

TR-IT-0039

SSS-ToolKit (Ver3.0) ユーザーズ・マニュアル

SSS-ToolKit (Ver3.0) User's Manual

鷹見 淳一

Jun-ichi TAKAMI

1994.2

概要

逐次状態分割法 (Successive State Splitting: SSS) を用いた隠れマルコフ網 (Hidden Markov Network: HMnet) の自動生成、混合連続 HMM の生成、移動ベクトル場平滑化手法 (Vector Field Smoothing: VFS) 等、主要な技術のプログラムの他、それらを使用して研究を進めていく上で有用ないくつかのプログラムを含む SSS-ToolKit.ver3 の使用方法について示す。

©ATR 音声翻訳通信研究所

©ATR Interpreting Telecommunications Research Labs.

目次

1	はじめに	1
2	SSS-ToolKit.ver3 の概要	5
3	モデルファイルのフォーマットについて	7
4	音声サンプルのフォーマットについて	11
5	各プログラムの使用法	13
5.1	HMnet の自動生成 : Exe.train_HMnet	13
5.2	HMnet の構造抽出 : Exe.topology_HMnet	16
5.3	HMnet の再学習 : Exe.retrain_HMnet	17
5.4	HMnet の環境補充 : Exe.fill_HMnet	18
5.5	HMnet の初期構造生成 : Exe.init_HMM	19
5.6	混合連続分布 HMM の作成 : Exe.train_HMM	21
5.7	HMnet のファイル合成 : Exe.compose_HMnet	22
5.8	HMnet の話者適応 : Exe.adapt_HMnet	23
5.9	HMnet の構造解析 : Exe.analysis_HMnet	24
5.10	HMnet の出力尤度評価 : Exe.evaluate_HMnet	25
5.11	HMnet の状態経路探索 : Exe.path_HMnet	26
5.12	HMnet の展開 : Exe.expand_HMnet	27
5.13	HMnet の状態継続時間算出 : Exe.viterbi_HMnet	28
5.14	HMnet のカテゴリ識別性能評価 : Exe.recognize_HMnet	30
6	SSS-ToolKit.ver3 の動作確認	33
6.1	Exe.train_HMnet の動作確認	35
6.2	Exe.topology_HMnet の動作確認	35
6.3	Exe.retrain_HMnet の動作確認	35
6.4	Exe.fill_HMnet の動作確認	35
6.5	Exe.init_HMM の動作確認	36

6.6	Exe.train_HMM の動作確認	36
6.7	Exe.compose_HMnet の動作確認	36
6.8	Exe.adapt_HMnet の動作確認	37
6.9	Exe.analysis_HMnet の動作確認	37
6.10	Exe.evaluate_HMnet の動作確認	37
6.11	Exe.path_HMnet の動作確認	37
6.12	Exe.expand_HMnet の動作確認	38
6.13	Exe.viterbi_HMnet の動作確認	38
6.14	Exe.recognize_HMnet の動作確認	38
7	おわりに	39
8	参考文献	41
A	SSS-ToolKit.ver3 で使用されるモデルファイルの例	45
B	HMnet 生成時のログファイルの例	49

第 1 章

はじめに

この SSS-ToolKit.ver3 は、逐次状態分割法 (Successive State Splitting: SSS) や、それにより生成される隠れマルコフ網 (Hidden Markov Network: HMnet) について研究するためのソフトウェアです。SSS-ToolKit.ver1 および SSS-ToolKit.ver2 に対して改良を加え、より使い易くしたものとなっています。

SSS-ToolKit.ver1 からの改良点は以下の通りです。

- 音声サンプルのフォーマットを全てのプログラムで共通化した。(SSS-ToolKit.ver1 では、処理の内容によって音素サンプルフォーマットと単語サンプルフォーマットを使い分ける必要がありました。)
- SSS による HMnet のトポロジー生成、およびその分布形状変更のための再学習の際に、連結学習を行なうことが可能となった。(これにより、音素境界の与えられていないデータに対しても SSS が適用できます。)
- モデル読み込みルーチンを改良し、モデルファイル中に記述されているブロック ([キーワード] で始まる一連の情報、3章参照) の出現順序に制約がなくなった。(SSS-ToolKit.ver1 では、各ブロックの記述順序を変更した場合、そのモデルは読み込めませんでした。)
- 上記改良のため、モデルファイルのフォーマットに以下の変更を加えた。
キーワードの変更: #max_segment (旧) → #max_path_length (新)、
フォーマットの変更: #element (旧) → #element [factor_index] (新)。
- 話者適応プログラム (Exe.adapt_HMnet) に、適応の対象となるモデルパラメータを指定する機能を加えた。(SSS-ToolKit.ver1 では、平均値ベクトルのみの適応が可能でした。)
- オプション名の一部を変更した。(この改良は本質的なものではなく、旧バージョンのプログラムとの間で混乱を招く恐れもあるため、その変更についてはいろいろと考慮しましたが、オプションの持つ意味をそのオプション名にできる限り反映させ覚えやすくするため

に、本バージョンでは、あえて変更致しました。旧バージョン用のシェルスクリプトを用いて実行する場合には十分に注意して下さい。)

また、SSS-ToolKit.ver2からの改良点は以下の通りです。

- HMnetの自動生成用プログラム(Exe.Train_HMnet)のアルゴリズムの一部を変更した。(この変更は、状態分割の可能性を判定する方法に関するものです。これまでのアルゴリズムでは、ある状態の分割が可能か否かの判定を、環境の分割可能性のみに着目して行なっていました。従って、ある状態での環境が細分化され、各環境要因の要素数が全て1になると、たとえ時間方向への状態分割の可能性が残されていても、その状態はそれ以上分割されませんでした。今回の改良ではこの点を修正し、環境方向と時間方向、さらに3D-SSSの場合には共通化環境方向まで含めて、いずれか一つにでも分割可能性が残されている間は状態分割を継続するようにしました。以上のような変更は、分割がある程度進み、環境がかなり細分化されてからのHMnetの構造に影響を与えるため、ここで得られるHMnetの構造はSSS-ToolKit.ver2で生成したものと一致しない場合がありますが、実用上の精度に大きな差はありません。)
- 未定義環境の補充(Exe.fill_HMnet)、初期モデルの構造生成(Exe.init_HMnet)、ラベル列に対応する状態経路の検索(Exe.path_HMnet)、HMnetの展開(Exe.expand_HMnet)、状態継続時間の算出(Exe.viterbi_HMnet)といった、幾つかの新しい機能を追加した。(特に、未定義環境の補充機能については、これまでのHMnetのように、未定義環境の存在を常に意識しなければならないといった厄介な問題を解消することのできる有効な機能だと考えています。未定義環境を補充するためには、学習用サンプル中に現れていない環境での音響パターンの振舞いを予測(予想)するという、非常に難しい問題を解く必要があります。今回開発したアルゴリズムで十分に高い精度が得られている保証はありません。しかし、実用上はほとんど問題なく、未定義環境が存在するために処理が繁雑になったり計算量が増加したりするといった弊害を避けるためにも、生成直後のHMnetには未定義環境の補充処理を行なうことをお勧めします。)
- HMnetの学習(再学習)や話者適応など、学習用サンプルを保存するために多くのメモリが必要となる可能性のあるプログラムについて、学習用サンプルの音響パラメータをメモリ上ではなくディスク上のファイルに置くことによって、メモリ使用量を節約する機能を追加した。(ただし、この機能を使用する場合、処理時間は多少増加します。)
- HMnetの学習(再学習)や話者適応など、自動的にログファイルが形成されるタイプのプログラムに、ログファイルに書かれる内容と同じものを標準エラーにも出力させるモニタ機能(オプションにより選択可能)を追加した。
- 各プログラムの実行時に生じたエラーの内容をユーザーに知らせるための表示機能を充実させた。

- その他、使い易さを向上させるための細かい部分での変更を行なった。

第 2 章

SSS-ToolKit.ver3 の概要

この SSS-ToolKit.ver3 は、以下の処理を行なうためのプログラムを含んでいます (図 2.1 参照)。

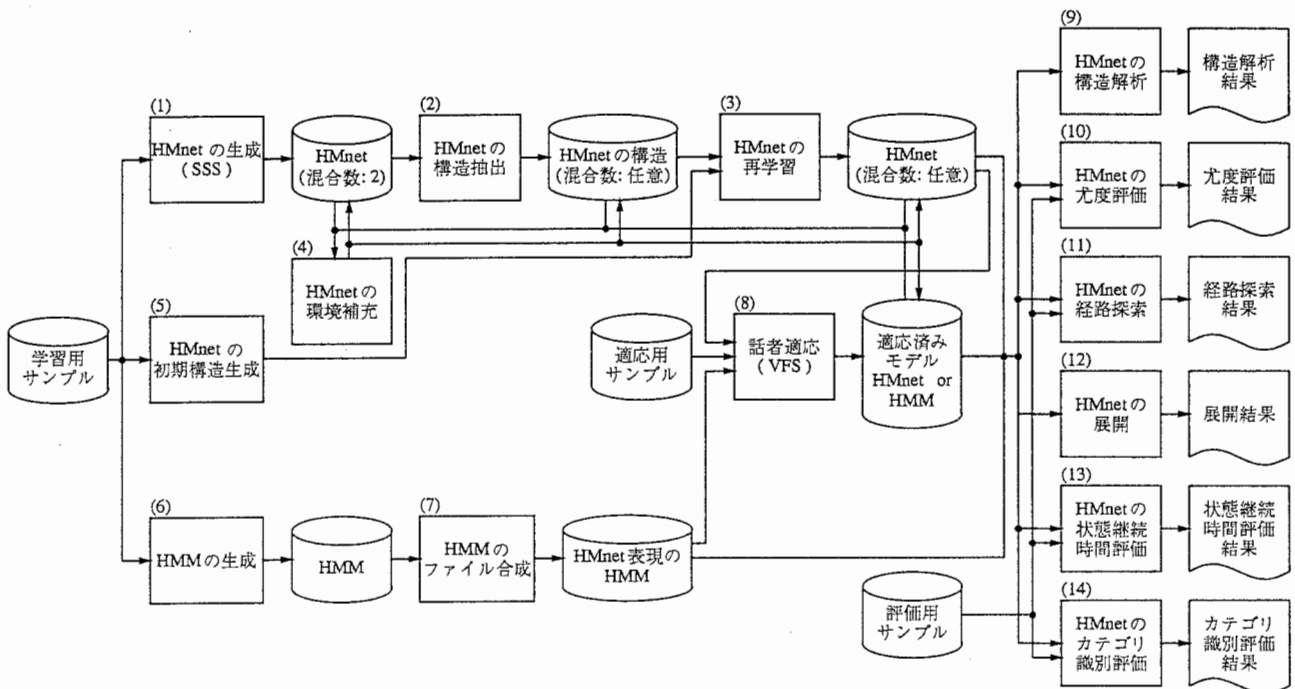


図 2.1: SSS-ToolKit.ver3 の処理内容

1. HMnet の自動生成 : Exe.train_HMnet

逐次状態分割法 (SSS, 3D-SSS, SP-SSS のいずれか) により、混合数 2 の混合ガウス分布を出力確率分布として持つ HMnet(original) を作成します。連結学習も可能です。

2. HMnet の構造抽出 : Exe.topology_HMnet

HMnet (original) からその構造を抽出して、出力分布の混合数を最終的に希望するものに変更した HMnet(topology) を作成します。

3. HMnet の再学習 : Exe.retrain_HMnet

HMnet(topology) を再学習し、最終的な HMnet を作成します。連結学習も可能です。

4. HMnet の環境補充 : Exe.fill_HMnet

未定義の環境に関する補充 (穴埋め) を行ない、全環境の直積空間内で表現可能な全ての環境に対応できる HMnet を作成します。

5. HMnet の初期構造生成 : Exe.init_HMnet

仕様ファイルで指定された HMnet(topology) の初期構造を生成します。

6. 混合連続分布 HMM の作成 : Exe.train_HMM

比較実験などに使用するための、通常の混合連続分布 HMM (ただし、モデルフォーマットは HMnet と共通であるため、他のプログラムでは HMnet と全く同様に扱うことができます) を作成します。連結学習も可能です。

7. HMnet のファイル合成 : Exe.compose_HMnet

それぞれ独立して作成された HMnet を、一つの大きなモデルとして合成します。

8. HMnet の話者適応 : Exe.adapt_HMnet

移動ベクトル場平滑化法 (VFS) により、HMnet の話者適応を行ないます。

9. HMnet の構造解析 : Exe.analysis_HMnet

HMnet の構造上のスペック (異なり経路数、延べ状態数、モデル化効率など) を解析します。

10. HMnet の出力尤度評価 : Exe.evaluate_HMnet

入力音声に対して得られる HMnet の出力尤度を計算します。

11. HMnet の状態経路検索 : Exe.path_HMnet

入力音声のラベル列に対応する HMnet 上の状態経路を検索します。

12. HMnet の展開 : Exe.expand_HMnet

HMnet を全展開し、それぞれのモデルがどのような環境を表現し、どのような状態列で構成されているかを調べます。

13. HMnet の状態継続時間算出 : Exe.viterbi_HMnet

入力音声に対して Viterbi 経路を計算し、HMnet の各状態での継続時間を求めます。

14. HMnet のカテゴリ識別性能評価 : Exe.recognize_HMnet

HMnet を用いたカテゴリ識別実験を行ないます。

第 3 章

モデルファイルのフォーマットについて

この SSS-ToolKit.ver3 では、HMnet や通常の HMM を、全て共通のファイルフォーマットとして表現することによって、話者適応やカテゴリ識別など、ほとんどのプログラムを共通に使用することができるように考慮されています。(SSS-ToolKit.ver2 からの変更はありません。)

モデルフォーマットは、“#[キーワード]”で始まるブロックと、“;”で始まるコメント行(3 タイプのコメントがあります)によって記述されています。コメント行以外の各ブロックは、必ず“#”で始まり空行で終わります。コメントのタイプと各キーワードの意味を以下に示します。

- ; (コメント type1)

一般的なコメント行を表します。主にファイル中で各キーワードの意味を示すために使用されています。

- ;; (コメント type2)

このコメント行は、主にファイルの処理に関わる条件等を保存するために使用されています。SSS-ToolKit.ver3 中のプログラムによって作成、あるいは変更されたファイルでは、そのときの生成、変更条件等を自動的にこのタイプのコメントとしてファイル中に保存します。このタイプのコメント行は、一つのファイルに対して複数の処理を行なった場合にその処理の順に累積されていきます。そのため、処理の履歴などを保存ために便利なものとなっています。

- ;;; (コメント type3)

このコメント行は、主にヘッダとしての意味が強いものです。SSS-ToolKit.ver3 中のプログラムによって作成されたモデルファイルの先頭には必ず入ります。

- #type

モデルフォーマットのタイプを表すブロックです。現在は“HMnet”の指定のみ可能です。将来の拡張用です。

- #covariance”

共分散行列の表現法を表すブロックです。現在は“diag”の指定のみ可能です。将来の拡張用です。

- **#dimension**
音響パラメータの次元数を表すブロックです。
- **#max_mix**
このフォーマットでは、各出力確率分布毎に混合数を任意設定できます。これは、そのうちの最大の混合数を表すブロックです。
- **#total_dist**
モデル全体で使用されている多次元単一ガウス分布の総数を表すブロックです。
- **#total_outp**
モデル全体で使用されている出力確率分布の総数を表すブロックです。
- **#dist**
個々の多次元単一ガウス分布のパラメータを表すブロックです。
- **#outp**
個々の出力確率分布のパラメータを表すブロックです。出力確率分布は幾つかの多次元単一ガウス分布の重み付き加算によって構成されます。
- **#total_state**
モデル全体で使用されている隠れマルコフ状態の総数を表すブロックです。
- **#max_path_length**
一つのHMM(経路) 当たりに使用される隠れマルコフ状態の数の最大値を表すブロックです。
- **#factor_num**
環境要因数を表すブロックです。
- **#element**
各環境要因の全要素を表すブロックです。通常は、音素ラベルなどがこの要素となります。
- **#tied_element**
3領域逐次状態分割法(3D-SSS) や話者並列逐次状態分割法(SP-SSS) を行なう際の、共通化すべき環境の要因(以降、共通化環境要因と呼ぶ)の全要素を表すブロックです。通常は話者名などがこの要素となります。この場合、話者適応を行なうとモデルの生成や話者適応に関わった話者名がここに列記されます。
- **#state**
個々の隠れマルコフ状態のパラメータを表すブロックです。ここには、その状態で受理される環境要素や、その状態に先行 / 後続する状態の番号、そこで使用される出力確率分布の

番号、その状態の推定に使用された音声サンプル数、状態遷移確率などの情報が記述されています。

付録 A に、HMnet ファイルの例を示します。

第4章

音声サンプルのフォーマットについて

この SSS-ToolKit.ver3 中のプログラムで使用される音声サンプルは、全て共通のフォーマットで表現されています。

音声サンプルフォーマットの内容を図 4.1 に示します。

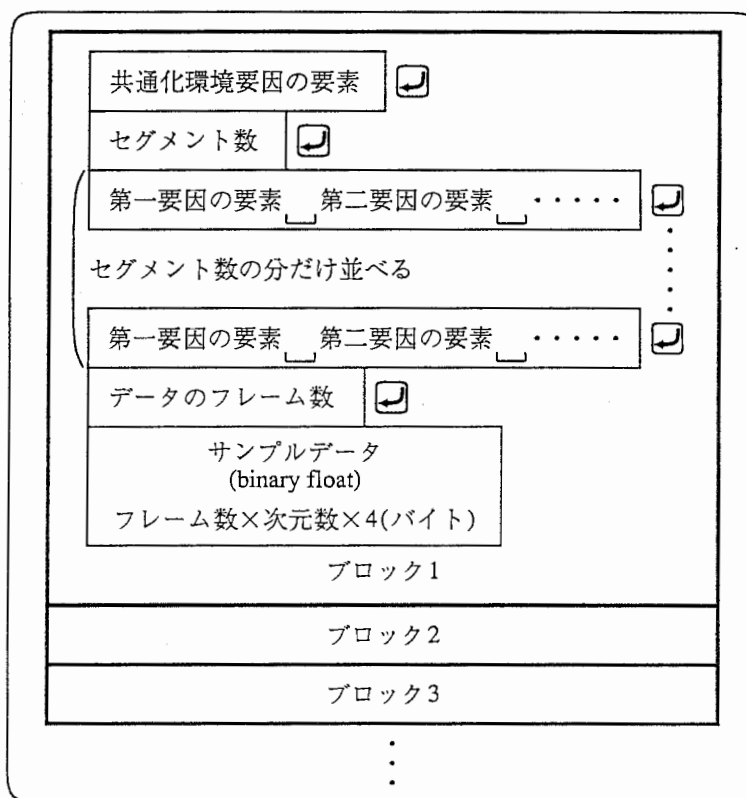


図 4.1: 音声サンプルのフォーマット

なお、このフォーマットでは、“セグメント数”を1とすることで孤立サンプルを表現することができます。HMnet の自動生成時や HMnet の再学習時に、入力音声サンプル中にこの“セグメント数”が2以上のものが1つでも存在する場合には、入力された音声サンプルが孤立サンプルではなく、融合ラベルを含むサンプルであるとみなし、自動的に連結学習モードに切替えられます。

第 5 章

各プログラムの使用法

5.1 HMnet の自動生成 : Exe.train_HMnet

このプログラムは、SSS によって HMnet(original) を作成するためのものです。

一話者の音声サンプルからその話者に専用の HMnet(original) を作成する SSS の他、複数話者の音声サンプルから、話者に共通の HMnet(original) を作成するための 3D-SSS や SP-SSS なども、オプションの切替えによって行なうことができます。

実行は、

Exe.train_HMnet [オプション] < “音声サンプルファイル”

によって行ないます。

このプログラムでは、音声サンプルファイルが孤立サンプルと見なせる場合 (音声サンプルファイル中の全てのブロックで“セグメント数”が1となっている場合) には切り出し学習モードとなり、それ以外の場合には連結学習モードに入ります。

連結学習モードでは、初期モデルの段階で、ある程度のセグメンテーション性能を確保する必要があるため、初期モデルとしては全サンプルを1つのモデルで表したものの代わりに、中心環境 (“-mf” オプションで指定した要因) の要素数分のモデルを使用します。従って、この場合の初期モデルの状態数は、 $[\text{中心環境の要素数}] \times [\text{初期モデルの状態数: “-ip” オプションで指定}]$ となります。

オプションとしては、以下のものが許されています。

- -cm (省略可能)

作成されるモデルファイル中に保存したいコメントがある場合、このオプションに続けて “...” の形で指定します。このオプションは、一回につき一行のみの記入が可能です。コメントが複数行に渡る場合には、このオプションを複数回指定して下さい。

- -cf (デフォルト: 0)

何らかの理由により計算の途中で中断してしまった HMnet(original) の生成を途中から再開したい場合、このオプションで1を指定します。中断直前の分割数の HMnet から学習を続けることができます。途中の HMnet は全て ASCII ファイルとして保存しており、精度が十分でないため、中断せずに作成したモデルのパラメータとの間で多少の誤差が出る場合がありますが、実用上はほとんど問題ありません。

- -di (デフォルト: 34)

音響パラメータの次元数を指定します。このプログラムに与える音声サンプルに依存します。

- -dm (デフォルト: NULL)

学習用サンプルの音響パラメータを格納する一時ファイルのリストを指定します。これによって計算時に使用されるメモリ量を節約することができます。このファイルのフォーマットは、ファイルの名前とそのサイズの上限(単位: byte)を、間に空白を挟んで記述したものです。複数の一時ファイルを使用したい場合には、このような行を、間に改行を入れながら複数行記述してください。なお、ここで作成される一時ファイルは、プログラムが正常に終了した場合(正常な動作を行なった後での終了、および“@ERROR ...”のメッセージを出して終了した場合)や、シグナル SIGINT(インタラプトシグナル: 通常は ctrl-C などで発生するもの)を受けとった場合には自動的に消去されますが、それ以外の異常終了時には、消去されずに残されたままになります。この場合には、残っている一時ファイルをマニュアルで消去して下さい。

- -fn (デフォルト: 3)

環境の要因数を指定します。このプログラムに与える音声サンプルに依存します。

- -ip (デフォルト: 1)

初期モデルの状態数(連結学習モードでは一モデル当たりの状態数)を指定します。特に連結学習を行なう場合には、初期モデルの段階である程度のセグメンテーション性能を確保する必要があるため、この値をある程度大きく設定して下さい。

- -mf (デフォルト: 1)

全環境要因の中で、中心となる要因のインデックスを指定します。このプログラムに与える音声サンプルに依存します。例えば、環境要因として先行音素、当該音素、後続音素の3要因をこの順で与えた場合には、先行音素がインデックス0、当該音素がインデックス1、後続音素がインデックス2となり、中心となるインデックスは1となります。このプログラムに与える音声サンプルに依存します。

- -mp (デフォルト: 4)

一本の経路当たりの状態数の上限を指定します。これによって SSS における時間方向への状態分割数に制約を与えることができます。SSS-ToolKit.ver3 で扱っているモデルフォー

マットでは、状態間のヌル遷移が認められていないため、ある経路を構成する状態数より短いフレーム数を持つ音声サンプルは、その経路上で評価することができません。そのため、通常は、認識時に使用する音声サンプルの最小フレーム数以下にこの値を設定して下さい。

- -ms (デフォルト: 600)
最大の状態分割数を指定します。ここで指定した状態数に達すると、プログラムは自動的に終了します。なお、一旦完成した HMnet(original) の状態数をさらに増やしたい場合は、“-cf” オプションによって分割を再開することができます。
- -of (デフォルト: “out_file”)
ここで作成される HMnet(original) を書き込むためのファイル名を指定します。実際には、ここで指定したファイル名の後に、何らかの拡張子を付けたものが使用されます。これについては、後で詳しく説明します。
- -pr (デフォルト: NULL)
環境を分割していく際の、各要因に対する分割の許可や優先順位を指定します。要因のインデックスと “/”、“!” および “,” の記号を含む文字列により指定します。[要因 A]/[要因 B] と指定した場合、要因 A は要因 B に優先して分割されます。また、[要因 A]! とした場合には、要因 A は分割されなくなります。複数の条件は “,” をはさんで列記して下さい。(“1/0, 1/2” や、“1/0, 2!” のように指定します。文字間のスペースは無視されます。)
- -ss (デフォルト: 10)
状態分割の過程で生成される HMnet(original) を保存する頻度を指定します。状態数をここで与えた数で割った時の余り(剰余)が 0 の場合に、その HMnet を保存します。実際には、ここで指定される状態数の HMnet の他に、“-cf” オプションによる途中再開に備えて直前の分割数の HMnet(original) と、“-sn” で指定された最大状態数の HMnet(original) が保存されます。
- -ty (デフォルト: “SSS”)
アルゴリズムのタイプを指定します。通常は、共通化環境要因の要素数が 1 の場合には “SSS” そうでなければ “3D-SSS” または “SP-SSS” とします。
- -vb (デフォルト: 0)
このオプションで 1 を設定することにより、ログファイルに書かれる内容と同一のものを標準エラーにも出力することができます。実行状況を監視するために使います。

このプログラムによって、モデルファイルおよびログファイルが作られます。

モデルファイルは、test.600 のように、“-of” で指定したファイル名の後に “.状態数” という拡張子が付けられたファイル名として書き込まれます。また、HMnet 生成の途中再開に備えて

分割の度に書き込まれる一時ファイルには、ファイル名の後に“.current”という拡張子が付けられます。このファイルは、“-cf”オプションで1を指定した場合に使用されます。なお、ここで作成されているHMnet(original)の出力確率分布は、SSSの制約上、全て混合数2の混合ガウス分布が割り当てられます。

ログファイルには、“-of”で指定したファイル名の後に“.log”という拡張子が付けられます。このログファイルは、状態分割がどのように行なわれたかを示す貴重な情報が保存されています。“-cf”オプションによりHMnet生成を途中再開する場合や、Exe.fill_HMnetにより、未定義環境の補充を行なう場合には、このログファイルが必要です。“-cf”オプションにより学習を途中から始める場合、学習再開直後にこのログファイルは一旦“ファイル名.log.bak”というファイル名に書き換えられ、“ファイル名.log”には新たに学習再開後のログが書き込まれます。(このような手続きを行なうために、途中再開したプロセスがエラーなどにより異常終了した場合、正しいログファイルは“ファイル名.log.bak”の方に移っており、“ファイル名.log”の内容は既に壊されています。この直後に再びHMnetの途中再開を行なうと、正しい内容が書き込まれている“ファイル名.log.bak”が壊され、復旧が不可能になりますので、このような場合には、必ずマニュアル操作で“ファイル名.log.bak”を“ファイル名.log”に戻して下さい。)

付録Bに、このプログラムから出力されるログファイルの例を示します。

5.2 HMnetの構造抽出：Exe.topology_HMnet

このプログラムは、Exe.train_HMnetによって作成された混合数2の出力確率分布を持つHMnet(original)の混合数を最終的に使用したいものに変更するための前処理として、HMnet(original)の構造を取り出し、その混合数の部分をオプションで与えた数に変更するためのものです。特に、3D-SSSやSP-SSSで作られたHMnet(original)に対しては、話者に共通な構造を抽出するために必要な処理です。

実行は、

Exe.topology_HMnet [オプション]

によって行ないます。

オプションとしては、以下のものが許されています。

- -cm (省略可能)

Exe.train_HMnetと同じ。

- -if (デフォルト: “in_file”)

構造を抽出したいHMnet(original)のファイル名を指定します。通常は、Exe.train_HMnetで作成されたモデルファイルを指定します。

- -mn (デフォルト: 1)
各状態の出力確率分布の混合数を指定します。
- -of (デフォルト: "out_file")
処理後の HMnet(topology) を書き込むファイル名を指定します。

このプログラムによって作成される HMnet(topology) は、まだパラメータ推定が行なわれていません。そのため、この HMnet(topology) のパラメータにはダミーの数値が入れられていますので、次の再学習の処理を終えるまで実際に使用することはできません。

5.3 HMnet の再学習 : Exe.retrain_HMnet

このプログラムは、Exe.topology_HMnet や Exe.init_HMnet などによって作成された HMnet(topology) を始め、モデルのファイルフォーマットの規則に従って記述された任意のモデルを再学習するためのものです。

実行は、

Exe.retrain_HMnet [オプション] < “音声サンプルファイル”

によって行ないます。

オプションとしては、以下のものが許されています。

- -cm (省略可能)
Exe.train_HMnet と同じ。
- -di (デフォルト: 34)
Exe.train_HMnet と同じ。
- -dm (デフォルト: NULL)
Exe.train_HMnet と同じ。
- -fn (デフォルト: 3)
Exe.train_HMnet と同じ。
- -if (デフォルト: "in_file")
パラメータを再学習したいモデルを指定します。通常は、Exe.topology_HMnet で作成された HMnet(topology) を指定しますが、モデルのファイルフォーマットの規則に従って記述されたものであれば、任意のモデルを指定することができます。
- -of (デフォルト: "out_file")
再学習したモデルを書き込むファイル名を指定します。

- -rp (デフォルト: 1)

処理の開始時にモデルのパラメータを初期化するか否かを指定します。Exe.topology_HMnetで作成したHMnet(topology)は、モデルパラメータにダミーの数値が記述されているため、必ずパラメータを初期化する必要があります。ただし、連結学習時に音素切り出し学習により作成されたモデルを初期モデルとして用いて学習精度の向上を図る場合などは、このオプションで0を指定して下さい。これによって、モデルパラメータの初期化を行なうことなく、“-if”で指定したモデルのパラメータを直接初期値として使用することができるようになります。また、半連続HMMのように、複数の状態が1つのガウス分布を共有しているような構造のモデルを再学習する場合、本プログラムのパラメータ初期化機能は正常に働きません。従って、この場合にもこのオプションで0を指定して下さい。(このようなモデルに対しては、他の何らかの手法によって、事前にモデルパラメータの初期値を作成しておくことが必要になります。)

- -vb (デフォルト: 0)

Exe.train_HMnetと同じ。

このプログラムによって最終的なHMnetが完成します。また、“-of”で指定したファイル名に“.log”という拡張子を付けた名前のログファイルが、このモデルファイルと同じディレクトリに書き込まれます。

5.4 HMnetの環境補充: Exe.fill_HMnet

このプログラムは、HMnetで定義されていない環境(環境ホール、あるいは環境ギャップ)を補充するためのものです。これによって、環境の直積空間内の全ての環境に対応する状態経路がHMnet上に定義されます。

実行は、

Exe.fill_HMnet [オプション]

によって行ないます。

オプションとしては、以下のものが許されています。

- -cm (省略可能)

Exe.train_HMnetと同じ。

- -if (デフォルト: “in_file”)

環境を補充したいHMnetのファイル名を指定します。これは、(Exe.train_HMnet)で作

成された HMnet を基にしたものであれば、構造抽出や再学習、話者適応などの処理を行なった後のものであってもかまいません。

- -lf (デフォルト: "log_file")
"-if" オプションで指定した HMnet の元となっている HMnet を Exe.train_HMnet によって作成した時に得られているログファイル名を指定します。このログファイルが既に失われている場合には、未定義環境の補充はできません。
- -of (デフォルト: "out_file")
未定義環境補充後の HMnet を書き込むファイル名を指定します。
- -vb (デフォルト: 0)
Exe.train_HMnet と同じ。

このプログラムでは、各状態への環境要素の追加と状態間の経路の追加によって未定義環境が補充されます。そのため、新たな状態の追加や、モデルパラメータの変更は行なわれません。

5.5 HMnet の初期構造生成 : Exe.init_HMM

このプログラムは、HMnet の初期構造を生成するためのものです。これによって、仕様ファイルによって指定された構造を持つ HMnet(topology) が生成されます。

実行は、

Exe.init_HMM [オプション]

によって行ないます。

オプションとしては、以下のものが許されています。

- -cm (省略可能)
Exe.train_HMnet と同じ。
- -di (デフォルト: 34)
Exe.train_HMnet と同じ。
- -fn (デフォルト: 3)
環境の要因数を指定します。このプログラムに与える仕様ファイルに依存します。
- -of (デフォルト: "out_file")
生成された HMnet の構造を書き込むファイル名を指定します。
- -sf (デフォルト: "spec_file")
モデル構造の仕様を記述したファイル名を指定します。

ここで生成される HMnet(topology) は、Exe.topology_HMnet で抽出される HMnet(topology) の構造と同様に、まだパラメータ推定が行なわれていません。そのため、この HMnet(topology) のパラメータにはダミーの数値が入れられていますので、Exe.retrain_HMnet による再推定を終えるまで実際に使用することはできません。

以下に仕様ファイルのフォーマットを示します。

```
#tied_element
MHT

#element 0
a i u e o

#element 1
a i u e o

#element 2
a i u e o

#spec
a,i a *
3 2 3

#spec
!a,i a *
2 2 2 3

.....
.....
```

#tied_element と #element という、2種類のキーワードは省略できません。#tied_element としては、例えば話者名、#element 0, #element n には、各環境要因の全要素を記述して下さい。環境要因数 (n の値) は、オプション “-fn” で指定する値に依存します。

#spec というキーワード1つで、1つのモデル (状態経路) を指定することができます。

次の行には環境要因数だけの環境要素を、間に空白を挟みながら列記します。各要因内で複数の要素を指定する時には “,” で区切って下さい。また、特殊な記号として、“!” はそれ以降で指定されている要素以外のものを表すものです。実際には、#element で指定された全要素から、ここで指定されている要素を取り除いたものとなります。また、“*” は、全要素を表すものです。#element で指定された全要素がそのまま指定されます。

その次の行で、各状態の混合数を指定します。間に空白を挟みながら、状態数分だけ列記します。

例えば、上の例で、

```
#spec
!a,i a *
2 2 2 3
```

と記述されている部分は、第1環境要因の要素が u, e, o のいずれか、第2環境要因の要素が a、第3環境要因の要素が a, i, u, e, o のいずれかであるような音声サンプルに対応する4状態の

モデルを、第1状態の混合数を2、第2状態の混合数を2、第3状態の混合数を2、第4状態の混合数を3として形成することを意味しています。

なお、このプログラムでは、状態の共有化については指定することができません。全ての状態は独立なものとして生成されます。

5.6 混合連続分布 HMM の作成 : Exe.train_HMM

このプログラムは、通常の混合連続分布 HMM を作成するためのものです。

実行は、

Exe.train_HMM [オプション] < “音声サンプルファイル”

によって行ないます。

このプログラムも、Exe.train_HMnet と同様に、音声サンプルファイルが孤立サンプルを表現したものと見なせる場合には切り出し学習モードとなり、それ以外の場合には連結学習モードに入ります。

本プログラムに与えた音声サンプルが孤立サンプルである場合には、与えられた音声サンプルを全て使用して一つの HMM を学習します。そのため、特定のカテゴリの HMM だけを学習するためには、与える音声サンプルに含まれるカテゴリが学習したい HMM のカテゴリのものだけとなるよう、予め分別しておく必要があります。

オプションとしては、以下のものが許されています。

- -cm (省略可能)
Exe.train_HMnet と同じ。
- -di (デフォルト: 34)
Exe.train_HMnet と同じ。
- -dm (デフォルト: NULL)
Exe.train_HMnet と同じ。
- -fn (デフォルト: 3)
環境の要因数を指定します。このプログラムに与える音声サンプルに依存します。ただし、このタイプのモデルでは、環境情報は利用されません。ここでの指定は、音声サンプルを読み込む時にのみ使用されます。
- -mf (デフォルト: 1)
Exe.train_HMnet と同じ。

- -mn (デフォルト: 1)
Exe.topology_HMnet と同じ。
- -of (デフォルト: "out_file")
学習された HMM を書き込むファイル名を指定します。
- -pl (デフォルト: 4)
HMM の状態数を指定します。
- -vb (デフォルト: 0)
Exe.train_HMnet と同じ。

このプログラムによって HMM が学習されます。“-of” で指定したファイル名に “.log” という拡張子を付けた名前のログファイルが、このモデルファイルと同じディレクトリに書き込まれます。

5.7 HMnet のファイル合成 : Exe.compose_HMnet

このプログラムは、複数のモデルファイルを一つの大きなモデルに合成するためのものです。Exe.train_HMM によって作成される HMM はカテゴリ毎に独立したモデルとなっています。そこでこのプログラムによって全てのカテゴリの HMM を一つの大きなファイルとしてまとめることにより、HMnet と全く同様の扱いをすることが可能となります。さらに、未定義環境の補充を行っていない HMnet と HMM を合成し、HMnet 上で特定の環境に対応する経路が見つからない場合の処理を通常の HMM に自動的に切替えるようにすることも可能です。

実行は、

Exe.compose_HMnet [オプション]

によって行ないます。

オプションとしては、以下のものが許されています。

- -cm (省略可能)
Exe.train_HMnet と同じ。
- -il (デフォルト: "list")
合成するモデルファイルのリストを指定します。このリストは、合成の対象となるファイル名を、間に改行を挟みながら並べたものです。
- -of (デフォルト: "out_file")
合成後のモデルを書き込むファイル名を指定します。

このプログラムによって、複数のファイルが一つのモデルとして合成されます。なお、このモデルを使用する際の経路の探索は、“-il”で指定されるリストに記述されたモデルの順で行なわれます。そのため、HMnet と HMM を合成して、未定義環境の補充を行っていない HMM を HMnet で経路が見つからなかった場合の控えとして使用したい場合には、必ず HMnet のファイル名が HMM のファイル名よりも先に記述されたリストを作成して下さい。

5.8 HMnet の話者適応 : Exe.adapt_HMnet

このプログラムは、移動ベクトル場平滑化法 (Vector Field Smoothing: VFS) によって、少数の音声サンプルを用いた HMnet の話者適応を行なうためのものです。

実行は、

Exe.adapt_HMnet [オプション] < “音声サンプルファイル”

によって行ないます。

オプションとしては、以下のものが許されています。

- -cm (省略可能)
Exe.train_HMnet と同じ。
- -di (デフォルト: 34)
Exe.train_HMnet と同じ。
- -dm (デフォルト: NULL)
Exe.train_HMnet と同じ。
- -fn (デフォルト: 3)
Exe.train_HMnet と同じ。
- -if (デフォルト: “in_file”)
話者適応したいモデルファイル名を指定します。
- -gl (デフォルト: NULL)

“-if”で指定したモデルが、未定義環境の補充を行っていない HMnet である場合、適応用単語サンプル中の特定の環境に対応する経路が存在しない場合があります。そのような場合のガベージモデルとして、通常の HMM を用いることができます。この場合は、このオプションでガベージモデルとして用いる HMM のリストを指定して下さい。なお、ここで指定されるモデルは処理の途中で一時的に用いられるだけで、話者適応後のモデルには影響を与えません。(なお、Exe.fill_HMnet によって未定義環境が補充された HMnet を用いる場合には、このオプションは必要ありません。)

- -kn (デフォルト: 6)
移動ベクトルの平滑化を行なう際に考慮する近傍数を指定します。
- -of (デフォルト: "out_file")
話者適応後のモデルを書き込むファイル名を指定します。
- -sr (デフォルト: 20.0)
移動ベクトルの平滑化を行なう際の平滑化パラメータです。正の値を指定して下さい。この値が大きい程強い平滑化が行なわれます。適応に使用するの音声サンプル数に応じて適当に決めて下さい。
- -tm (デフォルト: 1)
学習モードを示す値 (10 進数) を指定します。この内容は以下の通りです。
 - 001 平均値ベクトルを移動ベクトル場平滑化法により適応する。
 - 002 分散を Baum-Welch 法により適応する。(ただし、少サンプルによる適応を基本とするため、適応前より分散が大きくなる場合に限り分散を変更する。)
 - 004 混合成分の重みを Baum-Welch 法により適応する。(ただし、適応対象モデルが混合ガウス分布モデルである場合のみ有効。)
 - 008 状態遷移確率を Baum-Welch 法により適応する。
 なお、複数のモデルパラメータを同時に適応したい場合には、各値の和を指定して下さい。(例: 平均と分散を適応したい場合 → "-tm 3")
- -vb (デフォルト: 0)
Exe.train_HMnet と同じ。

このプログラムによって話者適応が実施されます。“-of”で指定したファイル名に“.log”という拡張子を付けた名前のログファイルが、このモデルファイルと同じディレクトリに書き込まれます。

5.9 HMnet の構造解析 : Exe.analysis_HMnet

このプログラムは、HMnet の構造を解析するためのものです。

実行は、

Exe.analysis_HMnet [オプション]

によって行ないます。

オプションとしては、以下のものが許されています。

- -if (デフォルト: "in_file")

構造を解析したいモデルファイル名を指定します。

このプログラムによってモデルの構造が解析されます。結果は標準出力に書き込まれます。これによって、次のような結果が得られます。

```

;; Wed Feb  9 01:51:53 1994
;;
;; <<< Analysis HMnet >>>
;;
;; Current Working Directory      : /tmp_mnt/home/atrh12/jun/SSS/Exe.sample
;;
;; [HMnet]
;; Input Model File Name         : Test.HP-UX/Fill_HMnet/HMnet.600
;;
# Actual State Number : 600
# Allophone Model Number : 1816
# Accumulative State Number : 7224
# Modeling Efficiency : 12.040000

```

これらの結果は、上から順に、• 解析したモデルファイル名、• モデル中で使用されている全状態数、• このモデルで表現されている異なり経路数、• 全モデルで使用されている延べ状態数、• [延べ状態数]/[実際の状態数] で計算されるモデル化効率、をそれぞれ表しています。

5.10 HMnet の出力尤度評価 : Exe.evaluate_HMnet

このプログラムは、与えられた音声サンプルに対する出力尤度を計算するためのものです。実行は、

Exe.evaluate_HMnet [オプション] < “音声サンプルファイル”

によって行ないます。

オプションとしては、以下のものが許されています。

- -di (デフォルト: 34)
Exe.train_HMnet と同じ。
- -dm (デフォルト: NULL)
Exe.train_HMnet と同じ。
- -fn (デフォルト: 3)
Exe.train_HMnet と同じ。
- -if (デフォルト: "in_file")
出力尤度計算に使用するモデルファイル名を指定します。

これによって、以下のような結果が標準出力より出力されます。

```

;; Wed Feb  9 01:52:03 1994
;;
;; <<< Evaluate HMnet >>>
;;
;; Current Working Directory      : /tmp_mnt/home/atrh12/jun/SSS/Exe.sample
;;
;; [HMnet]
;; Input Model File Name         : Test.HP-UX/Adapt_HMnet/HMnet.600
;;
;; [Data]
;; Parameter Dimension           : 34
;; Total Data Number             : 25
;; Total Sample Number           : 226
;; Total Frame Length            : 4468
;; Min. Frame Length             : 132
;;
;; # Cut Data Number : 0
;; # Total Probability : 2.286254e+05

```

この中で、“Cut Data Number”とは、評価しようとした音声サンプルの中で、その環境を表現する経路が見つからない、あるいはその経路を構成している状態数よりも音声サンプルのフレーム数の方が短くて評価することができない、などの理由により、無視されたサンプルの数を表しています。従って、実際の評価に使用されたサンプル数は、“Total Data Number”で表示されている全入力音声サンプル数からこの“Cut Data Number”の値を引いたものとなります。

5.11 HMnet の状態経路探索 : Exe.path_HMnet

このプログラムは、与えられたラベル系列に対応する HMnet 上の状態経路を探索するためのものです。

実行は、

Exe.path_HMnet [オプション] < “ラベル系列ファイル”

によって行ないます。

オプションとしては、以下のものが許されています。

- -fn (デフォルト: 3)
Exe.train_HMnet と同じ。
- -if (デフォルト: “in_file”)
状態経路探索に使用するモデルファイル名を指定します。

ここで使用する“ラベル系列ファイル”のフォーマットは、3章で示した音声サンプルのフォーマットの中から、状態系列の探索に不必要な“共通化環境要因の要素”や“データのフレーム数”、“サンプルデータ”を取り除いたものとなっています。このラベル系列ファイルの例を次に示します。

```

11
--k
--ko
kong
ongn
ngni
nich
ichi
chiw
iwa
wa-
a--

10
--s
--sa
saj
joo
oon
ona
nar
ara
ra-
a--

```

これによって、以下のような結果が標準出力より出力されます。

```

;; Wed Feb  9 01:52:06 1994
;;
;; <<< Find Path >>>
;;
;; Current Working Directory      : /tmp_mnt/home/atrh12/jun/SSS/Exe.sample
;;
;; [HMnet]
;; Input Model File Name          : Test.HP-UX/Fill_HMnet/HMnet.600
;;
# Data 0
- : 418 323 370 158
k : 286 312 214 464
o : 483 503 226 281
ng : 285 240 552 35
n : 25 40 336
i : 338 246 362 123
ch : 333 51 317 6
i : 396 390 439 306
w : 160 265 294 9
a : 233 223 84 541
- : 567 533 22 106

# Data 1
- : 256 104 32 259
s : 595 413 138 402
a : 156 318 352 105
j : 66 340 64 348
o : 437 475 499 120
n : 511 40 33
a : 485 280 42 500
r : 549 150 555 539
a : 272 14 84 142
- : 567 533 22 106

```

5.12 HMnet の展開 : Exe.expand_HMnet

このプログラムは、HMnet を全展開し、各モデルがどのような環境を表現し、どのような状態系列で構成されているかを調べるためのものです。

実行は、

Exe.expand_HMnet [オプション]

によって行ないます。

オプションとしては、以下のものが許されています。

- -if (デフォルト: "in_file")

展開したいモデルファイル名を指定します。

これによって、以下のような結果が標準出力より出力されます。

```

;; Wed Feb  9 01:52:20 1994
;;
;; <<< Expand HMnet >>>
;;
;; Current Working Directory   : /tmp_mnt/home/atrh12/jun/SSS/Exe.sample
;;
;; [HMnet]
;; Input Model File Name      : Test.HP-UX/Fill_HMnet/HMnet.600
;;
# Allophone 0
-
-
ts
1 79 364 583

# Allophone 1
- a i k j o zh z u d m g ch ng r ts s e b q t sh w n p h
p
a i - k j o zh z u d m g ch ng r sh e s b q t w n p ts h
5 81 76 110

# Allophone 2
r
u
k z s
7 196 291 41

# Allophone 3
r
u
zh g
7 196 579 36

(以下、省略)

```

5.13 HMnet の状態継続時間算出 : Exe.viterbi_HMnet

このプログラムは、入力音声に対して Viterbi 経路を計算し、各状態での継続時間を求めるためのものです。

実行は、

Exe.viterbi_HMnet [オプション] < “音声サンプルファイル”

によって行ないます。

オプションとしては、以下のものが許されています。

- -di (デフォルト: 34)
Exe.train_HMnet と同じ。
- -dm (デフォルト: NULL)
Exe.train_HMnet と同じ。
- -fn (デフォルト: 3)
Exe.train_HMnet と同じ。
- -if (デフォルト: "in_file")
状態継続時間算出の対象となるモデルファイル名を指定します。

これによって、以下のような結果が標準出力より出力されます。

```
;; Wed Feb  9 01:52:11 1994
;;
;; <<< Viterbi HMnet >>>
;;
;; Current Working Directory      : /tmp_mnt/home/atrh12/jun/SSS/Exe.sample
;;
;; [HMnet]
;; Input Model File Name          : Test.HP-UX/Adapt_HMnet/HMnet.600
;;
;; [Data]
;; Parameter Dimension            : 34
;; Total Data Number              : 25
;; Total Sample Number           : 226
;; Total Frame Length            : 4468
;; Min. Frame Length             : 132
;;
# Data 0
- : 23 ( 125 : 10 ) ( 200 : 6 ) ( 185 : 2 ) ( 91 : 5 )
i : 12 ( 65 : 3 ) ( 151 : 1 ) ( 362 : 1 ) ( 123 : 7 )
k : 23 ( 455 : 3 ) ( 258 : 6 ) ( 128 : 4 ) ( 12 : 10 )
o : 20 ( 576 : 6 ) ( 151 : 1 ) ( 564 : 1 ) ( 572 : 12 )
o : 25 ( 227 : 6 ) ( 115 : 10 ) ( 347 : 6 ) ( 121 : 3 )
i : 27 ( 349 : 6 ) ( 246 : 5 ) ( 363 : 6 ) ( 195 : 10 )
- : 28 ( 212 : 3 ) ( 47 : 7 ) ( 17 : 4 ) ( 260 : 14 )

# Data 1
- : 26 ( 125 : 16 ) ( 200 : 1 ) ( 185 : 3 ) ( 91 : 6 )
i : 20 ( 65 : 2 ) ( 151 : 11 ) ( 82 : 6 ) ( 306 : 1 )
j : 7 ( 66 : 1 ) ( 270 : 1 ) ( 64 : 1 ) ( 348 : 4 )
o : 15 ( 227 : 3 ) ( 475 : 4 ) ( 347 : 5 ) ( 121 : 3 )
i : 30 ( 349 : 8 ) ( 246 : 17 ) ( 82 : 4 ) ( 306 : 1 )
j : 8 ( 66 : 1 ) ( 270 : 1 ) ( 64 : 2 ) ( 348 : 4 )
o : 24 ( 227 : 5 ) ( 393 : 5 ) ( 399 : 7 ) ( 522 : 7 )
- : 20 ( 330 : 1 ) ( 331 : 12 ) ( 172 : 6 ) ( 171 : 1 )
```

(以下、省略)

この結果の見方は、例えば、“- : 23 (126 : 15) (212 : 1) (186 : 3) (93 : 4)”とある場合には、音素 /-/ の長さは全体で 23 フレームであり、その環境に対応するモデルの第 1 状態である状態 126 には 15 フレーム、第 2 状態である状態 212 には 1 フレーム、第 3 状態である状態 186 には 3 フレーム、第 4 状態である状態 93 には 4 フレーム、それぞれ滞在したことを表しています。

5.14 HMnet のカテゴリ識別性能評価 : Exe.recognize_HMnet

このプログラムは、モデルのカテゴリ識別性能を評価するためのものです。

実行は、

Exe.recognize_HMnet [オプション] < “音声サンプルファイル (孤立サンプル)”

によって行ないます。

このプログラムは、孤立した音声サンプルに対するカテゴリ識別性能を評価するためのものです。従って、与える音声サンプルとしては、孤立サンプル (音声サンプルファイル中の全てのブロックで“セグメント数”が1となっているもの) を使用して下さい。

オプションとしては、以下のものが許されています。

- -an (デフォルト: “???”)
正解カテゴリを指定します。このプログラムに与える音声サンプルに依存します。
- -cn (デフォルト: 5)
累積認識率を計算する範囲を指定します。
- -di (デフォルト: 34)
Exe.train_HMnet と同じ。
- -dm (デフォルト: NULL)
Exe.train_HMnet と同じ。
- -fn (デフォルト: 3)
Exe.train_HMnet と同じ。
- -if (デフォルト: “in_file”)
カテゴリ識別実験に使用するモデルファイル名を指定します。
- -mf (デフォルト: 1)
Exe.train_HMnet と同じ。
- -ob (デフォルト: “object”)
識別の対象となるカテゴリ名のリストを指定します。このリストには、識別の対象となるカテゴリ名を空白を入れながら列記して下さい。
- -re (デフォルト: “result”)
カテゴリ識別実験結果を書き込むファイル名を指定します。

このプログラムによって累積識別性能や、他のカテゴリとの間のコンフュージョンマトリクスが得られます。

これによって、以下のような結果ファイルが作られます。

```

;; Wed Feb  9 01:49:21 1994
;;
;; <<< Recognize HMnet >>>
;;
;; Current Working Directory      : /tmp_mnt/home/atrh12/jun/SSS/Exe.sample
;;
;; [HMnet]
;; Input Model File Name          : Test.HP-UX/Adapt_HMnet/HMnet.600
;;
;; [Data]
;; Parameter Dimension            : 34
;; Total Data Number              : 232
;; Total Sample Number           : 232
;; Total Frame Length            : 2838
;; Min. Frame Length             : 4
;;
;; [Condition]
;; Answer                         : b
;; Candidate Number               : 5
;; Object Category                : b d g p t k m n ng s sh h z ch ts zh r w j a i u e o q -
;;
# Data Number
232

# Performance
189 211 222 227 229

# Confusion Matrix
189 16 1 1 0 0 2 3 0 0 0 0 0 0 0 0 0 0 19 1 0 0 0 0 0 0 0 0

```

これらの結果は、上から順に、● 評価に使用された音声サンプル数、● 左から順に、第 n 位までに正解が含まれていたサンプルの数、● 各カテゴリ (順序は、“-ob” オプションで指定したリストに書かれているカテゴリの順で並んでおり、これはヘッダ部の “Object Category” の内容で確認することができます) に対するコンフュージョンマトリクス、をそれぞれ表しています。

この例の場合、音素 /b/ に対する第一位正解率は $185/232 = 0.7974$ (79.74%)、また第三位までの累積正解率は $224/232 = 0.9655$ (96.55%) のようにして求めることができます。

またカテゴリ識別実験中は、標準出力に識別状況を監視するための表示が出力されます。

第 6 章

SSS-ToolKit.ver3 の動作確認

SSS-ToolKit.ver3 には、実際の動作を確認するためのサンプルデータと、事前に得られている幾つかのサンプルモデル、および本章で示されている一連の動作確認用コマンドを実際に実行して得られたリファレンスが付録として付けられています。これらを用いて各プログラムの動作チェックを行なうことができます。ここでは、その方法について説明します。

動作確認の実行は、ディレクトリ “SSS/Exe.sample” の下で行ないます。

本章で示されている一連の動作確認用コマンドで作成されるモデルファイルやログファイル、実行結果などは、全て “SSS/Exe.sample/test” (ただし、環境変数 “\$OS” が設定されている場合には、“SSS/Exe.sample/TEST.\$OS”) の下に書き込まれるようにしてあります。このディレクトリの内容は、これと同じ操作を事前に行なって作成した “SSS/Exe.sample/Test.reference” に保存されている内容と一致、もしくは近い値となるはずですが (浮動小数点演算の結果が計算機によって微妙に異なる場合があり、全く同じ結果にはならない場合もあります)。各項目について動作を確認する度に、そこで作られるファイルの内容と比較してみてください。大きな違いが見られないようなら動作は正常です。なお、“SSS/Exe.sample/Test.reference” の下の各モデルや、本章で示されている実行結果の例は、全て HP9000/755 上で、オペレーティングシステムに HP-UX ver 9.01、コンパイラに gcc ver 2.4.5 を使用して得られたものとなっています。

動作確認のために使用する音声サンプルデータは、一部にバイナリ形式の浮動小数表現を使用しているため、big endian (HP、Sun などはこちらのタイプ) と little endian (DEC などはこちらのタイプ)2 通りを用意しました。動作確認試験を行なう前に、使用する計算機がどちらのタイプの浮動小数表現を使用しているかを確認した上、以下の操作を行なって下さい。

(byte ordering が big endian である場合)

```
ln -s Data.big data
```

(byte ordering が little endian である場合)

```
ln -s Data.little data
```

ただし、環境変数“\$OS”が設定されている場合には、“data”の代わりに“DATA.\$OS”を指定して下さい。

また、各プログラムの実行形式は、環境変数“\$OS”が設定されていなければ“SSS/bin”、“\$OS”が設定されていれば“SSS/BIN.\$OS”の下のものにあるものが使用されます。

以上の説明に多く現れている環境変数“\$OS”は、複数の計算機環境下でのプログラムの管理を簡単化するために使用されています。様々な計算機環境が混在する場合には、各環境毎にユニークな“\$OS”を設定することによって、実行形式や動作確認用サンプルなどを、自動的に使い分けることができるようになります。

本章では、環境変数“\$OS”が設定されていないものとして説明を進めます。“\$OS”が設定されている場合には、“SSS/data”および“SSS/Exe.sample/test”という2つのディレクトリ名を、それぞれ“SSS/DATA.\$OS”および“SSS/Exe.sample/TEST.\$OS”に置き換えて考えて下さい。

なお、これ以降に示す一連の動作確認のためのコマンドは、ディレクトリ“SSS/Exe.sample”の下で“ALL.sh”を実行することにより、全て自動的にこなすこともできます。

6.1 Exe.train_HMnet の動作確認

学習用サンプルとして男性話者一名 (MHT) の /b/, /d/, /g/ の 3 音素 (“SSS/data/data1”) を用いて HMnet を生成します。

“Train_HMnet.sh” を実行して下さい。

“SSS/Exe.sample/test/Train_HMnet” というディレクトリの下に、“HMnet.状態数” や “HMnet.current”、“HMnet.log” などのファイルが作成されます。

6.2 Exe.topology_HMnet の動作確認

6.1で作成した “SSS/Exe.sample/test/Train_HMnet/HMnet.20” から、その構造を取り出し、同時に混合数を 1 に変更します。

“Topology_HMnet.sh” を実行して下さい。

“SSS/Exe.sample/test/Topology_HMnet” というディレクトリの下に、“HMnet.20” というファイルが作成されます。

6.3 Exe.retrain_HMnet の動作確認

6.2で抽出した 20 状態の HMnet の構造に対して、HMnet 生成時に用いたものと同じ学習用サンプルを用いてパラメータを再学習します。

“Retrain_HMnet.sh” を実行して下さい。

“SSS/Exe.sample/test/Retrain_HMnet” というディレクトリの下に、“HMnet.20” というモデルファイルと、“HMnet.20.log” というログファイルが作成されます。

6.4 Exe.fill_HMnet の動作確認

“SSS/Model.sample/” のディレクトリの下に予め用意した、男性話者一名 (MHT) の音声で学習した 600 個の状態を持つ 26 音素用の HMnet (“HMnet.600”) に対して、未定義環境の補充を行いません。

“Fill_HMnet.sh” を実行して下さい。

“SSS/Exe.sample/test/Fill_HMnet” というディレクトリに、“HMnet.600” というモデルファイルと、“HMnet.600.log” というログファイルが作成されます。

6.5 Exe.init_HMM の動作確認

“List/list.init_HMnet” に記された仕様に従って、語頭の /b/ と語中の /b/ の HMM (それぞれ 3 状態、混合数 3) を生成します。

“Init_HMnet.sh” を実行して下さい。

“SSS/Exe.sample/test/Init_HMnet” というディレクトリの下に、“HMM.b” というモデルファイルが作成されます。

6.6 Exe.train_HMM の動作確認

男性話者一名 (MHT) の音素 /b/ のサンプル (“SSS/data/data2”) を用いて混合数 5 の混合連続分布 HMM を生成します。

“Train_HMM.sh” を実行して下さい。

“SSS/Exe.sample/test/Train_HMM” というディレクトリの下に、“HMM.b” というモデルファイルと、“HMM.b.log” というログファイルが作成されます。

6.7 Exe.compose_HMnet の動作確認

6.6 で作成した音素 /b/ 用の HMM と、“SSS/Model.sample” のディレクトリの下に予め用意した、他のカテゴリに対応する 25 通りの HMM とを、1 つの大きなモデルファイルとして合成します。

“Compose_HMnet.sh” を実行して下さい。

“SSS/Exe.sample/test/Compose_HMM” というディレクトリに、“HMM.26phone” というモ

デルファイルが作成されます。

6.8 Exe.adapt_HMnet の動作確認

男性話者一名 (MAU) の 25 単語音声サンプル (“SSS/data/data3”) を用いて、6.4 で作成した未定義環境補充後の 26 音素用 HMnet の話者適応を行ないます。

“Adapt_HMnet.sh” を実行して下さい。

“SSS/Exe.sample/test/Adapt_HMM” というディレクトリの下に、“HMnet.600” というモデルファイルと、“HMnet.600.log” というログファイルが作成されます。

6.9 Exe.analysis_HMnet の動作確認

6.4 で作成した未定義環境補充後の 26 音素用 HMnet に対して、その構造解析を行ないます。

“Analysis_HMnet.sh” を実行して下さい。

“SSS/Exe.sample/test/Analysis_HMM” というディレクトリの下に、“Result” という結果のファイルが作成されます。

6.10 Exe.evaluate_HMnet の動作確認

6.8 で使用した話者 MAU の 25 単語音声サンプル (“SSS/data/data3”) に対して、話者適応により得られたモデルから得られる出力尤度の総和を求めます。

“Evaluate_HMnet.sh” を実行して下さい。

“SSS/Exe.sample/test/Evaluate_HMM” というディレクトリの下に、“Result” という結果のファイルが作成されます。

6.11 Exe.path_HMnet の動作確認

6.4 で作成した未定義環境補充後の 26 音素用 HMnet に対して、状態系列を探索します。

“Path_HMnet.sh” を実行して下さい。

“SSS/Exe.sample/test/Path_HMM” というディレクトリの下に、“Result” という結果のファイルが作成されます。

6.12 Exe.expand_HMnet の動作確認

6.4で作成した未定義環境補充後の26音素用HMnetを展開します。

“Expand_HMnet.sh” を実行して下さい。

“SSS/Exe.sample/test/Expand_HMM” というディレクトリの下に、“Result” という結果のファイルが作成されます。

6.13 Exe.viterbi_HMnet の動作確認

6.8で使用した話者MAUの25単語音声サンプル(“SSS/data/data3”)に対して、話者適応により得られたモデルを適用した場合の状態継続時間を算出します。

“Viterbi_HMnet.sh” を実行して下さい。

“SSS/Exe.sample/test/Viterbi_HMM” というディレクトリの下に、“Result” という結果のファイルが作成されます。

6.14 Exe.recognize_HMnet の動作確認

話者MAUの音素/b/のサンプル(“SSS/data/data4”)に対して、6.8で作成したHMnetのカテゴリ(音素)識別試験を行ないます。

“Recognize_HMnet.sh” を実行して下さい。

“SSS/Exe.sample/test/Recognize_HMM” というディレクトリの下に、“Result.b” という結果のファイルと、“Result.b.log” というログファイルが作成されます。

第 7 章

おわりに

ここで紹介した SSS-ToolKit.ver3 は、SSS および HMnet を中心とした音声認識用音響モデルの研究用ツールとして構築したものです。特にこのバージョンでは、多くの研究員の方々の御意見を参考に、SSS-ToolKit.ver1 や SSS-ToolKit.ver2 に改良を加え、実際の研究に役立つ様々な機能が追加されています。その結果、これまでのものに比べ、プログラムの信頼性や使い易さを大幅に向上させることができました。

本プログラムの開発および改良に当たり、内容に関するアドバイス、移植性を高めるためのプログラム変更、プログラムの管理など、多方面からサポートして頂いた Harald Singer 研究員に深く感謝致します。また、実際に使用する中で生じた様々な問題に関する貴重な御意見を頂いた小坂哲夫研究員を始めとする ATR 音声翻訳通信研究所の皆様感謝いたします。

第 8 章

参考文献

SSS と HMnet の原理やアルゴリズムに関するもの

1. 鷹見 淳一, 嵯峨山 茂樹: “逐次状態分割法 (SSS) による隠れマルコフネットワークの自動生成,” 音講論集, 2-5-13, pp. 73-74 (1991.10).
2. 鷹見 淳一, 嵯峨山 茂樹: “音素コンテキストと時間に関する逐次状態分割による隠れマルコフ網の自動生成,” 信学技報, SP91-88, pp. 57-64 (1991.12).
3. 鷹見 淳一: “音声認識における HMM とその精度向上のための手法,” 信学技報, SP92-49, pp. 17-24 (1992.9).
4. 鷹見 淳一, 小坂 哲夫, 嵯峨山 茂樹: “話者方向を加えた逐次状態分割法 (SSS) による話者共通隠れマルコフ網の生成,” 音講論集, 3-1-8, pp. 155-156 (1992.10).
5. 鷹見 淳一, 永井 明人, 嵯峨山 茂樹: “「aka(赤) と aki(秋) の /k/ は同じ音 ?」 — 前後の音素を考慮した高精度音声認識 —,” ATR Journal, No. 12, pp. 12-17 (1992.11).
6. 鷹見 淳一, Harald Singer, 松村 壮史, 嵯峨山 茂樹: “連結学習を可能とした逐次状態分割法による隠れマルコフ網の音節単位学習,” 音講論集, 3-8-6, pp. 135-136 (1993.10).
7. 鷹見 淳一, 嵯峨山 茂樹: “逐次状態分割法による隠れマルコフ網の自動生成,” 信学論, Vol. J76-D-II, No. 10, pp. 2155-2164 (1993.10).

VFS の原理やアルゴリズムに関するもの

1. 服部 浩明, 嵯峨山 茂樹: “少量学習データを用いたコードブックマッピングによる話者適応化,” 音講論集, 1-5-23, pp. 49-50 (1991.3).
2. 鷹見 淳一, 嵯峨山 茂樹: “隠れマルコフ網 (HM-Net) を用いた話者適応,” 音講論集, 1-1-8, pp. 15-16 (1992.3).

3. 大倉 計美, 杉山 雅英, 嵯峨山 茂樹: “混合連続分布 HMM を用いた移動ベクトル場平滑化話者適応方式,” 信学技報, SP92-16, pp. 23-28 (1992.6).
4. Jun-ichi Takami, Akito Nagai, Shigeki Sagayama: “Speaker Adaptation of the SSS (Successive State Splitting)-Based Hidden Markov Network for Continuous Speech Recognition,” Proc. of SST'92 (Australia), pp. 437-442 (1992.12).
5. 鷹見 淳一, 嵯峨山 茂樹: “隠れマルコフ網のための話者適応法,” 信学技報, SP93-50, pp. 9-16 (1993.8).

HMnet や VFS の応用に関するもの

1. 鷹見 淳一, 嵯峨山 茂樹: “逐次状態分割法 (SSS) により自動生成した隠れマルコフ網の性能評価,” 音講論集, 2-1-8, pp. 65-66 (1992.3).
2. Jun-ichi Takami, Shigeki Sagayama: “A Successive State Splitting Algorithm for Efficient Allophone Modeling,” Proc. of ICASSP'92, 66.6 (San Francisco) (1992.3).
3. 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 (1992.6).
4. 小坂 哲夫, 鷹見 淳一, 嵯峨山 茂樹: “話者混合 SSS による不特定話者音声認識と話者適応,” 信学技報, SP92-52, pp. 17-24 (1992.9).
5. 鷹見 淳一, 永井 明人, 嵯峨山 茂樹: “逐次状態分割法 (SSS) と LR パーザを統合した SSS-LR 連続音声認識手法における話者適応の性能評価,” 音講論集, 2-5-5, pp. 127-128 (1992.10).
6. 永井 明人, 鷹見 淳一, 嵯峨山 茂樹: “逐次状態分割法 (SSS) と LR パーザを統合した SSS-LR 連続音声認識手法の文音声による評価,” 音講論集, 2-1-8 (1992.10).
7. A. Nagai, J. Takami, S. Sagayama: “The SSS-LR Continuous Speech Recognition System: Integrating SSS-derived Allophone Models and a Phoneme-Context-Dependent LR Parser,” ICSLP'92 (Canada) (1992.10).
8. 小坂 哲夫, 鷹見 淳一, 嵯峨山 茂樹: “話者混合 SSS による不特定話者音声認識,” 音講論集, 2-5-9, pp. 135-136 (1992.10).
9. Akito Nagai, Jun-ichi Takami, Shigeki Sagayama: “The SSS-LR Continuous Speech Recognition System: Integrating SSS-derived Allophone Models and a Phoneme-Context-Dependent LR Parser,” Proc. of ICSLP'92 (Canada) (1992.10).
10. 鷹見 淳一, 宮沢 康永, 永井 明人, 嵯峨山 茂樹: “話者適応型 SSS-LR 連続音声認識方式における標準話者予備選択の効果,” 音講論集, 2-Q-16, pp. 115-116 (1993.3).

11. 鷹見 淳一, 宮沢 康永, 永井 明人, 嵯峨山 茂樹: “話者適応型 SSS-LR 連続音声認識方式における標準話者予備選択の効果,” 音講論集, 2-Q-16, pp. 115-116 (1993.3).
12. 小坂 哲夫, Edward Willems, 鷹見 淳一, 嵯峨山 茂樹: “複数の話者適応法に基づく動的話者適応,” 音講論集, 2-4-9, pp. 35-36 (1993.3).
13. 嵯峨山 茂樹, 鷹見 淳一, 永井 明人, Harald Singer, 谷戸 文廣, 鈴木 雅実, 森元 暉, 樽松 明: “自動翻訳電話実験システム ASURA の概要,” 音講論集, 3-4-17, pp. 83-84 (1993.3).
14. Harald Singer, Jun-ichi Takami, Shoichi Matsunaga: “Mora Duration Models for SSS-LR Continuous Speech Recognition,” 音講論集, 2-7-3, pp. 75-76 (1993.10).
15. 小坂 哲夫, 松永 昭一, 嵯峨山 茂樹: “木構造話者クラスタリングを用いた話者適応,” 音講論集, 2-7-14, pp. 97-98 (1993.10).
16. 宮沢 康永, 鷹見 淳一, 松永 昭一, 嵯峨山 茂樹: “全音素エルゴディック隠れマルコフ網 (HM-net) による教師なし話者適応,” 音講論集, 2-Q-27, pp. 201-202 (1993.10).
17. 大脇 浩, Harald Singer, 鷹見 淳一, 樽松 明: “音素配列構造の制約を用いた音素タイプライタ,” 信学技報, SP93-113, pp. 71-78 (1993.12).
18. 永井明人, 鷹見淳一, 嵯峨山茂樹, Harald Singer: “隠れマルコフ網と一般化 LR 構文解析を統合した連続音声認識,” 信学論, Vol. J77-D-II, No.1, pp. 9-19 (1994.1).
19. 宮沢 康永, 鷹見 淳一, 嵯峨山 茂樹, 松永 昭一: “全音素エルゴディック隠れマルコフ網を用いた教師なし話者適応方式,” 信学技報, SP93-134, pp. 47-52 (1994.1).
20. 鷹見 淳一, Harald Singer, 大脇 浩: “語彙や構文に制約のない音声認識手法における音響モデルの性能評価,” 音講論集, 2-P-21 (1994.3).

付録 A

SSS-ToolKit.ver3 で使用されるモデルファイルの例

話者 MHT の音素サンプルより生成した 3 音素 (/b/, /d/, /g/) 用の HMnet (状態数: 5)

```

***
* <<< SSS-ToolKit Version 3.1 >>>
* Copyright(C) 1993 Jun-ichi TAKAMI
* ATR Interpreting Telecommunications Research Laboratory
* Department 1
* 2-2 Hikaridai Seika-cho Soraku-gun Kyoto 619-02
* Tel : 07749-5-1301
* (Direct): 07749-5-1379
* Fax : 07749-5-1308
* E-mail : jun@itl.atr.co.jp
*****
Wed Feb 9 01:29:54 1994
<<< Generate HMnet by SSS Algorithm >>>
Current Working Directory : /tmp_mnt/home/atrh12/jun/SSS/HP-UX/Train_HMnet/HMnet
[HMnet]
Algorithm Type : SSS
Max. State Number : 20
Initial Path Length : 1
Max. Path Length : 4
Factor Splitting Priority : 1/0,1/2
Tied Element : MHT
Factor Number : 3
Main Element : b d g
Output Model File Name : Test_HP-UX/Train_HMnet/HMnet
Model Save Step : 5

[Data]
Parameter Dimension : 34
Total Data Number : 718
Total Sample Number : 718
Total Frame Length : 9971
Min. Frame Length : 4

=====
#type [ model type ]
#covariance [ covariance type ]
#dimension [ param dimension ]
#max_mix [ max mixture num ]
#total_dist [ total dist num ]
#total_outp [ total outp num ]
#type HMnet
#covariance diag
#dimension 34
#max_mix 2
#total_dist 10
#total_outp 5
; ===== Distribution =====
; #dist [ dist index ]
; [ mean ] [ var ]
; .....
#dist 0
7.351477e+00 1.635113e+00
8.709206e-01 2.072252e-01
5.202673e-01 9.565680e-02
5.410694e-01 4.780594e-02
3.845737e-01 5.164911e-02
2.034372e-01 2.271131e-02

#dist 1
6.061195e+00 2.427879e+00
6.486789e-01 1.260009e-01
4.155360e-01 4.038290e-02
4.380012e-01 2.431800e-02
5.330097e-01 4.537463e-02
2.923742e-01 1.225628e-02
2.197266e-01 3.533159e-02
-1.242983e-01 5.398082e-02
2.074811e-01 1.736759e-02
2.333374e-01 1.324016e-02
1.021140e-01 1.068734e-02
-6.621758e-02 2.665641e-02
3.752727e-02 6.232005e-03
2.941346e-03 1.483902e-02
-1.121153e-02 8.312015e-03
-6.687385e-02 5.546011e-03
-5.108694e-02 4.420158e-03
5.989290e-01 1.344843e-01
1.557692e-01 6.154190e-03
4.291582e-02 1.922962e-03
3.970122e-02 8.889971e-04
3.612496e-02 1.961009e-03
1.207308e-02 5.467149e-04
-1.684773e-02 1.151794e-03
-2.261376e-02 1.929727e-03
-5.055969e-03 5.928715e-04
1.792730e-02 5.563539e-04
-1.729253e-03 3.788859e-04
-9.036450e-03 7.166087e-04
-4.446399e-03 2.952902e-04
-2.465942e-03 3.634913e-04
3.649782e-04 3.319944e-04
-6.726234e-03 2.394008e-04
-9.995306e-03 1.632405e-04

#dist 2
1.166466e+01 7.047920e-01
8.387966e-01 6.291148e-02
5.399318e-03 5.166297e-02
4.541014e-02 4.025535e-02

```



```

6.845395e-01 3.857015e-02 -1.794919e-03 1.816851e-04
8.184279e-02 3.645373e-02 8.521314e-04 1.410199e-04
3.826040e-01 4.283068e-02 3.283495e-03 7.890359e-05
-2.900319e-01 2.283608e-02
-1.529119e-01 1.791394e-02
3.236136e-01 1.501296e-02
9.077743e-02 1.345348e-02
-1.807360e-01 1.111971e-02
-7.423920e-02 9.173731e-03
-1.078214e-02 6.708212e-03
-8.305237e-02 4.968713e-03
-8.048127e-03 5.297367e-03
-1.014063e-01 4.449473e-03
5.276691e-02 1.395087e-02
-5.281225e-03 1.687719e-03
-1.805326e-02 1.262393e-03
1.320433e-03 1.084450e-03
-1.323096e-02 1.389178e-03
-1.232074e-02 9.722441e-04
-6.378938e-03 7.440109e-04
8.933043e-03 7.062598e-04
-3.217657e-03 4.643956e-04
-7.393678e-03 5.202959e-04
-3.485418e-03 3.690013e-04
2.117177e-03 2.932473e-04
-4.552748e-03 2.337412e-04
2.194063e-05 2.107995e-04
-3.689916e-04 1.262408e-04
-5.832823e-04 1.357658e-04
2.324758e-03 1.316813e-04

#dist 3
1.149685e+01 8.513393e-01
7.674430e-01 1.116313e-01
-5.553988e-03 5.566364e-02
1.597209e-01 7.059818e-02
6.032628e-01 4.294956e-02
7.174430e-02 4.790799e-02
4.071987e-01 4.775403e-02
-2.989329e-01 5.683910e-02
-1.463987e-01 2.910618e-02
2.616293e-01 3.278361e-02
9.309446e-02 2.226278e-02
-1.446232e-01 1.669071e-02
-6.522981e-02 1.229691e-02
-3.964318e-02 1.032491e-02
-5.377217e-02 8.038353e-03
-7.089017e-03 5.524332e-03
-9.071655e-02 7.127051e-03
-1.116046e-01 1.892543e-02
2.032801e-02 3.537423e-03
1.803672e-02 2.303411e-03
-1.290943e-02 1.990839e-03
1.813720e-02 1.980924e-03
1.598937e-02 1.930628e-03
2.085234e-02 1.116140e-03
-1.663138e-02 1.370023e-03
7.010854e-03 8.609190e-04
1.529064e-02 9.046204e-04
5.309394e-03 5.197566e-04
-1.593359e-03 4.654228e-04
5.314980e-03 4.00628e-04
-4.593864e-04 3.824808e-04
-5.578128e-03 2.383069e-04
1.102558e-03 1.917268e-04
-4.695630e-03 2.559660e-04

#dist 4
7.441997e+00 8.862970e-01
1.147194e+00 6.259048e-02
6.634336e-01 4.417102e-02
6.292877e-01 2.261353e-02
5.713712e-01 3.556051e-02
2.449414e-01 1.998849e-02
-6.404883e-03 1.921649e-02
-1.184464e-01 2.842882e-02
1.286417e-02 1.126676e-02
1.426051e-01 1.621042e-02
1.257090e-01 1.378800e-02
8.084253e-02 1.976957e-02
5.053393e-02 6.420058e-03
1.542194e-02 1.576535e-02
4.237513e-02 1.306571e-02
3.778759e-03 7.209075e-03
-6.365490e-02 4.829637e-03
-3.259869e-02 1.665177e-02
-2.474678e-02 2.082023e-03
1.244969e-02 2.388372e-03
1.284387e-03 1.026414e-03
-7.813316e-03 1.036529e-03
6.925455e-03 8.603570e-04
-5.178523e-03 6.594617e-04
1.184371e-02 5.445549e-04
-2.223611e-03 3.117535e-04
-7.237842e-03 3.237713e-04
1.061541e-03 2.288719e-04
2.104707e-03 2.800939e-04
-1.160097e-03 1.950515e-04
5.556055e-03 2.628884e-04

#dist 5
8.920906e+00 5.574213e+00
3.184051e-01 2.707826e-01
-6.091732e-02 2.111829e-01
3.996906e-01 9.720258e-02
1.819585e-01 2.369708e-01
5.919379e-02 1.007765e-01
1.909702e-01 3.540278e-02
-4.298482e-02 4.139183e-02
2.831195e-02 2.703532e-02
5.771242e-02 1.586627e-02
5.109858e-02 1.679743e-02
-1.433305e-02 1.412962e-02
-5.125656e-02 1.112091e-02
-1.999103e-02 6.803204e-03
-1.186088e-03 7.589752e-03
-1.940123e-02 6.027761e-03
-4.701601e-02 4.588152e-03
5.715671e-01 1.173232e-01
-5.247754e-02 8.588850e-03
-9.429610e-02 5.719260e-03
-1.312943e-02 4.193640e-03
-2.470086e-02 1.438445e-02
-5.667100e-02 3.368548e-03
-1.602311e-02 1.866976e-03
-4.926757e-03 1.474925e-03
-4.141040e-03 1.170606e-03
-2.401318e-02 5.313977e-04
-8.378208e-03 4.724439e-04
7.369990e-03 6.243629e-04
-2.534774e-03 5.475316e-04
-1.195882e-02 2.657291e-04
-5.441646e-04 3.316358e-04
5.679155e-06 2.373653e-04
2.778126e-03 2.780260e-04

#dist 6
8.843653e+00 1.084251e+00
1.272860e+00 1.074784e-01
3.266067e-01 1.334473e-01
4.744913e-01 7.188593e-02
4.303334e-01 4.460279e-02
1.701096e-03 7.609994e-02
1.267044e-01 5.178479e-02
-6.317082e-02 3.318859e-02
6.973947e-02 2.555096e-02
1.359615e-01 2.513237e-02
1.144510e-01 1.681987e-02
2.117875e-02 2.621165e-02
-1.726984e-02 1.794683e-02
-2.884286e-02 1.799892e-02
-3.406123e-03 1.226582e-02
-3.966552e-02 1.195870e-02
-8.367332e-02 5.190467e-03
-6.433897e-01 7.182212e-02
5.609998e-02 6.581336e-03
1.318546e-01 4.705865e-03
3.507444e-02 2.996104e-03
7.689263e-03 2.677077e-03
7.755766e-02 2.673992e-03
-2.605197e-02 1.729159e-03
2.594413e-02 1.380908e-03
7.944467e-03 9.291965e-04
1.802456e-02 8.429603e-04
1.179746e-02 6.941549e-04
1.820310e-02 8.803098e-04
1.487195e-02 7.646122e-04
1.699277e-02 5.380640e-04
5.304407e-03 5.917278e-04
3.683225e-03 3.400461e-04
6.883316e-03 3.247524e-04

#dist 7
6.091804e+00 7.816690e-01
6.172204e-01 5.094878e-02
4.382918e-01 3.219034e-02
5.829664e-01 2.745685e-02
6.204961e-01 1.924904e-02
4.645875e-01 1.422066e-02
2.462419e-01 1.563598e-02
-4.485929e-02 3.012150e-02
6.859719e-02 1.460173e-02
2.076709e-01 1.516340e-02
3.894049e-02 9.766683e-03
-1.165197e-01 8.139835e-03
-5.346617e-02 6.551810e-03
1.449122e-02 6.367717e-03
1.413468e-02 4.900170e-03
-5.878666e-02 5.045001e-03
-1.130098e-01 3.313330e-03
1.911914e-01 1.136998e-01
3.662796e-02 7.094014e-03
1.142230e-02 2.469314e-03
1.679888e-02 1.706258e-03
4.529798e-03 3.154541e-03

```

```

2.078954e-03 1.453971e-03
3.237034e-03 4.068524e-04
2.563418e-03 1.149962e-03
-8.788273e-03 3.863837e-04
-4.537316e-03 6.650789e-04
3.834685e-03 3.000126e-04
-2.163096e-03 5.249395e-04
-5.627714e-03 2.525736e-04
2.803694e-04 1.809133e-04
1.139108e-04 1.845072e-04
-3.163567e-04 1.957878e-04
-1.851176e-03 2.296436e-04

```

```

#dist 8
1.005954e+01 2.539114e+00
4.389464e-01 1.796631e-01
-1.170110e-01 6.745974e-02
2.294027e-01 5.426629e-02
7.728097e-02 4.386358e-02
-3.232499e-02 4.624293e-02
1.381446e-01 4.379093e-02
9.787327e-05 3.415179e-02
4.884301e-02 1.778012e-02
4.345350e-02 1.546054e-02
3.043819e-02 1.576603e-02
1.604341e-02 2.194553e-02
-5.349475e-02 1.400239e-02
-5.955709e-02 1.138534e-02
-2.668870e-02 7.933105e-03
-4.608196e-02 6.625697e-03
-7.636894e-02 6.529650e-03
9.890963e-01 5.739010e-02
-1.572476e-02 7.663406e-03
-1.263402e-01 6.842539e-03
-3.948657e-02 2.811045e-03
4.901626e-04 3.459620e-03
-8.862891e-02 2.922550e-03
-3.395425e-03 3.058117e-03
-4.349846e-02 1.254798e-03
-1.943580e-02 9.525873e-04
-2.319276e-02 7.729821e-04
-2.248027e-02 7.252822e-04
-2.913454e-02 8.594850e-04
-2.039939e-02 5.307304e-04
-1.556884e-02 5.117039e-04
-2.365825e-03 3.666651e-04
-7.276458e-03 4.000917e-04
-1.131821e-02 4.109822e-04

```

```

#dist 9
6.702361e+00 1.080437e+00
9.006195e-01 9.387314e-02
6.326792e-01 3.555728e-02
5.184745e-01 2.122361e-02
4.328576e-01 1.993603e-02
2.566330e-01 1.609128e-02
1.563300e-01 2.034964e-02
1.063134e-01 1.220399e-02
2.362369e-01 9.019439e-03
2.408388e-01 1.197638e-02
5.670889e-02 9.893353e-03
-1.239189e-01 1.250510e-02
-7.835069e-02 1.071713e-02
2.484791e-02 6.114309e-03
-4.391410e-03 9.330450e-03
-4.717435e-02 6.480177e-03
-4.924523e-02 2.294312e-03
-3.599774e-02 3.165447e-02
-4.144309e-02 5.537765e-03
-4.842710e-03 3.807481e-03
2.576792e-03 9.386520e-04
-3.086968e-02 1.122019e-03
-2.392578e-03 7.474847e-04
5.669761e-03 6.064854e-04
1.452982e-02 5.050353e-04
-5.104310e-03 5.071947e-04
-7.282554e-03 4.013770e-04
-3.706491e-03 3.651826e-04
1.515213e-02 3.985712e-04
-9.439287e-04 3.717395e-04
-7.597599e-03 2.026380e-04
4.001857e-03 2.095349e-04
7.388024e-03 1.551652e-04
1.929608e-03 8.941702e-05

```

```

; ===== Output Probability Density Distribution =====
; #outp [ outp index ]
; [ mix num ]
; [ dist index ] [ mix weight ]
; .....
;
#outp 0
2
0 4.864573e-01
8 5.135427e-01

```

```

#outp 1
2
1 3.662454e-01

```

```
6 6.337546e-01
```

```

#outp 2
2
2 7.355008e-01
5 2.644992e-01

```

```

#outp 3
2
3 5.034610e-01
7 4.965390e-01

```

```

#outp 4
2
4 5.610860e-01
9 4.389140e-01

```

```

; ===== HMnet Condition =====
; #total_state [ total state num ]
; #max_path_length [ max state path length ]
; #factor_num [ factor num ]
;
#total_state 5
#max_path_length 4
#factor_num 3

```

```

; ===== Factor Element =====
; #element
; [ element list ]
;
#element 0
u a o i n g e - q

```

```

#element 1
b d g

```

```

#element 2
i a u e o j

```

```

; ===== Tied Factor Element =====
; #tied_element
; [ tied element list ]
;
#tied_element
MHT

```

```

; ===== State Attribute =====
; #state [ state index ]
; [ element list ]
; .....
; [ preceding state list ]
; [ succeeding state list ]
; [ tied outp num ]
; [ outp index ] [ sample num ]
; [ tied elmenet list ]
; .....
; [ self-loop prob ] [ transition prob ]
;
#state 0
u a o i n g e - q
b d
i a u e o j
4
-1
1
0 425
MHT
7.970145e-01 2.029855e-01

```

```

#state 1
u a o i n g e - q
b d
i a u e o j
-1
4
1
1 425
MHT
7.314700e-01 2.685300e-01

```

```

#state 2
u a o i n g e - q
i a u e o j
3
-1
1
2 293
MHT
8.804138e-01 1.195862e-01

```

```

#state 3
u a o i n g e - q
g
i a u e o j

```

```
-1
2
1
3 293
MHT
8.414399e-01 1.585601e-01

#state 4
u a o i n g e - q
b d
```

```
i a u e o j
1
0
1
4 425
MHT
7.871328e-01 2.128672e-01
```

付録 B

HMnet 生成時のログファイルの例

話者 MHT の 3 音素 (/b/, /d/, /g/) を使用した HMnet 生成時に得られたログファイル

```

*** State : 0
*** <<< SSS-ToolKit Version 3.1 >>> [Element0] : u a o i n g e - q
*** Copyright(C) 1993 Jun-ichi TAKAMI [Element1] : b d g
*** ATR Interpreting Telecommunications Research Laboratories [Element2] : i a u e o j
*** Department 1 Outp : 0 ( Sample Number : 718 )
*** 2-2 Hikaridai Seika-cho Soraku-gun Kyoto 619-02 [Element*] : MHT
*** Tel : 07749-5-1301 # State : 1
*** (Direct): 07749-5-1379 [Element0] : u a o i n g e - q
*** Fax : 07749-5-1308 [Element1] : b d g
*** E-mail : jun@itl.atr.co.jp [Element2] : i a u e o j
*** ***** [Element*] : MHT
Wed Feb 9 01:25:42 1994 Outp : 1 ( Sample Number : 718 )
*** [Element*] : MHT
*** <<< Generate HMnet by SSS Algorithm >>> # Used Data : 100.000 %
*** 0 3.657299e+05 0.000000e+00
*** Current Working Directory : /tmp_mnt/home/atrh12/jun/SSS/Exe/9971 [Element*] : MHT
*** [HMnet] # Iteration Times : 11
*** Algorithm Type : SSS # Total Probability : 3.791239e+05
*** Max. State Number : 20 # Allophone Model Number : 1
*** Initial Path Length : 1 # Accumulative State Number : 2
*** Max. Path Length : 4 # Modeling Efficiency : 1.000000
*** Factor Splitting Priority : 1/0,1/2 # CPU Time : 16.000 sec
*** Tied Element : MHT # Total State : 3 / Total Outp : 3
*** Factor Number : 3 # Date : Wed Feb 9 01:26:58 1994
*** Main Element : b d g # Splittee State : 0
*** Output Model File Name : Test.HP-UX/Train_HMnet/HMnet [Element*] : MHT
*** Model Save Step : 5 # State : 0
*** # State : 0
*** [Data] # State : 0
*** Parameter Dimension : 34 [Element0] : u a o i n g e - q
*** Total Data Number : 718 [Element1] : b d
*** Total Sample Number : 718 [Element2] : i a u e o j
*** Total Frame Length : 9971 Outp : 0 ( Sample Number : 425 )
*** Min. Frame Length : 4 [Element*] : MHT
*** # State : 2
*** # Used Data : 100.000 % [Element0] : u a o i n g e - q
*** 0 3.450414e+05 0.000000e+00 [Element1] : g
*** 1 3.545981e+05 2.695089e-02 [Element2] : i a u e o j
*** 2 3.556484e+05 2.953136e-03 Outp : 2 ( Sample Number : 293 )
*** 3 3.557149e+05 1.868009e-04 [Element*] : MHT
*** 4 3.557280e+05 3.706450e-05 # Used Data : 100.000 %
*** 5 3.557363e+05 2.307809e-05 # State : 0
*** 6 3.557508e+05 4.091669e-05 # Iteration Times : 11
*** 7 3.557836e+05 9.214843e-05 # Total Probability : 3.566693e+05
*** 8 3.558576e+05 2.079938e-04 # Allophone Model Number : 1
*** 9 3.560137e+05 4.383148e-04 # Accumulative State Number : 1
*** 10 3.562935e+05 7.852939e-04 # Modeling Efficiency : 1.000000
*** 11 3.566693e+05 1.053724e-03 # CPU Time : 6.610 sec
*** # Total State : 1 / Total Outp : 1
*** # Date : Wed Feb 9 01:25:42 1994
*** # State : 0
*** [Element0] : u a o i n g e - q
*** [Element1] : b d g
*** [Element2] : i a u e o j
*** Outp : 0 ( Sample Number : 718 )
*** [Element*] : MHT
*** # State : 2
*** [Element0] : u a o i n g e - q
*** [Element1] : g
*** [Element2] : i a u e o j
*** Outp : 2 ( Sample Number : 293 )
*** [Element*] : MHT
*** # Used Data : 100.000 %
*** 0 3.759599e+05 0.000000e+00
*** 1 3.832626e+05 1.905421e-02
*** 2 3.847943e+05 3.980454e-03
*** 3 3.862757e+05 3.835039e-03
*** 4 3.881013e+05 4.703935e-03
*** 5 3.889307e+05 2.132690e-03
*** 6 3.894095e+05 1.229515e-03
*** 7 3.899909e+05 1.490759e-03
*** 8 3.906999e+05 1.814657e-03
*** 9 3.913001e+05 1.533801e-03
*** 10 3.916404e+05 8.688820e-04
*** 11 3.918015e+05 4.113874e-04
*** # Iteration Times : 11
*** # Total Probability : 3.918015e+05
*** # Allophone Model Number : 2
*** # Accumulative State Number : 4
*** # Modeling Efficiency : 1.333333
*** # CPU Time : 22.310 sec

```

```

# Total State : 4 / Total Outp : 4
# Date : Wed Feb 9 01:28:08 1994
# Splittee State : 1
# Split State : 1 / 3 ( Factor : 1 )
# State : 1
[Element0] : u a o i n g e - q
[Element1] : b d
[Element2] : i a u e o j
Outp : 1 ( Sample Number : 425 )
[Element*] : MHT
# State : 3
[Element0] : u a o i n g e - q
[Element1] : g
[Element2] : i a u e o j
Outp : 3 ( Sample Number : 293 )
[Element*] : MHT
# Used Data : 100.000 %
0 3.876204e+05 0.000000e+00
1 3.942262e+05 1.675624e-02
2 3.963582e+05 5.379066e-03
3 3.985838e+05 5.583761e-03
4 4.005871e+05 5.000766e-03
5 4.017926e+05 3.000452e-03
6 4.026555e+05 2.143002e-03
7 4.032473e+05 1.467441e-03
8 4.035864e+05 8.404143e-04
9 4.037677e+05 4.488803e-04
10 4.038596e+05 2.275980e-04
11 4.039120e+05 1.297143e-04
# Iteration Times : 11
# Total Probability : 4.039120e+05
# Allophone Model Number : 2
# Accumulative State Number : 4
# Modeling Efficiency : 1.000000
# CPU Time : 23.790 sec

# Total State : 5 / Total Outp : 5
# Date : Wed Feb 9 01:29:15 1994
# Splittee State : 0
# Split State : 4 - 0
# State : 0
[Element0] : u a o i n g e - q
[Element1] : b d
[Element2] : i a u e o j
Outp : 0 ( Sample Number : 425 )
[Element*] : MHT
# State : 4
[Element0] : u a o i n g e - q
[Element1] : b d
[Element2] : i a u e o j
Outp : 4 ( Sample Number : 425 )
[Element*] : MHT
# Used Data : 59.192 %
0 4.025357e+05 0.000000e+00
1 4.076613e+05 1.257310e-02
2 4.087723e+05 2.718062e-03
3 4.102390e+05 3.575069e-03
4 4.112076e+05 2.355638e-03
5 4.117482e+05 1.312880e-03
6 4.123127e+05 1.368979e-03
7 4.130259e+05 1.726761e-03
8 4.135036e+05 1.155376e-03
9 4.137473e+05 5.888815e-04
10 4.138798e+05 3.202614e-04
11 4.139683e+05 2.138365e-04
# Iteration Times : 11
# Total Probability : 4.139683e+05
# Allophone Model Number : 2
# Accumulative State Number : 5
# Modeling Efficiency : 1.000000
# CPU Time : 15.210 sec

# Total State : 6 / Total Outp : 6
# Date : Wed Feb 9 01:30:00 1994
# Splittee State : 1
# Split State : 5 - 1
# State : 1
[Element0] : u a o i n g e - q
[Element1] : b d
[Element2] : i a u e o j
Outp : 1 ( Sample Number : 425 )
[Element*] : MHT
# State : 5
[Element0] : u a o i n g e - q
[Element1] : b d
[Element2] : i a u e o j
Outp : 5 ( Sample Number : 425 )
[Element*] : MHT
# Used Data : 59.192 %
0 4.078932e+05 0.000000e+00
1 4.139286e+05 1.458075e-02
2 4.151872e+05 3.031553e-03
3 4.165906e+05 3.368774e-03
4 4.173623e+05 1.848907e-03
5 4.175823e+05 5.268085e-04

6 4.178052e+05 5.334052e-04
7 4.181006e+05 7.066047e-04
8 4.182783e+05 4.248350e-04
9 4.184312e+05 3.655376e-04
10 4.185943e+05 3.895257e-04
11 4.188020e+05 4.959340e-04
# Iteration Times : 11
# Total Probability : 4.188020e+05
# Allophone Model Number : 2
# Accumulative State Number : 6
# Modeling Efficiency : 1.000000
# CPU Time : 18.480 sec

# Total State : 7 / Total Outp : 7
# Date : Wed Feb 9 01:30:55 1994
# Splittee State : 5
# Split State : 5 / 6 ( Factor : 1 )
# State : 5
[Element0] : u a o i n g e - q
[Element1] : d
[Element2] : i a u e o j
Outp : 5 ( Sample Number : 203 )
[Element*] : MHT
# State : 6
[Element0] : u a o i n g e - q
[Element1] : b
[Element2] : i a u e o j
Outp : 6 ( Sample Number : 222 )
[Element*] : MHT
# Used Data : 59.192 %
0 4.147136e+05 0.000000e+00
1 4.205157e+05 1.379743e-02
2 4.208289e+05 7.444497e-04
3 4.209462e+05 2.784590e-04
4 4.210312e+05 2.019582e-04
5 4.211060e+05 1.776655e-04
6 4.211800e+05 1.757088e-04
7 4.212672e+05 2.070325e-04
8 4.213830e+05 2.748142e-04
9 4.215195e+05 3.238195e-04
10 4.216683e+05 3.527656e-04
11 4.218677e+05 4.728024e-04
# Iteration Times : 11
# Total Probability : 4.218677e+05
# Allophone Model Number : 3
# Accumulative State Number : 10
# Modeling Efficiency : 1.428571
# CPU Time : 18.770 sec

# Total State : 8 / Total Outp : 8
# Date : Wed Feb 9 01:31:51 1994
# Splittee State : 0
# Split State : 0 / 7 ( Factor : 1 )
# State : 0
[Element0] : u a o i n g e - q
[Element1] : b
[Element2] : i a u e o j
Outp : 0 ( Sample Number : 222 )
[Element*] : MHT
# State : 7
[Element0] : u a o i n g e - q
[Element1] : d
[Element2] : i a u e o j
Outp : 7 ( Sample Number : 203 )
[Element*] : MHT
# Used Data : 59.192 %
0 4.164410e+05 0.000000e+00
1 4.223954e+05 1.409678e-02
2 4.228964e+05 1.184483e-03
3 4.231335e+05 5.604145e-04
4 4.232939e+05 3.789629e-04
5 4.233858e+05 2.171195e-04
6 4.234332e+05 1.119118e-04
7 4.234628e+05 6.994128e-05
8 4.235042e+05 9.762529e-05
9 4.236023e+05 2.315632e-04
10 4.237609e+05 3.742679e-04
11 4.238936e+05 3.130839e-04
# Iteration Times : 11
# Total Probability : 4.238936e+05
# Allophone Model Number : 3
# Accumulative State Number : 10
# Modeling Efficiency : 1.250000
# CPU Time : 18.210 sec

# Total State : 9 / Total Outp : 9
# Date : Wed Feb 9 01:32:47 1994
# Splittee State : 4
# Split State : 4 / 8 ( Factor : 1 )
# State : 4
[Element0] : u a o i n g e - q
[Element1] : d
[Element2] : i a u e o j
Outp : 4 ( Sample Number : 203 )
[Element*] : MHT
# State : 8

```

```

[Element0] : u a o i n g e - q
[Element1] : b
[Element2] : i a u e o j
Outp : 8 ( Sample Number : 222 )
[Element*] : MHT
# Used Data : 59.192 %
0 4.182995e+05 0.000000e+00
1 4.239493e+05 1.332680e-02
2 4.249404e+05 2.332135e-03
3 4.256211e+05 1.599340e-03
4 4.260759e+05 1.067359e-03
5 4.264147e+05 7.946884e-04
6 4.266518e+05 5.556896e-04
7 4.268028e+05 3.538229e-04
8 4.268995e+05 2.265402e-04
9 4.269853e+05 2.009033e-04
10 4.271117e+05 2.958060e-04
11 4.272618e+05 3.513385e-04
# Iteration Times : 11
# Total Probability : 4.272618e+05
# Allophone Model Number : 3
# Accumulative State Number : 10
# Modeling Efficiency : 1.111111
# CPU Time : 18.350 sec

# Total State : 10 / Total Outp : 10
# Date : Wed Feb 9 01:33:42 1994
# Splittee State : 1
# Split State : 1 / 9 ( Factor : 1 )
# State : 1
[Element0] : u a o i n g e - q
[Element1] : b
[Element2] : i a u e o j
Outp : 1 ( Sample Number : 222 )
[Element*] : MHT
# State : 9
[Element0] : u a o i n g e - q
[Element1] : d
[Element2] : i a u e o j
Outp : 9 ( Sample Number : 203 )
[Element*] : MHT
# Used Data : 59.192 %
0 4.214651e+05 0.000000e+00
1 4.272781e+05 1.360463e-02
2 4.279616e+05 1.597178e-03
3 4.284385e+05 1.113108e-03
4 4.287998e+05 8.424773e-04
5 4.291338e+05 7.785213e-04
6 4.295155e+05 8.886142e-04
7 4.298206e+05 7.097482e-04
8 4.299837e+05 3.794628e-04
9 4.301049e+05 2.817084e-04
10 4.302211e+05 2.701832e-04
11 4.303505e+05 3.006848e-04
# Iteration Times : 11
# Total Probability : 4.303505e+05
# Allophone Model Number : 3
# Accumulative State Number : 10
# Modeling Efficiency : 1.000000
# CPU Time : 18.460 sec

# Total State : 11 / Total Outp : 11
# Date : Wed Feb 9 01:34:36 1994
# Splittee State : 6
# Split State : 6 / 10 ( Factor : 0 )
# State : 6
[Element0] : u a o i n g e
[Element1] : b
[Element2] : i a u e o j
Outp : 6 ( Sample Number : 161 )
[Element*] : MHT
# State : 10
[Element0] : -
[Element1] : b
[Element2] : i a u e o j
Outp : 10 ( Sample Number : 61 )
[Element*] : MHT
# Used Data : 30.919 %
0 4.266604e+05 0.000000e+00
1 4.301743e+05 8.168442e-03
2 4.306701e+05 1.151244e-03
3 4.311724e+05 1.164878e-03
4 4.317097e+05 1.244655e-03
5 4.322973e+05 1.359298e-03
6 4.326470e+05 8.082513e-04
7 4.328639e+05 5.011206e-04
8 4.330005e+05 3.154049e-04
9 4.331213e+05 2.788436e-04
10 4.332434e+05 2.819528e-04
11 4.333352e+05 2.117709e-04
# Iteration Times : 11
# Total Probability : 4.333352e+05
# Allophone Model Number : 4
# Accumulative State Number : 14
# Modeling Efficiency : 1.272727
# CPU Time : 9.880 sec

# Total State : 12 / Total Outp : 12
# Date : Wed Feb 9 01:35:07 1994
# Splittee State : 3
# Split State : 3 / 11 ( Factor : 0 )
# State : 3
[Element0] : u a o i n g e
[Element1] : g
[Element2] : i a u e o j
Outp : 3 ( Sample Number : 220 )
[Element*] : MHT
# State : 11
[Element0] : -
[Element1] : g
[Element2] : i a u e o j
Outp : 11 ( Sample Number : 73 )
[Element*] : MHT
# Used Data : 40.808 %
0 4.330550e+05 0.000000e+00
1 4.362964e+05 7.429381e-03
2 4.374052e+05 2.534964e-03
3 4.381269e+05 1.647270e-03
4 4.388237e+05 1.587841e-03
5 4.395170e+05 1.577358e-03
6 4.400059e+05 1.111213e-03
7 4.402615e+05 5.804288e-04
8 4.404195e+05 3.589588e-04
9 4.405350e+05 2.621831e-04
10 4.406109e+05 1.721727e-04
11 4.406617e+05 1.153759e-04
# Iteration Times : 11
# Total Probability : 4.406617e+05
# Allophone Model Number : 5
# Accumulative State Number : 16
# Modeling Efficiency : 1.333333
# CPU Time : 9.040 sec

# Total State : 13 / Total Outp : 13
# Date : Wed Feb 9 01:35:35 1994
# Splittee State : 2
# Split State : 2 / 12 ( Factor : 0 )
# State : 2
[Element0] : u a o i n g e
[Element1] : g
[Element2] : i a u e o j
Outp : 2 ( Sample Number : 220 )
[Element*] : MHT
# State : 12
[Element0] : -
[Element1] : g
[Element2] : i a u e o j
Outp : 12 ( Sample Number : 73 )
[Element*] : MHT
# Used Data : 40.808 %
0 4.366001e+05 0.000000e+00
1 4.401014e+05 7.955556e-03
2 4.413947e+05 2.930142e-03
3 4.423475e+05 2.154014e-03
4 4.432497e+05 2.035350e-03
5 4.437918e+05 1.221611e-03
6 4.440297e+05 5.357890e-04
7 4.441579e+05 2.884764e-04
8 4.442393e+05 1.832939e-04
9 4.442873e+05 1.081220e-04
10 4.443341e+05 1.052962e-04
11 4.443941e+05 1.348766e-04
# Iteration Times : 11
# Total Probability : 4.443941e+05
# Allophone Model Number : 5
# Accumulative State Number : 16
# Modeling Efficiency : 1.230769
# CPU Time : 9.080 sec

# Total State : 14 / Total Outp : 14
# Date : Wed Feb 9 01:36:01 1994
# Splittee State : 0
# Split State : 0 / 13 ( Factor : 2 )
# State : 0
[Element0] : u a o i n g e - q
[Element1] : b
[Element2] : i u e j
Outp : 0 ( Sample Number : 126 )
[Element*] : MHT
# State : 13
[Element0] : u a o i n g e - q
[Element1] : b
[Element2] : a o
Outp : 13 ( Sample Number : 96 )
[Element*] : MHT
# Used Data : 30.919 %
0 4.389921e+05 0.000000e+00
1 4.431725e+05 9.432885e-03
2 4.443980e+05 2.757591e-03
3 4.448271e+05 9.647613e-04
4 4.451472e+05 7.190662e-04

```

```

5 4.455207e+05 8.383485e-04
6 4.460467e+05 1.179071e-03
7 4.464537e+05 9.117447e-04
8 4.466532e+05 4.466807e-04
9 4.467992e+05 3.266221e-04
10 4.469254e+05 2.824033e-04
11 4.470242e+05 2.212042e-04
# Iteration Times : 11
# Total Probability : 4.470242e+05
# Allophone Model Number : 7
# Accumulative State Number : 24
# Modeling Efficiency : 1.714286
# CPU Time : 9.940 sec

# Total State : 15 / Total Outp : 15
# Date : Wed Feb 9 01:36:31 1994
# Splittee State : 5
# Split State : 5 / 14 ( Factor : 0 )
# State : 5
[Element0] : u a o i n g e
[Element1] : d
[Element2] : i a u e o j
Outp : 5 ( Sample Number : 134 )
[Element*] : MHT
# State : 14
[Element0] : - q
[Element1] : d
[Element2] : i a u e o j
Outp : 14 ( Sample Number : 69 )
[Element*] : MHT
# Used Data : 28.273 %
0 4.440356e+05 0.000000e+00
1 4.468522e+05 6.303054e-03
2 4.472970e+05 9.944660e-04
3 4.475740e+05 6.189720e-04
4 4.477987e+05 5.016238e-04
5 4.479916e+05 4.307143e-04
6 4.481706e+05 3.993971e-04
7 4.484331e+05 5.853916e-04
8 4.488400e+05 9.064552e-04
9 4.491259e+05 6.367299e-04
10 4.492504e+05 2.769218e-04
11 4.493333e+05 1.845339e-04
# Iteration Times : 11
# Total Probability : 4.493333e+05
# Allophone Model Number : 8
# Accumulative State Number : 28
# Modeling Efficiency : 1.866667
# CPU Time : 8.770 sec

# Total State : 16 / Total Outp : 16
# Date : Wed Feb 9 01:36:58 1994
# Splittee State : 7
# Split State : 7 / 15 ( Factor : 2 )
# State : 7
[Element0] : u a o i n g e - q
[Element1] : d
[Element2] : i a o
Outp : 7 ( Sample Number : 175 )
[Element*] : MHT
# State : 15
[Element0] : u a o i n g e - q
[Element1] : d
[Element2] : e
Outp : 15 ( Sample Number : 28 )
[Element*] : MHT
# Used Data : 28.273 %
0 4.447208e+05 0.000000e+00
1 4.475435e+05 6.307111e-03
2 4.479793e+05 9.727756e-04
3 4.482473e+05 5.979919e-04
4 4.484641e+05 4.833374e-04
5 4.486580e+05 4.322705e-04
6 4.488412e+05 4.079935e-04
7 4.491343e+05 6.527552e-04
8 4.495469e+05 9.177583e-04
9 4.498285e+05 6.259320e-04
10 4.499525e+05 2.755162e-04
11 4.500387e+05 1.915862e-04
# Iteration Times : 11
# Total Probability : 4.500387e+05
# Allophone Model Number : 10
# Accumulative State Number : 36
# Modeling Efficiency : 2.250000
# CPU Time : 8.770 sec

# Total State : 17 / Total Outp : 17
# Date : Wed Feb 9 01:37:25 1994
# Splittee State : 6
# Split State : 6 / 16 ( Factor : 0 )
# State : 6
[Element0] : u a o i e
[Element1] : b
[Element2] : i a u e o j
Outp : 6 ( Sample Number : 149 )
[Element*] : MHT
# State : 16
[Element0] : ng
[Element1] : b
[Element2] : i a u e o j
Outp : 16 ( Sample Number : 12 )
[Element*] : MHT
# Used Data : 30.919 %
0 4.429082e+05 0.000000e+00
1 4.472947e+05 9.806695e-03
2 4.486298e+05 2.975857e-03
3 4.491106e+05 1.070602e-03
4 4.494399e+05 7.327457e-04
5 4.497397e+05 6.666182e-04
6 4.500930e+05 7.848807e-04
7 4.503090e+05 4.797406e-04
8 4.504734e+05 3.648856e-04
9 4.505897e+05 2.581211e-04
10 4.506543e+05 1.433165e-04
11 4.506965e+05 9.372997e-05
# Iteration Times : 11
# Total Probability : 4.506965e+05
# Allophone Model Number : 12
# Accumulative State Number : 44
# Modeling Efficiency : 2.588235
# CPU Time : 10.080 sec

# Total State : 18 / Total Outp : 18
# Date : Wed Feb 9 01:37:55 1994
# Splittee State : 7
# Split State : 7 / 17 ( Factor : 0 )
# State : 7
[Element0] : u a o n g - q
[Element1] : d
[Element2] : i a o
Outp : 7 ( Sample Number : 147 )
[Element*] : MHT
# State : 17
[Element0] : i e
[Element1] : d
[Element2] : i a o
Outp : 17 ( Sample Number : 28 )
[Element*] : MHT
# Used Data : 28.273 %
0 4.459035e+05 0.000000e+00
1 4.487661e+05 6.378873e-03
2 4.492019e+05 9.703272e-04
3 4.494913e+05 6.436550e-04
4 4.497180e+05 5.040596e-04
5 4.499059e+05 4.178624e-04
6 4.500437e+05 3.060442e-04
7 4.502835e+05 5.324828e-04
8 4.506957e+05 9.146322e-04
9 4.510025e+05 6.803281e-04
10 4.511927e+05 4.216205e-04
11 4.513711e+05 3.950822e-04
# Iteration Times : 11
# Total Probability : 4.513711e+05
# Allophone Model Number : 13
# Accumulative State Number : 48
# Modeling Efficiency : 2.666667
# CPU Time : 8.810 sec

# Total State : 19 / Total Outp : 19
# Date : Wed Feb 9 01:38:24 1994
# Splittee State : 8
# Split State : 8 / 18 ( Factor : 2 )
# State : 8
[Element0] : u a o i n g e - q
[Element1] : b
[Element2] : i a e j
Outp : 8 ( Sample Number : 117 )
[Element*] : MHT
# State : 18
[Element0] : u a o i n g e - q
[Element1] : b
[Element2] : u o
Outp : 18 ( Sample Number : 105 )
[Element*] : MHT
# Used Data : 30.919 %
0 4.460148e+05 0.000000e+00
1 4.505657e+05 1.010037e-02
2 4.519942e+05 3.160438e-03
3 4.525298e+05 1.183562e-03
4 4.527984e+05 5.932622e-04
5 4.529203e+05 2.691251e-04
6 4.530136e+05 2.057717e-04
7 4.530495e+05 7.943657e-05
8 4.530728e+05 5.128506e-05
9 4.530859e+05 2.899385e-05
10 4.530948e+05 1.960157e-05
11 4.531003e+05 1.223038e-05
# Iteration Times : 11
# Total Probability : 4.531003e+05
# Allophone Model Number : 19
# Accumulative State Number : 72
# Modeling Efficiency : 3.789474
# CPU Time : 10.000 sec

```

```

# Total State : 20 / Total Outp : 20
# Date : Wed Feb 9 01:38:55 1994.
# Splittee State : 2
# Split State : 2 / 19 ( Factor : 2 )
# State : 2
  [Element0] : u a o i n g e
  [Element1] : g
  [Element2] : i u o j
  Outp : 2 ( Sample Number : 96 )
  [Element*] : MHT
# State : 19
  [Element0] : u a o i n g e
  [Element1] : g
  [Element2] : a e
  Outp : 19 ( Sample Number : 124 )
  [Element*] : MHT
# Used Data : 30.641 %
  0 4.513163e+05 0.000000e+00

```

```

1 4.534412e+05 4.686322e-03
2 4.539700e+05 1.164760e-03
3 4.541694e+05 4.390744e-04
4 4.543019e+05 2.915412e-04
5 4.544927e+05 4.198695e-04
6 4.548460e+05 7.768051e-04
7 4.551948e+05 7.662925e-04
8 4.553983e+05 4.467082e-04
9 4.555255e+05 2.793365e-04
10 4.556034e+05 1.710115e-04
11 4.556507e+05 1.038777e-04
# Iteration Times : 11
# Total Probability : 4.556507e+05
# Allophone Model Number : 20
# Accumulative State Number : 74
# Modeling Efficiency : 3.700000
# CPU Time : 6.010 sec

```