

**24.09 Minds and Machines**  
**spring 2006**

- evaluations, final exam questions on Monday
- today: Levine and Stoljar
- more Stoljar on Monday

---

**type-A and type-B materialism revisited**

- “epistemological” premise: zombies are conceivable, Mary can’t know what it’s like to see red, etc.
- from which the “metaphysical” conclusion that physicalism is false is supposed to follow
- the type-A materialist denies the epistemological premise
- the type-B materialist accepts the premise, but denies that the conclusion follows

---

**resisting the knowledge argument**

1. imprisoned Mary knows all the physical facts  
   hence: type-B
2. if physicalism is true, Mary (before her release) knows all the facts
3. after her release, Mary learns something—something she couldn’t have known while imprisoned  
   hence (from 3, 4):
4. if Mary learns something, she learns a fact  
   hence (from 3, 4):
5. Mary learns a fact  
   hence (from 2, 5):
6. physicalism is false

---

**Minds and Machines**  
**spring 2006**

- Levine and the explanatory gap

---

**“Materialism and qualia: the explanatory gap”**

- Levine’s position is similar to Nagel’s: physicalism is (or may well be) true, but we don’t understand how it’s true
- Levine arrives at this conclusion by transforming Kripke’s argument from a metaphysical one into an epistemological one
- Levine would accuse Jackson of “transforming an epistemological argument into a metaphysical one”

---

**“Materialism and qualia: the explanatory gap”**

- Kripke argues that pain ≠ c-fibers firing (a “metaphysical” conclusion)
- Levine thinks this argument does not work, but a related argument shows that “psycho-physical identity statements leave a significant explanatory gap”
- although Levine’s argument does not show that physicalism is false, it does “constitute a problem for materialism”
• according to Levine, psychophysical identities (e.g. ‘pain=c-fibers firing’) are “gappy”, unlike others (e.g. ‘heat=mke’)
• he thinks the gappiness of psychophysical identities is due to the absence of certain “conceptual” truths
• …and that the absence of such truths is shown by Kripke’s argument against the identity theory

Levine’s objection: deny (1)

1. if I can clearly and distinctly conceive a proposition $p$ to be true, then $p$ is possible

“Since epistemological possibility [i.e. clear and distinct conceivability] is not sufficient for metaphysical possibility, the fact that what is intuitively contingent turns out to be metaphysically necessary should not bother us terribly. It’s to be expected.”

o one might think this reply is a bit dismissive—surely conceivability is a good guide to possibility, so why does it fail in this case? (see Nagel’s note 11 and the paper by Hill for one suggestion)

explanatory and “gappy” identities

1) pain = c-fibers firing
2) heat = molecular kinetic energy
3) pain = functional state F

“Statement (2), I want to say, expresses an identity that is fully explanatory, with nothing crucial left out. On the other hand, statements (1) and (3) do seem to leave something crucial unexplained, there is a “gap” in the explanatory import of such statements.”

explanation and reduction

“The basic idea is that a reduction should explain what is reduced, and the way we tell whether this has been accomplished is to see whether the phenomenon to be reduced is epistemically necessitated by the reducing phenomenon…I claim we have this with the chemical theory of water but not with a physical or functional theory of qualia” (“On leaving out what it’s like”)
reminder from the philosophical toolkit:

- **a priori** and **a posteriori** truths (true propositions)
- an **a priori** truth is one knowable independently of experience
- an **a posteriori** truth is one knowable only on the basis of experience

(not implausible) examples of **a priori** truths

- mathematical truths: there is no highest prime, there are uncomputable functions, e is irrational,…
- logical truths: either it’s snowing or it isn’t, if Fred is rich and unhappy then he’s unhappy,…
- “analytic” truths: bachelors are unmarried, vixens are foxes, if something is red it’s colored,…
- these are all examples of necessary truths
- are the categories of the a priori and the necessary the same?
- Kripke argued that they aren’t; in fact, we have already seen some examples of necessary **a posteriori** truths (e.g., water=H2O)

an explanatory reduction of water to H2O

1) water = the clear odorless liquid that falls as rain and flows in the rivers and streams (e.g.) [a priori or “conceptual” truth]
2) H2O = the clear odorless liquid that falls as rain and flows in the rivers and streams [a posteriori or empirical truth]; hence
3) water = H2O

- this is both a justification of (3) and an explanation of it

is there an explanatory reduction of pain to c-fiber firing?

- pain = the state that is caused by bodily damage, causes avoidance behavior, etc. [a priori or “conceptual” truth??]
- c-fiber firing = the state that is caused by bodily damage, causes avoidance behavior, etc. [a posteriori or empirical truth]
- hence
- pain = c-fiber firing
- “there is more to our concept of pain than its causal role, there is its qualitative character, how it feels…” (so the first premise is not a priori)

summing up

- in order to explain why pain = c-fiber firing, we need to deduce this identity from conceptual truths plus empirical claims about c-fibers
- Kripke’s argument illustrates why we can’t do this: if we could, then c-fibers firing without pain (or vice versa) wouldn’t be conceivable
- since the missing explanation is also the only way we can justify (or justify more-or-less conclusively) that pain = c-fiber firing, this identity claim is “epistemologically inaccessible…a very undesirable consequence of materialism” (p. 359)

Levine accepts the **explanatory argument** (see Chalmers, 248)

1) physical accounts explain at most structure and function
2) explaining structure and function does not suffice to explain consciousness; so
3) no physical account can explain consciousness

n.b. a few more slides on Levine will be posted on the website
but: is it really a priori that water = the clear odorless liquid…? (see Block and Stalnaker)

1) water = the clear odorless liquid that falls as rain and flows in the rivers and streams (e.g.) [a priori or “conceptual” truth?]
2) if it isn’t, then Levine’s account of the difference between ‘water = H₂O’ and ‘pain = c-fiber firing’ is mistaken
3) another example: is it a priori that pencils are artifacts?
4) arguably not: pencils could turn out to be organism (see Putnam, 591)

= or =?

so how can identities be explained?

“identities don’t have explanations (though of course there are explanations of how the two terms can denote the same thing)” (Block and Stalnaker, 382)

Mark Twain
born in Florida, Missouri, on November 30, 1835…died in Connecticut on April 21, 1910

Samuel Clemens
born in Florida, Missouri, on November 30, 1835…died in Connecticut on April 21, 1910

an inconsistent tetrad

1) if physicalism is true, a priori physicalism is true
2) a priori physicalism is false
3) if physicalism is false, epiphenomenalism is true
4) epiphenomenalism is false
argument for (1)

• out of type-A and type-B materialism, the former is much more plausible than the latter
  • see Chalmers against type-B materialism
  • hence: if physicalism (materialism) is true, a priori physicalism (type-A materialism) is true

argument for (2)

• both the conceivability argument and the knowledge argument show that “knowledge of every physical property a person has cannot by itself suffice to know which qualia, if any, his or her experiences instantiate”
  • hence a priori physicalism (type-A materialism) is false

argument for (3)

• type-E dualism is much more plausible than type-D dualism
  • hence if physicalism is false (and so dualism is true), type-E dualism (and so epiphenomenalism) is true

argument for (4)

• obviously qualia are causally efficacious with respect to physical events, otherwise we wouldn’t have any reason to think that there are any qualia
  • hence epiphenomenalism is false

(1)-(4) are individually plausible, but at least one must be false

1) if physicalism is true, a priori physicalism is true
2) a priori physicalism is false
3) if physicalism is false, epiphenomenalism is true
4) epiphenomenalism is false

t-physicalism and o-physicalism

• P is a t-physical property iff P is (i) the sort of property that a (true) physical theory tells us about or (ii) a property which metaphysically supervenes on properties that satisfy (i)
  • so: having +ve charge and having mass are t-physical properties (by (i); the theories of electromagnetism/gravity)
  • either having +ve charge or being green is a t-physical property (by (ii); necessarily, if x is positively charged, x is either positively charged or green)
• P is an o-physical property iff P is (i) the sort of property required by a complete account of the nature of paradigmatic physical objects or (ii) a property which metaphysically supervenes on properties that satisfy (i)
• so: having +ve charge and having mass are o-physical properties (by (i): needed for a complete account of sticks and stones)
• either having +ve charge or being green is an o-physical property (by (ii): necessarily, if x is positively charged, x is either positively charged or green)

something new from the philosophical toolkit:

**categorical** properties
• a special kind of property
• not a dispositional property
• the kind of property the possession of which explains the possession of a dispositional property
• in the case of a fragile vase, a property “whose instantiation makes it the case that the vase is fragile”

**thesis 1**: physical theory tells us only about dispositional properties
• (roughly) to be positively charged is to be disposed to be attracted by electrons, repelled by protons, etc.
• (roughly) to have mass is to be disposed to warp space-time
• let’s grant thesis 1 for the sake of the argument

**thesis 2**: if an object x has a dispositional property D, x has a categorical property C that explains why x has the dispositional property
• why would this vase break when struck (i.e. why is it fragile)?
• there must be an explanation, and if the explanation is in terms of more dispositional properties, we will need an explanation of why the vase has these dispositional properties
• so this chain of explanations must bottom out in an explanation in terms of categorical properties
• let’s grant thesis 2 for the sake of the argument

**conclusion from theses 1 and 2**
• paradigmatic objects have categorical properties (thesis 2)
• these categorical properties are not t-physical properties (thesis 1)
• but they are o-physical properties (by the definition of ‘o-physical’)
• so, some o-physical properties are not t-physical properties
Minds and Machines
spring 2006

* read Stoljar, McGinn