Informatics 1 Cognitive Science

Lecture 9: Bayesian Modeling

Frank Keller

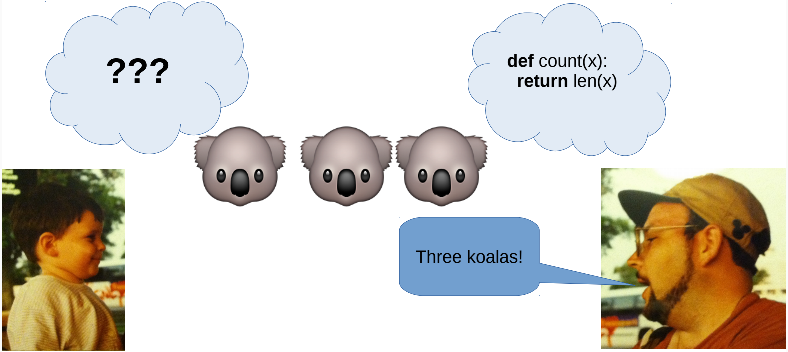
4 February 2025

School of Informatics
University of Edinburgh
keller@inf.ed.ac.uk

Slide credits: Frank Mollica, Chris Lucas, Mirella Lapata

Overview

Recap



- We observe input and infer the underlying cognitive technology.
- For example: we listen to speech and infer the underlying words.
- More generally: We experiment, we observe data and we update our beliefs.
- Last time, we looked at the use of transitional probabilities $P(y|x)$.
- Today, we will generalize this: Bayesian belief update.

Priors and Posteriors

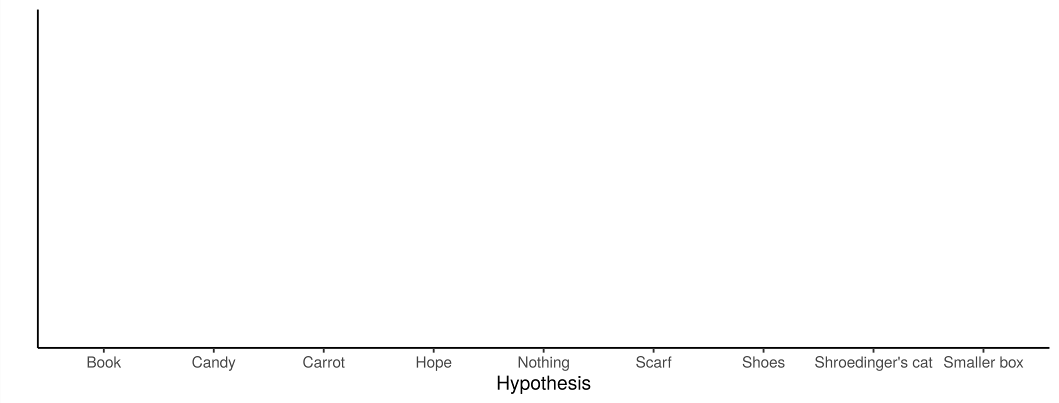
What's in the box?

What's in the Box?

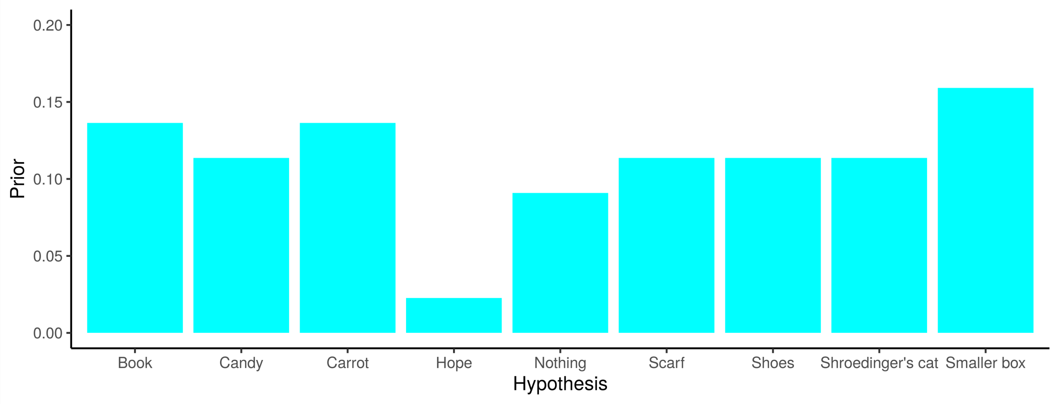
Time for a short quiz on Wooclap!

https://app.wooclap.com/YFCYTE

Hypothesis Space



Prior Beliefs



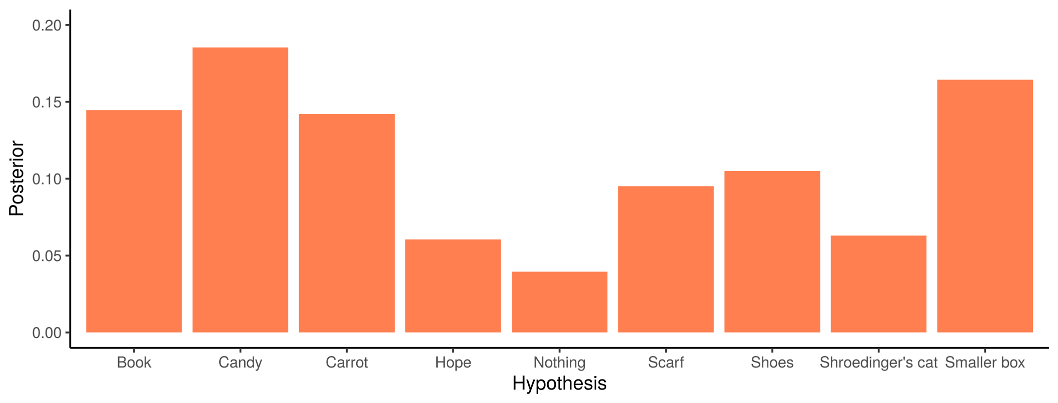
Evidence

Let's watch a short video about the box:

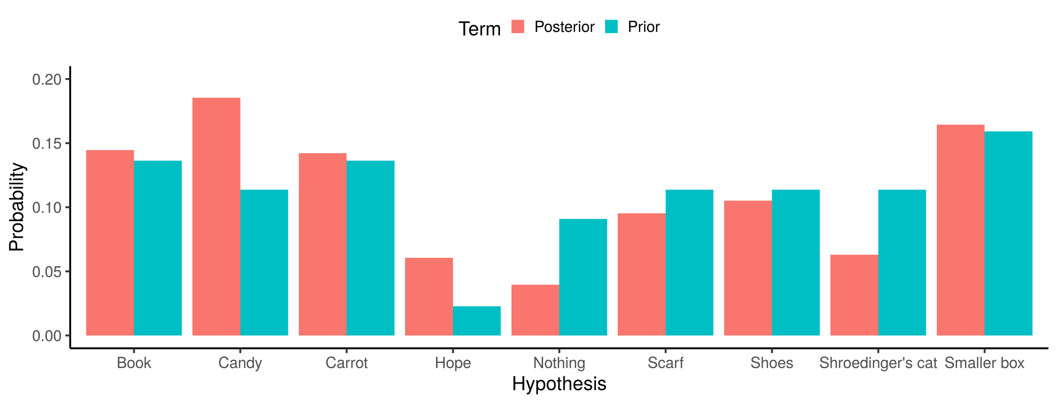
https://groups.inf.ed.ac.uk/teaching/cogsci/course/tutorials/Evidence.mp4

- The video provides additional evidence (data, observations).
- We use this to infer what's in the box.
- Specifically, for each hypothesis, if this hypothesis is true, how likely would it generate that noise?

Posterior Beliefs



Posterior Beliefs



Bayes Rule

Bayes Rule

is the calculus of belief updating:

$$P(\mathcal{H}|D) = \frac{P(D|\mathcal{H})P(\mathcal{H})}{P(D)}$$

- $P(\mathcal{H}|D)$ — **Posterior beliefs:** what you believe about hypotheses after seeing data
- $P(D|\mathcal{H})$ — **Likelihood:** how likely would a given hypothesis result in this data
- $P(\mathcal{H})$ — **Prior beliefs:** what you believe about hypotheses before seeing data
- $P(D)$ — **Evidence:** how likely is this data in general

Bayes Rule

is the calculus of belief updating:

$$P(\mathcal{H}|D) = \frac{P(D|\mathcal{H})P(\mathcal{H})}{P(D)}$$

Often we can't compute $P(D)$ directly, but we can obtain it by summing over all the hypotheses $h_i \in \mathcal{H}$:

$$P(D) = \sum_{h_i} P(D|h_i)P(h_i)$$

Example 1: Reference Resolution

Reference Resolution Example

As an example, let's look at a game. You need say a word to refer to one of the smileys.

Which word would you say so that a hearer can unambiguously pick out the smiley you are referring to?

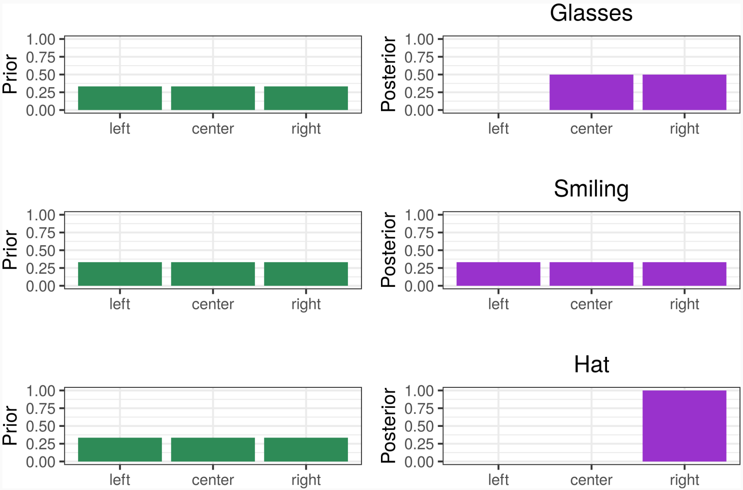
Reference Resolution Example

As an example, let's look at a game. You need say a word to refer to one of the smileys.

Which word would you say so that a hearer can unambiguously pick out the smiley you are referring to?

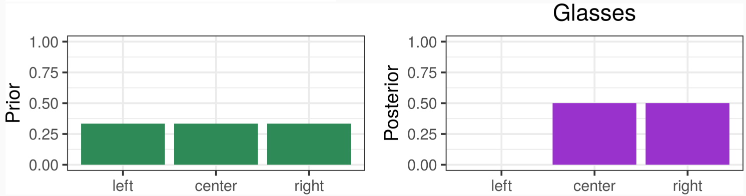
A priori, all three smileys are equally likely. When you say a word, you provide evidence based on which the hearer can choose correctly.

Reference Resolution Example

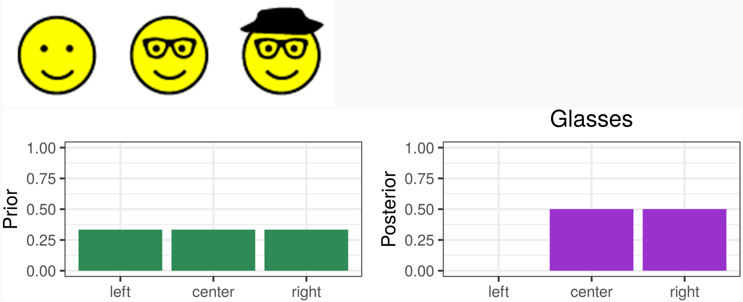


Depending on which word we say, the hearer infers different posterior beliefs.

Reference Resolution Example



Reference Resolution Example

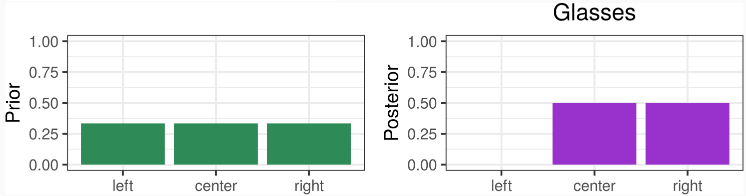


Assume the hearer uses Bayes Rule to choose the correct smiley:

$$P(h|d) = \frac{P(d|h)P(h)}{\sum_{h_i} P(d|h_i)P(h_i)}$$

h: hypothesis (here: which smiley); d: data (here: which word)

Reference Resolution Example

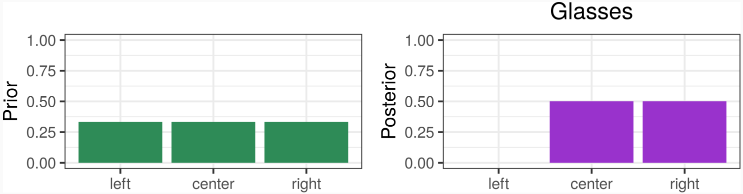


Assume the hearer uses Bayes Rule to choose the correct smiley:

$$P(h|d) = \frac{P(d|h)P(h)}{\sum_{h_i} P(d|h_i)P(h_i)} = \frac{0 \cdot \frac{1}{3}}{0 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3}} = 0$$

$h =$ left; $d =$ "Glasses"

Reference Resolution Example

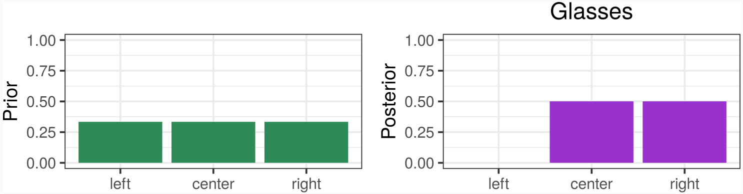


Assume the hearer uses Bayes Rule to choose the correct smiley:

$$P(h|d) = \frac{P(d|h)P(h)}{\sum_{h_i} P(d|h_i)P(h_i)} = \frac{1 \cdot \frac{1}{3}}{0 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3}} = \frac{1}{2}$$

h = center; d = "Glasses"

Reference Resolution Example



Assume the hearer uses Bayes Rule to choose the correct smiley:

$$P(h|d) = \frac{P(d|h)P(h)}{\sum_{h_i} P(d|h_i)P(h_i)} = \frac{1 \cdot \frac{1}{3}}{0 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3}} = \frac{1}{2}$$

$h = \text{right}$; $d =$ "Glasses"

Bayes Rule Terminology

Time for a short quiz on Wooclap!

https://app.wooclap.com/YFCYTE

Example 2: Decision Making

Decision Making Example

Problem

Peter and John are on vacation in a country where they don't know the local language. The city they are visiting has these awesome hot springs with monkeys and they want to go. Unfortunately, the hot springs do not allow clothing and Peter is really shy about bathing with women. Fortunately, the hot springs have one day a week for men and one day a week for women. Problem: This place is lo-tech and there is no signage to find out which day is for men only.

Solution?

Peter has an idea! Every day, they will pop by the springs and see who is leaving. Peter and John can use that information to solve their problem.

Decision Making Example: Hypothesis Space

	Hypothesis
h_1	Today is a women only day.
h_2	Today has no restriction.
h_3	Today is a men only day.

Decision Making Example: Prior

What do Peter and John know before they even leave the house?

	Hypothesis	Prior
h_1	Today is a women only day.	$\frac{1}{7}$
h_2	Today has no restriction.	$\frac{5}{7}$
h_3	Today is a men only day.	$\frac{1}{7}$

Decision Making Example: Likelihood

How likely would it be to see a man leave under each hypothesis?

	Hypothesis	**Prior**	**Like$_M$**
h_1	Today is a women only day.	$\frac{1}{7}$	0
h_2	Today has no restriction.	$\frac{5}{7}$	$\frac{1}{2}$
h_3	Today is a men only day.	$\frac{1}{7}$	1

Decision Making Example: Likelihood

How likely would it be to see a woman leave under each hypothesis?

	Hypothesis	**Prior**	**Like**$_M$	**Like**$_W$
h_1	Today is a women only day.	$\frac{1}{7}$	0	1
h_2	Today has no restriction.	$\frac{5}{7}$	$\frac{1}{2}$	$\frac{1}{2}$
h_3	Today is a men only day.	$\frac{1}{7}$	1	0

Bayesian Update

Decision Making Example: Update

So Peter and John arrive at the hot springs. They wait five minutes and see a man exit. Peter wants to be 90% sure it's men only before going in. Should they leave, go in, or wait longer?

$$P(\mathcal{H}|D) = \frac{P(D|\mathcal{H})P(\mathcal{H})}{P(D)}$$

	Hypothesis	Prior	Like$_M$	Post. Score
h_1	Today is a women only day.	$\frac{1}{7}$	0	$(0)(\frac{1}{7}) = 0$
h_2	Today has no restriction.	$\frac{5}{7}$	$\frac{1}{2}$	$(\frac{1}{2})(\frac{5}{7}) = \frac{5}{14}$
h_3	Today is a men only day.	$\frac{1}{7}$	1	$(1)(\frac{1}{7}) = \frac{1}{7}$

Decision Making Example: Update

$$P(\mathcal{H}|D) = \frac{P(D|\mathcal{H})P(\mathcal{H})}{P(D)}$$

	Hypothesis	Prior	Like$_M$	Post. Score
h_1	Today is a women only day.	$\frac{1}{7}$	0	$(0)(\frac{1}{7}) = 0$
h_2	Today has no restriction.	$\frac{5}{7}$	$\frac{1}{2}$	$(\frac{1}{2})(\frac{5}{7}) = \frac{5}{14}$
h_3	Today is a men only day.	$\frac{1}{7}$	1	$(1)(\frac{1}{7}) = \frac{1}{7}$

Evidence

$$P(D) = \sum_{h_i} P(D|h_i)P(h_i)$$

Decision Making Example: Update

$$P(\mathcal{H}|D) = \frac{P(D|\mathcal{H})P(\mathcal{H})}{P(D)}$$

	Hypothesis	**Prior**	**Like$_M$**	**Post. Score**
h_1	Today is a women only day.	$\frac{1}{7}$	0	$(0)(\frac{1}{7}) = 0$
h_2	Today has no restriction.	$\frac{5}{7}$	$\frac{1}{2}$	$(\frac{1}{2})(\frac{5}{7}) = \frac{5}{14}$
h_3	Today is a men only day.	$\frac{1}{7}$	1	$(1)(\frac{1}{7}) = \frac{1}{7}$

Evidence

$$P(D) = \sum_{h_i} P(D|h_i)P(h_i) = \frac{5}{14} + \frac{1}{7} = \frac{7}{14} = \frac{1}{2}$$

Decision Making Example: Update

$$P(\mathcal{H}|D) = \frac{P(D|\mathcal{H})P(\mathcal{H})}{P(D)}$$

	Hypothesis	Prior	Like$_M$	Post. Score	Posterior
h_1	Women only day.	$\frac{1}{7}$	0	0	0
h_2	No restriction.	$\frac{5}{7}$	$\frac{1}{2}$	$\frac{5}{14}$	$\frac{5}{7}$
h_3	Men only day.	$\frac{1}{7}$	1	$\frac{1}{7}$	$\frac{2}{7}$

Evidence

$$P(D) = \sum_{h_i} P(D|h_i)P(h_i) = \frac{1}{2}$$

Decision Making Example: Second Update

They wait another five minutes and see five more men exit. Peter wants to be 90% sure it's men only before going in. Should they leave, go in, or wait longer?

$$P(\mathcal{H}|D) = \frac{P(D|\mathcal{H})P(\mathcal{H})}{P(D)}$$

	Hypothesis	**Prior**	**Like$_M$**	**Post. Score**
h_1	Today is a women only day.	$\frac{1}{7}$	0	$(0)^6(\frac{1}{7}) = 0$
h_2	Today has no restriction.	$\frac{5}{7}$	$\frac{1}{2}$	$(\frac{1}{2})^6(\frac{5}{7}) = \frac{5}{448}$
h_3	Today is a men only day.	$\frac{1}{7}$	1	$(1)^6(\frac{1}{7}) = \frac{1}{7}$

Decision Making Example: Second Update

$$P(\mathcal{H}|D) = \frac{P(D|\mathcal{H})P(\mathcal{H})}{P(D)}$$

	Hypothesis	Prior	Like$_M$	Post. Score
h_1	Today is a women only day.	$\frac{1}{7}$	0	$(0)^6(\frac{1}{7}) = 0$
h_2	Today has no restriction.	$\frac{5}{7}$	$\frac{1}{2}$	$(\frac{1}{2})^6(\frac{5}{7}) = \frac{5}{448}$
h_3	Today is a men only day.	$\frac{1}{7}$	1	$(1)^6(\frac{1}{7}) = \frac{1}{7}$

$$P(D) = \frac{5}{448} + \frac{1}{7} = \frac{69}{448}$$

Decision Making Example: Second Update

$$P(\mathcal{H}|D) = \frac{P(D|\mathcal{H})P(\mathcal{H})}{P(D)}$$

	Hypothesis	Prior	Like$_M$	Post. Score	Posterior
h_1	Women only day.	$\frac{1}{7}$	0	0	0
h_2	No restriction.	$\frac{5}{7}$	$\frac{1}{2}$	$\frac{5}{448}$	$\frac{5}{69} = 0.072$
h_3	Men only day.	$\frac{1}{7}$	1	$\frac{64}{448}$	$\frac{64}{69} = 0.928$

$$P(D) = \frac{5}{448} + \frac{1}{7} = \frac{69}{448}$$

Summary

- Bayes Rule is a way of calculating the probability of a hypothesis, given some evidence (data).
- It includes four terms: posterior $P(\mathcal{H}|D)$; likelihood $P(D|\mathcal{H})$; prior $P(\mathcal{H})$; evidence $P(D)$.
- The prior encodes how probable a hypothesis is without any evidence, the likelihood encodes how probable it is to observe this evidence if the hypothesis is true.
- The posterior is the result of combining the two.
- Bayesian update: we can re-compute the posterior as we obtain more evidence for or against the hypothesis.
- Bayes rule provides a way of analyzing learning and inference.