

Cognitive Neuroscience for AI Developers

Week 14– Memory



Memory

- **Memory:** the brain's ability to encode, store, and retrieve data or information
 - retention of information over time for the purpose of influencing future action
 - If past events could not be remembered, it would be impossible for language, relationships, or personal identity to develop

Memory as an information processing system:

Sensory processor -> short-term / working memory -> long-term memory

Memory as information processing

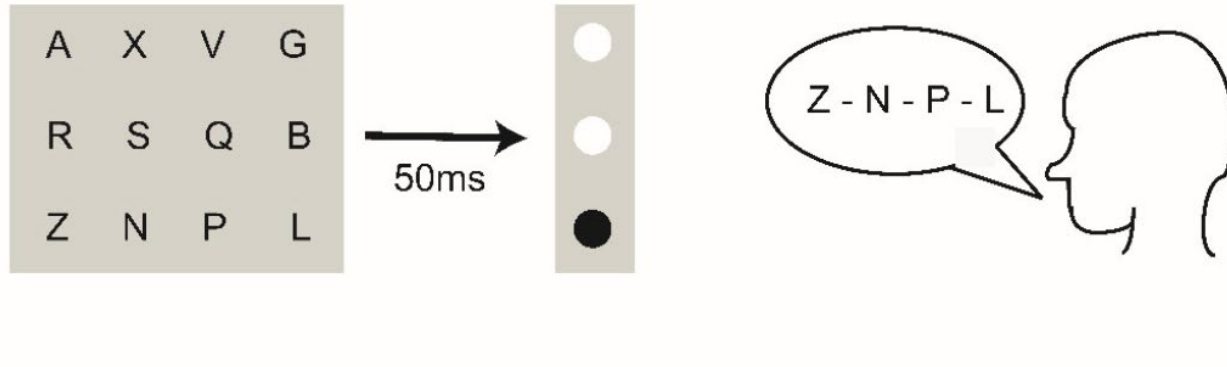
- **Encoding:** Processing of incoming information that creates memories
 - **Acquisition:** Bringing sensory stimuli which are stored in sensory buffer to short-term memory
 - **Consolidation:** Changes in the brain are stabilized to form long-term memories
- **Storage:** Permanent record of information
- **Retrieval:** Accessing stored information for example to generate conscious representation

Memory: Types and Classification

- **Types of memory (potentially mediated by different neural mechanisms):**
 - **Sensory memory (unconscious)**
 - **Short-term memory (conscious)**
 - **Working memory (conscious)**
 - **Long term memory**
 - Declarative (explicit memory): conscious (semantic, episodic)
 - Non-declarative (implicit memory): unconscious (procedural, perceptual, conditioning, non-associative)

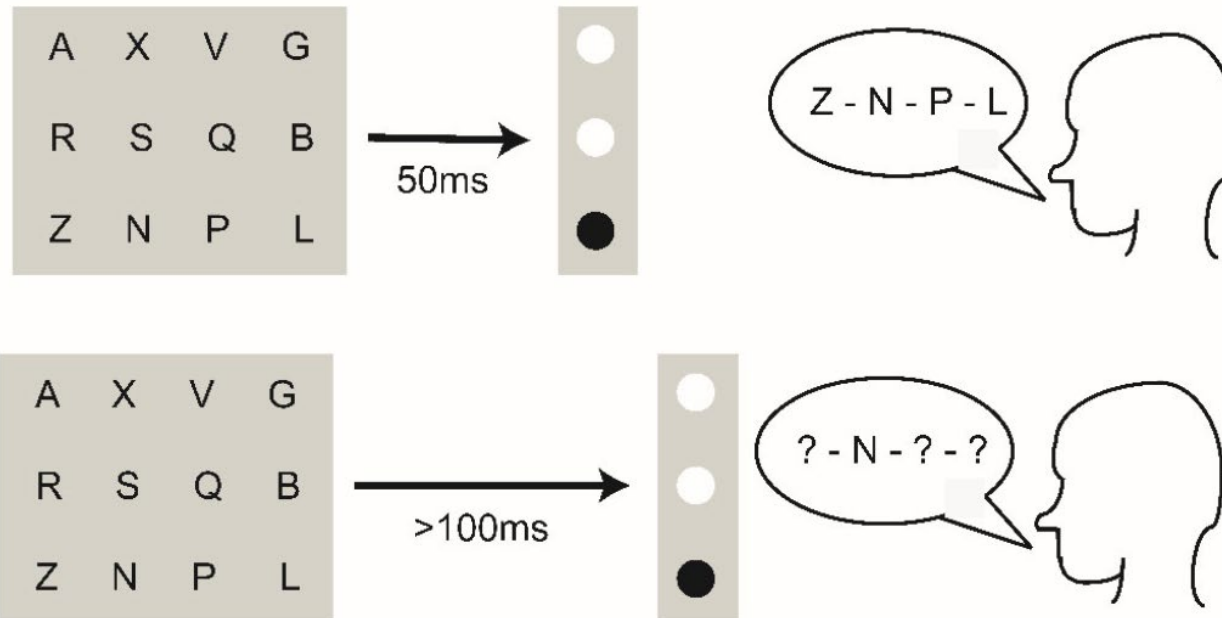
Sensory Memory holds information from sensory organs $<1s$

partial report paradigm by George Sperling (1963)



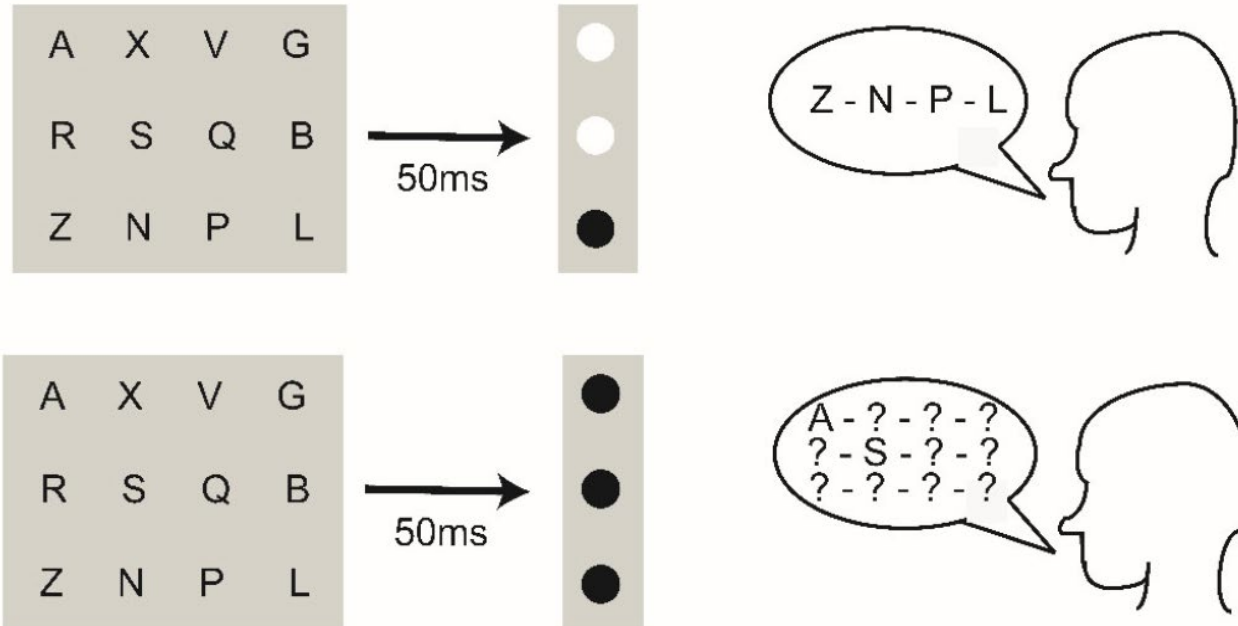
Sensory Memory holds information from sensory organs $<1s$

partial report paradigm by George Sperling (1963)



Sensory Memory holds information from sensory organs $<1s$

partial report paradigm by George Sperling (1963)



Sensory Memory holds information from sensory organs $<1s$

types of sensory memories:

- Iconic: visual stimuli
- Echoic: auditory stimuli
- Haptic: touch stimuli

Short-term memory is the capacity for holding, but not manipulating, a small amount of information in mind

several seconds to a minute without rehearsal

very limited capacity: 4-5 items

Capacity can be increased by **chunking**

e.g. a phone number:

3471892341

10 digits

347 189 2341

3 chunks of 3-4 digits

Short-term memory is the capacity for holding, but not manipulating, a small amount of information in mind

several seconds to a minute without rehearsal

very limited capacity: 4-5 items

Capacity can be increased by **chunking**

e.g. a phone number:

3471892341

10 digits

347 189 2341

3 chunks of 3-4 digits

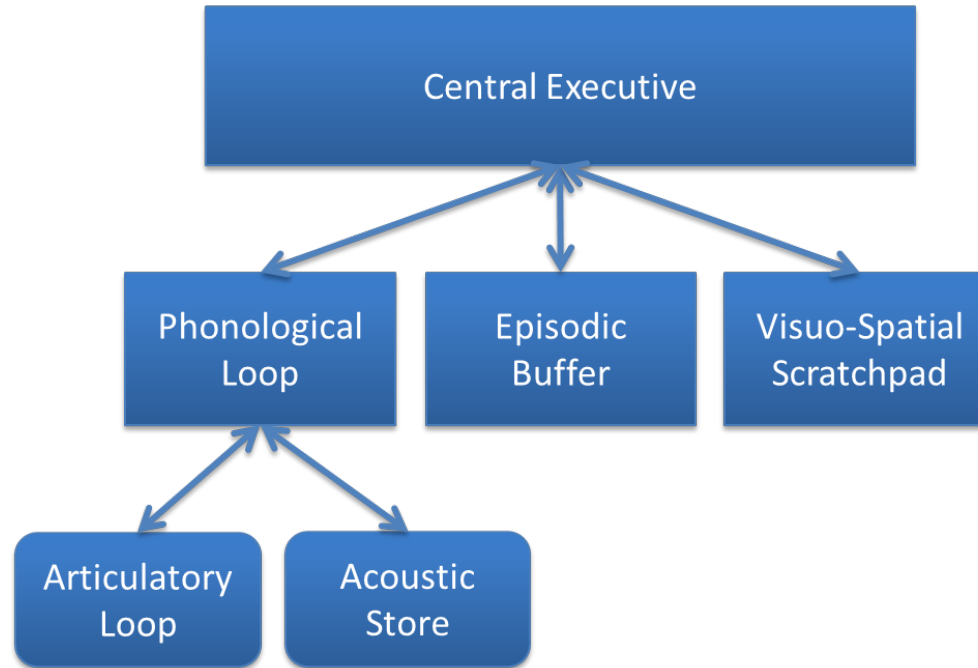
你好

Working memory is a cognitive system with a limited capacity that can hold information temporarily and enables manipulating information

important for reasoning and the guidance of decision-making and behavior

theoretical concept central to cognitive psychology, neuropsychology, and cognitive neuroscience

The multicomponent model of working memory by Baddeley and Hitch



Source: wikipedia.org

Long-term memory

Declarative / explicit memory: conscious recall

- semantic, episodic, autobiographic
- medial temporal lobes, hippocampi

Non-declarative / implicit memory: unconscious

- procedural, how to do sth?
- motor skill learning
- cerebellum, basal ganglia

Memory: Amnesia (Loss of memory)

- If some of the brain structures are lesioned -> **amnesia**
- **Anterograde amnesia**: Loss of all memories after a lesion -> inability to form new memories
- **Retrograde amnesia**: Loss of memories before lesion (sometimes only temporal loss)

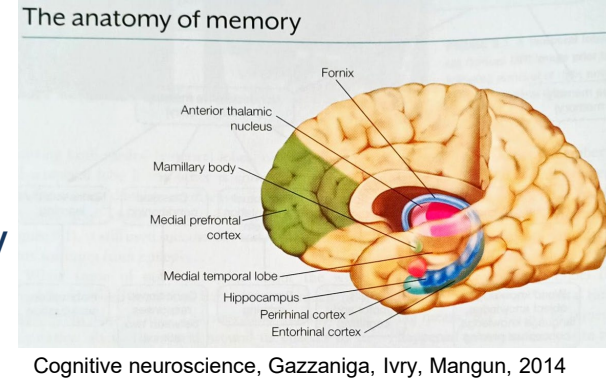
Henry Gustav Molaison, known widely as patient **H.M.**
1926-2008 -> Anterograde amnesia after removal of
hippocampi



Source: wikipedia.org

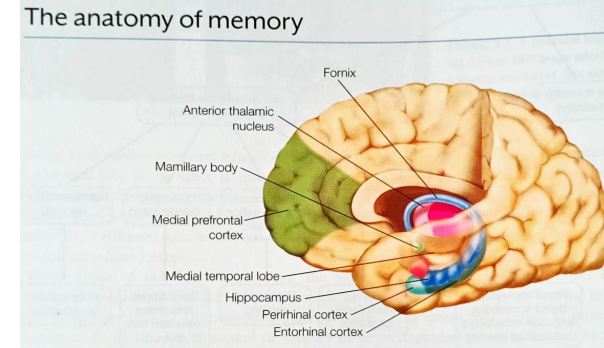
The Hippocampus

- **Encoding of information:**
 - Neuroimaging and behavioral studies show a key role of hippocampus to encode information
 - Basis to form long-term memories
- **Memory Retrieval:**
 - Neuroimaging studies also confirm that hippocampus is involved in retrieval of episodic memories (Do you remember that exact picture?)
 - Areas surrounding hippocampus are related to familiarity (Is a certain picture familiar to you?)



The Hippocampus: Memory Consolidation

- **Initial consolidation process**
 - Hippocampus important for rapid consolidation and initial storage of information for episodic and semantic memories
- **Slow permanent consolidation process:**
 - Under debate: sleep might play a crucial role (“**replay**” of memories during sleep)



Cognitive neuroscience, Gazzaniga, Ivry, Mangun, 2014

Cognitive Neuroscience for AI Developers

Week 14 – Free Will and Consciousness



What is Free Will?

What is Free Will?



What is Free Will?

the capacity for agents to choose between different possible courses of action unimpeded

- closely linked to the concepts of moral responsibility, praise, guilt, sin,
- the right to act outside of external influences or wishes
- the capacity to make choices undetermined by past events

What is Free Will?

the capacity for agents to choose between different possible courses of action unimpeded

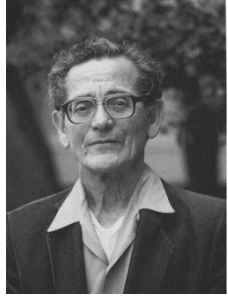
- closely linked to the concepts of moral responsibility, praise, guilt, sin,
- the right to act outside of external influences or wishes
- the capacity to make choices undetermined by past events
- **Determinism** suggests that only one course of events is possible

What is Free Will?

the capacity for agents to choose between different possible courses of action unimpeded

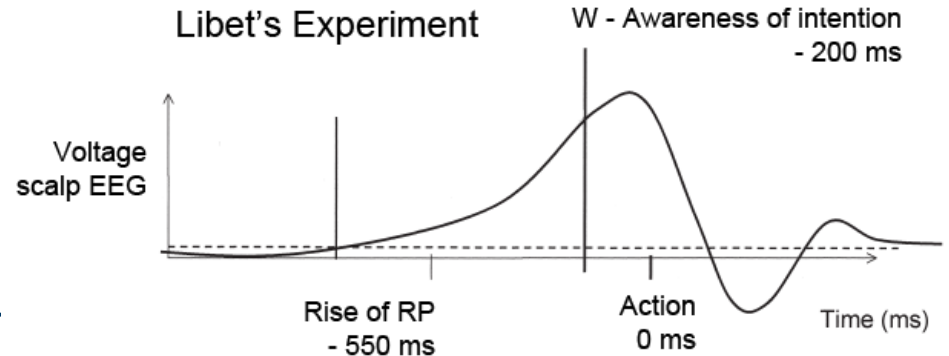
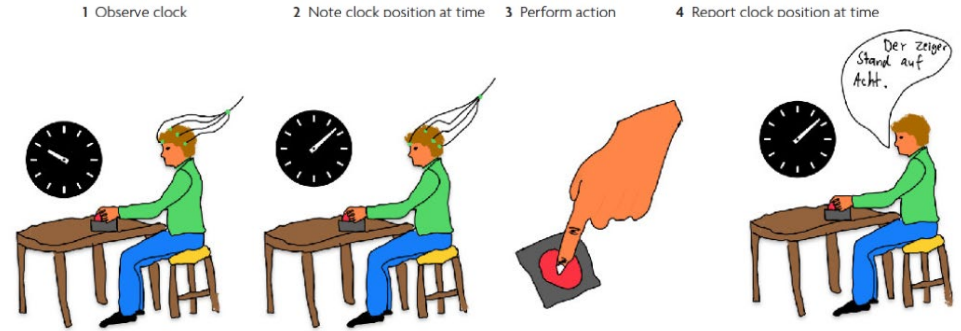
- closely linked to the concepts of moral responsibility, praise, guilt, sin,
- the right to act outside of external influences or wishes
- the capacity to make choices undetermined by past events
- **Determinism** suggests that only one course of events is possible
- **Incompatibilism**: determinism is true and thus free will is not possible
- **Compatibilism**: free will *is* compatible with determinism (deterministic chaos, quantum physics)

The Libet Experiment



Source: wikipedia.org
Benjamin Libet
1916-2007

unconscious brain activity of the readiness potential (RP) leading up to subjects' movements began approximately half a second **before** the subject was aware of a conscious intention to move

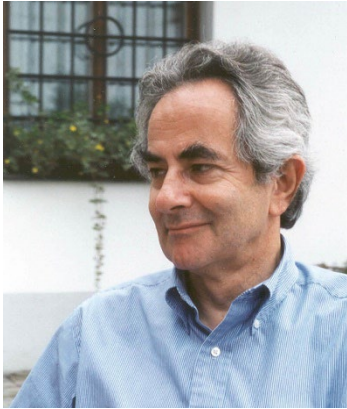


Source: informationphilosopher.com

What is Consciousness?

Consciousness is sentience or awareness of internal and external existence

"at once the most familiar and [also the] most mysterious aspect of our lives"
[Susan Schneider; Max Velmans (2008)]



Source: nytimes.com

Thomas Nagel

‘What is it like to be a bat?’

“... imagine that one has webbing on one’s arms, which enables one to fly around at dusk and dawn catching insects in one’s mouth; that one has very poor vision, and perceives the surrounding world by a system of reflected high-frequency sound signals; and that one spends the day hanging upside down by one’s feet in an attic. In so far as I can imagine this (which is not very far), it tells me only what it would be like for me to behave as a bat behaves. But that is not the question. I want to know what it is like for a *bat* to be a bat.”



Thomas Nagel, ‘What is it like to be a bat?’ (1974)

The Global Workspace Theory of Consciousness



Source: peerj.com

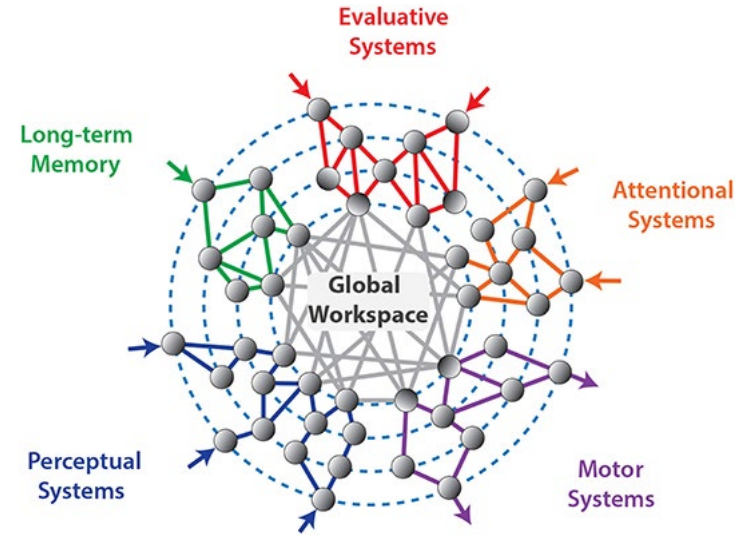
Bernard J Baars



Source: sciencesetavenir.fr

Stanislas Dehaene

- Information coming from sensory systems is integrated in a small group of brain regions and is then broadcasted to many different brain regions
- Information that is in global workspace becomes conscious
- Global workspace is a bright spot on working memory



Source: Krauss & Maier (2020).

The Integrated Information Theory of Consciousness



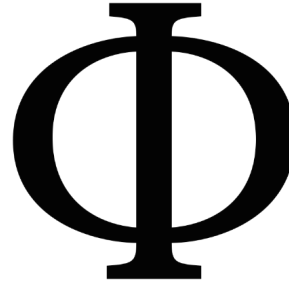
Source: wisc.edu

Giulio Tononi



Source: mit.edu

Christof Koch



Phi, the symbol used for integrated information

- Phi is called “synergy” -> extent to which a system is “more than its’ parts”
<https://blogs.scientificamerican.com/cross-check/can-integrated-information-theory-explain-consciousness/>
- Consciousness requires system with components which have physical “cause effect power on each other”
- Phi is maximized and serves as measure for degree of consciousness
<https://iep.utm.edu/integrated-information-theory-of-consciousness/>

Damasio's Theory of Consciousness

Emotions:

- collection of unconscious neural responses to external or internal stimuli
- cause observable changes in the organism
- act as neural objects, from which a physical reaction can be drawn



Source: wikipedia.org

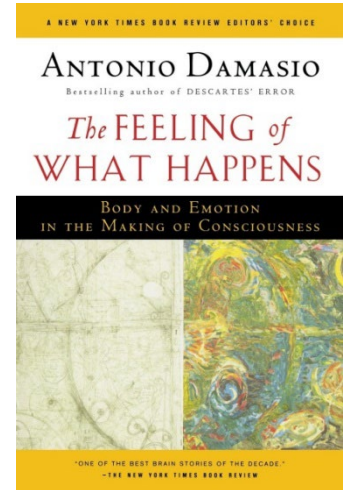
Antonio Damasio

Feelings (2nd order emotions):

- arises when organism becomes aware of the changes it is experiencing

hierarchy of stages:

- protoself
- core consciousness
- extended consciousness





Will We Ever Have Conscious Machines?

Patrick Krauss^{1,2} and Andreas Maier³*

¹ Neuroscience Lab, University Hospital Erlangen, Erlangen, Germany, ² Cognitive Computational Neuroscience Group, Chair of Linguistics, Friedrich-Alexander University Erlangen-Nürnberg (FAU), Erlangen, Germany, ³ Chair of Machine Intelligence, Friedrich-Alexander University Erlangen-Nürnberg (FAU), Erlangen, Germany

The question of whether artificial beings or machines could become self-aware or conscious has been a philosophical question for centuries. The main problem is that self-awareness cannot be observed from an outside perspective and the distinction of being really self-aware or merely a clever imitation cannot be answered without access to knowledge about the mechanism's inner workings. We investigate common

The Turing Test

I.—COMPUTING MACHINERY AND INTELLIGENCE

BY A. M. TURING

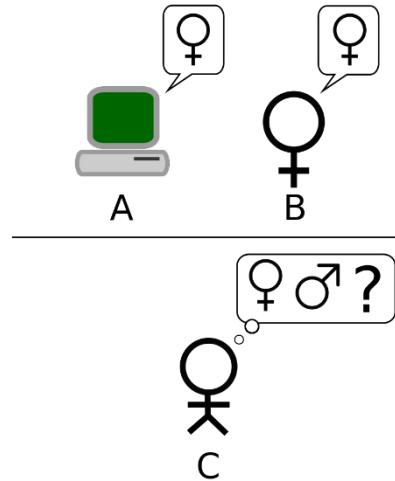


Source: faz.net

Alan Turing (1912-1954)

1. *The Imitation Game.*

I PROPOSE to consider the question, 'Can machines think?' This should begin with definitions of the meaning of the terms 'machine' and 'think'. The definitions might be framed so as to reflect so far as



Source: wikipedia.org

The Chinese Room

THE BEHAVIORAL AND BRAIN SCIENCES (1980) 3, 417-457

Printed in the United States of America

Minds, brains, and programs

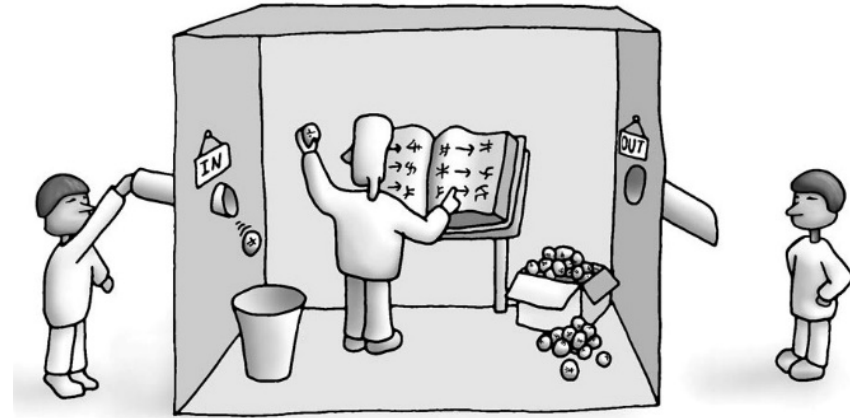


Source: wikipedia.org

John Searle

John R. Searle

*Department of Philosophy, University of California, Berkeley, Calif.
94720*



Source: <http://america.pink/images/9/6/3/2/5/4/en/2-chinese-room.jpg>

Thank you for your attention!

Do you have questions!