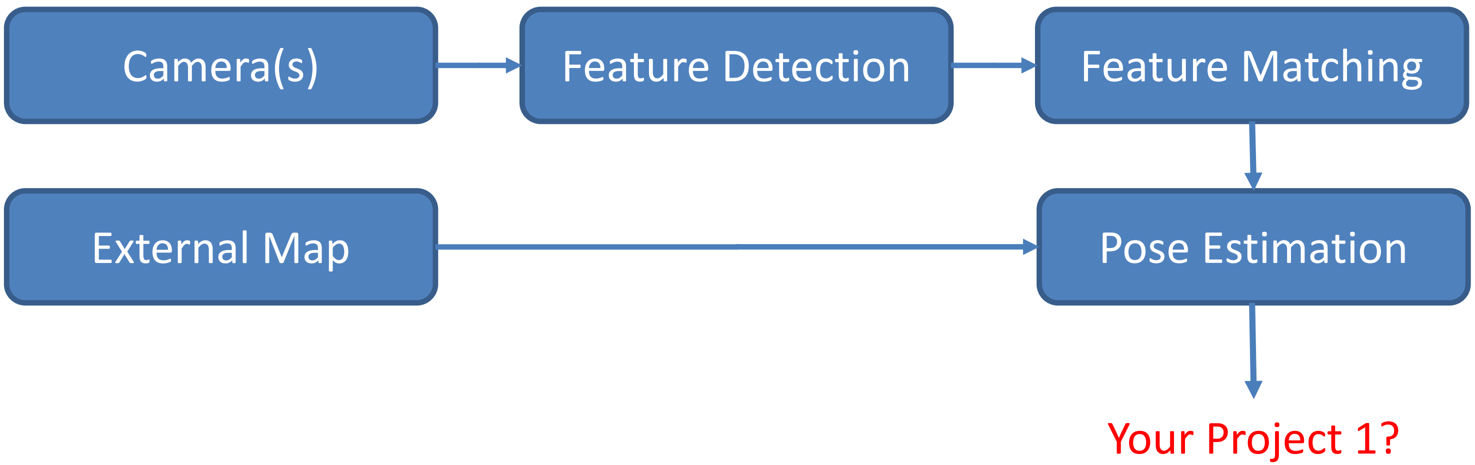
香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY
電子及計算機工程學系
DEPARTMENT OF ELECTRONIC & COMPUTER ENGINEERING

Introduction to Aerial Robotics
Lecture 8

Shaojie Shen
Associate Professor
Dept. of ECE, HKUST

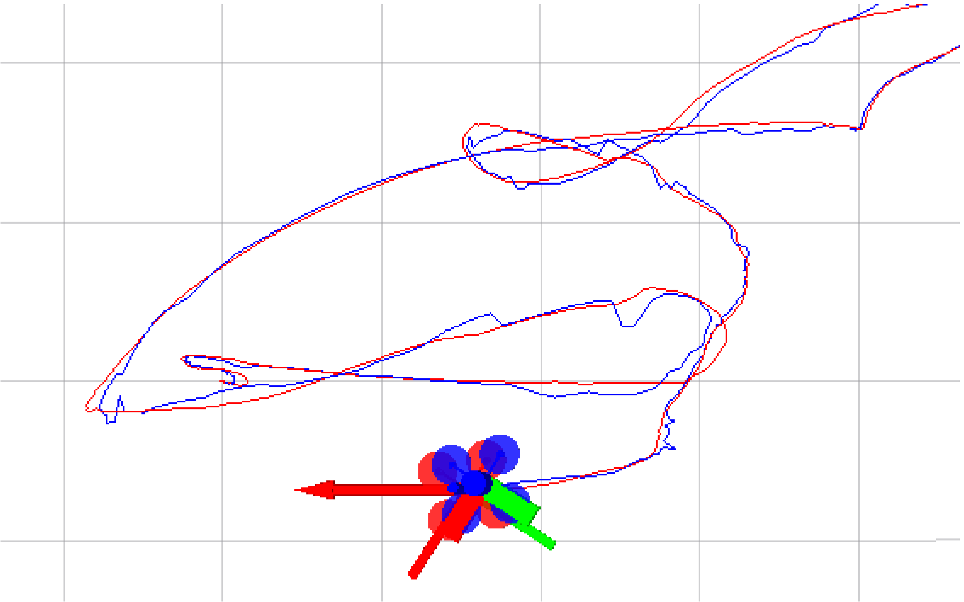
30 March 2021

Vision-based Navigation Pipeline



Why Sensor Fusion?

- Vision/GPS-only state estimation is too noisy, slow, and delayed for feedback control of agile aerial robots
- To improve robustness with multiple sensors and handle sensor failures
- To estimate quantities that are unobservable using single sensors



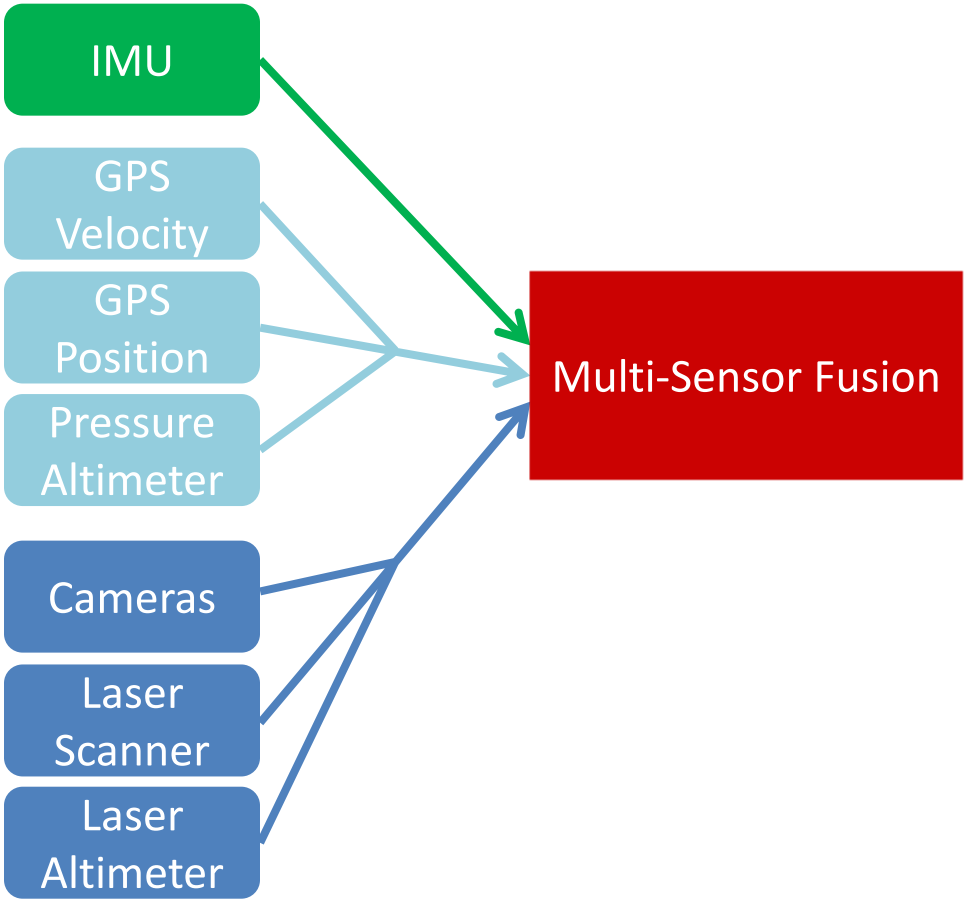
Red: Vision+IMU Fusion
Blue: Vision-only

Design Considerations…

- Accuracy
- Frequency
- Latency
- Sensor synchronization & timestamp accuracy
- Delayed and out-of-order measurements
- Estimator initialization
- Sensor calibration
- Different measurement models with uncertainties
- Robustness to outliers
- Computational efficiency

What to Fuse?

- IMU centric fusion
  - High frequency
  - Low latency
  - (Almost) always available
  - (Usually) large drift
- Absolute measurements
- Relative measurements



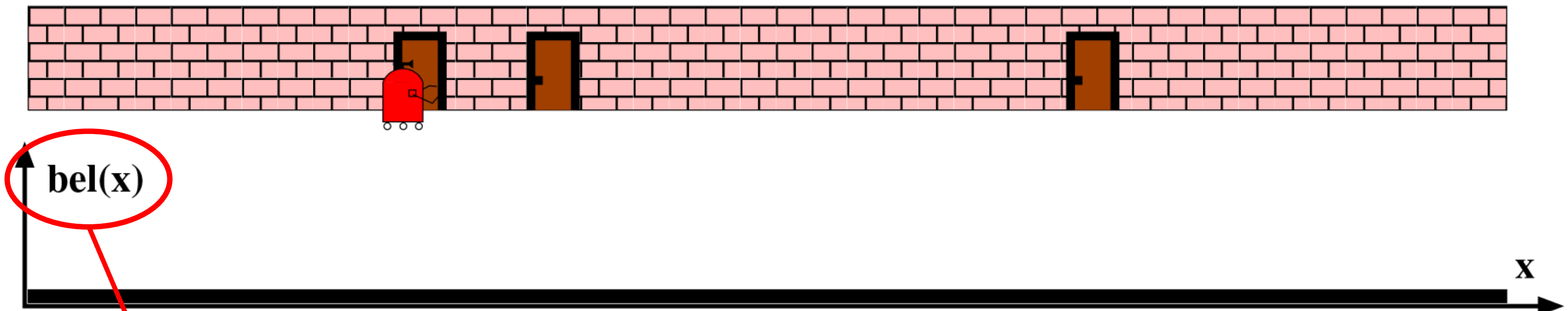
Outline

- Bayesian Filtering
  - Introduction to Probability
  - Bayes' Filter
- Kalman Filtering
  - Gaussian Random Variables
  - The Kalman Filter
  - Continuous Time Systems

Bayesian Filtering

Example Problem

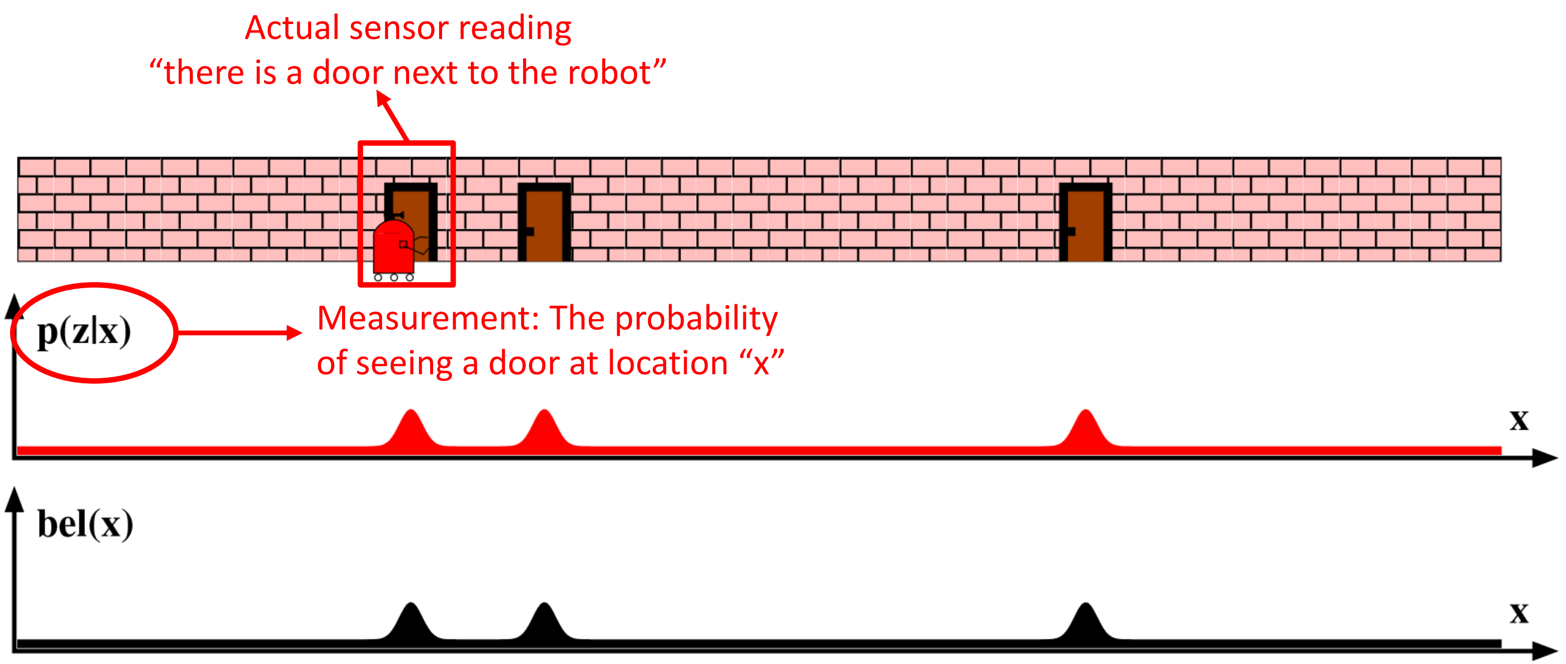
Known map



Use “belief” (or probability distribution) to represent robot state (pose)

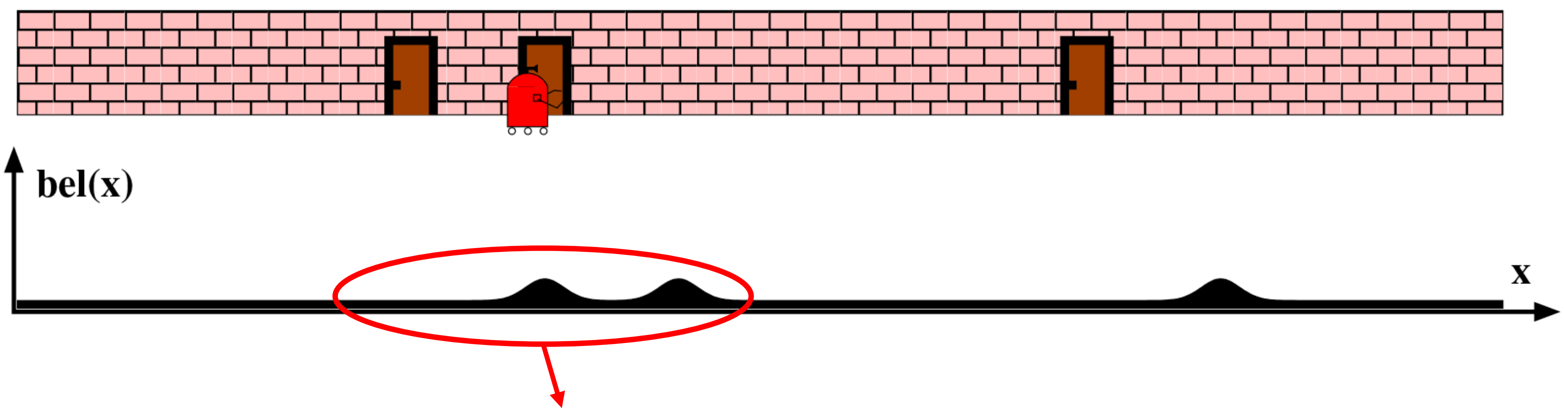
[Thrun, Burgard, and Fox. *Probabilistic Robotics*. 2005]

Example Problem



[Thrun, Burgard, and Fox. *Probabilistic Robotics*. 2005]

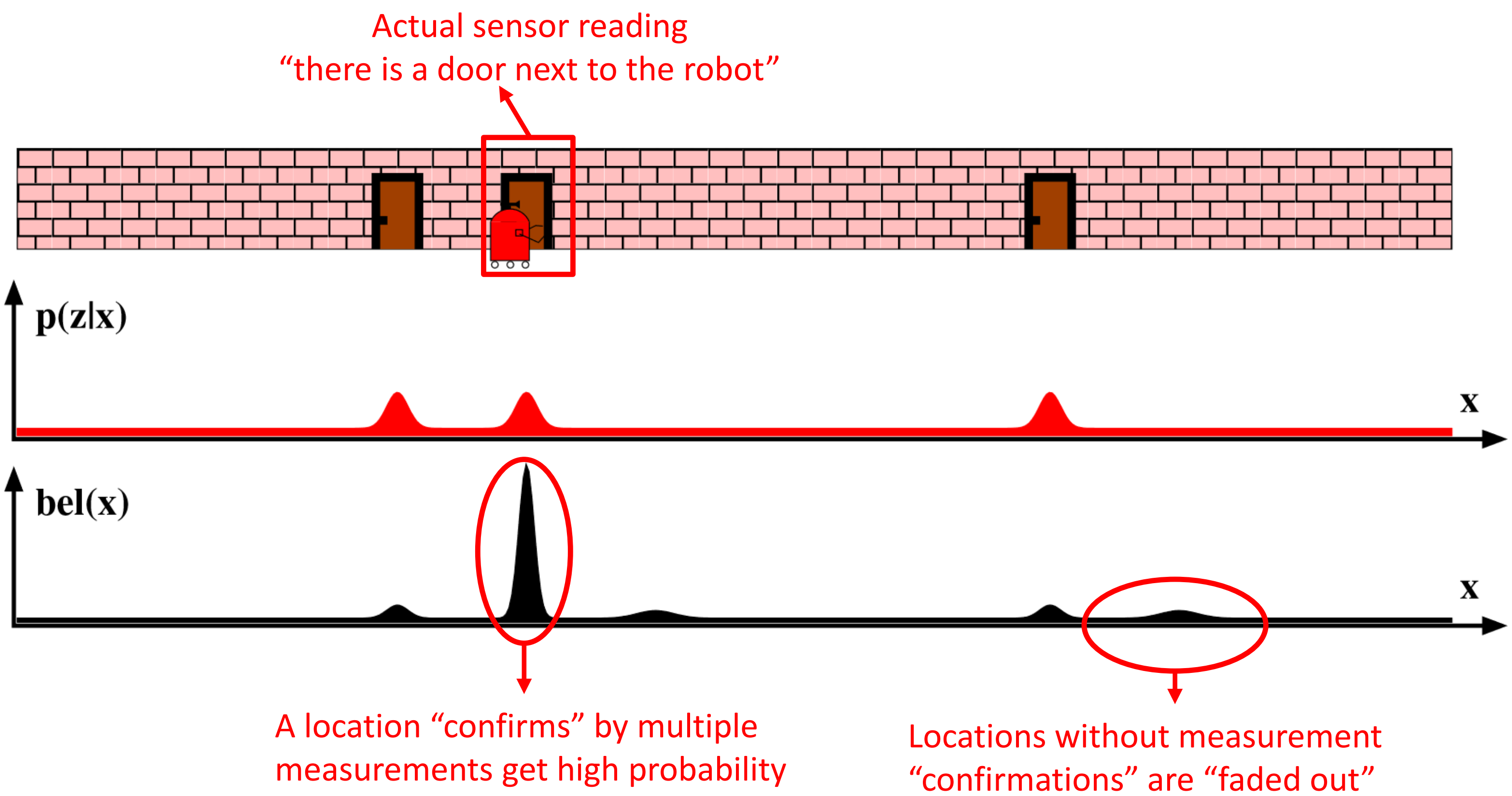
Example Problem



Robot movement "shifts" and "flatten" the belief

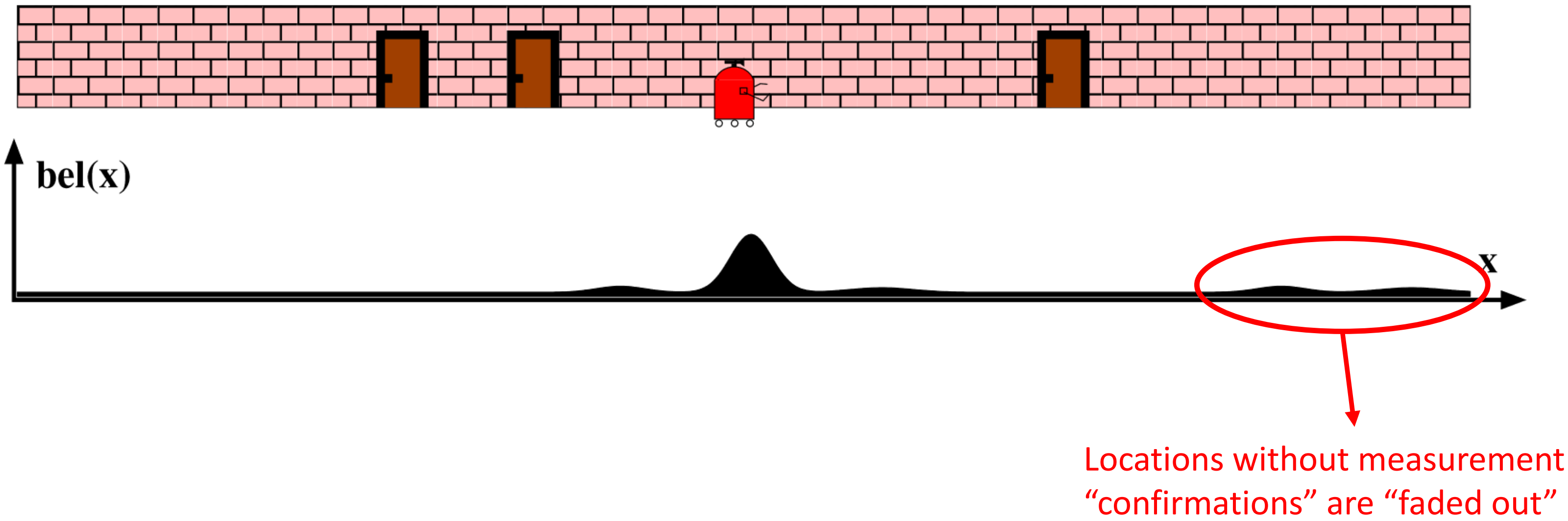
[Thrun, Burgard, and Fox. *Probabilistic Robotics*. 2005]

Example Problem



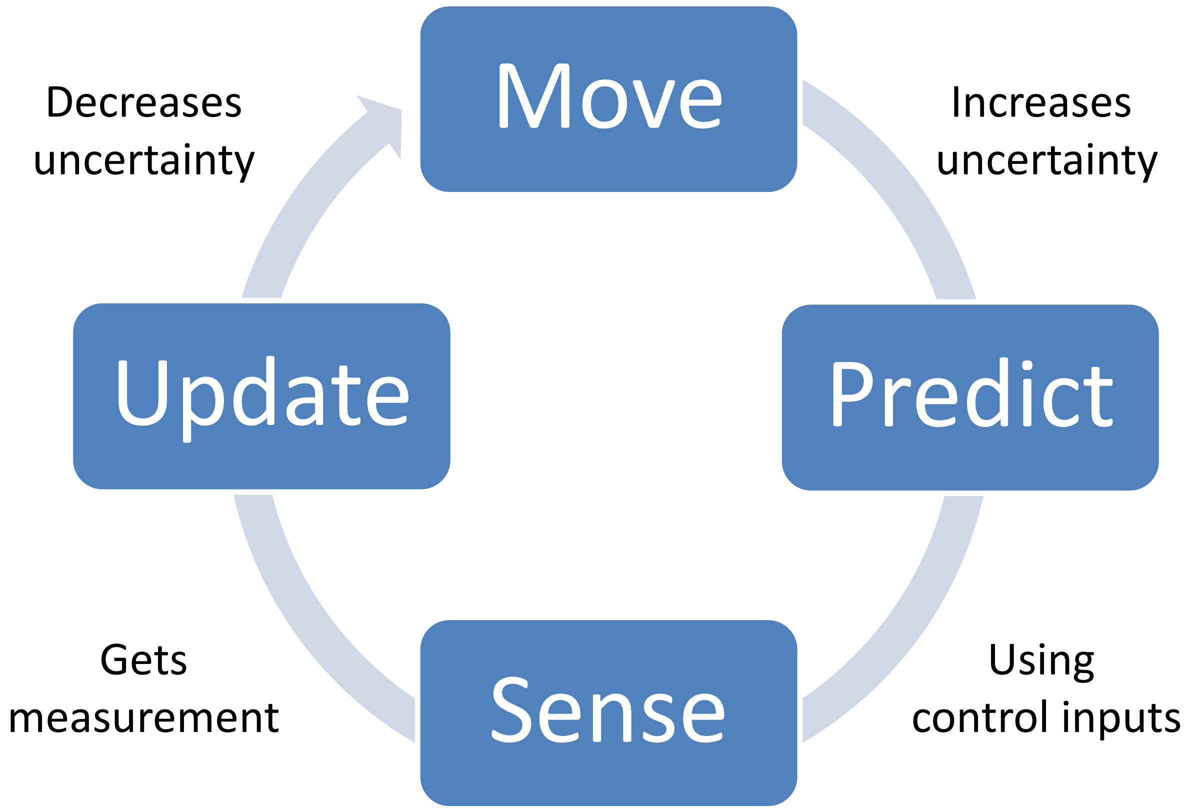
[Thrun, Burgard, and Fox. *Probabilistic Robotics*. 2005]

Example Problem



[Thrun, Burgard, and Fox. *Probabilistic Robotics*. 2005]

Problem Overview



Questions

- What are sources of uncertainty?
- How do we mathematically represent uncertainty in the system?
- How do we use collected evidence to update our belief?
- What can we observe?
- What can we not observe?

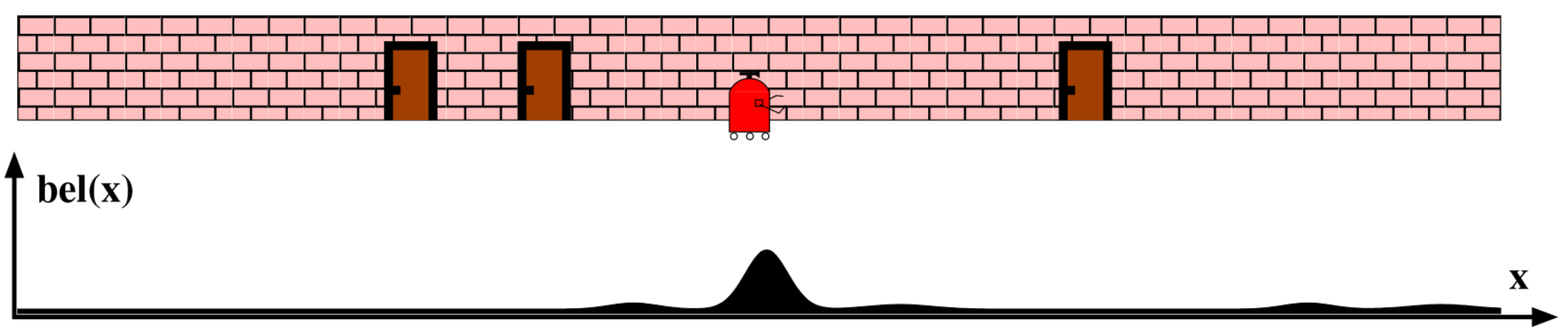
Introduction to Probability

Random Variables

- **Definition:** Variable whose value is subject to change due to randomness or chance
- **Properties:**
  - Can be continuous (e.g., position in 3D) or discrete (e.g., roll of a die)
  - Observed values of random variables are called ***realizations***
- **Example:**
  - Pose of a robot, $p(X = x)$, or value of a rolled die, $p(D = d)$

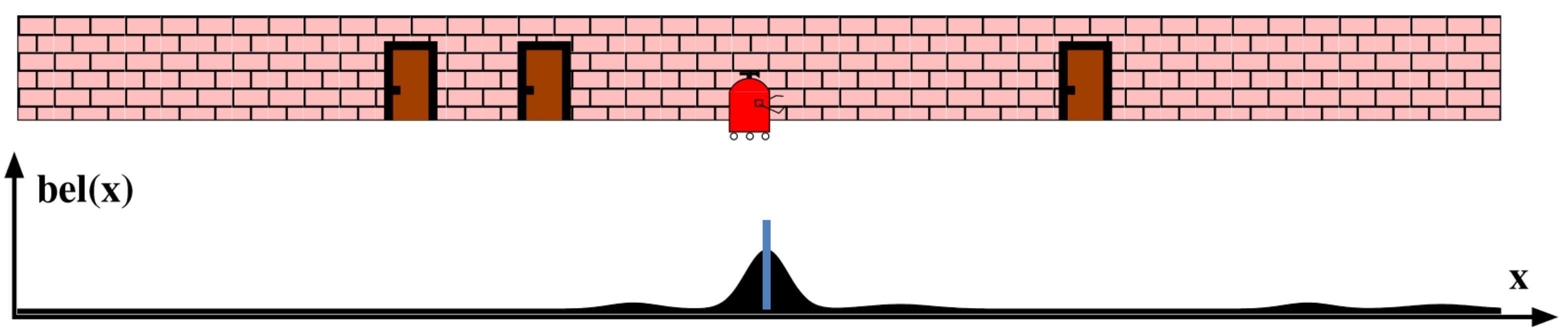
Probability Density Function

- **Definition:** Function describing the likelihood that a random variable X will take on a particular value x
- **Properties:**
  - Total probability is 1, $\int p(X = x)dx = 1$, $\sum_x p(X = x) = 1$
  - Non-negative, $p(X = x) \geq 0$



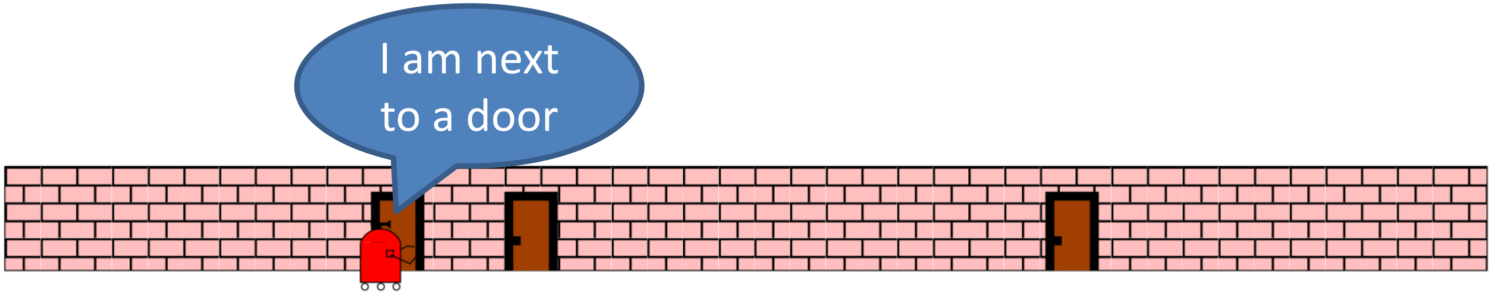
Expected Value

- **Definition:** Probability-weighted average value
  - $E[X] = \int p(X = x)\, x\, dx$
- **Intuition:** "Center of mass" of the probability distribution



Joint Probability Distribution

- **Definition:** The probability density function of a set of two or more random variables
- Also called a ***multivariate distribution***
- **Example:**
  - $p(X = x, Z = z) =$ a robot having a pose x and receiving a measurement z



Covariance

- **Definition:** A measure of how two random variables change together
  - $\sigma(X,Y) = E[(X - E[X])(Y - E[Y])]$
- The ***variance*** is a special case where the two random variables are identical
  - $\sigma^2(X) = \sigma(X,X)$
- **Intuition:** The "moment of inertia" of the probability distribution

Covariance Matrix

- For a multivariate distribution over $\boldsymbol{X} = [X_1, X_2, \dots, X_n]^T$ we define the ***covariance matrix*** to be
- $\boldsymbol{\Sigma} = E[(\boldsymbol{X} - E[\boldsymbol{X}])(\boldsymbol{X} - E[\boldsymbol{X}])^T]$

$$= \begin{bmatrix} \sigma^2(X_1) & \sigma(X_1, X_2) & \cdots & \sigma(X_1, X_n) \\ \sigma(X_2, X_1) & \sigma^2(X_2) & \cdots & \sigma(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ \sigma(X_n, X_1) & \sigma(X_n, X_2) & \cdots & \sigma^2(X_n) \end{bmatrix}$$

- The covariance matrix is symmetric and positive semi-definite

Independence

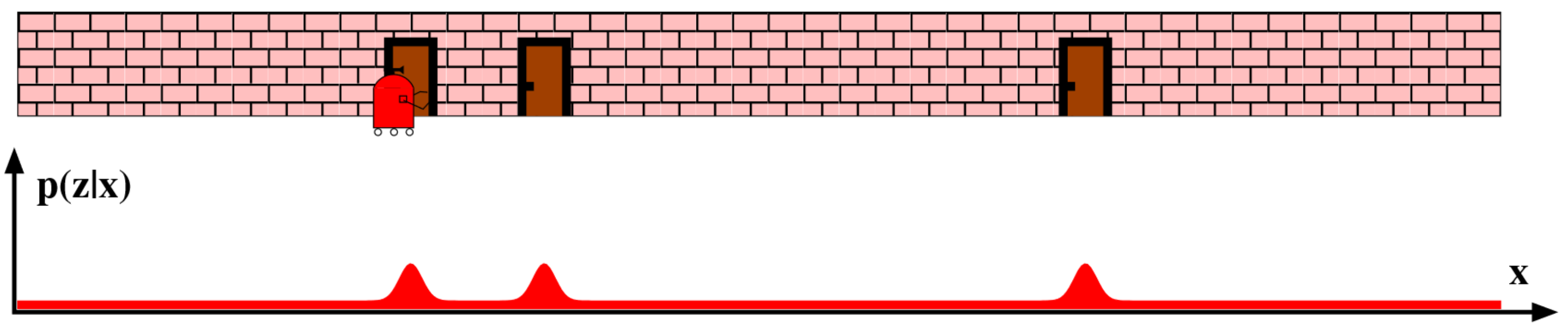
- **Definition:** Two random variables are independent if the outcome of one has *no effect* on the outcome of the other
- $p(x, z) = p(x)\, p(z)$
- **Example:**
  - If X, Z are the outcomes of two dice rolls
- **Properties:**
  - Independent random variables are ***uncorrelated***, $\sigma(X, Z) = 0$
  - Uncorrelated random variables are ***not*** necessarily independent

Example:

- X=U[-1,1] (uniform distribution between -1 and 1)
- $Y = X^2$
- X and Y are uncorrelated but clearly dependent

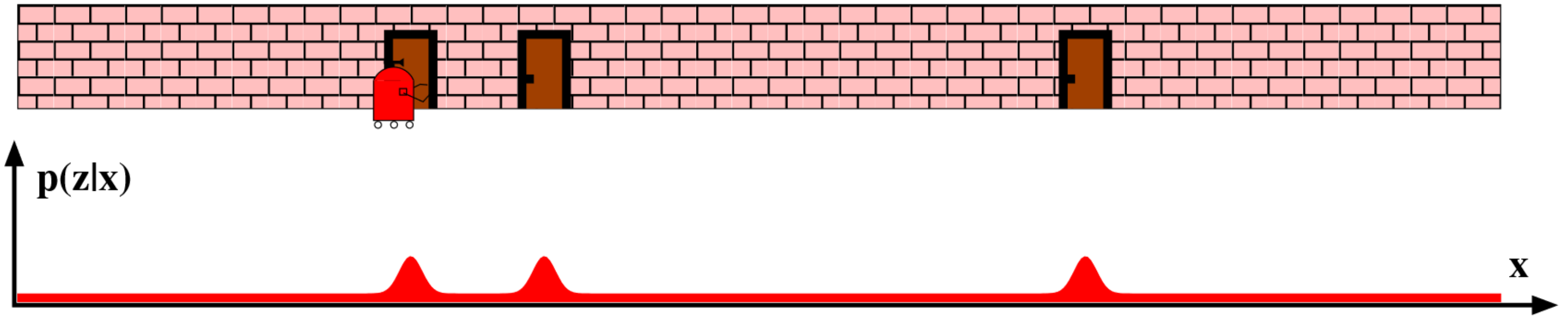
Conditional Probability

- **Definition:** Probability of an event z occurring conditioned on another event x occurring
- $p(z \mid x) = \frac{p(x,z)}{p(x)} \quad \Leftrightarrow \quad p(x,z) = p(z \mid x)\, p(x)$
- **Example:**
  - z = {there is a door next to the robot}



Conditional Independence

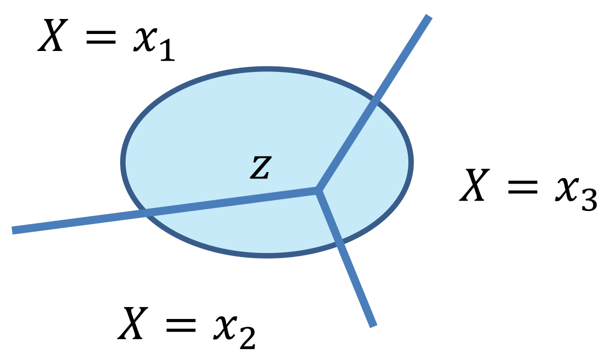
- **Definition:** Two random variables are ***conditionally independent*** if the outcome of one has ***no effect*** on the outcome of the other when conditioned on the outcome of a third random variable
- $p(z_1, z_2 \mid x) = p(z_1 \mid x)\, p(z_2 \mid x)$
- **Example:**
  - Let Z_1, Z_2 be two measurements taken from the same place
  - Z_1, Z_2 are conditionally independent given X



- **Question:** Are Z_1, Z_2 are independent?

Marginal Distribution

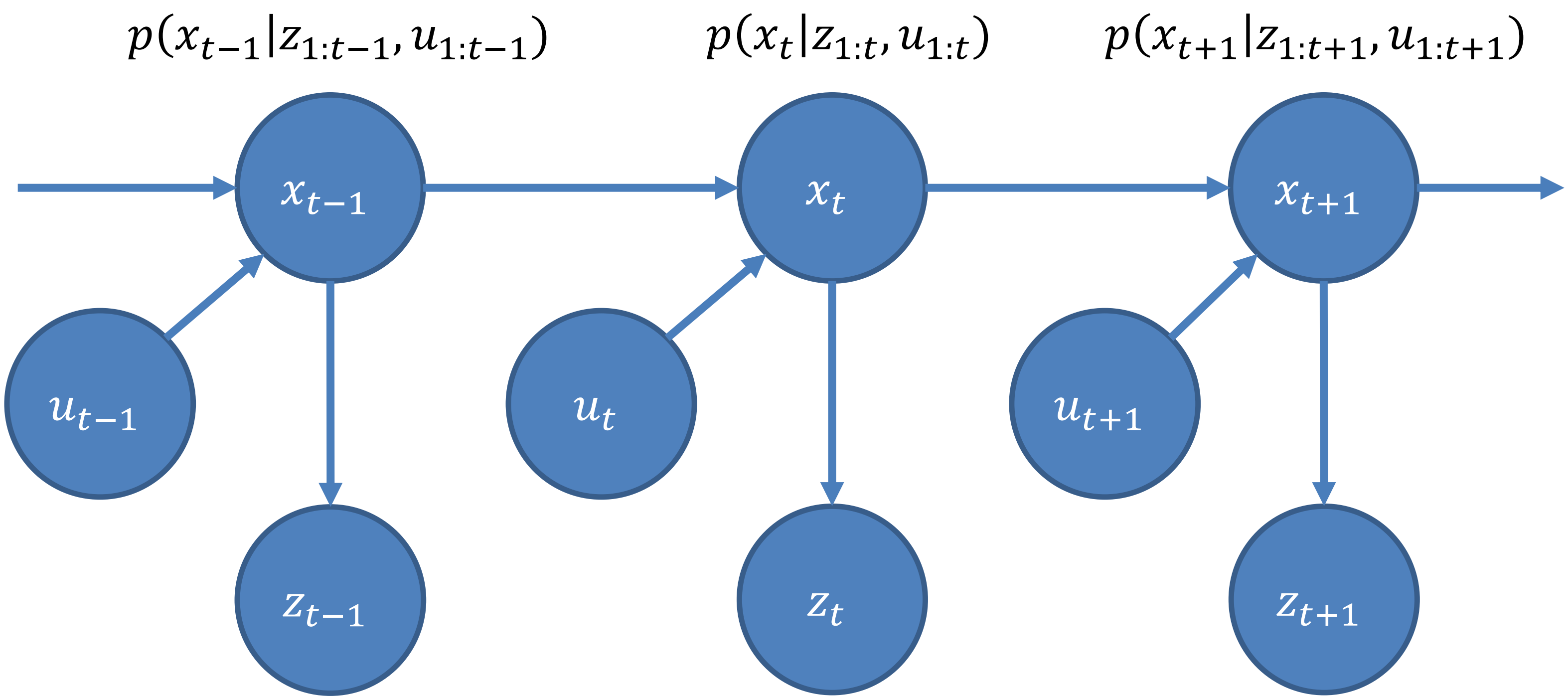
- **Definition:** The probability distribution of the subset of a collection of random variables
- $p(z) = \int p(x, z)\, dx$
- Also known as the ***Law of Total Probability***



$$p(z) = \sum_{i=1}^{3} p(z \mid X = x_i)\, p(X = x_i)$$

Bayes' Filter

Bayesian Filter



x: state
u: control signal
z: measurement

Bayes' Theorem

- $p(x \mid z) = \frac{p(z \mid x)\, p(x)}{p(z)} = \frac{p(z \mid x)\, p(x)}{\int p(z \mid x')\, p(x') dx'}$

- **Intuition:** Describes how the belief about a random variable X should change to account for the collected evidence (measurement) z

- **Derivation:**
  - $p(x, z) = p(z \mid x)\, p(x) = p(x \mid z)\, p(z)$

Markov Property

- **Definition:** The future state of the system is conditionally independent of the past states given the current state
  - $p(x_{t+1} | x_{0:t}) = p(x_{t+1} | x_t)$
  - $p(z_t \mid x_t, z_{1:t-1}, u_{1:t}) = p(z_t \mid x_t)$
  - $p(x_t \mid x_{t-1}, z_{1:t-1}, u_{1:t}) = p(x_t \mid x_{t-1}, u_t)$
- **Question:**
  - Which of the following satisfy the Markov assumption?
    - A first order system with $x = [\text{position}], u = [\text{velocity}]$
    - A second order system with $x = [\text{position}], u = [\text{acceleration}]$
      - How about with $x = [\text{position, velocity}], u = [\text{acceleration}]$

Bayes' Filter Derivation

- **Goal:** Want to update the probability distribution of the robot pose using the realizations of the control input and measurement
- $p(x_t \mid z_{1:t}, u_{1:t}) = \dfrac{p(z_t \mid x_t, z_{1:t-1}, u_{1:t})\, p(x_t \mid z_{1:t-1}, u_{1:t})}{p(z_t \mid z_{1:t-1}, u_{1:t})}$
- **Note:** The measurement is *conditionally independent* of the past measurements and control inputs given the current state of the robot
- $p(z_t \mid x_t, z_{1:t-1}, u_{1:t}) = p(z_t \mid x_t)$
- **Note:** The denominator can be found as a *marginal distribution* of the numerator
- $p(z_t \mid z_{1:t-1},\ u_{1:t}) = \int p(x_t, z_t \mid z_{1:t-1}, u_{1:t})\, dx_t$

Process Model

- Also known as the ***transition model*** or ***motion model***
- $p(x_t \mid z_{1:t-1}, u_{1:t})$
- **Note:** Can find the current pose via marginalization
- $p(x_t \mid z_{1:t-1}, u_{1:t}) = \int p(x_t, x_{t-1} \mid z_{1:t-1}, u_{1:t})\, dx_{t-1}$
- $\quad = \int p(x_t \mid x_{t-1}, z_{1:t-1}, u_{1:t})\, p(x_{t-1} \mid z_{1:t-1}, u_{1:t})\, dx_{t-1}$
- **Note:** The future state is *conditionally independent* of the past measurements and control inputs given the current state and input
- $\underbrace{p(x_t \mid z_{1:t-1}, u_{1:t})}_{\text{Prediction}} = \int \underbrace{p(x_t \mid x_{t-1}, u_t)}_{\text{Process model}}\, \underbrace{p(x_{t-1} \mid z_{1:t-1}, u_{1:t-1})}_{\text{Prior}}\, dx_{t-1}$

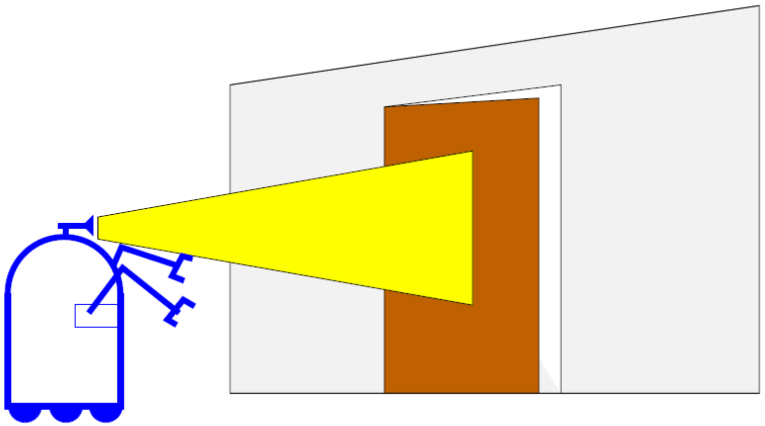
Bayes' Filter

- **Prior:** $p(x_0)$
- **Process model:** $f(x_t \mid x_{t-1}, u_t)$
- **Measurement model:** $g(z_t \mid x_t)$

- **Prediction step:**
- $p(x_t \mid z_{1:t-1}, u_{1:t}) = \int f(x_t \mid x_{t-1}, u_t)\, p(x_{t-1} \mid z_{1:t-1}, u_{1:t-1})\, dx_{t-1}$
- **Update step:**
- $p(x_t \mid z_{1:t}, u_{1:t}) = \dfrac{g(z_t \mid x_t)\, p(x_t \mid z_{1:t-1}, u_{1:t})}{\int g(z_t \mid x_t')\, p(x_t' \mid z_{1:t-1}, u_{1:t})\, dx_t'}$

Measurements

- Robots collect noisy information using sensors
  - Exteroceptive
    - Laser scanner
    - 3D depth sensor
    - Magnetometer
    - *Camera*
  - Proprioceptive
    - Motor encoder
    - *Gyroscope*
    - *Accelerometer*



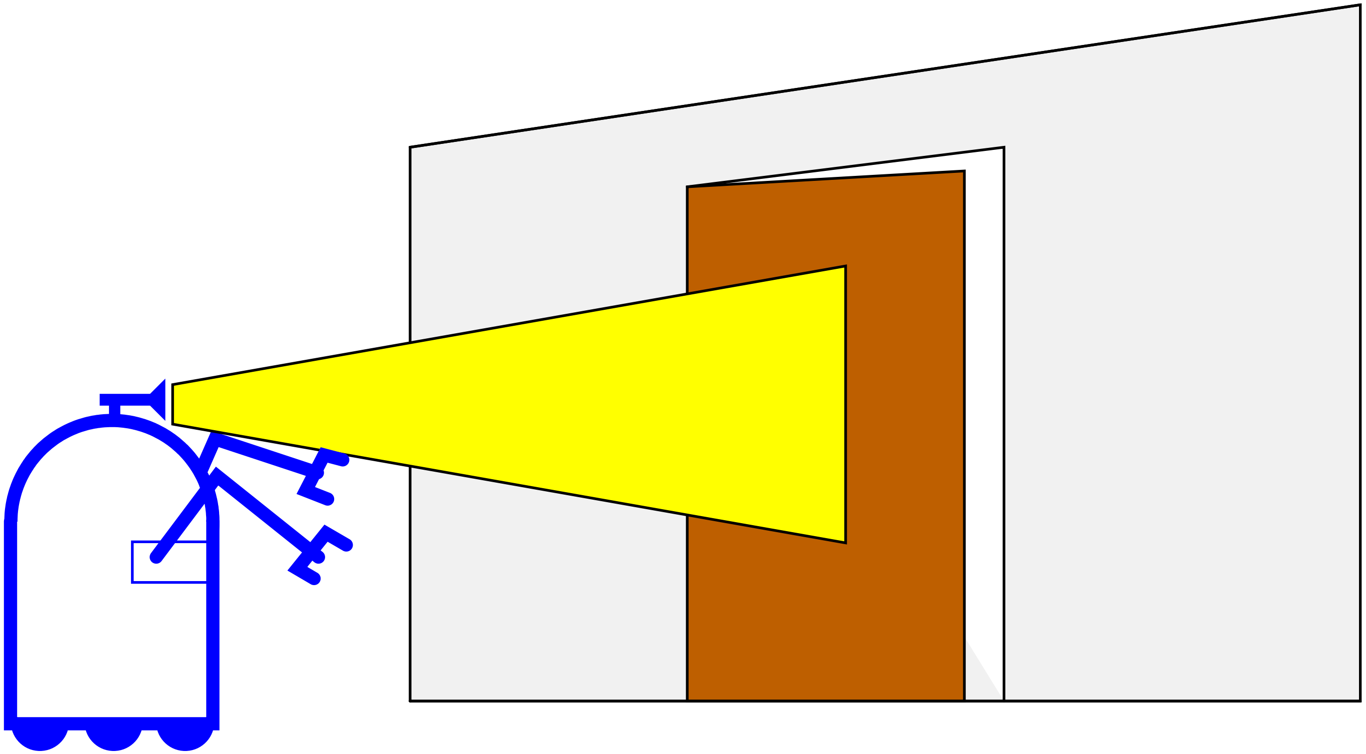
Courtesy Sebastian Thrun, Wolfram Burgard and Dieter Fox

Applications

- Can use Bayesian filtering in many other domains
  - Map building
  - Simultaneous localization and mapping (SLAM)
  - Feature tracking
  - Pose estimation
  - Target tracking

Simple Example of State Estimation

- Suppose a robot obtains measurement *z (e.g. brightness)*
- What is $P(open|z)$?



Courtesy Sebastian Thrun, Wolfram Burgard and Dieter Fox

Causal vs. Diagnostic Reasoning

- *P(open|z)* is diagnostic.
- *P(z|open)* is causal
  - Light sensor: If the door is open, what's the likelihood that the sensor receives this amount of light
  - Robot localization: Given a map, what's the likelihood that the sensor (camera/laser/etc.) gets this measurement
- Often causal knowledge is easier to obtain.
- Bayes rule allows us to use causal knowledge:

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

Courtesy Sebastian Thrun, Wolfram Burgard and Dieter Fox

Example

- $\boxed{P(z|open) = 0.6 \qquad P(z|\neg open) = 0.3}$ → Likelihood
- $\boxed{P(open) = P(\neg open) = 0.5}$ → Prior

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

Law of Total Probability

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

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

Courtesy Sebastian Thrun, Wolfram Burgard and Dieter Fox

Combining Measurement

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

Courtesy Sebastian Thrun, Wolfram Burgard and Dieter Fox

Bayesian Update

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

z_n is independent of $z_1, \ldots, z_{n-1}$ if we know x:

Conditional Independence

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

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

$$= \eta_{1 \ldots n} \prod_{i=1 \ldots n} P(z_i \mid x)\, P(x)$$

Courtesy Sebastian Thrun, Wolfram Burgard and Dieter Fox

Example: Second Measurement

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

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

Conditional Independence

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

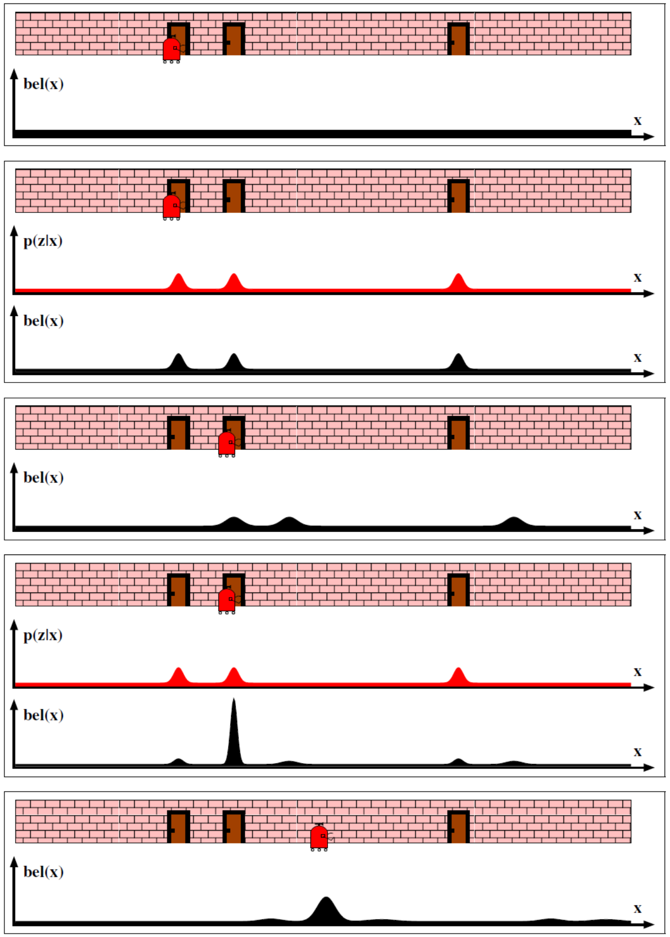
- z_2 lowers the probability that the door is open.

Courtesy Sebastian Thrun, Wolfram Burgard and Dieter Fox

Kalman Filter

Motivation

- Real systems have uncertainty
  - Initial conditions
  - Un-modeled effects
    - Aerodynamics
    - Friction
  - Disturbances
    - Wind gust
    - Wheel slip
- Errors will compound over time if not corrected



Bayes' Filter

State

Control input

- **Prior:** $p(x_0)$
- **Process model:** $f(x_t \mid x_{t-1}, u_t)$
- **Measurement model:** $g(z_t \mid x_t)$

Measurement

- **Prediction step:**
- $p(x_t \mid z_{1:t-1}, u_{1:t}) = \int f(x_t \mid x_{t-1}, u_t)\, p(x_{t-1} \mid z_{1:t-1}, u_{1:t-1})\, dx_{t-1}$
- **Update step:**
- $p(x_t \mid z_{1:t}, u_{1:t}) = \dfrac{g(z_t \mid x_t)\, p(x_t \mid z_{1:t-1}, u_{1:t})}{\int g(z_t \mid x_t')\, p(x_t' \mid z_{1:t-1}, u_{1:t})\, dx_t'}$

Markov Property

- **Definition:** The future state of the system is conditionally independent of the past states given the current state
  - $p(x_{t+1}|x_{0:t}) = p(x_{t+1}|x_t)$
  - $p(z_t \mid x_t, z_{1:t-1}, u_{1:t}) = p(z_t \mid x_t)$
  - $p(x_t \mid x_{t-1}, z_{1:t-1}, u_{1:t}) = p(x_t \mid x_{t-1}, u_t)$
- **Question:**
  - Which of the following satisfy the Markov assumption?
    - A first order system with $x = [\text{position}], u = [\text{velocity}]$
    - A second order system with $x = [\text{position}], u = [\text{acceleration}]$
      - How about with $x = [\text{position, velocity}], u = [\text{acceleration}]$

Assumptions

- The prior state of the robot is represented by a Gaussian distribution
  - $p(x_0) \sim N(\mu_0, \Sigma_0)$
- The process model $f(x_t \mid x_{t-1}, u_t)$ is linear with additive Gaussian white noise
  - $x_t = A_t\, x_{t-1} + B_t\, u_t + n_t$
  - $n_t \sim N(0, Q_t)$
  - $x_t, n_t \in \mathbf{R}^n, u_t \in \mathbf{R}^m$, $A_t, Q_t \in \mathbf{R}^{n \times n}$, and $B_t \in \mathbf{R}^{n \times m}$
- The measurement model $g(z_t \mid x_t)$ is linear with additive Gaussian white noise
  - $z_t = C_t\, x_t + v_t$
  - $v_t \sim N(0, R_t)$
  - $z_t, v_t \in \mathbf{R}^p, C_t \in \mathbf{R}^{p \times n}$, and $R_t \in \mathbf{R}^{p \times p}$

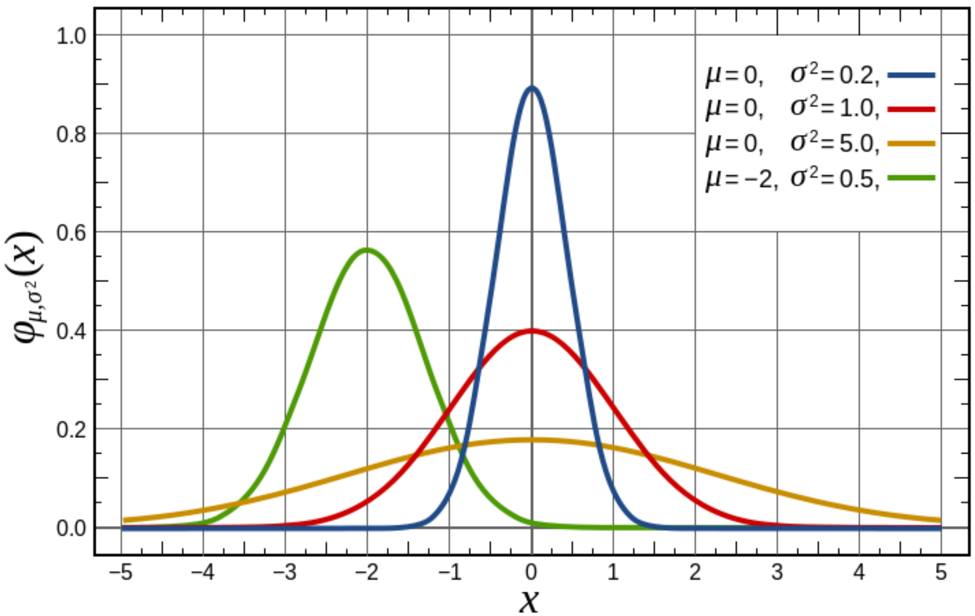
Gaussian Random Variables

Multivariate Normal (Gaussian) Distribution

- Let X be a vector of n random variables
- A multivariate normal distribution takes the form
- $f_X(x) = \frac{1}{(2\pi)^{n/2}\sqrt{\det(\Sigma)}} e^{\frac{-(x-\mu)^T\Sigma^{-1}(x-\mu)}{2}}$
- where $\mu \in \mathbf{R}^n$ and $\Sigma \in \mathbf{R}^{n\times n}$

Mean Covariance

- Fully parameterized by μ, Σ

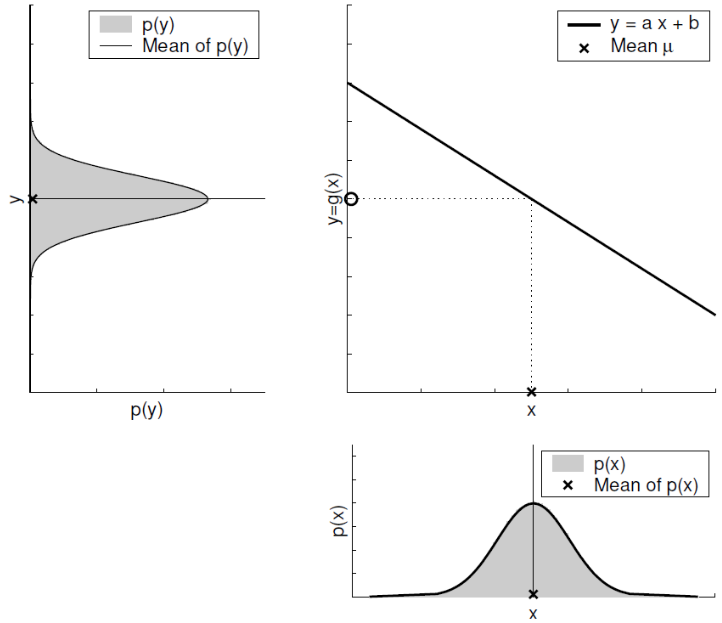


[http://en.wikipedia.org/wiki/Normal_distribution]

Affine Transformations

- Affine transformation of Gaussian distributions are Gaussian
- If $X \sim N(\mu_X, \Sigma_X)$ and $Y = AX + b$ then $Y \sim N(\mu_Y, \Sigma_Y)$ where
- $\mu_Y = A\,\mu_X + b$ and $\Sigma_Y = A\,\Sigma_X\,A^T$

- **Example:**
- $x_t = A_t\,x_{t-1} + B_t\,u_t + n_t$



Affine Transformations

- **Fact:**
  - Expectation is a linear operator of x
  - $E[X] = \int p(x)\, x\, dx$

$$\begin{aligned} \mu_Y &= E[Y] \\ &= E[AX + b] \\ &= A\, E[X] + b \\ &= A\, \mu_X + b \end{aligned}$$

$$\begin{aligned} \Sigma_Y &= E[(Y - \mu_Y)(Y - \mu_Y)^T] \\ &= E[(AX + b - A\mu_X - b)(AX + b - A\mu_X - b)^T] \\ &= E[(A(X - \mu_X))(A(X - \mu_X))^T] \\ &= A\, E[(X - \mu_X)(X - \mu_X)^T]\, A^T \\ &= A\, \Sigma_X\, A^T \end{aligned}$$

Independence

- Let $X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ where X_1, X_2 are uncorrelated, i.e., the covariance is of the form $\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$ where $\Sigma_{12} = \Sigma_{21} = 0$
- Then X_1, X_2 are independent and $f_X(X) = f_{X_1}(X_1)\, f_{X_2}(X_2)$
- **Note:** The converse is always true, i.e., if two random variables are independent then they are uncorrelated
- **Example:** We assume that the noise is independent of the state of the system

Sum of Independent Gaussians

- Let X, Y be independent multivariate Gaussian random variables with mean μ_X, μ_Y and covariance Σ_X, Σ_Y
- The sum $Z = X + Y$ is also Gaussian with mean $\mu_Z = \mu_X + \mu_Y$ and covariance $\Sigma_Z = \Sigma_X + \Sigma_Y$

- **Example:**
  - $x_t = x_{t-1} + n_t$
  - $z_t = x_t + v_t$

Jointly Normal Random Vectors

- Let X be a multivariate Gaussian random variable and let $X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$
- Then X_1, X_2 are both (multivariate) Gaussian random variables and are jointly normally distributed

- **Note:** If X_1, X_2 are both (multivariate) Gaussian random variables then it does *not* necessarily imply that $X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ is also Gaussian
- **Note:** If X_1, X_2 are independent (multivariate) Gaussian random variables then $X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ is also Gaussian

Conditional Distributions

- Let $X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ be a multivariate Gaussian with mean $\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$ and covariance $\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$
- Then the conditional density $f_{X_1|X_2}(x_1|X_2 = x_2)$ is a multivariate normal distribution with
  - mean $\mu_{X_1|X_2} = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2)$
  - covariance $\Sigma_{X_1|X_2} = \Sigma_{11} - \Sigma_{12}\,\Sigma_{22}^{-1}\,\Sigma_{21}$
- **Note:** $\Sigma_{X_1|X_2}$ is the Schur complement of Σ_{22}

Further readings: http://fourier.eng.hmc.edu/e161/lectures/gaussianprocess/node7.html

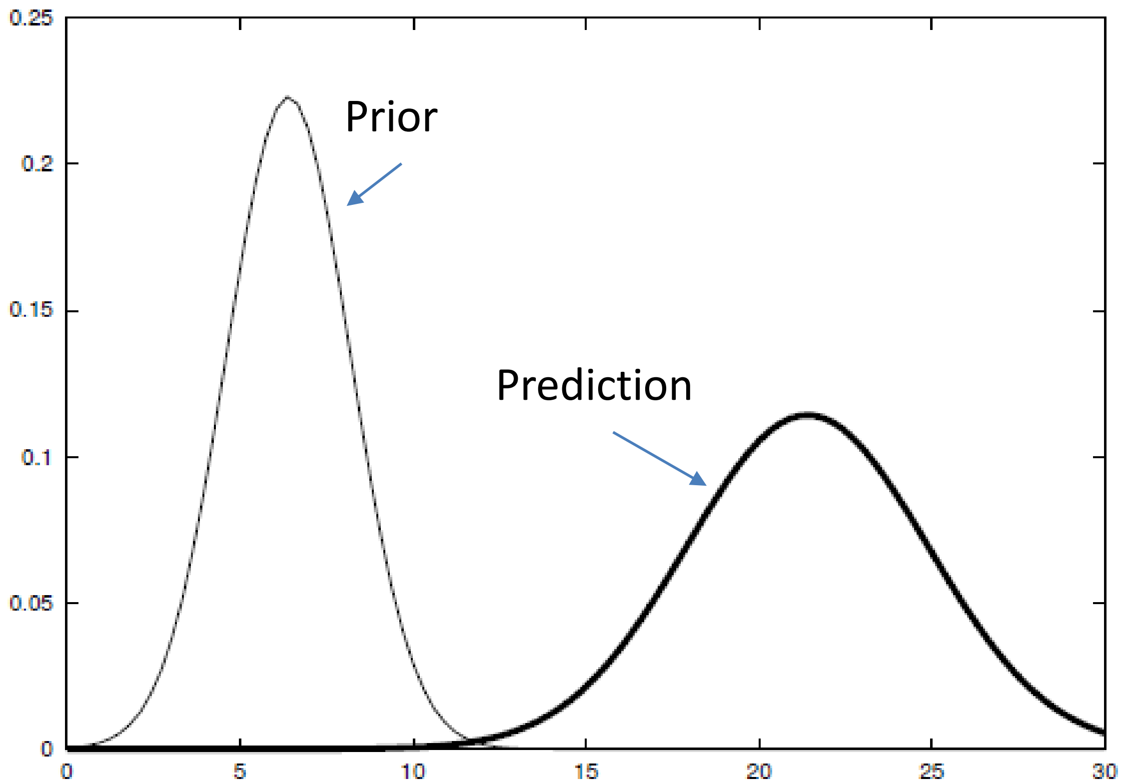
The Kalman Filter

System Model

- The prior state of the robot is represented by a Gaussian distribution
  - $p(x_0) \sim N(\mu_0, \Sigma_0)$
- The process model $f(x_t \mid x_{t-1}, u_t)$ is linear with additive Gaussian white noise
  - $x_t = A_t\, x_{t-1} + B_t\, u_t + n_t$
  - $n_t \sim N(0, Q_t)$
- The measurement model $g(z_t \mid x_t)$ is linear with additive Gaussian white noise
  - $z_t = C_t\, x_t + v_t$
  - $v_t \sim N(0, R_t)$

Kalman Filter – Prediction

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



[Thrun, Burgard, and Fox. *Probabilistic Robotics*. 2005]

Kalman Filter – Prediction

- Bayes:
  - $p(x_t \mid z_{1:t-1}, u_{1:t}) = \int f(x_t \mid x_{t-1}, u_t)\, p(x_{t-1} \mid z_{1:t-1}, u_{1:t-1})\, dx_{t-1}$
- $x_t = A_t\, x_{t-1} + B_t u_t + n_t$
- $n_t \sim N(0, Q_t)$
- Prior: $p(x_{t-1} \mid z_{1:t-1}, u_{1:t-1}) \sim N(\mu_{t-1}, \Sigma_{t-1})$
- Prediction:
  - $\bar{\mu}_t = A\, \mu_{t-1} + B\, u_t$
  - $\bar{\Sigma}_t = A\, \Sigma_{t-1}\, A^T + Q$

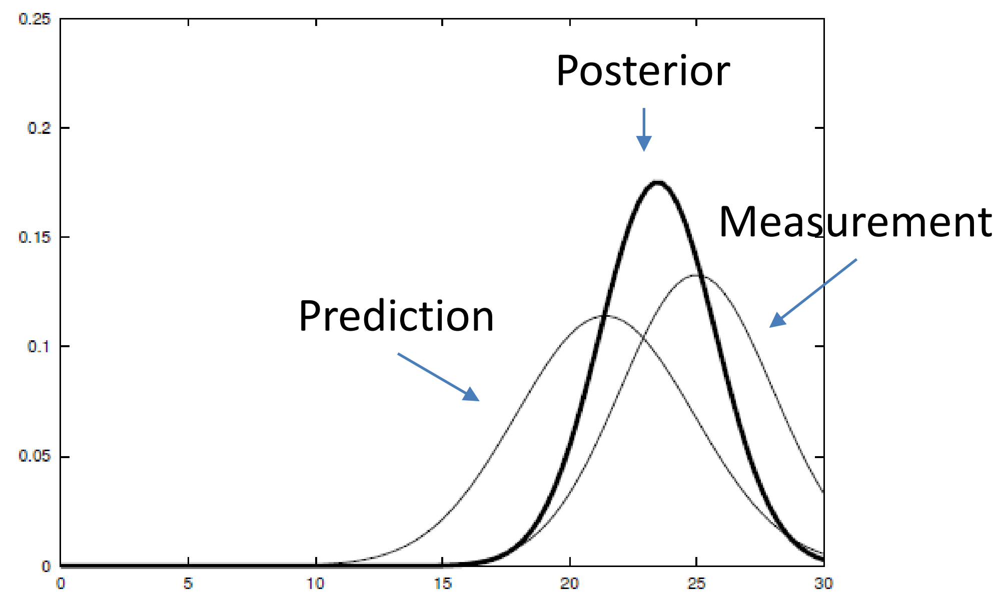
Kalman Filter – Update

- Bayes: $p(x_t \mid z_{1:t}, u_{1:t}) = \dfrac{g(z_t \mid x_t)\, p(x_t \mid z_{1:t-1}, u_{1:t})}{\int g(z_t \mid x_t')\, p\left(x_t' \mid z_{1:t-1}, u_{1:t}\right) dx_t'}$
- The observation model is $z_t = C_t\, \bar{x}_t + v_t, \quad v_t \sim N(0, R_t)$
- The best update without a measurement is to set $x_t = \bar{x}_t$
- $\begin{bmatrix} x_t \\ z_t \end{bmatrix} = \begin{bmatrix} I & 0 \\ C & I \end{bmatrix} \begin{bmatrix} \bar{x}_t \\ v_t \end{bmatrix}$
- **Question:** Is this a jointly normal distribution?
- $\mu = \begin{bmatrix} \bar{\mu}_t \\ C\bar{\mu}_t \end{bmatrix}$
- $\Sigma = \begin{bmatrix} I & 0 \\ C & I \end{bmatrix} \begin{bmatrix} \bar{\Sigma}_t & 0 \\ 0 & R \end{bmatrix} \begin{bmatrix} I & C^T \\ 0 & I \end{bmatrix} = \begin{bmatrix} \bar{\Sigma}_t & \bar{\Sigma}_t C^T \\ C\bar{\Sigma}_t & C\bar{\Sigma}_t C^T + R \end{bmatrix}$

Kalman Filter – Update

- The distribution of x_t conditioned on z_t is thus normal with
- $\mu_{x_t|z_t} = \bar{\mu}_t + \bar{\Sigma}_t C^T (C\bar{\Sigma}_t C^T + R)^{-1}(z_t - C\bar{\mu}_t)$
- $\Sigma_{x_t|z_t} = \bar{\Sigma}_t - \bar{\Sigma}_t C^T (C\bar{\Sigma}_t C^T + R)^{-1} C\bar{\Sigma}_t$

- Define the Kalman gain K_t
- $K_t = \bar{\Sigma}_t C^T (C\bar{\Sigma}_t C^T + R)^{-1}$
- $\mu_t = \bar{\mu}_t + K_t(z_t - C\bar{\mu}_t)$
- $\Sigma_t = \bar{\Sigma}_t - K_t C\bar{\Sigma}_t$



[Thrun, Burgard, and Fox. *Probabilistic Robotics*. 2005]

Kalman Gain

- $K_t = \bar{\Sigma}_t\, C^T\, (C\, \bar{\Sigma}_t\, C^T + R)^{-1}$
- **Intuition:** How much to trust the sensor vs. the prediction
- **Example:**
  - Perfect sensor $R = 0$
    - $K_t = \bar{\Sigma}_t\, C^T\, (C\, \bar{\Sigma}_t\, C^T + R)^{-1} = C^{-1}$
    - $\mu_t = \bar{\mu}_t + K_t(z_t - C\, \bar{\mu}_t) = C^{-1} z_t$
    - $\Sigma_t = \bar{\Sigma}_t - K_t\, C\, \bar{\Sigma}_t = 0$
  - Horrible sensor $R \rightarrow \infty$
    - $K_t = \bar{\Sigma}_t\, C^T\, (C\, \bar{\Sigma}_t\, C^T + R)^{-1} \rightarrow 0$
    - $\mu_t = \bar{\mu}_t + K_t(z_t - C\, \bar{\mu}_t) \rightarrow \bar{\mu}_t$
    - $\Sigma_t = \bar{\Sigma}_t - K_t\, C\, \bar{\Sigma}_t \rightarrow \bar{\Sigma}_t$

Kalman Filter

- Prior:
  - $p(x_0) \sim N(\mu_0, \Sigma_0)$
- Process model:
  - $x_t = A_t\, x_{t-1} + B_t\, u_t + n_t$
  - $n_t \sim N(0, Q_t)$
- Measurement model:
  - $z_t = C_t\, x_t + v_t$
  - $v_t \sim N(0, R_t)$

- Prior:
  - μ_{t-1}, Σ_{t-1}
- Prediction:
  - $\bar{\mu}_t = A_t\, \mu_{t-1} + B_t\, u_t$
  - $\bar{\Sigma}_t = A_t\, \Sigma_{t-1}\, A_t^T + Q_t$
- Update:
  - $\mu_t = \bar{\mu}_t + K_t(z_t - C_t\, \bar{\mu}_t)$
  - $\Sigma_t = \bar{\Sigma}_t - K_t\, C_t\, \bar{\Sigma}_t$
  - $K_t = \bar{\Sigma}_t\, C_t^T\, (C_t\, \bar{\Sigma}_t\, C_t^T + R_t)^{-1}$

Example Problem

$$x_t = x_{t-1} + u_t + n_t$$

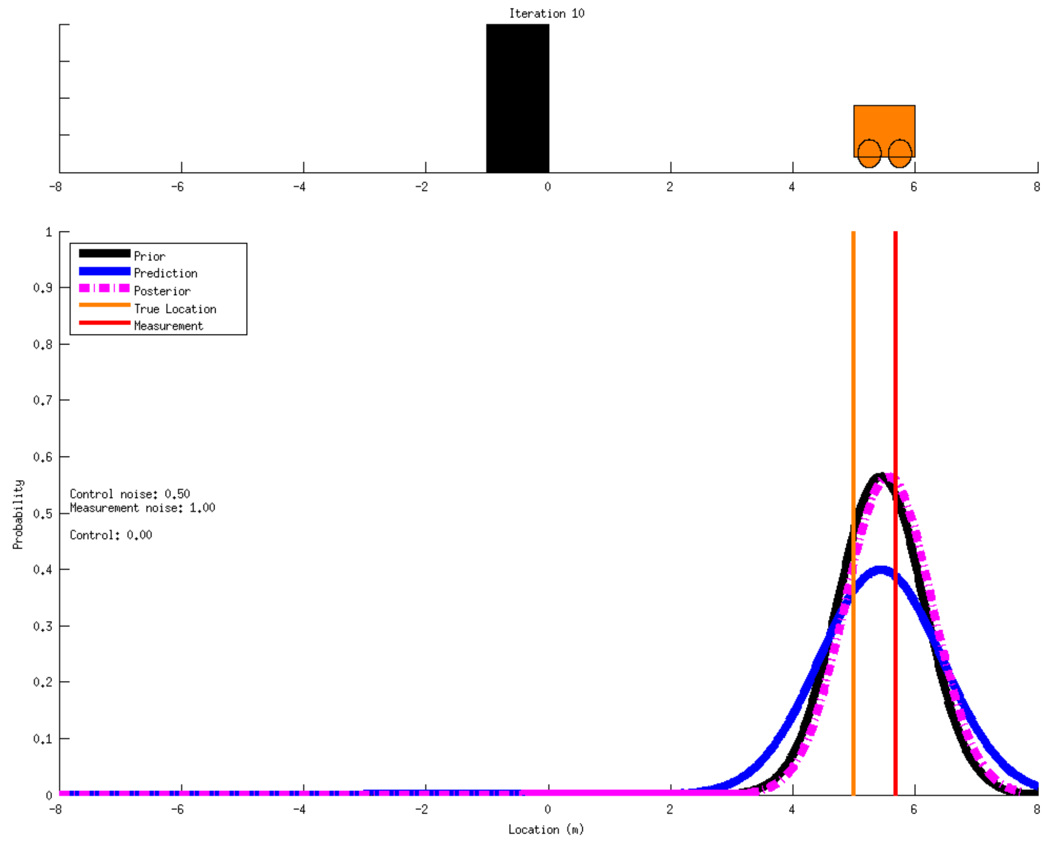
$$Q_t = 0.5$$

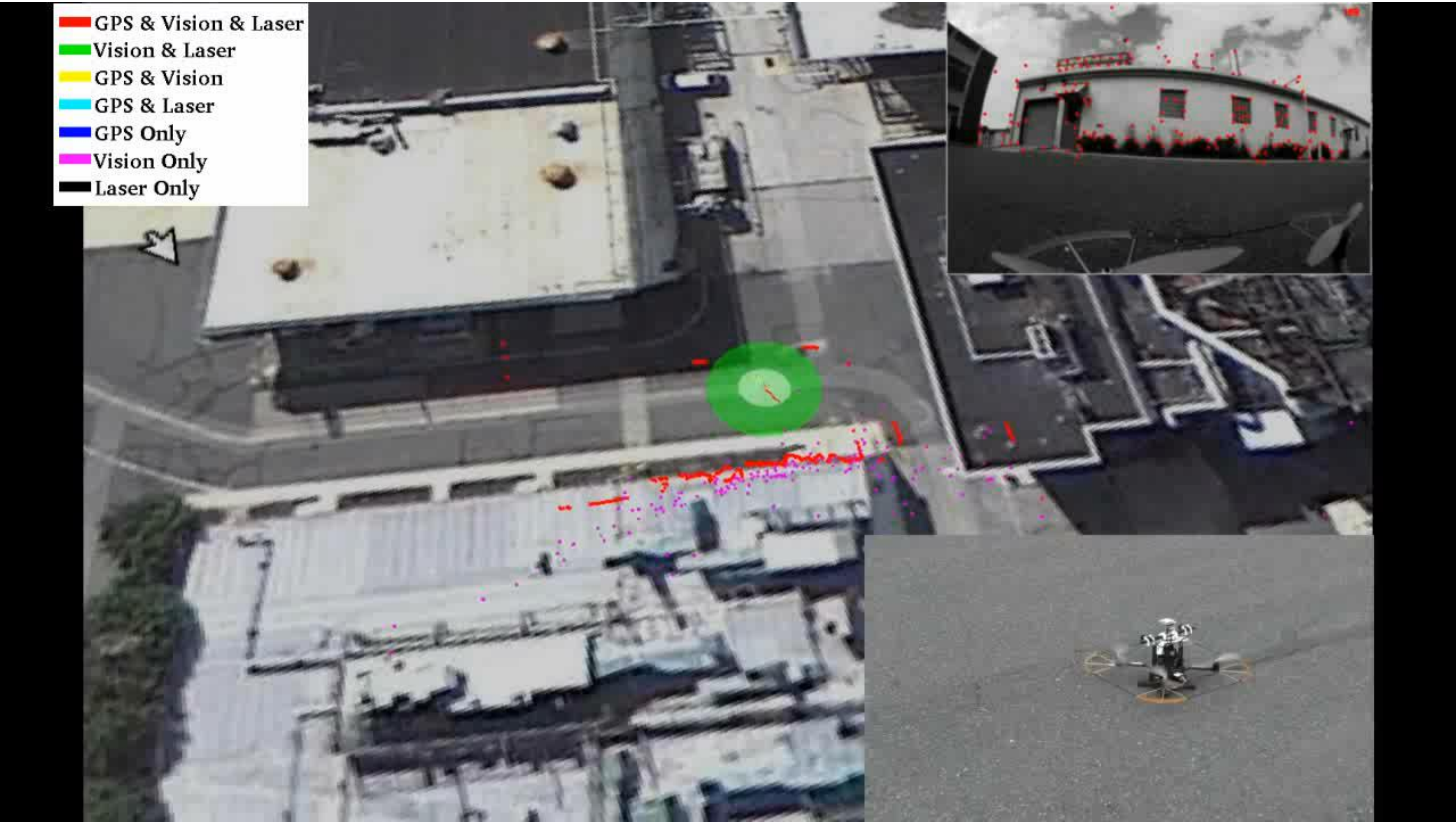
$$A_t = B_t = 1$$

$$z_t = x_t + v_t$$

$$R_t = 1.0$$

$$C_t = 1$$




GPS & Vision & Laser
Vision & Laser
GPS & Vision
GPS & Laser
GPS Only
Vision Only
Laser Only

Kalman Filter Facts

- If the distribution is not Gaussian, the Kalman filter is the minimum variance linear estimator
  - The noise must be uncorrelated with the initial state x_0
- The variance never increases due to receiving a measurement
- The variance update is independent of the measurement realization
- Prediction and update can happen in arbitrary order as long as they are temporally sorted

Continuous Time Systems

Discrete vs. Continuous Time

Discrete Time

- Events occur at discrete points in time
- Time intervals often evenly spaced

- Example:
  - Kinematic cart
  - $x_t = x_{t-1} + u_t + n_t$

Continuous Time

- Events may occur infinitesimally close to each other in time

- Example:
  - Ballistic motion
  - $\ddot{x} = -g + u + n$

Continuous Time Systems

- There is a continuous time version of the Kalman Filter
  - Continuous dynamics
  - Continuous observations
- Often called the Kalman-Bucy Filter
- Much less commonly used
- *Not* covered in this course

Continuous Dynamics

- $\dot{x} = f(x, u, n) = A\,x + B\,u + U\,n$
- **Question:** How do we turn this into a discrete time system?
  - State-transition matrix
  - Numerical integration
- One-step Euler integration
  - $x_t = x_{t-1} + f(x_{t-1}, u_t, n_t)\,\delta t$
  - $x_t = (I + \delta t\,A)\,x_{t-1} + (\delta t\,B)\,u_t + (\delta t\,U)\,n_t$
  - $x_t = F\,x_{t-1} + G\,u_t + V\,n_t$
- Prediction:
  - $\bar{\mu}_t = F\,\mu_{t-1} + G\,u_t$
  - $\bar{\Sigma}_t = F\,\Sigma_{t-1} F^T + V\,Q\,V^T$

Example Problem

- Second order system $\mathbf{x} = [s, \dot{s}]^T$
- Input is a force $\ddot{s} = u$
- $\dot{\mathbf{x}} = f(\mathbf{x}, u, n) = A\,\mathbf{x} + B\,u + U\,n$
- $F = (I + \delta t\,A)$
- $G = \delta t\,B$
- $V = \delta t\,U$
- Prediction:
  - $\bar{\mu}_t = F\,\mu_{t-1} + G\,u_t$
  - $\bar{\Sigma}_t = F\,\Sigma_{t-1} F^T + V\,Q\,V^T$

- $\dot{\mathbf{x}} = \begin{bmatrix} \dot{s} \\ \ddot{s} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u + n$
- $F = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \delta t \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & \delta t \\ 0 & 1 \end{bmatrix}$
- $G = \delta t \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ \delta t \end{bmatrix}$
- $V = \delta t$

Recap

Bayes' Filter

State

Control input

- **Prior:** $p(x_0)$
- **Process model:** $f(x_t \mid x_{t-1}, u_t)$
- **Measurement model:** $g(z_t \mid x_t)$

Measurement

- **Prediction step:**
- $p(x_t \mid z_{1:t-1}, u_{1:t}) = \int f(x_t \mid x_{t-1}, u_t)\, p(x_{t-1} \mid z_{1:t-1}, u_{1:t-1})\, dx_{t-1}$
- **Update step:**
- $p(x_t \mid z_{1:t}, u_{1:t}) = \dfrac{g(z_t \mid x_t)\, p(x_t \mid z_{1:t-1}, u_{1:t})}{\int g(z_t \mid x'_t)\, p(x'_t \mid z_{1:t-1}, u_{1:t})\, dx'_t}$

Assumptions

- The prior state of the robot is represented by a Gaussian distribution
  - $p(x_0) \sim N(\mu_0, \Sigma_0)$
- The process model $f(x_t \mid x_{t-1}, u_t)$ is linear with additive Gaussian white noise
  - $x_t = A_t\, x_{t-1} + B_t\, u_t + n_t$
  - $n_t \sim N(0, Q_t)$
- The measurement model $g(z_t \mid x_t)$ is linear with additive Gaussian white noise
  - $z_t = C_t\, x_t + v_t$
  - $v_t \sim N(0, R_t)$

Kalman Filter

- Prior:
  - $p(x_0) \sim N(\mu_0, \Sigma_0)$
- Process model:
  - $x_t = A_t \, x_{t-1} + B_t \, u_t + n_t$
  - $n_t \sim N(0, Q_t)$
- Measurement model:
  - $z_t = C_t \, x_t + v_t$
  - $v_t \sim N(0, R_t)$

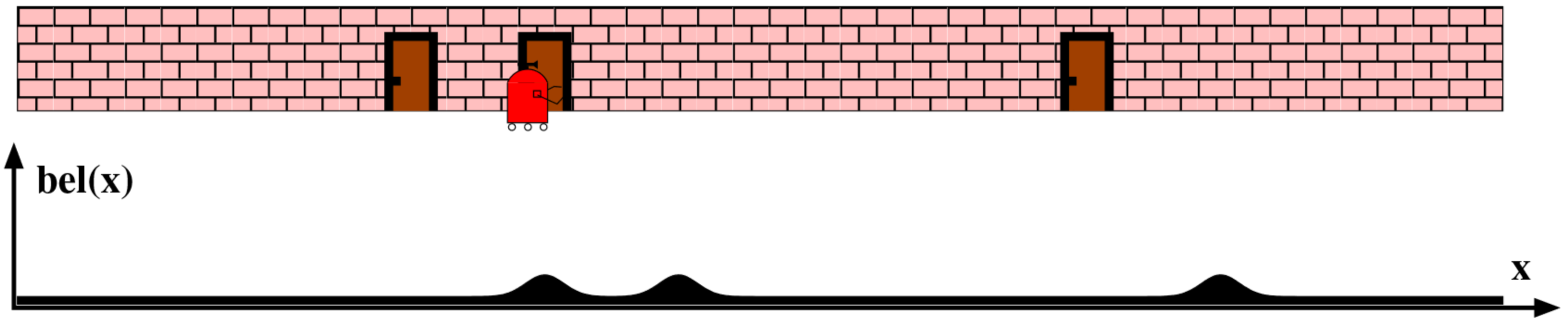
- Prior:
  - μ_{t-1}, Σ_{t-1}
- Prediction:
  - $\bar{\mu}_t = A_t \, \mu_{t-1} + B_t \, u_t$
  - $\bar{\Sigma}_t = A_t \, \Sigma_{t-1} \, A_t^T + Q_t$
- Update:
  - $\mu_t = \bar{\mu}_t + K_t(z_t - C_t \, \bar{\mu}_t)$
  - $\Sigma_t = \bar{\Sigma}_t - K_t \, C_t \, \bar{\Sigma}_t$
  - $K_t = \bar{\Sigma}_t \, C_t^T \, (C_t \, \bar{\Sigma}_t \, C_t^T + R_t)^{-1}$

Continuous Dynamics

- Can convert continuous time systems
- $\dot{x} = f(x, u, n) = A\,x + B\,u + U\,n$
- Into discrete time systems using one-step Euler integration
- $x_t = F\,x_{t-1} + G\,u_t + V\,n_t$
- $F = (I + \delta t\,A), G = \delta t\,B,\; V = \delta t\,U$
- This will introduce some error, but the observations can help correct it
- Prediction:
  - $\bar{\mu}_t = F\,\mu_{t-1} + G\,u_t$
  - $\bar{\Sigma}_t = F\,\Sigma_{t-1}F^T + V\,Q\,V^T$

Kalman Filter Discussion

- **Advantages:**
  - Simple
  - Purely matrix operations
    - Computationally efficient, even for high dimensional systems
- **Disadvantages:**
  - Assumes everything is linear and Gaussian
  - Unimodal distribution
    - Cannot handle multiple hypotheses



Reading

- "Probabilistic Robotics", Sebastian Thrun, Wolfram Burgard, and Dieter Fox, Chapter 2, Chapter 3

Logistics

- Nothing due this week, do spend more time on your visual odometry project