Probability: Review

Pieter Abbeel

UC Berkeley EECS

Many slides adapted from Thrun, Burgard and Fox, Probabilistic Robotics

Why probability in robotics?

- Often state of robot and state of its environment are unknown and only noisy sensors available
  - Probability provides a framework to fuse sensory information
  - → Result: probability distribution over possible states of robot and environment
- Dynamics is often stochastic, hence can't optimize for a particular outcome, but only optimize to obtain a good distribution over outcomes
  - Probability provides a framework to reason in this setting
  - → Result: ability to find good control policies for stochastic dynamics and environments

Example 1: Helicopter

- State: position, orientation, velocity, angular rate
- Sensors:
  - GPS : noisy estimate of position (sometimes also velocity)
  - Inertial sensing unit: noisy measurements from
    - (i) 3-axis gyro [=angular rate sensor],
    - (ii) 3-axis accelerometer [=measures acceleration + gravity; e.g., measures (0,0,0) in free-fall],
    - (iii) 3-axis magnetometer
- Dynamics:
  - Noise from: wind, unmodeled dynamics in engine, servos, blades

Example 2: Mobile robot inside building

- State: position and heading
- Sensors:
  - Odometry (=sensing motion of actuators): e.g., wheel encoders
  - Laser range finder:
    - Measures time of flight of a laser beam between departure and return
    - Return is typically happening when hitting a surface that reflects the beam back to where it came from
- Dynamics:
  - Noise from: wheel slippage, unmodeled variation in floor

Axioms of Probability Theory

- $$0 \le \Pr(A) \le 1$$
- $$\Pr(\Omega) = 1 \qquad\qquad \Pr(\phi) = 0$$
- $$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$$

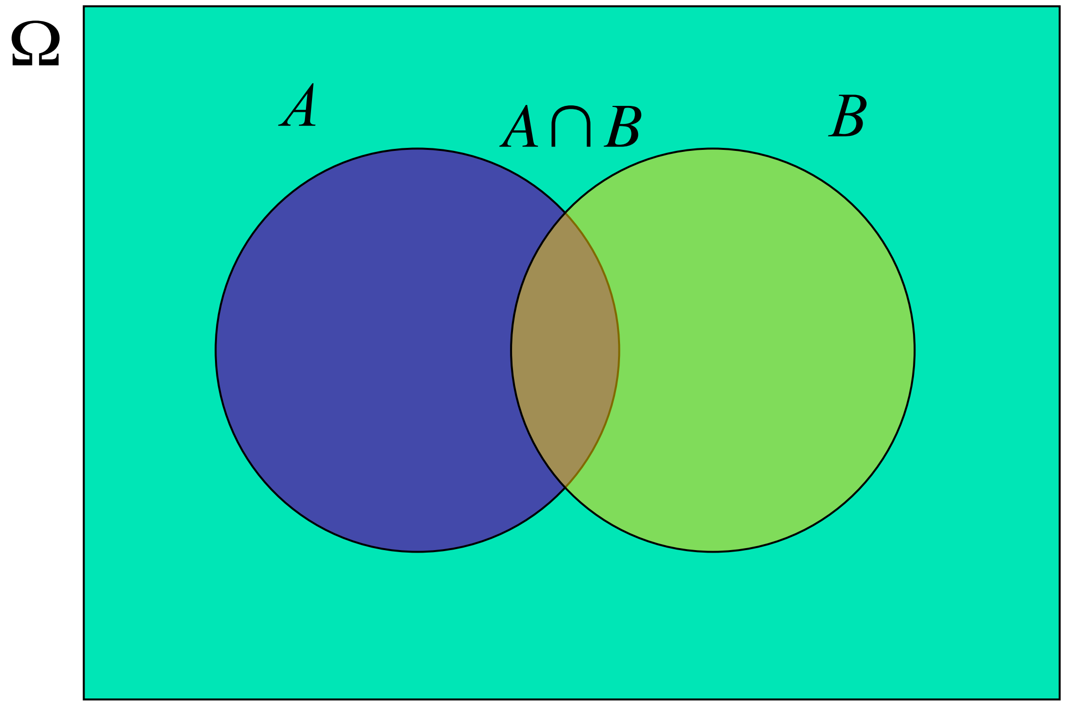
Pr*(A)* denotes probability that the outcome ω is an element of the set of possible outcomes *A*. *A* is often called an event. Same for *B.*

Ω is the set of all possible outcomes.

ϕ is the empty set.

A Closer Look at Axiom 3

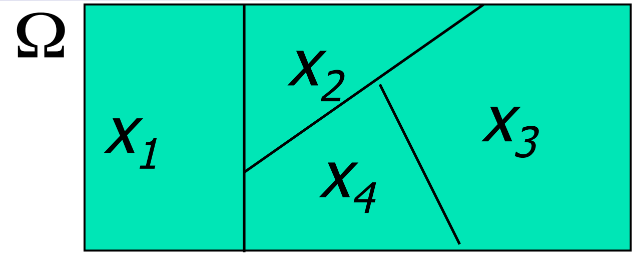
$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$$



Using the Axioms

$$
\begin{aligned}
\Pr(A \cup (\Omega \setminus A)) &= \Pr(A) + \Pr(\Omega \setminus A) - \Pr(A \cap (\Omega \setminus A)) \\
\Pr(\Omega) &= \Pr(A) + \Pr(\Omega \setminus A) - \Pr(\phi) \\
1 &= \Pr(A) + \Pr(\Omega \setminus A) - 0 \\
\Pr(\Omega \setminus A) &= 1 - \Pr(A)
\end{aligned}
$$

Discrete Random Variables



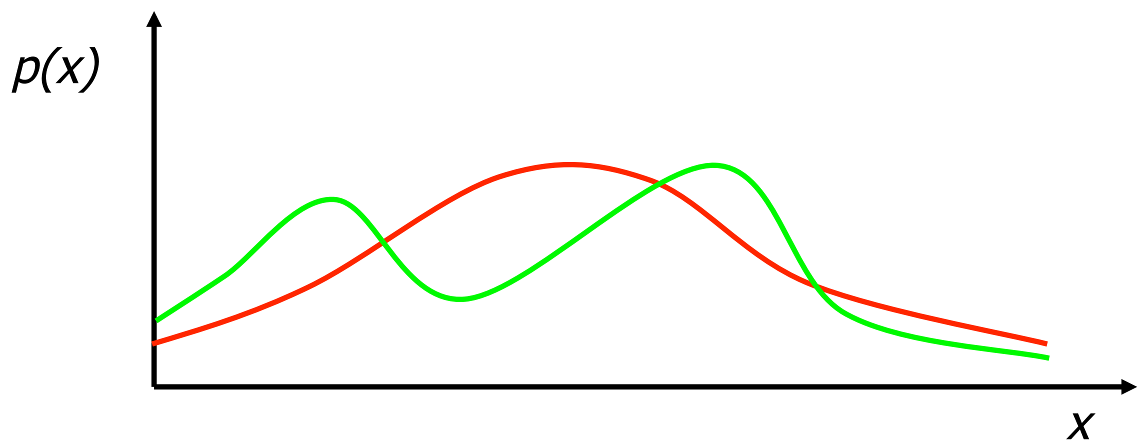
- X denotes a random variable.
- X can take on a countable number of values in $\{x_1, x_2, \ldots, x_n\}$.
- $P(X=x_i)$, or $P(x_i)$, is the probability that the random variable X takes on value x_i.
- $P(\cdot)$ is called probability mass function.
- *E.g., X models the outcome of a coin flip*, x_1 = head, x_2 = tail, $P(x_1) = 0.5$, $P(x_2) = 0.5$

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

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

- *Same for probability densities, just* $P \rightarrow p$

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

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

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

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

Bayes Formula

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

$$\Rightarrow$$

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

Normalization

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

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

Algorithm:

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

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

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

Conditioning

- Law of total probability:

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

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

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

Bayes Rule with Background Knowledge

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

Conditional Independence

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

equivalent to

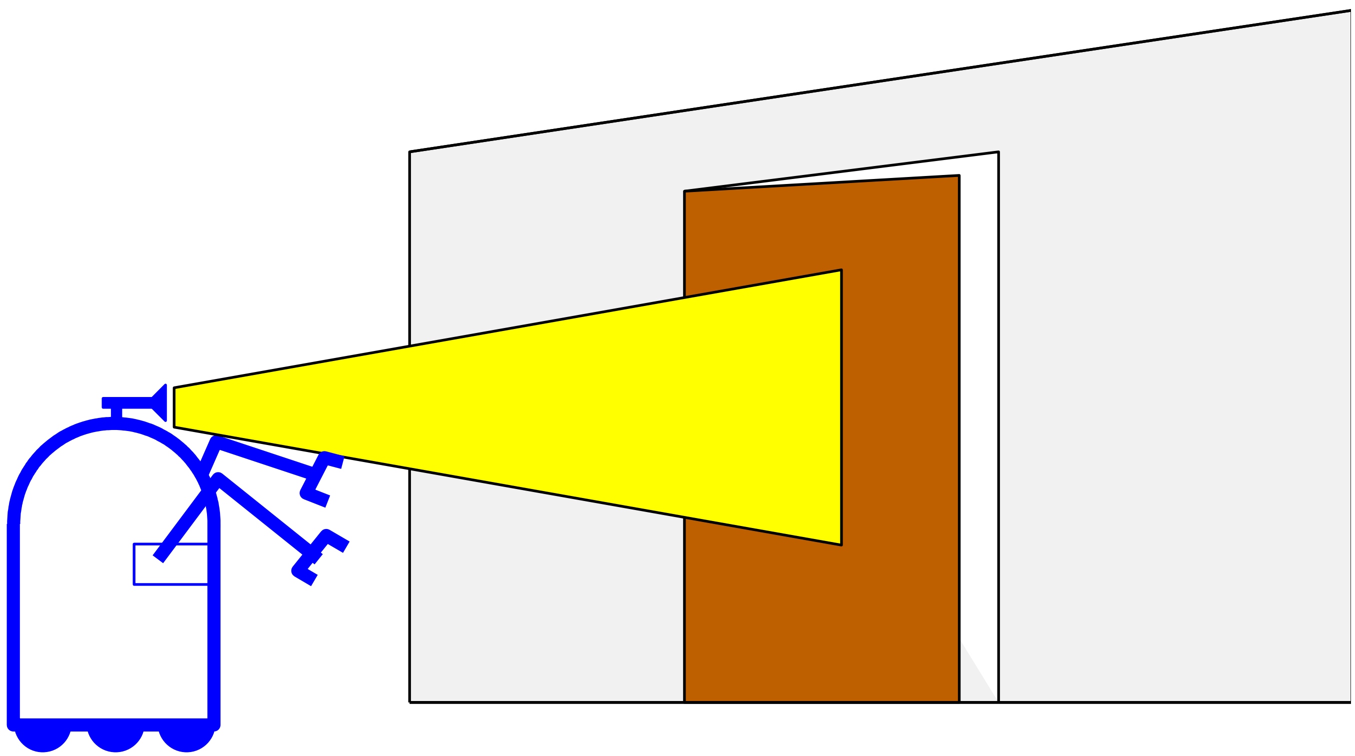
$$P(x|z)=P(x|z,y)$$

and

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

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 ca

count frequencies!

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

Example

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

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

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

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

- *z* raises the probability that the door is open.

Combining Evidence

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

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.*

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

Example: Second Measurement

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

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

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

• z_2 lowers the probability that the door is open.

A Typical Pitfall

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