

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4359087号
(P4359087)

(45) 発行日 平成21年11月4日(2009.11.4)

(24) 登録日 平成21年8月14日(2009.8.14)

(51) Int.Cl.

F I

G 1 O L 13/08 (2006.01)

G 1 O L 13/08 1 1 O A

G 1 O L 13/08 1 2 7 D

G 1 O L 13/08 1 3 1 B

請求項の数 10 (全 17 頁)

(21) 出願番号 特願2003-191289 (P2003-191289)
 (22) 出願日 平成15年7月3日(2003.7.3)
 (65) 公開番号 特開2004-102242 (P2004-102242A)
 (43) 公開日 平成16年4月2日(2004.4.2)
 審査請求日 平成18年5月25日(2006.5.25)
 (31) 優先権主張番号 特願2002-205201 (P2002-205201)
 (32) 優先日 平成14年7月15日(2002.7.15)
 (33) 優先権主張国 日本国(JP)

(出願人による申告) 国等の委託研究の成果に係る特許出願(平成14年度、通信・放送機構、「高齢者・障害者のための機能代行・支援通信システム技術の研究開発」委託研究、産業活力再生特別措置法第30条の適用を受けるもの)

(73) 特許権者 000005821
 パナソニック株式会社
 大阪府門真市大字門真1006番地
 (74) 代理人 100105050
 弁理士 鷺田 公一
 (72) 発明者 ▲廣▼瀬 良文
 大阪府門真市大字門真1006番地 松下
 電器産業株式会社内
 (72) 発明者 山上 勝義
 大阪府門真市大字門真1006番地 松下
 電器産業株式会社内
 (72) 発明者 脇田 由実
 大阪府門真市大字門真1006番地 松下
 電器産業株式会社内

最終頁に続く

(54) 【発明の名称】 音声合成装置

(57) 【特許請求の範囲】

【請求項1】

複数の単語が連鎖してなる複合語を含むテキスト情報から当該複合語を抽出する抽出手段と、

抽出された複合語に、当該複合語を構成する複数の単語の意味を示す複合語意味情報を付与する付与手段と、

複合語を構成する複数の単語を含むデータを予め記憶するデータベースと、

前記データの複数の単語の少なくともいずれか一つを分離位置の決定に影響を及ぼさない単語とするワイルドカードを用いて、複合語意味情報および当該複合語意味情報に対応する分離位置を示す分離規則を構築する構築手段と、

少なくとも一つの前記分離規則を記憶する記憶手段と、

付与された複合語意味情報に基づいて、前記記憶手段に記憶された分離規則の一つを適用して、抽出された複合語を分離する位置を決定する決定手段と、

決定された分離位置に基づいてテキスト情報の韻律を生成する生成手段と、

を有することを特徴とする音声合成装置。

【請求項2】

複数の単語が連鎖してなる複合語を含むテキスト情報から当該複合語を抽出する抽出手段と、

抽出された複合語に、当該複合語を構成する複数の単語の意味を示す複合語意味情報を付与する付与手段と、

複合語を構成する複数の単語の少なくともいずれか一つを分離位置の決定に影響を及ぼさない単語とするワイルドカードを用いて構築された、複合語意味情報および当該複合語意味情報に対応する分離位置を示す分離規則を記憶する記憶手段と、

付与された複合語意味情報に基づいて、前記記憶手段に記憶された分離規則の一つを適用して、抽出された複合語を分離する位置を決定する決定手段と、

決定された分離位置に基づいてテキスト情報の韻律を生成する生成手段と、
を有することを特徴とする音声合成装置。

【請求項 3】

前記決定手段は、

前記記憶手段に記憶された複数の分離規則が適用可能である場合、用いられたワイルドカード数に基づいて、当該複数の分離規則の中から適用する分離規則を選択する、

ことを特徴とする請求項 1 又は 2 請求項記載の音声合成装置。

10

【請求項 4】

複数の単語が連鎖してなる複合語を含むテキスト情報から当該複合語を抽出する抽出手段と、

抽出された複合語に、当該複合語を構成する複数の単語の意味を示す複合語意味情報を付与する付与手段と、

複合語の分離位置を含むデータを予め記憶するデータベースと、

前記データベースに記憶されたデータから一つのデータを選択し、選択されたデータと同一の複合語意味情報を含むデータを前記データベースに記憶されたデータの中から収集し、収集されたデータの数に対する選択されたデータと同一の分離位置を含むデータの数の割合の信頼度を取得し、取得された割合に基づいて前記データを規則化し、規則化されたデータを用いて、複合語意味情報および当該複合語意味情報に対応する分離位置を示す分離規則を構築する構築手段と、

20

少なくとも一つの前記分離規則を記憶する記憶手段と、

付与された複合語意味情報に基づいて、前記記憶手段に記憶された分離規則の一つを適用して、抽出された複合語を分離する位置を決定する決定手段と、

決定された分離位置に基づいてテキスト情報の韻律を生成する生成手段と、

を有することを特徴とする音声合成装置。

30

【請求項 5】

前記決定手段は、

前記記憶手段に記憶されている複数の分離規則が適用可能である場合、取得された割合の高い規則から、当該複数の分離規則の中から適用する分離規則を選択する、

ことを特徴とする請求項 4 記載の音声合成装置。

【請求項 6】

前記決定手段は、

付与された複合語意味情報と完全一致する複合語意味情報が前記記憶手段に記憶されていない場合、前記記憶手段に記憶されかつ付与された複合語意味情報と部分一致する複合語意味情報に対応する分離位置を、抽出された複合語の分離位置とする、

ことを特徴とする請求項 1 から 5 のいずれかに記載の音声合成装置。

40

【請求項 7】

前記決定手段は、

付与された複合語意味情報と完全一致する複合語意味情報が前記記憶手段に記憶されていない場合、付与された複合語意味情報の一部を任意の意味に置換することにより、前記記憶手段に記憶されかつ付与された複合語意味情報と部分一致する複合語意味情報を検索する、

ことを特徴とする請求項 6 記載の音声合成装置。

【請求項 8】

複数の単語が連鎖してなる複合語を含むテキスト情報から当該複合語を抽出する抽出ステップと、

50

抽出した複合語に、当該複合語を構成する複数の単語の意味を示す複合語意味情報を付与する付与ステップと、

付与した複合語意味情報に基づいて、データベースに記憶された複合語を構成する複数の単語の少なくともいずれか一つを分離位置の決定に影響を及ぼさない単語とするワイルドカードを用いて構築された、複合語意味情報および当該複合語意味情報に対応する分離位置を示しかつ記憶手段に記憶された分離規則の一つを適用して、抽出した複合語を分離する位置を決定する決定ステップと、

決定した分離位置に基づいてテキスト情報の韻律を生成する生成ステップと、
を有することを特徴とする音声合成方法。

【請求項 9】

コンピュータに、

複数の単語が連鎖してなる複合語を含むテキスト情報から当該複合語を抽出する抽出機能と、

抽出した複合語に、当該複合語を構成する複数の単語の意味を示す複合語意味情報を付与する付与機能と、

付与した複合語意味情報に基づいて、データベースに記憶された複合語を構成する複数の単語の少なくともいずれか一つを分離位置の決定に影響を及ぼさない単語とするワイルドカードを用いて構築された、複合語意味情報および当該複合語意味情報に対応する分離位置を示しかつ記憶手段に記憶された分離規則の一つを適用して、抽出した複合語を分離する位置を決定する決定機能と、

決定した分離位置に基づいてテキスト情報の韻律を生成する生成機能と、
を実現させるための音声合成用プログラム。

【請求項 10】

コンピュータに、

複数の単語が連鎖してなる複合語を含むテキスト情報から当該複合語を抽出する抽出機能と、

抽出した複合語に、当該複合語を構成する複数の単語の意味を示す複合語意味情報を付与する付与機能と、

付与した複合語意味情報に基づいて、データベースに記憶された複合語を構成する複数の単語の少なくともいずれか一つを分離位置の決定に影響を及ぼさない単語とするワイルドカードを用いて構築された、複合語意味情報および当該複合語意味情報に対応する分離位置を示しかつ記憶手段に記憶された分離規則の一つを適用して、抽出した複合語を分離する位置を決定する決定機能と、

決定した分離位置に基づいてテキスト情報の韻律を生成する生成機能と、
を実現させるための音声合成用プログラムが記録された記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、入力されたテキスト情報から、テキスト情報に応じた音声を合成する音声合成装置に関する。

【0002】

【従来の技術】

例えば新聞記事などのテキスト情報においては、複数の名詞が連続して並べられることにより 1 つの名詞を構成する複合語が頻繁に登場する。テキスト情報を音声情報に変換して音声を出力する音声合成装置において、複合語全体で 1 つのアクセント句を形成した場合、複合語をひと続きに発声するため、合成音が聞き取り難くなる。

【0003】

このため、従来の音声合成装置では、複合語の構成要素（複合語を構成する単語）の間に「助詞」または「ポーズ」を挿入する、ということを行っている（例えば、特許文献 1 参照）。

10

20

30

40

50

【 0 0 0 4 】

【特許文献 1】

特開平 1 0 - 2 2 8 4 7 1 号公報

【 0 0 0 5 】

【発明が解決しようとする課題】

しかしながら、従来の音声合成装置では、複合語に「助詞」または「ポーズ」を挿入する規則は、(1) 目的語の後には「を」を挿入したり(2) 副詞の後には「に」を挿入したりするなどの単純な規則であるが、実際には、目的語のあとに必ず「を」を挿入するとは限らず、また、すべての目的語のあとに「ポーズ」が挿入されるとは限らないので、複合語の表現を変換する精度が、十分ではないという問題があった。

10

【 0 0 0 6 】

本発明は、かかる点に鑑みてなされたものであり、十分な精度で複合語を表現することができる音声合成装置を提供することを目的とする。

【 0 0 0 7 】

【課題を解決するための手段】

本発明の音声合成装置は、複数の単語が連鎖してなる複合語を含むテキスト情報から当該複合語を抽出する抽出手段と、抽出された複合語に、当該複合語を構成する複数の単語の意味を示す複合語意味情報を付与する付与手段と、複合語意味情報および当該複合語意味情報に対応する分離位置を示す少なくとも一つの分離規則を予め記憶する記憶手段と、付与された複合語意味情報に基づいて、前記記憶手段に記憶された分離規則の一つを適用して、抽出された複合語を分離する位置を決定する決定手段と、決定された分離位置に基づいてテキスト情報の韻律を生成する生成手段と、を有する構成を採る。

20

【 0 0 0 8 】

この構成によれば、テキスト情報から抽出された複合語を構成する複数の単語の意味を示し複合語に付与された複合語意味情報に基づいて、抽出された複合語を分離する位置を決定し、決定された分離位置に基づいて、テキスト情報の韻律を生成するため、テキスト情報に含まれる複合語を適切な位置でアクセント句毎に分離することができ、複合語の構造を理解しやすい合成音を生成することができ、十分な精度で複合語を表現することができる。

【 0 0 0 9 】

本発明の音声合成方法は、複数の単語が連鎖してなる複合語を含むテキスト情報から当該複合語を抽出する抽出ステップと、抽出した複合語に、当該複合語を構成する複数の単語の意味を示す複合語意味情報を付与する付与ステップと、付与した複合語意味情報に基づいて、複合語意味情報および当該複合語意味情報に対応する分離位置を示しかつ記憶手段に記憶された分離規則の一つを適用して、抽出した複合語を分離する位置を決定する決定ステップと、決定した分離位置に基づいてテキスト情報の韻律を生成する生成ステップと、を有するようにした。

30

【 0 0 1 0 】

この方法によれば、テキスト情報から抽出された複合語を構成する複数の単語の意味を示し複合語に付与された複合語意味情報に基づいて、抽出された複合語を分離する位置を決定し、決定された分離位置に基づいて、テキスト情報の韻律を生成するため、テキスト情報に含まれる複合語を適切な位置でアクセント句毎に分離することができ、複合語の構造を理解しやすい合成音を生成することができ、十分な精度で複合語を表現することができる。

40

【 0 0 1 1 】

本発明の音声合成用プログラムは、コンピュータに、複数の単語が連鎖してなる複合語を含むテキスト情報から当該複合語を抽出する抽出機能と、抽出した複合語に、当該複合語を構成する複数の単語の意味を示す複合語意味情報を付与する付与機能と、付与した複合語意味情報に基づいて、複合語意味情報および当該複合語意味情報に対応する分離位置を示しかつ記憶手段に記憶された分離規則の一つを適用して、抽出した複合語を分離する

50

位置を決定する決定機能と、決定した分離位置に基づいてテキスト情報の韻律を生成する生成機能と、を実現させるようにした。

【 0 0 1 2 】

このプログラムによれば、テキスト情報から抽出された複合語を構成する複数の単語の意味を示し複合語に付与された複合語意味情報に基づいて、抽出された複合語を分離する位置を決定し、決定された分離位置に基づいて、テキスト情報の韻律を生成するため、テキスト情報に含まれる複合語を適切な位置でアクセント句毎に分離することができ、複合語の構造を理解しやすい合成音を生成することができ、十分な精度で複合語を表現することができる。

【 0 0 1 3 】

本発明の記録媒体は、コンピュータに、複数の単語が連鎖してなる複合語を含むテキスト情報から当該複合語を抽出する抽出機能と、抽出した複合語に、当該複合語を構成する複数の単語の意味を示す複合語意味情報を付与する付与機能と、付与した複合語意味情報に基づいて、複合語意味情報および当該複合語意味情報に対応する分離位置を示しかつ記憶手段に記憶された分離規則の一つを適用して、抽出した複合語を分離する位置を決定する決定機能と、決定した分離位置に基づいてテキスト情報の韻律を生成する生成機能と、を実現させる音声合成用プログラムが記録されるようにした。

【 0 0 1 4 】

この記録媒体によれば、テキスト情報から抽出された複合語を構成する複数の単語の意味を示し複合語に付与された複合語意味情報に基づいて、抽出された複合語を分離する位置を決定し、決定された分離位置に基づいて、テキスト情報の韻律を生成するため、テキスト情報に含まれる複合語を適切な位置でアクセント句毎に分離することができ、複合語の構造を理解しやすい合成音を生成することができ、十分な精度で複合語を表現することができる。

【 0 0 1 5 】

【発明の実施の形態】

以下、本発明の実施の形態について、図面を参照して詳細に説明する。

【 0 0 1 6 】

本発明においては、意味の類似した複合語はその複合語の構造の面においても類似しているという特徴を利用している。各要素単語の意味情報の連鎖パターンを用いて、複合語のアクセント句を分離する（すなわち、ポーズを挿入する）規則の自動構築を行う。

【 0 0 1 7 】

（実施の形態 1）

本発明の実施の形態 1 に係る音声合成装置の構成を、図 1 を参照しながら説明する。

【 0 0 1 8 】

1 1 は入力されたテキスト情報を単語に分割する形態素解析部であり、1 2 は分割された単語列から複合語を形成する単語列を抽出する複合語抽出部であり、1 3 は単語に意味情報を付与する意味情報付与部であり、1 4 は分離規則を参照して、入力された複合語の分離位置を決定する規則適用部であり、1 5 は分離位置（分離されたアクセント句）にしたがって韻律を生成する韻律生成部であり、1 6 は生成された韻律に基づいて入力テキストの合成音を生成する音声合成部であり、1 7 は入力されたテキスト情報を単語単位に分割するための単語辞書部であり、1 8 は単語の意味情報を保持するシソーラス辞書部であり、1 9 は分離規則辞書部である。

【 0 0 1 9 】

以上のように構成された本発明の実施の形態 1 に係る音声合成装置の動作について、2 つの動作例を挙げて説明する。

【 0 0 2 0 】

第 1 の動作例では、説明を容易にするため、具体的なテキスト情報として、「日本政府当局者は、1 日...」を本実施の形態に係る音声合成装置において合成して読み上げる場合を挙げて説明する。

10

20

30

40

50

【 0 0 2 1 】

まず、形態素解析部 1 1 では、入力されたテキスト情報を、単語辞書部 1 7 を用いて形態素解析を行い単語に分割する。「日本政府当局者は、1 日...」というテキストに対して形態素解析を行う場合、「日本・政府・当局・者・は・、・1・日・...」(ここで、「・」は単語分離を表す。)というように、単語単位に分割する。なお、形態素解析の具体的な方法は特に限定されない。

【 0 0 2 2 】

このように単語単位に分割されたテキスト情報は、次に複合語抽出部 1 2 に送られる。複合語抽出部 1 2 では、単語単位に分割されたテキスト情報に含まれる複合語を抽出する。抽出方法としては、例えば、名詞、接頭語、又は接尾語が連続してなる単語列を抽出する
10
方法が挙げられる。上記の例では、「日本・政府・当局・者」という複合語が抽出される。以下、抽出された複合語を「テキスト内複合語」と言う。

【 0 0 2 3 】

意味情報付与部 1 3 では、テキスト内複合語に、このテキスト内複合語を構成する各単語(ここでは、「日本」「政府」「当局」「者」)が有する意味情報を付与する。

【 0 0 2 4 】

より詳しく説明すると、シソーラス辞書部 1 8 には、「日本」という単語には「国名」という意味情報、「政府」という単語には「政府機関」という意味情報、「当局」という単語には「政府機関」という意味情報、「者」という単語には「接尾語」というように、単語と、この単語が有する意味情報とが対応づけられて記憶されている。
20

【 0 0 2 5 】

意味情報付与部 1 3 では、シソーラス辞書部 1 8 を参照しながら、テキスト内複合語を構成する各単語(「日本」「政府」「当局」「者」)の意味情報を拾い上げて複合語意味情報を生成する。この場合、「国名」「政府機関」「政府機関」「接尾語」という複合語意味情報が生成される。以下、この複合語意味情報を「テキスト内複合語意味情報」と言う。なお、意味情報の抽象度は段階的に変更できるような意味辞書を用いても良い。

【 0 0 2 6 】

テキスト内複合語意味情報は、規則適用部 1 4 に送られる。規則適用部 1 4 では、テキスト内複合語意味情報(この例では「国名」「政府機関」「政府機関」「接尾語」)が、どの意味情報と意味情報との間で分離されるべきかという分離位置(この分離位置を、「テキスト内複合語分離位置」という)を決定する。
30

【 0 0 2 7 】

これについてより詳細に説明すると、例えば、図 2 に示す通り、連続する 4 つの意味情報を順に並べて構成される複合語意味情報「進退」「運営」「接尾語」「計画」であれば、「接尾語」と「計画」との間(すなわち、3 番目と 4 番目の意味情報の間)を、連続する 4 つの意味情報を順に並べて構成される複合語意味情報「国」「会議」「接尾語」「役職」であれば、「接尾語」と「役職」との間(すなわち、3 番目と 4 番目の意味情報の間)を、連続する 4 つの意味情報を順に並べて構成される複合語意味情報「交わり」「原理」「関係」「接尾語」であれば、「原理」と「関係」との間(すなわち、2 番目と 3 番目の意味情報の間)を、また、連続する 4 つの意味情報を順に並べて構成される複合語意味情報「集合」「軍人」「会議」「文章」であれば、「会議」と「文章」との間(すなわち、3 番目と 4 番目の意味情報の間)を、それぞれ分離位置とする。
40

【 0 0 2 8 】

分離規則辞書部 1 9 には、上記の通り、意味情報が連続して並べられてそれぞれ構成されている複数個の複合語意味情報と、各複合語意味情報のうちどの意味情報と意味情報との間で複合語が分離されるべきかという分離位置とが対応づけられて記憶(記録)されている。規則適用部 1 4 では、このような分離規則辞書部 1 9 を参照しながら、テキスト内複合語分離位置の決定を行う。

【 0 0 2 9 】

分離規則辞書部 1 9 に、連続する 4 つの意味情報を順に並べて構成される複合語意味情報
50

「国名」「政府機関」「政府機関」「接尾語」であれば、分離位置は「政府機関」と「政府機関」との間（すなわち、分離位置は意味情報の2番目と3番目との間）であることが記憶（記録）されていれば、規則適用部14は、「国名」「政府機関」「政府機関」「接尾語」からなるテキスト内複合語意味情報に対応するテキスト内複合語分離位置が「政府機関」と「政府機関」との間にあるということを決定する。

【0030】

最後に、このようにして分離位置が決定された複合語を韻律生成部15に送る。韻律生成部15は入力されたテキスト情報に適切な韻律を付与し、音声合成部16に韻律情報を送る。音声合成部17は、韻律生成部16が生成した韻律に基づいて音声を合成する。この音声合成時にテキスト内複合語分離位置においてポーズがかかる。すなわち、上記の例であれば、「にほんせいふ（ポーズ）とうきょくしゃ」という音声合成され、発音される。

10

【0031】

このようにして、テキスト情報中に含まれる複合語を適切な位置でアクセント句分離可能となり、複合語の構造を理解しやすい合成音を生成できる。

【0032】

次に、第2の動作例では、説明を容易にするため、具体的なテキスト情報として、「アメリカ連邦捜査局は、1日…」を本実施の形態に係る音声合成装置において合成して読み上げる場合を挙げて説明する。

【0033】

上記第1の動作例では、分離規則辞書部19に記憶（記録）されている複合語意味情報と、テキスト内複合語意味情報とが完全一致する場合について説明したが、ここで説明する第2の動作例では、分離規則辞書部19に記憶（記録）されている複合語意味情報と、テキスト内複合語意味情報とが部分的にしか一致しない場合について説明する。

20

【0034】

形態素解析部11により単語に分割されたテキスト情報からテキスト内複合語を複合語抽出部12により抽出し、当該テキスト内複合語の各単語が有する意味情報を並べてテキスト内複合語意味情報を意味情報付与部13が生成することまでは第1の動作例と同様である。上記の例であれば、テキスト内複合語は「アメリカ」「連邦」「捜査」「局」となり、シソーラス辞書部18に「アメリカ」の意味情報は「国名」、「連邦」の意味情報は「政府機関」、「捜査」の意味情報は「調査」、「局」の意味情報は「接尾語」であることが記憶（記録）されていれば、意味情報生成部13は、テキスト内複合語意味情報として「国名」「政府機関」「調査」「接尾語」を生成する。

30

【0035】

すなわち、分離規則辞書部19には、「国名」「政府機関」「調査」「接尾語」というテキスト内複合語意味情報と完全一致する複合語意味情報が存在しない。

【0036】

この場合は、まず、テキスト内複合語意味情報の第1番目の意味情報（上記の例では「国名」）を「任意」にして、分離規則辞書部19に「任意」「政府機関」「調査」「接尾語」という複合語意味情報が存在するかどうかを判断する。説明を容易にするため、ここでは、このような複合語意味情報が存在しなかったとする。なお、本明細書および図面では、この「任意」の意味を「*」または「(*)」により表すことがある。

40

【0037】

そして、第1番目の意味情報を「国名」に戻して、第2番目の意味情報を「任意」にする。そして、分離規則辞書部19に「国名」「(*)」「調査」「接頭語」という複合語意味情報が存在するかどうかを判断する。説明を容易にするため、ここでは、このような複合語意味情報が存在しなかったと仮定する。

【0038】

そして、第2番目の意味情報を「政府機関」に戻して、第3番目の意味情報を「任意」にする。そして、分離規則辞書部19に「国名」「政府機関」「(*)」「接頭語」という

50

複合語意味情報が存在するかどうかを判断する。

【0039】

上記第1の動作例にて説明したように、分離規則辞書部19に、連続する4つの意味情報を順に並べて構成される複合語意味情報「国名」「政府機関」「政府機関」「接尾語」の分離位置は「政府機関」と「政府機関」との間であるということが記憶（記録）されているので、「国名」「政府機関」「(＊)」「接頭語」というテキスト内複合語意味情報は、分離規則辞書部19に記憶（記録）されている複合語意味情報「国名」「政府機関」「政府機関」「接尾語」と部分一致している。

【0040】

分離規則辞書部19には、複合語意味情報「国名」「政府機関」「政府機関」「接尾語」の分離位置が「政府機関」と「政府機関」との間であることが記憶（記録）されているので、「国名」「政府機関」「(＊)」「接頭語」という複合語意味情報も「政府機関」と「(＊)」との間にあると決定する。これ以後の動作は、上記第1の動作例と同様である。

10

【0041】

したがって、第2の動作例によれば、「あめりかれんぼう（ポーズ）そうさきよく」という音声合成され、発音される。

【0042】

この第2の動作例によれば、分離規則辞書部19にテキスト内複合語意味情報と完全に一致する複合語意味情報が存在しない場合であっても、おおよそ正しい分離位置を推定することができる。このように、第2の動作例においても、テキスト情報中に含まれる複合語を適切な位置でアクセント句毎に分離可能となり、複合語の構造を理解しやすい合成音を生成することができる。

20

【0043】

なお、テキスト内複合語意味情報の1つの意味情報を「(＊)」にしても、分離規則辞書部19に記憶（記録）されている複合語意味情報の中から意味情報が合致する複合語意味情報が存在しない場合には、「(＊)」を2つ乃至3つに増やし、テキスト内複合語意味情報の複数の意味情報を「(＊)」にして、分離規則辞書部19に記憶（記録）されている複合語意味情報の中から意味情報が合致する複合語意味情報と合致するかどうかを判断しても良い。

30

【0044】

また、テキスト内複合語意味情報の1つの意味情報を「(＊)」にすることにより、分離規則辞書部19に記憶（記録）されている複合語意味情報の中から意味情報が合致する複合語意味情報が複数個存在する場合には、これらの合致する複数個の複合語意味情報の中の一つを適宜選択しても良い。

【0045】

また、テキスト内複合語意味情報の全部の意味情報を「(＊)」にしない限り、分離規則辞書部19に記憶（記録）されている複合語意味情報の中から意味情報が合致する複合語意味情報が存在しない場合がある。この場合には、致し方ないが、「分離位置なし」としてポーズなしで音声を合成し、発音させることにする。または、予め決定しておいた位置で分離し、音声を合成しても良い。

40

【0046】

このように、本実施の形態によれば、テキスト情報中に含まれる複合語を適切な位置でアクセント句毎に分離可能となり、複合語の構造を理解しやすい合成音を生成することができる。

【0047】

（実施の形態2）

図3は、本発明の実施の形態2に係る音声合成装置の構成を示す図である。なお、本実施の形態に係る音声合成装置は、実施の形態1において説明した音声合成装置と同様の基本的構成を有しており、同一の構成要素には同一の参照符号を付し、その説明を省略する。

50

【 0 0 4 8 】

図 3 に示す音声合成装置は、図 1 に示す音声合成装置の構成要素に加えて、複合語のアクセント句分離規則を構築する分離規則構築部 2 1 と、複合語のアクセント句分離規則学習用の学習データベース (D B) 2 2 とを有する。

【 0 0 4 9 】

ここでは、本実施の形態の特徴である、音声合成装置の分離規則構築部 2 1 における分離位置の学習方法について、説明する。

【 0 0 5 0 】

分離位置を学習する際には、学習 D B 2 2 において、図 2 に示すような D B が用いられる。この D B に記憶される各データは、複合語を構成する単語と、単語が保有する意味情報 (図中括弧内の情報) と、分離位置を示す分離位置情報とを持つ。例えば分離すべきでない複合語には、分離位置情報として「 0 」が付与される。

10

【 0 0 5 1 】

なお、分離位置情報は 1 つの複合語に対して複数の分離位置情報を保持しても良い。

【 0 0 5 2 】

分離規則の自動構築は、この学習 D B 2 2 を用いて行われる。構築方法を以下に説明する。ここで、本発明における規則とは、複合語を構成する各単語の意味情報の連鎖パターンによってアクセント句毎に分離する分離規則である。例えば、「 C 1 」 「 C 2 」 「 C 3 」 「 C 4 」 という意味情報の連鎖パターンにおいて (C 1 ~ C 4 は意味情報を表す) 、 「 C 2 」 と 「 C 3 」 の間でアクセント句毎に分離する、という規則がある。

20

【 0 0 5 3 】

分離規則構築のフローを図 4 に示す。まず、ステップ S 2 0 0 では、学習 D B 2 2 から 1 つのデータを選択する。そして、ステップ S 3 0 0 では、ステップ S 2 0 0 で選択した意味情報の連鎖パターンと同一のパターンを持つデータを収集する。

【 0 0 5 4 】

そして、ステップ S 4 0 0 では、ステップ S 3 0 0 でデータ数が増加したか否かを判断する。この判断の結果として、データ数に変化がない場合 (S 4 0 0 : N O) は、ステップ S 2 0 0 に戻る。一方、データ数が増加していた場合 (S 4 0 0 : Y E S) は、ステップ S 5 0 0 に進む。

【 0 0 5 5 】

30

ステップ S 5 0 0 では、ステップ S 3 0 0 で集めたデータのうち、ステップ S 2 0 0 で選んだデータと同一の位置に分離位置があるデータの割合 (選択されたデータにおける分離位置の信頼度) を算出し、算出された割合 (信頼度) を予め設定された閾値と比較する。この比較の結果として、同一の分離位置である割合が閾値以上である場合 (S 5 0 0 : Y E S) は、ステップ S 6 0 0 に進む。一方、閾値未満である場合 (S 5 0 0 : N O) は、ステップ S 7 0 0 に進む。

【 0 0 5 6 】

ステップ S 6 0 0 では、ステップ S 2 0 0 で選択したデータにおける意味情報の連鎖パターンを規則化する。そして、ステップ S 7 0 0 では、学習 D B 2 2 における全てのデータについてステップ S 2 0 0 からステップ S 6 0 0 を処理したか否かを判断し、まだ処理していないデータが存在する場合 (S 7 0 0 : N O) は、ステップ S 2 0 0 に戻る。一方、全てのデータを処理した場合 (S 7 0 0 : Y E S) は、分離規則構築を終了する。

40

【 0 0 5 7 】

以上のように構築した分離規則の適用例を説明する。図 5 の (A) に示すように、「韓国政府当局者」と「日本政府当局者」というデータが学習 D B 2 2 に存在した場合、各データの意味情報の連鎖パターンである「 (国名) (政府機関) (政府機関) (接尾語) 」に対する分離位置が全て「 2 」であるため、「 (国名) (政府機関) / (政府機関) (接尾語) 」を規則化する。ここで、「 / 」は分離位置を表す。一方、図 5 の (B) に示すように、未知データとして「アメリカ政府当局者」という複合語が入力された場合には上記規則化された「 (国名) (政府機関) / (政府機関) (接尾語) 」における意味情報の連鎖パ

50

タンと一致するため、この規則を適用し、「アメリカ政府 / 当局者」というように「アメリカ政府」の直後で適切に分離することが可能となる。

【 0 0 5 8 】

なお、このような規則を構築する場合、最も精度が高い分離規則は、あらゆる組合せの意味情報の連鎖パターンについて全て網羅的に規則化することにより得られる。しかし、網羅的に規則化することはデータ量から見て容易ではない。例えば、4 単語から構成される複合語の場合、 N_c^4 パターン存在し（意味情報の種類数を N_c とする）、このような規則を学習するために学習 DB 2 2 を構築することは容易ではない。

【 0 0 5 9 】

そこで、本発明では、以下に示す方法によりアクセント句分離の決定に影響を及ぼさない単語を、全ての単語を含む意味クラス（以下「ワイルドカード」と言う）と見なす。このワイルドカードを用いることにより、高い精度を保ちながら規則の適用範囲を拡張することが可能となる。

【 0 0 6 0 】

ワイルドカードを導入した場合、複合語を構成する単語毎に、ワイルドカード化するかどうかを判断するため、 N 個の単語の連鎖（以下「 N 連鎖」と言う）の複合語の場合、 $(2^N - 1)$ 通りの適用のパターンがある。図 6 に、 $N = 3$ 、 4 、 5 のときのワイルドカード適用パターンについて示す。なお、 $N > 5$ の場合も同様にパターンを設定することが可能である。例えば、ワイルドカードを適用した場合、「 C_1 」「 C_2 」「 $*$ 」「 $*$ 」（先頭が C_1 、2 番目が C_2 で、3 番目、4 番目はどの単語でも良い）という連鎖パターンを持つ複合語は C_1 の後で分離する、という規則を構築することになる。

【 0 0 6 1 】

なお、ワイルドカード適用のパターン数を削減するために、ワイルドカード化するパターンに制約を設けても良い。より具体的には、連続した単語をワイルドカードとみなさない、という制約を設けても良い。例えば、「 $*$ 」「 C_1 」「 C_2 」「 $*$ 」はワイルドカード適用パターンとして有効であるが、「 $*$ 」「 $*$ 」「 C_1 」「 C_2 」は無効である。ここで、記号「 $*$ 」はワイルドカードを適用した単語を示す。

【 0 0 6 2 】

以下、ワイルドカードを用いて N 連鎖の複合語の分離規則を構築する動作について、図 7 を参照しながら説明する。

【 0 0 6 3 】

まず、ステップ S 1 0 0 では、図 6 に示すワイルドカードの適用パターンの中から 1 つを選び、選択されたワイルドカードを学習 DB 2 2 の全データに適用する。

【 0 0 6 4 】

例えば、0 個のワイルドカードを適用するときは、学習 DB 2 2 に記憶されている各データにおけるどの単語にも適用しない。また、1 個のワイルドカードを適用するときは、学習 DB 2 2 に記憶されている各データにおけるいずれか 1 つの単語に適用する。つまり、4 連鎖の複合語のデータの場合は、4 つのパターンの適用が可能である。また、2 個のワイルドカードを適用するときは、学習 DB 2 2 に記憶されている各データにおけるいずれか 2 つの単語に適用する。つまり、4 連鎖の複合語のデータの場合は、6 つのパターンの適用が可能である。

【 0 0 6 5 】

そして、ステップ S 2 0 0 では、学習 DB 2 2 から 1 つのデータを選択する。例えば、「（国名）（政府機関） / （政府機関）（接尾語）」という意味情報の連鎖パターンを有するデータが選択される。

【 0 0 6 6 】

そして、ステップ S 3 0 0 では、ステップ S 2 0 0 で選択した意味情報の連鎖パターンと同一のパターンを持つデータを収集する。

【 0 0 6 7 】

例えば、図 8 に示す通り、データ # 1 ~ # 4 を記憶している学習 DB 2 2 において 0 個の

10

20

30

40

50

ワイルドカードが適用されている場合、上記選択されたデータと同一のパターンを持つデータは、データ# 1、# 2である。一方、1個のワイルドカードが先頭の単語に適用されている場合はデータ# 1 ~ # 3が、また、2個のワイルドカードが先頭および3番目の単語に適用されている場合はデータ# 1 ~ # 4が、それぞれ上記選択されたデータと同一のパターンを持つ。

【0068】

そして、ステップS400では、ステップS300でデータ数が増加したか否かを判断する。この判断の結果として、データ数に変化がない場合(S400:NO)は、ステップS200に戻る。一方、データ数が増加していた場合(S400:YES)は、ステップS500に進む。

10

【0069】

例えば、図8において、1個のワイルドカードが適用されたときは、0個のワイルドカードが適用されたときと比べて、収集されたデータ数が1つ増加している。また、2個のワイルドカードが適用されたときは、0個のワイルドカードが適用されたときと比べて、収集されたデータ数が2つ増加している。

【0070】

ステップS500では、ステップS300で集めたデータのうち、ステップS200で選んだデータと同一の位置に分離位置があるデータの割合(選択されたデータにおける分離位置の信頼度)を算出し、算出された割合(信頼度)を予め設定された閾値と比較する。

【0071】

20

例えば、図8において、1個のワイルドカードが適用されているときに収集された各データ# 1 ~ # 3の分離位置が上記選択されたデータの分離位置と一致する割合、つまり選択されたデータの分離位置の信頼度は100%である。また、2個のワイルドカードが適用されているときに収集された各データ# 1 ~ # 4の分離位置が上記選択されたデータの分離位置と一致する割合、つまり選択されたデータの分離位置の信頼度は100%である。

【0072】

この比較の結果として、同一の分離位置である割合が閾値以上である場合(S500:YES)は、ステップS600に進む。一方、閾値未満である場合(S500:NO)は、ステップS700に進む。

【0073】

30

ステップS600では、ステップS200で選択した意味情報の連鎖パターンを規則化する。

【0074】

ステップS700では、学習DB22における全てのデータについてステップS200からステップS600を処理したか否かを判断し、まだ処理していないデータが存在する場合(S700:NO)は、ステップS200に戻る。一方、全てのデータを処理した場合(S700:YES)は、ステップS800に進む。

【0075】

ステップS800では、図6に示したワイルドカードの適用パターンの全てについて処理を実行したか否かを判断する。判断の結果として、まだ適用していないワイルドカードのパターンが存在する場合(S800:NO)にはステップS100を実行する。一方、全てのワイルドカードのパターンについて処理した場合(S800:YES)は、動作を終了する。

40

【0076】

なお、図6では、N連鎖の複合語のワイルドカード適用パターンを($2^N - 1$)通りとしているが、全ての構成要素にワイルドカードを適用し、どの規則にも適用されなかった複合語に対する分離位置を設定するようにしても良い。

【0077】

また、規則の信頼度を高めるために、規則化するときに最低限必要なデータの出現頻度を設定するようにしても良い。例えば、「データ中に3回以上出現した連鎖パターンのみ規則

50

化する」というような制約を設けても良い。

【 0 0 7 8 】

このようにして、N連鎖の複合語であればどのような複合語であっても、規則構築を行い、規則化することが可能である。

【 0 0 7 9 】

ワイルドカードを導入することにより、図9の(A)および(B)に示すように意味情報の連鎖パタンの一部が異なるような複合語が入力された場合にも、精度良く規則を適用することが可能となる。

【 0 0 8 0 】

図9の(A)には、学習DB22に存在する「韓国政府当局者」と「ロシア外務省筋」という複合語から「(国名)(政府機関)/(*)(接尾語)」という分離規則を構築した例が示されている。図9の(B)に示すように、未知データである「アメリカ連邦捜査局」という複合語が入力された場合には、この複合語の意味情報の連鎖パタンは「(国名)(政府機関)/(調査)(接尾語)」である。上記二つの連鎖パタンにおいてそれぞれ3番目に位置している意味情報は(調査)と(政府機関)であり互いに異なるが、ワイルドカードの適用により「アメリカ連邦」の直後で適切に分離し、「アメリカ連邦/捜査局」とすることが可能となる。

10

【 0 0 8 1 】

ところで、ワイルドカードを導入した場合、1つの連鎖パタンに対して適用可能な規則が複数存在することがある。このとき、2つの基準に基づいて、適用可能な複数の規則の中から1つを選択する。

20

【 0 0 8 2 】

ここで用いられる第1の基準は、ワイルドカードを適用した単語数が少ない方の規則を優先的に選択する、という基準である。図10に示す通り、一般的には、適用されるワイルドカード数が少ない規則の方が高い精度を有する一方、適用されるワイルドカード数が多い規則の方が高いカバレッジを有する。そこで、例えば、「(国名)(政府機関)/(*)(*)」という規則と「(国名)/(*)(*)(*)」という規則が適用可能である場合、ワイルドカードが少ない前者の規則を選択し、適用する。

【 0 0 8 3 】

また、ここで用いられる第2の基準は、規則を構築したときの信頼度が高い方の規則を優先的に選択する、という基準である。例えば、信頼度が80%である「(*)(政府機関)/(*)(*)」という規則と、信頼度が75%である「(国名)/(*)(*)(*)」という規則が適用可能である場合、信頼度が高い前者の規則を選択し、適用する。

30

【 0 0 8 4 】

上記の基準を予め設定しておくことにより、1つの連鎖パタンに対して適用可能な規則が複数存在する場合においても最適な分離規則を適用することができる。

【 0 0 8 5 】

なお、これらの基準に基づく分離規則の優先順位は、各分離規則を構築するときに決定されるものである。このため、全ての分離規則をこの優先順位に基づいて予め整列しておき、分離規則が適用可能な否かを、優先順位の高い規則から順に判断することにより、最初に適用可能と判断された分離規則を最適な分離規則として適用することが可能になる。この場合、各分離規則の信頼度を保持しておく必要性がなくなる。

40

【 0 0 8 6 】

以上の説明では、複合語のアクセント句の分離規則について述べたが、本実施の形態により、複合語のアクセント句の分離位置を決定するだけでなく、複合語を構成する各単語間に挿入する語句(例えば、助詞など)を決定する規則も構築することが可能である。

【 0 0 8 7 】

以下に、単語間に語句を挿入する例について説明する。学習DB22として、図11に示す通り、分離位置情報とその分離位置に挿入すべき語句を記述したDBを用意する。例えば、「国家・評議・会・議長」という複合語の場合、「国家評議会の議長」というように

50

助詞を挿入するという変換を行うことができるように、分離位置（３番目）に助詞「の」を挿入するということを示す情報を持つ学習ＤＢ２２を構築する。

【００８８】

規則の構築方法は、前述の方法と同じ方法で構築が可能であるが、上記フローのステップＳ２０４において、分離位置だけでなく挿入される助詞の種類も一致する割合（信頼度）を計算するように変更すれば、上記のＤＢを構築することが可能である。

【００８９】

なお、本発明は上記の実施の形態に限定されず、様々な実施の形態に実施することが可能である。例えば、上記実施の形態では、音声合成方法を音声合成装置で実行しているが、コンピュータ上でソフトウェアとして実行しても良い。

【００９０】

例えば、上記実施の形態で説明した音声合成方法を実行するプログラムを予め例えばＲＯＭ（Read Only Memory）などの記録媒体に記録しておき、そのプログラムをＣＰＵ（Central Processor Unit）によって動作させるようにしても良い。

【００９１】

【発明の効果】

以上説明したように、本発明によれば、上記の複合語分離方法を用いることにより、音声合成装置において、テキスト情報中に含まれる複合語を適切な位置でアクセント句毎の分離を行うことが可能となり、複合語の構造を理解しやすい合成音を生成することができ、十分な精度で複合語を表現することができるので実用上の価値は大きい。

【図面の簡単な説明】

【図１】本発明の実施の形態１に係る音声合成装置の構成を示す図

【図２】本発明の実施の形態１に係るテキスト内複合語分離位置を説明するための図

【図３】本発明の実施の形態２に係る音声合成装置の構成を示す図

【図４】本発明の実施の形態２において、分離規則構築の動作を説明するためのフロー図

【図５】本発明の実施の形態２において、構築された分離規則を未知データに適用した例を説明するための図

【図６】本発明の実施の形態２に係るワイルドカード適用パターンを示す図

【図７】本発明の実施の形態２において、ワイルドカードを用いた分離規則構築の動作を説明するためのフロー図

【図８】本発明の実施の形態２において、ワイルドカードを用いた分離規則構築の動作の具体例を説明するための図

【図９】本発明の実施の形態２において、ワイルドカードを用いて構築された分離規則を未知データに適用した例を説明するための図

【図１０】本発明の実施の形態２において、適用する分離規則を選択するときの第１の基準を説明するための図

【図１１】本発明の実施の形態２に係るテキスト内複合語分離位置および挿入語句を説明するための図

【符号の説明】

１１ 形態素解析部

１２ 複合語抽出部

１３ 意味情報付与部

１４ 規則適用部

１５ 韻律生成部

１６ 音声合成部

１７ 単語辞書部

１８ シソーラス辞書部

１９ 分離規則辞書部

２１ 分離規則構築部

２２ 学習データベース

10

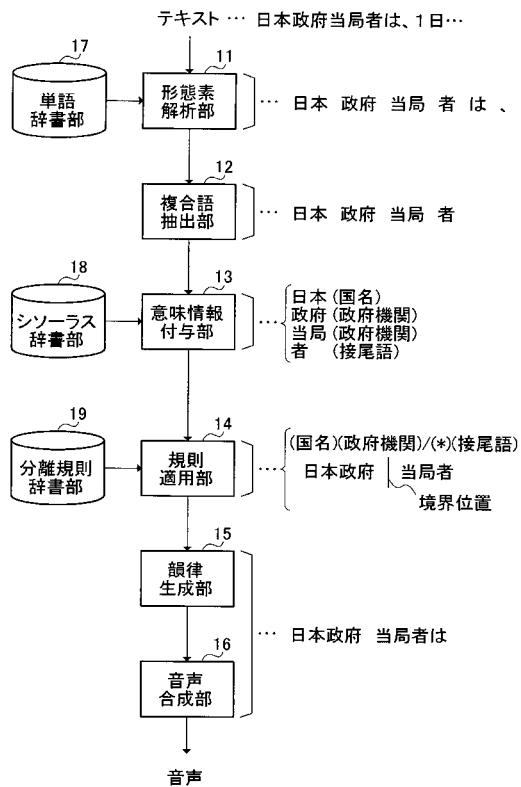
20

30

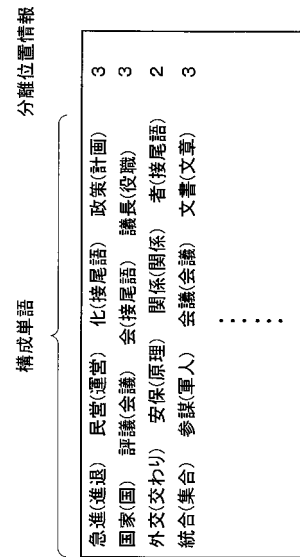
40

50

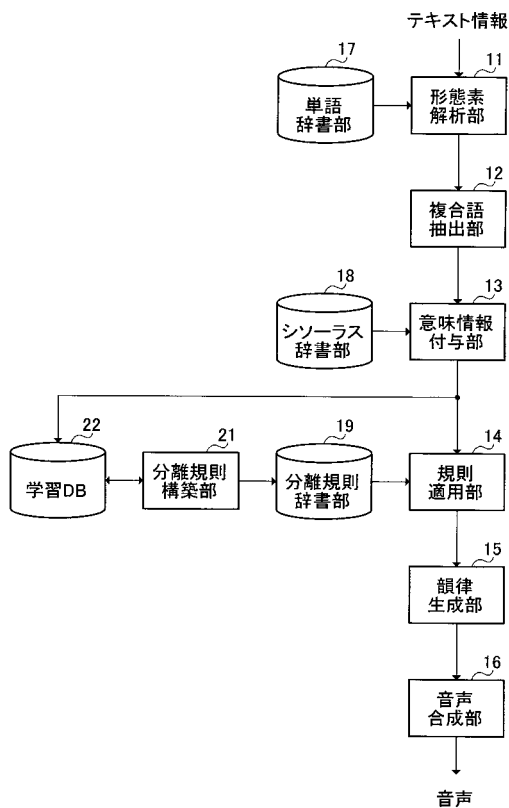
【図 1】



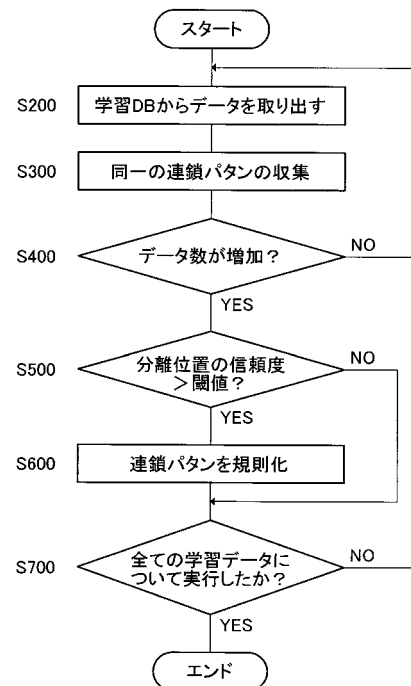
【図 2】



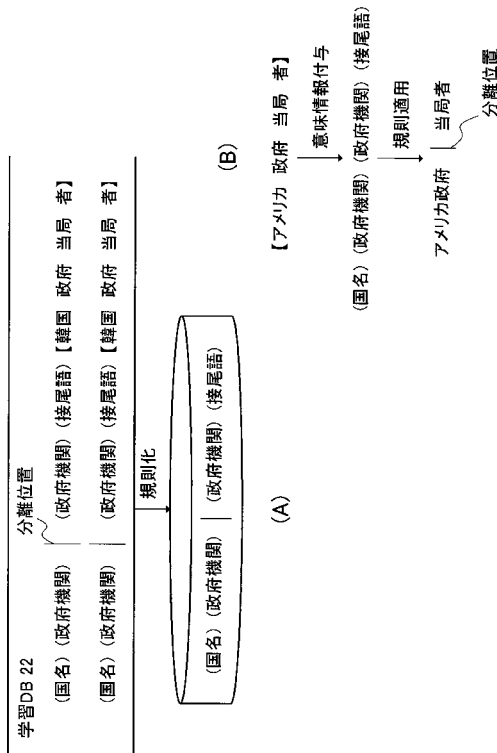
【図 3】



【図 4】



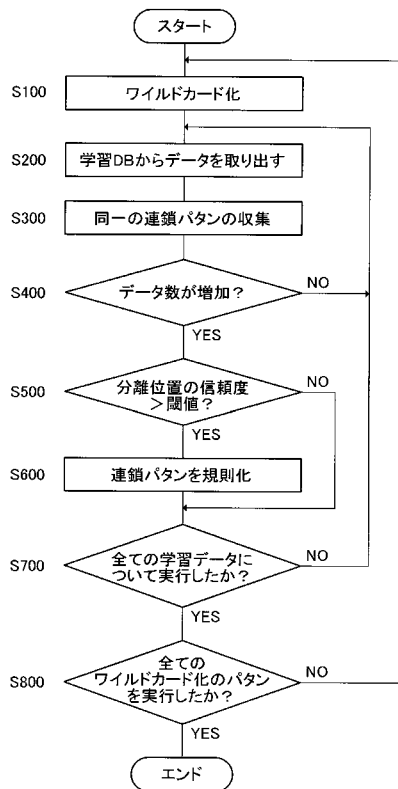
【図 5】



【図 6】

N=3 (7通り)	N=4 (15通り)	N=5 (31通り)
ABC	ABCD	ABCDE
AB*	ABC*	ABCD*
A*C	AB*D	ABC*E
*BC	A*CD	AB*DE
A**	*BCD	A*CDE
B	AB**	*BCDE
C	AD	ABC**
	**CD	*BCD*
	A*C*	**CDE
	*B*D	AB*D*
	BC	AB**E
	A***	A*CD*
	*B**	A*C*E
	C*	ADE
	***D	*BC*E
		*B*DE
		A***E
		A**D*
		*B**E
		A*C**
		*B*D*
		**C*E
		AB***
		*BC**
		**CD*
		***DE
		A****
		*B***
		C
		***D*
		****E

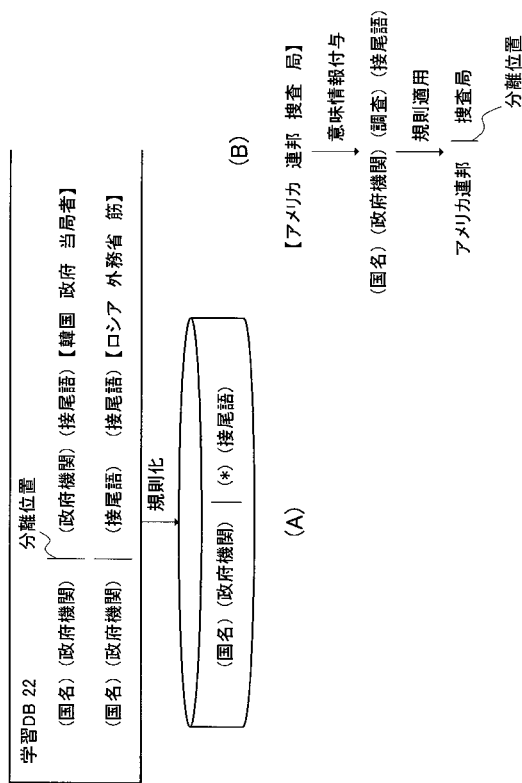
【図 7】



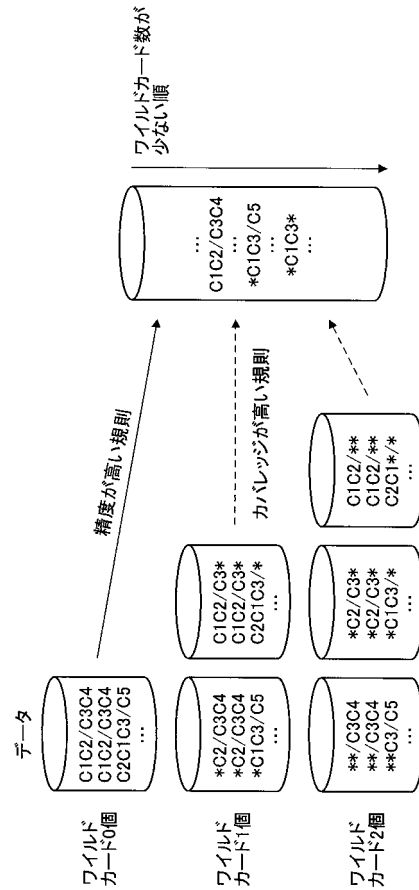
【図 8】

データ	データ#1: 日本 政府 / 当局 者 (国名)(政府機関)/(政府機関)(接尾語)	データ#2: 韓国 政府 / 当局 者 (国名)(政府機関)/(政府機関)(接尾語)	データ#3: 大阪 府庁 / 当局 者 (地域)(政府機関)/(政府機関)(接尾語)	データ#4: 東京 都庁 / 関係 者 (地域)(政府機関)/(関係)(接尾語)
ワイルドカード 0個	A	B	C	D
#1	(国名)	(政府機関)	(政府機関)	(接尾語)
#2	(国名)	(政府機関)	(政府機関)	(接尾語)
	(国名)(政府機関)/(政府機関)(接尾語)			
ワイルドカード 1個	*	B	C	D
#1	(国名)	(政府機関)	(政府機関)	(接尾語)
#2	(国名)	(政府機関)	(政府機関)	(接尾語)
#3	(地域)	(政府機関)	(政府機関)	(接尾語) ← 増加分
	*(政府機関)/(政府機関)(接尾語)			
ワイルドカード 2個	*	B	C	D
#1	(国名)	(政府機関)	(政府機関)	(接尾語)
#2	(国名)	(政府機関)	(政府機関)	(接尾語)
#3	(地域)	(政府機関)	(政府機関)	(接尾語)
#4	(地域)	(政府機関)	(関係)	(接尾語) ← 増加分
	*(政府機関)/ *(接尾語)			

【図 9】



【図 10】



【図 11】

構成単語		分離位置情報 (補入語句)
急進(進退)	民営(運営)	3(0)
国家(国)	評議(会議)	3(0)
外交(交わり)	安保(原理)	2(0)
統合(集合)	参謀(軍人)	3(0)
...		...
急進(進退) 民営(運営) 化(接尾語) 政策(計画)		3(0)
国家(国) 評議(会議) 会(接尾語) 議長(役職)		3(0)
外交(交わり) 安保(原理) 関係(関係) 者(接尾語)		2(0)
統合(集合) 参謀(軍人) 会議(会議) 文書(文章)		3(0)

フロントページの続き

審査官 山下 剛史

(56)参考文献 特開平 0 9 - 1 5 2 8 8 3 (J P , A)

特開平 0 4 - 0 3 6 7 9 9 (J P , A)

佐々木啓友他, " 複合名詞に対する複合語アクセント規則 ", 電子情報通信学会技術研究報告, SP2001-115 (2002-01) , Vol.101 , No.603 , pp.9-15

(58)調査した分野(Int.Cl. , D B 名)

G10L 13/00-13/08