

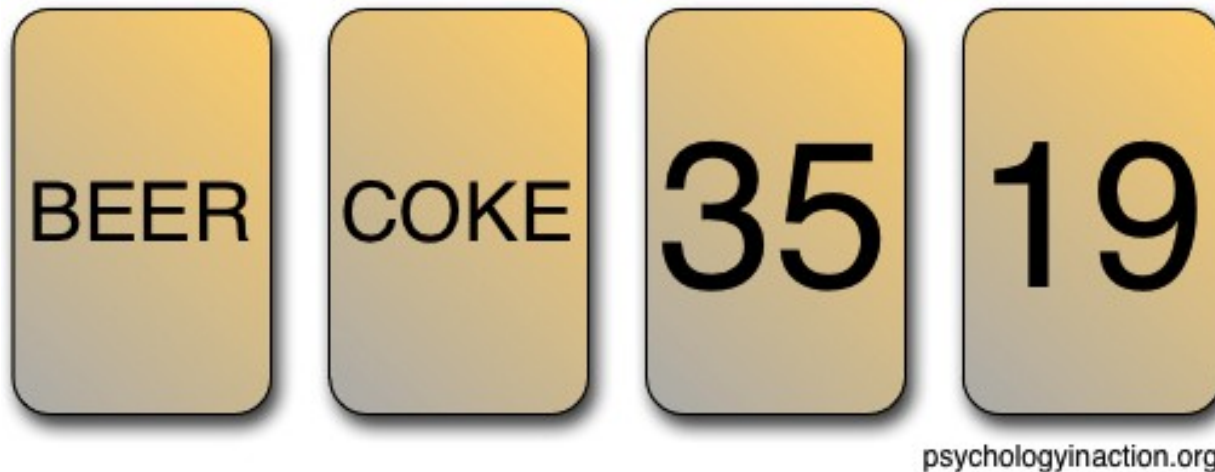


psychologyinaction.org

Each card has a letter on one side and a number on the other.

Proposed Rule: **If a card has a vowel on one side, then it has an even number on the other side.**

Which card(s) do you need to turn over in order to determine if the rule is true or false?



These four cards represent patrons in a bar, and each card has their drink on one side and their age in years on the other.

**Proposed Rule: If a patron is drinking a beer, then they must be 21 years or older.**

Which card(s) do you need to turn over in order to determine if the rule is being followed?

# Solutions (2024)

- Puzzle 1: **A 7**
  - Fully correct: 1/12 (**8%**)
  - **A 4** 6/12; **A** 2/12
  - Individual card-wise: 15/25 (60%)
- Puzzle 2: **Beer 19**
  - Fully correct: 2/14 (**14%**)
  - **Beer 35** 5/14; **Beer** 3/14
  - Individual card-wise: 18/26 (69%)

# Solutions (2023)

- Puzzle 1: **A 7**
  - Fully correct: 11/37 (**30%**)
  - **A 4** 10/37; **A** 6/37
  - Individual card-wise: 50/73 (68%)
- Puzzle 2: **Beer 19**
  - Fully correct: 26/38 (**68%**)
  - **Beer 35** 6/38; **Beer** 2/38
  - Individual card-wise: 63/73 (86%)

# Computation and Cognition

ELL457/HSL622

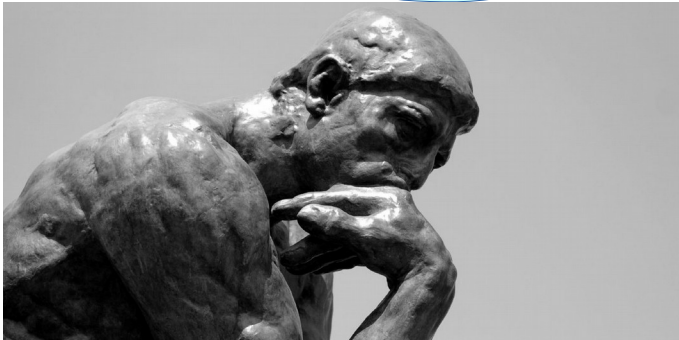
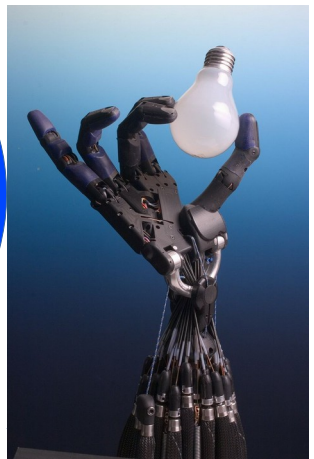
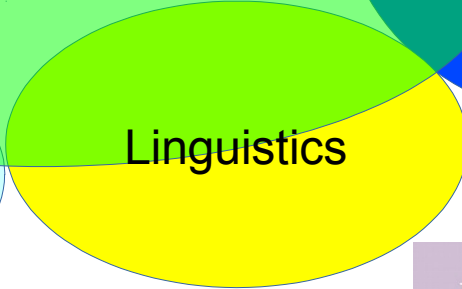
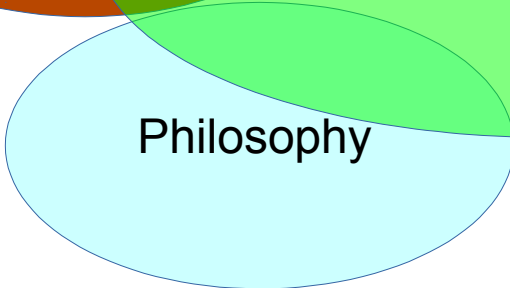
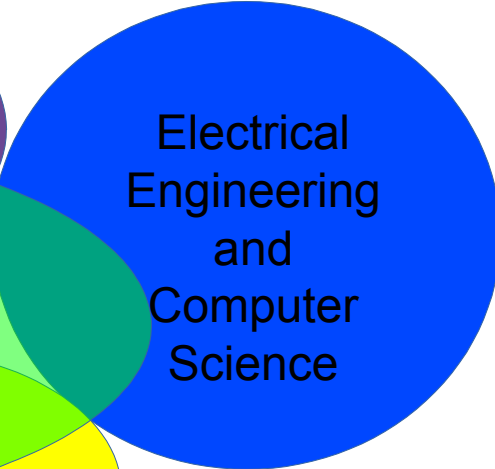
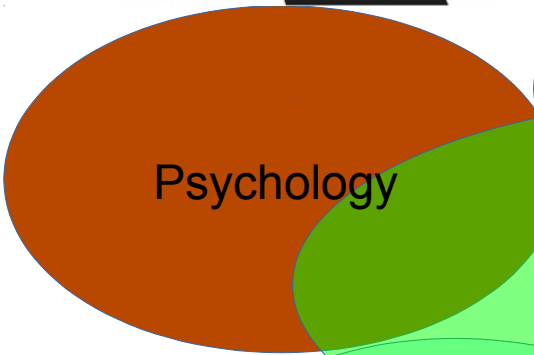
[Slides by Santanu Chaudhury, Hiranmay Ghosh, and  
Sumeet Agarwal]

[Material sourced from Friedenbergs and Silverman, 2006]

# Introduction: Philosophical & Psychological Perspective

*“The sciences have developed in an order the reverse of what might have been expected. What was most remote from ourselves was first brought under the domain of law, and then, gradually, what was nearer: first the heavens, next the earth, then animal and vegetable life, then the human body, and last of all (as yet very imperfectly) the human mind.”*

—Bertrand Russell, 1935





# Cognitive Science: Mind as Computer

- **Representation**

- Concepts

- Propositions

- Rules

- Analogies

- Digital (symbols)

- Analog (images)

- **Transformation/Processing/Computation**

# Intentionality

- **Intention:** A thing intended (aimed/planned for); purpose



- **Intension:** The internal content of a concept



So **intentionality** is about **symbols** and their ability to **represent** other things

Is this a *defining characteristic* of mind / the mental? Can bodies / physical objects have intentionality in themselves?

## Example: Analogical reasoning [Duncker '45]

Suppose you are a doctor faced with a patient who has a malignant tumor in his stomach. To operate on the patient is impossible, but unless the tumor is destroyed, the patient will die. A kind of ray, at a sufficiently high intensity, can destroy the tumor. Unfortunately, at this intensity the healthy tissue that the rays pass through on the way to the tumor will also be destroyed. At lower intensities the rays are harmless to healthy tissue, but will not affect the tumor.

How can the rays be used to destroy the tumor without injuring the healthy tissue?

## Analogy

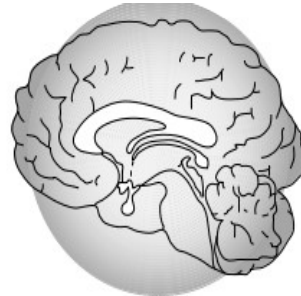
A small country was ruled from a strong fortress by a dictator. The fortress was situated in the middle of the country, surrounded by farms and villages. Many roads led to the fortress through the countryside. A rebel general vowed to capture the fortress. The general knew that an attack by his entire army would capture the fortress. He gathered his army at the head of one of the roads, ready to launch a full-scale direct attack. However, the general then learned that the dictator had planted mines on each of the roads. The mines were set so that small bodies of men could pass over them safely, since the dictator needed to move his own troops and workers to and from the fortress. However, any large force would detonate the mines. Not only would this blow up the road, but it would also destroy many neighboring villages. It seemed impossible to capture the fortress. However, the general devised a simple plan. He divided his army into small groups and dispatched each group to the head of a different road. When all was ready, he gave the signal and each group marched down a different road. Each group continued down its road to the fortress, so that the entire army arrived together at the fortress at the same time. In this way, the general captured the fortress and overthrew the dictator.

# Levels of Computation

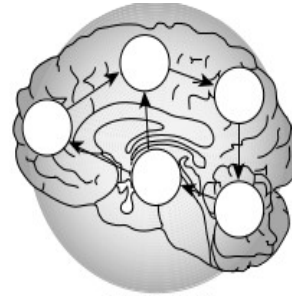
Any given information process, mental or artificial information-processing event can be evaluated on at least three different levels (Marr, 1982).

- The highest or most abstract level of analysis is the **computational level**. Tasks: What is this problem? Why is this process here in the first place?
- **Algorithm**, a formal procedure or system that acts on informational representations
- **Implementational level**: What is the information processor made of?

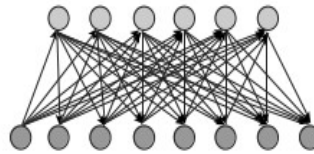
# Structural Levels in Neural Information Processing



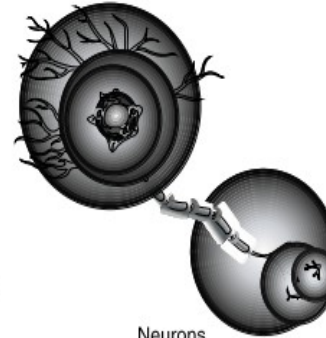
Brain



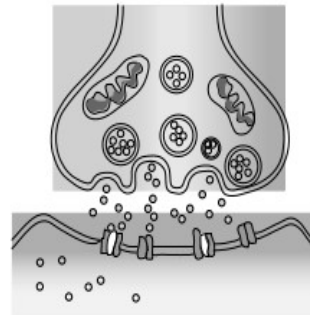
Brain regions



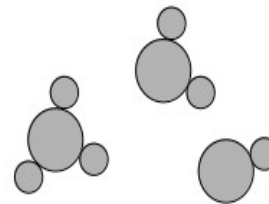
Neural networks



Neurons



Synapses



Molecules

# Daniel Dennett's '3 stances'

- Levels of (increasing) *abstraction* at which to explain or describe the behaviour of objects/systems/entities:
  - **Physical stance**  
(Physics, Chemistry: *Why does an apple fall from a tree to the ground?*)
  - **Design stance**  
(Biology, Engineering: *Why does the heart pump blood to the whole body?*)
  - **Intentional stance**  
(Psychology, Social Sciences: *Why do people invest so much time in preparing for IIT JEE?*)

# *The intentional stance* as predictive strategy

- How can we know about someone's *beliefs*? Are they objective physical states of brain/body? Or can they only be known subjectively, by interpretation?



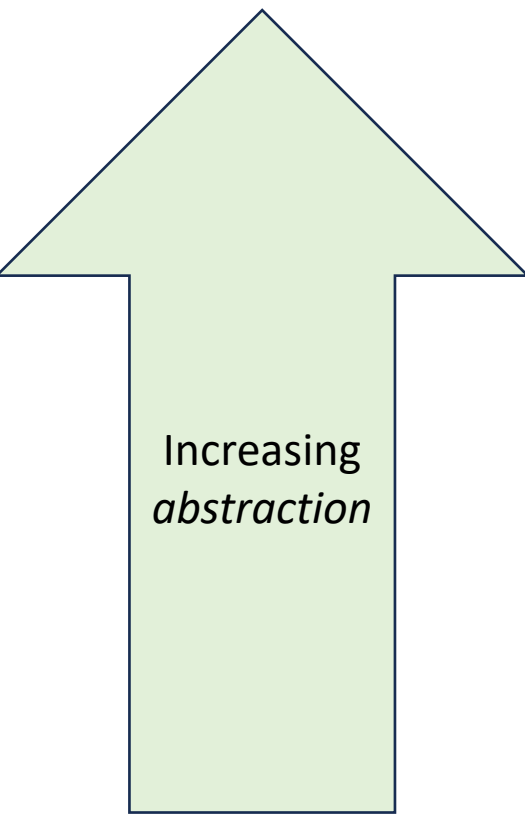
# *The intentional stance* as predictive strategy

- How can we know about someone's *beliefs*? Are they objective physical states of brain/body? Or can they only be known subjectively, by interpretation?
- The **intentional** strategy or stance is to treat the object of study as having beliefs, desires, and other mental states exhibiting **intentionality**; and as a rational **agent** with behaviour *predictable* from such states (crudely: the agent will act to fulfil its desires, in ways suggested by its beliefs)

# *The intentional stance* as predictive strategy

- How can we know about someone's *beliefs*? Are they objective physical states of brain/body? Or can they only be known subjectively, by interpretation?
- The intentional strategy or stance is to treat the object of study as having beliefs, desires, and other mental states exhibiting intentionality; and as a rational **agent** with behaviour *predictable* from such states (crudely: the agent will act to fulfil its desires, in ways suggested by its beliefs)
- If this strategy *works* in predicting the behaviour of the object/system, then the system is a *believer*, or an **intentional system**

# How does **intentionality** connect to **computation**?



Marr's tri-level hypothesis	Dennett's 3 stances	Basic gloss of what is being captured/described/explained about a process or behaviour
Functional/Computational level	Intentional stance	WHY it exists or occurs; what <i>purpose</i> is it serving for the system or <b>agent</b>
Algorithmic level	Design stance	HOW it is realised at the level of a formal/conceptual ( <b>representational</b> ) model or structure
Implementational level	Physical stance	HOW it is realised at the ( <b>grounded</b> ) level of material/physical objects and properties and phenomena

But take this correspondence with a pinch of salt: one of the big debates in the course is precisely around **whether computers can have intentionality or agency at all!**

# Classical and Connectionist Views of Computation

- In the classical view, knowledge is represented locally, in the form of symbols.
- In the connectionist view knowledge is represented as a pattern of activation or weights that is distributed throughout a network.
- The classical view has processing occurring in discrete stages.
- In connectionism, processing occurs in parallel through the simultaneous activation of nodes.

# Mind: Perspectives

- Philosophical
  - Philosophers of mind narrow their focus to specific problems concerning the nature and the characteristics of mind.
  - They ask questions like: What is mind? How do we come to know things? How is mental knowledge organized?
  - The primary method of philosophical inquiry is reasoning, both deductive and inductive.

# Mind: Perspectives

- Psychological
  - Psychologists apply the scientific method to both mind and behavior.
  - Attempt to understand internal mental phenomena, such as thoughts
  - Studies the external behaviors that these internal phenomena can give rise to.

# Mind: Perspectives

- Cognitive
  - Emphasis on the study of internal mental operations.
  - Adopted the computer as a metaphor for mind
  - Described mental functioning in terms of representation and computation.
  - Mind, like a computer, could be understood in terms of information processing.
  - Modularity
    - Modules are functionally independent mental units that receive inputs from other modules, perform a specific processing task, and pass the results of their computation onto yet additional modules.
  - Experimental Method and Computational Modelling

# Mind: Perspectives

- Neuroscience
  - Describe the biological “hard- ware” upon which mental “software” supposedly runs.
  - Provides multi-level analysis
  - Study the cell biology of individual neurons and of neuron-to-neuron synaptic transmission, the patterns of activity in local cell populations, and the interrelations of larger brain areas.



# Mind: Perspectives

- Evolutionary psychology
  - Modular approach to mind.
    - Modules correspond to “favoured” cognitive capacities that were used by ancestors successful at solving certain problems.
  - Evolutionary theories have been proposed to account for experimental results from categorization to memory, to logical and probabilistic reasoning, language, and cognitive differences between the sexes.

# **Philosophical Perspectives**

# Mind-Body Problem

- Addresses how psychological or mental properties are related to physical properties.
- The debate stems from a fundamental conception about what the mind is:
  1. Equivalent to the brain; material and physical; made up of substances that we can measure and understand.
  2. Something more, because
    - we can't equate our subjective conscious experiences, such as beliefs, desires, and thoughts, with something as mundane as the brain.
    - the mind is nonphysical and consists of something resembling a soul or spirit. The mind as a nonphysical entity inhabiting the brain or other physical entity is sometimes called "the ghost in the machine."

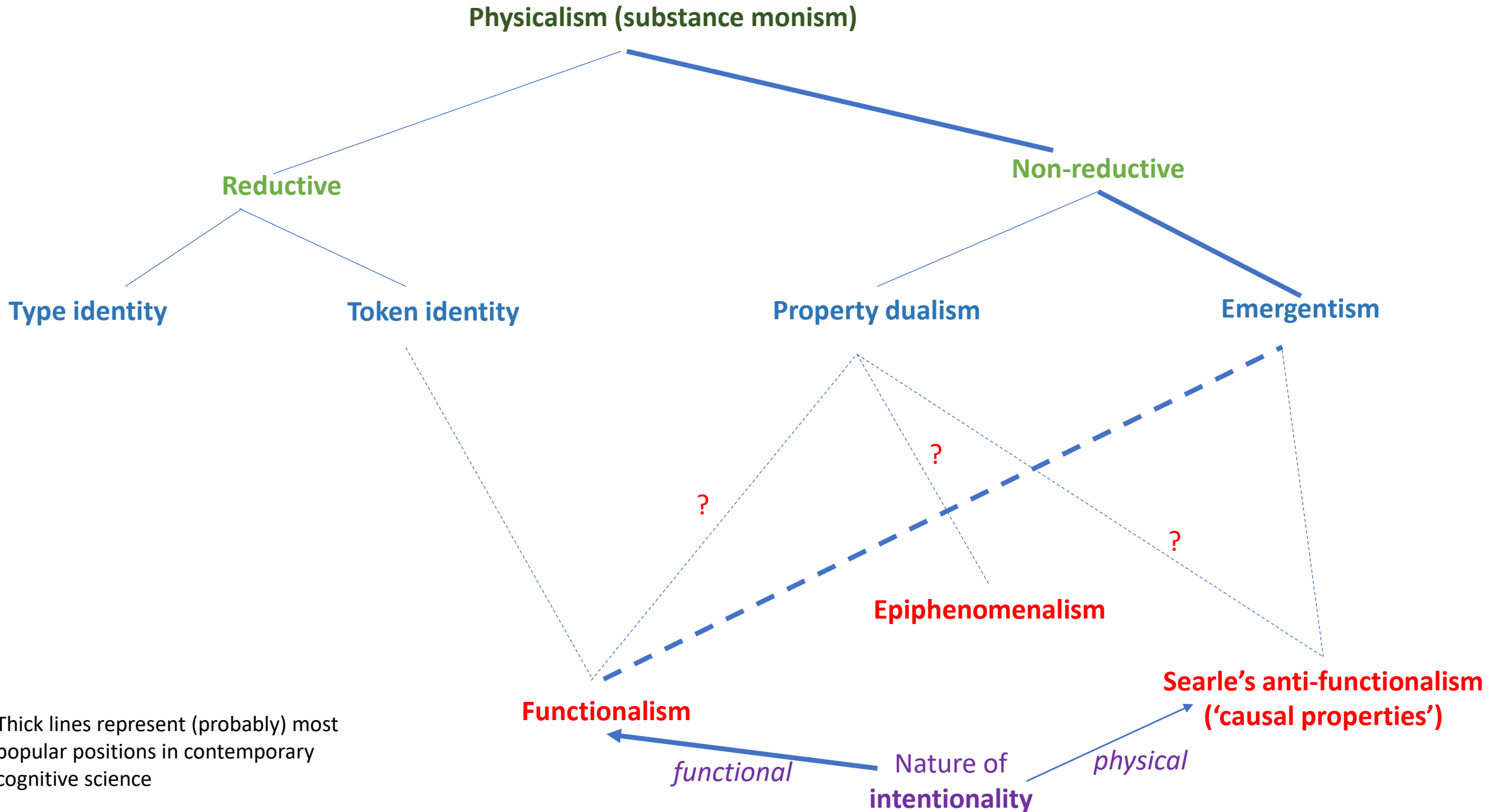
# Mind-Body Problem

- A second and more specific question:
  - If we assume that there are two such entities, then what is the causal relationship between them?
  - Does the mind control the body or does the body control the mind?

**Table 2.1** Different interpretations of the mind-body debate and the schools of thought associated with each

<i>Class of Theory</i>	<i>Name of Theory</i>	<i>Physical Universe</i>	<i>Causal Direction</i>	<i>Mental Universe</i>
Monism	Idealism/Solipsism	None	No causality	Mind
	Physicalism	Body	No causality	None
Dualism	Classical Dualism	Body	←	Mind
	Parallelism	Body	No causality	Mind
	Epiphenomenalism	Body	→	Mind
	Interactionism	Body	↔	Mind

Source: Adapted from Kitzis, S. N. (2002). *Mind and meaning: A flight of imagination, a voyage of discovery*. Pacific Grove, CA: Wadsworth Thompson.



# In (Mahayana) Buddhist thought (very roughly)

- From **metaphysical/ontological** perspective: non-essentialism, neither mind nor body foundational, mutual *dependence* of all phenomena, ontological *emptiness* <*Madhyamaka*>
- From **epistemological + phenomenological** perspective: subjective *idealism* <*Yogacara/Cittamatra*>

# Recap

- What are the two key ingredients of the cognitive approach to modelling the mind?
- What are the two main types of knowledge typically represented in cognitive systems; how do they differ?
- What are Marr's three levels of computation? What aspects of computation / information processing are they meant to capture?
- What is the Mind-Body problem? What are the prominent schools of thought on this problem in the history of philosophy?



# Functionalism

## Classifications of two kinds

– **Physical kinds** are identified by their material composition only:

jellyfish vs. carpets

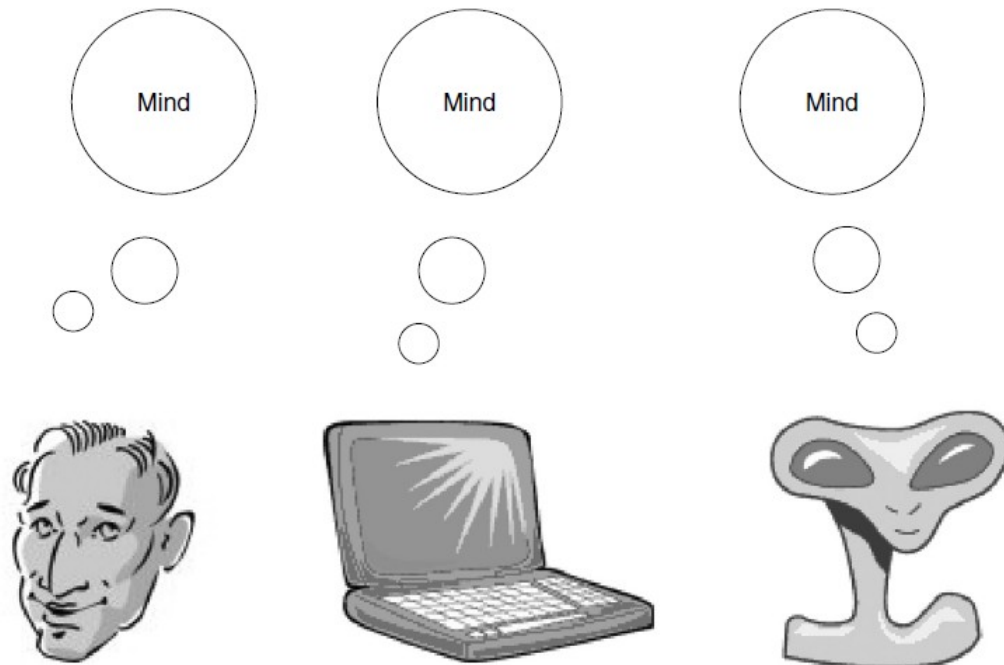
– **Functional kinds** are distinguished by their actions or tendencies:

all automobiles -- transport goods and people

# Functionalism (contd.)

- Mind as a physical thing would be same as Brain
- But can't computers develop minds? Can't there be alien species with minds with an entirely different chemical composition?
- Minds as *functional kinds*: defined by the sorts of processes they carry out rather than the stuff they're made of

- According to functionalism
  - Mental states are not just physical states, but also the functioning or operation of those physical states.
  - A mind could conceivably be implemented in any physical system, artificial or natural, capable of supporting the appropriate computation.



# Limitations of functionalism

- No evidence yet of Computers/Robots having mind of their own
- It cannot account for the felt or experienced character of mental states—a phenomenon known as **qualia**:
  - What it is like to feel **hungry**
  - to be **angry**
  - to see the color **red**
- It would seem that these kinds of experiences cannot be replicated as purely functional processes. A machine could be programmed to “see” the color **red**, even mimicking the same human functional process; but could it have the same experience of what it is like to see **red** that a person has?

# Subjectivity?

- Two individuals having the same conscious experience often do not experience it subjectively in the same way:
  - If asked to point out on a colour spectrum what pure **green** looks like, one person may select a **yellow-green**, another a **blue-green**;
  - Even though the functional operations of their respective brains as they view the colour are approximately equivalent.

# Determinism vs. Free Will

- **Determinism:** The view that all physical events are caused or determined by the sum total of all prior events and nothing else  
Can replicate and predict without ambiguity
- **Free Will:** Decisions are made autonomously and not under the influence of any preceding causal factors  
Complete free will violates causality
- Compatibilism? Choice and causes/constraints can co-exist

# Knowledge Acquisition: Nature vs. Nurture

- *Nature*, in this context, refers to traits that are genetically or biologically determined:
  - These are coded for in our genes and so are “hardwired,” meaning they are present at birth or appear at a specific time during development
- The term *nurture* refers to traits that are learned through experience and interaction with the environment
- In linguistics: Universal Grammar and Language Faculty as *nature*, to address poverty of stimulus problem for language acquisition

# Perspectives on Learning

- Nativism: a significant body of knowledge is innate or “built into” an organism; favours nature over nurture
- Rationalism: additionally emphasizes the existence of innate reasoning powers – including certain logical propositions, such as knowing that something cannot exist and not exist at the same time
- Empiricism: Knowledge as acquired through experience, interaction with an environment, learning; favours nurture over nature



# Perspectives on Learning

- Research supporting the notion that some forms of procedural knowledge are innate
  - Reflexes: physiology clear and established
  - Smell Preference
- Evolutionary Psychologists attributes such capacities to generations of selection pressures acting on a species
  - These pressures promote the development of adaptive (survival-related) cognitive abilities
  - Evolutionary psychologists further argue that these innate abilities are *domain-specific*

# Consciousness

- Individual subjective awareness of mental states
  - include sensation, perception, visual images, conscious thought processes, emotions, and sense of self
- Subjectively, our consciousness seems to be unitary
  - One recognizes himself or herself to be one person, experiencing things in the present moment
- When one studies the brain, though, one finds that there is no single place or even time where consciousness seems to happen: instead the brain in action is a case of activity going on all over the place.
  - Furthermore, the brain may even be processing different aspects of a single experience at different times

# Mind (re-visited)

- The **phenomenal concept of mind** is essentially the idea of mind as a conscious experience
  - Mental states in this view need to be explained in terms of how they feel
- The **psychological concept of mind** sees mental states only in terms of how they cause and explain behaviour
  - Mind is characterized by what it does—how it feels is irrelevant

# Consciousness

- The problem here is that science can only provide an objective account of a phenomenon and consciousness is an inherently subjective state
- As organisms capable of supporting consciousness, we can introspect and analyze what it is like to have or to experience a mental state
- Gap between an objective and a subjective description of mental phenomena goes by the name of the *explanatory gap*

# Mind as Emergent Memory

- Consciousness is an **emergent property** of the brain, i.e., realized through the interaction of the system's parts
  - If we have a given emergent system S, made up of elements a, b, c, and so on, then the features of S may not be the same as the features of a, b, c, and so on. This is because the features of S arise from the causal interactions of the parts
- Consciousness is a property of the brain but not of its parts. If we take neurons to be the relative parts, then they have their own properties.

# Neuroscience View

- Consciousness results from the coordinated activity of a population of neurons in the brain.
- Popper and Eccles (1981): emergent property of a large number of interacting neurons.
- Crick and Koch (1995): neurons specifically devoted to producing consciousness, located throughout the cortex and in associated areas.
- *Recurrence*: allows for feedback and learning. Recurrent activity in a network may sustain information over time and be the basis for conscious mental awareness.

# AI and Consciousness

- The **strong AI** view asserts that consciousness can arise from a purely physical process
  - As we create machines with greater complexity and computational power, we will see consciousness emerge in them
- Proponents of **weak AI** claim that consciousness is itself either not a physical process and so can never be reproduced; Or
- Is a physical process but such a complex one that we will never be able to duplicate it artificially

你好吗？



我是伏良！



# Psychological approaches

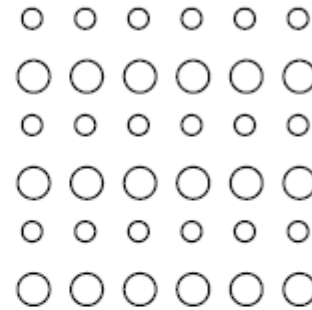
- Voluntarism
- Structuralism
- Functionalism
- Gestalt Movement
- Psychoanalysis
- Behaviourism
- **Cognitive Psychology**

# Gestalt and Pragnanz

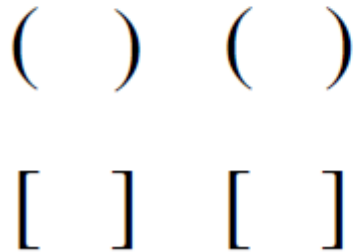
(a)



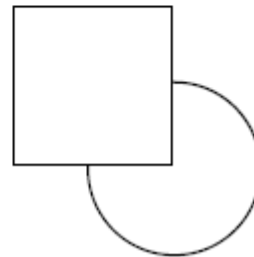
(b)

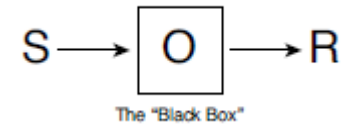
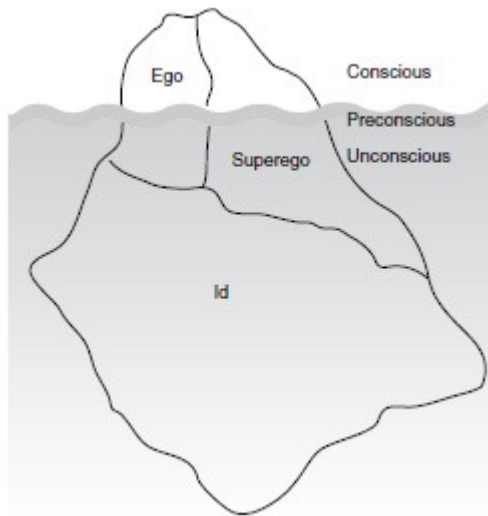


(c)



(d)





Behaviourism: Mind as black box

Psychoanalysis: Freud's iceberg model