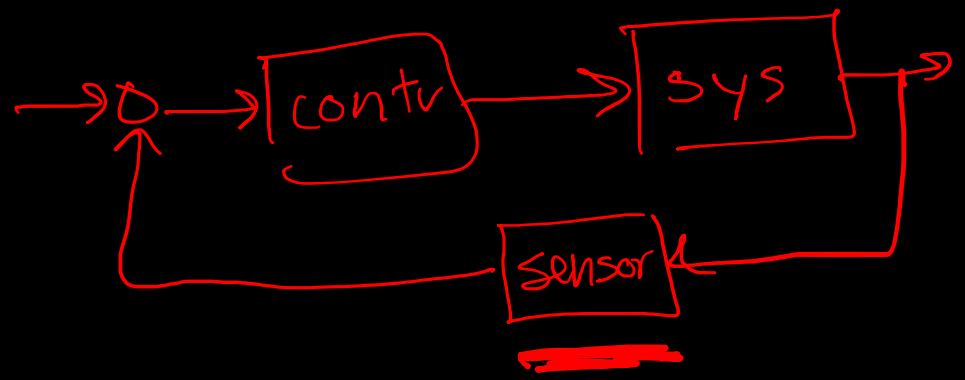


Lecture 11: Filtering & Localization

Professor Katie Driggs-Campbell

March 9, 2021

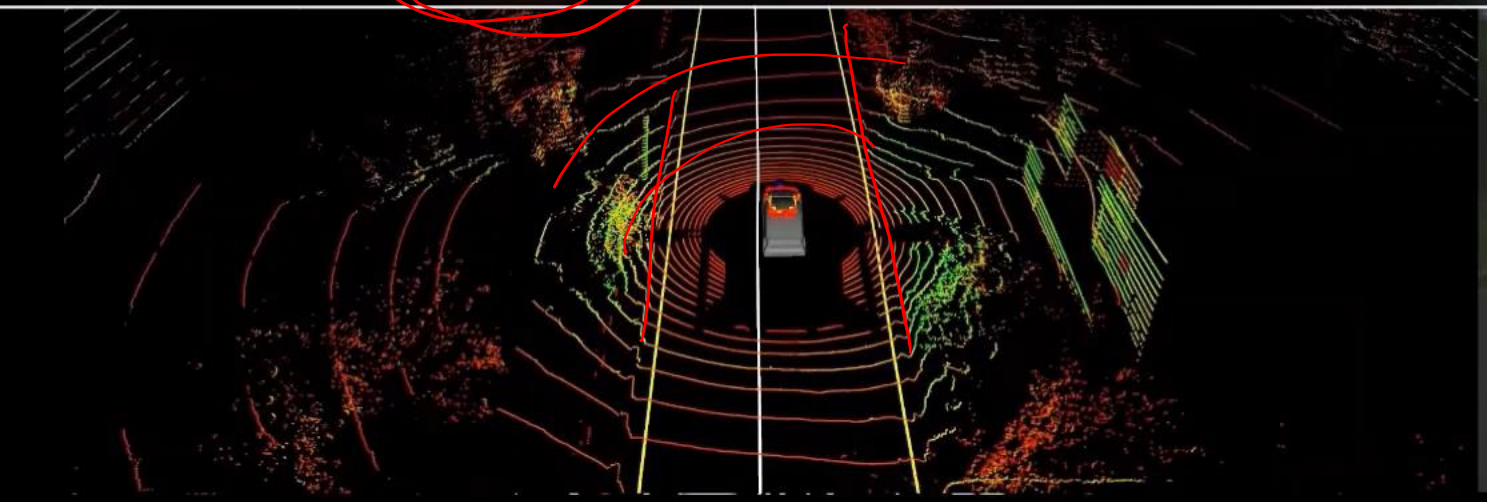
ECE484: Principles of Safe Autonomy

Administrivia

- MP2 tutorial on Thursday
  - Due 3/26
- Milestone report due Friday 3/19 by 5pm
  - Submitted via Gradescope
  - Rubric now online!

Thought Experiment

Navigating Intersections



SAIC

SISL
Stanford Intelligent Systems Laboratory

Filtering and Localization Use Cases

HD Maps

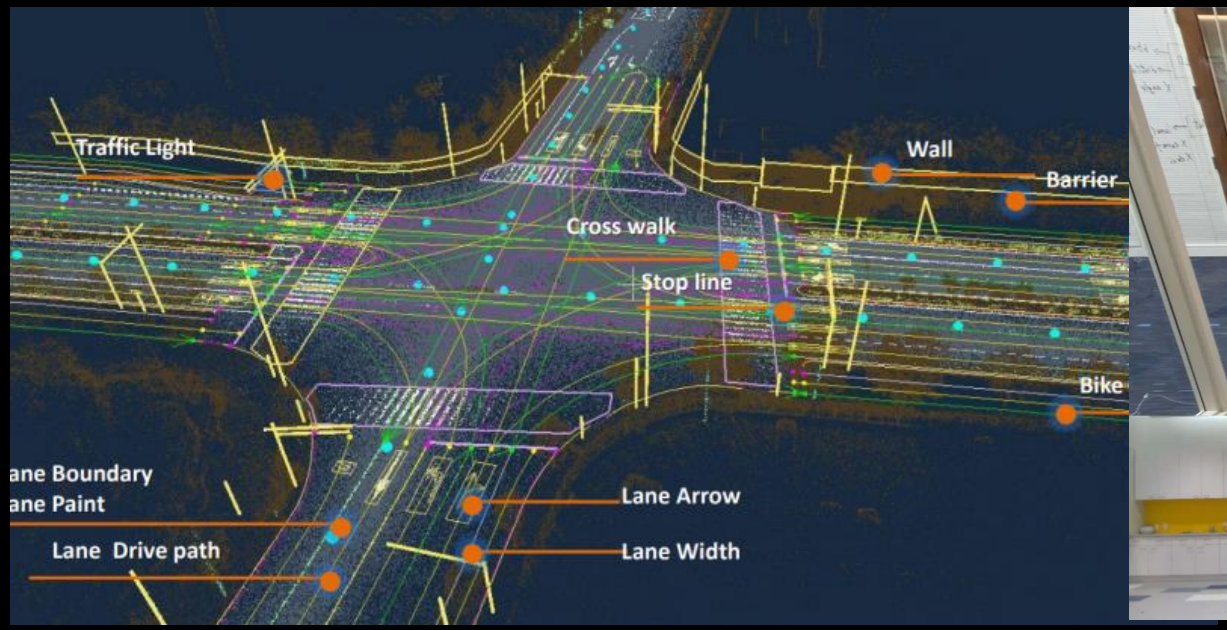


Image Credit: BMW China

Object Tracking

Today’s Plan

- What is filtering, mapping, and localization?
  - Probability review!
- Bayes Filters (discrete)
- Kalman Filters (continuous)
- Particle Filters

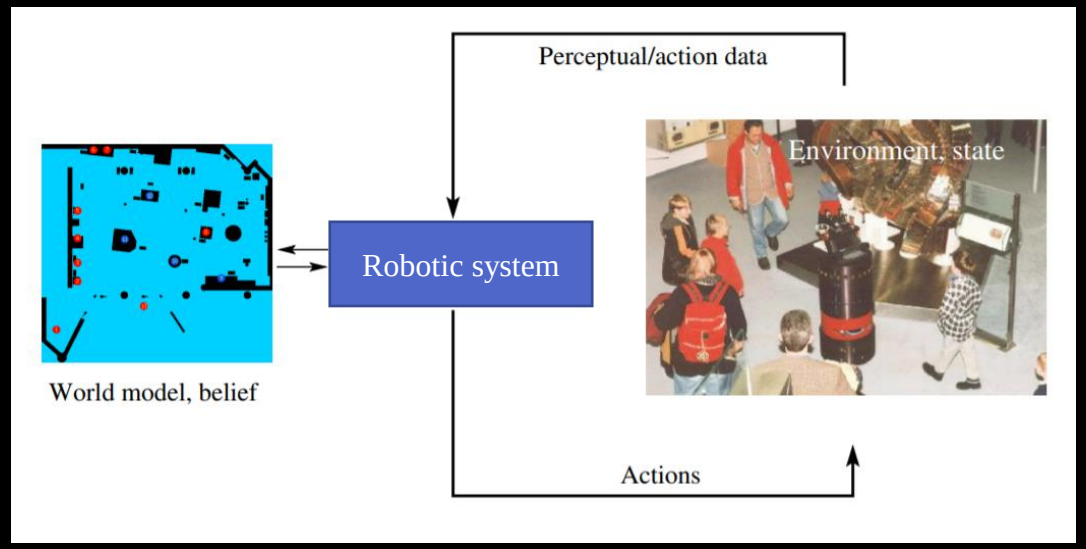
Today's Plan

- What is filtering, mapping, and localization?
  - Probability review!
- Bayes Filters (discrete)
- Kalman Filters (continuous)
- Particle Filters

Robot States and the Environment

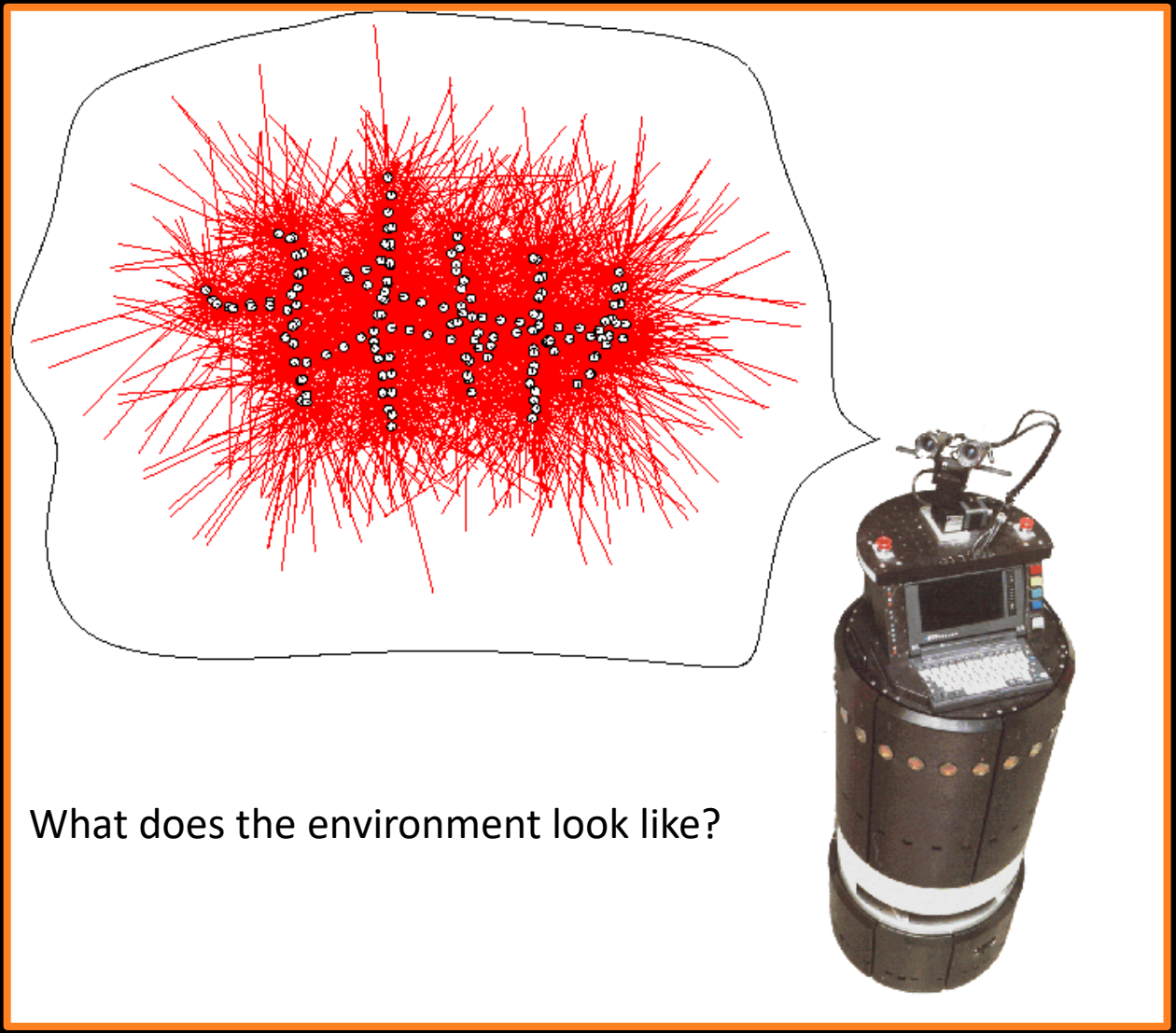
- **State** represents the environment as well as the robot, for example:
  - location of walls or objects (environment or static)
  - pose of the robot (physical or dynamical)
- **Environment interaction** comes in the form of
  - Sensor measurements
  - Control actions
- **Internal representation (or belief)** of the state of the world
  - In general, the state (or the world) cannot be measured directly
  - Perception is the process by which the robot uses its sensors to obtain information about the state of the environment

Maps and Representations



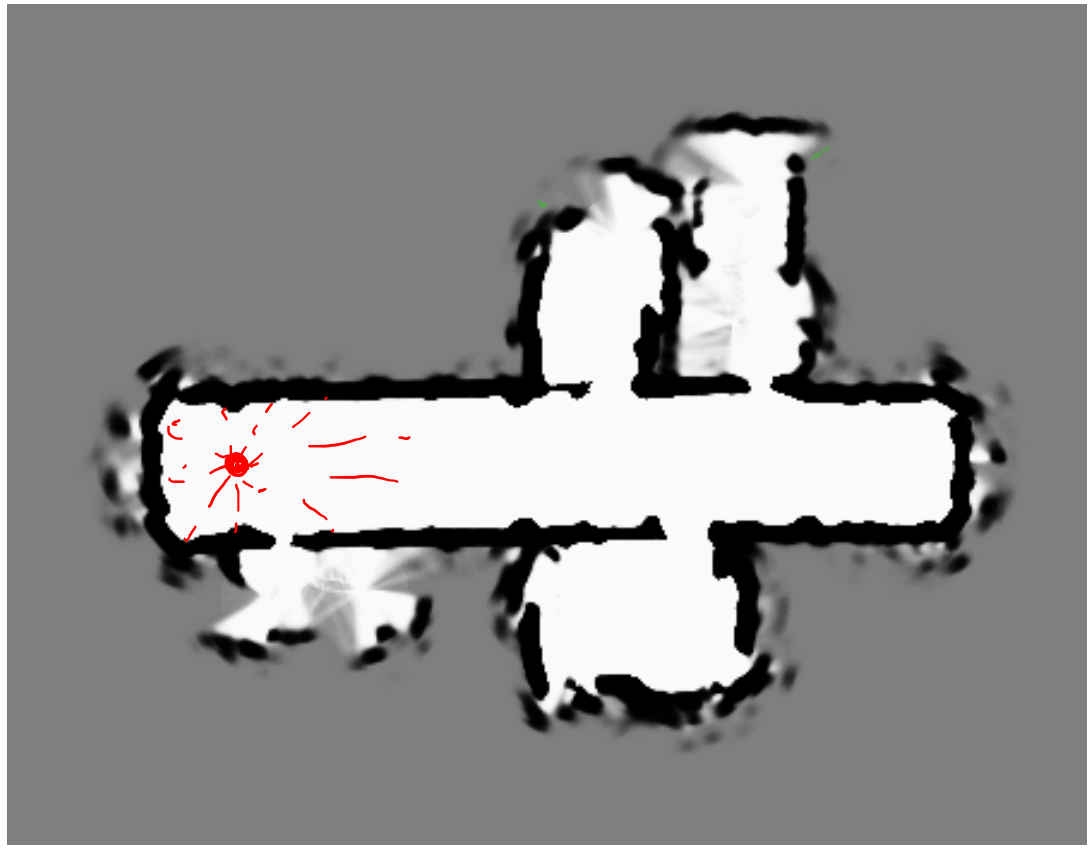
- **Mapping** is one of the fundamental problems in (mobile) robotics
- Maps allow robots to efficiently carry out their tasks and enable **localization**
- Successful robot systems rely on maps for localization, navigation, path planning, activity planning, control, etc.

The General Problem of Mapping



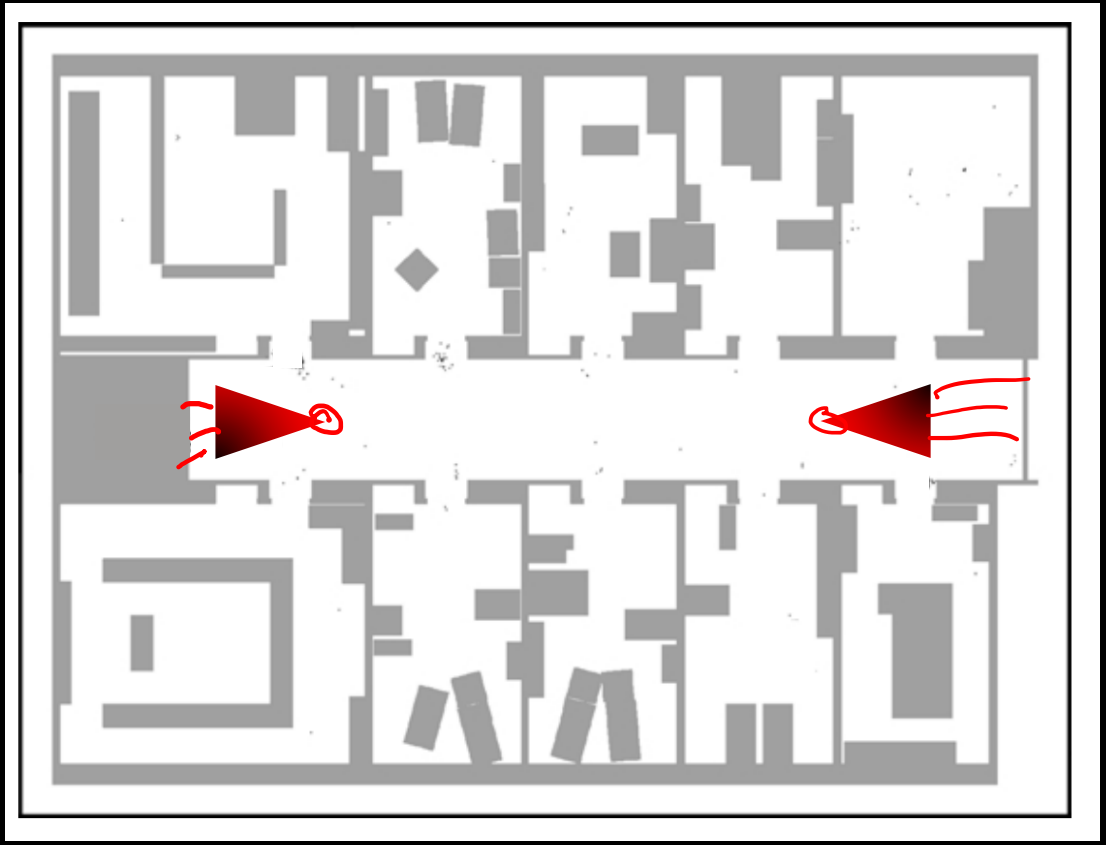
- Sensor interpretation
  - How do we extract relevant information from raw sensor data?
  - How do we represent and integrate this information over time?
- Robot locations must be estimated
  - How can we identify that we are at a previously visited place?
  - This problem is the so-called data association problem

Resulting Map Obtained with Ultrasound Sensors



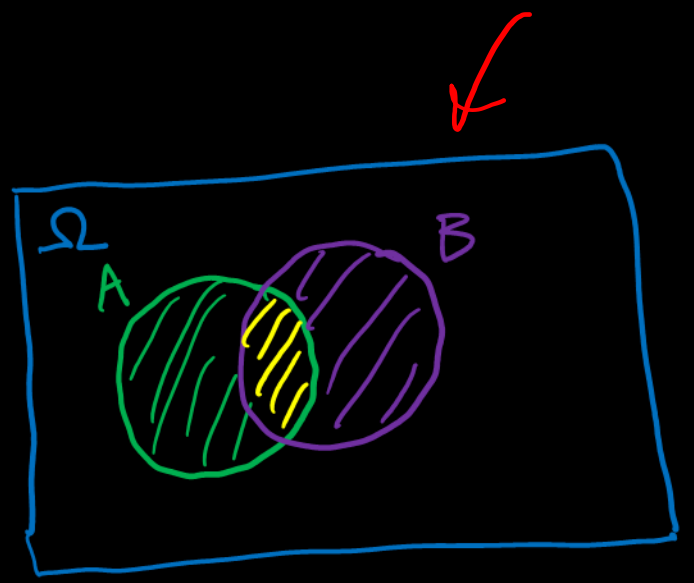
The Localization Problem

- Determine the pose (state) of the robot relative to the **given map** of the environment
- This is also known as **position or state estimation problem**
- Given uncertainty in our measurements and ambiguity from locally symmetric environment, we need to recursively update our estimate or belief



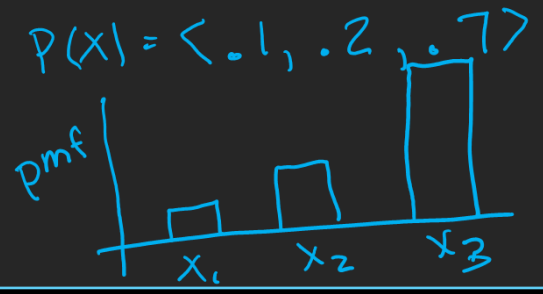
Probability review

- $P(A)$ denotes the probability that event A is true
  - $0 \leq \mathrm{P}(A) \leq 1, P(\mathrm{true}) = 1, P(\mathrm{false}) = 0$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$



Discrete Random Variables

- 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



Continuous Random Variables

- X takes on values in the continuum
- $p(X = x)$, or $p(x)$, is a probability density function
- $P(x \in (a, b)) = \int_a^b p(x)dx$

Joint, Conditional, and Total Probability

- Joint Probability: $P(X = x \text{ and } Y = y) = P(x, y)$
  - If x and y are independent, then $P(x, y) = P(x) \cdot P(y)$
- Conditional Probability: $P(x|y)$
  - $P(x, y) = P(x|y)P(y) \rightarrow P(x|y) = \frac{P(x,y)}{P(y)}$
  - If independent, then $P(x|y) = P(x)$
- Total Probability and Marginals

Discrete case:

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

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

Continuous case:

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

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

Bayes's Formula

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

often write:

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

normalization

posterior $\propto$ likelihood · prior

state to est given sensor/info

likelihood

prior

$$\underbrace{P(x|y)}_{\text{posterior}} = \frac{P(y|x)P(x)}{P(y)}$$

evidence

↳ cond. prob. assigned after update

P(door open | z) is diagnostic ✗

P(z | door open) is causal ✓

Robot's Belief over States

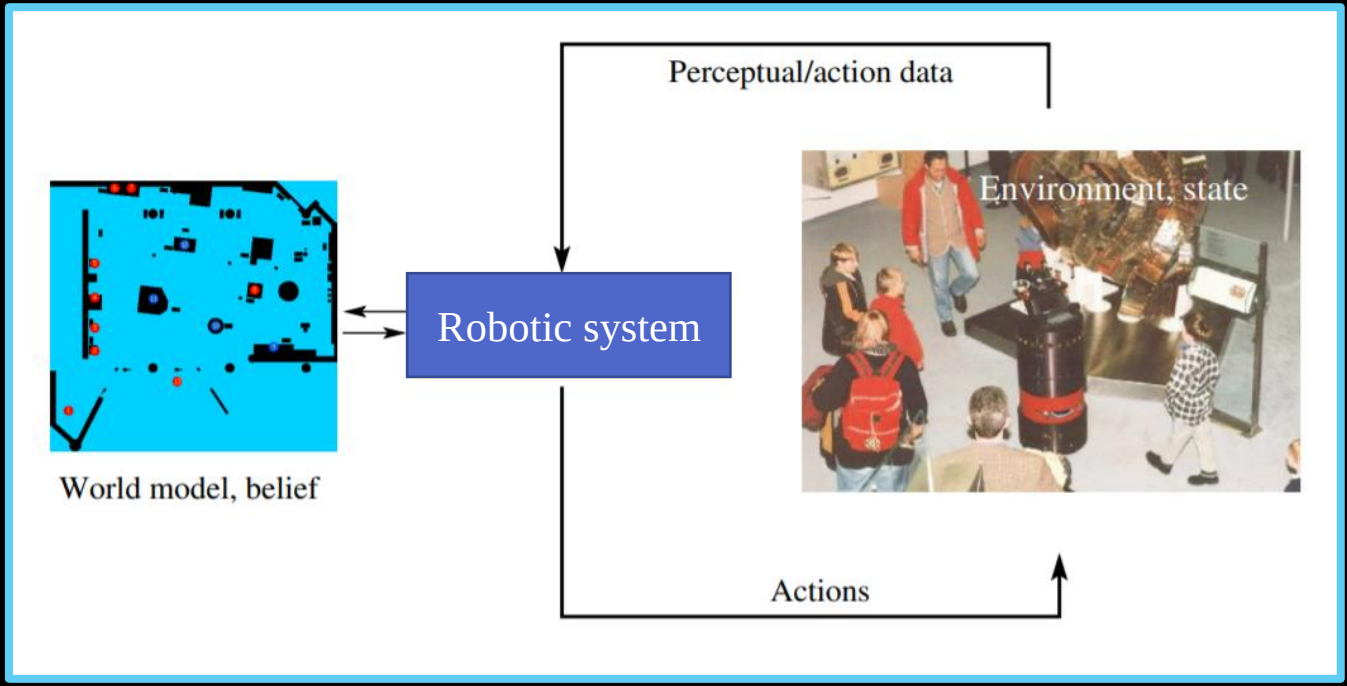
Belief: Robot's knowledge about the state of the environment

$$bel(x_t) = p(x_t | z_{1:t}, u_{1:t})$$

Posterior distribution over state at time *t* given all past measurements and control

Prediction: $\overline{bel}(x_t) = p(x_t | z_{1:t-1}, u_{1:t})$

Calculating $bel(x_t)$ from $\overline{bel}(x_t)$ is called **correction** or **measurement update**



Today's Plan

- What is filtering, mapping, and localization?
  - Probability review!
- Bayes Filters (discrete)
- Kalman Filters (continuous)
- Particle Filters

Notation and Definitions

- Discrete time model

$$x_{t_1:t_2} = x_{t_1}, x_{t_1+1}, x_{t_1+2}, \dots, x_{t_2} \text{ sequence of states } t_1 \text{ to } t_2$$

- Robot takes one measurement at a time

$$z_{t_{1:t_2}} = z_{t_1}, \dots, z_{t_2} \text{ sequence of all measurements from } t_1 \text{ to } t_2$$

- Control also exercised at discrete steps

$$u_{t_1:t_2} = u_{t_1}, u_{t_1+1}, u_{t_1+2}, \dots, u_{t_2} \text{ sequence control inputs}$$

State Evolution / Models

Evolution of the state and measurements are governed by probabilistic laws:

$p(x_t|x_{0:t-1}, z_{1:t-1}, u_{1:t})$ describes state evolution / motion model

If the state is *complete,* we can succinctly state:

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

Measurement process given by:

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

Similarly, if measurement is complete:

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

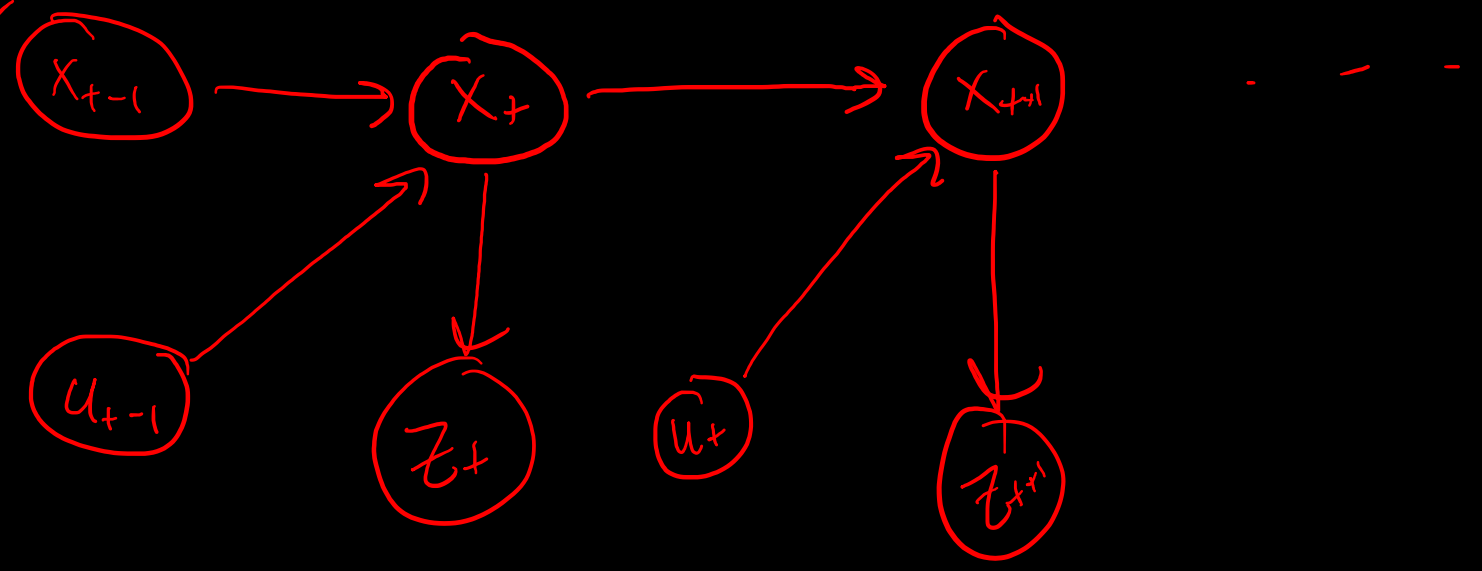
Discrete Bayes Filter Algorithm: Setup

- Evolution of the state is governed by probabilistic state transition:

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

- Measurement process given by:

$$p(z_t | x_t)$$



Robot's Belief over States

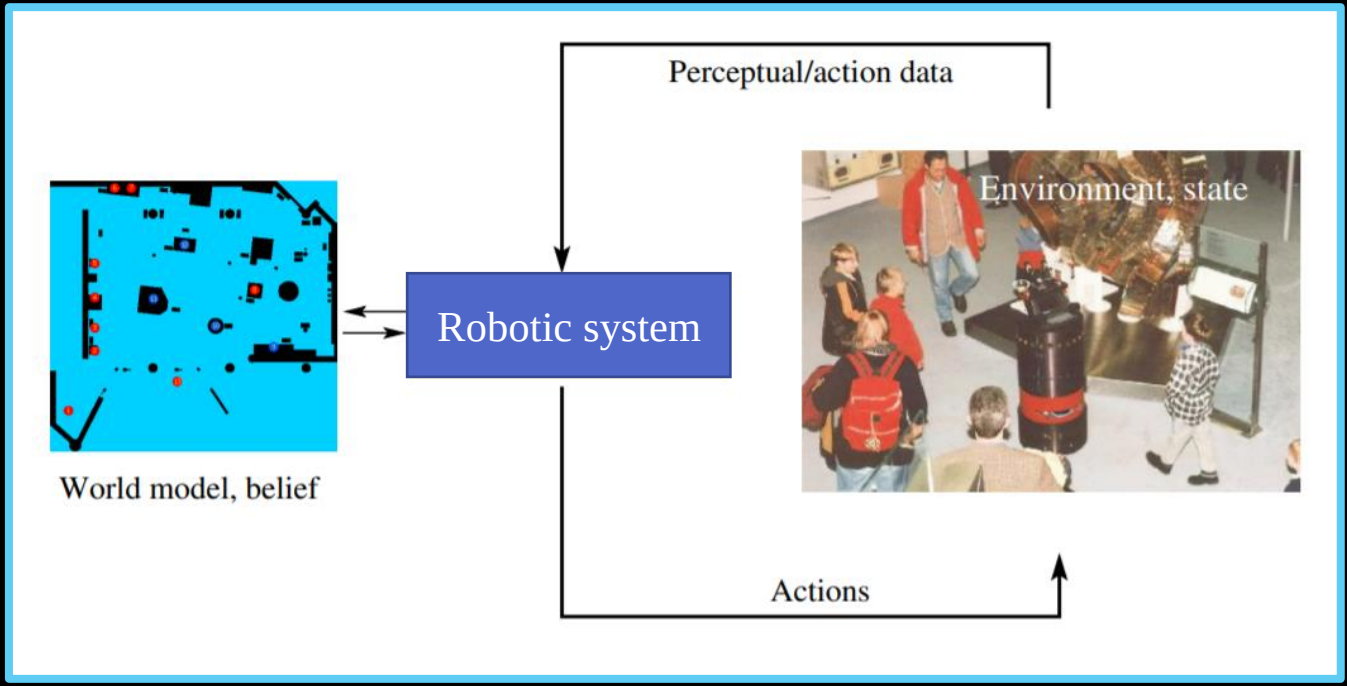
Belief: Robot's knowledge about the state of the environment

$$bel(x_t) = p(x_t | z_{1:t}, u_{1:t})$$

Posterior distribution over state at time *t* given all past measurements and control

Prediction: $\overline{bel}(x_t) = p(x_t | z_{1:t-1}, u_{1:t})$

Calculating $bel(x_t)$ from $\overline{bel}(x_t)$ is called **correction** **or** **measurement update**



Recursive Bayes Filter

Algorithm Bayes_Filter($bel(x_{t-1}), u_t, z_t$**)**

for all x_t do:

$$\overline{bel}(x_t) = \int p(x_t | u_t, x_{t-1}) bel(x_{t-1}) dx_{t-1}$$

$$bel(x_t) = \eta \, p(z_t | x_t) \, \overline{bel}(x_t)$$

end for

return $bel(x_t)$

Recursive Bayes Filter

Algorithm Bayes_Filter($bel(x_{t-1}), u_t, z_t$**)**

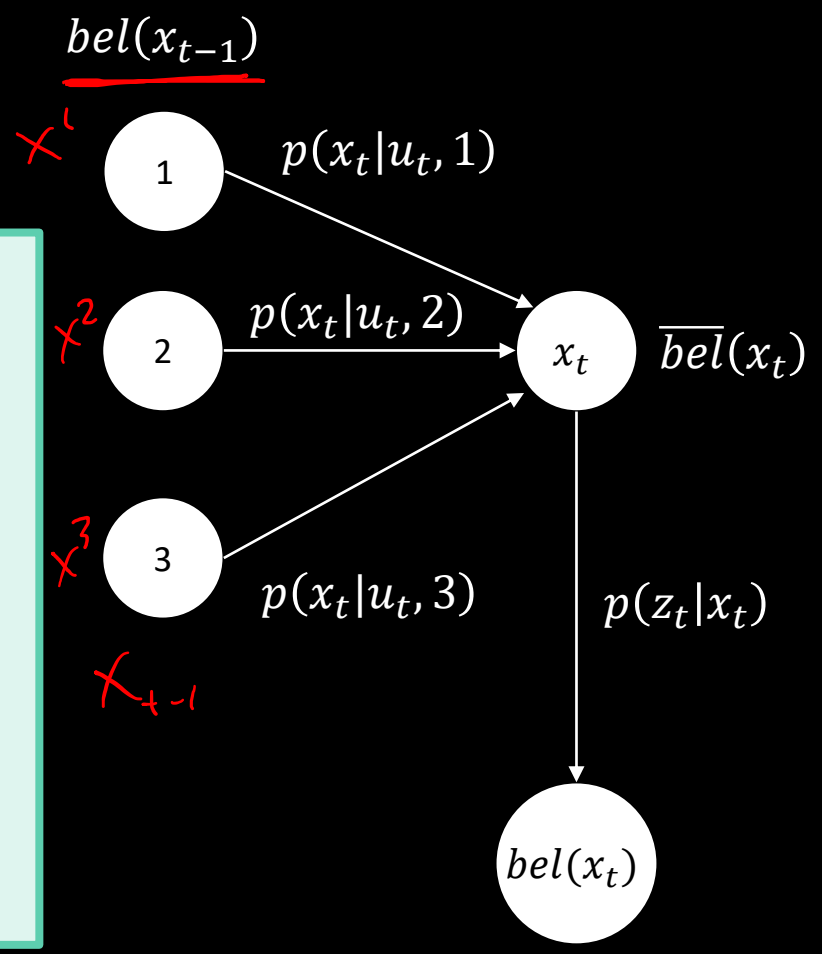
for all x_t do:

$$\overline{bel}(x_t) = \sum_{x_{t-1}} p(x_t|u_t, x_{t-1}) bel(x_{t-1})$$

$$bel(x_t) = \eta\, p(z_t|x_t)\, \overline{bel}(x_t)$$

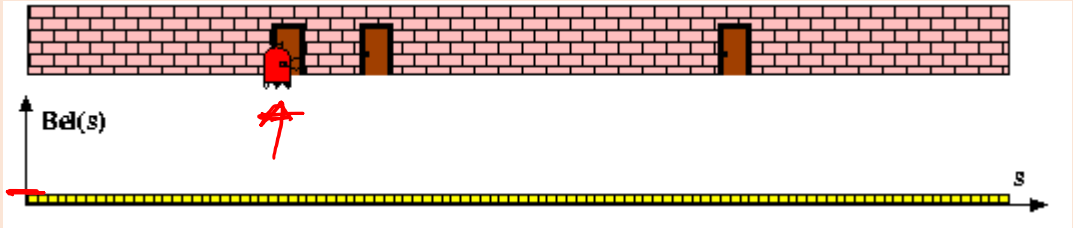
end for

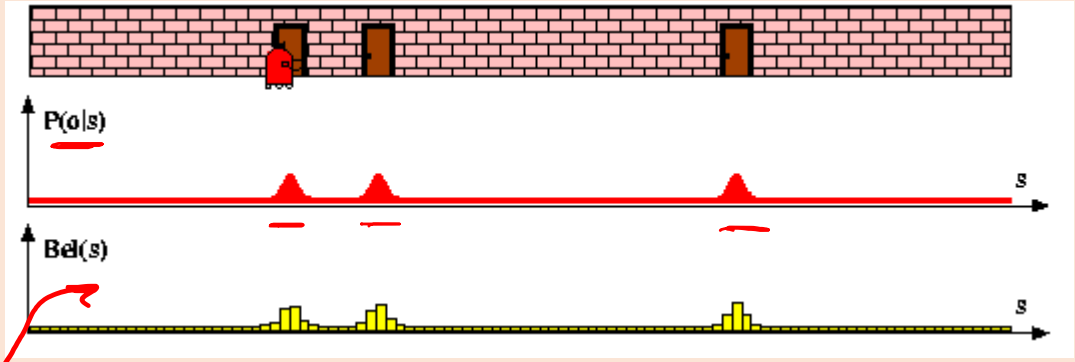
return $bel(x_t)$



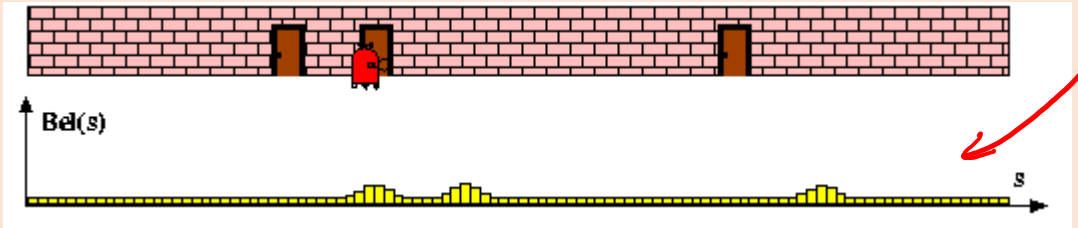
Discrete Bayes Filter - Illustration

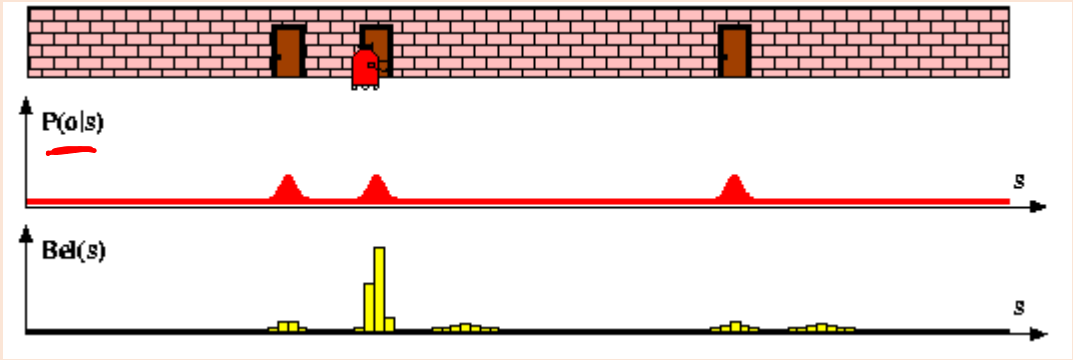
time = 0



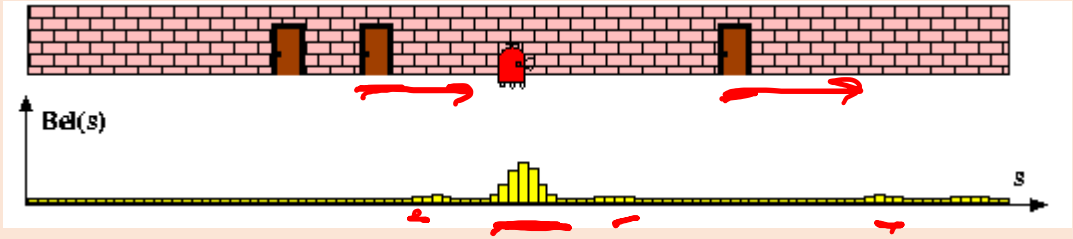


time = 1





time = 2



Today's Plan

- What is filtering, mapping, and localization?
  - Probability review!
- Bayes Filters (discrete)
- Kalman Filters (continuous)
- Particle Filters

Bayes Filter Reminder

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

- Prediction

$$\overline{bel}(x_t) = \int p(x_t|\, u_t, x_{t-1}) bel(x_{t-1}) dx_{t-1}$$

- Correction

$$bel(x_t) = \eta p(z_t|x_t) \overline{bel}(x_t)$$

What if we have a good model of our (continuous) system dynamics and we assume a Gaussian model for our uncertainty?

→ Kalman Filters!

Linear Systems with Gaussian Noise

Suppose we have a system that is governed by a linear difference equation:

$$x_t = A_t x_{t-1} + B_t u_t + \epsilon_t$$

with measurement

$$z_t = C_t x_t + \delta_t$$

Recall Multivariate Gaussians:

$$p(x) \sim N(\mu, \Sigma)$$

$$p(x) = \frac{1}{(2\pi)^{2d}|\Sigma|^{1/2}} e^{-1/2(x-\mu)^\top \Sigma^{-1}(x-\mu)}$$

$$Y = AX + B$$

$$Y \sim N(A\mu + B, A\Sigma A^\top)$$

A_t	a $n \times n$ matrix that describes how the state evolves from $t-1$ to t without controls or noise
B_t	a $n \times m$ matrix that describes how the control u_t changes the state from $t-1$ to t
C_t	a $k \times n$ matrix that describes how to map the state x_t to an observation z_t
ϵ_t & δ_t	random variables representing the process and measurement noise, and are assumed to be independent and normally distributed with zero mean and covariance Q_t and R_t, respectively

What is a Kalman Filter?

Suppose we have a system that is governed by a linear difference equation:

$$x_t = A_t x_{t-1} + B_t u_t + \epsilon_t$$

with measurement

$$z_t = C_t x_t + \delta_t$$

- Tracks the estimated state of the system by the mean and variance of its state variables -- minimum mean-square error estimator
- Computes the ***Kalman gain,*** which is used to weight the impact of new measurements on the system state estimate against the predicted value from the process model
- Note that we no longer have discrete states or measurements!

Summary

- Bayes filters are a probabilistic tool for estimating the state of dynamic systems
  - **They are everywhere!** Kalman filters, Particle filters, Hidden Markov models, Dynamic Bayesian networks, Partially Observable Markov Decision Processes (POMDPs), …
  - Bayes rule allows us to compute probabilities that are hard to assess otherwise
  - Recursive Bayesian updating can efficiently combine evidence
- Kalman filters give you the optimal estimate for linear Gaussian systems

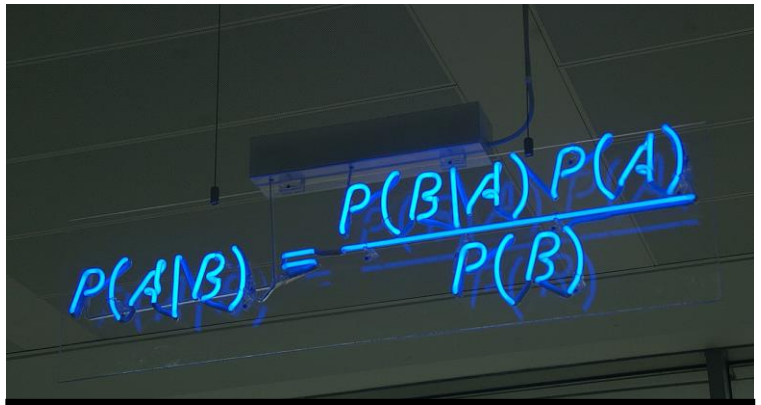
Extra Slides

Fun Fact: Who is Bayes?

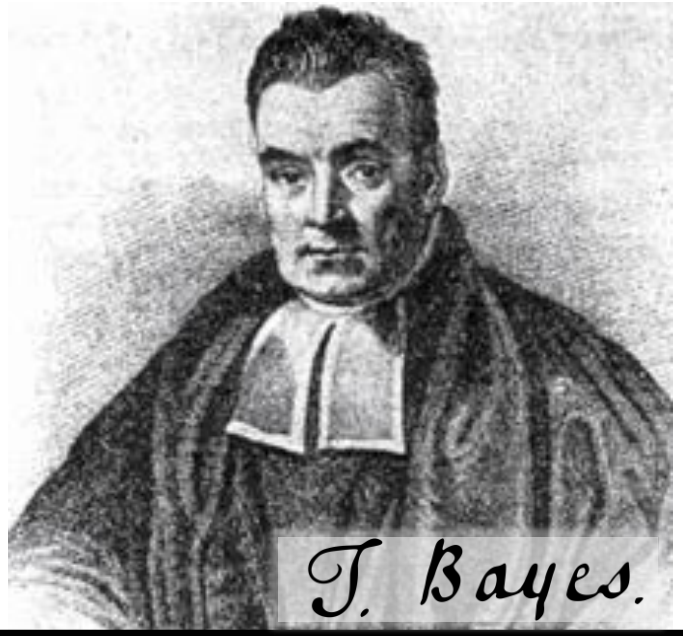
Bayes was an English **statistician**, philosopher, and minister who lived from 1701 to 1761, and is known for two works:

1. Divine Benevolence, or an Attempt to Prove That the Principal End of the Divine Providence and Government is the Happiness of His Creatures (1731)
2. An Introduction to the Doctrine of Fluxions, and a Defence of the Mathematicians Against the Objections of the Author of The Analyst (1736), in which he **defended the logical foundation of Isaac Newton's calculus** ("fluxions") against the criticism of George Berkeley, author of The Analyst

Bayes never published his most famous accomplishment **Bayes' Theorem**. These notes were edited and published after his death by Richard Price.



From the HP Autonomy Lab



Probably not Bayes

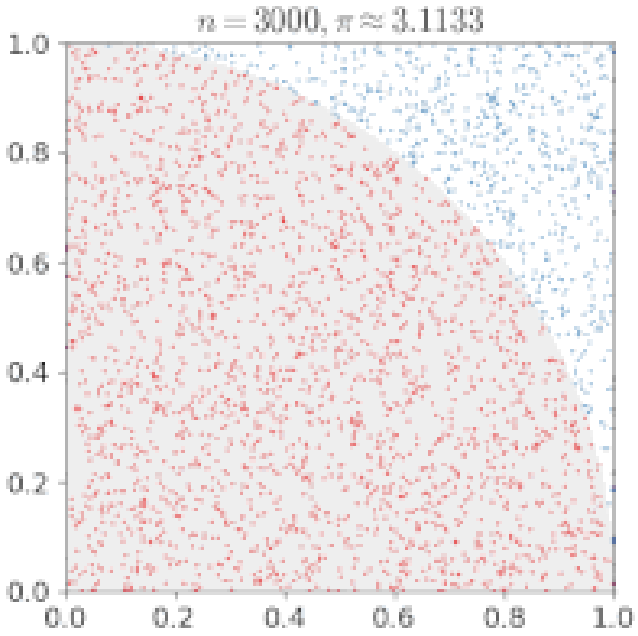
Fun facts about Monte Carlo Methods

… a question which occurred to me in 1946 as I was **convalescing from an illness and playing solitaires. The question was what are the chances that a Canfield solitaire laid out with 52 cards will come out successfully?** After spending a lot of time trying to estimate them by pure combinatorial calculations, I wondered whether a more practical method than "abstract thinking" might not be to **lay it out say one hundred times and simply observe and count the number of successful plays. This was already possible to envisage with the beginning of the new era of fast computers,** and I immediately thought of problems of neutron diffusion and other questions of mathematical physics, and more generally how to change processes described by certain differential equations into an equivalent form interpretable as a succession of random operations. Later [in 1946], I described the idea to John von Neumann, and we began to plan actual calculations. - *Stanislaw Ulam, ~1940s*

Monte Carlo methods vary, but tend to follow this pattern:

1. Define a domain of possible inputs
2. Generate inputs randomly from a probability distribution over the domain
3. Perform a deterministic computation on the inputs
4. Aggregate the results

Monte Carlo methods were central to the **simulations required for the Manhattan Project**, and more recently has been extended to Sequential Monte Carlo in advanced signal processing and Bayesian inference. Also, the extension **MC Tree Search enabled AlphaGo.**

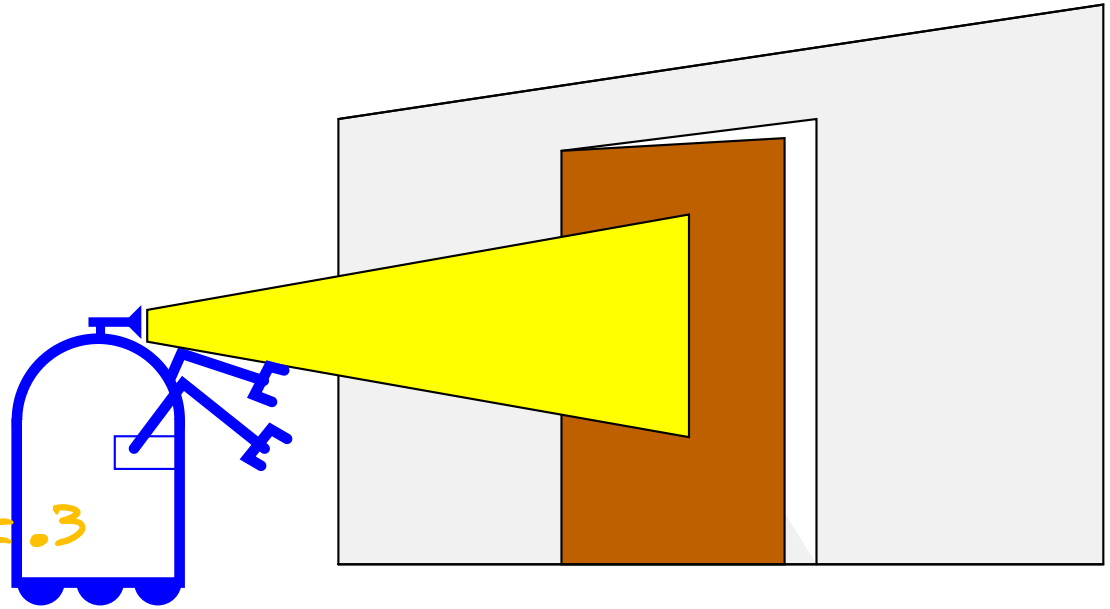


Who was Rudolf Kalman?

Image Credit: Wikipedia

- Kálmán was one of the most influential people on control theory today and is most known for his co-invention of the Kalman filter (or Kalman-Bucy Filter)
- The filtering approach was initially met with vast skepticism, so much so that he was forced to do the first publication of his results in mechanical engineering, rather than in electrical engineering or systems engineering
  - This worked out find as some of the first use cases was with NASA on the Apollo spacecraft
- *Kalman filters are inside every robot, commercial airplanes, uses in seismic data processing, nuclear power plant instrumentation, and demographic models, as well as applications in econometrics*

Door example of Bayes Rule



Suppose a robot obtains measurement z.
What is $P(\text{open}|z)$?

Given: $P(z|\text{open}) = .6$, $P(z|\text{closed}) = .3$

$P(\text{open}) = P(\text{closed}) = .5$

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

$$= \frac{P(z|\text{open})\,P(\text{open})}{P(z|\text{open})P(\text{open}) + P(z|\text{closed})\,P(\text{closed})}$$

$$= \frac{.6 \cdot .5}{.6 \cdot .5 + .3 \cdot .5} = 2/3 = \underline{.67}$$

Bayes Filters (1)

$bel(x_0 = \underbrace{open}_{o}) = .5$

$bel(x_0 = \underbrace{closed}_{c}) = .5$

$p(z_t \mid x_t)$ → open or closed

↳ sense_open (so)

sense_closed (sc)

x_t

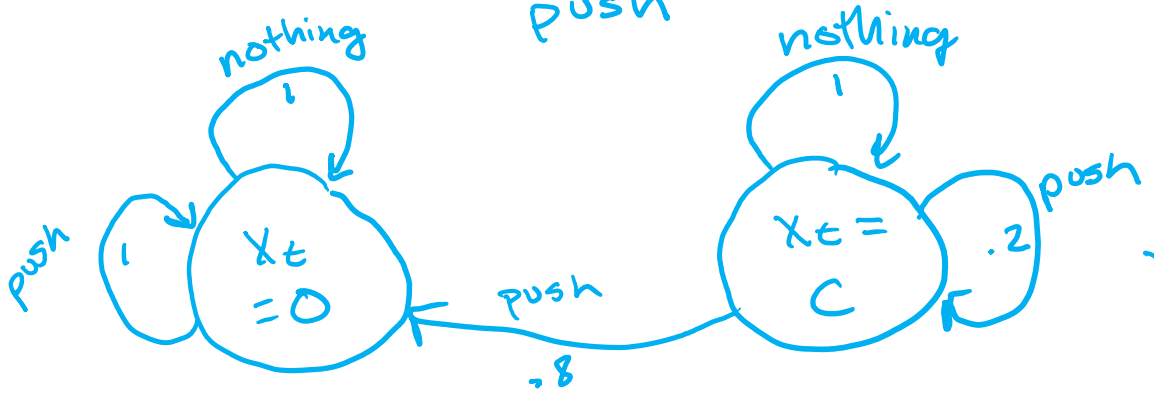
z_t	open	closed
so	.6	.2
sc	.4	.8

→ $p(z_t = so \mid x_1 = c)$

$p(x_t \mid u_t, x_{t-1})$

↳ nothing

push



$u_t = push$

x_{t-1}

x_t	o	c
o	1	.8
c	0	.2

$u_t = nothing$

x

x'	o	c
o	1	0
c	0	1

Bayes Filters (2)

at $t=1$, do nothing : $\rightarrow u_t = n$

$$\overline{bel}(x_1) = \sum_{x_0} p(x_1 \mid u_1, x_0)\, bel(x_0)$$

$$= p(x_1 \mid u_1 = n, x_0 = o)\, bel(x_0 = o) + p(x_1 \mid u_1 = n, x_0 = c)\, bel(x_0 = c)$$

suppose that $x_1 = o$, now plug it in!

$$\overline{bel}(x_1 = o) = p(x_1 = o \mid u_1 = n, x_0 = o)\, bel(x_0 = o) + p(x_1 = o \mid u_1 = n, x_0 = c)\, bel(x_0 = c)$$

$$= 1 \cdot .5 + 0 \cdot .5 = .5$$

Bayes Filters (3)

suppose $x_1 = c$

$$\overline{bel}(x_1 = c) = p(x_1 = c \mid u_1 = n, x_0 = o)\, bel(x_0 = o) + p(x_1 = c \mid u_1 = n, x_0 = c)\, bel(x_0 = c)$$

$$= 0 \cdot .5 + 1 \cdot .5 = .5$$

$$\overline{bel}(x_1 = o) = .5$$

$$\overline{bel}(x_1 = c) = .5$$

Bayes Filters (4)

recall: $bel(x_1) = \eta\, p(z_1 \mid x_1)\, \overline{bel}(x_1)$

suppose z_1 = sense_open (so)

1) $x_1 = 0$

$$bel(x_1=0) = \eta\, p(z_1 = so \mid x_1 = 0) \cdot \overline{bel}(x_1 = 0)$$
$$= \eta \cdot .6 \cdot .5 = \eta \cdot .3 = .75$$

2) $x_1 = c$

$$bel(x_1 = c) = \eta\, p(z_1 = so \mid x_1 = c)\, \overline{bel}(x_1 = c)$$
$$= \eta \cdot .2 \cdot .5 = \eta \cdot .1 = .25$$

normalizer: $\eta = \dfrac{1}{.3 + .1} = 2.5$

Bayes Filters (5)

suppose we push door + get new measurement

$u_2 = \text{push}$ (p) $\quad z_2 = \text{sense\_open}$ (so)

$\overline{bel}(x_2 = o) = 1 \cdot .75 + .8 \cdot .25 = .95$

$\overline{bel}(x_2 = c) = 0 \cdot .75 + .2 \cdot .25 = .05$

$bel(x_2 = o) = \eta \cdot .6 \cdot .95 \approx .983$

$bel(x_2 = c) = \eta \cdot .2 \cdot .05 \approx .017$