

複数音声認識候補の再順序付けのための各種ヒューリスティックの比較*

7M-3

田代敏久[†]ATR 音声翻訳通信研究所[‡]

1 はじめに

N-BEST インターフェースを採用した音声言語システムにおいて、言語処理部は、1) 音声認識候補のフィルタリング、2) 入力正しい意味解釈の決定、の2つの役割を持っている。音声認識候補のフィルタリングに必要な知識は、統語的・意味的な制約とヒューリスティックな選好とに分けられる。本報告では、音声認識候補のフィルタリングに有効と考えられる各種のヒューリスティックの性質を調べるために行なった音声認識候補の再順序付け実験の結果を述べる。

2 N-BEST インターフェース

最近の音声言語システムは、音声認識部と言語処理部のインターフェースとして N-BEST インターフェースを採用しているものが多い。N-BEST インターフェースとは、

- まず、計算コストが小さい知識(音響モデル、単純な統語モデル等)を用いて候補の絞り込みを行い、上位 N 個の候補を出力する。
- その後、より計算コストがかかる知識(複雑な syntax、semantics 等)を用いて、候補を再順序付け(reorder)する。

というテクニックである [1]。

N-BEST インターフェースの音声言語システムを作成する場合には、1) 音声認識部での N-BEST 探索手法、2) 言語処理部での N-BEST 候補の再順序付け手法、3) N-BEST 探索や再順序付けに用いる言語知識、等について研究する必要がある。本報告では、音声認識候補の再順序付けに必要な言語知識について考察を行なう。

3 言語処理部に必要な知識

音声言語システムの言語処理部に必要な知識には、統語的・意味的な制約とヒューリスティックな選好の2つがある。

統語的・意味的な制約は、入力の統語的・意味的な適格性を判断するために用いられる知識であり、不適格な入力を棄却するために用いることが可能である。

*Comparison of Heuristic Preferences for Reordering of N-best Speech Recognition Candidates

[†]Toshihisa TASHIRO

[‡]ATR Interpreting Telecommunication Research Laboratories

しかし音声認識候補の中には、統語的・意味的には正しいが不適當な候補も存在することが多い。次のような例では、どちらの文も統語的・意味的制約違反は起こしていないので、2番目の候補をよりもっともらしいと判断するための知識が必要となる。

- 1、来月お申し込みになりますと読まないんです
- 2、来月お申し込みになりますと四万円です

ヒューリスティックな選好には、言語学者の内省に基づくもの(Close Attachment やゼロ代名詞最小等)と、コーパスから自動的に抽出可能なもの(単語のマルコフモデルや確率文法、語の共起情報等)がある。上記の例では、内省に基づくヒューリスティックを用いる場合、「読まない」の必須格(主格及び目的格)が埋まっていない(ゼロ代名詞が多い)から不自然である、等と判断することが可能となる。一方、統計に基づくヒューリスティックを用いる場合には、訓練コーパスに「読まない」という語句がこのような文脈で出現することが少ない、ということにより2番目の候補をよりもっともらしいと判断するわけである。

4 音声認識候補の再順序付け実験

4.1 実験対象

国際会議の申し込みをタスクとする会話の音声認識結果(男性話者1名、計41発話、1発話につき最大5best候補、計187文)を実験対象とした。41発話中、29発話(70.7については1位候補が正解であり、39発話は5位以内に正解を含んでいた。音声認識には SSS-LR 音声認識システム [2] を利用した。

4.2 再順序付けのための言語知識

次のような言語知識を用いて音声認識候補を再順序付けした。

- 単一化文法
統語的・意味的な適格性を判断するために、単一化文法の枠組みで記述された日本語解析文法を用いた。文法規則の数は約1700(内、語彙規則約1500)である。この文法は「目的指向型」電話会話における日本語の基本表現の約90%の表現をカバーしている [3]。
- 単語の bigram
実験対象とは別の2000文(のべ単語数約15000語)

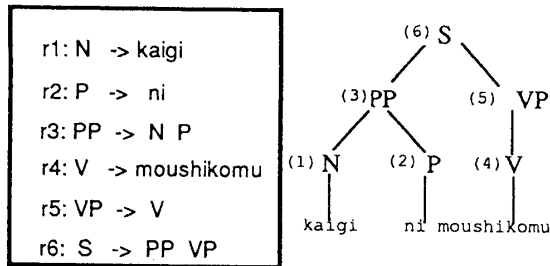


図 1: 最右導出の逆過程

から単語の bigram を学習した。異なり語の数は約 1500 である。

- 品詞の trigram
単語の bigram と同じ例文から品詞の trigram を学習した。品詞の数は 75 である。
- 文法規則の bigram
文脈依存の確率文法の一つである北 [4] により提案された言語モデル (ヒューリスティック) を、実験対象とは別の 800 文から学習した。

この言語モデルは、ある木の確率を、その木の最右導出の逆過程 (図 1 参照) を条件とした条件付確率の積、

$$P(T) = \prod_{i=2}^n P(r_i | r_1, r_2, \dots, r_{i-1}) \quad (1)$$

と定義し、これを次のような単純マルコフ過程 (bigram) で近似したものである。

$$P(T) \simeq \prod_{i=2}^n P(r_i | r_{i-1}) \quad (2)$$

このような確率文法の学習には、正しく解析されたデータが必要であるが、このようなデータを作成することは非常に手間のかかる作業である¹ため、学習に用いた例文の数が少なくなっている。

- ゼロ代名詞最小優先
必須格要素の欠落が少ないものを優先する内省に基づくヒューリスティックである。

4.3 結果と考察

音声認識候補の再順序付け実験の結果を表 1 に示す。まず、音声認識候補 41 発話 (187 文) に対し、単一化文法の枠組みで記述された統語的・意味的制約を加えた。その結果、音声認識で正しく 1 位に認識された候補を正しく解析したうえで、さらに 3 つの発話の不適切な候補を排除することにより、下位の正しい候補を選び出すことができた。

次に、統語的・意味的制約を満たした候補 (138 文) に対して、前述のヒューリスティックを用いてスコア計算し、候補を再順序付けした。CFG bigram による

¹Inside Outside アルゴリズム等を用いれば、教師なし学習も可能だが、この場合は学習時間が多くかかってしまう。

| | 一位正解 | 副作用 |
|--------------------|---------|-----|
| 音声認識結果 | 29(71%) | - |
| 単一化文法 | 32(78%) | - |
| 単一化文法 + 単語 bigram | 31(76%) | 5 |
| 単一化文法 + 品詞 trigram | 28(68%) | 5 |
| 単一化文法 + CFG bigram | 34(83%) | 3 |
| 単一化文法 + ゼロ代名詞 | 31(76%) | 1 |

表 1: 音声認識候補の再順序付け -1

| | 一位正解 | 副作用 |
|--------------------|---------|-----|
| 単一化文法 + 単語 bigram | 34(83%) | 2 |
| 単一化文法 + 品詞 trigram | 28(68%) | 5 |
| 単一化文法 + CFG bigram | 35(83%) | 2 |

表 2: 音声認識候補の再順序付け -2

再順序付けでは、音声認識結果の結果とくらべ 12% の正解率の向上が見られた。

表中の「副作用」の欄は、音声認識で正しく 1 位に認識された発話を、言語知識により評価を下げてしまった発話の数である。この数を減らすためには、ヒューリスティックを用いたスコア計算の際に、音声認識結果のスコアを適切に評価する必要がある。そこで、音声認識スコアとヒューリスティックのスコアを 3:1 (実験的に決定した) の重み付けで再順序付けを行ってみた²。表 2 に結果を示す。

単語 bigram と CFG bigram を用いた実験では、音声認識スコアを利用することにより正しい候補の評価を下げるものが少なくなった。しかし、品詞 trigram では、よい結果がでなかった。これは、音声認識で使用している言語モデルが文脈自由文法であるため、品詞 trigram で表現できるような知識は既に使われてしまっているためだと考えられる。

5 おわりに

本報告では、複数の音声認識候補を再順序付のためのヒューリスティックの比較実験の結果について述べた。今後は、さらに大量のデータによる実験を行なう予定である。

参考文献

- [1] Schwartz, R. M. and Chow, Y.: "The N-Best Algorithm: Efficient Procedure for Finding Top N sentence Hypotheses," ICASPP-90, pp.81-84, 1990
- [2] 永井明人、鷹見淳一、嵯峨山茂樹: "逐次状態分割法 (SSS) と音素コンテキスト依存 LR パーザを統合した SSS-LR 連続音声認識システム," 信学技報, SP92-33, (1992-06)
- [3] 浦谷則好ほか: "話し言葉の日英翻訳システムの評価法," 情報処理学会大 4 6 回全国大会, 6b-4, (1993-03)
- [4] Kita, K et al.: "Continuously Spoken Sentence Recognition by HMM-LR," ICSLP-92, pp.305-308, 1992

²「ゼロ代名詞最小優先」は、スコアが単なる整数値のため対象外とした。