24.09 Minds and Machines
spring 2007

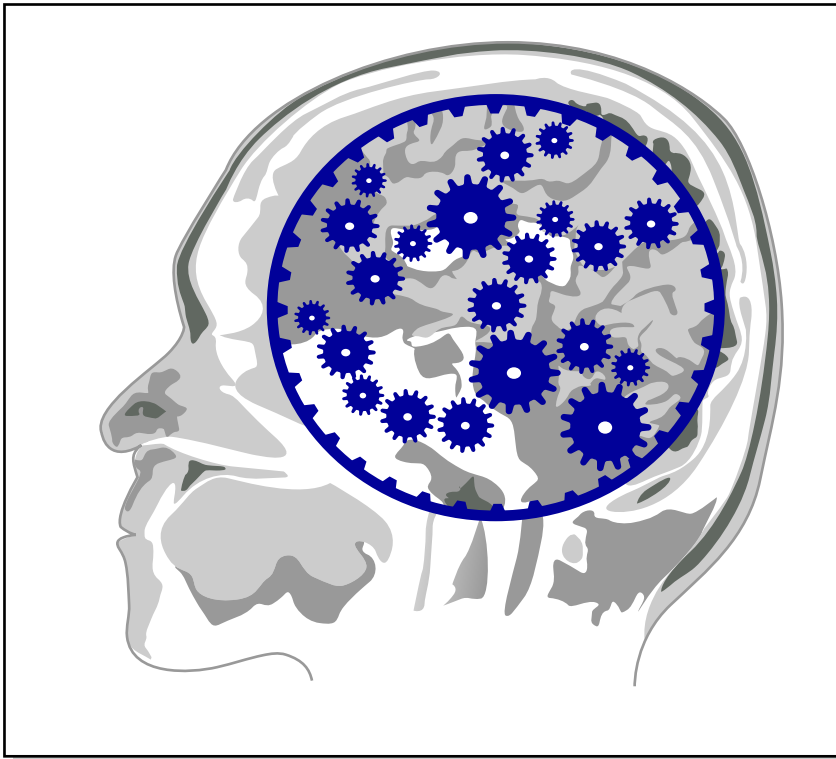- books
- problem set
- readings for this week
- add forms
- after class salon on wed

Figure by MIT OCW.

zombies

- not Hollywood zombies
- physical duplicates of us, but lacking conscious experience entirely
- zombies are imaginable, but are they possible—could zombies have existed?
- "materialists" say no; "dualists" say yes
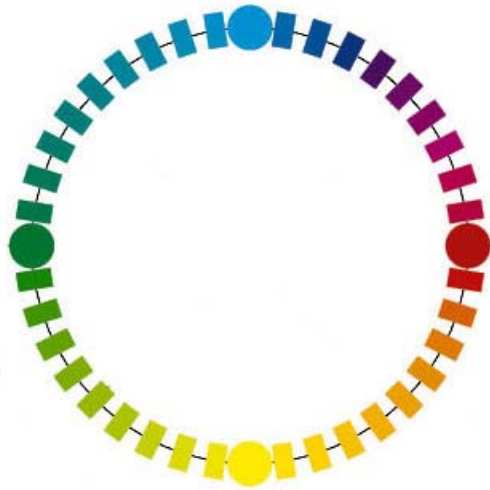- cf. a "digestion zombie"

Image removed due to copyright restrictions. Zombies!!! cover.

the "inverted spectrum"
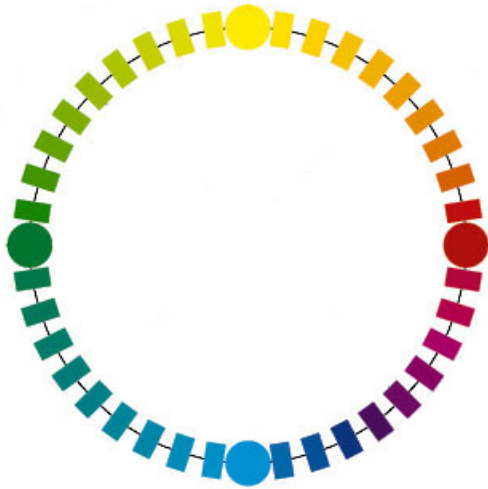
- devised by Locke, *An Essay Concerning Human Understanding* (1689)
- someone might have color experiences that are "inverted" with respect to yours

Invert's experiences

- arguably, an empirical possibility
- some philosophers think the thought experiment refutes various "materialist" theories of the mind
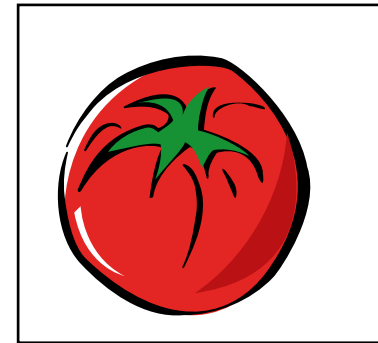
I wonder what it's like to see red?

black and white Mary

Figure by MIT OCW.

Figure by MIT OCW.

our first topic
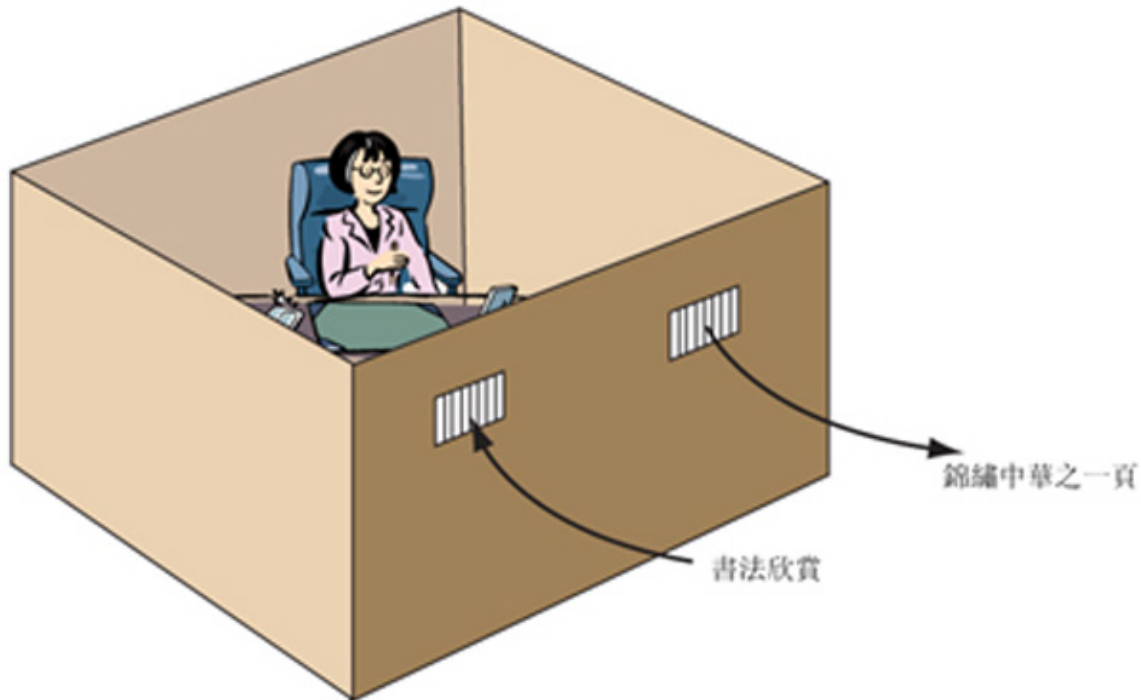
- Searle's "Chinese room argument"

書法欣賞

錦繡中華之一頁

Figure by MIT OCW.

arguments
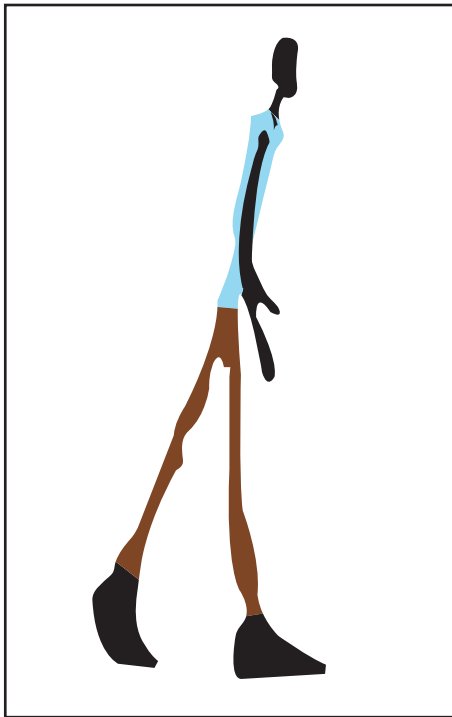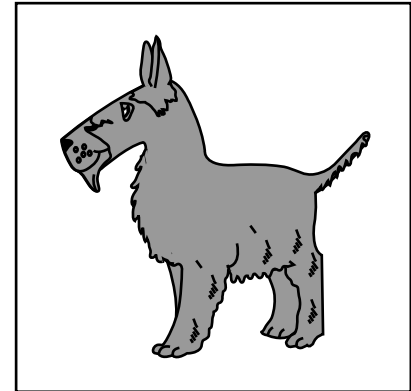

Figure by MIT OCW.

- an argument is a list of sentences, $P_1,...P_n$, C
- see <u>handout</u> on arguments (in study material section.)
- validity
- soundness

FIDO ES PELUDO

Fido is furry
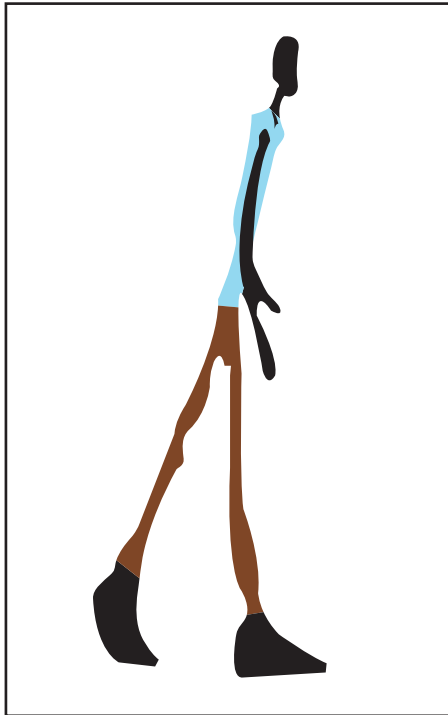
FIDO IS FURRY

Figures by MIT OCW.

intentionality

- that feature, possessed by (e.g.) words and mental states, of being "about"— representing, referring to—something. The belief that Fido is furry is a mental state that is about Fido. And the word 'Fido' refers to the dog, Fido. So these are two examples of intentionality.

- NB: don't confuse intentionality with intending something. The latter is just one example, along with believing and desiring, of an intentional mental state.

FIDO ES PELUDO

Fido is furry

FIDO IS FURRY
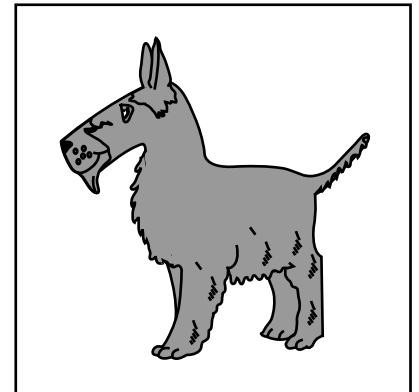
Figures by MIT OCW.

derived/underived intentionality

- something has <u>derived</u> intentionality just in case its intentionality can be explained in terms of the intentionality of something else (see Searle, 5)

  - plausibly, 'dog' refers to dogs in virtue of the beliefs, intentions, etc., of English speakers—hence 'dog' has derived intentionality; my belief that dogs have fur is an intentional state, and doesn't have its intentionality in virtue of the intentionality of anything else—hence my belief has <u>underived</u> (or original) intentionality.

  - if thinking is conducted in a language written in the brain, then the words of this language have underived intentionality

- Searle's Chinese room argument is directed against the claim that instantiating a computer program is sufficient for <u>underived</u> intentionality

can computers think?

Image removed due to copyright restrictions.

- according to "Strong AI", "the mind is to the brain, as the program is to the computer hardware"

- Searle's "Chinese room argument" purports to show that Strong AI is false

Sound by MIT OCW.

24.09 spring 07

WEAK AI

- the principle value of the computer in the study of the mind is that it gives us a very powerful tool—e.g. it enables us to simulate various kinds of mental processes

- cf. WEAK ARTIFICIAL METEOROLOGY

- WEAK AI is obviously correct (ditto WEAK AM)

Image removed due to copyright restrictions.

STRONG AI

- an appropriately programmed computer literally has mental states (in particular, cognitive states)

- cf. STRONG AM—an appropriately programmed computer literally has meteorological states

- STRONG AI is disputable, and disputed by Searle

- STRONG AM, at least, is obviously false

the Chinese room

- a program: an algorithm (mechanical recipe) for transforming symbols into symbols

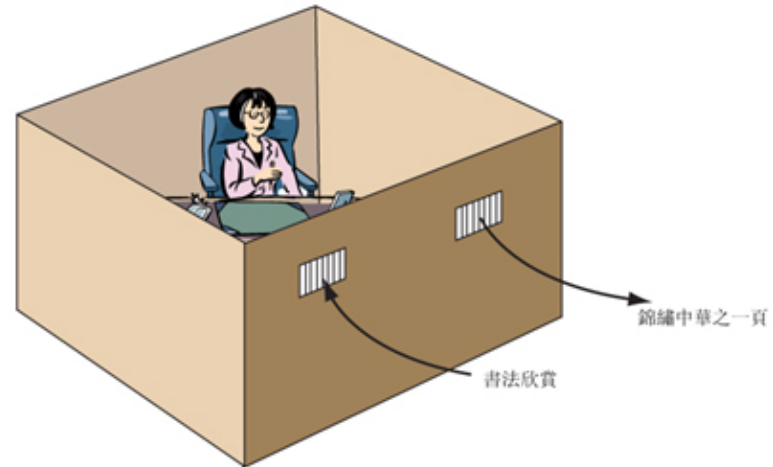- the thought experiment exploits the fact that computer programs can be "multiply realized"

Figure by MIT OCW.

"…you behave exactly as if you understood Chinese, but all the same you don't understand a word of Chinese. But if going through the appropriate computer program for understanding Chinese is not enough to giv you an understanding of Chinese, then it is not enough to give <u>any other digital computer</u> an understanding of Chinese"
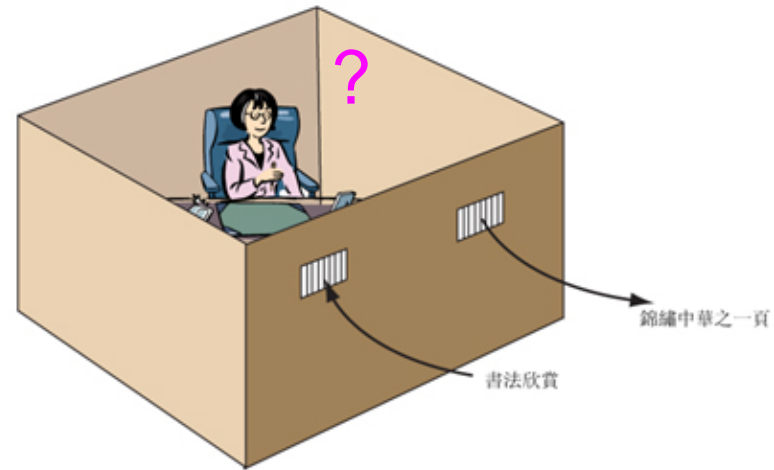
So, strong AI is false
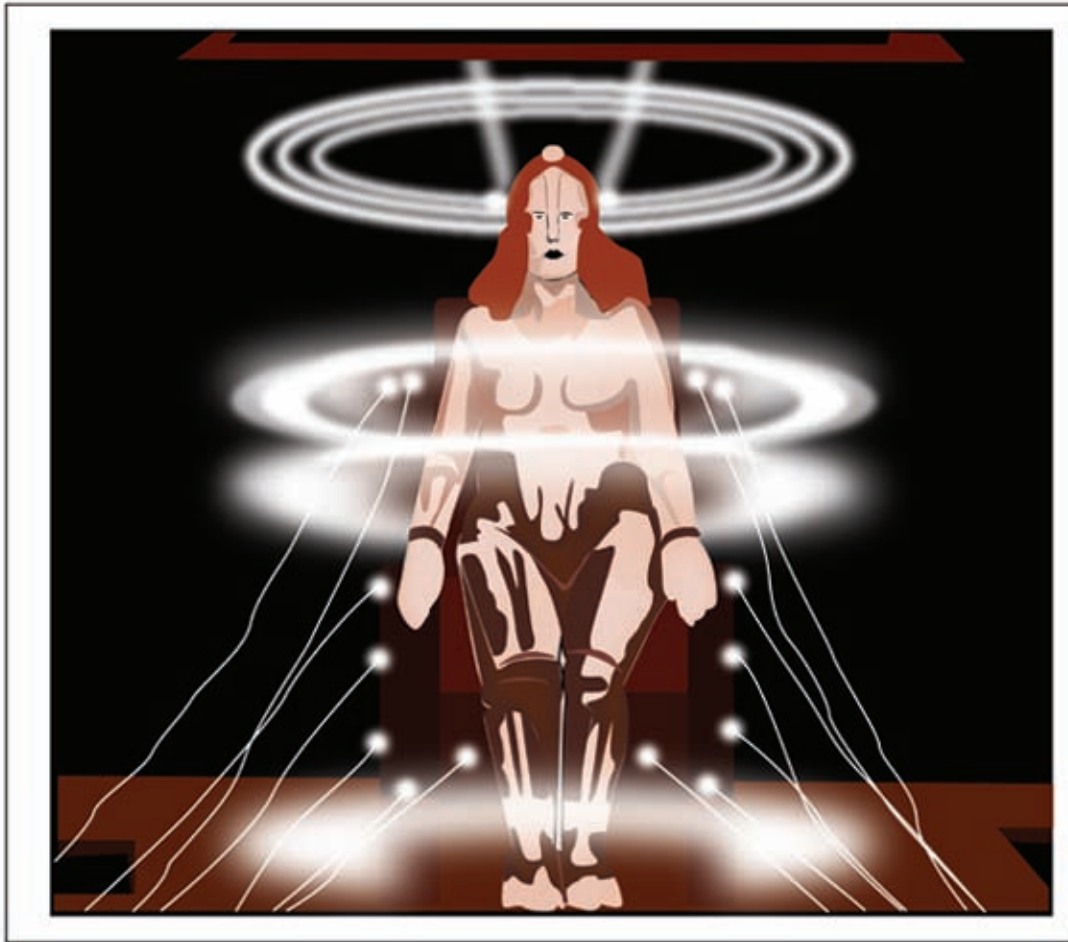
?

錦繡中華之一頁

書法欣賞

Figure by MIT OCW.

Minds and Machines
spring 2007

Figure by MIT OCW.

- read Descartes for wednesday

24.09 spring 07