W.S. McCULLOCH AND THE FOUNDATIONS OF AI

Mallory Selfridge

In 1948 before the Philosophical Club at the University of Virginia, Warren McCulloch serenely proclaimed that "for the first time in the history of science we know how we know and hence are able to state it clearly. This is actually a great heresy," he told them. "We are about to conceive of the knower as a computing machine." (1) He knew that to his audience of philosophers, metaphysicians and students this idea probably came as a shock. By way of reassurance he sketched for them the history of epistemological problems that since the days of the Pre-Socratics have worried philosophers, and he described how, after centuries of development of classical physics, for all his knowledge of the world the physicist has not been able to account for himself as a knower.

To solve this dilemma mathematicians, metaphysicians and physicists in modern contrapuntal themes have composed the theoretical sciences of Communication, Cybernetics, and Artificial Intelligence (AI), and have brought forth the invention of computers whose real business is not with things but with signals. The nervous system, surpassing all known systems for conveying signals has at the same time been intensively studied for the last century, receiving transfusions of knowledge from all other sciences. McCulloch and his colleagues contributed important theoretical ideas to neurophysiology, many of which have become fundamental to AI.

"Mind is the circuit action to be expected of a structure like the nervous system" (2) said McCulloch. The idea that a machine might be able to extract ideas from a mass of sensory input led to the study of knowledge by way of experiments in programming computers to see, to plan, and to use language. This persuaded those who followed McCulloch to believe that intelligence can be 580 SELFRIDGE

understood and that artificial intelligence is, in fact, possible. McCulloch's formulation of the brain as a computing machine enabled scientists to test theories of knowledge by building machines that can have knowledge.

Beyond the general idea that knowledge can be precisely described, McCulloch was instrumental in providing early examples of how knowledge might actually be acquired. He suggested mechanisms which could recognize forms and participated in experiments that demonstrated that similar mechanisms exist in the real world; for instance in the frog's eye and brain. These mechanisms inspired proposals as to how a brain could be aware that it was looking at a triangle of a face, and experiments which disclosed how a frog knows that it is looking at a flying bug. McCulloch said that these were prototypes of mechanisms for having an idea, or formulating an interval hypothesis about the nature of the outside world. Indeed, he used the term "abduction" which Charles Saunders Peirce used to describe the process of forming hypotheses. This term is in use in AI today. Thus McCulloch's work helped to provide AI with precise models of what it means for a machine to have an idea, and hence with the beginnings of mechanical models for epistemology.

McCulloch made a second contribution to AI, more difficult to describe. It concerns his personal influence on the people who are now active in AI. When McCulloch came to MIT in the early 50's the impact of the computer on the study of intelligence was beginning to be felt. AI had been in existence as a field for a few years, and was beginning to be known by that name. McCulloch's ideas and enthusiasm were absorbed by people like Marvin Minsky and Seymour Papert, whose contributions to AI are well known.

Philosophically McCulloch maintained a humanistic attitude toward the mechanisms of intelligence. He was concerned with humanness and wanted the study of intelligence to increase our understanding of what being human means. This tradition of McCulloch's philosophy seems to have persisted in those who studied with him. McCulloch was one of the first to raise the question "What is Knowledge?" in the context of intelligent

machines. This is one of the chief concerns of AI. But somehow there is more: rather than restricting human potential to function at the level of the machine, he insisted on the possibility of raising machines to the level of humanity. For the multi-sidedness of intelligence to be understood he felt it imperative that the mechanical description of the mind include both a sense of what has already been described and a sense of what is yet to be discovered. "We must not confuse a computer model of a person with a person," I seem to hear him say. "Rather, we must let our best ideas remind us of what we do not yet know about ourselves."

McCulloch helped ask the important questions and provided examples of what the answers might look like. He pointed out directions for research and gave the discipline a sense of why these were the right questions and directions. His modern heresy in Virginia has become commonplace. In this sense Warren McCulloch was one of the founders of Artificial Intelligence.

REFERENCES

- 1. Embodiments of Mind, W.S. McCulloch, MIT Press, 1965, 1988.
- 2. Collected papers. C. Hartshorn. P. Weiss and Aburks, editors 8 vols. etc.
- 3. Pople, H.E. In the Mechanization of Abductive Logic...
- 4. Pople, H.E. The Formation of Composite Hypotheses in Diagnosis...