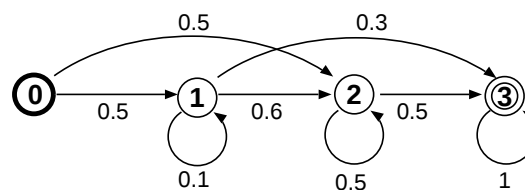


Automatic Speech Recognition

1. Exercise

Task 1.1 Hidden Markov Models

Consider the following HMM.



	1	2	3
sim	0.6	0.4	0
sa	0.2	0.2	0.3
la	0.1	0.4	0.3
bim	0.1	0	0.4

The HMM generates syllable sequences made up of the four syllables “sim”, “sa”, “la”, and “bim”. In the figure the transition probabilities are attached to the arcs of the automaton, while the emission probabilities are shown in the table. For example, the probability of generating the syllable “sim” when being in state 2 is given by $p(\text{“sim”} | 2) = 0.4$ and the probability of moving from state 2 to state 3 is $p(3 | 2) = 0.5$.

State 0 of the HMM is a virtual initial state and does not produce any output; 3 is the final state.

Look into the syllable sequence “sim-sa-la-bim”:

- Unfold the automaton along the position axis and draw the resulting grid for a sequence of four syllables. (2)
- Calculate the approximative probability for the syllable sequence “sim-sa-la-bim” using the Viterbi approximation, i.e. consider only the best path. (3)
- Calculate the probability for the syllable sequence “sim-sa-la-bim”, i.e. sum over all possible paths. (3)
- Compare and discuss the results from b) and c).
What are the advantages and disadvantages of using the Viterbi approximation in speech recognition? Remember, in efficient ASR systems the negative logarithms of the probabilities are used. (4)

Task 1.2 Repetition of Dynamic Programming in Speech Recognition

The application of dynamic programming to the speech recognition problem will be repeated in the exercise course on 27.04.2010 as preparation for the next exercise.