Cybernetics, AI, Cognitive Science and Computational Neuroscience: Historical Aspects

Péter Érdi
perdi@kzoo.edu

Henry R. Luce Professor
Center for Complex Systems Studies
Kalamazoo College
http://people.kzoo.edu/ perdi/

and

Institute for Particle and Nuclear Physics, Wigner Research Centre, Hungarian Academy of Sciences, Budapest
http://cneuro.rmki.kfki.hu/
1. Cybernetics
   - The Founding Fathers: Warren McCulloch and Norbert Wiener
   - John von Neumann: The Computer and the brain)
   - Second-order cybernetics
   - The dissolution/dissemination of cybernetic ideas

2. Artificial Intelligence: The physical symbol system hypothesis (Alan Newell and Herbert Simon)

3. Connectionism: From Perceptron to the Artificial Neural Network Movement

4. Hermeneutics of the brain

5. Schizophrenia - a broken hermeneutic circle

6. Back to Neurobiology: Computational Neuroscience

7. Cognitive Science

8. Cybernetics is back!
Cybernetics

The Founding Fathers: Warren McCulloch and Norbert Wiener

- logic-based physiological theory of knowledge
- the brain performs logical thinking ...
- ... which is described by logic
- therefore ... the operation of the brain could and should be described by logic
Cybernetics

The Founding Fathers: Warren McCulloch and Norbert Wiener

- “Behavior, Purpose and Teleology”

- Feedback control

“Control and Communication in the Animal and the Machine”

Voluntary nervous system may control the environment

- Theory of goal-oriented behavior: a new framework to understand the behavior of animals, humans, and computers
The Founding Fathers: Warren McCulloch and Norbert Wiener

The Macy conferences (1946 - 1953)

- Applicability of a Logic Machine Model to both Brain and Computer
- Analogies between Organisms and Machines
- Information Theory
- Neuroses and Pathology of Mental Life
- Human and Social Communication
Cybernetics

John von Neumann: The Computer and the Brain

- Set Theory
- Quantum Mechanics
- Game Theory
- Computer Architectures
- The Computer and the Brain
John von Neumann: Set Theory

- Russel - Whitehead: the crisis of mathematics

- Hilbert’s program: to formalize mathematics (the axiomatization of set theory)

- von Neumann-Bernays-Gödel set theory: (has only finitely many axioms)

- abandoned after Gödel’s theorems (1931)
  "no formal system powerful enough to formulate arithmetic could be both complete and consistent"
Von Neumann architecture

The IAS Computer was named for the Institute for Advanced Study, Princeton. The machine was built there under the direction of John von Neumann. It cost several hundred thousand dollars. The goal of developing the IAS was to make digital computer designs more practical and efficient. The project to build it began in 1946 and the computer was ready for use in 1952.

Designers of the IAS were required to make their plans available to several other government-funded projects, and several other machines were built along similar lines: JOHNNIAC, MANIAC, ORDVAC, and ILLIAC. -> also IBM 701 (from the website of the Smithsonian Museum)

- has three basic hardware subsystems: a CPU, a main-memory system and an I/O system
- is a stored-program computer
- carries out instructions sequentially
- has, or at least appears to have, a single path between the main memory system and the control unit of the CPU; this is often referred to as the von Neumann bottleneck
The Computer and the Brain
The Computer and the Brain in broader context

- McCulloch-Pitts and the Cybernetic movement
- Self-replicating automaton: the machine and its description
- Reliable calculation with unreliable elements
- Analog vs. digital machines
- Specialized memory unit
- Language of the brain: "Thus the outward forms of our mathematics are not absolutely relevant from the point of view of evaluating what the mathematical or logical language truly used by the central nervous system is"
From Cybernetics to AI and back?

- August 1955: coincidence:
- von Neumann was diagnosed with cancer
- A proposal for the Dartmouth summer research project on Artificial Intelligence
- mechanism vs. function
- zeros and ones vs. general symbol manipulation
From Cybernetics to AI and back?

While it seemed to be an analogy between Brain and Computer

+ at the elementary hardware level

+ at the level of mathematical (quasi)-equivalence

the Organization Principles are very different

The importance of the actual biological substrate:

Synaptic organization !!

"...Eccles has shown how excitation and inhibition are expressed by changes of membrane potential."
Cybernetics

Second-order cybernetics

- autonomous system, role of observer, self-referential systems
- Heinz von Foerster (1911–2002)
- radical constructivism
- knowledge about the external world is obtained by preparing models on it

It is difficult to reconstruct the story, but it might be true that a set of cyberneticians, who felt the irreducible complexity of the system-observer interactions, abandoned to build and test formal models, and used a verbal language using metaphors. They were the subjects of well-founded critics for not studying specific phenomena (Heylighen and Joslyn 2001)
The dissolution/dissemination of cybernetic ideas

At least four disciplines have crystallized from cybernetics

• Biological control theory
• Neural modeling
• Artificial Intelligence
• Cognitive psychology

Preliminary crystal seeds:

Turing’s paper on computing machinery and intelligence (Turing 1950)

Rashevsky’s continuous neural models (Rashevsky 1933).
Artificial Intelligence: The physical symbol system hypothesis (Alan Newell and Herbert Simon)

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

- Symbols are physical patterns
- Symbols can be combined to form complex symbol structures
- The system contains processes for manipulating complex symbol structures
- The processes for representing complex symbol structures can themselves be symbolically represented within the system
Artificial Intelligence: The physical symbol system hypothesis (Alan Newell and Herbert Simon)

• Thinking and the PSSS

• Problem solving: heuristic search

  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
Connectionism: From Perceptron to the Artificial Neural Network Movement

**Perceptron**

- delta rule: difference between the actual and the expected output
- adjusting threshold and weights
- linearly separable functions
- Minsky - Papert
- multilayer perceptron: backpropagation

"A hermeneutic process is introduced as a peculiar mechanism of an information processing of the brain. Concerning a pattern recognition, the difference between the functions of the brain and of the present artificial intelligence is discussed. The concept of metaconscious is introduced to explain the hermeneutic process. In this relation, it is suggested that the existence of a hidden dynamics is of importance for man’s cognition.”
Hermeneutics of the brain

Chaotic brain, chaotic itinerancy, hermeneutics

There is no UNIVERSAL answer to the question whether in a given physiological context chaos can be associated to healthy or rather with pathological behaviour?

+ novelty filter
+ memory searcher
+ dynamic store-room of LTM
+ catalyst for learning
+ non-linear pattern classifier
+ STM generator

"...Chaotic itinerancy is a closed-loop trajectory through high-dimensional state space of neural activity that directs the cortex in sequence of quasi-attractors. A quasi-attractor is a local region of convergent flows (attractant, absorbent) giving ordered, periodic activity and divergent flows (repellant, dispersive) giving disordered, chaotic activity between the regions."
Tsuda: "Chaos enhanced learning ability"

[basin boundaries of memory representation change in phase space and reorganization of phase space is consequently achieved. This may imply hierarchical organization and a reorganization of memory.]

So far so good. But what is hermeneutics?

"A physicist friend of mine once said that in facing death, he drew some consolation from the reflection that he would never again have to look up the word "hermeneutic" in the dictionary."

(Steve Weinberg)

Hermeneutics: branch of continental (i.e. mainland European) philosophy which treats the understanding and interpretation of texts.

Hermeneutic circle: definition or understanding of something employs attributes which already presuppose a definition or understanding of that thing. The method is in strong opposition of the classical methods of science, which does not allow such kinds of circular explanations.
Hermeneutics of the brain

Circular and network causality
Systems with feedback connections and the systems of these connected loops can be understood based on the concepts of circular and network causality.

Neural implementation of circular causality
Functional macro-network for associative memory
Schizophrenia - a broken hermeneutic circle

Understanding situations: needs hermeneutic interpretation

- logic, rule-based algorithms, and similar computational methods are too rigid to interpret ill-defined situations,

- hermeneutics, "the art of interpretation" can do it.

- hermeneutics: emphasize the necessity of self-reflexive interpretation and adopts circular causality

To understand other minds: i.e. to show empathy is to simulate other minds.

The neural basis of theory of mind related to mirror neurons, which is the key structure of imitation, and possibly language evolution (Michael Arbib).

A failure in interpreting self-generated action generated by the patient himself: (lack of ability to close the hermeneutic circle) can be characteristic for schizophrenic patients (Chris Firth). -> Neural basis: disconnection syndrome
Schizophrenia - a broken hermeneutic circle

Disconnection hypotheses of schizophrenia

Geschwind’s (general) disconnection syndromes (1965)

The pathways implicated in the principle syndromes described by Geschwind, classified into three types: sensory-limbic disconnection syndromes (dotted lines), sensory-motor disconnection syndromes (dashed lines); sensory-Wernicke’s area disconnection syndromes (solid lines).
Schizophrenia - a broken hermeneutic circle

Disconnection hypotheses of schizophrenia

• impairments in functional macro-networks in schizophrenia was suggested

• abnormal prefronto-hippocampal connectivity?

• changes in effective connectivity: (i) intrinsic connectivity of the network, (ii) input-dependent changes

• Task related functional connectivity: during object - location associative learning
Schizophrenia - a broken hermeneutic circle

Neural implementation of circular causality

Circular and network causality

Systems with feedback connections and the systems of these connected loops can be understood based on the concepts of circular and network causality.

Functional macro-network for associative memory

The glutamate - dopamine interplay (Carlsson)
Schizophrenia - a broken hermeneutic circle

Functional Disconnectivities

- impairments in functional macro-networks in schizophrenia was suggested

- abnormal prefronto-hippocampal connectivity?

- changes in effective connectivity: (i) intrinsic connectivity of the network, (ii) input-dependent changes

- Task related functional connectivity: during object - location associative learning

Which connections are significantly impaired during schizophrenia? Quantitative estimation for the degree of impairment
Schizophrenia - a broken hermeneutic circle

Functional Disconnectivities

Schizophrenia fMRI study: experiment and methods

Task: learning of object-location associations over repeated encoding and retrieval periods

Subjects: 11 diagnosed with schizophrenia and 11 healthy controls

DCM: generative model of the BOLD signal, parameters estimated by Bayesian statistics

\[ \dot{x} = (A + \sum_{j=1}^{N} u_j B^i) x + C u \]

\[ y = \lambda(x, \theta_L) \]

\[ p(\theta|y, M) = \frac{p(y|\theta, M) p(\theta|M)}{p(y|M)} \]

Model space: five areas involved, two sets defined by varying connections and the effects of conditions

Model selection: by the estimation of the Bayesian evidence
Schizophrenia - a broken hermeneutic circle

Functional Disconnectivities

Schizophrenia fMRI study: results

Parameter level comparison: connections between PFC and HPC and HPC and IT are impaired

Model comparison: top-down information flow and the modulatory effects of conditions are less likely to be present in schizophrenia

Slow learning: differentiated from the illness by model probability distribution

Schizophrenia - a broken hermeneutic circle

Functional Disconnectivities

Paradigm shift in interpreting dysfunctions!

- ADHD as a brain network dysfunction
- Functional network dysfunction in anxiety and anxiety disorders
- A network dysfunction perspective on neurodegenerative diseases (e.g. Alzheimer)
- Brain network dysfunction in bipolar disorder
Back to Neurobiology: Computational Neuroscience

Top down

• starts from behavioral data

• information processing circuit

• implementation of neural mechanisms

Bottom up

• starts from anatomical and physiological reality

• rhythms

• behavior

Figure 1: Brain as a multi-level system. Neural mechanisms → computational algorithms.

[I should leave something for 2023]
History of cognitive science: preliminary remarks

Fundamental **philosophical** questions since Aristotle and Plato:

1. What is the nature of mind? Metaphysics and nature of reality
2. What is the nature of knowledge? Epistemology and nature of knowledge

**Psychology:** 1890s. Behaviorism: can’t study what is in the mind
(from "philosophical psychology" towards "experimental psychology")
1950’s. Miller, etc.: mind has structure

3. How do we think? **Neuroscience:**
4. How does the brain make a mind?

**Artificial intelligence:** 1956. Minsky, Newell, Simon, McCarthy

5. How to construct mind?

**Linguistics:** 1956. Chomsky versus behaviorist view of language.

6. Innateness?

**Anthropology:** social, cultural aspects of knowledge

7. Is there any cultural difference in the thinking of people?

The need of **INTEGRATION:**
How can all these fields, with different histories and methodologies can be integrated to produce an understanding of mind?
Key concepts

Mental representation and Computational procedures

Thinking = Mental representations + computational procedures

more precisely: Thinking = representational structures + procedures that operate on those structures. Analogy between computation and thinking:

data structures   mental representations

+ algorithms   + procedures

= running programs   = thinking

Methodological consequence: study the mind by developing computer simulations of thinking
Cognitive Science

Beyond the classical paradigm of cognitive science

- Emotions
- Embodied cognition
- Consciousness
- Dynamical systems
- Integrative approaches
- Multilevelism
## Multilevelism

After Paul Thagard

<table>
<thead>
<tr>
<th>Level</th>
<th>Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Molecular biology</td>
<td>Biochemical experiments and genetic models</td>
</tr>
<tr>
<td>Neuroscience</td>
<td>Experiments and computational models</td>
</tr>
<tr>
<td>Distributed representations and processing</td>
<td>Connectionist models, psychological experiments</td>
</tr>
<tr>
<td>Symbol processing</td>
<td>Symbolic AI models, psychological experiments</td>
</tr>
<tr>
<td>Societies</td>
<td>Sociology, distributed AI, social epistemology, social Ψ</td>
</tr>
</tbody>
</table>
Budapest Semester in Cognitive Science (BSCS)

http://www.bscs-us.org/
Cybernetics is back!

+ Constructivism: is a reaction against the view that knowledge and perception are the results of sensation and maintains that (i) the nervous system, in order to be adaptive, must process available INFORMATION actively and CONSTRUCT an internal world (ii) these process be describable in terms of algorithms

CONSTRUCTIVISM: human learning is constructed, that learners build new knowledge upon the foundation of previous learning. This view of learning sharply constrasts with one in which learning is the passive transformer of information from one individual to another, a view in which reception, not construction is emphasized.

Neural Constructivism: the representational features of cortex are built from the dynamic interaction between neural growth mechanisms and environmentally derived neural activity.
Cybernetics is back!

Computer-brain disanalogy (Michael Conrad)

<table>
<thead>
<tr>
<th>Digital computers</th>
<th>Brains</th>
</tr>
</thead>
<tbody>
<tr>
<td>programmed from outside</td>
<td>self-organizing devices</td>
</tr>
<tr>
<td>structurally programmable</td>
<td>structurally non-programmable</td>
</tr>
<tr>
<td>low adaptability</td>
<td>high adaptability</td>
</tr>
</tbody>
</table>

Figure 2: discrete dynamics        discrete and continuous dynamics

physical implementation

irrelevant in principle          depends on biological substrate

information processing

at network level                network and intraneuronal levels
Figure 3: Otto Rössler: Universe can be described in two different ways, from the outside ("exophysics") and from within ("endophysics")

Biological systems contain their own descriptions, and therefore they need special methods.

[Hermeneutics: emphasize the necessity of self-reflexive interpretation]

Dennett: not only texts, [people and artifacts], but also biological organisms can be interpreted.
Cybernetics is back!

Objectivity, reality, hermeneutics and interpretation

Figure 4: Robert Rosen recalls the physicist’s approach: “Mind cannot be the object (or subject) of legitimate scientific study, since it cannot be identified with objective reality.”
Cybernetics is back!

Objectivity, reality, hermeneutics and interpretation

Rosen’s analysis:
- this kind of objectivity is narrowly understood and based on mechanistic notions
- biologists adopt a more narrow concept of objectivity: it should be independent not only from perceptive agents, but also from the environment. - ’to explain wholes from parts, that is ”objective”, but parts in terms of wholes, that is not’. - closed causal loops are forbidden in the ”objective” world
- the world of systems determined by linear (and only linear) causal relationships belongs to the class of ”simple systems” or mechanisms.
- the alternative is not a ”subjective” world, immune to science, but a world of complex systems, i.e., one which contains closed causal loops

Circular and network causality
Systems with feedback connections and the systems of these connected loops can be understood based on the concepts of circular and network causality

WHY ARE ORGANISMS DIFFERENT FROM MACHINES?