Course Announcements

Project Proposal due Nov 14th
Hidden Markov Model $\mathcal{M} = (Q, A, \Sigma, E)$

- Set of hidden states $Q$
  - Markov property
- Transition probabilities $A = [a_{ij}]$ on pairs of states
- Set of emitted symbols $\Sigma$
- Emission probabilities $E = [e_{ik}]$ on state-symbol pairs

Two decisions:
1. What symbol should I emit? [emission probabilities $E$]
2. What state should I move to next? [transition probabilities $A$]

**Fair Bet Casino**

$Q = \{F, B\}$

$A = \begin{pmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{pmatrix}$

$\Sigma = \{H, T\}$

$E = \begin{pmatrix} 0.5 & 0.5 \\ 0.75 & 0.25 \end{pmatrix}$

![Diagram of a Markov model with states F and B, transitions, and emissions.](image)
Three Questions

Question 1:
What is the most probable path \( \pi^* \) that generated observations \( \mathbf{x} \)?

Question 2:
What is probability of observations \( \mathbf{x} \) generated by any path \( \pi \)?

Question 3:
What is the probability of observation \( x_i \) generated by state \( s \)?
| Type        | Probability                                | Method                              | Solution                                              |
|-------------|--------------------------------------------|                                    |                                                      |
| Joint       | $\max_{\pi \in Q^n} \Pr(x, \pi)$          | Viterbi algorithm                   | $\Pr(x, \pi^*) = \max_{s \in Q} v[s, n]$             |
| Marginal    | $\Pr(x)$                                   | Forward algorithm, or backward      | $\Pr(x) = \sum_{s \in Q} f[s, n]$,                   |
|             |                                            | algorithm                           | $\Pr(x) = \sum_{s \in Q} a_{0,s} \cdot e[s, x_1] \cdot b[s, 1]$ |
| Posterior   | $\Pr(\pi_i = s \mid x)$                   | Forward algorithm, and              | $\hat{\pi}_i = \arg \max_{s \in Q} \frac{f[s,i] \cdot b[s,i]}{\sum_{t \in Q} f[t,n]}$ |

Table 1: Hidden Markov Models – Three different probabilities

<table>
<thead>
<tr>
<th>Type</th>
<th>Recurrence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Viterbi</td>
<td>$v[s, i] = \begin{cases} a_{0,s} e_{s,x_1}, &amp; \text{if } i = 1, \ e_{s,x_i} \max_{t \in Q} v[t, i - 1] \cdot a_{t,s}, &amp; \text{if } i &gt; 1. \end{cases}$</td>
</tr>
<tr>
<td>Forward</td>
<td>$f[s, i] = \begin{cases} a_{0,s} e_{s,x_1}, &amp; \text{if } i = 1, \ e_{s,x_i} \sum_{t \in Q} f[t, i - 1] \cdot a_{t,s}, &amp; \text{if } i &gt; 1. \end{cases}$</td>
</tr>
<tr>
<td>Backward</td>
<td>$b[s, i] = \begin{cases} 1, &amp; \text{if } i = n, \ \sum_{t \in Q} a_{s,t} \cdot e_{t,x_{i+1}} \cdot b[t, i + 1], &amp; \text{if } 1 \leq i &lt; n, \end{cases}$</td>
</tr>
</tbody>
</table>

Table 2: Hidden Markov Models – Three recurrences that each can be computed in $O(n|Q|^2)$ time and $O(n|Q|)$ space.
### Questions:
1. Compute most likely state path
2. Compute marginal probability
3. Compute posterior decoding
Summary

• Markov property – Current state depends only on previous state
• Hidden Markov Models: states are not given only emitted symbols
• Viterbi algorithm: Find the most likely sequence of states given a set of observations
• Baum-Welch algorithm: EM-algorithm to learn $A$ and $E$ from training set

Reading:
• Jones and Pevzner: Chapters 11.1-11.3
• Lecture notes