An example of context-dependent label format for HMM-based singing voice synthesis

HTS Working Group

December 25, 2015

$p_1@p_2\hat{\ }p_3-p_4+p_5=p_6\_p_7\%p_8\hat{\ }p_9\_p_{10}\sim p_{11}-p_{12}!p_{13}[p_{14}\$p_{15}]p_{16}$
$/\mathtt{A}:a_1-a_2-a_3@a_4\sim a_5\ /\mathtt{B}:b_1\_b_2\_b_3@b_4|b_5\ /\mathtt{C}:c_1+c_2+c_3@c_4\&c_5$
$/\mathtt{D}:d_1!d_2\#d_3\$d_4\%d_5|d_6\&d_7;d_8-d_9$
$/\mathtt{E}:e_1]e_2\hat{\ }e_3=e_4\sim e_5!e_6@e_7\#e_8+e_9]e_{10}\$e_{11}|e_{12}[e_{13}\&e_{14}]e_{15}=e_{16}\hat{\ }e_{17}\sim e_{18}\#e_{19}\_e_{20};e_{21}\$e_{22}\&e_{23}\%e_{24}[e_{25}|e_{26}]e_{27}-e_{28}\hat{\ }e_{29}+e_{30}\sim$
$e_{31}=e_{32}@e_{33}\$e_{34}!e_{35}\%e_{36}\#e_{37}|e_{38}|e_{39}-e_{40}\&e_{41}\&e_{42}+e_{43}[e_{44};e_{45}]e_{46};e_{47}\sim e_{48}\sim e_{49}\hat{\ }e_{50}\hat{\ }e_{51}@e_{52}[e_{53}\#e_{54}=e_{55}!e_{56}\sim e_{57}+e_{58}!e_{59}\hat{\ }e_{60}$
$/\mathtt{F}:f_1\#f_2\#f_3-f_4\$f_5\$f_6+f_7\%f_8;f_9$
$/\mathtt{G}:g_1\_g_2\ /\mathtt{H}:h_1\_h_2\ /\mathtt{I}:i_1\_i_2$
$/\mathtt{J}:j_1\sim j_2@j_3$

p_1	the language independent phoneme identity
p_2	the phoneme identity before the previous phoneme
p_3	the previous phoneme identity
p_4	the current phoneme identity
p_5	the next phoneme identity
p_6	the phoneme idendity after the next phoneme
p_7	the phoneme flag before the previous phoneme
p_8	the previous phoneme flag
p_9	the current phoneme flag
p_{10}	the next phoneme flag
p_{11}	the phoneme flag after the next phoneme
p_{12}	position of the current phoneme identity in the syllable (forward)
p_{13}	position of the current phoneme identity in the syllable (backward)
p_{14}	the distance from the previous vowel in the current syllable to the current consonant
p_{15}	the distance from the current consonant to the next vowel in the current syllable
p_{16}	undefined context
a_1	the number of phonemes in the previous syllable
a_2	position of the previous syllable identity in the note (forward)
a_3	position of the previous syllable identity in the note (backward)
a_4	the language of the previous syllable
a_5	the language dependent context of the previous syllable
b_1	the number of phonemes in the current syllable
b_2	position of the current syllable identity in the note (forward)
b_3	position of the current syllable identity in the note (backward)
b_4	the language of the current syllable
b_5	the language dependent context of the current syllable
c_1	the number of phonemes in the next syllable
c_2	position of the next syllable identity in the note (forward)
c_3	position of the next syllable identity in the note (backward)
c_4	the language of the next syllable
c_5	the language dependent context of the next syllable
d_1	the absolute pitch of the previous note (C0-G9)
d_2	the relative pitch of the previous note (0-11)
d_3	the key of the previous note (the number of sharp)
d_4	the beat of the previous note
d_5	the tempo of the previous note
d_6	the length of the previous note by the syllable
d_7	the length of the previous note by 0.01 second
d_8	the length of the previous note by one-third of the 32nd note
d_9	undefined context
e_1	the absolute pitch of the current note (C0-G9)
e_2	the relative pitch of the current note (0-11)
e_3	the key of the current note (the number of sharp)
e_4	the beat of the current note
e_5	the tempo of the current note
e_6	the length of the current note by the syllable
e_7	the length of the current note by 0.01 second
e_8	the length of the current note by one-third of the 32nd note
e_9	undefined context

e_{10}	position of the current note identity in the current measure by the note (forward)
e_{11}	position of the current note identity in the current measure by the note (backword)
e_{12}	position of the current note identity in the current measure by 0.01 second (forward)
e_{13}	position of the current note identity in the current measure by 0.01 second (backward)
e_{14}	position of the current note identity in the current measure by one-third of the 32nd note (forward)
e_{15}	position of the current note identity in the current measure by one-third of the 32nd note (backward)
e_{16}	position of the current note identity in the current measure by % (forward)
e_{17}	position of the current note identity in the current measure by % (backward)
e_{18}	position of the current note identity in the current phrase by the note (forward)
e_{19}	position of the current note identity in the current phrase by the note (backword)
e_{20}	position of the current note identity in the current phrase by 0.01 second (forward)
e_{21}	position of the current note identity in the current phrase by 0.01 second (backward)
e_{22}	position of the current note identity in the current phrase by one-third of the 32nd note (forward)
e_{23}	position of the current note identity in the current phrase by one-third of the 32nd note (backward)
e_{24}	position of the current note identity in the current phrase by % (forward)
e_{25}	position of the current note identity in the current phrase by % (backward)
e_{26}	whether slur or not in between the current note and the previous note
e_{27}	whether slur or not in between the current note and the next note
e_{28}	dynamic mark of the current note
e_{29}	the distance between the current note and the next accent by the note
e_{30}	the distance between the current note and the previous accent by the note
e_{31}	the distance between the current note and the next accent by 0.01 second
e_{32}	the distance between the current note and the previous accent by 0.01 second
e_{33}	the distance between the current note and the next accent by one-third of the 32nd note
e_{34}	the distance between the current note and the previous accent by one-third of the 32nd note
e_{35}	the distance between the current note and the next staccato by the note
e_{36}	the distance between the current note and the previous staccato by the note
e_{37}	the distance between the current note and the next staccato by 0.01 second
e_{38}	the distance between the current note and the previous staccato by 0.01 second
e_{39}	the distance between the current note and the next staccato by one-third of the 32nd note
e_{40}	the distance between the current note and the previous staccato by one-third of the 32nd note
e_{41}	position of the current note in the current crescendo by the note (forward)
e_{42}	position of the current note in the current crescendo by the note (backward)
e_{43}	position of the current note in the current crescendo by 1.0 second (forward)
e_{44}	position of the current note in the current crescendo by 1.0 second (backward)
e_{45}	position of the current note in the current crescendo by one-third of the 32nd note (forward)
e_{46}	position of the current note in the current crescendo by one-third of the 32nd note (backward)
e_{47}	position of the current note in the current crescendo by % (forward)
e_{48}	position of the current note in the current crescendo by % (backward)
e_{49}	position of the current note in the current decrescendo by the note (forward)
e_{50}	position of the current note in the current decrescendo by the note (backward)
e_{51}	position of the current note in the current decrescendo by 1.0 second (forward)
e_{52}	position of the current note in the current decrescendo by 1.0 second (backward)
e_{53}	position of the current note in the current decrescendo by one-third of the 32nd note (forward)
e_{54}	position of the current note in the current decrescendo by one-third of the 32nd note (backward)
e_{55}	position of the current note in the current decrescendo by % (forward)
e_{56}	position of the current note in the current decrescendo by % (backward)
e_{57}	pitch difference between the current and previous notes
e_{58}	pitch difference between the current and next notes
f_1	the absolute pitch of the next note (C0-G9)
f_2	the relative pitch of the next note (0-11)
f_3	the key of the next note (the number of sharp)
f_4	the beat of the next note
f_5	the tempo of the next note
f_6	the length of the next note by the syllable
f_7	the length of the next note by 0.01 second
f_8	the length of the next note by one-third of the 32nd note
f_9	undefined context
g_1	the number of syllables in the previous phrase
g_2	the number of phonemes in the previous phrase
h_1	the number of syllables in the current phrase
h_2	the number of phonemes in the current phrase
i_1	the number of syllables in the next phrase
i_2	the number of phonemes in the next phrase
j_1	the number of syllables in this song / the number of measures in this song
j_2	the number of phonemes in this song / the number of measures in this song
j_3	the number of phrases in this song