

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

Keiichiro Oura

May 14, 2010

$p_1 \hat{p}_2 - p_3 + p_4 = p_5$
 /A: $a_1 - a_2 - a_3$ /B: $b_1 - b_2 - b_3$ /C: $c_1 + c_2 + c_3$
 /D: $d_1 ! d_2 \# d_3 \$ d_4 \% d_5 | d_6 \& d_7 [d_8$
 /E: $e_1] e_2 \hat{e}_3 = e_4 \sim e_5 ! e_6 @ e_7 \# e_8 \% e_9 \$ e_{10} | e_{11} [e_{12} \& e_{13}] e_{14} = e_{15} \hat{e}_{16} \sim e_{17} \# e_{18} @ e_{19} ! e_{20} \$ e_{21} \& e_{22} \% e_{23} [e_{24} | e_{25}] e_{26} - e_{27} \hat{e}_{28} +$
 $e_{29} \sim e_{30} = e_{31} @ e_{32} \$ e_{33} ! e_{34} \% e_{35} \# e_{36} | e_{37} | e_{38} - e_{39} \& e_{40} \& e_{41} + e_{42} [e_{43} ; e_{44}] e_{45} ; e_{46} \sim e_{47} \sim e_{48} \hat{e}_{49} \hat{e}_{50} @ e_{51} ; e_{52} = e_{53} = e_{54} ! e_{55}$
 /F: $f_1 \# f_2 \# f_3 - f_4 \$ f_5 \$ f_6 + f_7 \% f_8$
 /G: $g_1 - g_2$ /H: $h_1 - h_2$ /I: $i_1 - i_2$
 /J: $j_1 \sim j_2 @ j_3$

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 identity after the next phoneme
a_1	the number of phonemes in the previous mora
a_2	position of the previous mora identity in the note (forward)
a_3	position of the previous mora identity in the note (backward)
b_1	the number of phonemes in the current mora
b_2	position of the current mora identity in the note (forward)
b_3	position of the current mora identity in the note (backward)
c_1	the number of phonemes in the next mora
c_2	position of the next mora identity in the note (forward)
c_3	position of the next mora identity in the note (backward)
d_1	the absolute interval of the previous note (C0-G9)
d_2	the relative interval 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 (SS: 1-75 SM: 76-90 SH: 91-105 MS: 106-120 MM: 121-135 MH 136-150 HS: 151-165 HM: 166-180 HH: 181-)
d_6	the length of the previous note by the mora
d_7	the length of the previous note by 0.1 second (1-99)
d_8	the length of the previous note by three thirty-second note (1-199)
e_1	the absolute interval of the current note (C0-G9)
e_2	the relative interval 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 (SS: 1-75 SM: 76-90 SH: 91-105 MS: 106-120 MM: 121-135 MH 136-150 HS: 151-165 HM: 166-180 HH: 181-)
e_6	the length of the current note by the mora
e_7	the length of the current note by 0.1 second (1-99)
e_8	the length of the current note by three thirty-second note (1-199)
e_9	position of the current note identity in the current measure by the note (forward, 1-49)
e_{10}	position of the current note identity in the current measure by the note (backward, 1-49)
e_{11}	position of the current note identity in the current measure by 0.1 second (forward, 1-49)
e_{12}	position of the current note identity in the current measure by 0.1 second (backward, 1-49)
e_{13}	position of the current note identity in the current measure by three thirty-second note (forward, 1-99)
e_{14}	position of the current note identity in the current measure by three thirty-second note (backward, 1-99)
e_{15}	position of the current note identity in the current measure by % (forward)
e_{16}	position of the current note identity in the current measure by % (backward)
e_{17}	position of the current note identity in the current phrase by the note (forward, 1-99)
e_{18}	position of the current note identity in the current phrase by the note (backward, 1-99)
e_{19}	position of the current note identity in the current phrase by 0.1 second (forward, 1-199)
e_{20}	position of the current note identity in the current phrase by 0.1 second (backward, 1-199)
e_{21}	position of the current note identity in the current phrase by three thirty-second note (forward, 1-499)
e_{22}	position of the current note identity in the current phrase by three thirty-second note (backward, 1-499)
e_{23}	position of the current note identity in the current phrase by % (forward)
e_{24}	position of the current note identity in the current phrase by % (backward)

e_{25}	whether tied (slur) or not in between the current note and the previous note (0: not tied, 1: tied)
e_{26}	whether tied (slur) or not in between the current note and the previous note (0: not tied, 1: tied)
e_{27}	dynamic mark of the current note
e_{28}	the distance between the current note and the next accent by the note (1-9)
e_{29}	the distance between the current note and the previous accent by the note (1-9)
e_{30}	the distance between the current note and the next accent by 0.1 second (1-99)
e_{31}	the distance between the current note and the previous accent by 0.1 second (1-99)
e_{32}	the distance between the current note and the next accent by three thirty-second note (1-99)
e_{33}	the distance between the current note and the previous accent by three thirty-second note (1-99)
e_{34}	the distance between the current note and the next staccato by the note (1-9)
e_{35}	the distance between the current note and the previous staccato by the note (1-9)
e_{36}	the distance between the current note and the next staccato by 0.1 second (1-99)
e_{37}	the distance between the current note and the previous staccato by 0.1 second (1-99)
e_{38}	the distance between the current note and the next staccato by three thirty-second note (1-99)
e_{39}	the distance between the current note and the previous staccato by three thirty-second note (1-99)
e_{40}	position of the current note in the current crescendo by the note (forward, 1-49)
e_{41}	position of the current note in the current crescendo by the note (backward, 1-49)
e_{42}	position of the current note in the current crescendo by 1.0 second (forward, 1-99)
e_{43}	position of the current note in the current crescendo by 1.0 second (backward, 1-99)
e_{44}	position of the current note in the current crescendo by three thirty-second note (forward, 1-499)
e_{45}	position of the current note in the current crescendo by three thirty-second note (backward, 1-499)
e_{46}	position of the current note in the current crescendo by % (forward)
e_{47}	position of the current note in the current crescendo by % (backward)
e_{48}	position of the current note in the current decrescendo by the note (forward, 1-49)
e_{49}	position of the current note in the current decrescendo by the note (backward, 1-49)
e_{50}	position of the current note in the current decrescendo by 1.0 second (forward, 1-99)
e_{51}	position of the current note in the current decrescendo by 1.0 second (backward, 1-99)
e_{52}	position of the current note in the current decrescendo by three thirty-second note (forward, 1-499)
e_{53}	position of the current note in the current decrescendo by three thirty-second note (backward, 1-499)
e_{54}	position of the current note in the current decrescendo by % (forward)
e_{55}	position of the current note in the current decrescendo by % (backward)
f_1	the absolute interval of the next note (C0-G9)
f_2	the relative interval 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 (SS: 1-75 SM: 76-90 SH: 91-105 MS: 106-120 MM: 121-135 MH 136-150 HS: 151-165 HM: 166-180 HH: 181-)
f_6	the length of the next note by the mora
f_7	the length of the next note by 0.1 second (1-99)
f_8	the length of the next note by three thirty-second note (1-199)
g_1	the number of moras in the previous phrase (1-99)
g_2	the number of phonemes in the previous phrase (1-99)
h_1	the number of moras in the current phrase (1-99)
h_2	the number of phonemes in the current phrase (1-99)
i_1	the number of moras in the next phrase (1-99)
i_2	the number of phonemes in the next phrase (1-99)
j_1	the number of moras in this song / the number of measures in this song (1-99)
j_2	the number of phonemes in this song / the number of measures in this song (1-99)
j_3	the number of phrases in this song (1-99)