

6M-8

ユーザの発話からの嗜好学習により話題を提供する 会話エージェントシステムに関する検討

畑中 聖二 加藤 誠巳
(上智大学理工学部)

1. まえがき

人が会話を行う場合、話している相手の知識や興味を念頭に置いて話しているのが普通である。しかし、従来の会話エージェントシステム^[1]では使用しているユーザが何に興味を持っているかなどを理解することがなかった。そのためユーザはすぐにエージェントに飽きてしまうという問題があった。この問題を解決するためにユーザの嗜好を学習し、学習した嗜好を利用してエージェントから話題を提供することでこの問題を解決することを考えた。以下にその嗜好学習、話題提供の方法について述べる。

2. システムの概要

今回はユーザの嗜好の学習に関し、主として検討した。エージェントはその会話によりユーザの好む事物を学習する。取得した情報を利用し、エージェントからその話題に関する会話が始められるようにシステムを組んだ。従来の会話システムにおける本システムの位置付けを図1に示す。また、本システムは音声認識されたものをテキストデータに変換して扱っているため文字列処理のみである。

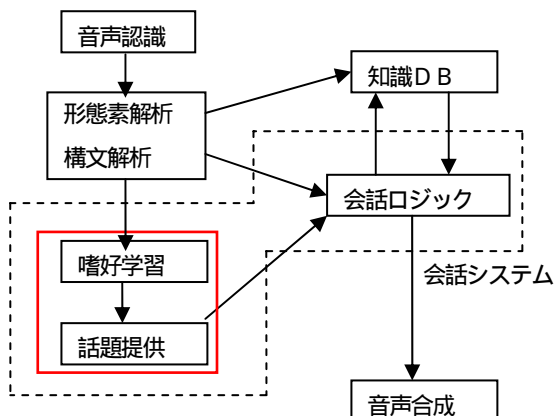


図1 本システムの位置付け

A Conversation Agent System Providing Favorite
Topics through Learning of User's Utterance
Seiji HATANAKA, Masami KATO
Sophia University

3. 嗜好の学習と話題の提供

3.1 ユーザの嗜好の判断方法

普段の生活で人間は他人の嗜好をどのように判断しているかを考えた。他人の嗜好の判断方法として以下の事を基準に判断することができると考えられる。

- ・楽しそうに話す事
- ・長く続いた話題
- ・深く内容を知っている事
- ・よく話題にする事

会話システムがこれらの情報を取得することができればユーザの嗜好が判断できる。今回の検討では「深く内容を知っている事」、「楽しそうに話す事」という二つの特徴に注目して嗜好の学習を行った。

3.2 嗜好の学習

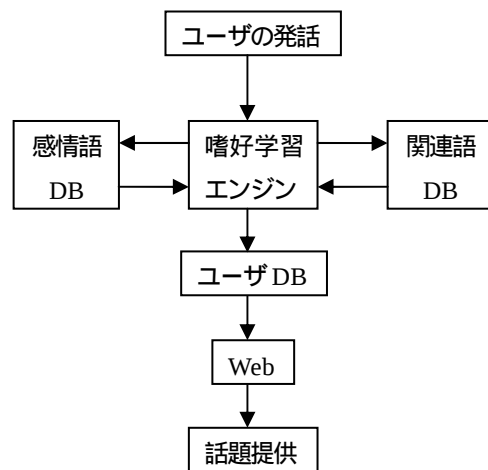


図2 本システムの流れ

ユーザの発話から話題の提供までの流れを図2を用いて説明する。

ではテキストデータになっているユーザの発話を嗜好学習エンジンに入力する。

「楽しそうに話す事」を利用するためである。日本語は言葉である程度の感情を判断することができ、話している事を嬉々として話しているのか、興味なく話しているのかを判定できる。感情語DBに

は人が喜びを表す言葉を用意しておく。例えば、ユーザの発話の中にこの言葉が現れたとき、同時に現れた名詞を取得する。ユーザ DB の中でその名詞と関連するカテゴリの評価点を加算する。また、他に「～したい」という希望を表す言葉がある。この言葉が現れたときも同様の操作を行った。

は「深く内容を知っている事」を利用するためである。各事物はそれぞれを連想させる関連した言葉を持っている。ユーザの発話の中にこの関連語が現れたときは、その関連語が連想させる事柄を好むと判断する。また、関連語は言葉によって重み付けされた値をもっていて同時にその値を取得する。重み付けの方法は、その事物について誰でも知っている言葉は値が小さく、本当に興味を持っている人のみが知る言葉は値が大きく設定している。ここで関連語 DB はその知識の量を測るものなので用意された関連語、重み付けされた値の取得は一度のみとした。

はユーザの発話から得た情報を記憶しておく DB を作る。で加算された値と で取得した値を加算し、この DB を作成していく。同時に で値を取得してきた名詞も登録される。このユーザ DB の中である閾値を超えた事物をユーザが好む物と学習する。

3.3 話題の提供

ユーザができるだけ飽きないように変化に富んだ発話を行うことが重要になってくる。そこで頻繁に情報が変化するインターネットを利用した。

は で得たユーザの情報をを用いて、インターネットの検索機能を用い、ユーザの好む事物に関連した Web サイトを探し URL を取得する。

は のサイトの本文を抜き出し、HTML タグや句点を用いて本文の中の一文を抽出する。抽出された文から記号や特殊文字などを除いた文にする。また、ユーザとの会話が進まなくなったときにユーザの入力を促すような発言をする。

4. 実行例

例えば以下のようなユーザの発話文があるとする。これらの発話文を本システムで実行したときの結果を図 3 に示した。このときのシステムの発話文はユーザ DB の中でスポーツを好むと学習していた場合

である。

User : 宮里藍って 19 歳で 1 億円獲得したんだって。

User : 田臥やっぱり解雇されたね。

User : NBA は厳しいよ。

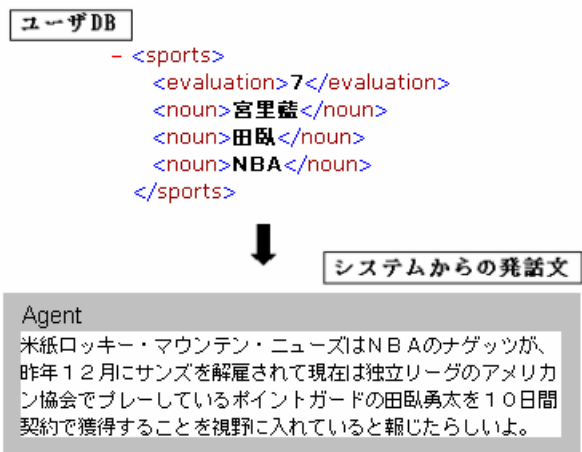


図 3 実行例

ユーザ DB は各カテゴリに対する値と会話中に現れたそのカテゴリに関連した言葉で構成されている。

5. 検討

関連語 DB を作成する際に、今回は人手で関連語 DB を作成した。しかし、事物に対する言葉は日々増加するので今後は自動化を検討していく必要がある。

話題の提供では Web サイトから文章を抜き出すときに短く、かつ内容をとらえている文章の取得方法を考慮していく必要がある。また、取得した文章を人が話すような自然な言葉にする必要がある。

6. むすび

ユーザの発話からの嗜好学習により話題を提供する会話エージェントシステムに関して検討した結果について述べた。好みの事物に対するユーザの知識の量によって嗜好を学習し、Web サイトからの情報を利用して話題を提供することにより本システムを実現した。

最後に、有益な御討論を戴いた本学 e-LAB/マルチメディア・ラボの諸氏に謝意を表する。

参考文献

- [1] 廣瀬、加藤：“ドライバの退屈・眠気回避を目的とした会話型 CG エージェントに関する検討,”