Robotics
Spring 2023

Abhishek Gupta

TAs: Yi Li, Srivatsa GS

Recap: Course Overview

- Filtering/Smoothing
- Localization
- Mapping
- SLAM

- Search
- Motion Planning
- TrajOpt
- Stability/Certification

- MDPs and RL
- Imitation Learning
- Solving POMDPs

Lecture Outline

Probability Review

↓

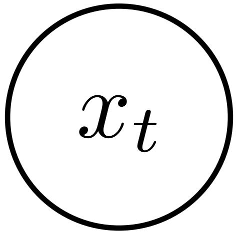
Bayesian Inference

↓

Bayesian Filtering

Why state estimation?

- "State" is an extremely hard thing to define and measure
  - Usually unobservable (only "measurements" are observable)
- State can be a choice
  - More detailed state, less uncertainty
  - Less detailed state, more uncertainty



Pose/velocity of the object

Position and momentum of all particles

Why **probabilistic** state estimation?

x_t Pose/velocity of the object

- When state is abstracted/incomplete, this manifests as noise/uncertainty
- Being probabilistic allows for:
  - Robustness to external noise
  - Exploration to get better/gather information
  - Dealing with inherently stochastic systems
  - Accounting for inaccurate hardware/software

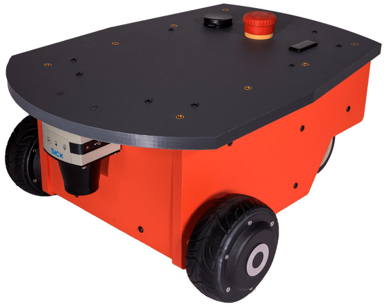
Probabilistic Robotics

Key idea: Explicit representation of uncertainty

(using the calculus of probability theory)

- Perception = state estimation
- Action = utility optimization

Example of Probabilistic Robotic Systems: Mobile Robot



- 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

Example of Probabilistic Robotic Systems: Robot Arm

- State: Joint encoders, object pose, object velocity
- Sensors:
  - Joint Encoders: Measure position and velocity at different joints
  - Camera images: Informs the position and semantics of objects in the scene
  - Depth images: Indicates the 3-D position and occupancy of object in the scene
- Dynamics:
  - Noise from: unmodeled contact dynamics, non-rigid contact or unmodeled friction

Fundamental Axioms of Probability

$$0 \le \Pr(A) \le 1$$

$$\Pr(\Omega) = 1 \qquad \Pr(\phi) = 0$$

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

- 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.

Useful Corollaries from Axioms

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

If A and B have no overlap then

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

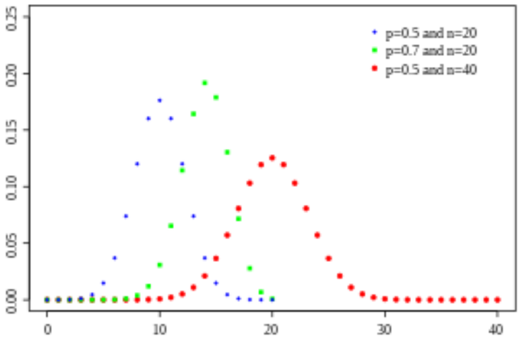
Discrete Random Variables

- ***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 $\boldsymbol{x}_i$, between [0, 1]
- ***P(·)*** is called probability mass function (sums to 1)
- E.g. $P(Room) = \langle 0.7, 0.2, 0.08, 0.02 \rangle$

Examples of Discrete Random Variables

Binomial

$$\binom{n}{k} p^k q^{n-k}$$

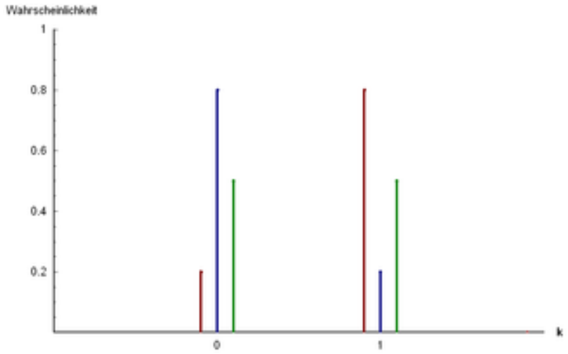


Multinomial

$$\frac{n!}{x_1! \cdots x_k!} p_1^{x_1} \cdots p_k^{x_k}$$

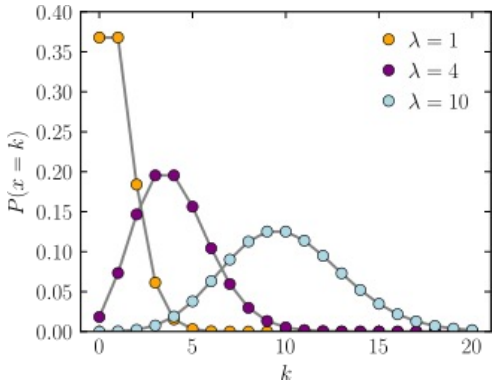
Bernoulli

$$\begin{cases} q = 1 - p & \text{if } k = 0 \\ p & \text{if } k = 1 \end{cases}$$



Poisson

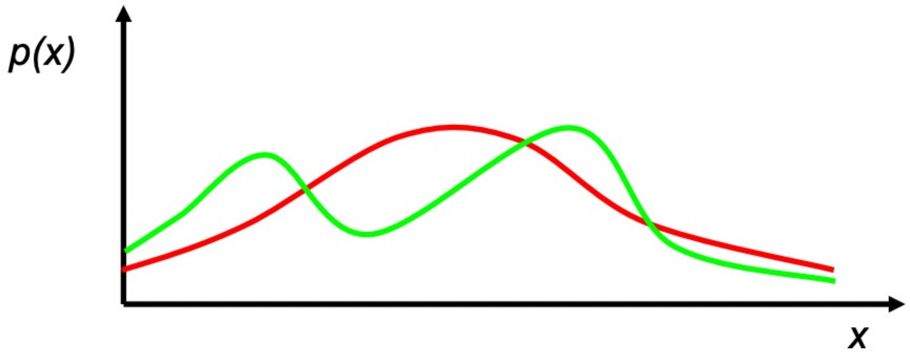
$$\frac{\lambda^k e^{-\lambda}}{k!}$$



Continuous Random Variables

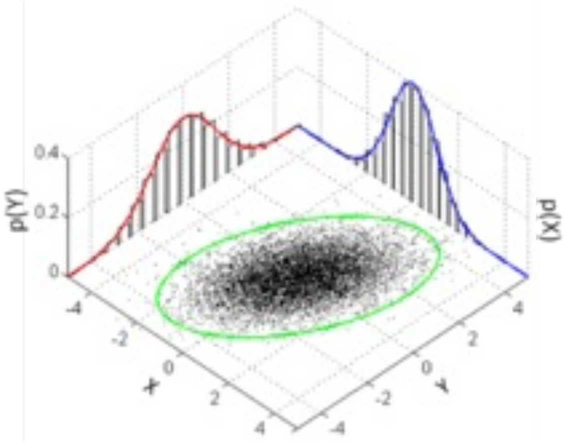
- ***X*** denotes a random variable.
- ***X*** can take on a continuum of values in the support of the probability density function
- ***P(X=x)***, or ***P(x)***, is the probability density function
  - Density function positive but not upper bounded by 1
  - Integrates to 1

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

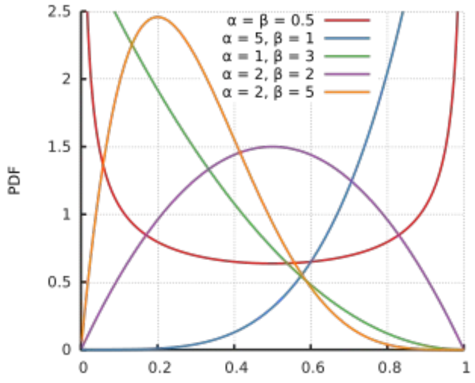


Examples of Continuous Random Variables

Multivariate Gaussian

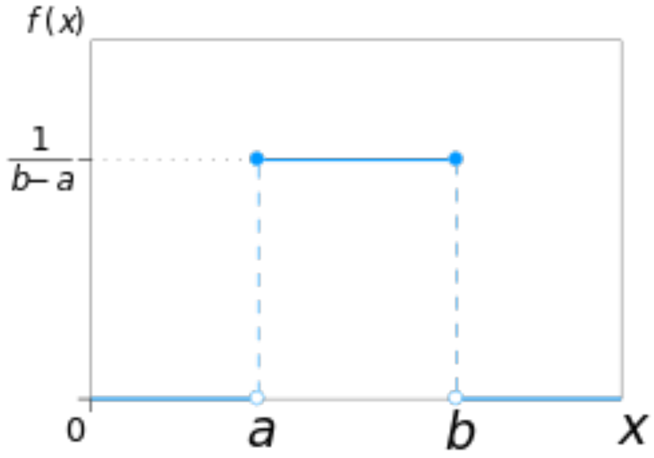


Beta Distribution



Uniform Distribution

$$\begin{cases} \frac{1}{b-a} & \text{for } x \in [a, b] \\ 0 & \text{otherwise} \end{cases}$$



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

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

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

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

Events

- P(+x, +y) ?
- P(+x) ?
- P(-y OR +x) ?
- Independent?

$P(X, Y)$

X	Y	P
+x	+y	0.2
+x	-y	0.3
-x	+y	0.4
-x	-y	0.1

Marginal Distributions

$P(X, Y)$

X	Y	P
+x	+y	0.2
+x	-y	0.3
-x	+y	0.4
-x	-y	0.1

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

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

$P(X)$

X	P
+x	
-x	

$P(Y)$

Y	P
+y	
-y	

Conditional Probabilities

$$P(X, Y)$$

X	Y	P
+x	+y	0.2
+x	-y	0.3
-x	+y	0.4
-x	-y	0.1

- P(+x | +y) ?
- P(-x | +y) ?
- P(-y | +x) ?

Lecture Outline

Probability Review

↓

Bayesian Inference

↓

Bayesian Filtering

Bayes Formula

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

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

Bayes Formula

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

$$P(y) = \sum_{x'} P(y \mid x')P(x')$$

$$\Downarrow$$

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

$$\eta = \frac{1}{\sum_x P(y, x)}$$

Can replace with integral

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

Example of Bayes Formula in Action

Cancer \ Symptom	Yes	No	Total
Yes	1	0	1
No	10	99989	99999
Total	11	99989	100000

Just because everyone with cancer has the symptom, doesn't mean everyone with the symptom has cancer

$$
\begin{aligned}
P(\text{Cancer}|\text{Symptoms}) &= \frac{P(\text{Symptoms}|\text{Cancer})P(\text{Cancer})}{P(\text{Symptoms})} \\
&= \frac{P(\text{Symptoms}|\text{Cancer})P(\text{Cancer})}{P(\text{Symptoms}|\text{Cancer})P(\text{Cancer}) + P(\text{Symptoms}|\text{Non-Cancer})P(\text{Non-Cancer})} \\
&= \frac{1 \times 0.00001}{1 \times 0.00001 + (10/99999) \times 0.99999} = \frac{1}{11} \approx 9.1\%
\end{aligned}
$$

Why Bayes Formula?

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

$$P(y) = \sum_{x'} P(y \mid x')P(x')$$

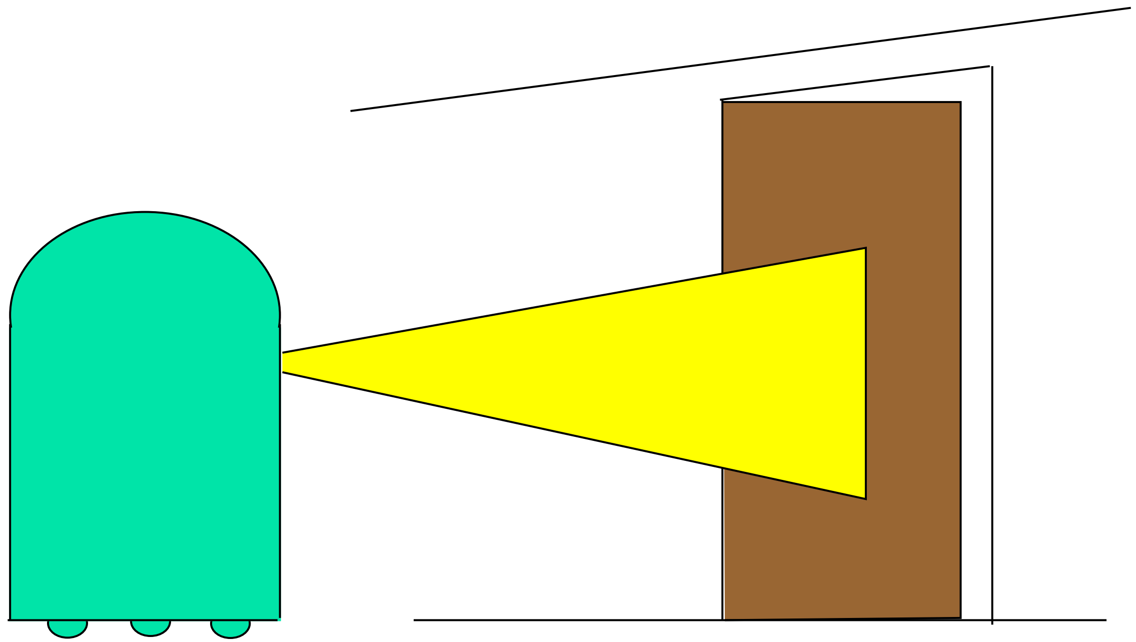
Diagnostic

Causal

- Causal knowledge may be easier to obtain/estimate/represent
- Which direction is causal is not always clear though!
- Allows us to estimate “beliefs” based on “measurements”

Simple Example of State Estimation

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



Example

$$P(z \mid open) = 0.6 \qquad P(z \mid \neg open) = 0.3$$

$$P(open) = P(\neg open) = 0.5$$

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

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

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

Conditioning

- Bayes rule and background knowledge:

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

$$P(x|y) \stackrel{?}{=} \int P(x|y,z)P(z)dz$$

$$\stackrel{?}{=} \int P(x|y,z)P(z|y)dz$$

$$\stackrel{?}{=} \int P(x|y,z)P(y|z)dz$$

Conditional Independence

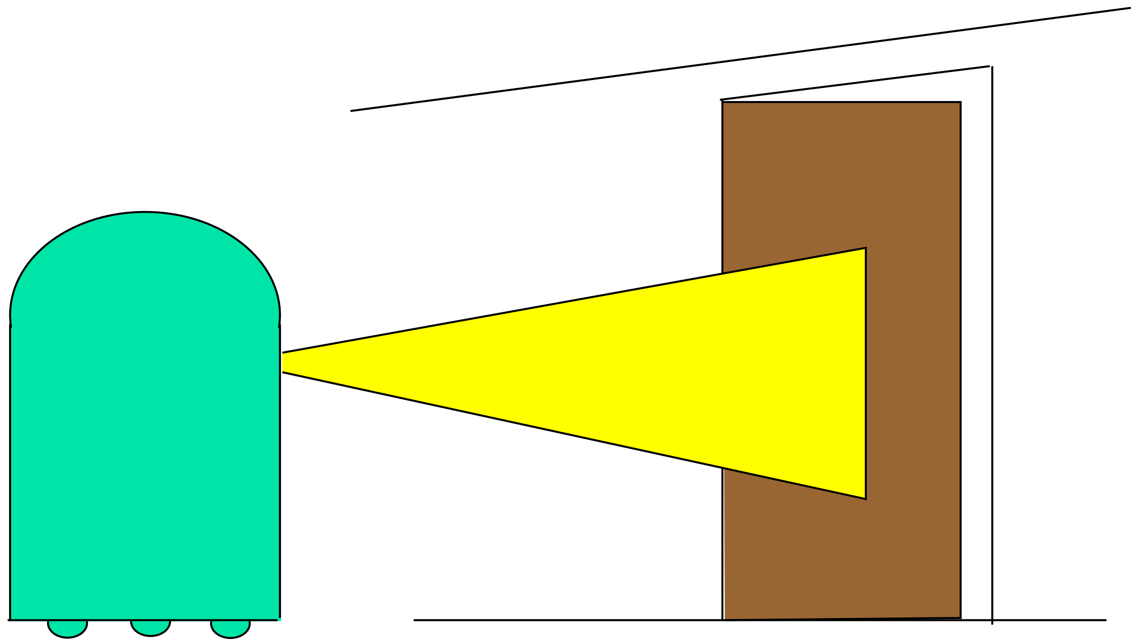
$$P(x, y \mid z) = P(x \mid z) P(y \mid z)$$

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

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

Second Simple Example of State Estimation

- Suppose our robot obtains another observation z_2.
- What is $\boldsymbol{P(open \mid z_1, z_2)}$?



Recursive Bayesian Updating

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

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

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

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

Example: Second Measurement

$$P(z_2 \mid open) = 0.5 \qquad P(z_2 \mid \neg open) = 0.6$$

$$P(open \mid z_1) = 2/3 \qquad P(\neg open \mid z_1) = 1/3$$

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

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

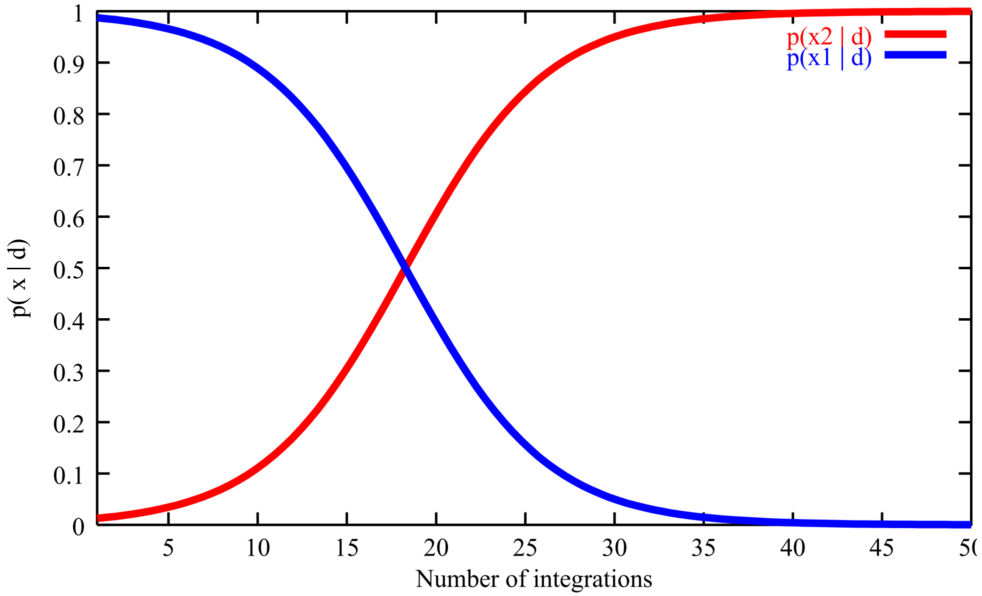
- z_2 lowers the probability that the door is open.

Effects of Incorrect Independencies

$$P(x|z_1,\ldots,z_n) = \eta_{1:n} \prod_{i=1,\ldots,n} P(z_i|x)P(x)$$

- If redundant sensors, $z_1 \ldots z_n$ are treated as independent, leads to double counting

  → overconfident predictions



Why is it challenging to be Bayesian?

Why is it challenging to be Bayesian?

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

$$P(y) = \sum_{x'} P(y \mid x')P(x')$$

$$p(y) = \int p(y|x')p(x')dx'$$

Difficult to compute analytically because of integral

How can we address this?

Partition function

$$p(y) = \int p(y|x')p(x')dx'$$

Markov Chain Monte Carlo

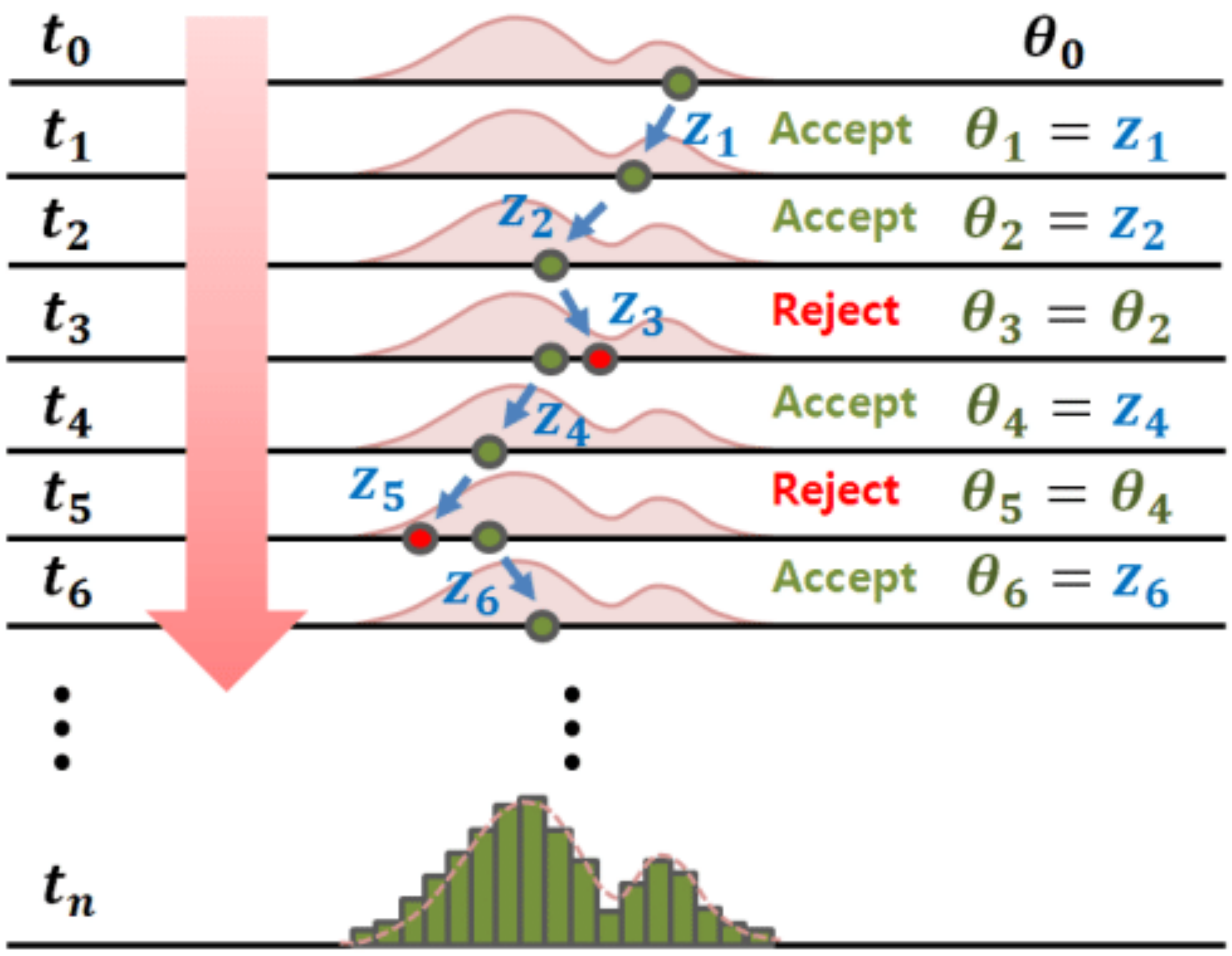
Conjugate Priors

Variational Inference

Discretization

Markov-Chain Monte Carlo

Construct a Markov chain with equilibrium distribution equal to joint
→ Inference via sampling



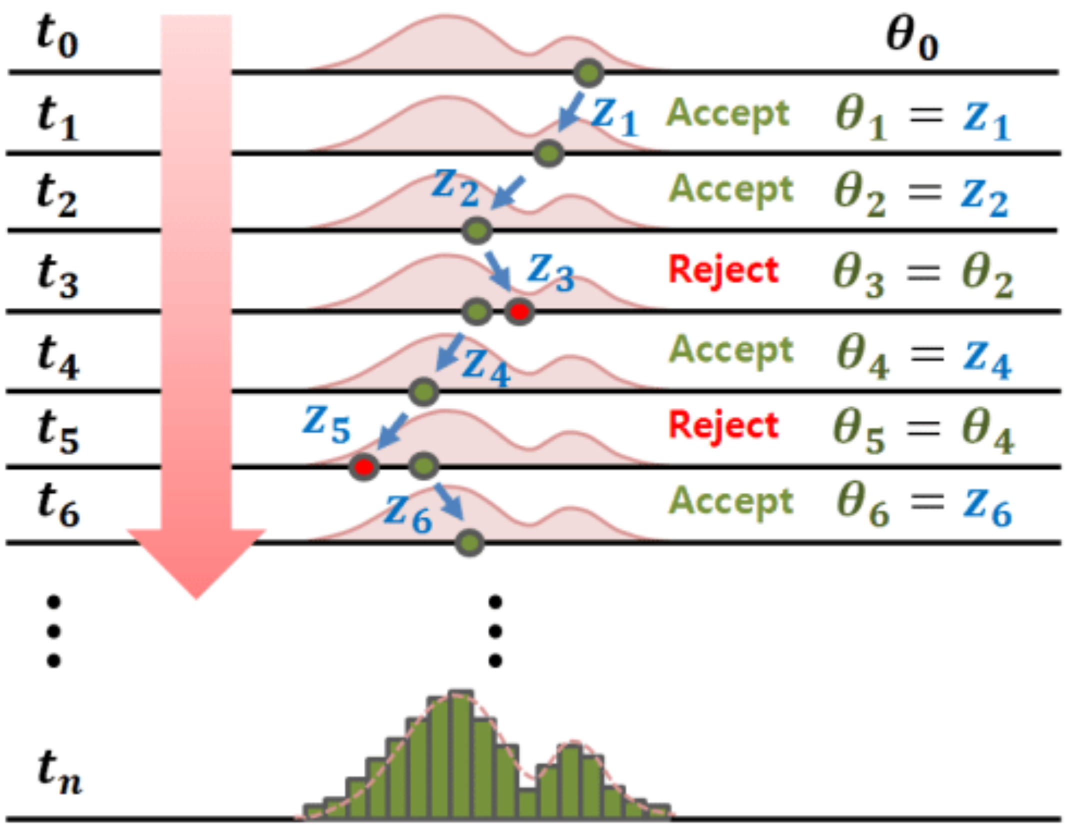
Markov-Chain Monte Carlo: Metropolis Hastings

A simple MCMC algorithm:
Assume access to an unnormalized p(x, y)

1. Start at some x, given a y
2. Propose a new x' according to some symmetric q(x'|x)
3. Compute acceptance ratio

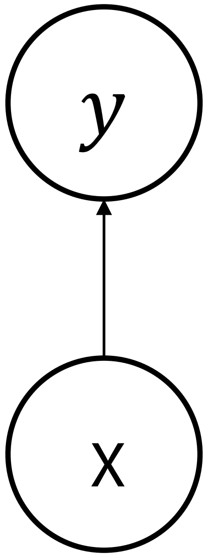
$$\alpha = \frac{p(x', y)}{p(x, y)}$$

4. Accept x' with likelihood α



Hard to converge in high dimensions

Variational Inference



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

$$p(y) = \int p(y|x')p(x')dx'$$

Instead of explicitly computing posterior, approximate with a tractable family

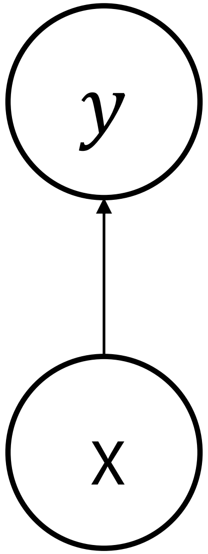
$$q(x|y) \longleftrightarrow p(x|y)$$

Gaussian approximation

$$\min_q D_{\mathrm{KL}}(q(x|y)||p(x|y))$$

Actually optimize a lower bound

Variational Inference: Evidence Lower bound



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

Instead of explicitly computing posterior, approximate with a tractable family

$$q(x|y) \longleftrightarrow p(x|y)$$

$$\min_q D_{\text{KL}}(q(x|y)||p(x|y))$$

$$\geq \mathbb{E}(p(y|x)p(x)) + \mathcal{H}(q(y|x))$$

Tractable optimization → inference becomes optimization

Conjugate Priors

	PDF
Binomial	$f(x) = \binom{n}{x} \theta^{x} (1-\theta)^{n-x}$ ↓ the function of x
Beta	$g(\theta) = \frac{1}{B(\alpha,\beta)} \theta^{\alpha-1} (1-\theta)^{\beta-1}$ ↓ the function of θ

$$P(\theta|X) = \frac{P(X|\theta) \cdot P(\theta)}{\int_{\theta} P(X|\theta) \cdot P(\theta)\, d\theta}$$

posterior: $P(\theta|X)$; sampling: $P(X|\theta)$; prior: $P(\theta)$; X ↓ data; ∫ → normalizing constant; θ ↓ for every possible θ

$$= \frac{\binom{n}{x} \theta^{x} (1-\theta)^{n-x} \cdot \frac{1}{B(\alpha,\beta)} \theta^{\alpha-1} (1-\theta)^{\beta-1}}{\int_0^1 \binom{n}{x} \theta^{x} (1-\theta)^{n-x} \cdot \frac{1}{B(\alpha,\beta)} \theta^{\alpha-1} (1-\theta)^{\beta-1}\, d\theta}$$

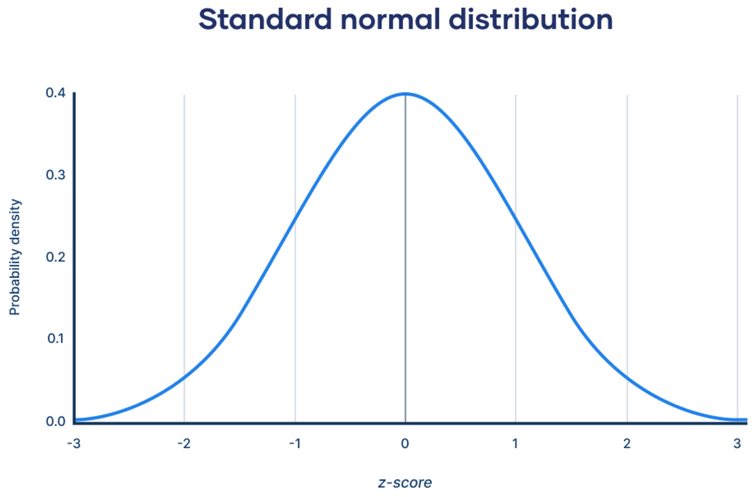
$$= \frac{\frac{{}_nC_x}{B(\alpha,\beta)} \theta^{x+\alpha-1} (1-\theta)^{n-x+\beta-1}}{\frac{{}_nC_x}{B(\alpha,\beta)} \underbrace{\int_0^1 \theta^{x+\alpha-1} (1-\theta)^{n-x+\beta-1}\, d\theta}_{B(x+\alpha,\, n-x+\beta)}}$$

$$= Beta(x+\alpha,\, n-x+\beta)$$

Normal posterior:

Normal prior * Normal likelihood → Normal posterior

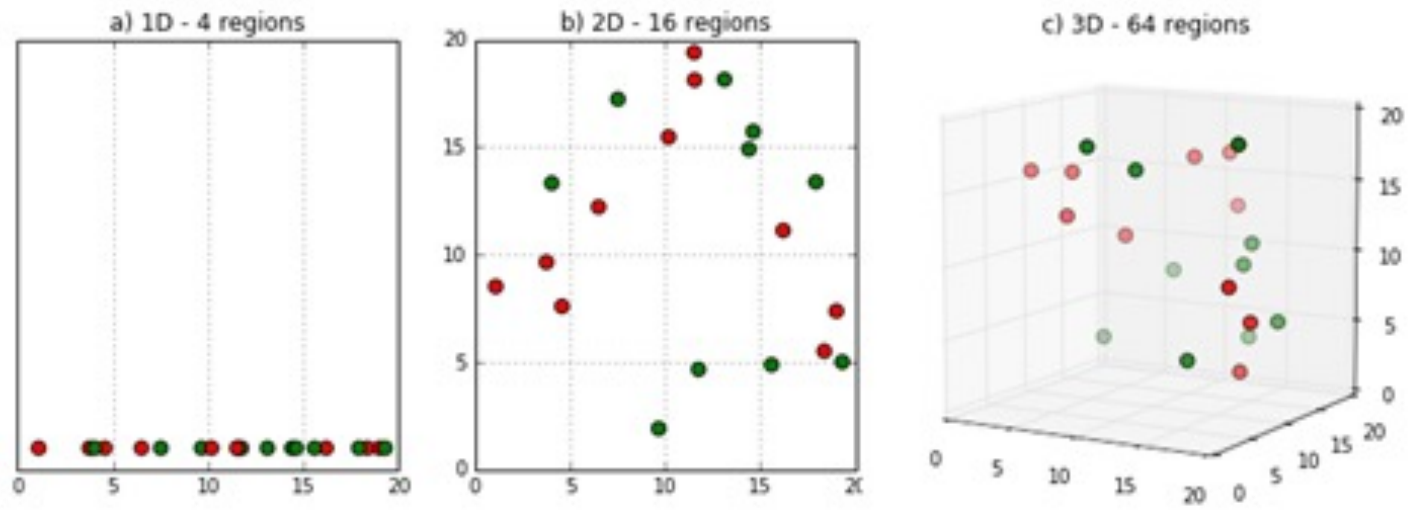
$$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \qquad \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \qquad \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$



Discretization

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

$$P(y) = \sum_{x'} P(y \mid x')P(x')$$



Grows exponentially with dimension!

Lecture Outline

Probability Review

↓

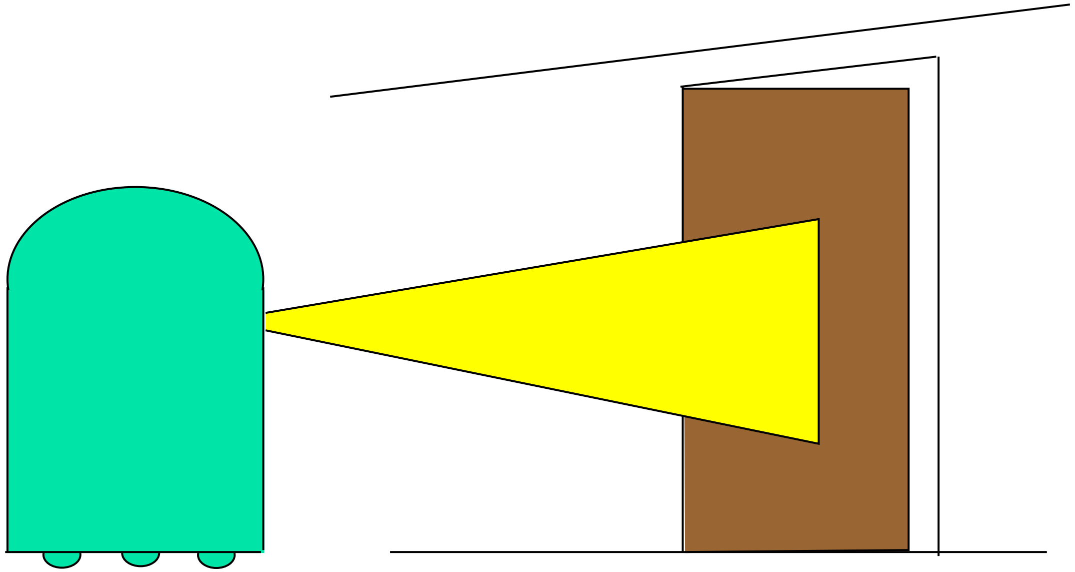
Bayesian Inference

↓

Bayesian Filtering

Let’s estimate “state” of our robot

- What affects uncertainty:
  - Robot actions (increase uncertainty typically)
  - Sensor measurements (decrease uncertainty typically)

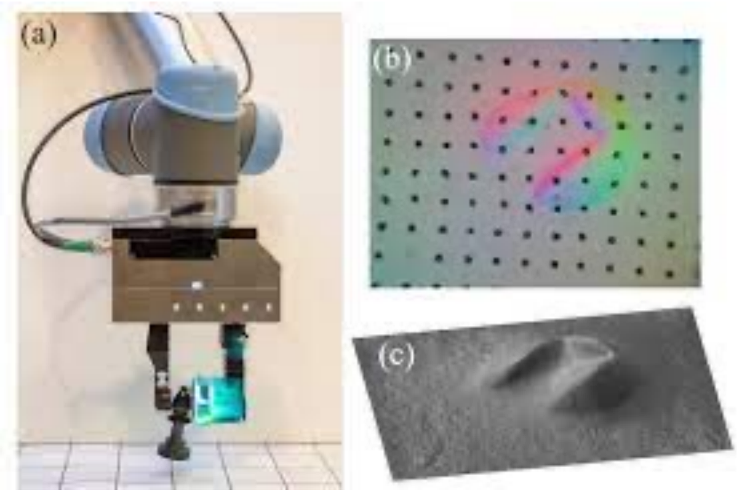


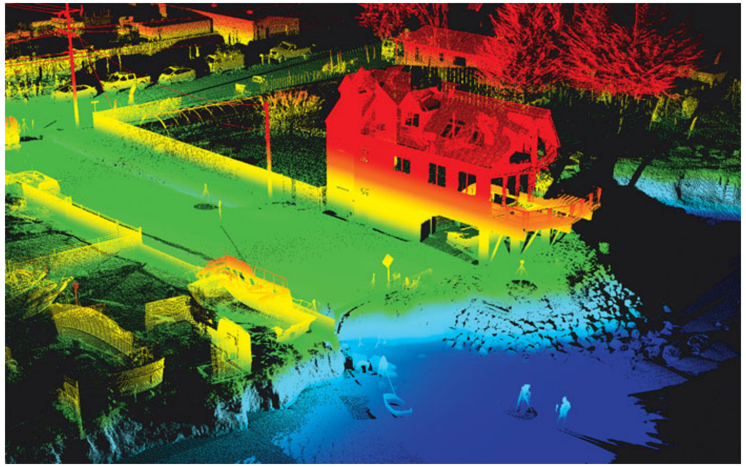
How do actions increase uncertainty?

- Actions transition the state of the system forward x → x′
  - But they may (and usually) do so with errors/noise!
- Robot wheels have slippage/noise, joints have stochasticity, environment introduces noise

How do sensors reduce uncertainty?

- Measurements usually convey more information about the state of the world
- Sensor readings can range from images to laser scans to tactile sensing, each of which has a different effect on uncertainty





Filtering

- Filtering is the process of making sense ("filtering") of sensor measurements and actions to estimate the system state
- Many different types of filters:
  - Matched filters (known signal)
  - Wiener filters (signal from noise)
  - Bayesian filters (bayesian state estimation)
    - Kalman
    - EKF / UKF
    - ….

Bayes Filters: Framework

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

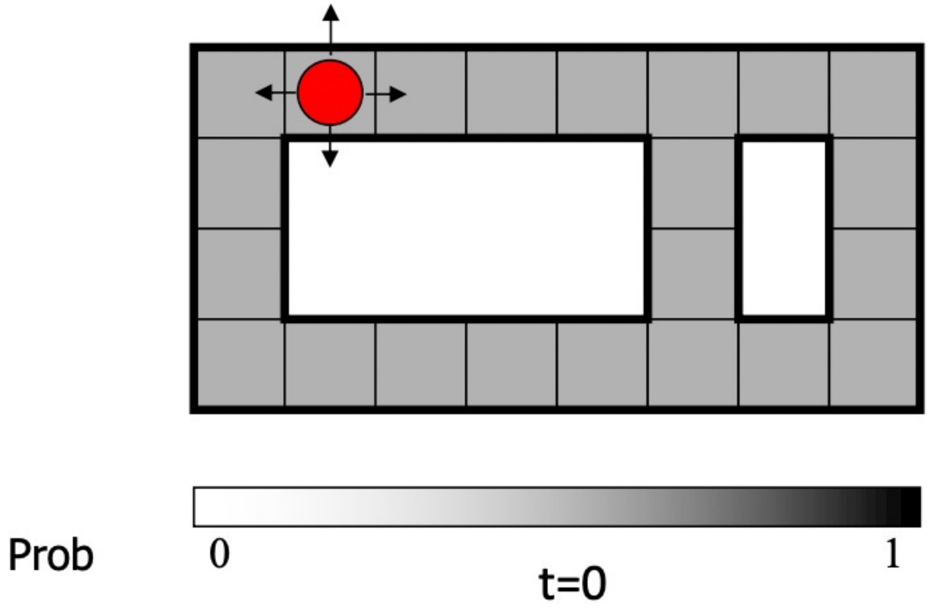
$$d_t = \{z_0, u_0, z_1, u_1, \ldots, z_t\}$$

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

$$Bel(x_t) = P(x_t | u_{0:t-1}, z_{0:t})$$

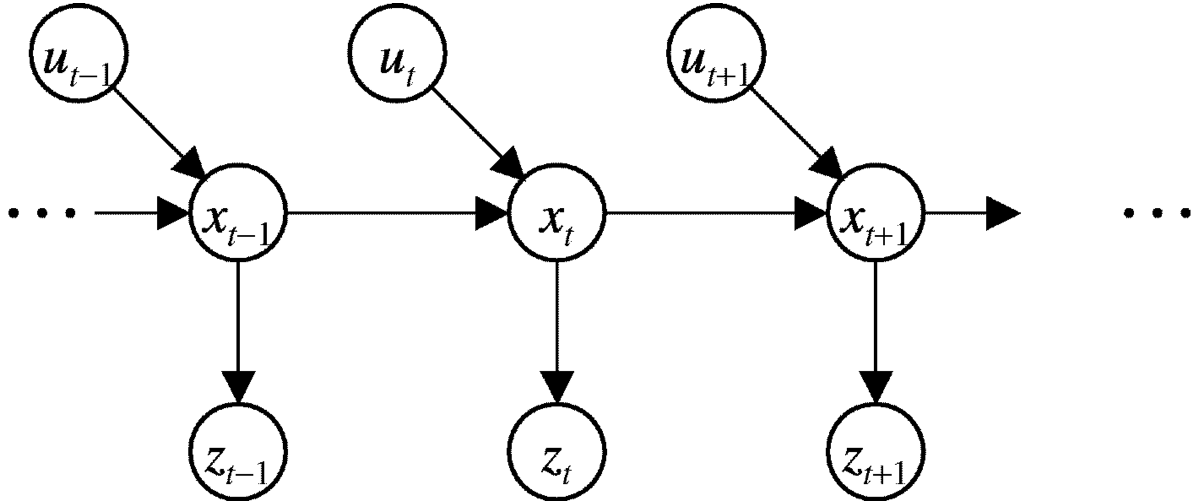
Example Situation for Filtering

"Where is my robot?"



- Sensor model: never more than 1 mistake
- Know the heading (North, East, South or West)
- Motion model: may not execute action with small prob.

Markov Assumption



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

Underlying Assumptions

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

Bayes Filters

z = observation
u = action
x = state

$$Bel(x_t) = P(x_t | u_{0:t-1}, z_{0:t})$$

We want to recursively express $Bel(x_t)$ in terms of three entities

$$p(z_t | x_t)$$

Measurement

$$p(x_t | x_{t-1}, u_{t-1})$$

Dynamics

$$Bel(x_{t-1})$$

Previous Belief

Bayes Filters: Intuition

z = observation
u = action
x = state

$$Bel(x_t) = P(x_t | u_{0:t-1}, z_{0:t})$$

We want to recursively express Bel(x_t) in terms of three entities

Integrate in effect of action

$$Bel(x_{t-1}) \quad + \quad p(x_t | x_{t-1}, u_{t-1}) \longrightarrow \overline{Bel}(x_t)$$

Previous Belief

Dynamics

With integration

Bayes Filters: Intuition

z = observation
u = action
x = state

$$Bel(x_t) = P(x_t | u_{0:t-1}, z_{0:t})$$

We want to recursively express Bel(x_t) in terms of three entities

Integrate in Measurement

$$\overline{Bel}(x_t) \quad + \quad p(z_t | x_t) \quad \longrightarrow \quad Bel(x_t)$$

Previous Belief

Measurement

With normalization

Bayes Filters

z = observation
u = action
x = state

$$Bel(x_t) = P(x_t | u_{0:t-1}, z_{0:t})$$

Bayes $$= \eta \, p(z_t | x_t, u_{0:t-1}, z_{0:t-1}) P(x_t | u_{0:t-1}, z_{0:t-1})$$

Remember: Bayes Rule

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

$$\eta = \frac{1}{\sum_x P(y, x)}$$

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

Bayes Filters

z = observation
u = action
x = state

$$Bel(x_t) = P(x_t | u_{0:t-1}, z_{0:t})$$

Bayes $$= \eta \, p(z_t | x_t, u_{0:t-1}, z_{0:t-1}) P(x_t | u_{0:t-1}, z_{0:t-1})$$

Markov $$= \eta \, p(z_t | x_t) P(x_t | u_{0:t-1}, z_{0:t-1})$$

Remember: Markov Property

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

Bayes Filters

z = observation
u = action
x = state

$$Bel(x_t) = P(x_t | u_{0:t-1}, z_{0:t})$$

Bayes $$= \eta \, p(z_t | x_t, u_{0:t-1}, z_{0:t-1}) P(x_t | u_{0:t-1}, z_{0:t-1})$$

Markov $$= \eta \, p(z_t | x_t) P(x_t | u_{0:t-1}, z_{0:t-1})$$

Total prob.

$$= \eta \, p(z_t | x_t) \int P(x_t | u_{0:t-1}, z_{0:t-1}, x_{t-1}) P(x_{t-1} | u_{0:t-1}, z_{0:t-1}) dx_{t-1}$$

Remember: Marginalization

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

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

Bayes Filters

z = observation
u = action
x = state

$$Bel(x_t) = P(x_t | u_{0:t-1}, z_{0:t})$$

Bayes $$= \eta \, p(z_t | x_t, u_{0:t-1}, z_{0:t-1}) P(x_t | u_{0:t-1}, z_{0:t-1})$$

Markov $$= \eta \, p(z_t | x_t) P(x_t | u_{0:t-1}, z_{0:t-1})$$

Total prob.

$$= \eta \, p(z_t | x_t) \int P(x_t | u_{0:t-1}, z_{0:t-1}, x_{t-1}) P(x_{t-1} | u_{0:t-1}, z_{0:t-1}) dx_{t-1}$$

Markov $$= \eta \, p(z_t | x_t) \int P(x_t | u_{t-1}, x_{t-1}) P(x_{t-1} | u_{0:t-1}, z_{0:t-1}) dx_{t-1}$$

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

Understanding Bayes Filters

z = observation
u = action
x = state

$$Bel(x_t) = P(x_t | u_{0:t-1}, z_{0:t})$$
$$= \eta \, p(z_t | x_t) \int P(x_t | u_{t-1}, x_{t-1}) Bel(x_{t-1}) dx_{t-1}$$

Step 1: Dynamics Update

Incorporate the effect of motion on uncertainty (typically increases)

Understanding Bayes Filters

z = observation
u = action
x = state

$$Bel(x_t) = P(x_t|u_{0:t-1}, z_{0:t})$$
$$= \eta \, p(z_t|x_t) \int P(x_t|u_{t-1}, x_{t-1}) Bel(x_{t-1}) dx_{t-1}$$

Step 2: Measurement Update

Incorporate the effect of new measurements on uncertainty (typically decreases)

Understanding Bayes Filters

z = observation
u = action
x = state

$$Bel(x_t) = P(x_t|u_{0:t-1}, z_{0:t})$$
$$= \eta\, p(z_t|x_t) \int P(x_t|u_{t-1}, x_{t-1}) Bel(x_{t-1}) dx_{t-1}$$

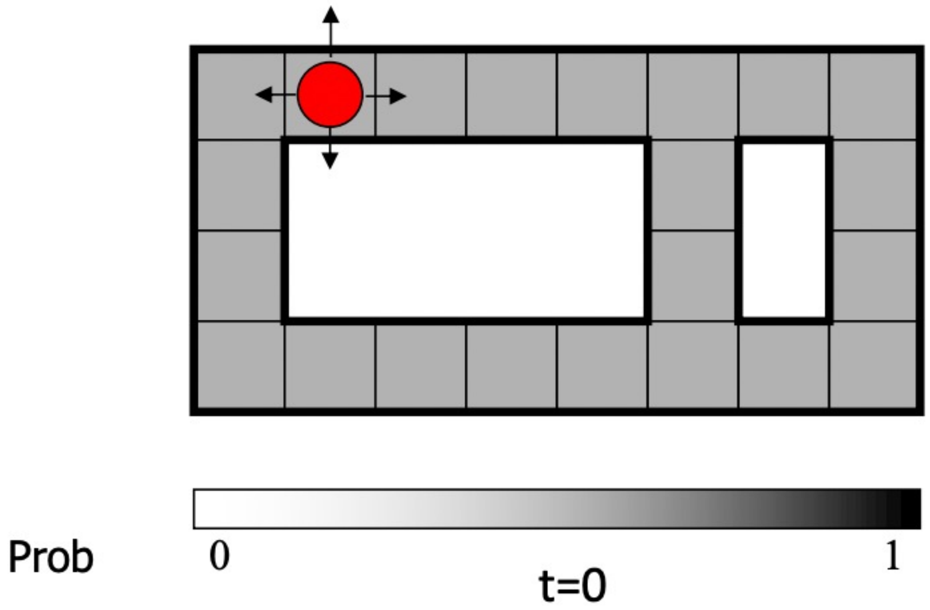
All Bayes filter iterate between performing the dynamics (prediction) step and the measurement (correction) step

Bayes Filter Algorithm

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

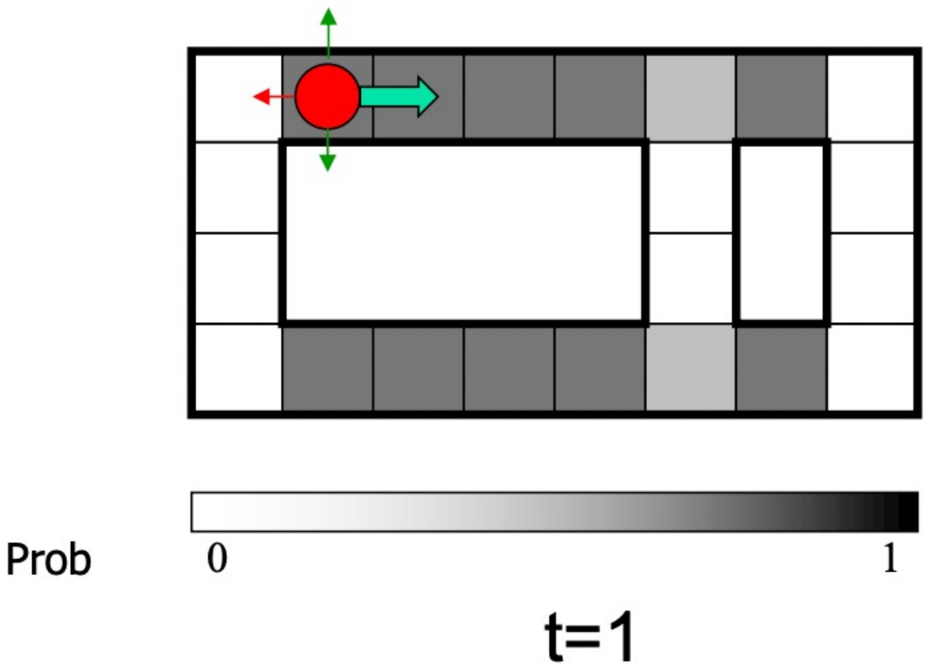
Algorithm **Bayes_filter**(*Bel(x),d*):
$n=0$
If *d* is a perceptual data item *z* then
    For all *x* do
        $Bel'(x) = P(z \mid x) Bel(x)$
        $\eta = \eta + Bel'(x)$
    For all *x* do
        $Bel'(x) = \eta^{-1} Bel'(x)$
Else if *d* is an action data item *u* then
    For all *x* do
  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

Example Run for Localization



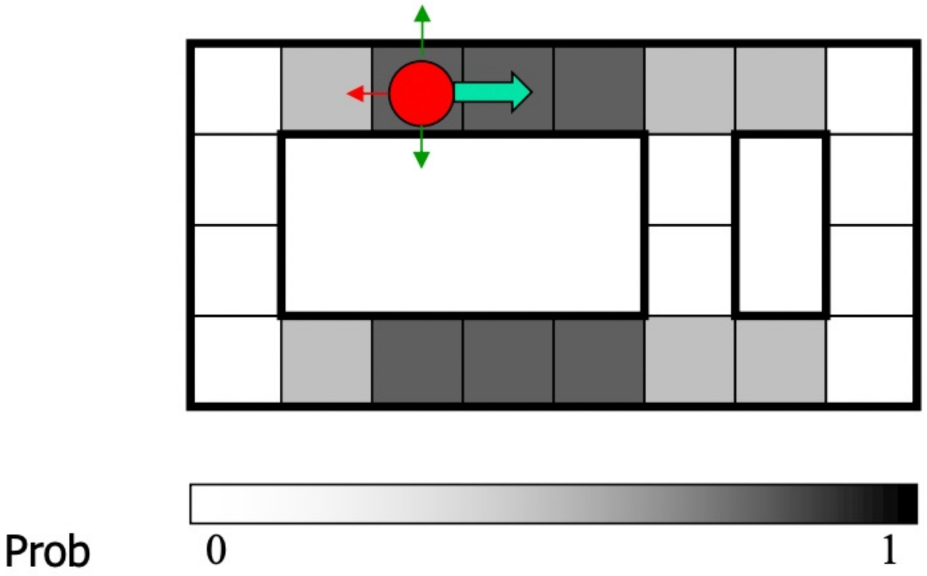
- Sensor model: never more than 1 mistake
- Know the heading (North, East, South or West)
- Motion model: may not execute action with small prob.

Example Run for Localization



- Sensor model: never more than 1 mistake
- Know the heading (North, East, South or West)
- Motion model: may not execute action with small prob.

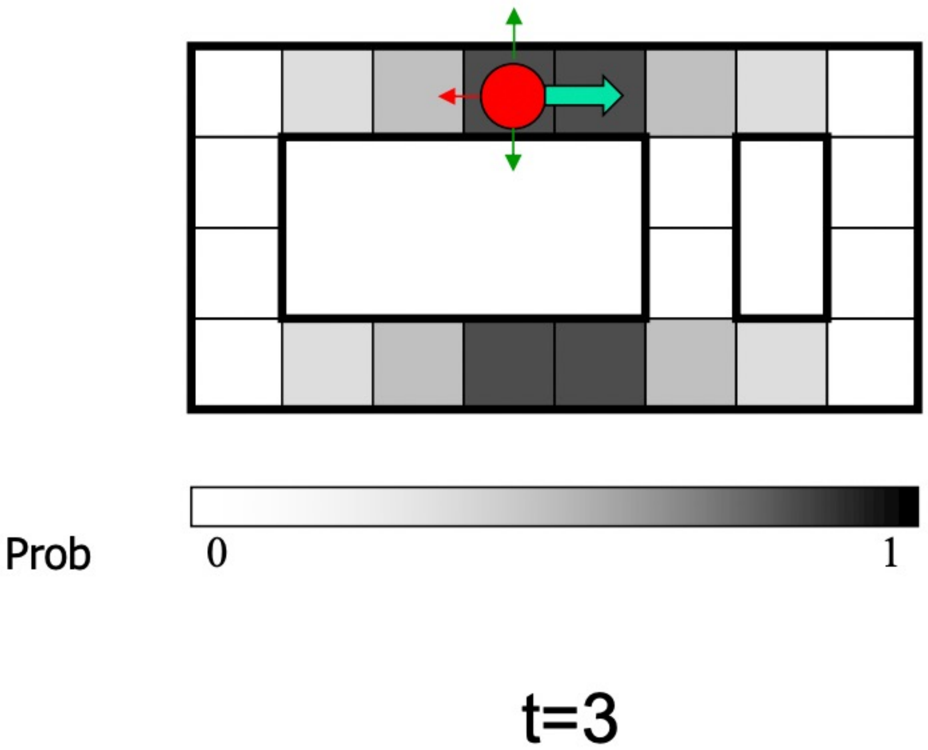
Example Run for Localization



t=2

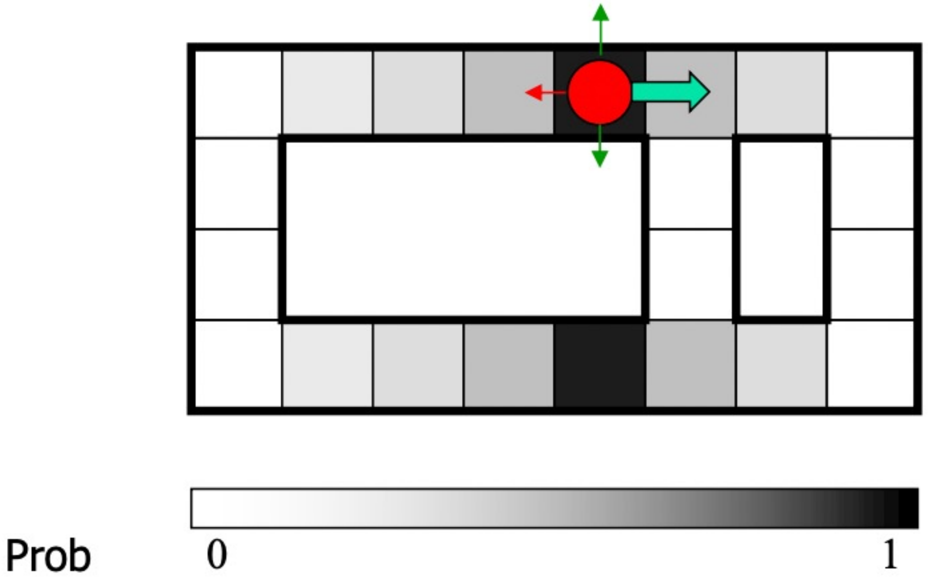
- Sensor model: never more than 1 mistake
- Know the heading (North, East, South or West)
- Motion model: may not execute action with small prob.

Example Run for Localization



- Sensor model: never more than 1 mistake
- Know the heading (North, East, South or West)
- Motion model: may not execute action with small prob.

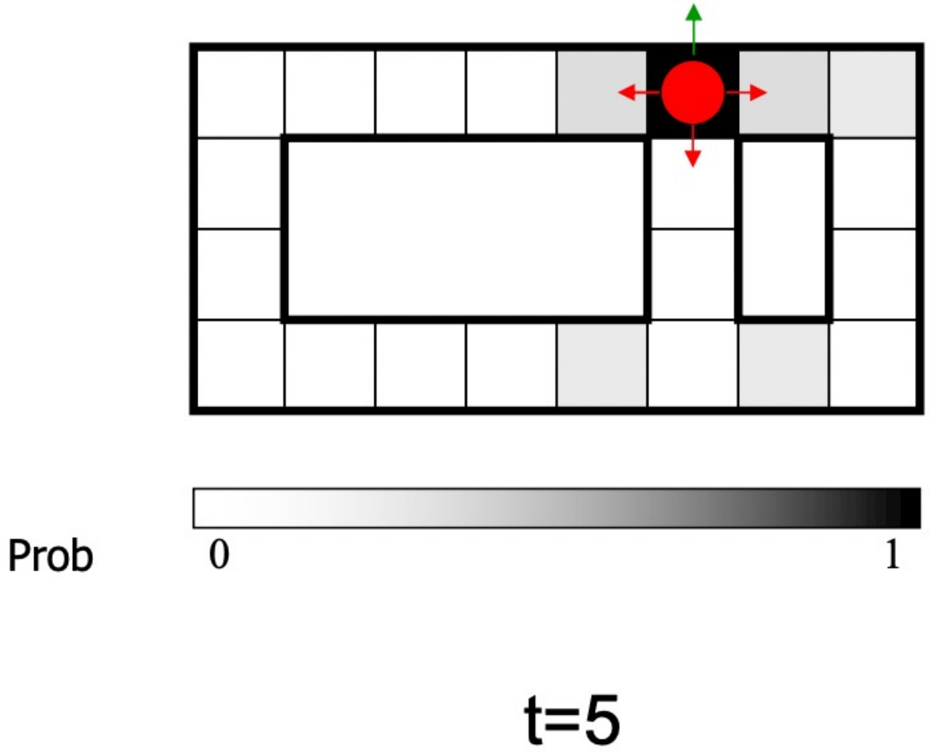
Example Run for Localization



t=4

- Sensor model: never more than 1 mistake
- Know the heading (North, East, South or West)
- Motion model: may not execute action with small prob.

Example Run for Localization



- Sensor model: never more than 1 mistake
- Know the heading (North, East, South or West)
- Motion model: may not execute action with small prob.

Representations for Bayesian Robot Localization

Discrete approaches (’ 95)
- Topological representation (’ 95)
  - uncertainty handling (POMDPs)
  - occas. global localization, recovery
- Grid-based, metric representation (’ 96)
  - global localization, recovery

Particle filters (’ 99)
- sample-based representation
- global localization, recovery

AI

Kalman filters (late-80s)
- Gaussians, unimodal
- approximately linear models
- position tracking

Robotics

Multi-hypothesis (’ 00)
- multiple Kalman filters
- global localization, recovery

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)

Why is this difficult?

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

Tractable Bayesian inference is challenging in the general case

We will work out the conjugate prior and discrete case,
leaving the MCMC/VI cases as an exercise

Lecture Outline

Probability Review

↓

Bayesian Inference

↓

Bayesian Filtering

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.