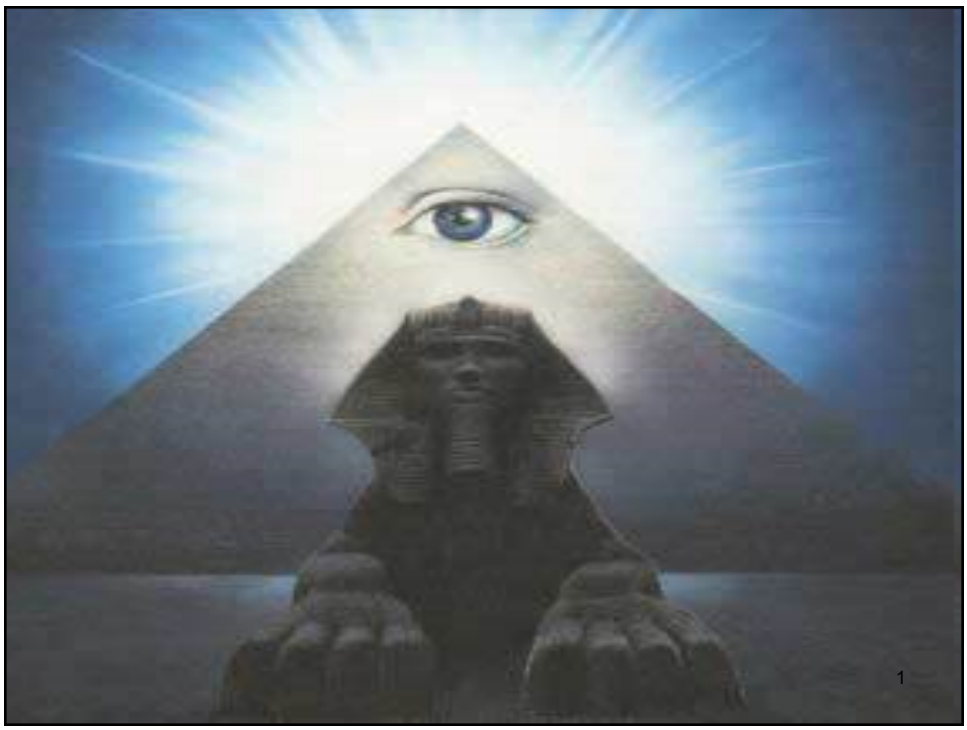


AI vs. human I

AI has seen great advances of many kinds recently, but there is one critical area where progress has been extremely slow: ordinary commonsense.

BY ERNEST DAVIS AND GARY MARCUS

- "The result is the pattern of progress often seen in AI: Rapid progress at the start of research up to a **mediocre level**, followed by slower and slower improvement."
- **Might we be missing something big?**

Current computer programs to carry out language tasks succeed to the extent the tasks can be carried out purely in terms of manipulating individual words or short phrases, without attempting any deeper understanding.

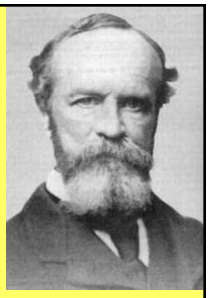
Introduction to Consciousness

in mammals and machines

1. Attention and consciousness: what are we talking about?
2. "Can machines be conscious?"
   1. Some functions NOT necessary for consciousness
   2. Meet the new Turing Test (same as the old)
   3. The importance of being integrated
3. Epiphenomenon or useful function?
4. Chinese Room Argument: is understanding (or consciousness in any form) reducible to an algorithm?

3

Attention chooses what enters consciousness



"Everyone knows what attention is. It is the taking possession by the mind, in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought. Focalization, concentration, of consciousness are of its essence. It implies withdrawal from some things in order to deal effectively with others…"

—William James

Principles of Psychology 1890

4

OK, but what's **consciousness?**

Giulio Tononi

Everybody knows what consciousness is: it is what vanishes every night when we fall into dreamless sleep and reappears when we wake up or when we dream. It is also all we are and all we have: lose consciousness and, as far as you are concerned, your own self and the entire world dissolve into nothingness.

Yet almost everybody thinks that understanding consciousness at the fundamental level is currently beyond the reach of science. The best we can do, it is often argued, is gather more and more facts about the neural correlates of consciousness—those aspects of brain function that change when some aspects of consciousness change—and hope that one day we will come up with an explanation. Others are more pessimistic: we may learn all about the neural correlates of consciousness and still not understand why certain physical processes seem to generate experience while others do not.

5

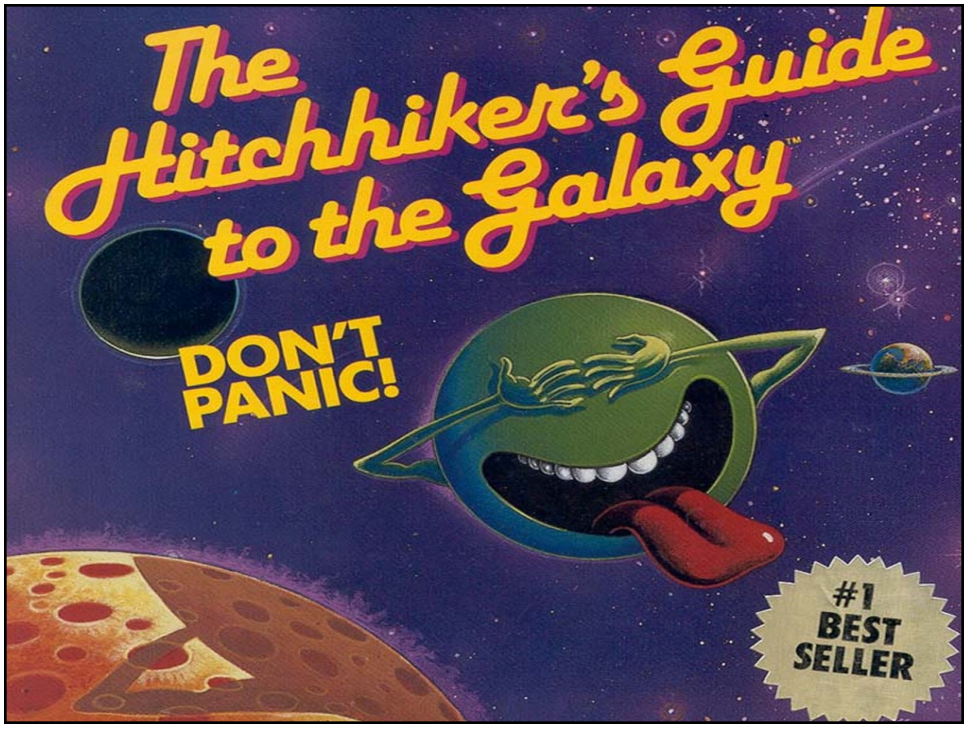
Synonyms and examples

- Awareness, visual awareness, seeing, hearing
- Experience, subjective experience
- Perception, sensory perception
- Thinking, feeling
- Remembering (explicit, declarative), imagining
- Being awake, dreaming
- Phenomenology, qualia (in philosophy)
- Mind? Cognition? Conscious vs unconscious mind

Motivations for trying to understand consciousness scientifically

- Technological: AI failures and limitations
- Medical/societal motivations:
  - Vegetative states
  - Anesthesia failures
  - Abortion debate
  - Animal (robot?) rights
- **Fundamental: what <u>are</u> we?**

7

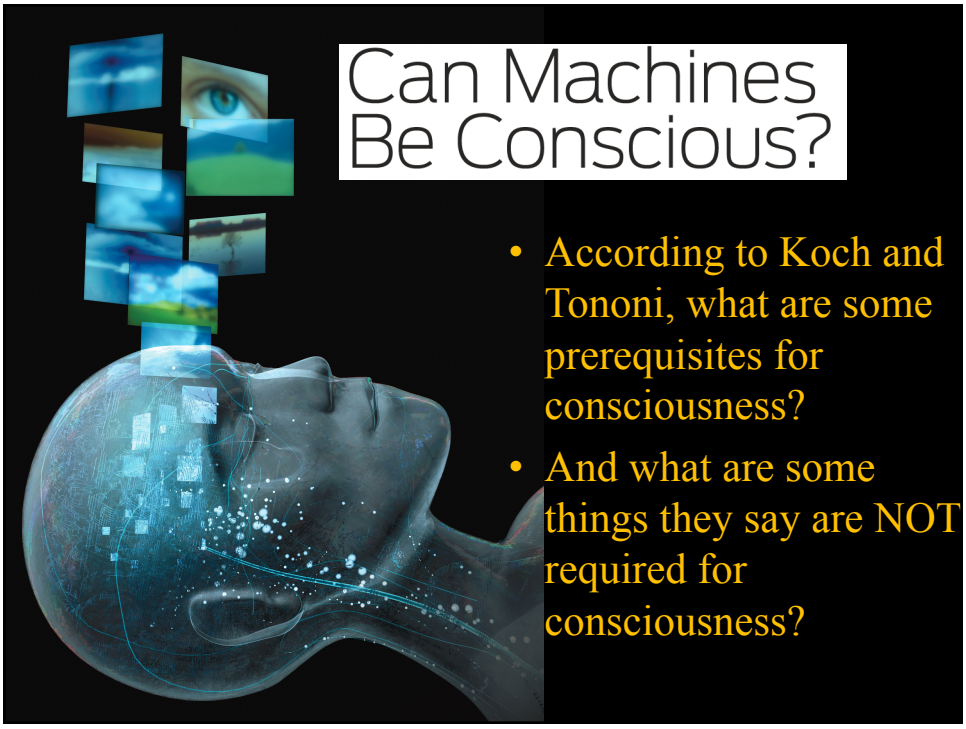


Introduction to Consciousness

in mammals and machines

1. Attention and consciousness: what are we talking about?
2. "Can machines be conscious?"
    1. Some functions NOT necessary for consciousness
    2. Meet the new Turing Test (same as the old)
    3. The importance of being integrated
3. Epiphenomenon or useful function?
4. Chinese Room Argument: is understanding (or consciousness in any form) reducible to an algorithm?

9



Information is:
how many possibilities get ruled out

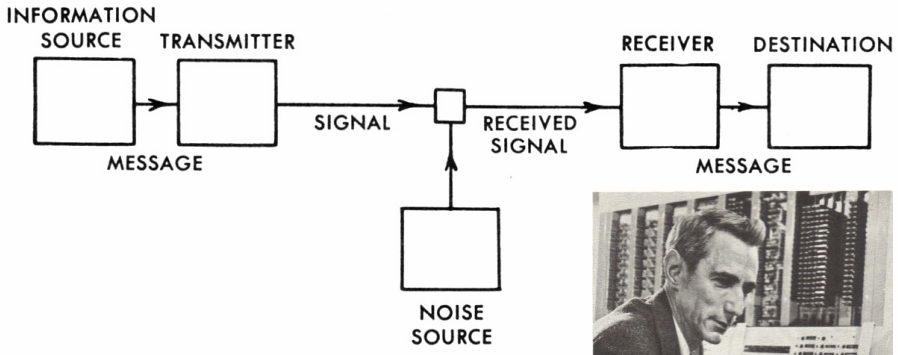
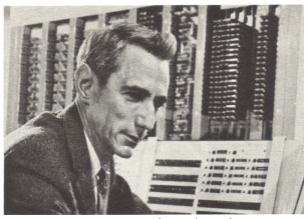


Fig. 1. — Schematic diagram of a general communication system.



Claude Shannon
1916-2001