)

前記対訳データを前記翻訳メモリに登録するに際し、前記翻訳対象文の翻訳に利用された対訳データに対して、前記予め設定された分野の翻訳に利用された回数を、前記分野を表す情報として関連付ける処理を、
20

コンピュータに実行させる付記 9 記載の翻訳支援プログラム。

(付記 11)

前記抽出された対訳データのいずれも翻訳に利用しない前記翻訳対象文の機械翻訳を行う処理と、

前記対訳データを前記翻訳メモリに登録するに際し、前記翻訳対象文と、前記機械翻訳により取得された当該翻訳対象文の翻訳文との組み合わせに、前記予め設定された分野を表す情報を関連付けて、前記翻訳メモリに登録する処理とを、

コンピュータに実行させる付記 9 または付記 10 記載の翻訳支援プログラム。
30

(付記 12)

前記対訳データを前記翻訳メモリに登録するに際し、前記翻訳対象文の翻訳に利用された対訳データに対して、前記予め設定された分野の翻訳に利用された回数を、前記分野を表す情報として関連付ける処理と、

前記対訳データを抽出するに際し、前記抽出した対訳データのうち、前記予め設定された分野の翻訳に利用された回数が最も多い対訳データを、前記翻訳対象文の翻訳に利用することに適した対訳データとして順位付ける処理とを、

コンピュータに実行させる付記 11 記載の翻訳支援プログラム。

【符号の説明】

【0051】

- 10 入出力制御部
- 20 翻訳制御部
- 21 翻訳文読み出し部
- 22 類似文抽出部
- 23 対訳データ登録部
- 30 機械翻訳部
- 40 分野設定部
- 41 分野情報記憶部
- 50 翻訳データ記憶部
- 60 翻訳メモリ

10

20

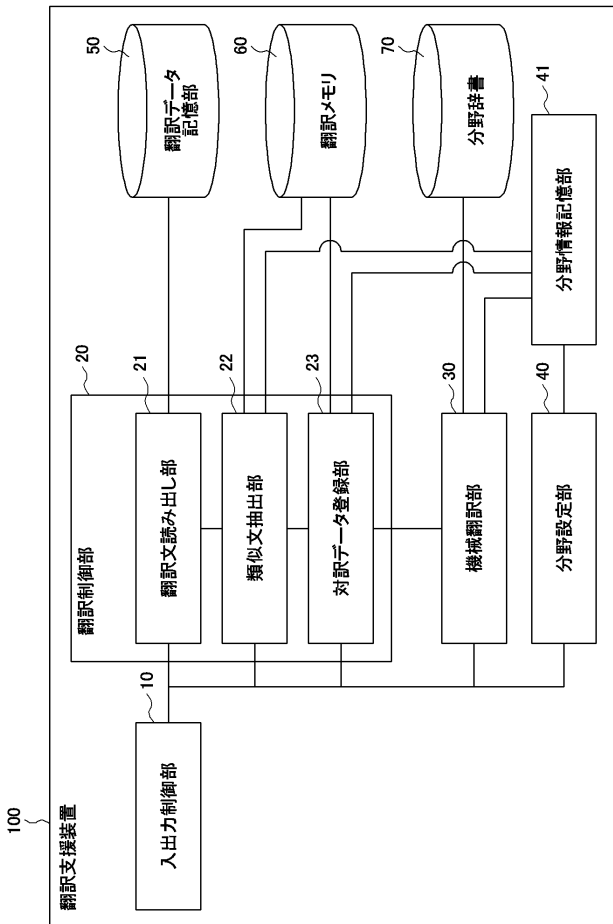
30

40

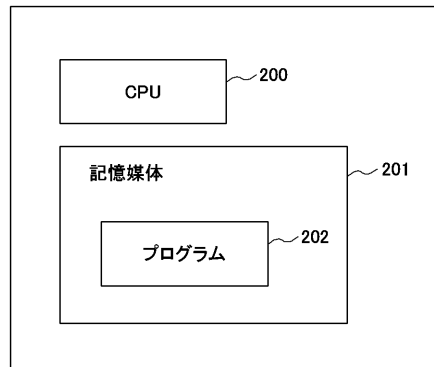
50

70 分野辞書

【図1】



【図2】



【 図 3 】

<input type="checkbox"/> 理学系統	<input type="checkbox"/> 化学	<input type="checkbox"/> 物理学	<input type="checkbox"/> 天文学
<input type="checkbox"/> 工学系統	<input type="checkbox"/> 植物学	<input type="checkbox"/> 生物学・バイオ	<input type="checkbox"/> 医学
<input type="checkbox"/> 建築土木	<input type="checkbox"/> 資源・エネルギー工学	<input type="checkbox"/> 発電	<input type="checkbox"/> 原子力
<input type="checkbox"/> 地下資源	<input type="checkbox"/> 金属工学	<input type="checkbox"/> 機械工学	<input type="checkbox"/> 自動車
<input type="checkbox"/> 鉄道工学	<input type="checkbox"/> 航空宇宙	<input type="checkbox"/> 電気工学	<input type="checkbox"/> 電子・通信・情報

【 図 5 】

対訳データ	
1	山田様自らあなたの履歴書をお届け下さり、あなたのことを高く評価していらっしゃいました。 He delivered your resume personally and recommended you highly.
2	貴殿の能力が正當に評価されたことを心からお祝い申し上げます。 I would like to extend my warmest congratulations to you on this well-deserved recognition of your ability.
3	貴社のプレゼンテーションが当方ではとても評価されている、ということをご報告しておきます。 This is just to let you know that your presentation went over very well with our people.
4	今までの地道な活動はきつと評価されること確信してしました。 I was certain that your steady record of Accomplishments would eventually be appreciated.

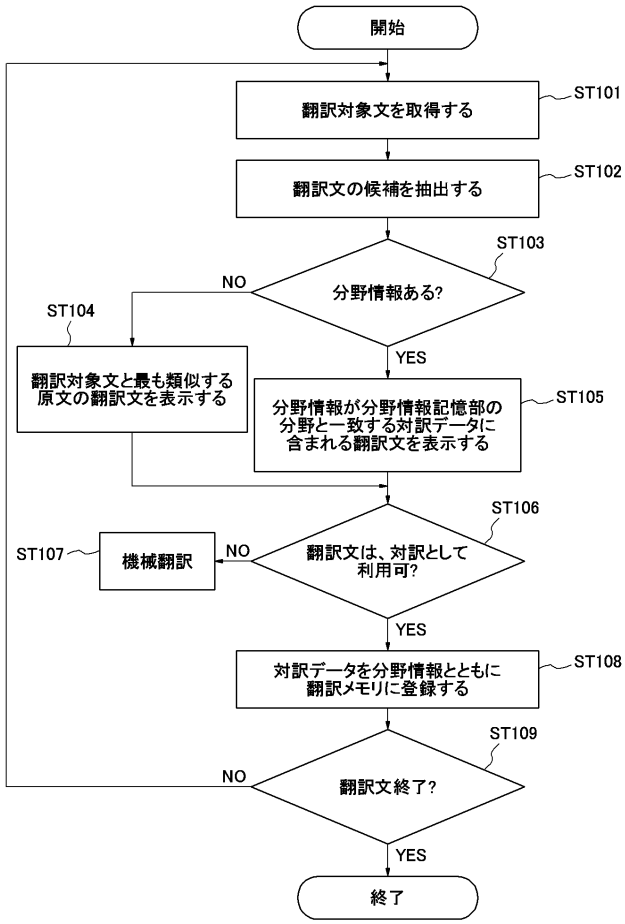
【 図 4 】



【 図 6 】

	対訳データ	分野		
		医学	経済	機械
1	山田様自らあなたの履歴書をお届け下さり、あなたのことを高く評価していらっしゃいました。 He delivered your resume personally and recommended you highly.	1	5	0
2	貴殿の能力が正當に評価されたことを心からお祝い申し上げます。 I would like to extend my warmest congratulations to you on this well-deserved recognition of your ability.	3	0	0
3	貴社のプレゼンテーションが当方ではとても評価されている、ということをご報告しておきます。 This is just to let you know that your presentation went over very well with our people.	2	0	1
4	今までの地道な活動はきつと評価されること確信してしました。 I was certain that your steady record of Accomplishments would eventually be appreciated.	1	0	0

【 図 7 】



【 図 8 】

