

# Phil. 2750

## Notes #1: Course Requirements, Introduction

To discuss today:

- This course: requirements, subject, guidelines
- What is philosophy?
- What does science fiction have to do with philosophy?
- What should you learn from this class?

### I. About this Class

Review syllabus. Some highlights:

- Who should take this class?
  - Class will contain:
    - Lots of arguments
    - Theoretical, abstract questions
    - Controversial ideas
  - Will not contain:
    - Directly practical knowledge
    - Touchy-feely, self-esteem-boosting material
- Course requirements. Quizzes, paper, participation
- Miscellaneous guidelines:
  - Come on time.
  - Come to office hours.
  - Participate.
- What should you do now?
  - Get books & reserve readings <<http://libraries.colorado.edu/search/p?SEARCH=huemer>>.
  - Read the syllabus.
  - Read the Pollock selection.

### II. What Is Philosophy?

#### A. The Subject Matter of Philosophy

- Philosophy studies some general, fundamental questions, about the nature of the world and our place in it.
- Three main branches:
  1. Metaphysics - Studies what sorts of things in general exist, and what sort of world this is. Examples: existence of God, free will vs. determinism, distinction between body and soul, the Ship of Theseus question.
  2. Epistemology - Studies the nature of knowledge - what is it and how do we know what we know?
  3. Ethics - Studies evaluative questions - what is good, bad, right, and wrong.
- Some smaller branches of philosophy:

4. Political philosophy - Studies the source of political authority, the best overall structure for society and the state, and related questions. (Can be seen as a branch of ethics.)
5. Aesthetics - Studies the nature of art, beauty, and related questions. (More generally: the nature of aesthetic qualities.)
6. Logic - Studies reasoning, esp. the principles of correct reasoning. Closely related to, but not the same as, epistemology.

#### B. The Methods of Philosophy

Philosophy in the Western tradition mainly relies on logical arguments & common experience.

#### C. An Example of Philosophy

The Ship of Theseus:

Theseus sailed around the Mediterranean for 10 years. During this time, he periodically had to replace pieces of the ship due to wear and tear. After 10 years, every plank in the ship had been replaced, one at a time. Q: Was it still the same ship?

Notice things about this question:

- A priori, not empirical.
- Far-reaching implications (identity of composite objects over time).
- Puzzling. Compelling arguments for incompatible positions.

### III. Science Fiction & Philosophy

- Science fiction is
  - a) Imaginative, yet
  - b) Realistic. (See how both of these are the case.)
- Provides material for thought experiments.
  - Exploring concepts through interesting cases on the borderline. Ex.:  
The concept of persons and: robots, alien species, symbiotic species, person-splitting, machine/organic hybrids.  
See Asimov's robots, Data, the Trill (from ST), "Second Chances" (ST:TNG), Borg
  - Exploring human nature through giving people special abilities. Ex.:  
Ability to rewrite the past, immortality, godlike powers over others.  
See ST:TOS, "Where No Man Has Gone Before"; Q from ST:TNG; various time travel stories.
- Some SF possibilities may actually be realized.
  - Androids, human/machine hybrids (like the Borg), nanotechnology (as in The Diamond Age), sophisticated virtual reality (Snow Crash), genetic engineering (Gattaca), colonies on the moon (The Moon Is a Harsh Mistress), radical restructuring of society (Diamond Age; Brave New World).
  - Only nerds will be ready.

### IV. What I Hope You Learn from this Class

#### A. Some Important Philosophical Ideas

- About free will, consciousness, personhood, time, knowledge, human nature, and the structure

of society.

## B. Thinking Skills

- Philosophy teaches us to think more clearly, to avoid common confusions.
- Philosophy teaches us to reason more cogently, to avoid common fallacies.
- Philosophy makes us aware of the fundamental questions.

## C. Philosophical Attitude

The Cardinal Rule of Philosophy: Truth comes first.

When doing philosophy, we are trying to identify what is true. That comes before personalities, feelings, and desires. The following rules are all consequences of this.

Four Principles of Philosophical Comportment:

### 1. Philosophers question:

- Question the claims of others.
- Question your own beliefs.
- This does not mean refusing to accept anything as true!

### 2. Philosophy is impersonal:

- The philosopher does not choose beliefs based on his personality or feelings.
- The philosopher does not take criticism of ideas personally.
- The philosopher does not accept or reject philosophical claims based on who says them.
- The philosopher does not go along with ideas because of personal or social consequences of criticizing them.

### 3. Philosophers are guided by reason:

- The philosopher has reasons for his beliefs.
- The philosopher asks for the reasons for others' beliefs.
- The philosopher is moved by good reasons presented to him.

### 4. Philosophers are open-minded and critical:

- Our ideas and arguments are open to criticism. The philosopher looks for objections to his beliefs.
- The ideas and arguments of others are also open to criticism.

## Phil. 2750

### Notes #2: Cartesian Skepticism

#### I. What Is Cartesian Skepticism?

Skepticism: Roughly, any philosophical view according to which some large class of things we normally believe (a) we do not know, or (b) we are not justified in believing.

Varieties of skepticism:

- External World Skepticism: We cannot have knowledge/justified belief about any contingent truths about the external world.
  - The external world: That which is independent of one's own mind.
  - Contingent truths: Things that could (conceivably) have been otherwise.
- Global Skepticism: We cannot have knowledge/justified belief about anything at all.
- In this unit, we consider external world skepticism regarding justified belief. This is the view that we have no justification for (contingent) claims about the external world.

Cartesian Skepticism:

- This is named after René Descartes. (See: Meditations on First Philosophy.)
- “Cartesian” skeptical arguments involve “skeptical scenarios”:
  - a) Scenario in which everything appears as it actually does, but
  - b) Your beliefs are radically mistaken.
- Examples:
  - The dream scenario
  - The deceiving God
  - The brain in a vat
  - The Matrix, etc.

#### II. The Brain in a Vat argument

1. If you know that P and P entails Q, then you can know that Q. (Premise: Closure Principle for knowledge.)
2. You can't know you're not a BIV. Argument for this:
  - a. Our sensory experiences are the only evidence we can have for claims about the external world. (Premise.)
  - b. If you were a BIV, you would have the same sort of sensory experiences as you actually have. (Premise.)
  - c. Your experiences are not evidence that you're not a BIV. (From b.)
  - d. You cannot have evidence that you're not a BIV. (From a, c.)
  - e. You can't know you're not a BIV. (From d.)
3. Therefore, you don't know (for example) that you have two hands. (From 1, 2.)  
(Implicit: Having 2 hands entails not being a BIV.)

In the following classes, think about which premise (if any) is denied by each response to skepticism.

### III. Moorean Responses to Skepticism

#### Moore's Proof of an External World

We all know there are external objects. But is it possible to prove this? Moore says it's easy:

1. Here is one hand. (Gesture.)
2. Here is another hand. (Gesture.)
3. Therefore, there are external objects.

#### Moore & Pollock's general response to skeptical arguments

Skeptic's argument:

1. If you know that P and P entails Q, then you can know that Q.
2. You can't know you're not a BIV.
3. Therefore, you don't know (for example) that you have two hands.

Moorean argument:

1. If you know that P and P entails Q, then you can know that Q.
2. I know I have two hands.
3. So, I can know I'm not a BIV.

Schematically:

Skeptic:

$P_1$

$P_2$

---

$\therefore \neg K$

Moore:

$P_1$

K

---

$\therefore \neg P_2$

Which is better?

- Both equally valid.
- Both equally circular or non-circular.
- Which has the more plausible premise?
- Neutral presentation:
  - Three propositions are jointly incompatible:  $P_1$ ,  $P_2$ , and K. At least one must be rejected.
  - Each initially seems true.
  - Solution: reject the one we are least confident of.
- Moore & Pollock's point generalizes to all skeptical arguments. Lesson:  
In a persuasive argument, each premise must be more initially plausible than the negation of the conclusion.

## Phil. 2750

### Notes #3: Are You Living in a Computer Simulation?

#### I. The Simulation Argument

An argument that we are probably living in a computer simulation

1. Functionalist theory of the mind. A sufficiently detailed simulation would be conscious.
2. An intelligent species has a good chance of reaching a “posthuman” stage.
3. There is a good chance that posthuman intelligent beings would run computer simulations.
4. If they did, there would be many more simulated beings than “real” beings.
5. The probability of being a simulated being = the fraction of all beings who are simulated beings. (The Pr. of being a real person = the fraction of all beings who are real people.)
6. Thus, there is a good chance that you are a simulated being.

Bostrom says: One of the following should be accepted:

- a) that our species will probably go extinct before it reaches a much more advanced state;
  - b) that almost none of the highly advanced beings who could do so would choose to run sophisticated computer simulations; or
  - c) that we're probably computer-simulated beings.
- This is the official thesis of the paper. He's not asserting (c).
  - He is definitely asserting premises (1), (4), and (5).

#### II. Which of the Three Possibilities Is Most Likely?

Why think (a)?

- Posthuman stage: A state in which people are extremely technologically advanced. So-called because these beings may be beyond human, e.g., genetically engineered, human/machine hybrids, etc.
- If technological progress continues, this stage will occur, perhaps in a few decades or a few centuries.
- If intelligent species in general are unlikely to reach a posthuman stage, then our species will probably not reach it.
- Why might we fail to become posthuman?
  - Species extinction—war, diseases, other dangerous technology.
  - Collapse of technological civilization.

Why think (b)?

- Maybe simulations would be considered unethical.
  - Too much suffering?
  - Is it wrong to create us??
- Maybe posthumans would have entirely different interests.

Why think (c)?

- Posthumans would have vast computing power.

- “Moore’s Law”: computing power doubles about every 2 years.
- This has some theoretical limits. But:
  - ☞ A planet-sized computer could do  $10^{42}$  operations per second.
  - ☞ Cost of a realistic simulation of human history:  $10^{33}$ - $10^{36}$  operations.
  - ☞ The planet-sized computer “could simulate the entire mental history of humankind . . . by using less than one millionth of its processing power for one second.”
- Thus, the vast majority of beings with minds like us could be simulated beings.

### III. Bostrom’s Premises

Bostrom is asserting these:

1. Functionalist theory of the mind. A sufficiently detailed simulation would be conscious.
4. If posthumans ran ancestor simulations, then there would be many more simulated beings than “real” beings.
5. The probability of being a simulated being = the fraction of all beings (with experiences like ours) who are simulated.

Why accept (5)?

- Analogy: Suppose 2/3 of the population has genetic condition S. Given no other information, what is the probability that you have S?
- Suppose 1/4 of all babies result from unintended pregnancies. Given just this information, what is the probability that you resulted from an unintended pregnancy?
- What if all beings placed bets on whether they were simulated or not? What would be the fair betting odds?
- If 100% of all beings are simulated, what is the probability that you are simulated?

Why accept (4)?

- Discussed above.

Why accept (1)?

- “Arguments for this thesis have been given in the literature, and although it is not entirely uncontroversial, we shall here take it as a given.” (2)
- “This attenuated version of substrate independence is quite widely accepted.” (2)

## Phil. 2750

### Notes #4: The Matrix as Metaphysics

#### I. Thesis: The Matrix/Simulation Scenario Is Not a Skeptical Scenario.

1. The Matrix Hypothesis “is equivalent to” the Metaphysical Hypothesis.
2. The Metaphysical Hypothesis is not skeptical.
3. So the Matrix Hypothesis is not skeptical.

Matrix: A computer simulation of a world.

The Matrix Hypothesis: We are living in a matrix.

The Metaphysical Hypothesis:

- a) The Creation Hypothesis: Physical spacetime and its contents were created by a being or beings outside spacetime.
- b) The Computational Hypothesis: Microphysical processes are constituted by computational processes.
- c) The Mind-Body Hypothesis: Our minds are outside physical spacetime but interact with processes in physical spacetime.
- d) The computational processes were designed by the creators as a computer simulation of a world.

#### II. The Metaphysical Hypothesis Is Not Skeptical

- a) The Creation Hypothesis is not skeptical.
  - Many people in fact believe it.
- b) The Computational Hypothesis is not skeptical.
  - It is just a theory about what constitutes physical particles.
  - Other theories about what constitutes observable objects are not skeptical.
  - Q: Could reality be fundamentally computational? Maybe not.
- c) The Mind-Body Hypothesis is not skeptical.
  - Many, perhaps most people already believe this (mind-body dualists).
- d) The combination isn't skeptical.

#### III. Objection: Simulation Is Not Reality

1. Matrix Hypothesis says that there is a computer simulation of physical reality.
2. This does not entail that physical objects exist.
3. Metaphysical Hypothesis entails that physical objects exist.
4. So Matrix Hypothesis does not imply Metaphysical Hypothesis.

Reply:

- a. “It is clearly possible that a computational level underlies real physical processes.”
- b. Two characteristics of computational processes: (i) the structure & complexity of the computation, (ii) how it is implemented.

- c. Any adequate computer simulation of reality “will have a rich enough causal structure that it could in principle underlie physics in our world.”
- d. It does not matter how the computation is implemented.
- e. So any adequate computer simulation would constitute a physical reality.

#### IV. Other Objections

Some beliefs of the brain in the vat that would seem false:

- a) “I am outside.”
- b) “I am in Tucson.”
- c) “I have hair.”
- d) “I have friends.”

Why these beliefs are really true:

- “I” (in [a]) refers to the brain’s virtual body.
- “Tucson” refers to Tucson\* (virtual Tucson).
- “hair” refers to hair\* (virtual hair).
- “friends” still refers to friends, but the BIV has actual friends (the other brains).

Chalmers’ Semantic Theory:

- Two kinds of words/concepts (pp. 24-5):
  - Semantically stable (“neutral”) terms and concepts: Retain reference when used by BIV’s and normal humans.
    - ☞ Includes: mental, causal, logical, mathematical, and “categorical” concepts. Also, “friend” and “action”.
  - Semantically unstable terms and concepts: Shift reference when used by BIV’s.
    - ☞ Refer to things they are causally related to.
    - ☞ Includes: “natural kind” concepts, proper names, spatial concepts, the concept “physical”.
    - ☞ Putnam’s example: Twin Earth
    - ☞ Chalmers’ example: Terry and Terry\* (16)
- So when the BIV says “physical object,” it refers to the virtual physical objects in the simulated world. BIV speaks truthfully in saying “there are physical objects around me,” etc.
- Chalmers claims that something along these lines is a consequence of his argument, not presupposed by it (25).

A Possible Objection:

- Maybe “physical”, or spatial terms, are semantically stable.
- Chalmers’ student Brad Thomson has argued against this.
- Related: Searle rejects the Twin Earth-style thought experiments.

## Phil. 2750

### Notes #5: A Direct Realist Response to Skepticism

#### I. Competing Theories of Perception

- Direct realism: Perception gives us direct (non-inferential) knowledge of the physical world.
- Indirect realism: We have inferential knowledge about the physical world, based on the character of our experiences.
- Skepticism: We can't know about the physical world.

#### II. Review of Dretske's and Klein's Responses to Skepticism

- Skeptic says: "BIV Hypothesis is an alternative to Real World Hypothesis. You can't rule out BIVH. So you don't know RWH."
- Dretske: "I can't rule out BIVH, but I still know RWH. BIVH is irrelevant (because not objectively possible)."
- Klein: "Maybe you can defend the closure principle, but then your claim that I can't rule out BIVH begs the question."

#### III. What Is Wrong with these Responses

##### The Courtroom Case

S is on trial for murder. The prosecution offers as evidence the fact that S's blood was found at the scene of the crime along with the victim's blood. They suggest that S got cut while stabbing the victim. The defense offers an alternative hypothesis: S is innocent, and the blood was planted at the crime scene by the police, seeking to frame S.

Dretske: For all we know, the defense is correct, and S was framed by the police. But anyway, we still know S is guilty.

Klein: Either Dretske is right, or the defense attorney's argument just begs the question.

##### The Scientific Case

Physicist A supports the Copenhagen Interpretation of quantum mechanics. Physicist B supports Bohm's Interpretation. Both interpretations explain all the same data, but they are incompatible with each other.

Dretske: I don't know whether Bohm's theory is right. But I know the Copenhagen theory is right.

Klein: Dretske is right, or Physicist B is begging the question.

#### IV. A Reformulation of the Argument & the DR Response

1. If E is some evidence and  $H_1$  and  $H_2$  are competing explanations of E, then S is justified in believing  $H_1$  on the basis of E only if S has independent grounds for rejecting  $H_2$ . (Preference Principle)
2. The BIV Hypothesis and the Real World Hypothesis are competing explanations of our sensory

experience.

3. So in order to believe RWH on the basis of our sensory experience, we must have independent grounds for rejecting BIVH. (From 1, 2.)
4. We have no such grounds.
5. Therefore, we are not justified in believing RWH on the basis of our sensory experience. (From 3, 4.)
6. Facts about sensory experience are the only justification we might have for RWH.
7. So we're not justified in believing RWH. (From 5, 6.)

Notes:

- This argument escapes Klein's and Dretske's responses, as it should.
- But it only refutes indirect realism.
- (6) is false. We might be foundationally justified in believing RWH, or we might be justified in believing RWH on the basis of facts about the physical world.
- (4) may be false. Facts about the physical world (of which we're directly aware) might be grounds for preferring RWH over BIVH.

## V. Objections

(i) Does the DR theory imply that perceptual beliefs are indefeasible?

No. See concept of "prima facie justification": perceptual beliefs are "presumed innocent until proven guilty," i.e., justified as long as there are no positive grounds for doubt.

(ii) Does the DR response 'beg the question'?

Two kinds of responses to skepticism:

- Aggressive response: positive argument, addressed to skeptic, that we have knowledge of the external world.
  - We have not provided one of these.
- Defensive response: response to skeptic's argument that we don't have knowledge of the external world.
  - We have provided one of these.

## Phil. 2750

### Notes #6: The Problem of Other Minds

#### I. Problem

- All you observe of other people is their physical behavior.
- Behavior does not logically entail any mental states.
- Hence, how do you know others are not mindless automata (without consciousness)?

#### II. Some Responses

The Argument by Analogy:

- You observe a correlation between
  - a. Your mental states and your resultant behavior.
  - b. Things that happen to you and your subsequent mental states.
- You observe other creatures who are outwardly similar to you.
- You infer (by 'analogy') that the same correlations probably hold for them. Hence, that they probably have similar mental states in similar circumstances.

Inference to the Best Explanation:

- You observe people's behavior.
- Hypotheses about their mental states would explain their behavior.
- You know of no other explanations.
- Hence, hypotheses about their mental states are the best available explanations for their behavior.
- Hence, you infer that they probably have mental states of various kinds.

Is It an Innate Capacity?

- Some say that we have an innate capacity to "perceive" others' mental states without reasoning. People's mental states are unconsciously revealed, and perceived by others, through their facial expressions, tone of voice, etc.
- Is this a possible account of our knowledge of other minds?

#### III. About Heinlein's Story, "They"

- Character claims to have evidence that most other people are mindless automata. Review: what is this evidence? Is it any good?
- What is his evidence for the conspiracy against him? Is his conclusion reasonable? If so, at what point does it become reasonable?