1 PROBABILISTIC REASONING OVER TIME

Deriving the expressions for filtering, prediction, and smoothing.

For this we use three simple formulas:

1. Chain rule:
$$P(A,B) = P(A|B)P(B)$$
2. Conditional probability:
$$P(A|B) = \frac{P(A,B)}{P(B)}$$
3. Marginal probability:
$$P(A) = \sum_{b \in B} P(A,b)$$

Also some simplifications from the definition of model

1. Transition model:
$$P(X_{t+1}|X_{0:t}, E_{1:t}) = P(X_{t+1}|X_t)$$
2. Sensor model:
$$P(E_t|X_{0:t}, E_{1:t-1}) = P(E_t|X_t)$$
3. Sensor model extended:
$$P(E_t|X_{0:k}, E_{1:k-1}) = P(E_t|X_k)$$
for $k \leq t$

2 FILTERING

Filtering is defined as $P(X_{t+1}|e_{1:t+1})$ and is derived as follows:

Using conditional probability

$$P(X_{t+1}|e_{1:t+1}) = \frac{P(X_{t+1}, e_{1:t+1})}{P(e_{1:t+1})} \tag{2.1}$$

separate $e_{1:t+1}$ into $e_{1:t}, e_{t+1}$

$$P(X_{t+1}, e_{1:t+1}) = P(X_{t+1}, e_{1:t}, e_{t+1}) \tag{2.2}$$

using the chain rule

$$P(X_{t+1}, e_{1:t+1}) = P(e_{t+1}|X_{t+1}, e_{1:t})P(X_{t+1}, e_{1:t}) \tag{2.3}$$

using the sensor model

$$P(X_{t+1}, e_{1:t+1}) = P(e_{t+1}|X_{t+1})P(X_{t+1}, e_{1:t}) \tag{2.4}$$

using the chain rule

$$P(X_{t+1}, e_{1:t+1}) = P(e_{t+1}|X_{t+1})P(X_{t+1}|e_{1:t})P(e_{1:t}) \tag{2.5}$$

Then using the marginal probability

$$P(X_{t+1}, e_{1:t}) = \sum_{x_t \in X_t} P(X_{t+1}, x_t, e_{1:t}) \tag{2.6}$$

applying the chain rule

$$P(X_{t+1}, e_{1:t}) = \sum_{x_t \in X_t} P(X_{t+1}|x_t, e_{1:t})P(x_t, e_{1:t}) \tag{2.7}$$

applying the chain rule once more

$$P(X_{t+1}, e_{1:t}) = \sum_{x_t \in X_t} P(X_{t+1}|x_t, e_{1:t})P(x_t|e_{1:t})P(e_{1:t}) \tag{2.8}$$

take $P(e_{1:t})$ outside the summation

$$P(X_{t+1}, e_{1:t}) = P(e_{1:t}) \sum_{x_t \in X_t} P(X_{t+1}|x_t, e_{1:t})P(x_t|e_{1:t}) \tag{2.9}$$

Then using conditional probability

$$P(X_{t+1}|e_{1:t}) = \frac{P(X_{t+1}, e_{1:t})}{P(e_{1:t})} \tag{2.10}$$

replace 2.9 into 2.10

$$P(X_{t+1}|e_{1:t}) = \frac{P(e_{1:t}) \sum_{x_t \in X_t} P(X_{t+1}|x_t, e_{1:t})P(x_t|e_{1:t})}{P(e_{1:t})} \tag{2.11}$$

cancel $P(e_{1:t})$ and use the transition model

$$P(X_{t+1}|e_{1:t}) = \sum_{x_t \in X_t} P(X_{t+1}|x_t)P(x_t|e_{1:t}) \tag{2.12}$$

then replace 2.12 into 2.5

$$P(X_{t+1}, e_{1:t+1}) = P(e_{t+1}|X_{t+1})P(e_{1:t}) \sum_{x_t \in X_t} P(X_{t+1}|x_t)P(x_t|e_{1:t}) \tag{2.13}$$

then replace 2.13 into 2.1

$$P(X_{t+1}|e_{1:t+1}) = \frac{P(e_{t+1}|X_{t+1})P(e_{1:t}) \sum_{x_t \in X_t} P(X_{t+1}|x_t)P(x_t|e_{1:t})}{P(e_{1:t+1})} \tag{2.14}$$

then we define $\alpha = \frac{P(e_{1:t})}{P(e_{1:t+1})}$ and finally obtain

$$P(X_{t+1}|e_{1:t+1}) = \alpha P(e_{t+1}|X_{t+1}) \sum_{x_t \in X_t} P(X_{t+1}|x_t)P(x_t|e_{1:t}) \tag{2.15}$$

3 Prediction

Prediction is defined as $P(X_{t+k}|e_{1:t})$ where $k > 0$ and is derived as follows:

Using conditional probability

$$P(X_{t+k}|e_{1:t}) = \frac{P(X_{t+k}, e_{1:t}))}{P(e_{1:t})} \tag{3.1}$$

using marginal probability

$$P(X_{t+k}, e_{1:t}) = \sum_{x_{t+k-1}} P(X_{t+k}, x_{t+k-1}, e_{1:t}) \tag{3.2}$$

using the chain rule

$$P(X_{t+k}, e_{1:t}) = \sum_{x_{t+k-1}} P(X_{t+k}|x_{t+k-1}, e_{1:t})P(x_{t+k-1}, e_{1:t}) \tag{3.3}$$

using the chain rule and the transition model

$$P(X_{t+k}, e_{1:t}) = \sum_{x_{t+k-1}} P(X_{t+k}|x_{t+k-1})P(x_{t+k-1}|e_{1:t})P(e_{1:t}) \tag{3.4}$$

taking $P(e_{1:t})$ outside the summation

$$P(X_{t+k}, e_{1:t}) = P(e_{1:t}) \sum_{x_{t+k-1}} P(X_{t+k}|x_{t+k-1})P(x_{t+k-1}|e_{1:t}) \tag{3.5}$$

replacing 3.5 into 3.1

$$P(X_{t+k}|e_{1:t}) = \frac{P(e_{1:t})\sum_{x_{t+k-1}} P(X_{t+k}|x_{t+k-1})P(x_{t+k-1}|e_{1:t})}{P(e_{1:t})} \tag{3.6}$$

canceling $P(e_{1:t})$

$$P(X_{t+k}|e_{1:t}) = \sum_{x_{t+k-1}} P(X_{t+k}|x_{t+k-1})P(x_{t+k-1}|e_{1:t}) \tag{3.7}$$

4 Smoothing

Smoothing is defined as $P(X_k|e_{1:t})$ where $k < t$ and is derived as follows:

Using conditional probability

$$P(X_k|e_{1:t}) = \frac{P(X_k, e_{1:t})}{P(e_{1:t})} \tag{4.1}$$

separating $e_{1:t}$ into $e_{1:k}, e_{k+1:t}$

$$P(X_k, e_{1:t}) = P(X_k, e_{1:k}, e_{k+1:t}) \tag{4.2}$$

using the chain rule

$$P(X_k, e_{1:t}) = P(e_{k+1:t}|X_k, e_{1:k})P(X_k, e_{1:k}) \tag{4.3}$$

using the chain rule again, and the transition model

$$P(X_k, e_{1:t}) = P(e_{k+1:t}|X_k)P(X_k|e_{1:k})P(e_{1:k}) \tag{4.4}$$

replacing 4.4 into 4.1

$$P(X_k|e_{1:t}) = \frac{P(e_{k+1:t}|X_k)P(X_k|e_{1:k})P(e_{1:k})}{P(e_{1:t})} \tag{4.5}$$

Using conditional probability

$$P(e_{k+1:t}|X_k) = \frac{P(e_{k+1:t}, X_k)}{P(X_k)} \tag{4.6}$$

expanding $e_{k+1:t}$ into $e_{k+1}, e_{k+2:t}$

$$P(e_{k+1:t}|X_k) = \frac{P(e_{k+1}, e_{k+2:t}, X_k)}{P(X_k)} \tag{4.7}$$

Using marginal probability

$$P(e_{k+1}, e_{k+2:t}, X_k) = \sum_{x_{k+1}} P(e_{k+1}, e_{k+2:t}, X_k, x_{k+1}) \tag{4.8}$$

using the chain rule

$$P(e_{k+1}, e_{k+2:t}, X_k) = \sum_{x_{k+1}} P(e_{k+1}|e_{k+2:t}, X_k, x_{k+1})P(e_{k+2:t}, X_k, x_{k+1}) \tag{4.9}$$

using the chain rule again

$$P(e_{k+1}, e_{k+2:t}, X_k) = \sum_{x_{k+1}} P(e_{k+1}|e_{k+2:t}, X_k, x_{k+1})P(e_{k+2:t}|X_k, x_{k+1})P(X_k, x_{k+1}) \tag{4.10}$$

using the chain rule again

$$P(e_{k+1}, e_{k+2:t}, X_k) = \sum_{x_{k+1}} P(e_{k+1}|e_{k+2:t}, X_k, x_{k+1})P(e_{k+2:t}|X_k, x_{k+1})P(x_{k+1}|X_k)P(X_k) \tag{4.11}$$

using the sensor model

$$P(e_{k+1}, e_{k+2:t}, X_k) = \sum_{x_{k+1}} P(e_{k+1}|x_{k+1})P(e_{k+2:t}|X_k, x_{k+1})P(x_{k+1}|X_k)P(X_k) \tag{4.12}$$

using the extended sensor model

$$P(e_{k+1}, e_{k+2:t}, X_k) = P(X_k) \sum_{x_{k+1}} P(e_{k+1}|x_{k+1})P(e_{k+2:t}|x_{k+1})P(x_{k+1}|X_k) \tag{4.13}$$

replacing 4.13 into 4.7

$$P(e_{k+1:t}|X_k) = \frac{P(X_k)\sum_{x_{k+1}} P(e_{k+1}|x_{k+1})P(e_{k+2:t}|x_{k+1})P(x_{k+1}|X_k)}{P(X_k)} \tag{4.14}$$

canceling $P(X_k)$

$$P(e_{k+1:t}|X_k) = \sum_{x_{k+1}} P(e_{k+1}|x_{k+1})P(e_{k+2:t}|x_{k+1})P(x_{k+1}|X_k) \tag{4.15}$$

replacing 4.15 into 4.5

$$P(X_k|e_{1:t}) = \frac{P(e_{1:k})P(X_k|e_{1:k})\sum_{x_{k+1}} P(e_{k+1}|x_{k+1})P(e_{k+2:t}|x_{k+1})P(x_{k+1}|X_k)}{P(e_{1:t})} \tag{4.16}$$

then we define $\alpha = \frac{P(e_{1:k})}{P(e_{1:t})}$

$$P(X_k|e_{1:t}) = \alpha P(X_k|e_{1:k}) \sum_{x_{k+1}} P(e_{k+1}|x_{k+1})P(e_{k+2:t}|x_{k+1})P(x_{k+1}|X_k) \tag{4.17}$$