Probability: Review

Pieter Abbeel
UC Berkeley EECS

Many slides adapted from Thrun, Burgard and Fox, Probabilistic Robotics

Why probability in robotics?

- Often state of robot and state of its environment are unknown and only noisy sensors available
  - Probability provides a framework to fuse sensory information
  - → Result: probability distribution over possible states of robot and environment
- Dynamics is often stochastic, hence can't optimize for a particular outcome, but only optimize to obtain a good distribution over outcomes
  - Probability provides a framework to reason in this setting
  - → Result: ability to find good control policies for stochastic dynamics and environments

Example 1: Helicopter

- State: position, orientation, velocity, angular rate
- Sensors:
  - GPS : noisy estimate of position (sometimes also velocity)
  - Inertial sensing unit: noisy measurements from
    (i) 3-axis gyro [=angular rate sensor],
    (ii) 3-axis accelerometer [=measures acceleration + gravity; e.g., measures (0,0,0) in free-fall],
    (iii) 3-axis magnetometer
- Dynamics:
  - Noise from: wind, unmodeled dynamics in engine, servos, blades

Example 2: Mobile robot inside building

- 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

Axioms of Probability Theory

- $0 \le \Pr(A) \le 1$
- $\Pr(\Omega) = 1$ $\qquad \Pr(\phi) = 0$
- $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$

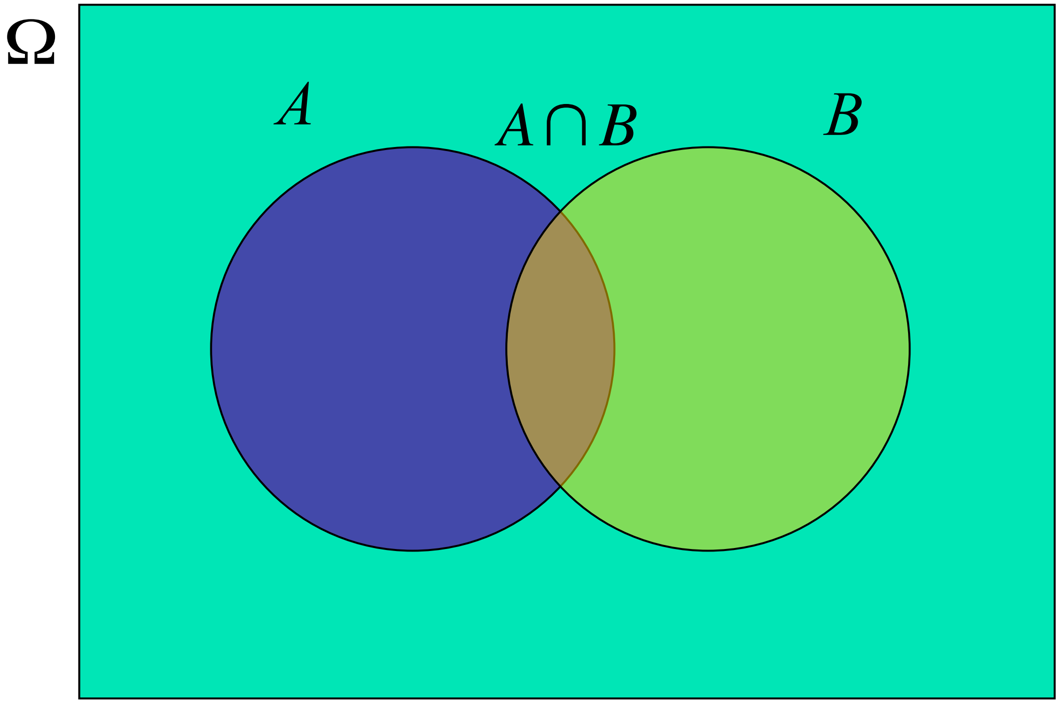
Pr*(A)* denotes probability that the outcome ω is an element of the set of possible outcomes *A*. *A* is often called an event. Same for *B.*

Ω is the set of all possible outcomes.

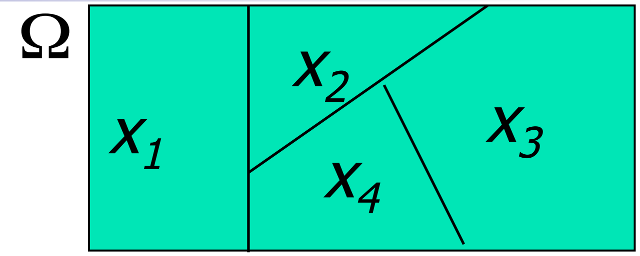
ϕ is the empty set.

A Closer Look at Axiom 3

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



Discrete Random Variables



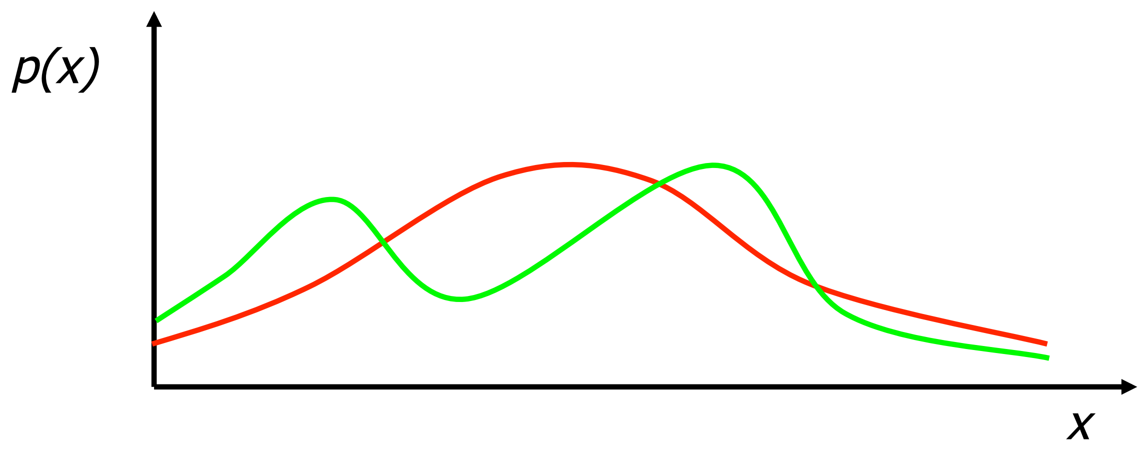
- *X* denotes a random variable.
- *X* can take on a countable number of values in $\{x_1, x_2, \ldots, x_n\}$.
- $P(X=x_i)$, or $P(x_i)$, is the probability that the random variable X takes on value x_i.
- $P(\cdot)$ is called probability mass function.
- *E.g., X models the outcome of a coin flip*, x_1 = head, x_2 = tail, $P(x_1) = 0.5$, $P(x_2) = 0.5$

Continuous Random Variables

- *X* takes on values in the continuum.
- *p(X=x)*, or *p(x)*, is a probability density function.

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

- E.g.



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

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

- *Same for probability densities, just* $P \rightarrow p$

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

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

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

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

Bayes Formula

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

$$\Rightarrow$$

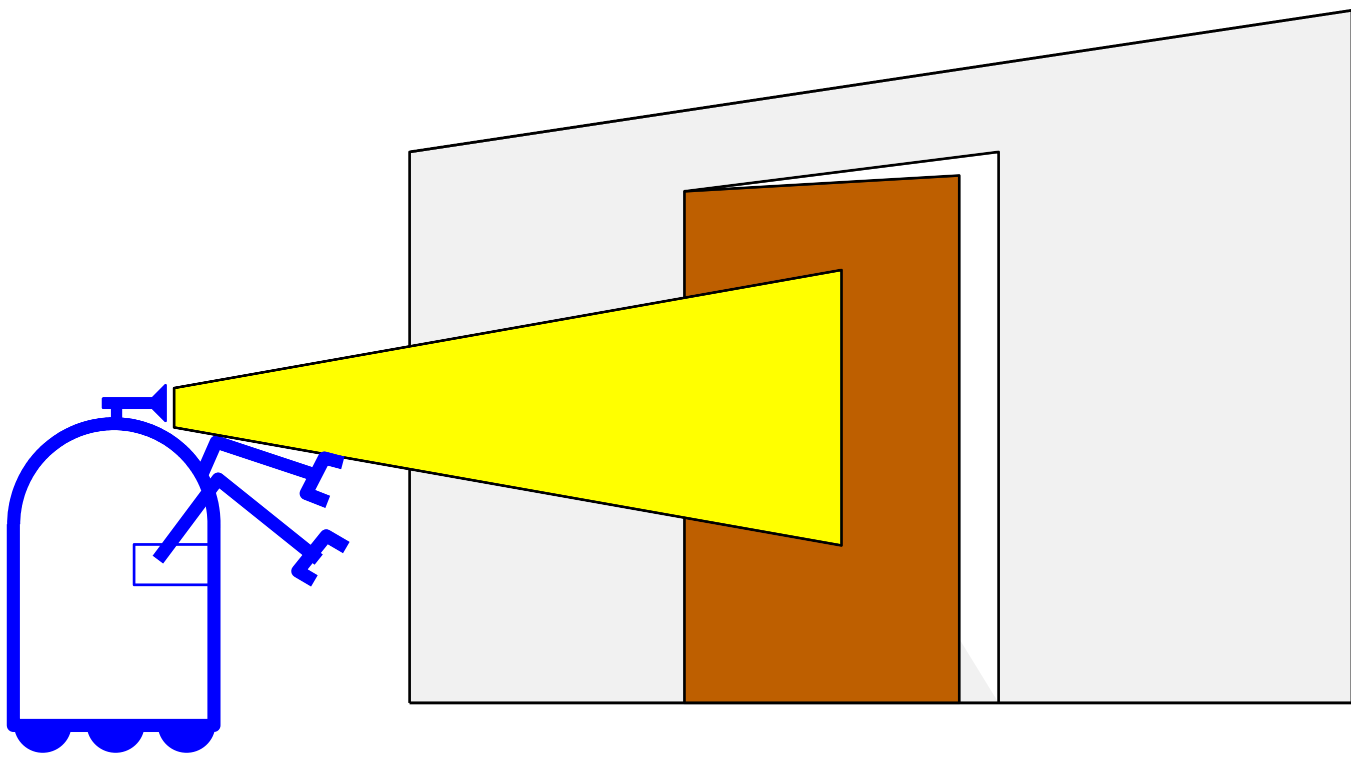
$$P(x \mid y) = \frac{P(y \mid x)\ P(x)}{P(y)} = \frac{\text{likelihood} \cdot \text{prior}}{\text{evidence}}$$

Bayes Rule with Background Knowledge

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

Simple Example of State Estimation

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



Causal vs. Diagnostic Reasoning

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

count frequencies!

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

Bayes Filters

Pieter Abbeel
UC Berkeley EECS

Many slides adapted from Thrun, Burgard and Fox, Probabilistic Robotics

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***.

Integrating the Outcome of Actions

Continuous case:

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

Discrete case:

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

Measurements

- Bayes rule

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

Bayes Filters: Framework

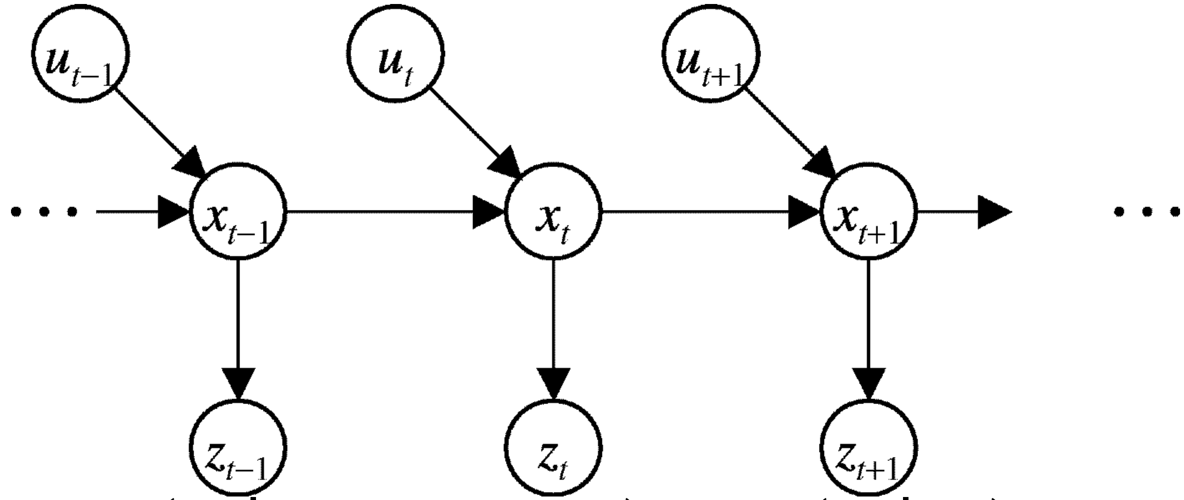
- **Given:**
  - Stream of observations z and action data *u:*

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

  - 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



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

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

Underlying Assumptions

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

Bayes Filters

z = observation
u = action
x = state

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

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

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

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

Algorithm **Bayes_filter**(*Bel(x),d*):

$\eta = 0$

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 Applications

- Robot localization:
  - Observations are range readings (continuous)
  - States are positions on a map (continuous)

- Speech recognition HMMs:
  - Observations are acoustic signals (continuous valued)
  - States are specific positions in specific words (so, tens of thousands)

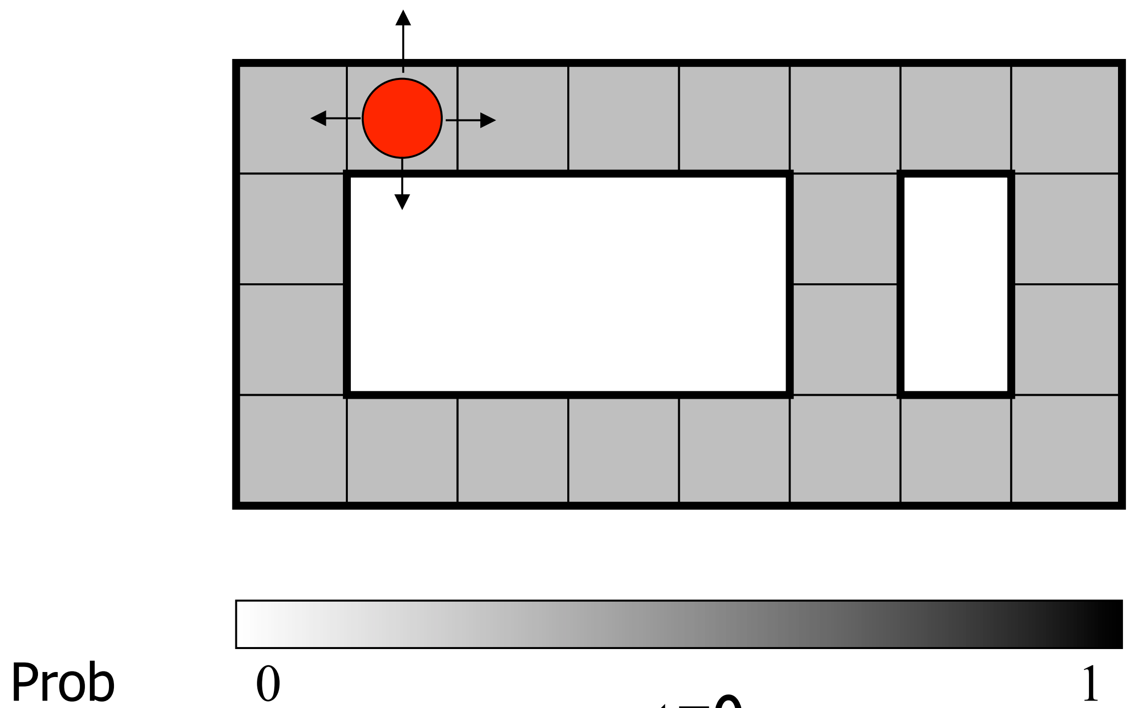
- Machine translation HMMs:
  - Observations are words (tens of thousands)
  - States are translation options

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.

Example: Robot Localization

Example from Michael Pfeiffer



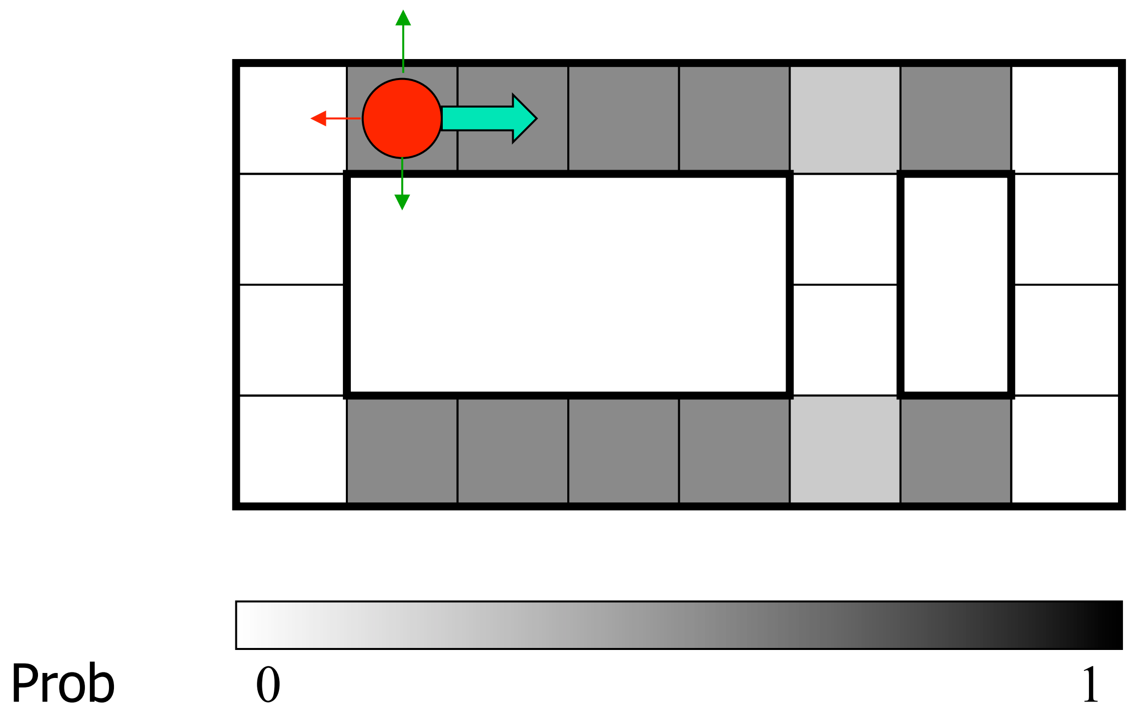
t=0

Sensor model: never more than 1 mistake

Know the heading (North, East, South or West)

Motion model: may not execute action with small prob.

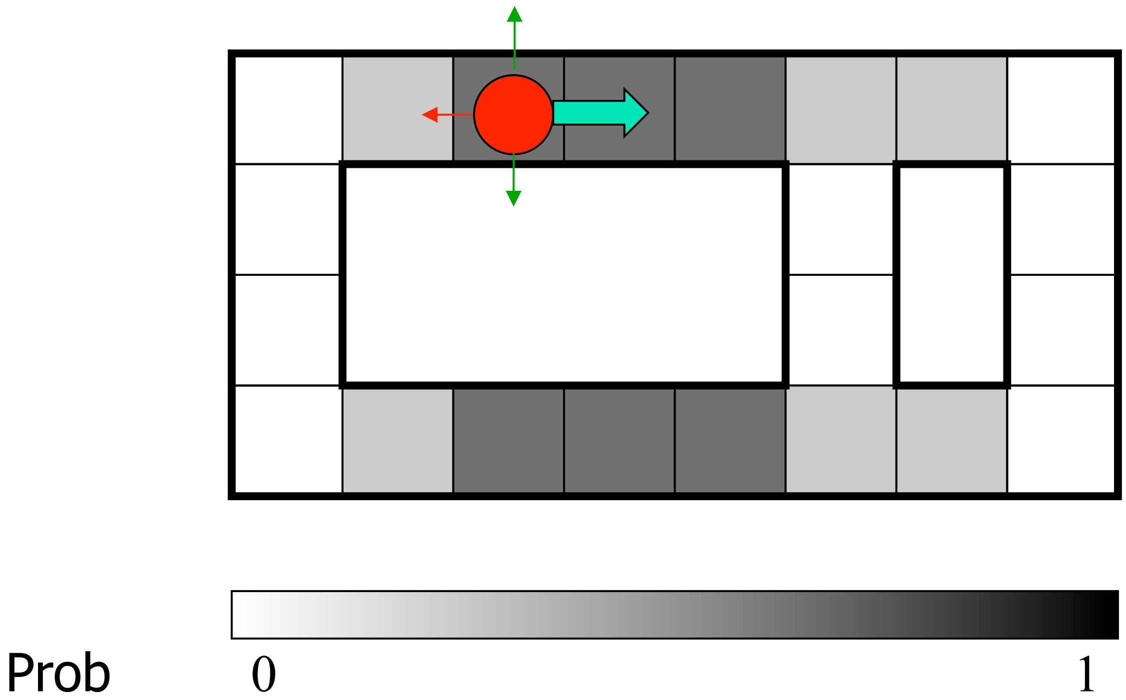
Example: Robot Localization



t=1

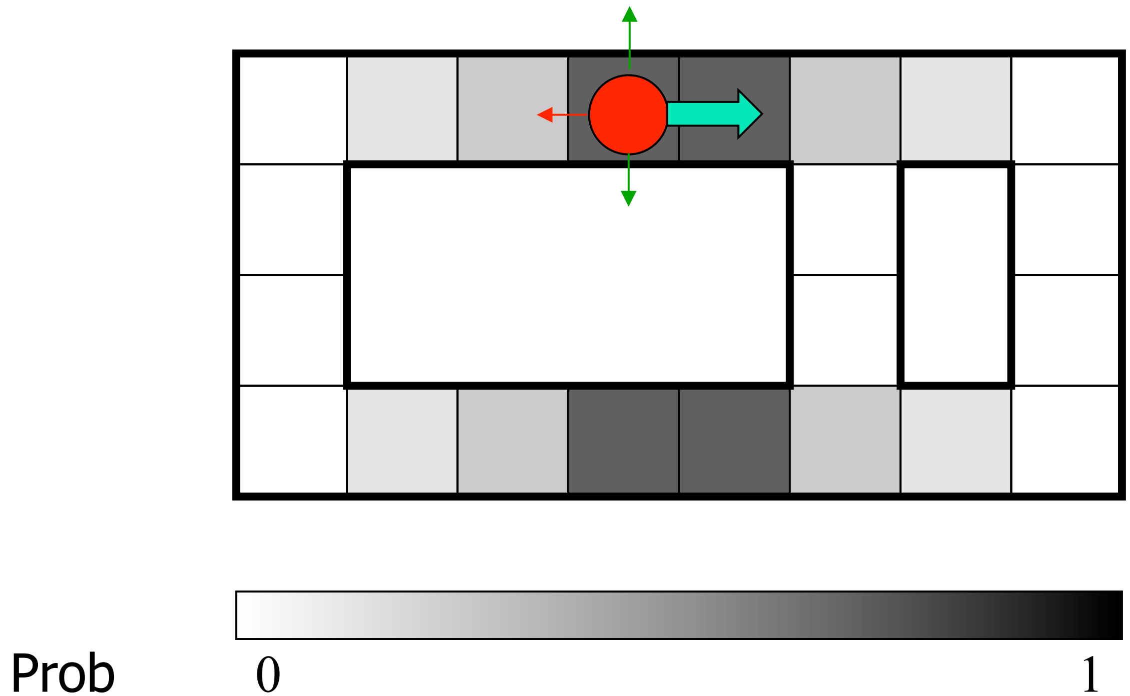
Lighter grey: was possible to get the reading, but less likely b/c required 1 mistake

Example: Robot Localization



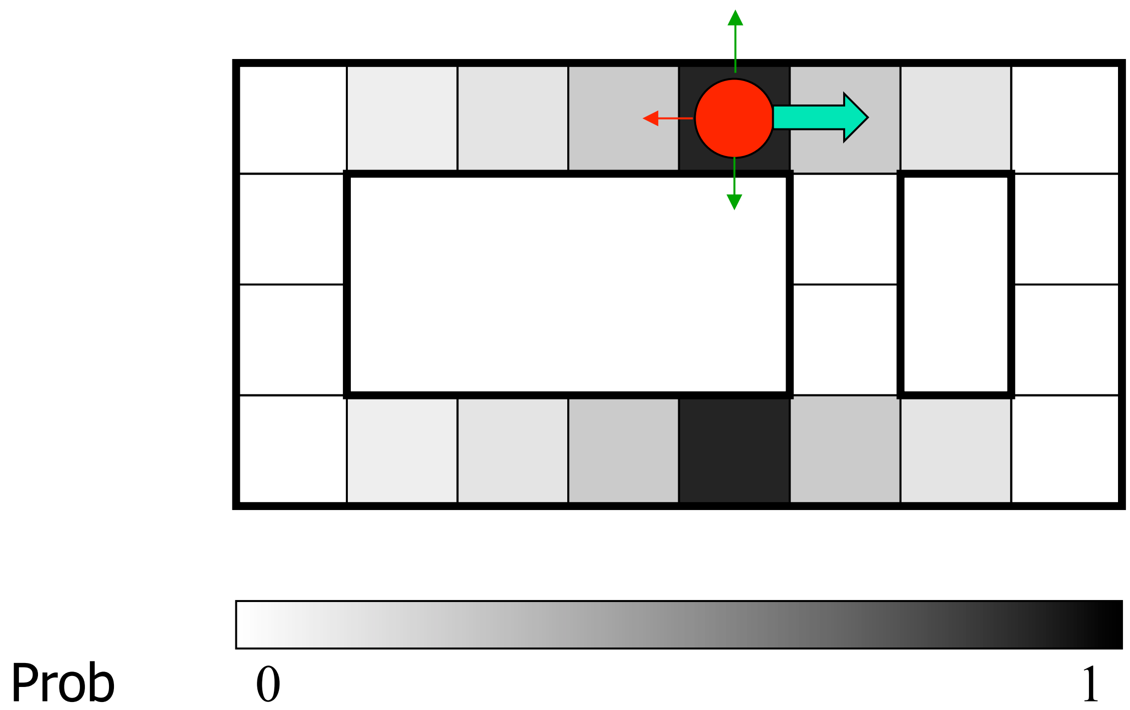
t=2

Example: Robot Localization



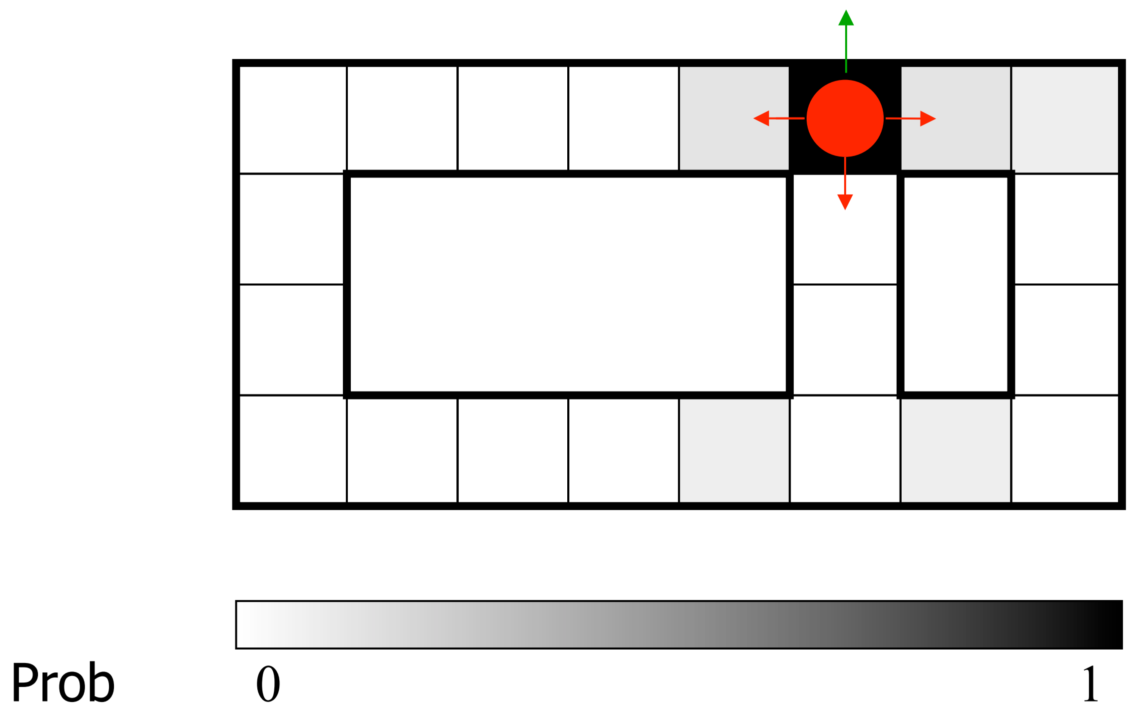
t=3

Example: Robot Localization



t=4

Example: Robot Localization



t=5