

## 7 章 対話

### 【本章の構成】

本章では、対話 (7-1 節) について述べる。

## 2 群 - 10 編 - 7 章

## 7-1 対話

(執筆者：河原達也)[2009年9月受領]

対話は、複数の人間・システム（通常 2 者）が自然言語でやり取りするものである。自然言語処理においては、人間どうしの対話の分析・モデル化（機械翻訳などへの応用も含む）を行う研究<sup>1)</sup>と、人間（= ユーザ）と対話するシステムを構築するための研究<sup>2)</sup>がある。ここでは、主に後者について解説する。なお近年は、3 名以上の多人数会話に関する分析も盛んに研究が行われている。

工学的あるいは実用的に意味のある対話システムを構築するためには、通常、何を対話システムか（= タスク；例えば検索や予約など）を設定して、そのタスクがどのくらい達成されたかで評価することが多い。その際に、対象とする範囲（= ドメイン；例えばフライト情報や観光情報など）を設定することも一般的である。

タスクドメインが極めて限定的（例えば切符予約など）であれば、システム主導で、必要な項目（例えば出発地・目的地など）を順番に質問していけばよい。しかし、必要な項目を定めることが難しい場合やそもそもタスクドメインが明確でない場合は、ユーザが自由に質問する枠組み（ユーザ主導）が必要である。また、両者を混在させた形式（混合主導）も考えられる。

対話システムの機能は、ユーザの入力発話をタスクドメインや文脈のもとで解釈し、適切な応答を生成することである。音声で対話するシステムの場合は、音声認識と音声合成をフロントエンドに含める。そのための方法論（= 対話モデル）においては、大きくわけて、一問一答型の拡張（表層的アプローチ）、対話の流れの記述・モデル化（状態遷移型アプローチ）、対話の目標や心的状態のモデル化（人工知能的アプローチ）などがある。これらについて、以下に説明する。

## 7-1-1 一問一答型の拡張（表層的アプローチ）

最も単純な対話システムは、想定されうる質問に対して、応答のテンプレートを用意する方式である。「ありがとう」に対して「どういたしまして」、「明日の天気は？」に対して「明日は{晴れ | 曇り | 雨 | 雪}です」のように答える。古典的な例として ELIZA<sup>3)</sup>が挙げられるが、現在でもロボットやエージェントなどのように対話を楽しむのが目的のシステムにおいて一般に用いられる。

最近では、質問応答などの情報検索技術を利用して、応答を生成することで、高度な対話を行うことも可能になってきている。この場合も基本的には、単語の表層的なマッチングに基づいており、意味表現などを介する言語理解は行わないのが特徴である。ただし、談話レベルの解析を行って、隣接ペア（例えば、依頼-承諾/拒否）をモデル化した方が見通しはよくなる。

また、単純な一問一答でなく、一連の対話として成立するためには、省略・照応解析や焦点管理などの処理が必要となる。以下に、京都の名所について音声対話で情報案内を行う「京都版ダイアログナビ」<sup>4)</sup>による対話例を示す。

- U 1 : 銀閣寺について教えてください ( 検索 )
- S 2 : 銀閣寺は京都市左京区にある東山文化を代表する臨濟宗相国寺派の寺院です .  
正式名は慈照寺 , 山号は東山です . 銀閣寺の歴史や見所について何か知りたいことはありますか ? ( 要約提示 )
- U 3 : いつ建てられましたか ? ( 質問 )
- S 3 : 銀閣寺の建立は 1 4 8 2 年です . ( 応答 )
- S 4 : では , 銀閣寺は誰が建てたか知っていますか ? ( システムからの情報推薦 )
- U 4 : 知りません . 誰でしょうか ?
- S 5 : 創立者は室町幕府第 8 代将軍の足利義政 , 開山は夢窓疎石とされています .

### 7-1-2 対話の流れのモデル化 ( 状態遷移型アプローチ )

最も一般的な対話システムは , システム主導で順番に項目を質問していったり , ユーザの質問に対して詳細化を行い , 検索結果を提示するような流れをスクリプトで記述するものである . プログラミング言語のように記述したり , 有限状態オートマトンの形式で記述したりする . 状態遷移を定義・決定するために , 入力発話から意味表現を抽出し , システムのアクションも定義する .

近年 , システムやユーザの心的状態 ( 信念/情報状態 ) など組み込んだうえで , このような状態遷移のモデルを対話データから機械学習する枠組みが盛んに研究されている . タスク達成などに対して報酬を定義して , ( 各状態におけるアクションを ) 強化学習するのが一般的であり , 音声認識や言語理解などを誤りのある入力チャネルとみなす部分観測マルコフ決定過程 ( POMDP : Partially-Observable Markov Decision Process ) <sup>5)</sup> が主流になりつつある .

### 7-1-3 対話の目標や心的状態のモデル化 ( 人工知能的アプローチ )

より知的で複雑な対話を実現するために , 対話で遂行すべきタスクを階層的な構造を持つ表現 ( 木構造やフレームなど ) で記述し , 入力発話をそれに対応づけられる意味表現に変換したうえで , 対話を通してタスクを遂行するためのステップを推論・プランニングする方式が考えられている . 古典的な例として , 積み木の世界の操作を扱った SHRDLU <sup>6)</sup> が挙げられる .

あるいは , 対話エージェントの心的状態をモデル化したうえで , 人間のように振る舞うように設計する方式も検討されている .

### 7-1-4 多様な対話システムに向けて

タスクドメインが明確で , かつ各ユーザの目標が当初から明確 ( 例えば「明日の 13 時に東京に到着する電車を調べたい」) であれば , 正解の応答も定義できるので , できるだけ手際よく簡潔にこれを出力すればよい .

しかし , レストランの案内や観光地の紹介などにおいては , 当初は目標が漠然としていても , 徐々に選好が明確化して来たり , やり取りをしているうちに興味がわいて来たりするのが自然であり , この特徴こそが対話の本質とも言える . したがって , 対話が長時間続いた方が望ましいとも考えられる .

また , 「銀閣寺」について紹介するにしても , 歴史を勉強している小学生と , 実際に京都に

来ている大人に対しては、優先する内容や望ましい応答文の形式も異なってくるように、適切な応答というのはユーザや状況に依存する。

このような多様な対話システムに対して、評価を行える枠組みが必要である。対話システムのよさを複数の指標を用いて総合的に評価する枠組みとして、例えば PARADISE<sup>7)</sup>などが提唱されているが、基本的には目標指向対話に限定されている。

#### 参考文献

- 1) 石崎雅人, 伝 康晴: “談話と対話,” 東京大学出版会, 2001.
- 2) 河原達也, 荒木雅弘: “音声対話システム,” オーム社, 2006.
- 3) J. Weizenbaum: “ELIZA – a computer program for the study of natural language communication between man and machine,” Commun. ACM, vol.9, no.1, pp.36-45, 1966.
- 4) 翠 輝久, 河原達也, 正司哲朗, 美濃導彦: “質問応答・情報推薦機能を備えた音声による情報案内システム,” 情報処理学会論文誌, vol.48, no.12, pp.3602-3611, 2007.
- 5) S. Young: “Using POMDPs for Dialog Management,” Proc. IEEE/ACL workshop Spoken Language Technology (SLT), 2006.
- 6) T. Winograd: “Understanding Natural Language,” Academic Press, 1972.
- 7) M.A. Walker, D. Litman, C.A. Kamm, and A. Abella: “PARADISE: a general framework for evaluating spoken dialogue agents,” Proc of the ACL, pp.271-280, 1997.