

# 24.09 Minds and Machines

## spring 2007

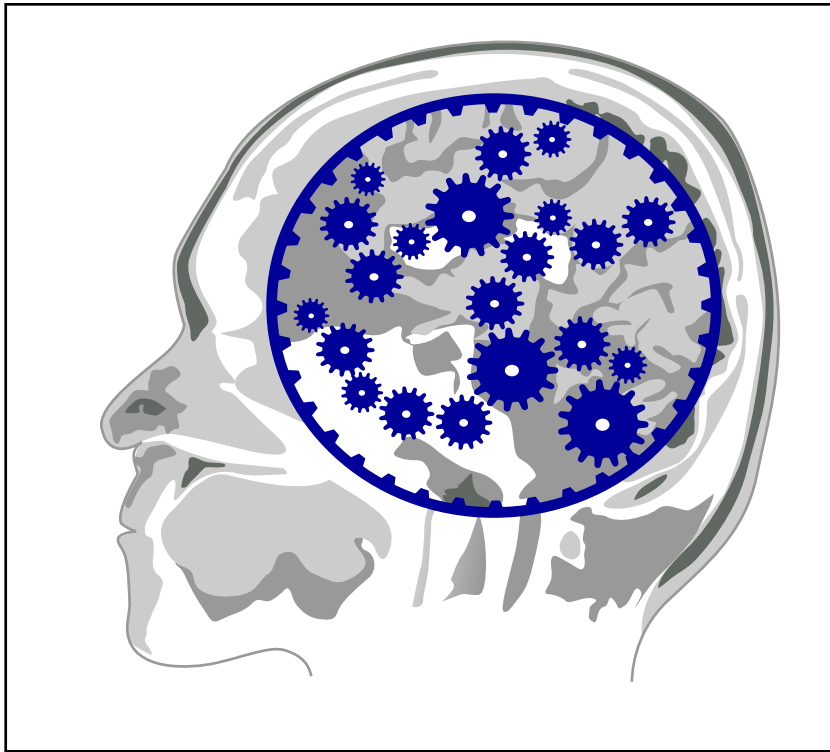


Figure by MIT OCW.

- the Churchlands on today's On Point (WBUR)
- Lewis, 'Mad pain...' in e-readings
  - you need to read this for Pset 5

we are  
here



# roadmap

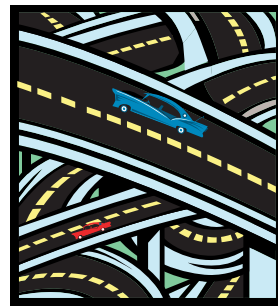
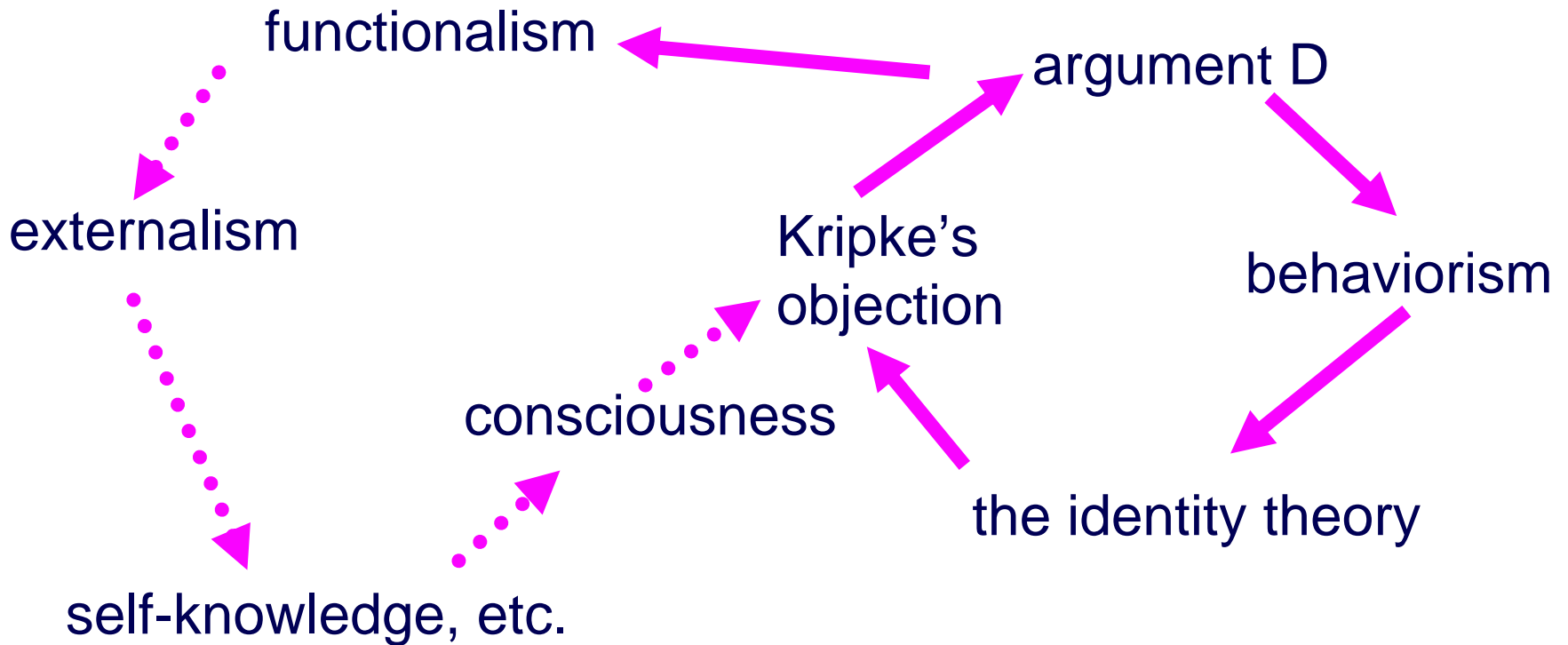


Figure by MIT OCW.



# multiple realization

Image removed due to copyright restrictions.

The [identity theorist] has to specify a physical-chemical state such that any organism (not just a mammal) is in pain if and only if (a) it possesses a brain of a suitable physical-chemical structure; and (b) its brain is in that physical-chemical state...it must be a state of the brain of any extra-terrestrial life that may be found that will be capable of feeling pain...

(Putnam, 77)

# “replacement” scenarios

- a gradual replacement, one neuron at a time, with a chip that has the same input-output profile
- is this any worse than a cochlear implant?

Image removed due to copyright restrictions.  
RoboCop movie poster.

Image removed due to copyright restrictions.  
The Six Million Dollar Man movie poster.

# lesson (?)

- minded creatures can in principle be made out of anything, provided certain structural constraints are met
- you couldn't make a mind out of Jello, but you could make one out of silicon chips

# the Turing test

- see “Computing machinery and intelligence” (1950)
- is passing the test sufficient for having a mind/being intelligent/etc.?
- the question is underdescribed
- who are the judges?
- how long is the test?
- what’s the subject matter?

Image removed due to copyright restrictions.  
Photograph of Alan Turing (1912-1954).

# the Turing test

- the claim that passing the test is sufficient for having a mind (etc.) should strike us as suspect
- it seems to conflate:
  - (a) we have/could have excellent evidence for  $p$
  - with
  - (b)  $p$  is true
- recall the distinction between “metaphysics” and “epistemology”, and the problem with behaviorism

# the Turing test

- in any event, for any Turing test, no matter how demanding, there is a machine that can pass it that seems not to have a mind
- this is shown by Block's "Aunt Bubbles" example



# the Aunt Bubbles machine

The machine works as follows. The judge goes first. Whatever the judge types in (typos and all) is one of  $A_1 \dots A_n$ . The machine locates the particular  $A$ , say  $A_{2398}$ , and then spits back  $B_{2398}$ , a reply chosen by the programmers to be appropriate to  $A_{2398}$ . The judge types another message, and the machine again finds it in the list of  $C$ s that sprout below  $B_{2398}$ , and then spits back the pre-recorded reply (which takes into account what was said in  $A_{2398}$  and  $B_{2398}$ ). And so on. Though the machine can do as well in the one hour Turing Test as Aunt Bubbles, it has the intelligence of a juke-box. Every clever remark it produces was specifically thought of by the programmers as a response to the previous remark of the judge in the context of the previous conversation.

# inside the Aunt Bubbles machine

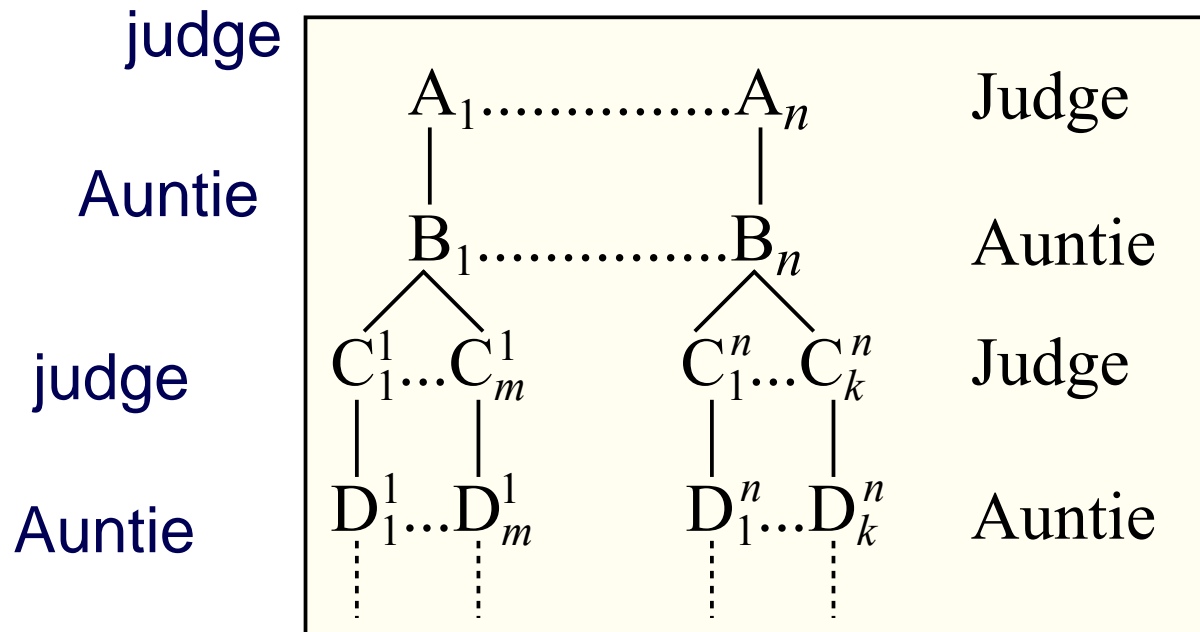


Figure by MIT OCW.

# lesson (?)

- minded creatures must have a certain sort of internal structure—being a giant lookup table isn't enough

# lessons (?) from our discussion of behaviorism

- mental states are the inner causes of behavior
- the correspondence between mental states and behavior is many-many
- (a) mental state M may produce different sorts of behavior, depending on the creature's other mental states
- (b) the same behavior may be produced by different mental states

# the mousetrap

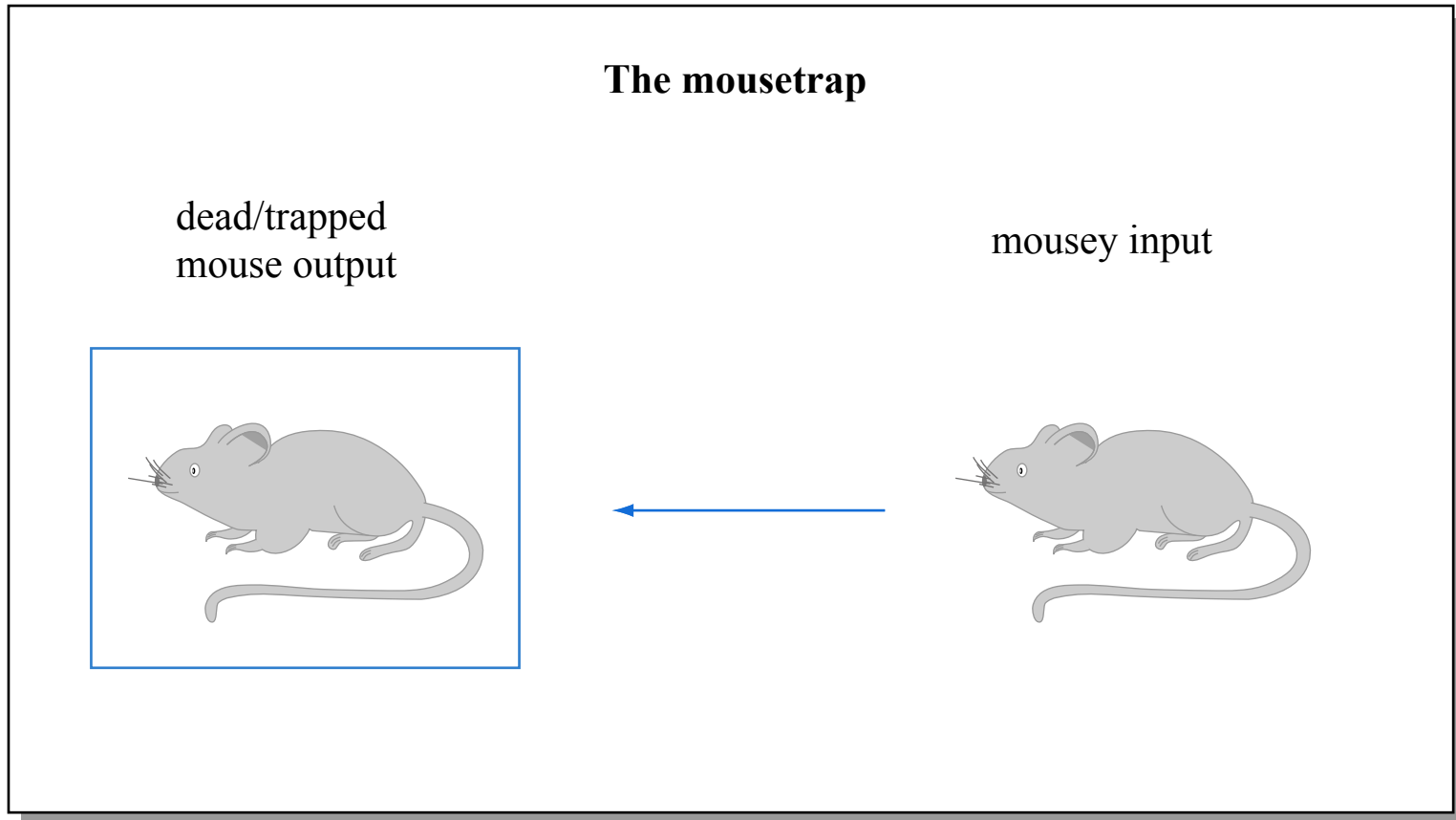


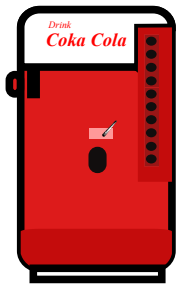
Figure by MIT OCW.

# mousetraps can be “multiply realized”

Images removed due to copyright restrictions.  
Mousetrap.

# the mousetrap

- the simplest sort of functional kind
- since no constraints are placed on its inner organization, it is also a behavioral kind



# the 3-Coke vending machine

Figure by MIT OCW.

initial  
state  
s

output  
for  
25¢ input

next  
states




$M_3 I_0$	$M_2 I_0$	$M_1 I_0$	$M_3 I_1$	$M_2 I_1$	$M_1 I_1$
"25¢"	"25¢"	"25¢"			
$M_3 I_1$	$M_2 I_1$	$M_1 I_1$	$M_2 I_0$	$M_1 I_0$	shut down

Figure by MIT OCW.



# the 3-Coke vending machine

- it can be multiply realized
- a functional but not behavioral kind
- what are the states  $M_3$ ,  $I_2$ , etc?  
—the table tells the whole story

# a toy functionalist theory of pain

input \ state	toe-stubbing	icepack on toe
P	P, "Ow!"	R, "Phew!"
R	P, "Ow!"	R, no output

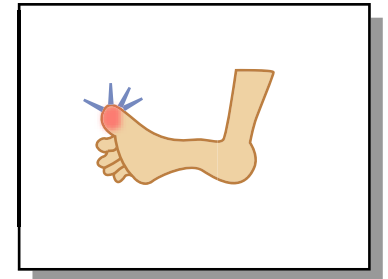


Figure by MIT OCW.

# the toy functionalist theory

- the state of being in pain (or being in a state of relief) just is being in P (R)
- S is in P iff S is in the first of two states X, Y, that are related to one another and to the possible inputs and outputs of S as follows:

being in X and stubbing its toe causes S to remain in X and emit “Ow!”; being in Y and stubbing its toe causes S be in X and emit “Ow!”; being in X and having an icepack on the toe causes S to be in Y and emit “Phew!”; being in Y and having an icepack on the toe causes S to remain in Y and produce no output

# functionalism

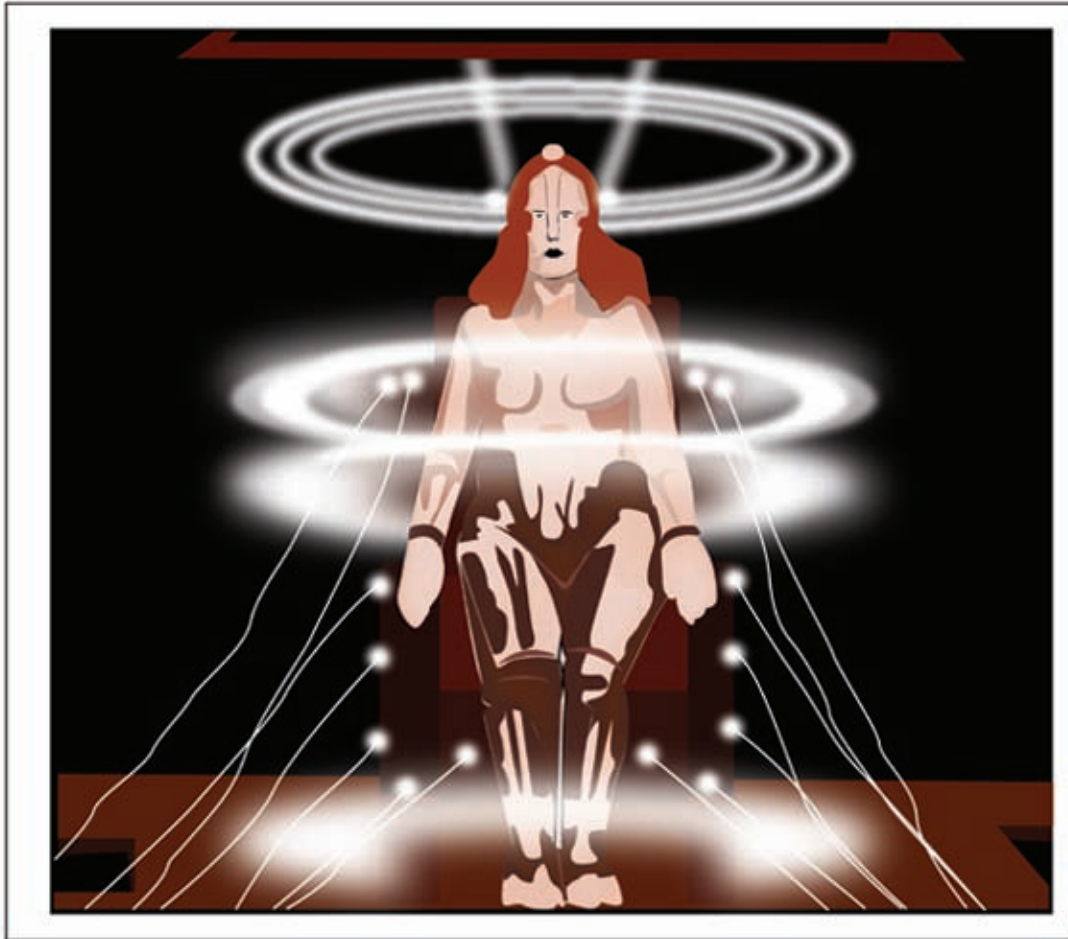
- the view that mental states are functional states—states specified in terms of their causal relations to inputs, outputs, and other states

# functionalism and the lessons (?)

- mental states can be multiply realized
- input-output isn't enough: it's what's in between that matters (recall the Aunt Bubbles machine)
- mental states are the inner causes of behavior
- the correspondence between mental states and behavior is many-many
- functionalism accommodates all the above

# Minds and Machines

spring 2007



- read Block, 'Troubles...' for next time

Figure by MIT OCW.