

String Edit Distance

$D(i, j) = d(i, j) + \min [D(i-1, j), D(i-1, j-1), D(i, j-1)]$ for $i=1.....M$ and $j = 1.....N$

$\text{Path}(i, j) = \arg \min [D(i-1, j), D(i-1, j-1), D(i, j-1)]$

Initial condition path starts at $D(1,1)$

Final condition path ends at $D(M,N)$

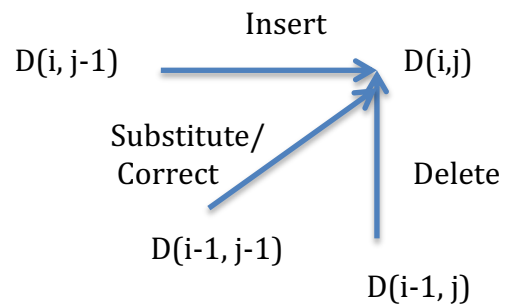
if $\text{str}(i) = \text{str}(j)$

$d(i, j) = 0$

else

$d(i, j) = 1$

end if



Backtrack the path to find the number of insertions, deletion, and substitution.

K	4					
R	3					
A	2					
P	1					
	0	1	2	3	4	5
		S	P	A	K	E

Word Error Rate

Reference transcript: *she had a dark suit in greasy water all year*

ASR output: *he had a dark grey suit in water all there*

WER = (# Ins + # Del + # Sub)/(# of words in reference transcript)

<i>year</i>	10										
<i>all</i>	9										
<i>water</i>	8										
<i>greasy</i>	7										
<i>in</i>	6										
<i>suit</i>	5										
<i>dark</i>	4										
<i>a</i>	3										
<i>had</i>	2										
<i>she</i>	1										
	0	1	2	3	4	5	6	7	8	9	10
		<i>he</i>	<i>had</i>	<i>a</i>	<i>dark</i>	<i>grey</i>	<i>suit</i>	<i>in</i>	<i>water</i>	<i>all</i>	<i>there</i>

Computation of log-likelihood of a HMM given an Observation Sequence
using Viterbi algorithm

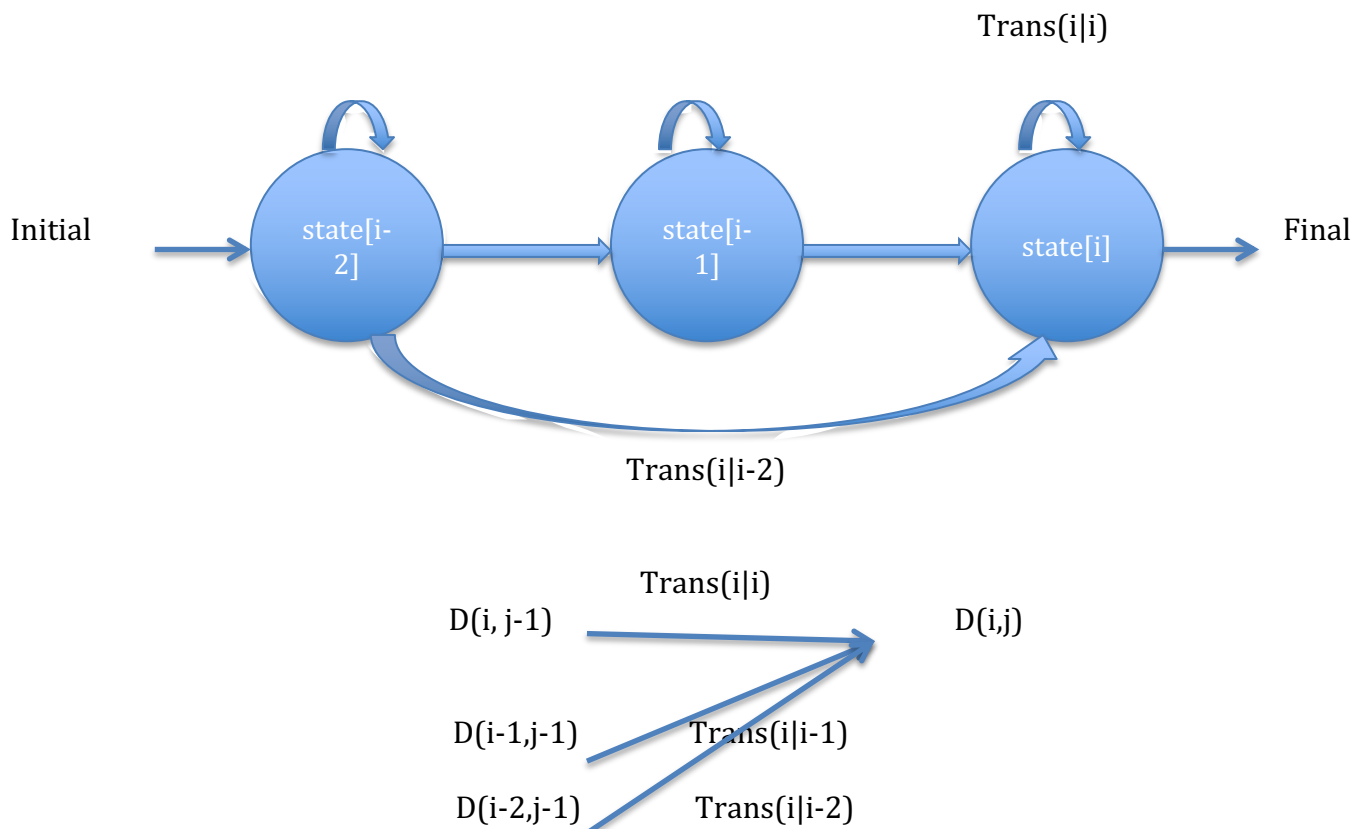
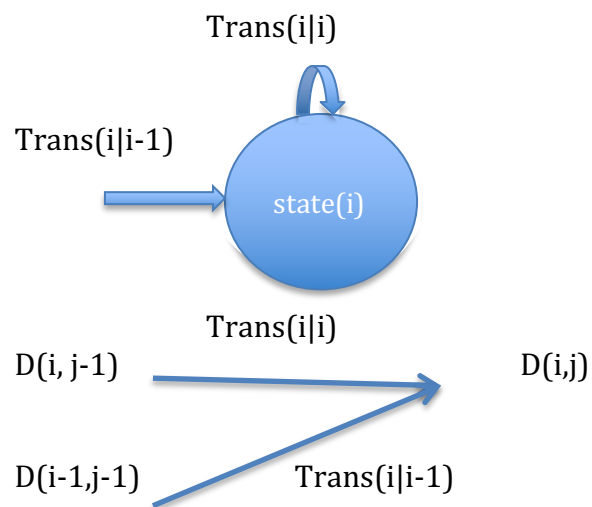
$$D(i, j) = d(i, j) + \min [(D(i-1, j-1) + \text{Trans}(i|i-1)), (D(i, j-1) + \text{Trans}(i|i))]$$

$$\text{Path}(i, j) = \arg \min [(D(i-1, j-1) + \text{Trans}(i|i-1)), (D(i, j-1) + \text{Trans}(i|i))]$$

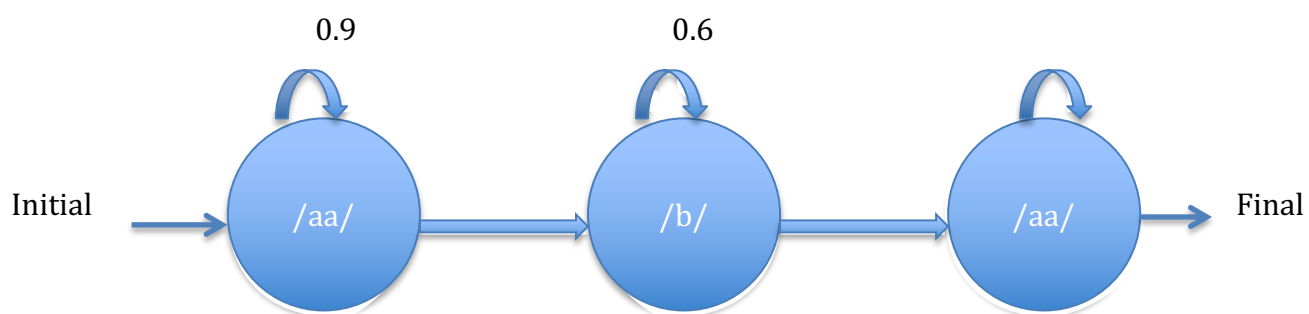
$$d(i, j) = -\log(P(\text{obs}(j)|\text{state}(i)))$$

$$\text{Trans}(i|i-1) = -\log(P(\text{state}(i)|\text{state}(i-1)))$$

NOTE: Local constraints are based on HMM topology



Example with single dimension observation



Initial and Final are non-emitting states

State /aa/:

mean = 15.1

standard deviation = 3.2

State /b/:

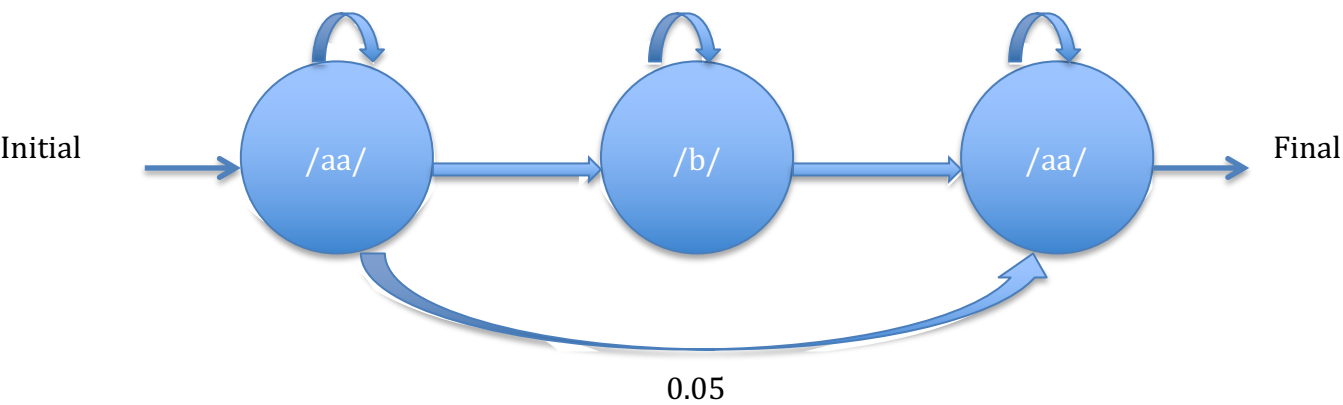
mean = 10.3

standard deviation = 4.4

Observation sequence: 16.0, 13.7, 12.0, 9.8, 10.0, 11.8, 14.2, 17.2, 12.5

/aa/	Inf									
/b/	Inf									
/aa/	Inf									
	0.0	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf
		16.0	13.7	12.0	9.8	10.0	11.8	14.2	17.2	12.5

Compute the log-likelihood for the following HMM topology. Note down the key changes in the computation that are to be made w.r.t the previous topology.



/aa/	Inf									
/b/	Inf									
/aa/	Inf									
	0.0	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf
		16.0	13.7	12.0	9.8	10.0	11.8	14.2	17.2	12.5