Bayesian Estimation

EE468/CE468: Mobile Robotics

Dr. Basit Memon

Electrical and Computer Engineering
Habib University

November 1, 2023

Table of Contents

Table of Contents

A trash sorting robot

- The robot is fed items of trash, one at a time, by a conveyor belt.
- Depending on the category of the item (state), e.g., paper, metal, etc., the robot's task is to move the item (actions) into the appropriate bin.
- Sensors are used to measure certain characteristics (sensing) of a given item of trash, and these sensor measurements are then used to infer the item's category (simple perception).
- Using these inferences, the robot chooses an action (planning) to place the item in an appropriate bin.

Prior belief about state

- For every 1000 pieces of trash,
  - cardboard: 200
  - paper: 300
  - cans: 250
  - scrap metal: 200
  - bottle: 50

$$p_C(c) = \begin{cases} 0.2 & \text{if c = cardboard,} \\ 0.3 & \text{if c = paper,} \\ 0.25 & \text{if c = can,} \\ 0.2 & \text{if c = metal,} \\ 0.05 & \text{if c = bottle} \end{cases}$$

Sensors for sorting trash are imperfect

- a conductivity sensor, outputting true|false
- a camera with three detection algorithms: bottle, cardboard, paper
- a scale, outputting continuous value in kg

Sensor uncertainty is captured as conditional PDFs.

- a conductivity sensor, outputting true|false

$$p(Z^1|C) =$$

Category	false	true
cardboard	0.99	0.01
paper	0.99	0.01
can	0.1	0.9
scrap metal	0.15	0.85
bottle	0.95	0.05

Bayesian framework computes posterior PDF

$$p(C|Z^1) = \frac{p(Z^1|C)p(C)}{p(Z^1)}$$
$$= \frac{p(Z^1|C)p(C)}{\sum_{c \in C} p(Z^1|C)p(C)}$$

- If sensor returns TRUE(1), then the value of PDF $p(C|1)$ for $C =$ cardboard can be computed as:

$$\frac{(0.01)(0.2)}{(0.01)(0.2) + (0.01)(0.3) + (0.9)(0.25) + (0.85)(0.2) + (0.05)(0.05)}.$$

Determine state using a loss function based on posterior PDF.

$$p(C|1) = \begin{cases} 0.005 & \text{if c = cardboard,} \\ 0.0075 & \text{if c = paper,} \\ 0.559 & \text{if c = can,} \\ 0.4224 & \text{if c = metal,} \\ 0.0062 & \text{if c = bottle} \end{cases}$$

- We can decide the state using some rule, e.g. MAP rule

$$\hat{C} = \arg\max_c p(C|Z^1).$$

Revisiting the posterior PDF expression

$$p(C|Z^1) = \frac{p(Z^1|C)p(C)}{p(Z^1)} = \eta\, p(Z^1|C)\, p(C)$$

- Notice that the denominator does not depend on unknown state and is simply playing the role of scaling factor, i.e.

$$p(C|Z^1) \propto \underbrace{p(Z^1|C)}_{\text{Likelihood Function}} \quad p(C),$$

  so it can be ignored in our optimization.

- The scaling factor η, can be explicitly computed as it is number that makes sum of PMF equal to 1.

Incorporate the camera as well.

- a camera with three detection algorithms: bottle, cardboard, paper

$$p(Z^2|C) =$$

Category	**bottle**	**cardboard**	**paper**
cardboard	0.02	0.88	0.1
paper	0.02	0.2	0.78
can	0.33	0.33	0.34
scrap metal	0.33	0.33	0.34
bottle	0.95	0.02	0.03

Multiple sensors require independence.

$$p(C|Z^1, Z^2) = \frac{p(Z^1, Z^2|C)p(C)}{p(Z^1, Z^2)}$$
$$\propto p(Z^1, Z^2|C)\, p(C)$$

- In order to proceed further, we have to assume conditional independence of Z^1 and Z^2, conditioned on knowing the state. Is it reasonable?

$$p(Z^1, Z^2|C) = p(Z^1|C)\, p(Z^2|C)$$
$$p(C|Z^1, Z^2) \propto p(Z^1|C)\, p(Z^2|C)\, p(C)$$

Weight sensor is a continuous measurement.

- a scale, outputting continuous value in kg

- In this case, $p(Z^3|C)$ is a different PDF for each value of C. Note that C is still discrete, but values of Z^3, the weight, are continuous.

- Posterior belief about the state is determined in a similar way:

$$p(C|Z^1, Z^2, Z^3) \propto p(Z^1|C)\, p(Z^2|C)\, p(Z^3|C)\, p(C).$$

- Note that $p(C|Z^1, Z^2, Z^3)$ is discrete PMF.

Vacuum Cleaning Robot

- State of the robot is continuous.
- State at time t is x_t, a random variable.
- Prior distribution of the state is known, i.e. $p(x_1)$.
- State is only considered at discrete instants in time, i.e. $x_1, x_2, x_3, \cdots$
- State at time $t+1$ depends on the state at time t and the action/control at time t.

Outcome of actions is uncertain and measurements are noisy.

- Uncertainty in state transitions is captured through conditional distributions in our motion model, i.e. x_{t+1} is not deterministic, given the past states and actions.

- At each time step, the robot receives measurement z_t from its sensors.

Evolution of state of robot system

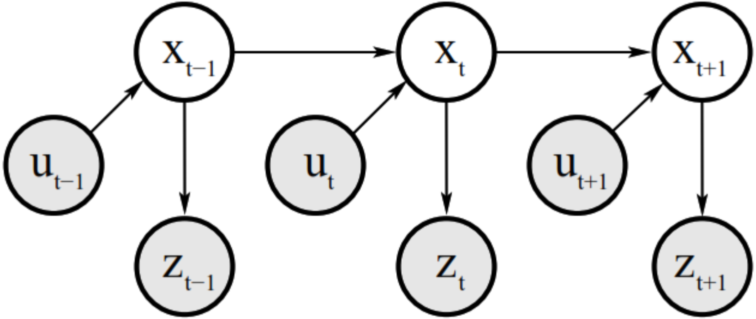


Figure: Relationship between state, action, measurements [1]

Extending Bayes: Continuous state and multiple measurements

- Can we incorporate measurements as they arrive, instead of processing as a batch? Iterative scheme?

Belief over state is given by conditional probability distribution.

$$
\begin{aligned}
bel(x_t) &= p\left(x_t \middle| z_{1:t}, u_{1:t}\right) \\
&= p\left(x_t \middle| z_1, z_2, \cdots, z_t, u_1, u_2, \cdots, u_t\right)
\end{aligned}
$$

- Belief gives us the **probability distribution, p**, over state $\mathbf{x_t}$ at time t, given entire sequences of past actions, $\mathbf{u_{1:t}}$, and measurements, $\mathbf{z_{1:t}}$.
- How do you compute this belief?

Bayes' Filter Derivation (1)

$$
\begin{aligned}
bel(x_t) &= p(x_t | z_{1:t}, u_{1:t}) \\
&= \underbrace{\frac{p(z_t | x_t, z_{1:t-1}, u_{1:t}) p(x_t | z_{1:t-1}, u_{1:t})}{p(z_t | z_{1:t-1}, u_{1:t})}}_{\text{Conditional Bayes: } p(A|B,C) = \frac{p(B|A,C)p(A|C)}{p(B|C)}} \\
&= \eta \, p(z_t | x_t, z_{1:t-1}, u_{1:t}) \, p(x_t | z_{1:t-1}, u_{1:t})
\end{aligned}
$$

Since denominator doesn't depend on x, we can simply treat it as a normalizing factor, η, needed to make integral of numerator equal to 1.

Markov or Complete State assumption

- Knowing the state x_t at time t is enough to predict the measurement z_t at time t.
  - The past measurements or control don't provide any additional information.
- Knowing the state x_{t-1} at time $t-1$ and control u_t at time t is enough to predict the state x_t.
- Conditional independence given state:

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

Bayes' Filter Derivation (2)

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

$$= \eta \underbrace{p(z_t \mid x_t)}_{\text{Markov assumption}} p(x_t \mid z_{1:t-1},\, u_{1:t})$$

$$= \eta\, p(z_t \mid x_t) \underbrace{\int p(x_t, x_{t-1} \mid z_{1:t-1},\, u_{1:t})\, dx_{t-1}}_{\text{By definition of joint pdf and marginal pdf, } \int_{b_i} p(A,B \mid C)\, db_i = p(A \mid C)}$$

$$bel(x_t) = \eta\, p(z_t \mid x_t) \underbrace{\int p(x_t \mid x_{t-1}, z_{1:t-1}, u_{1:t})\, p(x_{t-1} \mid z_{1:t-1}, u_{1:t}) dx_{t-1}}_{\text{Conditional Bayes, again}}$$

$$= \eta\, p(z_t \mid x_t) \int \underbrace{p(x_t \mid x_{t-1}, u_t)}_{\text{Markov assumption}}\, p(x_{t-1} \mid z_{1:t-1}, u_{1:t})\, dx_{t-1}$$

$$= \eta\, p(z_t \mid x_t) \int p(x_t \mid x_{t-1}, u_t)\, p(x_{t-1} \mid z_{1:t-1}, \underbrace{u_{1:t-1}}_{\text{Independence}})\, dx_{t-1}$$

State at time $t-1$ is not affected by control at time t.

Bayes' Filter Derivation (4)

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

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

$$bel(x_t) = \underbrace{\eta\, p(z_t|\, x_t) \underbrace{\int p(x_t|\, x_{t-1},\, u_t)\, bel(x_{t-1})\, dx_{t-1}}_{\text{Belief after prediction step}}}_{\text{Belief after update step}}$$

Recursive Bayes Algorithm

- Start with initial belief: $bel(x_{t-1})$
- **Prediction Step:** Push belief through system dynamics, given action:

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

- **Update Step:** Apply correction, based on measurement:

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

Ingredients of Recursive Bayes Algorithm

- **Initial Probability Distribution** $bel(x_0)$**.**
  - Initial robot location is unknown $\rightarrow$ uniform distribution over all poses.
  - Perfectly known $\rightarrow$ Delta function
- **Probabilistic Motion Model.**
  - Deterministic model from kinematics. $x_t = f(x_{t-1}, u_t)$.
  - Model error distribution over x_{t-1} and u_t, and then transform using f to obtain distribution for x_t, i.e. $p(x_t|u_t, x_{t-1})$.
  - Example of this is how we transformed covariance matrix earlier. Gaussian distribution was assumed.

- **Probabilistic Measurement Model.**
  - Deterministic model: $z_t = h(x_t)$
  - Typically, involves change of coordinate frames.
  - Probabilistic model is obtained by modeling noise on top of deterministic model to get $p(z_t|x_t)$.
  - As simple as $z_t = h(x_t) + n_t$

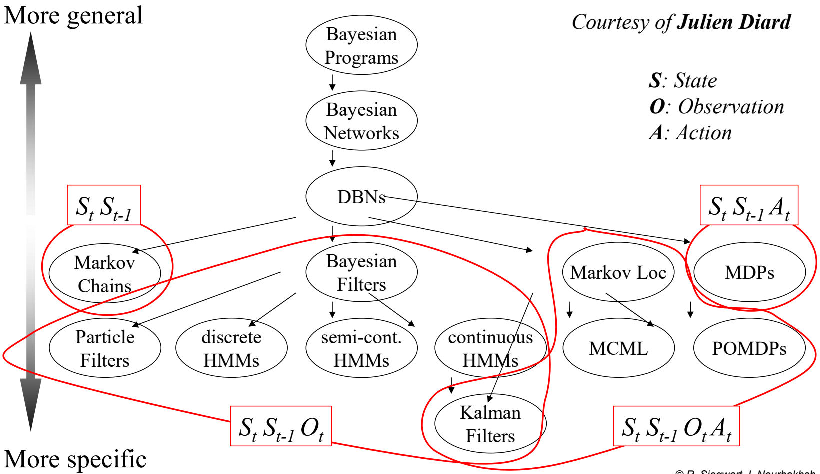
When does Markov assumption fail?

- Unmodeled dynamics in the environment not included in x_t, e.g. moving people.
- Inaccuracies in the motion and measurement models.
- Approximation errors when using approximate representations of belief functions, e.g. it is difficult to evolve entire distributions.
- Surprisingly, Bayes filters are found to be robust to such violations.

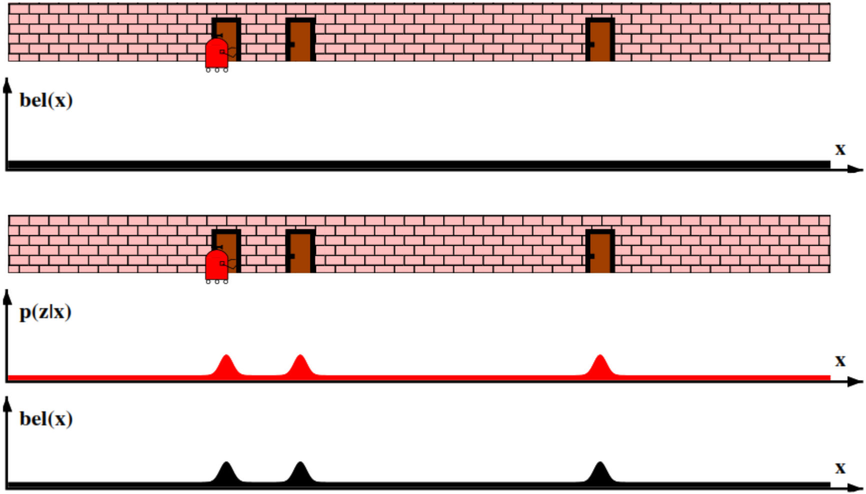
Observations about prediction step

- Computation of belief requires an integral (sum). This is because the belief has to be computed looking at all possible paths that could be taken to reach x_t.
- Integral needs to be computed for each x_t, i.e. computation needs to be done as many times as possible robot poses.
- Integral can be seen as a convolution, which could convince us prediction step increases uncertainty, i.e. belief function spreads.

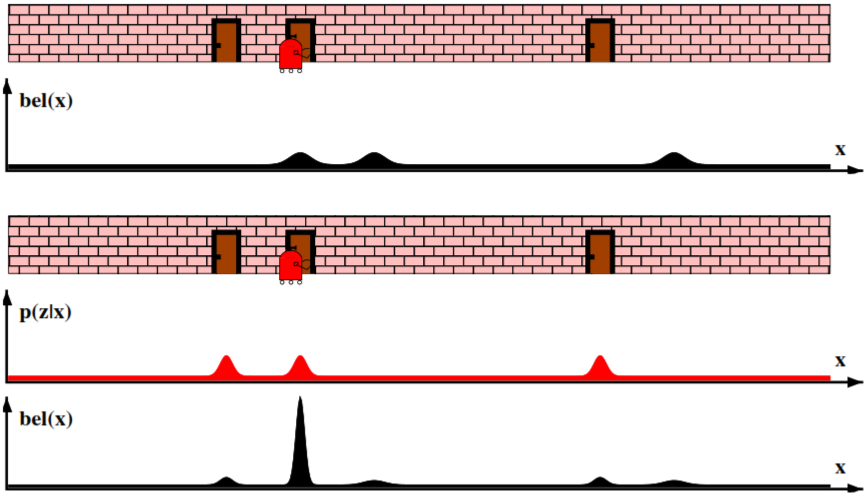
Bayesian Approach: A taxonomy of probabilistic models



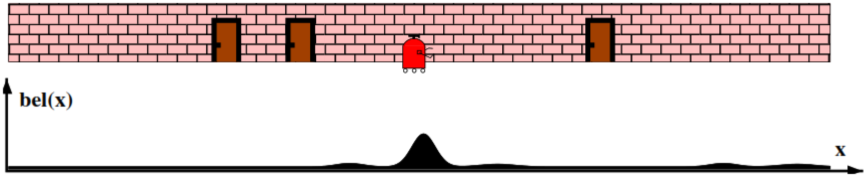
Bayes' view of localization: Belief updated given measurement



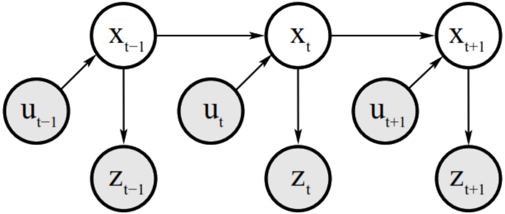
Bayes' view of localization: Belief updated after action



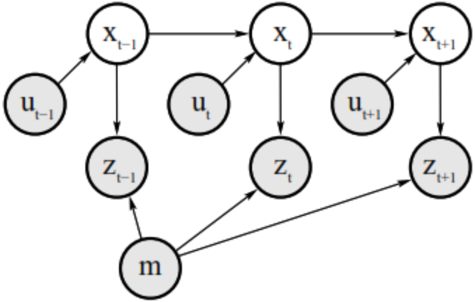
Bayes' view of localization: Belief provides good pose estimate



Pose Fixing requires maps.



$p(z_t | x_t)$



$p(z_t | x_t, \mathbf{m})$

Table of Contents

[1] **Sebastian Thrun, Wolfram Burgard, and Dieter Fox.**
Probabilistic robotics.
2006.