

# Minds and Machines

spring 2003

Content:  
psychosemantics, contd.

# preliminaries

- this week: anomalous monism and mental causation
- next week (and following): consciousness

# a different approach: “True believers”

- Dretske gives a *reductive* account of thought—an account that does not help itself to any mental ingredients
- Dennett, like Dretske, wants to show how a merely physical system could have thoughts
- but Dennett’s account is *nonreductive*—it does use mental ingredients

Daniel Dennett

# three predictive strategies

- the *physical stance*
- the *design stance*
- the *intentional stance*

# the physical stance

- use the system's physical properties and the laws of physics to predict its behavior

# the design stance

- assume that the system is designed to do such-and-such, and predict its behavior on this basis

# the intentional stance

- treat the system as a rational agent, figure out what beliefs and desires it ought to have, given its place in the world

“A little practical reasoning from the chosen set of beliefs and desires will in many—but not all—instances yield a decision about what the agent ought to do; that is what you predict the agent will do” (558)

# the power of the intentional stance

- “a fact largely concealed by our typical concentration on the cases in which it yields dubious or unreliable results”
- the prediction contest: Martian super-physicists vs. Earthlings



Where will *you* be in exactly one week





# the “perverse claim”

- “*all there is* to being a true believer is being a system whose behavior is reliably predictable via the intentional strategy, and hence *all there is* to really and truly believing... $p$  (for any proposition  $p$ ) is being an intentional system for which  $p$  occurs as a belief in the best (most predictive) interpretation”
- this is a “nonreductive” account of believing  $p$  because ‘being an intentional system for which...’ contains mental vocabulary

# the “perverse claim”

- but won't thermostats have beliefs about temperature on this view?



- no, because there are numerous different “best interpretations”

# the “perverse claim”

- this is a kind of behaviorism
- Commander Data, Block’s homunculi head, Block’s Aunt Bubbles machine (or a version of it suitably hooked up to a robot) are all “reliably predictable via the intentional strategy”
- to “really and truly” believe  $p$ , according to Dennett, the system just has to have the right kind of behavioral dispositions

# Minds and Machines

spring 2003

Anomalous monism

# “Mental events”

- a defense of “anomalous monism”
- like Dennett, Davidson is a materialist (hence, “monism”, as opposed to “dualism”)
- like Dennett, Davidson gives a nonreductive account of the mental (hence, “anomalous monism”)

Donald Davidson

# three principles

- **causal interaction**  
“every mental event is the cause or effect of some physical event”
- **the nomological character of causality**  
if c causes e, then there is a (strict) law of the form “A-type events are followed by B-type events”, where c is of type-A, and e is of type-B
- **the anomalism of the mental**  
there are no strict laws on the basis of which mental events can be predicted and explained

# anomalous monism

- **the token identity theory**  
“Every mental event...is a physical event”

(see the “demonstration of identity” on p. 124)

and:

- **the anomalism of the mental**  
there are no strict laws on the basis of which  
mental events can be predicted and explained

# lawlike statements

- **lawlike statements**

“general statements that support counterfactual and subjunctive claims, and [that] are supported by their instances”

e.g., ‘all swans are white’; evidence for this is also evidence for the “counterfactual conditional” ‘if this had been a swan, it would have been white’ (cf. ‘if this dime had been in my pocket, it would have been a quarter’)

- ***ceteris paribus* lawlike statements**

lawlike statements “qualified by generous escape clauses”—‘in normal conditions’, ‘other things being equal’, etc.



# laws

- **laws**  
true lawlike statements
- ***ceteris paribus* laws**  
true *ceteris paribus* lawlike statements  
(e.g., ‘normally, if demand increases faster than supply, prices rise’; ‘typically, people tend to avoid extreme pain’)
- **strict laws**  
laws that aren’t *ceteris paribus*—exceptionless true lawlike statements  
e.g., ‘nothing can be accelerated beyond  $c$ ’

# using laws to predict and explain

- all samples of copper expand when heated
- this bar is a sample of copper

so:

- this bar will expand when heated



# the anomalism of the mental

“There are no strict psychophysical laws because of the disparate commitments of the mental and physical schemes. It is a feature of physical reality that physical change can be explained by laws that connect it with other changes and conditions physically described. It is a feature of the mental that the attribution of mental phenomena must be responsible to the background of reasons, beliefs, and intentions of the individual. There cannot be tight connections between the realms if each is to retain allegiance to its proper source of evidence.”  
(p. 123)

# are the three principles consistent?

- **causal interaction**

“every mental event is the cause or effect of some physical event”

- in particular, suppose  $e_p$  causes  $e_m$

- **the nomological character of causality**

if  $c$  causes  $e$ , then there is a (strict) law of the form “A-type events are followed by B-type events”, where  $c$  is of type-A, and  $e$  is of type-B

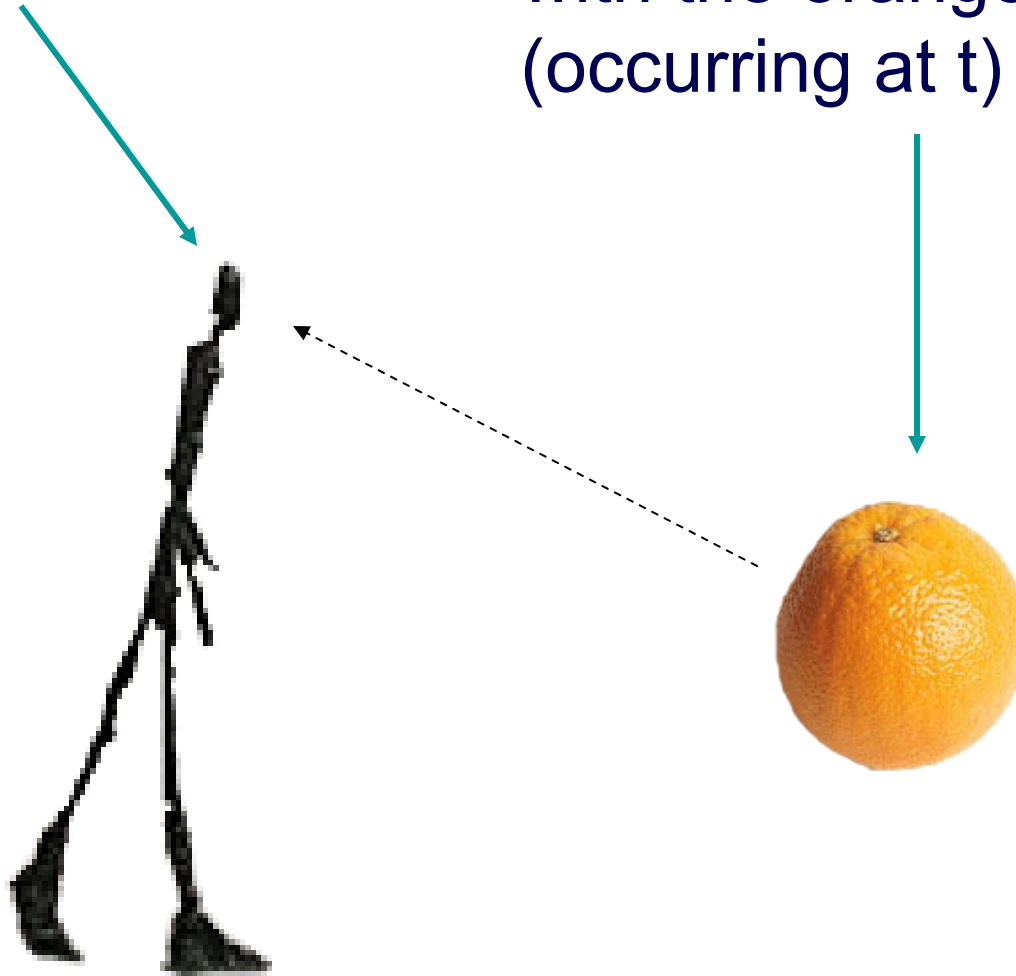
- so, there is a (strict) law of the form “A-type events are followed by B-type events”, where  $e_p$  is of type-A, and  $e_m$  is of type-B
- but then presumably this law can be used to predict and explain  $e_m$

# are the three principles consistent?

- in which case
- **the anomalism of the mental**  
there are no strict laws on the basis of which mental events can be predicted and explained
  - is false
- according to Davidson, this reasoning fails at the last step because “the principle of the anomalism of the mental concerns events described as mental” (p. 119)

$e_m$  = the experience  
(occurring at  $t+$ )

$e_p$  = interaction of light  
with the orange  
(occurring at  $t$ )



# are the three principles consistent?

- the required “strict law” will be couched in purely physical vocabulary  
it might look something like this:  
*if such-and-such electromagnetic events occur [and such-and-such complicated physical system is in the vicinity, in such-and-such physical state] then a few milliseconds later a neural event of type N will occur*
- this law cannot be used to predict and explain  $e_m$  “as such”, or “under a mental description”, because (according to Davidson) although  $e_m$  is a neural event of type N, not every neural event of type N is a visual experience as of an orange (or even a visual experience of any kind)

# Minds and Machines

spring 2003

- read Davidson, Kim