Backward algorithm

- Updates previous probabilities to take into account new evidence.
- Calculates $P(X_k \mid e_{1:t})$ for $k < t$
  - aka **smoothing**. (not the same kind of smoothing as in Naïve bayes)

Backward algorithm

- Algorithm generates a *backward vector* b for every timestep t.
  - This vector is based on the observation at time k and the *next day's* backward vector.

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

  - The initial backwards vector is for day t+1 and is a column vector of all 1's.

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

Backwards algorithm

- Each backward vector is used to *scale* the previous day's forward vector.
- After normalization, this is the updated probability for day k.

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

- (Remember, that multiplication above is an item by item multiplication, not a matrix multiplication.)

Backward matrices

- Main equations:

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

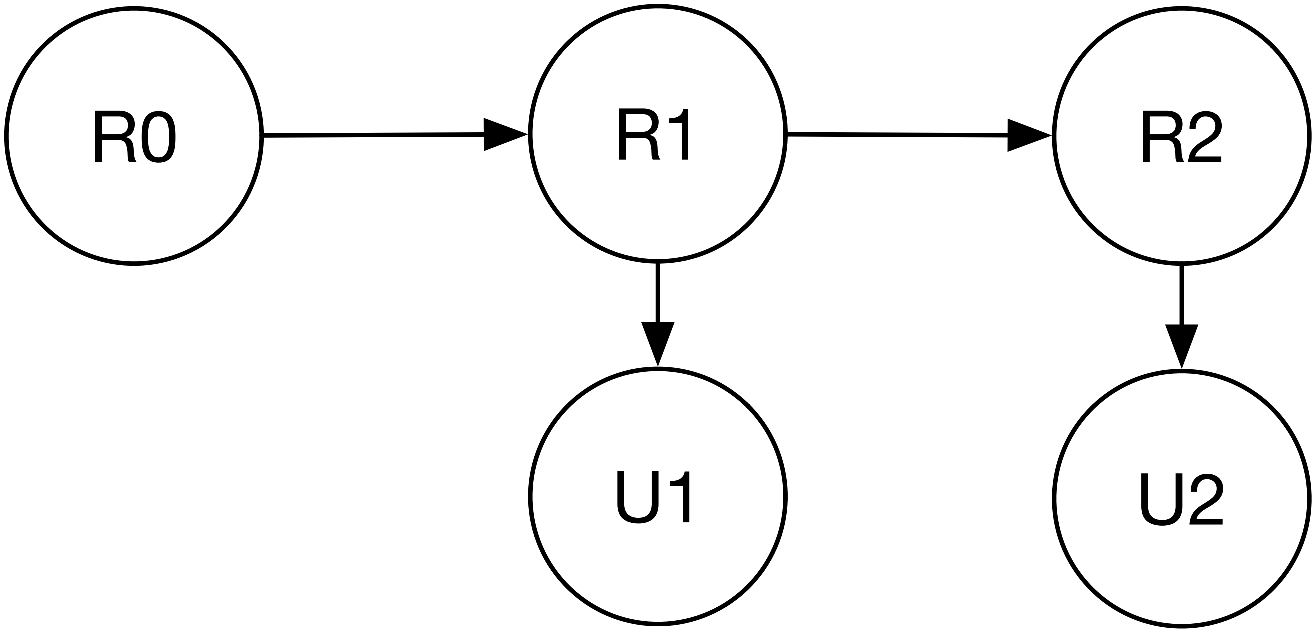
$$b_{t+1:t} = [1; \cdots ; 1]$$ (column vector of 1s)

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

f1:0=[0.5, 0.5] f1:1=[0.75, 0.25] f1:2=[0.846, 0.154]

b1:2=[0.4509, 0.1107] b2:2=[0.69, 0.27] b3:2=[1; 1]

mult=[0.803, 0.197] mult=[0.885, 0.115]




```
T  = [0.7, 0.3]
     [0.1, 0.9]

O1 = [0.9, 0.0]
     [0.0, 0.2]

O2 = [0.9, 0.0]
     [0.0, 0.2]
```

b3:2 = [1; 1]

b2:2 = T * O2 * b3:2 = [0.69, 0.27]

P(R1 | u1, u2) = α f1:1 x b2:2 = α[0.5175, 0.0675] = [0.885, 0.115]

b1:2 = T * O1 * b2:2 = [0.4509, 0.1107]

P(R0 | u1, u2) = α f1:0 x b1:2 = α[0.2255, 0.0554] = [0.803, 0.197]

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.