Suppose there are 2 states in the world, sunny and rainy. Further, suppose

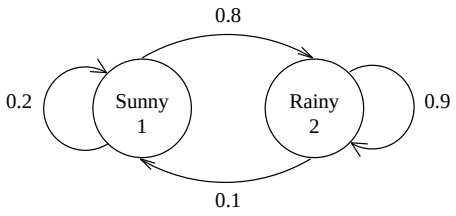
- If it is sunny today, it will still be sunny tomorrow with probability 0.2, and it will be rainy tomorrow with probability 0.8.
- If it is rainy today, it will still be rainy tomorrow with probability 0.9, and it will be sunny tomorrow with probability 0.1.

We are interested in the following questions:

1. What is the probability of a day being sunny?
2. If today is rainy, on average how long does it take to have a sunny day?

One powerful way to answer such questions is by Markov analysis.

We can express the above state transitions in a state diagram:



Note that for each state, the sum of the labels of the outgoing edges is 1.
Now let p_i be the probability that we are in state i. We want to calculate p_1 and p_2.
There are two ways we can go to state 1, either from state 1 or from state 2:

- From state 1, we go to state 1 with probability 0.2.
- From state 2, we go to state 1 with probability 0.1.

So,

$$p_1 = 0.2p_1 + 0.1p_2.$$

Similarly,

$$p_2 = 0.8p_1 + 0.9p_2.$$

It seems by solving the equations, we could get the values of p_1 and p_2.

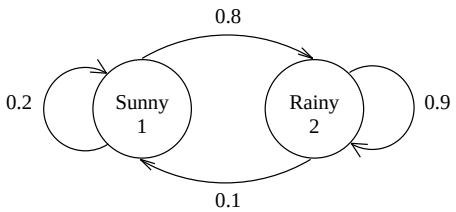
$$p_1 = 0.2p_1 + 0.1p_2.$$

$$p_2 = 0.8p_1 + 0.9p_2.$$

However, the two equations don't have a unique solution. They are dependent (homogeneous). $p_1 = p_2 = 0$ is one solution, but not the one we want. Also, any multiple of a solution is also a solution. To get the right answer, we need one more equation.

$$p_1 + p_2 = 1.$$

Now we have 2 independent equations in 2 unknowns. We can solve them and get $p_1 = 1/9$ and $p_2 = 8/9$. (Too good to be true!)



Now, we want to answer the question: If today is rainy, on average how long does it take to have a sunny day?

Let e_i be the expected time until entering state 1 given we are now in state i. Clearly $e_1 = 0$. What is e_2?

$$e_2 = 1 + 0.9e_2 + 0.1e_1.$$

Solving, we have $e_2 = 10$, which is the answer.

This is called a (discrete time, finite state) *Markov chain.*

The Markov property is often informally stated as follows: *given the present, the future becomes independent of the past.*

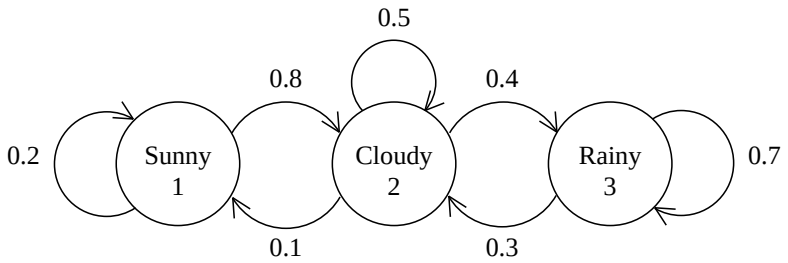
In the first problem, we looked for the *equilibrium distribution.* **We looked at the** *in*-**edges.**
In the second problem, we looked for the *expected time to absorption.* **We looked at the** *out*-**edges.**

Another way to represent Markov transitions is by a one-step Markov transition matrix. For example

$$\begin{array}{c} \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{cc} 1 & 2 \end{array} \\ \begin{pmatrix} 0.2 & 0.8 \\ 0.1 & 0.9 \end{pmatrix} \end{array}$$

Note that each row sums to 1.
Then calculating the equilibrium distribution and expected time to absorption involve solving a linear system with the one-step Markov transition matrix.

To make the problem less trivial, let's introduce one more state.



Suppose we want to answer similar questions as before. Let p_i **and** e_i **have the same meanings as before. To calculate the equilibrium distribution:**

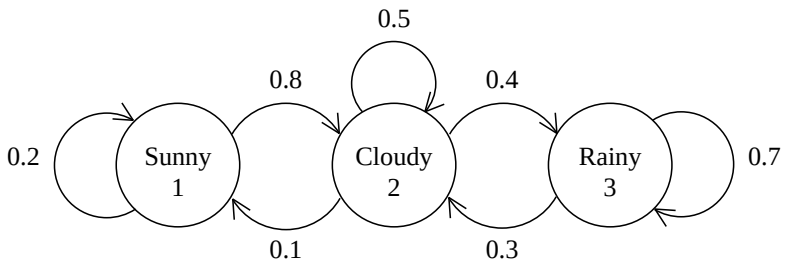
$$p_1 = 0.2p_1 + 0.1p_2.$$

$$p_2 = 0.8p_1 + 0.5p_2 + 0.3p_3.$$

$$p_3 = 0.4p_2 + 0.7p_3.$$

$$p_1 + p_2 + p_3 = 1.$$

Again the first 3 are homogeneous, so we need the fourth equation. The answer is left as an exercise.



We can calculate the e_i as before:

$$e_1 = 0$$

$$e_2 = 1 + 0.1e_1 + 0.5e_2 + 0.4e_3$$

$$e_3 = 1 + 0.3e_2 + 0.7e_3$$

Given today is not sunny, what is the expected time it takes to have sunny day? Recall that e_i is the expected time until entering state 1 given we are now in state i. So the answer to the question is neither e_2 nor e_3, since we do not know which state we are in. But we know the equilibrium distribution.
The answer is the weighted mean of e_2 and e_3,

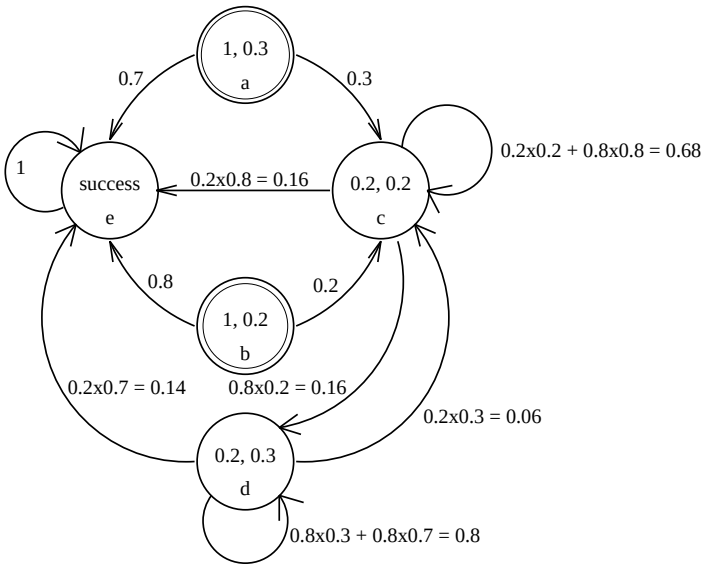
$$\frac{p_2}{p_2 + p_3} e_2 + \frac{p_3}{p_2 + p_3} e_3.$$

The Markov chain model is very general and has a lot of applications:

- Population model
- Genetic model
- Customer arrival
- **Network analysis**

Recall the taxi-driver problem in homework. Each driver says something with probability 0.3. If a colission occurs, both try to retransmit in the successive slots with probability 0.2 until a successful transmission. (After a successful transmission, the driver again says something with probability 0.3.) What is the expected time required to have one successful transmission.

We can model it as a Markov chain. There are probably many ways you can do this. I find the following chain easier to explain:



Each state (except the success state) is a pair of probabilities. The first is the probability I say something, and the second is the probability the other driver says something. Initial states are $(1, 0.3)$ and $(1, 0.2)$.

Define: e_i is the expected time to go to state e given the current state is i. We can calculate e_i as before:

$$e_a = 1 + 0.7e_e + 0.3e_c.$$

$$e_b = 1 + 0.8e_e + 0.2e_c.$$

$$e_c = 1 + 0.16e_e + 0.68e_c + 0.16e_d.$$

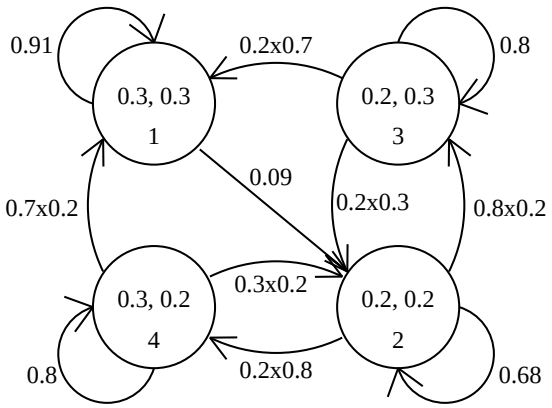
$$e_d = 1 + 0.14e_e + 0.06e_c + 0.8e_d.$$

$$e_e = 0.$$

The answers I got are $e_a = 2.985, e_b = 2.324, e_c = 6.618, e_d = 6.985, e_e = 0.$ (not guaranteed to be correct)

But again, we don't know the exact initial state. It can either be state a or state b.

We need to calculate the equilibrium distribution by analyzing another chain.



As before, solve for p_1, p_2, p_3 and p_4. I got $p_1 = 112/229, p_2 = 45/229, p_3 = p_4 = 36/229$.

The final answer is then just

$$\frac{p_1}{p_1 + p_4} e_a + \frac{p_4}{p_1 + p_4} e_b = 2.824.$$