CS 4610/5335 – Lecture 9
Localization
Kalman Filtering (1-D)

Lawson L.S. Wong
Northeastern University
2/16/22

Material adapted from:
1. Robert Platt, CS 4610/5335
2. Dan Klein & Pieter Abbeel, UC Berkeley CS 188
3. Sebastian Thrun, Wolfram Burgard, & Dieter Fox, Probabilistic Robotics

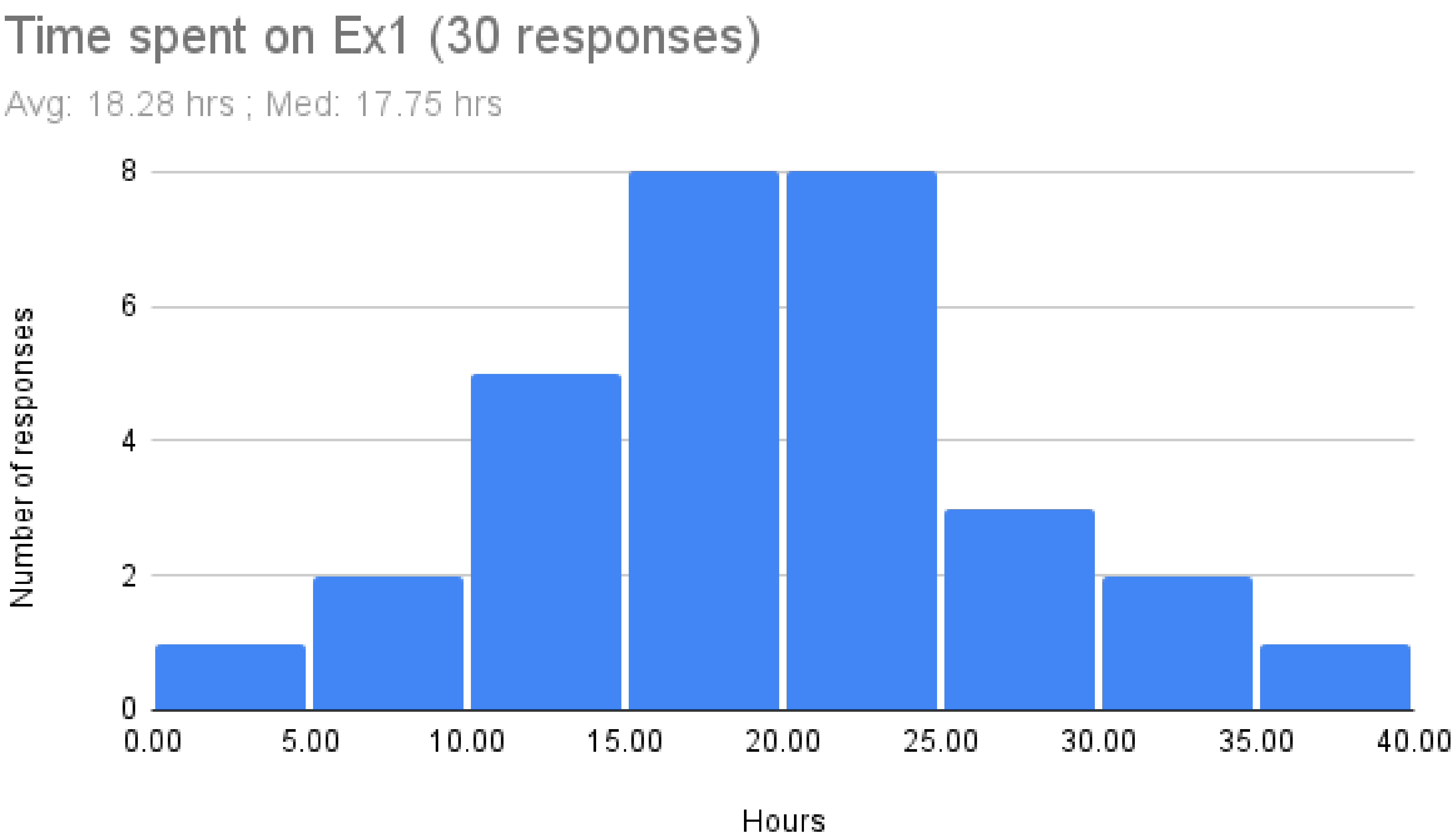
Announcements

Please continue gradually working on project thread

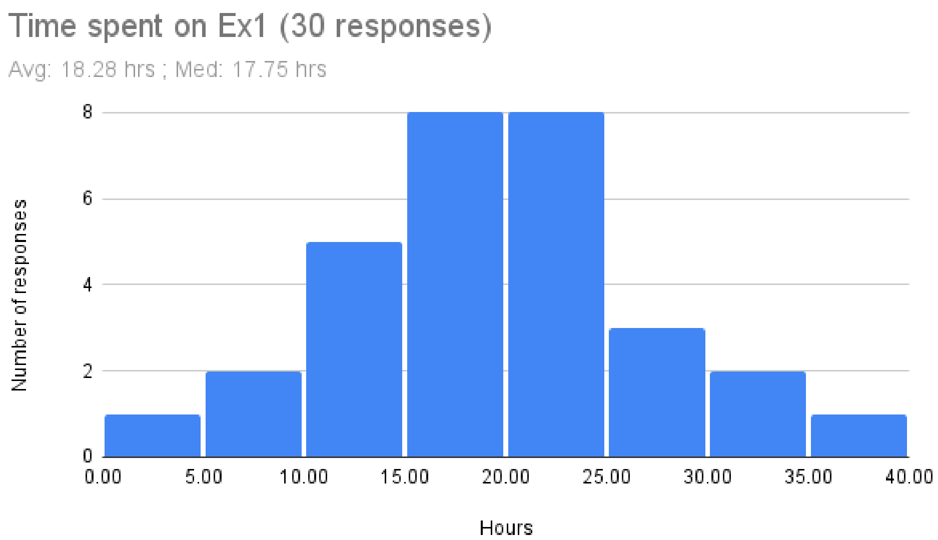
- Verify names (some missing / duplicates / typos)

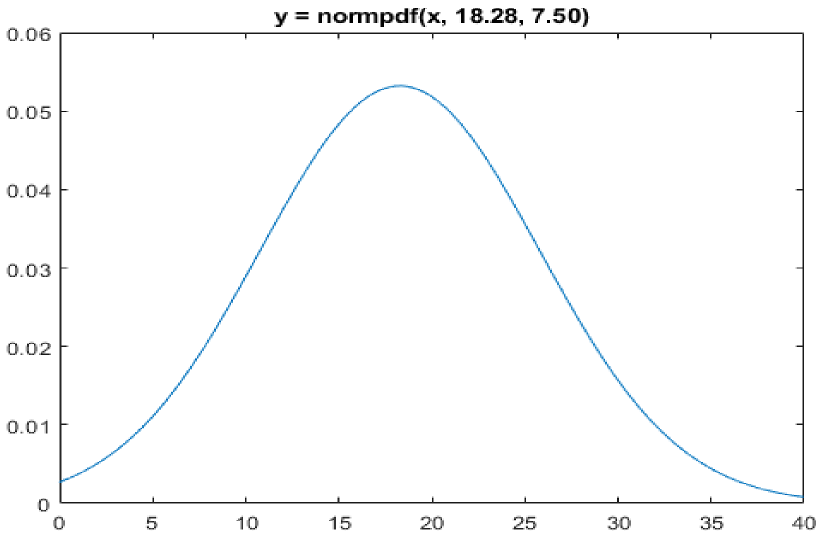
- Think even more about what hardware is necessary / acceptable
  - Will try to start ordering next week

Time spent on Ex1



Time spent on Ex1





Outline

Overview of localization

Discrete case (1-D)

Continuous case: Kalman filter (1-D)

1-D example

From “Probabilistic Robotics” (Ch. 7-8)
by Sebastian Thrun, Wolfram Burgard, & Dieter Fox

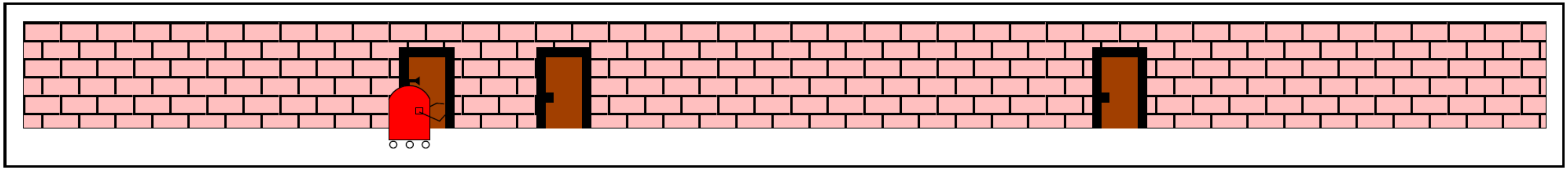


Figure 7.4 Example environment used to illustrate mobile robot localization: One-dimensional hallway environment with three indistinguishable doors. Initially the robot does not know its location except for its heading direction. Its goal is to find out where it is.

1-D example

From “Probabilistic Robotics” (Ch. 7-8)
by Sebastian Thrun, Wolfram Burgard, & Dieter Fox

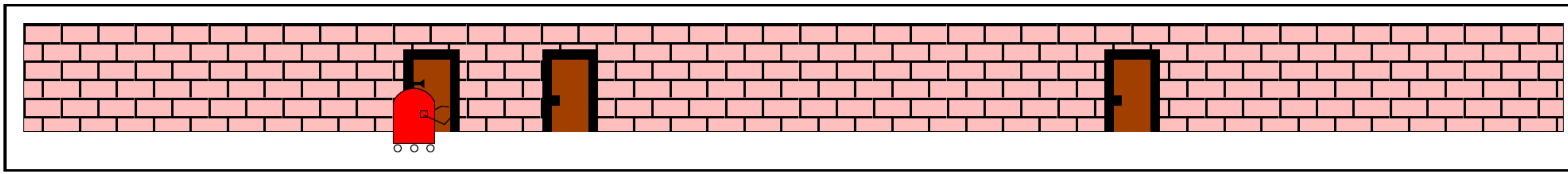


Figure 7.4 Example environment used to illustrate mobile robot localization: One-dimensional hallway environment with three indistinguishable doors. Initially the robot does not know its location except for its heading direction. Its goal is to find out where it is.

Challenge: Robot can only sense
wall/door at its <u>current</u> location

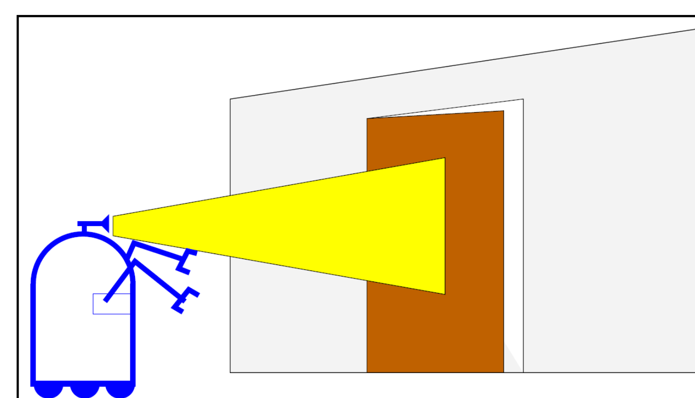


Figure 2.3 A mobile robot estimating the state of a door.

1-D example

From "Probabilistic Robotics" (Ch. 7-8)
by Sebastian Thrun, Wolfram Burgard, & Dieter Fox

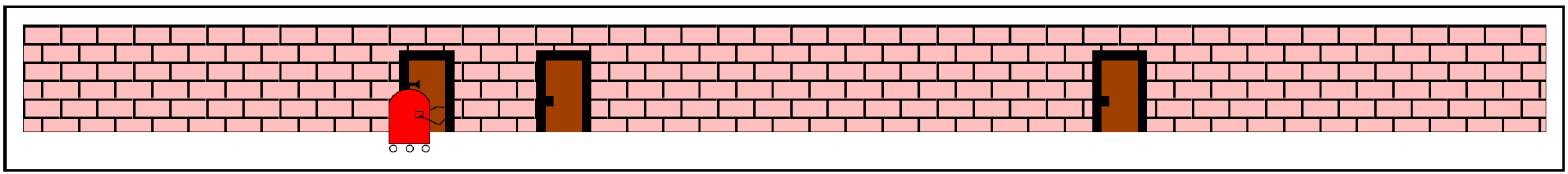


Figure 7.4 Example environment used to illustrate mobile robot localization: One-dimensional hallway environment with three indistinguishable doors. Initially the robot does not know its location except for its heading direction. Its goal is to find out where it is.

Challenge: Robot can only sense
wall/door at its <u>current</u> location

Robot starts somewhere random,
keeps on heading right,
detects Wall, Door, Wall, Wall, …
Where is it now?

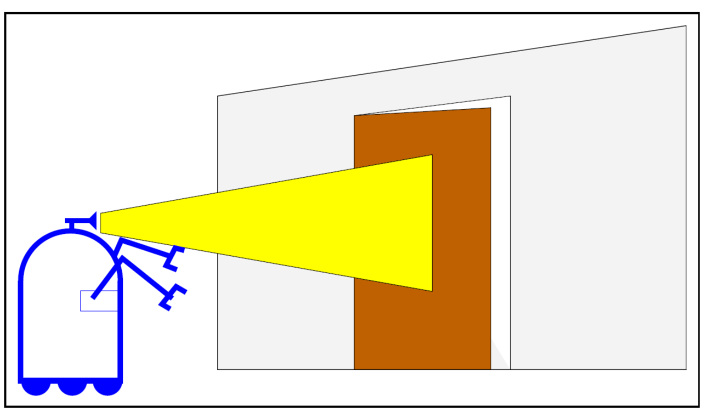


Figure 2.3 A mobile robot estimating the state of a door.

1-D example

From “Probabilistic Robotics” (Ch. 7-8)
by Sebastian Thrun, Wolfram Burgard, & Dieter Fox

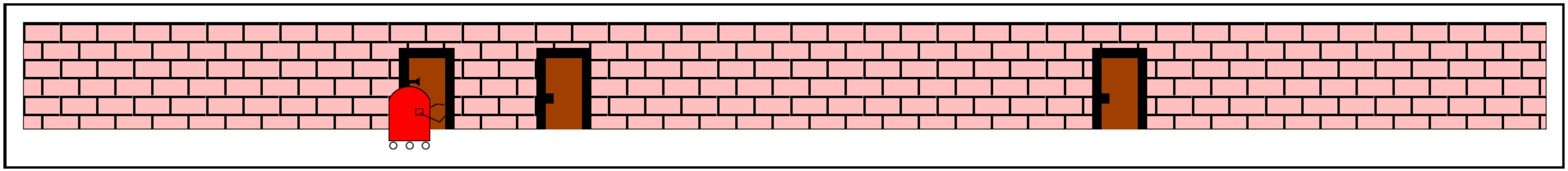


Figure 7.4 Example environment used to illustrate mobile robot localization: One-dimensional hallway environment with three indistinguishable doors. Initially the robot does not know its location except for its heading direction. Its goal is to find out where it is.

Robot starts somewhere random, keeps on heading right.
Where is it after:

W W D W D W W W
W W W D W W W

1-D example

From “Probabilistic Robotics” (Ch. 7-8)
by Sebastian Thrun, Wolfram Burgard, & Dieter Fox

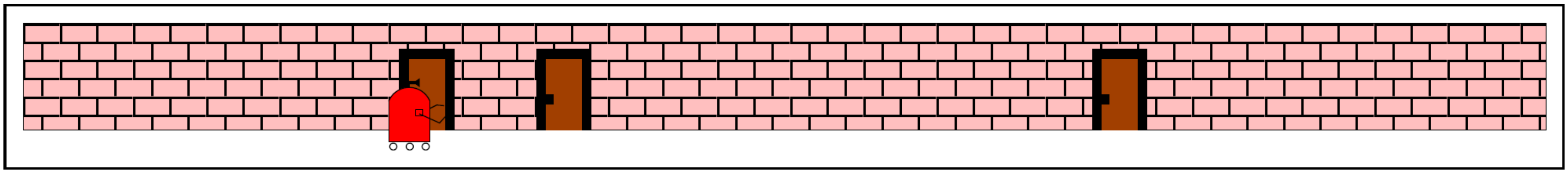


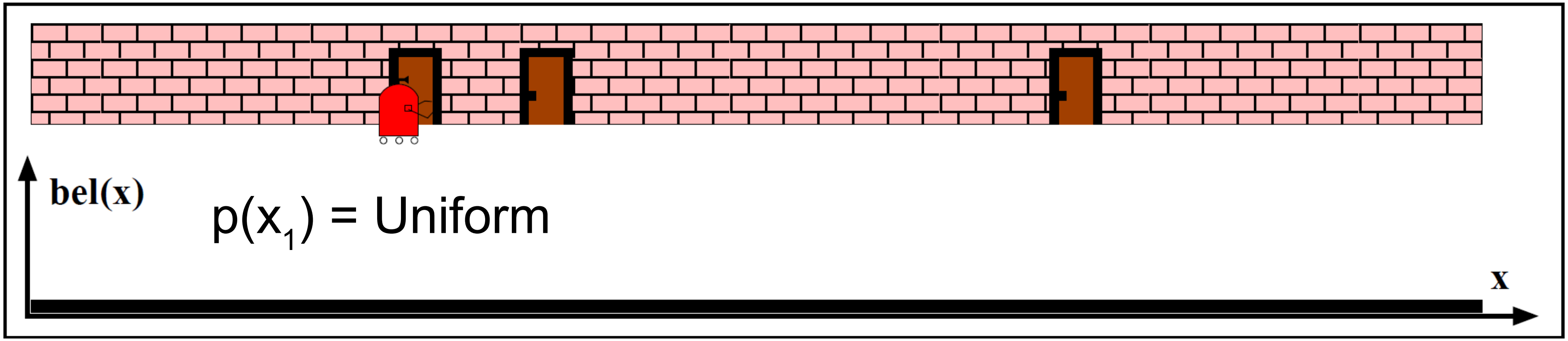
Figure 7.4 Example environment used to illustrate mobile robot localization: One-dimensional hallway environment with three indistinguishable doors. Initially the robot does not know its location except for its heading direction. Its goal is to find out where it is.

Robot starts somewhere random, keeps on heading right.
Where is it after:

W W D W D W W W
W W W D W W W
W W W D D D W W W
W D W D D W W D D D W W W

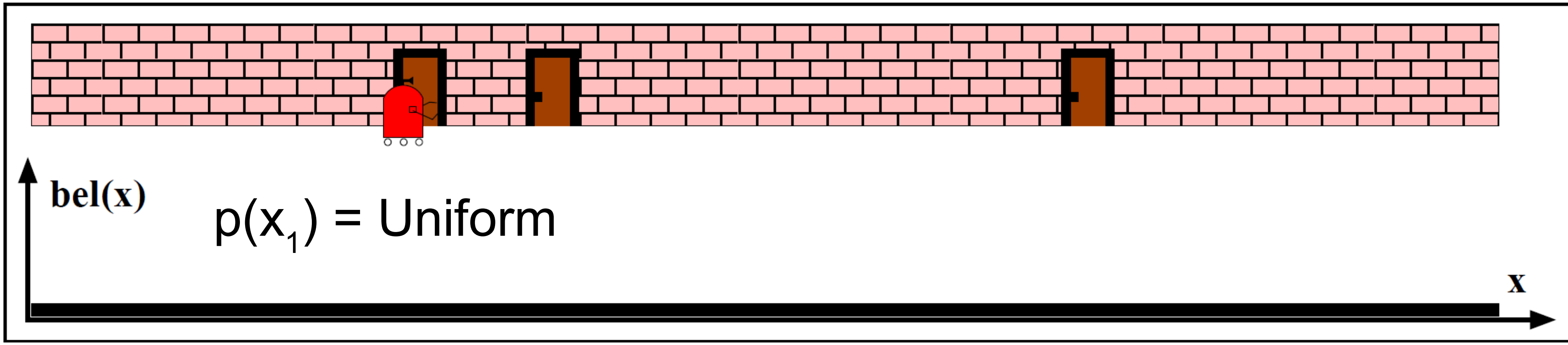
1-D example (Init)

(a)

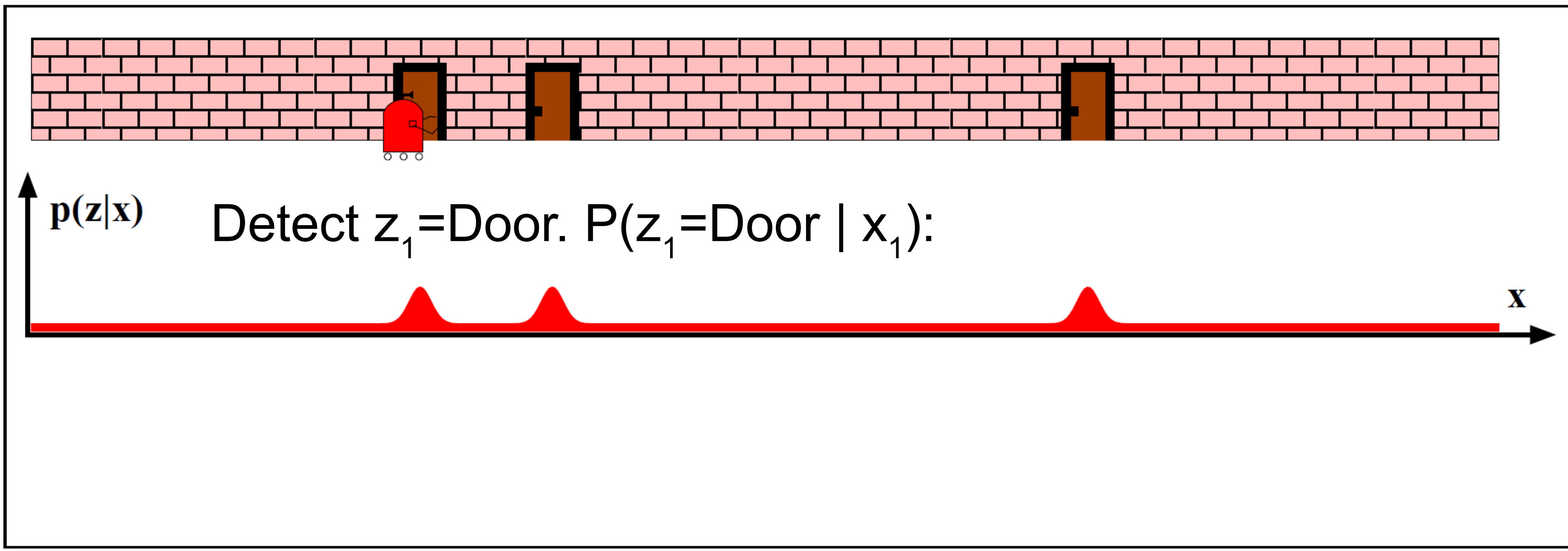


1-D example (Observation)

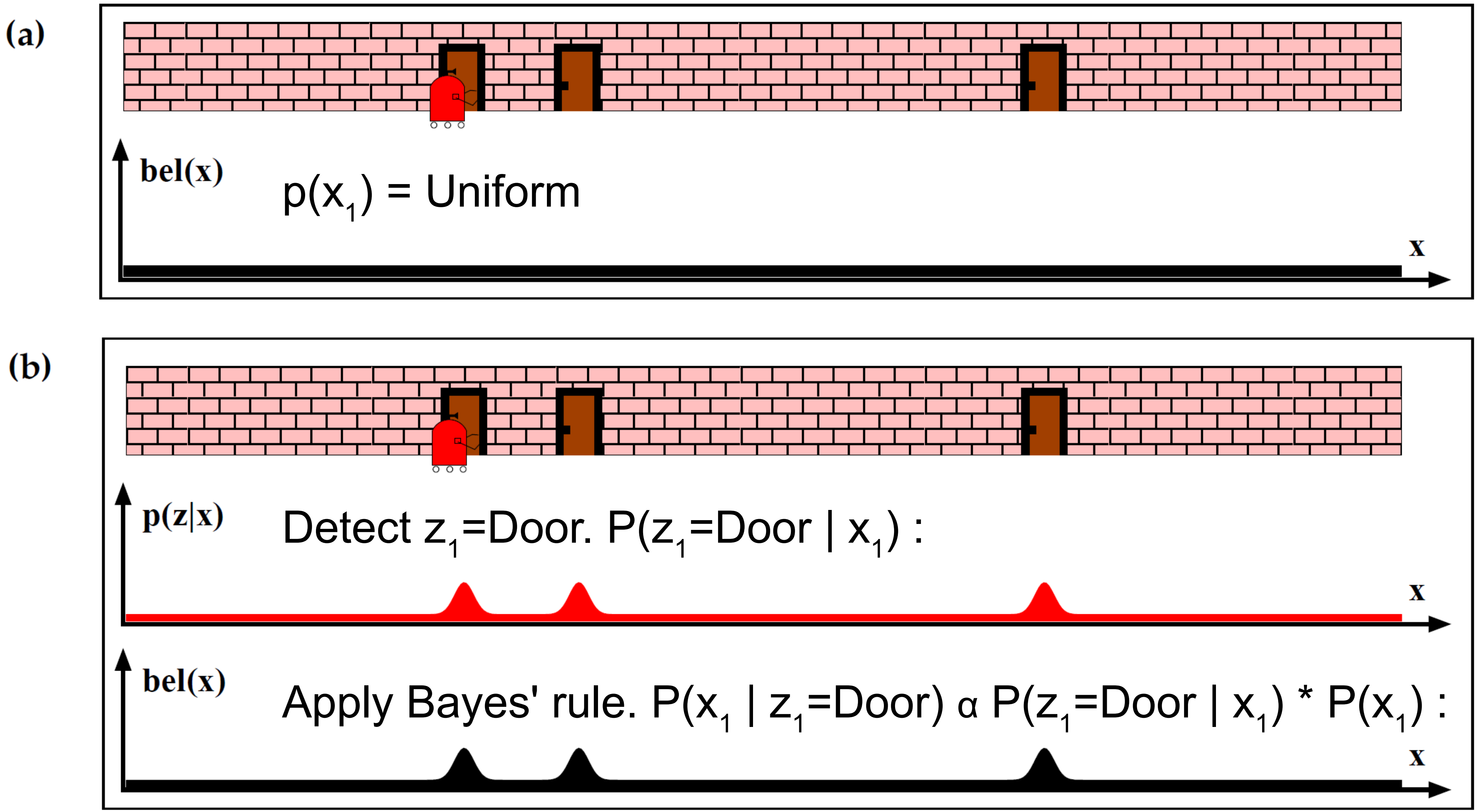
(a)



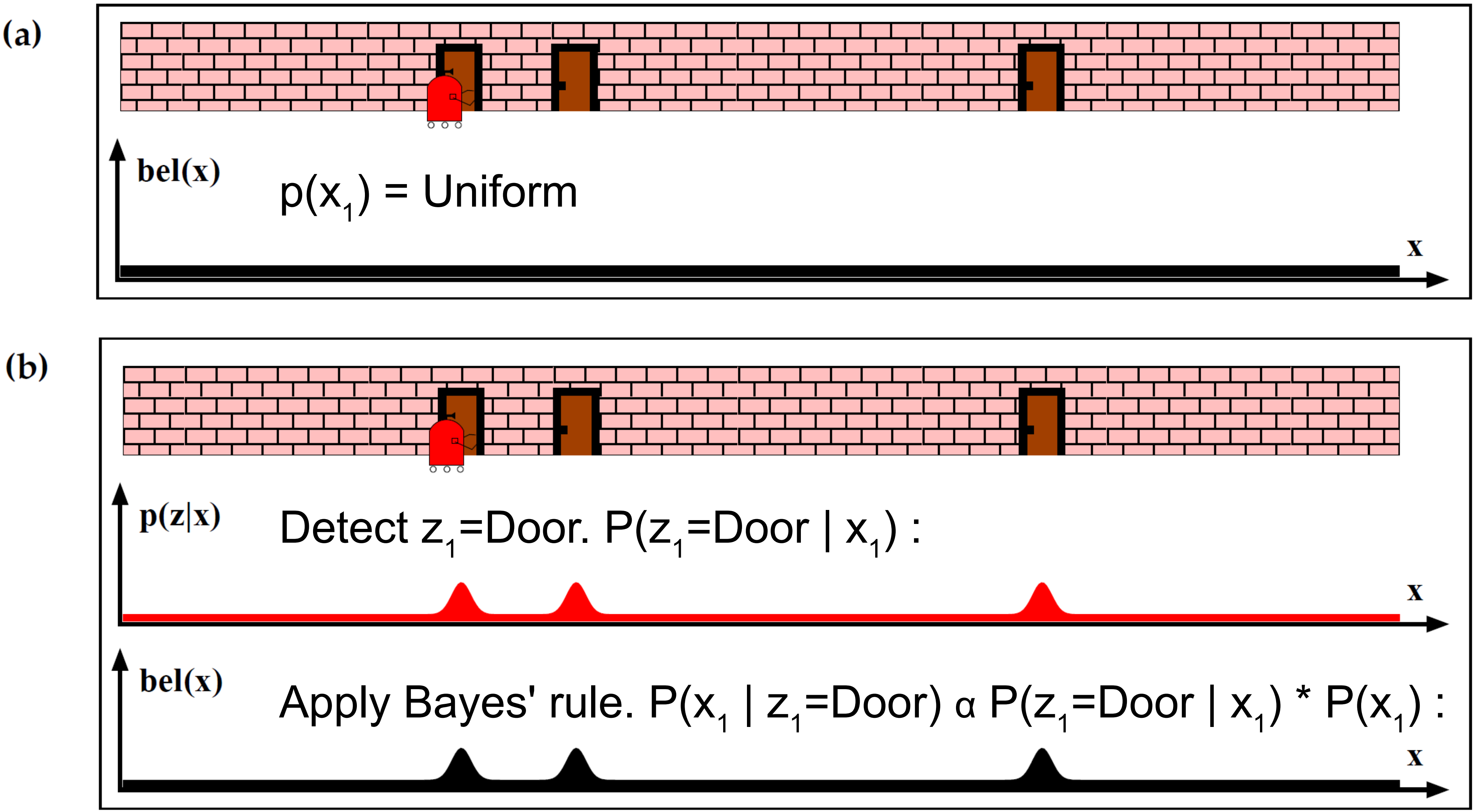
(b)



1-D example (Posterior update)



1-D example (Posterior update)

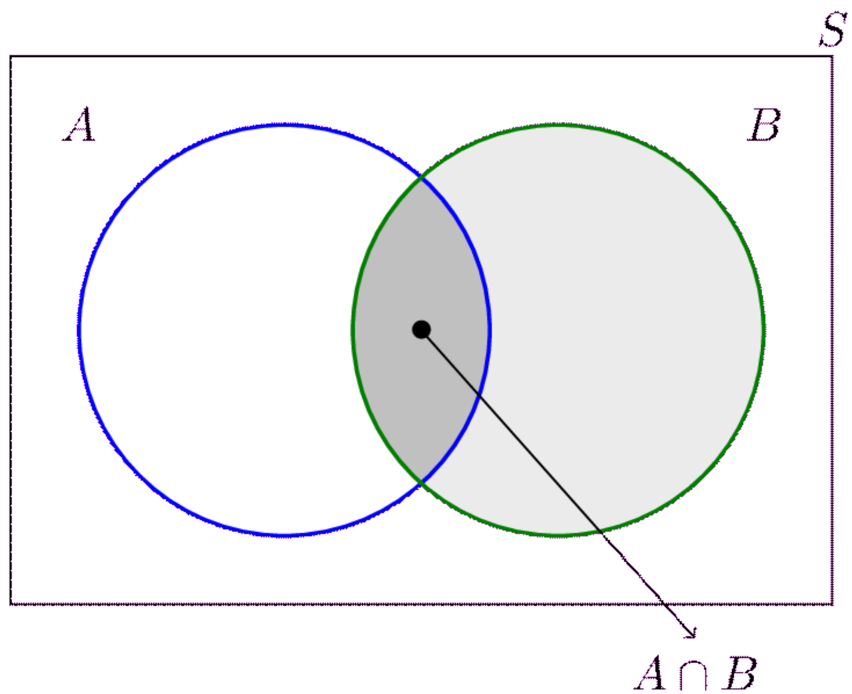


Normalization step not shown explicitly:

- Calculate RHS for all values of x_1
- Divide RHS by the sum (such that posterior sums to 1)

Probability reminder: Conditional Probabilities

<u>Conditional probability</u>: $P(A \mid B) = \dfrac{P(A,B)}{P(B)}$ (if P(*B*)>0)



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

Probability reminder: Bayes' Rule

$$P(a|b) = \frac{P(b|a)P(a)}{P(b)}$$

It's easy to derive from the product rule:

$$P(a,b) = P(b|a)P(a) = P(a|b)P(b)$$

Solve for this

Bayesian inference

Interpretation:

$$\underset{P(\text{state} \mid \text{obs})}{\textbf{Posterior}} = \frac{\overset{\textbf{Likelihood * Prior}}{P(\text{obs} \mid \text{state})\, P(\text{state})}}{\underset{\textbf{Evidence}}{P(\text{obs})}}$$

Bayesian inference

Interpretation: **Likelihood * Prior**

Posterior P(obs | state) P(state)

P(state | obs) = ---------------------------------------

P(obs)

Evidence

$$P(\text{state} \mid \text{obs}) = \frac{P(\text{obs} \mid \text{state})\, P(\text{state})}{P(\text{obs})}$$

E.g., State = Is it raining outside right now?

- Prior: It rains on 30% of days in Boston
- P(rain) = 0.3 ; P(not rain) = 1-0.3 = 0.7

Bayesian inference

Interpretation:

$$\underbrace{P(\text{state} \mid \text{obs})}_{\textbf{Posterior}} = \frac{\overbrace{P(\text{obs} \mid \text{state})\, P(\text{state})}^{\textbf{Likelihood * Prior}}}{\underbrace{P(\text{obs})}_{\textbf{Evidence}}}$$

E.g., State = Is it raining outside right now?

- Prior: It rains on 30% of days in Boston
- P(rain) = 0.3 ; P(not rain) = 1-0.3 = 0.7

Observation = People are carrying wet umbrellas

- P(obs | rain) = 0.9 ; P(obs | not rain) = 0.2

Bayesian inference

Interpretation:

$$\underbrace{P(\text{state} \mid \text{obs})}_{\textbf{Posterior}} = \frac{\overbrace{P(\text{obs} \mid \text{state})\, P(\text{state})}^{\textbf{Likelihood * Prior}}}{\underbrace{P(\text{obs})}_{\textbf{Evidence}}}$$

E.g., State = Is it raining outside right now?

- Prior: It rains on 30% of days in Boston
- P(rain) = 0.3 ; P(not rain) = 1-0.3 = 0.7

Observation = People are carrying wet umbrellas

- P(obs | rain) = 0.9 ; P(obs | not rain) = 0.2

Bayes' rule: P(rain | obs) α P(obs | rain) * P(rain)

- P(rain | obs) α 0.9 * 0.3 = 0.27
- P(not rain | obs) = ?

Bayesian inference

Interpretation:

$$\underbrace{P(\text{state} \mid \text{obs})}_{\textbf{Posterior}} = \frac{\overbrace{P(\text{obs} \mid \text{state})\, P(\text{state})}^{\textbf{Likelihood * Prior}}}{\underbrace{P(\text{obs})}_{\textbf{Evidence}}}$$

E.g., State = Is it raining outside right now?

- Prior: It rains on 30% of days in Boston
- P(rain) = 0.3 ; P(not rain) = 1-0.3 = 0.7

Observation = People are carrying wet umbrellas

- P(obs | rain) = 0.9 ; P(obs | not rain) = 0.2

Bayes' rule: P(rain | obs) α P(obs | rain) * P(rain)

- P(rain | obs) α 0.9 * 0.3 = 0.27
- P(not rain | obs) α P(obs | not rain) * P(not rain)
  = 0.2 * 0.7 = 0.14

Bayesian inference

Interpretation:

$$\underbrace{P(\text{state} \mid \text{obs})}_{\textbf{Posterior}} = \frac{\overbrace{P(\text{obs} \mid \text{state})\, P(\text{state})}^{\textbf{Likelihood * Prior}}}{\underbrace{P(\text{obs})}_{\textbf{Evidence}}}$$

E.g., State = Is it raining outside right now?

- Prior: It rains on 30% of days in Boston
- P(rain) = 0.3 ; P(not rain) = 1-0.3 = 0.7

Observation = People are carrying wet umbrellas

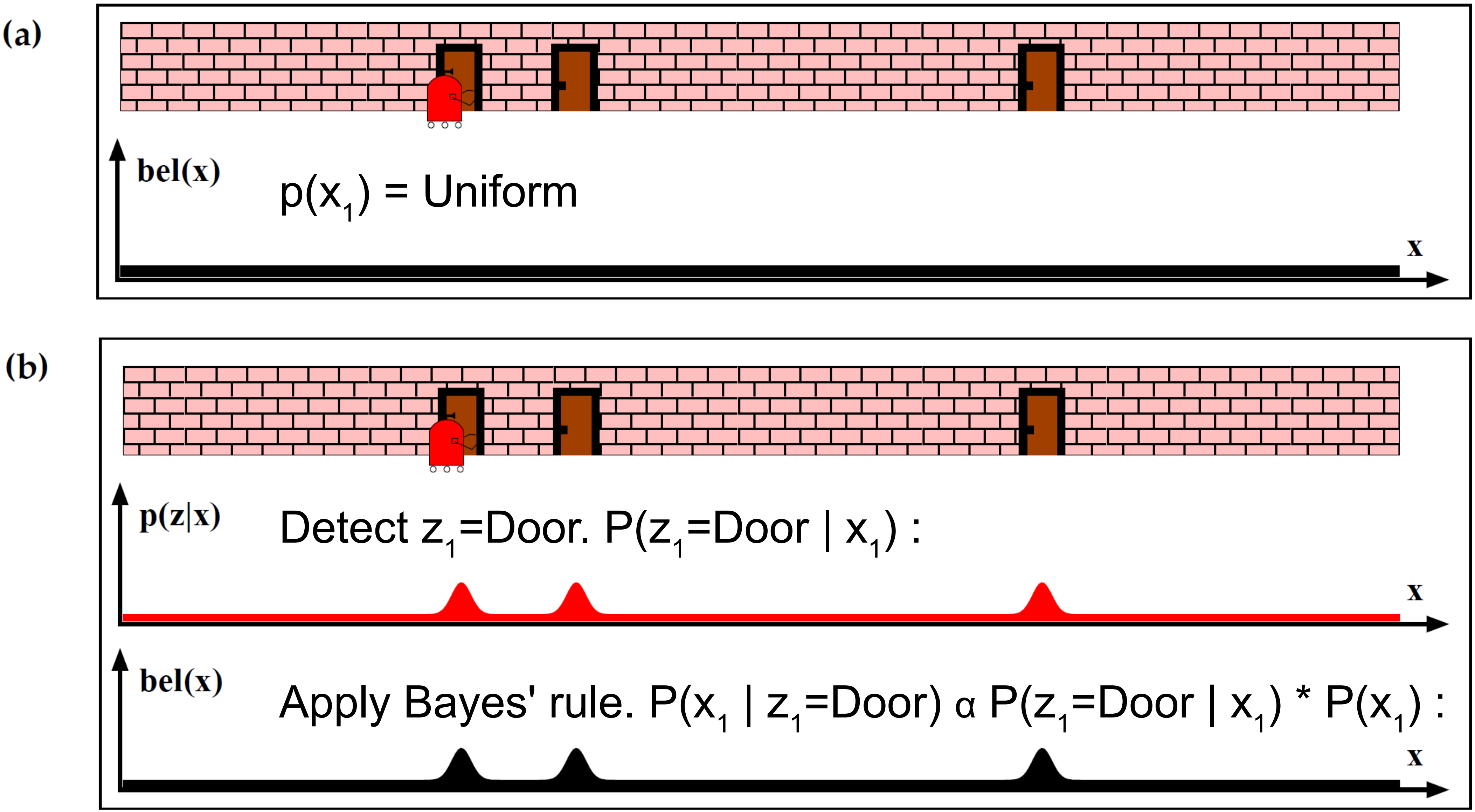
- P(obs | rain) = 0.9 ; P(obs | not rain) = 0.2

Bayes' rule: P(rain | obs) α P(obs | rain) * P(rain)

- P(rain | obs) α 0.9 * 0.3 = 0.27
- P(not rain | obs) α P(obs | not rain) * P(not rain) = 0.2 * 0.7 = 0.14

<u>Normalize</u> to get answer: P(rain | obs) = 0.27 / (0.27+0.14) ≈ 0.66

1-D example (Posterior update)

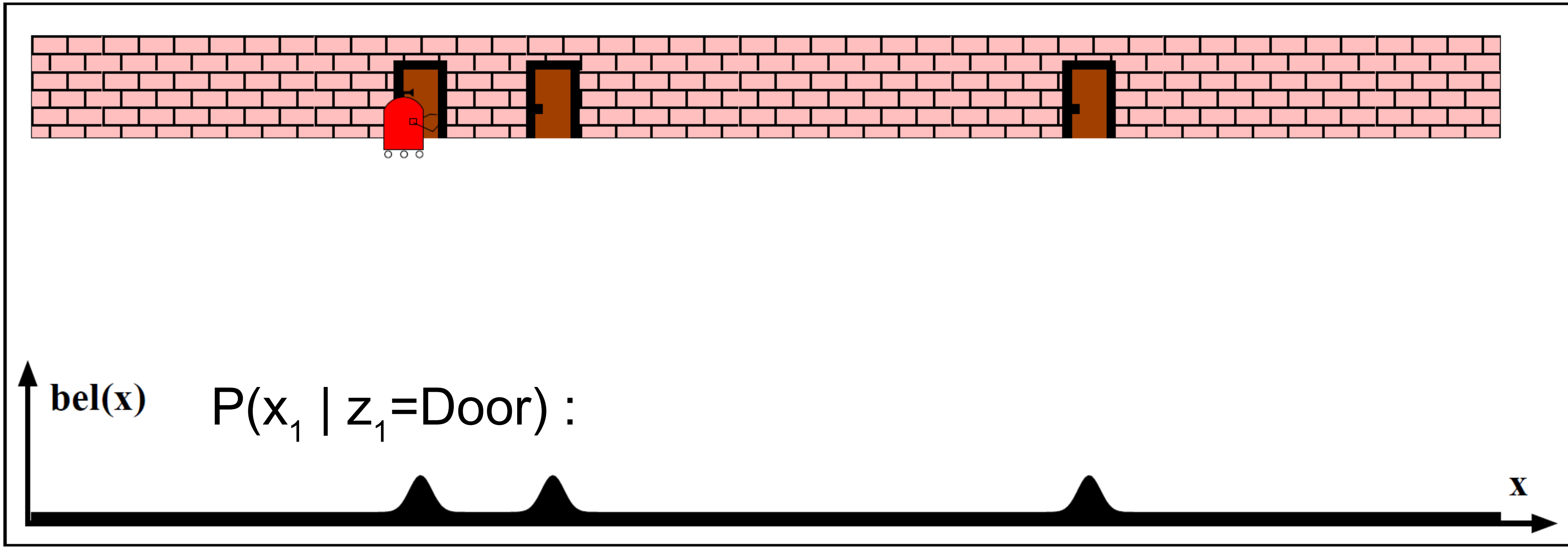


Normalization step not shown explicitly:

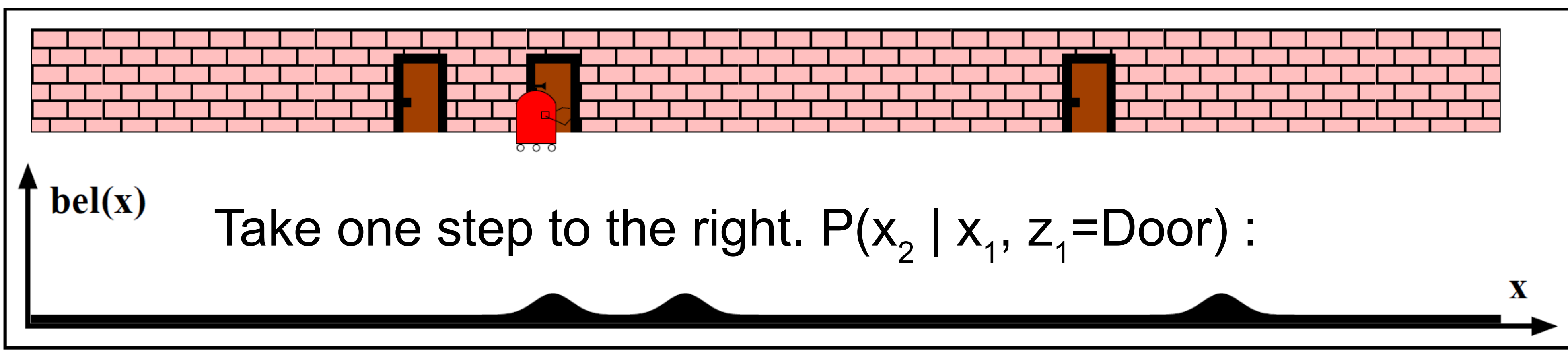
- Calculate RHS for all values of x_1
- Divide RHS by the sum (such that posterior sums to 1)

1-D example (Transition)

(b)

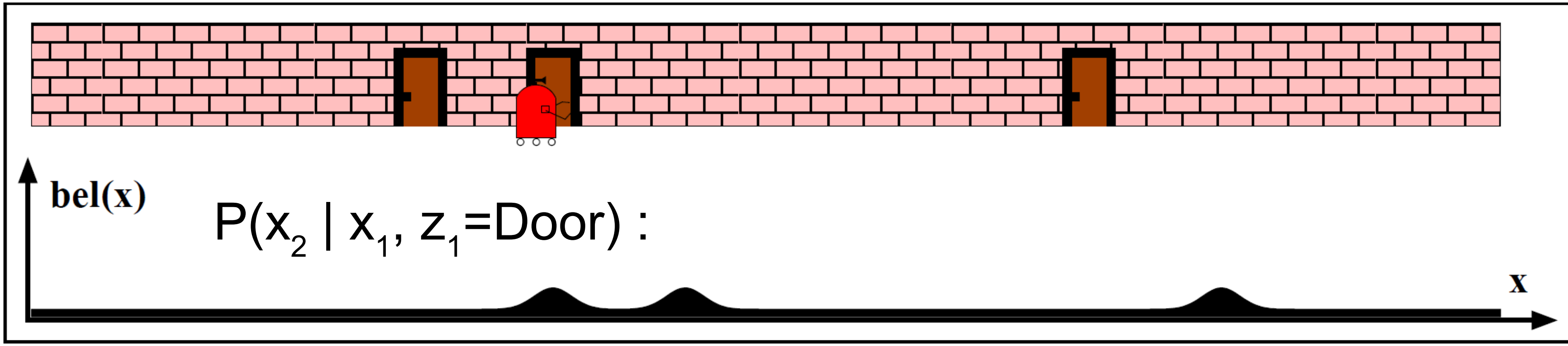


(c)

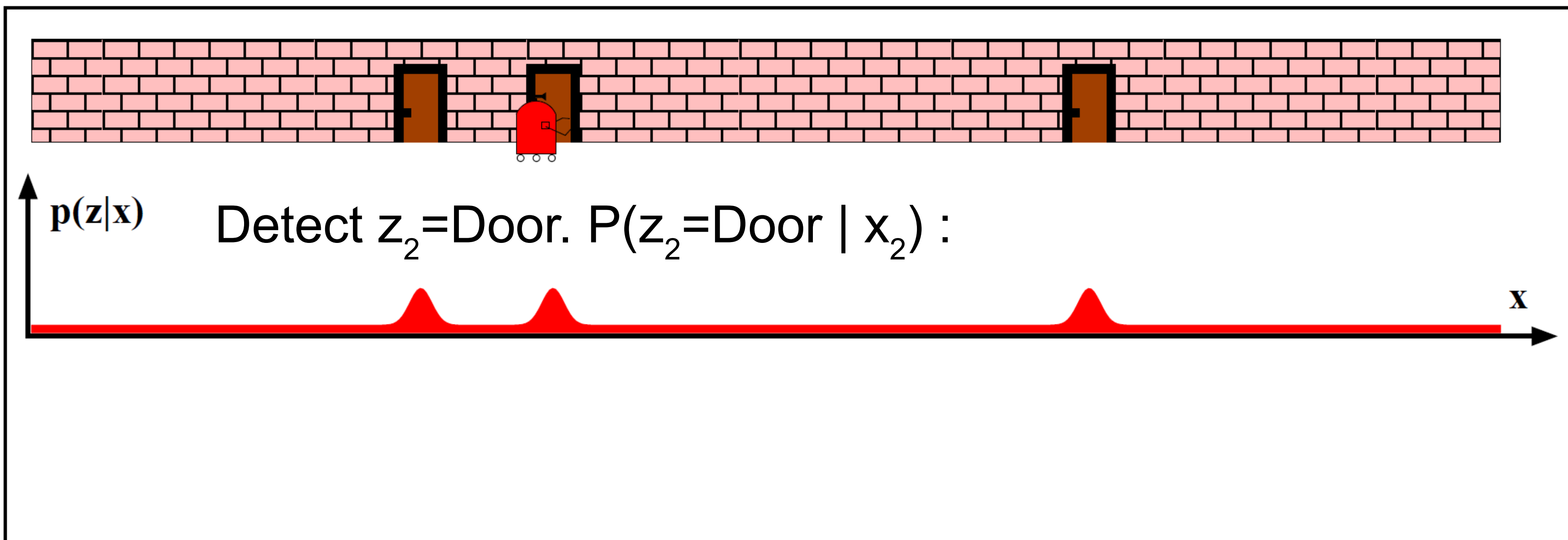


1-D example (Observation)

(c)

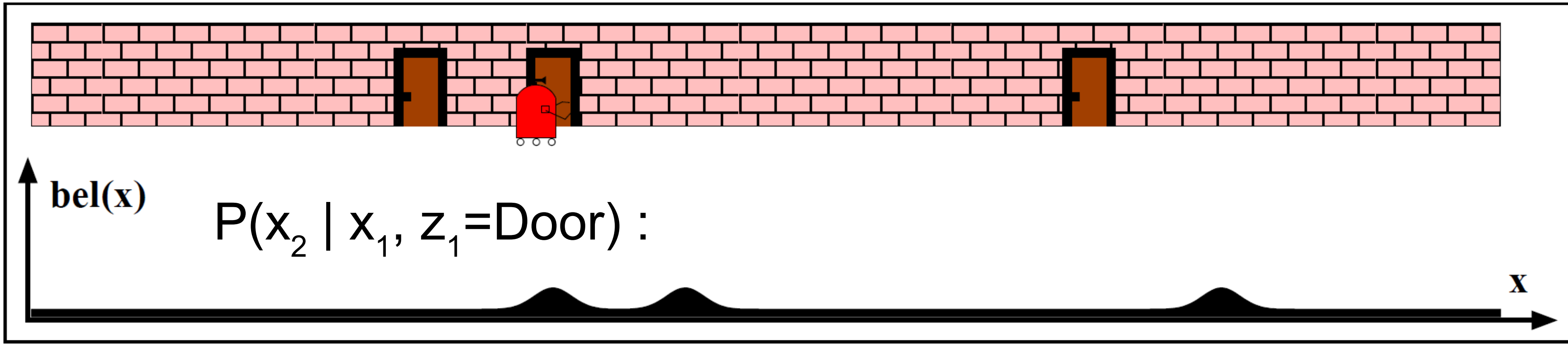


(d)

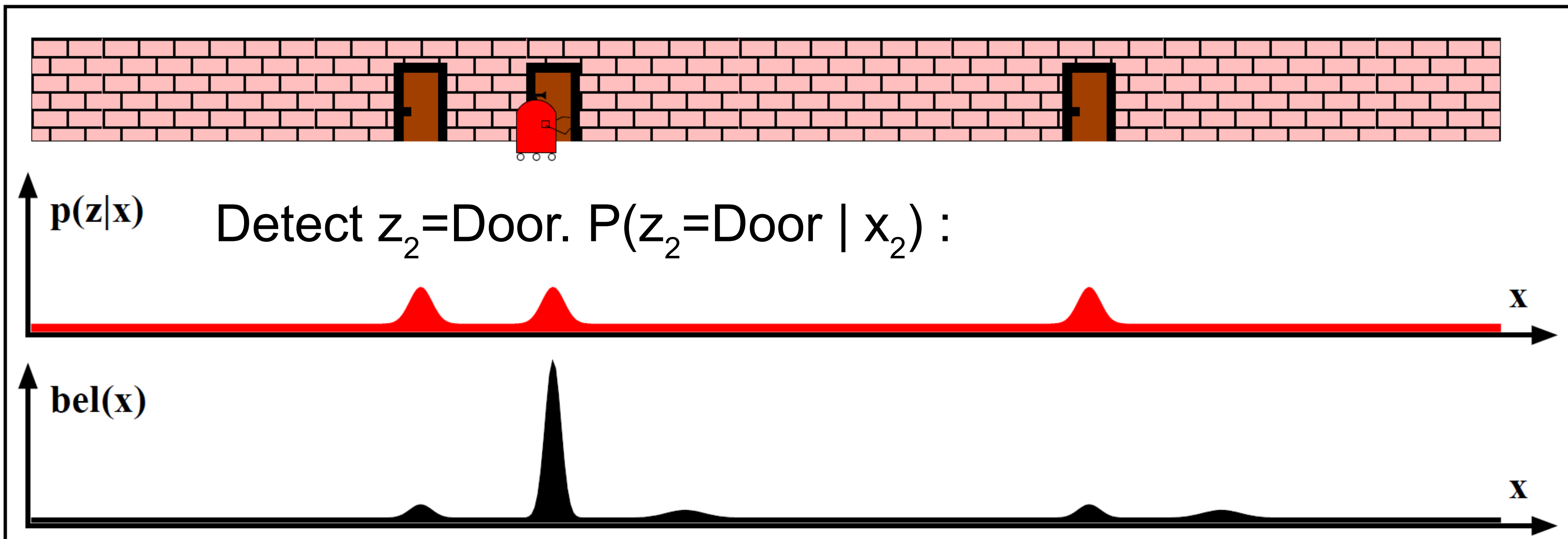


1-D example (Posterior update)

(c)



(d)

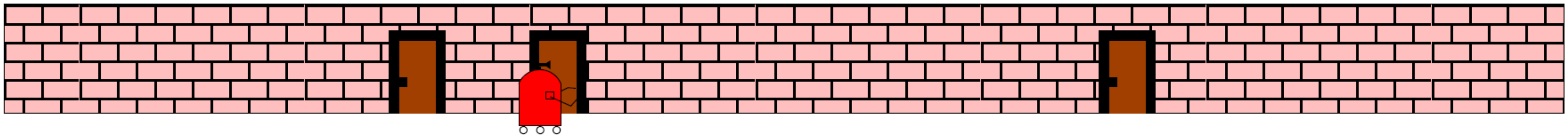


Apply Bayes' rule and <u>normalize</u> (not shown explicitly):

$P(x_2 \mid z_2=\text{Door}, z_1=\text{Door}) \propto P(z_2=\text{Door} \mid x_2) * P(x_2 \mid x_1, z_1=\text{Door})$

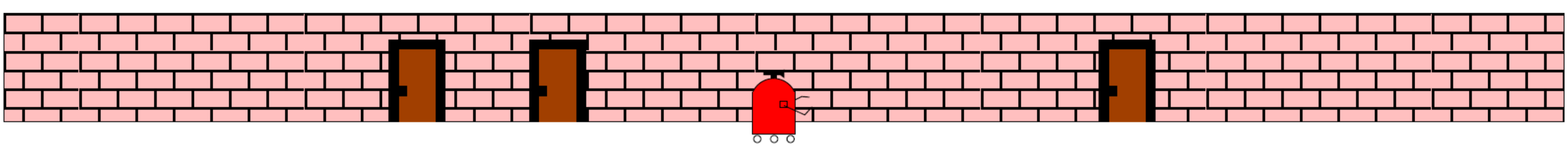
1-D example (Transition)

(d)

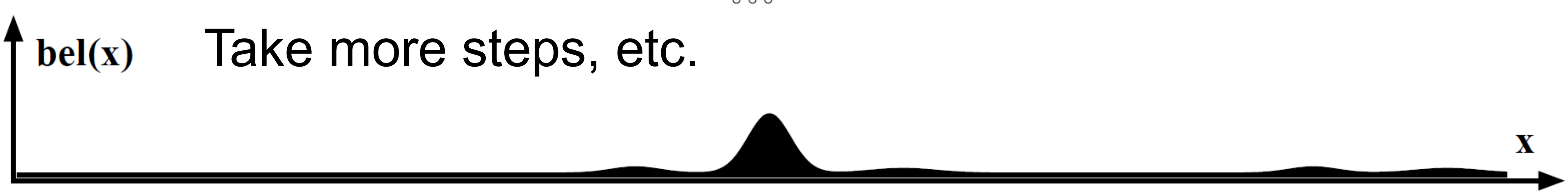


P(x_2 | z_2=Door, z_1=Door) :

(e)

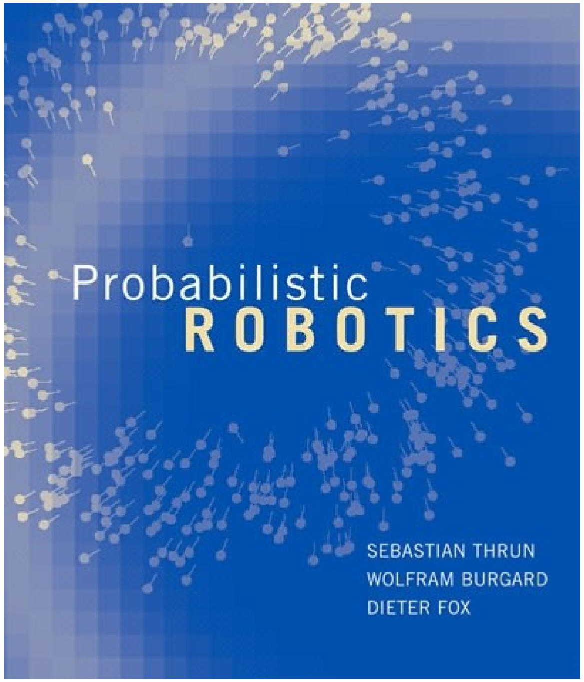


Take more steps, etc.



Mobile robot localization (Hallway example)

From
"Probabilistic Robotics"
(Ch. 7-8)
by
Sebastian Thrun,
Wolfram Burgard,
& Dieter Fox



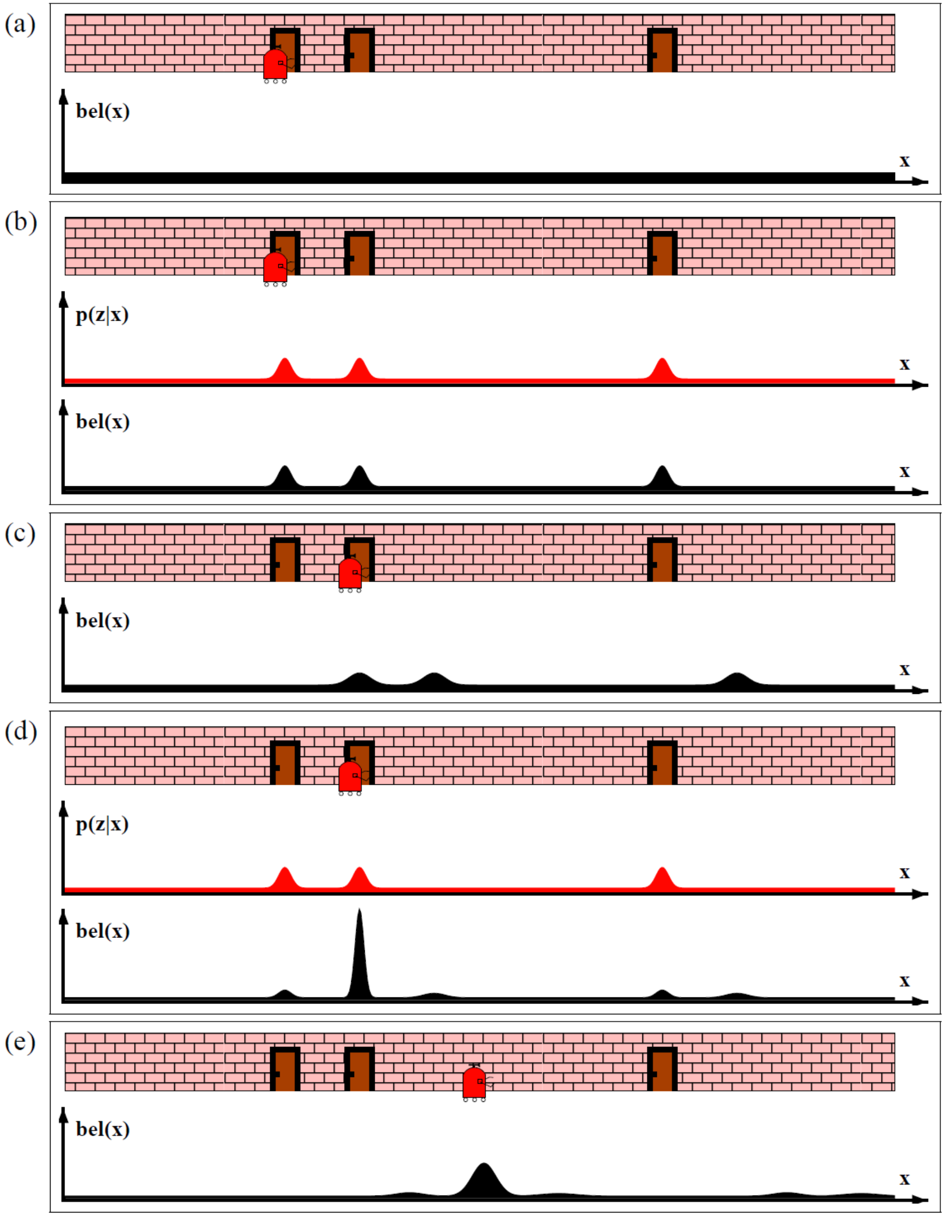


Figure 7.4 Illustration of the Markov localization algorithm. Each picture depicts the position of the robot in the hallway and its current belief $bel(x)$. (b) and (d) additionally depict the observation model $p(z_t \mid x_t)$, which describes the probability of observing a door at the different locations in the hallway.

Mobile robot localization (Hallway example)

From
“Probabilistic Robotics”
(Ch. 7-8)
by
Sebastian Thrun,
Wolfram Burgard,
& Dieter Fox

This is an important problem!
- Ch. 2-4 (half of Part 1)
  is about inference
- Ch. 7-8 (entire Part 2!)
  is about localization
- There are 4 parts,
  17 chapters in total

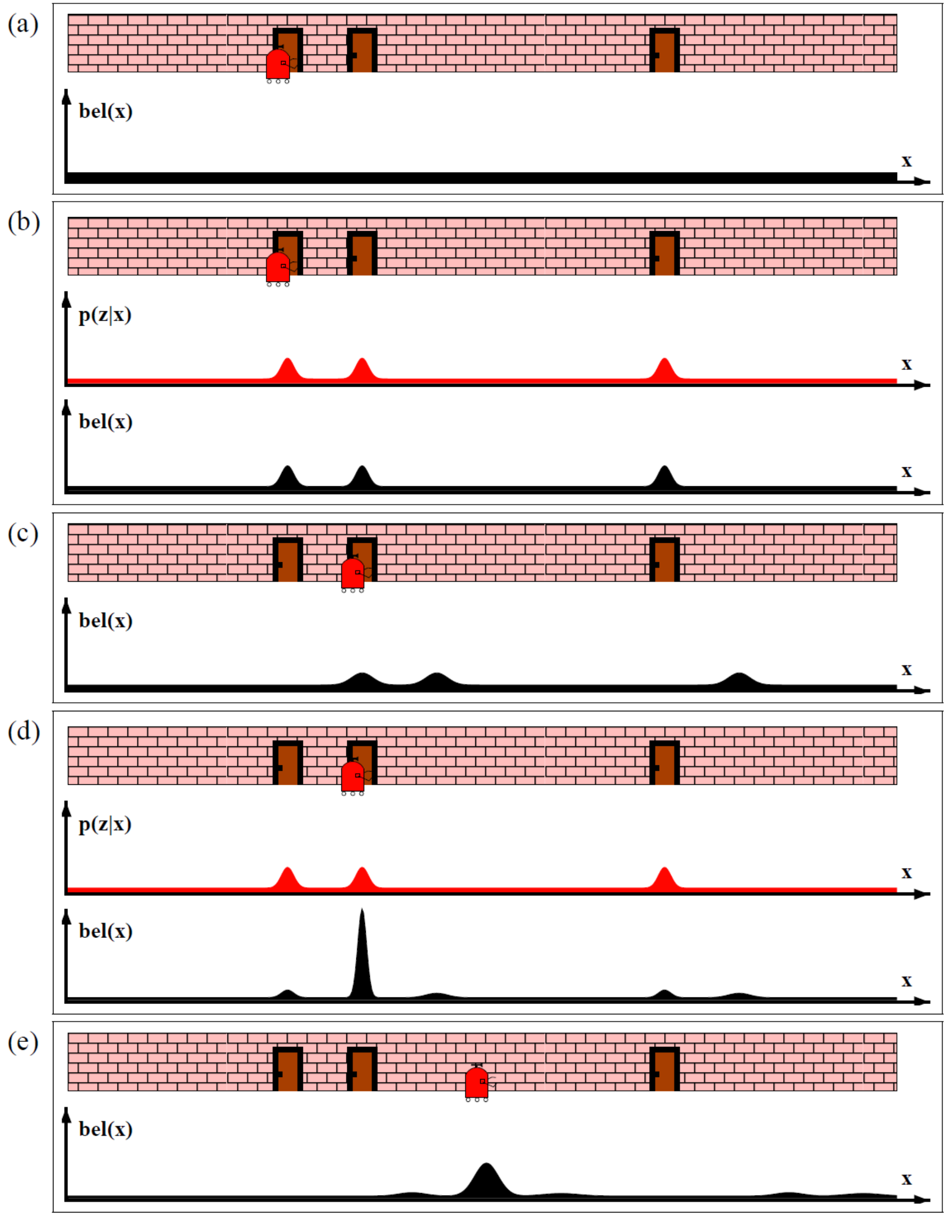


Figure 7.4 Illustration of the Markov localization algorithm. Each picture depicts the position of the robot in the hallway and its current belief $bel(x)$. (b) and (d) additionally depict the observation model $p(z_t \mid x_t)$, which describes the probability of observing a door at the different locations in the hallway.

Mobile robot localization (Hallway example)

From
“Probabilistic Robotics”
(Ch. 7-8)
by
Sebastian Thrun,
Wolfram Burgard,
& Dieter Fox

Continuous case;
Ideal outcome

We will study this extensively with the Kalman filter

– Get familiar with Gaussian distributions!

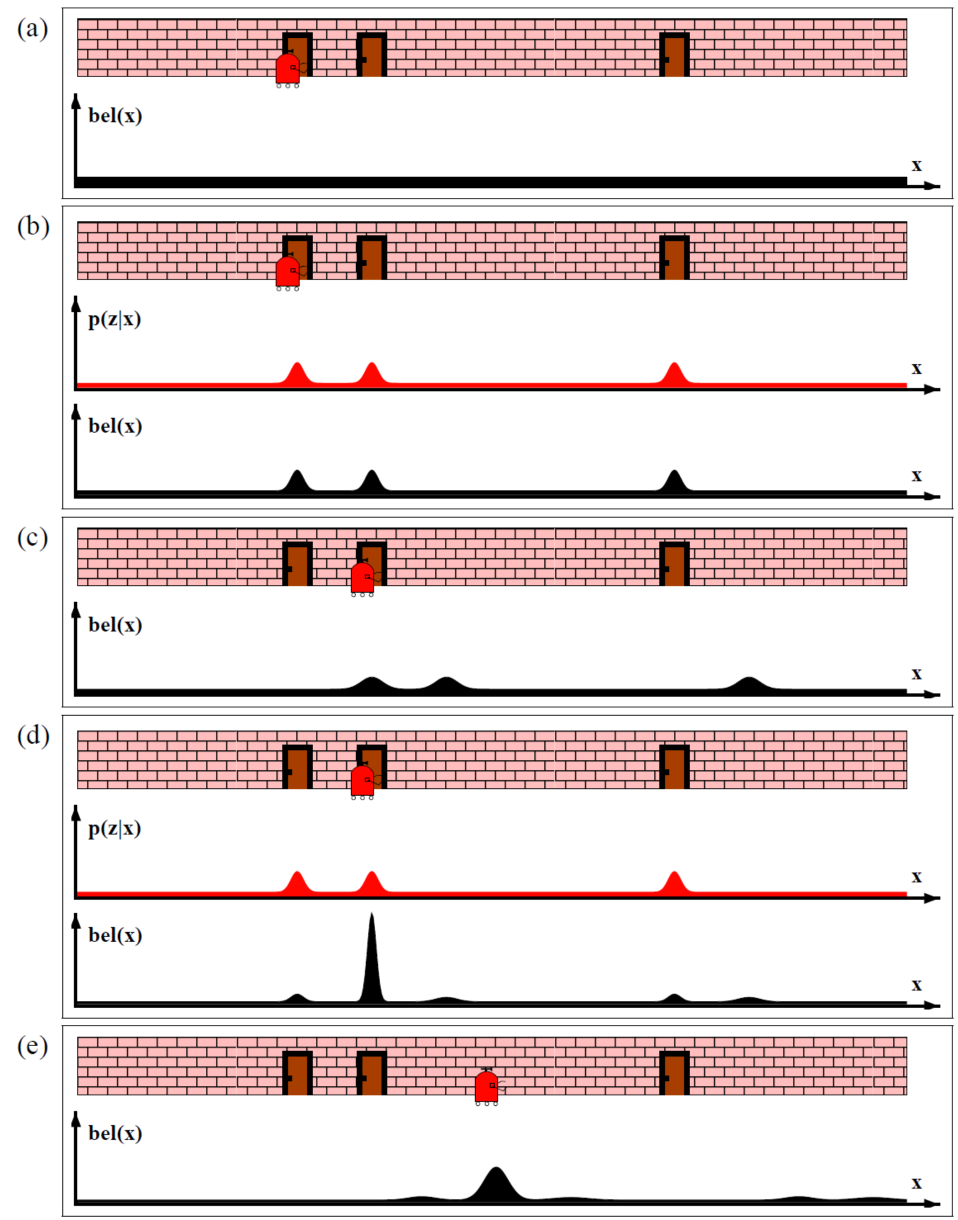


Figure 7.4 Illustration of the Markov localization algorithm. Each picture depicts the position of the robot in the hallway and its current belief $bel(x)$. (b) and (d) additionally depict the observation model $p(z_t \mid x_t)$, which describes the probability of observing a door at the different locations in the hallway.

Mobile robot localization (Hallway example)

From
"Probabilistic Robotics"
(Ch. 7-8)
by
Sebastian Thrun,
Wolfram Burgard,
& Dieter Fox

Up next:
Discrete case (1-D)

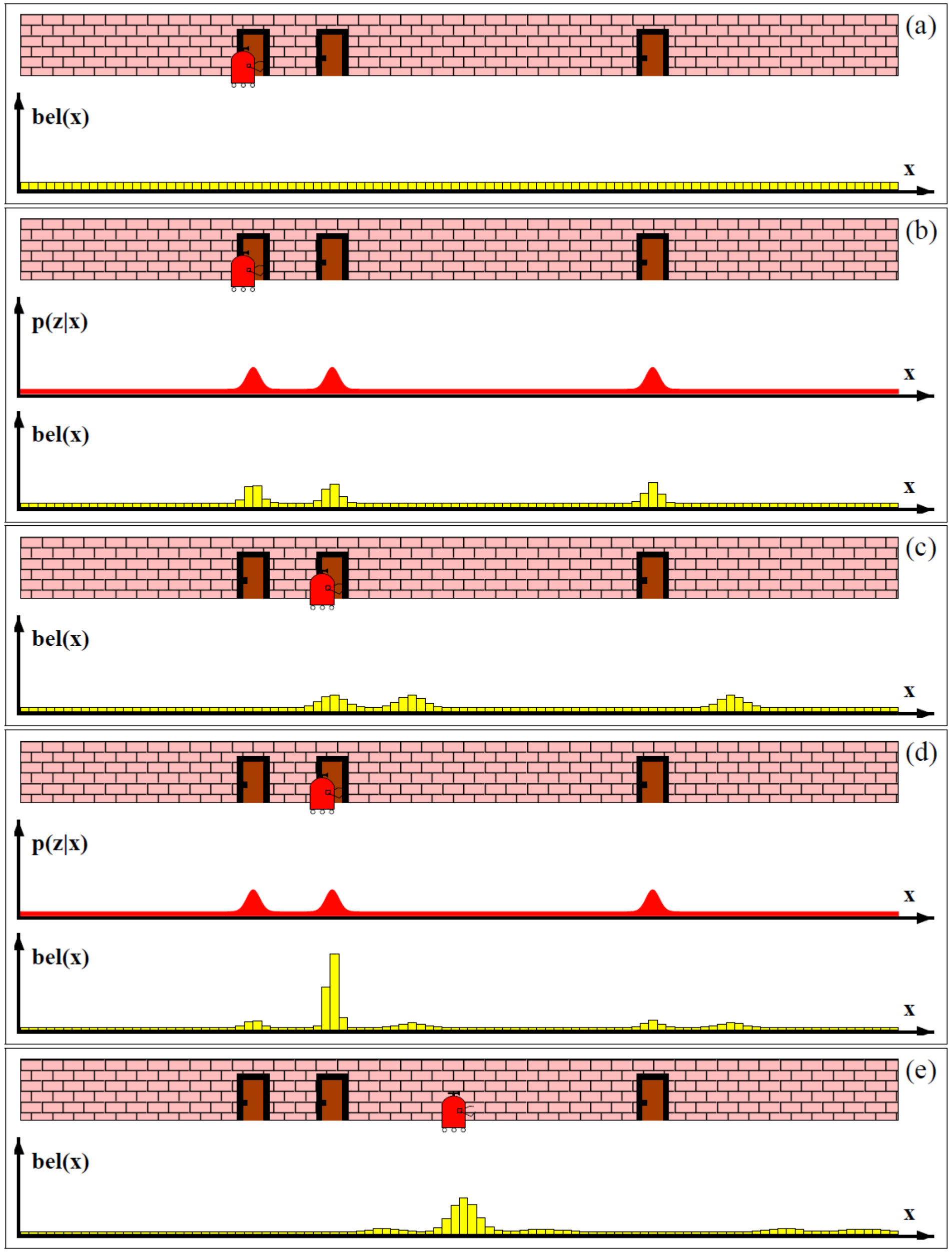


Figure 8.1 Grid localization using a fine-grained metric decomposition. Each picture depicts the position of the robot in the hallway along with its belief $bel(x_t)$, represented by a histogram over a grid.

Outline

✔ Overview of localization

Discrete case (1-D)

Continuous case: Kalman filter (1-D)

Mobile robot localization (Hallway example)

From
“Probabilistic Robotics”
(Ch. 7-8)
by
Sebastian Thrun,
Wolfram Burgard,
& Dieter Fox

Localization:
Discrete case (1-D)

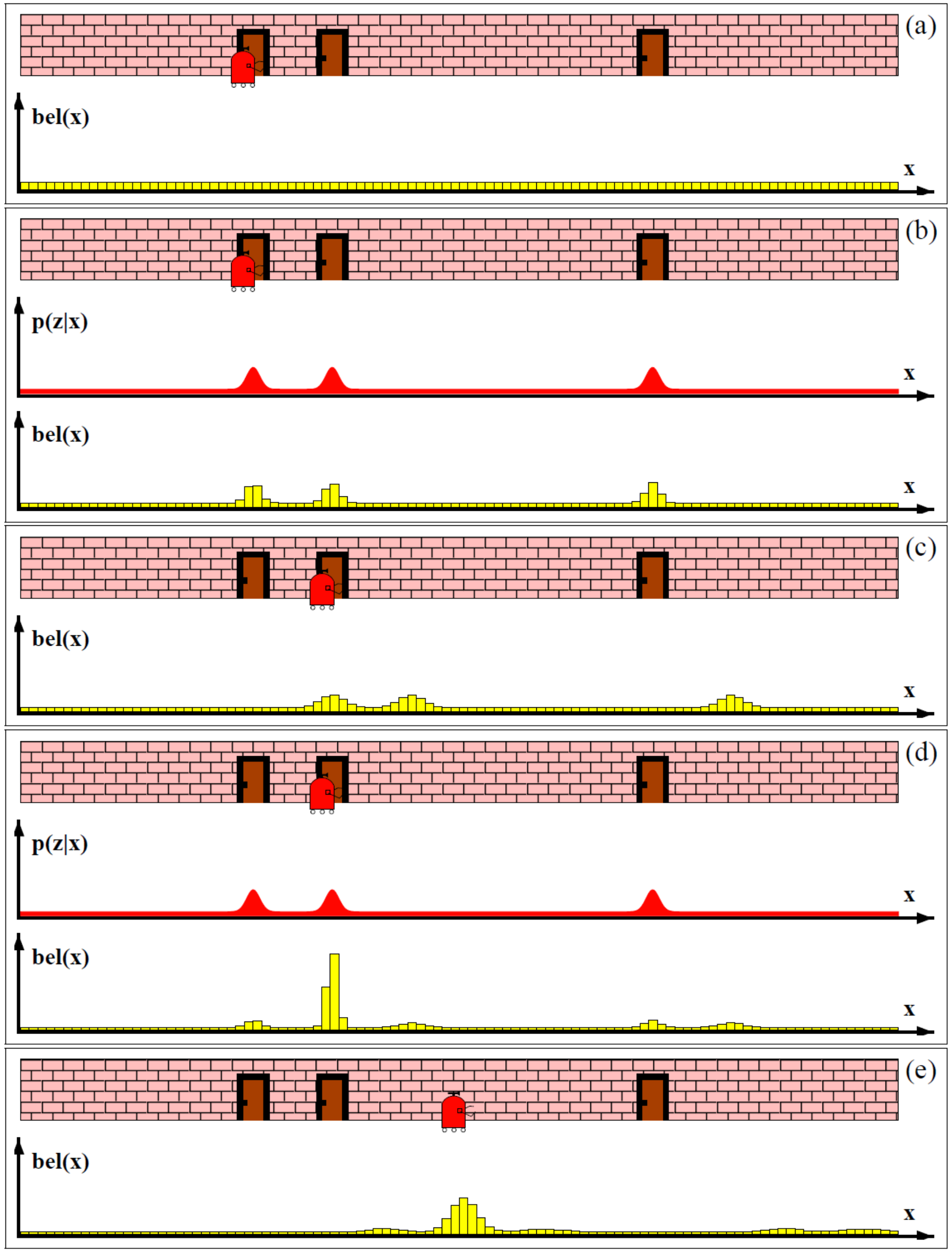
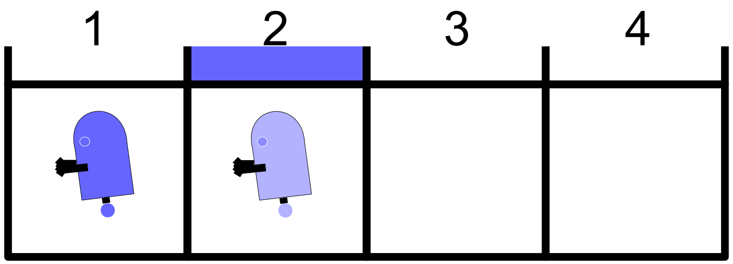


Figure 8.1 Grid localization using a fine-grained metric decomposition. Each picture depicts the position of the robot in the hallway along with its belief $bel(x_t)$, represented by a histogram over a grid.

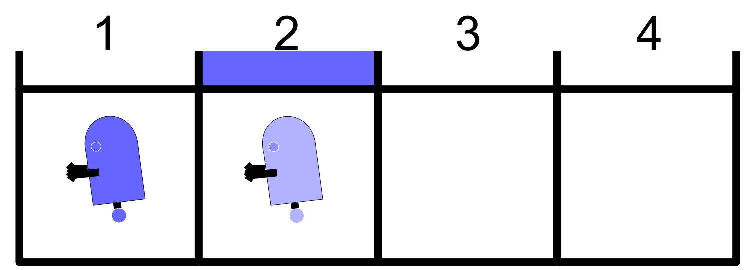
1-D example



Consider a 1-D grid world
The agent can only move right

Single action has <u>stochastic outcomes</u>
P(Actually moves right) = 0.8 P(Stays still) = 0.2

1-D example



Consider a 1-D grid world
The agent can only move right

Single action has <u>stochastic outcomes</u>

P(Actually moves right) = 0.8 P(Stays still) = 0.2

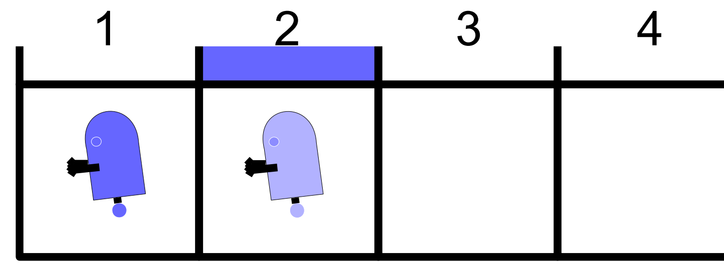
What if agent does not know where it is exactly?

- **New: State uncertainty**

<u>Belief</u>: Probability <u>distribution</u> over states

P(1) = 0.8 P(2) = 0.2 P(3) = 0.0 P(4) = 0.0

1-D example



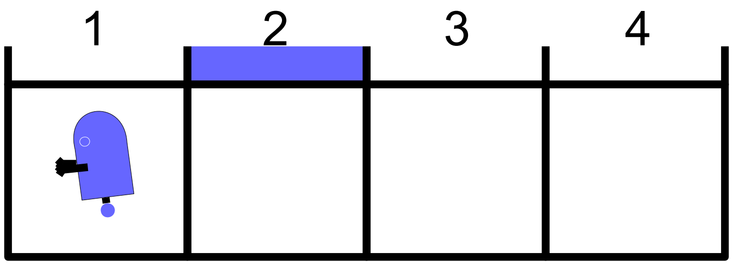
Consider a 1-D grid world
The agent can only move right

Single action has <u>stochastic outcomes</u>
P(Actually moves right) = 0.8 P(Stays still) = 0.2

What if agent does not know where it is exactly?
- **New: State uncertainty**
<u>Belief</u>: Probability <u>distribution</u> over states
P(1) = 0.8 P(2) = 0.2 P(3) = 0.0 P(4) = 0.0

What information can we use to infer our state?
- **New: Noisy observations**
Robot detects whether upper cell is door/empty
P(Empty | In state 1) = 0.9 P(Door | In state 1) = 0.1
P(Empty | In state 2) = 0.1 P(Door | In state 2) = 0.9
...

The first transition

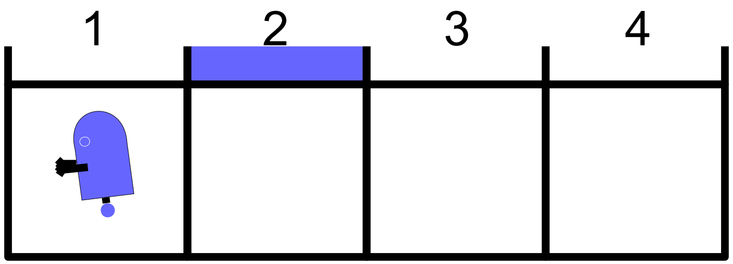


Consider a 1-D grid world
The agent can only move right

Suppose we know that $x_0 = 1$

$P(x_1) = ?$

The first transition

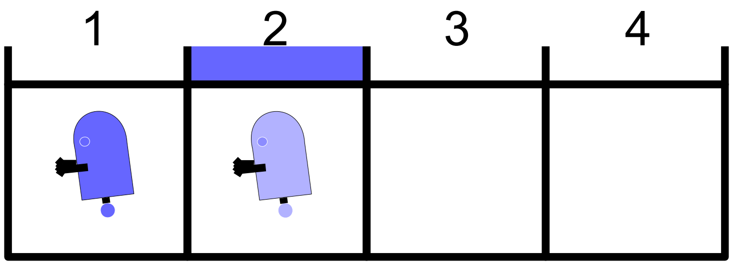


Consider a 1-D grid world
The agent can only move right

Suppose we know that $x_0 = 1$

$P(x_1)$: $P(x_1 = 1) = 0.2$ $P(x_1 = 2) = 0.8$

The first transition



Consider a 1-D grid world
The agent can only move right

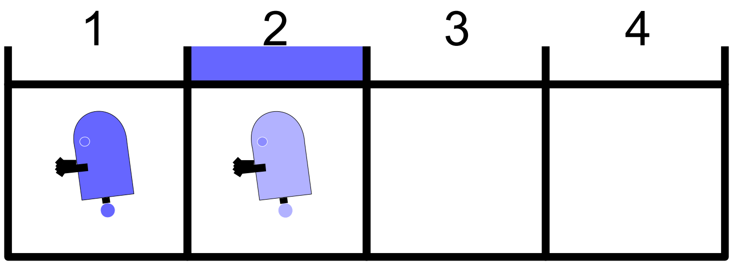
Suppose we know that $x_0 = 1$

$P(x_1)$: $P(x_1 = 1) = 0.2$ $P(x_1 = 2) = 0.8$

What if we <u>did not know</u> x_0, but only had a <u>***belief***</u> $P(x_0)$?

$P(x_0)$: $P(x_0 = 1) = 0.8$ $P(x_0 = 2) = 0.2$

The first transition



Consider a 1-D grid world
The agent can only move right

Suppose we know that $x_0 = 1$

$P(x_1)$: $P(x_1 = 1) = 0.2$ $P(x_1 = 2) = 0.8$

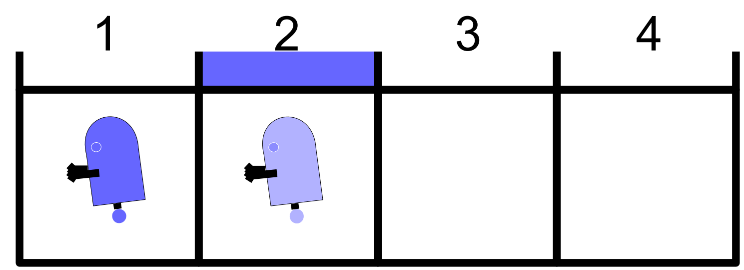
What if we <u>did not know</u> x_0, but only had a ***belief*** $P(x_0)$?

$P(x_0)$: $P(x_0 = 1) = 0.8$ $P(x_0 = 2) = 0.2$

With prob. 0.8, we start in $x_0 = 1$, then:

- With prob. 0.8, we end up in $x_1 = 2$ $\rightarrow p = 0.8 * 0.8 = 0.64$
- With prob. 0.2, we end up in $x_1 = 1$ $\rightarrow p = 0.8 * 0.2 = 0.16$

The first transition



Consider a 1-D grid world
The agent can only move right

Suppose we know that $x_0 = 1$

$P(x_1)$: $P(x_1 = 1) = 0.2$ $P(x_1 = 2) = 0.8$

What if we <u>did not know</u> x_0, but only had a <u>***belief***</u> $P(x_0)$?

$P(x_0)$: $P(x_0 = 1) = 0.8$ $P(x_0 = 2) = 0.2$

With prob. 0.8, we start in $x_0 = 1$, then:

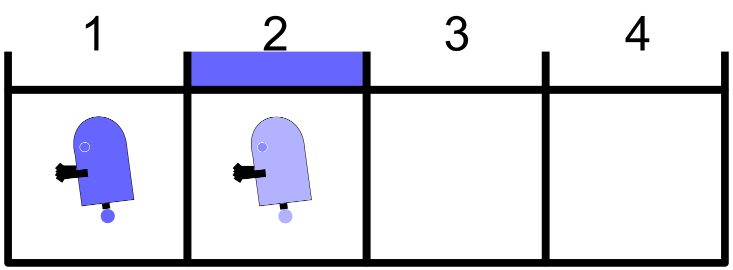
- With prob. 0.8, we end up in $x_1 = 2$ → $p = 0.8 * 0.8 = 0.64$
- With prob. 0.2, we end up in $x_1 = 1$ → $p = 0.8 * 0.2 = 0.16$

With prob. 0.2, we start in $x_0 = 2$, then:

- With prob. 0.8, we end up in $x_1 = 3$ → $p = 0.2 * 0.8 = 0.16$
- With prob. 0.2, we end up in $x_1 = 2$ → $p = 0.2 * 0.2 = 0.04$

The first transition

Consider a 1-D grid world
The agent can only move right



Suppose we know that $x_0 = 1$

$P(x_1)$: $P(x_1 = 1) = 0.2$ $\quad$ $P(x_1 = 2) = 0.8$

What if we <u>did not know</u> x_0, but only had a ***<u>belief</u>*** $P(x_0)$?

$P(x_0)$: $P(x_0 = 1) = 0.8$ $\quad$ $P(x_0 = 2) = 0.2$

With prob. 0.8, we start in $x_0 = 1$, then:

- With prob. 0.8, we end up in $x_1 = 2$ $\quad \rightarrow p = 0.8 * 0.8 = 0.64$
- With prob. 0.2, we end up in $x_1 = 1$ $\quad \rightarrow p = 0.8 * 0.2 = 0.16$

With prob. 0.2, we start in $x_0 = 2$, then:

- With prob. 0.8, we end up in $x_1 = 3$ $\quad \rightarrow p = 0.2 * 0.8 = 0.16$
- With prob. 0.2, we end up in $x_1 = 2$ $\quad \rightarrow p = 0.2 * 0.2 = 0.04$

$P(x_1 = 1) = 0.16$ $\quad$ $P(x_1 = 2) = 0.64 + 0.04 = 0.68$ $\quad$ $P(x_1 = 3) = 0.16$

The first transition

What if we <u>did not know</u> x_0, but only had a ***<u>belief</u>*** $P(x_0)$?

$P(x_0)$: $P(x_0 = 1) = 0.8$ $P(x_0 = 2) = 0.2$

With prob. 0.8, we start in $x_0 = 1$, then:

- With prob. 0.8, we end up in $x_1 = 2$ $\rightarrow p = 0.8 * 0.8 = 0.64$
- With prob. 0.2, we end up in $x_1 = 1$ $\rightarrow p = 0.8 * 0.2 = 0.16$

With prob. 0.2, we start in $x_0 = 2$, then:

- With prob. 0.8, we end up in $x_1 = 3$ $\rightarrow p = 0.2 * 0.8 = 0.16$
- With prob. 0.2, we end up in $x_1 = 2$ $\rightarrow p = 0.2 * 0.2 = 0.04$

$P(x_1 = 1) = 0.16$ $P(x_1 = 2) = 0.64 + 0.04 = 0.68$ $P(x_1 = 3) = 0.16$

$$\mathbb{P}(x_1 = 2) = \mathbb{P}(x_0 = 1) * \mathbb{P}(x_1 = 2 | x_0 = 1) + \mathbb{P}(x_0 = 2) * \mathbb{P}(x_1 = 2 | x_0 = 2)$$

The first transition

What if we <u>did not know</u> x_0, but only had a ***<u>belief</u>*** $P(x_0)$?

$P(x_0)$: $P(x_0 = 1) = 0.8$ $P(x_0 = 2) = 0.2$

With prob. 0.8, we start in $x_0 = 1$, then:

- With prob. 0.8, we end up in $x_1 = 2$ → $p = 0.8 * 0.8 = 0.64$
- With prob. 0.2, we end up in $x_1 = 1$ → $p = 0.8 * 0.2 = 0.16$

With prob. 0.2, we start in $x_0 = 2$, then:

- With prob. 0.8, we end up in $x_1 = 3$ → $p = 0.2 * 0.8 = 0.16$
- With prob. 0.2, we end up in $x_1 = 2$ → $p = 0.2 * 0.2 = 0.04$

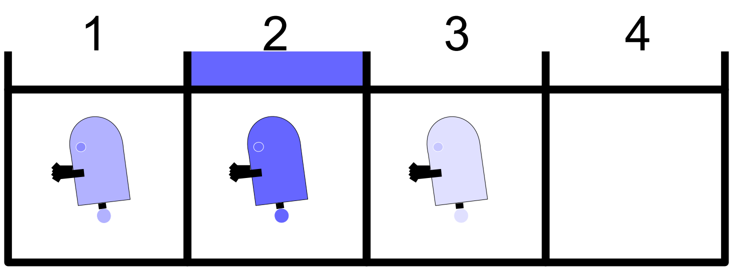
$P(x_1 = 1) = 0.16$ $P(x_1 = 2) = 0.64 + 0.04 = 0.68$ $P(x_1 = 3) = 0.16$

$$\mathbb{P}(x_1 = 2) = \mathbb{P}(x_0 = 1) * \mathbb{P}(x_1 = 2 | x_0 = 1) + \mathbb{P}(x_0 = 2) * \mathbb{P}(x_1 = 2 | x_0 = 2)$$

$$\mathbb{P}(x_1) = \sum_{x_0} \mathbb{P}(x_0, x_1) \qquad \text{Law of total probability}$$

$$= \sum_{x_0} \mathbb{P}(x_1 | x_0)\, \mathbb{P}(x_0) \qquad \text{Product rule of probability}$$

The first observation



Consider a 1-D grid world
The agent can only move right

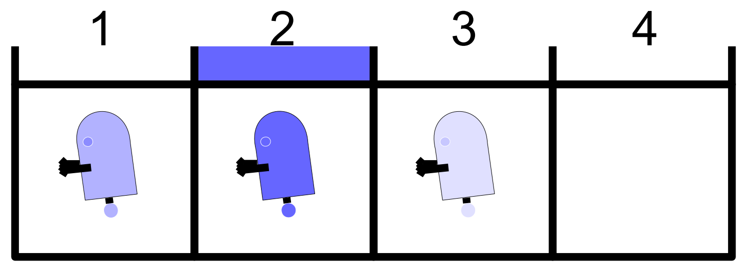
What if we <u>did not know</u> x_0, but only had a ***belief*** $P(x_0)$?

$P(x_0)$: $P(x_0 = 1) = 0.8$ $\quad$ $P(x_0 = 2) = 0.2$

$P(x_1 = 1) = 0.16$ $\quad$ $P(x_1 = 2) = 0.64 + 0.04 = 0.68$ $\quad$ $P(x_1 = 3) = 0.16$

What information can we use to infer our state?

The first observation



Consider a 1-D grid world
The agent can only move right

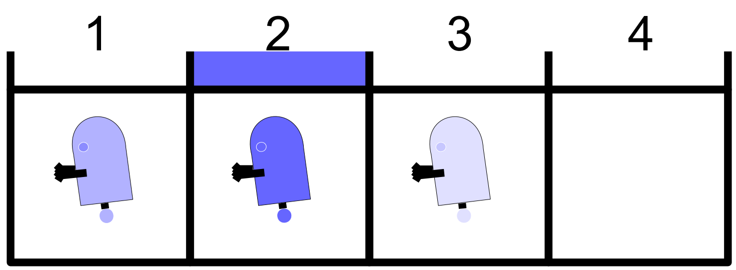
What if we <u>did not know</u> x_0, but only had a ***belief*** $P(x_0)$?

$P(x_0)$: $P(x_0 = 1) = 0.8$ $P(x_0 = 2) = 0.2$

$P(x_1 = 1) = 0.16$ $P(x_1 = 2) = 0.64 + 0.04 = 0.68$ $P(x_1 = 3) = 0.16$

What information can we use to infer our state?
Suppose we detect that the upper cell is a <u>door</u>
<u>Observation</u> at state x_1: y_1 = Door

The first observation



Consider a 1-D grid world
The agent can only move right

What if we <u>did not know</u> x_0, but only had a ***<u>belief</u>*** $P(x_0)$?

$P(x_0)$: $P(x_0 = 1) = 0.8$ $\quad P(x_0 = 2) = 0.2$

$P(x_1 = 1) = 0.16$ $\quad P(x_1 = 2) = 0.64 + 0.04 = 0.68$ $\quad P(x_1 = 3) = 0.16$

What information can we use to infer our state?
Suppose we detect that the upper cell is a <u>door</u>
<u>Observation</u> at state x_1: y_1 = Door

How do we use y_1 = Door to update $P(x_1)$?

What do you expect to happen?

Recall: Bayes' Rule

$$P(a|b) = \frac{P(b|a)P(a)}{P(b)}$$

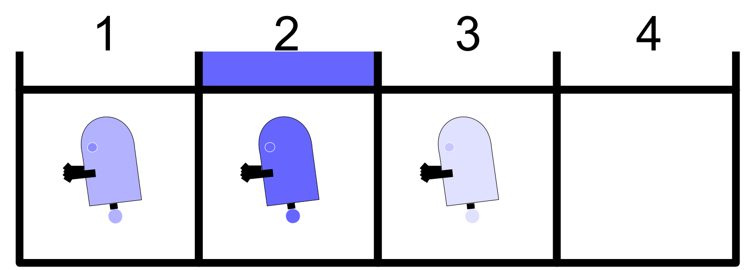
One interpretation:

$$P(\text{cause} \mid \text{effect}) = \frac{P(\text{effect} \mid \text{cause})\, P(\text{cause})}{P(\text{effect})}$$

Another interpretation:

$$\underbrace{P(\text{state} \mid \text{obs})}_{\textbf{Posterior}} = \frac{\overbrace{P(\text{obs} \mid \text{state})\, P(\text{state})}^{\textbf{Likelihood * Prior}}}{\underbrace{P(\text{obs})}_{\textbf{Evidence}}}$$

The first observation



Consider a 1-D grid world
The agent can only move right

What if we <u>did not know</u> x_0, but only had a ***<u>belief</u>*** $P(x_0)$?

$P(x_0)$: $P(x_0 = 1) = 0.8$ $P(x_0 = 2) = 0.2$

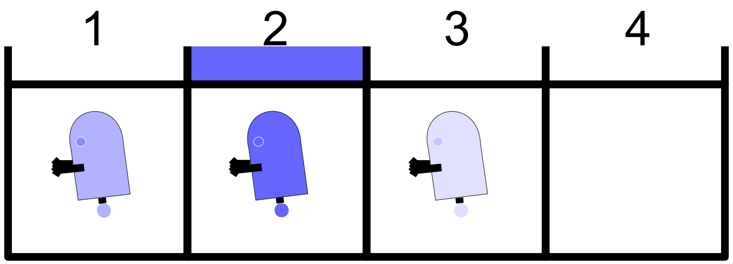
$P(x_1 = 1) = 0.16$ $P(x_1 = 2) = 0.64 + 0.04 = 0.68$ $P(x_1 = 3) = 0.16$

<u>Observation</u> at state x_1: y_1 = Door

How do we use y_1 = Door to update $P(x_1)$?

$$P(\text{state} \mid \text{obs}) = \frac{P(\text{obs} \mid \text{state})\, P(\text{state})}{P(\text{obs})}$$

The first observation



Consider a 1-D grid world
The agent can only move right

What if we <u>did not know</u> x_0, but only had a ***belief*** $P(x_0)$?

$P(x_0)$: $P(x_0 = 1) = 0.8$ $P(x_0 = 2) = 0.2$

$P(x_1 = 1) = 0.16$ $P(x_1 = 2) = 0.64 + 0.04 = 0.68$ $P(x_1 = 3) = 0.16$

<u>Observation</u> at state x_1: y_1 = Door

How do we use y_1 = Door to update $P(x_1)$?

$$P(x_1 \mid y_1 = \text{Door}) = \frac{P(y_1 = \text{Door} \mid x_1)\, P(x_1)}{P(y_1 = \text{Door})}$$

The first observation

What if we <u>did not know</u> x_0, but only had a ***belief*** $P(x_0)$?

$P(x_0)$: $P(x_0 = 1) = 0.8$ $P(x_0 = 2) = 0.2$

$P(x_1 = 1) = 0.16$ $P(x_1 = 2) = 0.64 + 0.04 = 0.68$ $P(x_1 = 3) = 0.16$

<u>Observation</u> at state x_1: y_1 = Door

$$P(x_1 \mid y_1 = \text{Door}) = \frac{P(y_1 = \text{Door} \mid x_1)\, P(x_1)}{P(y_1 = \text{Door})}$$

The first observation

What if we <u>did not know</u> x_0, but only had a ***<u>belief</u>*** $P(x_0)$?

$P(x_0)$: $P(x_0 = 1) = 0.8$ $P(x_0 = 2) = 0.2$

$P(x_1 = 1) = 0.16$ $P(x_1 = 2) = 0.64 + 0.04 = 0.68$ $P(x_1 = 3) = 0.16$

<u>Observation</u> at state x_1: y_1 = Door

$$P(x_1 \mid y_1 = \text{Door}) = \frac{P(y_1 = \text{Door} \mid x_1)\, P(x_1)}{P(y_1 = \text{Door})}$$

Notice that y_1 = Door is fixed – we cannot change this

Posterior probability varies with x_1

but denominator does not!

The first observation

What if we <u>did not know</u> x_0, but only had a ***belief*** $P(x_0)$?

$P(x_0)$: $P(x_0 = 1) = 0.8$ $P(x_0 = 2) = 0.2$

$P(x_1 = 1) = 0.16$ $P(x_1 = 2) = 0.64 + 0.04 = 0.68$ $P(x_1 = 3) = 0.16$

<u>Observation</u> at state x_1: y_1 = Door

$$P(x_1 \mid y_1 = \text{Door}) \;\alpha\; P(y_1 = \text{Door} \mid x_1)\, P(x_1)$$

Notice that y_1 = Door is fixed – we cannot change this

Posterior probability varies with x_1

but denominator does not!

The first observation

What if we <u>did not know</u> x_0, but only had a ***<u>belief</u>*** $P(x_0)$?

$P(x_0)$: $P(x_0 = 1) = 0.8$ $P(x_0 = 2) = 0.2$

$P(x_1 = 1) = 0.16$ $P(x_1 = 2) = 0.64 + 0.04 = 0.68$ $P(x_1 = 3) = 0.16$

<u>Observation</u> at state x_1: y_1 = Door

$$P(x_1 \mid y_1 = \text{Door}) \;\alpha\; P(y_1 = \text{Door} \mid x_1)\, P(x_1)$$

Notice that y_1 = Door is fixed – we cannot change this

Posterior probability varies with x_1

but denominator does not!

Compute numerator product for each x_1, then ***<u>normalize</u>***

The first observation

What if we <u>did not know</u> x_0, but only had a ***belief*** $P(x_0)$?

$P(x_0)$: $P(x_0 = 1) = 0.8$ $P(x_0 = 2) = 0.2$

$P(x_1 = 1) = 0.16$ $P(x_1 = 2) = 0.64 + 0.04 = 0.68$ $P(x_1 = 3) = 0.16$

<u>Observation</u> at state x_1: y_1 = Door

Compute numerator product for each x_1, then ***normalize***

$$P(x_1 \mid y_1 = \text{Door}) \ \alpha \qquad P(y_1 = \text{Door} \mid x_1)\, P(x_1)$$

$P(x_1 = 1 \mid y_1 = \text{Door})\ \alpha\ P(y_1 = \text{Door} \mid x_1 = 1)\, P(x_1 = 1) = 0.1 * 0.16 = 0.016$

The first observation

What if we <u>did not know</u> x_0, but only had a ***belief*** $P(x_0)$?

$P(x_0)$: $P(x_0 = 1) = 0.8$ $P(x_0 = 2) = 0.2$

$P(x_1 = 1) = 0.16$ $P(x_1 = 2) = 0.64 + 0.04 = 0.68$ $P(x_1 = 3) = 0.16$

<u>Observation</u> at state x_1: y_1 = Door

Compute numerator product for each x_1, then ***normalize***

$$P(x_1 \mid y_1 = \text{Door}) \;\alpha\; P(y_1 = \text{Door} \mid x_1)\, P(x_1)$$

$P(x_1 = 1 \mid y_1 = \text{Door}) \;\alpha\; P(y_1 = \text{Door} \mid x_1 = 1)\, P(x_1 = 1) = 0.1 * 0.16 = 0.016$

$P(x_1 = 2 \mid y_1 = \text{Door}) \;\alpha\; P(y_1 = \text{Door} \mid x_1 = 2)\, P(x_1 = 2) = 0.9 * 0.68 = 0.612$

$P(x_1 = 3 \mid y_1 = \text{Door}) \;\alpha\; P(y_1 = \text{Door} \mid x_1 = 3)\, P(x_1 = 3) = 0.1 * 0.16 = 0.016$

The first observation

What if we <u>did not know</u> x_0, but only had a ***belief*** $P(x_0)$?

$P(x_0)$: $P(x_0 = 1) = 0.8$ $P(x_0 = 2) = 0.2$

$P(x_1 = 1) = 0.16$ $P(x_1 = 2) = 0.64 + 0.04 = 0.68$ $P(x_1 = 3) = 0.16$

<u>Observation</u> at state x_1: y_1 = Door

Compute numerator product for each x_1, then ***normalize***

$$P(x_1 \mid y_1 = \text{Door}) \;\alpha\; P(y_1 = \text{Door} \mid x_1)\, P(x_1)$$

$P(x_1 = 1 \mid y_1 = \text{Door})\ \alpha\ P(y_1 = \text{Door} \mid x_1 = 1)\, P(x_1 = 1) = 0.1 * 0.16 = 0.016$

$P(x_1 = 2 \mid y_1 = \text{Door})\ \alpha\ P(y_1 = \text{Door} \mid x_1 = 2)\, P(x_1 = 2) = 0.9 * 0.68 = 0.612$

$P(x_1 = 3 \mid y_1 = \text{Door})\ \alpha\ P(y_1 = \text{Door} \mid x_1 = 3)\, P(x_1 = 3) = 0.1 * 0.16 = 0.016$

$0.016 + 0.612 + 0.016 = 0.644 \neq 1$ → That's okay, we ***normalize***

The first observation

What if we <u>did not know</u> x_0, but only had a ***<u>belief</u>*** $P(x_0)$?

$P(x_0)$: $P(x_0 = 1) = 0.8$ $P(x_0 = 2) = 0.2$

$P(x_1 = 1) = 0.16$ $P(x_1 = 2) = 0.64 + 0.04 = 0.68$ $P(x_1 = 3) = 0.16$

<u>Observation</u> at state x_1: y_1 = Door

Compute numerator product for each x_1, then ***<u>normalize</u>***

$$P(x_1 \mid y_1 = \text{Door}) \;\alpha\; P(y_1 = \text{Door} \mid x_1)\, P(x_1)$$

$P(x_1 = 1 \mid y_1 = \text{Door}) \;\alpha\; P(y_1 = \text{Door} \mid x_1 = 1)\, P(x_1 = 1) = 0.1 * 0.16 = 0.016$

$P(x_1 = 2 \mid y_1 = \text{Door}) \;\alpha\; P(y_1 = \text{Door} \mid x_1 = 2)\, P(x_1 = 2) = 0.9 * 0.68 = 0.612$

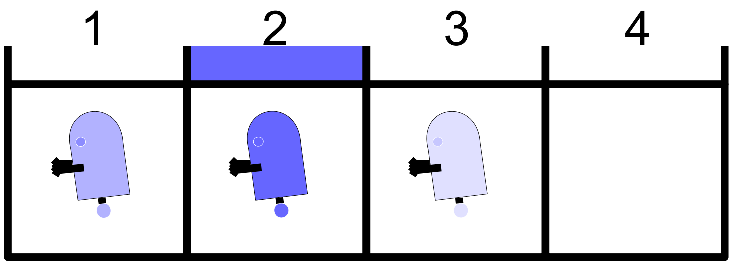
$P(x_1 = 3 \mid y_1 = \text{Door}) \;\alpha\; P(y_1 = \text{Door} \mid x_1 = 3)\, P(x_1 = 3) = 0.1 * 0.16 = 0.016$

$0.016 + 0.612 + 0.016 = 0.644 \neq 1$ $\rightarrow$ That's okay, we ***<u>normalize</u>***

$P(x_1 = 1 \mid y_1 = \text{Door}) = 0.016 / (0.016 + 0.612 + 0.016) \approx 0.025$

$P(x_1 = 2 \mid y_1 = \text{Door}) \approx 0.950$ $P(x_1 = 3 \mid y_1 = \text{Door}) \approx 0.025$

The first observation



Consider a 1-D grid world
The agent can only move right

What if we <u>did not know</u> x_0, but only had a ***belief*** $P(x_0)$?

$P(x_0)$: $P(x_0 = 1) = 0.8$ $P(x_0 = 2) = 0.2$

$P(x_1 = 1) = 0.16$ $P(x_1 = 2) = 0.64 + 0.04 = 0.68$ $P(x_1 = 3) = 0.16$

What information can we use to infer our state?
Suppose we detect that the upper cell is a <u>door</u>
<u>Observation</u> at state x_1: y_1 = Door

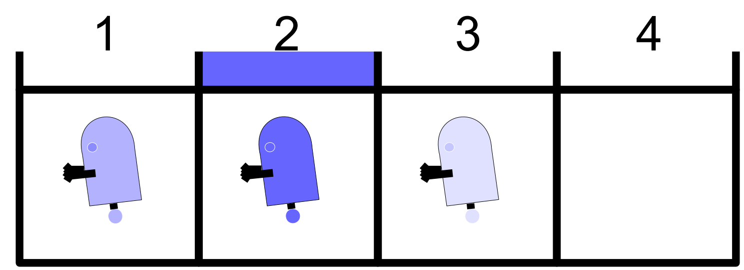
How do we use y_1 = Door to update $P(x_1)$?

$P(x_1 = 1 \mid y_1 = \text{Door}) \approx 0.025$

$P(x_1 = 2 \mid y_1 = \text{Door}) \approx 0.950$

$P(x_1 = 3 \mid y_1 = \text{Door}) \approx 0.025$

Summary: Discrete 1-D example



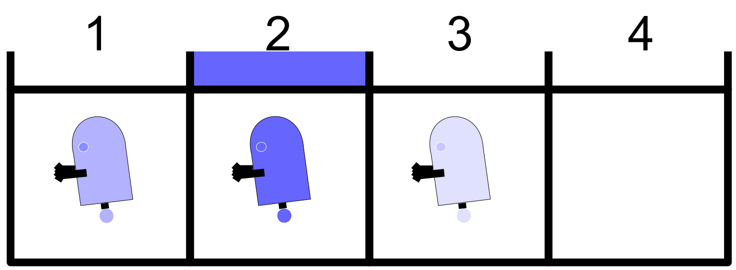
Consider a 1-D grid world
The agent can only move right

What if agent does not know where it is exactly?
- ***Belief***: Probability distribution over states
What information can we use to infer our state?
- **Noisy observations**

$\mathbb{P}(x_0)$	Initial / prior belief
$\mathbb{P}(x_1) = \sum_{x_0} \mathbb{P}(x_0, x_1)$	Law of total probability
$= \sum_{x_0} \mathbb{P}(x_1 \mid x_0)\, \mathbb{P}(x_0)$	Product rule of probability
$\mathbb{P}(x_1 \mid y_1) = \frac{\mathbb{P}(y_1 \mid x_1)\, \mathbb{P}(x_1)}{\mathbb{P}(y_1)}$	Bayes' rule
$\propto \mathbb{P}(y_1 \mid x_1)\, \mathbb{P}(x_1)$	Normalize to get posterior distribution

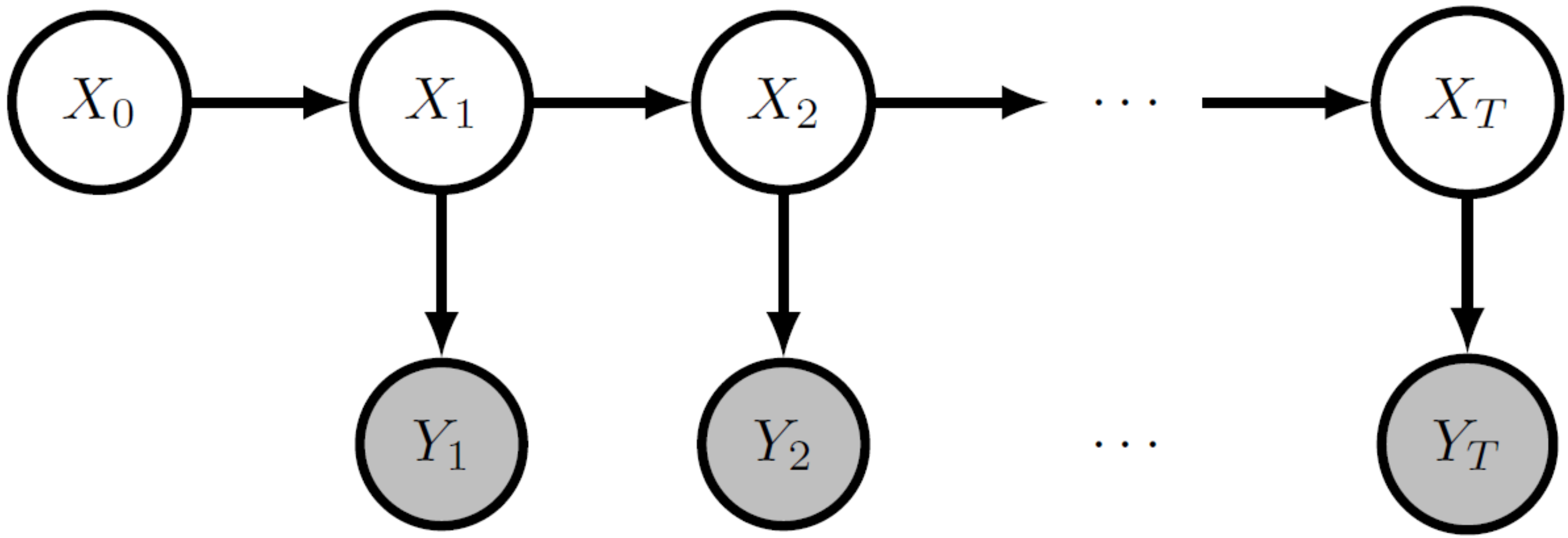
Hidden Markov Models

Consider a 1-D grid world
The agent can only move right



A Hidden Markov Model (HMM) is specified by 3 things:

- <u>Initial</u> belief $P(x_0)$
- <u>Transition</u> model $P(x_{t+1} \mid x_t)$
- <u>Observation</u> model $P(y_t \mid x_t)$ (or <u>emission</u> model $P(e_t \mid x_t)$)

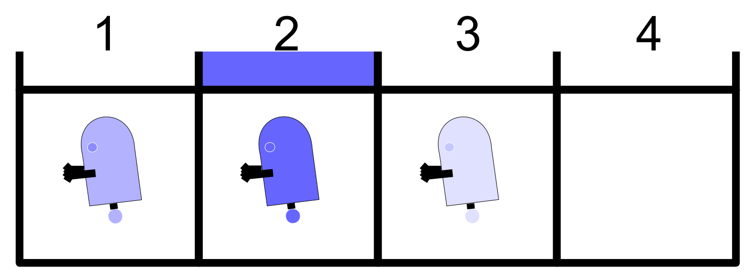


Recall: Rules of Probability

Name of rule	Formal definition
Probability	$\mathbb{P}(X) =$ How likely event X will occur
Conditional probability	$\mathbb{P}(Y\|X) = \frac{\mathbb{P}(X,Y)}{\mathbb{P}(X)} =$ "Probability of Y given X"
Product rule	$\mathbb{P}(X,Y) = \mathbb{P}(X\|Y)\,\mathbb{P}(Y)$ $\mathbb{P}(X,Y) = \mathbb{P}(Y\|X)\,\mathbb{P}(X)$
Bayes' rule	$\mathbb{P}(X\|Y) = \frac{\mathbb{P}(Y\|X)\,\mathbb{P}(X)}{\mathbb{P}(Y)} \propto \mathbb{P}(Y\|X)\,\mathbb{P}(X)$
Law of total probability	$\mathbb{P}(X) = \sum_Y \mathbb{P}(X,Y) = \sum_Y \mathbb{P}(X\|Y)\,\mathbb{P}(Y)$
Chain rule of probability	$\mathbb{P}(X_1,\ldots,X_n) = \mathbb{P}(X_n\|X_{n-1},\ldots,X_1)\,\mathbb{P}(X_{n-1}\|X_{n-2},\ldots,X_1)\,\ldots\,\mathbb{P}(X_2\|X_1)\,\mathbb{P}(X_1)$ $= \prod_{i=1}^{n} \mathbb{P}(X_i\|X_{i-1},\ldots,X_1)$

Recursive forward filtering in HMMs

Consider a 1-D grid world
The agent can only move right



A Hidden Markov Model (HMM) is specified by 3 things:
- Initial belief $P(x_0)$
- Transition model $P(x_{t+1} \mid x_t)$
- Observation model $P(y_t \mid x_t)$ (or emission model $P(e_t \mid x_t)$)

$$\mathbb{P}(x_0) \qquad \text{Initial / prior belief}$$

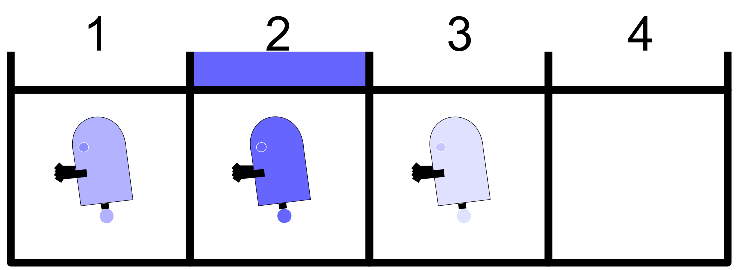
$$\mathbb{P}(x_1) = \sum_{x_0} \mathbb{P}(x_0, x_1) \qquad \text{Law of total probability}$$

$$= \sum_{x_0} \mathbb{P}(x_1|x_0)\, \mathbb{P}(x_0) \qquad \text{Product rule of probability}$$

$$\mathbb{P}(x_1|y_1) = \frac{\mathbb{P}(y_1|x_1)\, \mathbb{P}(x_1)}{\mathbb{P}(y_1)} \qquad \text{Bayes' rule}$$

$$\propto \mathbb{P}(y_1|x_1)\, \mathbb{P}(x_1) \qquad \text{Normalize to get posterior distribution}$$

Recursive forward filtering in HMMs



Consider a 1-D grid world
The agent can only move right

A Hidden Markov Model (HMM) is specified by 3 things:
- Initial belief $P(x_0)$
- Transition model $P(x_{t+1} \mid x_t)$
- Observation model $P(y_t \mid x_t)$ (or emission model $P(e_t \mid x_t)$)

More generally, at every time step:

$$\mathbb{P}(x_{t-1} | \underline{y_1, \ldots, y_{t-1}})$$ Current belief

$$\mathbb{P}(x_t | \underline{y_1, \ldots, y_{t-1}}) = \sum_{x_{t-1}} \mathbb{P}(x_t | x_{t-1})\, \mathbb{P}(x_{t-1} | y_1, \ldots, y_{t-1})$$ Prediction step

$$\mathbb{P}(x_t | \underline{y_1, \ldots, y_{t-1}, y_t}) \propto \mathbb{P}(y_t | x_t)\, \mathbb{P}(x_t | y_1, \ldots, y_{t-1})$$ Update step

Recursive forward filtering in HMMs

A Hidden Markov Model (HMM) is specified by 3 things:

- Initial belief $P(x_0)$
- Transition model $P(x_{t+1} | x_t)$
- Observation model $P(y_t | x_t)$ (or emission model $P(e_t | x_t)$)

$$\mathbb{P}(x_{t-1}|y_1,\ldots,y_{t-1}) \qquad \text{Current belief}$$

$$\mathbb{P}(x_t|y_1,\ldots,y_{t-1}) = \sum_{x_{t-1}} \mathbb{P}(x_t|x_{t-1})\, \mathbb{P}(x_{t-1}|y_1,\ldots,y_{t-1}) \qquad \text{Prediction step}$$

$$\mathbb{P}(x_t|y_1,\ldots,y_{t-1},y_t) \propto \mathbb{P}(y_t|x_t)\, \mathbb{P}(x_t|y_1,\ldots,y_{t-1}) \qquad \text{Update step}$$

b_0 = Initial belief

For t = 1 to T:

  Compute **prediction** step

    Compute $b_{t-1 \to t}$ using b_{t-1} and transition model

  Compute **update** step

    Compute b_t using $b_{t-1 \to t}$ and observation model

Outline

✔ Overview of localization

✔ Discrete case (1-D)

Continuous case: Kalman filter (1-D)

Mobile robot localization (Hallway example)

From
“Probabilistic Robotics”
(Ch. 7-8)
by
Sebastian Thrun,
Wolfram Burgard,
& Dieter Fox

Previously:
Discrete case (1-D)

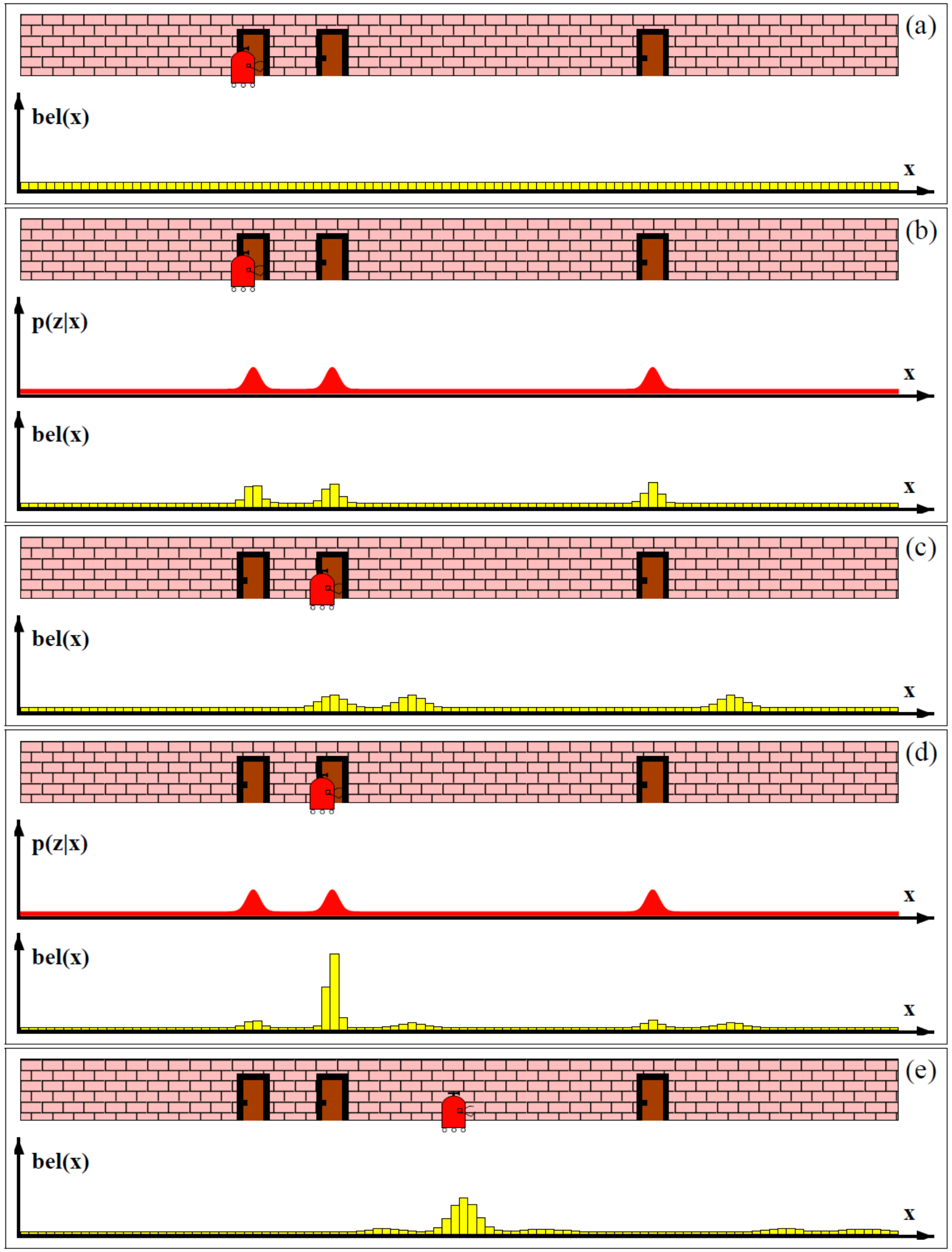


Figure 8.1 Grid localization using a fine-grained metric decomposition. Each picture depicts the position of the robot in the hallway along with its belief $bel(x_t)$, represented by a histogram over a grid.

Mobile robot localization (Hallway example)

From
“Probabilistic Robotics”
(Ch. 7-8)
by
Sebastian Thrun,
Wolfram Burgard,
& Dieter Fox

Now:
Continuous case (1-D)

We will study this
extensively with the
Kalman filter

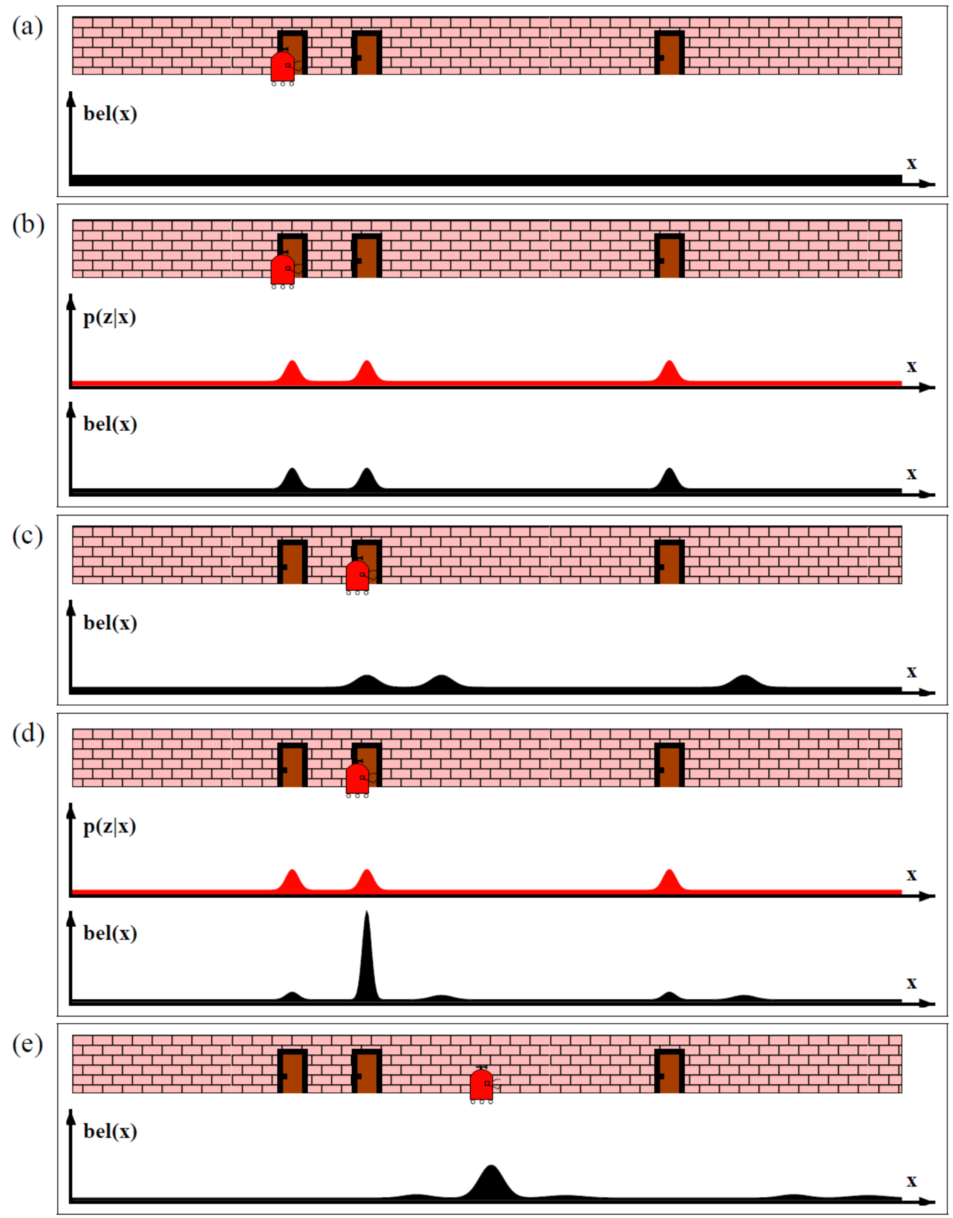
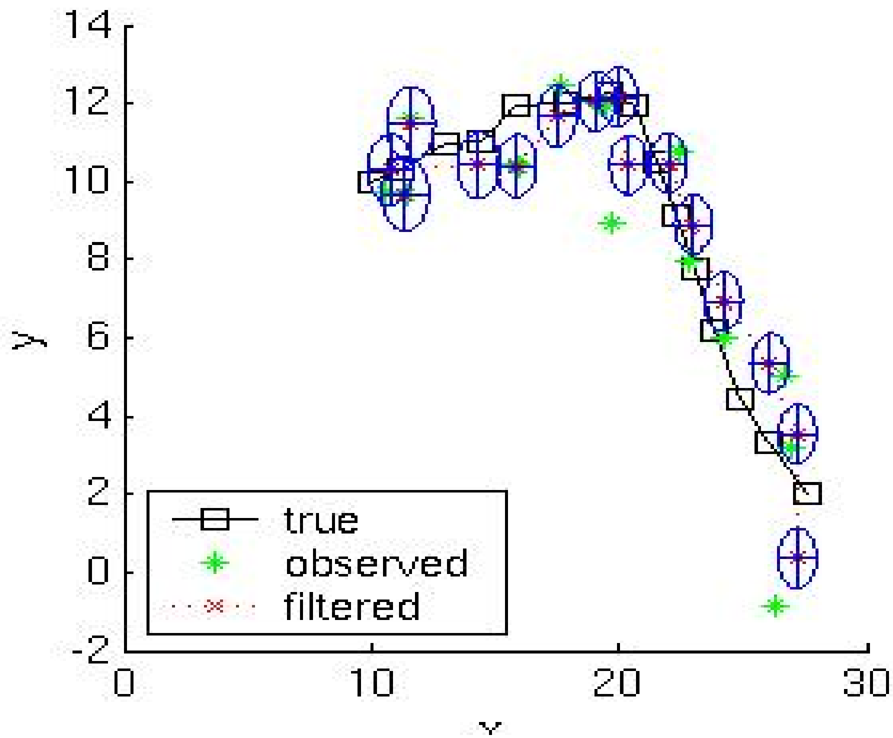


Figure 7.4 Illustration of the Markov localization algorithm. Each picture depicts the position of the robot in the hallway and its current belief $bel(x)$. (b) and (d) additionally depict the observation model $p(z_t \mid x_t)$, which describes the probability of observing a door at the different locations in the hallway.

Kalman Filtering



Sequential (recursive) Bayesian filtering
in continuous state spaces
- Represent beliefs as Gaussian distributions
- First developed in the early 1960s,
used in Apollo program

Image: UBC, Kevin Murphy Matlab toolbox

Kalman Filtering

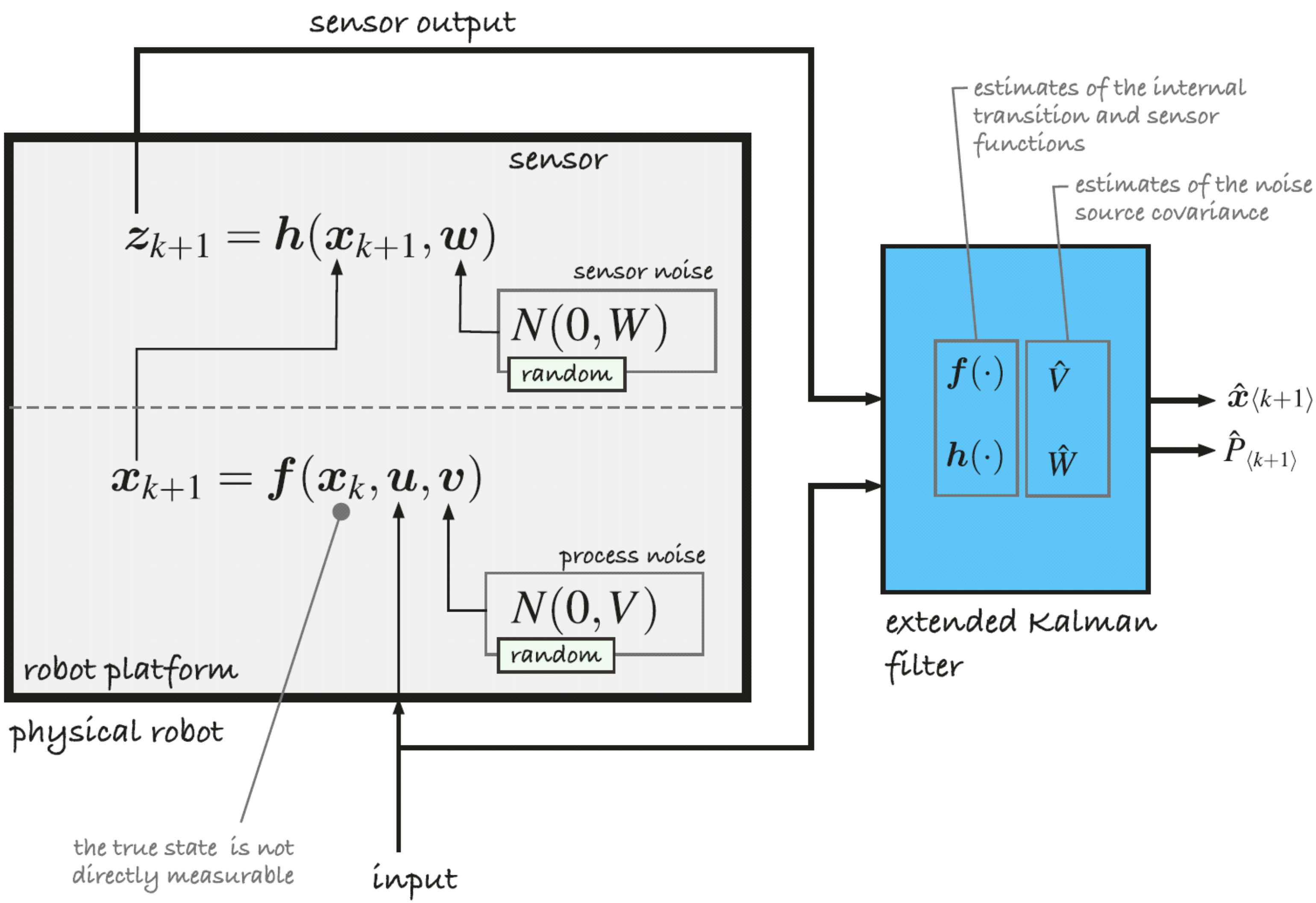
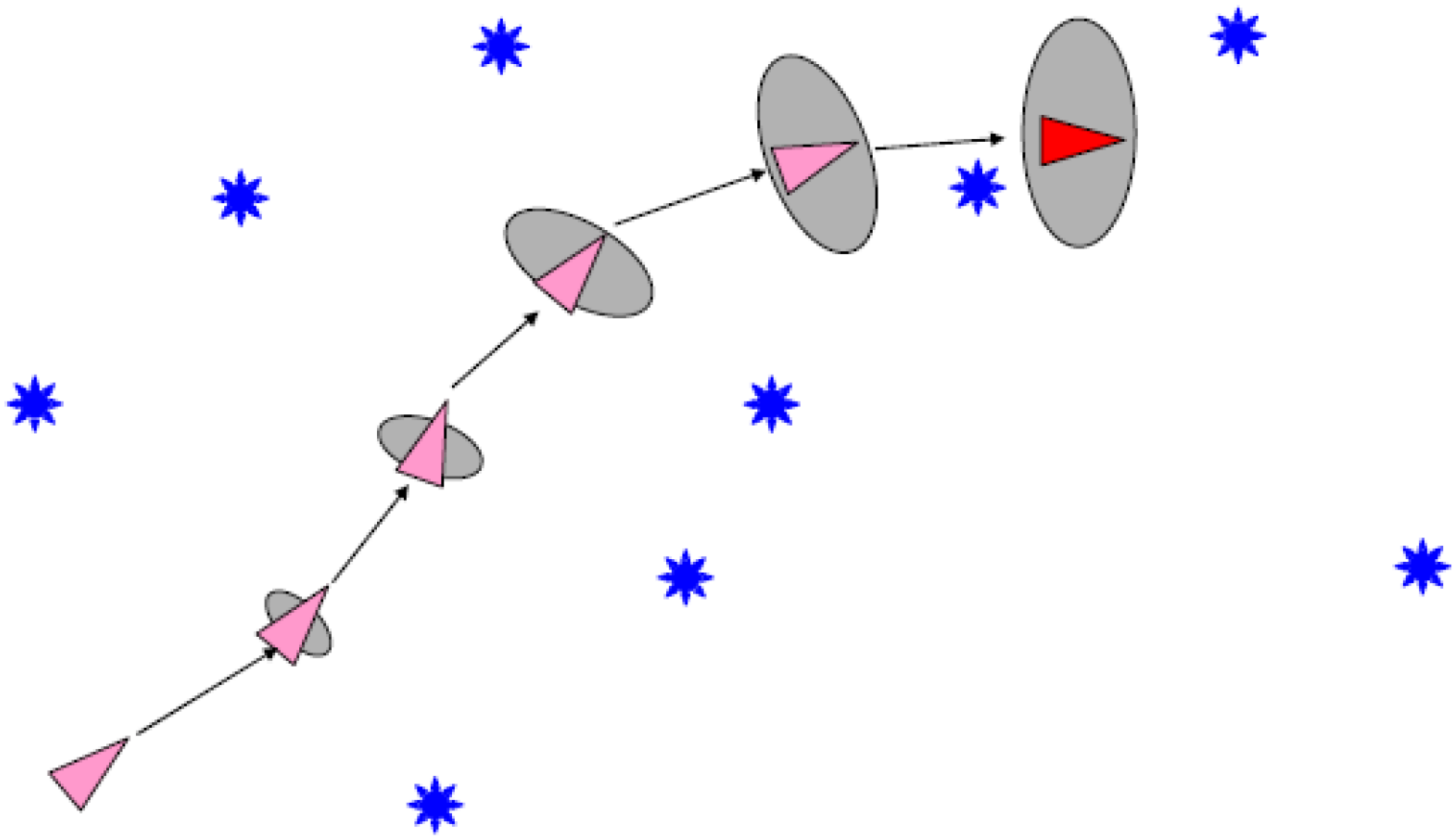
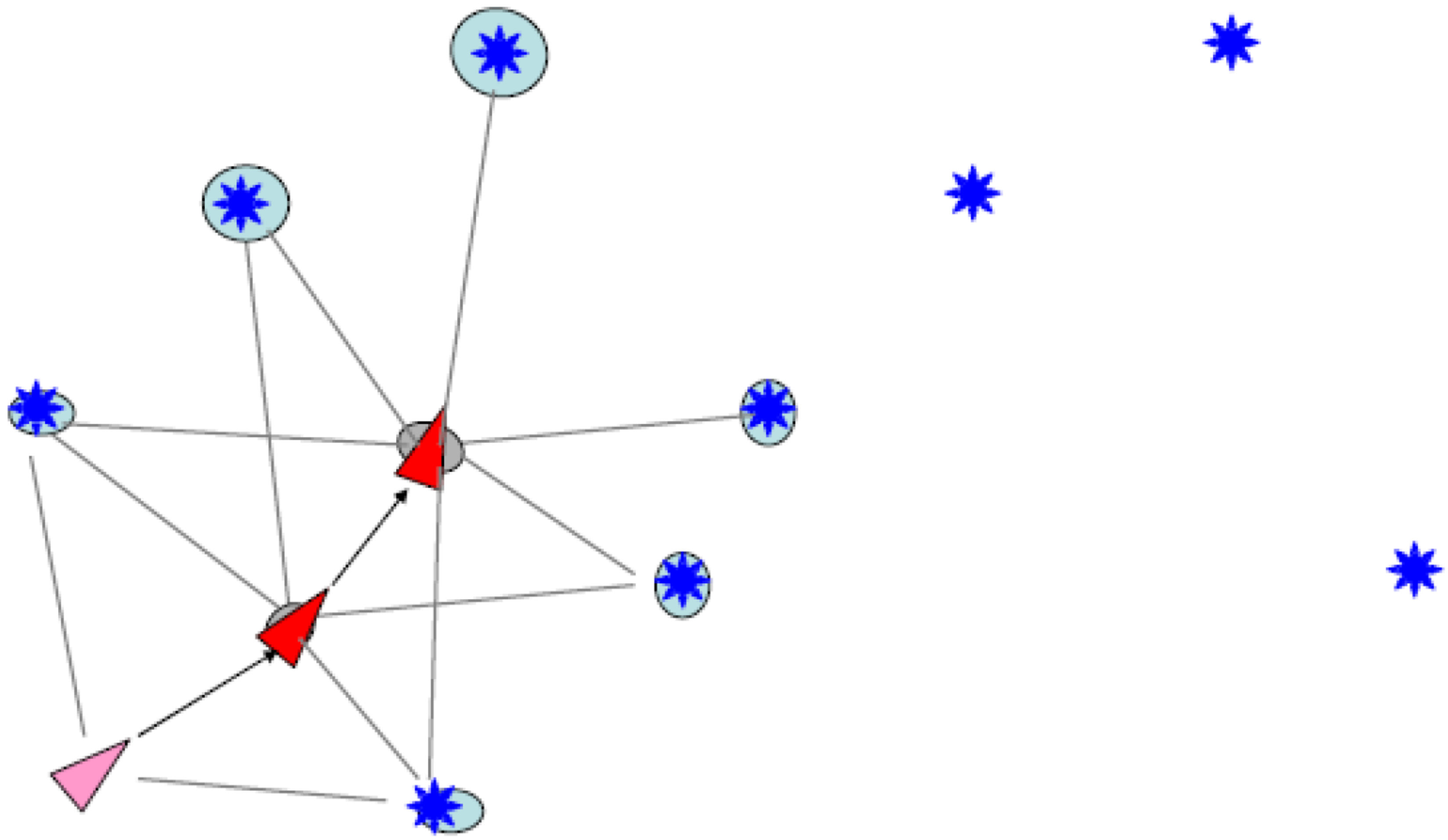


Fig. H.1.
The physical robot on the left has a true state that cannot be directly measured, however we gain a clue from the sensor output which is a function of this unknown true state

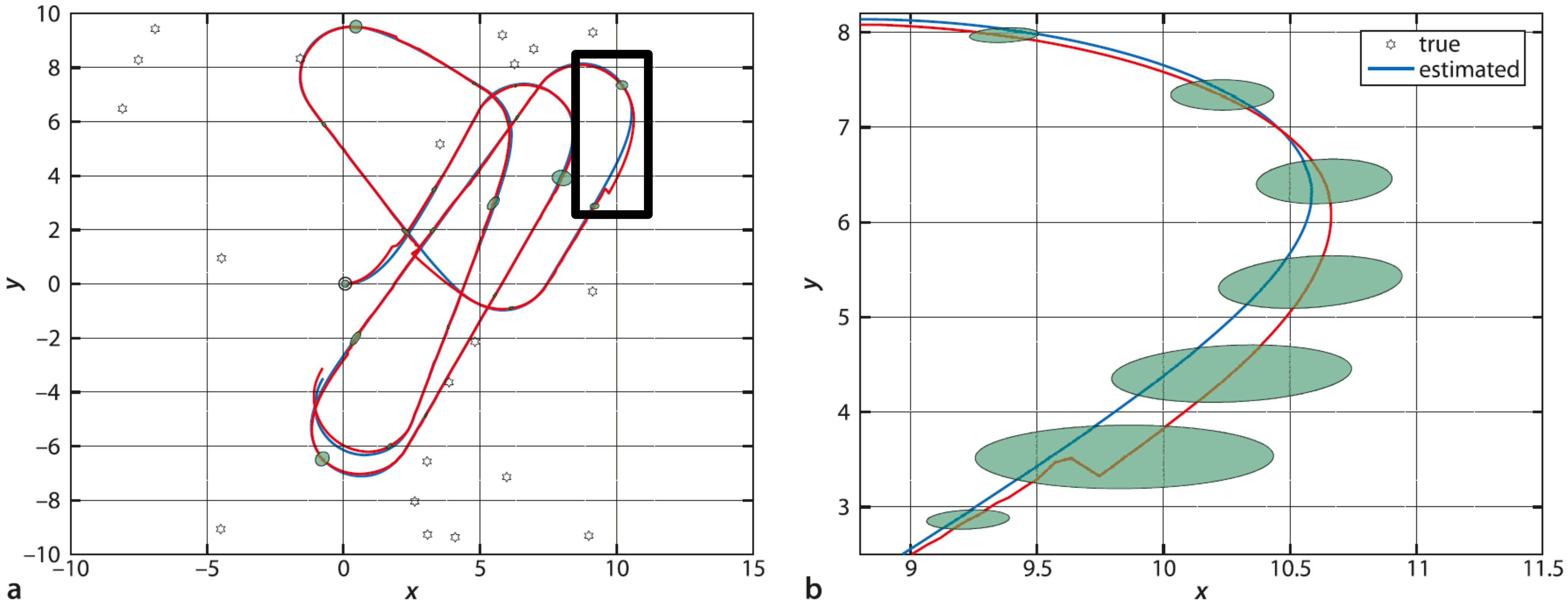
Application: Localizing against a known map



Application: Localizing against a known map



Application: Localizing against a known map



Kalman Filter (2-D)

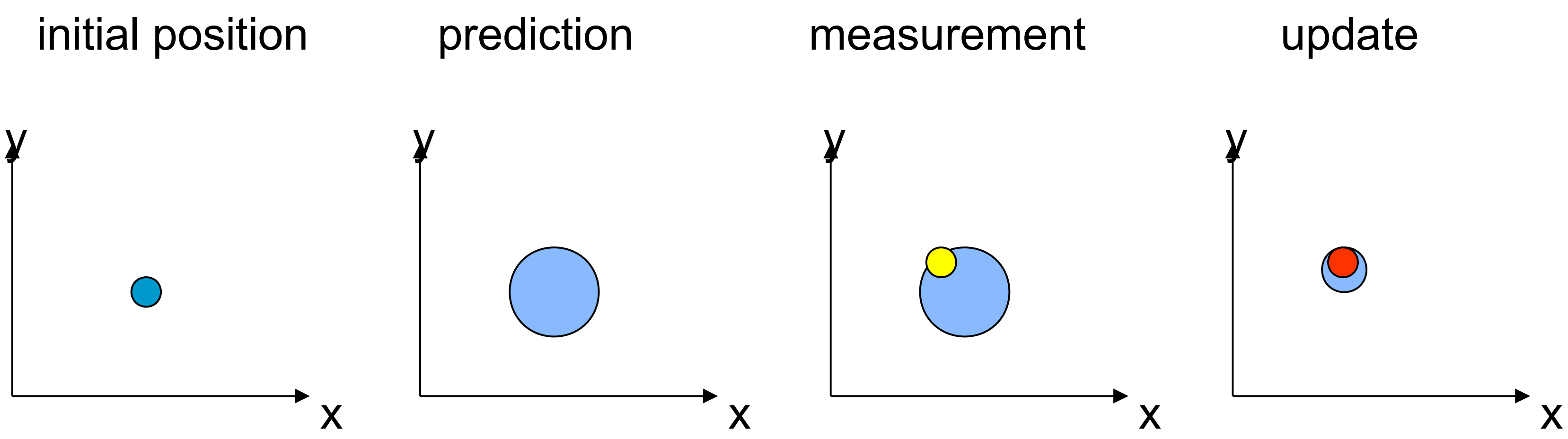


Image: Thrun *et al.*, *CS223B course notes*

Kalman Filter (1-D)

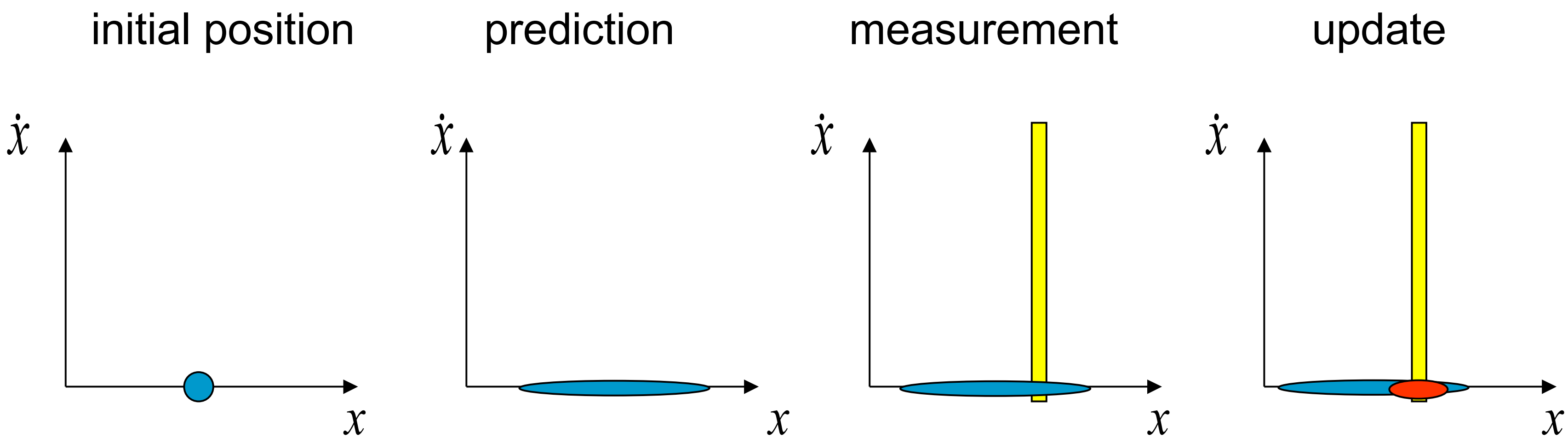
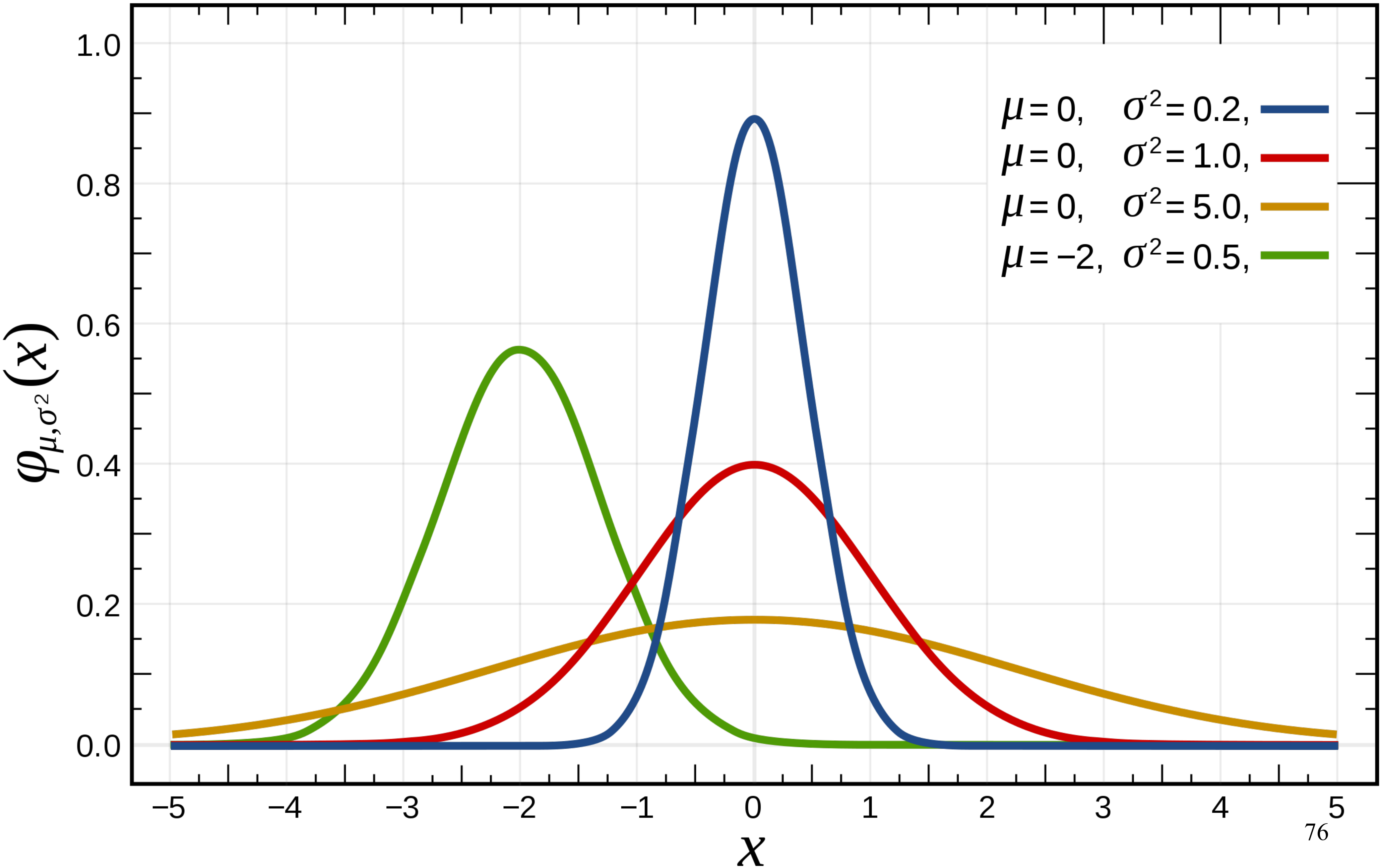


Image: Thrun *et al.*, *CS223B course notes*

Gaussian distribution / Normal distribution (1-D)



Kalman Filter (1-D)

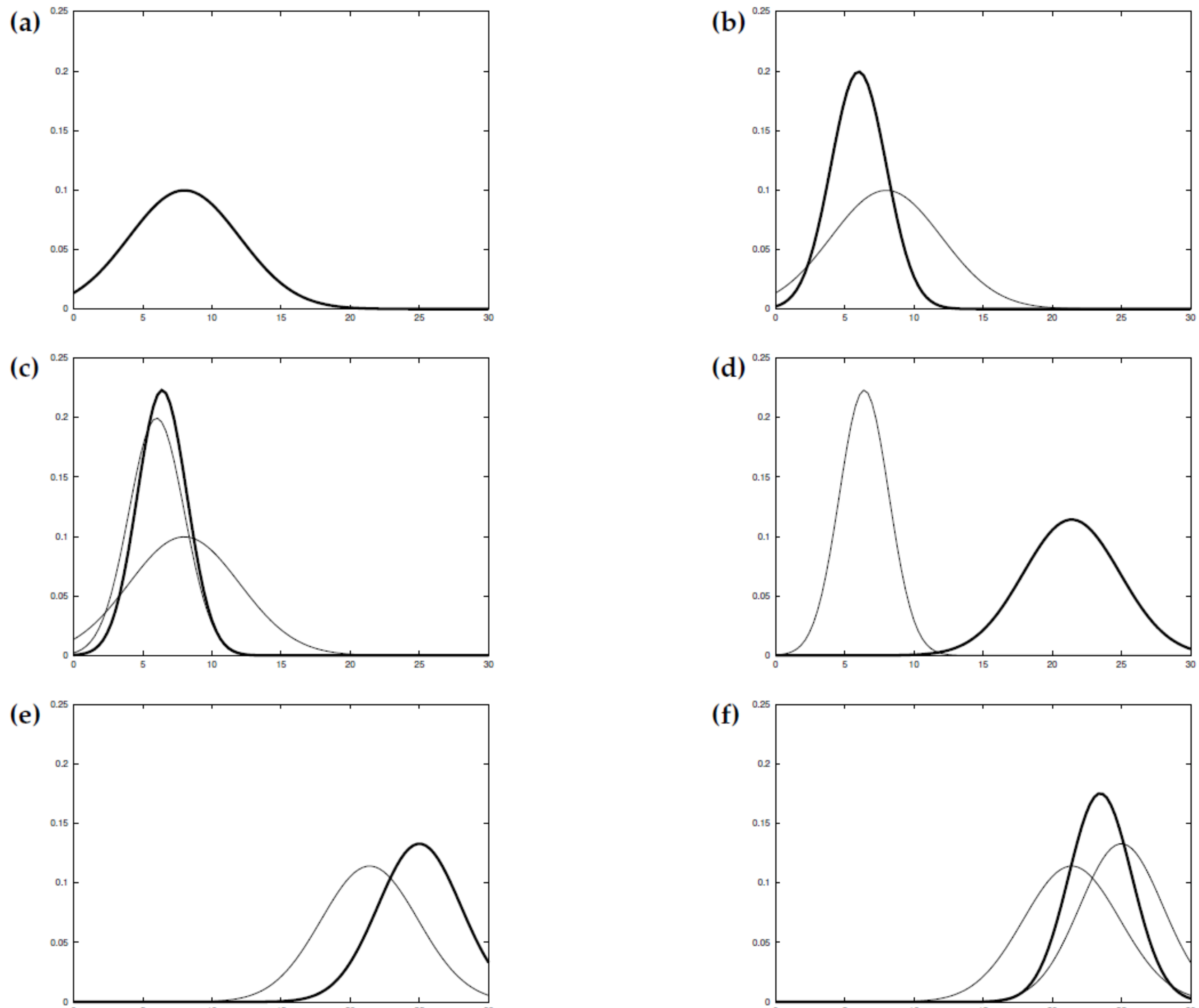
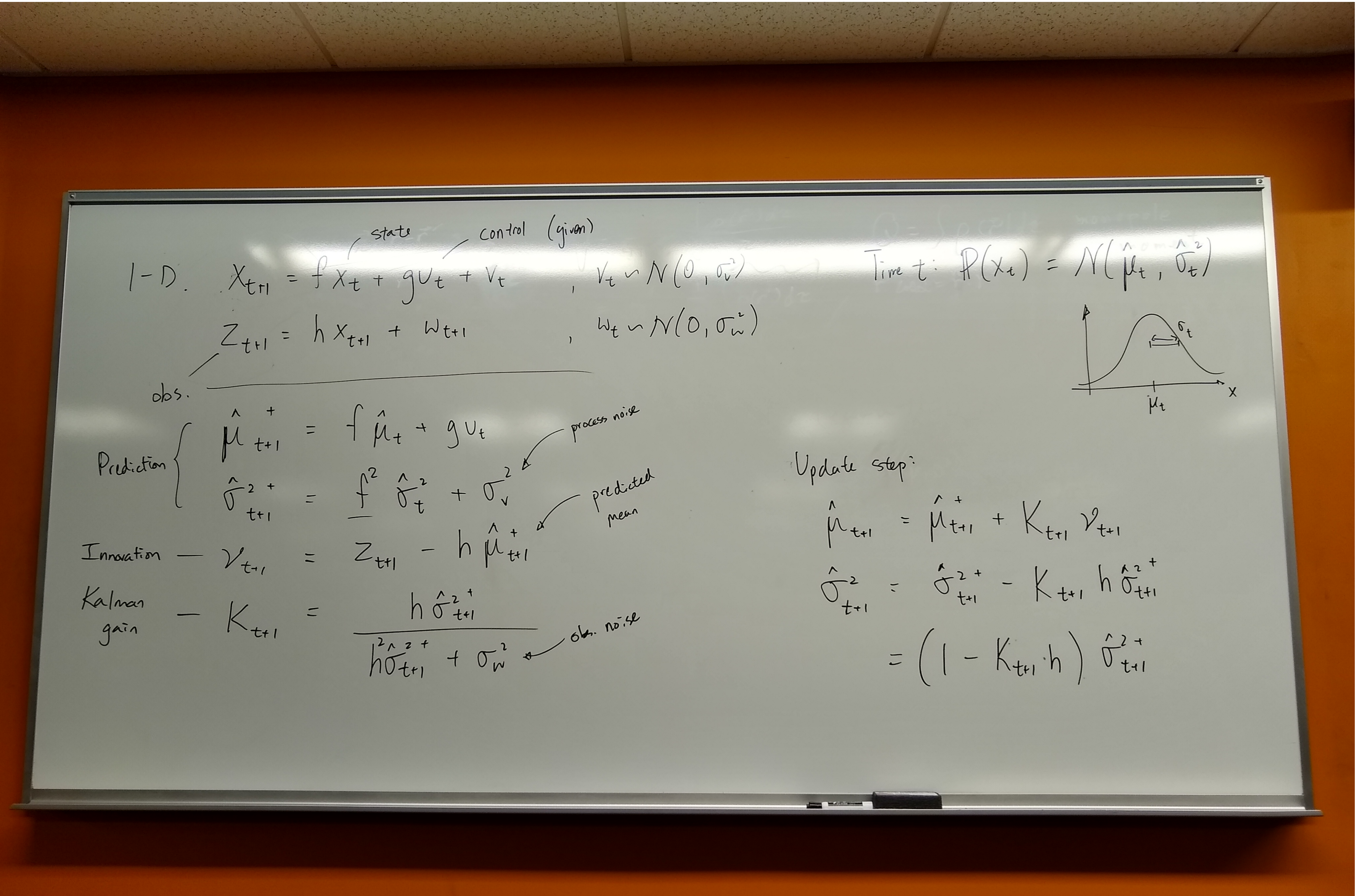


Figure 3.2 Illustration of Kalman filters: (a) initial belief, (b) a measurement (in bold) with the associated uncertainty, (c) belief after integrating the measurement into the belief using the Kalman filter algorithm, (d) belief after motion to the right (which introduces uncertainty), (e) a new measurement with associated uncertainty, and (f) the resulting belief.

On the board: 1-D Kalman Filter



On the board: 1-D Kalman Filter

System: $x_{t+1} = f x_t + g u_t + v_t \quad , \quad v_t \sim N(0, \sigma_v^2)$

$z_{t+1} = h x_{t+1} + w_{t+1} \quad , \quad w_t \sim N(0, \sigma_w^2)$

Kalman filter:

$$\hat{\mu}^{+}_{t+1} = f \hat{x}_t + g u_t$$

$$\hat{\sigma}^{2+}_{t+1} = f^2 \hat{\sigma}_t^2 + \sigma_v^2$$

$$\nu_{t+1} = z_{t+1} - h \hat{\mu}^{+}_{t+1}$$

$$K_{t+1} = \frac{h \hat{\sigma}^{2+}_{t+1}}{h^2 \hat{\sigma}^{2+}_{t+1} + \sigma_w^2}$$

$$\hat{\mu}_{t+1} = \hat{\mu}^{+}_{t+1} + K_{t+1} \nu_{t+1}$$

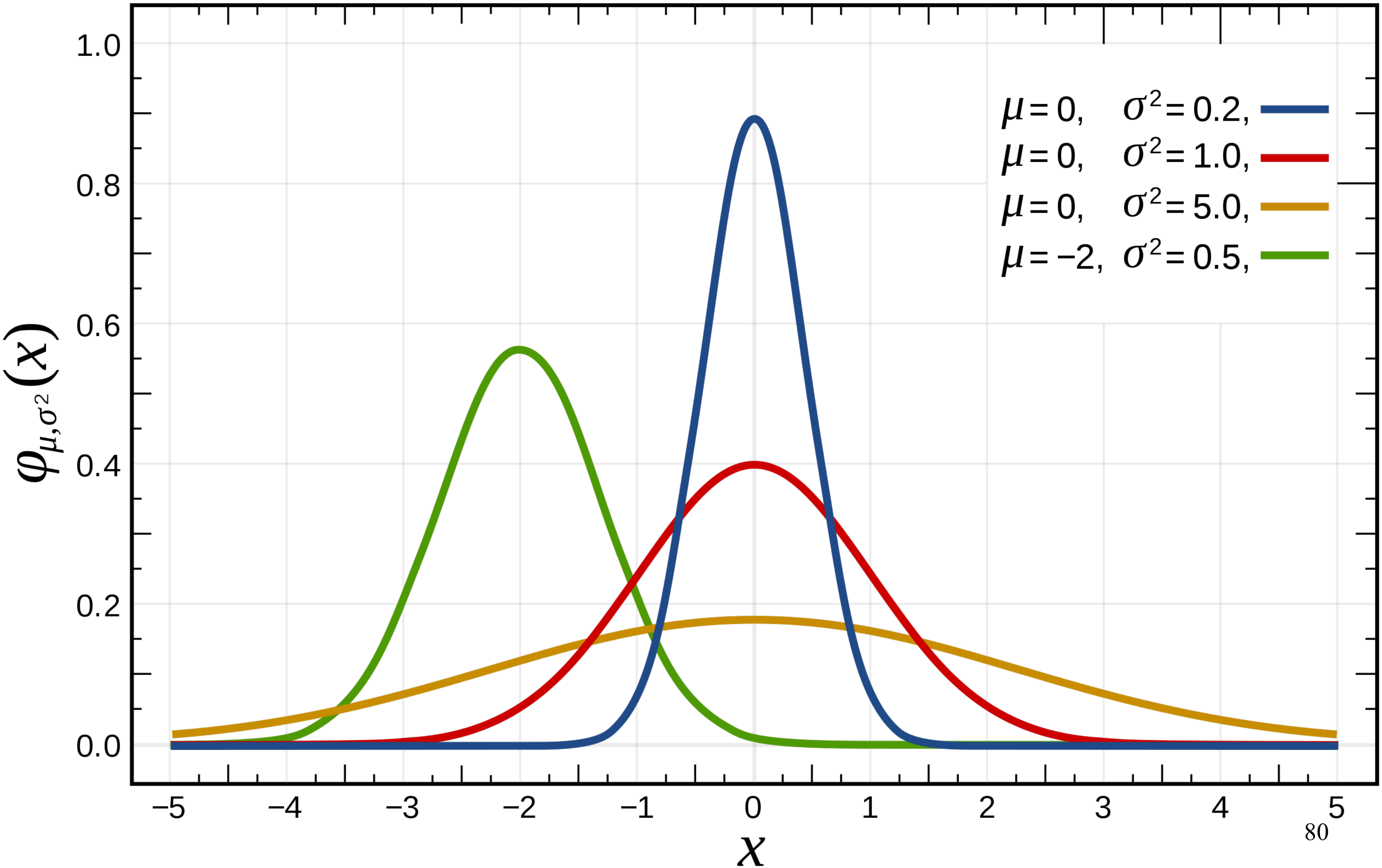
$$\hat{\sigma}^2_{t+1} = \left(1 - \frac{h^2 \hat{\sigma}^{2+}_{t+1}}{h^2 \hat{\sigma}^{2+}_{t+1} + \sigma_w^2}\right) \hat{\sigma}^{2+}_{t+1} = \frac{\sigma_w^2 \hat{\sigma}^{2+}_{t+1}}{h^2 \hat{\sigma}^{2+}_{t+1} + \sigma_w^2}$$

Belief:

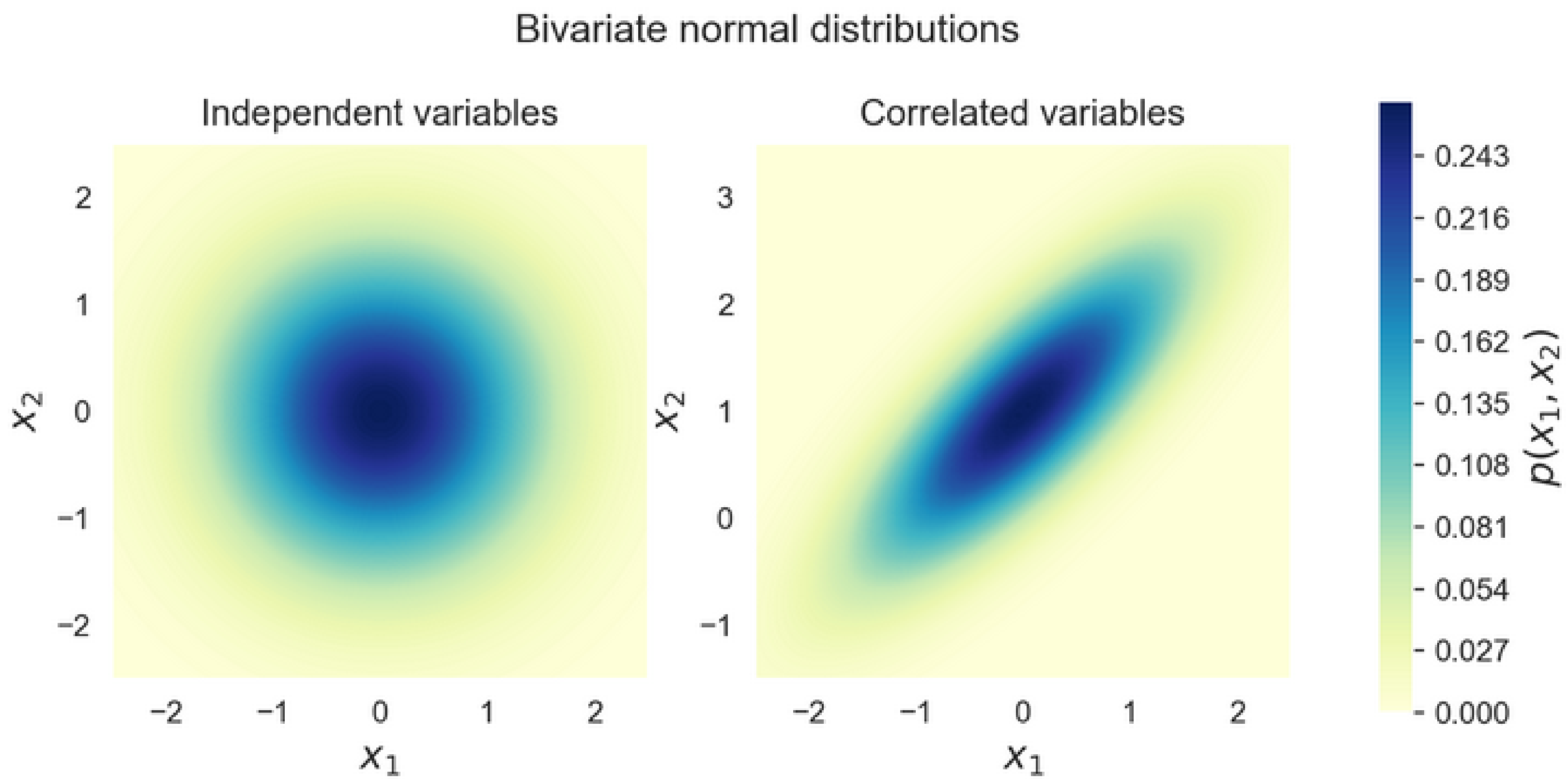
$$x_t \sim N(\hat{\mu}_t, \hat{\sigma}_t^2)$$

$$f x_t + g u_t \sim N\Big(\underbrace{f \hat{\mu}_t + g u_t}_{\hat{\mu}^{+}_{t+1}}, \; f^2 \sigma_t^2\Big)$$

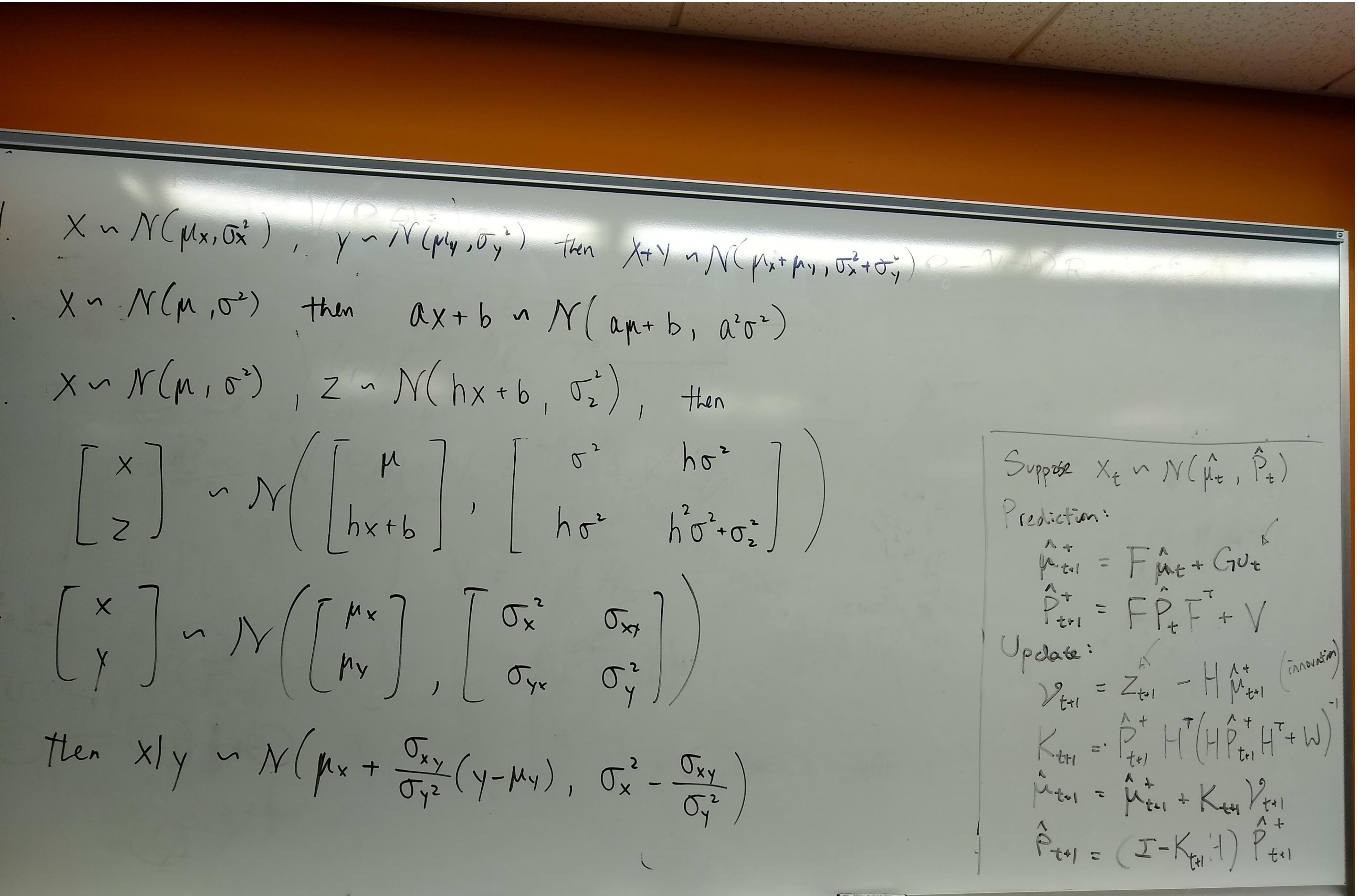
Gaussian distribution / Normal distribution (1-D)



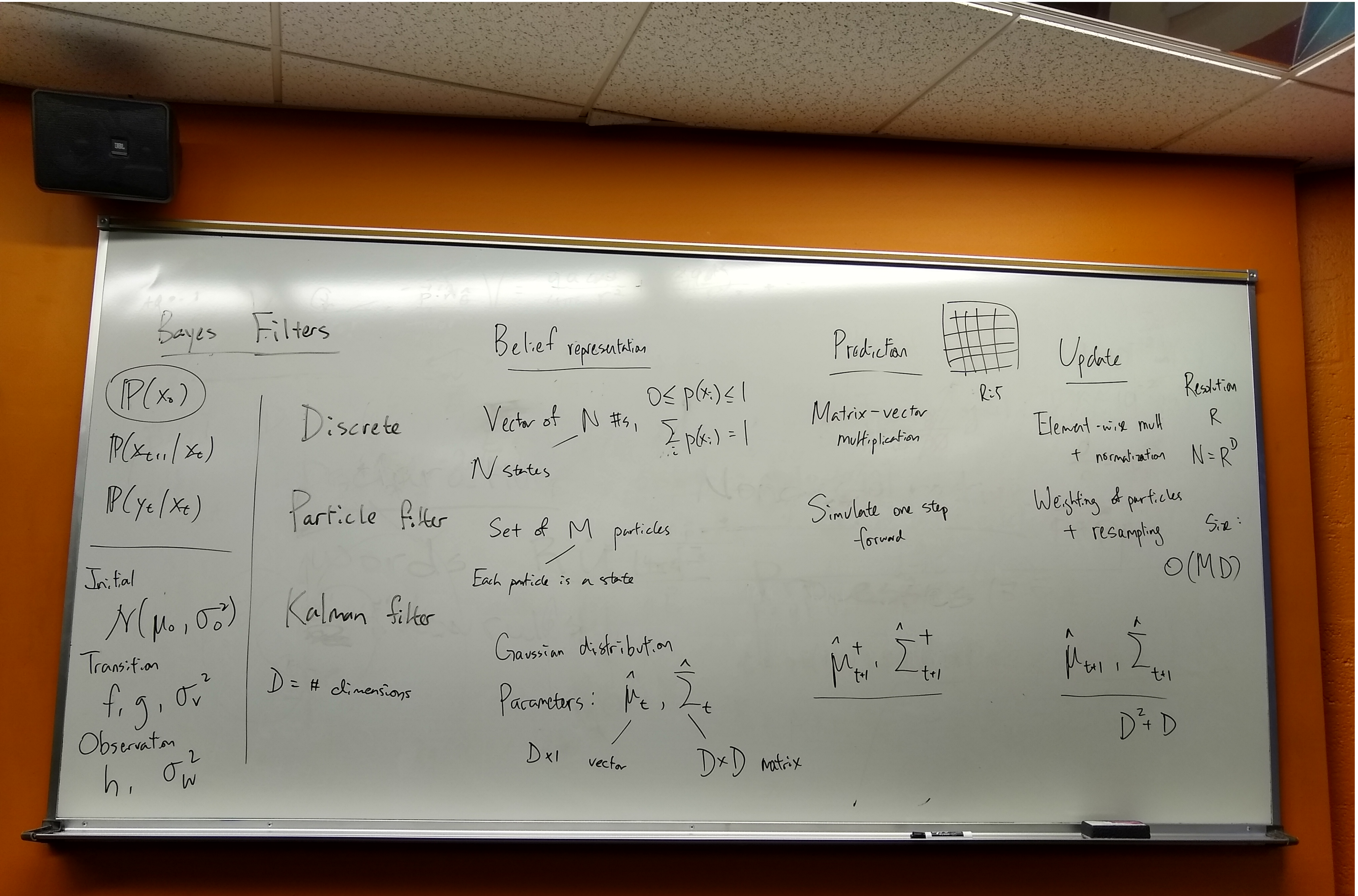
Gaussian distribution / Normal distribution (2-D)



On the board: Properties of Gaussian distributions



On the board: Bayesian filtering



Feedback

Piazza thread: 2/16 Lec 09 Feedback

Please post your answers to the following anonymously.

1. What did you like so far?

2. What was unclear?

3. Liking Gaussian distributions?

4. Any additional feedback / comments?