

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第3952009号
(P3952009)

(45) 発行日 平成19年8月1日(2007.8.1)

(24) 登録日 平成19年5月11日(2007.5.11)

(51) Int. Cl.

G06F 17/28 (2006.01)

F I

G06F 17/28

B

G06F 17/28

U

請求項の数 11 (全 20 頁)

(21) 出願番号 特願2003-403459 (P2003-403459)
 (22) 出願日 平成15年12月2日(2003.12.2)
 (65) 公開番号 特開2005-165652 (P2005-165652A)
 (43) 公開日 平成17年6月23日(2005.6.23)
 審査請求日 平成16年12月16日(2004.12.16)

(73) 特許権者 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 100103090
 弁理士 岩壁 冬樹
 (74) 代理人 100124501
 弁理士 塩川 誠人
 (72) 発明者 今野 清孝
 東京都港区芝五丁目7番1号 日本電気株
 式会社内

審査官 和田 財太

最終頁に続く

(54) 【発明の名称】 翻訳メモリシステム、翻訳方法、および翻訳メモリシステム用プログラム

(57) 【特許請求の範囲】

【請求項1】

第1の言語で記述される第1言語文字列を記憶する第1言語文字列記憶手段と、

第1言語文字列が第2の言語に翻訳された文字列である第2言語文字列を、当該第1言語文字列が記憶される第1言語文字列記憶手段の記憶領域と対応する記憶領域に記憶するとともに、文字列の表示態様としてフォントの種類、フォントサイズ、下線の有無、表示色の一部または全部を定めた文字属性情報を個々の第2言語文字列毎に記憶する第2言語文字列記憶手段と、

新規の第1言語文字列が入力される入力手段と、

前記新規の第1言語文字列と同一の第1言語文字列を第1言語文字列記憶手段が記憶する第1言語文字列の中から検索する検索手段と、

検索手段によって検索された第1言語文字列が記憶される第1言語文字列記憶手段の記憶領域と対応する記憶領域に記憶された第2言語文字列と、当該第2言語文字列の表示態様を定めた文字属性情報とを第2言語文字列記憶手段から抽出する抽出手段と、

抽出手段によって抽出された第2言語文字列および文字属性情報を出力する出力手段とを備えた

ことを特徴とする翻訳メモリシステム。

【請求項2】

入力手段には、新規の第1言語文字列と、当該第1言語文字列の表示態様を定めた文字属性情報とが入力され、

出力手段は、検出手段が前記新規の第 1 言語文字列と同一の第 1 言語文字列を検索できなかった場合に、前記新規の第 1 言語文字列および前記文字属性情報を出力する

請求項 1 に記載の翻訳メモリシステム。

【請求項 3】

第 1 言語文字列記憶手段は、個々の第 1 言語文字列を記憶した順番を特定可能に個々の第 1 言語文字列を記憶し、

検索手段は、新規の第 1 言語文字列と同一の第 1 言語文字列が第 1 言語文字列記憶手段の複数の記憶領域に記憶されている場合には、前記複数の記憶領域に記憶されている各第 1 言語文字列のうち、最後に記憶された第 1 言語文字列を検索する

請求項 1 または請求項 2 に記載の翻訳メモリシステム。

10

【請求項 4】

文字属性情報により定まる表示態様で文字列を表示する表示手段と、

表示手段が複数の文字列を表示した場合に、前記複数の文字列のうちの一つの文字列をオペレータから指定される指定受付手段とを備え、

第 1 言語文字列記憶手段は、文字属性情報を個々の第 1 言語文字列毎に記憶し、

検索手段は、新規の第 1 言語文字列と同一の第 1 言語文字列が第 1 言語文字列記憶手段の複数の記憶領域に記憶されている場合には、前記複数の記憶領域に記憶されている各第 1 言語文字列を、各第 1 言語文字列毎の文字属性情報により定まる表示態様で表示手段に表示させ、指定受付手段がオペレータから指定された第 1 言語文字列を、第 1 言語文字列記憶手段が記憶する第 1 言語文字列の中から検索する

20

請求項 1 または請求項 2 に記載の翻訳メモリシステム。

【請求項 5】

第 1 言語文字列記憶手段は、文字属性情報を個々の第 1 言語文字列毎に記憶し、

入力手段には、新規の第 1 言語文字列と、当該第 1 言語文字列の表示態様を定めた文字属性情報とが入力され、

検索手段は、前記新規の第 1 言語文字列と同一の第 1 言語文字列であって、文字属性情報が入力手段に入力された文字属性情報と同一である第 1 言語文字列を検索する

請求項 1 または請求項 2 に記載の翻訳メモリシステム。

【請求項 6】

入力手段には、文字属性情報により表示態様が定められた新規の第 1 言語文字列を記述したファイルが入力される請求項 1 から請求項 5 のうちのいずれか 1 項に記載の翻訳メモリシステム。

30

【請求項 7】

入力手段は、入力されたファイルのコピーを作成し、入力されたファイルから新規の第 1 言語文字列を読み込み、

検索手段は、入力手段がファイルから読み込んだ新規の第 1 言語文字列と同一の第 1 言語文字列を第 1 言語文字列記憶手段が記憶する第 1 言語文字列の中から検索し、

抽出手段は、検索された第 1 言語文字列が記憶される第 1 言語文字列記憶手段の記憶領域と対応する記憶領域に記憶された第 2 言語文字列と、当該第 2 言語文字列の表示態様を定めた文字属性情報とを第 2 言語文字列記憶手段から抽出し、

40

出力手段は、入力手段によってコピーされたファイルにおける前記新規の文字列が記述された位置で、前記新規の第 1 言語文字列を、抽出手段によって抽出された文字属性情報により表示態様が定められた第 2 言語文字列に置換するようにして、前記コピーされたファイルに前記第 2 言語文字列を記述する

請求項 6 に記載の翻訳メモリシステム。

【請求項 8】

入力手段には、文字属性情報により表示態様が定められた新規の第 1 言語文字列を記述するとともに画像データを含んでいるファイルが入力される請求項 7 に記載の翻訳メモリシステム。

【請求項 9】

50

入力手段は、入力されたファイルから新規の第 1 言語文字列を読み込み、

検索手段は、入力手段がファイルから読み込んだ新規の第 1 言語文字列と同一の第 1 言語文字列を第 1 言語文字列記憶手段が記憶する第 1 言語文字列の中から検索し、

抽出手段は、検索された第 1 言語文字列が記憶される第 1 言語文字列記憶手段の記憶領域と対応する記憶領域に記憶された第 2 言語文字列と、当該第 2 言語文字列の表示態様を定めた文字属性情報とを第 2 言語文字列記憶手段から抽出し、

出力手段は、入力手段に入力されたファイルとは異なるファイルに、抽出手段によって抽出された文字属性情報により表示態様が定められた第 2 言語文字列を記述する

請求項 6 に記載の翻訳メモリシステム。

【請求項 10】

10

第 1 の言語で記述される第 1 言語文字列を記憶する第 1 言語文字列記憶手段と、第 1 言語文字列が第 2 の言語に翻訳された文字列である第 2 言語文字列を、当該第 1 言語文字列が記憶される第 1 言語文字列記憶手段の記憶領域と対応する記憶領域に記憶するとともに、文字列の表示態様としてフォントの種類、フォントサイズ、下線の有無、表示色の一部または全部を定めた文字属性情報を個々の第 2 言語文字列毎に記憶する第 2 言語文字列記憶手段と、新規の第 1 言語文字列を記述したファイルが入力される入力手段とを備えた翻訳メモリシステムに適用される翻訳方法であって、

入力手段が、入力されたファイルのコピーを作成し、入力されたファイルから新規の第 1 言語文字列を読み込み、

検索手段が、前記新規の第 1 言語文字列と同一の第 1 言語文字列を第 1 言語文字列記憶手段が記憶する第 1 言語文字列の中から検索し、

20

抽出手段が、検索手段によって検索された第 1 言語文字列が記憶される第 1 言語文字列記憶手段の記憶領域と対応する記憶領域に記憶された第 2 言語文字列と、当該第 2 言語文字列の表示態様を定めた文字属性情報とを第 2 言語文字列記憶手段から抽出し、

出力手段が、入力手段によってコピーされたファイルにおける前記新規の第 1 言語文字列が記述された位置で、前記新規の第 1 言語文字列を、抽出手段によって抽出された文字属性情報により表示態様が定められた第 2 言語文字列に置換するようにして、前記コピーされたファイルに前記第 2 言語文字列を記述する

ことを特徴とする翻訳方法。

【請求項 11】

30

第 1 の言語で記述される第 1 言語文字列を記憶する第 1 言語文字列記憶手段と、第 1 言語文字列が第 2 の言語に翻訳された文字列である第 2 言語文字列を、当該第 1 言語文字列が記憶される第 1 言語文字列記憶手段の記憶領域と対応する記憶領域に記憶するとともに、文字列の表示態様としてフォントの種類、フォントサイズ、下線の有無、表示色の一部または全部を定めた文字属性情報を個々の第 2 言語文字列毎に記憶する第 2 言語文字列記憶手段と、新規の第 1 言語文字列を記述したファイルが入力される入力手段とを備えたコンピュータに、

入力手段に入力されたファイルのコピーを作成し、入力されたファイルから新規の第 1 言語文字列を読み込む処理、

前記新規の第 1 言語文字列と同一の第 1 言語文字列を第 1 言語文字列記憶手段が記憶する第 1 言語文字列の中から検索する処理、

40

検索された第 1 言語文字列が記憶される第 1 言語文字列記憶手段の記憶領域と対応する記憶領域に記憶された第 2 言語文字列と、当該第 2 言語文字列の表示態様を定めた文字属性情報とを第 2 言語文字列記憶手段から抽出する処理、および

コピーされたファイルにおける前記新規の第 1 言語文字列が記述された位置で、前記新規の第 1 言語文字列を、抽出手段によって抽出された文字属性情報により表示態様が定められた第 2 言語文字列に置換するようにして、前記コピーされたファイルに前記第 2 言語文字列を記述する処理

を実行させるための翻訳メモリシステム用プログラム。

【発明の詳細な説明】

50

【技術分野】

【0001】

本発明は、翻訳メモリによって翻訳を行う翻訳メモリシステム、翻訳方法および翻訳メモリシステムに適用される翻訳メモリシステム用プログラムに関する。

【背景技術】

【0002】

翻訳メモリによって翻訳を行う翻訳メモリシステムが知られている。翻訳メモリとは、既に翻訳されている原文およびその訳文を一組にして記憶装置に記憶しておき、新規に入力された原文の訳文を、記憶装置の中から抽出する機能のことを指す。この機能によって翻訳を行うシステムを翻訳メモリシステムと呼ぶ。

10

【0003】

翻訳メモリシステムとして、例えば、特許文献1に記載された翻訳システムがある。特許文献1に記載された翻訳システムは、文章記録部に原文とその訳文とを記録している。そして、入力部から与えられる新規の原文が、文章記録部内に記録された原文と等しいか否かを判定し、等しい場合には、その原文（予め記録された原文）と対になる訳文を新規原文の翻訳結果として使用する。

【0004】

また、翻訳メモリによらずに翻訳を行う翻訳システムも知られている。そして、このような翻訳システムの中には、入力される原文に応じた表示態様で訳文を出力する翻訳システムもある（例えば、特許文献2参照）。特許文献2に記載された翻訳システムは、入力された原文を解析して目的言語の訳文に変換する。そして、変換結果に対して、原文の文字属性に応じた変換を行い表示する。このとき、例えば、『原文において「下線付き体」で表示されていた語は、訳文において「斜字体」で表示する』等の規則を定めた文字属性変換表を用いて、翻訳結果に対する文字属性の変換を行う。

20

【0005】

また、特許文献3には、原文のフォントサイズやキャラクタースタイルなどを記憶する原文イメージ保存装置が記載されている。この原文イメージ保存装置は、辞書、文法規則および知識ベースを備える機械翻訳装置とともに使用され、機械翻訳装置は、訳文を表示する際、原文イメージ保存装置が記憶する情報に基づいて、訳文の表示を行う。例えば、「royal」という語がゴシック体で表現されているという情報を原文イメージ保存装置が記憶していた場合には、「王室の」という訳語もゴシック体で表現される。

30

【0006】

【特許文献1】特開平5-135095号公報（第3-5頁）

【特許文献2】特開平2-139666号公報（第2頁、第2図、第3図）

【特許文献3】特開平3-9472号公報（第2-3頁、第1-4図）

【発明の開示】

【発明が解決しようとする課題】

【0007】

特許文献2, 3に記載されているように、翻訳メモリによらずに翻訳を行う翻訳システム（原文の解析により翻訳を行ったり、辞書、文法規則、知識ベース等により翻訳を行う翻訳システム）では、訳文を原文の文字属性に応じた表示態様で表示することが可能となっている。

40

【0008】

しかし、既に翻訳されている原文およびその訳文を一組にして記憶装置に記憶させておく従来の翻訳メモリシステムでは、記憶装置に記憶した訳文を抽出して出力するだけであるので、原文の文字属性に応じた表示態様の訳文を出力することができなかった。そのため、訳文を、原文の文字属性に応じた表示態様にするために、原文をディスプレイ装置に表示させたり、印刷したりして、原文の表示態様を参照しながら、オペレータが訳文の文字属性を変更しなければならなかった。したがって、オペレータの作業負担が大きく、また、オペレータのミスにより訳文を原文と完全に同一な表示態様で表示できない場合もあ

50

った。

【0009】

そこで、本発明は、原文に応じた表示態様で表示される訳文を得ることができる翻訳メモリシステム、翻訳方法、およびその翻訳メモリシステムに適用される翻訳メモリシステム用プログラムを提供することを目的とする。

【課題を解決するための手段】

【0010】

本発明による翻訳メモリシステムは、第1の言語で記述される第1言語文字列を記憶する第1言語文字列記憶手段と、第1言語文字列が第2の言語に翻訳された文字列である第2言語文字列を、当該第1言語文字列が記憶される第1言語文字列記憶手段の記憶領域と対応する記憶領域に記憶するとともに、文字列の表示態様としてフォントの種類、フォントサイズ、下線の有無、表示色の一部または全部を定めた文字属性情報を個々の第2言語文字列毎に記憶する第2言語文字列記憶手段と、新規の第1言語文字列が入力される入力手段と、新規の第1言語文字列と同一の第1言語文字列を第1言語文字列記憶手段が記憶する第1言語文字列の中から検索する検索手段と、検索手段によって検索された第1言語文字列が記憶される第1言語文字列記憶手段の記憶領域と対応する記憶領域に記憶された第2言語文字列と、当該第2言語文字列の表示態様を定めた文字属性情報とを第2言語文字列記憶手段から抽出する抽出手段と、抽出手段によって抽出された第2言語文字列および文字属性情報を出力する出力手段とを備えたことを特徴とする。

【0011】

入力手段には、新規の第1言語文字列と、当該第1言語文字列の表示態様を定めた文字属性情報とが入力され、出力手段が、検出手段が新規の第1言語文字列と同一の第1言語文字列を検索できなかった場合に、新規の第1言語文字列およびその文字属性情報を出力するように構成されていてもよい。そのような構成によれば、第1言語文字列を検索できず、その結果、第2言語文字列およびその文字属性情報を抽出できなかった場合に、第2言語文字列等の代わりに、新規の第1言語文字列およびその文字属性情報を出力することができる。

【0012】

例えば、第1言語文字列記憶手段が、個々の第1言語文字列を記憶した順番を特定可能に個々の第1言語文字列を記憶し、検索手段が、新規の第1言語文字列と同一の第1言語文字列が第1言語文字列記憶手段の複数の記憶領域に記憶されている場合には、複数の記憶領域に記憶されている各第1言語文字列のうち、最後に記憶された第1言語文字列を検索するように構成されていてもよい。

【0013】

また、例えば、文字属性情報により定まる表示態様で文字列を表示する表示手段と、表示手段が複数の文字列を表示した場合に、複数の文字列のうちの一つの文字列をオペレータから指定される指定受付手段とを備え、第1言語文字列記憶手段が、文字属性情報を個々の第1言語文字列毎に記憶し、検索手段が、新規の第1言語文字列と同一の第1言語文字列が第1言語文字列記憶手段の複数の記憶領域に記憶されている場合には、複数の記憶領域に記憶されている各第1言語文字列を、各第1言語文字列毎の文字属性情報により定まる表示態様で表示手段に表示させ、指定受付手段がオペレータから指定された第1言語文字列を、第1言語文字列記憶手段が記憶する第1言語文字列の中から検索するように構成されていてもよい。そのような構成によれば、各第1言語文字列を、各第1言語文字列毎の文字属性情報により定まる表示態様で表示手段に表示させ、オペレータに対し、オペレータが意図している検索対象の第1言語文字列の指定を促すことができる。

【0014】

また、例えば、第1言語文字列記憶手段が、文字属性情報を個々の第1言語文字列毎に記憶し、入力手段には、新規の第1言語文字列と、当該第1言語文字列の表示態様を定めた文字属性情報とが入力され、検索手段が、新規の第1言語文字列と同一の第1言語文字列であって、文字属性情報が入力手段に入力された文字属性情報と同一である第1言語文

10

20

30

40

50

字列を検索するように構成されていてもよい。

【0015】

入力手段には、文字属性情報により表示態様が定められた新規の第1言語文字列を記述したファイルが入力されてもよい。

【0016】

入力手段が、入力されたファイルのコピーを作成し、入力されたファイルから新規の第1言語文字列を読み込み、検索手段が、入力手段がファイルから読み込んだ新規の第1言語文字列と同一の第1言語文字列を第1言語文字列記憶手段が記憶する第1言語文字列の中から検索し、抽出手段が、検索された第1言語文字列が記憶される第1言語文字列記憶手段の記憶領域と対応する記憶領域に記憶された第2言語文字列と、当該第2言語文字列の表示態様を定めた文字属性情報とを第2言語文字列記憶手段から抽出し、出力手段が、入力手段によってコピーされたファイルにおける新規の文字列が記述された位置で、新規の第1言語文字列を、抽出手段によって抽出された文字属性情報により表示態様が定められた第2言語文字列に置換するようにして、コピーされたファイルに第2言語文字列を記述するように構成されていてもよい。そのような構成によれば、入力されるファイルを表示したときのレイアウトと、第2言語文字列が記述されるファイルを表示したときのレイアウトを揃えることができる。

10

【0017】

入力手段には、文字属性情報により表示態様が定められた新規の第1言語文字列を記述するとともに画像データを含んでいるファイルが入力されてもよい。画像データもコピーされるので、第2言語文字列が記述されるファイル内に、入力されたファイルに含まれる画像データをそのまま保持することができる。

20

【0018】

入力手段が、入力されたファイルから新規の第1言語文字列を読み込み、検索手段が、入力手段がファイルから読み込んだ新規の第1言語文字列と同一の第1言語文字列を第1言語文字列記憶手段が記憶する第1言語文字列の中から検索し、抽出手段が、検索された第1言語文字列が記憶される第1言語文字列記憶手段の記憶領域と対応する記憶領域に記憶された第2言語文字列と、当該第2言語文字列の表示態様を定めた文字属性情報とを第2言語文字列記憶手段から抽出し、出力手段が、入力手段に入力されたファイルとは異なるファイルに、抽出手段によって抽出された文字属性情報により表示態様が定められた第2言語文字列を記述するように構成されていてもよい。

30

【0019】

また、本発明による翻訳方法は、第1の言語で記述される第1言語文字列を記憶する第1言語文字列記憶手段と、第1言語文字列が第2の言語に翻訳された文字列である第2言語文字列を、当該第1言語文字列が記憶される第1言語文字列記憶手段の記憶領域と対応する記憶領域に記憶するとともに、文字列の表示態様としてフォントの種類、フォントサイズ、下線の有無、表示色の一部または全部を定めた文字属性情報を個々の第2言語文字列毎に記憶する第2言語文字列記憶手段と、新規の第1言語文字列を記述したファイルが入力される入力手段とを備えた翻訳メモリシステムに適用される翻訳方法であって、入力手段が、入力されたファイルのコピーを作成し、入力されたファイルから新規の第1言語文字列を読み込み、検索手段が、新規の第1言語文字列と同一の第1言語文字列を第1言語文字列記憶手段が記憶する第1言語文字列の中から検索し、抽出手段が、検索手段によって検索された第1言語文字列が記憶される第1言語文字列記憶手段の記憶領域と対応する記憶領域に記憶された第2言語文字列と、当該第2言語文字列の表示態様を定めた文字属性情報とを第2言語文字列記憶手段から抽出し、出力手段が、入力手段によってコピーされたファイルにおける前記新規の第1言語文字列が記述された位置で、前記新規の第1言語文字列を、抽出手段によって抽出された文字属性情報により表示態様が定められた第2言語文字列に置換するようにして、前記コピーされたファイルに前記第2言語文字列を記述することを特徴とする。

40

【0020】

50

また、本発明による翻訳メモリシステム用プログラムは、第 1 の言語で記述される第 1 言語文字列を記憶する第 1 言語文字列記憶手段と、第 1 言語文字列が第 2 の言語に翻訳された文字列である第 2 言語文字列を、当該第 1 言語文字列が記憶される第 1 言語文字列記憶手段の記憶領域と対応する記憶領域に記憶するとともに、文字列の表示態様としてフォントの種類、フォントサイズ、下線の有無、表示色の一部または全部を定めた文字属性情報を個々の第 2 言語文字列毎に記憶する第 2 言語文字列記憶手段と、新規の第 1 言語文字列を記述したファイルが入力される入力手段とを備えたコンピュータに、入力手段に入力されたファイルのコピーを作成し、入力されたファイルから新規の第 1 言語文字列を読み込む処理、新規の第 1 言語文字列と同一の第 1 言語文字列を第 1 言語文字列記憶手段が記憶する第 1 言語文字列の中から検索する処理、検索された第 1 言語文字列が記憶される第 1 言語文字列記憶手段の記憶領域と対応する記憶領域に記憶された第 2 言語文字列と、当該第 2 言語文字列の表示態様を定めた文字属性情報とを第 2 言語文字列記憶手段から抽出する処理、およびコピーされたファイルにおける前記新規の第 1 言語文字列が記述された位置で、前記新規の第 1 言語文字列を、抽出手段によって抽出された文字属性情報により表示態様が定められた第 2 言語文字列に置換するようにして、前記コピーされたファイルに前記第 2 言語文字列を記述する処理を実行させることを特徴とする。

10

【発明の効果】

【0021】

本発明によれば、第 2 言語文字列とともに、新規の第 1 言語文字列に応じて、その第 2 言語文字列の文字属性情報も出力させることができる。したがって、オペレータは、出力される第 2 言語文字列の属性を変更する作業を行う必要がなく、オペレータの作業負担は軽減される。

20

【発明を実施するための最良の形態】

【0022】

以下、本発明を実施するための最良の形態を、図面を参照して説明する。なお、以下の説明において、「文字属性情報」とは、文字列の表示態様を定めた情報である。文字属性情報は、例えば、フォントの種類、フォントサイズ、下線の有無等を定める。なお、ここに挙げたフォントの種類、フォントサイズ、下線の有無等は、文字属性情報が定める項目の例示であり、これら以外の項目について定めていてもよい。例えば、文字属性情報が、文字列の表示色を定めていてもよい。また、文字属性情報は、フォントの種類、フォントサイズ、下線の有無、表示色等、多くの項目について定めていることが好ましいが、例えば、例えば、フォントの種類だけというように、1 つの項目について定めるものであってもよい。

30

【0023】

また、以下の説明において、翻訳前の言語（第 1 の言語）で記述された文字列を原文と記す。原文を目的の言語（第 2 の言語）に翻訳した文字列を訳文と記す。

【0024】

図 1 は、本発明による翻訳メモリシステムの構成の例を示すブロック図である。翻訳メモリシステムは、データ処理装置 2 と、記憶装置 3 とを備える。さらに、データ処理装置 2 は、文字列取得部 2 1 と、文字列検索部 2 2 と、ファイル出力部 2 3 と、登録情報取得部 2 4 とを備える。また、記憶装置 3 は、原文記憶部 3 1 と、訳文記憶部 3 2 とを備える。

40

【0025】

原文記憶部 3 1 は、原文となる文字列、およびその文字列の表示態様を定めた文字属性情報を記憶する。訳文記憶部 3 2 は、訳文となる文字列、およびその文字列の表示態様を定めた文字属性情報を記憶する。本例では、原文記憶部 3 1 および訳文記憶部 3 2 は、1 つの文毎に、文字列および文字属性情報を記憶するものとする。なお、1 つの文を構成する各文字が全て同じ表示態様で表示されるとは限らない。例えば、1 つの文を構成する文字の一部のフォントが、他の文字のフォントと異なっていたり、一部の文字に下線が付されていたりすることがある。この場合、1 つの文を構成する各文字のうち、どの文字がど

50

のフォントで表されるのか、どの文字に下線が付されるのか等を文字属性情報として記憶する。

【 0 0 2 6 】

登録情報取得部 2 4 には、システム管理者によって作成された登録用文書ファイル 1 3 が入力される。登録用文書ファイルには、原文を記述した原文ファイル 1 3 a と、その原文の訳文を記述した訳文ファイル 1 3 b とが含まれる。原文ファイル 1 3 a には、文字属性情報によって表示態様が定められた原文が記述されている。同様に、訳文ファイル 1 3 b には、文字属性情報によって表示態様が定められた訳文が記述されている。対になる原文と訳文の組み合わせにおいて、例えば、原文と訳文の表示態様が同じになるように定められる。なお、表示態様が同じとは、フォントの種類、フォントサイズ、下線の有無等の文字属性が同じということである。ただし、対になる原文と訳文の組み合わせにおいて、原文と訳文の表示態様が異なるように定められていてもよい。

10

【 0 0 2 7 】

図 2 (a) は、原文ファイル 1 3 a に記述された文字列の例を示す説明図であり、図 2 (b) は、訳文ファイル 1 3 b に記述された文字列の例を示す説明図である。図 2 では、原文ファイル 1 3 a は日本語で記述され、訳文ファイル 1 3 b は英語で記述される場合の例を示している。

【 0 0 2 8 】

原文ファイルの第 1 文は、「雨時々晴れ。」という文である。この文を構成する文字列の文字属性情報は、文字「雨」および文字「晴れ。」がゴシック体であり、文字「時々」が太字のゴシック体、かつ斜字体であり、下線付きであるということを定めている。訳文ファイルの第 1 文は、「Showers with sunny intervals.」という文である。この文を構成する文字列の文字属性情報は、文字「with」が太字のゴシック体、かつ斜字体であり、下線付きであり、他の文字がゴシック体であることを定めている。原文と訳文の第 1 文同士を比較すると、原文の「時々」に相当する「with」は、原文の「時々」と同じ表示態様になるように定められ、他の文字は全てゴシック体で表示されるように定められている。したがって、第 1 文では、原文と訳文の表示態様が同じになるように定められている。

20

【 0 0 2 9 】

原文ファイルの第 2 文は、「晴れのち雨。」という文である。この文を構成する文字列の文字属性情報は、文字「晴れ」および文字「雨。」がゴシック体であり、文字「のち」が太字のゴシック体、かつ斜字体であり、下線付きであるということを定めている。訳文ファイルの第 2 文は、「Rain after sunshine.」という文である。この文を構成する文字列の文字属性情報は、文字「after」が太字のゴシック体、かつ斜字体であり、下線付きであり、他の文字が明朝体であることを定めている。原文と訳文の第 2 文同士を比較すると、原文の「のち」に相当する「after」は、原文の「のち」と同じ表示態様になるように定められ、他の文字は全て明朝体で表示されるように定められている。したがって、第 2 文では、原文と訳文の表示態様が一部異なるように定められている。

30

【 0 0 3 0 】

登録情報取得部 2 4 は、原文ファイル 1 3 a に含まれる文字列と、その文字列の表示態様を定める文字属性情報を原文記憶部 3 1 に記憶させる。また、登録情報取得部 2 4 は、原文の文字列および文字属性情報を記憶させた原文記憶部 3 1 内の記憶領域と対応する訳文記憶部 3 2 内の記憶領域に、その原文と対になる訳文の文字列および文字属性情報を記憶させる。既に説明したように、本例では、原文記憶部 3 1 および訳文記憶部 3 2 は、1 つの文毎に文字列および文字属性情報を記憶する。したがって、登録情報取得部 2 4 は、原文に含まれる 1 つの文を構成する文字列、およびその文字列の文字属性情報を原文記憶部 3 1 内の記憶領域に記憶させる。そして、その記憶領域と対応する訳文記憶部 3 2 内の記憶領域に、その文の訳文となる 1 つの文を構成する文字列、およびその文字列の文字属性情報を記憶させる。

40

【 0 0 3 1 】

なお、原文記憶部 3 1 および訳文記憶部 3 2 は、原文記憶部 3 1 内の記憶領域と、その

50

記憶領域に対応する訳文記憶部 3 2 の記憶領域とが隣り合うように構成されていてもよい。図 3 は、そのような構成における対応する記憶領域を示す説明図である。図 3 に示すように、原文記憶部 3 1 のある記憶領域に、原文に含まれる 1 つの文を構成する文字列、およびその文字列の文字属性情報が記憶される。そして、その隣りの記憶領域に、その文の訳文となる 1 つの文を構成する文字列、およびその文字列の文字属性情報が記憶される。その隣の記憶領域には、また原文の文字列および文字属性情報が記憶される。図 3 に示すように、原文記憶部 3 1 の記憶領域と、訳文記憶部 3 2 の記憶領域とが交互に並び、隣り合う 1 組の記憶領域に 1 対の原文および訳文の文字列などを記憶する構成であってもよい。

【 0 0 3 2 】

10

登録情報取得部 2 4 が、対になる原文および訳文を文字属性情報とともに原文記憶部 3 1 および訳文記憶部 3 2 に記憶させることによって、新規に入力される原文に対する訳文を訳文記憶部 3 2 から抽出することが可能になる。

【 0 0 3 3 】

文字列取得部 2 1 には、新規の文書ファイル 1 1 が入力される。新規の文書ファイル 1 1 は、登録情報取得部 2 4 が原文の文字列および文字属性情報と、訳文の文字列および文字属性情報とを記憶装置 3 に記憶させた後に入力されるファイルである。新規の文書ファイル 1 1 には、文字属性情報によって表示態様が定められた新規の原文が記述されている。文字列取得部 2 1 は、新規の文書ファイル 1 1 から、文字属性情報によって表示態様が定められた新規の原文を読み込むことにより、文字属性情報および新規の原文を取得する。

20

【 0 0 3 4 】

文字列検索部 2 2 は、新規の文書ファイル 1 1 に記述された原文に含まれる各文毎に、その文を構成する文字列と同一の文字列を、原文記憶部 3 1 が記憶する文字列の中から検索する。さらに、文字列検索部 2 2 は、検索した文字列の訳文となる文字列、およびその文字列の文字属性情報を訳文記憶部 3 2 から抽出する。

【 0 0 3 5 】

ファイル出力部 2 3 は、文字列検索部 2 2 によって抽出された文字列およびその文字列の文字属性情報を出力する。ここで、文字属性情報により表示態様が定められた文字列を出力することも、文字列および文字属性の出力態様の一つである。ファイル出力部 2 3 は、文字列検索部 2 2 によって抽出された文字属性情報により表示態様が定められた文字列を訳文ファイル 1 2 に記述（出力）する。なお、訳文ファイル 1 2 は、図 1 に示した記憶装置 3 に記憶されても、図示していない他の記憶装置に記憶されてもよい。

30

【 0 0 3 6 】

以下の説明では、文字列取得部 2 1 が、新規の文書ファイル 1 1 のコピーを作成するものとする。そして、ファイル出力部 2 3 が、コピーされたファイルにおける新規の原文が記述された位置で、その新規の原文を、文字列検索部 2 2 によって抽出された文字属性情報により表示態様が定められた訳文に置換するようにして、コピーされたファイルに訳文の文字列を記述するものとして説明する。この場合、訳文ファイル 1 2 は、最初に、新規の文書ファイル 1 1 のコピーとして作成されることになる。

40

【 0 0 3 7 】

文字列取得部 2 1、ファイル出力部 2 3 および登録情報取得部 2 4 は、例えば、プログラムにしたがって動作する CPU および情報記憶媒体に対してデータの読み書きを行うドライブ装置（フロッピー（登録商標）ディスクドライブ、CD-ROM ドライブ等）によって実現される。また、文字列検索部 2 2 は、例えば、プログラムにしたがって動作する CPU によって実現される。このプログラムは、例えば、RAM 等の記憶装置に記憶される。

【 0 0 3 8 】

このプログラムは、第 1 の言語で記述される第 1 言語文字列を記憶する第 1 言語文字列記憶手段と、第 1 言語文字列が第 2 の言語に翻訳された文字列である第 2 言語文字列を、

50

当該第 1 言語文字列が記憶される第 1 言語文字列記憶手段の記憶領域と対応する記憶領域に記憶するとともに、文字列の表示態様としてフォントの種類、フォントサイズ、下線の有無、表示色の一部または全部を定めた文字属性情報を個々の第 2 言語文字列毎に記憶する第 2 言語文字列記憶手段と、新規の第 1 言語文字列を記述したファイルが入力される入力手段とを備えたコンピュータに、入力手段に入力されたファイルのコピーを作成し、入力されたファイルから新規の第 1 言語文字列を読み込む処理、新規の第 1 言語文字列と同一の第 1 言語文字列を第 1 言語文字列記憶手段が記憶する第 1 言語文字列の中から検索する処理、検索された第 1 言語文字列が記憶される第 1 言語文字列記憶手段の記憶領域と対応する記憶領域に記憶された第 2 言語文字列と、当該第 2 言語文字列の表示態様を定めた文字属性情報とを第 2 言語文字列記憶手段から抽出する処理、およびコピーされたファイルにおける前記新規の第 1 言語文字列が記述された位置で、前記新規の第 1 言語文字列を、抽出手段によって抽出された文字属性情報により表示態様が定められた第 2 言語文字列に置換するようにして、前記コピーされたファイルに前記第 2 言語文字列を記述する処理を実行させるための翻訳メモリシステム用プログラムである。

10

【 0 0 3 9 】

次に、動作について説明する。

まず、登録情報取得部 2 4 が、登録用文書ファイル 1 3 に基づいて、対になる原文および訳文を記憶装置 3 に記憶させる動作について説明する。図 4 は、登録情報取得部 2 4 が、対になる原文および訳文を記憶装置 3 に記憶させる動作の処理経過の例を示すフローチャートである。

20

【 0 0 4 0 】

登録情報取得部 2 4 は、原文ファイル 1 3 a および訳文ファイル 1 3 b を含む登録用文書ファイル 1 3 が入力されると、原文ファイル 1 3 a から、文字列、およびその文字列の文字属性情報を取得する（ステップ S 1 1）。ステップ S 1 1 において、登録情報取得部 2 4 は、1 つの文毎に、文を構成する文字列、およびその文字列の文字属性情報を取得する。登録情報取得部 2 4 は、文の切れ目に用いられる記号の存在によって、文の切れ目を判定すればよい。

【 0 0 4 1 】

例えば、日本語の場合、文の切れ目には句点が用いられる。したがって、登録情報取得部 2 4 は、原文の最初の文字から句点までが 1 つの文になっていて、以降、句点の次の文字から次の句点までが 1 つの文になっていると判定すればよい。また、例えば、英語の場合、文の切れ目にはピリオドが用いられる。したがって、登録情報取得部 2 4 は、原文の最初の文字からピリオドまでが 1 つの文になっていて、以降、ピリオドの次の文字から次のピリオドまでが 1 つの文になっていると判定すればよい。ここでは、日本語と英語の場合を例に説明したが、他の言語の場合も同様である。

30

【 0 0 4 2 】

登録情報取得部 2 4 は、ステップ S 1 1 の後、訳文ファイル 1 3 b から、文字列、およびその文字列の文字属性情報を取得する（ステップ S 1 2）。ステップ S 1 2 においても、登録情報取得部 2 4 は、1 つの文毎に、文を構成する文字列、およびその文字列の文字属性情報を取得する。既に説明したように、登録情報取得部 2 4 は、文の切れ目に用いられる記号の存在によって、文の切れ目を判定すればよい。

40

【 0 0 4 3 】

ステップ S 1 1 , S 1 2 で取得したそれぞれの文字列は、必ず 1 つの文を構成しているとは限らない。例えば、原文や訳文の最初にタイトルを示す文字列が記述され、タイトルの後で改行され、文が始まる場合がある。この場合、タイトルの最初の文字から、句点やピリオド等の記号までが、1 つの文であると判定されることになる。しかし、この場合、タイトルの部分は、後に続く文とは分けて、1 つの文とみなすことが好ましい。また、英語の場合、ピリオドは、文の切れ目だけでなく、語の省略を表す場合にも用いられる。したがって、登録情報取得部 2 4 は、文の途中で用いられたピリオドの位置で文が終了したと判定してしまう場合がある。

50

【 0 0 4 4 】

そこで、ステップ S 1 1 , S 1 2 の後、登録情報取得部 2 4 は、ディスプレイ装置等の表示手段（図 1 において図示せず。）に、ステップ S 1 1 , S 1 2 で取得したそれぞれの文字列を表示させて、その文字列が原文および訳文として対になっていることの確認をオペレータに促す。そして、表示した文字列が原文および訳文として対になっていることの確認を、例えば、キーボードやマウス等の入力デバイス（図 1 において図示せず。）を介して、オペレータから入力される（ステップ S 1 3 ）。

【 0 0 4 5 】

なお、表示した文字列を見たオペレータによって、2つの文であるべき文字列が1つの文として取得されていると判断された場合、登録情報取得部 2 4 は、表示した文字列において文の切れ目となる位置を、入力デバイスを介して、オペレータから入力される。2つの文であるべき文字列が1つの文として取得されている場合というのは、例えば、1つの文であるとみなすべきタイトルと、そのタイトルの後に続く文とが1つの文として取得されている場合である。この例の場合、登録情報取得部 2 4 は、タイトルの最後の文字が文の切れ目になるという指定をオペレータから受け付けばよい。また、オペレータは、そのような指定を行えばよい。登録情報取得部 2 4 は、オペレータに指定された文の切れ目より前の部分を、ステップ S 1 1 あるいはステップ S 1 2 で取得した文字列とみなす。その文の切れ目より後ろの部分は、再度ステップ S 1 1 あるいはステップ S 1 2 の処理を行うときに再び取得する。

【 0 0 4 6 】

また、表示した文字列を見たオペレータによって、1つの文の途中までしか取得されていないと判断された場合、登録情報取得部 2 4 は、オペレータの操作にしたがい、さらに文字列を取得する。1つの文の途中までしか取得されていない場合というのは、例えば、文の途中に用いられたピリオドまでが1つの文として取得された場合である。この例の場合、登録情報取得部 2 4 は、入力デバイスを介してオペレータから、次のピリオド（文の切れ目に用いられる記号）までの文字列をさらに取得する旨の指示を受け付けばよい。また、オペレータは、そのような指示を入力すればよい。登録情報取得部 2 4 は、この指示に応じて、さらに原文ファイル 1 3 a または訳文ファイル 1 3 b から文字列、およびその文字列の文字属性情報を取得する。そして、登録情報取得部 2 4 は、既にステップ S 1 1 あるいはステップ S 1 2 で取得した文字列と、さらに取得した文字列とが1つの文を構成していると判定する。

【 0 0 4 7 】

以上のような処理の結果、表示した文字列が原文および訳文として対になっていることの確認が、オペレータから入力されたならば、登録情報取得部 2 4 は、ステップ S 1 4 の処理を行う。

【 0 0 4 8 】

ステップ S 1 4 において、登録情報取得部 2 4 は、原文ファイル 1 3 a から取得した文字列（1つの文を構成する文字列）、およびその文字列の文字属性情報を、原文記憶部 3 1 に記憶させる。さらに、登録情報取得部 2 4 は、その記憶領域に対応する訳文記憶部 3 2 の記憶領域に、訳文ファイル 1 3 b から取得した文字列（1つの文を構成する文字列）、およびその文字列の文字属性情報を記憶させる。

【 0 0 4 9 】

登録情報取得部 2 4 は、ステップ S 1 4 の後、原文ファイル 1 3 a および訳文ファイル 1 3 b からの文字列の取得が終了したか否かを判定する（ステップ S 1 5 ）。原文ファイル 1 3 a および訳文ファイル 1 3 b にまだ取得していない文字列が存在するならば、文字列の取得が終了していないと判定し、ステップ S 1 1 以降の処理を繰り返す。また、原文ファイル 1 3 a および訳文ファイル 1 3 b から全ての文字列を取得しているならば、文字列の取得が終了したと判定し、処理を終了する。

【 0 0 5 0 】

例えば、図 2 に例示する文字列を記述した原文ファイル 1 3 a および訳文ファイル 1 3

10

20

30

40

50

bが入力されたとする。すると、図4に示す登録情報取得部24の処理によって、「雨時々晴れ。」という文字列が原文記憶部31に記憶される。また、文字「雨」および文字「晴れ。」がゴシック体であり、文字「時々」が太字のゴシック体、かつ斜字体であり、下線付きであるということを定めた文字属性情報も、原文記憶31に記憶される。そして、その記憶領域に対応する訳文記憶部32の記憶領域に、「Showers with sunny intervals .」という文字列が記憶される。また、文字「with」が太字のゴシック体、かつ斜字体であり、下線付きであり、他の文字がゴシック体であることを定めた文字属性情報も記憶される。原文の第2文、訳文の第2文についても、同様に記憶される。

【0051】

なお、ある表示態様の原文を翻訳するときに、原文とは異なる表示態様の訳文を得ることをシステム管理者が意図している場合には、対になる原文と訳文において、文字属性情報が異なるようにしておけばよい。例えば、図2(a)の第2文(文字「雨」等はゴシック体)の翻訳結果として、図2(b)の第2文(文字「Rain」等は明朝体)を出力することをシステム管理者が意図している場合には、図2の第2文で例示したように、文字属性情報が異なるようにして、原文ファイル13aおよび訳文ファイル13bを作成しておけばよい。

10

【0052】

次に、新規の文書ファイル11に記述された原文の訳文を記述した訳文ファイル12を出力する動作について説明する。図5は、この動作の処理経過の例を示すフローチャートである。

20

【0053】

文字列取得部21は、新規の文書ファイル11が入力されると、そのファイルのコピーを作成する。なお、コピーされたファイルは、訳文ファイル12として用いられ、翻訳メモリシステムは、コピーされたファイルに記述されている原文を訳文に置換するようにして、訳文を訳文ファイル12に記述していく。

【0054】

また、文字列取得部21は、新規の文書ファイル11に記述されている原文から、文字列、およびその文字列の文字属性情報を取得する(ステップS21)。ステップS21において、文字列取得部21は、1つの文毎に、文を構成する文字列、およびその文字列の文字属性情報を取得する。文字列取得部21は、文の切れ目に用いられる記号の存在によ

30

【0055】

ただし、既に説明したように、文の切れ目に用いられる記号のみによって判定を行うと、2つの文であるべき文字列が1つの文として取得されたり、1つの文の途中までが1つの文として取得される場合が生じる。したがって、文字列取得部21は、ステップS13の処理と同様に、取得した文字列をディスプレイ装置等の表示手段(図1において図示せず。)に表示させる。表示した文字列を見たオペレータによって、2つの文であるべき文字列が1つの文として取得されていると判断された場合、文字列取得部21は、表示した文字列において文の切れ目となる位置を、キーボードやマウス等の入力デバイス(図1において図示せず。)を介して、オペレータから入力される。文字列取得部21は、オペレータに指定された文の切れ目より前の部分を、ステップS21で取得した文字列とみなす。その文の切れ目より後ろの部分は、再度ステップS21の処理を行うときに再び取得する。また、表示した文字列を見たオペレータによって、1つの文の途中までしか取得されていないと判断された場合、文字列取得部21は、オペレータの指示にしたがい、さらに文字列、およびその文字列の文字属性情報を取得する。そして、文字列取得部21は、既に取得した文字列と、さらに取得した文字列とが1つの文を構成していると判定する。オペレータの操作にしたがって、1つの文を確定するまでの動作は、ステップS13において1つの文を確定する場合の動作と同様である。

40

【0056】

以上のような処理の結果、表示した文が1つの文になっていることの確認がオペレータ

50

から入力されたならば、文字列検索部 2 2 によるステップ S 2 2 の処理に移行する。

【 0 0 5 7 】

なお、文字列取得部 2 1 は、ステップ S 2 1 において、文字列を取得する際、その文字列が記述されている位置（例えば、新規の文書ファイル 1 1 の内容を表示した場合におけるその文字列の表示位置）を判定する。

【 0 0 5 8 】

ステップ S 2 1 では、1 つの文毎に、文を構成する文字列、およびその文字列の文字属性情報を取得するので、以下に説明するステップ S 2 1 ~ S 2 5 の処理も、1 つの文の文字列ごとに行われる。

【 0 0 5 9 】

ステップ S 2 2 において、文字列検索部 2 2 は、文字列取得部 2 1 が取得した文字列（1 つの文を構成する文字列）と同一の文字列を、原文記憶部 3 1 が記憶する文字列の中から検索する（ステップ S 2 2）。文字列取得部 2 1 が取得した文字列と同一の文字列が原文記憶部 3 1 に 1 つしか記憶されていないならば、文字列検索部 2 2 は、その文字列を検索すればよい。

【 0 0 6 0 】

ただし、文字列取得部 2 1 が取得した文字列と同一の文字列が、原文記憶部 3 1 の複数の記憶領域に記憶されている場合もあり得る。例えば、原文記憶部 3 1 のある記憶領域に記憶されている文字列と、原文記憶部 3 1 の他の記憶領域に記憶されている文字列とが同一であり、それぞれの文字列の文字属性情報が異なっている場合があり得る。このように、文字列取得部 2 1 が取得した文字列と同一の文字列が原文記憶部 3 1 の複数の記憶領域に記憶されている場合、文字列検索部 2 2 は、その各記憶領域に記憶されている文字列を、それぞれディスプレイ装置等の表示手段（図 1 において図示せず。）に表示させる。このとき、文字列検索部 2 2 は、各文字列を、各文字列の文字属性情報によって定まる表示態様で表示させる。このように、表示態様が異なる同一の文字列をそれぞれ表示して、オペレータに対し、オペレータが意図している検索対象の文字列の指定を促す。そして、キーボードやマウス等の入力デバイス（図 1 において図示せず。）は、表示された複数の文字列のうちの一つの文字列を、オペレータから指定される。文字列検索部 2 2 は、入力デバイスを介してオペレータから指定された文字列を検索する。

【 0 0 6 1 】

文字列検索部 2 2 は、文字列取得部 2 1 が取得した文字列と同一の文字列の検索に成功したか否かを判定する（ステップ S 2 3）。文字列取得部 2 1 が取得した文字列と同一の文字列をステップ S 2 2 で検索できたならば、文字列検索部 2 2 は、検索した文字列と対になる訳文の文字列、およびその文字列の文字属性情報を訳文記憶部 3 2 から抽出する。このとき、検索した文字列の記憶領域に対応する訳文記憶部 3 2 の記憶領域から訳文の文字列および文字属性情報を抽出すればよい。そして、ファイル出力部 2 3 が、その文字列および文字属性情報を出力する（ステップ S 2 4）。ステップ S 2 4 において、ファイル出力部 2 3 は、抽出された文字属性情報によって、抽出された文字列の表示態様を定め、表示する場合にはその表示態様で表示されるように、抽出された文字列を訳文ファイル 1 2 に書き込む。なお、訳文ファイル 1 2 は、新規の文書ファイル 1 1 のコピーとして作成されたファイルであるので、訳文ファイル 1 2 には、ステップ S 2 1 において取得された文字列（原文）が記述されている。ファイル出力部 2 3 は、訳文ファイル 1 2 におけるその原文が記述された位置で、その原文を、文字列検索部 2 2 によって抽出された文字属性情報により表示態様が定められた訳文に置換するようにして、訳文を訳文ファイル 1 2 に記述する。なお、訳文ファイル 1 2 内で置換対象となる原文が記述された位置は、ステップ S 2 1 で判定された位置と同じ位置である。

【 0 0 6 2 】

文字列取得部 2 1 が取得した文字列と同一の文字列を文字列検索部 2 2 がステップ S 2 2 で検索できなかったならば、ファイル出力部 2 3 は、文字列取得部 2 1 が取得した文字列、およびその文字列の文字属性情報を出力する（ステップ S 2 5）。なお、文字列取得

10

20

30

40

50

部 2 1 が取得した文字列と同一の文字列をステップ S 2 2 で検索できなかった場合というのは、文字列取得部 2 1 が取得した文字列が原文記憶部 3 1 に記憶されていない場合である。ステップ S 2 5 において、ファイル出力部 2 3 は、ステップ S 2 4 で訳文を記述する場合と同様に、ステップ S 2 1 で取得した文字列を訳文ファイル 1 2 に記述する。すなわち、訳文ファイル 1 2 には、ステップ S 2 1 において取得された文字列（原文）が記述されている。ファイル出力部 2 3 は、訳文ファイル 1 2 におけるその原文が記述された位置で、その原文を、ステップ S 2 1 で取得した文字属性情報により表示態様が定められた文字列に置換するようにして、ステップ S 2 1 で取得した文字列を訳文ファイル 1 2 に記述する。なお、この場合、置換前の文字列と置換後の文字列とは、同一である。

【 0 0 6 3 】

10

ステップ S 2 4 またはステップ S 2 5 の処理が終了したならば、文字列取得部 2 1 は、新規の文書ファイル 1 1 からの文字列の取得が終了したか否かを判定する（ステップ S 2 6）。新規の原文ファイル 1 1 にまだ取得していない文字列が存在するならば、文字列の取得が終了していないと判定し、再度ステップ S 2 1 の処理を実行し、文字列検索部 2 2 およびファイル出力部 2 3 は、再度ステップ S 2 2 ～ステップ S 2 4（またはステップ S 2 5）の処理を実行する。新規の文書ファイル 1 1 から全ての文字列を取得しているならば、文字列の取得が終了したと判定し、処理を終了する。この結果、原文に応じた表示態様で訳文が記述された訳文ファイル 1 2 が得られる。

【 0 0 6 4 】

また、新規の文書ファイル 1 1 には、画像データ（線図で描かれた図形データも含む。）が含まれていてもよい。新規の文書ファイル 1 1 に含まれる画像データは、そのまま訳文ファイル 1 2 にコピーされる。また、ファイル出力部 2 3 は、コピーされたファイル（訳文ファイル 1 2）において、原文が記述された位置で、その原文を訳文に置換するようにして、訳文を記述する。従って、新規の文書ファイル 1 1 に含まれる図形データを、訳文ファイル 1 2 においてもそのまま保持することができる。また、新規の文書ファイル 1 1 を表示した場合のレイアウトを、訳文ファイル 1 2 においてもそのまま保持することができる。

20

【 0 0 6 5 】

以下に、ステップ S 2 1 ～ S 2 6 の処理の具体例を示す。ただし、図 2（a）に示す原文に基づいて文字列および文字属性情報が原文記憶部 3 1 に記憶され、図 2（b）に示す訳文に基づいて文字列および文字属性情報が訳文記憶部 3 2 に記憶されているものとする。図 6（a）は、新規の文書ファイル 1 1 に記述されている原文の例を示す。図 6（a）に示す原文が記述された新規の文書ファイル 1 1 が文字列取得部 2 1 に入力された場合、文字列取得部 2 1 は、図 6（a）に示す第 1 文の文字列「雨時々晴れ。」という文字列と、その文字属性情報を取得する（ステップ S 2 1）。この文字属性情報は、文字「雨」および文字「晴れ。」がゴシック体であり、文字「時々」が太字のゴシック体、かつ斜字体であり、下線付きであるということを定めている。また、文字列取得部 2 1 は、新規の文書ファイル 1 1 のコピーを作成する。

30

【 0 0 6 6 】

次に、文字列検索部 2 2 は、ステップ S 2 1 で取得された文字列「雨時々晴れ。」と同一の文字列を、原文記憶部 3 1 が記憶する文字列の中から検索する（ステップ S 2 2）。本例では、原文記憶部 3 1 は、「雨時々晴れ。」という文字列を記憶している（図 2（a）参照）。したがって、文字列検索部 2 2 は、「雨時々晴れ。」という文字列を原文記憶部 3 1 から検索することができる（ステップ S 2 3 の Y）。したがって、ステップ S 2 4 に移行する。

40

【 0 0 6 7 】

文字列検索部 2 2 は、検索した文字列「雨時々晴れ。」が記憶される記憶領域に対応する訳文記憶部 3 2 の記憶領域から訳文の文字列および文字属性情報を抽出する。文字列「雨時々晴れ。」の記憶領域に対応する訳文記憶部 3 2 の記憶領域には、「Showers with sunny intervals.」という文字列と、その文字列の文字属性情報が記憶されている。この

50

文字属性情報は、文字「with」が太字のゴシック体、かつ斜字体であり、下線付きであり、他の文字がゴシック体であることを定めている（図2（b）参照。）。したがって、文字列検索部22は、「Showers with sunny intervals.」という文字列、およびその文字属性情報を抽出する。そして、ファイル出力部23は、抽出した文字属性情報によって表示態様が定められた文字列「Showers with sunny intervals.」を、訳文ファイル12（新規の文書ファイル11のコピー）に記述する。このとき、「雨時々晴れ。」が記述されていた位置で、「雨時々晴れ。」を「Showers with sunny intervals.」に置換するようにして、「Showers with sunny intervals.」を記述する。この結果、訳文ファイルには、第1文として、図6（b）に示す第1文「Showers with sunny intervals.」が書き込まれる。

10

【0068】

次に、文字列取得部21は、新規の文書ファイル11からの文字列の取得が終了したか否かを判定する（ステップS26）。新規の文書ファイル11には、まだ取得されていない第2文「晴れの日。」が残っているので（図6（a）参照。）、文字列取得部21は、文字列の取得が終了していないと判定する（ステップS26のN）。そして、文字列取得部21は、図6（a）に示す第2文の文字列「晴れの日。」という文字列と、その文字列の文字属性情報を取得する（ステップS21）。この文字属性情報は、文字「晴れ」および文字「日」がゴシック体であり、文字「の」が太字のゴシック体、かつ斜字体であり、下線付きであることを定めている。

【0069】

20

次に、文字列検索部22は、ステップS21で取得された文字列「晴れの日。」と同一の文字列を、原文記憶部31が記憶する文字列の中から検索する（ステップS22）。本例では、原文記憶部31は、「晴れの日。」という文字列を記憶していない（図2（a）参照。）。したがって、文字列検索部22は、「晴れの日。」という文字列を原文記憶部31から検索することができない（ステップS23のN）。したがって、ステップS25に移行する。

【0070】

ファイル出力部23は、ステップS21で取得した文字属性情報によって表示態様が定められた文字列「晴れの日。」（この文字列も、ステップS21で取得されたものである。）を、訳文ファイル12（新規の文書ファイル11のコピー）に記述する。このとき、「晴れの日。」が記述されていた位置で、「晴れの日」を「晴れの日」に置換するようにして、記述する。この場合では、置換前の文字列と置換後の文字列とは、同一である。この結果、訳文ファイルには、第2文として、図6（b）に示す第2文「晴れの日。」が書き込まれる。

30

【0071】

次に、文字列取得部21は、新規の文書ファイル11からの文字列の取得が終了したか否かを判定する（ステップS26）。新規の文書ファイル11からは、全ての文字列を取得しているので（ステップS26のY）、処理を終了する。以上の処理によって、図6（b）に示す文字列が記述された訳文ファイル12が出力される。なお、第2文「晴れの日。」については、原文と訳文の組が記憶装置3に記憶されていないために、翻訳できなかったことになる。この場合、オペレータが、訳文ファイル12を編集し、翻訳および翻訳後の文字列の属性変更を行えばよい。オペレータは、訳文記憶部31から抽出された文字列の属性については変更する必要がなく、図6（b）の第2文のように、翻訳されなかった文に対してのみ編集を行えばよい。したがって、オペレータの編集作業負担は軽減される。

40

【0072】

従来の、翻訳メモリシステムは、予め、対になる原文と訳文のみを記憶し、文字属性情報を記憶していなかった。そのため、出力される訳文を表示したときの表示態様は、デフォルトで定められた表示態様となり、入力される原文の表示態様に応じたものにはなっていなかった。したがって、従来の翻訳メモリシステムに、例えば、図7（a）に示す文字

50

列を入力したとしても、出力結果は、図 7 (b) に示すような予め定められた表示態様で表される文字列となり、入力される原文に応じた表示態様にならない。

【 0 0 7 3 】

本発明によれば、訳文記憶部 3 2 は、訳文の文字列とともにその文字列の文字属性情報も記憶する。したがって、表示時に原文に応じた表示態様で表示されるように、訳文の文字列を訳文ファイル 1 2 に書き込むことができる。その結果、オペレータは、原文に応じて表示態様が定められた文字列を記述した訳文ファイル 1 2 を得ることができる。オペレータは、その文字列の属性を変更する操作を行う必要はなく、翻訳されなかった文に対してのみ編集作業を行えばよい。したがって、オペレータの作業負担を軽減させることができる。

10

【 0 0 7 4 】

また、原文記憶部 3 1 は、新規に入力された原文の文字列と同一の文字列を原文記憶部 3 1 から検索できなかった場合に、新規に入力された文字列を、表示した場合における表示態様がそのまま保持されるようにして、訳文ファイル 1 2 に書き込む。したがって、訳文を出力できない場合に、入力された文字列をそのままの表示態様で出力することができる。

【 0 0 7 5 】

また、文字列取得部 2 1 は、新規の文書ファイル 1 1 のコピーを作成し、ファイル出力部 2 3 は、コピーされたファイルである訳文ファイル 1 2 に対して訳文を記述する。そして、ファイル出力部 2 3 は、訳文ファイル 1 2 において、原文が記述された位置で、その原文を訳文に置換するようにして、訳文を記述する。従って、新規の文書ファイル 1 1 に含まれる図形データを、訳文ファイル 1 2 においてもそのまま保持することができる。また、新規の文書ファイル 1 1 を表示した場合のレイアウトを、訳文ファイル 1 2 においてもそのまま保持することができる。

20

【 0 0 7 6 】

また、原文記憶部 3 1 は、原文の文字列とともに、その文字列の文字属性情報も記憶する。したがって、同一の文字列が原文記憶部 3 1 に複数記憶されている場合、その複数の同一文字列を、それぞれ文字属性情報に応じた表示態様で表示し、オペレータに対し、オペレータが意図している検索対象の文字列の指定を促すことができる。

【 0 0 7 7 】

30

以上の説明では、ステップ S 2 2 において、文字列取得部 2 1 が取得した文字列と同一の文字列が原文記憶部 3 1 の複数の記憶領域に記憶されているとき、オペレータに検索対象文字列の指定を促す場合を説明した。オペレータに検索対象文字列の指定を促すのではなく、文字列検索部 2 2 が、複数の記憶領域に記憶された同一の文字列の中から一つの文字列を特定してよい。例えば、同一の文字列が原文記憶部 3 1 の複数の記憶領域に記憶されている場合、文字列検索部 2 2 が、最後に記憶された文字列を検索してもよい。この場合、文字列検索部 2 2 は、複数の同一文字列をそれぞれの表示態様で表示して、オペレータに指定を促すことなく文字列検索を行う。したがって、複数の同一文字列をそれぞれの表示態様で表示する必要がないので、この場合には、原文記憶部 3 1 に文字属性情報を記憶させなくてもよい。ただし、原文記憶部 3 1 は、個々の文字列を記憶した順番を特定できるように個々の文字列を記憶する。例えば、文字列を記憶する記憶領域に予め順番を設けておき、登録情報取得部 2 4 が文字列を原文記憶部 3 1 に記憶させる場合、文字列の順番に応じた記憶領域に文字列を記憶させればよい。したがって、同一の文字列が原文記憶部 3 1 の複数の記憶領域に記憶されている場合、文字列検索部 2 2 は、その複数の記憶領域のうち最後の順番が定められた記憶領域に記憶されている文字列を検索してもよい。また、原文記憶部 3 1 が、個々の文字列を記憶した順番を特定できるように個々の文字列を記憶する構成として、文字列と、その文字列が記憶された時刻とを記憶するようにしてもよい。

40

【 0 0 7 8 】

また、ステップ S 2 2 の検索処理において、検索部 2 2 は、文字列取得部 2 1 が取得し

50

た文字列と同一の文字列であって、文字属性情報が、文字列取得部 2 1 が取得部した文字属性情報と同一である文字列を抽出してもよい。すなわち、検索部 2 2 は、文字列の同一性だけでなく文字属性情報の同一性も判定し、文字列と文字属性情報の双方において同一性を確認した場合に、原文記憶部 3 1 に記憶されているその文字列を検索対象として特定してもよい。

【0079】

また、上記の説明では、ファイル出力部 2 3 が、新規の文書ファイル 1 1 のコピーに対して訳文を記述する場合を示した。新規の文書ファイル 1 1 のコピーではなく、新規の文書ファイル 1 1 とは別の新たなファイルに対して訳文を記述してもよい。ファイル出力部 2 3 は、ステップ S 2 2 の検索が成功した場合には、文字列検索部 2 2 によって抽出された文字属性情報により表示態様が定められた訳文（検索部 2 2 によって抽出された文字列）を新たなファイルに記述すればよい。ステップ S 2 2 の検索が成功しなかった場合には、ステップ S 2 1 で取得した文字属性情報により表示態様が定められた原文（ステップ S 2 1 で取得された原文）を新たなファイルに記述すればよい。

10

【0080】

また、上記の説明では、訳文をファイルとして出力する場合を説明したが、ディスプレイ装置等の表示手段（図 1 において図示せず。）に表示させてもよい。すなわち、訳文を表示出力させてもよい。この場合、ファイル出力部 2 3 は、文字属性情報によって定められる表示態様で訳文をディスプレイ装置等に表示させればよい。また、新規の文書ファイル 1 1 に記述されている原文とともに訳文を表示させてもよい。さらに、入力デバイス（図 1 において図示せず。）を介して、オペレータから、表示させた文字列の編集操作を受け付けてもよい。

20

【0081】

また、上記の説明では、原文記憶部 3 1 および訳文記憶部 3 2 に 1 文毎に文字列を記憶させ、ステップ S 2 1 ～ S 2 5 の処理を 1 文毎に行う場合を示した。1 文ではなく、句単位や単語単位に処理を行ってもよい。その場合、原文記憶部 3 1 および訳文記憶部 3 2 に句（または単語）毎に文字列を記憶させ、ステップ S 2 1 ～ S 2 5 の処理を句（または単語）毎に行えばよい。また、原文ファイル 1 3 a、訳文ファイル 1 3 b および新規の文書ファイル 1 1 において、句（または単語）の切れ目を示す記号を記述しておき、登録情報取得部 2 4 および文字列取得部 2 1 は、その記号によって、句の切れ目（または単語の切れ目）を判定すればよい。

30

【0082】

以上の説明において、原文記憶部 3 1 は、第 1 言語文字列記憶手段に相当する。訳文記憶部 3 2 は、第 2 言語文字列記憶手段に相当する。文字列取得部 2 1 は、入力手段に相当する。文字列検索部 2 2 は、検索手段および抽出手段に相当する。ファイル出力部 2 3 は、出力手段に相当する。また、ディスプレイ装置（図 1 において図示せず。）は、表示手段に相当する。また、キーボードやマウスなどの入力デバイス（図 1 において図示せず。）は、指定受付手段に相当する。

【産業上の利用可能性】

【0083】

本発明は、翻訳メモリによって翻訳を行う翻訳メモリシステムに適用可能である。また、原文と訳文とを並べて表示して、表示した原文と訳文の編集操作を可能とする対訳エディタシステム等にも適用可能である。

40

【図面の簡単な説明】

【0084】

【図 1】本発明による翻訳メモリシステムの構成の例を示すブロック図である。

【図 2】原文ファイルおよび訳文ファイルに記述された文字列の例を示す説明図である。

【図 3】対応する記憶領域の例を示す説明図である。

【図 4】原文および訳文を記憶装置に記憶させる動作の処理経過の例を示すフローチャートである。

50

【図 5】訳文を記述した訳文ファイルを出力する動作の処理経過の例を示すフローチャートである。

【図 6】新規の文書ファイルおよび出力される訳文ファイルに記述される文字列の例を示す説明図である。

【図 7】従来の翻訳メモリシステムにおいて出力される訳文の表示態様の例を示す説明図である。

【符号の説明】

【 0 0 8 5 】

2 データ処理装置

3 記憶装置

10

1 1 新規の文書ファイル

1 2 訳文ファイル

1 3 登録用文書ファイル

1 3 a 原文ファイル

1 3 b 訳文ファイル

2 1 文字列取得部

2 2 文字列検索部

2 3 ファイル出力部

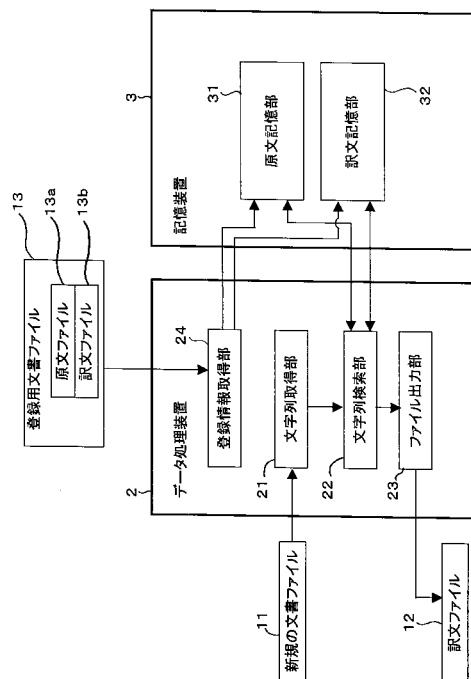
2 4 登録情報取得部

3 1 原文記憶部

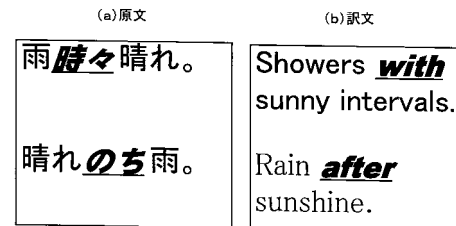
20

3 2 訳文記憶部

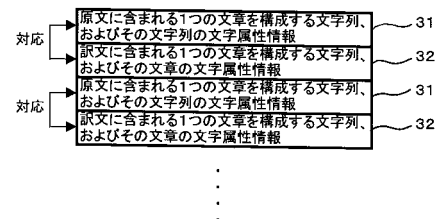
【図 1】



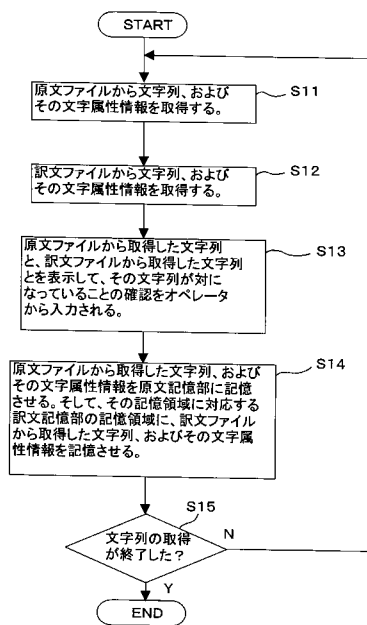
【図 2】



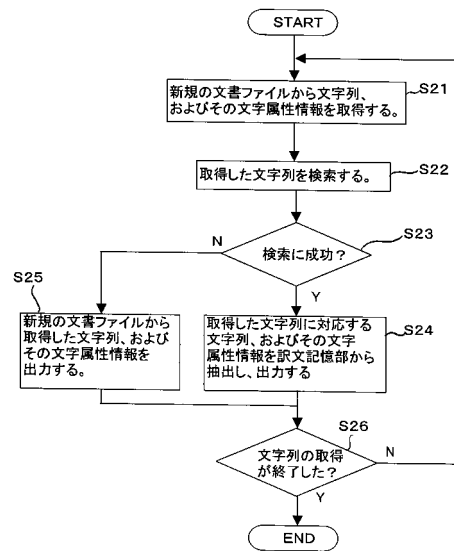
【図 3】



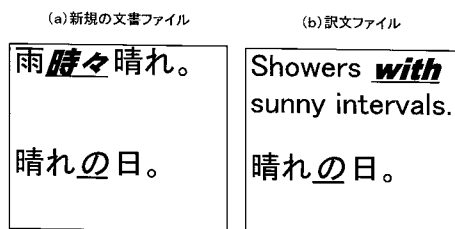
【図 4】



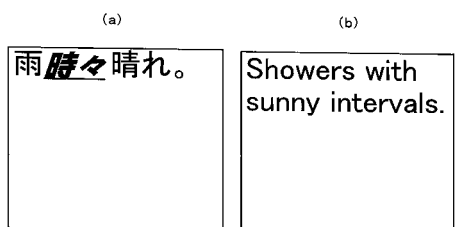
【図 5】



【図 6】



【図 7】



フロントページの続き

(56)参考文献 特開平02 - 139666 (JP, A)
特開平03 - 009472 (JP, A)
特開平05 - 135095 (JP, A)

(58)調査した分野(Int.Cl., DB名)
G06F 17/27 - 17/28