بسم الله الرحمن الرحیم

# علوم شناختی

## جلسه ۹

# فرضیه‌ی سیستم نماد فیزیکی

## The Physical Symbol System Hypothesis

کاظم فولادی قلعه
دانشکده مهندسی، دانشکدگان فارابی
دانشگاه تهران

http://courses.fouladi.ir/cogsci

# PART 2:
# MODELS AND TOOLS

CAMBRIDGE
UNIVERSITY PRESS

# Chapter 4:
## Physical Symbol Systems and the Language of Thought

CAMBRIDGE
UNIVERSITY PRESS

# Chapter 4.1: The physical symbol system hypothesis

# 1975 Turing Award

Given by Association of Computing Machinery (ACM) to Allen Newell and Herbert Simon – pioneers of AI

- **Logic Theory Machine** (1957)
- **General Problem Solver** (1956)

Newell and Simon used their Turing lecture to deliver a manifesto about the basic principles for studying intelligent information-processing.

# Laws of qualitative structure

Basic principles governing individual sciences

- **Biology**:
  The cell is the basic building block of organisms

- **Geology**:
  Geological activity results from the movement of a small number of huge plates

- **AI/Cognitive Science**:
  The physical symbol system hypothesis

# The physical symbol system hypothesis

A physical symbol system has
the necessary and sufficient means for intelligent action

- ***Necessity*:** Anything capable of intelligent action is a physical symbol system

- ***Sufficiency*:** Any (sufficiently sophisticated) PSS is capable of intelligent action

# Four basic ideas

(1) Symbols are physical patterns

(2) Symbols can be combined to form complex symbol structures

(3) The system contains processes for manipulating complex symbol structures

(4) The processes for representing complex symbol structures can themselves be symbolically represented within the system

# Thinking and the PSSS

- The essence of intelligent thinking is the ability to solve problems
- Intelligence is the ability to work out, when confronted with a range of options, which of those options best matches certain requirements and constraints
- Problem-solving is relative to a problem-space

# Specifying a problem in AI

Basic components of a representation

- description of given situation
- operators for changing the situation
- a goal situation
- tests to determine whether the goal has been reached

Problem space =
branching tree of achievable situations defined by potential application of operators to initial situation

- [e.g. chess]

# Problem-solving

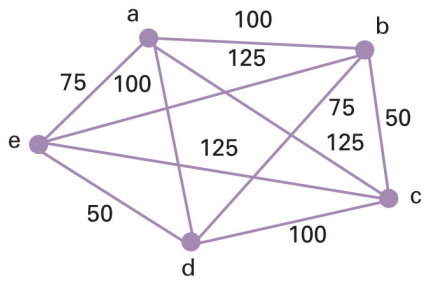
Problem-spaces are generally too large to be searched exhaustively (brute force algorithms)

Search must be selective ⇒ **heuristic** search rules

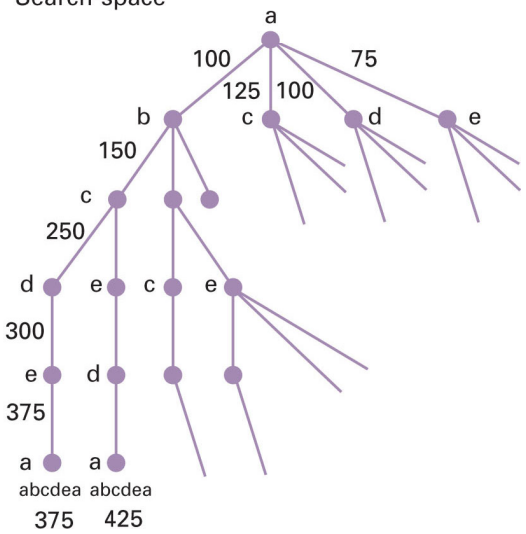
- effectively close off branches of the tree
  - e.g. in chess: "ignore branches that start with a piece being lost without compensation"

# Combinatorial explosion!

An instance of the traveling salesman problem




Search-space




- With *n* connected cities there are (*n* – 1)! possible paths through the search space
- This can be reduced to $2^n$
- But it would take a computer processing 1,000,000 possibilities per second over 30 years to solve a 50 city TP problem by brute force search

# Heuristic search hypothesis

Problems are solved by generating and modifying symbol structures until a solution structure is reached

- GPS starts with symbolic descriptions of the start state and the goal state
- aims to find a sequence of admissible transformations that will transform the start state into the goal state

# Heuristic search and algorithms

The PSSH is a reductive characterization of intelligence

- It is only illuminating if physical symbol systems are not themselves intelligent
- This means that the physical symbol systems must function algorithmically

# Illustration: Missionary and cannibals

Symbolic representation of state as *mcb*

*m* = number of missionaries on starting bank

*c* = number of cannibals on starting bank

*b* = number of boats on starting bank

Start state = 331

Goal state = 000

Permissible transformations?

# Permissible transformations

$m_i c_i b_k \Rightarrow m_{i+1} c_{i+1} b_{1-k}$ *where*

*either* difference between $m_i$ and $m_{i+1} = 2$ and
difference between $c_i$ *and* $c_{i+1} = 0$

*or* difference between $m_i$ and $m_{i+1} = 1$ and
difference between $c_i$ *and* $c_{i+1} = 1$

*or* difference between $m_i$ and $m_{i+1} = 0$ and
difference between $c_i$ *and* $c_{i+1} = 2$

*or* either the difference between $m_i$ and $m_{i+1}$ $= 1$ or the difference between $c_i$ *and* $c_{i+1} = 1$

# Impermissible states

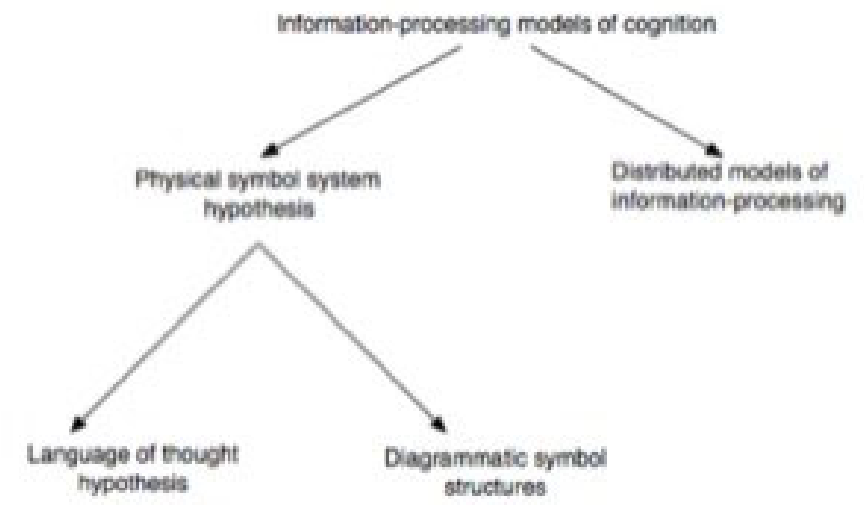
A branch ends if it reaches a state *mcb* where

$c > m$ [more cannibals than missionaries on R bank]

$(3 - c) > (3 - m)$ [more cannibals on L bank]

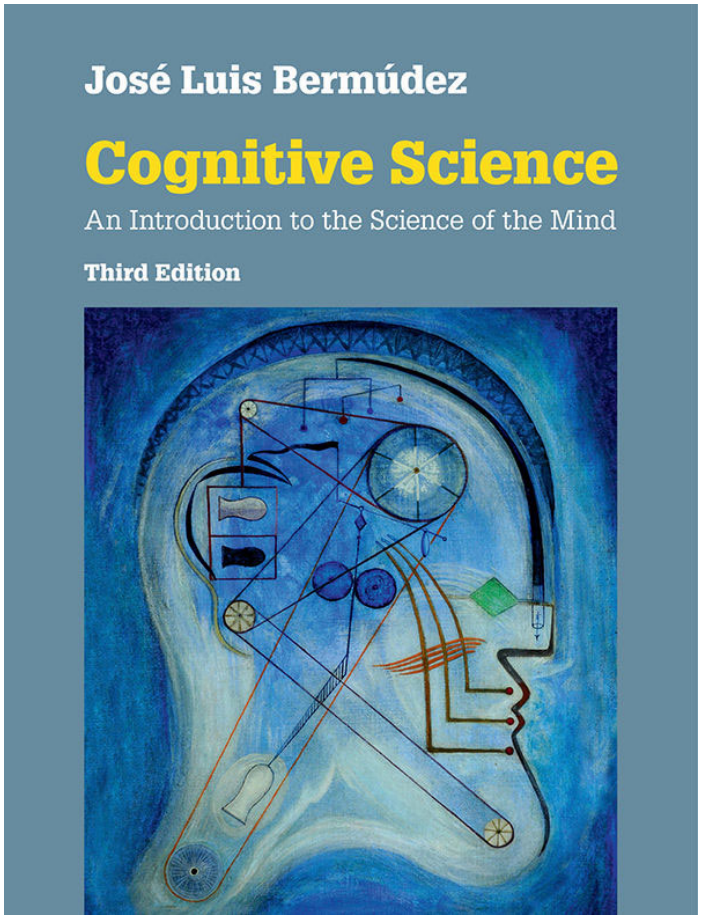
*mcb* has already appeared earlier in the tree

# The overall lie of the land

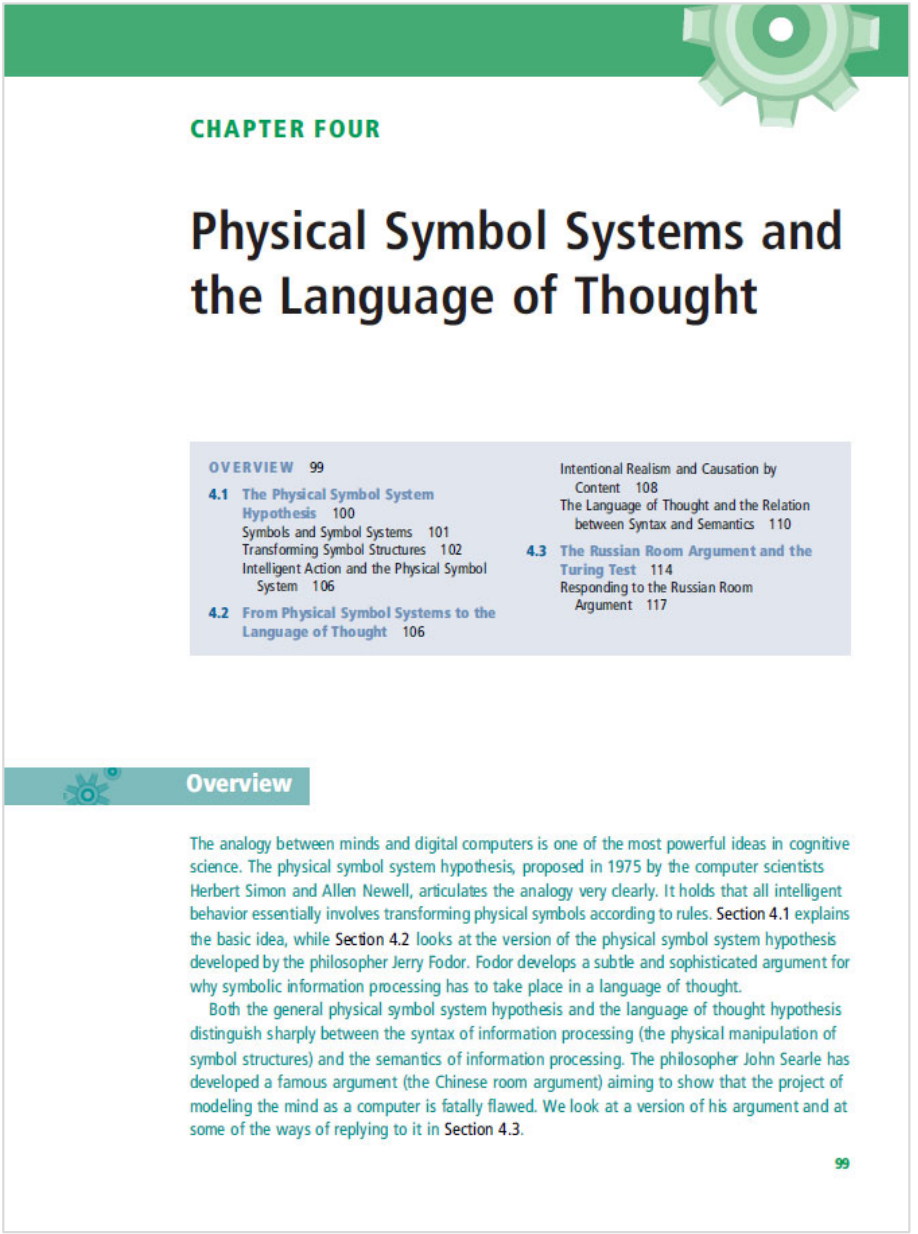



- The language of thought hypothesis is a specific proposal for developing the PSSH
- The example of WHISPER shows that symbol structures can be pictorial
- The contrast class for the PSSH is the class of neural network (connectionist) models

# منبع اصلی




José Luis Bermúdez,
**Cognitive Science:**
**An Introduction to the Science of the Mind,**
3rd ed., Cambridge University Press, 2020.
**Chapter 4 (Section 4.1)**



CHAPTER FOUR

## Physical Symbol Systems and the Language of Thought



### Overview

The analogy between minds and digital computers is one of the most powerful ideas in cognitive science. The physical symbol system hypothesis, proposed in 1975 by the computer scientists Herbert Simon and Allen Newell, articulates the analogy very clearly. It holds that all intelligent behavior essentially involves transforming physical symbols according to rules. Section 4.1 explains the basic idea, while Section 4.2 looks at the version of the physical symbol system hypothesis developed by the philosopher Jerry Fodor. Fodor develops a subtle and sophisticated argument for why symbolic information processing has to take place in a language of thought.

Both the general physical symbol system hypothesis and the language of thought hypothesis distinguish sharply between the syntax of information processing (the physical manipulation of symbol structures) and the semantics of information processing. The philosopher John Searle has developed a famous argument (the Chinese room argument) aiming to show that the project of modeling the mind as a computer is fatally flawed. We look at a version of his argument and at some of the ways of replying to it in Section 4.3.

99