LEVELS IN COGNITIVE SCIENCE

Phil/Psych 256

Chris Eliasmith
Theories of Mind

- Dualism (‘Cartesian interactionist dualism’)
  - Minds and bodies are two different kinds of substances
    - minds are essentially thinking
    - bodies are essentially spatial (extended)
  - They are independent, so each can exist without the other.
  - Classical/default view (philosophical, religious, scientific, political, etc.)
  - Some contemporary thinkers argue for this by asking:
    - How can 3rd person information (science) possibly explain 1st person experience?
Theories of Mind (cont.)

“Mind-body Problem”

– How do such different things have causal interactions (as they clearly seem to)?

Materialism:

– There is only one kind of substance, matter (i.e. 'bodies').

– Majority view in the cognitive sciences
Materialism (cont.)

Two Versions:

1. Identity theory: Churchland makes this position misleadingly acceptable by defining it as 'mental states [are] in fact identical to states of the brain'. Two possibilities:

   - Mental tokens are brain tokens [aka token identity theory]: which is standardly agreed to since it’s the same as materialism; and

   - Mental types are identical to brain types [aka type identity theory]: this is contentious.
Token vs. Type Identity Theory

Token Identity

Type Identity

M1
M2
M3
M4

N1
N2
N3
N4

Pain
Cat
Mad

N43
N121
N79
Mental state

- Pain
- Cat
- Mad

Neural state

- N43
- N121
- N79

Token Identity

Type Identity
2. Theory dualism (aka property dualism): Neuroscientific and psychological theories (or properties) are independent (Fodor; standard view). Reasons:

1. Neuroscience is hard

2. Psychological states are *multiply realizable*

3. Psychological explanations are intentional but neuroscientific explanations are causal.

1. **Computational** (aka 'semantic' or 'content'): What is the goal of the computation, why is it appropriate, and what is the logic of the strategy by which it can be carried out?

2. **Algorithmic** (aka 'syntactic' or 'form'): How can this computational theory be implemented? In particular, what is the representation for the input and output, and what is the algorithm for the transformation?

3. **Implementational** (aka 'physical' or 'medium'): How can the representation and algorithm be realized physically?
Computation

Independent from

Algorithm

Independent from

Implementation
Levels of Mind

- Why do C&S think this approach to mind is wrong?
  - Science doesn’t work that way.

- What would it mean to say psychology can be 'reduced to' neuroscience?

- C&S suggest this should mean integration, co-evolution and mutual constraint.
  - Examples: light, electricity, heat, life, gene, etc.
  - C&S think data at all 'levels' needs to be accounted for and constrain what goes on at other levels

- Critics generally think it means a strict translation from one discourse to another.
  - They are worried that those like C&S want to eliminate folk psychological terms (this is not surprising)
  - C&S say ‘the reduction of theories does not mean that the reduced phenomena somehow disappear or are discredited.’ Careful to talk of ‘phenomena’ instead of ‘terms’
Levels of Mind

What do those considerations mean for theory dualism?

1. Neuroscience is hard, but we've made lots of progress that seems relevant to the 'higher' level (e.g. how networks learn, how information is stored, what neural representations are like).

2. Multiple realizability doesn't mean reduction won't occur:
   - It just means that the reduction will be domain relative (i.e. relative to a species). This would be a great theory on the way to a more general 'all of psychology' kind of theory. For example, if we learned how to explain human cognition in terms of brain states, this doesn't mean such a system couldn't be made of silicon, but it goes a long way to understanding which properties will have to be maintained in a silicon version and which won't.
Against RTM

3. The standard meaning/intentionality story is RTM.
   Problems:
   - Hundred step rule
   - Parallel architecture, not Von Neumann machine
   - Storage and processing in the same location
   - Some tasks are more natural for some architectures
   - Hardware/software analogy fails because of learning
   - Levels in computers and brains don’t match up
   - Can not explain nonverbal animals/infraverbal humans

   C&S don't have a clear alternative (something like analyzing dynamic, high-dimensional vectors in certain abstract spaces?)

   Instead, they point out that it is generally unsafe to assume that the structure of a phenomena is similar to the structure of the cause of the phenomena (e.g. homuncular theory of reproduction).
Critique of Marr’s Levels

- Recall NETtalk (Sejnowski and Rosenberg):
  - NETtalk is constrained by information from both higher and lower levels.
  - Higher levels: information about phonemic structure is used to encode the output and determine the function being computed.
  - Lower levels: the network is constrained to use certain kinds of processing (i.e., parallel simple processors).

- Shape from shading:
  - Train a network to recognize curvature regardless of lighting, etc. Get receptive fields like V1 simple cells
  - Marr’s levels of analysis don’t map neatly onto understanding a biological structure like the brain because both models use similar constraints and approaches, but they should be mapped to different levels of analysis of the brain

- Levels are highly interdependent (e.g., brain/implementation knowledge contributes to algorithms (network design), and characterization of the problem (computational level))