

ユーザとの対話から学習する非タスク指向型対話エージェントの設計

*稲葉通将⁽¹⁾, 鳥海不二夫⁽¹⁾, 石井健一郎⁽¹⁾

(1) 名古屋大学大学院 情報科学研究科

〒464-8603 愛知県 名古屋市 千種区 不老町

E-mail : inaba@kishii.ss.is.nagoya-u.ac.jp,

{tori,kishii}@is.nagoya-u.ac.jp

Abstract: 近年, コンピュータの高性能化に伴い, コンピュータによるコミュニケーション支援の研究が活発化している. その一つが, 人間と対話のできるコンピュータ, すなわち対話エージェントに関する研究である. 本稿では, 雑談などのオープンドメインな対話を行う非タスク指向型対話エージェントの実現のために行ってきた研究を紹介する. そして, それらを元にした, ユーザと対話を繰り返すことにより学習する対話エージェントの設計手法について述べる.

1. はじめに

近年, コンピュータの高性能化に伴い, コンピュータによるコミュニケーション支援の研究が活発化している. その一つが, 人間と対話のできるコンピュータ, すなわち対話エージェントに関する研究である.

対話エージェントには大きく分けて2種類あり, 1つはタスク指向型対話エージェント, 他の1つは非タスク指向型対話エージェントである. 前者は, チケット予約や道案内など特定のタスク達成を行う対話エージェントであり, 後者は対話そのものを目的とし, 雑談などのオープンドメインな対話を行う対話エージェントである. 我々の身の回りのコミュニケーションの大半は非タスク指向型コミュニケーションであるにもかかわらず, これまでの研究の多くはタスク指向型対話エージェントに関するものであり, 非タスク指向型対話エージェントに関する研究は少ない. そこで我々は, 人間の話し相手になり, 対話を盛り上げる対話エージェントを作成することを目標にする.

本稿では, 我々がこれまでに提案してきた

- 対話エージェントの定量的・客観的な性能評価法
- 対話エージェントのための統計的発話選択法

の2つを統合し, ユーザとの対話から学習を行う, 新たな非タスク指向型対話エージェントの設計について述べる.

本研究で提案する対話エージェントは, 統計的発話選択法によって, 予め作成しておいた発話集から適切な発話を選択することによって対話を進める. 対話エージェントは, ユーザと対話を行った後, その対話ログを, 対話評価法を用いて自動的に評価する. そして, それを統計的発話選択法の学習データとして用いる. これにより, ユーザと対話を行えば行うほど性能を向上させることが可能な対話エージェントの実現が期待できる.

なお、本研究で扱う対話はテキスト対話に限定する。これは問題を言語処理に特化するためである。

2. 非タスク指向型対話エージェント評価法

本章では非タスク指向型対話エージェントの定量的・客観的な評価法[1]について述べる。

非タスク指向型対話エージェントの評価尺度としては、対話エージェントの発話の意味的な正しさや、対話の自然さ、対話の楽しさ等、様々な尺度が考えられる。その中で本研究では、発話の意味的な正しさに着目する。発話の意味的な正しさとは、「ある発話が直前の発話からみても、前後の文脈からみても論理的に矛盾がなく、意味的に自然であること」と定義する。

本研究では意味的な正しさを評価するために、語彙的結束性に着目した。語彙的結束性は意味的な繋がりを持つ語の出現により、テキストの意味的なつながりを明示する表層的な情報である。語彙的結束性は無作為に文を繋げたものではない、意味のあるテキストに出現する特性である。

そこで本研究では、対話において語彙的結束性を有する部分は意味のある、意味的に正しいやりとりが行われていると仮定し、その仮定に基づいて、対話エージェントがどの程度の割合で意味的に正しい発話を生成可能かという性能評価法を提案する。つまり、対話エージェントのある発話が前の文脈と語彙的結束性を有する場合、その発話は前の文脈を踏まえた上での適切な発話であると考えられる。また、その発話に対する返答を含む後の文脈と語彙的結束性を有する場合、対話エージェントの発話が人間の話者にとって理解可能なものであり、何らかの事柄についての対話が継続されたということであるので、発話は意味的に正しいと考えられる。図1に評価例を示した。語彙的結束性の判定に必要な意味的な繋がりを持つ語は、Web 日本語 N グラム[2]から抽出した。

本手法を用いて、2種類の対話エージェント A, B と、参考として人間の性能評価実験を行った。事前に行った人手による評価では、人間が最も性能が高く、次に A, 最も性能が悪かったのは B であった。評価実験の結果、評価法による評価値も、人手による評価と全く同じ傾向を示したことから、本手法の妥当性が確認できた。また、人手と評価法による評価値の相関は 0.50 であり、統計的に有意な相関が確認された。

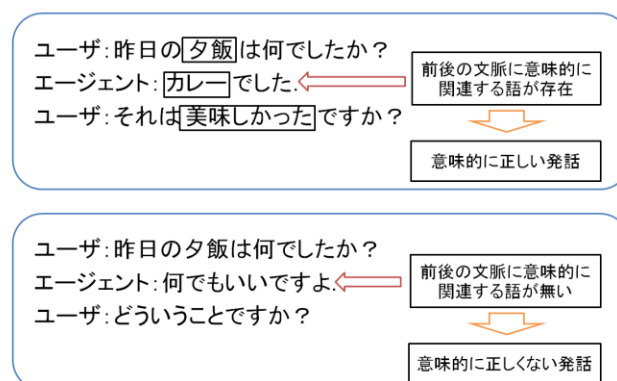


図1 非タスク指向型対話エージェント評価法による評価例

3. 統計的発話選択法

本章では、統計的発話選択法[3]について述べる。

優秀な対談番組の司会者は、対話相手に関する情報を基に質問を練り上げる等、周到に準備して対話に臨んでいる。本手法は、対話相手の情報から大量の発話集を用意して対話することにより、このような司会者をコンピュータ上に実現することを目指したものである。このような対話エージェントを実現するためには、大量の発話候補(スクリプト集)から最適な応答の自動選択する手法を確立する必要がある。なお、スクリプト集は予め用意されているものとする。

本手法では、対話の1時点の状態 s 、発話候補の集合を $A=\{a_1, a_2, \dots, a_m\}$ と定義する。対話の状態集合を S とし、対話の1状態 $s(s \in S)$ と $a(a \in A)$ に特定の処理を施すことによって生成される n 次元の特徴ベクトルを

$$\Phi(s, a) = (x_1(s, a), x_2(s, a), \dots, x_n(s, a))$$

とする。ここで、 $x_j(s, a)$ は特徴量を表す。例えば s の最後の発話と a に注目し、特定の単語または品詞、さらにこれらの組み合わせの有無などを $\{0,1\}$ の特徴量とすることができる。

さらに、特徴ベクトルの評価値を返す関数 $f(\Phi(s, a))$ を定める。ここで、 f は $x_j(s, a)$ に関する線形な関数とし、以下のように表す。

$$f(\Phi(s, a)) = \sum_{j=1}^n w_j x_j(s, a)$$

ここで、 w_j は $x_j(s, a)$ に対する重みを表すパラメータであり、パラメータベクトル $\mathbf{w}$ として以下のように表す。

$$\mathbf{w} = \{w_1, w_2, \dots, w_n\}$$

評価関数 f を用いて、対話の状態 s に対する最適な発話 $\hat{a}$ は、以下のように求めることができる。

$$\hat{a} = \arg \max_{a \in A} f(\Phi(s, a))$$

最適なパラメータは、最大エントロピー法などで求めることができる。

4. ユーザとの対話から学習する対話エージェントの設計

ユーザと対話を繰り返すことで性能が向上する対話エージェントを実現するため、エージェント設計におけるフレームワークを提案し、それにこれまで述べた手法を応用する。このフレームワークは、①発話選択、②発話評価の2つの機能から成り立つ。概念図を図2に示す。①発話選択は、スクリプト集の中から、状況ごとに適した発話を選択し、実際のエージェントの発話として出力する機能であり、ここに統計的発話選択法を適用する。②発話評価は①で出力された発話が、適切な発話であったかを対話終了時に評価し、評価結果を①発話選択の新たな学習データとしてフィードバックする機能である。2章で説明した非タスク指向型対話評価法は、発話ごとの正誤判定が可能であることから、ここに適用可能である。この機能により、ユーザとの対話から学習が可能となる。

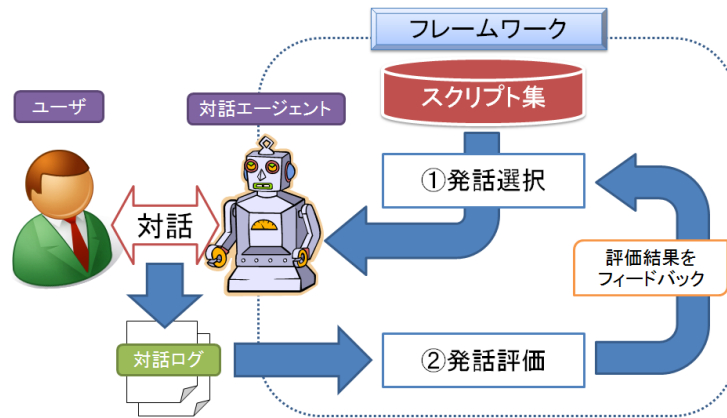


図 2 フレームワーク概念図

本研究で提案するフレームワークは、各機能の中身を入れ替えたり、付け加えたりすることで、性能を向上させることが可能である。例えば、上で述べた実装では、仮に全てが成功したとしても、対話が破綻しないというだけで、楽しい対話や盛り上がる対話がうまく出来るようにはならない。しかし、フレームワークの中身を入れ替えることにより、それは実現可能である。つまり、盛り上がり进行评估できる手法を新たに考案し、それを②発話評価に適用すればよい。盛り上がりの評価結果が①発話選択にフィードバックされ、盛り上がりを考慮した発話選択が可能となる。また、フレームワーク中の1つの機能に、2つ以上の要素を共存させることもできる。対話が盛り上がるためには、対話が破綻しないことは最低条件であるので、①発話選択に、意味的に正しい発話を選択する手法と、盛り上がる発話を選択する手法を共存させ、それらが選択した発話の共通項をとればよい。これで対話が破綻せず盛り上がる対話が可能となる。

5. おわりに

我々がこれまでに提案した手法である、非タスク指向型対話エージェントに対する定量的・客観的な評価法と、統計的発話選択法を紹介した。そして、その2つを応用した、ユーザと対話を繰り返すことで性能が向上する対話エージェントの設計について述べた。

今後の課題は、実際に対話エージェントを実装することである。ただし、実際にユーザと対話を行う際にはリアルタイム性が大きな問題になるため、それを解決する必要がある。例えば、統計的発話選択法では、スクリプト集のサイズに比例して、計算時間が大きくなる。そこで、複数のスクリプト集を用意し、状況ごとにそれを切り替えるといった前処理の検討を行う必要がある。

参考文献

- [1] 稲葉通将, 鳥海不二夫, 石井健一郎. 語の共起情報を用いた非タスク指向型対話エージェントの評価 情報処理学会研究報告 Vol.2009-MPS-75 No.18(2009)
- [2] 工藤拓, 賀沢秀人. Web 日本語 N グラム第1版. 言語資源協会発行.
- [3] Naoki Isomura, Fujio Toriumi, Kenichiro Ishii. Statistical Utterance Selection using Word Co-occurrence for a Dialogue Agent Springer Berlin / Heidelberg, Lecture Notes in Computer Science, Vol.5925/2009 pp.68-79(2009)