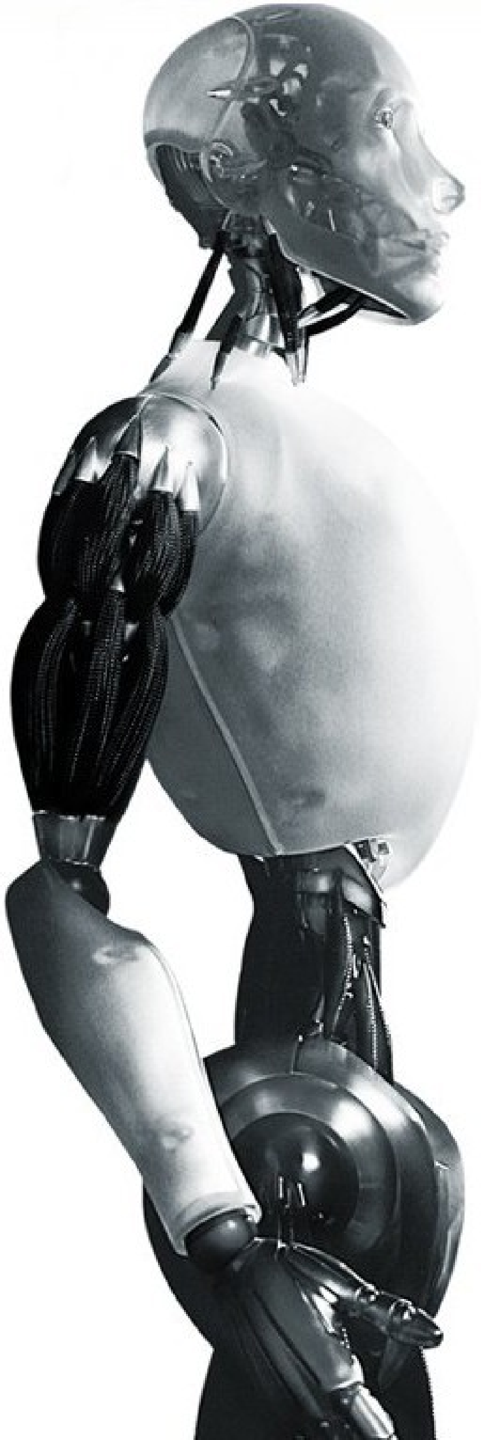
Disclaimer

These slides are intended as presentation aids for the lecture. They contain information that would otherwise be to difficult or time-consuming to reproduce on the board. But they are incomplete, not self-explanatory, and are not always used in the order they appear in this presentation. As a result, these slides should not be used as a script for this course. I recommend you take notes during class, maybe on the slides themselves. It has been shown that taking notes improves learning success.



Robotics

Monte Carlo localization

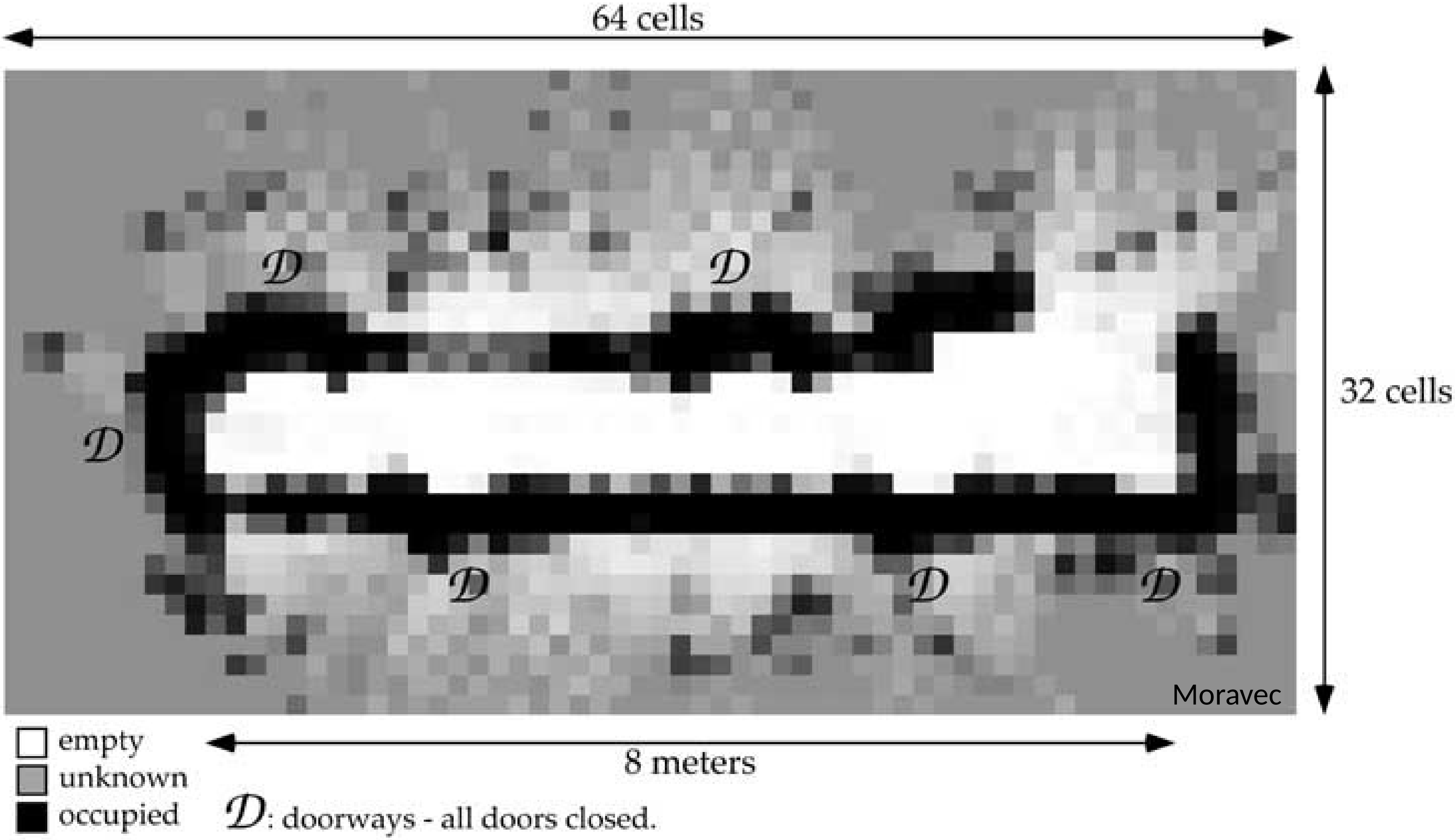
TU Berlin
Oliver Brock

Reading for this set of slides

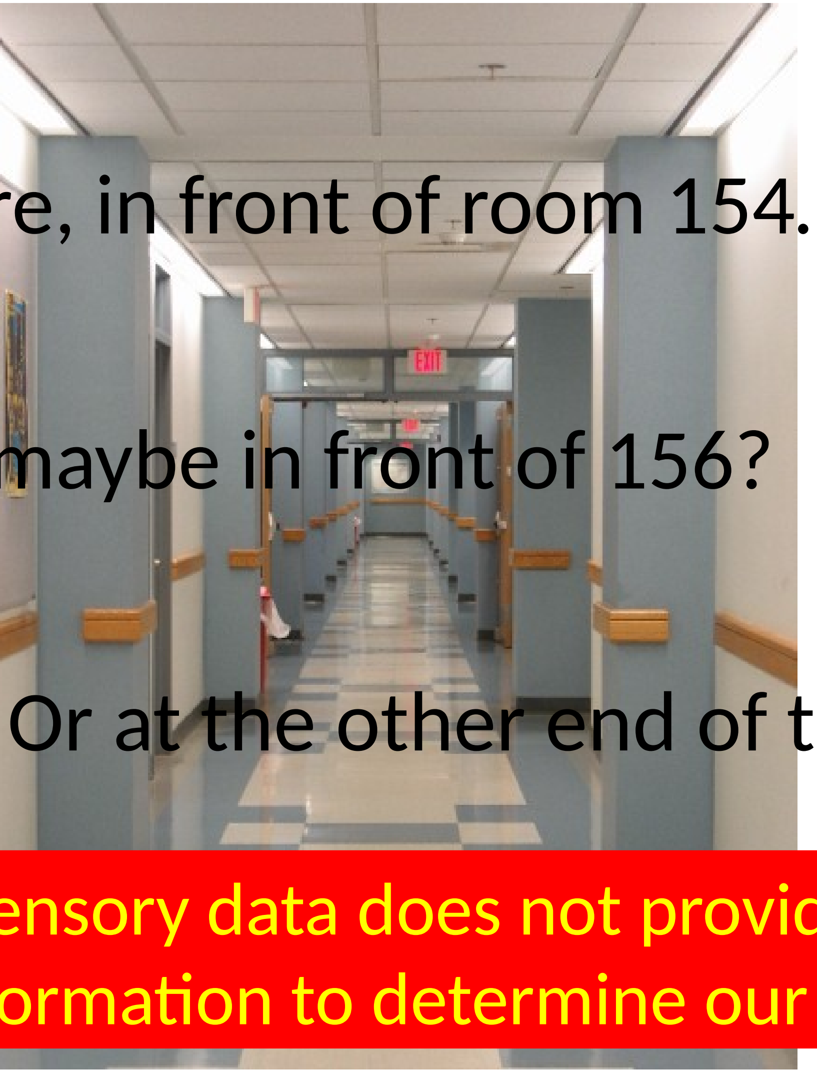
- Probabilistic Robotics
  - Chapters 1-4, 7, 8-10 (please match the level of detail from the lectures, not all the material in these chapters is required)

Please note that this set of slides is intended as support for the lecture, not as a stand-alone script. If you want to study for this course, please use these slides in conjunction with the indicated chapters in the text books. The textbooks are available online or in the TUB library (many copies that can be checked out for the entire semester. There are also some aspects of the lectures that will not be covered in the text books but can still be part of the homework or exam. For those It is important that you attend class or ask somebody about what was covered in class.

Occupancy Grid – A Map



Where are you?



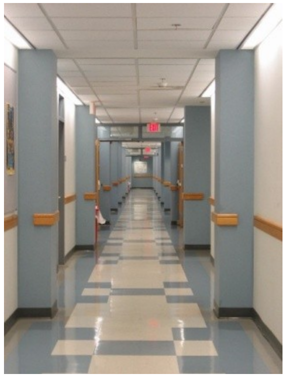
Pretty sure, in front of room 154….

But maybe in front of 156?

Or at the other end of the hall?

Our sensory data does not provide sufficient information to determine our position

How can we deal?

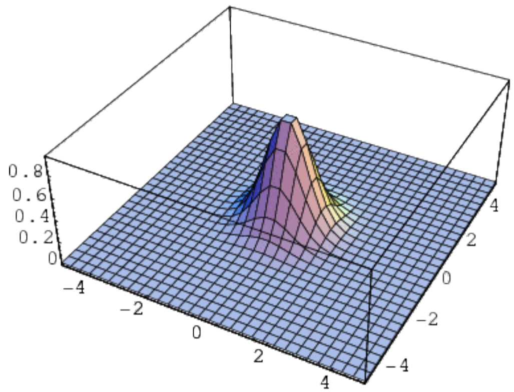


What is the probability of being in front of room 154, given we see what is shown in the image?

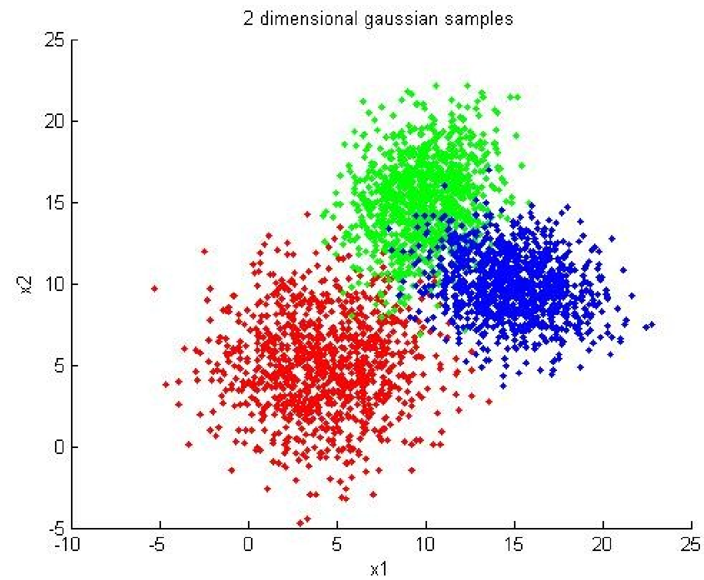
What is the probability given that we were just in front of room 156?

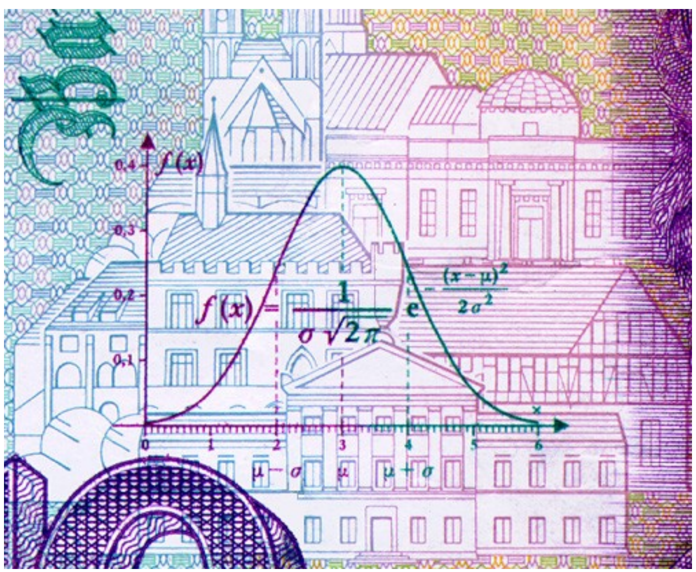
What is the probability given that we were in front of room 156 and moved 15 meters?

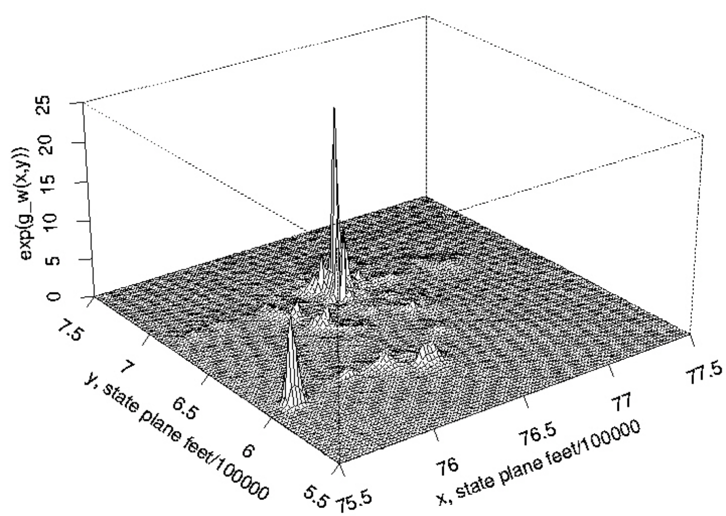
Parametric vs. Nonparametric



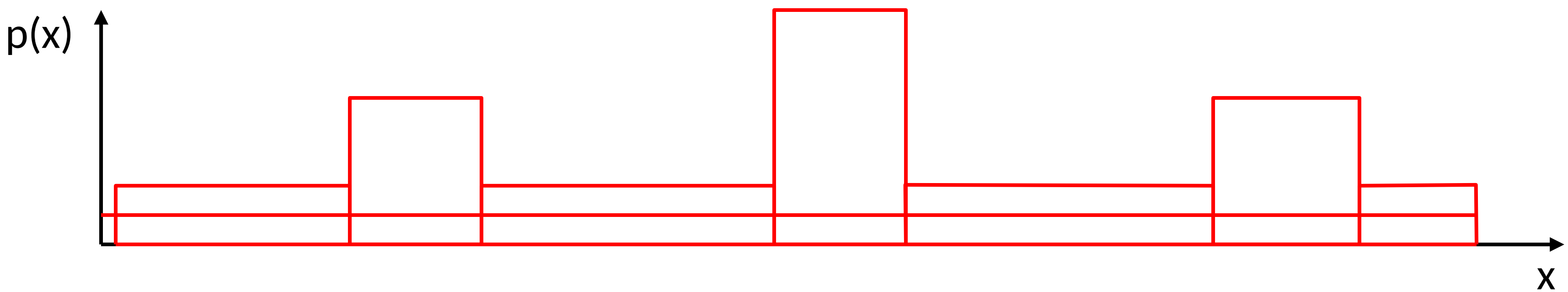
$$\frac{1}{\sigma\sqrt{2\pi}}\ \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$



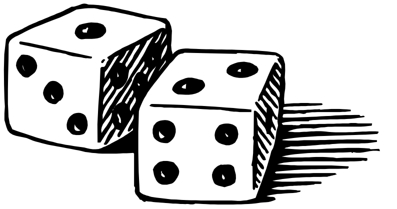




Intuition



Probabilistic Model



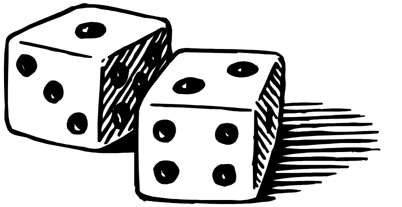
Sample Space $\Omega = \{1, 2, 3, 4, 5, 6\}$

Probability Law $L \ : \ A \subset \Omega \rightarrow 0 \leq \mathbf{P}(A) \leq 1$

$$
\begin{aligned}
\mathbf{P}(\{1\}) &= \frac{1}{6} \\
\mathbf{P}(\{2\}) &= \frac{1}{6} \\
\mathbf{P}(\{3\}) &= \frac{1}{6} \\
\mathbf{P}(\{1,2,3\}) &= \frac{1}{2} \\
\mathbf{P}(\{1,2,3,4,5,6\}) &= 1
\end{aligned}
$$

For a textbook on probability see *Introduction to Probability* by Dimitri P. Bertsekas and John N. Tsitsiklis

Probability Axioms



Nonnegativity $\forall A \subseteq \Omega \ : \ \mathbf{P}(A) \geq 0$

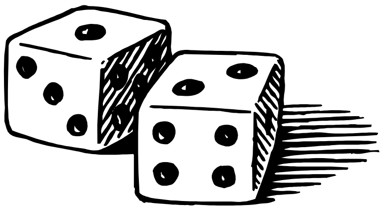
Additivity $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B)$

if A Å B = ∅

Normalization $\mathbf{P}(\Omega) = 1$

Probability Law

(Properties)



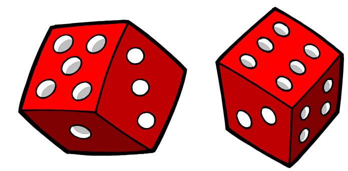
1. $A \subset B \;\Rightarrow\; \mathbf{P}(A) \leq \mathbf{P}(B)$

2. $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B)$

3. $\mathbf{P}(A \cup B) \leq \mathbf{P}(A) + \mathbf{P}(B)$

4. $\mathbf{P}(A \cup B \cup C) = \mathbf{P}(A) + \mathbf{P}(\bar{A} \cap B) + \mathbf{P}(\bar{A} \cap \bar{B} \cap C)$

Conditional Probability



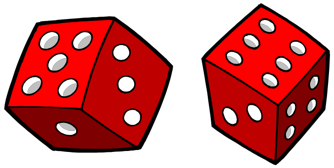
Conditional Probability allows us to reason about the outcome of an experiment based on partial information.

Given: $\underset{?}{r_1} + \underset{?}{r_2} = 9$ $\qquad \mathbf{P}(r_1 = 6) = ?$

$$A_{\Sigma=9} = \{(3,6),(4,5),(5,4),(6,3)\}$$

$$\mathbf{P}(1^{\text{st}}\text{of two rolls is 6} \mid \text{sum of 2 rolls is 9}) = \frac{1}{4}$$

More Formally



$$\mathbf{P}(1^{\text{st}}\text{of two rolls is 6} \mid \text{sum of 2 rolls is 9}) = \frac{1}{4}$$

$$\mathbf{P}(A_{r_1=6} \mid A_{\Sigma=9}) = \frac{1}{4}$$

$$A_{r_1=6} = \{(6,1),(6,2),(6,3),(6,4),(6,5),(6,6)\}$$

$$A_{\Sigma=9} = \{(3,6),(4,5),(5,4),(6,3)\}$$

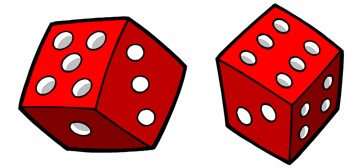
$$A_{r_1=6} \cap A_{\Sigma=9} = \{(6,3)\}$$

$$\mathbf{P}(A|B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)} = \frac{\mathbf{P}(A_{r_1=6} \cap A_{\Sigma=9})}{\mathbf{P}(A_{\Sigma=9})} = \frac{\frac{1}{36}}{\frac{4}{36}} = \frac{1}{4}$$

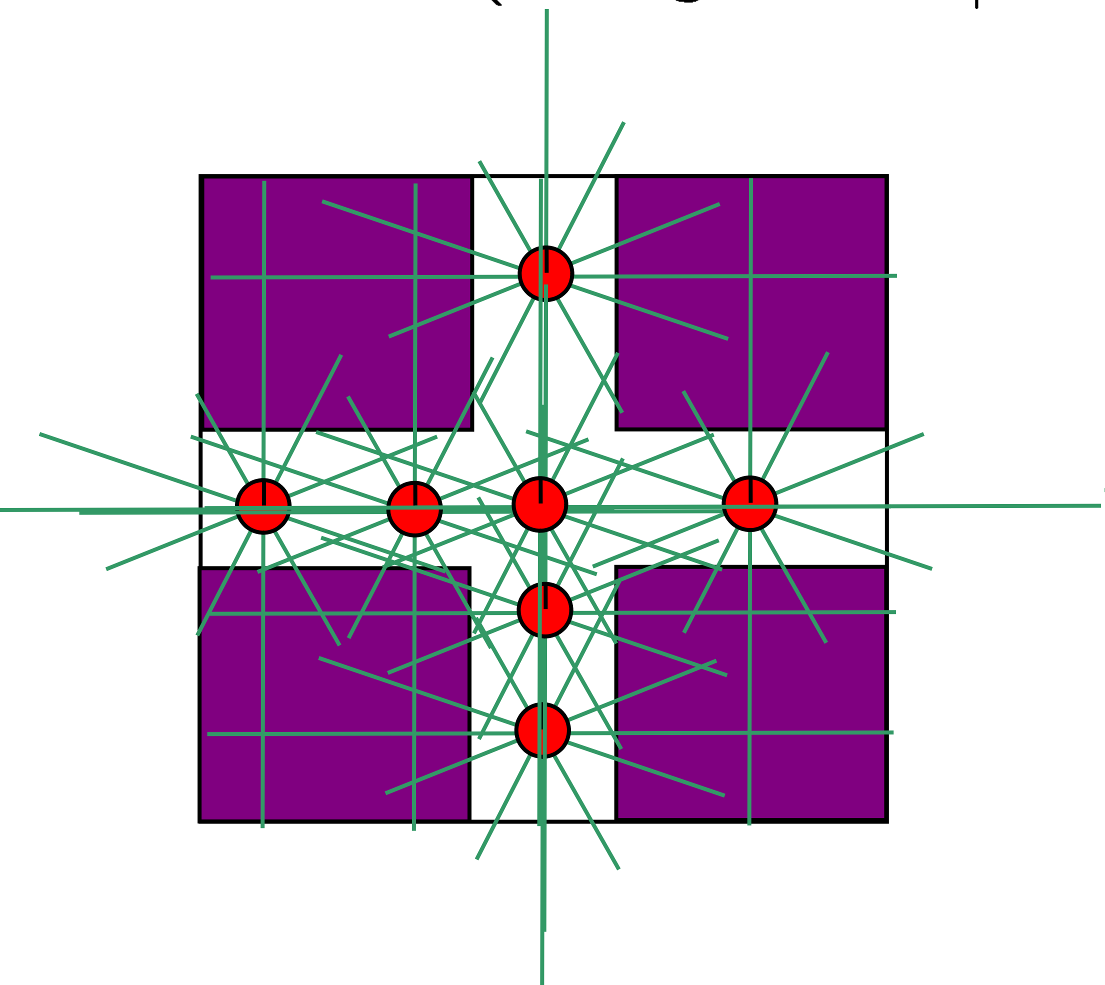
Another Example

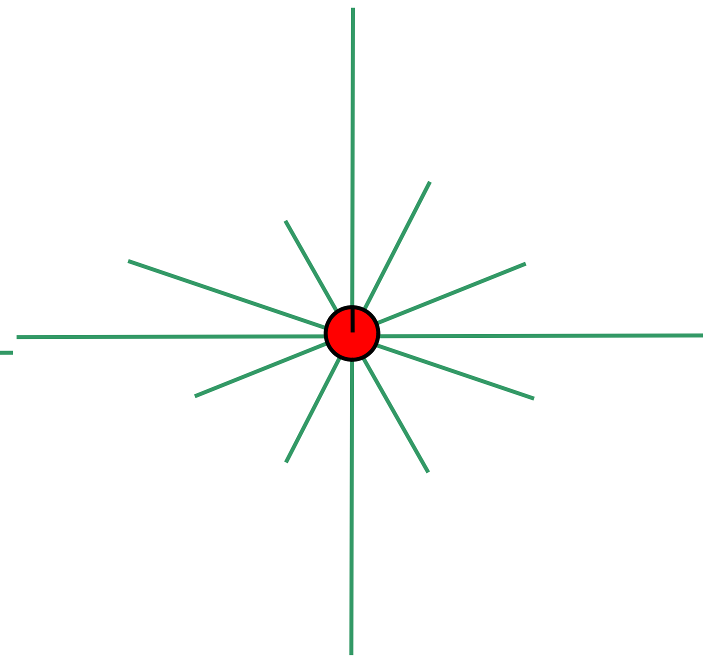
- Fair coin is thrown 3 times
- A = {more heads than tails come up}
- B = {1st toss is a head}
- P(A|B)?
- P(B) = 4/8
- P(A Å B) = 3/8
- P(A|B) = 3/4

Applying it...

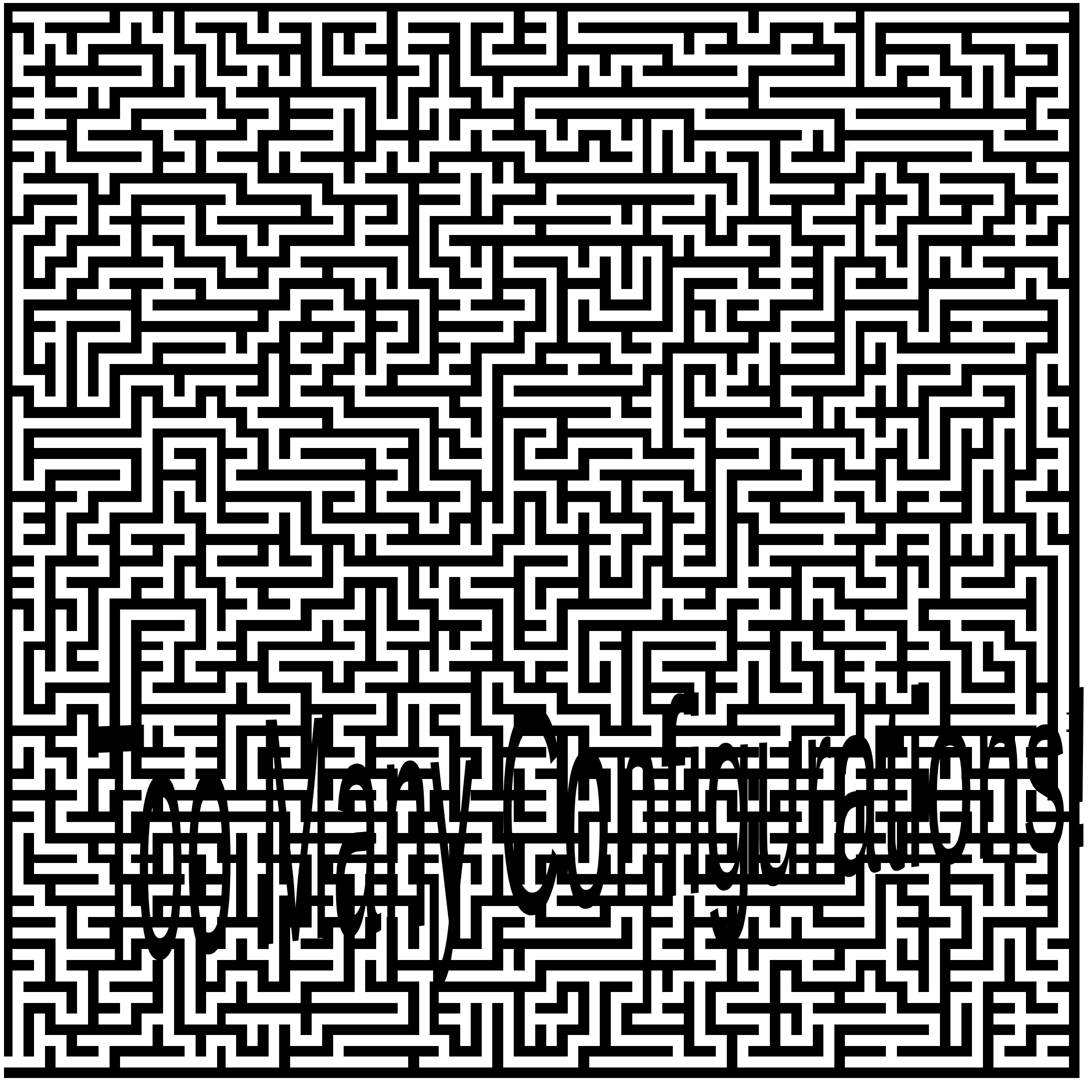


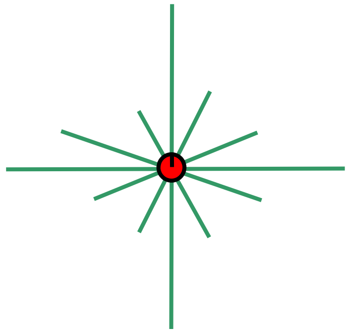
$$\mathbf{P}(\text{configuration} \mid \text{sensory information})$$





Problem!





$\mathrm{P}(\text{configuration} \mid \text{sensory information})$

Derivation of Bayes' Rule

Definition of Conditional Probability

$$\mathbf{P}(A|B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)} \qquad \mathbf{P}(B|A) = \frac{\mathbf{P}(B \cap A)}{\mathbf{P}(A)}$$

Multiplying both sides with denominator

$$\mathbf{P}(A \cap B) = \mathbf{P}(A|B)\,\mathbf{P}(B) \qquad \mathbf{P}(B \cap A) = \mathbf{P}(B|A)\,\mathbf{P}(A)$$

Set intersection is commutative

$$\mathbf{P}(A \cap B) = \mathbf{P}(B \cap A)$$

We equate the equations…

$$\mathbf{P}(B)\,\mathbf{P}(A|B) = \mathbf{P}(A)\,\mathbf{P}(B|A)$$

And devide by P(B) to arrive at Bayes' formula

$$\mathbf{P}(A|B) = \frac{\mathbf{P}(A)\,\mathbf{P}(B|A)}{\mathbf{P}(B)}$$

Interpretation of Bayes' Rule

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

$$P(\text{belief}|\text{sensory input}) = \frac{P(\text{sensory input}|\text{belief})P(\text{belief})}{P(\text{sensory input})}$$

$$P(\text{model}|\text{data}) = \frac{P(\text{data}|\text{model})}{P(\text{data})}P(\text{model})$$

Reversing the Condition

$$\mathbf{P}(A|B) = \frac{\mathbf{P}(A)\,\mathbf{P}(B|A)}{\mathbf{P}(B)}$$

$\boxed{\text{Not all configurations!}}$

(causal)

$$\mathbf{P}(\text{config}|\text{sensor}) = \frac{\mathbf{P}(\text{config})\,\mathbf{P}(\text{sensor}|\text{config})}{\mathbf{P}(\text{sensor})}$$

(diagnostic)

Summary

- Sample Space
- Probability Law
- Conditional Probability

$$\mathbf{P}(A|B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)}$$

- Bayes' Rule

$$\mathbf{P}(A|B) = \frac{\mathbf{P}(A)\ \mathbf{P}(B|A)}{\mathbf{P}(B)}$$

Localization

- Assumption:
  - we have a map!
- Approaches
  - Markov localization
  - Kalman filters
  - Hidden Markov models
  - Dynamic belief networks
  - Monte Carlo localization
  - Particle filters
  - Condensation methods

based on probabilistic representation

Monte Carlo Localization

- Represent continuous probability distribution by discrete set of samples *S* (particles)
- Samples have *importance factor*
- Initialize:
  - *m* samples with importance factor m^{-1}
- Iterate:
  - generate new samples based on the motion and the sensor information

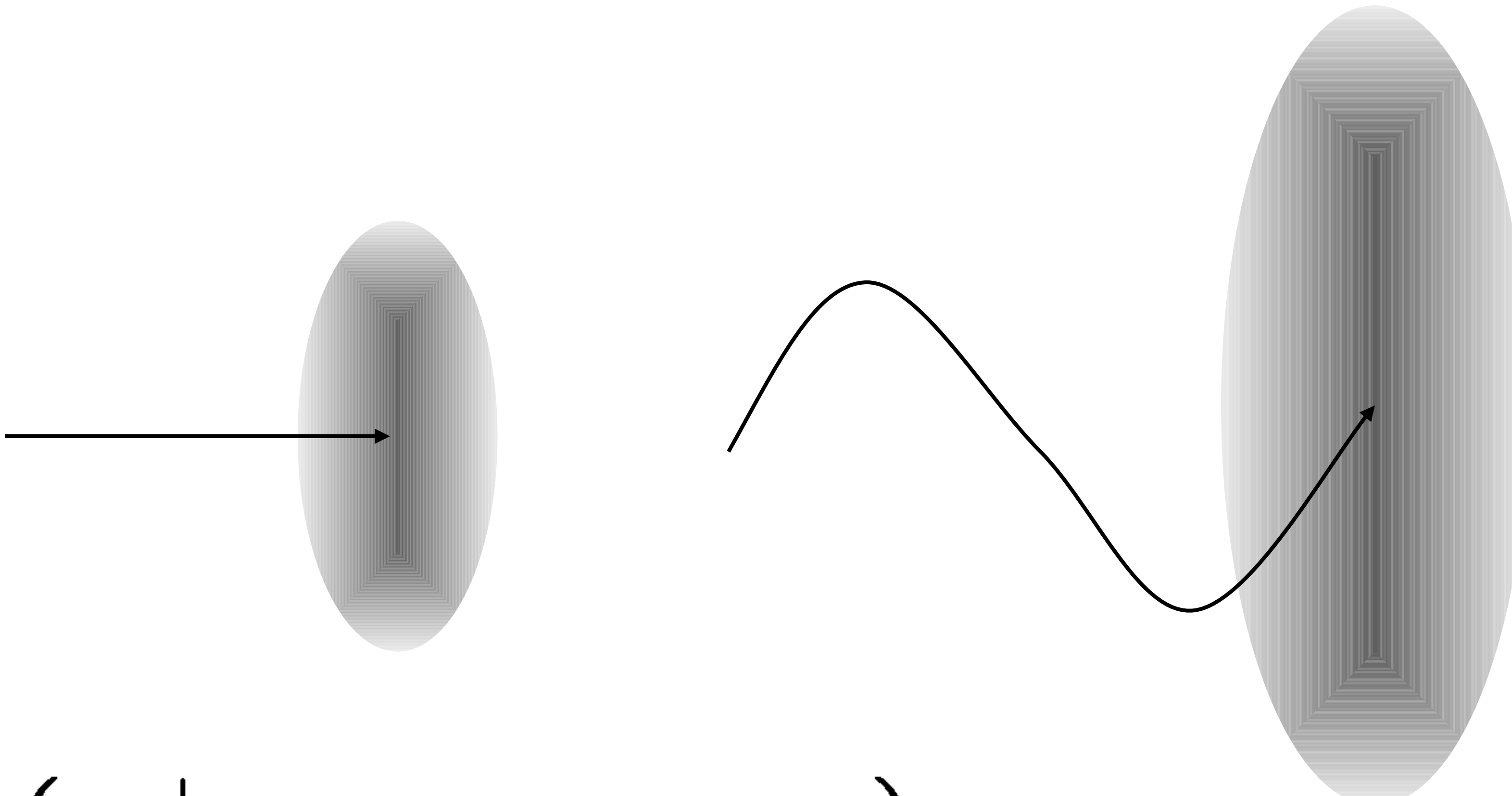
Iteration at time *t*

1. For each sample s_{t-1}
   1. Guess s_t based on motion model
   2. Assign importance factor for s_t based on sensor model
2. Repeat |S| times
   1. Pick s_t' with a probability proportional to its importance factor in P_{t-1}

Derivation

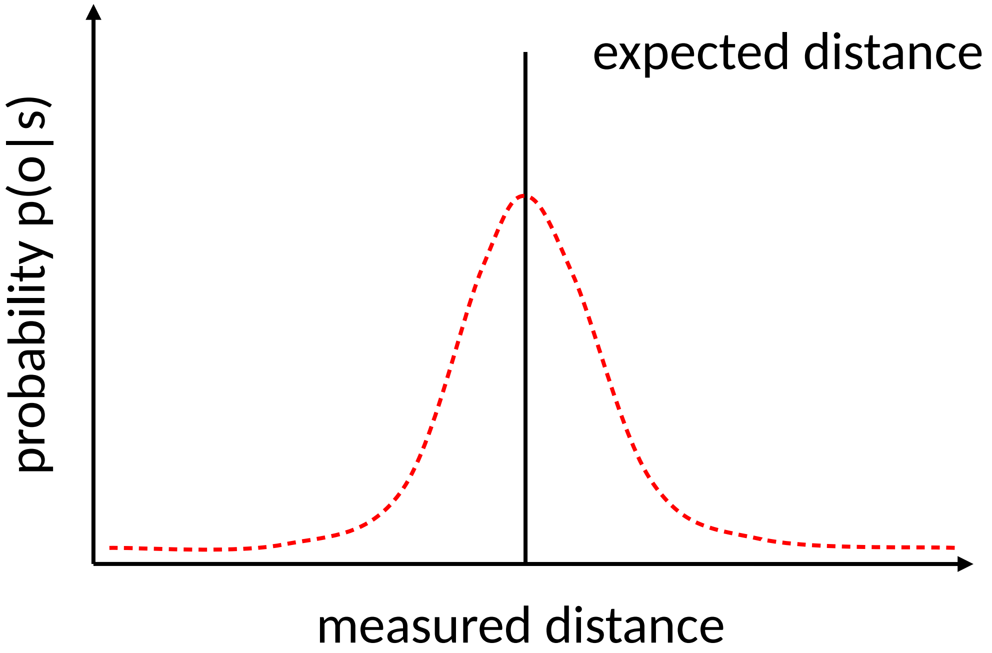
$$
\begin{aligned}
p_t(s_t) &= p(s_t \,|\, o_0, \cdots, a_{t-1}, o_t, m) \\
&= \mu_t \; p(o_t \,|\, o_0, \cdots, a_{t-1}, s_t, m) \;\cdot \\
&\qquad\qquad p(s_t \,|\, o_0, \cdots, a_{t-1}, m) \\
&= \mu_t \;\; p(o_t \,|\, s_t, m) \;\cdot \\
&\qquad\qquad p(s_t \,|\, o_0, \cdots, a_{t-1}, o_t, m) \\
&= \cdots \\
&= \mu_t \;\; p(o_t \,|\, s_t, m) \;\cdot \\
&\qquad\qquad \int p(s_t \,|\, a_{t-1}, s_{t-1}, m) \; p_{t-1}(s_{t-1}) \; ds_{t-1}
\end{aligned}
$$

Model for Motion



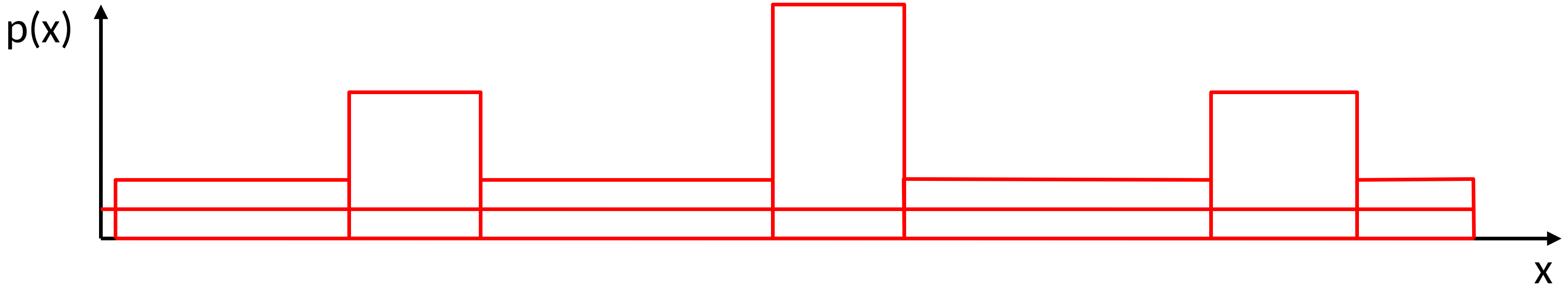
$$p(s_t \mid a_{t-1},\, s_{t-1},\, m)$$

Model for Sensing

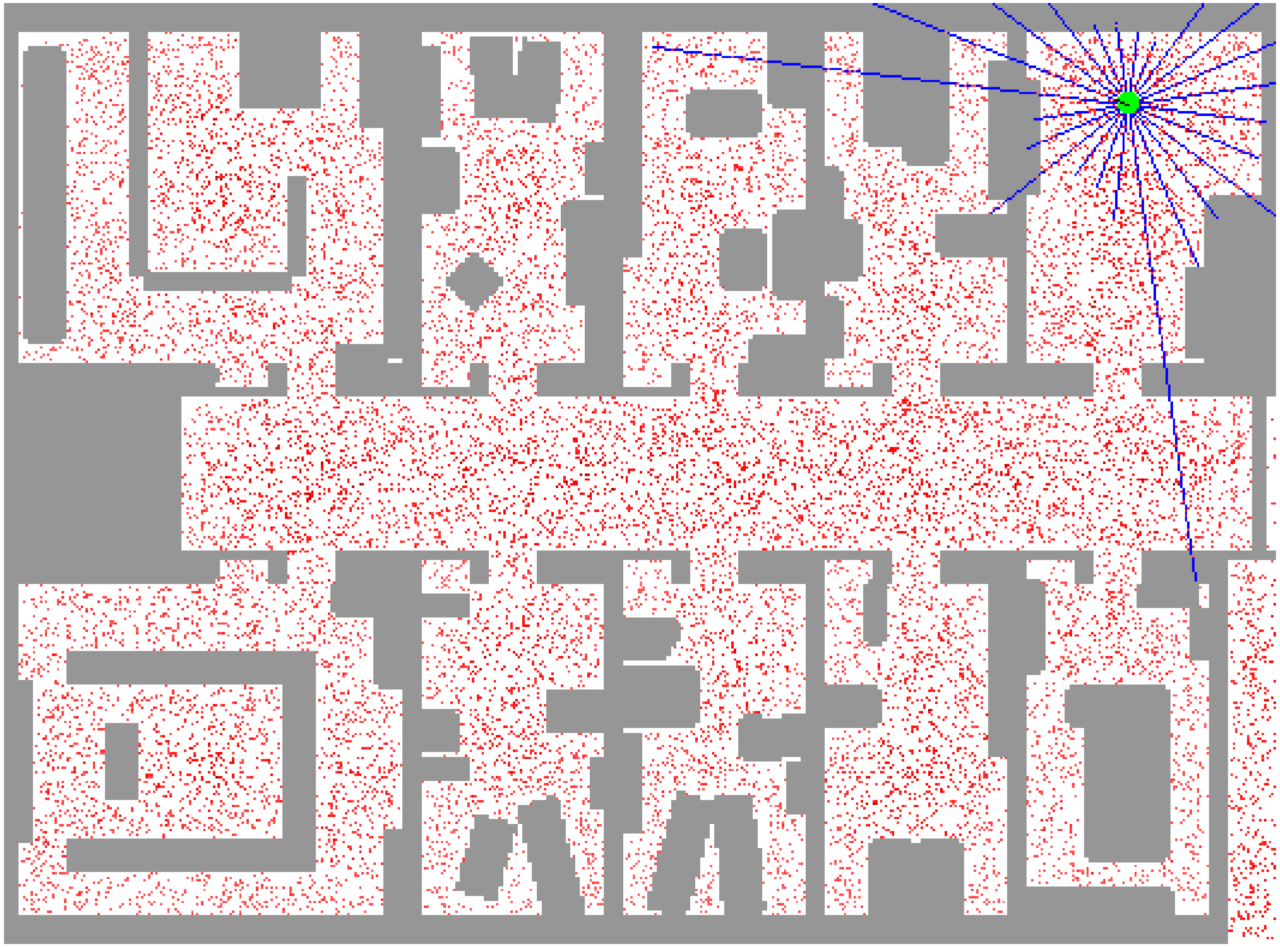


$$p(o \mid s_t, m)$$

Intuition



Putting it together…

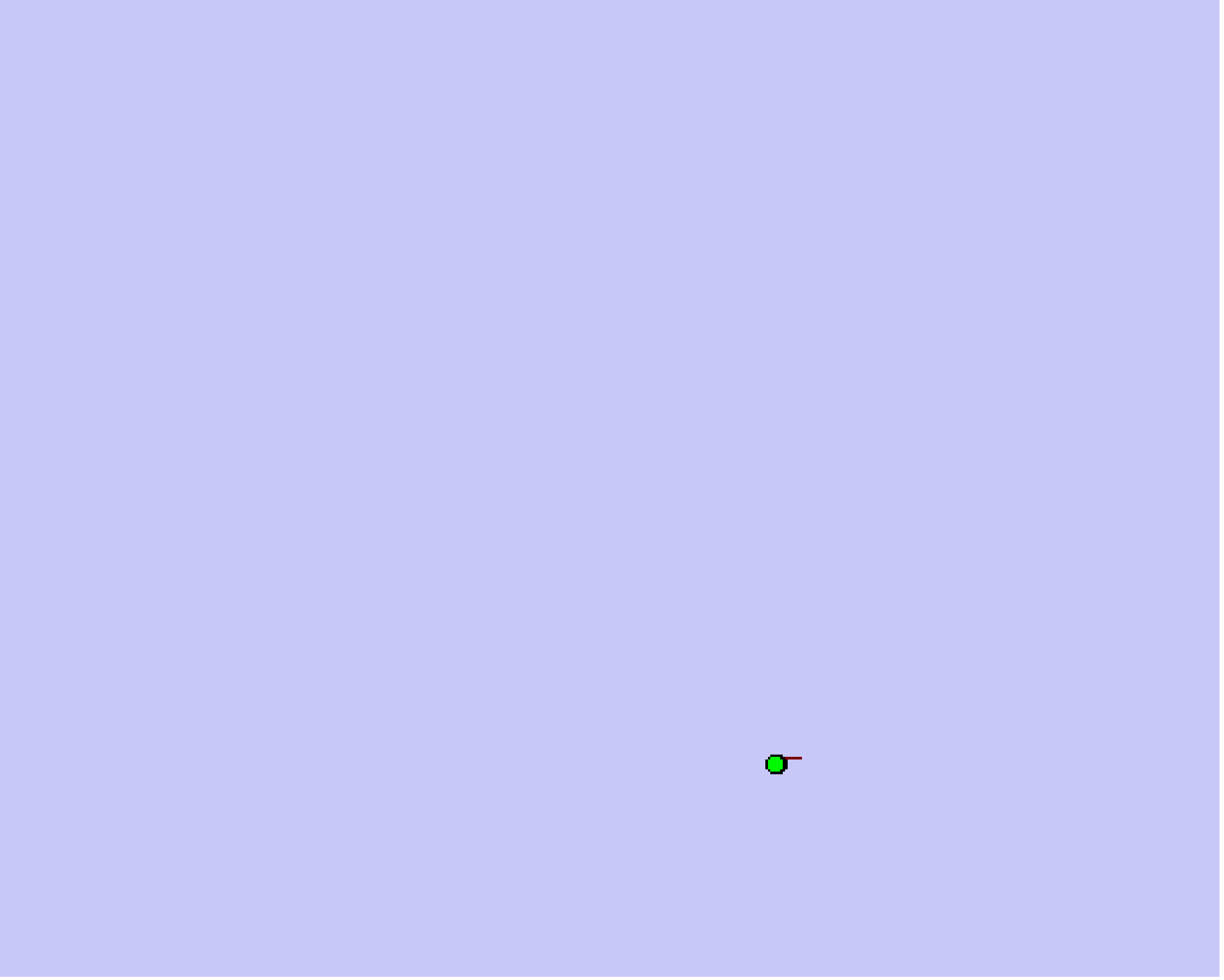


Fox et al. U Washington

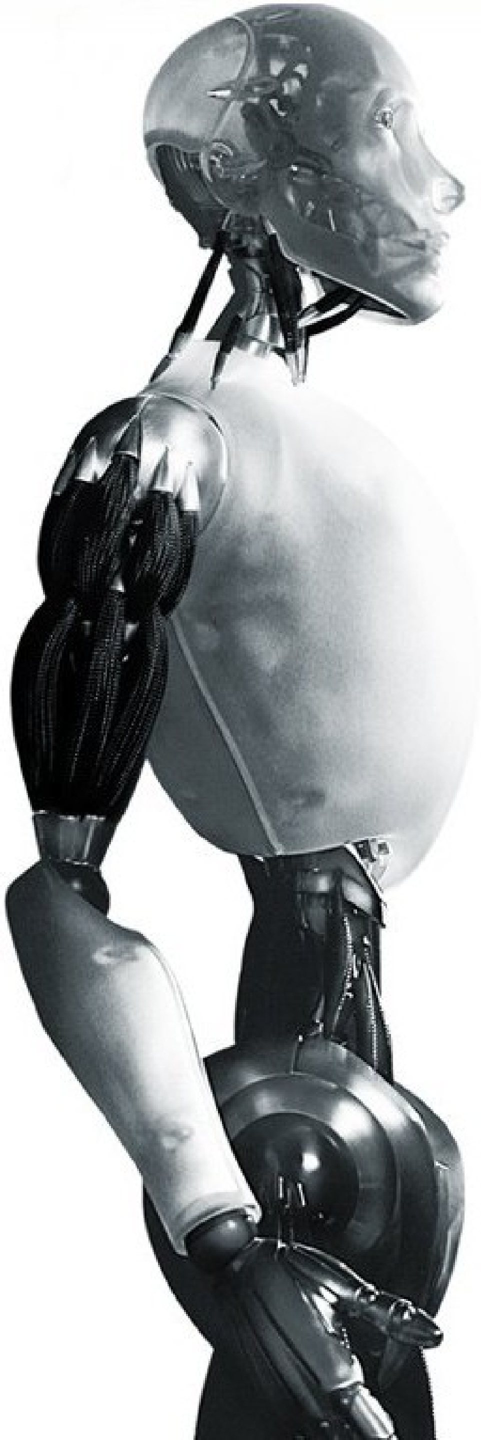
Fox et al. U Washington

Fox et al. U Washington

Fast-SLAM



Fox et al. U Washington



Robotics

Monte Carlo localization (Derivation)

TU Berlin

Oliver Brock

Bayes' Rule

Definition of Conditional Probability

$$\mathbf{P}(A|B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)} \qquad \mathbf{P}(B|A) = \frac{\mathbf{P}(B \cap A)}{\mathbf{P}(A)}$$

Multiplying both sides with denominator

$$\mathbf{P}(A \cap B) = \mathbf{P}(A|B)\,\mathbf{P}(B) \qquad \mathbf{P}(B \cap A) = \mathbf{P}(B|A)\,\mathbf{P}(A)$$

Set intersection is commutative

$$\mathbf{P}(A \cap B) = \mathbf{P}(B \cap A)$$

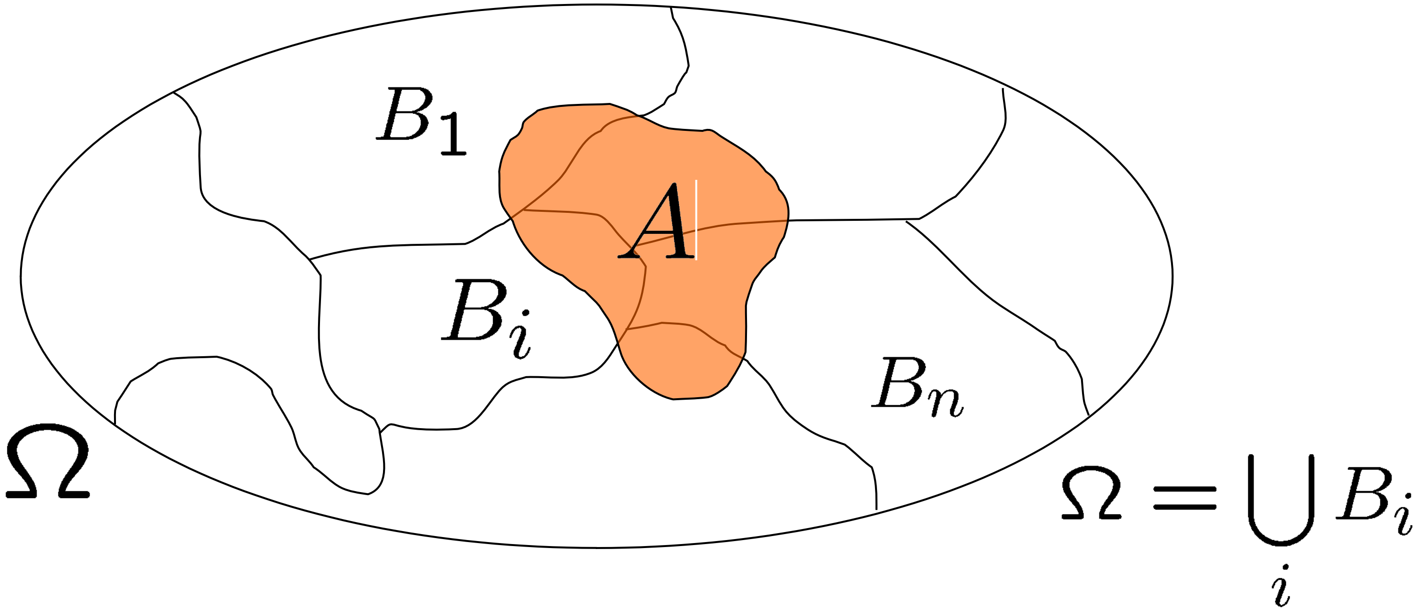
We equate the equations…

$$\mathbf{P}(B)\,\mathbf{P}(A|B) = \mathbf{P}(A)\,\mathbf{P}(B|A)$$

And devide by P(B) to arrive at Bayes' formula

$$\mathbf{P}(A|B) = \frac{\mathbf{P}(A)\,\mathbf{P}(B|A)}{\mathbf{P}(B)}$$

The Law of Total Probability



$$\mathbf{P}(A) = \sum_n \mathbf{P}(A \cap B_n)$$

Terminology

$$b_t(s_t) = p(s_t \mid o_0, \cdots, a_{t-1}, o_t, m)$$

$b_t(s_t)$ is the belief to be at time t in state s_t

o_t is the obersvation at time t

a_t is the action taken at time t

m is the map

Derivation: Step 1

$$
\begin{aligned}
b_t(s_t) &= p(s_t \mid o_0, \cdots, a_{t-1}, o_t, m) \\
&= \mu_t \, p(o_t \mid o_0, \cdots, a_{t-1}, s_t, m) \cdot p(s_t \mid o_0, \cdots, a_{t-1}, m)
\end{aligned}
$$

Using:

$$
\mathbf{P}(A|B) = \frac{\mathbf{P}(A)\,\mathbf{P}(B|A)}{\mathbf{P}(B)}
$$

Derivation: Step 2

$$
\begin{aligned}
b_t(s_t) &= p(s_t \mid o_0, \cdots, a_{t-1}, o_t, m) \\
&= \mu_t \, p(o_t \mid o_0, \cdots, a_{t-1}, s_t, m) \cdot \\
&\quad\; p(s_t \mid o_0, \cdots, a_{t-1}, m) \\
&= \mu_t \, p(o_t \mid s_t, m) \cdot p(s_t \mid o_0, \cdots, a_{t-2}, o_{t-1}, m)
\end{aligned}
$$

Using the *Markov Property*: the observation o_t does only depend on the current state s_t, but not on any states before that.

Derivation: Step 3

$$
\begin{aligned}
b_t(s_t) &= p(s_t \mid o_0, \cdots, a_{t-1}, o_t, m) \\
&= \mu_t \, p(o_t \mid o_0, \cdots, a_{t-1}, s_t, m) \cdot p(s_t \mid o_0, \cdots, a_{t-1}, m) \\
&= \mu_t \, p(o_t \mid s_t, m) \cdot p(s_t \mid o_0, \cdots, a_{t-2}, o_{t-1}, m) \\
&= \mu_t \, p(o_t \mid s_t, m) \cdot \\
&\quad \int p(s_t \mid o_0, \cdots, a_{t-1}, s_{t-1}, m) \, p(s_{t-1} \mid o_0, \cdots, a_{t-1}, m) \, ds_{t-1}
\end{aligned}
$$

Using: $$\mathbf{P}(A) = \sum_n \mathbf{P}(A \cap B_n)$$

Derivation: Step 4

$$
\begin{aligned}
b_t(s_t) &= p(s_t \,|\, o_0, \cdots, a_{t-1}, o_t, m) \\
&= \mu_t \; p(o_t \,|\, o_0, \cdots, a_{t-1}, s_t, m) \cdot p(s_t \,|\, o_0, \cdots, a_{t-1}, m) \\
&= \mu_t \; p(o_t \,|\, s_t, m) \cdot p(s_t \,|\, o_0, \cdots, a_{t-1}, m) \\
&= \mu_t \; p(o_t \,|\, s_t, m) \cdot \\
&\quad \int p(s_t \,|\, o_0, \cdots, a_{t-1}, s_{t-1}, m) \; p(s_{t-1} \,|\, o_0, \cdots, a_{t-2}, o_{t-1} m) \; ds_{t-1} \\
&= \mu_t \; p(o_t \,|\, s_t, m) \cdot \\
&\quad \int p(s_t \,|\, a_{t-1}, s_{t-1}, m) p(s_{t-1} \,|\, o_0, \cdots, o_{t-1}, m) \; ds_{t-1}
\end{aligned}
$$

Using the Markov Property once more.

Derivation: Step 5

$$
\begin{aligned}
b_t(s_t) &= p(s_t \,|\, o_0, \cdots, a_{t-1}, o_t, m) \\
&= \mu_t \; p(o_t \,|\, o_0, \cdots, a_{t-1}, s_t, m) \cdot p(s_t \,|\, o_0, \cdots, a_{t-1}, m) \\
&= \mu_t \; p(o_t \,|\, s_t, m) \cdot p(s_t \,|\, o_0, \cdots, a_{t-1}, m) \\
&= \mu_t \; p(o_t \,|\, s_t, m) \cdot \\
&\quad \int p(s_t \,|\, o_0, \cdots, a_{t-2}, o_{t-1}, s_{t-1}, m) \; p(s_{t-1} \,|\, o_0, \cdots, a_{t-1}, m) \; ds_{t-1} \\
&= \mu_t \; p(o_t \,|\, s_t, m) \cdot \\
&\quad \int p(s_t \,|\, a_{t-1}, s_{t-1}, m) \; p(s_{t-1} \,|\, o_0, \cdots, o_{t-1}, m) \; ds_{t-1} \\
&= \mu_t \; p(o_t \,|\, s_t, m) \cdot \\
&\quad \int p(s_t \,|\, a_{t-1}, s_{t-1}, m) \; b_{t-1}(s_{t-1}) \; ds_{t-1}
\end{aligned}
$$

Towards Implementation

$$b_t(s_t) = \mu_t \, p(o_t \,|\, s_t, m) \cdot \int p(s_t \,|\, a_{t-1}, s_{t-1}, m) \, b_{t-1}(s_{t-1}) \, ds_{t-1}$$

Sensor Model

$$b_t(s_t) = \mu_t \boxed{p(o_t \,|\, s_t, m)} \cdot \sum_{\text{samples}} \boxed{p(s_t \,|\, a_{t-1}, s_{t-1}, m)} \, \boxed{b_{t-1}(s_{t-1})}$$

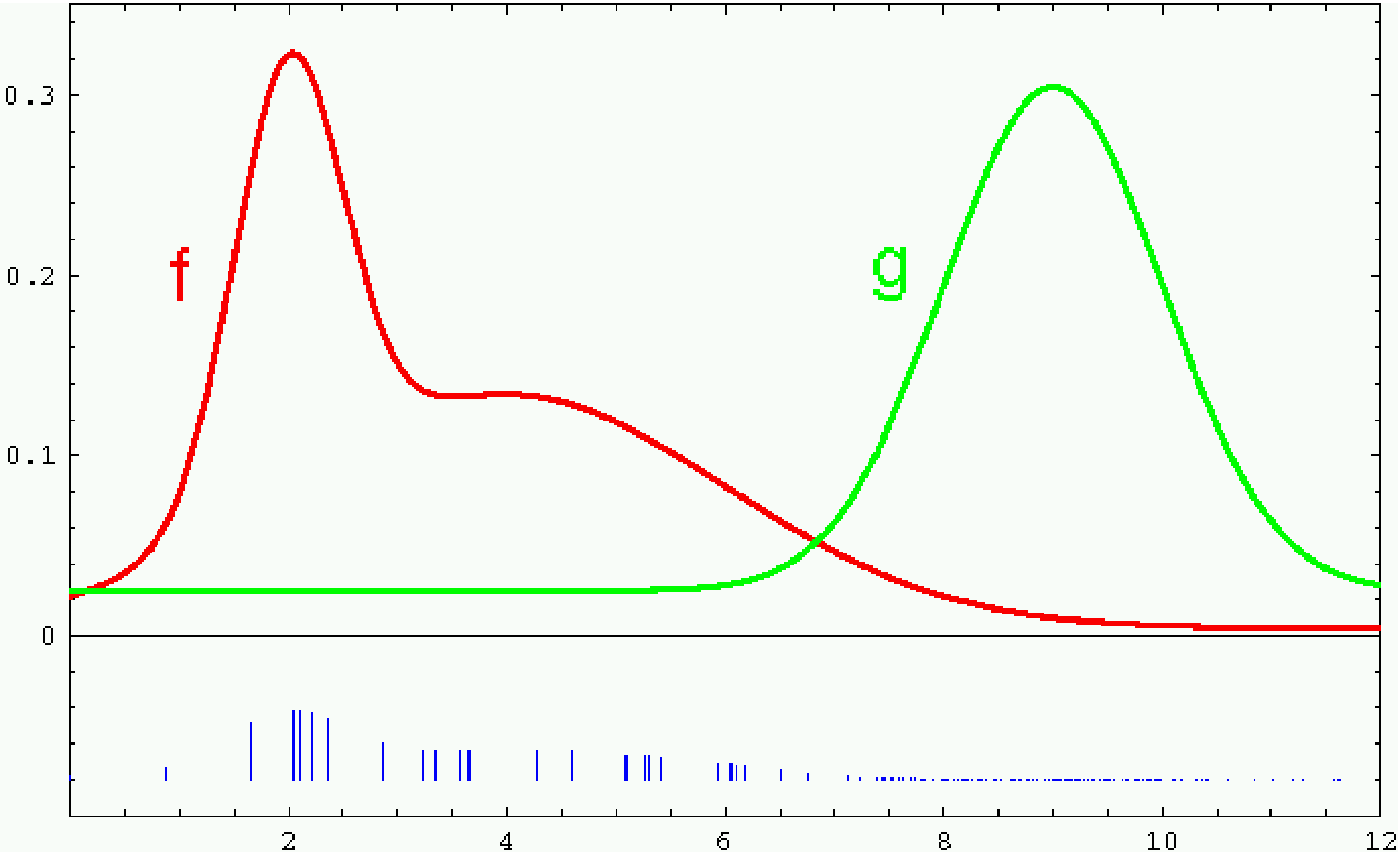
Motion model

Importance factor of sample s_{t-1}

Particle Filter: one iteration

1. draw random sample s_{t-1} from $b_{t-1}(s_{t-1})$ with
2. for s_{t-1} create s_t based on $p(s_t \mid a_{t-1}, s_{t-1}, m)$
3. compute importance factor $p(o_t \mid s_t, m)$
4. repeat 1-3 for n samples
5. normalize importance factors to sum to 1
6. Resample to get b_t

Key: Importance Sampling



Weight samples: $w = f / g$

Sampling for expected value of a function relative to a distribution

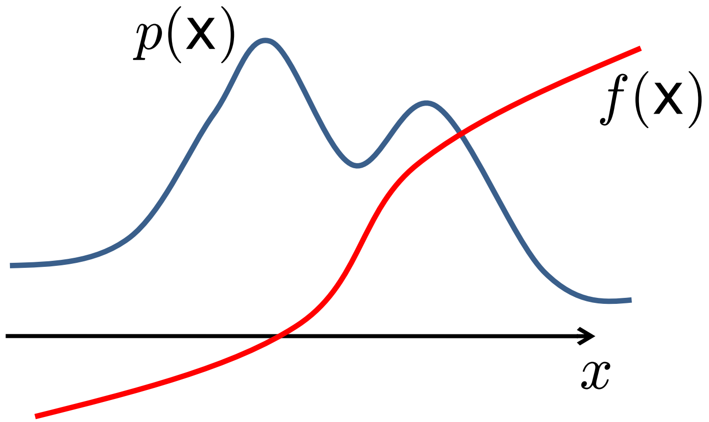
$$\mathbb{E}[f] = \int f(\mathbf{z})p(\mathbf{z})d\mathbf{z}$$

$$\hat{f} = \frac{1}{L}\sum_{l=1}^{L} f(\mathbf{z}^{(l)})$$

$z^{(l)}$ are drawn from $p(z)$

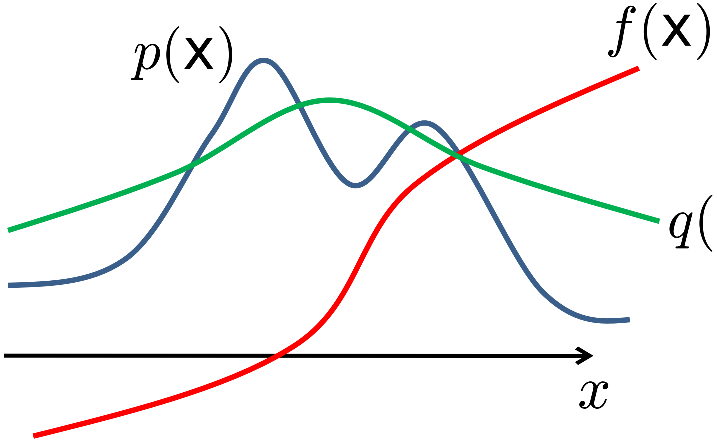
$$\mathbb{E}[f] \simeq \sum_{l=1}^{L} f(\mathbf{z}^{(l)})p(\mathbf{z}^{(l)})$$

$z^{(l)}$ are drawn from regular grid



Importance Sampling

$$\begin{aligned}\mathbb{E}[f] &= \int f(\mathbf{z})p(\mathbf{z})d\mathbf{z} \\ &= \int f(\mathbf{z})\frac{p(\mathbf{z})}{q(\mathbf{z})}q(\mathbf{z})d\mathbf{z} \\ &\simeq \frac{1}{L}\sum_{l=1}^{L} \underbrace{\frac{p(\mathbf{z}^{(l)})}{q(\mathbf{z}^{(l)})}}_{\text{importance weight}} f(\mathbf{z}^{(l)})\end{aligned}$$



Bayes Filter

Bayes Filter $(b(s_{t-1}), a_{t-1}, o_t)$

for all s_t do

prediction $\boxed{b'(s_t) = \int p(s_t \,|\, a_{t-1}, s_{t-1}) b(s_{t-1}) ds_{t-1}}$

measurement update $\boxed{b(s_t) = \mu \, p(o_t \,|\, s_t) \cdot b'(s_t)}$

end for

return $b(s_t)$

Nonparametric: Discrete Bayes

Discrete Bayes Filter $(\{p_{k,t-1}\}, a_{t-1}, o_t)$

for all k do

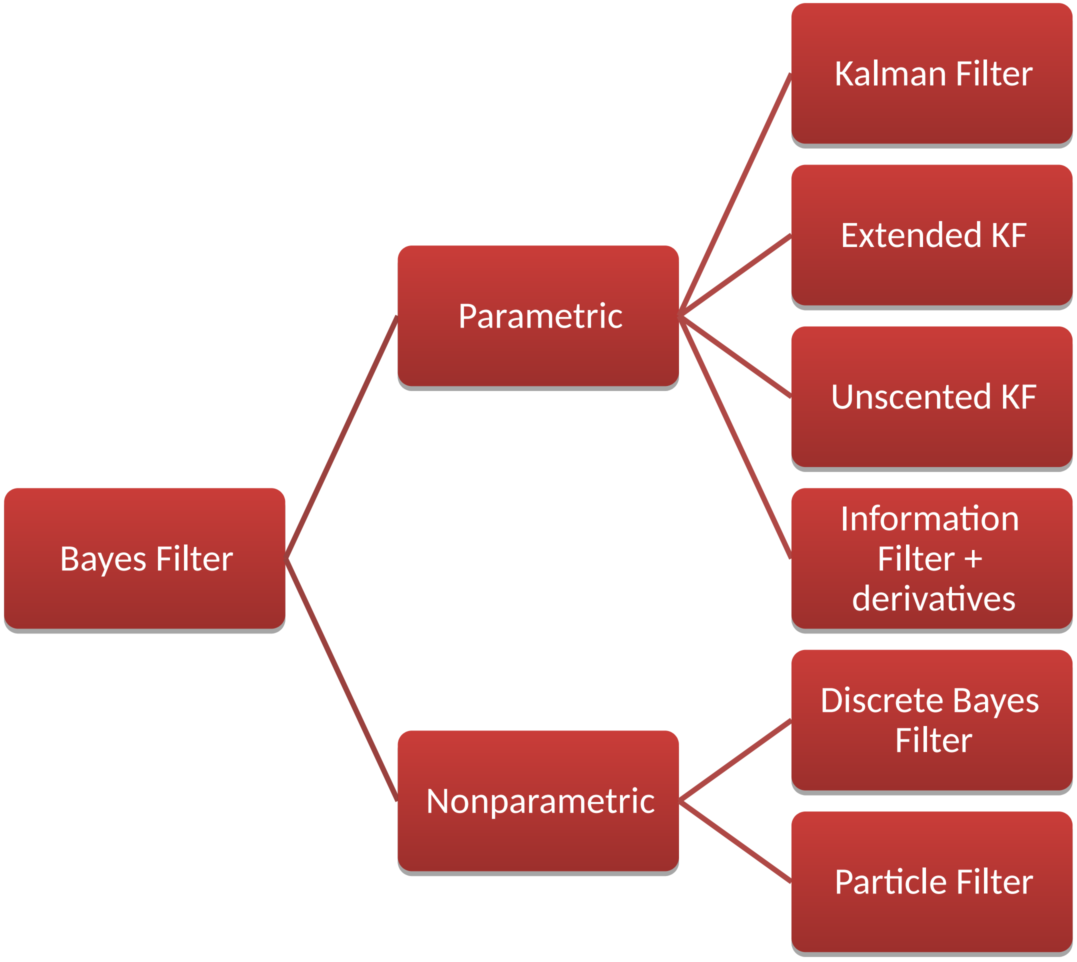
$$p'_{k,t} = \sum_i p(X_t = x_k \,|\, a_{t-1}, X_{t-1} = x_i) p_{i,t-1}$$

$$p_{k,t} = \mu \, p(o_t \,|\, X_t = x_k) \boxed{\cdot \, p'_{k t-1}}$$

end for

return $\{p_{k,t}\}$

A family of methods



Gaussian Bayesian Filters

- Kalman Filter
  - linear update of states based on action
    $$p(s_t|a_{t-1}, s_{t-1})$$
  - linear sensor model
    $$p(o_t|s_t)$$
  - belief can be described by a normal distribution
  - computationally efficient and elegant

Extended Kalman Filter (EKF)

- state transition probability non required to be linear any more
- neither is the measurement probability
- linearization via Taylor expansion

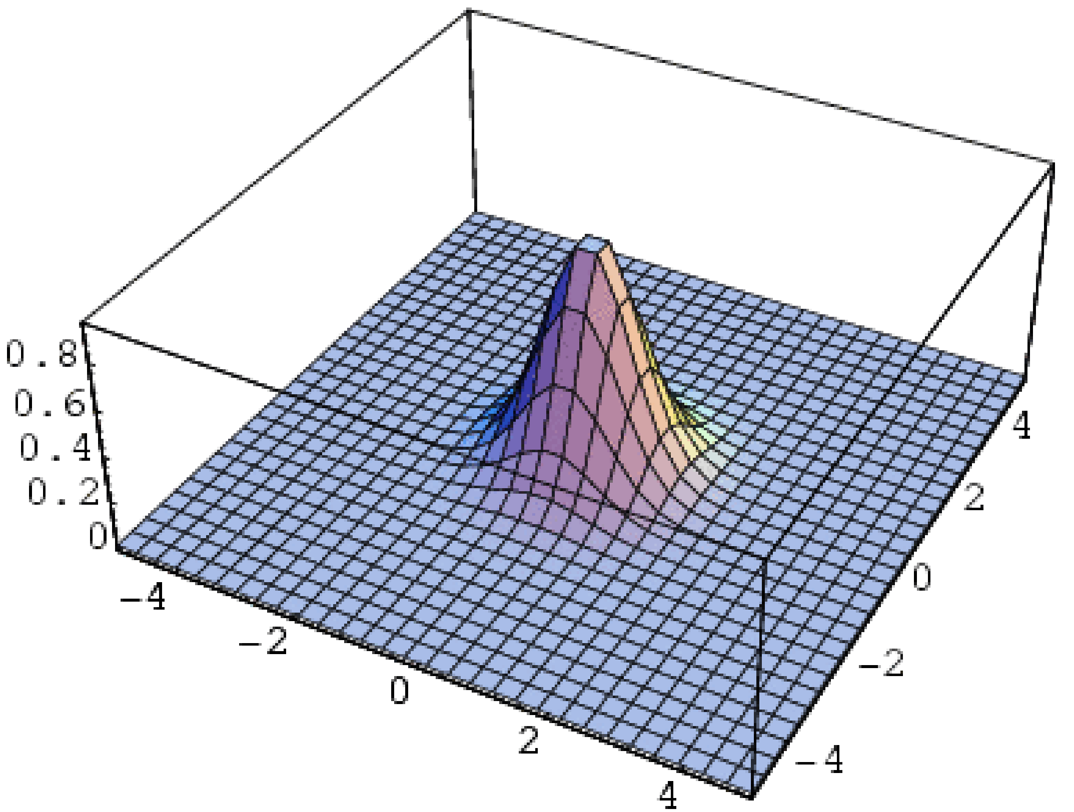
Unscented Kalman Filter (UKF)

- linearization through linear regression

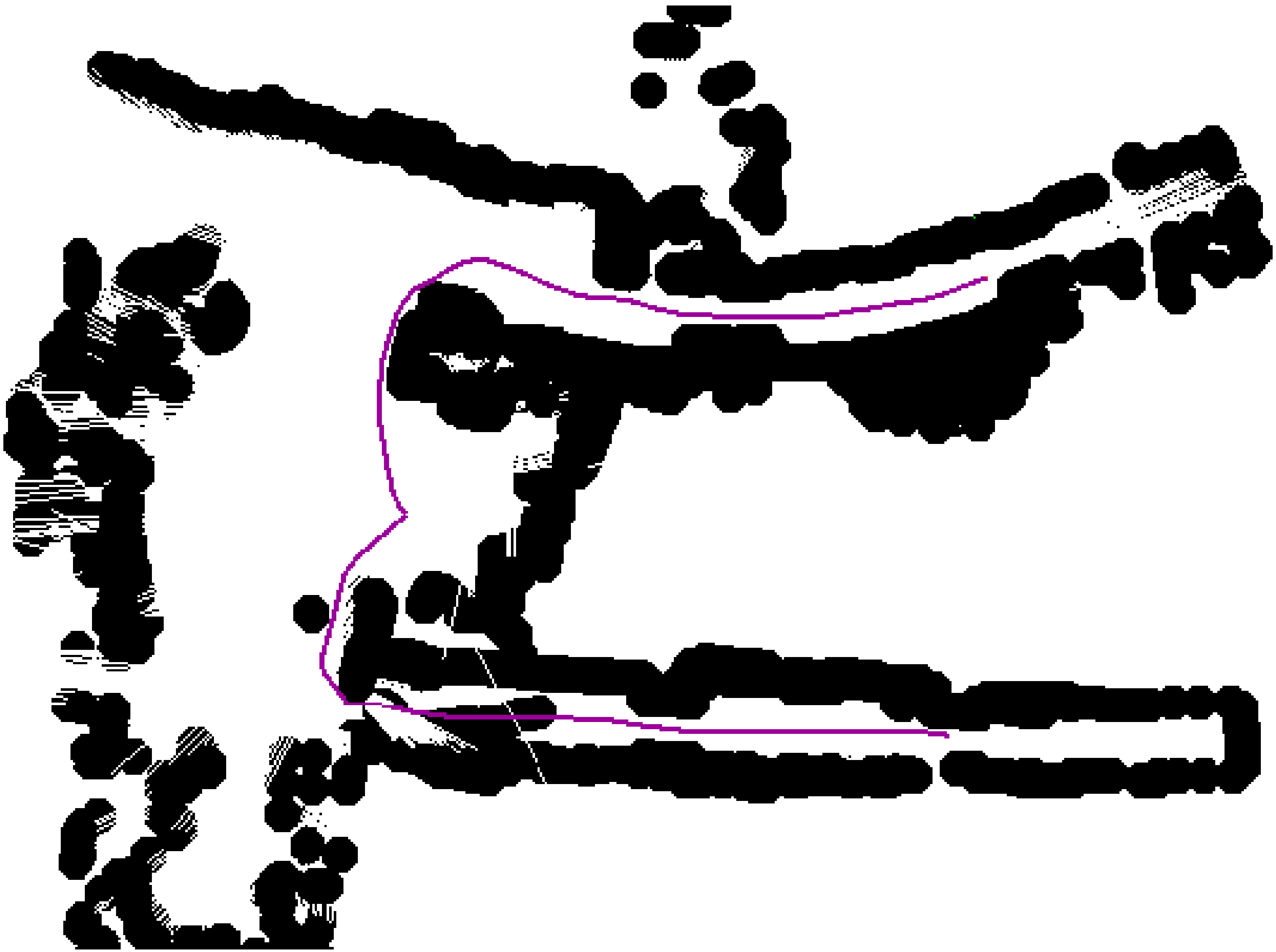
Information Filter

- Dual of KF
- Also Extended Information Filter

Histogram Filter

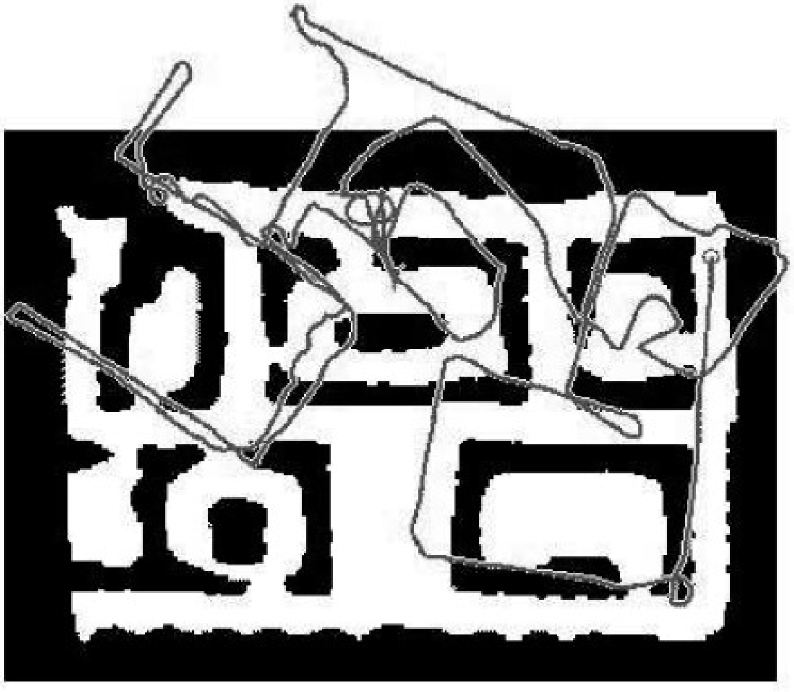


Occupancy Grid

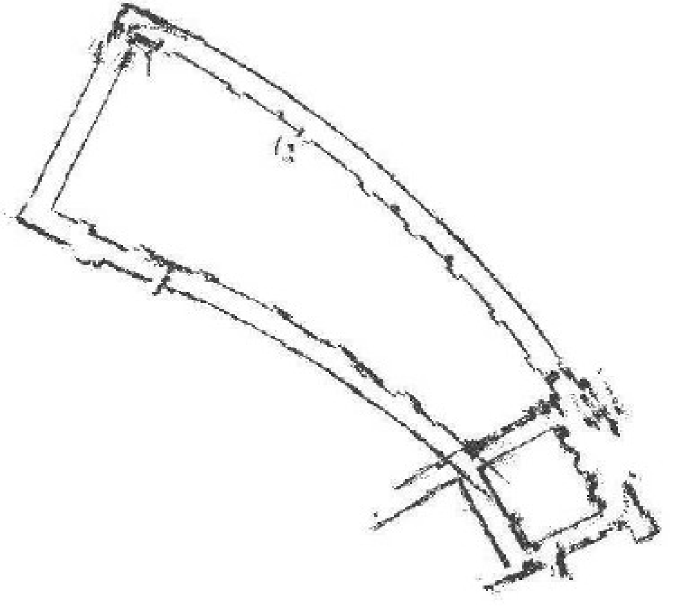


Odometry Error

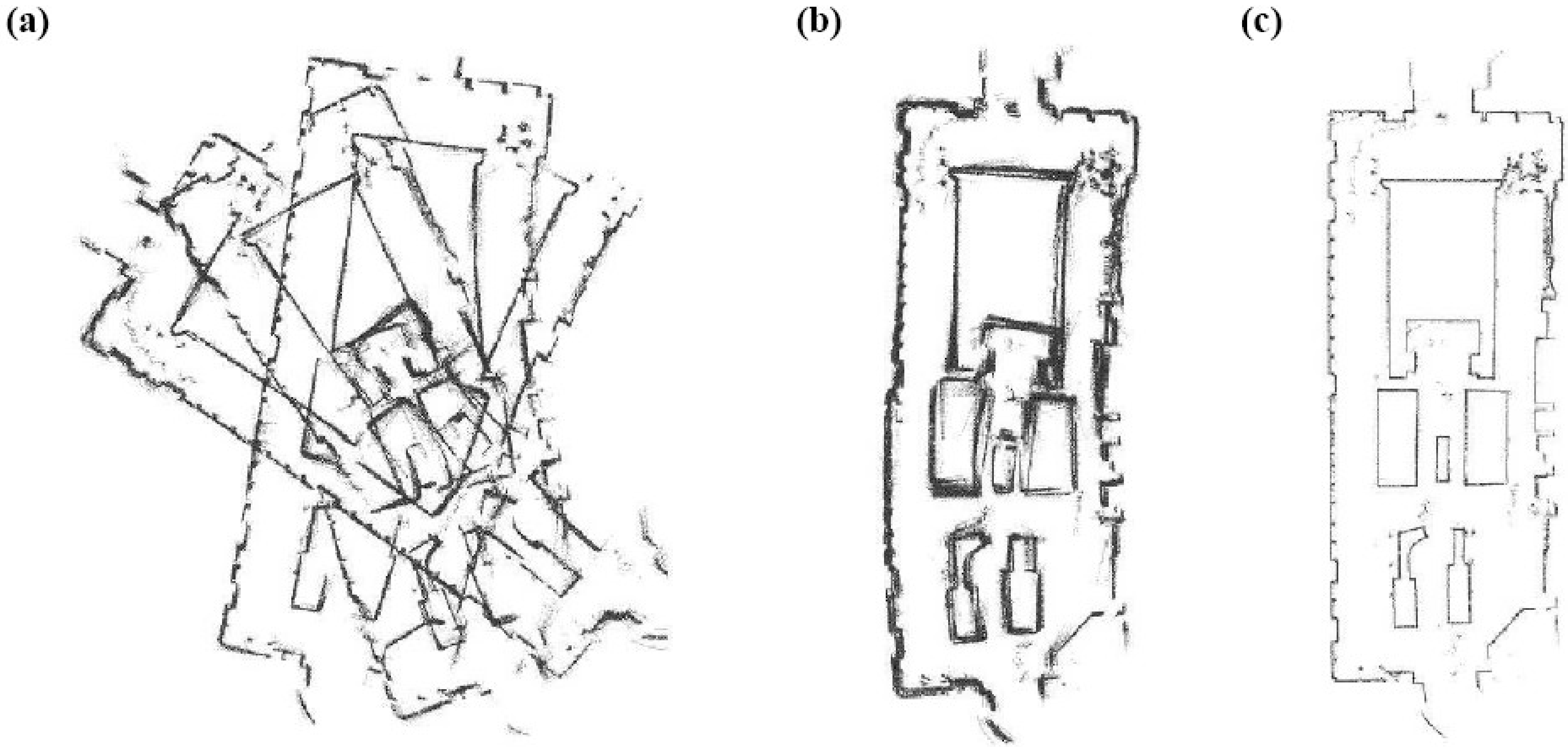
(a)



(b)



What we can do in spite of it…

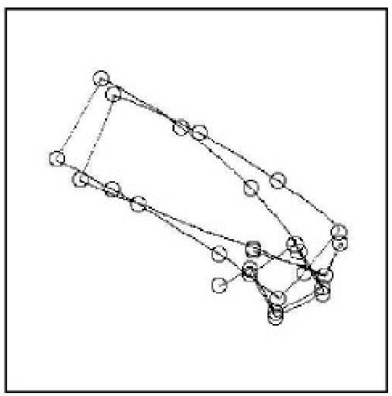


EM

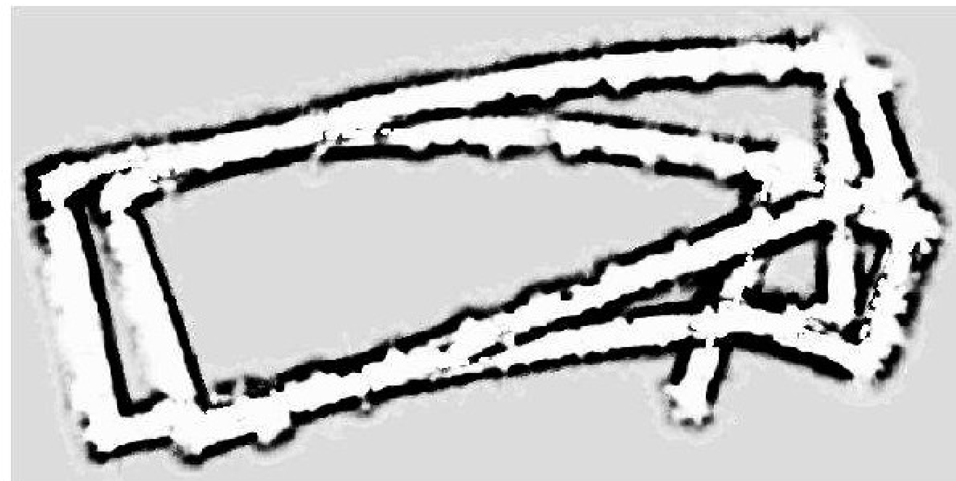
More advanced “EM”

Expectation Maximization (EM)

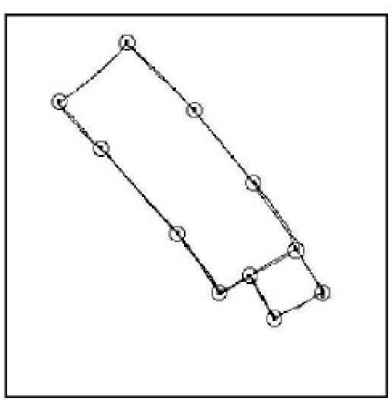
(a)



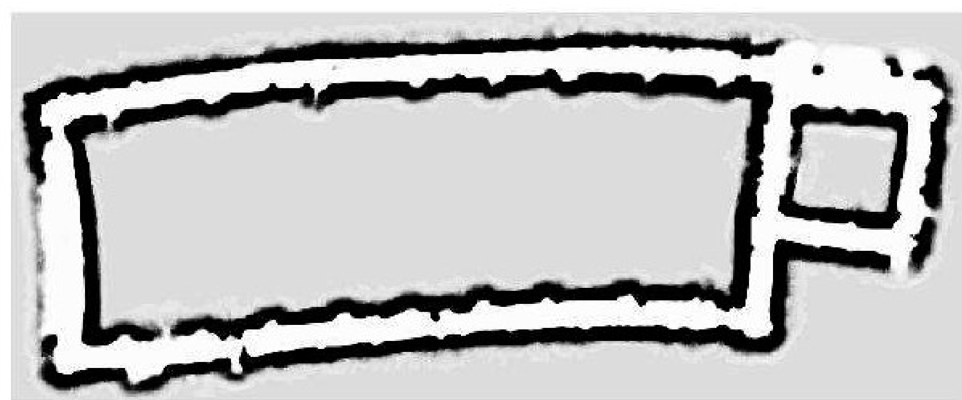
(b)



(c)



(d)



Map and Blueprint

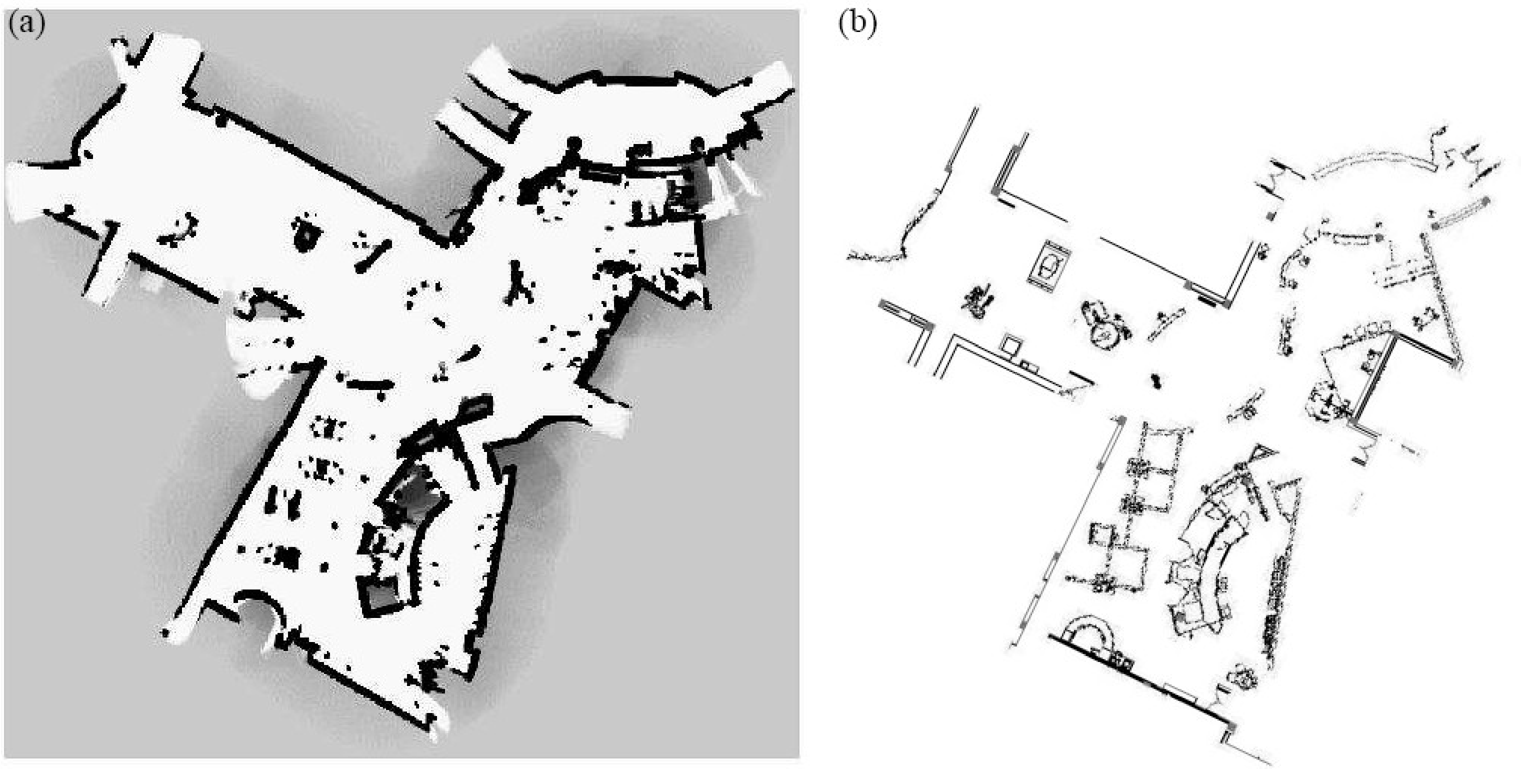


Figure 9: (a) Occupancy grid map and (b) architectural blueprint of a recently constructed building. The blueprint is less accurate than the map in several locations.

SLAM

- Simultaneous Localization and Mapping
- Size of hypothesis space
- Chicken and egg problem

$$p(s_t, m \mid a_{1:t}, o_{1:t})$$

$$p(s_{1:t}, m \mid a_{1:t}, o_{1:t})$$