The Limits of Artificial Intelligence

The Limits of Artificial Intelligence

What does it mean to think or to feel?
What is mind? Does mind really exist?

To what extent are minds functionally dependent upon the physical structures with which they are associated?

Are minds subject to the law of physics?
If so, what are the laws of physics?

Two kinds of Errors

(a) Errors of functioning

(b) Errors of conclusion

Logically Possibility of Machines with Mind

Empirically Impossibility Machines with Mind

Searle's Argument Against AI

Consciousness is central to the mental phenomena. We think of ourselves as conscious, mindful, rational agents in the world, but science tells us that the world consists entirely of mindless physical particles.

How can we match these two conceptions?

Can it be the case that the world contains nothing but unconscious physical particles, and yet that it also contains consciousness?

Can an essentially meaningless world contain meanings?

How should we interpret the recent work in computer science and artificial intelligence aimed at making intelligent machines?

Does the digital computer give us the right picture of the human mind?

What is the relation between the ordinary, commonsense explanations of people's behaviour and its scientific modes of explanation?

Searle offers a biological naturalism explanation of the mind.

Searle says that mental events and processes are as much part of our biological natural history as digestion, mitosis, meiosis, or enzyme secretion.

The biological naturalism raises many questions of its own.

What about the great variety of our mental life-pains, desires, tickles, thoughts, visual experiences, beliefs, tastes, smell, anxiety, fear, love, hate, depression and elation?

What exactly is consciousness and how exactly do conscious mental phenomena relate to the unconscious?

What are the special features of the ‘mental’, phenomena such as consciousness, intentionality, subjectivity, and mental causation? And how exactly do they function?

What are the causal relations between ‘mental’ phenomena and ‘physical’ phenomena?

Can we characterize those causal relations in a way that avoids epiphenomenalism?

Searle's biological naturalism provides an effective counter argument to the currently fashionable computational theory of mind according to which, the mind is a computer program.

The brain is just a digital computer and the mind is just a computer program. Searle call it 'strong artificial intelligence' or 'strong-AI'-by saying that the mind is to the brain, as the program is to the computer hardware.

Symbols have no meaning. They have no semantic content, they are not about anything. They have to be specified purely in terms of their formal or syntactical structure. By definition, our internal mental states have certain sorts of contents.

The mind has more than a syntax, it has a semantics. The reason that no computer program can ever be a mind is simply that a computer program is only syntactical, and minds are more than syntactical. Minds are semantic in the sense that they have more than a formal structure, they

Searle presents a thought experiment about a Chinese Room for refuting the possibility of AI. This is called the Chinese Room Argument.

The Imitation Game

MAN	WOMAN
“I’m the woman”	“I’m the woman”
INTERROGATOR	

~~The Imitation Game~~

The Turing Test

~~MAN~~ **CMPUTER** "I'm the woman"	**WOMAN** "I'm the woman"
INTERROGATOR	

~~The Turing Test #2~~
The Imitation Game

~~MAN~~
COMPUTER

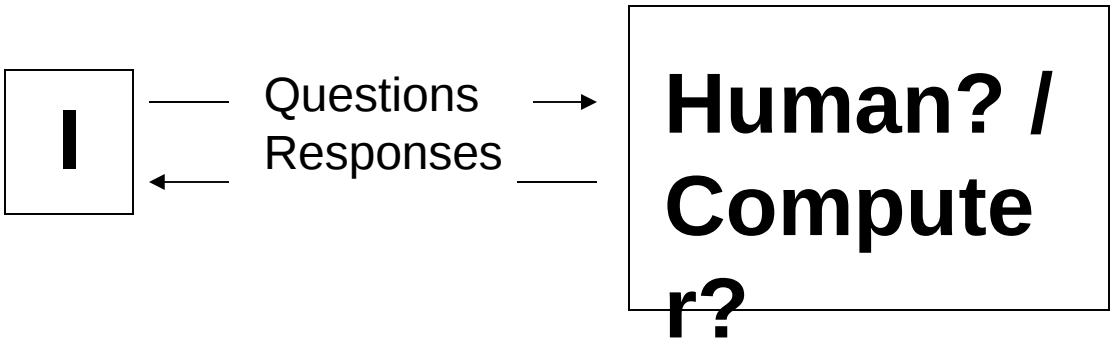
"I'm the ~~woman~~"
man

~~WOMAN~~
MAN

"I'm the ~~woman~~"
man

INTERROGATOR

The Turing Test

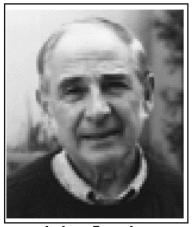


"I believe that at the end of the century the use of words and general educated opinion will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted."

- Turing 1950

The Chinese Room

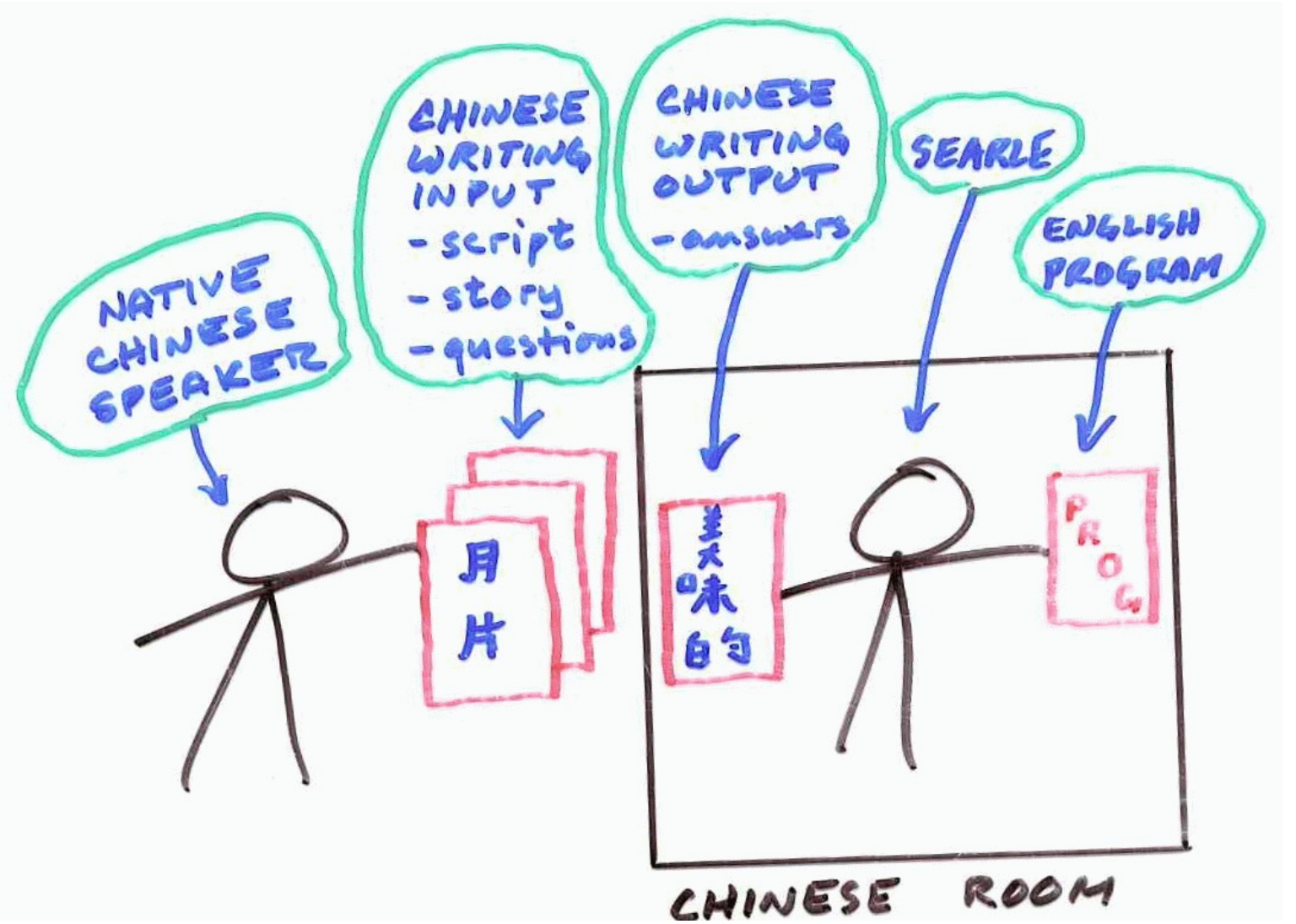
3 John Searle, 1980a, 1980b, 1990b
The Chinese Room argument. Imagine that a man who does not speak Chinese sits in a room and is passed Chinese symbols through a slot in the door. To him, the symbols are just so many squiggles and squoggles. But he reads an English-language rule book that tells him how to manipulate the symbols and which ones to send back out. To the Chinese speakers outside, whoever (or whatever) is in the room is carrying on an intelligent conversation. But the man in the Chinese Room does not understand Chinese; he is merely manipulating symbols according to a rule book. He is instantiating a formal program, which passes the Turing test for intelligence, but nevertheless he does not understand Chinese. This shows that instantiation of a formal program is not enough to produce semantic understanding or intentionality. **Note:** For more on Turing tests, see Map 2. For more on formal programs and instantiation, see the "Is the brain a computer?" arguments on Map 1, the "Can functional states generate consciousness?" arguments on Map 6, and sidebar, "Formal Systems: An Overview," on Map 7.



John Searle

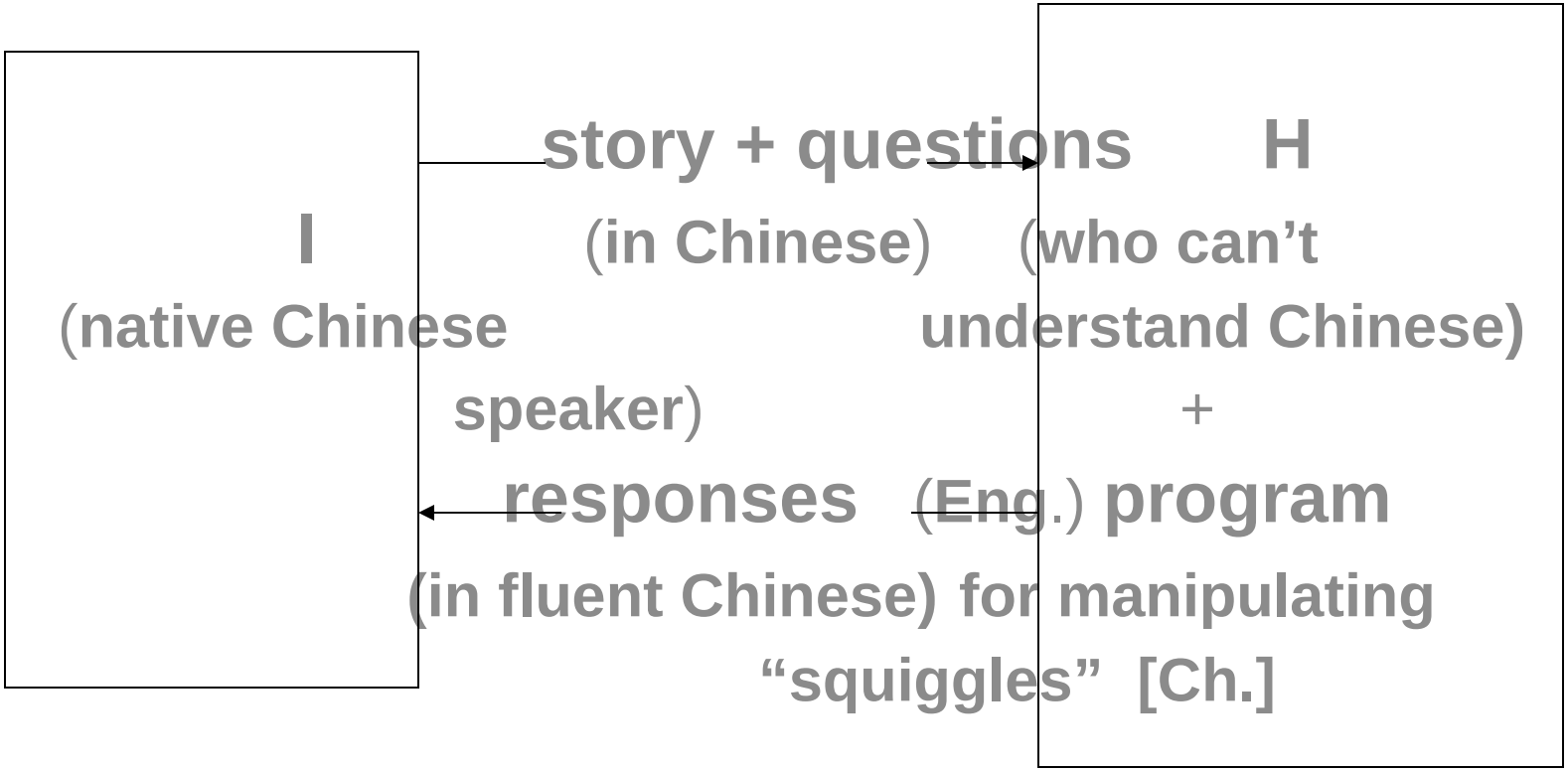
in · ten · tion · al · it · y: The property (in reference to a mental state) of being directed at a state of affairs in the world. For example, the belief that Sally is in front of me is directed at a person, Sally, in the world. Intentionality is sometimes taken to be synonymous with representation, understanding, consciousness, meaning, and semantics. Although there are important and subtle distinctions in the definitions of "intentionality," "understanding," "semantics," and "meaning," in this debate they are sometimes used synonymously.

© MacroVU Press


NATIVE CHINESE SPEAKER
CHINESE WRITING INPUT
- script
- story
- questions
CHINESE WRITING OUTPUT
- answers
SEARLE
ENGLISH PROGRAM
月片
美味的
PROG
CHINESE ROOM

The Chinese-Room Argument

It's possible to pass Turing Test, yet not (really) think



Argument from biology:

(a) Computer programs are non-biological

(b) Cognition is biological

(c) ∴ No non-biological computer program can exhibit biological cognition.

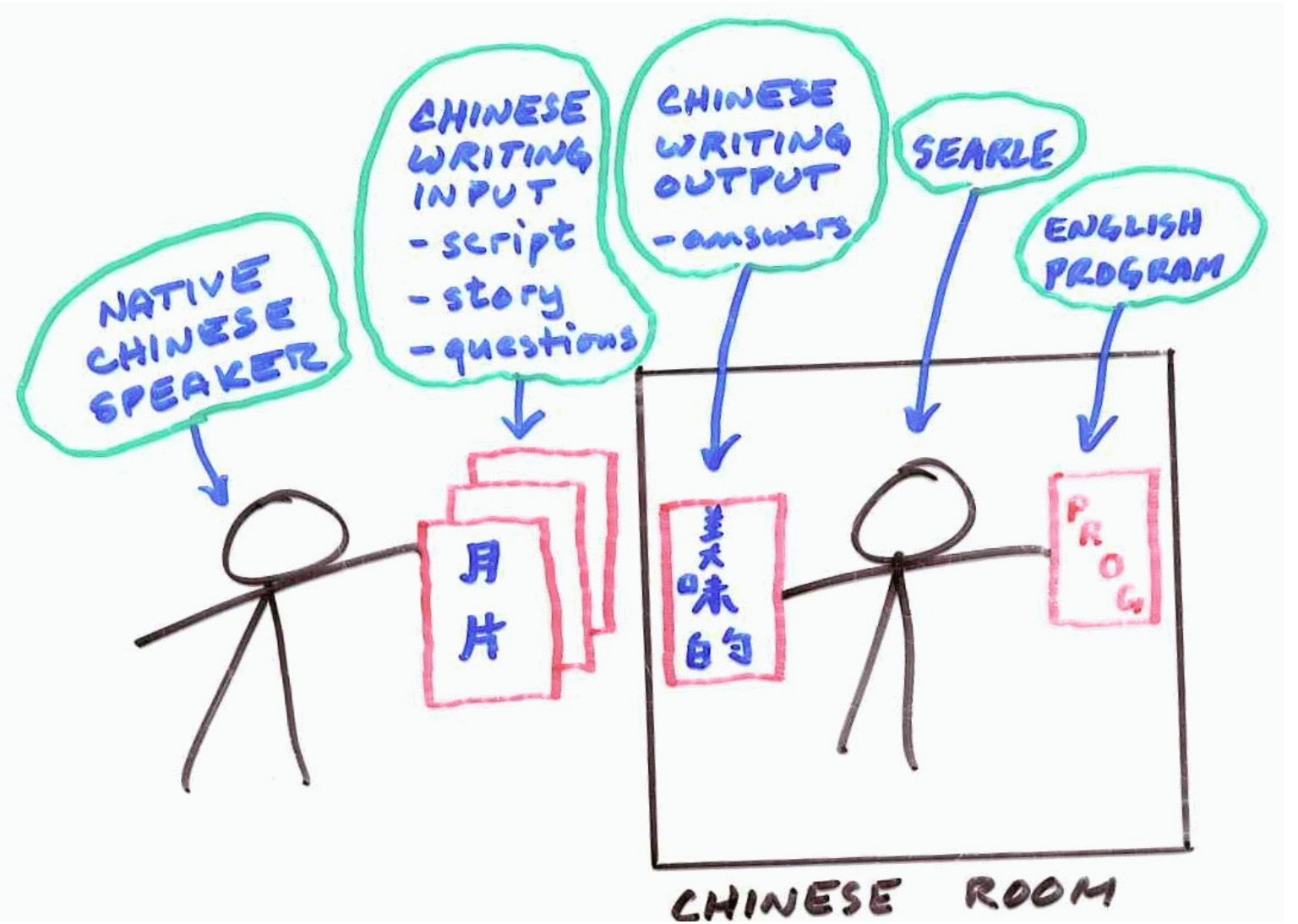
Argument from semantics:

(a) Computer programs are purely syntactic

(b) Cognition is semantic

(c) Syntax alone is not sufficient for semantics.

(d) ∴ No purely syntactic computer


NATIVE CHINESE SPEAKER
CHINESE WRITING INPUT
- script
- story
- questions
CHINESE WRITING OUTPUT
- answers
SEARLE
ENGLISH PROGRAM
月片
美味的
PROG
CHINESE ROOM