From previous class:

① Let's do this mathematically now. We want \( P(s_k | m_{1:n}) \leftarrow \text{offline} \)

② Let's start with a basic result for \( P(s_{1:n} | m_{1:n}) \)

\[
P(s_{1:n} | m_{1:n}) = \frac{P(s_{1:n}, m_{1:n})}{P(m_{1:n})} \propto \frac{P(s_{1:n}, m_{1:n})}{P(m_{1:n})}
\]

\[
P(s_{1:n}, m_{1:n}) = \underbrace{P(m_n | m_{1:n-1}, s_{1:n}) P(s_{1:n-1} | m_{1:n-2}, s_{1:n}) \ldots}_{\text{Markovian}} \underbrace{P(m_1 | s_{1:n}) P(s_{1:n-1} | s_{1:n-1}) \ldots P(s_2 | s_1) P(s_1)}_{\text{transition prob. matrix}}
\]

\[
P(s_{1:n}, m_{1:n}) = P(m_1 | s_1) P(s_1) \prod_{i=2}^{n} P(m_i | s_i) P(s_i | s_{i-1})
\]

③ Now we want \( P(s_k | m_{1:n}) \)

\[
P(s_k | m_{1:n}) \propto P(s_k, m_{1:n}) = P(s_k, )
\]

\[
= P(m_{KH:n} | ) P( )
\]

\[
= P(m_{KH:n} | ) P(s_k | m_{1:k}) P( )
\]

\[
= P( ) P( )
\]

\[
\downarrow
\]

\[
\downarrow
\]

Probability that is at \( s_k = \text{green st} \) given surveillance camera measurements of main street \( \rightarrow \) Wright street \( \rightarrow \) 6th street

\[
\downarrow
\]

\[
\downarrow
\]

Probability that Neil st. \( \rightarrow \) Kirby road \( \rightarrow \) Lincoln drive \( \rightarrow \) University Avenue, given \( s_k = \text{green street} \).
Walking

<table>
<thead>
<tr>
<th>L₁</th>
<th>L₂</th>
<th>L₃</th>
<th>...</th>
<th>Lᵣ</th>
</tr>
</thead>
<tbody>
<tr>
<td>t₁</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>t₂</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>t₃</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>t₄</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>t₅</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Let's look at the component \( P(s_k | m_{1:k}) \)

\[
P(s_k | m_{1:k}) = \frac{P(s_k, m_{1:k})}{P(m_{1:k})} = \sum_{s_{k-1}} P(s_k | s_{k-1}) P(s_{k-1}, m_{1:k-1})
\]

By marginalizing \( \text{RHS} = \sum_{s_{k-1}} P(s_k | s_{k-1}) P(s_k | s_{k-1}) P(s_{k-1}, m_{1:k-1}) \)

Initial condition \( P(s_1, m_1) \) needs to be known.

Example:

\[
\begin{array}{c}
\text{Error} \\
\downarrow
\end{array}
\begin{array}{c}
s \\
\uparrow
\end{array}
\]
Now let's look at the backward part:

\[
P(m_{k+1:n} \mid s_k) = P(m_{k+1:n}, s_k)
\]

\[
= \frac{1}{P(s_k)} \sum_{s_{k+1}} P(\ldots)
\]

\[
= \frac{1}{P(s_k)} \sum_{s_{k+1}} P(\ldots, P(\ldots)
\]

\[
\sum_{s_{k+1}} P(\ldots) P(\ldots)
\]

Say LHS = \(P(m_{k+1:n} \mid s_k) = \sum_{s_{k+1}} P(\ldots) P(\ldots) P(\ldots)
\]

\[
\beta_k = \sum_{s_{k+1}} P(\ldots)
\]

How should we initialize this \(\beta_k\)?

\[
\beta_{n-1} = P(\ldots) = \frac{1}{\sum_{s_n} P(m_n \mid s_n, s_{n-1})}
\]

\[
\beta_{n-1} = \sum_{s_n} P(\ldots) P(s_n \mid s_{n-1})
\]

\[
\beta_{n-2} = \sum_{s_n}
\]
Recall original goal: $P(s_k | m_1:n) \Rightarrow$ offline version

$$P(s_k | m_1:n) = P(s_k | m_1:k) P(m_{k+1:n} | s_k)$$

$$\alpha_k = \sum_{s_{k-1}} P(m_k | s_k) P(s_k | s_{k-1}) \alpha_{k-1}$$

$$\beta_k = \sum_{s_{k+1}} P(m_{k+1} | s_{k+1}) P(s_{k+1} | s_k)$$

HMM's \Rightarrow identifying the most likely value of $a$ from a huge space of

\Rightarrow Possible to also compute the full trajectory $\Rightarrow$ called
Some other applications of HMM (informal discussion) in smartphone keyboard.

\[ P(m_1 | s_1) = P(\text{ }) = \]

\[ P(s_2 | s_1) = P(\text{ }) = \]

Decodes to

Similar application in

\[ P(m_k | s_k) = P(\text{ }) \]