

音声対話システムの協調的応答に関する一評価

6H-5

伊藤敏彦 中川聖一

豊橋技術科学大学 情報工学系

1 はじめに

自然言語による対話システムにおいては、システムがユーザと協調的に対話を進めていくことは重要である[1]。データベース検索における協調的応答生成に関しては質問の答以外に付加的な情報を与えたり、失敗した質問に対する理由や代案を提示するものが多い[2]。例えば、ユーザの質問文に検索に必要な情報が含まれていなかったり、検索結果の数が多い場合などはユーザへの質問を行ったり、ユーザの望む検索結果が得られなかった場合、それに代わる代案を提供する。このようなユーザへの協調的応答によってユーザにかかる負担や不安を軽減することを我々は試みている[3]。

本稿では、我々が協調的応答生成に関して改良した音声対話システムについて、「システムの使い勝手の良さ」、「協調的応答」に着目して行なった評価実験について述べる。

2 協調的応答生成システム[3]

ユーザの発話した発話文は言語理解システムによって、等価な意味ネットワーク(意味表現)に変換され、最初に対話制御部に入力される。

対話制御部は対話の流れの決定、文脈情報の管理、必要な情報(条件)の質問などを行なっている。対話制御部は対話の流れを決定するために、ユーザの対話の意図(焦点)を抽出する意図解析部にユーザの発話意味ネットワークを送る。意図解析部は入力されたユーザの発話意味ネットワークからユーザの対話の意図(焦点)とユーザの提示した検索条件を抽出する[3]。そして対話制御部は獲得したユーザの対話の意図(焦点)から対話の流れを決定し、これまでの発話から獲得された情報で現在の発話にも移行できる情報(文脈的な情報)があればそれを文脈情報として利用する。

次に知識データベースを検索するための問題解決器にはユーザの発話意味ネットワークと先ほど述べた文脈情報を入力する。問題解決器はそれらのデータをもとに知識データベースを検索する。この時、問題解決器はユーザの発話意味ネットワークに含まれている検索条件と文脈情報を活用して検索するが、もし検索結果が得られなかった場合は検索条件の一部を同じ概念の検索条件に変更し検索をやり直す。そこで検索結果が得られたなら、代案としてその検索結果を対話制御部に送る。

対話制御部は獲得された検索結果を調べ、検索結果の数が多い場合には対話制御部はユーザへ検索結果選択のための質問を行なう。その時の質問はユーザの対

話の意図(焦点)に沿ったまだ獲得されていない検索条件に関して行なう。検索結果や代案のデータ数が適当であるなら検索された情報をユーザへ提示するために、対話制御部は応答文生成部へユーザの発話意味ネットワークと検索結果(代案結果も含む)を送る。

応答文生成部では、入力された発話意味ネットワークと検索結果からどのような形で応答すれば良いかを決定し、それに従い応答文意味ネットワークを形成する。それから応答文意味ネットワークを通常の文字列に変換し、ユーザへの応答として音声合成で出力する。

3 言語処理部改良に関する評価実験について

3.1 実験の形式・条件

実験は、被験者らに「1泊2日の富士山周辺への研究室旅行」を、システムとの対話で得られる情報をもとに計画してもらうという形式で行なった。決定してもらう内容は、1、2日目の目的地とそこでのプラン、及び、宿泊先の施設の所在、種類、料金、宿泊施設名の計8項目である。被験者は、音声対話システムに関して全く知識のない本大学生(学部2~3年)6人であり、特にシステムの使用法は教示せず、音声認識モデルの話者適応化用に発声してもらったタスクに関する50文を発声例として示した。さらに別の実験で同一のシステムを一時間程度使用してもらっている。

実験では以下の2種類のシステムで計画を立ててもらった。

1. 従来の対話システム[4](system1)
2. 言語処理部(意味理解部、応答生成部)改良後の対話システム(system2)

system2ではsystem1と比べ、疑問文以外の願望・依頼といった発話や副詞や形容詞を含んだ文も正しく意味理解できよう、さらに指示詞データベースの使用により、応答文や自分の発話に対する指示詞の使用もある程度正しく処理できるよう意味理解部の改良が行なわれている。

音声認識システムはどちらも同一のものを使用している。実験に使用した認識条件はケプストラム、回帰係数使用、ビーム幅5である(認識を高速に行なうためにビーム幅を狭くし、また音声の継続時間分布制御も使用していない)。またHMMには研究室内の雑音環境に適応化した不特定話者モデルを、実験前に被験者に発話してもらった富士山観光案内に関する適応化文50文で話者適応化し、使用している。音声認識用文法は単語数275、パープレキシティ約100、である。自然な発話を受理するように文法にロバスト性をもたしたためパープレキシティは非常に大きくなった。しかしながら、音声認識処理に有する時間は発話時間を含

Evaluations of cooperative responses in spoken dialog system

Toshihiko Ito, Seiichi Nakagawa

Toyohashi University of Technology, Department of Information and Computer Sciences

め実時間の1.5~2倍程度(HP-C160使用)で、比較的時間的にはスムーズな対話が続行できるようになった。

実験の手順としては、最初に実験の内容と決定した項目を記入してもらう紙を渡し、熟読してもらった。その後、system1とsystem2のそれぞれとの対話(最初に対話してもらうシステムは被験者毎に変更)で旅行計画を2案立ててもらった。実験終了後、使用感や応答に関するアンケートに答えてもらった。

3.2 実験の結果

まず、この6人の項目達成率(必要な8項目を決定できた割合)は、system1で平均94%、system2で平均100%であった。その理由としてsystem1の対話システムでは宿泊施設の料金に関する情報を聞き出すことができなかつた人が多かつたことによる。全発話数に関してはどちらのシステムも大きな差はでなかつた。system1(103文)とsystem2(101文)での発話数に大きな差がでなかつた理由として、第一にはsystem2の場合、システムの質問(9回)に回答する発話数が含まれていることがある。第二にはあまりにタスクが簡単すぎたという理由も考えられる。しかし一項目当りに必要な発話数を見るとsystem1で2.3、system2で2.1であり、改良による労力軽減の効果が現れているとは言える。

次に表1にシステムの評価結果を示す。正解音声認識率は正しく認識された割合を示している。認識文理解率は、(終)助詞落ち・誤りなどを許した場合の音声認識率である。正解意味理解率は言語理解部で正しい意味表現に変換できた割合である。情報提供率はシステムが正しい情報を提供した割合である。システム質問率はシステムが質問をした割合、代案提供率はシステムが代案を提示した割合である。正解応答率は情報提供率、システム質問率、代案提供率の他に意味表現作成に失敗した割合(システムは発話理解に失敗したことを応答する)を加えたものである。情報検索失敗率は、データベースの不備などによって応答に失敗した割合である。

表1: 評価結果

評価	system1	system2
全発話数	103文	101文
正解音声認識率	32文(31.0%)	21文(20.8%)
認識文理解率	83文(80.6%)	56文(55.4%)
正解意味理解率	63文(61.2%)	56文(55.4%)
正解応答率	63文(61.2%)	81文(80.2%)
情報提供率	49文(47.6%)	43文(42.6%)
システム質問率	0文(0%)	9文(9.0%)
代案提供率	0文(0%)	3文(3.0%)
情報検索失敗率	14文(13.6%)	4文(4.0%)

正解音声認識率、認識文理解率は発話された文内容の豊富さの違いから、同じ音声認識システムを使用しているにも関わらずシステムによって大きな差がでている。そこで認識文理解率に対する正解意味理解率、情報提供率等を比べると、今回の改良によってかなり性能が向上していることが分かる(認識文理解率-正解

意味理解率が小さいほど意味理解部の性能が向上、正解応答率-正解意味理解率が大きいほど応答生成部の性能は向上)。また、情報検索失敗率も減少していることから応答生成部の改良の効果が表れていることが分かる。しかし、システムの質問に対するユーザの応答が誤認識されるとシステムは誤った情報を提供したり、発話の誤認識によって文脈知識データが誤って更新されたりすることがあった。これらの失敗に対する対応(ユーザへの確認等)が必要であると考えられる。

音声認識率100%として被験者の発話を書き起こしたものを入力(テキスト入力)とした場合の評価では、system2は発話された文の種類が豊富なものにも関わらず、正解意味理解率(system1:70.9%,system2:89.1%)、正解応答率(system1:67.0%,system2:86.1%)共にsystem1よりsystem2の方が大きく上回っている。次にシステムがユーザにタスク達成のために有益な何らかの情報を提供する割合がsystem1の57.3%からsystem2の68.4%と大きく上昇している。この割合の上昇は今回改良した協調的応答によるもので、タスク達成に費やす労力を大きく軽減した。

3.3 使用感に関するアンケート

実験終了後、言語処理部に関するアンケートを被験者に書いてもらい使用感などを調査した。計画の立てやすさ、システムの便利さに関して、質問や代案を提示したり、処理できる文の種類が豊富なsystem2の評価が高かつた。

次にそれぞれのシステムでうまくいった対話例を書き起こして示し、システムの便利さ、システムの自然さ、次発話のし易さの項目を質問した場合も、全員system2の方を支持した。

これらのアンケートの結果からも、今回の改良による労力軽減の効果はあったと考えられる。

4 おわりに

本研究では、ユーザに協調的に応答する言語処理部の改良と評価実験を行ない、言語処理部の改良によって、タスク達成への労力の軽減が見られた。

今後は現在のシステムに対して、評価実験によって明らかになったシステム全体の不備な点を改良していく予定である。

参考文献

- [1] Kaplan, S.J.: Cooperative Responses from a Portable Natural Language Database Query System, Brady, M. and Berwick R.C. (eds.), Computational Models of Discourse, pp.167-208, MIT Press(1983).
- [2] Webber, B.: Question Answering, Shapiro, S.C.(ed.), Encyclopedia of artificial intelligence, pp.814-822, New York: Wiley(1987)
- [3] 伊藤, 中川: 「音声対話システムにおける協調的応答」, 情報処理学会, 音声言語情報処理研究会報告, 96-SLP-10, pp.105-110 (1996)
- [4] 山本, 伊藤, 肥田野, 中川: 「人間の理解手法を用いたロボトな音声対話システム」, 情報処理学会論文誌, VOL.37, No.4, (1996.4).