LECTURE 4: REASONING

[Disclaimer: These informal lecture notes are not intended to be comprehensive - there are some additional ideas in the lectures and lecture slides, textbook, tutorial materials etc. As always, the lectures themselves are the best guide for what is and is not examinable content. However, I hope they are useful in picking out the core content in each lecture.]

Introduction

The structure of the lecture comes in three parts...

- Deductive reasoning
- Inductive reasoning
- Informal reasoning

... but as usual I've motivated it with a William James quote:

> "We talk of man (sic) being the rational animal; and the traditional intellectualist philosophy has always made a great point of treating the brutes as wholly irrational creatures. Nevertheless, it is by no means easy to decide just what is meant by reason"

... and as usual there's quite a lot going on here.

What *Is* Reasoning?

Firstly, there's the question of what exactly do we *mean* when we talk about humans reasoning?

- I see that "lemons are yellow"
- I say that "lemons taste a little bit like oranges"
- I claim that "lemons and oranges are both citrus"
- I argue that "because a slice of lemon goes well with a summer beer, I think that a slice of orange will also go well with a summer beer?"

It seems obvious that the last of these is indeed reasoning, and the first is probably not, but where is the dividing line, and what is it that makes the last one count as "reasoning"?

Intuitively, it seems to be something like "using what we do know (about lemons in beer) to make a reasonable inference about something that we do not know (about orange in beer)". Yet that seems a little unsatisfying... after all, you've just been through a whole series of lectures on perception, and it's

pretty obvious that the visual system actually does some pretty smart inference, using observable cues to make sensible inferences about the world. Why don't we count "vision" as a form of reasoning? In some ways, there's not that much of a difference between perceptual inferences (which we agree that many animals do just as well as we do) and "higher" reasoning (which we seem to think is special to humans, and perhaps to some extent genuinely is unique to us).

James' quote is touching on some of this tension, and the lecture goes on to highlight this tension in a few places, by noting some parallels between "higher" reasoning in humans and "lower" perceptual phenomena that seem to be pretty widespread.

What Counts As *Good* Reasoning?

The other undercurrent in the James quote, one that I don't really resolve in the lecture, is the question of what counts as *good* reasoning? In much of the historical literature on reasoning, there's a very strong presumption that some reasoning problems have clear "right answers and wrong answers", yet it doesn't always work out that way in real life. Something that looks like "good reasoning" from one perspective can often look frightfully silly from another perspective. Some care is required when deciding who is reasoning "well" and who is not!

These issues notwithstanding, it does seem like we can say something about the kinds of questions that we should be concerned with in the study of human reasoning:

- How is the truth of a claim established?

- What should we believe?

- Are there rules we should follow?

- What are these rules? (And do we follow them?)

Deduction Vs Induction

One of the big distinctions we draw between kinds of reasoning refers to the difference between *induction* and *deduction*.

- Deductive reasoning: Using facts to reach a "logically certain" conclusion

- Inductive reasioning: Using facts to reach a "plausible" conclusion (allows room for doubt)

Historically, deduction was considered the "highest" form of reasoning and induction considered less impressive. In more recent times, this is no longer the case -- because in real life the problems we tend to have to solve are inductive problems!

Part 1: Deduction

For the most part, the study of reasoning focuses on how we make *arguments* to support a particular conclusion, and in everyday life an argument can be constructed in all sorts of ways. We write opinion pieces in newspaper columns, we make snarky tweets, we thump on tables to prove a point etc.

On Syllogisms

However, much of the literature (at least in the European and Anglophone traditions...) derives from the study of **syllogisms**, which try to "formalise" logical arguments. For instance, we might write a syllogism like this:

```
All men are mortal
Socrates is a man
----------------------------
Therefore, Socrates is mortal
```

The two statments above the line are the **premises**, and these are the facts that we *assume* to be true. Then, if we assume those two things are true, the argument asks us to accept that the **conclusion** (the statement below the line) is also true.

In this particular argument, we have a **major premise** (all men are mortal) that asserts a general rule, a **minor premise** (Socrates is a man) that asserts a specific fact, and from these two premises we are asked to judge the truth or falseness of the conclusion.

In many contexts in psychology, we tend to reframe the major premise as a conditional statement (i.e., an "if X then Y" statement), so we might write the same argument like this:

```
If Socrates is a man, then he is mortal
Socrates is a man
----------------------------
Therefore, Socrates is mortal
```

When written in this form, we can treat the major premise as having two parts, the **antecedent** (If Socrates is a man...), and the **consequent** (...then he is mortal). Depending on what the minor premise says, we can construct four distinct arguments:

- "Affirming the antecedent" (*modus ponens*, MP) occurs when the minor premise asserts that the antecedent is true (e.g., Socrates is a man). The argument above is MP.

- "Denying the consequent" (*modus tollens*, MT) occurs when the minor premise asserts that the consequent is false. An MT argument might look like this:

```
    If Socrates is a man, then he is mortal
```

```
    Socrates is not mortal
    ----------------------------
    Therefore, Socrates is not a man
```

The critical point here is that MP and MT are both *valid* arguments: if you accept that the major and minor premises are both true, it is impossible for the conclusion to be false.

In contrast, suppose we were to try the other way around?

- "Denying the antecedent" (DA) occurs when the minor premise assers that the antecedent is false. For instance, I might state that "Socrates is a rock". That would give us this argument:

```
If Socrates is a man, then he is mortal
Socrates is a not a man (e.g., a rock)
----------------------------
Therefore, Socrates is not mortal
```

Clearly, we have an *invalid* argument here, insofar as the conclusion can be (and is!) false even though the premises are true. However, it's really important to notice that if I said "Socrates is a cat", the resulting argument would still be invalid:

```
If Socrates is a man, then he is mortal
Socrates is my cat
----------------------------
Therefore, Socrates is mortal
```

Although cats are indeed mortal and so in this case it turns out that the conclusion is true, there is nothing about the *structure* of the argument that proves it to be true. The fact that Socrates is mortal is entirely incidental to the argument, and as such the argument is deductively invalid

- "Affirming the consequent" (AC) occurs when the minor premise asserts that the consequent is true, which also yields an invalid argument:

```
If Socrates is a man, then he is mortal
Socrates is mortal
----------------------------
Therefore, Socrates is a man
```

As the example of "my cat" illustrates, this argument doesn't work. There are mortal creatures that are not human, so this argument is deductively invalid.

On The Psychology Of Deductive Reasoning

There is a large literature looking at how people evaluate deductive arguments, and the lecture only covers the basics. The very first study we talked about in the lecture, Barrouillet et al (2000), hinted that there are some differences between how kids reason about these problems and how adults do. In that study, adults seemed to be very good when the minor premise makes reference to the antecedent: so we notice that MP is valid and DA is invalid. In contrast, we are less sure when the minor premise refers to the consequent. In general, people are less sure whether MT arguments are valid, and less sure whether AC arguments are invalid. In contrast, children tend to adopt a simpler approach, and will agree with any argument that "affirms" either part of the major premise: so they tend to judge MP and AC as valid, and judge MT and DA as invalid. In short, even from the beginning of the lecture we had some hints that reasoning is tricky...

The Wason Selection Task

Much of the literature stems from the "selection task" introduced by Wason (1968). In this task, people are told that they need to "test" a conditional rule. Imagine a set of cards with a letter on one side, and a number on the other. Suppose I assert that:

> If there is an R on one side of the card, then there is a 2 on the other

Your job is to determine whether this rule is true. You are then shown four cards, one showing an R, another showing a G, the third showing a 2 and the last showing a 7. Which cards need to be turned over to test the rule?

When these experiments are run, people tend to select the R and the 2. Intuitively, this feels sensible. Yet, suppose we were to write these as syllogisms. Here's what we get:

```
If there is an R then there is a 2
There is an R
----------------------
Therefore, there is a 2
```

This argument is "modus ponens" and it is deductively valid. If we were to turn over the R card and were to see something other than a 2, it *must* be the case that the rule were false. So this is a good card to choose.

What about the 2 card? Well, if we write this as an argument we get this:

```
If there is an R then there is a 2
There is a 2
----------------------
Therefore, there is an R
```

This argument "affirms the consequent" and it is deductively invalid. If I turn over a 2 and see something other than an R, that *doesn't* tell me anything. Even if I did see an R, it doesn't *prove* anything. It's not immediately clear how selecting the 2 card is helpful... so it's initially somewhat puzzling that people tend to select it

It's even more puzzling when you consider the fact that selecting the 7 card will produce a modus tollens argument, which is also deductively valid:

```
If there is an R then there is a 2
There is a 7
-----------------------
Therefore, there is not an R
```

Because this argument is valid, we can potentially use it to *falsify* the rule -- that is, if we were to turn over the 7 card and discover an R there, we would be able to show that the rule is untrue.

So what does this mean? There are three main ideas...

Interpretation #1: People Are Stupid

The original interpretation of the Wason task was to suggest that people just aren't very good at deductive reasoning. After all, the "hypothesis testing" (rule testing) puzzle seems like the sort of deductive reasoning that we should be engaged in -- and indeed when he first proposed it Wason was influenced by prominent philosophers of science like Karl Popper who argued that scientists should be trying to falsify hypotheses -- and yet we don't do it. So maybe we're just not very smart...

Interpretation #2: Deductive Reasoning Is Stupid

An alternative possibility (which I didn't talk about much in the lecture but is mentioned in the textbook) was suggested by Mike Oaksford and Nick Chater in the late 1990s, arguing that the problem isn't with human reasoning, but with Wason's assumption that we *ought* to be doing deductive reasoning in this task. They argue that what people are actually trying to do is solve an *inductive* reasoning problem, in which the goal is to obtain the "most possible information" about whether the rule is true or false. There's a lengthy argument in their papers, and it involves more maths than is really appropriate for this class, but the short version is that under most "real world" assumptions about rules, it actually makes a lot of sense to pick the R and 2 cards - if you think of the task as a kind of "question and answer game with nature", the R and 2 cards actually correspond to the most *informative* queries you can make. In any case, because the lecture doesn't talk much about this particular interpretation, I should move on to talk about...

Interpretation #3: The Task Is Stupid

One thing that is very obvious about the Wason task is that it feels a bit abstract and unnatural. In real life I don't usually reason about rules that pertain to cards that have letters and numbers printed on them. However, what I do spend a lot of time doing is checking to see if people are following social rules. So what happens if we reframe the selection task more in terms of a "cheater detection" problem? (Sperber & Girotto 2002)

More specifically, consider two different kinds of rules:

- **indicative conditionals** (e.g., if it is Monday then I wear black) state an abstract regularity about the world that may or may not be true

- **deontic conditionals** (e.g., if it is Monday then I MUST wear black) assert a social norm that I am expected to follow

In lecture we talked about an example of a deontic task. Suppose you're working at a bar, and you need to ensure that no underage drinking of alcohol takes place. Here are four people. Who do you need to check?

- A 15 year old drinking something (but you don't know what)

- A 65 year old drinking something (but you don't know what)

- Someone drinking a cup of tea (but you don't know their age)

- Someone drinking a beer (but you don't know their age)

It is immediately obvious in this situation that the people you need to be most worried about are the 15 year old and the beer drinker. When framed as arguments, we get...

```
If Ms Polly is under-18 she must not drink beer
Ms Polly is 15
---------------------
Therefore Ms Polly must not drink beer
```

Deciding to check her ID seems to be a form of modus ponens. Similarly, checking on the beer drinker appears to rely on modus tollens:

```
If Mr Tim is under-18 he must not drink beer
Mr Tim is drinking a beer
---------------------
Therefore Mr Tim must be 18 or over
```

Again, we have a deductively valid argument that appears to drive our behaviour.

The key point here is that when framed in terms of a problem that we're very familiar with, people seem to do much better at the task.

So it might be that the original version of the problem underestimates people's actual reasoning abilities?

Mini Summary

Logical reasoning

- Definitions of deductive and inductive reasioning
- Syllogisms and how they work
- Definitions of valid and invalid reasoning
- Four argument types: MP, MT, DA and DC

Empirical evidence

- Developmental changes?
- Wason selection task
- Indicative vs deontic versions

Part 2: Inductive Reasoning

Inductive arguments rely on limited evidence to make a conclusion seem more plausible. We can again write them in the form of syllogisms...

```
Socrates was mortal
Aristotle was mortal
Cicero was mortal
----------------
Therefore all humans are mortal
```

Obviously, this isn't a deductively valid argument. We've got three specific examples to work from, and there's no way we can *prove* a general rule from such little evidence. Even so, it does seem like a reasonable inference, especially if we have many many examples of humans that are mortal. However, there's never any guarantee that it will work. For instance, this...

```
Socrates was male
Aristotle was male
Cicero was male
----------------
Therefore all humans are male
```

... is rather silly.

Phenomena In Inductive Reasoning

As mentioned in the lecture, and hinted at in my commentary on the William James quote above, the lecture on similarity did touch on inductive reasoning already: the fact that we use similarity to make judgments about the properties of novel objects (e.g., this tomato is edible because other tomatos were edible) is a form of inductive reasoning.

In my experience, reading the literature on inductive reasoning has a slightly different "feel" to deductive reasoning. A lot of the way we talk about deductive problems tends to focus on whether people are doing it "right". However, the main reason (in my personal opinion, at least) we can do this is that deductive reasoning is so simple that it's actually possible for us to *know* what is or is not "correct" deductive reasoning. However, almost every real world problem (e.g., what colour pants should I wear? What will the lecturer want me to write on an exam paper?) is an inductive problem, and it's never possible to know for sure what the right answer is. Inductive problems are *hard*... and as a consequence, the focus is less on trying to determing whether people are doing it right or wrong (though there is indeed quite a lot of that) the focus tends to be more on trying to *describe* the kind of inductions that people consider reasonable.

In the lecture I referred to a study by Osherson et al (1990), a classic paper in the field that provided a long list of inductive phenomena. Three of them in particular we talked about in class:

- **Premise-conclusion similarity**. An inductive argument feels stronger when the premise and the conclusion are similar. For instance, it feels more plausible to argue that "*dolphins* possess the TH4 gene, therefore *seals* (probably) possess the TH4 gene" than to argue that "*dolphins* possess the TH4 gene, therefore *mice* (probably) possess the TH4 gene". As mentioned in the similarity lecture, this is what we use similarity for: things that are perceived to be similar are assumed to have more properties in common, so we are more willing to accept an inductive argument that pertains to two very similar entities.

- **Premise diversity**. An inductive argument feels stronger when it is based on diverse evidence. If I tell you that *dolphins*, *whales* and *seals* all possess the TH4 gene, you would probably be unwilling to guess that *cats* possess the TH4 gene. In contrast, if I tell you that *dolphins*, *elephants* and *mice* all possess the TH4 gene, you might be more willing to endores the idea that *cats* have the gene. Again, this makes sense: in the first argument it seems most likely that TH4 is specific to marine mammals, whereas in the second argument it seems more likely to be a general property of mammals.

- **Premise monotonicity**. An inductive argument feels stronger when it is based on more evidence. If I tell you that *cows* have the TH4 gene, you might be willing to conclude that *horses* do too, but you'd probably be even more convinced if I told you that *elephants*, *sheep*, *zebras* and *dogs* all possess the TH4 as well. The more widespread the property seems to be, the more willing we become to believe that a newly encountered entity will share that property.

Of course, these rules don't always hold. Not only are there exceptions to them, there are systematic

exceptions, and there are some interesting tricks you can do to explore what those exceptions tell us about human reasoning (see next lecture!) But all else being equal these rules seem to hold in most cases.

Mini Summary

- Difference between induction and deduction
- Phenomena in inductive reasoning
    - Premise-conclusion similarity
    - Premise diversity
    - Premise monotonicity

Part 3: Informal Reasoning

A different line of research into human reasoning has tended to focus on the kinds of "informal argumentation" that we do in everyday life. That is, while some "reasoning fallacies" occur because people fail to follow deductive logic (e.g., denial of the antecedent is a logical fallacy), there are many other kinds of argument (e.g., argument from personal incredulity, argument ad hominem) that are fallacies for rather different reasons! In the third part of the lecture, we focused on two kinds of informal argument that are logically invalid, but nevertheless can be convincing to people sometimes... and not necessarily for bad reasons!

The lecture here draws heavily from a paper by Ulrike Hahn and Mike Oaksford in 2007.

Arguments From Ignorance

The argument from ignorance asserts that

> "X must be true because you can't prove X is false"

Clearly this is deductively invalid. But does it have some degree of inductive strength? Can it be persuasive sometimes? And if people do find it persuasive sometimes, does that happen for a sensible reason?

To illustrate that arguments from ignorance seem to vary in their persuasiveness, compare the following:

- "Ghosts exist because there is no scientific proof that they do not"
- "There's no Hatfield train stop in Sydney, because it's not shown the Metro map"

These are both arguments from ignorance. The ghosts argument (arguably correctly) states that there is no positive proof in the scientific literature of the impossibility of ghosts, and then asks us to believe that

this constitutes evidence for the existence of ghosts. Most people find it implausible.

The train stop argument states (correctly) that there is no postitive evidence of Hatfield on the Metro map, and then asks us to believe that this constitutes evidence for the non-existence of a Hatfield stop. Most people find it plausible.

There are several differences between the two arguments, but one of the main ones is the **epistemic closure** assumption. The key idea here pertains to the "knowledgeability" or "completeness" of the evidence source. An epistemically closed source is one that contains all relevant facts... the Metro map is indeed epistemically closed (or at least we expect it to be). Transit authorities construct the maps in order to list *all* stops on the network... so the absence of a stop named "Hatfield" on the map actually is evidence that there is no such stop in the world. If there *were* such a stop in Sydney, we'd expect it to be on the map (epistemic closure), so the argument from ignorance is actually pretty compelling.

In contrast, consider the ghosts argument. The source here is the "scientific literature on ghosts", and it's not really epistemically closed. There isn't a very large literature on ghosts, and even if there were, I'm not sure I'd expect it to contain a proper "proof" that ghosts are impossible. So the absence of any such proof really doesn't tell me very much.

In the lecture, I mention a study (Oaksford & Hahn 2004) that illustrates that people really are sensitive to the kinds of things that we "ought" to be sensitive to (e.g., epsitemic closure) when evaluating arguments from ignorance, suggesting that there is often a very sensible basis for these arguments.

Circular Arguments

A circular argument involves

> "Assuming that X is true in order to prove that X is true"

and it's surprisingly common in everyday life. In lecture, I listed three examples:

- God exists because the Bible says so, and the Bible is the word of God

- Inductive reasoning is justified because it has worked in the past, so it will work in the future

- Electrons exist because we can see 3-cm tracks in a cloud chamber, and 3-cm tracks in a cloud chamber are the signatures of electrons

In the first case I assume Christian theology to prove Christian theology; in the second I assume inductive reasoning works in order to prove that inductive reasoning works; and in the third case I use the theory of particle physics to prove the theory of particle physics. All of these are "circular". So why do we make these arguments all the time?

Part of the answer can be found by noting that our own perceptual system "falls" for circular arguments. Consider the Kanizsa triangle illusion (see slide 93 for the picture). If we wrote it out as an argument, it would look like this...

- There is a white triangle because it is blocking the black circles and the black triangle ... and we assume there's a black triangle and black circles because there's a white triangle blocking them

... which is so obviously a circular argument that there must be a more sensible answer to "why do we do this?" than merely to assert "people aren't good at reasoning". After all, our visual systems are rather good at "seeing", so maybe some of what we're doing with circular arguments is kind of smart?

What Hahn & Oaksford (2007) argue is that a circular argument is implicitly an "appeal to an explanatory system". Whenever we're asked to accept a circular argument, what we're really being asked to do is believe that the "explanatory system" in question provides the best account of the observed facts. For the religious and scientific examples above, the arguments can be rewritten like this:

- (God and Bible) are part of an explanatory system called "Christianity", and the argument asks you to accept that the Christian faith provides the best explanation for the observations under discusion

- (Particle physics and cloud chamber experiments) are part of an explanatory system called "physics", and the argument asks you to accept that physics provides the best explanation for the observations under discussion

Neither of these is self-evidently wrong (or right). They are inductive arguments, and they should be evaluated as such. One test of the idea that this is how people actually think about circular arguments is to supply people with a plausible (or implausible) alternative to the circular argument. If there is a *better* explanation available, then the circular argument should seem weak. If there is no plausible alternative, then the circular argument should seem more reasonable. Hahn & Oaksford tested this by presenting people with dialogues such as this one:

```
John:   I think there's a thunderstorm
Anne:   What makes you think that?
John:   I just heard a loud noise that could have been thunder
Anne:   That could have been an airplane
John:   I think it was thunder, because I think it's a thunderstorm
Anne:   Well, it has been really muggy around here today
```

Clearly, John's argument is circular, but do people find it reasonable or acceptable? In one condition of the experiment, participants were told that *"John and Anne are in their trailer home near the airport"*, which makes Anne's airplane argument seem more plausible; in the other condition, they were told that *"John and Anne are in their camper van at their woodland campsite"* which makes Anne's argument seem much less plausible. What they found is that people rated John's argument as stronger when Anne's alternative was implausible.

Mini-Summary

- "Rational" explanations of fallacies?
- Examples:

- Argument from ignorance (epistemic closure)
- Circular arguments (appeal to explanatory system)