

Minds and Machines (Lecture 1): General Theories of Mind

Prof. Ioannis Votsis

Philosophy Discipline

ioannis.votsis@nulondon.ac.uk

www.votsis.org

Structure of the course

- **Topics:**

week 1: General Theories of Mind

week 2: Extending the Mind

week 3: Human and Machine Intelligence

week 4: Computationalism

week 5: Science and AI

week 6: Creativity and AI

week 7: ***READING WEEK***

week 8: Emotion and AI

week 9: Superintelligence

week 10: Scepticism and Computer Simulations

week 11: Consolidation Session

week 12: Collections

week 13: ***EASTER BREAK***



Introduction

The mind-body problem

- General (philosophical) theories of mind are theories that seek to understand the nature of the mind.
- More precisely, such theories tend to focus on the puzzling relationship between the mind and body (or brain).

The Mind-Body Problem:

What is the relation between the mind (or mental stuff) and the body (the physical stuff)?

Ontological and interaction questions

- To answer the aforementioned question, we need to answer a number of other questions. These include:

Questions:

- * What are mental states and processes?
- * What are body states and processes?
- * Are the two identical or distinct?

Ontology

} **Constitution**

- * Do the mental and physical interact?
- * If so, how exactly?

} **Interaction**

Monism vs. dualism

- There are two families of theories that give answers to the aforementioned questions.
- **Monism:** Only one of the two kinds (either mind stuff *or else* body stuff) exists.

Clearly, if one is a monist, then the interaction questions disappear (or get replaced with 'explaining away' questions).

- **Dualism:** Both mind and body stuff exist.

Answers to the interaction questions vary in accordance with the specific dualism espoused.



Dualism

Background

- The view that mental stuff is distinct from physical stuff and at least ontologically on par with it goes back to the ancients.
- Aristotle, for example, held the view that the intellect (and more broadly the soul) is immaterial.
- Dualism is also clearly associated with religious/spiritual views where the spirit or soul transcends the physical world.

NB: Some religious/spiritual views are in fact monist in that they deny the existence or reality of the physical realm.
- Hereafter, we restrict our discussion to the philosophical varieties of dualism.

Dualism: Constituents

- One useful distinction in getting to grips with dualism is that between the ontological categories we can be dualists about.

Substance Dualism: There exist two kinds of substances, namely mental and physical ones.

Property Dualism: There exist two kinds of properties, namely mental and physical ones.

- Those who endorse property dualism claim that there is only one kind of substance, i.e. physical substance.
- They then argue that physical substances have both physical and non-physical, i.e. mental, properties.

Dualism: Interactions

- Another useful distinction is that between the kinds of interactions that exist between the mental and the physical.

Interactionists: Minds and bodies are both causes and effects to each other, e.g. Descartes.

Epiphenomenalists: Minds are causally affected by bodies but not vice-versa, e.g. T.H. Huxley.

Occasionalists (Parallelists): Minds and bodies are not causally related but they are correlated, e.g. Malebranche.

Jackson's Mary in the black-and-white room

- Jackson (1982; 1986), an epiphenomenalist, presents a highly influential argument against physicalism.
- Physicalism, roughly, is the view that that only physical things exist and facts about them can be fully captured by physics.

NB: Where physics here means the *final* theory of physics.

- The experiment is meant to demonstrate the existence and knowability of non-physical categories like qualia.
- What is a quale (the singular form of 'qualia')? It is a quality of what it feels like to experience something.

Jackson's Mary: The thought experiment

- “Mary is confined to a black-and-white room... she learns everything there is to know about the physical nature of the world. She knows all the physical facts about us and our environment, in a wide sense of ‘physical’ which includes everything in *completed* physics, chemistry, and neurophysiology, and all there is to know about the causal and relational facts consequent upon all this... If physicalism is true, she knows all there is to know... It seems, however, that Mary does not know all there is to know. For when she is let out of the black-and-white room or given a color television, she will learn what it is like to see something red, say. This is rightly described as learning – she will not say ‘ho, hum’. Hence, physicalism is false” (Jackson 1986, p. 291) [original emphasis].



Monism

Mental vs physical monism

- **Mental monism:** Only mind(s) exist(s).

A number of variants have been posited including:

- > idealism / certain types of phenomenalism
- > solipsism

- **Physical monism:** Only body/bodies exist(s).

Again, a number of variants including:

- > identity theories (type and token)
- > functionalism
- > behaviourism

Mental vs physical monism

- **Mental monism:** Only mind(s) exist(s).

A number of variants have been posited including:

- > idealism / certain types of phenomenalism
- > solipsism

- **Physical monism:** Only body/bodies exist(s).

Again, a number of variants including:

- > identity theories (type and token)
- > functionalism*
- > behaviourism

Our focus



Identity Theories

Identifying the mind with the brain

- Roughly speaking, this is the view that mental states and processes are identical to brain states and processes.

NB: This is obviously a strong assertion when compared to the assertion that the two are correlated/causally related.

- Such identities may be conceded without conceding that the (whole of the) mind is identical to the brain.

NB: Polger (2009) holds that such theories typically focus on sensations and conscious mental states.

- The view has been characterised as ‘chauvinism’ as it insists that *biological* brains are required for the mental.

An IBE argument for the identity theory

- The identity theory offers the best explanation for:

“the correlations observed between mental and brain states, the success of psychological and neuropsychological theories, the success of neuroanatomy or cognitive neuroscience, etc.” (Polger 2009: 829).

where the best explanation is the most simple, unified, empirically adequate, fruitful, etc.

Types vs. tokens

- To understand type and token identity, we need to understand the distinction between types and tokens.
- Types are, roughly speaking, classes of things and tokens as the elements of those classes. Here are some examples:

Type

Homo sapiens

Chairs

Numerals

Pains

Books

...

Token

Ioannis, Alice, Christoph, ...

chair a, chair b, chair c, ...

1, 2, 3, ...

muscle pain, joint pain, ...

Brainstorms, Superintelligence, ...

...

Types vs. tokens: Two qualifications

(1) The term 'type' is often reserved for classes that have some internal uniformity. Intuitively:

	<u>Class/Set</u>	<u>Type/Kind</u>
the things in this room	Yes	No
human beings	Yes	Yes

(2) Tokens may (but need not) themselves be types. Any attribution of type-hood or token-hood is thus relativized.

Type:

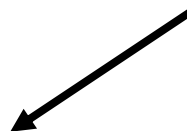
Books

Superintelligence

Token:

Superintelligence

my copy of Superintelligence



Type identity

- **Type identity theory:** Types of mental things are identical to types of brain things.

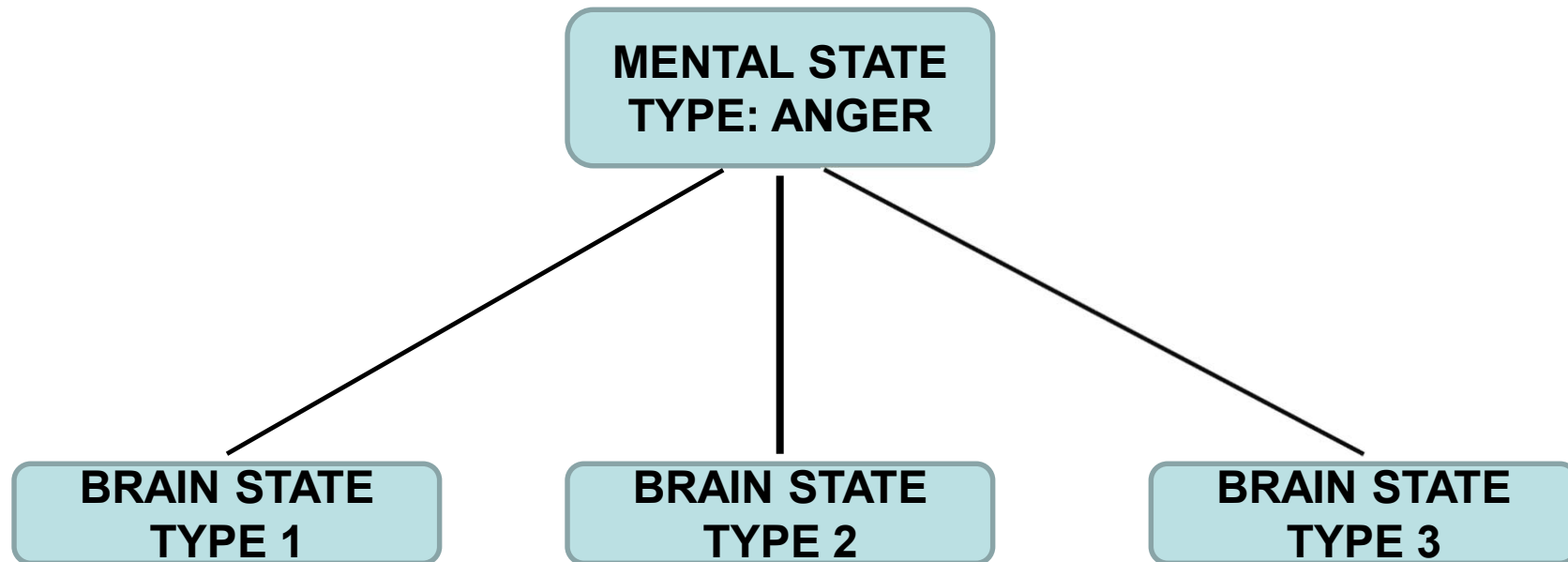
where these 'things' may include one or more of the following: events, processes, properties and states.

Example:

- Mental Type: Pain
 - Brain Type: Nerve fibres (e.g. C fibres) firing in the brain.
- Prominent advocates: D. Armstrong, D. Lewis, H. Feigl, U.T. Place and J.J.C. Smart.

Multiple realizability: A knockdown argument?

- Mental state types cannot be identified with brain state types for the former *can be had* by neurologically diverse creatures.



- Thus, according to multiple realizability (MR), the relationship is not one-to-one but rather one-to-many.

NB: MR originates in Fodor (1965) and Putnam (1967).

Counters to multiple realizability

- Identity theorists have given the following kinds of replies to the multiple realizability objection:

(1) The objection relies on a mismatch of types; once removed, the 1-1 correspondence gets re-established:

(a) apply a more fine-grained conception of mental types

- Anger in humans is identical to brain state type *X*.
- Anger in dogs is identical to brain state type *Y*.

(b) apply a more coarse-grained conception of mental types

- Anger is identical to *any* brain state *satisfying constraints C*.

(2) Doubts over the extent of neurological diversity, e.g. uniformity due to common descent or convergent evolution.

Token identity

- **Token identity theory:** Tokens of mental things are identical to tokens of brain things.

where these 'things' may include one or more of the following: events, processes, properties and states.

Example:

- Mental Token: Ann's headache right now.
 - Brain Token: Specific nerve fibres firing in her brain now.
- Prominent advocates: L.F. Barrett, J. Fodor, W. James and C. Taylor.

Anomalous monism

- Psychological behaviour cannot be explained by physics (anomalous) but thoughts and actions are physical (monism).

NB: The view is the brainchild of Davidson (1970).

- The anomalousness presupposes mental causation and the monism assumes a token identity theory about events.
- The rationale behind the anomalousness of psychology is the belief that, unlike physics, it cannot yield exception-less laws.
- Objection: Events need no more than one sufficient cause; hence no mental causation (the causal exclusion problem).

Zombies

- Chalmers (1996) argues that the identity theory could not be correct via a conceivability argument.

NB: This argument targets physicalism more generally.

- It is conceivable that there are creatures like us but without consciousness. In other words, zombies of a kind.
- If that's the case, then the brain types/tokens would be instantiated w/out the instantiation of mental types/tokens.
- Reply: The plausibility of identity theories needs to be determined empirically, not via conceivability arguments.



Functionalism

Role-based identification

- **Functionalism:** Mental states and processes are identifiable via the function (typically causal) they play in a system.

Example:

- Mental State: Pain
 - Function/Role: That state which is caused by physical injury and which causes certain beliefs (I am in an undesirable state) and desires (I need to remove myself from this state).
- Prominent advocates: D. Armstrong, F. Dretske, J. Fodor, D. Lewis, W. Lycan, B. McLaughlin and H. Putnam.

NB: They often endorse MR, as for them many different physical structures could play the same functional role.

The Chinese nation thought experiment

- A well-known objection to functionalism is the Chinese nation thought experiment (Block 1980).
- Imagine that the entire Chinese nation was recruited to simulate the human brain, where each person is a neuron.
- That is, they would be tasked to relay messages to each other in the way that neural networks in brains do so.
- Although the resulting 'Blockhead' would be functionally identical to a brain, it would lack mental states with qualia.
- Reply: Blockheads do have 'qualia'; it's just hard for us to imagine that they do.



Reductionism and Supervenience

Reductionism

- Identity theories endorse *some* form of reductionism.
NB: Polger (2009) claims that the two are not equivalent as identity theory is compatible w/a number of reductionisms.
- Roughly speaking, to reduce *B* to *A* is to:
 - * explain *B* law-like relations with *A* law-like relations
 - * expand (deepen and broaden) those explanationsand either:
 - * purge the ontology of *B* (eliminative reductionism) or retain it (non-eliminative reductionism).
- The ontological categories at stake in this case are mental types or tokens and brain (thus physical) types or tokens.

An even softer form of reduction?

- A set of states Y **supervenes** on a set of states X *iff* there can be no difference in Y without a difference in X .
- Applied to the current context:
Mental states M **supervene** on brain states B *iff* there can be no difference in M without a difference in B .
- It allows us to explain mental behaviour in terms of brain states *w/out necessarily* eliminating the mental states.
- In a nutshell, this notion allows some people, viz. epiphenomenalists, to have their cake and eat it too!



What Philosophers Think

The philpapers survey

- Carried out in Nov 2009 with 3226 respondents (1803 philosophy faculty members and/or PhDs).
- The following are the results when restricted to the only phil. of mind question and to 'Philosophy Faculty or PhD':

Mind: physicalism or non-physicalism?

Accept or lean toward: physicalism	981/1803 (54.4%)
Accept or lean toward: non-physicalism	521/1803 (27.1%)
Other	301/1803 (16.4%)

- The meta-survey indicated that professional philosophers predicted the trend on this question quite well.



The End