

# Philosophy of the Human Person

## Lecture #23

### Thought

#### Background

two kinds of artificial intelligence (AI) research

an engineering project—programming computers to emulate intelligent behavior

a philosophical project—attempting to understand human intelligence by analyzing how computers work, on the principle that

the mind is to the brain

as a program is to a computer

in terms of the outline arguments

AI researchers attempt to respond to the argument

No machines can think

Some organisms can think

So, Some organisms are not machines.

by arguing that some machines (namely, properly programmed digital computers) can think.

for the argument, see below

## 1. The Case for Computers that can Think

A standard argument for the possibility of machines thinking

	Premises	Intermediate Conclusions
Church's thesis	(1) Every rule-governed input-output function <sup>1</sup> can be computed in a finite time by a simple symbol-manipulating machine <sup>2</sup>	
	(2) Human response to conversational questions and remarks is a rule-governed input-output function.	So, (3) A human response to conversational questions and remarks can be computed in a finite time by a simple symbol-manipulating machine.
	(4) Anything that can be computed in a finite time by a simple symbol-manipulating machine can be performed by a digital computer	So, (5) A human response to conversational questions and remarks can be performed by a digital computer.  = A (properly programmed) digital computer can give a human response to conversational questions and remarks
Definition of the Turing Test	(6) Anything that can give a human response to conversational questions and remarks can pass the Turing Test.	So, (7) A (properly programmed) digital computer can pass the Turing Test.
Adequacy of the Turing Test	(8) Anything that can pass the Turing Test is intelligent.	So, (9) A (properly programmed) digital computer is intelligent.

<sup>1</sup> technically, every effectively computable function

<sup>2</sup> e.g., by a Turing machine

## Comments on the premises

## On P1, Church's thesis

Church's thesis is not formally provable, but it is reasonable and is generally accepted

what is a function?

an effectively computable function is a function for which there is a rote procedure for determining the output (for each input) in a finite amount of time

a Turing machine (see description in Kemeny) is an example of a maximally simple symbol-manipulating machine

it has been shown that

- (i) for any effectively computable function, one can build a Turing machine that can compute that function, and
- (ii) one can build a single Turing machine that is able to compute every effectively computable function

P2 is an invitation to see human conversation in a certain way

P4 is a fact about digital computers, based on the definition of digital computers  
digital computer—"a device for calculating and data processing that handles numbers & other information as collections of individual digits or letters"<sup>3</sup>

it is and is not controversial

## On the Turing Test

background

if the question is whether a machine can think, we need a criterion for thought; the Turing Test proposes one

what counts as "giving a human response"? giving responses so good that a human being cannot tell whether the responses are coming from another human being or are generated by a computer

P6 is a definition of the Turing Test

P8 says that the Turing Test is an adequate test for thought

## evaluation of the argument

its proponents say that it shows that, even though no computer can at present pass the Test (and therefore none can be said to think) there is no reason to think that there is a reason (in principle) that a computer could not do so

its critics

Searle rejects P8—Some things that can pass the Turing Test are not intelligent

---

<sup>3</sup> *Encyclopaedia Britannica* (1976), III-545

Dreyfus rejects P2—Human response to conversational questions and remarks is a rule-governed input-output function.

since computers can only handle rule-governed input-output functions, they cannot handle real human conversation as a human being does

### 3. Searle's Objection

Searle's objection to the argument

contra P8: *Some* machines that can pass the Turing Test must *not* be regarded as intelligent.

his argument (by counter-example & thought experiment)

(1) The Chinese Room passes the Turing Test

(2) The Chinese Room does not understand what it is saying

So (3) some things that pass the Turing Test don't understand what they're saying

(4) Anything that doesn't understand what it's saying doesn't display intelligence

So, (5) some things that pass the Turing Test do not display intelligence

Searle's general argument against "strong AI"

strong AI—(1) the brain is a digital computer; & (2) the mind is a computer program

hence, there is nothing distinctively biological about the human mind [i.e., it's not tied to life processes?]

(1) Anything that is not able to handle semantics (i.e., to attach meanings to symbols) does not display intelligence.

(2) Anything that can only do syntactic transformations (i.e., rearrangements of the *order* of the symbols) is not able to handle semantics (i.e., to attach *meanings* to symbols)

So (3) Anything that can only do syntactic transformations does not display intelligence.

(4) Some things that can pass the Turing test can only do syntactic transformations.

So, (5) some things that pass the Turing Test do not display intelligence

an elaboration on the distinction between syntax & semantics

if syntax is sufficient for semantics, there's always a single proper interpretation for symbol-manipulation

but there is not always a single proper interpretation

(1) suppose three people see a student writing  $A \oplus B = B \oplus A$  on a chalkboard

one thinks he's doing algebra and that  $\oplus$  is just the addition symbol

another thinks he is doing set theory, with  $\oplus$  standing for the union of two sets

another thinks he is doing Boolean logic, with  $\oplus$  standing for "or"

who's right? one can ask the student

he might just be manipulating symbols

but he probably is doing some kind of mathematics & knows what he is doing

he wouldn't be thinking the same thing if he were doing set theory as he would if he were doing Boolean algebra

(2) suppose the same three see a computer printing  $A \oplus B = B \oplus A$

one thinks he's doing algebra and that  $\oplus$  is just the addition symbol

another thinks it's doing proofs in set theory

another thinks it's doing proofs in Boolean logic

who's right?

one can't ask the computer; it's just manipulating symbols

one can ask the programmer what he meant the computer to be doing

but the computer is doing the same thing regardless of whether the programmer wants set theory or Boolean algebra

Another use of Searle's insights to distinguish computers & minds

(1) All digital computers are devices whose operations can be specified purely formally

this is the feature that makes them so powerful—the 1's & 0's can be about anything, hence they can do computations, graphics, word processing, &c.

(2) Any device whose operations can be specified purely formally ("syntactically") is a device which operates exclusively on symbols which have no meaning or content (i.e., no semantics).

So, (3) All digital computers are devices which operate exclusively on symbols which have no meaning.

(4) No minds are devices which operate exclusively on symbols which have no meaning.

since: All minds are devices which operate on symbols which have content.

or, the objects of our thoughts have semantic, as well as syntactic, aspects

So, (5) No minds are digital computers.

what exactly does the argument prove (according to Searle)?

not that machines can't think—we are all machines

nor that artefacts can't—it depends on what kinds of artefacts we can build

rather, that a digital computer can't think, or at least that something that is *only* a digital computer can't think

## 2. Dreyfus' Objection

Dreyfus objection to the argument that computers are intelligent

he denies P2: Human response to conversational questions and remarks is *not* a rule-governed input-output function

his fundamental distinction

intuitive thinking vs. "rational" thinking

intuitive thinking is based on

a large stock of situations & appropriate responses in memory

the ability to recognize the similarity of the situation of interest to one of the situations in memory

"rational" thinking is based on

a stock of rules

the ability to apply rules to cases

application of the distinction to computers

<u>Computers</u> computers are limited to these operations	<u>Thinking Things</u> thinkers can <i>also</i> do these
Reasoning from principles to particular applications of those principles	Intuitive movement from paradigms to similar situations
Syllogistic reasoning	Pattern recognition (analogical reasoning)
Strictly following rules	Guided by common sense

the distinction as applied to P2:

(1) Human response to conversational questions is an activity that requires intuitive thinking.

(2) No activity that requires intuitive thinking is a rule-governed input-output function.

So, (3) Human response to conversational questions and remarks is *not* a rule-governed input-output function.

in contrast to Searle, he accepts (at least for the sake of argument) the Turing Test as an adequate test of thought, but denies that any computer could pass it

the argument just made can thus be rephrased as follows:

(1) Passing the Turing Test requires intuitive thinking (he might have said, common sense)

(2) No computer is capable of intuitive thinking (or, has common sense)

So, (3) no computer can pass the Turing Test

the distinction as applied to directly to the question of whether computers can think

(1) All thinking things have a capacity for intuitive thought.

(2) No computers have the capacity for intuitive thought, since

(3) All computers are strict rule-followers

(4) No strict rule-followers have the capacity for intuitive thought

So (from 1 & 2), no computers can think.

with respect to the Turing test

he accepts the Turing Test as an adequate test of thought, but denies that any computer could pass it

his argument

Passing the Turing Test requires intuitive thinking (he might have said, common sense)

No computer is capable of intuitive thinking (or, has common sense)

So, no computer can pass the Turing Test