(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号
特許第3575242号
(P3575242)
(45)発行日　平成16年10月13日(2004.10.13)　(24)登録日　平成16年7月16日(2004.7.16)

(51)Int.Cl.[7]　F I
G06F 17/30　G06F 17/30 210A
G06F 17/30 170A

請求項の数 2　(全 15 頁)

(21)出願番号　特願平9-245578
(22)出願日　平成9年9月10日(1997.9.10)
(65)公開番号　特開平11-85766
(43)公開日　平成11年3月30日(1999.3.30)
審査請求日　平成13年6月4日(2001.6.4)

(73)特許権者　000004226
日本電信電話株式会社
東京都千代田区大手町二丁目３番１号
(74)代理人　100070150
弁理士　伊東　忠彦
(72)発明者　別所　克人
東京都新宿区西新宿三丁目１９番２号　日本電信電話株式会社内
(72)発明者　岩瀬　成人
東京都新宿区西新宿三丁目１９番２号　日本電信電話株式会社内

審査官　野崎　大進

最終頁に続く

(54)【発明の名称】キーワード抽出装置

(57)【特許請求の範囲】
【請求項１】
自然文からキーワードを抽出するキーワード抽出装置であって、
自然文を入力文として入力する入力手段と、
単語の文法情報及び意味情報を格納する単語情報記憶手段と、
前記単語情報記憶手段の情報を参照して、前記入力文を形態素解析する形態素解析手段と、
形態素解析された文の構文構造を求める構文解析手段と、
前記構文解析手段により構文解析された文の中の動詞に対する名詞句の役割及び句間の修飾関係を表現した意味構造を求める意味解析手段と、
ユーザの要求を表す句である条件句であり、かつ、動詞であるもの(以下、条件動詞)に対する名詞句の役割(以下、深層格)の種類及び該深層格に格納されるキーワードの種類の組み合わせからなるキーワード抽出規則を格納するキーワード抽出規則記憶手段と、
前記入力文から、条件句を動詞に対する名詞句の役割及び句間の修飾関係に基づいて特定する条件句抽出手段と、
前記条件句抽出手段で特定された前記条件句が条件動詞である場合に、前記キーワード抽出規則を用いて、前記入力文中の動詞及び名詞句に対してキーワードの種類を特定するキーワード抽出手段と、を有することを特徴とするキーワード抽出装置。
【請求項２】
自然文からキーワードを抽出するキーワード抽出装置であって、

自然文を入力文として入力する入力手段と、
単語の文法情報及び意味情報を格納する単語情報記憶手段と、
前記単語情報記憶手段の情報を参照して、前記入力文を形態素解析する形態素解析手段と、
形態素解析された文の構文構造を求める構文解析手段と、
前記構文解析手段により構文解析された文の中の動詞に対する名詞句の役割及び句間の修飾関係を表現した意味構造を求める意味解析手段と、
ユーザの要求を表す句である条件句であり、かつ、名詞句であるもの（以下、条件名詞句）を修飾するキーワードの種類の組み合わせからなるキーワード抽出規則を格納するキーワード抽出規則記憶手段と、
前記入力文から、条件句を動詞に対する名詞句の役割及び句間の修飾関係に基づいて特定する条件句抽出手段と、
前記条件句抽出手段で特定された前記条件句が条件名詞句である場合に、前記キーワード抽出規則を用いて、前記入力文中の動詞及び名詞句に対してキーワードの種類を特定するキーワード抽出手段と、を有することを特徴とするキーワード抽出装置。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】

本発明は、キーワード抽出装置に係り、特に、情報案内サービスにおいて、利用者からの問い合わせが自然文である場合に、検索キーとして使用するキーワードを入力文から抽出することを可能とする情報検索システムにおけるキーワード抽出装置に関する。

【０００２】

【従来の技術】

自然文からのキーワード抽出においては、形態素解析によって得られた単語集合から、指定した品詞、意味を持つ単語を抽出している。
また、別の手法としては、特開平９－１１４８４０に開示されているように、文献等のテキスト文に出現する頻度の多い単語をキーワードとして抽出する統計的手法を用いている。

【０００３】

【発明が解決しようとする課題】

しかしながら、形態素解析によって得られた単語集合から、指定した品詞、意味を持つ単語を抽出する方法では、その単語の文中での役割を考慮していないため、入力文においては重要でない単語もキーワードとして抽出してしまうという問題がある。

【０００４】

また、このような重要でない単語を抽出しないようにするため、指定する品詞、意味の条件を制限すると、逆に抽出すべき単語が抽出されないという問題がある。
例えば、「Ｉ ｗａｎｔ ｔｏ ｂｕｙ ａ ｃａｒ ｔｏ ｇｏ ｔｏ ｔｈｅ ｏｆｆｉｃｅ 」という問い合わせ文に対し、「ｃａｒ ｄｅａｌｅｒ」という分野が案内対象として対応するが、抽出キーワードの条件を動詞と名詞にすると、分野に対応するキーワードとして重要である「ｂｕｙ 」や「ｃａｒ 」のみならず、重要でない「ｇｏ」や「ｏｆｆｉｃｅ」まで抽出される。条件を名詞に制限すると、逆にキーワードとして重要な「ｂｕｙ 」が抽出されない。

【０００５】

また、出現頻度を用いる統計的手法では、検索システムへの問い合わせ文のように多くても２，３文程度の短文では、単語の母数が少ないため、適切にキーワードを抽出できないという問題がある。
本発明は、上記の点に鑑みなされたもので、キーワード抽出のノイズを無くすことができ、検索システムの性能の向上を図ることが可能なキーワード抽出装置を提供することを目的とする。

【０００６】

【課題を解決するための手段】
【０００８】
図１は、本発明の原理構成図(その１)である。
【０００９】
本発明は、自然文からキーワードを抽出するキーワード抽出装置であって、
自然文を入力文として入力する入力手段７０と、
単語の文法情報及び意味情報を格納する単語情報記憶手段１０と、
単語情報記憶手段１０の内容を参照して、入力文を形態素解析する形態素解析手段２０と、
形態素解析された文の構文構造を求める構文解析手段３０と、
構文解析手段３０により構文解析された文の中の動詞に対する名詞句の役割及び句間の修飾関係を表現した意味構造を求める意味解析手段４０と、
ユーザの要求を表す句である条件句であり、かつ、動詞であるもの（以下、条件動詞)に対する名詞句の役割(以下、深層格)の種類及び該深層格に格納されるキーワードの種類の組み合わせからなるキーワード抽出規則を格納するキーワード抽出規則記憶手段５０と、
入力文から、条件句を動詞に対する名詞句の役割及び句間の修飾関係に基づいて特定する条件句抽出手段６１と、
条件句抽出手段６１で特定された条件句が条件動詞である場合に、キーワード抽出規則を用いて、入力文中の動詞及び名詞句に対してキーワードの種類を特定するキーワード抽出手段６２と、を有する。
【００１２】
図２は、本発明の原理構成図(その２)である。
本発明は、自然文からキーワードを抽出するキーワード抽出装置であって、
自然文を入力文として入力する入力手段７０と、
単語の文法情報及び意味情報を格納する単語情報記憶手段１０と、
単語情報記憶手段１０の内容を参照して、入力文を形態素解析する形態素解析手段２０と、
形態素解析された文の構文構造を求める構文解析手段３０と、
構文解析手段３０により構文解析された文の中の動詞に対する名詞句の役割及び句間の修飾関係を表現した意味構造を求める意味解析手段４０と、
ユーザの要求を表す句である条件句であり、かつ、名詞句であるもの（以下、条件名詞句）を修飾するキーワードの種類の組み合わせからなるキーワード抽出規則を格納するキーワード抽出規則記憶手段５０と、
入力文から、条件句を動詞に対する名詞句の役割及び句間の修飾関係に基づいて特定する条件句抽出手段６１と、
条件句抽出手段６１で特定された条件句が条件名詞句である場合に、キーワード抽出規則を用いて、入力文中の動詞及び名詞句に対してキーワードの種類を特定するキーワード抽出手段６２と、を有する。
【００１６】
上記のように、本発明は、文中において、どの係り受け関係にある単語が文全体において重要である単語かを考慮した上で、キーワードを抽出するので、必要なキーワードのみを洩れなく抽出することが可能である。また、対象としている分野毎に、抽出する規則を変えることにより、分野毎に適切なキーワードの抽出を行うことができる柔軟性を持つ。
【００１７】
【発明の実施の形態】
図３は、本発明のキーワード抽出装置の構成を示す。同図に示すキーワード抽出装置１００は、単語の文法情報、意味情報を格納する単語情報記憶部１０、単語情報記憶部１０に格納されている情報を用いて入力文を単語分割する形態素解析部２０、形態素解析結果に基づいて入力文の構文構造を求める構文解析部３０、構文構造に基づいて入力文の意味構造を求める意味解析部４０、キーワードを抽出するキーワード抽出規則を格納するキーワ

ード抽出規則記憶部５０及び、キーワード抽出規則を用いて意味構造からキーワードを抽出する意図理解部６０から構成される。

【００１８】

次に、本発明のキーワード抽出の動作を説明する。

図４は、本発明のキーワード抽出の動作を示すフローチャートである。

ステップ１０１）　自然文が入力されると、形態素解析部２０は、単語情報記憶部１０の単語情報を参照して入力文を単語分割する。

ステップ１０２）　形態素解析結果に基づいて、構文解析部３０は、入力文の構文構造を求める。

【００１９】

ステップ１０３）　意味解析部４０は、構文構造に基づいて入力文の意味構造を求める。

ステップ１０４）　意図解析部６０は、キーワード抽出規則記憶部５０に格納されているキーワード抽出規則を用いて、意味構造からキーワードを抽出する。

【００２０】

【実施例】

以下、本発明の一実施例を図面と共に説明する。

［第１の実施例］

以下、第１の実施例を前述の図４に示すフローに従って説明する。

まず、入力文を

“Ｉ　ｗａｎｔ　ｔｏ　ｂｕｙ　ａ　ｃａｒ　ｔｏ　ｇｏ　ｔｏ　ｔｈｅ　ｏｆｆｉｃｅ”

とする。

【００２１】

上記の入力文が形態素解析部２０に入力されると、単語情報記憶部１０の内容を参照して、入力文を単語分割する。上記の入力文の例は、英文なので、各単語は予めスペースで分割されている。従って、当該形態素解析部２０では、各単語の品詞を同定する処理を行う。これにより、以下のような形態素解析結果が得られる。

I／N want／V to／PREP buy／V a／DET car／N
to／PREP go／V to／PREP the／DET office／N

構文解析部３０は、上記の形態素解析結果を入力として、句構造規則を用いて、入力文の構文構造を出力する。図５にその例を示す。

【００２２】

次に、意味解析部４０は、図５に示す構文構造を入力として、各文法規則を用いて、格フレームで表現された意味構造を出力する。動詞の深層格と、句間の修飾関係を解析する。図６（ａ）にその例を示す。

次に、意図理解部６０は、図６（ａ）に示す意味構造を入力として、キーワード抽出規則記憶部５０の内容を参照して、キーワードを抽出する。

【００２３】

図６（ｂ）は、キーワード抽出規則の例である。当該規則は、図６（ｂ）において表されている意味構造に対しては、所定の位置（太枠）の動詞はユーザの要求を表すものだとして動詞キーワードに、その対象格である物、事、人、職業の意味をもつ名詞句は、商品キーワードにすることを示している。

意図理解部６０は、キーワード抽出規則記憶部５０に格納されているキーワード抽出規則のうち、図６（ｂ）の意味構造が入力文の意味構造（図６（ａ））と一致することを判断し、当該抽出規則に基づいて「ｂｕｙ」を動詞キーワード、「ｃａｒ」を商品キーワードとして抽出する。その結果を図６（ｃ）に示す。

【００２４】

［第２の実施例］

次に、第２の実施例として、対象としている分野に応じて入力文から抽出するキーワード

の種類を特定する方法について説明する。
図７は、本発明の第２の実施例の対象とする分野に応じて入力文から抽出するキーワードの種類を特定する動作のフローチャートである。
【００２５】
以下のフローチャートは、入力文として、
『８月に電車で神社めぐりに行きたい』
が入力された場合を例に説明する。
ステップ２０１～ステップ２０３）　形態素解析部２０による形態素解析処理、構文解析部３０による構文解析処理、意味解析部４０による意味解析処理を経て入力文の意味構造が出力される。この例を図８（ａ）に示す。
【００２６】
ステップ２０４）　意図理解部６０における意図理解過程１では、分野によらないキーワードである商品、動詞、職業のキーワードを抽出する。即ち、意味構造を入力として、キーワード抽出規則記憶部５０から分野によらないキーワード抽出規則１（図８（ｂ））を検索し、意味構造（図８（ａ））とマッチングすることにより、キーワード抽出結果（図８（ｃ））を出力する。
【００２７】
ステップ２０５）　意図理解部６０における分野推論過程は、キーワード抽出結果（図８（ｃ））を入力として、検索システムが案内する対象を分類する分野名を推論する。図８（ｄ）はその推論結果である。
ステップ２０６）　意図理解部６０における意図理解過程２は、分野に依存するキーワードを抽出する。即ち、分野推論結果（図８（ｄ））を入力として、キーワード抽出規則記憶部５０から、分野推論結果（図８（ｄ））の分野『観光案内』に対応するキーワード抽出規則を求める。当該処理は、図９（ａ）に示す意味構造と、図９（ｂ）に示すキーワード抽出規則２をマッチングさせることにより、図９（ｃ）に示すキーワード抽出結果を出力する。
【００２８】
［第３の実施例］
次に、第３の実施例として、条件句が動詞（条件動詞）である場合に、条件動詞の深層格の種類及びその深層格に格納されるキーワードの種類の組み合わせからなるキーワード抽出規則を用いて、入力文の動詞や名詞句に対してキーワードの種類を特定する例について説明する。
【００２９】
図１０は、本発明の第３の実施例の意図理解部の構成を示す。同図に示す意図理解部６０は、条件句抽出部６１とキーワード抽出部６２から構成される。
ここで、条件句とは、質問文の中心をなす語句で、その他のキーワードを抽出する処理の基点となるものである。
以下に条件句抽出部６１の動作を説明する。
【００３０】
図１１は、本発明の第３の実施例の条件句抽出部の動作のフローチャートである。
以下の条件句抽出処理では、日本語を入力した場合の条件句を抽出する場合の処理を示す。ここで、「知りたい」「教えてほしい」など、検索システムに店の情報を要求するのに用いられる動詞を要求動詞と定義する。
【００３１】
ステップ３０１）　最初に入力文の最終文節の動詞が要求動詞か否かを判定し、要求動詞でない場合にはステップ３０２に移行し、要求動詞である場合にはステップ３０３に移行する。
ステップ３０２）　要求動詞でない場合（例えば、「買いたい」「食べたい」等）には、その最終文節の動詞を条件句とする。
【００３２】

ステップ３０３）　要求動詞である場合には、最終文節の動詞の対象格の意味が物、事、人かを判断し、物、事、人の場合（例えば、入力文「ピアノを教えてほしい」）には、ステップ３０４に移行し、そうでない場合にはステップ３０５に移行する。
ステップ３０４）　物、事、人の場合には、その最終文節の動詞を条件句とする。
【００３３】
ステップ３０５）　最終文節の動詞の対象格が職業または職業相当の代名詞（店、所、業者等）かを判断し、職業または、職業相当の代名詞の場合（例えば、入力文「レストランを教えてほしい」）には、ステップ３０６に移行し、そうでない場合にはステップ３０７に移行する。
ステップ３０６）　職業または、職業相当の代名詞である場合には、職業が重要な句なので、職業を条件句とする。ここで、「レストランの住所を教えてほしい」の「住所」や「電話番号」のように店の属性を表す項目を主題とする。
【００３４】
ステップ３０７）　最終文節の動詞の対象格が主題であり、主題を連体修飾する職業があるかを判定し、ある場合にはステップ３０８に移行し、そうでない場合にはステップ３０９に移行する。
ステップ３０８）　主題を連体修飾する職業がある場合には、この連体修飾している職業が重要な句なので、職業を条件句とする。
【００３５】
ステップ３０９）　これ以外の場合には、条件句は無しとする。
上記のフローチャートにおいて、「テレビ」のように、入力文の最終文節が動詞でなく、商品（物、事、人）の場合は、情報案内システムへの質問であることを仮定しているので、意味解析部４０において、「（商品）の店を知りたい」のような補完処理をして意味構造を出力する。これにより、意図理解部６０の条件句抽出部６１では、「店」は、職業相当の代名詞なので、ステップ３０６において、「店」が条件句となる。
【００３６】
「レストラン」のように、入力文の最終文節が動詞でなく職業の場合や、「レストランの住所」のように主題である場合は、やはり情報案内システムへの質問であることを仮定し、意味解析部４０において、入力文の後に「を知りたい」を補完して意味構造を出力する。これにより、意図理解部６０の条件句抽出部６１では、ステップ３０５または、ステップ３０７がマッチし、職業が条件句となる。
【００３７】
次に、キーワード抽出部６２の動作について説明する。
図１２は、本発明の第３の実施例のキーワード抽出規則の例であり、当該規則は、キーワード抽出規則記憶部５０に格納されている。条件動詞の深層格の種類及びその深層格に格納されるキーワードの種類の組み合わせの規則の例である。
ここで、入力文
「レストランでパーティーをしたい」
が入力された時の意味構造を図１３に示す。
【００３８】
意図理解部６０は、条件句抽出部６１により、動詞「する」を条件句と特定する。キーワード抽出部６２では、条件句が動詞である抽出規則について順次入力の意味構造と照合していく。図１２の規則１は、条件動詞の対象格が物、事、人、職業の意味をもつ名詞句ならば、その名詞句を商品キーワードに、条件動詞を動詞キーワードにすることを表している。意図理解部６０は、対象格の意味構造が規則１の構造と一致するので、「パーティー」を商品キーワードに、「する」を動詞キーワードにする。
【００３９】
また、図１２の規則２では、条件動詞の場所格が職業の意味を持つ名詞句ならば、その名詞句を職業キーワードにすることを表している。意図理解部６０は、場所格の意味構造が規則２の構造と一致するので、「レストラン」を職業キーワードにする。

入力文『インターネットを利用できる機器を買いたい』の意味構造を、図１４に示す。意図理解部６０は、条件句抽出部６１により、動詞「買う」を条件句と特定する。図１２の規則３は、条件動詞の対象格が物語尾、事語尾の意味をもつ名詞句であり、その名詞句に修飾する動詞があり、さらにその動詞の対象格が物、事、人、職業の意味をもつ名詞句の場合のキーワードの種類の特定の規則を表している。

【００４０】

意図理解部６０は、対象格の意味構造が図１２の規則３の構造と一致するので、「インターネット（事）を利用する（動詞）機器（物語尾）」を商品キーワードに、「買う」を動詞キーワードにする。この場合、商品キーワードは「インターネット」＋「利用する」という構造を持っており、分野推論過程（ステップ２０５）では、「インターネット」＋「利用する」はどういうものかをさらに推論することになる。

【００４１】

次に、英語を入力したときの条件句抽出部６１の処理を説明する。図１５は、本発明の第３の実施例の英語を入力したときの条件句抽出処理のフローチャートである。図１６は、キーワード抽出規則の例であり、図１７は、意味構造の例を示す。英語入力文に対しては、意味解析において動詞の深層格、句間の修飾関係、句間の主辞関係（主辞とは句に含まれる、文法的に中心的な役割を果たす句のことである）を表す意味構造が出力される。例えば、図１７は、入力文“Ｉ ｗａｎｔ ｔｏ ｂｕｙ ａ ｃａｒ ｔｏ ｇｏ ｔｈｅ ｏｆｆｉｃｅ ”の意味構造であるが、入力文全体の主辞は動詞「ｗａｎｔ」である。

【００４２】

以下、図１７の入力文を用いて処理フローを説明する。

まず、入力文の主辞の動詞「ｗａｎｔ」を、条件句抽出処理の対象とする（ステップ４０１）。

対象とする動詞が要求動詞であるかを判定する。ここでは、対象とする動詞「ｗａｎｔ」は、検索システムへの要求を表し得るので、要求動詞である（ステップ４０２）。対象とする動詞が要求動詞でない場合には、対象とする動詞を条件句とする（ステップ４０３）。

【００４３】

対象とする動詞の対象格が動詞であるかを判定し、動詞である場合にはステップ４０５に移行する。この場合の「ｗａｎｔ」の対象格は動詞「ｂｕｙ」であるので（ステップ４０４）、動詞「ｂｕｙ」を新たな対象とする（ステップ４０５）。対象とする動詞「ｂｕｙ」は、要求動詞ではないので、「ｂｕｙ」を条件句とする（ステップ４０３）。ステップ４０６以降の処理は、図１１のステップ３０３以降の処理と同様である。

【００４４】

「ｔｅｌｅｖｉｓｉｏｎ」のように、入力文の主辞が動詞でなく商品の場合は、日本語と同様に、意味解析部４０において、「ｋｎｏｗ ｔｈｅ ｓｈｏｐ ｏｆ （商品）」のような補完処理をして意味構造を出力する。これにより、意図理解部６０の条件句抽出部６１では、「ｓｈｏｐ」は、職業相当の名詞なので、ステップ４０８においてマッチし、「ｓｈｏｐ」が条件句となる。

【００４５】

「ｒｅｓｔａｕｒａｎｔ」のように、入力文の最終文節が動詞でなく、職業の場合や、「ｔｈｅ ａｄｄｒｅｓｓ ｏｆ ａ ｒｅｓｔａｕｒａｎｔ 」のように、主題である場合も、日本語と同様に、意味解析部４０において、入力文の前に「ｋｎｏｗ」を補完して意味構造を出力する。これにより、意図理解部６０の条件句抽出部６１では、ステップ４０８とステップ４１０においてマッチし、職業が条件句となる。

【００４６】

次に、英語を入力した場合のキーワード抽出部６２の処理を説明する。

意味構造として、図１７に示すものを用いて説明する。意図理解部６０は、条件句抽出部６１により動詞「ｂｕｙ」を条件句と特定する。キーワード抽出部６２では、キーワード

抽出規則記憶部５０に格納されている、条件句が動詞である抽出規則について順次入力の意味構造と照合していく。意図理解部６０は、対象格の意味構造が図１６に示すキーワード抽出規則の構造と一致するので、「ｃａｒ」を商品キーワードに、「ｂｕｙ」を動詞キーワードにする。条件動詞の目的格である動詞をキーワードにする抽出規則は、キーワード抽出規則記憶部５０にはないので、「ｇｏ」はキーワードにならない。

【００４７】

［第４の実施例］

次に、第４の実施例として、条件句が職業を意味する名詞句である場合に、条件名詞句を修飾するキーワードの種類の組み合わせの規則を用いて、入力文中の動詞や名詞句に対してキーワードの種類を特定する例を説明する。

図１８は、本発明の第４の実施例のキーワード抽出規則の例を示す。

【００４８】

同図に示すキーワード抽出規則は、キーワード抽出規則記憶部５０に格納されている。条件名詞句に修飾するキーワードの種類の組み合わせの規則の例である。入力文『ステーキの食べられるレストラン』の意味構造を図１９（ａ）に示す。意図解析部６０の条件句抽出部６１により職業「レストラン」を条件句と特定する。キーワード抽出部６２では、条件句が職業であるキーワード抽出規則について順次入力の意味構造と照合していく。図１８の規則２では、条件名詞句に修飾する動詞があり、さらに、その動詞の対象格が物、事、人、職業の意味を持つ名詞句の場合のキーワードの種類の特定の規則を表している。意図理解部６０は、入力の意味構造が図１８の規則２の構造と一致するので、「ステーキ」を商品キーワードに、「食べる」を動詞キーワードに、「レストラン」を職業キーワードにする。

【００４９】

入力文

『ｒｅｓｔａｕｒａｎｔ　ｆａｍｏｕｓ　ｆｏｒ　ｓｔｅａｋ　』

の意味構造を図１９（ｂ）に示す。意図理解部６０は、条件句抽出部６１により、職業「ｒｅｓｔａｕｒａｎｔ」を条件句と特定する。図１８の規則３では、条件名詞句に修飾する形容詞があり、さらに、その形容詞に修飾する物、事、人、職業の意味をもつ名詞句がある場合のキーワードの種類の特定の規則を表している。意図理解部６０は、入力の意味構造が図１８の規則３の構造と一致するので、「ｓｔｅａｋ　」を商品キーワードに、「ｒｅｓｔａｕｒａｎｔ」を職業キーワードにする。

【００５０】

このように、本発明は、入力文に対して形態素解析、構文解析、意味解析の処理を行うことにより、動詞の深層格及び、句間等の修飾関係を表した意味構造を得る。その意味構造を入力として、単語間の係り受け関係を考慮したキーワード抽出規則に基づいてキーワードを抽出することが可能となる。

また、本発明をプログラム（ソフトウェア）で構築し、キーワード抽出を行うコンピュータに接続されるディスク装置やフロッピーディスク、ＣＤ－ＲＯＭ等の可搬記憶媒体に格納し、必要に応じてインストールすることにより汎用的なキーワード抽出が可能となる。

なお、本発明は、上記の実施例に限定されることなく、特許請求の範囲内で種々変更・応用が可能である。

【００５１】

【発明の効果】

上述のように、本発明によれば、文中において、どの係り受け関係にある単語が文全体において重要である単語かを表現した規則を用いてキーワードを抽出するので、必要なキーワードのみを洩れなく抽出することができる。このため、キーワード抽出のノイズを無くすことができ、検索システムの性能向上の効果をもたらす。

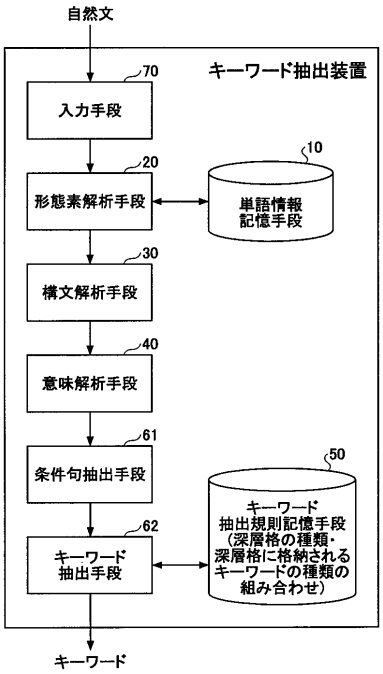
【００５２】

また、対象としている分野に応じて、抽出する規則を変えることにより、分野毎に適切なキーワードの抽出を行うことができ、検索システムの保守の上でも柔軟性をもたらす。

【図面の簡単な説明】
【図１】本発明の原理構成図(その１)である。
【図２】本発明の原理構成図(その２)である。
【図３】本発明のキーワード抽出装置の構成図である。
【図４】本発明のキーワード抽出の動作を示すフローチャートである。
【図５】本発明の第１の実施例の構文解析結果の例である。
【図６】本発明の第１の実施例のキーワード抽出処理を説明するための図である。
【図７】本発明の第２の実施例の対象とする分野に応じて入力文から抽出するキーワードの種類を特定する動作のフローチャートである。
【図８】本発明の第２の実施例のキーワード抽出処理を説明するための図(その１)である。
【図９】本発明の第２の実施例のキーワード抽出処理を説明するための図(その２)である。
【図１０】本発明の第３の実施例の意図理解部の構成図である。
【図１１】本発明の第３の実施例の条件句抽出部の動作のフローチャートである。
【図１２】本発明の第３の実施例のキーワード抽出規則の例(その１)である。
【図１３】本発明の第３の実施例の意味構造の例(その１)である。
【図１４】本発明の第３の実施例の意味構造の例(その２)である。
【図１５】本発明の第３の実施例の英語を入力としたときの条件句抽出処理のフローチャートである。
【図１６】本発明の第３の実施例のキーワード抽出規則の例(その２)である。
【図１７】本発明の第３の実施例の意味構造の例(その３)である。
【図１８】本発明の第４の実施例のキーワード抽出規則の例である。
【図１９】本発明の第４の実施例の意味構造の例である。
【符号の説明】
１０　単語情報記憶手段、単語情報記憶部
２０　形態素解析手段、形態素解析部
３０　構文解析手段、構文解析部
４０　意味解析手段、意味解析部
５０　キーワード抽出規則記憶手段、キーワード抽出規則記憶部
６０　意図理解部
６１　条件句抽出手段、条件句抽出部
６２　キーワード抽出手段、キーワード抽出部
７０　入力手段

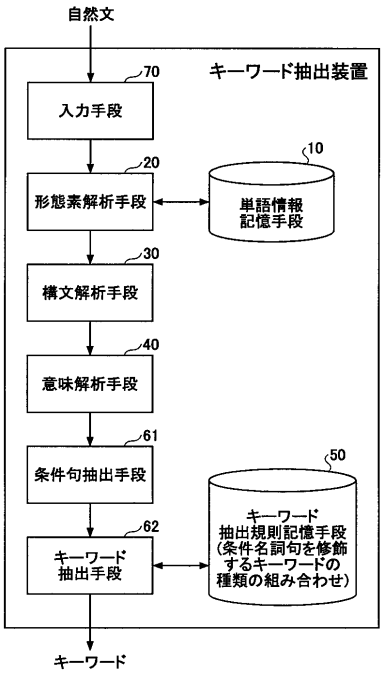
【図１】

本発明の原理構成図(その1)



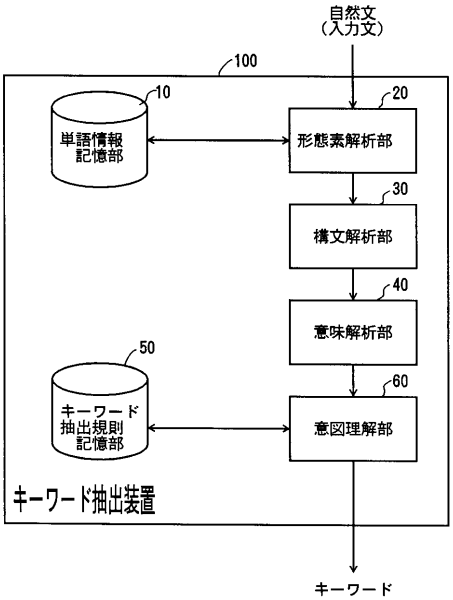
【図２】

本発明の原理構成図(その2)



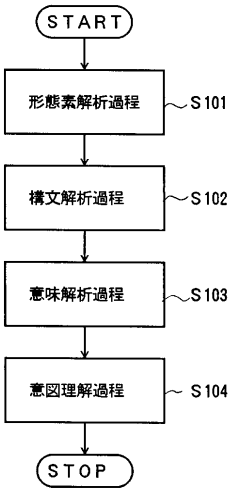
【図３】

本発明のキーワード抽出装置の構成図



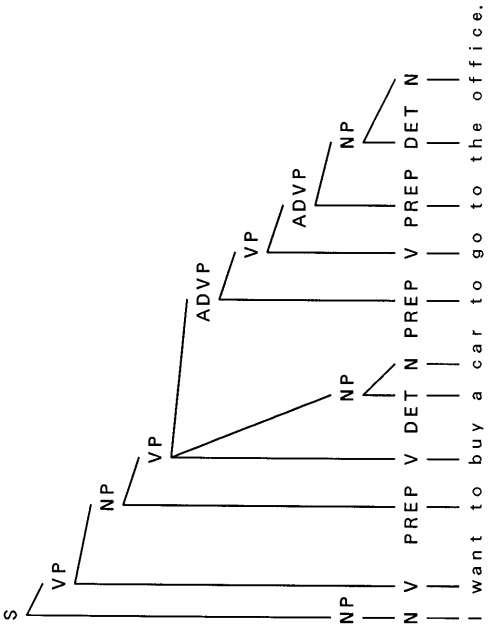
【図４】

本発明のキーワード抽出の動作を示すフローチャート



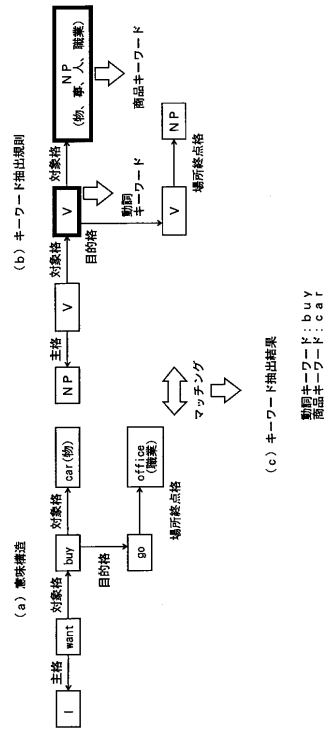
【図５】

本発明の第１の実施例の構造解析結果の例



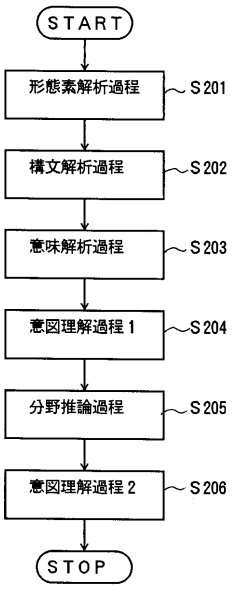
【図６】

本発明の第1の実施例のキーワード抽出処理を説明するための図



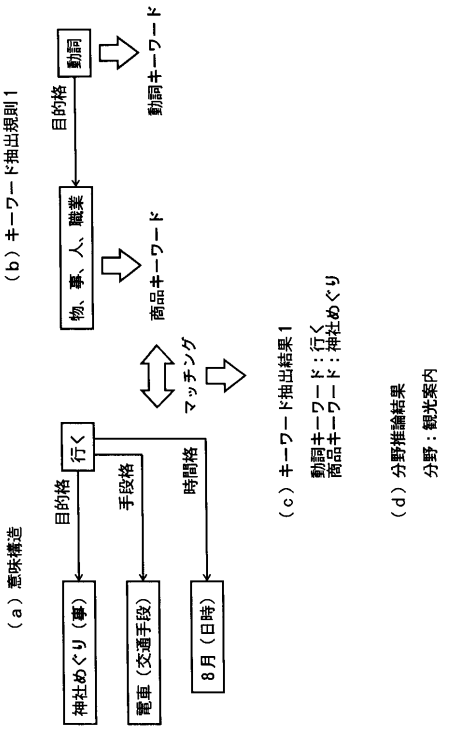
【図７】

本発明の第２の実施例の対象とする分野に応じて入力文から抽出するキーワードの種類を特定する動作のフローチャート



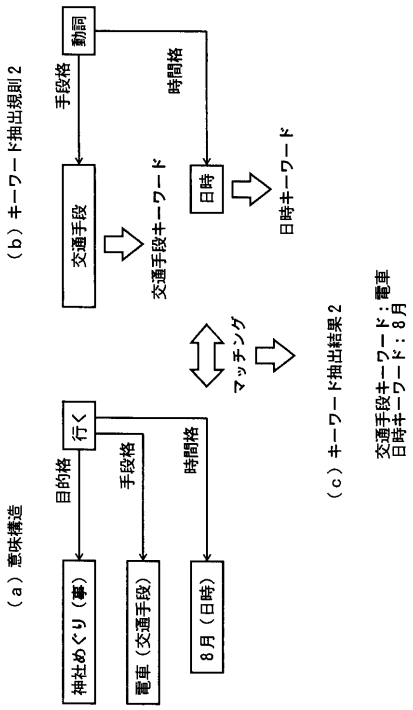
【図８】

本発明の第２の実施例のキーワード抽出処理を説明するための図（その１）



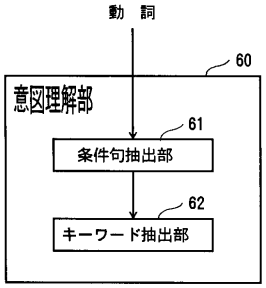
【図９】

本発明の第２の実施例のキーワード抽出処理を説明するための図（その２）



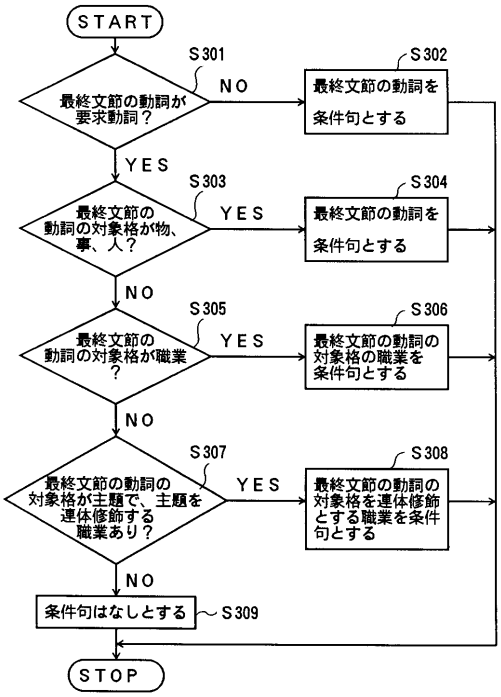
【図１０】

本発明の第３の実施例の意図理解部の構成図



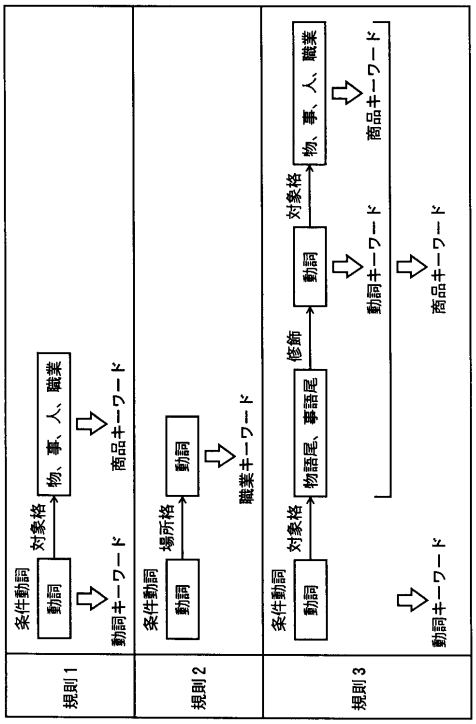
【図１１】

本発明の第３の実施例の条件句抽出部の動作のフローチャート



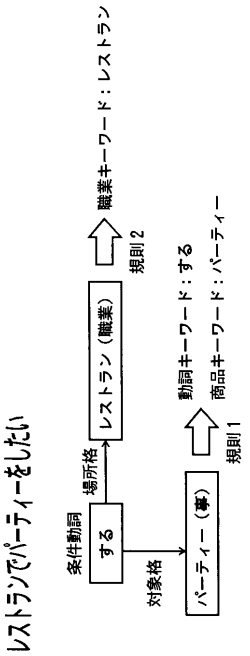
【図１２】

本発明の第３の実施例のキーワード抽出規則の例（その１）



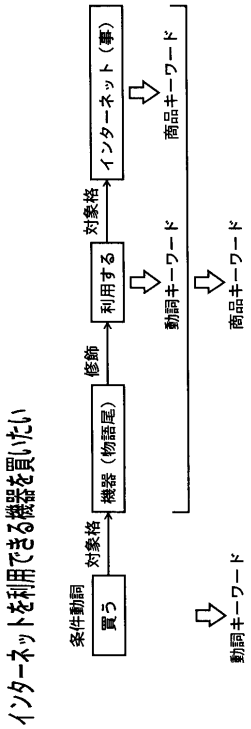
【図１３】

本発明の第３の実施例の意味構造の例（その１）



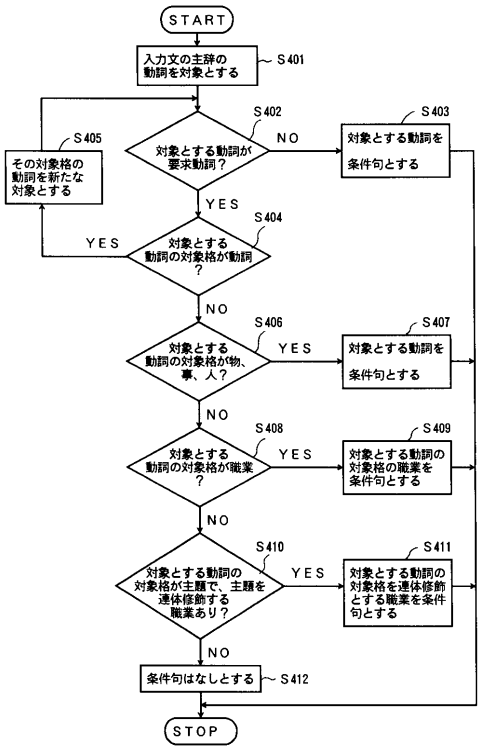
【図１４】

本発明の第３の実施例の意味構造の例（その２）



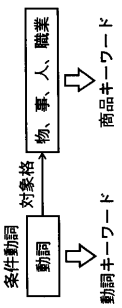
【図１５】

本発明の第３の実施例の英語を入力したときの
条件句抽出処理のフローチャート



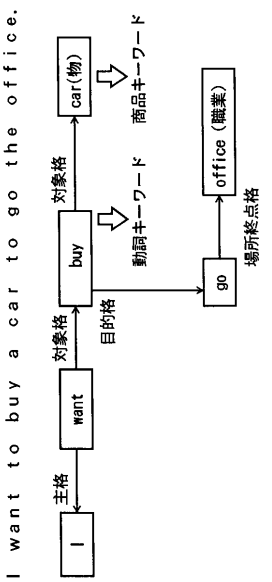
【図１６】

本発明の第３の実施例のキーワード抽出規則の例（その２）



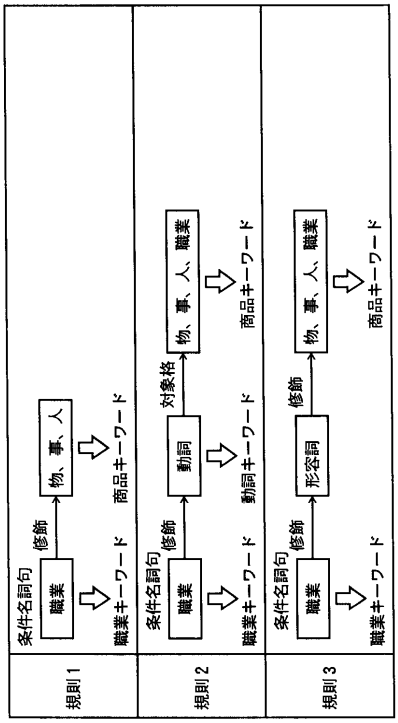
【図１７】

本発明の第３の実施例の意味構造の例（その３）



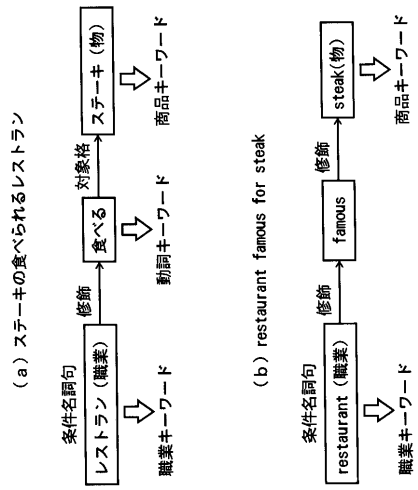
【図１８】

本発明の第４の実施例のキーワード抽出規則の例



【図１９】

本発明の第４の実施例の意味構造の例



フロントページの続き

(56)参考文献 特開平０４－１８２８７１（ＪＰ，Ａ）
特開昭６２－０８２４２９（ＪＰ，Ａ）
特開平０６－１９５３８３（ＪＰ，Ａ）
特開平０６－２３１１７８（ＪＰ，Ａ）
特開平０７－１８２３７０（ＪＰ，Ａ）
特開平０９－０６２６８４（ＪＰ，Ａ）
戸部 美春 他，インテリジェントページシステム，ＮＴＴ Ｒ＆Ｄ，日本，社団法人 電気通信協会，１９９６年 5月１０日，Vol.45,No.5，pp.455-461.
岩瀬 成人 他，通信目的の理解方式の一検討，電子情報通信学会技術研究報告，日本，社団法人 電子情報通信学会，１９８７年 5月２２日，Vol.87,No.37(SE87-27)，pp.67-71.
岩瀬 成人 他，自然言語処理技術を用いた職業別電話帳検索の高度化，電子情報通信学会論文誌，日本，社団法人 電子情報通信学会，１９９１年 9月２５日，Vol.J74-D-II,No.9，pp.1255-1263.
須之内 美幸 他，フルテキストの構造化に基づく検索システム，情報処理会研究報告，日本，社団法人 情報処理学会，１９９２年 9月１１日，Vol.92,No.71(92-DBS-90-7)，pp.55-64.
岸本 行生 他，テキストの構造化に基づく検索システム，情報処理学会論文誌，日本，社団法人 情報処理学会，１９９４年 5月１５日，Vol.35,No.5，pp.908-916.
別所 克人 他，イエローページにおける業種分類推論方式，情報処理学会第５３回全国大会講演論文集，日本，社団法人 情報処理学会，１９９６年 9月 4日，Vol.3,No.3T-10，pp.3-207～3-208.

(58)調査した分野(Int.Cl.^{7}，ＤＢ名)
G06F 17/30 210
G06F 17/30 170
JICSTファイル(JOIS)