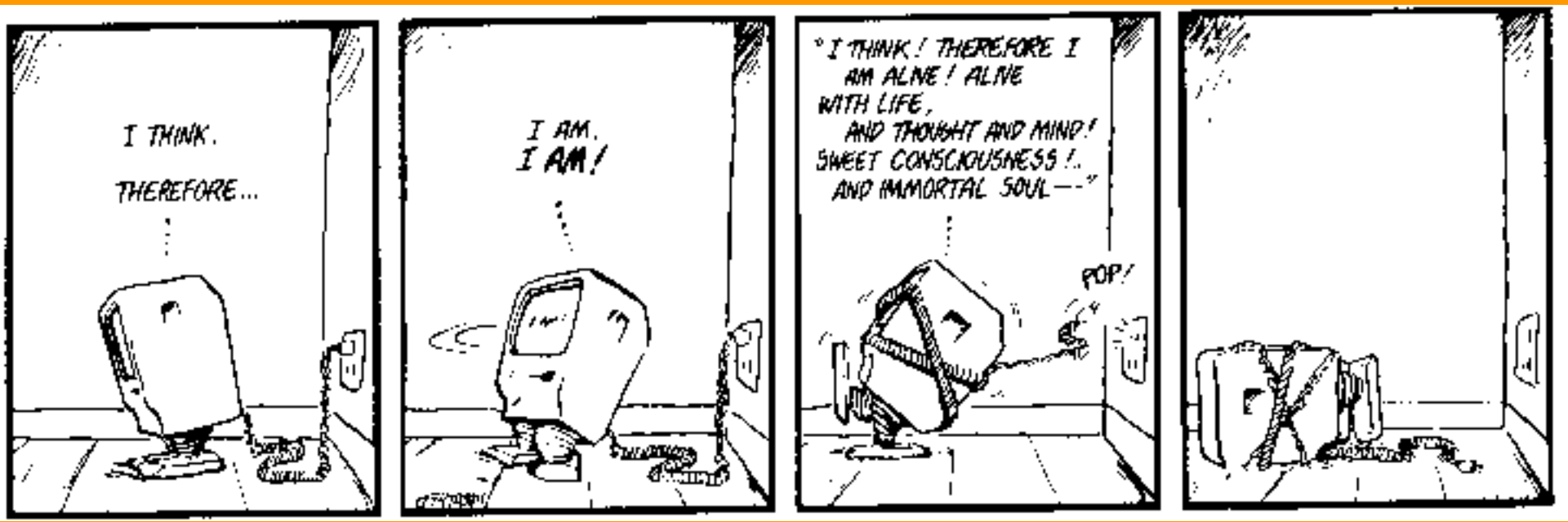
Bloom County on Strong AI



THE CHINESE ROOM

- **Searle's target: "Strong AI"**
  - **An appropriately programmed computer *is* a mind—capable of understanding and other propositional attitudes.**
    - **Schank's "story understander" program *really understands* the stories.**

The Gedankenexperiment

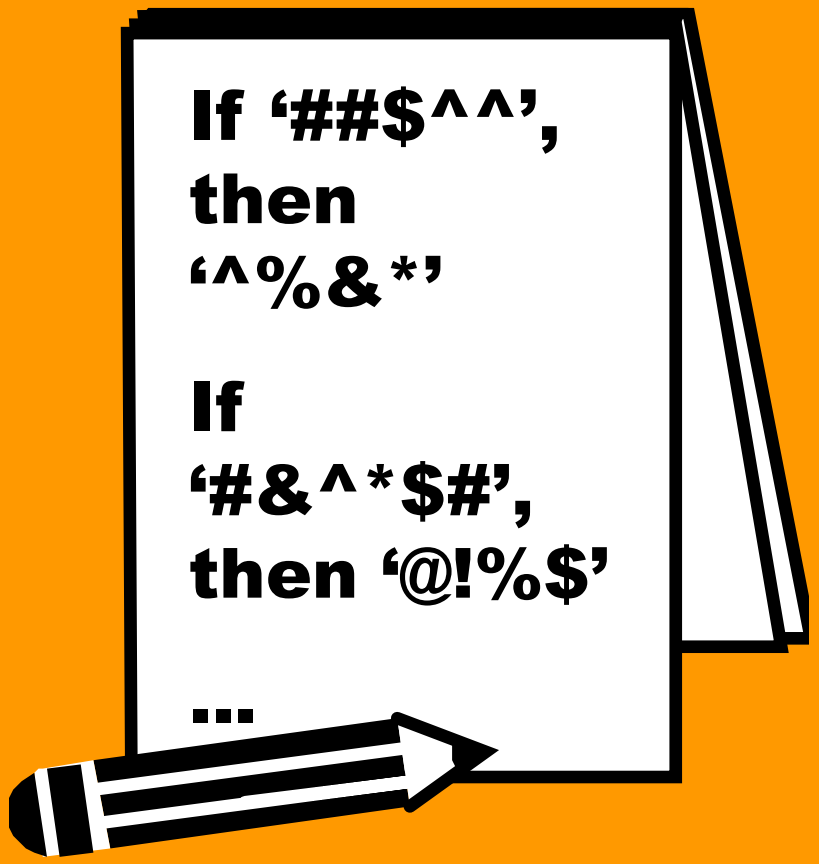
- **Searle, who knows no Chinese, is locked in a room with an enormous batch of Chinese script.**

The Gedankenexperiment

- **Slips of paper with still more Chinese script come through a slot in the wall.**

The Gedankenexperiment

- Searle has been given a set of rules in English for correlating the Chinese script coming through with the batches of script already in the room.



The Gedankenexperiment

- Searle is instructed to push back through the slot the Chinese script with which the scripts coming in through the slot are correlated according to the rules.

The Gedankenexperiment

- But Searle, remember, knows no Chinese; he identifies the scripts coming in and going out on the basis of their shapes alone. And following the rules requires only this ability.

The Gedankenexperiment

- **Suppose that those outside the room call the scripts going in ‘the questions’, the scripts coming out ‘the answers’, and the rules that Searle follows ‘the program’.**

The Gedankenexperiment

- Suppose also that the program writers get so good at writing the program and Searle gets so good at following it that Searle's answers are indistinguishable from those of a native Chinese speaker.

The result

- **It seems clear that Searle nevertheless does *not* understand the questions or the answers; he is as ignorant as ever of Chinese.**

"Neih hou ma?"

The result

- **But Searle is behaving just a computer does, “performing computational operations on formally specified elements”**

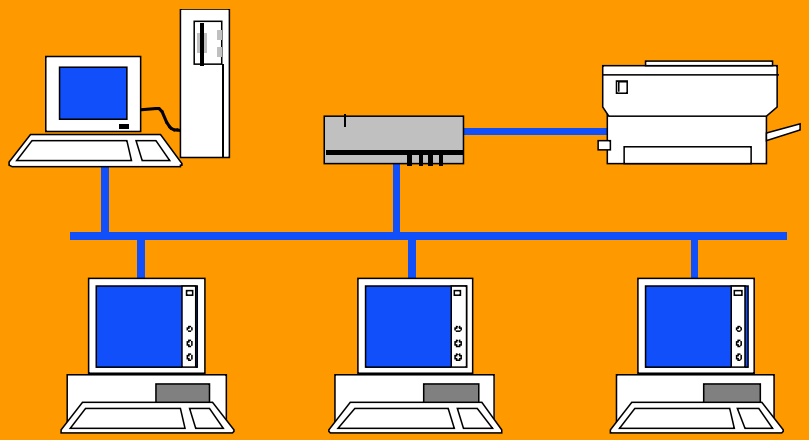
The result

- **Hence, manipulating formal symbols—which is just what a computer running a program does—is not sufficient for understanding or thinking.**

On the way to the replies

- **‘Understanding’—it comes in degrees, it’s vague…**
- **Searle: But I’m a *clear case* of understanding English sentences, and a *clear case* of failing to understand Chinese ones.**

The systems reply



- Of course the guy in the room doesn't understand. But the guy is just part of a larger system which does understand.

Systems reply: Searle's 1st rejoinder

- Let the guy (Searle) internalize all the elements of the system.
- But the guy still doesn't understand and neither does the system because there is nothing in the system that is not in him—the system is just a part of him.

Systems reply: a response to the 1st rejoinder

- Searle's rejoinder seems to rely on the following invalid principle: If I'm not F, and x is a part of me, then x is not F either.
- But I am not less than 5lbs in weight, but my heart, which is part of me, might be.

Systems reply: Searle's 2nd rejoinder

- **The idea is that, although the guy doesn't understand Chinese, somehow the *conjunction* of the guy plus the rulebook, the batches of script, the slot, and the room does understand. That's ridiculous.**

Systems reply: Searle's 2nd rejoinder

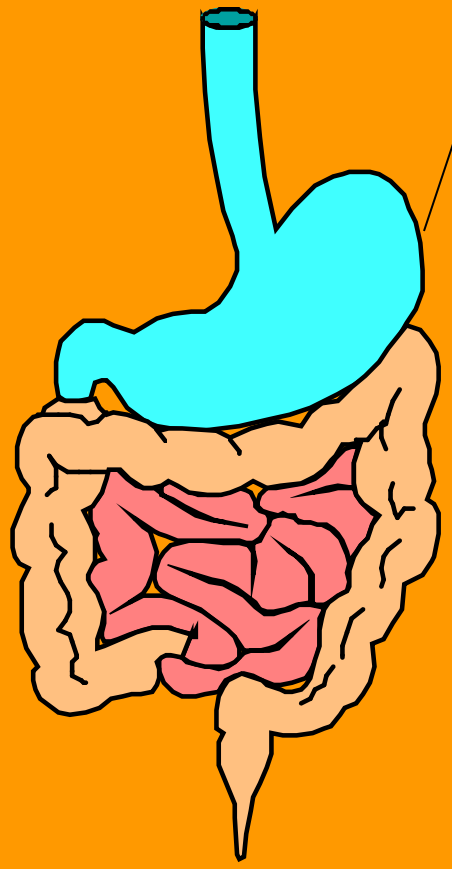
- Knowing what 'Do you want a hamburger?' means.
- One necessary condition: You have to know that 'hamburger' refers to hamburgers.
- How on earth does the Chinese Room system come to know facts like these?

Systems reply: Searle's 3rd rejoinder

- If the system counts as cognitive simply because it has certain inputs and outputs and a program in between, then all sorts of non-cognitive systems are going to count as cognitive.

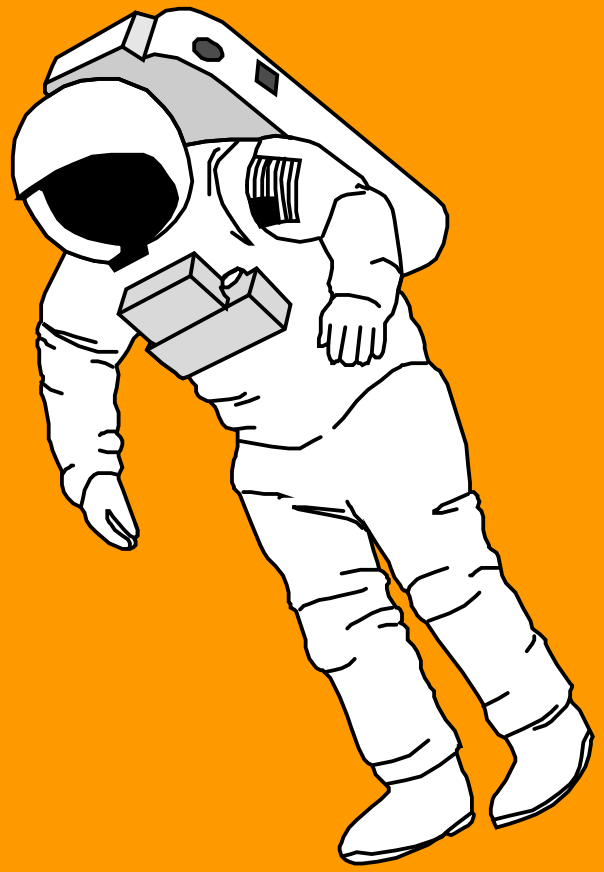
Systems reply: Searle's 3rd rejoinder

- **Just an interesting result of the strong AI model?**
- **Can't then explain what *makes* the mental mental.**



I understand too? Yippee!

The robot reply



- **Put a computer in the head of a robot giving the computer "perceptual" and motor capacities—this will bestow understanding and intentionality.**

Robot reply: Searle's rejoinder

- Note that the reply concedes that manipulating formal symbols does not add up to understanding.
- Put the Chinese Room in the head of a robot. Still no understanding.

Robot reply: Fodor's defense

- **"Given that there are the right kinds of causal linkages between the symbols that the device manipulates and things in the world, it is quite unclear that intuition rejects ascribing propositional attitudes to it."**

Robot reply: Fodor's defense

- **What Searle has shown (though we already knew it) is that *some* kinds of causal linkages are not the right kinds.**
- **It is a fallacy to suppose from this that *no* kinds of causal linkages could bestow understanding.**

Robot reply: Fodor's defense

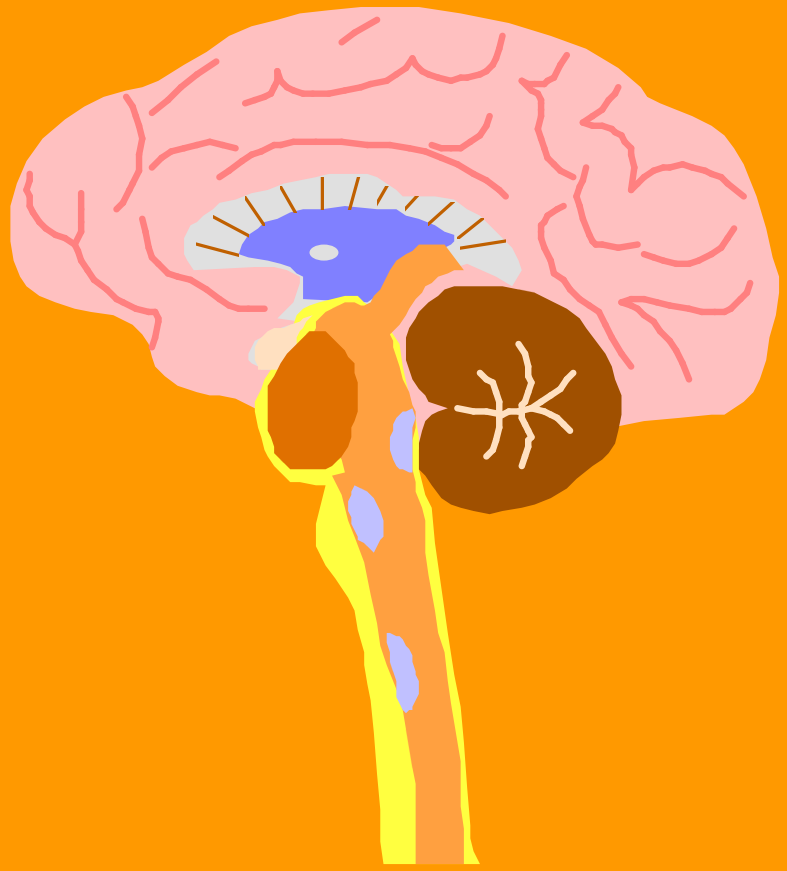
- The analogy with perception: 'A perceives x' is true only if there is some sort of causal relation btw x and A.
- The demonstration that some causal relations are not the required sort is irrelevant.

Robot reply: Searle's response to Fodor

The symbols can be connected to the world *any way you please* short of intentionality being essentially involved in the connection—still no understanding or propositional attitudes.

Brain simulator reply

- **Design a program that simulates the actual sequence of neuron firings that occur in a native Chinese speaker when he/she is understanding a story--wouldn't this do it?**

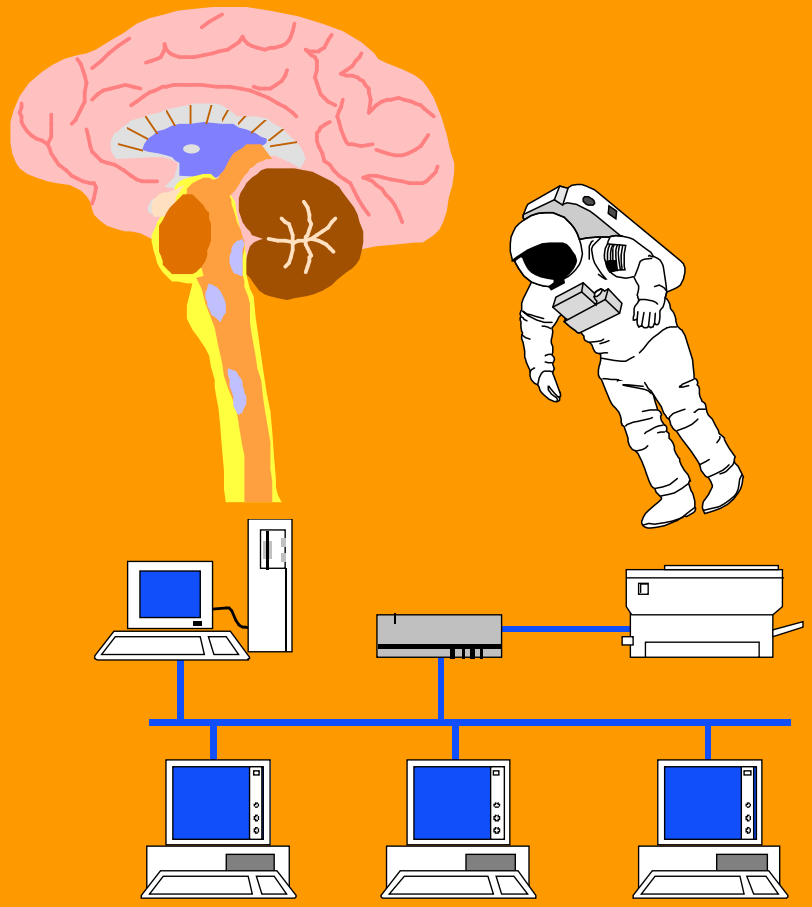


Brain simulator reply: Searle's rejoinders

- The reply concedes something important: that we might need to know how the brain works in order to know how the mind works. But this seems like an abandonment of Strong AI.
- Chinese plumbing in the Chinese room: Still no understanding.

Combination reply

- **Simulated brain in the head of a robot comprising a thinking system.**



Combination reply: Searle's rejoinder

- We would find it "irresistable" to attribute intentionality to such an entity. But that has entirely to do with its behavior. Once we learned how the entity functioned we'd withdraw our attributions.

Other minds reply

- If we refuse to attribute intentionality to the robot, won't we have to refuse to attribute intentionality to other people?
- Rejoinder: This confuses an epistemological problem with a metaphysical one.

Many mansions reply

- **Searle's argument is helped along by the fact that present technology is limited. Surely *some day* we'll be able to artificially reproduce intelligence and cognition**

Many mansions reply: Searle's rejoinder

- **This trivializes the project of Strong AI.**