

**MERGING THE NATURAL WITH THE ARTIFICIAL:  
THE NATURE OF A MACHINE AND THE COLLAPSE OF  
CYBERNETICS.**

**ALCIBIADES MALAPI-NELSON**

**A DISSERTATION SUBMITTED TO  
THE FACULTY OF GRADUATE STUDIES  
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS  
FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY**

**GRADUATE PROGRAM IN PHILOSOPHY  
YORK UNIVERSITY  
TORONTO, ONTARIO**

**APRIL 2015**

**© ALCIBIADES MALAPI-NELSON 2015**

## Abstract

This thesis is concerned with the rise and fall of cybernetics, understood as an inquiry regarding the nature of a machine. The collapse of this scientific movement, usually explained by external factors such as lack of funding, will be addressed from a philosophical standpoint.

Delving deeper into the theoretical core of cybernetics, one could find that the contributions of William Ross Ashby and John von Neumann shed light onto the particular ways in which cybernetics understood the nature and behavior of a machine. Ross Ashby offered an account of the nature of a machine and then extended the scope of “the mechanical”. This extension would encompass areas that will later be shown to be problematic for mechanization, such as learning and adaptation. The way in which a machine-ontology was applied would trigger effects seemingly contrary to cybernetics’ own distinctive features. Von Neumann, on the other hand, tinkered with a mechanical model of the brain, realizing grave limitations that prompted him to look for an alternative for cybernetics to work on. The proposal that came out of this resulted in a serious blow against the theoretical core of cybernetics.

Why did cybernetics collapse? The contributions coming from both thinkers, in their own ways, spelled out the main tenets of the cybernetic proposal. But these very contributions led to cybernetics’ own demise. The whole story can be framed under the rubric of a serious inquiry into the metaphysical underpinnings of a machine. The rise and fall of cybernetics could thus help us better understand what a machine is from a philosophical standpoint.

Although a historical component is present, my emphasis relies on a philosophical consideration of the cybernetic phenomenon. This metaphysical dissection will attempt to clarify how a machine-based ontology remained at the core of cybernetics. An emerging link will hopefully lead towards establishing a tri-partite correlation between cybernetics’ own evolution, its theoretical core, and its collapse. It will hopefully show how cybernetic inquiries into the nature of a machine might have proved fatal to the very enterprise at large, due to unsolvable theoretical tensions.

## **Dedications**

I would like to dedicate this dissertation to my supervisor, Jagdish Hattiangadi, whose unique wisdom, selfless guidance and sound advice made this work possible.

I would also like to dedicate this dissertation to my mother, Martha Nelson. Her enduring patience and relentless support always stood as a beacon of hope in the most challenging times.

## **Acknowledgements**

I would like to express my gratitude to Ian Jarvie, for his illuminating insights and meticulous reading of the draft, to Daniel McArthur, for his very positive support of my work, and to Steve Fuller, for his generous feedback on an earlier version of the draft.

I would also like to thank Erica Swanson for proofreading parts of this thesis and Ulyana Savchenko for her support.

I will always sincerely appreciate the time and effort they put into reading my work.

## TABLE OF CONTENTS

Abstract	ii
Dedications	iii
Acknowledgements	iv
Table of Contents	v
List of Figures	vii
Introduction	1
I. Cybernetics: The beginnings, the founding articles and the first meetings.	4
1. The problem of anti-aircraft weaponry.	4
2. The AA-Predictor.	9
3. The pre-meetings and founding articles.	25
4. The Macy Conferences.	34
II. Cybernetics: The book, the club and the decline.	48
1. Norbert Wiener's <i>Cybernetics</i> .	48
2. The Ratio Club.	57
3. The decline.	62
4. Traditional explanations for the collapse of cybernetics.	68
III. Pre-cybernetic context:	
An early 20th century ontological displacement of the machine.	84
1. The "Foundational Crisis of Mathematics" and the response from Formalism.	84
2. A machinal understanding of an algorithm and the material liberation of the machine.	92
3. Alan Turing's strange reversal regarding the question of Artificial Intelligence.	101
4. Cybernetics as a possible missing link regarding Turing's change of heart.	110

IV. Cybernetic tenets: Philosophical considerations.	116
1. Pinning down the core of cybernetics.	116
2. Machines can be teleological.	124
3. Machines can be immaterial.	133
4. Machines are embodied theories.	137
V. Extending the scope of a machine ontology.	143
1. William Ross Ashby's nature-machine equalization.	143
2. The homeostat: A living machine.	152
3. The Macy presentation: Keeping cybernetics honest.	157
4. The backlash: Unforeseen consequences of a behavior-based ontology.	162
5. Un-cybernetic DAMS: From behavior into structure.	169
VI. Emphasizing the limits of a machine epistemology.	175
1. John von Neumann's appropriation of McCulloch and Pitts' networks.	175
2. The letter to Wiener: Post hoc, ergo propter hoc?	180
3. The Hixon Symposium:	
Strengths and weaknesses of the McCulloch & Pitts' networks.	189
4. The Illinois Lectures and the kinematic model:	
Cybernetic materialization of a theory.	197
5. The Theory of Automata: From embodiment to abstraction.	203
VII. Cybernetic tensions: Further philosophical considerations.	207
1. Consequences of the irrelevancy of materiality for the nature of a machine.	207
2. Consequences of a machine's isomorphism with highly complex entities.	215
3. An ontological slippery slope: From complexity to disembodiment.	219
Concluding remarks:	
The nature of a machine and the collapse of cybernetics.	223
Bibliography.	231

## LIST OF FIGURES

Figure 1	147
Figure 2	148
Figure 3	148
Figure 4	154
Figure 5	155

## **Introduction.**

This thesis is concerned with the rise and fall of cybernetics, understood as an inquiry regarding the nature of a machine.

Mechanical devices have certainly been the focus of attention in the past. This thesis is not concerned with those --interesting in their own right, but tangentially related to the notion of a machine. In fact, there seems to be so far no treatise which directly addressed the nature of a machine *per se*. Cybernetics is a relatively recent phenomenon in the history of science, and it is my thesis that it was fundamentally concerned with the notion of a machine. Furthermore, cybernetics might stand as the first time in history where the nature of a machine *tout court* was taken up to constitute the hinge upon which the whole enterprise revolved. This scientific movement, which has arguably not yet received adequate attention from a *philosophical* standpoint, will be addressed at fair length and depth, using it as a guiding instance for the questioning of the nature of a machine.

Delving deeper into the theoretical core of cybernetics, one could find that the contributions of William Ross Ashby and John von Neumann shed light onto the particular ways in which cybernetics understood the essence of a machine. However, both these contributions have been relatively relegated to a secondary role in historical treatments, leaving their philosophical import somewhat in the dark. Ross Ashby pondered about the useful outcomes of extending the scope of “the mechanical” after offering an account of the nature of a machine. This extension would later encompass realms that will show to be problematic for mechanization, and the way in which a machine-ontology was applied would trigger effects seemingly contrary to cybernetics’ own distinctive features. Von Neumann, on the other hand,



tinkered with a mechanical model of the brain, realizing grave limitations that prompted him to look for an alternative for cybernetics to soldier on. The proposal that came out of it might have signified a serious blow against cybernetics' theoretical core.

The contributions coming from both thinkers, in their own ways, spelled out the main tenets of the cybernetic proposal, consistent with their original intentions for further advancing it. But these contributions –likely against the will of the contributors themselves—might have led to cybernetics' own demise. The whole story can be framed under the rubric of a serious inquiry into the metaphysical underpinnings of a machine. The rise and fall of cybernetics could thus help us better understand what a machine is from a philosophical standpoint.

This thesis is made up by two major parts: a historical and a philosophical one. The historical part is itself unfolded in two stages: a) the history of cybernetics proper: its beginnings in the early 1940s, flourishing, and decline a decade later; and b) the pre-history that lead to the cybernetic enterprise. What I tried to do in the latter was addressing the historical and theoretical context that put together the stage for cybernetics to emerge. Indeed a particular development in mathematics during the first half of the 20<sup>th</sup> century, which ended up in Alan Turing's contribution to the notion of computation, seems to have triggered a questioning close to the nature of a machine. Accordingly, an exposition of this evolving context will precede the articulation of the cybernetic main theoretical tenets in the following part.

Although an updated historical account of cybernetics can be worthy in and of itself, my emphasis relies on a philosophical consideration of the cybernetic phenomenon. This latter account is comprised by four parts: a) an articulation of the metaphysical tenets underlying cybernetics (chapters III and IV); b) the contribution of William Ross Ashby (chapter V); c) the contribution of John von Neumann (chapter VI) and d) a philosophical account on the effect of

these contributions (chapter VII). This metaphysical dissection will attempt to clarify how a machine-based ontology lay at the core of the cybernetic edifice. An emerging link will hopefully lead towards establishing a tri-partite correlation between cybernetics' own evolution, its theoretical core, and its collapse.

The first two chapters give a historical overview of cybernetics; they depict its origins, rise, climax and subsequent downfall. They aim to reveal how deeply intertwined the project was since its inception with an epistemology that directly stems from experimentation with machine behavior. The third chapter deals with the preceding events that lead to the emergence of cybernetics, mainly in recent history of mathematics. This chapter will aim to show how a previous transformation of the notion of a machine paved the way for cybernetics to appear. The fourth chapter aims at philosophically dissecting the main tenets of cybernetics, in order to unveil the metaphysical program behind the enterprise at large. This chapter will attempt to show that the entirety of the cybernetic ethos was underpinned by a framework constituted by bold conjectures regarding the nature of a machine. Chapters five and six address the contributions of Ross Ashby and John von Neumann – respectively. These contributions should help uncover relatively hidden tensions in the cybernetic project --both stemming from the nature of a machine. Chapter seven ties up both Ashby's and von Neumann's contributions with the main pillars of cybernetics. Hopefully it will show how cybernetic inquiries into the nature of a machine might have proved fatal to the very enterprise at large, due to unsolvable theoretical tensions. The concluding remarks will briefly recapitulate the thread of the argument underpinning this thesis.

## **I. Cybernetics: The beginnings, the founding articles and the first meetings.**

### **1. The problem of anti-aircraft weaponry.**

On September 1<sup>st</sup> 1939, the Nazis invaded Poland. Two days later, both France and Great Britain declared war on Germany. Soon thereafter, constant aerial raids were launched from both sides. Current anti-aircraft guns were proven reasonably effective against the massive air bombers that began to cast immense shadows over Europe. They flew at a relatively slow pace and in a fairly straight path. But the ground guns' effectiveness against the escorting fighter planes was wanting. These aircraft were smaller, faster and capable of snappily avoiding predictable paths with ease and surprise<sup>1</sup> –the celebrated and sometimes legendary skills of a pilot were largely based upon these extreme maneuverability capacities. Since a shot shell would take up to 20 seconds to reach the desirable altitude for explosion,<sup>2</sup> at a point 2 miles after the original position of the flying target,<sup>3</sup> a very concrete problem for the military emerged.

At that point, inheriting the legacy of the knowledge gained from the first aerial strikes at the end of World War I, the human decision factor was being gradually merged with the relative reliability of a machine designed to fulfill a task. An early proto-computer (a “predictor”) would receive visual information from soldiers on the ground, regarding the possible size of the target, its possible speed, the height of its path, the wind situation at the moment, etc. The analog machine would then plot the direction towards which the soldier had to aim the gun. If all worked well, the soldier had to worry mainly about constantly and accurately pointing at the moving target and re-charging ammunition --while the predictor would take care of the angle of

---

<sup>1</sup> Mindell 1995, p. 78

<sup>2</sup> Conway & Siegelman 2005, p. 110

<sup>3</sup> Mindell 2002, p. 86

shooting. Ideally, this synergy of man and machine would create a cloud of exploding shells around the target's flying path. Shrapnel should sufficiently damage the aircraft's fuselage for it to go down, become an easier target due to damage, or abort its mission and retreat.

It is now widely agreed that the first air raids over the United Kingdom, the "Battle of Britain", was the first military failure for the German advance over Europe. Further, British retaliation against Germany proved to be devastating for the Nazis. The English air forces inflicted a serious blow to the Nazi leadership's image, coordinating bombings to happen precisely while important political speeches were taking place –topping it with one on Hitler's birthday. The German morale was however unbroken, and the retaliation against Britain, in terms of air strikes, was massive --it is referred to as the "Blitz" ("Lightning" in German). Despite the fact that the German mission was more successful this time, it still failed to accomplish its main goal: to cripple the military response capacity of the United Kingdom. In fact, raids over Germany, now with help from the Allies, increased, and overall the destruction ended up being worse on the German side.

Although the United States chose to stay out of the war in the beginning, it gradually modified its stance, starting by helping its soon-to-be-Allies with supplies –a position captured in President Franklin D. Roosevelt's famed "Arsenal of Democracy" catchphrase. American elites were, expectably, following the events closely. Despite an alleged majority opposing joining the war, sentiments of preparation and readiness for the dramatic possibility were strong. In fact, there already were secret talks taking place between Great Britain and the United States regarding possible courses of action in the eventuality that the latter would enter the conflict. Scientists and academicians were not excluded from the patriotic sentiment of becoming useful for their nation given the distinct possibility of war. Quite the contrary.

Norbert Wiener (1894–1964) was in his forties an already famed mathematics professor at the Massachusetts Institute of Technology (MIT). Of Russian Jewish ancestry, he married a bride brought from Germany, had two daughters, and ended up attending a Christian Unitarian church with his family. A child prodigy, Wiener already knew the alphabet at 2 years old, finished his mathematics bachelor's degree at Tufts University at 14, and after studying philosophy both at Cornell and Harvard universities, he ended up obtaining a Ph.D. in mathematics from Harvard University at 18. After further studying in Europe under Bertrand Russell and David Hilbert, he finally landed a position at MIT, where he remained for the rest of his career. Between his return from Europe and the MIT professorship, World War I was still being fought. Wiener attempted to enlist in the military, but he was rejected, twice –his eyesight was remarkably poor. The third time he applied, this time as a lowly infantry soldier, he was accepted. WWI ended however just days later and to his chagrin he was quickly sent back home.

World War II broke out while Wiener was teaching at MIT. This time, equipped with prestigious degrees and academic fame, he approached the military via a proxy more related to his skills, addressing the scientific bureaucratic body set up to face the war. Vannevar Bush (1890-1974), a doctor in electrical engineering also from MIT and President Roosevelt's consultant in all matters scientific, insisted on creating a special unit in charge of linking military and civilian knowledge. WWI allegedly showed that these two remarkably distinct and traditionally opposed human groups are not naturally inclined to work in tandem. Collaboration and synchronization between both had to be fostered –even enforced. Thus, the National Defense Research Committee (NDRC) was created, with Bush as its Chairman. Soon later, the NDRC, which had authority for organizing applied scientific research, but *not* for funding and developing projects, was superseded by the Office of Scientific Research and Development

(OSRD). This newer office had the power to give the go-ahead to projects, enjoying access to practically unlimited resources for funding. The NDRC, which ended up being the research branch within the bigger OSRD, was divided in five divisions:

Division A - Armor and Ordnance

Division B - Bombs, Fuels, Gases, & Chemical Problems

Division C - Communication and Transportation

Division D - Detection, Controls, and Instruments

Division E - Patents and Inventions.

Division D, the one that concerns this chapter, was itself divided in four sections:

Section 1 - Detection

Section 2 - Controls

Section 3 - Instruments

Section 4 - Heat Radiation<sup>4</sup>

Thus, the division and section to which Wiener had to apply was D-2, chaired by the doctor in engineering and mathematician Warren Weaver. In his letter, Wiener made sure that the upper echelon of the newly appointed scientific administration would acknowledge his eagerness to be at the service of his country through the military, clearly stating his wish: “I hope... you can find some corner of activity in which I may be of use during the emergency”.<sup>5</sup> Not long after, to Wiener’s delight, he received word from NDRC letting him know that the military was interested in one of his proposals.<sup>6</sup>

---

<sup>4</sup> Mindell 2002, ch. 7

<sup>5</sup> Galison 1994, p. 228

<sup>6</sup> Indeed Wiener proposed several projects. Among them, he was advocating for the positioning of immense aerial balloons to be massively ignited in a prolonged manner once the enemy would fly close-by. This rejected proposal might have been refused due to its “passive defense” character. There already were projects that came to fruition which placed static “barrage balloons” tethered to the ground. This would provoke two desired outcomes. The enemy planes would have to fly higher, thus becoming more available as target --which is more difficult when they

The United Kingdom had secretly developed a “cavity magnetron” –a type of radar far more advanced than anything similar developed anywhere else. This was probably Britain’s best kept “secret” at the time (the Germans eventually found out). The British were willing to share this ground-breaking weapon accessory with the Americans in exchange for help in quickly developing anti-aircraft weapon systems –with which such an invention should interface. Vannevar Bush was aware that American anti-aircraft technology was not precisely cutting edge, having barely improved after WWI. Closely following the way the war was unfolding in Europe, he understood that such was precisely the kind of weapon improvement that America urgently needed in case the war reached the Pacific coast. He agreed to the exchange. America would be helping the UK against the advancement of the “Axis” (Germany, Italy and Japan) while dramatically improving its own anti-aircraft defense.

Early on, Wiener took issue with a perceived inadequacy of contemporary anti-aircraft weapon systems, thus wishing to approach the problem from a different angle –from a perspective closer to his own fields of expertise. In the past, Wiener had spent years working on the mathematical tractability of diverse chaotic phenomena, trying to find recognizable patterns via formulaic equations. For instance, he devised a numerical “filter” for the allegedly random movement of dust particles (Brownian motion) --which was later known as the ‘Wiener filter’. Given the relative ineffectiveness of the current anti-aircraft predictors, he saw a probable application of similar principles for the construction of an improved machine –one that could predict the position of the flying target by extrapolating its likely future location, statistically deducing it from its previous locations. Taking into consideration all the relevant variables

---

fly at a low altitude. Also, if they would refuse to fly higher, they would risk becoming entangled with the wires and crash. As it turned out, pilots soon easily overcame this difficulty, and these defenses stopped being used (Galison 1994, pp. 228-229).

available and finding an “average”, the machine would predict where the enemy object would be in the near future, thus producing its best possible shot at neutralizing it.

## **2. The AA-Predictor.**

The utmost importance of interfacing pure theory and embedded practice was always in Wiener’s mind --and in a personally dramatic manner. His plans for earning a Ph.D. in biology were trumped by his utter manual clumsiness while making his way through the mandatory laboratory coursework.<sup>7</sup> Once a professor, his poor sight led him to famously walk down the campus hallways with one hand always touching the wall, allegedly following the belief that tracking a maze’s wall would eventually take him to the exit. Wiener’s inability to shoot and hit even a big immobile object, “a barn among a flock of barns”,<sup>8</sup> spelled for him, much to his dismay, non-admittance to military service. The reason why he famously walked with his head tilted back was that he instructed his optometrist to divide his thick glasses in two: the upper half would allow him to read, whereas the lower half to see afar. Thus he used to walk towards and approach people in a manner of looking down on them --which gained for him some antipathy. He was just severely myopic.<sup>9</sup>

It is then unsurprising that as soon as he received funding for his project, finding a person with the appropriate hands-on skills was the first task. Nobody without at least passing knowledge in applied electronics (e.g. radio building) would even be considered as a potential assistant. “There is nothing in a drawing in abstract algebra or topology... which would prepare

---

<sup>7</sup> Hayles 1999, p. 99

<sup>8</sup> Conway & Siegelman 2005, p. 124

<sup>9</sup> For the famous anecdotes about his legendary blindness and eccentricities, see Jackson 1972.



one in a way to cooperate with engineering design”.<sup>10</sup> From now on, the scientist involved in wartime projects would have to get out from the comfort zone of “clean” and pure science and “get dirty”, so to speak, with real-world constraints.

If I could not find these talents joined together in a single man, I would be forced to assemble a team of people each with particular talents in one field and a general knowledge of the others. In this team I would probably be the only mathematician thus the project as a whole would concern other engineering groups as much as the American Mathematical Society and it would be necessary for me to cross ordinary professional lines.<sup>11</sup>

In January 1941, Wiener hired Julian Bigelow, a talented and down to earth electrical engineer from MIT. Most relevantly, he was an active airplane pilot as well. He joined Wiener’s team for exclusively working in the circuit design of the predictor, but his skills would later prove to be essential for a much bigger project.<sup>12</sup> As alluded to above, Wiener wished to utilise knowledge gained from his previous mathematical investigations into chaotic motion in order to improve the capabilities of the current predictors –so that the machine could better foretell the future position of a flying enemy target. But that part was just half the story. Such an approach, even if potentially of great value, had to be connected with the unavoidable “real world” constraints in order to bring something truly useful --for the military at any rate-- to fruition. Fortunately, Wiener’s knowledge proved to be capable of modeling and articulating Bigelow’s grounded know-how expertise as a pilot in interesting ways.

Bigelow emphasized to Wiener that when a pilot is shot at, the scope of possible sudden movements he would try in order to avoid enemy fire was not boundless. The dynamics of

---

<sup>10</sup> Galison 1994, p. 235

<sup>11</sup> Ibid.

<sup>12</sup> Bigelow would go on later to fulfil an important role as an irreplaceable collaborator for similarly inspired projects that came after this first one, which spanned a new way of thinking about science, engineering, and even philosophy, which Wiener later baptised as “cybernetics” (more below).

contact between the vessel and air, the resistance of the latter against the former's fuselage, the torsion forces that the airplane's body can maximally handle... all those aspects form the physical landscape that limit the amount of actual 'moves' and 'tricks' available to the pilot. Indeed a wrong sudden movement could end up breaking the plane. Wiener grasped the nature of this limited range of actions crisply. He then stated that there was a way to insert this limitation into the operation of the machine. Statistics was to provide the way: since the limited range of movement was known, one could find an average for the likelihood of a certain move to occur. In fact, the same reasoning could also be applied to the infamous sluggishness of the heavy anti-air guns operating from the ground. The predictor to be built was going to dramatically improve its prognostications using the power of statistics. In Wiener's mind, that was the feature that was going to provide this machine its unique capabilities.

As mentioned above, towards the end of WWI interest in anti-aircraft weaponry substantially grew. Accordingly, the notion of a "servomechanism" was given a good deal of attention, and thus it was right from the start at the crux of Wiener's designs. Indeed the patent for which he applied was firmly within the realm of predictors --themselves paradigmatic examples routinely articulated in treatises about servomechanisms. Although most work had been done "empirically" --without a "school of thought" or anything pertaining to a mature technology-- there were two main writings circulating around the American war laboratories. One was by Hendrik Bode --a doctor in engineering working from Bell Laboratories-- and another one, less exhaustive but more encompassing, by LeRoy MacColl --a mathematician from Columbia University. Wiener chose MacColl's treatise<sup>13</sup> and soon introduced it to Bigelow.

The common root between "servomechanism" and "servant" is not coincidental. Commonly regarded as an Anglicism, the noun "servant" was adopted in its unchanged form

---

<sup>13</sup> MacColl 1945

from the French language. The French *servant* denotes an entity fulfilling a role, but it has the connotation of being hired --and thus, remunerated. However the older Latin *servus* (from where the French word itself derives) has a more direct referent to the notion of a “slave”. A *servomoteur* (French) is an engine subservient to a greater, more encompassing mechanical task. Its enslavement is firmly put in place by means of a phenomenon within control theory whose rediscovery was in itself, for Wiener, a sort of revelation: Negative Feedback. This was a type of corrective input that would fix possible deviations from the system’s designed task. As we will see later, negative feedback as a major useful element entered Wiener’s mind after facing the very real-world problems he was choosing to confront. Of the two classical writings on servos alluded to above, the one chosen by Wiener (McCull’s) firmly treated servo-mechanisms as paradigmatically “enslaved” by means of negative feedback. Thermostats and self-guided torpedoes would be prime examples of devices reliant on these servomechanisms. In fact, Wiener would not regard his own system in a fundamentally different manner than these two devices. He theoretically expanded the main tenets underlying one self-regulated system (that of a thermostat), to envision another similar one (that of an anti-aircraft weapon system). Seen from afar, both systems, thermostat and anti-aircraft weapon system, would be one and the same.

The attentive reader will notice that while in the thermostat all the comprising parts of the closed system are inanimate and mechanical in nature (gears, levers, etc.), Wiener’s system will have human beings at two points (the pilot and the gunner) interlaced with plane and gun. It did not take long for Wiener to see, after he began working out the systems’ circuits, that those two entities would be the source of irregularities for the predictor. “Irregularities”, however, were not a scary notion for Wiener. He had been dealing with phenomena of this sort while studying the allegedly random (Brownian) motion of dust –hence the confidence in his own usefulness for the

cause of the war. Hence Wiener proceeded to treat them mathematically. This mathematization would pre-require some sort of measured report of these irregularities, captured within a controlled environment.

As advanced above, Bigelow's input would prove itself essential for the construction and operation of the new predictor –and more. As his airplane piloting know-how was in fact complementing his remarkable engineering skills, he designed and built a secondary machine at Wiener's laboratory that would begin producing the data needed for the principal machine, the predictor. Bigelow's machine, once constructed, began to receive test subjects in a laboratory-controlled environment, obtaining valuable data from their behavior.

The “gunners” recruited to use Bigelow's machine had to align, via the manipulation of a delayed-effect joystick, two large light dots projected against a wall –a white one representing the enemy plane and a red one being the gunner's beam. The recreated situation aimed to simulate the experience of a gunner aiming his heavy and slow gun turret towards the flying target and successfully hitting it with exploding shells. As mentioned before, airplane fighting made of the avoidance skills a matter of celebrative tradition –after all, the survival of the pilot was the positive outcome. This skilled irregular flying pattern had to be somehow represented in the movement of the white beam. Accordingly, Bigelow had the ingenious idea of projecting the white dot on the four walls of the classroom in a circular manner. When the dot would pass from one side of the corner to the other, it “jumped”, mimicking the sudden and deceptive avoidance behavior of a trained pilot. The “gunmen” reported how difficult this sudden move made the task of aligning their beam to the target, having to move the rigged joystick well ahead before the white spot would arrive at the predicted target location. Wiener was happy with the machine's

simulation, and the data being produced by the gunners' behavior was later input into Wiener's predictor.<sup>14</sup>

Around the same time, and unbeknownst to Wiener, Warren Weaver was (due to war urgency) funding another anti-aircraft weapons system project via NDRC's Fire Control section - of which he was chairman. Of special priority was the quick construction of a predictor that could be interfaced with Britain's shared radar machine. At the time, up to fourteen people were required to provide the visual input and manual direction of the heavy and slow gun turrets. In Wiener's mind, this process was "fascinating, absurdly cumbersome and ripe for improvement".<sup>15</sup> Given the relentless German bombing over England, automation of the whole system indeed carried a sense of urgency.

Neither Wiener nor Bigelow, despite their fairly high security clearances, knew at the time of the Radiation Laboratory<sup>16</sup> functioning in a close-by building at MIT itself. The British secret invention was being studied there. Both were invited, however, to the Bell Laboratories, where a simpler predictor, soon to be coupled with the British radar, was being constructed under the supervision of that other theorist of servomechanisms, Hendrik Bode. The engineers at Bell were dismissive of Wiener's project, feeling that theirs was less complicated and still able to

---

<sup>14</sup> Peter Galison has a different view on the nature of Bigelow's machine, which allegedly recreates the kinaesthetic dissociation present in a pilot's manoeuvres. This is how Wiener describes the psychological phenomenon: "... the pilot's kinaesthetic reaction to the motion of the plane is quite different from that which his other senses would normally lead him to expect, so that for precision flying, he must disassociate his kinaesthetic from his visual sense." (Galison 1994, p. 236). The air resistance against the plane's steering parts provides a feedback to which the pilot's own visual input has to eventually fit. It might give the sensation that the machine itself is resisting its direct control, and learning how to dominate it requires skill and experience. In fact the tradition of giving proper names to flying machines might actually stem from this "wrestling" between man (what the eye gives) and machine (what the physical controls give back in response). For Galison, Bigelow's machine represented not the enemy pilot (white dot) and the friendly gunman (red dot), but the enemy pilot (white dot) and the friendly pilot (red dot). Galison uses a vague recounting from a later book of Wiener (1956, p. 250-255), and a report by the American mathematician George Stibitz. Conway and Siegelman use declassified records from 1941, where Bigelow himself explains the machine workings at a Bell conference. The amount of quotes of Bigelow himself seems to make Conway's understanding the right one. See Conway & Siegelman 2005, pp. 110-115.

<sup>15</sup> Conway & Siegelman 2005, p. 110

<sup>16</sup> Rad Lab, in short. This nickname was a *non sequitur* from the original one, in order to trigger confusion and maintain its secrecy (Conway & Siegelman 2005, p. 115).

accomplish the same task. Wiener and Bigelow were unmoved, pointing out the almost simplistic and unsophisticated nature of the machine built by these engineers. They did not take the statistical considerations pointed out by Wiener for its construction, and it was exclusively based on calculating the future angle of the plane using its previous location. When the predictor was coupled to an anti-aircraft gun and tested in the field, it did fairly well --but far from perfectly. However, it did not have to be particularly accurate, since the shells were altitude-sensitive, and they would likely explode close to the plane fuselage anyway, damaging it with shrapnel. Given the urgency of the situation, Wiener and Bigelow gave a cold approval of the project, realizing that the Bell predictor had to be deployed anyway given the Axis' ongoing devastation.

Not long after Wiener was finally notified of the secretive Rad Lab operating very close to his own, and was asked to train the group of young scientists gathered in there. Since Wiener's ultimate reason for working on this whole project was patriotic in nature, he put understandable professional rivalries aside and agreed to help. However, schooling young and eager Ivy League minds would not be an easy task. From the point of view of scientific practice *vis à vis* different disciplines, there was a feature that lay at the center of Wiener's operational concerns. If we follow a traditional distinction between the notions of multidisciplinary, interdisciplinarity and transdisciplinarity,<sup>17</sup> one could arguably make the point that the scientists recruited by the bureaucratic overseeing institution (NDRC) would likely come equipped with a multidisciplinary perspective --open to what other disciplines would have to offer, but ultimately unwilling to substantially cross their own field boundaries. Wiener expected for his project (and indeed for

---

<sup>17</sup> "Multidisciplinary draws on knowledge from different disciplines but stays within their boundaries. Interdisciplinarity analyzes, synthesizes and harmonizes links between disciplines into a coordinated and coherent whole. Transdisciplinarity integrates the natural, social and health sciences in a humanities context, and transcends their traditional boundaries." (Choi 2006, p. 351)

any *working* project) in contrast, a radical interdisciplinarity –based upon what other fields could provide, holistically creating a separate entity. Partly due to adherence to already familiar ways and reluctance to change (a usual signature of large administrative bodies), this pre-requirement proved to be particularly difficult to fulfill –so difficult in fact that the project was at some points almost brought to a permanent halt. Wiener complained...

New members of the staff of your Laboratory are recruited from the theoretical physicists or mathematicians of the country, or indeed anywhere except from among the ranks of communication engineers in the strictest and narrowest sense of the term...to turn such an individual loose in your laboratory without special training, no matter what a big shot he may be in his own subject, is like ordering a corn-doctor to amputate a leg.<sup>18</sup>

The anchoring in reality by means of machine-construction should provide a hub for the encounter of heretofore separated disciplines. They all had to converge in one physically embodied project. This was an instance where an applied interdisciplinarity had to be the norm.<sup>19</sup> Realizing that it was not going to happen, Wiener renounced this project. With renewed motivation the scientific duo returned to its own project, aware that theirs had the potential to engender a much better prognosticator. A considerable amount of data was produced by Bigelow's machine, and the results finally began to be entered and crunched by Wiener's new predictor.<sup>20</sup> Soon the mathematician grasped two identifiable phenomena. One seemed positive for the project, the other one less so.

---

<sup>18</sup> Galison 1994, pp. 240-241

<sup>19</sup> As we will see later, a similarly inspired transdisciplinarity –the encounter of several disciplines, natural and social, that would as a whole transcend boundaries—was about to occur just a few years down the road under the banner of “cybernetics”.

<sup>20</sup> The machine would be fed with data produced by observation of aircraft behavior. Beforehand, Wiener's statistics would predict where the target would be 1-2 seconds before it reaches its position, extrapolating from observations taken 10-20 seconds in the past. All this information would be input into the machine, which would update itself every time new “real life” input is fed into it, rendering in advance the precise spot where and when the gun should fire. Since Wiener's machine compressed time by a factor of 4-5, a 1 second prediction would amount to 10 seconds

On the one hand, pilots were exceedingly consistent while in the simulator. This finding seemed to match Wiener's assumption that humans under stress tend to act repetitively. For Wiener, this also cast a light on the apparatus that made it look like a precursor to a learning machine. The predictor seemed to be good at learning how to accomplish a pilot's death in a rather "personalized" manner. Speculating, the machine could be successful at knowing how to effectively kill *one* individual, statistically extrapolating data based upon his past behavior. In the area where the predictor was shining, it seemed that it was indeed managing to foretell the future. As the visiting supervisor George Stibitz put it in his working diary, referring to the impression the machine caused on Warren Weaver...

It simply must be agreed that, taking into account the character of the input data, their statistical predictor accomplishes miracles... For a 1-second lead the behavior of their instrument is positively uncanny. WW [Warren Weaver] threatens to bring along a hack saw on the next visit and cut through the legs of the table to see if they do not have some hidden wires somewhere.<sup>21</sup>

Indeed some technicians jokingly made allusions to the "gremlins" that were surely inhabiting and controlling the machine –venting the feeling of "uncanniness" present when the non-living

---

in real-life. And since a shell would take 20 seconds to explode from the moment it is shot, it would seem that Wiener's machine was indeed in the right track. Wiener's lay man explanation of his machine goes as follows:

The proposed project is the design of a lead or prediction apparatus in which, when one member follows the actual track of an airplane, another member anticipates where the airplane is to be after a fixed lapse of time. This is done by a linear network into which information is put by the entire past motion of the airplane and which generates a correction term indicating the amount that the airplane is going to be away from its present position when a shell arrives in its neighborhood (Masani 1990, p. 182).

David Mindell explains it with somewhat more detail:

Wiener and Bigelow... turned to statistics and designed a new predictor based on "a statistical analysis of the correlation between the past performance of a function of time and its present and future performance." The network calculated a future position of the target based on the statistical characteristics of its past performance (its autocorrelation). It then continually updated its own prediction as time passed, comparing the target's flight path with previous guesses. A feedback network converged on guesses that minimized this error (Mindell 1990, p. 278)

For a description of Wiener's statistical prediction, see Galison 1994, p. 237. For an accessible explanation of the machine's inner workings, see Mindell 2002, ch. 11. For a technical account of Wiener's machine, see Bennett 1993, ch. 14.

Given the later fate of this machine, the reader should bear in mind that its importance for cybernetics is not anchored in the details of its inner workings. Instead, its relevance is based upon the insights it provided Wiener while trying to "mechanize" the behavior of both the pilot and gun-man, making it overlap with that of the machine.

<sup>21</sup> Galison 1994, p. 243



behaves in a life-like manner. Warren Weaver, after acknowledging the important result, pointed out that now the challenge remained regarding interfacing it with the rest of the weapons system in order to deploy it in the field. After all, it still had to be hooked to the British radar invention and the complicated manned gun turret. This proved to have many real world details that presented concrete problems for Wiener and Bigelow.

Coupling machines in order to accomplish a task was of course not a new feat. In fact, history registers several accomplishments where inter-regulated interface mechanisms worked successfully –from the Ancient Greeks on.<sup>22</sup> However, there were some instances where they seemed to not work in tandem --and for somewhat mysterious reasons. For example, gun turrets placed in European naval vessels during the 18<sup>th</sup> century sometimes tended to engage in recurrent “spasms”, swinging erratically and out of control for periods of time.<sup>23</sup> A similar occurrence was later recognized in sound engineering. A microphone and a speaker in close range would trigger a high sound loop --the so-called “howling” phenomenon--, when the amplification of both sources went unlimited.<sup>24</sup> As Bigelow reminded them, these instances of overcompensation in machine interfaces were gradually fixed by “negative feedback”: capturing a portion of the output and feeding it back into the stream, balancing it.<sup>25</sup> Wiener was heavily

---

<sup>22</sup> See Mayr 1970, ch. II

<sup>23</sup> See Bissell 2009

<sup>24</sup> Ibid.

<sup>25</sup> If the output of a system suffers variations that preclude the system from completing its task (due to changes in its input), a portion of the output could be feedback into the input, thus balancing its output. In Wiener’s words, negative feedback can fix the system’s “defective behavior and its breaking into oscillation when it is mishandled or overloaded”. It is “negative” because, “the information fed back to the control center tends to oppose the departure of the controlled from the controlling quantity” (Wiener 1948, pp. 97-98). Ashby uses the following explicative example:

A specially simple and well known case occurs when the system consists of parts between which there is feedback, and when this has the very simple form of a single loop. A simple test for stability (from a state of equilibrium assumed) is to consider the sequence of changes that follow a small displacement, as it travels round the loop. If the displacement ultimately arrives back at its place of origin with size and sign so that, when added algebraically to the initial displacement, the initial displacement is diminished, i.e. is (commonly) stable. The feedback, in this case, is said to be “negative” (for it causes an eventual *subtraction* from the initial displacement) (Ashby 1956, p. 80)

impressed by the importance of this negative feedback in mechanisms with a purpose or task built-in. He weighed the revolutionary importance that this engineering insight had for the industrial revolution<sup>26</sup> and did not hesitate in asking help from an old friend, who allegedly found similar occurrences in living organisms.

Arturo Rosenblueth, a Mexican Jew of Hungarian heritage who studied medicine in Paris and Berlin, was the protégé of Walter Cannon –a famed biologist who pioneered work with self-regulating mechanisms in living tissue (homeostasis). Wiener met Rosenblueth a decade before at the Harvard Medical School informal dinners organized by the latter. These dinners were an environment for free intellectual discussion on scientific matters. Both men being avid talkers who liked to convey their ideas and research with passion, Wiener clicked with Rosenblueth at once, and kept uninterrupted contact after that.

Upon being told of the problem Wiener was dealing with, Rosenblueth found it remarkably similar to a reported phenomenon occurring in some human patients: “purpose tremor”. This situation would manifest in the person when she would like to perform a simple, isolated, modular, task – like picking up an object from a table. The person’s arm would instead engage in uncontrollable swinging, back and forth, missing the object. If the person were handed a piece of candy, and was instructed to put it in his mouth, a similar occurrence would happen, missing the target –in this case, her mouth. Rosenblueth offered his insights to Bigelow and Wiener, and at least in theory, they seemed to have helped them find a solution. However, the implementation of these new self-regulatory aspects in the actual, physical instances of the machinery would prove to be a big challenge. Nevertheless, they soldiered on –as the war and devastation demanded.

---

<sup>26</sup> Embodied in the important role of James Watson’s mechanical “governor”, later mathematically treated by James C. Maxwell (Maxwell 1868). More on this in chapter II, section 1. For a brief history of control servomechanisms in the last century, see Bennett 1996.

There was another realm that emerged as fundamental for accomplishing the task of building the new predictor. Although this aspect at first looked like a problem just for Wiener's project, it planted the seed of a far-reaching change of outlook. In fact, it expanded Wiener's framework to areas never before successfully treated mechanically. It was said above that Wiener found the pattern of each pilot's behavior consistent with itself on an individual basis. However, the distressed behavior from one pilot to another did not seem to reveal any common pattern. The machine was falling short at capturing a recognizable behavior pertaining to humans *in general* --in order to wipe them out effectively, which was the main point of the military funded project in the first place. The human element in the system (the gunner and the pilot) was the main source of irregularities. Thus, the issue of "behavior" in general suddenly occupied the center of Wiener's concerns. Given his general background in undergraduate and graduate studies in philosophy and biology, Wiener was well aware of the school of "behaviorism" in psychology. Considering his own struggles regarding the rigorous articulation of behavior in general, he referred to the contemporary scholarship on the subject as having missed the point:

Behaviorism as we all know is an established method of biological and psychological study but I have nowhere seen an adequate attempt to analyze the intrinsic possibilities of types of behavior.

This has become necessary to me in connection with the design of apparatus to accomplish specific purposes in the way of the repetition and modification of time patterns.<sup>27</sup>

To be sure, preoccupation with human behavior as a source of randomness and unpredictability --potentially fatal when present in the chain of command-- was naturally at the core of the military mind. In the line of attack or defense, the human factor was sometimes regarded as less reliable than the mechanized parts of the weapon system --itself less prone to break, more

---

<sup>27</sup> As expressed in a letter to his friend the British geneticist John Scott Haldane (Kay 2000, p. 80).

predictable, and hence more reliable. As a veteran Admiral put it to a group of NDRC scientists when they were invited aboard a war naval vessel:

Twenty-five hundred officers and men: gentlemen, twenty-five hundred sources of error.<sup>28</sup>

Although the problem of (human) behavior was part and parcel of the nature of military command, now Wiener had this problem affecting the very design, construction and probability of success for his predicting machine. Wiener realized that the notion of *behavior* was at the crux of both his empirical and theoretical dilemma. Given the state of psychological knowledge regarding this issue, where “no behaviorist ha[d] ever really understood the possibilities of behavior”,<sup>29</sup> and considering the urgency inherent to war-related enterprises, Wiener exercised a philosophical extension that would later prove to be prescient.

It does not seem even remotely possible to eliminate the human element as far as it shows itself in enemy behavior. Therefore, in order to obtain as complete a mathematical treatment as possible of the over-all control problem, it is necessary to assimilate the different parts of the system to a single basis, either human or mechanical. Since our understanding of the mechanical elements of gun pointing appeared to us to be far ahead of our psychological understanding, we chose to try to find a mechanical analogue of the gun pointer and the airplane pilot.<sup>30</sup>

Given the state of darkness regarding the study of behavior in general and human behavior in particular, extending the realm of the mechanizable seemed like the logical course of action. Surely much needed to be understood regarding the new realm of control theory. But it struck Wiener as evident that such knowledge, even if limited, was still greater than knowledge about

---

<sup>28</sup> Edwards 1996, p. 206

<sup>29</sup> Galison 1994, p. 242

<sup>30</sup> Wiener 1956, pp. 251-252

man. Thus the source of irregularities (namely, the human element in the system) could be identified with a tractable aspect within the realm of machinery...

We realized that the "randomness" or irregularity of an airplane's path is introduced by the pilot; that in attempting to force his dynamic craft to execute a useful manoeuver, such as straight-line flight or 180 degree turn, the pilot behaves like a servo-mechanism, attempting to overcome the intrinsic lag due to the dynamics of his plane as a physical system, in response to a stimulus which increases in intensity with the degree to which he has failed to accomplish his task.<sup>31</sup>

Identifying that the pilot's behavior is comparable with (as in "is amenable to be subsumed under the study of") servomechanisms expanded the power of what could be mechanically tractable. From that point on, the usage of the term *behavior* (and its derivations) by Wiener and colleagues went well beyond the usual understanding of the notion as an exclusively Skinnerian concept –which also referred to a human being in machine terms, after equating it with a black box. However, as George Stibitz wrote in his report, Wiener's ideas, even if enriched, were still unmistakably anchored in clear behavioristic roots. He acutely pointed out that Wiener's "extension" of what behaves machine-like was based not upon the structure of the entity, but upon its behavior.<sup>32</sup>

W[iener] points out that their equipment is probably one of the closest mechanical approaches ever made to physiological behavior. Parenthetically, the Wiener predictor is based on good behavioristic ideas, since it tries to predict the future actions of an organism not by studying the structure of the organism but by studying the past behavior of the organism.<sup>33</sup>

---

<sup>31</sup> Galison 1994, p. 236

<sup>32</sup> This dichotomy will eventually be crucial for the enterprise that was about to emerge –cybernetics (more of this on chapter IV, where its philosophical backbone is articulated).

<sup>33</sup> Galison 1994, p. 243

Regardless of the project's outcome, Wiener saw a new and powerful paradigm beginning to take shape, capable of encompassing realms heretofore banned from a purely mechanical discourse. He did not hide his enthusiasm towards what was emerging before his eyes.

At this point, the machine was tested only in the laboratory and with data obtained *ad intra*. Let us recall that so far the data was produced by Bigelow's simulator in the carefully measured environment of the small MIT laboratory assigned to Wiener. Before the predictor could be deployed to a field test (like the one Wiener and Bigelow attended for testing the Bell predictor), they needed to input real-world data. For this, both men toured several proving ground facilities, getting all the information they could lay hands on regarding aircraft behavior. A particular piece of data was particularly useful. Flights 303 and 304 from the North Carolina proving ground registered telemetry that reported aircraft behavior every second. This was exactly what Wiener and Bigelow needed. Equipped with it, they crunched the data with the predictor for the next 5 months.

In practical terms, the results were not too encouraging. Wiener's predictor was compared with the mentioned simpler --and tested in front of him-- machine constructed under Bell's Hendrick Bode. Bode's predictor relied solely on a calculation of the target's location based upon input angles and distances, and a re-computation every 10 seconds to correct the aiming. It did not have statistical analysis, unlike Wiener's, but it was based upon readily available technology --and hence it was comparatively easier to construct and deploy. Wiener's machine with all its bells and whistles performed barely better than Bode's for flight 303 and even slightly worse for 304. The attack on Pearl Harbor had just occurred and imminent practical results were urgently needed: The United States had finally entered the global conflict. Wiener and Bigelow realized that hundreds of telemetry samples would need to be input into the

machine in order to make it practically functional, and the chances of having it finished before the end of the war were slim. They had to agree that Bode's machine was the one to be favored by the military, and thus abstained from recommending their own machine for construction.

Given his relentless patriotic spirit, Wiener was obviously disappointed in failing to provide tangible help by means of this project. However, he did substantially collaborate, in the larger scheme of things, to the cause of the war. Wiener was acknowledged --by both the engineers behind Bode's predictor and the scientists at the Rad Lab-- as responsible for improving the machine that the latter had already constructed. He would even go on to help "on the field" when this "rival" predictor would malfunction (e.g. aboard a naval vessel).<sup>34</sup> Furthermore, he produced by this time (1942) a manuscript summarizing the evolution of the theory of servomechanisms up to that point, improved with his insights on negative feedback, entitled "The Extrapolation, Interpolation and Smoothing of Stationary Time Series: With Engineering Applications".<sup>35</sup> The piece was immediately labelled as "classified" the moment Warren Weaver received it. The "Yellow Peril" --as it was nicknamed due to both its staggering complexity and the yellow cover proper to classified documents at the time-- became a legend of sorts among military engineering circles.<sup>36</sup>

Importantly for this thesis, the apparent failure by Wiener to accomplish the task of a new, substantially enhanced predictor yet opened up new routes of thinking. In the attempt at constructing this ambitious predictor, Wiener had had a profound insight --one that laid the foundations of a new science that would soon be regarded as revolutionary on a global scale.

---

<sup>34</sup> Conway & Siegelman 2005, p. 124

<sup>35</sup> Later declassified and published under the same title (Wiener 1949).

<sup>36</sup> Conway & Siegelman 2005, pp. 116ss. Wiener's contribution to the war effort and science in general was later clearly acknowledged by his government. He received the National Medal of Science from the hands of President Lyndon B. Johnson in 1964.

### **3. The pre-meetings and founding articles.**

Before the predictor project came to its conclusion, Wiener already had grand plans stemming out of his experience. Indeed it was reported that at the time he was working on the machine, his mind was seemingly already on something else --as reported by his direct project supervisor, Warren Weaver. One should also take into account Wiener's outspoken dislike for war-related secrecy --parameters which, to his understanding, run against the very notion of scientific practice. Wiener needed a freer environment to develop his new insights.

The Josiah Macy Jr. Foundation is a funding institution rooted in a wealthy family of merchants, interested in the interdisciplinary advancement of sciences --particularly medical sciences.<sup>37</sup> In 1942 it organized a gathering of 20 scientists to take place in New York, on May 13 and 14. Wiener seized the opportunity.

The conference was supposed to deal with issues of hypnosis and conditioned reflexes. In its spirit of interdisciplinarity, related approaches were welcome, and so Wiener's findings would allegedly fit. Wiener himself could not attend: the deadline for presenting his predictor was coming close. The person in charge of conveying his newly found ideas was a good choice. Arturo Rosenblueth, the medical scientist referred to above, who helped Wiener theoretically dissect the disconcerting endless loops, was to channel these ideas to the world for the first time --although to a small group, to begin with. This aspect was in fact the remarkable feature of the presentation, since much of the punch of Wiener's novelty had precisely to do with these circular loops disentangled with Rosenblueth's help. And indeed Rosenblueth did deliver. Although no

---

<sup>37</sup> It is still operating. See [macyfoundation.org](http://macyfoundation.org)



written records remain of this “Cerebral Inhibition Meeting”, personal recounts of the gathering described it as intense and powerful.<sup>38</sup>

Rosenblueth carefully conveyed what was possible without breaking the war-related oath of secrecy --e.g., he left aside much of the technicalities of Wiener’s “Yellow Peril”. He managed to deliver the core notions that the three men dealt with during work on the predictor. Rosenblueth reported that there was a novel notion of causality in nature that should be understood as “circular”. More specifically, given that some sub-parts of relatively complex systems sometimes seem to be caught in endless loops, there seem to be cases where overcompensating forces push the system “out of whack” (e.g., wild swinging of a person’s arm when she wants to take a bite of food, with tremor illness, missing her own mouth, or an anti-aircraft cannon wildly oscillating back and forth without apparent cause). However, there is a countervailing input, a “negative feedback”, that seems to correct this potential error, aligning the system back on track. A typical example would be that of a torpedo guiding itself, making the necessary corrections, so that it keeps aiming at its target (a ship or submarine) by following the noise of its machinery or the magnetic force exerted by its hull. Both the homeostatic self-regulation of body temperature in animals, and the phototropic behavior of a plant, are no different in substance. Rosenblueth thus stated that approaching these phenomena without understanding these occurrences within the framework of a “circular causality” would miss the mark for appropriately understanding the array of causes at play. This remark raised a few eyebrows.

Some of the attendees were the neurophysiologists Warren McCulloch and Rafael Lorente de Nó, the neurologist and psychoanalyst Lawrence Kubie, and the anthropologists

---

<sup>38</sup> The invited anthropologist Margaret Mead reportedly broke her tooth in this meeting, and she did not realize it until it was over, due to the sheer excitement. See Heims 1991, pp.14-17.

Gregory Bateson and Margaret Mead –a married couple. What Rosenblueth was saying felt for some of them like an escape valve out of years of frustration in their research and the scholarly literature of their own fields. Lawrence Kubie turned from neurology to psychoanalysis precisely due to what he would regard as intolerable reductions of the human psyche. The Spanish-American neuroscientist Lorente de Nó had found evidence for continuous loops in neural activity some years before –but no paradigm in which he could make sense of them. Bateson and Mead decried the lack of a sufficiently rich model to understand cultural exchanges and developments. In the case of Warren McCulloch, it could even be said that he found light after years of darkness. McCulloch, as Lorente de Nó, found that neural activity regularly engaged in these aforementioned endless loops --without being able to make heads and tails of them either. With one aggravating detail. For decades McCulloch was trying to build up a model that would correlate the all-or-nothing character of neuronal firing with the all-or-nothing character of logical processes --trying to establish a hylomorphism between a neuron and a logical gateway. These “loops” were a major obstacle against this otherwise sophisticated mapping. And now there was a glimmer of light at the end of the tunnel.<sup>39</sup>

The organizer, Frank Fremont-Smith, happy to see this positive response, asked the foundation for a follow up –probably as a regular recurring gathering. It turns out that Warren McCulloch was already lobbying for it. Nevertheless, the pressures of the war, in terms of both contracted commitments and imposed secrecy, made everyone agree that it should be postponed until it was over. However, the wheels for something much greater in scope at many levels were already in motion. The next year, Wiener, Rosenblueth and Bigelow published an iconic paper in the journal *Philosophy of Science*, largely based upon the presentation delivered by Rosenblueth,

---

<sup>39</sup> I will come back to this later in this chapter.

entitled “Behavior, Purpose and Teleology”.<sup>40</sup> There Wiener and company went the extra mile into controversy. After reconfirming what was said in the meeting, they advocated for the need to revive the notion of a final cause, which allegedly is scientifically perceived in nature’s doings. Indeed, they proposed the appropriation of the ages old notion of “teleology” in order to understand this reality scientifically. The proposal, available this time to a wider audience, raised many more eyebrows, but now the effects were polarizing.

In Aristotle’s theory of causality, of the four causes that explain the being of things (material, formal, efficient and final), the final cause (*telos*) was the most important for understanding things in the universe. His whole ethics stems from the notion of *telos* applied to the thing “human” (*eudaimonia*). This *telos*, after having received a profound theological connotation during the Middle Ages, was later rejected by the spirit of the Scientific Revolution –relying now on Bacon’s *experimentum crucis* as the sole filter and judge for an advanced hypothesis. Later on, philosophy (in its *logical positivist* version) took upon itself the furthering of this expulsion of anything non-empirically evidential from scientific discourse. This development rendered the very idea of metaphysics inimical to science –a prejudice to get rid of. Ultimately, behaviorism (a dominant trend in psychology at Wiener’s time, whose reductionist dicta he was very aware of)<sup>41</sup> bracketed the mind as a metaphysical entity itself, focusing instead on the inputs and outputs to a “black box”. In other words, the very mention of teleology was grave heresy for the scientific and philosophical orthodoxy of the time –even if, in this case, what was labelled a ‘final cause’ made no use of any metaphysical ‘force’, merely referring to

---

<sup>40</sup> Rosenblueth, Wiener & Bigelow 1943

<sup>41</sup> See Galison 1994, p. 243 (n). Wiener’s disappointment with the poverty of “behaviorism” as a viewpoint in psychology is indicated above. Although the word “behavior” --and its derivatives-- will be utilized often, it will no longer refer to the traditional psychological position. Instead, it will make a much broader referencing to the mode of being or process of coping of an entity –organic or artificial. More on this in the chapter concerning Ross Ashby’s work.

the goal to which the whole system aims, making use of corrective negative feedback. Still, Wiener went ahead and pushed for the use of the term, stating that “teleological” can perfectly be predicated of the behavior of a self-guided missile. As we will see later, while some saw in this revival attempt the sign of a science of the future, some others did not take it positively.

Once McCulloch went back to his home institution (the University of Illinois), he re-engaged into his experiments with renewed impetus, equipped with the providential insights from Wiener as faithfully conveyed by Rosenblueth. Norbert Wiener had met the psychiatrist McCulloch several years before, introduced by Rosenblueth at one of his Harvard evening meetings. McCulloch was already occupied with theorizing artificial neurons as early as 1927. The main idea, which had been haunting him for years, pivoted around the possibility of understanding the nervous system as a physical instantiation of a certain realm hitherto acknowledged as *a priori*, clear and untarnished: Logic. This pure realm stood in sharp contradistinction with the chaotic messiness of the human mind. However both realms are somehow uneasily embedded in the brain. Thus, the enigma arose: How is it possible that a fallible and unpredictable entity, such as the human mind, could possibly grasp the exactitudes of mathematics? Where is the link that connects and fixes the “messy” with the “exact”? These questions were already present in McCulloch’s early years as a college student, right in the midst of deciding whether to enter the theological seminary for becoming an ordained minister – faithful to his family’s wishes.<sup>42</sup> His spiritual mentor at the time asked him about what he would like to do with his life. McCulloch told him that he would like to know “What is a number, that a

---

<sup>42</sup> Puritan Protestantism ran deep in Warren McCulloch’s ancestry. His father remarried a substantially younger southern lady who was deeply religious. Eventually, Warren’s sister carried the religious family torch, becoming a committed Bible preacher. In fact, his ways strayed away from Puritanical Christianity in a manner that would ultimately carry consequences for his scientific endeavors, as we will see below when addressing the rift with Wiener (Heims 1991, pp. 31-32)

man may know it, and man, that he may know a number?”<sup>43</sup> He recounts his intellectual pilgrimage as follows...

I came, from a major interest in philosophy and mathematics, into psychology with the problem of how a thing like mathematics could ever arise-what sort of a thing it was. For that reason, I gradually shifted into psychology and thence, for the reason that I again and again failed to find the significant variables, I was forced into neurophysiology. The attempt to construct a theory in a field like this, so that it can be put to any verification, is tough.<sup>44</sup>

But the mathematical fleshing out of his developing insights was persistently evading him.

This wait came to an end, however, after McCulloch met someone whose genius might have been one of the most underreported biographical stories in 20<sup>th</sup> century history of science.

Jerome Lettvin, a pre-med student who caught notice of the powerful medical experimental facility recently assigned to McCulloch, introduced him to the high school dropout Walter Pitts (1923-1969). Lettvin recounted that Pitts was once escaping from a group of bullies and had to hide in a library. While in his “refuge”, Pitts stumbled upon a particular book, and once he began to read it, he felt drawn towards it. It was Bertrand Russell and Alfred North Whitehead’s three volumes of the *Principia Mathematica*.<sup>45</sup> Pitts spent the subsequent week devouring the first tome, and found what he thought was an error in the work triggering a letter that he sent to Russell. The British philosopher, impressed upon reading the missive, invited him to come to the UK to do graduate work by his side. Pitts was at the time 13 years old, and had no way of

---

<sup>43</sup> Warren McCulloch recalls the emergence of this question, which allegedly haunted him for the rest of his life, in the following manner:

In the fall of 1917, I entered Haverford College with two strings to my bow – facility in Latin and a sure foundation in mathematics. I “honored” in the latter and was seduced by it. That winter [the Quaker philosopher] Rufus Jones called me in. “Warren,” said he, “what is thee going to be?” And I said, “I don’t know.” “And what is thee going to do?” And again I said, “I have no idea; but there is one question I would like to answer: What is a number, that a man may know it, and a man, that he may know a number?” He smiled and said, “Friend, thee will be busy as long as thee lives.” (McCulloch 1960, p. 1)

<sup>44</sup> Von Neumann 1948, p. 32

<sup>45</sup> Whitehead & Russell 1927

accepting such an offer. Two years later, Pitts --who was routinely beaten up by his own father, a plumber-- finally escaped from his home in Detroit and headed to Chicago, in the hope of finding some intellectual comprehension and acceptance. At the time Russell was visiting professor at the University of Chicago, and upon reconnecting with Pitts, he suggested the latter to attend the lectures of the logician Rudolph Carnap --which he did. Again, Pitts rather quickly pointed out to the logician what he thought were inconsistencies in a certain piece of his work -- to the surprise of Carnap. He was 15 years old at this time, and the university took notice of the boy's powerful mind, routinely seen wandering around campus and sitting in on advanced graduate classes. Sensing the potential for an unusually bright future in academia, the university arranged for a modest amount of regular income for Pitts to literally survive --he was almost always homeless.<sup>46</sup>

When Lettvin introduced Pitts to McCulloch, the latter was profoundly impressed by his genius and soon practically adopted him, bringing both Pitts and Lettvin to live in his house. Not long after Walter Pitts had ironed out with passion the mathematical wrinkles that McCulloch's model was suffering from. The outcome of this conjoint work was the 1943 seminal paper "A Logical Calculus of Ideas Immanent in Nervous Activity".<sup>47</sup> The proposal contained there pivoted around the notion of very simplified and idealized "neurons" heavily interconnected with each other, in such a way that after a massive conjoint operation --following the rules of Boolean logic-- some behavioral outcomes in an organism could be accounted for. In other words, the mechanism taking place in animal cognition was subsumed under the realm of logical operations. The mind was underpinned by a real logical structure after all. It would seem that McCulloch was finally answering the haunting question of his youth. However, the paper failed to garner

---

<sup>46</sup> Conway & Siegelman 2005, p. 115

<sup>47</sup> McCulloch & Pitts 1943

attention from its main target: psychologists and neurologists. In turn, an unlikely group of people took a keen interest in it: the incipiently formed community of computer scientists. As we will see, this “productive misappropriation” would entail interesting consequences for the scientific enterprise that was about to be formed.

The mathematician John von Neumann (1903–1957), a Jewish Hungarian émigré who escaped to America fleeing the rising Nazi terror, was well acquainted with Wiener’s work since at least a decade earlier. He was also a consultant for the Aberdeen Proving Ground where Wiener was working. Although Wiener’s proposal for a computing machine delivered to Vannevar Bush was shelved at the time,<sup>48</sup> it propelled Wiener to be a consultant in computation for both the American Mathematical Society and the Mathematical Association of America.<sup>49</sup> Indeed the paths of both men were seemingly bound to intersect. In 1943, a key year for all these developments,<sup>50</sup> von Neumann joined the ultra-secret Manhattan Project, suddenly finding himself in need of a powerful computing machine that would crunch numbers with a complexity far beyond anti-aircraft firing tables. Mainly for this reason, he went to England to be acquainted with the work that the mathematician Alan Turing (1912–1954) was doing regarding computing machinery. When he came back, von Neumann confessed to being already obsessed with the enticing possibilities of computing power. Although he had access to Vannevar Bush’s second analog computing machine, the 100-ton Rockefeller Differential Analyzer,<sup>51</sup> this machine was already overbooked for war-related tasks. Von Neumann thus became part of the consultant team for the improvement of Bush’s successor computer, the ENIAC. Von Neumann was set to improve dramatically its shortcomings. McCulloch and Pitts’ 1943 article was the model he used

---

<sup>48</sup> Mindell 2002, p. 275

<sup>49</sup> Conway & Siegelman 2005, p. 145

<sup>50</sup> Both foundational papers were published in 1943 --“Behavior, purpose and teleology” and “A logical calculus of the ideas immanent in nervous activity”.

<sup>51</sup> Bowles 1996, p. 8

for his “First Draft” for the construction of still a more powerful successor, the EDVAC --whose design philosophy has prevailed until today under the name of “von Neumann architecture”.<sup>52</sup>

In this context, Wiener called for an exclusive meeting to be held at Princeton, where von Neumann was professor. The invitation was sent out in 1944 to seven people, including McCulloch, Pitts, Lorente de Nó, and Rosenblueth (who in turn could not attend this time). The meeting was co-organized by von Neumann and Howard Aiken, a Harvard mathematician. In 1945, the small “Teleological Society” meeting took place, and it served as an opportunity for von Neumann to reveal on the one hand the material feasibility of a working mechanical model of the brain, and on the other the state of computer science in general –at the time still a novel and fairly exotic area of knowledge. Wiener furthered his still forming views on the new science of communication. McCulloch and Lorente de Nó spoke about their model of the brain that was serving as a model for von Neumann’s machine. The meeting was a success and it further secured their desire to continue getting together on a regular basis. Von Neumann maintained close contact with Wiener from 1945 on. But the real launchpad for all these novel scientific approaches was just about to happen the following year.

---

<sup>52</sup> The “von Neumann architecture”, these days more commonly known as “stored-program computer”, provided the advantage over the earlier computer machines in that the set of instructions could be stored as software in the machine –as opposed to physically changing hardware switches, as it was customary up until that time. The downside, however, is that since both instructions and data share the same bus (data communication device), performance could be compromised – effect known as the “von Neumann bottleneck”. There were other differences that set it apart from the machine that Turing built in the UK. Martin Davis described it like this:

Like the tape on that abstract device, the EDVAC would possess a storage capability--von Neumann called it "memory"-- holding both data and coded instructions. In the interest of practicality, the EDVAC was to have an arithmetic component that could perform each basic operation of arithmetic... in a single step, whereas in Turing's original conception these operations would need to be built up in terms of primitive operations such as "move one square to the left." Whereas the ENIAC had performed its arithmetic operations on numbers represented in terms of the ten decimal digits, the EDVAC was to enjoy the simplicity made possible by binary notation. The EDVAC was also to contain a component exercising logical control by bringing instructions to be executed one at a time from the memory into the arithmetic component. This way to organize a computer has come to be known as the von Neumann architecture, and today's computers are for the most part still organized according to this very same basic plan...(Davis 2000, p. 182)



#### 4. The Macy conferences.

The war finally ended in 1945 and the time seemed appropriate to reconvene after the 1942 meeting. Frank Fremont-Smith was still in charge of the Macy Foundation's arm to fund medical sciences, and he promptly agreed, after McCulloch and Wiener's persuasion, to launch a recurring series of conferences. These series, for which they did not yet have a name, were to occur twice a year. The group was now richer, having merged both Macy's 1942 first list of attendants and the 1945 Princeton gathering of scientists and mathematicians. The original idea was to discuss theory at the meeting itself, and then test it in each of the attendants' own realms of practice, coming back later with a report. Subsequently, the general theory could be improved based upon the feedback from these reports. After the fourth conference the frequency of meetings was cut in half to just once a year –it was deemed more appropriate for giving sufficient time for implementation. The gatherings spanned the time period between 1946 and 1953 and they did not produce an official transcript until the sixth meeting in 1949. During the first conferences a mechanical device for audio recordings was used, but the results were far from satisfactory.<sup>53</sup>

The first conference was entitled “Feedback Mechanisms and Circular Causal Systems in Biological and Social Systems” and it took place in the Beekman Hotel, in New York.<sup>54</sup> The first lecture was delivered on the morning of March 8 1946 by John von Neumann, joined by Lorente de Nó. They spoke about the purported correlations between the artificial networks of computing machinery and the mind's logical structures embedded in real neuronal wetware –realms just

---

<sup>53</sup> “We tried some sessions with and some without recording, but nothing was printable. The smoke, the noise, the smell of battle are not printable” (McCulloch 1974, p. 12). See also Hayles 1999, p. 81.

<sup>54</sup> All of the conferences took place in the same hotel, except for the last one, which was moved to Princeton, New Jersey, in order to accommodate von Neumann –who nevertheless ended up not attending.

linked in 1943 courtesy of the McCulloch and Pitts' paradigm for the calculus of neural activity. Von Neumann addressed the audience from the standpoint of the mathematician-turned-computer scientist, providing a panoramic landscape of the state of computer theory and construction --including his own. Lorente de Nó, as a neurophysiologist, correlated von Neumann's description of the logical processes in the metal with human physiology, pointing out possible instantiations where such outlook could fulfill an explanatory one --such as the yes/no character of neuronal firing. The proposal of an actual mechanism underlying mental processes, where a mechanical model of a realm that had forever escaped the grasp of science was being put forward, reportedly put the audience in a state of positive consternation. As with the 1942 conference, enthusiastic discussion ensued --and this was the format kept throughout the rest of the conferences to come.

In the afternoon, Norbert Wiener finally himself delivered what he could not do in person the year before at the science and mathematics gathering at Princeton. He gave a report on the new science that was being formed out of the realization of circular causality as a more complete tool for understanding goal-oriented systems --both natural and artificial. Wiener delivered a historical introduction of the evolution of mechanisms of control, from the time of Ancient Greece, going through Watson's mechanical governors and Maxwell's mathematical treatment of them, all the way to his "teleological machine" built (conjointly with Bigelow) for anti-aircraft weaponry. Rosenblueth was his lecture-partner. Faithful to Wiener's ideas as he was the previous year when delivering Wiener's keynote, Rosenblueth took care to provide examples of the feedback mechanisms underlying the very survival of living systems. By this time, Wiener was more aware of the utmost importance of the notion of "communication" for the overall outlook

of this emerging science –one that successfully encompassed both living organisms and artificial machinery:

[I]t had already become clear to Mr. Bigelow and myself that the problems of control engineering and of communication engineering were inseparable, and that they centered not around the technique of electrical engineering but around the much more fundamental notion of the message, whether this should be transmitted by electrical, mechanical, or nervous means... [T]he group of scientists about Dr. Rosenblueth and myself had already become aware of the essential unity of the set of problems centering about communication, control, and statistical mechanics, whether in the machine or in living tissue.<sup>55</sup>

The second day, Warren McCulloch explained the proposal put forward in his cryptic article co-written with Pitts. As mentioned, this article failed to gain the attention of the audience to which it was originally intended --neurologists and psychologists-- but in an ironic twist in the history of science, it served as the footprint upon which von Neumann designed a powerful computer machine –one whose fundamental structure remains pervasive to our day. The yes/no nature of neural firing, since it seemingly has a fairly direct correlation with the yes/no character of logical gateways, can provide a way to disentangle complex mental processes --as long as one is able to rigorously describe its smallest units in absolute detail. Von Neumann applied this thought to a physical contraption, giving it empirical instantiation, and his machine was born. McCulloch, on the other hand, was firmly on the side of finding an explanation for mental processes, and he thought that he found a paradigmatic one. The messiness of the human mind, however capable of

---

<sup>55</sup> Wiener 1948, pp. 8, 11. That was how Wiener recalled in his later book his earlier cybernetic experiences. The issue of communication would later become a contentious one, particularly regarding the issue of the possibility of transferring “meaning” over a mathematized articulation of communication –more amenable for exclusively syntactic conveyance. For more on this issue, see below in this chapter (8<sup>th</sup> Conference) and chapter IV, section 2 – including the notes.

dealing with the exactness of math and logic could now be demystified: it is all a matter of complexity, which can ultimately be parsed into primigenial pieces of a network, each of them tractable in a yes/no fashion. Propositional calculus could provide a tool to understand nervous activity.

In all likelihood, half of the first two lectures of the first cybernetic conference were delivered with a Spanish accent –by Arturo Rosenblueth and Lorente de Nó.<sup>56</sup> The rest of the day was dedicated to the social sciences.<sup>57</sup> Once the conference was over, scientists went back to their own realms of research, invigorated by these novel ideas that promised to revolutionize science itself, largely by means of accurately expanding the realm of physics to areas heretofore considered inherently outside its grasp –like life and mind (*pace* behaviorism, as we saw above).<sup>58</sup>

Wiener began to work with Rosenblueth on nervous activity in living flesh --commuting between Mexico City and Boston one semester every other year. He was trying to make sense of a problem he previously saw during his work involving gun turrets. As reported above, when the mounted gun received undamped positive feedback, it would engage in a wild swinging until

---

<sup>56</sup> If one adds Santiago Ramón y Cajal to the duet, one could perceive a lineage of Spanish neuroscientific influence at the beginning of cybernetics (Ramón y Cajal’s views on discrete neural structures inspired McCulloch’s). This cybernetic “Spanish connection” is largely untreated in the literature. For Ramón y Cajal’s influence on McCulloch, see Boden 2006, p. 188

<sup>57</sup> There are those who maintain that an important role of the Macy conferences was the spreading of cybernetic ideas to the wider realms of society and politics, particularly if one takes into consideration the insistence of Gregory Bateson at the first two Macy gatherings (1942 and 1946) for including the social sciences. Indeed the social scientists present in the conferences were not unhappy about the amount of media attention that the conferences began to garner. And there is also the famed case of project Cybersyn: The democratically elected president of Chile Salvador Allende hired the cybernetician Stafford Beer in order to run a socialist Chile –before General Augusto Pinochet gave the *coup de etat*. Be as it may, the present writing is not concerned with the social extensions of the cybernetic proposal, but rather, with its core theoretical tenets, profoundly intertwined with the nature of a machine. For an account inclusive of the social aspects of the conferences, see Heims 1991. For an account of the failed attempt to establish a cybernetic South American government, see Medina 2006 and Medina 2011.

<sup>58</sup> They might have even seen the “defense” of physics as one of the noble roles of the cybernetic mission. In fact, Dupuy wants to put more emphasis on the “extension of physics” aspect of the enterprise, somewhat against the notion of cybernetics as a “new science.” However, not only are there authors who disagree, but Wiener himself calls cybernetics a “new science” in non-vague terms. See Dupuy 2000, pp. 44-49; Conway & Siegelman 2005, p. 127; Wiener 1948, pp. 14, 28.

some corrective (negative) feedback was applied. If his intuitions were correct, the same phenomenon of wild overcompensation observed in AA machinery should take place in living organisms<sup>59</sup> –known in the latter case as *clonus*. Indeed, after Wiener began experimenting with living animals, when an extra electrical impulse would be applied to a main nerve in a cat's leg, its limb would engage in spastic convulsions. Wiener and Rosenblueth were looking for a pattern. But results were evading them –and Wiener was vocal about his frustration. Working with living tissue greatly elevated the complexity of the study. After months of perseverance, however, they arrived to a startling finding. The increase of nervous activity responsible for the animal's limb movement was not additive, but logarithmic in nature<sup>60</sup> –not dissimilar with what he encountered when augmenting positive feedback during his experience with anti-aircraft systems. In other words, the clonic spastic movement observed in the cat's femoral muscle kept an exponential (not linear) correlation with the amount of electricity input in the animal's leg. The effect of the electric input seemed to be multiplicative, rather than additive, as it was the case in the swing and oscillating phenomenon found in gun turrets.<sup>61</sup> The significance of this

---

<sup>59</sup> Wiener and Rosenblueth were already operating within a common framework for both animal and machine in no uncertain terms:

We observed this pattern of contraction, paying attention to the physiological condition of the cat, the load on the muscle, the frequency of oscillation, the base level of the oscillation, and its amplitude. These we tried to analyze as we should analyze a mechanical or electrical system exhibiting the same pattern of hunting. We employed, for example, the methods of MacColl's book on servomechanisms (Wiener 1948, pp. 19-20).

<sup>60</sup> In the words of Wiener:

...the notions of facilitation and inhibition are much more nearly multiplicative than additive in nature. For example, a complete inhibition means a multiplication by zero, and a partial inhibition means a multiplication by a small quantity. It is these notions of inhibition and facilitation which have been used in the discussion... Furthermore, the synapse is a coincidence-recorder, and the outgoing fiber is stimulated only if the number of incoming impulses in a small summation time exceeds a certain threshold. If this threshold is low enough in comparison with the full number of incoming synapses, the synaptic mechanism serves to multiply probabilities, and that it can be even an approximately linear link is possible only in a logarithmic system (Wiener 1948, p. 20).

<sup>61</sup> Speaking of the significance of this logarithmic basis for overcompensating movements in both animal and machine, Wiener remarked:

The most striking point is that on this logarithmic basis, and with data obtained from the conduction of single pulses through the various elements of the neuromuscular arc, we were able to obtain very fair approximations to the actual periods of clonic vibration, using the technique already developed by the servo

finding was going to be spelled out shortly after. It meant reaching the zenith of Wiener's ideas regarding teleological mechanisms, since it tied in with research he had been involved with decades ago. There certainly was stuff to report in the next upcoming meeting.

The second conference, entitled "Teleological Mechanisms and Circular Causal Systems" took place in October, 1946. Although it was mostly devoted to the possible social extensions of feedback mechanisms, it did have some science-related news. Walter Pitts, now an MIT student due to Wiener's influence, was to begin a PhD tailored to Pitts alone –under the supervision of Wiener himself. The aim of the doctoral degree was going to be the construction of a three dimensional model of neuronal structures --transcending the so-far two-plane representation of neurons for a model more similar to the actual object, hoping that the functions of the neuron would be better captured and subsequently understood. The degree of complexity sent waves of elation throughout the audience, especially among those well versed in mathematics, triggering McCulloch to declare that "the mathematics for the ordering of the [random] net has yet to be evolved".<sup>62</sup>

Right after this conference was over, another sub-conference started at the New York Academy of Sciences, and this one balanced things back to the mathematics and science side of things. Here Wiener presented to the attendees the conclusion to his evolution of ideas, tested under the fire of machine-building (the servo-mechanical anti-aircraft systems), and later, by his experiments on living tissue with Rosenblueth in Mexico.<sup>63</sup> Wiener had been noticing that there was a connection between what he was witnessing and thermo-dynamics. The "message" brought order to chaos –the default tendency in nature known as "entropy"—in both organic and

---

engineers for the determination of the frequencies of hunting oscillations in feedback systems which have broken down (Wiener 1948, p. 20).

<sup>62</sup> Conway & Siegelman 2005, p. 162

<sup>63</sup> Wiener 1948b

inorganic realms. It would seem that beyond the general notion of the message, there lay the more fundamental notion of *information*. Information indeed “in-formed”, giving shape to the realm in question, fighting off the primordial downward spiral of nature towards disorder and oblivion –the said entropy. Whenever there is information, entropy recedes –and vice versa. Entropy is the opposite of information, which in turn can be understood as negative entropy –or negentropy. Life is an informational force, negating entropy within the vastness of a merciless universe. Further, a working, “surviving” machine carries information that shapes its existence away from entropic decay.<sup>64</sup>

What Wiener was proposing was breathtaking, and the audience, by now reasonably adjusted to the excitement of these meetings, was once again in awe. McCulloch was sure to send copies of this talk’s notes to several prominent scientists and mathematicians who were not present –among them, to the American mathematician Claude Shannon (1916–2001). Several years earlier, Shannon, a young MIT graduate and recently hired star employee at Bell Labs, was closely acquainted with Wiener and the war-related work of the latter –in a master-disciple kind of way. Picking up the theme shared here, but also as a culmination of long rationalizations regarding communication and information, he was to publish two years later (in 1948) the paper “A Mathematical Theory of Communication”,<sup>65</sup> which stands to this day as a seminal document for “communication theory” –a field cross-listed between information and mathematics. Shannon became a regular attendee from the 7<sup>th</sup> conference on. The rest of the conferences will be alluded to throughout this thesis, some more than others, according to their relevance to this work. However, an outline of them follows.<sup>66</sup>

---

<sup>64</sup> For some possible theoretical extensions of Wiener’s view on entropy, see Hayles 1999, pp. 100-108.

<sup>65</sup> Shannon 1948

<sup>66</sup> A reconstruction of the first five conferences, without written record, have been put together mainly via personal recollections of some of the attendees. There have been two main attempts to historically reconstruct the cycle of

3<sup>rd</sup> Conference: “Teleological Mechanisms and Circular Causal Systems”

(March 1947)

This conference was devoted mainly to psychology and psychiatry, somewhat to the displeasure of the members coming from the hard sciences. Von Neumann, having experienced a change of heart regarding the feasibility of using McCulloch and Pitts’ networks to map the mind,<sup>67</sup> began to lobby for having the German biophysicist Max Delbrück invited to the meetings as a core member.

4<sup>th</sup> Conference: “Circular Causal and Feedback Mechanisms in Biological and Social Systems”

(October 1947)

A further clash occurred between those members drawn to psychology and psychoanalysis, and those who were behind the more mathematical approach of McCulloch and Pitts’ networks --in regards to mental activity. In particular, long-winded discussions ensued after the proposed distinction of Gestalt “forms”, considered to be analogical, and McCulloch and Pitts’ networks’ inadequacy to capture them, due to the claimed digitality of the latter.<sup>68</sup>

---

conferences. First, historically, by Steve Heims (Heims1991) and then thematically, by Jean-Pierre Dupuy (Dupuy 2000). There is also a summary that McCulloch wrote about the first three conferences in a letter written to von Foerster (McCulloch 1947).

<sup>67</sup> I will address this reversal in the chapter concerning von Neumann’s contribution to cybernetics.

<sup>68</sup> In fact, the question regarding the possibility of finding the Gestalt “forms” of perception (shapes, sizes, patterns, etc.) was posed right at the first conference in 1946 by the neurologist-turned-psychoanalyst Heinrich Kluver. Some of the attendees were amicable to the plea for help, like Warren McCulloch –perhaps in his role of Chair of all ten of the Macy conferences. But some attendees found the whole issue unrigorous and ultimately intractable as a scientific issue –like Walter Pitts, openly hostile right from the start. By this 4<sup>th</sup> conference, McCulloch also defined this position as squarely foreign to scientific discourse at best –and ultimately senseless at worst. By the 6<sup>th</sup> conference, McCulloch emerged very aggressively against the whole of psychoanalysis. By the 8<sup>th</sup> conference, Rosenblueth, having in mind McCulloch and Pitts’ networks, asserts that mental events either occur, or they do not – a feature that leaves the unconscious memories of psychoanalysis, according to him, outside the realm of facticity and coherence.



5<sup>th</sup> Conference: “Circular Causal and Feedback Mechanisms in Biological and Social Systems”

(Spring 1948)

Max Delbrück attended the meeting, to the joy of von Neumann --who, by this time, was seriously interested in genetic self-replication, and expected Delbrück to become part of the group core. Unlike any other meeting until the end, the entire first day was devoted to language. This bad timing might have played a role in Delbrück’s disappointment. The second day Wiener spoke about chaos and Pitts about a formal modelling of chicken pecking. Delbrück never came back.<sup>69</sup>

The Hixon Symposium: “Cerebral Mechanisms in Behavior”<sup>70</sup>

(September 1948)

A separate conference outside the Macy cycles took place that same year in Pasadena, at the California Institute of Technology. This event was remarkable due to the fact that von Neumann now clearly attacked McCulloch and Pitts’ networks –both McCulloch and Lorente de Nó were present. Von Neumann provided an alternative course of action for the cybernetic grand-enterprise: to shift focus towards the investigation of the possibility of constructing artificial cellular automata.<sup>71</sup>

---

<sup>69</sup> Heims 1991, p. 95

<sup>70</sup> Published in 1951 under the same title (Jeffress 1951).

<sup>71</sup> This is the reason why the contemporary field of Artificial Life regards it as its birthplace. See Langton 1989.

6<sup>th</sup> Conference: "Cybernetics: Circular Causal and Feedback Mechanisms in Biological and Social Systems"<sup>72</sup> (March 1949)

Von Neumann, who could not attend the conference, sent a message to the attendees, which was read out loud. The audience took this missive seriously, and a further discussion regarding the inadequacy of McCulloch and Pitts' networks was triggered. This time, the issue was the insufficient capability of accounting for human behavior if relying on arrays of neurons alone –there were not enough of them. McCulloch defended its own model and Pitts eventually spotted an error in von Neumann's calculation regarding the quantity of neurons in a human brain, bringing the discussion to a halt.

The Austrian physicist Heinz von Foerster (1911–2002) was named editor in chief of the Macy conferences –partly in order to help improve his rudimentary English. He almost immediately proposed settling for one name for the conferences, *Cybernetics*, after Wiener's book, which was published the previous year. The acceptance of the proposal was unanimous. From that year on that title remained, and the previous mouthful was relegated to a subtitle.

7<sup>th</sup> Conference: "Cybernetics: Circular Causal and Feedback Mechanisms in Biological and Social Systems"<sup>73</sup> (March 1950)

Once again, lively discussion emerged regarding whether or not an analog model of the mind would be more suitable than a digital one --e.g. the McCulloch and Pitts' model. No side seemed to get the upper hand. Also, a renewed attack against psychoanalysis as a coherent paradigm ensued, this time not only coming from the hard sciences (e.g., Pitts) but from the social sciences as well (e.g., Bateson).

---

<sup>72</sup> Von Foerster, Mead & Teuber 1950

<sup>73</sup> Von Foerster, Mead & Teuber 1951

Concern was voiced regarding unwarranted and overoptimistic accounts given by the popular media (e.g., *Scientific American*, *Popular Science*, etc.) when referring to the conferences as a “cohesive group” securely marching towards the mechanization of the human being. The environment of discussion, although exciting and lively, was far from embodying a “group mentality”. McCulloch recalled an occasion with Wiener asking, while delivering a presentation, “May I finish my sentence?” to which an angry interlocutor replied “Don’t stop me when I’m interrupting!”<sup>74</sup> In that vein, a word of advice was thrown to those present coming from the social sciences: Not to fall for the allure of the exactness of mathematics and the hard sciences, which seemed to begin enjoying uncritical acceptance in their fields.

8<sup>th</sup> Conference: "Cybernetics: Circular Causal and Feedback Mechanisms in Biological and Social Systems"<sup>75</sup> (March 1951)

Both Wiener and von Neumann had by this time withdrawn from the conferences –for different reasons.<sup>76</sup> McCulloch remained the Chair, Pitts and Rosenblueth persevered and new people joined (e.g., Shannon). But the pair’s absence was noticeable –particularly Wiener’s. The British physicist and Christian apologist Donald Mackay (1922-1987) criticized Shannon’s views --well known by now. Shannon’s position, where the message had to be stripped from any aspect of “meaning” in order to account for an accurate and engineeringly feasible conveyance,<sup>77</sup> would have, for Mackay, disastrous consequences for the proper understanding of the intended

---

<sup>74</sup> McCulloch gave further detail of the kind of milieu where the discussions were taking place:

The first five meetings were intolerable... The smoke, the noise, the smell of battle were not printable... Nothing that I have ever lived through... has ever been like... those meetings... You never have heard adult human beings, of such academic stature, use such language to attack each other. I have seen member after member depart in tears, and one never returned (McCulloch 1974, p.12).

<sup>75</sup> Von Foerster, Mead & Teuber 1952

<sup>76</sup> These reasons will be addressed in chapter II, sections 3 and 4.

<sup>77</sup> Boden 2006, p. 204. More on this in chapter IV, section 2, where the philosophical underpinnings of cybernetics will be addressed.

message.<sup>78</sup> Donald Mackay also entertained the possibility of constructing a learning machine with randomness factors embedded within, in order to account for a more human behavioral outcome. The American statistician Leonard Savage vigorously objected to it, claiming that adding randomness would help in no way to mimic human behavioral processes. Savage gave a presentation of his own, where he proposed a calculation procedure for decision-making based upon statistical factors. This time McCulloch objected to Savage, advancing the idea that the multilevel realm of decision making is inherently contextual, and thus escapes the “linearity” of statistical analysis.

The American experimental psychologist Herbert Birch was aware of both Wiener’s treatment of communication as homogeneously pervading the realms of animal and machine, and of Shannon’s removal of semantics as essential for the rigorous treatment of communication. In face of this, he attempted to introduce a quality distinction. Coming from the area of research in animal communication itself, Birch differentiated the communication found in animals from the “real communication” found in humans. According to Birch, we speak of animal communication using a relaxed version of the notion of communication. The cyberneticists<sup>79</sup> reacted very negatively against this observation, alluding to the mentalisms and unwarranted metaphysical

---

<sup>78</sup> To be sure, Donald Mackay was not the only one concerned about this exorcization of semantics from a functional theory of communication:

Soviet critics charged that Shannon’s theory of communication reduced the human being to a “talking machine” and equated human speech with “just a ‘flow’ of purely conditional, symbolic ‘information,’ which does not differ in principle from digital data fed into a calculating machine.” Wiener’s formula, “information is information, not matter or energy,” provoked a philosophical critique of the concept of information as a non-material entity. Repeating Lenin’s criticism of some philosophical interpretations of relativity physics in the early twentieth century, Soviet authors castigated cyberneticians for replacing material processes with “pure” mathematical formulae and equations, in which “matter itself disappears.” Cybernetics was labeled a “pseudo-science produced by science reactionaries and philosophizing ignoramuses, prisoners of idealism and metaphysics.” (Mindell 2003, p. 82).

<sup>79</sup> Both “cyberneticist” and “cybernetician” are still used to refer to those who were part of the cybernetic project. Henceforth both words will be used interchangeably.

tenets underpinning such distinction. This occurrence is remarkable in its irony, given the event that was about to occur with Ross Ashby's presentation at the next conference.<sup>80</sup>

9<sup>th</sup> Conference: "Cybernetics: Circular Causal and Feedback Mechanisms in Biological and Social Systems"<sup>81</sup> (March 1952)

This was the only conference attended by the British medical doctor and psychiatrist William Ross Ashby (1903 –1972) from the Ratio Club.<sup>82</sup> Ashby presented two papers where he proposed the change of behavior via random trials as an appropriate articulation of the phenomenon of learning in natural or artificial systems. The cyberneticists objected to this, in a fashion not in tune with the cybernetic mandate –to some extent borrowing Herbert Birch's context for argumentation alluded to above.<sup>83</sup>

Warren McCulloch voiced complaints about the governmental constraints gradually imposed upon several of the conference members. Indeed the Cold War had already started and an environment of increasing secrecy let itself be felt. These remarks might have also served to deflect attention from the previous withdrawal of the core members (Wiener and von Neumann).

10<sup>th</sup> Conference: "Cybernetics: Circular Causal and Feedback Mechanisms in Biological and Social Systems"<sup>84</sup> (April 1953)

McCulloch acknowledged the decade of criticisms against his (and Pitts') model – particularly coming from von Neumann, who nevertheless built upon it arguably the most

---

<sup>80</sup> More on this in chapter V, where the contribution of William Ross Ashby is addressed.

<sup>81</sup> Von Foerster, Mead & Teuber 1953

<sup>82</sup> More about the Ratio Club in chapter II, section 2.

<sup>83</sup> More attention will be given to this theoretical reversal towards the end of the chapter dealing with Ashby (V).

<sup>84</sup> Von Foerster, Mead & Teuber 1955

advanced computer of its time. McCulloch conceded that it is honorable to stay within the fine tradition of scientific refutability.

The discussion, despite its good humor, lacked “content” according to the German psychologist Hans Teuber, the assistant editor. Since he was just starting his academic career, in his view he had much to lose if such registration of events were to become public. Teuber was to offer his resignation if the transactions were to be published.<sup>85</sup> He later agreed to publish only the documents, without the ensuing discussions.

This was the only meeting that did not take place in New York. They moved it to Princeton, New Jersey, in order to accommodate von Neumann –who ended up not attending.

But this is just half of the cybernetics’ story...

---

<sup>85</sup> Hayles 1999, p. 75

## II. Cybernetics: The book, the club and the decline.

### 1. Norbert Wiener's *Cybernetics*.

France might have been destined to be intertwined with the development of cybernetics right from its inception, as is shown by the fact that one of the first science historians who took notice of the movement was Pierre de Latil.<sup>86</sup> In that country both engineers and the intelligentsia took early notice of Wiener's ideas –while in America interest was more circumscribed to circles of engineering and the military, partly due to the ambience of secrecy owing to the Cold War.

In 1946, Wiener was invited to give a talk in Paris, at the *Collège de France*. After the talk, he went for a coffee with the Mexican-French publisher Enrique Freymann, and they hit it off –Wiener later spoke of him as one of the most interesting individuals he had ever met. Freymann asked Wiener if he would like to publish his ideas in a book. Wiener, surprised, answered that he could not think of anyone who would be interested in publishing them. After it was clear that Freymann was referring to himself, Wiener promised to submit a manuscript in three months. He subsequently went back to Boston, and then flew to Mexico City, to spend the period corresponding to his funded work with Rosenblueth. It is in this city that Wiener, now fluent in Spanish with a Latin accent, wrote the entirety of the book --in the early hours of the morning, before heading to his laboratory work for the rest of the day with his Mexican colleague.

Wiener was on a roll; he wrote relentlessly. It seemed that he had been accumulating all these new ideas for the last years, and they were now boiling and eager to come out. Freymann received the write up in the mail after ninety days --much to his surprise, since he actually did

---

<sup>86</sup> See De Latil 1956.

not take seriously Wiener's agreement to deliver a manuscript promptly. He rushed to print out a first draft to send it back to Wiener for revision. In the meantime, MIT caught wind of Wiener's publication project and tried to persuade the publisher to concede the copyrights. To no avail. After intense negotiation, both publishers agreed to publish the book conjointly in the US, but Freymann retained the rights for the French version (also in English) which was the one distributed internationally.<sup>87</sup>

In the book, after acknowledging with gratitude the influence and help received from Bush, McCulloch, Turing, von Neumann, Shannon and others, Wiener did his due diligence in order to find a name for the new science he was advancing. Given that he was very well acquainted with the Greek and Latin classics –a legacy from his strict father, a Russian philologist-- it did not take long for him to engage in a thorough survey of possible terms for this novel discipline. At first, since “communication” lay at the core of the theory, he thought that the word *messenger* should be appropriate: *angelos* in Greek. However, the semantic overtones that this noun had historically acquired, that of an “angel” (*messenger* of the “good news” – *evangelion*) would obviously carry a connotation that Wiener was not trying to convey. So he went on to explore the possibility of borrowing an image that had a long history of rich connotations for Greek thought, in fact used by Plato on more than one occasion: that of a “steersman” –*kubernetes*. The Ancient Greek κυβερνήτης, steersman or helmsman, is itself a construction formed out of two words: κυβερ (pilot) and ναυτος (ship). Plato used the term κυβερνητική in the *Georgias*, the *Laws* and *the Republic*, in order to refer to the “art of navigation” (or proper steering) of a community. Plato was referring to the political art of governance. Unbeknownst to Wiener at the time, the French physicist André-Marie Ampère did

---

<sup>87</sup> Conway & Siegelman 2005, p. 175



use “la cybernétique” in 1834, referring also to the art of governance, but expanded to include elements of religion.<sup>88</sup>

Later on, the Latin derivation of the Greek word, *gubernator*, was acknowledged as the root of the English word *governor*, a device constructed by James Watt in the 1800s to be attached to steam engines for speed regulation –and later mathematically articulated. The British “natural philosopher” James Clerk Maxwell wrote in 1868 a mathematical treatise on these apparatuses –“On Governors”.<sup>89</sup> This mechanical device is considered to have played an important role at the very core of the Industrial Revolution. Even if Maxwell did not use the term “cybernetics” per se, his 1867 treatise showed some prescient elements of control theory and was thus heralded by Wiener as the founding document for cybernetic thinking.<sup>90</sup> The mathematization of physical phenomena showing corrective negative feedback indeed lay at the core of cybernetics. Leibnitz, due to his idea of a universal language of calculus, was proclaimed by Wiener as cybernetics’ patron saint.<sup>91</sup> Such illustrious lineage was reason enough for Wiener to coin *cybernetics* as the word that would name from now on his newly unveiled science.

---

<sup>88</sup> This is how Ampère defined “la cybernétique”:

Sans cesse il a à choisir entre diverses mesures celle qui est la plus propre à atteindre le but; et ce n'est que par l'étude approfondie et comparée des divers élémens que lui fournit, pour ce choix, la connaissance de tout ce qui est relatif à la nation qu'il régit, à son caractère, ses mœurs, ses opinions, son histoire, sa religion, ses moyens d'existence et de prospérité, son organisation et ses lois, qu'il peut se faire des règles générales de conduit, qui le guident dans chaque cas particulier. Ce n'est donc qu'après toutes les sciences qui s'occupent de ces divers objets qu'on doit placer celle dont il est ici question et que je nomme *Cybernétique*, du mot κυβερνητική, qui, pris d'abord, dans une acception restreinte, pour l'art de gouverner un vaisseau, reçut de l'usage, chez les Grecs même, la signification, tout autrement étendue, de l'*art de gouverner en général* (Ampère 1838, p. 141. Italics original).

<sup>89</sup> See Maxwell 1868

<sup>90</sup> Wiener had this to say about the importance of feedback for cybernetic ancestry:

We have decided to call the entire field of control and communication theory, whether in the machine or in the animal, by the name Cybernetics, which we form from the Greek κυβερνήτης or steersman. In choosing this term, we wish to recognize that the first significant paper on feedback mechanisms is an article on governors, which was published by Clerk Maxwell in 1868, and that governor is derived from a Latin corruption of κυβερνήτης. We also wish to refer to the fact that the steering engines of a ship are indeed one of the earliest and best-developed forms of feedback mechanisms (Wiener 1948, pp. 11-12).

<sup>91</sup> Wiener wrote the following about Leibnitz as the “spiritual father” of cybernetics:

If I were to choose a patron saint for cybernetics out of the history of science, I should have to choose Leibniz. The philosophy of Leibniz centers about two closely related concepts --that of a universal

It would seem fair to say that the uniqueness of Wiener's book relies on the fact that he managed to put under one cover many connected ideas regarding communication and information that were current on both sides of the Atlantic --mainly due to war-related research, where the United States officially invested. In fact, the British hub of cyberneticists, the *Ratio Club*,<sup>92</sup> entertained the implication that such was the reason for Wiener's in particular, and America's in general, lead in the new field. Wiener had come up with a perfectly timed book that not only gave these ideas wide circulation, but which served as a hub for the cybernetic insights already floating around. These ideas were in fact emerging in several parts of the world, including France, Sweden, the Soviet Union and the UK. War-related research was the main motivation for all of them. The British and the rest, without official support, missed that opportunity. Wiener, who had formal government back up, seized it.

One could safely assume, however, that Wiener himself had the main cybernetic ideas boiling and looking for a release valve. For starters, he already had three written pieces amounting to a solid cybernetic core: The top-secret "Yellow Peril" of 1942,<sup>93</sup> his co-authored 1943 article on teleological mechanisms,<sup>94</sup> and the 1946 lecture at the New York Academy of Sciences.<sup>95</sup> Wiener also started with the acknowledgment of the need for a common vocabulary for the emerging discipline, after witnessing the slippery slope of borrowing terms from one field

---

symbolism and that of a calculus of reasoning. From these are descended the mathematical notation and the symbolic logic of the present day. Now, just as the calculus of arithmetic lends itself to a mechanization progressing through the abacus and the desk computing machine to the ultra-rapid computing machines of the present day, so the *calculus ratiocinator* of Leibniz contains the germs of the *machina ratiocinatrix*, the reasoning machine. Indeed, Leibniz himself, like his predecessor Pascal, was interested in the construction of computing machines in the metal. It is therefore not in the least surprising that the same intellectual impulse which has led to the development of mathematical logic has at the same time led to the ideal or actual mechanization of processes of thought. (Wiener 1948, p. 12. Italics original.)

<sup>92</sup> More on this group of scientists below.

<sup>93</sup> See Wiener 1949

<sup>94</sup> See Rosenblueth 1943

<sup>95</sup> See Wiener 1948b

to another, hence imposing one worldview onto the other and missing the substance of the novel approach.

Wiener's newly found insights, regarding the poverty of the notion of behaviorism after his experience with the predictor, were spelled out. The utmost importance of negative feedback, after Rosenblueth's input from physiology, was carefully articulated. The existence of a common process aimed at survival, common to both animal (homeostasis) and machine (self-regulation) was established: Both were teleological mechanisms. Wiener's startling realization of the defining importance of "information", as the carrier of order, into and against the natural tendency towards chaos in reality—entropy-- was given due treatment. Beyond control theory and communication, information (or "negative entropy") was the way nature regulated itself, in the organic and inorganic, as already hinted in the discipline of physics by the second law of thermodynamics—an area with which he was intimately acquainted, from decades ago, due to his studies of chaotic behavior. Reality's inherent tendency towards decay is counterbalanced by information, "life" being a paradigmatic example—but also, a machine, as an island of will surviving in the midst of, and against, chaos.

Wiener provided cases which, linked via mathematical treatment, convincingly showed that the notions referred to above equally applied to both the living and non-living—and that a common methodology of study and experimentation was now emerging.<sup>96</sup> For example, he linked the case of a machine breaking down due to an unrestrained overflow of positive feedback with that of a brain overwhelmed with information when no more network storage is available. In both cases the lack of a regulative negative feedback would result in disastrous consequences—in the former one due to a lack of balance that breaks the system (probably a perennial loop), in the latter due to a circular (vicious) reasoning that results in neurosis, and eventually, insanity.

---

<sup>96</sup> Johnston 2008, p. 29

Wiener also ruminated about the social consequences of the new science, and the consequent moral responsibility for the cybernetician. Up to that moment he was keeping those thoughts to himself. Two events, one past and one on-going, may have triggered his cautious view. Firstly, the dropping of two nuclear bombs that killed more than 150,000 people –mostly civilians. Secondly, the growing ambience of secrecy that was effectively crippling the otherwise free flow of communication between scientists –part and parcel of the very essence of scientific practice, in Wiener’s view. Expanding on the foreseeable consequences of a dramatically augmented capacity for pervasive control, he predicted a social and economic change that could be regarded as disturbingly prescient by some. These days, Wiener saw an ugly head gradually rearing out of the newly found capacity of communication engineering that does not distinguish between man and machine.

“Information is information, not matter or energy”,<sup>97</sup> Wiener would say. Leave alone just the “content” of a message. The delivery of order would be, quite literally, the most powerful force in the universe: the very impetus that fights against reality’s own downward spiral into oblivion –the latter having been reason for hopelessness in more than one scientific mind. In a “cybernetically organized mankind”<sup>98</sup> the manipulation of information, now understood and appropriated, would have the pervasive capacity of profoundly and substantially modifying the physical and the non-physical, bringing the capacity of reality-control and modification to a level never seen in history. If the *industrial* revolution downgraded the “human arm by the

---

<sup>97</sup> Wiener 1948, p. 132.

<sup>98</sup> This is Martin Heidegger’s term. The German philosopher was closely following the developments of cybernetics. For him, it embodied the culmination of the unfolding of the history of metaphysics –or the essence of technology. More on this in the next chapter.

competition of machinery”,<sup>99</sup> the *cybernetic* revolution “is similarly bound to devalue the human brain, at least in its simpler and more routine decisions”.<sup>100</sup> And that would be just the beginning.

The scientist or engineer vaguely familiar with cybernetic ideas was certainly expected to have a “field trip” of enjoyment in reading the piece. For one, engineers finally made the connection, only perceived before by vague intuition, between what they were doing at the level of electronic signals transmitted via networks (e.g., telephone), and the nature of servomechanisms at large. There was no articulated theory of feedback that would equally apply to noise channels (in order to guarantee the accurate conveyance of a message) and to industrial machinery (in need of self-regulating devices for smooth behavior). Leave aside the startling demonstration that such regulating circular causality had always been present in nature itself – e.g., homeostasis. Several areas of science took notice at roughly the same time, while various engineering departments rushed to formalize the adoption of these views –which were not entirely new, but for the first time articulated with precision.<sup>101</sup>

Wiener’s book was riddled, after the passionately written “Introduction”, with complex equations --which in the first edition turned out to contain some mistakes.<sup>102</sup> However, it also contained powerful statements, sometimes slightly aphoristic, which revealed a mind that had been occupied for years in the developing of these notions. Some reviewers dismissed it, partly due to its complex mathematics and partly due to the sometimes non-linear way of argumentation. But what happened not long after its publication by far exceeded everyone’s

---

<sup>99</sup> Wiener 1948, p.27

<sup>100</sup> Ibid.

<sup>101</sup> Conway & Siegelman 2005, p. 184

<sup>102</sup> Wiener blamed Pitts for having misplaced the corrected manuscript for a while. This was a source of uncomfortable tension between them. Once the document reappeared, it turns out that someone mistakenly swapped the corrected version with the original one, sending the unchecked one back to France. This is the one that was printed and published --carrying some errors (Conway 2005, pp. 203-204).

expectations –not only the author’s and the publisher’s, but also the “cybernetic group”<sup>103</sup> of Macy at large. *Cybernetics: Or Control and Communication in the Animal and the Machine* became a best-seller on both sides of the Atlantic. The readership exceeded far beyond those human groups circumscribed to mathematics, engineering or hard sciences. The social and political scientist, the philosopher and the humanities scholar, the middle-class intelligentsia, the entrepreneur and businessman, the typical university educated layman at large... all began to read *Cybernetics*.

Finally this legendary group, led by this mysterious genius-mathematician, was producing a manifesto, relatively accessible to a wider audience –the Introduction to the book is an enthusiastic narrative that can give a very good idea to any reasonably intelligent reader of what the project was all about. The popular media *ipso facto* noticed the shockwave and global buzz, and Wiener soon was the subject of articles in such magazines as *Scientific American*, *Newsweek* and *Time*. Later this wave was joined by *Business Week*, *The New York Times* and *The Times Book Review*. Later on, Wiener was portrayed in *Life* and *The New Yorker*, and *Time* had a cover story dedicated to him. Suddenly he was being invited by the most prestigious universities and societies for keynote addresses –e.g., the 1950 International Congress of Mathematicians, at Harvard University.

Cybernetics was having a palpable impact in science, philosophy and beyond on a global scale –to the point of extending its influence to that part of the world that was, due to the Cold War, regarded as being outside Western reach: The Soviet Union.<sup>104</sup> Exchanges between cyberneticians coming from both the political, military and economic rival sides of the globe

---

<sup>103</sup> This is Steve Heims’ coinage. See the title of Heims 1991.

<sup>104</sup> Norbert Wiener engaged in active professional exchange with his Soviet colleagues when it came to cybernetic research, insights and conferences. In the Soviet Union, cybernetics enjoyed a form of *placet* from the ruling Communist party, probably due to its aura of “control.” These rapprochements were expectably not celebrated by governments of either side of the “Iron Curtain.” (Conway & Siegelman 2005, ch. 16)

were relatively fruitful, indeed to the point of raising some eyebrows among surveillance authorities on both sides. Interestingly (albeit perhaps expectably, due to socialism's canonical attraction towards all aspects of control), the Soviet cyberneticians counted on the tacit but firm support of the state,<sup>105</sup> in sharp contrast with their British counterparts, who had to rely mostly on the practitioner's own free time and economic means.

Within the West, cybernetics' reach went well beyond the English speaking world, overcoming the centuries-old multi-level historic animosity between Continental Europe and its Anglo-American occidental partner. Indeed as an enterprise whose core practitioners exclusively used English as their language of communication and practice, cybernetics found enthusiastic reception in France (e.g. Lacan<sup>106</sup>) and in Germany (e.g. Heidegger<sup>107</sup>). As mentioned, the first international edition of *Cybernetics*, to the chagrin of Wiener's own home institution printing press (MIT Press), was published by a French publishing house –following up the conferences in France to which he was invited.

Furthermore, it was one of the few scientific movements in modern history that managed to garner widespread fame –indeed well beyond scientific and academic circles. Mainstream media was very receptive to cybernetic developments, arguably due to the bold proposals (for the time) of understanding mind as machine and life as mechanical. Reasons for this existence in the spotlight can also be found, however, in the aura formed around the visually engaging spectacle that cybernetic robots used to provide –a veritable “gallery of monsters”.<sup>108</sup> This almost surreal world –a pictorial reminder of the post-modern times that were being lived-- was filled with

---

<sup>105</sup> Ibid. For the Soviet development of cybernetics –a story still fairly unknown in the West-- see Gerovitch 2002.

<sup>106</sup> For an account of Lacan's appropriation of cybernetic thinking applied to psychoanalysis, see Johnston 2008, ch. 2.

<sup>107</sup> See Heidegger 1966 and Heidegger 1969.

<sup>108</sup> This is Andrew Pickering's coinage. Cf. Pickering 2005

reported anecdotes of autonomous robots chasing women's legs,<sup>109</sup> and children being surprisingly (again, for the time) prone to interact with these “mechanical animals”.<sup>110</sup>

Here it is relevant to indicate that the role of Britain in cybernetics emerges as fairly substantial when we refer to these mechanical contraptions. The British collaboration with the cybernetic movement has been lately acknowledged with increased emphasis.<sup>111</sup> To be fair, despite Wiener's somewhat belittling remarks after a visit to the UK,<sup>112</sup> some members coming from across the Atlantic did play an important role in the unfolding of the whole cybernetic enterprise –in particular towards the end. In line with the more informal cybernetic development coming from British soil, they had their own scaled down version of the reportedly fancy Macy gatherings<sup>113</sup> –indeed, in the form of a small but exclusive “club”.

## 2. The Ratio Club

The Ratio Club was a small community of British scientists that gathered in London on a regular basis to discuss all things cybernetic. Just as the Macy meetings, it served as a hub where scientists gathered to share and discuss their insights, cognizant that there was no other forum at

---

<sup>109</sup> This reportedly happened with the phototropic tortoises. The bright type of stockings used by women at the time used to reflect light. Spotted by the machine, it would resolutely aim towards the feminine lower limbs while avoiding men's trousered legs (Hayward 2001, p. 624).

<sup>110</sup> De Latil's book shows a photograph of Grey Walter, his wife and their son playing with a “tortoise”. The description below reads: “Vivian Dovey and Grey Walter have two offspring: Timothy, a human baby and Elsie, the tortoise, of coils and electronic valves. Timothy is very friendly with his mechanized sister.” (De Latil 1956, p. 34)

<sup>111</sup> First by Husbands & Holland (2008); later by Pickering (2010).

<sup>112</sup> Wiener characterized his 1947 visit to England thus:

I found the interest in cybernetics about as great and well informed in England as in the United States, and the engineering work excellent, though of course limited by the smaller funds available... I did not find, however, that as much progress had been made in unifying the subject and in pulling the various threads of research together as we had made at home in the States (Wiener 1948, p. 23).

<sup>113</sup> Conway & Siegelman, in describing the first Macy meeting, write that “They gathered for two days and two nights around a great circle of tables at the Beekman Hotel on Manhattan's Upper East Side, with their lodging, meals, and the cocktails Wiener would not touch paid for by the Macy Foundation.” (Conway & Siegelman 2005, p. 155)



the time where they could discuss cybernetics. Expectably, they were filled with enthusiasm, and members would always be looking forward to the next meeting. Nevertheless, there were some remarkable differences from the Macy conferences as well. Although the Macy gatherings appreciated the spontaneous discussions that followed the presentations, the London meetings had a decided bent on keeping the gatherings more informal than in New York. The pace was set by talking over drinks, with no records ever produced of the discussions. They were to meet once a month, at night, in the basement of a hospital –arranged with chairs, food and beer. These kept going in a relatively uninterrupted fashion from 1949 to 1955. In order to maintain the discussion in total freedom, without affectation from an “authority figure”, no professors were allowed. Also, if someone would become someone else’s boss, it was expected he would resign. Following this vein, the one-time proposal for a follow up in a more professional setting, probably with a journal, was flatly rejected.<sup>114</sup>

The origin of its name can be traced back to the semantic implications stemming from the fact that the Latin root *ratio* underpinned several concepts important for the group.<sup>115</sup> Leibnitz is mentioned by Wiener as a precursor to cybernetic thinking, when he acknowledged that his *machina ratiatrix* was anteceded by Leibnitz’ *calculus ratiator*.<sup>116</sup> The club members were also interested in the notions of ratiarium (statistical account) and ratiocinatus (argumentation). *Ratio*, the common root for all these notions, seemed like the right choice.

Over and above the mentioned differences existing between the Ratio gatherings and Macy’s, there was a feature that set them apart in a qualitatively distinct way. Despite the highly

---

<sup>114</sup> This usually was Ross Ashby’s idea. He was more than amenable to the prospective of constituting a more formal, open society that could reverberate into a journal. Other members vehemently disagreed. (Husbands & Holland 2008, p. 128)

<sup>115</sup> *Ratiocinatio* is St. Thomas Aquinas’ term for discursive reasoning, namely, the natural reasoning used by humans when a truth is sought –as opposed to the angels’ intuitive knowledge of things, or God’s (Husbands & Holland 2008, p. 102)

<sup>116</sup> See above note on cybernetics’ patron saint according to Wiener.

abstract origin of its name, the club was largely about device making, and the reason for this lay partly in its recent contemporary history –indeed a sad history that for some of the members was still very fresh. One member died just before making it to the first meeting: Kenneth Craik (1914–1945). Craik’s “synthetic philosophy”, which will be addressed in the next chapter, in fact strongly informed and defined an essential aspect of the group: its emphasis was on device-construction as a means of explanation. In fact, at least one Ratio member referred to what was going on in the American side as “thinking on very much the same lines as Kenneth Craik did, but with much less sparkle and humour”.<sup>117</sup>

One main criterion for recruiting members (which were always kept constant at around twenty, between core members and guests) was that they should have had, in John Bates words, “Wiener’s ideas before Wiener’s book appeared”.<sup>118</sup> There was substantial emphasis placed on this mandate, since there was the conviction that the American side, to a great extent, was enjoying the fruits of having merely put cybernetic ideas together with impeccable timing. As advanced above, many of those ideas were allegedly already floating around before Wiener and company --and adding insult to injury, many were coming from Britain. In fact, the Ratio group did have prominent members that went down in history as prominent cyberneticians or thinkers in their own right<sup>119</sup> --such as Alan Turing, Grey Walter,<sup>120</sup> Donald Mackay and Ross Ashby.

---

<sup>117</sup> Husbands & Holland 2008, p. 14

<sup>118</sup> This criterion was several times repeated in letters throughout the formation of the club. The neurologist John Bates, who came up first with the idea of these gatherings, wrote in one of his invitation letters: “I know personally about 15 people who had Wiener’s ideas before Wiener’s book appeared and who are more or less concerned with them in their present work and who I think would come.” (Husbands & Holland 2012, p. 242)

<sup>119</sup> They did acknowledge as “honorary members” McCulloch, Pitts, Shannon and either Wiener or Weaver. A list of the club membership in a member’s correspondence reads ‘Mc’, ‘P’, ‘S’ and a letter that could be either a ‘U’ or a ‘W’. Husbands speculates that it could refer to either Norbert Wiener or Warren Weaver (Husbands & Holland 2008, p. 98).

<sup>120</sup> William Grey Walter (1910–1977) was an American (naturalized British) roboticist and neurophysiologist, largely self-taught while working at hospitals. He is best remembered by his famous mechanical “tortoises”, displaying animal-like behavior at festivals and media gatherings. He called them both *machina speculatrix* (and each, Elmer and Elsie) because the way in which they would find their way around environments seemed to suggest that they were “deciding” about possible routes.

These almost patriotic statements might up to some point be understood as reactive to some of Wiener's remarks regarding the state of cybernetics in England. As mentioned above, Wiener acknowledged in his book the originality and enthusiasm that British scientists demonstrate, but pointed out, probably not without reason, that they were not as developed and organized as in America.<sup>121</sup> After all, the very nature of the Ratio Club pointed to a hobby-style, "indie" type of pursuing of cybernetic insights,<sup>122</sup> without an official backup coming from the British government.<sup>123</sup> Wiener's remarks might have been exacerbated in their effects after McCulloch's talk at the Ratio Club. There was a unanimous feeling that it fell flat. In addition, due to McCulloch's well-known penchant for flowery language and grandiose remarks –after all, he was heavily trained in the Classics—a member voiced the feeling that they found "Americans less clever than they appear to think themselves".<sup>124</sup> Alan Turing referred to him as "a charlatan" straight up.<sup>125</sup>

In fact, there seemed to be an undercurrent of wounded pride throughout the British cybernetic development.<sup>126</sup> As an example, J. O. Wisdom's article "The Hypothesis of Cybernetics" might fall within this context.<sup>127</sup> Wisdom was intrigued by both the audacity of the cybernetic proposals, and the rigor --even elegance-- of their theoretical premises and procedures. However, Wisdom somewhat reduced the cybernetic hypothesis to a sophisticated

---

<sup>121</sup> Wiener 1948, p. 23

<sup>122</sup> Pickering 2010, p. 10

<sup>123</sup> In fact, the opposite could be said as having been the case regarding the support of cybernetic pursuits by the British government. As a telling example, at some point British Intelligence began to be concerned about the amount of military surplus that Grey Walter was purchasing, and instructed someone to find out, in a covert but official capacity. Walter was just scavenging for useful parts to build his 'tortoises' (Husbands & Holland 2008, p. 127).

<sup>124</sup> Husbands & Holland 2008, p. 114

<sup>125</sup> Hodges 1983, p. 411

<sup>126</sup> Perhaps present until today's accounts of what happened back then. British philosopher Andrew Pickering does not hide that he wants to "set the record straight" regarding the absence in history of the British contribution to cybernetics. In fact he is clear in stating that his book is not a treatment of cybernetics, but of an aspect of British cybernetics (Pickering 2010, pp. 3-4).

<sup>127</sup> See Wisdom 1951

mechanistic view of the nervous system. And the reason for this might be partly historical in nature.

Not long after the role of negative feedback in machinery had been carefully examined, the hypothesis that claimed that the nervous system was, in fact, a machine with negative feedback, saw the light. This might have been influenced by the fact that, particularly in the UK, many biology researchers were recruited, due to the war efforts, to work in radar research and development.<sup>128</sup> These biologists became thus familiarized with electronics and mechanics, and biological realms began to be framed under a mechanistic guise. Conversely, urgently needed research on radar technology was an incentive to think of it as an artificial extension of human sensoria. This British emphasis on radar development, as seen in the previous chapter regarding American predictors being coupled with English radars, was an important aspect of the context of development of the Ratio Club.

Wisdom was ultimately critical of the scope and limits of the cybernetic enterprise, equating it to a mechanical hypothesis of the nervous system --which was but just *one* aspect, even if important, of the cybernetic epistemology. Wisdom seemed to have missed the crux of the theoretical pillars of the enterprise --namely, that of the assumption of the perfect possibility, nay, the hitherto unacknowledged existence, of an entire class of non-material (disembodied) machines. And the patriotic context mentioned above might have had something to do with this. It could be reasonably entertained that Wisdom might have wished to point out that the tradition of research on the nervous system within the context of computation was on the British side at least as good as on the American one --as depicted by von Neumann and McCulloch.

---

<sup>128</sup> John Bates described the core membership of the club as “half primarily physiologists though with ‘electrical leanings’ and half communication theory and ex-radar folk with biological leanings” (Pickering 2010, p. 59).

Be that as it may, the group enjoyed a frequency that Macy, only once a year, did not. Also, at least reportedly, the meetings were less messy and conflictual than at Macy –as reported above. Rather, members acknowledged that a true interdisciplinary discussion (at least in spirit) used to emerge.<sup>129</sup> Most remembered the Ratio meetings as bringing intense academic and intellectual fulfillment to its members –in fact, for some, those meetings gave the most insightful and helpful insights for their own, later prominent careers.<sup>130</sup> From 1949 to 1951, the meetings were as frequent as scheduled; however, attendance began to dwindle after 1952. There were only four meetings in 1953, three in 1954 and two in 1955. There was one more, sort of “final reunion” in 1958. And the club never met again.<sup>131</sup>

### 3. The decline.

From the pre-Macy conferences up until half the lifespan of the Macy cycle, most cyberneticians were in a state of a nearly constant adrenaline rush. But some began to grow worried regarding the grand-promises that this new science was to deliver to the world, as reported by the popular

---

<sup>129</sup> As opposed to the chaos of Macy where some –mostly social scientists—used to be lost. This is how Dupuy describes it:

The cyberneticians showed no reservation about entering into technical discussion with the widest variety of specialists, examining with equal ease the results of a psychoacoustic experiment and a theory of the conditions under which the words of a language acquire a specific individual meaning. The few generalists present, the most notable of whom was Gregory Bateson, often found themselves lost. In their frustration at being unable to follow the discussion, they were apt to beg their colleagues not to lose sight of the universalist vocation of cybernetics. Careful examination of the transactions of the conferences makes it painfully clear how "out of it" Bateson actually was (Dupuy 2000, p. 88).

<sup>130</sup> Husbands & Holland 2008, pp. 138-141; Holland & Husbands 2011, pp. 120-122

<sup>131</sup> Reasons for its slow decline were diverse. For one thing, it was too far for many, having to return home very late at night, to the annoyance of their respective wives. On the other hand, it was fairly expensive to get there, and for some the costs were not covered by their institutions. However, it would seem that the most important factor was that by the early 50's cybernetics had already reached a fairly mainstream status, and so the ambience of originality, uniqueness, and to some point compelling mystery and scientific heroism was no longer there. Allegedly members lost interest as a consequence (Husbands & Holland 2008, p. 129). Needless to say, these are just some of the reasons as reckoned by Husbands and Holland. There might be others that shall be gradually unveiled later in time, after researchers take interest in the fate of this interesting group.

media. This conflicting reception can be noticed when seeing from an eagle's view both the unanimous motion of approval for having the series of Macy conferences' title changed to Cybernetics *and* the words of cautious advice voiced by its own members.

Despite being the target of popular attention (or perhaps, precisely due to it), Norbert Wiener was right from the beginning relentless in trying to calm down what he regarded as an unwarranted hype around the cybernetic project. Not only would he shun claims that could be deemed plain excessive, but indeed more importantly, Wiener was particularly skeptical of the social applications of the powerful cybernetic concept. Bridging heretofore ontologically incompatible realms, such as man, machine, and life –thus extending the realm of physical science—was difficult enough. And he felt understandably proud about such a feat. But extending the mechanical umbrella to portions of reality where the amount of variables was qualitatively over and above the tractable was the place where he would draw the line. The unpredictability of social behavior, for Wiener, inherently escaped numerical framing.<sup>132</sup>

These are relevant observations, since they point to a web of internal tensions already present within the project right from the start –tensions that have been traditionally highlighted as part and parcel of the eventual fate of cybernetics. Almost right from their beginnings, the Macy conferences witnessed sociologists and anthropologists as some of the most enthusiastic 'converts' to the cybernetic view --to the discreet discomfort of some cyberneticians, Wiener being prominent among them.<sup>133</sup> Seemingly, some scholars emanating from the “soft” sciences

---

<sup>132</sup> While Norbert Wiener was voicing his cautious outlook on applying a cybernetic framework to society, just a year after his seminal 1943 paper, fellow cybernetician John von Neumann was publishing a book with Oskar Morgenstern on *Theory of Games and Economic Behavior* (1944), which launched within economics --itself traditionally regarded as a social science— entire subfields heavily based upon mathematics. There are those who believe that von Neumann contributed to the theoretical backbone of contemporary economics more than what he is usually credited for. For an example of this view, see Mirowski 2002.

<sup>133</sup> This is how Wiener would characterize the import of cybernetics over the social sciences:

I mention this matter because of the considerable, and I think false, hopes which some of my friends have built for the social efficacy of whatever new ways of thinking this book may contain. They are certain that

were more than eager to adopt a heuristic tool that would bring the scientific validation that was eluding them right from the foundation of their own fields –including psychology, *pace* Skinner.<sup>134</sup>

The tension witnessed between these two grand realms of knowledge within cybernetics may well point to another source of unease: the radical interdisciplinarity of the movement. Part of the allure of cybernetics was the deep realization, common across fields, that this neo-mechanistic understanding of life and mind could encompass a novel unifying language, structured upon the strengths of physics and mathematics. This approach, which necessarily entailed the interdisciplinary gatherings of Macy, was for some members one of the most attractive features of the cybernetic proposal.

However, for some others, particularly coming from the “hard” sciences, the attempt at dialoguing with other disciplines was coming at a high price, entailing a substantial dissolving of their hard earned, specialized, field-centric semantics –and all that just for hooking up with not-so-neighboring fields anyway. This alleged theoretical bridge would be founded, according to its critics, more on vagueness than on substance –and that was too much of a hard pill to swallow. Some indeed never came back.<sup>135</sup> All tensions and early defections notwithstanding, the overlapping consensus among cyberneticians was that something deeply important –indeed, a type of 2<sup>nd</sup> scientific revolution of sorts-- was in the making, and thus chose to be patient and stick to the cybernetic project. The opportunity for a profoundly positive outcome for science

---

our control over our material environment has far outgrown our control over our social environment and our understanding thereof. Therefore, they consider that the main task of the immediate future is to extend to the fields of anthropology, of sociology, of economics, the methods of the natural sciences, in the hope of achieving a like measure of success in the social fields. From believing this necessary, they come to believe it possible. In this, I maintain, they show an excessive optimism, and a misunderstanding of the nature of all scientific achievement (Wiener 1948, p. 162).

<sup>134</sup> See the first section of the previous chapter (I. 1.) for Wiener’s complain against the poverty of traditional behaviorism.

<sup>135</sup> One striking example of this attitude was Max Delbrück’s refusal to return to the conferences, mentioned in the previous chapter. I will come back to this occurrence in the chapter pertaining to von Neumann’s work (VI).

would be too great to dismiss, and just due to some wrinkles in the way it was being laid out.

Probably most importantly, and this aspect has arguably not received enough attention, is the impact that cybernetics had on its members, underpinning later advances in technology and science. In the case of American cybernetics, the fields of information theory, computer science, genetics, artificial intelligence and artificial life, could be traced back to its cybernetic roots in a fairly straightforward manner –although this uncompleted task has not received enough attention from the scholarship.<sup>136</sup> In the case of the Ratio Club, the testimony of some of its members regarding their indebtedness in this regard to cybernetics is particularly striking.<sup>137</sup> The difficulty in pinning down what exactly it is about cybernetics that manages to carry until now an aura of fascination might be deeply intertwined with this holistic and fluid aspect of its scientific ethos. All these historical circumstances, theoretical features and even spin-offs emanating from the cybernetic enterprise, stand as witness to the greatness of the project according to its founders, its profound role displayed upon its members, and its well acknowledged influence at the time on the scientific community at large.

However, after an intense decade of enthusiastic interaction between scientists, mathematicians and engineers of the first order --reciprocally connecting in one novel theoretical

---

<sup>136</sup> In fact, Gualtiero Piccinini's review of Dupuy (2000) points to this lack, fulfillment allegedly somewhat promised in the book's preface (Piccinini 2002). Still Dupuy (2000, ch. 2) could be a good start for connecting cybernetics with the cases of artificial intelligence and cognitive science. The legacy of cybernetics in the areas of genetics and information theory is briefly covered by Kay (2000, ch. 3). Langton (1989) addresses von Neumann's foundational role for artificial life. For the cybernetic ancestry of computer science, see Boden (2006, pp. 157-162, 195-198).

<sup>137</sup> It would seem that the strong emphasis on the informality of the club was precisely what made some of its members cherish it even after its extinction. Donald Arthur Sholl, a member of the club whose research in neurology is consulted until now (Sholl 1956) said this of the impact of the discussions in his professional life:

I consider membership of the Club not only as one of my more pleasant activities but as one of the most important factors in the development of my work. I have stressed before how valuable I find the informality and spontaneity of our discussion and the fact that one does not have to be on one's guard when any issue is being argued. At the present time we have a group of workers, each with some specialised knowledge and I believe that the free interchange of ideas which has been so happily achieved and which, indeed, was the basis for the founding of the Club, largely results from the fact that questions of academic status do not arise (Husbands & Holland 2008, p. 129).



structure of common knowledge-- a sudden general vacuum seemed to have taken over. Despite the success and fame that the movement enjoyed from 1943 on, just a decade later it came to a relatively abrupt halt. By 1951 two of its main contributors, von Neumann and Wiener, permanently dropped out from the Macy conferences. By 1953 the last Macy conference took place. Not even informal gatherings ever happened again in America. Whatever happened between 1943 and 1953 that took the world by storm, way beyond the exclusive realms of academic and scientific circles and into the educated average population at large, was not there anymore.

Their British counterpart, the Ratio Club, continued to meet sporadically, gradually dying out during a span of five years, until 1958. By the time the penultimate Ratio Club reunion was taking place in 1955, the American side of the cybernetic movement was practically dissolved. By that time, having morphed into another project, which did not have the flare, allure, tenets and members of the original project, cybernetics ended up in a relatively obscure program in an American university for one more struggling decade. It was led by one of the scientists who was in charge of publishing the proceedings of the Macy meetings, Heinz von Foerster, and provided institutional cover for one of the last raunchily self-proclaimed cyberneticians, the British Ross Ashby. Aside from this attempt to keep it alive in a substantially different guise, the ground shaking ideas firing up bold statements in the media and elsewhere were all but gone.<sup>138</sup> For all intents and purposes classical cybernetics was effectively dissolved and its electrifying vibe dead by 1954.

---

<sup>138</sup> This “post-cybernetic” approach started when the Ratio Club officially ended, in 1958, at the Biological Computer Laboratory at the University of Illinois in Urbana Campaign –mostly funded by the military. This follow-up movement went by the name of “2<sup>nd</sup> order cybernetics” (or “cybernetics of cybernetics”). It was led by Heinz von Foerster, the physicist who was given the task of recording the proceedings of the Macy Conference, and it housed the cyberneticist Ross Ashby. Due to lack of funding, and to von Foerster’s retirement, it shut down in 1970. It always stood as a relatively obscure enterprise without any of the flavor or cache that characterized cybernetics two decades before –only known to hardcore “neo-cybernetic” followers. For an account of the fate of this laboratory, see Müller & Müller 2007.

This is an astonishing development, even today, after the fact. Six decades later, we are still asking ourselves what happened. How can one explain this? How can a movement that merited the appraisals (sometimes tinged with fear) of thinkers such as Heidegger and Lacan, that was touted as the science of the future by eminent thinkers, that received substantial funding by the state (at least on the American side),<sup>139</sup> and that was hailed by media as the next step for humanity, come to such an abrupt and unremarkable end?

Several attempts at answering this inquiry have emerged –each tackling one or more levels of explanation. Indeed there have been attempts at articulating this demise coming from several regions of knowledge. Each one has its own strength, and each one points to different aspects of the troubled project towards the end. Some bring economic reasons for cybernetics’ death. Competing cognitive proposals soaked up all the available funding for this type of research, effectively leaving cybernetics to dry out. Some others point to the inherent tensions alluded to above. Some even pin it to reasons of a personal nature. Without dismissing reasons of funding, uneasy interdisciplinarity, or even conflict of personalities, a philosophical assessment has not been sufficiently laid out so far to seek the reasons for the demise of cybernetics.

Indeed entering deeper into core dimensions of cybernetic theory one could find enlightening signs pointing to what could have happened, at the level of the solidity of its theoretical framework. Most importantly for this thesis, these cues revolve around the development and evolution of the nature of a machine. This evolving notion was embedded within the cybernetic mindset deeper than many cyberneticians probably realized. Plunging deep into the investigation of the nature of a machine does provide angles to understand better

---

<sup>139</sup> British cyberneticians pursued their endeavor more in a hobby-like style, putting their ideas to work in their free time. During the war, funding was funneled to research on computation –having as a prime aim the decryption of the enemies’ communication (as opposed to American research on computation for building tables for anti-aircraft weaponry, and eventually, the atomic bomb). After the war, the rebuilding of Great Britain was the main target of funding. Government funding was never officially provided to the British side of cybernetics.

whatever happened to cybernetics. As well, it opens up doors for understanding later and current approaches to science that carry the cybernetic signature.

But first, I will provide an updated overview of the typical reasons so far adduced for the death of cybernetics.

#### **4. Traditional explanations for the collapse of cybernetics.**

Perhaps the most well-known version attempting to explain cybernetics' demise is that it simply ran out of funding. This story emphasizes the alleged fierce competition that streamed from some of the people that were ironically under the influence of senior cyberneticians, but that later had antagonistic proposals of their own. The story line links cybernetics to artificial neural networks in an essential fashion –courtesy of having both McCulloch and Pitts' networks as a common ancestor. The gist of this explanation relies on the alleged fact that a “symbolic architecture” approach was shown to be capable of dealing with some problems, such as theorem proving and chess playing. This dominant change of perspective, coming from the founders of what became later known as Artificial Intelligence (AI), supposedly funneled to itself most of the available funding. This money --still mainly being granted by military offices-- went to fund the symbolic approach, leaving cybernetics out in the cold.

Marvin Minsky and Seymour Papert, both disciples of McCulloch, published a book severely criticizing Frank Rosenblatt's “perceptron”<sup>140</sup> –a core element of neural networks, and allegedly, by extension, a core element of cybernetics. It supposedly showed the inherent limitations of single neurons in sophisticated thought processes, a weakness particularly shown in the area of problem solving. The implication was that if the artificial neuron was not capable

---

<sup>140</sup> Rosenblatt 1958

in small realms (lacking as a standalone or in simple networks), then it could not be good in the big ones (more complex networks would never be able to develop higher processes). This report<sup>141</sup> was regarded by many as the final nail in the coffin for network-related research—and a fortiori, for the distancing from cybernetics as an acknowledged influence.<sup>142</sup>

At least two main criticisms can be launched against such an account. Firstly, as it has been already advanced (and will be further discussed later), although neural networks were part of the core of cybernetics (as we saw above, McCulloch and Pitts' article was indeed a foundational one), it is clear that the networks' approach was not the only one that cybernetics had at the core of its own theory. In fact, criticisms against neural networks were launched by von Neumann within the very heart of the cybernetic enterprise, advocating for an alternative view, probably even more cybernetic in spirit—namely, the possibility of building self-reproducing machines. One could even safely assume that Wiener's lack of reaction to von Neumann's severe criticism of neural networks might be explained by the fact that for Wiener the core acceptance of networks was not *a sine qua non* condition for the cybernetic endeavor to continue and flourish. For him there were other more important elements whose widespread acknowledgement was more essential for the cybernetic insight, such as the nature of the message, information and control.<sup>143</sup>

The other criticism that can be pointed out against this typical account of cybernetics' demise is more historical in nature. By the time Minsky was criticizing the “perceptron” (1969),

---

<sup>141</sup> Minsky & Papert 1969

<sup>142</sup> This approach put emphasis on symbol manipulation. On the one hand, it began to heavily rely on the advance of computers—instead of neurology—and on the other it proposed a systematic structure of mental states, where these symbols were manipulated. It later developed into both Artificial Intelligence and the Computational Theory of the Mind underpinning it. Eventually it was philosophically sanctioned first by Hilary Putnam and Noam Chomsky, and later by Jerry Fodor—who self-admittedly was very hostile to artificial neural networks as a general model of cognition. For an account of this traditional view, both regarding the advantages of the perceptron and the aftermath for connectionism after its criticism, see McCorduck (2004, pp. 102ss); Dupuy (2000, pp. 102ss); Boden (2006, pp. 903ss) and Johnston (2008, pp. 287ss).

<sup>143</sup> See Wiener 1948, Introduction. Also Conway & Siegelman 2005, ch. 6.

cybernetics was largely already dead. This traditional account is thus best located within the context of the origins of cognitive science, particularly the area concerning early neural networks --not within cybernetics at large. And in any case, even if what little remained of cybernetics at the Biological Computer Laboratory at the University of Illinois actually competed for funding with the new and eager generation of AI researchers, it would be inaccurate to affirm that what such “neo-cybernetic” group was asking, namely, funding, equated to cybernetics competing for funding.

In fact, there is the view that the very existence of the Biological Computer Laboratory was the living proof that cybernetics was officially dead.<sup>144</sup> The so called “Second order cybernetics” era, or “Cybernetics of cybernetics”, attempted to include the observer into the system, dissolving the whole mechanizing cybernetic view of control. It effectively nullified the bold proposals being advanced by classical cybernetics a decade before, leaving its strong ontological stance.<sup>145</sup> It switched to an epistemological preoccupation with the observer of the system, partly inspired in quantum mechanics’ concern with the pervasive role of the observer in any possible measurement within subatomic phenomena. Von Foerster was indeed a physicist by training and was keenly interested in quantum problems right from the moment he entered the first Macy conference. All these “autopoietic” insights, which brought Humberto Maturana and Francisco Varela to the laboratory, escaping from Pinochet’s Chile and infused with “the Buddha and the French phenomenologist, Maurice Merleau-Ponty”,<sup>146</sup> as well as with some allegedly Heideggerian insights, could have evolved into something unique and interesting –and maybe it

---

<sup>144</sup> See Asaro 2010

<sup>145</sup> For Andrew Pickering, cybernetics was substantially more about ontology than about epistemology, instantiated in the mechanical theatre of “monsters” (Ashby’s homeostat, Grey’s tortoises, Shannon’s rats) that it produced. See Pickering 2002.

<sup>146</sup> Dennett 1993

did.<sup>147</sup> But what is clear is that it resembles little to the original cybernetic enterprise. So if it is indeed the case that it replaced it, then one could conclude that cybernetics was, in a very real sense, dead.

What is more, a further detail can shed some light on the misconception of identifying the advance of AI as spelling doom for cybernetics. The generation of disciples that followed the cybernetic masters went the extra-mile to deny acknowledgement of having its parental roots squarely set in cybernetics. After the attack of AI to networks research, anything related to the latter –including cybernetics-- would suffer being side-stepped for funding. This “guilty by association” context might have had deeper reasons than just playing safe for obtaining grants. In fact, it is an area of history of science that has been largely unexplored.<sup>148</sup> Again, this explanation, situated at the beginning of cognitive science (1960s) accounts for a time when cybernetics was already dissolved. Denying the cybernetic origin of cognitive science is an issue worth studying in its own right, but it gives little in terms of explaining why the theoretical parent experienced atrophy in the first place.

Still related to theoretical issues, there is another one that has lately attracted more attention, but that is nevertheless intertwined with personal issues. It has to do with what occurred after von Neumann’s letter to Wiener, regarding the inadequacy of McCulloch and Pitts’ neural networks for articulating the human mind --and von Neumann’s subsequent proposal for exercising a “back to the basics” move, instantiated in the study of the self-

---

<sup>147</sup> Cf. Varela, Thompson & Rosch 1991. This book, which combined phenomenology, cognitive science and Buddhism, enjoyed a moment of fame within the areas of philosophy of mind and cognitive science itself. Roberto Cordeschi’s assessment of this proposal is relevant:

It is still unclear whether these positions bring any advancement in our understanding of cybernetic related phenomena. On the other hand, many important and legitimate requirements underlying these positions seem to be already fulfilled by the very tradition that they are challenging, whenever the latter is not caricatured or oversimplified (Cordeschi 2008, p. 195)

<sup>148</sup> Gualtiero Piccinini’s critical review of Jean Pierre Dupuy’s book (2000) on cybernetics points out to this absence. See Piccinini 2002.

reproductive capabilities of virus. This (in)famous letter was sent after von Neumann delivered the opening talk at the first Macy meeting --and actively participated in the second one. It set up a meeting with Wiener that supposedly took place in December of the same year (1946).

Whatever happened in that meeting set out to discuss the discomfort vented in the letter, it was evident for many that a rift between the two heretofore close-friends had occurred --and this rift was expressed in almost eccentric ways. It was reported that Wiener would loudly snore at von Neumann's talks. Von Neumann, in turn, would loudly flip the pages of the New York Times, seated in the front row, while Wiener was lecturing.<sup>149</sup>

Whatever the effect the letter might have had, there were reasons for Wiener's reaction -- even if delayed. For starters, when the friendship between both men was at its pinnacle, Wiener's lobbying at MIT managed to procure for von Neumann a position at the mathematics department. MIT had a long tradition of engineering, while Princeton had a number of Nobel Laureates. Wiener was probably interested in bringing that kind of flare to the cybernetic hub -- which for the lack of an official institutional setting, was de facto occurring at MIT. Von Neumann used this important offer as leverage to persuade Princeton to give him funding for building his computer machine --which he obtained, promptly refusing MIT's invitation.

As if that maneuver was not enough, von Neumann asked Wiener to "lend" him his chief engineer, Julian Bigelow, whose genius was so crucial for building his predictor --and Wiener agreed! The double irony of this move is that the machine to be constructed had as its main aim the aiding of von Neumann's secret assignment --the building and improvement of atomic bombs. All these while Wiener had already become deeply disturbed on account of the

---

<sup>149</sup> See Heims 1980, p. 208. In all fairness, Wiener's loud snoring is part of the collection of his eccentricities -- mainly due to the fact that once awoken, he would provide the most penetrating insights regarding what was said during his sleep (Heims 1980, pp. 206-207). The difference is that this feature was no longer taken as a picturesque occurrence; rather, von Neumann took offence from it, triggering his conscious noisy behavior while Wiener was lecturing.

destructive capabilities of nuclear power –and hence the grave moral responsibilities of men of science.

Von Neumann was an active advocate of a pre-emptive strike against the Soviet super power. In sharp contrast, Wiener was severely affected after the bombing of Hiroshima and Nagasaki; his stance regarding war changed 360 degrees, making him a vociferous peace supporter. In fact, this newly acquired stance, which was represented in his resolute decision to walk away from all government and military contracts, was seen with suspicion by the McCarthyism of the Cold War era. This development has been identified as one more reason why cybernetics fell out of grace. Adding insult to injury, Wiener kept a close relationship with his Russian peers in science –who had developed a keen interest in cybernetics, and were largely backed by the Communist Party. This might have triggered a negative aura around Wiener and his project --a “cautious” view promoted by the government.<sup>150</sup>

As to what strained his relations with von Neumann, we have no proof, but Wiener might have caught wind of von Neumann’s side agenda. Or he might have had a delayed reaction regarding von Neumann’s Janus-faced “move” to obtain funding for his computer machine at Princeton –at the expense of Wiener’s good intentions. Or perhaps it was a combination of the above.

Not all outward resentment came from the side of Wiener, however. It is acknowledged that when Wiener made the link between his cybernetic incursions into control, communication and negative feedback on the one hand, and thermodynamics and chaos on the other (thus identifying information with negative entropy) von Neumann took it personally --and rather badly. Von Neumann had been also studying chaos for years, and he made a name for himself from his earlier years in the realm of the mathematics of quantum mechanics. However, he did

---

<sup>150</sup> Conway & Siegelman 2005, chs. 12 & 13



not elaborate his insights into chaos and entropy in connection with information to the point that Wiener did. In fact, there might have always been a rivalry particularly coming from him, who probably felt that Wiener always had the upper hand in mathematics –after all, the latter had been recognized as a veritable genius since his childhood. Von Neumann became jealous that Wiener established the startling correlation and not him. So what might have started as friendly academic competition could have developed into mutual hostility.<sup>151</sup>

There was still a further detail that might have played a role, above and beyond the distancing with von Neumann –in fact, with his distancing from the whole cybernetic group: an occurrence that has lately received more attention in terms of its possible influence on the dissolution of the whole enterprise.

Margaret Wiener, who was literally “shipped in” from Germany to marry Norbert, belonged to a severely conservative Protestant Puritan tradition. Also, she was a fairly overt admirer of Adolf Hitler, and her family, proudly Nazi. In fact, one of her brothers ran a concentration camp. Barbara, their first daughter, once got in trouble in elementary school because she candidly talked about the (Nazi) literature her mother avidly read at home. Apparently, the secularism of Wiener’s parents might have been sufficiently profound for Margaret to ignore Norbert’s Semitic roots. Rather, it was clear that for Margaret, the fame and stature of her husband on the world stage, and hence her social position, was the absolute priority. The whole family, Norbert included, went to a church of her denomination in a disciplined manner.

Since it is until now unclear what happened between both men, it is not out of the realm of possibility that Margaret, at a certain moment, turned her husband’s mind against von Neumann --a Jew (despite the fact that his family adopted Catholicism in the 1930s). They

---

<sup>151</sup> Ibid., pp. 165-166

already half-mocked von Neumann's impeccable dressing and ultra-formal manners.<sup>152</sup> She knew how to get Norbert Wiener to steer things her way --as it will be shown below. Margaret's ideas were likely already beginning to be distilled, following her own motto, despising the ways of a liberal life.<sup>153</sup> Von Neumann divorced and remarried after his first wife ran away with a graduate student of his.

If Wiener grew apart from von Neumann, chances are that the latter was not profoundly affected, since von Neumann always kept some distance from the cybernetic group anyway. He was in the group, without being part of it, so to speak. This might be explained, beyond his own way of being --a cosmopolitan man gifted with political maneuverability--with the sort of secret appointments that he took for the most part of his career. But the distance that occurred between Wiener and the rest of the group certainly had dramatic consequences. In 1951, three years after the publication of the book, Norbert Wiener severed all ties with his close friends and cyberneticians, particularly with Warren McCulloch and Walter Pitts. This episode has not attracted enough attention until lately,<sup>154</sup> and a brief exposition of what occurred seems fitting, in order to assess the degree of influence that this episode had in the entire cybernetic endeavor.

From the moment that Wiener and McCulloch had met, a gradual but eventually close friendship was formed; it was only to grow more intense in the subsequent years after the publication of the article in 1943 and the years of the Macy conference. McCulloch acquired a ranch (with a farm) in Old Lyme, Connecticut, which he renovated and put to work. This place was a hub for all the intellectuals, scientists and academicians that were around for a conference,

---

<sup>152</sup> Ibid., p. 144

<sup>153</sup> Margaret Wiener wrote down on a piece of paper what would probably summarize her outlook on social life:

One way to arrive at the aristocracy  
if you aren't born there  
is to eschew all forms of liberalism (ibid., p. 335)

<sup>154</sup> Ibid., pp. 213-234

or who just wanted to drop by and mingle with the McCullochs for a while. In fact, the place was big enough for guests to spend the night –even several days.

Reportedly, Norbert Wiener always was in the happiest of moods while he was at the ranch. The lake, the nature, the late chats over drinks over entire weekends, all these were a relaxing antidote for Wiener’s frequently stressed mind. The McCulloch’s were a remarkably liberal couple and such things as swimming at the lake in the nude was customary –for them and their guests. At least one person recalls the usual sight of Norbert Wiener floating on the lake, huge belly up, cigar in hand, loudly talking while floating away. Such was the happy, if permissive, ambience at the McCulloch’s ranch.

Needless to say, Margaret Wiener would not have ever approved of their conduct. In fact, unlike the other wives, she never accompanied Norbert to any of his weekend getaways at the McCulloch’s. Warren McCulloch’s pomposity, eccentricity, and for the time, markedly liberal ways, were more than what a conservative Puritan German could take --and the open repulsion against Wiener’s friend was immediate from the first time they met. Wiener’s daughters, Barbara and Peggy, recall that with absolute certainty, their mother Margaret never found out about this “debauchery”. Her reaction would have been, in all probability, nothing less than cataclysmic. Margaret Wiener’s disgust for this circle of friends triggered a long-term plan that seemed to have worked out at the end.

In 1951, three years after the publication of *Cybernetics*, and having published another successful book,<sup>155</sup> Wiener was trying to publish his memoirs --to no avail. Publishers found the manuscript to be too vitriolic and ridiculing of some people who had been Wiener’s mentors. Even Wiener’s portrayal of himself was embarrassing. The intended book was coming across as

---

<sup>155</sup> *The Human Use of Human Beings* (Wiener 1950), unlike *Cybernetics* (Wiener 1948), was intended to serve as an introduction to the new science for the layman. As previously mentioned, the 1948 book became a best-seller against all odds.

a personal rant against both the author himself and people who were still alive. Publishing companies did not find this material attractive. Some found it even parochial.<sup>156</sup> The uniform rejection, even from the publishing house at his own institution which previously fought for co-publishing *Cybernetics* (MIT), was taking a toll on Wiener's state of mind. In fact, it was exacerbating the intense depression into which he submerged himself out of writing about his own (not particularly happy) childhood. Wiener was as vulnerable as a seriously depressed person can get. His wife Margaret saw the opportunity to advance her plan of secession.

Some years before, when Barbara had begun her undergraduate studies in Chicago, Warren McCulloch expectably offered his house for her to stay, since it was close to her school. Wiener was grateful for the offer, accepted it, and sent Barbara to stay there for quite some time. The McCulloch's were so attentive about her staying with them, that once they even confronted her regarding a medical student she began dating. She was 19 years old and they felt responsible. It was all innocent and nothing came out of it, but history turned that episode into an irony.

Recall that this was the same period during which McCulloch gave shelter to Pitts and Lettvin. Margaret Wiener told Norbert, at the lowest point of his depression, that during the time when Barbara was staying with the McCulloch's, "the boys" (as the young cyberneticists used to be called) seduced their young daughter more than once. And more than one... Norbert Wiener went into shock and rage ensued.

Margaret Wiener had an extreme fixation with the perverse (for her) nature of sex, having submitted her daughter Barbara to several humiliations –all sex-related, and all implying she did something gravely inappropriate, when she did not. The Puritanical rattle was sometimes so

---

<sup>156</sup> As a publisher put it while rejecting the manuscript, "a book of almost wholly American interest." (Conway & Siegelman 2005, p. 218)

evident, that even Norbert Wiener, despite his legendary absent-mindedness,<sup>157</sup> on more than one occasion came to the rescue of his daughter from the intrigues of the mother.<sup>158</sup> However, by this time, Wiener's psyche was at an all-time low,<sup>159</sup> allowing Margaret to instill venom in his defenseless mind. Given the (sexual) nature of the accusation, and given the social ways of the time, Norbert was never going to further stir things and damage his daughter's reputation by inquiring about the truth of this revelation –outcome that Margaret likely knew quite well, successfully sealing her master plan.

Around the same time, Walter Pitts and Jerome Lettvin were in the best of moods. The Radiation Laboratory (Rad Lab) mentioned in the previous chapter was morphed into the Research Laboratory of Electronics after the war ended. This new lab had more than sufficient funds for serious research, and had Warren McCulloch as its head –for which he left his permanent professorship at the University of Illinois in Chicago. As if that was not good enough, the laboratory was going to have research bent towards human cognition inserted in the “bigger picture” of military research and development. Pitts and Lettvin, with their new fancy machines and bright prospects of adventurous investigation, were ecstatic and overjoyed. In accordance with their mood, they wrote a bombastic letter to Wiener (and Rosenblueth) the language of which was supposed to be taken as a prank. It started with the words “Know, o most noble, magnanimous and puissant lords...” In all likelihood, Wiener would have found it amusing, had he found himself in a “normal” state of mind. Margaret's “revelations” of his daughter's permanent “stain” at the hands of these “boys” had occurred just the night before. Wiener's

---

<sup>157</sup> Jackson 1972

<sup>158</sup> Conway & Siegelman 2005, pp. 199-200

<sup>159</sup> A psychologist friend of Wiener recounted how he would cry uncontrollably when he would begin to talk about his past --even while in the midst of a meal with him at a restaurant (Ibid., p. 218).

answer was terse in the extreme. The telegram that he sent in response to a colleague of theirs, so that he passed on the message, read (caps original):

IMPERTINENT LETTER RECEIVED FROM PITTS AND LETVIN. PLEASE INFORM  
THEM THAT ALL CONNECTION BETWEEN ME AND YOUR PROJECTS IS  
PERMANENTLY ABOLISHED. THEY ARE YOUR PROBLEM. WIENER<sup>160</sup>

As for McCulloch, Wiener opted instead for applying a permanent “silent treatment” --while at the same time vituperating his persona, academically and professionally, among common colleagues. Both McCulloch and Pitts (and Lettvin) were first in shock, then in denial. For a while, they thought that it was going to be a passing phase,<sup>161</sup> that things would later come back to normal, and that they would find out what happened. They simply did not (and could not) understand what took place. Gradually, however, it became clear that Wiener’s decision had a permanent character, and a dark cloud settled in for everyone.

McCulloch took it badly. Although he would continue with a “business as usual” attitude in his life and work, it was clear for those surrounding him that he was hurting and merely “pushing through” --surviving. The years of joyous friendship, of meeting for the most fulfilling conferences and gatherings, of seeing eye-to-eye on fundamental ideas about the world... these were not going to go away easily. Wiener himself confessed later that the break up with his colleagues and friends was taking a toll of him, not only mentally, but with psychosomatic consequences as well. He alluded to having contracted a painful heart condition due to this episode. By all accounts, however, the one who took it the worst, and by far, was Walter Pitts.

Pitts regarded Norbert Wiener as the father he never had. Even if McCulloch exercised a *de facto* “adopting” of him by letting him stay at his house when he was homeless, McCulloch

---

<sup>160</sup> Ibid., p. 219

<sup>161</sup> Wiener had short episodes of crisis in the past, overwhelmed by work and responsibilities. James Killian Jr., president of MIT at the time, recalled having in his desk drawer a folder with all the letters of resignation Wiener handed over the years (Ibid., pp. 220-223).

did the same for several other students. McCulloch was, to put it simply, just a generous man who cared in a very concrete sense for his students' well-being. With Wiener, however, it was different. Wiener brought Pitts to MIT, vouched for him, put him under his direct supervision, and lobbied for scholarships, grants and job positions for him. Wiener only had two daughters – as smart and intelligent as they were, given the times, he might have desired to have his legacy passed through to a son. In pressing Pitts to perform his duties and accomplish his goals (Pitts had a severe habit of procrastination), Wiener was very likely re-enacting the rigor that his own father had with him for his own good. Pitts, who escaped an uneducated environment where he used to be beaten, found in a very literal way, a father.

After Wiener dispatched the above missive (and knowing that Wiener was mouth trashing him and McCulloch), Pitts entered a spiral of self-destruction. He burnt all the research and writing that was going to grant him his PhD on a three-dimensional model of neural networks – a proposal that baffled mathematicians at the time with admiration, as mentioned above. He isolated himself from people. He stopped talking. He began to drink heavily. And he remained in that deplorable state until his death, which was alcohol-related (cirrhosis).<sup>162</sup> Tellingly, Pitts died only five years after Wiener, despite being considerably younger than him.<sup>163</sup>

Pitts' friend, Jerome Lettvin, the last cybernetician alive, died in 2011. Later in his life, he was asked to give an account of Norbert Wiener at a conference in Genoa, which he did with most empathy and in the best manner. At almost 80 years old, he was still hurting from the break

---

<sup>162</sup> See Smalheiser 2000

<sup>163</sup> It may be fair to say that the story of Walter Pitts is until now one of those great untold stories in the history of science. It might be the lack of sufficient written records about him, or that he did not formally accomplish much himself. But great minds around him reported the reliance of their own upon his help. He remains a sort of myth surrounding cybernetics: the man whose mind baffled Norbert Wiener (a genius himself), Bertrand Russell and Rudolf Carnap, and who ended up in total obscurity (See Easterling 2001).

up. Ten years after this episode, Arturo Rosenblueth finally told him what happened, recounting the machinations that Margaret Wiener put in place to secure their separation.<sup>164</sup> In spite of it all, Lettvin respected and admired Wiener until the end. Wiener's widow, Margaret, was at the same conference, and approached Lettvin to thank him for his kind words about her late husband. This is how Lettvin recalls the encounter:

I prepared a very careful, very adulatory talk and, afterwards. Mrs. Wiener came up to congratulate me and offered her hand –you know, she was a slight woman—but I really wanted to hit her as hard as I could, because I knew that she had contrived the break.<sup>165</sup>

Jerome Lettvin did not want to open up much regarding this painful episode for him and his friends, but being the last cybernetician alive, he did set the record straight before his recent death in 2011. These revelations have spanned a new round of possible explanations for cybernetics' decline. Given the powerful effects that passion and human emotions may have on historical occurrences of long lasting consequences, one could be tempted to reduce the whole development of cybernetics implosion to a unique, selfish and hidden act of a sole person. After all, Henry VIII's infatuated lust for Anne Boleyn in the 17<sup>th</sup> century ended up in the subsequent breakup of the entire United Kingdom from Rome, effectively creating an unsurmountable wedge in a divided West.<sup>166</sup> Could it be that one person's "matters of the heart" affect in such a

---

<sup>164</sup> Rosenblueth was present when Margaret told the "revelation" to Wiener, during dinner. He kept it for himself for several years (Conway & Siegelman 2005, pp. 224-225).

<sup>165</sup> *Ibid.*, p. 223

<sup>166</sup> In a paragraph brilliant in its succinctness, Elizabeth Eisenstein summarized the dramatic split of the West thus: Sixteenth-century heresy and schism shattered Christendom so completely that even after religious warfare had ended, ecumenical movements led by men of good will could not put all the pieces together again. Not only were there too many splinter groups, separatists, and independent sects who regarded a central church government as incompatible with true faith; but the main lines of cleavage had been extended across continents and carried overseas along with Bibles and breviaries. Within a few generations, the gap between Protestant and Catholic had widened sufficiently to give rise to contrasting literary cultures and lifestyles. Long after Christian theology had ceased to provoke wars, Americans as well as Europeans were separated from each other by invisible barriers that are still with us today (Eisenstein 2012, pp. 172-173).



profound way the future of an entire worldview? Indeed Oliver Selfridge<sup>167</sup> said of this gossip that

It really fucked up cybernetics...because here you've got the guy who invented the term and invented the idea right there with you, but there was no interaction at all with Norbert, which was a crying shame.<sup>168</sup>

Without attempting to subtract the causal power from the passionate act of a woman with twisted intentions upon a man's overreaching actions (be it Anne Boleyn or Margaret Wiener) one could however recall the traditional advice given to any first year philosophy student: A cause for an effect is seldom exhaustive. One particular situation may have several valid explanations intertwined, spanning across several levels of discourse and reality. In this section some of these levels have been laid out –theoretical, social, economic, political and personal. But projects that advance bold ontological (and epistemological) proposals in a robust way tend to supersede their own founders. In the words of Wiener himself,

We have contributed to the initiation of a new science which... embraces technical developments with great possibilities for good and evil. They belong to the age, and the most any of us can do by suppression is to put the development of the subject into the hands of the most irresponsible and the most venal of our engineers.<sup>169</sup>

It is reasonable to entertain the possibility that there is space for a further philosophical interpretation of what happened with cybernetics at the theoretical level. This is particularly appropriate given the fact that after the breakup, some cyberneticians were not only still active with a classical cybernetic outlook on things, but indeed took the project to the next level in different ways --indeed away from “irresponsible and venal engineers”. This was precisely the

---

<sup>167</sup> One of the latest and youngest additions to the cybernetics group before the break up, he died in 2002.

<sup>168</sup> Conway & Siegelman 2005, p. 233

<sup>169</sup> Wiener 1948, p. 28

case with William Ross Ashby and John von Neumann. Identifying what went on in each case will, I hope, further contribute towards the understanding of a notion of machine, and with that, shed some light on what possibly occurred with this movement. But first, I will lay out in the next chapter the context that allowed for the rise of cybernetics in the first place.

### **III. Pre-cybernetic context: An early 20<sup>th</sup> century ontological displacement of the machine.**

...understanding by a purely mechanical process  
one which could be carried out by a machine.<sup>170</sup>

#### **1. The “Foundational Crisis of Mathematics” and the response from Formalism.**

In the second half of the 19<sup>th</sup> century, the German mathematician George Cantor (1845-1918) engaged in a series of investigations regarding infinite numbers. Some of his observations were considered disturbing by some schools of thought in mathematics, and even more so the conclusions that were being reached from them. A typical Cantorian reasoning would lead to a paradox involving operations with transfinite numbers. For instance, if we have an infinite set of natural numbers, and a set conformed by all the sub-sets of natural numbers, which set would be larger? The notion of a cardinal number is introduced to account for the situation in which two infinite sets would still give an infinite set. This seemed self-contradictory and the tension was denounced, since it was occurring at the foundation of arithmetics --at the very core of mathematics.

The association of the notion of the “infinite” with the idea of God, put in place since early medieval philosophy, was still present in its own nuanced way, at the end of the 19<sup>th</sup> century. Indeed Cantor thought of himself as being the receptor of the mathematical insight regarding infinite numbers as a message delivered by God. In an occurrence that shows how controversial were his investigations at the time, Leopold Kronecker, Cantor’s own mentor,

---

<sup>170</sup> Turing 1939, p. 150

reportedly stated that “God made the natural numbers; all else is the work of man”.<sup>171</sup> Adding these metaphysical underpinnings to the extent of the consequences of his musings, might help clarify the reasons behind the subsequent furor against Cantor --who eventually broke down psychologically, ending his days in a mental ward.

This whole episode is prelude to what was known as the “Foundational Crisis of Mathematics” (*Grundlagenkrise der Mathematik*). The four main schools of thought in mathematics (logicism, formalism, intuitionism and Platonism) reacted to the claimed paradoxes. There is widespread consensus that the first two, logicism and formalism, which were substantially related, got the upper hand in responding to the crisis. Each displayed its own particular emphasis in the treatment of the problematic situation, and each triggered a different (although again, related) effect on mathematics, philosophy and science.

The logicist attempt at giving a solid foundation to mathematics resulted in efforts to reduce it to an extension of logic. But this route had some pre-conditional issues to resolve. Immanuel Kant famously stated that analytical judgments do not advance knowledge. Following the principle of non-contradiction, analytical propositions merely contemplate their own concepts as their object of understanding, finding the idea that was already embedded in them (e.g., a triangle has 180 degrees). Logic is the analytic discipline *par excellence*: obeying the law of non-contradiction, it engages in the contemplation of itself as an object of study. Although it provides no augmentation of knowledge, its own nature made it remain stable, not having changed since the time of Aristotle:

---

<sup>171</sup> The quote in German uses the word “integers” instead of “natural numbers,” however the latter is the way in which the quote is traditionally remembered (Bell 1937, p. 527).

For the advantage that has made it so successful logic has solely its own limitation to thank, since it is thereby justified in abstracting – is indeed obliged to abstract – from all objects of cognition and all the distinctions between them; and in logic, therefore, the understanding has to do with nothing further than itself and its own form.<sup>172</sup>

Thus, the rescuing strategy for attaining a solid body of mathematics would somehow have to base mathematical knowledge on the secure foundations of logic. However, the mere mentioning of the notion of mathematical *knowledge*, which implies an epistemic gain, already suggests that we are facing a stumbling block --since there is no augmentation of knowledge from strictly analytical grounds. However, mathematics, by most accounts, indeed provides an increase of knowledge. And this goes against, to repeat, the very nature of an analytical enterprise --such as logic. After all, logic is secure, but at the price of not having changed in millennia; whereas mathematics has clearly advanced. In fact, Kant viewed mathematics (specifically of arithmetic, geometry and algebra) as being synthetic *a priori* --just as the judgments that form the natural sciences (e.g., physics) are, where the advances after Newton were staggering. Kant defended arithmetic as synthetic in one of the passages that has garnered the most dissatisfaction among his readers, ever...

We might, indeed at first suppose that the proposition  $7 + 5 = 12$  is a merely analytical proposition, following (according to the principle of contradiction) from the conception of a sum of seven and five. But if we regard it more narrowly, we find that our conception of the sum of seven and five contains nothing more than the uniting of both sums into one, whereby it cannot at all be cogitated what this single number is which embraces both... We must go beyond these conceptions, and have recourse to an intuition which corresponds to one of the two—our five fingers, for example... For I first take the number 7, and, for the conception of 5 calling in the aid

---

<sup>172</sup> Kant 1787, B ix

of the fingers of my hand as objects of intuition, I add the units, which I before took together to make up the number 5, gradually now by means of the material image my hand, to the number 7, and by this process, I at length see the number 12 arise. That 7 should be added to 5, I have certainly cogitated in my conception of a sum =  $7 + 5$ , but not that this sum was equal to 12. Arithmetical propositions are therefore always synthetical, of which we may become more clearly convinced by trying large numbers. For it will thus become quite evident that, turn and twist our conceptions as we may, it is impossible, without having recourse to intuition, to arrive at the sum total or product by means of the mere analysis of our conceptions.<sup>173</sup>

However, the German mathematician Gottlob Frege (1848-1925) took upon himself the task of showing that arithmetic, which does provide augmentation of knowledge, is in fact an analytic discipline. The question is how to provide the solidity of logic without removing the capability of furnishing new knowledge? In his *Foundations of Arithmetic*, he severely criticizes the idea instantiated in the paragraph above –which to Frege’s credit, did trigger much discomfort not long after Kant provided his “visual” explanation. Frege’s attack, to some degree commonsensical, involved the multiplication of very large numbers, the premise being that one does not refer to spatial objects (e.g., fingers) when operating with those...

[I]s it really self-evident that  $135664+37863=173527$ ? It is not; and Kant actually urges this as an argument for holding these propositions to be synthetic... Kant thinks he can call on our intuition of fingers or points for support, thus running the risk of making these propositions appear to be empirical, contrary to his own expressed opinion... [H]ave we, in fact, an intuition of 135664 fingers or points at all? If we had, and if we had another of 37863 fingers and a third of 173527 fingers, then the correctness of our formula, if it were unprovable, would have to be evident right away, at least as applying to fingers; but it is not. Kant obviously was thinking of

---

<sup>173</sup> Ibid., B 205

small numbers. So that for large numbers the formula would be provable, though for small numbers they are immediately self-evident through intuition.<sup>174</sup>

Having accused Kant of underestimating the power of analysis (in the operationally fruitful sense), the ground was ripe for declaring that arithmetic is *both* a knowledge-augmenting discipline *and* an analytic practice. What was being accomplished was no small feat, as Frege was keenly aware. He was one small step closer to the ambitious Leibnitzian dream of subsuming everything under heaven to the power of *calculus rationalis*, getting thus closer to the sort of idealized *visio Dei* always longed for by philosophers. Now that arithmetic could be spoken of as analytical (without destroying its capacity for knowledge-rendering), the path was clear for mapping it upon that other discipline known by its secure foundations: logic. Being analytical no longer was a cardinal sin against the augmentation of knowledge.

As indicated above, there was more than one response to the foundational crisis that mathematics was experiencing. Besides the logicist account outlined above, what came to be known as the formalist approach exercised an impact in the history of mathematics that spilled over to neighboring areas of possible research, which gradually morphed into fields of their own. The German mathematician David Hilbert (1862–1943) envisioned the construction of the whole mathematical edifice not based on logic, as Frege, Whitehead and Russell would like. He rather preferred anchoring mathematics upon foundational conventions to which mathematicians would arrive after sorting out some basic issues that were marring, according to Hilbert, the solid web of deductive truisms characteristic of mathematical knowledge.

Accordingly, at the International Congress of Mathematicians of 1900, convoked in Paris to celebrate the turn of the century, Hilbert addressed the audience in an original manner. Instead of delivering a presentation that would contribute to the increase of mathematical

---

<sup>174</sup> Frege 1884, p. 6

knowledge, he produced a set of 23 mathematical problems that according to him would set the pace and provide the motivation for the development of mathematics in the following century. Only 10 of these problems were actually presented at the conference –the rest were made available later. Further clarification in the articulation of these problems evolved during the following three decades, and two of them (problems 2 and 10) begot what came later to be known as the “Decision problem”. Originally, the 2<sup>nd</sup> problem (named “The compatibility of the arithmetical axioms”) which asked whether the axioms of arithmetic were consistent, was stated in the form of a question, asking...

*Whether, in any way, certain statements of single axioms depend upon one another, and whether the axioms may not therefore contain certain parts in common, which must be isolated if one wishes to arrive at a system of axioms that shall be altogether independent of one another.*<sup>175</sup>

The 10<sup>th</sup> problem (entitled “Determination of the solvability of a Diophantine equation”) asked for a “procedure” that would allow for the discovery of whether an equation is true or false –at the time, Hilbert circumscribed the task only to “Diophantine equations”:

*Given a diophantine equation with any number of unknown quantities and with rational integral numerical coefficients: To devise a process according to which it can be determined by a finite number of operations whether the equation is solvable in rational integers.*<sup>176</sup>

In 1928 at Bologna, in another meeting of the Internal Congress of Mathematicians, Hilbert expanded the reach of both problems into a more general and relevant set of questions affecting the whole of mathematics. Although Hilbert presented them as open questions, he was of the idea that the answer to each one of them would be in the positive --this effort constituted what came to be known as “Hilbert’s Program”. These questions, each under the banner of their identifying issue, were thus:

---

<sup>175</sup> Hilbert 1902, p. 447 (Italics original)

<sup>176</sup> Ibid., p. 458 (Italics original)



- 1) Completeness: Are the true propositions of mathematics all provable?
- 2) Consistency: Is the set of true propositions in mathematics free of contradiction?
- 3) Decidability: Is there a procedure by which every true proposition could be provable?<sup>177</sup>

The Austrian mathematician Kurt Gödel (1906-1978) was in all likelihood present at this 1928 talk, which also served as a retirement speech for Hilbert. They never met in person, but two days later, at the same conference, during a round table, Gödel gave a presentation regarding certain aspects of the *Principia Mathematica*. By this time, Gödel had already reached his “First Incompleteness Theorem”. The threat to Hilbert’s goal was noticed by the previously mentioned mathematician John von Neumann, who took him aside to talk. The next year von Neumann communicated to Gödel that he followed up on his presentation and reached a second incompleteness theorem. By this time Gödel had already done so as well. In 1931, only one year after the completion of his Doctorate, at 25 years old, Kurt Gödel published “On Formally Undecidable Propositions of Principia Mathematica and Related Systems I”.<sup>178</sup> After this paper, quick and wide (but not absolute) consensus was reached in asserting that, at least in the way he framed Hilbert’s first and second new problems (completeness and consistency), Gödel showed that the answer was in fact in the negative. Or rather, that the very idea of wanting a solid system that is both complete and consistent, was ill formed.

---

<sup>177</sup> The text of the conference has been translated as “Problems of the grounding of mathematics” (Mancosu 1998, pp. 266-273). Stephen Hawking’s somewhat canonical abbreviation of the triad problem reads like this:

1. To prove that all true mathematical statements could be proven, that is, the *completeness* of mathematics.
2. To prove that only true mathematical statements could be proven, that is, the *consistency* of mathematics,
3. To prove the decidability of mathematics, that is, the existence of a *decision procedure* to decide the truth or falsity of any given mathematical proposition (Hawking 2005, p. 1121).

<sup>178</sup> Gödel 1931. There was supposed to be a follow up paper by Gödel, which never occurred –hence the “I” at the end of the title.

Gödel's "first theorem" asserted that if in a sufficiently robust arithmetic system the set of axioms is consistent, then there will be at least one true proposition that cannot be proved.<sup>179</sup> The "second theorem" showed that if a system is consistent, such consistency cannot be proved from within the system.<sup>180</sup> Arithmetic, and by extension mathematics, if understood as founded upon a strong body of propositions, is incomplete if it is found to be consistent: there will be at least one axiom that, even if true, will not be provable from within the system.

These findings were devastating for Hilbert --who never formally replied-- and by extension to Russell and Whitehead --whose logicist approach was partly shared by Hilbert for number theory, but was later rejected. In fact, at a certain point, Hilbert *was* in principle amenable to relying on *Principia Mathematica* to formalize arithmetic, since a mechanized way of proving the whole mathematical edifice would be close to his goal of making mathematics an orderly and tightly fit structure. The effect of the incompleteness theorems for mathematics in

---

<sup>179</sup> Kurt Gödel's article is famous not only due to his mathematical rigor (accessible only to the trained mathematician), but also due to the fact that its translation into English is an issue in itself. One could safely state that he never was fully satisfied with any of the several translations that took place throughout the years while he was still alive. In a translation by Elliott Mendelson, the one which he was the least unhappy with, Gödel's very technical text excerpt regarding what came to be known as his "first theorem" --in fact, the sixth theorem in the paper-- reads:

Theorem VI: For every  $\omega$ -consistent recursive class  $\kappa$  of FORMULAS, there exists a recursive CLASS EXPRESSION- $r$ , such that neither  $\nu$  Gen nor Neg ( $\nu$  Gen  $r$ ) belongs to Flg ( $\kappa$ ) (where  $\nu$  is the FREE VARIABLE of  $r$ ) (Gödel 1965, p. 24).

In layman terms, this formulation could be paraphrased as:

the first theorem... asserts for any formal theory... rich enough to include all the formulas of formalized elementary number theory... that if it is consistent... then it is incomplete (Stoll 1979, p. 446).

By all the formulas of elementary number theory we can understand the whole of arithmetic. It follows that:

the entirely natural idea that we can give a complete theory of basic arithmetic with a tidy set of axioms is wrong (Smith 2007, p. 13).

It would thus seem that if the set of all arithmetic propositions is complete, then it is inconsistent --and if consistent, then it is incomplete.

<sup>180</sup> The surprisingly clear textual reference in Gödel's paper appears as follows:

Theorem XI. Let  $\kappa$  be any recursive consistent class of FORMULAS; then the SENTENCE which asserts that  $\kappa$  is consistent is not  $\kappa$ -PROVABLE; in particular, the consistency of  $P$  is unprovable in  $P$ , assuming that  $P$  is consistent (in the contrary case, of course, every statement is provable) (Gödel 1931, p. 36).

Here Gödel would deduce a truism that follows from the previous (so-called "first") theorem. If there is a proposition that would prove the consistency of a set of propositions, then this proposition would lie outside the system of such set.

what regards its decidability was certainly foreseen by Gödel. However, as we will see, Hilbert's third problem took on a life of its own.

## **2. A machinal understanding of an algorithm and the material liberation of the machine.**

The Cambridge mathematician Max Newman had a student in his 1935 class, recipient of a fellowship for King's College, at the same institution. A young Alan Turing (1912-1954), after completing a dissertation in Gaussian theory for his undergraduate degree, joined professor Newman's course, which dealt with fundamental mathematical problems. Noticing that the "Decision problem" was so far deemed unsolved, Turing embarked on its solution. Professor Newman had in his lectures identified the quest for an "effective procedure" for deciding whether or not a mathematical equation is true. What an "effective procedure" stands for will show itself to be an issue soon.

The notion of "mechanical" was not foreign in mathematical logic, particularly among logicians. An entity "affecting" something implies a certain action, a change. A procedure that affects a change in a deterministic manner could be spoken of as being mechanical. Alan Turing was receptive to this qualification. Further equalizing the notion of a mechanical procedure to that of computability, he aimed to show in his paper that not the whole of mathematics is in fact computable –thus answering Gödel's third problem (whether there is an effective procedure to prove any true mathematical proposition) in the negative.

Just before publishing the paper, Turing (and Newman) found out that both Gödel and the American logician Alonzo Church (1903-1995) were separately working on the Decision problem at roughly the same time. In fact, Church had already published his findings in 1935.

Expectably, Newman was at first reticent in supporting Turing publishing his paper, but later he was convinced that the angle was sufficiently different, giving Turing the get go –which happened in early 1936. The name of this historic paper was “On Computable Numbers, with an Application to the Entscheidungsproblem”.<sup>181</sup>

Turing took the “mechanical procedure” seriously enough as to devise an abstract machine that would function *qua* algorithm –the very definition of which was being sought after by Gödel (via the notion of recursive function) and Church (by means of his own mathematical construction, the “lambda calculus”). Key notions in Turing’s imaginary engine were determinism and the “brute” following of instructions. The Turing machine (as it quickly began to be called) was constituted by an endless tape, a reading “head” (which can also write and erase), and a set of instructions. It works in an ingeniously simple way. The endless tape is divided in squares (just as the old film tapes), and each square can have written the symbols 0 or 1 in it; or it can be empty. The instructions will be constituted by a set of commands that will instruct the head what to do if it finds, say, a “1” (erase, write “0”, move to the right) or a “0” (erase, move to the left) or an empty square (write “0”, move to the right), or another “1” (move to the left), and so on.<sup>182</sup>

Turing was aware of what he was accomplishing. He provided, as Church put it later, an intuitive notion of computation that becomes immediately clear to the reader –as opposed to, say, the one given by his own lambda-calculus. Turing exhausted the realm of possibilities of what a

---

<sup>181</sup> Turing 1937

<sup>182</sup> This is how Turing described it:

The machine is supplied with a “tape” (the analogue of paper) running through it, and divided into sections (called “squares”) each capable of bearing a “symbol”... We may call this square the “scanned square”. The symbol on the scanned square may be called the “scanned symbol”. The “scanned symbol” is the only one of which the machine is, so to speak, “directly aware”... In some of the configurations in which the scanned square is blank (i.e. bears no symbol) the machine writes down a new symbol on the scanned square: in other configurations it erases the scanned symbol. The machine may also change the square which is being scanned, but only by shifting it one place to right or left (Turing 1937, p. 231).

computation could do—at least in this intuitive way—by this machine: “It is my contention that these operations include all those which are used in the computation of a number”.<sup>183</sup> Further, he also gave an algorithm (the set of instructions) the absolute power over its effected procedure: it would completely determine the behavior of the chosen parsed area of reality...

We may compare a man in the process of computing a real number to a machine which is only capable of a finite number of conditions  $q_1, q_2, \dots, q_R$  which will be called “m-configurations”...

The possible behaviour of the machine at any moment is determined by the m-configuration  $q_n$  and the scanned symbol  $S(r)$ . This pair  $q_n, S(r)$  will be called the “configuration”: thus the configuration determines the possible behaviour of the machine.<sup>184</sup>

To recall, the main challenge was to find out whether there was an effective procedure by which one could demonstrate, in finite time, whether a given mathematical proposition is provable from a finite set of axioms. It soon became evident that the notion of an “effective procedure” should be defined --or at least clarified. As mentioned above, Alonzo Church identified an effective procedure with an “algorithm”, thus making inroads towards its definition. And we have seen how Turing’s insights into the nature of an algorithm, within the broader aim of answering Hilbert’s challenge, coincidentally encountered the parallel development of Church’s work. As also mentioned, Church beat Turing to the punch in publishing similar findings some months before the latter was given the *placet* by his supervisor, Max Newman, for submitting his own piece in 1936 –and published the next year. However, by recommendation of his own supervisor, Turing was by the end of the same year at Princeton University, pursuing graduate studies under Alonzo Church himself.

---

<sup>183</sup> Ibid., p. 232

<sup>184</sup> Ibid., p. 231. The reader may have noticed the introduction of the notion of “man” for the first time, alluding to ideas of intelligence and behavior. The issue of whether or not a machine can be intelligent (*contra* Turing at this point) was to emerge still sometime in the future. Let us recall that computers as we know them did not yet exist at this time. These issue will however come back below –sections 3 & 4 of this chapter.

Two years were enough for Turing to obtain his Ph.D., with a dissertation entitled “Systems of Logic Based on Ordinals”.<sup>185</sup> This monograph, often historically overlooked, built upon the idea of the noncomputability of certain realms of mathematics. Within this framework, the thesis left important cues regarding Turing’s own later theoretical evolution. More relevantly to the issue that concerns the present work, Turing’s dissertation contains seminal insights into what he would consider to be a machine. In this thesis, in an excerpt that could be taken as joint statement regarding the equivalence between the definitions of computability reached by Gödel, Church and himself, Turing wrote:

A function is said to be "effectively calculable" if its values can be found by some purely mechanical process. Although it is fairly easy to get an intuitive grasp of this idea, it is nevertheless desirable to have some more definite, mathematically expressible definition. Such a definition was first given by Gödel at Princeton in 1934... These functions were described as "general recursive" by Gödel. We shall not be much concerned here with this particular definition. Another definition of effective calculability has been given by Church..., who identifies it with lambda-definability. The author has recently suggested a definition corresponding more closely to the intuitive idea...<sup>186</sup>

Turing then puts emphasis on the identification of effective calculability with a mechanical process, equation that permits the other two definitions (Gödel’s and Church’s) to also emerge. Here Turing came close to identifying what a machine might stand for,<sup>187</sup> doing it in a reverse manner: if a process is said to be “purely mechanical” then one is to expect that there is a

---

<sup>185</sup> Turing 1939

<sup>186</sup> Turing 1939, p. 166

<sup>187</sup> Close, but not quite there. Hodges acknowledges that the “central thrust of Turing’s thought was that the action of any machine would indeed be captured by classical computation”, and that is the closest that Turing gets to define a machine (Hodges 2008, p. 88). Hodges however reduces Turing’s probable meaning of “machine” to “Turing machine” (Hodges 2008, p. 77).

machine carrying that process. “Purely mechanistic” implies a machine doing the job.<sup>188</sup> Turing’s thesis states that

“a function is effectively calculable if its values can be found by some purely mechanical process”. We may take this statement literally, understanding by a purely mechanical process one which could be carried out by a machine.<sup>189</sup>

This machine, bounded by its own nature, exhaustively generates something germane to its own being: a mechanical process. In fact, the analytical relation between a concept and an idea already contained in the concept (e.g., a triangle has three sides) is not foreign to this relation. A mechanical procedure is already contained in the idea of a machine. The machine, emanating what is to be understood as mechanical *per se*, completely determines what the mechanical procedure will be.

The suggestion of having a model of machines --a paradigm of what can exhaustively perform a mechanical process—where its physicality is irrelevant, did not go unnoticed. Indeed the “purely mechanical process” here is instantiated by an algorithm<sup>190</sup> (a set of commands) – “a machine”. What makes a machine ultimately a machine, even if it is not yet defined, is just tangentially related to its possible physicality. The latter is not referred to as a *sine qua non*

---

<sup>188</sup> When Turing talks of a machine, right from his 1937 paper onwards, he is implicitly talking about an ‘automatic’ machine –as opposed to a ‘choice’ machine (a machine whose instructions and outcome leave room for human tweaking) or an ‘oracle’ machine (a fabled machine which processes uncomputable realms). Each would be named a-machine, c-machine and o-machine respectively. He stated that when he will refer to a “machine” he will always be referring to an a-machine. So he conventionally dropped the apposition “a-” and begun to call them simply “machines” (Turing 1937, p. 232).

<sup>189</sup> The quote continues, making explicit the reliance on the articulation of an effective calculability (computability) by means of making it amenable to being instantiated as a mechanical procedure:

It is possible to give a mathematical description, in a certain normal form, of the structures of these machines. The development of these ideas leads to the author's definition of a computable function, and to an identification of computability with effective calculability. It is not difficult, though somewhat laborious, to prove that these three definitions are equivalent (Turing 1939, p. 166).

<sup>190</sup> “Turing’s work gives an analysis of the concept of “mechanical procedure” (alias “algorithm” or “computation procedure” or “finite combinatorial procedure”).” (Gödel 1934, p. 72)

condition pertaining to its nature. Instead, what makes a machine to be itself is its inherent capability of producing a determinate and effective procedure, a “purely mechanical” outcome.

There had certainly been in history several metaphysical appreciations of mechanical entities. The Scientific Revolution emphasized mechanisms, beyond experimentation, as a sure-proof way for building up grounded theoretical edifices.<sup>191</sup> Rene Descartes embodied this lemma with an intensity that is often overshadowed by his legacy of “substance dualism”. A less often noticed side of his philosophy was his faithful reliance on a literally machinal view of organisms. Historically recalling, Descartes divided reality in two qualitatively distinct substances: *res extensa* and *res cogitans*. The first one referred to things that had materiality (physical extension), e.g., rocks, plants, animals, human bodies. The second one referred to a substance, equally existent and real as the first one, but which lacked physicality: minds, angels and God.<sup>192</sup> The theological benefits of this separation have been often pointed out as a purported move to “protect the soul”, given the advances in the new sciences. But another alleged advantage gained from this was the possibility of denying outright any animal participation of soul properties, avoiding the ontological hassle of distributing watered down versions of the logos throughout the animal kingdom (from primates all the way down to insects and beyond). In this way the special place of man in the universe, as image and likeness of God, gets removed from an otherwise dangerous situation of being superior only as a matter of degree, and not in kind.

---

<sup>191</sup> Paradigmatic in this reliance upon a mechanical element of knowledge is Francis Bacon’s praising of the “mechanical arts,” collapsing the Greek legacy of a distinction between practical and theoretical knowledge. Wisdom shall now be merged with mechanical mastery, culminating in Bacon’s *experimentum crucis*. Indeed this distinction (particularly the Hellenic frowning upon τέχνη) is what allegedly crippled the development of science --as Bacon understood it-- among the Greeks...

Whereas in the mechanical arts, which are founded on nature and the light of experience, we see the contrary happen, for these (as long as they are popular) are continually thriving and growing, as having in them a breath of life, at first rude, then convenient, afterwards adorned, and at all times advancing (Bacon 1620, § 74)

<sup>192</sup> “I deny that true extension, as it is usually understood by everyone, is found either in God or in angels or in our mind, or, in short, in any substance that is not a body” (Descartes 1649, p. 293).



Another alleged benefit was the continuation of the tradition of guiltlessly practicing vivisection –following the Greek physician Galen and Leonardo Da Vinci, who reported important insights into physiology as an outcome.<sup>193</sup>

The soul is far removed from this picture in its quality of *res cogitans*, and thus, safe. However both animals and human bodies pertain to the realm of an organic *res extensa*, and thus, they are totally subjected to the deterministic laws of mechanics, which could now be pursued. On his *Treatise of Man*, Descartes refers to the intricacies of the human body, with all its vegetative and non-vegetative functions, in the following terms:

I desire, I say, that you should consider that these functions in the machine naturally proceed from the mere arrangement of its organs, neither more nor less than do the movements of a clock, or other automaton, from that of its weights and its wheels; so that, so far as these are concerned, it is not necessary to conceive any other vegetative or sensitive soul, nor any other principle of motion or of life, than the blood and the spirits agitated by the fire which burns continually in the heart, and which is no wise essentially different from all the fires which exist in inanimate bodies.<sup>194</sup>

---

<sup>193</sup> The fact that these were advances coming from Protestant realms was probably also part of his motivation. There is a received view that Protestant colleges were putting emphasis on the development and teaching of the sciences and “liberal arts” and trades, whereas the more traditional Catholic universities were still emphasizing the teaching of ancient languages, philosophy, theology and the arts. The Church allegedly preferred a unified Europe under its tutelage, uniformly communicating via Latin. Descartes’ academic formation however came from a school set up in the context of the Counter-Reformation: the Jesuits instilled in him a deep admiration for mathematics and sciences, while studying at the Collège Royal Henry Le Grand at La Flèche. Descartes’ Catholicism, although not particularly orthodox, was likely profound. The abdication of Queen Christina of Sweden, his pupil, in order to become a Catholic, is often brought up as an instance of Descartes’ Catholicism all the way to his death (Swedish law, as in today’s United Kingdom, prohibited a Catholic to become a monarch). For a recent study on Descartes’ intellectual life, including the context of his education and religious beliefs, see Gaukroger 1995. For an account of the educational context of Europe after the Reformation, see Eisenstein 2012, ch. 7.

<sup>194</sup> Descartes enumerates in the previous paragraph what functions he is referring to. “All the functions which I have attributed to this machine (the body), as the digestion of food, the pulsation of the heart and of the arteries; the nutrition and the growth of the limbs; respiration, wakefulness, and sleep; the reception of light, sounds, odours, flavours, heat, and such like qualities, in the organs of the external senses; the impression of the ideas of these in the organ of common sensation and in the imagination; . . .the retention or the impression of these ideas on the memory; the internal movements of the appetites and the passions; and lastly the external movements of all the limbs, which follow so aptly, as well the action of the objects which are presented to the senses, as the impressions which meet in the memory, that they imitate as nearly as possible those of a real man.” (Descartes 1662, § 202)

As one can readily notice, what is relentlessly mechanized is still the realm of material entities – *res extensa*. The demystification that is taking place is circumscribed by stripping the organic area of *res extensa* from any purported vitalism inherent to the phenomenon of “life”. Clocks are as much automata as animals and human bodies are. But the mind, inhabiting the realm of *res cogitans*, is still perfectly safe –anticipating Kant’s anxiety regarding human freedom for almost 150 years.<sup>195</sup> In fact, the mechanization of something other than *res extensa* would probably not have made much sense for Descartes.<sup>196</sup>

In contrast, with Alan Turing we have the notion of a non-physical entity as a certain type of machine --a non-material one. This never happened before. It had not happened because the notion of a machine had not yet been developed to a point at which it would give up its physicality as a non-essential feature. Although the horizon of this insight might have been already simmering for a while, it was not made explicit until the above addressed evolution of logic and mathematics took place. That development was the articulation of the notion of an algorithm --explicitation deemed necessary for fully clarifying the basic notion of an effective computation. As previously mentioned, this definition was in turn necessary for addressing Hilbert’s dramatic question regarding the possibility of producing an algorithm that would solve any mathematical problem within a finite time and within a finite set of axioms. Both Turing and Church attempted to define such a notion (of an algorithm), and the isomorphic correspondence to a familiar object in the physical world was soon evinced. That object, to rehash, was a machine. More specifically, a physical machine is but one instantiation of what a machine is *per*

---

<sup>195</sup> Descartes originally wrote his *Treatise on Man* in 1633, but it was posthumously published in 1662. Kant wrote his first two *Critiques* in 1781 and 1789, where he laid out his grand project of saving freedom by restricting the realm of a rightfully mechanistic modern science to the *phenomenal* –sparing the *noumenal* realm for our exercise of free will.

<sup>196</sup> Somewhat anachronistically Michael Wheeler speculates that Descartes could have expanded his notion of the mechanical to the non-physical (*res cogitans*), but that his view of a machine was, due to his 17<sup>th</sup> century context, somewhat crude (Wheeler 2008). The mechanization of the whole of man (not only of his body), a related but different issue --where Julien Offray De La Mettrie (1709 –1751) would be relevant-- lies ahead in the near future.

*se*, since physicality is *not* a *sine qua non* feature of a machine *qua* machine. With this new understanding of a machine, devoid of a heretofore deemed necessary physicality, previous realms hitherto beyond the scope of a rigorously mechanistic explanation were suddenly amenable to being encompassed. Such as life and mind.

Something remarkable –over and beyond the response to an admittedly important mathematical problem-- was thus accomplished in Turing’s “On Computable Numbers”. Although no definition of machine was given in the article (or in any other writing by Turing for that matter),<sup>197</sup> the seeds of an insight that would prove powerful for subsequent scientific, technological and philosophical endeavors in decades to come were already there. Turing stopped short of defining a machine, but the rigorous explicitation of what stands for a “mechanical procedure” demanded from him the abstract construction of a machine that would do precisely that. This abstract machine, capable of exhaustively (and effectively) performing any possible computation (as construed by Turing), became a paradigmatic machine. An ideal machine, model of all machines. Again, nothing had so far put forward the notion of a non-physical machine. Hence this observation was rather quickly noticed by some scientists, mathematicians, psychologists and thinkers alike, expanding their foreseeable theoretical exercises and scientific playgrounds. Here we come to a realization of something that, although probably still unaware of its full impact, might have changed the course of recent history. This is one of the pillars upon which a scientific movement that lasted a decade (1942-1952) was built. Turing’s insight into the nature of an algorithm, an essential feature of which being that it maintains an isomorphic relation with an entity in the physical world --namely, a machine-- can

---

<sup>197</sup> Turing might have been close to disclose a full-fledge theory of machines; but he never did. See the previous section (2).

be regarded as the theoretical foundational bedrock upon which this was devised as a general scientific theory. The name of this scientific outlook went under the name of *cybernetics*.

### **3. Alan Turing's strange reversal regarding the question of Artificial Intelligence.**

We have seen how the need for an explicitation of the notion of an algorithm was a pre-condition for defining the idea of a computation, with the goal of addressing Hilbert's problem. We have also seen how this task opened up insights into the nature of a machine, first captured, although not fully worked out, by Alan Turing in his "On Computable Numbers". There is another realm, however, in which cybernetics and Turing might have been deeply theoretically intertwined as well. The question of Turing's reversal regarding the possibility of a machine possessing intelligence has been tangentially addressed, but the puzzle regarding this dramatic turn seems to remain. The puzzle would involve finding out what went through Turing's mind to make him change his view on the philosophically important issue of Artificial Intelligence.

The development of cybernetics might shed some light on what went through Turing's mind for the span of a decade, probably pointing to the reason why he evolved into an early advocate of AI, epitomized in his 1950 article "Computing Machinery and Intelligence".<sup>198</sup> At some point between 1936 and 1946, Turing's views regarding what a machine could do underwent a strange reversal. In fact, the contrast between what Turing would say regarding the possibility of machine intelligence<sup>199</sup> in 1936, and then almost 15 years later is startling. Let us

---

<sup>198</sup> Turing 1950

<sup>199</sup> "Machine intelligence", in this context, refers to straightforward human intelligence inhabiting an artificial machine --as we will see below. The nature and evolution of the notion of intelligence (which allowed its usage in, for example, contemporary electronic engineering) is beyond the scope of this work.

return to the interesting quote of the 1936 article:<sup>200</sup>

We may compare a man in the process of computing a real number to a machine which is only capable of a finite number of conditions  $q_1, q_2, \dots, q_R$  which will be called “m-configurations”... The possible behaviour of the machine at any moment is determined by the m-configuration  $q_n$  and the scanned symbol  $S(r)$ . This pair  $q_n, S(r)$  will be called the “configuration”: thus the configuration determines the possible behaviour of the machine.<sup>201</sup>

The reader will surely note the quiet comparison between a man and a machine. This analogy has triggered some confusion in the past. Let us be reminded that the existence of computers, as we know them now, still lay somewhat far in the future, at least one decade later. At the time, however, the task of computing numbers already existed –in fact, it always did.<sup>202</sup> The task of crunching numbers with a pen and paper was part of the idiosyncratic picture of available jobs at the time (indeed usually held by women). This job, infamous for its tediousness, required the blind following of a rule, absolutely and without deviation. In this vein, Wittgenstein has a relevant remark that reeks of mysticism only if taken without the above context: “Turing’s ‘Machines’. These machines are *humans* who calculate”.<sup>203</sup> The *machinic*<sup>204</sup> character of jobs of such nature (e.g. the endless and monotonous repetition of a small task or movement), where “mechanical” and “mindless” were referred to as synonymous, was a common theme in the

---

<sup>200</sup> Reproduced above, p. 94.

<sup>201</sup> Turing 1937, p. 231

<sup>202</sup> There is no universally agreed definition of computation. The intuitive depiction of it by Turing let it gain traction in what regards the future instantiations of computing machines. But certainly there are other ways to set a computational process. That raises the question whether the existence of “man” is a necessary prerequisite for a computation to take place at all –someone has to perform the computation, e.g., make up the set of instructions and run it. After the arrival of the “genetic revolution” some would contend that computation occurs in nature itself – after all, it is a “mindless” procedure. Some would suspect that it may even subsume the entirety of physics into a realm too immense for us to grasp at this time. For an attempt to lay out a panoramic vista of theory of computation in our days –friendly to the view just mentioned – see Smith 2002.

<sup>203</sup> Wittgenstein 1980, §1096

<sup>204</sup> The term –at least as applied to this context-- is John Johnston’s (Johnston 2008).

popular culture of the time.<sup>205</sup> In fact, the media began to refer to computing machines as “computors” in order to differentiate them from “computers”, the latter being humans who had the hard and boring job of building up number tables. As machines gradually (and literally) took over the job, they were eventually referred to as “computers”.<sup>206</sup> Turing’s 1937 paper, therefore, should not be regarded yet as a precursor to what later came to be known as “Artificial Intelligence”. Not only were there no computers (in the modern sense) at the time, with which humans could be compared, but the very notion of “mechanical” carried an epithet of “mindlessness” inherently attached to it. This pejorative adjective ascribed to a machine was accepted and defended by Turing himself –which would make the notion of “machine intelligence” an oxymoron, and by extension “artificial intelligence” senseless.

Further, the paper had as its original aim answering Hilbert’s Decision problem, and since this answer was given in the negative, the paper showed that some mathematical propositions are in fact not computable –hence the answer *must* be in the negative. Some incomputable realms of mathematics *a fortiori* cannot be treated by a machine. This limitation as to what a mechanical procedure can do puts a machine further away from a possible comparison with the human mind, canonically agreed to be capable, by Turing’s own standards, of dealing with non-mechanical realms, as it is the case with “intuition” –indeed a familiar occurrence reported among mathematicians like him while engaged in problem solving. Turing’s change of mind, giving a machine a qualitatively new kind of cognitive powers, still lay ahead in the future.

After Turing had defended his dissertation, he was offered a position at Princeton, but he refused it in order to go back to Britain. The shadow of war was already looming and it is quite

---

<sup>205</sup> Tellingly, Charles Chaplin’s iconic film “Modern Times” was released in 1936. A vestigial sentiment of repulsion triggered by the dehumanizing social effects of the Industrial Revolution might have been at play.

<sup>206</sup> Thus Wittgenstein’s remark could be transduced as “Computers are computers who calculate”, or simply, computers do the mechanically mindless job of a human, faster.

possible that he was moved by patriotic considerations. From 1939 (year in which his thesis was published) to 1945, Turing was submerged in war-related tasks –the secrecy of which partly explaining a publishing hiatus. By 1945 the design for an early electronic computer, the Automatic Computing Engine (ACE) was finished, a task that demonstrated, beyond his mathematical brilliance, also his engineering skills.

Tinkering with machines was not an altogether new experience for Turing –it is reported that he had an uncanny attraction towards machinery since his childhood (for example, with the nature of the typewriter). During the span of this decade, something occurred in Turing’s mind regarding the status of machines, the computable and intelligence. In 1946, in a report on the ACE, he suddenly (and mellifluously) introduced for the first time the idea of “mechanical intelligence” –a notion regarded at the time as radically senseless. But he did it in such a tangential, almost cryptic way, that it seemingly did not raise many eyebrows (which could be partly understood given the ambience of secrecy in that era)...

We stated... that the machine should be treated as entirely without intelligence. There are indications however that it is possible to make the machine display intelligence at the risk of its making occasional serious mistakes.<sup>207</sup>

Advancing an insightful premise, it was suggested that making mistakes could be a sign of intelligence. Moreover, one could certainly program a computer in a way that it could “make the machine display intelligence”.<sup>208</sup> Turing did not elaborate much on the bone thrown there. Why would mistake-making be a reproducible sign of intelligence? In 1947, in a presentation for the London Mathematical Society, Turing expanded on the bizarre idea of “machine intelligence”, again touching the issue of what mistake-committing might entail for a machine. Connecting

---

<sup>207</sup> Turing 1946, p. 16

<sup>208</sup> Ibid.

with previous work done in the foundations of mathematics, he described what the outcome is for both a machine and a mathematician dealing with paradoxes encountered in logic and arithmetic. By the very nature of mathematics, according to the recent discoveries by Gödel, Church and himself, there cannot be an algorithm that would find out whether *any* mathematical proposition is true or false --indirectly answering Hilbert's problem. That entails that a machine will eventually stop in its attempt to prove a certain equation, in a particular system --one in which there will be at least one true proposition that is not provable...

It has for instance been shown that with certain logical systems there can be no machine which will distinguish provable formulae of the system from unprovable, i.e. that there is no test that the machine can apply which will divide propositions with certainty into these two classes. Thus if a machine is made for this purpose it must in some cases fail to give an answer. On the other hand if a mathematician is confronted with such a problem he would search around and find new methods of proof, so that he ought eventually to be able to reach a decision about any given formula.<sup>209</sup>

Given the undecidability of mathematics (even given the non-facile nature of mathematics, period), humans commit errors in looking for solutions to complex problems. But we still call them "intelligent". The "intelligent" mathematician will look for other ways, attempt other strategies, in order to continue the task. However, when it comes down to machines, the pre-conceived notion exists that "if a machine is expected to be infallible, it cannot also be intelligent".<sup>210</sup> Turing denounces that this is "unfair play" to the machine.

I would say that fair play must be given to the machine. Instead of it sometimes giving no

---

<sup>209</sup> Turing 1947, p. 497

<sup>210</sup> Ibid. We will see that the emanation of something as exact as *mathematics*, from something as messy as the *mind*, disturbed other thinkers. Superimposing the former -- amenable of instantiation by a machine--upon the latter, in order to explain mental behavior, will have far-reaching implications. See the case of Warren McCulloch in chapter II, section 3 and chapter IV, section 3 of this work.



answer we could arrange that it gives occasional wrong answers. But the human mathematician would likewise make blunders when trying out new techniques. It is easy for us to regard these blunders as not counting and give him another chance, but the machine would probably be allowed no mercy.<sup>211</sup>

If, somewhat ironically, the mistake-committing human is regarded as intelligent, whereas the infallible machine is not, then let us make the machine become fallible. A machine that commits mistakes will then display a feature so far reserved only for humans --hence Turing's 1946 remark that "it is possible to make the machine display intelligence at the risk of its making occasional serious mistakes".<sup>212</sup> This allowance for mechanical mistakes opens up another realm of theorizing about the possible intelligence of a mechanical entity. After technology advances a fair bit (Turing wonders about a future with machines bearing a substantially expanded "storage capacity") nothing would in principle preclude the machine from changing its own set of instructions. Therefore, the machine would be allowed to commit errors so that it could be able to reorganize itself, even if for this "the machine must be allowed to have contact with human beings in order that it may adapt itself to their standards".<sup>213</sup> This suggestion, namely, that of a machine self-changing its internal parameters, carves up still another area of comparison between a machine and a human being --this time with a child, in the context of the infant's learning process:

Let us suppose we have set up a machine with certain initial instruction tables, so constructed that these tables might on occasion, if good reason arose, modify those tables. One can imagine that after the machine had been operating for some time, the instructions would have altered out of all recognition, but nevertheless still be such that one would have to admit that the machine was still

---

<sup>211</sup> Turing 1947, p. 497

<sup>212</sup> Turing 1946, p. 16

<sup>213</sup> Turing 1947, p. 497

doing very worthwhile calculations. Possibly it might still be getting results of the type desired when the machine was first set up, but in a much more efficient manner. In such a case one would have to admit that the progress of the machine had not been foreseen when its original instructions were put in. It would be like a pupil who had learnt much from his master, but had added much more by his own work. When this happens I feel that one is obliged to regard the machine as showing intelligence.<sup>214</sup>

The picture of a creature surprising its creator with its behavior has been a common Golem-type theme among humanists worried about the possibility of “playing God”.<sup>215</sup> Self-modifying algorithms could in principle, eventually, remove from the machine the stigma of being a mere instruction-following brute. However, this would be just a first step towards the mechanical appropriation of the notion of “intelligence”. Turing was aware of this, and he went on to address the notion of “learning”, revealing quite evident behavioristic cues.<sup>216</sup> Thus in his 1948 report “Intelligent Machinery”, the iconoclastic comparison of a “learning machine” with a “learning child” was further elaborated.

Firmly entrenched into the grand-idea of a Mechanistic Continuum<sup>217</sup> underpinning much of modern science and philosophy, Turing was well aware of the foundational experiments in behaviorism performed by the Russian physiologist Ivan Pavlov (1849-1936) and his followers. The phenomenon of “learning”, when broken down into pieces, reveals a “mechanism” (as anything amenable of scientific treatment, for that matter): a building and reinforcement of neural pathways, conditioned by external stimuli. This mechanism can be ascribed across the natural spectrum to the lower forms of life, as well as to the higher vertebrates.<sup>218</sup> Any spiritual,

---

<sup>214</sup> Ibid., p. 496

<sup>215</sup> This is a main theme in Wiener 1964.

<sup>216</sup> See Shanker 1995.

<sup>217</sup> Ibid., p. 53

<sup>218</sup> Ibid.

irreducible, epiphenomenal understanding of ‘learning’ that recoils against this ‘empirical’ explanation of what actually occurs behind the phenomenon (namely, its ‘mechanics’), would merely be the death rattle of a metaphysically-ridden Medieval understanding of man and nature.

In that vein, for Turing a child’s cortex “is an unorganized machine, which can be organized by suitable interfering training”.<sup>219</sup> The training of the child, the “teaching”, is not different in principle than those occurring in Pavlov’s dogs, laboratory worms, and “evolving” machines. Learning pathways get created, and then reinforced, when associations between a certain state and a stimulus of pleasure (or pain) are repeatedly allowed to occur. The “unorganized machine” with the capacity of altering its own instructions (with a storage capacity that would provide enough “memory”), would be equally suitable for showing off real learning, after a period of productive interaction with a (probably human) master who has delivered the appropriate punishments and rewards. Three years later, in a talk similarly entitled as the report, albeit with a bolder twist --“Intelligent Machinery, A Heretical Theory”—Turing was more specific regarding the behavioral training of this “child machine”:

I suggest that there should be two keys which can be manipulated by the schoolmaster, and which can represent the ideas of pleasure and pain. At later stages in education the machine would recognize certain other conditions as desirable owing to their having been constantly associated in the past with pleasure, and likewise certain others as undesirable.<sup>220</sup>

The next year Turing began to write the seminal AI paper “Computing Machinery and Intelligence” (to appear in 1950 in *Mind*, the only philosophy journal where he published throughout his career --a testament for how his ideas got wide circulation, attracting the attention of the philosophical community). This paper has been one of the most quoted philosophy papers

---

<sup>219</sup> Turing 1948, p. 424

<sup>220</sup> Turing 1996, p. 259. Although the talk was given in 1951, it was published posthumously in 1996.

from the second part of the 20<sup>th</sup> century on, so I will not dwell on it here. Suffice it to say that to a great extent this paper put together in an organic manner the ideas put forward by Turing since 1936 –but especially since after the war, where his commitment to a total mechanization of intelligence was clear. For instance, the famed “Turing Test”, subject of much philosophical debate up to these days, can already be foreseen in the 1946 paper (the ACE report). There Turing writes about paying the price for “seeing” intelligence in the machine by allowing it to have errors, suggesting that “it is possible to make the machine display intelligence at the risk of its making occasional serious mistakes”.<sup>221</sup>

We are here in front of a full 360 degree turn in what concerns the possibility of a machine being intelligent --from its denial, based upon the machinal incapacity of dealing with intuition, to its defense, based upon the possibility of a machine displaying human features, such as mistake-committing and learning. Since arguably much of contemporary cognitive science pivots upon the notion of “artificial intelligence” as articulated in this 1950 paper, the question still remains regarding Turing’s change of mind. What made him go from the position where mechanical procedures merely were those underlying the accomplishment of tasks by instruction-following brutes, to the position where these very mechanical processes could expose true intelligence in those bearing them? One can rephrase the question factoring in the historical framework. What went on with Turing behind his gradual dismissal of the previously vigorously defended<sup>222</sup> “oracle-machine” --the fabled machine that could compute the uncomputable? This first position was changed later on, where not only the discrete areas of thought dealing with mathematical problems, but the whole human intelligence, was to be understood as being essentially computational in nature. Oracle machines were no longer necessary to account for

---

<sup>221</sup> Turing 1946, p. 16

<sup>222</sup> In his doctoral dissertation (Turing 1939).

those areas that cannot be subsumed to a machine, since everything was machinal in one form or another, e.g., even via subconscious computations. Now, intelligence could be one day artificially constructed.

#### **4. Cybernetics as a possible missing link regarding Turing's change of heart.**

There are those who find reasons of a personal character regarding Turing's change of heart. Alan Turing had an early love in his teenage years who died prematurely of tuberculosis. This death haunted him for the rest of his life. It has been speculated that a subconscious wish of bringing back the beloved one (or making a beloved one never die) was behind the relentless desire for making the whole of intelligence mechanizable –and thus, constructible and retrievable. Turing handwrote in 1932 a short essay on the “Nature of the Spirit” where he wonders after his friend's death whether a detached spirit can re-inhabit another adequate vessel:

... as regards the actual connection between spirit and body I consider that the body by reason of being a living body can ‘attract’ and hold on to a ‘spirit’, whilst the body is alive and awake the two are firmly connected. When the body is asleep I cannot guess what happens but when the body dies the ‘mechanism’ of the body, holding the spirit is gone and the spirit finds a new body sooner or later perhaps immediately.<sup>223</sup>

There seems to be the hope that a mind could exist without its biological brain, probably amenable to be re-instantiated once a proper holding mechanism is achieved.<sup>224</sup> This would not seem too farfetched from Turing's mechanistic views if we understand both realms, the material and immaterial, as being equally machinal –and hence, with a nature determined not by physicality, but by mechanicity. Turing's later strong correlation between the notion of an

---

<sup>223</sup> Hodges 1983, p. 63

<sup>224</sup> Jean Lassègue proposes that the later Turing Test had this psychoanalytical motivation (Lassègue 1996).

algorithm and that of a machine might have been a cornerstone towards a further enrichment of his notion of a machine –culminating in the oracle machine (below) where determinate-behavior as an inherent feature of a machine is superseded.<sup>225</sup>

There are others, more lenient towards a “Science and Technology Studies” view, that give a prime role to Turing’s confrontation with engineering feats during the war, which allegedly affected his subsequent views. Turing saw first-hand the power of machines, designed and constructed by him to break the German code of the sinister U-Boats.<sup>226</sup> For Winston Churchill, this accomplishment greatly contributed towards Great Britain’s victory,<sup>227</sup> acknowledging Turing’s essential contribution towards the Empire’s victory<sup>228</sup> –a victory that might not have happened if the calculations would have been performed solely by humans. Thus, Turing witnessed the rapid improvement in the construction of these all-important machines, going from entirely re-wiring the machine to vacuum tubes and from there to relays –one design material making the machine faster and smaller than the next one. What could the future hold in 50 years from then? That is the question that Turing poses towards the end of his 1950 paper, and he is of the idea that intelligence will be mechanical.<sup>229</sup> And this would not be our intelligence, but the machine’s: it would be *artificial intelligence*.

---

<sup>225</sup> Andrew Hodges disagrees (Hodges 2000). “On Computable Numbers” marks for Hodges the end of any dualistic remnant in Turing’s mind, committing the mathematician for good to a deterministic and materialistic stance. It is relevant to remark that the case could be made that Turing’s 1) later oracle-machine; and 2) insights into the machinal as anchored in mechanicity and not materiality (as depicted in the 1937 article and his 1939 dissertation) could attenuate Hodges’ objections.

<sup>226</sup> Hodges shares this view –although he does not particularly articulate the alleged link between being impressed by powerful machines and the possibility of machine intelligence (Hodges 2000).

<sup>227</sup> Turing and his fellow co-workers wrote a conjoint letter directly to Churchill, where they vented their frustrations due to the slowly bureaucratic pace at which their logistic demands were met. Upon receiving the letter, Churchill wrote a terse and short note to his Chief of Staff: “ACTION THIS DAY. Make sure they have all they want on extreme priority and report to me that this had been done.” (Copeland 2004, pp. 336-340).

<sup>228</sup> “According to Winston Churchill, Turing made the single biggest contribution to Allied victory in the war against Nazi Germany.” (Spencer 2009).

<sup>229</sup> Turing 1950, pp. 455-456

Still others, more philosophically oriented, in a view not altogether unrelated to the previous one, see in Turing's change a gradual but firm realization of his own place in the grand-scheme of the above mentioned Mechanistic Continuum. If he was to bring his own insights -- already firmly set in such a tradition-- to their rightful completion, then thinking machines would necessarily follow. Learning, or the acquisition of knowledge, as indicated above, can be broken down into pieces that are so small that they can be found neatly interacting with the overall machinery of survival, even in the simplest organisms. Relevant connections for survival are forged and reinforced via trial and error, the environment producing stimuli in the form of reward or punishment. Certainly the learning process of a worm and of a classical composer can be regarded as vastly different, but this difference would not be one of kind, but of degree. The latter is greatly more *complex*, but the principle of productive change upon which both rest is largely the same.<sup>230</sup> This realization might have further laid out in Turing the foundation for the hope of an artificial intelligence to emerge.

All these interesting reasons notwithstanding, there nevertheless is a venue of possible explanation for Turing's reversal that has been less explored --but which popped out as a recognizable clarifying scheme during this writer's research. This explanatory path goes through the historical cross-roads that one could identify within the decade mentioned above (1936-1946). Between Turing's insights characterizing a machine as being inherently incapable of intelligence, and the now ubiquitous AI-friendly classical computationalism stemming from Turing's later position, there is a missing link. To explore this suggestion, we have to go back to

---

<sup>230</sup> "Herein lies the thrust of the behaviourist version of the continuum *picture*... The crucial point is the idea that *learning* consists in the formation of stimulus-response connections which themselves require no intelligence... The 'higher' forms of learning... are distinguished from these lower forms by the complexity of the stimulus-response connections forged in the organism's brain. But the mechanical nature of the atomic associations which form the basis for all levels of learning remain identical. This provides the rationale for describing what had hitherto been regarded as disparate phenomena... as constituting a *learning continuum* ranging from simple negative adaptation, habituation, accommodation, and tropisms, through animal and infant learning, to the highest reaches of education and scholarship" (Shanker 1995, p. 63).

the time frame occurring between Turing's 1937 paper and his subsequent thesis dissertation under Alonzo Church. Laying out this venue of explanation, which emerged in the midst of my own investigation into the nature of a machine, will hopefully open up the context for introducing a theory in the history of science that revolved around the notion of a machine *tout court*—cybernetics.

Turing's paper "On Computable Numbers", as previously mentioned, contains the seed of a radical view of a machine, which seized the attention of a handful of scientists—who were soon to engage in the cybernetic enterprise. Turing, to recall, further elaborated the notion of a machine for his doctoral dissertation under Alonzo Church. However, instead of staying in America, he chose to return to England, likely motivated to help his country *vis à vis* the looming war on Europe.<sup>231</sup> Once back in England, Turing was a regular attendee at the British version of the Macy Conferences—the much smaller and informal Ratio Club.<sup>232</sup> Both gatherings, Macy's and Ratio's, exchanged some speakers—mostly coming from the US to the UK. That's how Wiener and McCulloch showed up in England on occasion. Turing, later on, while engaged in the war effort, would go to the US to share knowledge regarding the construction of the novel computing machines. Even if he did not ever attend a Macy meeting, he was in constant contact with the cyberneticians—especially with von Neumann, due to the common task of building the first computers-- who would act as hosts during his stay in America.

In other words, although Turing left aside further developing the notion of a machine *per se*--and engaged instead in implementational endeavors emanating from it, such as breaking the German encrypted communications by means of the urgent construction of a computer machine-- he was still in proxy contact with the cybernetic enterprise, running strong in the US. Ten years

---

<sup>231</sup> One cannot help wondering about the counterfactual of Turing having stayed in America, and in all likelihood having joined the cybernetic group.

<sup>232</sup> See chapter II, section 2.



later, when Turing wrote the 1946 report where he muses about the possibility of a thinking machine, cybernetics was reaching its own pinnacle. The cybernetic view of a machine was much further elaborated, and the notion of a de-physicalized mechanical entity already had full and rightful existence.<sup>233</sup> The mind was already a machine. So when someone like Turing would ask, in his own terms, the heretical question “can a machine think?”, the answer, coming from cybernetics, could only have been in the positive: they obviously can --they already do! The existence of our mind is there for proof. Ours *is* a mechanical intelligence.

This slow but firm parallel development, which ended in a reversal of the question, might have dynamically exercised influence upon Turing himself, in order for him to state his famous rhetorical question.<sup>234</sup> It is not farfetched to suppose, then, that cybernetics own nuanced development of the nature of a machine (encompassing non-material entities through and through) might have played a non-minor role in Turing’s evolution regarding the capabilities of mechanical entities --indeed particularly regarding thinking. Just as the computer was an implementational spin off from the attempt to effectively logicise the mind,<sup>235</sup> cybernetics was just as much a spin off from the attempt to understand the nature of computability --articulating the nature of a machine. But while Turing seemingly wanted to give machines anthropocentric features, cybernetics wanted to make of man a machine.

Artificial Intelligence left aside further investigation into the nature of a machine --and marched forward instead with the *anthropomorphization* of a machine (they would want a

---

<sup>233</sup> In all fairness, due in part to pre-cybernetic British theoretical influence. See the section on Kenneth Craik in chapter IV, section 4 of this work.

<sup>234</sup> Jean Pierre Dupuy sees a connection between AI and cybernetics in this respect, but he does not place it specifically on Turing (Dupuy 2000, chapter 4). Indeed Piccinini criticizes Dupuy’s book for failing to fully flesh out the alleged direct (albeit fairly unrecognized) cybernetic ancestry of cognitive science as a whole (AI included) – which is seemingly promised at the beginning of the book (Piccinini 2002).

<sup>235</sup> John von Neumann based his famed “First Draft” upon McCulloch and Pitts’ 1943 article for designing the architecture that underlies most computers today, known now as the “von Neumann architecture. This claim is not free from controversy. More on this in chapter VI.

machine to show a human feature: thinking). A key for understanding the theoretical launch of AI would thus be Turing's late change of mind in terms of what machines could do. First, they could not ever show capacities reserved only for humans --such as intuition. Later, in principle, they could do just that. The subsequent development in Turing's thoughts could help understand why not only cybernetics ended up being a somewhat "fringe" occurrence in science history, but could also help clarifying the reason behind the overemphasis on classical computationalism in today's cognitive science.<sup>236</sup>

A novel science, cybernetics, might have deeply influenced the mathematician Alan Turing towards a profound change in perspective regarding a metaphysical and ontological issue: the possibility of an artificial, mechanical construction possessing human-like thinking capabilities. It is thus fitting to attempt approaching cybernetics from the angle of the philosophical tenets on which it was based, and the philosophical conclusions that it entailed.

---

<sup>236</sup> Although not fully articulated, Jean Pierre Dupuy correctly recalls some historical pointers that lead towards the fairly orphan status of cognitive science as a whole --reluctant to acknowledge its cybernetic origins (Dupuy 2000, ch. 2). In all fairness, exposing this historical (and philosophical) ancestry might merit a dissertation on its own. This issue is not unrelated to the historical episode depicting an incipient research in artificial intelligence aggressively competing for funding at the beginning of the Cold War (see chapter II, section 4 of this work).

#### IV. Cybernetic tenets: Philosophical considerations.

...there is considerable reason to suppose we have not yet defined the nature of "machine" with sufficient precision or abstraction.<sup>237</sup>

##### 1. Pinning down the core of cybernetics.

Even if the proper definition of cybernetics is an issue without closure,<sup>238</sup> key ideas underpinning the enterprise were present right from its inception –such as circular causality, negative feedback, teleology, information, self-regulation, self-replication and complexity. These semantically dense notions have prompted some minds to find cybernetic signatures embedded in nature as early as in pre-Socratic times. Indeed it is claimed that even the early philosopher Heraclitus of Ephesus already had “proto-cybernetic” insights of sorts. Some arguably recognizable instances of a cybernetic flair could be mentioned:<sup>239</sup>

Circular causality and feedback:

The way up and the way down is one and the same (DK 60)<sup>240</sup>

Upon those who are stepping into the same rivers different and again different waters flow

(DK 12)<sup>241</sup>

---

<sup>237</sup> Shaw & McIntyre 1974, p. 310

<sup>238</sup> The *American Society for Cybernetics* lists 46 serious definitions of the term. See [www.asc-cybernetics.org/foundations/definitions.htm](http://www.asc-cybernetics.org/foundations/definitions.htm)

<sup>239</sup> The connection between cybernetics and Heraclitus has been mentioned time and time again in academic circles, but almost at a colloquial level. A rigorous fleshing out of these possible links escapes the aim of this thesis --and remains to be done. The philosopher who addressed the possibility of a strong connection between them was Martin Heidegger, but he suggested that such connection is too deep for us to understand it at the present time. More on this below.

<sup>240</sup> Marcovich 1967, p. 171

Cybernetics as control:

Thunderbolt steers all things (DK 64)<sup>242</sup>

Wisdom is one thing:

to know the Thought...

by which all things are steered through all... (DK 41)<sup>243</sup>

Information:

This world-order, the same for all...:

an ever-living fire, kindling in measures and going out in measures (DK 30)<sup>244</sup>

If you have heard... not me but the Logos,

It is wise to agree that all things are one (DK 50)<sup>245</sup>

Connexions:

...out of every thing there can be made a unity,

and out of this unity all things are made (DK 53)<sup>246</sup>

Invisible connexion is stronger than visible (DK 54)<sup>247</sup>

The real constitution of each thing is accustomed to hide itself (DK 123)<sup>248 249</sup>

Self-organization and equilibrium (Ashby):<sup>250</sup>

War

is father of all... (DK 53)<sup>251</sup>

---

<sup>241</sup> Marcovich 1967, p. 206

<sup>242</sup> Marcovich 1967, p. 424

<sup>243</sup> Marcovich 1967, p. 449

<sup>244</sup> Marcovich 1967, p. 268

<sup>245</sup> Marcovich 1967, p. 113

<sup>246</sup> Marcovich 1967, p. 105

<sup>247</sup> Marcovich 1967, p. 36

<sup>248</sup> Marcovich 1967, p. 33

<sup>249</sup> The last two quotes might also be read as circular causality being more explanatory powerful than linear causality –which, of course, is itself intimately related with information.

<sup>250</sup> Chapter V, section 2.

Martin Heidegger, who in a later period of his life closely followed cybernetic developments on both sides of the Atlantic, addressed the claimed connection between Heraclitus and cybernetics in his famed *Heraclitus' Seminar*.<sup>252</sup> Prima facie, it would seem that the notion of *steering*, fundamental for cybernetics to the point of conforming its very etymology, and pervasive in some Heraclitean fragments, would point towards a deep shared core between the prescient pre-Socratic thinker and the novel science.<sup>253</sup> However, Heidegger was quick to point out that we should not confuse the Greek primal imagery of Zeus affecting nature (φύσις) using the lightning bolt as a stirring hand --a paradigmatic Heraclitean figure<sup>254</sup>-- with a nascent cybernetic epiphany. In this case, for instance, it would lack the cybernetic feature of a dynamically involved systemic action, Zeus not being “stirred back” or modified in any way, by the affected φύσις. There is an absence of negative feedback, thus rendering the Heraclitean insights as not properly cybernetic. Even if there is a deep connection between Heraclitus and cybernetics, which Heidegger does not deny, it would be beyond our grasp, given our poor current knowledge.<sup>255</sup> Still, some persist in seeing a proto-cybernetic epistemology in Heraclitus –and probably not without reason.<sup>256</sup>

---

<sup>251</sup> Marcovich 1967, p. 145

<sup>252</sup> Heidegger & Fink 1993

<sup>253</sup> “At present, we reflect on the phenomenon of steering. This phenomenon has today, in the age of cybernetics, become so fundamental that it occupies and determines the whole of natural science and the behavior of humans so that it is necessary for us to gain more clarity about it.” (Heidegger & Fink 1993, p.12)

<sup>254</sup> Heraclitean fragment cited above as DK 64 (Marcovich 1967, p. 424).

<sup>255</sup> “I don't want to allow a misunderstanding to arise from my allusion to modern cybernetics in the course of the discussion about what steering is. Misunderstanding would arise if we restricted ourselves to what is said about steering in Frs. 64 and 41, and if we constructed a connection between Heraclitus and cybernetics. This connection between Heraclitus and cybernetics lies much deeper hidden and is not so easy to grasp. It goes in another direction that we could not discuss in the context of our present awareness of Heraclitus.” (Heidegger & Fink 1993, p. 16)

<sup>256</sup> “The deepest intuitions concerning real-life complex systems date back already to Heraclitus.” Hyötyniemi goes on to state three core ideas that in his view are evidently Heraclitean:

Everything changes, everything remains the same.

Everything is based on hidden tensions.

Everything is steered by all other things.

Heidegger did have a deep grasping of the cybernetic ethos, referring to it on one occasion as the “metaphysics of the atomic age”.<sup>257</sup> In an interview he granted to the German magazine *Der Spiegel* (with the explicit condition that it was not to be made public until after his death) he referred to cybernetics in a rather obscure manner. Five days after Heidegger passed away, it was accordingly published.<sup>258</sup> In the translation done by his lifetime friend, Father William Richardson S.J., one can find a dramatic view regarding the role that cybernetics was to assume in mankind:

**Heidegger:** ...The role philosophy has played up to now has been taken over by the sciences today... Philosophy dissolves into the individual sciences: psychology, logic, and political science.

**Der Spiegel:** And what takes the place of philosophy now?

**Heidegger:** Cybernetics.<sup>259</sup>

In order to attempt to further understand the gist of Heidegger’s remarks, it is helpful to cradle them within the broader context of the gradual embodiment of *Machenschaft* --the gigantic “all-pervasive and totalizing ‘makeability’ of everything”.<sup>260</sup> This notion, almost exclusively Heideggerian, has gained renewed attention in the Anglo-American world after the translation in 1996 of *Contributions to Philosophy (From Enowning)*,<sup>261</sup> considered by some to be Heidegger’s *magnus opus* after *Being and Time*. “Machination” as it is sometimes translated, is a process bound to eventually convert everything thinkable into the buildable, and is a parallel outcome of the unfolding of technology --which in Heidegger’s thought amounts to the evolution of

---

Although he does not mention this, one could relate the first one to DK12, the second one to DK 54 & 123, and the third one to DK 41 & 64 (See Hyötyniemi 2006, p. 7).

<sup>257</sup> Dupuy 2000, p. 90. It would seem that this is Dupuy’s personal translation of a part of the Spiegel interview I transcribe below.

<sup>258</sup> On May 31<sup>st</sup>, 1976. The English translation appeared in 1981 (Heidegger 1981).

<sup>259</sup> Heidegger 1981, p. 59

<sup>260</sup> Cited in MacDonald 2008, p. 177.

<sup>261</sup> Heidegger 1999

metaphysics itself. In this later book Heidegger further explicates this heretofore somewhat cryptic notion.

Machination is the domination of making and what is made. But in this regard one is not to think of human dealings and operating but rather the other way around; such [human activity] is only possible, in its unconditionally and exclusivity, on the basis of machination... At the same time machination contains the Christian-biblical interpretation of beings as *ens creatum* –regardless of whether this is taken in a religious or secular way.<sup>262</sup>

Heidegger wrote a treatise on the metaphysical foundations of logic, where he referred to Leibnitz as a major force behind “framing”<sup>263</sup> and machination --due to Leibnitz’ ideal of a universal language based upon logic and calculus.<sup>264</sup> As Wiener himself did, he found in Leibnitz the same “patron saint” for this omni-encompassing machination. However, for Heidegger the impetus of this “machinization” (as *Machenschaft* is sometimes translated), goes well beyond the 17<sup>th</sup> century. For him machination is nothing less than just the other side of the coin regarding Western thought itself. For Heidegger “we know too little of it, even though it dominates the history of being in western philosophy up to now, from Plato to Nietzsche”.<sup>265</sup> It is so profoundly engrained into who we are that

It is the double and contradictory process of ‘humanizing’ of everything, pioneered by the Cartesian conception of the ‘de-humanizing’, typical for an age which has been submitted to the total ‘machinization’, in order to achieve the absolute grip on being, in a word, an age totally enslaved by planetary technology.<sup>266</sup>

Now one can be more equipped for further grasping the reason for such darkness in Heidegger’s

---

<sup>262</sup> Heidegger 1999, § 67

<sup>263</sup> The word that Heidegger used is “enframing”, which gives a more active role to the otherwise more passive connotation of “frame”. See Heidegger 1977, p. 20.

<sup>264</sup> Heidegger 1984

<sup>265</sup> Heidegger 1989, § 61

<sup>266</sup> Quoted in François 2007, p. 432.

remarks regarding cybernetics. Heidegger has some rather unsettling remarks concerning the future of mankind in what pertains to the consequences of the advance of contemporary science for the human condition *vis à vis* its undisclosed cybernetic identity:

No prophecy is necessary to recognize that the sciences now establishing themselves will soon be determined and guided by the new fundamental science which is called cybernetics...Cybernetics transforms language into an exchange of news. The arts become regulated-regulating instruments of information. The development of philosophy into the independent sciences which, however, interdependently communicate among themselves ever more markedly, is the legitimate completion of philosophy. Philosophy is ending in the present age.<sup>267</sup>

Even if this impulse towards framing and machinization --which reached its zenith in cybernetics-- already was with us for as long as there is metaphysics (with its own impetus for understanding, manipulating and controlling), it was Alan Turing who let the machine finally break free. His inquiry into the nature of an algorithm, in order to properly address the nature of computability, let the door open. And even if there is consensus in that a formal and finished definition of an algorithm is still lacking,<sup>268</sup> it exercised a clear influence in the progress of technology.<sup>269</sup> Probably more importantly, as in many intellectual endeavours, Turing's

---

<sup>267</sup> Heidegger 1977b, p. 376. The quote continues:

[Philosophy] has found its place in the scientific attitude of socially active humanity. But the fundamental characteristic of this scientific attitude is its cybernetic, that is, technological character. The need to ask about modern technology is presumably dying out to the same extent that technology more definitely characterizes and regulates the appearance of the totality of the world and the position of man in it.

<sup>268</sup> "When algorithms are defined rigorously in Computer Science literature (which only happens rarely), they are generally identified with abstract machines... this does not square with our intuitions about algorithms and the way we interpret and apply results about them... This problem of defining algorithms is mathematically challenging, as it appears that our intuitive notion is quite intricate and its correct, mathematical modeling may be quite abstract." (Moschovakis 2001, p. 919).

<sup>269</sup> The recognized abstract nature of an algorithm makes it ineligible for patenting:

Determining whether the claim falls within one of the four enumerated categories of patentable subject matter recited in 35 U.S.C. 101 (i.e., process, machine, manufacture, or composition of matter) does not end the analysis because claims directed to nothing more than abstract ideas (such as mathematical algorithms), natural phenomena, and laws of nature are not eligible for patent protection. (United States Patent and Trademark Office 2014, § 2106, II)

In that vein, the patenting of software (arguably a conjunction of algorithms) remains controversial, as one can



tangential (to the nature of a machine) but dramatic insight opened up a new realm of thinking.

The rediscovered awareness regarding this normally overseen nature of a machine, namely, that it did not need to be materially instantiated to be such, is what lies at the core of the cybernetic impetus. This quintessentially cybernetic tenet has been usually overseen when articulating cybernetics as a historical event. Precisely this liberation from the machine's heretofore physical constraints is what made the machinal idea amenable of instantiation in previously mechanistically-unfriendly realms. The enhanced notion of *machine* substantially extended the realm of what could be tractable under a straightforwardly mechanical approach -- or, in the view of the incipient cybernetic scientific community, subsumable under *science*, period.

Arguably a major contribution from Turing to the cybernetic enterprise (and eventually probably to science at large), is the insight that the existence of a mechanism *necessarily implies* the existence of a machine, namely "understanding by purely mechanical process one which could be carried out by a machine".<sup>270</sup> Behaving machine-like is the outcome of a mechanical structure: A machine *acts* like a machine. Indeed pondering on this observation seems to begin to resolve some conundrums in scientific explanation. To the question of whether or not the

---

witness in current news regarding the mutual lawsuits launched between global technology companies (Google, Apple, Samsung, etc.). However, certain uses of an algorithm, on entities that could qualify as "processes", are patentable. The USPTO designates four realms of reality that could receive patent protection:

i. Process – an act, or a series of acts or steps... ("A process is a mode of treatment of certain materials to produce a given result. It is an act, or a series of acts, performed upon the subject-matter to be transformed and reduced to a different state or thing."...) (Ibid., § 2106, I).

Interestingly, the second realm patentable (a machine), is thus defined:

ii. Machine – a concrete thing, consisting of parts, or of certain devices and combination of devices... This includes every mechanical device or combination of mechanical powers and devices to perform some function and produce a certain effect or result. (Ibid.)

It would seem that the USPTO, not without reason, has avoided attaching intrinsic materiality to a machine ("a concrete *thing*") and has essentially linked mechanicity to an underpinning machine-structure ("this includes every mechanical device or combination of mechanical powers").

The other two areas of reality captured by the USPTO for patenting protection are iii. Manufacture and iv. Composition of matter (Ibid.).

<sup>270</sup> Turing 1939, p. 150

recognition of a mechanical process in a phenomenon (natural or not) entails the recognition of a machine, one should, if switching our metaphysical dampers offline, answer in the positive.

Physical entities previously regarded as fundamentally different from machines, are now amenable to being treated mechanically. This occurs due to the fact that physical processes canonically regarded as pertaining to living organisms only --such as self-adaptation to internal and external (environmental) change-- were now found to be mechanizable. When the theoretical (mainly metaphysical) divide between the natural and the artificial collapses, one could start the endeavor of understanding and developing one theory of “control” for both animals (humans included) and machines. By “machines” we could thus refer indistinctly to both the traditional machines and to living organisms. As radical as this view may sound, it would however hardly qualify as radically *new*. The novel aspect of cybernetics, the one that sets it apart, is an aspect that was qualitatively distinct –and that the culmination of the crisis of mathematics indicated elsewhere. Cybernetics now referred to immaterial entities as machines.

One of the objects of study that thereby became scientifically approachable was the human mind. Animal cognition and life in general were now amenable of being treated at the same epistemological level, as complex artificial machines: both display features of successful coping with environments – what in biology was already called “homeostasis”. Recognizing that the very nature of what a machine is lies at the very foundation of cybernetic thinking might put Wiener’s experience with the predictor as midwifery to what was already in the making. At the very least, it qualifies re-reading cybernetics from the standpoint of its unique tenets regarding the nature of a machine.

There have been several ways of classifying the main theoretical tenets of cybernetics, each based upon its founders, its founding papers, its philosophical tenets, its scientific tenets, its

subsumed disciplines and even its subsequently founded disciplines.<sup>271</sup> This writer proposes still a new classification, based upon the conceptualization that, for our purposes, best encompasses the multifarious, multilevel theoretical signature that cybernetics came to be known for. Ross Ashby, the British cyberneticist who might have had cybernetic insights even before Wiener did, spoke of cybernetics as “the domain of ‘all possible machines’”.<sup>272</sup> Departing from this approach, one could get adequately equipped for digging into the metaphysical assumptions behind cybernetics, in order to unveil, tackle and articulate its underlying ontology. Accordingly, this writer found appropriate to predicate of cybernetics as relying on three main theoretical pillars themselves cradled on fundamental aspects of the nature of a machine. I will proceed to spell out the cybernetic hypothesis subsuming its different foundational elements under this proposed triad.

**2. Machines can be teleological:** *Machines are islands of order in a chaotic world. To speak of a deterministic system aiming at an ultimate purpose is not contradictory.*

The notion of a final cause had been cast out from mechanistic explanations due to the alleged fact that it locates “the cause after the effect”, so to speak, in the natural phenomenon being studied. The cybernetic proposal of “teleological machinery” was, thus, controversial to say the least, and for some, just plain scandalous. However, this proposal was justified by retaining an important aspect of such teleological explanation, namely, that it describes a physical process as aiming towards a goal to which such process is projected or attracted, aided by corrections exerted through “negative feedback” --this last being a characteristic feature of cybernetic

---

<sup>271</sup> Dupuy 2000, ch. 2

<sup>272</sup> Ashby 1957, p. 2

parlance.

It has been previously mentioned how the better understanding of the interface between man and machine came to be a crucial priority for anti-aircraft weapon development in the early 1940's. As an important aspect of this effort, the notion of a "machine" had to be meticulously dissected, in order to situate "man" as a smooth moving gear within the whole mechanical system --in fact an integral *part* of the bigger machine (the anti-aircraft gun). A salient feature of this mechanical operation is the role of feedback for the proper functioning of the weapon. Negative feedback<sup>273</sup> allows a machine to engage in self-correction in order to maintain stability toward a certain goal--in this case, hitting the airplane. Information had to be fed into the machine in such a way that it could reliably plot an extrapolation of the future position of the flying target, in order to shoot projectiles in the right direction at the right time. This situation was compared to a strategic movement found in nature itself, namely, that of a cat hunting a mouse. The feline would supposedly not run behind the rodent; instead, it foresees the future position of the prey, and then the hunter runs *directly* towards the precise point where the nutritious moving target would arrive.

As we saw in chapter I, Norbert Wiener (after sharing his thoughts and enticing the audience at the "Cerebral Inhibition Meeting" in 1942) further articulated his claims in a 1943 article under the title "Behavior, Purpose and Teleology" --published in conjunction with the Mexican cardiologist Arturo Rosenblueth and the MIT engineer Julian Bigelow. Publishing this article presented to a larger audience the notion of a teleological machine: a machine whose operation towards a set goal is kept in check by negative feedback, without human intervention. Negative feedback in a physical process had as its aim to regulate the mechanical operation in

---

<sup>273</sup> In contradistinction, positive feedback occurs when feedback has a reverberating effect, reinforcing the process towards the goal, without correction. An example would be a receptor and transmitting radio, which picks up the signal of what it already transmitted, processing it again and broadcasting it stronger (See chapter I, section 2).

such a manner that the disturbances that would direct the processes away from its goal are corrected --checked and put back on track. A paradigmatic example of this dynamical interaction occurs in the famed steam governor mentioned in the first chapter. The three cybernetic authors asserted in this article that a self-regulated anti-aircraft gun machine, a self-guided torpedo, and an airplane's auto-pilot, are all examples of teleological machinery.

Although the article was primarily aimed to a scientific audience, these statements expectably caught the attention of the philosophical community. Indeed, since the idea of a "final cause" always was a problematic issue in philosophy and science, the notion of a teleological machine attracted the interest of some philosophers of science, with the expectable sharp criticism. Attacks against the usage of the term "teleological" were soon targeting the cybernetic authors --in a typically severe Anglo-Analytic manner. In 1950, the then young philosopher Richard Taylor<sup>274</sup> vigorously attacked the cyberneticists' choice of "teleology" as a wishful scientific neologism, claiming that it shared only a superficial resemblance with the classical term used in philosophy from times immemorial. But precisely keeping this term was indeed keenly desired by the cyberneticists, arguing that, with some modification,<sup>275</sup> it was a quintessential aspect of the whole cybernetic stance. They would not want to give it up. Disregarding this, Taylor claimed that, in principle, a torpedo being pulled by a ship by means of a wire or rope attached to its nose could also qualify as a "teleological machine", since its "aim" (eventually reaching and hitting the vessel) would also be accomplished --and indeed even by means of correcting possible disturbances in its trajectory. He trivialized Wiener's view of teleological behavior.

---

<sup>274</sup> Taylor 1950

<sup>275</sup> E.g., foregoing the notion of a final cause, but retaining the notion of a pattern of behavior that aims towards a goal.

Rebuttals came from both sides.<sup>276</sup> Taylor's criticism was solid and well argued. Nevertheless, Wiener and Rosenblueth firmly stood by their ideas. Probably their strongest point against Taylor was that he was still operating within a notion of physics that could be understood as thoroughly Newtonian –the so called “philosopher’s physics”. Causality in complex systems occurs in non-linear ways, and they were referring to those. Also, Wiener and Rosenblueth hinted at some of the problems raised by Taylor as being purely “verbal” in nature, lacking any substance. Finally, they just dismissed some of the philosopher’s concerns as simply not worth addressing. Wiener published in 1945 a piece on the “The Roles of Models in Science”<sup>277</sup> to justify his choice of retaining the somewhat modified notion of teleology, so that it could be applied to mechanical systems --and he referred the reviewer (Taylor) to this article in his response to the critical piece. It is worth noting that of the 10 conferences held by the Macy foundation, a professional philosopher was invited only to the first one –Filmer Northrop (1893–1992). Interestingly, however, most cyberneticians pursued formal studies in classical philosophy for some time earlier in their academic lives.<sup>278</sup>

There was, however, another episode that would point out to a later deep rift in scientific worldviews among cyberneticians regarding the idea of behavior-with-an-aim. This rift might have partly spelled out the implosion of the enterprise itself at the theoretical level. This could indeed be one of the first instances where a new paradigm of science, which is reportedly advancing in the 21<sup>st</sup> century,<sup>279</sup> trumpeted across the legacy of a previous one, bringing the latter into oblivion. Cybernetics might have inadvertently been one of its first victims. This assertion requires some explanatory context.

---

<sup>276</sup> Rosenblueth & Wiener 1950, Taylor 1950b

<sup>277</sup> Rosenblueth & Wiener 1945

<sup>278</sup> Particularly so in the cases of Norbert Wiener and Warren McCulloch.

<sup>279</sup> Prigogine & Stengers 1984

Just as Wiener's ideas were already known in relevant circles before the publication of his book in 1948, a younger mathematician and engineer was, around the same time, beginning to gain notoriety. Since his years as an MIT electrical engineering graduate student, Claude Shannon (1916–2001) was already interested in Wiener's work, knowing that the latter was doing important research at his same institution. While pursuing a Master's degree, Shannon became an assistant for Vannevar Bush' second differential analyzer mentioned above.<sup>280</sup> Once he landed a permanent position at Bell Laboratories, war-related work put both Wiener and him in direct contact, since Wiener had to deal with the Bell Company on a frequent basis (both Bell Labs and MIT were funded under the referred above subsection D-2). Shannon started to pay Wiener frequent visits at his research lab, seeking inspiration and advice on relevant topics. These tours became so frequent indeed, that at a certain point –and aware of the closely related research occurring at Bell Labs—Wiener began to be concerned about the possibility of having his ideas siphoned out without his due acknowledgement.<sup>281</sup> In fact, the same year that Wiener published *Cybernetics*, Shannon published the famed “A Mathematical Theory of Communication”--a long article delivered in two installments, soon acquiring legendary status within engineering circles.

Warren Weaver, still supervising the military-related research of section D-2,<sup>282</sup> endorsed the paper by later expanding it into a joint book –coauthored by Shannon and himself— and

---

<sup>280</sup> See chapter I, section 3. Shannon's Master's thesis was an application of Boolean algebra to the workings of such machine. He went on to pursue a relatively quick Ph.D. in mathematics, also at MIT, proposing an algebra applicable to Mendelian genetics. It was right after earning this degree that he began working for Bell Laboratories.

<sup>281</sup> Conway 2005, p.186

<sup>282</sup> Although Warren Weaver was not directly immersed in the cybernetic project, partly due to his role as science administrator (rather than practitioner), he was by proxy firmly related to the circle. In a 1934 Annual Report of the Rockefeller Foundation, one can already notice a cybernetic incipient signature regarding the future of humanity:

Important questions are: Can we obtain enough knowledge of the physiology and psychobiology of sex so that man can bring this aspect of his life under rational control?... Can we develop so sound and extensive a genetics that we can hope to breed in the future superior men? Can we solve the mysteries of the various vitamins, so that we can nurture a race sufficiently healthy and resistant? Can psychology be shaped into a

given the more ambitious title *The Mathematical Theory of Communication*.<sup>283</sup> The monograph included Shannon's paper and Weaver's simplified explanation and assessment of the theory's foreseen import, taking advantage of the latter's gift for conveying otherwise cryptic concepts – both as a people's person and in writing. The paper acknowledged very clearly Wiener's influence, and Wiener was thankful and respectful of such homage.<sup>284</sup> Shannon carried on Wiener's insight on the digitality of information, asserting logarithm at base 2 at the core of the general equation of information. In fact, even if Shannon's Bell Labs colleague John Tukey first coined the term "bit" for "binary digit",<sup>285</sup> Shannon established the term as a core part of information theory's structure.<sup>286</sup> However, Shannon's proposal had, as Wiener publicly praised, original features of its own.

As advanced above,<sup>287</sup> the paper started putting forward (or "clarifying", in Shannon's view), the controversial claim that information has nothing to do with meaning –or rather, that the problem of semantics should be bracketed as irrelevant to the fundamental problem of communication –namely, "that of reproducing at one point either exactly or approximately a message selected at another point".<sup>288</sup> As shown in the first chapter<sup>289</sup> several people at Macy were concerned about this move away from the "significance" of the message conveyed --among those being Donald Mackay and the editor in chief, Heinz von Foerster. However, much of this unease might had been partly triggered by a misunderstanding fostered by the intense enthusiasm

---

tool effective for man's everyday use? In short, can we rationalize human behavior and create a new science of man? (Rockefeller Foundation 1933, p. 199)

<sup>283</sup> Shannon & Weaver 1949

<sup>284</sup> Later on Wiener would become more assertive, and even protective, of the role that he himself played in developing "Shannon's theory". Cf. Conway 2005, pp. 187-188

<sup>285</sup> Shannon 1948, p. 380

<sup>286</sup> Boden 2006, pp. 204-205; Hayles 1999, pp. 50-57

<sup>287</sup> Chapter I, section 4.

<sup>288</sup> Shannon 1948, p. 379

<sup>289</sup> Concerning the 8<sup>th</sup> Macy Conference



of cybernetics itself.<sup>290</sup> In any case, due to its simplifying character, more suitable for engineering purposes, Shannon's take prevailed as the approach of choice.<sup>291</sup>

An interesting shift away from Wiener's views, however, occurred in Shannon's paper. Wiener's views on entropy, as previously pointed out,<sup>292</sup> largely came to him out of the realisation of deep correlations found between on the one hand his early works on chaos and thermodynamics, and on the other the nature of the message occurring between internal structures and their environment in surviving systems. This context led Wiener to equalize information with negative entropy. Information provided organization to a reality that inherently tends towards chaos. But Shannon stated that information was just the opposite: that information was entropy itself --and not negative entropy. This reversal carries relevance since it would come back later under disguise with important consequences for the cybernetic enterprise.

As Weaver explained Shannon's view of information, "information is associated with the amount of freedom of choice we have in constructing messages".<sup>293</sup> If the message conveyed happens to bring to the receptor what the receptor already has, then little information has been conveyed. If, on the other hand, there is a high degree of uncertainty regarding the possibility of reconstructing the message at the receptor's end, then one can talk of a greater amount of information at stake. So it would seem that information would amount to being a function of the

---

<sup>290</sup> A recorded conversation between Claude Shannon and his wife Betty, towards the end of his life, seems to point in that direction:

Betty: In the first place, you called it a theory of communication... You didn't call it a theory of information.

Claude: Yes, I thought that communication is a matter of getting bits from here to here, whether they're part of the Bible or just which way a coin is tossed...

Betty: It bothered you along the way a few times but by that time it was out of your hands.

Claude: The theory has to do with just getting bits from here to here... That's the communication part of it, what the communication engineers were trying to do. Information, where you attach meaning to it, that's the next thing, that's a step beyond, and that isn't a concern of the engineer, although it's interesting to talk about (Conway & Siegelman 2005, pp. 189-190).

<sup>291</sup> Not without some still clinging for a while longer to Mackay's more cumbersome (for engineering) approach (Hayles 1999, p. 56).

<sup>292</sup> Chapter I, section 4.

<sup>293</sup> Weaver 1949, p. 12

amount of randomness in the conveyance of a message --or as Weaver would put it, “a measure of one’s freedom of choice in selecting a message”.<sup>294</sup> Indeed, Weaver would cleverly capture the case where a good degree of order is present, and the amount of choice is poor, saying that this “situation is highly organized, it is not characterized by a large degree of randomness or of choice, the information (or the entropy) is low”.<sup>295</sup>

In contrast, Wiener stopped short of identifying entropy with “evil;” we are after all destined to turn to ashes at the end, due to reality’s tendency towards chaos and disorganization. Should we thus consider Shannon’s approach as inimical to that ultimate challenger of entropy, “life” –that most excellent embodiment of information? That would not seem to be the case, given Shannon’s keen interest in synthetically recreating instances of living systems, as portrayed in the construction of his famous robotic “maze-solving rats”.<sup>296</sup> It would also be relevant to point out that the scientific and intellectual environment regarding the physical and its semantic implications for chaos were about to begin undergoing revision. Entropy, or the randomness present in a physical environment started to be seen as a fruitful environment for complexity to arise. Physical systems, organic or not, had to engage in a certain degree of self-organization in order to be able to cope with that changing physical environment. Lack of flexibility would entail destruction –and death.<sup>297</sup>

Wiener’s view on entropy as contrary to life and organization largely comes from the observation of an isolated system’s behavior under the 2<sup>nd</sup> law of thermodynamics. Since energy is not created or destroyed, in an exchange of it something has to give, and that would be the

---

<sup>294</sup> Shannon & Weaver 1949, p. 9

<sup>295</sup> Ibid., p. 13

<sup>296</sup> Hayles 1999, pp. 63-65

<sup>297</sup> The received view of entropy as the spiraling into chaos and oblivion inherent to nature gets questioned as the default state of affairs in things (Prigogine & Stengers 1984). More on this below.

order (information) that maintains the system together.<sup>298</sup> This inverse relational status between entropy and life was sanctioned by Schrodinger<sup>299</sup> –which Wiener approved. Once the environment is factored into the picture, it would seem that the emergence of complexity (partly due to random changes in order to cope) is now the friendly context for life to emerge. Indeed it seems as if self-organization would need such an explosion of variables to successfully match a changing milieu and reach homeostasis.

Claude Shannon, in that respect, advanced what was to be later given empirical plausibility by the 1977 Nobel Prize in chemistry Ilya Prigogine (1917–2003) for his study of self-organization and thermodynamics. It would seem that order in nature (and life itself) are necessary processes of reciprocal coping balances that are actually bound to happen –*contra* the traditional view regarding the consequences of entropy.<sup>300</sup>

Matter near equilibrium behaves in a "repetitive" way. On the other hand, far from equilibrium there appears a variety of mechanisms corresponding to the possibility of occurrence of various types of dissipative structures. For example, far from equilibrium we may witness the appearance of chemical clocks, chemical reactions which behave in a coherent, rhythmical fashion. We may

---

<sup>298</sup> “Every process, event, happening –call it what you will; in a word, everything that is going on in Nature means an increase of the entropy of the part of the world where it is going on... An isolated system or a system in a uniform environment... increases its entropy and more or less rapidly approaches the inert state of maximum entropy. We now recognize this fundamental law of physics to be just the natural tendency of things to approach the chaotic state...” (Schrödinger 2012, p. 71).

<sup>299</sup> “[A] living organism continually increases its entropy --or, as you may say, produces positive entropy-- and thus tends to approach the dangerous state of maximum entropy, which is of death. It can only keep aloof from it, i.e. alive, by continually drawing from its environment negative entropy... What an organism feeds upon is negative entropy. Or, to put it less paradoxically, the essential thing in metabolism is that the organism succeeds in freeing itself from all the entropy it cannot help producing while alive... Thus the device by which an organism maintains itself stationary at a fairly high level of orderliness (= fairly low level of entropy) really consists [in] continually sucking orderliness from its environment.” (Ibid.)

<sup>300</sup> “At all levels, be it the level of macroscopic physics, the level of fluctuations, or the microscopic level, *nonequilibrium is the source of order. Nonequilibrium brings "order out of chaos."*” (Prigogine & Stengers 1984, pp. 286-287. Italics original)

also have processes of self-organization leading to nonhomogeneous structures to nonequilibrium crystals.<sup>301</sup>

The consequences of this surreptitious semantic change regarding information as a fundamental element of cybernetics might have entailed a deep displacement in the ontological commitment underlying the project. I will come back to these effects in the chapter regarding the work of William Ross Ashby.<sup>302</sup>

**3. Machines can be immaterial:** *In virtue of this feature, they can be “extended” to encompass, beyond the organic realm (life), non-physical entities as well (mind).*

It has been mentioned that McCulloch was tinkering with the notion of idealized neurons, in connection with the premise that there should be a logical way of transducing cognitive processes --so that we can have a glimpse of hope for relating the exactness of mathematics with the chaos of thinking. Indeed McCulloch already had for a while a nagging insight which suggested that the nervous system was nothing but the ‘wetware’ instantiation of a higher, more universal, abstract logical machine. In fact his artificial neurons would stand for the logical gateways existent thanks to Boolean logic (‘and’, ‘or’, etc.). It is in this context that Turing’s observations regarding an “ideal” machine importantly shaped McCulloch’s own educated musings. He recalls that

it was not until I saw Turing's paper that I began to get going the right way around, and with Pitts' help formulated the required logical calculus... The important thing was, for us, that we had to take a logic and subscript it for the time of the occurrence of a signal (which is, if you will, no more than a proposition on the move). This was needed in order to construct theory enough to be

---

<sup>301</sup> Prigogine & Stengers 1984, p. 13

<sup>302</sup> Chapter V.

able to state how a nervous system could do anything. The delightful thing is that the very simplest set of appropriate assumptions is sufficient to show that a nervous system can compute any computable number.<sup>303</sup>

This model, which re-instantiated a pristine logical platform into an object of reality, thus “purifying” the latter, was a major impulse for cybernetics, which was now holding a mechanical explanation for the mind’s hitherto scientifically intractable operations. Indeed the article turned out to be a theoretical pillar not only for cybernetics, but also for the start-up of the field of computer science. John von Neumann’s famed *First Draft of a Report on the EDVAC*<sup>304</sup> used, to the puzzlement of the engineers hired to build these early computers, McCulloch’s vocabulary and symbols of “neurons” throughout his paper. He was open about having written the draft based upon McCulloch and Pitts’ proposal.<sup>305</sup> Von Neumann’s paper served as a blueprint for the construction of one of the first “stored-program” computers (the EDVAC<sup>306</sup>) to be delivered, as it was expected given the urgencies of the time, for ballistics research at the Aberdeen Proving Ground.<sup>307</sup> The model can be explained in the following way.

For instance, in the case of an organism (e.g., a bird) “deciding” whether or not eating an object (e.g. a blueberry), something like the following procedure would take place. The animal’s neuron would receive two possible inputs, one regarding shape (roundness) and the other regarding color (purple). When the bird’s biological detector perceives “roundness”, it would fire a signal (symbolized by a 1). If it does not --if the object is squared--, then there would be no firing (symbolized by a 0). By the same vein, if the detector perceives “purple”, then 1 would

---

<sup>303</sup> Von Neumann 1951, p. 32

<sup>304</sup> Von Neumann 1945

<sup>305</sup> Von Neumann 1945, p. 5.

<sup>306</sup> **E**lectronic **D**iscrete **V**ariable **A**utomatic **C**omputer

<sup>307</sup> Situation which, as we will see below, signaled the different orientation that Americans, in contrast with its British Allies, had regarding the war efforts --which reverberated in different ways of practicing cybernetic science. More on this in the next section (4).

occur --if it perceives blue, then 0 would happen. The situation that we have thereafter is that when the bird perceives a blueberry, a threshold would be formed in the ideal neuron. This threshold would be represented by the number 2, formed as a result of detectors firing due to the perception of both roundness (1) and purpleness (2):  $1 + 1 = 2$ . If this threshold is reached, the neuron will this time itself fire a response, represented by 1 if the command is “eat”, and by 0 if it is “don’t eat”. Only if a threshold that is equal or more than 2 is reached, then the positive command will occur. Since both the roundness and purpleness of a blueberry causes the detectors to send overall signals that reach the 2 threshold, then a 1 signal will be fired and the bird will eat it. If on the other hand, the bird sees a golf ball, there would be roundness (1) but no purpleness (0), forming a total output of just 1 ( $1+0=1$ ). Since the threshold requires a minimum of 2 for the neuron to fire “eat” (1), the bird would not eat. A similar situation would happen with a violet, where there is purpleness (1) but not roundness (0), and so the total output would still be below the minimum required threshold of 2 ( $1+0=1$ ). Still in the case of a hotdog, lacking this food both roundness and purpleness, the outcome will fall further below the minimum required to reach the 2 threshold --since  $0+0=0$ . And so on and so forth.<sup>308</sup> Expectably, McCulloch and Pitts’ networks get quickly more complicated once more features are added into the operation.

Part of the fame of this article resides in its almost epic difficulty. John von Neumann was quick in pointing out that its cryptic nature was in great part due to the usage of Rudolph Carnap’s *sui generis* and somewhat convoluted symbolic nomenclature. This unfortunate choice --in the eyes of von Neumann-- was probably unavoidable anyway, since Carnap was one of Pitts’ admired mentors. Despite its inaccessibility (which partly almost cost it its publication), it became a theoretical bedrock for cybernetics. It was not, however, devoid of problems.

---

<sup>308</sup> Marsalli 2006

The complexity that is rapidly attained after more elements enter the cognitive picture turned out to be particularly problematic in a two-fold manner. On the one hand the networks, even if complex, were still considered too simple for occurrences of higher cognition. On the other, even if immensely more complex cognitive processes were to be mapped, the constructability of the model would get severely compromised. Since cybernetics had as a signature feature the buildability of its models,<sup>309</sup> this also brought about theoretical unease. In fact, there were later developments of the model, largely side-stepping the constructability mandate, (e.g., the perceptron), to some extent betraying the cybernetic mandate of material grounding.

Pointing out the importance of this paper for the theoretical genesis of computer science is indeed relevant for a better understanding of what later occurred to the cybernetic enterprise. Right from the start, von Neumann, who certainly saw its usefulness --even built an impressive physical device out of it-- insightfully saw its shortcomings as well. These weaknesses became evident for him when the model was applied to what was supposed to understand in the first place: the mind. When the neatly abstract and disembodied machine is instantiated back into materiality, a remodeling of the studied object seems to occur, severely complicating it right into intractability, evincing deep theoretical problems in the model. Thus it is from insights into the feasibility of the McCulloch and Pitts' model into realms other than computer machines, that the most severe criticisms against the core of cybernetics would be launched.<sup>310</sup>

---

<sup>309</sup> See the section below. Also Pickering 2005.

<sup>310</sup> The problematic epistemological strategy that the construction of McCulloch and Pitts' networks entailed will be further addressed in the chapter pertaining to the contribution of John von Neumann (VI).

#### **4. Machines are embodied theories:** *Physical constructability should act as theory grounding.*

*Theoretical investigation ought to be constrained by a concrete empirical buildability*

Probably the most famously peculiar characteristic of the cybernetic movement was its emphasis on the necessary constructability of whatever was proposed –instantiated in cybernetics’ legendary autonomous robots. Certainly this leniency towards the engineering aspect of a theory has older roots.<sup>311</sup> Constructability turns out to be the strongest sign of a model being truthful and reliable: knowledge entails buildability. True epistemic assets entail engineering success, our model enjoying an equilateral isomorphism with the modeled.<sup>312</sup>

It is relevant to point out that the cybernetic collaboration coming from Britain had this distinctive feature, which hitherto lacked, at least in such measure, in the American side of cybernetic research: machine building as theory testing.<sup>313</sup> In fact, it could be maintained that if the English were not as actively engaged as Americans in cybernetic research (as both

---

<sup>311</sup> Duns Scotus already proposed a “univocity of being”, where our difference from the Divine would be one of quantity, not of quality --thus suggesting that true knowledge would imply the necessity of building, since the Creator knows while/by/in creating (Scotus 1300, p. 20). Immanuel Kant later suggested that God’s intuitive knowledge (the divine equivalent of the kind of knowledge generated by our “pure intuitions of sensibility”) would necessarily imply creation (Kant 1787, B72). But it is Giambattista Vico who most clearly stated, regarding the essential reliance of knowledge on the best possibly constructed model, that “Verum et factum convertuntur”: The Truth and the made converge --or in more colloquial language, “if you know it, build it”. Only a Creator could fully know its own creation (Vico 1710, pp. 45-48). But since Scotus defends that humans epistemically share the divine light in a lower degree --due to the incarnation as a full human in Christ--, we too can in principle also know in such light. This emphasis on constructability as criterion of truth has inspired later thinkers more concerned with the tractability of physically embedded problems. Vico’s prescient *Verum Factum* seems to forward the Scotian *Univocatio Entis* into a culmination of knowledge attainment that might metaphysically underpin contemporary technological research to an important extent. His claim on the future “certain sciences” is telling: “The most certain sciences are those that wash away the blemish of their origin and that become similar to divine science through their creative activity in as much as in these sciences that which is true and that which is made are convertible” (Vico 1710, p. 52). It should not come as a surprise that a contemporary laboratory in the Netherlands is named “The Giambattista Vico Institute for Cybernetics and Applied Epistemology”. For an assessment on Medieval and early Modern insights present in contemporary science and technology, see Malapi-Nelson 2014.

<sup>312</sup> The psychoanalytic assessment of what supposedly lurks behind the mostly male-filled area of Artificial Intelligence, the so-called “womb envy” syndrome, could in fact be subsumed under this more primogenial philosophical rubric referred to in the previous note. See McCorduck 2004, ch. 5.

<sup>313</sup> This likely is the legacy of experimentation of the Royal Society since Bacon, Boyle and Newton. I thank my supervisor for pointing this out.



McCulloch and Wiener asserted after visiting their English counterparts),<sup>314</sup> the British definitely compensated for that lack with their emphasis on machine building. This aspect of British cybernetics might have been partly the result of the particular orientation of their military research, due to their own urgencies of war: less theoretical investigation and more quickly-built defensive gadgetry. Such a circumstance would not only put a constraint in terms of time and secrecy (which the Americans also had), but also foster a hands-on approach to scientific research. In addition, since cybernetics did not have in the UK the official financial backing that it enjoyed in the US, much of the British cybernetic developments took place as side jobs, even as hobbies, being performed in the scientists' kitchens and garages during their spare time.<sup>315</sup>

The physical embodiment of an otherwise non-material machine, not only as a heuristic tool, but as the ultimate ontology of a living object, was an idea already present in a premonitory way in early 20<sup>th</sup> century British psychology. Both the Ratio Club and a thinker who never made it to his first meeting due to untimely death, become here highly relevant. Kenneth Craik (1914–1945) was a young philosopher with an intense interest in the workings of the mind and a confessed passion for machinery. After studying philosophy, he began pursuing a Ph.D. in psychology. In 1943, five years before Wiener's *Cybernetics*, he published a book entitled *The Nature of Explanation*. The thesis in this book had profound implications for cybernetics on both sides of the Atlantic.

Craik advanced the idea that one of the most startling capacities of the mind was the possibility of putting together prediction models about the state of affairs of reality. That is largely what thinking would amount to. In principle, Craik stated, a machine should be capable of doing the same: be able to predict situations in reality. In fact, the mind was itself, a type of

---

<sup>314</sup> See chapter II, section 2.

<sup>315</sup> Ibid. Also Pickering 2010, p. 10.

machine capable of constructing models of the world. This manner of understanding reality, by means of actual model-building, was called the “synthetic method”.<sup>316</sup> And this idea began to make a profound impression in those who would later become British cyberneticists. Emphasis in robot making as a heuristic tester for psychological theories was his explanatory signature. However, just two years after the publication of his book, and just before the first Ratio Club meeting, he was killed in a bicycle traffic accident –at only 31 years of age.

It has been previously pointed out that the Ratio Club, in an important way was marked by this legacy.<sup>317</sup> In fact, the group, which deeply lamented this loss, considered naming itself the “Craik Club”. Craik’s emphasis in his “synthetic method” as a way to understand physical phenomena (i.e., gadget building) allegedly set the British cyberneticians apart from their American counterparts --mainly focused in mathematical expansions of what could be physicalized by means of extending the notion of what can be understood mechanistically. Also, the social implications of cybernetic thinking, which allegedly gave the American side a strong footing in terms of mainstream attention and founding, was entirely avoided by the English group. In fact, Grey Walter, who later became a ‘media personality’ of sorts due to his artificial tortoises, not only openly acknowledged receiving inspiration from Craik, but furthermore, characterized the American cybernetic movement as embodying Craik’s ideas --minus the fun.<sup>318</sup>

Having in consideration that Ross Ashby himself --who was keen to convey the plausibility of his ideas by means of showing, in the metal, constructed devices-- kept active correspondence with Craik, one would be tempted to agree with such remark.<sup>319</sup> In fact, Craik was already publishing in 1943 what would provide important cybernetic insights later –the same

---

<sup>316</sup> Cordeschi 2008a

<sup>317</sup> Chapter II, section 2.

<sup>318</sup> “...thinking on very much the same lines as Kenneth Craik did, but with much less sparkle and humour” (Husbands & Holland 2008, p. 14).

<sup>319</sup> Husbands & Holland (ibid.) share such such opinion.

year when Wiener was just publishing his foundational paper with Rosenblueth and Bigelow, and full five years before the publication of his book in 1948. These timeline considerations could point out to a prescient philosophy that could concede to Craik the position of a cybernetic pioneer. However, in all fairness to American cybernetics, one would have to also consider the machines built in the US, and recognize that they did not play a small role in the development of cybernetic thinking. A list of relevantly built cybernetic machines, both British and American, would be:

- Ross Ashby's Homeostat –a machine that will have substantial implications for cybernetic thinking.<sup>320</sup>
- Grey Walter's "tortoises" –phototropic, self-charging robots that launched cybernetics in popular culture.<sup>321</sup>
- Wiener's (and Bigelow's) Predictor –the machine that inspired the 1943 foundational article.<sup>322</sup>
- Von Neumann's computer –mapped upon McCulloch and Pitts' article, and then the source of further cybernetic thinking.<sup>323</sup>
- Claude Shannon's "rats" –maze-solving small mobile robots.<sup>324</sup>

The first two machines were British. The remaining three were American. An important aspect of the operation of anti-aircraft weaponry that jumpstarted the cybernetic movement in America was the need to make the calculations for future positions of the target aircraft, tractable in a timely way, for the expectably urgent reasons that characterize war times. If done by hand, these procedures could have taken months. There was a pressing need to develop a mechanized way to

---

<sup>320</sup> Referred to in chapter V, section 2 of this work.

<sup>321</sup> Pickering 2010, pp. 37-54

<sup>322</sup> Referred to in chapter I, section 2 of this work.

<sup>323</sup> Referred to above and in chapter VI, section 1.

<sup>324</sup> Hayles 1999, pp. 63-65

*compute* these numbers. This set America on a different path from Britain concerning the orientation and purpose of their computing research. After Turing's contribution with his *On Computable Numbers* to the understanding of the notion of a machine, by means of spelling out what an algorithm stands for, John von Neumann, almost in parallel with the British allies, embarked upon the goal of constructing a machine capable of doing algorithmic computations -- among other tasks. The British were instead focused on cracking the coded German communications. Meanwhile, the American interest in computer building morphed into von Neumann's involvement with the EDVAC project, a computing machine to be delivered to the Ballistics Research Laboratory of the Aberdeen Proving Ground in Maryland --the same facility where Norbert Wiener was doing ballistics research, applying his studies of Brownian motions.

As pointed out in chapter I, it has to be said for the American side of cybernetics regarding the anchoring of theoretical advancement in physical substrata, that Wiener was right from the beginning of his projects vehemently pushing for required engineering practical skills for his working team --given his past own misfortunes with the hands-on aspects of his theoretical musings.<sup>325</sup> In fact, he politely scolded von Neumann at some point for not having sufficiently performed a transition from designing computing architecture to immersion into control theory.<sup>326</sup> He looked for scientists with hands-on knowledge in at least basic electrical engineering --even as a hobby (e.g., radio construction). This type of approach underpinned Wiener's whole project of the Anti-Aircraft Predictor, as also seen by his enthusiastic support of Julian Bigelow's ingenious fighter simulator. Hence, there was a clear and unambiguous emphasis in cradling mathematics and physics in the engineering-conscious, nitty-gritty sensitive, aspect of machinery. In Wiener's mind, proper modeling was of utmost importance in

---

<sup>325</sup> As seen on chapter III, section 2 of this work. Also Galison 1994, pp. 233-252

<sup>326</sup> Galison 1994, p. 249

the advancement of scientific knowledge, as both him and Rosenblueth emphasized in their 1945 paper on “The Role of Models in Science”,<sup>327</sup> --stating the peculiar claim that “the best material model for a cat is another, or preferably the same cat”.<sup>328</sup>

It is suggested that the cybernetic emphasis on this useful theoretical constraint --the need to ground it in physical operability-- is what made cybernetics a qualitatively different proposal in the history of science. It allegedly shifted the usual theoretical focus from being epistemologically-based to one founded on the ontology of objects.<sup>329</sup> Indeed, testing how a certain philosophical or mathematical insight would work-out by means of observing the coping of a machine (metallic or organic) with the environment, fundamentally mattered. This approach in fact went well beyond a mere methodological choice for scientific research. As seen above, Wiener regarded the reality of a “machine” as “islands of hope”, as pockets of intention and will, in the midst of the deadly indifference of entropy --the ever-present dark reminder that reality spirals down into oblivion.<sup>330</sup>

Considering the importance that machine-building had for the cybernetic ethos, there is something to be said regarding the way in which these instantiations affected its own theoretical cores. The next two chapters will attempt to show the influence that the relevant constructed machinery had on cybernetics’ evolution itself. One machine will be British, by Ross Ashby, and the other(s)<sup>331</sup> American, by von Neumann. This machinery provoked, on the one hand, the unveiling of problematic commitments that were not in the open when they still remained solely in the realm of theory. On the other, it triggered retreat into realms rejected by cybernetics in the first place as part of its fundamental mandate, going against ultimately cybernetic standpoints.

---

<sup>327</sup> Rosenblueth & Wiener 1945

<sup>328</sup> Rosenblueth & Wiener 1945, p. 320

<sup>329</sup> Pickering 2010, ch. 2

<sup>330</sup> See chapter I, section 4 and chapter II, section 1 of this work.

<sup>331</sup> One physically embodied, the other immaterial.

## V. Extending the scope of a machine ontology.

[Cybernetics] takes as its subject-matter  
the domain of ‘all possible machines’<sup>332</sup>

### 1. William Ross Ashby’s nature-machine equalization

If one is to make a philosophical survey of what the notion of *machine* stands for, one will be quickly surprised by the fact that there is little written on the topic per se --"surprised" because philosophers and scientists have been extensively using its derivatives (mechanistic, mechanical, machinal, etc.) for at least three centuries by now. Indeed, upon tracing back the history of modern science, one cannot find a rigorous attempt at a definition of a machine.<sup>333</sup>

There have been good studies on the notions of automaton,<sup>334</sup> artifact,<sup>335</sup> and even gadget,<sup>336</sup> which go all the way back to Ancient Greece.<sup>337</sup> However, these refer, once again, to derivatives of the more general and substantial notion of a machine, and not to the object, or the idea of the object, in itself. This reason alone (namely, the lack of a historical factum pointing

---

<sup>332</sup> Ashby 1956, p. 2

<sup>333</sup> The closest attempt at a definition found by this writer is Heinrich Hertz’—although he does not dwell on the nature of a machine. Hertz briefly states the following, when exposing the main elements of his *Principles of Mechanics*: “A system whose masses are considered vanishingly small in comparison with the masses of the systems with which it is coupled, is called a machine.” (Hertz 1899, § 531)

<sup>334</sup> Chapuis & Gelis 1928

<sup>335</sup> Margolis & Laurence 2007; Read 2009

<sup>336</sup> Berryman 2003

<sup>337</sup> Ancient Greeks report the construction of mechanical marvels with unique functions. Their attraction to mechanical devices was mostly directed towards the aspect of clever and useful tools. Archimedes’ inventions could be brought up as illustrative examples. Medieval times witnessed an idiosyncratic infatuation with machinery, epitomized in the construction of incipient logical machines. Ramon Lull’s rotating concentric circles and *volvelles* in general could be representative instances. Modern science was the locus for a renewed interest in machines, this time chiefly taking in consideration the heuristic capacities that they allegedly carry in terms of explanatory power. Rene Descartes and Leonardo Da Vinci could be pointed out as prime examples of thinkers who shared a strong interest in machinery.

towards a rigorous treatment of the idea of a machine) would already justify addressing the philosophical underpinnings of what a machine stands for. But there is a more relevant, compelling reason for squarely tackling the idea of what a machine means. Ashby attempted to define cybernetics as the theory of all possible machines.<sup>338</sup> This machine-based approach to the understanding of cybernetics gained strength after it was gradually realized that for understanding animal (man) and machine under the same mechanistic laws,<sup>339</sup> a clarification of what a “machine” actually meant was desirable –not in an altogether different fashion as to what occurred a couple of decades earlier, concerning the need for articulating the notion of an algorithm, in order to realize what can and cannot be computable.<sup>340</sup>

The scientist who somewhat mysteriously asserted this preeminence --of the machine as lying at the hard core of cybernetics-- went by the name of William Ross Ashby. This British medical doctor and psychiatrist was a founding member of the Ratio Club: the famed “study group” that stood as the (more humble) English parallel to the Macy gatherings in the US, and which included Alan Turing and Grey Walter as regular members.<sup>341</sup> Ashby’s theoretical developments were imbued, pretty much right from the beginning, with the characteristic cybernetic signature. Already in 1940, a full three years before both foundational articles by Wiener, McCulloch and company saw day light,<sup>342</sup> Ashby published a distinctly cybernetic article that would go on to serve as the cornerstone of his own subsequent plethora of ideas.<sup>343</sup>

Ashby confessed that for several years an apparently insurmountable difficulty regarding the possibility of understanding a living organism as a machine haunted him. There seemed in

---

<sup>338</sup> Quote above (Ashby 1956, p. 2).

<sup>339</sup> Implicit in the very subtitle of Norbert Wiener’s *Cybernetics: Or, the Theory of Control in Animals and Machines* (Wiener 1948).

<sup>340</sup> Chapter III, section 2

<sup>341</sup> Chapter II, section 2.

<sup>342</sup> Rosenblueth, Wiener & Bigelow 1943; McCulloch & Pitts 1943

<sup>343</sup> Ashby 1940

place a default view of an utter incompatibility between the notions of a machine and, say, a brain --due to the notorious adaptive capacities of the latter, allegedly inherently lacking in the former. Candidly referring to himself in the 3<sup>rd</sup> person, he shared:

The author has been concerned for some years with attempting to reconcile the fact of the occurrence of this type of adaptation with the usual hypothesis that the brain is a machine, i.e., a physico-chemical system. In the opinion of some, there is a fundamental impossibility that any isolated machine could show behaviour of this type...<sup>344</sup>

Ashby thought that what may lie at the crux of this claimed incompatibility might be a poor understanding of the entity to which the nervous system is being compared. He asserted that “[i]t seemed possible, however, that a more elaborate study of the essentials of ‘machines’ might show possibilities hitherto overlooked”.<sup>345</sup> Accordingly, he proceeded to gradually break down the task at hand --namely, to find out what a machine is-- into pieces, carefully re-building the notion, along with the isomorphic characteristics allegedly shared with the nervous system, in a meticulous and sequential fashion. His suggestion for a “more elaborate study” took him almost a decade.

As stated above, Ashby’s cybernetic argumentation started in 1940. In an article entitled “Adaptiveness and Equilibrium”, he began by addressing the very feature that supposedly sets the nervous system irreconcilably apart from a machine, advanced in its title: *adaptiveness*. He proposed equalizing it with the familiar and recurring physical phenomenon of *equilibrium*. This

---

<sup>344</sup> Ashby 1947, p. 44. Ashby was more specific regarding this problematic duality five years later, at the beginning of the 1952 –first edition—of his *Design for a Brain* book:

On the one hand the physiologists have shown in a variety of ways how closely the brain resembles a machine: in its dependence on chemical reactions, in its dependence on the integrity of anatomical paths, and the precision and determinateness with which its component parts act on one another. On the other hand, the psychologists and biologists have confirmed with full objectivity the layman’s conviction that the living organisms behaves typically in a purposeful and adaptive way. These two characteristics of the brain’s behavior have proved difficult to reconcile, and some workers have gone so far as to declare them incompatible (Ashby 1952, p. 1).

<sup>345</sup> Ashby 1947, p.44



should, for him, provide an objective approach to adaptiveness, void of subjective appreciations and metaphysical preconceptions. For this, he began scrutinizing the said notion of equilibrium, incrementally building up on top of some basic propositions.

For heuristic purposes, Ashby encouraged the reader to imagine three physical objects, each showing a distinct type of equilibrium: a *cube* lying on one of its sides (stable), a non-moving *sphere* (neutral) and an inverted *cone* resting on its tip (unstable). If we try to tilt the cube, it will resist before it reaches 45 degrees; after that, it will topple over. If we push the sphere, it will roll. If we touch the inverted cone, it will fall on its side. In regards to the resulting distribution of forces, one could say that in the cube's case the disturbing force acts against the resultant force; in the case of the sphere the resultant force *is* the disturbing force (there is only disturbing force); and in the case of the inverted cone the resultant force acts along with the disturbing force.

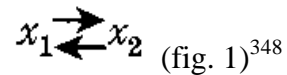
This quantitative approach to what equilibrium stands for suggests two important features: 1) Equilibrium does not imply immobility: a perfect pendulum, with a frictionless hinge, stands for a system in equilibrium. 2) Equilibrium is inherently dynamic: if the three objects alluded above were unaffected by any disturbing force, ascribing any type of equilibrium present would hardly make sense --one could say at most that Newton's First Law is being witnessed.<sup>346</sup>

At this point, Ashby was ready to more clearly specify what equilibrium is, advancing that "a variable is in stable equilibrium if, when it is disturbed, reactive forces are set up which act back on the variable so as to *oppose* the initial disturbance. If they go with it then the variable

---

<sup>346</sup> "Every body persists in its state of being at rest or of moving uniformly straight forward, except insofar as it is compelled to change its state by force impressed" (Newton 1687, p. 416).

is in unstable equilibrium”.<sup>347</sup> Once he accomplished this clarification, Ashby made an interesting move. He proposed that every system in equilibrium implies a circuit. How could this be so? The system of opposing forces reveal a circuital flow: a variable that is affected by the disturbing force affects back the disturbing force itself. Graphically, it could be displayed as something like this:



Moreover, Ashby further stated that the reverse is also true: that the existence of a circuit entails an existing equilibrium! And this is so because the flow of disturbance to the variables also denotes a system with a reciprocal interaction of forces.  $X_1$  affects  $X_2$ , which in turn affects  $X_1$ ; once we compare the range of change, we can find a stable equilibrium --or not, in which case it would be an unstable one.

In continuing with this line of reasoning, and building upon this found range of variability, Ashby stated that all systems in equilibrium have a neutral point. This is pointed out as an important claim, since the stable equilibrium of a system would depend on the amount of variation away from this neutral point: if the disturbance pushes the system out of the said neutral threshold, then it becomes unstable. In fact, this ontological leash, within which the system can still be said to be in stable equilibrium, is given by Ashby the name of “range of stability”.<sup>349</sup>

Since the whole point of the argument is to adequately articulate the notion of a machine in order to defend its claimed isomorphism with the nervous system (or any other living organism), Ashby readily acknowledged that one likely criticism would be that such circuitry is

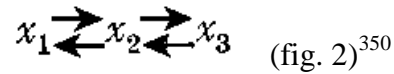
---

<sup>347</sup> Ashby 1940, p. 479

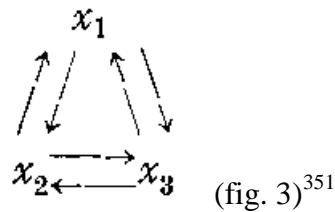
<sup>348</sup> Ibid., p. 480

<sup>349</sup> Ibid., p. 481

way too simple for it to have any hope of being compared to a living entity. To this, he advanced the response that such “simple” circuits can very quickly become complex --and in an extreme manner. Consider the following case:



Or even...



One could readily notice that the modification of just one variable will alter the whole system. And so, if we have more variables, and each one of them is itself modified, the compound circuitry can quickly engage in dramatic complexity. The adaptive circuitry of a natural machine (a living organism), thus, should no longer be regarded as belonging to a very high complexity of a different *kind* --it is rather different in a matter of *degree*. Ashby was reassured that what he found has serious implications for our properly mechanistic understanding of life. He exhorts us to see that...

there is a point of fundamental importance which must be grasped. It is that stable equilibrium is necessary for *existence*, and that systems in unstable equilibrium inevitably destroy themselves. Consequently, if we find that a system *persists*, in spite of the usual small disturbances which affect every physical body, then we may draw the conclusion with absolute certainty that the system must be in stable equilibrium. This may sound dogmatic, but I can see no escape from this deduction.<sup>352</sup>

---

<sup>350</sup> Ibid., p. 481

<sup>351</sup> Ibid., p. 481

<sup>352</sup> Ibid., p. 482

A system's equilibrium has to remain within the neutral point for it to remain stable, so that the "range of variability" occurring in its circuitry would allow for its proper adaptation to the changing environment. Ashby is well aware that when such process occurs in the natural world, it is called by a familiar name: *survival*. The mechanistic bridge with living organisms has now been established. The most important feature exposed by living entities, the very feature that accounts for its survival, would seem to be a characteristic of matter itself --as long as it is a coping system that remains stable. In Ashby's own words...

The moment we see that "adaptiveness" implies a circuit and that a circuit implies an equilibrium, we can see at once that this equilibrium must be of the stable type, for any unstable variable destroys itself. And it is precisely the main feature of adaptive behaviour that it enables the animal to continue to exist... Vast numbers of variables associated with the animal are all in stable equilibrium. Not only is this so as an observed fact, but it is clear that it must be so because any variable or system in unstable equilibrium inevitably destroys itself.<sup>353</sup>

Having realized this organism-machine identity in terms of structure, Ashby went on to deal more specifically with the characteristics of its coping behavior. In a subsequent article five years later,<sup>354</sup> he hypothesized about the possible processes that might take place when an organism copes with an altering environment in order to survive. He proposed that a living entity engages in a trial-and-error operation, altering variables in its inner circuit until it finally finds the right neutral 'sweet spot', where it can remain stable --*alive*-- exactly as when a coping machine reaches equilibrium.

The trial and error coping in the animal portrays the following characteristics: 1) It is called upon only when the environment provides a hostile milieu for the organism; 2) Each tried option remains for a determinate period of time, until it is no longer satisfactory; 3) When the

---

<sup>353</sup> Ibid., p. 483

<sup>354</sup> Ashby 1945

hostile environment remains, the trial and error dynamic does not stop; 4) The next attempted option is randomly chosen, not necessarily being a better one; 5) When a tried choice provides an adequate coping, then the system stops ‘hunting’ and remains stable. This would entail an animal successfully surviving.<sup>355</sup>

Ashby is here introducing an astonishing suggestion. This trial-and-error procedure, he proposes, is not unique to living entities; quite on the contrary, “it is an elemental and fundamental property of all matter”.<sup>356</sup> In fact, this biological operation finds its mechanical counterpart in the phenomenon of a machine “breaking”. To push this point through, Ashby deemed it necessary to squarely tackle the notion of a machine, carefully picking it apart. This might well be the first time in history that the concept of a machine *per se* was directly addressed.<sup>357</sup>

Ashby indeed had at this stage an incipient definition of a machine --one that would be increasingly fine-tuned throughout the subsequent years. A machine, he claims, is “a collection of parts which (a) alter in time, and (b) which interact on one another in some determinate and known manner”.<sup>358</sup> We are in front of characteristically machine-like features of change, reciprocal interaction, determinism and measurability. The parts can be called variables, since they are measurable. The alteration of these variables in themselves does not necessarily modify the overall constant of the system in equilibrium; however, a change in the organization of the

---

<sup>355</sup> Quoting Herbert Spencer Jennings, Ashby subscribes to the idea that

Organisms do those things that advance their welfare. If the environment changes, the organism changes to meet the new conditions... If the mammal is cooled from without, it heats from within, maintaining the temperature that is to its advantage... In innumerable details it does those things that are good for it (Jennings 1906, p. 338).

Ashby would also acknowledge thinking on the same vein as John Scott Haldane, who recognizes homeostatic endeavors of organisms in his studies on respiration: “Biology must take as its fundamental working hypothesis the assumption that the organic identity of a living organism actively maintains itself in the midst of changing external circumstances” (Haldane 1922, p. 391).

<sup>356</sup> Ashby 1945, p. 13

<sup>357</sup> See note above in this section on Hertz.

<sup>358</sup> Ashby 1945, p. 14

network of parts does. Thus the identity of a machine has to do less with changes in its parts, and more with the change in the constants of the overall array of variables—in other words, with its internal *organization*.

When a machine “breaks”, Ashby reminds us, the entity is still a machine --but another type of machine, since its internal organization has changed. A “break” is thus a change in the constant of the overall network of variables that conjointly forms the machine. And this break has a reason to exist. When a coping machine, due to an external disturbance, is pushed out of its neutral point of stability (equilibrium), its variables will individually change so that in group they compensate for the attempted alteration of the network. This compensation brings back the machine to its stable equilibrium, effectively steering it away from the overall effect of the disturbance:

A machine which has available an indefinitely large number of breaks depending on configurations closely similar to one another will inevitably change its internal organization spontaneously until it arrives at an organization which has an equilibrium with the special property that it avoids those configurations.<sup>359</sup>

When the disturbing force is stronger than what the variables are able to compensate --and thus collectively push back-- then a change in the overall network occurs. The constant behind the equilibrium point then changes to *another constant* --constituted by the overall change of altered parts—so that the machine can persist in stable equilibrium. Each break will send the array of parts hunting for a newly organized internal structure, until another constant is reached and a stable equilibrium is again found. If another constant is not eventually reached in a finite timeline, the machine does not break: it gets destroyed. It does not survive.

---

<sup>359</sup> Lack of ‘commas’ in the original. Ashby did not seem to be particularly fond of generous punctuation.

Now we can better see the ontological bridge with a living organism firmly set in place. If we recall, an animal has to re-wire its internal circuitry in order to cope with an altering environment. This is done by means of a trial and error procedure, its variables individually changing until the overall network of parts finds a stable equilibrium with the modified environment. This constitutes the characteristic animal adaptiveness. The attained equilibrium will persist until another change will make the organism “break”, once again until a new successfully coping self-organization is achieved, and so on. This is survival. And this is the manner in which a successfully coping machine persists. Organism and machine, in virtue of both having one quintessentially defining behavior in common --self-organization in order to survive-- can be said to be, effectively, one and the same.

## **2. The homeostat: A living machine.**

In an article published two years later,<sup>360</sup> another previously suggested point of no small importance for Ashby’s subsequent proposal is reinforced. A machine’s ontological counterpart in nature is, according to him, a living organism *and* its environment --the animal alone is an incomplete machine. *Both* organism and environment, engaged in dynamical interaction, constitute a machine. In fact, that will be a key distinction for Ashby’s next conjecture regarding the fact that a certain type of living organism (the nervous system) successfully adapts to change --rather than being caught up in an endless internal chaotic loop, entailing what certainly occurs at times: death.

Having already established the isomorphism between working machine and living being, Ashby did not waste time in expressing a keen interest in having a particular type of organic

---

<sup>360</sup> Ashby 1947

entity understood as a machine: the nervous system (and its environment). In it, he advances, “the ‘variables’ are mostly impulse-frequencies at various points in the nerve-network. And as the impulses at one point affect in various ways the impulse-frequencies at other points, the nervous system is a dynamic system par excellence”.<sup>361</sup> Citing empirical studies of his time that in his mind were suggestively friendly to his view, Ashby felt that he succeeded in providing a full machine-based explanation for how (at least basic) living entities survive.

Apparently not entirely content with what was accomplished theoretically, in elegantly cybernetic fashion Ashby had only one year later (1948) publicly acknowledged having constructed a machine that would instantiate his hypothesis “in the flesh”.<sup>362</sup> He built what he baptized as the Homeostat --with obvious reference to his intellectual mentor Walter Canon’s studies in biological homeostasis. This machine, which was to become a celebrity among hardcore cyberneticians,<sup>363</sup> was built from WW2 spare military parts.

Echoing a salient characteristic of British cybernetic pursuits, Ashby’s machine was gradually built during his spare time, partly at home and partly at the storage area of the hospital where he worked. There was no formal institutional funding behind its construction.<sup>364</sup> The apparatus consisted of four Royal Air Force bomb control switch gear kits<sup>365</sup> mounted on four aluminum boxes, each of which connected to a pivoting needle with a concave trough filed with conducting liquid (water). Each control gear switch, or uniselector, was previously hardwired

---

<sup>361</sup> Ashby 1947, p. 57

<sup>362</sup> “In the metal”, to be precise: Ashby started constructing it in 1946 and presented the machine at a meeting of the Electroencephalographic Society of Britain in 1948 (Ashby 1949).

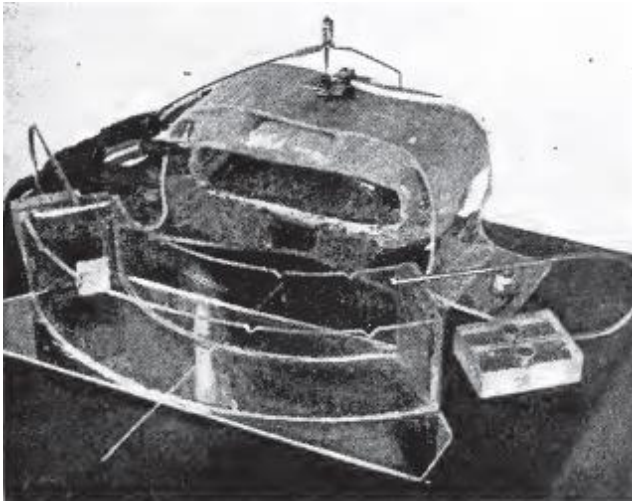
<sup>363</sup> Margaret Boden’s view, as opposed to the famed tortoises build by Grey Walter, which enjoyed widespread fame among the general public (Boden 2006, pp. 228-230).

<sup>364</sup> Eventually Ashby was given a professorship at the University of Illinois at Urbana-Campaign, but at this time it is safe to say that “classical” cybernetics, was already spiraling into oblivion. However, funding for his second machine (section 5, below) was secured (Pickering 2010, p. 310).

<sup>365</sup> Ashby 1928-1972, p. 2432. These are stepping switches or uniselectors –electromechanical devices that allow for the input and output of electric current via a chosen built-in channel at a time.



with random values,<sup>366</sup> constituting the particular threshold that the current has to reach –or not—in order to be let through. The four units were electrically interconnected among themselves. Each had 25 possible combinations. When the current of electricity passed through them, the overall result was that each of the uniselectors' floating needles would reach a middle point in the trough, and remain there. The pictures below show, in order, each uniselector, and the four of them inter-connected forming the whole homeostat.<sup>367</sup>



(fig. 4)<sup>368</sup>

---

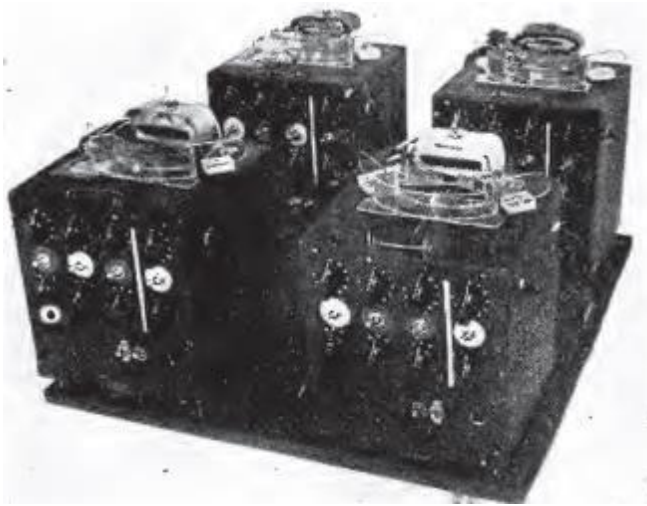
<sup>366</sup> These were taken straight from a table of random numbers that Ashby found in Fisher and Yates's *Statistical Tables*, originally published in 1938 (Fisher & Yates 1963, pp. 134-139).

<sup>367</sup> Ashby describes the construction of his machine in the 2nd –substantially revised-- edition of his book *Design for a Brain* as follows:

The Homeostat... consists of four units, each of which carries on top a pivoted magnet... The angular deviations of the four magnets from the central positions provide the four main variables. Its construction will be described in stages. Each unit emits a D.C. output proportional to the deviation of its magnet from the central position. The output is controlled in the following way. In front of each magnet is a trough of water; electrodes at each end provide a potential gradient. The magnet carries a wire which dips into the water, picks up a potential depending on the position of the magnet, and sends it to the grid of the triode... If the grid-potential allows just this current to pass through the valve, then no current will flow through the output. But if the valve passes more, or less, current than this, the output circuit will carry the difference in one direction or the other...

Next, the units are joined together so that each sends its output to the other three; and thereby each receives an input from each of the other three... But before each input current reaches its coil, it passes through a commutator..., which determines the polarity of entry to the coil, and through a potentiometer..., which determines what fraction of the input shall reach the coil. As soon as the system is switched on, the magnets are moved by the currents from the other units, but these movements change the currents, which modify the movements, and so on (Ashby 1960, pp. 100-102).

<sup>368</sup> *Ibid.*, p. 101



(fig. 5)<sup>369</sup>

The peculiarity of the machine relied on the following feature. If one were to modify the connections, say “re-wire” the machine in another way –inverting the current, connecting the uniselectors in different order, messing with the needles, turning off one of the uniselectors, etc.—the needles would appear to go hunting for a different position, trying out different states, until they would all settle back into each one of their middle points in the trough. Ashby’s colleagues would find themselves frantically rewiring the machine several times, in different ways, just to witness somewhat in awe how the machine would “re-organize” itself, gradually returning to its initial point of stable equilibrium at the center –equidistant from either extreme, 45 degrees at each side. Ashby would later succinctly describe its behavior like this:

On top of each box is magnet [*sic*] which can turn on a pivot. The central position counts arbitrarily as “optimum”; at 45° on either side is the “lethal” state the brain must avoid, and it is at that position that the relay closes to make the unselector change to a new position. If it is

---

<sup>369</sup> Ibid., p. 101

stable, the needles come to the center; if they are displaced, they will all fluctuate but they will come back to the center. That behavior corresponds to the behavior of the adapted organism.<sup>370</sup> As we can see, Ashby expectably regarded this behavior as being one and the same with that of a living entity *surviving*. Following his previous understanding of a dynamical machine as conformed by both animal *and* environment, part of the homeostat's configuration would be the organism per se, and the other would be its threatening milieu.<sup>371</sup> In fact, the titles of the first mainstream publications that featured the machine were "The Electronic Brain"<sup>372</sup> and the "The Thinking Machine"<sup>373</sup> --in *Radio Electronics*, a contemporary popular electronics magazine and the famed *Time*, respectively.

Three years later, Ashby's book *Design for a Brain* (first edition 1952) gave a preeminent role to his machine, locating it almost at the axis of how and why we should see the machine-based hypothesis as very likely true: It is presented as an instantiation of a theorem. That same year, equipped with his "cybernetic monster",<sup>374</sup> he was finally invited by the "cybernetic group" to the Macy conferences in New York. This occasion will turn out to be the first and last time that he had the opportunity to attend the cybernetic Mecca of intellectual exchange. As mentioned in chapter II, of the ten conferences that took place in a span of less than a decade (1946-1953), only the last five enjoyed published proceedings. Since this was the penultimate one, we have a fairly detailed account of what happened on the day of his presentation.<sup>375</sup> It is helpful to complement it with Ashby's personal diary entries,<sup>376</sup> in an effort to shed light on what

---

<sup>370</sup> Ashby 1953, p. 97

<sup>371</sup> What part of the machine would be the animal and what part would be the environment was entirely up to the operator. We could regard, say, one uniselector as a small animal and the remaining three as a big environment, or vice versa. In fact Ashby referred to the Homeostat as "a machine within a machine" (ibid, p. 96).

<sup>372</sup> Ashby 1949b

<sup>373</sup> Ashby 1949c

<sup>374</sup> Pickering 2005

<sup>375</sup> The meeting took place on Thursday, March 20<sup>th</sup>, 1952 --and it was subsequently published as Ashby 1953.

<sup>376</sup> Ashby 1928-1972, pp. 3732-3736, 3738-3744.

took place that day. What happened in there represents a somewhat strange and unique occurrence in what regards 20<sup>th</sup> century science gatherings.

### **3. The Macy presentation: Keeping cybernetics honest.**

Right from the opening remarks of his talk, Ashby did not hide both the bold epistemic status of his proposal –namely, understanding the essence of life as identical with machine behavior – and the audaciously attributed ontological status of his machine –as inhabiting a middle-ground between the living and the non-living...

I must say unambiguously what I mean by ‘alive’. I assume that if the organism is to stay alive, a comparatively small number of essential variables must be kept between physiologic limits. Each of these variables can be represented by a pointer in a dial... This, of course, is in no way peculiar to living organisms. An engineer, seated at the control panel in a ship, has exactly the same task to do. He has a row of dials some of which represent essential variables in the ship, and it is his business to run things so that the needles always stay within their proper limits. The problem, then, is uniform between the inanimate and the animate.<sup>377</sup>

Ashby went to great lengths in trying to explain what he was about to show. He first wanted to set the proper context so that when the homeostat was introduced, the participants were already familiar with the issue to be addressed --namely “how the organism manages to establish homeostasis in the larger sense”.<sup>378</sup> This was not an easy feat. Ashby had to carefully convey, in a very succinct way, roughly the gist of the research mentioned at the beginning of this chapter. An important aspect of this aim was to establish the feasibility of an unambiguous setting of the

---

<sup>377</sup> Ashby 1953, p. 73

<sup>378</sup> Ibid.

problem –task which some participants clearly saw and agreed upon, while others did not quite get.<sup>379</sup>

Ashby started asking whether or not an animal's changing environment can be specified in a clear and rigorous way. The emphasis would not be placed on the environment's complexity, but on its degree of determinateness. This would allow him to go on with his proposal. But for this to happen, Ashby had to first agree on that point with the audience. And despite the common cybernetic credo of the conferees, it encountered some resistance at first,<sup>380</sup> probably due, to some extent at least, to the difficulty in understanding where Ashby was actually going with all these. Eventually they settled in agreeing that a sharp and clear cut structuring of the problem was indeed possible...

*Ashby:* Let the environment be represented by the operator, E. The organism's problem is to convert the brain into an operator, which might be represented by  $E^{-1}$ ... The formulation with E and  $E^{-1}$  is merely meant to state the problem in a form that is clear physically and mechanistically...  $E^{-1}$  is fairly straightforward, if  $E^{-1}$  does not exist, the organism cannot find a solution and dies. If it does exist, then the animal may perhaps find several solutions. But what I want to know is whether E can be given a fully rigorous definition.<sup>381</sup>

---

<sup>379</sup> Let us recall that the Macy conferences were radically interdisciplinary, its participants coming from several distinct fields encompassing mathematics, engineering, biology, psychology, anthropology and others. Thus although all were sharing a cybernetic interest, they were not all comfortable with Ashby's terminology. Throughout the presentation, one could notice some participants struggling with getting the conceptualization of the homeostatic problem right. Ashby was merely preparing the context for the introduction of the machine, and complications seemed already happening (See Ashby 1953, pp. 73-75).

<sup>380</sup> Although eventually the majority of the audience agreed with Ashby on the ontological feasibility of accepting the variables E and E-1, one could perceive that the trained engineer Julian Bigelow was going to be Ashby's nemesis throughout his talk. Regarding Ashby's plea to accept the ontological feasibility of the variables E and E-1 in order to continue with the presentation of his proposal, Bigelow had this to say:

*Ashby:* The chief thing that I'm interested in is the question, if this system is grossly non-linear and if there are delays... can this operator E be given a meaningful and rigorous form?

*Bigelow:* In general, no (Ashby 1953, p. 75).

<sup>381</sup> One should guard oneself from being misled into thinking that Ashby means 'mechanistic' in a solely heuristic way. In accordance to the rest of his philosophy, when an *explanans* was rigorously mechanical, the *explanandum*

*Wiesner:* <sup>382</sup> That is an operation we would agree it can be done.

*Pitts:* It is certain that it could be.

*Ashby:* You think it could be?

*Wiesner:* By definition, yes. <sup>383</sup>

Subsequently Ashby shared the reason why focusing right from the start on the adequate addressing of the environment was a mandatory requirement. Without rigorously treating, articulating and then factoring in the environment little would afterwards be intelligible when talking about the brain. Thus, a deep study of the organ ought to have an equally profound study of its corresponding containing milieu. He claimed that

there can't be a proper theory of the brain until there is a proper theory of the environment as well... Without an environment, 'right' and 'wrong' have no meaning. It is not until you say, 'Let's give it  $E_{47}$ , let's see what it will do against that', it is not until you join the brain on to a distinctly formulated environment that you begin to get a clear statement of what will happen. <sup>384</sup>

Ashby's position regarding the nature of environment would appear quite reasonable to some, while utterly unrealistic to others –depending on one's philosophical leniencies. The animal environment's elements were found by Ashby to be existent, at least temporarily, in a physically discrete and non-continuous manner. However, the Humean chunks he was dividing reality into were not to be understood as effectively isolated from each other; rather, they were to be regarded as *not totally and continually* interconnected among themselves...

There is a form intermediate between the environment in which everything upsets everything else

---

happened to be a machine. The behavior of the entity, and not its structure, determines whether or not that thing is a machine. That is part of the crux of his contribution. And that is also the source of the problems to be addressed below.

<sup>382</sup> Jerome Wiesner (1915–1994) received his PhD in Electrical Engineering from the University of Michigan and went to work at MIT's above mentioned Rad Lab. He was a science consultant for the Kennedy administration. Later in his life, much like Norbert Wiener, he became a vociferous advocate against war.

<sup>383</sup> Ashby 1953, pp. 75-76.

<sup>384</sup> *Ibid.*, p. 86

and the environment which is cut into parts. This intermediate type of environment is common and is of real significance here; it is an environment that consists of parts that are temporarily separable, and yet by no means permanently separable...<sup>385</sup>

This created some tension in the audience. Bigelow replied that “[o]ur environment doesn’t consist of that sort of phenomena in very many ways. The statement you are making amounts to something that mathematically sounds independent but does not exist in the real world”.<sup>386</sup> It may have been the case that Ashby was not being properly understood, since he did make the remark (above) that environmental factors, even if we can dissect them to some extent, are “yet by no means permanently separable”.<sup>387</sup> And this misunderstanding may have not been exclusive to Bigelow. For instance, Heinrich Klüver<sup>388</sup> politely interrupted to say:

Your scheme should be of particular interest to psychologists, because unlike other models we have encountered here, it stresses environmental factors and the interrelations between animal and environment. I did not quite follow what was implied in your statement that constancies may function as barriers.<sup>389</sup>

As addressed at the beginning of this chapter, once the coping machine, in trying to accommodate itself to the changing environment, reaches an overall stable equilibrium, this attained state remains constant, until another change is such that the parts/variables are obliged to go hunting again --in which case we can say that the machine goes through a “break”. But until then, this threshold, this “ontological leash” that these parts enjoy, allow them to tolerate and push back the changes, until the overall pressure is such that a break has to occur --so that the animal can again rewire itself. Also, having in consideration that in order to have a flow of

---

<sup>385</sup> Ibid., p. 82

<sup>386</sup> Ibid., p. 84

<sup>387</sup> Ibid., p. 82 (quoted above).

<sup>388</sup> Heinrich Klüver (1897–1979), German-American psychologist who received his PhD from Stanford University.

<sup>389</sup> Ibid., p. 86

information, “change” must be present, a variable locked into an achieved constancy will effectively block the flow of information.<sup>390</sup> Accordingly --and putting emphasis on the hunting parts-- each trying to be adequate to the changing circumstances, Ashby patiently answered that...

If a system is composed of variables that often get locked constant, it tends to cut the system into functionally independent sub-systems which can join and separate. Instead of being a totally interlaced system with everything acting on everything else, it allows sub-systems to have temporary independencies.<sup>391</sup>

And hopefully having cleared out the point, Ashby went on with his presentation, just to be interrupted with clarifying questions by other participants every so often, which he patiently addressed. The most engaged into the discussion were three authors already familiar to us, (McCulloch, Pitts, and Bigelow), and Wiesner. Ashby, probably acutely self-aware of the next bolder step he was about to take, did not hesitate in answering each and every question -- sometimes briefly, other times lengthily, but always firmly.

What could be deemed in contemporary academic ambiances as a hostile milieu for debating was seemingly just part of a typical heat of the discussion in the 1950s.<sup>392</sup> In fact, not all remarks showed puzzlement, on the spirit of “what is it that he is talking about”. Some were enthusiastic about the wild cues that were seemingly being suggested. Gerhardt von Bonin,<sup>393</sup> for example, in referring to Klüver’s above participation and Ashby’s response to it, added sympathetically...

*Von Bonin:* To pick up the remark Klüver made a moment ago, the animal appears to break down

---

<sup>390</sup> As seen in chapter IV, section 2, Claude Shannon’s foundational paper (Shannon 1948), despite the secrecy that surrounded it when it was first written, was already very well known in the cybernetic circle.

<sup>391</sup> Ashby 1953, p. 86

<sup>392</sup> Ashby wrote down of this day in his diary as being “highly successful” (Ashby 1928-1972, p. 3732).

<sup>393</sup> Gerhardt von Bonin (1890-1979) was a German-American neuroscientist, professor at the University of Illinois.



the environment into certain patterns which seem to develop in his mind. Something goes on in its brain and then that structures the environment. It perceives the environment in a certain pattern, set by its brain, so that it can deal with it.

*Ashby*: I assume that the brain works entirely for its own end.

*Von Bonin*: Oh surely.

#### **4. The backlash: Unforeseen consequences of a behavior-based ontology.**

Some of the members attending the talk were taken aback by the claims of the homeostat behaving isomorphically to an animal in reciprocal interaction with its environment. It was not only that unlike Grey Walter's tortoises and Claude Shannon's mice this machine did not formally resemble any actual animal. What worried them was what was being suggested as entailing from the machine's random looking for a stable equilibrium after its uniselectors were tampered with. Even if they would be willing to accept that some primitive organisms do indeed resort to random trials in order to cope with a changing milieu, it seemed that the whole image of four interlocking uniselectors mounted on a table behaving animal-like was more than what they were willing to accept. Bigelow, somewhat in confusion expressed that the homeostat "may be a beautiful replica of something, but heaven only knows what".<sup>394</sup> Others sided with him...

*Hutchinson*:<sup>395</sup> Do you interpret the ordinary sort of random searching movement of one of the lower animals as something that is comparable to the uniselector?

*Ashby*: Yes.

*Hutchinson*: It doesn't seem to me that the setup gives any particular clear suggestion that this is comparable to the way invertebrates do behave.

---

<sup>394</sup> Ashby 1953, p.95

<sup>395</sup> George Evelyn Hutchinson (1903–1991) was a British zoologist and ecologist who spent most of his professional life teaching at Yale University –although he never earned a PhD.

*Bigelow: Agreed.*<sup>396</sup>

It would seem that some metaphysical attachment to the nature of living beings was still not being let go. And this situation was happening among those very individuals whose aim, as proper cyberneticians that they were, was to dismantle any metaphysical remnants behind a vitalist impetus that might differentiate in kind --and not degree-- a living organism from a machine. Ashby seemed to have noticed it. Without a sign of dubitation, he was firm in eroding piece by piece the seemingly still entrenched epistemic attitudes being witnessed at work. Faithful to “cybernetically correct” belief, Ashby did not lose focus and rehashed on *that* aspect of both machine and organism that is core of cybernetics parlance: *behavior*. For all ends and purposes, an organism behaves in the way the homeostat behaves --literally.

Beginning from Bigelow’s (and company) 1943 foundational article,<sup>397</sup> that pillar of cybernetics was never shaken. Ashby merely rehearsed what they already knew and indeed, defended<sup>398</sup> –albeit probably unaware of the ultimate consequences that would naturally emerge.<sup>399</sup> Randomness in animal adaptation was probably *the* characteristic that helped the most in establishing the ontological link between machine and living being. That was the way in which a living organ, as adaptationally sophisticated as a brain, could be said to behave in a machine-like manner. Machines can certainly adapt: the name for that in the physical world is equilibrium, as Ashby was careful to establish in his first published cybernetic article mentioned

---

<sup>396</sup> *Ibid.*, p. 98

<sup>397</sup> Rosenblueth, Wiener & Bigelow 1943

<sup>398</sup> Dupuy 2000, pp. 148-155

<sup>399</sup> As seen above in this chapter, it would seem that Ashby had faithfully adhered to cybernetic ideas in a manner that preceded Wiener himself, judging from his 1940 article (Ashby 1940). However, importantly for this work, Ashby does acknowledge that a profound, fundamental insight into the nature of a machine is first given by Wiener:

I see that my idea that maths is a dynamic process, and that the Russelian and other paradoxes are simply the machine going to a cycle, has already been stated by Wiener. Maybe I got the idea from him.

Ashby goes on to quote a line from Wiener’s *Cybernetics*: “...the study of logic must reduce to the study of the logical machine,...” (Ashby 1928-1972, p. 6023).

above.<sup>400</sup> And so, he pushed the idea of randomness as properly animal, and properly homeostatic, until the end...

*Hutchinson:* I am not altogether convinced that that searching [an ant looking for sugar cubes dropped on a table] is in any way comparable to what we have seen here.

*Ashby:* But don't you agree that if you disturb a living organism with some vital threat and the organism takes steps to bring itself back to normal physiological conditions, that is perfectly typical of vital activity?

*Hutchinson:* Yes, but I think when it does that, there is not the fluttering oscillation that you first described. Isn't that inherent in the system?

*Ashby:* Well, some sort of random trial is necessary. Surely, random trying in difficult situations has been described over and over again from Paramecium upwards.

(...) *Quastler:*<sup>401</sup> I believe Dr. Ashby would not claim that random searching was the only mechanism by which an organism finds a normal or favorable condition.

*Ashby:* I don't know of any other way an organism can do it.<sup>402</sup>

A further point of contention derived from the above remarks was the issue of learning as an outcome of a random process. They seemed to be uncomfortable with the idea that a change of behavior --a resulting accommodation-- coming out of a random search could count as learning proper. Once again, it seemed that the very proponents of the treatment of machines and animals under the same physical laws could not foresee the metaphysical backlash that could follow. In cybernetics, let us recall, behavior is the sole and privileged communicator of any change of a state in a given living entity. In Bigelow's (and company) 1943 article, one of its principal goals is indicated as attempting to "define the behavioristic study of natural events and to classify

---

<sup>400</sup> Ashby 1940

<sup>401</sup> Henry Quastler (1908-1963) was an Austrian-American physician and radiologist at the University of Illinois. After WW2 he did pioneering research on the effect of radiation in living tissue.

<sup>402</sup> Ashby 1953, p. 99

behavior”.<sup>403</sup> This most important notion is subsequently referred to as “any change of an entity with respect to its surroundings”.<sup>404</sup> Further clarifying, it is claimed that “any modification of an object, detectable externally, may be denoted as behavior”,<sup>405</sup> enriching the otherwise reductive understanding of behavior coming from psychology.<sup>406</sup> A perfectly feasible outcome from this established framework is accepting as meaningful the variation from state A to state B in a given organism after it deals with a change of its environment. It did not take much for associating the semantics of a learning process –itself already understood behaviorally, independently of cybernetics—to the meaningful change of state of a machine. Change of behavior after coping with a modifying milieu *is* learning. And this necessary development is what makes the following lines somewhat striking in their naïveté, and fascinating when confronted with Ashby’s calm, perfectly cybernetic-abiding answers:

*Bigelow:* Sir, in what way do you think of the random discovery of an equilibrium by this machine as comparable to a learning process?

*Ashby:* I don’t think it matters. Your opinion is as good as mine.

*Bigelow:* But I’m asking, if you place an animal in a maze, he does something which we call learning. Now, in what way does this machine do something like that?

(...) *Wiesner:* I think that [the mechanical mouse of] Shannon and Marceu came close to having a learning machine.

*Bigelow:* That is the point. This machine finds a solution, I grant you –and please call it anything you want; I’m not trying to criticize your language. I merely wonder why finding a

---

<sup>403</sup> Rosenblueth, Wiener & Bigelow 1943, p. 18

<sup>404</sup> Ibid.

<sup>405</sup> More specifically, it is stated:

Given any object, relatively abstracted from its surroundings for study, the behavioristic approach consists in the examination of the output of the object and of their relations of this output to the input... The above statement of what is meant by the behavioristic method of study omits the specific structure and the intrinsic organization of the object (ibid.).

<sup>406</sup> See chapter I, section 2 of this work.

solution necessarily implies that it learns anything, in your opinion.

*Ashby*: I think that the word learning, as we understand it in the objective sense, without considering anything obtained introspectively, is based on observations of this sort of thing happening.<sup>407</sup>

Ashby seemed perfectly aware that the pillars upon which he was relying were as strong as they could get, in terms of being flawlessly consistent with essential cybernetic tenets. It is as if he is clarifying to the cyberneticists themselves what their own beliefs entail –besides letting them know at the same time how perfectly comfortable he is with those underlying frameworks.

Further, Ashby seems willing to champion these tenets and apply them all the way down, regardless of the consequences. He does not look in the least afraid of forever losing a metaphysical anchor that might had been in some non-scientific way beneficial for the sub-group of machinery of which he is himself part. Faced with the awkward complaints being generated by the cybernetic group itself, Ashby cut clean through their malaise, and confronted them with the expectable effect of their own machine-based ontology and behavior-based epistemology...

*Young*:<sup>408</sup> [The homeostat] doesn't show change of state. The essence of learning is that the system that has been through a procedure has different properties than those it had before.

*Ashby*: Would you agree that after an animal has learned something, it behaves differently?

*Bigelow*: Yes.

*Ashby*: Well, the homeostat behaves differently.<sup>409</sup>

---

<sup>407</sup> Ashby 1953, p. 103

<sup>408</sup> John Zachary Young (1907–1997) was a British zoologist and then neurophysiologist, who applied research in animal nervous system to injured humans after WW2.

<sup>409</sup> Ashby 1953, p. 103. Bigelow was, however, still reticent right until the quoted above behavioristic gauntlet was thrown. He told Ashby:

...if you have a box with a hole in the bottom and a ball bearing inside and keep banging at it, the bearing will fall through. That is in your sense, a way of learning (Ashby 1953, p. 104).

To this, Pitts interjected:

The homeostat is a bit better than that (ibid.)

This seemed to do it. The change in attitude was almost immediate. Despite the painful remarks thrown around that marred almost the entirety of Ashby's talk, agreement became suddenly consensual, to the point that the audience, almost leaving Ashby aside, took the role of vindicating his proposal. As heated as debates can come, the audience took up the torch and continued engaged in conversation among themselves, not even allowing the speaker to formally end the presentation.<sup>410</sup> All things considered, and checked out with the core of cybernetic mandates, they had to agree with the claimed facts that 1) the behavior of the homeostat is identical to that of an organism coping with its changing environment; and 2) the homeostat, for all intents and purposes (literally), learns. And so they became an audience geared into acceptance mode:

*Bigelow*: ...Do you consider this a learning device as it stands?

*Pitts*: Yes, I should say so.

*McCulloch*: Would you consider Shannon's mechanical mouse a learning device?

*Bigelow*: Surely.

*McCulloch*: Then this is a learning device in just that sense Shannon's mouse learns how to run  
*one* maze.<sup>411</sup>

(...) *Bateson*: I should like to put a question to our ecologist: if an environment consists largely of organisms, or importantly of organisms, is not the learning characteristic of Ashby's machine approximately the same sort of learning as that which is shown by the ecological system?

*Hutchinson*: Yes, definitely it is.<sup>412</sup>

These were the last interventions. Ashby wrote in his diary that the presentation was "highly successful"<sup>413</sup> –although acknowledging the distress of some.<sup>414</sup> From a historical perspective, it

---

<sup>410</sup> As per the transactions (Ashby 1953, p. 108).

<sup>411</sup> Ashby 1953, p. 105. Italics added.

<sup>412</sup> Ibid., p. 106

could be said that this conflicted exchange of ideas set up the environment for what could *effectively* count as the last meeting for the cybernetics group. The next meeting lacked so many key members and was so scattered and out of focus, that the editor of the proceedings refused to publish them --eventually, they were indeed published but only the papers, without the discussions.<sup>415</sup> One cannot help but wonder if that meeting was a (probably more than) symbolic *cout d'état* for a body of ideas that was already receiving a battery of attacks from different flanks.<sup>416</sup> What happened that day might have been just the tip of the iceberg regarding the deeply problematic ontology on which cybernetics was resting. That ontology was already leaking concerns at various levels of the cybernetic enterprise --however they were not identified as a whole, and anyhow there never was a unified effort to patch them.

Four years later Ashby would give a more detailed account of the mechanism behind the homeostat, paying special attention to the problematic feature of its random search. As footnoted above, this feature used a published table of random numbers by Fisher and Yates. The fact that the picked numbers from the table were fixed in the machine, occasioned some trouble in the audience, in regards to the veracity of the claim of the machine searching randomly. However, this fixation was also an essential part of the necessarily deterministic nature of the machine, as

---

<sup>413</sup> Note above (Ashby 1928-1972, p. 3732). Ashby seemed to be particularly delighted with Walter Pitts' siding with him in regards to the proper framing of the homeostatic problem specified above that his machine seemed to portray:

"Walter Pitts assured me that E is certainly well defined if it is assumed to be some neutral physical mechanism, for the latter's factuality ensures proper definition" (ibid., p. 3738).

<sup>414</sup> For instance, he says this of Julian Bigelow's intervention: "At the Macy Conference, Bigelow was very insistent that there was no virtue in random searching as against systematic searching" (ibid., p. 3744).

<sup>415</sup> See chapter I, section 4, towards the end.

<sup>416</sup> As one could induce from the conflictual nature of the Macy debates (Chapter I, section 4, regarding the 7<sup>th</sup> Macy Conference). Admittedly, it could also be seen as the inherently difficult nature of the interdisciplinary trading of ideas.

Ashby indicated at some point.<sup>417</sup> But it was clear that the idea of order achieved out of chaos was not sitting well among the cyberneticians present at that gathering.<sup>418</sup>

## **5. Un-cybernetic DAMS: From behavior into structure.**

Another issue that Ashby would go on to treat was related to the criticism of the homeostat being effectively too primitive to mimic higher vertebrates' behavior—specifically, mammals. One that elicited his agreement was the claim that the homeostat probably was indeed capable of autonomously reaching self-equilibrium due to the fact that, after all, 4 uniselectors combinable in 25 different ways among themselves (and accounting for a bit more than 390,000 possible combinations), would still fall very short of representing the extreme adaptive richness of higher animals. Ashby acknowledged the input and granted this point. In fact, he already was considering constructing something capable of more complexity than the homeostat—with more possible internal combinations than the homeostat's less than 400,000—although operating under the same principles. Ashby was aware that he needed to construct a machine with more complexity: one that, if possible, should show a behavior that would genuinely surprise its designer—show something that was not put there by the inventor. Indeed, almost two years before of the presentation, he already began thinking about constructing a more complicated machine.

---

<sup>417</sup> Jerome Wiesner took the “random search” for stability performed in the homeostat as grounds for stating that the machine was “randomly operated”—which it was not. There were clear thresholds in each stepping switch, determined by hardwired random numbers that had to be reached, in order for the current to go through to the next uniselector, until they would all return to a stable state of equilibrium. Thus, Ashby interjected: “Just a minute. I don't like the term ‘randomly operated.’ The machine's behavior depends, determinately, on whether its essential variables are inside or outside their proper limits” (Ashby 1953, p. 96).

<sup>418</sup> However, this epistemic position might have already crawled within cybernetics, as it is noted by Ashby's total comfort with the idea. See chapter IV, section 2 (towards the end).



I have to note a critical day. Ever since the homeostat was built (May '47) I have been thinking about the next machine, trying to find a suitable form for a more advanced machine.<sup>419</sup>

Its longer name was going to be *Dispersive And Multi-stable System* (DAMS). The report on the construction of such a machine appeared at the footnote of the first edition of his *Design for a Brain*—and disappeared in later editions.<sup>420</sup> Not much is known about this machine, other than the entries recorded in his diary.<sup>421</sup> The idea behind giving DAMS a more complex nature relied on the possibility of letting the machine isolate some of its parts when they would have reached adaptation, and letting the rest of the machine keep hunting for other “friendly” variables. What would be gained from this would be that it would no longer be necessary to re-wire the whole machine to prepare it for the next task. As seen in the previous section, given the benefit of the doubt, Ashby was granted that the homeostat portrayed the same level of learning behavior of the cybernetic mice, which could sort out a maze *once*, before getting ready to get re-wired for another task.<sup>422</sup> The neon tubes installed in DAMS allowed electric current to go through just when they would reach a certain threshold, remaining inert and uncondutive when the threshold was not met. This was an improvement over the homeostat’s wires --which were always conductive-- and affected the whole system. Always. Truer learning could allegedly be achieved when already reached adaptations were kept and other portions of the organism kept coping. However, finding the right form of design for this more sophisticated device was not an easy feat

---

<sup>419</sup> Ashby 1928-1972, p. 2950

<sup>420</sup> Pickering 2010, pp. 122-123

<sup>421</sup> Ashby 1928-1972, Other Index, 4: DAMS.

<sup>422</sup> The excerpt of the relevant dialogue portrayed above reads thus:

McCulloch: Would you consider Shannon’s mechanical mouse a learning device?

Bigelow: Surely.

McCulloch: Then this is a learning device in just that sense Shannon’s mouse learns how to run *one* maze (Ashby 1953, p. 105. Italics added).

(The reason why I previously italicized the word “one” should now become apparent)

for Ashby –in fact, it eventually proved to be more than what he could handle.<sup>423</sup> At a certain point, believing that being stuck in the “design stage” of the construction ironically betrays the whole purpose of a self-accommodating ultra-stable system, he foregoes the design aspect, puts the pieces together, and waits to see what happens...

Today it occurred to me that this is not the proper way. According to my own ‘theory’ all this design is irrelevant; what I should do is to throw the design to the winds, join all at random, and trust. I should not attempt to follow the consequences of the design right thorough; for if I do that I’m forcing the design down to an elementary level. If I am to break right through into a new world of performance I must find what my theory really needs, attend only to that, and then let the rest of the machine be assembled literally at random.<sup>424</sup>

However, once that approach was put into the metal, almost right from the start, the lack of recognition of an identifiable pattern was gradually evident, acknowledging that “DAMS has reached the size of ten valves, and it has proved exceedingly difficult to understand”.<sup>425</sup> Ashby patiently observed DAMS’ behavior for two years, waiting to see if it would organize itself in an intelligible way. No pattern was to be found, the machine behaving just plainly chaotic, without achieving any stability by itself --ever.

It is almost painful as it is fascinating to trace how Ashby patiently tries to tinker with parts of the machine here and there, pulling it apart and constructing it again, so that some kind of pattern of stability would obtain –and the subsequent ruminations that he has about the workings of the biological world after each tinkering.<sup>426</sup> It would show that his personal views on

---

<sup>423</sup> Ashby wrote in his diary that:

...to find the form was difficult. I thought about a number of possibilities but found it difficult to satisfy myself that they were right; for the possibilities got so devilish complicated that I found them difficult to trace them right through from design to performance (Ashby 1928-1972, p. 2950).

<sup>424</sup> Ibid.

<sup>425</sup> Ashby 1928-1972, p. 3148

<sup>426</sup> Pickering suggests that reconstructing the painstaking story of this machine would be a serious task worth pursuing (Pickering 2010, p. 110).

fundamental issues, such as Darwinian evolution, were being affected by it. For instance, within the framework of “throw design to the winds”<sup>427</sup> in frustration with the crippling complexity of designing a working DAMS, he finds solace in articles that would back up such a sort of approach:

A good example of a system of knowledge, i.e. of action, that has been worked out almost entirely by Darwinian methods of trial, error, small improvements, and always following whatever seemed better is the technique of histological staining, especially as shown in the variety of methods worked out by Cajal and the Spanish school. Many of their methods were developed to a high degree of effectiveness, and without the slightest rational foundation.<sup>428</sup>

It became increasingly obvious for Ashby that certain “fixes”, “ancillary regulations” had to be applied for the ultra-interconnected machine to show some kind of stable state. However, it would seem that one would be reverting back to the issue of “design”. There is no way to account for the needed “gate” that would control the passage of information (the flow of electric current), whenever it is required, in a sufficiently complex process of adaptation. But this gate allegedly has to exist, since complex (biological) organisms do adapt. The question regarding the entity that controls this gate seems to remain.<sup>429</sup> After all, the earlier homeostat’s random search was also an issue for some of the cyberneticians, for whom the idea of order as naturally emerging out of chaos due to inherent properties of matter was an uneasy concept – and unfaithful to Wiener’s understanding of entropy.<sup>430</sup> Ashby would later develop an answer that could be seen in stark contrast with the quote above...

The provision of the ancillary regulations thus demands that a process of selection, of appropriate

---

<sup>427</sup> Ashby 1928-1972, p. 2950 (quoted above).

<sup>428</sup> Ashby 1928-1972, pp. 3526-3527

<sup>429</sup> The issue of “design” in nature seems to be far from settled, and it shall not come as a surprise that Ashby’s name appears in contemporary treatments of the “science and religion” debate. See Huchingson 2005, p. 304; Rolston 2006, pp. 114-115.

<sup>430</sup> See chapter I, section 4.

intensity, exist. Where shall we find this process? The biologist, of course, can answer the question at once; for the work of the last century, and especially of the last thirty years, has demonstrated beyond dispute that natural, Darwinian, selection is responsible for all the selections shown so abundantly in the biological world. Ultimately, therefore, these ancillary regulations are to be attributed to natural selection... There is no other source.<sup>431</sup>

This biological “Maxwell’s Demon” in charge of a necessary-for-survival gateway operation is, according to Ashby, “Darwinian evolution”. As it is now known, Darwinism, beyond “natural selection”, was eventually improved via Mendelian genetics. In fact this improvement accounts for much of Darwinism’s explanatory *mechanisms* that would account for the mutation of a surviving organism in the first place. In face of this, Ashby’s radical *tabula rasa* epistemology (anti-nativist to the point of ascribing life characteristics to matter itself, and thus pushing the burden of causality to environmental influence alone) seems to collapse here. Evolution would have provided then these life-saving *homunculi* via genetic inheritance. There *is* something already furnished with nature that allows for living things to survive. If Ashby was not caught in a metaphysical web in conceding that the gateways would be the turf of Mother Nature, he did at least accept that whatever takes place *ad intra* the organism does complement behavioral cues. The “structure” of the living system, in particular its nervous system --and most interestingly, its brain—ultimately matters.

Models based upon measurable behavior are not enough past a certain point of complexity. After that point is reached, it would seem that chaos ensues. Ashby might have robbed cybernetics of the impetus that the enterprise had as one of its most important strengths for growth and development. But if he did, he did it in accordance to cybernetic dogma and by concluding a downward spiral that was already there in potential. One of these seeds pointed to a

---

<sup>431</sup> Ashby 1960, pp. 229-230

more troubling aspect of cybernetics –from the point of view of their scientific stand regarding studied phenomena and the role of models in science. The early alert regarding this phenomenon was given, in fact, by von Neumann. The next chapter will show how those concerns were indeed pointing to features of cybernetics' own undeclared metaphysics that would eventually prove to be harmful for the consistency of the project. A treatment of that aspect of the cybernetic development towards its own collapse is the content of the next chapter.

## **VI. Emphasizing the limits of a machine epistemology.**

...one could draw within plausible time limitations a fictitious ‘nervous network’ that can carry out all these functions... It is possible that it will prove to be too large to fit into the physical universe.<sup>432</sup>

### **1. John von Neumann’s appropriation of McCulloch and Pitts’ networks.**

One of the fiercest proponents of cybernetics even before the time it became widely known, John von Neumann was a prominent Hungarian Jewish mathematician educated in an elite Protestant school. He fled the growing Nazi terror and landed in the US, progressively becoming, probably partly due to émigré effusiveness, profoundly involved in secret “nation-identity-building” military projects, such as the development of the atomic bomb.<sup>433</sup> The role of von Neumann in the cybernetic enterprise has been, up until now, regarded as somewhat secondary. Yet, he actively engaged in most of the Macy conferences, having kept close contact with its main protagonists: Norbert Wiener, Warren McCulloch and Walter Pitts. In fact, the constructed computer machine whose internal design became known as the “Von Neumann architecture” was directly inspired by cybernetic discussion.<sup>434</sup> Probably more importantly, some of his insights surmounted deep philosophical issues within cybernetics that might have not been fully grasped

---

<sup>432</sup> Von Neumann 1951, pp. 33-34

<sup>433</sup> Having chosen Hiroshima and Nagasaki as the most convenient cities to drop the bomb, and tirelessly advocating a preventive nuclear strike against the Soviet Union, von Neumann had the Chiefs of all the American Forces (Air, Naval, and Army) next to his bed while sick and close to his death, asking for his strategic advice. (Rédei 2005, p. 7).

<sup>434</sup> Dupuy 2000, pp. 65-69.

at the time by its own members, but that nevertheless might have contributed to cybernetics' own implosion.

Originally interested in quantum mechanics and in the incipient conceptualization of information in physics, von Neumann maintained a close friendship with Norbert Wiener from the mid-1930s on. As previously mentioned, both Wiener's insights into information and entropy, and the philosophical and scientific consequences coming out of them, were very close to his inner interests. Indeed he probably felt having been beaten by Wiener into showing first the import and relevance of these connections.<sup>435</sup> The beginning of the 1940's, however, marked a distinct moment in his intellectual journey. Having read Turing's paper "On Computable Numbers",<sup>436</sup> and while being deeply impressed by it, he still could not get around the idea of applying such an outlook to the human mind. For him<sup>437</sup> the apparent ontological rivalry between the rich, creative and messy human mind, versus the dry, exact and rigorous logic that came out of it, entailed an enduring interrogation mark. Indeed, he referred to Turing's brilliant paper, in his own words, as being very "anti-neurological"<sup>438</sup>--until he came across McCulloch and Pitts' article.<sup>439</sup> This paper, itself heavily influenced by Turing's insights, struck a chord in von Neumann for two important reasons. One of them, of a more practical nature: he needed a machine to perform massive number crunches in order to advance the secret development of a nuclear weapon. McCulloch and Pitts' paper, which in his own words was needlessly convoluted due to the use of Carnap's logical nomenclature, was nevertheless the basis for his EDVAC

---

<sup>435</sup> Chapter II, section 4.

<sup>436</sup> Turing 1937

<sup>437</sup> As for Warren McCulloch (See chapter IV, section 3).

<sup>438</sup> Von Neumann 2005, p. 278

<sup>439</sup> McCulloch & Pitts 1943

project –what became the world’s first stored-program computer, whose architecture still in use in most computers was referred above.<sup>440</sup>

Von Neumann’ secret and active engagement in atomic weaponization seemingly made of him a cyberneticist with his own non-disclosed agenda. Although several cyberneticists were involved in military projects (cybernetics itself having come out of an effort for facilitating complex military tasks), von Neumann went the extra mile in using cybernetic resources for his own military endeavors. As previously mentioned, he made Wiener buy into believing his willingness for coming to MIT to work with him, provided there would be adequate funding for constructing a remarkably expensive computing machine. Once the offer by MIT became concrete (thanks to Wiener’s intense lobbying), Princeton University, von Neumann’s home institution, took serious notice, and in turn offered funding to build the machine there. Thus von Neumann never went to MIT –as he probably never intended to do in the first place, merely using MIT’s offer as negotiation leverage with Princeton. Furthermore, he even convinced Wiener to send his own close collaborator, Julian Bigelow, to work with him --to which Wiener

---

<sup>440</sup> Chapter IV, section 3. Although beyond the scope of this thesis, the popular ascription of “father of the computer” to von Neumann is lately disputed mainly by British scholars who resent that Turing is not widely acknowledged as responsible for today’s computers. The traditional claim is that the American side, with all its resources, managed to “once again” kidnap an original British idea, exploit it and reap its benefits. A report by Stanley Frankel, a colleague of von Neumann at Los Alamos, is brought up, seemingly remarking von Neumann’s reliance on Turing’s 1937 “On Computable Numbers”. Indeed, US mathematicians and engineers allegedly were “brute forcing” by means of complicated and expensive hardware the instantiation of von Neumann’s fairly vague ideas laced with biological terminology (more on this below). Also, the engineer in charge of translating von Neumann’s manuscript into a physical machine, Harry Huskey, refers to it as being “of little help to those of us working on the project.” It is alleged that Turing’s approach, in contrast, mindful of the scarcity of material resources, would come up with clever solutions requiring the least possible hardware. On the American side, a report by Julian Bigelow is brought up, where he advances that the merit truly relies on von Neumann’s acumen to flesh out Turing’s philosophical insights into the metal. In fact, Turing’s mentor, Max Newman, spoke of von Neumann as being an “applied mathematician at heart” –despite the latter’s brilliance in articulating theoretical issues (e.g., in quantum physics, at the beginning of his career). Since both British and American governments keep de-classifying WWII documents, the issue is far from closed –in fact, it is gaining momentum among technology historians, particularly from the British side. See Hodges 1983, ch. 5; Copeland 2004, pp. 21-27; Huskey 2005, p. 283; Copeland 2012, ch. 8; Copeland & Proudfoot 2005, pp. 112-117.



candidly agreed. And so Bigelow became chief engineer for the building of EDVAC.<sup>441</sup> All these moves were performed while portraying the building of the computer as a genuinely cybernetic aim –which in part, it was, but only in part, as we will see. For the time being, the secret task of collaborating towards a nuclear weapon was conveniently left unspoken.

In the famous *First Draft*<sup>442</sup> for building the EDVAC, it was mentioned that more than one engineer found it odd that von Neumann used biological terms for referring to the parts of the future machine --such as neurons, synapses and organs.<sup>443</sup> This oddity points towards the less practical and more philosophical second reason behind von Neumann's deep appreciation of the McCulloch and Pitts' paper. He was impressed by the apparently successful effort of linking the two allegedly naturally incompatible entities --mind and logic, the latter having strangely come out of the former. The 1943 paper gave a mechanistic account of how a nervous system in general, and the mind in particular, might work.<sup>444</sup> This was no small feat. Von Neumann enthusiastically jumped into the cybernetic wagon and only three years later he undertook the honor of giving the first ever formal talk on cybernetics at the Macy conferences in March 1946.<sup>445</sup> At this opening event he recounted and explained how his computer machine became a reality, clearly acknowledging his theoretical debt to McCulloch and Pitts' work (This was followed, within the same presentation, by Lorente de Nó's suggested application of the model to brain physiology).

---

<sup>441</sup> These strategically motivated moves, which showed von Neumann's acute political wit, might have contributed later to tensions with Wiener, as depicted in chapter IV, section 3.

<sup>442</sup> Von Neumann 1945

<sup>443</sup> There are those who believe that keeping the biological vocabulary in an engineering paper allowed von Neumann to circumvent the veil of secrecy that would be otherwise put by the military –and so he would be able to share it, discuss it and eventually improve it. Let us not forget that this was a tool being built in order to expedite the construction of atomic weapons. See Tedre 2006, pp. 211-215.

<sup>444</sup> See chapter IV, section 3.

<sup>445</sup> See chapter I, section 4.

However, after von Neumann's 1943 'epiphany' while working on the construction of the computing machine (with an undisclosed eye firmly set on developing the bomb), he remained genuinely curious regarding the actual correlation between the McCulloch & Pitts' model on the one hand and real biological substrata on the other. Von Neumann, as much as McCulloch and Pitts,<sup>446</sup> defended the heuristic value of axiomatization and the resulting use of idealized, black-boxed neurons, only reactive in a yes/no fashion. Armed with this methodology, von Neumann engaged in extensive connections with the biological and medical community, exchanging ideas and discussing views.<sup>447</sup> It did not take long for him to begin showing warning signs regarding the feasibility of the newly devised artificial networks in what pertains to brain matters –at the same time that the computer based upon the same model was being duplicated by dozens (since he did not copyright the now legendary *First Draft*).<sup>448</sup>

Tracing von Neumann's gradual coming to terms with the extreme complexity of the structure they were bracketing for the sake of theoretical operability --the nervous system and the brain-- is historically fascinating. Below is a critical account of what progressively grew into a rift within the cybernetic movement –a rift which opened up hidden assumptions whose implications are arguably letting themselves be felt up to our times.

---

<sup>446</sup> In all fairness to Walter Pitts, having undertaken himself a journey that started with logic and became progressively close to biology, he was always aware and cautious of such idealizations. Dupuy remarks that in a "curious exchange of roles, while the logician Pitts abandoned formal models to pursue experimental research on the nervous system, the neurophysiologist McCulloch was to devote all his efforts to theory and logic" (Dupuy 2000, p. 137).

<sup>447</sup> Ironically, there are more records left of these biomedical meet ups, than those concerned with cybernetics. See Aspray 1990, pp. 298-301.

<sup>448</sup> Situation that occasioned a bitter rift among those who were involved in the EDVAC project. Herman Goldstine was one of them, and he recounts the story in his book (Goldstine 1993, pp. 211-224).

## 2. The letter to Wiener: Post hoc, ergo propter hoc?

As already mentioned, two years before the *First Draft* for building EDVAC,<sup>449</sup> and enthused by the possibilities opened up by McCulloch and Pitts' networks, von Neumann actively engaged in dialogue with the biomedical community. By November 1946, towards the end of the very same year when he inaugurated the 1<sup>st</sup> Macy Conference with the first talk given in March, von Neumann wrote a revealing letter to Norbert Wiener. Its contents are of considerable importance if one takes it as a probe into what was going on in his mind regarding the feasibility of the cybernetic enterprise for modeling the brain. The letter starts with a dramatic tone, putting forward the issues that disturbed him. He then elaborates on them, in order to end up suggesting a change of focus for their future cybernetic endeavors—a change that he eventually undertook as a prime task of his intellectual journey, indeed until the end of his days, one decade later.

Von Neumann started off by acknowledging that some issues were deeply troubling him for the most part of the year. One could safely assume that he already was entertaining these unsettling thoughts *while* giving his presentation at the 1<sup>st</sup> Macy Conference, nine months earlier. He mentioned his reticence for letting him (Wiener) know of these troubles up until then, due to the lack of a positive alternative –which he introduced towards the end of the letter. The general problem is stated as cybernetics having chosen, in his view, probably the absolutely most difficult object of study: the brain. Cybernetics relied on theories which can certainly give some kind of explanation, in a substantially mechanical manner. That much it is granted. But the situation in which the whole enterprise lies now, he continues, is not better than the moment in which it started. In fact, he believed that “after the great positive contribution of Turing-cum-

---

<sup>449</sup> Von Neumann 1945

Pitts-and-McCulloch is assimilated, the situation is rather worse than better than before”.<sup>450</sup> The cause for this rests *both* in the extreme complexity of the subject at hand, and the theoretical models that have been used.

More specifically, von Neumann somewhat resented that McCulloch and Pitts’ network provided results in such a self-contained, exact but general manner, that it left no way to confront it with the object that they were supposed to understand in the first place. If the model ends up enjoying a space beyond the realm of testability (and thus beyond scientific verifiability) it might still be heuristically good for something, but likely not for a nervous system. This other “something” was of course already having a physical instantiation at the hands of von Neumann. Indeed, the weakness in the model’s refutability aspect did not remove from it its proved productive capacity, as the construction of his computing machine was amply evincing. However, once we step back and begin to really pay attention to the structure of the object that we wanted to originally parcel and explain, the complexity is such that one is shocked into a sort of epistemic paralysis:

After these devastatingly general and positive results one is therefore thrown back on microwork and cytology –where one might have remained in the first place... Yet, when we are in that field, the complexity of the subject is overawing.<sup>451</sup>

Von Neumann gave picturesque analogies of the situation in which he felt they were while trying to study the brain with the best of the cybernetic constructs. For him to “understand the brain with neurological methods seems to me about as hopeful as to want to understand the ENIAC... with no methods of intervention more delicate than playing with a fire hose... or dropping

---

<sup>450</sup> Von Neumann 2005, p. 278

<sup>451</sup> Von Neumann 2005, p. 278

cobblestones into the circuit”.<sup>452</sup> He then proceeded to expand on the nature of the problem they were facing, which revealed a second discouraging aspect. If one is to scale down the choice of degree of complexity of the nervous system to study, say from a human one to that of an ant, much of what can be grasped by the McCulloch and Pitts’ model goes down the drain, since such simplicity in turn comes packaged with several blind spots: “As the digital (neural) part simplifies, the analogy (humoral) part gets less accessible, the typical malfunctions less known, the subject less articulate, and our possibilities of communicating with it poorer and poorer in content”.<sup>453</sup> However, thereafter he tried to turn the negativity around and finally produced, all things considered, a radical, short and revealing suggestion, signaling the venue that cybernetics had to pursue in order to render more feasible results: “I feel that we have to turn to simpler systems”.<sup>454</sup>

Von Neumann tried to expose the flawed nature of the argument behind the cybernetic need for studying only multi-cellular entities. The idea was that since a neuron is just one cellular entity, only a large number of them networked together could render something meaningful. He pointed out to the complexity of that one cell, which shows reproductive features, and in some cases, it appears to be a perfectly self-contained living organism. Having in consideration these characteristics, which point out to a rich and likely relevant complexity in the unicellular organism, von Neumann cast a dark shadow upon McCulloch and Pitts’ networks: “This in itself should make one suspicious in selecting the cells as the basic ‘undefined’ concepts of an axiomatism”.<sup>455</sup> Without specifically mentioning either scientists or their model, the reference to

---

<sup>452</sup> Von Neumann 2005, p. 279

<sup>453</sup> Von Neumann 2005, p. 279

<sup>454</sup> Von Neumann 2005, p. 279

<sup>455</sup> Von Neumann 2005, p. 279

McCulloch and Pitts' networks is clear: The "black-boxology"<sup>456</sup> that brackets the somatic aspect of the neuron for heuristic purposes might be paying a high price. So high indeed, that both the studied baby and the epistemic bathwater might be, in fact, being disposed of.

Furthermore, the whole issue was probably being approached backwards --or to be more precise, upside down. Von Neumann questioned the wisdom of targeting the object of study (particularly if it is regarded as a highly complex machine) from the perspective of understanding its finished "whole" --itself prohibitively sophisticated and already displaying a behavior whose complexity explodes off the charts. Common sense would indicate that the way to approximate such a complex object would be via its simpler, relatively isolated parts, and then reconstruct the entire machinery from the ground up. In his words, the venue currently pursued by cybernetics did not amount to a reasonable methodology of study --once again, due to the uniquely complex nature of its object of study...

Consider, in any field of technology, the state of affairs which is characterized by the development of highly complex "Standard components", which are at the same time individualized, well suited to mass production, and (in spite of their "standard" character) well

---

<sup>456</sup> This black box approach was present in cybernetic gatherings right from the start, and it played a role beyond a typical behavioristic metaphor. In fact, understanding an unknown "enemy machine" would use this heuristic --all the more relevant due to the ongoing war. The first cybernetic gathering at Princeton, as recalled by Warren McCulloch, shows that von Neumann was quite comfortable with a black box approach, when the situation merited it...

Lorente de Nó and I, as physiologists, were asked to consider the second of two hypothetical black boxes that the allies had liberated from the Germans. No one knew what they were supposed to do or how they were to do it. The rust box had been opened and exploded. Both had inputs and outputs, so labelled. The question was phrased unforgettably: "This is the enemy's machine. You always have to find out what it does and how it does it. What shall we do?" By the time the question had become that well defined, Norbert was snoring at the top of his lungs and his cigar ashes were falling on his stomach. But when Lorente and I had tried to answer, Norbert rose abruptly and said: "You could of course give it all possible sinusoidal frequencies one after the other and record the output, but it would be better to feed it noise --say white noise-- you might call this a Rorschach." Before I could challenge his notion of a Rorschach, many engineers' voices broke in. Then, for the first time, I caught the sparkle in Johnny von Neumann's eye. I had never seen him before and I did not know who he was. He read my face like an open book. He knew that a stimulus for man or machine must be shaped to match nearly some of his feature filters, and that white noise would not do. There followed a wonderful duel: Norbert with an enormous club chasing Johnny, and Johnny with a rapier waltzing around Norbert --at the end of which they went to lunch arm in arm. (McCulloch 1974, p.11).

suited to purposive differentiation. This is clearly a late, highly developed style, and not the ideal one for a first approach of an outsider to the subject, for an effort towards understanding. For the purpose of understanding the subject, it is much better to study an earlier phase of its evolution preceding the development of this high standardization-with-differentiation. I.e. to study a phase in which these “Elegant” components do not yet appear. This is especially true, if there is reason to suspect already in the archaic stage mechanisms (or organisms) which exhibit the most specific traits of the simplest representatives of the above mentioned “late” stage.<sup>457</sup>

Furthermore, von Neumann suggested that they should still go one big notch down in what concerns the optimal selection of an object of study. In going after elementary parts to treat, once they are better understood, cybernetics could have a better grasp of how the whole thing works. This methodology, von Neumann hoped, would in turn become a firmer building block for the improved cybernetic endeavor.

Von Neumann specified that they could draw the line in terms of how far “down” they should go once they reached the realm inhabited by *virus*. He reminded Wiener that such entities do show features that one could reasonably ascribe to full-blown living organisms, given that they “are self-reproductive and they are able to orient themselves in an unorganized milieu, to move towards food, to appropriate it and use it”.<sup>458</sup> Von Neumann confidently went on to express that “a ‘true’ understanding of these organisms may be the first relevant step forward and possibly the greatest step that may be at all required”.<sup>459</sup> The word “true” was purposely put by him inside quotations marks, as he went on to specify what he understood by it in this context. His clarification of what is “true” no longer rings as being in communion with the “behavior-with-a-purpose” cybernetic article of faith --namely, that something is a machine because it

---

<sup>457</sup> Von Neumann 2005, pp. 279-280

<sup>458</sup> Von Neumann 2005, p. 280

<sup>459</sup> Von Neumann 2005, p. 280

*behaves* machine-like. This last move in effect entailed an important epistemological displacement in cybernetic dogma. Now the cybernetic take on “true” (understanding) would approximate that which was strategically bracketed off precisely for the sake of theory extension in the first place:

I would, however, put on “true” understanding the most stringent interpretation possible: That is, understanding the organism in the exacting sense in which one may want to understand a detailed drawing of a machine, i.e. finding out where every individual nut and bolt is located, etc. It seems to me that this is not at all hopeless... I suppose (without having done it) that if one counted rigorously the number of “elements” in a locomotive, one might also wind up in the high ten thousands. Consequently this is a degree of complexity which is not necessarily beyond human endurance.<sup>460</sup>

The suggestion of turning their focus into the structure of the object, into the stuff that makes up the actual body of the organism probably had a bitter taste for Wiener. Avoiding opening and tinkering with the innards of the black box was to a considerable degree what helped launching the cybernetic totalizing worldview in the first place. Recall Wiener’s struggle with a reductionist notion of “behavior” and subsequent musings at expanding it.<sup>461</sup> What von Neumann is suggesting goes frontally against what Wiener had so effusively proposed in his 1943 founding paper. Towards the end, that article leaves no doubt regarding the more-than-heuristic role that a behavioristic approach has in the study of animal and machine as one sole entity –different in degree of complexity, not in kind.<sup>462</sup> Wiener maintained that

a uniform behavioristic analysis is applicable to both machines and living organisms, regardless of the complexity of the behavior. It has sometimes been stated that the designers of machines

---

<sup>460</sup> Von Neumann 2005, p. 280

<sup>461</sup> While constructing his AA-Predictor (Chapter I, section 2).

<sup>462</sup> For the historical and philosophical roots of this approach, see the note on Duns Scotus and Giambattista Vico in chapter IV, section 4.



merely attempt to duplicate the performances of living organisms. This statement is uncritical. That the gross behavior of some machines should be similar to the reactions of organisms is not surprising. Animal behavior includes many varieties of all the possible modes of behavior and the machines devised so far have far from exhausted all those possible modes. There is, therefore a considerable overlap of the two realms of behavior... A further comparison of living organisms and machines leads to the following inferences. The methods of study for the two groups are at present similar. Whether they should always be the same may depend on whether or not there are one or more qualitatively distinct, unique characteristics present in one group and absent in the other. Such qualitative differences have not appeared so far.<sup>463</sup>

Furthermore, Ross Ashby, the theorizer of cybernetic fundamentals --always in intellectual sync with Wiener-- defined cybernetics as the science that “treats... not things but ways of behaving. It does not ask ‘what is this thing?’ but ‘what does it do?’ ...It is thus essentially functional and behaviouristic”.<sup>464</sup> And we saw how steadfastly holding to this tenet cost him a high price.<sup>465</sup> All in all, we can see that an enriched behavioristic approach laced cybernetics from the beginning to the end, from the American and the British side.

Let us recall that von Neumann’s critical observations are coming the very same year he was giving the inaugural Macy conference presentation. Expectably, the letter was not sent without some prudent hesitation. He consulted with Walter Pitts before sending it off.<sup>466</sup> What exactly happened after it was delivered remains largely unknown. We do know that a meeting between Wiener and von Neumann likely took place after the letter, in December --but we do not know what transpired there. We also know that Norbert Wiener did not mention anything about this letter to anyone --not in writing nor orally-- ever. However, as has been mentioned, right

---

<sup>463</sup> Rosenblueth, Wiener & Bigelow 1943, p.22

<sup>464</sup> Ashby 1961, p. 2

<sup>465</sup> I.e., the case of DAMS (Chapter V, section 4).

<sup>466</sup> Macrae 1992, p. 107

after those two months people began to notice a discernible distance between these two hitherto close old friends. Was one fact caused by the other, given that one occurred after the other? Beyond merely correlating events, one can only speculate a causal connection between both. And even if the correlation seems to be fairly strong, we now know that there was more going on.<sup>467</sup> Still, there is a curious epilogue regarding this letter, which might partially explain von Neumann's posterior ardent impetus for further developing his critique against the (at the time) current methodological philosophy of cybernetics.

As mentioned earlier, from 1943 on, due to McCulloch and Pitts' bold hypothesis regarding the brain, von Neumann developed an active interest in biology and physiology, taking advantage of fruitful contacts in the biomedical community. In fact McCulloch regarded von Neumann's labors as the reason why the biological community took interest in his paper. Less than a year after the controversial letter, at the 3<sup>rd</sup> Macy Conference,<sup>468</sup> von Neumann successfully lobbied to bring the German biophysicist Max Delbrück (1906–1981) as invited guest. Von Neumann had in fact mentioned Delbrück's work in embryology and bacteriophage virus in the letter to Wiener. All the intentions were set in place for christening Delbrück as a permanent member of the cybernetic group. For von Neumann, Delbrück would be the perfect glove for the cybernetic hand: a physicist turned biologist, trying to scientifically (and thus, for them, mechanically) understand life, paying due attention to the most elementary building blocks of a living system. Once the invitation was officially arranged, the events that followed it took a somewhat bizarre turn.

---

<sup>467</sup> See chapter II, section 4.

<sup>468</sup> March 1947

Max Delbrück, now formally invited, could not attend the 4<sup>th</sup> Macy Conference, but he did make it to 5<sup>th</sup> one.<sup>469</sup> It can ostensibly be assumed that von Neumann was quite excited about this. After all, the scientist who symbolized the field of research he thought to be the next platform for cybernetics to take off was showing up at the cybernetic hub. As it turns out, right after the conference ended, Delbrück made it clear that he was never going to come back again.

Delbrück was indeed so disappointed, that he was quoted some years later characterizing the Macy conference as “vacuous in the extreme and positively inane”.<sup>470</sup> In all fairness, it might have been a case of spectacular bad timing. The first day that Delbrück attended practically the whole day was devoted to issues of language. He might have felt of it, literally, as a waste of his time. Delbrück was a genuine physicist turned biologist, accustomed to the rigors of hardcore physics --despite the cybernetic mandate, there never was even one physicist in the group. He might have been turned off by the purposely “flexible” language, which always tried to accommodate the vastly, even disparately different disciplines being loosely matched at the gatherings --from mathematics and engineering to psychology and anthropology.<sup>471</sup> Or it could be that both reasons perfectly conflated into the disastrous (for von Neumann) outcome. Be that as it may, it is impossible to know with exactitude what kind of effect this reaction had on him. But one can safely assume that he did not take it lightly. The subsequent increasing sharpness of von Neumann’s criticism against cybernetics might further fuel that assumption.<sup>472</sup>

---

<sup>469</sup> Spring 1948

<sup>470</sup> As personally conveyed to Steve Heims years later (Heims 1991, p. 95).

<sup>471</sup> Dupuy 2000, pp. 76-80

<sup>472</sup> Max Delbrück went on to make substantial inroads in what later became known as “molecular biology”. However, this field was not entirely developed outside the cybernetic aura. A decade earlier, Warren Weaver was the head of the Natural Sciences section of the Rockefeller Foundation. Weaver helped bringing Delbrück to America in 1938, securing founding for him at Caltech. In fact, Weaver coined the very term “molecular biology”. Furthermore, around the time that Delbrück would be refusing to join the “cybernetic group”, Warren Weaver was writing to Norbert Wiener about a feasible extension of Claude Shannon’s 1948 famous paper, namely, the possibility of Machine Translation --which drew substantial military interest from the government, an already familiar parallel effect of cybernetics (See Hutchins 2000).

### 3. The Hixon Symposium: Strengths and weaknesses of the McCulloch and Pitts' networks.

Only three months later, von Neumann was again standing behind a podium giving a talk on cybernetics, this time at the famed Hixon Symposium.<sup>473</sup> By this time, two years after the infamous letter, he had his plan for the future development of cybernetics more in focus. Von Neumann took particular interest in the logic behind features that portray life in “very simple” organisms. A feature he was particularly interested in was reproduction. It was his understanding that thanks to Turing’s “On Computable Numbers” it was no longer a mystery how a more complex entity can be begotten from a less complex one --something which apparently goes against common sense, but nevertheless is allegedly a key feature of evolution. As von Neumann read Turing, a universal Turing machine embedded with sufficient instructions can produce (the behavior of) any other machine, provided that rigorous and unambiguous instructions are furnished. Von Neumann remarked that this is no more mysterious than the capacity we have for reading pretty much anything after the rules of grammar are firmly entrenched in our cognitive realm.

However, the logic required for the reproduction of artificial entities would have to be attentive to details pertaining to their structure and its elements. Such future logic should take into consideration issues of heat, material fatigue, friction, etc. It would be a logic acquainted with issues of thermodynamics. That much von Neumann had clear by that time. And the way he introduced “The general and logical theory of automata”<sup>474</sup> was by means of pointing out the deficiencies of current cybernetic models. There had to be, after all, a core reason why

---

<sup>473</sup> September 1948. Published as Jeffress 1951.

<sup>474</sup> Von Neumann 1951

cyberneticists should turn their attention to his new proposal.<sup>475</sup> Hence, the McCulloch and Pitts' model received a gloves-off attack whose entire array of consequences might have not been thoroughly evaluated until these days.<sup>476</sup>

Warren McCulloch might have already been aware of von Neumann's Janus-faced attitude regarding his (and Pitts') model. It can be safely assumed that by this time he understood well that von Neumann was cheerful about his model only in so far as it opened up possibilities for constructing computing machines based upon it --machines that would perform certain tasks that until that moment were done by humans<sup>477</sup> in an incomparably faster manner, and practically flawlessly. McCulloch likely was indeed aware that von Neumann was less than happy, in fact probably disturbed, by the association of such model to an actual nervous system.

McCulloch gave a talk at the very same conference entitled "Why the Mind is in the Head".<sup>478</sup> There he made a reference to von Neumann's computers, somewhat jokingly pointing to their inherent limitations. In praising the practical efficiency of the brain's neural networks, McCulloch said that if von Neumann's machines would attempt to come any closer to the overall behavior of the brain, then the kind of resources demanded would exceed current practical feasibility...

Neurons are cheap and plentiful. If it cost a million dollars to beget a man, one neuron would not cost a mill. They operate with comparatively little energy. The heat generated raises the blood in

---

<sup>475</sup> Later on von Neumann regarded his subsequent work in theory of automata as going beyond being a solution for an ill-oriented cybernetics. He saw his developing theory as a future cornerstone for the revision of not only thermodynamics, but of logic and mathematics in general. Thus the theory of automata was meant to be his most important legacy --correspondingly speaking of it with solemnity until the moment of his death, in 1957. Such might have been the reason why he chose to pursue it alone, unlike any other project he was involved with up to that point. See Shannon 1958, p. 123; Von Neumann & Burks 1966, p. 28; Goldstine 1993, p. 285.

<sup>476</sup> Some people besides McCulloch did not like the criticism. Marvin Minsky, a doctoral student of von Neumann and later pioneer of AI, did not regard it well. Although the newly formed wave of AI researchers saw symbolic logic as superior to artificial neural networks, they were unfriendly to the idea that neural networks (and probably by extension, their own symbolic models) were that far removed from what actually occurs in the brain (See Dupuy 2000, pp. 65-69).

<sup>477</sup> Called computers, or number crunchers --a remarkably tedious job. See chapter III, section 3.

<sup>478</sup> McCulloch 1951

passage about half a degree, and the flow is half a liter per minute, only a quarter of a kilogram calorie per minute for  $10^{10}$ , that is, 10 billion neurons. Von Neumann would be happy to have their like for the same cost in his robots. His vacuum tubes can work a thousand times as fast as neurons, so he could match a human brain with 10 million tubes; but it would take Niagara Falls to supply the current and the Niagara River to carry away the heat.<sup>479</sup>

These extraordinarily problematic outcomes would be happily acknowledged by von Neumann, with the following twist. These problems are in fact expectable, since McCulloch and Pitts' artificial networks are good for constructing computing machines, not for telling us how our nervous system, in fact *any* nervous system, works. Further, von Neumann went the extra-mile in the discussion after his own talk at the conference, pointing to even more mind-blowing outcomes if one is to attempt modeling the nervous system with such networks. The problem is not his machines: the problem is the model.

Von Neumann started by pointing out the methodology that validated the formal networks in the first place. Faced with the extreme complexity characteristic of natural organisms, scientists can be justified (given their mechanistic leniencies) in understanding the object of study as constituted by smaller, fairly isolated parts. This view in turn can subdivide the problem in two sub-tasks: 1) The attempt to understand “the structure and functioning of such elementary units, individually;”<sup>480</sup> and 2) “Understanding how these elements are organized into a whole, and how the functioning of the whole is expressed in terms of these elements”.<sup>481</sup>

The first task was dismissed by von Neumann as extremely (sometimes impossibly) demanding, being the core of nitty-gritty scientific experimentation, in which humans will

---

<sup>479</sup> McCulloch 1951, p. 54

<sup>480</sup> Von Neumann 1951, p. 2

<sup>481</sup> Von Neumann 1951, p. 2

engage for likely a long time to come.<sup>482</sup> He is concerned with the second task. A person with mathematical or logical inclinations can find fertile land here. How? By means of “axiomatizing”, putting into “black boxes” all the chemico-biological elements underlying the basic units of the organic system --in this case, the neuron. The job concerning what takes place *ad intra* these heuristically convenient brackets is left to the scientist underdog --that is, task 1).

Once this convention is put in place, the work gets greatly simplified. We assume of these “neurons” as being “automatisms, the inner structure of which need not be disclosed, but which are assumed to react to certain unambiguously defined stimuli, by certain unambiguously defined responses”.<sup>483</sup> These basic units admit only of a yes/no type of reaction, allegedly being the biological instantiation of a logical gateway, and thus amenable of being mapped upon good old logical connectives (not, and, or, if...then). The result is that “any functioning in this sense which can be defined at all logically, strictly and unambiguously in a finite number of words can also be realized by such formal neural networks”.<sup>484</sup> Von Neumann acknowledged that the implications of this methodological ‘move’ are as tremendous as they are admirable. In a sentence that resembles one of Ross Ashby’s ultra-cybernetic lines of reasoning, he evaluated and appreciated it...

McCulloch and Pitts' important result is that any functioning in this sense which can be defined at all logically, strictly, and unambiguously in a finite number of words can also be realized by such a formal neural network. It is well to pause at this point and to consider what the implications are. It has often been claimed that the activities and functions of the human nervous system are so complicated that no ordinary mechanism could possibly perform them. It has also been attempted to name specific functions which by their nature exhibit this limitation. It has

---

<sup>482</sup> Note that this is precisely what von Neumann suggested Wiener to do in the letter, two years earlier. See section 2, above.

<sup>483</sup> Von Neumann 1951, p. 2

<sup>484</sup> Von Neumann 1951, p. 22

been attempted to show that such specific functions, logically, completely described, are per se unable of mechanical, neural realization. The McCulloch-Pitts result puts an end to this. It proves that anything that can be exhaustively and unambiguously described, anything that can be completely and unambiguously put into words, is ipso facto realizable by a suitable finite neural network.<sup>485</sup>

There would finally be a glimpse of an answer to the age old question regarding how both exact disciplines and rigorous thinking can emerge from heretofore the anti-thesis of mechanical exactitude: the mind. This was not only the question that kept Warren McCulloch out of the Protestant Theology Seminary. It also made von Neumann plunge into the cybernetic project, later coming out with a full-blown computer. In fact, this insight was probably the one that, along with Wiener's views on teleology, drove scientists of the time into the cybernetic movement in the first place.

However, that was as far as von Neumann would go in terms of praising the cash value and heuristic justification of the neural model. Subsequently, he picked up where he left off regarding such a remarkable achievement –namely, making feasible of being run by these networks every behavior that can be rigorously and unambiguously described. Von Neumann began to explore the downside of this methodological mandate. Specifically, he remarked on two basic questions that emerge; 1) “whether that network can be realized within a practical size, specifically, whether it will fit into the physical limitations of the organism in question;”<sup>486</sup> and 2) “whether every existing mode of behavior can really be put completely and unambiguously into words”.<sup>487</sup>

---

<sup>485</sup> Von Neumann 1951, p. 23

<sup>486</sup> Von Neumann 1951, p. 23

<sup>487</sup> Von Neumann 1951, p. 23



Regarding the first one, once again von Neumann left it as the stuff for painstaking hard science tinkering –however, he would later make a clear reference to this issue within the discussion that followed suit right after his presentation.<sup>488</sup> Regarding the second question, he tackled the issue by means of an example. What would it take to understand how we can know a triangle? Prima facie, applying McCulloch and Pitts’ networks we can describe clearly the characteristics of a triangle (3 sides, 180 degrees). However, we soon realize that we are also able to recognize circular triangles, deformed triangles, upside down triangular figures, and so on. Suddenly, we realize that the more accurate our description is, the lengthier our report becomes. Nevertheless, even if it gets very long, it is still feasible... That is, until we clash into a second aspect of the issue, whose chances of tractability look as bleak as they look discouraging:

All of this [the set of descriptions], however, constitutes only a small fragment of the general concept of identification of analogous geometrical entities. This in turn is only a microscopic piece of the general concept of analogy.<sup>489</sup>

Von Neumann questioned the plausibility, in terms of practicality, of totally and finally describing the behavior of an organ as the good route to take –even if it is feasible in principle. Here he re-introduced what was advanced to Wiener in his letter two years before: the very un-cybernetic idea of opening the “black box” and describing the actual “connections of the visual brain”.<sup>490</sup> In what amounts to a radical change of modeling strategy, von Neumann suggested swapping the model for the modeled as the most realistic avenue to pursue, given the explosively unhandleable nature of describing all possible behaviors:

---

<sup>488</sup> Von Neumann 1951, pp. 32-41

<sup>489</sup> Von Neumann 1951, p. 24

<sup>490</sup> Von Neumann 1951, p. 24

It is not at all certain that in this domain a real object might not constitute the simplest description of itself, that is, any attempt to describe it by the usual literary or formal-logical method may lead to something less manageable and more involved.<sup>491</sup>

After a certain threshold of complexity in the studied organ, it becomes more feasible to directly describe the very structure of the object, rather than accounting for all the possible behaviors of the said organ. This renders McCulloch and Pitts' networks good for very simple organisms, but utterly inadequate for something whose degree of complexity will entail that a constructed model would require the dynamic and cooling forces of the Niagara Falls --as McCulloch sarcastically suggested. Von Neumann did not spare any opportunity to clearly mention the "McCulloch-Pitts results" as no longer useful when we attempt to describe a more complex organ, especially the brain --the most complex of them all. The discussion that ensued expectably had Warren McCulloch as the first respondent. What he said within his response was a hard confession that points to an acknowledgment of the critique:

I taught neuro-anatomy while I was in medical school, but until the last year or two I have not been in a position to ask any neuroanatomist for the precise detail of any structure. I had no physiological excuse for wanting that kind of information. Now we are beginning to need it.<sup>492</sup>

After the behavioral neurophysiologist Ralph Gerard (1900-1974) asked for further explanations regarding the issue of rigorously describing complex behavior in a model, von Neumann took the occasion to relentlessly continue his critique. He accentuated the very practical problem of having to be able, when dealing with a complex object, to come up with a description that at least "can be read in a lifetime".<sup>493</sup> Furthermore, in connection with the two first questions

---

<sup>491</sup> Von Neumann 1951, p. 24

<sup>492</sup> Von Neumann 1951, p. 24

<sup>493</sup> Von Neumann 1951, p. 33

“dismissed” above,<sup>494</sup> and metaphorically vastly surpassing McCulloch’s Niagara Falls’ sarcastic reference, von Neumann asserted that using McCulloch and Pitts’ networks to map the human nervous system would “turn out to be much larger than the one we actually possess”.<sup>495</sup> *Much larger* is here used as a charitable understatement, as he then clarifies that “it is possible that it will prove to be too large to fit into the physical universe”.<sup>496</sup> Naturally, von Neumann consequently asked “What then? Haven’t we lost the true problem in the process?”<sup>497</sup>

As we can see, both “first questions”, initially dismissed as the stuff pertaining to the typical tasks of the hard sciences, returned in full force, since the consequence of modeling an extremely complex structure upon the possible behavior of such structure incurs into a growth explosion in terms of model size. Once the threshold of becoming factually non-manipulable is reached, we seem to have betrayed the whole point of the endeavor, namely, attempting to understand the organism. This reversal might have been profoundly influential for cybernetics’ own fate, instantiated in implications that failed to be recognized in full. A direct outcome, however, is below.

---

<sup>494</sup> Regarding the “structure” of a studied object. See previous section (3).

<sup>495</sup> Von Neumann 1951, p. 34

<sup>496</sup> Von Neumann 1951, p. 34

<sup>497</sup> Von Neumann 1951, p. 34

#### 4. The Illinois Lectures and the kinematic model: Cybernetic materialization of a theory.

Von Neumann reassumed the attack during a series of five lectures he gave at the University of Illinois in March 1949.<sup>498</sup> It is relevant to point out, however, that just as in the Hixon Symposium, the criticism of McCulloch and Pitts' networks was not the point of his presentation. Rather, the criticism was introduced as the point of departure upon which his "Theory of Self-Reproducing Automata"<sup>499</sup> would soon be developed, advancing it as "the way to go" precisely because of the unfeasibility of using McCulloch and Pitts' networks for modeling very complex organisms. The second of the five lectures is the one where von Neumann addressed McCulloch and Pitts' networks squarely, in the introductory manner just indicated. In this lecture, he gave some more details pointing to the inadequacy of the model in what pertains to the modeling the brain. Von Neumann would later on be more specific about what is being lost if one is to choose this neural model as the instrument for studying the nervous system. The line of criticism was consistent with his previous remarks on the issue.

---

<sup>498</sup> Nine months earlier (and six months after the Hixon Symposium), at the Sixth Macy conference in March 1949, no one other than Warren McCulloch was in charge of delivering at the opening presentation a message from von Neumann—who could not attend. McCulloch faithfully conveys what he calls "von Neumann's warning" to the audience. The gist of the warning has to do with suspected inadequacy of taking for granted that  $10^{10}$  neurons is not enough for accounting for the human brain. A signaling fact towards that assessment would be that the mere 300 neurons of the ant cannot afford to explain its behavior. There were some expressions of discomfort regarding such change in spirit:  $10^{10}$  was regarded as an adequate estimate for neuronal population just at the previous Fifth Macy conference.

It became gradually apparent that there was a confusion between synapses and actual neurons, made evident by the parallel confusion regarding whether to ascribe the analogy of computer's vacuum tubes to synapses or neurons. Each analogy relation would of course lead to different outcomes. Walter Pitts finally brought the discussion to a halt, proposing that von Neumann got his calculations wrong: Seemingly von Neumann was referring to synapses (although McCulloch did mention "neurons"). A neuron has  $2^n$  firing possibilities. "Therefore,  $2^n$ ,  $n$  is the number of distinct afferents which go to it, is the maximum. I don't see how you can get up to 1010 in a possible cell." That is how the artificial networks were conceived in the first place. It would also seem to follow that it is safe to assume that ants in fact have more than 300 "thinking" neurons.

This minor bleep did not stop von Neumann—it merely stopped *that* discussion. He remained undeterred, as we will see below (Von Foerster, Mead & Teuber 1950, p. 17).

<sup>499</sup> Published posthumously, almost a decade after von Neumann's untimely death in 1957, as Von Neumann 1966. His student Arthur Burks took the task of gathering, heavily editing and even completing von Neumann's manuscripts on his theory of automata.

Von Neumann started off discreetly, just letting slip at this point that he is developing a slightly dismissive stance regarding the implication of having McCulloch and Pitts' networks as a plausible counterpart for the nervous system's natural networks. He expressed that their neural model "has a meaning which concerns me at this moment a little less... They wanted to discuss neurons".<sup>500</sup> The way McCulloch and Pitts proceed to go about this, von Neumann continued, is by means of using what is "known in mathematics as the axiomatic method, stating a few simple postulates and not being concerned with how nature manages to achieve such a gadget".<sup>501</sup> But he to some extent justified this strategy, pointing out the already mentioned remarkable achievement of showing how anything amenable to rigorous description can be instantiated by the model. Indeed von Neumann cautiously endorsed the heuristic simplification attained after the encapsulating power of such idealized neurons, given that "one gets a quick understanding of a part of the subject by making this idealization".<sup>502</sup>

But then von Neumann, in stronger language, started to gradually convey the problems being brought up as side effects of such accomplishment. He asserted that the scientific duo "believed that the extremely amputated, simplified, idealized object which they axiomatized possessed the essential traits of the neuron, and that all else are incidental complications which in a first analysis are better forgotten".<sup>503</sup> Von Neumann wondered whether this move --namely, leaving aside the details that constitute a real neuron for the sake of the model's operability-- will ever reach wide acceptance. He doubted it.

Still in another instance, displaying a half-sympathy for the axiomatic strategy, von Neumann went on to describe a neuron as having "two states: It's excited or not. As to what

---

<sup>500</sup> Von Neumann 1966, p. 43

<sup>501</sup> Von Neumann 1966, p. 43

<sup>502</sup> Von Neumann 1966, p. 44

<sup>503</sup> Von Neumann 1966, p. 44

excitation is, one need not tell”.<sup>504</sup> In what seems as approvingly allowing for the bracketing of an otherwise important detail, von Neumann faithfully described McCulloch and Pitts’ neuron, specifying that “its main characteristic is its operational characteristic and that has a certain circularity about it: its main trait is that it can excite other neurons”.<sup>505</sup>

The last nod of approval towards McCulloch and Pitts’ networks is given right before von Neumann begins to point out its shortcomings –although from a different angle than what was remarked at the Hixon Symposium. Graciously mentioning another time the rigorous descriptive parsing accomplished by the model, he followed it with still another benefit, not without remarking that these outcomes carry a philosophical twist:

May I point out what follows from this from a philosophical point of view, and what does not follow. It certainly follows that anything that you can describe in words can also be done with the neuron method. And it follows that the nerves need not be supernaturally clever or complicated. In fact, they needn’t be quite as clever and complicated as they are in reality, because an object which is a considerably amputated and emasculated neuron, which has many fewer attributes and responds in a much more schematic manner than a neuron, already can do everything you can think up.<sup>506</sup>

Now the requirement of intellectual honesty had been fulfilled, having given the best possible depiction of the model “on trial”. Von Neumann went on to begin listing, in an almost accusative tone, the high prices that were paid for the theoretical package purchased, and the great losses obtained in return. In his words, next to the great benefits attained by the McCulloch and Pitts’ networks, “what is not demonstrated by the McCulloch and Pitts result is equally important”.<sup>507</sup>

There are three major issues left unsolved, or even further problematized, after assuming the

---

<sup>504</sup> Von Neumann 1966, p. 44

<sup>505</sup> Von Neumann 1966, p. 44

<sup>506</sup> Von Neumann 1966, p. 46

<sup>507</sup> Von Neumann 1966, p. 46

McCulloch and Pitts' networks. What von Neumann recounted here summarized the discomfort he felt for three years: 1) There is no proof anywhere that McCulloch and Pitts' networks actually occur in nature.<sup>508</sup> 2) There is no assurance that whatever was left aside in the axiomatization of the neuron is not of the utmost importance. 3) Most disturbingly...

It does not follow that there is a considerable problem left just in saying what you think is to be described. Let me try to put this in another way. If you consider certain activities of the human nervous system, you find that some of them are such that all parts of them can be described, but one is flabbergasted by the totality of what has to be described.<sup>509</sup>

Once he justified plenty why McCulloch and Pitts' networks were not the way to go for a cybernetic articulation in general --and of a nervous system in particular --, von Neumann advocated for the line of research proposed to Wiener in the 1946 letter. Primitive organic systems should be the first ones to be understood, putting special attention on bacteriophages. The feature of primitive life that one ought to focus on is self-reproduction. If one is able to recreate such an essentially biological feature, von Neumann would consider it the first qualitatively distinct step towards a practical understanding of a living system feasibly reproducible in a model. Recreating self-reproduction would be a concrete task that could give a distinctly profound insight into the nature of life. In faithfully cybernetic manner, he thought of a real world, physical model of this primitive system: he specifically called for an automaton occupying Euclidean three-dimensional space. The primitive automaton would arbitrarily measure some decimeters, and it would be constituted by the following elements:

---

<sup>508</sup> This points out to the devastatingly simple question, relevant these days (for example, in cognitive science), whether a proposed model, whichever it is, actually *occurs* (namely, it is at least part of a physically instantiated mechanism). It may be argued that despite the simplicity and relevancy of the inquiry, its sole mentioning generally fosters confusion. The question about the ontological status of models is one that has largely gone underground. This issue is not unrelated to the one dealing with entities sharing a common fundamental trait, differentiated quantitatively or qualitatively. The note on Scotus and Vico mentioned above applies (Chapter IV, section 4).

<sup>509</sup> Von Neumann 1966, p. 46

- + Four computing elements (three logical and one processing):
  - Disjunctive organ (for the function *or*)
  - Conjunctive organ (for the function *and*)
  - Inhibitory organ (for the function *not and*)
  - Sensing element (stimuli producing)
- + Four mechanical elements
  - A fusing element; this organ would weld parts together.
  - A cutting element; this organ would separate parts.
  - A muscle organ; this organ would provide motion.
  - A rigid organ; which would just provide structural rigidity.<sup>510</sup>

The automaton would operate in the following way. The “mother” automaton, in possession of the whole set of instructions for a model of itself, is to be left freely floating in a liquid substance, surrounded by physical parts of itself, freely floating as well. When the automaton would bump into a part, tests would be performed within itself in order to identify what kind of part it is (a muscle contracts, a rigid bar does not change, etc.). It was not specified how this contact process would precisely operate, but von Neumann suggested that there might be a “sensing” device, actuated by the muscle element. The muscle organ would also provide motion to the fusing and cutting organs. The automaton would gradually put all the parts welded

---

<sup>510</sup> Von Neumann gathered a small group of people at the Institute for Advanced Studies at Princeton University, likely as a preparation for his presentation at the Hixon Symposium. What was said in those talks, unfortunately never recorded or published, gave the most details regarding his early theory of automata –the “kinematic model”, as Burks baptized it. Arthur Burks attempted to reconstruct what transpired there. In terms of the “elements” of the self-reproducing automaton, von Neumann reportedly named them as follows:

*A stimulus organ* receives and transmits stimuli; it receives them disjunctively, that is, it realizes the truth function “p or q.” *A coincidence organ* realizes the truth-function “p and q.” *An inhibitory organ* realizes the truth-function “p and not-q.” *A stimuli producer* serves as a source of stimuli. The fifth part is a *rigid member*, from which a rigid frame for an automaton can be constructed... These connections are made by a *fusing organ* which, when stimulated, welds or solders two parts together... Connections may be broken by a *cutting organ*, which, when stimulated, unsolders a connection. The eighth part is a *muscle*, used to produce motion (Von Neumann 1966, p. 81. See also Burks 1969, pp. 2-3).



together, and as the last step, it would transfer the set of instructions of itself,<sup>511</sup> for the “daughter” automaton to continue the reproductive process. In such way, the “miracle” of life would be a step closer to be fully artificially recreated in the metal, thus providing the kind of cybernetic true knowledge only possible when breaking down the problem into an engineering issue.

A number of issues can readily be identified in the previous description of artificial self-reproduction. Firstly, von Neumann conspicuously avoided touching upon the issue of an energy source –all the more telling, after advocating for a future logic that should be sensitive to issues of thermodynamics (the amount of steps to reach an answer would have practical consequences in friction, decay, erosion, heat dissipation, etc.). Probably, he was going to add this feature later.<sup>512</sup> However, it is already signaling the strategy that he is going to be using: axiomatizing elements and black-boxing issues that otherwise would throw paralyzing obstacles for the experimental research to continue. After all, he did justify the same strategy used by McCulloch and Pitts’ networks to *some* extent.

Also, there is no elaboration as to how the grasping of the floating parts would be performed. Presumably by means of a movable mechanical “arm”, helped by both the muscle organ and the energy source, the part would be grabbed and put in the right spot. The same goes for the welding and cutting procedure; there is no mention of the way in which these would

---

<sup>511</sup> Let us recall Turing’s efforts to make explicit the notion of an algorithm by means of rigorously describing its behavior –which facilitated the Turing machine (chapter III, section 2). Von Neumann recognizes here the value of Turing’s call for rigorously and completely describing parts –i.e., the elements of his Turing machine-- later embraced by McCulloch and positively recognized by him (section 3, above). In order to circumvent the problem of a machine abstracting a pattern for the structure of the “baby” machine, it would be more efficient to just transmit the clear set of instructions:

In any conceivable method invented by man, an automaton which produces an object by copying a pattern will go first from the pattern to a description and then from the description to the object. It first abstracts what the thing is like and then it carries it out. It’s therefore simpler not to extract from a real object its definition, but to start from the definition. To proceed in this matter one must have axiomatic descriptions of automata. You see, I’m coming quite close to Turing’s trick with universal automata, which also started with a general formal description of automata (Von Neumann 1966, p. 83).

<sup>512</sup> Such is Arthur Burks’ speculation (Von Neumann 1966, p. 82).

work. Or even for the all-important gadget that would function as a “recognizing device”, which would presumably provide information for the proper accommodation of the found part. There is practically nothing said about this.<sup>513</sup>

## **5. The Theory of Automata: From embodiment to abstraction.**

Despite the cautious justification for obviating these empirical aspects for the sake of the advancement of the theory, one can safely assume that this might have been a cold comfort for von Neumann. He was not only keenly aware of the cybernetic tradition of demonstrating knowledge by means of “recreating” it in the metal, as showed by Grey’s tortoises, Shannon’s rat, Ashby’s homeostat and even the early Wiener’s uncanny AA predictor. Von Neumann demonstrated a personal leniency, unlike many mathematicians, to implement his ideas in the flesh, being the primal example what came to be known as the von Neumann-architecture computer.<sup>514</sup> Leaving aside the truly mechanical aspects of the automaton was likely not something he cherished. In fact, it is recalled that he would be seen walking around Princeton University with a smiley face and big boxes filled with Meccano pieces of tinker toys.<sup>515</sup> After

---

<sup>513</sup> This is how von Neumann would scantily describe the self-reproduction of the automaton, to the best recollection of Burks:

The constructing automaton floats on a surface, surrounded by an unlimited supply of parts. The constructing automaton contains in its memory a description of the automaton to be constructed. Operating under the direction of this description, it picks up the parts it needs and assembles them into the desired automaton. To do this, it must contain a device which catches and identifies the parts that come in contact with it... Two stimulus units protrude from the constructing automaton. When a part touches them tests can be made to see what kind of part it is (Von Neumann 1966, p. 82).

Arthur Burks later attempts to explain how this might work, admittedly infusing this explanation with his own educated guess for what von Neumann might have had in mind (Burks 1969, pp. 3-6).

<sup>514</sup> See the note above (section 1) on the dispute regarding calling von Neumann “the father of the computer.”

<sup>515</sup> Herman Goldstine recalls a happy-faced von Neumann buying the biggest set of Meccano toys available, and enjoying tinkering with them at the university. Once his (probably only true) friend Stanislaw Ulam convinced him of the mathematical advantage of switching his model to a two-plane reality, he gave away the tinker toys to Oskar Morgenstern’s son (Goldstine 1993, p. 278).

all von Neumann did have the cybernetic mandate of knowledge as re-creating as an integral part of his epistemology.

Conversation with colleagues gradually convinced him to leave aside the “problematic” approach to knowledge-gathering having physical substrata as a grounding factor. The Polish mathematician Stanislaw Ulam (1909–1984) gradually convinced him to leave aside bodily three-dimensionality to re-state the problem in a two-dimensional plane. Being an outstanding mathematician himself, von Neumann acknowledged that his three-dimensional “kinematic” model lacked the mathematical “elegance” that might be required for future generations to build upon it. With the new proposal, void of a balancing physical checkpoint, it was expected that he would be able to give attention to the elaboration of a logic of reproduction that he claimed was non-existent.

What von Neumann developed after the suggestion of his colleague is what came to be known as the theory of cellular automata.<sup>516</sup> It consisted of two-dimensional squares, called lattices, which would --under specific instructions based upon the condition of the contiguous squares (occupied or not) -- “grow” and “reproduce”. These two-plane automata would show a fairly complex behavior stemming from simple rules --which would allegedly fulfill the evolutionary evidence of complex structures coming from simpler ones (as it is shown in the biological reproduction of entities, commencing in unicellular beings). All this was to occur on a flat geometrical plane, completely void of Euclidean mass and thus free from the extra (and implementationally and operationally “crippling”) complications of the three-dimensional world.

The details of this historic theory are not the concern of this work. What is important to remark is the effective outcome of von Neumann’s change of methodology. As indicated above,

---

<sup>516</sup> Or just the “theory of automata,” since the kinematic model was, as we saw, discarded. This new two-dimensional approach is dealt with in the second part of von Neumann’s posthumous book (Von Neumann 1966).

von Neumann thought of this later work as the most fundamental of his whole career. He used to talk about it with solemnity, effectively devoting to it the last decade of his life --to the chagrin of government officials and colleagues who wanted him more involved in work with computing machines. However, despite the effort of a few disciples,<sup>517</sup> the theory of cellular automata fell into oblivion,<sup>518</sup> after von Neumann died.<sup>519</sup> It was attempted to be rescued<sup>520</sup> but it remains an effort whose punchline never arrived.<sup>521</sup>

Most importantly perhaps, cybernetics missed a fundamental arm in its own body of theoretical tenets: the utmost importance of the physical instantiations of their theoretical pursuits. Von Neumann effectively ended up taking the same path that he criticized just a few years earlier, provoking a focus in the abstracted model as if it would be the object of study itself. The importance of the structure of the object was shown to be particularly relevant and helpful when the mapping of possible behavior became extremely complex and beyond feasible management in practical terms. But once one takes a peek into the structure itself, the difficulty found is “flabbergasting” –as von Neumann would like to refer to it. And so the nuanced epistemic trip taken by von Neumann followed these stages: 1. the study of virus; 2. the attempt to recreate “primitive” self-reproduction; 3. the paralyzing effects of the physical details of the mechanical model; 4. the retreat into the non-physical realm. And this last one is precisely what

---

<sup>517</sup> Mainly Arthur Burks. He published about von Neumann’s automata theory in Burks 1959; 1960; 1963; 1963b, 1966, 1970; 1974; 1975 and of course Von Neumann 1966, where Burks did the heavy editing and completion.

<sup>518</sup> Boden 2006, pp. 890-892

<sup>519</sup> Von Neumann reverted to the Catholic faith that he adopted when he married his first wife. He was accompanied by a Benedictine priest the last days of his life, at the end of which he received the Last Rites (McRae 1992, pp. 378-379).

<sup>520</sup> Probably the most serious attempt was NASA’s commission to study the feasibility of constructing and using self-replicating machines for space exploration. This technical report has a section on asteroid mining that has lately re-gained some momentum, due to the renewed interest in such potentially pecuniary enterprise (See Freitas & Gilbreath 1980).

<sup>521</sup> The relatively novel field of Artificial Life does not take sides with the kinematic model as opposed to the cellular model, but it does acknowledge the former as the first model that fully grasped the essence of ALife –a fortiori, a successfully self-reproducing automaton would qualify as *living* for this discipline (Langton 1989). Theorizers of nanotechnology do recognize von Neumann’s kinematic model as conforming the root of their envisaged “molecular assembler” (See Drexler 1986, ch. 4).

happened to McCulloch and Pitts, which von Neumann was quick and effective in pointing out – to the eventual resigned acceptance of McCulloch. Seemingly, von Neumann’s suggestion to focus on the structure of very simple and primitive entities was confronted with the reality that they were not simple enough for mechanical re-creation. Once physical instantiation is out of the equation, cybernetics is left with little that would make it different and unique, let alone revolutionary.

## VII. Cybernetic tensions: Further philosophical considerations.

### 1. Consequences of the irrelevancy of materiality for the nature of a machine.

Let us recall that the origin of cybernetics is situated in the context of research on anti-aircraft defense --the nature of the mechanically operated anti-aircraft gun, the soldier's almost symbiotic relation with the operating machine, both processing in sync the information available about the enemy target. The question "what kind of machine have we placed in the middle?"<sup>522</sup> was attacked by Wiener by means of a behavioral understanding of both the machines *and* the human in between them, under one and the same theoretical framework. This was made possible thanks to a more extended view of what a machine stands for, taking cues from Turing's suggestions regarding the mechanical essence of an algorithm --or the algorithmic essence of a machine.<sup>523</sup> Shannon's work on information theory anchored both realms even closer, both seamlessly crossed-through by information flow. Considering these black boxes as displaying active "aims" was the feature that distinguished cybernetic talk from the already entrenched behavioral psychology of the time.

The approach seemed promising, and to some extent, it flourished. Understanding adaptive --even intelligent-- behavior from an exclusively external point of view (leaving inquiries into the nature of the working structures aside) produced striking results. There were the mice and the tortoises of Shannon and Walter, which attracted attention beyond cybernetic circles, occupying pages in mainstream magazines and booths in science festivals. There was McCulloch and Pitts' networks, which mathematized certain purported behaviors of the brain

---

<sup>522</sup> Edwards 1996, p. 197

<sup>523</sup> Chapter III, section 2

and the nervous system, serving as the platform for von Neumann's construction of the first stored-program computer. They seemed to have hit the sweet epistemological spot regarding the behavior of entities in nature, which would otherwise be riddled with vitalist assumptions. Rigorous behavioral descriptions could be unambiguously measured. It seemed to be the right path chosen for understanding man, faithful to (and consistent with) the mechanistic tradition. After all, for the cyberneticians such tradition still effectively was the strongest one (if not the only one) in modern science, if we are to maintain that "to understand" a thing means "to provide the mechanism" for such thing.

The homeostat's creation was a necessary outcome of such a position --it was just a matter of time. Cybernetic tenets pointed in that direction right from the start. But cyberneticians might not have gone *that* extra mile. When Ashby entered the picture, it probably rang the death knell of a project whose radically materialistic ontology might not have shown itself with all its colors to the cyberneticists themselves. Mechanical mice and electric tortoises, wandering autonomously, avoiding obstacles and "figuring out" their way out of a maze, were creatures that immediately were predicated of as being essentially the same as the ones that nature boasts --just simpler, but with possibilities of growth in complexity. The homeostat was another story. With this machine, the learning process of a kitten (who is first attracted to a flame, but after being burnt, begins to avoid it) does not qualitatively stand out from the situation of a plastic ball finding its way out of a moving box with a hole.<sup>524</sup> Under the hood, both situations were underpinned by random processes that would lock and capitalize on those outcomes that benefit the system --to the dismay of Julian Bigelow, who saw no virtue in that.<sup>525</sup>

---

<sup>524</sup> Chapter V, section 4

<sup>525</sup> Chapter V, section 4

Let us recall the damning examples of the young philosopher Richard Taylor when attacking Bigelow, Wiener and Rosenblueth's 1943 article.<sup>526</sup> The behavior of a missile with a rope attached to its nose, being pulled by a submarine, would also count as behavior with an aim (along with that of a self-guided missile, the example given in the criticized article).

Trivialization seemed to have been a looming outcome always present. In the case of the homeostat at Macy, trivialization, although it showed its face, seemed to have worked its way to those very foundations –at least in what pertains to the reasons for an important part of its success and its intentions-- which lured brilliant minds into cybernetics. The cybernetic project was partly appealing because of the promise of being able to construct artifacts that would not only emulate, but eventually replace those constructed by nature –thanks to the discovered knowledge of a machine-basis common to both, the natural and the artificial. That promise proved to be exciting and extremely attractive (probably more so in the context of a Cold War threatening the existence of natural life as we knew it). But then came a cyberneticist who faithful to cybernetic tenets *in extremis*, proclaimed that the very features that we understand as the ultimate pointers for life, namely self-organization and adaptation, are in no way particular to living entities –in fact, they are a necessary condition of existence for material objects coping, such as crystals.<sup>527</sup> And randomly so at that. And he had a machine to show it. Ashby, however, was not without a tradition of thought that heavily relied on experimentation and physical recreation.

It has been indicated that scientists in the United Kingdom were very interested in cybernetics right from the beginning, influencing their own careers and even forming a small,

---

<sup>526</sup> Chapter IV, section 2

<sup>527</sup> Chapter IV, section 2.



invitation only, regular discussion gathering.<sup>528</sup> Ross Ashby, producing his own insights squarely within cybernetic thinking at the start of the 1940s<sup>529</sup> –almost a decade before Wiener’s catalyzing book<sup>530</sup>—seemingly performed a tune-up of the enterprise’s theoretical core. Some years later, he recounted in his own book<sup>531</sup> the intellectual adventure into the solidification of cybernetic science, providing valuable insight into his life’s work --which was to some extent, as previously shown, seemingly more concerned with cybernetic *dicta* than that of the original cyberneticians themselves.

A crucial point that Ashby found as fundamental to the whole enterprise was the notion of *control*. We have seen that Norbert Wiener was well aware of its importance, identifying James C. Maxwell’s 19<sup>th</sup> century mathematical treatment of governors<sup>532</sup> as the theoretical foundational document for cybernetics.<sup>533</sup> Wiener’s famous book carried the noun “control” in its subtitle. However, Ashby identified that a previous, more fundamental notion, was the one that made possible control in the first place; this more primal notion was that of a *machine*. The strategic move that extended the notion of machine into realms heretofore untouched by it is located here. This extension is ingeniously articulated by means of choosing geometry as a cybernetically suitable theoretical counterpart. Ashby described how this mathematical branch went from being ruled (and constrained) by physical instantiations, to actually containing them later on. This extension gained for geometry unprecedented *control*...

There was a time when “geometry” meant such relationships as could be demonstrated on three-dimensional objects or in two-dimensional diagrams... In those days a form which was suggested by geometry but which could not be demonstrated in ordinary space was suspect or unacceptable.

---

<sup>528</sup> The Ratio Club (chapter 2, section 2).

<sup>529</sup> Ashby 1940

<sup>530</sup> Wiener 1948

<sup>531</sup> Ashby 1952

<sup>532</sup> Maxwell 1868

<sup>533</sup> Chapter II, section 1

Ordinary space dominated geometry. Today the position is quite different... Today it is geometry that contains the terrestrial forms, and not vice versa, for the terrestrial forms are merely special cases in an all-embracing geometry. The gain achieved by geometry's development hardly needs to be pointed out. Geometry now acts as a framework on which all terrestrial forms can find their natural place, with the relations between the various forms readily appreciable. With this increased understanding goes a correspondingly increased power of control.<sup>534</sup>

Cybernetics follows this path: it has the same relation to an actual machine that geometry has to an actual, 'terrestrial' object. Hence, cybernetics "takes as its subject-matter the domain of 'all possible machines,' and is only secondarily interested if informed that some of them have not yet been made, either by Man or by Nature".<sup>535</sup> In such way, Ashby managed to identify the crucial elements that make a machine what it is --in plain philosophy, its 'essence'-- thus furnishing its particular capacities. And the mechanization of entities entails the exertion of controlling power over them. For Ashby, such articulation was isomorphic with scientific explanation, in so far as one understands "explanation" in this context as isomorphic to providing an "explanatory *mechanism*". A mechanism entails a machine --and we managed to find it in nature. Modern science is on the right path.

Aware that this accomplishment is something remarkable in terms of extending powers of tractability to realms previously beyond mechanistic reach, Ashby claimed that it is "one of the substantial advances of the last decade [1950s] that we have at last identified the essentials of the 'machine in general'".<sup>536</sup> Here Ashby is in fact referring to himself, as he was correctly aware that "[m]any a book has borne the title 'Theory of Machines', but it usually contains information

---

<sup>534</sup> Ashby 1956, p.2

<sup>535</sup> Ashby 1956, p. 2

<sup>536</sup> Ashby 1962, p. 260

about mechanical things, about levers and cogs”.<sup>537</sup> It is he who comes up first with a rigorous definition of what a machine is:

[A] machine is that [whose] internal state, and the state of its surroundings, defines uniquely the next state it will go to.<sup>538</sup>

This over-encompassing definition, the embodiment of a merge between a theory of machine behavior and negative entropy, gained unprecedented “control” for cybernetics by means of extending the reach of what can be understood as a full-blown machine. Ashby confided that in order to see clearly a machine for what it is, a behavior-based premise had to gain precedence, so as to dispel probable metaphysical or epistemological misunderstandings of what essentially constitutes a machine. And so he expressed that

Before the essentials could be seen, we had to realize that two factors had to be excluded as irrelevant. The first is “materiality” –the idea that a machine must be made of actual matter, or the hundred or so existing elements. This is wrong, for examples can readily be given... showing that what is essential is whether the system, of angels and ectoplasm if you please, behaves in a law-abiding and machine-like way.<sup>539</sup>

By musing about machines made up of “angels and ectoplasm”, Ashby wanted to push the point across unambiguously: the materiality of a machine is entirely circumstantial. Not only has this observation closed the case regarding whether something purporting to be a mechanism is a machine --the answer being obviously in the positive. It also displaces the importance we have traditionally (both scientifically and philosophically) attributed to physicality --in the sense of

---

<sup>537</sup> Ashby 1956, p. 1. Also, see the introductory remarks to chapter V.

<sup>538</sup> This might be, in the view of this writer, the only rigorous definition of a machine ever advanced. History of science does not seem to show any precedent. Alan Turing and Heinrich Hertz might have come the closest, but they still fell short of providing a full-blown definition (See chapters III, section 2; and V, section 1, respectively). Ashby 1962, p. 261.

<sup>539</sup> The quote continues: “Also to be excluded as irrelevant is any reference to energy, for any calculating machine shows that what matters is the regularity of the behavior –whether energy is gained or lost, or even created, is simply irrelevant” (Ashby 1962, p. 260).

materiality. In this sense the physical sciences themselves might see their strong theoretical attachments to materiality seriously affected. The definition of the physically real would no longer entail material measurability. This seems to be a historic instance where an epistemological backhand move was surreptitiously passed. And it could have resulted in deep tensions within the cybernetic ontology --committed since the start to a physical grounding in a radically Viconian manner.<sup>540</sup>

Furthermore, it could have also signified the coming of an epistemic turn --still in the making<sup>541</sup> --inimical to a classical cybernetic epistemology: life as necessary order coming out of chaos, or negative entropy as the natural course of things coping in the world; both against what Wiener would defend.<sup>542</sup> Let us recall Ross Ashby's heated exchange at the 1942 Macy conference.<sup>543</sup> There, some in the audience were taken aback by Ashby's assertion that his homeostat displayed exactly the same learning and survival that a cat displays when it learns and survives. Ashby reminded them that "to learn", in cybernetic lingo, necessarily entailed, "objectively" speaking --without retorting to psychologisms or introspections-- a *consequent change of behavior*. An "animal" can be observed to "learn" in order to "survive": In strictly physicalist terms, that translates to a system adapting in order to reach and maintain equilibrium. To lose the capacity to adapt means to die. Seemingly, some cyberneticians (e.g., Julian Bigelow) were not too comfortable with the kind of conclusions that Ashby showed they were compelled to reach, if they were to still hold on to the basic tenets of cybernetics. To recall, Ashby indeed bridged the artificial inanimate with the organic living in no vague terms.<sup>544</sup>

---

<sup>540</sup> Chapter IV, section 4.

<sup>541</sup> Chapter IV, section 2

<sup>542</sup> Chapter IV, section 2

<sup>543</sup> Chapter V, section 5

<sup>544</sup> Recalling the relevant quote by Ashby:

The moment we see that "adaptiveness" implies a circuit and that a circuit implies an equilibrium, we can see at once that this equilibrium must be of the stable type, for any unstable variable destroys itself. And it

Ashby was relentlessly taking a machine-ontology based upon observable behaviors to its ultimate consequences, confident of its epistemological strength and methodological maneuverability. He showed this unmovable confidence throughout the years in his own theoretical development –if anything, the certainty grew stronger. Hence, at the Macy conference, he convincingly showed to the cyberneticians that if behavior, *tout court*, is what tells us what a thing is, then they have to concede that his presented machine is alive –or, at the very least, that living entities live and survive in the same way that this machine does. If one considers that a cat behaving differently after training can be said to have learned, one is compelled to accept that the machine has learned, given that it behaved differently after training.<sup>545</sup> This was in fact a crucial aspect of cybernetics.

Ashby was however somewhat encouraged by the criticism he received regarding the relative simplicity of the exposed entity (his machine), of which it could be said, through a behavior-only lens, that it was indeed alive and learning –but not precisely “alive and kicking”. More was needed. Mindful of the feedback received at the conference, Ashby allowed for higher layers of complexity in the structure of the machine to be an issue to be dealt with –so that it could be more justifiably comparable with the behavior of biological entities. In fact, he already had embarked upon the construction of DAMS. And as we saw,<sup>546</sup> this more complex automaton did not show any pattern of self-organizing behavior whatsoever, to Ashby’s disappointment. For a moment, while the homeostat showed its glory, one could witness the argument for a state of

---

is precisely the main feature of adaptive behavior that it enables the animal to continue to exist... I must say unambiguously what I mean by ‘alive’. I assume that if the organism is to stay alive, a comparatively small number of essential variables must be kept between physiologic limits. Each of these variables can be represented by a pointer in a dial... This, of course, is in no way peculiar to living organisms (Ashby 1953, p. 73)

<sup>545</sup> The “post hoc ergo propter hoc” fallacy was seemingly not flagged, possibly due to the context of the necessity of an exclusively mechanical explanation –which severely reduces the possibility of entertaining alternative explanations.

<sup>546</sup> Chapter V, section 5

nature that coping mechanisms inevitably reach, through trial and error towards self-organization, bringing order out of chaos --and eventually, life. This, most likely, would not have been to Wiener's liking.

## **2. Consequences of a machine's isomorphism with highly complex entities.**

An exclusively behavioral-based understanding of a coping system was explanatorily adequate when the entity did not exceed a certain threshold of complexity. Once that threshold was crossed, then explanatory models would explode towards the practically unfeasible. During the Hixon Symposium,<sup>547</sup> McCulloch considered the possibility of having his networks utilized for mapping thought processes, accepting that the adequate implementation might now be in its infancy, but mostly due to an issue of availability of materials --and not to the theory itself. Thus, we saw how McCulloch referred somewhat compassionately to von Neumann's physical instantiations of his networks --namely, the first stored-program computing machines.<sup>548</sup>

Von Neumann would more than concur with McCulloch's hyperbole. In fact, building the computing machines using McCulloch's (and Pitts') model helped him notice precisely such inadequacy, and in productive ways. Von Neumann saw in the very nature of the model -- namely, one that maps network behaviors rather than structures-- an inherent incapacity for capturing a very complex entity. It was a model surely amenable to be instantiated in a

---

<sup>547</sup> Jeffress 1951

<sup>548</sup> To recall, McCulloch said that

Neurons are cheap and plentiful. If it cost a million dollars to beget a man, one neuron would not cost a mill. They operate with comparatively little energy. The heat generated raises the blood in passage about half a degree, and the flow is half a liter per minute, only a quarter of a kilogram calorie per minute for 10<sup>10</sup>, that is, 10 billion neurons. Von Neumann would be happy to have their like for the same cost in his robots. His vacuum tubes can work a thousand times as fast as neurons, so he could match a human brain with 10 million tubes; but it would take Niagara Falls to supply the current and the Niagara River to carry away the heat (McCulloch 1951, p. 54).

computing machine –as he efficiently showed-- but thoroughly inadequate to map a brain. Being the latter an object of very high complexity, encompassing all the possible behaviors of such system would render the outcome of the model realistically intractable --even if *in principle* feasible. Von Neumann thus gave an image-rich (negative) verdict of McCulloch and Pitts’ networks in regards to being a model suitable for understanding the nervous system.<sup>549</sup>

John von Neumann also expressed this dissatisfaction in the aforementioned infamous letter to Norbert Wiener<sup>550</sup> at two main levels. First, McCulloch and Pitts’ networks seemed to be a model unfit for studying a highly complex object --knowledge likely acquired and/or reinforced with his own experience while building a computer based upon it (and which already occupied a physical space of considerable dimensions). Second, leaving the model aside, studying the nervous system *at all* seemed too far-fetched and even hubristic, given its extraordinary --probably the most extreme-- complexity.

Putting the two criticisms together, von Neumann fleshed out a proposal: Going “back to the basics” --also at two main levels. First, choosing the simplest living system, and start from that; he suggested a virus. Second, to forego McCulloch and Pitts’ networks of behavioral mapping-- they have proven their efficacy for other realms, such as automated computing processes, relieving humans from such tedious task-- and to focus instead on the very *structure* at hand. To recall, for von Neumann, after a certain threshold of complexity, it is more feasible to describe the structure of the object than its *behavior* --hence effectively reversing the modeling

---

<sup>549</sup> Recalling the relevant quote by von Neumann...

I think that it is quite likely that one may give a purely descriptive account of the outwardly visible functions of the central nervous system in a humanly possible time. This may be 10 or 20 years --which is long, but not prohibitively long. Then, on the basis of the results of McCulloch and Pitts, one could draw within plausible time limitations a fictitious "nervous network" that can carry out all these functions. I suspect, however, that it will turn out to be much larger than the one that we actually possess. It is possible that it will prove to be too large to fit into the physical universe. What then? Haven't we lost the true problem in the process? (Von Neumann 1951, p. 34).

<sup>550</sup> Von Neumann 2005

roles. Attaining a full description of the elements constituting the studied entity would still be indeed difficult, but now it remains within the realm of feasibility –and that is a huge gain. However, there he introduces a very un-cybernetic idea for a “true” understanding of a complex object. It is worth taking another look at what he said...

I would, however, put on “true” understanding the most stringent interpretation possible: That is, understanding the organism in the exacting sense in which one may want to understand a detailed drawing of a machine, i.e. finding out where every individual nut and bolt is located, etc. It seems to me that this is not at all hopeless... I suppose (without having done it) that if one counted rigorously the number of “elements” in a locomotive, one might also wind up in the high ten thousands. Consequently this is a degree of complexity which is not necessarily beyond human endurance.<sup>551</sup>

The analogy here with a machine is not casual, and more so coming from an unapologetic cybernetician –which he certainly was, despite his vocal disagreements. The immediate usefulness of the machine imagery here is the entailed amenability of being taken apart piece by piece --literally, one by one-- as one could do, despite the daunting task, with an airplane or a ship. Cybernetics certainly aimed to understand living organisms as machines; however, the extension of the notion of machine upon them was clearly based on the *behavior* of the former, and not upon the *structure* of the latter. As an outcome, for von Neumann, the still gigantic difficulty of the job would be now quantitative and not qualitative in nature --and that is where the ‘huge gain’ is located. The complex entity would now be at least tractable --which is not the case with the unworkable amount of variables resulting from the exact description of the

---

<sup>551</sup> Von Neumann 2005, p. 280



behavior of a very complex object (like the brain). Such is the gain of turning into an “exacting sense”<sup>552</sup> of rigorously understanding something *qua machina*.

In order to have a clear and mechanical model that could be neatly disassembled (at least theoretically), the object itself has to be sufficiently away from that threshold of complexity that stunned von Neumann in the first place. And because cybernetics aimed at understanding living beings in a full-blown mechanical way, von Neumann elected a virus as the entity to start with. However, we know that right in the very letter of complaint against McCulloch and Pitts’ networks (as a model for a brain), he also expressed his shocked amazement at the complexity of even the simplest entities of the biological world.<sup>553</sup>

Once we shift attention from a complex entity’s behavior to its structure, the employed model risks entering into a one to one mapping with the objects’ features, potentially bringing an unsurmountable intractability. Realizing the staggering complexity of even the simplest living entities, von Neumann not only settled for the study of viruses, but within that realm, he advocated focusing on its capacity for reproduction. It would seem that it is more reasonable to identify a feature of an organism that at least looks amenable to being recreated. All the more so if such feature paradigmatically characterizes an entity as living: self-replication --certainly present in a virus. That should provide us with a still audacious, but more realistic, task to accomplish.

True to his cybernetic spirit –the same spirit that used the McCulloch and Pitts’ networks for building the first computing machines-- von Neumann embarked upon the task of recreating it *mechanically*. His “kinematic model” came to fruition, where an eventually constructed entity

---

<sup>552</sup> Von Neumann 2005, p. 280

<sup>553</sup> To recall, he candidly confessed that...

After these devastatingly general and positive results [from the McCulloch and Pitts networks] one is therefore thrown back on microwork and cytology –where one might have remained in the first place... Yet, when we are in that field, the complexity of the subject is overawing (Von Neumann 2005, p. 278).

should be able to navigate through a host of floating resources, and use them to make a copy of itself. However, this task was proven to be exceedingly more difficult than expected –arguably, until today. Issues of solderability, separation and energy sources crippled the problem right from the start. This pushed von Neumann to consider an alleged equivalent of this self-reproducing entity, but now entirely living in logical space. He was finally persuaded to switch to a disembodied model, in the form of a logic of lattices –a two dimensional model that by its very nature avoids the problems entailed by the actual physicality of a self-reproducing automaton. Although von Neumann died before completing what for him was the most important undertaking of his life, the project was somewhat briefly followed up by a few disciples, only to end up in relative obscurity ever since.

### **3. An ontological slippery slope: From complexity to disembodiment.**

As previously pointed out, cybernetics' perch for model building was not radically new. Although it became right from the start a sort of defining feature for the scientific enterprise, the material recreation of a studied object has deep roots in the Western tradition of philosophy. The manner in which cybernetics assumed this feature (model buildability), was based upon the behavior of the object –readily pointed out by Wiener on the American side and by Ashby on the British side. The strategy proposed by von Neumann, however, pushed for a preceding “return” to the structure of the object per se. The basis for this, to repeat, was that after a certain threshold of complexity is reached, it is more attainable to map the object's structure directly, rather than comprising the gigantic plethora of all its possible behaviors. Still within a cybernetic spirit, von Neumann nevertheless pushed ahead with the necessary buildability of the obtained modeled

object, however now based upon its structure, and not its behavior. Put back by the seemingly intractable complexity of micro-biology, not even a virus seemed amenable to be structurally cracked for epistemological purposes via a one-to-one descriptive mapping and subsequent recreation. Von Neumann tried to focus at least on reproduction alone, but he ended up in the logical space of two-dimensional lattices. That must have signified a rift in the cybernetic embodied epistemology.

Although von Neumann and Ashby differed in regards to their approach to a behavior-based understanding of a surviving entity, both did accept it as a legitimate and appraisable cybernetic methodological tool, as presented and depicted by Wiener, Rosenblueth and Bigelow's 1943 founding article –where they distinguished it from mere behaviorism *tout court*.<sup>554</sup> When referring to the accomplishment of McCulloch and Pitts' networks in demonstrating that behavior can be totally and rigorously described, Von Neumann had positive thoughts about it –so much so that he went off to build a computing machine, the first of its kind, based upon this approach.

However, von Neumann applied theoretical “breaks” when confronted with the attempt at applying these behavioristic lenses to something extraordinarily more complex than a computer machine --namely, the nervous system and the brain. Once there, McCulloch's somewhat mocking suggestion (that one would need the Niagara Falls to provide the electric current and the Niagara River to dissipate the heat) was actually topped by von Neumann. He suggested instead --surpassing McCulloch's picturesque rhetoric-- that a description using McCulloch and Pitts' model may take more time than that of a human's lifetime and its material dimensions exceed those of the physical universe. In face of this, a look into the actual structure of the

---

<sup>554</sup> Rosenblueth, Wiener & Bigelow 1943, p.22. Also chapter VI, section II.

system, whose description might be less complex than its behavior, seemed like a way out – eventually to no avail, as we now know.

Hence, albeit with different technical motivations but with the same cybernetic intentions, both thinkers suffered the backlash of tinkering with actual structures –as opposed to manipulating models of their behavior. It is important to signalize that at this point both had already crossed the methodological line in the sand, entering into non-cybernetic territory. The investigation of the nitty-gritty stuff operating *under the hood* in the ‘teleologically behaving’ mechanical systems is the task of the traditional science ‘underdog,’ not of the new science of cybernetics, equipped with such an impressive novel theoretical edifice. Even if cybernetics did not have as an aim to transcend or replace physics, it did want to extend it –in its own eyes, to actually ‘defend’ it.<sup>555</sup> The price to pay for this trespassing came from cybernetics itself.

It has been several times pointed out that reliance on physical instantiations of its own theoretical proposals was part and parcel of the cybernetic impetus. This was untenable after an anti-cybernetic ‘heresy’ was committed. Once the ‘harm’ was done, it became evident for both cases that the cybernetic mandate for physically grounding the theory, complicated things to the point of no solution. Where Ashby had to concede even on his own terms an augmentation of complexity within the structure behind the homeostat’s behavior (to be able to account for a more life-like comparison, thus giving birth to the doomed project of DAMS, where no order was to be attained), von Neumann had to plunge into the structure of complex entities (bypassing McCulloch and Pitts’ networks, engaging into the construction of automata according to cybernetic dictum, only to then retreat into the non-physical due to the very complexity that made him turn away from the artificial neural networks in the first place).

---

<sup>555</sup> Chapter I, section 4 –regarding the first Macy conference.

When the cybernetic mandate for behaviorally articulating the studied object fell short, then entering the realm of structure to better understand the operating entity seemed to be the next, scientifically motivated, viable step. But this move clashed with the cybernetic mandate of instantiating theories *in concreto*. More precisely, the mandate for buildability cancelled the possible import of going out of the cybernetic way into structural insights, given that this very need for physicality rendered the recreated structures as caricatures of the modeled object. After von Neumann's kinematic model failed he had to retreat into logical research. After Ashby's DAMS failed he ended up studying systems. Unexpectedly, the last cybernetic move was to find refuge in the realm of the disembodied, betraying a main column (arguably its strongest differentiator) of this "new science".<sup>556</sup>

---

<sup>556</sup> Wiener 1948, p. 28

## **Concluding remarks: The nature of a machine and the collapse of cybernetics.**

The main aim of the preceding chapters has been to establish the strong correlation between the evolving role of the notion of a machine, its role in scientific explanation and the subsequent outcome in the scientific pursuits it fostered, with particular emphasis on cybernetics –a scientific endeavor whose very core pivoted around the ontology of a machine.

In an attempt to nest the theoretical pillars of cybernetics within the evolution of ideas, the context that preceded the movement was laid out, starting with identifying an impetus for addressing certain fundamental mathematical problems presented at the beginning of the previous century by David Hilbert. Several European thinkers, ranging from Gottlob Frege to Kurt Gödel, engaged into an attempt of founding mathematics upon logical and/or formalistic grounds, triggering the foundational question for the rigorous definition of the notion of an algorithm. Alan Turing picked up the task, devising an abstract machine whose behavior would exactly define what an algorithm is supposed to do --and hence, what it is supposed to *be*. This definition of an entity based upon its behavior, carried some far-reaching consequences, as it was indicated later. Turing's puzzling reversal regarding the possibility of intelligent machinery allows to show a possible link between cybernetics and the evolution of Turing's thought.

The chapters pertaining to cybernetics proper (the core of this work) were intended to be, beyond a mere scholarly updated description of what the movement stood for, rather an articulated account of its theoretical pillars, based upon its different takes on the notion of a machine –namely, the latter being possibly teleological, possibly non-material and ultimately the instantiation (accordingly, materially or not) of a theory. This reading of cybernetics' structure and goals allowed me to show the subsequent internal tensions that arguably led to its demise.

Coupled with the implosion of the Macy conferences and the Ratio club, two reported experiences occurring during the mid-1940s and lasting for less than a decade, those of William Ross Ashby and John von Neumann, can serve as historical markers pointing to the time when classical cybernetics was effectively over --around the mid 1950's.

However, beyond historical clarifications, a theoretical issue of philosophical import seems to emerge from the scrutiny of both experiences. Both scientists, after struggling with physical instantiations of their advocated proposals, ended up retreating from the material anchoring of these proposals into a disembodied realm --the very realm inhabited by those positions they were set to dispel in the first place, be it psychologisms or metaphysical truisms. These *ad intra* conflicting cybernetic developments are exposed and treated in the two chapters that followed, one pertaining to von Neumann's contributions and the other to Ross Ashby's -- two cyberneticists whose influence on the enterprise has been so far relatively relegated to a secondary role, probably due in part to the widespread non-philosophical accounts of cybernetics' atrophy.

William Ross Ashby's role in the cybernetic enterprise was characterized by a brilliant mind that combined a philosophical underlayer of psychology with the utilization of serious engineering skills, and whose British empiricist heritage led him to be faithful to the materialistic tenets of cybernetics until the end. Ashby's role might have given him the image of the underlaborer of cybernetics, also caring more about cybernetic principles than what the founding cyberneticists themselves did. In that spirit, Ashby's inroads into cybernetics commenced with his preoccupation with the underdeveloped notion of a machine --notion that necessarily had to grow in sophistication if it was destined to encompass living entities, as cybernetics clearly intended. For this, he embarked in what could stand as the only exclusive treatment of what a

machine stands for, including arriving at a definition of it. Ashby did this by substantially relating on the one hand the notions of equilibrium in physical systems, and on the other the adaptation by self-organization in living entities. He then equated this relationship with *survival* itself. Having bridged the organic and the inorganic by means of a fundamental feature of life present in both –namely, adaptability--, he then built a peculiar machine to show how it can do what was hitherto believed as being contrary to its own machinal nature: adapting to its changing environment. A machine which literally survived. Ashby's homeostat, acting as an embodied heuristic device, uncomfortably cut through the not entirely metaphysics-free cybernetic minds, showing them how far the cybernetic mechanistic commitment goes, testing whether they were willing to pay the epistemological and ontological price.

Although Ashby was reasonably satisfied with the outcome of the exposition of his ideas by means of his device, he accepted criticisms, granting that the machine was embedded with too little a number of variables as it is, in its attempt to resemble a living entity. In fact he already had built another, more complicated one –DAMS. However, this last one never managed to attain the kind of self-organizing adaptive behavior of the simpler homeostat. Instead its chaotic behavior never seemed to reach, let alone settle in, a coping pattern --despite 2 years of careful observation. Ashby, discouraged, abandoned the project and focused on the theoretical exploration of the notion of “cybernetics of cybernetics” (or second order cybernetics), where the “observer becomes part of the system”, engaging in projects without physical grounding and effectively going under the radar (both scientifically and publicly) until the end of his life.

John von Neumann's part in the cybernetic movement was in turn characterized by a genius put in check by Continental skepticism. Despite his praise of a bold and exciting proposal about the nature of life and mind –cybernetics--, he reminded us of the majesty of the “objects”



of study at hand. The networks constructed by McCulloch and Pitts, a model that, as we saw, was to become a major theoretical column for cybernetics, begun to be heavily criticized by von Neumann on a number of grounds. Despite his acknowledgment of the heuristic utility of the model, he pointed out to its excessive removal from the reality it was supposed to model – by means of necessary simplification. Being able to have a “working” model of what occurs inside the brain was certainly a plausible achievement, but the quantity of essential details left aside (for the model to work) was a source of concern for him. It might be the case that the very details that would otherwise give a richer and more revealing account, were being left aside. Von Neumann in fact advanced the idea that likely the most important aspects of our neural grid were indeed being left aside. Bringing this criticism to its logical outcome, he observed that if we are to take into consideration such a model for mapping real cognitive processes, the model would turn out to be of a degree of complexity that would absolutely escape our handling capabilities. In realistic terms, von Neumann would ask, what would the use for it be, if modeling a cognitive process would take someone’s entire life time, or if the scheme rendered would occupy the entire physical universe?

After the stark criticism against the verifiability of results given by McCulloch and Pitts’ networks –one of the main pillars of cybernetics-- von Neumann proposed an epistemological turn. He encouraged shifting away from the observations of the behavior of the object in order to focus instead on its structure, given that after a certain point of complexity, the latter could be more amenable of tractability than the former. Since the ultimate structure of a complex entity is complex itself (the model approaching a one to one mapping upon the object), von Neumann recommended retargeting the cybernetic attention towards simpler organisms. In fact, the simplest: a virus. Despite the limbo area that it occupies in terms of fully counting as a living

entity, a virus does display a feature essential to living things: reproductive capacities. Precisely this capability for self-replication is what von Neumann recognized and extrapolated as worthy to focus on and --in true cybernetic fashion—literally *build*. His resulting kinematic model, however, was fraught with issues pertaining to details inherent to material embodiment --to the way in which three-dimensional entities inhabit the world. He subsequently found asylum in logical representations of lattices --sidestepping the problems of physical embodiment altogether. After jumping out of the physical grounding into the abstract realm (calling for the developing of a “new logic” more suitable for automata reproduction), the project failed to garner attention. Shortly after, von Neumann died.

Both epic missions seemingly eventuated in negative outcomes. As advanced in this thesis, the tension between the mandate for model construction and the seemingly perennial temptation for abandoning all grips with physical reality might have proven to be devastating for the cybernetic endeavor. And more so if what ultimately characterized cybernetics was its impetus for control, in order to enhance and eventually complete its overarching articulating power. In the case of von Neumann, his kinematic model proved to be fraught with a plethora of practical issues, such as weight, density, friction, heat, material fatigue, etc. Being physical is hard, someone would say. In a move to avoid the problematic entailments of the physical construction of a fairly complex entity, he resorted to leaving physical grounding aside, concentrating instead on the possible future logic that would allow such entities, his automata, to exist. In the case of Ashby, the acknowledgement of the need of augmentation of variables, for an entity to more convincingly qualify as alive and learning, was realized in the construction of DAMS. However, one could witness via the dramatic entries in his diary how Ashby grows

increasingly discouraged in the span of two years, observing a machine that only renders chaotic behavior.

If one approaches the experience of classical cybernetics from the theoretical framework of an inquiry into the nature of a machine, one might be able to shed more light into what went down with such peculiar enterprise –one that manages to carry allure until these days. The cybernetic articulation of a machine might have been ahead of its time, particularly in what regards the role of entropy in the physical world, and the relevance of materiality for a physical entity to exist –such as a machine. Martin Heidegger might have been somewhat prescient in his usually unattended preoccupation of cybernetics. He was aware that the technological mandate of purposeful manipulation, paradigmatically present in cybernetics, was purportedly pursued for the sake of the improvement of the human condition –as Wiener would proclaim in later years. However, as we saw, Heidegger recognized in cybernetics an epistemology that was going to wrap up, for him, the very history of metaphysics. Cybernetics was not only going to transcend modern science as it was practiced, bringing us one step closer to a Leibnizian dream of a universally subsumed reality under the power of calculus. Cybernetics was also destined to replace philosophy itself, by means of the ultimate mechanization of anything understandable -- precisely *because* it is understandable.

With cybernetics, an enriched understanding of what a machine is, was brought to the fore, reaching its zenith with Ross Ashby. Ashby brought to fruition Turing’s hint at a definition of a machine not fundamentally relying on its materiality, but rather, on its behavior. That much would be in sync with Norbert Wiener’s enriched understanding of a machine. However, seemingly an ontological displacement was taking place at the same time. While Wiener was talking about “teleological machines”, Ashby was interested in the machine’s capacity for *self-*

*organization* –after all, if living entities truly were machines, they should show this capability. While Wiener would recognize negative-feedback-fed machines with a purpose in both the organic and inorganic realms, Ashby would see an inherent capacity of self-organization towards equilibrium. Both could *prima facie* be seen as compatible. With one difference: Wiener saw *chaos* as the fundamental inclination of physical reality. Ashby saw *order* as the forming undercurrent of this reality. The consequences of this subterranean displacement would come later. When Ashby, true to the cybernetic spirit, attempted to show a physical model of the machine –which should consequently blur the difference between the organic living and the artificial, inert contraption—he ran into problems. The instantiation of a self-organizing machine would work only with very simple and few variables (e.g., the homeostat). Once more complexity was introduced—which is required if one is to ever mimic, let alone re-create, life—the physical instantiation of the model would no longer work. Ashby put his machines aside and dedicated himself to the disembodied study of systems, until the end of his career.

In the case of John von Neumann, the aspect of this new and richer understanding of a machine that he wanted to put emphasis on was *self-replication*. This approach emerged out of a frustration taking place after attempting to model a nervous system or a brain upon artificial neural networks. Von Neumann showed that they would certainly work for far less complex machines –such as the first program-stored computer that he built. But it would be hopeless for mapping anything substantially more complex –such as the brain. In such case, the gain of being able to rigorously describe a behavior simply disappears, since the complexity is so tremendous that the rigorous behavior description would be translated into a model of such dimensions that it would inherently escape any possibility of human tractability. In such case, it would be “simpler” to in effect turn our attention into the *structure* of the entity –a very uncybernetic move, just as

much as Ashby's stance regarding a necessary order coming out of chaos. Von Neumann's approach, still riding the cybernetic spirit, had to be instantiated in the flesh, and such attempt is remembered as the *kinematic model*. "Real world" problems precluded this model from attaining success, and in a manner that ultimately is not dissimilar from Ashby's late reaction, he dropped the physical instantiation of the model and opted for a theoretical project towards a new and enriched logic –leaving physical instantiations aside.

Once cybernetics lost its unique methodological and ontological stance –namely, knowing by means of constructing and letting this construction lead to more knowledge—not much was left that was unique to cybernetics. Considering the renewed and sophisticated understanding of what a machine is --which constituted a cybernetic cornerstone at a time-- it might come as somewhat expectable. After all, a machine does not have to be materially instantiated to be a machine, and so a physical embodiment of a machine is only tangentially important for the machine to be itself. But physical instantiations were fundamental for the cybernetic enterprise from its inception. It would seem that a profound insight eventually led to its own demise. Under this perspective, one could see as the positive outcome of the cybernetic implosion a richer and deeper understanding of the nature of a machine.

## **Bibliography.**

- Abraham, T. H. (2003). From theory to data: Representing neurons in the 1940s. *Biology and Philosophy*, 18(3), 415-426.
- Adkins, B. M. (1951). The homeostat. *British Journal for the Philosophy of Science*, 2 (7), 248
- Ampère, A. M. (1838). *Essai sur la philosophie des sciences*. Paris, France: Bachelier.
- Asaro, P. (2006). On the origins of the synthetic mind: Working models, mechanisms, and simulations. (Ph.D., University of Illinois at Urbana-Champaign).
- Asaro, P. (2008). From Mechanisms of Adaptation to Intelligence Amplifiers: The Philosophy of W. Ross Ashby. In P. Husbands, O. Holland & M. Wheeler (Eds.), *The Mechanical Mind in History*. Cambridge, MA: MIT Press, pp. 149-184.
- Asaro, P. (2010). Whatever Happened to Cybernetics? In G. Friesinger, J. Grenzfurthner, T. Ballhausen, and V. Bauer (Eds.), *Geist in der Maschine. Medien, Prozesse und Räume der Kybernetik* (pp. 39-49). Vienna, Austria: Verlag Turia & Kant.
- Ashby, W. R. (1928-1972). *Journal*. The W. Ross Ashby Digital Archive, 2008. Retrieved from [www.rossashby.info/journal](http://www.rossashby.info/journal)
- Ashby, W. R. (1940). Adaptiveness and equilibrium. *Journal of Mental Science*, 86, 478-483.
- Ashby, W. R. (1945). The physical origin of adaptation by trial and error. *Journal of General Psychology*, 32, 13-25.
- Ashby, W. R. (1947). The nervous system as physical machine: With special reference to the origin of adaptive behaviour. *Mind*, 56(1): 44-59.
- Ashby, W. R. (1947). Principles of the self-organizing dynamic system. *Journal of General Psychology*, 37(2): 125-125.

- Ashby, W. R. (1949). Experimental homeostat. *Electroencephalography & Clinical Neurophysiology*, 1, 116-117.
- Ashby, W. R. (1949b, March). The Electronic Brain. *Radio-Electronics*, 77-80.
- Ashby, W. R. (1949c, January). The Thinking Machine. *Time*, 66.
- Ashby, W. R., (1952) *Design for a Brain*. New York, NY: John Wiley & Sons.
- Ashby, W. R. (1952b). Can a mechanical chess-player outplay its designer? *British Journal for the Philosophy of Science*, 3, 44-57.
- Ashby, W. R., (1953) Homeostasis. In H. Von Foerster, M. Mead & H. L. Teuber (Eds.). *Cybernetics: Circular Causal and Feedback Mechanisms in Biological and Social Systems. Transactions of the Ninth Conference, New York, March 20-21, 1952* (pp. 73-108). New York, NY: Josiah Macy, Jr. Foundation.
- Ashby, W. R. (1956). *An Introduction to Cybernetics*. London, UK: Chapman and Hall, New York, NY: John Wiley & Sons.
- Ashby, W. R. (1960). *Design for a brain: The origin of adaptive behavior* (2<sup>nd</sup> revised ed.). New York, NY: John Wiley & Sons.
- Ashby, R. W. (1962). Principles of the self-organizing system. In H. von Foerster and G. W. Zopf, Jr. (Eds.), *Principles of self-organization: Transactions of the University of Illinois Symposium* (pp. 255-278). London, UK: Pergamon Press.
- Ashby, W. R. (1981). *Mechanisms of intelligence: Ross Ashby's writings on cybernetics* (R. Conant, Ed.). Salinas, CA: Intersystems Publications.
- Aspray, W. (1990). The origins of John von Neumann's theory of automata. In J. Glimm, J. Impagliazzo & I. Singer (Eds.). *Proceedings of Symposia in Pure Mathematics, Volume 50: The Legacy of John von Neumann*. Providence, RI: American Mathematical Society.

- Bacon, F. (1620). *The New Organon*. In: R. S. Sargent (Ed.), *Francis Bacon: Selected Philosophical Works*. Indianapolis, IN: Hackett, 1999.
- Bell, E. T. (1937) *Men of Mathematics* (2<sup>nd</sup> ed., Vol. 2). London, England: Penguin Books.
- Bennett, S. (1993). *A History of Control Engineering, 1930-1955*. London, UK: Peter Peregrinus Inc.
- Bennett, S. (1996). A Brief History of Automatic Control. *IEEE Control Systems*, 16(3), 17–25.
- Bissell, Christopher (2009). A history of automatic control. In S. Y. Nof (Ed.), *Springer handbook of automation* (pp. 53–69). Springer handbook series (LXXVI). Heidelberg, Germany: Springer Verlag.
- Boden, M. A. (2006). *Mind as machine: A history of cognitive science*. Oxford; New York: Clarendon Press; Oxford University Press.
- Bowles, M. (1996) U.S. Technological Enthusiasm and British Technological Skepticism in the Age of the Analog Brain. *IEEE Annals of the History of Computing* 18(4), 5-15.
- Bryant, J. (1991). *Systems theory and scientific philosophy: An application of the cybernetics of W. Ross Ashby to personal and social philosophy, the philosophy of mind, and the problems of artificial intelligence*. Lanham, MD: University Press of America.
- Burks, A. W. (1959). The Logic of Fixed and Growing Automata. *Proceedings of an International Symposium on the Theory of Switching, 2-5 April 1957* (part I, pp. 147-188). Cambridge, MA: Harvard University Press.
- Burks, A. W. (1960). Computation, Behavior, and Structure in Fixed and Growing Automata. In M. Yovits & S. Cameron (Eds.). *Self-Organizing Systems* (pp. 282-311, 312-314). New York, NY: Pergamon Press.



- Burks, A. W. (1963). Programming and the Theory of Automata. In P. Braffort and D. Hirschberg (Eds.). *Computer Programming and Formal Systems* (pp. 100-117). Amsterdam, Holland: North Holland Publishing Company.
- Burks, A. W. (1963b). Toward a Theory of Automata Based on More Realistic Primitive Elements. In C. M. Popplewell (Ed.). *Information Processing, Proceedings of IFIP Congress 62* (pp. 379-385). Amsterdam, Holland: North Holland Publishing Company.
- Burks, A. W. (1966). Automata Models of Self-Reproduction. *Information Processing* 7 (May, 1966): 121-123. (Information Processing Society of Japan)
- Burks, A. W. (Ed.). (1970). *Essays on Cellular Automata*. Urbana, IL: University of Illinois Press.
- Burks, A. W. (1974). Cellular Automata and Natural Systems. In W. D. Keidel, W. Handler & M. Spreng (Eds.). *Cybernetics and Bionics (Proceedings of the 5th Congress of the Deutsche Gesellschaft für Kybernetik, Nürnberg, March 28-30, 1973)*, (pp. 190-204). Munich, Germany: R. Oldenbourg.
- Burks, A. W. (1975). Logic, Biology, and Automata - Some Historical Reflections. *International Journal of Man-Machine Studies* 7: 297 -312.
- Burks, A. W & Wang, H. (1955). The Logic of Automata. *Journal of the Association for Computing Machinery* 4 (April, 1955):193-218 and 4 (July, 1955): 279-297.
- Chapuis, A & Gélis, E. (1928). *Le Monde des Automates: Étude Historique et Technique* (Vols. 1-2). Genève, Switzerland: Editions Slatkine.
- Choi B.C. & Pak A.W. (2006). Multidisciplinarity, interdisciplinarity and transdisciplinarity in health research, services, education and policy: 1. Definitions, objectives, and evidence of effectiveness. *Clinical Investigative Medicine*, 29(6), 351-64.

- Christensen, W. (1996). A complex systems theory of teleology. *Biology and Philosophy*, 11(3), 301-320.
- Conway, F. & Siegelman, J. (2005). *Dark hero of the information age: In search of Norbert Wiener, the father of cybernetics*. New York, NY: Basic Books.
- Copeland, J. (Ed.). (2004). *The Essential Turing: Seminal Writings in Computing, Logic, Philosophy, Artificial Intelligence, and Artificial Life plus The Secrets of Enigma*. New York, NY: Oxford University Press.
- Copeland, J. (2005) Introduction. In: J. Copeland et al., *Alan Turing's Electronic Brain: The Struggle to Build the ACE, the World's Fastest Computer* (pp. 1-14). Oxford, UK: Oxford University Press.
- Copeland, J. (2012). *Turing: Pioneer of the Information Age*. Oxford, UK: Oxford University Press.
- Copeland, J. & Proudfoot D. (2005) Turing and the computer. In: J. Copeland et al., *Alan Turing's Electronic Brain: The Struggle to Build the ACE, the World's Fastest Computer* (pp. 107-148). Oxford, UK: Oxford University Press.
- Cordeschi, R. (2002). *The discovery of the artificial: Behavior, mind, and machines before and beyond cybernetics*. Dordrecht; Boston: Kluwer.
- Cordeschi, R. (2008a). Steps Toward the Synthetic Method: Symbolic Information Processing and Self-Organizing Systems in Early Artificial Intelligence Modeling. In P. Husbands, O. Holland & M. Wheeler (Eds.), *The mechanical mind in history* (pp. 219-258). Cambridge, MA: MIT Press.
- Cordeschi, R. (2008b). Cybernetics. In L. Floridi (Ed.), *The Blackwell guide to the philosophy of computing and information* (pp. 186-196). Oxford, UK: Blackwell.

- Davis, M. (2000). *The universal computer: The road from Leibniz to Turing*. New York, NY: W. W. Norton & Company Inc.
- De Latil, P. (1956). *Thinking by Machine: A Study of Cybernetics*. Boston, MA: Houghton Mifflin.
- Dennett, D. (1993). Review of *The Embodied Mind*, by F. Varela, E. Thompson and E. Rosch. *American Journal of Psychology*, 106, 121-126.
- Descartes, R. (1662). *Treatise on Man* (S. Gaukroger, Trans.) In: S. Gaukroger, (Trans., Ed.). *Rene Descartes: The World and Other Writings* (pp. 99-169). Cambridge, UK: Cambridge University Press, 1998.
- Descartes, R. (1649). Reply to More (February) (M. Grene & R. Ariew, Trans.). In: R. Ariew (Trans., Ed.). *Rene Descartes: Philosophical Essays and Correspondence* (pp. 292-297). Indianapolis, IN: Hackett, 2000.
- Drexler, K. E. (1986). *Engines of Creation: The Coming Era of Nanotechnology*. New York, NY: Anchor Books.
- Dupuy, J. P. (2000). *The mechanization of the mind: On the origins of cognitive science*. Princeton, N.J: Princeton University Press.
- Dupuy, J. (2007). Some pitfalls in the philosophical foundations of nanoethics. *The Journal of Medicine and Philosophy*, 32(3), 237.
- Easterling, K. (2001). Walter Pitts. *Cabinet*, 5(02)
- École polytechnique (France). Centre de recherche épistémologie et autonomie. (1985). *Généalogies de l'auto-organisation*. Paris: "Cahiers du C.R.E.A".

- Edwards, P. N. (1996). *The Machine in the Middle: Cybernetic Psychology and World War II*. In *The closed world: Computers and the politics of discourse in cold war America* (pp. 174-207). Cambridge, MA: MIT Press.
- Eisenstein, E. (2012). *The Printing Revolution in Early Modern Europe* (2<sup>nd</sup> ed.). New York, NY: Cambridge University Press.
- Ewald, W. B. (Ed.) (1996). *From Kant to Hilbert: A Source Book in the Foundations of Mathematics* (Vol. 2). New York, NY: Oxford University Press.
- Findler, N. V. (1988). The debt of artificial-intelligence to John von Neumann. *Artificial Intelligence Review*, 2(4), 311-312.
- Fisher, R. A. & Yates, F. (1963) *Statistical Tables for Biological, Agricultural and Medical Research*. London, UK: Oliver and Boyd. (Original work published in 1938)
- Floridi, L. (2004). *The Blackwell guide to the philosophy of computing and information*. Malden, MA: Blackwell Pub.
- Franchi, S., & Güzeldere, G. (2005). *Mechanical bodies, computational minds: Artificial intelligence from automata to cyborgs*. Cambridge, MA: MIT Press.
- François, D. (2007) *The Self-Destruction of the West*. Paris, France: Éditions Publibook
- Frege, G. (1884). *The Foundations of Arithmetic: A Logico-Mathematical Enquiry Into the Concept of Number* (J. L. Austin, Trans.). Evanston, IL: Northwestern University Press, 1980.
- Freitas, R. & Gilbreath, W. (Eds.). (1980). *Proceedings of the 1980 NASA/ASEE Summer Study: Advanced Automation for Space Missions*. Washington, DC: US Government Printing Office.

- Galison, P. (1994). The ontology of the enemy: Norbert Wiener and the cybernetic vision. *Critical Inquiry*, 21(1), 228-266.
- Gaukroger, S. (1995) *Descartes: An Intellectual Biography*. Oxford, UK: Oxford University Press.
- Gerovitch, S. (2002). *From Newspeak to Cyberspeak: A History of Soviet Cybernetics*. Cambridge, MA: MIT Press.
- Glimm, J., Impagliazzo, J., & Singer, I. (Eds.). (1990). Proceedings of Symposia in Pure Mathematics, Volume 50: *The Legacy of John von Neumann*. Providence, RI: American Mathematical Society.
- Gödel, K. (1931). On Formally Undecidable Propositions of Principia Mathematica and Related Systems I (E. Mendelson, Trans.). In M. Davis (Ed.) (1965) *The Undecidable: Basic Papers on Undecidable Propositions, Unsolvability Problems and Computable Functions* (pp. 4-38). Hewlett, NY: Raven.
- Gödel, K. (1934). On Undecidable Propositions of Formal Mathematical Systems. In M. Davis (Ed.) (1965) *The Undecidable: Basic Papers on Undecidable Propositions, Unsolvability Problems and Computable Functions* (pp. 39-74). Hewlett, NY: Raven.
- Goldstine, H. H. (1993). *The Computer from Pascal to von Neumann*. Princeton, NJ: Princeton University Press. (Original work published 1972)
- Goujon, P. (1994). From logic to self-organization - the learning of complexity. *Cybernetica*, 37(3-4), 251-272.
- Haldane, J. S. (1922) *Respiration*. New Haven, CT: Yale University Press.
- Hawking, S. W. (Ed.). (2005) *God Created the Integers: The Mathematical Breakthroughs that Changed History*. Philadelphia, PA: Running Press.

- Hayles, N. K. (1999). *How we became posthuman: Virtual bodies in cybernetics, literature, and informatics*. Chicago, IL: University of Chicago Press.
- Hayward, R. (2001). The tortoise and the love-machine: Grey Walter and the politics of electroencephalography. *Science in Context* 14(4), 615-641.
- Heidegger, M. (1977) The question concerning technology. In W. Lovitt (Trans.), *The Question Concerning Technology and Other Essays* (pp. 3-35). New York, NY: Harper & Row. (Original work published 1954)
- Heidegger, M. (1977b). The end of philosophy and the task of thinking. In D. F. Krell (Ed.), *Basic Writings. From Being and Time (1927) to The Task of Thinking (1964)* (pp. 373-392). New York, NY: Harper & Row. (Original work published 1969)
- Heidegger, M. (1981). "Only a God Can Save Us": The Spiegel Interview. In T. Sheehan (Ed.), *Heidegger: The Man and the Thinker* (pp. 45-67) Chicago, IL: Precedent. (Original work published 1966)
- Heidegger, M. (1984). *The Metaphysical Foundations of Logic* (M. Heim, Trans.). Bloomington, IN: Indiana University Press. (Original work published 1978)
- Heidegger, M. (1999) *Contributions to Philosophy (From Enowning)*. (P. Emad & K. Maly, Trans.). Bloomington, IN: Indiana University Press. (Original work published 1989)
- Heidegger, M. & Fink, E. (1993) *Heraclitus Seminar*. Evanston, IL: Northwestern University Press. (Original work published 1970)
- Heims, S. J. (1980). *John von Neumann and Norbert Wiener: From mathematics to the technologies of life and death*. Cambridge, MA: MIT Press.
- Heims, S. J. (1991). *The cybernetics group*. Cambridge, MA: MIT Press.

- Hertz, H. (1899) *The Principles of Mechanics Presented in a New Form*. (D. E. Jones & J. T. Walley, Trans., P. Lenard, Ed.). London, UK: Macmillan. (Original work published in 1894)
- Heylighen, F. (2000). Foundations and methodology for an evolutionary world view: A review of the principia cybernetica project. *Foundations of Science*, 5(4), 457-490.
- Hilbert, D. (1902). Mathematical Problems (M. W. Newson, Trans.). (Original work published in 1900). *Bulletin of the American Mathematical Society*, 8(10), 437-479
- Hodges, A. (1983). *Alan Turing: The Enigma of Intelligence*. London, UK: Unwin Publishing Company.
- Hodges, A. (2000) Turing. In: R. Monk & F. Raphael (Eds.), *The Great Philosophers: From Socrates to Turing* (pp. 493-540). London, U.K.: Phoenix.
- Hodges, A. (2008). What Did Alan Turing Mean by “Machine”? In P. Husbands, O. Holland & M. Wheeler (Eds.), *The mechanical mind in history* (pp. 75-90). Cambridge, MA: MIT Press.
- Holland, O. & Husbands, P. (2011) The origins of British cybernetics: the Ratio Club. *Kybernetes*, 40(1/2), 110-123.
- Huchingson, J. E. (2005). *Religion and the Natural Sciences: The Range of Engagement*. Eugene, OR: Wipf and Stock. (Original work published 1993)
- Husbands, P. & Holland, O. (2008). The Ratio Club: A Hub of British Cybernetics. In P. Husbands, O. Holland & M. Wheeler (Eds.), *The mechanical mind in history* (pp. 91-148). Cambridge, MA: MIT Press.
- Husbands, P. & Holland, O. (2012). Warren McCulloch and the British Cyberneticians. *Interdisciplinary Science Reviews*, 37(3), 237–253.

- Husbands, P., Holland, O., & Wheeler, M. (Eds.). (2008). *The mechanical mind in history*. Cambridge, MA: MIT Press.
- Husky, H. D. (2005). The ACE Test Assembly, the Pilot ACE, the Big ACE, and the Bendix G15. In J. Copeland et al., *Alan Turing's Electronic Brain: The Struggle to Build the ACE, the World's Fastest Computer* (pp. 281-296). Oxford, UK: Oxford University Press.
- Hutchins, W. J. (2000). Warren Weaver and the launching of MT: brief biographical note. In W. J. Hutchins (Ed.). *Early years in machine translation: Memoirs and Biographies of Pioneers*. (pp. 17-20). Amsterdam, Netherlands: John Benjamins.
- Hyötyniemi, H. (2006, August). *Neocybernetics in Biological Systems*. Espoo, Finland: Helsinki University of Technology, Control Engineering Laboratory. Retrieved from [www.control.hut.fi/rpt/r151isbn9789512286133.pdf](http://www.control.hut.fi/rpt/r151isbn9789512286133.pdf)
- Jackson, B. (1972). The Greatest Mathematician in the World: Norbert Wiener Stories. *Western Folklore*, 31(1), 1-22.
- Jeffress, L. A. (Ed.). (1951). *Cerebral mechanisms in behavior: The Hixon symposium*. New York, NY: John Wiley & Sons.
- Jennings, H. S. (1906). *Behavior of the Lower Organisms*. New York, NY: Columbia University Press.
- Johnston, J. (2008). *The allure of machinic life: Cybernetics, artificial life, and the new AI*. Cambridge, MA: MIT Press.
- Kant, I. (1787). *Critique of Pure Reason*. (P. Guyer & A. W. Wood, Trans. & Eds.). New York, NY: Cambridge University Press, 1998.



- Kay, L. E. (2000). *Who Wrote the Book of Life?: A History of the Genetic Code*. Stanford, CA: Stanford University Press.
- Kleene, S. C. (1967). *Mathematical logic*. New York, NY: John Wiley & Sons.
- Langton, C. (1989). Artificial Life. In C. Langton (Ed.), *Artificial Life* (pp. 1-44). Reading, MA: Addison-Wesley Publishing Company.
- Lassègue, J. (1996) What Kind of Turing Test Did Turing Have in Mind? *Tekhnema* 3, 37-58
- MacColl, L. A. (1945) *Fundamental theory of servomechanisms*. New York, NY: D. Van Nostrand.
- MacDonald, I. (2008). *Adorno and Heidegger: Philosophical Questions*. Stanford, CA: Stanford University Press.
- Malapi-Nelson, A. (2014). Humanities' Metaphysical Underpinnings of Late Frontier Scientific Research. *Humanities* 3(4), 740-765.
- Mancosu, P. (Ed.). (1998) *From Brouwer to Hilbert: The Debate on the Foundations of Mathematics in the 1920s*. New York, NY: Oxford University Press.
- Marciszewski, W., & Murawski, R. (1995). *Mechanization of reasoning in a historical perspective*. Amsterdam, Netherlands; Atlanta, GA: Rodopi.
- Marcovich, M. (1967). *Heraclitus. Greek text with a short commentary. Editio maior*. Merida, Venezuela: Universidad de los Andes Press.
- Margolis, E. & Laurence, S. (Eds.). (2007). *Creations of the Mind: Theories of Artifacts and Their Representation*. Oxford, UK: Oxford University Press.
- Marsalli, M. (2006). *McCulloch-Pitts Neurons* (National Science Foundation Grants #9981217 and #0127561). Retrieved from [www.mind.ilstu.edu/curriculum/modOverview.php?modGUI=212](http://www.mind.ilstu.edu/curriculum/modOverview.php?modGUI=212)

- Masani, P. R. (1990). *Norbert Wiener, 1894–1964*. Basel, Switzerland: Birkhäuser.
- Masani, P. R. (1994). The scientific methodology in the light of cybernetics. *Kybernetes*, 23(4), 5-132.
- Maxwell, J. C. (1868). On Governors. *Proceedings of the Royal Society of London*, 16, 270-283.
- Mayr, O. (1970). *The Origins of Feedback Control*. Cambridge, MA: MIT Press.
- McCarthy, J., Minsky, M. L., & Shannon, C. E. (1955). A proposal for the Dartmouth summer research project on artificial intelligence - August 31, 1955. In R. Chrisley & S. Begeer (Eds.). (2000). *Artificial intelligence: Critical concepts* (Vol. 2, pp. 44-53). London, UK & New York, NY: Routledge.
- McCorduck, P. (2004). *Machines Who Think: A Personal Inquiry into the History and Prospects of Artificial Intelligence*. Natick, MA: A K Peters.
- McCulloch, W. S. (1947). *An Account of the First Three Conferences on Teleological Mechanisms*. New York, NY: Josiah Macy, Jr. Foundation.
- McCulloch, W. S. (1951). Why the Mind is in the Head [and Discussion]. In L. A. Jeffress (Ed.), *Cerebral Mechanisms in Behavior: The Hixon Symposium* (pp. 42-74). New York, NY: John Wiley & Sons.
- McCulloch, W. S. (1960). *General Semantics Bulletin*, 26/27, 7-18.
- McCulloch, W. S. (1965). *Embodiments of mind*. Cambridge, MA: MIT Press.
- McCulloch, W. S. (1974). Recollections of the Many Sources of Cybernetics. *ASC Forum*, VI(2), 5-16.
- McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5, 115-133.

- McCulloch, W. S. & Pitts, W. (1948). The statistical organization of nervous activity. *Biometrics*, 4, 91-99.
- Macrae, N. (1992). *John von Neumann: The Scientific Genius who Pioneered the Modern Computer, Game Theory, Nuclear Deterrence, and Much More*. Providence, RI: American Mathematical Society.
- Medina, E. (2006). Designing freedom, regulating a nation: Socialist cybernetics in Allende's Chile. *Journal of Latin American Studies*, 38(3), 571-606.
- Medina, E. (2011). *Cybernetic Revolutionaries: Technology and Politics in Allende's Chile*. Cambridge, MA: MIT Press.
- Mindell, D. (1995). Automation's Finest Hour: Bell Laboratories' Control Systems in World War II. *IEEE Control Systems* 15(6), 72-80.
- Mindell, D. (2002). *Between Human and Machine: Feedback, Control, and Computing before Cybernetics*. Baltimore, MD: Johns Hopkins University Press.
- Mindell, D., Segal, J., Gerovitch, S. (2003). Cybernetics and Information Theory in the United States, France and the Soviet Union. In M. Walker (Ed.), *Science and Ideology: A Comparative History* (pp. 66-95). London, UK: Routledge.
- Minsky M. & Papert S. (1969). *Perceptrons*. Cambridge, MA: MIT Press.
- Mirowski, P. (2002). *Machine Dreams: Economics Becomes a Cyborg Science*. Cambridge, UK: Cambridge University Press.
- Moschovakis, Y. N. (2001). What is an algorithm? In B. Engquist, & W. Schmid (Eds.), *Mathematics Unlimited — 2001 and beyond* (pp. 919-936). Berlin, Germany: Springer.
- Müller A. & Müller K. (Eds.). (2007) *An Unfinished Revolution? Heinz von Foerster and the Biological Computer Laboratory (BCL), 1958–1976*. Vienna, Austria: Edition Echoraum.

- Newton, I. (1687). *The Principia: Mathematical Principles of Natural Philosophy*. (B. Cohen, A. Whitman, & J. Budenz, Trans.). Oakland, CA: University of California Press, 1999.
- Owens, L. (1986). Vannevar Bush and the Differential Analyzer: The Text and Context of an Early Computer. *Technology and Culture*, 27(1), 63-95.
- Piccinini, G. (2002). Review of *The Mechanization of Mind: On the Origins of Cognitive Science*, by J. P. Dupuy. *Minds and Machines* 12(3), 449-453.
- Piccinini, G. (2004). The first computational theory of mind and brain: A close look at McCulloch and Pitts' "Logical calculus of ideas immanent in nervous activity". *Synthese*, 141(2), 175-215.
- Pickering, A. (2002). Cybernetics and the mangle: Ashby, Beer and Pask. *Social Studies of Science*, 32(3), 413-437.
- Pickering, A. (2005). A Gallery of Monsters: Cybernetics and Self-Organisation, 1940-1970. In S. Franchi, G. Güzeldere (Eds.), *Mechanical Bodies, Computational Minds: Artificial Intelligence from Automata to Cyborgs* (pp. 229-245). Cambridge, MA: MIT Press.
- Pickering, A. (2010). *The Cybernetic Brain: Sketches of Another Future*. Chicago, IL: University of Chicago Press.
- Polanyi, M. (1952). The hypothesis of cybernetics. *British Journal for the Philosophy of Science*, 2, 312-314.
- Prigogine, I. & Stengers, I. (1984) *Order Out of Chaos: Man's New Dialogue with Nature*. New York, NY: Bantam Books.
- Poundstone, W. (1985). *The recursive universe: Cosmic complexity and the limits of scientific knowledge* (1st ed.). New York, NY: William Morrow & Co.

- Read, D. (2009) *Artifact Classification: A Conceptual And Methodological Approach*. Walnut Creek, CA: Left Coast Press.
- Rédei, M. (2005). Introductory Comments. In J. von Neumann (2005), *John von Neumann: Selected letters*. M. Rédei (Ed.) (pp. 1-40). Providence, R.I: American Mathematical Society.
- Rockefeller Foundation (1933). *Annual Report*. New York, NY: W. Weaver.
- Rolston III, H. (2006). *Science and Religion: A Critical Survey*. West Conshohocken, PA: Templeton Foundation. (Original work published 1987)
- Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6), 386-408.
- Rosenblueth, A., Wiener, N., & Bigelow, J. (1943). Behavior, purpose and teleology. *Philosophy of Science*, 10, 18-24.
- Rosenblueth, A., & Wiener, N. (1945). The role of models in science. *Philosophy of Science*, 12, 316-321.
- Rosenblueth, A., & Wiener, N. (1950). Purposeful and non-purposeful behavior. *Philosophy of Science*, 17, 318-318.
- Schlatter, M., & Aizawa, K. (2008). Walter Pitts and "A logical calculus". *Synthese*, 162(2), 235-250.
- Schrodinger, E. (2012). *What is Life?: With Mind and Matter and Autobiographical Sketches*. Cambridge, UK: Cambridge University Press. (Original work published 1944)
- Scotus, D. (1300). *Philosophical Writings: A Selection*. Indianapolis, IN: Hackett, 1987.
- Shanker, S. G. (1987). *The decline and fall of the mechanist metaphor*. New York: St. Martin's.

- Shanker, S. G. (1987a). *Wittgenstein and the turning-point in the philosophy of mathematics*. Albany, NY: SUNY Press.
- Shanker, S. G. (1987b). Wittgenstein versus Turing on the nature of Church's thesis. *Notre Dame Journal of Formal Logic*, 28, 615-649.
- Shanker, S. G. (1988). The dawning of (machine) intelligence. *Philosophica (Belgium)*, 42, 93-144.
- Shanker, S. G. (1995). Turing and the origins of AI. *Philosophia Mathematica*, 3(1), 52-85.
- Shannon, C. E. (1948). A Mathematical Theory of Communication. *The Bell System Technical Journal*, 27, 379-423.
- Shannon, C. E. & Weaver, W. (1949). *The Mathematical Theory of Communication*. Champaign, IL: University of Illinois Press.
- Shannon, C. E. (1953). Computers and automata. *Proceedings of the IRE*, 41(10), 1234-1241.
- Shannon, C. E. (1958). Von Neumann's contributions to automata theory. *Bulletin of the American Mathematical Society* 64(3), 123-129.
- Shannon, C. E., McCarthy, J., & Ashby, W. R. (1956). *Automata studies*. Princeton, NJ: Princeton University Press.
- Shaw, R.E. & McIntyre, M. (1974). Algoristic foundations to cognitive psychology. In W.B. Weimer & D.S. Palermo (Eds.) *Cognition and the Symbolic Processes* (pp. 305-362). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Sholl, D. A. (1956) *The organization of the cerebral cortex*. New York, NY: John Wiley & Sons.
- Simon, H. A. (1954). The axiomatization of classical mechanics. *Philosophy of Science*, 21, 340-343.
- Simon, H. A. (1981). *The sciences of the artificial* (2nd ed.). Cambridge, MA: MIT Press.

- Smalheiser, N. R. (2000). Walter Pitts. In *Perspectives in Biology and Medicine*, 43 (2), 217-226.
- Smith, B. C. (2002) The Foundations of Computing. In: M. Scheutz (Ed.), *Computationalism: New Directions*. Cambridge, MA: MIT Press.
- Smith, P. (2007). *An Introduction to Gödel's Theorems*. Cambridge, UK: Cambridge University Press.
- Spencer, C. (2009, 11 September). Profile: Alan Turing. *BBC News* (available at [news.bbc.co.uk/1/hi/uk/8250592.stm](http://news.bbc.co.uk/1/hi/uk/8250592.stm))
- Stoll, R. R. (1979). *Set Theory and Logic*. Mineola, NY: Dover.
- Tamburrini, G., & Datteri, E. (2005). Machine experiments and theoretical modelling: From cybernetic methodology to neuro-robotics. *Minds and Machines*, 15(3-4), 335-335.
- Taylor, R. (1950). Comments on a mechanistic conception of purposefulness. *Philosophy of Science*, 17(4), 310-317.
- Taylor, R. (1950b). Purposeful and non-purposeful behavior: A rejoinder. *Philosophy of Science*, 17(4), 327-332.
- Tedre, M. (2006). *The Development of Computer Science: A Sociocultural Perspective* (Unpublished doctoral dissertation). University of Joensuu, Finland.
- Turing, A. M. (1937). On Computable Numbers, with an Application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, 42(1): 230-265.
- Turing, A. M. (1938). On Computable Numbers, with an Application to the Entscheidungsproblem - A Correction. *Proceedings of the London Mathematical Society*, 43(1): 544-546.

- Turing, A. M. (1939). Systems of Logic Based on Ordinals. *Proceedings of the London Mathematical Society*, 45(1), 161–228.
- Turing, A. M. (1946). Proposed electronic calculator. [paper, 48 sh. in envelope]. *The Papers of Alan Mathison Turing* (ref.: DSIR 10/385). King's College Archive Centre, Cambridge, U.K. (Xerox TS available at turingarchive.org; ref.: AMT/C/32)
- Turing, A. M. (1947) Lecture to the London Mathematical Society on 20 February 1947. In: S. B. Cooper & J. van Leeuwen (Eds.) (2013) *Alan Turing: His Work and Impact* (pp. 486-497). Amsterdam, Netherlands: Elsevier.
- Turing, A. M. (1948) Intelligent Machinery. In: J. Copeland (Ed.). (2004). *The Essential Turing: Seminal Writings in Computing, Logic, Philosophy, Artificial Intelligence, and Artificial Life plus The Secrets of Enigma* (pp. 395-432). New York, NY: Oxford University Press.
- Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59, 433-460.
- Turing, A. M. (1996). Intelligent Machinery, A Heretical Theory. *Philosophia Mathematica*, 3(4), 256-260.
- United States Patent and Trademark Office (2014, March) *Manual of Patent Examining Procedure*. Alexandria, VA: U.S. Government Printing Office.
- United States Office of Naval Research (1962). Transactions of the University of Illinois Symposium on Self-Organization: *Principles of self-organization* (H. Von Foerster & G. W. Zopf, Eds.). New York, NY: Pergamon Press.
- Varela, F., Thompson, E., & Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.
- Vico, G (1710). *On the Most Ancient Wisdom of the Italians: Unearthed from the Origins of the Latin Language*. Ithaca, NY: Cornell University Press, 1988.



- Von Bertalanffy, L. (1968). *General System theory: Foundations, Development, Applications*.  
New York, NY: George Braziller.
- Von Foerster, H., Mead M., & Teuber H. L. (Eds.). (1950). *Cybernetics: Circular Causal and Feedback Mechanisms in Biological and Social Systems. Transactions of the Sixth Conference, New York, March 24-25, 1949*. New York, NY: Josiah Macy, Jr. Foundation.
- Von Foerster, H., Mead M., & Teuber H. L. (Eds.). (1951). *Cybernetics: Circular Causal and Feedback Mechanisms in Biological and Social Systems. Transactions of the Seventh Conference, New York, March 23-24, 1950*. New York, NY: Josiah Macy, Jr. Foundation.
- Von Foerster, H., Mead M., & Teuber H. L. (Eds.). (1952). *Cybernetics: Circular Causal and Feedback Mechanisms in Biological and Social Systems. Transactions of the Eighth Conference, New York, March 15-16 1951*. New York, NY: Josiah Macy, Jr. Foundation.
- Von Foerster, H., Mead M., & Teuber H. L. (Eds.). (1953). *Cybernetics: Circular Causal and Feedback Mechanisms in Biological and Social Systems. Transactions of the Ninth Conference, New York, March 20-21, 1952*. New York, NY: Josiah Macy, Jr. Foundation.
- Von Foerster, H., Mead M., & Teuber H. L. (Eds.). (1955). *Cybernetics: Circular Causal and Feedback Mechanisms in Biological and Social Systems. Transactions of the Tenth Conference, Princeton, April 22, 23 and 24, 1953*. New York, NY: Josiah Macy, Jr. Foundation.
- Von Foerster, H. (2003). *Understanding understanding: Essays on cybernetics and cognition*.  
New York: Springer.
- Von Neumann, J. (1945). *First Draft of a Report on the EDVAC* (Contract No. W-670-ORD-4926, Between the United States Army Ordinance Department and the University of Pennsylvania). Philadelphia, PA: Moore School of Electrical Engineering.

- Von Neumann, J. (1951). The general and logical theory of automata. In L. A. Jeffress (Ed.), *Cerebral Mechanisms in Behavior: The Hixon Symposium* (pp. 1-41). New York, NY: John Wiley & Sons.
- Von Neumann, J. (1961). *John von Neumann. Collected works* (A. H. Taub, Ed.). New York, NY: Pergamon Press.
- Von Neumann, J., Aspray, W., & Burks, A. W. (Eds.). (1987). *Papers of John von Neumann on computing and computer theory*. Cambridge, MA; Los Angeles, CA: MIT Press; Tomash Publishers.
- Von Neumann, J. (1966). *Theory of Self-Reproducing Automata* (A. W. Burks, Ed.). Urbana, IL: University of Illinois Press.
- Von Neumann, J. (2005). Letter to Wiener. In M. Rédei (Ed.). *John von Neumann: Selected letters* (pp. 277-282). Providence, R.I: American Mathematical Society.
- Weaver, W. (1949). The Mathematics of Communication. *Scientific American* 181(1)
- Wheeler, M. (2008). God's Machines: Descartes on the Mechanization of Mind. In: P. Husbands, O. Holland & M. Wheeler (Eds.), *The mechanical mind in history* (pp. 307-330). Cambridge, MA: MIT Press.
- Whitehead, A. N., & Russell, B. (1927). *Principia Mathematica, 3 Volume Set* (2<sup>nd</sup> ed.). Cambridge, UK: Cambridge University Press.
- Wiener, N. (1948). *Cybernetics: Or Control and Communication in the Animal and the Machine*. Paris, France: Hermann et Cie.; Cambridge, MA: MIT Press. (2<sup>nd</sup> rev. ed. 1961).
- Wiener, N. (1948b). Time, communication and the nervous system. In: *Annals of the New York Academy of Sciences* 50, 197–220.

- Wiener, N. (1949). *The Extrapolation, Interpolation and Smoothing of Stationary Time Series: With Engineering Applications*. Cambridge, MA: M.I.T. Press.
- Wiener, N. (1950) *The Human Use of Human Being: Cybernetics and Society*. Boston, MA: Houghton Mifflin Co. (2<sup>nd</sup> rev. ed. 1954)
- Wiener, N. (1956). *I Am Mathematician: The later life of a prodigy*. New York, NY: Doubleday.
- Wiener, N. (1964). *God and Golem, Inc: A Comment on Certain Points where Cybernetics Impinges on Religion*. Cambridge, MA: M.I.T. Press.
- Wisdom, J. O. (1951). The hypothesis of cybernetics. *British Journal for the Philosophy of Science*, 2, 1-24.
- Wittgenstein, L. (1980) *Remarks on the Philosophy of Psychology (Vol. 1)*. (G. E. M. Anscombe, Trans., Ed. & G. H. von Wright, Ed.). Chicago, IL: University of Chicago Press.