# The Heuristic Reasoning Manifesto

**Praveen K. Paritosh (paritosh@cs.northwestern.edu)**
Qualitative Reasoning Group, Northwestern University, Evanston IL 60208.

**Abstract**

We argue for heuristic reasoning as a solution to the brittleness problem. Heuristic reasoning methods exploit the information processing structure of the reasoning system and the structure of the environment to produce reasonable answers when knowledge and/or computational resources for finding the perfect correct answer might not exist. Capturing all the heuristics to generate reasonable answers might not be as colossal of a project as it might first seem: we conjecture that there are about fifteen *heuristic domains*, and each of them have approximately ten *heuristic methods*.

## 1 Introduction

Brittleness is a serious problem for most AI programs, and perhaps software in general. We propose *heuristic reasoning* as a solution to the brittleness problem. We are inspired by the everyday human ability to generate educated guesses, reasonable explanations and ballpark estimates when we run into situations where knowledge and/or cognitive resources are lacking. Human reasoning is more flexible and scales better than any artificial reasoning system built. For example, in contrast to most systems, we get more fluent, not slower, as we know more about a domain. There is more than thirty years of research in psychology of judgment and decision making (Tversky and Kahneman, 1974; Hammond et al., 1980; Gigerenzer et al., 1999) which has produced a rich catalog of heuristics used by humans. Heuristic reasoning methods are not limited to these psychological heuristics, but at this stage, we conjecture that the psychological heuristics are an important subset. Equipping our programs with heuristic reasoning can make them less brittle, and help them guide and double-check other reasoning processes. There are two overlapping goals of this work:

1) Computational: Build flexible reasoning systems that can make educated guesses when other methods fail.
2) Cognitive: By implementing the psychological heuristics, explore the architectural, structural and representational assumptions underlying them.

We begin with a discussion of the brittleness problem and ground it in the context of knowledge-based systems. The next section presents a brief review of related work. We then present an analysis of heuristic methods: domains where they might or might not work, and our hypotheses about how and why they work. We present nine domains where we believe heuristic methods can be leveraged, one of which has been explored in our earlier work (Paritosh and Forbus, 2001, 2004, 2005). We end with a discussion and conclusions.

## 2 Brittleness in Knowledge-based Systems

The two common manifestations of brittleness are: 1) the software cannot find an answer, because of gaps in the knowledge base, or because of a lack of required computational resources; and 2) the software comes up with an unreasonable answer, possibly because of inaccuracies in its knowledge base. For instance, in an evaluation of question-answering programs that mine text for answers, one program came up with 360 tons as the amount of Folic acid that an expectant mother should have per day, and 14 feet as the diameter of the earth![1]

We focus on knowledge-based systems. Knowledge-based systems consist of reasoning mechanisms that use an explicit *knowledge base*, a database of facts, to answer queries. However, these arguments might apply even more broadly. Figure 1 shows a highly simplified view of a knowledge-based system. The reasoning mechanisms might consist of forward and backward chaining, planning, analogy, spatial reasoning, and special-purpose procedural attachments to handle specific tasks. Many of these reasoning methods are computationally complex, and in theory can take unbounded amounts of time. However, a crucial bottleneck for these reasoning mechanisms is the knowledge base. If the knowledge base has gaps, i.e., lacks relevant knowledge, then there is no hope of being able to find an answer.

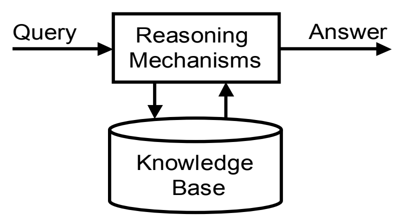


**Figure 1: A simplified schematic of Knowledge-based Systems**

Cyc, the largest knowledge representation effort, consists of over 3 million assertions represented in predicate calculus. Yet, the *brittleness bottleneck* (Lenat et al., 1986) is far from overcome. One premise of the Cyc project is that by explicitly representing the commonsense knowledge that a six-year old has, we can build more flexible systems, where commonsense fills in the gaps when the system comes to a point where it would otherwise exhibit brittle behavior. Although a test comparing a six-year old to Cyc has not been carried out, based on our experiences with using it, we conjecture that

[1]The question is from TREC9, and this was reported in the IBM TJ Watson AQUAINT Briefing.

it is not yet close to the flexibility of the six-year old; even though it might be capable of making very sophisticated inferences in some domains.

Openmind Commonsense[1], another such effort, consists of 800,000 assertions in English authored by volunteers on the web (Singh et al., 2002). The innovative idea in this project is that by lowering the barrier to knowledge authoring, it might be possible to quickly build a large collection of commonsense knowledge. However, the problem of inaccuracies in the knowledge base is a serious problem as about 30% of the knowledge is “garbage” (Lieberman, personal communication). Furthermore, it supports very weak notions of reasoning with these facts, if at all.

A broad commonsense knowledge base is necessary for building robust programs. However, commonsense might be much vaster than imagined, and approaches to building large databases of knowledge is not enough by itself. One solution for brittleness is building reasoning methods that scale with respect to the amount of relevant knowledge available. The heuristic reasoning approach presented in this paper is about operationalizing patterns of reasoning that can flexibly handle gaps in the knowledge base at the cost of being right most, not all, of the time.

## 3 Related Work

Let's begin with the most popular senses in which heuristics has been discussed in the literature. George Polya (1945) popularized heuristics in his book as various possible steps one could take while solving mathematical problems. Some of his heuristics included drawing a figure, working backward from what is to be proved and considering a more general version of the problem. The final output in such reasoning is sound mathematical statements, however the heuristics help explore the space in a clever way. Herb Simon coined the notion of *Bounded Rationality* and *Satisficing* (1957). In this approach, reasoning is still governed by laws of rationality and realistic resource constraints are placed on it. Newell and Simon (1963) proposed weak methods, e.g., means-end analysis, generate and test, etc., as the basis of intelligence.

Doug Lenat's AM and Eurisko (1982) systems made scientific discoveries in the domain of mathematics, device physics, games, and heuristics itself, among others, armed with a library of hundreds of heuristics. Lenat called for a formal study of the science of heuristics, heuretics. However, Lenat's notion of heuristics is different from ours. The goal of his systems was to make interesting scientific conjectures, and his heuristics guided exploring the space. For example, one of his heuristics would suggest that if a function $f(x,y)$ takes two arguments, then its worth the time and effort to define and explore the behavior of $g(x)=f(x,x)$, that is, to see what happens when the arguments coincide. If $f$ is multiplication, this new function $g$ is squaring; if $f$ is union or intersection, then $g$ is the identity function, and so on. His notion of heuristics was ways to branch out and explore the space in some guided ways. This is different from the way we are framing heuristic reasoning: our goal is to be focused and generate answers quickly.

[1] http://openmind.media.mit.edu/

In the 1970s, the psychologists Amos Tversky and Daniel Kahneman started the *Heuristics and Biases* program. The goal in this program was to use peoples' systematic biases in judgment under uncertainty to reveal the heuristics they use. This led to a large body of literature in Psychology exploring various aspects of intuitive reasoning. Recently, Gerd Gigerenzer and his team (1999) have made compelling arguments for *fast and frugal heuristics*, in which they view the mind having an adaptive toolbox of heuristics that work because of the way the environment is structured. One of their heuristics is the *recognition* heuristic: something that you can recognize is likely more important than something you don't. In a study where both a sample of German and US students were asked questions about cities like “Which is bigger: San Antonio or San Diego?” they showed that Germans performed significantly better than Americans on American cities and vice versa for German cities. Their argument is that with lesser knowledge of American cities, German students can invoke the recognition heuristic to pick the answer that is most likely going to be right, while American students cannot use that heuristic as they probably have heard of both cities. However, their focus is on populating this toolbox and not on figuring out how this might be integrated with other cognitive functions.

## 4 Heuristic Reasoning

Heuristic reasoning exploits the information processing architecture of the reasoning system (in the case of psychological heuristics, the human mind), and the structure of the world to generate reasonable answers. When does heuristic reasoning work, and how many heuristics are there? We begin with some definitions. A *heuristic domain* is a reasoning task that is amenable to heuristic reasoning. A *heuristic method* is a specific pattern of reasoning that yields a reasonable inference in its heuristic domain.

Let's consider an example of a heuristic method. Suppose you were asked, “What American company sells the most greeting cards?” One way to answer the question might be to look up statistics about sales of various greeting card companies. However, a typical answer might look more like the following:

> “Let's see... Hallmark comes to mind. I have seen Hallmark cards all over the place. In fact, I can't think of any other major greeting card manufacturer, so I bet it's Hallmark.”

The above answer and rationale appear reasonable to most people, and in most circumstances such reasoning is right[2].

[2] Hallmark's revenue is approximately $5 billion, its rival American Greetings' revenue is around $2 billion.

It exploits an important fact about human memory: the ease with which we can recall instances of something is usually correlated with the frequency of that thing in the world, and unheard-of things are often not very important.

Reasoning tasks where there are multiple answers and/or processes to arrive at the answer, with varying degrees of correctness or quality are heuristic domains. On the other hand, questions like “What two US biochemists won the Nobel prize in 1992?” or “What is the scientific name of Viagra?” are examples for which it is less likely to have reasonable guesses – you either know the answer or don't. Both of these questions are from the TREC[1] corpus, which places more emphasis on such questions than on those that require reasoning/inference. Figure 2 shows an abstract characterization of various reasoning domains by plotting the quality of an answer with varying amounts of knowledge and computational resources. Note that such a graph cannot be really drawn, as we will rarely have enough data points, and both X and Y axes represent complicated multi-dimensional concepts: it is simply used here to indicate the nature of heuristic reasoning processes. R1 denotes a reasoning task from a brittle domain, where one can either produce an answer, or completely fail. Both R2 and R3 represent heuristic domains, where with decreasing resources, one can still produce answers, though of decreasing quality. Note the difference between R2 and R3: R2 has a sweet spot, and with much less resources produces a high quality answer, while R3 doesn't. One of the goals of this project is to have a deeper understanding of heuristic domains in such terms.

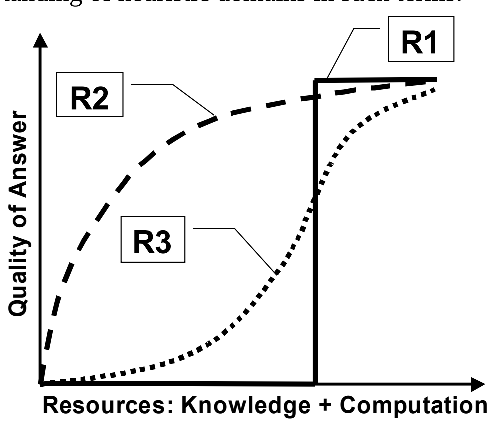


Figure 2: Characterization of different types of reasoning tasks.

The goal of the proposed research program is to build a complete set of heuristic domains and heuristic methods. Another way to express this goal will be to say that we want to codify all the processes that underlie the human ability of making educated guesses and coming up with reasonable answers. Because of the way heuristic domains and methods carve up the reasoning processes and their level of abstraction, we make the following conjecture about the magnitude of this program:

*There are approximately fifteen heuristic domains, and each has about ten heuristic methods that achieve broad coverage in that domain.*

The figure of fifteen is not sacrosanct, it is based on our efforts to build an exhaustive list from analysis of problem solving in multiple domains and the literature on psychology of human problem solving, judgment and decision making. We don't believe that the list of heuristic methods and domains that we later present are all there is, but lest the reader consider this to be an inexhaustible set, the above conjecture concretizes our belief about the magnitude of this research program. The next section presents the first heuristic domain, back of the envelope reasoning, the domain of making rough quantitative estimates, where we have achieved broad coverage with a small set of heuristic methods.

## 4.1 Back of the Envelope Reasoning: A Heuristic Domain

Back of the envelope (BotE) reasoning involves generating quantitative answers in situations where exact data and models are unavailable, and where available data is often incomplete and/or inconsistent. Such reasoning is a key component of commonsense reasoning about everyday physical situations. In our previous work, we presented arguments for why BotE reasoning is important and practical (Paritosh and Forbus, 2001). We presented the design of BotE-Solver, a general-purpose problem solving framework that uses estimation strategies, the ResearchCyc knowledge base, and keeps track of its problem solving progress in an AND/OR tree (Paritosh and Forbus, 2004). The power of BotE-Solver comes from its strategies that enable it to come up with an answer even when none can be found using standard methods. A strategy transforms a given question into other, possibly easier questions. A key contribution of this work is that a core set of seven strategies provides broad coverage, and is possibly the complete set of back of the envelope problem solving strategies. There is twofold support for this hypothesis: 1) an empirical analysis of all problems (n=44) on Force and Pressure, Rotation and Mechanics, Heat, and Astronomy from Clifford Swartz’s (2003) book, “Back-of-the-Envelope Physics,” and 2) an analysis of problems solved by BotE-Solver.

A BotE question asks for an estimate of a quantity for some object, which can be abstractly stated as (Q O ?V), where Q is the quantity, O the object, and ?V the unknown value. This suggests three syntactic transformations, namely, transforming the object, quantity, or both. An example of an object-based strategy is the *ontology strategy*, which suggests going up the class hierarchy to generate an estimate. For example, while estimating the height of Jason Kidd, one could use the information that he is a point guard[2], or even that he is a basketball player. An

[1]http://trec.nist.gov/

[2]The point guard is one of the standard positions in a regulation basketball game. Typically one of the smallest players on the team, the point guard’s job is to pass the ball to other players who are responsible for making most of the points.

example of a quantity-based strategy is the *density strategy*, which suggests estimating a quantity by using a density (e.g., average, rate, per capita income) and multiplying by its extent. When asked a question, BotE-Solver first tries to see if the answer is available in the knowledge base. Failing that, it tries to find similar examples for which answers are available. This is the *analogy strategy* (Paritosh and Klenk, 2006), an important object-based strategy. If no analogues are found, then other applicable strategies are applied. For a complete list of these strategies and more details, the reader is referred to Paritosh and Forbus (2005). BotE reasoning is a heuristic domain, and the ontology strategy, density strategy, analogy strategy are examples of heuristic methods.

## 4.2 Other Heuristic Domains

In this section we present a list of heuristic domains, and some hypotheses about heuristic methods that might work in those domains. The numbering begins with H2, as the first heuristic domain was covered in the last section.

### H2. Temporal Estimation: When did X happen?

Even when we do not know the exact date when something happened, research in autobiographical memory (Thompson et al., 1996) suggests that by recalling landmark events and constructing a local temporal scale, people can generate reasonable estimates. Allen's temporal interval calculus (1983) presents a neat set of relationships that could be used to organize the heuristic methods in this domain. For example, consider various ways to answer “When was Mark Twain born?” If you happened to know that Mark Twain wrote an account of his participation[1] in the American Civil War, which went on from 1861 to 1865, then you might guess that he was probably born around 1830.

### H3. Comparison*:* Is X larger than Y along dimension D? Who is the maximum/minimum of a class/set along dimension D?

These questions involve making comparisons between two or more objects along some scalar dimension. At first glance, this might look like solving a few back of the envelope problems and comparing the results. However, it is often easier to answer the comparative question. For example, it is easier to say that Microsoft research spending is more than Apple's than it is to estimate their respective spendings and compare them. One heuristic method here is *projection*. If we are comparing X and Y along dimension D, and we know another dimension E that is qualitatively proportional to D, then we can project the ordinal result along E on to D. Qualitative representations and techniques of comparative analysis (Weld 1987) might play an important role in this heuristic domain. An important psychological heuristic method is the *availability* heuristic (Tversky and Kahneman, 1973). According to the availability heuristic, the ease with which instances come to mind is used as indicator of the size or frequency of the class. For example, when asked the question, “Do homicides or suicides cause more deaths in the US?” most people erroneously answer homicides, as it is easier to recall examples of homicides than suicides. Tversky and Kahneman's goal was to highlight the heuristic by pointing out when it leads to systematic errors. However, the availability heuristic is a useful one, and how often it is right is an empirical question. An interesting implication is the idea of “ease of recall”– for most knowledge based systems, fact lookup will take roughly the same amount of time, irrespective of the fact in question. Can it be useful (performance and/or efficiency-wise) to have a model of “ease of recall”? What would it look like?

[1] “The Private History of a Campaign That Failed” also made into a movie.

### H4. Probability: How likely is X? Is X more likely than Y?

There are some fundamental differences about the interpretation of the question: the frequentist approach says that in order to compute the probability of X, we need a well-defined random experiment where frequencies of various events including X can be counted. This approach would not be willing to define probability for a new event like the death of a specific person. However, people make judgments and decisions based on the likelihood of various events, for example, the author of a scientific paper might consider: “What is the likelihood that my paper will get accepted by a certain conference or journal?” One can generate a reasonable guess about which of two journals are more likely to accept the paper without knowing detailed joint probability distributions. It might be possible to answer the question without knowing a priori all the relevant variables affecting acceptance. One psychological heuristic method to answer these questions is the *representativeness* heuristic (Tversky and Kahneman, 1972) that guides people's estimates of such likelihood. The representativeness heuristic says that people judge the probability that P is a member of category C on the basis of the similarity of P to our concept of a prototypical member of C. Models of analogy and generalization (Falkenhainer, Forbus and Gentner, 1989; Kuehne et al., 2000) could be used to model the representativeness heuristic. Recent work by Halstead (2005) has incorporated probability into the structured models of generalization.

### H5. Classification: Does X belong to the class Y? Does X satisfy property P?

Allan Collins' seminal work on plausible reasoning (1989) gives us a set of strategies used by people in answering such questions, based on an analysis of verbal protocols used by people in answering such questions. Consider questions like: Is Somalia a developing nation? Do they grow coffee in Russia? One could use Somalia's similarity to other instances of developing nation as evidence for answering the question in the affirmative. By noticing the dissimilarities between Russia and other coffee growing countries like Ethiopia, Brazil, Kenya, India, etc., one might conclude that Russia doesn't grow coffee. The representativeness heuristic is useful in answering classification questions as well. Gigerenzer (1999) has proposed the *take-the-first* and *take-the-best* heuristics, which suggest that even though we need to know information along various dimensions to predict if a country is a developing nation, usually we can make a decision based on just one dimension. This is owing to the *non-compensatory* nature of cues in the world, which says that the classification made using the most important dimension is likely to be right, as that dimension usually dominates all the other dimensions.

### H6. Choice, evaluation, decision making: Is X good? Is X better than Y? What is the best course of action?

At first blush, this might look like H3, the comparison domain above. However, a key idea in choice and decision making is that of evaluating a situation for how good it is. In Economics, this idea of evaluation is captured by *utility*. Prospect theory (Kahneman and Tversky, 1979) is the psychological version of the utility theory. Based on studying firefighters, pilots, nurses in Neonatal Intensive Care Units, and such people who constantly are making decisions with important consequences, Gary Klein (1999) has developed the Recognition-primed decision model, which is essentially an analogical approach. Consider questions like: Is Toyota Corolla the right car for me? Should we hire X or Y? Similarity and experiential knowledge are key elements of the heuristic methods in this domain.

### H7. Prediction: What will happen if X?

Qualitative representations and methods of qualitative reasoning are a crucial part of making predictions in the face of incomplete knowledge. Consider: What will happen if the price of gasoline increases? What will happen to the outside temperature if it is snowing? The former involves identifying the causally related quantities to the price of gasoline, and might be explained to a large extent by first-principles qualitative reasoning. However, a more reasonable account of how people might answer the latter question is with experience: we know that it gets relatively warmer after snowing, but might not have a full causal account of the phenomenon. This hybrid explanation of qualitative mental models: relying on mostly similarity-based reasoning and only a little on first-principles based reasoning (Forbus and Genter, 1997) is currently being explored by Yan and Forbus (2004).

### H8. Explanation: Why X?

This is another forte of qualitative reasoning. As qualitative representations make causal relationships and modeling assumptions explicit, they naturally provide the grist for generating explanations (Bouwer and Bredeweg, 1999). Consider a question like: Why are hybrid cars more fuel efficient?

### H9. Sanity checking: Does X make sense? Is X reasonable?

This is a meta-heuristic domain of sorts, where rather than answering a question, we are given a question and a candidate answer, and we use all the above methods to figure out if the answer sounds reasonable. It might be possible to do sanity checking for reasoning domains for which we don't even have heuristic methods. For example, the question in the introduction that asked for the scientific name of Viagra: we can easily reject “Cialis,” “sex,” or “42” as being obviously incorrect. The first step in sanity checking is *typechecking* – making sure that the candidate answer is of expected class. Maintaining some global sense of various scales is another important aspect of sanity checking. For example, it is easy to reject 14ft as the diameter of Earth. All of the heuristic methods above can be then used to generate a plausible answer and compare it with the candidate answer to conclude if something makes sense or not.

## 3.3 Psychological Heuristic Methods

We have presented nine heuristic domains, out of which one, back of the envelope reasoning, has been tackled. This is likely an incomplete list, and there are probably more heuristic domains still to be found. Some of the heuristic methods come from an analysis of the structure of the domain. For example, methods of qualitative reasoning, although inspired in part by the human ability to reason without differential equations are not psychologically faithful. Other heuristics like availability and representativeness are psychological heuristics, which we believe might be generally useful in heuristic reasoning. An incidental benefit of exploring such psychological heuristics computationally is that it could lead to a better understanding of the architectural and representational assumptions underlying those heuristics, something which hasn't been much explored by psychologists. D. Kahneman has hypothesized a dual system architecture of the human mind: System 1 (Intuition) is fast, automatic, effortless, associative, slow-learning and emotional; System 2 (Reasoning) is slow, controlled, effortful, rule governed, flexible, and emotionally neutral. Although such models are popular in psychology, few AI systems are built in these ways. We believe that the heuristic reasoning approach will let us explore both goals at the same time: to build flexible systems, and to understand how the mind works.

# 5 Discussion

What kind of guarantees can be provided for heuristic reasoning? In this section we talk about the soundness, completeness and complexity of heuristic reasoning.

Instead of soundness, heuristic reasoning is concerned with *reasonableness*, best described by whether humans will find such answers acceptable. In some cases like numeric estimates, being in the correct order of magnitude might be used as a crude metric for how correct the answer is. Reasonableness in such reasoning will come from: 1) Finding multiple answers using different methods and seeing how close they are, and 2) Experiential statistics accumulated over many different problems solving episodes for keeping track of which heuristic methods give more reasonable answers. Instead of completeness in the logical sense, heuristic reasoning is concerned with *task-completeness* – can we answer all or most of the questions in a reasonable way? Regarding complexity: our hypothesis is that when heuristic methods succeed, then the solution is obtained by a shallow search. For example, in the BotE reasoning domain, the depth of the solution AND/OR trees is never more than ten.

Another issue concerns representation of heuristic methods. In our BotE work, we used *suggestions* and *procedural attachments* to represent the heuristic methods. As the library of heuristic methods and domains grows, more abstract and declarative representations will be crucial. Most of the heuristic methods we have seen until now can be described at the level of operations they provide for manipulating knowledge required to answer the question. Suppose that the knowledge required to answer the question accurately is K. The two primitive operations provided by heuristic methods are:

1. *Subset* methods: specify how to use K' which is a subset of K to generate a reasonable answer.
2. *Proxy* methods: specify how to use K" which is a proxy for K. K" contains information that is correlated in such a way with K, that it leads to an answer that is similar to what would be obtained if K was available

This suggests that we need to represent heuristic methods at the level of knowledge operations they perform. We believe that these heuristic methods and domains are compositional.

## 6 Conclusions

While an ambitious proposal, the decomposition into heuristic domains suggests a tractable approach towards building a comprehensive theory and implementation of heuristic reasoning. There are many interesting questions about the nature of heuristic domains that this research program hopes to answer: Which domains and tasks are inherently brittle, and which domains are heuristic? Are there different types of heuristic domains? We believe that this approach to heuristic reasoning will lead to software that's less brittle, and help us understand the aspects of intuitive reasoning in human minds.


## Acknowledgments

This research is supported by the Artificial Intelligence Program of the Computer Science Division of the Office of Naval Research. I would like to thank Ken Forbus for insightful comments on this paper and supporting these ideas. I would also like to thank Julie Saltzman, Dan Halstead and Matt Klenk for reading and discussing various drafts of this paper.


## References


Allen, James F. 1983. Maintaining knowledge about temporal intervals. *Communications of the ACM* 26(11) pp.832-843.

Bouwer, A. and Bredeweg, B. 1999. Explanation and Qualitative Reasoning, In *Proceedings of International Workshop on Qualitative Reasoning*.

Collins, A. and Michalski, R. 1989. The Logic of Plausible Reasoning: A Core Theory. *Cognitive Science*, 13, 1-49.

Falkenhainer, B., Forbus, K. D., & Gentner, D. (1989). The structure-mapping engine: Algorithm and examples. *Artificial Intelligence*, 41, 1-63.

Forbus, K. and Gentner, D. (1997).Qualitative mental models: Simulations or memories? *Proceedings of the Eleventh International Workshop on Qualitative Reasoning*, Cortona, Italy, June 3-6, pp. 97-104.

Gigerenzer, G., Todd, P. M. & the ABC Research Group. (1999). Simple heuristics that make us smart. New York: Oxford University Press.

Halstead, D. and Forbus, K. 2005. Transforming between Propositions and Features: Bridging the Gap. *Proceedings of AAAI-2005*. Pittsburgh, PA.

Hammond, K.R., McClelland, G.H. & Mumpower, J. (1980). Human Judgement and Decision Making. New York: Praeger.

Kahneman, D., and Tversky, A. Subjective probability: A judgment of representativeness. *Cognitive Psychology*, 1972, 3, 430-454.

Kahneman, D., and Tversky, A. 1979. Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47, 263-292.

Klein, G. 1999. Sources of Power: How People Make Decisions. MIT Press.

Kuehne, S., Forbus, K., Gentner, D. and Quinn, B.(2000) SEQL: Category learning as progressive abstraction using structure mapping. *Proceedings of CogSci 2000*.

Lenat, D.B. 1982. The nature of heuristics, *Artificial Intelligence*, Volume 19, Issue 2, Pages 189-249

Lenat D.B., Prakash, M. and Sheperd M., 1986. CYC: Using Common Sense Knowledge to Overcome Brittleness and Knowledge Acquisition Bottlenecks. *AI Magazine*.

Newell A. and Simon H. A.. 1963. GPS, a program that simulates human thought. In Edward A. Feigenbaum and Julian Feldman, editors, Computers and Thought, pages 279--296. McGraw-Hill, New York.

Paritosh, P.K and Forbus, K.D. (2001). Common Sense on the Envelope, In *Proceedings of the 15th International Workshop in Qualitative Reasoning*, San Antonio, TX.

Paritosh, P.K. and Forbus, K.D. (2004). Using Strategies and AND/OR Decomposition for Back of the Envelope Reasoning. In *Proceedings of the 18th International Workshop on Qualitative Reasoning*, Evanston.

Paritosh, P.K. and Forbus, K.D., (2005). Analysis of Strategic Knowledge in Back of the Envelope Reasoning, *In Proceedings of the 20th National Conference on Artificial Intelligence (AAAI-05)*, Pittsburgh, PA.

Polya, G. 1945. How to Solve It, Princeton University Press, Princeton, NJ.

Simon, H. A. 1957. A Behavioral Model of Rational Choice, in Models of Man.

Singh, Push, Lin, Thomas, Mueller, Erik T., Lim, Grace, Perkins, Travell, & Zhu, Wan Li (2002). Open Mind Common Sense: Knowledge acquisition from the general public. In Robert Meersman & Zahir Tari (Eds.), Lecture Notes in Computer Science: Vol. 2519. On the Move to Meaningful Internet Systems 2002, (pp. 1223-1237). Heidelberg: Springer-Verlag.

Swartz, C. E., 2003. Back-of-the-Envelope Physics. Johns Hopkins University Press, Maryland.

Thompson, C.P., Skowronski, J.J., Larsen, S.F., and Betz, A.L., 1996. Autobiographical Memory: Remembering What and Remembering When, Lawrence Erlbaum, NJ.

Tversky, A., & Kahneman, D. Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 1973, 5, 207-232.

Tversky, A., and Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases, *Science*, 185, pp 1124-1131.

Weld, D.S. 1987. Comparative Analysis. In *Proceedings of IJCAI*, 959-965.

Yan, J. and Forbus, K. 2004. Similarity-based qualitative simulation: A preliminary report. In *Proceedings of the 18th International Qualitative Reasoning Workshop*, Evanston.