Abstract

This is where the abstract will go. I guess I'll mention a thing or two about the contents, what I plan to discuss and what my analysis shows.

Cognitive Science and Artificial Intelligence: An Interwoven Approach

Supervisor: Francesco Bianchini

Paolo Marzolo

June 24, 2021

1 Introduction

This is Paolo Marzolo's bachelor thesis, written as part of the three-year program in computer science at University of Bologna. The stated objective of this document is to analyze the history of Cognitive Science and Artificial Intelligence and identify how influences between the two disciplines and others led to a partially shared evolution in the overarching research topics throughout their lifespans. Other similarities will be pointed out, and some of the algorithms and concepts contained throughout the sections will be explained in detail, in order to give the reader a more complete understanding.

The structure of the document will be as follows: after this introduction, a brief glossary will introduce some of the terms that will be used in this document with a short definition; this has been included to avoid having "foundational" terms be constrained by a specific philosophy or line of research. Then, the rest of the document will develop parallel to the history of the disciplines. In the final section, a bird's-eye-view will provide additional insight, and a brief discussion of the roles of symbols will conclude the contents. **TODO: MAYBE the discussion on symbols**

2 Terms and Definitions

Before defining our glossary, it is important to understand the reasoning behind why we chose to include it: when discussing researchers' understanding of human thought, it is nearly impossible to avoid using terms that have a strong past history. As an example, simply the mention of "thought" could already be considered too far from a behaviorist point of view. A further example is a recent discussion that took place after a somewhat controversial paper by Nunez was published [39], questioning the multidisciplinarity of Cognitive Science as a discipline (and journal) and declaring "The prospect launched by the cognitive revolution of a unified and coherent interdisciplinary seamless cognitive science did not materialize"; this questioned the existence of the field of study itself: we will discuss our usage of the term in the first entry of this glossary.

Cognitive Science. As we will see in following sections, saying "definitions of Cognitive Science have evolved throughout the years" would be a massive understatement. Nonetheless, its multidisciplinary nature is clear in what the International Encyclopedia of Social & Behavioral Sciences [19] reports 'may have been the first published use of the term cognitive science':

> The concerted efforts of a number of people from ... linguistics, artificial intelligence, and psychology may be creating a new field: cognitive science

. At the same time, even the "essential original features" identified by Gardner in 1987 [14] (summarized here as (1) necessity to speak about mental representation as a separate layer of analysis from the biological, (2) faith that the computer is central to the understanding of the human mind and (3) de-emphasizing

3

factors such as emotions or cultural factors) would be completely or partially thrown out by contemporary scholars.

In a more recent publication[5], Cognitive Science is characterized as

> ... better defined as the study of 'mind as machine' ... More precisely, cognitive science is the interdisciplinary study of mind, informed by theoretical concepts drawn from computer science and control theory.

.

Not only was its definition cloudy and unstable ("cognitive science is ... a perspective, rather than a discipline in any conventional sense" [48]), but as Nunez points out, its disciplines have varied wildly, and how represented they are in the Cognitive Science enterprise shifted throughout the years as well. Because of the reasons outlined here, far removed from the subject of this document, we will avoid using the term "Cognitive Science", and prefer the acronym "DCS".

Descriptive Cognitive Sciences (DCS). As we mentioned, the disciplines which make up Cognitive Science are not only multiple, but subject to interpretation as well. Since the objective of this work is to compare it to the history of Artificial Intelligence, we will from this point on use the acronym "DCS", for Descriptive Cognitive Sciences, as an alternate approach to the Constructive one taken by Artificial Intelligence researchers. This is not to say that a psychologist cannot take a constructive approach to the explanation of consciousness: the only reason we chose this is because we found it to be an intuitive use of the term.

Mind. Once again, although we take notice of the history of the term, we have to select a few terms to use in our language. Hereafter, we consider the mind as the non-physical correlate of human brains: "the complex of faculties involved in perceiving, remembering, considering, evaluating, and deciding. Mind is in some sense reflected in such occurrences as sensations, perceptions, emotions, memory, desires, various types of reasoning, motives, choices, traits of personality, and the unconscious."[32].

whatever else will come up

3   A History of Influences

As mentioned in the introduction, our approach will follow the historical sequence of events, although some references or explanations may be anachronistic for clarity. In order to give a general view, we split the histories of these disciplines into broad periods: one for (more or less) every substantial shift in approach and views. Generally, every time period will mention two sides of the story: one of them will focus on DCS, and the other on AI and Computer Science. At the same time, the two fields are in close relation: because of this, it becomes hard to neatly define the two fields, so some paragraphs will be somewhere in the middle (as they should be!).

4

3.1 Landscape before 1950

Although the official birth of the "Cognitive Science" institutions is in the late 1970s, reasoning about thought has been a staple in philosophical research for centuries. Because of the scope of this document, we will focus on a few important concepts, and use them to set the stage for the first large shift of ideas.

3.1.1 Mathematics and Computer Science

Some of the most relevant contributions to the "reasoning as computation" line of research come from Mathematics and what would later become Theoretical Computer Science. We will outline some of them here, while we trace part of the history of conceiving of thought as computation, first, and computers as devices for computation, second. In this respect, the following step is to be expected: can we use devices for the computation that thoughts "work" with?

Boole's Laws of Thought and Boolean Algebra. To avoid going too deep in mathematical concepts for our purposes, we can think of Boolean algebra as the branch of algebra where the variables can be either true or false (1 and 0), and the main operations on its variables are conjuction (and, $\wedge$), disjunction (or, $\vee$), negation (not, $\neg$). Through these, logical operations can be described. In "An Investigation of the Laws of Thought on Which are Founded the Mathematical Theories of Logic and Probabilities", one of the author's two monographs on algebraic logic, George Boole, then mathematics professor in Ireland, introduces Boole's algebra as an extension to Aristotle's logic. In it, Boole provides Aristotle's algebra with mathematical foundations, and expands it from two-term to any-term. Boole's algebra differs from modern Boolean algebra (in Boole's algebra *uninterpretable* terms exist) and cannot be intepreted as set operations; still, its introduction marked a step towards the formalization of laws of thought and a possible bridge between mathematical research and thinking processes (even the title of the book it was introduced in gives a very clear direction). Boolean algebra would instead be developed by Boole's successors (Jevons, Peirce, Schroder and Huntington in particular); this work allows boolean algebra to now be defined by the Stanford Encyclopedia as

> the algebra of two-valued logic with only sentential connectives, or equivalently of algebras of sets under union and complementation.

KEY CONCEPTS: laws of thought can be modeled in mathematics, using algebra.

Automata theory. The study of how automatic calculators (more properly, abstract machines **TODO: mi sbaglio?** or automata) can be used to compute and solve problems is a part of theoretical computer science research. The history of Automata Theory is especially interesting, as it will let us meet some important researchers: it features two neurophysiologists, Warren McCulloch and Walter Pitts, and is thus born from the desire of modeling human thought itself. The first model was proposed in 1943 [30], in a seminal paper that also introduced other research themes we will come back to later. A little over twelve

years later, two computer scientists, Mealy and Moore, generalized the theory to more powerful machines, "finite-state machines". The general idea behind them is this: starting from an input and a set of states, a *transition function* maps the current state and an input to an output together with the next state. They do not have any memory, and as such can only "solve" simpler problems: if used to recognize languages, they can only recognize regular ones.

More powerful abstract machines had already been proposed: Turing had introduced "Turing machines" in 1937 [55], as part of his proof of the Entscheidungsproblem. ~~The relationship between automata "expressive power" and language complexity will be explained in later chapters.~~ **TODO: undecidability problem already puts a stop to modeling thought as maths?**

KEY CONCEPTS: Modeling algebra, so thought, is possible through some computational structure

Cybernetics. Although in recent years the term "cybernetic" has been used to mean futuristic/sci-fi technology, Cybernetics is a transdisciplinary discipline that studies regulatory systems. The core of the discipline are feedback loops (or circular causality), where the result of action is taken as input for (choosing) future actions. Cybernetics isn't bound to any particular usage, so its applications include biology, sociology, computer science, robotics and many others. Its flexible approach led to many different definitions: two early ones are the one used in Macy cybernetics conferences, "the study of circular causal and feedback mechanisms in biological and social systems"[52], and the definition by Norbert Wiener, considered the originator of cybernetics, "the scientific study of control and communication in the animal and the machine"[62]. Although the word itself was used by Plato to signify the governance of people, our interest resides in contemporary cybernetics, born in the 1940s. Before the aforementioned paper by McCulloch and Pitts, the study of feedback was considered by Anokhin in 1935 [2] (physiologist). In the same year as the McCulloch-Pitts paper was published, Wiener, together with Rosenblueth and Bigelow, published "Behavior, Purpose and Teleology" [44]: these three researchers, together with McCulloch, Turing, Grey Walter and Ross Ashby, would go on to establish the discipline of cybernetics. Wiener coined the term to denote "teleological mechanisms".

An important addition to the field would be the Von Neumann cellular automata: these are another model of computation part of automata theory. A cellular automata is a grid of cells (of any dimensions, but for clarity, consider a 2-dimensional one first), where each cell has a finite number of states it can be in; the cellular automata evolves by moving from generation zero ($t = 0$) to the next generation ($t = 1$) following mathematical rules: the state of every cell is determined by its past state and the surrounding cells. Without going into the specific rules Von Neumann determined, this is relevant to us because it introduces the concept of self replication, soon adopted by cybernetics as a core concept. Another fundamental contribution from cybernetics is the creation of Artificial Neural Networks, introduced in the same McCulloch-Pitts paper we mentioned earlier.

KEY CONCEPTS: Study of feedback is subject-agnostic, self replication,

artificial neural networks.

Information theory and technical advances. As we have seen, theoretical advances were many and varied, but the technical advances were what drove the ability to put those in practice. Among those, we have to mention the move from electromechanical devices to vacuum tube-based computers, which gave birth to a device for controlling the connections between telephone exchanges, thanks to Flowers, in 1934. The record for the first general-purpose stored-program (as in, controlled by wires, the opposite of a stored-program computer) went to Konrad Zuse, with the Z3 machine. This machine also used a binary system, but it was not a universal computer. In 1944, the Bletchley Park cryptanalysts started using Colossus. The first Turing-complete (i.e. with the same computing ability as the Turing machine) computer was completed in 1945. It used over 18.000 vacuum tubes. The first stored-program computer, built as a testbed for new technology and design, was the Manchester Baby, ran in June 1948[10].

As part of the advances of this period, we must mention the birth and development of Information Theory. Information Theory encompasses the study of quantification, storage and communication of information, in digital form. After being introduced by Nyquist and Ralph [40], the field was firmly established by Shannon's "A Mathematical Theory of Communication" in 1948. Without going into details, its main influences include the bit as a unit of informationa and the necessity of redundancy of a source when using unreliable communication channels.

Lastly, we note that neuroscience had new tools at his disposal: electrophysiological techniques, such as brain stimulation, single cell recording and EEG recording [19] were instrumental to the research into localization studies (such as deficits derived from brain lesions) approached by Geschwind in the 1950s.

3.1.2 DCS

The DCS landscape around 1950 was strongly rooted in Behaviorism, with hints of the revolution that was soon to come. Some of the larger influences from the Computer Science side, such as the McCullough Pitts artificial neural network we mentioned, would in fact have a relatively small impact and be re-discovered later.

Behaviorism. Behaviorism emerged as the dominant school in Western psychology as a reaction to depth psychology and other forms of psychology that did not fit well with scientific experimental verification. That is not to say it was unprecedented: Thorndike presented the law of effect (using consequences to strengthen or weaken behaviour) in 1898. Still, behaviorism was introduced as "methodological behaviorism" by a 1924 publication by John Watson [58], and then further expanded by many researchers, of which we must mention B. F. Skinner.

Behaviorism, more than a way to impose empirical constraints on studying psychology, is a doctrine of how to do behavioral science itself. The Stanford

Encyclopedia identifies three claims as the roots of behaviorism (as a doctrine):

- Psychology is the science of behavior. Psychology is not the science of the inner mind – as something other or different from behavior.

- Behavior can be described and explained without making ultimate reference to mental events or to internal psychological processes. The sources of behavior are external (in the environment), not internal (in the mind, in the head).

- In the course of theory development in psychology, if, somehow, mental terms or concepts are deployed in describing or explaining behavior, then either (a) these terms or concepts should be eliminated and replaced by behavioral terms or (b) they can and should be translated or paraphrased into behavioral concepts.

These fundamental truths identify three of the various flavours behaviorism is studied in. Skinner, mentioned above, was the first to suggest that covert behavior, such as cognition and emotions, is governed by the same controlling variables as observable behavior: although focused on the third "truth", his philosophy combines all three mentioned pillars, and is described as *radical behaviorism* by skinner himself [49].

One can easily see how the philosophy itself forced the practitioners into a state of absolute experimental dependency, which constrained the concept explored to the scientific realm. At the same time, its complete rejection of mental processes (or at least their relevance to scientific study) is the complete opposite of the assumptions that were made on the "CS" side of comprehension. Other behaviorists, though, were less radical: Clark Hull was willing to put drive inbetween stimulus and response, but only to create a corresponding theory that explained it in terms of behavior[18]; Edward Tolman, instead, proposed rats navigate a maze following a mental map[54].

Cognitive signs. Just like Tolman, other cognitive-leaning psychologists proposed ideas that did not fit with the behavioral narrative. Among them, we mention some relevant ones. The Gestalt psychology refused the behavioristic assumption that conscious experience could be considered by reducing it to the sum of it parts, and proposed the principle of totality; it also proposed the principle of psychophysical isomorphism, which meant the cerebral activity was correlated to conscious activity [57]. Vygotsky and Luria pioneered "cultural-historical psychology", which noted the role of culture and language in the development of higher pschological functions; Luria, alone, also published research on individuals' thought processes as his doctoral dissertation.

Lastly, we mention Miller, who was just a trainee at Stevens's Psychoacoustic Laboratory at Harvard: he will soon become relevant, as part of the 1956 cognitive revolution.

In our exploration of the state of disciplines around 1950, it is clear that Computer Science was firmly en route to a first attempt at thought modeling though mathematical "symbols": if, as they suspected, thought was to be

considered a use of (or better yet, possible to model with) algebra, then once physical computers were capable enough they would be capable of thought. On the other side of the fence, DCS was still firmly rooted in behaviorism: in their view, the entire discussion would be based on false premises which were in turn based on wrongful research; the roots of human behavior were to be found in human behavior itself, and assuming otherwise was not only useless but unscientific, as it would lead to unprovable theories and impossible experiments. At the same time, cognitive suggestions were starting to appear, challenging the general (or at the very least American) current view.

3.2   1956: A Pivotal Year

As we have mentioned in the previous section, there were various lines of research into thought modeling: automata theory was focused on what problems were possible to model, cybernetics took (analog and biology-based) feedback and self replication as founding pillars, while information theory dedicated itself to information storage and transmission. Instead, DCS was still mostly led by behaviorist views, but cognitive-oriented proposals started to emerge. This trend would continue in 1956, and spike in 1957 with a publication by Noam Chomsky that would change the field.

3.2.1   CS

The most relevant event of 1956 (and quite possibly of the history of AI research) is the Darmouth College Workshop, a sort of convention that connected researchers from diverse fields interested on similar topics. This is also the context in which the term "Artificial Intelligence" was attached to the field.

Dartmouth College Workshop. As we said, at the start of the 1950s thinking machines were being inspected by a few different disciplines; in 1955 John McCarthy, an Assistant Professor at Dartmouth, proposed a conference to organize and fertilize such disciplines. He proposed the name "Artificial Intelligence" because, unlike today, it was still neutral; Wikipedia reports that avoiding cybernetics was partly due to 'him potentially having to accept the assertive Norbert Wiener as guru or having to argue with him'. The project was formally proposed in September, by four of those who would become (if they weren't already) prominent researchers in the field: McCarthy himself, Marvin Minsky, Nathaniel Rochester and Claude Shannon. Among the extraordinary attendees we mention: Minsky (who will become very relevant in the next section), Bigelow (co-author of the seminal paper "Behavior, Purpose and Teleology." on cybernetics), Solomonoff (inventor of algorithmic probability), Holland (pioneer of genetic algorithms, was invited but did not end up attending), Ross Ashby (psychiatrist and cybernetics pioneer), McCulloch (who we've already mentioned), Nash (prolific mathematician, also known for his work on game theory), Samuel (creator of what is considered the first AI program, a checkers program) and finally Allen Newell and Paul Simon, who presented their recently completed "Logic Theorist". Although the discussions were not

9

directed, many of the topics would have a long-lasting impact on the field, like the rise of symbolic methods and limiting domains (which would lead to expert systems).

Logic Theorist. The Logic Theorist was created in 1955 by Newell and Simon, helped by the systems programmer John Shaw. In order to prove a theorem, the simplest strategy is to start from the theory's postulates and create new theorems by combining them; then continue by combining every theorem with every postulate and every theorem again, exploring the entire truth spectrum. Although this may seem obvious, this is part of the first important concept introduced by the Logic Theorist: seeing the truth space as a tree, that started with the hypothesis and aimed at the proposition to prove; envisioning *reasoning as search*. Of course, exploring the entire tree is impractical, because of the time it takes to explore the entire truth space (as it grows exponentially); when considered from the point of view of "modeling the human thought process", this solution would not be useful even if it was practical, because this is not how human theorem-provers work. In order to solve this problem, the Logic Theorist introduced the second important factor: employing *heuristics* to ignore branches that were unlikely to lead to the goal. The last important factor is technical: in order to implement the Logic Theorist, the authors implemented IPL, a programming language that used symbolic *list processing* in the same way as the following, fundamental, Lisp.

3.2.2 DCS

In this section, we'll talk about some of the important findings that seemed difficult to integrate with behaviorism and the important actors behind them.

Miller. George Miller, before becoming one of the founders of cognitive psychology, was of the behaviorist school (although he later wrote of one of his works on language 'By Skinner's standards, my book had little or nothing to do with behavior'[31]). After slowly moving to the cognitive side, driven by similar thinkers ('Peter Wason, Nelson Goodman and Noam Chomsky had the most influence on my thinking at that time'). In 1956 he published a paper that had a sizable impact: "The magical number seven, plus or minus two". In it, he observed tht various experimental findings revealed that, on average, human can hold seven items in short-term capacity. We note that it is not the finding that goes against behavioral philosophy and psychology, but the framework in which it is put in general: such attention to mental processes would be irrational, when seen from a behavioral point of view, hence disregarding mental processes as a whole. **TODO: is this right? check with mom**.

Bruner, Goodnow, Austin, and the basis of cognitive science. Another important book published in 1956 is "A Study of Thinking", by Bruner, Goodnow and Austin. The book is focused on using categories for concept formation, or how human beings group the world of particulars into classes, together with the results of relevant experiments. Before such experiment on cognition, Bruner had dedicated himself to the study of perception: two rele-

vant studies we report are the one on estimating the sizes of coins or similarly sized wooden sticks (the first were significantly overestimated), and another one on slowing reaction times while playing cards in connection with reversed suit symbols. These two experiments are relevant because of the focus on the internal interpretation of external stimuli. Other foundational ideas of cognitive science, developed in the years following the Miller publication, include the application of the scientific method to human cognition (if anything was to come after behaviorism, it could not avoid its history of scientific "rigor"), the interest towards information processing and storage, and as we will see in the next paragraph, a degree of possible innateness.

Chomsky and the final departure. In this last section, we move further than 1956. Nonetheless, it is extremely relevant to the subject discussed, and represents the strongest blow (in purely historical terms; this document has no psychological authority to express an evaluation of *any* theory) to behaviorism. In 1957, Skinner published "Verbal Behavior". In it, he describes the controlling elements of verbal behavior, and attempts to form a hypothesis about the behavioral framework with which verbal behavior is to be understood. In it, he uses specific terminology for his analysis, using both existing words and neologisms; in his own words: 'The emphasis [in Verbal Behavior] is upon an orderly arrangement of well-known facts, in accordance with a formulation of behavior derived from an experimental analysis of a more rigorous sort. The present extension to verbal behavior is thus an exercise in interpretation rather than a quantitative extrapolation of rigorous experimental results'[50].

In the same year, Chomsky proposed a different model for understanding language; in "Syntactic Structures" he argued two important points that would have a large impact on the field of linguistics:

1. **Syntax vs semantics.** The first point he makes is the clear distinction between syntax and semantics: '...such semantic notions as reference, significance, and synonymity played no role in the discussion.'

2. **Generative grammars.** His approach to syntax was formal, and followed both his teacher's (Zellig Harris) and notions advanced by Danish linguist Louis Hjelmslev: language was to be understood as a generative grammar, which bound by "phrase structure rules" (producing new sentences) and "transformations" (modifying exising sentences).

This seminal paper would soon be interpreted as an argument for a mentalistic, *innate* view of language production, but this intepretation was not originally put forth in the book itself: '[Chomsky's generative system of rules] was more powerful that anything ... psycholinguists had heretofore had at their disposal. [It] was of special interest to these theorists. Many psychologists were quick to attribute generative systems to the minds of speakers and quick to abandon ... Behaviorism'[53].

Two years later, Chomsky published a scathing (this time both in historical terms and considering the tone of the paper) review [8] of the book which had a widespread effect of the decline of behaviorism's influence. In it, one of the

points he argued was that children are not taught the rules of language, and the amount of input they receive is not sufficient to derive them. This argument would later be called the "Poverty of the Stimulus" argument, and to this day represent a very controversial issue of linguistics and language acquisition.

In the words of Newmeyer:

> Chomsky's review has come to be regarded as one of the foundational documents of the discipline of cognitive psychology, and even after the passage of twenty-five years it is considered the most important refutation of behaviorism. Of all his writings, it was the Skinner review which contributed most to spreading his reputation beyond the small circle of professional linguists.

[38]. The review has been criticized by other writers, such as MacCorquodale [24], but its effect cannot be ignored.

As we have explored in this section, 1956 was both the culmination and start of a cognitively-inspired revolution across the DCS. As behaviorism grew less popular, cognitive findings and research drew more interest. At the same time, one of the very first AI programs was presented, and it tackled a purely symbolical problem with a purely symbolical approach: the trend was clear, and it was pushing towards a cognitive approach.

3.3   1960-1970: Great Promise

The years after 1956 are considered by many the "golden years" of AI research: thanks to considerable successes and a general wave of optimism, money was poured into the field, which thankfully generated more results and increased the hopes again. Although some interest was generated towards neural networks, this was completely shut down by a Minsky critique in 1969 (analogous to the Chomsky critique of "Verbal Behavior"). Psychology saw the rise of research into representations, categories and memory, as we will briefly overview in this section.

3.3.1   Computer Science and Artificial Intelligence

For what concern Computer Science and Artificial Intelligence, this period saw various directions, inspired by some of the previous research we've touched on in the previous sections. Most of them were focused on symbolic AI; at the same time, the "perceptron" proposal from McCulloch and Pitts saw some interest, before being shut down for more than ten years.

Reasoning and the General Problem Solver. As we mentioned, an important paradigm was introduced with the Logic Theorist: seeing reasoning as search. Newell and Simon, in 1959, developed what they hoped could become a general version of the LT, the General Problem Solver. Although the paradigm of reasoning as search was maintained, the GPS did not prune paths that were unlikely to lead to the goal, but used *means-ends analysis* to limit search. When

following MEA, a system chooses, given a current state, an action that reduces the difference between the current state and the goal state. By focusing on the difference between current and goal state, MEA improves on brute-forcing all possible choices. In addition, if knowledge about the relative importance of differences is available, the goal-seeking system can follow the path which decreases the difference most, further pruning the possible choices. The correspondance difference-action, also called operator, must be given as an input, and represents "a priori knowledge" of the problem. This separation between problem-specific knowledge and strategy of how to solve it is a relevant feature of the project, and an important point when compared with the following paradigm, expert systems.

Symbols and successes. Following the GPS, other symbol and knowledge-based systems led to great successes. In 1958, the same year in which he invented Lisp, McCarthy published "Programs with Common Sense"; in it, he described a hypothetical program that used general knowledge to search for solutions to problems (such as generating a plan to drive to the airport). Called the Advice Taker, it also allowed for additional knowledge (axioms) to be introduced during the course of operation. As such, it embodied the a simple principle of knowledge representation: manipulating a formal representation of the world and its workings as a mean to solving problems. For later purposes, we mention the Shakey project at Stanford, which used subgoals (like GPS) and logic to control a robot. Minsky, who moved to MIT in 1958, started supervising students who tackled limited problems that seemed to require intelligence to solve: these would become known as microworlds. Two of these were Daniel Bobrow's STUDENT (1967), which could solve high school algebra word problems and Tom Evans's ANALOGY (1968), that solved geometric analogy problems from IQ tests. Research based on microworld continued throughout the 1970s.

Finally, we mention the different perspective taken by Joseph Weizenbaum, then MIT professor: between 1962-1964 he created ELIZA, a natural language processing program that mimicked conversational ability while following a simple script, the most popular of which was the "Rogerian" DOCTOR. Although the creation of it was meant to show how superficial interaction between machines and people really is, it gave (although briefly) the impression of an intelligent interaction; it did not even have any storage, so links between sentences were impossible. Its relevancy is now both historical, as it was the first attempt at creating the illusion of intelligence through human-machine interaction, and ethical, as its creation led to ethical questions regarding its usage as a therapeutic tool.

Genetic algorithms and perceptrons. Between the late 1950s and early 1960s [13], Friedberg started researching machine evolution (later called genetic algorithms), with scarce success. The basic idea was to make a series of small modifications to a program, then select the best-performing variant and repeat the process until the result was good enough. Unfortunately, due to how immature representation research and because of computing power constraints, these showed very limited success and the program was dropped.

Following the work of McCulloch and Pitts, research on neural networks picked up. Bernie Widrow researched his adalines[61], while Rosenblatt researched perceptrons. In addition, in 1962 Block showed, with the perceptron convergence theorem, that if a pattern of connection strenghts that matches a certain input data exists, then the learning algorithm can adjust the strengths correctly[45].

3.3.2 DCS

DCS research in this period was mainly focused on modeling higher-brain function, such as memory, within the new framework of cognitive psychology. At the same time, behaviorism shifted from the strict theoretical research and found itself evolving into Applied Behavior Analysis as a scientific discipline used in therapy.

Behaviorism shifts closer to its current form. In a study from 1959, "The psychiatric nurse as a behavioral engineer"[3], the authors demonstrated the effectiveness of using a token economy to reinforce adaptive behavior for patients with schizophrenia and intellectual disability. The practical application of behavioral research grew throughout the years: a journal, the "Journal of Applied Behavior Analysis" was founded in 1868; the "Behavior Analysis" subdivision in the American Psychological Association was introduced in 1964; the "Applied Animal Behaviour Science" was founded in 1975.

Memory research. A series of studies on memory by different researchers helped clear the picture: Sperling focused, throught the 1960s, on sensory memory, starting with his PhD thesis at Harvard in 1959, and papers like "The information available in brief visual presentations"[51] and "Successive approximations to a model for short-term memory" (1967). Peterson worked on short term memory, with "Short-term retention of individual verbal items."[42], and Waugh studied the difference between short-term memory and long-term memory in "Primary memory."[59]. This allowed Atkinson and Shiffrin to propose, in 1968, the Atkinson-Shiffrin memory model: it viewed memory as a tripartite system, split between sensory memory, short term memory and long term memory. Although the idea of tripartite systems wasn't novel [20] and some of the concepts it included, like rehearsal as the transfer mechanism, have been criticized by later research, it sparked additional interest in the area of memory.

Milestones and Neuroscience. In addition, 1960 was the inauguration year for Miller-Bruner center for Cognitive Studies. At the same time, neuroscience was developing: in 1962, the FitzHugh-Nagumo model was presented, as a semplification of the previous Hodgkin–Huxley model. These models gave a formal background to the activation behavior of neurons (once the stimulus reached a threshold, the system is briefly excited before going back to resting state). In the same period, Katz modeled neurotrasmission across synapses [22][21], and from 1966 Kandel and others started examining biochemical reactions to learning and memory in Aplysia (a genus).

Lastly, we mention a line of research by Shepard and his student Cooper and

Metzler, in which they showed that reaction time in subjects asked to determine whether a transformation was a rotation or a reflection increased linearly with rotation degree: this suggested an internal image that was being rotated.

Throughout the 1960s, cognitive research started to appear, and was soon to be institutionalized. Artificial Intelligence programs focused mainly on symbolic systems, but research on neural networks increased, with some important theoretical findings supporting the hopes of the ideators. This was to come to a screeching halt, with, once again, a critical (and later partly controversial) review of the literature by Minsky and funding issues.

3.4   1970-1985: Symbols and Knowledge

During this period, AI left the connectionism world behind, following a Minsky literature review in 1969. Meanwhile, symbolic research continued, and found its new paradigm: using domain-specific knowledge to solve bigger, more complicated tasks in a *specific* field. The DCS domain expanded on their previous views, and some important points of contact were explored.

3.4.1   Computer Science

As we mentioned, research in this period mainly focused on symbolic systems. The addition of context and domain specific knowledge shifted interest from imitating human or semi-human intelligence to strictly solving problems as well as possible. Although these *expert systems* worked well, a series of companies and projects that overpromised advances were born, which lead to what is called "AI winter", when companies and nations realized their hope was, sometimes, unfounded.

Minksy and perceptrons. As we mentioned in the last section, research on connectionist models, although less popular than symbolic models, kept going. One of the most relevant techniques for training neural networks was introduced in 1969[7], although the original research was about optimal control instead of machine learning. This is considered the first description of modern back-propagation[23], but their version was never applied to machine learning, where backpropagation would be rediscovered in the 1980s. Instead, research on perceptrons slowed down considerably after 1969. In the words of the standard textbook on the topic "The subsequent demise of early perceptron research efforts was hastened (or, the authors later claimed, merely explained) by the book Perceptrons, which lamented the field's lack of mathematical rigor ... and noted the lack of effective learning algorithms for multilayer networks". The book in question, Perceptrons[35], also noted other theoretical limitations of perceptrons, some of which may have been misinterpreted and so contributed to the general feeling towards perceptrons in general. Admittedly, in the last chapter the authors mention that multilayer neural nets are a "sterile" extension.

Integrating knowledge. The approach so far (consider for example the efforts behind the GPS) had been to solve problems by modeling human thinking

processes in their most general form, to be then applied to a specific problem. With DENDRAL, a research project that started in 1965, the creators explored a different approach: it used a large number of special-purpose rules, extracted from analytical chemists, to infer molecular structure of molecules from mass spectrometer data and the elementary formula of the molecule. As such, it became the first knowledge-intensive system. It was soon to be followed by other expert systems, such as MYCIN, which diagnosed blood infections, and gave rise to a new, very successful paradigm. Its successes brought AI into the commercially viable technologies, and the increase of demand brought an increased interest in knowledge representation schemes: in particular, the two main approaches were logic-based and frame-based. Frames were a Minsky proposal[33] to organize facts about objects and events into a large hierarchical taxonomy.

Natural language understanding. In the previous section, we introduced research on microworlds. The most famous microworld was the block world, a set of blocks on a tabletop (real or virtual), which had to be rearranged, one block at a time, according to instructions. The success of this microworld (SHRDLU was the name of the successful program [64]) derived from the cooperation of many different researchers, as "Artificial Intelligence: a Modern Approach" reports:

> The blocks world was home to the vision project of David Huffman (1971), the vision and constraint-propagation work of David Waltz (1975), the learning theory of Patrick Winston (1970), the natural-language-understanding program of Terry Winograd (1972), and the planner of Scott Fahlman (1974).

[45] As a fitting consequence for such successes, in 1976 Newell and Simon formulated the Physical Symbol System Hyposthesis [37]. It states that "a physical symbol system has the necessary and sufficient means for general intelligent action"; which means that any system possessing intelligence must operate by manipulating symbols. This statement would later be challenged by many researchers.

Natural language understanding was another area in which domain knowledge had vast effects: the success of the block world-natural language program was in fact due to its specificity, and a series of programs followed it[47][63][46], which all focused on understanding natural language by reasoning with the knowledge required.

3.4.2 DCS

During the 1970s, Cognitive Science went from a collection of studies with similar intentions to a true discipline, complete with relevant courses, journal and grants for research. Points of contact with existing AI research became more common, as did researchers working in both fields. Some philosophers, possibly accusing the weight of the wrongful predictions of AI research, raised critiques towards the field in general, while others began developing collective theories of mind, from Fodor's functionalism to the Computational Theory of Mind.

Points of contact. In this paragraph, we will highlight two relevant contact points during this period. The first is the result of a collaboration between Gordon Bower and one of his students, John Anderson. Bower had gone from learning theory and animal testing, to mathematical models of learning to, finally, cognitively oriented work about mental representation, such as the study of chunking for short-term memory usage[15]. Their cooperation would give rise to a semantic network model named HAM, later described in their 1973 boook Human Associative Memory[1]. Anderson would keep working on it, later adding a production system, increasing the types of nodes and links between nodes, and explaining the time it takes to perform a task as due the matching for the production system, until he would publish the ACT-R architecture, still in research today.

The second important connection was a clash between the author of ELIZA, mentioned in the last section, and the psychiatrist Kenneth Colby. Colby expanded on the work of Weizenbaum[9], and wrote what he considered was a "computer program which can conduct psychotherapeutic dialogue", with which Weizenbaum clearly disagreed. Later, in 1976, Weizenbaum published "Computer Power and Human Reason"[60], where he declares that computers should never be allowed to make important decisions as they would always lack compassion and wisdom.

Putnam, Fodor, and the Computational Theory of Mind. In this section, our aim is to give a brief introduction to the philosophical path before the Computational Theory of Mind (CTM), a family of theories and views which hold that the human mind is an information processing system, and as such cognition and consciousness (sometimes not both, according to the specific theory) are a form of computation. Although it was introduced in 1961 by Putnam [17], it was developed by Fodor throughout the following decades. The Stanford Encyclopedia defines the CTM as combining "an account for reasoning with an account of the mental states". Of these, the second one (Representational Theory of Mind), argues that intentional mental states, such as beliefs, are relations between "a thinker and symbolic representations of the content of the states"('I believe there is a book on the table' would be the functional *belief* relation between me and the mental, symbolic representation of 'there is a book on the table'). The first, instead, maintains that reasoning involves the symbolic representations *only* in their non-semantic, syntactic properties. As such, this process can be considered a formal symbol manipulation, which qualifies as computation.

The relational character of mental states was initially introduced by Fodor in 1978[11], where he identified mental states as a three-way relationship between the individual, representations and propositional contents.

We believe that, seen in this context, the Computational Theory of Mind is a coherent theory for a few different viewpoints in the history of AI. At the same time, as we have seen in this section, AI research steered away from such open-ended questions in favor of technical and engineering achievemnts, based on restricted domains, background knowledge and specific rules: a far cry from

the cognitive model the CTM would imply, and detached from the connectionist possibilities it considered just a few years prior.

3.5   1987-1993: Bodies as the Key to Minds

Unfortunately, the enthusiasm for expert systems did not last long. As it had happened before, the hype behind AI research was too great for its own good, and towards the end of the 80s the market suffered a few serious blows. As always, though, a new wind was blowing, with several new paradigms. One in particular was very closely related to DCS research: embodied cognition.

3.5.1   Computer Science, AI and engineering

The shortcomings of expert systems, serious as they may be, had their main effect on the economic side, as research continued. Here, we will go over the reasons behind the economic crash and the new perspective in research.

Expert Systems and the Hype. The downfall of expert system was foreseen by some in the research community: from a 1984 article[56],

> Yet Minksy and Schank contend that today's systems are largely based on 20-year-old programming techniques that have merely become practical as computer power got cheaper. Truly significant advances in computer intelligence, they say, await future breakthoughs in programming.

Although their argument was a theoretical one, the practical implications had a large impact on the market: in the late 1980s, desktop computers were slowly overtaking specialized and expensive Lisp machines. In 1987, the reasons to buy them simply ended, and a large industry fell overnight.

The expert systems themselves started to show their flaws: they couldn't be updated, could not learn, and made large mistakes when given unusual innputs. Some issues with them had been shown years earlier, like the qualification problem (the inability of listing all the necessary preconditions for an action in the real world to have its intended effect). They worked in very specific scenarios, but were not as successful a recipe as they had been presented. Some of the initiatives launched were retracted, and funding dwindled[29].

Robotics.   As a direct consequence of the "lowering the mind into the body", robotics went back to the forefront of AI research. In 1990, Brooks published "Elephants Don't Play Chess"[6], in which he argued that symbols are unnecessary for cognition, because

> the world is its own best model. It is always exactly up to date. It always has every detail there is to be known. The trick is to sense it appropriately and often enough.

This is obviously against the Physical Symbol System Hypothesis, and represents a general awakening towards robotics-based approaches. In fact, symbol-sustaining researchers such as Minsky felt similarly; in 1986 Minsky writes "In

general, we're least aware of what our minds do best, [...] we're more aware of simple processes that don't work well than of complex ones that work flawlessly"[34]. The comparative difficulty of sensorimotor skills compared to reasoning is considered the Moravec's paradox:

> it is comparatively easy to make computers exhibit adult level performance on intelligence tests or playing checkers, and difficult or impossible to give them the skills of a one-year-old when it comes to perception and mobility

[36] This general feeling led to research into Behavior-Based Robotics (robots that exhibit complex behavior while having little internal modeling state) and Nouvelle AI, pioneered by Brooks himself, working on situated (robots interacting with their sensors and their environment) robots with intelligence close to one of an insect.

Connectionism returns! Around the mid-1980s, following the 'downfall' of expert system and symbolic approaches in general, the connectionist approach based on *neural networks* made a strong resurgence. A new algorithm for training such networks was rediscovered by at least four different groups[45]: it was the same algorithm found in 1969 by Bryson and Ho. We presented this field of research as it is usually presented: as if in antithesis to symbolic approaches. As we will see in a following paragraph, though, research was beginning to shift to a cooperative view, using complementary approaches to explaining cognition. The field went on to bifurcate into a CS and engineering-focused strand (exploring possible neural architectures, determining their properties) and a neuroscience/empirical-focused strand (modeling biological neurons as accurately as possible).

Probabilistic reasoning. We mentioned how brittle expert systems were found to be when they were applied in the real world: this drove researchers to a more scientific approach, aiming at reproducible experiments instead of philosophical claims, building on existing theories instead of constantly introducing new approaches, and including probability instead of Boolean logic. This has been called the victory of the *neats*, but it would not stand forever; the recent dive into deep learning has shown impressive results from an overall *scruffies*-based philosophy.(preferring to try new ideas without completing the mathematical background first). Probabilistic reasoning was pushed forward by a 1988 piece by Judea Pearl[41], who introduced an efficient formalism for dealing with uncertain data as well as practical algorithms.

3.5.2 DCS: between philosophy, psychology and neuroscience

The disillusion towards expert systems led some researchers to advocate a new approach, based on the physical world and robotics. They considered abstract thinking to be the least interesting human skill, and argued for "lowering" the mind into the body. The approach was not new: we remind the reader about the impact of cybernetics on the birth of the field of AI.

Marr and computational neuroscience. David Marr had a similar approach towards vision, about a decade earlier. With papers in 1969, 1970 and 1971 he proposed computational theories on cerebellum[28], neocortex[25], and hippocampus (which he called 'archicortex')[26]. Afterwards, he focused on vision, together with the Italian researcher Tomaso Poggio. To them, vision was to be understood 'bottom-up', focusing on the physical level before any symbolic processing. They considered vision an information processing system, to be analyzed at three levels[27]:

- *Computational level.* What the system does and why.

- *Algorithmic level.* How the system does what it does, and with which processes.

- *Physical level.* How the system is implemented.

The system may seem very simple; nonetheless, the idea of levels of analysis and its similarity to computational approaches signal the resurgence of interest in computational neuroscience that was on the horizon.

Fodor and modularism. The interest toward lower level systems wasn't new, especially in the philosophical side of research. Fodor had been advocating for a different notion to understand the mind, also very reminiscent of technical paradigms in Computer Science: modularism. As it was introduced in 1983[12] and developed in the decades since, it considered a system (i.e. the mind) modular if it was at least partially composed by subunits, innate neural structures with distinct, evolutionary-developed functions. The definition of module changed, but the initial proposal contained 9 features that characterize such systems; of these, we mention:

1. **Information inaccessibility and encapsulation**. The direction of information flow is restricted; for example, althuogh you may be aware of perceptional issues while watching an optical illusion, the perception will not change.

2. **Speed and superficiality**. Modular systems are mostly fast (Fodor considers roughly half a second) and concern superficial concept: in Fodor's book this may be interpreted as computationally simple (few calculations) or informationally simple (general); both may be true, and Fodor generally excludes theoretical concepts such as "turbine" or "proton"

3. **Dissociability**. A system is functionally dissociable if it can be damaged without significant impairment to other systems.

4. **Innateness**.

Fodor considers relatively low-level systems of the mind to be modular (like perception or language), while high-level systems are not to be understood by modularism. Following thinkers would go on to expand the idea to "Massive Modularity", arguing all elements of the mind are in fact modular.

As we saw in this section, the crisis of expert system, while it crippled the overly excited AI market, coincided with a renewed interest towards embodied and situated systems. This was a common thread between AI and DCS research. We close this chapter by mentioning that a companion theory to situated robotics (closely related to Nouvelle AI) in the DCS was Situated Cognition, which slowly emerged at the end of the twentieth century. Because it argues for learning as and individual's increasingly effective performance across situations, with cognition inseparable from the context, it is closer to Skinner's behavior analysis than previous storage-and-retrieval-based theories.

3.6   1993-2010: Agents and Cooperation

All the paradigms we introduced in the last chapter were developed throughout this period as well, so we won't get into them again here. Still, this newfound period of success had two important features: the rekindled interest in the general problem of AI, and the availability of very large data sets. On the DCS side, the 90s brought fundamental new technologies, extensions of previous research and the use of genetic techniques.

3.6.1   Artificial Intelligence

As we mentioned, the (academic, first) success of the new paradigms pushed researchers to solving the "whole agent" problem again. In particular, the new context in which agents had to learn to operate was the Internet: AI algorithms started to act as the foundations behind, for example, search engines and recommender systems. Clearly, the process of merging previous results in separate tasks had its own share of issues, but the ideas we mentioned in the previous section allowed for a more complete picture: sensory systems (whether that was speech recognition or vision) were known to provide imperfect information, so planning systems had to handle them accordingly, using probabilistic approaches. Examples of this are the two challenges set by DARPA for autonomous driving, respectively 135 miles along an unknown desert trail, completed in 2005 by STANLEY, and 22 miles in an urban environment, completed in 2007 by BOSS. This approach is known as the "Intelligent Agents" approach; researchers hoped that a complete agent architecture (like Newell's SOAR, recently described[43]) would give researchers the tools to build intelligent systems from the interaction of agents.

Big Data. Still, the largest impact on the world of research was probably the new availability of very large datasets, thanks in no small part to the pervasive effects of the Internet. Researchers Banko and Brill[4] argued that the increase in the size of the dataset (two or three-fold) would outgrow any advantage that was to be found by tweaking the algorithm. This sentiment, which is not by any mean a formal proof, is echoed throughout the machine learning industry: another article by Norvig et al mentions, in the context of learning from text, "But invariably, simple models and a lot of data trump more elaborate models based on less data"[16].

Throughout the 2010s, AI research racked up a series of wins; apart from the ones already mentioned, we also note: autonomous planning and scheduling in space exploration, by REMOTE AGENT(2000, generated plans and monitored their execution), MAPGEN (2004, the previous one's successor, planned NASA's Mars Exploration Rover), MEXAR2 (2007, mission planning fo the European Space Agency's 2008 Mars mission); game playing, by DEEP BLUE in chess and Watson in Jeopardy!; logistics planning, by DART (DARPA's logistics planner for the 1991 Persian Gulf crisis). We note that DARPA mentioned the deployment of DART more than paid back their 30-year investment in AI.

3.6.2 DCS

the fMRI, the TMS, the current version of the PET scan

3.7 2010-now: New Perspectives and Hybrid Systems

3.7.1 gianandrea

4 Perception shifts

Should I merge these two?

4.1 Symbolism and Connectionism

4.2 Symbols and Subsymbols: Collect or Extract

5 Conclusion

References

[1] John R. Anderson and Gordon H. Bower. "Human Associative Memory". English. In: (1973). (Visited on 06/21/2021).

[2] P. K. Anokhin. "Problems of centre and periphery in the physiology of nervous activity". In: *Gorki, Gozizdat* (1935).

[3] Teodoro Ayllon and Jack Michael. "The psychiatric nurse as a behavioral engineer". In: *Journal of the Experimental Analysis of Behavior* 2.4 (Oct. 1959). tex.pmcid: PMC1403907, pp. 323–334. ISSN: 0022-5002. DOI: 10.1901/jeab.1959.2-323. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1403907/ (visited on 06/20/2021).

[4] Michele Banko and Eric Brill. "Scaling to very very large corpora for natural language disambiguation". In: *Proceedings of the 39th annual meeting of the Association for Computational Linguistics*. 2001, pp. 26–33.

[5] Margaret A. Boden. *Mind as Machine: A History of Cognitive Science*. English. Clarendon Press, June 2008. ISBN: 978-0-19-954316-8.

[6]   Rodney A Brooks. "Elephants Don't Play Chess". English. In: (1990), p. 12.

[7]   Arthur E. Bryson and Yu-Chi Ho. *Applied optimal control: optimization, estimation, and control*. A Blaisdell book in the pure and applied sciences. Waltham, Mass: Blaisdell Pub. Co, 1969.

[8]   Noam Chomsky. *A Review of BF Skinner's Verbal Behavior*. Harvard University Press, 2013.

[9]   Kenneth Mark Colby, James B. Watt, and John P. Gilbert. "A Computer Method of Psychotherapy: PRELIMINARY COMMUNICATION". en-US. In: *The Journal of Nervous and Mental Disease* 142.2 (Feb. 1966), pp. 148–152. ISSN: 0022-3018. URL: `https://journals.lww.com/jonmd/Citation/1966/02000/A_COMPUTER_METHOD_OF_PSYCHOTHERAPY__PRELIMINARY.5.aspx` (visited on 06/21/2021).

[10]   *Computer Resurrection Issue 20*. Jan. 2012. URL: `https://web.archive.org/web/20120109142655/http://www.cs.man.ac.uk/CCS/res/res20.htm#d#d` (visited on 06/17/2021).

[11]   J. A. Fodor. "Propositional Attitudes". In: *The Monist* 61.4 (Oct. 1978), pp. 501–524. ISSN: 0026-9662. DOI: `10.5840/monist197861444`. URL: `https://doi.org/10.5840/monist197861444` (visited on 06/22/2021).

[12]   Jerry A. Fodor. *The modularity of mind*. MIT press, 1983.

[13]   R. M. Friedberg. "A Learning Machine: Part I". In: *IBM Journal of Research and Development* 2.1 (Jan. 1958), pp. 2–13. ISSN: 0018-8646. DOI: `10.1147/rd.21.0002`.

[14]   Howard E. Gardner. *The Mind's New Science: A History of the Cognitive Revolution*. June 1987.

[15]   *Gordon H. Bower. - PsycNET*. URL: `https://content.apa.org/record/2007-00058-003` (visited on 06/21/2021).

[16]   Alon Halevy, Peter Norvig, and Fernando Pereira. "The Unreasonable Effectiveness of Data". en. In: *IEEE Intelligent Systems* 24.2 (Mar. 2009), pp. 8–12. ISSN: 1541-1672. DOI: `10.1109/MIS.2009.36`. URL: `http://ieeexplore.ieee.org/document/4804817/` (visited on 06/24/2021).

[17]   Steven Horst. "The Computational Theory of Mind". In: (July 2003). URL: `https://stanford.library.sydney.edu.au/archives/fall2015/entries/computational-mind/` (visited on 06/22/2021).

[18]   C. L. Hull. "Goal attraction and directing ideas conceived as habit phenomena". In: *Psychological Review* 38.6 (1931), pp. 487–506. ISSN: 1939-1471(Electronic),0033-295X(Print). DOI: `10.1037/h0071442`.

[19]   *International Encyclopedia of Social & Behavioral Sciences - 1st Edition*. URL: `https://www.elsevier.com/books/international-encyclopedia-of-social-and-behavioral-sciences/smelser/978-0-08-043076-8` (visited on 06/14/2021).

[20] William James et al. *The principles of psychology*. Vol. 1. 2. Macmillan London, 1890.

[21] B. KATZ. "The release of neural transmitter substances". In: *Liverpool University Press* (1969), pp. 5–39. URL: https://ci.nii.ac.jp/naid/10009658302/ (visited on 06/20/2021).

[22] B. Katz and R. Miledi. "Ionic Requirements of Synaptic Transmitter Release". English. In: *Nature* 215.5101 (Aug. 1967). tex.copyright: 1967 Nature Publishing Group, pp. 651–651. ISSN: 1476-4687. DOI: 10.1038/215651a0. URL: https://www.nature.com/articles/215651a0 (visited on 06/20/2021).

[23] Yann LeCun et al. "A theoretical framework for back-propagation". In: *Proceedings of the 1988 connectionist models summer school*. Vol. 1. 1988, pp. 21–28.

[24] Kenneth MacCorquodale. "On Chomsky's review of Skinner's Verbal behavior". In: *Journal of the experimental analysis of behavior* 13.1 (1970), p. 83.

[25] D. Marr and Giles Skey Brindley. "A theory for cerebral neocortex". In: *Proceedings of the Royal Society of London. Series B. Biological Sciences* 176.1043 (Nov. 1970), pp. 161–234. DOI: 10.1098/rspb.1970.0040. URL: https://royalsocietypublishing.org/doi/10.1098/rspb.1970.0040 (visited on 06/23/2021).

[26] D. Marr and Giles Skey Brindley. "Simple memory: a theory for archicortex". In: *Philosophical Transactions of the Royal Society of London. B, Biological Sciences* 262.841 (July 1971), pp. 23–81. DOI: 10.1098/rstb.1971.0078. URL: https://royalsocietypublishing.org/doi/10.1098/rstb.1971.0078 (visited on 06/23/2021).

[27] D. Marr and T. Poggio. "From Understanding Computation to Understanding Neural Circuitry". English. In: (May 1976). URL: https://dspace.mit.edu/handle/1721.1/5782 (visited on 06/23/2021).

[28] David Marr. "A theory of cerebellar cortex". English. In: *The Journal of Physiology* 202.2 (1969). tex.copyright: © 1969 The Physiological Society, pp. 437–470. ISSN: 1469-7793. DOI: 10.1113/jphysiol.1969.sp008820. URL: https://physoc.onlinelibrary.wiley.com/doi/abs/10.1113/jphysiol.1969.sp008820 (visited on 06/23/2021).

[29] Pamela McCorduck. *Machines Who Think: A Personal Inquiry into the History and Prospects of Artificial Intelligence*. AK Peters Ltd, 2004. ISBN: 978-1-56881-205-2.

[30] Warren S. McCulloch and Walter Pitts. "A logical calculus of the ideas immanent in nervous activity". English. In: *The bulletin of mathematical biophysics* 5.4 (Dec. 1943). tex.copyright: 1943 The University of Chicago Press, pp. 115–133. ISSN: 1522-9602. DOI: 10.1007/BF02478259. URL: https://link.springer.com/article/10.1007/BF02478259 (visited on 06/16/2021).

[31] George A Miller. "The cognitive revolution: a historical perspective". English. In: *Trends in Cognitive Sciences* 7.3 (Mar. 2003), pp. 141–144. ISSN: 13646613. DOI: 10.1016/S1364-6613(03)00029-9. URL: https://linkinghub.elsevier.com/retrieve/pii/S1364661303000299 (visited on 06/18/2021).

[32] *Mind.* English. URL: https://www.britannica.com/topic/mind (visited on 06/14/2021).

[33] M. Minsky. "A framework for representing knowledge". In: (1975).

[34] Marvin Minsky. *The Society of Mind.* 1986. URL: https://www.amazon.com/Society-Mind-Marvin-Minsky/dp/0671657135 (visited on 06/24/2021).

[35] Marvin Minsky and Seymour Papert. "Perceptron: an introduction to computational geometry". In: *The MIT Press, Cambridge, expanded edition* 19.88 (1969), p. 2.

[36] Hans Moravec. *Mind children: The future of robot and human intelligence.* Harvard University Press, 1988.

[37] Allen Newell and Herbert A. Simon. "Computer science as empirical inquiry: symbols and search". In: *Communications of the ACM* 19.3 (Mar. 1976), pp. 113–126. ISSN: 0001-0782. DOI: 10.1145/360018.360022. URL: https://doi.org/10.1145/360018.360022 (visited on 06/19/2021).

[38] Frederick J. Newmeyer. *The politics of linguistics.* English. 1986. ISBN: 978-0-226-57720-3. URL: https://dialnet.unirioja.es/servlet/libro?codigo=605873 (visited on 06/18/2021).

[39] Rafael Núñez et al. "What happened to cognitive science?" English. In: *Nature Human Behaviour* 3.8 (Aug. 2019). tex.copyright: 2019 Springer Nature Limited, pp. 782–791. ISSN: 2397-3374. DOI: 10.1038/s41562-019-0626-2. URL: https://www.nature.com/articles/s41562-019-0626-2 (visited on 06/13/2021).

[40] H. Nyquist. "Certain Topics in Telegraph Transmission Theory". In: *Transactions of the American Institute of Electrical Engineers* 47.2 (Apr. 1928), pp. 617–644. ISSN: 2330-9431. DOI: 10.1109/T-AIEE.1928.5055024.

[41] Judea Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.* San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1988. ISBN: 978-1-55860-479-7.

[42] Lloyd Peterson and Margaret Jean Peterson. "Short-term retention of individual verbal items". In: *Journal of Experimental Psychology* 58.3 (1959), pp. 193–198. ISSN: 0022-1015(Print). DOI: 10.1037/h0049234.

[43] The MIT Press. *The Soar Cognitive Architecture — The MIT Press.* en. Publisher: The MIT Press. 2012. URL: https://mitpress.mit.edu/books/soar-cognitive-architecture (visited on 06/24/2021).

[44] Arturo Rosenblueth, Norbert Wiener, and Julian Bigelow. "Behavior, Purpose and Teleology". In: *Philosophy of Science* 10.1 (Jan. 1943), pp. 18–24. ISSN: 0031-8248. DOI: 10.1086/286788. URL: https://www.journals.uchicago.edu/doi/abs/10.1086/286788 (visited on 06/16/2021).

[45] Stuart Russell and Peter Norvig. "Artificial intelligence: a modern approach". In: (2002).

[46] R. Schank and C. Riesbeck. "Inside Computer Understanding". English. In: *undefined* (1981). URL: https://www.semanticscholar.org/paper/Inside-Computer-Understanding-Schank-Riesbeck/41a18aead56314c2eedd5efb5c7beef6c5212fbb (visited on 06/21/2021).

[47] Roger C. Schank and Robert P. Abelson. *Scripts, plans, goals and understanding: An inquiry into human knowledge structures*. Scripts, plans, goals and understanding: An inquiry into human knowledge structures. Oxford, England: Lawrence Erlbaum, 1977. ISBN: 978-0-470-99033-9.

[48] Noel Sheehy and Antony J. Chapman. *Cognitive Science*. Sept. 1995.

[49] B. F. Skinner. *About behaviorism*. About behaviorism. Oxford, England: Alfred A. Knopf, 1974. ISBN: 978-0-394-49201-8.

[50] Burrhus Frederic Skinner. *Verbal behavior*. New York: Appleton-Century-Crofts, 1957.

[51] George Sperling. "The information available in brief visual presentations". In: *Psychological Monographs: General and Applied*. 1960, pp. 1–29. URL: http://www.cogsci.uci.edu/~whipl/staff/sperling/PDFs/Sperling_PsychMonogr_1960.pdf (visited on 06/20/2021).

[52] M. D. Steer. "Cybernetics: Circular Causal and Feedback Mechanisms in Biological and Social Systems. Transactions of the Seventh Conference, March 23-24, 1950, New York. Heinz von Foerster, Ed. New York: Josiah Macy, Jr. Foundation, 1951. 251 pp. $3.50". English. In: *Science* 115.2978 (Jan. 1952). tex.copyright: Copyright © 1952 by the American Association for the Advancement of Science, pp. 100–100. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.115.2978.100. URL: https://science.sciencemag.org/content/115/2978/100.1 (visited on 06/16/2021).

[53] Danny D. Steinberg, Hiroshi Nagata, and David P. Aline. *Psycholinguistics: Language, Mind and World*. English. Routledge, Oct. 2013. ISBN: 978-1-317-90056-6.

[54] Edward C. Tolman. "Cognitive maps in rats and men". In: *Psychological Review* 55.4 (1948), pp. 189–208. ISSN: 1939-1471(Electronic),0033-295X(Print). DOI: 10.1037/h0061626.

[55] A. M. Turing. "On Computable Numbers, with an Application to the Entscheidungsproblem". English. In: *Proceedings of the London Mathematical Society* s2-42.1 (1937). tex.copyright: © 1937 London Mathematical Society, pp. 230–265. ISSN: 1460-244X. DOI: `10.1112/plms/s2-42.1.230`. URL: `https://londmathsoc.onlinelibrary.wiley.com/doi/abs/10.1112/plms/s2-42.1.230` (visited on 06/16/2021).

[56] © Stanford University, Stanford, and California 94305. *Why Computers Can't Outthink the Experts*. English. 1984. URL: `https://exhibits.stanford.edu/feigenbaum/catalog/nr990gh3548` (visited on 06/23/2021).

[57] Johan Wagemans et al. "A Century of Gestalt Psychology in Visual Perception I. Perceptual Grouping and Figure-Ground Organization". In: *Psychological bulletin* 138.6 (Nov. 2012). tex.pmcid: PMC3482144, pp. 1172–1217. ISSN: 0033-2909. DOI: `10.1037/a0029333`. URL: `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3482144/` (visited on 06/17/2021).

[58] J. B. Watson. "The Unverbalized in Human Behavior". In: *Psychological Review* 31.4 (1924), pp. 273–280. ISSN: 1939-1471(Electronic),0033-295X(Print). DOI: `10.1037/h0071569`.

[59] Nancy C. Waugh and Donald A. Norman. "Primary memory". In: *Psychological Review* 72.2 (1965), pp. 89–104. ISSN: 1939-1471(Electronic),0033-295X(Print). DOI: `10.1037/h0021797`.

[60] Joseph Weizenbaum. *Computer power and human reason: From judgment to calculation*. Jan. 1976.

[61] Bernard Widrow and Marcian E. Hoff. "Associative Storage and Retrieval of Digital Information in Networks of Adaptive "Neurons"". English. In: *Biological Prototypes and Synthetic Systems: Volume 1 Proceedings of the Second Annual Bionics Symposium sponsored by Cornell University and the General Electric Company, Advanced Electronics Center, held at Cornell University, August 30–September 1, 1961*. Ed. by Eugene E. Bernard and Morley R. Kare. Boston, MA: Springer US, 1962, pp. 160–160. ISBN: 978-1-4684-1716-6. DOI: `10.1007/978-1-4684-1716-6_25`. URL: `https://doi.org/10.1007/978-1-4684-1716-6_25` (visited on 06/19/2021).

[62] Norbert Wiener. "Cybernetics : Control and Communication in the Animal and the Machine –2nd. ed". In: *212 p. Cambridge, Mass.: The MIT press, 1961. includes index*. CUMINCAD, 1961. URL: `http://cumincad.scix.net/cgi-bin/works/Show&_id=caadria2010_000&sort=DEFAULT&search=series:caadria/Show?_id=4e2e&sort=DEFAULT&search=%2Fseries%3A%22CADline%22&hits=808` (visited on 06/16/2021).

[63] Robert Wilensky. "Understanding goal-based stories." phd. USA: Yale University, 1978.

[64] Terry Winograd. "Procedures as a Representation for Data in a Computer Program for Understanding Natural Language". English. In: (Jan. 1971). URL: `https://dspace.mit.edu/handle/1721.1/7095` (visited on 06/21/2021).