

# Lecture notes for Philosophy of Cognitive Science

An upper level under-graduate course taught by:  
Kelly Alexandra Roe

2010

# Introduction

In 2010 I was subcontracted to teach (lecture, tutor, grade) around  $\frac{3}{4}$  of a course in Philosophy and Cognitive Science at Macquarie University, Sydney.

The tenured Professor had received an Australian Research Council grant to research psychopathy so she subcontracted most of the teaching tutoring and grading work to me while I waited for the ANU to get my research out for external examination.

I wrote lecture note outlines to accompany the following text-book (which had already been ordered):

Clarke, A. (2002). *Mindware: An introduction to the philosophy of cognitive science*. Oxford University Press.

I have adapted them from power-point presentations that were uploaded to the online learning platform, and distributed as handouts in class.

The notes are basic in outline. They served as a prompt for me and a reminder to students of the reading they had done in advance. They were distributed in class and by way of online learning platform.

I was told by the department chair not to assume that students had taken previous courses in psychology, philosophy of psychology, or philosophy of mind.

Thanks to students for attending, engaging in such high-quality discussion in both the lectures and tutorials, and writing such wonderful essays in response to the content of this course.

I believe that Macquarie retains the audio recordings of my presentation of the lectures.

# Contents

1. Introduction	1
2. Dualism	5
3. Behaviourism	8
4. Identity theory	11
5. Eliminativism	13
6. Functionalism	15
7. Mind as meatware	19
8. Consciousness	23
9. Possibility, conceivability, supervenience, and zombies	28
10. The possible world framework, correlation, and identity	33
11. Operationalization, neural correlates of consciousness, qualitative experience	37
12. Symbol systems, intelligence, the Turing test	41
13. Objections to symbol systems	45
14. Folk psychology, the life world, stances	49
15. Intentional stance, indeterminacy, real patterns, true believers	53

16. Connectionism and symbol systems	54
17. Connectionism: Features and problems	58
18. Coltheart's challenge	62

# 1. Introduction

## What is cognitive science?

- Thinking, reasoning, inferring, choosing, deciding, willing, intending, loving, fearing, hoping, wishing, imagining, seeing, hearing, smelling, tasting, feeling, experiencing, dreaming...
- Inter-disciplinary science of cognition
  - Philosophy
  - Cognitive psychology
  - Computer science (artificial intelligence and robotics)
  - Linguistics
  - Anthropology
  - Neuroscience
  - Researchers and theorists in related fields such as education, developmental psychology, ethology etc.

## What is the role of philosophy in cognitive science?

- This is controversial. Two main views:
  1. There is a radical divide between philosophy and the natural sciences
  2. There is a continuum between philosophy and the natural sciences

## 1. The radical divide

- Some theorists maintain that there is a radical divide between philosophy and the natural sciences
  - Soul vs matter
  - Mind vs matter
  - Value vs matter
  - Meaning vs matter
- The findings of science are not relevant to answering philosophical questions
- Science misses the point / changes the subject
  - If this is so it may be that it is because philosophical theories aren't testable

## 2. The continuum

- There is some kind of reciprocal relationship or continuity between philosophy and the natural sciences
- Look to science for data that is relevant for answering philosophical questions
  - If this is so science may be data collection for the philosophical theorist
- Look to philosophy for questions, hypotheses, and / or theories
  - If this is so there may be no more philosophy with scientific progress
- In doing science we have to start somewhere
- Better to start from some place plausible rather than from some place implausible
- So let us start with some 'common-sense' intuitions
  - I mean to say I will try and articulate some of the following, hopefully in a way that seems intuitive to you

## Mental states as propositional attitudes

- Each of these mental states seems to be ABOUT something (p):

Thinking that p, reasoning that p, inferring that p, choosing p, deciding p, willing p, intending p, loving p, fearing that p, hoping that p, wishing that p, imagining p, seeing p, hearing p, smelling p, tasting p, experiencing p, dreaming p...
- What they are about (p) is the CONTENT of the mental state
- Mental contents are thought to be PROPOSITIONAL ATTITUDES
- Propositions are (for our purposes) abstract meanings or informational contents
  - ‘The sun is hot’
  - ‘Hot, the sun is’
  - Sentences in other language that are synonymous
- Different sentences, phrases, expressions, thoughts that have the same meaning, informational, or propositional content

## Some features of the propositional view

- One person can entertain the very same propositional content at different points in time
- Different people can entertain the very same propositional content
- Speakers of different languages can entertain the same propositional content (say the same things or think the same thoughts)

## Folk psychology (aka belief-desire psychology or ‘common-sense psychology’)

- We can make fairly good predictions about what people will do by appealing to what they believe (the way they represent the world to be) and what they desire (the way they are motivated to alter the world)
- We also explain actions by appealing to what people believed (represented) and what people desired (were motivated to do)

- It seems common-sense that (roughly) ‘an agent will act so as to satisfy or obtain their strongest desire under the assumption that their beliefs are true’
- No other (scientific) theory seems to capture the relevant predictions

## Questions

- What enables us to attribute mental states to others / employ folk-psychology?
- What does this ability show us about the structure of the mind?
- What is the status of folk-psychology compared to scientific theories in physics, chemistry, biology, cognitive psychology etc?
- What does folk-psychology show us about the nature of mental states?
- Are (all?) mental states usefully thought of as propositional attitudes?



## 2. Dualism

### Reminder of the cognitive phenomenon

- Thinking, reasoning, inferring, choosing, deciding, willing, intending, loving, fearing, hoping, wishing, imagining, seeing, hearing, smelling, tasting, feeling, experiencing, dreaming...

### Preview

- Theories of the mind-body (mind vs matter) relationship that we will be looking at:
  1. Dualism
  2. Behaviourism
  3. Identity theory
  4. Eliminativism
  5. Functionalism

### Dualism (mind is distinct from matter)

- *Why* is dualism thought to be plausible?
  - Introspection. It just seems (upon introspection) that the mind (or perhaps the soul) is different or distinct from matter
  - How could... How could certain phenomenon (the way that my pain feels or inferential reasoning, for example) arise from ‘matter nicely orchestrated’?
- There are two main kinds of dualism

1. Substance dualism
2. Property dualism

## 1. Substance dualism

- The mind (soul, spirit) is a substance (kind of stuff, object, or thing) that is fundamentally or radically different from matter

## 2. Property dualism

- Mental states are non-physical (immaterial) properties (states) of the physical brain or body

## Objects vs Properties

- In order to understand the difference between substance and property dualism we will need to take a look at the difference between substance and property
- Water is thought to be a substance
- Liquidity is thought to be a physical property of water (and substances other than water)
- Mental states (according to the dualist) are different from water or from fluidity in being non-physical or immaterial

## Dualism and causal interaction

- There are three main views on how the mental (substance or property) causally interacts with or relates to material or physical substance or property
  1. Interactionism
    - Two-way causal interaction between body and mind (e.g., Descartes)
  2. Epi-phenomenalism

- Matter causes changes to matter but mind is itself causally impotent (the shadow analogy)

### 3. Paralellism

- No causal interaction between body and mind (e.g., Leibniz)

## Problems for dualism

- Providing an account of non-material stuff / properties
- Understanding how causal interaction is possible (on the interactionist views)
- A simpler explanation posits only one kind of stuff (or property) rather than two (Morgan's cannon, Ockham's razor)
- Analogy with mind and vital spirit as science progresses and common-sense is revised

## Currently

- No many defenders of substance dualism
- A few defenders of property dualism for consciousness (more on that later) e.g., Block, Chalmers
- Most would say that dualism has been abandoned as the result of scientific advances (taking the vitalism analogy very seriously indeed)
- We will return to this when we look at consciousness

### 3. Behaviourism

#### Behaviourism (mental states are behavioural dispositions)

- Two main kinds:
  1. Analytical behaviourism
  2. Methodological behaviourism

#### Analytical behaviourism

- Primarily a thesis about how we should analyze mental discourse
- E.g., to say ' $x$  is in pain' just is to say that 'if  $x$ ' were placed in these circumstances then  $x$  would be disposed to...'
- Since mental language refers to dispositions to behave, mental states just are behaviours or dispositions to behave

#### Methodological behaviourism

- Different psychoanalytic / psychodynamic theories of the mind seemed 'unscientific'
- In order to become a 'real science' 'just like physics' the best bet for psychology is to become the science of behaviour (e.g., Skinner, Watson)
- Strictly speaking, methodological behaviourists don't need to be analytical behaviourists

- In practice most justify their view by appealing to analytical behaviourism, however

## Why is behaviourism thought to be plausible?

- Mostly because of LEARNING
- We don't observe others mental states directly - so how do we learn to label our own?
- The importance of behaviour as a source of evidence for figuring out what mental state a person (including ourself) is in

## Problems for analytical behaviourism

- Providing a fuller account of the dispositional profiles
  - Problems cashing out the dispositional profile (concern it might be infinite or circular)
  - The thought that the disposition might best be thought of as an inner state of the brain
- Doesn't seem so plausible as an account of the felt quality of experience (qualia, p-consciousness)

## Problems for methodological behaviourism

- The birth of cognitive psychology and the cognitive revolution showed that psychology didn't have to be about behaviours / behavioural dispositions in order to be scientific!

## Currently...

1. Within psychology it is often thought that cognitive psychology *replaced* behaviourism as a methodological paradigm
2. Functionalism may be viewed as an *extension or* development of analytical behaviourism, however (as we shall see)

3. In practice methodological behaviourism is alive and well in certain areas (e.g., animal behaviour in ethology, behavioural change especially in children, and the intellectually handicapped in clinical psychology)

## 4. Identity Theory

**Identity theory (mental states are brain states)**

**Identity theory aka: The ‘Australian Thesis’**

- Two main kinds:
  1. Type-type
  2. Token-token (arose later - a strategic retreat)

**Why is the identity theory thought to be plausible?**

- Just as science showed us that lightening just is a certain kind of electrical discharge...
- Science is showing us that mental states just are states of the brain
- Neuro-science is appealing to many. It seems to ‘pop the hood’ on behavioural dispositions
- We see lots of pictures of the neural correlates of various mental activities in journals

**Type-type identity theory**

- Types of mental states are types of brain states (e.g., pain = C fibres firing)

## Problems with type-type identity theory

- If mental states = brain states then beings without brains (e.g., computers, robots, certain kinds of aliens, angels, god) can't have mental states
- If dolphins don't have C-fibres (or whatever brain state we are in when we are in pain) then they can't be in pain and this seems counter-intuitive to most
- Leibniz law objections
  - If you have x and y and want to know whether  $x=y$  (where = is the identity relation that each object bears only to itself then if you can find a property that x has that y lacks (or vice versa) then you can conclude that x does not = y. Some candidates:
    - \* My pain is in my toe but my brain is in my head
    - \* My beliefs can be true or false but my brain states can't be
  - Responses
    - \* It might *seem to you* that they have different properties
    - \* But that is a feature of *you* and you are *wrong or misguided*
    - \* E.g., Your brain *represents* bodily damage in your toe and that state of your brain just is pain. So the pain is in your brain but your brain *represents* the pain as being in your toe
    - \* Similarly, what your brain *represents* can be true or false (e.g., referred pain) and beliefs are just your brain *representing* things to be a certain way
  - Currently a number of neuroscientists / cognitive neuroscientists think that type-type identities have been or will be made between mental or cognitive states and neuro-physiological states
  - A number of other neuroscientists or cognitive neuroscientists think that actually type-type identities haven't been as forthcoming as we would have hoped
  - The later has fuelled the two further developments that we will look at - eliminativism and functionalism.



## 5. Eliminativism

### Eliminativism (there aren't any mental states)

- Folk-psychology and our common-sense understanding of mental states involves our committing to a certain view of their nature (e.g., that mental states are types of brain states)
- It turns out that (according to some neuroscientists) mental states aren't correlated with brain states
- Therefore, neuroscience has shown us that there aren't any mental states. Just like science showed us that there isn't any phlogiston
- Paul and Patricia Churchland think that as neuroscience matures the vocabulary of neuroscience will come to replace the vocabulary of folk-psychology / mentalistic discourse
- Neuroscience textbooks don't talk about (have a chapter on) 'belief' so it seems that some of our mentalistic discourse already has been eliminated from neuroscience, at least.

### Problems for eliminativism

- We can't just eliminate mentalistic discourse from our everyday lives and go on with business as usual.
- Neuroscience seems to be at the wrong grain to capture the predictions that can be made from folk-psychology (e.g., that people will turn up to class this week because of certain beliefs and desires they have. We will return to this.
- Perhaps any theory that commits us to concluding that 'there aren't any mental states' must be false

## Problems for mentalistic discourse

- But then isn't mentalistic discourse scientific if it isn't open to being falsified?
- What does that imply for a science of cognition?

## 6. Functionalism

**Functionalism (mental states are functional role states)**

### Preview

- Functionalism, functions, functional roles
- Machine tables as functional role characterisations
- Semantics, syntax, reducing semantics to syntax
- Kinds of functionalism
- A concern for machine functionalism
- There has been a pendulum between dualism and materialism through history
  - Functionalism is an attempts to avoid the pendulum
- Wouldn't it be nice if there were a theory to capture what is plausible in what went before while avoiding some of the problems?
  - Functionalism is an attempt to do so

### What are functions?

- Think of mathematical functions:
  - $-$ ,  $+$ ,  $=$ ,  $\times$  (mathematical operators)
- Or logical functions:
  - $\neg$ ,  $\vee$ ,  $\rightarrow$ , (logical operators)

- Transformations:
  - (Modus ponens, or disjunction introduction)
- Or syntactic functions:
  - Rules of combining words into sentences
  - Transformations (e.g., present to past tense)

## What is a functional role?

- In order for  $x$  to count as a state of belief  $x$  must play the functional role of belief
- Functional roles are abstract, structural, formal, or syntactical properties

## A Machine Table

- Example of a soda machine that takes 50c and 1\$ coins. Soda costs \$1.50 and the machine will give change
- 4 states of the machine:
  - State 0
    - \* If \$1 is input then goto state 1
    - \* If 50c is input then goto state 2
  - State 1
    - \* If 50c is input then output soda and goto state 0
    - \* If \$1 is input then output soda and output 50c and goto state 0
  - State 2
    - \* If \$1 is input then output soda and goto state 0
    - \* If 50c is input then goto state 3
  - State 3
    - \* If \$1 is input then output soda and output 50c and goto state 0

- \* If 50c is input then output soda and goto state 0

## Functionalism and machine tables

- The machine table specifies different (internal) states of the coke machine (0, 1, 2, and 3)
- Each state is defined by its abstract structural, formal, or syntactic relation to:
  1. Inputs (\$1 and 50c)
  2. Internal states (goto)
  3. Outputs (soda, change)
- Functionalism is thus a tripartite, or three-part theory

## Semantics vs syntax

- A semantics for a language is the meaning, or informational content that the syntax provides rules for manipulating
- A semantics for logic would replace content-less variables (p, q etc) with semantic contents (e.g., Socrates, man)
- Thus we have a distinction between content / meaning and rules that govern content / meaning transitions
- One (controversial) thought is that ‘if you take care of the syntax then the semantics will take care of itself’
- The thought is that semantic content (e.g., ‘dog’) can be characterised syntactically with respect to:
  - Typical input (dogs)
  - Inferences it licenses (is not a cat), the relation it bears to other states (e.g., desires)
  - The output that is produced (e.g., petting)
- We will return to look at machine intelligence
  - Can programming an appropriate syntax give machines content to think about (genuine understanding)?

## Kinds of functionalism.

- How do we specify the functional role of the different kinds of mental states?
  - Machine functionalism - look to logic / syntactic transformation rules
  - Analytic functionalism - look to common-sense folk-psychology
  - Empirical functionalism - look to science (e.g., cognitive psychology, biological psychology)
  -

## A concern

- ‘Chauvinism’ was an objection to the type-type identity theory
  - Beings with different brains or no brains could have mental states
- ‘Excessive liberalism’ is an objection to machine functionalism
  - A bucket of river water warming in the sun can probably be described as instantiating any computational description that you care to think of
- Getting the balance between these is tricky

# 7. Mind as meatware

## Mind as meatware

### Preview

- The role / realizer distinction
- Multiple realizability
- Software / hardware
- Mindware / wetware

### Roles vs realizers

- We have seen that functionalists think that mental states are functional roles
- The functional role that is thought to be relevant depends on the version of functionalism (whether the functional role is to be given by common-sense, computational specification, or by the empirical biological sciences)
- It is only in virtue of the state playing the relevant functional role that the state would count as being a mental state
- Functionalists identify mental states with the functional role being filled rather than with whatever it is that happens to fill the functional role
- That is what makes functionalism different from the view that mental states are to be identified with (are one and the same as) whatever it is that happens to realize the role (e.g., that they are brain states or immaterial states)

- Example: Consider a doorstep. A doorstep (let us suppose) is whatever it is that plays the doorstep role.
- A shoe, bag, block of wood, rolled up newspaper etc isn't a doorstep - except insofar as it is realizing or instantiating the doorstep role. That is to say it is being used as a doorstep.
- If you were to go a step further and say that that particular shoe really is a doorstep (even when it is not being used as a doorstep) then that would be token-token identity theory. That means to say this particular (token of) a shoe is (an 'is' of identity) a particular instance or token of a doorstep.
- If you were to go a step further and say that shoes are doorstops that would be type-type identity theory regarding the relationship between shoes and doorstops

## Multiple realizability

- While the role of the states can be specified by their inputs, internal relations to each other, and their outputs the realizers (particular things) that fill or instantiate the role can be made of glass, copper, tin, plastic, immaterial souls or ghosts etc
- Thus functionalism (strictly speaking) avoids the pendulum swing between our having to choose materialism or dualism by remaining neutral or agnostic as to the nature of the realizers
- The realizers could be neural states or silicon states of a computer or nitrogen hydraulics of aliens or immaterial states of a Cartesian soul or animus from the breath of gods
- Mental states are thus multiply realizable
- Which means that we can have a science of the mind / cognition without worrying about neurons or the nature of the hardware.

## Mind is to brain as software is to hardware

- The intuitive idea is that the same software programme (e.g., Microsoft word) can run on different hardware (e.g., PC, mac)



- Though they acknowledge that hardware constrains software (e.g., you can't run Microsoft word on water)
- We can consider features of the Microsoft word programme abstractly enough so that the different hardware is irrelevant
  - Critics maintain that differences in hardware make important differences to relevant features of the software (e.g., processing time)

## Mindware / wetware

- Similarly, while some cognitive psychologists maintain that the mind program can be studied in abstraction from the neural implementation...
- They allow that the neural basis does impose some constraints (e.g., on processing time) but they don't think that those constraints are particularly relevant for understanding the mind
  - Critics maintain that neurological differences will turn out to be crucially important and they cannot be ignored

## Multiple realizability and types

- So while multiple realizability is typically thought to be a feature of functionalism, critics maintain that it doesn't hold up. It is hard to know how the science will go.

## Types of states

- There is an issue around what types of mental states there are
  - E.g., folk psychology considers 'memory' a type of mental state but folk psychology considers 'iconic visual sensory register'. 'semantic memory'. 'episodic memory,' etc to be different types of mental states
- There is an issue around what types of brain states there are
  - E.g., activation, or a more particular kind of activation?
- Maybe if we got both of those right there would be type-type correlations

## To come:

- There are issues around whether a computer programmed to follow syntax could ever represent, have content, or understand
- And even if it could, could it have conscious experience (e.g., feel pain)?
- 
- Is it any more surprising that meat (or bundles of nerves) could represent, have content, understand, and have conscious experience?
- When we resume we will look at consciousness to see whether it poses a special problem
- Then we will return to the issue of machine intelligence

## 8. Consciousness

### Consciousness

- What is consciousness? Some candidates:
  - Awakeness
  - Self awareness
  - Availability for verbal report
  - Availability for the control of intentional action
  - Qualia (qualitative experience, phenomenal awareness)
- Used to be regarded as ‘off limits’ for science or scientific research
- Now if you read the cognitive neuroscience literature you might well think that the problem is solved
- Fashionable topic, currently
- Some people think that much of the science misses the point
- Other people think that the philosophical notion of consciousness needs to be rehabilitated else eliminated
- Distinction between A (access) consciousness and P (phenomenal) consciousness
- The (comparatively) ‘easy’ and ‘hard’ problems of consciousness

### Awakeness

- The distinction or difference between being asleep and being awake
- Relevant for anesthesiology

- ‘She’s unconscious’ and ‘she’s asleep’ seem to be used synonymously or treated as synonyms
- We are learning much about the (relatively) primitive brain structures that regulate sleep and wakefulness
- *But it seems that we can have conscious experiences while asleep - e.g., the experiences we have while dreaming.*

## Self awareness

- The capacity for ‘meta-cognition’ - to reflect on our cognitive states and experiences
- To have a sense of ourself as persisting through time, having different projects and preferences and dreams
- People with dissociative identity disorder (formerly known as multiple personality disorder) experience a ‘fragmented sense of personal identity’ (multiple selves). This is thought to be a disorder of consciousness
- *Small children and animals lack a sense of self-awareness but still have experiences*

## Availability for verbal report

- Often the best way to know what a person is experiencing is to ask them
- We seem to be able to report on our conscious experiences
- *Seems to suggest that those who lack verbal capacity (e.g., animals and small children) lack conscious experience insofar as their experiences are not available for verbal report*
- *Seems possible to promptly FORGET a conscious experience. E.g., would you rather take a drug for surgery that blocked the experience of pain or blocked the availability of the experience to verbal report?*

## Availability for control of intentional action

- People can often use their conscious experiences to guide a diversity of plans, projects, and goals

- E.g., psychological experiments where people press a key when they have certain experiences. These experiments are interpreted as showing us whether people have consciously experienced the stimuli
- It is unclear how much animals and small children are able to control intentional action yet it seems intuitive to most of us that they are conscious
- Would you rather lack the experience of pain or lack the availability of the experience of pain to guide your action?

## **Qualitative experiences or qualia (singular - 'quale')**

- Probably impossible to define
- Can gesture towards the phenomena... And hope that people get the intuition
  - Descartes - *cognito ergo sum*
  - The particular way that your experiences feel or seem to you from your first person point of view
  - The sum total of your experience right now is your 'phenomenal field'

## **Some scientific research on consciousness**

- The binding problem (the unity of the phenomenal field)
- Blindsight
- Dorsal vs ventral processing

## **Binding**

- The experiential field seems to be unified
  - Within modalities. E.g., objects near and far away seem to be present at the same time
  - Between modalities. E.g., vision and audition

- The binding problem has to do with how the brain manages to unify sensory modalities to underwrite the unity of experience
  - E.g., the 40hz thesis (when different regions fire in this frequency the information presents as bound in the phenomenal field)

## Blindsight

- Cortical damage results in a sciomata (blind spot) in the visual field
- Can present information to that region
- Patients report no conscious visual experience
- But when they are forced to ‘guess’ their guesses are above chance

## Dorsal vs ventral processing

- Dorsal stream - ‘where’ (spatial location) - guidance of action
- Ventral stream - ‘what’ (identification / recognition) - perception involved in visual awareness
- So action without awareness (ventral deficit) or awareness without appropriate action (dorsal deficits)

## Ned Block

- A consciousness (access consciousness) - information poised to control action and verbal report
- P consciousness (phenomenal consciousness) - qualia (qualitative experience or the felt quality of experience)

## David Chalmers

- The (comparatively) ‘easy problem’ of consciousness - learning about states that are poised to control action and verbal report
- The (comparatively) ‘hard problem’ of consciousness - learning about qualia

- ‘What is striking is that it is only that final target (qualia) that threatens to present any special kind of problem for our standard modes of cognitive scientific explanation and understanding’ (Textbook)
- 

## **Some questions for further thought / discussion**

- Do you have the intuition that an understanding / explanation of consciousness will involve understanding / explaining qualia or P consciousness?
- If so, then what has cognitive science shown us about consciousness thus far?
- If not, then should ‘Phenomenal consciousness’ go the way of vial spirits (i.e., be eliminated from the subject matter of science)?

## 9. Possibility, conceivability, supervenience, and zombies

### Possibility, conceivability, supervenience, and zombies

- Possibility (logical, physical, biological)
- Conceivability (seems possible to me for all I know - epistemic possibility)
- Supervenience (baldness supervenes on hair distribution)
- Zombies - (intended to show that consciousness does not supervene on material states)

### Possibility

- Think of 'possibility' as a space (the space of possibility)
- Different possible worlds (located within that space) are different ways the world might be
- E.g., there is a possible world in which Bush won the last US election
- The actual world that we inhabit - @
- Sometimes people talk about different kinds of possibility
  - Logical
  - Physical (or metaphysical)
  - Biological



- Each of these notions places different constraints on the limitations of the space / the worlds within that space

## Logical possibility

- Logical possibility sets the outer limits on what is possible
- Contradictions are necessarily false or logically impossible - which is just to say they are false in all possible worlds
- Tautologies are necessarily true or logically necessary - which is just to say they are true in all possible worlds
  - Logical or mathematical truths are thought to be tautologies (e.g.,  $1+1=2$ . Either  $p$  or not  $p$ ).

## Physical possibility

- Physical possibility is a subset of logical possibility because it has an additional constraint:
- Not only are contradictions ruled out - but worlds that are inconsistent with the laws of physics at @ (at the actual world) are ruled out
- So while it is logically possible that there are worlds with laws of physics that are different from ours (there is no contradiction in that)...
- It is not physically possible that there are worlds with laws of physics that are different from ours

## Possibility and conceivability

- Conjectures in math are either true or false
- If they are true they are necessarily true (there are no worlds in which they are false)
- If they are false they are necessarily false (false in all worlds)
- However, from my point of view it seems to me (in some sense) that it is 'possible' that it be true and also 'possible' that it is false - as I could conceive (in some sense) of the math turning out either way

## Conceivability / epistemic possibility

- What ‘seems possible to me given the state of my knowledge’ is what is conceivable to me, or what is epistemically possible for me - given what I know
- Conceivability has to do with what finite minds like ours can imagine (or what we think we can imagine)
- Conceivability seems to be subjective in the sense that different people could conceive or fail to conceive of different things

## Possibility

- Possibility has to do with what does or does not follow given certain constraints (e.g., non-contradiction or the laws of physics)
- Possibility is objective. What is and is not possible is independent of what any of us think is possible (e.g., ‘Goldbach’s conjecture is true’ is either necessarily true else it is impossible). So, in some sense, it is not possible that it turn out either way.
- By analogy, let us suppose that on day one God does two (and only two) things:
  - God fixes all the laws of fundamental physics
  - God fixes the nature and distribution of the fundamental units of the fundamental physics
- Question: On day two does God have more work to do? Does he have to fix the nature and distribution of the fundamental units of chemistry?
- Or does the chemistry just ‘fall out of’ (so to speak) the physics?

## Supervenience

- To say that A supervenes on B is to say: There cannot be a change in A without a change in B
  - E.g., to say that ‘baldness supervenes on hair distribution’ is to say that if two people are alike in hair distribution then they are alike in baldness

- It is important that supervenience is an a-symmetric relation. To say that A supervenes on B does not rule out a change in B without a change in A
  - E.g., to say that ‘baldness supervenes on hair distribution’ does not rule out two people being alike in their baldness but different in the way that hair is distributed (people can go bald in different places on their scalp)
- To say ‘chemistry supervenes on physics’ is to say that two worlds cannot be alike in physics without being alike in chemistry
  - So that God could have rested on day 2 if he had have chosen to have approached day 1 by fixing the laws of physics and the distribution of physics fundamental particles
  - It is not to rule out the possibility of two worlds that are alike in chemistry but different in physics
- This is similar to multiple realizability

## Zombies (finally!)

- Imagine (conceive of) a world that is a complete physical duplicate of this world (the actual world)
- On this world you have a counter-part (defined as an atom for atom duplicate of you)
- Your counter-part says the things you say and does the things you do...
  - Seems to be an A-consciousness duplicate of you
- Zombies are defined as physical duplicates of actual people that lack p-consciousness
- The possibility of zombies would mean that p-consciousness does not supervene on physics
- The possibility of zombies would show dualism to be true of p-consciousness
- Materialists counter that while we may think that zombies are possible we are mistaken and they aren’t possible at all.
- Materialists say that they are conceivable but not possible and one can’t infer possibility from conceivability (e.g., in the case of Goldbach’s

conjecture)

- Else they say they don't believe in p-consciousness (eliminativism)
- Next week: More on consciousness

# 10. The possible world framework, correlation, and identity

## Plan

- Last time we briefly looked at some of the ways that consciousness has been operationalized so that we could get underway with a science of consciousness
- Then we turned to the possible worlds framework to try and understand the difference between possibility and conceivability
- This time we will start by using the possible worlds framework to try and understand correlation and numeric identity ( $p=p$ )
- Then we will return to operationalizations, and the issue of whether discovering neural correlates of consciousness shows us that consciousness is one and the same as the physical correlates

## The possible worlds framework

- Braddon-Mitchell and Jackson, and Chalmers say to think of possible worlds as ‘universes’
- Modal realists think that possible worlds are objectively existing and concrete (though spatio-temporally and causally isolated from this world)
- Other theorists think that possible worlds are best understood as something along the lines of sets of ‘maximally complete sentences / propositions’

## Possible worlds

- Fictional worlds are incomplete insofar as there are truth value gaps
  - E.g., ‘Cinderella had 10,000 hairs on her head when she put on the shoe that fit’
- Possible worlds are maximally complete insofar as there are no truth value gaps
  - Either because there is a fact about the world (on a modal realist view)
  - Or because worlds are constructed by stipulating (consistent) truth values for sentences / propositions (on the view that they are sets of sentences / propositions)

## Correlation

- x and y are correlated in the actual world if whenever x occurs y occurs (and vice versa)
- Correlations can be contingent, however
- This is just to say that while x and y might be correlated in the actual world it might be possible that they not be correlated
- This is just to say that there are possible worlds (or there is a non-contradictory set of sentences describing a situation or world) in which they aren’t correlated

## Identity

- x and y are numerically identical if there is one object (substance, property etc) rather than two
- An object (substance, property etc) is numerically identical to itself
  - $p=p$ , or brain state x = brain state x
- If there is no correlation between x and y then x and y cannot be numerically identical (one and the same object, property, etc)
- Leibniz law describes this:

- If x has a property that y lacks (or vice versa) then x does not = y

## Contingency of correlation, necessity of identity

- Correlations may be contingent
  - If x and y are actually correlated it may be possible that they not be
- Identities are necessary, however
  - If x and y are numerically identical then they are in all possible worlds
- This is because an object is always identical to itself (p necessarily = p, brain state b necessarily = brain state b)
- So, while correlation in the actual world is necessary for identity it is not sufficient
- That is to say that if there is not a correlation there cannot be an identity but if there is a correlation this is not sufficient to establish identity

## Informational value

- To say that 'p=p' or 'brain state b = brain state b' seems uninformative
- To say that 'brain state b = mental state m' seems informative, however
  - Informativeness seems to do with the state of our knowledge
  - Conceivability was relative to the state of our knowledge
- But it is that we can conceive of things turning out either way rather than it being possible for things to turn out either way
- If the identity holds in the actual world it holds in all possible worlds (it is necessary)

## Gold

- Gold = 79 protons in the nucleus of the atoms
- if the above identity claim is true then it holds in the actual world and in all possible worlds
  - If we were able to remove a proton from the nucleus of the atoms of a sample then we would have transmuted the substance from gold to something else
  - Similarly, if there was a possible world in which the yellowy malleable valuable stuff turned out to have a different number of protons in the nucleus then that substance would not be gold

## Water

- Water =  $H_2O$
- If the above identity claim is true then it holds in the actual world and in all possible worlds
  - So, if the colorless, odorless stuff that falls from the skies and fills the lakes, the drinkable potable stuff is  $XYZ$  then it is not water

## Neural correlates of consciousness

- Thus both materialists and dualists can be interested in discovering the neural correlates of consciousness
  - Identity theorists think that the discovery of neural correlates is a discovery of the identity of conscious states
  - Dualists think that the discovery of neural correlates is nothing more than that
  - Functionalists think that the discovery of neural correlates that fill the functional role isn't a discovery of the identity of conscious states (because they identify conscious states with the role being filled rather than with the filler of the role)



# 11. Operationalization, neural correlates of consciousness, qualitative experience

## Plan

- Last time we briefly looked at some of the ways that consciousness has been operationalized so that we could get underway with a science of consciousness
- Then we turned to the possible worlds framework to try and understand the difference between possibility and conceivability
- This time we will start by using the possible worlds framework to try and understand correlation and numeric identity ( $p=p$ )
- Then we will return to operationalizations, and the issue of whether discovering neural correlates of consciousness shows us that consciousness is one and the same as the physical correlates

## The possible worlds framework

- Braddon-Mitchell and Jackson, and Chalmers say to think of possible worlds as ‘universes’
- Modal realists think that possible worlds are objectively existing and concrete (though spatio-temporally and causally isolated from this world)
- Other theorists think that possible worlds are best understood as something along the lines of sets of ‘maximally complete sentences / propositions’

## Possible worlds

- Fictional worlds are incomplete insofar as there are truth value gaps
  - E.g., ‘Cinderella had 10,000 hairs on her head when she put on the shoe that fit’
- Possible worlds are maximally complete insofar as there are no truth value gaps
  - Either because there is a fact about the world (on a modal realist view)
  - Or because worlds are constructed by stipulating (consistent) truth values for sentences / propositions (on the view that they are sets of sentences / propositions)

## Correlation

- x and y are correlated in the actual world if whenever x occurs y occurs (and vice versa)
- Correlations can be contingent, however
- This is just to say that while x and y might be correlated in the actual world it might be possible that they not be correlated
- This is just to say that there are possible worlds (or there is a non-contradictory set of sentences describing a situation or world) in which they aren’t correlated

## Identity

- x and y are numerically identical if there is one object (substance, property etc) rather than two
- An object (substance, property etc) is numerically identical to itself
  - $p=p$ , or brain state x = brain state x
- If there is no correlation between x and y then x and y cannot be numerically identical (one and the same object, property, etc)
- Leibniz law describes this:

- If x has a property that y lacks (or vice versa) then x does not = y

## Contingency of correlation, necessity of identity

- Correlations may be contingent
  - If x and y are actually correlated it may be possible that they not be
- Identities are necessary, however
  - If x and y are numerically identical then they are in all possible worlds
- This is because an object is always identical to itself (p necessarily = p, brain state b necessarily = brain state b)
- So, while correlation in the actual world is necessary for identity it is not sufficient
- That is to say that if there is not a correlation there cannot be an identity but if there is a correlation this is not sufficient to establish identity

## Informational value

- To say that 'p=p' or 'brain state b = brain state b' seems uninformative
- To say that 'brain state b = mental state m' seems informative, however
  - Informativeness seems to do with the state of our knowledge
  - Conceivability was relative to the state of our knowledge
- But it is that we can conceive of things turning out either way rather than it being possible for things to turn out either way
- If the identity holds in the actual world it holds in all possible worlds (it is necessary)

## Gold

- Gold = 79 protons in the nucleus of the atoms
- if the above identity claim is true then it holds in the actual world and in all possible worlds
  - If we were able to remove a proton from the nucleus of the atoms of a sample then we would have transmuted the substance from gold to something else
  - Similarly, if there was a possible world in which the yellowy malleable valuable stuff turned out to have a different number of protons in the nucleus then that substance would not be gold

## Water

- Water =  $H_2O$
- If the above identity claim is true then it holds in the actual world and in all possible worlds
  - So, if the colorless, odorless stuff that falls from the skies and fills the lakes, the drinkable potable stuff is  $XYZ$  then it is not water

## Neural correlates of consciousness

- Thus both materialists and dualists can be interested in discovering the neural correlates of consciousness
  - Identity theorists think that the discovery of neural correlates is a discovery of the identity of conscious states
  - Dualists think that the discovery of neural correlates is nothing more than that
  - Functionalists think that the discovery of neural correlates that fill the functional role isn't a discovery of the identity of conscious states (because they identify conscious states with the role being filled rather than with the filler of the role)

## 12. Symbol systems, intelligence, the Turing test

### Plan

- Last time we briefly looked at some of the ways that consciousness has been operationalized so that we could get underway with a science of consciousness
- Then we turned to the possible worlds framework to try and understand the difference between possibility and conceivability
- This time we will start by using the possible worlds framework to try and understand correlation and numeric identity ( $p=p$ )
- Then we will return to operationalizations, and the issue of whether discovering neural correlates of consciousness shows us that consciousness is one and the same as the physical correlates

### The possible worlds framework

- Braddon-Mitchell and Jackson, and Chalmers say to think of possible worlds as ‘universes’
- Modal realists think that possible worlds are objectively existing and concrete (though spatio-temporally and causally isolated from this world)
- Other theorists think that possible worlds are best understood as something along the lines of sets of ‘maximally complete sentences / propositions’

## Possible worlds

- Fictional worlds are incomplete insofar as there are truth value gaps
  - E.g., ‘Cinderella had 10,000 hairs on her head when she put on the shoe that fit’
- Possible worlds are maximally complete insofar as there are no truth value gaps
  - Either because there is a fact about the world (on a modal realist view)
  - Or because worlds are constructed by stipulating (consistent) truth values for sentences / propositions (on the view that they are sets of sentences / propositions)

## Correlation

- x and y are correlated in the actual world if whenever x occurs y occurs (and vice versa)
- Correlations can be contingent, however
- This is just to say that while x and y might be correlated in the actual world it might be possible that they not be correlated
- This is just to say that there are possible worlds (or there is a non-contradictory set of sentences describing a situation or world) in which they aren’t correlated

## Identity

- x and y are numerically identical if there is one object (substance, property etc) rather than two
- An object (substance, property etc) is numerically identical to itself
  - $p=p$ , or brain state x = brain state x
- If there is no correlation between x and y then x and y cannot be numerically identical (one and the same object, property, etc)
- Leibniz law describes this:

- If x has a property that y lacks (or vice versa) then x does not = y

## Contingency of correlation, necessity of identity

- Correlations may be contingent
  - If x and y are actually correlated it may be possible that they not be
- Identities are necessary, however
  - If x and y are numerically identical then they are in all possible worlds
- This is because an object is always identical to itself (p necessarily = p, brain state b necessarily = brain state b)
- So, while correlation in the actual world is necessary for identity it is not sufficient
- That is to say that if there is not a correlation there cannot be an identity but if there is a correlation this is not sufficient to establish identity

## Informational value

- To say that 'p=p' or 'brain state b = brain state b' seems uninformative
- To say that 'brain state b = mental state m' seems informative, however
  - Informativeness seems to do with the state of our knowledge
  - Conceivability was relative to the state of our knowledge
- But it is that we can conceive of things turning out either way rather than it being possible for things to turn out either way
- If the identity holds in the actual world it holds in all possible worlds (it is necessary)

## Gold

- Gold = 79 protons in the nucleus of the atoms
- if the above identity claim is true then it holds in the actual world and in all possible worlds
  - If we were able to remove a proton from the nucleus of the atoms of a sample then we would have transmuted the substance from gold to something else
  - Similarly, if there was a possible world in which the yellowy malleable valuable stuff turned out to have a different number of protons in the nucleus then that substance would not be gold

## Water

- Water =  $H_2O$
- If the above identity claim is true then it holds in the actual world and in all possible worlds
  - So, if the colorless, odorless stuff that falls from the skies and fills the lakes, the drinkable potable stuff is  $XYZ$  then it is not water

## Neural correlates of consciousness

- Thus both materialists and dualists can be interested in discovering the neural correlates of consciousness
  - Identity theorists think that the discovery of neural correlates is a discovery of the identity of conscious states
  - Dualists think that the discovery of neural correlates is nothing more than that
  - Functionalists think that the discovery of neural correlates that fill the functional role isn't a discovery of the identity of conscious states (because they identify conscious states with the role being filled rather than with the filler of the role)



# 13. Objections to symbol systems

## Plan

- Last time we briefly looked at some of the ways that consciousness has been operationalized so that we could get underway with a science of consciousness
- Then we turned to the possible worlds framework to try and understand the difference between possibility and conceivability
- This time we will start by using the possible worlds framework to try and understand correlation and numeric identity ( $p=p$ )
- Then we will return to operationalizations, and the issue of whether discovering neural correlates of consciousness shows us that consciousness is one and the same as the physical correlates

## The possible worlds framework

- Braddon-Mitchell and Jackson, and Chalmers say to think of possible worlds as ‘universes’
- Modal realists think that possible worlds are objectively existing and concrete (though spatio-temporally and causally isolated from this world)
- Other theorists think that possible worlds are best understood as something along the lines of sets of ‘maximally complete sentences / propositions’

## Possible worlds

- Fictional worlds are incomplete insofar as there are truth value gaps
  - E.g., ‘Cinderella had 10,000 hairs on her head when she put on the shoe that fit’
- Possible worlds are maximally complete insofar as there are no truth value gaps
  - Either because there is a fact about the world (on a modal realist view)
  - Or because worlds are constructed by stipulating (consistent) truth values for sentences / propositions (on the view that they are sets of sentences / propositions)

## Correlation

- x and y are correlated in the actual world if whenever x occurs y occurs (and vice versa)
- Correlations can be contingent, however
- This is just to say that while x and y might be correlated in the actual world it might be possible that they not be correlated
- This is just to say that there are possible worlds (or there is a non-contradictory set of sentences describing a situation or world) in which they aren’t correlated

## Identity

- x and y are numerically identical if there is one object (substance, property etc) rather than two
- An object (substance, property etc) is numerically identical to itself
  - $p=p$ , or brain state x = brain state x
- If there is no correlation between x and y then x and y cannot be numerically identical (one and the same object, property, etc)
- Leibniz law describes this:

- If x has a property that y lacks (or vice versa) then x does not = y

## Contingency of correlation, necessity of identity

- Correlations may be contingent
  - If x and y are actually correlated it may be possible that they not be
- Identities are necessary, however
  - If x and y are numerically identical then they are in all possible worlds
- This is because an object is always identical to itself (p necessarily = p, brain state b necessarily = brain state b)
- So, while correlation in the actual world is necessary for identity it is not sufficient
- That is to say that if there is not a correlation there cannot be an identity but if there is a correlation this is not sufficient to establish identity

## Informational value

- To say that 'p=p' or 'brain state b = brain state b' seems uninformative
- To say that 'brain state b = mental state m' seems informative, however
  - Informativeness seems to do with the state of our knowledge
  - Conceivability was relative to the state of our knowledge
- But it is that we can conceive of things turning out either way rather than it being possible for things to turn out either way
- If the identity holds in the actual world it holds in all possible worlds (it is necessary)

## Gold

- Gold = 79 protons in the nucleus of the atoms
- if the above identity claim is true then it holds in the actual world and in all possible worlds
  - If we were able to remove a proton from the nucleus of the atoms of a sample then we would have transmuted the substance from gold to something else
  - Similarly, if there was a possible world in which the yellowy malleable valuable stuff turned out to have a different number of protons in the nucleus then that substance would not be gold

## Water

- Water =  $H_2O$
- If the above identity claim is true then it holds in the actual world and in all possible worlds
  - So, if the colorless, odorless stuff that falls from the skies and fills the lakes, the drinkable potable stuff is  $XYZ$  then it is not water

## Neural correlates of consciousness

- Thus both materialists and dualists can be interested in discovering the neural correlates of consciousness
  - Identity theorists think that the discovery of neural correlates is a discovery of the identity of conscious states
  - Dualists think that the discovery of neural correlates is nothing more than that
  - Functionalists think that the discovery of neural correlates that fill the functional role isn't a discovery of the identity of conscious states (because they identify conscious states with the role being filled rather than with the filler of the role)

# 14. Folk psychology, the life world, stances

## Plan

- Last time we briefly looked at some of the ways that consciousness has been operationalized so that we could get underway with a science of consciousness
- Then we turned to the possible worlds framework to try and understand the difference between possibility and conceivability
- This time we will start by using the possible worlds framework to try and understand correlation and numeric identity ( $p=p$ )
- Then we will return to operationalizations, and the issue of whether discovering neural correlates of consciousness shows us that consciousness is one and the same as the physical correlates

## The possible worlds framework

- Braddon-Mitchell and Jackson, and Chalmers say to think of possible worlds as ‘universes’
- Modal realists think that possible worlds are objectively existing and concrete (though spatio-temporally and causally isolated from this world)
- Other theorists think that possible worlds are best understood as something along the lines of sets of ‘maximally complete sentences / propositions’

## Possible worlds

- Fictional worlds are incomplete insofar as there are truth value gaps
  - E.g., ‘Cinderella had 10,000 hairs on her head when she put on the shoe that fit’
- Possible worlds are maximally complete insofar as there are no truth value gaps
  - Either because there is a fact about the world (on a modal realist view)
  - Or because worlds are constructed by stipulating (consistent) truth values for sentences / propositions (on the view that they are sets of sentences / propositions)

## Correlation

- x and y are correlated in the actual world if whenever x occurs y occurs (and vice versa)
- Correlations can be contingent, however
- This is just to say that while x and y might be correlated in the actual world it might be possible that they not be correlated
- This is just to say that there are possible worlds (or there is a non-contradictory set of sentences describing a situation or world) in which they aren’t correlated

## Identity

- x and y are numerically identical if there is one object (substance, property etc) rather than two
- An object (substance, property etc) is numerically identical to itself
  - $p=p$ , or brain state x = brain state x
- If there is no correlation between x and y then x and y cannot be numerically identical (one and the same object, property, etc)
- Leibniz law describes this:

- If x has a property that y lacks (or vice versa) then x does not = y

## Contingency of correlation, necessity of identity

- Correlations may be contingent
  - If x and y are actually correlated it may be possible that they not be
- Identities are necessary, however
  - If x and y are numerically identical then they are in all possible worlds
- This is because an object is always identical to itself (p necessarily = p, brain state b necessarily = brain state b)
- So, while correlation in the actual world is necessary for identity it is not sufficient
- That is to say that if there is not a correlation there cannot be an identity but if there is a correlation this is not sufficient to establish identity

## Informational value

- To say that 'p=p' or 'brain state b = brain state b' seems uninformative
- To say that 'brain state b = mental state m' seems informative, however
  - Informativeness seems to do with the state of our knowledge
  - Conceivability was relative to the state of our knowledge
- But it is that we can conceive of things turning out either way rather than it being possible for things to turn out either way
- If the identity holds in the actual world it holds in all possible worlds (it is necessary)

## Gold

- Gold = 79 protons in the nucleus of the atoms
- if the above identity claim is true then it holds in the actual world and in all possible worlds
  - If we were able to remove a proton from the nucleus of the atoms of a sample then we would have transmuted the substance from gold to something else
  - Similarly, if there was a possible world in which the yellowy malleable valuable stuff turned out to have a different number of protons in the nucleus then that substance would not be gold

## Water

- Water =  $H_2O$
- If the above identity claim is true then it holds in the actual world and in all possible worlds
  - So, if the colorless, odorless stuff that falls from the skies and fills the lakes, the drinkable potable stuff is  $XYZ$  then it is not water

## Neural correlates of consciousness

- Thus both materialists and dualists can be interested in discovering the neural correlates of consciousness
  - Identity theorists think that the discovery of neural correlates is a discovery of the identity of conscious states
  - Dualists think that the discovery of neural correlates is nothing more than that
  - Functionalists think that the discovery of neural correlates that fill the functional role isn't a discovery of the identity of conscious states (because they identify conscious states with the role being filled rather than with the filler of the role)



15. Intentional stance,  
indeterminacy, real patterns,  
true believers

## 16. Connectionism and symbol systems

### Plan

- Last time we briefly looked at some of the ways that consciousness has been operationalized so that we could get underway with a science of consciousness
- Then we turned to the possible worlds framework to try and understand the difference between possibility and conceivability
- This time we will start by using the possible worlds framework to try and understand correlation and numeric identity ( $p=p$ )
- Then we will return to operationalizations, and the issue of whether discovering neural correlates of consciousness shows us that consciousness is one and the same as the physical correlates

### The possible worlds framework

- Braddon-Mitchell and Jackson, and Chalmers say to think of possible worlds as ‘universes’
- Modal realists think that possible worlds are objectively existing and concrete (though spatio-temporally and causally isolated from this world)
- Other theorists think that possible worlds are best understood as something along the lines of sets of ‘maximally complete sentences / propositions’

## Possible worlds

- Fictional worlds are incomplete insofar as there are truth value gaps
  - E.g., ‘Cinderella had 10,000 hairs on her head when she put on the shoe that fit’
- Possible worlds are maximally complete insofar as there are no truth value gaps
  - Either because there is a fact about the world (on a modal realist view)
  - Or because worlds are constructed by stipulating (consistent) truth values for sentences / propositions (on the view that they are sets of sentences / propositions)

## Correlation

- x and y are correlated in the actual world if whenever x occurs y occurs (and vice versa)
- Correlations can be contingent, however
- This is just to say that while x and y might be correlated in the actual world it might be possible that they not be correlated
- This is just to say that there are possible worlds (or there is a non-contradictory set of sentences describing a situation or world) in which they aren’t correlated

## Identity

- x and y are numerically identical if there is one object (substance, property etc) rather than two
- An object (substance, property etc) is numerically identical to itself
  - $p=p$ , or brain state x = brain state x
- If there is no correlation between x and y then x and y cannot be numerically identical (one and the same object, property, etc)
- Leibniz law describes this:

- If x has a property that y lacks (or vice versa) then x does not = y

## Contingency of correlation, necessity of identity

- Correlations may be contingent
  - If x and y are actually correlated it may be possible that they not be
- Identities are necessary, however
  - If x and y are numerically identical then they are in all possible worlds
- This is because an object is always identical to itself (p necessarily = p, brain state b necessarily = brain state b)
- So, while correlation in the actual world is necessary for identity it is not sufficient
- That is to say that if there is not a correlation there cannot be an identity but if there is a correlation this is not sufficient to establish identity

## Informational value

- To say that 'p=p' or 'brain state b = brain state b' seems uninformative
- To say that 'brain state b = mental state m' seems informative, however
  - Informativeness seems to do with the state of our knowledge
  - Conceivability was relative to the state of our knowledge
- But it is that we can conceive of things turning out either way rather than it being possible for things to turn out either way
- If the identity holds in the actual world it holds in all possible worlds (it is necessary)

## Gold

- Gold = 79 protons in the nucleus of the atoms
- if the above identity claim is true then it holds in the actual world and in all possible worlds
  - If we were able to remove a proton from the nucleus of the atoms of a sample then we would have transmuted the substance from gold to something else
  - Similarly, if there was a possible world in which the yellowy malleable valuable stuff turned out to have a different number of protons in the nucleus then that substance would not be gold

## Water

- Water =  $H_2O$
- If the above identity claim is true then it holds in the actual world and in all possible worlds
  - So, if the colorless, odorless stuff that falls from the skies and fills the lakes, the drinkable potable stuff is  $XYZ$  then it is not water

## Neural correlates of consciousness

- Thus both materialists and dualists can be interested in discovering the neural correlates of consciousness
  - Identity theorists think that the discovery of neural correlates is a discovery of the identity of conscious states
  - Dualists think that the discovery of neural correlates is nothing more than that
  - Functionalists think that the discovery of neural correlates that fill the functional role isn't a discovery of the identity of conscious states (because they identify conscious states with the role being filled rather than with the filler of the role)

# 17. Connectionism: Features and problems

## Plan

- Last time we briefly looked at some of the ways that consciousness has been operationalized so that we could get underway with a science of consciousness
- Then we turned to the possible worlds framework to try and understand the difference between possibility and conceivability
- This time we will start by using the possible worlds framework to try and understand correlation and numeric identity ( $p=p$ )
- Then we will return to operationalizations, and the issue of whether discovering neural correlates of consciousness shows us that consciousness is one and the same as the physical correlates

## The possible worlds framework

- Braddon-Mitchell and Jackson, and Chalmers say to think of possible worlds as ‘universes’
- Modal realists think that possible worlds are objectively existing and concrete (though spatio-temporally and causally isolated from this world)
- Other theorists think that possible worlds are best understood as something along the lines of sets of ‘maximally complete sentences / propositions’

## Possible worlds

- Fictional worlds are incomplete insofar as there are truth value gaps
  - E.g., ‘Cinderella had 10,000 hairs on her head when she put on the shoe that fit’
- Possible worlds are maximally complete insofar as there are no truth value gaps
  - Either because there is a fact about the world (on a modal realist view)
  - Or because worlds are constructed by stipulating (consistent) truth values for sentences / propositions (on the view that they are sets of sentences / propositions)

## Correlation

- x and y are correlated in the actual world if whenever x occurs y occurs (and vice versa)
- Correlations can be contingent, however
- This is just to say that while x and y might be correlated in the actual world it might be possible that they not be correlated
- This is just to say that there are possible worlds (or there is a non-contradictory set of sentences describing a situation or world) in which they aren’t correlated

## Identity

- x and y are numerically identical if there is one object (substance, property etc) rather than two
- An object (substance, property etc) is numerically identical to itself
  - $p=p$ , or brain state x = brain state x
- If there is no correlation between x and y then x and y cannot be numerically identical (one and the same object, property, etc)
- Leibniz law describes this:

- If x has a property that y lacks (or vice versa) then x does not = y

## Contingency of correlation, necessity of identity

- Correlations may be contingent
  - If x and y are actually correlated it may be possible that they not be
- Identities are necessary, however
  - If x and y are numerically identical then they are in all possible worlds
- This is because an object is always identical to itself (p necessarily = p, brain state b necessarily = brain state b)
- So, while correlation in the actual world is necessary for identity it is not sufficient
- That is to say that if there is not a correlation there cannot be an identity but if there is a correlation this is not sufficient to establish identity

## Informational value

- To say that 'p=p' or 'brain state b = brain state b' seems uninformative
- To say that 'brain state b = mental state m' seems informative, however
  - Informativeness seems to do with the state of our knowledge
  - Conceivability was relative to the state of our knowledge
- But it is that we can conceive of things turning out either way rather than it being possible for things to turn out either way
- If the identity holds in the actual world it holds in all possible worlds (it is necessary)



## Gold

- Gold = 79 protons in the nucleus of the atoms
- if the above identity claim is true then it holds in the actual world and in all possible worlds
  - If we were able to remove a proton from the nucleus of the atoms of a sample then we would have transmuted the substance from gold to something else
  - Similarly, if there was a possible world in which the yellowy malleable valuable stuff turned out to have a different number of protons in the nucleus then that substance would not be gold

## Water

- Water =  $H_2O$
- If the above identity claim is true then it holds in the actual world and in all possible worlds
  - So, if the colorless, odorless stuff that falls from the skies and fills the lakes, the drinkable potable stuff is  $XYZ$  then it is not water

## Neural correlates of consciousness

- Thus both materialists and dualists can be interested in discovering the neural correlates of consciousness
  - Identity theorists think that the discovery of neural correlates is a discovery of the identity of conscious states
  - Dualists think that the discovery of neural correlates is nothing more than that
  - Functionalists think that the discovery of neural correlates that fill the functional role isn't a discovery of the identity of conscious states (because they identify conscious states with the role being filled rather than with the filler of the role)

## 18. Coltheart's challenge

### Plan

- Last time we briefly looked at some of the ways that consciousness has been operationalized so that we could get underway with a science of consciousness
- Then we turned to the possible worlds framework to try and understand the difference between possibility and conceivability
- This time we will start by using the possible worlds framework to try and understand correlation and numeric identity ( $p=p$ )
- Then we will return to operationalizations, and the issue of whether discovering neural correlates of consciousness shows us that consciousness is one and the same as the physical correlates

### The possible worlds framework

- Braddon-Mitchell and Jackson, and Chalmers say to think of possible worlds as 'universes'
- Modal realists think that possible worlds are objectively existing and concrete (though spatio-temporally and causally isolated from this world)
- Other theorists think that possible worlds are best understood as something along the lines of sets of 'maximally complete sentences / propositions'

### Possible worlds

- Fictional worlds are incomplete insofar as there are truth value gaps

- E.g., ‘Cinderella had 10,000 hairs on her head when she put on the shoe that fit’
- Possible worlds are maximally complete insofar as there are no truth value gaps
  - Either because there is a fact about the world (on a modal realist view)
  - Or because worlds are constructed by stipulating (consistent) truth values for sentences / propositions (on the view that they are sets of sentences / propositions)

## Correlation

- x and y are correlated in the actual world if whenever x occurs y occurs (and vice versa)
- Correlations can be contingent, however
- This is just to say that while x and y might be correlated in the actual world it might be possible that they not be correlated
- This is just to say that there are possible worlds (or there is a non-contradictory set of sentences describing a situation or world) in which they aren’t correlated

## Identity

- x and y are numerically identical if there is one object (substance, property etc) rather than two
- An object (substance, property etc) is numerically identical to itself
  - $p=p$ , or brain state x = brain state x
- If there is no correlation between x and y then x and y cannot be numerically identical (one and the same object, property, etc)
- Leibniz law describes this:
  - If x has a property that y lacks (or vice versa) then  $x \neq y$

## Contingency of correlation, necessity of identity

- Correlations may be contingent
  - If x and y are actually correlated it may be possible that they not be
- Identities are necessary, however
  - If x and y are numerically identical then they are in all possible worlds
- This is because an object is always identical to itself (p necessarily = p, brain state b necessarily = brain state b)
- So, while correlation in the actual world is necessary for identity it is not sufficient
- That is to say that if there is not a correlation there cannot be an identity but if there is a correlation this is not sufficient to establish identity

## Informational value

- To say that 'p=p' or 'brain state b = brain state b' seems uninformative
- To say that 'brain state b = mental state m' seems informative, however
  - Informativeness seems to do with the state of our knowledge
  - Conceivability was relative to the state of our knowledge
- But it is that we can conceive of things turning out either way rather than it being possible for things to turn out either way
- If the identity holds in the actual world it holds in all possible worlds (it is necessary)

## Gold

- Gold = 79 protons in the nucleus of the atoms
- if the above identity claim is true then it holds in the actual world and in all possible worlds

- If we were able to remove a proton from the nucleus of the atoms of a sample then we would have transmuted the substance from gold to something else
- Similarly, if there was a possible world in which the yellowy malleable valuable stuff turned out to have a different number of protons in the nucleus then that substance would not be gold

## Water

- Water =  $H_2O$
- If the above identity claim is true then it holds in the actual world and in all possible worlds
  - So, if the colorless, odorless stuff that falls from the skies and fills the lakes, the drinkable potable stuff is  $XYZ$  then it is not water

## Neural correlates of consciousness

- Thus both materialists and dualists can be interested in discovering the neural correlates of consciousness
  - Identity theorists think that the discovery of neural correlates is a discovery of the identity of conscious states
  - Dualists think that the discovery of neural correlates is nothing more than that
  - Functionalists think that the discovery of neural correlates that fill the functional role isn't a discovery of the identity of conscious states (because they identify conscious states with the role being filled rather than with the filler of the role)