Dupuy, *The Mechanisation of the Mind,* Chapter 2, "A Poorly Loved Parent"

Two Papers published in 1943 - "Behavior, Purpose and Teleology," Arturo Rosenblueth, Norbert Wiener, Julian Bigelow & "A logical calculus of the ideas immanent in nervous activity," Warren McCulloch, Walter Pitts

Macy Conferences - 1946-1953 interdisciplinary conferences – Ashby, McCulloch, Pitts, von Neumann, Wiener, Bigelow, von Foerster, Shannon

Key Ideas in Cybernetics

teleology (final cause), goal-directedness, purposiveness

feedback loops
results of system output effect subsequent input

"modeled after a Turing machine, which is to say a machine provided with an internal state that it is capable of modifying as a function of input and the internal state of the preceding period."

error-driven learning

"learning to adjust its behavior in the light of the errors that it commits"

QUOTATIONS

"Man-made machines are not brains, but brains are a very ill-understood variety of computing machines. Cybernetics has helped pull down the wall between the great world of physics and the ghetto of the mind." — McCulloch

"McCulloch's distancing of himself from von Neumann was motivated by a strictly philosophical disagreement about the status of models. A model, for McCulloch, was not a simple instrument of calculation having a purely pragmatic value, determined by the answer to the question "Does it work?" It had an ontological reality."

"It is…absurd to say or to write, as is too often done, that the model of idealized neurons is only the product of a "formal" or "fictitious" neurology. One might as well say, as McCulloch remarked, that Maxwell's theory of electromagnetism (or any other formalized physical theory) is fictitious because it deals with idealized entities."

"The philosophical ambition of this article is considerable, since it at- tempted nothing less than to give a purely neuroanatomical and neuro- physiological basis for synthetic a priori judgments, and thus to ground a neurology of mind"

1969 - Minsky & Pappert deliver an unwarranted "death blow" to neural network models

More strange offspring of cybernetics:

autopoiesis
cybernetic governance
second-order cybernetics
game theory and decision theory

(Please see the following sheet for a summary of the paper itself)

- First (afaik) model of a neuron proposed, based on some basic observations about neurons, but not much of the biology was known back then
    - Leads to some “educated guesses” like the variance in thresholds from neuron to neuron, and the fact that one inhibitory synapse can stop whole neuron from firing
    - Essentially, very similar to perceptrons later on

- The frame of this paper is very “logic-oriented”, though they are inspired by the basic high-level idea of a neuron, they moreso use this idea to think about what kinds of propositions their neural net construction can handle on a logical basis
    - I think this is why they spend so much time on the recurrence (“nets with cycles”) section since this may set up a situation where “future” events influence “past” events
        - “the possibility that activity may be set up in a circuit and continue reverberating around it for an indefinite period of time, so that the realizable Pr may involve reference to past events of an indefinite degree of remoteness”
        - This section is quite dense so I didn’t try to understand it since it didn’t seem like a good use of time. They seem to resolve this apparent “contradiction” through modular arithmetic though

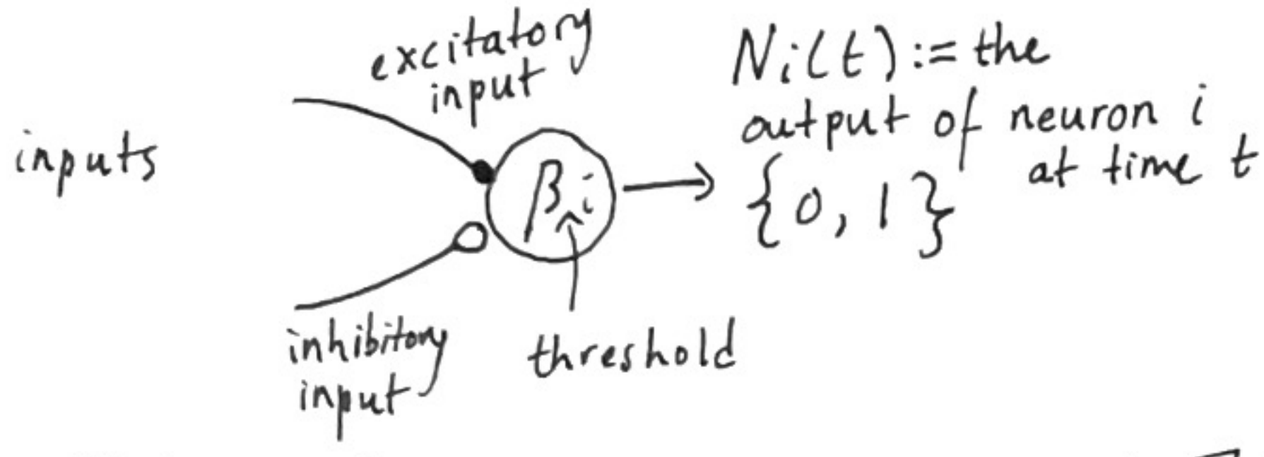
    - They do not handle modifications to the net (assume topology doesn’t change over time). They do acknowledge that this can happen though, but leave it for future work
        - This means that the focus is not ‘learning’ in this paper

- It seems like this was written optimistically at a time when fields were less disjoint, and the authors believed that their work would be helpful for all computational and psychological fields
    - “The psychiatrist may take comfort from the obvious conclusion concerning causality--that, for prognosis, history is never necessary. [...] Instead, diseased mentality can be understood without loss of scope or rigor, in the scientific terms of neurophysiology. “
    - “This is of interest as affording a psychological justification of the Turing definition of computability and its equivalents, Church's [lambda] definability and Kleene's primitive recursiveness: If any number can be computed by an organism, it is computable by these definitions, and conversely. “

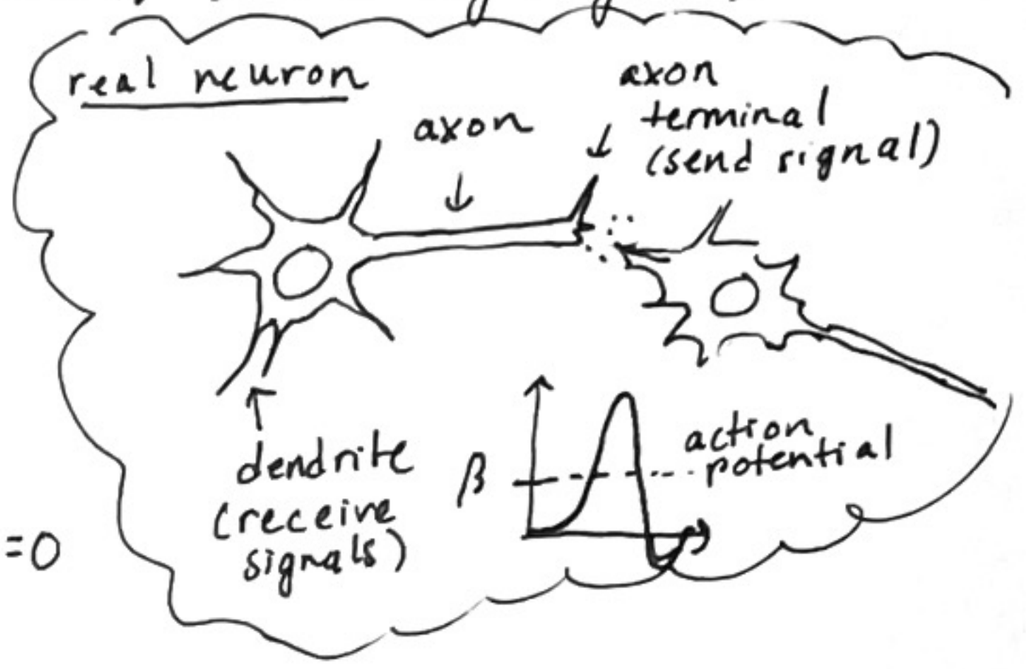
- The story behind these researchers (esp. Pitts) is actually very interesting and tragic as well (search it up if interested)

McCulloch / Pitts neuron model

- written at a time when less was known about anatomy
  - ↳ they take a primarily logical view (neuron is like a logic gate)



$N_i(t) :=$ the output of neuron i at time t $\in \{0, 1\}$



$$N_i(t) = \begin{cases} 1 & \text{if } \sum_{j \in E} N_j(t-1) \geq \beta_i \text{ and } \sum_{k \in I} N_k(t-1) = 0 \\ 0 & \text{otherwise} \end{cases}$$

where E is excitatory synapses and I is inhibitory synapses.

Assumptions about nets (acyclic)

1. Neuron activity is "all-or-nothing"
2. The threshold for each neuron is independent of prev. activity
3. The only delay is synaptic delay (timesteps)
4. Activity of inhibitory synapse absolutely prevents the neuron from firing
5. Structure of the net doesn't change over time

- solution of a net: set of logical sentences $\{S\}_{i=1}^{N}$ for the N neurons s.t. the sentence is true iff the neuron is firing (in terms of similar S_is for previous neurons and constant sentences)
  - ↳ realizability: for some $\{S\}_{i=1}^{N}$, we can construct such a net

- order of a net: largest # p of neurons that can be removed to make the net acyclic ⇒ acyclic is order 0

- temporal prepositional expr. (TPE) - type of recursively defined prop. with free time variable
  - ↳ predicate of one variable is a TPE (sub. from $t-1$?)
  - ↳ logical combinations of two TPEs with the same free variable is a TPE
  - ↳ nothing else is a TPE

Thm Every TPE can be realized by an acyclic net, and every acyclic net can be solved by a TPE

- they have a lot of discussion of cyclic (recurrent) nets but unfortunately I don't understand it...

Communicative behavior: *“modify their patterns of behavior on the basis of past experience so as to achieve specific anti-entropic ends.”*

s upuy writes, cybernetics does not recognize “true communication” as implying “anticipation, intentionality, and symbolization, and requir[ing] the capacities of learning, perceiving, and engaging in social life.” ommunication is simply the behavioral response of an organism to **feedback** *from its environment. This response to feedback separates these communicating organisms from eibniz s automata that are in “pre-established harmony” with the universe.*

iener casts a number of complex phenomena as instances of feedback driven control systems, and argues for the essential role of feedback in the functioning of these systems. These range in scope from small and local to vast and global

- *pplause is a crucial form of feedback for a performer, and completes a two-way channel between the performer and their audience*
- *ontrolling the anama anal requires feedback about how ships passing through respond to actions from the control room, and necessitates a two-way communication channel*
- *Officials leading any organization should have a two-way communication channel with those executing their policies, to ensure they are not misguided about what the facts on the ground are*
    - *True of armies at war, corporations, universities, etc.*

iener even extends this idea of feedback and communication to be a model of how society itself is structured.

iener writes about how social structure constrains channels of communication between people from non-hierarchical cooperation among “ skimos”, caste in India, feudal states, and even white supremacy in the S. e seems to use these to illustrate how the structure of communication determines reflects the structure of society. Some argue for a society with a more rigid structure, and rigid lines of communication, in opposition to a democratic system.

One society with a rigid “fascist” structure is that of ants. ach ant has a designated role that it performs as a product of its conditioning. iener contends that this organization of society is not suitable for humans because the way ant physiology constrains its intelligence, and consequently the social structure, is very different from the constraints posed by human physiology.

iener highlights two differences between human and ant physiology that he contends are key to explaining the difference in intelligence

- *Skeleton umans have internal skeletons that are bound to muscles by cartilage. This allows the skeleton to grow with the body. nts have external skeletons. This limits growth as it makes shedding the hard exoskeleton moulting a necessary step for growth. oulting impedes continuity in organization, and by extension impedes memory and learning.*

---

which he cites as a reason for the S falling short of the ideal of communication

- Respiration: Ants respire purely through diffusion. There is no “pump”. This limits the size to which ants can grow, and thus limits the number of neurons they can have in a brain. Even if the neurons are of the same type, they are far fewer in number, which significantly constrains ant intelligence.

Ants are constrained by their physiology to exhibit limited learning (or adaptation to feedback), and thus limited communication. This rigidity constrains the structure of ant society. If humans structured society as if they were as rigid, it would conflict with their physiology that allows for a far greater degree of learning.

Cybernetics takes the view that the structure of the machine or of the organism is an index of the performance that may be expected from it.

A further physiological quirk of humans is that adult humans resemble young apes more than adult apes – corresponding to a prolonged childhood. Humans spend up to 40% of their life as learners.[2]

Wiener then lays out an account of **learning**.

- *feedback is a method of controlling a system by reinserting into it the results of its past performance*
    - A guided missile estimating its trajectory based on the location of a moving target
- *If, however, the information which proceeds backward from the performance is able to change the general method and pattern of performance, we have a process which may well be called learning.*
    - A **change in taping**
    - A guided missile control system estimating how to estimate a trajectory for different types of targets through trial and error

If the plane were able to take a perfectly arbitrary evasive action, no amount of skill would permit us to fill in the as yet unknown motion of the plane between the time when the gun was fired and the time when the shell should arrive approximately at its goal.

- The inherent regularities in the world make learning possible
- We can identify these regularities from experience, using statistics

The advanced process of learning which we have here discussed is still limited by the mechanical conditions of the system in which it occurs, and clearly does not correspond to the normal process of learning in man. But from this process we can infer quite different ways in which learning of a complex sort can be mechanized.

The possibility of these inferences is suggested by the work of Locke and Pavlov. Locke put forth a model where the mind starts *tabula rasa*, and acquires ideas through experience. Based on how these ideas present themselves, they combine into complex ideas. However, Locke did

[2] It’s an interesting claim since Wiener arrives at this 40% number by counting years spent in the formal education system but what Wiener considers learning seems to be well beyond the scope of formal education.

not propose a mechanism for how atomic experiences combine into complex ideas. Pavlov's experiments with conditioning show how associations between ideas may form.

Wiener uses an electric fence as an example. A cow that receives a shock from the electric fence associates seeing the fence with the pain of the shock. The response to the shock is now transferring to a trigger situation (seeing the fence). In cybernetics, this constitutes a change in taping. Wiener speculates about how such a change in taping might be from an "undirected message" sent due to the original stimulus, that causes a change in synaptic thresholds.

Nevertheless, it is interesting to know that the sort of phenomenon which is recorded subjectively as emotion may not be merely a useless epiphenomenon of nervous action, but may control some essential stage in learning, and in other similar processes. I definitely do not say that it does, but I do say that those psychologists who draw sharp and uncrossable distinctions between man's emotions and those of other living organisms and the responses of the modern type of automatic mechanisms, should be just as careful in their denials as I should be in my assertions.

Wiener puts forth what appears to be a materialist account of intelligence, rejecting a non-materialist account such as Leibniz's. Leibniz rejected the possibility of materialism – "the thesis that perception and consciousness can be given mechanical (i.e. physical) explanations" – by arguing that if a mechanical system capable of perception existed, you would be able to scale it up, walk into it, and not be able to spot the consciousness. In Leibniz's writing,

> *One is obliged to admit that perception and what depends upon it is inexplicable on mechanical principles, that is, by figures and motions. In imagining that there is a machine whose construction would enable it to think, to sense, and to have perception, one could conceive it enlarged while retaining the same proportions, so that one could enter into it, just like into a windmill. Supposing this, one should, when visiting within it, find only parts pushing one another, and never anything by which to explain a perception. Thus it is in the simple substance, and not in the composite or in the machine, that one must look for perception.*[3]

[3] Kulstad, Mark and Laurence Carlin, "Leibniz's Philosophy of Mind", The Stanford Encyclopedia of Philosophy (Winter 2020 Edition), Edward N. Zalta (ed.), https://plato.stanford.edu/archives/win2020/entries/leibniz-mind/