

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第6067616号
(P6067616)

(45) 発行日 平成29年1月25日(2017.1.25)

(24) 登録日 平成29年1月6日(2017.1.6)

(51) Int.Cl.

F 1

G 1 O L 15/22 (2006.01)
G 1 O L 13/00 (2006.01)G 1 O L 15/22 3 O O U
G 1 O L 13/00 1 O O M

請求項の数 8 (全 15 頁)

(21) 出願番号 特願2014-90935 (P2014-90935)
 (22) 出願日 平成26年4月25日 (2014.4.25)
 (65) 公開番号 特開2015-210342 (P2015-210342A)
 (43) 公開日 平成27年11月24日 (2015.11.24)
 審査請求日 平成27年11月30日 (2015.11.30)

(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一丁目5番1号
 (74) 代理人 100121706
 弁理士 中尾 直樹
 (74) 代理人 100128705
 弁理士 中村 幸雄
 (74) 代理人 100147773
 弁理士 義村 宗洋
 (72) 発明者 目黒 豊美
 東京都千代田区大手町一丁目5番1号 日本電信電話株式会社内
 (72) 発明者 杉山 弘晃
 東京都千代田区大手町一丁目5番1号 日本電信電話株式会社内

最終頁に続く

(54) 【発明の名称】発話生成手法学習装置、発話生成手法選択装置、発話生成手法学習方法、発話生成手法選択方法、プログラム

(57) 【特許請求の範囲】

【請求項 1】

所定のユーザ発話文に対して、前記ユーザ発話文に対する応答であるシステム発話文を予め定めた複数の発話生成手法によって生成した場合の、前記発話生成手法ごとのシステム発話文の適切さに基づいて、前記ユーザ発話文に対して選択すべき発話生成手法を人手で指定した情報である指定情報と、前記指定情報に対応する前記ユーザ発話文とを対応付けて記憶する指定情報付与済み発話データ記憶部と、

前記指定情報付与済み発話データ記憶部から選択した前記ユーザ発話文を形態素解析する形態素解析部と、

前記形態素解析の結果から特徴量を生成する特徴量生成部と、

前記生成された特徴量に基づいて、前記選択されたユーザ発話文に対応する前記指定情報を出力するモデルを学習するモデル学習部と、

前記学習されたモデルを記憶するモデル記憶部と、を含む
発話生成手法学習装置。

【請求項 2】

所定のユーザ発話文に対して、前記ユーザ発話文に対する応答であるシステム発話文を予め定めた複数の発話生成手法によって生成した場合の、前記発話生成手法ごとのシステム発話文の適切さに基づいて、前記ユーザ発話文に対して選択すべき発話生成手法を人手で指定した情報である指定情報と、前記指定情報に対応する前記ユーザ発話文の形態素解析の結果から生成した特徴量と、に基づいて学習されたモデルを記憶するモデル記憶部と

、発話生成手法選択対象として入力されたユーザ発話文を形態素解析する形態素解析部と、
 前記形態素解析の結果から特徴量を生成する特徴量生成部と、
 前記生成された特徴量と、前記モデル記憶部に記憶されたモデルに基づいて、前記指定情報生成し、生成された指定情報を出力する発話生成手法選択部と、を含む
 発話生成手法選択装置。

【請求項3】

所定のユーザ発話文に対して、前記ユーザ発話文に対する応答であるシステム発話文を
 予め定めた複数の発話生成手法によって生成した場合の、前記発話生成手法ごとのシステム
 発話文の適切さに基づいて、前記ユーザ発話文に対して選択すべき発話生成手法を人手
 で指定した情報である指定情報と、前記指定情報に対応する前記ユーザ発話文と、前記各
 システム発話文とを対応付けて記憶する指定情報付与済み発話データ記憶部と、

10

前記指定情報付与済み発話データ記憶部から選択した前記ユーザ発話文、又は前記各シ
 ステム発話文を形態素解析する形態素解析部と、

前記形態素解析の結果から特徴量を生成する特徴量生成部と、

前記生成された特徴量に基づいて、前記選択されたユーザ発話文に対応する前記指定情
 報を出力するモデルを学習するモデル学習部と、

前記学習されたモデルを記憶するモデル記憶部と、を含む

発話生成手法学習装置。

20

【請求項4】

所定のユーザ発話文に対して、前記ユーザ発話文に対する応答であるシステム発話文を
 予め定めた複数の発話生成手法によって生成した場合の、前記発話生成手法ごとのシステム
 発話文の適切さに基づいて、前記ユーザ発話文に対して選択すべき発話生成手法を人手
 で指定した情報である指定情報と、前記指定情報に対応する前記ユーザ発話文、又は前記
 各システム発話文の形態素解析の結果から生成した特徴量と、に基づいて学習されたモ
 デルを記憶するモデル記憶部と、

発話生成手法選択対象として入力されたユーザ発話文、又は前記入力されたユーザ発話
 文から生成した各システム発話文を形態素解析する形態素解析部と、

前記形態素解析の結果から特徴量を生成する特徴量生成部と、

30

前記生成された特徴量と、前記モデル記憶部に記憶されたモデルに基づいて、前記指定
 情報を生成し、生成された指定情報を出力する発話生成手法選択部と、を含む
 発話生成手法選択装置。

【請求項5】

前記特徴量が、

前記ユーザ発話文または前記各システム発話文に含まれる単語N-gram、品詞N-gram、意味
 属性N-gram、単語N-gramを圧縮した意味ベクトル、別途用意されたテキスト文書から予
 め学習された言語モデルに対する前記ユーザ発話文または前記システム発話文のパープレ
 キシティ、ルールベース発話生成手法が用いるルールと前記ユーザ発話文の一致度、前記
 ユーザ発話文と前記システム発話文の類似度のうち何れか一つ以上のパラメータからなる
 請求項3又は4に記載の装置。

40

【請求項6】

所定のユーザ発話文に対して、前記ユーザ発話文に対する応答であるシステム発話文を
 予め定めた複数の発話生成手法によって生成した場合の、前記発話生成手法ごとのシステム
 発話文の適切さに基づいて、前記ユーザ発話文に対して選択すべき発話生成手法を人手
 で指定した情報である指定情報と、前記指定情報に対応する前記ユーザ発話文とを対応付
 けて記憶する指定情報付与済み発話データ記憶部を含む発話生成手法学習装置が実行する
 、発話生成手法学習方法であって、

前記指定情報付与済み発話データ記憶部から選択した前記ユーザ発話文を形態素解析す
 る形態素解析ステップと、

50

前記形態素解析の結果から特徴量を生成する特徴量生成ステップと、
前記生成された特徴量に基づいて、前記選択されたユーザ発話文に対応する前記指定情報
を出力するモデルを学習するモデル学習ステップと、を含む
発話生成手法学習方法。

【請求項 7】

所定のユーザ発話文に対して、前記ユーザ発話文に対する応答であるシステム発話文を
予め定めた複数の発話生成手法によって生成した場合の、前記発話生成手法ごとのシステム
発話文の適切さに基づいて、前記ユーザ発話文に対して選択すべき発話生成手法を人手
で指定した情報である指定情報と、前記指定情報に対応する前記ユーザ発話文の形態素解
析の結果から生成した特徴量と、に基づいて学習されたモデルを記憶するモデル記憶部を
含む発話生成手法選択装置が実行する、発話生成手法選択方法であって、

発話生成手法選択対象として入力されたユーザ発話文を形態素解析する形態素解析ステ
ップと、

前記形態素解析の結果から特徴量を生成する特徴量生成ステップと、
前記生成された特徴量と、前記モデル記憶部に記憶されたモデルに基づいて、前記指定
情報を生成し、生成された指定情報を出力する発話生成手法選択ステップと、を含む
発話生成手法選択方法。

【請求項 8】

コンピュータを、請求項 1 から 5 の何れかに記載の装置として機能させるためのプログ
ラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、雑談対話システムにおいて用いられる発話生成方法のうち、最適な発話生成
方法を選択するためのモデルを学習する発話生成手法学習装置、最適な発話生成方法を選
択する発話生成手法選択装置、発話生成手法学習方法、発話生成手法選択方法、プログラ
ムに関する。

【背景技術】

【0002】

雑談対話システムとは、ユーザからの発話に対して、システムが発話することによって
、ユーザとシステムの間での雑談対話を実現するシステムのことである。ここで、ユーザ
からの発話は、「雑談」と称されるように、特定の情報を要求する発話ではなく、明確な
目的のない発話であることが想定されている。以下、ユーザからの発話（音声信号）を「
ユーザ発話」、あるいは「ユーザの発話」と呼称し、「ユーザ発話」のテキスト情報を「
ユーザ発話文」と呼称する。「ユーザ発話文」は、「ユーザ発話」を音声認識によってテ
キスト情報に変換したものであってもよいし、ユーザが端末に入力したテキスト情報であ
ってもよい。また、上述の、システムの発話を「システム発話」、あるいは「システムの
発話」と呼称し、「システム発話」のテキスト情報を「システム発話文」と呼称する。シ
ステムからユーザへの発話はテキスト情報である「システム発話文」として得られるのが
通常である。従って、「システム発話」は、テキスト情報である「システム発話文」を音
声信号に変換した情報である。

【0003】

雑談対話システムを実現する上で最もよく用いられている発話生成手法は、ユーザの発
話に応答するルールを人手で大量に記述する手法である。以下、この手法を「ルールベ
ースの発話生成手法」、「ルールベース発話生成手法」、あるいは単に「ルールベース」と
呼称する。ルールベースの発話生成手法は、ユーザ発話に合致するルールがある場合、適
切な返答をすることができる、ルール作成者の直感や考えに沿った変更が容易であるとい
う利点がある。ルールベースの発話生成手法については、非特許文献 1 に開示されている。

【0004】

一方で近年、大規模なテキストデータから、統計的に発話を生成する手法が見出されて

10

20

30

40

50

いる。以下、この手法を「統計的発話生成手法」、「統計的な発話生成手法」あるいは単に「統計」と呼称する。統計的発話生成手法については、非特許文献2、3に開示されている。統計的発話生成手法は、大規模なデータから統計的にユーザ発話に関連性の高い文章を抽出して、システム発話とすることで、様々な話題のユーザ発話に対して返答を生成することができる。

【先行技術文献】

【非特許文献】

【0005】

【非特許文献1】Wallace, R. S.: The Anatomy of A.L.I.C.E., A.L.I.C.E. Artificial Intelligence Foundation, Inc. (2004)

10

【非特許文献2】Bessho, F., Harada, T., and Kuniyoshi, Y.: Dialog System Using Real-Time Crowdsourcing and Twitter Large-Scale Corpus, in Proc. SIGDIAL, pp. 227-231(2012).

【非特許文献3】柴田雅博、富浦洋一、西口友美：雑談自由対話を実現するためのWWW上の文書からの妥当な候補文選択手法、人工知能学会論文誌、Vol. 24, No. 6, pp. 507-519(2009)

【発明の概要】

【発明が解決しようとする課題】

【0006】

ルールベース発話生成手法によれば、適切なルールが存在する発話には適切な返答ができる。しかし、ルールベース発話生成手法では、事前に想定されていない発話に応答する場合や、部分的なルールのマッチを許容する場合などには不適切な応答がある。通常、ルールベース発話生成手法が用いるルールは、ユーザ発話文にマッチさせるためのパターン(pattern)と、そのパターンとマッチするテンプレート(template)が対となっている。patternとtemplateの例を以下に示す。

```
<pattern> * T O E I C *</pattern>
<template>難しそうですね</template>
```

【0007】

上述の<pattern></pattern>で挟まれた部分はユーザ発話とマッチさせるためのパターンである。パターン中の*は任意の文字列とマッチする。例えば、パターン<pattern> * T O E I C *</pattern>は、ユーザ発話文「明日T O E I C受けるんです」「T O E I C受けたことがありますか?」などとマッチする。<template></template>で挟まれた部分は、システム発話文の候補であるテンプレート(template)を表している。例えば、ユーザ発話文「明日T O E I C受けるんです」や「T O E I C受けたことがありますか?」と、<pattern> * T O E I C *</pattern>がマッチするため、システムは、<pattern> * T O E I C *</pattern>と対応づけられたテンプレートである<template>難しそうですね</template>を選択し、システム発話文として「難しそうですね」を決定し、システム発話文「難しそうですね」を返答する。

【0008】

この例の場合、ユーザ発話「明日T O E I C受けるんです」に対して、システムが「難しそうですね」と返答するのは、違和感は少ないと思われる。しかし、ユーザ発話「T O E I C受けたことがありますか?」に対してシステムが「難しそうですね」と返答するのは、ユーザの質問を無視しており、適切な返答とは言えない。

【0009】

一方、統計的発話生成手法を用いる場合は、大規模なテキストデータから、発話を生成するため、様々な話題、内容について返答を作成することが可能である。統計的発話生成手法には大きくわけて二つの生成手法がある。一つ目は、それまでの対話に出てきた発話に関連性の高い単語を用いて、テンプレートに当てはめる手法、二つ目は、人間が発話した大量の対話データ(例えば、SNSにおける対話データのログ)から関連性の高い発話を抽出する手法である。一つ目の手法では、文法上の間違いや、適切でない単語を含む発

20

30

40

50

話が生成される場合がある。また、二つ目の手法では、すべてのユーザ発話に対して適切な応答を選択することは難しく、適切でない発話が選択される場合がある。発明者らが先行研究において考案した統計的発話生成手法（詳細は、参考非特許文献1参照）では、ユーザ発話から発話の内容を表す係り受け構造（係り受け関係をもつ2文節の組）を抽出し、抽出された係り受け構造と係り受け関係を持つ係り受け構造とを合わせてテンプレートへ代入することで、ユーザ発話に関連した新しい情報を含む発話文を生成している。

（参考非特許文献1：杉山弘晃、目黒豊美、東中竜一郎、南泰浩：任意の話題を持つユーザ発話に対する係り受けを利用した応答文の生成、人工知能学会研究会、SIG-SLUD、pp. 55-60(2013)）

【0010】

係り受け構造x中の文節のうち少なくとも1つが係り受け構造y中の文節と係り受け関係にある場合、係り受け構造yは、前述の「係り受け構造と係り受け関係を持つ係り受け構造」に該当するものと定義する。例えば「お腹が空いた」というユーザ発話文についてシステム発話文を生成する場合、ユーザ発話文から「お腹 空いた」という係り受け構造を抽出し、コーパスから当該係り受け構造と係り受け関係を持つ係り受け構造を検索する。検索結果中に、「ご飯 食べる」という係り受け構造との係り受け関係が多く存在していた場合には、「 ますか？」というテンプレートに代入し、システム発話文「ご飯食べますか？」を取得することが出来る。上記のような手法でシステム発話文を生成した場合、ユーザ発話に関連した発話を返答することができるが、返答文としての適切さが保障されないという欠点がある。

【0011】

上述したように、ルールベース発話生成手法は、ルールとマッチするユーザ発話文には適切な返答ができるが、ルールにマッチしない発話文には適切な返答ができない。一方、統計的発話生成手法は、広い話題に対して返答が可能だが、適切ではない返答を行うことがある。

【0012】

そこで本発明では、入力されたユーザ発話が何れの発話生成手法で返答すべき発話であるかを識別する発話生成手法学習装置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】

【0013】

本発明の発話生成手法学習装置は、指定情報付与済み発話データ記憶部と、形態素解析部と、特徴量生成部と、モデル学習部と、モデル記憶部を含む。

【0014】

指定情報付与済み発話データ記憶部は、所定のユーザ発話文に対して、ユーザ発話文に対する応答であるシステム発話文を予め定めた複数の発話生成手法によって生成した場合の、発話生成手法ごとのシステム発話文の適切さに基づいて、ユーザ発話文に対して選択すべき発話生成手法を人手で指定した情報である指定情報と、指定情報に対応するユーザ発話文とを対応付けて記憶する。形態素解析部は、指定情報付与済み発話データ記憶部から選択したユーザ発話文を形態素解析する。特徴量生成部は、形態素解析の結果から特徴量を生成する。モデル学習部は、生成された特徴量に基づいて、選択されたユーザ発話文に対応する指定情報を出力するモデルを学習する。モデル記憶部は、学習されたモデルを記憶する。

【発明の効果】

【0015】

本発明の発話生成手法学習装置によれば、入力されたユーザ発話が何れの発話生成手法で返答すべき発話であるかを識別することができる。

【図面の簡単な説明】

【0016】

【図1】実施例1の発話生成手法学習装置、発話生成手法選択装置、発話生成手法選択学習装置の構成を示すブロック図。

10

20

30

40

50

【図2】指定情報付与済み発話データの生成過程を例示する図。

【図3】実施例1の発話生成手法学習装置の動作を示すフローチャート。

【図4】実施例1の発話生成手法選択装置の動作を示すフローチャート。

【図5】特徴量を例示する図。

【図6】ユーザ発話文の文例に対し単語、品詞、意味属性をunigramで抽出した例を示す図。

【図7】本発明において学習され用いられるモデルの識別精度を確認する実験の結果を示す図。

【発明を実施するための形態】

【0017】

以下、本発明の実施の形態について、詳細に説明する。なお、同じ機能を有する構成部には同じ番号を付し、重複説明を省略する。

【0018】

<ルールベース発話生成手法>

以下、実施例で用いるルールベース発話生成手法について概説する。実施例で用いるルールベース発話生成手法は、Artificial Intelligence Markup Languageという対話システムのルールを記述するためのマークアップ言語に則って実行される。このマークアップ言語では、ユーザ発話文とのマッチングに用いるパターンとシステム発話文の候補であるテンプレートは、例えば以下のように記述される。

[例1]

<pattern>* カラオケ* 行かれ* か</pattern>
<template>はい、よく行きます。ストレス発散にいいですよね。</template>

[例2]

<pattern>花火* 見* 行か* か</pattern>
<template>そうですね。手持ちより打ち上げが好き</template>

【0019】

ルールベース発話生成手法は、ユーザ発話文が与えられたときに、適切なルールを選択してシステム発話文を生成する手法である。実施例では、ルールを選択するにあたって、TF-IDFの重み付コサイン類似度によって、ユーザ発話文に最も類似しているパターンを選択する方法を用いることができる。具体的には、パターンに含まれる単語列ベクトルと、ユーザ発話文に含まれるすべての単語列ベクトルに含まれる単語にそれぞれTF-IDFで重み付けし、そのベクトル間のコサイン類似度を用いた。TF-IDFは、TF(単語の出現頻度)とIDF(逆文書頻度)の二つの指標にもとづいて計算される。IDFは一般語フィルタとして働き、多くのドキュメントに出現する語(一般的な語)は重要度が下がり、特定のドキュメントにしか出現しない単語の重要度を上げる役割を果たす。コサイン類似度は、ベクトル間の類似度を測る一般的な手法である。この類似パターン選択方法は、発話の近さなどを測る上で一般的な手法である。

【0020】

<統計的発話生成手法>

本実施例に用いる統計的発話生成手法として、例えば参考非特許文献1に開示された手法を用いることができる。

【実施例1】

【0021】

<発話生成手法学習装置1>

以下、図1、図2、図3を参照して本発明の実施例1の発話生成手法学習装置について説明する。図1は、本実施例の発話生成手法学習装置1の構成を示すブロック図である。図2は、指定情報付与済み発話データの生成過程を例示する図である。図3は、本実施例の発話生成手法学習装置1の動作を示すフローチャートである。図1に示すように、本実施例の発話生成手法学習装置1は、指定情報付与済み発話データ記憶部11と、形態素解析部12と、特徴量生成部13と、モデル学習部14と、モデル記憶部15を含む。指定

10

20

30

40

50

情報付与済み発話データ記憶部11には、指定情報付与済み発話データが記憶されている。指定情報付与済み発話データとは、1つのユーザ発話文に対して、システムが何れの発話生成手法で返答するかを指定する情報である指定情報を付与したデータのことである。

【0022】

<指定情報付与済み発話データ記憶部11>

以下、図2の例を参照して、指定情報付与済み発話データの生成手順について説明する。指定情報付与済み発話データの生成は人間の手で行われる。以下、指定情報付与済み発話データの生成に関する者をデータ生成者と総称する。データ生成者は、図2に示される指定情報付与済み発話データの例を次のように作成する。まずデータ生成者は、対話データから複数の発話文を「ユーザ発話文」として抽出する。データ生成者は、抽出された「ユーザ発話文」を、各発話生成手法を用いたモジュールに入力し、各モジュールから「システム発話文」を取得する。データ生成者は、取得した「システム発話文」一つ一つの「発話の適切さ」を評定する。図2の例では、「発話の適切さ」を評定する5段階の評定値である。データ生成者は、「発話の適切さ」により示される最も適切な発話生成手法を、対応するユーザ発話文に対して選択すべき発話生成手法であるとして指定する「指定情報」を決定し、対応するユーザ発話文に当該指定情報を付与する。ここで、複数の発話生成手法の「発話の適切さ」が同等であった場合には、何れかの発話生成手法が選択される例外ルールを定義しておいてもよい。例えばルールベース発話生成手法と統計的発話生成手法の評定値が同点の場合は、ルールベースを選んでもよい。

【0023】

図2の例では、データ生成者は、対話データから抽出されたユーザ発話文である「リーガルハイは見とけばよかったなあと後悔してる。」を、既存の統計的発話生成手法を用いたモジュールに入力し、当該モジュールからシステム発話文である「ダンダリンもリーガルハイだね」を取得する。同様に、データ生成者は、ユーザ発話文「リーガルハイは見とけばよかったなあと後悔してる。」を、既存のルールベース発話生成手法を用いたモジュールに入力し、当該モジュールからシステム発話文である「そうでしょうね」を取得する。データ生成者は、取得したシステム発話文「ダンダリンもリーガルハイだね」に、「発話の適切さ」を評定する5段階の評定点のうち、例えば2点を付与する。同様に、データ生成者は、システム発話文「そうでしょうね」に、「発話の適切さ」を評定する5段階の評定点のうち、例えば5点を付与する。データ生成者は、5段階の評定点により示される最も適切な発話生成手法である「ルールベース」を、対応するユーザ発話文「リーガルハイは見とけばよかったなあと後悔してる。」に対して選択すべき発話生成手法であるとして指定する「指定情報」を決定し、当該ユーザ発話文に当該指定情報を付与する。

【0024】

なお、上記の個別に評価値を付与してから適切な発話生成手法を選択する方法に加え、直接各発話生成手法から取得したシステム発話文を見比べることで、評価値を経由せずに適切な発話生成手法を選択し指定情報とする方法も考えられる。また、複数のシステム発話文を各システムに生成させ、それら全体の傾向を見比べることで、指定情報を決定してもよい。

【0025】

このように、指定情報付与済み発話データ記憶部11は、所定のユーザ発話文に対して、ユーザ発話文に対する応答であるシステム発話文を予め定めた複数の発話生成手法によって生成した場合の、発話生成手法ごとのシステム発話文の適切さに基づいて、ユーザ発話文に対して選択すべき発話生成手法を人手で指定した情報である指定情報と、指定情報に対応するユーザ発話文とを対応付けて記憶している。

【0026】

また、指定情報付与済み発話データ記憶部11は、所定のユーザ発話文に対して、指定情報に加え、各システム発話文を対応付けて記憶することとしてもよい。以下、本実施例の発話生成手法学習装置1の各構成要件の動作を説明する。

【0027】

10

20

30

40

50

<形態素解析部 1 2 >

形態素解析部 1 2 は、指定情報付与済み発話データ記憶部 1 1 から選択したユーザ発話文を形態素解析する (S 1 2)。指定情報付与済み発話データ記憶部 1 1 に各システム発話文も併せて記憶されている場合、形態素解析部 1 2 は、指定情報付与済み発話データ記憶部 1 1 から選択したシステム発話文を形態素解析することとしてもよい。例えば、ユーザ発話文「バイオハザード 5 (登録商標) はいつ放送するんです？」の形態素解析結果は以下のようになる。

バイオハザード 5 名詞: 固有バイオハザード5 [] [] [] オンライン百科事典 A

は 連用助詞はハ [] [] []

10

いつ 連用詞いつイツ [2670] [] []

放送 名詞: 動作放送ホウソウ [1540, 1552, 1120, 919] [] []

する 動詞接尾辞: 連体するスル [2050] [] []

ん 補助名詞ん [1] [] []

です 判定詞: 終止ですデス [] [] []

? 句点: 疑問符 ? [] [] []

EOS

【0028】

各行の先頭は単語毎に分けられた際の文中の単語そのものの表記、次に品詞、基本形、読み(存在する場合にはカタカナで表記)、[]内は、語彙大系の意味属性の番号を表している。例えば、2670は「時間」、1540は「報道」、2050は「実行」という意味を表している。複数の意味属性が付いている場合は、先頭のみを用いる。行末の「オンライン百科事典 A」は、オンライン百科事典 A の見出し語として存在する単語であることを表しており、これは単語に対する付加情報として我々が独自に出力している情報である。EOSは文末を表している。

20

【0029】

<特徴量生成部 1 3 >

特徴量生成部 1 3 は、形態素解析の結果から特徴量を生成する (S 1 3)。以下、ステップ S 1 3 で用いられる特徴量について、図 5 を参照して説明する。図 5 は、特徴量を例示する図である。図 5 に示すように、例えば特徴量 9 は、ユーザ発話文または各発話生成手法から取得された各システム発話文に含まれる単語N-gram 9 1、品詞N-gram 9 2、意味属性N-gram (語彙大系N-gram) 9 3、単語N-gram 9 1 をトピックモデルや行列分解を用いて圧縮・抽象化したベクトル列 (意味ベクトル) 9 1 '，別途用意されたテキスト文書から予め学習された言語モデルに対するユーザ発話文またはシステム発話文のパーセンテージ 9 4、ルールベース発話生成手法が用いるルールとユーザ発話文の一一致度 9 5、ユーザ発話文と各システム発話文の類似度 9 6 のうち何れか一つ以上のパラメータからなるものとすることができます。

30

【0030】

以下、それぞれのパラメータについて説明する。

40

【0031】

<単語N-gram 9 1、品詞N-gram 9 2、意味属性N-gram (語彙大系N-gram) 9 3 >

パラメータとして、単語N-gram 9 1、品詞N-gram 9 2、意味属性N-gram (語彙大系N-gram) 9 3 を用いることが出来る。

【0032】

語彙大系には、日本語の語彙30万語に約3000種類の意味属性が付与されている。上記の形態素解析の結果において [] で囲んで表記した数値が、それぞれの意味属性の番号を指している。例えば、品詞、単語、語彙大系の一つ組(unigram)を用いる場合には、例えば図 6 のようなパラメータとなる。図 6 では、ユーザ発話文「バイオハザード (登録商標) 5 はいつ放送するんです？」に対し、単語N-gram 9 1、品詞N-gram 9 2、意味属性N-gram

50

9 3 、のNを1とした場合の各パラメータについて表記している。各パラメータは「要素名：数値」の形式で表記されており、「要素名」はユーザ発話に含まれる単語、品詞、語彙大系の意味属性、を表しており、「数値」はユーザ発話中に出現している回数を表している。例えば、「バイオハザード（登録商標）5:1」は、単語「バイオハザード（登録商標）5」が1回出現していることを表している。

【0033】

図6の例では、N-gramのNが1である場合、つまりunigramのみを用いたが、Nは1以外でもよい。例えば、Nを2とすれば、「バイオハザード（登録商標）5は」の出現回数1回、というように、単語の二つ組の回数を用いることになる。

【0034】

また、単語N-gramを用いた特徴量は、一般に極めて疎である。この疎性を緩和する方法として、トピックモデルや行列分解による、単語情報圧縮手法が知られている（例えば参考非特許文献2、3）。これらの方針によって、単語を意味空間に写像し意味ベクトルを得ることができる。本発明では、単語N-gram 9 1から意味ベクトル 9 1'を得て、これらを 9 1 の代用、もしくは併用する形でパラメータの一つとすることができます。

（参考非特許文献2：D. Blei, A. Ng, and M. Jordan, "Latent Dirichlet Allocation", in Journal of Machine Learning Research, 2003, pp. 1107-1135.）

（参考非特許文献3：Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. Indexing by latent semantic analysis. Journal of the American Society for Information Science, 41, 391-407 (1990).）

【0035】

<パープレキシティ94>

ルールベース発話生成手法が持つルールは、事前に想定された発話に対して作られる。つまり、通常の対話で生起しやすいと思われる発話に対して作られることが多いため、通常の対話ではあまり生起しないような発話には、適切な返答を出力することが難しいと考えられる。パープレキシティ94は、言語モデルを学習したデータにおいて、その文章が生起しやすい発話であるか、生起しにくい発話であるかを表す尺度である。つまり、大量の自然文から学習した言語モデルに照らし合わせたとき、パープレキシティ94が低い発話は確率的に生起しにくい発話であり、ルールの中に適切な返答ができるものが含まれていない可能性が高い。逆に、パープレキシティ94が高い発話は確率的に生起しやすい発話であるため、適切なルールが存在する可能性が高い。このことから、パープレキシティ94を用いることで、ルールベース発話生成手法で適切な返答ができる発話か、そうでない発話であるか識別できると考えられる。そこで、言語モデルに対するユーザ発話のパープレキシティ94を、特徴量のパラメータの一つとして用いる。例えば所定の言語モデルを用いてパープレキシティ94を計算した場合、第1の発話文例「クックパッド（登録商標）見て余りもんでチャーハンをこさえるくらいはできますが料理を企画する力が全くおまへん。」ではパープレキシティ=63.4、第2の発話文例「こっちはさっきまで土砂降りでしたが、晴れました。」ではパープレキシティ=18.8であった。第1の発話文例のように、「クックパッド（登録商標）」のような珍しい単語や、「こさえるくらい」や「全くおまへん」といった、生起しにくい表現が含まれる発話文では、パープレキシティ94が低くなる傾向にある。一方、第2の発話文例「こっちはさっきまで土砂降りでしたが、晴れました。」のように生起しやすい表現が多く含まれる発話文では、パープレキシティ94が高くなる傾向がある。パープレキシティ94が高い発話は、通常の対話でも生起しやすい発話である可能性が高く、当該発話はルールが想定する発話である可能性が高いと期待できる。

【0036】

また、例えば、ルールベース発話生成手法がデータAに基づいて構築され、統計的発話生成手法がデータBに基づいて構築された場合、それぞれのデータにおけるパープレキシティを計算し、データAにおけるパープレキシティが低ければ、比較的ルールベース発話生成手法が想定した発話に近いと予想される。この直観を反映するため、パープレキシティ

10

20

30

40

50

イを複数の言語データから構成し、それらの値を用いる方法や、それらの値の差分値を用いる方法を用いても良い。

【0037】

<ルールとの一致度95>

前述のように、ルールベース発話生成手法を用いて発話を生成する際には、ユーザ発話とルール中のパターンとのTF-IDFの重み付コサイン類似度を求め、最も近いパターンを持つルールを選択することができる。ルールとの一致度が高く、よくマッチしているユーザ発話に対しては、システム発話としてルールベースの出力を採用すべきと考えられる。この一致度を反映させるため、ルールを選択する際に求めたコサイン類似度の値（ルールとの一致度95）を特徴量のパラメータの一つとする。

10

【0038】

<ユーザ発話文と各システム発話文の類似度96>

ルールとの一致度を計算するときと同様に、TF-IDFの重み付きコサイン類似度を求め、この値を特徴量のパラメータの一つとする。

【0039】

<モデル学習部14>

モデル学習部14は、生成された特徴量に基づいて、選択されたユーザ発話文に対応する指定情報を出力するモデルを学習する（S14）。モデル学習部14は、学習したモデルをモデル記憶部15に記憶する。学習方法としてSupport Vector machine (SVM)や、ロジスティック回帰や、決定木などの識別モデルを学習する手法を用いることができる。これらのモデル構築の手法は、機械学習を用いる上で一般的な手法である。学習されたモデルは、モデル記憶部15に記憶される。

20

【0040】

<モデル記憶部15>

前述したように、モデル記憶部15には、所定のユーザ発話文に対するシステム発話文を予め定めた複数の発話生成手法によって生成した場合の、発話生成手法ごとのシステム発話文の適切さに基づいて、ユーザ発話文に対して選択すべき発話生成手法を人手で指定した情報である指定情報と、指定情報に対応するユーザ発話文又は各システム発話文の形態素解析の結果から生成した特徴量と、に基づいて学習されたモデルが記憶されている。

30

【0041】

<発話生成手法選択装置2>

以下、引き続き図1、新たに図4を参照して本実施例の発話生成手法選択装置2について説明する。図4は、本実施例の発話生成手法選択装置2の動作を示すフローチャートである。図1に示すように、発話生成手法選択装置2は、発話生成手法学習装置1と共に構成要素である形態素解析部12と、特徴量生成部13と、モデル記憶部15を含み、発話生成手法学習装置1に含まれない発話生成手法選択部16を含む。

【0042】

モデル記憶部15には、前述の発話生成手法学習装置1により予め学習されたモデルが記憶済みであるものとする。発話生成手法選択装置2の形態素解析部12、特徴量生成部13は、上述と同様のステップS12、S13を実行する。すなわち、発話生成手法選択装置2の形態素解析部12は、発話生成手法選択対象として新たに入力されたユーザ発話文を形態素解析する（S12）。あるいは、発話生成手形態素解析部12は、新たに入力されたユーザ発話文から生成した各システム発話文を形態素解析してもよい。形態素解析部12は、発話生成手法学習装置1の構成要件として機能する場合には、指定情報付与済み発話データ記憶部11から抽出したユーザ発話文を入力とする一方、発話生成手法選択装置2の構成要件として機能する場合には、発話生成手法選択対象として新たに入力されたユーザ発話文を入力とする点に注意する。発話生成手法選択装置2の特徴量生成部13は、形態素解析の結果から特徴量を生成する（S13）。

40

【0043】

<発話生成手法選択部16>

50

以下、発話生成手法選択部 16 について説明する。発話生成手法選択対象として入力されたユーザ発話文、または各システム発話文に対して生成された特徴量と、モデル記憶部 15 に記憶されたモデルに基づいて、対応するユーザ発話文に対して選択すべき発話生成手法であるとして指定する指定情報を生成し、生成された指定情報を出力する (S16)。

【0044】

本実施例の発話生成手法選択装置 2 を対話システムに組み込んで動作させる際には、対話システムは、ルールベースで返答すべき旨の指定情報を受信した場合には、ルールベース発話生成手法でシステム発話を生成し、統計的発話生成手法で返答すべき旨の指定情報を受信した場合には、統計的発話生成手法システム発話を生成し、ユーザに返答する。

10

【0045】

なお、上述の発話生成手法学習装置 1 は、発話生成手法選択装置 2 と組み合わせて、図 1 に示すような発話生成手法選択学習装置 3 として構成することも可能である。

【0046】

<評価実験>

本実施例で用いられるモデルの精度について、下記のような実験を行った。入力用のユーザ発話として、発明者らが収集した雑談対話のコーパスから149文、Twitter (登録商標) 中の tweet80文を選択した。雑談対話のコーパスは、発明者らのルールベース対話システムのルール作成時に参考にしたデータであり、ルール作成者が事前に対話中に起こりやすいと想定した発話群であると言える。逆に、Twitter (登録商標) からは様々な話題や言い回しの発話が抽出されており、事前に想定が難しい発話であると言える。

20

【0047】

各ユーザ発話について、ルールベース発話生成手法と、統計的発話生成手法によって返答文を生成し、人手で選択すべき発話生成手法をラベル付けし、発話の特徴量から選択すべき発話生成手法を識別するモデルをSVMで学習し、このモデルの性能を10分割交差検定で評価した。結果は図 7 のようになった。

【0048】

図 7 は、本発明において学習され、用いられるモデルの識別精度を確認する実験の結果を示す図である。図 2 の % で示した値は、各パラメータを特徴量として用いた時の識別精度 (10分割交差検定) である。「unigram」カラム、「一致度」カラム、「パープレキシティ」カラムにはそれぞれ、各パラメータのみを用いた場合の識別精度が入力されている。「すべて」カラムには、すべてのパラメータ (91、92、93、94、95) を用いた場合の識別精度が入力されている。「ランダム」カラムには、ランダムに生成手法を選択した場合の識別精度が入力されている。雑談ローには、雑談対話のコーパスから選択したユーザ発話文に対する識別精度が入力されている。Twitterローには、Twitter (登録商標) 中から選択したユーザ発話文に対する識別精度が入力されている。

30

雑談対話のコーパスから選択したユーザ発話に対しては、すべての特徴量を用いた場合が最も精度が良く、「ランダム」の条件と比較して有意に精度が向上した(マクネマー検定: $p < .05$)。一方で、Twitter (登録商標) 中から選択したユーザ発話文に対しては、すべての特徴量を用いた場合の性能が、「ランダム」の条件における性能と変わらず、パープレキシティのみを用いた時が最も精度がよく、有意に分類精度が向上した(マクネマー検定: $p < .01$)。つまり、ユーザ発話が、珍しい発話かどうかで、ルールベースで返答すべきか否かが決定するということを表している。

40

【0049】

以上のように、雑談のようにルール作成時に事前に想定できる発話群においても、Twitter (登録商標) のように事前に想定が難しい発話群においても、本発明の手法を用いることで、適切な発話生成手法が選択できるということがわかった。

【0050】

本実施例の発話生成手法学習装置 1、発話生成手法選択装置 2、発話生成手法選択学習装置 3 によれば、入力されたユーザ発話が何れの発話生成手法で返答すべき発話であるか

50

を識別することができる。

【0051】

例えば、ユーザ発話に対する返答として、ルールベース発話生成手法で発話すべきか、統計的発話生成手法で発話すべきか識別することで、現状のどちらかの発話生成手法のみを用いる対話システムと比べ、より適切な返答をすることができる。それにより、さらに円滑な自然言語によるコミュニケーションが行えるコンピュータが実現できる。

【0052】

上述の各種の処理は、記載に従って時系列に実行されるのみならず、処理を実行する装置の処理能力あるいは必要に応じて並列的あるいは個別に実行されてもよい。その他、本発明の趣旨を逸脱しない範囲で適宜変更が可能であることはいうまでもない。

10

【0053】

また、上述の構成をコンピュータによって実現する場合、各装置が有すべき機能の処理内容はプログラムによって記述される。そして、このプログラムをコンピュータで実行することにより、上記処理機能がコンピュータ上で実現される。

【0054】

この処理内容を記述したプログラムは、コンピュータで読み取り可能な記録媒体に記録しておくことができる。コンピュータで読み取り可能な記録媒体としては、例えば、磁気記録装置、光ディスク、光磁気記録媒体、半導体メモリ等どのようなものでもよい。

【0055】

また、このプログラムの流通は、例えば、そのプログラムを記録したDVD、CD-R、OM等の可搬型記録媒体を販売、譲渡、貸与等することによって行う。さらに、このプログラムをサーバコンピュータの記憶装置に格納しておき、ネットワークを介して、サーバコンピュータから他のコンピュータにそのプログラムを転送することにより、このプログラムを流通させる構成としてもよい。

20

【0056】

このようなプログラムを実行するコンピュータは、例えば、まず、可搬型記録媒体に記録されたプログラムもしくはサーバコンピュータから転送されたプログラムを、一旦、自己の記憶装置に格納する。そして、処理の実行時、このコンピュータは、自己の記録媒体に格納されたプログラムを読み取り、読み取ったプログラムに従った処理を実行する。また、このプログラムの別の実行形態として、コンピュータが可搬型記録媒体から直接プログラムを読み取り、そのプログラムに従った処理を実行することとしてもよく、さらに、このコンピュータにサーバコンピュータからプログラムが転送されるたびに、逐次、受け取ったプログラムに従った処理を実行することとしてもよい。また、サーバコンピュータから、このコンピュータへのプログラムの転送は行わず、その実行指示と結果取得のみによって処理機能を実現する、いわゆるASP(Application Service Provider)型のサービスによって、上述の処理を実行する構成としてもよい。なお、本形態におけるプログラムには、電子計算機による処理の用に供する情報であってプログラムに準ずるもの(コンピュータに対する直接の指令ではないがコンピュータの処理を規定する性質を有するデータ等)を含むものとする。

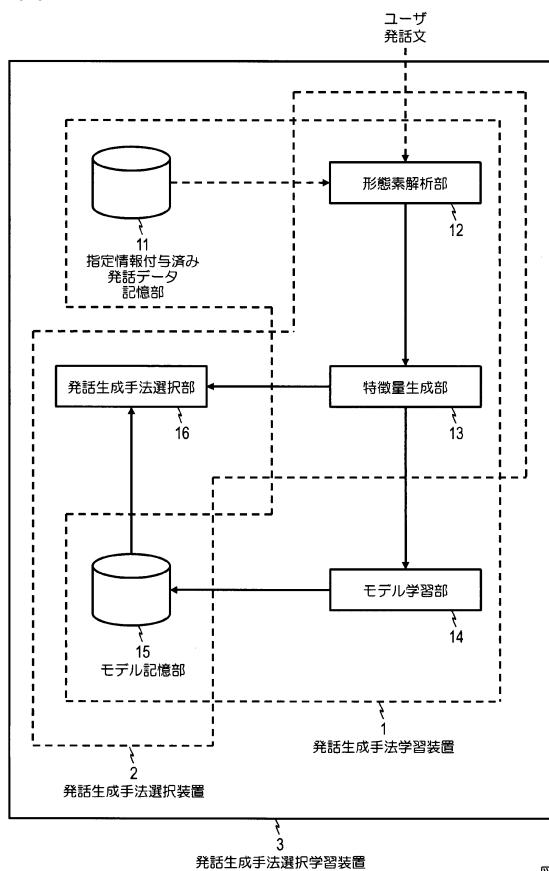
30

【0057】

また、この形態では、コンピュータ上で所定のプログラムを実行させることにより、本装置を構成することとしたが、これらの処理内容の少なくとも一部をハードウェア的に実現することとしてもよい。

40

【図1】

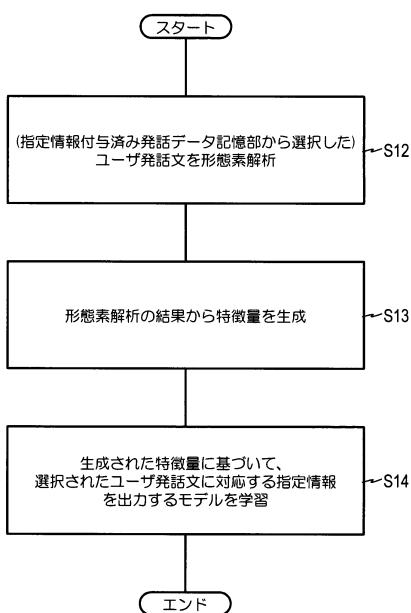


【図2】

ユーザ発話文	システム発話文	発話の適切さ評定点(5段階)	選択すべき発話生成方法(指定情報)
リーガルハイは見とけばよかったですあと後悔してる。	ダンダリンもリーガルハイだね。(統計)	2	ルールベース
	そうでしょうね。(ルールベース)	5	
おいしいものを食べすぎて痛風になんじゃないか。	なに美味しい?(統計)	2	統計
	恋人ですかね~。あなたは。(ルールベース)	1	
肉じゃがを煮るときに塩麹いれたらめちゃくちゃ美味しい。	肉じゃが食べたいです(統計)	5	ルールベース
	ああ、使ったことないです。どんな料理に使うですか?使います?(ルールベース)	5	

図2

【図3】



【図4】

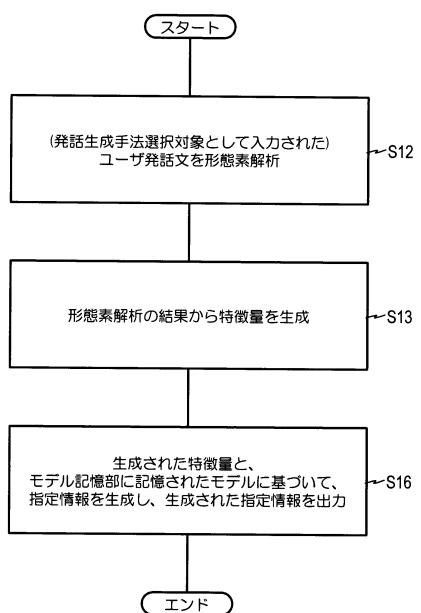


図3

図4

【図5】

特徴量 9	単語N-gram	91
	意味ベクトル	91'
	品詞N-gram	92
	意味属性N-gram	93
	パープレキシティ	94
	ルールとの一致度	95
	類似度	96

パラメータ	特徴量							
	バイオ ハザード 5 :1	は	いつ	放送	する	ん	です	?
単語 unigram		:1	:1	:1	:1	:1	:1	:1
品詞 unigram	名詞 :固有 :1	運用助詞	運用詞	名詞 :動作 :1	動詞接尾辞 :連体 :1	補助名詞	判定詞 :終止 :1	句点 :疑問符 :1
意味属性 unigram				2670 (時間) :1	1540 (報道) :1	2050 (実行) :1		

図5

図6

【図7】

入力文	ランダム	unigram	一致度	パープレキシティ	すべて
雑談	69.8%	71.3%	75.3%	58.7%	78.7%
Twitter	50.0%	53.8%	22.5%	62.5%	53.4%

図7

フロントページの続き

(72)発明者 東中 竜一郎
東京都千代田区大手町一丁目5番1号 日本電信電話株式会社内

(72)発明者 南 泰浩
東京都千代田区大手町一丁目5番1号 日本電信電話株式会社内

審査官 大野 弘

(56)参考文献 特開2013-167985(JP,A)

(58)調査した分野(Int.Cl., DB名)

G 10 L 15 / 22

G 10 L 13 / 00