7

The Automation of Pattern Recognition

As the industrial revolution concludes in bigger and better bombs, an intellectual revolution opens with bigger and better robots.

Warren McCulloch, Hixon symposium, 1948[1]

A new, essentially logical, theory is called for in order to understand high-complication automata and, in particular, the central nervous system. It may be, however, that in this process logic will have to undergo a pseudomorphosis to neurology to a much greater extent than the reverse.

John von Neumann, Hixon symposium, 1948[2]

The controversy about Gestalt perception

At the core of the intuition that paved the way for artificial neural networks lies an enduring controversy: whether or not human perception is an act of cognition that can be analytically represented and therefore mechanised. The confrontation about this issue flared up in the 1940s during the Macy conferences between the cyberneticians (who

---

1 Lloyd A. Jeffress (ed.), *Cerebral Mechanisms in Behavior: The Hixon Symposium*, New York: Wiley, 1951, 42.

2 Ibid., 24.

argued that the perceptual field as a whole can be computed by machines, such as simple electric relays) and the Gestalt school (who maintained that a machine would never be able to emulate the complex synthetic faculty of the human mind).[3] It constituted a new chapter in the ongoing Gestalt controversy from Europe; yet this time, exiled from Nazi Germany to the United States, its protagonists included mathematicians and engineers in addition to the previous cast of philosophers, psychologists, and neurologists.[4] It was in the aftermath of this debate, in fact, that the expression 'Gestalt perception' gradually morphed, in military reports and academic publications, into the more familiar 'pattern recognition' and gave momentum to the experiments with artificial neural networks.[5]

Cyberneticians such as Warren McCulloch and Walter Pitts proposed a dramatic simplification of the act of perception, at which any art historian would baulk. However, their research provoked a breakthrough in the 'automation of perception' that would unfold, half a century later, in the well-known exploits of deep learning.[6] Before cyberneticians reduced the faculty of perception to a simple binary act of classification (whether or not an image belongs to a given class), Gestalt scholars such as Max Wertheimer, Kurt Koffka, and Wolfgang Köhler had advanced a more complex theory of perception. Gestalt theory famously referred to the fact that one perceives the whole before its parts – a melody, for instance, remains discernible even when it is played in a different scale

---

3 See Claus Pias (ed.), *Cybernetics: The Macy Conferences 1946–1953; The Complete Transactions*, Berlin: diaphanes, 2016; Steve Heims, *The Cybernetics Group*, Cambridge, MA: MIT Press, 1991, chaps. 9–10. Steve Heims called it the 'Gestalten go to bits' controversy.

4 See Mitchell G. Ash, *Gestalt Psychology in German Culture, 1890–1967: Holism and the Quest for Objectivity*, Cambridge: Cambridge University Press, 1998.

5 The German term 'Gestalt' is often translated in English as 'form', although German possesses the distinctive word *Form*. The best English translation is probably 'configuration', which is a form understood in its general structure and internal implications. The term 'pattern recognition' was popularised by *Scientific American* in 1960: Oliver G. Selfridge and Ulric Neisser, 'Pattern Recognition by Machine', *Scientific American* 203, no. 2 (1960): 60–9. See also William D. Rowe, 'Gestalt Pattern Recognition with Arrays of Predetermined Neural Functions', in *International Joint Conference on Artificial Intelligence* (1969): 117–40.

6 Paul Virilio used the expressions 'automation of perception' and 'artificial vision' in relation to the neural network perceptron in *The Vision Machine*, Indiana University Press, 1994. See chapter 9 of this book.

and with different sounds. It pointed to diverse phenomena of image closure, such as the subconscious perception of a configuration in its totality in spite of limited information about its elements (e.g., the perception of a triangle in the abstract relation of three separated elements). Eventually, it was formalised in the principles of *Prägnanz* (proximity, similarity, continuity, connectedness, etc.) and encapsulated in the famous motto 'The whole is other than the sum of its parts'.

Textbooks on machine learning repeat the standard account that McCulloch and Pitts's invention of artificial neural networks was inspired by the neurophysiology of the brain, while overlooking the broader intellectual context that includes their confrontation with the Gestalt school. Though artificial neural networks may be regarded as a brilliant solution to a challenge of automation, it was the Gestalt school that established the perception of a visual form as a key issue to address. The specific problem that Gestalt theory urged cyberneticians to resolve through a computational architecture was the recognition of a pattern, such as an image. Is the recognition of an image equivalent to logical reasoning? McCulloch and Pitts thought so, therefore arguing that any visual pattern (even a complex, fuzzy, and incomplete one) can be fully defined by computing the relation of its elements within the visual field. In their view, the recognition (or classification) of a pattern could be resolved by an algorithm to compute a large input into a simple binary output (0 or 1), representing in this way a simple binary question: 'Does the image belong to a given class or not?'

After short skirmishes during the Macy conferences, the Gestalt controversy convened at a dedicated event: the Hixon symposium on 'Cerebral Mechanisms and Behavior', which took place at the California Institute of Technology, Pasadena, in 1948, and was attended by McCulloch, von Neumann, and Köhler, as well as famed psychologists Heinrich Klüver and Karl Lashley, among others.[7] The Hixon symposium was a watershed event in the history of computation: even John McCarthy admitted that his felicitous term 'artificial intelligence', which was coined for the Dartmouth conference in 1956, had been inspired by his attendance at this earlier symposium as a graduate student in mathematics.[8]

---

7 For the whole proceedings, see Jeffress, *Hixon Symposium*.

8 Nils John Nilsson, *The Quest for Artificial Intelligence*, Cambridge: Cambridge University Press, 2009, 52. McCarthy only cultivated an interest in propositional logic.

At the time, neither of the opposing factions won the controversy. As physicist and science historian Steve Heims noted, 'In a sense, both were romantics, Köhler in his holism and McCulloch in his mechanism.'[9] As a matter of fact, it was von Neumann who proposed a synthesis of Gestalt theories and computational logic that eventually progressed the field and inspired Rosenblatt's neural network perceptron, which implemented statistical calculus in place of McCulloch and Pitts's rigid logic. This chapter aims to re-examine the influence of Gestalt theory upon the history of AI, and to highlight, in particular, the transformation of artificial neural networks from mediums of logical reasoning into instruments for pattern recognition. The Gestalt controversy and the Hixon symposium are proposed, then, as an alternative history of AI that does not relitigate the (failed) enterprise of symbolic AI and its ambitions to mechanise deductive logic, but shifts the focus to the lineage of connectionism and the automation of inductive logic.

The Gestalt controversy is a cognitive fossil of unresolved problems that is today buried at the core of deep learning, and its study can help us to understand the logical form and limits that twenty-first-century AI has inherited.[10] The Gestalt controversy also points to an *optical unconscious* still operative in contemporary machine learning, as the technique of pattern recognition has been extended from visual to non-visual information datasets over the past few decades. In a peculiar twist of fate, it is the mechanisation of perception as pattern recognition that has come to be traded as the mechanisation of cognition, or artificial intelligence. Nevertheless, despite its origins in the automation of vision, the use of anthropomorphic metaphors of perception to describe the

In 1979, he was still keen to clarify that 'machines as simple as thermostats can be said to have beliefs'. John McCarthy, 'Ascribing Mental Qualities to Machines', in *Philosophical Perspectives in Artificial Intelligence*, ed. Martin Ringle, Atlantic Highlands, NJ: Humanities Press, 1979.

9 Heims, *Cybernetics Group*, 241.

10 Google researcher François Chollet has designed the Abstract and Reasoning Corpus (ARC) as a visual test that is deemed simple for humans but still difficult even for current large AI models to solve. Without acknowledging it, ARC covers tasks such as image completion with limited information that had previously been canonised by the Gestalt school as part of their *Prägnanz* principles. See François Chollet, 'On the Measure of Intelligence', preprint (2019), arxiv.org. See also Hannes Bajohr, 'The Gestalt of AI: Beyond the Holism-Atomism Divide', *Interface Critique* 3 (2021): 13–35.

operations of artificial neural networks, as well as today's deep learning, is misleading. As is often repeated, machine vision 'sees' nothing: what an algorithm 'sees' – that is, calculates – are topological relations among numerical values of a two-dimensional matrix. Ultimately, the breakthrough brought about by artificial neural networks was not so much biomorphic as, in this case, *topological*. In other words, it was not so much about imitating the structure of neural networks in the eye's retina as, essentially, about developing techniques of self-organising information to read the visual matrix.

The invention of artificial neural networks

As outlined in the previous chapter, the invention of artificial neural networks was canonised in a milestone paper by McCulloch and Pitts: 'A Logical Calculus of the Ideas Immanent in Nervous Activity' (1943). This was followed by another key text that was a direct response to the Gestalt controversy: 'How We Know Universals' (1947).[11] While the former introduced the idea of a network of artificial neurons to automate logical reasoning, the latter advanced its application to 'the perception of auditory and visual forms' (see fig. 7.1). The passage from the former to the latter marks a logical breakaway. Whereas the 1943 paper proposed neural networks as *deductive machines* for propositional calculus, the 1947 paper pointed towards *inductive machines* for automating pattern recognition. It is worth bearing in mind that McCulloch and Pitts's 1943 paper did not mention computers, because they were not then an established technology.[12]

---

11 On the theoretical background of McCulloch and Pitts, see Tara Abraham, '(Physio)logical Circuits: The Intellectual Origins of the McCulloch–Pitts Neural Networks', *Journal of the History of Behavioural Science* 38, no. 1 (2002): 3–25.

12 At the time of the Gestalt controversy, the digital computer was still an experimental idea and not a technical convention taken for granted. Statements such as 'the human brain is a computer' and 'a computer can think' had a different meaning then than they do now. The design of the computer was in fact resolved *after* the idea of artificial neural networks was introduced. Thomas Haigh and Mark Priestley have also noted that the influence of the Turing machine model on the actual design of the computers is overestimated. Von Neumann was among the first to propose a viable design for the Turing machine, which has since become known as the 'von Neumann architecture'. However, according to Haigh and Mark, von Neumann mostly copied

In order to put the history of AI in perspective, it is important at this point to clarify the difference between deductive and inductive logic. Since Leibniz's idea of a *calculus ratiocinator*, the modern project of pursuing machine intelligence has been founded on the postulate that human logic can be expressed by propositional logic ('if *x*, then *y* is true/false'), which is equivalent to Boolean logic (AND, OR, and NOT operators). A proposition expressed according to this formal logic can be easily encoded into a mechanism made of rotating gears, electric relays, or electronic gates, such as valves or transistors. On a closer look, this type of logic is pursuing a linear form of rationality that replicates the linearity of written language and symbol manipulation according to the rules of *deductive inference* – an approach well exemplified by the Turing machine and the way it executes instructions from a one-dimensional tape. Symbolic AI, expert systems, and inference engines are all examples of this tendency of deductive machines that continued until the 1980s. On the other hand, artificial neural networks – along with all machine learning algorithms – are examples of inductive machines. Whereas deductive logic is the application of a rule, reasoning from the general to the particular, inductive logic involves reasoning from the particular to the general, thereby forming rules of classification. The canonical example is the movement from the discovery that 'each human being dies' to the definition of the rule that 'all human beings are mortal'. This opposition between deductive and inductive logic is key to understanding not only the Gestalt controversy but also the rise of machine learning.

According to Heims's account, in preparation for the highly anticipated visit of the Gestalt scholar Köhler to the Macy conferences in New York, Klüver issued a challenge to the gathered scientists: to formulate a theory on 'how an automaton could perceive Gestalten'. Klüver's challenge was taken up by McCulloch and Pitts, who, with the financial support of the Macy and Rockefeller foundations, were seeking to move

McCulloch and Pitts's 'biological language' in describing the EDVAC computer in 1945. He called the main components of the EDVAC 'organs' and its storage 'memory', and deployed 'an abstract neuron-inspired notation to represent computing circuits'. See Thomas Haigh and Mark Priestley, 'Von Neumann Thought Turing's Universal Machine Was "Simple and Neat", but That Didn't Tell Him How to Design a Computer', *Communications of the ACM* 63, no. 1 (2019): 26–32; and John von Neumann, *Theory of Self Reproducing Automata*, Urbana: University of Illinois Press, 1966, 45.

LEFT
MEDIAL RECTUS

N III

RIGHT
LATERAL RECTUS

N VI

PLOT OF
VISUAL FIELD

+10°   0°
+30°   -10°
        -30°
L10°   -50°
        -60°
L20°
+50°   L30°
L30°   -40°
        -50°
L40°   -60°

ξ

ξ + dξ

PLOT OF
EYE POSITIONS
+30° 0° -30°

RIGHT SUPERIOR
COLLICULUS
(AFTER APTER)

SUP. RECT. N III INF. RECT.
SUP. OBL. N IV INF. OBL.

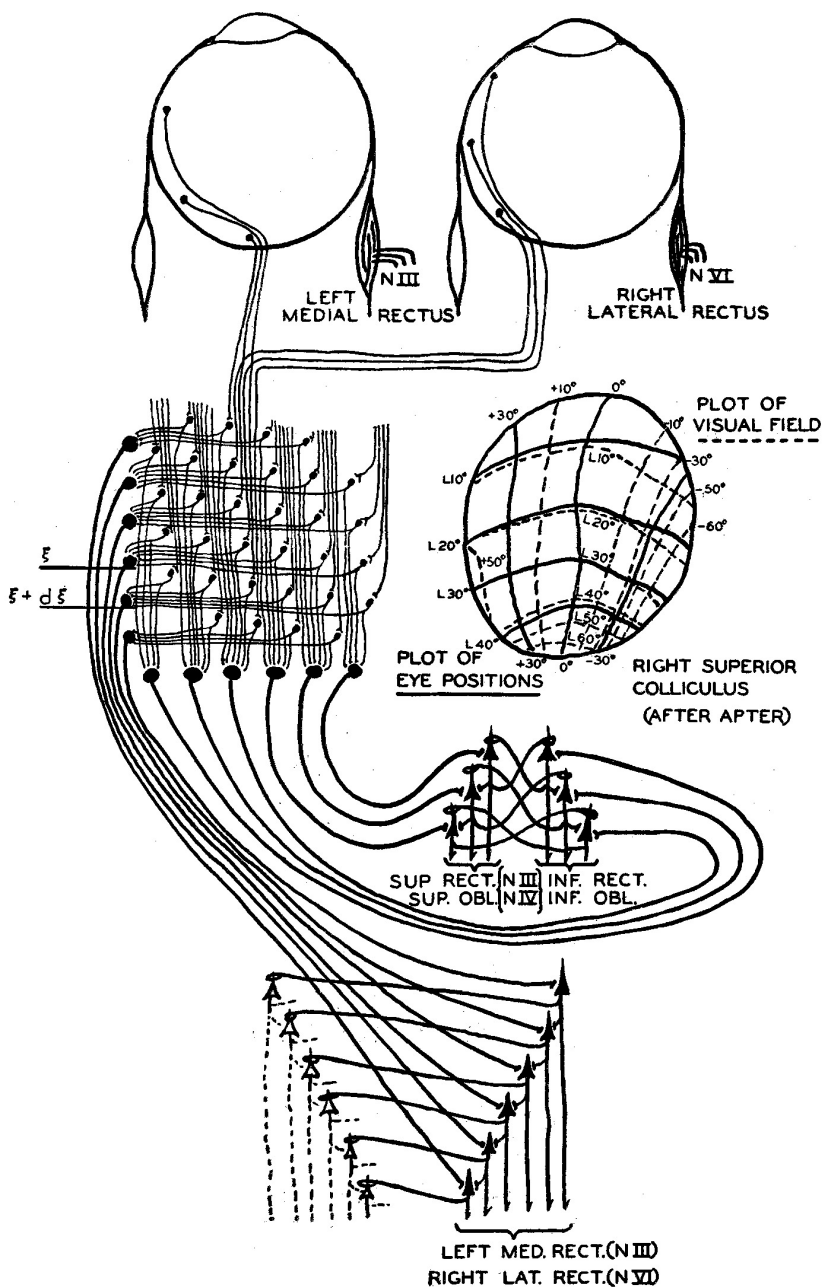LEFT MED. RECT. (N III)
RIGHT LAT. RECT. (N VI)

Figure 7.1. Diagram of the superior colliculus of the midbrain.
Warren McCulloch and Walter Pitts, 'How We Know
Universals: The Perception of Auditory and Visual Forms',
Bulletin of Mathematical Biophysics 9, no. 3 (1947): 141.

beyond their 1943 proof that a finite network of simplified model-neurons could compute anything that can be unambiguously stated in words. Their aim, this time, was to develop neural mechanisms for specific brain functions such as perception, ideally with enough precision to be tested experimentally.[13]

As a result of the discussion, in 1947, McCulloch and Pitts published their paper 'How We Know Universals: The Perception of Auditory and Visual Forms'.[14] Compared to the previous one from 1943, here the concern is not the computation of propositions according to a linear logic but, rather, the recognition of a pattern in a bidimensional matrix. McCulloch and Pitts conceived a 'device' by which 'numerous nets, embodied in special nervous structures, serve to classify information according to useful common characters'.[15] What would this 'device' look like? Given the technical development of the late 1940s, it had to be made, for instance, of photoreceptors that could turn a visual input into a digital image – that is, a grid of pixels measured by numerical values. Without going into detail about colour, presumably such a grid would have to have been encoded in binary digits: value 1 for black pixels, value 0 for white pixels. McCulloch and Pitts's device for pattern recognition was a computing network that would resolve a grid of binary numbers into the binary output 1 (true) if a pattern was recognised or 0 (false) if a pattern was not recognised.

In mathematical terms, a device for pattern recognition is an algorithm that computes the same output for any input that a human classifies as belonging to the same type of patterns. With their mathematical device, McCulloch and Pitts wanted to challenge the Gestalt school's belief in the uniqueness of human cognition and demonstrate that perception can be described algorithmically and automated. Their original exposition of an artificial neural network for pattern recognition reads:

> Numerous nets, embodied in special nervous structures, serve to classify information according to useful common characters. In vision

---

13 Heims, *Cybernetics Group*, 224, 231.

14 Warren McCulloch and Walter Pitts, 'How We Know Universals: The Perception of Auditory and Visual Forms', *Bulletin of Mathematical Biophysics* 9, no. 3 (1947): 127–47.

15 McCulloch and Pitts, 'How We Know Universals', 127.

they detect the equivalence of apparitions related by similarity and congruence, like those of a single physical thing seen from various places. In audition, they recognize timbre and chord, regardless of pitch. The equivalent apparitions in all cases share a common figure and define a group of transformations that take the equivalents into one another but preserve the figure invariant. So, for example, the group of translations removes a square appearing at one place to other places; but the figure of a square it leaves invariant. These figures are the *geometric objects* of Cartan and Weyl, the *Gestalten* of Wertheimer and Köhler. We seek general methods for designing nervous nets which recognize figures in such a way as to produce the same output for every input belonging to the figure.[16]

According to McCulloch and Pitts, it would be possible to compute the principles of Gestalt recognition even when they seem to involve a profound power of abstraction, as in the phenomenon of closure (when, for example, to use again the above-cited example, a triangle is recognised from incomplete lines and non-connected points). Although it may appear reductionist, McCulloch and Pitts in fact proposed a complex dimorphism of dimension between perception and cognition. Gestalt theorists maintained that perception and cognition exist along a continuous isomorphism: the spatial relations of an object are isomorphic with its percept, and thus with its mental image. In contrast, McCulloch and Pitts suggested the possibility of transforming the two-dimensional relations of an image into a one-dimensional representation – basically a code, or logical proposition – without undermining the content or quality of information.

This is a crucial passage in the history of connectionism and AI: McCulloch and Pitts argued that the isomorphism between an image and its logical representation is not necessary. The form of a triangle, for instance, does not need to be memorised as an isomorphic triangle in some parts of the brain but can be distributed across different dimensions (a topological transformation also termed 'internal representation' in deep learning).[17] The two scientists advanced this hypothesis against

---

16 Ibid.
17 It is also for this reason that Kölher accused McCulloch and Pitts's neural model of old-school 'dualism'. Jeffress, *Hixon Symposium*, 69.

the Gestalt school to argue that a visual manifold can be computed by a linear information machine such as that of Turing:

> This point is especially to be taken against the Gestalt psychologists, who will not conceive a figure being known save by depicting it topographically on neuronal mosaics, and against the neurologists of the school of Hughlings Jackson, who must have it fed to some specialized neuron whose business is, say, the reading of squares. That language in which information is communicated to the homunculus who sits always beyond any incomplete analysis of sensory mechanisms and before any analysis of motor ones neither needs to be nor is apt to be built on the plan of those languages men use toward one another.[18]

McCulloch and Pitts's paradigm of perception and cognition was as much as atomistic as multidimensional: it was based on the idea that in order to 'perceive' and 'understand', a biological or artificial neural network can dismember a manifold of two dimensions and project it onto a different configuration of dimensions than the original ones. In a visionary way, they challenged the famous dictum by Gestalt theory: 'The whole is other than the sum of its parts.'[19] They sought to replace Gestalt theory's holistic complexity with another kind of multidimensionality which was purely mathematical.

What the frog's eye tells the frog's brain

At the Hixon symposium, Köhler was the lone delegate representing the Gestalt school. He participated not to defend a mystic power of the mind but to present physical measurements of brain activity during visual perception, precisely 'records of electric currents which have been taken from the heads of human subjects under conditions of pattern vision'.[20] Köhler's study demonstrated that Gestalt theorists were not

---

18 McCulloch and Pitts, 'How We Know Universals'.

19 Originally formulated by Kurt Koffka. Köhler, *Principles of Gestalt Psychology*, New York: Harcourt & Brace, 1935, 176.

20 Wolfgang Köhler, 'Relational Determination in Perception', in Jeffress, *Hixon Symposium*, 200.

opposed to the quest for neural correlates of perception; they simply pursued a traditional method of scientific testing rather than the thinking by analogy of cyberneticians. As Heims has also noted:

> Gestalt psychologists did not capitulate to mysticism, political reaction and vitalism, all of which were connected with the hunger for wholeness. They insisted on empirical studies of phenomena and sought fundamental laws describing structural characteristics of experience, even as they supported the popular opposition to atomistic and mechanistic analyses of the world.[21]

Specifically, Köhler relied on the model of force fields, which was derived from physics and inspired, among others, by his close friend Max Planck, to explain the phenomenon of visual perception happening between eye and brain.

At the Hixon symposium, the *physical hypothesis* of force fields underlying Gestalt perception clashed with the *computational analogy* of brains as finite-state automata. Köhler attacked McCulloch's reductionist approach to the brain, arguing that nerve impulses, when seen and measured by a histologist or neurophysiologist, do not look like logical propositions but simply like . . . impulses![22] Against cybernetics' view of the brain as 'machine arrangements', Köhler proposed the theory of force fields to explain a structural continuity of form (isomorphism) between the stimulus of perception, its neural correlates, and the higher faculties of cognition (such as the faculty of insight):[23]

> Continuity is a *structural* trait of the visual field. It is also a *structural* fact that in this field circumscribed particular percepts are segregated as patches, figures and things. In both characteristics, we have found, the macroscopic aspect of cortical processes resembles visual

---

21 Heims, *Cybernetics Group*, 108.

22 See Köhler's comment to McCulloch in Jeffress, *Hixon Symposium*, 66.

23 Wolfgang Köhler, 'Review of *Cybernetics or Control and Communication in the Animal and the Machine* by Norbert Wiener', *Social Research* 18, no. 1 (1951): 127. On isomorphism, see Abraham Luchins and Edith Luchins, 'Isomorphism in Gestalt Theory: Comparison of Wertheimer's and Köhler's concepts', *Gestalt Theory* 21, no. 3 (1999): 208–34.

experience. To this extent, therefore, vision and its cortical correlate are isomorphic.[24]

In contrast with the gung-ho enthusiasm of cyberneticians for neural correlates resembling logical operators, Köhler was cautious about his own findings.[25] Moreover, contrary to what the term may suggest, Gestalt 'isomorphism' was not a monotonic notion but one that included principles of plasticity such as self-organisation and self-repair in the broader sense.[26] The Gestalt school considered the idea of the mind as a finite-state automaton unfit for describing these higher capacities of the mind for synthesis and abstraction through self-organisation.

On the other side of the camp, the most vocal and virulent anti-Gestaltist was Norbert Wiener, who called holism a 'pseudo-scientific bogy'. Accordingly, Wiener argued that 'if a phenomena can only be grasped as a whole and is completely unresponsive to analysis, there is no suitable material for any scientific description of it; for the whole is never at our disposal'.[27] Wiener shared McCulloch and Pitts's computational view of the mind ideologically, with no hesitation:

> It is a noteworthy fact that the human mind and animal nervous systems, which are known to be capable of the work of a computation

---

24 Wolfgang Köhler, *The Place of Value in a World of Facts*, New York: Liveright, 1938, 217. See Wolfgang Köhler, *Dynamics in Psychology*, New York: Liveright, 1940.

25 'If it were really demonstrated that pattern vision is a matter of continuity physics, then we should be seriously tempted to believe that this holds also for analogous facts in other parts of psychology. Just because I realize that we are dealing with a major problem of psychophysics in general, I feel inclined to move slowly. It may be that the cortical correlate of pattern vision has now become accessible to experimentation. But we can not yet be sure that it has.' See Köhler, 'Relational Determination in Perception', 229.

26 On organicism and plasticity see Kurt Goldstein, *Der Aufbau des Organismus: Einführung in die Biologie unter besonderer Berücksichtigung der Erfahrungen am kranken Menschen*, Den Haag: Nijhoff, 1934, translated as *The Organism: A Holistic Approach to Biology Derived from Pathological Data in Man*, New York: American Book Company, 1939.

27 Norbert Wiener, 'Some Maxims for Biologists and Psychologists' (1950), in *Collected Works with Commentaries*, vol. 4, *Cybernetics, Science, and Society*; *Ethics, Aesthetics, and Literary Criticism; Book Reviews and Obituaries*, ed. P. Masini, Cambridge, MA: MIT Press, 1985, 454–6.

system, contain elements which are ideally suited to act as relays. These elements are the so-called *neurons* or nerve cells.[28]

The idea of artificial neurons and the project of automating Gestalt perception had already been discussed when, in 1948, Wiener published *Cybernetics*, which provided the first summa of digital computation and information theories. In 1951, a few years after the Hixon symposium, Köhler wrote a harsh review of Wiener's book and its chapter about 'Gestalt and Universals', in which he argued that while it is true that machines are better than humans at calculation, such calculation does not constitute thinking or, indeed, capacity of insight (*Einsicht*) into a problem:

> In the relation of human beings to the computing machines, thinking in the proper sense of the term appears to remain the task of the former. Excellent engineers and mathematicians build into such instrument mechanically forced ways of operation which may serve as factually reliable substitutes for certain quantitative activities of the human mind such as addition and multiplication. It is not astonishing that, so far as speed is concerned, the substituted operations of the machines are far superior to anything that human brains can achieve. At the same time, these operations appear to be generically different from those of a human being who is occupied with a mathematical problem . . . The machines do not know, because among their functions there is none that can be compared with insight into the meaning of a problem.[29]

In the mid-1950s, neurophysiologist Jerome Lettvin and biologist Humberto Maturana invited McCulloch and Pitts to contribute to their studies on the neurons of the frog's eye. Their conclusion would be published in 1959 in the famous paper 'What the Frog's Eye Tells the Frog's Brain', which for many represents the final nail in the coffin of the Gestalt controversy.[30]

----

28 Norbert Wiener, *Cybernetics: Or Control and Communication in the Animal and the Machine*, Cambridge, MA: MIT Press, 1961 [1948], 120.

29 Köhler, 'Review of *Cybernetics*', 128.

30 J. Y. Lettvin et al., 'What the Frog's Eye Tells the Frog's Brain', *Proceedings of the IRE* 47, no. 11 (1959): 1940–51.

Their research moved from the simple observation that the optical nerve of higher animals is composed of fewer fibres than the retina. This suggests that the stimulus must undergo compression before it reaches the brain. The question that arises is what exactly happens to the stimulus between the retina, the optical nerve, and the visual cortex? The scientists took the frog as their model organism. They measured the stimuli in the optical nerve using an electrode, while presenting high-contrast shapes to the animal's eye. In this way, they recorded four types of neural behaviours, or 'operations': '1) sustained contrast detection; 2) net convexity detection; 3) moving edge detection; 4) net dimming detection'.[31] By way of illustration, let us consider the second operation – net convexity detection – which is perhaps the most intuitive to understand: it is about the perception of a small object roaming across the visual field, like a flying insect, which is crucial for food retrieval and the survival of the frog. As such, it was colloquially called the 'bug detector':

Against the Gestalt theory's primacy of brain functions, the authors argued that the eye already performs basic tasks of cognition, such as pattern recognition, and sends signals to the brain that are already well-formed *concepts* and not just *percepts*. They contended that the eye is already employing the language of 'complex abstractions' rather than being simply a medium of sensations. They concluded that 'the eye speaks to the brain in a language already highly organized and interpreted, instead of transmitting some more or less accurate copy of the distribution of light on the receptors'.[32] What the eye sends to the brain would not be of a form analogous to the percept (which would maintain *Gestalt* proportions) but rather a semantic proposition – a piece of information interpreting the pattern present in the percept. The collective article also integrated findings from McCulloch and Pitts's 1947 paper, including the observation that an image can be logically defined and 'perceived' by the calculation of the relations between its elements according to a spatial logic.

> By transforming the image from a space of simple discrete points to a congruent space where each equivalent point is described by the intersection of particular qualities in its neighborhood, we can then give the

---

31 Ibid., 1943.
32 Ibid., 1950.

image in terms of distributions of combinations of those qualities. In short, every point is seen in definite contexts. The character of these contexts, genetically built in, is the physiological synthetic a priori.[33]

A digital image is comprised not simply of discrete points but of spatial relations that can be described as 'distributions of combinations'. Notably, the authors seem to suggest here that it is not the neural network per se which comprises the 'physiological synthetic a priori' of thought but rather the spatial logic that the neural network computes.

The biomorphism of the early artificial neural networks (in this case, the imitation of biological neural networks) must eventually be questioned. The human organ that McCulloch and Pitts most often investigated and envisioned as a network of logical operators was not the brain but the eye; and, of the eye, their artificial neural networks implicitly inherited a specific hierarchy of behaviour. In McCulloch and Pitts's 1947 essay, the cooperation among the nodes of the artificial neural network was supposed to resolve the combinatorial geometry of the field of vision, not the propositional logic of a syllogism, as in their 1943 publication. In the Gestalt controversy, then, one does not find a model of cognition to be mechanised but rather several models of perception, since the elements of the perceptual field, rather than mental states, were represented in computational terms. The development of artificial neural networks continued by mostly studying the neurons of the eye rather than the brain: from Lettvin's frogs to David Hubel and Torsten Wiesel's cats (which inspired convolutional neural networks such as AlexNet), the history of connectionist AI is indebted to experiments on animals and their organs of vision.[34]

From a mathematical point of view, the Gestalt controversy was, on one hand, about the transformation of an image into a logical construct and, on the other, about the fact that a logical construct of human reasoning may share the probabilistic features of image perception.[35]

---

33 Ibid.

34 D. H. Hubel and T. N. Wiesel, 'Receptive Fields of Single Neurones in the Cat's Striate Cortex', *Journal of Physiology* 124, no. 3 (1959): 574–91.

35 Historian of science Peter Galison has investigated the relation between the 'mimetic' and 'logical' traditions in the construction of scientific images. See Peter Galison, *Image and Logic: A Material Culture of Microphysics*, Chicago: University of Chicago Press, 1997.

Reading between the lines, the Gestalt controversy was already pointing towards a logical form that would only emerge properly in twenty-first-century machine learning: the idea of a spatial and statistical dimension of information beyond the linear modality of the Turing machine. Logic gates and perceptual fields – the basic concepts of cybernetics and the Gestalt school respectively – were two polarised ideals that the evolution of AI eventually overcame and synthesised into more and more abstract constructs: today the inner logic of machine learning is described, for instance, through entities such as multidimensional vectors, latent spaces, and statistical distributions. The visual matrix (and, with it, the picture plane of modern iconology) has since then evolved into these advanced logical forms, into vast ramifications of statistical inference yet to be fully understood.

The language of the brain is not the language of mathematics

As the epistemologist David Bates has noted, the Gestalt school's scepticism about computational reductionism was nevertheless received by a central protagonist of US cybernetics: the polymath John von Neumann. A key figure of the Manhattan Project and fervent anti-Communist,[36] von Neumann occupied a role of mediator in the controversy. In so doing, he effected a synthesis of Gestalt theory and logical neural networks and, in this way, opened the terrain for Rosenblatt's idea of statistical neural networks. He maintained a strong interest in the mathematical analysis of biological neural networks without claiming either that the brain is a machine or that a machine can perfectly simulate a brain. Instead, he argued that, considering the number of their units (neurons in brains relative to switches in machines), the scale of human cognition cannot be compared with mechanical computation: 'The number of nerve cells in the human central nervous system has been estimated to be 10 billion . . . and nobody has seen a switching organ of 10 billion units; therefore one would be comparing two unknown

---

36 David Bates, 'Creating Insight: Gestalt Theory and the Early Computer', in *Genesis Redux: Essays in the History and Theory of Artificial Life*, ed. Jessica Riskin, Chicago: University of Chicago Press, 2007, 237–59.

objects.'[37] As we will see, ultimately von Neumann did not reduce the brain to the computer – the biological to the logical – but instead envisioned in-between models and implementations to translate one domain into the other.

Von Neumann agreed with McCulloch and Pitts that 'anything that you can describe in words can also be done with the neuron method' but disagreed 'that any circuit you are designing in this manner really occurs in nature'.[38] While he sympathised with McCulloch and Pitts's effort to axiomatise biological neurons into formal ones, he also agreed with the Gestalt school that the systemic nature of form perception is such that it cannot be easily described in propositional terms and deductive logic, unless by exploding the size of descriptions into lengthy and verbose code scripts. As von Neumann explains, arithmetic logic carries manifest limits of scale in its representation of the world:

> Suppose you want to describe the fact that when you look at a triangle you realize that it's a triangle, and you realize this whether it's small or large. It's relatively simple to describe geometrically what is meant: a triangle is a group of three lines arranged in a certain manner. Well, that's fine, except that you also recognize as a triangle something whose sides are curved, and a situation where only the vertices are indicated, and something where the interior is shaded and the exterior is not. You can recognize as a triangle many different things, all of which have some indication of a triangle in them, but the more details you try to put in a description of it the longer the description becomes. [With] respect to the whole visual machinery of interpreting a picture, of putting something into a picture, we get into domains which you certainly cannot describe in those [logical] terms.[39]

Contra McCulloch and Pitts, von Neumann contended that the brain can recognise the complexity of any image precisely because of its probabilistic logic. Arithmetic logic would be then a simplification and approximation of the brain's actual workings and could only partially

---

37 Von Neumann, *Theory of Self Reproducing Automata*, 37. See also John von Neumann, 'The General and Logical Theory of Automata', in Jeffress, *Hixon Symposium*, 3.

38 Von Neumann, *Theory of Self Reproducing Automata*, 46.

39 Ibid., 46–7.

grasp its features. As a result, von Neumann defined arithmetic logic as a metalanguage ('secondary language' or 'short code') that is efficient but not effective at fully describing the underlying probabilistic language of the brain ('primary language' or 'complete code'). He also observed, similarly to Köhler, that the nervous system speaks this primary language using the statistical properties of stimuli ('complete code'), rather than exact markers ('short code').

Given such probabilistic dynamics, von Neumann concluded, counter-intuitively, that in the nervous system a 'deterioration' of arithmetic precision can actually result in 'an improvement in logics'.[40] He high-lighted, among others, that biological neural networks operate with an error tolerance that would set any deterministic computing machine out of joint. A single mistake in a computer programme and the whole machine halts. However, such mistakes never trouble organisms. It is, then, the same probabilistic nature of organisms that make them fault redundant and able to perceive fuzzy and complex figures effortlessly:

> The fact that natural organisms have such a radically different attitude about errors and behave so differently when an error occurs is prob-ably connected with some other traits of natural organisms, which are entirely absent from our automata. The ability of a natural organism to survive in spite of a high incidence of error (which our artificial automata are incapable of) probably requires a very high flexibility and ability of the automaton to watch itself and reorganize itself. And this probably requires a very considerable autonomy of parts. There is a high autonomy of parts in the human nervous system. This auton-omy of parts of a system has an effect which is observable in the human nervous system but not in artificial automata.[41]

There is today consensus that biological neural networks operate in spite of errors, faults, and damages – and precisely also thanks to them.[42] Cyberneticians encountered this idea of neuroplasticity in the cognitive sciences of the time, in particular in the work of the neurologists Karl

----

40 John von Neumann, *The Computer and the Brain*, Silliman Memorial Lectures, New Haven, CT: Yale University Press, 1958, 80.

41 Von Neumann, *Theory of Self Reproducing Automata*, 73.

42 For the influence of this idea on the genesis of deep learning, see Yann LeCun et al., 'Optimal Brain Damage', in *Neural Information Processing Systems* 2 (1989): 598–605.

Lashley (who also presented at the Hixon symposium) and Kurt Goldstein (who had emigrated to the US as part of the Gestalt school diaspora). Lashley tested the effect of artificially induced brain injuries on the memory of laboratory rats, while Goldstein studied the capacity of World War I veterans to reorganise their memories after real brain traumas.[43] However, it was Goldstein rather than Lashley who provided a more systematic model of neuroplasticity. His theory of the brain's capacity to reorganise itself after a trauma was also closely related to the debate on memory localisation. A thesis key to a pioneer of holistic neurology, Constantin von Monakow, was that memory is distributed rather than localised, and this is why it can be recovered after a brain injury (see chapter 8). Von Neumann echoed such theories:

> The main difficulty with the memory organ is that it appears to be nowhere in particular. It is never very simple to locate anything in the brain, because the brain has an enormous ability to re-organize. Even when you have localized a function in a particular part of it, if you remove that part, you may discover that the brain has reorganized itself, reassigned its responsibilities, and the function is again being performed.[44]

It was this idea of a distributed memory in human brains that suggested the delocalisation of memory in machines. Neural networks, with their distributed architecture, were the perfect candidates to accomplish this. As Rosenblatt acknowledged in *Neurodynamics*, von Neumann's comment on distributed memory was a main inspiration for the perceptron (see chapter 9).[45]

---

43 The chapter on the 'pathologies of machines' in Wiener's *Cybernetics* is an echo of these debates on neuroplasticity. See also Matteo Pasquinelli (ed.), *Alleys of Your Mind: Augmented Intelligence Trauma,* Lüneburg: Meson Press, Leuphana University, 2015, introduction; David Bates, 'Unity, Plasticity, Catastrophe: Order and Pathology in the Cybernetic Era', in *Catastrophe: History and Theory of an Operative Concept*, ed. Andreas Killen and Nitzan Lebovic, Berlin: De Gruyter, 2014, 252; Matteo Pasquinelli, 'What an Apparatus Is Not: On the Archeology of the Norm in Foucault, Canguilhem, and Goldstein', *Parrhesia* 22 (May 2015): 79–89.

44 Von Neumann, *Theory of Self Reproducing Automata*, 49. See also *The Computer and the Brain*, 63–8.

45 Frank Rosenblatt, *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms*, report, Buffalo, NY: Cornell Aeronautical Laboratory, 1961, 10. On Lashley, see page 4.

Seen in perspective, von Neumann pursued a different method of inquiry compared to the other cyberneticians. Against the Platonism and intuitionism then popular also in engineering, von Neumann maintained a constructivist perspective on language, logic, and mathematics. He believed that these concepts were not inherent or innate, but rather products of historical development.

> It is only proper to realize that language is largely a historical accident. The basic human languages are traditionally transmitted to us in various forms, but their very multiplicity proves that there is nothing absolute and necessary about them. Just as languages like Greek or Sanskrit are historical facts and not absolute logical necessities, it is only reasonable to assume that logics and mathematics are similarly historical, accidental forms of expression. They may have essential variants, i.e. they may exist in other forms than the ones to which we are accustomed. Indeed, the nature of the central nervous system and of the message systems that it transmits indicate positively that this is so.[46]

This historical relativisation of mathematics and logic is crucial for understanding von Neumann's approach to the computational theory of the mind and the project of building thinking automata. Von Neumann held a more sophisticated position than functionalism, or the thesis of the *multiple realisability of the mind*, which maintains that the same operative model of intelligence can be implemented across different substrates made of neurons, relays, transistors, and so forth. Instead, von Neumann promoted a method of mutual implementation between artificial and natural worlds, that is, a method of *modelling*.[47] Of course, von Neumann was positive about using computers as model machines for studying the brain.

Even so, at the end of his life, when he was invited by Yale University to deliver the Silliman Lectures in 1956, he sought to clarify this point without ambiguity. His lecture series 'The Computer and the Brain' concludes with the declarative title 'The Language of the Brain Not the

---

46 Von Neumann, *The Computer and the Brain*, 82.

47 John von Neumann, 'The General and Logical Theory of Automata', in Jeffress, *Hixon Symposium*, 3.

Language of Mathematics'. In this lecture series, he recognised a key role to the 'secondary language' of mathematics in the knowledge of the 'primary language' of the brain. But, rather than reiterating the computationalism of McCulloch, Pitts, and Wiener, von Neumann – himself no romantic – made at the end a remarkable intervention: he reversed the relation between logic and nature, computer and brain, to the point of suggesting that the study of neurophysiology could one day reshape logic altogether. In the introduction to the lectures's anthology, shortly before his death in 1957, presciently, von Neumann wrote: 'I suspect that a deeper mathematical study of the nervous system [in fact] may alter the way in which we look on mathematics and logics proper.'[48]

Von Neumann's final thoughts were a recognition of the dialectical relation of any system of thought with nature and the external world. They came to terms with the bounds of any system of formalisation and acknowledged the material constraints, including historical ones, of technological and scientific abstractions. These insights surpassed the standard reductionism of the cybernetic mentality. Models of minds and machines do not need to match each other; they often operate through modelling the world rather than reducing it to fixed representations. An epistemology which would acknowledge the flexibility and limitations of modelling is recommended, and yet not a sufficient condition for a progressive philosophy of the mind. As we will see in the next chapter, in fact, the neoliberal economist Friedrich Hayek attempted to turn model thinking into a core principle of economic individualism.

---

48 Von Neumann, *The Computer and the Brain*, 2.