(Very) Brief Review to Probabilistic Theory

Joint and Conditional Probability

- $P(X=x \text{ and } Y=y) = P(x,y)$
- If X and Y are independent then
  $P(x,y) = P(x)\,P(y)$
- $P(x \mid y)$ is the probability of x given y
  $P(x,y) = P(x \mid y)\,P(y)$
  $P(x \mid y) = P(x,y) / P(y)$
- If X and Y are independent then
  $P(x \mid y) = P(x)$

2

Law of Total Probability, Marginals

Discrete case

$$\sum_x P(x) = 1$$

$$P(x) = \sum_y P(x,y)$$

$$P(x) = \sum_y P(x \mid y) P(y)$$

Continuous case

$$\int p(x)\,dx = 1$$

$$p(x) = \int p(x,y)\,dy$$

$$p(x) = \int p(x \mid y) p(y)\,dy$$

3

Bayes Rule

$$P(x,y) = P(x \mid y)P(y) = P(y \mid x)P(x)$$

$$\Rightarrow$$

$$P(x \mid y) = \frac{P(y \mid x)\ P(x)}{P(y)} = \frac{\text{likelihood} \cdot \text{prior}}{\text{evidence}}$$

Não depende de x
Difícil de calcular!

Solução: lei da probabilidade total!

$$P(y) = \sum_x P(y \mid x) P(x)$$

4

Bayes Rule

$$P(x \mid y) = \frac{P(y \mid x)\ P(x)}{P(y)} = \frac{\text{likelihood} \cdot \text{prior}}{\text{evidence}}$$

$$P(x \mid y) = \frac{P(y \mid x)\ P(x)}{\sum_x P(y \mid x) P(x)} = \eta\ P(y \mid x)\, P(x)$$

5

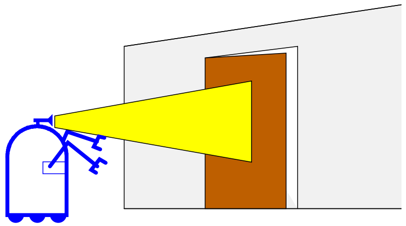
Bayes Rule with Background Knowledge

$$P(x \mid y, z) = \frac{P(y \mid x, z)\ P(x \mid z)}{P(y \mid z)}$$

6

Simple Example of State Estimation

- Suppose a robot obtains measurement z
- What is $P(open|z)$?



7

Causal vs. Diagnostic Reasoning

- $P(open|z)$ is diagnostic.
- $P(z|open)$ is causal.
- Often causal knowledge is easier to obtain.
- Bayes rule allows us to use causal knowledge:

count frequencies!

$$P(open \mid z) = \frac{P(z \mid open) P(open)}{P(z)}$$

8

Example

- $P(z|open) = 0.6$ $\qquad P(z|\neg open) = 0.3$
- $P(open) = P(\neg open) = 0.5$

$$P(open \mid z) = ??$$

$$P(x \mid y) = \frac{P(y \mid x)\ P(x)}{P(y)}$$

$$P(x \mid y) = \frac{P(y \mid x)\ P(x)}{\sum_x P(y \mid x) P(x)} = \eta\ P(y \mid x) P(x)$$

9

Example

- $P(z|open) = 0.6$ $\qquad P(z|\neg open) = 0.3$
- $P(open) = P(\neg open) = 0.5$

$$P(open \mid z) = \frac{P(z \mid open) P(open)}{P(z \mid open)\, p(open) + P(z \mid \neg open)\, p(\neg open)}$$

$$P(open \mid z) = \frac{0.6 \cdot 0.5}{0.6 \cdot 0.5 + 0.3 \cdot 0.5} = \frac{2}{3} = 0.67$$

- z raises the probability that the door is open.

10

Combining Evidence

- Suppose our robot obtains another observation z_2.
- How can we integrate this new information?
- More generally, how can we estimate $P(x|\ z_1 ... z_n)$?

11

Recursive Bayesian Updating

$$P(x \mid z_1, \ldots, z_n) = \frac{P(z_n \mid x, z_1, \ldots, z_{n-1})\ P(x \mid z_1, \ldots, z_{n-1})}{P(z_n \mid z_1, \ldots, z_{n-1})}$$

Markov assumption: z_n is independent of $z_1, \ldots, z_{n-1}$ if we know x.

$$P(x \mid z_1, \ldots, z_n) = \frac{P(z_n \mid x)\ P(x \mid z_1, \ldots, z_{n-1})}{P(z_n \mid z_1, \ldots, z_{n-1})}$$

$$= \eta\ P(z_n \mid x)\ P(x \mid z_1, \ldots, z_{n-1})$$

$$= \eta_{1 \ldots n} \prod_{i=1 \ldots n} P(z_i \mid x)\ P(x)$$

12

Example: Second Measurement

- $P(z_2|open) = 0.5$ $\quad$ $P(z_2|\neg open) = 0.6$
- $P(open|z_1) = 2/3$

$$P(open \mid z_2, z_1) = \frac{P(z_2 \mid open)\, P(open \mid z_1)}{P(z_2 \mid open)\, P(open \mid z_1) + P(z_2 \mid \neg open)\, P(\neg open \mid z_1)}$$

$$= \frac{\frac{1}{2}\cdot\frac{2}{3}}{\frac{1}{2}\cdot\frac{2}{3}+\frac{3}{5}\cdot\frac{1}{3}} = \frac{5}{8} = 0.625$$

- z_2 lowers the probability that the door is open.

Actions

- Often the world is **dynamic** since
  - **actions carried out by the robot**,
  - **actions carried out by other agents**,
  - or just the **time** passing by changes the world.

- How can we **incorporate** such **actions**?

Typical Actions

- The robot **turns its wheels** to move
- The robot **uses its manipulator** to grasp an object
- Plants grow over **time**…

- Actions are **never carried out with absolute certainty**.
- In contrast to measurements, **actions generally increase the uncertainty**.

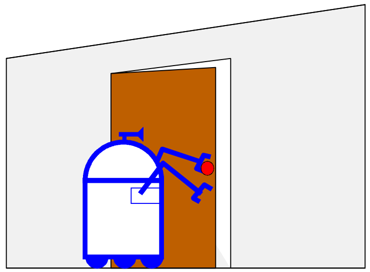
Modeling Actions

- To incorporate the outcome of an action *u* into the current "belief", we use the conditional pdf

P(x|u,x')

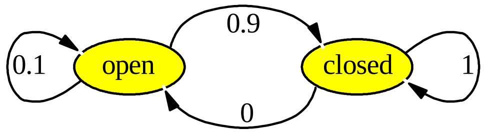
- This term specifies the pdf that **executing *u* changes the state from *x'* to *x*.**

Example: Closing the door



State Transitions

P(x|u,x') for *u* = "close door":



If the door is open, the action "close door" succeeds in 90% of all cases.

Integrating the Outcome of Actions

Continuous case:

$$P(x|u)=\int P(x|u,x')P(x')dx'$$

Discrete case:

$$P(x|u)=\sum_{x'} P(x|u,x')P(x')$$

Example: The Resulting Belief

$$P(closed|u)=\sum_{x'} P(closed|u,x')P(x')$$

$$P(open|u)=\sum_{x'} P(open|u,x')P(x')$$

Example: The Resulting Belief

$$\begin{aligned} P(closed|u) &= \sum_{x'} P(closed|u,x')P(x') \\ &= P(closed|u,open)P(open) \\ &\quad + P(closed|u,closed)P(closed) \\ &= \frac{9}{10}*\frac{5}{8}+\frac{1}{1}*\frac{3}{8}=\frac{15}{16} \\ P(open|u) &= \sum_{x'} P(open|u,x')P(x') \\ &= P(open|u,open)P(open) \\ &\quad + P(open|u,closed)P(closed) \\ &= \frac{1}{10}*\frac{5}{8}+\frac{0}{1}*\frac{3}{8}=\frac{1}{16} \\ &= 1-P(closed|u) \end{aligned}$$

Bayes Filters: Framework

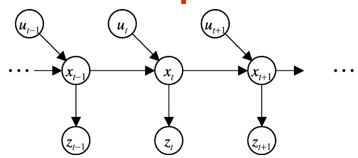
- **Given:**
  - Stream of observations *z* and action data *u:*

    $$d_t=\{u_1,z_1 \ldots,u_t,z_t\}$$

  - Sensor model *P(z|x).*
  - Action model *P(x|u,x').*
  - Prior probability of the system state *P(x').*
- **Wanted:**
  - Estimate of the state *X* of a dynamical system.
  - The posterior of the state is also called **Belief**:

    $$Bel(x_t)=P(x_t|u_1,z_1 \ldots,u_t,z_t)$$

Markov Assumption



$$p(z_t|x_{0:t},z_{1:t},u_{1:t}) = p(z_t|x_t)$$

$$p(x_t|x_{1:t-1},z_{1:t},u_{1:t}) = p(x_t|x_{t-1},u_t)$$

Underlying Assumptions

- Static world
- Independent noise
- Perfect model, no approximation errors

Bayes Filters

z = observation
u = action
x = state

$$Bel(x_t)=P(x_t|u_1,z_1 \ldots,u_t,z_t)$$

Bayes $$=\eta\ P(z_t|x_t,u_1,z_1,\ldots,u_t)\ P(x_t|u_1,z_1,\ldots,u_t)$$

Markov $$=\eta\ P(z_t|x_t)\ P(x_t|u_1,z_1,\ldots,u_t)$$

Total prob. $$=\eta\ P(z_t|x_t)\int P(x_t|u_1,z_1,\ldots,u_t,x_{t-1})\ P(x_{t-1}|u_1,z_1,\ldots,u_t)\ dx_{t-1}$$

Markov $$=\eta\ P(z_t|x_t)\int P(x_t|u_t,x_{t-1})\ P(x_{t-1}|u_1,z_1,\ldots,u_t)\ dx_{t-1}$$

Markov $$=\eta P(z_t|x_t)\int P(x_t|u_t,x_{t-1})\ P(x_{t-1}|u_1,z_1,\ldots,z_{t-1})\ dx_{t-1}$$

$$=\eta\ P(z_t|x_t)\int P(x_t|u_t,x_{t-1})\ Bel(x_{t-1})\ dx_{t-1}$$

$$Bel(x_t) = \eta \ P(z_t \mid x_t) \int P(x_t \mid u_t, x_{t-1}) \ Bel(x_{t-1}) \ dx_{t-1}$$

Algorithm **Bayes_filter**(*Bel(x),d*):
η=0
If *d* is a perceptual data item *z* then
    For all *x* do
        $Bel(x) = P(z \mid x) Bel'(x)$
        $\eta = \eta + Bel(x)$
    For all *x* do
        $Bel(x) = \eta^{-1} Bel(x)$
Else if *d* is an action data item *u* then
    For all *x* do
        $Bel(x) = \int P(x \mid u, x') \ Bel(x') \ dx'$
Return *Bel(x)*

Bayes Filters are Familiar!

$$Bel(x_t) = \eta \ P(z_t \mid x_t) \int P(x_t \mid u_t, x_{t-1}) \ Bel(x_{t-1}) \ dx_{t-1}$$

- Kalman filters
- Particle filters
- Hidden Markov models
- Dynamic Bayesian networks
- Partially Observable Markov Decision Processes (POMDPs)

Summary

- Bayes rule allows us to compute probabilities that are hard to assess otherwise.
- Under the Markov assumption, recursive Bayesian updating can be used to efficiently combine evidence.
- Bayes filters are a probabilistic tool for estimating the state of dynamic systems.