

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第3981729号

(P3981729)

(45) 発行日 平成19年9月26日(2007.9.26)

(24) 登録日 平成19年7月13日(2007.7.13)

(51) Int. Cl.

F I

G06F 17/30 (2006.01)
G06F 17/21 (2006.01)

G06F 17/30 380Z
G06F 17/30 170A
G06F 17/30 210A
G06F 17/21 590E

請求項の数 13 (全 28 頁)

(21) 出願番号 特願2003-67045 (P2003-67045)
(22) 出願日 平成15年3月12日(2003.3.12)
(65) 公開番号 特開2004-280176 (P2004-280176A)
(43) 公開日 平成16年10月7日(2004.10.7)
審査請求日 平成15年3月12日(2003.3.12)

(73) 特許権者 301022471
独立行政法人情報通信研究機構
東京都小金井市貫井北町4-2-1
(74) 代理人 100103827
弁理士 平岡 憲一
(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
(72) 発明者 竹内 和広
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

審査官 深津 始

最終頁に続く

(54) 【発明の名称】 キーワード強調装置及びプログラム

(57) 【特許請求の範囲】

【請求項1】

抽出の単位を設定する抽出単位設定手段と、
強調表示する抽出表現を設定する抽出表現設定手段と、
入力された1個の文書データの抽出領域の場所を設定する抽出領域設定手段と、
情報を格納する格納手段と、
抽出手段とを備え、
前記抽出手段は、前記抽出領域より前記抽出の単位に相当するものを抽出し、そのうち、
前記抽出表現に該当するものを前記格納手段に格納し、前記入力された1個の文書データ
を調べて現在の前記抽出の単位に相当するものと同じものが、前記格納手段に格納され
ている場合は、それを強調表示することを特徴としたキーワード強調装置。

10

【請求項2】

前記入力された1個の文書データを調べて、現在の抽出の単位に相当するものと同じものが
前記格納手段に格納されている場合で、一つ前の抽出の単位が強調表示されるもので
ある場合は、前記一つ前の抽出の単位と現在の抽出の単位に相当するものを特別強調表示
して表示することを特徴とした請求項1記載のキーワード強調装置。

【請求項3】

前記抽出表現設定手段で強調表示する抽出表現の設定は、少なくとも1つが名詞である
ことを特徴とする請求項1又は2記載のキーワード強調装置。

【請求項4】

20

前記抽出領域の場所として、前記入力された1個の文書データのタイトル部分とすることを特徴とした請求項1～3のいずれかに記載のキーワード強調装置。

【請求項5】

前記抽出領域の場所として、前記入力された1個の文書データのユーザが指定した部分とすることを特徴とした請求項1～3のいずれかに記載のキーワード強調装置。

【請求項6】

前記ユーザが指定した部分として、複数部分を指定し、該指定した部分により、異なる強調表示をすることを特徴とした請求項5記載のキーワード強調装置。

【請求項7】

前記入力された1個の文書データの差分を検出するために比較する領域の単位である検出領域を設定する検出領域設定手段と、

文書差分検出装置の格納手段と、

抽出手段とを有し、

前記抽出手段で、入力された1個の文書データの現在の検出領域において、前記文書差分検出装置の格納手段に格納されていない前記抽出単位に相当するものを抽出し、該抽出したものを前記文書差分検出装置の格納手段に格納することを繰り返す文書差分検出装置を備え、

前記ユーザが指定した部分として、前記文書差分検出装置の前記抽出手段が抽出した前記文書差分検出装置の格納手段に格納されていない前記抽出単位に相当するものを指定することを特徴とした請求項5又は6記載のキーワード強調装置。

【請求項8】

前記入力された1個の文書データの差分を検出するために比較する領域の単位である検出領域を設定する検出領域設定手段と、

文書差分検出装置の格納手段と、

抽出手段とを有し、

前記抽出手段で、前記入力された1個の文書データの現在の前記検出領域以外の領域から全ての前記抽出単位に相当するものを抽出して前記文書差分検出装置の格納手段に格納し、現在の前記検出領域において、前記文書差分検出装置の格納手段に格納されていない前記抽出単位に相当するものを抽出することを、前記検出領域ごとに繰り返す文書差分検出装置を備え、

前記ユーザが指定した部分として、前記文書差分検出装置の前記抽出手段が抽出した前記文書差分検出装置の格納手段に格納されていない前記抽出単位に相当するものを指定することを特徴とした請求項5又は6記載のキーワード強調装置。

【請求項9】

抽出の単位を設定する抽出単位設定手段と、

入力された1個の文書データの検出領域の場所を設定する検出領域設定手段と、

情報を格納する格納手段と、

抽出手段と、

前記抽出手段で、入力された1個の文書データの現在の検出領域において、文書差分検出装置の格納手段に格納されていない前記抽出単位に相当するものを強調表示して現在の検出領域の文書を出力し、前記強調表示したものを前記文書差分検出装置の格納手段に格納することを、繰り返す文書差分検出装置と、

抽出領域の場所として前記文書差分検出装置が強調表示した部分を設定する抽出領域設定手段とを備え、

前記抽出手段は、前記抽出領域より前記抽出の単位に相当するものを抽出して前記格納手段に格納し、前記入力された1個の文書データを調べて現在の前記抽出の単位に相当するものと同じものが、前記格納手段に格納されている場合は、それを強調表示することを特徴としたキーワード強調装置。

【請求項10】

抽出の単位を設定する抽出単位設定手段と、

10

20

30

40

50

入力された1個の文書データの検出領域の場所を設定する検出領域設定手段と、
情報を格納する格納手段と、
抽出手段と、

前記抽出手段で、前記入力された1個の文書データの現在の前記検出領域以外の検出領域から全ての前記抽出単位に相当するものを抽出して文書差分検出装置の格納手段に格納し、現在の前記検出領域において、前記文書差分検出装置の格納手段に格納されていない前記抽出単位に相当するものを強調表示して現在の検出領域の文書を出力することを、前記検出領域ごとに繰り返す文書差分検出装置と、

抽出領域の場所として前記文書差分検出装置が強調表示した部分を設定する抽出領域設定手段とを備え、

前記抽出手段は、前記抽出領域より前記抽出の単位に相当するものを抽出して前記格納手段に格納し、前記入力された1個の文書データを調べて現在の前記抽出の単位に相当するものと同じものが、前記格納手段に格納されている場合は、それを強調表示することを特徴としたキーワード強調装置。

【請求項11】

抽出の単位を設定する抽出単位設定手段と、

強調表示する抽出表現を設定する抽出表現設定手段と、

入力された1個の文書データの抽出領域の場所を設定する抽出領域設定手段と、

前記抽出領域より前記抽出の単位に相当するものを抽出し、そのうち、前記抽出表現に該当するものを前記格納手段に格納し、前記入力された1個の文書データを調べて現在の前記抽出の単位に相当するものと同じものが、前記格納手段に格納されている場合は、それを強調表示する抽出手段として、

コンピュータを機能させるためのプログラム。

【請求項12】

抽出の単位を設定する抽出単位設定手段と、

入力された1個の文書データの検出領域の場所を設定する検出領域設定手段と、

抽出手段で、前記入力された1個の文書データの現在の検出領域において、文書差分検出装置の格納手段に格納されていない前記抽出単位に相当するものを強調表示して現在の検出領域の文書を出力し、前記強調表示したものを前記文書差分検出装置の格納手段に格納することを、繰り返す文書差分検出装置と、

抽出領域の場所として、前記文書差分検出手段が強調表示した部分を設定する抽出領域設定手段と、

前記抽出領域より前記抽出の単位に相当するものを抽出して格納手段に格納し、前記入力された1個の文書データを調べて現在の前記抽出の単位に相当するものと同じものが、前記格納手段に格納されている場合は、それを強調表示する抽出手段として、

コンピュータを機能させるためのプログラム。

【請求項13】

抽出の単位を設定する抽出単位設定手段と、

入力された1個の文書データの検出領域の場所を設定する検出領域設定手段と、

抽出手段で、前記入力された1個の文書データの現在の前記検出領域以外の検出領域から全ての前記抽出単位に相当するものを抽出して文書差分検出装置の格納手段に格納し、現在の前記検出領域において、前記文書差分検出装置の格納手段に格納されていない前記抽出単位に相当するものを強調表示して現在の検出領域の文書を出力することを、前記検出領域ごとに繰り返す文書差分検出装置と、

抽出領域の場所として前記文書差分検出装置が強調表示した部分を設定する抽出領域設定手段と、

前記抽出領域より前記抽出の単位に相当するものを抽出して格納手段に格納し、前記入力された1個の文書データを調べて現在の前記抽出の単位に相当するものと同じものが、前記格納手段に格納されている場合は、それを強調表示する抽出手段として、

コンピュータを機能させるためのプログラム。

10

20

30

40

50

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、タイトル又はユーザが指定した領域に含まれる語を重要なキーワードとし、本文においてそれらキーワードの部分を強調表示するキーワード強調装置及びプログラムに関する。

【0002】

【従来技術】

従来、タイトルは、文書の中で重要なものとされている。例えば、タイトルの部分が重要とし、そこに現れるキーワードの得点を高くすることで情報検索精度の向上を計っている

10

【0003】

【非特許文献1】

村田真樹，馬青，内元清貴，小作浩美，内山将夫，井佐原均，
位置情報と分野情報を用いた情報検索，
自然言語処理（言語処理学会誌）
2000年4月，7巻，2号，p.141～p.160

【0004】

【発明が解決しようとする課題】

上記従来タイトルに現れるキーワードの得点を高くするのでは、本文中のどの部分が重

20

要なのか判断するのが難しかった。

【0005】

本発明は上記問題点の解決を図り、タイトル又はユーザが指定した領域に含まれる語を重要なキーワードとし、本文においてそれらキーワードの部分を強調表示することで、本文の中で重要部分が容易にわかるようにすることを目的とする。

【0006】

【課題を解決するための手段】

図1は本発明の原理説明図である。図1中、2は抽出手段、3aは格納手段、21は抽出単位設定手段、22は抽出表現設定手段、23は抽出領域設定手段である。

【0007】

本発明は、前記従来課題を解決するため次のような手段を有する。

30

【0008】

(1)：抽出の単位を設定する抽出単位設定手段21と、強調表示する抽出表現を設定する抽出表現設定手段22と、入力された1個の文書データの抽出領域の場所を設定する抽出領域設定手段23と、情報を格納する格納手段3aと、抽出手段2とを備え、前記抽出手段2は、前記抽出領域より前記抽出の単位に相当するものを抽出し、そのうち、前記抽出表現に該当するものを前記格納手段3aに格納し、前記入力された1個の文書データを調べて現在の前記抽出の単位に相当するものと同じものが、前記格納手段3aに格納されている場合は、それを強調表示する。このため、入力された1個の文書データ中で抽出領域の説明に該当する場所を容易に見つけることができる。

40

【0009】

(2)：前記(1)のキーワード強調装置において、前記入力された1個の文書データを調べて、現在の抽出の単位に相当するものと同じものが前記格納手段に格納されている場合で、一つ前の抽出の単位が強調表示されるものである場合は、前記一つ前の抽出の単位と現在の抽出の単位に相当するものを特別強調表示して表示する。このため、入力された1個の文書データ中で抽出領域の説明に該当する場所をより明確に見つけることができる。

【0010】

(3)：前記(1)又は(2)のキーワード強調装置において、前記抽出表現設定手段で強調表示する抽出表現の設定の少なくとも1つが名詞である。このため、名詞等の重要

50

なもののみ強調表示することができる。

【0011】

(4)：前記(1)～(3)のキーワード強調装置において、前記抽出領域の場所として、前記入力された1個の文書データのタイトル部分とする。このため、入力された1個の文書データ中で重要なものとされているタイトル部分の説明に該当する場所を容易に見つけることができる。

【0012】

(5)：前記(1)～(3)のキーワード強調装置において、前記抽出領域の場所として、前記入力された1個の文書データのユーザが指定した部分とする。このため、入力された1個の文書データ中でユーザが指定した部分の説明に該当する場所を容易に見つけることができる。

10

【0013】

(6)：前記(5)のキーワード強調装置において、前記ユーザが指定した部分として、複数部分を指定し、該指定した部分により、異なる強調表示をする。このため、入力されたデータ中でユーザが指定した複数部分の説明に該当する場所を容易に見つけることができる。

【0014】

(7)：前記(5)又は(6)のキーワード強調装置において、前記入力された1個の文書データの差分を検出するために比較する領域の単位である抽出領域を設定する抽出領域設定手段と、文書差分検出装置の格納手段と、抽出手段とを有し、前記抽出手段で、入力された1個の文書データの現在の抽出領域において、前記文書差分検出装置の格納手段に格納されていない前記抽出単位に相当するものを抽出し、該抽出したものを前記文書差分検出装置の格納手段に格納することを繰り返す文書差分検出装置を備え、前記ユーザが指定した部分として、前記文書差分検出装置の前記抽出手段が抽出した前記文書差分検出装置の格納手段に格納されていない前記抽出単位に相当するものを指定する。このため、入力された1個の文書データ中でユーザが指定した部分の説明に該当する場所をより明確に見つけることができる。

20

【0015】

(8)：前記(5)又は(6)のキーワード強調装置において、前記入力された1個の文書データの差分を検出するために比較する領域の単位である抽出領域を設定する抽出領域設定手段と、文書差分検出装置の格納手段と、抽出手段とを有し、前記抽出手段で、前記入力された1個の文書データの現在の前記抽出領域以外の領域から全ての前記抽出単位に相当するものを抽出して前記文書差分検出装置の格納手段に格納し、現在の前記抽出領域において、前記文書差分検出装置の格納手段に格納されていない前記抽出単位に相当するものを抽出することを、前記抽出領域ごとに繰り返す文書差分検出装置を備え、前記ユーザが指定した部分として、前記文書差分検出装置の前記抽出手段が抽出した前記文書差分検出装置の格納手段に格納されていない前記抽出単位に相当するものを指定する。このため、入力された1個の文書データ中でユーザが指定した部分の説明に該当する場所をより明確に見つけることができる。

30

【0016】

(9)：抽出の単位を設定する抽出単位設定手段21と、入力された1個の文書データの抽出領域の場所を設定する抽出領域設定手段と、情報を格納する格納手段3aと、抽出手段2と、前記抽出手段2で、入力された1個の文書データの現在の抽出領域において、文書差分検出装置の格納手段に格納されていない前記抽出単位に相当するものを強調表示して現在の抽出領域の文書を出力し、前記強調表示したものを前記文書差分検出装置の格納手段に格納することを、繰り返す文書差分検出装置と、抽出領域の場所として前記文書差分検出装置が強調表示した部分を設定する抽出領域設定手段22とを備え、前記抽出手段2は、前記抽出領域より前記抽出の単位に相当するものを抽出して前記格納手段3aに格納し、前記入力された1個の文書データを調べて現在の前記抽出の単位に相当するものと同じものが、前記格納手段3aに格納されている場合は、それを強調表示する。このた

40

50

め、入力された1個の文書データ中で初めて出てくる単語等の抽出の単位に相当する部分の説明に該当する場所を簡単な手段で明確に見つけることができる。

【0017】

(10)：抽出の単位を設定する抽出単位設定手段21と、入力された1個の文書データの抽出領域の場所を設定する抽出領域設定手段と、情報を格納する格納手段3aと、抽出手段2と、前記抽出手段2で、前記入力された1個の文書データの現在の前記抽出領域以外の抽出領域から全ての前記抽出単位に相当するものを抽出して文書差分抽出装置の格納手段に格納し、現在の前記抽出領域において、前記文書差分抽出装置の格納手段に格納されていない前記抽出単位に相当するものを強調表示して現在の抽出領域の文書を出力することを、前記抽出領域ごとに繰り返す文書差分抽出装置と、抽出領域の場所として前記文書差分抽出装置が強調表示した部分を設定する抽出領域設定手段22とを備え、前記抽出手段2は、前記抽出領域より前記抽出の単位に相当するものを抽出して前記格納手段3aに格納し、前記入力された1個の文書データを調べて現在の前記抽出の単位に相当するものと同じものが、前記格納手段3aに格納されている場合は、それを強調表示する。このため、入力された1個の文書データ中で初めて出てくる単語等の抽出の単位に相当する部分の説明に該当する場所を簡単な手段で明確に見つけることができる。

10

【0018】

【発明の実施の形態】

(1)：キーワード強調装置の説明

図2はキーワード強調装置の説明図である。図2において、キーワード強調装置には、入力手段1、抽出手段2、抽出物記憶装置3、出力手段4が設けてある。入力手段1は、キーボード、マウス、読み取り装置等の情報の入力を行うものである。抽出手段2は、入力された文書の差分を抽出するものである。抽出物記憶装置3は、単語、漢字、名詞句などの抽出物を格納する格納手段である。出力手段4は、表示装置、プリンタ等の情報の出力を行うものである。

20

【0019】

(2)：キーワード強調装置の処理の説明

図3はキーワード強調装置の処理フローチャートである。以下、図3の処理S1～S4-2に従って説明する。

【0020】

S1：入力手段1等により、予め抽出の単位(抽出単位)、抽出表現を定める。抽出の単位には、「単語」「漢字」「名詞句」などが考えられる。抽出表現には、「名詞」「動詞」などの主要な品詞などが考えられる。また、抽出表現は助詞、助動詞、空白などの主要でない品詞等を除くように設定することもできる。

30

【0021】

S2：あらかじめ抽出領域の場所を定める。抽出領域としては、タイトル、ユーザが指定した領域などがある。

【0022】

タイトルの場合は、タイトルの配置の構造からタイトル部分と認識(タイトル認識手段)する。例えば、ウェブ(web)テキストならタイトルを意味するSGML(Standard Generalized Markup Language)表現を利用する。他の場合でもタイトルの部分だけフォントが変わっているや、また、本文との間に改行があるものについては、改行までをタイトルと認識するという方法もとることができる。

40

【0023】

ユーザが指定した領域の認識(指定領域認識手段)は、例えば、ユーザの指定方法をマウスでドラッグして領域を反転させる方法を用いる場合は、反転している箇所を認識すればよい。

【0024】

S3：抽出手段2は、抽出領域より抽出の単位に相当するものを抽出し、そのうち、抽出表現に該当するものを抽出物記憶装置3に格納する。

50

【 0 0 2 5 】

このとき、抽出の単位が日本語の単語の場合は、単語の取得には形態素解析が必要である。また英語の単語の場合は、stemming（語幹化）のアルゴリズムが必要である。

【 0 0 2 6 】

また、抽出表現の指定に品詞などを用いる場合は形態素解析、もしくは品詞タグづけシステムが必要になる。

【 0 0 2 7 】

S 4：抽出手段 2 は、入力されたデータを左から調べて現在の抽出の単位に相当するものから処理 S 1 で定めた抽出の単位に相当するものごとに以下の処理 S 4 - 1 と処理 S 4 - 2 を繰り返す。

【 0 0 2 8 】

S 4 - 1：抽出手段 2 は、現在の抽出の単位に相当するものと同じものが抽出物記憶装置 3 に格納されている場合は、それを強調表示する。

【 0 0 2 9 】

S 4 - 2：抽出手段 2 は、現在の抽出の単位に相当するものと同じものが抽出物記憶装置 3 に格納されていない場合は、それを強調表示せずそのまま表示する。

【 0 0 3 0 】

1：形態素解析システムの説明

日本語を単語に分割するために、抽出手段 2 が行う形態素解析システムが必要になる。ここではChaSenについて説明する（奈良先端大で開発されている形態素解析システム茶筌 <http://chasen.aist-nara.ac.jp/index.html.jp> で公開されている）。

【 0 0 3 1 】

これは、日本語文を分割し、さらに、各単語の品詞も推定してくれる。例えば、「学校へ行く」を入力すると以下の結果を得ることができる。

【 0 0 3 2 】

学校	ガッコウ	学校	名詞 - 一般		
へ	へ	へ	助詞 - 格助詞 - 一般		
行く	イク	行く	動詞 - 自立	五段・力行促音便	基本型

E O S

このように各行に一個の単語が入るように分割され、各単語に読みや品詞の情報が付与される。ここで、分割された単語を抽出の単位に用い、付与された品詞が抽出表現の指定に用いられる。

【 0 0 3 3 】

2：英語のstemmer（ステマー）の説明

抽出手段 2 で単語を抽出するには、英語では単語はわかち書きされているので、単語を基本形式に戻すstemmingをするだけでよい。このstemmingをするアルゴリズムとしては有名なPorterのものがある。（Porter, M.F., 1980, An algorithm for suffix stripping, Program, 14(3) :130-137 参照）。

【 0 0 3 4 】

また、英語の文で単語に品詞を付けるシステム（英語の品詞タグ付けシステム）としては、Brill の文献が有名である。なお、出力の表現は、前記ChaSenと同様である。（Eric Brill, Transformation-Based Error-Driven Learning and Natural Language Processing: A Case Study in Part-of-Speech Tagging, Computational Linguistics, Vol.21, No.4, p.543-565, 1995. 参照）。

【 0 0 3 5 】

（例による説明）

以下、具体例により、タイトルを利用したキーワード強調を入力例と出力例により説明をする。（記事は毎日新聞より取得）。

【 0 0 3 6 】

（入力例 1）

10

20

30

40

50

[余録]今年こそ優しく生きよう

「私の大切にしたいのは/その国の大きさでも繁栄でもない/その国はごく小さくていいし/すこしは武器らしいものを持つが/誰(だれ)も使おうとしない」あれあれ、これからの日本のあるべき姿かなと思ったら、古代中国の哲学者、老子の言葉だった。「そこに住む人はみんな/生きることに死ぬことを大切にすることから/船や車で遠くとびだしたりしない」というくだりもある 詩人、加島祥造さんの訳した「老子」の一節だ。数年前、旅行先で何気なく英訳本を手にしたのが老子との出会いという。加島さんは信州・伊那谷の山荘にこもり、十数冊の英訳をもとに翻訳を進めた。古めかしい「老子」が驚くほど若返った。今日的になった 「あくまでも頑張る軍隊は全滅する/木も、堅く突立ったものは風に折れる/しなやかで、柔らかで/弱くて繊細なものこそ/上に位置を占めて/花を咲かせるべきなのだ」という文章もある。日本という国も少々突っ張りすぎた 加島さんの「老子」体験は雑誌「学燈」に連載された「伊那谷の『老子』」に詳しい。その老子は中国服の伝説的人物というより、レーンコートを着て、英語の本を携え、伊那谷を散歩する上品な老紳士のような。「大切なのは優しさ、柔らかさ」と老子は語る 思えばこれまで、優しさ、柔らかさより、強さや硬さを追いかけていたような気がする。「弱いように見えるものが強いものを従え/柔らかいものが固いものを征服する」という老子の予言が胸にずっしり響く。今年こそ優しく生きよう。伊那谷にこぼれる日差しのように。

10

【0037】

(出力例1)

《[余録]今年》こそ《優しく生きよ》う

「私の大切にしたいのは/その国の大きさでも繁栄でもない/その国はごく小さくていいし/すこしは武器らしいものを持つが/誰(だれ)も使おうとしない」あれあれ、これからの日本のあるべき姿かなと思ったら、古代中国の哲学者、老子の言葉だった。「そこに住む人はみんな/生きることに死ぬことを大切にすることから/船や車で遠くとびだしたりしない」というくだりもある 詩人、加島祥造さんの訳した「老子」の一節だ。数年前、旅行先で何気なく英訳本を手にしたのが老子との出会いという。加島さんは信州・伊那谷の山荘にこもり、十数冊の英訳をもとに翻訳を進めた。古めかしい「老子」が驚くほど若返った。今日的になった 「あくまでも頑張る軍隊は全滅する/木も、堅く突立ったものは風に折れる/しなやかで、柔らかで/弱くて繊細なものこそ/上に位置を占めて/花を咲かせるべきなのだ」という文章もある。日本という国も少々突っ張りすぎた 加島さんの「老子」体験は雑誌「学燈」に連載された「伊那谷の『老子』」に詳しい。その老子は中国服の伝説的人物というより、レーンコートを着て、英語の本を携え、伊那谷を散歩する上品な老紳士のような。「大切なのは優しさ、柔らかさ」と老子は語る 思えばこれまで、優しさ、柔らかさより、強さや硬さを追いかけていたような気がする。「弱いように見えるものが強いものを従え/柔らかいものが固いものを征服する」という老子の予言が胸にずっしり響く。《今年》こそ《優しく生きよ》う。伊那谷にこぼれる日差しのように。

20

30

。

【0038】

ここではタイトル中の助詞、助動詞、記号 - 空白以外をキーワードとした。形態素解析にはChaSenを用いた。「《」、「》」(二重山括弧)で囲っているものが強調表示されている。

40

本文では最後の方の「《今年》こそ《優しく生きよ》う。」の部分が強調表示されている。この部分の周辺が重要とわかる。したがって、この部分を中心に読めばよいとわかるので、便利である。

【0039】

(入力例2)

大統領官邸など炎上 首都中心部で激戦 チェチェン共和国

[モスクワ31日飯島一孝]ロシア南部チェチェン共和国の首都グロズヌイに進攻したロシア軍は三十一日、首都中心部を装甲車などで攻撃、大統領官邸など数カ所が炎上した。

50

ロシア側は首都制圧の最終段階に入ったとみられる。

グロズヌイからの報道では、ロシア軍は激しい空爆と砲撃を加えた後、装甲車部隊が大統領官邸付近に進出。同官邸前などでドゥダエフ政権部隊と激しい市街戦を展開している。一方、ドゥダエフ政権側の首都防衛司令官は同日夕、テレビを通じ、首都防衛はうまくいっており、ロシア軍の戦車五十両を破壊したと発表。また、ドゥダエフ大統領は現在、交渉中のロシア議会の代表団とともに防空ごうに避難しており、無事という。ドゥダエフ大統領は三十日夜、エリツィン・ロシア大統領に正月休戦を提案したが、ロシア側はこれを黙殺した。(この記事には図「クロヌイの中心街略図」があります)

【0040】

(出力例2)

《大統領官邸》など《炎上》《首都中心部》で《激戦 チェチェン共和国》[モスクワ31日飯島一孝]ロシア南部《チェチェン共和国》の《首都》グロズヌイに進攻したロシア軍は三十一日、《首都中心部》を装甲車などで攻撃、《大統領官邸》など数カ所が《炎上》した。ロシア側は《首都》制圧の最終段階に入ったとみられる。

グロズヌイからの報道では、ロシア軍は激しい空爆と砲撃を加えた後、装甲車部隊が《大統領官邸》付近に進出。同《官邸》前などでドゥダエフ政権部隊と激しい市街戦を展開している。

一方、ドゥダエフ政権側の《首都》防衛司令官は同日夕、テレビを通じ、《首都》防衛はうまくいっており、ロシア軍の戦車五十両を破壊したと発表。また、ドゥダエフ《大統領》は現在、交渉中のロシア議会の代表団とともに防空ごうに避難しており、無事という。ドゥダエフ《大統領》は三十日夜、エリツィン・ロシア《大統領》に正月休戦を提案したが、ロシア側はこれを黙殺した。(この記事には図「クロヌイの《中心》街略図」があります)

【0041】

この文書では、重要なキーワードの「チェチェン共和国」は第一段落にあることがすぐわかる。「チェチェン共和国」に興味があれば第一段落を中心に読むとよいことがわかる。

【0042】

(入力例3)

[列島ロジャー]あなたのまちで/1 寄付金拋出企業のリスト判明 長野五輪招致委
拋出した後、工事を受注 問われる情報公開

会計帳簿の紛失で、巨額な活動資金の詳細が不明だった長野冬季オリンピック招致委員会。県・市職員の出向者が大半のこの公的組織に対する寄付金の内訳を示す「企業・金額リスト」を三十一日、毎日新聞は入手した。寄付金約十億円余のうち最も多かったのはゼネコンなど建設業界の約三億三千万円。ほとんどの企業がその後に五輪競技施設や土木工事を受注していた。これまで招致委は寄付側のプライバシー保護を理由に企業名を非公開にしてきたが、リストには発注権を持つ自治体と企業との「もたれあいの構図」が浮かぶ。一方、使い道では、国際オリンピック委員会(IOC)委員への巨額の接待漬けのほか、宣伝用ビデオ作製に二億円もかけたことが分かった。税金や免税の寄付金を扱う自治体の

巨大イベントは、市民への情報公開が問われている。(社会面に関連記事)
招致委員会の事務局が置かれた長野市役所八階の会議室。一九九〇年四月、県、長野市、地元財界などの関係者数人が集まった。

【0043】

「こういう計画で金を集めますから」

県幹部が、出席者にコピーを配った。欄外に招致委の名称が印刷された便せんに、「事前資金計画案」とあった。「支出」「収入」「予算不足額」などの項目があり、それぞれ数字が羅列されていた。ひときわ目を引いたのは「不足額の財源手当て」という欄。「八日会 200」と書かれていた。単位は百万円、二億円だった。

【0044】

10

20

30

40

50

八日会は、県外の大手ゼネコンなど三十八社で作る親ぼく団体。「実態は県発注工事の受注調整をする談合組織」と建設関係者は語り、ゼネコン事件が起きた一昨年解散した。寄付に応じたゼネコン幹部は「(会議から三カ月後の)九〇年七月、県幹部から八日会に要請があった。ほとんどの社は翌九一年三月、一斉に支払った」と証言する。

【0045】

長野冬季オリンピック招致委員会は、八九年十月、任意団体として発足。長野県の吉村午良知事が会長に就任した。東京、札幌五輪は国家プロジェクトだったが、長野は自治体の自立の意味も込めて県市主導の招致活動を行った。名簿によると同委事務局五十一人の九割が長野県・市・関係町村職員。八日会への寄付金依頼について吉村知事は「ないと思う」と語った。

10

【0046】

毎日新聞が入手した寄付リストによれば、業態別には別表のようになる。トップの建設・土木の会社別では、ゼネコン十二社が各一千万円のほか、二千万円、五百万円、百五十万円がそれぞれ一社。寄付をした建設・土木・資材会社は中小合わせて延べ約六百社に上った。これら財界からの寄付は特定公益増進法人の日本体育協会を経由しており、免税措置を受けている。

【0047】

一方、五輪競技場建設の受注状況と寄付社の関係を見ると、すでに契約が交わされているフィギュア、スピードスケート、アイスホッケーA、ボブスレー・リュージュ会場、ジャンプ台(ラージ、ノーマル)、開閉会式会場の本体工事などを受注しているゼネコンは全社が数百万円から一千万円寄付していた。また、ある大手通信機器会社は、同業他社が横並びで三百万から五百万円の寄付だったところに、その倍の一千万円を拠出。この会社は、八九年から九二年にかけ、防災無線デジタル化など、他社より格段に多い総額三十億円に上る工事を受注した。

20

【0048】

同委は活動資金についてこれまで大枠の、収入約二十一億七千万円(県交付金六億百万円、関係市町村分担金二億三千万円、財界寄付金十億八千万円など)、支出十九億六千万円(内訳は広報費など五項目)以外は非公開としている》

膨大な報告書に、簡単な支出説明

長野五輪は九一年六月のIOCバーミンガム総会で米国ソルトレークシティなど四都市を退け、第十八回冬季大会(九八年二月)開催が決まった。招致委は約二億円の余剰金を、現在の長野五輪組織委員会に寄付し、九一年十月解散。二百六十八ページの膨大な「招致報告書」を出した。ところが肝心の招致費の支出については、簡単な決算状況に続き、「募金実績」として約十億円の数字と企業数が書いてある五ページだけ。県当局が市民に公表するには、余りにも少ない情報量だった。

30

【0049】

× × ×

戦後半世紀を迎える一九九五年が明けた。今年は地方分権推進法が国会に上程され、統一地方選が予定されている。地方自治を足元から見つめ直し、新たなあり方を探る年になりそうだ。地方自治を考える「列島ロジャー」の第一部では「あなたのまち」の現状、税金の無駄遣いや情報の非公開性などの問題点を、生活者、納税者の立場から点検する。

40

【0050】

主な業界別の寄付額(毎日新聞集計)

建設(ゼネコン、建設、橋りょうなど、)	約3億3000万円
開発、不動産・住宅	約7400万円
銀行・証券	約5300万円
食品	約4200万円
コンピューター・通信	約3400万円
大手電気設備	約3200万円
自動車関連	約2600万円

50

電機製造 約1300万円
 長野県オーナー会社 約7600万円

 【0051】

(出力例3)

《[列島ロジー]あなた》の《まち》で《/1》 《寄付金拠出企業》の《リスト判明
 長野五輪招致委》

《拠出》した後、工事を受注《 》問われる情報公開
 会計帳簿の紛失で、巨額な活動資金の詳細が不明だった《長野》冬季オリンピック《招致
 委》員会。県・市職員の出向者が大半のこの公的組織に対する《寄付金》の内訳を示す「 10
 《企業》・金額《リスト》」を三十一日、毎日新聞は入手した。《寄付金》約十億円余の
 うち最も多かったのはゼネコンなど建設業界の約三億三千万円。ほとんどの《企業》がそ
 の後に《五輪》競技施設や土木工事を受注していた。これまで《招致委》は《寄付》側の
 プライバシー保護を理由に《企業》名を非公開にしてきたが、《リスト》には発注権を持
 つ自治体と《企業》との「もたれあいの構図」が浮かぶ。一方、使い道では、国際オリン
 ピック委員会(IOC)委員への巨額の接待漬けのほか、宣伝用ビデオ作製に二億円もか
 けたことが分かった。税金や免税の寄付《金》を扱う自治体の巨大イベントは、市民への
 情報公開が問われている。(社会面に関連記事)

《招致》委員会の事務局が置かれた《長野》市役所八階の会議室。一九九〇年四月、県、
 《長野》市、地元財界などの関係者数人が集まった。 20

【0052】

「こういう計画で《金》を集めますから」

県幹部が、出席者にコピーを配った。欄外に《招致委》の名称が印刷された便せんに、「
 事前資金計画案」とあった。「支出」「収入」「予算不足額」などの項目があり、それぞ
 れ数字が羅列されていた。ひときわ目を引いたのは「不足額の財源手当て」という欄。「
 八日会 200」と書かれていた。単位は百万円、二億円だった。

【0053】

八日会は、県外の手ゼネコンなど三十八社で作る親ぼく団体。「実態は県発注工事の受
 注調整をする談合組織」と建設関係者は語り、ゼネコン事件が起きた一昨年解散した。《
 寄付》に応じたゼネコン幹部は「(会議から三カ月後の)九〇年七月、県幹部から八日会 30
 に要請があった。ほとんどの社は翌九一年三月、一斉に支払った」と証言する。

【0054】

《長野》冬季オリンピック《招致》委員会は、八九年十月、任意団体として発足。《長野
 》県の吉村午良知事が会長に就任した。東京、札幌《五輪》は国家プロジェクトだったが
 、《長野》は自治体の自立の意味も込めて県市主導の《招致》活動を行った。名簿による
 と同委事務局五十一人の九割が《長野》県・市・関係町村職員。八日会への《寄付金》依
 頼について吉村知事は「ないと思う」と語った。

【0055】

毎日新聞が入手した《寄付リスト》によれば、業態別には別表のようになる。トップの建
 設・土木の会社別では、ゼネコン十二社が各一千万円のほか、二千万円、五百万円、百五
 十万円がそれぞれ一社。《寄付》をした建設・土木・資材会社は中小合わせて延べ約六百
 社に上った。これら財界からの《寄付》は特定公益増進法人の日本体育協会を經由してお
 り、免税措置を受けている。 40

【0056】

一方、《五輪》競技場建設の受注状況と《寄付》社の関係を見ると、すでに契約が交わさ
 れているフィギュア、スピードスケート、アイスホッケーA、ボブスレー・リュージュ会
 場、ジャンプ台(ラージ、ノーマル)、開閉会式会場の本体工事などを受注しているゼネ
 コンは全社が数百万円から一千万円《寄付》していた。また、ある大手通信機器会社は、
 同業他社が横並びで三百万から五百万円の《寄付》だったところに、その倍の一千万円を
 《拠出》。この会社は、八九年から九二年にかけ、防災無線デジタル化など、他社より格 50

段に多い総額三十億円に上る工事を受注した。

【0057】

同委は活動資金についてこれまで大枠の、収入約二十一億七千万円（県交付《金》六億八千万円、関係市町村分担《金》二億三千万円、財界《寄付金》十億八千万円など）、支出十九億六千万円（内訳は広報費など五項目）以外は非公開としている

膨大な報告書に、簡単な支出説明

《長野五輪》は九一年六月のIOCバーミンガム総会で米国ソルトレークシティーなど四都市を退け、第十八回冬季大会（九八年二月）開催が決まった。《招致委》は約二億円の余剰《金》を、現在の《長野五輪》組織委員会に《寄付》し、九一年十月解散。二百六十八ページの膨大な「《招致》報告書」を出した。ところが肝心の《招致》費の支出については、簡単な決算状況に続き、「募金実績」として約十億円の数字と《企業》数が書いてある五ページだけ。県当局が市民に公表するには、余りにも少ない情報量だった。

10

【0058】

× × ×

戦後半世紀を迎える一九九五年が明けた。今年も地方分権推進法が国会に上程され、統一地方選が予定されている。地方自治を足元から見つめ直し、新たなあり方を探る年になりそうだ。地方自治を考える「《列島ロジ》」の第一部では「《あなた》の《まち》」の現状、税金の無駄遣いや情報の非公開性などの問題点を、生活者、納税者の立場から点検する。

【0059】

20

主な業界別の《寄付》額（毎日新聞集計）

建設（ゼネコン、建設、橋りょうなど、）	約3億3000万円	
開発、不動産・住宅	約7400万円	
銀行・証券	約5300万円	
食品	約4200万円	
コンピューター・通信	約3400万円	
大手電気設備	約3200万円	
自動車関連	約2600万円	
電機製造	約1300万円	
《長野》県人オーナー会社	約7600万円	30

【0060】

この文章では、キーワードが集中している第一段落が重要だとわかる。また、「毎日新聞が入手した《寄付リスト》によれば」で始まる段落は、「寄付リスト」に関する情報が重点的に書いてあることがわかり便利である。

ここでは、それぞれ一行目をタイトルと自動認識している。

【0061】

（3）：強調表示する単語が二単語連続している場合の説明

強調表示する単語が二単語連続している場合は、その部分は特に強調表示（特別強調表示）する。

【0062】

40

図4は強調表示する単語が二単語連続している場合、その部分は特に強調表示する処理フローチャートである。以下、図4の処理S11～S14-5に従ってキーワード強調装置の処理を説明する。

【0063】

S11：入力手段1等により、予め抽出の単位（抽出単位）、抽出表現を定める。抽出の単位には、「単語」「漢字」「名詞句」などが考えられる。抽出表現には、「名詞」などの品詞などが考えられる。なお、この抽出表現は助詞、助動詞、空白、記号などの主要でない品詞等を除くように設定することもできる。

【0064】

S12：あらかじめ抽出領域の場所を定める。抽出領域としては、タイトル、ユーザが指

50

定した領域などがある。

【 0 0 6 5 】

タイトルの場合、タイトルの配置の構造からタイトル部分と認識する。例えば、ウェブ（web）テキストならタイトルを意味する S G M L（Standard Generalized Markup Language）表現を利用する。他の場合でもタイトルの部分だけフォントが変わっているや、また、本文との間に改行があるものについては、改行までをタイトルと認識するという方法もとることができる。

【 0 0 6 6 】

ユーザが指定した領域の認識は、例えば、ユーザの指定方法をマウスでドラッグして領域を反転させる方法を用いる場合は、反転している箇所を認識すればよい。

10

【 0 0 6 7 】

S 1 3：抽出手段 2 は、抽出領域より抽出の単位に相当するものを抽出し、そのうち、抽出表現に該当するものを抽出物記憶装置 3 に格納する。

【 0 0 6 8 】

このとき、抽出の単位が日本語の単語の場合は、単語の取得には形態素解析が必要である。また英語の単語の場合は、stemming（語幹化）のアルゴリズムが必要である。

【 0 0 6 9 】

また、抽出表現の指定に品詞などを用いる場合は、形態素解析、もしくは品詞タグづけシステムが必要になる。

【 0 0 7 0 】

S 1 4：抽出手段 2 は、入力されたデータを左から調べて現在の抽出の単位に相当するものから処理 S 1 1 で定めた抽出の単位に相当するものごとに以下の処理 S 1 4 - 1 ~ 処理 S 1 4 - 5 を繰り返す。

20

【 0 0 7 1 】

S 1 4 - 1：抽出手段 2 は、現在の抽出の単位に相当するものと同じものが抽出物記憶装置 3 に格納されている場合で、一つ前の抽出の単位に相当するものと同じものが抽出物記憶装置 3 に格納されていない場合は、現在の抽出の単位に相当するものを第一強調物として記憶する。

【 0 0 7 2 】

S 1 4 - 2：抽出手段 2 は、現在の抽出の単位に相当するものと同じものが抽出物記憶装置 3 に格納されている場合で、一つ前の抽出の単位が第一強調物として記憶されている場合は、第一強調物と現在の抽出の単位に相当するものを特別強調表示して表示する。

30

【 0 0 7 3 】

S 1 4 - 3：抽出手段 2 は、現在の抽出の単位に相当するものと同じものが抽出物記憶装置 3 に格納されている場合で、一つ前の抽出の単位を特別強調表示した場合は、現在の抽出の単位に相当するものを特別強調表示して表示する。

【 0 0 7 4 】

S 1 4 - 4：抽出手段 2 は、現在の抽出の単位に相当するものと同じものが抽出物記憶装置 3 に格納されていない場合で、一つ前の抽出の単位が第一強調物として記憶されている場合は、第一強調物を通常強調表示して表示する。現在の抽出の単位に相当するものを強調表示せずにそのまま表示する。

40

【 0 0 7 5 】

S 1 4 - 5：抽出手段 2 は、現在の抽出の単位に相当するものと同じものが抽出物記憶装置 3 に格納されていない場合で、一つ前の抽出の単位が第一強調物として記憶されていない場合は、現在の抽出の単位に相当するものを強調表示せずにそのまま表示する。

【 0 0 7 6 】

（例による説明）

以下の例では、抽出の単位は単語、抽出表現は助詞、助動詞、空白記号を除くすべての単語である。すなわち、抽出表現は名詞、動詞等の主要な品詞の単語となる。

【 0 0 7 7 】

50

(入力例1)

大統領官邸など炎上 首都中心部で激戦 チェチェン共和国

[モスクワ31日飯島一孝]ロシア南部チェチェン共和国の首都グロズヌイに進攻したロシア軍は三十一日、首都中心部を装甲車などで攻撃、大統領官邸など数カ所が炎上した。ロシア側は首都制圧の最終段階に入ったとみられる。

【0078】

グロズヌイからの報道では、ロシア軍は激しい空爆と砲撃を加えた後、装甲車部隊が大統領官邸付近に進出。同官邸前などでドゥダエフ政権部隊と激しい市街戦を展開している。

【0079】

一方、ドゥダエフ政権側の首都防衛司令官は同日夕、テレビを通じ、首都防衛はうまくいっており、ロシア軍の戦車五十両を破壊したと発表。また、ドゥダエフ大統領は現在、交渉中のロシア議会の代表団とともに防空ごうに避難しており、無事という。ドゥダエフ大統領は三十日夜、エリツィン・ロシア大統領に正月休戦を提案したが、ロシア側はこれを黙殺した。(この記事には図「クロヌイの中心街略図」があります)

10

【0080】

(出力例1)

《《大統領官邸》》など《炎上》《《首都中心部》》で《《激戦 チェチェン共和国》》

[モスクワ31日飯島一孝]ロシア南部《《チェチェン共和国》》の《首都》グロズヌイに進攻したロシア軍は三十一日、《《首都中心部》》を装甲車などで攻撃、《《大統領官邸》》など数カ所が《炎上》した。ロシア側は《首都》制圧の最終段階に入ったとみられる。

20

【0081】

グロズヌイからの報道では、ロシア軍は激しい空爆と砲撃を加えた後、装甲車部隊が《《大統領官邸》》付近に進出。同《官邸》前などでドゥダエフ政権部隊と激しい市街戦を展開している。

【0082】

一方、ドゥダエフ政権側の《首都》防衛司令官は同日夕、テレビを通じ、《首都》防衛はうまくいっており、ロシア軍の戦車五十両を破壊したと発表。また、ドゥダエフ《大統領》は現在、交渉中のロシア議会の代表団とともに防空ごうに避難しており、無事という。ドゥダエフ《大統領》は三十日夜、エリツィン・ロシア《大統領》に正月休戦を提案したが、ロシア側はこれを黙殺した。(この記事には図「クロヌイの《中心》街略図」があります)

30

【0083】

ここでは、「通常強調表示」は「《》」(二重山括弧)で、「特別強調表示」は「《《》》」(二つの二重山括弧)で、囲っている。「通常強調表示」の箇所は多いが、「特別強調表示」の箇所はそれほど多くなく、目立つことになる。また、第三段落では「特別強調表示」はなく、それほど重要な段落でないといえる。なお、強調表示として、ここでは二重山括弧を用いたが他の強調表示を行うこともできる。例えば、通常の文字を「黒」とし、通常強調表示は「青」、特別強調表示は「赤」と色分けすることもできる。

40

【0084】

(入力例2)

[列島ロジャー]あなたのまちで/1 寄付金抛出企業のリスト判明 長野五輪招致委
抛出した後、工事を受注 問われる情報公開

会計帳簿の紛失で、巨額な活動資金の詳細が不明だった長野冬季オリンピック招致委員会。県・市職員の出向者が大半のこの公的組織に対する寄付金の内訳を示す「企業・金額リスト」を三十一日、毎日新聞は入手した。寄付金約十億円余のうち最も多かったのはゼネコンなど建設業界の約三億三千万円。ほとんどの企業がその後五輪競技施設や土木工事

50

を受注していた。これまで招致委は寄付側のプライバシー保護を理由に企業名を非公開にしてきたが、リストには発注権を持つ自治体と企業との「もたれあいの構図」が浮かぶ。一方、使い道では、国際オリンピック委員会（ＩＯＣ）委員への巨額の接待漬けのほか、宣伝用ビデオ作製に二億円もかけたことが分かった。税金や免税の寄付金を扱う自治体の巨大イベントは、市民への情報公開が問われている。（社会面に関連記事）

招致委員会の事務局が置かれた長野市役所八階の会議室。一九九〇年四月、県、長野市、地元財界などの関係者数人が集まった。

【 0 0 8 5 】

「こういう計画で金を集めますから」

県幹部が、出席者にコピーを配った。欄外に招致委の名称が印刷された便せんに、「事前資金計画案」とあった。「支出」「収入」「予算不足額」などの項目があり、それぞれ数字が羅列されていた。ひときわ目を引いたのは「不足額の財源手当て」という欄。「八日会 200」と書かれていた。単位は百万円、二億円だった。

【 0 0 8 6 】

八日会は、県外の大手ゼネコンなど三十八社で作る親ぼく団体。「実態は県発注工事の受注調整をする談合組織」と建設関係者は語り、ゼネコン事件が起きた一昨年解散した。寄付に応じたゼネコン幹部は「（会議から三カ月後の）九〇年七月、県幹部から八日会に要請があった。ほとんどの社は翌九一年三月、一斉に支払った」と証言する。

【 0 0 8 7 】

長野冬季オリンピック招致委員会は、八九年十月、任意団体として発足。長野県の吉村午良知事が会長に就任した。東京、札幌五輪は国家プロジェクトだったが、長野は自治体の自立の意味も込めて県市主導の招致活動を行った。名簿によると同委事務局五十一人の九割が長野県・市・関係町村職員。八日会への寄付金依頼について吉村知事は「ないと思う」と語った。

【 0 0 8 8 】

毎日新聞が入手した寄付リストによれば、業態別には別表のようになる。トップの建設・土木の会社別では、ゼネコン十二社が各一千万円のほか、二千万円、五百万円、百五十万円がそれぞれ一社。寄付をした建設・土木・資材会社は中小合わせて延べ約六百社に上った。これら財界からの寄付は特定公益増進法人の日本体育協会を経由しており、免税措置を受けている。

【 0 0 8 9 】

一方、五輪競技場建設の受注状況と寄付社の関係を見ると、すでに契約が交わされているフィギュア、スピードスケート、アイスホッケーA、ボブスレー・リュージュ会場、ジャンプ台（ラージ、ノーマル）、開閉会式会場の本体工事などを受注しているゼネコンは全社が数百万円から一千万円寄付していた。また、ある大手通信機器会社は、同業他社が横並びで三百万から五百万円の寄付だったところに、その倍の一千万円を拠出。この会社は、八九年から九二年にかけ、防災無線デジタル化など、他社より格段に多い総額三十億円に上る工事を受注した。

【 0 0 9 0 】

《同委は活動資金についてこれまで大枠の、収入約二十一億七千万円（県交付金六億百万円、関係市町村分担金二億三千万円、財界寄付金十億八千万円など）、支出十九億六千万円（内訳は広報費など五項目）以外は非公開としている》

膨大な報告書に、簡単な支出説明

長野五輪は九一年六月のＩＯＣバーミンガム総会で米国ソルトレークシティなど四都市を退け、第十八回冬季大会（九八年二月）開催が決まった。招致委は約二億円の余剰金を、現在の長野五輪組織委員会に寄付し、九一年十月解散。二百六十八ページの膨大な「招致報告書」を出した。ところが肝心の招致費の支出については、簡単な決算状況に続き、「募金実績」として約十億円の数字と企業数が書いてある五ページだけ。県当局が市民に公表するには、余りにも少ない情報量だった。

【 0 0 9 1 】

10

20

30

40

50

× × ×

戦後半世紀を迎える一九九五年が明けた。今年は地方分権推進法が国会に上程され、統一地方選が予定されている。地方自治を足元から見つめ直し、新たなあり方を探る年になりそうだ。地方自治を考える「列島ロジー」の第一部では「あなたのまち」の現状、税金の無駄遣いや情報の非公開性などの問題点を、生活者、納税者の立場から点検する。

【0092】

主な業界別の寄付額（毎日新聞集計）

建設（ゼネコン、建設、橋りょうなど、）	約3億3000万円	
開発、不動産・住宅	約7400万円	
銀行・証券	約5300万円	10
食品	約4200万円	
コンピューター・通信	約3400万円	
大手電気設備	約3200万円	
自動車関連	約2600万円	
電機製造	約1300万円	
長野県オーナー会社	約7600万円	

【0093】

（出力例2）

《《[列島]》《《ロジー]あなた》》の《まち》で《《ノ1》》《《寄付金》》《《抛出企業》》の《《リスト判明》》《《長野》》《《五輪招致委》》 20

《抛出》した後、工事を受注《 》問われる情報公開
会計帳簿の紛失で、巨額な活動資金の詳細が不明だった《長野》冬季オリンピック《《招致委》》員会。県・市職員の出向者が大半のこの公的組織に対する《《寄付金》》の内訳を示す「《企業》・金額《リスト》」を三十一日、毎日新聞は入手した。《《寄付金》》約十億円余のうち最も多かったのはゼネコンなど建設業界の約三億三千万円。ほとんどの《企業》がその後《五輪》競技施設や土木工事を受注していた。これまで《《招致委》》は《寄付》側のプライバシー保護を理由に《企業》名を非公開にしてきたが、《リスト》には発注権を持つ自治体と《企業》との「もたれあいの構図」が浮かぶ。一方、使い道では、国際オリンピック委員会（IOC）委員への巨額の接待漬けのほか、宣伝用ビデオ 30
作製に二億円もかけたことが分かった。税金や免税の寄付《金》を扱う自治体の巨大イベントは、市民への情報公開が問われている。（社会面に関連記事）

《招致》委員会の事務局が置かれた《長野》市役所八階の会議室。一九九〇年四月、県、《長野》市、地元財界などの関係者数人が集まった。

【0094】

「こういう計画で《金》を集めますから」

県幹部が、出席者にコピーを配った。欄外に《《招致委》》の名称が印刷された便せんに、「事前資金計画案」とあった。「支出」「収入」「予算不足額」などの項目があり、それぞれ数字が羅列されていた。ひときわ目を引いたのは「不足額の財源手当て」という欄。「八日会 200」と書かれていた。単位は百万円、二億円だった。 40

【0095】

八日会は、県外の手ゼネコンなど三十八社で作る親ぼく団体。「実態は県発注工事の受注調整をする談合組織」と建設関係者は語り、ゼネコン事件が起きた一昨年解散した。《寄付》に応じたゼネコン幹部は「（会議から三カ月後の）九〇年七月、県幹部から八日会に要請があった。ほとんどの社は翌九一年三月、一斉に支払った」と証言する。

【0096】

《長野》冬季オリンピック《招致》委員会は、八九年十月、任意団体として発足。《長野》県の吉村午良知事が会長に就任した。東京、札幌《五輪》は国家プロジェクトだったが、《長野》は自治体の自立の意味も込めて県市主導の《招致》活動を行った。名簿によると同委事務局五十一人の九割が《長野》県・市・関係町村職員。八日会への《《寄付金》》 50

》依頼について吉村知事は「ないと思う」と語った。

【0097】

毎日新聞が入手した《《寄付リスト》》によれば、業態別には別表のようになる。トップの建設・土木の会社別では、ゼネコン十二社が各一千万円のほか、二千万円、五百万円、百五十万円がそれぞれ一社。《寄付》をした建設・土木・資材会社は中小合わせて延べ約六百社に上った。これら財界からの《寄付》は特定公益増進法人の日本体育協会を經由しており、免税措置を受けている。

【0098】

一方、《五輪》競技場建設の受注状況と《寄付》社の関係をみると、すでに契約が交わされているフィギュア、スピードスケート、アイスホッケーA、ボブスレー・リュージュ会場、ジャンプ台（ラージ、ノーマル）、開閉会式会場の本体工事などを受注しているゼネコンは全社が数百万円から一千万円《寄付》していた。また、ある大手通信機器会社は、同業他社が横並びで三百万から五百万円の《寄付》だったところに、その倍の一千万円を《拠出》。この会社は、八九年から九二年にかけ、防災無線デジタル化など、他社より格段に多い総額三十億円に上る工事を受注した。

【0099】

《同委は活動資金についてこれまで大枠の、収入約二十一億七千万円（県交付《金》六億百万円、関係市町村分担《金》二億三千万円、財界《《寄付金》》十億八千万円など）、支出十九億六千万円（内訳は広報費など五項目）以外は非公開としている》

膨大な報告書に、簡単な支出説明

《《長野五輪》》は九一年六月のIOCバーミンガム総会で米国ソルトレークシティーなど四都市を退け、第十八回冬季大会（九八年二月）開催が決まった。《《招致委》》は約二億円の余剰《金》を、現在の《《長野五輪》》組織委員会に《寄付》し、九一年十月解散。二百六十八ページの膨大な「《招致》報告書」を出した。ところが肝心の《招致》費の支出については、簡単な決算状況に続き、「募金実績」として約十億円の数字と《企業》数が書いてある五ページだけ。県当局が市民に公表するには、余りにも少ない情報量だった。

【0100】

× × ×

戦後半世紀を迎える一九九五年が明けた。今年は地方分権推進法が国会に上程され、統一地方選が予定されている。地方自治を足元から見つめ直し、新たなあり方を探る年になりそうだ。地方自治を考える「《《列島ロジ》》」の第一部では「《あなた》の《まち》」の現状、税金の無駄遣いや情報の非公開性などの問題点を、生活者、納税者の立場から点検する。

【0101】

主な業界別の《寄付》額（毎日新聞集計）

建設（ゼネコン、建設、橋りょうなど、）	約3億3000万円
開発、不動産・住宅	約7400万円
銀行・証券	約5300万円
食品	約4200万円
コンピューター・通信	約3400万円
大手電気設備	約3200万円
自動車関連	約2600万円
電機製造	約1300万円
《長野》県人オーナー会社	約7600万円

【0102】

こちらの例もだいたいみやすくなる。例えば、真ん中あたりに、毎日新聞が入手した《《寄付リスト》》によれば、業態別には別表のようになる。トップの建設・・という文があるが、これのおかげで「寄付リスト」という重要なものがこのあたりにあり、さらにそれが別表にあるとわかる。

10

20

30

40

50

【 0 1 0 3 】

「寄付金」「招致委」「長野五輪」といったものが特別強調表示されているが、「列島ロジ」 という面白いものも特別強調表示されていることがわかり、表題にあった「列島ロジ」とはなにかと疑問をもっていた人も本文の特別強調表示されている「列島ロジ」の箇所を見て納得することにもなる。

【 0 1 0 4 】

(4) : ユーザの領域指定を利用した場合の例の説明

あらかじめ抽出領域の場所を定めるとき、抽出領域としてユーザが指定した領域とするものである。ユーザが指定した領域の認識は、例えば、ユーザの指定方法をマウスでドラッグして領域を反転させる方法を用いる場合は、反転している箇所を認識すればよい。強調表示する単語が二単語連続している場合は、その部分は特に強調表示する。

10

【 0 1 0 5 】

(具体例による説明)

(入力例 1) 特許文書の場合

〔請求項 1〕刃部材の先端の刃部を凹凸に形成し波状刃とするとともに螺旋状に湾曲させ、前記刃部材に取っ手を取り付けたことを特徴とする草取り鎌。

〔請求項 2〕取っ手の上部及び下部に滑り止め部を設けたことを特徴とする草取り鎌。

【 0 1 0 6 】

(中略)

次に、本発明について図面を参照して説明する。図 1 は本発明である草取り鎌の正面図、図 2 は本発明である草取り鎌の背面図、図 3 は、本発明である草取り鎌の右側面である。

20

〔 0 0 0 7 〕本草取り鎌 1 は、図 3 に示すように、刃部材 2 の刃部 2 b は当該先端の一面が波状の波状刃 5 に形成されているとともに背面が平坦に形成されている刃部材 2 と、取っ手 3 から構成されている。

〔 0 0 0 8 〕刃部材 2 は、図 1、図 2 及び図 3 に示すように、延長部 2 a があり取っ手 3 の約 2 倍程の長さがある。波状刃 5 の刃部 2 b は一方向に湾曲している。

〔 0 0 0 9 〕図 4 は本発明の草取り鎌の刃部の正面拡大図である。図に示すように、雑草を刈り取る刃部 2 b は、凸部 5 a と凹部 5 b が交互に存在し波状となっている。

〔 0 0 1 0 〕図 5 は本発明である草取り鎌の刃部の拡大図である。刃部 2 b を構成する凸部 5 a の先端はやや左方向に傾いている。これは、雑草をより引っ掛けて刈り取り易くするためである。

30

〔 0 0 1 1 〕図 6 は本発明である草取り鎌の刃部の湾曲状態を示した一部拡大図である。図に示すように、刃部 2 b の延長部 2 a より刃部 2 b の先端 2 c は垂直線 6 からより湾曲している。

〔 0 0 1 2 〕図 7 は、図中の A - A 線に沿った断面図である。刃部 2 b の上面 7 は傾斜し、凸部 5 a の先端 5 c は尖っている。そして、刃部 2 b 自体が湾曲するとともに螺旋している。

〔 0 0 1 3 〕図 8 は、本発明である草取り鎌の他の実施例の正面図、図 9 は本発明である草取り鎌の他の実施例の背面図、図 1 0 は本発明である草取り鎌の他の実施例の右側面図、図 1 1 は、本発明である草取り鎌の他の実施例の一部拡大図である。

40

〔 0 0 1 4 〕本例の草取り鎌 1 a は、刃部材 2 の延長部 2 a が短いとともに刃部 2 b の部分がやや大きく形成してある。

〔 0 0 1 5 〕また、取っ手 3 が長く、握り部 3 b の上に、握り部 3 b の径よりやや大きい径の上滑り止め部 3 a を設けるとともに、下にも同様に握り部 3 b より大径の下滑り止め部 3 c b を設けてある。

〔 0 0 1 6 〕図 1 0 に示すように、本例の草取り鎌 1 a の刃部 2 b も図 1 から図 7 までに示した草取り鎌 1 と同様に螺旋状に湾曲している。

〔 0 0 1 7 〕このように、先端部が螺旋状に湾曲させることにより、芝生等に生えている雑草を根こそぎ取り除くことが容易となる。

50

ユーザが、〔請求項2〕の文章のうち、「取っ手の上部及び下部に滑り止め部を設けたこと」の部分だけを領域指定したとする。そうすると以下の結果を得る。

【0107】

(出力例1)

〔請求項1〕刃部材の先端の刃《部》を凹凸に形成し波状刃とするとともに螺旋状に湾曲させ、前記刃部材に《取っ手》を取り付けたことを特徴とする草取り鎌。

〔請求項2〕《取っ手》の《《上部及び下部》》に《《滑り止め部》》を《設け》たことを特徴とする草取り鎌。

(中略)

次に、本発明について図面を参照して説明する。図1は本発明である草取り鎌の正面図、図2は本発明である草取り鎌の背面図、図3は、本発明である草取り鎌の右側面である。

10

〔0007〕本草取り鎌1は、図3に示すように、刃部材2の刃《部》2bは当該先端の一面が波状の波状刃5に形成されているとともに背面が平坦に形成されている刃部材2と、《取っ手》3から構成されている。

〔0008〕刃部材2は、図1、図2《及び》図3に示すように、延長《部》2aがあり《取っ手》3の約2倍程の長さがある。波状刃5の刃《部》2bは一方向に湾曲している。

〔0009〕図4は本発明の草取り鎌の刃《部》の正面拡大図である。図に示すように、雑草を刈り取る刃《部》2bは、凸《部》5aと凹《部》5bが交互に存在し波状となっている。

20

〔0010〕図5は本発明である草取り鎌の刃《部》の拡大図である。刃《部》2bを構成する凸《部》5aの先端はやや左方向に傾いている。これは、雑草をより引っ掛けて刈り取り易くするためである。

〔0011〕図6は本発明である草取り鎌の刃《部》の湾曲状態を示した一部拡大図である。図に示すように、刃《部》2bの延長《部》2aより刃《部》2bの先端2cは垂直線6からより湾曲している。

〔0012〕図7は、図中のA-A線に沿った断面図である。刃《部》2bの上面7は傾斜し、凸《部》5aの先端5cは尖っている。そして、刃《部》2b自体が湾曲するとともに螺旋している。

〔0013〕図8は、本発明である草取り鎌の他の実施例の正面図、図9は本発明である草取り鎌の他の実施例の背面図、図10は本発明である草取り鎌の他の実施例の右側面図、図11は、本発明である草取り鎌の他の実施例の一部拡大図である。

30

〔0014〕本例の草取り鎌1aは、刃部材2の延長《部》2aが短いとともに刃《部》2bの部分がやや大きく形成してある。

〔0015〕また、《取っ手》3が長く、握り《部》3bの上に、握り《部》3bの径よりやや大きい径の上滑り《《止め部》》3aを設けるとともに、下にも同様に握り《部》3bより大径の下《《滑り止め部》》3cbを《設け》てある。

〔0016〕図10に示すように、本例の草取り鎌1aの刃《部》2bも図1から図7までに示した草取り鎌1と同様に螺旋状に湾曲している。

〔0017〕このように、先端《部》が螺旋状に湾曲させることにより、芝生等に生えている雑草を根こそぎ取り除くことが容易となる。

40

【0108】

この特許文書で、特別強調表示がされている部分を探すと、「請求項2」がまず目につく。次に、段落「0015」が目につく。このことにより、「請求項2」に関連する内容が段落「0015」にあることがすぐにわかる。

【0109】

請求項を読んでいるとき、それに対応する実施例を探したいなと思うことがよくある。この発明を用いるとそれを容易に実現できる。

【0110】

ところで、「上滑り止め部」のうち、「止め部」の部分しか強調表示されていないが、こ

50

れは形態素解析システムの影響で「上滑り」を一単語と誤解析したためである。

【0111】

(5) : 文書差分検出装置を用いる場合の説明

図5は文書差分検出装置を用いる場合の説明図である。図5において、キーワード強調装置には、入力手段1、抽出手段2、抽出物記憶装置3、出力手段4、文書差分検出装置5が設けてある。入力手段1は、キーボード、マウス、読み取り装置等の情報の入力を行うものである。抽出手段2は、入力された文書の差分を抽出するものである。抽出物記憶装置3は、単語、漢字、名詞句などの抽出物を格納する抽出物記憶手段である。出力手段4は、表示装置、プリンタ等の情報の出力を行うものである。文書差分検出装置5は、入力された文章のうちで初めての文字列を強調表示するものである。

10

【0112】

図6は文書差分検出装置の説明図である。図6において、文書差分検出装置には抽出手段51、格納手段52が設けてあり、抽出手段51には抽出・検出領域設定手段53が設けてある。抽出手段51は、入力された文書の差分を抽出するものである。格納手段52は、単語、漢字、名詞句などの抽出物を格納する抽出物記憶手段である。抽出・検出領域設定手段53は、抽出の単位(抽出単位)や検出領域の単位を設定するものである。

【0113】

なお、抽出手段51として抽出手段2を用いることもでき、また、格納手段52として抽出物記憶装置3を用いることができる。

【0114】

この文書差分強調装置で、入力された文章のうちで初めての文字列を強調表示すべきと判断する手法は、次の手法1、2がある(特願2002-290946参照)。

20

【0115】

(手法1)

1 入力部1等により、予め抽出の単位(抽出単位)、検出領域の単位を定める。抽出単位とは、差分として出力する対象の単位である。抽出単位には、「単語」「漢字」「名詞句」などが考えられる。検出領域の単位とは、差分を検出するために比較する領域の単位のことである。検出領域の単位には、「文字」「単語」「文」「箇条書の項目」「段落」「特許の請求項」などが考えられる。

【0116】

2 抽出手段51は、すべての入力データを記憶手段(抽出手段51内の)に記憶させる。

30

【0117】

3 抽出手段51は、入力されたデータを左から調べて左の検出領域から1で定めた検出領域ごとに以下の処理4と処理5を繰り返す。

【0118】

4 抽出手段51は、現在の検出領域以外の領域すべてから、すべての抽出単位に相当するもの(例えば単語)を抽出し、それを格納手段52に格納する。

【0119】

5 抽出手段51は、現在の検出領域において、格納手段52に格納されていない抽出単位に相当するもの(例えば単語)を強調表示して現在の検出領域の文章を出力する。

40

【0120】

(手法2)

1 入力部1等により、予め抽出の単位(抽出単位)、検出領域の単位を定める。抽出単位とは、差分として出力する対象の単位である。抽出単位には、「単語」「漢字」「名詞句」などが考えられる。検出領域の単位とは、差分を検出するために比較する領域の単位のことである。検出領域の単位には、「文字」「単語」「文」「箇条書の項目」「段落」「特許の請求項」などが考えられる。

【0121】

2 入力1部から前記1で定めた検出領域ごとに入力データが入力され、抽出手段

50

5 1 は、以下の処理 3 と処理 4 を繰り返す。

【0122】

3 抽出手段 5 1 は、現在の検出領域において、格納手段 5 2 に格納されていない抽出単位に相当するもの（例えば単語）を強調表示して現在の検出領域の文章を出力する。ただし、格納手段 5 2 は最初は空である。

4 前記処理 1 で強調表示した表現を格納手段 5 2 に格納する。

【0123】

（具体例による説明）

・（手法 1 の例 1）特許文書の場合で、文書差分強調装置の上記手法 1 の出力結果が次のものであるとする。

【0124】

〔請求項 1〕《刃部材》の《先端》の《刃》部を《凹凸》に《形成し波状刃》とする《とともに螺旋状》に《湾曲させ、前記刃部材》に取っ手を《取り付け》たことを特徴とする草取り鎌。

〔請求項 2〕取っ手の《上部及び下部》に《滑り止め》部を《設け》たことを特徴とする草取り鎌。

【0125】

ここで、請求項 1 又は 2 だけに出てくる単語が強調表示（2 重山括弧で囲む）されている。これに対して、ユーザが領域指定したとして、例えば、次の請求項 2 だけの領域を指定したとする。

〔請求項 2〕取っ手の《上部及び下部》に《滑り止め》部を《設け》たことを特徴とする草取り鎌。

すると、抽出手段 2 で、文書差分強調装置の強調表示と領域指定の重なった「上部及び下部」「滑り止め」「設け」だけをユーザが指定した領域と判断して同じアルゴリズムを用いる。

【0126】

（出力例 1）

〔請求項 1〕刃部材の先端の刃部を凹凸に形成し波状刃とするとともに螺旋状に湾曲させ、前記刃部材に取っ手を取り付けたことを特徴とする草取り鎌。

〔請求項 2〕取っ手の《《上部及び下部》》に《《滑り止め》》部を《設け》たことを特徴とする草取り鎌。

（中略）

次に、本発明について図面を参照して説明する。図 1 は本発明である草取り鎌の正面図、図 2 は本発明である草取り鎌の背面図、図 3 は、本発明である草取り鎌の右側面である。

〔0007〕本草取り鎌 1 は、図 3 に示すように、刃部材 2 の刃部 2 b は当該先端の一面が波状の波状刃 5 に形成されているとともに背面が平坦に形成されている刃部材 2 と、取っ手 3 から構成されている。

〔0008〕刃部材 2 は、図 1、図 2 及び図 3 に示すように、延長部 2 a があり取っ手 3 の約 2 倍程の長さがある。波状刃 5 の刃部 2 b は一方向に湾曲している。

〔0009〕図 4 は本発明の草取り鎌の刃部の正面拡大図である。図に示すように、雑草を刈り取る刃部 2 b は、凸部 5 a と凹部 5 b が交互に存在し波状となっている。

〔0010〕図 5 は本発明である草取り鎌の刃部の拡大図である。刃部 2 b を構成する凸部 5 a の先端はやや左方向に傾いている。これは、雑草をより引っ掛けて刈り取り易くするためである。

〔0011〕図 6 は本発明である草取り鎌の刃部の湾曲状態を示した一部拡大図である。図に示すように、刃部 2 b の延長部 2 a より刃部 2 b の先端 2 c は垂直線 6 からより湾曲している。

〔0012〕図 7 は、図中の A - A 線に沿った断面図である。刃部 2 b の上面 7 は傾斜し、凸部 5 a の先端 5 c は尖っている。そして、刃部 2 b 自体が湾曲するとともに螺旋している。

10

20

30

40

50

〔 0 0 1 3 〕 図 8 は、本発明である草取り鎌の他の実施例の正面図、図 9 は本発明である草取り鎌の他の実施例の背面図、図 1 0 は本発明である草取り鎌の他の実施例の右側面図、図 1 1 は、本発明である草取り鎌の他の実施例の一部拡大図である。

〔 0 0 1 4 〕 本例の草取り鎌 1 a は、刃部材 2 の延長部 2 a が短いとともに刃部 2 b の部分がやや大きく形成してある。

〔 0 0 1 5 〕 また、取っ手 3 が長く、握り部 3 b の上に、握り部 3 b の径よりやや大きい径の上滑り《止め》部 3 a を設けるとともに、下にも同様に握り部 3 b より大径の下《滑り止め》部 3 c b を《設け》てある。

〔 0 0 1 6 〕 図 1 0 に示すように、本例の草取り鎌 1 a の刃部 2 b も図 1 から図 7 までに示した草取り鎌 1 と同様に螺旋状に湾曲している。

〔 0 0 1 7 〕 このように、先端部が螺旋状に湾曲させることにより、芝生等に生えている雑草を根こそぎ取り除くことが容易となる。

この出力だと、より明確に段落番号 0 0 1 5 が、請求項 2 に対応する箇所であるとよくわかる。

【 0 1 2 7 〕

・（手法 1 の例 2）特許文書の場合で、文書差分強調装置の上記手法 1 の出力結果が次のものであるとする。

【 0 1 2 8 〕

〔請求項 1〕《刃部材》の《先端》の《刃》部を《凹凸》に《形成し波状刃》とする《とともに螺旋状》に《湾曲させ、前記刃部材》に取っ手を《取り付け》たことを特徴とする草取り鎌。

〔請求項 2〕取っ手の《上部及び下部》に《滑り止め》部を《設け》たことを特徴とする草取り鎌。

【 0 1 2 9 〕

ここでは、請求項 1 又は 2 だけに出てくる単語が強調表示（2 重山括弧で囲む）されている。これに対して、ユーザが領域指定したとして、例えば、次の請求項 1、2 の 2 つの領域（指定領域 1、2）を指定したとする。

【 0 1 3 0 〕

（指定領域 1）

〔請求項 1〕《刃部材》の《先端》の《刃》部を《凹凸》に《形成し波状刃》とする《とともに螺旋状》に《湾曲させ、前記刃部材》に取っ手を《取り付け》たことを特徴とする草取り鎌。

（指定領域 2）

〔請求項 2〕取っ手の《上部及び下部》に《滑り止め》部を《設け》たことを特徴とする草取り鎌。

【 0 1 3 1 〕

すると、抽出手段 2 で、文書差分強調装置の強調表示と領域指定の重なった指定領域 1 の「刃部材」「先端」「刃」「凹凸」「形成し波状刃」「とともに螺旋状」「湾曲させ、前記刃部材」「取り付け」と指定領域 2 の「上部及び下部」「滑り止め」「設け」だけをユーザが指定した領域と判断して同じアルゴリズムを用いる。ここで指定領域 1 と指定領域 2 のキーワードでは異なる強調表示をする。

【 0 1 3 2 〕

（指定領域により異なる強調表示する出力例）

〔請求項 1〕《《刃部材》》の《先端》の《刃》部を《凹凸》に《《形成し》》《《波状刃》》とするとともに《《螺旋状》》に《《湾曲さ》》《《せ、》》《《前記刃部材》》に取っ手を《取り付け》たことを特徴とする草取り鎌。

〔請求項 2〕取っ手の 上部及び下部 に 滑り止め 部を 設け たことを特徴とする草取り鎌。

（中略）

次に《、》本発明について図面を参照《し》て説明する。図 1 は本発明である草取り鎌の

10

20

30

40

50

正面図《、》図2は本発明である草取り鎌の背面図《、》図3は《、》本発明である草取り鎌の右側面である。

〔0007〕本草取り鎌1は《、》図3に示すように《《、刃部材》》2の《刃》部2bは当該《先端》の一面が《波状》の《《波状刃》》5に《《形成さ》》れているとともに背面が平坦に《《形成さ》》れている《《刃部材》》2と《、》取っ手3から構成《さ》れている。

〔0008〕《《刃部材》》2は《、》図1《、》図2 及び 図3に示すように《、》延長部2aがあり取っ手3の約2倍程の長《さ》がある。《《波状刃》》5の《刃》部2bは一方向に《《湾曲し》》ている。

〔0009〕図4は本発明の草取り鎌の《刃》部の正面拡大図である。図に示すように《、》雑草を刈り取る《刃》部2bは《、》凸部5aと凹部5bが交互に存在《《し波状》》となっている。 10

〔0010〕図5は本発明である草取り鎌の《刃》部の拡大図である。《刃》部2bを構成する凸部5aの《先端》はやや左方向に傾いている。これは《、》雑草をより引っ掛けて刈り取り易くするためである。

〔0011〕図6は本発明である草取り鎌の《刃》部の《湾曲》状態を示した一部拡大図である。図に示すように《《、刃》》部2bの延長部2aより《刃》部2bの《先端》2cは垂直線6からより《《湾曲し》》ている。

〔0012〕図7は《、》図中のA-A線に沿った断面図である。《刃》部2bの上面7は傾斜《《し、》》凸部5aの《先端》5cは尖っている。そして《《、刃》》部2b自体が《湾曲》するとともに《《螺旋し》》ている。 20

〔0013〕図8は《、》本発明である草取り鎌の他の実施例の正面図《、》図9は本発明である草取り鎌の他の実施例の背面図《、》図10は本発明である草取り鎌の他の実施例の右側面図《、》図11は《、》本発明である草取り鎌の他の実施例の一部拡大図である。

〔0014〕本例の草取り鎌1aは《《、刃部材》》2の延長部2aが短いとともに《刃》部2bの部分がやや大きく《《形成し》》てある。

〔0015〕また《、》取っ手3が長く《、》握り部3bの上に《、》握り部3bの径よりやや大きい径の上滑り止め部3aを設けるとともに《、》下にも同様に握り部3bより大径の下滑り止め部3cbを設けてある。 30

〔0016〕図10に示すように《、》本例の草取り鎌1aの《刃》部2bも図1から図7までに示した草取り鎌1と同様に《《螺旋状》》に《《湾曲し》》ている。

〔0017〕このように《《、先端》》部が《《螺旋状》》に《《湾曲さ》》せることにより《、》芝生等に生えている雑草を根こそぎ取り除くことが容易となる。

【0133】

ここでは、請求項1から得られたキーワードによる強調表示では、「《、》」（二重山括弧）を用い、請求項2から得られたキーワードによる強調表示では、「、」（一重の山括弧）を用いている。また、それぞれの強調表示をまとめて表示していることになる。 40

【0134】

この出力だと、請求項1に関係するのは全体に広がっていて、なおかつ、請求項2に関係するのは段落番号0015だということがわかる。このことが、同時にわかるのである。便利である。

【0135】

なお、ここで指定領域は、指定領域1、2の二つの領域を指定したが三つ以上を指定することもできる。また、強調表示として、2重山括弧以外に、下線、色分け、背景の変更、字体の変更、点滅等他の強調表示を行うこともできる。

【0136】

(4)：プログラムインストールの説明

入力手段1、抽出手段2、抽出物記憶装置3、格納手段3a、52、出力手段4、文書差 50

分検出装置 5、抽出単位設定手段 2 1、抽出表現設定手段 2 2、抽出領域設定手段 2 3、抽出・検出領域設定手段 5 3 等は、プログラムで構成でき、主制御部（CPU）が実行するものであり、主記憶に格納されているものである。このプログラムは、一般的な、コンピュータで処理されるものである。このコンピュータは、主制御部、主記憶、ファイル装置、表示装置、キーボード等の入力手段である入力装置などのハードウェアで構成されている。このコンピュータに、本発明のプログラムをインストールする。このインストールは、フロッピィ、光磁気ディスク等の可搬型の記録（記憶）媒体に、これらのプログラムを記憶させておき、コンピュータが備えている記録媒体に対して、アクセスするためのドライブ装置を介して、或いは、LAN等のネットワークを介して、コンピュータに設けられたファイル装置にインストールされる。そして、このファイル装置から処理に必要なプログラムステップを主記憶に読み出し、主制御部が実行するものである。

10

【0137】

【発明の効果】

以上説明したように、本発明によれば、次のような効果がある。

【0138】

（1）：抽出手段で、抽出領域より抽出の単位に相当するものを抽出し、そのうち、抽出表現に該当するものを格納手段に格納し、入力された1個の文書データを調べて現在の前記抽出の単位に相当するものと同じものが、前記格納手段に格納されている場合は、それを強調表示するため、入力された1個の文書データ中で抽出領域の説明に該当する場所を容易に見つけることができる。

20

【0139】

（2）：入力された1個の文書データを調べて、現在の抽出の単位に相当するものと同じものが格納手段に格納されている場合で、一つ前の抽出の単位が強調表示されるものである場合は、前記一つ前の抽出の単位と現在の抽出の単位に相当するものを特別強調表示して表示するため、入力された1個の文書データ中で抽出領域の説明に該当する場所をより明確に見つけることができる。

【0140】

（3）：抽出表現設定手段で強調表示する抽出表現の設定の少なくとも一つが名詞であるため、名詞等の重要なもののみ強調表示することができる。

【0141】

（4）：抽出領域の場所として、入力された1個の文書データのタイトル部分とするため、入力された1個の文書データ中で重要なものとされているタイトル部分の説明に該当する場所を容易に見つけることができる。

30

【0142】

（5）：抽出領域の場所として、入力された1個の文書データのユーザが指定した部分とするため、入力された1個の文書データ中でユーザが指定した部分の説明に該当する場所を容易に見つけることができる。

【0143】

（6）：ユーザが指定した部分として、複数部分を指定し、該指定した部分により、異なる強調表示をするため、入力された1個の文書データ中でユーザが指定した複数部分の説明に該当する場所を容易に見つけることができる。

40

【0144】

（7）：抽出手段で、入力された1個の文書データの現在の検出領域において、文書差分検出装置の格納手段に格納されていない抽出単位に相当するものを抽出し、該抽出したものを前記文書差分検出装置の格納手段に格納することを繰り返す文書差分検出装置を備え、ユーザが指定した部分として、前記文書差分検出装置の前記抽出手段が抽出した前記文書差分検出装置の格納手段に格納されていない前記抽出単位に相当するものを指定するため、入力された1個の文書データ中でユーザが指定した部分の説明に該当する場所をより明確に見つけることができる。

【0145】

50

(8) : 抽出手段で、入力された 1 個の文書データの現在の検出領域以外の領域から全ての抽出単位に相当するものを抽出して文書差分検出装置の格納手段に格納し、現在の検出領域において、前記文書差分検出装置の格納手段に格納されていない前記抽出単位に相当するものを抽出することを、前記検出領域ごとに繰り返す文書差分検出装置を備え、ユーザが指定した部分として、前記文書差分検出装置の前記抽出手段が抽出した前記文書差分検出装置の格納手段に格納されていない前記抽出単位に相当するものを指定するため、入力された 1 個の文書データ中でユーザが指定した部分の説明に該当する場所をより明確に見つけることができる。

【 0 1 4 6 】

(9) : 文書差分検出装置の格納手段に格納されていない抽出単位に相当するものを強調表示して現在の検出領域の文書を出力し、前記強調表示したものを前記文書差分検出装置の格納手段に格納することを繰り返す文書差分検出装置を備え、抽出領域の場所として文書差分検出装置が強調表示した部分とし、抽出表現設定手段を用いなくて、抽出手段で、前記抽出領域より前記抽出の単位に相当するものを抽出して格納手段に格納し、前記入力された 1 個の文書データを調べて現在の前記抽出の単位に相当するものと同じものが、前記格納手段に格納されている場合は、それを強調表示するため、入力された 1 個の文書データ中で初めて出てくる単語等の抽出の単位に相当する部分の説明に該当する場所を簡単な手段で明確に見つけることができる。

【 0 1 4 7 】

(1 0) : 入力された 1 個の文書データの現在の検出領域以外の検出領域から全ての抽出単位に相当するものを抽出して文書差分検出装置の格納手段に格納し、現在の前記検出領域において、前記文書差分検出装置の格納手段に格納されていない前記抽出単位に相当するものを強調表示して現在の検出領域の文書を出力することを前記検出領域ごとに繰り返す文書差分検出装置を備え、抽出領域の場所として前記文書差分検出装置が強調表示した部分とし、抽出表現設定手段を用いなくて、前記抽出手段で、前記抽出領域より前記抽出の単位に相当するものを抽出して前記格納手段に格納し、前記入力された 1 個の文書データを調べて現在の前記抽出の単位に相当するものと同じものが、前記格納手段に格納されている場合は、それを強調表示するため、入力された 1 個の文書データ中で初めて出てくる単語等の抽出の単位に相当する部分の説明に該当する場所を簡単な手段で明確に見つけることができる。

【 0 1 4 8 】

(1 1) : 抽出の単位を設定する抽出単位設定手段と、強調表示する抽出表現を設定する抽出表現設定手段と、入力された 1 個の文書データの抽出領域の場所を設定する抽出領域設定手段と、前記抽出領域より前記抽出の単位に相当するものを抽出し、そのうち、前記抽出表現に該当するものを前記格納手段に格納し、前記入力された 1 個の文書データを調べて現在の前記抽出の単位に相当するものと同じものが、前記格納手段に格納されている場合は、それを強調表示する抽出手段として、コンピュータを機能させるためのプログラム又はプログラムを記録したコンピュータ読取可能な記録媒体とするため、このプログラムをコンピュータにインストールすることで、入力された 1 個の文書データ中で抽出領域の説明に該当する場所を容易に見つけることができるキーワード強調装置を容易に提供

【 0 1 4 9 】

(1 2) : 抽出の単位を設定する抽出単位設定手段と、入力された 1 個の文書データの検出領域の場所を設定する検出領域設定手段と、抽出手段で、前記入力された 1 個の文書データの現在の検出領域において、文書差分検出装置の格納手段に格納されていない前記抽出単位に相当するものを強調表示して現在の検出領域の文書を出力し、前記強調表示したものを前記文書差分検出装置の格納手段に格納することを、繰り返す文書差分検出装置と、抽出領域の場所として、前記文書差分検出手段が強調表示した部分を設定する抽出領域設定手段と、前記抽出領域より前記抽出の単位に相当するものを抽出して前記格納手段に格納し、前記入力された 1 個の文書データを調べて現在の前記抽出の単位に相当するも

10

20

30

40

50

のと同じものが、前記格納手段に格納されている場合は、それを強調表示する抽出手段として、コンピュータを機能させるためのプログラム又はプログラムを記録したコンピュータ読取可能な記録媒体とするため、このプログラムをコンピュータにインストールすることで、入力された1個の文書データ中でユーザが指定した部分の説明に該当する場所をより明確に見つけることができるキーワード強調装置を容易に提供することができる。

【0150】

(13)：抽出の単位を設定する抽出単位設定手段と、入力された1個の文書データの検出領域の場所を設定する検出領域設定手段と、抽出手段で、前記入力された1個の文書データの現在の前記検出領域以外の検出領域から全ての前記抽出単位に相当するものを抽出して文書差分検出装置の格納手段に格納し、現在の前記検出領域において、前記文書差分検出装置の格納手段に格納されていない前記抽出単位に相当するものを強調表示して現在の検出領域の文書を出力することを、前記検出領域ごとに繰り返す文書差分検出装置と、抽出領域の場所として前記文書差分検出装置が強調表示した部分を設定する抽出領域設定手段と、前記抽出領域より前記抽出の単位に相当するものを抽出して格納手段に格納し、前記入力された1個の文書データを調べて現在の前記抽出の単位に相当するものと同じものが、前記格納手段に格納されている場合は、それを強調表示する抽出手段として、コンピュータを機能させるためのプログラム又はプログラムを記録したコンピュータ読取可能な記録媒体とするため、このプログラムをコンピュータにインストールすることで、入力された1個の文書データ中で初めて出てくる抽出の単位に相当する部分の説明に該当する場所を簡単な手段で見つけることができるキーワード強調装置を容易に提供することができる。

10

20

【図面の簡単な説明】

【図1】本発明の原理説明図である。

【図2】実施の形態におけるキーワード強調装置の説明図である。

【図3】実施の形態におけるキーワード強調装置の処理フローチャートである。

【図4】実施の形態における強調表示する単語が二単語連続している場合、その部分は特に強調表示する処理フローチャートである。

【図5】実施の形態における文書差分検出装置を用いる場合の説明図である。

【図6】実施の形態における文書差分検出装置の説明図である。

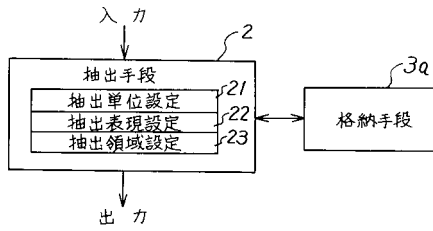
【符号の説明】

- 2 抽出手段
- 3 a 格納手段
- 2 1 抽出単位設定手段
- 2 2 抽出表現設定手段
- 2 3 抽出領域設定手段

30

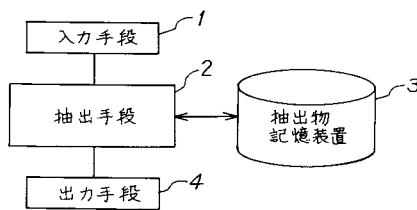
【図1】

本発明の原理説明図



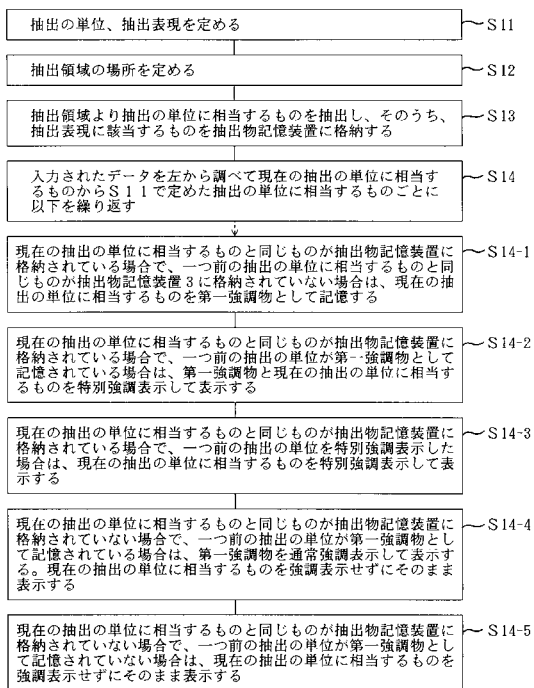
【図2】

キーワード強調装置の説明図



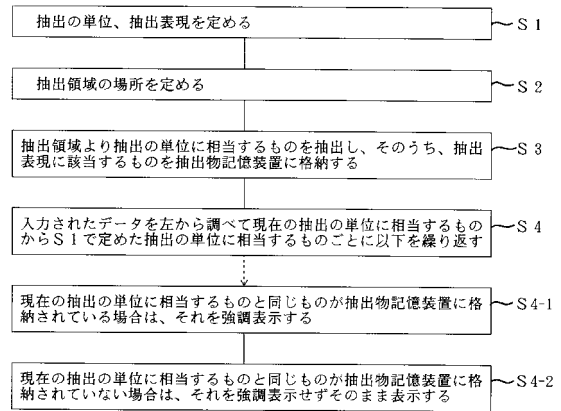
【図4】

強調表示する単語が二単語連続している場合、その部分は特に強調表示する処理フローチャート



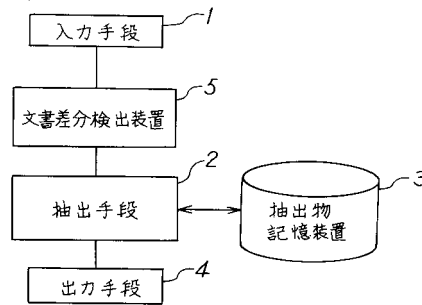
【図3】

キーワード強調装置の処理フローチャート



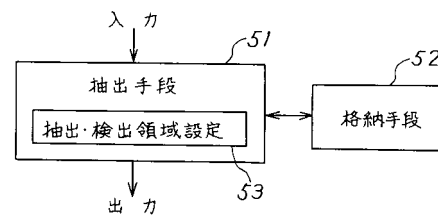
【図5】

文書差分検出装置を用いる場合の説明図



【図6】

文書差分検出装置の説明図



フロントページの続き

- (56)参考文献 特開平02 - 045876 (JP, A)
特開2001 - 134617 (JP, A)
特開2002 - 063197 (JP, A)

(58)調査した分野(Int.Cl., DB名)

G06F 17/21
G06F 17/30