

ま え が き

自然言語で対話する能力は、人間の知能の中でも最も重要なものの一つと考えられている。計算機科学の父といわれる数学者アラン・チューリングは、機械が本当に知能を持っているかどうかを調べるテストとして、チューリングテストと呼ばれるものを考案した。これは、相手が機械か人間かわからない状況で、人間に機械と自然言語で対話をさせるといふものであった。そして、対話を行った人間が、機械と対話しているのか人間と対話しているのか判別できなければ、その機械は知能を持っているとよいといふものである。このように、対話能力は知能の代名詞にもなっている。一般の人にとっても、機械と自然言語で対話をしたいという期待がある。これはSFに出てくる機械やロボットが人間と対話する能力を持っていることが多いことからわかる。

そのような期待に応え、古くから、人間と対話するシステムの研究が進められてきた。近年では、対話機能を持つ電話応答システム、ロボット、スマートフォンのアプリケーション、カーナビゲーションシステムが開発され、実用に供されている。しかしながら、対話システムの基本的な技術がどのようなものであるかは、一般の人のみならず、関連分野の研究者にもあまりイメージが持てないのではないだろうか。この理由の一つは、対話システム技術の解説や研究発表が、自然言語処理、音声情報処理、ヒューマンインタフェース、ロボット工学などの分野に分散していることにあると考えられる。

本書はこのような状況を踏まえて執筆したものであり、対話システムの主要技術と、それが個々の種類の対話システムの中でどのように用いられているかを解説することを目的とした。対象とする読者は、これから対話システムの研究開発を始めようとしている方、および、関連分野の研究・開発者で、対話システムに興味をお持ちの方である。特段の前提知識は仮定しないように努めたが、

情報科学や人工知能の基礎知識があると理解しやすいと考えている。本書を読むことで、簡単な対話システムを構築できるようになったり、最新の論文を理解できるようになったりすることを目標とした。対話システム研究者にとって常識となっていることは、最新の論文には書かれなことが多いが、そのような事柄もなるべく丁寧に解説するようにした。

本書の構成は以下のとおりである。まず、1章で対話システムを定義し、対話システムの研究および実用化の歴史について述べる。

2章では、人間どうしの対話の分析に基づく対話のモデルを解説する。すべての対話システムが人間どうしの対話のモデルに基づいているわけではないが、自然言語処理や音声情報処理が品詞や音素などの言語学の概念に基づいているのと同じように、対話システムの技術も対話のモデルが基礎になっている。また、対話のモデルは対話システム研究者共通の常識となっており、技術論文を理解するのに不可欠である。技術系の読者は読み進めづらい場合があるかもしれないが、その場合は、3章を先に読み始め、必要に応じて2章に戻っていたきたい。

3章では、対話システムの基礎技術を解説する。さまざまな対話システムに共通な技術をより統一的な方法で解説するように心がけた。技術の解説だけでなく、対話システム研究に共通の課題は何で、その課題がどのように取り組まれているのかということにも重点を置いた。

4章では、対話システムを設計し構築する手法について述べる。対話システムが扱うタスクやモダリティのバリエーションに応じてどのように対話システムを設計すべきかや、具体的な対話システム構築のプロセスについて解説する。

5章では、対話システムに関する統計的手法や、人間どうしの対話に近い自然な対話の実現など、発展的な技術を解説する。マルチモーダル対話システムや対話ロボット、複数のユーザと対話をするマルチパーティ対話システムなどについても説明する。

本書の執筆は以下の分担で行ったが、著者全員でコメントし合うことで統一のとれたものにするように心がけた。いくつかの節は共同で執筆した。専門用

語の定義や対話システムの構成などは、研究者間でもコンセンサスが得られていないものがあるが、著者間で議論することで、なるべく一般的に用いられているものを示すように努めた。

- 中野 幹生 1, 3.1, 3.2, 3.3, 3.5, 4.1, 4.2.1, 4.2.2, 4.2.4, 4.2.5, 4.3.1, 4.3.4, 4.5.1, 4.5.2, 4.5.2 のコラム (2), 4.5.3, 5.1.1, 5.3 のコラム, 5.7.1, A.1, A.3, A.4, A.5
- 駒谷 和範 1.2, 3.4, 4.4, 4.5.4, 5.1.3, 5.2, 5.3, 5.6.1 [2]
- 船越 孝太郎 2, 3.4, 5.1.2, 5.4, 5.6.1 [1], 5.7.2, A.2
- 中野 有紀子 4.2.3, 4.3.2, 4.3.3, 4.5.2 のコラム (1), 5.5, 5.6.1 [3], [4], 5.6.2

対話システム技術の発展のためには、実際に対話システムを作って、ユーザーに使ってもらい、課題を発見するというサイクルを回すことが重要である。多くの方々が対話システムの構築に携わることで、対話システム技術がより早く発展し、社会の役に立つと考えられる。本書がその一助になれば幸いである。

本書の出版にあたっては多くの方にお世話になった。中でも、監修をさせていただいた奥村 学先生、編集作業にご尽力いただいたコロナ社の各位、草稿を詳細に読んでコメントをくださった北岡 教英氏、東中 竜一郎氏、翠 輝久氏、杉山 貴昭氏に心より感謝する。

2014 年 12 月

執筆者一同

目 次

1. 対話システムの概要

1.1 対話システムとは	1
1.2 対話システムの類型	4
1.2.1 入出力のモダリティ	6
1.2.2 達成すべき目標の有無・種類	7
1.2.3 対話のドメイン	9
1.2.4 対話参加者の数	10
1.3 対話システム研究の歴史	11
1.4 対話システム実用化の歴史	16

2. 対話のモデル

2.1 対話のモデルとは	20
2.2 対話の基本構造	22
2.2.1 発 話 行 為	23
2.2.2 談 話 構 造	25
2.2.3 隣接ペア・交換	29
2.2.4 話 者 交 替	31
2.2.5 発 話 単 位	34
2.3 共通基盤と基盤化	36
2.3.1 共 通 基 盤	36

2.3.2	貢献に基づく基盤化のモデル	39
2.3.3	基盤化アクトに基づく基盤化のモデル	45
2.4	プランに基づく対話モデル	47
2.4.1	プランと問題解決	50
2.4.2	共有プランと協調的問題解決	56
2.4.3	談話義務	58
2.4.4	BDIモデル	59
2.5	対話の背景構造	60
2.5.1	共同注意	60
2.5.2	参加構造	61
2.5.3	同調	62

3. 対話システムの構成と処理の概要

3.1	対話システムのアーキテクチャ	66
3.2	対話管理の基礎概念	73
3.2.1	内部状態が保持する情報	74
3.2.2	内部状態更新処理	77
3.2.3	行動選択処理	80
3.2.4	内部状態更新処理・行動選択処理が駆動されるタイミング	82
3.3	対話管理の基礎的な手法	83
3.3.1	ネットワークモデルに基づく対話管理	83
3.3.2	フレームに基づく対話管理	86
3.3.3	アジェンダに基づく対話管理	91
3.4	対話の主導権	92
3.4.1	システム主導	93
3.4.2	ユーザ主導	94

3.4.3 混 合 主 導	94
3.5 入力理解・出力生成	98
3.5.1 入力理解・出力生成の概要	98
3.5.2 言 語 理 解	100
3.5.3 言 語 行 動 生 成	109

4. 対話システムの設計と構築

4.1 対話システムのデザイン	114
4.2 対話のタスクと対話管理	115
4.2.1 フォームフィリング対話システム	116
4.2.2 データベース検索対話システム	116
4.2.3 説明対話システム	117
4.2.4 非タスク指向型対話システム	120
4.2.5 マルチドメイン対話システム	122
4.3 さまざまなモダリティの対話システム	125
4.3.1 音声対話システム	126
4.3.2 マルチモーダル対話システム	133
4.3.3 バーチャル対話エージェント	137
4.3.4 対 話 ロ ボ ッ ト	143
4.4 エラーハンドリング	147
4.4.1 基盤化の必要性	148
4.4.2 確認要求の方法	150
4.4.3 確信度に基づく確認要求	153
4.5 対話システムの開発と評価	159
4.5.1 対話システムの開発プロセス	159
4.5.2 対話システムの効率的な構築	161

4.5.3 ユーザスタディとデータ収集	169
4.5.4 対話システムの評価	171

5. 対話システムの発展技術

5.1 統計的意図理解	180
5.1.1 候補列挙法	181
5.1.2 ベイジアンネットワークに基づく意図理解	183
5.1.3 動的ベイジアンネットワークに基づく意図理解	189
5.2 適応的な対話管理・応答生成	192
5.2.1 対話戦略のオンライン変更	192
5.2.2 ユーザモデル	196
5.2.3 強化学習	201
5.2.4 ヘルプ生成	204
5.2.5 ユーザの振舞いのモデル	207
5.3 問題のある状況の検出	209
5.3.1 問題検出	209
5.3.2 訂正発話検出	211
5.4 話者交替	212
5.4.1 柔軟な話者交替	213
5.4.2 相づちの生成	217
5.5 マルチモーダル・マルチパーティ対話技術	218
5.5.1 マルチモーダルな話者交替	218
5.5.2 マルチパーティ対話システムにおけるフロアマネジメント	220
5.5.3 受話者推定	221
5.5.4 共通基盤確立におけるマルチモダリティ	222
5.6 人間と対話システムのインタラクション	226

x 目 次

5.6.1	システム表現の拡張	226
5.6.2	ユーザとの関係性の構築	230
5.7	今後の重要技術	230
5.7.1	対話システムの知識獲得	231
5.7.2	ロボットの移動と対話管理	234
付 録		237
A.1	対話システムのための機械学習	237
A.1.1	分 類	237
A.1.2	系列ラベリング	238
A.2	ベイジアンネットワークの基礎	239
A.3	対話システム構築のためのツール	241
A.4	対話システム研究開発のための対話コーパス	243
A.5	文献調査案内	244
引用・参考文献		246
索 引		275

1

対話システムの 概要

対話システムとは何だろうか。対話システムにはさまざまな技術分野との接点があるため、対話システムという用語の指す意味が人によって異なる。また、分野によっては、対話システムを別の用語で呼んでいる場合もある。本章では、まず、1.1節で対話システムとは何かを定義し、本書で扱う技術の範囲、およびほかの研究分野との関連を述べる。

1.2節では、対話システムの類型について述べる。本書では、ここで述べる類型が本書の説明の基礎となる。

1.3節では、対話システム研究の歴史を振り返る。対話システム研究には数十年の長い歴史があり、その中でさまざまな対話システムが作られ、問題点が発見され、新しい技術が開発されてきた。その過程では、人文科学における人間どうしの対話の分析的研究の成果や、計算機科学、人工知能分野のさまざまな技術が取り入れられている。

対話システムは、学術的研究のみならず、実用化・製品化が行われている。1.4節では、学術的研究の成果を実用化・製品化に結び付けてきた過程を概観する。

1.1 対話システムとは

自然言語でコミュニケーションを行い、情報を授受することを対話 (dialogue) という。人間と対話する機械またはソフトウェアを対話システム (dialogue system) と呼ぶ。

対話システムは人工知能 (artificial intelligence) である。それは、自然言語を用いて対話を行うということが、人間の知能に相当するものを必要とするか

2 1. 対話システムの概要

らである。もちろん、現状の技術では、人間と同じような対話を行うことはできない。しかし、対話システム研究が目指すのは、人間と同等の知的な処理を機械に行わせることである。自律性、および人間や他の人工知能との相互作用に焦点を当てるとき、人工知能をエージェント (agent) と呼ぶことがある。また、対話システムも対話エージェント (dialogue agent) と呼ぶことがある。

対話システムは自然言語を用いてコミュニケーションを行うが、自然言語のみを用いてコミュニケーションを行うものだけを対話システムと呼ぶのではない。コミュニケーション手段の一つに自然言語が用いられていれば対話システムと呼ぶ。例えば、スマートフォン上の対話システムは画面上にシステムからの情報を表示するし、対話ロボットはジェスチャを用いて情報を伝達することがある。さらに、システムがカメラ画像から人間の顔の向きや視線を認識して、だれに話しているかや、話し始めようとしているかどうかなど、人間の意図の推定に役立てることもある。

対話とよく似た用語に会話 (conversation) がある。両者の区別は曖昧^{あいまい}であるが、対話は目的がはっきりしているものを指し、会話は話をする事自体を目的とするものを指す場合が多い。したがって、対話システムというときには、予約や情報検索など目的があって用いるシステムを意味し、会話システム (conversational system) というときには、雑談のように目的のないやりとりを行うシステム (1.2.2 項で述べる非タスク指向型対話システム) を意味する場合がある。なお、二人で話をする場合を対話、三人以上の場合を会話といい、両者を区別する場合もある。その場合、会話システムは、二人以上のユーザと話すシステムや、二つ以上のシステムが一緒になってユーザと話すようなもの (1.2.4 項で述べるマルチパーティ対話システム) を指す。

対話システムと似た用語に対話型システム (interactive system) がある。ヒューマンコンピュータインタラクションの研究において、人間の操作に対して反応するシステムの意味で用いられるが、自然言語を用いるとは限らないため、対話システムとは定義が異なる。

対話システムと類似したシステムとして、質問応答システム (question-answering

system)がある。単純な質問応答システムでは、ユーザが発話して[†]（またはキーボードから一文を入力して）システムが答える、一問一答式のやりとりが行われる。これに対し、対話システムは、一度のやりとりだけではなく、複数回のやりとりを行う。したがって、対話システムは、対話の履歴 (dialogue history) を利用する。対話においては、発話の意味が対話の履歴によって変わる。また、どのような発話をすべきかは対話の履歴に応じて変える必要がある。

例えば、つぎのようなやりとりを考える。ここで、U はユーザ発話、S はシステム発話である。U、S のあとの数字は発話番号である。本書を通じてこの表記を用いる。

U1: 品川のベトナム料理店を教えてください
 S1: 品川のベトナム料理店は5軒あります。読みあげますか？
 U2: 五反田は？

ここで、U2の発話は「五反田のベトナム料理店を教えてください」という意味であると考えられる。また

U1: 個室がある大崎のイタリアン
 S1: 個室がある大崎のイタリア料理店は見つかりませんでした。ほかにご用件はございますか？
 U2: 五反田は？

のU2は「五反田の個室があるイタリア料理店を教えてください」という意味であると考えられる。このように発話の意味はそれまでの対話の内容によって変わってくる。

システム発話の内容も対話の履歴によって変えたほうがよいことがある。例えば、上記の例で、S1ですでに個室とイタリア料理店を理解したことはユーザに伝えているので、U2に対する応答は「個室がある五反田のイタリア料理店は8軒あります」よりも「五反田には8軒あります」のほうが簡潔でよい。

対話の履歴の利用の利点として、複数のやり取りからユーザの意図を理解で

[†] 発話の定義は2章の冒頭で行う。

4 1. 対話システムの概要

きるといふ点も挙げられる。以下の対話では、一度のユーザ発話からだけではユーザに情報提供を行うのに十分な情報が得られてなかったため、システムは情報要求を行っている。

U1: 個室がある中華料理店を教えてください
S1: どのエリアですか？
U2: 五反田

システムは U2 を理解した時点で、その時点までの履歴と合わせてユーザの意図が「個室がある五反田の中華料理店を調べる」ことであることがわかる。さらに、ユーザ発話が理解できなかつたときに聞き返したり確認したりするなど、履歴を用いて対話的にユーザの意図理解を行って、適切な発話を行うことができる。

対話の履歴に限らず、発話の意味に影響を与える要因を文脈 (context) と呼ぶ。文脈は、対話の履歴以外に、対話の行われている場所、時間、対話者の社会的関係など、さまざまな要因を含む。これらの要因が変化すれば、発話の解釈や望ましいシステム発話も異なる。対話の履歴以外の文脈のことを状況 (situation) と呼ぶ。どのように対話の履歴や状況を表現し利用するかは、対話システム研究の中心的課題であるといえる。

1.2 対話システムの類型

対話システムは多種多様であるため、特定の種類の対話システムに言及するときに便利のように、いくつかの観点から分類されている。本節では、以下の四つの観点による分類に基づき説明する。

- ① 入出力のモダリティ
- ② 達成すべき目標の有無・種類
- ③ 対話のドメイン
- ④ 対話参加者の数

表 1.1 に、さまざまなタイプの対話システムを、本節で述べる観点で分類した結果を示す。

表 1.1 さまざまなタイプの対話システム

システム例	入力 モダリティ	出力 モダリティ	抽象タスク	タスクドメイン	対話参加者数
ELIZA ³⁷⁵⁾	テキスト	テキスト	(非タスク指向)	オープン	2
SHRDLU ³⁸²⁾	テキスト	テキスト, 物体画像	コマンド& コントロール	クローズド (積木の移動)	2
Hearsay-II ⁸⁶⁾	音声	音声	データベース 検索	クローズド (文献抄録検索)	2
京都市バス運行 情報案内システム 197)	音声	音声	フォームフィ リング	クローズド (バス運行情報)	2
MATCH ¹⁶³⁾	音声, ペン 入力	音声, 地図 画像	データベース 検索	クローズド (街案内)	2
長尾らのシス テム ²⁵²⁾	音声	音声, 顔画 像	協調的問題 解決	クローズド (料理)	3
Rea ⁵⁷⁾	音声	音声, エー ジェントの 画像 (表情, 視線, ジェ スチャ)	説明	クローズド (不動産案内)	2
Wabot-2 ³¹⁴⁾	音声	音声, ロボッ トの移動	コマンド& コントロール	クローズド (移動)	2
Jijo-2 ¹⁷⁾	音声, 画像, 障害物位置 情報など	音声, ロボッ トの移動	データベー ス検索 + コ マンド&コ ントロール	クローズド (マルチドメイン)	2
ROBITA ²³¹⁾	音声, 画像	音声, ロボッ トの身体表 現	(非タスク指 向)	オープン	3

1.2.1 入出力のモダリティ

対話システムは、入出力のモダリティ(modality)により分類することができる。モダリティとは、人間が情報伝達を行う際の様式で、発話、表情(facial expression)、ジェスチャ(gesture)、視線(eye gaze)などが含まれる[†]。

入力のモダリティとしては、キーボードからのタイプ入力や音声入力がおもに用いられている。音声入力の場合には、音声をテキストに変換する音声認識が用いられる。これに加え、表情、身振り・手振り、視線などの情報を画像認識を用いて取得するシステムもある。

また、人間どうしのコミュニケーションのモダリティとは異なるが、ポインティングデバイス(マウスによるクリックや、タッチパネルへの接触)を入力として併用することもある。このようなシステムは、画面上のオブジェクトを指し示しながら行った発話を、発話内容とポインティング情報を統合して理解する。例えば、ユーザがディスプレイ上の地図の2点を順にクリックしながら「ここからここまでの距離を教えて」と発話したような場合にユーザの意図を理解する。

さらに、ユーザが用いている端末の位置情報や、ユーザとロボットが同時に見ている物体の情報などが、入力モダリティとして用いられる。

出力のモダリティとしては、テキスト出力、すなわち、文をテキストで画面に出力することや音声出力が用いられている。これに加え、画面上のキャラクターエージェントや、ロボットを用いるシステムでは、表情、身振り・手振り、視線などのモダリティを用いる場合がある。

入出力のモダリティがテキストだけのものをテキスト対話システム(text dialogue system)と呼ぶ。また、入出力のモダリティが音声だけのものを音声対話システム(spoken dialogue system)と呼ぶ。キャラクターエージェントを用いる対話システムをバーチャル対話エージェント(virtual dialogue agent)

[†] モダリティと似た用語にメディア(media)がある。音声、テキスト、画像などの入出力のチャンネルがメディアと呼ばれるが、これらもモダリティと呼ぶ場合もある。本書では特に区別をせずモダリティと呼ぶ。

または会話エージェント (conversational agent) などと呼ぶ (図 1.1)。また、対話機能を持つロボットを対話ロボット (dialogue robot) または会話ロボット (conversational robot) と呼ぶ。



図 1.1 バーチャル対話 (会話) エージェントの例

出力でも、入力と同じように、人間どうしのコミュニケーションとは異なるモダリティも用いられている。例えば、写真や地図などの画像を端末画面に表示させることがある。

テキスト対話システムや音声対話システムなど、入出力のモダリティが一つのものユニモーダル対話システム (unimodal dialogue system) と呼び、複数ものをマルチモーダル対話システム (multimodal dialogue system) と呼ぶ。

1.2.2 達成すべき目標の有無・種類

対話システムは、ユーザが対話を通じて達成したい具体的な目標 (goal) を持っているかどうかという観点からも分類できる。ここで目標とは、ユーザが対話を通じて達成しようとしていることを指す。例えば、ホテル検索システムでは、ユーザが持っているであろう条件を満たすホテルを見つけることが目標である。なお、上記の議論からは、この概念を目標指向やゴール指向 (goal-oriented) と呼ぶのが適当であるが、慣例に従い、本書ではこれをタスク指向 (task-oriented) と呼ぶ。

目標を持った対話を行うシステムをタスク指向型対話システム (task-oriented dialogue system) と呼び、それ以外のシステムを非タスク指向型対話システム (non-task-oriented dialogue system) と呼ぶ。タスク指向型対話システムが達成できる目標の集合を、そのシステムのタスク (task) と呼ぶ。

8 1. 対話システムの概要

タスク指向型対話システムは、タスクの種類によってさらに細分化できる。例えば、システムとユーザの間におもにどのような情報の流れがあるか、つまり、どちらがおもに目標達成に必要な知識を持っているかで、概ね以下の5種類の抽象タスク¹³⁾に分類できる(表1.2)。

- フォームフィリング型
- コマンド&コントロール型
- データベース検索型
- 協調的問題解決型
- 説明型

表 1.2 情報の流れと抽象タスク

情報の流れ	抽象タスク	タスク例
ユーザ → システム	フォームフィリング	テレフォンショッピング オンライン取引
	コマンド&コントロール	ロボット制御
ユーザ ↔ システム	データベース検索	書籍注文 文献検索
	協調的問題解決	列車運行計画立案
ユーザ ← システム	説明	地理案内 操作マニュアル

一般に、空欄がある文書のことをフォームと呼び、その空欄を埋めることをフォームフィリング (form-filling) と呼ぶ。このようなタスクを行う対話システムをフォームフィリング型の対話システムと呼ぶ。乗換案内や航空券の予約など、Web ページのフォームにテキストをタイプすることで目標を達成できるようなタスクがこれに相当する。この型の対話システムでは、システムがフォームの内容のうち必須のものを順に尋ねることで目標を達成できる。

コマンド&コントロール (command and control) は、ユーザがシステムに向かって命令をし、システムがその命令を実行する。基本的に、目標達成に必要な新情報はユーザからシステムに向けて流れる。一方で、命令が曖昧な場合やシステムにとって理解不能な場合にシステムが聞き返したり、命令の実行を完了した場合や実行不可能であった場合にそのことをシステムからユーザに伝

索引

【あ】	
相づち	32, 217, 226
曖昧性	234
アクションスケジューリング	142
アクセント句	36
アクト	30
アジェンダ	91
アジェンダに基づく対話管理	92
アノテーションスキーマ	24
アフォードダンス	204
アブダクション	64
誤り検出	148
暗黙の確認	150
【い】	
言い淀み	35
移行適格場所	32
意図	30, 48, 57, 59
意図構造	27
意図理解	71
意味表現	68
意味フレーム	69
意味文法	101
印象評価	172
インタラクティブ	
アラインメント	63
イントネーション	36
イントネーション句	36
イントネーションユニット	36
韻律	63, 131
韻律情報	215

【う】	
韻き	32
【え】	
エージェント	2, 59
エラーハンドリング	13, 148
エンタテインメント	62, 207
【お】	
応答	30
応答者	47
オープンドメイン	9
オブジェクト指向	
プログラミング	166
オペレータ	51
重み付き有限状態	
トランスデューサ	103
音響モデル	128
音声区間検出	126
音声対話システム	6, 126
音声認識	99
オントロジ	70, 183, 206, 232, 235
【か】	
開始	30
開始者	47
回復戦略	148
外部連携部	72
会話	2, 19
会話エージェント	7
会話サービスロボット	144
会話システム	2

会話分析	19, 30
会話ロボット	7, 143
学習データ	238
確信度	71, 153
確認要求	148
格率	48
隠れマルコフモデル	107, 121, 238
含意	48
環境理解部	72
関係の格率	49
頑健な言語理解	103
感情	131
間接発話内行為	24
間接発話行為	24, 51
願望	59
関連性	50
【き】	
記号接地	37
記述表現	77
記述論理	235
基盤化	37, 234
—の度合	157
基盤化アクト	45
基盤化状態	75
キーフレーズ	106
キーフレーズ抽出	106
客観評価	172
強化学習	201
協調的問題解決	9, 56
協調の原理	48
共通基盤	37
共同活動	56

——のデザイン	115	テストデータ	238	バックエンド	72
対話状態	67	データベース検索	9, 116	バックチャンネル	32
対話状態追跡	180	伝達行為	24	発語行為	23
対話ムーブ	24			発語内行為	23
対話ロボット	7, 15, 143			発語内効力	24
他者開始修復	42	【と】		発語媒介行為	23
他者修復	42	統計の文生成	112	発話	19, 23
タスク	7	同調	32, 62, 207	発話区間検出	126, 214
タスク指向	7	盗聴者	61	発話検出	126
タスク指向型対話システム	7	動的ベイジアン		発話検証	154
タスクプランニング部	72	ネットワーク	189, 241	発話行為	13, 23
ターン	31	頭部姿勢	138	発話行為タイプ	24
単語誤り率	130	ドメイン	9	発話修復	42
単語辞書	130	ドメイン外	185	発話選択	71
談話	19	ドメイン外発話	123	発話単位	35, 45, 70, 128
談話義務	30, 58, 75	ドメイン選択	123	発話分類	106
談話構造	27	ドメイン追跡	124	パラ言語情報	36
対話制御部	71	ドメインプラン	53, 75	判別器	237
談話セグメント	26, 79				
談話セグメント目的	27, 76	【な】		【ひ】	
談話表示理論	64	内部状態	67	引込み	32, 63
談話プラン	53, 75	内部状態更新	71	非言語情報	36, 139, 226
談話分析	30	内容生成	71	非タスク指向型	
談話目標	75, 76	長い単位	36	対話システム	7, 120
談話ユニット	45			ビッグファイブ	227
談話理解	71	【に】		標識	30
		二重課題法	172	表情	6
【ち】		入力理解部	67	表層生成	53, 72
逐次生成	83	認知発達ロボティックス	232		
逐次理解	82			【ふ】	
知識	115	【ね】		フォームフィリング	8, 116
チャットボット	12	ネットワークモデル	83	不気味の谷	227
注意状態	27, 75, 79	【の】		副対話	26, 42, 54, 76, 95
中国語の部屋	174	能動学習	233	含み	48
中断	33	【は】		符号体系	47
チューリングテスト	174	背景雑音	127	付帯的コミュニケー	
聴取者	61	漠然性	234	ション	42, 187
直示表現	77	バージョン	33	部分解析	103
		パーソナリティ	227	部分観測マルコフ	
【て】		バーチャルエージェント	137	決定過程	121, 190, 203
提示	40, 234	バーチャル対話		ブラックボックス	171
テキスト対話システム	6	エージェント	6	プラン	13, 27, 50, 59
デコーダ	130			——に基づくモデル	50

engagement	223	107, 121, 238	Julius	242
entailment	48	HMM	Juman	242
entrainment	62, 207			
environment understanding		[I]	[K]	
module	72	illocutionary act	kappa coefficient	176
error detection	148	illocutionary force	κ 值	176
exchange	30	implicature	keyphrase	106
explanation	9	implicit confirmation	keyphrase extraction	106
explicit confirmation	150	incremental generation	Kinect SDK	242
eye gaze	6	incremental understanding	knowledge	115
			KNP	242
[F]		indirect illocutionary	KyTea	242
facial expression	6	act		
false alarm	154	indirect speech act	[L]	
finite-state automaton		inference model	language and behavior	
model	83	information state	generation	99
finite-state transducer	101	information state	language generation	99
first pair part	29	update	language model	128
floor	33	initiation	language understanding	99
floor management	219	initiative	LDC	243
focus space	27	Initiator	Let's Go	243
follow-up	30	input understanding	likelihood	154
form-filling	8	ule	likelihood ratio	155
frame	69	intention	Likert scale	172
frame-based dialogue		intentional structure	Linguistic Data	
management	86	intention understanding	Consortium	243
FS	27	interaction	lip-sync	143
FST	101, 135	interactive alignment	listener	61
		interactive system	locutionary act	23
[G]		internal state	Loebner Prize	17, 174
gesture	6	internal state update	long-utterance unit	36
goal	7, 59	interruption		
goal-oriented	7	intonation	[M]	
grounding	37	intonational phrase	marker	30
grounding act	45	intonational unit	Markov decision	
GUS	12	IOB2 法	process	202
			MATCH	5
[H]		[J]	maxim	48
head pose	138	Jijo-2	maxim of manner	49
Hearsay-II	5, 14	joint action	maxim of quality	49
help message	205	joint activity	maxim of quantity	49
hesitation	35	joint attention	maxim of relation	49
hidden Markov model		joint focus of attention	MDP	202

- | | | | | | |
|----------------------------|--------|-----------------------------|---------------|-----------------------------|------------|
| Mecab | 242 | other-initiated repair | 42 | | |
| media | 6, 229 | other-repair | 42 | | [R] |
| mel-frequency cepstral | | output generation module | 72 | ratified participant | 61 |
| coefficients | 127 | | | Rea | 5 |
| mental state | 67 | out-of-domain | 185 | reactive agent | 59, 88 |
| MFCC | 127 | out-of-domain utterance | 123 | reactive planning | 59, 88 |
| misunderstanding | 42 | overhearer | 61 | reactive tokens | 32 |
| modality | 6 | | | real life | 229 |
| morphological analysis | 99 | | | recipe | 57 |
| move | 30 | [P] | | reference resolution | 77 |
| multi-domain dialogue | | pair type | 29 | referent | 77 |
| system | 10 | palmkit | 242 | referring expression | 77 |
| multi-party dialogue | | para-linguistic information | 36 | reinforcement learning | 201 |
| system | 10 | PARADISE | 175 | relevance | 50 |
| multilogue | 61 | partially-observable | | repair | 42 |
| multimodal dialogue | | Markov decision | | repair strategy | 148 |
| system | 7, 133 | process | 121, 190, 203 | Responder | 47 |
| multiparty dialogue | 61 | partial parsing | 103 | response | 30 |
| mutual belief | 37 | participation | 61 | rhetorical relation | 26 |
| | | participation structure | 61 | ROBITA | 5 |
| [N] | | perlocutionary act | 23 | robust language | |
| named entity | 107 | personality | 227 | understanding | 103 |
| named entity extraction | 107 | personalization | 197 | [S] | |
| | | plan | 50 | SD 法 | 172 |
| non-task-oriented dialogue | | planning | 50 | SDRT | 64 |
| system | 7 | plan recognition | 50 | Second Life | 242 |
| non-understanding | 42 | plan stack | 54 | second pair part | 29 |
| non-verbal information | 36 | plan synthesis | 50 | segmented discourse | |
| <i>N</i> -best | 130 | plan-based model | 50 | representation | |
| <i>N</i> -best 音声認識結果 | 130 | POMDP | 121, 190, 203 | theory | 64 |
| | | posture | 63 | self-repair | 42 |
| [O] | | predicate logic | 51 | semantic differential scale | |
| objective evaluation | 172 | presentation | 40 | method | 172 |
| object-oriented | | production rule | 88 | semantic frame | 69 |
| programming | 166 | prosody | 131 | semantic grammar | 101 |
| ontology | 70 | push-to-talk | 213 | semantic representation | 68 |
| open domain | 9 | | | sequence labeling | 107 |
| OpenCV | 242 | [Q] | | service robot | 143 |
| OpenFst | 242 | question-answering | | shared plan | 56 |
| OpenJtalk | 242 | system | 2, 177 | short-utterance unit | 36 |
| openSMILE | 242 | | | SHRDLU | 5, 12 |
| operation | 51 | | | side participant | 61 |
| operator | 51 | | | | |

— 監修者・著者略歴 —

奥村 学 (おくむら まなぶ)

1984年 東京工業大学工学部情報工学科卒業
1989年 東京工業大学大学院博士課程修了
(情報工学専攻), 工学博士
1989年 東京工業大学助手
1992年 北陸先端科学技術大学院大学助教授
2000年 東京工業大学助教授
2007年 東京工業大学准教授
2009年 東京工業大学教授
現在に至る

中野 幹生 (なかの みきお)

1988年 東京大学教養学部基礎科学科第一卒業
1990年 東京大学大学院理学系研究科修士課程修了
(相関理化学専攻)
1990年 日本電信電話株式会社
1998年 博士(理学)(東京大学)
2004年 株式会社ホンダ・リサーチ・インスティテュート・ジャパン
2011年 早稲田大学客員教授(兼務)
現在に至る

駒谷 和範 (こまたに かずのり)

1998年 京都大学工学部情報工学科卒業
2000年 京都大学大学院情報学研究所修士課程修了
(知能情報学専攻)
2002年 京都大学大学院情報学研究所博士後期課程修了
(知能情報学専攻), 博士(情報学)
2002年 京都大学助手
2010年 名古屋大学准教授
2014年 大阪大学教授
現在に至る

船越 孝太郎 (ふなこし こうたろう)

2000年 東京工業大学工学部情報工学科卒業
2002年 東京工業大学大学院情報理工学研究科修士課程修了
(計算工学専攻)
2005年 東京工業大学大学院情報理工学研究科博士課程修了
(計算工学専攻), 博士(工学)
2006年 株式会社ホンダ・リサーチ・インスティテュート・ジャパン
現在に至る

中野 有紀子 (なかの ゆきこ)

1988年 東京女子大学文理学部心理学科卒業
1990年 東京大学大学院教育学研究科修士課程修了
(教育心理学専攻)
1990年 日本電信電話株式会社
2002年 マサチューセッツ工科大学大学院修士課程修了
(Media Arts & Sciences 専攻)
2005年 博士(情報理工学)(東京大学)
2005年 東京農工大学特任助教授
2008年 成蹊大学准教授
2013年 成蹊大学教授
現在に至る

対話システム

Dialogue Systems

© Honda Motor Co., Ltd., K. Komatani, Y. Nakano 2015

2015年2月13日 初版第1刷発行

検印省略

監修者 奥村学
著者 中野幹生
駒谷和範
船越孝太郎
中野有紀子
発行者 株式会社 コロナ社
代表者 牛来真也
印刷所 三美印刷株式会社

112-0011 東京都文京区千石 4-46-10

発行所 株式会社 コロナ社

CORONA PUBLISHING CO., LTD.

Tokyo Japan

振替 00140-8-14844・電話(03)3941-3131(代)

ホームページ <http://www.coronasha.co.jp>

ISBN 978-4-339-02757-0 (新宅) (製本: 愛千製本所)

Printed in Japan



本書のコピー、スキャン、デジタル化等の無断複製・転載は著作権法上での例外を除き禁じられております。購入者以外の第三者による本書の電子データ化及び電子書籍化は、いかなる場合も認めておりません。

落丁・乱丁本はお取替えいたします