

対話システムシンポジウム10周年記念企画

対話システム研究の10年を振り返って —これまでとこれから—

SLUD研究会 主査

(株)ホンダ・リサーチ・インスティテュート・ジャパン

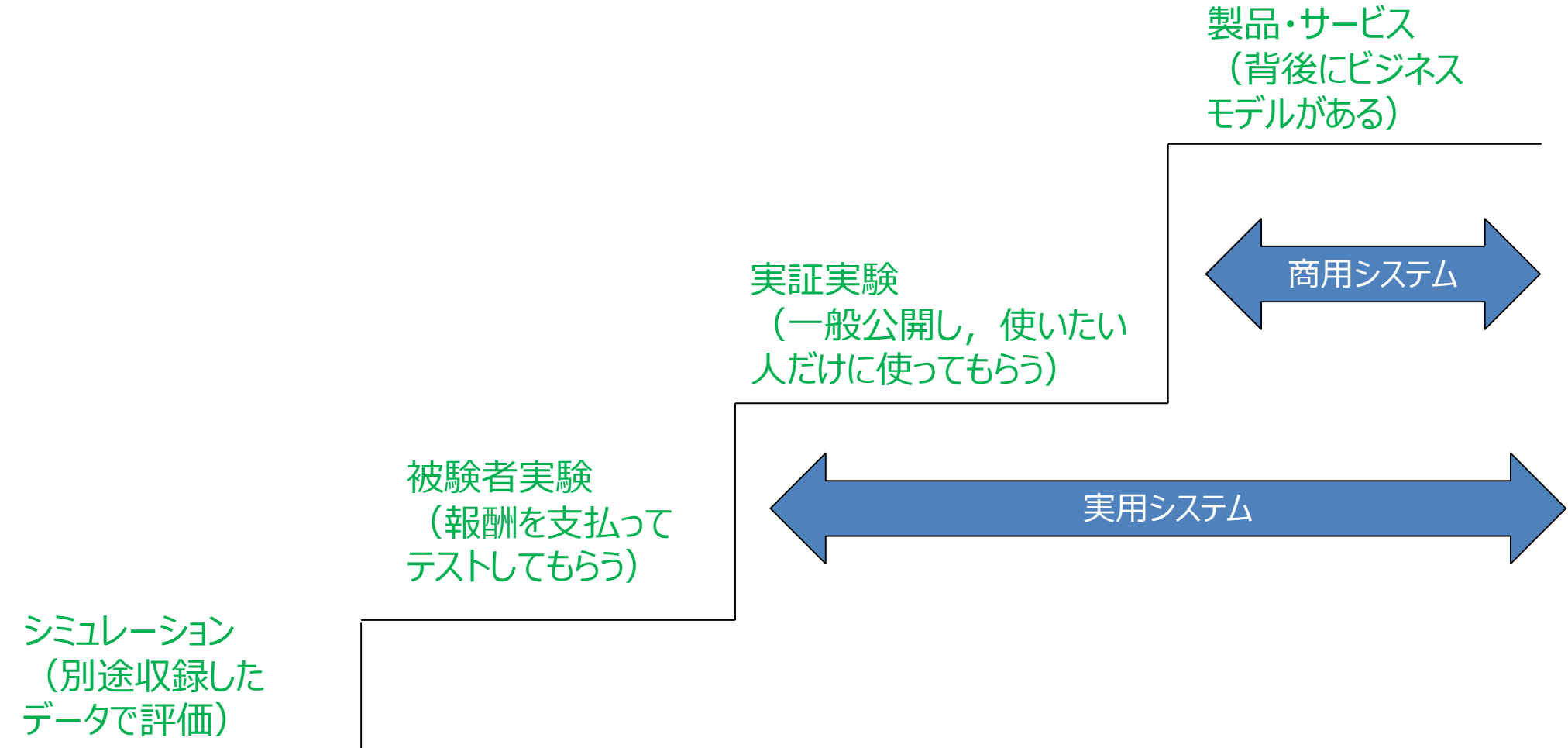
中野 幹生

nakano@jp.honda-ri.com

- 対話システムこの10年
- 対話システム研究のこれまでとこれから
 - 中野の意見 + アンケート結果

今日は「**实用・商用システムにつながる研究**」
という観点でお話します

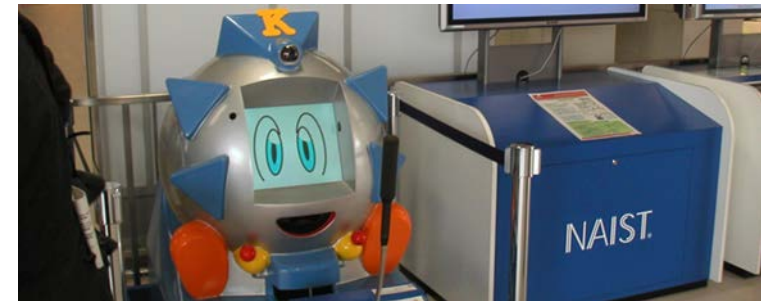
実用システム・商用システム



対話システムこの10年

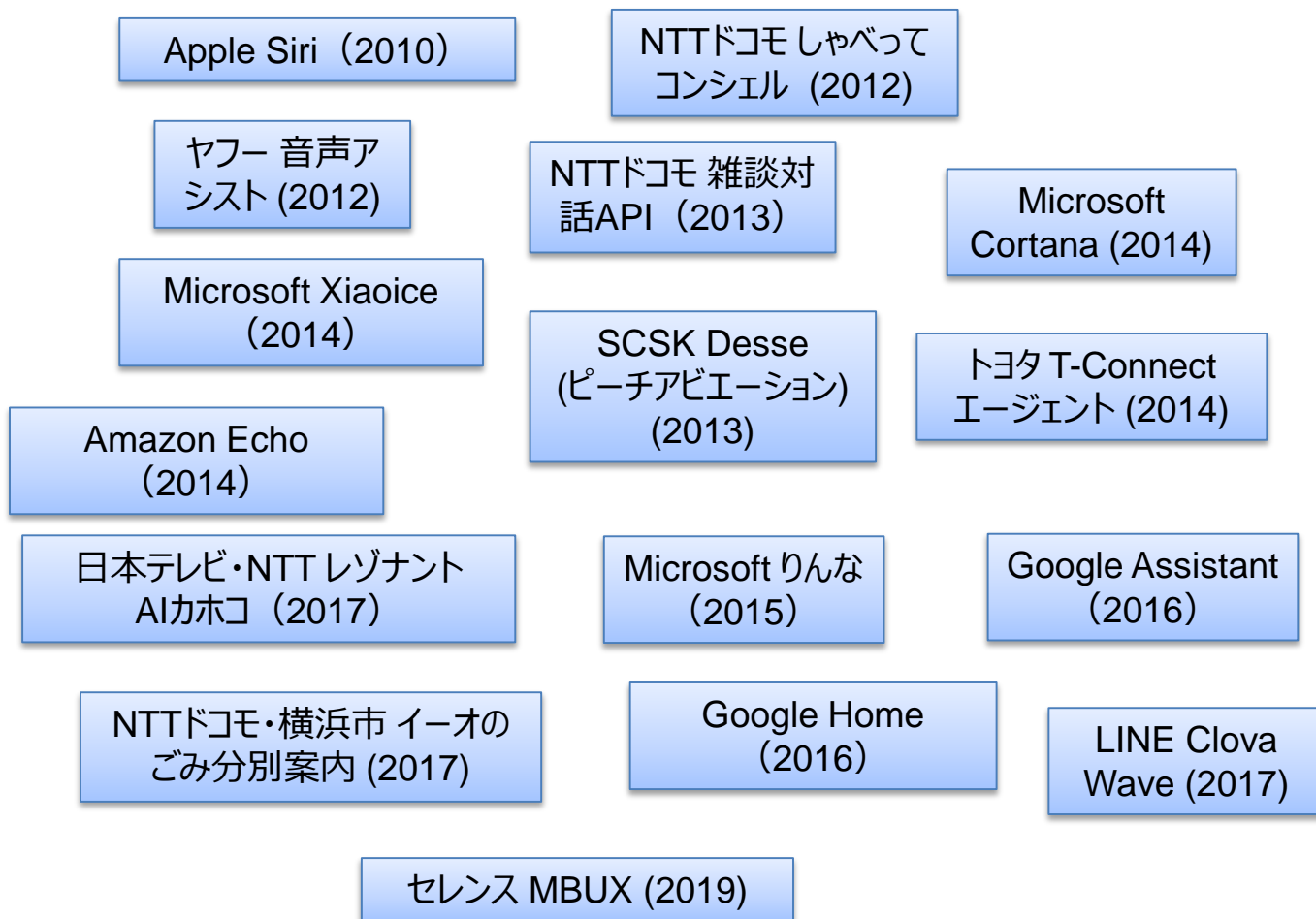
2010年以前の実用対話システムの例

- 商用システム
 - カーナビの音声対話
 - 小田急ロマンスカー予約 電話対話システム
 - IKEAのAnnaさん (web上のテキストチャットボット)
- 非商用実用システム
 - A.L.I.C.E.
 - Jabberwacky
 - NEC Papero
 - 京大 京都バス運行情報システム
 - 奈良先端大 たけまるくん, キタちゃん

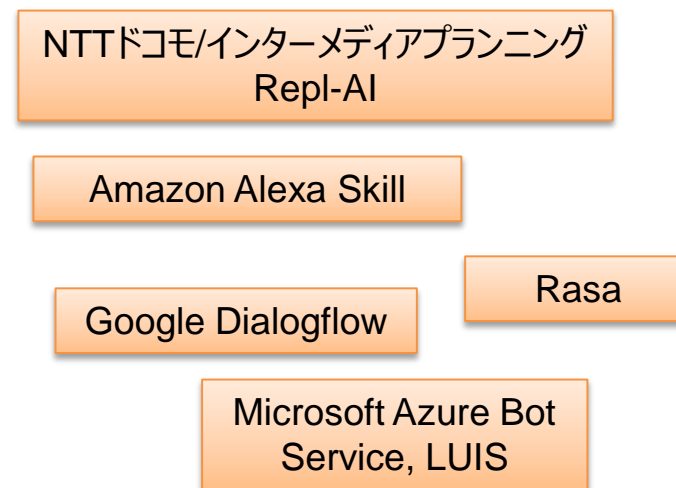


実用・商用システムこの10年

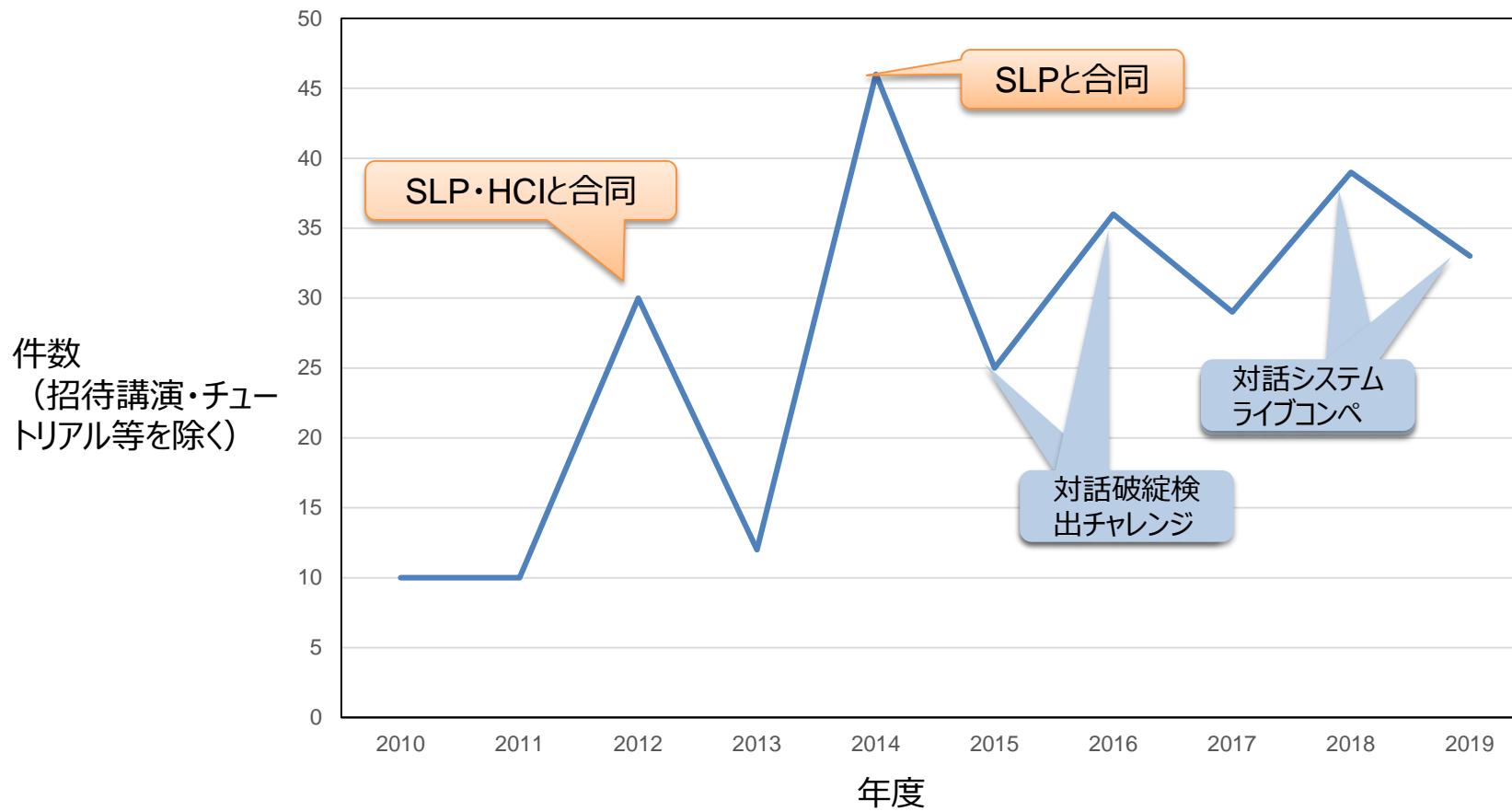
様々なシステムがサービスを開始



様々な開発ツールが登場



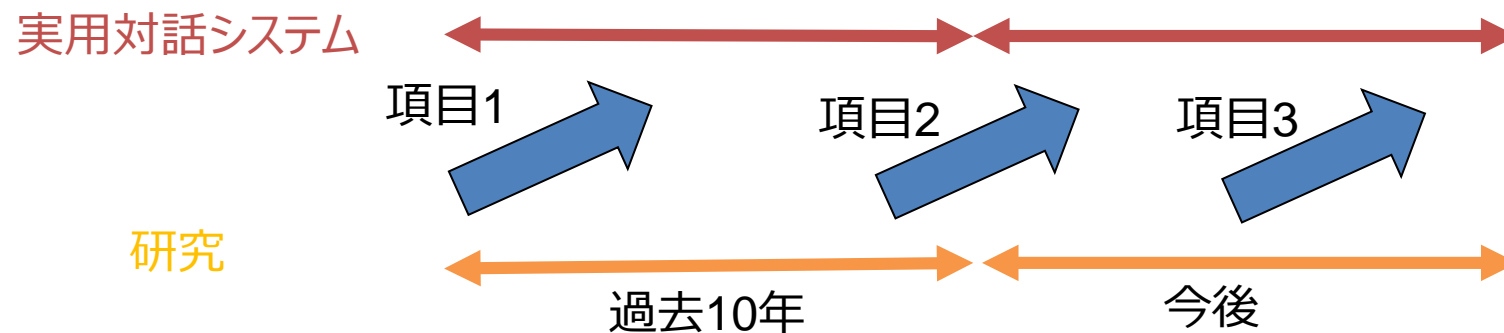
発表件数



これまでとこれからアンケート

アンケート項目

1. 過去10年(2010年-2019年)の実用対話システムの進化において重要な役割を果たした技術や知見は何でしょうか？
2. 過去10年の対話システム研究の成果で、今後の実用システム開発において重要となるとと思われるものは何でしょうか？
3. 今後の実用対話システムの発展のために重要となる研究分野は何でしょうか？
4. その他、過去10年の対話システムの研究に関して何か特筆すべきことがあればお書きください。



協力して頂いた方

- 35歳から55歳くらいの基礎研究者の皆様
 - 商用システム開発に携わっている方々
- } 計21名

【敬称略】

石川 泰 (Cerence Japan 株式会社)

吉村 健 (NTTドコモ)

Xianchao Wu (Microsoft)

徳久良子 (豊田中央研究所)

高橋 哲朗 (富士通研究所)

杉山 弘晃 (NTT)

河原 達也 (京都大学)

駒谷 和範 (大阪大学)

北岡 教英 (豊橋技術科学大学)

吉野 幸一郎 (NAIST)

岡田 将吾 (北陸先端科学技術大学院大学)

稲葉 通将 (電気通信大学)

松山 洋一 (早稲田大学)

+ 匿名希望の方 8名

どうもありがとうございました！

全回答は付録にあります。

過去10年の実用対話システムの進化において重要な役割を果たした技術や知見は何でしょうか？

講演者の意見と関連するアンケートの回答

- DNNを用いた音声認識・言語理解
- 応答抽出ベースの雑談対話手法
- ユーザスタディ・実証実験に基づく対話システム構築プロセス
- クライアント-サーバ連携に基づく対話システムアーキテクチャ
- クラウドベースの対話システム構築ツール
- クラウドソーシングによる対話知識構築・対話システム評価

- 音声認識の大語彙化（吉村）
- 音声認識の向上（石川, 徳久, 企業・研究, 企業・研究）, 杉山, 駒谷, 吉野, 岡田, 松山, 匿名, 匿名）
- 言語理解（吉野, 大学等, 岡田）

- クラウド上の計算機環境の充実（企業・研究）
- 音声認識のクラウド化（駒谷）
- クラウド上の計算機環境の充実（企業・研究）

- クラウドソーシング（稲葉, 企業・研究）

皆様の回答から

- Siri, Xiaoice, Echoなどの製品が世の中に出たこと（Wu, 企業・商用システム開発, 岡田）
- 大規模データ（石川, 杉山, 河原）
- ビジネスモデルの変化による, 利用拡大（石川）
- 特に人手不足によりボットの有用性が認められたこと（企業・商用システム開発）
- 音声合成（徳久, 大学等）
- POMDP・強化学習（企業・商用システム開発, 吉野, 北岡）
- （雑談）対話生成（徳久, 松山, 北岡）
- ニューラルネットワークを用いた文生成（高橋, 杉山）

項目2,3の回答は入り混じっていたので一緒にしました

過去10年の対話システム研究の成果で、今後の実用システム開発において重要となると思われるものは何でしょうか？

それ以外で今後の実用対話システムの発展のために重要となる研究分野は何でしょうか？

講演者の意見と関連するアンケートの回答 (項目 2)

- 様々な知識を用いる雑談対話システムの構成法
- Webやテキストコーパスからの対話知識獲得
- タスク指向型対話と非タスク指向型対話の融合
- 社会的信号処理
- パーソナライゼーション
- パーソナリティ表出

- 対話知識の自動獲得と構造化 (吉村)
- 知識獲得 (大学等)
- 知識構築 (高橋)

- Dialog platforms that integrate emotional cares and functional tasks (Wu)

- マルチモーダル (企業・研究, 徳久, 大学等, 吉野)
- パラ言語・非言語情報の活用, 多人数対話 (匿名)
- 頷きタイミングや視線の実践的なモデル (大学等)
- 非言語情報 (言語以外のモダリティの) の処理とユーザの内面状態 (感情や態度) の推定技術と対話システムとの統合 (岡田)

- システムの一貫性 (記憶・パーソナリティ) (匿名)

講演者の意見と関連するアンケートの回答（項目3）

- 個人情報保護

- 倫理, 悪用・悪意に対する頑健性 (杉山)

- 言語表現のDBや物理世界へのグラウンディング (IoTとの連携を含む)

- シンボルグラウンディング (石川, 企業・研究)
- 実世界理解 (画像・動画) (杉山)
- 動画や画像などのマルチメディアデータ処理 (匿名)

- 文脈・状況に応じた音声理解・言語生成・音声合成

- 対話文脈に限らず, 外界, ユーザ・システムの状況を広く活用する技術 (石川)

- 実時間システムのための逐次的な対話処理

- ターンテイキング (杉山, 吉野)

- 様々なステークホルダ (ユーザ, システム開発者, システム運営者など) を考慮した評価手法

- 対話システムの評価技術 (石川)
- 評価尺度の拡充 (企業・商用システム開発)
- 対話の自動評価 (杉山)

アンケートの回答から

- 機械学習
 - ニューラルネットワークで用例ベースの応答を文脈に合わせて書き換える技術（稲葉）
 - ニューラルネットワーク文生成（高橋, 杉山）
 - Seq2seq の対話技術（大学等）
- コーパス
 - 大規模な実対話DB（大学等）
 - 膨大な対話データからよい対話を選別・蒸留する技術（河原）
- 理論の導入
 - 共通基盤などの対話理論の工学的実現（企業・研究）
 - 談話構造, 語用論的な表現をニューラルネットに本格的に導入すること. 伝統的な各種会話理論（会話分析, 社会言語学等）の知見に基づく, ニューラルアーキテクチャの設計（演繹的・帰納的手法の統合）（松山）

アンケートの回答から

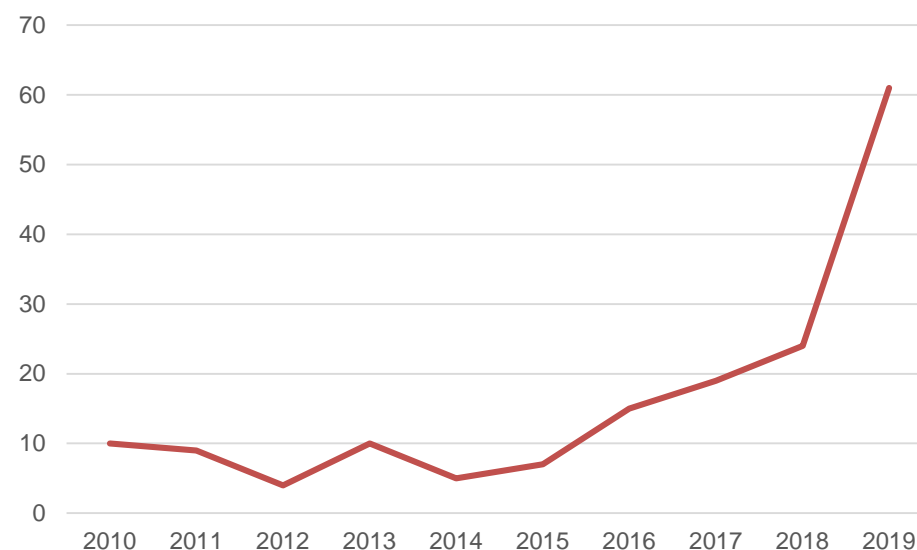
- タスク
 - 対話システムのカウンセリング・傾聴システムへの応用（岡田）
 - 現実的な利用シナリオ（駒谷）
 - タスク知識をどのように言語理解・対話管理と統合するか（匿名）
 - 実用サービスとしても基礎研究としても対話の重要な側面が現れるような
いタスクを選ぶ（松山）
- システム構築
 - 簡単で柔軟なシステム構築を可能とするアーキテクチャ（石川）

その他，過去10年の対話システムの研究に関して何か特筆すべきことがあればお書きください。

講演者の意見

- End-to-endの対話システム構築法の論文の増加
- チャレンジ・コンペティションの増加と様々な客観手法の提案
- 様々な会議での論文投稿数の激増と査読システムの混乱
- 今でもATISデータが用いられていること

ACLにおける“dialog”, “conversation”, “chatbot”をタイトルに含む論文数



アンケートの回答から（赤字は講演者）

- 機械翻訳や画像認識は深層学習とビッグデータによって劇的に性能が上がりましたが、**対話は話者の背景知識や話者間の関係性**において成り立つもので、必ずしも巷で得られるビッグデータからでは自然な対話の実現は難しいと実感しています。（吉村）
- この10年、アカデミックな世界から私自身は離れていますが、本分野の研究が精力的に行われていること、その成果、また、学会を牽引する皆様が以前から**実用化を強く意識され**、学術振興と実用化に貢献されてきたことが、分野の進展におおきな役割を果たしていると思います。ただ、実用化は海外ITメーカーを中心に大きく進んできています。**産学、海外とのさらなるコラボレーションが必要**だと考えています。この10年をこのように振り返り、将来を考えようとされていることこそが、特筆すべきことだと感じております。（石川）
- from ideas to systems and from systems to **platforms** (Wu)
- アカデミックには統計モデルを用いた対話システムが多いですが、商用対話システムはまだ**ルールベース**が多いです。ただ、これから先は統計モデルが主流になっていくと思いますので、**低コストで、ドメイン展開しやすい統計モデル**が必要になっていくと思います。（企業・商用システム開発）
- 10年前は難しかったことがどんどん実用化されたのは、音声認識の進化が大きいと思います。今後、本当に対話技術が必要とされていくのは、その**活用場面**が重要になると思います（例えば手術中などの手が使えない時に機械と対話や、字が書けない子どもとのコミュニケーションなど?）。今後、対話技術が自分たちの生活にはなくてはならない存在になれるように発展していくことを願います（徳久）
- Siriや Alexaに代表される製品が広く使われ始めているが、**対話研究との距離**が大きい（高橋）
- 思ったより進みが遅いように感じる。**一問一答以降**のことをもっと考えないといけない。（企業・研究）
- 「対話」の**キラーコンテンツ**がはっきりしていない（少なくとも社会に浸透していない）ことは大きな問題だと思います。（企業・研究）
- 音声認識が信頼できる精度まで高まったことと、**「雑談」**に焦点が当たり始めたことが極めて大きい変化だと思います。（杉山）
- POMDPなどの対話制御の機械学習の研究が盛んに行われたにも関わらず、**ほとんど役に立っていない**。現在のseq2seqの研究もそうならないか心配。（河原）

アンケートの回答から（赤字は講演者）

- 対話システムを使う人側が、（事実を伝達する対話においては、）詳しい正確な応答よりも、より早く**短い回答を好む**ようになってきているのかもしれない。（人どうしのコミュニケーションもTwitter化？）
（駒谷）
- POMDPから対話状態追跡にかけて、状態を離散的ではなく、確率的・連続的に捉えるやり方が定着しかけたのは良い兆候。しかし、**対話では応答出力が必要で、ここは言語的・シンボリック・離散的なものにならざるを得ない**。このギャップが埋められていないところで停滞している感がある。（北岡）
- これまでは総合格闘技、これからは異種格闘技。まだまだ難しい。世の中に知的情報処理の機運が高まり、対話研究に対する社会の要請も変化してきている昨今、まずは着実に、もう一歩手前の段階として、知的サービスやエージェントといった**「モノ以上・人間未満（と人間が認知する）存在」**と人間とのナチュラルなインターフェースとしての、使える技術としての対話技術の確立に、全力で取り組むべき。そして、そのためには、大量のデータに基づくデータドリブンな解析や、インタラクティブ・インターフェース・自然言語処理といった**各分野との「異種格闘技戦」**を行っていくべきである。（大学等）
- 対話システム研究が流行っているとは思いますが、人間の対話タスクを置き換えようというのは無理で、まだまだ研究すべきことが残っています。現在の研究は、人間の対話タスクとは異なる形で、必要なゴールに向けてタスクをデザインすることができれば商用システムを作れるところまでは関連分野の研究開発が成熟したと思います。ただ、**タスクをデザインする人よりも既存のタスクに乗って研究をする人が増えたというのを危惧しています**。タスクをデザインする能力、ノウハウなどが共有される必要があると思います。（吉野）
- チャットボット等、10年前には考えられなかった対話モデルが実現されてきました。しかし、それがそのまま人々に「実用」されるかということについては大きな疑問が残ります。道具立てが洗練されてきている今、改めて、対話システムが本質的に取り組むべき問題を整理する必要があるように感じています。
一般的に、狭いAIと広いAIのどちらの場合でも実現への確かな歩みを重ねるためには、**よいドメインと適切な評価関数**を探さなくてはならないと思います。音声言語処理や対話システムの研究分野は長く深い歴史を有しており、深層学習の出現により音声認識や言語理解・生成の要素技術はさらに発達してきましたが、どういう意義あるタスクの中でどういう方針で複雑な会話AIのコンポーネントを統合すれば良いかというグランドビジョンを改めて考える必要があるように思います。（松山）

- この10年，音声認識，機械学習の向上，対話システム構築に関する知見の集積，周辺技術の向上，ビジネスモデルの変化から実用システムは大きく進化した
- 重要だと思ふ技術は様々
- 生成系の機械学習モデル（強化学習，ニューラル対話生成）の価値に関して意見が分かれた
- 比較的多かった意見
 - 良いタスクを見つけることが重要
 - 他分野の技術を取り込むことが重要
 - 理論と統計の融合

- 方向性のバリエーションが大きい：分野の生存戦略としては良いが、少し効率が悪い？
 - 研究者が少ない
 - 基礎研究のテーマの選定にもビジネスモデルを考慮することが必要かもしれない
 - 実用と商用のギャップ：儲け - (開発コスト + 運用コスト) > 0?
 - 流行するデバイス, クラウド, 5G, IoTなど外的な要因
- 対話システムの別の利用法
 - 例：コンピュータサイエンス教育
 - 人工知能, 機械学習, 言語・音声処理, ソフトウェア工学, データベース, HCI
 - セキュリティ, プライバシー, 画像処理, CG, 教育応用, ゲーム, IoT, ...



(付録)

アンケート結果一覧(敬称略)

● 石川泰(Cerence Japan 株式会社)

(1) 過去 10 年(2010 年-2019 年)の実用対話システムの進化において重要な役割を果たした技術・手法・知見は何でしょうか？

以下の意見は、私個人の意見であり、所属組織を代表する意見ではありません。ご理解ください。

【サマリ】

特定の技術・手法・見識ではなく、以下の蓄積と相互作用であると考えます。

- 1)統計モデル, 機械学習技術に基づく音声認識・理解性能の向上
- 2)大規模データの収集と利用環境の充実, 成果・ツールの公開などのオープンな研究開発手法の進展
- 3)ビジネスモデルの変化による, 利用拡大

【説明】

この 10 年で実用対話システム(対話システムではない音声認識システムも含め)は大きく進展してきました。ただ、長年の技術の蓄積により実用化はすすんでおり、この 10 年で考えますと、進化に大きな役割を果たした「特定の新たな技術」があったわけではないと考えています。以下の各項目が相互に効果的に影響しあい、今日の進化を生んだと考えます。

1) 音声認識・理解技術

認識理解の性能が大きく向上していることは確かです。これには、1990 年代からの統計的手法、2000 年代(2001 年～2010)から活発化した DNN をはじめとする機械学習に基づいた手法、これらのたゆまぬ研究成果が基盤となったことは間違いありません。

2) データの活用と成果活用, 研究開発手法の変化

大規模なデータの活用が可能となったことに加え、過去の閉じた研究開発から、データの流通、大学関係者の成果の公開・相互活用、メーカを含めた成果・ツール類の公開が進みました。オープンな研究開発の手法が広がったことが、実用研究の推進に大きな役割を果たしていると考えています。

3) ビジネスモデルの変化

対話システムに限らず、情報サービスが、エンドユーザへの販売・課金を目的とするのではなく、広告などを含めた情報流通、蓄積データ(特に近年の帰納推論に基づく価値創造のもとで)を経営資源として販売すること、ブランド価値などにそのビジネスターゲットを移したことで、多くのユーザ

が無償(あるいは無償であるかのように)サービスを利用できるようになりました。これが 2), 1)をさらに加速し, 好循環が生じていると考えます。

(2) 過去 10 年の対話システム研究の成果で, 今後の実用システム開発において重要となると思われるものは何でしょうか?

【サマリ】

「対話技術」の実用化が進んだとは言えず, この 10 年の特定の研究成果の延長に将来の実用システム開発の進展があるとは思っておりません。研究領域としては, 古くからの領域ですが, 以下の領域で新たな成果が生まれることを期待しています。

- 1) 対話管理: 新規サービスや複数のサービス間の調整を含め, 決定的な対話状態遷移を記述せずに目的を達成する管理手法
- 2) 文脈理解・シンボルグラウンディング: 対話文脈に限らず, 外界, ユーザ・システムの状況を広く活用する技術

【説明】

多くのすぐれた研究がなされておりますし, それぞれの目的も異なりますので, 課題を明確にし, 新たなチャレンジをしているすべての成果は将来の実用化の基盤となるものと考えています。ただ, 私がかかわっているビジネス領域を考えますと, 人間懼口人間の音声対話, 複数の人間の談話などの基礎的な研究(ターンテーク, 対話の維持, 調整, そのためのパラ言語など)は興味深いものの, 実用化は先だと考えています。特に人間社会における人工物の目的についての議論が必要だと思っています。

実用的音声対話システムについていえば, 多くは一問一答で, 対話システムではその手続きを記述し, 決定的な状態遷移で管理をするものがほとんどです。ユーザの目的を支援する対話システムの実用化をさらに進めるに, この 10 年の特定の成果の延長として意見を述べることは困難ですが, 以下の領域(いずれも, 何をいまさらという古い言葉ですが)での研究の進展を期待しております。

1) 対話管理

POMDP などの検討がなされてから, 実用化のための対話管理はあまり重点が置かれていないように思います(不勉強ですが), しかし, 新規サービスの追加や, 複数サービス間の調整が必要なドメインでの対話システムを考えると, 再度考える必要があると考えます。

2) 文脈理解・シンボルグラウンディング

対話文脈に限らず, 言語と実世界情報とのマッピング, 情報世界, ユーザの行動, 他のユーザの行動などの文脈を利用する技術, その学習方法は, 今後の自律的システムとの間の対話システムの実用化の上で重要だと考えます。

(3) (2)にあげて頂いたもの以外で、今後の実用対話システムの発展のために重要となる研究分野は何でしょうか？

体系的に考えることが困難で、思い付きになりますが、以下が重要だと考えています。

【サマリ】

- 1) 対話システムのアーキテクチャ研究
- 2) 対話システムの評価技術:観測可能な人間と人工物の行動から評価・改良する技術
- 3) 人間と人工物の相互理解を上位目的とする対話研究

【説明】

1)アーキテクチャ

研究開発課題としては、少し考えにくいのかもかもしれませんし、ビジネスではデファクトスタンダードとなるのかもかもしれません。マルチモーダル IF 構築、サービス・アプリケーション両面で簡単に柔軟なシステム構築を可能とする SW, データ交換, モデル学習などのアーキテクチャを検討する必要があります。

2)評価技術

過去の専門家の発見, 利用者の主観に基づく対話システムの評価手法から, IoT をベースとして実運用時のユーザの行動が観測可能となったときの新たな評価方法が必要になるはず。評価方法, そのフィードバックにより簡単にアップデートが可能な方法, そのデータ流通のための個人情報保護のセキュリティの枠組みを検討する必要があります。

3)人工物と人間の相互理解と対話

人間が, 人工物を操作する, 情報世界から情報を抽出するという関係におけるインタフェース(有効性, 効率, 満足が尺度)から, 人間と自律動作する人工物(情報世界でも)とが協調動作する世界におけるインタフェース(人工物はあくまでも人工物であり, 人間のようにふるまう必要はない対象タスクを考えています)では, 人間と人工物の相互理解をゴールとするインタラクションが必要になります。そのモデリングや, 人工物が人間の要求や状況を理解するだけでなく, 人間が人工物の行動や状況を理解するための技術, 直近では状況を説明できる自然言語生成技術なども重要だと考えています。

(4) その他, 過去 10 年の対話システムの研究に関して何か特筆すべきことがあればお書きください。

この 10 年, アカデミックな世界から私自身は離れていますが, 本分野の研究が精力的に行われていること, その成果, また, 学会を牽引する皆様が以前から実用化を強く意識され, 学術振興と実用化に貢献されてきたことが, 分野の進展におおきな役割を果たしていると思います。ただ, 実用化は海外 IT メーカーを中心に大きく進んできています。産学, 海外とのさらなるコラボレーションが必

要だと考えています。この 10 年をこのように振り返り、将来を考えようとされていることこそが、特筆すべきことだと感じております。

- 吉村健(NTTドコモ)

(1) 過去 10 年(2010 年-2019 年)の実用対話システムの進化において重要な役割を果たした技術・手法・知見は何でしょうか？

音声認識の大語彙化、ルールベースの対話、文書分類、固有表現抽出

(3) (2)にあげて頂いたもの以外で、今後の実用対話システムの発展のために重要となる研究分野は何でしょうか？

対話知識の自動獲得と構造化、及びそれを活用した発話理解と発話生成

(4) その他、過去 10 年の対話システムの研究に関して何か特筆すべきことがあればお書きください。

機械翻訳や画像認識は深層学習とビッグデータによって劇的に性能が上がりましたが、対話は話者の背景知識や話者間の関係性において成り立つもので、必ずしも巷で得られるビッグデータからでは自然な対話の実現は難しいと実感しています。

- Xianchao Wu (Microsoft)

(1) 過去 10 年(2010 年-2019 年)の実用対話システムの進化において重要な役割を果たした技術・手法・知見は何でしょうか？

products such as Amazon's Echo, Alexa; Google Home; Microsoft's XIAOICE, Cortana and so on

(2) 過去 10 年の対話システム研究の成果で、今後の実用システム開発において重要となると思われるものは何でしょうか？

domain specific knowledge graph; user intention detection and reasoning; task-oriented multi-turn dialog; dialog platforms that integrate emotional cares and functional tasks.

(3) (2)にあげて頂いたもの以外で、今後の実用対話システムの発展のために重要となる研究分野は何でしょうか？

specific domains' task-oriented chatbots, such as finance, security, law, and so on

(4) その他、過去 10 年の対話システムの研究に関して何か特筆すべきことがあればお書きください。

from ideas to systems and from systems to platforms

● 匿名(40 歳以上, 企業・商用システム開発)

(1) 過去 10 年(2010 年-2019 年)の実用対話システムの進化において重要な役割を果たした技術・手法・知見は何でしょうか？

ディープリーニング、POMDP、強化学習、形態素解析、分散型コンピューティング

(2) 過去 10 年の対話システム研究の成果で、今後の実用システム開発において重要と思われるものは何でしょうか？

単語分散表現、attention

(3) (2)にあげて頂いたもの以外で、今後の実用対話システムの発展のために重要となる研究分野は何でしょうか？

転移学習、zero-shot learning, 音声認識との end2end

(4) その他、過去 10 年の対話システムの研究に関して何か特筆すべきことがあればお書きください。

アカデミックには統計モデルを用いた対話システムが多いですが、商用対話システムはまだルールベースが多いです。ただ、これから先は統計モデルが主流になっていくと思いますので、低コストで、ドメイン展開しやすい統計モデルが必要になっていくと思います。

● 匿名(企業・商用システム開発)

(1) 過去 10 年(2010 年-2019 年)の実用対話システムの進化において重要な役割を果たした技術・手法・知見は何でしょうか？

ニューラルネットワークが機械学習の主流となったこと。siri や小氷が世の中に出たこと。チャットアプリの普及によりボットに適したプラットフォームができたこと。特に人手不足によりボットの有用性が認められたこと

(2) 過去 10 年の対話システム研究の成果で、今後の実用システム開発において重要となると思われるものは何でしょうか？

機械学習

(3) (2)にあげて頂いたもの以外で、今後の実用対話システムの発展のために重要となる研究分野は何でしょうか？

言語資源や評価尺度の拡充

● 徳久良子(豊田中央研究所)

(1) 過去 10 年(2010 年-2019 年)の実用対話システムの進化において重要な役割を果たした技術・手法・知見は何でしょうか？

対話のための知識獲得や知識整備

タスク指向型対話の対話生成・制御技術

非タスク指向型対話(雑談対話)生成技術

音声認識技術

音声合成技術

(2) 過去 10 年の対話システム研究の成果で、今後の実用システム開発において重要となると思われるものは何でしょうか？

どれも重要だと思います

さらに、今後は、市場のニーズをもとに、対話システムを活用する場面や、応用システムの活用場所をきちんと整理することがとても重要になると思います。一問一答でいいのか、対話的なやりとりが必要なのかをきちんと整理する必要があると考えています。

(3) (2)にあげて頂いたもの以外で、今後の実用対話システムの発展のために重要となる研究分野は何でしょうか？

表情や音声や発話内容など、マルチモーダルな情報も含めたコミュニケーション技術

(4) その他、過去 10 年の対話システムの研究に関して何か特筆すべきことがあればお書きください。

10 年前は難しかったことがどんどん実用化されたのは、音声認識の進化が大きいと思います。今後、本当に対話技術が必要とされていくのは、その活用場面が重要になると思います(例えば手術中などの手が使えない時に機械と対話や、字が書けない子どもとのコミュニケーションなど?)。今後、対話技術が自分たちの生活にはなくてはならない存在になれるように発展していくことを願います

● 高橋哲朗(富士通研究所)

(1) 過去 10 年(2010 年-2019 年)の実用対話システムの進化において重要な役割を果たした技術・手法・知見は何でしょうか？

ニューラルネットワークを用いた文生成

(2) 過去 10 年の対話システム研究の成果で、今後の実用システム開発において重要と思われるものは何でしょうか？

ニューラルネットワークを用いた文生成

(3) (2)にあげて頂いたもの以外で、今後の実用対話システムの発展のために重要となる研究分野は何でしょうか？

知識処理(知識構築, 推論)

(4) その他、過去 10 年の対話システムの研究に関して何か特筆すべきことがあればお書きください。

Siri や Alexa に代表される製品が広く使われ始めているが、対話研究との距離が大きい

● 匿名(40歳以上, 企業・研究)

(1) 過去10年(2010年-2019年)の実用対話システムの進化において重要な役割を果たした技術・手法・知見は何でしょうか？

- ・音声認識精度の向上
- ・計算機の性能向上
- ・クラウド上の計算機環境の充実
- ・クラウドソーシング

(2) 過去10年の対話システム研究の成果で、今後の実用システム開発において重要と思われるものは何でしょうか？

- ・マルチモーダル処理

(3) (2)にあげて頂いたもの以外で、今後の実用対話システムの発展のために重要となる研究分野は何でしょうか？

- ・共通基盤などの対話理論の工学的実現
- ・フレーム問題やシンボルグラウンディング問題などの基本問題

(4) その他、過去10年の対話システムの研究に関して何か特筆すべきことがあればお書きください。

思ったより進みが遅いように感じる。一問一答以降のことをもっと考えないといけない。

● 匿名(40歳以上, 企業・研究)

(1) 過去10年(2010年-2019年)の実用対話システムの進化において重要な役割を果たした技術・手法・知見は何でしょうか？

音声認識

(2) 過去10年の対話システム研究の成果で、今後の実用システム開発において重要と思われるものは何でしょうか？

動画や画像などのマルチメディアデータ処理

(4) その他, 過去 10 年の対話システムの研究に関して何か特筆すべきことがあればお書きください.

「対話」のキラーコンテンツがはっきりしていない(少なくとも社会に浸透していない)ことは大きな問題だと思います。

● 杉山弘晃(NTT)

(1) 過去 10 年(2010 年-2019 年)の実用対話システムの進化において重要な役割を果たした技術・手法・知見は何でしょうか?

音声認識技術の進展, 利用可能なテキストの爆発的増加と質の変化(Twitter・FB 等での極めて日常的な事象への記述が増加), 莫大なデータを用いた深層学習による流暢な文の生成

(2) 過去 10 年の対話システム研究の成果で, 今後の実用システム開発において重要となると思われるものは何でしょうか?

音声認識技術, 莫大なデータを用いた深層学習による流暢な文の生成, 対話の自動評価, 強化学習, 実世界理解(画像・動画)

(3) (2)にあげて頂いたもの以外で, 今後の実用対話システムの発展のために重要となる研究分野は何でしょうか?

雑談対話のタスク的構造化, 雑談対話のユーザへの効果の追求, 対話用音声合成, ターンテイク, 倫理, 悪用・悪意に対する頑健性

(4) その他, 過去 10 年の対話システムの研究に関して何か特筆すべきことがあればお書きください.

音声認識が信頼できる精度まで高まったことと, 「雑談」に焦点が当たり始めたことが極めて大きい変化だと思います.

● 河原達也(京都大学)

(1) 過去 10 年(2010 年-2019 年)の実用対話システムの進化において重要な役割を果たした技術・手法・知見は何でしょうか？

音声対話システムが普及した一番の要因は音声認識精度の向上。
対話に関しては、データの蓄積と(ドメイン同定などの)機械学習

(2) 過去 10 年の対話システム研究の成果で、今後の実用システム開発において重要となると思われるものは何でしょうか？

深層学習

(3) (2)にあげて頂いたもの以外で、今後の実用対話システムの発展のために重要となる研究分野は何でしょうか？

膨大な対話データからよい対話を選別・蒸留する技術

(4) その他、過去 10 年の対話システムの研究に関して何か特筆すべきことがあればお書きください。

POMDP などの対話制御の機械学習の研究が盛んに行われたにも関わらず、ほとんど役に立っていない。現在の seq2seq の研究もそうならないか心配。

● 駒谷 和範(大阪大学)

(1) 過去 10 年(2010 年-2019 年)の実用対話システムの進化において重要な役割を果たした技術・手法・知見は何でしょうか？

・音声対話システムでは、音声認識のクラウド化(それを支える通信の高速化・コモディティ化)と、音声認識の性能が一線を越えたことが大きい。デスクトップ PC ではなく、モバイル端末で(最新の)音声認識が使えるようになり、それに伴って音声応答アプリ(一問一答型対話システム)が実現されたことは大きい。また音声認識の「手離れ」がよくなり、非専門家が使ってもそれなりに動くようになったのは素晴らしい成果である。

・テキスト対話システムは、テキスト検索の域を大きく超えていない。オートマトン型の状態遷移を持った対話システムがテキストチャットでリバイバルしているのは、これもモバイル端末上で(つまり入出力可能な情報が少ない場面で)こういうシステムを使いたいというニーズが増えているのかもしれない。

(2) 過去 10 年の対話システム研究の成果で、今後の実用システム開発において重要となると思われるものは何でしょうか？

・狭義の対話システム研究(対話管理など)は、実用システムにおいては、一問一答型アプリケーションの台頭により実用的にはほぼ進んでいない。

・一方で、言語理解や対話管理への機械学習の適用は、たくさんの研究事例が積み重ねられ、コモディティ化しつつある。つまり、雌伏の時期を経てこなれてきたことにより、次世代での実用システムの基礎となる素地が整いつつある。ただ「単にデータが大量にあれば学習できる」という機械学習の適用法から、問題の性質上、「データが少ないが、人の(ぼんやりした)知見はある」という状況から、いかに学習やチューニングを行えるか、という方法論が必要となると思われる。このためには実用的なシミュレーションが必要であり、それを行うための人間のふるまいや評価基準(感じ方)に関する知見や分析が必要となる。

・「賢い応答をする」対話システムという文脈では、自然言語処理などの隣接分野における成果を取り入れる時期に来ている。概念間の距離化を行う新たな道具立てが生まれているのは大きい。

(3) (2)にあげて頂いたもの以外で、今後の実用対話システムの発展のために重要となる研究分野は何でしょうか？

・「実用」されるのに必要なのは現実的な利用シナリオである。ここ 10 年の音声応答アプリの台頭は、スマホの出現による利用シーンの変化によるところが大きい。

・PC ではなく、スマホのように入出力可能な情報が限られた状況での利用が続くなら、20 年前の電話での音声対話システムのように、状況に応じた出力を行う対話管理や自然言語生成などの研究は重要性を増す。ただし、VR 技術のコモディティ化など、対話システムが利用されるシーンが変化すると、このような前提は大きく変わり得る。

(4) その他、過去 10 年の対話システムの研究に関して何か特筆すべきことがあればお書きください。

対話システムを使う人側が、(事実を伝達する対話においては、)詳しい正確な応答よりも、より早く短い回答を好むようになってきているのかもしれない。(人どうしのコミュニケーションも Twitter 化?)

● 北岡教英(豊橋技術科学大学)

(1) 過去 10 年(2010 年-2019 年)の実用対話システムの進化において重要な役割を果たした技術・手法・知見は何でしょうか？

POMDP(少し前ですが 2010 年には現役だったかと)、NN による対話状態追跡、End-to-end 応答生成

(2) 過去 10 年の対話システム研究の成果で、今後の実用システム開発において重要となると思われるものは何でしょうか？

正直言って分かりません。(少なくとも上記は何も残らないのではないかと考えてます)

(3) (2)にあげて頂いたもの以外で、今後の実用対話システムの発展のために重要となる研究分野は何でしょうか？

雑談対話システムの研究、そのための、1. 自然な応答スタイルの研究 と、2. 適切な応答内容の生成の研究、という古典的などころに戻るべきと考える。1. は最近、いくつかの研究例が復活し始めている気がする。2. に関しては、データサイエンス時代の知見を活かしたリアルタイムな知識ベースの構築・共有が必要ではないかと考える(例えば世界中のリアルタイムな天気・気温や天気予報、時事ネタなど、その他いろいろ)。これができれば、2. の研究はそれをいかに活用するかが研究の主眼になる。

(4) その他、過去 10 年の対話システムの研究に関して何か特筆すべきことがあればお書きください。

POMDP から対話状態追跡にかけて、状態を離散的ではなく、確率的・連続的に捉えるやり方が定着しかけたのは良い兆候。しかし、対話では応答出力が必要で、ここは言語的・シンボリック・離散的なものにならざるを得ない。ここのギャップが埋められていないところで停滞している感がある。

● 匿名(40 歳以上, 大学・公的研究機関)

(1) 過去 10 年(2010 年-2019 年)の実用対話システムの進化において重要な役割を果たした技術・手法・知見は何でしょうか？

音声認識・合成の性能向上、NN ベースの言語理解の頑健性向上、IBM Watson、アンドロイド ERICA

(2) 過去 10 年の対話システム研究の成果で、今後の実用システム開発において重要となると思われるものは何でしょうか？

Seq2seq の対話技術、大規模な実対話 DB、顔きタイミングや視線の実践的なモデル

(3) (2)にあげて頂いたもの以外で、今後の実用対話システムの発展のために重要となる研究分野は何でしょうか？

オープンプラットフォーム、マルチモーダル統合、異常検知、強化学習、知識獲得

(4) その他、過去 10 年の対話システムの研究に関して何か特筆すべきことがあればお書きください。

これまでは総合格闘技、これからは異種格闘技。

人間どうしの対話については、本質的なモデル化はもちろんデータに基づく表層的な再現すらまだまだ難しい。世の中に知的情報処理の機運が高まり、対話研究に対する社会の要請も変化して

きている昨今、まずは着実に、もう一歩手前の段階として、知的サービスやエージェントといった「モノ以上・人間未満(と人間が認知する)存在」と人間とのナチュラルなインタフェースとしての、使える技術としての対話技術の確立に、全力で取り組むべき。そして、そのためには、大量のデータに基づくデータドリブンな解析や、インタラクション・インタフェース・自然言語処理といった各分野との「異種格闘技戦」を行っていくべきである。

● 匿名(40 歳以上, 大学・公的研究機関)

(1) 過去 10 年(2010 年-2019 年)の実用対話システムの進化において重要な役割を果たした技術・手法・知見は何でしょうか？

クラウドコンピューティングの影響は大きいと思います。特に対話システムの分野は音声認識, 音声合成, 自然言語処理など, フロントエンドの時点で大きな計算量の複雑な処理を必要とするため, クラウドを通じてそれらを手軽に利用できるようになったのは大きいと思います。

関連して, スマートフォンからスマートスピーカ, 会話ロボットなどの流れでのハードウェアの発展(普及)も大きいかと思います。日常的に使われるというところまでは及ばないものの, 試そうと思えばいつでも試せる状態にあるかと思います。

また, 言うに及びませんが, 深層学習の発展も大きいと思います。特に音声認識や音声合成に関しては, 従来の問題設定における精度向上という観点での研究はほとんどのものが終わってしまったと言っても過言ではないと思います。

(2) 過去 10 年の対話システム研究の成果で、今後の実用システム開発において重要となると思われるものは何でしょうか？

(1) で書いたように、過去 10 年で多くの対話システムが実現されて普及した一方で、様々な問題が浮き彫りになってきたのだと思います。それらのいくつかは、音声対話システムの研究者にとっては 10 年以上前から認識されていた問題でありながら、現在の音声対話システムには組み込まれていないものがほとんどだと思います。

文字でやり取りをする対話システムと比較して、音声対話では「いつ喋るか」というタイミングの問題と、文字では伝わらない言語外の情報の扱いが重要と思われます。これらの扱いに関しては 2000 年代にかけて様々な取り組みがありましたが、(1) で述べた、ハードの普及と深層学習の発展を受けて再度注目を集める部分かと思います。

● 吉野 幸一郎 (NAIST)

(1) 過去 10 年(2010 年-2019 年)の実用対話システムの進化において重要な役割を果たした技術・手法・知見は何でしょうか？

音声認識, 言語理解, Belief Tracking (Deep な手法含む), (ニューラル)言語生成, 音声合成, 音響イベント検出 (VAD), embedding (言語表現)

(2) 過去 10 年の対話システム研究の成果で、今後の実用システム開発において重要となると思われるものは何でしょうか？

言語理解, Belief Tracking, Policy learning, 意味解析, 言語表現

(3) (2) にあげて頂いたもの以外で、今後の実用対話システムの発展のために重要となる研究分野は何でしょうか？

Computer vision, ターンテイクなどのリアルタイムイベントの利用, マルチモダリティ

(4) その他、過去 10 年の対話システムの研究に関して何か特筆すべきことがあればお書きください。

対話システム研究が流行っているとは思いますが、人間の対話タスクを置き換えようというのは無理で、まだまだ研究すべきことが残っています。現在の研究は、人間の対話タスクとは異なる形で、必要なゴールに向けてタスクをデザインすることができれば商用システムを作れるところまでは

関連分野の研究開発が成熟したと思います。ただ、タスクをデザインする人よりも既存のタスクに乗って研究をする人が増えたというのを危惧しています。タスクをデザインする能力、ノウハウなどが共有される必要があると思います。

● 岡田将吾(北陸先端科学技術大学院大学)

(1) 過去 10 年(2010 年-2019 年)の実用対話システムの進化において重要な役割を果たした技術・手法・知見は何でしょうか？

- ・大規模音声・言語データの収集と深層学習技術の発展に伴う、音声認識・言語処理精度の向上
- ・Alexa, Google home などのスマートスピーカの市場投入：
(一般ユーザが音声インターフェイスを使う機会が増え、機械に話しかけることが通常の生活の一部となったから。また、今後様々なサービスを提供する対話システムを搭載するためのプラットフォームになることが期待できるから。)

(2) 過去 10 年の対話システム研究の成果で、今後の実用システム開発において重要と思われるものは何でしょうか？

- ・非言語情報(言語以外のモダリティの)の処理とユーザの内面状態(感情や態度)の推定技術と対話システムとの統合 (SimSensei, Social Skill Trainer)
- ・Chatbot に代表される、「話すこと」自体を目的とする対話システムの構築

(3) (2)にあげて頂いたもの以外で、今後の実用対話システムの発展のために重要となる研究分野は何でしょうか？

- ・クロスモーダル、マルチモーダル情報の認識・生成
- ・対話システムのカウンセリング・傾聴システムへの応用
(高齢者との対話・心療内科向けのカウンセリングといった具体的なタスクを実行できる対話システム・エージェントの研究開発)

● 稲葉通将(電気通信大学)

(1) 過去 10 年(2010 年-2019 年)の実用対話システムの進化において重要な役割を果たした技術・手法・知見は何でしょうか？

技術ではないかもしれませんが、クラウドソーシングの普及は大きかったと思います。

(2) 過去 10 年の対話システム研究の成果で、今後の実用システム開発において重要となると思われるものは何でしょうか？

ニューラルネットワークで用例ベースの応答を文脈に合わせて書き換える技術は今後有力になっていく可能性があると思います。例えば <http://arxiv.org/abs/1809.05296>

松山洋一(早稲田大学)

(1) 過去 10 年(2010 年-2019 年)の実用対話システムの進化において重要な役割を果たした技術・手法・知見は何でしょうか？

深層学習の爆発的進展とそれに基づく音声認識、音声合成、Neural Response Generation(2015 年以來)等によって従来対話システムの研究フィールドと接点のなかった自然言語処理・機械学習の研究者・開発者が参加できるようになり、AI のグランドチャレンジの一つとしての対話処理が大きな意味を持ち始めた歴史的転換の 10 年でした。同時に、対話の本質的な難しさと面白さも露呈してきたという印象があります。

(2) 過去 10 年の対話システム研究の成果で、今後の実用システム開発において重要となると思われるものは何でしょうか？

同上

(3) (2)にあげて頂いたもの以外で、今後の実用対話システムの発展のために重要となる研究分野は何でしょうか？

・談話構造、語用論的な表現をニューラルネットに本格的に導入すること。伝統的な各種会話理論(会話分析、社会言語学等)の知見に基づく、ニューラルアーキテクチャの設計(演繹的・帰納的手法の統合)

・そもそも慎重に対話のドメインを選ぶこと。実用サービスとしても基礎研究としても対話の重要な側面が現れるような(とってつけたような対話ではない)いい問題(タスク)を選ぶべきだと思います。

(4) その他、過去 10 年の対話システムの研究に関して何か特筆すべきことがあればお書きください。

チャットボット等, 10 年前には考えられなかった対話モデルが実現されてきました. しかし, それがそのまま人々に「実用」されるかということについては大きな疑問が残ります. 道具立てが洗練されてきている今, 改めて, 対話システムが本質的に取り組むべき問題を整理する必要があるように感じています.

一般的に, 狭い AI と広い AI のどちらの場合でも実現への確かな歩みを重ねるためには, よいドメインと適切な評価関数を探さなくてはならないと思います. 音声言語処理や対話システムの研究分野は長く深い歴史を有しており, 深層学習の出現により音声認識や言語理解・生成の要素技術はさらに発達してきましたが, どういう意義あるタスクの中でどういう方針で複雑な会話 AI のコンポーネントを統合すれば良いかというグランドビジョンを改めて考える必要があるように思います.

● 匿名

(1) 過去 10 年(2010 年-2019 年)の実用対話システムの進化において重要な役割を果たした技術・手法・知見は何でしょうか?

音声認識の高精度化および統計的自然言語理解の進展によって, 音声対話システム・テキスト対話システムともに, 入力エラーを特に意識せずに対話システムのプロトタイピングが可能になったこと.

(2) 過去 10 年の対話システム研究の成果で, 今後の実用システム開発において重要と思われるものは何でしょうか?

事前学習言語モデルを活用した少数事例からの対話システムの開発

(3) (2)にあげて頂いたもの以外で, 今後の実用対話システムの発展のために重要となる研究分野は何でしょうか?

タスク知識をどのように言語理解・対話管理と統合するか

● 匿名

(1) 過去 10 年(2010 年-2019 年)の実用対話システムの進化において重要な役割を果たした技術・手法・知見は何でしょうか?

音声認識性能の改善

(2) 過去10年の対話システム研究の成果で、今後の実用システム開発において重要と思われるものは何でしょうか？

パラ言語・非言語情報の活用, 多人数対話

(3) (2)にあげて頂いたもの以外で、今後の実用対話システムの発展のために重要となる研究分野は何でしょうか？

システムの一貫性(記憶・パーソナリティ)