Maintaining a Balanced Perspective on Artificial Intelligence
by Paul E. Silvey
The MITRE Corporation
2021-02-25

The quest to build Artificial Intelligence (AI), at least narrowly defined as computer programs that can mimic human cognitive capabilities sufficiently well as to solve practical problems for us, has again captured the business world's attention. Government and private investment dollars are flowing into this sector, motivated by Machine Learning (ML) successes in performing object or event recognition and classification from perceptual data, at levels that can match and sometimes surpass the abilities of humans. The availability and affordability of massive amounts of data and computational resources, along with algorithmic innovations including large multi-layer artificial neural networks, has created a kind of gold rush phenomenon in the last ten years.

These impressive classifiers are known as Deep Learning systems, but they represent only a small subset of the technologies that have been researched, developed, and labeled as AI over its sixty-plus year history as a formal discipline. For example, there are many algorithmic approaches for building human decision aids, many of which have traditionally relied on knowledge programming vs. data-driven machine learning. Anyone interested in seeking productivity gains from AI should be aware of these differences in approaches, their suitability to one's specific problem, and their relative technological maturity.

Early efforts in AI focused on the challenge of hand-crafting knowledge in such a way that a computer program could reason its way to a solution from some problem starting state. This reasoning process involved sequential consideration (search) of possible ways to do things, or possible intermediate beliefs to establish and leverage. It was about problem solving from a position of expertise, without any automation of the learning process needed to acquire such expertise in the first place. One of the hallmarks of this kind of AI was the desire of their programmers to express knowledge in human-understandable forms, which was propositional and logical, that is, made up of symbols composed into linguistic forms that directly modeled rules of inference. The meanings of these symbols came in large part from humans interpreting them, and not from a deep understanding or grounding of their meaning by the computer itself. While it was successful in many problem domains, this Knowledge-Engineered kind of AI ultimately was considered too expensive and brittle for many candidate application areas.

After more than a decade of limited fanfare and attention to the field, the rise of ML shifted the emphasis dramatically, so much so that many of the lessons learned from the early days of AI are no longer even taught to students. The essence of this ML paradigm shift was from problem solving given appropriate knowledge, to using data and experience to learn how to recognize when a particular response or solution was applicable. Being data-driven also led these systems to favor probabilistic and statistical representations of their learned knowledge. Supervised Learning systems can be superb recognizers when fully trained, and their most common form is to work as simple question and answer systems. Show them a test case and they tell you how

to label it, based on many training examples of similar cases with correct category labels provided. But they can also be used to solve problems that involve multiple and sequential decision making, for example, as one does when playing a game. Many problems like this, however, don't have a single correct choice that the human trainers know enough to provide at every situation. But they can nevertheless be solved using a form of ML known as Reinforcement Learning (RL), where the system learns through many iterations of trial-and-error, often only being rewarded or punished after many sequential moves, such as when the game is finally won or lost.

Both problem-solving and solution-recognition types of AI can be seen as requiring search. The former uses search at run-time for a thread of reasoning to achieve a goal, whereas the latter searches at training-time for a memory structure that properly distinguishes and recognizes patterns in data. In Reinforcement Learning, this distinction has come to be known as model-based vs. model-free. Model-free learning in RL is like learning what to do in a situation without having a predictive model of what such an action will immediately lead to, so it can be smart but unable to explain itself. Model-based RL, on the other hand, learns a predictive or generative model of its actions in the environment, and so can do a kind of run-time search that compares nicely with Good Old-Fashioned AI.

Being able to reason with a generative model of the world, to understand cause and effect, and to be able to explain one's conclusions, are all important cognitive skills that simple model-free supervised learning is greatly challenged by today. Although there are many applications that only need the superb discrimination abilities of Deep Learning or model-free RL, there remain many problems for which our AI systems still seem pretty limited. Many researchers and practitioners of general AI today believe we need some kind of hybrid systems that combine the strengths of perceptual pattern recognition with higher level linguistic forms of reasoning -- to simultaneously be good at learning, problem-solving, and explanation. Reinforcement Learning is one approach that can both learn and reason, and it has the nice feature of being applicable for real-world robots who are embodied within their interactive environments, potentially reducing the cost of curating large training data sets. But RL has some of its own deficiencies, including its often high computational training cost, the difficulty of designing an appropriate reward function for multiple and varying goals, and the statistical non-stationarity of sentient multi-agent worlds in which we would like to deploy it. The interested reader is encouraged to seek further details regarding these issues in the literature.

MITRE's government sponsors are increasingly seeking nontraditional ways to address their hard problems, and the vibrant commercial marketplace of innovative ideas now includes lots of AI and ML activity. It is exciting to see the resurgence of interest in AI investments, but a cautiously optimistic stance and a willingness to continue to research the many still open problems of automating intelligence would be healthy for this industry, so as not to raise expectations above what the technology can actually deliver. As a trusted partner regarding technology and Systems Engineering, MITRE provides numerous ways to educate, connect, and collaboratively engage partners across government, academia, and commercial industry.

Artificial Intelligence is one important topic area that all of these players would like to use and to help advance.

Paul E. Silvey is a Complexity and Cognitive Systems Scientist with MITRE's Advanced Capabilities department. His focus is on contributing to novel advances and practical applications of Artificial Intelligence, Machine Learning, Data Analytics, and Distributed Systems. He has served as Department Chief Engineer and Division Technology Integrator, and is currently the Technical Lead for Artificial Intelligence for Bridging Innovation.