DOI:10.1145/3593007

Thomas Haigh

Historical Reflections

Conjoined Twins: Artificial Intelligence and the Invention of Computer Science

How artificial intelligence and computer science grew up together.

HYPE AND HANDWRINGING concerning artificial intelligence (AI) abound. Technologies for face recognition, automatic transcription, machine translation, the generation of text and images, and image tagging have been deployed on an unprecedented scale and work with startling accuracy. Optimists believe the promises of self-driving cars and humanoid robots; pessimists worry about mass unemployment and human obsolescence; critics call for ethical controls on the use of AI and decry its role in the propagation of racism.

Right now, AI refers almost exclusively to neural network systems able to train themselves against large datasets to successfully recognize or generate patterns. That is a profound break with the approaches behind previous waves of AI hype. In this column, the first in a series, I will be looking back to the origins of AI in the 1950s and 1960s. Artificial intelligence was born out of the promise that computers would quickly outstrip the ability of human minds to reason and the claim that building artificial minds would shed light on human cognition. Although the deep learning techniques underlying today's systems are relatively new, artificial intelligence was a key component in the emergence of computer science as an academic discipline.

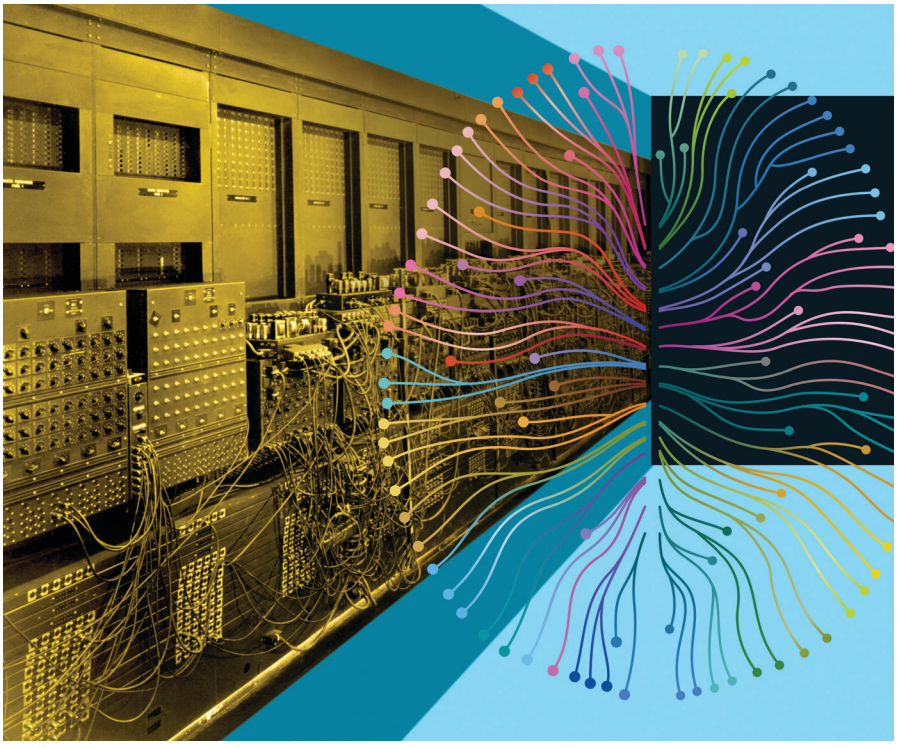

IMAGE BY ANDRIJ BORYS ASSOCIATES. USING PHOTO FROM EVERETT COLLECTION/SHUTTERSTOCK

Giant Cybernetic Brains

More than commonly realized, the modern computer was itself viewed as a thinking machine within the rich stew of what was about to be branded as *cybernetics*. The basic architecture of modern computers, centered on the retrieval of numerically coded instructions from an addressable high-speed store, was first described in John von Neumann's "First Draft of a Report on the EDVAC." As von Neumann wrote this material in early 1945 he was enmeshed in discussions with a group attempting to charter a "Teleological Society" to explore the radical idea that organisms and machines were substantively equivalent. Von Neumann described the building blocks of digital computer logic, later known as *gates*, with the biological term *neurons*. This was inspired by the work of Warren McCulloch and Walter Pitts, who had asserted that real neurons worked

as binary switches and so were functionally equivalent to Turing machines and to statements expressed in formal logic. Taking the biological metaphor further, von Neumann called the constituent parts of his planned computer *organs* and its internal storage unit *memory*.

The first popular book to describe computer technology, Edmund Berkeley's *Giant Brains: Or, Machines That Think*, doubled down on this perspective.[2] Berkeley, a former insurance executive, was the most active participant in the newly formed Association for Computing Machinery. He heralded the new machines as information processing devices whose capabilities would quickly match, and eventually outstrip, those of the human brain. He forecast applications for computers in areas that would later be thought of as part of AI, including machine translation, speech recognition, and automated psychological therapy.

Within this frame, however, it was not necessary for a computer to compose a sonnet or win a game of chess to be counted as a thinking machine. If brains and telephone switches, the topic of early work by Claude Shannon, were both equivalent to Boolean algebra then the electronic switching that took place when a computer executed a program was fundamentally the same as the neural switching that took place as humans thought. They differed in complexity, but not in essence. For Berkeley, the defining characteristic of a mechanical brain was simply that it "handles information, transfers information from one part of the machine to another, and has a flexible control over the sequence of its operations."[1,2] Thus readers could build their own thinking machine by following his instructions to build Simon, a simple device consisting of two light bulbs, some switching relays, and two homemade paper tape readers.

Berkeley's conception of thought seems unnatural to me, and probably to you, but that is only because his framing of number-crunching computers as artificial brains did not stick. We see nothing odd in talking about computer memory, which at the time was an equally jarring appropriation of biological terminology. The Teleological Society never happened, but its proponents successfully rebranded the effort as *cybernetics*. Both names referred the shared ability of biological and mechanical systems to steer themselves toward a goal.

As cybernetics developed through series of conferences sponsored by the Macy Foundation (1946–1953) the relationship of brains to computers and the ability of machines to learn from their environments was a central topic of debate.[7] British cyberneticists were particularly likely to describe simple lab-built learning mechanisms as brains, to the extent that philosopher Andrew Pickering called his history of that movement *The Cybernetic Brain*.[11] von Neumann himself avoided calling computers brains despite his appropriation of other biological terms. Succumbing to terminal cancer in 1955, he devoted the last of his intellectual energy to a lecture series on the relationship between the two. While his unfinished text still talks about computers having organs and memories, it asserted that biological neurons used a mix of analog and digital logic mechanisms and so were not truly equivalent to switches.[14,a]

Inventing AI

Strictly interpreted, the history of AI begins in 1955 when the term *artificial intelligence* appeared for the first time. John McCarthy, a newly arrived assistant professor at Dartmouth College, wrote a proposal to host a "Summer Research Project on Artificial Intelligence" the following year. Researchers including Claude Shannon and Marvin Minsky would spend up to eight weeks at Dartmouth, during which "an attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves."[b] The effort was stimulated by the arrival of the programmable electronic computer, able to automatically manipulate coded symbols with unprecedented speed and flexibility.

Six months before arriving at Dartmouth, Herb Simon had famously announced to his students that "Over Christmas, Al Newell and I invented a thinking machine."[12] He meant they had successfully simulated, on paper, the operation of a program to automate the work of logical deduction and thus demonstrated that computers could carry out the work of minds. The Logic Theorist program treated reasoning as a search tree whose branches represented sequences of possible operations. To keep its size manageable, the system pruned unpromising paths by applying *heuristic* rules.[c]

According to Berkeley, thinking machines were already a reality. Herb Simon himself was first exposed to computers when he read Berkeley's book and built a small digital logic kit sold mail-order as a "genuine brain machine."[12,d] In contrast, Alan Turing's 1950 paper "Computing Machinery and Intelligence" proposed the ability to successfully imitate a human in written conversation as a more tractable replacement for the question

The modern computer was itself viewed as a thinking machine within the rich stew of what was about to be branded as "cybernetics."

a The evolution of von Neumann's conceptualization of brains and automata is explored in detail in Aspray, W. *John von Neumann and the Origins of Modern Computing*. MIT Press, Cambridge, MA, 1990.

b See https://bit.ly/2GNE58J. On the Dartmouth event, see chapter 5 of Pamela McCorduck, *Machines Who Think*. A.K. Peters, Natick, MA, 2004.

c The Logic Theorist, originally called the Logic Theory Machine, is described, with a particular focus on its implementation for the RAND Corporation's JOHNNIAC computer, in Stephanie Dick, "Of Models and Machines: Implementing Bounded Rationality," *Isis 106*, 3 (2015). Dick suggests the linked list, a fundamental data type, was first developed in this project.

d The description of Berkeley's GENIAC comes from an advertisement, at https://bit.ly/3GSeygA

"can a machine think?"[13] Turing's paper described in detail the capabilities of powerful digital computers but asserted that another 50 years of progress would be needed to make the idea of thinking machines acceptable in "general educated opinion." Paradoxically, then, AI was defined not by the first claims that computers could think like brains but the first assertions that only specific, hard to program, activities like learning, forming concepts, manipulating language, or demonstrating creativity should be counted as thought. This was a retreat from the radical idea that every digital computer was already thinking.

Creating Computer Science

AI was a description of a general goal that left a lot of space for different topics and approaches. It began as a brand used by researchers at a small set of elite institutions to tie their work to lofty goals, win research support, and bolster their position within the emerging field of computer science. AI subsumed several preexisting research streams. Shannon was already working with McCarthy on a collection of papers on automata theory that covered much of the same ground. Research on neural networks was already established as part of cybernetics. AI was not the only such brand, for example, 1958 symposium that catalyzed British research in the area explored the *Mechanization of Thought Processes*, but it soon became dominant.

During the 1960s, AI developed under the auspices of computer science programs and affiliated laboratories. AI consolidated intellectually as its leaders drew firm boundaries to exclude some of the topics and approaches inherited from earlier identities such as cybernetics and automata studies. From that time onward, at least until the recent boom of AI within firms such as Google and Facebook, AI researchers were employed primarily within university computer science departments and most of those who entered the field received graduate training in computer science.

To say AI grew up within computer science is not to deny its interdisciplinarity. Computer science itself was, by necessity, an interdisciplinary field in its first decades. The first computer science programs were established by academics who had become fascinated by computer technology while laboring in campus computer centers or in departments of mathematics or electrical engineering. In such settings the study of computing could only be justified instrumentally, to support research in established disciplines. The early faculty members of computer science programs were thus intellectual immigrants with training in other fields. To study computing as an end in itself, they had to formulate an interdisciplinary synthesis out of the human and intellectual materials available. Each computer science program hammered together an assortment of coursework around bodies of craft knowledge that had grown up around the new machines: compiler creation, numerical analysis, computer architecture and engineering, and so on. AI loomed large as one of the most prestigious of these subfields.

To say AI grew up within computer science is not to deny its interdisciplinarity.

The first ACM A.M. Turing Award was awarded in 1966, just as computer science programs were beginning to graduate doctoral candidates. Both were important steps in defining and demarcating the new discipline. During the first decade of the awards program, 11 men were honored, among them Marvin Minsky, John McCarthy, Allen Newell, and Herb Simon. With four recipients, AI had beaten out all other areas of computer science (though like every Turing awardee prior to Thompson and Richie in 1983, none of the four had earned a degree in computer science).

The four early winners founded all three leading centers for AI research and graduate education in the U.S. Newell and Simon were based in the business school of what became Carnegie Mellon University. They founded a lab to research AI and were instrumental in the 1965 establishment of a computer science department, which retains a particular focus on AI and robotics. In 1958, McCarthy, who left Dartmouth after a single year, and Minsky jointly established what eventually became the MIT AI Lab. It was one of the central constituents of MIT's computer science community, which was spread over several projects and labs rather than being consolidated in a conventional department. By 1962, McCarthy had moved again, to Stanford, where he established an AI project that found a physical home as the Stanford AI Lab (SAIL) in 1966. This process was contemporaneous with Stanford's establishment of a computer science department in 1965.[10] Theirs is a powerful legacy: MIT, Stanford, and Carnegie Mellon are still ranked as the top three academic programs in the U.S., not just for AI but for computer science itself.

Demarcating a Field

Even insider histories depict AI not as a coherent whole but as a set of techniques and approaches that have risen and fallen in credibility over time. The work of the four award winners embodied the most important first-generation approaches to AI, at least in the eyes of the distinguished computer scientists who selected the winners. Each had attended McCarthy's 1956 workshop at Dartmouth. Each had come, at least by the time of the awards, to believe that techniques based on the logical manipulation of symbols would be applicable across many problem areas and shed light on human thought patterns.

As Newell himself told the story, he and his fellow awardees had defined AI around *symbolic* approaches based on the manipulation and processing of encoded symbols, over approaches such as neural networks inspired more directly by cybernetic conceptions of brain function. Newell called the rival approach *continuous*, in a nod to analog control systems, but today it is more often called *connectionist* because of the focus on manipulating patterns of connections between simulated neurons. Echoing the search trees he used to model problem solving, Newell interpreted the

history of his field as a succession of binary choices in which, for example, "the continuous-system folks ended up in electrical-engineering departments; the AI folk ended up in computer-science departments."[9]

Marvin Minsky had made early experiments with neural networks but soon turned to the symbolic representation of knowledge, most famously with his "theory of frames." The shift was celebrated in a book Minsky coauthored to highlight the limitations of a simple neural network model, championed by Frank Rosenblatt, that had been popular with early researchers.[8] Analysis of citation patterns confirms connectionist work was more prominent through the 1950s and into the early 1960s but was cited much less often than symbolic work during the 1970s and 1980s.[3]

Edward Feigenbaum, a student of Simon's who would later receive his own Turing award, co-edited *Computers and Thought*, an influential collection of papers intended to provide students with a convenient bundle of the interdisciplinary contributions from which the new field would be assembled.[6] The book's sections defined the central topics of the new field as game playing, theorem proving, answering questions posed in natural language, computer vision ("pattern recognition"), automatic learning, and decision making. One of its central contributions was an extensive bibliography compiled by Minsky.[e] In the absence of textbooks, their editorial decisions must have exerted an outsized influence on what appeared in the AI courses offered by less experienced instructors and at lower-tier universities.

Back at the Dartmouth event in 1956, Newell had claimed to have already solved many of the problems others were still considering. He had not exactly solved them, but the reduction of intelligence to formal reasoning and of reasoning to search did come to dominate AI. As the first

Because AI researchers were trying to program things that pushed the limits of computing they needed the biggest and most expensive computers available.

lecture of an introductory AI course delivered at Oxford University in 2000 concluded: "We have indicated that two important ingredients in AI are search and knowledge representation. The two appear, often inextricably entwined, in one guise or another in every AI problem."[f] The techniques used by Simon and Newell worked startlingly well in proving textbook logic theorems, but disappointed when they were extended (in a sequel, the General Problem Solver) and applied to other situations.

Simon's belief that a computer was thinking when it proved a theorem, but not thinking when is solved a differential equation numerically or processed a payroll, was equally significant in shaping the field. AI researchers defined their shared task as figuring out how to program tasks that computers were bad at but which humans, or at least humans of the kind viewed by Newell and his friends as intelligent, were good at. The techniques they invented would, hopefully, prove to be analogs of the processes at work within the human mind. Influential projects set specific goals that could be presented as a step toward developing such capabilities, like understanding speech inputs or navigating a robot through a space. The logic theorist was the first in a series of celebrated AI demonstration systems that lingered in classrooms and textbooks for generations as apparent proofs of

e The significance of this bibliography is explored in Jonathan Nigel Ross Penn. *Inventing Intelligence: On the History of Complex Information Processing and Artificial Intelligence in the United States in the Mid-Twentieth Century.* Ph.D. dissertation. University of Cambridge, 2020, 186–190.

f See https://bit.ly/3AQhLK1

concept for techniques that worked only in very specific situations. The same dialogue transcripts and anecdotes appeared again and again, as a substitute for generalized and robust methods that could be applied to real-world situations.

Some projects of the 1960s, most notably robot-building efforts at the University of Edinburgh and the Stanford Research Institute, did try to tie together capabilities such as vision, planning, and natural language. But most researchers dug deep into specific techniques, to the extent the designation of work as falling into AI or into some other area of computing could have more to do with branding, institutional affiliations, and accidents of history than any inherent relationship to cognition.

Likewise, much AI work focused on optimizing search, which mathematically was not so different from much other work carried out on algorithms or optimization. For example, the SRI team programming Shakey the robot to navigate needed a way for it to plan its route. In 1968, they published a path-searching algorithm, A*, which quickly became one of the standard route-finding methods. Its three creators, Peter Hart, Nils Nilsson, and Bertram Raphael became prominent figures in the AI community. Yet other algorithms for the same job, including an earlier and closely related algorithm created by Edsger Dijkstra, were developed by researchers interesting in optimizing systems with no connection to AI. In a similar way George Danzig's work on linear programming, which became a central technique of operations research, was never branded as AI despite having a broad usefulness when optimizing systems too complex to search in full.

Early AI enthusiasts hoped techniques developed in programming computers to do one thing traditionally associated with intelligence, such as playing a board game, could be used to automate many other intellectual processes. This turned out not to be true. Stephanie Dick has shown that later generations of automatic theorem provers were far more capable than Newell and Simon's original but used different techniques that made no claim to mirror human cognition and were not branded as AI.[4] Likewise, the first chess programs could barely complete a game, but thanks to decades of programming effort and hardware improvements programs running on smartphones can now beat human grandmasters. Yet the methods they use are alien to any processes that might plausibly be at work in human minds.[5] Rather than declare computers intellectually superior to us, we have collectively agreed that intelligence is not needed to play chess or plan a route.

These shifting boundaries parallel the founding premise of AI that computerized thought included logical proofs but not numerical mathematics: if automating thought meant programming a computer to do something that could not be achieved with existing methods then the domain of AI shrank every time conventional methods advanced.

The Legacy of Early AI

Because AI researchers were trying to program things that pushed the limits of computing they needed the biggest and most expensive computers available. Their approach emphasized tangible accomplishments visible in running code, which aligned AI closely with efforts to develop interactive computer systems, tools, and languages. The biggest contributions of AI to the development of computing came not replicating human thought but building infrastructure.

For example, John McCarthy had an influence on computing far beyond his technical contributions to logic programming and AI theory or his institution-building achievements. To support his projects McCarthy invented the concept of garbage collection, designed the widely used and influential Lisp programming language, and proposed the introduction of recursion into Algol from whence it spread into other procedural languages. McCarthy was an early proponent of timesharing as way of making interactive computer access feasible at a time when computers were enormously expensive. This led to crucial projects at MIT. McCarthy advocated for the "utility computing," the idea that computing power would, like electrical power, be most efficiently generated in huge central facilities serving thousands of users. That idea failed in the 1960s but has become a reality in the era of cloud computing.

Looking at that list of accomplishments it would be easy to argue that, despite the failure of early AI to achieve any of its primary goals, the incidental and infrastructural achievements of AI researchers represented a good return on money invested.[g] By the early 1970s, however, concern over broken promises was making military and governmental funders reluctant to continue their support for some of the highest-profile AI efforts. In subsequent columns, I will be following the story of AI into the 1970s and 1980s, looking particularly at sources of funding for AI work, a new emphasis on the representation of knowledge, and the rise of expert systems and knowledge based systems as alternatives to the tainted brand of artificial intelligence. C

g See https://bit.ly/3mUHPQG

References

1. Akera, A. Edmund Berkeley and the origins of ACM. *Commun. ACM 50*, 5 (May 2007).
2. Berkeley, E.C. *Giant Brains or Machines That Think*. John Wiley & Sons, NY, 1949.
3. Cardon, J.-P.C. and Mazieres, A. Neurons spike back: The invention of inductive machines and the artificial intelligence controversy. *Réseaux*, 211 (2018)
4. Dick, S.A. After Math: (Re)configuring Minds, Proof, and Computing in the Postwar United States. Ph.d. dissertation. Harvard University, 2015.
5. Ensmenger, N. Is chess the drosophila of artificial intelligence? A social history of an algorithm. *Social Studies of Science 42*, 1 (2011).
6. Feigenbaum, E.A. and Feldman, J. Eds. *Computers and Thought*. McGraw-Hill, NY, 1963.
7. Kline, R. *The Cybernetics Moment, Or Why We Call Our Age the Information Age*. Johns Hopkins University Press, 2015.
8. Minsky, M. and and Papert, S. *Perceptrons: An Introduction to Computational Geometry*. MIT Press, Cambridge, MA, 1969.
9. Newell, A. Intellectual issues in the history of artificial intelligence. In *The Study of Information: Interdisciplinary Messages*. F. Machlup and U. Mansfield, Eds. John Wiley and Sons, NY, 1983).
10. November, J. "Geroge Forsythe and the Creation of Computer Science As We Know It. In Communities of Computing: Computer Science and Society in the ACM, ed. Thomas J. Misa, Ed. Morgan & Claypool, 2017.
11. Pickering, A. *The Cybernetic Brain: Sketches of Another Future*. University of Chicago Press, Chicago, 2011.
12. Simon, H.A. *Models of My Life*. Basic Books, New York, 1991.
13. Turing, A. Computing machinery and intelligence. *Mind LIX*, 236 (1950).
14. von Neumann, J. *The Computer and the Brain*. Yale University Press, New Haven, CT (1958).

Thomas Haigh (thomas.haigh@gmail.com) is a professor of history at the University of Wisconsin—Milwaukee, WI, USA, and a Comenius visiting professor at Siegen University, Germany.

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)-Project-ID 262513311-SFB 1187 Media of Cooperation.