VIEWING HISTORY OF SCIENCE AS COMPILED HINDSIGHT

Lindley Darden
Committee on the History and Philosophy of Science
1131 Skinner Hall
University of Maryland
College Park, MD 20742

ABSTRACT

This paper argues that the history of science can be used as a source for constructing abstract theory types to aid in solving recurring problem types. Two theory types are discussed, selection theories and instructive theories, which aid in forming hypotheses to solve fitness problems. Providing cases from which to construct theory types is one of several ways in which the history of science can be viewed as "compiled hindsight."

SUMMARY OF LECTURE

History of the natural sciences is a rich source of information about past attempts by humans to deal with four of the major areas of research in Artificial Intelligence (AI): knowledge, reasoning, representation, and learning. (Although it should be noted that little work directly addresses the problem of representation of knowledge for computer manipulation, other than in mathematical models.)

Science has been one of the most successful knowledge gaining enterprises to date. One of the features that marks its history is that it has removed common sense misconceptions, often at the expense of producing counter-intuitive, though well-confirmed, theories. Much of Aristotle's world view was a common sense one: objects stop moving if no mover is pushing them; the earth is at the center of the universe and does not move; animal species perpetuate themselves eternally and act according to teleological goals; substances have essential and accidental properties. Science, or natural philosophy as it was long-called, has served to correct these and many other misconceptions. Thus, science is an important storehouse of knowledge about the natural world. Storing this knowledge and finding ways of making it readily accessible to AI systems is an important task.

But science is not only a body of knowledge. Science can also be characterized by its methodologies, which have themselves developed over time. Baconian induction, hypothetico-deductivism, comparative and experimental methods—these are a few of the reasoning methods developed by science as means of acquiring knowledge about the natural world. AI has begun a study of these and will continue to find the history of science a rich source of real world (as opposed to toy) examples on which to run these strategies.

Philosophers of science have studied the logic of confirmation of theories extensively. However, reasoning in the discovery of new theories has received comparatively little attention; in fact, philosophers, such as Karl Popper, have argued that no logic of discovery exists. If the goal is to find something weaker than a "logic," namely heuristics for discovery, then use of the history of science provides a means of doing so. By following the incorrect as well as the correct twists and turns along the pathways of science and devising strategies that are sufficient for producing the plausible laws and theories along those pathways, the means of making new discoveries (let me boldly assert) can be found.

A number of strategies used in science for learning or hypothesizing are beginning to receive attention in AI; for example, abduction, learning from a previously successful deduction, reasoning by analogy, reasoning by interrelations stronger than analogy, and reasoning by use of an abstract type of theory. The latter will be focus of attention here.

A powerful idea developed by AI is that of abstraction and instantiation. Except for devising mathematical models, little effort has been made to extract from the history of science abstract theory types that can be instantiated in new problem situations. The key thesis to be argued here is the following: <u>the history of science can serve as a source of compiled hindsight for constructing useful abstractions</u>. These will be abstractions of theory types and may prove useful in new problem situations.

A selection type theory is an example of a successful theory type in biology. Darwinian natural selection and clonal selection for antibody production are two instantiations. The abstraction of variation, interaction, and perpetuation of the most fit variant is an important theory type available only since the middle of the nineteenth century.

Historically, an alternative to selection theories for solving fitness problems that predated selection is the instructive type theory in which information is used to generate, not a random set of variants, but ones that already are fit. Neo-Lamarckian inheritance of acquired characteristics is an example of a form adapting and that form being perpetuated without random variation and selection. The template theory of antibody formation is another example of an instructive type theory. Neither of the instructive theories was confirmed historically. Since the disconfirmed theories provided a theory type, this case illustrates what historians of science have long been arguing: the study of disconfirmed, as well as confirmed, theories, the incorrect reasoning paths as well as the successful ones, is important.

The compiled hindsight from the study of selection and instructive theories is the following: find a current problem in which one thing is to change over time to fit another thing; devise good abstractions of the theory types for fitness problems, namely selection and instructive theories; instantiate those abstractions in that problem situation to provide solutions; test the solutions; note whether the selection theory is again successful.

In conclusion, a research program emerges: study the history of science to find recurring problem types and theory types; devise computationally useful abstractions for them; and build AI systems to use such compiled hindsight. Success with such a program of research would show the usefulness of interfield interactions between the history of science and artificial intelligence.

ACKNOWLEDGMENTS

Lindley Darden gratefully acknowledges the support of the University of Maryland Institute for Advanced Computer Studies for this work. The phrase "compiled hindsight" is taken from (Lenat, 1983).

REFERENCES

Darden, Lindley, and Nancy Maull. 1977. "Interfield Theories." Philosophy of Science 44:43-64.

Darden, Lindley. 1983. "Reasoning by Analogy in Scientific Theory Construction." Proceedings of the International Machine Learning Workshop, Monticello, Illinois, pp.32-40.

Darden, Lindley, and Rada, Roy. 1986. "Hypothesis Formation via Interrelations." Proceedings of Analogica '85, Rutgers University, New Brunswick, New Jersey, forthcoming.

Lenat, Douglas. 1983. "The Role of Heuristics in Learning by Discovery: Three Case Studies." Machine Learning. Edited by R. Michalski, J. Carbonell, and T. Mitchell. Palo Alto, California: Tioga Publishing Co. Pages 243-306.

Langley, Pat, Jan Zytkow, Gary Bradshaw, Herbert Simon. 1983. "Three Facets of Scientific Discovery." Proceedings of the Eighth International Joint Conference on Artificial Intelligence, Karlsruhe, West Germany, v.1. Pages 465-468.

Popper, Karl. 1965. The Logic of Scientific Discovery. New York: Harper Torchbooks.