

特集「音声対話システムの実用化に向けて」にあたって

花沢 健

(NEC 情報・メディアプロセッシング研究所)

駒谷 和範

(名古屋大学)

香山 健太郎

(情報通信研究機構)

近年、国内外で音声インタフェースを利用した音声対話システムの実用化が進んでいる。音声インタフェースは人間にとって最も自然なインタフェースの一つであり、それを利用した音声対話技術は機器操作や情報取得などさまざまな用途に利用可能である。身体性をもつロボットとのインタフェースとしても活用されている音声対話技術およびその重要な構成要素である音声認識・音声合成技術は、過去に盛衰を繰り返してきた歴史がある。2000 年頃には、統計モデルによる音声認識の技術向上に伴い、電話によって情報取得や機器操作を行う IVR (Interactive Voice Response, 音声自動応答) を中心に実用化が盛んになった。しかしその後、携帯電話からのインターネット接続が一般的になると、情報取得のためのインタフェースとしてはキー入力と画面表示が主流となり、音声対話技術の実用化は一時沈静化している。2010 年頃からは、スマートフォンの普及と大規模データの利活用技術の発展に伴い、スマートフォン上のインタフェースあるいはアプリケーションとして利用が促進されるようになってきている。このように、情報取得あるいは機器操作を音声対話・音声応答によって効率良く行おうという流れが、現在定着しつつある。また、効率を求めるだけでなく、音声による対話の体験そのものをエンターテインメント領域などに活かそうという動きもある。

そこで今回、音声対話に関する特集を企画した。本特集では、現在の音声対話技術の研究開発動向・業界動向を俯瞰し、すでに実績がある音声対話システムを軸として、特に実運用面での困難や今後の課題を整理することを目的とした。音声対話は、音声認識・言語理解・音声合成・対話処理などの技術を高次元で組み合わせて実現するものであり、各技術が必ずしも完全ではないことから、組み合わせた場合の難易度も高い。また、音声対話システムを実環境・実現場にもって行くと、実験室環境とは異なることからさまざまな課題が噴出することもわかっている。特に音声のような物理現象をインタフェー

スとして直接扱う場合には、環境雑音の問題や入力の実用性の問題、音声認識精度・処理速度の不足などが顕著になる。これらは、これまで音声対話システムが爆発的に普及することを妨げてきた要因である。各解説においては、現在稼働している各種音声対話システムを紹介いただき、実システム稼働時の課題や、今後の技術の方向性を議論する機会としたい。

執筆いただいた内容について、以下に簡単に紹介する。

1. 「音声対話システムの進化と淘汰

—歴史と最近の技術動向—

(河原達也 (京都大学))

河原氏には、音声対話システムの歴史をひもとくところから解説いただく。音声対話システムの 20 年以上の盛衰の過程を、同じ時を刻んできている氏の 20 年の研究活動から分析し、「進化と淘汰」というやや刺激的なタイトルにて解説されている。これから音声対話の研究活動を始めようとする学生から、第一線の研究者・開発者にまで、わかりやすく整理された内容となっている。

2. 「10 年間の長期運用を支えた音声情報案内システム「たけまるくん」の技術」

(西村竜一 (和歌山大学)・原直 (岡山大学)・川波弘道 (奈良先端科学技術大学院大学)・李晃伸 (名古屋工業大学)・鹿野清宏 (奈良先端科学技術大学院大学))

西村氏には、音声情報案内システム「たけまるくん」について解説いただく。「たけまるくん」は 2002 年に生駒市北コミュニティセンターに設置された、一問一答型の情報提供エージェントである。その後順調に稼働を続け、今年 2012 年で 10 周年を迎えた。長期にわたる運用がもたらした成果は多く、関連技術研究や別エージェント (キタちゃん/キタロボ) への発展が得られている。

3. 「キャンパスの公共空間におけるユーザ参加型双方向音声案内デジタルサイネージシステム」

(大浦圭一郎・山本大介・内匠 逸・李 晃伸・徳田恵一 (名古屋工業大学))

大浦氏らには、名古屋工業大学で稼働している音声対話技術を利用したデジタルサイネージシステムについて解説いただく。音声インタラクションシステム構築ツールキットである“MMDAgent”を活用し、大学キャンパスの情報提供を行う双方向の音声案内システムを設置・運用している。対話シナリオも含めたコンテンツの制作を第三者 (例えば学内の学生) でも行えることが特長であり、コンテンツ制作や登録のインタフェースも用意した CGM 的アプローチが興味深い。

4. 「音声対話による観光案内システムの開発と多言語化—音声対話システム AssisTra の研究開発から得られた知見と課題—」

(翠 輝久・水上悦雄・堀 智織・柏岡秀紀 (情報通信研究機構))

翠氏らには、NICT でサービスされている「AssisTra」および、その英語版の開発について解説いただく。AssisTra は、スマートフォンをクライアントとし、対話システムサーバと連携して音声対話により観光案内 (観光に関する情報提供) を行うシステムである。観光案内タスクでは、ただ問われた質問に答えるだけでな

く、ユーザがもつ曖昧な要求に対して推薦も含めた柔軟な応答が求められる。本解説では、このための言語理解や対話制御について、実運用により得られた課題とともに解説されている。

5. 「実サービスにおける音声認識と自然言語インタフェース技術」

(辻野孝輔・栄藤 稔・磯田佳徳・飯塚真也 (NTT ドコモ))

辻野氏らには、NTT ドコモでサービスされている「しゃべってコンシェル」について解説いただく。「しゃべってコンシェル」は、スマートフォンをクライアントとし、クラウド連携により音声対話を行う情報コンシェルジェである。2012 年 3 月のサービス開始以降、話題となっている。強力な音声認識・言語理解により、ユーザの要求を自由な表現で受理可能であること、さまざまな要求に答えられることが特長である。本解説では、この実現を支える技術について解説されている。

以上 5 編の解説は、現行の音声対話技術・音声対話システムをマクロ (大局的) およびミクロ (局所的・具体的) に捉えるのに最適であり、研究者・開発者の方々に有益な情報となると信じる。また、今回取り上げた音声対話システムが、今後 (今度こそ) 定着・発展していくことを期待したい。