MEN791
Autonomous Unmanned Vehicles

Bayes Filter

CONTACT

Ulsan National Institute of Science and Technology

Address 50 UNIST-gil, Ulju-gun, Ulsan, 44919, Korea
Tel. +82 52 217 0114 **Web.** www.unist.ac.kr

School of Mechanical, Aerospace, and Nuclear Engineering

5th Engineering Building Room 801-4

Tel. +82 52 217 3048 **Web.** https://sites.google.com/site/aslunist/

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.

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

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. $P(Room) = \langle 0.7, 0.2, 0.08, 0.02 \rangle$

“Probability Sums up to One”

Discrete case

$$\sum_x P(x) = 1$$

Continuous case

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

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

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

Law of Total Probability

Discrete case

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

Continuous case

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

Marginalization

Discrete case

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

Continuous case

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

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:

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

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

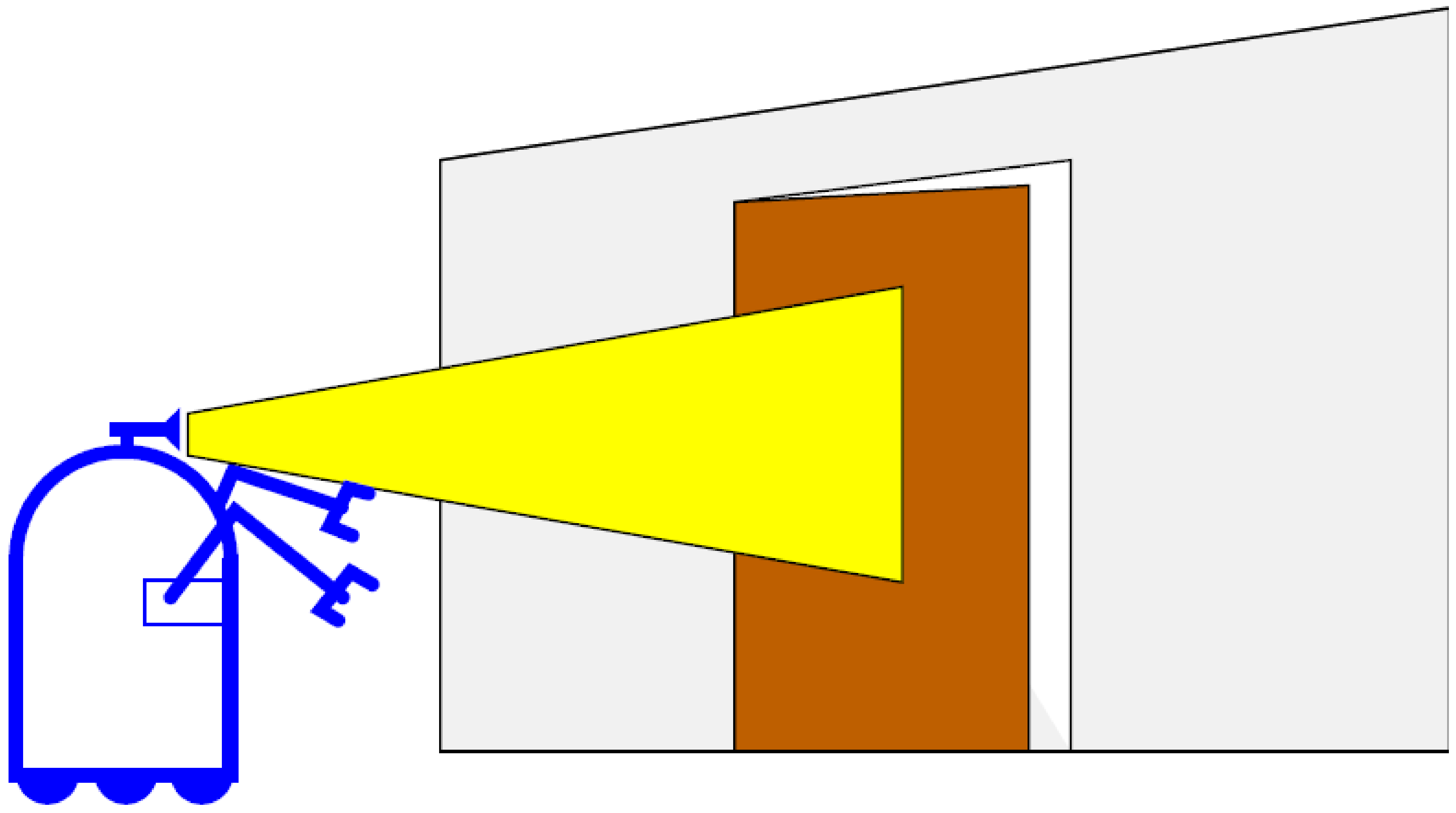
$$\forall x : P(x \mid y) = \eta\ \mathrm{aux}_{x|y}$$

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

count frequencies!

- Bayes rule allows us to use causal knowledge:

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

Example

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

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

$$P(open \mid z) = \frac{0.6 \cdot 0.5}{0.6 \cdot 0.5 + 0.3 \cdot 0.5} = \frac{0.3}{0.3 + 0.15} = 0.67$$

- *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

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

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

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

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

Example: Second Measurement

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

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

$$= \frac{\frac{1}{2}\cdot\frac{2}{3}}{\frac{1}{2}\cdot\frac{2}{3}+\frac{3}{5}\cdot\frac{1}{3}} = \frac{\frac{1}{3}}{\frac{1}{3}+\frac{1}{5}} = \frac{\frac{1}{3}}{\frac{8}{15}} = \frac{5}{8} = 0.625$$

- z_2 lowers the probability that the door is open

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**…

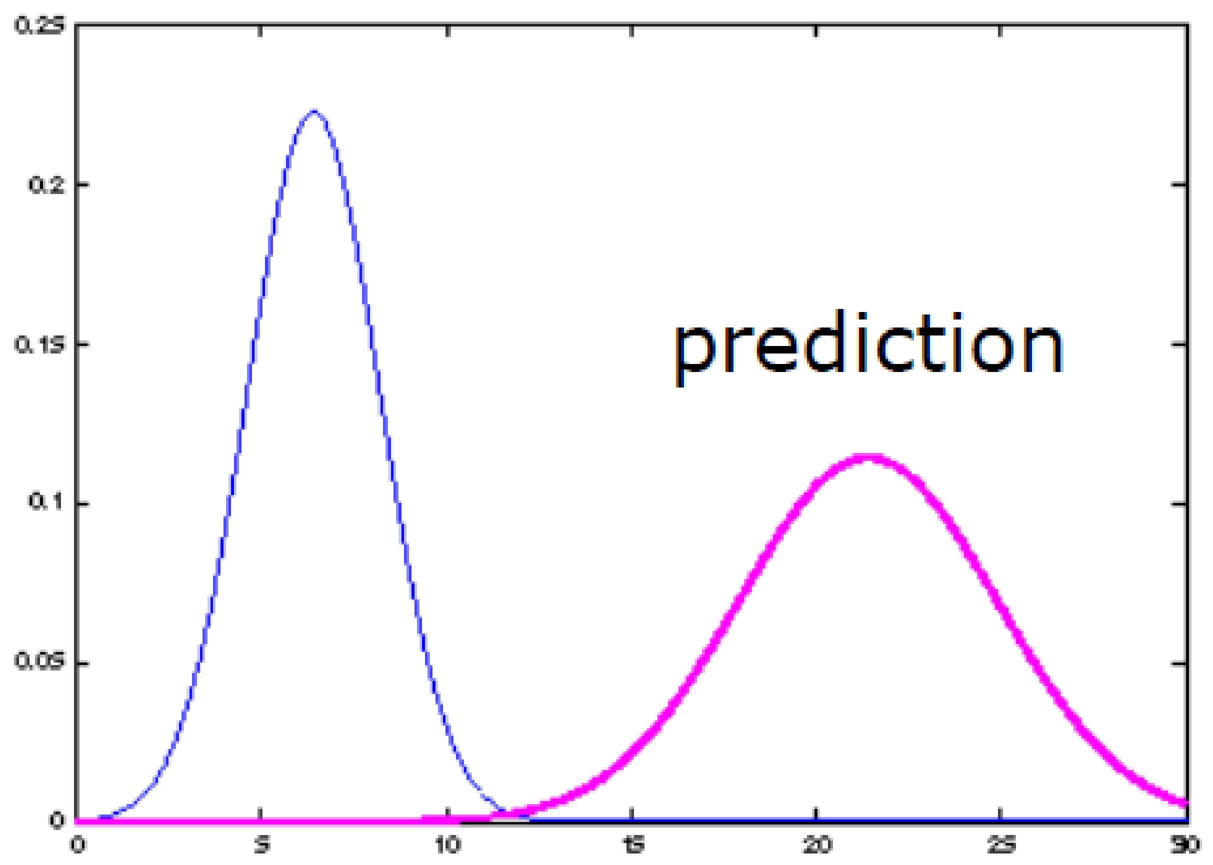
- Actions are **never carried out with absolute certainty**
- In contrast to measurements, **actions generally increase the uncertainty**

Modeling Actions

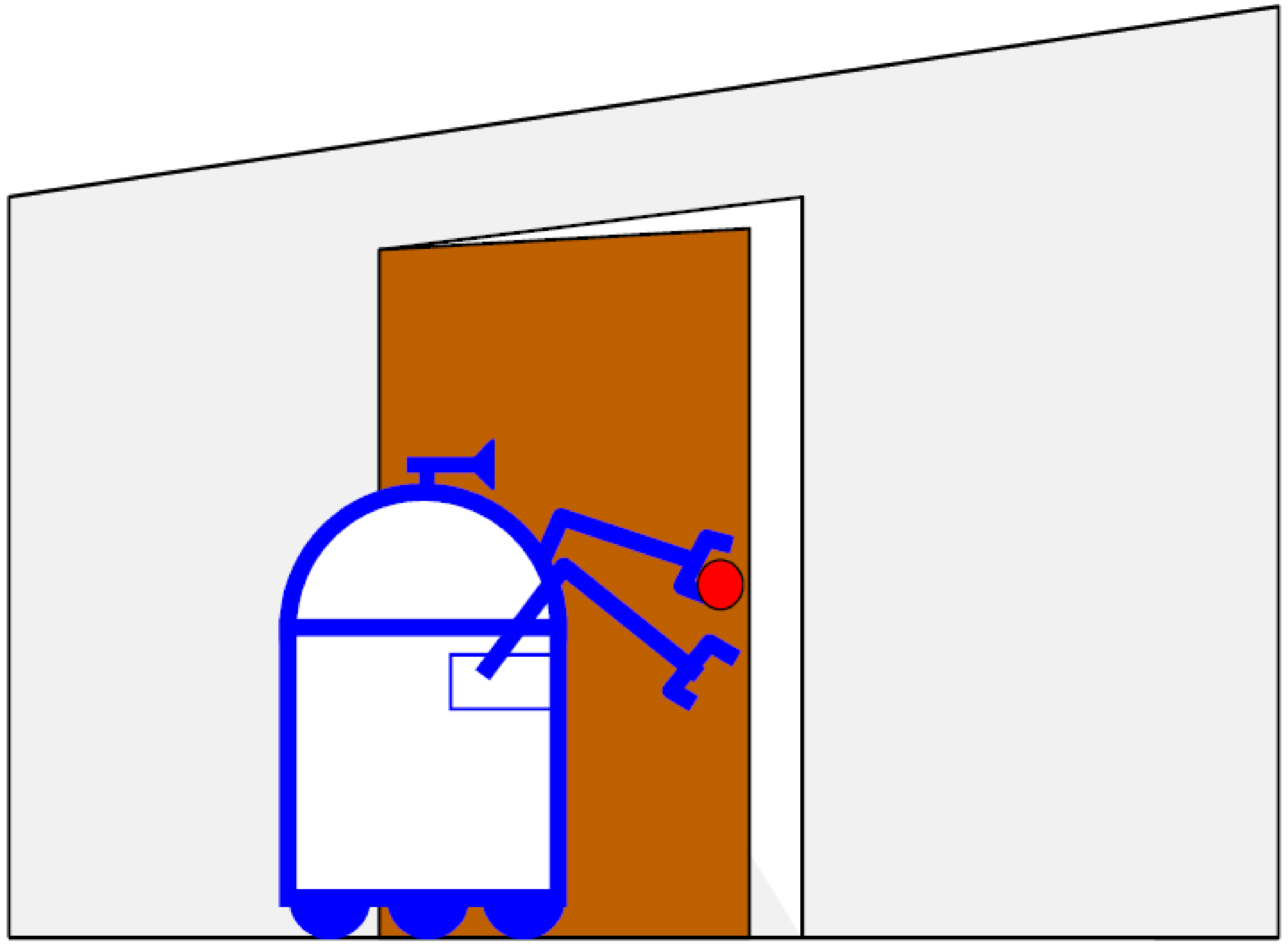
- To incorporate the outcome of an action *u* into the current "belief", we use the conditional pdf

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

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

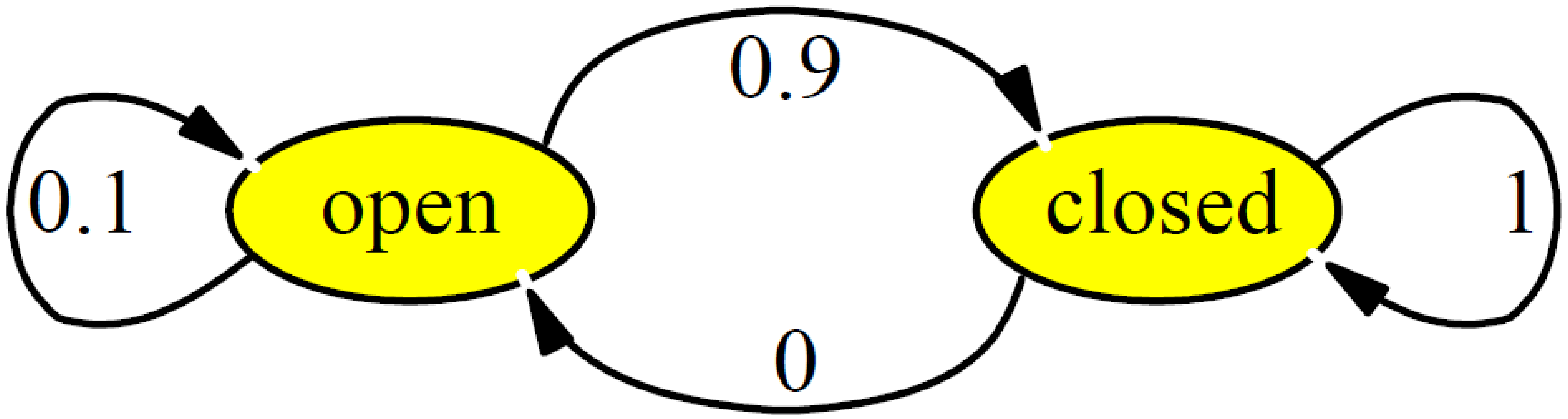


Example: Closing the door



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**:

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

$$bel(x_t) \quad = \quad p(x_t \mid z_{1:t}, u_{1:t}) \; . \tag{2.35}$$

This posterior is the probability distribution over the state x_t at time t, conditioned on all past measurements $z_{1:t}$ and all past controls $u_{1:t}$.

The reader may notice that we silently assume that the belief is taken *after* incorporating the measurement z_t. Occasionally, it will prove useful to calculate a posterior *before* incorporating z_t, just after executing the control u_t. Such a posterior will be denoted as follows:

$$\overline{bel}(x_t) \quad = \quad p(x_t \mid z_{1:t-1}, u_{1:t}) \tag{2.36}$$

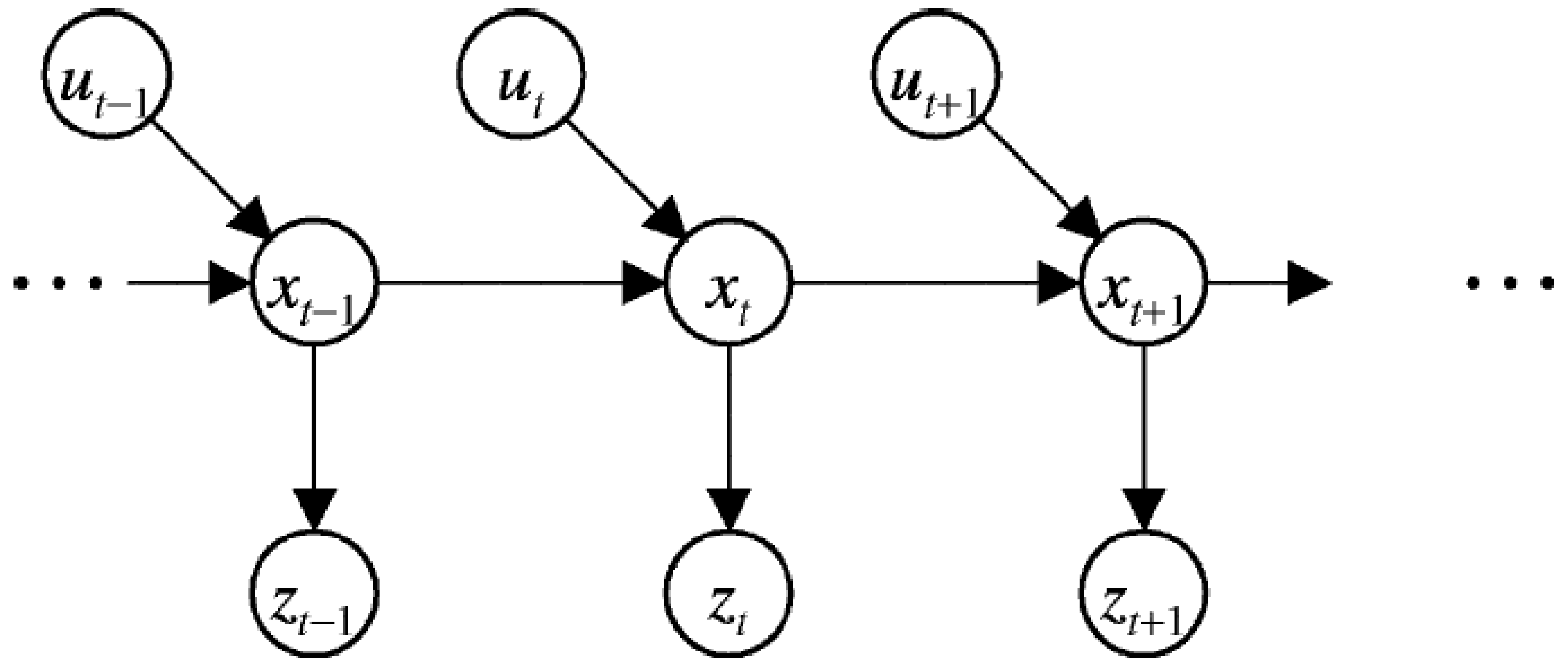
This probability distribution is often referred to as *prediction* in the context of probabilistic filtering. This terminology reflects the fact that $\overline{bel}(x_t)$ predicts the state at time t based on the previous state posterior, before incorporating the measurement at time t. Calculating $bel(x_t)$ from $\overline{bel}(x_t)$ is called *correction* or the *measurement update*.

A state x_t will be called *complete* if it is the best predictor of the future. Put differently, completeness entails that knowledge of past states, measurements, or controls carry no additional information that would help us to predict the future more accurately. It it important to notice that our definition of completeness does not require the future to be a *deterministic* function of the state. The future may be stochastic, but no variables prior to x_t may influence the stochastic evolution of future states, unless this dependence is mediated through the state x_t. Temporal processes that meet these conditions are commonly known as *Markov chains*.

A word is in order on the Markov assumption, or the complete state assumption, since it plays such a fundamental role in the material presented in this book. The Markov assumption postulates that past and future data are independent if one knows the current state x_t. To see how severe an assumption this is, let us consider our example of mobile robot localization. In mobile robot localization, x_t is the robot's pose, and Bayes filters are applied to estimate the pose relative to a fixed map. The following factors may have a systematic effect on sensor readings. Thus, they induce violations of the Markov assumption:

- Unmodeled dynamics in the environment not included in x_t (e.g., moving people and their effects on sensor measurements in our localization example),
- inaccuracies in the probabilistic models $p(z_t \mid x_t)$ and $p(x_t \mid u_t, x_{t-1})$,

Markov Assumption



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

Underlying Assumptions

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

If x_t is complete,

i) x_t is sufficient to predict the (potentially noisy) measurement z_t;

ii) It is sufficient summary of all that happened in the previous time steps. In particular, x_{t-1} is a sufficient statistic of all previous controls ad measurements up to this point in time

Bayes Filters

z = observation
u = action
x = state

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

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

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

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

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

Algorithm Bayes_filter($bel(x_{t-1}), u_t, z_t$**):**
  for all x_t do
    $\overline{bel}(x_t) = \int p(x_t \mid u_t, x_{t-1})\, bel(x_{t-1})\, dx$
    $bel(x_t) = \eta\, p(z_t \mid x_t)\, \overline{bel}(x_t)$
  endfor
  return $bel(x_t)$

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.

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY
2009

THANK YOU