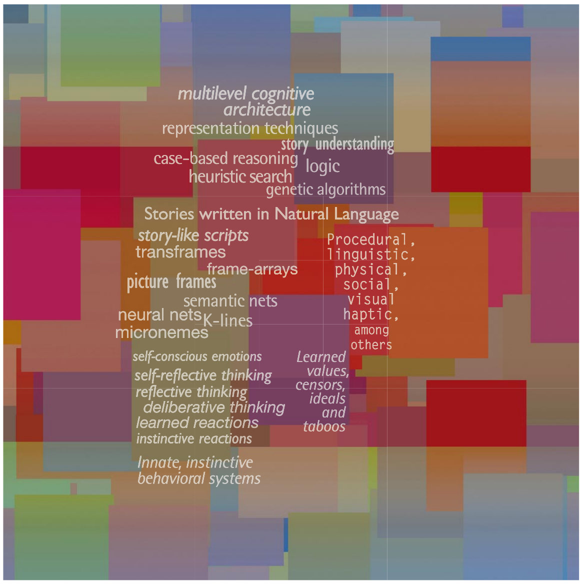


An architecture of diversity for commonsense reasoning

Although computers excel at certain bounded tasks that are difficult for humans, such as solving integrals, they have difficulty performing commonsense tasks that are easy for humans, such as understanding stories. In this Technical Forum contribution, we discuss commonsense reasoning and what makes it difficult for computers. We contend that commonsense reasoning is too hard a problem to solve using any single artificial intelligence technique. We propose a multilevel architecture consisting of diverse reasoning and representation techniques that collaborate and reflect in order to allow the best techniques to be used for the many situations that arise in commonsense reasoning. We present story understanding—specifically, understanding and answering questions about progressively harder children's texts—as a task for evaluating and scaling up a commonsense reasoning system.

In the fall of 2001, a proposal was developed by Marvin Minsky, Erik Mueller, Doug Riecken, Push Singh, Aaron Sloman, and Oliver Steele for a project to develop a human-level commonsense reasoning system. The basic proposal was (1) to develop certain ideas of Minsky and Sloman about a multilevel cognitive architecture, and (2) to develop the system in a way that would exploit many existing artificial intelligence techniques for commonsense reasoning and knowledge representation, such as case-based reasoning, logic, neural nets, genetic algorithms, and heuristic search.

We proposed to organize a meeting at which we would bring together many of the major established researchers in the area of commonsense knowledge and reasoning. Riecken organized a preliminary meeting at the IBM Thomas J. Watson Research Center in March 2002, at which many IBM researchers were invited to discuss and react to this general subject as well as to present their own ideas. Afterwards, the specific proposal was discussed in more detail by specialists in commonsense knowledge and reasoning at a meeting held on St. Thomas, Virgin Islands, in April 2002, and hosted by Jeffrey Epstein. This Technical Forum contribution focuses on the preliminary meeting, but also contains some material presented at the April meeting, including some material from Minsky's forthcoming book *The Emotion Machine*.[1]

At the IBM meeting, a broad consensus was reached on three main points. First, there was agreement that the community should strive toward solving a nontrivial problem that would require a level of knowledge, and a capability of reasoning with that knowledge, beyond what is demonstrated by current systems. The problem put forward was that of story understanding. An important advantage of the story understanding task is that standardized tests are available to evaluate students on their reading comprehension skills. Moreover, these tests require the use of commonsense reasoning skills. It is thus possible to evaluate the performance of any story understanding system against that of students at different reading levels.[2]

Second, there was consensus that the story understanding task provides a strong testbed for evaluating a commonsense reasoning system. Not only does such a system need several different forms of reasoning, representation, and learning, but it also needs them to work in conjunction with each other. In addition, the task highlights the importance of using and reasoning with common sense. This is illustrated by a sentence from a story about a child and her grandfather: "He gently takes my elbow as we walk so that I can help show him the path." Knowledge of the fact that the grandfather is blind, and the commonsense facts that people ordinarily use their sight to find paths and that blind people are unable to see, enable the inference that the child is guiding the grandfather and not merely pointing out the path, another frequent sense of the word "show." Absence of this commonsense knowledge could lead to the incorrect interpretation of the word "show."

Third, there was agreement on the need to develop a testbed architecture for representation and reasoning that allows different systems and representations to work with each other. Researchers often try to solve a problem using just one form of representation and reasoning. But such an approach does not work well for sufficiently complex problems such as story understanding. In contrast, enabling various techniques to collaborate will allow the best techniques to be used for a given situation. Any such architecture must provide metalevel control and knowledge that will enable different techniques to determine whether or not they are suited for a given task, to decide what other techniques may be better for the task, and to communicate information and share partial results with each other.

What makes commonsense reasoning difficult

Commonsense reasoning—the sort of reasoning we would expect a child to do easily—is difficult for computers to do. Certainly, the relative paucity of results in this field does not reflect the considerable effort that has been expended, starting with McCarthy's paper "Programs with Common Sense."[3] Nevertheless, the problem remains unsolved. What is it about commonsense reasoning that makes it difficult to automate? Various explanations have been suggested, some of which we discuss in this section.

McCarthy's commonsense informatic situation. The knowledge needed to solve a commonsense reasoning problem is typically much more extensive and general than the knowledge needed to solve difficult problems. McCarthy points out that the knowledge needed to solve well-formulated problems in fields such as physics or mathematics is bounded.[4] In contrast, there are no *a priori* limitations to the facts that are needed to solve commonsense problems: the given knowledge may be incomplete; one may have to use approximate concepts and approximate theories; one will generally have to use nonmonotonic reasoning to reach conclusions; and one will need some ability to reflect upon one's own reasoning processes. Morgenstern provides an example of the commonsense informatic situation in the problem of two friends arranging to meet for dinner at a restaurant.[5]

Explicit vs implicit knowledge. Commonsense knowledge is often implicit, whereas the knowledge needed to solve well-formulated difficult problems is often explicit. For example, the knowledge needed to solve integrals can be found in explicit form in a standard calculus textbook. However, the knowledge needed to arrange a dinner meeting exists in vague, implicit form. Implicit knowledge must first be made explicit, which is a time-consuming task requiring a serious knowledge engineering effort.

Domain knowledge. A huge amount of knowledge is needed to do even simple forms of commonsense reasoning. For example, to figure out what sorts of objects will work as stakes in a garden—a reasoning task that seemingly demands no effort—requires knowledge of plant materials, how plants grow, flexibility and hardness, shapes of plants, soil texture, properties of wind, spatial reasoning, and temporal reasoning.[6] Although there have been a number of efforts to capture large amounts of world knowledge, most notably the Cyc** project,[7] we are not at this point aware of any knowledge base that contains the information necessary to reason about stakes in a garden or about fumbling for an object in one's pocket.

This Technical Forum piece does not present a solution to these difficulties. Rather, we are attempting to see how far we can progress on an important commonsense reasoning problem even in the presence of such difficulties.

Story understanding as a vehicle for studying commonsense reasoning

Story understanding requires addressing the commonsense informatic situation. A story understanding system should be able to read and understand a story, and demonstrate its understanding by (1) answering questions about the story, (2) producing paraphrases and summaries of the story, and (3) integrating the information the story contains into a database. Further, useful results from this work will have a direct impact on many business products and services.

A brief history of story understanding systems. Starting in the 1960s,[8] researchers have studied story understanding and have built systems that can read and answer questions about simple stories. An early system built by Charniak[9] used a single mechanism, test-action demons, for making inferences in understanding. In the 1970s, Schank and Abelson[10] proposed scripts, plans, and goals as knowledge structures for understanding. These knowledge structures were incorporated into the SAM[11] and PAM[12] story understanding systems.

In the 1980s, knowledge structures for emotions, story themes, and spatial/temporal maps were incorporated into BORIS.[13] AQUA[14] used case-based reasoning to retrieve and apply explanation patterns in order to answer questions raised by anomalies encountered while reading a story. CRAM[15] used a connectionist approach to story understanding.

Recent story understanding systems have adopted the approach of understanding a story by building and maintaining a simulation that models the mental and physical states and events described in the story, as demonstrated in ThoughtTreasure.[16] The advantage of this approach is that it is easy to answer questions about the story simply by examining the contents of the simulation.

Critical problems for story understanding systems. The story understanding systems built so far work only on the particular stories they are designed to handle. For example, SAM[11] handles five stories, BORIS[13] three, AQUA[14] five, and ThoughtTreasure[16] three. What prevents story understanding systems from scaling up to hundreds of previously unseen stories?

We contend that story understanding research is blocked on three critical problems: (1) complexity of the structure of natural language, (2) necessity for large commonsense knowledge bases, and (3) combinatorial explosion in the understanding process.

Complexity of the structure of natural language. Rare is the simple subject-verb-object sentence that maps into a simple proposition. More typically, text contains numerous language phenomena such as adverbials, compound nouns, direct and indirect speech, ellipsis, genitive constructions, and relative clauses.[17] Present-day syntactic and semantic parsers have trouble producing accurate parses of typical story sentences.

Necessity for large commonsense knowledge bases. Understanding even simple stories requires knowing a huge number of facts. For example, understanding the first paragraph of *The Cat in the Hat* requires knowing about children's play, how children can be affected by winter weather, their relationship to their parents, and notions of discipline, boredom, surprise, and risk. Similarly, as Davis[18] points out, the first paragraph of *The Tale of Benjamin Bunny* assumes familiarity with concepts of quantity, space, time, physics, goals, plans, needs, and communication.

Combinatorial explosion in the understanding process. Multiple possible interpretations arise at all levels of language. Words are ambiguous as to part of speech and word sense. Sentences are syntactically ambiguous. There are several possible explanations for any action of a story character, several possible explanations for those explanations, and so on. We get a combinatorial explosion: the understanding process must search an extremely large space of possibilities.

Approaches to critical problems in story understanding. What can be done? We propose a three-pronged approach. First, to deal with the complexity of the structure of natural language, we make a major cut in complexity by going back to books for early readers. Second, to deal with the necessity for large commonsense knowledge bases, we propose to identify the domains most frequently used in a restricted set of stories and to address these first. Last, to deal with the combinatorial explosion in the understanding process, we propose a new paradigm for commonsense reasoning: an architecture of diversity.

Early readers. Early reader texts are designed for preschool and kindergarten students. These texts employ a small or controlled vocabulary, short sentences, and limited language constructions. Working with early reader texts will enable us to effectively solve the language front-end problem using existing research techniques.

Text annotation for domain identification. We cannot hope to deal with the commonsense informatic situation head-on. The point of McCarthy's 1996 paper[4] is that any domain can be relevant to a particular problem: when reading a story, any area of knowledge may be necessary for comprehension. This is less true for stories designed for very young readers; although, as our examples above show, a great many concepts and domains are still needed for full comprehension even of early reader texts. Nevertheless, we believe we can make progress by choosing to address those domains that most frequently turn up in children's stories. Such an approach would, we hope, make the problem tractable.

We thus propose the following corpus-based approach. We start with a corpus of stories at the preschool and kindergarten levels and divide the corpus into a development set and a test set. We manually annotate each story in the development set with an informal inventory of what domains of commonsense knowledge and reasoning must be addressed in order to understand the story. We sort the domains by their frequency and attempt to develop methods to understand the domains that occur most frequently. We start with the most frequent domain, proceeding to the next most frequent domain, and so forth. Development proceeds on the development set, and a final evaluation of the generality of the system is conducted on the previously unseen test set. We iterate this process on successively higher reading levels, progressing to stories designed for Grades 1, 2, and 3. This approach, based on an incremental series of experiments, will enable a significant research focus at each step on an architecture of diversity.

To demonstrate how this approach would work, we formed a corpus of 15 early reader stories and annotated them as to the domains of common sense necessary for understanding them. The vocabulary size was 561 words. The top 10 domains of common sense are shown in Table 1. This provides us with a path for research in understanding the story corpus: focus on handling the most frequently appearing domains of common sense.

Dealing with these concepts is by no means trivial. We plan to leverage the extensive work that has been done in these areas. Such work includes: ThoughtTreasure,[16] NETL2,[19] Cyc,[7] Shanahan's formalization of time,[20] the RCC formalization of space,[21] and Kuipers's Spatial Semantic Hierarchy.[22] We will also employ rapid knowledge formation techniques such as Open Mind.[23]

Table 1 Early reader corpus: top 10 domains of common sense

Domain	Number of Stories	Percentage of Stories
space—location	14	93.3
space—motion	11	73.3
math—counting	10	66.6
attitude—positive	9	60.0
speech act	9	60.0
space—size	8	53.3
space—grasping	7	46.6
sound—speech	7	46.6
logic—universal quantification	7	46.6
space—housing	6	40.0

An architecture of diversity

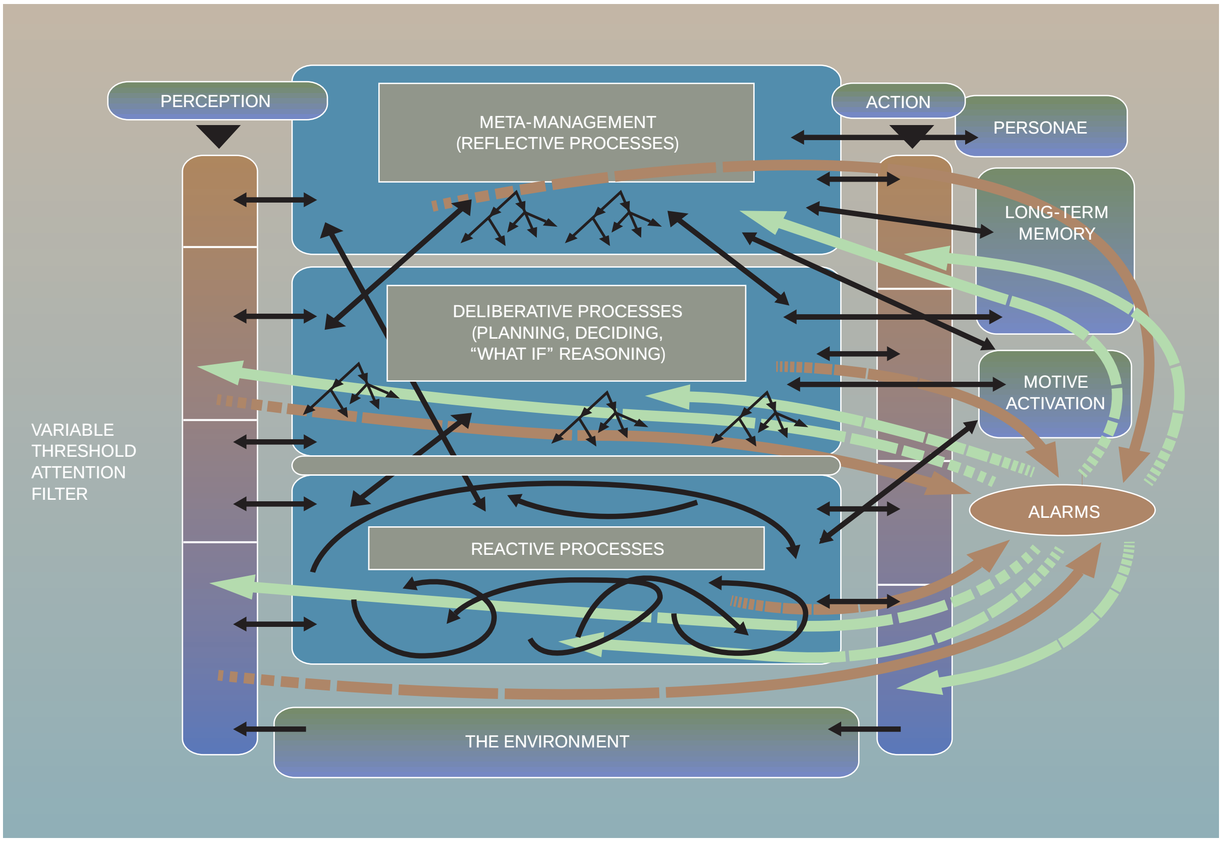
Many attempts to build intelligent computers have hunted for a single mechanism (such as universal subgoaling, propagation rules, logical inference, probabilistic reasoning) or representation (such as production rules, connectionist networks, logical formulas, causal networks) that would serve as a basis for general intelligence. Why have these approaches so far failed to achieve human-level common sense?

We believe that the problem is too large to solve using any single approach. Human versatility must emerge from a large-scale architecture of diversity in which each of several different reasoning mechanisms and representations can help overcome the deficiencies of the other ones.[24,1] Our hypothesis is that such an architecture can overcome the combinatorial explosion problem in story understanding.

Multilevel cognitive architecture. We conjecture that the information processing architecture of a human is something like the three-level architecture developed by Sloman in the Cognition and Affect project[25] (H-Cogaff), shown in Figure 1. This conjecture is based on evidence of many kinds from several disciplines, and constraints on evolvability, implementability in neural mechanisms, and functionality.[26]

Reactive processes are those in which internal or external states detected by sensors immediately trigger internal or external responses. Deliberative processes are those in which alternative possibilities for action can be considered, categorized, evaluated, and selected or rejected. More generally a deliberative mechanism may be capable of counterfactual reasoning about the past and present and hypothetical reasoning about the future. The depth, precision, and validity of such reasoning can vary. Meta-management processes add the ability to monitor, evaluate, and to some extent control processes occurring within the system in much the same way as the whole system observes and acts on the environment. The three layers operate concurrently and do not form a simple dominance hierarchy. Arrows represent flow of information and control, and boundaries need not be sharp in all implementations.

Figure 1 The H-Cogaff three-level architecture



A. Sloman, "Beyond Shallow Models of Emotion," *Cognitive Processing,* Vol. 1, No. 1 (2001).

The reactive and deliberative layers differ in that the deliberative layer evolved much later and requires a far more sophisticated long-term memory, as well as symbolic reasoning capabilities using a short-term reusable memory. The meta-management layer may have evolved at a still later time and requires explicit use of concepts referring to states of an information processing architecture. The earliest organisms, such as most existing organisms, were totally reactive. Deliberative and meta-management layers evolved later. Adult humans appear to have all three types of processing, which is probably rare among other animals.

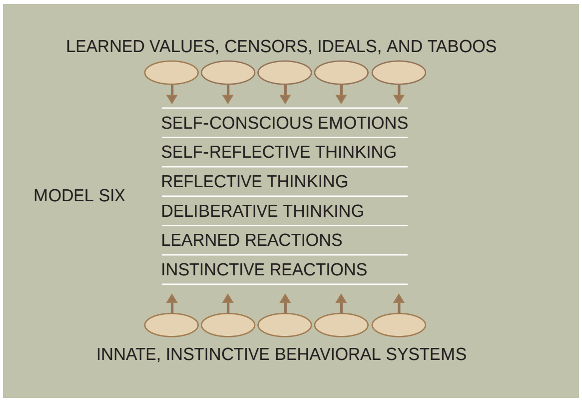
One of the key features that gives H-Cogaff its generality is the fact that different components, instead of forming parts of simple pipelines, can concurrently send information of various kinds to arbitrarily many other components, allowing a wide variety of feedback mechanisms and triggering mechanisms.

In story understanding, the meta-management level may control the deliberative level in a number of ways.

- If the deliberative level is spending too much time considering certain details and those details are not crucial to the story, the meta-management level will make the deliberative level stop.
- If the deliberative level is spending too much time on a task that does not relate to the goal of reading the story, the meta-management level will make the deliberative level stop.
- If the deliberative level becomes confused, the meta-management level will tell it to go back and reread. The deliberative level may have ruled out a possibility earlier that needs to be reconsidered in light of new information.

Minsky further elaborates the H-Cogaff architecture into the six-level architecture called "Model Six" shown in Figure 2.[1] At its bottom lies a "zoo of instinctive subanimals" built upon ancient, ancestral systems that still maintain our bodies and brains. These include systems for feeding, breathing, heating, sleeping, and other systems that keep us alive. The deliberative and reflective levels are engaged to solve more difficult kinds of problems. The self-reflective level is engaged when the problems involve our relationships with our past and future selves. At the top lies machinery that we acquire from our societies, such as suppressors and censors, imprimers and values, and our various kinds of self-ideals.

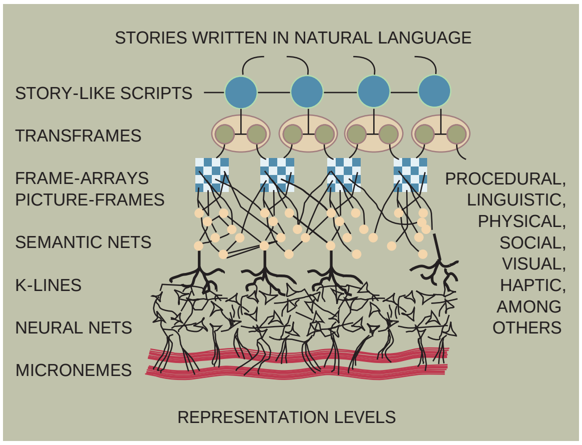
Figure 2 The Model Six six-level architecture



M. Minsky, *The Emotion Machine*, Pantheon, New York (forthcoming).

Multiple reasoning and representation schemes and levels. An architecture of diversity would embed representations from natural language to micronemes[27,1] as depicted in Figure 3. The representations depicted include frames, transframes, frame-arrays, K-lines, and micronemes. A frame is a representation based on a set of slots to which other structures can be attached.[28] Each slot is connected to a default assumption that is easily displaced by more specific information. A transframe is a particular type of frame representing the causal trajectory between the initial and resulting states representing a situation that a legal action was performed on. A frame-array is a collection of frames that share the same slots, making it easy to change perspective with respect to physical viewpoint or other mental realms. A knowledge-line or K-line is a wirelike structure that attaches itself to whichever resources are active in solving a problem. The K-line simplifies activation of those same resources when solving a similar problem in the future. Micronemes are low-level features for representing the many cognitive shades and hues of a context. In Figure 3, new evolved structures are made from older lower-level ones, and the tower shown might be a plausible Darwinian brain-development scheme.

Figure 3 Multiple reasoning and representation schemes and levels



From M. Minsky, "Common Sense-Based Interfaces," *Communications of the ACM,* Vol. 43, No. 8, 67-73 (2001). Copyright 2001 ACM. Reprinted by permission.

Table 2 shows just a few of the diverse representation and reasoning schemes useful for domains of story understanding.

We propose to address the commonsense reasoning problem starting with stories for very young readers. However, to demonstrate all of the different ways we think when understanding a story, and what we would eventually expect a commonsense story understanding system to be able to handle, consider the following adult story (the discussion here is condensed from Reference 1).

Joan heard a ring and picked up the phone. Charles was answering her question about how to use a certain technique. He suggested she read a certain book, which he would soon bring to her since he had planned to be in her neighborhood. Joan thanked him and ended the call. Soon Charles arrived and gave her the book.

Following are a few of the understandings an adult reader would have after hearing the story.

Table 2 Diverse schemes for story understanding domains

Domain	Representation/Reasoning Schemes
space	frame, generalized cylinder model, interval logic, occupancy grid
time, action effects	causal model, event calculus, situation calculus, transframe
reactivity	neural net, production system, subsumption architecture
schemas, scripts	finite automaton, frame, frame-array, generalized Petri net
subgoaling	first-order logic, K-line, marker passing, semantic net
emotions, attitudes	microneme, neural net, temporal modal logic

- *Joan heard a ring*. She recognizes it as a telephone bell and feels the need to respond quickly. She knows how to use the telephone.
- *She picked up the phone*. She is subsequently holding the phone to her ear.
- *Charles was answering her question*. Charles and Joan are not in the same room. Charles also knows how to use the telephone.
- *He suggested she read a certain book*. Joan probably now feels some relief, since she knows where to find the knowledge she needs.
- *He had planned to be in her neighborhood*. Joan will not be surprised when he arrives, because she will remember that he said he would come.
- *He gave her the book*. Will she have to give it back? The story does not tell us that.

These conclusions are based on reasoning and representations in many realms, as follow.

The physical realm. In this realm, *give* might mean the motion of the book through space. This could be represented as a transframe that starts with Charles's hand holding the book and ends with Joan's hand carrying it. One must know a lot about physical things and how they behave in space and time.

The social realm. In this realm, *give* may signify social acts that can alter the relationships of the actors. What were Charles's motives or his attitudes? Clearly, he was not returning a loan. Was he hoping to ingratiate himself? Or was he just being generous? How will Joan feel about Charles after he gives her the book? One must know a lot about what people are, and a certain amount about how people work.

The dominion realm. Given *Charles gave Joan the book*, one infers not only that Joan is holding the book, but also that, at least for a time, she possesses the right to use it.

The conversational realm. How do conversations work? Consider how many elaborate skills are involved in a typical verbal exchange. One has to keep track of what is being discussed, what one has previously told the listener, and what the listener knows. Thus conversations are partly based on knowledge of how human memories work and what is commonly known in one's culture. One has to make sure the listener has understood what was said and why it was said. One certainly needs to know how to speak and to understand some of what one may hear.

The procedural realm. How does one make a telephone call? One must first find a phone and dial a number. Then once the connection has been established, one says hello, talks a bit, and eventually leads into why one called. At the end, one says goodbye and hangs up the phone. Generally, such scripts have certain steps that are specified, while other steps provide for more room to improvise.

The sensory and motor realms. Each of the above steps raises questions. For example, it takes only one second or so for one's arm to reach out in order to pick up the phone. How can one do that so quickly?

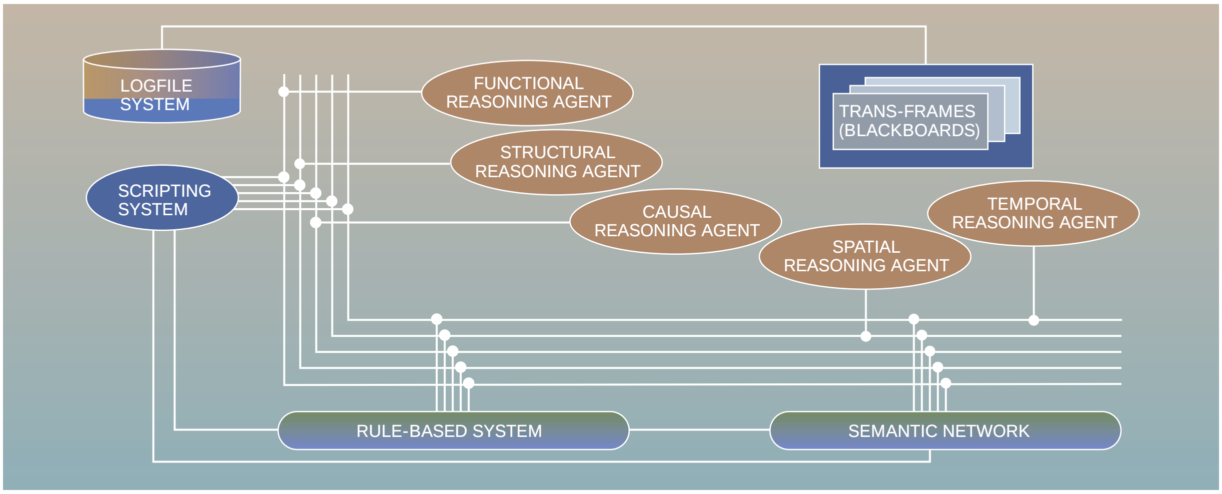
The kinesthetic, tactile, and haptic realms. Using a telephone or any other physical object engages a great base of body-related knowledge and skills. One anticipates how the phone will feel against one's ear or sandwiched between shoulder and cheek. One expects certain haptic sensations such as the feel of the phone's weight. One strengthens one's grip when the phone starts to slip.

The temporal realms. People have elaborate models of time where events are located in futures and pasts that are represented in relation to other times and events or in anecdotal stories.

The economic realm. People know and reason about the costs incurred by each action or transaction in terms of money, energy, space, or time.

The reflective realm. People know about themselves. One knows to some degree what one can or cannot do, what kinds of problems one can solve, how one's thinking and memory works, and what sorts of things one is able to learn.

Figure 4 The M system



From D. Riecken, "An Architecture of Integrated Agents," *Communications of the ACM,* Vol. 37, No. 7, 107-116 (1994). Copyright 1994 ACM. Reprinted by permission.

Along with these positive kinds of knowledge, one also has negative knowledge about what might go wrong when using a phone. One must know what to do if one gets a wrong number, if there is no answer, or if a modem or intercept recording is reached.

Example system with architecture of diversity. Thus far, the Sloman and Minsky architectures are theoretical constructs and have not yet been implemented. However, there are examples of working systems that capture the spirit of such architectures. One such example is the M system depicted in Figure 4.[29] M integrates multiple reasoning processes and representations to serve as an assistant to a user collaborating with other workers within a virtual meeting room that hosts multimedia desktop conferencing. M serves to recognize and classify the actions performed by the participants as well as the objects upon which the actions are applied; example actions and objects are brainstorming on a whiteboard, coauthoring a document, and creating and working with other artifacts.

Next steps

The two recent meetings held in March 2002 at the IBM Thomas J. Watson Research Center and in April 2002 on St. Thomas indicate that there is a dedicated group of recognized researchers interested in working together on a project to develop a solution to commonsense reasoning. We are now planning to undertake some of the next steps in a plan for such a project. The inspiration for this work comes from Minsky's past and forthcoming work. We close with his thoughts on how such a project might be realized, as follows.

Our goal is to aim toward a critical "change of phase" that will come when we cross a threshold at which our systems know how to improve themselves. This is something that all young children can do, but we do not know enough about how they do it; so one goal of the project must be to develop better models of how normal people think.

We will start by trying to implement some of the architectures proposed over the past decade. There already exist many useful schemes for representing and using knowledge mostly of a factual nature for use on what we call the deliberative level. However, there has not been enough work on the higher reflective and self-reflective levels that humans use, as they learn to improve their thinking itself. Any such system, we claim, will need additional kinds of metaresources, which will include systems that manage, criticize, and modify the already operating parts of the structure.

In the field of AI we already have many resources related to this, for example, neural networks, formal logic, relational databases, genetic programs, statistical methods, and of course the heuristic search, planning, and case-based reasoning schemes of earlier years. However, our goal is not to discuss which method is best. Instead we will try to develop a plan of how to incorporate into one system the virtues of many different approaches. Of course, each such scheme has deficiencies and our hope is that our system can escape from these by using higher-level, more reflective schemes that understand what each of those other schemes can do and in what context they are most effective.

**Trademark or registered trademark of Cycorp, Inc.

Cited references and notes

1. M. Minsky, *The Emotion Machine*, Pantheon, New York (forthcoming). Several chapters are on line at http://web.media.mit.edu/~minsky.
2. The use of reading comprehension tests as a metric for evaluating story understanding systems was previously proposed in L. Hirschman, M. Light, E. Breck, and J. Burger, "Deep Read: A Reading Comprehension System," *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics*, College Park, MD, June 1999, Association for Computational Linguistics (1999).
3. J. McCarthy, "Programs with Common Sense," *Proceedings of the Symposium on Mechanisation of Thought Processes*, Her Majesty's Stationery Office, London (1958), pp. 77–84.
4. J. McCarthy, "From Here to Human-Level Intelligence," *Proceedings of the Fifth International Conference on Principles of Knowledge Representation and Reasoning* (KR'96), Cambridge, MA, November 1996, Morgan Kaufmann, San Mateo, CA (1996), pp. 640–646.
5. L. Morgenstern, "A Formal Theory of Multiple Agent Nonmonotonic Reasoning," *Proceedings of the Eighth National Conference on Artificial Intelligence*, AAAI Press, Menlo Park, CA (1990), pp. 538–544.
6. E. Davis, "The Naive Physics Perplex," *AI Magazine* **19**, No. 4, 51–79 (1998).
7. D. Lenat, "Cyc: A Large-Scale Investment in Knowledge Infrastructure," *Communications of the ACM* **38**, No. 11, 32–38 (1995).
8. More details can be found in E. T. Mueller, "Story Understanding," to appear in *Encyclopedia of Cognitive Science*, Nature Publishing Group, London (2002).
9. E. Charniak, *Toward a Model of Children's Story Comprehension*, Technical Report AITR-266, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA (1972).
10. R. C. Schank and R. P. Abelson, *Scripts, Plans, Goals, and Understanding*, L. Erlbaum Associates, Hillsdale, NJ (1977).
11. R. E. Cullingford, *Script Application: Computer Understanding of Newspaper Stories*, Technical Report YALE/DCS/tr116, Computer Science Department, Yale University, New Haven, CT (1978).
12. R. Wilensky, *Understanding Goal-Based Stories*, Technical Report YALE/DCS/tr140, Computer Science Department, Yale University, New Haven, CT (1978).
13. M. G. Dyer, *In-Depth Understanding*, MIT Press, Cambridge, MA (1983).
14. A. Ram, *Question-Driven Understanding: An Integrated Theory of Story Understanding, Memory, and Learning*, Technical Report YALE/DCS/tr710, Computer Science Department, Yale University, New Haven, CT (1989).
15. C. Dolan, *Tensor Manipulation Networks: Connectionist and Symbolic Approaches to Comprehension, Learning, and Planning*, Technical Report 890030, Computer Science Department, University of California, Los Angeles, CA (1989).
16. E. T. Mueller, *Natural Language Processing with ThoughtTreasure*, Signiform, New York (1998), full text of book available on line at http://www.signiform.com/tt/book/.
17. L. G. Alexander, *Longman English Grammar*, Longman, London (1988).
18. E. Davis, *Representations of Commonsense Knowledge*, Morgan Kauffman, San Mateo, CA (1990).
19. S. E. Fahlman, *NETL: A System for Representing and Using Real-World Knowledge*, MIT Press, Cambridge, MA (1979).
20. M. Shanahan, *Solving the Frame Problem*, MIT Press, Cambridge, MA (1997).
21. D. A. Randell, Z. Cui, and A. G. Cohn, "A Spatial Logic Based on Regions and Connection," *Proceedings of the Third International Conference on Knowledge Representation and Reasoning*, Morgan Kaufmann, San Mateo, CA (1992), pp. 165–176.
22. B. Kuipers, "The Spatial Semantic Hierarchy," *Artificial Intelligence* **119**, 191–233 (2000).
23. P. Singh, "The Public Acquisition of Commonsense Knowledge," *Proceedings of the AAAI Spring Symposium on Acquiring (and Using) Linguistic (and World) Knowledge for Information Access*, Palo Alto, CA, March 2002, American Association for Artificial Intelligence (2002).
24. M. Minsky, *The Society of Mind*, Simon & Schuster, New York (1985).
25. A. Sloman, "Beyond Shallow Models of Emotion," *Cognitive Processing* **1**, No. 1 (2001).
26. A. Sloman, "Architectural Requirements for Human-Like Agents both Natural and Artificial," K. Dautenhahn, Editor, *Human Cognition and Social Agent Technology*, John Benjamins, Amsterdam (2000), pp. 163–195.
27. M. Minsky, "Common Sense-Based Interfaces," *Communications of the ACM* **43**, No. 8, 67–73 (2001).
28. M. Minsky, "A Framework for Representing Knowledge," AI Laboratory Memo 306, Artificial Intelligence Laboratory, Massachusetts Institute of Technology (1974), reprinted in *The Psychology of Computer Vision*, Patrick Winston, Editor, McGraw-Hill, New York (1975).
29. D. Riecken, "An Architecture of Integrated Agents," *Communications of the ACM* **37**, No. 7, 107–116 (1994).

Accepted for publication May 17, 2002.

J. McCarthy
Stanford University
Stanford, California

M. Minsky
Massachusetts Institute of Technology
Cambridge, Massachusetts

A. Sloman
University of Birmingham
Birmingham, UK

L. Gong
IBM Research Division
Hawthorne, New York

T. Lau
IBM Research Division
Hawthorne, New York

L. Morgenstern
IBM Research Division
Hawthorne, New York

E. T. Mueller
IBM Research Division
Hawthorne, New York

D. Riecken
IBM Research Division
Hawthorne, New York

M. Singh
IBM Research Division
Hawthorne, New York

P. Singh
Massachusetts Institute of Technology
Cambridge, Massachusetts