

Basics of Probability Theory in Robotics

Slides Courtesy of Prof. Thrun et al.

Dr. Ramviyas Nattanmai Parasuraman,
Asst. Professor, Computer Science, UGA

08/23/2018



Department of Computer Science

Franklin College of Arts and Sciences

UNIVERSITY OF GEORGIA

Probabilistic Robotics

Key idea:

Explicit representation of uncertainty using the calculus of probability theory

- Perception = state estimation
- Action = utility optimization

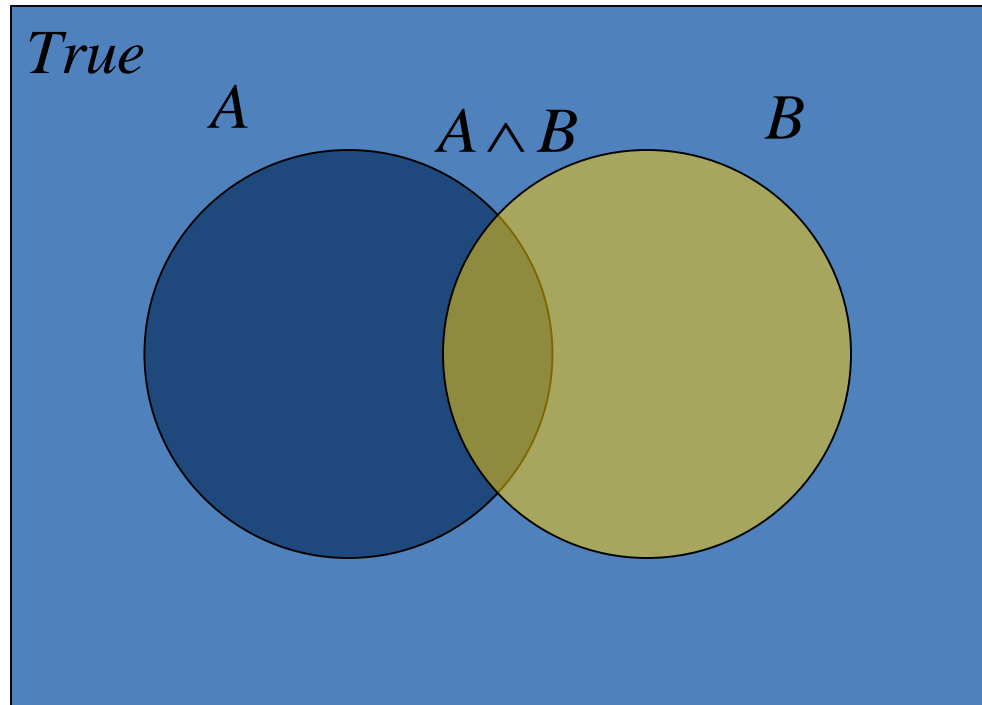
Axioms of Probability Theory

$\Pr(A)$ denotes probability that proposition A is true.

- $0 \leq \Pr(A) \leq 1$
- $\Pr(\textit{True}) = 1$ $\Pr(\textit{False}) = 0$
- $\Pr(A \vee B) = \Pr(A) + \Pr(B) - \Pr(A \wedge B)$

A Closer Look at Axiom 3

$$\Pr(A \vee B) = \Pr(A) + \Pr(B) - \Pr(A \wedge B)$$



Using the Axioms

$$\Pr(A \vee \neg A) = \Pr(A) + \Pr(\neg A) - \Pr(A \wedge \neg A)$$

$$\Pr(\textit{True}) = \Pr(A) + \Pr(\neg A) - \Pr(\textit{False})$$

$$1 = \Pr(A) + \Pr(\neg A) - 0$$

$$\Pr(\neg A) = 1 - \Pr(A)$$

Discrete Random Variables

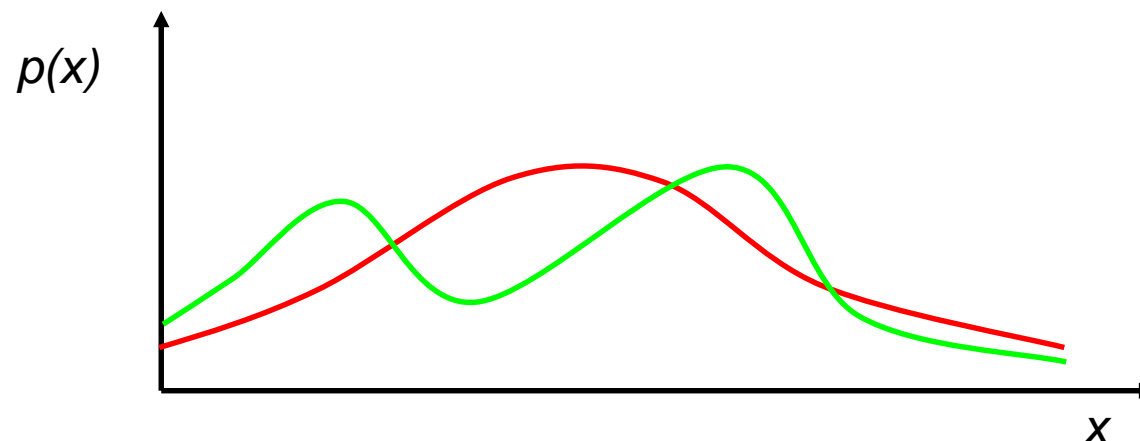
- X denotes a **random variable**.
- X can take on a countable number of values in $\{x_1, x_2, \dots, x_n\}$.
- $P(X=x_i)$, or $P(x_i)$, is the **probability** that the random variable X takes on value x_i .
- $P(\cdot)$ is called **probability mass function**.
- E.g. $P(Room) = \langle 0.7, 0.2, 0.08, 0.02 \rangle$

Continuous Random Variables

- X takes on values in the continuum.
- $p(X=x)$, or $p(x)$, is a probability density function.

$$\Pr(x \in (a, b)) = \int_a^b p(x) dx$$

- E.g.



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
$$P(x \mid y) = P(x,y) / P(y)$$
$$P(x,y) = P(x \mid y) P(y)$$
- If X and Y are independent then
$$P(x \mid y) = P(x)$$

Law of Total Probability, Marginals

Discrete case

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

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

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

Continuous case

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

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

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

Bayes Rule

$$P(x, y) = P(x | y)P(y) = P(y | x)P(x)$$

\Rightarrow

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

Normalization

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

$$\eta = P(y)^{-1} = \frac{1}{\sum_x P(y | x) P(x)}$$

Algorithm:

$$\forall x : \text{aux}_{x|y} = P(y | x) P(x)$$

$$\eta = \frac{1}{\sum_x \text{aux}_{x|y}}$$

$$\forall x : P(x | y) = \eta \text{aux}_{x|y}$$

Conditioning

- Law of total probability:

$$P(x) = \int P(x, z) dz$$

$$P(x) = \int P(x | z) P(z) dz$$

$$P(x | y) = \int P(x | y, z) P(z | y) dz$$

Bayes Rule with Background Knowledge

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

Conditioning

- Total probability:

$$P(x) = \int P(x, z) dz$$

$$P(x) = \int P(x | z) P(z) dz$$

$$P(x | y) = \int P(x | y, z) P(z) dz$$

Conditional Independence

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

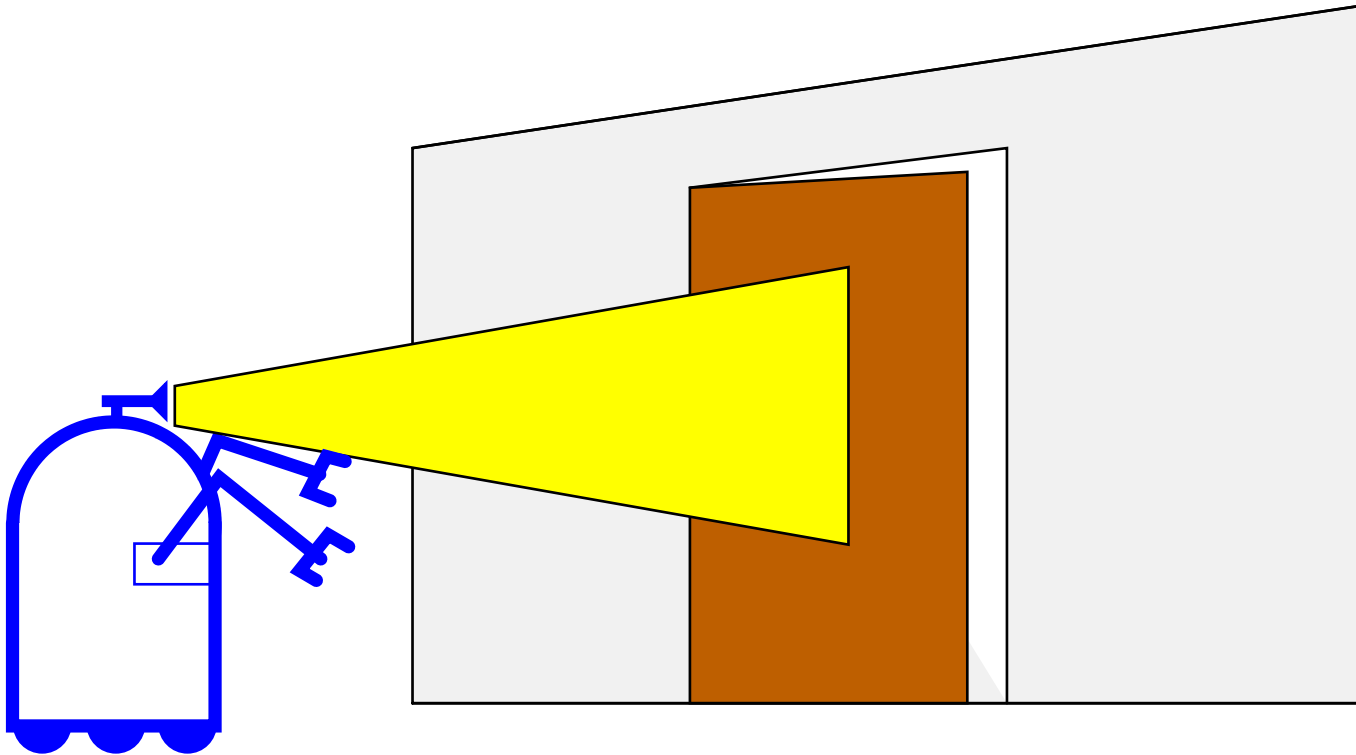
equivalent to

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

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

Simple Example of State Estimation

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



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!

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

Example

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

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

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

- z raises the probability that the door is open.

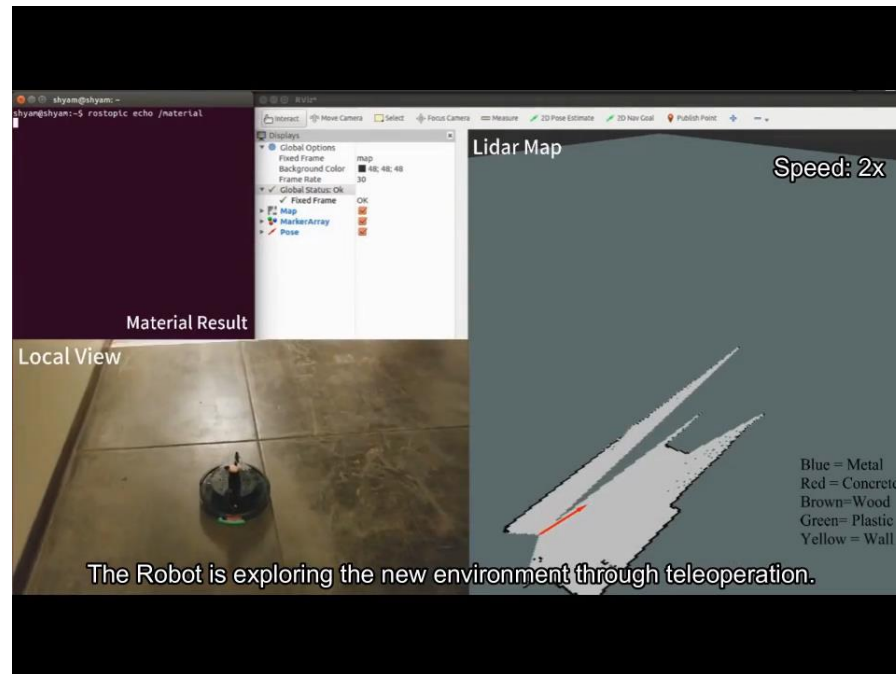
Combining Evidence

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

Robot Learning – Active Perception

Objective: Perception of materials in the environment

Approach: Active tapping of objects and machine learning of sounds to classify materials.



Wonse Jo, Shyam Sundar Kannan, **Ramviyas Parasuraman**, and Byung-Cheol Min, "Use of Tapping Sounds to Recognize Contact Material Types for Mobile Robots", *IROS 2018*, Madrid, Spain, October 1 -5, 2018. (Under review).

Recursive Bayesian Updating

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

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

$$\begin{aligned} P(x \mid z_1, \dots, z_n) &= \frac{P(z_n \mid x) P(x \mid z_1, \dots, z_{n-1})}{P(z_n \mid z_1, \dots, z_{n-1})} \\ &= \eta P(z_n \mid x) P(x \mid z_1, \dots, z_{n-1}) \\ &= \eta_{1..n} \prod_{i=1..n} P(z_i \mid x) P(x) \end{aligned}$$

Example: Second Measurement

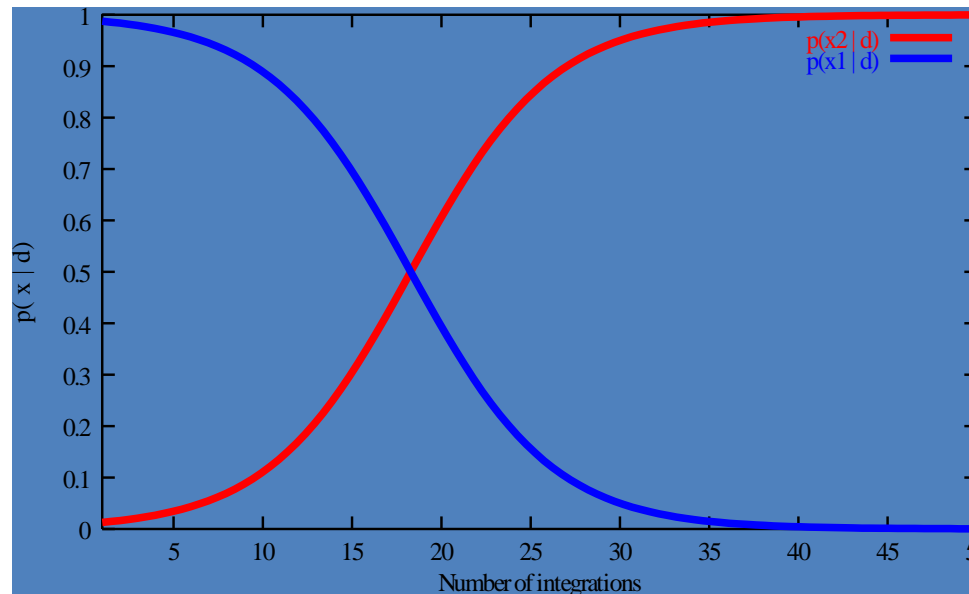
- $P(z_2/open) = 0.5$ $P(z_2/\neg open) = 0.6$
- $P(open/z_1) = 2/3$

$$\begin{aligned} P(open | z_2, z_1) &= \frac{P(z_2 | open) P(open | z_1)}{P(z_2 | open) P(open | z_1) + P(z_2 | \neg open) P(\neg open | z_1)} \\ &= \frac{\frac{1}{2} \cdot \frac{2}{3}}{\frac{1}{2} \cdot \frac{2}{3} + \frac{3}{5} \cdot \frac{1}{3}} = \frac{5}{8} = 0.625 \end{aligned}$$

- z_2 lowers the probability that the door is open.

A Typical Pitfall

- Two possible locations x_1 and x_2
- $P(x_1)=0.99$
- $P(z|x_2)=0.09$ $P(z|x_1)=0.07$



Actions

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

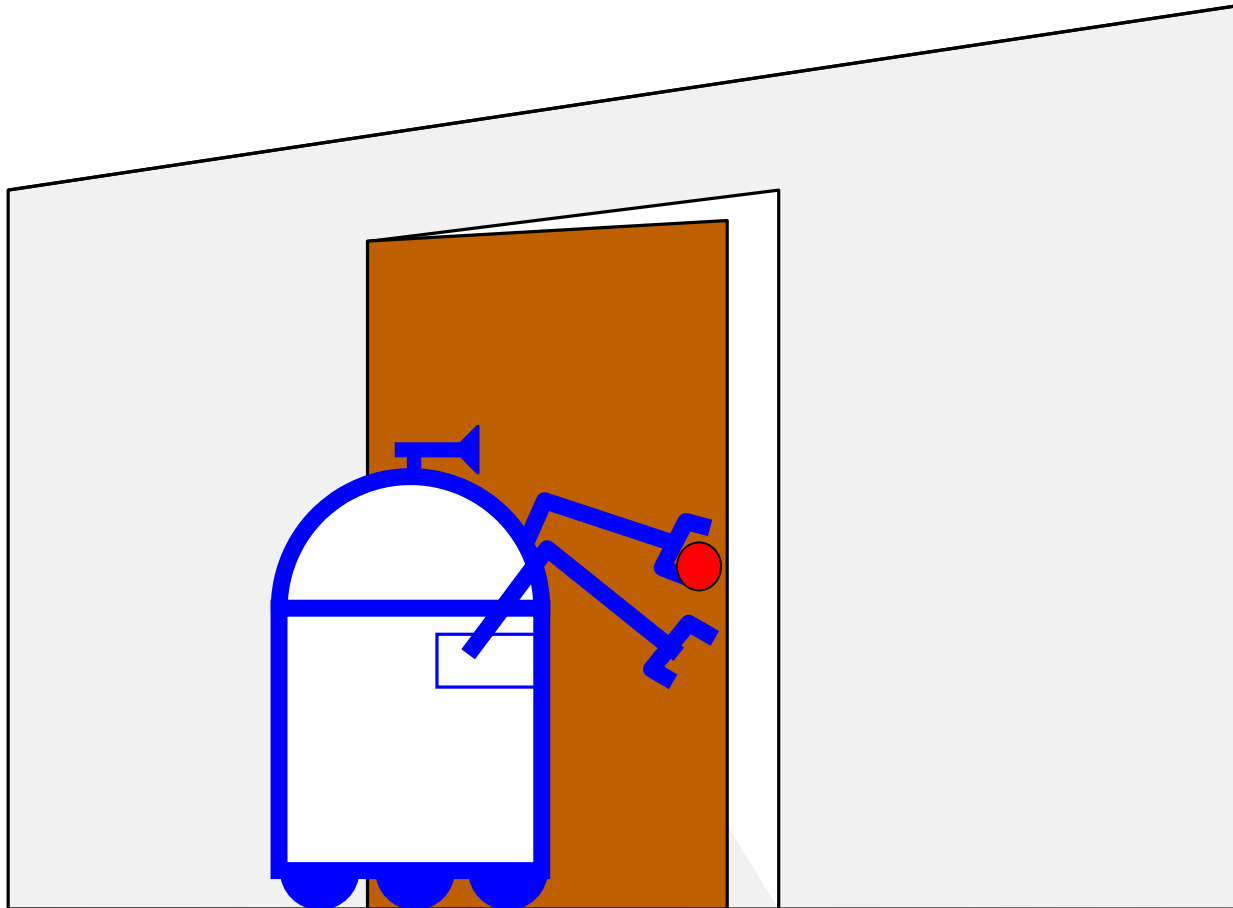
Modeling Actions

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

$$P(x|u, x')$$

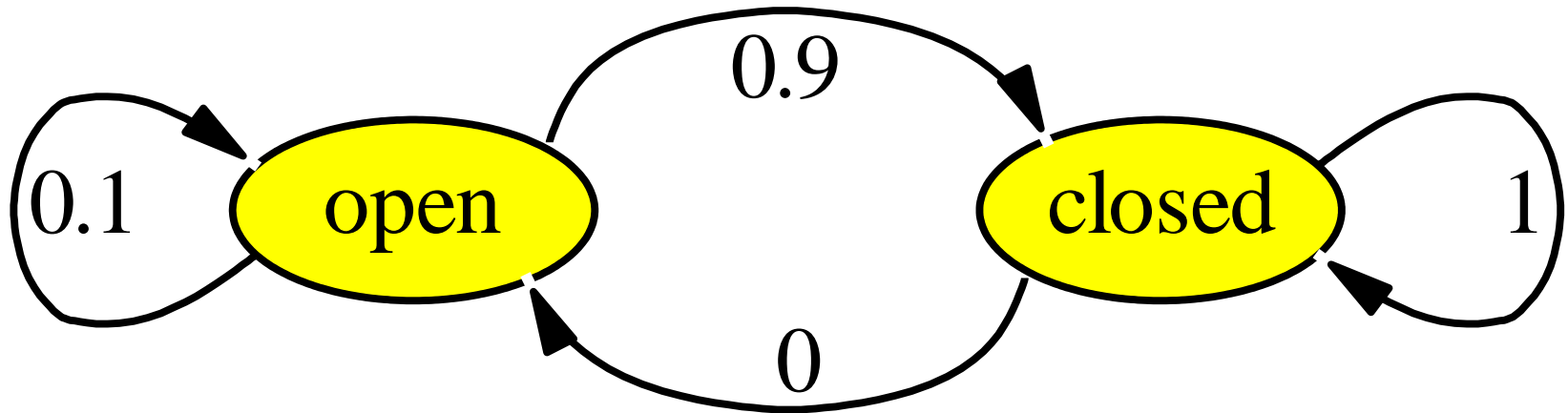
- This term specifies the pdf that **executing u changes the state from x' to x .**

Example: Closing the door



State Transitions

$P(x/u, x')$ for $u = \text{"close door"}:$



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:

$$P(x | u) = \sum P(x | u, x') P(x')$$

Example: The Resulting Belief

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

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

Bayes Filters: Framework

- **Given:**

- Stream of observations z and action data u :

- **Sensor model** $P(z_t | x_t, d_t^x) = \{u_1, z_1, \dots, u_t, z_t\}$

- **Action model** $P(x_t | u_t, x_{t-1})$.

- **Prior** probability of the system state $P(x)$.

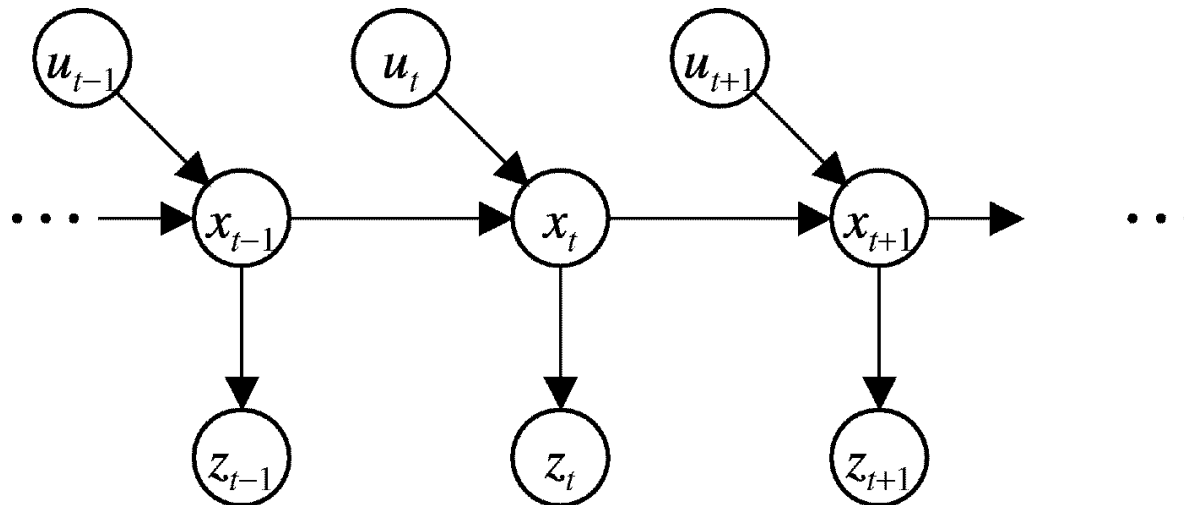
- **Wanted:**

- Estimate of the state X of a **dynamical system**.

- The posterior of the state is also called **Belief**:

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

Markov Assumption



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

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

Underlying Assumptions

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

Bayes Filters

$$\boxed{Bel(x_t)} = P(x_t | u_1, z_1, \dots, u_t, z_t)$$

z = observation
u = action
x = state

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

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

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

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

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

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

Bayes Filter Algorithm

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

1. Algorithm **Bayes_filter**($Bel(x), d$):
2. $\eta = 0$
3. If d is a **perceptual** data item z then
4. For all x do
5. $Bel'(x) = P(z | x) Bel(x)$
6. $\eta = \eta + Bel'(x)$
7. For all x do
8. $Bel'(x) = \eta^{-1} Bel'(x)$
9. Else if d is an **action** data item u then
10. For all x do
11. $Bel'(x) = \int P(x | u, x') Bel(x') dx'$
12. Return $Bel'(x)$

Bayes Filters are Familiar!

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

- 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.