Probabilistic Robotics*

Probability Basics

Probabilities, Bayes rule, Bayes filters

Prof. Mohan Sridharan
Chair in Robot Systems
University of Edinburgh, UK
https://homepages.inf.ed.ac.uk/msridhar/
m.sridharan@ed.ac.uk

*Revised original slides that accompany the book: PR by Thrun, Burgard and Fox.

Probabilistic Robotics

Key idea:

Explicit representation of uncertainty using the calculus of probability theory

- Perception = state estimation
- Action = utility optimization

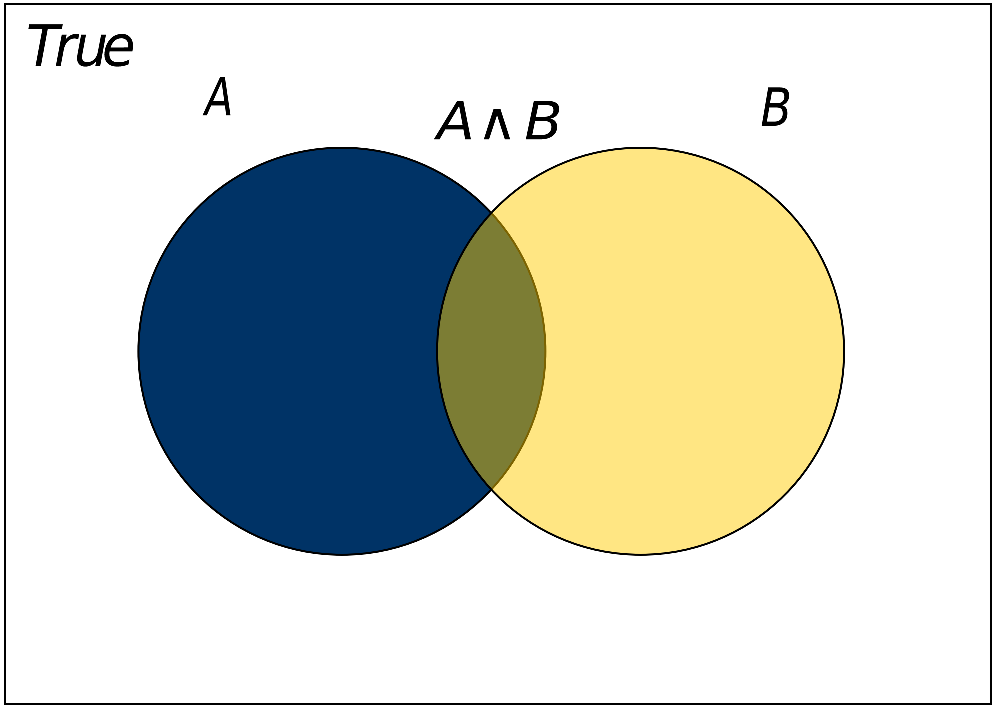
Axioms of Probability Theory

Pr*(A)* or P(*A*) denotes probability that proposition *A* is true.

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

A Closer Look at Axiom 3

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



Using the Axioms

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

$$\Pr(True) \qquad \Pr(A)+\Pr(\neg A)-\Pr(False)$$

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

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

Discrete Random Variables

- *X* denotes a random variable.
- *X* can take on a countable number of values in:

  $\{x_1, x_2, \ldots, x_n\}$.

- $P(X=x_i)$, or $P(x_i)$, is the probability that the random variable *X* takes on value x_i.
- $P(.)$ is called probability mass function.
- E.g. $P(Room) = \langle 0.7, 0.2, 0.08, 0.02 \rangle$

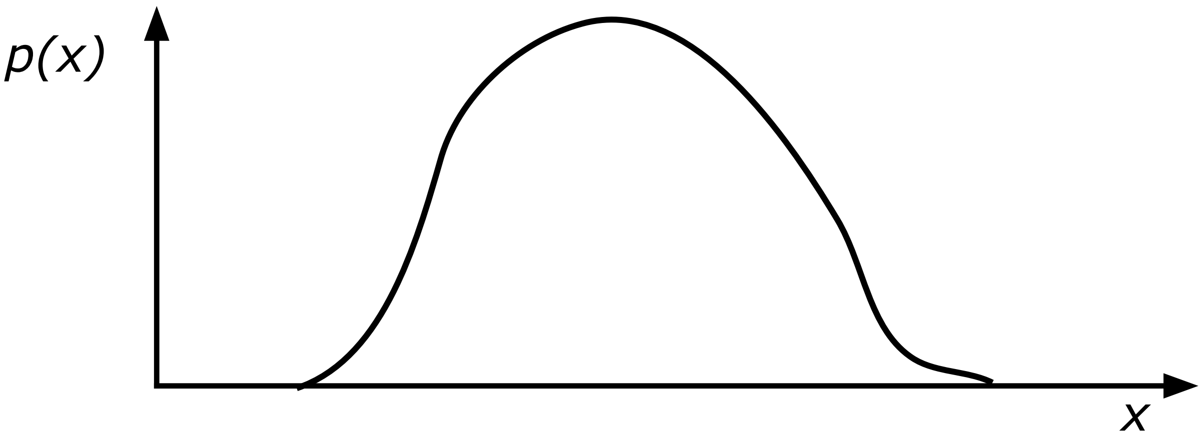
Continuous Random Variables

- *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

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

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

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

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

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

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

Normalization

- Denominator of Bayes rule is a “normalizer”.

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

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

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

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

- Bayes rule can take into account background knowledge:

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

- Essential condition on background knowledge.

Conditional Independence

- X and Y conditionally independent given Z:

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

equivalent to:

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

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

Conditional Independence Example

- Two coins; one fair, one biased (always shows heads).
- Pick coin at random and toss twice.

- Define three events:
  - X = Heads on first throw.
  - Y = Heads on second throw.
  - Z = first (fair) coin was selected.

- Compute the following:
  - P(X|Z), P(Y|Z), P(X, Y|Z), P(X), P(Y), P(X, Y).

Formal Definitions (Section 2.3, PR)

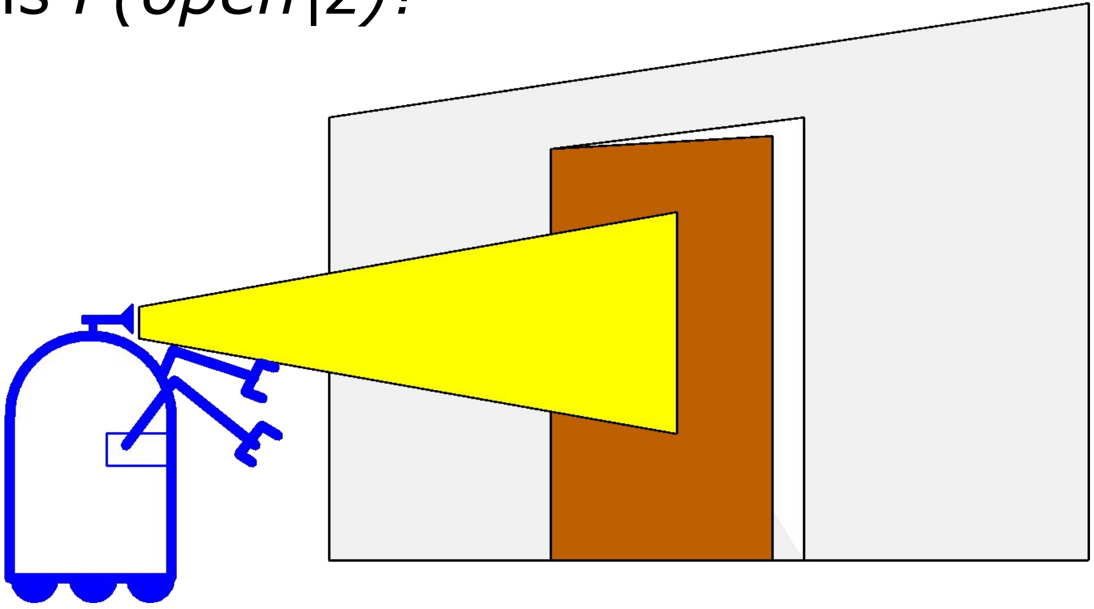
- State: all aspects of robot and environment that can impact the future (*x* or *s*).
- Static and dynamic state; complete state. Discrete and continuous state.
- Pose: position + orientation.

- Markov assumption: past and future data independent given current state.
- Environment interaction:
  - Sensor measurements (z or o). *Increase knowledge.*
  - Control actions (u or a). *Increase uncertainty.*

- Belief (or belief/information state) $bel(x_t) = p(x_t | z_{1:t}, u_{1:t})$

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.

count frequencies?

- Often causal knowledge is easier to obtain.
- Bayes rule allows us to use causal knowledge:

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

Example

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

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

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

Measurement z raises probability that the door is open.

Combining Evidence

- Suppose robot obtains another observation z_2.
- How can we integrate this new information?
- How can we estimate the result of a series of measurements/observations?

$$P(x| z_1 \ldots z_n) = ?$$

Recursive Bayesian Updating

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

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

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

Example: Second Measurement

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

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

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

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 world changes over the passage of time.

- How can we incorporate such actions?

Typical Actions

- The robot turns its wheels to move.
- Robot uses its manipulator to grasp an object
- Plants grow over time … ☺
- Actions are never carried out with certainty.
- In contrast to measurements, actions generally increase uncertainty.

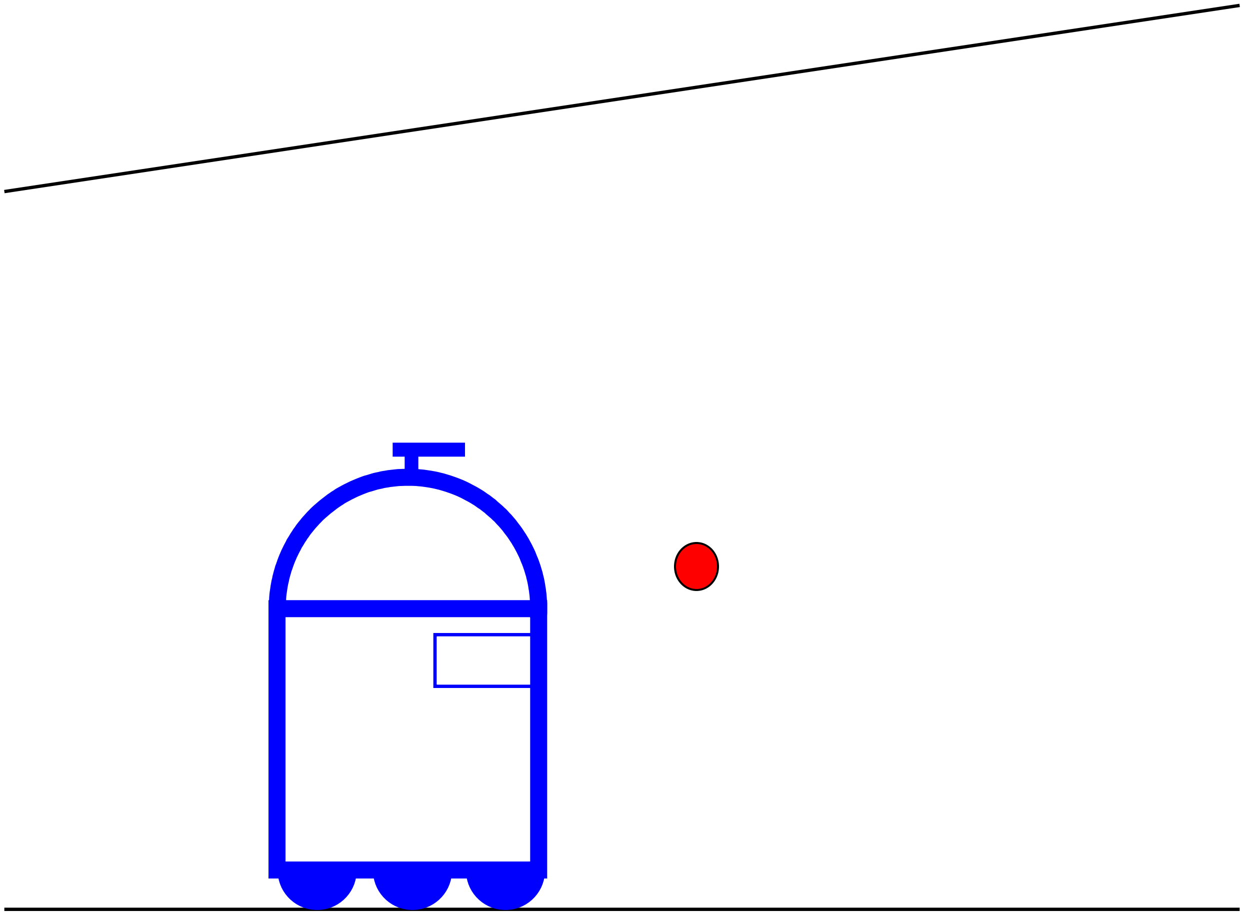
Modeling Actions

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

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

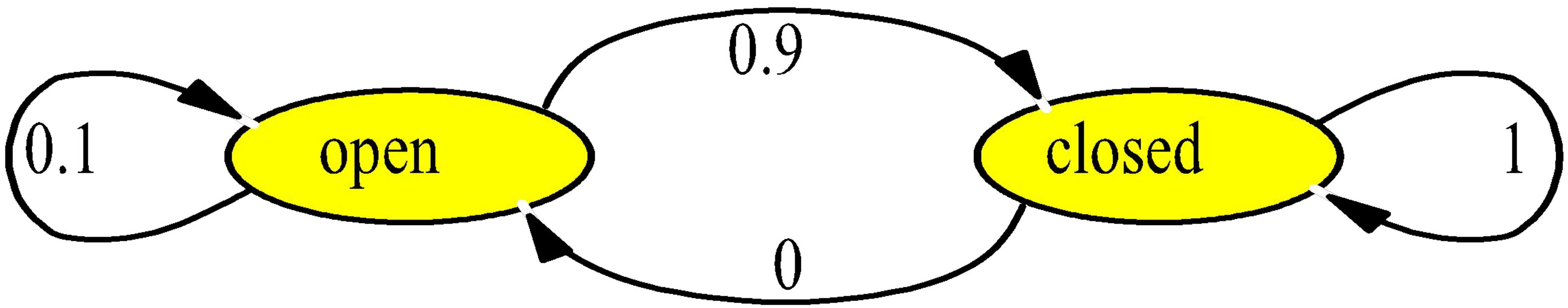
- This term specifies the pdf that executing *u* changes the state from *x' to x.*

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 the 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

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

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

Another Example: Four Rooms

- Four rooms arranged in a square; four actions (up, down, left, right). Simple transition probabilities:

$$P(x|u,x') = 0.8/0.2 \text{ for valid actions}$$
$$= 0 \text{ otherwise}$$

- How do we compute updated probabilities given *u=up* has been executed?

Bayes Filters: Framework

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

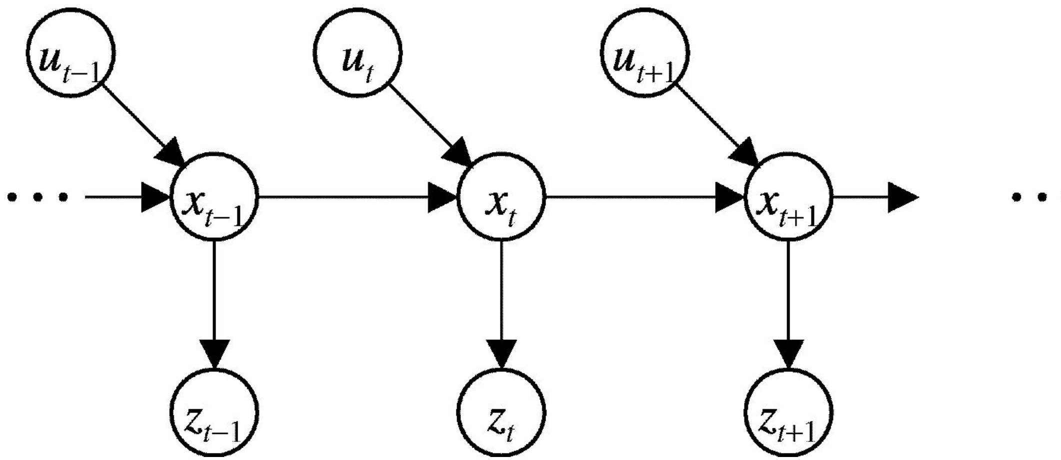
$$d_t = \{u_1, z_1 \ldots, u_t, z_t\}$$

  - 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



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

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

<u>Underlying Assumptions:</u>

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

Bayes Filters

z = observation
u = action
x = state

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

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

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

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

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

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

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

Bayes Filter Algorithm

1. Algorithm **Bayes_filter**($Bel(x_{t-1}), u_t, z_t$):
2. For all x_t do
3. $\overline{Bel}(x_t) = \int p(x_t \mid u_t, x_{t-1}) Bel(x_{t-1}) dx_{t-1}$
4. $Bel(x_t) = \eta \; p(z_t \mid x_t) \overline{Bel}(x_t)$
5. End for $Bel(x_t)$
6. Return

Two key steps: *prediction* and *correction*.

Also known as control update and measurement update.

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

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.

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.