Markov Process: Time Invariance, Finite State, Transition Matrix

Definition: Markov process

A stochastic process is a Markov process (or: Markov chain) if for all $n \in \mathbb{N}$,

$$X_1 \to X_2 \to \cdots \to X_n.$$

In a Markov process, the value of each step can only depend on the previous value, similarly to the three-variable Markov chains from earlier. The infinite sequence of coin tosses from earlier is an example of a Markov process: you can see this by inspecting the definition of $P_{X_{n+1}|X_1 \cdots X_n}$, and noting that it is indeed independent of the values for X_1 to X_{n-1}. The process in this example even fulfills a stronger property: it is time invariant.

Definition: Time-invariant Markov process

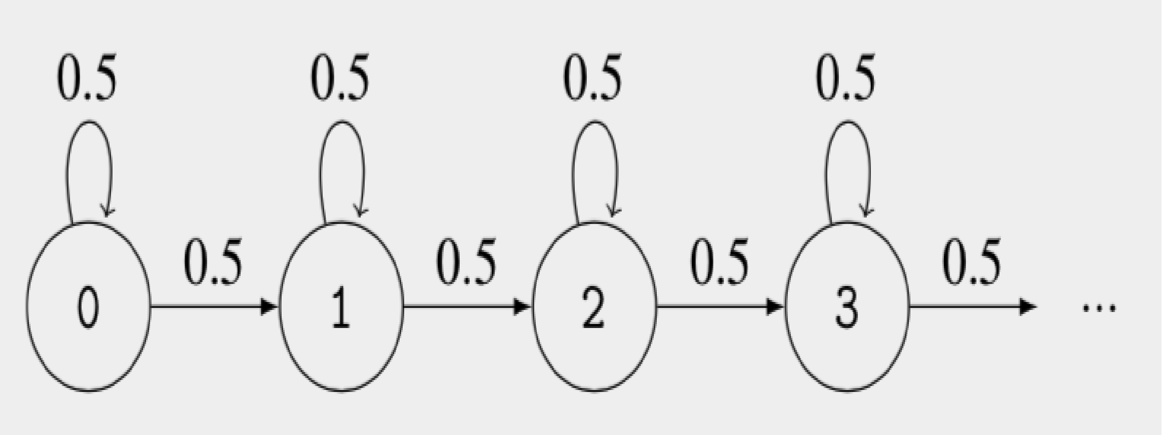
A Markov process is time invariant if for all $n \in \mathbb{N}$, and for all $a, b \in \mathcal{X}$,

$$P_{X_{n+1}|X_n}(a|b) = P_{X_2|X_1}(a|b) \qquad \text{whenever } P_{X_n}(b) > 0 \text{ and } P_{X_1}(b) > 0.$$

Time-invariant Markov processes can be nicely visualized using state diagrams, where the **states** represent the values in $\mathcal{X}$, and the labels on the arrows represent the conditional probabilities. A time-dependent Markov process could also be visualized in this way, but the labels on the arrows would have to be functions of the time step i. In both cases, one also needs to specify P_{X_1}, the **initial distribution**, in order to completely describe the stochastic process.

Example 1: Repeatedly tossing a fair coin, continued

The infinite sequence of coin tosses is represented by the following state diagram:



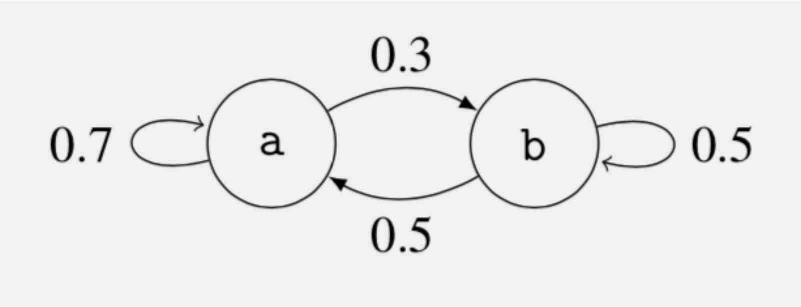
For example, the arrow from state 2 to state 3, represents the probability $P_{X_{n+1}|X_n}(3|2) = 0.5$. The initial distribution is $P_{X_1}(0) = P_{X_1}(1) = 0.5$: the process is equally likely to start out in state 0 (if the first toss is a tails) or state 1 (if the first toss is a heads).

Definition: Finite-state Markov process

A Markov process is finite-state if $|\mathcal{X}|$ is finite.

Example 2: A finite-state time-invariant Markov process

Consider the following state diagram, for a Markov process with initial distribution $P_{X_1}(\mathtt{a}) = 1$ and $P_{X_1}(\mathtt{b}) = 0$:



The process starts in state a with probability one. A possible run of the process would be $\mathtt{aabaaabbabaa}\cdots$.

Note that in the state diagrams, the probabilities of the outgoing arrows add up to 1 for every state. This is necessary for a well-defined Markov process. A time-invariant Markov process can alternatively be represented by its initial distribution combined with a **transition matrix** R, where the entry R_{ij} represents the transition probability $P_{X_{n+1}|X_n}(j|i)$. For a finite-state process, the transition matrix is finite. In the transition matrix, the row entries have to sum up to 1. When the current state is given by a row vector v_n, the state after one step of the Markov process is given by $v_{n+1} = vR$. If you (like me) don't like multiplying vectors from the left with matrices, you can also compute $v_{n+1} = \left(R^T v^T\right)^T$ where $(\cdot)^T$ denotes the transposition of matrices and vectors.

Example 2: A finite-state time-invariant Markov process, continued

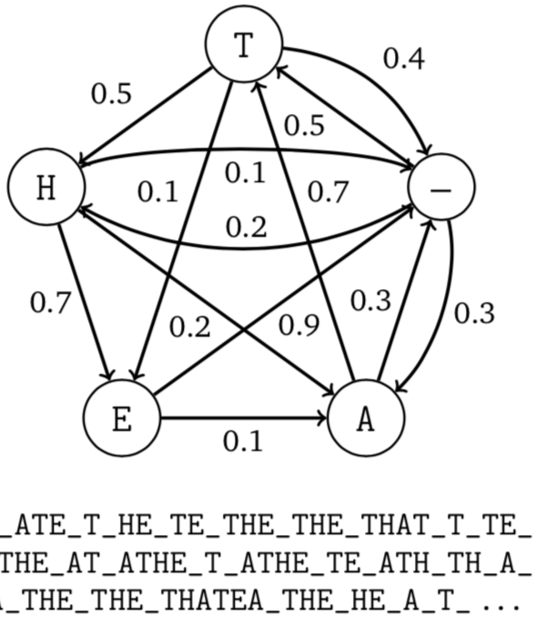
The matrix representation of the process above is given by

$$R = \begin{bmatrix} 0.7 & 0.3 \\ 0.5 & 0.5 \end{bmatrix}$$

where the state `a` is represented by the first row/column, and the state `b` by the second. Verify that the row entries indeed sum up to 1. The initial distribution is still given by $P_{X_1}(\mathtt{a}) = 1$ and $P_{X_1}(\mathtt{b}) = 0$.

Another example: primitive language model

Observe how this very simple 5-state Markov chain produces samples that resemble English language. This is a first hint of how surprisingly powerful Markov models can be.



(image by Mathias Madsen)