

2. エッセイ集

## 8 対話システム研究の動向

### —対話システムは次世代のインタフェースになるか—

東中竜一郎 (NTTメディアインテリジェンス研究所)

#### 対話システムの利用シーンの拡大

対話システムは対話によって所定のタスクを遂行するタスク指向型対話システムと、いわゆるおしゃべりを行う非タスク指向型対話システム(雑談対話システムとも呼ぶ)に分けられる<sup>1)</sup>。

タスク指向型対話であれば、一定程度のタスクを対話によって行うことができるようになってきている。また、非タスク指向型対話についても、技術が成熟していないとはいえ、簡単な会話であれば可能になってきた。それに加え、音声認識はこのところ急速に精度が改善してきている。このような背景から、近年では多くの対話システムが商品化され、多くのユーザに使われるようになってきた。Apple社のSiriをはじめとし、NTTドコモ社の「しゃべってコンシェル」、Yahoo!の「音声アシスト」などは多くのユーザに利用されている。家庭内にも対話システムは進出している。米国ではAmazon Echoが好評であり、Google Homeも発表された。また、まだ発売はされていないが、Jiboへの期待は高い。国内でも、パーソナルロボットとして、タカラトミー社が「OHaNAS」を発売している。さらに家庭内にとどまらず、コミュニケーションロボットも出現してきている。ソフトバンク社のPepperは街中で見かけることも多くなってきた。

このように、対話システムはスマートフォン上のみならず、我々の社会に入り込もうとしている。これらの背後にあるのは、対話インタフェースへの期待であろう。これまで、キーボードやGUIによって我々はコンピュータを操作してきたが、日常生活のさまざまな場面でこれらのデバイスは必ずしも使い勝手が良いわけではない。コンピュータと対話することで、さまざまな問題解決をするようになることの利便性は高いだろう。そうなったとき、データを蓄積し、ユーザに、より深くリーチできるようにするための熾烈な競争が現在繰り広げられていると言える。なお、スマートフォン上でも、同様の理由で、チャットボットが流行の兆しを見せている。

たとえば、リクルートジョブズ社の「パン田一郎」やMicrosoft社の「りんな」「XiaoIce」「Tay」などがある。すべてを対話で解決することはまだまだできないが、自然言語による簡単なやりとりで楽しみや癒しを得たり、課題が解決できたりする経験をユーザが重ねていくことによって、今後は対話が、人間にとって、コンピュータとのやりとりの主要なメディアになっていくと思われる。

#### 雑談対話システム

国立国語研究所の調査によれば、人間の会話のうち60%程度は雑談である<sup>2)</sup>。よって、対話システムの雑談能力を高めていくことは、人間に寄り添うインタフェースにとって非常に重要である。

そもそも、人間が自然に行っている雑談であるが、機械にとっては難しい。なぜなら、雑談における話題や内容は多岐にわたり、また、タスク指向型対話システムとは違い、正解を決めることは難しく、工学的な問題として落とし込むことが困難なためである。雑談対話システムの実現方法は主として3つある。1つ目の方法は、ルールベースによる方法である。これは、ユーザの入力に対する出力を手作業によりルールとして書くという方法である。対話システムの元祖とも言えるELIZAはこの方法で実現されている。ルールを作成するコストは高いものの、質の高い発話を生成できるため、いまだに商用サービスではこの手法によるところが大きい。2つ目の方法は抽出ベースによる方法である。TwitterなどのConsumer Generated Media (CGM)上では大勢のユーザが対話を行っている。この大量のやりとりのデータから現在の対話文脈にあるものをうまく抜粋して対話に利用するという手法である。インターネット上のコンテンツであるので多くの話題について応答できるシステムが実現できるものの、CGMはノイズが多く高精度に良い発話を抽出することは困難である。3つ目の方法は、現在最も多く研究されている

もので、生成ベースの方法である。これは、抽出ベースで用いるような大量の対話データから、生成モデルを、主に深層学習（特に、encoder-decoder）の手法によって実現しようとするものである。機械翻訳の分野において、大量の対訳データから生成モデルを実現することで人手を超える性能を達成したことが契機になっていると言える。データが大量にあれば、対話の生成もできるのではないかと考えたわけである。この分野では、Facebook 社、Google 社、Microsoft 社が活発に研究発表を行い、Neural Conversational Model や Memory Networks といった新しい手法が提案されている。しかしながら、現在のところ、データが大量にあるからといって、雑談がスムーズにできるようになったというわけではない。機械翻訳では一文を一文に翻訳することを行うが、対話では、これまでの文脈を受けて何を生成するかを決めなくてはならない。これは決定的な違いであり、対話研究の難しさでもある。おそらく、深層学習の単純な適用ではブレークスルーにはつながらないだろう。これまでの対話研究で扱われてきた重要な要素、すなわち、発話行為論、計画立案・認識、共有信念、基盤化などの概念を扱えるようにしていかなければならない。また、会話を盛り上げるためにはユーザの状態やユーザの置かれた状況も理解しておく必要があり、これらの研究も並行して進めていく必要がある。

雑談対話技術の学会的な動きについても触れておくと、信号処理の国際会議である APSIPA ASC において、2年連続でチャットボットの特別セッションが開催されている。また、チャットボットの国際ワークショップ (WOCHAT) も今年から始まっており、2016年9月には2回目が開かれる。国内では、人工知能学会 SLUD 研究会における対話システムシンポジウムや人工知能学会全国大会におけるオーガナイズドセッション「知的対話システム」においても近年多くの関連発表がされている。

## 共有タスク

近年の対話技術の急速な進展は共有タスク (Shared Task) によるところも大きい。共有タスクとは、共通のデータセットについて、共通の評価尺度を用いて、複数の参加者が取り組む課題のことである。

タスク指向型対話システムにおける共有タスクで

有名なものに Dialogue State Tracking Challenge (DSTC) がある。これは、対話における話者の意図を発話から推定する共有タスクである。対象となるドメインは、バス案内、レストラン案内、観光案内などである。DSTC-1 から DSTC-3 までは人間とシステムの対話についてユーザの意図を推定する問題に取り組んでいたが、今年開催された DSTC-4 では、初めて人間同士の対話が対象となった。これは、人間同士の対話という、より難しい対象にチャレンジするとともに、世の中にある大量の人間同士の対話データ（たとえば、コールセンターのデータ）に対する処理もビジネス的に重要と認識されるようになってきたことによる。

非タスク指向型対話システムにおける共有タスクには2つある。1つ目は、対話システムシンポジウムにて開催している対話破綻検出チャレンジ<sup>3)</sup>である。これは、人間とシステムの雑談対話において、対話破綻（システムの不適切な発話によって対話が継続できなくなってしまう箇所）を検出するタスクである。2015年では6チームが参加し、その精度を競った。もう1つは、国立情報学研究所が主催する NTCIR で開催している、Short Text Conversation (STC)<sup>4)</sup>である。これは、大規模な対話のリポジトリ (STC は Weibo と Twitter を利用) があるときに、対話システムへの入力に対する応答を、リポジトリから検索するというタスクである。これは、抽出ベースの手法に関する共有タスクと言える。日中合わせて23チームが参加した。なお、2017年の冬に開催予定の STC-2 は、生成ベースの手法を中心とした共有タスクとなる予定である。深層学習の進展によって、どの程度コンピュータと雑談ができるようになるものか非常に見ものである。

### 参考文献

- 1) 中野幹生, 他: 対話システム, コロナ社 (2015).
- 2) 小磯花絵, 他: 大規模日常会話コーパスの構築に向けた取り組み—会話収録法を中心に—, SIG-SLUD, Vol. B5, No. 01 (2015): 37-42.
- 3) 東中竜一郎, et al.: 対話破綻検出チャレンジ, SIG-SLUD, Vol. B5, No. 02 (2015): 27-32.
- 4) Shang, L., et al.: Overview of the NTCIR-12 Short Text Conversation Task, Proceedings of NTCIR-12 (2016). (2016年7月19日受付)

東中竜一郎 (正会員) Higashinaka.ryuichiro@lab.ntt.co.jp

2001年慶應義塾大学大学院政策・メディア研究科修士課程、2008年博士課程修了。博士(学術)。2001年日本電信電話(株)入社。しゃべってコンシェルにおける質問応答技術や雑談対話システムの研究開発に携わる。