The Computational Gauntlet of Human-Like Learning

Pat Langley

Institute for the Study of Learning and Expertise
2164 Staunton Court, Palo Alto, CA 94306 USA
Center for Design Research, Mechanical Engineering
Stanford University, Stanford, CA 94305 USA

Abstract

In this paper, I pose a major challenge for AI researchers: to develop systems that learn in a human-like manner. I briefly review the history of machine learning, noting that early work made close contact with results from cognitive psychology but that this is no longer the case. I identify seven characteristics of human behavior that, if reproduced, would offer better ways to acquire expertise than statistical induction over massive training sets. I illustrate these points with two domains – mathematics and driving – where people are effective learners and review systems that address them. In closing, I suggest ways to encourage more research on human-like learning.

Background and Motivation

Despite its modest beginnings, machine learning has come to play a dominant role in AI research and engineering. Recent years have seen impressive results on problems long viewed as challenging. We now have systems that recognize object classes in images with high accuracy, translate passages reliably between languages, and outperform experts on difficult games. However, the field has become remarkably uniform and shares critical assumptions about the representation, use, and acquisition of expertise. The AI community should question these tenets and devote more energy to another paradigm that holds even greater potential. In particular, it should pursue more research on computational approaches to *human-like* learning, which operates very differently than mainstream techniques for machine induction.

Consider a familiar domain in which many people gain broad expertise through the educational system: mathematics. This requires students to acquire concepts and skills at many levels. Children learn to distinguish digits with some ease, although writing them takes more practice. Once they master digits, students learn arithmetic tables, followed by procedures for multicolumn addition, subtraction, and the like. Later they acquire high-level procedures for simplifying fractions and solving algebraic equations. This curriculum takes years, but people do not require thousands of cases for each concept. Rather, teachers combine explicit instruction, worked-out examples, and practice problems to cement student understanding. The trajectory of human learning differs radically from how we currently train machines.

Now turn to a second domain that has received more attention in the statistical learning community: driving a vehicle. Many people master driving in their teenage years, after they have gained substantial knowledge about the physical world, but they must still learn specialized concepts and skills for the new setting. These include categories for roads, lanes, intersections, signs, and signals, along with skills for changing lanes, passing others, and turning corners. They must also internalize laws and social norms, such as obeying signs and rights of way. These require training and practice, but most drivers are reasonably proficient after a short course. Unlike statistical learning systems, they do not need millions of miles' experience to acquire basic competence.

In the sections that follow, I address the growing chasm between human and machine learning. I begin with a brief history of the latter field, which originally made close contact with ideas from cognitive psychology. Next I consider high-level features of human learning, which I cast as a gauntlet of challenges that the discipline should traverse. Finally, I give examples of AI systems that learn like people, respond to some possible objections, and propose ways the community can foster progress along these lines. I intend this as a 'blue sky' paper that will stimulate the community to pursue an important line of research. Human-like learning is not a new goal, and I cite early work in this tradition, but in recent decades this aim has fallen into disfavor, and it deserves far more resources than it has received to date.

A Brief History of Machine Learning

The field of machine learning was launched at a 1980 workshop that produced an influential edited volume (Michalski et al. 1983). Other workshops followed in 1983, 1985, and 1987, which led in turn to a refereed conference in 1988 and to the journal *Machine Learning*.[1] Initial membership came from two communities – AI and cognitive psychology – with the second group being a sizable fraction. Some researchers desired to automate construction of expert systems, while others focused on modeling high-level learning in humans.

Research in machine learning's first decade emphasized methods that created *symbolic structures*, such as condition-action rules, decision trees, and grammars, with parameter-

[1]There was research on computer learning before these events, but little sense of a coherent field with its own aims and methods.

fitting techniques viewed as entirely distinct. However, a growing interest in experimental evaluation led, by 1990, to redefining the field as the study of all computational artifacts that improve performance based on experience. This encouraged researchers to borrow ideas from areas like pattern recognition and to consider statistical and probabilistic methods, including neural networks. Initial empirical comparisons of the paradigms were controversial, but the community soon recognized they were different but comparable.

Early applications of machine learning (Langley and Simon 1995), both symbolic and parametric, were deployed, providing evidence that the field offered a viable path to automated creation of expert systems. These results fostered the data-mining movement, which launched a new conference in 1995 and a journal shortly thereafter. This paradigm emphasized large data sets and efficient processing, and it led to more application successes, but it also discouraged use of human learning as inspiration for new techniques. Instead, data mining drew on ideas from statistics, which buttressed new work on learning with Bayesian networks, support vector machines, and ensembles of models. Efforts on computer simulation in psychology continued, but focused increasingly on fitting quantitative results from specific studies, rather than being consistent with qualitative phenomena.

Progress on learning in neural networks, which also relied on large training sets and high-performance computing, encouraged this trend further. Despite common claims that such networks behave like people, their widespread adoption has taken the field even further away from human learning, at least concerning high-level abilities that distinguish us from other species. The research community applauds when systems reach human-level expertise by training on millions of images or game-playing runs, but they seldom express concern that, to reach this level, they need orders of magnitude more samples than people. Arguably, machine learning has succeeded at the aim of automating construction of expert systems, but only at the cost of abandoning its links to psychology, which still has many insights to offer.

Constraints on Human-Like Learning

Before the AI community can develop systems that acquire expertise in a human-like manner, we must first identify the important features of human learning. I am not referring to quantitative results of psychology experiments, which all too often are fit using modern variants of nonparametric regression. Rather, I am talking about high-level regularities that hold for human learning in many settings. These correspond to what Newell and Simon (1976) called *laws of qualitative structure*, which serve to bound more detailed models.

Insights about the character of human learning provide strong constraints on the AI systems that aim to mimic them. These should serve as steps in a computational *gauntlet* – a passage lined with armed adversaries – that candidates for human-like learning must traverse. Many of these features have been documented in the psychological literature, but some are so obvious that it is difficult to provide references. In this section, I enumerate these characteristics, in each case contrasting them with mainstream machine learning and citing work that has addressed them. I divide these properties into two groups: those concerning the representation and use of learned expertise and those involving the mechanisms that transform training experiences into knowledge.

One basic feature of human learning, shared by many psychological theories (Bower 1981) and supported by many empirical studies, concerns the nature of acquired content:

- Learning involves the acquisition of **modular cognitive structures**.

This statement does not specify details about these structures, which manifest differently in competing theories, but only that expertise consists of discrete mental elements. Candidates from the literature include concepts, production rules, exemplars, chunks, and even stimulus-response pairs, but each contrasts sharply with the idea that learning produces a single large structure, whether a multilayer neural network, complex decision tree, or ensemble of models.

The second characteristic, related to use of expertise, is enabled by the first one and often associated with it closely:

- Learned cognitive structures can be **composed during performance**.

In other words, the elements of expertise are accessed and combined as needed to produce behavior. This is often seen as a stepwise process, as when one rule creates elements that enable other rules to apply, but the idea can take on different forms. Theories of generative grammar in linguistics offered early examples of such frameworks, and the idea is central to cognitive architectures (Langley et al. 2009). Composable elements differ radically from the large structures (e.g., neural networks or decision trees) produced by most statistical methods, which are created and stored at learning time.

The two illustrative domains offer compelling examples of both principles. Mathematics education clearly involves the acquisition of distinct mental structures that are reflected in separate lessons. Moreover, test problems often require students to combine these elements to find solutions. Similarly, driving expertise comprises a large set of identifiable concepts and skills. People acquire many of these elements separately, from instruction and practice, but they must later combine them dynamically to make extended trips.

Most early efforts on machine learning focused on the acquisition of modular, composable structures. Symbolic rules were a common encoding of knowledge that satisfied these criteria. They appeared in supervised concept learning, but their use as composable structures was key to research on learning in problem solving (e.g., Minton 1990) and language (e.g., Berwick 1980). Later work on analogical reasoning demonstrated composition of retrieved cases (e.g., Veloso and Carbonell 1993). In contrast, modern statistical approaches use modular building blocks but effectively compile them into monolithic structures at learning time.

Another representational characteristic, somewhat more specific, concerns the form of elements that humans learn:

- Many learned cognitive structures are **relational**, in that they refer to connections among entities or events.

This claim appears frequently in computational theories of human cognition. Concepts often have this character (Kotovsky and Gentner 1996), and both production systems (Langley et al. 1987) and analogical models (Gentner and Forbus 1991) adopt relational encodings. These contrast with notations favored by statistical learning, which focuses on nonrelational tasks or mimics relations with schemes like convolutional processing. Not all human learning involves relations, but it is common enough to include as a constraint.

Both mathematics and driving are inherently relational domains. In multicolumn subtraction, whether one borrows depends on the relative size of digits, and actions in algebra are determined by symbols' positions relative to the equals sign. Driving requires one to maintain appropriate relations to lanes, pedestrians, and other vehicles. These have quantitative aspects, but the difference between ahead and behind or left and right is essentially qualitative. These examples suggest relational encodings will be useful in both arenas.

Early research on machine learning relied on relations to represent both experience and expertise. This held for categorization (e.g., Winston 1975), problem solving (e.g., Minton 1990), and language (e.g., Berwick 1980), but by 1990 they had become less widespread than attribute-value schemes. Two important exceptions are inductive logic programming (Muggleton 1999) and statistical relational learning (Getoor and Taskar 2007), which remain active but receive less attention than attribute-value methods. Convolutional neural networks can encode certain spatial relations, but recent efforts on graph convolutional networks (e.g., Li et al. 2018) have also addressed abstract relations.

Other characteristics concern how people process training experiences and create structures during learning. One important feature, linked to the modularity idea, is that:

- Expertise is acquired in a **piecemeal** manner, with one element being added at a time.

In other words, people do not learn complex models *en masse*, as done by most statistical systems. They first acquire one structure and then another, continuing until they have achieved broad coverage. This does not mean they never revisit elements created earlier to revise content or adjust numeric annotations, but each structure is learned in a reasonably independent way. An exception is discrimination learning (Medin 1976), which requires distinguishing one class of situations from others, but even this occurs in a piecemeal manner, rather than via massive statistical analysis.

A second processing constraint, which focuses on training cases rather than on acquired knowledge elements, is that:

- Learning is an **incremental** activity that processes one experience at a time.

This is connected to the notion of on-line learning, which denotes sequential presentation of training cases, but it also requires processing these stimuli only once or at least rarely. Many neural network methods update weights after examining each instance, but revisit them repeatedly on successive epochs. Incremental processing is often associated with piecemeal learning, but the two features are orthogonal. Cascade correlation (Fahlman and Lebiere 1990) and bottom-up induction of context-free grammars (Wolff 1980) add new structures one at a time but they process data in batches. Naive Bayes is incremental but not piecemeal, since it updates global statistics on classes and predictive attributes.

Our domains provide clear examples of both piecemeal and incremental processing. Mathematics students acquire knowledge elements sequentially from instructions and sample solutions. These experiences may remind them of earlier ones, but there is no evidence they retrieve and reprocess more than a few prior cases. Similarly, novice drivers learn from situations encountered one at a time that lead to new structures or that update existing ones. They do not collect endless instances of stop signs or left turns and analyze them statistically for regularities, as done in modern approaches.

The early literature on supervised learning emphasized incremental methods that added or updated knowledge elements after each experience (e.g., Winston 1975; Mitchell et al. 1986). Research on concept formation (e.g., Fisher 1987) and learning for problem solving (e.g., Minton 1990) also had this piecemeal, incremental character. But by the late 1990s, batch processing was the default, with incrementality linked to data-stream mining (Gama 2010) and cognitive architectures (Langley et al. 2009). Reinforcement learning is often incremental but updates statistics rather than creating new structures, and deep learning has been used only rarely in on-line settings (Ren et al. 2021).

This dependence on previous experience leads naturally to a sixth key characteristic of the acquisition process:

- Learning is **guided by knowledge** that aids the interpretation of new experiences.

Because the acquisition of expertise is piecemeal and incremental, it takes place in the presence of structures added earlier. This content provides context that modulates processing of new training cases. The influence can take different forms that depend on the types of structures being created:

- *Taxonomy construction* introduces subcategories that are conditioned on more general ones already in memory;
- Acquisition of *composite structures* builds upon existing components in a cumulative fashion;
- *Explanations* of observations in terms of known content are cached in new structures that *compile* them.

Such knowledge-guided learning receives little attention in the data-intensive paradigm, which idolizes knowledge-free induction. Methods for 'k-shot' learning (Wang et al. 2020) are exceptions but assume prior training on large data sets. Recent work on 'physics-informed learning' (Karniadakis et al. 2021) is a better analog in the statistical community.

Our two illustrative domains offer clear examples of how knowledge aids human learning. In mathematics, we tackle algebra only after mastering basic numeracy skills, since they serve as building blocks, and students who explain sample solutions to themselves learn more effectively than ones who do not (Chi and VanLehn 1991). Novice drivers acquire basic categories for road configurations, signs, and vehicles, which they later refine into a taxonomy, and they construct complex skills in terms of simpler ones learned previously.

The early literature on machine learning reserved a central place for knowledge. During the 1980s, 'explanation-based' approaches (e.g., Mitchell et al. 1986) were a dominant theme, but most methods relied on deduction, while human explanations are often abductive and make plausible assumptions (Lombrozo 2012). Research on concept formation (e.g., Fisher 1987) used an existing taxonomy to categorize instances, which in turn refined this knowledge. Work on cumulative learning was less common, but appears in papers on structural transfer (e.g., Könik et al. 2007). However, using knowledge to guide acquisition is rare in modern efforts, and this aspect of human behavior merits attention.

A final important characteristic of human learning, enabled by the combination of features covered earlier, is that:

- Cognitive structures are acquired and refined **rapidly**, from small numbers of training cases.

This postulate interacts with the earlier notion that learned content is modular. We do *not* acquire *all* expertise in a domain from only a few instances, but we do learn each structural element quite rapidly. Learning curves (Thorndike 1927), which plot performance as a function of experience, are often very steep in humans. This differs from statistical induction's reliance on thousands or millions of items. People need more training cases to distinguish more classes, but still far fewer per category than modern learning methods.

Again, our two domains provide examples of rapid learning. Mathematics takes years for typical students to master, but they acquire many concepts and skills, each of which takes modest effort. Driving takes less time, presumably because students transfer many structures from locomotion, bicycle riding, and other physical activities. Human drivers fine tune their skills with practice, but they acquire the core elements quickly. This does not resemble the way companies currently train self-driving cars on massive data sets.

Many of the approaches developed early in the field's history supported rapid learning, including techniques for acquiring concepts (e.g., Winston 1975; Fisher 1987), search heuristics (e.g., Minton 1990), and grammatical knowledge (e.g., Berwick 1980). This was often explained in terms of a strong but appropriate *inductive bias*. In contrast, methods like deep neural networks make much weaker commitments, so they require many more training cases. One response is to specify a declarative bias, as in recent work on probabilistic programming, which can learn much more rapidly than neural networks on given tasks (e.g., Lake et al. 2015).

These do not exhaust the qualitative characteristics of human learning,[2] but they clarify how far machine learning has strayed from its original vision. Early work addressed these features, but the promising results were preliminary and addressing them more fully should become a priority for the field. This will lead to a new generation of AI systems that acquire expertise more effectively than current methods, while also giving insight into the nature of human cognition.

Examples of Human-Like Learning

The previous section cited research on human-like learning, but it will be useful to consider a few examples in more detail, some from the early literature and others more recent:

- Fisher's (1987) COBWEB acquires conceptual categories incrementally from unsupervised instances. The system encodes expertise as a taxonomic hierarchy of concepts, with terminal nodes storing training cases and nonterminal nodes denoting probabilistic summaries of their children. Learning is interleaved with categorization, which sorts each new case downward through the hierarchy, updates probabilities on the way, and occasionally alters its structure. COBWEB not only learns rapidly in an incremental, piecemeal way, but also explains typicality and basic-level effects from psychology.
- Minton's (1990) PRODIGY improves its ability to find plans based on previous experience. The architecture encodes expertise as relational rules that select operators, goals, and states, which it uses to guide means-ends analysis, a problem-solving strategy observed widely in humans. The system uses an abstract theory of planning to explain successes and failures, which it turns into control knowledge that reduces search on future tasks. PRODIGY also collects statistics on rules' usefulness to determine which ones to retain. Together, these incremental mechanisms lead to substantial speed up on novel problems.
- McLure et al.'s (2015) SAGE acquires complex concept descriptions from a sequence of training cases. Each description includes a set of relational literals with associated probabilities. Given a new case, the system retrieves similar descriptions and invokes structural analogy to select the best mapping for use in inference. If the similarity is high enough, SAGE merges the case with the retrieved structure, updates probabilities, and adds new relations as needed; if not, it stores a new description based on the case. The system learns geographical concepts, musical genres, and shapes of everyday objects, which it acquires from far fewer examples than statistical methods.
- Muggleton et al. (2018) report a novel technique for visual learning that combines logical reasoning with the ability to define new predicates. Acquired expertise takes the form of a multi-level logic program that, given results from low-level pixel processing, creates a parse tree that gives the types of objects in an image and their parameters. Learning generates a set of abductive explanations for an image in terms of generic background knowledge, evaluates candidates based on fit to data, and selects the best option. This relational approach induces a variety of visual concepts from noisy images, using fewer than five percent of the samples needed by statistical approaches.

Each of these systems satisfies the constraints described earlier, although they are certainly not the only such examples. Some research on deep neural networks has started to address these issues, but data-intensive, statistical induction of monolithic models, which ignores them, still remains the dominant approach within modern machine learning.

Points and Counterpoints

Hopefully, most readers have found the earlier arguments compelling, but a few may remain unconvinced. For instance, some might ask why we should bother changing approaches when deep learning, random forests, and other sta-

[2] In addition, human learning is often interactive, causal, and goal oriented, but we lack the space to discuss these facets here.

tistical techniques have been so successful? One response is that this success is questionable, in that humans are far more data efficient and they learn effectively in many more settings. Another is that, scientifically, we should understand the full range of learning methods, not just one corner of that space; the dangers of sampling bias are well known. Finally, from an engineering perspective, we should hesitate to rely too heavily on one class of solutions, which can lead to local optima. Developing systems that learn like humans will take the field into a very different region of the design space.

Another critique, often heard in AI circles, is that airplanes do not fly like birds, so why should computers think or learn like people? However, at some level of abstraction, airplanes *do* fly like birds in that both respond to issues of lift, thrust, weight, and drag, although in different ways. Moreover, early designs for flying machines were modeled more closely on birds; the materials and power sources at the time made them impractical, but advances have now produced robotic birds that flap their wings. Also, we design airplanes to carry people and payloads, which have different constraints on operation than those for avians. Mainstream induction techniques may be preferable for some problems, while human-like learning methods may be better for others.

A third rejoinder is that the distinction I have proposed is not between data-intensive statistical systems and human-like ones, but rather between fitting numeric parameters and acquiring symbolic structures. Modern work in machine learning, especially on neural networks, relies centrally on parameter estimation, but it is often applied to areas like language processing and scene interpretation that appear to involve structure. The question is whether human-like approaches that directly acquire modular structures in an incremental, piecemeal manner will fare better on such problems. In addition, many tasks that require parameter estimation, such as equation discovery in scientific domains, also involve creating symbolic structures (e.g., Langley and Arvay 2015), so the two paradigms are not mutually exclusive.

Finally, note that I have not claimed human learning *never* relies on statistics. Like other species, Homo sapiens exhibits background learning that collects and analyzes stimuli over time, typically in an unconscious manner. But the ability to acquire new cognitive structures rapidly, from few experiences, is what distinguishes us from other species, or at least from nonprimates. Any science of machine learning that ignores this amazing capacity, and its potential, is necessarily incomplete. It is also noteworthy that statistical analysis was invented not to *generate* hypotheses but to *test* them. Humans excel at forming candidate structures from limited data, but they often use additional samples to evaluate them. This complementarity does not reduce the benefits of rapid structural learning, but rather highlights its importance.

Fostering Work on Human-Like Learning

The characteristics described earlier place strong constraints on computer artifacts designed to exhibit human-like learning. These were once accepted widely by AI researchers and there were no compelling reasons to abandon them. They offer a paradigm for machine learning that remains as viable today as three decades ago, and it deserves much more attention than it has received in recent years. Thus, it seems natural to ask how we might restore efforts in this direction and encourage its wider adoption within the AI community.

One important step involves *broadening education* in machine learning. Most courses focus on statistical approaches and ignore older methods with links to cognitive psychology. Few graduate students read papers more than ten years old, so they are not exposed to the classic literature. Instead, we need more representative instruction that cuts across the paradigms, examines their strengths and weaknesses, and reviews key findings on human learning. Some courses cover a few early publications, but we can and should do better.

Another response is to *expand research funding* for computational systems that learn like humans. We cannot rely on data-centric companies like Google and Amazon, whose business models emphasize gigantic repositories. However, government funding agencies like DARPA and ONR have good reasons to desire more effective learning in data-sparse settings. Government support for the paradigm will speed progress, but this depends on committed and informed people joining the relevant agencies as program officers.

Naturally, we will also require places to *publish new results* on human-like learning. A few meetings, such as Advances in Cognitive Systems, welcome research that draws inspiration from psychology, but mainstream conferences like AAAI and IJCAI have a strong implicit bias against the paradigm. We might introduce a special track, but it would only help temporarily and we need more permanent solutions. This can only happen if program chairs take seriously the need for intellectual diversity and ensure that submissions from outside the mainstream receive fair treatment.

This raises the issue of *evaluation* for systems designed to support human-like learning. The UCI Repository helped transform machine learning into an empirical field, but it also encouraged an obsession with performance metrics. Such measures remain valuable, but we must also dare researchers to demonstrate that their learning systems behave like people. We can make such claims *operational* by borrowing analytical tools from psychology, such as learning and transfer curves, to measure the degree to which artifacts exhibit human-like learning. We should encourage AI scientists to tackle audacious problems, such as mastering mathematics and driving, but also to run their systems through the gauntlet of constraints listed above and show they can overcome each obstacle they encounter along the way.

The preceding pages have echoed analyses by other authors (e.g., Fahlman 2012; Marcus and Davis 2021) about the broader AI landscape, but focused on human learning, which displays qualities seldom addressed by work on data-intensive induction. These include a reliance on modular structures, often relational, that are composed during performance, along with acquisition in a piecemeal, incremental manner guided by knowledge to produce rapid improvement. Early research on machine learning took these constraints seriously. The statistical movement has largely abandoned them, but the challenge remains: to develop intelligent systems that survive the computational gauntlet of these traits and that acquire expertise as effectively as humans.

Acknowledgements

This essay was supported by Grant N00014-20-1-2643 from the Office of Naval Research and by Grant FA9550-20-1-0130 from the Air Force Office of Scientific Research, which are not responsible for its contents. I thank J. Laird, M. Pazzani, P. Rosenbloom, and the reviewers for useful comments.

References

Berwick, R. C. 1980. Computational analogues of constraints on grammars: A model of syntactic acquisition. In *Proceedings of the Eighteenth Annual Meeting of the Association for Computational Linguistics*, 49–53. Philadelphia, PA: ACL.

Bower, G. H. 1981. *Theories of learning*. Englewood Cliffs, NJ: Prentice-Hall.

Chi, M. T. H., and VanLehn, K. 1991. The content of physics self-explanations. *The Journal of the Learning Sciences* 1: 69–106.

Fahlman, S. E. 2012. Beyond idiot-savant AI. *Advances in Cognitive Systems* 1: 15–21.

Fahlman, S. E., and Lebiere, C. 1990. *The cascade-correlation learning architecture*. In *Advances in Neural Information Processing Systems 2*, 524–532. Denver, CO: Morgan Kaufmann.

Fisher, D. H. 1987. Knowledge acquisition via incremental conceptual clustering. *Machine Learning* 2: 139–172.

Gama, J. 2010. *Knowledge discovery from data streams*. Boca Raton, FL: CRC Press.

Gentner, D., and Forbus, K. 1991. MAC/FAC: A model of similarity-based retrieval. In *Proceedings of the Thirteenth Annual Conference of the Cognitive Science Society*, 504–509. Chicago, IL: Lawrence Erlbaum.

Getoor, L; and Taskar, B. 2007. *Introduction to statistical relational learning*. Cambridge, MA: MIT Press.

Karniadakis, G. E., Kevrekidis, I. G., Lu, L. Perdikaris, P., Wang, S., and Yang, L. 2021. Physics-informed machine learning. *Nature Reviews Physics* 3: 422–440.

Könik, T., Choi, D., Shapiro, D., Park, C., Nejati, N., Langley, P., and Stracuzzi, D. 2007. Structural transfer of cognitive skills. In *Proceedings of the Eighth International Conference on Cognitive Modeling*. Ann Arbor, MI.

Kotovsky, L., and Gentner, D. 1996. Comparison and categorization in the development of relational similarity. *Child Development* 67: 2797–2822.

Lake, B. M., Salakhutdinov, R., and Tenenbaum, J. B. 2015. Human-level concept learning through probabilistic program induction. *Science* 350: 1332–1338.

Langley, P., and Arvay, A. 2015. Heuristic induction of rate-based process models. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, 537–544. Austin, TX: AAAI Press.

Langley, P., Laird, J. E., and Rogers, S. 2009. Cognitive architectures: Research issues and challenges. *Cognitive Systems Research* 10: 141–160.

Langley, P., and Simon, H. A. 1995. Applications of machine learning and rule induction. *Communications of the ACM* 38: 55–64.

Li, R., Wang, S., Zhu, F., and Huang, J. 2018. Adaptive graph convolutional neural networks. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence*, 3456–3553. New Orleans, LA: AAAI Press.

Lombrozo, T. 2012. Explanation and abductive inference. In K. J. Holyoak and R. G. Morrison (Eds.), *The Oxford handbook of thinking and reasoning*. New York, NY: Oxford University Press.

Marcus, G., and Davis, E. 2021. Insights for AI from the human mind. *Communications of the ACM* 64: 38–41.

McLure, M. D., Friedman, S. E., and Forbus, K. D. 2015. Extending analogical generalization with near-misses. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, 565–571. Austin, TX: AAAI Press.

Medin, D. L. 1976. Theories of discrimination learning and learning set. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes*. Hillsdale, NJ: Lawrence Erlbaum.

Michalski, R. S., Carbonell, J. G., and Mitchell, T. M. (Eds.) 1983. *Machine learning: An artificial intelligence approach*. Palo Alto, CA: Tioga Press.

Minton, S. N. 1990. Quantitative results concerning the utility of explanation-based learning. *Artificial Intelligence* 42: 363–391.

Mitchell, T. M., Keller, R. M., and Kedar-Cabelli, S. 1986. Explanation-based generalization: A unifying view. *Machine Learning* 1: 47–80.

Muggleton, S. 1999. Inductive logic programming: Issues, results and the challenge of learning language in logic. *Artificial Intelligence* 114: 283–296.

Muggleton, S., Dai, W.-Z., Sammut, C., Tamaddoni-Nezhad, A., Wen, J., and Zhou, Z.-H. 2018. Meta-interpretive learning from noisy images. *Machine Learning* 107: 1097–1118.

Neches, R., Langley, P., and Klahr, D. 1987. Learning, development, and production systems. In D. Klahr, P. Langley, and R. Neches (Eds.), *Production system models of learning and development*. Cambridge, MA: MIT Press.

Newell, A., and Simon, H. A. 1976. Computer science as empirical enquiry: Symbols and search. *Communications of the ACM* 19: 113–126.

Ren, M., Iuzzolino, M. L., Mozer, M. C., and Zemel, R. S. 2021. Wandering within a world: Online contextualized few-shot learning. *Proceedings of the Ninth International Conference on Learning Representations*.

Thorndike, E. L. 1927. The law of effect. *The American Journal of Psychology* 39: 212–222.

Veloso, M. M., and Carbonell, J. G. 1993. Derivational analogy in PRODIGY: Automating case acquisition, storage, and utilization. *Machine Learning* 10: 249–278.

Wang, Y., Yao, Q., Kwok, J. T., and Ni, L. M. 2020. Generalizing from a few examples: A survey on few-shot learning. *ACM Computing Surveys* 53: 1–34.

Winston, P. H. 1975. Learning structural descriptions from examples. In P. H. Winston (Ed.), *The psychology of computer vision*. New York, NY: McGraw-Hill.

Wolff, J. G. 1980. Language acquisition and the discovery of phrase structure. *Language and Speech* 23: 255–269.