

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第5412981号  
(P5412981)

(45) 発行日 平成26年2月12日 (2014. 2. 12)

(24) 登録日 平成25年11月22日 (2013. 11. 22)

(51) Int. Cl.

F I

G O 6 F 19/00 (2011. 01)

G O 6 F 17/30 (2006. 01)

G O 6 F 19/00 1 3 0

G O 6 F 17/30 2 2 0 Z

G O 6 F 17/30 1 7 0 A

G O 6 F 17/30 2 1 0 A

請求項の数 9 (全 10 頁)

(21) 出願番号 特願2009-148704 (P2009-148704)  
 (22) 出願日 平成21年6月23日 (2009. 6. 23)  
 (65) 公開番号 特開2011-8335 (P2011-8335A)  
 (43) 公開日 平成23年1月13日 (2011. 1. 13)  
 審査請求日 平成24年5月18日 (2012. 5. 18)

早期審査対象出願

(73) 特許権者 000005496  
 富士ゼロックス株式会社  
 東京都港区赤坂九丁目7番3号  
 (74) 代理人 110000154  
 特許業務法人はるか国際特許事務所  
 (72) 発明者 外池 昌嗣  
 神奈川県足柄上郡中井町境430 グリー  
 ンテクなかい 富士ゼロックス株式会社内  
 (72) 発明者 大熊 智子  
 神奈川県足柄上郡中井町境430 グリー  
 ンテクなかい 富士ゼロックス株式会社内  
 (72) 発明者 梅基 宏  
 神奈川県足柄上郡中井町境430 グリー  
 ンテクなかい 富士ゼロックス株式会社内

最終頁に続く

(54) 【発明の名称】 情報処理装置及びプログラム

(57) 【特許請求の範囲】

【請求項 1】

同一文書の中から、対象の第1の状態を表す第1の状態表現、及び前記対象の前記第1の状態より後の第2の状態を表す第2の状態表現を抽出する状態表現抽出手段と、

前記抽出される第1及び第2の状態表現の内容に基づいて、前記対象の前記第1の状態から前記第2の状態への変化を表す変化情報を出力する変化情報出力手段と、

を含むことを特徴とする情報処理装置。

【請求項 2】

前記変化情報出力手段は、前記複数の状態表現のうち、前記文書内に表れる順序が後の状態表現を、前記対象の変化の内容を示す前記変化情報で置き換える

ことを特徴とする請求項1記載の情報処理装置。

【請求項 3】

前記変化情報出力手段は、前記抽出される各状態表現の内容に基づいて、前記対象の変化の極性を評価し、当該評価の結果を示す情報を前記変化情報として出力する

ことを特徴とする請求項1又は2記載の情報処理装置。

【請求項 4】

前記対象の状態表現となりうる複数の候補文字列と、当該各候補文字列により表される状態に対する評価情報と、を含む状態名辞書を記憶する辞書記憶手段をさらに含み、

前記変化情報出力手段は、前記状態名辞書を用いて、前記抽出される各状態表現により表される各状態の評価情報を取得し、当該取得した各評価情報の変化に基づいて、前記対

象の変化の極性を判定する

ことを特徴とする請求項 3 記載の情報処理装置。

【請求項 5】

前記対象に影響する作用を表す作用表現を抽出する作用表現抽出手段と、  
前記出力される変化情報と、前記抽出される作用表現と、に基づいて、前記作用による  
前記対象の変化に関する結果情報を出力する結果情報出力手段と、  
をさらに含むことを特徴とする請求項 1 から 4 のいずれか一項記載の情報処理装置。

【請求項 6】

文書群の中から、対象の第 1 の状態を表す第 1 の状態表現、及び前記対象の前記第 1 の  
状態より後の第 2 の状態を表す第 2 の状態表現を抽出する状態表現抽出手段と、  
前記抽出される第 1 及び第 2 の状態表現の内容に基づいて、前記対象の前記第 1 の状態  
から前記第 2 の状態への変化を表す変化情報を出力する変化情報出力手段と、  
を含むことを特徴とする情報処理装置。

10

【請求項 7】

前記変化情報出力手段は、前記変化情報として、前記対象の前記第 1 の状態から前記第  
2 の状態への変化を表す文字列を出力する  
ことを特徴とする請求項 1 から 6 のいずれか一項記載の情報処理装置。

【請求項 8】

同一文書の中から、対象の第 1 の状態を表す第 1 の状態表現、及び前記対象の前記第 1  
の状態より後の第 2 の状態を表す第 2 の状態表現を抽出する状態表現抽出手段、及び  
前記抽出される第 1 及び第 2 の状態表現の内容に基づいて、前記対象の前記第 1 の状態  
から前記第 2 の状態への変化を表す変化情報を出力する変化情報出力手段、  
としてコンピュータを機能させるためのプログラム。

20

【請求項 9】

文書群の中から、対象の第 1 の状態を表す第 1 の状態表現、及び前記対象の前記第 1 の  
状態より後の第 2 の状態を表す第 2 の状態表現を抽出する状態表現抽出手段、及び  
前記抽出される第 1 及び第 2 の状態表現の内容に基づいて、前記対象の前記第 1 の状態  
から前記第 2 の状態への変化を表す変化情報を出力する変化情報出力手段、  
としてコンピュータを機能させるためのプログラム。

【発明の詳細な説明】

30

【技術分野】

【0001】

本発明は、情報処理装置及びプログラムに関する。

【背景技術】

【0002】

近年、複数の文書を収集し、その中から、何らかの対象に関する評価を表す表現を抽出  
する技術が研究されている（例えば特許文献 1、非特許文献 1 及び 2 参照）。このような  
技術の適用対象となる文書においては、通常、対象に対する評価が多様な表現で記述され  
ている。また、一つの対象に対する評価の観点にも多様性がある。そこで、例えば非特許  
文献 1 及び 2 では、対象に対する評価の観点として属性を定義し、ある対象のある属性に  
関する評価を、対象・属性・評価値の組として収集する技術が提案されている。

40

【先行技術文献】

【特許文献】

【0003】

【特許文献 1】特開 2005 - 235014 号公報

【非特許文献】

【0004】

【非特許文献 1】小林のぞみ、乾健太郎、松本裕治、立石健二、福島俊一、「意見抽出の  
ための評価表現の収集」、自然言語処理、言語処理学会、2005 年 7 月、Vol.12、No.3  
、pp.203-222

50

【非特許文献2】飯田龍、小林のぞみ、乾健太郎、松本裕治、立石健二、福島俊一、「意見抽出を目的とした機械学習による属性 - 評価値対同定」、情報処理学会研究報告・自然言語処理研究会報告、情報処理学会、2005年1月、Vol.2005、No.1、pp.21-28

【発明の概要】

【発明が解決しようとする課題】

【0005】

本発明は、文書内の記述から、対象に生じた変化を評価することのできる情報処理装置及びプログラムを提供することを目的とする。

【課題を解決するための手段】

【0006】

10

請求項1記載の発明は、情報処理装置であって、文書の中から、対象の状態を表す複数の状態表現を抽出する状態表現抽出手段と、前記抽出される各状態表現の内容に基づいて、前記対象の変化に関する変化情報を出力する変化情報出力手段と、を含むことを特徴とする。

【0007】

請求項2記載の発明は、請求項1記載の情報処理装置であって、前記変化情報出力手段は、前記複数の状態表現のうち、前記文書内に表れる順序が後の状態表現を、前記対象の変化の内容を示す前記変化情報で置き換えることを特徴とする。

【0008】

請求項3記載の発明は、請求項1又は2記載の情報処理装置であって、前記変化情報出力手段は、前記抽出される各状態表現の内容に基づいて、前記対象の変化の極性を評価し、当該評価の結果を示す情報を前記変化情報として出力することを特徴とする。

20

【0009】

請求項4記載の発明は、請求項3記載の情報処理装置であって、前記対象の状態表現となりうる複数の候補文字列と、当該各候補文字列により表される状態に対する評価情報と、を含む状態名辞書を記憶する辞書記憶手段をさらに含み、前記変化情報出力手段は、前記状態名辞書を用いて、前記抽出される各状態表現により表される各状態の評価情報を取得し、当該取得した各評価情報の変化に基づいて、前記対象の変化の極性を判定することを特徴とする。

【0010】

30

請求項5記載の発明は、請求項1から4のいずれか一項記載の情報処理装置であって、前記対象に影響する作用を表す作用表現を抽出する作用表現抽出手段と、前記出力される変化情報と、前記抽出される作用表現と、に基づいて、前記作用による前記対象の変化に関する結果情報を出力する結果情報出力手段と、をさらに含むことを特徴とする。

【0011】

請求項6記載の発明は、文書の中から、対象の状態を表す複数の状態表現を抽出する状態表現抽出手段、及び前記抽出される各状態表現の内容に基づいて、前記対象の変化に関する変化情報を出力する変化情報出力手段、としてコンピュータを機能させるためのプログラムである。

【発明の効果】

40

【0012】

請求項1及び6記載の発明によれば、文書内における対象の状態に関する記述に基づいて、対象に生じた変化を評価するための情報を出力できる。

【0013】

請求項2記載の発明によれば、文書中における対象の状態を表す表現を、対象に生じた変化を表す表現に置き換えることができる。

【0014】

請求項3記載の発明によれば、文書内における対象の状態に関する記述に基づいて、対象に生じた変化の極性を表す情報を出力できる。

【0015】

50

請求項４記載の発明によれば、対象の状態が定量的でない態様で記述される場合であっても、対象に生じた変化の極性を表す情報を出力できる。

【００１６】

請求項５記載の発明によれば、作用によって生じる対象への影響を評価するための情報を出力できる。

【図面の簡単な説明】

【００１７】

【図１】本発明の実施の形態に係る情報処理装置の構成例を示す図である。

【図２】本発明の実施の形態に係る情報処理装置が実現する機能の一例を示す機能ブロック図である。

10

【図３】本発明の実施の形態に係る情報処理装置による処理の対象となる文書の一例を示す図である。

【図４】本発明の実施の形態に係る情報処理装置による処理の対象となる文書の別の例を示す図である。

【図５】状態表現を変化表現に置換する処理を実行した結果の一例を示す図である。

【図６】変化表現を極性表現に置換する処理を実行した結果の一例を示す図である。

【図７】変化表現を極性表現に置換する処理を実行した結果の別の例を示す図である。

【発明を実施するための形態】

【００１８】

以下、本発明の実施の形態について、図面を参照しながら説明する。

20

【００１９】

図１は、本発明の一実施形態に係る情報処理装置１の構成例を示す図である。情報処理装置１は、例えばパーソナルコンピュータ等であって、制御部１１と、記憶部１２と、操作部１３と、表示部１４と、を含んで構成される。

【００２０】

制御部１１は、例えばＣＰＵ等であって、記憶部１２に格納されるプログラムに従って各種の情報処理を実行する。本実施形態において制御部１１が実行する処理の具体例については、後述する。

【００２１】

記憶部１２は、例えばＲＡＭ等のメモリ素子やハードディスク等を含んで構成される。記憶部１２は、制御部１１によって実行されるプログラムや、各種のデータを保持する。また、記憶部１２は、制御部１１のワークメモリとしても動作する。

30

【００２２】

操作部１３は、例えばキーボードやマウス等であって、利用者の指示操作を受け付けて、当該指示操作の内容を制御部１１に出力する。表示部１４は、例えば液晶ディスプレイ等であり、制御部１１からの指示に従って、画像の表示を行う。

【００２３】

以下、本実施形態に係る情報処理装置１が実現する機能について、説明する。情報処理装置１は、機能的に、図２に示すように、文書取得部２１と、状態表現抽出部２２と、変化情報出力部２３と、作用表現抽出部２４と、結果情報出力部２５と、を含んで構成される。これらの機能は、例えば制御部１１が記憶部１２に格納されるプログラムを実行することによって実現される。このプログラムは、例えばインターネット等の通信手段を介して提供されてもよいし、光ディスク等の各種のコンピュータ読み取り可能な情報記憶媒体に格納されて提供されてもよい。

40

【００２４】

文書取得部２１は、情報処理装置１による処理の対象となる処理対象文書Ｄを取得する。処理対象文書Ｄは複数の文を含んでおり、何らかの対象（以下、評価対象という）に対して影響を及ぼす作用や、当該評価対象のある時点における状態（すなわち、評価対象に関する事象）などが記録された文書である。ここでは処理対象文書Ｄは、特定の分野に関する文書であって、評価対象に影響する作用や評価対象の状態などが時系列に沿って記録

50

されているものとする。具体例として、処理対象文書Dは、患者に対する診療等の内容が記述された医療文書や、製品に対する修理作業等に関する修理記録文書などであってよい。なお、文書取得部21は、処理対象文書Dを各種の情報記憶媒体から読み出して取得してもよいし、インターネット等の通信手段を介して外部の情報処理装置から取得してもよい。

#### 【0025】

状態表現抽出部22は、文書取得部21が取得した処理対象文書Dの中から、評価対象の状態を表す状態表現を複数個抽出する。具体的に、状態表現抽出部22は、予め定められた評価対象を表す文字列（評価対象表現）を処理対象文書D内から検索し、当該評価対象表現とともに用いられている状態を表現する文字列を、状態表現として抽出する。

10

#### 【0026】

以下、状態表現抽出部22が状態表現を抽出する処理の具体例について、説明する。本実施形態では、評価対象を表す文字列が、予め記憶部12に記憶された対象名辞書に格納されている。さらに、記憶部12には、評価対象の状態を表す状態表現となりうる文字列を定義する情報が、評価対象と関連づけて記憶されている。具体的に、例えば状態が数値情報で表される場合、状態表現は数字列（または数字列＋単位を表す文字列）であることが定義されている。また、評価対象の種別によっては、状態表現となりうる語（候補文字列）が複数個記憶部12内の状態名辞書に格納されてもよいし、状態表現に用いられる文字種（数字や記号、カタカナなど）が定義されていてもよい。

#### 【0027】

20

状態表現抽出部22は、対象名辞書に格納されている評価対象を表す文字列を、処理対象文書D内から検索する。そして、当該評価対象を表す文字列に対して、予め定められた条件を満足するように処理対象文書D内に含まれている文字列を、評価対象の状態を表す状態表現として抽出する。ここで、状態表現が満たすべき条件は、例えば、当該文字列が状態名辞書に登録されており、かつ、処理対象文書D内において評価対象を表す文字列と同じ文や同じ行内に含まれる、などの条件であってよい。

#### 【0028】

一例として、処理対象文書Dが図3に示すような内容の医療に関する記録文書である場合について説明する。ここでは、評価対象は検査項目の一種である「CA125」であり、その状態は検査結果の数値によって表されることとする。すなわち、対象名辞書には「CA125」という語が含まれており、この「CA125」の状態表現は数字列であることが定義されている。この場合、状態表現抽出部22は、対象名辞書に含まれる対象名「CA125」を処理対象文書Dの2行目及び5行目から抽出し、さらに、当該評価対象を表す文字列に続いて処理対象文書D内に表れる数字列（2行目においては「10」、5行目においては「5」）を、評価対象の状態を表す状態表現として抽出する。なお、ここでは処理対象文書D内には1種類の対象名しか含まれないこととしているが、対象名辞書に含まれる複数の対象名が処理対象文書D内に含まれる場合もあり得る。その場合、対象名ごとに、その評価対象の状態を表す状態表現が順に抽出される。

30

#### 【0029】

また、別の例として、処理対象文書Dが図4に示すようなハードディスク装置の修理記録文書である場合について説明する。例えばこの文書の1行目は、ハードディスクから「カッコンカッコン」という音が鳴っている状態を表している。この例では、対象名辞書に「機器から聞こえる音」という語が含まれており、ハードディスクから生じる音を表す擬音語が状態名辞書に複数個含まれていることとする。そして、状態表現抽出部22は、状態名辞書に含まれており、かつ、「機器から聞こえる音」という文字列に続いて処理対象文書D内に表れる語を状態表現として抽出する。

40

#### 【0030】

なお、状態表現抽出部22は、処理対象文書D内の文章に対して公知の構文解析や構文意味解析などを行い、このような解析の結果に基づいて、定められた条件を満たす状態表現を抽出してもよい。具体的に、例えば状態表現抽出部22は、評価対象の状態を表す文

50

字列が否定を表す語句とともに用いられているか否かを構文解析の結果に基づいて判定し、否定を表す語句とともに用いられている場合には状態表現としての条件を満たさないと判定してもよい。こうすれば、状態表現の定義に一致する表現であっても、文書中で実際には評価対象の状態を表現するために用いられていないと推定される場合、そのような表現は状態表現として抽出されなくなる。

#### 【 0 0 3 1 】

変化情報出力部 2 3 は、状態表現抽出部 2 2 によって抽出された各状態表現の内容に基づいて、評価対象の変化に関する変化情報を出力する。本実施形態では、処理対象文書 D 内において、評価対象の状態が時系列に記述されていることを想定している。そこで、変化情報出力部 2 3 は、状態表現抽出部 2 2 が抽出した複数の状態表現のうち、処理対象文書 D 内に表れる順序が後の状態表現（以下、単に「後の状態表現」という）を、評価対象に生じた変化の内容を示す変化情報（変化表現）で置き換えることとする。すなわち、ある状態 S 1 を表す状態表現 P 1 と、別の状態 S 2 を表す状態表現 P 2 と、が処理対象文書 D 内においてこの順に表れる場合、後の状態表現 P 2 を、状態 S 1 から状態 S 2 への変化を示す「P 1 P 2」といった変化表現に置き換えることとする。図 5 は、図 3 に示す処理対象文書 D に対して、このような文字列置換処理を実行した結果得られる文書の内容を示している。さらに、処理対象文書 D 内において、状態表現 P 2 の後に別の状態 S 3 を表す状態表現 P 3 が含まれる場合、変化情報出力部 2 3 は、状態表現 P 3 を「P 2 P 3」のような変化表現に置換する。こうして、処理対象文書 D 内において評価対象のある時点における状態を示す状態表現が、前の状態から当該状態への変化を示す変化表現に置き換えられる。

#### 【 0 0 3 2 】

さらに、変化情報出力部 2 3 は、単に状態 S 1 から状態 S 2 への変化をそのまま示すのではなく、状態表現抽出部 2 2 が抽出した各状態表現の内容に基づいて、評価対象の変化の極性を評価し、この評価結果を示す情報（極性表現）を変化情報として出力してもよい。ここで、極性表現は、変化の方向性を表す表現であり、具体例として、状態が定量的に表される場合にはその変化が数値の上昇か下降かを表す情報であってよい。また、変化情報出力部 2 3 は、変化の極性として、その変化が「改善」、「悪化」、又は「無効」のいずれかを評価することとしてもよい。この場合、「改善」は一般的な評価において好ましい変化に、「悪化」は好ましくない変化に、「無効」は変化がないか又は意味のない変化であることに、それぞれ対応する。

#### 【 0 0 3 3 】

具体的に、状態が数値で表される場合、変化情報出力部 2 3 は、各状態表現の数字列を数値に変換し、後の状態を表す数値から前の状態を表す数値を減じることによって、数値の変化量を算出する。そして、算出された変化量の数値が正の値であれば極性を「上昇」と評価し、負の値であれば極性を「下降」と評価する。その後、変化情報出力部 2 3 は、後の状態表現を、この評価結果を示す極性表現（例えば「上昇」であれば「+」、「下降」であれば「-」）で置き換える。これにより、処理対象文書 D の内容は、変化の態様により容易に把握される表現に変換される。図 6 は、図 5 に示す文書中において「状態表現 状態表現」というパターンで表される変化表現に対して、このような極性の評価、及び文字列置換処理が行われた結果得られる文書の内容を示している。なお、変化情報出力部 2 3 は、以上説明したような後の状態表現から変化表現、変化表現から極性表現という 2 段階の置換処理を行うのではなく、状態表現抽出部 2 2 が抽出した状態表現に基づいて、後の状態表現を直接極性表現に置き換えてもよい。

#### 【 0 0 3 4 】

また、評価対象の状態が状態名辞書に含まれる複数の候補文字列のいずれかによって表現される場合、変化情報出力部 2 3 は、状態名辞書に含まれる各状態の評価情報を用いて、極性を判定することとしてもよい。この場合、状態名辞書には、状態表現の候補文字列のそれぞれに関連づけて、当該候補文字列により表される状態に対する評価を示す評価情報が含まれることとする。例えば前述したハードディスクの音の例では、「カッコンカッ

コン」という音は異常、「コリコリ」は正常など、各状態が正常か異常かを示す評価情報が候補文字列とともに状態名辞書に含まれることとする。変化情報出力部23は、状態表現抽出部22が抽出した各状態表現により表される状態が正常か異常かを、この状態名辞書に含まれる評価情報を取得することによって判定する。そして、変化表現により示される状態の変化が「正常」「異常」の場合には、変化表現を「悪化」という極性表現に置換する。また、「異常」「正常」の変化であれば「改善」という極性表現に、「正常」「正常」又は「異常」「異常」であれば「無効」という極性表現に、それぞれ変化表現を置換する。図7は、図4に例示される処理対象文書Dに対して、以上説明したような後の状態表現から変化表現、変化表現から極性表現への置換処理が実行された結果を示している。

10

#### 【0035】

ここで、状態名辞書において状態表現の候補文字列に関連づけられる評価情報は、処理対象文書Dと同種の文書群の内容を解析することによって収集されることとしてもよい。具体例として、処理対象文書Dが修理記録文書であり、これと同種の文書からなる文書群の中に「ハードディスクからカッコンカッコンという音が鳴り始めたので交換です。」という文が含まれていたとする。また「交換」という文字列が「異常」という評価と結びつくことが予め定義されているとする。この場合、本実施形態に係る情報処理装置1は、上述した「カッコンカッコン」という状態表現と「交換」という文字列が含まれる文を文書群の中から発見したことに応じて、「カッコンカッコン」という状態表現の候補文字列に「異常」という評価情報を関連づけた状態名辞書を生成する。

20

#### 【0036】

作用表現抽出部24は、処理対象文書Dの中から、評価対象に影響する作用を表す作用表現を抽出する。ここで作用とは、評価対象に影響を及ぼす可能性のある人の行為（診療や修理などの対処）や現象などである。作用表現抽出部24は、状態表現の場合と同様に、記憶部12に記憶された作用名辞書に含まれる作用表現の候補文字列を処理対象文書D内から検索するなどの方法で、作用表現を抽出する。

#### 【0037】

結果情報出力部25は、作用表現抽出部24が抽出した作用表現と、変化情報出力部23が出力する変化情報と、に基づいて、当該作用表現により表される作用によって評価対象に生じる変化（すなわち、作用による評価対象への影響や効果）に関する結果情報を出力する。具体的に、例えば結果情報出力部25は、予め定められた条件を満足する位置関係で処理対象文書D内に含まれる作用表現、評価対象表現、及び変化表現の3つの要素からなる組情報を、結果情報として出力する。このような処理は、非特許文献1などに開示される対象、属性、評価値の組を特定、収集する手法と同様の手法によって実現されてよい。図4に例示される処理対象文書Dに対してこのような処理が実行されることにより、「HDDの交換」という作用（対処）によって「機器から聞こえる音」という評価対象が「改善」されたことを示す結果情報が、例えば記憶部12内のデータベースに出力される。

30

#### 【0038】

あるいは、結果情報出力部25は、作用表現抽出部24が抽出する作用表現と、変化情報出力部23が出力する極性表現と、を結びつけて、当該作用表現によって表される作用による効果を示す結果情報として出力してもよい。

40

#### 【0039】

なお、結果情報出力部25は、変化情報出力部23が出力する変化情報によらずに、作用表現抽出部24が抽出する作用表現と、処理対象文書Dと同種の文書群から収集された情報と、に基づいて、抽出された作用表現によって表される作用の効果を推定する処理を実行してもよい。このような推定処理の具体例について、以下、説明する。

#### 【0040】

複数の作用表現が連続して抽出された場合、結果情報出力部25は、これらの作用表現と同様の作用表現が同様の順序で含まれ、かつ、これらの作用表現の間に「無効」を表す

50

表現が含まれる文書を、文書群の中から検索する。具体例として、処理対象文書 D に下記のような記載が存在するとする。

「 0 7 / 0 3 / 0 1 抗がん剤 B を投与

0 7 / 0 4 / 0 1 抗がん剤 C を投与」

また、ここでは「抗がん剤 B」及び「抗がん剤 C」が作用表現として作用名辞書に含まれているものとする。

【 0 0 4 1 】

この場合、結果情報出力部 2 5 は、「抗がん剤 B」、「無効」を表す表現、及び「抗がん剤 C」がこの順序で含まれる文書を、検索対象の文書群の中から検索する。ここで、「無効」を表す表現は、予め辞書に記述された固定的な表現であってもよい。また、前述した状態名辞書において、「異常」を示す評価情報が関連づけられた状態表現や、過去の解析結果において「悪化」を示す極性表現に置換された状態表現であってもよい。一例として、文書群に含まれる文書の中に、

「 0 5 / 0 8 / 0 2 抗がん剤 B を投与

0 5 / 0 9 / 0 2 C E A が上昇したため、抗がん剤 C に変更」

といった記載が存在したとする。ここで、C E A (Carcinoembryonic Antigen、がん胎児性抗原) の上昇は抗がん剤の効果が十分でないことを示しており、前述した「無効」を表す表現として連続する「C E A」及び「上昇」の語が定義されていることとする。この文書からは、抗がん剤 B の投与後に抗がん剤 C を投与するのは、抗がん剤 B の効果がなかった場合であることが推定される。そこで、結果情報出力部 2 5 は、このような「抗がん剤 B」、「C E A」+「上昇」、「抗がん剤 C」が連続する記載を含む文書の検索結果に基づいて、処理対象文書 D における「抗がん剤 B」という作用(対処)による変化(効果)を示す結果情報として、「無効」であることを示す情報を出力する。

【 0 0 4 2 】

なお、以上の説明においては、処理対象文書 D に対する処理を実行する際に、結果情報出力部 2 5 が連続する作用表現を文書群から検索することとしたが、これに限らず、例えば予め文書群から複数の作用表現と、その間の「無効」を表す表現と、を検索し、検索結果を辞書として記憶しておくこととしてもよい。この場合、結果情報出力部 2 5 は、連続する複数の作用表現を処理対象文書 D 内から抽出した際には、当該作用表現の並びがこの辞書内で定義されている場合に、前の作用表現によって表される作用の効果が無効であると判定する。

【符号の説明】

【 0 0 4 3 】

1 情報処理装置、1 1 制御部、1 2 記憶部、1 3 操作部、1 4 表示部、2 1 文書取得部、2 2 状態表現抽出部、2 3 変化情報出力部、2 4 作用表現抽出部、2 5 結果情報出力部。

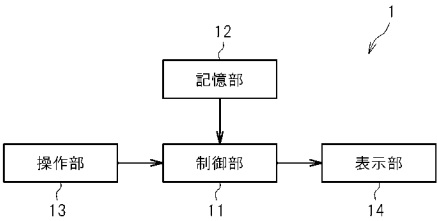
10

20

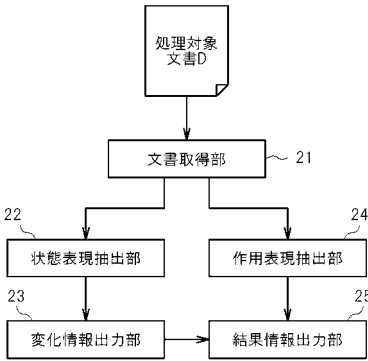
30



【図 1】



【図 2】



【図 3】

03/09/05	抗がん剤A(80mg/m2=120mg/body)
03/10/02	CA125 10に低下
03/10/03	抗がん剤A(80mg/m2=120mg/body)
03/10/30	抗がん剤A(80mg/m2=120mg/body)
03/11/11	CA125 5

【図 4】

機器から聞こえる音:	カッコンカッコン
HDDを交換	
機器から聞こえる音:	コリコリ

【図 7】

機器から聞こえる音:	カッコンカッコン
HDDを交換	
機器から聞こえる音:	改善

【図 5】

03/09/05	抗がん剤A(80mg/m2=120mg/body)
03/10/02	CA125 10に低下
03/10/03	抗がん剤A(80mg/m2=120mg/body)
03/10/30	抗がん剤A(80mg/m2=120mg/body)
03/11/11	CA125 10 → 5

【図 6】

03/09/05	抗がん剤A(80mg/m2=120mg/body)
03/10/02	CA125 10に低下
03/10/03	抗がん剤A(80mg/m2=120mg/body)
03/10/30	抗がん剤A(80mg/m2=120mg/body)
03/11/11	CA125 —

---

フロントページの続き

(72)発明者 増市 博

神奈川県足柄上郡中井町境430 グリーンテクなかい 富士ゼロックス株式会社内

審査官 鈴木 和樹

(56)参考文献 特開2006-146567(JP,A)

特開2000-315207(JP,A)

特開2007-183903(JP,A)

国際公開第2007/043322(WO,A1)

斉藤公一、外5名、数値情報をキーとした新聞記事からの情報抽出、情報処理学会研究報告(98-NL-125)、日本、社団法人情報処理学会、1998年5月29日、第98巻、第48号、p.63-70

阿部秀尚、外2名、テキストマイニングによる退院時サマリーからの時系列関係知識抽出、医療情報学、日本、有限責任中間法人日本医療情報学会 株式会社篠原出版新社、2007年5月30日、第27巻、第1号、p.33-38

Eiji ARAMAKI、外5名、TEXT2TABLE: Medical Text Summarization System based on Named Entity Recognition and Modality Identification, Proceedings of the Workshop on BioNLP [online], Association for Computational Linguistics, 2009年6月5日、p.185-192, URL, <http://aclweb.org/anthology//W/W09/W09-1324.pdf>

(58)調査した分野(Int.Cl., DB名)

G06F 19/00

G06F 17/30

G06F 17/20-28