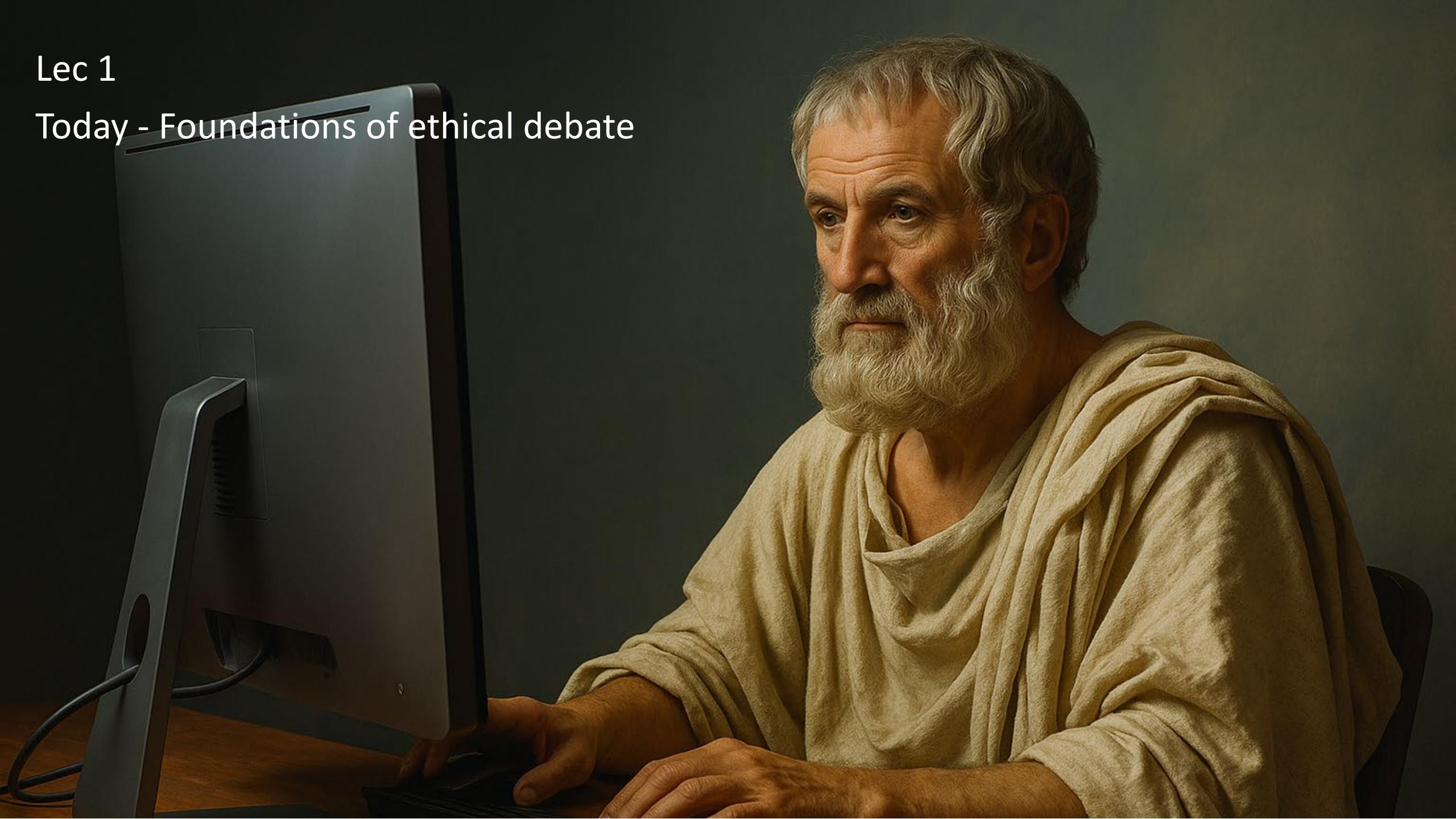


Lec 1

Today - Foundations of ethical debate



Instructor: Vasco Castela

- BA Comm Studies (Lisbon)
- MA Philosophy of AI/Cognitive Science (@ COGS – Sussex, UK)
- PHD Ethics (Manchester, UK)
- vcastela@bcit.ca
- SW2-255 – DTC 617 - Office hours – by appointment (in person or Teams)

How the class will run

- You don't need to buy a textbook. All course material is available online (course notes + assigned readings + videos + ...)
- Lecture: ethical theory primarily
- Labs: revisions and application of ethical theory to cases in computing
- Lab instructors: Shulamit and Simon
- (lab instructors do all the grading)
- Laptop use in class is not allowed
- You can take notes with anything that doesn't have a physical keyboard

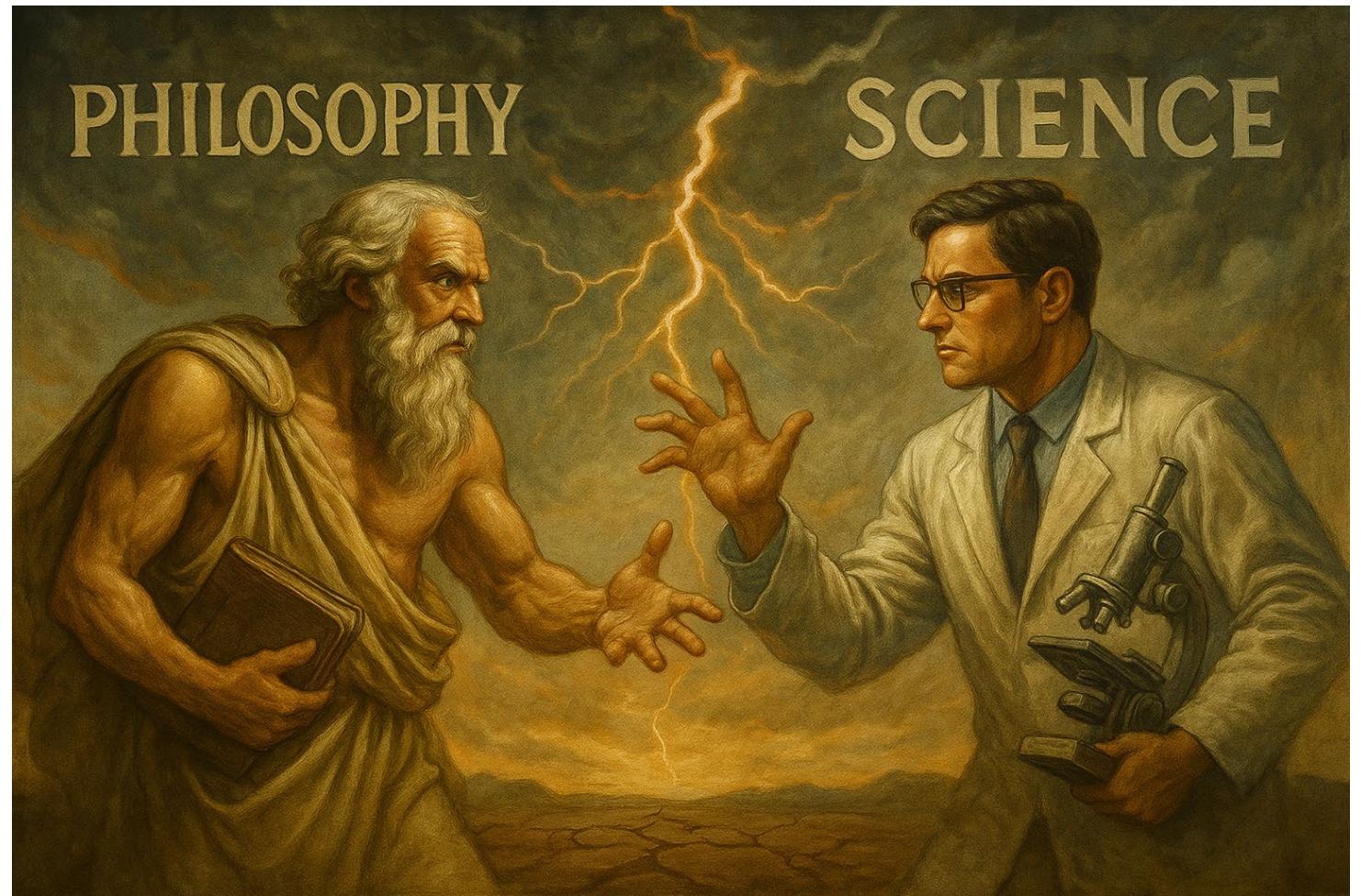
Grading components

	%	Comments
Essay + Essay Prep	10	Essay prep 1&2 (2) + 800-word Essay 1 to be written in lab (8)
Essay 2	15	Essay 2 - 800 words - in lab
Midterm	20	Lec 1-5
Participation	12	Tutorials only - frequency and quality of participation
Attendance	12	Lectures only (note: sign attendance sheet in+out and take weekly quiz to get attendance mark)
Quiz	1	Final exam prep quiz - 1 for completion (reasonable attempt required)
Final exam	30	short answer + essay (case study)
Total	100	

Course schedule

Week	Wed	Lec	Lecture - theory	Applied reading
1	07-Jan	1	Introduction to ethics	Employment I - Resume Fraud/Padding
2	14-Jan	2	Conventional morality + essay prep 1 (lab)	Employment II - Conflicts of interest vs loyalty
3	21-Jan	3	Rejecting ethical relativism + essay prep 2 (lab)	Cybercrime - Is it all a matter of perspective?
4	28-Jan	4	Act utilitarianism I	Internet Fraud I - Hacking
5	04-Feb	5	Act utilitarianism II + essay 1 in lab	Internet Fraud II - Dark Patterns
6	11-Feb	6	Kant I	Property rights I - Copyright, patents, trade secrets
7	18-Feb	7	Kant II	Property rights II - Open source
8	25-Feb	-	Midterm week	
9	04-Mar	8	Virtue Ethics I	Remote work I - Trust, honesty and productivity
10	11-Mar	9	Virtue Ethics II + essay 2 in lab	Remote work II - Remote workplace bullying
11	18-Mar		Reading week	
12	25-Mar	10	Rawls and Nozick	Contracts - Non-compete clauses and NDAs
13	01-Apr	11	Friedman - Shareholder theory	Data I - Fiduciary obligations
14	08-Apr	12	Solomon/Stakeholder theory	Data II - Privacy and Negligence
15	15-Apr		Revisions	
16	22-Apr		Final exams	

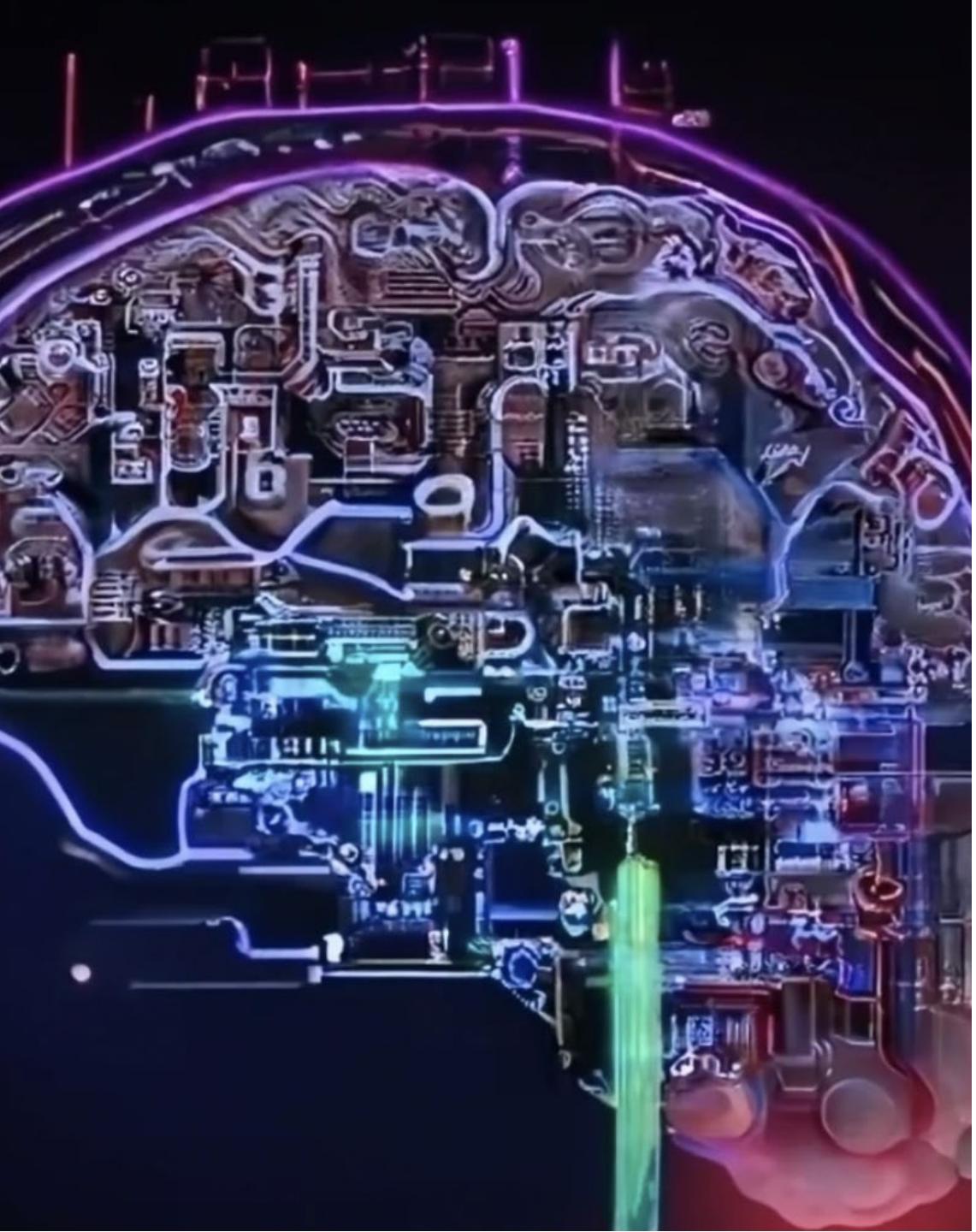
- What is philosophy?
- Do we still need it now that we have science?
- What type of problems does it deal with?



Philosophy

- In some cases, philosophy just makes problems more tractable. What is life? Why do things move? In many cases, sciences arose from the work of philosophers (e.g. Aristotle developed the basis of biology, mechanics and economics)
- Once empirical research can deal with it, then science takes over the problem or set of problems
- **Questions remain that cannot be settled by empirical research *alone*.** “Is it wrong to lie to my boss? Why?”
- At the edge of science, philosophers are still at work, so we have philosophy of mathematics, philosophy of physics, etc. along with the traditional areas of philosophical research that continue to resist pure empirical analysis
- Right, first automaton, Archytas (400 BC), mechanical “pigeon”





Examples of questions that are examined by different branches of philosophy

- What is knowledge? (this is studied by the area of study in philosophy called **Epistemology**).
- What constitutes valid reasoning? (this is discussed in **Logic**) –
- What is time? Is there causation? (**Metaphysics** – nature of reality that is not physical)
- What makes an action morally right? (this is discussed in **Ethics**)
- There's also Philosophy of Cognitive Science, Philosophy of Science, Aesthetics,, ...

Ethics

- **meta-ethics** discusses the meaning of terms and statements (e.g. What do we mean when we say that an action is "wrong"? Are we just expressing a subjective feeling (e.g. this action disgusts me – "yuck"), or are we perhaps making a more objective statement?)
- **normative ethics** explores what moral norms one should follow (e.g. Is it wrong to lie? In some cases or all cases? Why?)
- **applied ethics** examines what one ought to do in particular cases (e.g. Is there a moral duty to protect Hannah's private data? If so, is this a duty to her because she's a friend, or to all humans?).





Philosophy is not giving your personal opinion

- Philosophy is much less subjective than what you may think.
- Think of how arguments are presented in a court of law
- The aim of the argumentative essays you will write is to persuade the reader that your thesis (e.g. a statement about what is moral) is correct by using rational arguments.
- You will be evaluated on the **strength of those arguments** and **knowledge of the theories and concepts** you use to construct them. **The arguments should be your own (not just a repetition of what was said in class).**
- The point is *not* to express your personal opinions or share your feelings (e.g. “I ~~believe that~~ X”, “I ~~don’t feel~~ that’s right”).
- Note: See my notes on essay writing in Content > Orientation

Discussing ethics

Common sense vs ethical theory

- We could discuss resume fraud now - by using common sense ethics
- But... arguments strictly from common sense/ based on intuition rely on shared values and are often poorly justified or not justified at all. Intuitions may be wrong.
- In this course we aim to add ethical theories to our portfolio of tools
- Intuition remains important, e.g. to evaluate the success of theories – Is a recommended action and its justification intuitive or not?
- Example: utilitarianism could possibly support hacking the computers of military Russian hospitals to make them fail, helping Russia lose to end the war quicker. But is this right?
- **Using only or primarily intuition in your essays will produce weak arguments** and will not demonstrate knowledge of the theories and concepts – see Content>Orientation>Essay Grading Criteria

AI-generated



The class today

Some basics about moral reasoning,
including validity and soundness of
arguments

**An essential question: Is there genuine
moral action?**

Example: Most people lie on their resumes.
Digital badges can help this. Would you lie
if you knew you can't be caught? (e.g. if
you could perfectly fake a digital badge)

Moral reasoning – three basics

- 1. There is “**no ought without can**”: there can be no moral requirement to do something unless this is possible
- 2. “**No ought from is**” – Moral value (what ought to be done) cannot ever be established from empirical facts (what is) *alone*. Consequence: every moral argument must build on a mutual acceptance of some moral value.
- 3. Psychological **explanation** is not the same as moral **justification** (e.g. “I drove over your foot *because* I thought that would be funny”). A psychological/motivational account of the action is distinct from moral justification.
- In other words, explanations provide psychological/motivational **reasons** for action (e.g. “I stabbed him *because I was in the mood for murder*”), but not necessarily good/**moral** ones.



“Who wants to be rescued first?”

Another distinction useful for moral reasoning

Intrinsic vs instrumental value

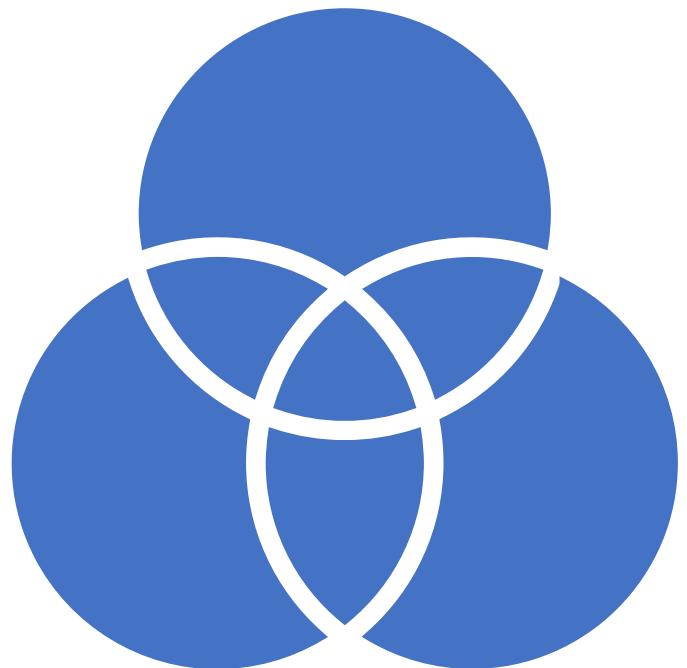
1. Provide an example of something that has **instrumental value**. Can it also have intrinsic value?



2. Provide an example of something that has **intrinsic value**. What does it mean to say that something is intrinsically right/wrong?



3. Can something have both intrinsic and instrumental value?

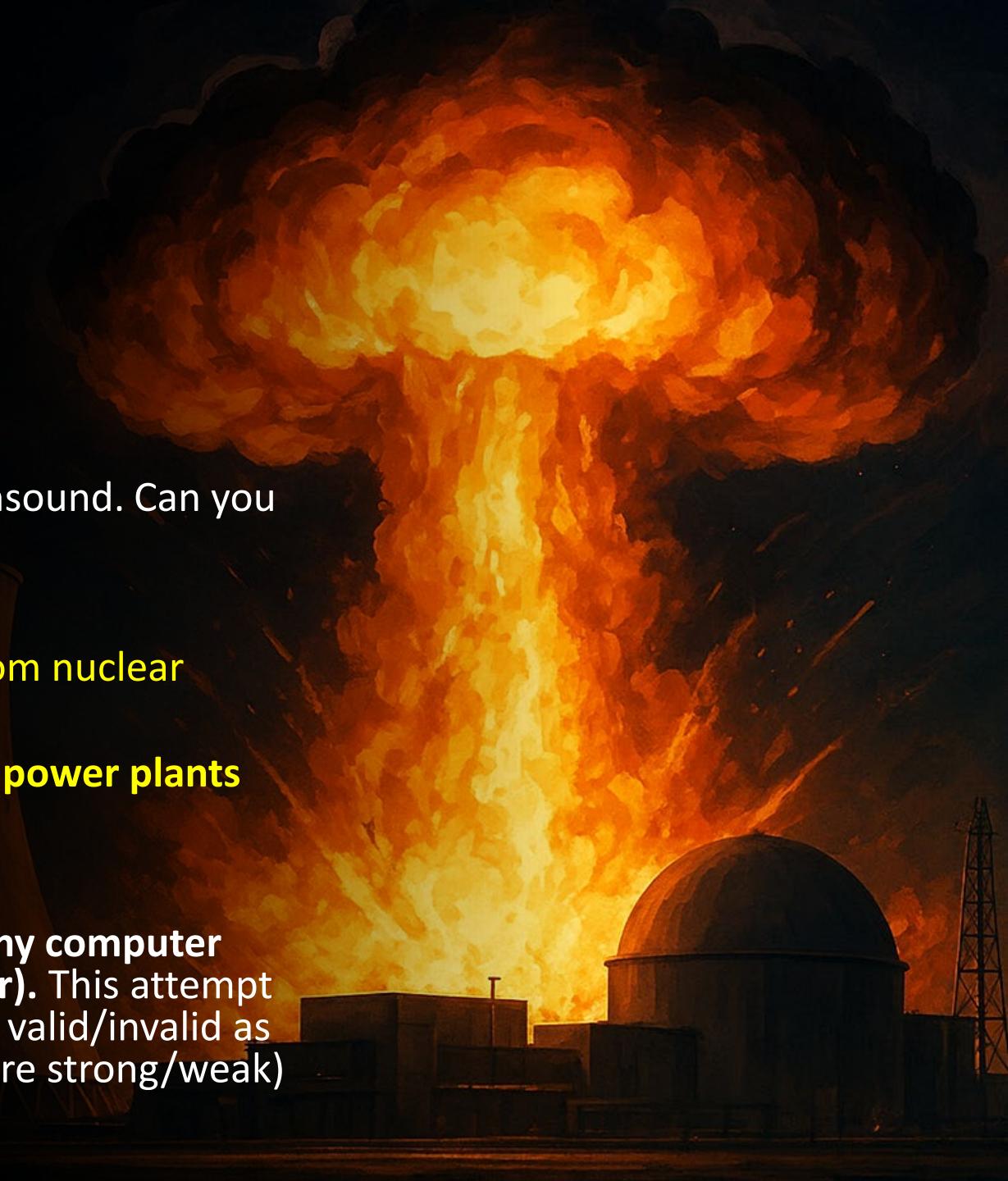


The structure of an argument

- **Arguments** are persuasive *structures* containing *statements* with truth values: one or more premises and a conclusion
- **Premises** are presented as reasons to believe the conclusion.
- The structure itself (the argument) does not have a truth value as it's not a statement
- There are no true or false arguments

Deductive and inductive arguments

- **Deductive** arguments can be valid/invalid and sound/unsound. Can you guess whether this a **valid** argument?
- P1 Lockbit is an example of ransomware
- P2 Using ransomware is a fun way to obtain ransoms from nuclear power plants and hospitals
- **C: Lockbit is a fun way to obtain ransoms from nuclear power plants and hospitals**
- **Inductive argument:** This anti-virus has been keeping my computer safe for years, so it's *likely* I will not get a virus this year). This attempt to generalize is called an inductive leap. They cannot be valid/invalid as there is never a guarantee the conclusion is true (they are strong/weak)



Validity of (deductive) arguments

- Three ways of defining a valid argument:
- 1. An argument is **valid** if the conclusion follows from the premises
- 2. In other words, the truth of the premises guarantees the truth of the conclusion.
- 3. When an argument is valid, we say that it is truth-preserving (if the premises are true, the conclusion *must* be true)
- An argument can be valid regardless of whether the premises or conclusion are true.

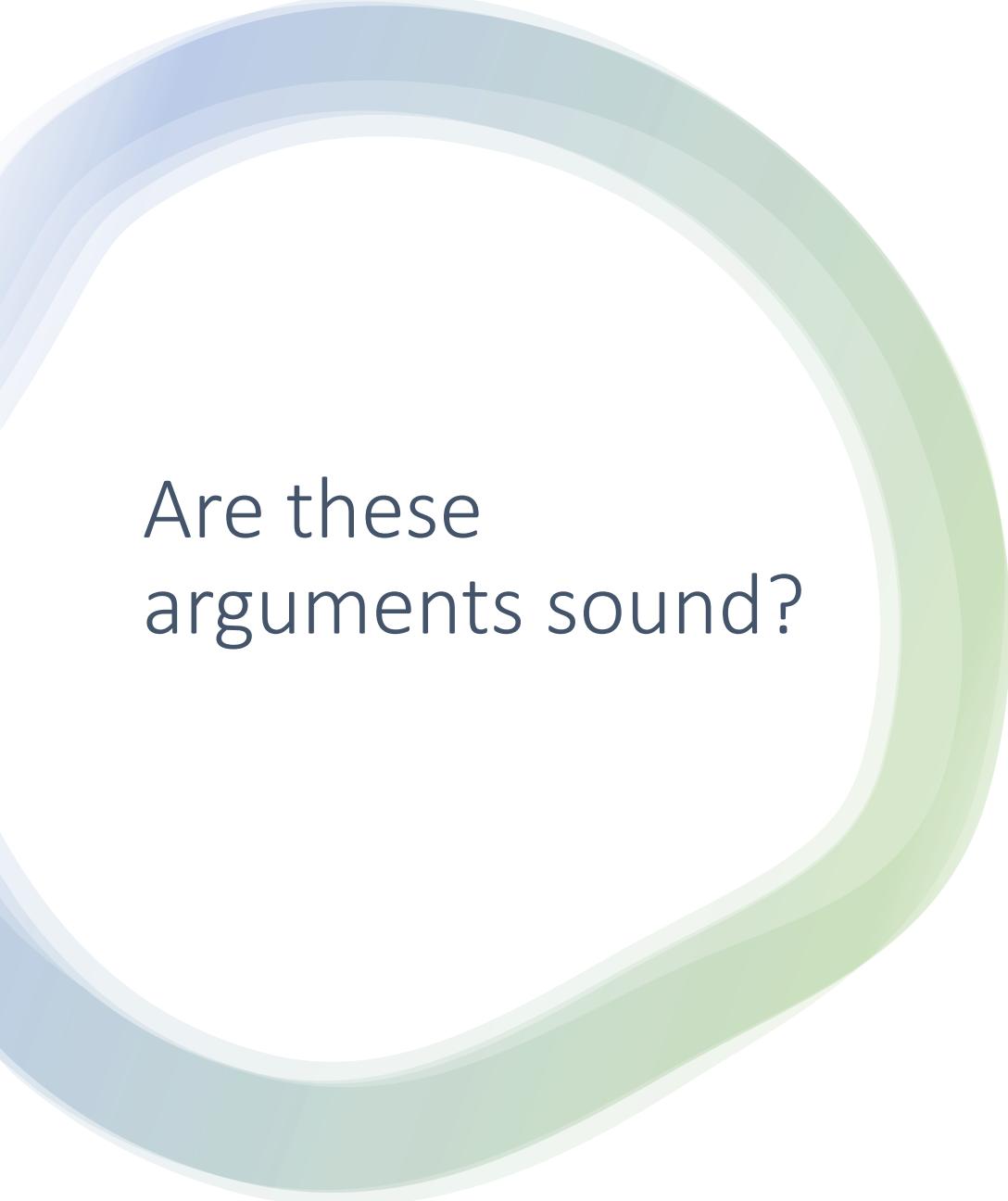


Is this argument valid?

- P1 - All computers are intelligent
 - P2 - Intelligence makes you miserable
 - C - Therefore, all computers are miserable.
-
- Yes, the argument is valid. Look at the structure:
 - P1 - All X are Y.
 - P2 - All Y are Z
 - C - Therefore all X are Z.

The soundness of arguments

- Arguments can then said to be sound **if and only if** these two conditions are satisfied:
 - **A)** the argument is valid
 - **B)** all the premises are true
- Q: Why do we say “if and only if”?
 - A: Because these two conditions are not just **necessary but also sufficient** for an argument to be sound. Nothing else is required.
- Validity, for instance, is a necessary but not sufficient condition for soundness



Are these arguments sound?

- P1 - All computing students either already know or can learn how to hack a computer
- P2 – BCIT has many computing students
- C - Therefore, BCIT is nothing but a training camp for hackers

- The premises may be true, but the above argument is invalid (the conclusion **doesn't follow** from the premises). Therefore, it is unsound.

- P1 – Michelle is a coder from Germany
- P2 – All good coders are from Germany
- C – Michelle is a good coder

- Invalid argument and one false premise - therefore unsound

Summary – validity and soundness of arguments

- So now we know that *deductive* arguments can be:
 - Valid/invalid
 - Sound/Unsound
 - They can't be anything else!
-
- Note that only the premises and the conclusion, being claims, are truth-apt (i.e. we can say of them that they are true or false)
 - Arguments are not claims, so they are not truth-apt. They cannot ever be said to be true or false.

True or false?

- A valid argument can have false premises.
- A valid argument can have a false conclusion.
- A deductively valid argument where all premises are true must have a true conclusion.
- In a deductively valid argument, if the conclusion is true so too must be the premises.
- All sound arguments have nothing but true premises and are valid.

Are these arguments valid?

- 1.1. OpenAI's ChatGPT conversation bot is conscious (has experience/feelings)
1.2. Every computer that can run ChatGPT is conscious. My laptop is running ChatGPT.

1.3. My laptop is conscious.

- 2.1 Musk is God and Tesla can drive by itself and its advanced visual system makes all crashes impossible
2.2 My car is a Tesla and I have a tattoo of Elon's face

2.3 My car can drive by itself and it will never crash

- 3.1 BCIT is either in BC or in Kursk
3.2 BCIT is not in BC

3.3 BCIT is in Kursk

Are these arguments valid?

- 4.1 Peter Dinklage is taller than Shaquille O'Neil
4.2. Shaquille O'Neill is taller than Emilia Clarke

4.3 Peter Dinklage is taller than Emilia Clarke
- 5.1 All software developers apply for well-paying jobs.
5.2. Jane is a person who has a developer job she applied for.

5.3 Jane has a well-paying job.
- 6.1 Consumer-level drones can potentially be flown into the flight path of a commercial airplane
6.2 A plane can crash if it collides with a drone and a turbine gets damaged

6.3 Drones are dangerous and shouldn't be flown.

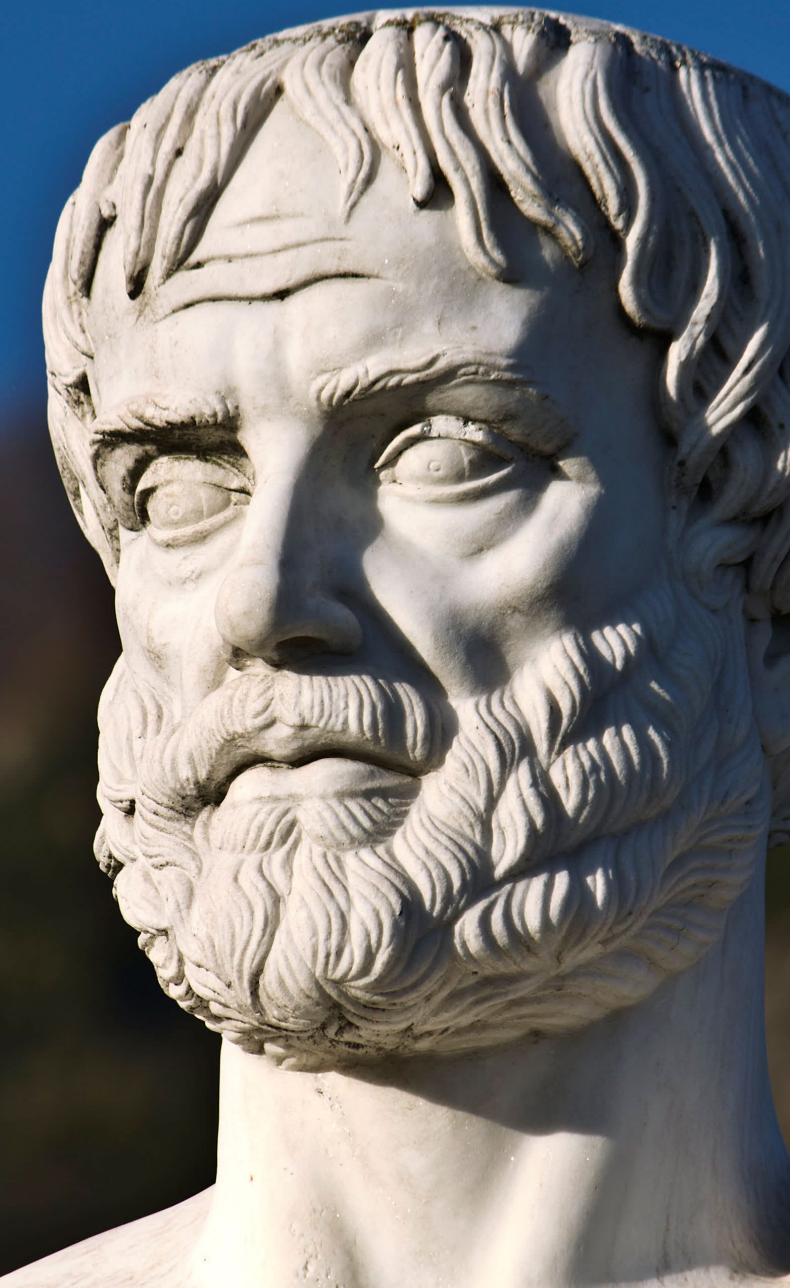


Types of claims.
First, distinguish
value claims from
empirical claims

- “With quantum computers, we can code more information into a smaller computer”
- “What’s exciting about quantum computers is that we can code more information into a smaller computer”
- “Justin Trudeau said that what’s exciting about quantum computing is that we can code more information into a smaller computer”
- Value claim – it claims that something has value (e.g. it’s good, it’s bad, should be done, should not be done) – in a way that cannot be settled by using sensory tests alone (*exclusively* evaluating empirical information)
- Empirical claim – If we can use sensory tests alone to determine the truth of the claim

Empirical claims vs. value claims vs conceptual claims

- Empirical claims - Can we use sensory tests alone to determine the truth of the claim?
 - Value claims – Is there a judgment of approval/disapproval that cannot be reduced to an empirical claim? Moral and aesthetic claims are examples of types of value claims.
 - Conceptual claims - Can we use the meaning of the words alone to determine the truth of the claim?
-
- This function is well designed and so are you.
 - Hitler actually made awesome soup. He was the first soup Nazi.
 - Jeffrey Dahmer is a psychopath and so are you.
 - Those pants make you look really fat
 - There is a UFO parked outside.
 - I have noticed that methane emissions into the atmosphere have increased since you walked in
 - Vitamin D increases one's resistance to Covid making vaccination completely unnecessary. Drinking bleach does the same.
 - There are no four-sided triangles in Manitoba.
 - Mammals are animals.
 - Some brothers have no siblings.



Discussing ethics - Value arguments

- According to Aristotle, this is what a value argument looks like:
- Value premise: All humans **should** exercise
- Empirical premise: **I am** a human being
- Conclusion: **I should** exercise
- A value argument contains **at least one value claim** (often associated with use of words “should/ought”), one or more empirical claims, and a value conclusion.
- In other words: it is **not possible** to get to a value conclusion by using only empirical statements – “no ought from is”
- Why is this important? Don’t be fooled by attempts to get to value conclusions by using empirical statements.

Value argument

- “**no ought from is**” - we cannot justify value statements by using empirical claims alone
- Suppose that someone says:
 - P1 Lies hurt competitors (empirically testable)
 - P2 Your resume includes lies
 - C: It’s wrong to lie on your resume
- The argument is incomplete and is therefore invalid...
- Lies do hurt the competitors, sure, but that’s just an *empirical fact*. It wasn’t established that hurting people is wrong. So the consequence doesn’t **follow**.
- A complete (**valid**) argument is therefore:
 - P1 Lies hurt people (empirical)
 - P2 There is no moral justification for hurting competitors to get ahead unfairly (value claim)
 - C: Lying is wrong in this case (borrowed value from P2)



Plato's *Republic* - Gyges' Ring

- Many have argued that genuine moral action is actually an illusion. They claim we act in ways that *look* moral because we're trying to get something from it (instrumental value), maybe:
 - 1. money
 - 2. reciprocal favours or avoiding punishment
 - 3. the admiration of others
 - 4. pleasure of believing in the illusion we're pure kind people
- According to Glaucon, someone who had a magic ring that allows us to disappear and reappear at will would never act according to moral rules. He uses the story as a thought experiment to support his view that we are “just” “of necessity” (he means: we are moral because we have to).
- **Is there genuine moral action?**



Activity - discuss

- 1. Is there genuine moral action? – Consider whether *everyone* would lie on their resume to get a job if they knew they could never be caught? Or lie, steal, maim, kill to get ahead?
- (From the reading, suppose you can produce a perfect digital badge to support your fake resume and hack the database of the company producing the badges so it can be checked)
- 2. Are there things you wouldn't do to get ahead even if you could never get caught?
- 3. If you still do what is right at least sometimes, are you doing it just to *feel good* about yourself?

Is there genuine moral action?

- Whatever “genuine moral action” might mean (there is no definition everyone agrees on, but we can try to come close to a definition, and eliminate a few poor ones...), we need to clarify these concepts in order to come up with our own conception:
- Selfishness vs Self-interest
- Psychological Egoism and why we should reject it

Are we all selfish? – the theory of psychological egoism

- “[Psychological] Egoism maintains that when we care about what happens to others, we do so only as a means to increasing our own welfare [...] In all our social interactions, we are driven by a single question – ‘What’s in it for me?’” Sober & Wilson, *Unto Others* (1998: 2).
- **Psychological** egoism: the theory/view that **everyone** is self-interested at all times (an empirical claim)
- According to the theory of psychological egoism, our only ultimate human motive is always self-interest (i.e. we either do something for the sake of self-interest directly, or value actions that instrumentally serve our self-interest, so ultimately all is done for the sake of self-interest)
- **Normative** egoism: The theory/view that **everyone ought to** always act out of self-interest (a value/normative claim)



Hobbes famously subscribed to psychological egoism

- Hobbes puts it like this: “[O]f the voluntary acts of every man, the [ultimate] object is some good to himself” (1660: XIV). Because of this, Hobbes reduces every motivation that appears to be moral to self-interest.
- A good example is Hobbes’s account of pity: “Grief for the calamity of another is pity; and ariseth from the imagination that the like calamity may befall himself” (1660: VI).

A counter-argument by Hume

- Hume: Yes, *sometimes* we do deceive ourselves and tell ourselves we are nice when we are being self-interested or selfish. But always? No!
- “A man, who has lost a friend and patron, may flatter himself, that all his grief arises from generous sentiments, without any mixture of narrow or interested considerations
- But a man, that grieves for a valuable friend, who needed his patronage and protection; how can we suppose that his passionate tenderness arises from some metaphysical regards to a self-interest, which has no foundation or reality? [...]
- What interest can a fond mother have in view, who loses her health by assiduous attendance on her sick child, and afterwards languishes and dies of grief, when freed, by his death, from the slavery of that attendance?”



We can find (a degree of) benevolence in every human being

- Benevolence is a natural tendency of human beings, says Hume
- Hume doesn't mean we always have it to the right degree and in the right circumstances. He means we all have *some* degree of benevolence (even the most vicious among us), and it's not distributed equally (mostly directed to those close to us). It's a common moral mistake to overly favour those we love.
- **We often thus fail to be moral not by being selfish, but by directing our benevolence incorrectly**
- **So the dichotomy moral/selfish is a false dichotomy.**
- **It's false or at least inaccurate to say, without qualification, that everyone is selfish** (everyone is selfish sometimes, but everyone is also benevolent at least sometimes)



Another counter-argument to psychological egoism

- Someone may find the following argument irrefutable: ‘I do only what I desire. That is, I always act so as to satisfy *my* desires. It is clear, then, that all my motivations are selfish.’
- Williams: The stated premises are (trivially) right, but the conclusion is wrong. To desire is not the same as to be selfish. To be selfish is rather to have selfish desires.
- Williams: What makes someone an egoist is not that this is a person who has desires, or that the desires are theirs. What matters is *what they desire!* – things for themselves only or also things for others?





To understand Williams's point, we must distinguish selfishness from self-interest

- Self-interest is common, but not all self-interest is **selfish**
- Everyone cares for their self-interest.
- You eat and drink
- You are not selfish if you tie your shoes to avoid tripping, or sleep because you're tired
- But what if I, not being particularly thirsty, have two glasses of water, when someone next to me is dying of thirst?
- Selfishness is then *excessive* concern with one's self-interest.
- Not all self-interested actions are selfish

Conclusion

- When humans fail to be moral, it's often out of (a flawed sense of) loyalty/love to others or ourselves that we consider appropriate/moral. We mean well
- It's inaccurate to say that everyone is selfish
- Genuine moral action does exist



Is there “genuine moral action”?

- ***Do the readings and read the course notes (lec 2 folder) before lecture 2. See you next week!***
- **Summary**
- If Hume is right, we are all benevolent, at least sometimes, with some people. So at least some of the time we are not acting exclusively to benefit ourselves.
- If Hume is right, often when we fail to act morally, we are not actually acting selfishly (either/or fallacy – it's a false dichotomy to say we must be either moral or selfish). We may be acting benevolently when we should have been fair, for instance (e.g. nepotism)
- If Bernard Williams is right, it's OK to *always* act on one's desires (of course we always do), so long as they are not all directed towards benefitting oneself
- It's not necessarily selfish to take care of oneself. We must be entitled to take care of ourselves at least to some degree (or we would die, need someone's assistance, etc), so it's clearly OK to care for one's self-interest, so long as we don't do that *exclusively or in excess*