An Integrated Account of Explanation and Question Answering

Ben Meadows (bmea011@aucklanduni.ac.nz)
Richard Heald (rhea335@aucklanduni.ac.nz)
Pat Langley (patrick.w.langley@gmail.com)
Department of Computer Science, The University of Auckland
Private Bag 92019, Auckland 1142, New Zealand

Abstract

Many high-level cognitive tasks involve *understanding* – the mechanisms by which an agent attempts to construct accurate mental representations of its world. In this paper, we discuss two such processes: explanation and question answering. We propose four theoretical assumptions about representation and processing that arise in these tasks: both involve inference, this inference requires making default assumptions, it occurs in an incremental manner, and it produces structures that can be expressed as directed graphs of conceptual ground literals. We analyze two models of explanation and question answering in terms of these commonalities and evaluate experimental claims about them using reading comprehension passages. In closing, we discuss our findings in light of related research.

Keywords: abductive inference; cognitive systems; explanation; question answering; symbolic reasoning; understanding

Introduction

Understanding is a primary component of high-level cognition. Many cognitive tasks require an agent to update its model of the world based on inferences it makes about connections among its sensory inputs, existing beliefs, and knowledge. *Explanation* is one understanding task, in which an agent assembles its perceptions of the world into structures that integrate coherently into some model. *Question answering* is a similar task, in which an agent maps a query about the world onto a model, extending it as needed to produce a plausible answer. In this paper we report an integrated account of behavior on these two tasks.

As a motivating example, consider this passage from a first grade reading comprehension book (Liscinsky, 2010):

```
[1]  The spider wants food.
[2]  She likes to eat bugs.
[3]  So she makes a trap.
[4]  The spider spins a sticky web.
[5]  Then she waits.
[6]  The web is hard for bugs to see.
[7]  Whap!  A bug flies into the web.
[8]  The web shakes when the bug lands.
[9]  The bug is stuck.
```

We can elicit several ideas from this extract. Some facts are not overtly stated, but rather implied by others (e.g., the bug has wings). The same holds for causal content (e.g., the spider is going to eat the bug) and constraints (e.g., the conditions under which a spider can ensnare a bug).

This suggests some reasonable questions to ask about the domain. *What is the spider going to eat? How did the spider catch the bug? Does the bug have wings?* Furthermore, it suggests plausible structural components of explanation. Content words (e.g., 'food', 'spins') indicate concepts and relations that are organized in recognizable patterns which constitute prior knowledge (e.g., '*?x* is an instance of *?y* being stuck to *?z*'). We can picture the elements as nodes and knowledge as sets of labeled edges in a directed graph.

We build on these ideas in the following sections. First we propose some theoretical principles about the connections between explanation and question answering as complementary forms of understanding. We then describe two systems that address these tasks, make empirical claims, and test them experimentally. We draw comparisons with related research and conclude with plans for future work.

Theory and Commonalities

Take the illustrative example of the spider building a web to catch a fly. Suppose an agent possesses a piece of structural knowledge *K*, stating "a mobile entity that cannot detect an object may touch it by accident". The agent may infer that the bug flew into the web by accident. Making the inference requires the agent to (a) identify the rule *K* in its knowledge base as one it can usefully apply, (b) instantiate the rule correctly (e.g., *?entity=bug* & *?object=web*) using known values, and (c) introduce the assumption that the bug is mobile.

Alternatively, we may ask the question: How did the bug come to fly into the web by accident? The components – a high-level query and a predictable output (a pattern of concepts and events) – match the same piece of knowledge, *K*. The agent may decide it is reasonable to make a single supporting assumption, then assume that the bug is mobile, instantiate the argument, and return the result as an answer.

The first example involves explanation: interpreting observations in terms of what is already known, filling in gaps as necessary, and thereby incrementally adding higher-level patterns. The second concerns question answering: interpreting evidence to fit a query by using the question as input and undergoing an analogous process. Both are directed explorations that produce interconnected patterns by applying knowledge to beliefs (when there are repeated iterations beyond the single step contained in the example above). We can now define these tasks more precisely:

- *Explanation* is a cognitive task in which an agent is:
  - *Given* a sequence of input observations and a corpus of hierarchically organized knowledge, and
  - *Finds* an explanation represented as a directed graph of concepts and rule instances.

- *Question answering* is a task in which an agent is:
  - *Given* a set of query elements, a corpus of hierarchically organized knowledge, and zero or more initial facts, and
  - *Finds* an elaboration of the query that expands an existing explanation to provide a coherent answer.

We have claimed that these tasks do not just bear surface similarities, but have parallel structures. We will expand on this idea by discussing their representational and processing components – first defining the cornerstones of a computational approach that incorporates these commonalities and then describing specific implementations. To this end, we have composed four theoretical tenets that we argue are shared by question answering and explanation.

1. *Both tasks involve inference* in that they generate a set of interconnected beliefs about facts. Information is often incomplete and must be extended by using knowledge. In our example, the answer to the question "why did the spider make a web?" is not stated, yet is 'obvious'. In explanation, inference is a largely bottom-up process of interpreting low-level facts and observations. Question answering instead is a top-down process that reconciles high-level question elements with existing beliefs.
2. *Inferred structures are organized into directed graphs of working memory elements.* Understanding requires the ability to cache information in a short-term belief store that tracks basic conceptual elements. Long-term knowledge, comprising rules (patterns of generalized elements), is organized in a hierarchical manner. A rule relevant to the bug's flight into the web might be

   is-a(?x, accidental-contact) ⇐
   is-a(?x, contact) &
   attribute(?x, agent, ?a) & *attribute(?x, recipient, ?r)* &
   attribute(?x, intentional, false) &
   is-a(?y, cannot-see) &
   attribute(?y, agent, ?a) & *attribute(?y, recipient, ?r)* &
   attribute(?a, mobile, true)

   where each attribute literal is a terminal node, *contact* and *cannot-see* denote relations specified by other rules, and so on. Both explanation and question answering produce the same type of interconnected rule instances.
3. *Inference requires the introduction of default assumptions.* Deductive proofs are typically viable only in abstract or closed-world scenarios. In the real world, there may be substantial but not *conclusive* support for a belief. We hold that human understanding employs a form of everyday reasoning that involves abductive inference (Bridewell & Langley, 2011). Assumptions should be *reasonable* in that they meet criteria such as consilience, parsimony, and coherence. *The spider is alive* is more consistent with other beliefs than *the spider is a lifelike robot*.
4. *Inference occurs in an incremental, on-line manner.* This reflects the idea of a cognitive cycle (Young, 2001) that involves processing of a single structure, such as application of a rule, so that elements are added to working memory incrementally. Rule instances become interconnected, with later questions being answered using inferences or assumptions from earlier queries. For example, the element *the spider notices the bug is trapped in its web* should reduce the search required for the *the spider eats the bug*.

Together, these four suppositions provide a theoretical framework for understanding that supports both explanation and question answering. We will now turn to two specific instances of this theory that model behavior on these tasks, describing their representation and mechanisms.

Two Complementary Models of Understanding

We have developed two computational models that address different but complementary aspects of understanding: UMBRA (Meadows, Langley, & Emery, 2013, 2014), an account of explanation, and PHOS, a model of question answering. We developed UMBRA to address the data-driven construction of explanations that occurs naturally as one observes a stream of events, and we implemented PHOS to handle the process of question answering that builds upon these explanations.[1] The two systems share the theoretical underpinnings just described, including representational assumptions that we will review before discussing their mechanisms.

Representations for Understanding

We designed UMBRA and PHOS to operate over the same representational structures. Each system has a working memory that contains beliefs encoded as relational structures like '*x is an instance of y*', '*j has the attribute-value pair* <*a,b*>', or '*n and m are not equal*'. These elements may contain constants, identifiers for other elements, or skolemized placeholders for unknown values. In our scenario, we can express the bug becoming stuck with the elements *is-a(skolem1, trapped) & attribute(skolem1, agent, bug) & attribute(skolem1, object, web) & attribute(skolem1, actor, spider)*, much as in a semantic network (Gentner, 1975). Moreover, working memory organizes facts, inferences, and assumptions into connected structures that we call *explanations*.

Questions are encoded in a similar manner, as sets of connected, partially instantiated elements in working memory.[2] For example, {*is-a(?s, spider)*} denotes 'what spiders exist?', while {*is-a(?c, cannot-see), attribute(?c, agent, ?a), attribute(?c, recipient, web1), attribute(?c, cause, ?y)*} encodes 'who cannot see web1, and why?' A question typically involves some informative elements (e.g., 'spider', 'recipient') and some unknown 'answer' values (e.g., '?s', '?y'). Answers are instantiations of questions whose constant values appear throughout the explanation graph produced by the

[1] *Phos* is the Greek word for light, which reveals what lies in the shadows but also cast new shadows in the process.

[2] Our work does not focus on natural language processing; we assume that questions have already been translated into internal structures expressed as connected sets of elements. However, we do not assume that word sense is provided, so our system must utilize knowledge to interpret questions it encounters.

question-answering process. These answers may be partly ungrounded if a complete answer is not found.

UMBRA and PHOS also draw on a common long-term memory that contains relational knowledge structures. Knowledge is encoded in a conceptual hierarchy in which higher-level nodes are specified in terms of lower-level nodes. Rule components include the patterns *is-a(?x, ?type)*, *attribute(?x, ?label, ?y)*, *(?a≠?b)*, and *(?a=?b)*. We assume that neither system will encounter attributes with multiple values, such as *attribute(?1, label, value1)* & *attribute(?2, label, value2)* & *(?value1≠?value2)* ⇒ *(?1≠?2)*. The previous section presented a simplified example of a conceptual rule for *accidental-contact*.

A Model for Generating Explanations

UMBRA models the generation of explanations. We provide the system with conceptual knowledge and a sequence of facts and observations in the notation just described. Explanations are directed graphs comprising domain literals that include input elements as well as other, similar, ones that have been inferred or assumed. These elements may be only partially instantiated, and they are connected by instances of rules that have been applied. For example, the literal *is-a(web1, location)* may appear in several different rule instances – in defining a web, or waiting at a place, or a locus of entrapment. In this way, an explanation can be structurally cohesive.

UMBRA constructs its explanations through an incremental, data-driven form of abductive inference (Meadows et al., 2014). This process operates through a sequence of high-level 'observation' cycles, each of which begins with the acquisition of new beliefs from observations and then expands the explanation graph through a number of inference steps. The system has a resource bound that limits the inference steps it carries out on each observational cycle. Rules that require more assumptions have higher cost, and the system stops chaining when it exceeds a threshold.

On each inference step, UMBRA generates a set of candidate rule instances that match against at least one element in working memory. It calculates the additional inferences and assumptions needed for each rule, using this as the basis for a cost metric. The system uses high-level control knowledge to prune candidates that violate existing constraints, or that would not sufficiently improve explanation consilience. It evaluates each remaining rule instance according to the cost metric and selects the least-cost remaining candidate. UMBRA applies this rule instance, introducing new elements (default assumptions) and extending the explanation.

A Model of Question Answering

We developed PHOS to model the process of question answering. We provide the system with conceptual knowledge, a set of existing beliefs organized into an explanation structure, and a question. It produces an elaborated explanation with additional inferences that connects the question to prior beliefs and provide an answer. An input explanation is not strictly necessary. The system may receive a set of beliefs that are not linked by rule instances or, in extreme cases, it may receive no initial beliefs at all.

Table 1: Summaries of the five sample scenarios. Page numbers from the source (Liscinsky, 2010) are given in *italics*.

- ○ A hungry spider spins a web and a bug blunders into it. *(44)*
- ○ A zoo is described as containing various animals and activities, including mythical or anthropomorphized ones. *(108)*
- ○ A hippo stays in water during the day when the sun is hot, then comes out at night to eat. *(51)*
- ○ Insects and spiders are compared and contrasted. *(104)*
- ○ A pig enjoys rolling around in a cool mud puddle. *(17)*

PHOS operates in a series of high-level query cycles, each of which represents a discrete attempt to find a coherent answer. These in turn comprise a number of inference steps. A query cycle begins by initializing a set of candidate literals – the *fringe* – with the contents of the original question. The system then incrementally extends the current query graph. On each inference step, PHOS selects an element from the fringe to focus on, then enumerates the rule instances, assumptions, or unifications with which it could be supported. The system calculates a cost for each candidate. A coherence heuristic uses these costs to select a candidate at random with probabilities proportional to their fitness, thus providing search control that guides expansion of the query graph. The inference step ends by adding the candidate's elements to working memory. In contrast to UMBRA, this process operates in a top-down manner. However, whenever the system extends the explanation, new instantiations are propagated back up the graph to the original query elements.

If PHOS processes every element in the fringe, then it has found an answer and the query cycle ends. Otherwise, it halts when the accrued total cost becomes too high. The system retains the working memory elements and rule instances involved in the answer's elicitation. Importantly, this includes the original questions, which very often carry implicit meaning (e.g., asking "Did the hungry spider sleep?" introduces the idea that there *is* some spider that is hungry), and thus are useful resources for later, related questions. If a question is repeated later, PHOS can unify its components with the answer elements from the existing structure, resulting in pattern-based retrieval.

We have shown how these two models of high-level understanding instantiate the theoretical basis that we outlined. UMBRA uses an incremental, data-driven form of abduction that introduces default assumptions in an effort to produce a cohesive explanation. PHOS uses an incremental, query-driven form of abductive inference to generate meaningful answers that it incorporates into a new or existing explanation. Now that we have shown how these models align with our core theory, we turn to analyzing their behavior, in particular how the processes of explanation generation and question answering interact.

Empirical Studies

UMBRA and PHOS are computational models that instantiate the theoretical tenets we presented earlier. As such, we can examine their behavior in particular scenarios to reveal interactions between explanations and question answering. Our aim is not to fit quantitative measures like error rates or reaction times. Instead, we desire to demonstrate that, taken together, the two models produce behavior that is qualitatively similar to that observed in humans. In summary, we adopt a *cognitive systems* approach (Langley, 2012) that studies the behavior of integrated computational artifacts.

Part of our evaluation involves running the explanation system on observations and then asking questions about the altered memory state. This approach follows Zelle et al. (1994), who measured the performance of an integrated system for language understanding in terms of its final outputs. By measuring answer accuracy, we can quantify the success of our two models operating in tandem. We are interested in how explanation generation and question answering interact. We make three main empirical claims:

1. **Explanation construction and question answering are complementary and commutative.** Computational resources spent on explanation can reduce the cost of subsequent question answering activities and vice versa.
2. **Question answering centrally involves inference.** Retrieval processes are crucial to high-level understanding, which in turn relies on processes like abductive inference.
3. **Interference impacts understanding, but can be overcome.** Confounding information escalates the cost of processing, increasing the cycles or resources required, but one can still disregard superfluous facts.

Although we cannot provide details here, these claims follow directly from our two models and they also parallel known aspects of human understanding.

Experimental Design

Reading comprehension is a good task for evaluating understanding, in that passages contain elided information, utilize conceptual knowledge, and have relevance to human cognition. We therefore worked with five scenarios from a first-grade text for reading comprehension (Liscinsky, 2010).[3] Table 1 summarizes these passages.

We translated the five vignettes into logical literals, producing scenarios with varying characteristics. For instance, the *comparing insects to spiders* vignette describes only domain rules, with no initial facts, while another had 127 ground facts; the mean was 45.4. We extracted content from the text manually, encoded it using these literals, and then generated a set of questions. For example, "Is there any insect with eight legs?" translates to the conjunction *is-a(?x, insect) & is-a(?p, has-property) & attribute(?p, entity, ?x) & attribute(?p, property, limbs-legs) & attribute(?p, number, 8)*. We generated three different questions for each scenario.

We measured PHOS's performance by counting cases in which the correct answer was returned and cases in which spurious inferences occurred. We calculated precision and recall scores from those metrics. We also estimated computation time per answer with the abstract cost thresholds used by both UMBRA and PHOS. We measured cognitive cycles per answer, but resource consumption predicted this metric very closely, so we do not report it here.

Target answers varied in size from a single element ("no") to graph structures containing more than 50 elements. We repeated each question twice to compensate for minor effects from the nondeterministic heuristics, although in practice we saw very few differences across repeated runs. In all cases, we provided both systems with the full complement of 60 domain rules derived from all scenarios.

Claim: Understanding Tasks are Complementary

Our first claim was that the work done by explanation reduces the amount of effort question answering requires and vice versa. Intuitively, making more initial inferences may reduce the effort required to find an answer later. This hypothesis is informed by theories of different types of elaboration in human reasoning (e.g., Bradshaw & Anderson, 1982). In the extreme case, question answering works without prior explanation as long as it has sufficient computational resources.

To test this claim, we ran UMBRA on each domain, then ran PHOS on its outputs. We systematically varied the resources allocated to each system, running UMBRA at zero, scarce, or plentiful levels, and PHOS with scarce or plentiful levels.[4] We ran the systems in sequence a total of 180 times, ignoring UMBRA's incorrect inferences (because these may sometimes translate into inputs that PHOS regards as inconsistent, in which case the system halts). We found that, when PHOS had scarce computational resources, recall scores increased as UMBRA's resources increased, from 0.17 to 0.33 to 0.37. When PHOS had plentiful resources, recall remained at 0.55, consistently higher than the scarce case. We also found that increasing UMBRA's resources reduced mean cost per answer by 40 to 60 percent. Together, these results support our first claim.

Claim: Question Answering involves Inference

Our second claim was that inference is central to the question-answering process; understanding depends on the organization of ground facts and supporting assumptions into known patterns, so that reasonable answers can be produced by common-sense reasoning where information is elided. We tested this premise by removing PHOS's abductive inference capability, eliminating the step at which the model can choose to construct a subquery by matching an element from the fringe with a rule head. We ran it on the test scenarios with

[3] We chose five passages that the author categorized differently – 'compare and contrast', 'fantasy and reality', 'prediction', 'main idea', and 'what and how' – to maximize the range of scenarios.

[4] The systems' internal notions of 'processing resources' are not identical, but they are directly analogous.

plentiful resources, expecting the lesioned system would be unable to function effectively. In previous work, we have reported similar studies with the UMBRA model in isolation (Meadows et al., 2013).

In this case, we found that the system only succeeded on 13 percent of the runs, and then only because it defaulted to 'false', the correct answer for two questions. This is evidence that it could not answer the questions with straightforward retrieval. PHOS was also able to perform more direct retrievals after running UMBRA. These results support our hypothesis that inference is crucial to effective question answering.

Claim: Interference Effects can be Overcome

Our final claim stated that confounding information necessitates more processing to find answers. Intuitively, we expect that trying to compensate for noise will impose overheads: consider making the decision to ignore an improbable outlier or separating relevant and irrelevant ground facts. To test this premise, we combined the 227 initial elements from all the scenarios. Running PHOS with plentiful resources, we found that confounders reduced the mean recall score from 0.53 to 0.37, often due to finding an answer with faulty low-level bindings rather than wrongly instantiating the top-level query elements or not answering. This means the model sometimes incorporates extra information in an unreasonable way.

The mean cost per answer, surprisingly, decreased from 1499.5 to 762.3, contradicting our hypothesis. Analysis revealed that rather than requiring more cognitive cycles before it found a suitable answer, PHOS often managed to incorporate confounders in ways that seemed consistent, reducing the need for expensive default assumptions. For example, a sparsely-grounded conceptual relation $?x$ might be inferred to be an instance of not only *eating*, but also *having a goal* – an inappropriate combination leading to a wrong answer, but not specified in knowledge as a contradiction. Our third claim, then, is partially refuted: the model was nontrivially affected by confounders. We plan to ameliorate this effect by providing it with more discriminating conceptual knowledge.

Remarks on the Evaluation

Overall, our combined model of explanation generation and question answering behaved as expected. Our experiments with UMBRA and PHOS supported our first two claims about interactions among these two processes, although our third study uncovered some surprises. Nevertheless, taken together, they suggest that our models provide a viable instance of our computational theory of understanding.

Note that our measurements were conservative: we only scored an answer as right or wrong, so a single low-level error produces a negative score even if the overall answer is plausible. Top-level query elements could be perfect matches and still be considered 'misses'. Indeed, we observed that almost every spurious answer incorporated some literals from the canonical answer; a finer-grained scoring system would have reported higher accuracy.

Related Research

Our framework shares elements with previous research but also has distinctive features. For example, work in the paradigm of plan recognition (Goldman, Geib, & Miller, 1999; Bridewell & Langley, 2011) has focused on generating explanations of observed behavior, often inferring agents' goals using abductive mechanisms on hierarchical knowledge structures, but it has not addressed the related task of question answering. Winston's (2012) research on story understanding has a similar flavor, encoding knowledge as rules and explanations as elaboration graphs much like our structures, but, again, has not addressed question answering.

At the other extreme, some computational models of human memory include accounts of question answering but do not touch on explanation generation. Anderson and Bower (1980) offer a detailed account for the retrieval of facts from memory in response to questions, but their storage process involves no inference. Graesser et al. (1991) report a more sophisticated model of question answering that incorporates criteria similar to ours, such as coherence. However, their work assumes that all content used in this process is already stored in memory. Waldinger et al. (2011) describe an applied system that, given access to online databases, combines language processing with deduction to answer medical questions. Narayanan and Harabagiu (2004) present an alternative that uses probabilistic inference to answer questions given predicate-argument descriptions of a large corpus of sentences from the Wall Street Journal.

Research on natural language processing during the 1970s and 1980s dealt with both capabilities, but with somewhat different emphases. Lehnert (1978) and Dyer (1983) both described models that constructed explanations of narratives and that chained over the resulting structures to answer questions, but the latter processes did little to extend the explanation.[5] They adopt script-based representations for knowledge and explanations that provide more structure than does our formalism, but their systems also utilize a constrained variety of abductive inference to generate explanations. Kolodner's (1983) model of reconstructive memory comes even closer to our own, in that it makes many knowledge-based inferences at the time of question answering. More recently, Barbella and Forbus (2011) report a system that uses analogical reasoning to generate inferences when answering questions; this provides a form of abductive inference that operates over cases constructed during reading rather than over rules.

In summary, the literature includes many efforts on generating explanations and on answering questions, but only a small number of computational models that address both of these cognitive tasks. Of these, even fewer carry out substantial inference during question answering, and those draw upon different representations and processes than we have proposed in our framework.

[5] However, Dyer mentions in passing that answering a question, his model can modify memory as an unintended side effect.

Concluding Remarks

In this paper, we argued that the cognitive tasks of generating explanations and answering questions are two complementary aspects of understanding that share many facets. We introduced four theoretical assumptions about the representations and processes that underlie them, then described computational models for behavior on each task that incorporate these tenets. In addition, we reported experiments on reading comprehension scenarios, showing that our abductive mechanisms produce effective understanding and testing claims for interactions between explanation and question answering.

In future work, we plan to build upon our initial results by extending PHOS to better handle confounding information, thus improving scalability, and to interact more dirctly with UMBRA. Specific areas of interest include modeling word sense disambiguation and the influence of inferences generated when answering an early question on responses to later ones. We should also develop an approach to answering open-ended questions, such as "what happens next?", which our current knowledge structures cannot handle.

Acknowledgments

This research was supported in part by Grant N00014-10-1-0487 from the Office of Naval Research. Pat Langley is also affiliated with the Institute for the Study of Learning and Expertise. We thank Paul Bello, Will Bridewell, and Miranda Emery for useful discussions that influenced the approach we have reported here.

References

Anderson, J. R., & Bower, G. H. (1980). *Human associative memory: A brief edition*. Hillsdale, NJ: Lawrence Erlbaum Publishers.

Barbella, D., & Forbus, K. D. (2011). Analogical dialogue acts: Supporting learning by reading analogies in instructional texts. In *Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence* (pp. 1429–1435). San Francisco: AAAI Press.

Bradshaw, G. L., & Anderson, J. R. (1982). Elaborative encoding as an explanation of levels of processing. *Journal of Verbal Learning and Verbal Behavior*, *21*, 165–174.

Bridewell, W., & Langley, P. (2011). A computational account of everyday abductive inference. In *Proceedings of the Thirty-Third Annual Meeting of the Cognitive Science Society* (pp. 2289–2294). Boston: The Cognitive Science Society.

Dyer, M. G. (1983). *In-depth understanding: A computer model of integrated processing for narrative comprehension*. Cambridge, MA: MIT press.

Goldman, R. P., Geib, C. W., & Miller, C. A. (1999). A new model of plan recognition. In *Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence* (pp. 245–254). San Francisco: Morgan Kaufmann.

Graesser, A. C., Lang, K. L., & Roberts, R. M. (1991). Question answering in the context of stories. *Journal of Experimental Psychology: General*, *120*, 254.

Kolodoner, J. (1983). Reconstructive memory: A computer model. *Cognitive Science*, *7*, 281–328.

Langley, P. (2012). The cognitive systems paradigm. *Advances in Cognitive Systems*, *1*, 3–13.

Lehnert, W. (1978). *The process of question answering*. Hillsdale, NJ: Lawrence Erlbaum Publishers.

Liscinsky, C. (2010). *Reading comprehension*. Monterey, CA: Evan-Moor.

Meadows, B., Langley, P., & Emery, M. (2013). Seeing beyond shadows: Incremental abductive reasoning for plan understanding. In *Proceedings of the 2013 AAAI Workshop on Plan, Activity, and Intent Recognition* (pp. 24–31). Bellevue, WA: AAAI Press.

Meadows, B., Langley, P., & Emery, M. (2014). An abductive approach to understanding social interactions. *Advances in Cognitive Systems*, *3*, 87–106.

Narayanan, S., & Harabagiu, S. (2004). Question answering based on semantic structures. In *Proceedings of the Twentieth International Conference on Computational Linguistics* (pp. 693–702). Geneva: Association for Computational Linguistics.

Waldinger, R., Bobrow, D. G., Condoravdi, C., Richardson, K., & Das, A. (2011). Accessing structured health information through English queries and automatic deduction. In *AI and Health Communication: Papers from the 2011 AAAI Spring Symposium.* Stanford, CA: AAAI Press.

Winston, P. H. (2012). The right way. *Advances in Cognitive Systems*, *1*, 23–36.

Young, R. M. (2001). Production systems in cognitive psychology. In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social and behavioral sciences.* Oxford, UK: Pergamon Press.

Zelle, J. M., Mooney, R. J., & Konvisser, J. B. (1994). Combining top-down and bottom-up techniques in inductive logic programming. In *Proceedings of the Eleventh International Conference on Machine Learning* (pp. 343–351). New Brunswick, NJ: Morgan Kaufmann.