

Internal Use Only (非公開)

TR-SLT-0032

日本語及び中国語音声合成のためのHMMによる韻律生成部の設計

HMM-based prosody modeling and synthesis
for Japanese and Chinese speech synthesis

全 炳河
Heiga Zen

陸 金林
Jinlin Lu

倪 晋富
Jinfu Ni

徳田 恵一
Keiichi Tokuda

河井 恒
Hisashi Kawai

2002年12月6日

概要

一般にテキスト音声合成システムは、テキスト処理部、韻律生成部、波形生成から構成されている。自然な韻律(F_0 パターン)を生成することは、高品質な合成音を得るために不可欠であると考えられる。近年、HMMを用いた F_0 パターンのモデル化及び生成手法が提案され、自然音声に近い高品質な F_0 パターンが生成できる事が報告されている。この手法では、言語情報や音韻情報を含んだコンテキスト依存ラベルの仕様及び決定木によるコンテキストクラスタリング時に用いる分割条件が言語依存である。生成される F_0 パターンの自然性は、コンテキスト依存ラベルに F_0 パターンの制御に必要な言語・音韻情報が含まれているか、適切な分割条件が用意されているかに依存する。本研究では、HMMによる F_0 パターンのモデル化・生成手法を日本語及び中国語に適用するため、コンテキスト依存ラベル及び分割条件の設計を行った。

(株) 国際電気通信基礎技術研究所
音声言語コミュニケーション研究所

〒619-0288 「けいはんな学研都市」 光台二丁目2番地2 TEL : 0774-95-1301

Advanced Telecommunication Research Institute International
Spoken Language Translation Research Laboratories
2-2-2 Hikaridai "Keihanna Science City" 619-0288, Japan
Telephone: +81-774-95-1301
Fax : +81-774-95-1308

©2002 (株) 国際電気通信基礎技術研究所

©2002 Advanced Telecommunication Research Institute International

目次

1	はじめに	1
2	HMM による日本語韻律合成のためのコンテキスト依存ラベルの設計	2
2.1	導入したコンテキスト	2
(2.1.1)	音素に関するコンテキスト	2
(2.1.2)	モーラに関するコンテキスト	2
(2.1.3)	形態素に関するコンテキスト	3
(2.1.4)	韻律語に関するコンテキスト	3
(2.1.5)	呼気段落に関するコンテキスト	4
(2.1.6)	文章に関するコンテキスト	4
(2.1.7)	コンテキスト依存ラベルの仕様	5
2.2	コンテキストクラスタリング時の質問	6
(2.2.1)	音素に関する質問	6
(2.2.2)	モーラに関する質問	7
(2.2.3)	形態素に関する質問	7
(2.2.4)	韻律語に関する質問	11
(2.2.5)	呼気段落に関する質問	12
(2.2.6)	文章に関する質問	12
3	HMM による中国語韻律合成のためのコンテキスト依存ラベルの設計	13
3.1	導入したコンテキスト	13
(3.1.1)	半音節に関するコンテキスト	13
(3.1.2)	音節に関するコンテキスト	14
(3.1.3)	形態素に関するコンテキスト	14
(3.1.4)	呼気段落に関するコンテキスト	15
(3.1.5)	文章に関するコンテキスト	16
(3.1.6)	コンテキスト依存ラベルの仕様	16
3.2	クラスタリング時に用いる質問	17
(3.2.1)	半音節に関する質問	17
(3.2.2)	音節に関する質問	19
(3.2.3)	品詞に関する質問	20
(3.2.4)	呼気段落に関する質問	21
(3.2.5)	文章に関する質問	21
4	おわりに	22

1 はじめに

近年, コーパスベースの音声合成手法に関する研究が盛んに行われている. この方式は, 従来の規則に基づいた合成方式の多くが発見的な手法に基づいているのに対し, 大量のデータを用いた自動学習や音声単位選択に基づいているため, 高品質で自然性の高い音声を合成しやすい, システムの自動学習が容易である, 音声データ提供話者の個人性, 更には発話様式が合成音に反映される, などの特徴を持つ.

隠れマルコフモデル (Hidden Markov Model:HMM) は音声認識において広く用いられている手法の 1 つである. HMM を F_0 パターンの生成に用いる試みはいくつか行われている [1] [2].

HMM に基づく音声合成システム [3] では, F_0 パターンを有声を表す 1 次元空間の出力と無声を表す 0 次元空間からの出力が時間的に混在した系列としてとらえ, 多空間上の確率分布に基づく HMM(multi-space probability distribution HMM: MSD-HMM) [4] を用いてモデル化し, 尤度最大化基準に基づき滑らかな F_0 パターンを生成する手法 [5] を用いて F_0 パターンのモデル化・生成を行っている. この手法は HMM の枠組みの中で便宜的な近似を行わずに F_0 パターンを取り扱うため, 幾つかの有用な手法 (決定木に基づくコンテキストクラスタリング, 話者適応など) を利用することができる. また, 実音声に非常に近い F_0 パターンを生成することができる.

HMM に基づく音声合成システムでは, 音韻情報や韻律情報, 言語情報を, コンテキスト (変動要因) として組み込んだコンテキスト依存ラベルを用いて HMM を学習する. コンテキスト依存ラベルの仕様及びクラスタリング時の分割条件のみ言語依存であり, その他の部分は独立であるため, 多言語に容易に適用できる点が大きな特徴である. これまで日本語 [3] 及び英語 [6] に適用され, コンテキスト依存ラベルとクラスタリング時の分割条件を各言語に対して適切に設定する事で同じ枠組みで音声合成システムを構築できることが確認されている.

我々の研究グループでは, 日本語及び中国語の単位選択型音声合成システムの研究を行っており, 単位選択の為の韻律予測に HMM による音声合成システムを利用しようと考えている. 日本語における HMM 音声合成システム [3] で利用しているコンテキストの詳細は, [7] に公開されているが, 位置を表現する際に前からカウントした位置の情報しか利用していない, カウントする単位がモーラのみであるなど, 韻律を詳細にモデル化するには, 情報が不足していると思われる.

そこで本研究では, 高精度な韻律合成を行うため, 日本語のコンテキスト依存ラベルの仕様を見直し, より精度の高い韻律合成器を設計する. また, 同様に中国語のコンテキスト依存ラベルを設計する.

2 HMM による日本語韻律合成のためのコンテキスト依存ラベルの設計

2.1 導入したコンテキスト

日本語における HMM 音声合成システムでは、1 つの HMM は 1 つの音素に対応している。コンテキストとしては音素を最小単位とし、モーラ、形態素、韻律語 (アクセント句)、呼吸段落、文章の順で階層的にコンテキスト情報を取り扱う。各階層におけるコンテキスト情報を以下に示す。

(2.1.1) 音素に関するコンテキスト

音素セットを以下のように定義する。

- 音素用
a,i,u,e,o (母音)
A,I,U,E,O (無声化母音)
k,s,sh,t,ch,ts,n,h,f,m,y,r,w,g,z,j,d,b,p,ky,ny,hy,my,ry,gy,by,py,dy (子音)
N (撥音)
Q (促音)
- 文中小休止用
pau
- 文頭、文末無音用
sil

音素階層における変動要因としては、前後の音素環境が考えられる。音声認識では、前後 1 つの音素環境だけでなく、前後 2 つの音素環境を考慮することにより、認識率が向上することが報告されている [8]。以上を考慮して、音素階層における変動要因として、以下に述べるものを導入する。

- p_{L2} 先先行音素
- p_{L1} 先行音素
- p_C 当該音素
- p_{R1} 後続音素
- p_{R2} 後後続音素

(2.1.2) モーラに関するコンテキスト

モーラ階層における変動要因として、以下に述べるものを導入する。

- a_1 アクセント核とモーラ位置との差 (単位: モーラ)
- a_2 前からカウントした当該韻律語内での当該モーラ位置 (単位: モーラ)
- a_3 後からカウントした当該韻律語内での当該モーラ位置 (単位: モーラ)

(2.1.3) 形態素に関するコンテキスト

形態素階層における重要な変動要因として、品詞、活用型、活用形が考えられる。品詞体系としては様々なものが提案されている。文献 [3] では、ATR 日本語音声データベース b-set に附属している言語韻律情報ファイルの品詞体系を利用している。音声合成において、品詞や活用形といった情報は、大まかな分類でよいのか、それとも詳細な分類が必要であるかは確認されていない。そこで、本研究では、[9] で提案されている階層的な品詞情報を利用する。コンテキスト依存ラベル中には詳細な情報を記述し、決定木によるクラスタリングの際に大まかな分類や詳細な分類を分割条件として導入することで、適切な分類が行われることが期待できる。

品詞、活用形、活用型は [9] で示されている品詞体系を数値に変換し、これをコンテキスト依存ラベル中に記述する。

形態素階層における変動要因として、以下に述べるものを導入する。

- b_1 先行形態素の品詞
- b_2 先行形態素の品詞の活用形
- b_3 先行形態素の品詞の活用型
- b_4 先行形態素の境界の種別
- c_1 当該形態素の品詞
- c_2 当該形態素の品詞の活用形
- c_3 当該形態素の品詞の活用型
- c_4 当該形態素の境界の種別
- d_1 後続形態素の品詞
- d_2 後続形態素の品詞の活用形
- d_3 後続形態素の品詞の活用型
- d_4 後続形態素の境界の種別

(2.1.4) 韻律語に関するコンテキスト

韻律語階層における変動要因として、以下に述べるものを導入する。

- e_1 先行韻律語の長さ (単位: モーラ)
- e_2 先行韻律語のアクセント型
- e_3 先行韻律語の属性 (強調, 疑問など)
- e_4 先行韻律語と当該韻律語間のポーズの有無
- f_1 当該韻律語の長さ (単位: モーラ)
- f_2 当該韻律語のアクセント型
- f_3 当該韻律語の属性 (強調, 疑問など)
- f_4 前からカウントした当該呼気段落での韻律語の位置 (単位: 韻律語)
- f_5 後からカウントした当該呼気段落での韻律語の位置 (単位: 韻律語)
- f_6 前からカウントした当該呼気段落での韻律語の位置 (単位: モーラ)
- f_7 後からカウントした当該呼気段落での韻律語の位置 (単位: モーラ)
- g_1 後続韻律語の長さ (単位: モーラ)
- g_2 後続韻律語のアクセント型
- g_3 後続韻律語の属性 (強調, 疑問など)
- g_4 後続韻律語と当該韻律語間のポーズの有無

(2.1.5) 呼気段落に関するコンテキスト

呼気段落階層における変動要因として, 以下に述べるものを導入する.

- h_1 先行呼気段落の長さ (単位: 韻律語)
- h_2 先行呼気段落の長さ (単位: モーラ)
- i_1 当該呼気段落の長さ (単位: 韻律語)
- i_2 当該呼気段落の長さ (単位: モーラ)
- i_3 前からカウントした文中での当該呼気段落の位置 (単位: 呼気段落)
- i_4 後からカウントした文中での当該呼気段落の位置 (単位: 呼気段落)
- i_5 前からカウントした文中での当該呼気段落の位置 (単位: 韻律語)
- i_6 後からカウントした文中での当該呼気段落の位置 (単位: 韻律語)
- i_7 前からカウントした文中での当該呼気段落の位置 (単位: モーラ)
- i_8 後からカウントした文中での当該呼気段落の位置 (単位: モーラ)
- j_1 後続呼気段落の長さ (単位: 韻律語)
- j_2 後続呼気段落の長さ (単位: モーラ)

(2.1.6) 文章に関するコンテキスト

文章階層における変動要因として, 以下に述べるものを導入する.

- k_1 文の長さ (単位: 呼気段落)
- k_2 文の長さ (単位: 韻律語)
- k_3 文の長さ (単位: モーラ)

(2.1.7) コンテキスト依存ラベルの仕様

前述の変動要因を HTK [10] で取り扱うため、以下のような仕様のラベルを音素単位で作成する。決定木によるコンテキストクラスタリング時のパターンマッチングに要する時間を極力短縮するため、複雑な仕様になっている。

$$\begin{aligned}
& PL2 \hat{=} PL1 - PC + PR1 = PR2 \\
& /A: a_1 + a_2 + a_3 /B: b_1 \# b_2 \# b_3 \# b_4 /C: c_1 \# c_2 ! c_3 \# c_4 /D: d_1 \# d_2 ; d_3 \# d_4 \\
& /E: e_1 \# e_2 \$ e_3 \# e_4 /F: f_1 \# f_2 \# f_3 @ f_4 \# f_5 | f_6 \# f_7 /G: g_1 \# g_2 \% g_3 \# g_4 \\
& /H: h_1 \# h_2 /I: i_1 \# i_2 @ i_3 + i_4 \& i_5 \# i_6 | i_7 + i_8 /J: j_1 \# j_2 /K: k_1 + k_2 \# k_3
\end{aligned}$$

上記の“アルファベット+数字”の変数はそれぞれ、2.1 のコンテキスト情報に対応する。コンテキスト依存ラベルの例を以下に示す。

文章：「あらゆる現実を全て自分の方へねじ曲げたのだ。」

```

x`x-sil-a+r/A:~x+x#x/B:~x#x#x/C:~x!x#x/D:0000_x!x!D:1500_x;x#E:~x#x#x/F:~x_x#x#x|G:4.3%#x/H:~x#x/I:~x#x#x#x#x|J:2.9/K:2+6-26
x`sil-a+r=A:~2+1+4/B:~x#x#x/C:0000_x!x!D:1500_x;x#E:~x#x#x/F:4.3_x#1_2|1.9/G:5.0%#x/H:~x#x/I:2-9#1+2#1-6|1+26/J:4.17/K:2+6-26
sil`a-r+a=y/A:~1+2+3/B:~x#x#x/C:0000_x!x!D:1500_x;x#E:~x#x#x/F:4.3_x#1_2|1.9/G:5.0%#x/H:~x#x/I:2-9#1+2#1-6|1+26/J:4.17/K:2+6-26
a`r-a+y=u/A:~1+2+3/B:~x#x#x/C:0000_x!x!D:1500_x;x#E:~x#x#x/F:4.3_x#1_2|1.9/G:5.0%#x/H:~x#x/I:2-9#1+2#1-6|1+26/J:4.17/K:2+6-26
r`a-y+u=r/A:~0+3+2/B:~x#x#x/C:0000_x!x!D:1500_x;x#E:~x#x#x/F:4.3_x#1_2|1.9/G:5.0%#x/H:~x#x/I:2-9#1+2#1-6|1+26/J:4.17/K:2+6-26
a`y-u+r=r/A:~0+3+2/B:~x#x#x/C:0000_x!x!D:1500_x;x#E:~x#x#x/F:4.3_x#1_2|1.9/G:5.0%#x/H:~x#x/I:2-9#1+2#1-6|1+26/J:4.17/K:2+6-26
y`u-r+u=r/A:~1+4+1/B:~x#x#x/C:0000_x!x!D:1500_x;x#E:~x#x#x/F:4.3_x#1_2|1.9/G:5.0%#x/H:~x#x/I:2-9#1+2#1-6|1+26/J:4.17/K:2+6-26
u`r-u+g=e/A:~1+4+1/B:~x#x#x/C:0000_x!x!D:1500_x;x#E:~x#x#x/F:4.3_x#1_2|1.9/G:5.0%#x/H:~x#x/I:2-9#1+2#1-6|1+26/J:4.17/K:2+6-26
r`u-g+e=N/A:~1+1+5/B:~0000_x#x!C:1500_x!x#x/D:9630_x;x#E:4.3%#x/F:5.0_x#2_1|5.5/G:3.1%#x/H:~x#x/I:2-9#1+2#1-6|1+26/J:4.17/K:2+6-26
u`g-e+N=j/A:~1+1+5/B:~0000_x#x!C:1500_x!x#x/D:9630_x;x#E:4.3%#x/F:5.0_x#2_1|5.5/G:3.1%#x/H:~x#x/I:2-9#1+2#1-6|1+26/J:4.17/K:2+6-26
g`e-N+j=i/A:~2+3+3/B:~0000_x#x!C:1500_x!x#x/D:9630_x;x#E:4.3%#x/F:5.0_x#2_1|5.5/G:3.1%#x/H:~x#x/I:2-9#1+2#1-6|1+26/J:4.17/K:2+6-26
N`j-i+ts=A:~3+3+3/B:~0000_x#x!C:1500_x!x#x/D:9630_x;x#E:4.3%#x/F:5.0_x#2_1|5.5/G:3.1%#x/H:~x#x/I:2-9#1+2#1-6|1+26/J:4.17/K:2+6-26
N`j-i+ts=U/A:~3+3+3/B:~0000_x#x!C:1500_x!x#x/D:9630_x;x#E:4.3%#x/F:5.0_x#2_1|5.5/G:3.1%#x/H:~x#x/I:2-9#1+2#1-6|1+26/J:4.17/K:2+6-26
j`i-t+s=U/A:~4+4+2/B:~0000_x#x!C:1500_x!x#x/D:9630_x;x#E:4.3%#x/F:5.0_x#2_1|5.5/G:3.1%#x/H:~x#x/I:2-9#1+2#1-6|1+26/J:4.17/K:2+6-26
i`t-s+u=pau/A:~4+4+2/B:~0000_x#x!C:1500_x!x#x/D:9630_x;x#E:4.3%#x/F:5.0_x#2_1|5.5/G:3.1%#x/H:~x#x/I:2-9#1+2#1-6|1+26/J:4.17/K:2+6-26
ts`u-o+pau=s/A:~5+5+1/B:~1500_x#x#x/C:9630_x!x#x/D:1100_x;x#E:4.3%#x/F:5.0_x#2_1|5.5/G:3.1%#x/H:~x#x/I:2-9#1+2#1-6|1+26/J:4.17/K:2+6-26
u`o-pau+s=D/A:~x+x#x/B:~9630_x#x#x/C:~x!x#x/D:1100_x;x#E:5.0%#x/F:~x_x#x#x|G:3.1%#x/H:~2.9/I:~x#x#x#x#x|J:~x#x/K:2+6-26
o`pau-s+U=b/A:~0+1+3/B:~9630_x#x#x/C:1100_x!x#x/D:1500_x;x#E:5.0%#x/F:3.1_x#01_4|1.17/G:4.0%#x/H:~2.9/I:4-17#2+1#3-4|10+17/J:~x#x/K:2+6-26
pau`S-U+U=b=A:~0+1+3/B:~9630_x#x#x/C:1100_x!x#x/D:1500_x;x#E:5.0%#x/F:3.1_x#01_4|1.17/G:4.0%#x/H:~2.9/I:4-17#2+1#3-4|10+17/J:~x#x/K:2+6-26
s`U-b+e=t/A:~1+2+2/B:~9630_x#x#x/C:1100_x!x#x/D:1500_x;x#E:5.0%#x/F:3.1_x#01_4|1.17/G:4.0%#x/H:~2.9/I:4-17#2+1#3-4|10+17/J:~x#x/K:2+6-26
U`b-e+e=t/A:~1+2+2/B:~9630_x#x#x/C:1100_x!x#x/D:1500_x;x#E:5.0%#x/F:3.1_x#01_4|1.17/G:4.0%#x/H:~2.9/I:4-17#2+1#3-4|10+17/J:~x#x/K:2+6-26
b`e-t+e=j/A:~2+3+1/B:~9630_x#x#x/C:1100_x!x#x/D:1500_x;x#E:5.0%#x/F:3.1_x#01_4|1.17/G:4.0%#x/H:~2.9/I:4-17#2+1#3-4|10+17/J:~x#x/K:2+6-26
e`t-e+j=i/A:~2+3+1/B:~9630_x#x#x/C:1100_x!x#x/D:1500_x;x#E:5.0%#x/F:3.1_x#01_4|1.17/G:4.0%#x/H:~2.9/I:4-17#2+1#3-4|10+17/J:~x#x/K:2+6-26
j`i-b+u=N/A:~1+1+4/B:~1100_x#x#x/C:1500_x!x#x/D:9600_x;x!E:3.1%#x/F:4.0_x#2_3|4.14/G:3.1%#x/H:~2.9/I:4-17#2+1#3-4|10+17/J:~x#x/K:2+6-26
e`j-i-b+u=N/A:~1+1+4/B:~1100_x#x#x/C:1500_x!x#x/D:9600_x;x!E:3.1%#x/F:4.0_x#2_3|4.14/G:3.1%#x/H:~2.9/I:4-17#2+1#3-4|10+17/J:~x#x/K:2+6-26
j`i-b+u=N/A:~2+2+3/B:~1100_x#x#x/C:1500_x!x#x/D:9600_x;x!E:3.1%#x/F:4.0_x#2_3|4.14/G:3.1%#x/H:~2.9/I:4-17#2+1#3-4|10+17/J:~x#x/K:2+6-26
i`b-u+N=N/A:~2+2+3/B:~1100_x#x#x/C:1500_x!x#x/D:9600_x;x!E:3.1%#x/F:4.0_x#2_3|4.14/G:3.1%#x/H:~2.9/I:4-17#2+1#3-4|10+17/J:~x#x/K:2+6-26
b`u-N+u=N/A:~3+3+2/B:~1100_x#x#x/C:1500_x!x#x/D:9600_x;x!E:3.1%#x/F:4.0_x#2_3|4.14/G:3.1%#x/H:~2.9/I:4-17#2+1#3-4|10+17/J:~x#x/K:2+6-26
u`N-n+o=i/A:~4+4+1/B:~1500_x#x#x/C:9600_x!x#x/D:1500_x;x#E:3.1%#x/F:4.0_x#2_3|4.14/G:3.1%#x/H:~2.9/I:4-17#2+1#3-4|10+17/J:~x#x/K:2+6-26
N`n-o+i=o/A:~4+4+1/B:~1500_x#x#x/C:9600_x!x#x/D:1500_x;x#E:3.1%#x/F:4.0_x#2_3|4.14/G:3.1%#x/H:~2.9/I:4-17#2+1#3-4|10+17/J:~x#x/K:2+6-26
n`o-i+o=o/A:~0+1+3/B:~9600_x#x#x/C:1500_x!x#x/D:9630_x;x!E:4.0%#x/F:3.1_x#03_2|8.10/G:7.3%#x/H:~2.9/I:4-17#2+1#3-4|10+17/J:~x#x/K:2+6-26
o`h-o+o=e/A:~1+2+3/B:~9600_x#x#x/C:1500_x!x#x/D:9630_x;x!E:4.0%#x/F:3.1_x#03_2|8.10/G:7.3%#x/H:~2.9/I:4-17#2+1#3-4|10+17/J:~x#x/K:2+6-26
h`o-o+e=e/A:~1+2+3/B:~9600_x#x#x/C:1500_x!x#x/D:9630_x;x!E:4.0%#x/F:3.1_x#03_2|8.10/G:7.3%#x/H:~2.9/I:4-17#2+1#3-4|10+17/J:~x#x/K:2+6-26
o`e-e+e=A:~2+3+1/B:~1500_x#x#x/C:9630_x!x#x/D:3100_02;32_x/E:4.0%#x/F:3.1_x#03_2|8.10/G:7.3%#x/H:~2.9/I:4-17#2+1#3-4|10+17/J:~x#x/K:2+6-26
o`e-e+e=j/A:~2-1+7/B:~9630_x#x#x/C:3100_02;32_x/D:8000_1f;09_x/E:3.1%#x/F:7.3_x#04_1|11.7/G:~x#x#x#x#x/H:~2.9/I:4-17#2+1#3-4|10+17/J:~x#x/K:2+6-26
e`n-e+j=i/A:~2-1+7/B:~9630_x#x#x/C:3100_02;32_x/D:8000_1f;09_x/E:3.1%#x/F:7.3_x#04_1|11.7/G:~x#x#x#x#x/H:~2.9/I:4-17#2+1#3-4|10+17/J:~x#x/K:2+6-26
n`e-j+i=m/A:~1-2+6/B:~9630_x#x#x/C:3100_02;32_x/D:8000_1f;09_x/E:3.1%#x/F:7.3_x#04_1|11.7/G:~x#x#x#x#x/H:~2.9/I:4-17#2+1#3-4|10+17/J:~x#x/K:2+6-26
e`j-i+m=a/A:~1-2+6/B:~9630_x#x#x/C:3100_02;32_x/D:8000_1f;09_x/E:3.1%#x/F:7.3_x#04_1|11.7/G:~x#x#x#x#x/H:~2.9/I:4-17#2+1#3-4|10+17/J:~x#x/K:2+6-26
j`i-m+a=g/A:~0+3+5/B:~9630_x#x#x/C:3100_02;32_x/D:8000_1f;09_x/E:3.1%#x/F:7.3_x#04_1|11.7/G:~x#x#x#x#x/H:~2.9/I:4-17#2+1#3-4|10+17/J:~x#x/K:2+6-26
i`m-a+g=e/A:~0+3+5/B:~9630_x#x#x/C:3100_02;32_x/D:8000_1f;09_x/E:3.1%#x/F:7.3_x#04_1|11.7/G:~x#x#x#x#x/H:~2.9/I:4-17#2+1#3-4|10+17/J:~x#x/K:2+6-26
m`a-g+e=t/A:~1+4+4/B:~9630_x#x#x/C:3100_02;32_x/D:8000_1f;09_x/E:3.1%#x/F:7.3_x#04_1|11.7/G:~x#x#x#x#x/H:~2.9/I:4-17#2+1#3-4|10+17/J:~x#x/K:2+6-26
a`g-e+t=a/A:~1+4+4/B:~9630_x#x#x/C:3100_02;32_x/D:8000_1f;09_x/E:3.1%#x/F:7.3_x#04_1|11.7/G:~x#x#x#x#x/H:~2.9/I:4-17#2+1#3-4|10+17/J:~x#x/K:2+6-26
g`e-t+a=n/A:~2+5+3/B:~3100_02;32_x/C:8000_1f;09_x/D:9300_x;x#E:3.1%#x/F:7.3_x#04_1|11.7/G:~x#x#x#x#x/H:~2.9/I:4-17#2+1#3-4|10+17/J:~x#x/K:2+6-26
e`t-a+n=D/A:~2+5+3/B:~3100_02;32_x/C:8000_1f;09_x/D:9300_x;x#E:3.1%#x/F:7.3_x#04_1|11.7/G:~x#x#x#x#x/H:~2.9/I:4-17#2+1#3-4|10+17/J:~x#x/K:2+6-26
t`a-n+o=d/A:~3+6+2/B:~8000_1#09_x/C:9300_x!x#x/D:8000_03;09_x/E:3.1%#x/F:7.3_x#04_1|11.7/G:~x#x#x#x#x/H:~2.9/I:4-17#2+1#3-4|10+17/J:~x#x/K:2+6-26
a`n-o+d=a/A:~3+6+2/B:~8000_1#09_x/C:9300_x!x#x/D:8000_03;09_x/E:3.1%#x/F:7.3_x#04_1|11.7/G:~x#x#x#x#x/H:~2.9/I:4-17#2+1#3-4|10+17/J:~x#x/K:2+6-26
n`o-d+a=sil/A:~4+7+1/B:~9300_x#x#x/C:8000_03;109_x/D:~x;x#E:3.1%#x/F:7.3_x#04_1|11.7/G:~x#x#x#x#x/H:~2.9/I:4-17#2+1#3-4|10+17/J:~x#x/K:2+6-26
o`d-a=sil=A:~4+7+1/B:~9300_x#x#x/C:8000_03;109_x/D:~x;x#E:3.1%#x/F:7.3_x#04_1|11.7/G:~x#x#x#x#x/H:~2.9/I:4-17#2+1#3-4|10+17/J:~x#x/K:2+6-26
d`a-sil-x=A:~x+x#x/B:~8000_03#09_x/C:~x!x#x/D:~x;x#E:7.3%#x/F:~x_x#x#x|G:~x#x#x#x#x/H:~4.17/I:~x#x#x#x#x|J:~x#x/K:2+6-26

```


2.2 コンテキストクラスタリング時の質問

HMM 音声合成では, 決定木によるコンテキストクラスタリング [11] を用いて分布の共有化を行い, 学習データ量の問題に対応している. また, 決定木によるコンテキスト補完によって, 学習データに存在しないコンテキストに対応するモデルを合成している.

従来, 音声合成において韻律情報を統計モデルで制御する場合には, 数量化 I 類, 回帰木, それらを拡張した CTR などが用いられるが, その意味では, HMM による韻律生成は決定木により制御を行っているため, 回帰木を用いるものに分類される.

以下に, 前章で設計したコンテキスト依存ラベルに対して設計した, 分割条件を列挙する.

(2.2.1) 音素に関する質問

- { 先先行, 先行, 当該, 後続, 後後続 } 音素が有声音
- { 先先行, 先行, 当該, 後続, 後後続 } 音素が母音
- { 先先行, 先行, 当該, 後続, 後後続 } 音素が無声化母音
- { 先先行, 先行, 当該, 後続, 後後続 } 音素が前舌母音
- { 先先行, 先行, 当該, 後続, 後後続 } 音素が後舌母音
- { 先先行, 先行, 当該, 後続, 後後続 } 音素が顎響母音
- { 先先行, 先行, 当該, 後続, 後後続 } 音素が顎反響母音
- { 先先行, 先行, 当該, 後続, 後後続 } 音素が半母音
- { 先先行, 先行, 当該, 後続, 後後続 } 音素が摩擦音
- { 先先行, 先行, 当該, 後続, 後後続 } 音素が破擦音
- { 先先行, 先行, 当該, 後続, 後後続 } 音素が破裂音
- { 先先行, 先行, 当該, 後続, 後後続 } 音素が鼻音
- { 先先行, 先行, 当該, 後続, 後後続 } 音素が唇音
- { 先先行, 先行, 当該, 後続, 後後続 } 音素が歯茎音
- { 先先行, 先行, 当該, 後続, 後後続 } 音素が口蓋音
- { 先先行, 先行, 当該, 後続, 後後続 } 音素が声門音
- { 先先行, 先行, 当該, 後続, 後後続 } 音素が無音
- { 先先行, 先行, 当該, 後続, 後後続 } 音素が有声摩擦音
- { 先先行, 先行, 当該, 後続, 後後続 } 音素が無声摩擦音
- { 先先行, 先行, 当該, 後続, 後後続 } 音素が有声破裂音
- { 先先行, 先行, 当該, 後続, 後後続 } 音素が無声破裂音
- { 先先行, 先行, 当該, 後続, 後後続 } 音素が A
- { 先先行, 先行, 当該, 後続, 後後続 } 音素が E
- { 先先行, 先行, 当該, 後続, 後後続 } 音素が I
- { 先先行, 先行, 当該, 後続, 後後続 } 音素が N
- { 先先行, 先行, 当該, 後続, 後後続 } 音素が O
- { 先先行, 先行, 当該, 後続, 後後続 } 音素が U
- { 先先行, 先行, 当該, 後続, 後後続 } 音素が a
- { 先先行, 先行, 当該, 後続, 後後続 } 音素が b
- { 先先行, 先行, 当該, 後続, 後後続 } 音素が by
- { 先先行, 先行, 当該, 後続, 後後続 } 音素が ch
- { 先先行, 先行, 当該, 後続, 後後続 } 音素が Q
- { 先先行, 先行, 当該, 後続, 後後続 } 音素が d
- { 先先行, 先行, 当該, 後続, 後後続 } 音素が dy
- { 先先行, 先行, 当該, 後続, 後後続 } 音素が e
- { 先先行, 先行, 当該, 後続, 後後続 } 音素が f

{ 先先行, 先行, 当該, 後続, 後後続 } 音素が g
 { 先先行, 先行, 当該, 後続, 後後続 } 音素が gy
 { 先先行, 先行, 当該, 後続, 後後続 } 音素が h
 { 先先行, 先行, 当該, 後続, 後後続 } 音素が hy
 { 先先行, 先行, 当該, 後続, 後後続 } 音素が i
 { 先先行, 先行, 当該, 後続, 後後続 } 音素が j
 { 先先行, 先行, 当該, 後続, 後後続 } 音素が k
 { 先先行, 先行, 当該, 後続, 後後続 } 音素が ky
 { 先先行, 先行, 当該, 後続, 後後続 } 音素が m
 { 先先行, 先行, 当該, 後続, 後後続 } 音素が my
 { 先先行, 先行, 当該, 後続, 後後続 } 音素が n
 { 先先行, 先行, 当該, 後続, 後後続 } 音素が ny
 { 先先行, 先行, 当該, 後続, 後後続 } 音素が o
 { 先先行, 先行, 当該, 後続, 後後続 } 音素が p
 { 先先行, 先行, 当該, 後続, 後後続 } 音素が pau
 { 先先行, 先行, 当該, 後続, 後後続 } 音素が py
 { 先先行, 先行, 当該, 後続, 後後続 } 音素が r
 { 先先行, 先行, 当該, 後続, 後後続 } 音素が ry
 { 先先行, 先行, 当該, 後続, 後後続 } 音素が s
 { 先先行, 先行, 当該, 後続, 後後続 } 音素が sh
 { 先先行, 先行, 当該, 後続, 後後続 } 音素が sil
 { 先先行, 先行, 当該, 後続, 後後続 } 音素が t
 { 先先行, 先行, 当該, 後続, 後後続 } 音素が ts
 { 先先行, 先行, 当該, 後続, 後後続 } 音素が u
 { 先先行, 先行, 当該, 後続, 後後続 } 音素が w
 { 先先行, 先行, 当該, 後続, 後後続 } 音素が y
 { 先先行, 先行, 当該, 後続, 後後続 } 音素が z

(2.2.2) モーラに関する質問

当該モーラと当該韻律語のアクセント核との差分 = N
 当該モーラと当該韻律語のアクセント核との差分 $\leq N$
 当該モーラの韻律語内での { 前, 後 } から数えた音節位置 = N
 当該モーラの韻律語内での { 前, 後 } から数えた音節位置 $\leq N$

但し, N は 0 以上.

(2.2.3) 形態素に関する質問

{ 先行, 当該, 後続 } 形態素の品詞が自立語
 { 先行, 当該, 後続 } 形態素の品詞が付属語
 { 先行, 当該, 後続 } 形態素の品詞が活用あり
 { 先行, 当該, 後続 } 形態素の品詞が性質を表す
 { 先行, 当該, 後続 } 形態素の品詞が活用なし
 { 先行, 当該, 後続 } 形態素の品詞が体言
 { 先行, 当該, 後続 } 形態素の品詞が修飾語
 { 先行, 当該, 後続 } 形態素の品詞が内容語
 { 先行, 当該, 後続 } 形態素の品詞が連体詞
 { 先行, 当該, 後続 } 形態素の品詞が名詞
 { 先行, 当該, 後続 } 形態素の品詞が名詞-副詞可能

- { 先行, 当該, 後続 } 形態素の品詞が名詞-数詞
- { 先行, 当該, 後続 } 形態素の品詞が名詞-固有名詞
- { 先行, 当該, 後続 } 形態素の品詞が名詞-固有名詞-地域
- { 先行, 当該, 後続 } 形態素の品詞が名詞-固有名詞-地域-国
- { 先行, 当該, 後続 } 形態素の品詞が名詞-固有名詞-地域-一般
- { 先行, 当該, 後続 } 形態素の品詞が名詞-固有名詞-組織
- { 先行, 当該, 後続 } 形態素の品詞が名詞-固有名詞-人名
- { 先行, 当該, 後続 } 形態素の品詞が名詞-固有名詞-人名-名
- { 先行, 当該, 後続 } 形態素の品詞が名詞-固有名詞-人名-姓
- { 先行, 当該, 後続 } 形態素の品詞が名詞-固有名詞-人名-一般
- { 先行, 当該, 後続 } 形態素の品詞が名詞-固有名詞-一般
- { 先行, 当該, 後続 } 形態素の品詞が名詞-形状詞可能
- { 先行, 当該, 後続 } 形態素の品詞が名詞-一般
- { 先行, 当該, 後続 } 形態素の品詞が名詞-サ変可能
- { 先行, 当該, 後続 } 形態素の品詞が副詞
- { 先行, 当該, 後続 } 形態素の品詞が動詞
- { 先行, 当該, 後続 } 形態素の品詞が動詞-非自立
- { 先行, 当該, 後続 } 形態素の品詞が動詞-自立
- { 先行, 当該, 後続 } 形態素の品詞が代名詞
- { 先行, 当該, 後続 } 形態素の品詞が代名詞-縮訳
- { 先行, 当該, 後続 } 形態素の品詞が代名詞-一般
- { 先行, 当該, 後続 } 形態素の品詞が接尾辞
- { 先行, 当該, 後続 } 形態素の品詞が接尾辞-名詞的
- { 先行, 当該, 後続 } 形態素の品詞が接尾辞-名詞的-副詞可能
- { 先行, 当該, 後続 } 形態素の品詞が接尾辞-名詞的-地域
- { 先行, 当該, 後続 } 形態素の品詞が接尾辞-名詞的-人名
- { 先行, 当該, 後続 } 形態素の品詞が接尾辞-名詞的-一般
- { 先行, 当該, 後続 } 形態素の品詞が接尾辞-名詞的-サ変可能
- { 先行, 当該, 後続 } 形態素の品詞が接尾辞-動詞的
- { 先行, 当該, 後続 } 形態素の品詞が接尾辞-助数詞
- { 先行, 当該, 後続 } 形態素の品詞が接尾辞-形容詞的
- { 先行, 当該, 後続 } 形態素の品詞が接尾辞-形状詞的
- { 先行, 当該, 後続 } 形態素の品詞が接頭辞
- { 先行, 当該, 後続 } 形態素の品詞が接頭辞-名詞接続
- { 先行, 当該, 後続 } 形態素の品詞が接頭辞-動詞接続
- { 先行, 当該, 後続 } 形態素の品詞が接頭辞-数接続
- { 先行, 当該, 後続 } 形態素の品詞が接頭辞-形容詞接続
- { 先行, 当該, 後続 } 形態素の品詞が接続詞
- { 先行, 当該, 後続 } 形態素の品詞が助動詞
- { 先行, 当該, 後続 } 形態素の品詞が助詞
- { 先行, 当該, 後続 } 形態素の品詞が助詞-副助詞
- { 先行, 当該, 後続 } 形態素の品詞が助詞-副助詞-並列
- { 先行, 当該, 後続 } 形態素の品詞が助詞-副助詞-一般
- { 先行, 当該, 後続 } 形態素の品詞が助詞-特殊
- { 先行, 当該, 後続 } 形態素の品詞が助詞-接続助詞
- { 先行, 当該, 後続 } 形態素の品詞が助詞-準体助詞
- { 先行, 当該, 後続 } 形態素の品詞が助詞-終助詞
- { 先行, 当該, 後続 } 形態素の品詞が助詞-係助詞
- { 先行, 当該, 後続 } 形態素の品詞が助詞-格助詞
- { 先行, 当該, 後続 } 形態素の品詞が助詞-格助詞-連体化
- { 先行, 当該, 後続 } 形態素の品詞が助詞-格助詞-副詞力
- { 先行, 当該, 後続 } 形態素の品詞が助詞-格助詞-引用

- { 先行, 当該, 後続 } 形態素の品詞が助詞-格助詞-一般
- { 先行, 当該, 後続 } 形態素の品詞が形容詞
- { 先行, 当該, 後続 } 形態素の品詞が形容詞-非自立
- { 先行, 当該, 後続 } 形態素の品詞が形容詞-自立
- { 先行, 当該, 後続 } 形態素の品詞が形状詞
- { 先行, 当該, 後続 } 形態素の品詞が形状詞-非自立
- { 先行, 当該, 後続 } 形態素の品詞が形状詞-自立
- { 先行, 当該, 後続 } 形態素の品詞が感動詞
- { 先行, 当該, 後続 } 形態素の品詞がフィラー
- { 先行, 当該, 後続 } 形態素の品詞がその他
- { 先行, 当該, 後続 } 形態素の品詞がその他-記号
- { 先行, 当該, 後続 } 形態素の品詞がその他-記号-読点
- { 先行, 当該, 後続 } 形態素の品詞がその他-記号-数字
- { 先行, 当該, 後続 } 形態素の品詞がその他-記号-空白
- { 先行, 当該, 後続 } 形態素の品詞がその他-記号-句点
- { 先行, 当該, 後続 } 形態素の品詞がその他-記号-括弧閉
- { 先行, 当該, 後続 } 形態素の品詞がその他-記号-括弧開
- { 先行, 当該, 後続 } 形態素の品詞がその他-記号-一般
- { 先行, 当該, 後続 } 形態素の品詞がその他-記号-アルファベット

- { 先行, 当該, 後続 } 形態素の活用型がア行下二段-一般
- { 先行, 当該, 後続 } 形態素の活用型がア行下二段-得
- { 先行, 当該, 後続 } 形態素の活用型がカ行下二段
- { 先行, 当該, 後続 } 形態素の活用型がカ行五段-イ音便
- { 先行, 当該, 後続 } 形態素の活用型がカ行五段-促音便
- { 先行, 当該, 後続 } 形態素の活用型がカ行五段-特殊
- { 先行, 当該, 後続 } 形態素の活用型がカ行四段
- { 先行, 当該, 後続 } 形態素の活用型がカ行変格-クル
- { 先行, 当該, 後続 } 形態素の活用型がカ行変格-来ル
- { 先行, 当該, 後続 } 形態素の活用型がガ行下二段
- { 先行, 当該, 後続 } 形態素の活用型がガ行五段
- { 先行, 当該, 後続 } 形態素の活用型がガ行四段
- { 先行, 当該, 後続 } 形態素の活用型がサ行下二段
- { 先行, 当該, 後続 } 形態素の活用型がサ行五段
- { 先行, 当該, 後続 } 形態素の活用型がサ行四段
- { 先行, 当該, 後続 } 形態素の活用型がサ行変格-ースル
- { 先行, 当該, 後続 } 形態素の活用型がサ行変格-ーズル
- { 先行, 当該, 後続 } 形態素の活用型がサ行変格-スル
- { 先行, 当該, 後続 } 形態素の活用型がザ行下二段
- { 先行, 当該, 後続 } 形態素の活用型がタ行下二段
- { 先行, 当該, 後続 } 形態素の活用型がタ行五段
- { 先行, 当該, 後続 } 形態素の活用型がタ行四段
- { 先行, 当該, 後続 } 形態素の活用型がダ行下二段
- { 先行, 当該, 後続 } 形態素の活用型がダ行上二段
- { 先行, 当該, 後続 } 形態素の活用型がナ行下二段
- { 先行, 当該, 後続 } 形態素の活用型がナ行五段
- { 先行, 当該, 後続 } 形態素の活用型がハ行下二段
- { 先行, 当該, 後続 } 形態素の活用型がハ行四段
- { 先行, 当該, 後続 } 形態素の活用型がハ行上二段
- { 先行, 当該, 後続 } 形態素の活用型がバ行下二段
- { 先行, 当該, 後続 } 形態素の活用型がバ行五段
- { 先行, 当該, 後続 } 形態素の活用型がバ行四段
- { 先行, 当該, 後続 } 形態素の活用型がマ行下二段

- { 先行, 当該, 後続 } 形態素の活用型がマ行五段
- { 先行, 当該, 後続 } 形態素の活用型がマ行四段
- { 先行, 当該, 後続 } 形態素の活用型がヤ行下二段
- { 先行, 当該, 後続 } 形態素の活用型がラ行下二段
- { 先行, 当該, 後続 } 形態素の活用型がラ行五段-一般
- { 先行, 当該, 後続 } 形態素の活用型がラ行五段-特殊
- { 先行, 当該, 後続 } 形態素の活用型がラ行四段
- { 先行, 当該, 後続 } 形態素の活用型がラ行変格
- { 先行, 当該, 後続 } 形態素の活用型がワ行下二段
- { 先行, 当該, 後続 } 形態素の活用型がワ行五段-一般
- { 先行, 当該, 後続 } 形態素の活用型がワ行五段-特殊
- { 先行, 当該, 後続 } 形態素の活用型が一段-ウル
- { 先行, 当該, 後続 } 形態素の活用型が一段-クレル
- { 先行, 当該, 後続 } 形態素の活用型が一段-一般
- { 先行, 当該, 後続 } 形態素の活用型が形容詞-アウオ男
- { 先行, 当該, 後続 } 形態素の活用型が形容詞-イ段
- { 先行, 当該, 後続 } 形態素の活用型が助動詞-キ
- { 先行, 当該, 後続 } 形態素の活用型が助動詞-ケム
- { 先行, 当該, 後続 } 形態素の活用型が助動詞-ケリ
- { 先行, 当該, 後続 } 形態素の活用型が助動詞-ゴトシ
- { 先行, 当該, 後続 } 形態素の活用型が助動詞-ジャ
- { 先行, 当該, 後続 } 形態素の活用型が助動詞-ズ
- { 先行, 当該, 後続 } 形態素の活用型が助動詞-タ
- { 先行, 当該, 後続 } 形態素の活用型が助動詞-タイ
- { 先行, 当該, 後続 } 形態素の活用型が助動詞-タリ
- { 先行, 当該, 後続 } 形態素の活用型が助動詞-ダ
- { 先行, 当該, 後続 } 形態素の活用型が助動詞-ツ
- { 先行, 当該, 後続 } 形態素の活用型が助動詞-デス
- { 先行, 当該, 後続 } 形態素の活用型が助動詞-ナイ
- { 先行, 当該, 後続 } 形態素の活用型が助動詞-ナリ
- { 先行, 当該, 後続 } 形態素の活用型が助動詞-ベシ
- { 先行, 当該, 後続 } 形態素の活用型が助動詞-マジ
- { 先行, 当該, 後続 } 形態素の活用型が助動詞-マス
- { 先行, 当該, 後続 } 形態素の活用型が助動詞-ヤ
- { 先行, 当該, 後続 } 形態素の活用型が助動詞-ラシイ
- { 先行, 当該, 後続 } 形態素の活用型が助動詞-リ
- { 先行, 当該, 後続 } 形態素の活用型が助動詞-ル
- { 先行, 当該, 後続 } 形態素の活用型が不変化型

- { 先行, 当該, 後続 } 形態素の活用形が仮定形
- { 先行, 当該, 後続 } 形態素の活用形が仮定形-一般
- { 先行, 当該, 後続 } 形態素の活用形が仮定形-強縮約
- { 先行, 当該, 後続 } 形態素の活用形が仮定形-口語表記
- { 先行, 当該, 後続 } 形態素の活用形が仮定形-縮訳
- { 先行, 当該, 後続 } 形態素の活用形が仮定形-特殊
- { 先行, 当該, 後続 } 形態素の活用形が仮定形-文語

- { 先行, 当該, 後続 } 形態素の活用形が基本形
- { 先行, 当該, 後続 } 形態素の活用形が基本形-工音便
- { 先行, 当該, 後続 } 形態素の活用形が基本形-一般
- { 先行, 当該, 後続 } 形態素の活用形が基本形-口語表記
- { 先行, 当該, 後続 } 形態素の活用形が基本形-特殊
- { 先行, 当該, 後続 } 形態素の活用形が基本形-文語

- { 先行, 当該, 後続 } 形態素の活用形が語幹

- { 先行, 当該, 後続 } 形態素の活用形が未然形
- { 先行, 当該, 後続 } 形態素の活用形が未然形-ウ接続
- { 先行, 当該, 後続 } 形態素の活用形が未然形-ヌ接続
- { 先行, 当該, 後続 } 形態素の活用形が未然形-レル接続
- { 先行, 当該, 後続 } 形態素の活用形が未然形-一般
- { 先行, 当該, 後続 } 形態素の活用形が未然形-文語
- { 先行, 当該, 後続 } 形態素の活用形が未然形-撥音便
- { 先行, 当該, 後続 } 形態素の活用形が命令形
- { 先行, 当該, 後続 } 形態素の活用形が命令形-イ
- { 先行, 当該, 後続 } 形態素の活用形が命令形-エ
- { 先行, 当該, 後続 } 形態素の活用形が命令形-ヨ
- { 先行, 当該, 後続 } 形態素の活用形が命令形-ロ
- { 先行, 当該, 後続 } 形態素の活用形が連体形
- { 先行, 当該, 後続 } 形態素の活用形が連体形-一般
- { 先行, 当該, 後続 } 形態素の活用形が連体形-口語表記
- { 先行, 当該, 後続 } 形態素の活用形が連体形-特殊
- { 先行, 当該, 後続 } 形態素の活用形が連体形-文語
- { 先行, 当該, 後続 } 形態素の活用形が連体形-撥音便
- { 先行, 当該, 後続 } 形態素の活用形が連用形
- { 先行, 当該, 後続 } 形態素の活用形が連用形-イ音便
- { 先行, 当該, 後続 } 形態素の活用形が連用形-ウ音便
- { 先行, 当該, 後続 } 形態素の活用形が連用形-一般
- { 先行, 当該, 後続 } 形態素の活用形が連用形-促音便
- { 先行, 当該, 後続 } 形態素の活用形が連用形-特殊
- { 先行, 当該, 後続 } 形態素の活用形が連用形-文語
- { 先行, 当該, 後続 } 形態素の活用形が連用形-撥音便
- { 先行形態素と当該形態素, 当該形態素と後続形態素 } の境界種別が引用境界
- { 先行形態素と当該形態素, 当該形態素と後続形態素 } の境界種別が係り受け
- { 先行形態素と当該形態素, 当該形態素と後続形態素 } の境界種別が節境界
- { 先行形態素と当該形態素, 当該形態素と後続形態素 } の境界種別が複合語内
- { 先行形態素と当該形態素, 当該形態素と後続形態素 } の境界種別が並列
- { 先行形態素と当該形態素, 当該形態素と後続形態素 } の境界種別が並列始め

(2.2.4) 韻律語に関する質問

- { 先行, 当該, 後続 } 韻律語のモーラ単位で数えた長さ = N
- { 先行, 当該, 後続 } 韻律語のモーラ単位で数えた長さ $\leq N$
- { 先行, 当該, 後続 } 韻律語のアクセント型 = N
- { 先行, 当該, 後続 } 韻律語のアクセント型 $\leq N$
- { 先行, 当該, 後続 } 韻律語の属性 (強調, 疑問等)
- 当該韻律語の当該呼気段落内での { 前, 後 } から { モーラ, 韻律語 } 単位で数えた位置 = N
- 当該韻律語の当該呼気段落内での { 前, 後 } から { モーラ, 韻律語 } 単位で数えた位置 $\leq N$
- { 先行韻律語と当該韻律語, 当該韻律語と後続韻律語 } 間のポーズの有無

但し, N は 0 以上

(2.2.5) 呼気段落に関する質問

{ 先行, 当該, 後続 } 呼気段落の { モーラ, 韻律語 } 単位で数えた長さ = N

{ 先行, 当該, 後続 } 呼気段落の { モーラ, 韻律語 } 単位で数えた長さ $\leq N$

当該呼気段落の文章内での { 前, 後 } から { モーラ, 韻律語, 呼気段落 } 単位で数えた位置 = N

当該呼気段落の文章内での { 前, 後 } から { モーラ, 韻律語, 呼気段落 } 単位で数えた位置 $\leq N$

但し, N は 0 以上

(2.2.6) 文章に関する質問

{ モーラ, 韻律語 } 単位で数えた文章の長さ = N

{ モーラ, 韻律語 } 単位で数えた文章の長さ $\leq N$

但し, N は 0 以上

3 HMM による中国語韻律合成のためのコンテキスト依存ラベルの設計

3.1 導入したコンテキスト

中国語における HMM 音声合成システムでは, 1 つの HMM は 1 つの半音節に対応している. コンテキスト (変動要因) としては半音節を最小単位とし, 音節, 形態素, 呼気段落, 文章の順で階層的にコンテキスト情報を取り扱う. 各階層におけるコンテキスト情報を以下に示す.

(3.1.1) 半音節に関するコンテキスト

半音節セットを, 以下のように定義する.

表 1: 中国語 HMM 音声合成における半音節 (声母:子音に対応)

	塞音		塞擦音		擦音		鼻音	辺音
	清音		清音		清音	濁音	濁音	濁音
	不送気	送気	不送気	送気				
双唇音	b	p					m	
唇齒音					f			
舌尖前音			z	c	s			
舌尖中音	d	t					n	l
舌尖后音			zh	ch	sh	r		
舌面音			j	q	x			
舌根音	g	k			h			

表 2: 中国語 HMM 音声合成における半音節 (韻母:母音に対応)

	開口呼	齊齒呼	合口呼	撮口呼
単韻母	a, o, e, er, re	i, ic, ich	u	v
複韻母	ai, ei, ao, ou	ia, ie	ua, uo	ve
結合韻母		iao, iu	uai, ui	
声随韻母	an, en, ang, eng	ian, in, iang, ing	uan, un, uang, ong	van, vn, iong

文頭・文末無音	sil
文中小休止	pau

半音節階層における変動要因として, 以下に述べるものを導入する.

- p_{L2} 先先行半音節
- p_{L1} 先行半音節
- p_C 当該半音節
- p_{R1} 後続半音節
- p_{R2} 後後続半音節

(3.1.2) 音節に関するコンテキスト

中国語は、音節単位で韻律が変化する声調言語である。よって、音節階層における変動要因に声調に関する情報を含める。

標準中国語に声調は4つの種類がある。これを四声という。また、中国語は言葉の調子、または音節のつなぎ具合によって、声調が本来の基本声調から変わることがある。この声調の変化を変調または調変という。

今回は、基本の声調を TONE, 調変により変化し、実際に発声された声調を NEW TONE とし、ラベル中に組み込むことにする。TONE 及び NEW TONE には1から4の値が入り、それぞれ第一声、…、第四声に対応する。

音節階層における変動要因として、以下に述べるものを導入する。

- a_1 先行音節の TONE
- a_2 先行音節の NEW TONE
- a_3 先行音節と当該音節間のポーズの有無
- b_1 当該音節の TONE
- b_2 当該音節の NEW TONE
- b_3 前からカウントした形態素内での当該音節の位置 (音節単位)
- b_4 後からカウントした形態素内での当該音節の位置 (音節単位)
- c_1 後続音節の TONE
- c_2 後続音節の NEW TONE
- c_3 当該音節と後続音節間のポーズの有無

(3.1.3) 形態素に関するコンテキスト

形態素は単語に対応するものである。中国語は日本語と異なり、活用が存在しない。形態素の情報としては、単語の品詞が重要と思われる。以下に、中国語における品詞分類を示す。

形態素階層における変動要因として、以下に述べるものを導入する。

表 3: 品詞の分類

実詞	体詞	数詞 量詞 名詞	CD, OD, DT MC NN, NT, NR, PN
	述詞	形容詞 形容動詞 動詞	JJ, VA VA VC, VE, VV
虚詞	前置詞		BA, BE, CC, PP
	不変化詞		DEC, DEG, DER, DEV, MSP, SP, AS, ETC, LC
その他			ON, IJ, FW, PU

- d_1 先行形態素の品詞
- d_2 先行形態素の長さ (音節単位)
- d_3 先行形態素と当該形態素間のポーズの有無
- e_1 当該形態素の品詞
- e_2 当該形態素の長さ (音節単位)
- e_3 前からカウントした呼気段落内での当該形態素の位置 (音節単位)
- e_4 後からカウントした呼気段落内での当該形態素の位置 (音節単位)
- e_5 前からカウントした呼気段落内での当該形態素の位置 (形態素単位)
- e_6 後からカウントした呼気段落内での当該形態素の位置 (形態素単位)
- f_1 後続形態素の品詞
- f_2 後続形態素の長さ (音節単位)
- f_3 当該形態素と後続形態素間のポーズの有無

(3.1.4) 呼気段落に関するコンテキスト

中国語では、疑問文による韻律変化が呼気段落単位で生じることが多い。そこで、呼気段落階層における変動要因として、疑問であるかの情報を加える。

呼気段落階層における変動要因として、以下に述べるものを導入する。

- g_1 先行呼気段落の長さ (音節単位)
- g_2 先行呼気段落の長さ (形態素単位)
- g_3 先行呼気段落の疑問の有無 (0:なし, 1:あり)
- h_1 当該呼気段落の長さ (音節単位)
- h_2 当該呼気段落の長さ (形態素単位)
- h_3 当該呼気段落の疑問の有無 (0:なし, 1:あり)
- h_4 前からカウントした文章内での当該呼気段落の位置 (音節単位)
- h_5 後からカウントした文章内での当該呼気段落の位置 (音節単位)
- h_6 前からカウントした文章内での当該呼気段落の位置 (形態素単位)
- h_7 後からカウントした文章内での当該呼気段落の位置 (形態素単位)
- h_8 前からカウントした文章内での当該呼気段落の位置 (呼気段落単位)
- h_9 後からカウントした文章内での当該呼気段落の位置 (呼気段落単位)
- i_1 後続呼気段落の長さ (音節単位)
- i_2 後続呼気段落の長さ (形態素単位)
- i_3 後続呼気段落の疑問の有無 (0:なし, 1:あり)

(3.1.5) 文章に関するコンテキスト

文章階層における変動要因として、以下に述べるものを導入する。

- j_1 文章の長さ (音節単位)
- j_2 文章の長さ (形態素単位)
- j_3 文章の長さ (呼気段落単位)
- j_4 文型 (0:平叙文, 1:疑問文)

(3.1.6) コンテキスト依存ラベルの仕様

コンテキスト依存ラベルの仕様を以下のように定義する。

$$PL2 \hat{=} PL1 - PC + PR1 = PR2$$

$/A: a_1 a_2 a_3 / B: b_1 b_2 @ b_3 b_4 / C: c_1 c_2 c_3$
 $/D: d_1 d_2 \# d_3 / E: e_1 e_2 @ e_3 e_4 | e_5 e_6 / F: f_1 f_2 \# f_3$
 $/G: g_1 g_2 \% g_3 / H: h_1 h_2 \% h_3 @ h_4 h_5 \& h_6 h_7 | h_8 h_9 / I: i_1 i_2 \% i_3$
 $/J: j_1 j_2 j_3 ; j_4$

上記の“アルファベット+数字”の変数はそれぞれ、3.1のコンテキスト情報に対応する。コンテキスト依存ラベルの例を以下に示す。

他说，中泰两国议会关系友好，彼此间的往来也十分频繁。

```
X`X-sil-t+=a/A:X_X/B:X_X/C:1.1_0/D:X_X/E:X_X/F:PN_1#0/G:X_X/H:X_X/I:2.250/J:23.16.4;0
X`sil-t+=sh/A:X_X/B:1.101_1/C:1.1_0/D:X_X/E:PN_101_2|1.2/F:VV_1#0/G:X_X/H:2.25001_23&1.16|1.4/I:10.750/J:23.16.4;0
sil`-t+=shuo/A:X_X/B:1.101_1/C:1.1_0/D:X_X/E:PN_101_2|1.2/F:VV_1#0/G:X_X/H:2.25001_23&1.16|1.4/I:10.750/J:23.16.4;0
t`a-sh+uo=pau/A:1.1_0/B:1.101_1/C:1.1_0/D:PN_1#0/E:VV_102_1|2.1/F:NR_1#1/G:X_X/H:2.25001_23&1.16|1.4/I:10.750/J:23.16.4;0
a`sh-uo=pau=zh/A:1.1_0/B:1.101_1/C:1.1_0/D:PN_1#0/E:VV_102_1|2.1/F:NR_1#1/G:X_X/H:2.25001_23&1.16|1.4/I:10.750/J:23.16.4;0
sh`uo-pau+zhong/A:1.1_0/B:X_X/C:1.1_0/D:VV_1#0/E:X_X/F:NR_1#0/G:2.250/H:X_X/I:10.750/J:23.16.4;0
uo`pau-zh+ong/A:1.1_0/B:1.101_1/C:4.4_0/D:VV_1#1/E:NR_101_10|1.7/F:NR_1#0/G:2.250/H:10.75003_21&3_14|2.3/I:6.450/J:23.16.4;0
pau`zh-ong+t-ai/A:1.1_0/B:1.101_1/C:4.4_0/D:VV_1#1/E:NR_101_10|1.7/F:NR_1#0/G:2.250/H:10.75003_21&3_14|2.3/I:6.450/J:23.16.4;0
zh`ong-t+ai=A:1.1_0/B:4.401_1/C:3.3_0/D:NR_1#0/E:NR_102_9|2.6/F:CD_1#0/G:2.250/H:10.75003_21&3_14|2.3/I:6.450/J:23.16.4;0
ong`-t+ai=iang/A:1.1_0/B:4.401_1/C:3.3_0/D:NR_1#0/E:NR_102_9|2.6/F:CD_1#0/G:2.250/H:10.75003_21&3_14|2.3/I:6.450/J:23.16.4;0
t`ai-l+iang=A:4.4_0/B:3.301_1/C:2.2_0/D:NR_1#0/E:CD_103_8|3.5/F:NN_1#0/G:2.250/H:10.75003_21&3_14|2.3/I:6.450/J:23.16.4;0
ai`l-iang+uo/A:4.4_0/B:3.301_1/C:2.2_0/D:NR_1#0/E:CD_103_8|3.5/F:NN_1#0/G:2.250/H:10.75003_21&3_14|2.3/I:6.450/J:23.16.4;0
l`iang-g+uo=A:3.3_0/B:2.201_1/C:4.4_0/D:CD_1#0/E:NN_104_7|4.4/F:NN_2#0/G:2.250/H:10.75003_21&3_14|2.3/I:6.450/J:23.16.4;0
iang`g-uo+i=h/A:3.3_0/B:2.201_1/C:4.4_0/D:CD_1#0/E:NN_104_7|4.4/F:NN_2#0/G:2.250/H:10.75003_21&3_14|2.3/I:6.450/J:23.16.4;0
uo`i-h+ui=A:2.2_0/B:4.401_2/C:4.4_0/D:NN_1#0/E:NN_205_6|5.3/F:NN_2#0/G:2.250/H:10.75003_21&3_14|2.3/I:6.450/J:23.16.4;0
ui`i-h+ui=g/A:4.4_0/B:4.402_1/C:1.1_0/D:NN_1#0/E:NN_205_6|5.3/F:NN_2#0/G:2.250/H:10.75003_21&3_14|2.3/I:6.450/J:23.16.4;0
i`h-ui+g=uan/A:4.4_0/B:4.402_1/C:1.1_0/D:NN_1#0/E:NN_205_6|5.3/F:NN_2#0/G:2.250/H:10.75003_21&3_14|2.3/I:6.450/J:23.16.4;0
h`ui-g+uan=A:4.4_0/B:1.101_2/C:4.4_0/D:NN_2#0/E:NN_206_5|6.2/F:VA_2#0/G:2.250/H:10.75003_21&3_14|2.3/I:6.450/J:23.16.4;0
ui`g-uan+x=i/A:4.4_0/B:1.101_2/C:4.4_0/D:NN_2#0/E:NN_206_5|6.2/F:VA_2#0/G:2.250/H:10.75003_21&3_14|2.3/I:6.450/J:23.16.4;0
g`uan-x+i=iu/A:1.1_0/B:4.402_1/C:3.2_0/D:NN_2#0/E:NN_206_5|6.2/F:VA_2#0/G:2.250/H:10.75003_21&3_14|2.3/I:6.450/J:23.16.4;0
uan`x-i+iu=h/A:1.1_0/B:4.402_1/C:3.2_0/D:NN_2#0/E:NN_206_5|6.2/F:VA_2#0/G:2.250/H:10.75003_21&3_14|2.3/I:6.450/J:23.16.4;0
x`i-uh+ao/A:4.4_0/B:3.302_1/C:3.3_0/D:NN_2#0/E:VA_207_4|7.1/F:PN_2#1/G:2.250/H:10.75003_21&3_14|2.3/I:6.450/J:23.16.4;0
i`iu-h+ao=pau/A:3.2_0/B:3.302_1/C:3.2_0/D:NN_2#0/E:VA_207_4|7.1/F:PN_2#1/G:2.250/H:10.75003_21&3_14|2.3/I:6.450/J:23.16.4;0
iu`h-ao+pau=h/A:3.2_0/B:3.302_1/C:3.2_0/D:NN_2#0/E:VA_207_4|7.1/F:PN_2#1/G:2.250/H:10.75003_21&3_14|2.3/I:6.450/J:23.16.4;0
h`ao-pau+i=A:3.3_0/B:X_X/C:3.2_0/D:VA_2#0/E:X_X/F:PN_2#0/G:10.750/H:6.450013_11&10_7|3.2/I:5.350/J:23.16.4;0
ao`pau-b+i=c/A:3.3_0/B:3.201_2/C:3.3_0/D:VA_2#1/E:PN_201_6|1.4/F:LC_1#0/G:10.750/H:6.450013_11&10_7|3.2/I:5.350/J:23.16.4;0
pau`b-i+c=i/A:3.3_0/B:3.201_2/C:3.3_0/D:VA_2#1/E:PN_201_6|1.4/F:LC_1#0/G:10.750/H:6.450013_11&10_7|3.2/I:5.350/J:23.16.4;0
b`i-c+i=j/A:3.2_0/B:3.302_1/C:1.1_0/D:VA_2#1/E:PN_201_6|1.4/F:LC_1#0/G:10.750/H:6.450013_11&10_7|3.2/I:5.350/J:23.16.4;0
i`c-i+c=j=ian/A:3.2_0/B:3.302_1/C:1.1_0/D:VA_2#1/E:PN_201_6|1.4/F:LC_1#0/G:10.750/H:6.450013_11&10_7|3.2/I:5.350/J:23.16.4;0
c`i-c-j=ian=d/A:3.3_0/B:1.101_1/C:0.0_0/D:PN_2#0/E:LC_102_5|2.3/F:DEG_1#0/G:10.750/H:6.450013_11&10_7|3.2/I:5.350/J:23.16.4;0
i`c-j-ian=d=e/A:3.3_0/B:1.101_1/C:0.0_0/D:PN_2#0/E:LC_102_5|2.3/F:DEG_1#0/G:10.750/H:6.450013_11&10_7|3.2/I:5.350/J:23.16.4;0
j`ian-d=e=uang/A:1.1_0/B:0.001_1/C:3.3_0/D:LC_1#0/E:DEG_103_4|3.2/F:NM_2#0/G:10.750/H:6.450013_11&10_7|3.2/I:5.350/J:23.16.4;0
ian`d-e=uang=l/A:1.1_0/B:0.001_1/C:3.3_0/D:LC_1#0/E:DEG_103_4|3.2/F:NM_2#0/G:10.750/H:6.450013_11&10_7|3.2/I:5.350/J:23.16.4;0
d`e-uang=l=ai/A:0.0_0/B:3.301_2/C:2.2_0/D:DEG_1#0/E:NN_204_3|4.1/F:AD_1#1/G:10.750/H:6.450013_11&10_7|3.2/I:5.350/J:23.16.4;0
e`uang-l=ai=pau/A:3.3_0/B:2.202_1/C:3.3_0/D:DEG_1#0/E:NN_204_3|4.1/F:AD_1#1/G:10.750/H:6.450013_11&10_7|3.2/I:5.350/J:23.16.4;0
uang`l-ai+pau=i/A:3.3_0/B:2.202_1/C:3.3_0/D:DEG_1#0/E:NN_204_3|4.1/F:AD_1#1/G:10.750/H:6.450013_11&10_7|3.2/I:5.350/J:23.16.4;0
i`ai-pau+i=sh/A:2.2_0/B:X_X/C:3.3_0/D:NN_2#0/E:X_X/F:AD_1#0/G:6.450/H:X_X/I:5.350/J:23.16.4;0
ai`pau-i+sh=i/A:2.2_0/B:3.301_1/C:2.2_0/D:NN_2#1/E:AD_101_5|1.3/F:AD_2#0/G:6.450/H:5.350019_5&14_3|4.1/I:X_X/J:23.16.4;0
pau`ie-sh+i=ch/A:3.3_0/B:2.201_2/C:1.1_0/D:AD_1#0/E:AD_202_4|2.2/F:VA_2#0/G:6.450/H:5.350019_5&14_3|4.1/I:X_X/J:23.16.4;0
ie`sh-i+ch=f/A:3.3_0/B:2.201_2/C:1.1_0/D:AD_1#0/E:AD_202_4|2.2/F:VA_2#0/G:6.450/H:5.350019_5&14_3|4.1/I:X_X/J:23.16.4;0
sh`i-ch+f=en/A:2.2_0/B:1.102_1/C:2.2_0/D:AD_1#0/E:AD_202_4|2.2/F:VA_2#0/G:6.450/H:5.350019_5&14_3|4.1/I:X_X/J:23.16.4;0
ich`f-en+p=in/A:2.2_0/B:1.102_1/C:2.2_0/D:AD_1#0/E:AD_202_4|2.2/F:VA_2#0/G:6.450/H:5.350019_5&14_3|4.1/I:X_X/J:23.16.4;0
f`en-p+in=A:1.1_0/B:2.201_2/C:2.2_0/D:AD_2#0/E:VA_203_3|3.1/F:X_X/G:6.450/H:5.350019_5&14_3|4.1/I:X_X/J:23.16.4;0
en`p-in+an/A:1.1_0/B:2.201_2/C:2.2_0/D:AD_2#0/E:VA_203_3|3.1/F:X_X/G:6.450/H:5.350019_5&14_3|4.1/I:X_X/J:23.16.4;0
in`f-an=sil/A:2.2_0/B:2.202_1/C:X_X/D:AD_2#0/E:VA_203_3|3.1/F:X_X/G:6.450/H:5.350019_5&14_3|4.1/I:X_X/J:23.16.4;0
sil`f-an=sil=X/A:2.2_0/B:2.202_1/C:X_X/D:AD_2#0/E:VA_203_3|3.1/F:X_X/G:6.450/H:5.350019_5&14_3|4.1/I:X_X/J:23.16.4;0
f`an-sil=X=X/A:2.2_0/B:X_X/C:X_X/D:VA_2#0/E:X_X/F:X_X/G:5.350/H:X_X/I:X_X/J:23.16.4;0
```

3.2 クラスタリング時に用いる質問

決定木によるコンテキストクラスタリングを行う際に用いる分割条件として、以下に示すような質問を用いた

(3.2.1) 半音節に関する質問

- { 先先行, 先行, 当該, 後続, 後後続 } 半音節が有声音
- { 先先行, 先行, 当該, 後続, 後後続 } 半音節が母音
- { 先先行, 先行, 当該, 後続, 後後続 } 半音節が単母音
- { 先先行, 先行, 当該, 後続, 後後続 } 半音節が二重母音
- { 先先行, 先行, 当該, 後続, 後後続 } 半音節が三重母音
- { 先先行, 先行, 当該, 後続, 後後続 } 半音節が複合母音
- { 先先行, 先行, 当該, 後続, 後後続 } 半音節が鼻音+母音
- { 先先行, 先行, 当該, 後続, 後後続 } 半音節が前母音
- { 先先行, 先行, 当該, 後続, 後後続 } 半音節が後母音
- { 先先行, 先行, 当該, 後続, 後後続 } 半音節が中母音

{ 先先行, 先行, 当該, 後続, 後後続 } 半音節が音素 ich
 { 先先行, 先行, 当該, 後続, 後後続 } 半音節が音素 ic
 { 先先行, 先行, 当該, 後続, 後後続 } 半音節が音素 in
 { 先先行, 先行, 当該, 後続, 後後続 } 半音節が音素 ing
 { 先先行, 先行, 当該, 後続, 後後続 } 半音節が音素 iong
 { 先先行, 先行, 当該, 後続, 後後続 } 半音節が音素 iu
 { 先先行, 先行, 当該, 後続, 後後続 } 半音節が音素 j
 { 先先行, 先行, 当該, 後続, 後後続 } 半音節が音素 k
 { 先先行, 先行, 当該, 後続, 後後続 } 半音節が音素 l
 { 先先行, 先行, 当該, 後続, 後後続 } 半音節が音素 m
 { 先先行, 先行, 当該, 後続, 後後続 } 半音節が音素 n
 { 先先行, 先行, 当該, 後続, 後後続 } 半音節が音素 o
 { 先先行, 先行, 当該, 後続, 後後続 } 半音節が音素 ong
 { 先先行, 先行, 当該, 後続, 後後続 } 半音節が音素 ou
 { 先先行, 先行, 当該, 後続, 後後続 } 半音節が音素 p
 { 先先行, 先行, 当該, 後続, 後後続 } 半音節が音素 pau
 { 先先行, 先行, 当該, 後続, 後後続 } 半音節が音素 q
 { 先先行, 先行, 当該, 後続, 後後続 } 半音節が音素 r
 { 先先行, 先行, 当該, 後続, 後後続 } 半音節が音素 s
 { 先先行, 先行, 当該, 後続, 後後続 } 半音節が音素 sh
 { 先先行, 先行, 当該, 後続, 後後続 } 半音節が音素 sil
 { 先先行, 先行, 当該, 後続, 後後続 } 半音節が音素 sp
 { 先先行, 先行, 当該, 後続, 後後続 } 半音節が音素 t
 { 先先行, 先行, 当該, 後続, 後後続 } 半音節が音素 u
 { 先先行, 先行, 当該, 後続, 後後続 } 半音節が音素 ua
 { 先先行, 先行, 当該, 後続, 後後続 } 半音節が音素 uai
 { 先先行, 先行, 当該, 後続, 後後続 } 半音節が音素 uan
 { 先先行, 先行, 当該, 後続, 後後続 } 半音節が音素 uang
 { 先先行, 先行, 当該, 後続, 後後続 } 半音節が音素 ui
 { 先先行, 先行, 当該, 後続, 後後続 } 半音節が音素 un
 { 先先行, 先行, 当該, 後続, 後後続 } 半音節が音素 uo
 { 先先行, 先行, 当該, 後続, 後後続 } 半音節が音素 v
 { 先先行, 先行, 当該, 後続, 後後続 } 半音節が音素 van
 { 先先行, 先行, 当該, 後続, 後後続 } 半音節が音素 ve
 { 先先行, 先行, 当該, 後続, 後後続 } 半音節が音素 vn
 { 先先行, 先行, 当該, 後続, 後後続 } 半音節が音素 x
 { 先先行, 先行, 当該, 後続, 後後続 } 半音節が音素 z
 { 先先行, 先行, 当該, 後続, 後後続 } 半音節が音素 zh

(3.2.2) 音節に関する質問

{ 先行, 当該, 後続 } 音節の Tone == N
 { 先行, 当該, 後続 } 音節の Tone ≤ N

{ 先行, 当該, 後続 } 音節の NewTone == N
 { 先行, 当該, 後続 } 音節の NewTone ≤ N

当該音節の形態素内での { 前, 後 } から数えた音節位置 == N
 当該音節の形態素内での { 前, 後 } から数えた音節位置 ≤ N

先行音節と当該音節間のポーズの有無
 当該音節と後続音節間のポーズの有無

但し, N は 0 以上.

(3.2.3) 品詞に関する質問

{ 先行, 当該, 後続 } 形態素の品詞が実詞
 { 先行, 当該, 後続 } 形態素の品詞が虚詞
 { 先行, 当該, 後続 } 形態素の品詞がその他
 { 先行, 当該, 後続 } 形態素の品詞が体詞
 { 先行, 当該, 後続 } 形態素の品詞が述詞
 { 先行, 当該, 後続 } 形態素の主品詞が形容詞
 { 先行, 当該, 後続 } 形態素の主品詞が数詞
 { 先行, 当該, 後続 } 形態素の主品詞が名詞
 { 先行, 当該, 後続 } 形態素の主品詞が不変化詞
 { 先行, 当該, 後続 } 形態素の主品詞が前置詞
 { 先行, 当該, 後続 } 形態素の主品詞が動詞
 { 先行, 当該, 後続 } 形態素の主品詞がその他

{ 先行, 当該, 後続 } 形態素の次品詞 AD
 { 先行, 当該, 後続 } 形態素の次品詞 AS
 { 先行, 当該, 後続 } 形態素の次品詞 BA
 { 先行, 当該, 後続 } 形態素の次品詞 BE
 { 先行, 当該, 後続 } 形態素の次品詞 CC
 { 先行, 当該, 後続 } 形態素の次品詞 CD
 { 先行, 当該, 後続 } 形態素の次品詞 CS
 { 先行, 当該, 後続 } 形態素の次品詞 DEC
 { 先行, 当該, 後続 } 形態素の次品詞 DEG
 { 先行, 当該, 後続 } 形態素の次品詞 DER
 { 先行, 当該, 後続 } 形態素の次品詞 DEV
 { 先行, 当該, 後続 } 形態素の次品詞 DT
 { 先行, 当該, 後続 } 形態素の次品詞 ETC
 { 先行, 当該, 後続 } 形態素の次品詞 FW
 { 先行, 当該, 後続 } 形態素の次品詞 IJ
 { 先行, 当該, 後続 } 形態素の次品詞 JJ
 { 先行, 当該, 後続 } 形態素の次品詞 LB
 { 先行, 当該, 後続 } 形態素の次品詞 LC
 { 先行, 当該, 後続 } 形態素の次品詞 MC
 { 先行, 当該, 後続 } 形態素の次品詞 MSP
 { 先行, 当該, 後続 } 形態素の次品詞 NN
 { 先行, 当該, 後続 } 形態素の次品詞 NR
 { 先行, 当該, 後続 } 形態素の次品詞 NT
 { 先行, 当該, 後続 } 形態素の次品詞 OD
 { 先行, 当該, 後続 } 形態素の次品詞 ON
 { 先行, 当該, 後続 } 形態素の次品詞 PN
 { 先行, 当該, 後続 } 形態素の次品詞 PP
 { 先行, 当該, 後続 } 形態素の次品詞 PU
 { 先行, 当該, 後続 } 形態素の次品詞 SB
 { 先行, 当該, 後続 } 形態素の次品詞 SP
 { 先行, 当該, 後続 } 形態素の次品詞 VA
 { 先行, 当該, 後続 } 形態素の次品詞 VC
 { 先行, 当該, 後続 } 形態素の次品詞 VE
 { 先行, 当該, 後続 } 形態素の次品詞 VV

{ 先行, 当該, 後続 } 形態素の音節単位で数えた長さ == N

{ 先行, 当該, 後続 } 形態素の音節単位で数えた長さ ≤ N

当該形態素の呼気段落内での { 前, 後 } から { 音節, 形態素 } 単位で数えた位置 == N

当該形態素の呼気段落内での { 前, 後 } から { 音節, 形態素 } 単位で数えた位置 ≤ N

先行形態素と当該形態素間のポーズの有無 当該形態素と後続形態素間のポーズの有無

但し, N は 0 以上.

(3.2.4) 呼気段落に関する質問

{ 先行, 当該, 後続 } 呼気段落の { 音節, 形態素 } 単位で数えた長さ == N

{ 先行, 当該, 後続 } 呼気段落の { 音節, 形態素 } 単位で数えた長さ ≤ N

当該呼気段落の文章内での { 前, 後 } から { 音節, 形態素, 呼気段落 } 単位で数えた位置 == N

当該呼気段落の文章内での { 前, 後 } から { 音節, 形態素, 呼気段落 } 単位で数えた位置 ≤ N

{ 先行, 当該, 後続 } 呼気段落の疑問の有無

但し, N は 0 以上

(3.2.5) 文章に関する質問

{ 音節, 形態素, 呼気段落 } 単位で数えた文章の長さ == N

{ 音節, 形態素, 呼気段落 } 単位で数えた文章の長さ ≤ N

文型

但し, N は 0 以上

4 おわりに

本研究では, HMM による韻律生成のためのコンテキスト依存ラベル及びクラスタリング時の分割条件の設計を行った.

今後は, 新しく設計したコンテキスト依存ラベル及び分割条件で学習を行った HMM より生成した F_0 パターンの評価, 波形選択型音声合成システムへの組み込み等を行う予定である.

参考文献

- [1] A. Ljolje and F. Fallside, "Synthesis of natural sounding pitch contours in isolated utterances using hidden Markov models," *Proc. of IEEE Trans. Acoust., Speech & Signal Process.*, ASSP-34, pp.1074–1080, 1986.
- [2] T. Fukada, Y. Komori, T. Aso and Y. Ohora, "A study of pitch pattern generation using HMM-based statistical information," *Proc. of ICSLP*, pp.723–726, 1994.
- [3] 吉村 貴克, 徳田 恵一, 益子 貴史, 小林 隆夫, 北村 正, "HMM に基づく音声合成におけるスペクトル・ピッチ・継続長の同時モデル化", *信学論 (D-II)*, vol.J83-D-II, no.12, pp.2099–2107, Nov. 2000.
- [4] 徳田 恵一, 益子 貴史, 宮崎 昇, 小林 隆夫, "多空間上の確率分布に基づいた HMM", *電子情報通信学会論文誌*, vol.J83-D-II, no. 7, pp.1579–1589, July 2000.
- [5] 益子 貴史, 徳田 恵一, 宮崎 昇, 小林 隆夫, "多空間確率分布 HMM によるピッチパターン生成", *電子情報通信学会論文誌*, vol.J83-D-II, no. 7, pp.1600–1609, July 2000.
- [6] K. Tokuda, H. Zen and A. W. Black, "An HMM-based speech synthesis system applied to English," *Proc. of IEEE Speech Synthesis Workshop*, Sep. 2002.
- [7] 小林 隆夫, 徳田 恵一, 山下 洋一, 山田 篤, "音声対話擬人化エージェント音声合成プログラムの基本設計書", <http://www.lang.astem.or.jp/Agent/sekkei/gousei-spec.pdf>, pp.19–23.
- [8] P. C. Woodland, C. J. Leggetter, J. J. Odell, V. Valthev, and S. J. Young, "The 1994 HTK large vocabulary speech recognition system," *Proc. of ICASSP*, pp.73–76, 1995.
- [9] 伝 康晴, 宇津呂 武仁, 山田 篤, 浅原 正幸, 松本 裕治 "話し言葉研究に適した電子化辞書の設計", 第 2 回「話し言葉の科学と工学」ワークショップ, Feb. 2002.
- [10] S. Young, G. Evermann, D. Kershaw, G. Moore, J. Odell, D. Ollason, D. Povey, V. Valtchev and P. Woodland, "The HTK Book for HTK Version 3.2", Cambridge University Engineering Department, Dec. 2002.
- [11] J. J. Odell, "The Use of Context in Large Vocabulary Speech Recognition," PhD dissertation, Cambridge University, 1995.

Appendix : HMM による音声合成システムの学習手順

本章では、HMM による音声合成システムの学習手順について記す。

学習データの準備

HTK 用特徴量ファイルとそのリスト

スペクトルに関する特徴量 (STRAIGHT-MelCepstrum 及びその 1 次及び 2 次の動的特徴からなるベクトル) と基本周波数に関する特徴ベクトル (有声区間では対数基本周波数及びその 1 次及び 2 次の動的特徴量からなる 3 次元のベクトル, 無声区間では -10^{10} からなる 3 次元のベクトル) をフレーム単位で結合したデータを作成する。

`/data/GOUSEI_data/xhzen/mkdata.pl` 中の音声が存在するディレクトリ, F_0 が存在するディレクトリ, 出力ディレクトリを適宜変更し, 実行することで自動的に学習データを作成できる。

上記の Perl スクリプトを実行し, 作成された学習データのファイル名及びパスを列挙したりリストを作成する。

言語韻律情報ファイル

学習データに対応した言語韻律情報ファイルを準備する。

ラベルファイルとリストファイル

学習データに対応した以下のラベルファイルを準備する。

- 音素境界情報付きコンテキスト独立ラベルファイル
- コンテキスト依存ラベルファイル

これらのファイルは, 以下のプログラムを用いて作成できる。

- 日本語

`/home/xhzen/work/inf2lab/japanese/plinfo/mklab.sh` 中の言語韻律情報ファイルの存在するディレクトリ, 音素境界ファイル (mk ファイル) の存在するディレクトリ, 出力ディレクトリを適宜変更して実行する。

- 中国語

`/home/xhzen/work/inf2lab/chinese/mklab.sh` 中の言語韻律情報ファイルの存在するディレクトリ, 出力ディレクトリを適宜変更して実行する。

リストファイルは, ラベルファイル中に出現したコンテキスト依存ラベルを全て列挙したものである。

決定木によるコンテキストクラスタリングのための分割条件

2.2 で述べた分割条件が/home/xhzen/work/questions/japanese/questions_qn001.hed にある.

3.2 で述べた分割条件が/home/xhzen/work/questions/chinese/questions_qn001.hed にある.

HMM のトポロジー定義ファイル

HMM のトポロジー定義ファイルは以下ようになっており, HTK のフォーマットを拡張した型式になっている.

HMM 音声合成用トポロジー定義ファイル

```

'o <VecSize> 78 <USER> <MSDINFO> 4 0 1 1 1 <StreamInfo> 4 75 1 1 1
<BeginHMM>
<NumStates> 7
<State> 2
<Stream> 1
  <Mean> 75
    0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
    0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
    0.0 0.0 0.0 0.0 0.0 0.0
    0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
    0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
    0.0 0.0 0.0 0.0 0.0
    0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
    0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
    0.0 0.0 0.0 0.0 0.0
  <Variance> 75
    1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0
    1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0
    1.0 1.0 1.0 1.0 1.0
    1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0
    1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0
    1.0 1.0 1.0 1.0
    1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0
    1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0
    1.0 1.0 1.0 1.0 1.0
  <Stream> 2
  <NumMixes> 2
  <Mixture> 1 0.5000
    <Mean> 1
      0.0
    <Variance> 1
      1.0
  <Mixture> 2 0.5000
    <Mean> 0
    <Variance> 0
  <Stream> 3
  <NumMixes> 2
  <Mixture> 1 0.5000
    <Mean> 1
      0.0
    <Variance> 1
      1.0
  <Mixture> 2 0.5000
    <Mean> 0
    <Variance> 0
  <Stream> 4
  <NumMixes> 2
  <Mixture> 1 0.5000
    <Mean> 1
      0.0
    <Variance> 1
      1.0
  <Mixture> 2 0.5000
    <Mean> 0
    <Variance> 0
  <State> 3
  .
  .
  <TransP> 7
    0.000e+0 1.000e+0 0.000e+0 0.000e+0 0.000e+0 0.000e+0 0.000e+0
    0.000e+0 6.000e-1 4.000e-1 0.000e+0 0.000e+0 0.000e+0 0.000e+0
    0.000e+0 0.000e+0 0.000e+0 6.000e-1 4.000e-1 0.000e+0 0.000e+0
    0.000e+0 0.000e+0 0.000e+0 6.000e-1 4.000e-1 0.000e+0 0.000e+0
    0.000e+0 0.000e+0 0.000e+0 0.000e+0 6.000e-1 4.000e-1 0.000e+0
    0.000e+0 0.000e+0 0.000e+0 0.000e+0 0.000e+0 6.000e-1 4.000e-1
    0.000e+0 0.000e+0 0.000e+0 0.000e+0 0.000e+0 0.000e+0 0.000e+0
  <EndHMM>

```

第1ストリームがスペクトルに関する特徴量, 第2ストリームが F_0 の静的特徴量, 第3ストリームが F_0 の1次動的特徴量, 第4ストリームが F_0 の2次動的特徴量に対応する。

この例では, 25次の STRAIGHT-MelCepstrum を用いているため, 第1ストリームの次数は $75(25 \times 3)$ になっているが, 次数を変更したい場合は, $\langle \text{VecSize} \rangle$, $\langle \text{StreamInfo} \rangle$ 及び第1ストリームの $\langle \text{Mean} \rangle$, $\langle \text{Variance} \rangle$ の次数を変更すれば良い。

HMM の学習

前章で準備したファイルを用いて学習を行う。手順を以下に示す。

実行手順

1. 初期学習 (HInit) :
音素境界情報を用いて, Viterbi 学習によりコンテキスト独立モデルを学習する.
2. 再学習 (HRest) :
音素境界情報を用いて, Baum-Welch アルゴリズムにより HMM のパラメータを更新する.
3. コンテキスト依存モデルの作成 (HHEd) :
コンテキスト依存モデルの中心音素に対応するコンテキスト独立モデルをコピーして, コンテキスト依存モデルを作成する.
4. 連結学習 (HERest) :
各発話に対応する文章 HMM に対して Baum-Welch アルゴリズムを適用し, コンテキスト依存モデルのパラメータを更新する. これを数回繰り返す.
5. 決定木によるコンテキストクラスタリング (HHEd) :
コンテキスト依存モデルに対し, スペクトル, 基本周波数それぞれ別に決定木によるコンテキストクラスタリングを適用し, パラメータの共有を行う.
6. 連結学習 (HERest) :
パラメータの共有を行った HMM に対して 4 と同様に Baum-Welch アルゴリズムを適用する. 最後の繰り返しの際に, コンテキスト依存状態継続長分布モデルを作成する.
7. 決定木によるコンテキストクラスタリング (HHEd) :
コンテキスト依存状態継続長モデルに対し, 決定木によるコンテキストクラスタリングを適用し, パラメータの共有を行う.
8. 未知コンテキストモデルの作成 (HHEd) :
学習データに含まれないモデルを, 構築した決定木によるコンテキスト補完により作成する.

`/home/xhzen/work/Training.pl` 中の各ファイル名, ディレクトリを適宜変更し実行することで, 自動的に学習を行うことができる.

音声の合成

合成したい文章のラベル列を用意し, `/home/xhzen/work/Synthesis.pl` 中のラベル列の存在するディレクトリ, 出力ディレクトリ等を適宜設定し実行することで, 音声合成される.