Managing Massive Multiplayer Online Games
SS 2017

Exercise Sheet 8: Temporal Behavior
Discussion: June 28th, 2017

Exercise 8-1    Suffix Trees

The alphabet $A = \{A, B, C, D, N\}$ is given.

(a) Insert the sequence $G_1 = \{B, A, N, A, N, A\}$ into an empty suffix tree $ST$
(b) Additionally insert the sequence $G_2 = \{C, A, N, A, D, A\}$ into $ST$.
(c) Find the subsequence $S_1 = \{N, A, N, A\}$. Which sequence contains $S_1$?
(d) Which is the longest common subsequence of $G_1$ and $G_2$?
(e) Which extension would be necessary to support finding the most frequent subsequence of length $n$ (or longer)?

Exercise 8-2    Levenshtein Distance

Compute the Levenshtein Distance between the sequences $BANANA$ and $CANADA$.

Exercise 8-3    Markov Chains

The Markov Chain $M$ is given below as a graph. Nodes represent states, edges possible transitions and edge labels transition probabilities.

(a) Specify $M$ in matrix notation. For that assume that the starting states are uniformly distributed and that sequences can only end after state 3 with a probability of 50%.
(b) What is the probability to observe the sequence 3 – 1 – 2 – 3?
(c) What is the probability to observe the sequence 2 – 3 – 2 – 1 – 2?

**Exercise 8-4  HMM: Calculation exercise**

Regard the Markov model below.

(a) Specify the set of states $A$ and the set of observations $B$. Deduce the transition matrix $D$ and the output matrix $F$ from the model. Assume that the starting probabilities are uniformly distributed and that the probability that sequences end in a state correspond to the values which are missing to the sum 1.

(b) Calculate the probability that the observation $O_1 = \{H, E, L, L, O\}$ is generated by the HMM.

(c) Which sequence $(s_1, s_2, \ldots, s_k)$ with $s_i \in A$ explains the observation $O_2 = \{H, E, L, L, O\}$ best?

**Exercise 8-5  HMM: Evaluation / Detection**

The Hidden Markov Model (HMM) $M = \{S, B, D, F\}$ with $S = \{A, B, C\}$, $B = \{\spadesuit, \heartsuit, \diamondsuit\}$ is given.

$$D = \begin{pmatrix}
0 & 1/3 & 1/3 & 1/3 \\
1/4 & 1/4 & 1/4 & 1/4 \\
1/4 & 0 & 1/4 & 1/2 \\
1/4 & 1/4 & 1/2 & 0
\end{pmatrix}$$

$$F = \begin{pmatrix}
A & 1/4 & 3/4 & 0 \\
B & 0 & 0 & 1 \\
C & 0 & 1/4 & 3/4
\end{pmatrix}$$

(a) Compute the probability of the observation $\spadesuit, \heartsuit, \diamondsuit$ without algorithmic procedures. Tag the most probable sequence of states for the observation.
(b) Compute the probability of the observation ♠, ♥, ♦ inductively with help of the forward-variable

\[ \alpha_j(1) = d_{-j} f_{j,o_1} \quad \alpha_j(t + 1) = \left( \sum_{i=1}^{[A]} \alpha_i(t) d_{i,j} \right) f_{j,o_{t+1}} \]

(c) Determine with help of the Viterbi-Algorithm which sequence of states most probably produced the observation ♠, ♥, ♦

\[ \delta_j(1) = d_{-j} f_{j,o_1} \quad \delta_j(t + 1) = \left( \max_i \delta_i(t) d_{i,j} \right) f_{j,o_{t+1}} \]

\[ \psi_j(1) = 0 \quad \psi_j(t + 1) = \arg \max_i \delta_i(t) d_{i,j} \]