

BOOK NOTE

MIND, MACHINE, & METAPHOR: AN ESSAY ON ARTIFICIAL INTELLIGENCE AND LEGAL REASONING

By Alexander E. Silverman¹

Bolder, Colorado: Westview Press. 1993.

Pp. 144. \$52.50. (hard).

"[Law] might . . . be imagined as a Gibsonian cyberspace, a lived consensual hallucination" (p. 88). In *Mind, Machine, and Metaphor: An Essay on Artificial Intelligence and Legal Reasoning*, Alexander E. Silverman presents a number of new ways of looking at the law, many of which are as radical as the one just cited. The author hopes to induce the reader to reexamine his or her preconceptions of law and the legal system. "At the most general level, our metaphor of law matters. What we mean by the rule of law, and our beliefs about how best to achieve it, may change if we allow ourselves to see the rule of law as something other than the law of rules" (p. 94). Accordingly, Silverman presents the reader with many new metaphors to law including law as connection machine, law as robot, and as mentioned above, law as cyberspace. While all of the metaphors are entertaining and some inspire insight into the legal process, others appear less inspiring and less accurate. *Mind, Machine, & Metaphor* takes a wide-ranging stroll through AI and the law, occasionally straying too far afield but nonetheless provoking a reexamination of one's conception of the law.

SUMMARY

Mind, Machine, & Metaphor begins its journey with a survey of current artificial intelligence (AI) techniques. The author notes that AI research has multiple goals. Some researchers seek to enable computers to perform specific tasks which had previously required human direction (p. 6). Others want to understand the mechanics of human intelligence

1. Associate with the firm of Townsend and Townsend Kourie and Crew, Palo Alto, California.

through modeling neurons on a computer (p. 6). Some want to create a system which has independent consciousness (p. 6). Often, researchers will be seeking some combination of these goals (p. 6). The techniques described in this chapter reappear in various forms throughout the book and provide a springboard for the author's many metaphors.

The most common AI systems are "classical," or "expert" systems. Such systems have a set of rules defining relationships between antecedent facts in a fixed manner, e.g., "if A and B are both true, then C is also true." By repeatedly applying its rules to a set of inputs, the system infers a number of facts from the initial data. The "knowledge" of a classical system is provided by a programmer who talks to experts in the field which the computer seeks to model. Early classical AI researchers had great success, but classical AI systems have proven too brittle for many real-world problems. Classical systems are also prone to catastrophic failure when their inputs deviate from the system designer's anticipated norms (p. 7). Classical systems seem well suited only to problems which require conscious deductive reasoning, which correspond to thought processes which take more than a second in human beings (p. 4).

In the past decade, researchers significantly expanded experimentation with "connectionist" AI machines, also known as neural networks (p. 3). Connectionist machines seek to model the biology of the brain by simulating "neurons" which are connected in such a way that the firing of one neuron may precipitate the firing of another if certain preconditions are met. The "intelligence" in connectionist AI machines is "learned," either through a training process in which the system is presented with "correct" input/output pairs, or as part of the machine's functioning, in which case the machine learns as it performs tasks. Their knowledge is stored in the connections between neurons and not in any localized area (p. 4). Complicated behavior can emerge from the connection of a large number of simple neurons. Connectionist systems have proven useful for tasks, such as pattern recognition, which occur in the human brain in time scales of less than a second (pp. 5, 7).

A third technique which may be seen as a subset of connectionist AI is "interactionist" AI. In interactionist systems, the real world is seen as a friend to the system rather than a foe. Interactionist machines always interact with an outside "world," whether that world is a video game,² or

2. One of the researchers taught a computer to play the video game Pengo (p. 28).

a laboratory.³ (pp. 25, 28). Advocates of interactionist approaches see consciousness as deriving as much from one's body and environment as from one's brain.

Silverman describes several legal AI systems which are under development or currently in use. He describes how classical systems have been used to analyze fact patterns in trade secret misappropriation cases (p. 19), and to predict damage awards in automobile accidents (p. 20). Silverman also notes a connectionist system for legal research which describes links between cases depending on the strength of their doctrinal similarity or difference and which learns from patterns of a specific attorney's usage of the system.⁴

In Chapter Three, the author uses Lon Fuller and H.L.A. Hart's hypothetical "No vehicles in the park" ordinance to highlight the limits of both classical AI and mechanical jurisprudence (p. 36).⁵ A hypothetical ordinance forbids vehicles from a public park. Silverman presents several challenges to the ordinance as written. Is a bicycle a vehicle? Are emergency vehicles exempt? What if a group of veterans decides to mount a tank in the park as a statue? How would you distinguish that monument from an abandoned car? (p. 36). The author would develop a program in a hypothetical classical AI language which attempts to determine if any vehicles are violating the ordinance.⁶ He shows how it is nearly impossible *ex ante* to define the "essence" of a vehicle which should be prohibited, and that most classical AI systems must rely on the programmer to provide ad-hoc rules to deal with special cases such as the emergency vehicles in this problem. The open texture⁷ of law prohibits a complete solution to the problem, demonstrating that classical AI systems will be unable to solve this problem and that mechanical jurisprudence must fail.

Just as legal scholars realized this inevitable failure of mechanical jurisprudence, and moved to a model which is more flexible and responsive to context and equity, the author claims that researchers in AI

3. Another interactionist group built a robot which roams the room and steals soda cans from desks (p. 25).

4. This system is called SCALIR and is being developed at U.C. San Diego (p. 20).

5. The author avoids the term "mechanical jurisprudence" because he (correctly) notes that the term carries pejorative connotations which would be unwarranted if the "machine" were more human (p. 35). I use the term here to mean law based on rigid, unbending rules.

6. In an appendix, the author presents a PROLOG version of the program (p. 111).

7. Open texture is a phrase commonly used in linguistic theory which indicates that a concept cannot be defined in a deterministic manner.

are now moving toward a new paradigm for artificial intelligence: one in which connectionist and interactionist techniques provide flexibility and stability to an AI system. Silverman discusses Kuhn's theory of "paradigm shifts," in which one's way of seeing the world is replaced by another (pp. 67-70). Kuhn developed his theory to describe scientific "revolutions" in which scientists in a given field collectively change their perspective. Silverman asserts that such shifts are currently occurring in both artificial intelligence and jurisprudence. In general, one cannot subscribe to two paradigms at the same time: they are characterized by "gestalt flips," (p. 69) as in the famous drawing which can be a pair of faces or a vase but not both at once.

After tracking parallel developments in legal theory and artificial intelligence, Silverman suggests that metaphors can lead to fresh perspectives on our legal structure. First, and most importantly, he introduces the mind-machine metaphor, a two-way correspondence which shows us how we think of our minds as computers, and also how we see our computers as having minds (p. 23). Indeed, connectionist AI systems seek to mimic the human brain with its large number of interconnected neurons (p. 6).

Next, the author notes that a connectionist AI machine can be approximated by a function: certain inputs (pictures, text, etc.) are matched with outputs (identification of the picture, analysis of the text, etc.) (p. 5). Similarly, the law matches case fact patterns (the input) with decisions (the output) (p. 80). The law's function may have multiple values, as where precedents conflict, and may have a steep slope where the law is uncertain. Legal theories model this law function, and they can be seen as "best-fit" approximations to the function (pp. 81-83).

Since law and AI can both be seen as functions and both are tools for solving problems or disputes, Silverman seeks an analogy between AI and the law (p. 80). In this metaphor, the legal actors (e.g., judges, attorneys and law professors) are nodes in a large neural network. The many channels of communication (e.g., case reporters, law reviews, legal instruction in the classroom) are links between the nodes. The law network thus constructed has its knowledge stored as much in its connections as in its nodes. It is not completely controlled by any one actor. Legal inertia and synergy between the Supreme Court and other courts limit the degree to which the system may be peremptorily modified, in addition to the constraints imposed by the Constitution. Doctrine in such a system builds from the lower courts: the distributed

nodes. With this analogy, the author captures and reconceives much of the received knowledge regarding the common law. Common law builds from specific cases, and develops traditions based on practice and not abstract theory. The common law is strong because it allows judges to build on past experience in the form of precedent while allowing development to fit changing times. By comparing law to an AI network, the author calls attention to the importance of the connections to the emergent properties of a legal system. Rather than law as an imposed set of rules (mechanical jurisprudence) or law as judge's whim, he sees law as a dance in which legal actors interact in a complex yet orderly manner (See pp. 39-41).

In the final chapter, Silverman returns to more humble ground. He discusses the role of judgment in law, and provides an intriguing explanation for seemingly arbitrary decisions in some cases. Humans by their nature are unable to completely describe the process of thinking (p. 99). Silverman provides an example: "What is your mother's maiden name? Now *how* did you recall that? 'It just came to me.'" (p. 101). Any attempt to explain further would quickly reduce to absurdity. The author asserts that we are unable to explain our actions in formal terms when they have their bases in thought processes which occur on a time span of less than a second (p. 101). Just as one cannot completely describe a connectionist AI machine without printing out the entire contents of the nodes and connections, a person cannot describe decisions which have their basis in low-level knowledge. "You just know" (p. 101). Legal rules may help to *justify* such decisions, but they will not be the actual bases for them (p. 102). In contrast, if a judgment is based on thought processes which take more than a second, it is quite amenable to description in formal terms. Not incidentally, classical AI systems excel at making these types of judgments.

Through this process, the author shows us that we cannot expect judges to explain all their decisions completely.⁸ Some decisions will inevitably be based on intuition. Silverman limits the broad scope of this claim by noting that judges will often change their initial hunches after they attempt to explain them in formal terms. Rules do matter, but not as much as classical legal theory assumes.

8. The author does not mention this, but this is likely to be the basis for the "clearly erroneous" standard of review of lower court fact findings.

ANALYSIS OF STYLE

While *Mind, Machine, & Metaphor* inspires the reader to reconsider his or her conception of the legal process, the book is weakened by editing problems. The author overextends his metaphors, presents irrelevant or poorly constructed examples, and spends too much time developing peripheral argument.

Silverman often describes a well-designed metaphor and then spends several pages creating superfluous parallels. He notes that law and society shape each other and that when law changes too fast, societal upheaval results (e.g., immediately after *Brown v. Board of Education*⁹ society was thrown into turmoil after the sudden, forced integration of the schools). Silverman describes this process as “[l]aw and society resonat[ing] like a pair of damped, driven, coupled mechanical oscillators” (p. 77).¹⁰ In another overextended metaphor, he compares an AI system’s output function to a quilt and then extends the metaphor to the law:

The grid is the outline of squares; the threads within each square create a dense, richly colored texture; the whole, though composed of many disparate patches, embodies a coherent design Imagine next the law network function projected as a virtual reality The virtual landscape might even be imagined as a Gibsonian cyberspace, a lived consensual hallucination (replete with rogue AI denizens?) in which the law itself is the stuff of both intelligences and environment. (p. 88)

Silverman appears to be rationalizing his metaphor rather than developing a useful line of comparison. Furthermore, law does not seem much like a “lived consensual hallucination,” except perhaps to a first year associate at a Wall Street law firm (p. 88).

The author also provides some unnecessary examples which detract

9. 347 U.S. 483 (1954).

10. Indeed, even if one accepts that one needs a complete description for this analogy to be valuable, it is not necessary that the oscillator be mechanical. Electronic oscillators obey this same behavior.

from his otherwise, forceful metaphors. For example, during his discussion of vehicles in the park in which he explores the limits of classical AI, he describes an interactionist robot's implementation of the hypothetical ordinance. He posits that an interactionist robot would place barricades at the entrance of the park to block vehicles and that all vehicles which got past the barricades would not be considered vehicles for the purposes of the ordinance. Police in the park could catch all "vehicles" which managed to circumvent the barricades. This, of course, is cheating. The classical AI system was not allowed the option of failing to catch vehicles. The author specifically disparaged classical systems for their failure in special cases.

Later in the book, the author again invokes cyberspace:

[If the mental world is viewed as cyberspace, theory takes on the added dimension of active intelligence. Not only is the theory-world populated by active human agents, but theory itself becomes one or more active agents. The "intelligence" of these agents may be seen in their power to create illusions in human beings and to affect the larger world. (p. 66)

Is a virus an active intelligence because it infects others and may cause certain people to have fevered hallucinations? Silverman consistently explains that the emergent behavior from a connectionist system depends on the interaction of connections (theories) and nodes (people). How is it that in this case, the connections take on an intelligence of their own? In what does the theory's intelligence reside?

Finally, the book is hampered by its extensive discussion of peripheral issues. While the survey of current AI techniques is necessary to understand the later metaphors, the detailed exposition of the vehicle in the park program seems unnecessary. Also, the author devotes one sixth of the book to a discussion of Kuhnian paradigm shifts (pp. 59-76). Much of this discussion is an abstract defense of Kuhn's philosophy which seems irrelevant to the book's thesis (pp. 67-73). Indeed, it seems that he is only citing Kuhn for the rather unremarkable proposition that one must recognize "the contingent nature of what seems at times to be objective truth" and "the tendency of a dominant paradigm to become entrenched" (p. 74). In reading this section, one loses sight of both AI and law, and becomes mired in a discussion of Kuhn. Such excessive

peripheral discussions make the thesis of the book hard to follow.

STRENGTH OF THE ARGUMENT

Putting these criticisms aside, the book has some intriguing suggestions. The author cogently argues that judges should explicitly pay attention to their intuition because intuition often reflects a valuable body of knowledge which could never be set forth in concrete form.

The author also notes that attempting to create AI systems which model or predict the law tells us something about both AI and law:

[T]he experience of organizing a body of law into a form readily accessible to a computer may illuminate previously hidden aspects of the law's structure. Areas of law where "further definition may be appropriate or necessary" become clear Legal AI systems may provide a convenient mechanism for testing intersections of proposed and current law. (p. 21)

The metaphors of law as connectionist AI machine have a similar flavor to traditional descriptions of the common law, thus analyzing how AI systems fail in the laboratory may give us insight into why legal regimes fail in the real world. His argument explains the failure of mechanical jurisprudence in the same terms as classical AI: We do not live in a world which can be accurately modeled with rigid rules. Should we discover general principles about new techniques to enhance the effectiveness of artificial intelligence networks, the metaphor might allow us to apply them to the legal system. Similarly, the legal system presents an example of a network which has evolved in the real world and which seems to work fairly well. Perhaps AI researchers who read this book will turn to the "law network" for insight.

Silverman does not advocate any concrete modifications to the legal system, but rather proposes a set of provocative metaphors and leaves them there for the reader's perusal. The reader is left to follow the author's instruction with respect to the ideas presented: "*Play with them. See what comes up*" (p. 109).

Jeffrey D. Osterman