Scienze Cognitive Computazionali e Intelligenza Artificiale

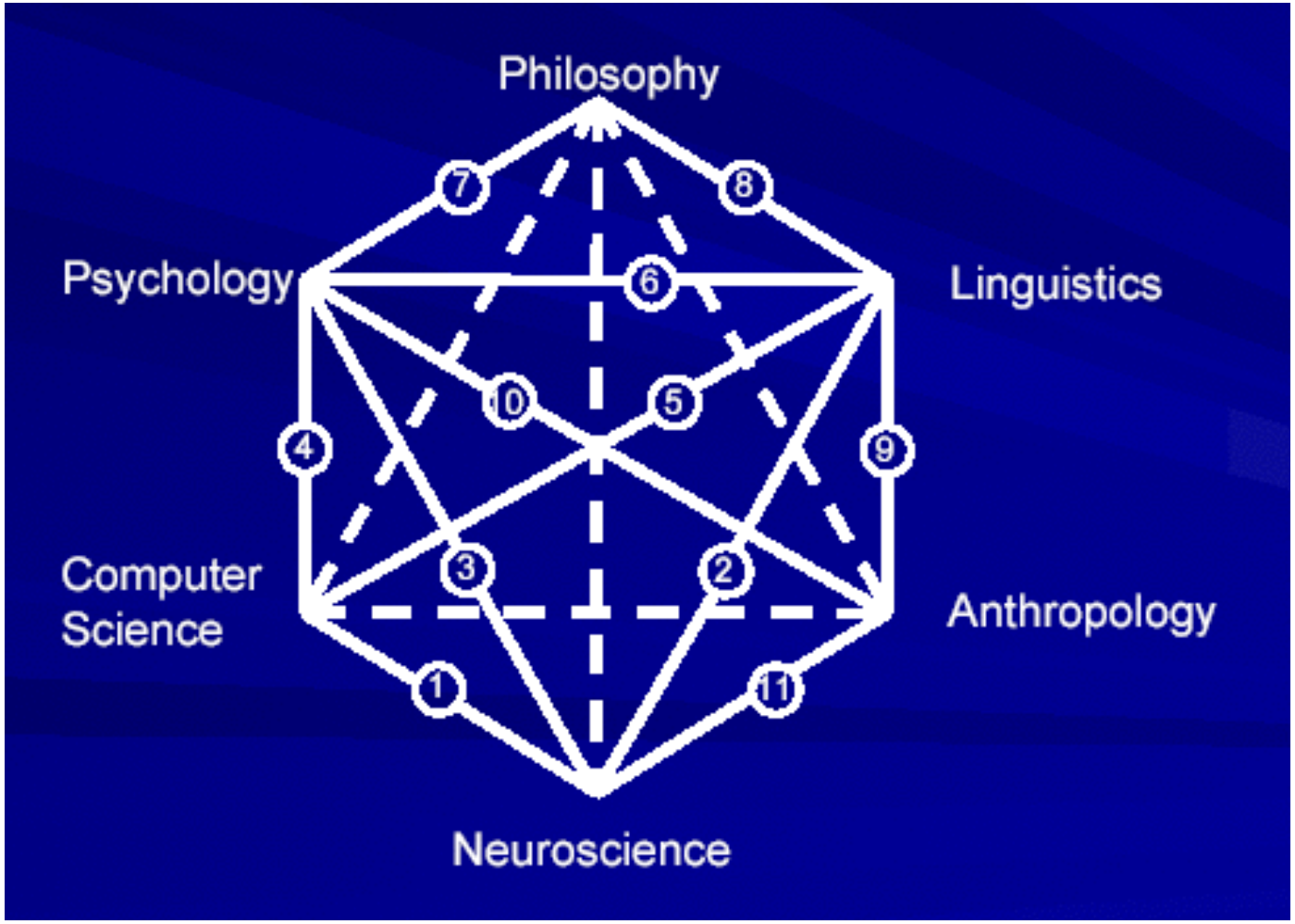
Antonio Lieto

Università di Torino, Dipartimento di Informatica, Italy
ICAR - CNR, Palermo, Italy

http://www.di.unito.it/~lieto/

Le scienze cognitive

- (Almeno) sei discipline diverse:
  - **Intelligenza artificiale**
  - Filosofia
  - Neuroscienza
  - Linguistica
  - Psicologia
  - Antropologia


Philosophy
Psychology
Linguistics
Computer Science
Anthropology
Neuroscience
1
2
3
4
5
6
7
8
9
10
11

Cosa hanno in comune?

- Interesse comune per la mente, la percezione, l'intelligenza, la capacità di risolvere problemi, di apprendere, comunicare, di ricordare degli esseri umani
- La scienza cognitiva nasce quando le diverse discipline capiscono di avere degli interessi in comune
- Ma questo non basta…

2 Ways to replicate cognitive behaviours

1) Invent a novel **computational** solution

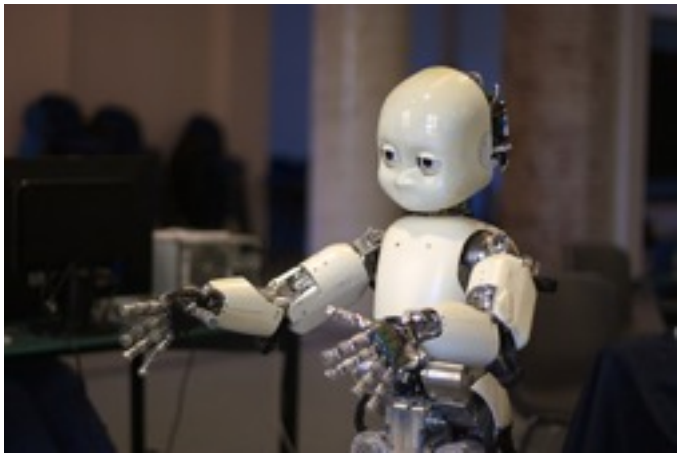
2) Taking inspiration from the nature (**cognitive or bio-inspired systems**)

The first modelling element of the previous list deals with the problem of defining the balance of the computational/cognitive bio-inspired spectrum (**1st aspect**).

Problem if we go for the second option: we need a **MODEL** of at the cognitive/biological level of the system taken as source of inspiration. Such models are very **COMPLEX**.

From Human to Artificial Cognition (and back)

Inspiration



Historical Digression

Cognitive Science and Artificial Intelligence (AI) are, nowadays, scientific fields each endowed with a specific autonomy and research agenda.

E.g. **Cognitive Science** (understanding the mind); **AI** (main goal: build intelligent systems).

Despite the current different focuses and objectives of their research, however, these two disciplines share the idea of studying the "mind", its emergent properties and its functioning, in natural and artificial systems respectively.

In the early days of A**rtificial Intelligence** (whose origin dates back to the "**Dartmouth Workshop**" in 1956), the research on intelligent machines was **strongly and explicitly inspired** by that one coming from the experimental research in Psychology. Some examples in the next slides.

“Natural/Cognitive” Inspiration and AI

Early AI

Cognitive Inspiration
for the Design of “Intelligent Systems”

Modern AI

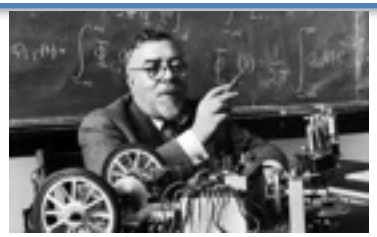
“Intelligence” in terms of optimality of a performance (narrow tasks)

A. Newell

M. Minsky

mid‘80s

e.g. IBM Watson…



N. Wiener

H. Simon

R. Shank

e.g. IBM Deep Blue…

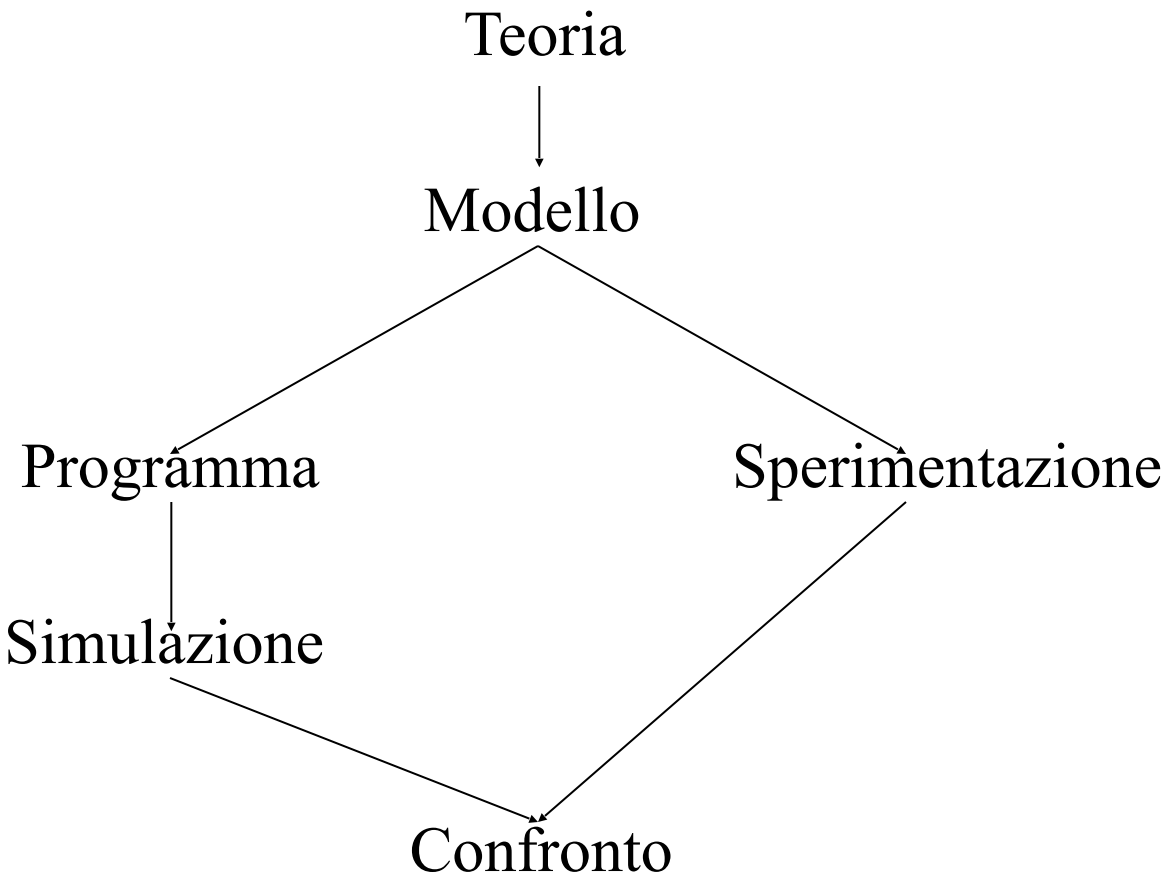
L’intelligenza artificiale

- Dartmouth College Seminar 1956 con J. McCarthy, M. Minsky, A. Newell, H. Simon
- “ogni aspetto dell’apprendimento e ogni altra caratteristica dell’intelligenza puo’ essere descritto in modo tanto preciso da renderlo simulabile da una macchina”

- “My hypothesis then is that thought models, or parallels, reality -- that its essential feature is not ‘the mind’, ‘the self’, ‘sense data’ nor ‘propositions’, but is symbolism, and that this symbolism is largely of the same kind which is familiar to us in mechanical devices which aid thought and calculation” (Craik, 1943, p. 57).

Kenneth Craik (1914-1945)

Validazione di una teoria



Teorie, modelli e programmi

- teoria cognitiva

– postula in linguaggio informale un insieme di strutture di rappresentazione e un insieme di procedure

- modello

– rende più precisi strutture e procedure in analogia con specifiche di strutture dati e algoritmi

- programma

– implementazione del modello attraverso un linguaggio di programmazione

- La piattaforma può contribuire alla specificazione del modello

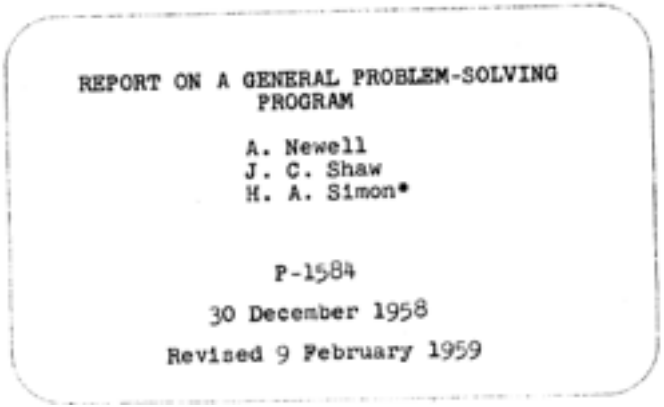
GPS (General Problem Solver)

A system able to **demonstrate simple logic theorems** whose decision strategies were explicitly inspired by human verbal protocols (**Simon, Shaw, Newell, 1959)**.

Nobel Prize
"**bounded rationality**"

Idea: the computer system had to approximate the decision operations described by the humans in their **verbal descriptions** as closely as possible.

In particular, the **GPS** system **was able to implement a key mechanism in human problem solving: the so called means-ends analysis** (or M-E heuristic). In M-E analysis the problem solver compares the current situation with the goal situation; computes the difference between the two states; finds in memory an operator that experience has taught reduces differences of this kind; and applies the operator to change the current situation until the goal state is reached..


REPORT ON A GENERAL PROBLEM-SOLVING PROGRAM

A. Newell
J. C. Shaw
H. A. Simon*

P-1584

30 December 1958

Revised 9 February 1959

Semantic Networks

Ross Quillian (1968) *developed a* a psychologically plausible model of human semantic memory implemented in a computer system. The idea of Quillian was that human memory was associative in nature and that concepts were represented as sort of nodes in graphs and activated through a mechanism of "spreading activation" allowing to propagate information through the network to determine relationships between objects.

Early AI Systems – Ex.Frames

Frame 1	
Concept 1	
Attribute 1	Value 1
Attribute 2	Value 2
Attribute 3	Value 3
…	…

(Minsky M., 1975)

This framework was used for representing, in artificial systems, common-sense knowledge (including *default and typical* knowledge) about the external world (e.g *Birds usually fly*). This type of knowledge organization proposed in the Frames enabled the first AI systems to extend their automated reasoning abilities from the classical deduction to more complicated forms of common-sense and defeasible reasoning.

Scripts (Shank and Abelson, (1977)

R. Shank

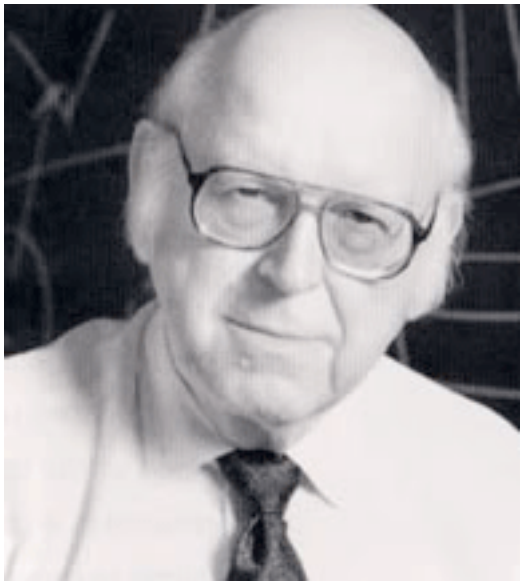
Data structure for r**epresenting knowledge of common sequences of events** (e.g. the sequence of events used to go out for a dinner) and used in *natural language processing systems* as way to enable intelligent question-answering about simple stories.

Script Restaurant: A script representing the restaurant situation is a data structure that would record the typical events associated to this scenario: e.g. entering in the restaurant, asking for a table, sitting down, consulting a menu, eating the food, pay the check etc.

Scripts enabled early AI system to answer questions to simple stories.

Example: "**Mary went to a restaurant, ordered a salmon. When she was paying she noticed that she was late for the next appointment**", Question: "**Did Mary eat dinner last night?**"). This information is not explicitly provided in the story. Answering to these types of questions was possible trough the use of a script of the restaurant situation.

Cognitive Architectures



A **cognitive architecture** (Newell, 1990) implements the **invariant structure of the cognitive system**.

The work on such systems started in the ‘80s (**SOAR** (Newell, Laird and Rosenbloom, 1982)

It captures the underlying **commonality** between different **intelligent agents** and provides a **framework** from which intelligent behavior arises.

The **architectural approach** emphasizes the **role of memory** in the **cognitive process**.

Allen Newell (1990)
Unified Theory of Cognition

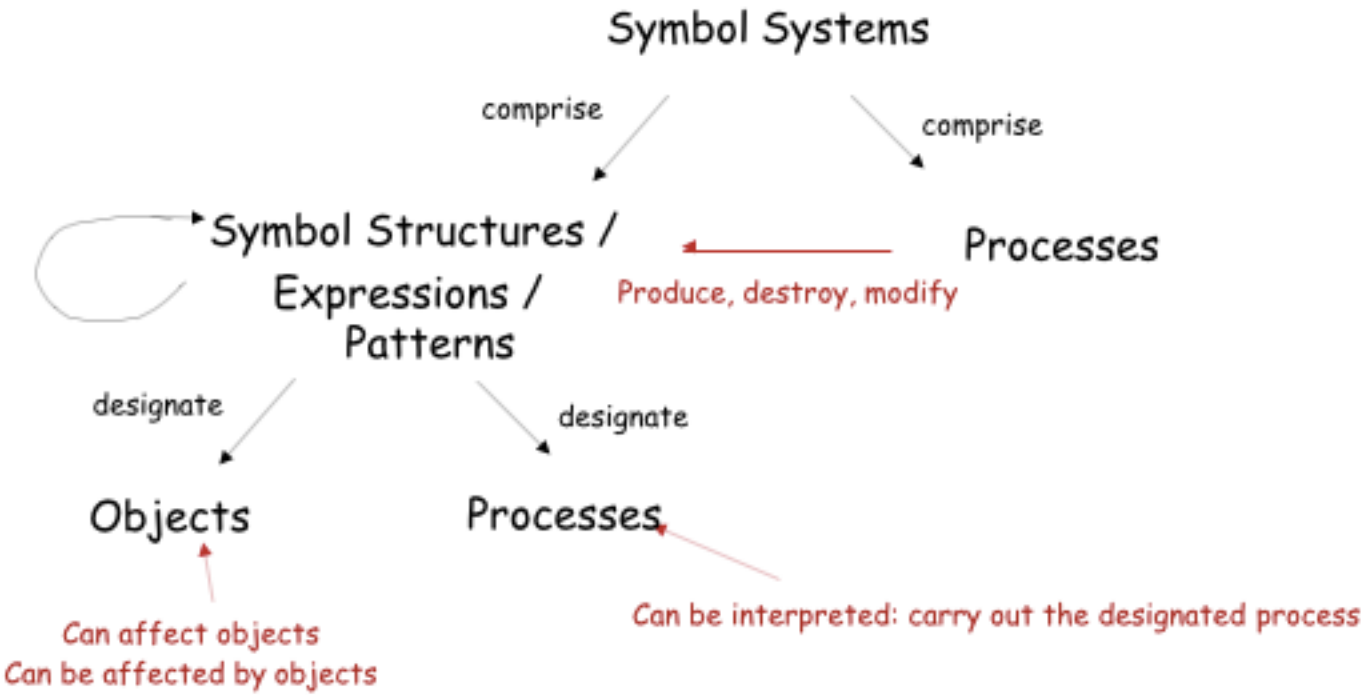
Cognitivism & Artificial Intelligence

- Physical Symbol Systems
  - Symbols are abstract entities that can be instantiated as tokens
  - A physical symbol system has [Newell 90]:
    - Memory (to contain the symbolic information)
    - Symbols (to provide a pattern to match or index other symbols)
    - Operations (to manipulate symbols)
    - Interpretations (to allow symbols to specify operations)
    - Capacities for
      - Composability
      - Interpretability
      - Sufficient memory
  - Symbol systems can be instantiated but … behaviour is independent of the particular form of the instantiation

Cognitivism & Artificial Intelligence

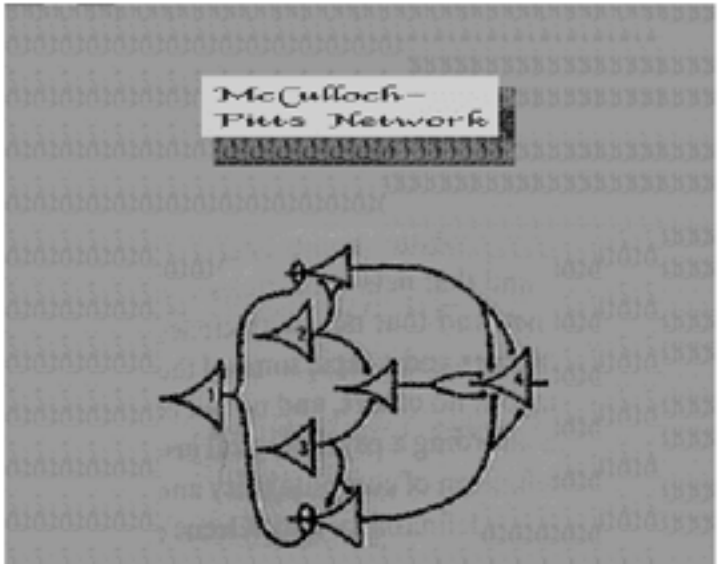
- Physical Symbol Systems [Newell and Simon 1975]
  - The Physical Symbol System Hypothesis
    - A physical symbol system has the necessary and sufficient means for general intelligent action
    - Any system that exhibits general intelligence is a physical symbol system
    - A physical symbol system is ‘a machine that produces through time an evolving collection of symbol structures’

Cognitivism & Artificial Intelligence



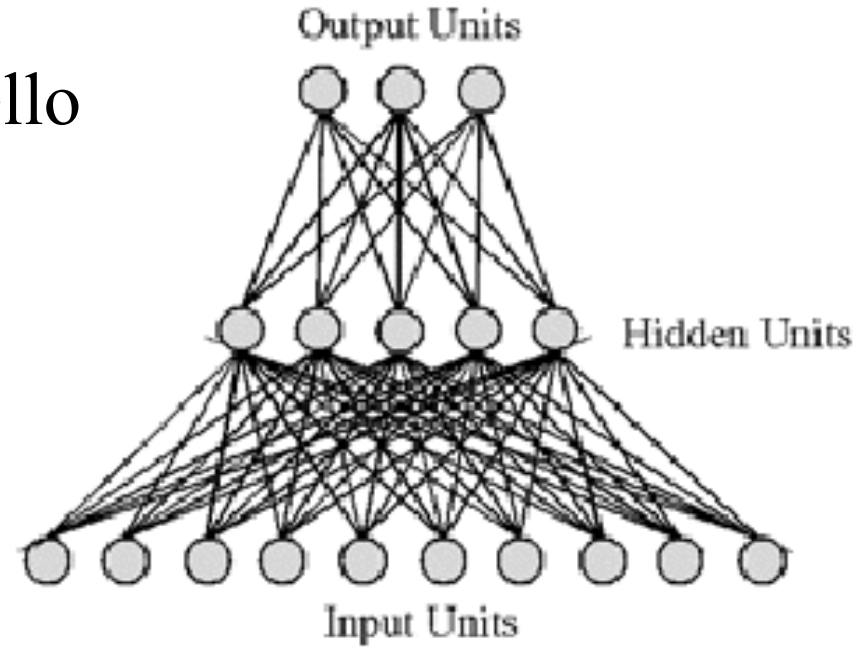
Reti neurali

- McCulloch e Pitts, *A logical calculus of ideas immanent of nervous activity,* 1943
- Qualunque funzione computabile può venir realizzata da una rete opportuna di neuroni ideali



“reti neurali”

- D. Rumelhart e J. McLelland *Parallel distributed processing* 1986
- Al posto di una architettura Von Neumann, rete di unità semplici che inviano segnali eccitatori o inibitori a quelle vicine.
- Simulo neuroni del cervello
- Computazione parallela
- Apprendimento per back propagation



Cybernetics tradition of the AI

This approach to the study of the artificial did not came out *ex-abrupto*. It borrowed its original inspiration – from a historical perspective– from the methodological apparatus developed by the scholars in Cybernetics (Cordeschi, 1991).

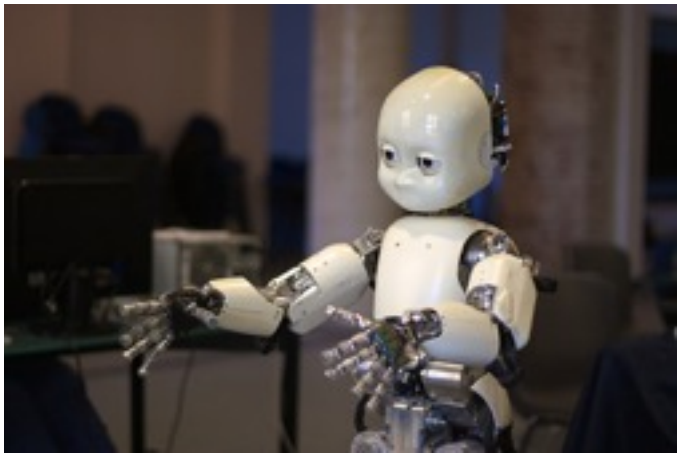
The origins of cybernetics are usually traced back to the middle of the 1940s, with the release of the 1948 book of Norbert Wiener entitled "**Cybernetics: Or Control and Communication in the Animal and the Machine**".

One of underlying idea of cybernetics was that one of building mechanical models to simulate the adaptive behavior of natural systems.

As indicated in Cordeschi (Cordeschi, 2002): **"the fundamental insight of cybernetics was in the the proposal of a unified study of organisms and machines**".

From Human to Artificial Cognition (and back)

Inspiration



Explanation

Lieto and Radicioni, Cognitive Systems Research, 2016

Explanatory power

When a biologically inspired **computational system/ architecture** has an **explanatory power** w.r.t. the **natural system** taken as source of inspiration ?

Which are the **requirements** to consider in order to **design** a **computational model of cognition with** an **explanatory power** ?

- **Cognitive AI Paradigms:** some methodological and technical considerations.

- **Functionalist vs Structuralist Design** Approach.

- Preliminary introduction to the **Cognitivist** and the **Connectionist** Paradigms (we will see more in details such paradigms in the following lectures) and the **design implications** for the level of abstraction considered

Cognitive AI

Attention to the **heuristics-based** solutions adopted by humans (e.g. Gigerenzer & Todd, 1999) for combinatorial problems ("**bounded rationality heuristics**").

Heuristics realize/implement some **cognitive functions** and are responsible of the macroscopic external behaviour of an agent.

Heuristics

The terms comes from “Eureka” that means “help/shortcut for the discovery of a solution” (of a given problem)

This term is used with different meanings from the beginning of the AI research:

1) detailed simulation of **human cognitive processes** (**Information Processing Psychology**, Newell and Simon, 1972) by using a computer program

2) the most efficient (and efficacious) performance possible from computer programs, by allowing also for typically non-human procedures (e.g. **brute-force approaches**).

In the so called Cognitive AI we refer to the first meaning of the expression.

A focus shift in AI

Vision the **early days of AI**: “Understanding and reproducing, in computational systems, the full range of intelligent behavior observed in humans” (**P. Langley, 2012**).

This view was abandoned. Why?

- Emphasis on **quantitative results** and metrics of performance: (“**machine intelligence”**: achieving results and optimize them !)

- Renewed attention since “The gap between natural and artificial systems is still enormous” (**A. Sloman, AIC 2014**).

2 Main Perspectives

"**Cognitive Systems**" (Brachman and Lemnios, 2002): "designs, constructs, and studies computational artifacts that exhibit the full range of human intelligence". [**Cognitivist approach**, Vernon 2014].

"**Nouvelle AI**" (e.g. *Parallel Distributed Processing* (Rumhelarth and McLelland, 1986) based on **bio-plausibility** modelling techniques allowing the functional reproduction of heuristics in artificial systems (neglecting the physical and chemical details). [**Emergent approach**, Vernon 2014].

Cognitivism	Nouvelle AI
Focus on **high level** cognitive **functions**	Main focus only on **perception**
Assuming **structured representations** (*physical symbol system*, Simon and Newell, 1976)	Assuming **unstructured representation** (e.g. such as neural networks etc.) and also integration with symbolic approaches.
Architectural Perspective (integration and interaction of all cognitive functions	**System perspective** (not necessary to consider a whole architectural perspective).
Inspiration from human cognition (**heuristic-driven approach**)	Bio-inspired computing, **bottom-up approach** (for learning etc.).

A Matter of Levels

- Both the “classical” and “novuelle” approach can realize, ***in principio***, “**cognitive artificial systems**” or “**artificial models of cognition**” provided that their models operate at the “right” level of description.

- A debated problem in AI and Cognitive Science regards the **legitimate level of descriptions** of such models (and therefore their **explanatory power**).

Functionalist vs Structuralist Models

Functionalism

- Functionalism (introduced by **H. Putnam**) postulates a **weak equivalence** between cognitive processes and AI procedures.
- AI procedures have the functional role (“**function as**”) human cognitive procedures.
- **Multiple realizability** (cognitive functions can be implemented in different ways).
- **Equivalence** on the **functional macroscopic properties** of a given intelligent behaviour (based on the same input-output specification).
- This should produce **predictive models** (given an input and a set of procedures functionally equivalent to what is performed by cognitive processes then **one can predict a given output**).

Problems with Functionalism

- If the equivalence is so ***weak*** it is not possible to interpret the results of a system (e.g. interpretation of the **system failure**…).

- A pure **functionalist model** (posed **without structural constraints**) is a **black box where** a predictive model with the same output of a cognitive process can be obtained with no explanatory power.

Birds and Jets

- Both a **Bird** and a **Jet** can fly but a jet is not **a good explanatory model** of a bird since its flights mechanisms are different from the mechanism of bird.

- Purely **functional models/systems** are **not** "computational models of cognition" (they have no explanatory power w.r.t. the natural system taken as source of inspiration).

Modern “Functional” Systems in AI

They are very good artificial systems but they have no **explanatory power** with respect to how humans solve/face the same problems.

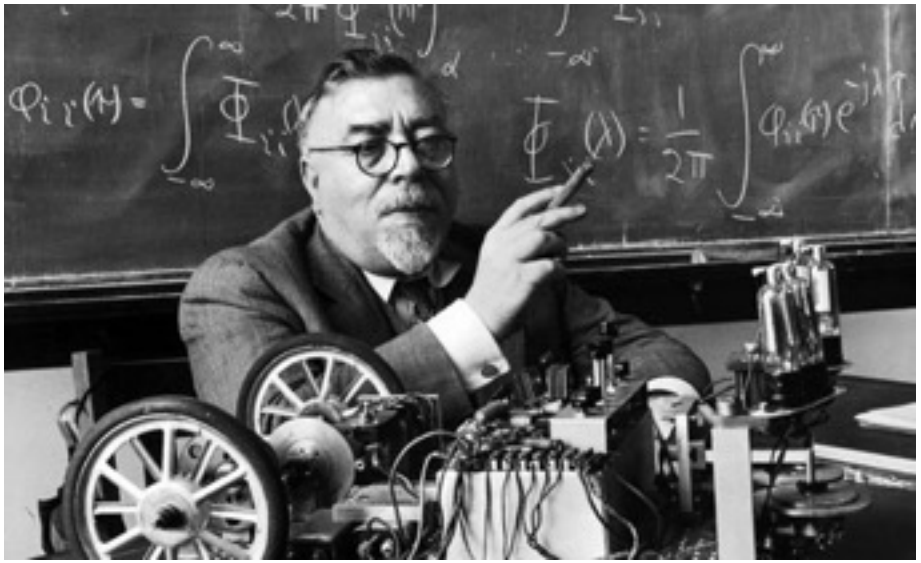
In this sense they are not **cognitive ! (e.g. despite IBM claims)**

Structuralism

- **Strong equivalence** between cognitive processes and AI procedures (Milkowski, 2013).

- Focus not only on the **functional organization** of the processes but also on the **human-likeliness** of a model (**bio-psychologically plausibility**).

Wiener's "Paradox"



"*The best material model of a cat is another or possibly the same cat*"

- Difficulty of realizing models of a given natural system.
- Need of ***proxy-models*** *(i.e. good approximations)*

A Design Problem

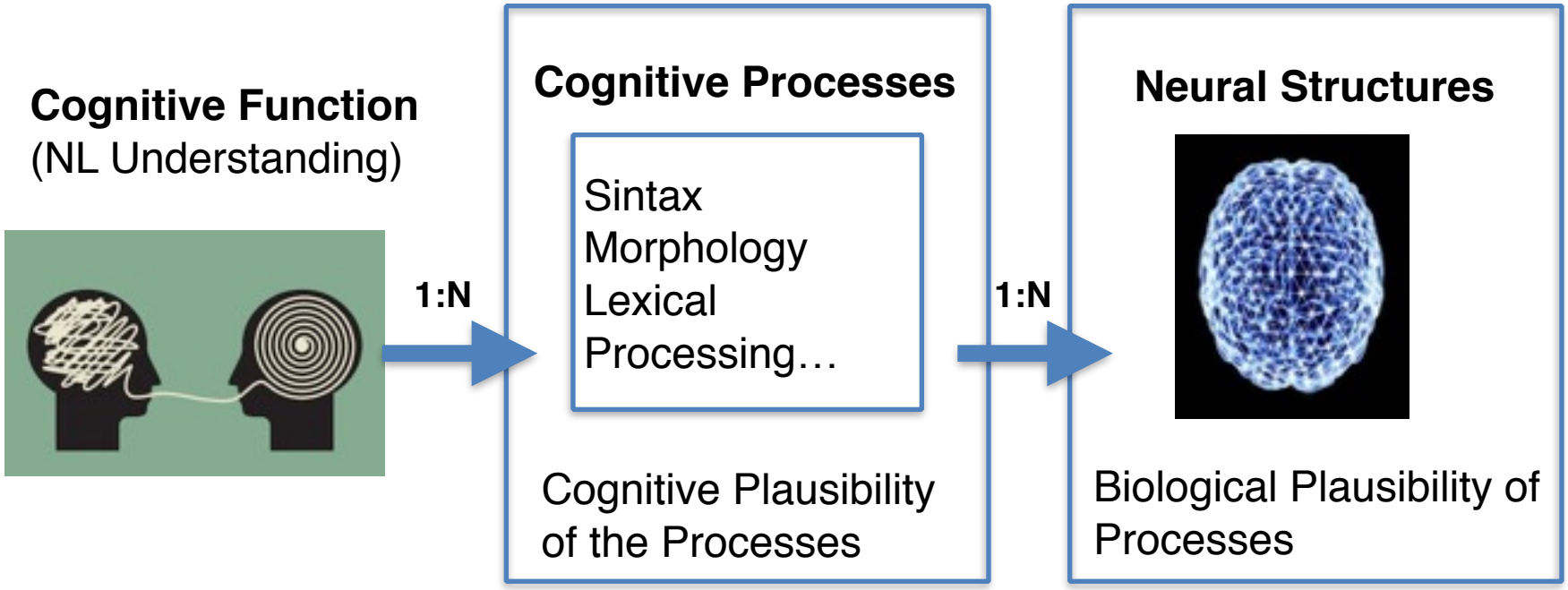
Z.Pylyshyn (’79): “*if we do not formulate any restriction about a model we obtain the functionalism of a Turing machine. If we apply all the possible restrictions we reproduce a whole human being*”

- Need for looking at a **descriptive level** on which to enforce the constraints in order to carry out a **human-like** computation.

- A design perspective: between the **explanatory level of functionalism** (based on the macroscopic stimulus-response relationship) and the **mycroscopic one of fully structured models** (reductionist materialism) we have, in the middle, a lot of **possible structural models.**

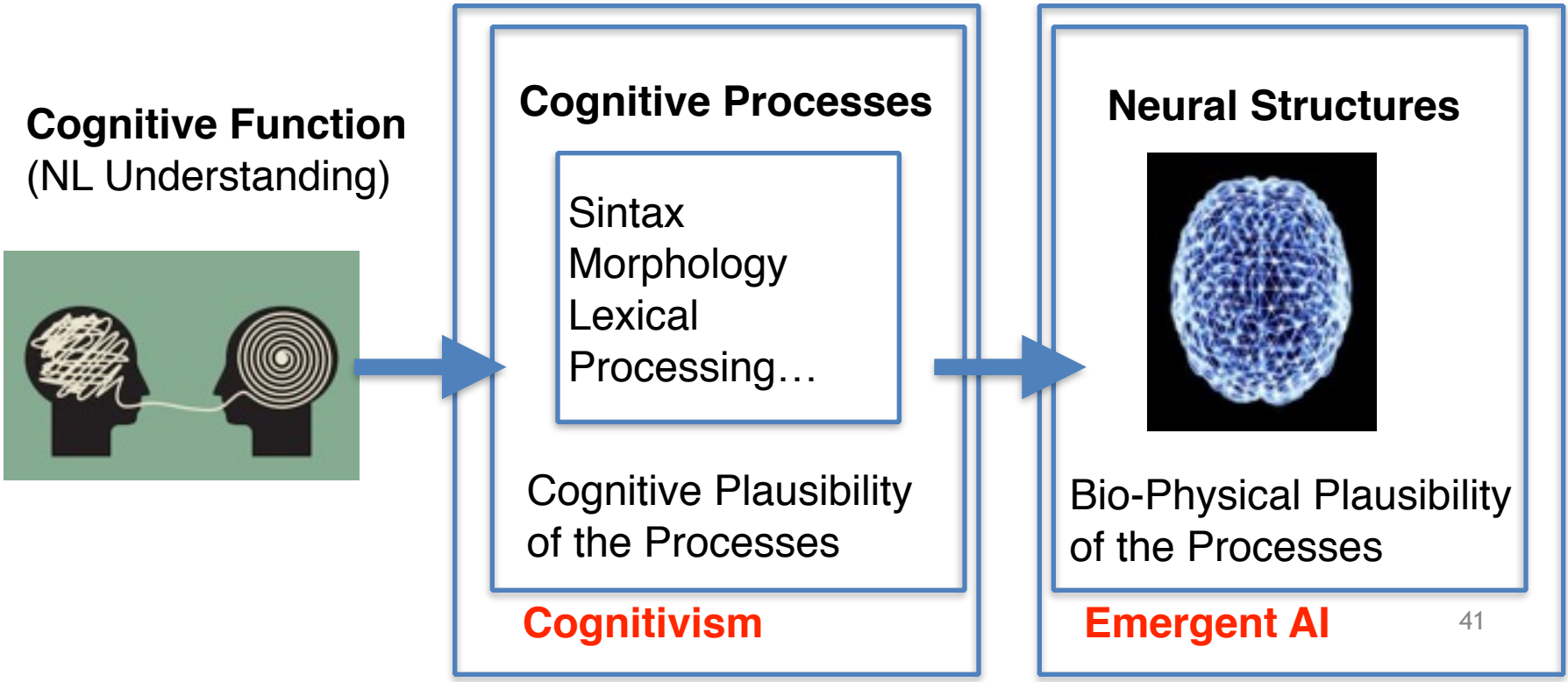
Many Structural Models

Both the presented AI approaches may build **structural models of cognition** at different levels of details (having an **empirical adequacy**).



Many Structural Models

Both the presented AI approaches may build **structural models of cognition** at different levels of details (having an **empirical adequacy**).



Take home message (part 1)

- **Cognitive Artificial Models** have an explanatory power **only if** they are **structurally valid models** (realizable in different ways and empirically adequate).

- **Cognitive Artificial Systems** built with this design perspective have an **explanatory role** for the **theory** they implement and the "**computational experiment**" can provide results useful for refining of rethinking theoretical aspects of the **natural inspiring system**.

Tesi

- Gli argomenti trattati a lezione sono possibile oggetto di tesi
- Sono postate proposte per progetti di tesi (su questi argomenti) a questo link: http://www.antoniolieto.net/InfPsych.html