Formal Systems of Neurons in Artificial Intelligence

Molly Graham

Dec 13, 2023

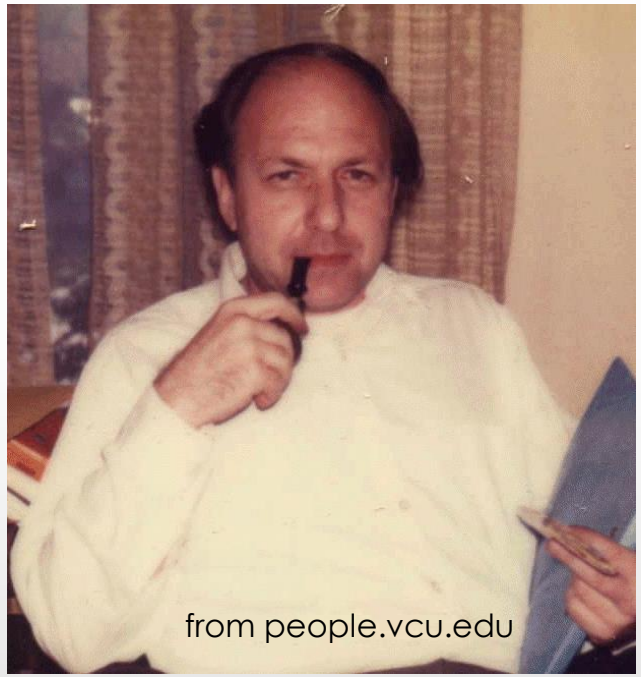
CONTEXT

- Chapter 1 – Social Robots for the Future
- Chapter 2 – Analyzing Empathy
- Chapter 3 – The Limitations of Developmental Robotics
  1. History of Artificial Intelligence
  2. Introducing iCub
  3. **iCub's Failure to Empathize**
     - aka ***Formal Systems of Neurons in AI***
- Chapter 4 – Modelling Life for Behaviour

Robert Rosen

(1934 – 1998)

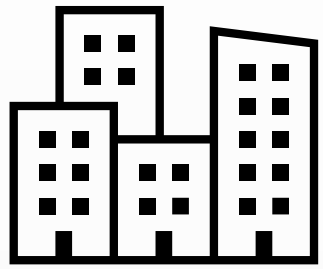
- Theoretical biologist
  - Relational biology: functional organization
- Prof. of Biophysics at Dalhousie in NS (1975)
- Publications:
  - *Anticipatory Systems* (1985)
  - *Life Itself* (1991)
  - *Essays on Life Itself* (2001)

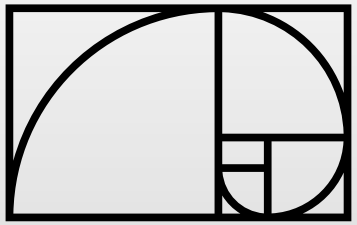


from people.vcu.edu

Natural Systems

- Physical/natural world
- Human constructs
  - cars, factories, cities, etc.





Formal Systems

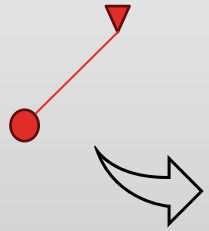
- Mathematical objects
- Other formal systems
  - Propositions (language)
  - Input-output relations

$$(\exists x)(Fx \rightarrow (\sim Gy \vee z)) \wedge (\exists x)(Gy \rightarrow \sim Hxy))$$

Modelling Relations

- Connecting properties of NS to entities in FS
- Encode from NS to FS
  - Modelled in a FS
- Decode from FS to NS
  - Prediction in the NS

$$\frac{d^2\theta}{dt^2} + \frac{g}{\ell}\sin\theta = 0$$



Computer Code as a Formal System

- Steps to be taken to accomplish a task (algorithm)
  - Pure *syntax*, no *semantics*
  - Software is independent of hardware
- Structures of inference or entailment
  - If x, then *y*
  - $p \rightarrow q$

McCulloch-Pitts neuron:

- 0 = inactive neuron
- +1 = active neuron
  - e = excitatory input
  - i = inhibitory input
  - θ = firing threshold
  - s(t) = state at time t
  - (t – 1) = previous state

McCulloch-Pitts neuron:

$$s(t) = +1 \quad \text{iff} \left[\sum_{k=1}^{m} e_k(t-1) + \sum_{k=1}^{n} i_k(t-1) \right] \geq \theta;$$

$$= 0 \quad \text{otherwise.}$$

(AS 187)

Computer Code as Artificial Intelligence

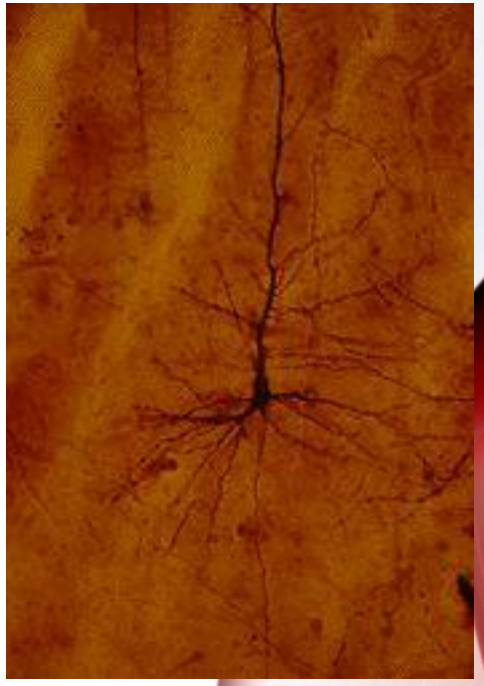
- Early computers inspire ideas about AI
  - Turing machine in 1936
  - von Neumann architecture in 1945
  - '*artificial intelligence*' in 1956

- Programs created for tasks:
  - Symbolic (if-then)
  - Connectionist (neural networks)

- How to add semantics?
  - Developmental robotics for symbol grounding
  - iCub *seems* capable of learning…

Istituto Italiano di Tecnologia

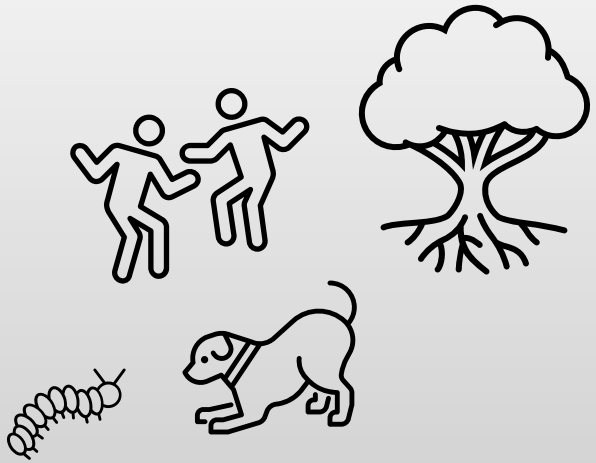
Santiago Ramón y Cajal
nobelprize.org/prizes/medicine/1906/cajal/article/



Simulated Behaviour

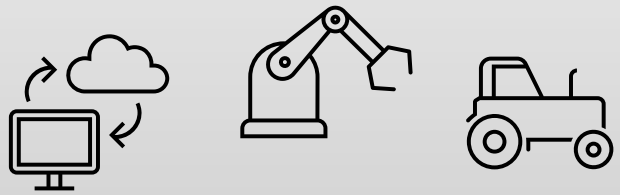
- Behavioural limitations will inevitably rise from the physical distinctions between NS and FS
  - Anticipatory systems: organisms like humans
  - Passive artifacts: computers and machines

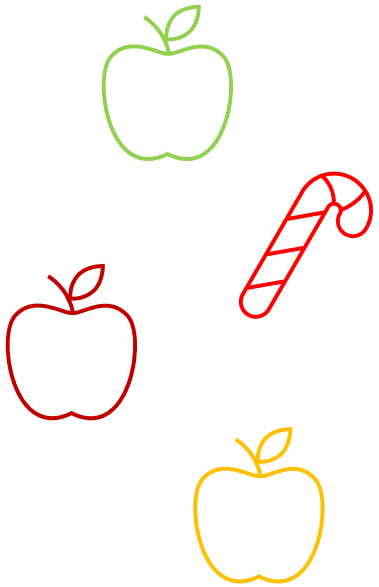
- Semantics *from* biological structures
  - Can the functionality of semantics be separated?
  - No: cannot fully encapsulate semantic values from syntax alone



$$s(t) = +1 \quad \text{iff} \left[\sum_{k=1}^{m} e_k(t-1) + \sum_{k=1}^{n} i_k(t-1) \right] \geq \theta;$$

$$= 0 \quad \text{otherwise.}$$

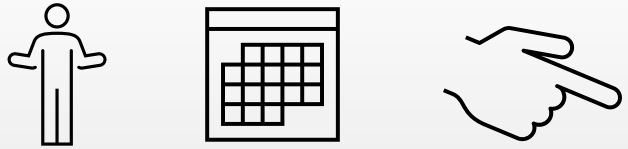




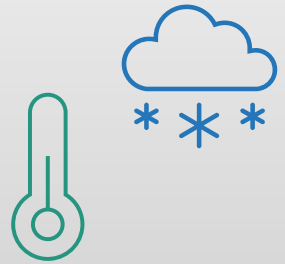
Simulated Behaviour – Continued

- Rosen appeals to Gödel's *Incompleteness Theorem:*
  - There exist true propositions which cannot be proven by axioms
  - No set of truths that we can derive *all* math from, there is a hole
  - **Self-referential** statements are a problem for syntax

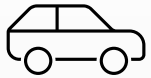
- In computers, no semantics required but self-reference relies on semantic values
  - 'me', 'today', 'that', etc.

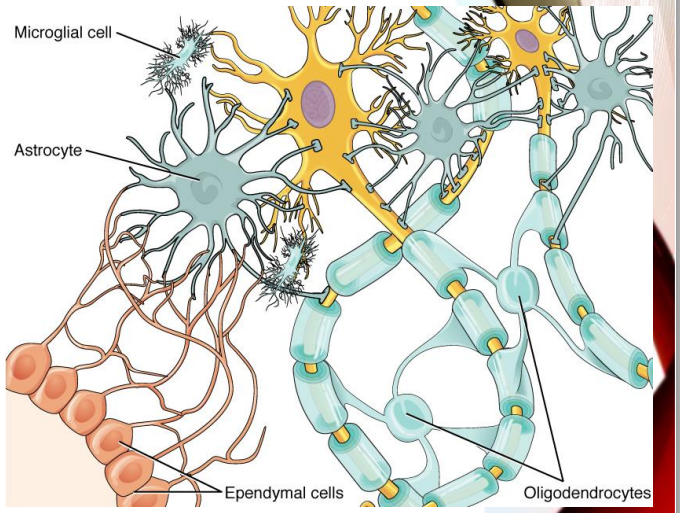


- Semantics cannot be fully described or accounted for by purely syntactical structures
  - Infinite regress: "what do you mean by…?"
    - Numerical values must be given
    - Which values? For a purpose…



Social [Computerized] Robots

- Communication involves more than just words
  - E.g. body language, tone, context, etc.
- May appear intelligent but it's a simulation of understanding
  - Associations between stimuli are created
    - "Car" → 
  - But the word is *meaningless* to the computer/robot
    - Only "knows" or "learns" associations, not semantic information
      - New car smell
      - Feeling of the open road
      - Acceleration, gear shifting, winter driving
  - Associations ≠ meaning
    - "Why are icy roads bad?"

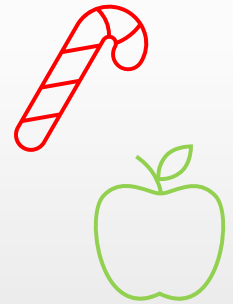


CONCLUSION

- Formal systems *model* natural systems
  - Reductive, an idealized model
  - NS incompletely captured by FS

$$s(t) = +1 \quad \text{iff} \left[\sum_{k=1}^{m} e_k(t-1) + \sum_{k=1}^{n} i_k(t-1) \right] \geq \theta;$$
$$= 0 \quad \text{otherwise.}$$

- Computerized robots like iCub only use syntax
  - No semantics, unlike biological organisms
  - Can simulate semantics with syntax *to a degree*
  - Socialization and empathy requires semantics



- A physical distinction between types of behaviour
  - Human, animal, plant (natural biological systems)
  - iCub (embodied formal system)
- Why does this matter? Anthropomorphization…

THANK YOU!

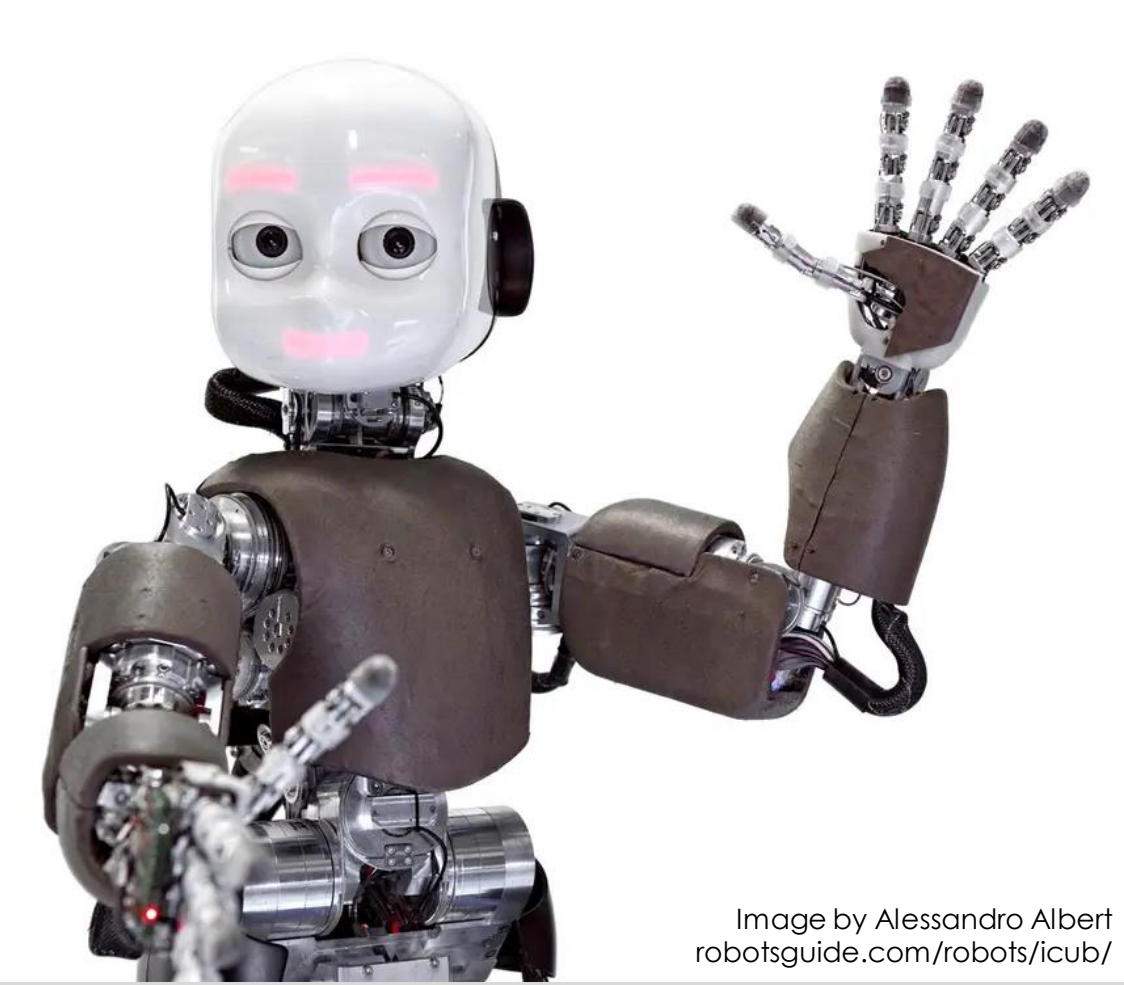


Image by Alessandro Albert
robotsguide.com/robots/icub/