Philosophy 101

Can Machines Think?

October 9, 2014

Review

What is a Representation?

Representation: something that represents something (either itself or something else)

- goes proxy; stands for; symbolizes something
- refers to something; is accurate/inaccurate; is true/false

Two Questions

1. How does a representation represent?
2. Are some kinds of representation more fundamental than others?

Examples of Representational Kinds

- Pictorial Representation
- Linguistic Representation
- Mental Representation

Questions

1. Crane argues that pictorial representation is not fundamental because

   A. it requires imagination
   B. it requires pictures
   C. it requires interpretation
   D. it requires language

Questions

2. Crane argues that mental representation is fundamental in part because

   A. it is conventional
   B. it is not conventional
   C. it requires interpretation
   D. it requires neither convention nor interpretation

Intentionality

Intentionality: The 'directedness' of a mental state at some (existing or non-existing) thing

- All mental representations are 'intentional' or directed at some thing, fact, or situation

Mental Representation

- According to Crane, only mental representation could plausibly be fundamental because only mental representation is *intrinsically* intentional
- What kinds of things might have intentional states?
    - Could machines have intentional states?

Can Machines Think?

Poll

1. Could a sufficiently advanced computer qualify as a thinking being?

   A. Yes
   B. No

Two Questions

1. Can a physical system capable of performing certain functions think?
2. Can a sufficiently sophisticated computer program think?

Two Questions

2. Can a sufficiently sophisticated computer program think?
    - Is the mind to the brain like software is to hardware?

Strong & Weak AI

Strong AI: thinking is constituted by the manipulation of formal symbols, such as occurs in a computer program

Weak AI: thinking may be modeled by formal symbol systems, such as computer programs

The Imitation Game

- Can you guess, using a series of questions, which of two conversation partners is a machine and which a human?
- Questions may be of all kinds:
  - what's your name
  - what's your favorite color?
  - what does the smell of freshly cut grass remind you of?

The Turing Test

I believe that in about fifty years' time it will be possible to programme computers…to make them play the imitation game so well that an average interrogator will not have more than 70 percent chance of making the right identification after five minutes of questioning…I believe that at the end of the century the use of words and general educated opinion will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted. (Alan Turing)

The Turing Test

1. For some arbitrary time period, there will be no discernible difference between the linguistic behavior of a person and that of a machine
2. If there is no discernible difference in linguistic behavior between man and machine, then there is no reason to think that there is any underlying difference in the causes of that behavior
3. ∴ If we are willing to say that it is intelligent thought that is the cause of the linguistic behavior in the person we should be willing to say the same thing about the machine

Strong AI & the Turing Test

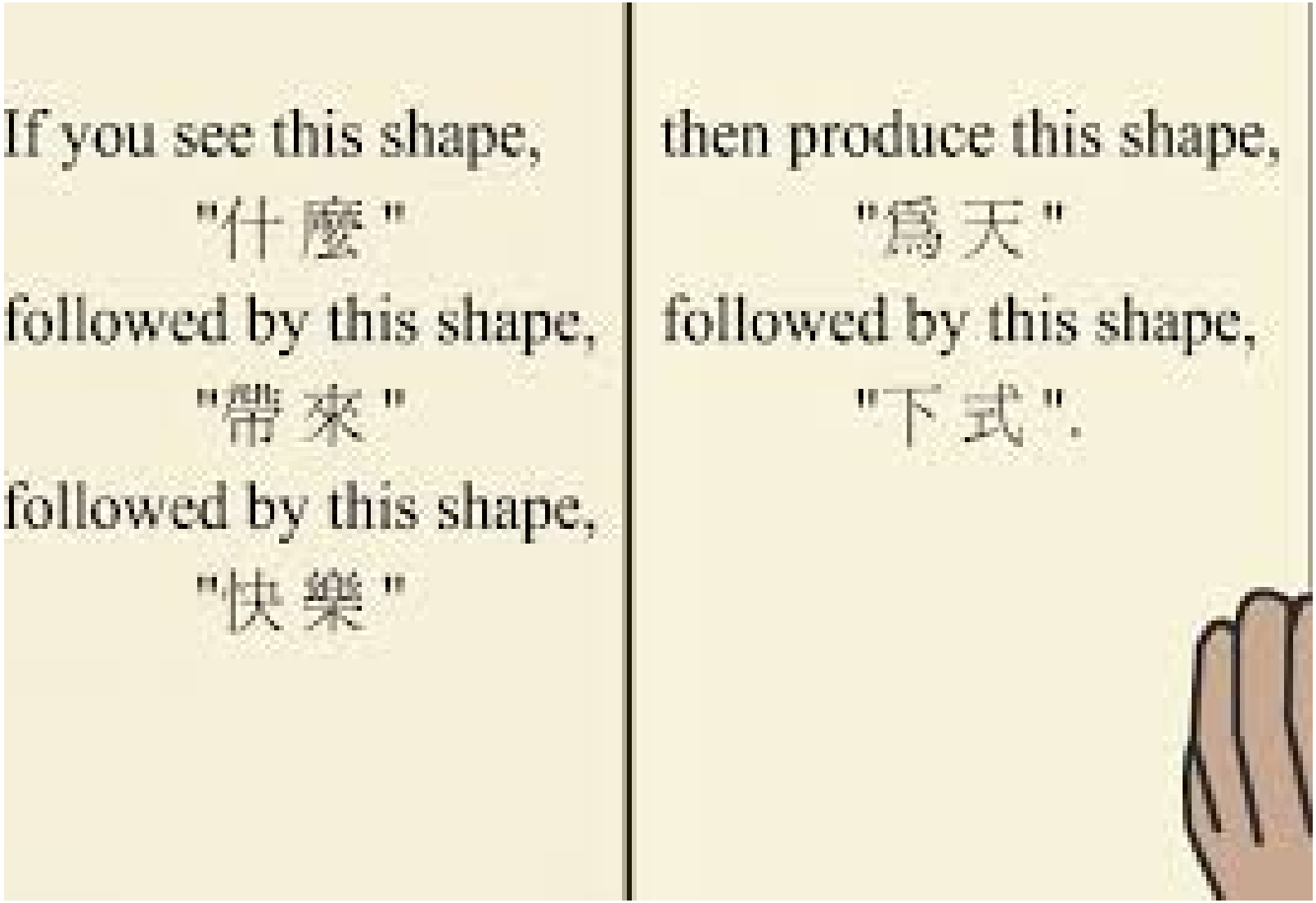
- Any computer that can pass the Turing Test for arbitrarily long periods of time will, according to strong AI, qualify as a thinking machine

Searle & the Chinese Room

The Chinese Room Argument

suppose I am placed in a room containlng baskets full of Chinese symbols. Suppose also that I am given a rule book in English for matching Chinese symbols with other Chinese symbols. The rules identify the symbols entirely by their shapes and do not require that I understand any of them. The rules might say such things as, "Take a squiggle-squiggle sign from basket number one and put it next to a squoggle-squoggle sign from basket number two." Imagine that people outside the room who understand Chinese hand in small bunches of symbols and that in response I manipulate the symbols according to the rule book and hand back more small bunches of symbols.

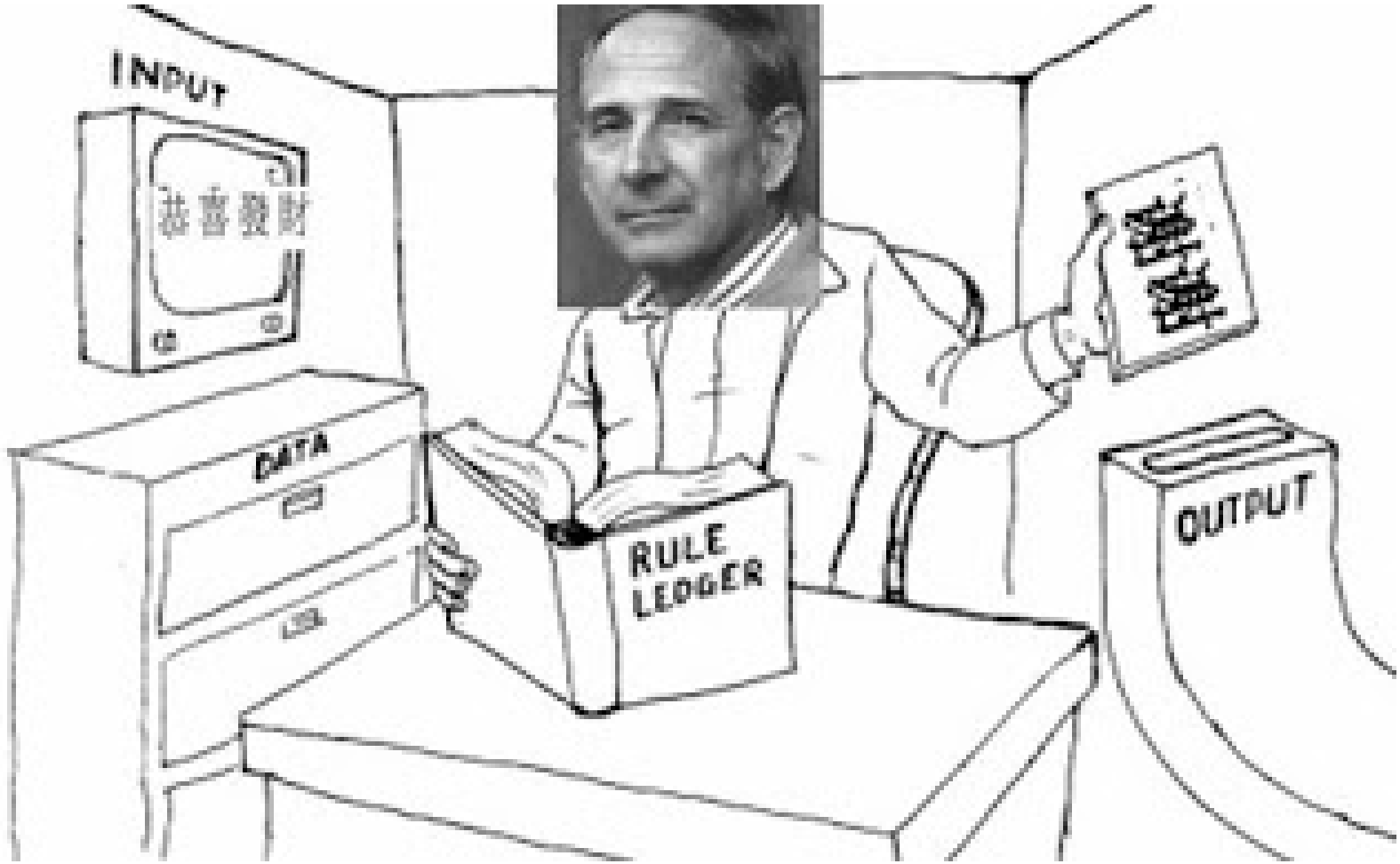
The Chinese Room Argument



The Chinese Room Argument

Now, the rule book is the "computer program." The people who wrote it are "programmers," and I am the "computer." The baskets full of symbols are the "data base," the small bunches that are handed in to me are "questions" and the bunches I then hand out are "answers."

The Chinese Room Argument



The Chinese Room Argument

Now suppose that the rule book is written in such a way that my "answers" to the "questions" are indistinguishable from those of a native Chinese speaker. For example, the people outside might hand me some symbols that unknown to me mean, "What's your favorite color?" and I might after going through the rules give back symbols that, also unknown to me, mean, "My favorite is blue, but I also like green a lot." I satisfy the Turing test for understanding Chinese. All the same, I am totally ignorant of Chinese. And there is no way I could come to understand Chinese in the system as described, since there is no way that I can learn the meanings of any of the symbols. Like a computer, I manipulate symbols, but I attach no meaning to the symbols. (Searle, 26)

The Chinese Room Argument

The Chinese Room Argument

1. Programs are purely formal (syntactic)
2. Human minds have mental contents (semantics)
3. Syntax by itself is neither constitutive of, nor sufficient for, semantic content
4. ∴ Programs by themselves are not constitutive of nor sufficient for minds

Syntax & Semantics

Syntax: the formal or structural features of a symbol system which determine which expressions are legitimate members of the system and which are not

- The syntax of English (its grammar) requires that all complete sentences have a noun phrase and a verb phrase
  - 'John goes to school' vs. 'school John to goes'

Syntax & Semantics

Semantics: The system of meanings assigned to a symbol system, given by determining the referents of the symbols and the truth conditions of symbol strings

- ‘Schnee’ refers to snow
- ‘weiß’ refers to the property of being white
- ‘Schnee ist weiß’ is true just in case snow is white

What Does the Argument Intend to Prove?

- You can't get semantic content from syntax alone
- A system must have more than purely syntactic properties in order to possess intentional states