**Professor Phillips and the Squirrel in the Attic**

Recently, Prof Phillips discovered that a squirrel had gotten into his attic and was living up there and making a lot of noise. Help Prof Phillips discover where in the attic the squirrel is.

| | | |
|---|---|---|
| A | B | C |

The attic is divided into three sections, which we will call A, B, and C.
The squirrel is always in exactly one of the A/B/C locations. Every minute, the squirrel may choose to move to an adjacent location within the attic or stay where it is. The squirrel will choose to stay where it is with probability 0.3, and will choose to move to an adjacent location with probability 0.7. If there are two possible adjacent locations, the 0.7 will be divided equally between them.

Because the squirrel is in the attic, Prof Phillips can't know exactly where it is. However, he hears the squirrel best when it's at location B, and doesn't hear it as well when it's at location A or C. We will say that if the squirrel at location B, he hears it with probability 0.9. If the squirrel is at location A or C, he hears it with probability 0.4.

Assume that for at minute 1, Prof Phillips hears the squirrel, and at minutes 2 and 3, he doesn't hear the squirrel. Assume that the squirrel begins (at minute 0) in location A (because that's where Prof Phillips found a hole leading to the outside).

1. What would this problem look like modelled with a Markov Chain? What states are hidden, and what evidence is known?
2. Use the forward algorithm to find the most likely location of the squirrel at minute 3. (Begin by drawing the state transition diagram, the transition matrix, and the observation matrices.)

3. Use the forward-backward algorithm to find the most likely location of the squirrel at minutes 0, 1, 2, and 3.

Pseudocode

**$f_{t:0}$** is the probability distribution moving from time 0 to time t (the forward part of the algorithm).
**$b_{t+1:t}$** is the probability distribution moving from time t + 1 to time t (the backward part of the algorithm).
**$b_{t+1:t}$** starts as a column of all 1s and then changes each time you move back a state.
**T** is the probability matrix that governs the transitions between states.
**$O_t$** is the observation matrix given based on the evidence we observe at time t.

Some notes:
-For matrices, multiplication order matters!
-To derive accurate probability, you need both the forward distribution and the backward distribution.
-Note that the matrices these formulas produce will need to be normalized in order to remove alpha.

# Forward-backward algorithm

$$f_{1:0} = P(X_0)$$

$$f_{1:t+1} = \alpha f_{1:t} \cdot T \cdot O_{t+1}$$

Compute these forward from $X_0$ to wherever you want to stop ($X_t$)

$$b_{t+1:t} = [1; \cdots ; 1]$$

$$b_{k:t} = T \cdot O_k \cdot b_{k+1:t}$$

$$P(X_k \mid e_{1:t}) = \alpha f_{1:k} \times b_{k+1:t}$$

Compute these backwards from $X_{t+1}$ to $X_0$.