Advances in Cognitive Systems 8 (2019) 3–12 Submitted 12/2019; published 12/2019

Authoring Papers on Cognitive Systems Research

Pat Langley PATRICK.W.LANGLEY@GMAIL.COM
Institute for the Study of Learning and Expertise, Palo Alto, California 94306 USA
Department of Computer Science, University of Auckland, Private Bag 92019, Auckland 1142 NZ

Abstract

In this essay, I discuss the role of communication in scientific research and its relevance to publications on cognitive systems. In particular, I discuss the content that our field's papers should cover in order to convey their contributions to readers. This material should include a clear statement of the cognitive task under study and the target behaviors or phenomena one wants to reproduce. Articles should also present a high-level theory of these behaviors and distinguish it from a more detailed system or model that instantiates the theory. In addition, authors should make explicit behavioral claims and support them with evidence, empirical or analytical. Other topics should include related research, limits of the approach described, and plans to remedy the latter in future work.

1. The Importance of Scientific Communication

Science is a communal enterprise. The image of the solitary scientist who works in isolation is a misleading myth. Even researchers who do not participate in explicit teams spend time tracking others' progress and incorporating selected insights into their own efforts. Communication is central to this process, as it lets scientists exchange ideas, identify points of agreement and disagreement, and build on earlier results. Newton (1675) noted the importance of previous work in physics, writing "If I have seen further it is by standing on the shoulders of Giants". However, this dependence holds equally well for other natural sciences and even for 'sciences of the artificial' (Simon, 1969), which includes cognitive systems. Such cumulative progress is only possible through effective communication among researchers.

Scientists communicate in many different ways – in conference talks, during informal breaks, over the telephone, and through electronic mail – but the classic medium is published articles. This makes the art of writing such papers a central element of success in any field. As Winston (in press) discusses, good writing depends on many factors, some of them independent of a discipline's content, and I encourage readers to follow his sage advice.[1] However, each scientific field also has special concerns, and cognitive systems is distinct enough that it merits its own style of research paper. Unlike mainstream AI, our discipline does not focus on narrowly defined technical schemes that are familiar to specific subcommunities (e.g., statistical learning or automated planning). Rather, the broad scope and inherent diversity of cognitive systems means that articles must provide substantial background and state their ideas and their contributions clearly and explicitly.

1. Winston also discusses communicating ideas through formal talks and lectures, which raise many similar challenges.

In this essay, I discuss subject matter that authors should consider including when they report progress in the cognitive systems paradigm. First I note the importance of specifying the task under study and the target behaviors that one hopes to produce or explain. After this, I review the need to state clear theoretical postulates, describe an implementation within this framework, and keep the two contributions distinct. Next I discuss presentations of claims and the evidence that supports them, research on related topics, limitations of the approach, and plans to address them in future work. Not all papers on cognitive systems, including this essay, will fit this template, but it is relevant to most publications about computational artifacts that aim to produce human-like behavior. I encourage authors to consider it seriously when drafting their future manuscripts.

2. Define the Cognitive Task

In most cognitive systems papers, the first step is to describe the task being addressed. This may involve any facet of cognition, from low-level activities like categorization and decision making to high-level ones like problem solving, natural language understanding, and explanation generation. The research may also address the integration of multiple abilities, such as combined generation, execution, and monitoring of plans (e.g., Langley et al., 2017). Another classic example is dialogue processing (e.g., Allen et al., 1996), which combines sentence parsing, conceptual inference, and reactive control. Research on such integrated computational artifacts is a central concern of the cognitive systems movement.

We can specify any cognitive task in terms of the information provided and the information produced. This characterization should describe the type of content for each one, but it need not commit to any particular representations or data structures. Such details are important and interesting, but they are usually distinct from the specification of the task itself, which can be addressed in different ways. In some cases, it is natural to specify a problem in terms of inputs and outputs. For categorization, one is given a description of some entity or event, along with knowledge about candidate classes, and one selects the most appropriate category. For planning, the inputs include an initial state S, a goal description G, and a set of operators or actions, while the output is a sequence of operators that transform S into a state that satisfies G.

However, many cognitive activities require continuous processing over time, making their 'outputs' difficult to characterize. For instance, integrated plan generation, execution, and monitoring requires the same inputs as simple plan generation, but the three interacting mechanisms may continue indefinitely. The same holds for dialogue systems, which involve repeated turns among the conversing parties. In such cases, we can specify the component steps in terms of inputs, but the overall behavior does not fit into this scheme. However, we can state these steps clearly and also clarify that they may be repeated indefinitely or at least until the cognitive system encounters some conditions for termination.

For papers about learning, it is crucial to specify both performance and learning tasks, since the purpose of the latter is to improve behavior on the former. For example, if the performance task involves categorization or classification, the input to learning will include training cases, often with associated classes, and the output will include expertise for future classification. Similarly, if the performance task is plan generation, then the learning task will involve acquiring knowledge (e.g., search heuristics, macro-operators, or hierarchical task networks) that influence the planning process. Note that two distinct learning tasks may have the same output specifications. Learning from search (e.g., Sleeman, Langley, & Mitchell, 1982; Minton, 1988) receives only a set of training problems and operators. In contrast, learning from observation (e.g., Nejati, Langley, & Könik, 2006; Hogg, Muñoz-Avila, & Aha, 2008) also receives sample solutions for each training case.

Nonincremental induction, as widely adopted by the machine learning community, is easily stated in terms of inputs (a set of training examples) and outputs (learned expertise). However, cognitive systems researchers are more typically interested in incremental mechanisms that, like humans, process one training experience at a time. Research on such methods was common (Langley, 1995) until the advent of the data-mining movement in the mid-1990s. In such cases, as with integrated performance tasks, we can only specify the inputs and outputs for each step of learning, along with stating that the process continues indefinitely. Here the stepwise input is a training instance and the current knowledge, whereas the stepwise output is the updated knowledge. A similar situation applies when specifying sequential transfer (e.g., Könik et al., 2007), in which the results of prior acquisition serve as inputs to later learning tasks.

3. Specify Target Behaviors and Phenomena

The next stage in a cognitive systems paper is to state the target behaviors or the phenomena that one aims to explain. These may be high-level capabilities associated with human intelligence that we want to mimic in machines, even if the latter operate quite differently, or they may be empirical phenomena observed in people that we hope to reproduce. For example, we may desire a categorization system that classifies stimuli with approximately the same accuracy as humans, but this may not be enough. We may also want the system to improve its classification accuracy at roughly the same rate as people and thus produce comparable learning curves. Alternatively, we may desire a cognitive system that not only plays chess as well as human experts, but that achieves this level of proficiency by searching no further ahead in the game than do people.

This description of target behaviors should build on the earlier task specification, but it should also introduce further constraints or criteria for success. These additions may involve content, such as the ability to use certain types of knowledge, or they may concern processing, such as the ability to revise candidates during planning. The purpose is to rule out some approaches to the task or to provide a means of ranking different methods based on their behavior. For instance, a system that generates low-quality plans with loops might be seen as unacceptable, and a mechanism that carries out large amounts of search might be less desirable than one that finds the same solutions with less effort. Similarly, we may want a syntactic processor that not only generates acceptable parses for most English sentences, but that also covers rare but important types of utterances. Researchers should decide for themselves the aims of their research, selecting ones they deem of scientific interest, rather than simply adopting criteria that have become popular in the community.

Some authors confound the task specification with behavioral criteria, as in common attempts to define problems in terms of finding 'optimal' solutions,[2] but it is important to keep such elements distinct. The first declares the problem under study, whereas the second plays a key role in

2. Simon (1993) offers compelling arguments for why such efforts are not only misplaced but ill defined, but they remain popular in many AI circles, despite the field's early emphasis on heuristic methods and satisficing.

empirical evaluation or formal analyses. This separation also clarifies that we can select different targets for system behavior, and thus measures of success, for the same task. For instance, work on categorization is usually associated with measures of accuracy, but it can also address processing time. Analogously, research on planning usually focuses on the amount of search or processing time, but it can also be concerned with plan quality. Folding such issues into task definitions eliminates this option and thus rules out the possibility of tradeoffs among behavioral criteria, which are widespread in human behavior.

4. Present Theoretical Postulates

Once the authors have described the aims of their research in terms of the cognitive task and target behaviors, they can present the key theoretical ideas they will use to address them. These are often the most important part of a paper, as they state its core contributions to the research community. Such theoretical postulates are necessarily high level and abstract, but this makes it even more essential that authors present them clearly and succinctly.[3] Papers may specify these tenets formally, say in logic or equations, but this is not essential, and there are many examples from the history of science that were stated informally. The motion theory of heat, the oxygen theory of combustion, and the germ theory of disease all introduced important new ideas to their research communities in natural language. Even mathematical accounts, such as the theory of gravitation, included informal statements to provide crucial context.

Scientific theories usually distinguish between structures and processes. The former specify the types of entities, their associated attributes, and the relations in which they participate. The latter state the activities or mechanisms that operate over and alter these structures. For example, the theory of chemical reactions states that physical substances are composed of many small molecules, each comprising one or more atomic elements, while reactions transform some substances into others by decomposing molecules and recombining their constituents. A theory's structures and processes may be described in qualitative terms (e.g., Pasteur's germ theory) or quantitative ones (e.g., Newton's theory of gravitation). Also, note that abstract theories are distinct from more concrete models stated within them, which we will discuss shortly.

Elsewhere (Langley, 2018), I have given four examples of theories from the cognitive systems literature, some of them dating back to AI's earliest days. These included physical symbol systems (Newell & Simon, 1976), production systems (Neches, Langley, & Klahr, 1987), heuristic search (Newell & Simon, 1976), and the HPS architecture (Langley, Barley, & Meadows, 2018). For instance, the theory of physical symbol systems posits a set of symbols (stable patterns in a physical medium) and symbol structures (organized sets of symbols) that designate other entities, along with processes for interpreting, creating, and modifying the structures during extended operation. The theory of production systems postulates a dynamic working memory of specific symbol structures and a production memory of generic condition-action rules. These interact during a recognize-act cycle that repeatedly matches rule conditions against working memory elements, selects a set of matched rules, and applies them to update working memory.

3. One way to achieve this result is to use itemized lists that emphasize the theory's quintessential assumptions.

Each of these theories distinguishes between mental structures (memories and their contents) and processes that operate over them. They involve neither formal axioms nor equations, but they offer a set of clear postulates that explain an impressive range of high-level cognitive behaviors. Also note that scientific theories are never drawn from whole cloth; they invariably build on earlier ideas that enable and constrain them. For instance, the theory of heuristic search borrows from and extends the notion of physical symbol systems, and the HPS architecture elaborates on both frameworks. Similarly, production systems are a variety of symbol systems, but posit more specific forms of cognitive structures and introduce new mechanisms for performance and learning. Papers in our discipline should be clear about which postulates their theory adopts from earlier accounts and which ones are novel contributions to the cognitive systems literature.

5. Describe Implementations and Models

Theories offer an intellectual framework but they are too abstract to be tested directly. Thus, after describing a set of theoretical postulates, a cognitive systems paper should present one or more concrete models stated within them. These will specify additional modeling assumptions, constrained by the theory but not central to it, that produce or predict behavior. Examples of models abound in the history of science. Newton's theory of gravitation could not predict orbital motions around the sun without commitments about its mass and about the positions of planets. Similarly, Dalton's atomic theory was not operational without assumptions about the elements that make up particular compounds, and Pasteur's germ theory remained vague without statements linking microorganisms to particular diseases. Papers on cognitive systems should report analogous modeling assumptions.

For research on high-level cognition, models typically take the form of implemented computer programs and authors should describe them with some care. Research articles do not have space to describe every detail of such systems, but they can discuss the most important elements. Just as theories make commitments about structures and processes, so implementations incorporate specific data structures to encode content and particular mechanisms to manipulate it. Often, implementations involve multiple levels that merit separate descriptions. For instance, a theory of the cognitive architecture must be implemented in software, but it will be analogous to a programming language. To generate behavior, one must also write programs in that notation, which require their own descriptions. Papers on cognitive systems should explain these facets of the implementation and illustrate them with examples from relevant domains. They should also discuss domain knowledge, heuristics, and other structures that influence behavior.

For instance, an article on planning should provide basic information about the software implementation of the theory presented earlier. This would include not only the programming language (e.g., Lisp or Python) used to construct it, but also assumptions that it adopts for the sake of tractability or convenience (e.g., agents carry out only one action at a time), provided they are not central to the theory. The paper should also describe content specific to the domains used in demonstrations or tests. For planning, these would include the predicates used to describe beliefs, goals, and states, the actions available to the agent, and domain-specific heuristics or constraints used to guide search. The authors should also report the complexity of states, goals, and plans that occur in test problems, although it may be more natural to present this information in a later section on evaluation.

There has been considerable discussion in the AI community about the desirability of replicating results and encouragement to publish software that can be run repeatedly to support this process. However, replicability in the natural sciences is a quite different matter, where multiple laboratories attempt to reproduce experimental results reported earlier by another group. When such replications are successful, they are compelling not because the original article provided enough details to run an identical study, but because the new experiments necessarily vary in many details but nevertheless produce similar results. Effectively, the different laboratories have operated with different models of the same theory, which provides stronger support for the latter. The analog in cognitive systems would involve researchers creating different implementations of a given theory and demonstrating that they produce similar target behaviors.

6. Report Claims and Evidence

Once the authors have presented their theory and an associated model or implementation, they should also demonstrate whether the latter acts as intended. The most basic step is to show that the implemented system carries out the specified task, ideally on a set of test cases that is diverse enough to substantiate its generality. However, they should also examine whether the system produces the target behaviors they outlined earlier in the paper. Preferably, the authors should make explicit claims or hypotheses to this effect and then present convincing evidence that either corroborates these claims or contradicts them. Negative results that undermine hypotheses can be as valuable to the research community as positive ones that support them, as long as readers gain insights into the reasons they occurred.

Empirical studies are the most common approach to providing such evidence in cognitive systems research.[4] In some cases, qualitative demonstrations of the desired behaviors may be enough to support the authors' claims. To this end, they may devise scenarios or test problems for each target ability, run the implemented system on them, and report whether they behave as intended. For example, McShane, Nirenberg, and English (2018) show their approach to language understanding covers many phenomena (e.g., nominal compounds, verb phrase ellipsis, indirect speech acts) that pose processing challenges. Similar, Langley et al. (2017) demonstrate that their architecture for plan generation, execution, and monitoring deals with four types of anomalies arising during goal-directed activity. Both provide qualitative evidence for coverage of particular cognitive abilities.

Another widespread form of empirical study involves controlled experiments. Here one states testable claims or hypotheses about how dependent variables – typically measures of system performance – are influenced by one or more independent factors. Most readers will have encountered experiments that compare the behavior of some new system to the behavior of more established ones on commonly used test problems or domains. Experimental studies of this variety typically claim to demonstrate 'progress' when the new system outperforms its predecessors, sometimes even by only minor amounts. Unfortunately, such 'bakeoffs' seldom reveal insights into the reasons for any observed behavioral differences, as both the systems and domains vary in so many ways that few meaningful conclusions are possible.

4. Formal analyses often introduce similar claims that relate behavior to factors that make the task more or less difficult. They differ mainly in that support for these conjectures is analytical rather than empirical.

Langley and Kibler (1991) have described other types of experiments that offer potential for deeper scientific insights. They identified two broad categories of independent variables that can affect behavior – system characteristics and domain features. They further distinguished among ways to manipulate the system being studied (e.g., parametric experiments, lesion studies) and to alter the domain (e.g., increasing target complexity, introducing noise, adding irrelevant features). Their analysis of experimental designs focused on machine learning, but Langley and Messina (2004) have argued that similar issues arise in the empirical study of integrated intelligent systems. Such fine-grained studies, when motivated by well-crafted hypotheses, provide far more information than global comparisons of complete systems' behavior on completely different domains.

Ideally, authors should motivate controlled experiments with specific hypotheses related to the target behaviors given earlier in their paper. For instance, they might predict that, on average, their system's search heuristics will let it find plans in time that is linear in the solution length. In this case, a plausible experiment would vary the form of heuristic guidance and the plan length, then measure the time taken to find solutions. Similarly, authors might claim that the learning curves for their incremental learning method will be slowed only modestly by introduction of irrelevant attributes. These examples focus on the effect of domain factors, whereas lesion and parametric studies provide information about which components of a system are responsible for producing target behaviors. Experiments that vary system and domain characteristics can identify both sources of power and challenges to overcome in future work.

7. Review Related Research

Authors of any scientific report should present the intelluectual context for their effort, so it is important to review related research. However, this poses a communicative challenge, as there are two distinct kinds of related research. One involves prior results that directly motivated the authors' approach, say because the latter builds directly on the former or responds to its identified limitations. The other involves research that has similar aims or elements as the current one, but that did not motivate it directly. Unfortunately, many authors attempt to cover both types of work in a single section. If this combined treatment appears late in the paper, then readers will not acquire the background necessary to understand the authors' contribution; if it appears too early, then authors will have no way to compare and contrast the second type of work to their own approach.

The obvious solution, suggested by Winston (personal communication), is to divide the coverage of related research into two parts. Motivational work should appear early in the paper, say in the introduction, where it makes sense to describe the drawbacks of earlier systems or to review promising ideas that inspired the current effort. This will provide context necessary for readers to understand the reasons for the new work, whose description should refer back to this material. Prior work that did not directly motivate the current research can appear later in the manuscript, after the authors have covered their theoretical postulates, system implementation, and empirical or analytical results. This will give readers the information needed to understand its similarities to, and differences from, other approaches to related problems.

One way to organize the latter material is using a chronological list, with each piece of prior work described in turn. However, it is not enough for authors to review previous systems; in each case, they should explain the ways in which the predecessor is similar to their own artifact and the ways in which it differs. Most will have addressed same task but adopt a novel approach to it, focusing on different target behaviors, adopting another theoretical framework, or devising a new empirical evaluation. When discussing differences, authors may choose to argue that their approach is superior, but such comments are not required and they may be a distraction. The current approach need not be a radical departure from earlier ones. Often new research retains key ideas from previous work but also reports some extension that enables new or improved functionality.

However, presenting a list of systems is neither exciting to readers nor especially insightful. A more promising approach is to organize discussion of previous work around the authors' main theoretical tenets, as these provide the core of their intellectual contribution. In this scheme, for each postulate, they should give credit to earlier researchers who introduced or adopted the idea. If a theoretical statement is truly novel, then the authors should note this fact, although the incremental character of science means it will be reasonably rare. In fact, some papers will not incorporate any new postulates at all, but instead make a contribution by integrating existing ideas in some novel way. The systems-level nature of our discipline makes this form of scientific progress a more common situation than in fields that focus on isolated components.

8. Discuss Limitations and Responses

Science is an ongoing process that is never fully complete, which means that articles should always discuss limitations of the work to date, as well as plans to address them in the future. In cognitive systems research, such limitations often involve simplifications about representational structures, processes that operate on them, or system evaluations. Without comments about these drawbacks, readers may conclude that the reported research has achieved all of its aims or, more likely, they may infer that the authors are not being entirely honest about its failings. Although it is natural to discuss these issues late in a paper, too often authors treat them as mere afterthoughts and include only a few sentences about them in their conclusions. In contrast, such drawbacks really deserve an entire section, or at least multiple paragraphs, of thoughtful discussion.

Because limitations of the current research serve as motivation for additional work, it makes sense to discuss them before turning to potential remedies. Typically these weaknesses are linked to the target behaviors identified earlier, as the system will carry out the specified task, but may not do it as well as desired. Authors of cognitive systems papers should attempt to localize these drawbacks, specifying whether they reside in their theoretical tenets, in their system implementation, or in their evaluation regimen. Every theory can be extended to address new phenomena, implementations always make simplifying assumptions that can be relaxed, and computational artifacts can always be tested more carefully and thoroughly. This is especially true for research on cognitive systems, which often reports prototypes that demonstrate new functionality rather than mature software that has been refined extensively.

Authors should also outline how they plan to address these limitations in future work, typically immediately after presenting each of them. Researchers need not provide great detail about extensions they intend for their system or additional studies they plan to carry out, but they should say enough about them to convince readers they have plausible ideas that offer promising responses. For example, a paper on problem solving might suggest better ways to guide heuristic search in large spaces or one on language processing might propose improved ways to disambiguate word senses or handle complicated syntax. In rare cases where authors lack any hypothesis about the cause of some problem, they should admit the situation and make the generation of an explanation a high priority for future work. Thoughtful discussion of such issues will lead to a well-balanced and satisfying research article on cognitive systems.

9. Concluding Remarks

Finally, readers often find it useful when authors summarize the main points of their paper in a closing section. For cognitive systems articles of the type discussed here, this should review the task under study, the desired target behaviors, the theoretical postulates about structures and processes, and the impleented system that instantiates this theory. The authors should also reiterate their empirical claims or hypotheses, the results of studies designed to test them, the limitations or drawbacks of their work to date, and their plans to address them in future ventures. In addition, they should discuss prior work that motivated the current effort and, separately, other approaches that have adopted similiar ideas or addressed related issues. Readers should come away with insights about the main intellectual contributions of the research.

As noted earlier, science is a communal endeavor, and for centuries written articles have served as the connective tissue that holds the enterprise together. Even the most solitary researchers obtain ideas from their colleagues' reports and incorporate them into their own efforts; this adoptive process underlies the cumulative character of science. Authors who communicate their contributions effectively are more likely to pass on their intellectual genes and thus influence their discipline's trajectory. Clear presentation of target behaviors, theoretical postulates, concrete models, and empirical results will improve the chances that these elements will survive in the competitive landscape of science. Even the most valuable breakthroughs will have little impact if they are not conveyed to the community in understandable and accessible terms.

Nevertheless, authors of cognitive systems papers face challenges that those in other paradigms seldom encounter. Their research focuses on high-level processing over structured representations, which can be difficult to describe succinctly. Also, they typically develop integrated systems that require specification not only of component mechanisms but also how they interact. Empirical evaluation often requires new techniques, especially when researchers address a problem that has received little attention in the literature. Even discussion of related work can be nontrivial because the new research incorporates ideas from different paradigms and disciplines. Nevertheless, effective communication of tasks and behaviors, theories and implementations, and evaluation results is essential to intellectual progress in our field, and I encourage authors to pursue this task as diligently as they labor onother facets of their research.

Acknowledgements

This essay was supported by Grant N00014-17-1-2434 from the Office of Naval Research, which is not responsible for its contents. I would like to dedicate it to Patrick Henry Winston, who was not only a master communicator but who shared his ideas about effective communication with others.

References

Allen, J. F., Miller, B. W., Ringger, E. K., & Sikorski, T. (1996). A robust system for natural spoken dialogue. *Proceedings of the Thirty-Fourth Annual Meeting of the Association for Computational Linguistics* (pp. 62–70). Santa Cruz, CA: Morgan Kaufmann.

Hogg, C., Muñoz-Avila, H., & Aha, D. (2008). HTN-MAKER: Learning HTNs with minimal additional knowledge engineering required. *Proceedings of the Twenty-Third National Conference on Artificial Intelligence* (pp. 950–956). Chicago: AAAI Press.

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

Langley, P. (1995). Order effects in incremental learning. In P. Reimann & H. Spada (Eds.). *Learning in humans and machines: Towards and interdisciplinary learning science*. Oxford: Elsevier.

Langley, P. (2018). Theories and models in cognitive systems research. *Advances in Cognitive Systems*, *6*, 3–16.

Langley, P., Barley, M., & Meadows, B. (2018). Adaptive search in a hierarchical problem-solving architecture. *Advances in Cognitive Systems*, *6*, 251–270.

Langley, P., Choi, D., Barley, M., Meadows, B., & Katz, E. P. (2017). Generating, executing, and monitoring plans with goal-based utilities in continuous domains. *Proceedings of the Fifth Annual Conference on Cognitive Systems*. Troy, NY.

Langley, P., & Kibler, D. (1991). *The experimental study of machine learning*. Unpublished manuscript, AI Research Branch, NASA Ames Research Center, Moffett Field, CA.

Langley, P., & Messina, E. (2004). Experimental studies of integrated cognitive systems. *Proceedings of the Performance Metrics for Intelligent Systems Workshop*. Gaithersburg, MD.

Minton, S. (1988). Quantitative results concerning the utility of explanation-based learning. *Proceedings of the Seventh National Conference on Artificial Intelligence* (pp. 564–569). St. Paul, MN: Morgan Kaufmann.

McShane, M., Nirenburg, S., & English, J. (2018). Multi-stage language understanding and actionability. *Advances in Cognitive Systems*, *6*, 119–138.

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

Nejati, N., Langley, P., & Könik, T. (2006). Learning hierarchical task networks by observation. *Proceedings of the Twenty-Third International Conference on Machine Learning* (pp. 665–672). Pittsburgh.

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

Simon, H. A. (1969). *The sciences of the artificial*. Cambridge, MA: MIT Press.

Simon, H. A. (1993). Decision making: Rational, nonrational, and irrational. *Educational Administration Quarterly*, *29*, 392–411.

Sleeman, D., Langley, P., & Mitchell, T. (1982). Learning from solution paths: An approach to the credit assignment problem. *AI Magazine*, *3*, 48–52.

Winston, P. H. (in press). *Communication*. Cambridge, MA: MIT Press.