Artificial Intelligence

Review of “Knowing Our World: An Artificial Intelligence Perspective”, by George F. Luger, Springer, 2021.

--Manuscript Draft--

Manuscript Number:	
Article Type:	Book Review
Section/Category:	Example Review Section
Corresponding Author:	Daniel G Schwartz Tallahassee, FL United States
First Author:	Daniel G Schwartz
Order of Authors:	Daniel G Schwartz
Suggested Reviewers:	
Opposed Reviewers:	

Review of "Knowing Our World: An Artificial Intelligence Perspective", by George F. Luger, Springer, 2021.

Contributed by Daniel G. Schwartz, Department of Computer Science, Florida State University, Tallahassee, Florida 32306, USA, schwartz@cs.fsu.edu.

This book is a "must read" for anyone that has an interest in Artificial Intelligence and epistemologically related issues. Besides offering new insights into the philosophical foundations of epistemology, it is a veritable encyclopedia covering the history and core aspects of contemporary AI and Cognitive Science.

Chapter 1 begins with a discussion of epistemology and its relevance to AI. Epistemology is the study of how people know their world. It can be described as the study of human understanding, knowledge, and meaning. This is relevant to AI in that "creating any program for a computer requires selecting symbols and program instructions, called *algorithms*, to 'capture' the task at hand." (p. 4) This task-driven selection of symbols and programming instructions is called an *epistemic commitment* or *stance*. In effect, the program embodies the understanding that the programmer holds with respect to the aspects of the world that the program is intended to simulate or represent.

Chapter 1 continues with a discussion of the general concept of computation, Turing machines, Post production systems, the role of high-level computing languages, issues of expressivity and precision of knowledge representations, the basic issues of state-space representations of problem domains, and solutions provided by state-space search, e.g., as in expert systems and computer games.

Chapter 2 delves into the historical development of ideas and philosophical themes that started with the early Greeks and led up to the invention of the computer. Three different lines of thought are identified: rationalist, empiricist, and pragmatist. The rationalist view is that knowledge of the world is properly represented via the formalism provide by mathematics. Some names associated with this perspective are Copernicus, Bacon, Schickard, Pascal, and Descartes. The empiricists added the observation that human knowledge begins with experience. It is only through interacting with the world that concepts are formed and understanding of the world is attained. Some names associated with this view are Hobbes, Locke, and Hume. The pragmatists added to this the observation that our understanding of the world and the conceptual constructs that we create to represent this understanding are driven by their having some designated purpose, and the validity of those constructs is determined by whether or how well the intended purpose is achieved. Some names associated with this view are Pierce, James, and Dewey.

There also is discussion of two philosophers that served to synthesize the rationalist and empiricist traditions, namely Spinoza and Kant. Immanuel Kant in particular is well-known for distinguishing between two types of knowledge, the *a priori* and the *a posteriori*. The former is regarded as innate and comprises the knowledge that enables one to construct the formalisms employed by mathematics and logic. The later comprises the experiential knowledge that results from interaction with the world. As such the two forms of knowledge are integrated into a unified system.

Chapter 2 continues with discussion of the mathematical foundations of computation and the contributions of Hobbes, Leibnitz, Babbage, Boole, Frege, Whitehead and Russell, and Tarski. These works laid the foundation for the von Neuman computer and offered formal logic as an important knowledge representation methodology. The chapter concludes with a discussion of the Turing test and the birth of AI.

Chapter 3 begins by recounting three AI success stories: IBM's Deep Blue chess playing program, IBM's Watson question answering system, and Google's Alpha Go. It then reviews some early works in AI: the Logic Theorist created by Newell and Simon, the geometry theorem prover by Gelernter, and Samuel's checker playing program. Samuel's work is interesting especially in that it was the first instance of what is now called "reinforcement learning". The chapter next revisits the famous 1956 Dartmouth summer workshop where the term "Artificial Intelligence" was coined, and the core objectives of AI research were articulated. These have withstood the test of time; they are as valid today as when they were first formulated. Chapter 3 concludes with some of the early controversies within the field of AI, the birth of Cognitive Science, and the symbolic versus connectionist themes in AI practice.

Chapters 4, 5, and 6 survey the three major themes of contemporary AI: (1) symbol-based knowledge representation and reasoning, (2) the connectionist approach, providing neural nets and ''deep learning", and (3) evolutionary computation with genetic algorithms. Chapter 4 might have benefited from a more detailed discussion of automated reasoning and its applications, such as theorem proving, robot planning, and program verification. Also deserving mention is fuzzy sets theory as providing a model of human reasoning with imprecise linguistic information and its widespread application in the form of fuzzy logic controllers. This notwithstanding, however, each of these chapters is a fairly comprehensive overview of its respective topic.

Chapter 7 presents Luger's contribution to epistemology. This serves to synthesize the three philosophical lines of thought discussed in the foregoing. As mentioned, Kant effectively unified the rationalist and empiricist views. However, Luger proposes that this additionally requires that the human agent, using *a priori* knowledge, develops a conceptual framework through which the *a posteriori* knowledge is received and interpreted. Such a conceptual framework would comprise a model or schema. In this regard, Luger references the work by cognitive scientists Maturana and Varela which purports that "human agents have no 'direct access' to anything, including their own epistemic dialect". In other words, the knowledge that is obtained, either through interaction with the world or through introspection, is obtained only by mediation through some conceptual construct that enables semantic interpretation of what is being perceived. In addition, Luger argues that such schemata are always constructed to fulfill some objective or serve some practical purpose. Thus, this folds in the pragmatist view.

The fact that the creation of such schemata is necessary for knowing is what Luger refers to by calling this new epistemology "constructivist". With regard to such conceptual constructions, moreover, Luger argues that these are necessarily subject to *inductive bias*. This term appears in the machine learning literature as referring "to a set of (explicit or implicit) assumptions made by a learning algorithm in order to perform induction, that is, to generalize a finite set of observation (training data) into a general model of the domain." (cf. [1]) But Luger uses this term in a more general sense of pertaining not only to machine learning but to all aspects of human and machine cognition. His definition is contained in the segment quoted in the following paragraph.

Another important feature of Luger's epistemology is that the conceptual constructions thus created are subject to ongoing revision by the perceiving agent. This view is inspired by the works of Piaget concerning the stages of a human child's cognitive development. In Luger's words (p. 178):

> "Perceived information, Kant's a posteriori knowledge, rarely fits precisely into our preconceived and a priori schemata. From this tension to comprehend and act, the schema-

> based biases a subject uses to organize experience are strengthened, modified, or replaced. Attempts to *accommodate* in the context of unsuccessful interactions with the environment drive a process of cognitive *equilibration*. The constructivist epistemology is one of cognitive evolution and continuous model refinement. An important consequence of constructivism is that the interpretation of any perception-based situation involves the imposition of the observer's unique concepts and categories on what is perceived. This constitutes an *inductive bias*. When Piaget first proposed a constructivist approach to a child's understanding of the world, he called it a *genetic epistemology*. When encountering new phenomena, the lack of a comfortable fit of current schemata to the world "as it is" creates a cognitive tension. This tension drives a process of schema revision. Schema revision, Piaget's *accommodation*, is the continued evolution of the agent's understanding towards *equilibration*.
>
> I contend that schema revision and continued movement toward equilibration is a genetic predisposition of an agent for an accommodation to the constraints of self, society, and the world."

Chapter 7 concludes with a list of five assumptions and eight conjectures comprising the foundation of Luger's epistemological stance. The first of these is (p. 182):

> "Assumption 1: Survival is the motivation or driving force of all living humans."

Thus, according to Luger, the ultimate underlying pragmatic objective of human behavior (and cognition) is human survival. The other assumptions and conjectures build upon this.

Another key aspect of Luger's vision is that "all knowledge of the world is best represented probabilistically". This is because, as just noted, "we have no direct access to anything, so our perceptions of phenomena can best be seen as samplings from distributions of phenomenon."

This leads in Chapter 8 into an analysis of Perl's theory of Bayesian networks. The argument is made that this theory provides an appropriate formal tool for representing and managing the uncertainty that accompanies all human perceptions. An important contribution of Perl's work in this respect is it's providing a computational, and tractable, model of material causation.

Chapter 9 concludes the book with a review of the key points made in earlier chapters; a discussion of model building in AI as a process of exploration, with numerous examples from work in robotics; a return to and more detailed discussion of the work of Piaget as it pertains to model revision and adaptation; prospects for the future of research in AI; and further discussion of meaning, truth, and the proposed foundation for a modern epistemology. The latter section includes a discussion of neo-pragmatism, which comprises a modern continuation of the earlier pragmatist views noted previously, and a discussion of how findings in neuroscience support Luger's understanding of human perception. Finally, the book concludes with a discussion of the differences between machines and humans, and a reiteration of the potential role for Bayesian networks as providing "*sufficient* models for aspects of human perception, knowledge, and performance." (p. 232) In the very last sentence (p. 236), the proposed epistemic stance is given as "an *active, pragmatic, model-revising realism*."

The book in this manner presents an interesting and compelling new vision of how humans know their world and how this vision is rooted in modern research in Cognitive Science and AI. For my own purposes, however, I would have appreciated some explanation of how this new epistemic stance might

pertain to the AI enterprise itself, i.e., what role it might play in the development of AI systems and artifacts. It seems reasonable that the idea of constructivism for human cognitive processes should apply equally to the construction of models employed by AI systems. Similarly, the notion of subjective bias surely applies to such models; the models are constructed with some pragmatic purpose; and they can undergo ongoing revision and improvement. While this may be implicit in Luger's discussion, it would seem that he has been remiss in not explicitly pointing this out.

One point I would take issue with is the assertion in Assumption 1 that survival is "the" driving force of human behavior. I would say rather that is it "a" driving force, but not necessarily the only one. In the 1940's, Abraham Maslow introduced a motivational model composed of a "hierarchy of needs", where the levels of the hierarchy bear titles such as "physiological", "safety", "belonging and love", "esteem", "self-actualization", and "transcendence". Note that "physiological", which occurs at the bottom of the hierarchy, corresponds to the notion of physical survival. Thus, whether one accepts the details of Maslow's hierarchy, it seems incontrovertible that humans have motivations beyond mere survival. Having said this, however, it does seem that Luger's epistemic stance still applies if one simply replaces the pragmatic objective of survival with any of the other higher-level needs. The effort to fulfill such needs would employ the same constructivist, model-revising cognitive processes.

[1] Hüllermeier, E., Fober, T., Mernberger, M. (2013). Inductive Bias. In: Dubitzky, W., Wolkenhauer, O., Cho, KH., Yokota, H. (eds) Encyclopedia of Systems Biology. Springer, New York, NY. https://doi.org/10.1007/978-1-4419-9863-7_927

[2] https://en.wikipedia.org/wiki/Maslow%27s_hierarchy_of_needs

Declaration of interests

☒The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: