A Computational Account of Everyday Abductive Inference

Will Bridewell
Center for Biomedical Informatics Research
Stanford University, Stanford, CA 94305

Pat Langley
Computing Science and Engineering
Arizona State University, Tempe, AZ 85287

Abstract

In this paper, we review the main qualitative characteristics of everyday inference in human cognition and present a computational account that is consistent with them. This includes both a representational framework and associated processes that produce abductive explanations in a flexible, incremental, and efficient manner. We clarify our approach with examples and report some initial empirical results. In closing, we examine related work and suggest directions for future research.

Keywords: conceptual inference, abduction, plausible inference, cognitive architecture

Background and Motivation

The ability to draw inferences from knowledge and observations is one of the distinguishing features of human cognition. However, although people can, with some deliberation, carry out the forms of deductive reasoning associated with traditional logic, their inference often appears to operate quite differently, in a spontaneous and almost effortless fashion. Such reasoning underlies much of our ability to understand language, but it also supports many other aspects of human behavior. We will refer to this cognitive activity as *everyday inference* to distinguish it from deliberate deductive reasoning.

In this paper, we present a novel computational account of such everyday inference. Following Cassimatis, Bello, and Langley (2008), our goal is not to match the details of behavior in specific experimental studies, but rather to offer a high-level explanation of this ability in humans that is consistent with all of its main qualitative characteristics. This seems appropriate given that, to our knowledge, no existing computational models satisfy this basic criterion.

We intend to embed our computational account of inference in Icarus, a theory of the human cognitive architecture that we have described at length elsewhere (e.g., Langley, Choi, & Rogers, 2009). This framework includes modules for conceptual inference, skill execution, problem solving, and skill acquisition. We will not review Icarus in detail here, but we should recount three key assumptions relevant to the current work:

- humans always operate in some environment that provides information about their situation;
- human cognition occurs over time, with each cycle drawing on inference to guide skill execution and problem solving; and
- the inference process combines environmental input with conceptual knowledge to produce beliefs.

These assumptions place important constraints on the architecture's account of conceptual inference. For instance, most reasoning is driven by observations of the world rather than by queries. Moreover, on each cycle Icarus must update beliefs in ways that incorporate recent inputs. Both constraints differ from those usually placed on computational treatments of reasoning.

Although we have developed Icarus agents that operate in a variety of simulated environments and have shown their qualitative behavior is similar to humans' in many ways, the architecture's inference module has always been its weakest link. Although data-driven and automatic, its mechanisms for generating beliefs are implausible on a number of fronts. Two drawbacks are that it supports only deductive reasoning and that it works in an exhaustive manner. We intend the approach reported in this paper to address these and other limitations.

In the next section, we review some qualitative characteristics of everyday inference that our account should reflect. After this, we describe the representational and organizational structures that our framework uses to support reasoning, and then present the mechanisms that operate over them to produce beliefs. We clarify these processes with illustrative examples and discuss scenarios on which we have tested them empirically. We conclude by discussing related research on plausible inference, noting limitations and directions for future work, and summarizing our contributions.

Characteristics of Everyday Inference

We should begin by describing the inference-related phenomena that we desire to explain. Again, our concern is not with details like reaction time or error rate on specific tasks but with the broad characteristics that humans exhibit in their everyday inference. We view these as similar to what Newell and Simon (1976) refer to as 'laws of qualitative structure', in that they provide a framework within which to cast specific models. We will treat these characteristics as constraints on mechanisms that could support human-like abilities to understand incomplete, ambiguous information from complex environments.

- *Everyday inference deals with understanding or interpreting experience*. The primary aim is not to prove that some statements follow from others, but rather to make sense of observed facts and events. Like deduction, this process combines general rules with specific beliefs to infer other beliefs, but everyday inference has an explanatory character that attempts to connect observations into a coherent whole. This suggests that everyday reasoning is *abductive* in nature, since the acknowledged purpose of abduction is to construct explanations (Peirce, 1878).

- *Everyday inference relies on plausible reasoning*, in that it does not depend on taking deductively valid steps which guarantee that the premises imply the conclusion. Rather, it draws conclusions about the situation consistent with its general knowledge but not required by it. This feature lets humans make inferences unavailable to purely deductive reasoners.
- *Everyday inference relies on flexible processing*. Although people can utilize general rules of the sort identified with deductive reasoning, they apply these rules in a fluid manner, chaining forward or backward as needed. In this way, they can handle unpredictable inputs which may include some predicates that do not appear in any rule consequents and others that do not appear in any antecedents. This fluid access to relevant knowledge lets humans reason in situations that confound unidirectional approaches to inference.
- *Everyday reasoning regularly makes default assumptions*. This feature interacts with flexible operation by activating partially matched rules and enabling inference with incomplete information. The introduction of deductively unfounded but abductively useful beliefs has many applications, one being inference about others' mental states—their beliefs, goals, and intentions. Such plausible assumptions support additional inferences that can explain future observations.
- *Everyday inference typically involves constructing a single explanation*. Although abduction is often cast as search through a space of competing hypotheses, in most cases humans generate a single account that covers most observations at hand. Even when anomalies lead to belief revisions, people usually retain one explanatory structure, rather than framing alternatives and selecting one after evaluating them.
- *Everyday reasoning operates in an on-line, incremental manner*. Humans process new facts or observations as they arrive, typically incorporating them into an existing account. Inference also has an anytime character, in that the reasoner produces at least shallow explanations rapidly, but can generate richer ones given additional time. Together, these suggest that human reasoners incrementally refine and elaborate explanations as they process new information.
- *Everyday inference relies on very efficient operations*. Processing time appears unaffected by size of the knowledge base, suggesting a mechanism that relies on local refinement in response to new results. These features do not imply any guarantees of completeness. Many plausible inferences are never made, but those that are occur in an almost effortless manner.

Taken together, these characteristics of everyday inference suggest a model that differs radically from the usual computational accounts of reasoning, which take their inspiration from logical deduction. Earlier models of abductive inference satisfy some constraints but not others, indicating the need for a new approach. In the next two sections, we present a computational model of everyday reasoning that is consistent with all the constraints, starting with its representational assumptions and then discussing the mechanisms that operate over them.

Representations for Everyday Inference

As noted earlier, we intend our computational account of everyday inference to replace the current inference module in the ICARUS architecture. The existing module has clear drawbacks, but it also makes some representational and organizational commitments that we wish to retain. One such assumption is the clear distinction between generic concepts and concrete beliefs. Another is the hierarchical organization of conceptual knowledge. A third supposition concerns the relational character of concepts and beliefs, which often describe environmental configurations among two or more entities.

The new inference framework retains all these commitments. As in ICARUS, conceptual knowledge takes the form of Horn clauses that specify a generalized consequent in terms of generalized antecedents. For instance, Table 1 shows a number of such conceptual rules related to diseases and research projects. These rules are definitional in the sense that an antecedent serves mainly as shorthand to summarize situations described by the antecedents. However, as we will see, the logic-like notation does *not* mean that the inference mechanism must interpret the rules as deductive implications. It is a common misconception that logical formalisms can be used in only an inflexible, deductive manner.

As the example illustrates, conceptual rules take a relational form. The consequent involves a predicate with associated arguments, whereas the antecedent is a set of such structures. Terms like ?person and ?s1 are pattern-match variables that must bind to constants during matching. The framework supports hierarchy by letting predicates in the antecedent of some rules appear in the consequents of others, much as in languages like Prolog. Beliefs take the form of instantiated literals that consist of a predicate and its constant arguments. The framework assumes three types of belief: ones that come from external perception or communication, like those in Table 2; ones inferred from a rule's consequent; and default assumptions based on a rule's antecedents.

A third, albeit inherently transient, type of mental structure is the rule instance. These objects take the same form as generic conceptual rules but have domain or Skolem constants in place of variables. Thus, rather than making general claims about the nature of the

Table 1: Some conceptual rules that support everyday inference about illness.

```
(has-flu ?person) ⇐
      (has-symptom ?person ?s1) (fever ?s1)
      (has-symptom ?person ?s2) (cough ?s2)
(has-food-poisoning ?person) ⇐
      (has-symptom ?person ?s1) (fever ?s1)
      (has-symptom ?person ?s2) (vomiting ?s2)
(has-lung-cancer ?person) ⇐
      (has-symptom ?person ?s1) (cough ?s1)
      (has-symptom ?person ?s2) (yellow-teeth ?s2)
(caught-flu ?person1 ?person2) ⇐
      (at-meetings ?person1 ?project) (has-flu ?person1)
      (at-meetings ?person2 ?project) (has-flu ?person2)
(project ?project) ⇐
      (member-of ?person1 ?project)
      (paid-from ?person1 ?project)
      (at-meetings ?person1 ?project)
```

world, they make specific claims about a particular situation. Such rule instances serve as hypothetical connections among candidate beliefs, some of which do not yet appear in short-term memory. These structures are created, evaluated, and often discarded during inference.

The final type of mental element is the justification, which is a longer-lived variant of a rule instance. Similarly, these structures contain literals that are grounded with domain and Skolem constants but, in this case, they serve to connect the beliefs that appear in them. In effect, justifications act as the glue that binds beliefs to each other, creating a supporting lattice of observations and assumptions. Notably, our framework does not distinguish between justifications that are valid deductions from observations and those that involve assumptions.

We can view the collection of beliefs, conceptual rules, and justifications as a tentative explanation for a collection of facts. The resulting structure is a lattice of beliefs attached to the window of the world by observations. We will see shortly that, as the explanation expands to cover more observations and increases its internal connectivity, it becomes more cohesive and, hopefully, more coherent.

Mechanisms for Everyday Inference

Now that we have described the structures on which everyday inference is based, we can describe the mechanisms that operate over them. Recall that the purpose of inference is not only to infer new beliefs from other beliefs, but also to explain how they relate to each other. Also remember that we are concerned with agents like humans that exist over time, observing a few facts and processing them before encountering additional ones. Moreover, the inference process must support plausible, flexible reasoning in partially observable settings.

Our computational framework, which we call AbRA, posits that inference operates in cycles which alternate between selecting a current belief on which to focus and chaining off this belief through a rule instance. On each

Table 2: Initial observations driving an example of everyday inference in the illness domain.

1. (member-of Ann muri-project)
2. (member-of Bob muri-project)
3. (has-symptom Ann s1)
4. (fever s1)
5. (has-symptom Bob s2)
6. (cough s2)

cycle the inference mechanism has access to a set of beliefs, some originating from outside as observed facts and others inferred on previous rounds. These literals constitute the contents of the agent's *working memory*. We hypothesize that the inference process does not match rules against all these elements, as in production-system frameworks like Soar (Laird et al., 1987), but that it selects one of them as the focus of cognitive attention.[1] For example, given the literals in Table 2, the system might select *(has-symptom Ann s1)* as the current focus. For now, we will assume this choice is arbitrary, but later we will consider ways to guide the selection process.

Once AbRA has selected some literal L from working memory, it uses this element as an anchor to drive rule instantiation and application. To this end, the system finds all rules that have one or more antecedent or consequent that unify with L. For instance, the *has-flu*, *has-food-poisoning*, and *has-lung-cancer* rules in Table 1 include the *has-symptom* predicate that unifies with the focus. If a rule unifies with the focus literal in multiple ways, then AbRA considers each possibility. For each candidate rule R, it also finds existing literals that are identical, after bindings substitutions, with R's other antecedents or consequents. For example, one can extend the match of the *has-flu* rule to include the literal *(fever s1)*, which is identical to the existing belief.

After the inference mechanism has found all rule instances that connect with the current focus of attention, it selects one of these candidates. As before, we will assume for now that this choice is arbitrary. AbRA applies the selected rule instance to carry out an inference step. This involves generating new literals for antecedents or consequents that do not exactly match existing beliefs after the instantiation phase. When an antecedent or consequent includes unbound variables, the system uses Skolem constants for those terms; if the same unbound variable appears in multiple rule elements, it uses the same Skolem for each occurrence. In our example, AbRA infers three new beliefs—*(has-flu Ann)*, *(has-symptom Ann sk1)*, and *(cough sk1)*—only two of which include Skolems. Note that some inferences correspond to antecedents, while others relate to consequents.

[1]This idea bears some resemblance to ACT-R's (Anderson et al., 2004) reliance on buffers that hold single elements, but we have been influenced more by Cassimatis et al.'s (2010) Polyscheme, which controls inference in a similar manner.

Having selected a justification from the available candidates, the inference system adds it to the growing explanation. This step involves storing the justification, adding new beliefs, and creating links that connect these memory elements. New inferences become explicit beliefs in short-term memory, whereas justifications appear as separate structures used primarily by heuristics that we discuss shortly. To this end, AbRA creates pointers between each justification and beliefs that it supports.

After the system has generated one or more new beliefs and their associated justifications, it continues the process. On the next cycle, it selects another literal as the focus of attention, finds rules with antecedents or consequents that unify with the focus, selects a rule instance to apply, and so on. This continues until the process can generate no other beliefs or until time runs out. Because AbRA assumes the agent operates in an external environment, literals that correspond to observations may enter memory on any cycle, providing material to drive inference. In principle, it can reach a quiescent state in which no further inferences arise, but this will not occur as long as novel content arrives from outside.

If the inference system operated entirely in the fashion just described, working memory would overflow with Skolems, each representing some new object. For this reason, another mechanism flushes literals once their Skolems unify with domain constants. Returning to our example, suppose AbRA has focused on *(cough s2)*. Not only would the rules for *has-flu* and *has-lung-cancer* apply, but so would the justification from the previous step, which yielded *(cough sk1)*. Suppose the system chooses to unify the focus with this prior justification, producing the literals *(has-symptom Ann s2)* and *(cough s2)*. During this step, it recognizes that the new candidate specializes an existing justification, which it then removes. Other justifications that supported the beliefs *(cough sk1)* or *(has-symptom Ann sk1)* are then transferred to their more specific counterparts.

The mechanism we have just described incorporates all the constraints outlined earlier. The set of generated beliefs together with the justifications that link them constitute an explanation that indicates the system's understanding of the facts. The resulting inferences are not deductively valid, but each reasoning step is nevertheless plausible. The process uses rules flexibly, chaining off either antecedents or consequents, and it introduces default assumptions as necessary. Moreover, the approach creates and extends a single explanation in an incremental manner that, because it relies on local computations, is efficient and scalable.

Heuristics for Everyday Inference

The mechanisms described in the previous section rely on heuristics to guide the decisions that drive inference. Since there are often many potential inferences of varying value, we need ways to identify promising ones. Although there is a growing movement to characterize cognition as a statistical process, we hold that there are other psychologically plausible heuristics that inform reasoning. These come into play when selecting a focus of attention and selecting a rule to chain off it.

Our current implementation uses two main heuristics to guide the focus of attention. First, since AbRA works incrementally and assimilates new observations, it prefers observations or inferred belief that are more recent. Even though older beliefs can influence new inferences, they tend not to drive them. Second, we adapt the idea of essential explanations (Fischer et al., 1991), in that the system prefers beliefs that unify with fewer rules in long-term memory. The intuition is that, given a single candidate, rule selection is trivial and having fewer options means there are fewer wrong choices. Moreover, beliefs created by earlier inference steps provide a richer context for evaluating and deciding among later ones.

Our abductive inference mechanism also incorporates heuristics for selecting which rule instance to chain off the focal belief. The main technique is inspired by Thagard's (2007) theory of explanatory coherence. After finding all rule instances that unify with the current focus, AbRA scores these candidates in terms of the average coherence of existing beliefs that match its antecedent and consequents. We define a belief's coherence as the number of existing justifications that include it, plus a boost if it was observed rather than inferred.

To illustrate this measure, consider the domain we introduced in Tables 1 and 2. Suppose AbRA focuses on *(has-symptom Ann s1)*. The rules for (1) *has-flu*, (2) *has-food-poisoning*, and (3) *has-lung-cancer* are all potential candidates. Each of these will create their own set of assumptions, such as *(has-flu Ann)*, *(cough sk)*, and *(has-symptom Ann sk)*, where *sk* is a placeholder for a Skolem constant. However, instances (1) and (2) will also include *(fever s1)*, which is also a fact. If we award observations one point, then the instances (1) and (2) will score 2/5 and (3) will score 1/5. Suppose the system selects (1) as one of the best-scoring candidates. Then during the next cycle, the beliefs *(has-flu Ann)*, *(has-symptom Ann sk)*, and *(cough sk)* will be worth one point, whereas the other two beliefs will be worth two. As beliefs become more connected, their scores increase, which gives them more influence over time.

Experience suggests that locally calculated coherence is insufficient to reliably produce plausible explanations. Often several candidate rule instances have similar scores, especially during the initial stages of inference, leading to ill-informed choices before various strands of an explanation are connected. In response, we introduced another heuristic which favors rule instances that are on a path that links the focus to other beliefs. This *lookahead* procedure starts by finding all rules that could match the focus. If they contain literals that might in turn unify with existing beliefs, then AbRA scores each such rule as if that second unification had occurred. Otherwise, it branches out through rules that might unify with other literals in that rule, recurring until connecting with an existing belief. The rule instance that chained off the focus then receives the coherence score of the leaf of that path. Ties are resolved by extending the process up to a maximum depth. This method has consistently improved the plausibility of explanations.

Empirical Results

We have tested our approach to inference on a variety of domains. One of the least complex involves the rules in Table 1, which specify a folk theory of disease. Given the six facts from Table 2, AbRA reliably generates an explanation that infers both Ann and Bob have the flu. The reasoning involved is nontrivial, and earlier versions would infer different sets of diseases, sometimes assigning multiple diseases to the same person. Introduction of the lookahead procedure coupled with the heuristic for essential explanations stabilized the ability to generate the most plausible solution.

Exploring AbRA's behavior on more complex problems required a larger set of domain rules. For this purpose, we used Ng and Mooney's (1992) knowledge base about a storytelling domain, stated as 107 rules for interpreting events like "Bill gave the bus driver a token" and "Bill went to a liquor store and pointed a gun at the owner." The number of observations per story varies, but each involves around ten literals. Whereas early versions of our system tended to confound events (e.g., inferring that Bill stopped at a restaurant before robbing the liquor store), the version with the stronger heuristics just described produce plausible, well supported beliefs across several examples.

We have also obtained preliminary results on plan recognition, where the system infers the plan that produced a sequence of observed actions when given a hierarchical plan library and those actions. For this evaluation, we selected 100 cases from the Monroe Plan Corpus (Blaylock & Allen, 2005) and encoded the plan library as 50 rules, roughly one for each method. We measured AbRA's ability both to reconstruct the entire original plan and to recover the top-level literal (predicate and arguments) that generated each sequence. To determine how early it can infer a plan during execution, we gave AbRA the first 25%, 50%, 75%, and 100% of the trace. Table 3 reports precision and recall on full plan recovery for each percentage, along with accuracy on recovering the top-level plan. The table also presents results from Raghavan and Mooney (2010) for Markov logic and Bayesian abductive logic programs. These are not directly comparable, being based on a larger sample, but they suggest the three approaches have similar abilities.

Table 3: Precision, recall, and accuracy scores on Monroe plan recognition corpus. *Accuracy* scores for Markov logic and Bayesian abductive logic are based on recall averaged over 1000 cases, with credit awarded for partial matches (Raghavan & Mooney, 2010).

	100%	75%	50%	25%
AbRA (precision)	74%	74%	64%	64%
AbRA (recall)	85%	74%	50%	29%
AbRA (accuracy)	70%	53%	31%	11%
Markov logic networks	79%	37%	17%	7%
Bayesian abductive logic	92%	57%	25%	9%

In summary, our computational model of inference reliably produces plausible inferences in a number of domains using mechanisms that are consistent with the constraints outlined earlier in the paper. This suggests the approach provides a reasonable qualitative account for the main features of everyday reasoning in humans.

Related Research on Abduction

Throughout this paper, we have cast everyday inference in the mold of abductive reasoning (Josephson & Josephson, 1996). This has been a relatively small but important area of cognitive science for nearly four decades. One of the earliest medical expert systems, INTERNIST-I, relied on specialized abductive mechanisms (Miller et al., 1982). Later systems combined the representational power of predicate logic with a weighted form of abduction, where assumptions incur a cost, with applications including language understanding (Hobbs et al., 1993) and plan recognition (Appelt & Pollack, 1992). A number of more recent efforts have replaced these weights with probabilities on rules and assumptions.[2] For instance, Charniak and Goldman (1991) have used Bayesian inference to carry out plan recognition, whereas Kate and Mooney (2009) have adapted Markov logic and developed their own approach, Bayesian abductive logic, for the same task.

Most of these systems address the first four constraints identified earlier, in that they generate explanations through plausible, flexible forms of reasoning that produce default assumptions when needed. However, they all carry out extensive search through the space of explanations, and thus violate the single-explanation constraint. Moreover, they are invariably provided with all observations at the outset, so they do not model the incremental character of human abductive inference. In addition, because these methods carry out global evaluation of candidate explanations, their run times scale poorly with the size of the knowledge base and thus lack

[2] Inference over Bayesian networks involves a form of abduction, but this framework does not support the relational representations required for many reasoning tasks.

the efficiency of human inference. In contrast, our approach satisfies all seven constraints mentioned earlier.

We should note that our focus on coherence rather than on posterior probabilities has been influenced by two earlier efforts. Thagard's (2007) model is guided by this metric, but it compiles candidate explanations into an influence network before evaluating them, so it does not satisfy the constraint of incremental processing. Ng and Mooney's (1992) ACCEL system, which also uses a form of coherence, operates incrementally, but it maintains multiple explanations rather than one account and evaluates them globally rather than locally, which raises concerns about efficiency.

Concluding Remarks

In this paper, we reviewed the main qualitative characteristics of everyday human inference and proposed a computational account of this process. Our framework includes a set of representational assumptions about conceptual knowledge, beliefs, and justifications that relate them. The account posits a control mechanism that alternates between selecting a belief on which to focus cognitive attention and selecting a rule instance that produces new beliefs. In addition, it includes heuristics to guide selection of beliefs and rule instances in ways that drive inference toward coherent explanations. The framework supports a flexible form of plausible reasoning that generates default assumptions; the mechanism incrementally extends a single explanation using local evaluation criteria that support efficient processing. To our knowledge, it is the first computational theory that is consistent with all of these key qualitative features of everyday inference in humans.

Nevertheless, our account remains incomplete in some important ways. The mechanisms cannot detect inconsistencies between previous beliefs and new facts or inferences, and they cannot revise beliefs in ways that repair such problems. Neither can they operate over temporal constraints that specify orderings on components of events. Most important, they cannot represent or reason about the beliefs, goals, and intentions of other agents, which is required by many instances of natural language. However, we believe the framework lends itself naturally to extensions that handle these challenges, and we are working actively to extend our mechanisms to address them. The result will be a more complete account of how everyday inference underlies human cognition.

Acknowledgements

This material is based on research sponsored by ONR under agreement N00014-09-1-1029. The views and conclusions contained herein are the authors' and should not be interpreted as representing the official policies or endorsements of ONR or the U. S. Government, either expressed or implied.

References

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, *111*, 1036–1060.

Appelt, D. E., & Pollack, M. E. (1992). Weighted abduction for plan ascription. *User Modeling and User-Adapted Interaction*, *2*, 1–25.

Blaylock, B., & Allen, J. (2005). Generating artificial corpora for plan recognition. *Proceedings of the Tenth International Conference on User Modeling* (pp. 179–188). Edinburgh: Springer.

Cassimatis, N. L., Bello, P., & Langley, P. (2008). Ability, breadth and parsimony in computational models of higher-order cognition. *Cognitive Science*, *32*, 1304–1322.

Charniak, E., & Goldman, R. (1991). A probabilistic model of plan recognition. *Proceedings of the Conference of the American Association for Artificial Intelligence* (pp. 160–165). Anaheim, CA: AAAI Press.

Fischer, O., Goel, A., Svirbely, J. R., & Smith, J. W. (1991). The role of essential explanation in abduction. *Artificial Intelligence in Medicine*, *3*, 181–191.

Hobbs, J. R., Stickel, M. E., Appelt, D. E., & Martin, P. (1993). Interpretation as abduction. *Artificial Intelligence*, *63*, 69–142.

Kate, R. J., & Mooney, R. J. (2009). Probabilistic abduction using Markov logic networks. *Proceedings of the Workshop on Plan, Activity, and Intent Recognition.*

Laird, J. E., Newell, A., & Rosenbloom, P. S. (1987). Soar: An architecture for general intelligence. *Artificial Intelligence*, *33*, 1–64.

Langley, P., Choi, D., & Rogers, S. (2009). Acquisition of hierarchical reactive skills in a unified cognitive architecture. *Cognitive Systems Research*, *10*, 316–332.

Miller, R. A., Pople, H. E. J., Myers, J. D. (1982). Internist-1: An experimental computer-based diagnostic consultant for general internal medicine. *New England Journal of Medicine*, *307*, 468–476.

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

Ng, H. T., & Mooney, R. J. (1992). Abductive plan recognition and diagnosis: A comprehensive empirical evaluation. *Proceedings of the Third International Conference on Principles of Knowledge Representation and Reasoning* (pp. 499–508). Cambridge, MA.

Peirce, C. S. (1878). Deduction, induction, and hypothesis. *Popular Science Monthly*, *13*, 470–482.

Raghavan, S., & Mooney, R. J. (2010). Bayesian abductive logic programs. *Proceedings of the AAAI-10 Workshop on Statistical Relational AI* (pp. 82–87).

Thagard, P. (2007). Coherence, truth, and the development of scientific knowledge. *Philosophy of Science*, *74*, 28–47.