P R E F A C E

About a year and a half ago, while I was visiting Stanford University, Penny Nii suggested to me the similarities between journalism and a curious new field called "knowledge engineering." Nii is a computer scientist whose specialty is artificial intelligence, the effort to make machines that think. Knowledge engineers interview experts—in medicine, for example—and translate some of what they know into a form that can be fed into a computer. Then, if all goes well, the computer will be able to make decisions as good as or better than those of its human counterparts.

That's the theory, anyway. Knowledge engineering is still in its infancy; it remains more a craft than a science. As I explored various artificial-intelligence laboratories around the country, I found that the design of so-called expert systems is only a tiny corner of a sprawling, wonderfully diverse attempt to use computers and programs to lay bare the mysteries of thought and create intelligence outside of the human body. Though my research took me further and further from the more mundane realm of knowledge engineering, I found that my thoughts kept returning to that conversation with Nii.

Knowledge engineers, she had explained, must absorb themselves in a field they don't belong to and emerge with enough understanding to write a proper program. They must be able to describe a subject without learning every detail, to capture expertise without devoting the years it takes to become an expert.

Journalists find themselves in much the same situation, though their aim is to communicate not with a computer but with an audience of curious nonspecialists who hope to get a feel for what scientists do without becoming scientists themselves. The knowledge engineers do their work using tools such as data structures, algorithms, programming languages—the greater and lesser arcana of the field of computer science. In the case of the journalist, the tools are metaphor, description, example, narrative technique—in short, the English language. Using these devices, one must try to take the strange and relate it to the familiar, communicate an unknown subject by building on what people already know. A journalist who is writing a book about artificial intelligence cannot become a computer scientist. No matter how much you read or how many people you interview, you are always in the frustrating position of knowing less than the people you are writing about—experts with Ph.D.s and years or decades of experience.

How then is science writing possible? The answer, I believe, is that the journalist has another kind of expertise, the ability to cut through the details of a subject and see its essence, to identify the important themes and personalities of a field and synthesize them into a story that a general reader will find enticing and clear.

One of the most engaging things I found about artificial intelligence is that it drives you to introspect, to turn your powers of observation inward and wonder how you do what you do—and how you would describe it precisely enough to get a computer to do it. Penny Nii challenged me to introspect about journalism, to think of myself as a computer programmed to go about the land interviewing scientists, reading papers and books, attending conferences, and eventually writing a book about artificial intelligence. What are the procedures I would follow?

In computer science, researchers often talk about what they call levels of abstraction. Anything can be described from many different levels. At the least abstract, most detailed level—the lowest level of abstraction—a human being can be thought of as an aggregate of whirling electrons, protons, neutrons, and other subatomic particles. But such a description is so complex that it is almost useless. Describe things at too low a level and the details overwhelm. On a slightly higher level, all these particles can be seen as parts of atoms, and at the next level up, atoms as parts of molecules, and then molecules as parts of cells. As, level by level, we climb the ladder of abstraction, the descriptions become less detailed, more general. We see that cells form tissues, which form organs, which are crucial components of the body.

Depending on our needs and interests, we can describe an organism from any one of these vantage points, using the languages of physics, chemistry, molecular biology, cell physiology, anatomy, zoology, or psychology. If we want to move to even higher planes and consider the human as part of a culture or society, we can use the languages of anthropology and sociology. Each of these levels rests on the ones below it, but each has its own rules. Psychologists can view the whole creature as a thing exhibiting behaviors that seem completely unrelated to atomic physics, but it is atoms from which we are made.

It is because things can be described from a number of different levels that journalists and knowledge engineers are able to do their jobs. In researching a book, a writer must constantly move up and down between levels of abstraction. At times you fly high overhead looking down on the contours below. When you see something that catches your fancy, you swoop down for a closer look. As you drop lower and lower, the details become more pronounced. At some point they become too formidable, so you must decide whether to stay long enough to make sense of them all or fly skyward again and find another part of the landscape to explore.

This same jumping between levels occurs while writing. You can use only a fraction of what you have learned. So, during most of the narrative you give the reader a bird's-eye view. You concentrate on the themes and implications—the big picture. But occasionally it is important to descend earthward and muck around in the details. Only then can you get a real feel for what is going on.

One of the hardest things about science writing is striking the right balance between the high and low levels, between the general and the detailed. Scientists often write books that are too technical for a general reader. On the other hand, they have been warned so many times to keep things simple that sometimes they overdo it. They generalize too much, underestimating the willingness of their readers to learn. One gets a taste of the subject but leaves the book hungry for the meat.

Morally, philosophically, economically, and scientifically, the implications of artificial intelligence are so great that I believe it is important for us, as outsiders, to have a solid understanding of the field. Is it really possible to make machines that think? Much writing on the subject begs the question. The standard line of argument goes something like this:

(1) A computer is a device that can do anything that can be precisely described. (2) Anything, including thinking, can be precisely described. (3) Therefore, computers can do anything, including thinking, and since they are faster than brains and will soon have bigger memories, inevitably they will be smarter than us. It's just a matter of working out the technical details. Laymen can assume that these are being taken care of. Meanwhile we can sit back and ponder a wonderful new future, or a brave new world. This kind of reasoning has led to some writing about computers that, at its worst, is half science, half science fiction.

Artificial intelligence is a field that lends itself to grand (and sometimes wild) speculation. That is part of its allure. In this book, scientists will describe future scenarios as fantastic as those we've seen in the movies. They really believe them, and, as experts, they are qualified to speculate with a certain amount of authority. But for those of us who are not computer scientists to take these predictions on more than faith we must understand the research behind them. We shouldn't be dazzled by looking at things from too high a plane. Unless we believe that there is something magical about human intelligence that can never be captured by a machine, we will probably concede that artificial intelligence is at least theoretically possible. Perhaps it *is* just a matter of working out the details, but as any scientist would agree, the details are everything.

"I've never heard any convincing argument that humans represent any kind of level of intelligence that it is impossible to go to with machines," Patrick Winston, director of MIT's artificial-intelligence laboratory, once told me. "But the interesting issue is not whether machines can be made smarter but if humans are smart enough to pull it off. A raccoon obviously can't make a machine as smart as a raccoon. I wonder if humans can."

This book is very much grounded in the present, in the fantastic things that are happening now. How *do* you get a machine to understand language, both written and spoken? How do you get a computer to make sense of what it sees, to learn from experience, to make discoveries and create on its own? How close are we to making machines that do these things? How far can we go, and how soon? It is one thing to say that computers can embody any step-by-step process and that thinking proceeds in a step-by-step manner, but it is quite another to show how such subtle and complex procedures actually might be mechanized. What do we mean when we say "understand," "learn," "discover," "create"? If we can describe these vague notions precisely enough to program them, then perhaps we'll have a more solid grasp of what goes on inside the mind.

In writing this book, I am assuming that most people know little about computers. They should find all that they need in these pages. Computers are becoming so pervasive that someday it won't be necessary to explain them anew each time a book such as this is written. But that day has not yet arrived. And even a veteran computer hand may find that research in artificial intelligence presents old ideas about computing in a surprising new light.

One might suppose, after flipping through a textbook or journal devoted to machine intelligence, that it would be impossible for a layman to understand a subject like this, except in the most superficial terms; that it would require laborious training, or at least a close familiarity with computers. I don't believe that is true. Some subjects, like quantum physics, are impossible to penetrate without knowing a great deal of mathematics; the general reader is stuck forever at the fringe. In artificial intelligence the concepts are more intuitive. By using the tools of analogy, metaphor, and example, and by viewing the key concepts from several different angles, it is possible to home in on the subject through a sort of literary triangulation. Then it can be examined in enough detail to leave the reader satisfied.

At the same time, I recognize that readers of popular science books are divided between those who savor the occasional detailed explanation for the sense of mastery it gives them and those who would rather stick to the broader bird's-eye view. I've made efforts to accommodate both tastes. In the prefaces to their eleven-volume history of the world, *The Story of Civilization*, Will and Ariel Durant often issued a disclaimer: "Certain especially dull passages, not essential to the story, are indicated in reduced type." I can't muster such courage (anyway, I think the details are the most interesting part), but in a few places where I am about to launch into a fairly involved description—for example, how a program called Hearsay-II understands a spoken sentence, or how the Automated Mathematician rediscovered arithmetic—I will try to drop a subtle hint. Those willing to take it on faith that the programs live up to their billings can skim for a few pages. But I've worked hard to make these descriptions accessible—and I hope compelling—to anyone with a natural curiosity for how things work. In the course of my research I found it exciting that I, a nonscientist, could understand these things. This book is an effort to share that sense of discovery.

Finally, I am writing not just about how machines work, but how scientists work as well. I've chosen not to write about the people on the level of detail appropriate to a profile or a biography, but from a higher level of abstraction. Ultimately, what I hope emerges is a portrait of a profession, with all its politics, infighting, and other blemishes as well as its triumphs and discoveries. I've tried to capture not only what it is like to do science, but to be at the beginning of a very young science, one that may lead to a whole new way of looking at the world.

Artificial intelligence is often compared to molecular biology. Both fields have grown rapidly; both are suddenly becoming commercialized. But there is more to the similarity than that. Each of these sciences is about patterns, structures, and codes. In molecular biology, scientists work to explain how four nucleotides form a language that tells a cell how to arrange twenty amino acids into the complex system we call life. Artificial intelligence is about how symbols representing things in the world can be woven into the language of the mind, the patterns we call thinking.

As I wrote this book, the work I held in the back of my mind was *The Eighth Day of Creation*, Horace Freeland Judson's grand and panoramic history of molecular biology. The field of artificial intelligence is too new for so definitive a treatment. It is still awaiting a breakthrough as fundamental as Watson and Crick's discovery of the structure of DNA and all that it said about reproduction, inheritance, and other mysteries of life. For me, Judson's work was not so much a model as an inspiration, a demonstration that it is possible to write about a science in a manner that is entertaining and accessible, while rendering it with enough precision to make it come alive.