

Computing and Informatics, Vol. 22, 2003, 521–543

EXAMINING THE SOCIETY OF MIND

Push SINGH

*Media Lab
Massachusetts Institute of Technology
20 Ames Street
Cambridge, MA 02139
United States
e-mail: push@mit.edu*

Revised manuscript received 4 November 2003

Abstract. This article examines Marvin Minsky’s *Society of Mind* theory of human cognition. We describe some of the history behind the theory, review several of the specific mechanisms and representations that Minsky proposes, and consider related developments in artificial intelligence since the theory’s publication.

Keywords: The Society of Mind, multiagent systems, commonsense reasoning

1 INTRODUCTION

The functions performed by the brain are the products of the work of thousands of different, specialized sub-systems, the intricate product of hundreds of millions of years of biological evolution. We cannot hope to understand such an organization by emulating the techniques of those particle physicists who search for the simplest possible unifying conceptions. Constructing a mind is simply a different kind of problem – of how to synthesize organizational systems that can support a large enough diversity of different schemes, yet enable them to work together to exploit one another’s abilities. [1]

What is the human mind and how does it work? This is the question that Marvin Minsky asks in *The Society of Mind* [2]. He explores a staggering range of issues, from the composition of the simplest mental processes to proposals for the largest-scale architectural organization of the mind, ultimately touching on virtually every

important question one might ask about human cognition. How do we recognize objects and scenes? How do we use words and language? How do we achieve goals? How do we learn new concepts and skills? How do we understand things? What are feelings and emotions? How does ‘common sense’ work?

In seeking answers to these questions, Minsky does not search for a ‘basic principle’ from which all cognitive phenomena somehow emerge, for example, some universal method of inference, all-purpose representation, or unifying mathematical theory. Instead, to explain the many things minds do, Minsky presents the reader with a theory that dignifies the notion that the mind consists of a great diversity of mechanisms: every mind is really a ‘Society of Mind’, a tremendously rich and multifaceted society of structures and processes, in every individual the unique product of eons of genetic evolution, millennia of human cultural evolution, and years of personal experience.

Minsky introduces the term *agent* to refer to the simplest individuals that populate such societies of mind. Each agent is on the scale of a typical component of a computer program, like a simple subroutine or data structure, and as with the components of computer programs, agents can be connected and composed into larger systems called *societies of agents*. Together, societies of agents can perform functions more complex than any single agent could, and ultimately produce the many abilities we attribute to minds.

Minsky’s vision of the mind as a society gives the reader a familiar yet powerful metaphor for organizing the great complexity of the human mind, for a society is almost by definition not a uniform or unified system, and instead, is composed of a great many different types of individuals each with a different background and different role to play. Yet we must be careful – the societies of The Society of Mind should not be regarded as very much like human communities, for individual humans are ‘general purpose’, and individual agents are quite specialized. So while the concept of a society is a familiar notion, this metaphor is only a starting point, and the theory raises a host of questions about how societies of mind might actually be organized.

This article examines the Society of Mind theory. We first give the reader a brief sense of the history of the development of the theory. From where did these ideas originate? Then, we will return to the questions that the theory raises, and describe some of the mechanisms that it proposes. What are agents? How do they work? How do they communicate? How do they grow? Finally, we consider related developments in Artificial Intelligence since the the publication of the Society of Mind.

2 THE EARLY DEVELOPMENT OF THE SOCIETY OF MIND THEORY

The Society of Mind theory was born in discussions between Marvin Minsky and Seymour Papert in the early 1970s at the MIT Artificial Intelligence Lab. One of

the world's leading AI labs, its explorations encompassed diverse strands of research including machine learning, knowledge representation, robotic manipulation, natural language processing, computer vision, and commonsense reasoning. As a result, it was perhaps clearer to this small community than any other at the time the true complexity of cognitive processes.

The severity of this issue may have been confronted for the first time in the famous 'copy-demo' project. Toward the end of the 1960s, Minsky and Papert and their students built one of the first autonomous hand-eye robots. Its task was to build copies of children's building-block structures it saw through cameras using a robotic hand. Minsky recalls:

Both my collaborator, Seymour Papert, and I had long desired to combine a mechanical hand, a television eye, and a computer into a robot that could build with children's building-blocks. It took several years for us and our students to develop Move, See, Grasp, and hundreds of other little programs we needed to make a working Builder-agency... It was this body of experience, more than anything we'd learned about psychology, that led us to many ideas about societies of mind. [2, Section 2.5]

In particular, Minsky and Papert found that no single algorithm or method was adequate for solving even the simplest-seeming problems like assembling towers of blocks.

In trying to make that robot see, we found that no single method ever worked well by itself. For example, the robot could rarely discern an object's shape by using vision alone; it also had to exploit other types of knowledge about which kinds of objects were likely to be seen. This experience impressed on us the idea that only a society of different types of processes could possibly suffice. [2, Postscript and Acknowledgement]

Ultimately, these kinds of experiences led Minsky and Papert to become powerful advocates of the view that intelligence was not the product of any simple recipe or algorithm for thinking, but rather resulted from the combined activity of great societies of more specialized cognitive processes. However, there were few ideas at the time for how to understand and build systems that engaged in thousands of heterogeneous cognitive computations. The conventional view within AI for how problem-solving systems should be built could well be summarized by this statement by Allen Newell from 1962:

The problem solver should be a single personality, wandering over a goal net much as an explorer wanders over the countryside, having a single context and taking it with him wherever he goes. [3]

But in Minsky and Papert's experience, this 'explorer' was being overwhelmed by the sheer magnitude of tasks and subtasks encountered in ordinary commonsense activities such as seeing, grasping, or talking. The emergence of this unanticipated procedural complexity demanded a theory for how such systems could be built.

Some of Minsky's early thoughts about how to approach this problem appear in his famous paper *A Framework for Representing Knowledge* [4], in which he considers a variety of ideas for how to organize the collections of the procedural and declarative knowledge needed to solve many commonsense problems such as recognizing visual scenes and understanding natural language. He summarizes the motivation for frames as follows:

The 'chunks' of reasoning, language, memory, and 'perception' ought to be larger and more structured, and their factual and procedural contents must be more intimately connected in order to explain the apparent power and speed of mental activities.

In the original MIT AI Lab Memo form of its publication, he includes a section written by his student Scott Fahlman, to which Minsky later added the following introduction:

The following essay was written by Scott Fahlman (in 1974 or 1973?), when a student at MIT. It is still one of the clearest images of how societies of mind might work. I have changed only a few terms. Fahlman, now a professor at Carnegie-Mellon University, envisioned a frame as a packet of related facts and agencies – which can include other frames. Any number of frames can be aroused at once, whereupon all their items – and all the items in their sub-frames as well – become available unless specifically canceled. The essay is about deciding when to allow such fragments of information to become active enough to initiate yet other processes.

Systems of frames with attached procedures were thus the ancestor of the concept of an agent. So while the Society of Mind theory had not yet been given a name in 1974, the roots of the theory were clearly present in the work that was going on at the time. Related notions in the programming language community, such as the development of Smalltalk and other object-oriented programming languages, were inspiring people to discover the advantages of new, more 'cellular', ways to think about organizing programs [5].

Computational expressions of the idea of societies of interacting agents began to emerge in the middle of the 1970s. An early articulation of these ideas took form in Carl Hewitt's concept of 'Actors' [6], a computational model where sets of concurrently active agents solved problems by exchanging messages. Hewitt observed that this new model seemed to greatly simplify the control structure of complex programs:

As a result of this work, we have found that we can do without the paraphernalia of "hairy control structure" (such as possibility lists, non-local gotos, and assignments of values to the internal variables of other procedures in CONNIVER.) . . . The conventions of ordinary message-passing seem to provide a better structured, more intuitive, foundation for constructing the communication systems needed for expert problem-solving modules to cooperate effectively.

However, Minsky and Papert's specific ideas about societies of agents were rather different from the systems that their students were developing. They were anticipating a range of problems that to this day have not been fully addressed by the AI community. For example, they regarded agents to be too simple to understand the languages of most other agents, and most messages too complicated to be passed around like currency. Instead, as we will discuss later, in the Society of Mind agents use forms of communication that are often more indirect.

Minsky and Papert's ideas continued to develop, and soon they decided that they would write a book about their emerging Society of Mind theory:

In the middle 1970s Papert and I tried together to write a book about societies of mind but abandoned the attempt when it became clear that the ideas were not mature enough. The results of that collaboration shaped many earlier sections of this book. [2, Postscript and Acknowledgement]

While they ended up abandoning the book, the draft was by 1976 quite substantial, and included the following early descriptions of a Society of Mind:

The mind is a community of "agents". Each has limited powers and can communicate only with certain others. The powers of mind emerge from their interactions for none of the Agents, by itself, has significant intelligence. [...] Everyone knows what it feels like to be engaged in a conversation with oneself. In this book, we will develop the idea that these discussions really happen, and that the participants really "exist". In our picture of the mind we will imagine many "sub-persons", or "internal agents", interacting with one another. Solving the simplest problem – seeing a picture – or remembering the experience of seeing it – might involve a dozen or more – perhaps very many more – of these agents playing different roles. Some of them bear useful knowledge, some of them bear strategies for dealing with other agents, some of them carry warnings or encouragements about how the work of others is proceeding. And some of them are concerned with discipline, prohibiting or "censoring" others from thinking forbidden thoughts. [7]

This early draft was never published, and eventually Papert turned towards developing novel theories of education that built on these ideas, and Minsky continued to develop the Society of Mind theory on his own. Aspects of the theory emerged in pieces throughout the late 1970s in a series of papers. In [8], Minsky describes the Society of Mind as follows:

The present paper is in part a sequel to my paper (Minsky 1974) and partly some speculations about the brain that depend on a theory being pursued in collaboration with Seymour Papert. In that theory, which we call "The Society of Minds", Papert and I try to combine methods from developmental, dynamic, and cognitive psychological theories with ideas from AI and computational theories. Freud and Piaget play important roles. In this theory, mental abilities, both "intellectual" and "affective" (and we ultimately reject the distinction) emerge from interactions between "agents" organized

into local, quasi-political hierarchies. Overall coherency of personality finally emerges, not from any clear and simple cybernetic principles, but from the interactions, under elaborate genetic control, of communities of do-ers, “critics” and “censors”, culminating in almost Freudian agencies for self-discipline that compare one’s behavior with fragments of “self-images” acquired at earlier stages of development.

And later, in [10], Minsky included the following tantalizing description of a Society of Mind:

One could say but little about “mental states” if one imagined the Mind to be a single, unitary thing. But if we envision a mind (or brain) as composed of many partially autonomous “agents” – a “Society” of smaller minds – then we can interpret “mental state” and “partial mental state” in terms of subsets of the states of the parts of the mind. To develop this idea, we will imagine first that this Mental Society works much like any human administrative organization. On the largest scale are gross “Divisions” that specialize in such areas as sensory processing, language, long-range planning, and so forth. Within each Division are multitudes of subspecialists – call them “agents” – that embody smaller elements of an individual’s knowledge, skills, and methods. No single one of these little agents knows very much by itself, but each recognizes certain configurations of a few associates and responds by altering its state.

In all of these papers Minsky did not offer a detailed specification of the theory, but rather took a higher-level perspective:

Some workers in artificial intelligence may be disconcerted by the “high level” of discussion in this paper, and cry out for more lower-level details. [...] There are many real questions about overall organization of the mind that are not just problems of implementation detail. The detail of an AI theory (or one from psychology or from linguistics) will miss the point, if machines that use it can’t be made to think. Particularly in regard to ideas about the brain, there is at present a poverty of sophisticated conceptions, and the theory below is offered to encourage others to think about the problem. [8]

We encourage the reader to read these early papers to gain more insight into the Society of Mind theory, for in our view the theory cannot be fully understood without studying the course of its development.

3 HOW DO SOCIETIES OF MIND WORK?

Despite the great popularity of the book *The Society of Mind*, there have been few attempts to implement very much of the theory. One difficulty is that Minsky presents the theory in fragments and at a variety of levels, and the more ‘mechanical’ aspects of the theory are largely distributed throughout the text, and only specially distinguished in the glossary. To support those interested in implementing Societies of Mind, this section reviews several of the specific mechanisms and representations that Minsky describes.

3.1 What Are Agents?

Minsky sees the mind as a vast diversity of cognitive processes each specialized to perform some type of function, such as expecting, predicting, repairing, remembering, revising, debugging, acting, comparing, generalizing, exemplifying, analogizing, simplifying, and many other such ‘ways of thinking’. There is nothing especially common or uniform about these functions; each agent can be based on a different type of process with its own distinct kinds of purposes, languages for describing things, ways of representing knowledge, methods for producing inferences, and so forth.

To get a handle on this diversity, Minsky adopts a language that is rather neutral about the internal composition of cognitive processes. He introduces the term ‘agent’ to describe any component of a cognitive process that is simple enough to understand, and the term ‘agency’ to describe societies of such agents that together performs functions more complex than any single agent could.¹ Many agencies can be seen as unitary agents by ignoring their internal composition and considering only their external effects. Minsky notes:

I recycled the old words “agent” and “agency” because English lacks any standardized way to distinguish between viewing the activity of an “agent” or piece of machinery as a single process as seen from outside, and analyzing how that behavior functions inside the structure or “agency” that produces it. [14]

In the Society of Mind, mental activity ultimately reduces to turning individual agents on and off. At any time, only some agents in a society of mind are active, and their combined activity constitutes the ‘total state’ of the mind. However, there may be many different activities that are going on at the same time in different agencies, and Minsky introduces the term ‘partial state of mind’ to describe the activities of subsets of the agents of the mind.

3.2 What Are the Simplest Agents?

To Minsky, agents are the building blocks of the mind. He describes several ‘primitive’ agents out of which larger agencies can be constructed.

K-lines. K-lines are the most common agent in the Society of Mind theory. The purpose of a K-line is simply to turn on a particular set of agents, and because agents have many interconnections, activating a K-line can cause a cascade of effects within a mind. Many K-lines are formed by ‘chunking’ the net effects of a problem solving episode, so that the next time the system faces a similar problem, it not only has the previous solution as a starting point, but also the experience of deriving that solution, which includes memories of false starts,

¹ Curiously, the term ‘agency’ has seen little use, and people often use ‘society of agents’ instead. A ‘society of mind’, then, is the ‘total’ set of agents in a mind.

unexpected discoveries, and other lessons from the previous experience that aren't captured by the final solution alone. Thus K-lines cause a Society of Mind to enter a particular remembered configuration of agents, one that formed a useful society in the past. K-lines are a simple but powerful mechanism for disposing a mind towards engaging relevant kinds of problem solving strategies, forms of knowledge, types of goals, memories of particular experiences, and the other mental resources that might help a system solve a problem.²

Nomes and Nemes. Minsky also uses K-lines in a somewhat different way, as the basis for a more structured or computer-like model of how information is represented and processed in a Society of Mind. He introduces two general classes of K-lines he calls 'nemes' and 'nomes', which are analogous to the data and control lines in the design of a computer. Nemes are concerned with *representing* aspects of the world, and nomes are concerned with *controlling* how those representations are processed and manipulated.

Nemes. Nemes invoke representations of things, and are mostly produced by learning from experience. Minsky gives examples of a few types of nemes, including *polynemes* and *micronemes*.

Polynemes invoke partial states within multiple agencies, where each agency is concerned with representing some different aspect of a thing. For example, recognizing an apple arouses an 'apple-polyneme' that invokes certain properties within the color, shape, taste, and other agencies to mentally manufacture the experience of an apple, as well as brings to mind other less sensory aspects such as the cost of an apple, places where apples can be found, the kind of situations in which one might eat an apple, and so forth. Polynemes support the idea that a thing's 'meaning' is best expressed not in terms of any single representation, but rather in a distributed way across multiple representations.

Micronemes provide 'global' contextual signals to agencies all across the brain. Often, they describe aspects of our situation that are too subtle for words and for which we otherwise have no more determinate concepts, such as specific smells, colors, shapes, or feelings. Micronemes are also used to refer to aspects of a situation that are difficult to attach to any particular thing (such as a particular object or event), and are instead more diffuse or indefinite in their reference.

Minsky suggests that nemes are organized into 'ring-closing networks', great societies of nemes and recognizer-agents for those nemes in which activity spreads in both bottom-up and top-down directions, so as to engage in context-sensitive

² In Minsky's forthcoming book *The Emotion Machine* [9], K-lines have been elevated to the role of selectors for emotional states. Each such 'selector K-line' turns on a set of resources in the mind, disposing it towards using certain representations, selecting or prioritizing sets of goals, retrieving particular fragments of knowledge, invoking certain strategies for using this knowledge, and all the other aspects of a particular 'cognitive-emotional state'.

pattern matching processes for recognizing objects, parsing sentences, producing plans, reducing ambiguities, and so on. Because these kinds of parsing and recognition processes are subject to garden path phenomena and other forms of getting stuck, Minsky suggests that many additional agents are involved in regulating the ring-closing process to ‘weed out’ problems in the network’s activity.

Nomes. Nomes control how representations are manipulated. Minsky gives examples of a few types of nomes, including *isonomes*, *pronomes*, and *paranomes*.

Isonomes signal to different agencies to perform the same uniform type of cognitive operation. For example, they may cause a set of agencies to save their current state to short-term memory and load in a different state, or cause them to begin training a new long-term K-line to reproduce the current state, or cause them to imagine the consequences of taking a certain action.

Pronomes are isonomes that control the use of short-term memory representations. A pronome is often associated with a specific ‘role’ in a larger situation or event, such as the actor who takes an action, or the location an event occurs. Some pronomes connect to very restricted types of short-term memories that can store only very specific types of knowledge, for example, places, shapes, or paths. Other pronomes may be much more general-purpose and influential, reaching most of the agencies of the brain. Minsky calls these ‘IT’ pronomes, and they connect to a great many representations that together can describe virtually anything. There may be only just a few of these ‘IT’ pronomes, because of the great many connections required to implement them.

Paranomes are sets of pronomes linked to each other so that assignments or changes made by one pronome to some representation produce corresponding assignments or changes by the other pronomes to related representations. Minsky introduces the concept of a paranome to describe how knowledge represented in different ways could nonetheless be related and treated together in a somewhat uniform manner. For example, a ‘location’ paranome might be connected to location pronomes attached to two different representations of spatial location, one in terms of an egocentric or ‘body-centered’ coordinate system, and the other in terms of an external or ‘third-person’ coordinate system. Using paranomes, one can coordinate the use of these multiple representations.³

3.3 How Can We Combine Agents to Build Larger Agencies?

Minsky gives several examples in the Society of Mind of how one might combine primitive agents into larger agencies that can do more complex things. Several of these examples are concerned with how larger representational systems might be

³ In *The Emotion Machine* [9], Minsky uses the term ‘panalogy’ to refer to the collection of linked partial states across each of these multiple representations, the set of analogous descriptions controlled by the pronomes of the paranome.

built from simpler agents, and in particular, he describes how several types of frames might be built from simple elements like pronomes and recognizer-agents.

Frames. Frames are a form of knowledge representation concerned with representing a thing and all the other things or properties that relate to it in certain particular ways, which are attached to the ‘slots’ of the frame. Minsky describes how simple frames may be built from sets of pronomes that control the attachments to the slots of the frame. These pronomes are bound together so that when the frame is invoked, those pronomes cause their associated representations to invoke partial descriptions of aspects of the thing being described.

Frame-arrays. We can describe a thing richly by using not just one frame, but rather a collection of frames where each frame describes the thing from some particular perspective or point of view. Frame-arrays are collections of frames that have slots or pronomes in common. Minsky gives the example of representing the appearance of a cube from multiple different viewpoints, where each of those viewpoints is described with its own frame, but whose common parts (e.g. the six sides of the cube) are linked across different frames. The reason for sharing slots is that when one frame description is inadequate for solving some problem or representing some situation, it is easy to switch to one of the other frames, because some of their slots are already attached to relevant information. The shared slots of frame-arrays are the ancestors of paranomes.

Transframes. Transframes are a central form of knowledge representation in the Society of Mind theory. Transframes represent events and all of the entities that were involved with or related to the event. They may have slots for representing the *origin* and *destination* of a change (the before and after states), who or what caused the event, the motivation behind the event or the goal it is intended to achieve (if there was indeed an intention), what objects are affected and how, when it occurred, what tools or objects were involved in producing the change, and other important aspects of the event. Because so much of ordinary life concerns the relationship between events of various types, the transframe representation is central to everyday thinking.

Other types of frames. In addition to transframes, Minsky describes several other types of frames, including story-frames that represent structured collections of related events, and picture-frames that represent the spatial layout of objects within scenes. Presumably these are only a few of the many types of frames that are required to represent and organize the world as well as cognitive processes themselves.

3.4 How Can Agents Solve Problems?

While Minsky argues that there no single method that societies of agents use to solve problems, he does suggest several ways one might build or organize problem solving agencies.

Difference-Engines. What does it mean to ‘solve’ a problem? Solving a problem can be regarded as reducing or eliminating the important differences between the current state and some desired goal state. Minsky proposes a simple machine called a *difference-engine* that embodies this problem solving strategy. Difference-engines operate by recognizing differences between the current state and the desired state, and acting to reduce each difference by invoking K-lines that turn on suitable solution methods. The difference-engine idea is based on Newell and Simon’s early ‘GPS’ problem solver [11]. Minsky elevates the GPS idea to a central principle, and one might interpret Minsky as suggesting that we view the mind as a society of such difference-reducing machines that populate the mind at every level. But ultimately, there is no single mechanism for building difference-engines, because there is no single way to compare different representations.

Censors and Suppressors. No method of problem solving or reasoning will always work, especially when it comes to ordinary, commonsense reasoning. Thus Minsky proposes that in addition to knowledge about problem solving methods themselves, we also have much knowledge about how to avoid the most common bugs and pitfalls with those methods. He calls this type of knowledge *negative expertise* [12]. In the Society of Mind he describes this knowledge as embodied in form of *censor* and *suppressor* agents. *Censors* suppress the mental activity that precedes unproductive or dangerous actions, and *suppressors* suppress those unproductive or dangerous actions themselves. Minsky suggests that such negative expertise could even form the bulk of what we know, yet remain invisible because knowledge about what *not* to do does not directly manifest itself. Further, he suggests that there is an intimate connection between humor and such negative expertise; when we laugh at a joke, we may be learning about a particular type of problem or pitfall with ordinary common sense reasoning! Minsky discusses these ideas extensively in [13].

A-brains and B-brains. Some types of unproductive mental activity are not specific to any particular method, such as ‘looping’ or ‘meandering’, which might occur in any problem solving method that engages in search. Minsky introduces the notion of the ‘B-brain’ whose job is not so much to think about the outside world, but rather to think about the world inside the mind (the ‘A-brain’), so as to be able to notice these kinds of errors and correct them. This division of the mind into ‘levels of reflection’ is an idea that has become even more central in Minsky’s more recent theories.⁴

⁴ In *The Emotion Machine* [9], Minsky gives these levels separate names: the ‘reactive’, ‘deliberative’, ‘reflective’, ‘self-reflective’, ‘self-conscious’, and ‘self-ideals’ levels, where each level is concerned with representing and responding to problems in the levels beneath.

3.5 How Do Agents Communicate with Each Other?

In the Society of Mind, agents use different internal representations, and so they must interact with each other without knowing very much about how the others work. Minsky recognized early the difficulty in trying to formulate a theory of cognition that assumed consistent meanings for cognitive signaling:

For agents to use symbols that others understand, they would need a body of conventions. We want to bypass the need for dispersed but consistent symbol definitions. [8]

In fact, as Minsky points out, the simplicity of agents only makes this more challenging:

The smaller two languages are, the harder it will be to translate between them. This is not because there are too many meanings, but because there are too few. The fewer things an agent does, the less likely that what another agent does will correspond to any of those things. [2, Section 6.12]

The mechanisms that Minsky proposes for agent communication in the Society of Mind are designed with these considerations in mind.

K-lines. The simplest method of communication is for an agent to just arouse some other sets of agents. An agent can turn on a polyneme to arouse agents that think about some particular object, event, or situation; or it may turn on micronemes that cause other agents to think about some general context; and so forth. For this type of communication it is not necessary for the sender agent itself to know how to express an idea in terms of the representations available to the recipient agent. Rather, that information is stored in the intervening K-lines that result from successful past communications.

Connection lines. Many agents are not directly connected to each other but rather communicate via ‘connection lines’, buses or bundles of wires that transmit signals to other agents attached to the bus. These wires can be thought of as simple agents in themselves, and while they are initially meaningless, over time the individual wires begin to take on local significance, that is, they come to acquire dependable and repeatable ‘meanings’. Agents communicate over these connection lines by connecting K-lines to random subsets of the bus wires. This strategy of random connection, first invented by Calvin Mooers, allows a relatively small bus to simultaneously represent a great many independent symbols or states with a low probability of collision between them. On the other side, agents can observe these connection lines to begin to recognize patterns and signals in the wires. Just as the senders do not initially know what the wires mean, neither do the recipients, and so the agents that connect to the bus need to make guesses and hypotheses to come to understand their meanings.

Internal language. For agencies that need to communicate more complex, structured descriptions of things, as in a system of linked frames, Minsky proposes

a more elaborate communication mechanism in [14], modeled after his ‘re-duplication’ theory of how people communicate with each other through natural language. If one agent wishes to convey a complex idea to another, it attempts to *re-construct* the idea expressed in its own representational system by a sequence of frame retrieval and instantiation operations. For each of these operations there is an associated ‘grammar-tactic’ that produces words or signals that are sent to the recipient agent. ‘Inverse-grammar-tactics’ in the recipient agent perform analogous construction operations, but now in terms of the different representations used by the recipient agent. This method of communication requires that the two communicating agents agree well enough on the meanings of these ‘words’ or representation-construction operations.

Paranomes. Perhaps the most common method of communication in the Society of Mind is actually for there to be *no active communication* at all. Instead, agents often find that the information they need is already available when they need it. This is the result of the use of paranomes. As described earlier, when one pronome of a paranome produces a particular state in terms of one representation, the other pronomes simultaneously update their representations so that they enter corresponding states. Here, communication happens not by sending explicit messages, but rather at all times different agencies are separately looking at the same property, object, event, situation, or other type of thing from their own unique perspective, and any changes to one representation are immediately reflected in the other corresponding representations.

Ambiguity. While this is not so much a communication mechanism in and of itself, an important consideration in how agents communicate in a Society of Mind is that precise communication may be unnecessary, and in fact, may be impossible, between different agents.

We often find it hard to “express our thoughts” – to summarize our mental states or put our ideas into words. It is tempting to blame this on the ambiguity of words, but the problem is deeper than that. Thoughts themselves are ambiguous! [...] The significance of any agency’s state depends on how it is likely to affect the states of other agencies. [...] It is an illusion to assume a clear and absolute distinction between “expressing” and “thinking”, since expressing is itself an active process that involves simplifying and reconstituting a mental state by detaching it from the more diffuse and variable parts of its context. [...] We can tolerate the ambiguity of words because we are already so competent at coping with the ambiguity of thoughts. [2, Section 20.1]

Thus, while all these methods of communication no doubt produce misunderstandings, both because there is no agreed upon convention for the meanings of symbols, and because meanings themselves are not stable entities, Minsky argues that this state of affairs really cannot be avoided, and we must find ways to build societies of agents that are tolerant to some degree of vagueness and ambiguity in both their thoughts and communications.

3.6 The Growth of Mental Societies

How do societies of mind ‘grow up’? Minsky suggests that mental societies are constructed over time, and that the trajectory of this process differs from person to person. He offers several potential mechanisms for growth.

Protospecialists. In an infant mind, the first functional large-scale agencies are *protospecialists* – highly evolved agencies that produce behaviors providing initial solutions to problems like locomotion, obtaining food and water, staying warm, defending yourself from predators, and so forth. While these initial protospecialists are fairly unskilled, over time their performance may improve by exploiting agencies that develop later.

Predestined learning. Related to the notion of a protospecialist is the idea that complex behaviors need not be fully pre-specified nor fully learned, and instead can result from a mixture of more partial such influences. Minsky suggests many of the kinds of abilities that are shared among all people, e.g. language or walking, are the result of *predestined learning*, learning that develops under just enough internal and external constraints that the end result is more or less guaranteed.

Types of learning. Minsky describes several important forms of learning: accumulating, uniframing, transframing, and reformulating. *Accumulating* is the simplest form of learning, where you simply remember each example or experience as a separate case. *Uniframing* amounts to finding a general description that subsumes multiple examples. *Transframing* is forming an analogy or some other form of bridge between two representations. *Reformulation* amounts not to acquiring fundamentally new knowledge per se, but finding new ways to describe existing knowledge.

Learning from attachment figures. A very important type of learning is concerned with the question of how we learn our goals in the first place. This form of learning is not so much about how to acquire the specific representations and processes needed to achieve some goal, but rather how to learn when a particular goal should be adopted and how it should be prioritized relative to our other goals. Minsky suggests that we learn many of our goals through interactions with our ‘attachment figures’, special people in our lives, such as our parents, whom we respect. Praise and censure from our attachment figures results in ‘goal learning’ as opposed to ‘skill learning’; when we are praised by an attachment figure for solving some problem, we choose to seek out more such problems in the future, and when we are censured by them, we choose to avoid such problems.

Learning mental managers. As a mind grows up, it acquires not only increasingly sophisticated models of its environment, but also builds increasingly sophisticated cognitive processes for making use of those models – knowledge about *when* and *how* to use knowledge. As we accumulate our mental societies of

agents, we build ‘mental managers’ to regulate them, processes for delegating, controlling, repairing, and selecting more specific knowledge and problem solving techniques. This idea led Minsky to coin *Papert’s Principle*:

Some of the most crucial steps in mental growth are based not simply on acquiring new skills, but on acquiring new administrative ways to use what one already knows. [2, Section 10.4]

These sorts of managerial hierarchies often make use of a conflict resolution strategy that Minsky calls the *Principle of Non-Compromise*. When two agents disagree, rather than employing some simple method such as averaging or weighting, mental managers can instead use the conflict as a sign that they need to reformulate the problem by bringing in some third agent with a different perspective, one that can see the problem and its potential solution with greater clarity. This goes hand in hand with the idea of distributing authority in a mind; the job of the manager is not so much to directly resolve the conflict, but rather to find other points of view that cause the conflict simply to disappear from the new vantage point.

Developmental stages. Ultimately, nothing as complex as a mind can be built in a single stage of construction. This idea was on Minsky’s mind from his earliest thoughts about societies of mind:

The proper form of a “theory of mind” cannot focus only on end-products; it must describe principles through which: An earlier stage directs the construction of a later stage. Two stages can operate compatibly during transition. The construction skill itself can evolve and be passed on. [8]

Minsky discusses the idea that the mind develops in multiple stages, and in particular, that these stages can be regarded as training each other. The mind can be seen as the result of a sequence of teaching operations where at each stage there is a ‘teacher’ that teaches the ‘student’, and where the student becomes the teacher at the next stage, and so on.

4 RECENT DEVELOPMENTS

Since the publication of the Society of Mind there has been much work that is similar in spirit to Minsky’s ideas, some of which was directly inspired by his theory. This section samples some of this research.

4.1 Combining Symbolic and Connectionist Methods

Since the Society of Mind was published there has been much work on developing ‘connectionist’ systems that recognize patterns, make inferences, and solve problems by distributing the solution process across societies of simple computational

elements. One of Minsky's purposes in developing the Society of Mind theory was to establish a framework which naturally incorporated both symbolic and connectionist notions, yet despite the compatible presentation Minsky gives in the Society of Mind, the debate between symbolic and connectionist approaches continues to rage to this day. Minsky reviews the apparent false dichotomy in [1].

There have been a number of attempts since to combine the advantages of these two kinds of approaches. In his review of the Society of Mind, Dyer presents several alternatives to Minsky's ideas for how to represent frames in a connectionist framework [15], including his own work on parallel distributed semantic networks [16], and the SHRUTI architecture [17] of Shastri. A different approach, focusing more on planning than general purpose inference, was suggested by Maes, who implemented a connectionist, distributed agent based planning system that was directly inspired by the Society of Mind [19].

4.2 The Rise of Statistical Inference

Statistical inference methods, for example, reasoning with Bayesian networks, are currently the most popular technique for reasoning under uncertainty. Typically, such methods operate on probabilistic graphical models, networks of random variables linked by their conditional or joint probabilities. There is a strong similarity between such graphical models and the networks of neme-agents described in the Society of Mind. Both are kinds of structures that represent the dependency relationships between things and their properties and parts. While Minsky does not give the arousal of his agents a probabilistic interpretation, his ring-closing mechanism for recognition, which combines top-down and bottom-up constraints, is similar in some ways to Pearl's popular message-passing belief propagation procedure for inferencing in Bayesian networks [20].

Despite these similarities, there are many important differences between the agencies of the Society of Mind and graphical models; in our view, many of these differences point to serious deficiencies in current statistical systems, ones that may cause them to eventually adopt more ideas from the Society of Mind. For example, there are no 'nomes' in graphical models, short-term memory units that can temporarily capture fragments of state during reasoning. These may be needed to engage in more complex kinds of parsing and interpretation processes, where complex hypotheses can be quickly formulated and dismissed, and to allow portions of the network to be reused for different purposes in the same larger computation. Also, there is no special procedural control knowledge in graphical models for how to propagate probabilities through the network, other than a small set of belief updating rules that are applied uniformly throughout the network. Such procedural control knowledge, which in the Society of Mind is embodied in mental managers, is useful for guiding inference systems towards answers in large search spaces, and is likely critical for scaling statistical inference methods to networks with millions of nodes and links connected in elaborate topologies, the scale that is encountered in building human-like commonsense reasoning systems.

4.3 Procedural Circuits of Simple Computational Elements

The connectionist and statistical movements have led to many ideas about how inference systems might be understood and implemented as networks of simple computational elements. However, there has been less work on building *programming languages* based on the Society of Mind, ones that take higher level descriptions of cognitive processes and compile them into circuits of simple computational elements. Most researchers instead develop their AI systems using existing programming languages such as Lisp, or using rule-based systems. One attempt to build something like a Society of Mind programming language is Brooks's Behavior Language [21]. Brooks's early robots had control structures composed of networks of simple computational elements, and to make it easier to program these robots he developed the Behavior Language, which on the surface resembles Lisp, but compiles programs into finite state automata which are then run directly. A more recent attempt to build such a language is Hearn's K-line Language, which uses K-lines as a primitive element and builds on his prior work, a behavior system for a simulated robot that used polynemes, pronomes, picture-frames, transframes, and other simple elements from the Society of Mind to direct the actions of the robot [22].

The major difficulty with such approaches is that it has been hard to find good 'higher level' abstractions for agencies, ones that make it easy to describe agent systems that use higher level representations and methods for problem solving, and can map these abstractions to simple elements like K-lines and recognizer-agents. It remains an open question whether a programming language at the level of the simplest computational elements will ultimately be the language in which developers program, or more like a 'machine language' to which some even higher level programming language compiles down to. In either case, however, it is clear that developing such a language is a useful endeavor.

4.4 Case-Based Reasoning

There are many questions about the details of how to implement K-lines. How do you retrieve an appropriate K-line? How do you then adapt the prior solution retrieved by the K-line to fit the current problem situation? How do you store and index these adaptations when you are done? How are successes and failures recorded? Since the publication of the Society of Mind, the field of *case-based reasoning* has grown to study these kinds of questions [23]. In fact, an important early development in case-based reasoning, derivational analogy [24], came directly from the K-line theory. Derivational analogy (now called *replay* in the case-based reasoning field) operates by retrieving not just the final solution to a problem solved in the past, but rather by retrieving and re-applying the entire search tree that recorded the process of discovering that past solution, including the failed branches. While expressed in less connectionist terms than Minsky's K-line theory, it captures some important aspects of the theory.

Of the techniques that have been developed in the AI community for reasoning and learning, the methods that have been developed by the case-based reasoning community seem to be the most similar in spirit to Minsky's ideas about K-lines. However, no one has combined a connectionist representation scheme and an elaborately structured case-based reasoning architecture to produce a literal implementation of Minsky's K-lines.

4.5 Blackboard Systems

Societies of Mind make use of multiple representations and methods of reasoning to solve problems. *Blackboard systems* [25] have long been used to build problem solvers that integrate heterogeneous AI techniques. A blackboard system consists of multiple 'knowledge sources' that observe a 'blackboard' that is read from and written to by these knowledge sources. Each knowledge source can make use of a different type of knowledge, style of representation, or method of reasoning.

While the blackboard model does indeed let you use multiple representations, building effective blackboard systems is not a simple matter of just programming the various knowledge sources and letting them run concurrently. There have been serious challenges in finding good ways to control the inferencing that happens on the blackboard [26]. While the blackboard metaphor may work when there are only a few agents using the blackboard, by the time there are hundreds of agents, let alone thousands or millions, the image of them huddled around a blackboard is no longer reasonable, and in fact no one has built a blackboard system of this scale. The difficulty is both that most agents do not understand the representations of other agents, and when agents do happen to be able to interact, there are often conflicts of various kinds. The blackboard model offers no special solutions to these kinds of problems, but these were precisely the kinds of problems that Minsky and Papert had in mind when they were first formulating the Society of Mind theory. These problems may not be solvable without making use of more knowledge about which types of knowledge to apply when, that is, without more 'managerial' expertise within the system about problem-solving itself.

4.6 The Cyc Knowledge Base

To cope with the everyday world, a person's Society of Mind must embody a great deal of knowledge about how the world works. Much of this knowledge is shared by most people in our culture, for example, that it is darker in the night than in the day, that drinking water will quench one's thirst, or that an object can't be in two places at once. This is what we call 'commonsense' knowledge. The Cyc project [27] is the largest and most ambitious attempt to build a database of such commonsense knowledge, and today Cyc contains over a million facts and rules about the everyday world. Cyc's knowledge is organized into 'microtheories', resembling Minsky's agencies, that each use their own knowledge representation scheme and make their own sets of assumptions. The microtheory idea is derived

from McCarthy's notion of contexts [28] and is a powerful way to organize the Cyc knowledge base into quasi-independent regions that are internally consistent but can be mutually inconsistent. These microtheories are linked via 'lifting rules' that allow translation and communication of expressions between microtheories. Perhaps the greatest contribution of the Cyc project is not so much its knowledge base, but the vast ontology of terms and predicates that it uses to express its knowledge; Cyc likely has what is the most expressive such ontology in existence. Any Society of Mind with common sense will need at least as diverse a variety of representations.

While Cyc is undoubtedly a great accomplishment, it is not yet a Society of Mind in the sense Minsky describes. Cyc focuses on the kinds of knowledge that might populate the A-brain of a Society of Mind, that is, it knows a great deal about the kinds of objects, events, and other entities that exist in the external world. However, it knows far less about *cognitive processes* themselves, knowledge about how to learn, reason, and reflect. As Cyc begins to incorporate more such knowledge, it will come much closer to being a Society of Mind.

Another commonsense reasoning system that more intimately mixes declarative with procedural knowledge, and which was directly inspired by the Society of Mind, is Mueller's ThoughtTreasure system [29], a story understanding system with a great variety of commonsense knowledge about how to read and understand children's stories.

4.7 The Soar Cognitive Architecture

The Soar cognitive architecture, described in Allen Newell's magnum opus *Unified Theories of Cognition* [30], in a sense represents the opposite of what Minsky desired with his Society of Mind theory. Rather than seeking to celebrate the diversity of the mind, Newell instead seeks to ground human intelligence in terms of a small set of basic mechanisms and underlying representations. However, in many ways the Soar theory is remarkably similar to the Society of Mind theory. The best summary of these similarities was probably by Minsky himself, in his review of Newell's theory [31]. Minsky notes some striking correspondences, for example, the similarity between Soar's chunking mechanism and the Society of Mind's production of K-lines, or between Soar's impasse mechanism for selecting new problem spaces and the Society of Mind's selection by mental managers of new agents to invoke given conflicts between currently active agents.

But despite these similarities, there are many differences between these theories. The major difference is almost philosophical: While Soar seeks the commonalities between different aspects of cognition, the Society of Mind instead seeks the differences. Thus even though one could potentially build on top of Soar a great many varied devices – societies of agents, even – typically one does not. To the developers of Soar, the interesting question is what are the least set of basic mechanisms needed to support the widest range of cognitive processes [32]. The opposing argument of the Society of Mind theory is that the space of cognitive processes is so broad that no particular set of mechanisms has any special advantage; there will always be some

things that are easy to implement in your cognitive architecture and other things that are hard. Perhaps the question we should be asking is not so much how do you unify all of AI into one cognitive architecture, but rather, how do you get several cognitive architectures to work together? That is the question that Minsky poses in trying to find ways for diverse agents and agencies to live together in a mind.

However, beyond even these differences between Soar and the Society of Mind, there is one most curious difference. While Soar has seen a series of implementations, the Society of Mind theory has not. Minsky chose to discuss many aspects of the theory but left many of the details for others to fill in. This, however, has been slow to happen.

4.8 Multiagent Systems

The modern field of multiagent systems has made progress in answering many of the kinds of questions one might ask about how one might build a Society of Mind. Researchers have proposed many ideas about how agents should communicate, how they might coordinate their different goals, how they might work together to plan solutions to problems, and so forth. There are now a wide variety of architectural ideas about how to build multiagent systems (see for example [33] for one summary of recent work.)

However, while much progress has been made in understanding and finding ways to build multiagent systems, there has been less progress on the specific kind of multiagent system that Minsky proposes in the Society of Mind: a richly heterogeneous architecture that uses multiple representations and methods, has special support for reasoning about commonsense domains (such as time, space, causality, physics, other minds, and so on), that can make use of an initial endowment of commonsense knowledge, that can reflect upon and improve its own behavior, that grows and develops in stages, and that is based upon not any one architectural arrangement of agents but rather is capable of reconfiguring itself to produce many distinct ‘ways to think’ [34].

5 THE FUTURE

The Society of Mind summarizes a lifetime of work by a founder of and perhaps the most influential single individual in the field of artificial intelligence. It provides the first articulation of a computational theory of mind that takes seriously the full range of things that human minds do. In this article we could only examine a small fraction of the full Society of Mind theory, and we encourage the reader to go back to the book and read it cover to cover. *The Society of Mind* is more than just a collection of theories – it is a powerful catalyst for thinking about thinking. Minsky encourages the reader to ponder questions about the mind that they may never have thought to ask, and provides hundreds of examples of how one might start down the road towards answering such questions. In our experience it has been

useful to return to the book once each year, as it is written at level of abstraction where each reading brings out new ideas and reflections in the mind of the reader.

It remains a curious fact that the AI community has, for the most part, not pursued Society of Mind-like theories. It is likely that that Minsky's framework was simply ahead of its time, in the sense that in the 1980s and 1990s, there were few AI researchers who could comfortably conceive of the full scope of issues Minsky discussed – including learning, reasoning, language, perception, action, representation, and so forth. Instead the field has shattered into dozens of subfields populated by researchers with very different goals and who speak very different technical languages. But as the field matures, the population of AI researchers with broad perspectives will surely increase, and we hope that they will choose to revisit the Society of Mind theory with a fresh eye.

The theory itself has not stopped developing. *The Society of Mind* represents a snapshot in time of Minsky's ideas, and he will soon publish a sequel, *The Emotion Machine* [9], that will describe the many ideas that Minsky has had about the mind since the late 1980s. However, we predict that *The Society of Mind* will still be read decades from now, when other AI books have long become outdated, and that it will continue to inspire and challenge future generations. And in time, we expect that the fundamental hypothesis of the book will be finally proved:

What magical trick makes us intelligent? The trick is that there is no trick.
The power of intelligence stems from our vast diversity, not from any single,
perfect principle. [2, Section 30.8]

Acknowledgements

We would like to thank Marvin Minsky and Seymour Papert for discussing some of the early history of the Society of Mind theory, and for providing us with the fascinating but as yet unpublished early 'Brazil' drafts of the Society of Mind.

REFERENCES

- [1] MINSKY, M.: Logical vs. Analogical or Symbolic vs. Connectionist or Neat vs. Scruffy. In: Artificial Intelligence at MIT, Expanding Frontiers, Patrick H. Winston (Ed.), Vol. 1, MIT Press, 1990. Reprinted in AI Magazine, 1991.
- [2] MINSKY, M.: The Society of Mind. Simon and Schuster, New York, 1986.
- [3] NEWELL, A.: Some Problems of Basic Organization in Problem-Solving Programs. Rand Corporation Memorandum RM-3283-PR, December, 1962.
- [4] MINSKY, M.: A Framework for Representing Knowledge. MIT AI Lab Memo 306, 1974.
- [5] KAY, A.: The Early History of Smalltalk. In Proceedings of 2nd ACM SIGPLAN History of Programming Languages Conference, published in ACM SIGPLAN Notices, Vol. 28, 1993, No. 3, pp. 69–75.

- [6] HEWITT, C.: Viewing Control Structures as Patterns of Passing Messages. MIT AI Lab Memo 410, 1976.
- [7] MINSKY, M.—PAPERT, S: The Society of Mind—Unpublished “Brazil” drafts, Oct. 24, 1976.
- [8] MINSKY, M: Plain Talk about Neurodevelopmental Epistemology. In Proceedings of IJCAI-1977, Cambridge, MA, 1977.
- [9] MINSKY, M: The Emotion Machine. Forthcoming.
- [10] MINSKY, M: K-lines: A Theory of Memory. *Cognitive Science*, Vol. 4, 1980, pp. 117–133.
- [11] NEWELL, A.—SHAW, J. C.—SIMON, H. A.: Report on a General Problem-Solving Program. In Proceedings of the International Conference on Information Processing, 1960.
- [12] MINSKY, M.: Negative Expertise. *International Journal of Expert Systems*, Vol. 7, 1994, No. 1, pp. 13–19.
- [13] MINSKY, M.: Jokes and the Cognitive Unconscious. MIT AI Lab Memo 603, 1980.
- [14] MINSKY, M.: A Response to Four Reviews of the Society of Mind. *Artificial Intelligence*, Vol. 48, 1991, pp. 371–396.
- [15] DYER, M.: A Society of Ideas on Cognition: Review of Marvin Minsky’s *The Society of Mind*. *Artificial Intelligence*, Vol. 48, 1991, No. 3, pp. 321–334.
- [16] SUMIDA, R. A.—DYER, M.: Storing and Generalizing Multiple Instances while Maintaining Knowledge-Level Parallelism. In Proceedings of IJCAI-89, Detroit, MI, 1989.
- [17] SHASTRI, L: Temporal Synchrony, Dynamic Bindings, and SHRUTI: A Representational but Non-Classical Model of Reflexive Reasoning. *Behavioral and Brain Sciences*, Vol. 19, 1996, No. 2, pp. 331–337.
- [18] NARAYANAN, S.: Moving Right Along: A Computational Model of Metaphoric Reasoning about Events. In Proceedings of AAAI-99, Orlando, Florida, 1999.
- [19] MAES, P.: How to Do the Right Thing. *Connection Science Journal*, Vol. 1, 1989, No. 3, pp. 291–323.
- [20] PEARL, J.: Probabilistic Reasoning in Intelligent Systems. Morgan Kaufmann, San Mateo, CA, 1988.
- [21] BROOKS, R. A.: The Behavior Language; User’s Guide. MIT AI Lab Memo 1227, 1990.
- [22] HEARN, R. A.: Building Grounding Abstractions for Artificial Intelligence Programming. (Masters Thesis). Department of Electrical Engineering and Computer Science, MIT, 2001.
- [23] AAMODT, A.—PLAZA, E.: Case-Based Reasoning: Foundational Issue, Methodological Variations, and System Approaches. *AI Communications*, Vol. 7, 1994, No. 1, pp. 39–59.
- [24] CARBONELL, J.: Derivational Analogy and its Role in Problem Solving. In Proceedings of AAAI-83, Washington, D. C., 1983.
- [25] NIL, H. P.: Blackboard Systems: The Blackboard Model of Problem Solving and the Evolution of Blackboard Architectures. *AI Magazine*, Vol. 7, 1986, No. 2, pp. 38–53.

- [26] CARVER, N.—LESSER, V.: Evolution of Blackboard Control Architectures. *Expert System with Applications*, Vol. 7, 1994, pp. 1–30.
- [27] LENAT, D.: CYC: A Large-Scale Investment in Knowledge Infrastructure. *Communications of the ACM*, Vol. 38, 1995, No. 11, pp. 33–38.
- [28] MCCARTHY, J.: Notes on Formalizing Context. In *Proceedings of IJCAI-93*, Chambéry, France, 1993.
- [29] MUELLER, E. T.: *Natural Language Processing with ThoughtTreasure*, Signiform. New York, 1998.
- [30] NEWELL, A.: *Unified Theories of Cognition*. Harvard University Press, Cambridge, MA, 1990.
- [31] MINSKY, M.: Review of “Allen Newell, *Unified Theories of Cognition*”. *Artificial Intelligence*, Vol. 59, 1993, pp. 343–354.
- [32] LAIRD, J. E.—ROSENBLUM, P. S.: The Evolution of the Soar Cognitive Architecture. In D. M. Steier and T. M. Mitchell, (Eds.), *Mind Matters: A Tribute to Allen Newell*, pp. 1-50. Erlbaum, Mahwah, NJ, 1996.
- [33] WEISS, G. (Ed.): *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*. MIT Press, Cambridge, MA, 2000.
- [34] SINGH, P.—MINSKY, M.: In *Proceedings of the International Conference on Knowledge Intensive Multi-Agent Systems*, Cambridge, MA, 2003.



Push Singh is a doctoral candidate in MIT’s Department of Electrical Engineering and Computer Science. His research is focused on finding ways to give computers human-like common sense, and he is presently collaborating with Marvin Minsky to develop an architecture for commonsense thinking that makes use of many types of mechanisms for reasoning, representation, and reflection. He started the Open Mind Common Sense project at MIT, an effort to build large-scale commonsense knowledge bases by turning to the general public, and has worked on incorporating commonsense reasoning into a variety of real-world applications. Singh received his B.S. and M.Eng. in Electrical Engineering and Computer Science from MIT.