

## 学習用テキストからの認識探索空間作成能力の評価

1 E-7

松永昭一 鹿野清宏

NTT ヒューマンインタフェース研究所

## 1 はじめに

従来、連続音声認識では、大規模な学習テキストを用いて  $n$ -gram モデルなどの統計的言語モデルを作成し、認識性能を向上させてきた [1]。また、構文も併用され、学習テキストと構文に関する知識を元に、人為的に作られていた。構文は、認識探索空間を制約し、処理速度、認識性能を向上させる。特に、大語彙や未知語、比較的自由的な構文などを扱う Dictation システムの場合には探索空間が非常に大きくなり、構文の記述は重要な要素である [2]。一方、認識探索空間を自動的に制約することは、学習テキストのみでも可能である。しかし、この場合には、学習テキストに含まれない未知の構文(単語系列)が入力された場合、処理不能となる。これを回避するためには、さらに膨大な学習テキストを収集すれば良いが、非常にコストのかかる仕事である。従来は、人手による構文作成時に、他の文例等の知識を利用して回避していた。そこで、我々は既存の学習テキストとその統計量から、学習テキストに含まれず、かつ、出現確率が高いと推定されるテキストを生成し、そのテキストと既存の学習テキストから探索空間を限定した認識法を提案した [3]。即ち、未知の構文に対処するために、テキストを生成した。本稿では、この認識法の探索範囲制約の程度を変えたときの、処理量と認識性能について検討する。

## 2 テキストの自動生成と探索空間の制約

学習テキストとその文字 trigram を用いて以下の手順で自動生成する [3]。

1. 文節を二つに分ける任意の分割点で、文節の前半(a)と後半(b)の文字 trigram の尤度の和が最大となる分割点を選ぶ。
2. 学習文節の中で、前半が(a)と同じ文節の後半と、後半が(b)と同じ文節の前半を組合せ、その文字 trigram の尤度が閾値以上のものを新規文節として生成する。
3. 上記分割点を除いて、1,2の操作を繰り返す。

医用所見に関する 15000 文節の学習テキストを用いて新規文節を生成し、評価用 12000 文節の文字三つ組の被覆個数の増加(15000 学習文節を用いた時の被覆個数からの増加分)を、生成文節数に応じて調べた(図1)。この結果、10000 文節程度生成した段階で、生成できる三つ組の大部分が被覆される。また、人手で収集した場合(破線)と比較すると、5000 文節収集した場合(計 20000 文節)に、15000 文節を自動生成した場合と被覆数が同じになる。自動生成は、人手による収集には及ばないが、被覆能力は大きいことがわかる。

## 3 連続音声認識実験

認識実験では、文節単位発声の Dictation システム [4] を用いた。文字 trigram を作成する学習テキストは 15000 文節であり、これより新規のテキストを約 60000 文節作成し、探索空間を学習、新規テキストに出現する文字  $n$  組のみに制約した。

まず、話者独立の 222 文節の認識を文字 tri-gram を用いて行なった。文節認識率、HMM 照合回数、処理時間の割合を探索範囲制約 (1,2,3,4 組) に応じて図 2 に示す。組数が 0 の場合は探索制約がないことを示す (探索空間は学習テキストの 920 文字の組合せ)。この結果、4 組制約を用いる場合、文節認識性能を劣化させることなく (75.7% → 78.8%)、処理時間を約 1/4 とした。また、人手による構文と文字 trigram (72000 文節) で認識した結果 [5] を図 2 の右端に併記する。

一方、(a) 学習テキスト 15000 文節のみ、(b) 評価テキストを含む学習テキスト 500 文節のみ (共に自動生成なし) で制約した場合の HMM 照合回数、処理時間の割合を図 3 に示す。図 2 と (a) の差が、生成テキストにより増加した探索量となる。また、図 3 により、本手法の削減処理量の上限がわかる。これらの結果、テキストを膨大に生成し被覆率をあげても認識処理量が極端に増加することはないことがわかる。

4 まとめ

既存の学習テキストとその統計的量から、学習テキストに含まれず、かつ、出現確率が高いと推定されるテキストを生成する方法と、それを元に探索空間を限定した認識法について検討を行なった。

謝辞 日頃御指導、御討論頂く古井室長、坪井主任研究員、山田社員を初めとする皆様に感謝致します。

- <文献> [1] L. R. Bahl et al., "A maximum likelihood approach to continuous speech recognition", PAMI-5, 2, (1983)
- [2] J. Hosaka and T. Takezawa, "Construction of corpus-based syntactic rules for accurate speech recognition", COLING (1992)
- [3] 松永他: "学習用テキストからの認識探索空間の自動作成", 音学講論 (1992.10)
- [4] T. Yamada, et al., "Japanese dictation system using character source modeling", ICASSP (1992)
- [5] S. Matsunaga, et al., "Continuous speech recognition for medical diagnoses using a character trigram model", ICSLP (1992)

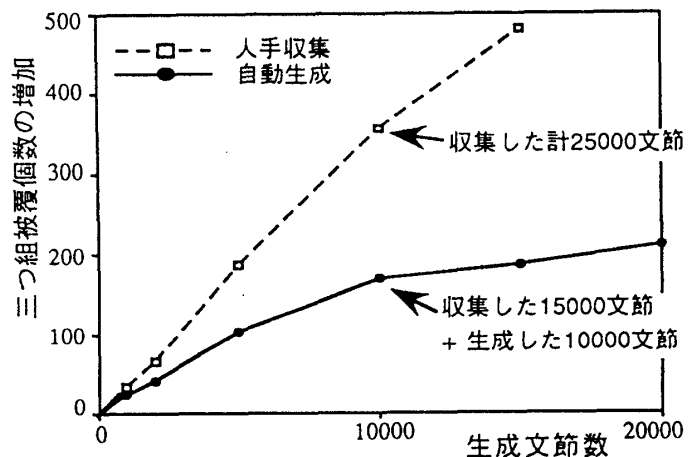


図 1. 文節の自動生成における三つ組被覆個数の増加

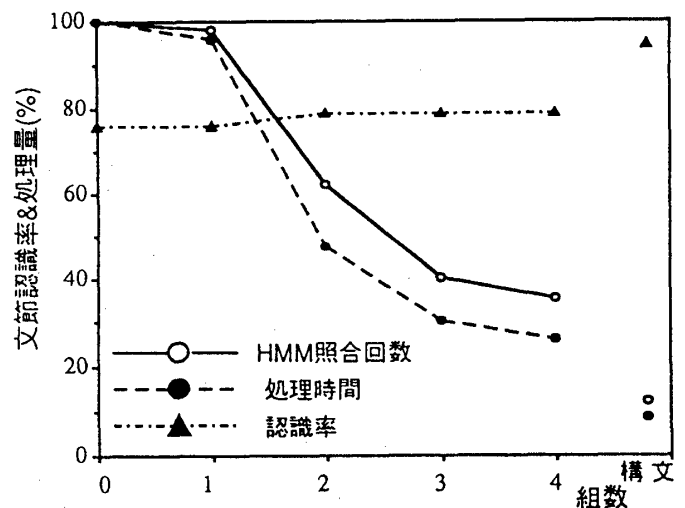


図 2. 探索範囲制約と認識率、処理量の関連 (15000 学習文節、文節生成あり)

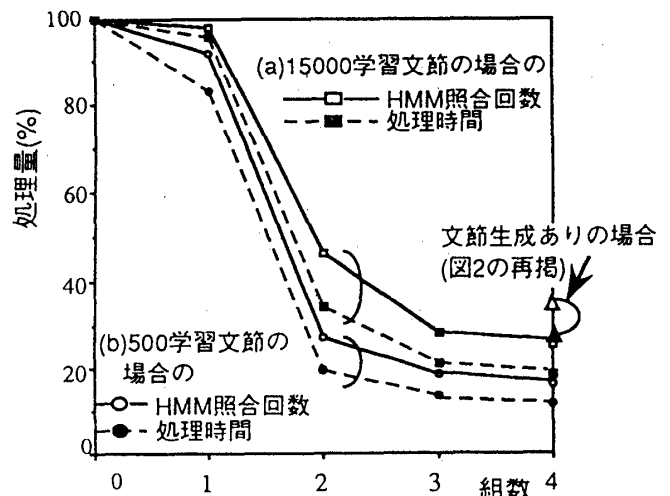


図 3. 探索範囲制約と処理量の関連 (文節生成なし)