An example of context-dependent label format for HMM-based speech synthesis in English

Keiichiro Oura

May 3, 2011

$p_1\hat{\ }p_2-p_3+p_4=p_5@p_6\_p_7$
$/\texttt{A}:a_1\_a_2\_a_3$ $/\texttt{B}:b_1-b_2-b_3@b_4-b_5\&b_6-b_7\#b_8-b_9\$b_{10}-b_{11}!b_{12}-b_{13};b_{14}-b_{15}|b_{16}$ $/\texttt{C}:c_1+c_2+c_3$
$/\texttt{D}:d_1\_d_2$ $/\texttt{E}:e_1+e_2@e_3+e_4\&e_5+e_6\#e_7+e_8$ $/\texttt{F}:f_1\_f_2$
$/\texttt{G}:g_1\_g_2$ $/\texttt{H}:h_1=h_2\hat{\ }h_3=h_4|h_5$ $/\texttt{I}:i_1\_i_2$
$/\texttt{J}:j_1+j_2-j_3$

p_1	the phoneme identity before the previous phoneme
p_2	the previous phoneme identity
p_3	the current phoneme identity
p_4	the next phoneme identity
p_5	the phoneme after the next phoneme identity
p_6	position of the current phoneme identity in the current syllable (forward)
p_7	position of the current phoneme identity in the current syllable (backward)
a_1	whether the previous syllable stressed or not (0: not stressed, 1: stressed)
a_2	whether the previous syllable accented or not (0: not accented, 1: accented)
a_3	the number of phonemes in the previous syllable
b_1	whether the current syllable stressed or not (0: not stressed, 1: stressed)
b_2	whether the current syllable accented or not (0: not accented, 1: accented)
b_3	the number of phonemes in the current syllable
b_4	position of the current syllable in the current word (forward)
b_5	position of the current syllable in the current word (backward)
b_6	position of the current syllable in the current phrase (forward)
b_7	position of the current syllable in the current phrase (backward)
b_8	the number of stressed syllables before the current syllable in the current phrase
b_9	the number of stressed syllables after the current syllable in the current phrase
b_{10}	the number of accented syllables before the current syllable in the current phrase
b_{11}	the number of accented syllables after the current syllable in the current phrase
b_{12}	the number of syllables from the previous stressed syllable to the current syllable
b_{13}	the number of syllables from the current syllable to the next stressed syllable
b_{14}	the number of syllables from the previous accented syllable to the current syllable
b_{15}	the number of syllables from the current syllable to the next accented syllable
b_{16}	name of the vowel of the current syllable
c_1	whether the next syllable stressed or not (0: not stressed, 1: stressed)
c_2	whether the next syllable accented or not (0: not accented, 1: accented)
c_3	the number of phonemes in the next syllable
d_1	gpos (guess part-of-speech) of the previous word
d_2	the number of syllables in the previous word
e_1	gpos (guess part-of-speech) of the current word
e_2	the number of syllables in the current word
e_3	position of the current word in the current phrase (forward)
e_4	position of the current word in the current phrase (backward)
e_5	the number of content words before the current word in the current phrase
e_6	the number of content words after the current word in the current phrase
e_7	the number of words from the previous content word to the current word
e_8	the number of words from the current word to the next content word
f_1	gpos (guess part-of-speech) of the next word
f_2	the number of syllables in the next word
g_1	the number of syllables in the previous phrase
g_2	the number of words in the previous phrase
h_1	the number of syllables in the current phrase
h_2	the number of words in the current phrase
h_3	position of the current phrase in this utterance (forward)
h_4	position of the current phrase in this utterance (backward)
h_5	TOBI endtone of the current phrase
i_1	the number of syllables in the next phrase
i_2	the number of words in the next phrase
j_1	the number of syllables in this utterance
j_2	the number of words in this utterance
j_3	the number of phrases in this utterance