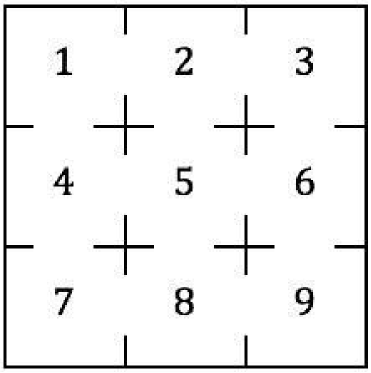

Analytical Solution

Pi - probability that AIM reaches state 3 before state 7, given that AIM started in compartment i, i ∈ {1,2,3,4,5,6,7,8,9}

Here are the immediate solutions to Pi

- **P3 = 1**
- **P7 = 0**
- **P1,P5,P9 = 0.5** (because of symmetry, the random experiments will converge towards this value. Assumption here is that AIM reaches either 3 or 7 in all or most of the experiments)
- **P2 = P6**, **P4 =P8** and (**P2 = 1 - P4**) (symmetry)
- I don't have any concrete estimate of i ∈ {2,4,6,8} at this point

Simulation

I assumed that AIM takes a random walk with equal probabilities in each possible move. E.g. it will move to 1,5 or 3 from 2 with Probability of ⅓ each. Similarly, it will move to either 2 or 4 from 1 with Probability of ½ each. I have Markov Chain for the simulation. You can find the code in attachments of the email (script and jupyter notebook).

Big-O Notation = O(max_steps . number_of_runs)
(Assumption - I'm using a fixed size array for state_dict as I've not written a scalable solution)

Here's how the output looks like -

```
P 1 = 0.48
P 2 = 0.66
P 3 = 1.0
P 4 = 0.32
P 5 = 0.51
P 6 = 0.68
P 7 = 0.0
P 8 = 0.35
P 9 = 0.5
```

P2 and P6 are ⅔ whereas P4 and P8 are ⅓.

Analytical Solution Continued

Analytical Solution for P where $i \in \{2,4,6,8\}$

We can split the problem down by moving one random step from the desired cell.
For example,

P2 = ⅓ * (P1 + P5 + P3)
P2 = ⅓ * (0.5+0.5+1)
P2 = ⅔

Similarly, P6 can be derived to be ⅔ as well. And as P7 = 0, P4 and P8 can be derived to be ⅓ each. This completes the analytical solution.

Part (3)

Question : Suppose we drop AIM in some fixed compartment $i \in \{1,2,4,5,6,8,9\}$ 100 times and record the number of times that AIM reaches state 3 before state 7. How would you evaluate if AIM is actually learning to go through this maze? Provide as much detail you feel necessary to justify you answer.

Answer:

1. Repeat the experiment with 100 repetitions or we can split the experiment into multiple smaller experiments (such as 5 experiments with 20 runs or 2 with 50 etc)
2. Calculate the Pi for each of the experiment
3. We should be able to observe increase in the probability for successive experiments. For example $P_{in} < P_{i(n+1)}$

where i = the compartment number, n = number of experiment

For this experiment we'll have to exclude first step for i=7 as our simulation will always end at step 0 with negative outcome.

To fix this issue, we can modify the way we run the simulation. We'll run simulation until the mouse reaches 3 and we'll evaluate each simulation by seeing if it has passed through 7 before reaching 3 or not (or how many times). But I'm not proposing at the main solution because it can be computationally more expensive and we might need to increase the max_steps.