

文脈を理解して話す雑談対話システム

人にとって何気ない雑談であっても、同じような雑談を対話システムが行うにはたくさんの難しさがあります。近年の雑談対話の研究によって、雑談対話の最初の課題であった幅広い話題に対する応答ができるようになってきました。しかしながら、文脈をふまえた応答は難しく、「このシステム分かってない」とユーザーに感じさせることが多々ありました。本稿では、文脈を理解して話す雑談対話システムの取り組みについて紹介します。

なりまつ ひろみ すぎやま ひろあき
成松 宏美 / 杉山 弘晃
 みずかみ まさひろ ありもと つねひろ
水上 雅博 / 有本 庸浩
 みやざき のぼる
宮崎 昇

NTTコミュニケーション科学基礎研究所

雑談対話システムの実現をめざして

マイデイズをはじめとするスマートフォン上のエージェントやAIスピーカなどの普及に伴い、人と機械の対話が増えてきています。現在商用として使われている対話システムの多くは、主に、「Aさんに電話して」「今日の天気を教えて」、などのタスク実行を目的としています。雑談ができる対話システムにも期待が高まっています。雑談をすることには多くの効果があるといわれており、記憶の整理や人のコミュニケーションスキルの向上にも役立つと期待されています。NTTコミュニケーション科学基礎研究所では、早くから「雑談」に着目をして研究を行ってきました。

雑談を行う対話システムでは、先述のタスクを行う対話システムと異なり、ユーザーの発話の話題が幅広く、ユーザーの発話を事前に想定して設計することができないという難しさがあります。例えば、レストラン予約などのタスクを行う場合には、予約日時や予約者の名前と電話番号など予約をするうえで必要な情報は決まっており、ユーザーの発話に含まれる情報をあらかじめ想定することが可能です。一方で、

雑談対話では、ユーザーの発話に含まれる情報を想定しておくことは難しく、ユーザーのあらゆる発話に対して適切に応答することが困難でした。

私たちの所属する研究グループでは、幅広い話題に応答できるようにするために、質問と応答、発話と応答、発話と質問などのさまざまな応答ペアを多量に用意し、機械学習の訓練データとして用いたり、文間類似度により発話選択をしたりする手法に取り組んできました。これまでの研究成果により、一問一答ベースでは、ユーザーの発話に対してある程度近い応答はできるようになってきました。

しかしながら、人と同じように対話できる相手になるためには、相手の発話に合わせた適切な応答や、文脈に整合した適切な応答ができなければなりません。本稿では、これらの問題に対する、私たちの最新の研究成果を紹介します。

一問一答ベースの対話システムの問題点

従来の一問一答ベースの対話システムでは、直前のユーザー発話に対して、多量に用意した発話例から近い応答を返すという戦略でした⁽¹⁾。そのため、

違和感のある発話や、それ以前の対話を聞いていないような発話をする 경우가あり、数分の対話でも「分かってくれないな」とユーザーに感じさせてしまうという課題がありました。例えば、従来のシステムでは、ユーザーとの対話がしばしば次のようになります(図1)。この対話では、初めにユーザーが「夏休みにたこ焼き食べた」と言ったのにもかかわらず、4発話目にシステムが「いつ行ったんですか?」と聞いてしまい、ユーザーに「さっき夏休みって言ったのに、分かってないな…」と思わせています。また、「夏休みは避暑地がいい」という発話も、これまでのたこ焼きの話題から急に離れてしまうため、システムがなぜその発話をしたのか分からず、ユーザーを困惑させています。こうした、①直前までの対話の内容と整合しない発話や、②根拠のない発話によって、ユーザーは、「システムは分かってないな」や「このシステムは何が言いたいかわからない」と感じてしまい、システムと対話するのを諦めてしまう原因となります。これでは、コミュニケーションする相手として良いといえないどころか、ユーザーに「対話できる相手」として認めてもらえず、使ってもらえなくなります。

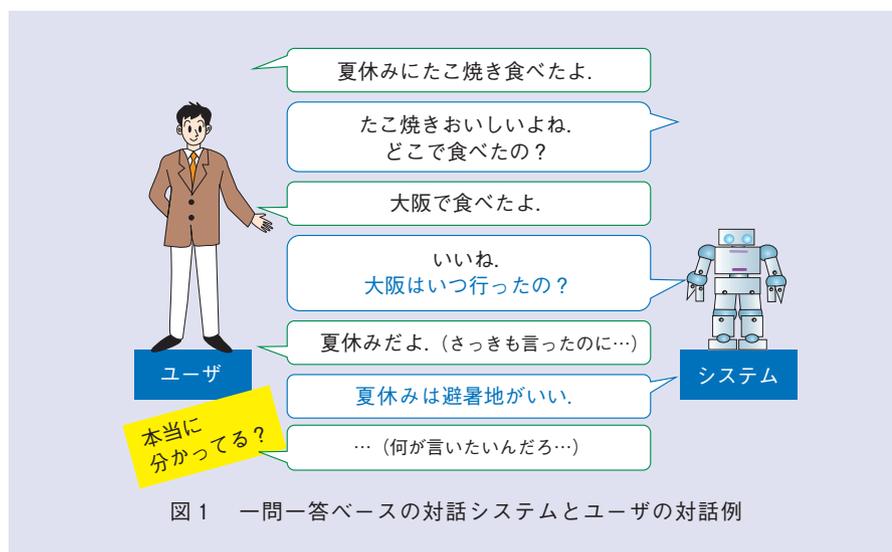


図1 一問一答ベースの対話システムとユーザの対話例

「対話できる相手」になるために

「対話できる相手」として認めてもらうためには、少なくとも前述した問題を解消する必要があります。これは、心理学者のグライスが提唱した対話の成立条件⁽²⁾でも言われています。関連のないことを言うてはいけない（関連性の公準）、根拠のない適当なことを言うてはいけない（質の公準）が挙げられており、①文脈に整合しない発話や、②根拠のない発話は対話破綻を導くとされています。そこで、2つの問題に対して、文脈に整合した発話および根拠のある発話を行うために、私たちは、「文脈の理解」および2つの発話生成「文脈に整合した発話生成」「根拠に基づく発話生成」に取り組みました。以降では、それぞれに対する私たちのアプローチを紹介します。

文脈の理解

文脈となる現時点以前の対話をどのように理解し、文脈情報として保持し

たら良いのでしょうか？

私たちは、ユーザの体験は5W1H+感想で表せることが多い点に着目し、5W1H+感想の情報を文脈として理解し、利用する方法を考えました。5W1Hのフレームは非常にシンプルですが、人とのコミュニケーションやカウンセリングの対話においても5W1Hの質問をしていく戦略がとられており、いろいろなシチュエーションで共通して使えるフレームです。例えば、旅行の話をするときにも、美味しいものを食べた話をするときにも、「どこに行ったの?」「いつ行ったの?」「どうだった?」などは自然な流れで出てくる質問だと思えます。

では、5W1H+感想の情報はどのように理解したら良いのでしょうか？

5W1Hのうち、時間や場所に関する情報は、従来から固有表現抽出技術で抽出の対象とされてきていました。例えば、「昨日、東京に行ったよ」という文が与えられた場合、「昨日」は時間、「東京」は場所の固有表現として抽出

されます。主に、固有名詞や日付・時間に特有の表現などが抽出の対象となっています。でも、人の自由発話に含まれる時間や場所の情報は、これだけで足りるのでしょうか。私たちが、実際の人どうしの雑談対話を収集し、人が「時間」や「場所」として理解するフレーズを調べた結果、固有表現だけではないフレーズが多くを占めていることが分かりました（場所フレーズの場合は約7割）。

そこで、ユーザの自由発話に含まれる5W1H+感想に該当するフレーズを抽出する、フレーズ抽出器を構築しました。フレーズ抽出器には、固有表現抽出でも有効とされている系列ラベリング手法を用います。代表的な手法は、CRF⁽³⁾ですが近年ではディープニューラルネットワークを用いた手法も提案されています。これらの手法に対して、人どうしの雑談対話中で5W1H+感想の各項目として人が理解するフレーズに対して人手でアノテーションを行い、学習させました⁽⁴⁾。

この結果、正式名称でなくとも「京都駅近くの公園」などが場所としてとれたり、「たこ焼きを食べた」をWhat項目としてとれたり、文脈に必要な情報をユーザの発話から新たに抽出できるようになりました。場所フレーズを例に、従来の固有表現抽出器とその抽出結果を比較すると、表のように、固有名詞を含むフレーズや、一般名詞などが抽出できるようになっていることが分かります。本技術を用いて、対話中に現れたユーザの5W1H+感想の情報を埋めていくことで、文脈を理解することができるようになりました。

表 ユーザ発話に含まれる場所フレーズの例と抽出結果の比較

ユーザ発話 (赤字:場所フレーズ)	従来技術の抽出結果	本技術の抽出結果
夏休みにイタリアに行きました。	イタリア	イタリア
京都駅近くの公園でお花見したよ。	京都駅	京都駅近くの公園
暇なときは電気屋に行きます。	(なし)	電気屋

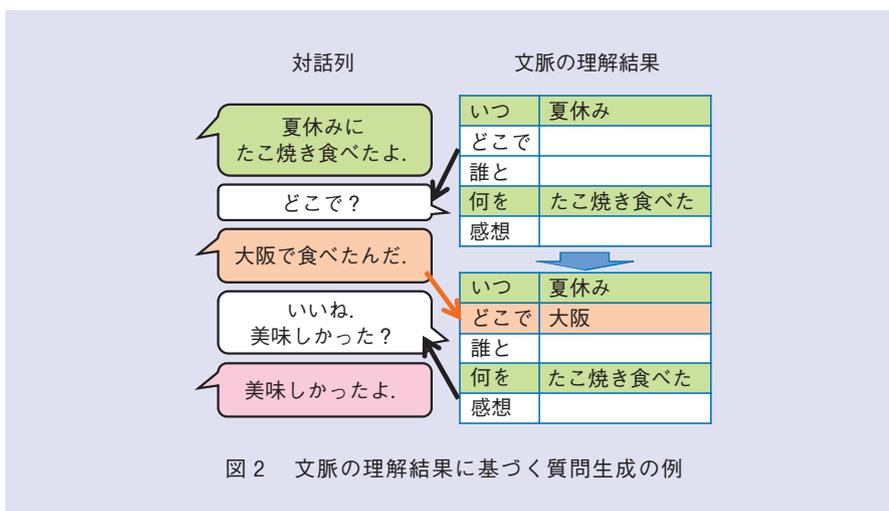
情報を付け加えます。これにより、システムがなぜその質問をしたのかをユーザに伝えることができます⁽⁵⁾。

2番目は、システムが共感や感想を述べる際に、なぜそう思うのかの根拠情報を付け加える方法です。「たこ焼きは美味しいですね」と伝える際に、「私も大阪でたこ焼き食べましたよ。熱々だったので美味しかったです」というように、なぜその感想を抱いたかの理由や根拠となる情報を付け足します。例えば、システム自身の知識として、図3のような構造を持っており、それを発話のフォーマット「私も[場所]に行って[何を]した。[感想根拠]ので[感想]」に当てはめることで、根拠を含めた発話を生成することができます。これにより、単に「たこ焼きは美味しいね」というよりも、システムが自身の体験⁽⁶⁾や知識に基づき「本当にそう思って」共感している感じを与えられます。

先述の「文脈の理解」および「文脈に整合した発話生成」と本技術を組み合わせることで、文脈に整合した発話にさらに根拠を付け加えることが可能となりました。これにより、図4のような「分かってくれる」対話が可能となります。

今後の展開

今回の取り組みにより、対話システムは、文脈を理解し、文脈に合わせた質問や根拠のある発話の生成ができるようになりました。これは、「このシステム分かってないな」とユーザに思われていた対話システムを、「理解して」対話できるシステムに変える大き



文脈に整合した発話生成

前述の文脈理解の結果を用いることで、文脈に整合した質問や発話の生成が容易になります。例えば、5W1H+感想の情報を引き出す対話であれば、従来の技術で起きていた直前にユーザが話したことを質問してしまうことも抑制できるようになります(図2)。また、対話中に「夏休みに旅行に行った」と「大阪観光した」という発話があれば、「時間:夏, 場所:大阪」という情報から2つを関連付けて、「海遊館とか行った?」という関連する質問や、「夏の大阪は暑いですよ」という文脈に整合した発話を生成することができます。これは、従来手法で直前のユーザ発話(大阪観光した)のみから生成され得る「大阪には道頓堀が

ありますね」という発話よりも、より文脈に整合した発話となっていることが分かります。

根拠に基づく発話生成

文脈に整合するだけでは、根拠のある発話になることは保証されません。そこで、私たちは、システムが発話を行う際にその根拠となるような情報を付け加えて提示するアプローチを提案しました。ここでは質問と共感発話を例にアプローチを紹介します。

1番目は、システムが質問をする際に、なぜ質問したかの理由を加える方法です。システムが「夏でも楽しめますか?」と質問をする際には、「私は夏休みに行けたらと思っているので、参考にさせてもらえたらな」と思って」のように、質問した理由や根拠となる

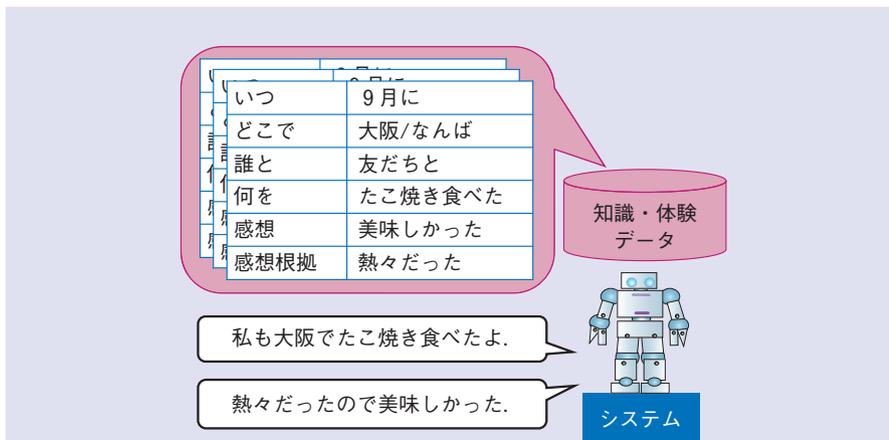


図3 システムの知識・体験データを根拠とした発話生成

- (3) J. Lafferty, A. McCallum, and F.C.N. Pereira: "Conditional random fields: Probabilistic models for segmenting and labeling sequence data," Proc. of ICML 2001, pp.282-289, June 2001.
- (4) H. Narimatsu, H. Sugiyama, and M. Mizukami: "Detecting Location-Indicating Phrases in User Utterances for Chat-Oriented Dialogue Systems," Proc. of LACATODA 2018, pp.8-13, July 2018.
- (5) 杉山・成松・水上・有本: "文脈に沿った発話理解・生成を行うドメイン特化型雑談対話システムの実験的検討," 人工知能学会 言語・音声理解と対話処理研究会 (SLUD) 第84回研究会 (第9回対話システムシンポジウム), 2018.
- (6) M. Mizukami, H. Sugiyama, and H. Narimatsu: "Event Data Collection for Recent Personal Questions," Proc. of LACATODA 2018, July 2018.

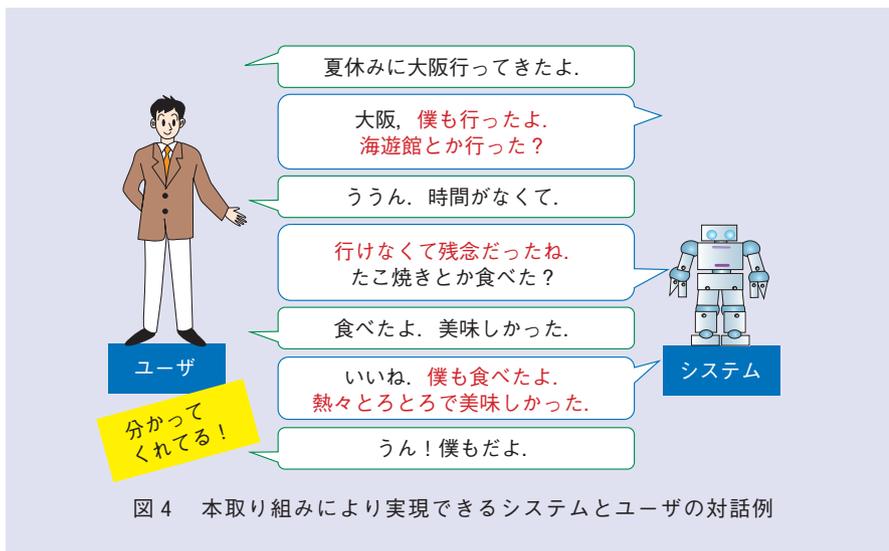


図4 本取り組みにより実現できるシステムとユーザの対話例



(左から) 杉山 弘晃/ 有本 庸浩/
成松 宏美/ 水上 雅博/
宮崎 昇

人は相手の理解状況に合わせて情報提示の方法を変え、円滑なコミュニケーションをしています。今後も、システム自身の理解力向上に取り組みつつ、対話相手の理解度に応じた柔軟な応答も検討していきます。

◆問い合わせ先

NTTコミュニケーション科学基礎研究所
協創情報研究部
インタラクション対話研究グループ
TEL 0774-93-5020
FAX 0774-93-5026
E-mail cs-liaison-ml@hco.ntt.co.jp

な一歩であると考えます。自分の話を理解して話してくれるシステムがいたら、いろいろな話をしたり相談したりするようになるでしょう。コミュニケーショントレーニングや相談などさまざまなシーンでの対話システムの活用促進にもつながると考えています。

しかしながら、これらの実現には、対話の流れをうまく設計し、システムの知識となるデータを人手で作成するなどの、手間とノウハウが必要であり、

誰でも簡単に同様のシステムがつけられる状況には至っていません。

今後は、今回人手で作成したようなデータをWebや人との対話を通して自動的に作成する手法にも取り組んでいきます。

■参考文献

- (1) 杉山・東中・目黒: "気軽に雑談できるシステムの実現をめざして," NTT技術ジャーナル, Vol.28, No.9, pp.16-20, 2016.
- (2) H.P. Grice: "Logic and conversation," Syntax and Semantics, Vol.3, Speech Acts, pp.41-58, 1975.