

対話サービスプラットフォームの開発

Development of Dialog Service Platform

山上 勝義
Katsuyoshi Yamagami

牛尾 貴志
Takashi Ushio

遠藤 充
Mitsuru Endo

藤原 菜々美
Nanami Fujiwara

史 宏 杰
Honjie Shi

堀井 則 彰
Noriaki Horii

要 旨

対話技術は従来のユーザーインターフェースの問題を克服する技術として有力である。対話技術の実用化を加速するため、対話技術群を統合した対話サービスプラットフォームを開発した。共通の対話処理の枠組みの提供により、対話サービスの開発を効率化できる。対話サービスプラットフォームの利用により対話サービスの開発工数を大幅に低減した。

Abstract

A dialogue system is a powerful technology to overcome poor interactivity in conventional user interfaces. To accelerate practical dialogue-based service development, we integrated underlying technologies into a common framework called Dialogue Service Platform (Dialogue SPF). This paper reports specific cases of dialogue-based service developments using the platform, and its effectiveness for development cost reduction.

1. はじめに

機器やサービスのIT化が急速に進むなか、GUI (Graphical User Interface) 操作を意思疎通の主軸とする従来のユーザーインターフェースでは、ユーザーが行うべき操作の煩雑化が深刻な問題となりつつある。この問題を解決するべく、筆者らは、人間にとって最も意思表示が容易にできる「自然言語」を意思疎通の手段とする対話インターフェース技術の開発を進めてきた。

さまざまなアプリケーション向けの対話インターフェース (対話サービス) の開発を効率化するため、対話処理の要素技術を統合した対話サービスプラットフォーム [1] (以後、対話SPFと表記) を開発した。対話SPF開発の狙いは主に次の3つである。

1. 対話サービス共通の対話処理の枠組みの提供
2. 対話サービス開発のコスト低減の枠組みの提供
3. 対話サービス品質担保の枠組みの提供

1は、共通の対話処理の枠組み提供により、対話サービスごとの開発量そのものを削減し、類似の対話サービスの開発を効率化するためである。2は、対話サービス開発で避けられない、自然言語という表現の自由度が高い入力を扱うことに起因する開発コストを低減するためである。3は、2と同様に自然言語を入力することに起因する、対話サービス品質テストの負担を緩和するためである。

対話サービスの動向として、AIスピーカなどの音声コマンドを理解する音声対話端末の普及の兆しがある[2]。その1つであるEcho^(注1)では、ユーザーの発話をあらかじめ設定した有限個の意図 (コマンド) に割り当てるタ

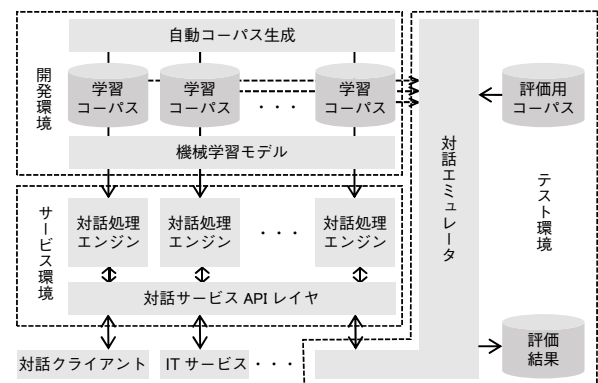
イプの対話機能が提供される[3]。これは音声認識による機器操作と同等レベルの対話機能の提供である。

筆者らの対話SPFでは、Echoが提供する対話機能をカバーする一問一答型の対話機能に加え、複数回の対話からユーザーの意図を推定する意図推定型の対話機能という、人が行う対話的業務に近いレベルの機能を提供する。

本稿では、2章で対話SPFの概要について、3章~5章で対話SPFの技術特徴について、6章で実証実験システムについて、7章で対話SPFの効果について述べる。

2. 対話サービスプラットフォームの概要

第1図に対話SPFの構成図を示す。対話SPFは、対話処



第1図 対話サービスプラットフォームの構成

Fig. 1 Structure of Dialog Service Platform

(注1) Amazon Technologies, Inc.の米国およびその他の国における登録商標または商標。

理を実行するサービス環境、対話サービスの開発をサポートする開発環境、および、テスト環境の3つの環境から構成される。以降、各環境の役割について述べる。

2.1 サービス環境

対話処理とは、ユーザーが入力した発話文に対して、応答文を返す処理を言う。サービス環境は、ユーザーが使用する対話クライアント、あるいは、他のITサービスからの対話処理要求に応じて対話処理を行う。対話処理は対話処理エンジンが担う。対話サービスAPI (Application Programming Interface) レイヤは、対話処理エンジンに対話処理を要求するためのAPIを提供し、対話クライアントがAPIで指定した対話処理エンジンに対話処理要求を振り分ける。

業務用途の対話サービスで共通性の高い次の2つのタイプの対話処理エンジンを開発した。

〔1〕 一問一答型対話処理エンジン

本エンジンは、FAQ検索のように、ユーザーの1回の発話文入力に対して応答文を返す対話処理を行う。ユーザーの発話文の「発話主題」を推定し、推定した発話主題に最も適合する応答文を選択する。3章で本エンジンの処理の詳細について述べる。

〔2〕 意図推定型対話処理エンジン

本エンジンは、ネットショッピングでの膨大な商品候補からのおすすめ提案のように、ユーザーとの複数回の対話によりユーザーの意図を理解し、その意図に沿った結果を返す対話処理を行う。

一般的なユーザーは、自身の意図(要望)を1つの発話文としての確に表現できるとは限らない。自身の意図に関連する断片的な情報を複数回発話することにより、自身の意図の全体像を伝えようとする(例えば、販売員との会話で徐々に意図を伝えるような状況)。本エンジンは、このようにユーザーから断片的な情報を引き出しつつ、ユーザーの意図の候補を絞り込む対話を実現する。4章で本エンジンの処理の詳細について述べる。

2.2 開発環境

開発環境の役割は、対話処理エンジンが担う対話処理の開発の効率化である。具体的には、対話処理エンジン内部の機械学習で用いる学習モデル構築に必要な学習データ(学習コーパス)を自動的に拡張する(自動コーパス生成)。これにより、学習コーパス作成のコストを低減する。5章で自動コーパス生成について詳細を述べる。

2.3 テスト環境

テスト環境は、対話処理エンジンに対して、実サービ

ス運用時にユーザーが入力する発話文を網羅的に想定したテストを行う。不正な発話文も含めた発話文セット(評価用コーパス)を、ユーザーの対話処理要求を模倣する対話エミュレータを通じて、サービス環境(対話処理エンジン)に与え、その動作をテストする。これにより対話サービスの開発者が自前でテスト環境を整備するコストを低減可能とした。

3. 一問一答型対話処理エンジン

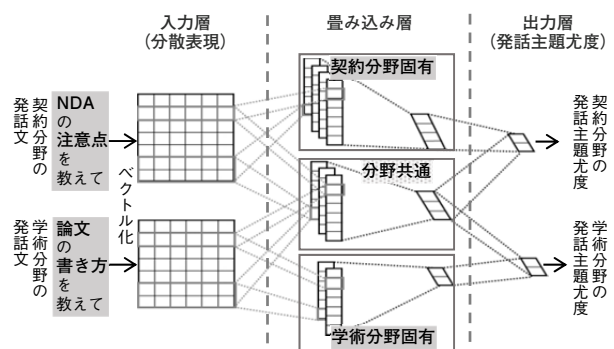
3.1 複数分野CNNによる発話主題推定

畳み込みニューラルネットワーク(CNN)によりユーザーの発話主題推定を行う方式[4]を開発した。

第2図に示すように、ユーザーの発話文を単語列に分割し、単語ごとに分散表現(注2)と呼ばれる固定次元のベクトルに変換した後、畳み込み層に入力する。畳み込み層で、発話文中の単語とその組み合わせといった言語特徴量と発話主題との相関モデルを学習し、出力層において複数の発話主題ごとに尤度(ゆうど)を計算する。最も尤度の高い発話主題を推定結果とする。

本方式のポイントは、複数の異なるタスクの分野における発話主題推定において、分野固有の畳み込み層に学習される相関モデル情報の不足を、分野共通の畳み込み層に学習される相関モデル情報で補完することで発話主題推定性能を向上した点にある。

対話状態推定の国際チャレンジタスクDSTC4[5]において、本方式が1位の性能を獲得した結果を第1表に示す。F-measureは適合率と再現率の調和平均であり、トレードオフ関係にある適合率・再現率を総合した評価値である。



第2図 複数分野CNNによる発話主題推定

Fig. 2 Estimation of utterance topic by multi-domain CNN

(注2) 意味が類似した単語のベクトル間距離が近くなるように構成したベクトル表現。同じ意味の文の表現の違いを吸収する効果が期待できる。

第1表 DSTC4における対話状態推定性能結果

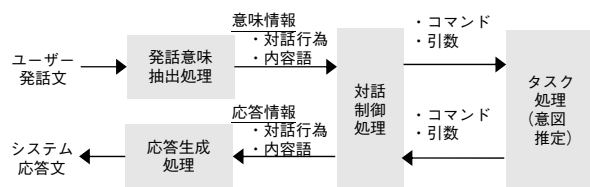
Table 1 Result of dialog state estimation in DSTC4

参加チーム	F-measure
Team-1 (Panasonic)	0.48
Team-2	0.47
Team-3	0.35
Team-4	0.43
Team-5	0.42

4. 意図推定型対話処理エンジン

4.1 意図推定型対話処理エンジンの対話処理の流れ

意図推定型対話処理エンジン[6]では、ユーザーから複数回入力される各発話文から意味情報を抽出し、その意味情報を用いて、段階的にユーザー意図の候補を絞り込む形でユーザーの意図を推定する。第3図に本エンジンの構成を示す。



第3図 意図推定型対話処理エンジンの構成

Fig. 3 Structure of dialog engine for intention estimation

発話意味抽出処理は、ユーザー発話文から意味情報（「対話行為（次節で述べる）」と「内容語」）を抽出する。内容語は発話文において意味の中心を担うキーワードである。例えば、後述のレシピ提案対話では、料理のジャンル名（和風、洋風、中華風…）や食材名（牛肉、にんじん…）などが内容語である。

対話制御処理は、意味情報から対話状態^(注3)を判断し、タスク処理（意図推定処理）で実行すべき処理の「コマンド」と「引数」を出力する。

タスク処理では、ユーザーの発話から得られた情報（「コマンド」と「引数」）により、ユーザーの意図の候補を絞り込む。後述のレシピ提案対話では、ユーザーの意図はユーザーの好みに合うレシピの候補として表現される。絞り込んだ候補の多少に応じて、対話制御処理でどのような応答をすべきかを判断するための「コマンド」「引数」を出力する。

対話制御処理は、タスク処理部からの「コマンド」「引数」により、応答すべき内容を指定する応答情報（「対

(注3) 対話における受け答えの状況をシステム内部の表現として表したもの。例えば、ユーザーの発話文が質問文であった場合、対話システムは回答文を返すべき状態であることを内部状態として表現する。

話行為」「内容語」）を生成する。

応答生成処理は応答情報からシステム応答文を生成する。この一連の処理過程でのデータのの流れの例を第2表に示す（上から下に処理が進行）。この例は、ユーザーの「和風がいい」（和風料理が好きである）という発話から意味情報の抽出を経て、意図の絞り込みを行うタスク処理に「和風」を肯定する条件を追加するコマンドと引数（AddCond, +和風）を発行し、タスク処理は「和風」での絞り込み結果の候補数が多いため、さらにユーザーから条件を引き出すためのコマンドと引数（NeedCond, Other）を発行し、ユーザーに他の条件を問う応答文が返される流れを表している。

第2表 意図推定型対話処理エンジンの内部処理

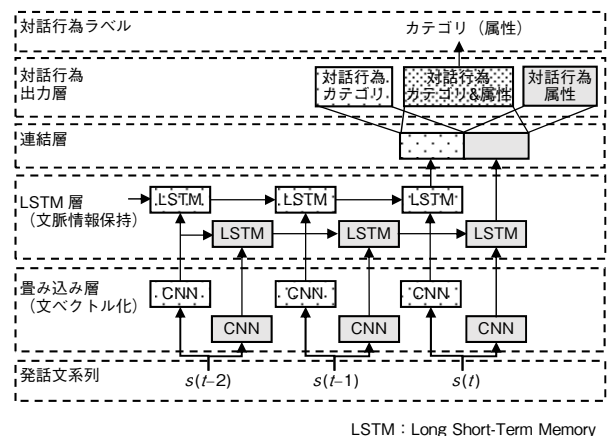
Table 2 Internal process of intention estimation dialog engine

データの種別	データの例
ユーザー発話文	「和風がいい」
意味情報（対話行為、内容語）	State(Food), 和風
コマンド、引数	AddCond, +和風
コマンド、引数	NeedCond, Other
応答情報（対話行為、内容語）	Qst, Other
システム応答文	「他に条件はありますか？」

4.2 共起関係モデル学習による対話行為推定

対話行為とは、対話のなかで発話文が果たす機能の分類を言う。本対話行為推定方式[7]では、タスク非依存の対話行為をカテゴリ、タスク依存の対話行為を属性と定義し、両者を同時に推定する。例えば「和風がいい」という発話文の場合は、「食事について述べる」という機能を果たす対話行為 State (Food) が推定すべき対話行為である。Stateはカテゴリの例、Foodは属性の例である。

第4図に本方式の概要を示す。発話文系列の $s(t)$ は、対話行為推定対象の発話文であり、 $s(t-1)$ は $s(t)$ の直前、 $s(t-2)$



LSTM : Long Short-Term Memory

第4図 共起関係モデル学習による対話行為推定

Fig. 4 Estimation of dialog act by learning co-occurrence model

は $s(t-1)$ の直前の発話文である。 $s(t-2)$ より前の発話文は図示を省略した。

発話文を畳み込み層 (CNN) により文単位で固定次元ベクトルに変換した後、時系列入力を逐次的・累積的に学習するリカレントニューラルネットの1種であるLSTM (Long Short-Term Memory) 層に入力する。 $s(t)$ のベクトルが入力されたLSTM層は、 $s(t)$ までに至る過去の発話文系列の情報 (文脈情報) も保持している。このLSTM層の出力は、連結層を経て対話行為出力層に入力され、対話行為が推定される。

本方式のポイントは、カテゴリと属性の共起関係を学習できるようにした点にある。カテゴリ推定用のLSTM層 (第4図のLSTM層の上段) と、属性推定用のLSTM層 (第4図のLSTM層の下段) をそれぞれ独立に構成し、各々のLSTM層の出力ベクトルを連結層で連結し、カテゴリと属性の共起関係を表現し学習可能とした。これにより、カテゴリと属性、それぞれ単独で推定した結果を組み合わせるよりも推定精度を高めることができた。

対話状態推定の国際チャレンジタスクDSTC5[8]において、本方式が対話行為推定タスクの部門の1位の性能を獲得した結果を第3表に示す。DSTC5は旅行ガイドと旅行者の対話を題材にしたタスクであり、対話行為推定のベンチマークは旅行ガイド発話、旅行者発話に分かれて実施され、どちらも1位の性能であった。

第3表 DSTC5における対話状態推定性能結果

Table 3 Result of dialog state estimation in DSTC5

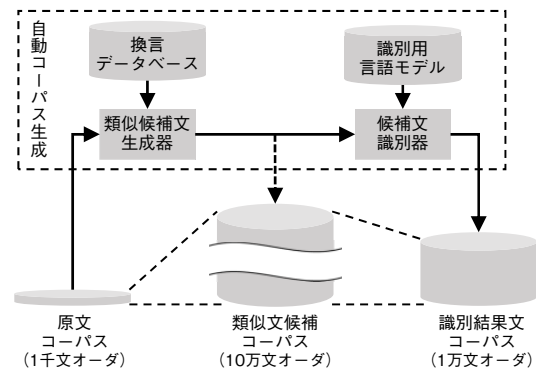
参加チーム	F-measure	F-measure
	旅行ガイド発話	旅行者発話
Team-0	0.321	0.245
Team-2 (Panasonic)	0.464	0.530
Team-3	0.395	0.441
Team-5	0.419	0.474
Team-7	0.373	0.457

5. 自動コーパス生成

対話技術を含む自然言語処理での機械学習の性能向上の鍵は、学習コーパスの量と質である。学習コーパスは、同じ意味の発話文であっても、ユーザーが発話する可能性のある表現を複数かつ網羅的に含んでいることが望ましい。しかし、それを人手作業で達成するのは作業コストが高い。

筆者らは解決策として、表現の異なる発話文 (類似文) を自動生成して学習コーパスの量を増やす方式 (自動コーパス生成方式) [9]を開発した。

第5図に本方式の構成と働きを示す。人手で作成した



第5図 自動コーパス生成の構成と働き

Fig. 5 Process flow of automatic corpus generation

比較的少量の原文コーパスを、類似候補文生成器により、表現の一部を変換して類似文候補コーパスに拡張する。表現の変換は、換言データベースが保持する複数の換言ルールを複数・選択的に適用して行う。

生成された類似文候補コーパスには、表現変換に起因する、言語表現として違和感のある文も含まれる。候補文識別器は、別途大量の言語コーパスから学習した識別用言語モデル (注4) を用いて、言語表現の自然さのスコアを計算し、スコアが閾 (しきい) 値を下回る類似文候補を違和感のある文として削除する。削除した結果を識別結果文コーパスとする。

本方式は、元々機械翻訳システムの学習コーパス生成向けに開発したもので、現段階では対話処理エンジン向けの学習コーパス生成には適用していない。本稿では、機械翻訳システムの学習コーパス作成に用いた際の評価結果を示す。

工場での会話を収集したコーパス約500文の原文コーパスのみで学習した機械翻訳システムの翻訳性能と、本方式で拡張した学習コーパスで学習した機械翻訳システムの翻訳性能の比較結果を第4表に示す (それぞれの学習時に旅行会話コーパス16万文を追加している)。機械翻訳性能を表すBLUE値スコアの上昇が14.82と大幅な翻訳性能向上に結びついている。機械翻訳性能のBLUE値から換算したコーパス数に基づいて原文コーパスに対してどれだけコーパスを新規に生成できたかの指標であるコーパスの生成効率では10.53 (倍) と従来のコーパス生

第4表 コーパスの種類による翻訳性能の結果

Table 4 Result of translation performance for each corpus

コーパスの種類	翻訳性能 (BLUE値)
原文コーパスのみ	7.21
本方式での拡張コーパス	22.03

(注4) 言語表現としての自然さを単語の系列の出現確率で表したものを。

成方式[10]の生成効率1.45（倍）と比較して大幅に向上している。

6. 実証実験システム

現在、対話SPFを利用して開発した対話サービスの利便性の検証を目的とした実証実験を行っている。2つの対話実証実験システムについて述べる。

6.1 FAQ検索実証実験システム

一問一答型対話処理エンジンを用いて、当社法務部門向けの「NDA契約締結に関するFAQ検索システム」を開発した。NDA契約に関する質問文を入力すると、質問内容に対応する回答が表示される。約150件のFAQを対象に検索できる。2017年10月の運用開始から本稿執筆時点（2017年12月）までに累計約3000件の検索に回答した。

本実証実験システムは、一問一答型対話処理エンジンが機械学習ベースの方式であるという特徴を活かし、運用中にユーザー自身が質問文を学習コーパスに追加できる機能をもつ。これによりシステムの運用を通じて、多様な表現の質問文を獲得しFAQの検索性能を向上できるという特長をもつ。実際、ユーザーからは「質問の表現が違っていても回答が精度良く出力される」といったフィードバックが得られている。

6.2 レシピ提案対話実証実験システム

意図推定型対話処理エンジンを用いて、レシピ提案対話の実証実験システムを開発した。一般公開にて運用中（注5）である。

本システムはチャット形式の対話サービスである。システムからの「どんな料理が好みですか？」という問いかけに対して「中華がいい」「豚肉が使いたい」などの発話によりユーザーの要望を具体的に伝えることができる。具体的な要望がない場合は、システムからの「炒（いた）め物でいいですか？」などの問いかけに「はい」、「いいえ」、「どちらでも良い」と答えることで、ユーザーの好み（ユーザー意図）に合ったレシピを絞り込んで提案することができる。第6図に本システムでの対話の様子を示す。

本システムの運用開始（2017年1月）から4週間で得られたユーザーの評価の一部を紹介する。チャットの最後にユーザーの満足度を問うアンケート機能を実装し、ユーザーの回答を収集した。

ユーザーの回答結果を第5表に示す。本実験システム



第6図 レシピ提案対話実証実験システム（操作画面）

Fig. 6 Prototype for dialog-based recipe recommendation (GUI)

第5表 レシピ提案対話のユーザー満足度に関する回答

Table 5 Response regarding user satisfaction with recipe recommendation dialog

質問内容	回答（そう思う）
Q1：好みにあったレシピが提案できたか？	60.3%
Q2：また使おうと思うか？	70.3%
自由回答（好意的な回答）	
<ul style="list-style-type: none"> ・思いついた言葉でレシピが探せる ・チャットを通じて徐々にレシピを絞り込んでいける 	
自由回答（課題を指摘する回答）	
<ul style="list-style-type: none"> ・味付けや温度の要望が理解されない ・「あっさり」などの感覚的な表現が理解されない 	

では提案できるレシピの候補数が500未満で少ないことがQ1の評価に影響したと思われる。しかしQ2で「また使おうと思う」と回答したユーザーの割合が約7割であったことから、利便性の観点からは対話サービスの品質として比較的良好であったと考えている。

ユーザー回答でも指摘されている「理解されない発話内容の課題」に対しては、現在、食関連の知識ベース構築[11]による発話理解能力向上の取り組みを産総技術総合研究所・京都大学と共同で進めている。機械学習ベースの対話処理方式である特徴を活かし、構築した知識ベースから発話を機械的に生成し学習コーパスに追加することで、効率的に発話理解性能を向上できる見込みである。

7. 対話サービスプラットフォームの効果

対話SPFを開発したことによる効果について述べる。同じ条件での比較ではないが、対話SPF開発の以前と以後での対話サービス開発事例の比較を第6表に示す。あくまで参考値であるが、対話サービス当たりの開発工数

（注5） <https://www.weekcook.jp/trial/cookchat/lp.html>, 参照 Apr. 20, 2018.

第6表 対話SPF開発前後での対話サービス開発事例の比較

Table 6 Comparison of dialog service development before/after Dialog SPF

	対話SPF開発以前	対話SPF開発以後
対話サービス	家電制御	FAQ検索, レシピ提案
開発工数	約120人月 (約20人×6箇月)	約24人月 (4人×約6箇月)
対話サービス数	1 (実運用なし)	4 (実運用中)
他用途展開	なし	複数検討中

で見ると約20倍の効率化を達成したことになる。

対話SPF開発以前では1つの対話サービスに特化したソフトウェア構造であったため、他用途への展開もできなかった。一方、対話SPF開発以後では、対話SPFの共通の対話処理の枠組みが活用可能で、上記の実運用中以外に他の複数の対話サービスの開発が検討中である。ソフトウェア開発投資の面から見ても従来の効率を大幅に改善したと言える。

8. まとめ

対話サービスプラットフォームの開発により、共通の対話処理の枠組みを実現し、対話サービス開発を効率化した。機械学習ベースの対話処理方式採用により、対話処理性能向上の効率化も実現した。

今後は、自動コーパス生成を学習コーパス生成に適用しさらなる開発効率の向上を目指す。さらに、対話サービスの運用を通じて蓄積された対話データから有効な発話文を抽出し、学習コーパスへの追加による対話処理性能向上、評価用コーパスへの追加によるテスト環境の充実を図っていく。これらの取り組みにより対話サービスの開発から運用までを支えるプラットフォームの確立を目指す。

本対話SPFの開発に際し、ソフトウェア品質担保の観点から開発マネジメントに関して、多大な支援を頂いたSLAD (株) の井上健司氏に感謝する。

参考文献

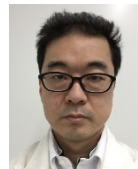
- [1] 山上勝義 他, “対話サービスプラットフォームの開発,” 人工知能学会全国大会 (第31回) 論文集, 2017.
- [2] 野澤哲生, “音声対話が世界を揺るがす,” 日経エレクトロニクス, 2016年8月号, pp. 26-50, 2016.
- [3] Amazon.com, Inc., “チュートリアルとテンプレート|Alexa Skill Kit,” <https://developer.amazon.com/ja/alexa-skills-kit/start>, 参照 Apr. 19, 2018.
- [4] H. Shi et al, “Convolutional Neural Networks for Multi-topic Dialog State Tracking,” Dialogue with Social Robots, Springer, pp. 451-463, 2016.

- [5] S. Kim et al, “The Fourth Dialog State Tracking Challenge,” Dialogue with Social Robots, Springer, pp. 435-449, 2016.
- [6] 遠藤充 他, “発話とコマンドの系列を制御する対話エージェント,” 人工知能学会全国大会 (第31回) 論文集, 2017.
- [7] T. Ushio et al, “Recurrent convolutional neural networks for structured speech act tagging,” Spoken Language Technology Workshop (SLT), 2016 IEEE, 2016.
- [8] S. Kim et al, “The fifth dialog state tracking challenge,” Spoken Language Technology Workshop (SLT), 2016 IEEE, 2016.
- [9] 藤原菜々美 他, “自動コーパス生成による少量対訳コーパスからの統計的機械翻訳,” 言語処理学会第22回年次大会発表論文集, pp.219-222, 2016.
- [10] M. Yuval, et al, “Distributional Phrase Paraphrase Generation for Statistical Machine Translation,” ACM Trans. Intell. Syst. Technol. 4, 3, Article 39, 2013.
- [11] 清丸寛一 他, “料理レシピとクラウドソーシングに基づく基本料理知識ベースの構築,” 言語処理学会第24回年次大会発表論文集, pp.662-665, 2018.

執筆者紹介



山上 勝義 Katsuyoshi Yamagami
テクノロジーイノベーション本部
パイオニアリングリサーチセンター
Pioneering Research Center, Technology Innovation Div.
(2017年3月より (国研) 産業技術総合研究所 出向中)



遠藤 充 Mitsuru Endo
ビジネスイノベーション本部
AIソリューションセンター
AI Solution Center, Business Innovation Div.



史 宏杰 Honjie Shi
ビジネスイノベーション本部
AIソリューションセンター
AI Solution Center, Business Innovation Div.



牛尾 貴志 Takashi Ushio
ビジネスイノベーション本部
AIソリューションセンター
AI Solution Center, Business Innovation Div.



藤原 菜々美 Nanami Fujiwara
ビジネスイノベーション本部
AIソリューションセンター
AI Solution Center, Business Innovation Div.



堀井 則彰 Noriaki Horii
ビジネスイノベーション本部
AIソリューションセンター
AI Solution Center, Business Innovation Div.