

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2006-244102
(P2006-244102A)

(43) 公開日 平成18年9月14日(2006.9.14)

(51) Int. Cl. F I テーマコード(参考)
G06F 17/30 (2006.01) G06F 17/30 180A 5B075
 G06F 17/30 180B

審査請求 未請求 請求項の数 6 O L (全 13 頁)

<p>(21) 出願番号 特願2005-58390 (P2005-58390)</p> <p>(22) 出願日 平成17年3月3日(2005.3.3)</p> <p>(特許庁注：以下のものは登録商標)</p> <p>1. Linux</p>	<p>(71) 出願人 504171134 国立大学法人 筑波大学 茨城県つくば市天王台一丁目1番1</p> <p>(74) 代理人 100089635 弁理士 清水 守</p> <p>(72) 発明者 藤井 敦 茨城県つくば市天王台一丁目1番1 国立 大学法人筑波大学内</p> <p>(72) 発明者 石川 徹也 茨城県つくば市天王台一丁目1番1 国立 大学法人筑波大学内</p> <p>(72) 発明者 三原 英理 茨城県つくば市天王台一丁目1番1 国立 大学法人筑波大学内</p> <p>Fターム(参考) 5B075 NR02 PP30 QP01 UU40</p>
--	---

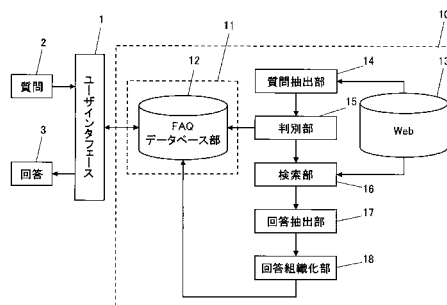
(54) 【発明の名称】 質問応答システム

(57) 【要約】

【課題】 行動を問う質問に対して、領域を限定することなく回答することができる質問応答システムを提供する。

【解決手段】 よくある質問とそれに対する回答を事前に記憶するFAQデータベース部12と、ウェブ13に接続され、ウェブ13上に頻出する質問を抽出する質問抽出部14と、この質問抽出部14に接続され、前記よくある質問であるか否かの判別部15と、前記判別部15に接続されるとともに、ウェブ13に接続され、前記判別部15でよくある質問でないと判定された場合、検索エンジンを用いて質問に関連するページを網羅的に検索する検索部16と、この検索部16に接続され、検索されたページから行動表現を含む段落を抽出する回答抽出部17と、この回答抽出部17に接続され、含まれる行動表現によって段落を分類する回答組織化部18とを備え、前記回答組織化部18で分類された回答と質問を対にして、前記FAQデータベース部11に蓄積する回答部19とを具備する。

【選択図】 図1



【特許請求の範囲】**【請求項 1】**

(a) 質問を入力し、回答が出力されるユーザインターフェース装置と、
(b) 該ユーザインターフェース装置に接続されるとともに、よくある質問とそれに対する回答を事前に記憶する F A Q データベース部と、
(c) ウェブに接続され、ウェブ上に頻出する質問を抽出する質問抽出部と、
(d) 該質問抽出部に接続され、前記よくある質問であるか否かを判別する判別部と、
(e) 前記判別部に接続されるとともに、ウェブに接続され、前記判別部でよくある質問でないとは判定された場合、検索エンジンを用いて質問に関連するページを網羅的に検索する検索部と、
(f) 該検索部に接続され、検索されたページから行動表現を含む段落を抽出する回答抽出部と、
(g) 該回答抽出部に接続され、含まれる行動表現によって段落を分類する回答組織化部と、
(h) 前記回答組織化部で分類された回答と質問を対にして、前記 F A Q データベース部に蓄積する回答部とを具備することを特徴とする質問応答システム。

10

【請求項 2】

請求項 1 記載の質問応答システムにおいて、前記回答抽出部は検索されたページ群から回答となる段落を抽出し、スコア付けを行うことを特徴とする質問応答システム。

【請求項 3】

請求項 2 記載の質問応答システムにおいて、前記段落の抽出には、文字数に関する閾値を併用し、120 文字以上 200 文字以下とすることを特徴とする質問応答システム。

20

【請求項 4】

請求項 2 記載の質問応答システムにおいて、前記スコア付けについては、検索されたページに出現する個々の行動表現にスコアを付け、段落内に含まれる行動表現のスコアを総和することを特徴とする質問応答システム。

【請求項 5】

請求項 4 記載の質問応答システムにおいて、前記行動表現のスコアは、回答として適切な行動表現を選択するため、

- (a) 係り受けの距離が近い
- (b) 推奨表現と禁止表現を伴っている
- (c) 抽出元ページの検索における順位が高い
- (d) 質問に含まれる行動表現との距離が近い
- (e) 質問者(ユーザ)がすべき行動である

30

上記 5 つの基準を用いることを特徴とする質問応答システム。

【請求項 6】

請求項 1 記載の質問応答システムにおいて、前記回答組織化部による回答組織化によって、行動の順序関係や依存関係を抽出するための手がかりを得ることを特徴とする質問応答システム。

【発明の詳細な説明】

40

【技術分野】**【0001】**

本発明は、何らかの困った状況にあるユーザからの質問に対して、解決手段となる行動を回答する質問応答システムに関するものである。

【背景技術】**【0002】**

インターネットの普及により、Web(ウェブ)[World Wide Webの略]上には多種多様かつ膨大な情報が存在する。そのため、日常的な疑問に対し、検索エンジンを用いてウェブ上の情報から回答を探す機会が増えている。しかし、既存の検索エンジンを使うためには、質問の内容をキーワードで的確に表現する必要がある。さらに、検

50

索された膨大なページから回答となる語句や文章をユーザ自身が探さなければならない。

【0003】

ところで、自然言語による質問文に対して回答するために、質問応答システムが研究されている。既存の研究は、限定された領域に関する知識や規則によって推論を行うもの（推論型システム）と、組織化されていない文書集合から回答を抽出するもの（抽出型システム）との2種類に大別される。近年は後者の抽出型システムが主に研究されている。

【0004】

以下、上記した従来技術について詳細に説明する。

【0005】

人工知能分野における質問応答の研究は1960年代に遡る。具体的には、エキスパート・システムや対話システムであり、推論によって回答する「推論型システム」である。これらは、システムとユーザとの対話を通して質問・回答を洗練させたり、推論によって複数の事実から回答を導き出したりするという特長がある。しかし、領域知識や推論規則を手で構築する必要があるため、回答できる分野が限定され、拡張性に乏しいといった問題があった。

10

【0006】

一方、自然言語処理や情報検索分野では、1990年代後半からTREC (Text Retrieval Conference: <http://trec.nist.gov/>) やNTCIR (情報検索システム評価用テストコレクション構築プロジェクト) で質問応答の研究が盛んになった。人工知能分野における推論型システムと異なり、ここでは、「抽出型システム」が主な研究対象である。組織化されていない文書集合を情報源として用いるため、領域知識や推論規則の構築が不要な点で推論型システムとは異なる。

20

【0007】

このようなTRECやNTCIRにおける質問応答システムの回答対象は客観的事実 (factoid) である。ユーザが入力した、who、when、where、whatを中心とするWH疑問文に対し、名詞で回答する。客観的事実を回答とするシステムとしては、goolabo日本語自然文検索 (<http://labs.nttrd.com/>) やMURAX (下記非特許文献1参照) がある。MURAXは入力された質問文を解析することで、何についての質問で、回答はどのようなものかという仮説を立てる。その仮説に基づいてクエリ構築や検索を行い、文書集合から該当箇所を探して回答する。

30

【0008】

また、下記非特許文献2では、行動を問う「how型質問」への回答を目的とし、知識獲得手法が提案されている。linuxユーザのメーリングリストに投稿された質問メール・回答メールの本文から、重要文を抽出することで、方法や対処法を問う質問に回答するための知識を得ている。しかし、本手法では対象となる領域が限定されている。

40

【非特許文献1】Julian Kupiec. MURAX: Finding and Organizing Answers from Text Search. In T. Strzalkowsky, editor, Natural Language Information Retrieval, pp. 311-332. Kluwer Academic Publishers, 1999.

【非特許文献2】渡辺靖彦, 園和也, 岡田至弘. 「メーリングリストを利用した質問応答システムのための知識獲得」, 情報処理学会研究報告, 2004-NL-162, pp. 131-138, 2004.

【非特許文献3】徳永健伸, 岩山真, 乾健太郎, 田中穂積. 「日本語語順の推定モデルとその応用」, 情報処理学会研究報告, 91-NL-81, pp. 9-16, 1991.

【発明の開示】

【発明が解決しようとする課題】

【0009】

しかしながら、上記抽出型システムは、who、when、where、whatを中心とするWH疑問文に対し、名詞で回答するものが主流であるため、必ずしもユーザが知

50

りたい回答（特に行動表現）と一致するとは限らない。また、行動を問うhow型質問に対応した研究は、コンピュータの利用法などに対象領域が限定されている。

【0010】

本発明は、上記状況に鑑みて、抽出型の質問応答システムをウェブに適用し、行動を問うhow型質問に領域を限定することなく回答することができる質問応答システムを提供する。つまり、何らかの困った状況にあるユーザからの質問に対して、取るべき行動を示唆するヘルプデスク指向の質問応答システムを提供する。

【課題を解決するための手段】

【0011】

本発明は、上記目的を達成するために、

〔1〕質問応答システムにおいて、質問を入力し、回答が出力されるユーザインターフェイス装置と、このユーザインターフェイス装置に接続されるとともに、よくある質問とそれに対する回答を事前に記憶するFAQデータベース部と、ウェブに接続され、ウェブ上に頻出する質問を抽出する質問抽出部と、この質問抽出部に接続され、前記よくある質問であるか否かを判別する判別部と、前記判別部に接続されるとともに、ウェブに接続され、前記判別部でよくある質問でないとは判定された場合、検索エンジンを用いて質問に関連するページを網羅的に検索する検索部と、この検索部に接続され、検索されたページから行動表現を含む段落を抽出する回答抽出部と、この回答抽出部に接続され、含まれる行動表現によって段落を分類する回答組織化部と、前記回答組織化部で分類された回答と質問を対にして、前記FAQデータベース部に蓄積する回答部とを具備することを特徴とする。

10

20

【0012】

〔2〕上記〔1〕記載の質問応答システムにおいて、前記回答抽出部は検索されたページ群から回答となる段落を抽出し、スコア付けを行うことを特徴とする。

【0013】

〔3〕上記〔2〕記載の質問応答システムにおいて、前記段落の抽出には、文字数に関する閾値を併用し、120文字以上200文字以下とすることを特徴とする。

【0014】

〔4〕上記〔2〕記載の質問応答システムにおいて、前記スコア付けについては、検索されたページに出現する個々の行動表現にスコアを付け、段落内に含まれる行動表現のスコアを総和することを特徴とする。

30

【0015】

〔5〕上記〔4〕記載の質問応答システムにおいて、前記行動表現のスコアは、回答として適切な行動表現を選択するため、(a)係り受けの距離が近い、(b)推奨表現と禁止表現を伴っている、(c)抽出元ページの検索における順位が高い、(d)質問に含まれる行動表現との距離が近い、(e)質問者(ユーザ)がすべき行動である、上記5つの基準を用いることを特徴とする。

【0016】

〔6〕上記〔1〕記載の質問応答システムにおいて、前記回答組織化部による回答組織化によって、行動の順序関係や依存関係を抽出するための手がかりを得ることを特徴とする。

40

【発明の効果】

【0017】

本発明によれば、以下のような効果を奏することができる。

【0018】

(a)領域を限定することなく、how型質問に回答することができる。つまり、何らかの困った状況にあるユーザに対して、取るべき行動を示唆するヘルプデスク指向の質問応答システムを提供する。

【0019】

(b)よくある質問とそれに対する回答を事前に記憶するFAQデータベース部を有す

50

るため、よくある質問に対しては迅速なる回答を提示することができる。

【発明を実施するための最良の形態】

【0020】

本発明の質問応答システムは、質問を入力し、回答が出力されるユーザインタフェース装置と、このユーザインタフェース装置に接続されるとともに、よくある質問とそれに対する回答を事前に記憶するFAQデータベース部と、ウェブに接続され、ウェブ上に頻出する質問を抽出する質問抽出部と、この質問抽出部に接続され、前記よくある質問であるか否かを判別する判別部と、前記判別部に接続されるとともに、ウェブに接続され、前記判別部でよくある質問でないとは判定された場合、検索エンジンを用いて質問に関連するページを網羅的に検索する検索部と、この検索部に接続され、検索されたページから行動表現を含む段落を抽出する回答抽出部と、この回答抽出部に接続され、含まれる行動表現によって段落を分類する回答組織化部とを備え、前記回答組織化部で分類された回答と質問を対にして、前記FAQデータベース部に蓄積する回答部とを具備することを特徴とする。

10

【実施例】

【0021】

以下、本発明の実施の形態について詳細に説明する。

【0022】

本発明のシステムは、入力された質問文に関連するウェブページを検索し、抽出した回答を分類して提示する。対象とするのは、what-if、what-to-do、how-to-do型の行動を問う質問である。これらの質問に回答するため、名詞句と動詞からなる「行動表現」に着目する。例えば、「蜂に刺されたらどうすればよい？」という質問には「針を抜く」、「流水で洗う」、「軟膏を塗る」などの行動表現と、その行動表現を含む段落を回答する。

20

【0023】

図1は本発明の質問応答システムの構成図である。

【0024】

この図において、1はユーザインタフェース装置、2は質問（質問文：質問情報）、3は回答（回答文：回答情報）、10は事前処理部、11は回答部、12はFAQデータベース部、13はウェブ、14は質問抽出部、15はよくある質問か否かを判別する判別部、16は検索部、17は回答抽出部、18は回答組織化部である。

30

【0025】

本システムでは、検索されたページを解析して回答を抽出することになるため、通常のウェブ検索よりも応答に時間がかかる。そこで、本発明では、聞かれそうな質問とそれに対する回答を事前にFAQ(Frequently Asked Questions: 頻りに尋ねられる質問)としてデータベース化することで、迅速な回答を可能にする。

【0026】

まず、FAQデータベースを構築する事前処理(オフライン処理)について説明する。ここでは、質問抽出部14による「質問抽出」、検索部16による「検索」、回答抽出部17による「回答抽出」、回答組織化部18による「回答組織化」の処理を段階的に行う。

40

【0027】

まず、質問抽出部14は、ウェブ13上に頻出する質問を抽出する。次に、検索部16は、既存の検索エンジンを用いて、質問抽出部14により抽出された質問に関連するページを網羅的に検索する。次に、回答抽出部17は、検索部16によって検索されたページから行動表現を含む段落を抽出する。次に、回答組織化部18は、含まれる行動表現によってこの抽出された段落を分類し、分類された回答と質問を対にしてFAQデータベース部12に蓄積する。

【0028】

次に、ユーザがシステムを利用する場合のオンライン処理について説明する。ユーザが

50

ユーザインタフェース装置 1 に質問 2 を入力すると、FAQ データベース部 1 2 内に同じ質問がないか判別部 1 5 で検索する。同じ質問があれば、それに対応する回答を FAQ データベース部 1 2 より取り出し提示する。同じ質問がなければ、事前処理と同様に、検索部 1 6 で検索、回答抽出部 1 7 で回答抽出、回答組織化部 1 8 で回答組織化を行い、動的に作成された回答を回答部 1 1 に生成して提示する。

【0029】

図 2 はその回答を示す画面例である。

【0030】

その回答は、図 2 に示すような画面で提示する。これは「蜂に刺されたらどうすればいい？」という質問に対する回答画面例である。「病院へ行く」という見出し表現は組織化後のグループを代表する行動表現で、その下の「針を抜く」、「診察を受ける」などは、このグループに含まれる他の行動表現である。それらの行動表現を含む段落と、抽出元のページへのリンクを併記し、詳細情報を取得しやすくしている。

10

【0031】

すなわち、本システムの「回答」とは、「病院へ行く」のような行動表現の句であり、また同時に、行動表現を含む段落のような文章でもある。ユーザは利用環境などによって、行動表現のグループ一覧やグループ内の段落一覧など、表示内容を選択することができる。

【0032】

以下、質問抽出、検索、回答抽出、回答組織化の各処理について説明する。

20

【0033】

(1) 質問抽出

ウェブ 1 3 上には、FAQ サイトに用意された質問や、個人の日記サイトやウェブログに書かれた素朴な疑問まで、様々な質問表現が潜在するため、これらを収集する。

【0034】

「 が したら、どうする?」、「 を したら?」のような表現を単語や品詞で規則化し、該当するフレーズや文を抽出する。具体的には『名詞の後に助詞(が・に・を)、動詞(～したら)、と続き、さらに疑問符や「(、) どうする」が続く』という規制を満たす表現を抽出する。

【0035】

このとき、形態素も用いるため、表層表現に一致しない語句も抽出することができる。他方において、質問以外の表現(ノイズ)が検出される可能性がある。しかし、仮に FAQ にノイズが含まれても、ユーザが当該ノイズを含む同様の質問をしなければ実害はない。そこで、質問抽出では精度(precision)よりも再現率(recall)を重視する。

30

【0036】

抽出元データには NTCIR (<http://research.nii.ac.jp/ntcir/index-ja.html>) のウェブコレクションを使用した。これは J P ドメインから 11,038,720 のページ(約 100 GB)を収集したコレクションである。その結果、6961 のフレーズが抽出された。

40

【0037】

(2) 検索

ウェブ 1 3 検索には検索サイト Google を利用し、入力された質問文をそのままクエリとして用いる。しかし、実験の規模を拡張するためには、質問応答に適した検索エンジンを独自に開発する必要がある。

【0038】

(3) 回答抽出

回答抽出部 1 7 は、検索部 1 6 で検索されたページ群から行動表現を含む回答となる段落を抽出し、スコア付けを行う。実際のウェブ 1 3 上には PDF ファイルなども存在するが、本発明では本文のレイアウト解析を行う必要があるため HTML ファイルのみを対象

50

とする。

【0039】

以下、詳しい抽出手順を説明する。

【0040】

まず、検索された各ページから、一定の長さを持つ段落を抽出する。これは、単語や文の単位で抽出した回答では、行動を説明するための情報が少なく、逆にページ単位では不要な情報が含まれやすいためである。段落の判別には、HTMLタグの中でも<P>と
を用いた。

【0041】

タグの使い方はページ作成者によって様々であり、タグを判別に用いるだけでは段落の長さにはばらつきが出てしまう。例えば、段落を示す<P>を改行代わりに用いているページでは、各文が段落となってしまう情報が不足する。逆に、非常に長い段落を読むことはユーザの負担が大きい。そこで、文字数に関する閾値を併用し、120文字以上200文字以下を段落とした。

【0042】

次に、回答としての適切さによって段落にスコア付けを行う。本発明のシステムでは行動表現が回答の最小単位となるため、行動を表す表現がより多く含まれている段落が回答として重要である。そのため、行動表現を形式化し、スコア付けに用いた。

【0043】

段落のスコア付けでは、まず、検索されたページに出現する個々の行動表現にスコアを付け、段落内に含まれる行動表現の各スコアを総和する。

【0044】

ここで、行動表現について説明する。

日本語係り受け解析器Cabocha (<http://chasen.org/~taku/software/cabocha/>)

【0045】

を用いて、検索されたページのテキスト部分について係り受け解析を行う。そこから「名詞+助詞+動詞」という係り受け構造を抽出し、行動表現の最小単位とする。例えば「薬を患部に塗る。」という文の場合は「薬+を+塗る」と「患部+に+塗る」の2フレーズが得られる。ここで「薬を塗る」ならば「薬：塗る」のように助詞を省き、これを行動表現とする。

【0046】

助詞を省くのは「蜂に人が刺される」と「蜂が人を刺す」のように、格交代により助詞が変化しても「誰が何をどうする」という行動内容は変わらず、同じ行動表現とみなすためである。

【0047】

本手法は、「薬を患部に塗る」のように複数の名詞句(名詞+助詞)が動詞に係った表現を行動表現として扱うこともできる。しかし、名詞句が複数あった場合「薬を塗る」と「薬を患部に塗る」が別の行動表現とみなされるという問題が生じる。そこで、実験では名詞句をひとつに限定した。

【0048】

行動表現のスコア付けでは、回答として適切な行動表現を選択するため、次の5つの基準を用いる。

【0049】

(a) 係り受けの距離が近い。

【0050】

係り受けの距離とは、名詞句と動詞の間にある形態素数のことである。日本語の場合、文末に近い要素ほど動詞の意味を限定する性質がある(上記非特許文献3参照)。また、距離が近いほど係り受け解析の誤りが少ないという利点もある。

【0051】

10

20

30

40

50

(b) 「～すること」、「～しましょう」といった推奨表現と、「～してはいけない」といった禁止表現を伴っている。

【0052】

推奨表現は問題解決に有効な対処法を述べる時に用いられる。また、禁止表現は行ってはならない対処法を述べる時に用いられるため、推奨表現と同等に有用である。例えば、「蜂に刺されたら？」の質問例では「薬を塗りましょう」、「病院へ行くこと」などが推奨表現であり、「アンモニアをかけてはいけない」、「毒袋をつぶさないでください」などが禁止表現である。このような表現を正規表現によって照合し、検出する。

【0053】

(c) 抽出元ページの検索における順位が高い。

10

【0054】

抽出元ページとは、行動表現が抽出されたページである。現在、検索にはGoogleを用いているため、抽出元ページの順位とはGoogleの検索順位を意味する。GoogleはPageRankにより、他ページから多くリンクされているページは高順位とする。そのため、信頼性を部分的に取り入れることができる。なお、このスコアは以下の式で計算する。

【0055】

(検索ページ数 - 順位) / 検索ページ数

(d) 段落内で質問に含まれる行動表現との距離が近い。

【0056】

これは行動表現中の動詞と質問中の動詞の間にある形態素数から判断する。行動表現中の動詞が質問中の動詞に近いほど質問との関連性が強く、対処法として重要な行動表現が出現しやすい。

20

【0057】

(e) 質問者(ユーザ)がすべき行動である。

【0058】

質問者が動作主となる行動を取り出すため、格助詞「ガ」(以下、ガ格)によって主語が明示されている行動表現はスコアを下げる。例えば「蜂が巣を守る。」という文であれば動作主は「蜂」である。この文に出現する行動表現「蜂が守る」、「巣を守る」はともに蜂が行う行動表現であり、質問者への回答として適切でない。

30

【0059】

各行動表現のスコアは、上記(a)～(e)の基準を下記式(1)によって結合し、計算する。

【0060】

【数1】

$$s(x) = \sum \left(\log \frac{1}{a_x} + b_x + c_x + \frac{1}{d_x} \right) \cdot e_x \quad \dots (1)$$

ここで、xは段落に出現する行動表現である。axは係り受けの距離、dxは質問との距離である。bxは推奨・禁止表現の有無によって1か0を取る。cxは上記(c)で計算される値である。exはガ格の有無によって1か0を取る。

40

【0061】

また、段落Pのスコアは下記式(2)により、Pに含まれる行動表現のスコアを総和して求める。

【0062】

【数2】

$$s(P) = \sum_{x \in P} s(x) \quad \dots (2)$$

(4) 回答組織化

回答組織化部18は、上記の回答抽出の手法で抽出された複数の段落を分類する。こ

50

では、同じ行動表現を含む段落を同じグループにまとめる。ただし、スコアの低い行動表現はノイズになるので、事前に削除する。現在は経験的に、スコアが2以上の行動表現を使用して分類を行っている。ユーザに回答を提示する際に、グループごとに最も高いスコアの段落と行動表現を提示することで、類似した複数の段落を読む手間を省くことができ、どのような回答グループがあるのかを概観しやすくなる。

【0063】

また、同一グループに含まれる行動表現は依存関係にあり、逆に、同じグループに属さない行動表現は排他的な関係にある。このような行動表現同士の関係は、行動の順序関係や依存関係を抽出するための手がかりになる。

【0064】

(評価実験)

以下、本発明の質問応答システムの評価実験を行ったので説明する。

【0065】

how型の質問応答システムはほとんど研究されていないため、評価方法が確立されていない。そこで、人手でテストコレクションを作成し、本発明のシステムで最も重要な機能といえる回答抽出の機能、中でも行動表現のスコア付けに焦点を絞って評価を行った。

【0066】

これは、本発明のシステムが、高スコアの行動表現を元に段落をグループ化し、グループ内で最高スコアの行動表現を代表(すなわち、図2における「病院へ行く」としてユーザに提示するため、回答として適切な行動表現のスコアが他よりも高くなることが重要となるからである。そこで、行動表現のスコア付けについて、先に提案した(a)~(e)の5つの基準が有効であるかという観点で評価を行った。

【0067】

質問は、質問抽出で抽出されたフレーズのうち、質問頻度が高い以下の13フレーズを使用した。

【0068】

(1)パスワードを忘れたら、(2)事故にあったら、(3)迷子になったら、(4)急病にかかったら、(5)盗難にあったら、(6)蜂に刺されたら、(7)怪我をしたら、(8)ニキビができたら、(9)発作がおきたら、(10)パスポートを紛失したら、(11)セクハラを受けたら、(12)やけどをしたら、(13)火事になったら

これらの質問に対して、抽出された行動表現の適否を人手で判定した。さらに、適切な表現については、禁止表現かどうかも判定し、正解として扱った。

【0069】

スコアによって順位付けた行動表現について、以下に示す精度と再現率を用い、F値を計算した。

【0070】

ここで、F値(F-measure)とは、再現率(recall)と精度(precision)の調和平均である。再現率と精度の両方を考慮した評価尺度であり、再現率と精度が大きいほど、F値も大きくなる。

【0071】

今、ある質問に対する全ての可能な行動表現に対して、「システムが回答として出力した行動表現かどうか?」、「その質問に対する正しい行動表現かどうか?」という観点で分類すると、以下のような分割表(contingency table)を作ることができる。

【0072】

10

20

30

40

【表 1】

		システムが回答として出力したか？	
		Y e s	N o
正しい行動 表現か？	Y e s	A	B
	N o	C	D

ここで、A ~ D は該当するセルの行動表現件数であり、 $A + B + C + D$ は全行動表現数になる。また、A ~ C のセルは以下のケースに対応する。

10

【0073】

A : 成功、B : 失敗 (抽出漏れ)、C : 失敗 (抽出誤り)

上記の分割表に基づいて、再現率と精度はそれぞれ以下の式で計算される。

【0074】

再現率 = $A / (A + B)$

精度 = $A / (A + C)$

F 値は再現率 (R) と精度 (P) の調和平均なので、以下で計算される。

まず、R と P の逆数に対して算術平均を計算する。

【0075】

【数 3】

$$\frac{1}{2} \times \left(\frac{1}{R} + \frac{1}{P} \right) = \frac{1}{2} \times \left(\frac{A+B}{A} + \frac{A+C}{A} \right) = \frac{2A+B+C}{2A} \quad \dots (3)$$

20

さらに、逆数を取ると、 $F \text{ 値} = 2A / (2A + B + C)$

すなわち、F 値は、「成功と失敗 (抽出漏れ、抽出誤り)」に対する「成功」の比率を表す。

【0076】

まとめると、

精度 = 出力した正解数 / 出力した全行動表現数

再現率 = 出力した正解数 / 全行動表現中の正解数

F 値 = $2 \times \text{再現率} \times \text{精度} / (\text{再現率} + \text{精度})$

(結果と考察)

30

回答数による F 値の推移を調べるため、行動表現を段階的に増やしていき、上位 200 件までの行動表現について F 値を計算した。評価に用いた全質問について、行動表現数を変化させた各点における F 値を計算し、その平均値を用いて比較評価した。

【0077】

結果を図 3 に示す。グラフの縦軸は平均 F 値、横軸は行動表現数を表す。それぞれの系列は以下の条件で評価した値を用い、全質問の平均を取っている。

【0078】

a l l : 基準 (a) ~ (e) をすべて用いたスコアによる順位付け

P a : 基準 (a) 「係り受けの距離」を使用したスコアによる順位付け

P b : 基準 (b) 「推奨・禁止表現の有無」を使用したスコアによる順位付け

P c : 基準 (c) 「検索における順位」を使用したスコアによる順位付け

P d : 基準 (d) 「質問に含まれる行動表現との距離」を使用したスコアによる順位付け

40

け

P e : 基準 (e) 「ガ格の名詞がかかる行動表現」を使用したスコアによる順位付け

また、比較のために、

f r e q : 出現頻度による順位付け

という条件で評価した結果も示す。

【0079】

50

まず、freqとその他を比較すると、基準(d)以外に関しては、いずれかの基準を用いることで平均F値が上がり結果が良くなっており、基準(d)以外は単独で用いても有効な基準といえる。

【0080】

また、allとPa~Peを比較すると、出力する行動表現の数によらず、allの方が良い結果であった。すなわち、基準の組み合わせによる相乗効果があることが分かった。

【0081】

上記評価実験ではhow型質問についての調査を示したが、本発明によれば、原理的にはwhere-to-do質問やwhen-to-do質問にも応用可能であり、「行動を問う質問」に対して幅広い観点から回答することができる。

【0082】

なお、本発明は上記実施例に限定されるものではなく、本発明の趣旨に基づき種々の変形が可能であり、これらを本発明の範囲から排除するものではない。

【産業上の利用可能性】

【0083】

本発明の質問応答システムは、行動を問う質問に対して、領域を限定することなく回答することができる質問応答システムとして利用可能である。

【図面の簡単な説明】

【0084】

【図1】本発明の実施例を示す質問応答システムの構成図である。

【図2】本発明の実施例を示す質問応答システムにおける回答画面例を示す図である。

【図3】本発明の実験例における各条件によるF値グラフである。

【符号の説明】

【0085】

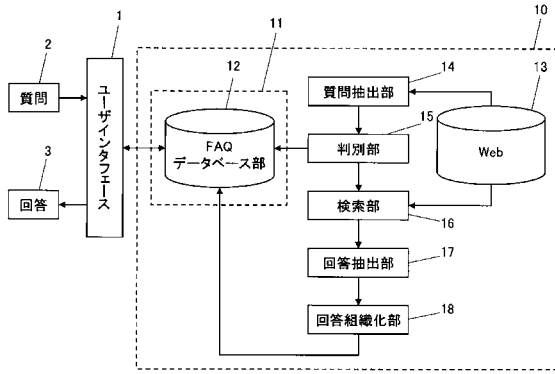
- 1 ユーザインタフェース装置
- 2 質問(質問文:質問情報)
- 3 回答(回答文:回答情報)
- 10 事前処理部
- 11 回答部
- 12 FAQデータベース部
- 13 ウェブ
- 14 質問抽出部
- 15 判別部
- 16 検索部
- 17 回答抽出部
- 18 回答組織化部

10

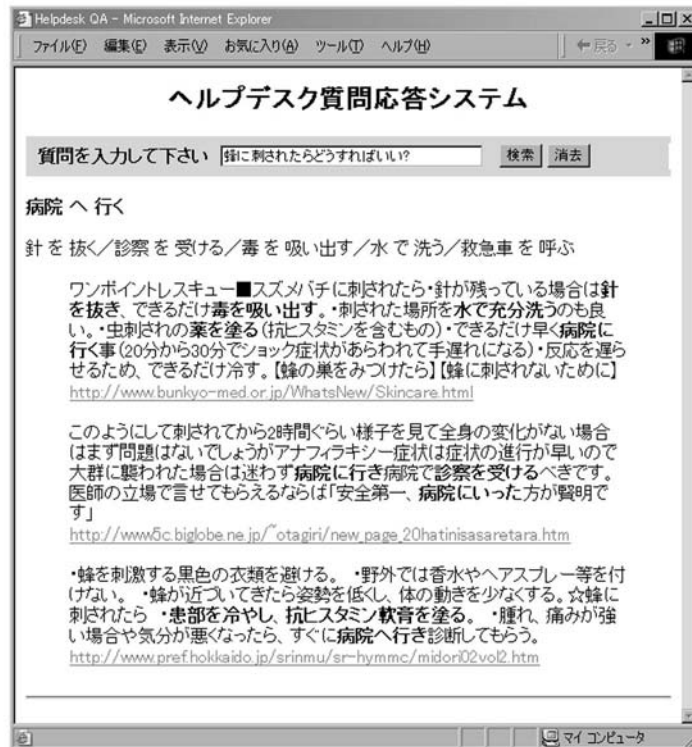
20

30

【 図 1 】



【 図 2 】



【 図 3 】

