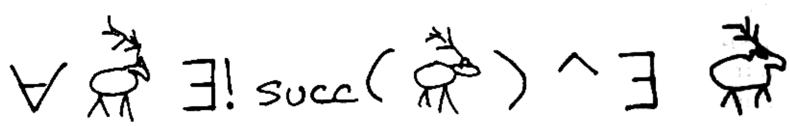
AI Magazine Volume 10 Number 2 (1989) (© AAAI)

The Mind at AI: Horseless Carriage to Clock

William C. Hill

Commentators on AI converge on two goals they believe define the field: (1) to better understand the mind by specifying computational models and (2) to construct computer systems that perform actions traditionally regarded as mental. We should recognize that AI has a third, hidden, more basic aim; that the first two goals are special cases of the third; and that the actual technical substance of AI concerns only this more basic aim. This third aim is to establish new computation-based representational media, media in which human intellect can come to express itself with different clarity and force. This article articulates this proposal by showing how the intellectual activity we label AI can be likened in revealing ways to each of five familiar technologies.

0738-4602/89/$3.50 © 1989 AAAI.



Chic, Modern, Advanced Representational Technology.

Idiot Cro-Magnon Representational Technology.

AI is not about building artificial intelligences, nor is it about understanding the human mind or any other kind of mind. A more fundamental human urge is playing out, that of seeking new expressive power through new expressive means.

I've stated this thesis with a strength that belies the case. However, I state it this way to clarify a position I find myself moving toward rather than one I hold firmly and defend easily. The thesis fits my technical experiences with AI, equips me with a new ensemble of criteria for assessing what's significant and insignificant about new computational proposals, and serves as a common analytic point of view for the many AIs that now exist (Papert 1988). Obviously, some researchers who label their work as AI claim they are constructing intelligences, and others model the brain or mind in efforts to understand it. However, if we doubt whether the former is desirable or even possible as some do and just continue to call the latter biology, philosophy, or psychology, why bother with AI at all? Despite this route of dismissal, the vitality of uncovering the expressive facilities and limits of new models of computation continues and I contend that this is what AI research is really about.[1] New models of computation are likely to emerge in the future, and where these models lead, questions about their relevance for representing—or, more tendentiously, hosting—mentality will follow.

Intrigue with the possibilities of representing or hosting mentality has long motivated AI effort. Thus, commentators on AI converge on two goals they see defining the field: (1) to better understand the (human) mind by specifying computational models and (2) to construct computer systems that perform actions traditionally regarded as mental. However, the consensus expressed by the recognition of these two aims is not the only way to regard AI and, perhaps in the long run, not even the most significant. We should recognize that AI has a third, hidden, more basic aim; that the first two goals are special cases of the third; and that the actual technical substance of AI concerns only this more basic aim. This third goal is establishing new computation-based representational media, media in which human intellect can come to express itself with a different clarity and force.

By *representational media*, I mean physical means of recording with their associated conventions, for example, ochre on cave wall pictographs, wet clay and stylus cuneiform, vellum and pen calligraphy, paper and pencil writing, canvas and paint likenesses and abstractions, stone and chisel sculpture, printing and spelling, photography and photographic composition, scoring and musical composition, vinyl discs, album organization and album notes, blueprinting and architectural diagramming, film and cinematic convention, videotape and video style, programming languages and programming style, logical notations and their calculi, and compact discs with their new search and sequencing options.[2] These examples share the common role of enabling representational tasks. As such, a common representational analysis applied to them would demonstrate how the ways in which they differ are tuned to the tasks for which they are designed. This article proceeds from the premise that AI techniques fall within this category of representational media and that such an analysis of AI techniques as representational media calls for and organizes many significant observations concerning AI, computers, and the mind.[3]

Why propose that AI is primarily concerned with the establishment of new representational means and conventions? It is because these means and conventions exactly comprise AI's research vehicles and eventual impacts, although recognizing them as such is not a particularly popular vocation. On the whole, cognitive scientists and AI researchers prefer to see their contributions in terms of their aspirations—to understand or create intellect. Regardless of aspirations, systematic acts of arranging computation to simulate or represent intellectual processes or states cannot help but effect new representational means and conventions. In this regard, consider GRAPES (Sauers 1982), GOMS (Card, Moran, and Newell 1983), SOAR (Laird 1986), PUPS (Anderson and Thompson 1987), PPS (Bovair, Kieras, and Polson 1988), genetic rule populations (Holland 1975, 1986), Hopfield nets (Hopfield and Tank 1985), and back propagation (Rumelhart, McClelland, and the PDP Research Group 1986). Furthermore, this perspective raises questions about the representational powers of computations that tend to remain submerged when considering AI solely in terms of its prospects for mimicking or constructing intelligences.

This proposal originates in view of humankind as a representing animal—that what it means to be **humanly** intelligent is tightly bound to the use of external representations in language and other media. In a phrase, homo sapien means homo depictor (Hacking 1983). In the sense that this view locates intelligence both inside and outside the cranium, this perspective shares ground with Simon's "ant on the beach" argument as he applied it to humankind (Simon 1969, p. 23 ff.) (that is, apparent cognitive complexity reflects environmental complexity), and Norman's (1988) "knowledge in the world" psychology of everyday things. We now turn to some of computation's unique representational properties to set the stage for its analysis.

Representational Properties of Computation Exploited by AI Efforts

New representational media established on the ground of computation exhibit novel representational properties, in particular, new levels of dynamics and depth and new forms of facility and interactivity. What do these properties entail? First, by indicating dynamics as a representational property, I mean how computation realized as activity in electronic circuits corresponds along the dimension of time with models of activity in the world, thereby empowering their straightforward, moment-for-moment representation. With respect to this dynamics property, consider painting or, more generally, two-dimensional, static images as contrasting examples. Painting is not particularly adequate for conveying dynamics but good for still life. Of course, two-dimensional, static images are often used to convey dynamics by mapping time onto a spatial or other quality, but the moment-for-moment advantages are lost. In contrast, by exploiting moment-for-moment correspondence, along with film, video, and audio, computation records and conveys dynamics well.

Second, by depth as a representational property, I mean how computation can be arranged to hide or make vivid aspects of its processing, thus conveying the impression of surface and depth, outside and inside, behavior and internal mechanism—surface behavior for the aspects highlighted and depth of internal mechanism for the aspects hidden. Unlike previous mechanical or electronic media, what is new here is the ease with which one can move aspects from the hidden to highlighted category and back again. Third, by facility as a representational property, I mean the perceived match between thoughts about some phenomenon a user intends to represent and the repertoire of operations and displays afforded by a computational regime. As an example of this facility property, consider how the operations and displays afforded by an electronic spreadsheet program facilitate financial planning. While planning finances, one is thinking about time periods, dollar figures, and expenditure categories. A paper-and-pencil ledger allows one to organize these kinds of elements nicely. Yet, an electronic version greatly eases the calculation and recalculation over the ledger according to the column and row equations one specifies. It also eases spatial reorganization of the ledger. These operations and displays afford a new level of representational

What is new is the greater extent of computation's capability for representing dynamics, depth and interactivity and, in particular, that the means by which computation facilitates representation are themselves extensible.

facility to support financial planning. Fourth, by interactivity as a representational property, I mean how computation can be arranged to allow its context to influence its trajectory. For example, users communicate with a program using a keyboard and mouse.

Given these properties of dynamics, depth, facility, and interactivity, computation is particularly handy for representing change through time, giving the impression of an outside and an inside, abstracting operations and operands constructively to coincide with thoughts about the way the world is and affording continued interaction with ongoing computation.[4] However, we should remember that these novelties, which AI efforts build on, are only the recent technical possibilities in a long, rich, variegated representational tradition.[5]

Thus, the nature of computation affords new representational possibilities. However, what is really new in the mentioned properties? The fact that computation is capable of mediating dynamics, depth, facility, and interactivity does not alone distinguish it from previous mechanical and electronic media, media in which devices such as looms, organ works, piano mechanisms, clockworks, jack hammers, combustion engines, phonographs, radios, and televisions also mediate dynamics and interactivity with an impression of depth. What is new is the greater extent of computation's capability for representing dynamics, depth, and interactivity and, in particular, that the means by which computation facilitates representation are themselves extensible. This last property makes computation a medium like no other. For the first time, it is easy not only to be deliberate about representational facility but easy to create and compare various configurations of representational facility.[6] For example, one can create various forms of representational languages, one on another; choose some form of defeasibility as a desirable representational property and do something about it; build up a new inference-control architecture; or simulate models of agent societies, rule ecologies, neuron bundles, or whatever. After all, aren't AIers, as they huddle around whiteboards and terminals, actually being deliberate about representational facility? This extensible representational facility property of computation acts as the essential practical core fueling and refueling AI imaginations and efforts.

The point deserves emphasis. In technical terms, AI research exhibits a profound gravity that continuously attracts even the best planned efforts back toward tweaking and playing around with representational facility. For this reason, AI dissertations begin with grand themes about the mind, brain, intelligence, or personhood and then slowly, as the tougher questions about the nature of intellect resist answers, end up primarily contributing a new representational regime with different facility for the particular portion of the fabric of intellect under study. This tendency in AI research explains why when a new representational scheme catches on in the AI community, bestowing renown on its creators, it's rarely because of any adequacy as a theory of mind but because one really can do something new and nifty in the scheme, or there's a proof about its limits (that is, news that one could never do something nifty, old, and hoped-for in the scheme). For this reason, efforts to axiomatize common sense never get around to axiomatizing much common sense but evolve to concentrate on tweaking representational facility to handle this or that new case.

AI Is Ordinary

Considering AI techniques and programs as recent steps in a long representational tradition emphasizes their place in the world as results of human intellectual activity—more in their role as representational products of the mind than as direct representations of mind, per se. In this light, they are not significantly different from mapmaking techniques and maps, painting techniques and paintings, or dancing techniques and dances. Representation is an ordinary human activity and so is AI. I make this point because it remains fashionable to portray AI as somehow exceptional but essentially divorced from its context as an ordinary human activity.

The adoption of this emphasis on AI techniques and programs as ordinary intellectual products rather than extraordinary machine realizations of intellectual activity nudges some traditionally central issues about AI and the mind to the periphery and bumps others to the center. For example, the acclaimed questions "What aspects of human intellect are in principle computable?" and "What aspects of human intellect will as a matter of practicality be computed?" remain open—evocative but peripheral. Other questions move toward the center; for example, "How do representational properties of particular AI techniques serve and shape intellectual goals and products?" "What is human cognition individually and socially such that we might know best how to facilitate its expression in computation?" "What exactly are people doing when they, as researchers, engineers, and users, show up at their workplaces to create, modify, and use new representational regimes mediated through computation?" I suggest that addressing these latter questions will start us down a productive path precisely because they presume the relevance of the human dimension in AI.

A Five-Point Representational Analysis

To start down this path, one might examine the ways to clarify the powers and limits of AI technology as a representational medium. Such clarification requires focusing attention on the representational qualities of its techniques. Here, we make five general observations about representational media and pursue their relevance to AI. First, the powers and uses of new representational media are initially misconceived in terms of older, familiar media, leading to the use of the new media to follow the forms prevalent in earlier media. Second, all representational media make some representational tasks easy and others difficult. Third, all representational media establish two characters of experience, one for onlookers and one for workers in the medium. Fourth, all representational media communicate their users' views and values, albeit imperfectly, but with specific and regular manners of imperfection. Fifth, all representational media influence their users' thoughts, and because representational media are typically shared, these influences are usually social.

Consider an example of how these points play in a real AI research effort. They are illustrated in reports of the Imperial College Logic Programming Group's experimental work at the University of London, representing the British Nationality Act of 1981 in Prolog (Sergot et al. 1986).[7] Consider the first point: Representing a law text as a program realizes the essence of the recapitulating form prevalent in earlier media. Indeed, the group worked to preserve the textual and organizational qualities of the act in the program, although perhaps a better use of logic programming more in tune with its powers would be a what if workbench environment for drafting legislation, an idea the team later came on.

Regarding the second point, through long effort the group learned that much of the act could be easily and perspicuously rendered as Horn clauses, which make up the fragment of first-order logic that permits at most one hopeful conclusion and any number of hopeful conditions. However, some of it could not—exactly those parts which exhibit negative conclusions, classically negated conditions, counterfactual conditionals, nonmonotonic conclusions, and judgment. To address some of these difficulties, the team hacked the program's fundamental objects (read "tweaked representational facility") to handle some representational extensions required by the nature of the act, specifically, negative conclusions, some domain-specific negation predicates, and some tedious expansions of counterfactual conditionals. These adjustments were not wholly satisfactory, and the difficulties they were intended to remedy remained.

On the third point, the resulting Prolog program clearly established two characters of experience, one for members of the British Home Office who wrote the legislation and interacted with the program's interface on demonstration occasions and another for the builders of the program. How so? The program builders took care to arrange for the rule code to resemble the original English text and preserve the act's organization. In this attempt, they succeeded. The Home Office officials could also order texts composed by the program to explain why it requested the particular information it did or how it reached its conclusions. Thus, the team's efforts succeeded in establishing an intelligible experience for the onlooking Home Office officials. In contrast, the builders themselves intimately knew the program as a complex compromise between their articulate thoughts about the act and the expressive power of the representational means and conventions at hand (that is, an experimental extension of Prolog).

On the fourth point, the resulting logic program computes and communicates official views and values concerning British citizenship expressed in the act. The computation gives a rigid, mechanistic character to the act, thus communicating values different from what might be experienced with a person applying the act. On the fifth point, the British Nationality Act of 1981 itself was a product of the drafters in the British Home Office group intellect. Another group of six at the university produced its logic program representation with a common tool, and the two groups interacted. During the interaction, members of the programming team pointed out sets of conditions not anticipated by the drafters, and the drafters observed how such programs would be helpful not only as delivery systems for applying the rules but as workbench tools for drafting and redrafting legislation (Sergot et al. 1986). The representational medium fostered these influences because without it, they would not have occurred. The remainder of this article details the relevance of these five representational points for the articulation between AI techniques and intellectual tasks. Within each point, I use illustrations of one or more familiar non-AI examples.

The First Point: The Horseless Carriage

Powers and uses of new representational media are initially misconceived in terms of older, familiar media. The phrase artificial intelligence aligns thinking in such a way as to call attention to superficial resemblances to familiar notions and effectively diverts attention from its technical substance. This diversion remains effective because of the relative historical novelty of computation-based representational media and our consequent lack of experience with them. To date, we know only the first chapter in the story of how people got smarter about using more and more computation for purposes of representation beginning in the second half of the twentieth century. Before the 1970s, no one had rigorously stated the fundamental notion of NP-completeness (Cook 1971, 1973; Garey and Johnson 1979), which now plays such a critical role in estimating the computational resources required by various representational schemes (that is, determining what is computable in a reasonable amount of time or within memory store limitations) and which has played such a salutary role in supressing naive overenthusiasm for AI's prospects. Only in the 1980s have hidden node learning algorithms been acknowledged as capable of producing computationally complete networks, meaning the door is open for net computations to find more uses than previously envisioned on the basis of the early perceptron theorems (Minsky and Papert 1969). Another way of saying what these considerations illustrate is that we are in AI's infancy; which amounts to taking seriously that we do not yet have many inspiring models of computation, that we have not yet piled together much computational "stuff," and that we have not yet explored even a small fraction of the space of representational facility which computation affords.

An interesting parallel can be drawn here. During its own infancy, the printing craft's early authorities, the early printers themselves, referred to printing as artificialiter scribere, or artificial writing (Buhler 1960, p. 16), which it was not, except in naive, superficial terms. This coincidence illustrates how a new technology's founding authorities tend to view their creation in terms later outmoded and retrospectively naive.

What further light might the "new media are misconceived" thesis shed? It suggests that AI is not necessarily about building intelligences but about minds making sense of their worlds for particular purposes, with new representational media as tools, and using these tools to externalize what's in or on the minds in question. It is the mind **at** AI that matters. What deserves emphasis here is that as a culture, we are not yet conceiving of these new representational techniques in terms of their proper powers and limits but rather in familiar terms wholly inadequate, even misleading, for the task of perceiving and then leveraging their lasting significance. In this sense, the phrase artificial intelligence is a *pristechnolocution*, that is, an earlier technology phrase (such as horseless carriage, wireless telegraph, iron horse, glass teletype, or artificial writing), a phrase which describes a new technology wrongly in terms of an old familiar one, a phrase that directs attention toward unimportant resemblances and decades later sounds anachronistically naive.

The AI label offends doubly in this regard. It invites thinking about new technology in familiar terms inadequate to the task. Even worse, it casts new technology not in terms of previous technological phenomena but in terms of previous mental phenomena, raising at no extra cost ferocious issues from the philosophy of mind, which although profound in their own right can do little to clarify the proper powers and limits of emerging, computation-based, representational technology. Not surprisingly, arguments and counterarguments that stroke broadly, purely, and resoundingly over the ground originally covered by Turing's (1950) test or Searle's (1980) Chinese room offer meager technical advice. These arguments address big AI questions about smart AI fictions.

However, the little AI questions about dumb AI realities do more for understanding how to get on with finding new expressive power through new expressive means. Ritchie and Hanna's (1984) analysis of AM, Lenat and Brown's (1984) response, David Chapman's (1987) analysis of planning techniques, and chapter three of *Mind Over Machine* (Dreyfus and Dreyfus 1986) are examples of informative approaches to AI realities conducted in this spirit. These analyses succeed because the authors are careful to level detailed criticism at specific techniques. In this same spirit, we turn to making the point that AI techniques make some tasks easy and others difficult.

AI research exhibits a profound gravity that continuously attracts even the best planned efforts back toward tweaking and playing around with representational facility.

The Second Point: Dust Abacus and Wet Clay

The second point is that representational media make some tasks easier and others difficult. As an example, compare the dust abacus with wet clay for doing simple calculations. A dust abacus is a planar surface covered with wheat dust of a contrasting color. With a stylus, the user strokes numerals into the dust to perform arithmetic in styles similar to the paper and pencil arithmetic we learned in school. One can also perform similar calculations with wet clay and a stylus for stroking numerals into the clay. Now, what can be done in each medium after one commits an error by stroking in the wrong numeral?

The representational medium of the dust abacus makes it easy to correct errors. The user scratches out the stroke by covering it with dust. The stylus might even have an edge for smoothing over the dust. With clay, error correction is not as easy. Depending on the consistency and stickiness of the clay, one can mash and trowel over the offending strokes, possibly smoothing over the result, but the solution takes longer and the effects are not clean. The dust abacus is superior to wet clay for error correction.

Now, suppose we find we need to do lots of quick calculations; let's say we're busy merchants. Give the dust abacus a few good shakes, and we've got a blank dust abacus, ready to go.

Blanking the wet clay is additional work. Throughout the work day, the clay is drying out and getting harder to work with. Of course, we might keep lots of blank clays around, but then we need technology to keep them from drying out before we use them. We might also form clay tablets as we need them from a big lump, but this process is time consuming and disruptive. The dust abacus is again superior for lots of quick calculations.

Finally, suppose we find we need to do calculations and later carry them to our employer's house to be checked or to preserve them so that we can check them next year; let's say we're employees of some merchant or that we carry on seasonal trade. The representational medium of clay that holds its form immediately and later undergoes a transition from soft to hard in the open air is well suited for carrying around or consulting over extended time. Carrying a dust abacus around can disturb its pattern, depending on the distance, terrain, and wind. Additional isolation technology would be required to preserve dust abacus calculations over time. Wind, insects, birds, earthquakes, or accident-prone librarians could all do irreparable damage. The representational medium of clay is superior for movement and preservation.

In sum, the dust abacus and the clay tablet make the various representational tasks they serve easy or difficult. The dust abacus makes error correction and quick repetition easy. The clay tablet makes transportation and preservation easy. This pattern of trading off advantages is no less true of advanced representational techniques. With the dust abacus and wet clay example in mind, consider Horn-clause logic as an illustration of how AI techniques in general make some representational tasks easy and others difficult. To consider this logic, we more closely examine incidents from the experience of the University of London groups in their attempt to formalize the British Nationality Act of 1981 with Horn clauses. Although we carry out the point with respect to Horn clauses, it applies to AI techniques in general. The examples here come from their report. To understand the examples, the reader must know a little about Prolog, a programming language that uses Horn clauses.

Powers and uses of new representational media are initially misconceived in terms of older, familiar media.

Horn-clause logic is the subsystem of first-order logic that permits, at most, one possibly acceptable conclusion term and any number of possibly acceptable condition terms, where terms are atomic relations among individuals. Three types of clauses result: rules, queries, and facts. Rules have both conditions and a conclusion. Queries are degenerate rules that have only conditions and no conclusion. Facts are degenerate rules that state supplied conclusions without conditions. Thus, with the British Nationality Act having rules of the form:

```
Conclude x if conditions y and z and w
hold ,
```

it is straightforward to express the idea that a person is a British citizen if this person was born in the United Kingdom with at least one parent holding British citizenship or that being a parent means being a mother or father to some child:

```
Conclude x is a British Citizen
   if x was born in the U.K.
   and x was born on date y
   and y is after or on commencement
   and z is a parent of x
   and z is a  British Citizen on date y  .
z is a parent of x
   if z is a mother of x  .
z is a parent of x
   if z is a father of x  .
```

It is also easy to ask questions such as "According to some section of the Nationality Act, is Peter a British citizen as of January 16, 1986?" as in the following:

```
?- citizen(peter, (date 16 Jan 1984), SectionOfLaw)
```

You can also state facts relevant to the case, such as "Peter was born in the U.K." and "William is father of Peter":

```
bornInUK(peter).
fatherOf(peter william).
```

With this simple regime, the University of London group formalized a surprising amount of the British Nationality Act of 1981 into an axiomatic theory simple in structure, thus demonstrating that Horn clauses make formalizing some parts of the act easy. However, Horn clauses did not make it easy to formally express other parts of the act; for example, negative conclusions:

A person who . . . shall not become a British citizen under subsection 11.1 unless . . .

classically negated conditions:

A newborn infant, . . . found abandoned in the United Kingdom, shall, unless the contrary is shown, be deemed for purposes of subsection 1.2 . . .

counterfactual conditionals:

. . . became a British citizen by descent or would have done so but for his having died or ceased to be a citizen . . . (by) renunciation.

nonmonotonic conclusions:

If an abandoned infant acquires citizenship via subsection 1.2 and subsequently its parents are identified, neither parent being determined to be citizens or settled . . .

judgment:

If in the special circumstances of a particular case the Secretary of State thinks fit, he may treat (2) as if the reference to twelve months were a reference to six years.

A quote from the report illustrates the representational difficulties posed by these examples, in particular, doubly negated conditions.[8] How did the group cope with this thorny representational problem? The group finessed special, domain-specific negation predicates where the legislation explicitly specifies all the cases for which the predicate was presumed to hold. Producing these was tedious. I quote:

> In principle, we could continue with this kind of analysis, reasoning through the provisions of the act to construct explicit definitions for the predicates: "x was not a British citizen at the time of y's birth," and "x was not settled in the U.K. at the time of y's birth" .
>
> In practice, however, we could not construct such definitions; the act is too large, and there are too many separate possibilities to consider for this solution to be practical.

In summary, the representational fixes to Horn-clause logic investigated by the group were less than satisfactory. Some difficulties remained.

Other well-known examples of the difficulty-ease trade-off exist. Connectionist architectures make extensional programming easy—for example, providing supervised training on examples of an objective function—and make intensional programming difficult—for example, algorithm writing. The reverse is true for most symbolic approaches. Unlike traditional symbolic approaches, connectionist architectures enable gracefully degrading performance but lose the full measure of semantic transparency and the ability to debug that symbolic approaches retain. However, symbolic approaches ease the representation of exceptional cases. In contrast, because nets generalize, they offer the worst performance for infrequent cases. Genetic architectures offer an interesting blend of facilities by taking evolving populations of rules as an inspiring vision of computation (Holland 1975, 1986). Extensional programming and some modest intensional programming are trivial for genetic algorithms. Evolving rule populations exhibit graceful performance degradation, and at any point, one can stop the program and straightforwardly interpret them.

The Third Point: The Card Trick

The third point is that representational media establish two characters of experience, one for onlookers and one for workers in the medium. Let's illustrate this point first with a non-AI example.

Observing a clever and well-performed card trick, one experiences wonder and then perhaps curiosity about the technique, then maybe wonder again, and so on. The card trick seems like magic. In contrast, the performer of a card trick experiences no wonder, no magic, only technique. In this card trick case, does one experience magic or merely sleight of hand? The answer is, "It depends." Of both experiences, it is not the case that one is true and the other false. The experiences differ with respect to the emphasis they place on aspects of the phenomena at hand and derive from the previous experience and current role of the observer.

The experience of AI branches similarly, only more so. Observing the results of computations that mimic intellect and knowing the utterly syntactic pebble-pushing mechanisms at work engender both reactions of wonder and dismissal together and inconsistently. Consider Hitech, a recent state-of-the-art chess-playing program that relies on extensive parallel search (Berliner 1987). Reporting on Hitech's winning the Pennsylvania State Chess Championship, Berliner remarks on one of Hitech's particularly good moves in its first game winning against a player rated over 2400, "To me, it is amazing that Hitech is able to manage the technique to win this very difficult-to-win position." (Berliner 1987, p. 102). Berliner knows the simple, fast processes at work in Hitech, but this knowledge increases, rather than diminishes, his wonder. When observing Hitech's play at the tournament and experiencing it as amazing, Berliner is playing the onlooker role in a complex social setting. In another role, back in his laboratory with the proper computational tools, he might well analyze the amazing move into the relevant board conditions and decision criteria at work. In this case, the move would no longer seem amazing, merely the obvious consequence of an intricate but lengthy search. In this other role of worker in a medium, the character of experience for the same move would be highly technical. Both characters of experience are important, and it would be far too simple to say that only the technical internal experience is proper and that Berliner should be discouraged from future spells of amazement at chess tournaments.

Although different in feel, the onlooker and worker in the medium characters of experience are similar in the ways they arise. Each results from a system image (Norman 1986) that mediates character of experience. At a chess tournament, an official types an opponent's move, let's say, "Nc6," into a terminal. Some seconds or minutes later, Hitech displays its counter move, "Bb5." When all that appears to happen in the production of senior master level chess moves is this little exchange, no wonder it seems amazing. Witnessing this little exchange and understanding chess, one perceives the glory but not the guts. However, back in the laboratory, other system images of Hitech exist, including, I presume, tools for writing code and logic diagrams for the design of specialty chips. These system images draw one's thoughts in at the level of technical minutia, providing a useful medium for the tasks of assembling hardware and software details. Little wonder or amazement remains at this level, only technique and effort.

Types of Experiencing AI. A declension exists that usefully categorizes such AI experiences. The declension consists of first-, second-, and third-person experiences, analogous to the use of the pronouns I, you, he, she, and it.[9] Builders of AI programs arrange for their programs to exhibit system images that cause interaction to be experienced in primarily one person. First-person experiences of AI programs are those where I, as a user, feel like I am doing the intellectual work—exploring hypotheses, making inferences, drawing conclusions, and selecting alternatives—rather than instructing the machine to do it at my behest. Examples of systems that foster this intellectual prosthetic experience (Hutchins 1988a) include Rabbit (Williams 1984), graphic extensions to advanced Prolog debuggers such as Coda (Plummer 1987), Transparent Prolog Machine (Eisenstadt and Brayshaw 1986), and Decider (Farrell 1988).

Second-person experiences of AI programs are those where I, as a user, interact with the program as an other, a you to converse with. I ask a question of the computer, it asks me some questions, and then it answers my original question and the cycle repeats. Examples of systems that foster this other agent experience include question-answering systems such as Unix Consultant (Wilensky 1982, 1983; Wilensky, Arens and Chin 1984) or Knowledge and Modality Planner (Appelt 1983).

Third-person experiences of AI programs are those where the system image encourages me to experience it as an artifact of someone's creation. For third-person experiences to occur, the program's system image must mediate authorship. Here, Malone's intelligent electronic message system, Lens, is an example (Malone et al. 1987) or Stefik's (1986) vision of the next knowledge medium.

AI images in general social currency are second-person images, and second-person experiences dominate among the experiences of AI programs. In the extreme, these second-person experiences of programs misrepresent. The program behind the monitor screen is unlike a person in many significant ways. However, if in the extreme they misrepresent, in the norm they serve efficiently. Talking about a program as though it were an other that knew, believed, wanted, and acted is efficient and common. As Daniel Dennett (1987) has been pointing out, this habit, although often denigrated as folk psychology (especially when applied to the human psyche), can also be promoted to theory, as Dennett does with the idea of intentional stance (that is, 'adopting a successful and reliable predictive strategy that accounts for behavior by references to intentional concepts').

Where are these experiences of AI leading? What will experiences of future AI programs be like? At this point, I would like to speculate about future experiences of AI artifacts and tell the odd commingling of experience I suspect will emerge as the norm. Second-person experiences of AI artifacts will remain common because they are common now and efficient. Successful intentional strategies will strongly flavor these second-person experiences. However, extended interaction with these artifacts will also reveal the artifacts' mechanical regularities and stupidities, creating something of a logical conflict between the two interpretations. On the one hand, the programs will seem like yous, full of intentions. On the other hand, they will not seem this way at all. What will result from this experiential dilemma? I suggest that a peaceful and inconsistent blending of the perspectives is a likely result. At points during interaction with a particular program when the system image encourages an I-you feel, and the attribution of intention to the program predicts well, second-person experiences will dominate. At other times, when the program reveals the limited nature of its mechanisms, treating it like a dumb machine will dominate. Users will easily move between perspectives.

I imagine here that competence to experience AI artifacts will evolve to accommodate inconsistencies in the artifacts' system images, without any assumption that the resulting experience need be a consistent whole. I have in mind something akin to the ability to read or watch fiction. With fiction, if I want to experience the story, I suspend disbelief and enter in. When I do, in some nonimmediate way, I still know that it's just fiction, but I experience it as though it were immediate and real. However, I have no trouble jumping between the two perspectives at will, noticing perhaps how the author is clever about character sketching, plot development, and attention to detail or admiring the film director's attention to detail, continuity, and subtle use of lighting before jumping back into the story. I think users of AI artifacts will become similarly practiced at turning on and off an I-you intentional stance just as they are now at suspending and resuming disbelief in their stances toward fictional literature and film.[10]

The Fourth Point: The Mouse Trap

I now pursue the point that all representational media communicate the views and values their users hold and that they do this imperfectly, with specific and regular manners of imperfection.

Artifacts communicate the views and values their authors hold. For example, a mouse trap expresses the value that someone somewhere prefers a crushed mouse to an uncrushed mouse. However, as representations of values, computations are imperfect in specific and regular ways, as are all media. A mouse trap will snap a baby's hand, showing that although the mouse trap was intended to express some homeowner's values involving mice, it actually executes values about arbitrary objects which disturb its pressure-sensing trigger.

Consider again the British Nationality Act expressed as a logic program. The program mediated values regarding the privilege of citizenship expressed in the act itself. However, the way it mediated these values differs markedly from the ways an informed official would likely perform. The program follows a pure bureaucracy of relentless logic. The human official would more likely interpret the act; misremember its details; blend it with personal bias; and, in general, act as though aspects of the situation at hand, such as current case load or mood or closeness to holiday, mattered. To his or her credit, the official would be capable of bringing clever interpretation or appeal to higher authorities to bear on the case.

Furthermore, as mediators of values, computations demonstrate ill-specified conditions. The British Nationality Act program exemplifies an interesting twist with respect to this point. The program succeeds to a

I imagine here that competence to experience AI artifacts will evolve to accommodate inconsistencies in the artifacts' system images, without any assumption that the resulting experience need be a consistent whole.

large degree because the act's conditions, as stated, are performative; that is, the conditions stated are by law the correct conditions. In general, it is not so easy to specify conditions under which a value-laden action should be taken. This difficulty results from either a lack of expressivity in the representational medium or amorphous and open-ended values —that is, their conditions are unexpressible outside a concrete situation, and the boundless particularity of the world consistently proves any putatively final value expression inadequate—or both. This difficulty with the representation of value-laden decision procedures is also why, in the realm of law, case law is a necessary adjunct to statutory law. One never gets the statutes right, not even after much reworking. They are only approximations that require concrete situations to elicit the inadequacies of their conditions as stated. What is true of the rule by statute is true, a fortiori, of rulelike computation but without the buffers of appeal and case law.

One might be tempted to think it possible to fix this imperfection by turning to another model of computation better suited for representing nuances—let's say, for instance, weight networks tuned over tens of thousands of trials to recognize necessary conditions. However, this move only avoids one imperfection while it gains another, which, I hunch, remains the fate of all such moves. Consider this example; it is fictional but technically consistent with the state of the art in net computations, and it is sheer imagination with respect to blood chemistry. However, it makes a point.

Suppose that researchers at a large midwestern university gained permission to use their regional supercomputing facilities to feed 14,000 exhaustive blood chemistry workups from patients diagnosed with AIDS related complex (ARC) into the largest weighted network-recognition device ever attempted. Using information about which ARC patients progressed to the AIDS condition, they trained this network device to recognize with 91 percent accuracy the statistical properties of ARC blood chemistry that predict the onset of AIDS. Beyond some obvious regularities already known to medical science, the higher-order nonlinearities computed by this huge device defy interpretation, given the current state of blood chemistry knowledge. This situation would supply us with a trustworthy prediction procedure for deciding when ARC patients will likely become AIDS patients. However, we would not know why. With this type of value-laden computation, we would trade ill-specified, interpretable conditions for effective but uninterpretable ones, or one set of problems for another. It seems likely that computationally expressed values will continue to exhibit regular imperfections dictated by the model of computation used.

The Fifth Point: The Clock

"So long as we represent technology as an instrument, we remain fast in the will to master it. We press on past the essence of technology" (Heidegger 1955, p. 5).

All representational media influence their users' thoughts, and because representational media are usually shared, these influences are typically social. As community artifacts, AI programs recast what a co-representing community can be and do. By co-representing community, I mean any social group that creates, exchanges, and discusses the stature and utility of a common set of ideas mediated by a common set of representations. Such co-representing communities include, for example, physicians, choreographers, petroleum geologists, religious sects, telephone company repair persons, bookies, and serious baseball card collectors—any group in a process of agreeing how to consider the world by refining and exchanging shared representations. What is of interest to us is how co-representational technology at hand influences what the process will be like.

Clocks recast what a synchronized community could be and do. Printing recast what an interreading community could be and do. AI recasts what a co-representing community can be and do. Clocks allow the explicit co-measuring of time. This ability allows physically separated persons to coordinate activities in time with greater precision. Printing engenders the fast and wide spread of common texts among many physically separated persons, affording intellectual forums of previously impossible magnitudes. Knowledge representation technologies allow explicit co-representing of partial world views, and thus afford physically separated persons the ability to coordinate intellectual activity with greater precision. For example, since June of 1987, when the American Medical Association made DXplain available by way of personal computer and modem, a rural doctor in Parrish, Alabama, or anywhere can input any of 4700 medical terms and see a list of possible diagnoses consistent with the terms, based on recent medical updates. The National Cancer Institute's Physician Data Query, the National Library of Medicine's AI/Rheum, the hospital systems Help and Care are other programs that use techniques found 10 years ago only in AI laboratories. These programs are changing the way medical practitioners share representations of their arts and sciences. However, more is going on here than just coordination of intellectual activity.

Clocks brought more than the ability to measure time precisely. They made time into something divisible (Mumford 1934). Is human knowledge divisible into simple assertions or weight networks the way time is divisible into minutes and seconds? Even though AI has learned that in general knowledge can't be neatly and uniformly represented and that knowledge representation poses many wicked problems, the answer is yes. Time is divisible precisely because we divide it with clocks. In similar fashion, although human knowledge resists neat and uniform divisions into representable chunks, knowledge is divisible if we divide it, even if methodological inadequacies render the results less than hoped for. I find this prospect shocking, ironic, and realistic. Over the next decades, just as under the influence of clocks we changed our experience of time, so also under the influence of new representational media, our views of knowledge will evolve to accommodate these media.

Even while drawing away effort from theory, the new [representational] techniques will inspire new thoughts about what kind of event mentality might be.

New Socioergonomic Forces. Through the seconds and minutes they define, clocks exert new types of force in our social worlds; for example, workers punch in and out at time clocks, and competitors win or lose big in races timed to hundredths of seconds. Interreading communities enabled by printing technologies exhibit historically novel forces of editorial control. So, too, the computable forms of co-represented knowledge introduce new socioergonomic forms of force. By socioergonomic force, I refer to the means by which representational tools guide their users into particular patterns of use determined consciously or unconsciously by the tool designers, that is, how the new tools afford particular usage. Thus, AI techniques coalesce in new ways the power to officially categorize for the purposes of representation, to "carve the world at its joints," so to speak. Power to cast the world in which particular values thrive or die is, in and of itself, nothing new. For example, should we categorize homosexuality as an illness? Through its diagnostic standards manual (DSM), the American Psychiatric Association did for the majority of its tenure as an association. However, AI representational techniques introduce new twists, especially in the case of large, shared, persistent knowledge bases. Here are five new forms of force unleashed by AI co-representational techniques.

The first force is *stare decisis*, the legal term for "standing in the decision," the principle that lends weight to precedents in legal decisions. With respect to AI techniques, I mean the inescapable requirement of making a decision about primitive categories and sticking with the decision. In connectionist architectures, this stare decisis means deciding on the scheme for encoding training examples on input and output nodes, tying in the net to its world. In symbolic architectures, stare decisis means establishing categories, usually hierarchically, to develop an object world over which inferences can be made.

The stare decisis force appears in earnest when a categorization established for one purpose conflicts with a categorization established for another. My own most memorable example involved the difference between axiomatizing the act of creating (or moving) an icon as creation or movement given a palette of icons. Creating implies what is created didn't exist before. In this case, the palette icons refer to creatable iconic types. However, movement implies what's moved existed at the time of moving. In this case, the palette icons could be picked up and moved to another area of the display. The difference seems trivial, but if two programmers write two sets of inference procedures consulting the same model of the world, and one regards icon creation as creation and another regards it as movement, the stare decisis form of force becomes evident. You make a decision and stick with it. However, the merit of a representation is always relative to its use, and use evolves over time; so, unique co-representing difficulties continue to emerge.

I call the second force the "rock in the lawn" because it describes the way that shared knowledge bases acquire inertia. Here's the analogy: I find a rock in my lawn and decide to dig it up. The next morning, I start digging around it only to find the rock I originally observed is the tip of a larger rock. I keep digging. By early afternoon, I realize the rock is so big that I can't remove it even if I can uncover it; so, I give up and fill the lawn back in. Modifying a shared knowledge base can be a similar situation. It is difficult to estimate the work a proposed change entails. Not surprisingly, what seems to be a small reconceptualization can cascade into a myriad of changes necessary for carrying through the original vision. Often, some of the cascaded changes are undesirable. Rather than make all the necessary changes, especially when other people are relying on the current structure, it is easier to live with the status quo. In this way, large structures develop an inertia of their own.[11]

Third, implementing policies of privileged control and enforced coherence by locking out particular classes of modifications and performing consistency checks on proposed transactions are simple matters—you give power to those with privilege. However, power and privilege are not new. What is new here is the clarity and finality with which co-representational privileges can be extended and withheld with respect to considering the world in concert with others. Here lies an opportunity to give the concept of imprimatur a whole new meaning, rich with force.

Fourth, the phrase *terminological shift* refers to the process of buying into an established categorization of some part of the world. This process amounts to coming to believe an existing knowledge base. For example, if high in a hierarchically structured knowledge base an intangible object category exists, when I use this category term, it begins to take on the meaning of these objects represented as intangible objects. Over time, the process completes. Daily use of the categorization breeds familiarity, which shapes thinking. Once, for me, intangibility was, well, sort of intangible. Now I know exactly what intangible means—referents of those structures marked in the knowledge base as intangible. My use of terms and, therefore, my ground for thought have shifted to accommodate frequent interaction with the categorization.

The fifth force modifies what co-representation is and how it works. Most knowledge representation regimes shift satisfaction criteria for successful co-representing toward greater explicitness. You and I can have a meaningful professional conversation without being exceptionally explicit about the meanings of the terms we use. Our shared professional background permits us to use suitably vague but putatively commensurate terms. Vagueness is, thus, an asset in conversation; we don't have to spend time making everything explicit. We need not achieve consensus on terms because commensurate interpretations suffice. However, sharing knowledge bases does not afford the luxury of efficient vagueness. Representation regimes enforce a "tyranny of the explicit" as the regimes foster agreement and disagreement at levels of detail not previously attempted. Such greater explicitness often requires more representational decisions than one would care to make and means that additional ways exist in which conceptions can be seen as incompatible. This "tyranny" evolves into a force toward greater standardization as different parties adjust to a greater level of explicit agreement or rivalry while they continue their disagreements, which are fueled with details forced into joint visibility by the nature of the representational medium.

Conclusion: AI Literacy

If computer literacy exists, then there is probably such a thing as AI literacy. On the prospect of being literate about AI, the representational medium viewpoint has, I think, something to add. Considered in representational media terms, AI literacy might well include two fundamental points: (1) At one level, AI seeks new expressive power through new computation-based expressive means. This seeking of expressive power fastens it to an overarching concern with modifying representational facility—making some representational tasks easier and others harder. AI efforts cannot escape from this core. (2) As Margaret Boden observed (Bobrow and Hayes 1987, p. 41), as representations, computations are commitments to particular ways of thinking about the world and, thus, are challengeable with respect to the distinctions they make, their decision criteria, and the values they embody. This observation is particularly relevant to intelligent artifacts, which by virtue of their symbolic complexity and depth don't wear their presuppositions on their sleeves.

On the first point, it is just as hard to say what new representational techniques future models of computation will inspire as it is easy to critique the limits of existing techniques. Whatever the new representational techniques might be, it seems certain we will test their potential for fashioning likenesses of our own mentality. I have suggested that the representational technologies under test are likely to draw attention to themselves—attracting a large portion of effort into mucking about with the way in which the techniques facilitate particular representational tasks rather than into developing theory about the mind. I claim this fact is so because extensible representational facility is the central technical advance that models of computation enable. Even while drawing away effort from theory, the new techniques will inspire new thoughts about what kind of event mentality might be.

On the second point, it is currently quite difficult to challenge program presuppositions. The only avenues of approach are to get the source code and analyze it in depth or write another program. However, given the representational potential inherent in computation, this case need not be true. AI programs should appear more like what-if artifacts; their presuppositions should be apparent and modifiable. The nature of declarative representation invites making these presuppositions apparent and modifiable and we will see what comes of the invitation. David Owen, at the University of California at San Diego Cognitive Science Institute, has applied the phrase center of gravity to human computer interfaces to communicate the observation that interfaces gently coerce users into a particular level of analysis and action for a given activity which in fact has many levels. Although this concept of center of gravity is imprecise, I believe it is important that future representational media offered by the AI research community draw their users into a superior position, a high center of gravity, where it is expected and easy to recast so-called facts. This suggestion asks a good deal, but empowering users by such means retards the prospect of using the representational means of computation coercively.

Acknowledgments

I thank Ben Broderson for the opportunity to address the University of Kentucky's Symposium on the Human Dimension in Artificial Intelligence with these ideas. I also thank Jim Hollan, Dave Wroblewski, Tim McCandless, Loren Terveen, Jim Miller, Elaine Rich, Ed Hutchins, Don Norman, Peter Polson, Bob Joyce, and Rich Cohen for formative discussions concerning the ideas presented here and for assorted comments on various drafts and the Imperial College Logic Programming

Group at the University of London for its excellent report. I also want to thank Michael Arbib and Bob Sokolowski for encouraging me at just the right times.

References

Agre, P., and Chapman, D. 1987. From Reaction to Participation. Paper presented at DARPA Planning Workshop, Santa Cruz, Calif.

Anderson, J. R., and Thompson, R. 1987. Use of Analogy in a Production Architecture. Unpublished manuscript.

Appelt, D. 1983. Planning English Referring Expressions, Technical Report, 312, Computer Science and Technology Div., Artificial Intelligence Center., SRI International.

Berliner, H. 1987. HITECH Wins Chess Tourney. *AI Magazine* 8(4): 101–102.

Bobrow, D. G., and Hayes, P. J. 1987. Artificial Intelligence: Where Are We? *Abacus* 4(3): 8–41.

Bovair, S., Kieras, D. E., and Polson, P. G. 1988. The Acquisition and Performance of Text Editing Skill: A Production System Analysis, Technical Report, 28, Univ. of Michigan.

Buhler, C. 1960. *The Fifteenth Century Book*. Philadelphia: University of Pennsylvania Press.

Calvino, I. 1979. *If on a Winter's Night a Traveler.* New York: Harcourt Brace Jovanovich.

Card, S. K., Moran, T. P., and Newell, A. 1983. *The Psychology of Human-Computer Interaction*. Hillsdale, N.J.: Lawrence Erlbaum.

Chapman, D. 1987. Planning for Conjunctive Goals. *Artificial Intelligence* 32(3): 333–369.

Chomsky, N. 1956. Three Models for the Description of Language. *Information and Control* 2(2): 137–167.

Church, A. 1941. The Calculi of Lambda-Conversion. In *Annals of Mathematics Studies*, vol. 6., 1. Princeton, N.J.: Princeton University Press.

Church, A. 1936. An Unsolvable Problem of Elementary Number Theory. *American Journal of Mathematics* 58: 345–363.

Cook, S. A. 1973. A Hierarchy for Nondeterministic Time Complexity. *Journal of Computer and System Sciences* 7(4): 343–353.

Cook, S. A. 1971. The Complexity of Theorem Proving Procedures. In Proceedings of the Third Annual Association for Computing Machinery Symposium on the Theory of Computing, 151–158. New York: Association for Computing Machinery.

Curry, H. B. 1958. *Combinatory Logic*. Amsterdam: North Holland.

Dennett, D. C. 1987. *The Intentional Stance.* Cambridge, Mass.: MIT Press.

Dreyfus, H., and Dreyfus, S. 1986. *Mind over Machine.* New York: The Free Press.

Eisenstadt, M., and Brayshaw, M. 1986. The Transparent Prolog Machine (TPM), Technical Report, 21, Human Cognition Research Lab, The Open Univ.

Farrell, R. 1988. Facilitating Self-Education by Questioning Assumptive Reasoning. Paper presented at the Seventh National Conference on Artificial Intelligence, 22–26 August, St. Paul, Minn.

Fredkin, E., and Toffoli, T. 1982. Conservative Logic. *International Journal of Theoretical Physics* 21: 219–253.

Garey, M. R., and Johnson, D. S. 1979. *Computers and Intractability.* New York: Freeman.

Hacking, I. 1983. *Representing and Intervening.* Cambridge University Press.

Heidegger, M. 1955. *The Question Concerning Technology and Other Essays.* New York: Harper & Row.

Holland, J. 1986. Escaping Brittleness: The Possibilities of General Purpose Learning Algorithms Applied to Parallel Rule-Based Systems. In *Machine Learning: An Artificial Intelligence Approach,* vol. 2, eds. R. Michalski, J. Carbonell, and T. M. Mitchell, 593–623. Los Altos, Calif.: Morgan Kaufmann.

Holland, J. H. 1975. *Adaptation in Natural and Artificial Systems.* Ann Arbor: University of Michigan Press.

Holland, J. H., Rich, E., Hill, W. C., Wroblewski, D., Wilner, W., Wittenburg, K., Grudin, J., and members of the MCC Human Interface Laboratory. 1988. An Introduction to HITS: Human Interface Tool Suite, Technical Report, ACA-HI-406-88, Microelectronics and Computer Technology Corporation.

Hopfield, J. J., and Tank, D. 1985. "Neural" Computations of Decisions in Optimization Problems. *Biological Cybernetics* 52: 141–152.

Hutchins, E. L. 1988a. Metaphors for Interface Design. In *The Structure of Multimodal Dialogue,* eds. M. M. Taylor, F. Neel, and D. G. Bouwhuis. Forthcoming.

Hutchins, E. L. 1988b. The Technology of Team Navigation, Technical Report, 8804, Inst. of Cognitive Science, Univ. of California at San Diego.

Kleene, S. C. 1956. Representation of Events in Nerve Nets and Finite Automata. In *Automata Studies,* eds. C. Shannon and J. McCarthy, 3–41. Princeton, N.J.: Princeton University Press.

Kleene, S. C. 1936. General Recursive Functions of Natural Numbers. *Mathematische Annalen* 112: 727–742.

Kowalski, R. A. 1974. Predicate Logic as a Programming Language. *IFIP* 74: 569–574.

Laird, J. E. 1986. SOAR User's Manual, Technical Report, Xerox Palo Alto Research Center.

Laurel, B. 1986. Interface as Mimesis. In *User Centered System Design,* eds. D. A. Norman and S. W. Draper, 68–85. Hillsdale, N.J.: Lawrence Erlbaum.

Lenat, D., and Brown, J. S. 1984. Why AM and Eurisko Appear to Work. *Artificial Intelligence* 23(3): 269–294.

Malone, T. W., Grant, K. R., Turbak, F. A., Brobst, S. A., and Cory, M. D. 1987. Intelligent Information-Sharing Systems. *Communications of the ACM* 30(5): 390–402.

Margolus, N. 1984. Physics-Like Models of Computation. *Physica* 10(D): 81–95.

Mealey, G. H. 1955. A Method for Synthesizing Sequential Circuits. *Bell System Technical Journal* 34(5): 1045–1079.

Minsky, M., and Papert, S. 1969. *Perceptrons: An Introduction to Computational Geometry.* Cambridge, Mass.: MIT Press.

Moore, E. F. 1956. Gedanken-Experiments on Sequential Machines. In *Automata Studies*, eds. C. Shannon and J. McCarthy, 129–153. Princeton, N.J.: Princeton University Press.

Mumford, L. 1934. *Technics and Civilization*. New York: Harcourt, Brace and World.

Norman, D. A. 1988. *The Psychology of Everyday Things.* New York: Basic Books.

Norman, D. A. 1986. Cognitive Engineering. In *User Center System Design,* eds. D. A. Norman and S. W. Draper, 31–61. Hillsdale, N.J.: Lawrence Erlbaum.

Norman, D. A., and Hutchins, E. L. 1988. Computation via Direct Manipulation, Technical Report N00014-85-C-0133, Institute of Cognitive Science, University of California at San Diego.

Papert, S. 1988. One AI or Many? *Daedalus* 117(1): 1–15.

Plummer, D. 1987. CODA: An Extended Debugger for Prolog, Technical Report, Univ. of Texas at Austin.

Post, E. 1943. Formal Reductions of the General Combinatorial Decision Problem. *American Journal of Mathematics* 65: 197–215.

Post, E. 1936. Finite Combinatory Process-

Formulation. *Journal of Symbolic Logic* 1: 103–105.

Ritchie, G. D., and Hanna, F. K. 1984. AM: A Case Study in AI Methodology. *Artificial Intelligence* 23(3): 249–268.

Rumelhart, D. E., McClelland, J. L., and the PDP Research Group. 1986. *Parallel Distributed Processing.* Cambridge, Mass.: MIT Press.

Sauers, R. 1982. GRAPES User's Manual, Technical Report, Carnegie Mellon Univ.

Searle, J. 1980. Minds, Brains, and Programs. *Behavior and Brain Sciences* 3: 417–424.

Sergot, M. J., Sadri, F., Kowalski, R. A., Kriwaczek, F., Hammond, P., and Cory, H. T. 1986. The British Nationality Act as a Logic Program. *Communications of the ACM* 29(5): 370–386.

Simon, H. A., 1969. The Architecture of Complexity. In Proceedings of the American Philosophical Society, vol. 106, 467–499.

Smullyan, R. 1961. Theory of Formal Systems. In *Annals of Mathematics Studies,* vol. 47. Princeton, N.J.: Princeton University Press.

Stefik, M. 1986. The Next Knowledge Medium. *AI Magazine* 7(1): 34–46.

Turing, A. M. 1950. Computing Machinery and Intelligence. *Mind* 59 (new series) 236: 433–460.

Turing, A. M. 1936. On Computable Numbers, With an Application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society* 2(42): 230–265.

von Neumann, J. 1956. Probabilistic Logics and the Synthesis of Reliable Organisms from Unreliable Components. In *Automata Studies,* 43–98. Princeton, N.J.: Princeton University Press.

Wilensky, R. 1983. *Planning and Understanding: A Computational Approach to Human Reasoning.* Reading, Mass.: Addison-Wesley.

Wilensky, R. 1982. Talking to Unix in English: An Overview of UC. In Proceedings of the Second National Conference on Artificial Intelligence, 103–106. Menlo Park, Calif.: American Association for Artificial Intelligence.

Wilensky, R., Arens, Y., and Chin, D. 1984. Talking to Unix in English: An Overview of UC. *Communications of the ACM* 27:574–593.

Williams, M. D. 1984. What Makes RABBIT Run? *International Journal of Man-Machine Studies* 21: 333–352.

Notes

1. By model of computation, I mean specifications of formal calculi one can at least partially realize with finite-resource algorithms. Older examples are the Turing (1936) machine, the Post (1936, 1943) production system, Church's (1936, 1941) Lambda calculus, Kleene's (1936) general recursive functions, Curry's (1958) combinatory logic, Mealey's (1955) machine, Kleene's (1956) analysis of McCulloch and Pitts threshold neurons, von Neumann's (1956) cellular automata, Chomsky's (1956) context-free language type, and Moore's (1956) machine. Recent examples include Smullyan's (1961) formal systems, Holland's (1975) genetic algorithms, Procedural Horn-clause logic (Kowalski 1974), various hidden node connectionist architectures (Rumelhart, McClelland, and the PDP Research Group 1986), Fredkin and Toffoli's (1982) conservative logic, Margolus's (1984) ballistic model of computation, Chapman's (1987) planning paradigm, and Agre and Chapman's (1987) improvisation notion. Each model focuses attention on new facets of computation. No model has an a priori claim to preeminence. Each is more or less useful. Hence, fashion among these computational models matters.

2. Associated with each recording device are stylized usage conventions, arts of composing expressions in accordance with the conventions, and common ways of selectively attending to records. Such conventional usages are integral aspects of media (without associated usage conventions and rhetorical art, a recording device isn't a medium) and, thus, contribute to categorization as media.

3. Proposing a representational medium analysis of computation is not original. Notably, Edwin Hutchins (1988b), Jim Holland et al. (1988), and Norman and Hutchins (1988) point to the utility of this perspective and clarify the vocabulary of concepts with which such analyses can be fruitfully conducted.

4. This is not to say that these novel representational properties are used well or even at all. Right now designers tend to copy the prevalent content of earlier media in computation (for example, the desktop metaphor, rooms, and static icons). This content recapitulation is not exclusive to computation. As many commentators have pointed out, most early uses of new media exhibit this trend.

5. This possibility is the point of the introductory illustration of the successor axiom using the elk variable juxtaposed with the Cro-Magnon painting from La Mairie Cave. It is common to presume formal representations (such as the successor axiom) constitute a representational advance of such magnitude as to disdain categorizing them with more naive or primitive representations (such as the cave painting). Looking at AI as only a recent representational advance in a long history suggests the two might have more in common than a really modern (but provincial) outlook would perceive.

6. At this point, the bells and whistles should start chiming and hooting, and the fireworks should fire off. To really ease the activity of deliberately experimenting with representational facility, now that's something new and exciting in the world.

7. This effort was an experiment undertaken by a world-class research group. It did not propose or advocate final solutions to representational inadequacies in Prolog. I have chosen the work of this group as an example primarily because of the clarity and depth with which they reported this piece of their research.

8. For further detail, see Sergot et al. (1986).

9. Brenda Laurel (1986) introduced the human-computer interaction research community to the concepts of personness and first-person experiences of interactive computation.

10. For an interesting example of how an author plays explicity with "the fictional stance," read Italo Calvino's *If on a Winter's Night a Traveler* (Calvino 1979).

11. Imagine what it might mean to be stuck with a QWERTY knowledge base in the way we are now stuck with the QWERTY keyboard.

As a scientist in the Microelectronics and Computer Technology Corporation Human Interface Laboratory, 3500 West Balcones Center Drive, Austin, TX 78759, William Hill conducts research on mediating collaboration computationally using knowledge-intensive interface techniques. He is part of the team evolving Human Interface Tool Suite, a prototype tool environment for designing collaborative interfaces to high-functionality environments by incorporating visual programming, restricted natural language, gesture recognition, and task reasoning in desk-sized, multimodal worksurfaces. He received his Ph.D. from Northwestern University for his research on advice seeking, giving, and following of graphic interfaces.