

## 統計的手法による談話構造解析

6H-4

加藤 直人 森元 暉

(株)ATR音声翻訳通信研究所

### 1 はじめに

音声認識の精度を上げるためには言語的制約が必要であり、係り受け等による1文内の処理とともに、発話間の結束性という制約も重要となる。

本稿では、発話間の結束度を定義することによって談話構造を解析するとともに音声認識候補の制約となる手法について述べる。結束度はコーパスから計算される。また、発話対を構成した簡単な実験結果についても述べる。

### 2 背景

目的指向型対話は、局所的には質問・応答のような発話対にまとめられ、大局的には階層化された話題によって構造化される。

図1のような対話を考えよう。「ご希望の日にちをお願いいたします」と「八月十日から一二日です」は発話対としてまとめられる。これは、「日にち」と「八月十日」や「一二日」との名詞間の類似性や、「お願いいたします」と「です」のような質問・応答という発話意図間の関係において結束性が高いと判断されるからである。「何名様でしょうか」は「何名様」という名詞から人数という話題を推定することができ、宿泊予約という話題下においてそれまでの発話と結束性があることがわかる。

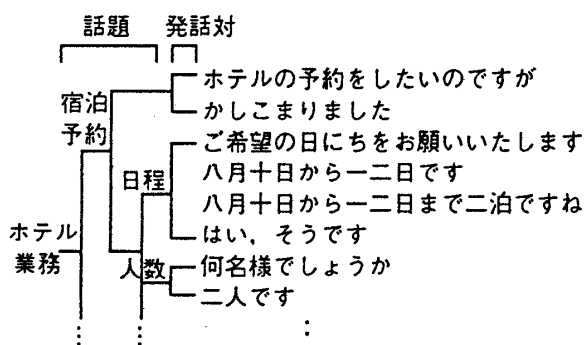


図1 目的指向型対話の例

Statistical Approach to Discourse Analysis  
Naoto Katoh and Tsuyoshi Morimoto  
ATR Interpreting Telecommunications Research Laboratories  
2-2 Hikaridai Seika-cho Soraku-gun Kyoto  
619-02, Japan

### 3 概要

一般に発話は命題、発話意図、手がかり語（例えば接続詞）で構成され、命題は動詞、名詞から構成される。また、発話はある話題に関して述べられ、話題は発話中の名詞から推定できる場合が多い。

そこで  $i$  番目の発話  $U$  を次の様にモデル化する。

$$U_i = (if_i, verb_i, noun_i, clue_i, topic_i),$$

$$noun_i = \{ noun_{i,1}, noun_{i,2}, \dots, noun_{i,n} \}$$

ここで、 $if$ は発話意図 (illocutionary force)、 $verb$ は動詞、 $noun$ は名詞の集合、 $clue$ は手がかり語 (clue word)、 $topic$ は話題を表わす。これらには重複して属してもよく、例えば、サ変名詞は  $verb$  に属すとともに、 $noun$  にも属すものとする。

発話間の結束性の有無は、手がかり語、発話対上の結束度、話題上の結束度の3つの観点から決定する。すなわち、 $i-1$  番目までの発話と  $i$  番目の発話間の結束度  $cohesion(U_{0,i-1}, U_i)$  を手がかり語  $clue_i$ 、発話対上の結束度  $cohesion\_inter(U_{i-1}, U_i)$ 、話題上の結束度  $cohesion\_topic(topic_{0,i-2}, topic_i)$  の3つの値からなるベクトルで定義する。

$$cohesion(U_{0,i-1}, U_i) = (clue_i, cohesion\_inter(U_{i-1}, U_i), cohesion\_topic(topic_{0,i-2}, topic_i))$$

手がかり語は前発話に関係なく、前発話との結束性を強制的に切る役目を持つものがある。また、話題上の結束度は[大川94]と同様、キーワード（特に名詞）から推定することを考えている。しかし、両者に関しては本稿では詳細は述べない。

発話対上の結束度は、発話意図間、動詞間、名詞間の結束度に重みづけした和で定義する。

$$cohesion\_inter(U_{i-1}, U_i)$$

$$= \lambda_1 cohesion\_if(if_{i-1}, if_i)$$

$$+ \lambda_2 cohesion\_verb(verb_{i-1}, verb_i)$$

$$+ \lambda_3 cohesion\_noun(noun_{i-1}, noun_i)$$

$$\lambda_1 + \lambda_2 + \lambda_3 = 1$$

以下では発話意図間の結束度を使って発話対の結束度の計算手法について述べるが、動詞、名詞間の結束度も同様である。

### 4 発話対の結束度

結束度はコーパスから計算するが、次を仮定する  
仮説：結束性が高いものは隣接し、頻度が高い。

## (1) 発話意図間の結束度

日本語の場合、発話意図は文末表現によってほぼ表わされるので、本手法では文末表現そのものを発話意図とする。しかし、本稿でいう文末表現とは、文末の助動詞の並びだけを指すのではなく、例えば「教えてください」のように、動詞「教える」を含んでいてもそれ全体を文末表現と呼ぶ。

文末表現は、あるしきい値以上の出現頻度をもつ文末の定型表現として抽出した。抽出した文末表現でコーパスから文末表現のbigramを作成した。

文末表現を  $expr_{i-1}$ ,  $expr_i$  と表すと、

$cohesion\_if(if_{i-1}, if_i) = cohesion\_if(expr_{i-1}, expr_i)$  と表われ、bigram や相互情報量によって計算される。しかしこれでは sparse data problem のため値が非常に小さいものになってしまう。そこで、次の2つの補間手法をそれぞれ用いる。

## (a) 小文末表現による補間

包含される文末表現（小文末表現と呼ぶ）で補間する。小文末表現をそれぞれ、 $\{expr_{i-1}^1, expr_{i-1}^2, \dots, expr_{i-1}^j\}$ ,  $\{expr_i^1, expr_i^2, \dots, expr_i^k\}$  とすると

$$\begin{aligned} cohesion\_if(if_{i-1}, if_i) \\ = \mu_1 cohesion\_if(expr_{i-1}, expr_i) \\ + \sum_{j,k} \mu_{jk} cohesion\_if(expr_{i-1}^j, expr_i^k) \\ \mu_1 + \sum_{j,k} \mu_{jk} = 1 \end{aligned}$$

と補間する。例えば、

$$\begin{aligned} cohesion\_if(ですか, です) \\ = \mu_1 cohesion\_if(ですか, です) \\ + \mu_{11} cohesion\_if(か, です) \end{aligned}$$

## (b) 文末表現のクラスタリング

「いただけますか」と「ください」は同じ発話意図を持つものであるが、(a)の方法では別々のものとして扱われてしまう。そこで文末表現のbigramを使って文末表現をクラスタリングすることにより文末表現のクラスを作り、クラスのbigramを使って補間する。

$$\begin{aligned} cohesion\_if(if_{i-1}, if_i) \\ = \mu_1 cohesion\_if(expr_{i-1}, expr_i) \\ + \mu_2 cohesion\_if\_class(class_a, class_b) \\ \mu_1 + \mu_2 = 1, \cdot expr_{i-1} \in class_a, \cdot expr_i \in class_b \end{aligned}$$

例えば、

$$\begin{aligned} cohesion\_if(教えていただけますか, です) \\ = \mu_1 cohesion\_if(教えていただけますか, です) \\ + \mu_2 cohesion\_if\_class(QUESTION, ANSWER) \\ 教えていただけますか \in QUESTION, \\ です \in ANSWER \end{aligned}$$

発話意図のクラスタリングは、コーパスから発話意図のタイプ (illocutionary force type) をボトムアップに自動作成していると考えられる。

発話意図が多義である文末表現の場合には英語の対訳による多義解消を考えているが、文末表現のクラスタリングについての詳細は別稿に譲る。

## 5 実験

(a) の小文末表現の補間による発話意図間の結束度のみを使って、発話対を求める実験を行った。ATR音声言語データベース[Morimoto 94]の約600対話から文末表現約200語を抽出した後、同一話者、話者交代の場合に分けてそれぞれの文末表現bigramを作り、相互情報量を使って結束度を計算した。しきい値より結束度の値が大きい場合を結束性有とした。人手による結果と一致した場合を正解とし、closed data, open data (それぞれ10対話) に対する結果を表1に示す。ここで比較のために、デフォルトでは結束性有の発話対の割合を示した。

表1 発話対の実験結果

実験データ	正解率	デフォルト
closed data	70.1%	62.5%
open data	69.8%	65.5%

## 6 おわりに

統計的手法による談話構造解析の概要、特に発話意図による結束度について述べた。本手法は従来の結束度を求める方法[工藤90][山岡90][Nagata 94]に比べ、人手による作業が少ないという利点がある。

## 参考文献

- [工藤90]工藤育男：文と文の結束性を捕らえるための知識、情報処理学会自然言語処理研究会報告、NL-76-7, pp.1-8 (1990).  
 [Morimoto 94] Morimoto, T., Uratani, N., Takezawa, T., Furuse, O., Sobashima, Y., Iida, H., Nakamura, A., Sagisaka, Y., Higuchi, N., Yamazaki, Y.: A Speech and Language Database for Speech Translation Research, Proc. of ICSLP-94, pp. 1791-1794 (1994).  
 [Nagata 94] Nagata, M. and Morimoto, T.: An Information-Theoretic Model of Discourse for Next Utterance Type Prediction, Trans. of IPSJ, Vol.35, No.6, pp. 1050-1061 (1994).  
 [大田94]大田雅彰, 山下洋一, 溝口理一郎：音声対話理解のための話題の決定について、電子情報通信学会音声研究会報告、SP-93-129, pp.9-16 (1994).  
 [山岡90]山岡 孝行, 飯田 仁：文脈を考慮した音声認識結果絞り込み手法、情報処理学会自然言語処理研究会報告、NL-78-16, pp.121-128 (1990).