Introduction

Mike Travers

> When your ideas seem inadequate, remember someone more expert at this, and imagine what that person would do.
>
> —Marvin Minsky (Essay 4)

We have some preconceived ideas about genius—that it is a divine mystery, a gift given to just a few, something both inexplicable and inaccessible to ordinary people. Marvin Minsky was a recognized genius, but it was the nature of his particular form of genius to question the very idea of innate talent and aptitude and the mystifications that surround it. To Marvin, the idea that the mind (or anything else) was beyond explanation was an affront and a challenge.

Marvin was brilliant in numerous ways aside from the work on artificial intelligence for which he is best known: he was an accomplished inventor, mathematician, and musician. But his big trick was to face squarely the mechanical nature of the human mind and not be alarmed by it. Indeed, he found it rather delightful and intriguing. This put him at odds with standard-issue humanists, which suited him just fine. But Marvin himself was not in any way inhuman, far from it. He was an extremely warm and welcoming individual, and always willing to engage

with anyone's open mind. For example, he was active in Usenet discussion groups (chiefly comp.ai.philosophy), where he would explain or argue his views on AI with all comers regardless of their academic qualifications.

Marvin had a unique talent for reflecting on mental processes and coming up with plausible mechanisms that might explain how they worked. He applied this ability to his own mind, of course, but also to the thinking of his students, mentors, colleagues, and friends. He delighted in the cleverness of mental machinery, and sought to understand it by modeling it and encouraging others to share in the task of self-understanding.

Minsky is usually identified with computer technology and artificial intelligence, but there is little in the following essays that is specifically about digital technology or, indeed, technology of any sort. Their focus is on the psychology of learning, and the nature of systems that are capable of learning. He hoped to build computers along these lines, but his inspiration was the human mind, and these essays are both reflections on how minds work and concrete suggestions for ways to reorganize education to better support them—in some cases by involving the intelligent deployment of computer technology, to be sure, but never as the central focus. Technology is merely a tool in pursuit of better understandings, and computer programming merely the best available language for expressing new ideas about how minds work.

Minsky's last two books, *Society of Mind* (SOM) and *The Emotion Machine* (TEM), were both crammed with simple, concrete, and powerful ideas for how minds—both natural and artificial—might be constructed, distilled from decades of work building computational models and guiding the work of others. Both were essentially technical books written in nontechnical language—

a choice that may have caused some problems with their reception, as people did not quite know how to read them. But in this and in many other ways, Marvin gleefully ignored the standard boundaries and rules.

Situating These Essays

Five of the essays collected here emerged from Marvin's participation in the One Laptop per Child (OLPC) project, a massive effort to put computational technology in the hands of the world's children.[1] Marvin saw this as an opportunity to fix some of the ingrained bad habits of the educational system. For instance, essay 5 opens with the suggestion that the educational focus on broad general education is misplaced, and children would be better served by a system that allowed them to specialize and dive into a single topic they cared about deeply.

Regardless of the practicality of this idea, it's notable in what it indicates about his approach to learning. To Marvin, students were not empty vessels to be filled with knowledge, nor poor approximations of fully developed adults, but instead understood as fully active agents and creators of their own minds. As such, they needed to hone their mental skills on the kind of demanding tasks that come with intense pursuit of some personal goal. Traditional education provides content, but creators require methods and tools.

Marvin was finely attuned to the power of ideas, both good and bad, and the OLPC project was a chance "*to provide our children with ideas they could use to invent their own theories about themselves*" (essay 6).This was a call for a model of education entirely in tune with the ancient Greek injunction to "know thyself." Marvin saw that computers and computational ideas have

the potential to give us much better tools for accomplishing this task than have ever existed before in history. The computational revolution in human self-knowledge has been enabled but not yet realized, and these essays may be considered efforts to move us toward that goal.

Themes

Marvin's thinking revolved around a number of tightly interconnected ideas, which pop up repeatedly in his writing here and elsewhere. Here I try to pull out and summarize a few of his most important recurring themes.

The Centrality of Goals

> We need to develop better ways to answer the questions that kids are afraid to ask, like "*What am I doing here, and why?*"
>
> —Marvin Minsky (Essay 2)

> You're almost always pursuing goals.
>
> —Marvin Minsky (Essay 4)

The importance of goals and the design of goal-directed machinery was one of the founding principles of AI and its ancestor, cybernetics. Essay 6 contains a short explanation of the General Problem Solver (GPS), an early goal-directed AI architecture. Marvin's psychological theories (SOM/TEM) may be seen as designs for how minds could contain and manage systems of goals: innate drives, learned goals, sub-goals, interpersonal goals, and meta-goals, all interoperating to produce intelligent behavior.

Goals have a special role in the educational context because an important and neglected key to learning a field of knowledge is *internalizing its goals*. For example, students often find it

hard to learn math because it's unconnected to their personal lives and goals. Teachers need to communicate not just the facts and techniques of a field of knowledge, but its goal structures as well—what questions the field is trying to answer, and what makes those questions interesting. This is something standard education typically only accomplishes accidentally when it accomplishes it at all. For instance, you can learn lots of facts about biology without learning to understand and share the deep motivations of the scientists who discover them. Good teachers can sometimes transmit the goals of a field, but more often than not this doesn't happen, leading to students who are alienated from school, dutifully memorizing knowledge detached from the motivations that led to its discovery, and with no ability to connect it to their own goals.

An emphasis on goals does not mean exhortation to mindless efforts—rather the opposite:

> One needs to learn not only what works, but also what to do when failure looms. I don't like that tale of "*The Little Engine that Could*" with its helpless injunction to simply repeat "*I think I can, I think I can*." A better motto would be to think "*perhaps it's time to try something else*" because every setback can offer a chance for a new phase of mental development. (Essay 6)

In Marvin's view of intelligence, skillful thinking requires not only having goals but being able to think about and modify them.

Construction from Parts

> The secret is in finding out how much can come from so few kinds of parts.
>
> —Marvin Minsky (Essay 1)

Minsky details some of his early fascination with the construction sets of his childhood: Tinkertoy, Meccano, and so on. From this came the central insight that simple parts could be used to construct arbitrarily complex structures, and that these structures had properties of their own, independent of the nature of their components. Brains, minds, and computers all share this compositional quality, although with the latter we have the advantage of being able to know exactly what the parts are, how they behave, and how they relate to each other and to the overall properties of the system.

Programming languages are also construction kits, with parts that children can recombine in novel ways. Systems like Logo and Scratch act as procedural Tinkertoys, enabling exploration, modeling, and discovery in a new and dynamic domain.

A computer made out of Tinkertoys can implement the exact same computation as one made out of circuitry—and so we assume a properly constructed computer could implement the same kinds of mental processes as the brain. With systems composed of parts, the important thing is the relationship between them, not the physical substrate they are made of.

The universality and substrate-independence of computation, the equivalence of computation to any effective process, and the equivalence of computation and mental activity are all deep, powerful, and in some cases controversial ideas. In an educational context, we do not expect to settle these questions—but giving children the tools to make models of their own thinking and behavior allows even the very young to engage with these fundamental issues.

Other Minds

Social processes are crucial to developing high-level goals, and Marvin has a number of insightful critiques and proposals in

this area. Consider a term he coined: *imprimers*—the people from whom one learns foundational goals and values. These can be parents, teachers, or peers, but in any case they play a key role in learning because the goals they impart serve to focus and drive everything else a mind does.

Marvin's emphasis on the social nature of learning might come as a surprise to those accustomed to the usual emphasis on the mechanisms of individual minds that is the default methodological stance of AI. Certainly the AI of Marvin's period of greatest activity did not pay a great deal of attention to the social embeddedness of learning and intelligence. But Marvin was not one to let the current limits of computation interfere with his forward-looking theories of mind. More recently, the social transmission of goals has resurfaced as the focus of attempts to mitigate the supposed existential risks of AIs by achieving "value alignment."[2]

Networks as Escape

The vision of OLPC was to build not only a computer for all the world's children but also a network that would connect them with each other and with the wider culture. Marvin saw this as an opportunity for the intellectually inclined student—so often neglected and bullied in mainstream school culture—to find remote mentors or peers on the net. This vision has been realized to some extent by the later development of online learning communities,[3] such as the one centered around the Scratch programming environment from the Lifelong Kindergarten Group at the MIT Media Lab (http://scratch.mit.edu).

Multiplicity

> Until you understand something more than one way, you don't really understand it.[4]

> It's also important to know multiple ways to represent things, so that if one method gets stuck, you can switch to another.
>
> —Marvin Minsky (Essay 5)

Individual minds are composed of a multiplicity of different parts that are skilled in different ways of thinking. And because everyone's mix is different, individual learners must necessarily develop their own cognitive styles.

One recurrent theme of Marvin's writing on learning is how wonderful it is to find new, nonstandard ways of solving problems—and how the standard model of education tends to suppress such methodological creativity in favor of teaching a single, supposedly right way to do things. The constructionist model of education upends this approach by giving learners a rich set of combinatorial parts that facilitates exploring a space of possible approaches.[5]

The Centrality of Reflection

> Human Minds think about what they're thinking about. … I'm convinced that these "self-reflective" processes are the principal ones that people use for *developing new ways to think*.
>
> —Marvin Minsky (Essay 6)

If there is one grand unifying idea to Marvin's techniques for generating insight and his proposals to fix education (and he might have denied that there was), it is the central importance of reflection: thinking about thinking. Everyone necessarily has to think about their own thinking—it's part of being a human being—but most of our ideas about ourselves are not that good and can be improved.

Minsky believed that computers could be a tool for reflection but also that reflexive heuristics could be identified, named, and

taught. The role of computers in education is not merely to be a substitute for teachers or libraries but to be a language and toolkit for creating models, and particularly self-reflective models. And the product of the attempts of Minsky and Papert and their many students to realize this vision is not improved math scores but children who can think deeply about processes, systems, and themselves.

Conclusion

Digital technology has revolutionized the world in many ways and is continuing to do so as it evolves. But one of the key insights of the early AI days seems in danger of getting lost—that computational ways of thinking are powerful intellectual tools, not just for building games and websites but for understanding complex systems, especially minds.

Computation, in other words, provides a sophisticated language for modeling, and one that is accessible to the young and naive. The trick is to encourage children to think of themselves as computers, and to think of computers as potentially human. It was obvious to Marvin (and Seymour) that this idea was not merely valid but enormously *generative* of new insights. Unsurprisingly, it ran counter to cultural prejudices, which viewed machinery as inherently antihuman. To call something mechanical, in common usage, carries strong connotations of mindlessness. Marvin fought against this dichotomy all of his life, and devoted most of his career to figuring out how to make mindful machines.

There are many reasons to pursue the dream of artificial intelligence, whether they be scientific, economic, or simply deriving from whatever force causes life and intelligence to try to

replicate themselves in new forms. But there is a subtler reason that I think is implicit in these essays, which is that almost all of our preconceived ideas about the mind are severely wrong and broken (including such ancient concepts as willpower, freedom, consciousness, and innate ability), and this limits us and causes us needless suffering. Computational ideas give us a radically new way to see ourselves, and their power as a tool for human self-reflection has barely been touched. Getting these ideas into the hands of more people, especially children and other learners, may be one of the most important things we can do for human progress.