Reasoning with Uncertainty

Mark van der Wilk

Department of Computing
Imperial College London

@markvanderwilk
m.vdwilk@imperial.ac.uk

January 16, 2023

Reasoning & Logic

To build “intelligent” systems, we need them to be able to “make statements” about the world.

Reasoning is all about

- going from knowledge that we already have, to
- statements about the task at hand.

Logic mechanises this process, e.g.

- Socrates is a human.
- All humans are mortal.

$\therefore$ Socrates is mortal.

Logic is a process

Logic is the **process** by which we derive new true statements given existing ones.

Reasoning & Logic

We can agree on a reasoning, but disagree about truth.

- Socrates is immortal.
- All humans are mortal ($\implies$ all immortals are not human).

$\therefore$ Socrates is not human.

Logic is a process

Logic is the **process** by which we derive new **beliefs** given existing ones.

- Beliefs in statements is subjective.
- The process to manipulate them is not.

Reasoning under Uncertainty

Knowledge is generally not completely certain:

- **Incomplete information** (epistemic)
  Someone has a secret, you just don't know it.
  A drug has an effectiveness, but we haven't determined it.
- **Inherent unpredictability** (aleatoric)
  Radioactive decay.
  Flipping a coin ... sort of.

E.g.:

- Dark clouds usually cause rain.
- My friend tells me there are dark clouds.

? How sure am I that it will rain?

Is there a well-defined **process**
for reasoning under uncertainty?

Requirements for Uncertainty-Aware Reasoning

A non-rigorous selection of requirements:

1. Degrees of belief (DoB) can be **ordered**,
   i.e. if $B(y) > B(x)$, and $B(z) > B(y)$, then $B(z) > B(x)$.
2. Reduces to **propositional calculus** (Boolean algebra) when dealing with full certainty.
   2.1 The DoB in a statement x and its negation $\bar{x}$ are related
   $$B(x) = f(B(\bar{x}))\,. \tag{1}$$
   2.2 The DoB in two statements simultaneously (x AND y) is related to the DoB in x given y is certainly true ($x|y$), and the DoB in y:
   $$B(x,y) = g(B(x|y), B(y)) \tag{2}$$
3. **Consistency** between different derivations of the same belief.

We take x, y, z to be logical statements, $\bar{x}$ to be the negation of the statement x, and $B(x)$ to be the belief in the statement x.

Probability satisfies requirements!

Example: Simple Implication

- Dark clouds usually cause rain.
- My friend tells me there are dark clouds.

? How sure am I that it will rain?

Taking binary variables R, C to indicate rainyness and cloudiness, our background knowledge $\mathcal{H}$ gives us beliefs in:

- $P(R = 1 \,|\, C = 0, \mathcal{H}) = r_0$, $P(R = 1 \,|\, C = 1, \mathcal{H}) = r_1$
- $P(C = 1|F, \mathcal{H}) = c$ (where F is what our friend told us)

We can derive our belief whether it will rain:

$$\begin{aligned} P(R = 1|F, \mathcal{H}) &= \sum_c P(R = 1|C = c, \mathcal{H}) P(C = c|F, \mathcal{H}) \\ &= r_1 c + r_0(1 - c) \end{aligned} \tag{3}$$

If we had full certainty about statements, e.g. $r_1 = 1, c = 1$, then this is *modus ponens*.

Example: Denying the Consequent

- Dark clouds usually cause rain.
- No water droplets are falling onto my head.

? Do I expect there to be dark clouds?

This time, we have defined our beliefs in:

- $P(R = 1 \,|\, C = 0, \mathcal{H}) = r_0$, $P(R = 1 \,|\, C = 1, \mathcal{H}) = r_1$
- $P(R = 0|\mathcal{H}) = d$

We can derive our belief of whether there will be clouds:

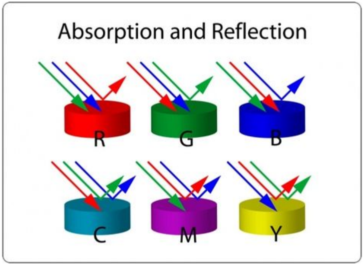
$$P(C = 1|\mathcal{H}) = \sum_r P(C = 1|R = r, \mathcal{H})P(R = r|\mathcal{H}) \tag{4}$$

$$P(C = 1|R = r, \mathcal{H}) = \frac{P(R = r|C = 1, \mathcal{H})P(C = 1|\mathcal{H})}{\sum_c P(R = r|C = c, \mathcal{H})P(C = c|\mathcal{H})} \tag{5}$$

Full certainty (e.g. $r_1 = 1, d = 1$) $\rightarrow$ *modus tollens*.
Note: we can't choose $P(C|\mathcal{H})$ independently of $P(R = 0|\mathcal{H})$, so this is to show the reduction to logic only.

Example: What colour is an object?

- We observe light L, reflected off an object with colour C, under illumination I.
- Given we observe light, what colour is the object?



$$P(C|L, \mathcal{H}_k) = \sum_i P(C|L, I = i) \, P(I = i|\mathcal{H}_k)$$

$$P(C|L, I = i) = \frac{P(L|C, I = i)P(C)}{P(L|I = i)}$$

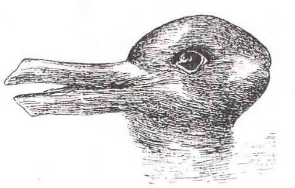
Representing uncertainty

White and Gold or Blue and Black?

- Lighting situation was **ambiguous**, so prior determined posterior.
- Prior is subjective, but we can agree on reasoning!

So what could explain this phenomenon?

- Our brains could have a biased prior,
- or the prior is 50-50, and we perceive our brain's best guess!



- Guess: Our brain jumps to one interpretation, even if the belief is only mildly stronger.
- Wouldn't it be better if we kept track of all possibilities?

Further Reading: Cox Axioms

We discussed that probability can represent **beliefs in statements**, and that in a certain way it extends propositional calculus to deal with uncertainty.

- Cox (1946, 1963) proposed **axioms** that such a way of "reasoning under uncertainty" should satisfy.
- Further claim: Probability theory is the **only** way to reason under uncertainty!
- This is controversial! Not everyone agrees to what degree this is proven, or which axioms are necessary.

See e.g. Van Horn (2003).

Beware of looking into this! You end up doing philosophy, which is interesting, but not helpful to get things done...

Further Reading: Dutch Books

Map beliefs to bets:
You take a bet if its expected value under your beliefs is positive.

- Example: For a coin for which you believe $P(\text{heads}) = p$, where you get a payout of **€1 if you win**, you will pay up to **€p to take the bet**.

A Dutch book is a

- set of bets that you **would take**,
- but that **always leads you to lose money**!

If you have **inconsistent beliefs**, then you a Dutch book exists, e.g.:

- Belief $P(\text{heads}) = 0.9$ **and** $P(\text{tails}) = 0.9$ (inconsistent)
- Bet 1: Pay €0.9 to bet on heads for a payout of €1.
- Bet 2: Pay €0.9 to bet on tails for a payout of €1.
- Result: You always lose €0.8.

Dutch book theorem: No Dutch book if you follow probability.

What is this course about?

Inference

Given my **understanding** of how the world works
Given **incomplete information** about the world
What do I know about what I don't observe?

- Use probability to represent subjective state of **uncertainty**.
- Reduction in uncertainty is **learning**!
- Rules of probability describe learning **process**.

What is this course about?

- Specifying models by **specifying beliefs** using probabilities.
- **Computing beliefs** about the world after seeing data.
- Using uncertainty when making **decisions**.

The course will be mathematical! (but it's not a maths course)

- We will discuss proofs (although focussing on the big picture).
- We will analyse the behaviour of algorithms.
- Exam will require you to demonstrate ability to apply mathematical principles.
- BUT, hopefully you gain an intuition into the principles too.

Course outline

I) Bayesian brainteasers (graphical models, tractable inference)
How do we put problems into the mathematical formalism?
What is a model? How do we formulate assumptions in models?

II) Gaussian processes
Specifying models, computing beliefs.

III) Decision Theory & Bayesian optimisation
Using uncertainty to take actions.

IV) Approximate inference
What happens if we cannot compute our beliefs exactly?

V) Modern applications, e.g. generative models
How is this used right now?

Why take this course?

Develop toolset of:

- (Bayesian) statistical methods[1]
  Useful to solve problems (next slide)
- Dealing with distributions in deep learning
  Generative modelling (e.g. VAEs, diffusion models)

Develop understanding of:

- Reasoning under uncertainty and ambiguity
  (Arguably): All good procedures must map onto Bayesian inference somehow
- The underlying process of what deep learning needs to solve
  Combining cues from disparate sources

[1] There is more to statistics than this course!

Why take this course?

I would recommend if you want to do

- data science, i.e. building models for careful prediction,
- development of new machine learning models / techniques,
- machine learning research.

Probably not necessary if you want to focus on implementing ML models, or ML infrastructure.

What problems can we solve?

- Low data prediction (how (un)certain am I?)
- Experiment design (what data should I gather next?)
- Data fusion (How should I combine information sources?)
- Learning (How should my belief change to match the world?)
- Decision making (Should I take a risk or play it safe?)

Prerequisites

- Good understanding of Mathematics for Machine Learning, e.g.:
  - Linear algebra (eigendecompositions etc)
  - Probability and basic statistics
  - Vector calculus
  - Gradient-based optimisation
- Python coding

`https://mml-book.com`

Expectations

What is expected for the exam:

- Knowledge of topics discussed
- Awareness of why topics are relevant
- Derive methods using mathematical concepts discussed
- Analyse methods using mathematical concepts discussed

How to study & revise:

- Join the lectures in person, engage, share your questions
- Think about how how theory applies in different settings
- Do the exercises

Highly Recommended Reading

Information Theory, Inference, and Learning Algorithms (MacKay, 2003)

- §2.1 (4pgs): Refresher of probability + notation we will use.
- §2.2 (1pg): Probability as belief.
- §2.3 (5pgs): Examples of Bayes rule. Exercises + solutions are very illustrative.

You really should read this.

See EdStem for links to books.

Practicalities

- Two assessed coursework (on Part II & III and IV).
  - Coding exercises assessed by unittests, designed to **teach**.
  - Exam is designed to **assess**.
- Q&A sessions most Fridays: Quickest way to get an answer to your question.
- Feedback: TAs+I are happy to give feedback on your solutions to exercises (e.g. if you want to know if you did enough steps) Friday Q&A session or on EdStem.
- I would like
- I will look at EdStem questions once per week.

Questions?

Hopefully this gives you an overview of the course.

Questions?

References I

Cox, R. T. (1946). Probability, frequency and reasonable expectation. *American journal of physics*, 14(1):1–13.

Cox, R. T. (1963). The algebra of probable inference. *American Journal of Physics*, 31(1):66–67.

MacKay, D. J. C. (2003). *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press.

Van Horn, K. S. (2003). Constructing a logic of plausible inference: a guide to cox's theorem. *International Journal of Approximate Reasoning*, 34(1):3–24.