

(51) Int.Cl. ⁷	識別記号	F I	テ-マコード [*]	(参考)
G06F 17/28		G06F 17/28	Z	5B009
17/21	550	17/21	L	5B091

審査請求 有 請求項の数 5 O L (全 7 頁)

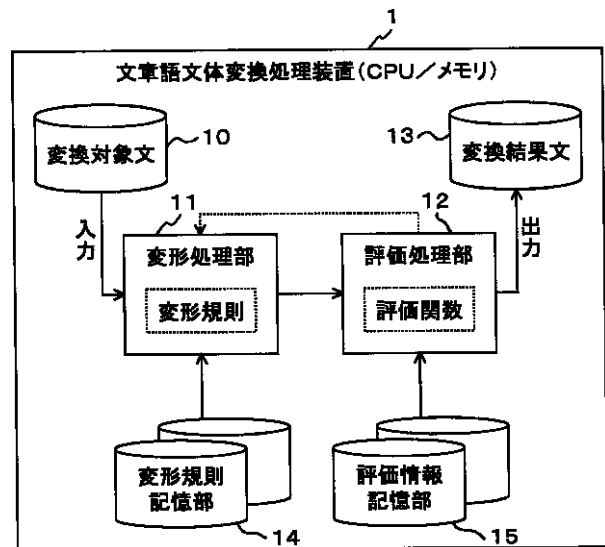
(21)出願番号	特願2001 - 205888(P 2001 - 205888)	(71)出願人	301022471 独立行政法人通信総合研究所 東京都小金井市貫井北町 4 - 2 - 1
(22)出願日	平成13年 7 月 6 日(2001.7.6)	(72)発明者	村田 真樹 東京都小金井市貫井北町 4 - 2 - 1 独立 行政法人通信総合研究所内
特許法第30条第 1 項適用申請有り 平成13年 3 月30日 言語処理学会発行の「言語処理学会第 7 回年次大会ワー クショップ論文集」に発表		(72)発明者	井佐原 均 東京都小金井市貫井北町 4 - 2 - 1 独立 行政法人通信総合研究所内
		(74)代理人	100097836 弁理士 福井 國敬 (外 2 名)
		F タ-ム(参考)	5B009 QA17 5B091 AA15

(54) 【発明の名称】 文章語文体変換システムおよび文章語文体変換処理プログラム

(57) 【要約】

【課題】 入力した文章語を目的とする文体に適切に変換できるようにする。

【解決手段】 第 1 の文字列を同義の第 2 の文字列に言い換える変形の規則を多数記憶する変形規則記憶部 1 4 と、文字列を変形した結果が目的とするふさわしい文体になっているかどうかを評価するための尺度を与える評価関数または評価規則を記憶する評価情報記憶部 1 5 とを用意し、自然言語で記述された文体変換対象の文字列を入力すると、変形処理部 1 1 によって、入力された文字列を変形規則を用いて変形し、変換の候補を生成する。次に、生成された変換の候補を、評価処理部 1 2 によって、評価関数または評価規則を用いて評価し、評価結果のよい表現を選択し、選択された表現の変換結果を、目的とする文体に変換された文章語として出力する。



【特許請求の範囲】

【請求項 1】 ある自然言語で記述された文章語を、同一の自然言語で記述された他の文体による文章語に変換するシステムであって、前記自然言語で記述された第 1 の文字列を同義の第 2 の文字列に言い換える変形の規則を記憶する変形規則記憶手段と、自然言語で記述された文体変換対象の文字列を入力する入力手段と、前記入力された文字列を前記変形規則記憶手段に記憶された変形の規則を用いて変形し、変換の候補を生成する変形処理手段と、文字列を変形した結果の表現が目的とする文体になっているかどうかを評価するための所定の評価尺度に基づいて、前記変形処理手段により生成された変換の候補を評価し、評価結果のよい表現を選択する評価処理手段と、前記選択された表現の変換結果を、目的とする文体に変換された文章語として出力する出力手段とを備えることを特徴とする文章語文体変換システム。

【請求項 2】 前記評価尺度は、変換先の文体による文章を多く含む文章集合からなる大量の用例における変換の候補の出現頻度または出現確率を含み、出現頻度または出現確率が大きいものに高い評価を与えるものであることを特徴とする請求項 1 記載の文章語文体変換システム。

【請求項 3】 前記評価尺度は、平易な文章を多く含む文章集合からなる大量の用例における変換の候補の出現頻度または出現確率を含み、出現頻度または出現確率が大きいものに高い評価を与えるものであり、前記文体の変換は、入力した難解文の文字列を平易文に変換するものであることを特徴とする請求項 1 記載の文章語文体変換システム。

【請求項 4】 前記評価尺度は、特定の個人の文章集合からなる大量の用例における変換の候補の出現頻度または出現確率を含み、出現頻度または出現確率が大きいものに高い評価を与えるものであり、前記文字列の変換は、入力した文字列の表現を前記特定の個人の文体に変換するものであることを特徴とする請求項 1 記載の文章語文体変換システム。

【請求項 5】 コンピュータによって、ある自然言語で記述された文章語を、同一の自然言語で記述された他の文体による文章語に変換するためのプログラムであって、自然言語で記述された文体変換対象の文字列を入力する処理と、前記入力された文字列を、前記自然言語で記述された第 1 の文字列を同義の第 2 の文字列に言い換える変形の規則を用いて変形し、変換の候補を生成する処理と、文字列を変形した結果の表現が目的とする文体になっているかどうかを評価するための所定の評価尺度に基づいて、前記変形によって生成された変換の候補を評価し、評価結果のよい表現を選択する処理と、前記選択された表現の変換結果を、目的とする文体に変換された文章語として出力する処理とを、コンピュータに実行させるための文章語文体変換処理プログラム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、ある自然言語で記述された文章語の文字列を、同一の自然言語で記述された他の文体による文章語の文字列に変換する文章語文体変換システムに関するものである。

【0002】

【従来の技術】自然言語で記述された文または文章に関する表現の変換処理として典型的なものは、機械翻訳である。機械翻訳では、ある国の自然言語で記述された文または文章を他の国の自然言語で記述された文または文章に変換する。

【0003】機械翻訳が他の国の言語に変換するのに対し、同一の自然言語間での文または文章の変換処理を行うシステムも用いられるようになってきている。例えば、要約文を自動生成したり、文章を推敲したりするシステムである。

【0004】一般に同一自然言語間での文の変換処理では、変換前の語・句・文などのパターンと変換後の語・句・文などのパターンとの対からなる変換規則を大量に用意し、いわゆるパターン・マッチングによって入力文中に現れる変換前のパターンを探し出し、該当するパターンがあれば、それを変換後の語・句・文などのパターンに置き換える処理を行っている。

【0005】また、同一自然言語間での文の変換処理として、要約文を自動生成したり、文章を推敲したりすることは行われていたが、例えば芥川龍之介が書いた小説を、夏目漱石の文体の小説に変換するというように、ある特定の個人の文体または一般的な文体の文章を、他の特定の個人の文体に変換するというようなことは行われていなかった。

【0006】

【発明が解決しようとする課題】従来の同一自然言語間での文または文章の変換処理では、一般に変換規則による一律な変換を行っており、変換結果の良し悪しについての評価は行われていなかった。そのため、実際により変換が行われるかどうかは、あらかじめ用意された変換規則の良し悪しに大きく依存し、適用された変換規則によっては、目的とする変換と異なる変換結果になってしまうことがあった。

【0007】また、変換の精度を良くしようとすると、変換規則として真に妥当なものだけを選別する必要があり、大量な変換規則を選別するのは、大変な困難を伴う作業であった。例えば変換規則の中に、「A」という文字列を「B」という文字列に変換する規則と、これとは逆に「B」という文字列を「A」という文字列に変換する規則とが混在すると、従来の技術では目的とする正しい変換結果は得られなかった。

【0008】そのため、特に従来技術では、個人文体の変換や難解な文を小学生にも理解しやすい平易文に変換

するというようなシステムを構築することは難しいという問題があった。

【 0 0 0 9 】本発明は上記問題点の解決を図り、文章語の表現を変える規則として、厳選された規則ではなく、例えばコンピュータにより自動獲得されたような多種多様な規則を用いた場合でも、入力した文章語を目的とする文体に適切に変換できるシステムを提供することを目的とする。

【 0 0 1 0 】

【課題を解決するための手段】本発明は、上記課題を解決するため、ある自然言語で記述された文章語を、同一の自然言語で記述された他の文体による文章語に変換するシステムにおいて、第 1 の文字列を同義の第 2 の文字列に言い換える変形の規則を多数記憶する変形規則記憶手段と、文字列を変形した結果が目的とする文体になっているかどうかを評価するための尺度を与える評価関数または評価規則を記憶する評価情報記憶手段とを備え、自然言語で記述された文体変換対象の文字列を入力すると、変形処理手段によって、入力された文字列を変形規則記憶手段に記憶された変形の規則を用いて変形し、変換の候補を生成する。次に、生成された変換の候補を、評価処理手段によって評価情報記憶手段に記憶された評価関数または評価規則を用いて評価し、評価結果のよい表現を選択し、選択された表現の変換結果を、目的とする文体に変換された文章語として出力する。

【 0 0 1 1 】例えば、評価関数または評価規則による評価の尺度を、平易な文章集合からなる大量の用例における変換の候補の出現頻度または出現確率を含むものとし、その出現頻度または出現確率が大きいものに高い評価を与えるようにすることにより、入力した難解文の文字列を平易文に変換する文章語文体変換システムを実現することができる。

【 0 0 1 2 】また、例えば評価関数または評価規則による評価の尺度を、特定の個人の文章集合からなる大量の用例における変換の候補の出現頻度または出現確率を含むものとし、その出現頻度または出現確率が大きいものに高い評価を与えるようにすることにより、入力した文字列の表現をある特定の個人の文体に変換する文章語文体変換システムを実現することができる。

【 0 0 1 3 】評価の尺度として、必ずしも前記出現頻度または出現確率の大小を用いる必要はないが、何らかの評価尺度によって、文章語の変形の後に目的に合致したふさわしい文体になっているかどうかを評価し、高い評価の変形を変換結果として選択するので、あらかじめ文体の変形規則として用意する情報は、基本的に文字列の同義性があれば十分であり、多種多様な変形規則を厳選することなく用いることができる。変形元と変形先というような変形規則の方向性についての考慮も不要である。すなわち、例えば変形規則の中に、「 A 」という文字列を「 B 」という文字列に変形する規則と、これとは

逆に「 B 」という文字列を「 A 」という文字列に変形する規則とが混在しているような場合でも、評価によって最終的に目的とする変換結果が得られることになる。したがって、変形規則の作成が容易であり、ある文体変換のために用意した変形規則を、他の目的の文体変換のために利用するようなことも可能である。

【 0 0 1 4 】以上の手段は、コンピュータと、そのコンピュータにインストールされ実行されるソフトウェアプログラムとによって実現することができ、そのプログラムは、コンピュータが読み取り可能な可搬媒体メモリ、半導体メモリ、ハードディスク等の適当な記録媒体に格納することができる。

【 0 0 1 5 】

【発明の実施の形態】図 1 は、本発明のシステム構成例を示す。図中、1 は CPU およびメモリなどからなるコンピュータによって実現される文章語文体変換処理装置であって、変形処理部 1 1、評価処理部 1 2、変形規則記憶部 1 4、評価情報記憶部 1 5 を備える。

【 0 0 1 6 】変換対象文 1 0 は、本システムにおける入力となる自然言語文である。以下、特に断らないが変換対象文 1 0 は必ずしも一文に限られるわけではなく、文章または句、節のようなものであってもよい。変換結果文 1 3 は、本システムの出力であって、変換対象文 1 0 を同一の種類の自然言語で元の文体と異なる文体に言い換えたものである。

【 0 0 1 7 】文章語文体変換処理装置 1 のモジュールは、基本的に変形処理部 1 1 と評価処理部 1 2 とから構成される。変形処理部 1 1 は、変形規則記憶部 1 4 に格納されている変形規則を用いて、変換の候補を獲得するモジュールである。評価処理部 1 2 は、変換の候補の良さ、すなわち目的のふさわしい文体であるかどうかを、あらかじめ評価情報記憶部 1 5 に記憶されている評価の尺度（評価関数など）によって評価し、最も評価の高い変換の候補を選択するモジュールである。

【 0 0 1 8 】変換対象文 1 0 が入力されると、変形処理部 1 1 は、変形規則を用いて変換の候補を挙げ、評価処理部 1 2 は、変形された文体の妥当性をチェックして、最も妥当であると判断されたものを選択し、その結果を変換結果文 1 3 として出力する。

【 0 0 1 9 】変形規則記憶部 1 4 に記憶する変形規則は、人手によってあらかじめ作成された規則であってもよいし、コンピュータによって大量の言語データから自動獲得したものでもよい。変形規則および変換の候補を評価する評価関数（評価尺度）の評価情報は、扱う問題ごとに適正なものを用意する。

【 0 0 2 0 】評価情報は、評価のための数値情報であってもよいし、関数群もしくはサブルーチン群などによる手続き的なものであってもよい。また、評価方法を記述した規則（ルール）であってもよい。これらの組み合わせで実現することも可能である。評価処理部 1 2 で用い

る評価の尺度の代表的な例としては、目的とする文体の文章集合からなる大量の言語データ中での出現頻度（または出現確率）を挙げることができる。

【 0 0 2 1 】例えば、文章語の文体を変える本システムにおいて、変形処理部 1 1 が使用する変形規則がすべて同義性を満足するものであるとする。この場合、文体を変換したいデータを、そのデータの出現（生起）確率が目的とする文体の文章集合中で高くなるように変形すると目的とする文体に非常に近い文章語となる。

【 0 0 2 2 】もう少し簡単な例でこれを説明すると、例えば入力した文章がいわゆる「です調」の文体であり、「～です」という文字列が多くあったとする。変形規則に「～です」を「～である」に変形する規則があったとしよう。目的とする文体が「である調」の文体であれば、その目的とする文体の文章データが大量に格納されたデータベースを用意して、評価のために用いる。そのデータベースにおいて「～です」と「～である」の出現回数を数える。「～である」の出現回数のほうが数が多い場合、「～である」のほうの評価を「～です」より高くする。この評価によって、「です調」の文体は、「である調」の文体に自動変換されることになる。

【 0 0 2 3 】ここで、出現頻度（または出現確率）を調べるコーパスをいろいろと変えることにより、さまざまな文体の変換の結果を得ることができる。例えば、入力データが法律関係の文のときに、コーパスとして平易な文章の集合を与えておくと、法律関係の難解な文章を平易な文章に変形させることが期待できる。

【 0 0 2 4 】また、ここで入力データとして適当に誰かが書いた小説の文章を入れて、コーパスとしてシェークスピアの小説をいれると、シェークスピアの文体の小説が新たに完成することになる。同様に、芥川龍之介の小説を夏目漱石の文体に変換するなどといったことも可能になる。

【 0 0 2 5 】本システムで用いる変形規則は、例えば次のようにしてコンピュータにより自動生成することも可能である。まず、複数の同一言語により記述された意味的な対応関係がある言語情報を抽出する（例えば、複数の異なる辞書ファイルを用意し、それらから同じ単語の説明文を抽出する）。次に、抽出した複数の言語情報を突き合わせ、その結果から同義語または同義フレーズを抽出する。抽出した同義語または同義フレーズから、第 1 の文字列を同義の第 2 の文字列に言い換える変形規則を自動生成する。

【 0 0 2 6 】図 2 は、変形処理部 1 1 の処理フローチャートである。変形処理部 1 1 は、まずステップ S 1 0 により、文体の変換対象として指定された変換対象文 1 0 を入力する。キーボードなどからの入力、ファイルからの入力、アプリケーションプログラムからの入力など、入力方法は問わない。

【 0 0 2 7 】ステップ S 1 1 では、変形規則記憶部 1 4

から変換に必要な変形規則を読み込む。既に読み込まれている場合には、ここでの読み込みは不要である。ステップ S 1 2 では、入力した変換対象文 1 0 を変形規則を用いて変形する。この変形した後の表現の候補を評価処理部 1 2 へ引き渡す。このとき、変形した後の表現の候補を一つずつ評価処理部 1 2 へ渡してもよいし、複数ある場合には複数まとめて渡してもよい。

【 0 0 2 8 】ステップ S 1 3 では、評価処理部 1 2 は、変形処理部 1 1 から変換対象文 1 0 を変形した後の表現の候補を受け取り、評価情報記憶部 1 5 に記憶されている評価情報を用いて、目的とする文体にふさわしい表現になっているかどうかを評価する。この評価情報は、評価処理部 1 2 から呼び出される評価関数のようなものでも、また評価関数が使用するパラメータのようなものでもよい。表現の各候補について、評価結果が数値（評価値）として算出されることになる。ステップ S 1 4 では、評価結果の最もよい変形後の表現を選択し、その変形した表現を変換結果文 1 3 として出力する。

【 0 0 2 9 】以下、各種の文章語文体変換処理システムへの具体的な適用例を説明する。

【 0 0 3 0 】（ A ）難解文変換システムへの適用例
図 3 は、難解文変換システムへの適用例を示している。図 3 に示す難解文変換システムでは、法律文章を平易な文に書き換えたり、難しい新聞の記事を小学生向けの易しい文に書き換えたりする処理を行う。

【 0 0 3 1 】例えば図 3 の例のように、変換対象文 1 0 として、「大臣を罷免する」という文が入力されたとする。変形処理部 1 1 は、この変換対象文 1 0 を、変形規則記憶部 1 4 にあらかじめ用意された変形規則を用いて、同義の異なる表現に言い換える。ここで、変形規則として、「罷免する」「やめさせる」・・・という規則があったとすると、変形処理部 1 1 は、変換対象文 1 0 に変形規則を適用することにより、「大臣を罷免する」という文から「大臣をやめさせる」という文を生成する。この他にも、種々の変形規則が存在し、多くの変形された文が候補として生成されることになる。これらの文を評価処理部 1 2 に渡す。なお、変形されなかった変換対象文 1 0 についても候補の一つとして評価処理部 1 2 に渡す。

【 0 0 3 2 】評価処理部 1 2 は、評価情報記憶部 1 5 にあらかじめ用意された評価情報（評価関数）を用いて、変形処理部 1 1 が変形した文を評価する。ここで評価の尺度は、例えば小学生向けというような低年齢層向けの文章集合での出現頻度または出現確率が大きくなる変換に高い評価を与えるものである。したがってこの例では、評価処理部 1 2 は、あらかじめ定められた範囲での低年齢層向けの文章集合における「大臣を罷免する」と「大臣をやめさせる」の出現頻度を求める。簡便な手法としては、変形した部分を含む小さい領域範囲の文字列が言語データで何回出現したかを数える。「大臣をやめ

させる」のほうが出現頻度が大きい場合、この表現のほうが低年齢層向けの易しい表現であるとわかる。これによりこの変形はよしとされ、変換結果文 1 3 として「大臣をやめさせる」が出力される。なお、出現頻度ではなく、出現（生起）確率を計算してもよい。

【 0 0 3 3 】また、評価の尺度としては、所定の文章集合での出現頻度や出現確率に限らず、他の何らかの尺度を用いることもできる。例えば、あらかじめ単語の結び付きや、構文解析結果から得られる文法上の言い回しに対して、評価ポイントを定めておき、それを用いて評価するようなことも可能である。また、変換後の文章集合から平易さを示す何らかの尺度を学習し、それによって評価する方法も考えられる。

【 0 0 3 4 】なお、低年齢層向けの文章集合において、評価の都度、変換の候補の出現頻度を数えるのではなく、あらかじめ文章集合に現れる各文、句、文節、単語などごとに、それらの出現頻度（出現確率）を求めておいたテーブルを用意しておき、そのテーブルを検索することによって出現頻度（出現確率）を求めるようにしてもよいことは言うまでもない。

【 0 0 3 5 】（ B ）個人文体変換システムへの適用例 図 4 は、個人文体変換システムへの適用例を示している。図 4 に示す個人文体変換システムでは、例えば芥川龍之介の小説を、夏目漱石の文体の小説に書き換えたり、ある無名の作家の小説をシェークスピアの文体の小説に書き換えたりする処理を行う。

【 0 0 3 6 】例えば図 4 の（ 1 ）の例のように、変換対象文 1 0 として、「大臣を罷免するなどを行った」という文が入力されたとする。変形処理部 1 1 は、この変換対象文 1 0 を、あらかじめ変形規則記憶部 1 4 に用意された変形規則を用いて、異なる同義の表現に言い換える。ここで、「といった」という表現を多用する人の文体への変形規則として、

「するなど」「するといったこと」
.....

という規則があったとすると、変形処理部 1 1 は、変換対象文 1 0 に変形規則を適用することにより、「大臣を罷免するなどを行った」という文から「大臣を罷免するといったことを行った」という文を生成する。この他にも、種々の変形規則が存在し、多くの変形された文が候補として生成されることになる。これらの文を評価処理部 1 2 に渡す。なお、変形されなかった変換対象文 1 0 についても候補の一つとして評価処理部 1 2 に渡す。

【 0 0 3 7 】評価処理部 1 2 は、あらかじめ評価規則記憶部 1 5 に用意された評価情報（評価関数）を用いて、変形処理部 1 1 が変形した文を評価する。ここで評価の尺度は、変換目的である特定個人の文章集合での出現頻度または出現確率が高くなるような表現に、高い評価を与えるものである。したがってこの例では、評価処理部 1 2 は、その特定個人の文章集合における「大臣を罷免

するなどを行った」という文や、「大臣を罷免するといったことを行った」という文の出現頻度を求める。なお、出現頻度は、必ずしも文全体の出現回数でなくともよく、変形した部分を含む小さい領域範囲の文字列が文章集合の中で何回出現したかでもよい。「大臣を罷免するといったことを行った」という文の出現頻度が大きい場合、評価処理部 1 2 は、変換結果文 1 3 として「大臣を罷免するといったことを行った」を出力する。

【 0 0 3 8 】また、「であろう」を多用する人への文体への変更の場合、例えば変形規則として、
「と思われる」「であろう」
.....

という規則を用い、評価の尺度として、その「であろう」を多用する特定個人の文章集合での出現頻度または出現確率が高くなるような表現をよしとするものを用いる。

【 0 0 3 9 】変形処理部 1 1 は、図 4 の（ 2 ）のように「大臣を罷免すると思われる」という変換対象文 1 0 を入力すると、この入力に対して変形規則を適用することにより、この文を「大臣を罷免するであろう」という表現に変形する。評価処理部 1 2 による評価によって、「大臣を罷免するであろう」という表現の評価値が最も高いことがわかると、評価処理部 1 2 はこの文を変換結果文 1 3 として出力する。

【 0 0 4 0 】なお、評価の尺度としては、所定の文章集合での出現頻度や出現確率に限らず、他の何らかの尺度を用いることができることは、前述したシステムの例と同様である。

【 0 0 4 1 】以上、難解文変換システムへの適用例と個人文体変換システムへの適用例を説明したが、本システムは、文章語の何らかの文体を変換するものであれば、同様に適用することができる。この場合、各変形規則は共用することも可能である。例えば作者 A の文体を作者 B の文体に変形する規則と、作者 B の文体を作者 A の文体に変形する規則とが混在している変形規則があった場合でも、評価情報を変えるだけで、同じ変形規則を用いて作者 A の文体から作者 B の文体への変換、これとは逆に作者 B の文体から作者 A の文体への変換を実現することができる。この点が従来技術と大きく異なる点である。もちろん、変形規則をそれぞれの文体の変換に適した規則に選別して用いてもよいことは言うまでもない。

【 0 0 4 2 】

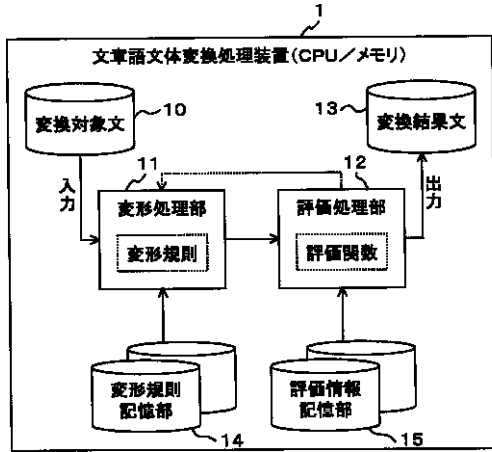
【発明の効果】以上説明したように、本発明によれば、種々の目的とする文体変換を自動的に行うことが可能になる。文体変換のための文字列の変形規則は、必ずしも目的とする文体に変換されるような方向性のある変形の規則でなくともよく、少なくとも同義性が満足されれば十分であるので、変形規則の収集・蓄積を容易に行うことが可能である。

【図面の簡単な説明】

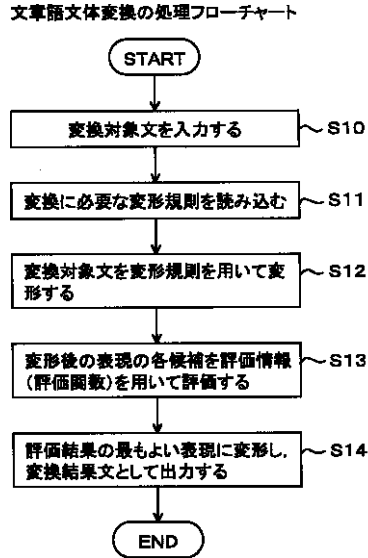
- 【図1】本発明のシステム構成例を示す図である。
- 【図2】文章語文体変換の処理フローチャートである。
- 【図3】難解文変換システムへの適用例を示す図である。
- 【図4】個人文体変換システムへの適用例を示す図である。
- 【符号の説明】

- 1 文章語文体変換処理装置
- 10 変換対象文
- 11 変形処理部
- 12 評価処理部
- 13 変換結果文
- 14 変形規則記憶部
- 15 評価情報記憶部

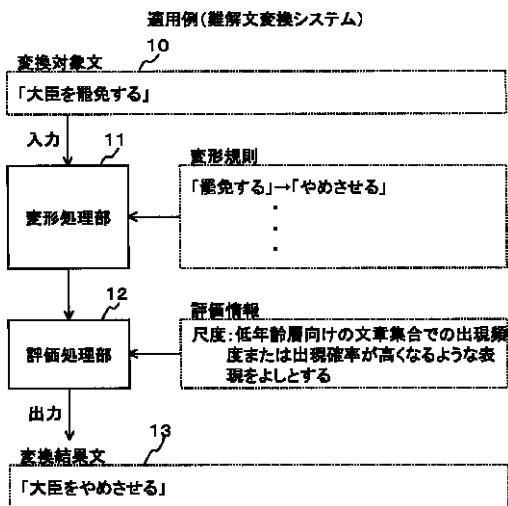
【図1】



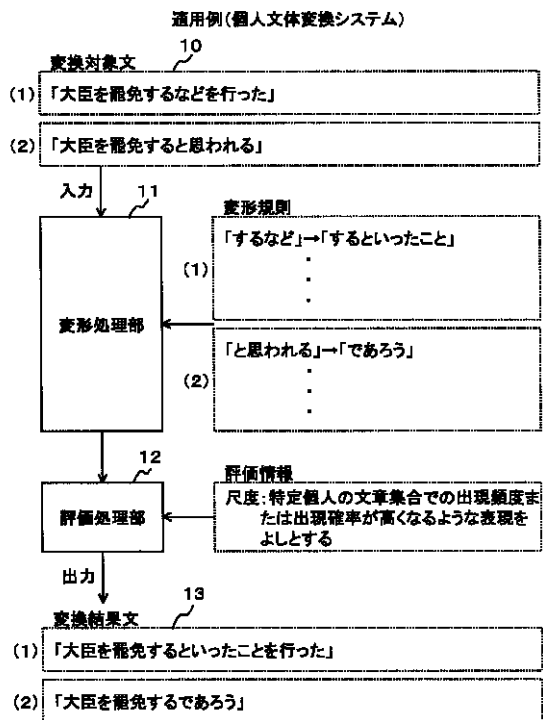
【図2】



【図3】



【図4】



【手続補正書】

【提出日】平成14年7月15日(2002.7.15)

【手続補正1】

【補正対象書類名】明細書

【補正対象項目名】特許請求の範囲

【補正方法】変更

【補正内容】

【特許請求の範囲】

【請求項1】ある自然言語で記述された文章語を、同一の自然言語で記述された他の文体による文章語に変換するシステムであって、前記自然言語で記述された第1の文字列を同義の第2の文字列に言い換える変形の規則を記憶する変形規則記憶手段と、文字列を変形した結果の表現が目的とする文体になっているかどうかを評価するための所定の評価尺度に基づく、数値情報、関数群もしくはサブルーチン群、評価方法を記述した規則、またはこれらを組み合わせた情報からなる評価情報を記憶する評価情報記憶手段と、自然言語で記述された文体変換対象の文字列を入力する入力手段と、前記入力された文字列を前記変形規則記憶手段に記憶された変形の規則を用いて変形し、複数の変換の候補を生成する変形処理手段と、前記評価情報記憶手段に記憶された評価情報に基づいて、前記変形処理手段により生成された複数の変換の候補を評価し、評価結果のよい表現を選択する評価処理手段と、前記選択された表現の変換結果を、目的とする文体に変換された文章語として出力する出力手段とを備えることを特徴とする文章語文体変換システム。

【請求項2】前記評価情報を規定する評価尺度は、変換先の文体による文章を多く含む文章集合からなる大量の用例における変換の候補の出現頻度または出現確率を含み、出現頻度または出現確率が大きいものに高い評価を与えるものであることを特徴とする請求項1記載の文章語文体変換システム。

【請求項3】前記評価情報を規定する評価尺度は、平易な文章を多く含む文章集合からなる大量の用例における変換の候補の出現頻度または出現確率を含み、出現頻度または出現確率が大きいものに高い評価を与えるものであり、前記文体の変換は、入力した難解文の文字列を平易文に変換するものであることを特徴とする請求項1記載の文章語文体変換システム。

【請求項4】前記評価情報を規定する評価尺度は、特定の個人の文章集合からなる大量の用例における変換の候補の出現頻度または出現確率を含み、出現頻度または出現確率が大きいものに高い評価を与えるものであり、前記文字列の変換は、入力した文字列の表現を前記特定の個人の文体に変換するものであることを特徴とする請求項1記載の文章語文体変換システム。

【請求項5】コンピュータによって、ある自然言語で記述された文章語を、同一の自然言語で記述された他の文体による文章語に変換するためのプログラムであって、自然言語で記述された文体変換対象の文字列を入力する処理と、あらかじめ変形規則記憶手段に記憶された、前記自然言語で記述された第1の文字列を同義の第2の文字列に言い換える変形の規則を用いて、前記入力された文字列を変形し、複数の変換の候補を生成する処理と、あらかじめ評価情報記憶手段に記憶された、文字列を変形した結果の表現が目的とする文体になっているかどうかを評価するための所定の評価尺度に基づく、数値情報、関数群もしくはサブルーチン群、評価方法を記述した規則、またはこれらを組み合わせた情報からなる評価情報に基づいて、前記変形によって生成された複数の変換の候補を評価し、評価結果のよい表現を選択する処理と、前記選択された表現の変換結果を、目的とする文体に変換された文章語として出力する処理とを、コンピュータに実行させるための文章語文体変換処理プログラム。