Decision Making
in Robots and Autonomous Agents

Probabilistic methods: Filtering and Exploration
(Following S. Thrun et al, Probabilistic Robotics)

Subramanian Ramamoorthy
School of Informatics

1 March 2019

Probabilistic Robotics

Key idea:
Explicit representation of uncertainty using the calculus of probability theory

Perception	= state estimation
Action	= utility optimization

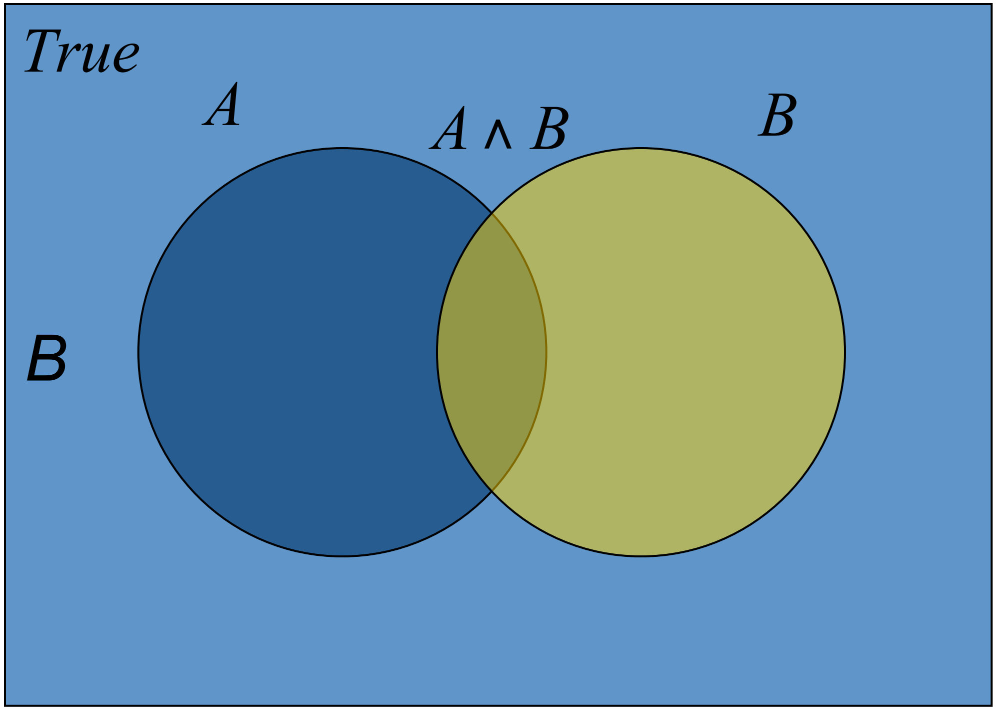
Axioms of Probability Theory

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

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

A Closer Look at Axiom 3

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



Using the Axioms

$$
\begin{aligned}
\Pr(A \vee \neg A) &= \Pr(A) + \Pr(\neg A) - \Pr(A \wedge \neg A) \\
\Pr(True) &= \Pr(A) + \Pr(\neg A) - \Pr(False) \\
1 &= \Pr(A) + \Pr(\neg A) - 0 \\
\Pr(\neg A) &= 1 - \Pr(A)
\end{aligned}
$$

Bayes Formula

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

$$\Rightarrow$$

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

Normalization

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

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

Algorithm:

$$\forall x : \mathrm{aux}_{x|y} = P(y \mid x)\ P(x)$$

$$\eta = \frac{1}{\sum_x \mathrm{aux}_{x|y}}$$

$$\forall x : P(x \mid y) = \eta\ \mathrm{aux}_{x|y}$$

Conditioning

- Law of total probability:

$$P(x) = \int P(x, z) dz$$

$$P(x) = \int P(x \mid z) P(z) dz$$

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

Bayes Rule with Background Knowledge

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

Conditioning

- Total probability:

$$P(x) = \int P(x, z)dz$$

$$P(x) = \int P(x \mid z)P(z)dz$$

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

Conditional Independence

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

equivalent to

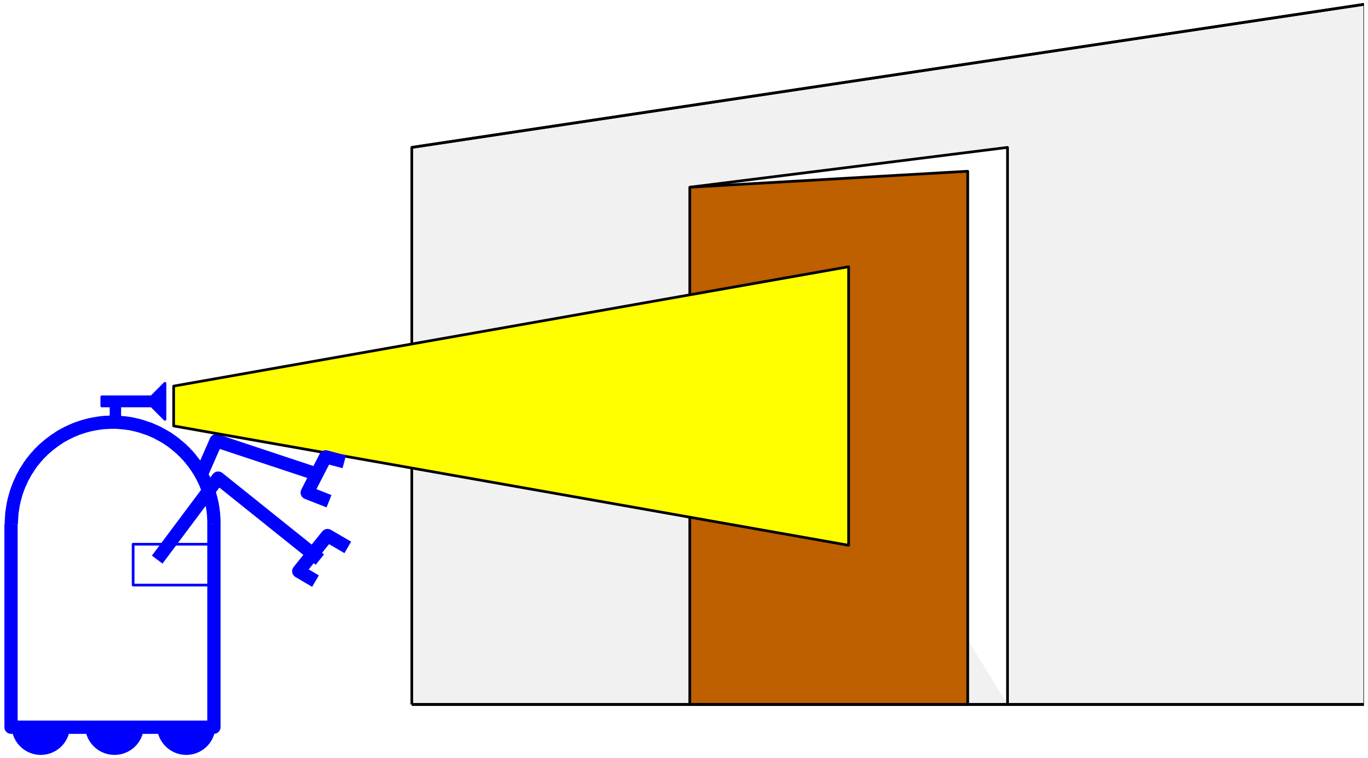
$$P(x \mid z) = P(x \mid z, y)$$

and

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

Simple Example of State Estimation

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



Causal vs. Diagnostic Reasoning

count frequencies!

- *P(open|z)* is diagnostic.
- *P(z|open)* is causal.
- Often causal knowledge is easier to obtain.
- Bayes rule allows us to use causal knowledge:

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

Example

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

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

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

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

Combining Evidence

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

Recursive Bayesian Updating

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

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

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

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

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

Example: Second Measurement

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

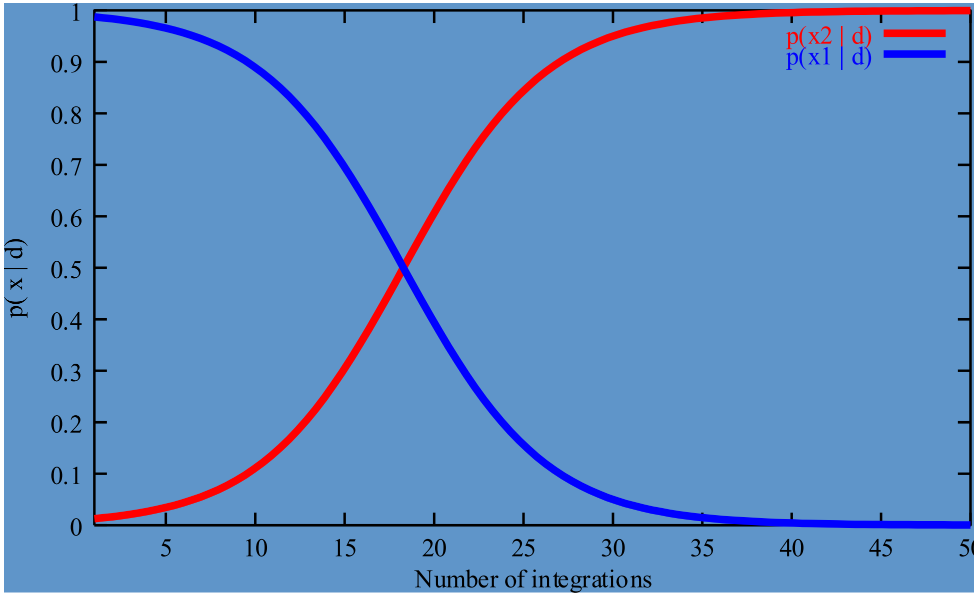
$$P(open \mid z_2, z_1) = \frac{P(z_2 \mid open)\, P(open \mid z_1)}{P(z_2 \mid open)\, P(open \mid z_1) + P(z_2 \mid \neg open)\, P(\neg open \mid z_1)}$$

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

- z_2 lowers the probability that the door is open.

A Typical Pitfall

- Two possible locations x_1 and x_2
- $P(x_1)=0.99$
- $P(z|x_2)=0.09$ $P(z|x_1)=0.07$



Actions

- Often the world is **dynamic** since
  - **actions carried out by the robot**,
  - **actions carried out by other agents**,
  - or just the **time** passing by

  change the world.

- How can we **incorporate** such **actions**?

Typical Actions

- The robot **turns its wheels** to move
- The robot **uses its manipulator** to grasp an object
- Plants grow over **time**…

- Actions are **never carried out with absolute certainty**.
- In contrast to measurements, **actions generally increase the uncertainty**.

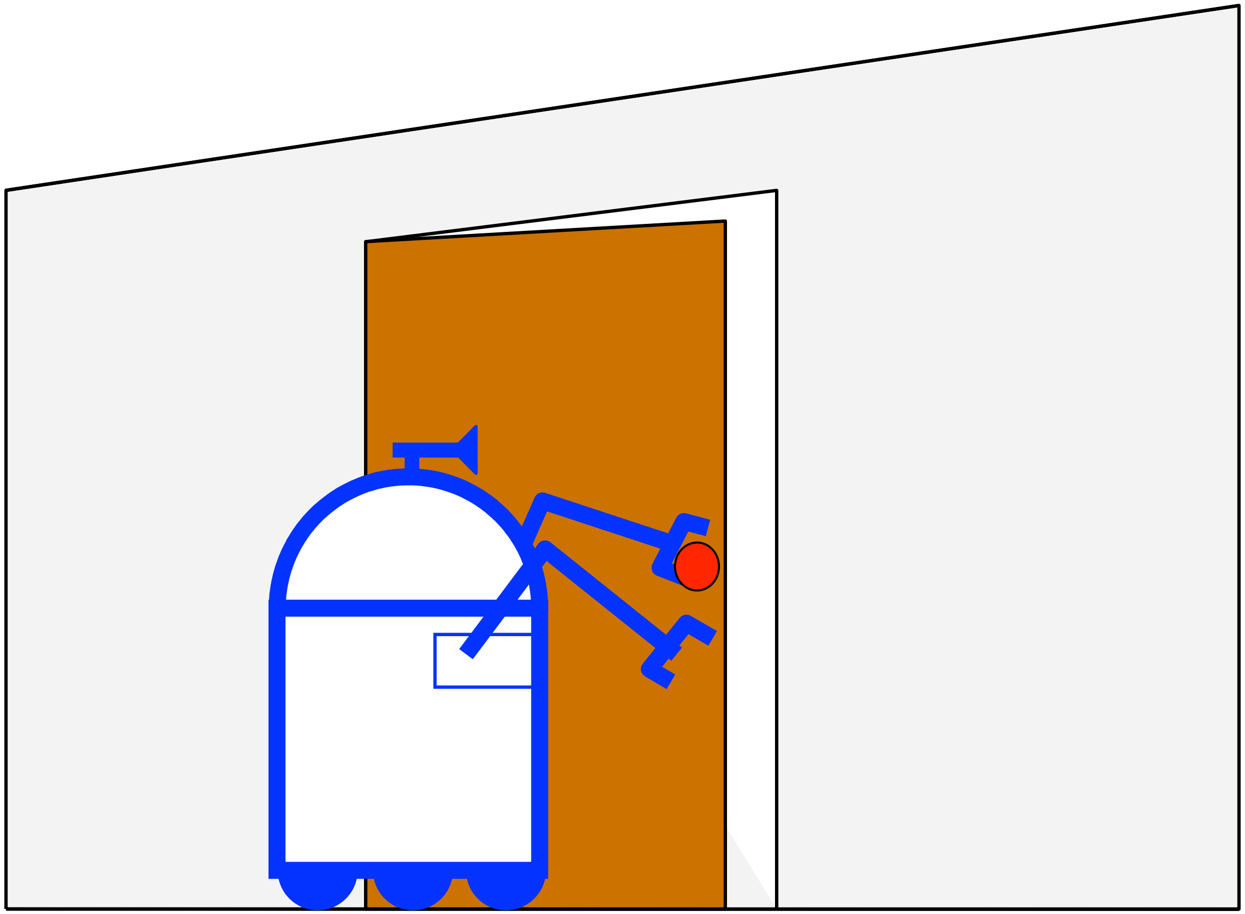
Modeling Actions

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

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

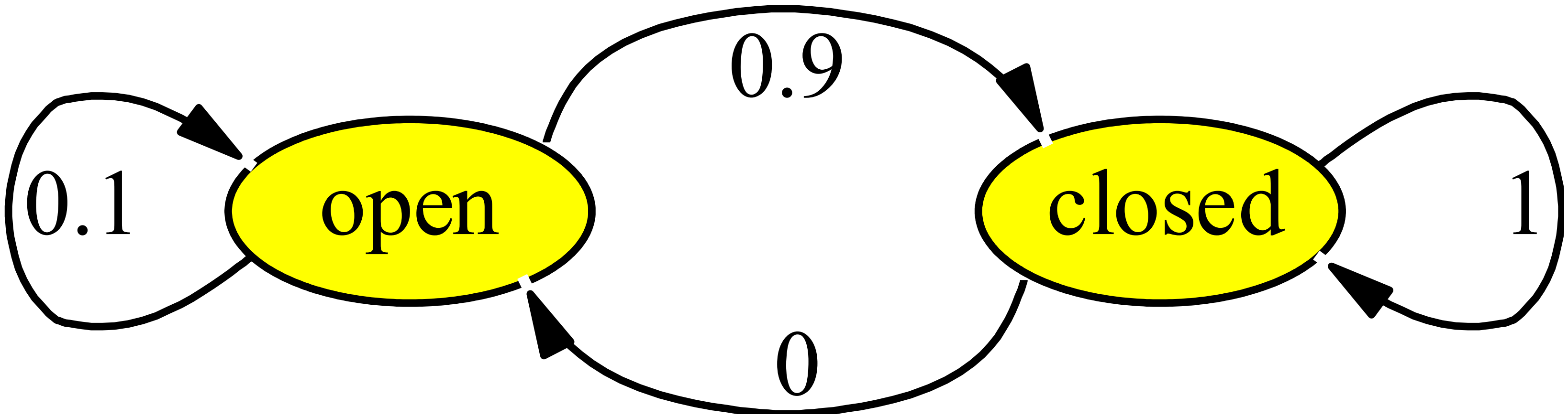
- This term specifies the pdf that **executing *u* changes the state from *x'* to *x*.**

Example: Closing the door



State Transitions

$P(x|u,x')$ for u = "close door":



If the door is open, the action "close door" succeeds in 90% of all cases.

Integrating the Outcome of Actions

Continuous case:

$$P(x \mid u) = \int P(x \mid u, x')P(x')\,dx'$$

Discrete case:

$$P(x \mid u) = \sum P(x \mid u, x')P(x')$$

Example: The Resulting Belief

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

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

Bayes Filters: Framework

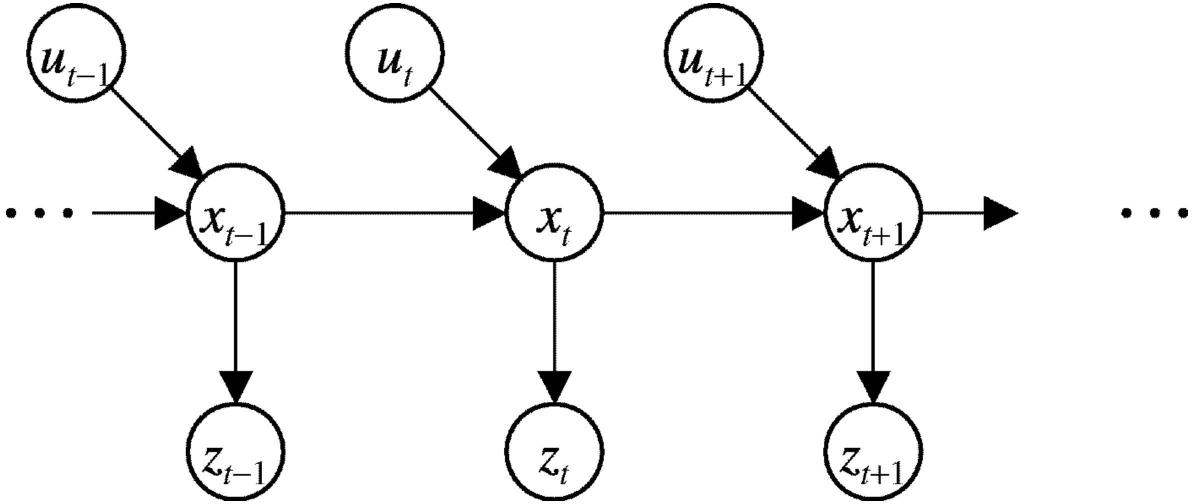
- **Given:**
  - Stream of observations *z* and action data *u:*

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

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

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

Markov Assumption



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

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

Underlying Assumptions

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

z = observation
u = action
x = state

Bayes Filters

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

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

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

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

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

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

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

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

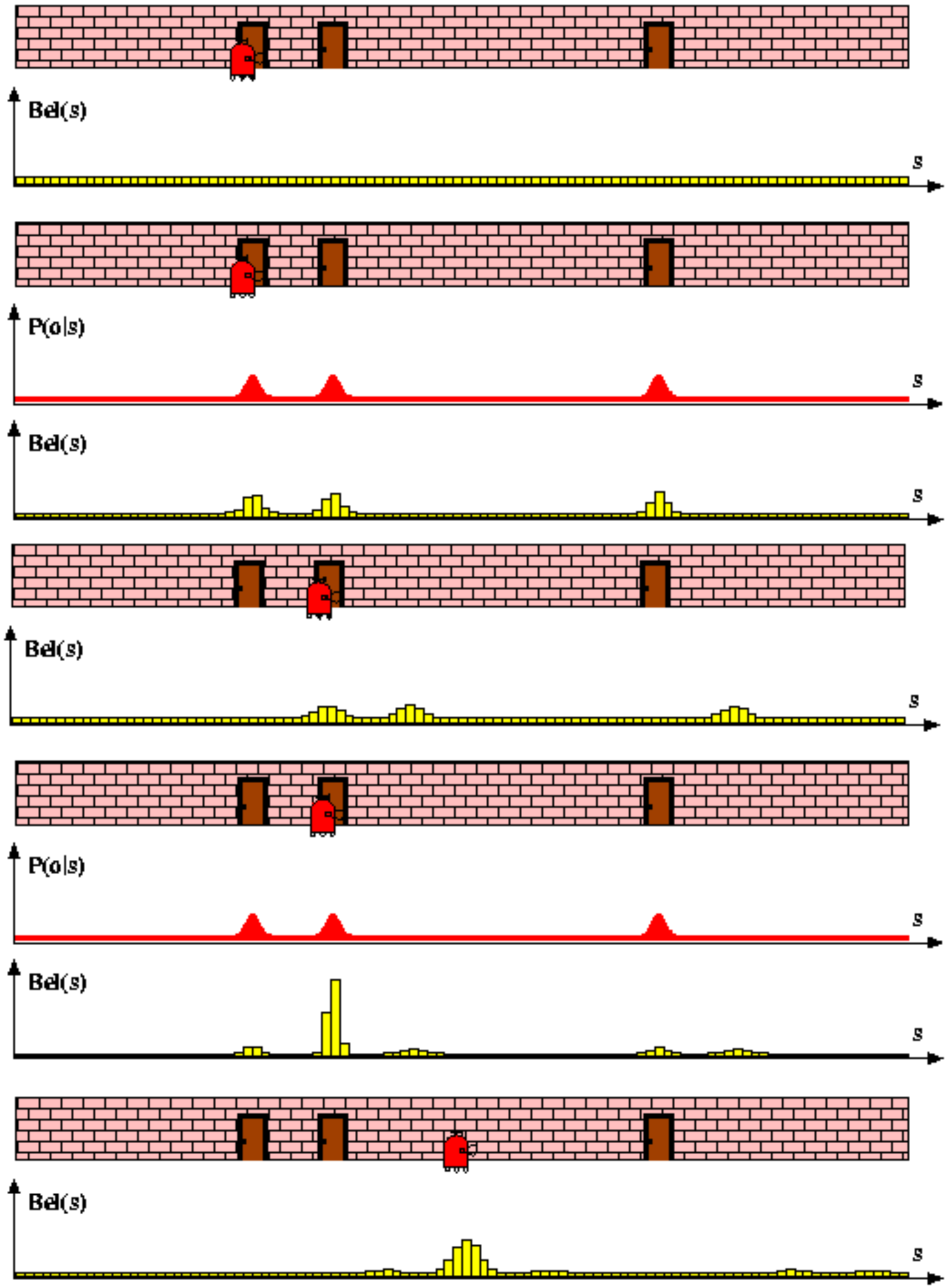
1. Algorithm **Bayes_filter**(*Bel(x),d*):
2. η=0
3. If *d* is a perceptual data item *z* then
4. For all *x* do
5. $Bel'(x) = P(z \mid x) Bel(x)$
6. $\eta = \eta + Bel'(x)$
7. For all *x* do
8. $Bel'(x) = \eta^{-1} Bel'(x)$
9. Else if *d* is an action data item *u* then
10. For all *x* do
11. $Bel'(x) = \int P(x \mid u, x')\ Bel(x')\ dx'$
12. Return *Bel'(x)*

Bayes Filters as a Family of Algorithms

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

- Kalman filters
- Particle filters
- Hidden Markov models
- Dynamic Bayesian networks
- Partially Observable Markov Decision Processes (POMDPs)

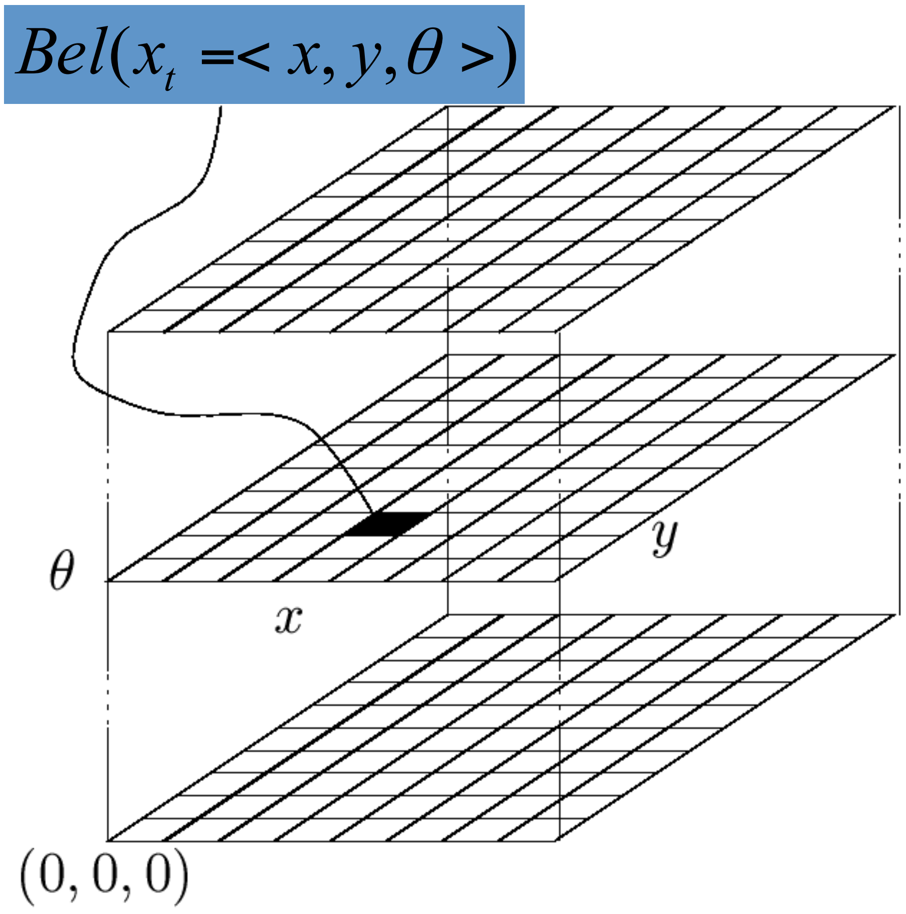
Discrete filter: Piecewise Constant case



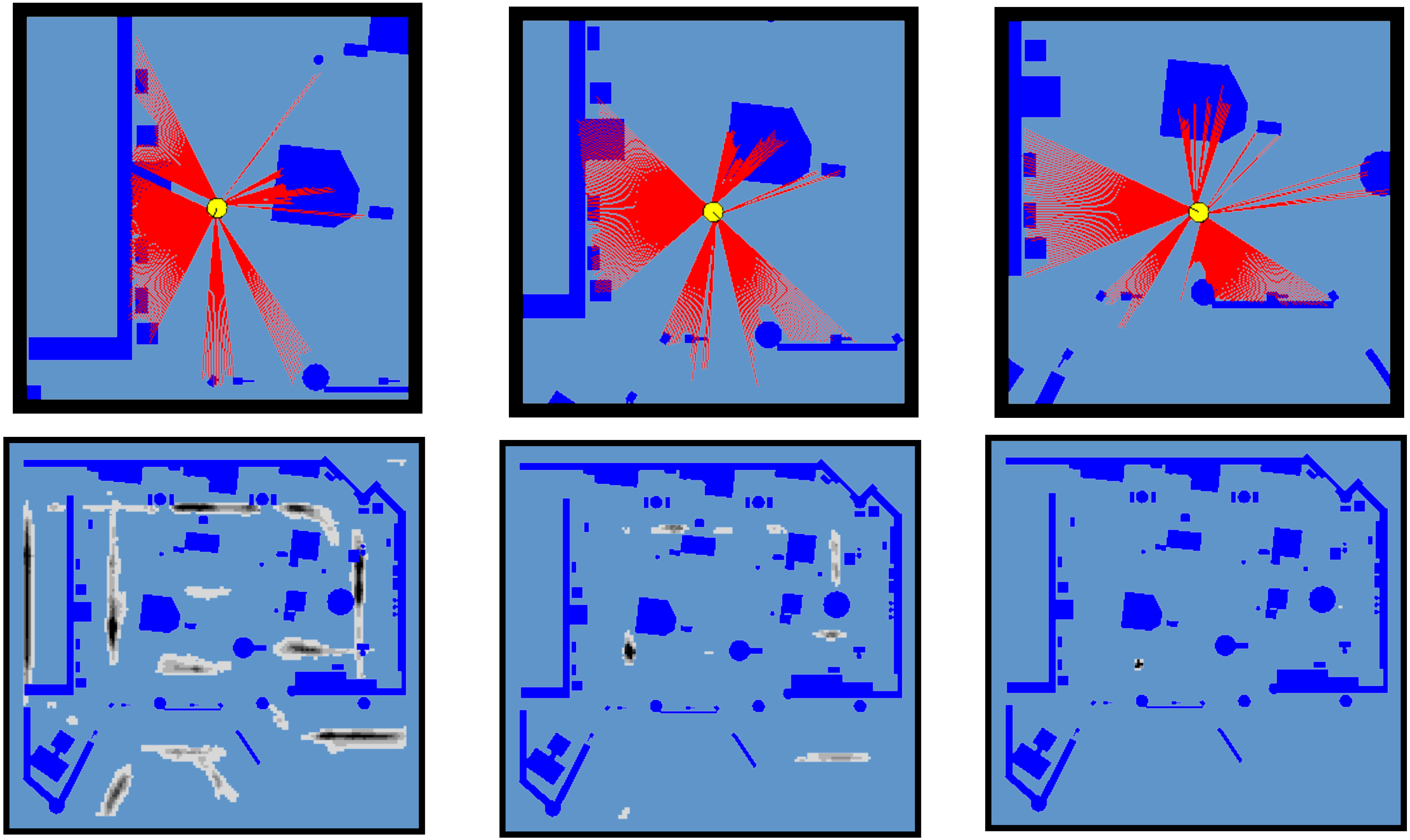
Discrete Bayes Filter Algorithm

1. Algorithm **Discrete_Bayes_filter**(*Bel(x),d*):
2. η=0
3. If *d* is a perceptual data item *z* then
4.   For all *x* do
5. $$Bel'(x) = P(z \mid x) Bel(x)$$
6. $$\eta = \eta + Bel'(x)$$
7.   For all *x* do
8. $$Bel'(x) = \eta^{-1} Bel'(x)$$
9. Else if *d* is an action data item *u* then
10.   For all *x* do
11. $$Bel'(x) = \sum_{x'} P(x \mid u, x')\, Bel(x')$$
12. Return *Bel'(x)*

Piecewise Constant Representation

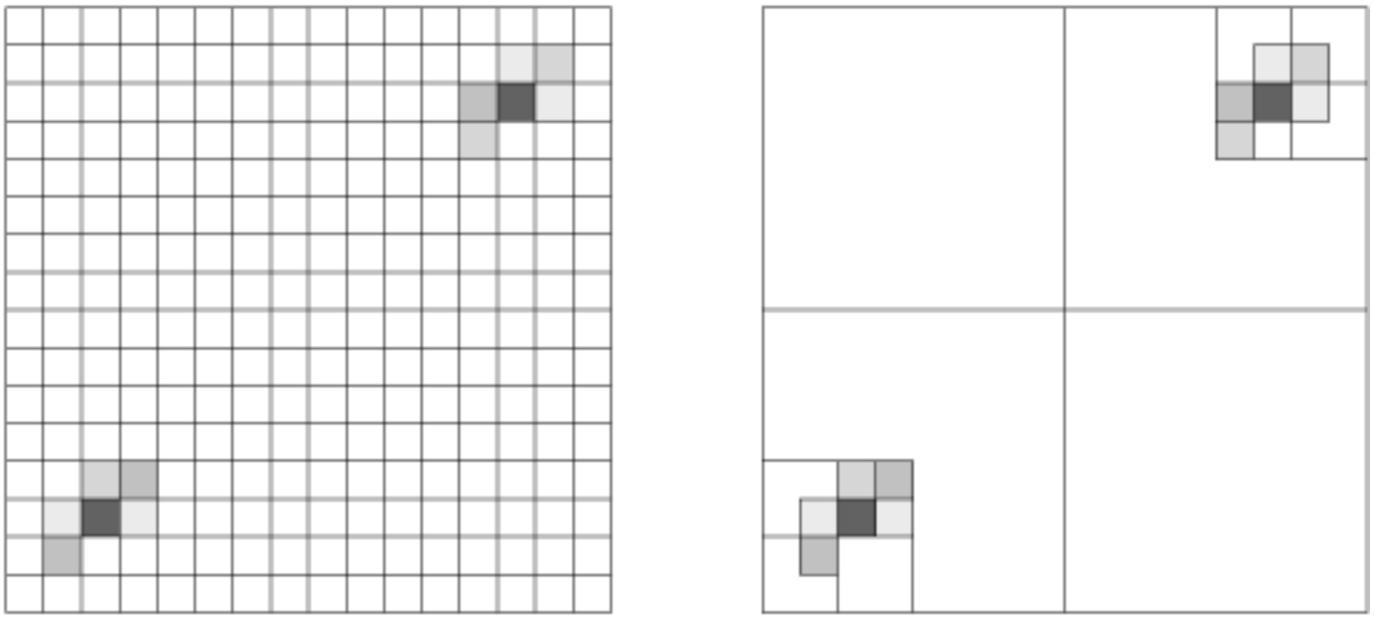


Example: Grid-based Localization



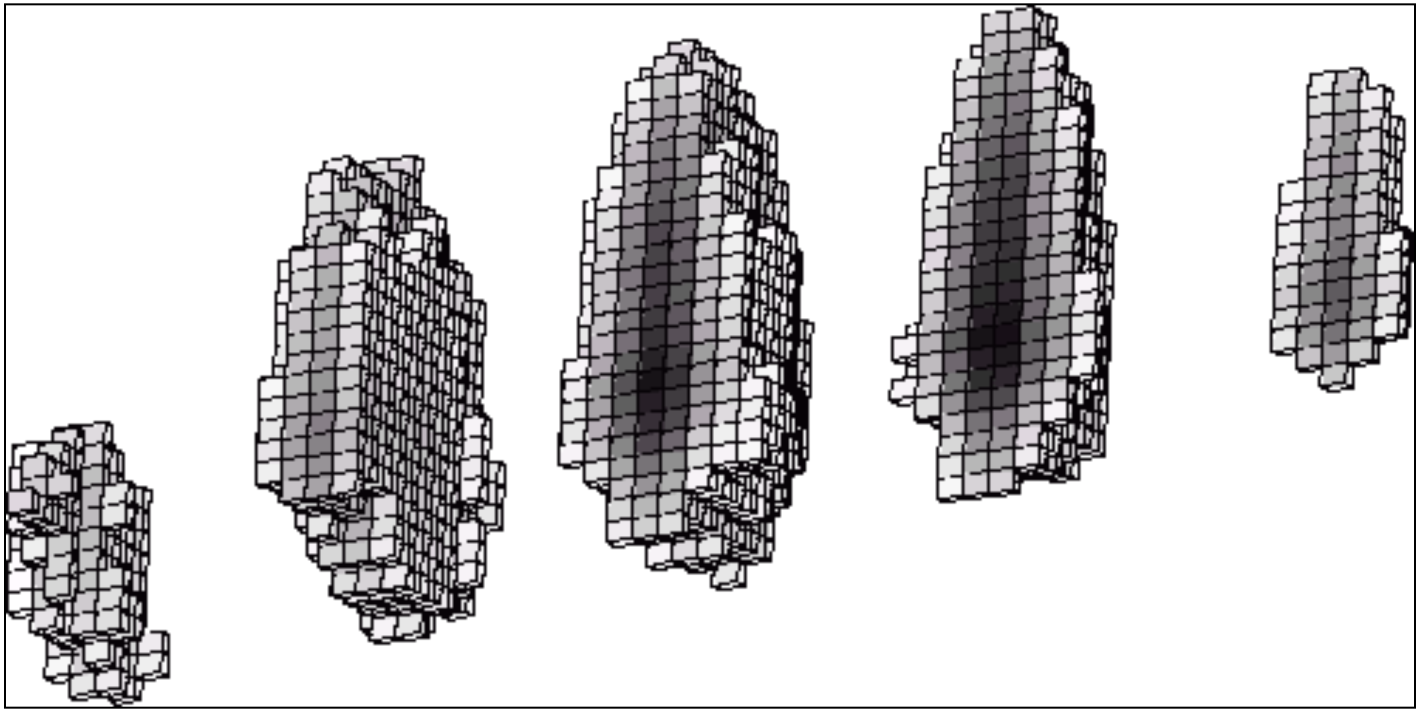
Tree-based Representation

Idea: Represent density using a variant of octrees



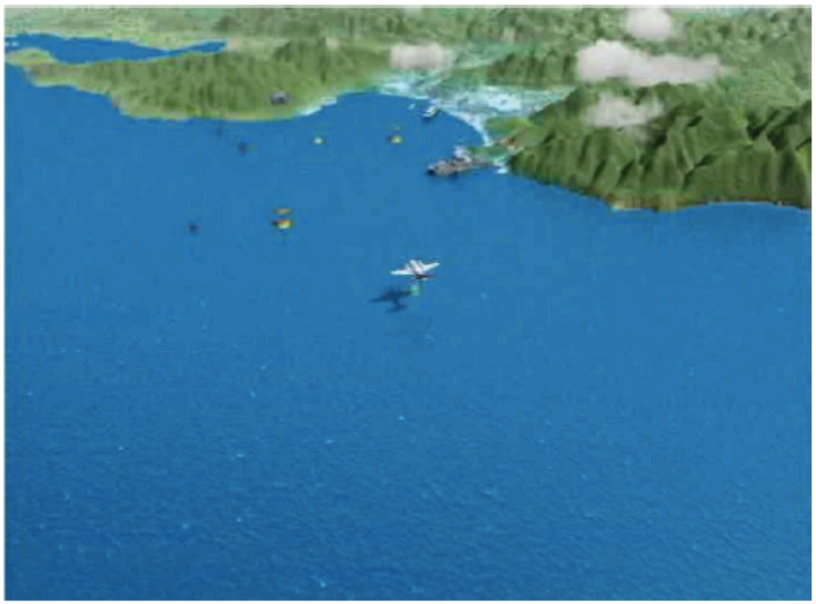
Tree-based Representations

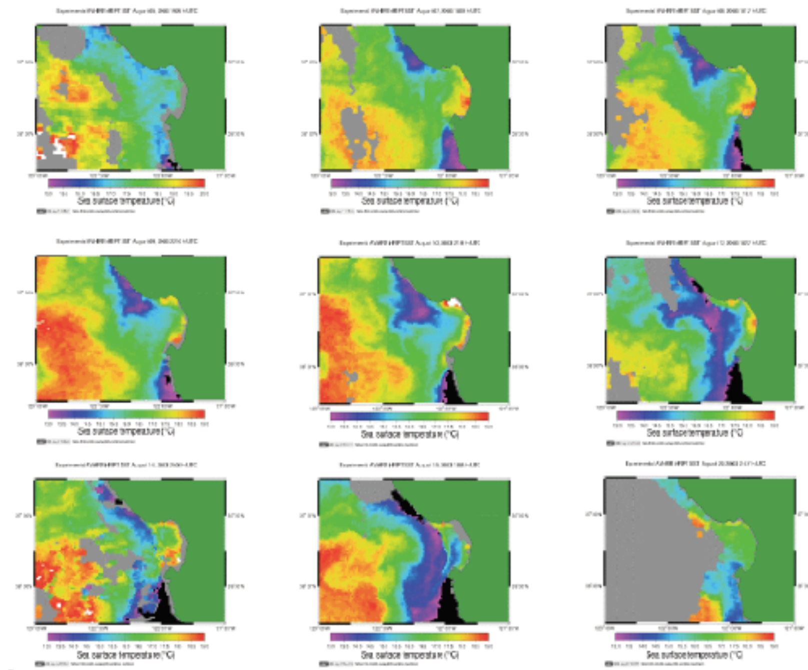
- Efficient in space and time
- Multi-resolution



Exploration with Mobile Robots

Example Application: Sampling Spatiotemporal Fields





Satellite Sea Surface Temperature (SST), Monterey Bay, CA, Aug 5-20, 2003

Questions for Ocean Sampling Example

- How to represent the objective that the goal of motion planning is to acquire information which is then used in model learning?
- Concretely, how to decide where and when to sample on the basis of this?

Exploration Problems

- Exploration: control a mobile robot so as to maximize knowledge about the external world
- Example: robot needs to acquire a map of a static environment. If we represent map as "occupancy grid", exploration is to maximise cumulative information we have about each grid cell
- POMDPs already subsume this function but we need to define an *appropriate payoff function*
- One good choice is information gain:

    Reduction in entropy of a robot's belief as a function of its actions

Exploration Heuristics

- While Partially Observable versions of MDPs (i.e., POMDPs) are conceptually useful here, we may not want to use them directly – state/observation space is huge
- We will instead try to derive greedy heuristic based on the notion of *information gain*.
- Limit lookahead to just one exploration action
  - The exploration action could itself involve a sequence of control actions (but logically, it will serve as one exploration action)
  - For instance, select a location to explore anywhere in the map, then go there

Information and Entropy

- The key to exploration is information.
- Entropy of expected information:

$$H_p(x) = -\int p(x) \log p(x) dx \qquad \text{or} \qquad -\sum_x p(x) \log p(x)$$

- Entropy is at its maximum for a uniform distribution, p
- Conditional entropy is the entropy of a conditional distrib.
- In exploration, we seek to minimize the expected entropy of the belief after executing an action
- So, condition on measurement z and control u that define the belief state transition

Conditional Entropy after Action/Observation

- With $B(b,z,u)$ denoting the belief after executing control u and observing z under belief b,
- Conditional entropy of state x' after executing action u and measuring z is given by,

$$H_b(x'|z,u) = -\int B(b,z,u)(x') \log B(b,z,u)(x') dx'$$

- The conditional entropy of the control is,

$$H_b(x'|u) = E_z[H_b(x'|z,u)]$$

$$= \int\int H_b(x'|z,u) p(z|x') p(x'|u,x) b(x) dz dx' dx$$

Greedy Techniques

- Expected information gain lets us phrase exploration as a decision theoretic problem.

- Information Gain is

$$
\begin{aligned}
I_b(u) &= H_p(x) - H_b(x'|u) \\
&= H_p(x) - E_z[H_b(x'|z,u)]
\end{aligned}
$$

Greedy Techniques

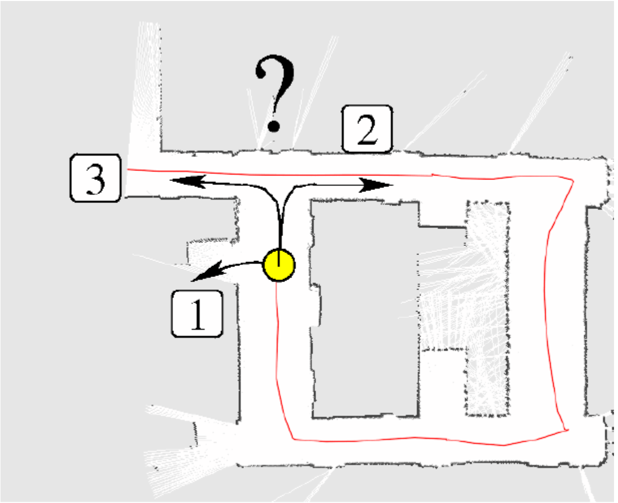
- If $r(x,u)$ is the cost of applying control u in state x (treating cost as negative numbers), then optimal greedy exploration for the belief b maximizes difference between information gain and cost,

$$\pi(b) = \arg\max_u \alpha\underbrace{(H_p(x) - E_z[H_b(x'|z,u)])}_{\text{Expected information gain (Original entropy – Cond. Entropy)}} + \underbrace{\int r(x,u)b(x)dx}_{\text{Expected cost}}$$

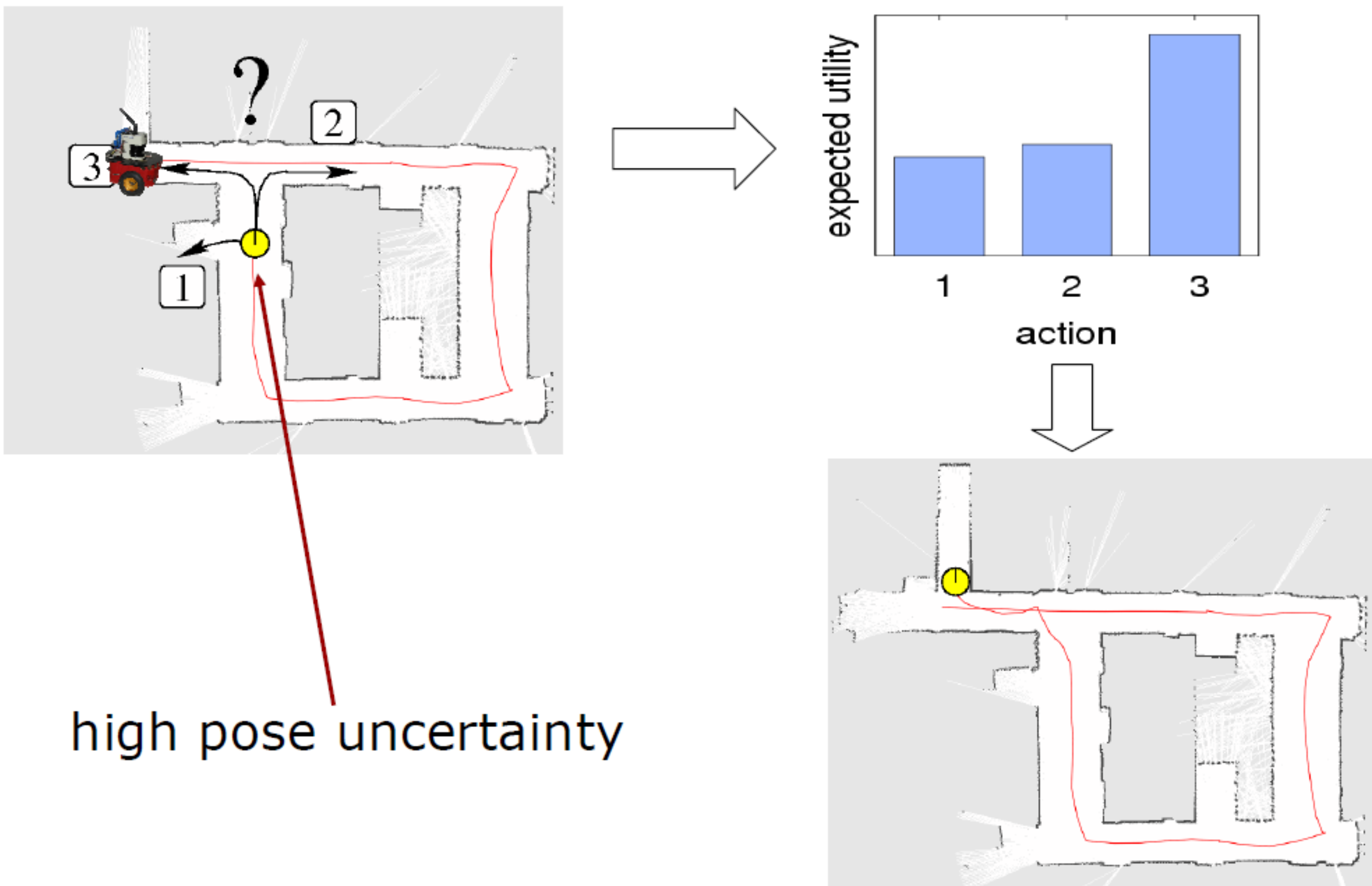
Example: Combining Exploration and Mapping

- By reasoning about control, the mapping process can be made much more effective

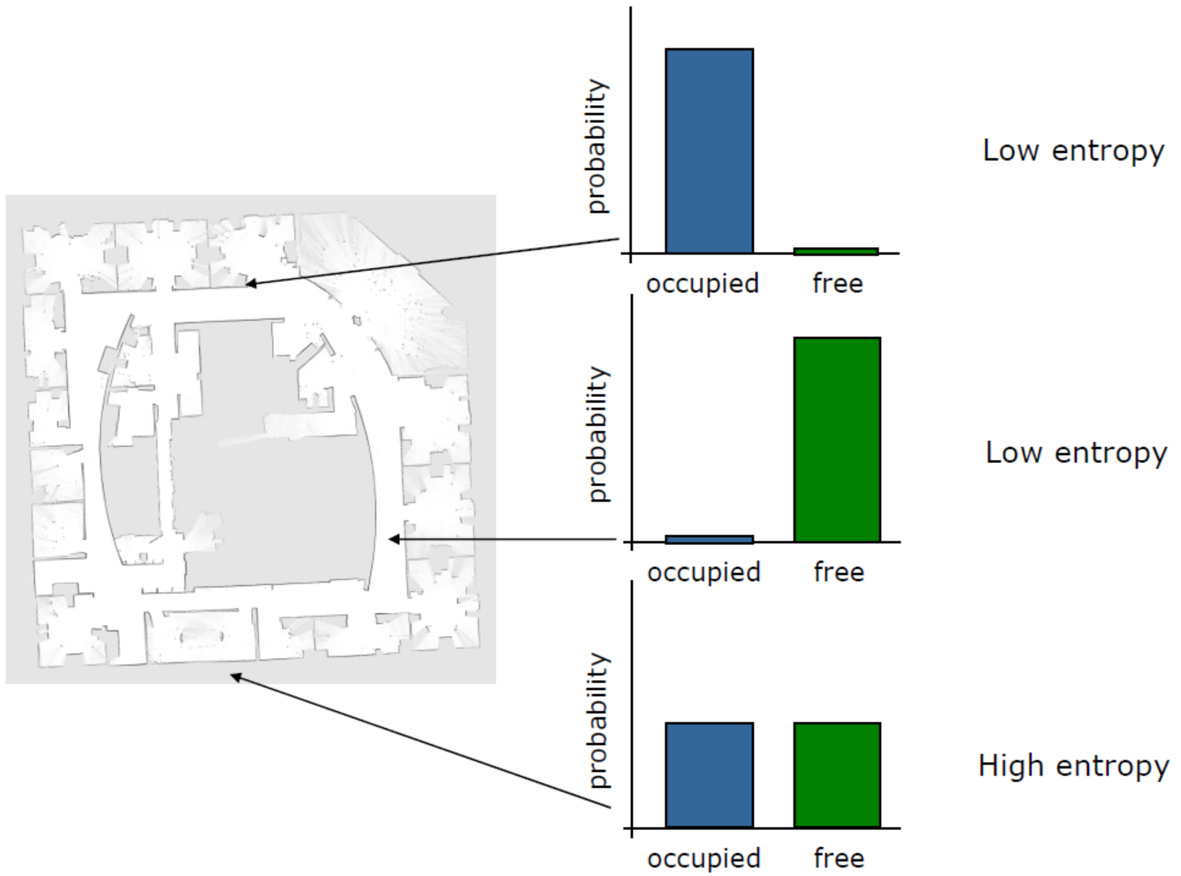
- Question: Where to move next in a map?



Exploration Problem: Visually



Map Entropy



The overall entropy is the sum of the individual entropy values