Chapter 5
Optimal Estimation

Part 4

5.4 Bayesian Estimation

Carnegie Mellon
THE ROBOTICS INSTITUTE

Outline

- 5.4 Bayesian Estimation
  - 5.4.1 Definitions
  - 5.4.2 Bayes' Rule
  - 5.4.3 Bayes' Filters
  - 5.4.4 Bayesian Mapping
  - 5.4.5 Bayesian Localization
  - Summary

Outline

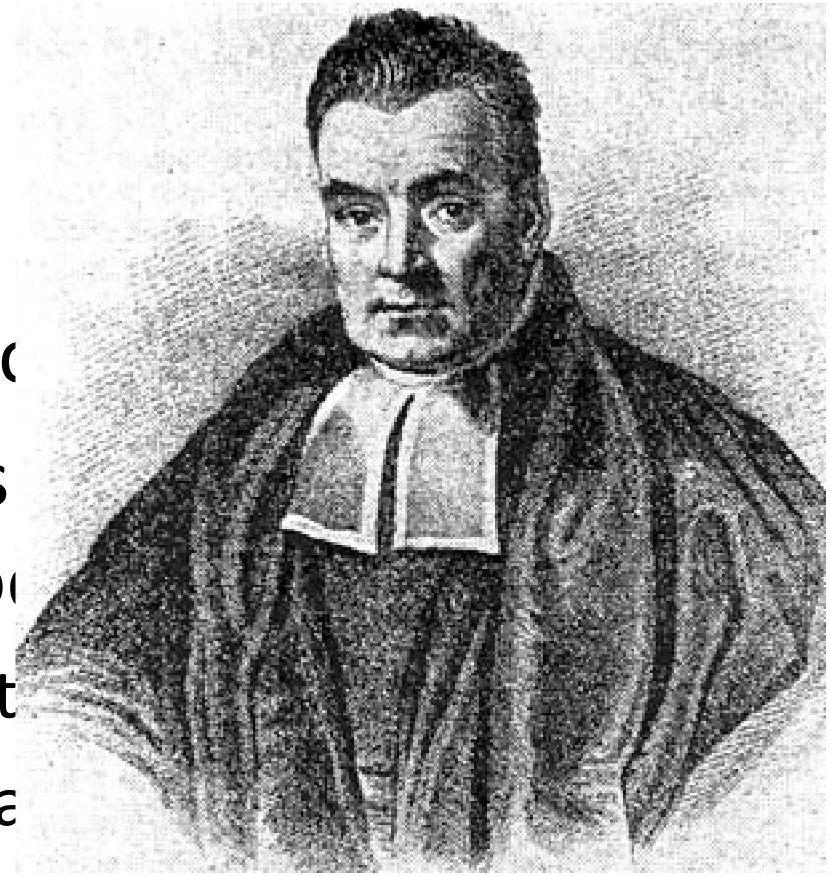
- 5.4 Bayesian Estimation
  - 5.4.1 Definitions
  - 5.4.2 Bayes' Rule
  - 5.4.3 Bayes' Filters
  - 5.4.4 Bayesian Mapping
  - 5.4.5 Bayesian Localization
  - Summary

Whats The Big Deal

- Can handle arbitrary (non Gaussian) distributions,
- Produces an arbitrary distribution as a result.
- Hence, computes the probability the robot is in every place. Solves the "kidnapped robot" problem.

Thomas Bayes

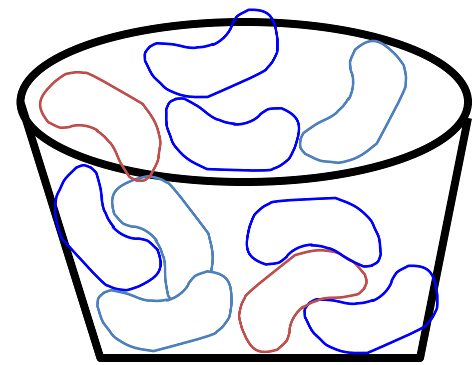
- Born London 1719
- Educated Edinburgh
  - Logic / Theology
- Known for tutorial on Newt
- Posthumously (1764) publis
  Solving a Problem in the Do
  - Addresses "inverse probabilit
  - Given some observations of a
    be said about the distribution.



Carnegie Mellon
THE ROBOTICS INSTITUTE

Probability

- Let A and B be discrete random variables.
- So long as A is a variable, P(A) is a function: P(A) : A → [0,1].
- For a specific value of A, like "red", P(red) is a number. By P(A=red) we mean
  - the probability that
  - the proposition that a red jelly bean was selected
  - is true.



Notation

- If f(x) = x2, f(y) usually means y2.
- Not so for probability.
- P(A) means an unspecified function over the domain of A and P(B) means a different function over the domain of B.
  - Concentrate on what's inside the ( ).
  - The A in P(A) determines the form of the function P.
- Could have P(A) = 1-A and P(B) = B2 {not 1-B}.

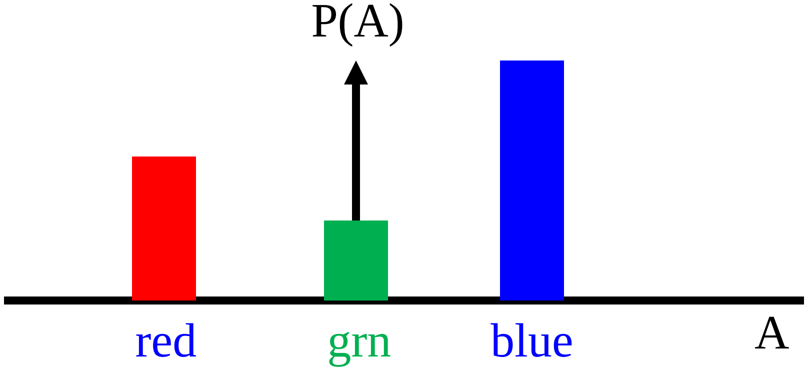
Carnegie Mellon THE ROBOTICS INSTITUTE

Variables and Events

- If A is a random variable, then an event is some statement about its value, like A=red.
- The variable A has a probability mass function or distribution like e^{-x^2}.
- The event A=red has a probability like 0.3.
- Sometimes we talk about several:
  - events (statements about values of variables)
  - variables (different random processes)

Predicates and Functions

- Sometimes P(A) means the probability that A takes a particular assumed value (like "true" if A is binary.
- Then, its best to write P(A=true).
- Other times P(A) means the entire distribution of probabilities for each possible value of A.



Negation

- $\overline{P}$(red) or P(¬A) is often used for the probability that the proposition a nonred bean was selected is true.
- Always:
  - P(red) = 1 – P(¬ red)

Odds

- Odds of event A:

$$O(A) = \frac{P(A)}{P(\bar{A})} = \frac{P(A)}{1 - P(A)} = \frac{1 - P(\bar{A})}{P(\bar{A})}$$

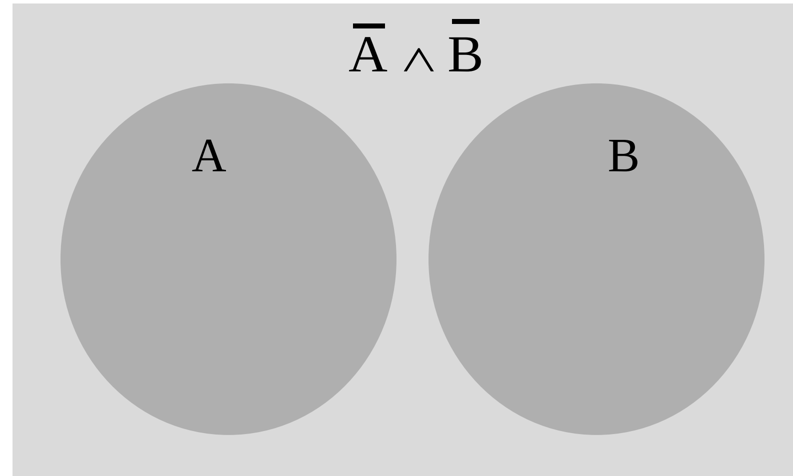
- Knowing any one of $P(A)$ $P(\bar{A})$ $O(A)$ determines the other two.

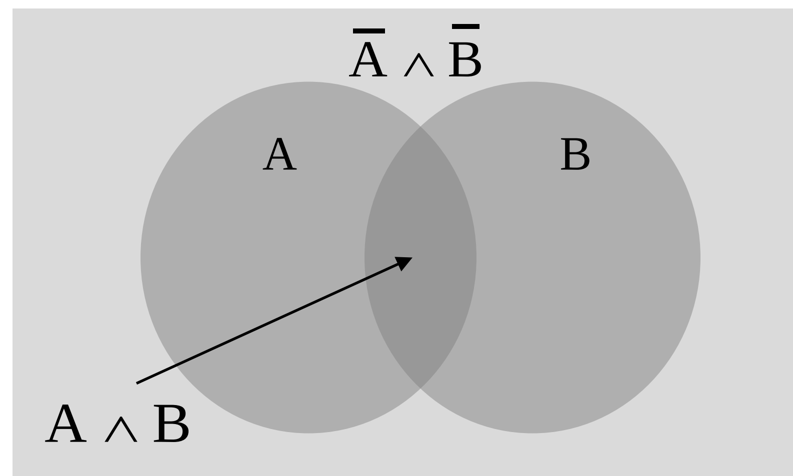
One equation
3 unknowns. Hmmmm

- Odds formulations and “log odds” = log[O(A)] can be very computationally efficient.

Venn Diagrams

- Imagine a process that selects <u>one</u> point in space with uniform probability.
- Label some regions.
  - The event “A” occurs when the point ends up in circle A etc.
- A point can only be in one place, so…..
- How many “disjoint” events can you list.
  - When the circles do not overlap, there are 3 elemental possibilities.
  - When the circles do overlap, there are 4 elemental possibilities.





Disjunction (OR)

- Probability of either A or B:

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

- When A and B are mutually exclusive (disjoint):

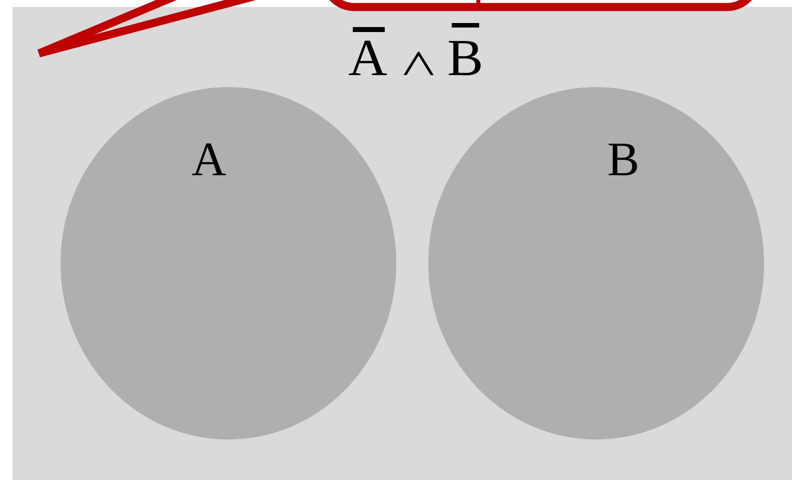
$$P(A \vee B) = P(A) + P(B)$$

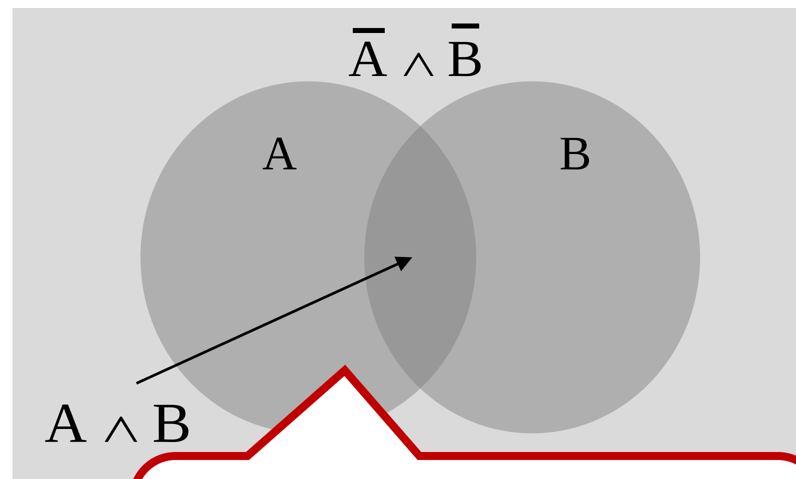
- Because:

$$P(A \wedge B) = 0$$

- Consider P(¬ grn v ¬ blu)
  – 2/3+2/3-1/3 = 1.

Prevents double counting any overlap





Exclusive means circles don’t overlap, so its impossible for both A and B to occur at same time for one point.

Conjunction (AND)

- When A and B are independent, the probability of both occurring is:

$$P(A \wedge B) = P(A) \times P(B)$$

- Which gives:

$$P(A) = \frac{P(A \wedge B)}{P(B)}$$

$$P(B) = \frac{P(A \wedge B)}{P(A)}$$

Dependence / Conditional

- P(A | B) means prob of A occurring "given that" B has occurred.
- Still talking about <u>one</u> point.

Its not about whether one point falls in A <u>after</u> another falls in B.

- When the circles overlap, the two events are dependent.
- If you know B is true, there is a slightly higher probability that A is true and vice versa.
- Dependence means…

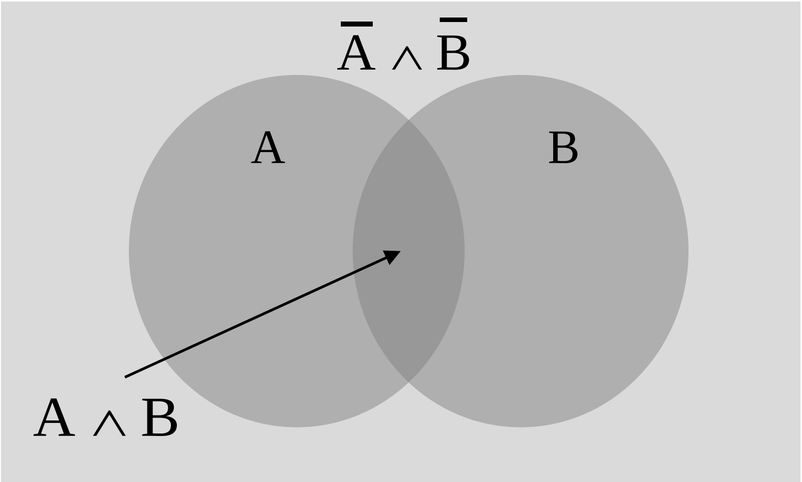
$$P(A|B) \neq P(A)$$

- Independence means

$$P(A|B) = P(A)$$

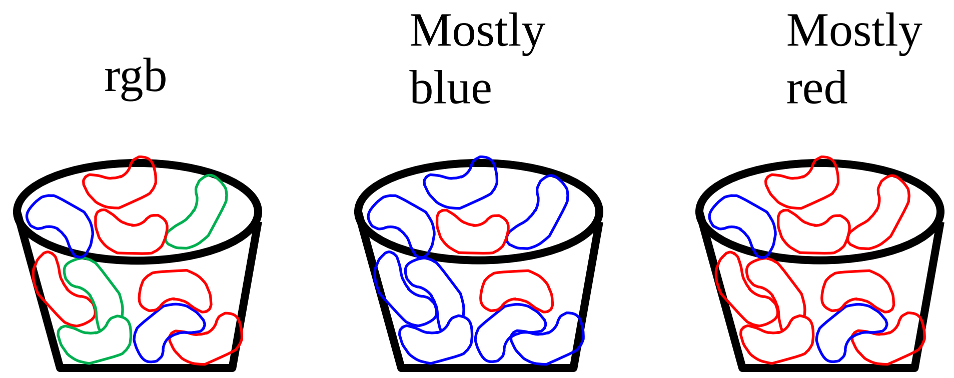
- Disjoint/Exclusive means:

$$P(A|B) = 0$$



Interpretation as a Staged Experiment

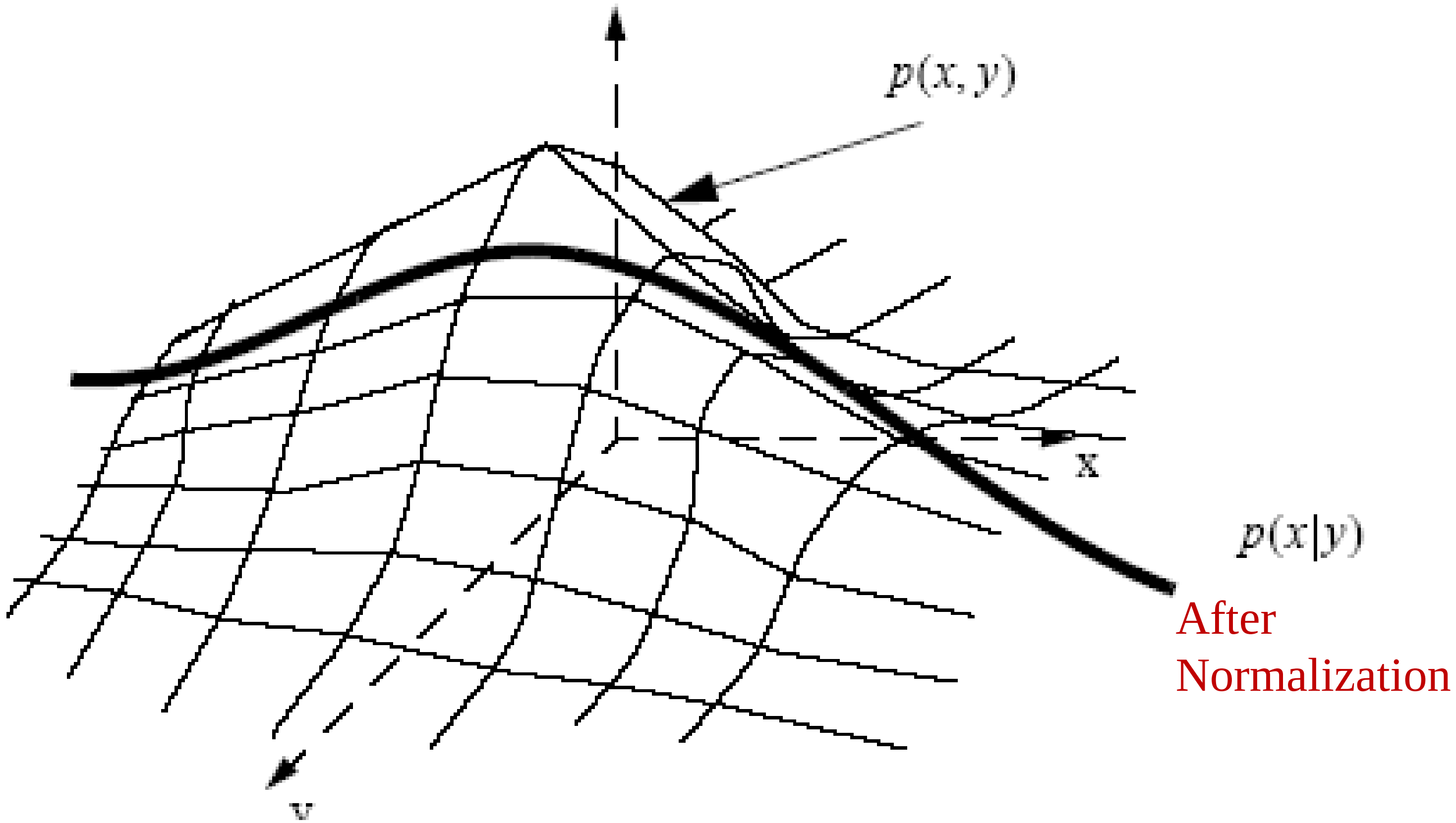
- P(A=red) means the prob a red bean is selected from any barrel.
- Stage 1:
  – P(B) means prob of selecting each barrel of jelly beans.
- Stage 2:
  – P(A=red|B) means prob of selecting a red bean from a specific barrel.



Three conditional prob functions

P(A|rgb)
P(A|mostly blue)
P(A| mostly red)

Interpretation as a Slice



- Variable y ceases to be random once its known.
- Then p(x) depends on it deterministically.

Carnegie Mellon THE ROBOTICS INSTITUTE

Interpretation as an Estimation Process

- Let A be the state of a system denoted x and let B be a measurement denoted z.
- Then:
  - $p(x)$ means the likelihood of the system being in state x.
  - $p(z)$ means the likelihood that a particular measurement is observed.
  - $p(x|z)$ means the likelihood of the system being in the state x if the measurement z is observed.
  - $p(z|x)$ means the likelihood of a measurement z being observed if the state is x
  - $p(x,z)$ means the likelihood of the system being in a state x and measurement z is observed.
- Every one of these is a different number $0<n<1$.

Total Probability and Marginalization

- Suppose n mutually exclusive events (like n different values of the variable B).

$$B_1 \ldots B_n$$

- If we know one (unknown) of these events Bi has occurred, then the probability of A is:

$$P(A|B_1 \ldots B_n) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \ldots + P(A|B_n)P(B_n)$$

- Consider: one of grn…blu has occurred..

$$P(\neg\, red|grn \ldots blue) = P(\neg\, red|grn)P(grn) + P(\neg\, red|blue)P(blu)$$

$$P(\neg\, red|grn \ldots blue) = (1)(1/2) + (1)(1/2) = 1$$

Not 1/3
Only grn or blu
were possible

- Integrates out a dimension in the PDF.

Carnegie Mellon
THE ROBOTICS INSTITUTE

Marginalization

4 Mutually Exclusive Events
All 4 add to 1

P(A,B)	A	¬A
B	A&B 0.1	¬A&B 0.2
¬B	A&¬B 0.2	¬A&¬ B 0.5

This row (B) is 0.3P(A|B)
P(B) = 0.3

This row (¬B) is 0.7P(A| ¬ B)
P(¬B) = 0.7

This col (A) is 0.3P(B|A)
P(A) = 0.3

This col (¬A) is 0.7P(B| ¬ A)
P(¬ A) = 0.7

Different Views

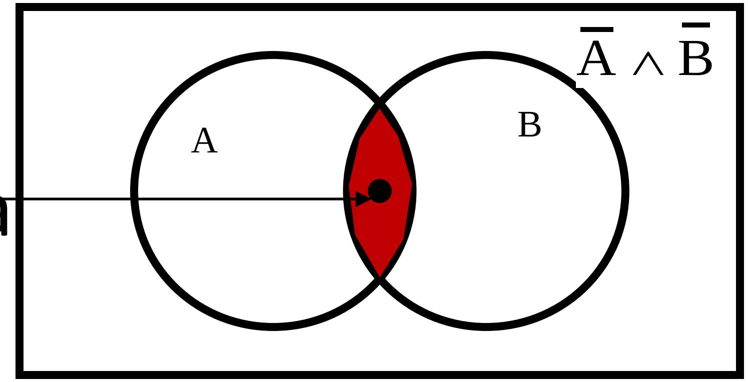
- $P(A|B)$ – conditional, A variable, B fixed
  - I happen to know B
  - Based on Bs value, how likely is each value of A.
- $P(A,B)$ – joint, A,B variable
  - Probability of each different pair of values (a,b)
- $P(A)$ – prior, A variable, B irrelevant
  - Probability of each value (a) if we knew nothing ("prior" to knowing something) about B.

Outline

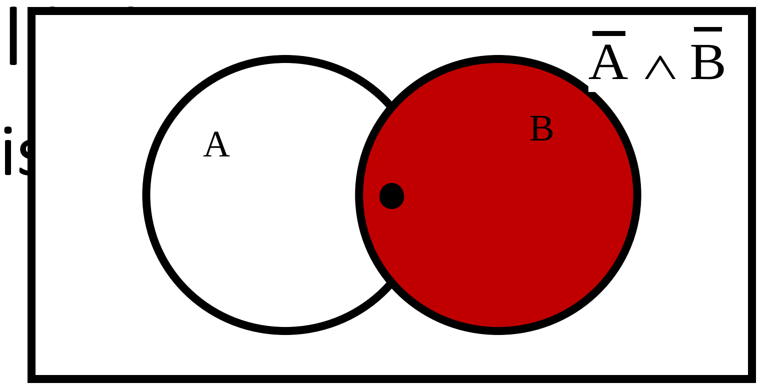
- 5.4 Bayesian Estimation
  - 5.4.1 Definitions
  - 5.4.2 Bayes' Rule
  - 5.4.3 Bayes' Filters
  - 5.4.4 Bayesian Mapping
  - 5.4.5 Bayesian Localization
  - Summary

Bayes Rule

- We can compute P(A|B).
- Given B has occurred, the ra
  in the overlap or not in it.



- If all points in B are equally l
  – Ratio of overlap to area of B i
    of falling in the overlap.



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

Other Forms

- Recall: $$P(A|B) = \frac{P(A \wedge B)}{P(B)}$$
- By symmetry: $$P(B|A) = \frac{P(A \wedge B)}{P(A)}$$
- Eliminate common expression:

$$P(A) \times P(B|A) = P(B) \times P(A|B)$$

- Rearrange:

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

Relates "Inverse" Probabilities

"Prior" prob of A

"Posterior" prob of A

- Computes P(X|Y) from P(Y|X).
- We care about:
  - P(x|z) → P(state | measurements)

Normalization

- In estimator notation:

$$P(X|Z) = \frac{P(Z|X)}{P(Z)} \times P(X)$$

- Can compute P(Z) using the total probability theorem.

$$P(Z) = \sum_{\text{all x}} P(Z|X)P(X) = \sum_{\text{all x}} P(X \wedge Z)$$

- Common notation is:

$$\eta(Z) = \frac{1}{P(Z)}$$

"Normalizer"

- Note that knowing P(Z|X) and P(X) completely determines P(Z).

Carnegie Mellon
THE ROBOTICS INSTITUTE

Outline

- 5.4 Bayesian Estimation
  - 5.4.1 Definitions
  - 5.4.2 Bayes' Rule
  - 5.4.3 Bayes' Filters
  - 5.4.4 Bayesian Mapping
  - 5.4.5 Bayesian Localization
  - Summary

Example Museum Tour Guide Robot

- Robot has sonar sensors.
- Sits idle until it detects someone in the room.
- Room has noisy fan nearby which corrupts sonar readings.
- Some visitors stand still for long periods.

Notation

- Use values of variables to imply the variables themselves.

$$P(X = \text{visitor}) \leftrightarrow P(vis)$$

$$P(X = \neg \text{visior}) \leftrightarrow P(\neg vis)$$

$$P(Z = \text{motion}) \leftrightarrow P(mot)$$

$$P(Z = \neg \text{ motion}) \leftrightarrow P(\neg \text{ mot})$$

Prior on X

- Room is empty 90% of the time.

X	P(X)
vis	0.1
¬ vis	0.9

Sensor Model P(Z|X)

- Fan noise makes the sensor barely effective.

P(Z|X)			
		Z	
		mot	¬ mot
X	vis	0.7	0.3
	¬ vis	0.6	0.4

0.7 only slightly higher than 0.6

0.3 only slightly lower than 0.4

Process First Measurement

- Bayes Rule:

$$P(X|Z_1) = \frac{P(Z_1|X)}{P(Z_1)} \times P(X)$$

$P(X|Z_1)*P(Z_1)$ or $P(Z_1|X)*P(X)$			
		Z	
		mot	¬ mot
X	vis	(0.7)(0.1)	(0.3)(0.1)
	¬ vis	(0.6)(0.9)	(0.4)(0.9)

P(Z|X)			
		Z	
		mot	¬ mot
X	vis	0.7	0.3
	¬ vis	0.6	0.4

X	P(X)
vis	0.1
¬ vis	0.9

Vector dot product !

Normally, only one column needs to be computed. The one for the Z you measured.

Normalizer

- Normalizer: $$P(Z_1) = \sum_{\text{all } x} P(Z_1|X)P(X)$$

Z_1	$P(Z_1)$
mot	0.61
¬ mot	0.39

$P(X\|Z_1)*P(Z_1)$ or $P(Z_1\|X)*P(X)$			
		Z	
		mot	¬ mot
X	vis	(0.07)	(0.03)
	¬ vis	(0.54)	(0.36)

Add up the columns

- Represents the prior likelihood of each sensor reading. What you would get if you:
  - measured continuously for a month,
  - paid no attention to whether there were visitors in the room or not, and
  - computed averages.

Normalize First Measurement

- Bayes Rule: $$P(X|Z_1) = \frac{P(Z_1|X)}{P(Z_1)} \times P(X)$$
- Divide by normalizer:

$P(X|Z_1)$ or $P(Z_1|X)*P(X)/ P(Z_1)$			
		Z	
		mot	¬ mot
X	vis	0.11	0.08
	¬ vis	0.89	0.92

0.1

0.9

$P(X|Z_1)*P(Z_1)$ or $P(Z_1|X)*P(X)$			
		Z	
		mot	¬ mot
X	vis	(0.07)	(0.03)
	¬ vis	(0.54)	(0.36)

Z_1	$P(Z_1)$
mot	0.61
¬ mot	0.39

- Normally only one column is computed

Prior on X is bumped up or down slightly to get $P(X|Z_1)$.

Recursive Bayesian Update

- First measurement is processed with:

$$P(X|Z_1) = \frac{P(Z_1|X)}{P(Z_1)} \times P(X)$$

- Suppose there is another measurement. P(X|Z1) (old posterior) becomes the new prior:

$$P(X|Z_1, Z_2) = \frac{P(Z_2|X, Z_1)}{P(Z_2|Z_1)} P(X|Z_1)$$

Baye's Rule for second measurement

- Put a Z1 after the | everywhere.

Markov Assumption

- Last slide had: $$P(X|Z_1, Z_2) = \frac{P(Z_2|X, Z_1)}{P(Z_2|Z_1)} P(X|Z_1)$$
- Assume Z2 is independent of Z1 when X is known (for any particular value of X).

$$P(Z_2|X, \cancel{Z_1}) = P(Z_2|X)$$

The Famous Markov Assumption

  – Intuitively, Z2 depends on X, but not on Z1.

- Baye's Rule becomes:

Now, we don't need a different table for every possible sensor measurement sequence.

$$P(X|Z_1, Z_2) = \frac{P(Z_2|X)}{P(Z_2|Z_1)} P(X|Z_1)$$

Process Second Measurement

- Bayes Rule: $$P(X|Z_1, Z_2) = \frac{P(Z_2|X)}{P(Z_2|Z_1)} P(X|Z_1)$$
- Consider only case of Z2=Z1 to avoid a 3D table.

$P(X|Z_1,Z_2)*P(Z_2|Z_1)$ or $P(Z_2|X)*P(X|Z_1)$			
		Z_1,Z	
		mot^2	¬ mot^2
X	vis	(0.7)(0.11)	(0.3)(0.08)
	¬ vis	(0.6)(0.89)	(0.4)(0.92)

$P(Z_2|X)=P(Z|X)$ (same)			
		Z_2	
		mot	¬ mot
X	vis	0.7	0.3
	¬ vis	0.6	0.4

$P(X|Z_1)$ (last result)			
		Z_1	
		mot	¬ mot
X	vis	0.11	0.08
	¬ vis	0.89	0.92

Normally, only one column needs to be computed. The one for the Z you measured.

Multiply these two element by element

Carnegie Mellon THE ROBOTICS INSTITUTE

Normalizer

- Normalizer:

$$P(Z_2|Z_1) = \sum_{\text{all } x} P(Z_2|X)P(X|Z_1)$$

$P(X|Z_1,Z_2)*P(Z_2|Z_1)$ or $P(Z_2|X)*P(X|Z_1)$			
		Z_1,Z	
		mot^2	$\neg mot^2$
X	vis	0.08	0.02
	¬ vis	0.53	0.37

Add up the columns

Z_1, Z_2	$P(Z_2|Z_1)$
mot^2	0.61
$\neg mot^2$	0.39

- Unchanged to 2 sig figs from last normalizer but different in general.

Carnegie Mellon THE ROBOTICS INSTITUTE

Normalize Second Measurement

- Bayes Rule: $$P(X|Z_1, Z_2) = \frac{[P(Z_2|X) \times P(X|Z_1)]}{P(Z_2|Z_1)}$$
- Divide by normalizer:

$P(X|Z_1, Z_2)$ or $P(Z_2|X)*P(X|Z_1)/ P(Z_2|Z_1)$

		Z	
		mot^2	$\neg\ mot^2$
X	vis	0.13	0.06
X	¬ vis	0.87	0.94

0.1

0.9

$P(X|Z_1, Z_2)*P(Z_2|Z_1)$ or $P(Z_2|X)*P(X|Z_1)$

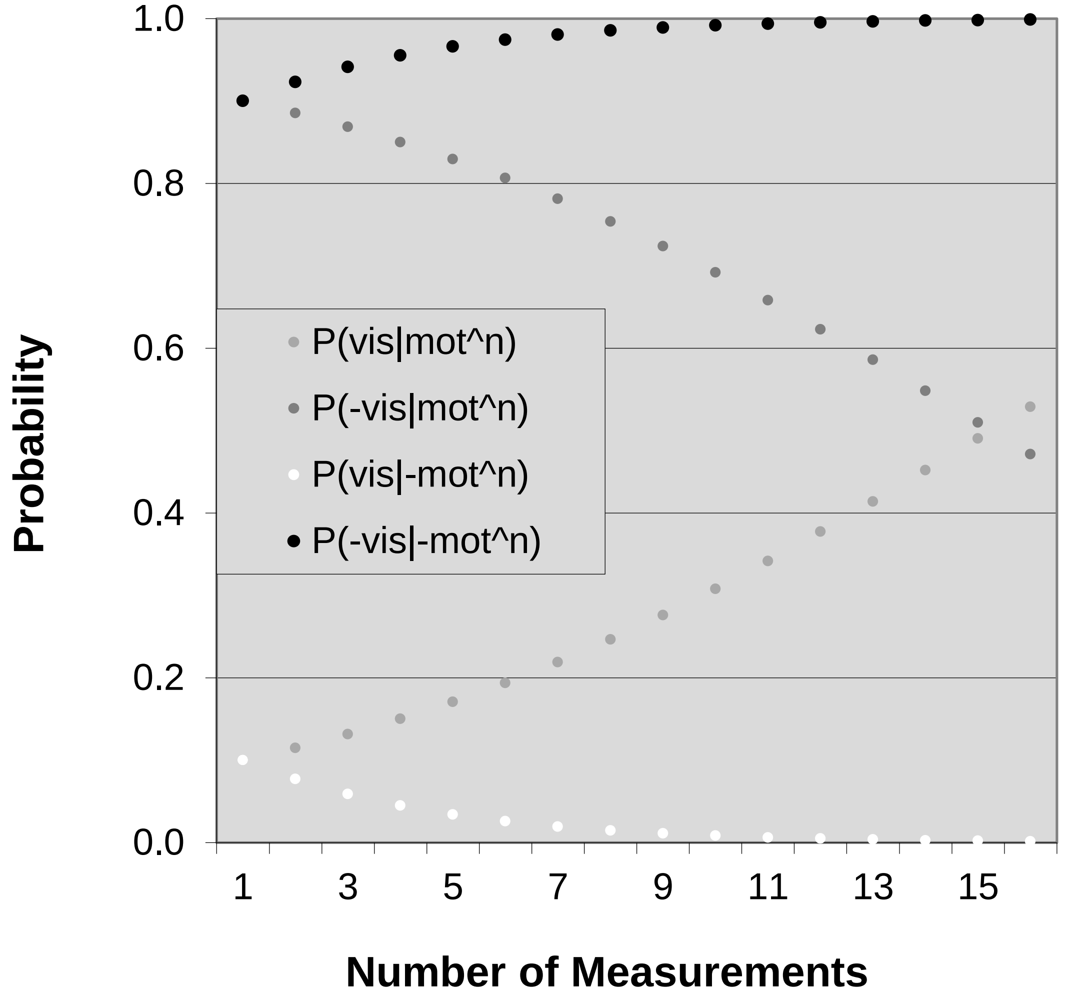
		Z	
		mot^2	$\neg\ mot^2$
X	vis	0.08	0.02
X	¬ vis	0.53	0.37

Z_1, Z_2	$P(Z_2\|Z_1)$
mot^2	0.61
$\neg\ mot^2$	0.39

- Normally only one column is computed.

Prior on X is bumped up or down slightly MORE.

Ad Infinitum



Multiple Measurements

- Last result is easy to generalize
  - Move all measurements so far after the |
- Denote all measurements so far as:

$$Z_{1,n} = Z_1, Z_2, \ldots, Z_n$$

- Bayes Rule in the multiple measurement form:

$$P(X|Z_{1,n}) = \frac{P(Z_n|X, Z_{1,n-1})}{P(Z_n|Z_{1,n-1})} P(X|Z_{1,n-1})$$

- With Markov assumption:

$$P(X|Z_{1,n}) = \left[\frac{P(Z_n|X)}{P(Z_n|Z_{1,n-1})}\right] P(X|Z_{1,n-1})$$

Multiple Measurement Normalizer

- Prior on Z is:

$$P(Z_n|Z_{1,n-1}) = \sum_{\text{all } x} P(Z_n|X, Z_{1,n-1}) P(X|Z_{1,n-1})$$

- Used in this form with Markov Assumption:

$$P(Z_n|Z_{1,n-1}) = \sum_{\text{all } x} P(Z_n|X) P(X|Z_{1,n-1})$$

- Can unwind the recursion to get this impressive result:

$$P(X|Z_{1,n}) = \left\{ \prod_{k=1}^{n} \left[\frac{P(Z_k|X)}{\sum_{\text{all } x} P(Z_n|X) P(X|Z_{1,n-1})} \right] \right\} P(X)$$

Π means Product

Carnegie Mellon THE ROBOTICS INSTITUTE

Bayesian Filters

- Define the "belief" function as the distribution over X given all evidence so far:

$$\mathrm{Bel}(X_n) = P(X|Z_{1,n})$$

- Then the normalizer is:

$$\eta(Z_{1,n}) = \{P(Z_n|Z_{1,n-1})\}^{-1}$$

- The normalizer is a constant scalar and the belief function is a distribution over X (a vector).

Carnegie Mellon THE ROBOTICS INSTITUTE

Bayesian Filter Algorithm

- Bayes_filter(Bel(X),Z):
- $\eta^{-1} = 0$
- For all x do
- $Bel'(X) = P(Z_n | X) \times Bel(X)$
- $\eta^{-1} = \eta^{-1} + Bel'(X)$ Accumulate normalizer
- For all x do
- $Bel'(X) = \eta^{-1} \times Bel'(X)$ Normalize
- return Bel'(x)

Carnegie Mellon
THE ROBOTICS INSTITUTE

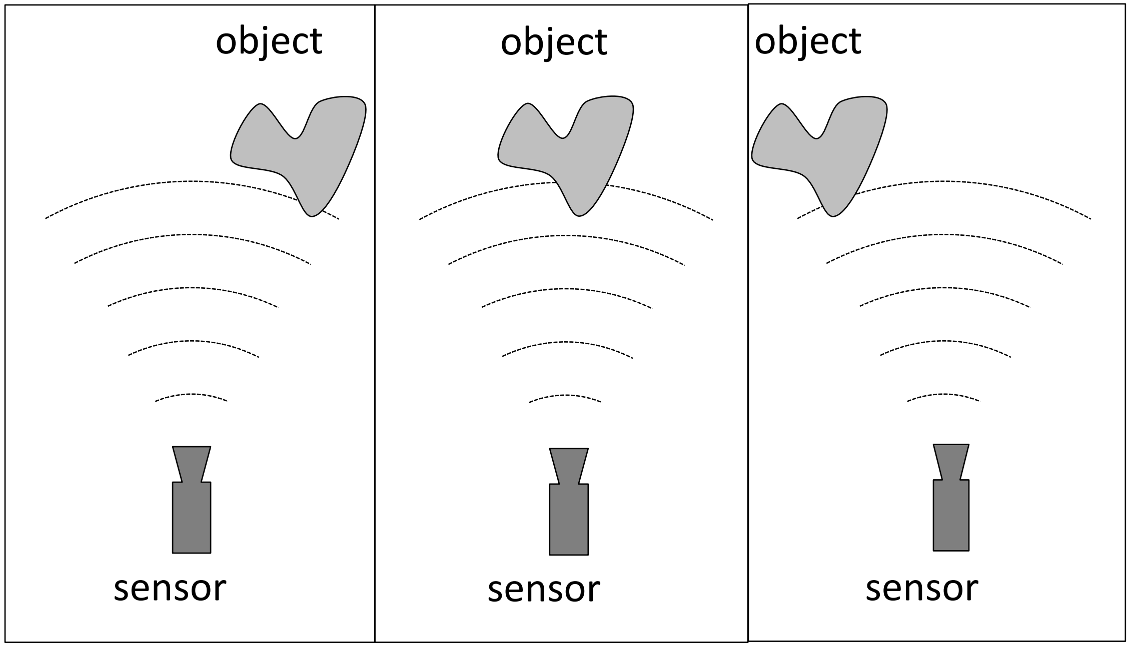
Outline

- 5.4 Bayesian Estimation
  - 5.4.1 Definitions
  - 5.4.2 Bayes' Rule
  - 5.4.3 Bayes' Filters
  - <u>5.4.4 Bayesian Mapping</u>
  - 5.4.5 Bayesian Localization
  - Summary

Certainty Grids

- Recursive Localizer:
  - Given sensor ranges and map, compute position
- Certainty Grid Mapper:
  - Given sensor ranges and position, compute map
- Originally proposed as a mechanism to deal with the poor angular resolution of sonar.

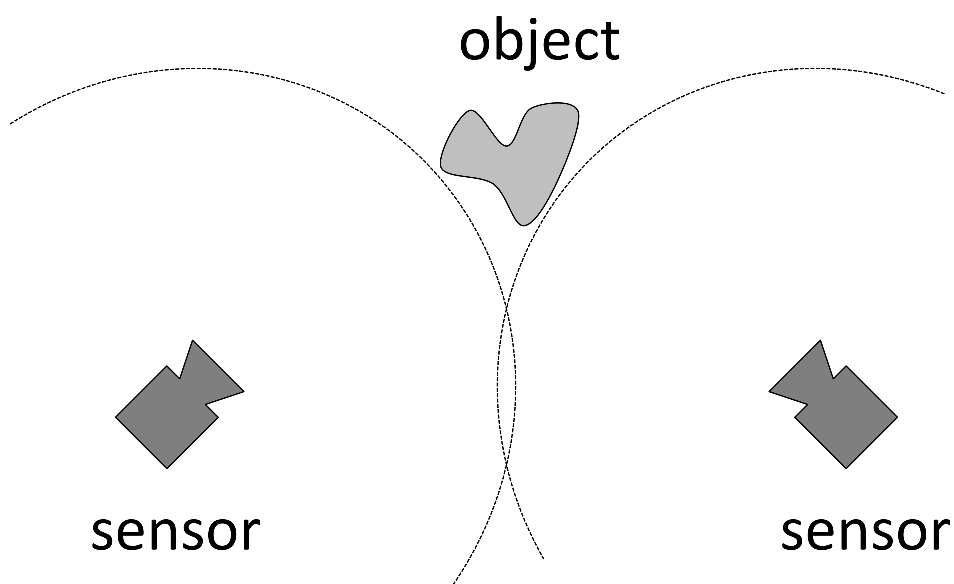
Sonar Angular Resolution



- The range to the object is known but either of the above three positions could generate the same range reading:
  - Angular resolution of a 30 degree sonar beam is poor.

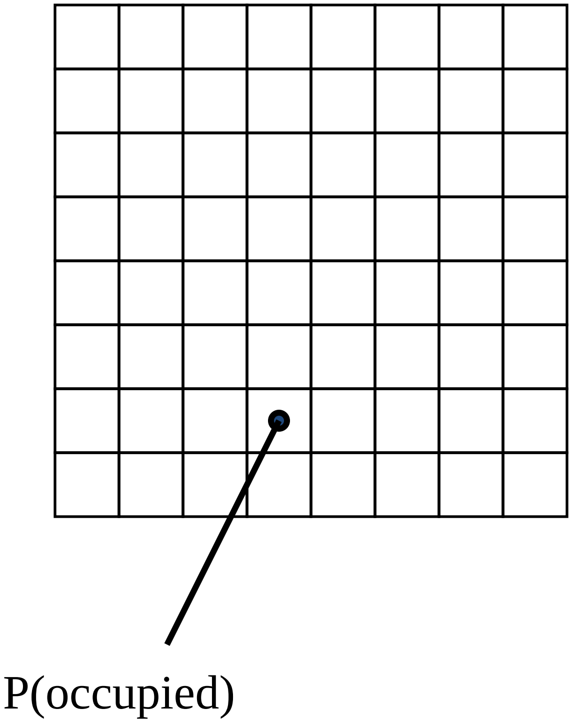
Synthetic Aperture

- Use of sensor motion (and accurate position) to achieve the improved angular resolution of a larger aperture (antenna radius).



Occupancy Grids

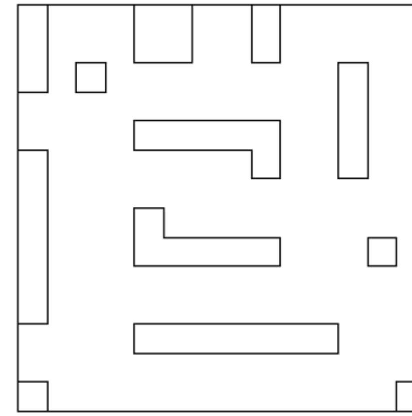
- Each cell encodes probability cell is occupied (by an obstacle).
- Really, a discrete approximation to a random field.



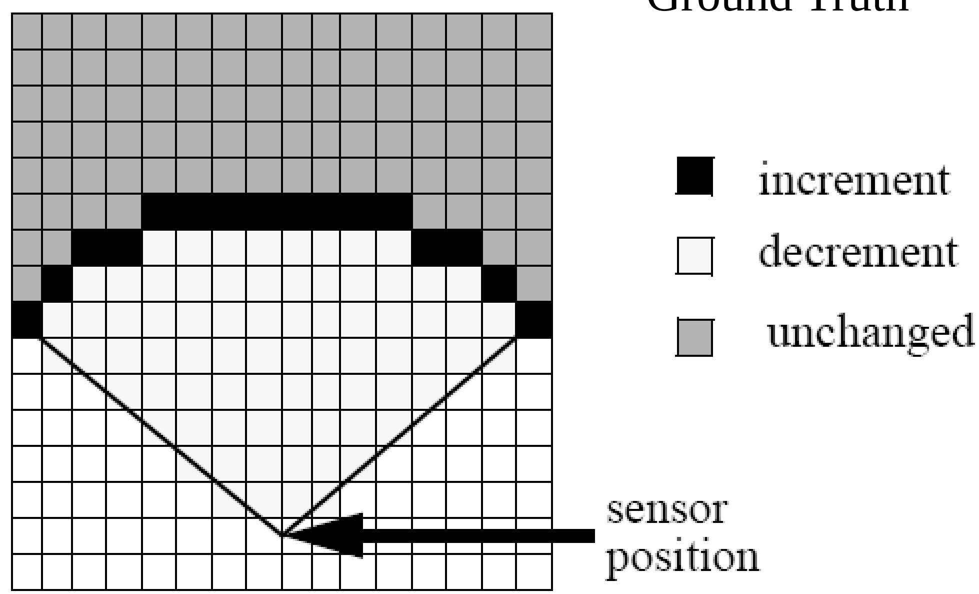
Carnegie Mellon
THE ROBOTICS INSTITUTE

Incorporating (Range) Measurements

- Simplest case is to count hits (maybe and
- Range reading R is evidence of:
  - Occupancy at R
  - No occupancy <R
- Tells you nothing beyond R.



Ground Truth



Bayesian Update

- Assume Independence:

$$P(occ[x_i, y_i] | occ[x_j, y_j]) = P(occ[x_i, y_i]) \qquad i \neq j$$

- Now can imagine a bank of Bayesian filters.
- P(X) for two values of X is just P(occ).
- Bayes Rule is:

Range Reading

$$P(occ|R_{1,k}) = \left[\frac{P(r_k|occ)}{P(r_k|R_{1,k-1})}\right] P(occ|R_{1,k-1})$$

- Also, for the other value of X.

$$P(\overline{occ}|R_{1,k}) = \left[\frac{P(r_k|\overline{occ})}{P(r_k|R_{1,k-1})}\right] P(\overline{occ}|R_{1,k-1})$$

Odds Update Formulation

- Take the ratio of the last two results:

$$\frac{P(occ|R_{1,k})}{P(\overline{occ}|R_{1,k})} = \left[\frac{P(r_k|occ)}{P(r_k|\overline{occ})}\right]\frac{P(occ|R_{1,k-1})}{P(\overline{occ}|R_{1,k-1})}$$

- Recall the definition of odds:

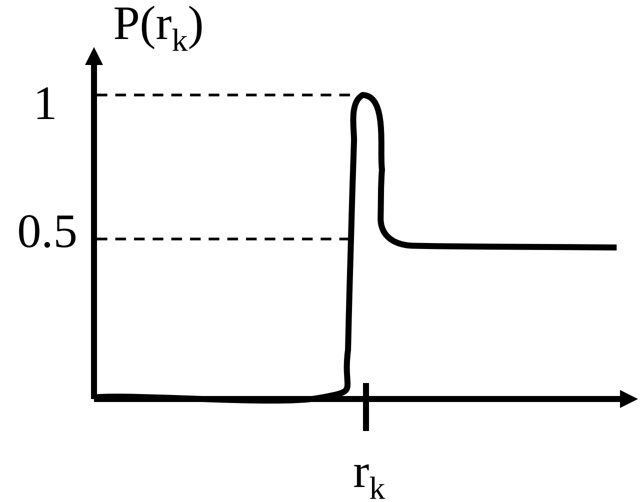
very computationally efficient

$$O(occ|R_{1,k}) = O(r_k|occ) \cdot O(occ|R_{1,k-1})$$

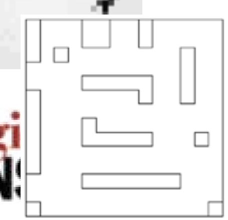
- Fill map initially with O(occ) prior and then multiply each cell by O(rk|occ) continuously.

Ideal Sensor Model

$$p(r_k|occ) = \begin{bmatrix} 0.5, \text{range(cell)} > r_k \\ 1.0, \text{range(cell)} = r_k \\ 0.0, \text{range(cell)} < r_k \end{bmatrix}$$

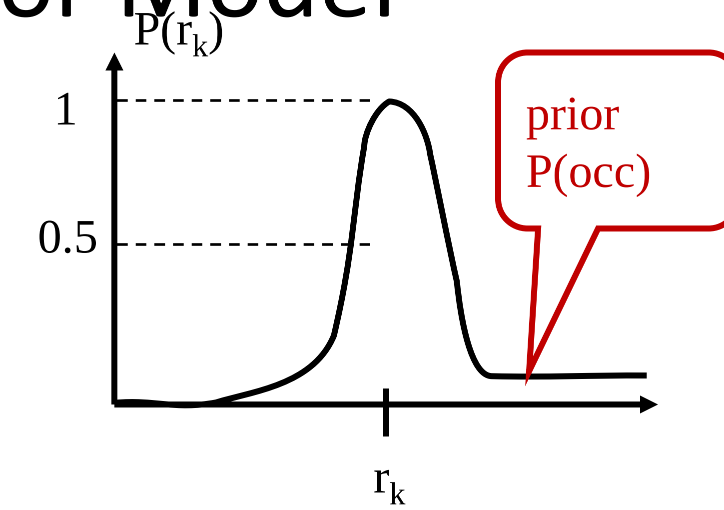


- Nothing can be said beyond th range.
- Trouble resolving corners.

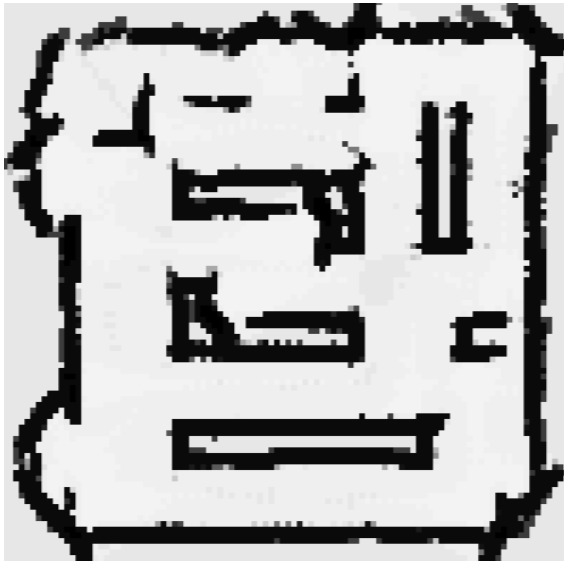


More Realistic Sensor Model

$$p(r_k|o) = \begin{bmatrix} 0.5, \ \rho > r_k \\ 0.99\exp\left[-\frac{1}{2}\left(\frac{\theta}{\sigma_r}\right)^2\right], \ (\rho - r_k) < \sigma_r \\ 0.05, \ \text{otherwise} \end{bmatrix}$$



- Allows for (models) sensor error.
- Models prior P(occ).
- Trouble resolving corners.



Modelling Dependence

- To capture dependence, use a sensor "map" of the inverse form.

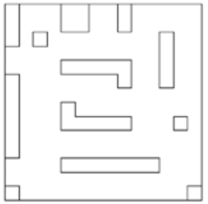
$$p(o|R_{k-1} \wedge r_k) = p(o|R_{k-1}) \times p(o|r_k) / p(o)$$

- Sensor model is sum of these two terms:

$$p_1 = \begin{bmatrix} 0.05\exp\left[-\frac{1}{2}\left(\frac{\rho - r_k}{\sigma_r}\right)\right] - 0.05,\ \rho < r_k \\ 0.0,\ \rho > r_k \end{bmatrix}$$

$$p_2 = 0.95\exp\left[-\frac{1}{2}\left(\frac{\rho - r_k}{\sigma_r}\right)^2\right]$$

- Better results.



Outline

- 5.4 Bayesian Estimation
  - 5.4.1 Definitions
  - 5.4.2 Bayes' Rule
  - 5.4.3 Bayes' Filters
  - 5.4.4 Bayesian Mapping
  - <u>5.4.5 Bayesian Localization</u>
  - Summary

NOTE

- Get the sonar processing stuff from the text.

Carnegie Mellon
THE ROBOTICS INSTITUTE

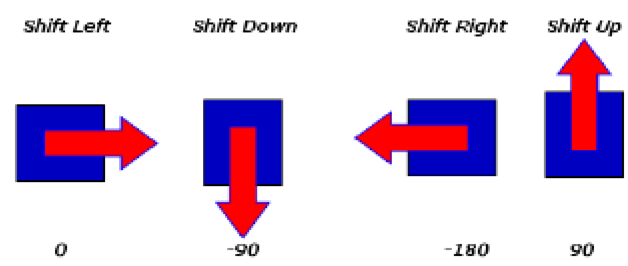
Bayesian Localization

- Sensor Model P(Z|X)
  - Probability of every range image given every state

$$P(Z_1 \mid X) = 1 - \left(\frac{1}{N \cdot r_{\max}} \left(\sum_{i=0}^{N} (r_i - \hat{r}_i)^2 \right)^{\frac{1}{2}} \right)$$

- Action Model

$$P(X_k \mid X_{k-1}, U_{k-1})$$



Bayesian Action Models

- Unlike measurements, actions tend to increase uncertainty.
  - None are executed perfectly
- Seek a pmf over state conditioned on the controls. Something like:

$$P(X_k | X_{k-1}, U_{k-1})$$

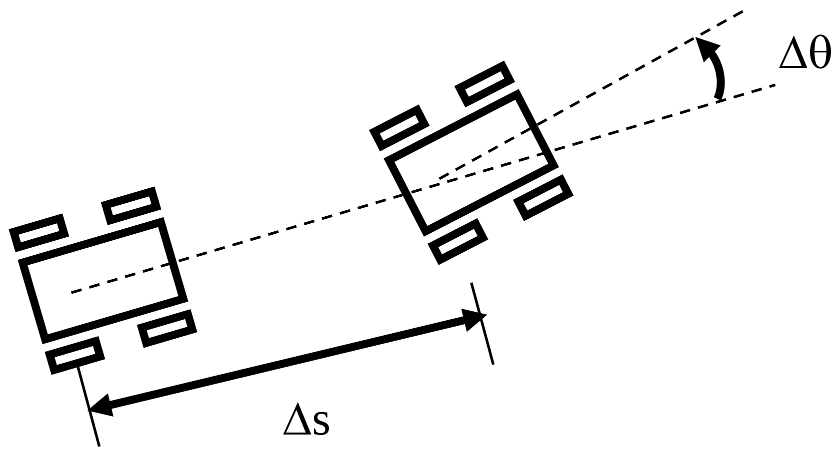
- This means the probability of
  - ending up in state Xk given that
  - the state was Xk-1 and
  - the control that was executed was Uk-1.

INPUTS FUNCTION JUST LIKE CONTROLS

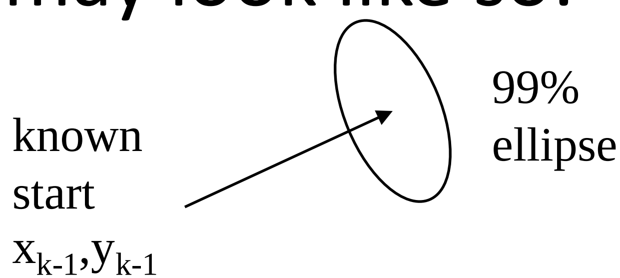
Carnegie Mellon THE ROBOTICS INSTITUTE

Action Uncertainty

- Suppose the actions are of the form:

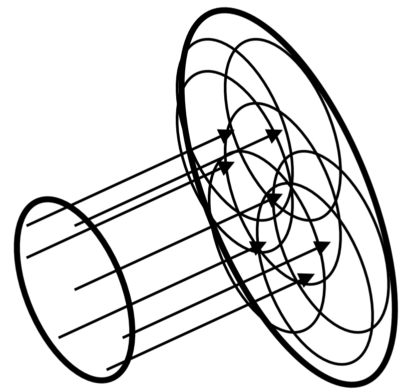


- If the start were known, the position part of the transition pmf may look like so:



Action Uncertainty

- If the start is unknown, project every possibility forward in time via marginalization over the previous state:



$$P(X_n|U_{1,n}) = \sum_{\text{all } X_{n-1}} P(X_n|X_{n-1}, U_{1,n-1})P(X_{n-1})$$

- Under Markov assumption:

$$P(X_n|U_{1,n}) = \sum_{\text{all } X_{n-1}} P(X_n|X_{n-1}, U_{n-1})P(X_{n-1})$$

Carnegie Mellon
THE ROBOTICS INSTITUTE

Action Uncertainty

- Output is a weighted smoothing operation on the input:

3: Compute likelihood of the transition given the input

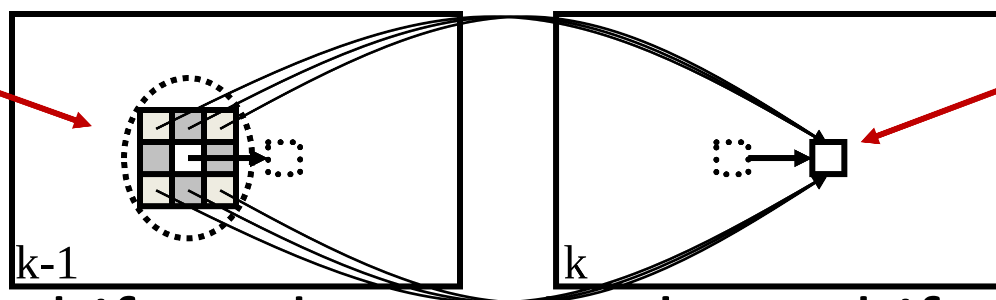
4: Weighted by likelihood of former state

$$P(X_n|U_{1,n}) = \sum_{\text{all } X_{n-1}} P(X_n|X_{n-1}, U_n)P(X_{n-1})$$

1: For each possible output state

2: For each possible former state

"All" possible input states



One possible Output state

- Intuition is shift-and-smooth where shift means shift the distribution by the nominal motion.

Bayesian Filter With Actions

- Algorithm Bayes_filter(Bel(X),D):
- $\eta^{-1} = 0$
- if D is a perceptual data item then:
- for all x do
- $Bel'(X) = P(Z_n | X) \times Bel(X)$
- $\eta^{-1} = \eta^{-1} + Bel'(X)$
- for all x do
- $Bel'(X) = \eta^{-1} \times Bel'(X)$
- else if is an action data item then:
- for all x do
- $Bel'(X) = \sum_{\text{all xn-1}} P(X_n | X_{n-1}, U_n) P(X_{n-1})$
- return Bel'(x)

Observe

Predict

Carnegie Mellon
THE ROBOTICS INSTITUTE

Outline

- 5.4 Bayesian Estimation
  - 5.4.1 Definitions
  - 5.4.2 Bayes' Rule
  - 5.4.3 Bayes' Filters
  - 5.4.4 Bayesian Mapping
  - 5.4.5 Bayesian Localization
  - Summary

Summary

- Bayesian estimation is more powerful than Kalman Filters
  - Can model arbitrary distributions.
- This generality comes at a computational cost.
- Can achieve impressive disambiguation through evidence accumulation.
  - Kidnapped robot problem
  - Localization in ambiguous, nearly repetitive environments.
- Still end up making assumptions in many cases.
  - Markov (zk independent of zk-1)
  - Spatial independence [P(x) independent of neighbors].