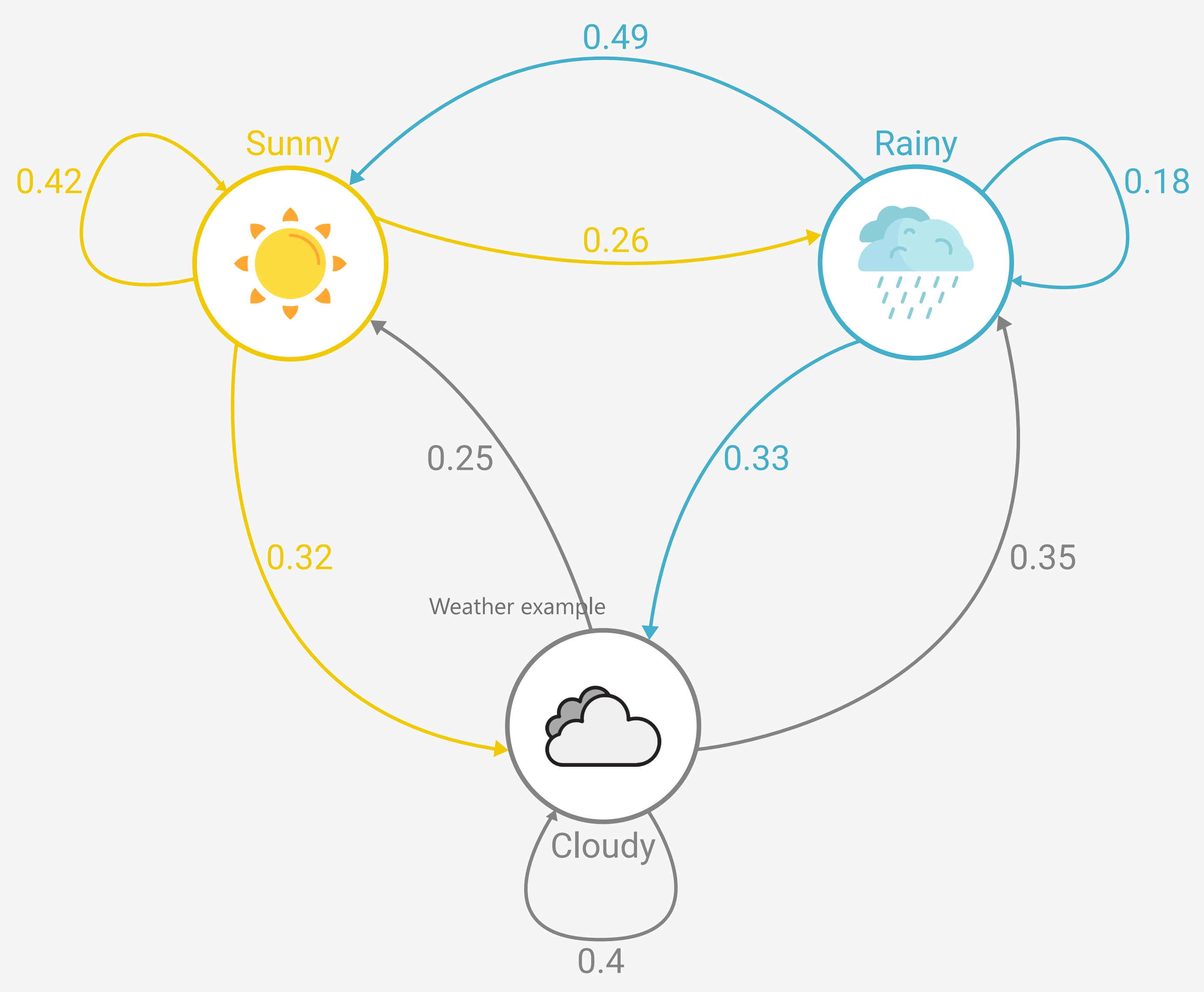
← Back

Weather example

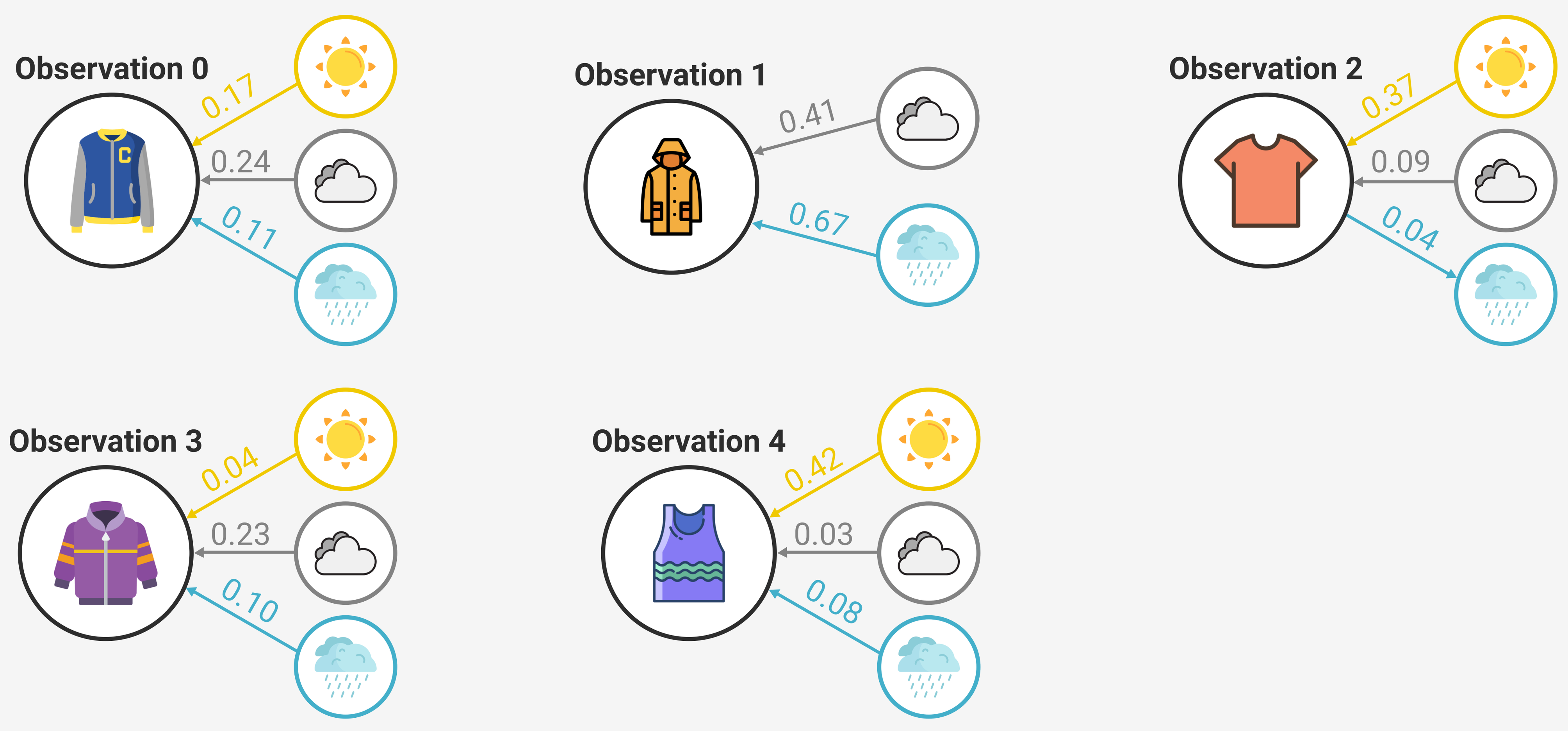
Probabilities of state transitions



Intuitive meaning

- If it's sunny at t_x, the probability of sun at t_{x+1} is 0.42, the probability of clouds at t_{x+1} is 0.32 and the probability of rain at t_{x+1} is 0.26.
- If it's rainy at t_x, the probability of sun at t_{x+1} is 0.49, the probability of clouds at t_{x+1} is 0.33 and the probability of rain at t_{x+1} is 0.18.
- If it's cloudy at t_x, the probability of sun at t_{x+1} is 0.25, the probability of clouds at t_{x+1} is 0.4 and the probability of rain at t_{x+1} is 0.35.

State probabilities given observation



Intuitive meaning

- If the HMM is in the sunny state, the probability of observing **observation 0** is 0.17, **observation 2** 0.37, **observation 3** 0.04, and **observation 4** 0.42.
- If the HMM is in the cloudy state, the probability of observing **observation 0** is 0.24, **observation 1** 0.41, **observation 2** 0.09, **observation 3** 0.23, and **observation 4** 0.03.
- If the HMM is in the rainy state, the probability of observing **observation 0** is 0.11, **observation 1** 0.67, **observation 2** 0.04, **observation 3** 0.10, and **observation 4** 0.05.

HMM parameters

- **Initial state distribution** As described in the *basic concepts* section, the initial state distribution declares the probabilities of being in a certain state at the initial time step t_0. In this example, the initial state distribution could be completely different depending on what knowledge we have about the weather. For now, let's just accept that we have the initial state distribution:

  $$\pi(i) = \begin{bmatrix} 0.435 & 0.125 & 0.440 \end{bmatrix}$$

  Since we haven't yet defined in what order the states should be considered, it's time we do that now! Let's say that **S** = {☀, ☁, 🌧}, that is, ☀ = S_0, ☁ = S_1 and 🌧 = S_2. Now we have the information we need to express the initial state distribution in the following way:

  $$\pi(☀) = \pi(S_0) = P(S_0 = S_0) = 0.435$$
  $$\pi(☁) = \pi(S_1) = P(S_0 = S_1) = 0.125$$
  $$\pi(🌧) = \pi(S_2) = P(S_0 = S_2) = 0.440$$

  So what does this mean? In English we read the above notations as: *"at the initial time step t_0 the probability that the weather is sunny is 0.435, cloudy 0.125, and rainy 0.440"*.

- **Transition matrix** We use the transition matrix A, or $a_{i,j}$, to describe the probabilities of transitioning between states. Note that it's possible for the HMM to transition from a certain state to itself. E.g. A[2][2] would be imply the probability $P(S_t = S_2 \mid S_{t-1} = S_2)$ which is a valid transition. In this weather example, we have the following transition matrix:

  $$a_{i,j} = \begin{bmatrix} 0.42 & 0.32 & 0.26 \\ 0.25 & 0.40 & 0.35 \\ 0.49 & 0.33 & 0.18 \end{bmatrix}$$

  To make things a little clearer, let's have a look at the following figure which also describes the transition probabilities:

A	☀	☁	🌧
☀	0.42	0.32	0.26
☁	0.25	0.40	0.35
🌧	0.49	0.33	0.18

  From this transition matrix we can see that $a_{2,1} = P(S_t = S_1 \mid S_{t-1} = S_2) =$ $= P(S_t =$ ☁ $\mid S_{t-1} =$ 🌧 $) = 0.33$. Also, note that all rows sum up to one!

- **Observation matrix** We use the observation matrix B, or $b_{j,k}$, to describe probabilities of making specific observations given that the hidden model is in a particular state. E.g. what is the probability that we observe a 🎽 given that the weather is sunny? In other words what's $P(O_t = O_4 \mid S_t = S_0) = P(O_t =$ 🎽 $\mid S_t =$ ☀ $)$? The answer is 0.42. By looking at all the different states and their probabilities in relation to possible observations we get the observation matrix B:

  $$b_{i,j} = \begin{bmatrix} 0.17 & 0.00 & 0.37 & 0.04 & 0.42 \\ 0.24 & 0.41 & 0.09 & 0.23 & 0.03 \\ 0.11 & 0.67 & 0.04 & 0.10 & 0.08 \end{bmatrix}$$

  For better understanding we can have a look at the following figure, which we use to clearify the observation matrix:

B	🧥	🧥	👕	🧥	🎽
☀	0.17	0.00	0.37	0.04	0.42
☁	0.24	0.41	0.09	0.23	0.03
🌧	0.11	0.67	0.04	0.10	0.08

  From this observation matrix we can see that $b_{1,0} = P(O_t = O_0 \mid S_t = S_1) =$ $= P(O_t =$ 🧥 $\mid S_t =$ ☁ $) = 0.24$. Also, note that all rows sum up to one!

- **Lambda** The initial state distribution π, the transition matrix A and the observation matrix B together describes how the weather behaves and what observations we can expect when being in some of the states. At last, we can gather the three in $\lambda = \{A, B, \pi\}$.

Food for thought *Why does the observation matrix need to be row-stochastic?* To grasp this question we must first accept that each HMM is a simplification of reality. We make simplifications because it allows us to control all the variables in the system we are interested in. When an HMM are in a certain state, we *know* that the model will make some observation, it's *defined* that way. We have also *defined* that all possible observations that can be made are **O** = {🧥, 🧥, 👕, 🧥, 🎽}. Therefore, in order for us to *know* that we make *some* observation, all separate possibilities for observations must together sum up to 1.0. The same reasoning can be made for the initial state distribution π and for the transition matrix A.