Nature, Nurture, and Knowledge Acquisition

Thomas R. Gruber

International Journal Human-Computer Studies, Vol. 71, Issues 2, February 2013, pp.191-194.
https://doi.org/10.1016/j.ijhcs.2012.10.004

Abstract

The nature vs. nurture dualism has framed the modern conversation in biology and psychology. There is an analogous distinction for Knowledge Acquisition and Artificial Intelligence. In the context of building intelligent systems, Nature means acquiring knowledge by being programmed or modeled that way. Nurture means acquiring knowledge by machine learning from data and information in the world. This paper develops the nature/nurture analogy in light of the history of Knowledge Acquisition, the current state of the art, and the future of intelligent machines learning from human knowledge.

Introduction

In his epic contribution to this issue (Gaines, 2012), Brian Gaines places the history and trajectory of the field of Knowledge Acquisition (KA) research in an appropriately broad context. Knowledge is a heady topic, even when studied in the embodied forms of media, expert systems, and the web. This community has always embraced the theoretical and practical together, drawing insights from history, philosophy, mathematics and other disciplines as well as the shiny objects of modern engineering. Brian offers a beautifully woven tapestry of these connections, placing the contributions of knowledge acquisition research in the context of the evolution of knowledge media and representation, information technology and artificial intelligence, and the collective body of human knowledge available online.

Within that contextual fabric, I see a familiar pattern, and it is not pareidolia. Squint hard, and ask yourself: in what field have we seen the following progression before?

- In the early days, most people thought it was magic. The established authorities said that it was supernatural, that it comes from some kind of divine or universal source.
- A few people could work the magic, and the people were impressed. The magic workers got status and power and could start charging for tickets.
- Then others wanted the magic, and started asking questions that the magic workers couldn't answer. So they started inventing their own tricks, and since they were not in power, they were willing to share them.
- With this new information available, others started to notice how some tricks worked. They made theories.
- At first, the main product of theory-making was a good story, and an audience developed for listening to good stories. It helped make sense of the magic, and it was entertaining.
- The audience for stories naturally diversified into groups, each of which is drawn to hear stories that confirm their own identity and mindset. Audiences who believe that things are predetermined liked the stories about the magic coming from fundamental, universal

structures. Audiences who believe that things are created by work liked the stories about the magic coming from the practiced craft of the magician.
- Meanwhile, people kept inventing tricks, and the stories and audience clusters helped spread the most impressive tricks, so that practices emerged. Each theoretical camp explained the successes with their best stories. The inventors liked the praise, and started using the theories to invent new tricks.
- With the synergy of theory and practice, inventing became engineering. The structural theories gave engineers building blocks with which to build impressive artifacts. The behavioral theories gave engineers a way to measure and analyze, to know how things behave in practice.
- And with time, very impressive things were created.

Brian gives a more elegant framework for this kind of progression, applied to the history of computer technology, information science, knowledge science, and beyond. What struck me was how this seems to happen in other fields.

Biology, Psychology, and AI

Consider biology and medicine. In the early days, people thought life and animate activity came from the gods or the equivalent in nature. A few people were good at creating or destroying life and influencing behavior, and they got to be priests, shaman, warlords, etc. Other people got tired of paying dues and figured out how to raise crops and animals to do medicine on their own. A lot of them died. (A lot of everybody died.) The independent thinkers started noticing patterns, some of them intermarried with practitioners, and camps of story/practice/knowledge transmission emerged (think Ayurvedic, Chinese, and European traditions in medicine). As best practices started to improve the quality of life, science and engineering emerged, and knowledge became power. Systems of knowledge acquisition and dissemination that worked better got more power. The structural theories of biological mechanisms such as germs and genes led to breakthrough therapies. At the same time, evidence-based medicine helped reward the practices that worked, even when not understood. Today, we have both systems working together. The combination has produced impressive artifacts, like the human genome database and therapies constructed by molecular engineering.

I see the same pattern in the history of psychology and cognitive science. Fast forward from early ideas about human nature and the craft of illusionists to the emergence of theories that begin to predict and prescribe. Social scientists got better at devising experiments; therapists got better at applying results. Again, camps of story/practice/knowledge transmission emerged. Structuralists explore genetic and biological determinants of behavior; behaviorists explore the nature of learning. With technology, we can now test structuralist theories of cognition in an MRI, and behavioralist theories on the web. Impressive indeed.

Now to bring this home, apply this to the field of artificial intelligence (AI) and knowledge science. Twenty five years ago, we were in awe of the magicians. It was amazing! Computer programs could play board games with people, diagnose infectious diseases, and talk as if they could see. We were amused by the appeal of ELIZA (Weizenbaum, 1966) and wanted our own professor's assistant from Apple's *Knowledge Navigator* video (1987)[1]. We had high priests and grand theories, but they didn't do the tricks. The tricks were where the action was, and the Knowledge Acquisition community

---

[1] http://www.youtube.com/watch?v=-jiBLQyUi38&feature=gv&hl=en

asked a lot of questions. Is there a man behind the curtain in that demo? If it's an expert system, how *do* you model expertise? If knowledge is power, why is a large program more brittle than a small one? How do we make an engineering discipline out of knowledge representation? How do we make tools for building systems that can reason? What are the tricks for machines learning from people?

As in other fields, the dance of theory and practice emerged. Generic methods for representation and reasoning led to tools for the same, with corresponding architectures for ontology-based classification, task-based problem solving, and evidence-based inference. While mainstream AI conferences rewarded increasingly dry-erase whiteboard results, the culture of the KA community always stayed close to practice. As Brian recounts, the workshops demanded equal facility at writing a peer reviewed article for a journal and giving a demo under a poster. The culture of working systems and application-grounded research gave the KA community momentum and adaptability. When AI went into winter recess, we had sled races. When the web happened, we embraced it - even though the architects of HTTP, HTML, XML, and the rest did not know about "proper" knowledge representation. They created magic, and we wanted to see what we could do with it. Soon, the architect of the Web was proposing a *semantic* web (Berners-Lee et al., 2001), adopting the tools of our trade. And AI people started building programs to learn from the "collected intelligence" of the web.

Where are we today in this developmental progression? Has the field of artificial intelligence and knowledge science created impressive things? Well, we have Watson[2] and Siri[3].

Watson is a master question-answering machine, better than most humans at answering factual questions. Watson doesn't "know" more than people. Rather, I think of Watson as a brilliant research librarian; it is an expert at understanding questions and finding potential answers from an array of sources. Although it is tuned for performance in the game of Jeopardy, its architecture is capable of absorbing content from a large variety of sources and adapting to new types of questions (Ferrucci et al., 2010). Watson demonstrates convincingly that intelligent performance can be achieved by harvesting human knowledge in written form.

Siri is a virtual personal assistant, realizing a longstanding dream of AI as a consumer product used by millions of people every day. If Watson is a research librarian, Siri is like a personal secretary or concierge: it helps users solve tasks such as making a schedule or reserving a restaurant. Siri (as described before acquisition by Apple, Gruber (2009))[4] is an integration of many technologies, including speech to text, natural language understanding, dialog generation, semantic data processing, service orchestration, task flow models, and domain models. Siri demonstrates that intelligent assistance is possible now, for everyone, on a platform of mobile devices connected to powerful cloud computing connected to the world of web-based services (Gruber, 2010).

While there have been many great applications of AI, I call out Watson and Siri as landmark moments in the historical progression of a field as outlined above. Led by teams of people with both research and business expertise, they are practical architectures built for flexible development and

---

[2] http://ibmwatson.com

[3] http://siri.com

[4] One might also find relevant patents published about this product, although they describe possible and not necessarily implemented technology.

scale. They are powered by knowledge of the human world, drawn from many disparate, independently developed sources. They are relatively robust in the face of noise and incomplete information. Neither is the result of a single breakthrough; they are large system integration projects combining the state of the art from many sources. They have made the transition from science to engineering.

Nature and Nature in Intelligent Systems

Returning to biology and psychology again, consider the distinction called nurture/nature. The nurture camp views the organism as a general-purpose machine and says that the functional behavior we observe comes from its experience in interacting with the environment. B. F. Skinner says that animals are mainly learning machines, and he applied this to human psychology. Modern psychology allows for internal state, such as the computational mechanisms posited by cognitive science, but the "nurture" camp maintains that a general machine is programmed by experience. The "nature" camp views the organism as strongly guided by internal mechanism and structure. In this view, the intelligence we see has evolved to adapt the to environment, rather than being determined by it. Evolutionary psychology is naturist, looking for inherited "human nature" in modern clothes. Noam Chomsky says that people are wired for language (Chomsky, 1986). They can learn a particular language, but the ability to learn language is innate.

I think there is a strong analogy for knowledge acquisition. In the context of building intelligent systems, nurture means acquiring knowledge by machine learning from data and information in the world. Nature means acquiring knowledge by being programmed or modeled that way. There are great stories for each camp.

At a high level, we can see Watson and Siri as representing the nature end of things. If Watson were born when it first encountered its operational environment (competing against humans), we would say that it was "born smart" about Jeopardy. It draws on large bodies of text, but is "wired" to index, match, rank, and apply passages from that text to answer questions. Its functional behavior of interest — answering questions in the game format — is *not* the result of raising a baby jeopardy player on a diet of encyclopedias and letting it play for 10,000 hours.

Siri was also "born smart" about the domains it covers. It is powered by explicit models of domains, tasks, and language. Unlike an omniscient agent or universal learning machine, the virtual assistant paradigm introduced by Siri achieves competence on a known set of tasks and domains. In the Chomskian sense it is "wired" to acquire the knowledge of new tasks and domains—and human languages. It is not a general purpose AI agent, armed with a quiver of web services, and fed a torrent of queries to figure out. It knows what it knows, and it knows who to ask when it doesn't know something. These external services, in turn, have models of their domains and the use cases they support.

In the nurture corner of the ring we can find search engines such as Google. Although many of its internal algorithms are kept secret, it is fair to say that Google's search engine is largely a learning machine. It systematically reads everything it can get, and continually updates an index that matches query phrases to sets of matching pages. The answers that it gives for a particular query are almost entirely driven by what it reads in its sweep through the web. As the content changes, so do the answers. Accounting for hysteresis to keep things stable, this is what you would expect from a *tabula rasa* intelligence.

Biology and psychology, now in a golden age of discovery and invention, have found gold in both nature and nurture. The frontier of knowledge is in the complex interplay of the two. An organism is predisposed for some things by its genetic structure. It lives in an environment that alters the expression of that genetic structure. It is not nature versus nurture; it is both. And as Brian Gaines discussed, our knowledge of such a system is itself a system that builds on itself. The more we learn about the structure of our genome, the more we find out ways in which the environment and other structures can modulate its expression.

I think AI is ready to make a similar transition into a new level of understanding about the nature (and nurture) of knowledge and learning. We now have the greatest source of data for experimentation ever imagined—the web. Google and others can do experiments in linguistics, psychology, and machine learning of fantastic scale in incredibly short order. We can test the limits of nature/nurture and probe the subtle interactions. For example, can a "blank slate" mind learn a language, and from what inputs? Google Translate and experiments like it are a fabulous environment for testing the limits of unsupervised learning of language. On the nature side, can machines be programmed for language with enough models? Or is there a hybrid approach? Modern speech recognition technology is based on acoustic and language models, but those models are generated from training data. Is there an equivalent for the natural language, semantics, task, and dialog capabilities of an assistant? Can new domains be learned from human teachers? Can domain models be built by humans *and* tuned with data?

Ultimately, we would ask: can machines bootstrap off the substrate of human knowledge to greater intelligence? It won't happen by sheer quantity of data and speed of computation, as some have predicted. It will require human-made structures to bias induction (desJardins and Gordon 1995) and steer the universal learning machine toward useful generalizations. For example, if concepts can be "discovered" in patterns of words, which concepts are to be discovered? It is not, and cannot be purely a function of the training data. The preprogrammed structure is an inductive bias that guides the choice of "interesting" patterns. If the program looks for correlations of sequences of words, then it will learn concepts denoted by sequences of words, as is the common practice in learning from text today. However, the discovery machine could just as easily look for words with interestingly shaped letters, defined by some "interestingness heuristic" (Lenat and Brown, 1984). For example, if X and S are interesting letters, it might discover "concepts" that are denoted by words containing a lot of X's and S's. Evidence for this concept might actually be found in the data - think of the unnatural frequency of S's and X's in the names of sexy luxury cars. The machine, however, will not have discovered a meaningful concept but rather would have stumbled over a marketing practice unrelated to the shapes of letters. Humans who design KA systems make choices about what patterns to aim their programs at. As computers get faster and bigger they can look through more data, but they are still being told what to look for. For knowledge acquisition systems, the inductive bias is the endowment from nature that guides how the agent adapts to its nurturing environment.

The Future

For our field, a fundamental question for knowledge acquisition might be this: *to learn from collective human knowledge, what kind of structure applied to what kind of knowledge in what form can lead to learning genuinely new knowledge?* Will the structure be in the patterns to look for, the relationships among things, the statistical properties of collections, or other inductive biases? Will the golden discoveries be found in the human generated content, structured data, graphical or auditory sources, query streams, or other breadcrumb trails of human experience? What will be the most useful measures of interestingness when learning from everything?

The answers to this inquiry will guide us into the next level, which is: can this process be applied recursively to itself? That is, can our machine learners apply what they learn to improve their own learning? Based on my intuition drawn from analogous fields, I would predict that mechanisms of knowledge discovery are not recursive. Experience from biology and psychology have shown that levels matter; fundamental processes by which function emerges from structure in an environment do not always operate the same way at other levels. Is there an epigenetics of knowledge? As we learn to build machines that can learn from us, we might invent new biasing structures to guide them to become better learners. Let's see how this plays out in the field of knowledge acquisition for the next 25 years.

References

Berners-Lee, T., Hendler, J., and Lassila, O. (2001). The Semantic Web. *Scientific American*, May 2001, pp. 34–43.

Chomsky, N. (1986). *Knowledge of language: Its origin, nature and use.* Westport: Greenwood.

desJardins, M., and Gordon, D. (1995). Evaluation and selection of biases in machine learning. *Machine Learning Journal*, 20 (1) 1-17.

Ferrucci, D., Brown, E., Chu-Carroll, J., Fan, J., Gondek, G., Kalyanpur, A., Lally, A., Murdock, J., Nyberg, E., Prager, J., Schlaefer, N., and Welty, C., 2010. Building Watson: An Overview of the DeepQA Project. AI Magazine 31(3): 59-79.

Gaines, B. (2012). Knowledge Acquisition: Past, Present and Future. *International Journal of Human-Computer Studies*, Special Issue on 25 Years of Knowledge Acquisition.

Gruber, T. (2009). Siri: A Virtual Personal Assistant. Keynote presentation at the Semantic Technologies conference, June 2009. http://tomgruber.org/writing/semtech09.htm

Gruber, T. Big Think Small Screen: How semantic computing in the cloud will revolutionize the consumer experience on the phone. Keynote presentation at Web 3.0 conference, January 27, 2010. http://tomgruber.org/writing/web30jan2010.htm

Lenat, D.; Brown, J.S. (1984). "Why AM and EURISKO appear to work". Artificial Intelligence 23 (3): pp. 269–294.
Weizenbaum, J. (1966). "ELIZA — A Computer Program For the Study of Natural Language Communication Between Man And Machine", *Communications of the ACM* 9 (1), pp. 36–45.