

## 対話処理システムにおける内的コミュニケーション

島津 明 小暮 潔 川森雅仁 堂坂浩二 中野幹生  
NTT 基礎研究所

### 1 はじめに

人間同士の対話の特徴付けるものは何であろうか。それを計算機で扱うには、どのようなことが必要であろうか。最近、話し言葉表現の特徴の分析が言語処理の立場から進められている<sup>2),3),8),10),11),13),15)</sup>。自然な対話システムの実現には、それらの研究に加えて、人間同士の対話という側面について、その特徴と処理の関連の分析が必要である<sup>1)</sup>。

従来、情報の検索、CAIなどを対象に、質問応答という形態で計算機と人間の対話が扱われている。従来の対話システムの対話と人間同士の対話とを比べると、言語表現、インタラクションの様式など、様々な点で違う。従来の対話システムの処理の流れは、入力を受けて、それを理解し、システムの発話を生成するというサイクルが基本であるが、人間のような対話を扱うためには、そのような処理の流れだけでは不相当である。

本稿では、人間同士の対話をシミュレートするのに、対話システムの構成要素の間にどのような情報のやりとりが必要であるかについて、人間同士の対話の特徴を踏まえて、考察する。

### 2 従来の人間と機械の対話

人間と計算機が対話する形態の一つとして、質問応答システムが従来から研究されている。質問応答システムは、一般に、理解、生成、問題解決の各部からなる(図1)。

この枠組みでは、一文毎に理解と応答を繰り返す。文が入力されるたびに、文を解析して、意味構造を抽出し、それに基づいて、情報検索などの問題解決をし、それに基づいて応答をする。より進んだシステムでは、入力文を処理するたびに、談話構造を構築していき、これに基づいて、文脈処理をする。

質問応答システムが入力を受けて理解し応答を出すまでに、システムのモジュールの間には以下の情報のやりとりがある。

- (1) 理解部は入力文(音声)を理解した結果を問題解決部に伝える。
- (2) 問題解決部は情報検索などをし、その結果を応答生成部に伝える。

従来のシステムは一般に入力に文を仮定している。情報検索の自然言語インタフェースなどには、

名詞句を入力として許すものがあるが、そのような場合、名詞句に述部が適当に付加され疑問文として解釈される。

質問応答システムでは、応答文の作成は簡単である。すなわち、質問内容をデータベースに照らし、結果に応じた応答文を作る。例えば、入力がYes/No疑問文のときに、質問内容がデータベースに合っていれば、肯定の答えを返す。WH疑問文であれば、不定詞に相当する未知変数にマッチした内容を答えとする。

80年頃からプラン理解、発話のプランニングなどの研究がなされるようになったが、その場合のシステムは、文を理解した後、目標選択をして、問題解決や文生成をする構成をとる。

- (1') 理解部は入力文を理解した結果を目標決定部に伝える。
- (2') 目標決定部は目標を決定し、問題解決をして、応答目標を生成部に伝える。

この場合、理解、生成の処理が高度化しているが、情報の流れは一般に前述のように理解と生成というサイクルの繰り返しになっている。

### 3 人間の対話の特徴

分析対象として、研究者と助手の電話による対話、電話を通じた道案内などを調べている<sup>1)9)14)</sup>。道案内対話を例に自然な対話の特徴を表1に示す。これらの特徴をシミュレートするのに必要と考える対話処理の要件を併せて記す。分析している対話の例を図2に示す。

話し言葉は、書き言葉の観点からみると非文法的であり、文という単位で発話されるわけではない<sup>1)9)</sup>。

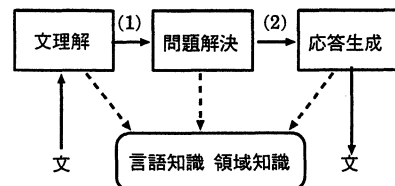


図1: 質問応答システムの構成

表 1: 人間の対話の特徴

特徴	処理
表現が多様	話し言葉表現の処理
発話単位が多様	文概念に代る処理単位
発話が漸次的	断片的表現系列の理解と生成
アイツチ, 並行発話	漸次的で実時間の理解と生成, 相互信念の形成
応答時間が短い	漸次的で実時間の理解と生成

断片的な語句が漸次的、連続的に繰り返される。しかし、無原則に発話されるわけではなく、そこには音調も含めた規則性がある<sup>6)9)</sup>。だからこそ、言葉が理解でき、コミュニケーションができると言える。

話し言葉による対話では、相手の発話が終わらないうちに、相手の発話と並行して、発話がなされたりする。アイツチも顕著である。アイツチなどの道具立てにより、話者と聴者が適当にコーディネートしながら対話が進められる。この過程を通して、話者と聴者の共有信念が形成されていく。

発話に対する相手の応答までの時間は短く一般に即時になされる。アイツチも、音響的に調べると、多くの場合、相手発話が終わらないうちになされることが分かる<sup>12)</sup>。不適切な間は発話権の放棄と見なされるためと考えられる<sup>1)</sup>。即時の応答はフィルターなどの表現を使ってでも、発話が即時になされる。このような漸次的で即時の発話も実際に実現することができる<sup>1)</sup>。

#### 4 従来の内的コミュニケーションの問題点

人間同士の対話をシミュレートすることを考えると、少なくとも以下の問題に対処するには、従来のように、文単位毎に一方方向に内的コミュニケーションを進めるのは不十分である。

##### ○発話単位は文ではない

一般の対話では、書き言葉にみられる文表現が必ずしも認められない。文ではない表現に対して返答が要求される。このため、従来のように、文単位毎に構文解析、プラン理解、問題解決、応答文生成などをした処理枠組みは不適切である<sup>7)</sup>。

例として、アイツチの生成がある。「はい」「うん」「ええ」のアイツチ発話を観察すると、アイツチは体言、格助詞、接続助詞などの後にも、多く現われることが分かる<sup>14)</sup>。アイツチは相手発話に対し、受理、理解、返答を表す<sup>4)5)14)</sup>。アイツチを打つためには、体言などの語句に対して、理解、目標選択、生成の処理をしなければならない。

##### ○発話は漸次的である

- M: [8]そこに行くーあの道順を教えていただきたいんですけども。  
 N: まずバスで [11]ー吉祥寺にでます=<=はい>でー吉祥寺から [12]ー井の頭(かしら)線のってですね  
 M: [11]はい [12]はい  
 M: 井の頭(かしら)せ [13]んは [14]い  
 N: [13]は [14]い。それで下北沢=<=はい>でそこからー小田急せ [15]んに乗り換えて [15\*]急行でいいです=<=はい> [16]それでー本 [16\*]、あの、本厚木の次の愛甲石田ってところで降ります=  
 M: [15]小田急線、はい [15\*] [16]急行で [16\*]=はい、あ急 [17]行が [18]とまるんですね  
 N: [17]で [18]え

図 2: 道案内の対話例

話し言葉では発話は漸次的である<sup>1)</sup>。システムが発話した部分的な語句に対し、対話相手がアイツチ、繰り返し、疑問などを発したりする。相手の部分的な発話を理解するためには、システムの発話と関連付けて理解する必要がある。システムの発話は必ずしも始めに意図した内容とは限らない。システムが発話を意図した内容、実際に発話したことなどを目標選択部、理解部は利用できることが望まれる。システムが発話を終り、その発話情報を文脈に積むだけでは不十分であろう。システムが実際に発話したことを一般には相手がどこまで実際に聞いたか理解したかも考慮する必要がある。

##### ○対話にはコーディネーションがある

対話は話者と聴者が適当にコーディネートしながら進められる。このコーディネーションには階層がある(表2)<sup>5)</sup>。適切なコーディネーションを実現するためには、各階層での理解、目標選択、生成の間の並行処理と即時の情報交換が必要と考えられる。

話し言葉では発話は漸次的であり、システムの発話中に、対話相手はアイツチ、繰り返し、疑問などを発したりする。対話のコーディネーションを成立させるためには、発話の実時間の処理が必要である。対話の現象を分析すると、同時発話を含め、対話参加者の発話の時間的重複がしばしば見られる。このような現象をみると、自己の発話生成と並行して相手発話の理解が進んでいることが伺われる。このような現象を捉えるためには、処理を並行して進める必要がある。このとき、システムの発話、相手の割り込み発話、その理解の間には依存関係があるから、生成、目標選択、理解の間にはタイムリーな

表 2: 対話のコーディネーションにおける階層

レベル	内容
発話	発話の番に従うなど、最低限、会話を成立させる規約に従っているレベル.
談話	呼掛けに答える, 質問に返答するなど, 対話を成立させる言語規約に従っているレベル.
意図	相手の意図を理解して, 対話を進行させるレベル.
協力	相手の意図に協力するよう, 対話を進行させるレベル.

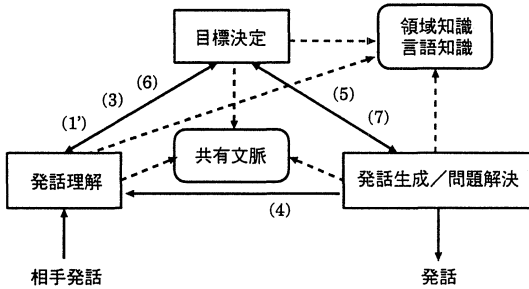


図 3: 対話システムの構成

通信が必要であろう。このような、生成と理解の並行処理、その間の情報交換は、従来の質問応答型の枠組では捉えきれない。

## 5 必要な内的コミュニケーション

人間同士がするように話し言葉による対話を進めるためには、上記の内的コミュニケーションに加え、システム内のコミュニケーションとして、少なくとも以下の情報伝達が必要と考えられる(図3)。

- (3) 目標選択部は、対話相手をもつと予想する行動目標を理解部に伝える。あるいは、予想される目標を共有メモリに書き込み、理解部が利用できるようにする。

これは、相手発話の解釈に、相手の行動目標が予測できている方が対応しやすいためである。

「緑の窓口」などの窓口の駅員は、客の行動目標として、切符の購入依頼、切符の価格の問い合わせ、空席状況の問い合わせ、経路の問い合わせ、などを予想する。これらの予測のもとで、発話領域に依存した解釈ができ、探索範囲が絞られるので解析速度が速くなる。例えば、「東京まで」と発話が始まれば、行為として、客の目標が東京までの切符の購入依頼、切符の値段の質問、所要時間の質問などであると予測することができる。これにより、システムは行動の準備ができ、探索範囲が絞られ、「東京まで」の後に続く発話を効率的に解釈できる。

この発話に対するシステムの振舞を考えてみる。「東京まで」は「東京まで切符をください」「東京までいくらですか」と曖昧であるが、「東京まで」は、目的地を指示しているという解釈はできる。そこで、「東京まで」に対して、アイツチを打つことができる。この場合、相手が目的地を指示したことを受理したことを表明することになる。

ここで、「東京まで」がどのような状況で発話されたかを考えると、客の目的は、多くの場合、切符を買うことである。すなわち、これをデフォルトとすると、「東京まで」という発話を「東京までの切符をください」と解釈し、切符を売ることに向けた準備体制に入れる。さらに、この後、特に発話がなされないなら、先のアイツチは、切符販売の依頼の受理、応答行動の表明ともみなせる。「東京まで」の後、「幾らですか」とか「行きたいのですが、席空いていますか」とかが続くと、先のデフォルト解釈は、棄却されることになる。従来の枠組では、入力が文であること、および、構文解析によって得られる論理表現にプラン理解が適用されることが前提であるために、「東京まで」のような断片的な语句に対する扱いは対象外となる<sup>7)</sup>。

プラン理解には、どこまで理解を掘下げるかという問題がある<sup>7)</sup>。目標決定部と情報を交わすことにより、無意味な推論の排除が可能になる。

- (4) 言語生成部は出力中の発話内容をプランとともに理解部に伝える。あるいは、共有メモリ上に記録し、理解部がこの情報を利用できるようにする。

「下北沢から小田急で本厚木まで」(U1)と発話を続けたところで、「急行ですか」(U2)と聞かれた場合、システムはU1をもたらししたプランを完了させる前に、U1に対してU2を関連付けて解釈することにより、適切な応答をすることができる。すなわち、U1は生成部が出力するものであり、U2は理解部が受けるものであるが、生成部は発話を中断し、理解部はU1という文脈のもとでU2を解釈することが望まれる。この例では、U1は、「下北沢から小田急で急行で本厚木まで行きます」というプランP1の一部に対するものである。そして、発話U2が、P1のもとで、U1につながるということで、システム

はU2を妥当な質問であると解釈することができ、P1の残りの部分の言語化を続行するよりも、P1のもとでU2に対する適切な応答を生成することができる。

このような例は、状況の変化に応じたモジュール間のコミュニケーションの必要性を示す。従来の枠組のように、“入力-解釈-応答”というサイクルを繰り返すだけでは、適切に対応できない。例のような割り込みのある対話のほか、質問者が逆に質問を受ける対話なども、単純な処理の流れでは、自然な対話は難しい。

- (5) 言語生成部は発話を出力したこと、あるいは、どこまで出力したかを目標決定部に伝える。

自分の発話が相手に伝わったという前提のもとで、発話による相手信念の変化が共有メモリに書き込まれ、それに基づき、次の相手発話が解釈される。

目標決定部が要請した発話内容を言語生成部がすべて発話するとは限らない。そのような場合、言語生成部は、要請された発話内容の出力した部分を目標決定部に伝えることになる。前述の(4)の例の場合、プランP1に対し、発話した部分U1を目標選択部に伝える。また、U2の発話を受けて、目標選択部は、発話が一時停止中であったP1が完了したとして、U2に対する肯定発話を言語生成部に指令する。

以下は、従来のような質問応答の繰り返しでは見られない、割り込みによる内的コミュニケーションである。

- (6) 相手から発話があると、理解部は発話を漸次解析中であることを目標選択部に知らせる。

理解部は生成部と並行して処理を進める。特に、発話中に相手発話が入ってくることを取り扱うために、このコミュニケーションが必要である。目標生成部は、システムが続行中の発話と相手発話との関連をみて、新たな目標設定をする。

- (7) 目標選択部は発話生成中でも、場合により、発話停止指令を発話生成部に送る。

前述のように、発話中に相手発話が入ってきた場合、目標選択部は発話生成部に発話の一時停止指令を出すか、発話を続行するかを判断し、一時停止の場合、その旨を言語生成部に送る。これを受けて、言語生成部は、発話の生成あるいは出力部への表現の送出を一時停止する。ただし、言語(音声)出力装置から出力中の表現は出力されてしまう。(4)の例の場合、U1の発話中にU2の発話が挿入される。入力が入ってきたことを理解部は目標選択部に伝える。

(4)の例のように、相手発話が出力中の発話に対する質問であり、発話を中断することが適切と判断されると、目標選択部は、一時停止指令を言語生成部に送り、言語生成部は発話を一時停止する。

## 6 おわりに

人間同士の対話をシミュレートするのに、対話システムの構成要素の間にもどのような情報のやりとりが必要であるかについて考察し、システムの構成要素間のコミュニケーションとして、従来にない情報の流れ(方向)が必要であり、入力を受けてから、理解、生成と単に順に処理するのではなく、理解、生成などの構成要素の並行処理が必要なことを述べた。

## 参考文献

- [1] 堂坂, 島津. 実時間対話における漸次的発話のプランニングモデル. NLC-95-40, 1995.
- [2] 伝. 制約と統計に基づく自然な発話の解析. 言語処理学会第1回大会, 1995.
- [3] Heeman, P. and Allen, J. Detecting and correcting speech repairs. *the 32th ACL*, 1994.
- [4] 片桐, 川森, 島津. あいづちの分散システムモデル. 言語処理学会第1回大会, 1995.
- [5] Kawamori, M., Shimazu, A., and Kogure, K. Roles of Interjectory Utterances in Spoken Discourse. *the Third ICSLP*, 1994.
- [6] 川森, 川端, 島津. 談話標識語の音韻的研究 (II). NLC-95-48, 1995.
- [7] 小暮, 島津, 中野. 対話状況依存型プラン解析. NLC-95-64, 1995.
- [8] Levelt, W. J. M. *Speaking*. The MIT Press, 1989.
- [9] Nakano, M., Shimazu, A., and Kogure, K. A Grammar and a Parser for Spontaneous Speech. *the 15th Coling*, 1994.
- [10] Nakatani, C. and Hirschberg, J. A speech-first model for repair detection and correction. *the 31th ACL*, 1993.
- [11] 大塚, 岡田. 自然な発話における漸時的精緻化について. NLC92-41, 1992.
- [12] 小坂. あいづちを中心とした会話音声の呼応関係の分析. SP87-107, 1987.
- [13] 佐川, 大西, 杉江. 対話文における誤りの自動修復. NL93-10, 1993.
- [14] 島津, 川森, 小暮. 対話の分析-問投詞的応答に着目して-. NLC-93-9 / NL-95-9, 1993.
- [15] 竹沢, 田代, 森元. 自然言語の言語現象と音声認識用法. SLP-6, 1995.