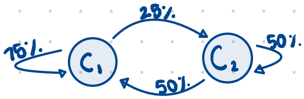
HIDDEN MARKOV MODELS

example scenario:

suppose Alice has 2 coins (C1 and C2) with different biases: $\begin{cases} \mathbb{P}_{C_1}[H] = 30\% \\ \mathbb{P}_{C_1}[T] = 70\% \end{cases}$ $\mathbb{P}_{C_2}[H] = 10\%$, $\mathbb{P}_{C_2}[T] = 90\%$

- hidden states: which coin Alice flips — i.e. $C_1\ C_1\ C_2\ C_1\ C_2\ C_2$
- observations: the resulting heads/tails — i.e. T H T T T T

there are also different probabilities of transitioning from 1 hidden state (coin) to another:



suppose we only see the observations. Given a long enough sequence of observations, we can use statistical algorithms & techniques to infer the underlying sequence of hidden states.

→ HMMS have powerful applications in genomics i.e. the hidden states can be various genes or genetic structures that we can infer from the observed sequence of bases

Markov chains

S = set of states

A = state transition probability matrix

ex| providence weather

$S = \{\text{sunny}, \text{rainy}\}$

$$A = \begin{array}{c|cc} \text{state 1} \backslash \text{state 2} & \text{sunny} & \text{rainy} \\ \hline \text{sunny} & 0.7 & 0.3 \\ \text{rainy} & 0.5 & 0.5 \end{array}$$

→ i.e. $\mathbb{P}$[sunny → rainy transition] = 0.3

$\mathbb{P}$[rainy → sunny transition] = 0.5

Hidden Markov Models: have an added component of different states having different observation/emission probabilities

S = set of states ; V = set of possible observation symbols

A = probability transition matrix → $a_{ij} = \mathbb{P}[q_{t+1} = S_j \mid q_t = S_i]$ (probability of state t+1 being S_j given state t is S_i)

B = emission probability matrix → $b_j(k)$ = probability of emitting symbol k when in state j

π = initial state probability distribution

Alice's coin flip example:

$S = \{C_1, C_2\}$

$V = \{H, T\}$

$$A = \begin{array}{c|cc} & C_1 & C_2 \\ \hline C_1 & 0.75 & 0.25 \\ C_2 & 0.5 & 0.5 \end{array} \qquad B = \begin{array}{c|cc} & H & T \\ \hline C_1 & 0.3 & 0.7 \\ C_2 & 0.1 & 0.9 \end{array} \qquad \pi = [0.5, 0.5]$$

Alice is equally likely to start by flipping coin 1 or 2

Fundamental Computational HMM Problems

▷ **Problem 1: The evaluation problem**

input: HMM defined by $\lambda = \{A, B, \pi\}$ and an observation sequence $O = o_1 o_2 o_3 \ldots o_T$

output: the probability of observing the sequence O given the HMM λ: $\mathbb{P}(O \mid \lambda)$

▷ **Problem 2: The decoding problem**

input: HMM defined by $\lambda = \{A, B, \pi\}$ and an observation sequence $O = o_1 o_2 o_3 \ldots o_T$

output: the most likely sequence of states $Q = q_1 q_2 \cdots q_T$ given the observation sequence O. The solution can be found using the Viterbi Algorithm

▷ **Problem 3: The Learning Problem** * not too worried about this problem in this class

input: observation sequence O

output: find the most likely HMM defined by $\lambda = \{A, B, \pi\}$ given O

Solution to problem 1

ex: observations: a a b c c c $\Rightarrow |O| = 6$

if we have N states, there are N^6 possible sequences of states underlying the observation sequence O

let's define a few fundamental probabilities:

$\mathbb{P}$[observing O given a sequence of states Q]: $\mathbb{P}[O \mid Q] = \mathbb{P}(o_1 \mid q_1)\, \mathbb{P}(o_2 \mid q_2) \ldots \mathbb{P}(o_T \mid q_T)$

$= b_{q_1}(o_1)\, b_{q_2}(o_2) \ldots b_{q_T}(o_T)$

$\mathbb{P}$[having a particular sequence of states Q]: $\mathbb{P}[Q] = \pi_{q_1} a_{q_1 q_2} a_{q_2 q_3} \ldots a_{q_{T-1} q_T}$

Joint probability of O and Q: $\mathbb{P}[O, Q] = \mathbb{P}[O \mid Q] \cdot \mathbb{P}[Q]$

To be continued...