

Garbage model と構文的拘束を用いたワードスポッティングの手法における

7 B-4

Garbage model の学習に関する検討*

小沼 知浩

慶應義塾大学 大学院 理工学研究科 計算機科学専攻

1 はじめに

自由発話文の認識において、不要語・未知語の処理は大きな問題である。不要語・未知語処理の方法として、Garbage model を用いた方法 [1] が提案され、未知語処理に関する実験 [2] が報告されている。Garbage model を用いた方法の応用として、Garbage model と単語の構文的拘束を用いた単語スポッティングの手法 [3] を提案した。文献 [3] では、手法の概要と特徴を知るための実験に対する検討を行なった。今回の発表では提案した手法の最も重要な部分である Garbage model の学習方法について報告する。

発表のながれとしては、まず本手法の概要を述べ、実験条件を示した後、Garbage model の学習方法についての結果の報告・検討を行なう。

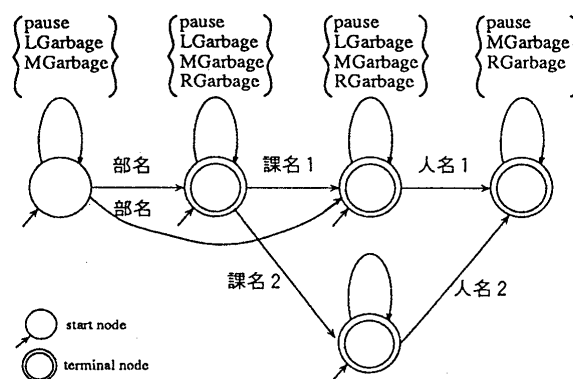


図1 構文的拘束

2 Garbage model と構文的拘束を用いた方法

この方法は、比較的自由に発話された連続音声の中の不要語・未知語をスキップし、タスク達成のための単語 (vocabulary word と呼ぶ) をスポッティングすることにより入力文を受理するものである。以下のものを用いる。

- vocabulary word HMM
各 vocabulary word に対応する HMM。
- Garbage model HMM
vocabulary word 以外の単語を受理する HMM。文頭、文中、文末の3つ (LGarbage, MGarbage, RGarbage) を用いる。
- 構文的拘束
図1のような有限状態遷移 (FSN) で表現され、正規文法により vocabulary word の出現順序と入力文の構文を規定する。

有限状態遷移のセルフループの弧の部分に Garbage model HMM と pause を表現する HMM を、その他の弧にそれぞれのカテゴリの vocabulary word HMM を置いて得られるネットワークに基き、Viterbi サーチを用いて認識を行う。

3 認識実験

Garbage model の学習方法の検討のため以下のような認識実験を行なった。

不要語・未知語に関するサンプル収集が不十分であったため、内線番号案内タスク [4] において、タスク達成に必要な単語 (『部名』、『課名』、『人名』) 以外を不要語・未知語とみなし実験を行なった。典型的な内線問い合わせ文は以下のようなものである。下線は vocabulary word を表している。

- 総務部の部長代理の渡辺さんをお願いします。
- はい。〈pause〉 業務の大久保さんです。
- いいえ。企画開発室です。〈pause〉

3.1 実験条件

ATR 音韻バランス文 4000 文 (80 名 × 50 文) で学習したコンテキスト依存 subphoneme HMM (382 状態数) を用いて vocabulary word HMM を構成した。また、Garbage model HMM は 5 状態の left-to-right HMM を 3 つ (文頭用 LGarbage, 文中用 MGarbage, 文末用 RGarbage) 用意し、タスク依存文 (O) (認識文以外の内線問い合わせ文) で学習した。特徴パラメータは、LPC メルケプストラム (1024) とデルタケプストラム (256) を用いた (括弧内はコードブックサイ

*The training method of concatenated Garbage models on Word-spotting using a FSN grammar with Garbage models
Tomohiro KONUMA
Keio University

ズ)。認識実験は14名(男・女7名)の発声による内線問い合わせ文(560文)を用いて不特定話者の認識を行った。さらに、認識実験ではタスク依存文を学習データとして用いた場合、MGarbageが vocabulary word を学習している可能性が高いという特性に基づき、MGarbageを用いていない。また、HMMの各状態に対して duration control を行なった。

3.2 Garbage model の学習方法

Garbage model の学習方法については、文献 [1] でも述べられているが、文献 [3] では構文的拘束を用いた手法に基づく、ポーズを目印とする Garbage model の連結学習の方法を提案した。以下の3つの連結方法を考案し、比較実験を行なった。

G-3 モデル (ポーズを考慮していないモデル)、pause-G-3 モデル (各 Garbage の前後にポーズを挿入したモデル)、pause-G network モデル (pause-G-3 モデルを入力文中のポーズ数変化に対応できるように変更したモデル) の3種 (図2) で行った。図2中の pause-G network モデルにおいて、枝わかれている状態遷移の各状態は、図中に破線で示すとおり状態の確率密度関数を共有している。

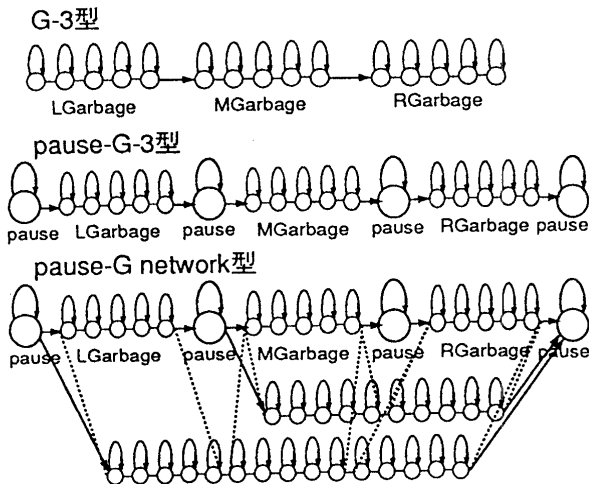


図2 学習時の連結法

3.3 実験結果と考察

各 Garbage model 連結方法の間に表1のような認識率の差異が認められた。特にポーズを考慮していない G-3 モデルと、pause-G-3 モデル、pause-G network モデルの間には、7~12%の認識率の改善が見られた。

表1 Garbage model 連結方法による認識率の差異

連結方法	認識率
G-3	56.8%
pause-G-3	63.8%
pause-G net	68.0%

学習データには、文頭、文中、文末の3つ Garbage model の学習区間を示すラベル付けを行なっていない。このような条件のもとで、3つの Garbage model の連結が期待されているような学習区間を選択し、適切な学習が行なわれるかどうかということが本実験の目的となっている。

結果は、各 Garbage model 学習時の適切な学習区間の自動的な選択は、暗には行なわれにくいということを示している。pause-G network のように、ポーズを3つ Garbage model の学習区間の割り当ての目印とすることは、認識率の大幅な改善となり得ることを示唆している。すなわち、ポーズは各 Garbage model の学習区間選択の大きな役割を果たしているということがいえる。ポーズのような Garbage model の学習区間選択の目印を探索し、組み合わせることにより適切な学習が行なわれ、認識率の向上が期待できる。

また、音響的に特徴を持っているポーズを Garbage model の学習対象から除き、同等に扱うことも認識率を改善していると思われる。

4 まとめ

Garbage model と構文的拘束を用いたワードスポットティングの手法における Garbage model の学習方法についての実験・考察を行なった。

実験の結果から、ポーズが Garbage model の学習区間の大きな役割を果たしているということが推測される。また、このような学習区間選択の目印をいくつか組み合わせることにより認識率の改善が期待できる。さらに、音響的に特徴を持つポーズは Garbage model とは分離し、単独で HMM を用意し Garbage model と同列に扱うことにより、認識率が改善されることがわかった。

今回明らかになった Garbage model 学習方法の性質をもとにさらにいくつかの特徴を調査し、今後は実際の自由発話文(及び音声対話)の枠組みへの適用を検討していく。

参考文献

- [1] J. G. WILPON, L. R. RABINER, C. H. LEE, E. R. GOLDMAN, *Automatic Recognition of Keywords in Unconstrained Speech Using Hidden Markov Models*, IEEE Trans. on ASSP, Vol.38, No.11, Nov. 1990, pp.1870-1878.
- [2] A. ASADI, R. SCHWARTZ, J. MAKHOUL, *Auto-matic Detec-tion of New Words in a Large Vocabulary Continuous Speech Recognition System*, ICASSP '90, pp125-128, April. 1990.
- [3] 小沼知浩, 武田一哉, Garbage model と構文的拘束を用いたワードスポットティングの検討, 日本音響学会秋季研究発表会(平成4年10月予定)
- [4] 黒岩眞吾, 武田一哉, 井ノ上直己, 庄境誠, 内線番号案内システムのための不特定話者連続音声認識, 音講論集 1-5-17(平成3年10月)