概 要

In this paper, we investigate statistical network grammar using Ergodic Hidden Markov Model(HMM).

There are two types of natural language modeling. One is the class of deterministic models like network grammar or context free grammar, that exploit some known specific properties of the language, and the other is the class of statistical models like bigram or trigram in which one tries to characterize the statistical properties of the corpus.These statistical models include stochastic context free grammar and Markov process, a sort of non-deterministic finite state automaton. A HMM effectively exploits language model as a random process. By choosing specific parameters of HMM, grammatical rule can be estimated in a well-defined manner as a transition network. HMM is very rich in mathematical structure so that language models are determined more precisely than that of stochastic network grammar or Markov process.

This paper includes the results obtained to statistical network grammar automatically from about 4000 words using Ergodic HMM. The resultant model indicates that some grammatical features exist even though the process is carried out automatically.

Ergodic HMM による確率つきネットワーク文法の獲得の可能性について

The Possibility for Acquisition of Statistical Network Grammar using Ergodic HMM

村上仁一　　山本寛樹　　嵯峨山茂樹

Jin'ichi Murakami　Hiroki Yamatomo　Shigeki Sagayama

ATR 自動翻訳電話研究所

ATR Interpretating Telephony Research Laboratories

1　はじめに

音声認識に利用される言語モデルには、ネットワーク文法や文脈自由文法に代表される構文モデル [1] や、bigram・trigram に代表される統計モデル [2][3] がある。

ネットワーク文法や文脈自由文法などの構文的な言語モデルは、自然言語処理の分野で実績があるが、言語に関する知識に基づいて構文規則を人間が記述するため多大な労力を要する。一方 bigram や trigram などの統計モデルは、簡単なモデルであるため音声認識の分野で言語モデルとして良く利用されている [2] が、このモデルは言語を表現するにはあまりにも単純である。そこで両モデルの問題点を補完するために、構文モデルに確率を加えた確率つきネットワーク文法や確率つき文脈自由文法などの研究がある [4][5]。

ところで音声認識の分野では隠れマルコフモデル (HMM) が良く利用されている [4]。HMM の種類の中で全状態間の遷移の許された離散型 Ergodic HMM の構造と確率つきネットワーク文法の構造は類似している。また HMM は Baum-Welch algorithm を用いることによって、学習データの生成尤度が最大になるように各パラメータを推定することができる。そこで言語モデルとして Ergodic HMM を考え、テキストデータを学習データとして Baum-Welch algorithm を利用することにより、確率つきネットワーク文法を自動的に獲得できる可能性がある。

なお、村瀬等 [7] はカテゴリーを学習データとして学習後のモデルのエントロピーを調べ、bigram や trigram と比較し、Ergodic HMM による言語のモデル化の可能性を報告している。田本等 [8] は品詞を学習データとして、学習後の Ergodic HMM を解析

し、その形態が従来使われているネットワーク文法と類似していることを報告している。また、英語では Ergodic HMM は確率つきネットワーク文法の獲得手段としてでなく [4][5]、形態素解析として研究されることが多かった [6]。この場合、品詞ラベルが付与された大量のテキストデータがあれば HMM のパラメータは直接計算できるため、品詞ラベルがないテキストデータから Baum-Welch の学習アルゴリズムを用いた大規模な実験はまだ行なわれていないようである。

本論文では日本語においてテキストデータを単語の系列と見なして Ergodic HMM を学習させた。この場合 Ergodic HMM はネットワーク文法と同時に、単語に対する新しい品詞体系を得ることができる可能性がある。この観点から学習後の Ergodic HMM のパラメータを調べた。また文音声認識における言語情報として用いたときの有効性なども調べた [9][10]。

2 HMM と確率つきネットワーク文法

2.1 HMM

HMM[11] は、出力シンボルによって一意に状態遷移先が決まらないという意味での非決定有限状態オートマトンとして理解できる。また確率的性質を持つ信号源が Markov 的に切替えられて全体として非定常信号源を表現しているモデルとも考えられる。このモデルにはいくつかの種類があり、図 1 のような left-right 型 HMM は音声認識の分野で音素モデルとして良く利用されているが、図 2 のように全状態が全状態に接続されているモデルは Ergodic HMM と呼ばれる。

HMM はパラメータとして状態遷移確率 A, シンボル生成確率 B, 初期状態確率 π を持つ。これらのパラメータに初期値を与えて、Baum-Welch algorithm を利用すると学習データに対して尤度が局所的最大になるように再推定することができるが [11]、必ずしも最大値に収束しないため初期パラメータの設定は重要である。

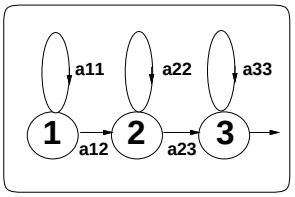

図 1: Left-right HMM の例 (3-state HMM)

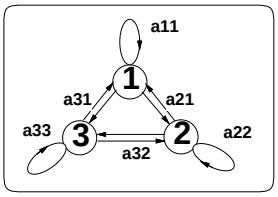

図 2: Ergodic HMM の例 (3-state HMM)

2.2 Ergodic HMM と確率つきネットワーク文法と品詞分類

ネットワーク文法は、言語を有限状態オートマトンでモデル化したもので、状態が遷移するときに単語が出力されるモデルである。このモデルに確率をつけた確率つきネットワーク文法は Ergodic HMM と同じ構造を持っている。したがって言語モデルとして Ergodic HMM を用いた場合、Baum-Welch algorithm によってテキストデータから尤度最大の条件で確率つきネットワーク文法を自動的に獲得できる可能性がある。

さらに単語列を学習データとした場合、このネットワークでの挙動が似た単語はネットワーク上の同一の状態遷移に集中する傾向があると思われる。一方、従来の品詞体系は基本的には構文的な観点で分類されている。したがってこのグループ化は品詞と強い関連があると思われる。そこで同一遷移に出現する複数の単語を同一の品詞と考えるような品詞分類と考えると、Ergodic HMM によりネットワーク文法の獲得と同時に、尤度最大の基準での新しい品詞体系を得る可能性がある。

3 Ergodic HMM による言語のモデル化

本論文では離散出力分布の状態遷移型 Ergodic HMM を大量のテキストの単語列により学習させ、確率つきネットワーク文法が獲得できるかどうか試みた。

3.1 言語データ

Ergodic HMM の学習に用いる言語データベースとして、ATR 対話データベース (ADD)[12] を用いた。文の例を表 1 に示す。この言語データベースは電話対話であるため、「はい」「もしもし」「わかりました」「そうですか」などの短い会話文が多い。単語

の定義は、日本語において曖昧な点があるが、このデータベースは人手で形態素解析がされていて品詞も付与されているため、これにしたがった。また、同じ表記でも読み方の異なる場合や、品詞や活用形や活用型の異なる場合は、別単語として扱った。その結果、異なり語彙数は 6418 種類、品詞は名詞、動詞などの基本形で 25 種類、活用形および活用型まで含めて 114 種類ある。

実験では、8000 文を奇数番目の文の set と偶数番目の文の set とに分けて使用した。以後、奇数番目の set を “odd4000”、偶数番目の set を “even4000” と呼ぶことにする。

表 1: 文の例

・はい もしもし
・えーっと そちら 第 1 回 の 通訳 電話 国際 会議 の 事務 局 で しょ う か
・はい そう です
・えーっと ちょっと その 会議 の こと でね
・はい どうぞ
・えーっと 今 手元 に あの 登録 用紙 が ある ん です けれども
・えーっと その 中 で ちょっと あの クレジットカード を ね
・あのー クレジットカード の 名前 と なん か ナンバー を 書く ところ が ある ん です が
・はい そう です
・えーっと それ を ちょっと クレジットカード を 持っ て い ない 者 が いる ん です けれども
・その 場合 は どう な ん でしょ う か

3.2 実験条件

前節で作成した対話データ odd4000 を学習データとし、状態数の異なる 4 種類の状態遷移出力型 Ergodic HMM (2 状態、4 状態、8 状態、16 状態) でモデル化の実験を行なった。Baum-Welch algorithm の学習終了条件として、シンボル系列を生成する確率の上昇率がある一定値以下になった場合に学習を終了させる方式を採用した。実験の条件を表 2 にまとめた。

表 2: 言語モデル生成実験の条件

HMM の構造	状態遷移出力型 Ergodic HMM
HMM の状態数	2 状態,4 状態,8 状態,16 状態
HMM の出力シンボル	単語
開始 · 終了状態	任意
初期状態遷移確率	ランダム
初期シンボル出力確率	ランダム
初期状態確率	均等
異なり語彙数	6418
学習データ set	odd4000
学習データ数	4000 文
単語総数	57354 単語
学習終了条件	生成確率上昇率 1%未満

4 Ergodic HMM の定性的特徴

品詞情報を付与しない単語列データにより学習された Ergodic HMM を定性的に評価した。この結果について述べる。

4.1 Ergodic HMM の解析方法

HMM は確率的信号源のクラスタリングとしての性質があるため、単語列を学習データとした場合、言語的挙動が類似した単語はネットワーク上の同一の状態遷移に大きな出力確率を持つようになる。このグループ化は従来の品詞と強い関連をもつことが考えられる。データベースにおいて単語の品詞は事前にわかっているため、次のような手順で Ergodic HMM を解析した。

1. 単語に品詞 (活用形を含む) のラベルをつける。
2. 各遷移について、単語のシンボル出力確率を品詞ごとにまとめる。
3. (状態遷移確率 × 出力確率) が 5%未満の品詞を削除してネットワークに表示する。

ネットワークの表示に用いた略号を表 3 に示す。なお以後示すネットワークの図において、状態の太細は初期確率の大小を、遷移の太細は遷移確率の大小を、品詞名、活用形の略号の右の数字は (状態遷移確率 × 出力確率) の値を示している。

表 3: 品詞・活用形の略号

品詞名	略号	品詞名	略号	品詞名	略号
普通名詞	普名	助動詞	助動	格助詞	格助
代名詞	代名	間投詞	間投	係助詞	係助
本動詞	本動	感動詞	感動	接続助詞	接助
補助動詞	補動	副詞	副詞	終助詞	終助
準体助詞	準助	接頭辞	接頭	接尾辞	接尾
未然形	未	終止形	終	連体形	体
連用形	用				

4.2 4 状態 Ergodic HMM の解析結果

4 状態の Ergodic HMM について解析した結果を図 3 に示す。この図からネットワーク上では単語が品詞ごとに偏って出力されているのがわかる。特に、活用しない品詞と活用する品詞にグループ分けされているのが見られる。以下に 4 状態 Ergodic HMM に見られる特徴を述べる。

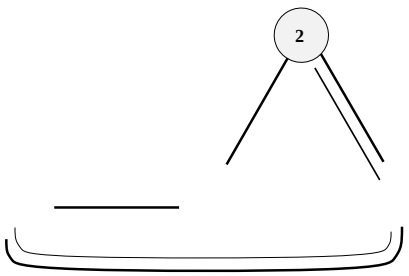

図 3: 4 状態 Ergodic HMM の解析結果

●文頭の間投詞、感動詞はイニシャルノード○ 1のループで出力されている。

●名詞は主として、状態○ 0に集まる遷移 (○ 2,○ 0⇒ ○ 0) で出力される。

●活用する品詞 (本動詞、補助動詞、助動詞、形容詞) は状態○ 2,○ 3に集まる遷移 (○ 2,○ 3⇒ ○ 2, ○ 0,○ 1,○ 2⇒ ○ 3) で出力されている。そして、連体形のものは状態○ 2に集まる遷移 (○ 2,○ 3⇒ ○ 2) で出力され、連用形のものは主に状態○ 3に集まる遷移 (○ 1,○ 2,○ 3⇒ ○ 3) で出力されている。

4.3 8 状態 Ergodic HMM の解析結果

8 状態の Ergodic HMM について解析した結果を図 4 に示す。8 状態の Ergodic HMM では 4 状態の場合と比較して、文法的に同じような役割を果たす (品詞、活用形が等しい) 単語が、より細かくグループ化されていることがわかる。また同一品詞が複数の遷移で出力されていることも多いが、この場合、遷移の起点または終点が意味を持っていることが多い。以下に、8 状態 Ergodic HMM から抽出される特徴を述べる。

●間投詞、感動詞について

間投詞 (あのー、えー、など) や感動詞 (もしもし、はい、など) がイニシャルノード (遷移を開始する状態) ○ 0からの遷移で出力されている。

●品詞の基本形と活用形について

本動詞、補助動詞、助動詞、形容詞のような活用する単語は品詞の基本形ではなく、活用形でグループ化されている。連体形は状態○ 3への遷移 (○ 4,○ 1⇒ ○ 3) で、連用形は状態○ 4への遷移 (○ 4,○ 5,○ 6⇒ ○ 4) で、助動詞の終止形は状態○ 6への遷移 (○ 1,○ 4,○ 5⇒ ○ 6) で、それぞれ出力されている。

名詞は主として状態○ 0,○ 3,○ 7から状態○ 1,○ 2に遷移する際に出力されている。しかし、各々の遷移には特徴がある。○ 3からの遷移では、形式名詞が他の名詞類よりもやや高い。また、○ 3⇒ ○ 1の遷移では準体助詞「の」が出力されている。状態○ 7からの遷移では、個々の単語出力の上位は代名詞が多い。状態○ 0からの遷移では、主として形式名詞「方」(ほう) が出力されている。また、名詞以外の単語では副詞の出力が多く見られた。

●格助詞について

格助詞は複数の遷移 (○ 2⇒ ○ 0, ○ 4, ○ 5, ○ 7), (○ 1⇒ ○ 5, ○ 7) で出力されているが、各遷移は別々のそれぞれ異なる単語を出力している。例を以下に示す。

∗状態○ 2⇒ 状態○ 0 「の」 (94%)
∗状態○ 2⇒ 状態○ 7 「から」 (28%) 「で」 (25%)
∗状態○ 2⇒ 状態○ 4 「と」 (94%)
∗状態○ 2⇒ 状態○ 5 「に」 (73%) 「で」 (18%)
∗状態○ 1⇒ 状態○ 5 「を」 (39%) 「が」 (27%)
∗状態○ 1⇒ 状態○ 7 「は」 (80%) 「も」 (7%)

4.4 16 状態 Ergodic HMM の解析結果

16 状態の Ergodic HMM について解析した結果を図 5 に示す。16 状態の Ergodic HMM では、多くの場合 1 つの状態遷移に 1 つの品詞が出力されている。また初期状態確率は 2 つの状態 (○ 2, ○ 3) で値を持っている。つまり、8 状態に比べ、より細かい品詞でグループ化がされていることがわかる、16 状態 Ergodic HMM から抽出される特徴を品詞ごとに述べる。

●間投詞、感動詞について

間投詞、感動詞は、4、8 状態の場合と同様に初期状態確率の高い状態 (○ 2, ○ 3) からの遷移で出力される。ただし、初期状態確率は状態○ 2,○ 3の 2 状態に値を持ち、初期状態確率が最も高い状態○ 2での自己ループは見られない。

●名詞について

名詞は主として、○ 5, ○ 6, ○ 7から○ 1, ○ 9, ○ 12への遷移で出力されている。

状態○ 6,○ 7からの遷移では主として形式名詞が出力され、状態○ 5からの遷移では、形式名詞以外の名詞が出力される。このように普通名詞と形式名詞の分化がおきている。なお、状態○ 5からの遷移では、出力確率が特に高い単語はなく、多種類の普通名詞の単語がこの遷移で出力されている。

●活用する品詞について

図 4: 8 状態 Ergodic HMM の解析結果

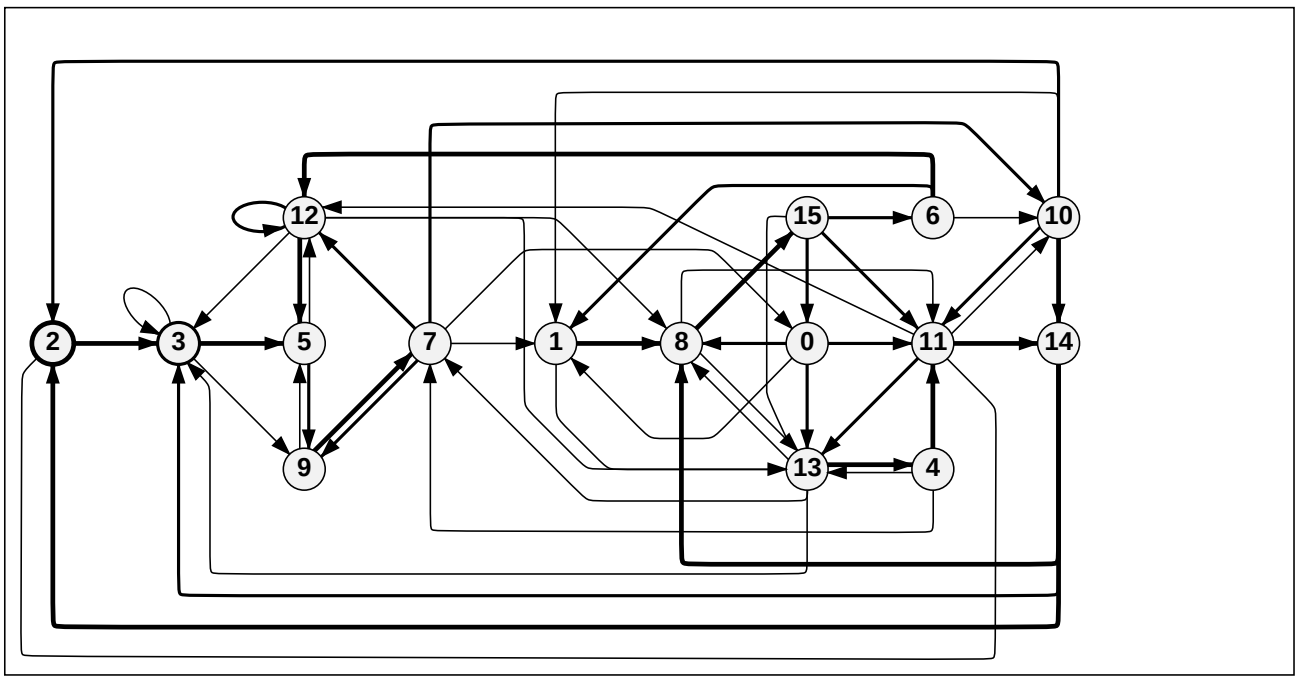

図 5: 16 状態 Ergodic HMM の解析結果

8 状態の Ergodic HMM では活用形でグループ化されていた。これに対し 16 状態では品詞の基本形でもグループ化されている。本動詞は状態○ 8○ 5からの遷移で、補助動詞は状態○ 0○ 4からの遷移で、助動詞は状態○ 0○ 1○ 1○ 1からの遷移でそれぞれ出力されている。

本動詞は複数の遷移において出力されているが、主に○ 8○ 1からの遷移で出力される。しかし、各遷移によって出力される単語に違いが見られる。また、本動詞の中でも連用形が他の活用形に比べ非常に高い。補助動詞は、○ 0○ 4からの遷移で出力される。本動詞と同様に連用形の出力が多く、助動詞をともなうことが多い。また。状態○ 0から出力される単語は表記が異なるものの意味的には「する」と同じものが多い。助動詞は、遷移の起点となる状態によって全く異なるを出力しており、8 状

態と同様に「です」「ます」がグループ化され、さらに 16 状態では「た」「たい」が他の単語からグループ化して出力されている。

●格助詞について

格助詞は、主に名詞を出力した遷移の集まる状態○8 ○9 ○91からの遷移で出力される。例を次に示す。8 状態では同じ遷移で出力されていた「で」と「に」、「が」と「を」が別々の遷移に分かれて出力されている。

∗状態○0⇒ 状態○8 「で」(86%)
∗状態○1⇒ 状態○8 「に」(87%)
∗状態○8⇒ 状態○15「と」(63%) 「お」(28%)
∗状態○9⇒ 状態○5 「から」(57%) 「で」(23%)
∗状態○9⇒ 状態○7 「の」(93%)
∗状態○12⇒ 状態○3 「は」(46%) 「ので」(35%)
∗状態○12⇒ 状態○5 「が」(55%) 「は」(24%)
∗状態○12⇒ 状態○8 「を」(75%)
∗状態○12⇒ 状態○12「を」(74%) 「について」(24%)

4.5 Ergodic HMM の定性的な評価結果のまとめ

以上のように学習された Ergodic HMM を評価するために、HMM を定性的に解析して獲得されたモデルの性質を調べた。この結果、状態遷移ごとの出力確率のパターンは、品詞対応にグループ化されている傾向が見られる。この意味で、Ergodic HMM はネットワーク文法と品詞の概念を同時に獲得する能力をある程度持つと言える。また、Ergodic HMM の状態数が多くなるに従い、品詞の活用形からグループ化が始まり、品詞の基本形に移行する傾向がある。

5 学習データ量と平均文生成確率の関係

ここでは、Ergodic HMM に必要な学習データ量を調べるために、学習データ量を変化させたときの平均文生成確率の変化を調べた。なお平均文生成確率とは各文について forward アルゴリズムによって計算された対数確率の平均値と定義した。実験ではデータベースの 8000 文を奇数番目の文の set と偶数番目の文の set とに分け、奇数番目の先頭から 1000 文の set、先頭から 2000 文の set、4000 文の set の計 3set を 2 状態、4 状態、8 状態の Ergodic HMM に学習させた。またテストデータとして偶数番目の 4000 文 (つまり text-open data) を用いた。なお、Baum-Welch の学習の結果、単語の出力確率が 0.0 になった場合、10^{-5} に置き換えた。

5.1 計算結果

学習データ量を変化させたときの平均文生成確率の変化を図 6 に示す。

図 6: 学習データ量と平均文生成確率との関係

図 6 から、学習データ量を増加するにしたがい平均文生成確率は上昇していることがわかる。仮に学習データ量が十分にあった場合、平均文生成確率に変化はないであろうと考えられるので、学習データの量はまだ不十分であると考えられる。

なお、テキストデータ中のトピックやバックグラウンドの変化などのため、言語データベースの前半部分と後半部分で対話文の内容が異なっていた。そのため、先頭から 2000 文で学習した Ergodic HMM と先頭から 4000 文で学習した Ergodic HMM では、ネットワークの定性的特徴も大きく異なっていた。これが平均文生成確率に変化を及ぼした可能性もある。

6 文音声認識への適用

Ergodic HMM で獲得した確率つきネットワーク文法を音声認識における言語モデルとして使用して、この有効性を調べた。その概要と結果を次に述べる。

6.1 文音声認識の実験条件

この実験では HMM+Viterbi サーチ [2] を基本アルゴリズムにした。音響尤度の計算には 5240 単語発声を用いて学習した混合連続分布型連続 HMM による音素モデルを用い、言語モデルとして odd4000 で学習した状態数 2,4,8 の Ergodic HMM を使用した。また言語モデルと音響モデルの結合の際、Ergodic HMM から得られる値を 16 倍して Viterbi algorithm で計算された対数音響尤度の値に加えた。認識実験ではテストデータとして学習データと同一タスクの 38 文を用いた。なお同一のテストデータを使って text-open および text-closed data の文音声認識実験を行なうために、text-open の実験には odd4000 で学習した

Ergodic HMM のパラメータを使用し、text-closed の実験では odd4000 で学習したパラメータを初期値とし、odd4000 にテストデータ 38 文を加えた 4038 文を再び学習させた Ergodic HMM のパラメータを使用した。その他の実験条件を表 4 に示す。

表 4: 文音声認識の実験条件

基本アルゴリズム	continous mixture HMM + word Ergodic HMM + beam search
音素モデル数	52 音素
音素音響モデル	4 状態 3 ループ混合分布型 HMM 混合数は音素ごとに異なる。 継続時間長制御なし。
話者	男性アナウンサー 1 名 (MAU)
音響パラメータ	log パワー + 16 次 LPC ケプストラム +Δlog パワー + 16 次 LPCΔ ケプストラム
音響分析条件	サンプリング周期 12kHz フレーム窓長 20ms フレーム周期 5ms
ビーム幅	4096
認識語彙数	435 単語
テストデータ	同一話者発声 (MAU) 38 文

6.2 文実験結果

図 7 に実験結果を示す。比較のため、言語モデルを用いない場合 (つまり音響モデルのみで認識した場合) と言語モデルとして単語 bigram を用いた場合を合わせて示す。実験から text-closed データ, text-closed データともに言語モデルがない場合に比べ高い認識率が得られることや状態数が増えるにつれ認識率が上昇することが確認された。この観点で Ergodic HMM は言語モデルとして有効であると言える。また、Ergodic HMM と単語 bigram を比較すると、text-closed data の実験では Ergodic HMM は単語 bigram に及ばなかったが、text-open data の実験では単語 bigram と同等か、やや上回る認識性能が得られた。

7 モデルの初期パラメータ依存性

Baum-Welch algorithm は学習データの尤度を極大値に収束するようにパラメータを学習するが、必ずしも最大値に収束しないためパラメータの初期値の設定が重要となる。このため異なる初期値の 8 状態 Ergodic HMM を 8 個用意し、平均文生成確率と文音声認識率の関係を調べた。学習の終了条件を再推定回数 20 回とし、その他の条件は以前の実験と同一とした。学習後の平均文生成確率と text-closed data に対する文認識率の関係を図 8 に示す。この図からパラメータの初期値によってモデルの平均文生成確率が異なること、また認識率の最高値と最低値の差は約 13%(5 文) あることがわかる。また、サンプルが 8 個しかないが平均文生成確率と認識率の間にはほぼ正の相関が見られる。

8 考察

8.1 Ergodic HMM による文法の自動獲得の可能性

Ergodic HMM の定性的な解析から HMM の状態遷移出力確率の分布に品詞のグループ化が観測された。このような意味で Ergodic HMM はネットワーク文法を自動的に獲得できる可能性があると考えられる。また、4 状態、8 状態、16 状態と Ergodic HMM の状態数が増えるごとに、より細かいグループ化が観測されたことから、今後さらに状態数を多くすることによって、より詳細な文法を獲得できる可能性がある。

なお、現在の品詞体系は基本的には構文的な観点で構築されている。しかし、Ergodic HMM において 1 つの遷移に出現する単語のグループを品詞とみなすことによって、尤度最大という基準での別の品

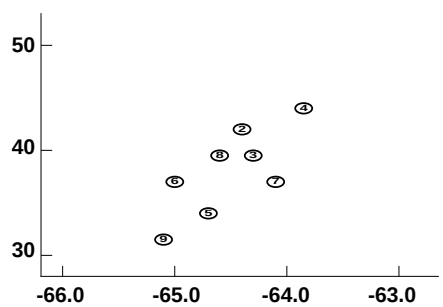

図 7: 文音声認識結果

図 8: 平均文生成確率と文認識率の関係

詞体系を自動構築する可能性がある。

8.2 学習データ量について

Ergodic HMM のパラメータの推定に必要な学習データ量の明確な基準は知られていないが、少なくとも HMM のパラメータ数と同程度以上の学習データ量が必要であると思われる。今回の実験では、8 状態 Ergodic HMM が持つパラメータ数は合計 410752 個あるのに対し、学習データは 57354 単語であるため、十分な学習量とは言えない。学習データ量に対する平均文生成確率の変化を見ても、学習データの内容が前半と後半では異なっている問題があるが、これを裏づけていると思われる。今後、学習データ量を増やしたときの平均文生成確率やエントロピーおよび HMM の定性的な構造の変化を検討する必要がある。

8.3 文音声認識実験の結果について

文音声認識の実験は、規模が小さいので再現性について確かめる必要があるが、現在の結果から考えられることを述べる。Text-open data、text-closed data 両方に対して、言語モデルとして Ergodic HMM を用いたほうが、用いない場合より高い認識性能を示した。この点では Ergodic HMM は一般的のある言語モデルをある程度獲得しているといえる。しかし、認識誤りを生じた文を見ると非文と思われる文が多いことから、8 状態の Ergodic HMM では十分に言語情報を獲得しているとは言えない。今後状態数を増加させたときの認識率の変化を追跡する必要があろう。また、Ergodic HMM と単語 bigram を比較すると、text-closed data の実験では Ergodic HMM は単語 bigram に及ばなかったが、text-open data の実験では単語 bigram と同等か、やや上回る値が得られた。text-open の実験では学習データとテストデータの選択の方法により、認識性能が大きく異なる可能性があるので、この実験だけからでは明言できないが、Ergodic HMM は bigram より言語モデルとして優れている可能性がある。最後に、音響尤度に対する言語尤度の重み (現在 16) は Ergodic HMM を言語モデルとしたとき認識性能が最大になるように実験的に選んだ値であるため、今後理論的な根拠を与える必要がある。

8.4 初期モデルについて

初期パラメータを変えた実験結果から Ergodic HMM は初期値によって収束したモデルの尤度が異なることが示された。最適な初期値の決定法は今のところ知られていないが、文の生成確率と認識率が正の相関を持っていることから有効な一つの方法として、いくつかの初期モデルで学習をある 1 定回数繰り返した後、最も高い尤度を持つモデルについて学習を続行する方法が考えられる。

9 結論

本論文では、Ergodic HMM と確率つきネットワーク文法が類似した構造を持ち同種のパラメータで表現されること、および Baum-Welch algorithm で HMM のパラメータを学習できることに着目し Ergodic HMM によるネットワーク文法の獲得の可能性を検討した。

Ergodic HMM にテキストデータの単語列を学習させ、学習後の HMM を解析した結果、Ergodic HMM は各遷移ごとに単語の品詞と関連してグループ化がされていることがわかった。また Ergodic HMM の状態数が増えるほど、細かくグループ化が行なわれていることがわかった。また、連続音声認識実験において認識率が向上することから Ergodic HMM は一般的な言語情報を獲得する可能性がある。

以上のように、Ergodic HMM を用いることによって大量のテキストデータから確率つきネットワーク文法が自動的に獲得できる可能性があること、同時に新しい品詞体系も得られる可能性があることが示唆された。

謝辞

本研究を進めるにあたって、ATR 自動翻訳電話研究所の皆様に懇切なる御指導を賜わりました。特に Ergodic HMM の解析に関しては、浦谷氏、竹沢氏の協力を得ました。心から感謝の意を表します。

なお本研究は村上、嵯峨山のアイデアの下に山本が実験をしたもので、より詳細なデータは [13] に載せているので、興味あるかたはこれを参照して下さい。

参考文献

[1] K. Kita, T. Kawabata, T. Hanazawa, “HMM speech recognition using stochastic language models”, ICASSP’90 J. Acoust. Soc. Jpn. (E) 12,3 *pp.99–105* (1991)

[2] K.F.Lee, “Large-Vovabulary Speaker Independent Continuous Speech Recognition: The SPHINX System,” 15213 CMU-CS-88-148 (April 18, 1988).

[3] 村上, 嵯峨山, “単語の trigram を用いた連続音声認識の一アルゴリズム,” 音学講論, 2-Q-7, pp.185-186 (1992-10).

[4] F. Jelinek, R.L.Mercer, S.Roukos, “Principle of lexical language modeling for speech recognition,” In S. Furui and M. M. Sondhi, editors, Advances in Speech Signal Processing, pp 651–699, Marcel Dekker, Inc., New York, New York. (1992).

[5] K.Lari, S.J.Young, “The estimation of stochastic context-free grammars using the Inside-Outside algorithm,” speech recoginition, Computer Speech and Language, vol.**4**, pp.35-56 (1990).

[6] J.Kupiec, “Robust part-of-speech tagging using a hidden Markov model,” Computer Speech and Language, vol.**6**, pp.225-242 (1992).

[7] 村瀬, 上田, 中川, “文脈自由文法と bigram,trigram による言語のモデル化の検討,” 電子情報通信学会技術研究報告, SP90-**75**, pp.49–56 (1990).

[8] 田本, 村上, 嵯峨山, “HMM を利用した言語獲得の可能性について,” 人工知能学会研究会資料, SIG-SLUD-9201-**6**, pp.47–54 (1992).

[9] 山本, 村上, 嵯峨山, “HMM を言語モデルに用いた連続音声認識の検討,” 情報処理学会第 45 回全国大会予稿集, vol.**3**, pp.227–228 (1992.10).

[10] 山本, 村上, 嵯峨山, “HMM を言語モデルに用いた連続音声認識の検討,” 日本音響学会秋季講演論文集, vol.**1**,pp.193–194 (1992.10).

[11] X.D.Huang, Y.Ariki, M.A.Jack, “Hidden Markov models for speech recoginition,” Edinburgh University Press (1990).

[12] 江原, 小倉, 森本, “電話対話ベースの構築,” 情報処理学会第 40 回全国大会予稿集, vol.**1**, pp.486–487 (1990).

[13] 山本寛樹, “HMM による言語モデルの自動獲得の検討,” 早稲田大学理工学研究科修士論文 (1993).