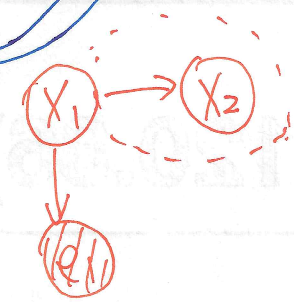
HMM



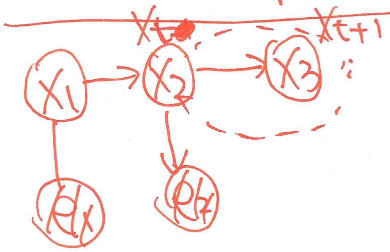
Given $P(X_t \mid e_{1:t})$

also $P(X_0) \cdot P(X_{t+1} \mid X_t) \cdot P(e_t \mid X_t)$

How to get $P(X_{t+1} \mid e_{1:t})$?

Base

$$P(X_1 \mid e_1) = \frac{P(X_1, e_1)}{P(e_1)} \propto P(X_1, e_1) = P(X_1)\, P(e_1 \mid X_1) \qquad P(X_2) = \sum_{x_1} P(X_1, X_2) = \sum_{x_1} P(X_1)\, P(X_2 \mid X_1)$$



$$P(X_{t+1} \mid e_{1:t}) = \sum_{x_t} P(X_{t+1}, x_t \mid e_{1:t})$$

$$= \sum_{x_t} \underline{P(X_{t+1} \mid x_t, e_{1:t})}\, P(x_t \mid e_{1:t})$$

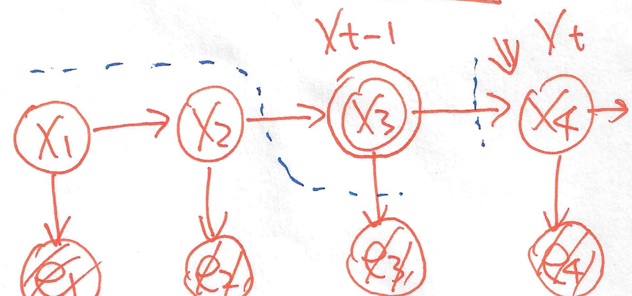
↓ HMM independence assumption

$$= \sum_{x_t} \underline{P(X_{t+1} \mid x_t)}\, P(x_t \mid e_{1:t})$$

Independence

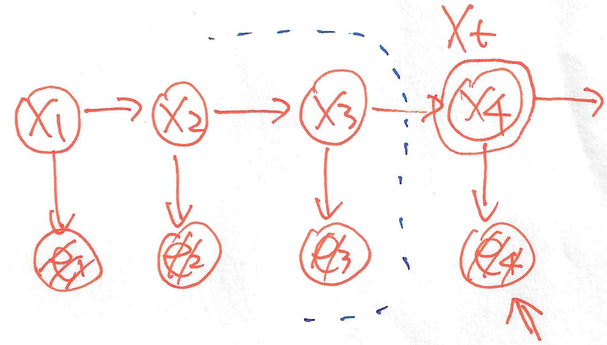
$X_t \perp\!\!\!\perp X_1, \cdots X_{t-2} \mid X_{t-1}$

$X_t \perp\!\!\!\perp E_1, \cdots E_{t-1} \mid X_{t-1}$



$E_t \perp\!\!\!\perp X_1, \cdots, X_{t-1} \mid X_t$

$E_t \perp\!\!\!\perp E_1, \cdots, E_{t-1} \mid X_t$



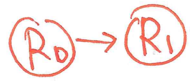
$P(X_t \mid X_{0:t-1}) = P(X_t \mid X_{t-1})$

$P(E_t \mid X_{0:t}; E_{0:t-1}) = P(E_t \mid X_t)$

Weather HMM

$B_0(+r, -r) = P(R_0) = \langle 0.5, 0.5 \rangle$ initialization

↓ time pass



$$B_0'(+r, -r) = P(R_1) = \sum_{r_0} P(R_1 | r_0) P(r_0)$$

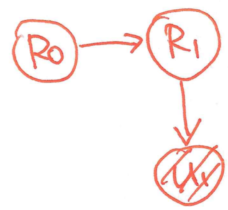
$$= \langle 0.7, 0.3 \rangle \times 0.5 + \langle 0.3, 0.7 \rangle \times 0.5$$

$$= \langle 0.5, 0.5 \rangle$$

↓ observe an evidence $U_1 = \text{true}$ (umbrella appears)

$$B_1(+r, -r) = P(R_1 | u_1) = \frac{P(R_1, u_1)}{P(u_1)}$$

$$\propto_{R_1} P(R_1, u_1)$$



$$= P(u_1 | R_1) P(R_1)$$

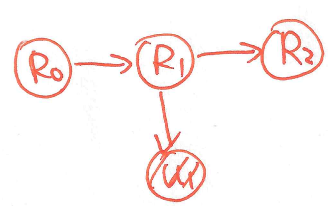
$$= \langle 0.9, 0.2 \rangle \langle 0.5, 0.5 \rangle$$

$$= \langle 0.45, 0.1 \rangle$$

normalize ↓

$$\approx \langle 0.818, 0.182 \rangle$$ so. sum to 1

↓ time pass



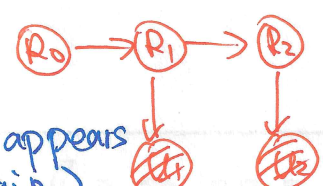
$$B_1'(+r, -r) = P(R_2 | u_1) = \sum_{r_1} P(R_2 | r_1) P(r_1 | u_1)$$

$$= \langle 0.7, 0.3 \rangle \times 0.818 + \langle 0.3, 0.7 \rangle \times 0.182$$

$$\approx \langle 0.627, 0.373 \rangle$$

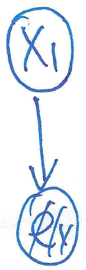
↓ observe another evidence

$U_2 = \text{true}$ (umbrella appears again)



$$B_2(+r, -r) = P(R_2 | u_1, u_2) = \ldots$$

Inference : Base Cases



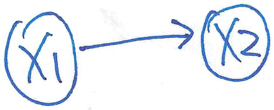
$P(X_1)$ to incorporate evidence.

$P(E_1|X_1)$

vector — a specific value — reinvented Baynes Rule

$$P(X_1|e_1) = \frac{P(X_1, e_1)}{P(e_1)} = \frac{P(e_1|X_1)\,P(X_1)}{P(e_1)}$$

$P(X_1)$ $P(X_2|X_1)$



to advance time

$$P(X_2) = \sum_{x_1} P(X_2, x_1) = \sum_{x_1} P(X_2|X_1)\,P(X_1)$$