

7N-4

統計的知識を利用した音声対話処理

五味和洋、西 宏之、小島順治
NTT電気通信研究所

1. はじめに

人間が、特別な知識や慣れを必要とせずに機械と音声による対話を行えるようにすることは、音声情報処理の研究における主要な目標の一つである。このため、従来から多くの研究が行われているが⁽¹⁾、十分な認識語数・認識率を達成することの困難さから、実用的なシステムの実現にはまだ至っていない。一方、着信直後の電話応対や閉店時の受付業務等においては、対話が一定のパターンで遷移していくことが多いので、相手音声を一言一句理解できなくても、直前の対話遷移から相手の発言内容を推定することができると考えられる。本稿では、電話着信直後の対話内容遷移を統計的に分析し、その結果を知識として利用することによって相手の発言内容を推定し自然な応対を行う対話処理について報告する。

2. 対話データ収集分析

当研究所へ着信があった際の電話応対を対話データとして収集した(データ数:112着信分)。この際、着呼者は必ず「はい、〇〇研究室です」と述べて応対を開始した。この第一応答メッセージから発呼者が用件を述べ始めるまでに交される対話内容に着目して、対話データを分析し遷移パターンの出現頻度としてまとめたものが図1である。ただし、図1の中で用いている記号は、表1に示すように、発呼者・着呼者各々の対話内容を分類し、記号化したものである。図1より以下の3点がわかる。①発呼者が最初から用件を述べるケースはまれである(出現率1%)。②着信直後の対話遷移パターンは8種類程度に集中する(約85%)。③発呼者第一メッセージは表2に示す4種類のパターンに分類できる。

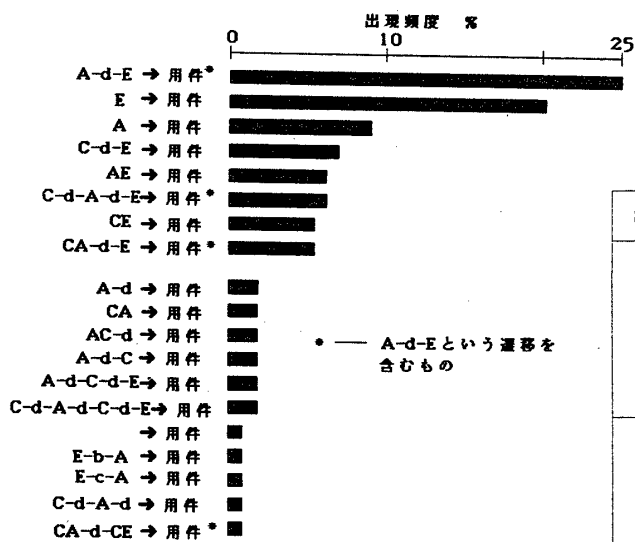


表1. 対話内容の分類

記号	発声者	意味	例
A	発呼者	発呼者名の名乗り	「佐藤ですが」
C		あいさつ	「もしもし」「恐れ入ります」
E		用件のある相手名の宣言	「鈴木さんお願いします」
b	着呼者	発呼者名の確認	「どちら様でしょうか」
c		あいさつ	「もしもし」「こんにちは」
d		相槌	「はい」「ええ」

図1. 対話遷移パターンの出現頻度

A man-machine conversation system using statistical knowledge
Kazuhiro GOMI, Hiroyuki NISHI, Junji KOJIMA
NTT Electrical Communications Laboratories

次に、音声長でこれらのパターンを識別することを考え、各パターンに属する発呼者第一メッセージの音声長を測定した結果を累積分布の形で図2に示す。

表2. 発呼者第一メッセージの分類

パターン	内容	出現頻度
1	AもEも含む	6.3%
2	Aを含みEを含まない	50.0%
3	Eを含みAを含まない	27.9%
4	AもEも含まない	16.2%

3. 相手メッセージ内容推定

図1, 2より、発呼者第一メッセージが表2のいずれのパターンに属するかは、各パターンの出現頻度とその音声長に関する知識から、確率的に推定・識別可能である。しきい値 T_H, T_L を、 T_H 以上のものがパターン1、 T_L 以下のものがパターン4である確率が最大になるように定めておき、実際の音声長としきい値の比較からパターン1、パターン4、それら以外のパターンのいずれであるかを識別する。図2における識別成功確率はパターン1:60%、パターン4:76%、それ以外のパターン:90%である。

一方、パターン2とパターン3の識別は、音声長に関する知識だけでは行えないので、対話内容遷移に関する知識を用いて行う。まず、発呼者第一メッセージに続いて相槌(d)「はい」を返す。もし発呼者第一メッセージがパターン2に属していれば、図1に示すようにその後“A d E”と遷移する対話例が多く存在することから(“A d”と遷移するパターンの81%)、相槌に対してすぐにEが応答されると考えられる。一方、パターン3の場合には、“E d”と遷移して用件に至る対話例がないため、相槌に対する応答は行われないと予想される。このことから、音声長が T_H と T_L の間である場合には、相槌に対する相手の反応の有無からパターンの識別を行うことができる。以上の相手メッセージ内容推定法を、図3に示す。

これらの識別結果及び音声の有無情報を利用して、次に送出すべきガイダンスを選択することにより、利用者の発声内容に対し柔軟に対応できる音声対話システムを実現できる。

4. まとめ

対話に関する統計的知識を利用して、相手のメッセージ内容を推定する処理について述べた。現在、汎用のパソコンを用いたシステムを試作し、以上述べた処理法の有効性確認を行っている。今後は、他の知識の利用や、音声認識技術との有機的な結合等により、更に複雑な遷移パターンにも対応できるシステムをめざしていく。

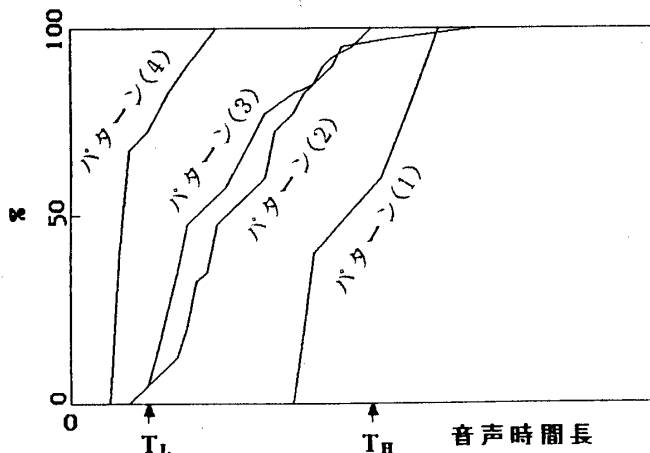


図2. 発呼者第一メッセージ音声長の累積分布

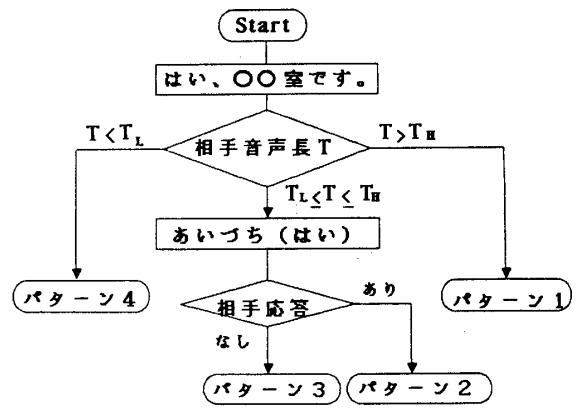


図3. 相手メッセージ内容推定法

[文献] (1)例えば、D.R.Reddy et al. "Speech understanding system-summary of 5-year effort at Carnegie-Mellon Univ.", CMU Tech.Rep., (1977)