ARTIFICIAL PERSONS

Anthony O. Putman

ABSTRACT

An alternative approach to the enterprise of artificial intelligence (AI) is presented. The paper divides naturally into four semi-autonomous sections. The first attempts to delineate the subject matter of AI; it offers specifications for what would qualify as an "artificial person". The second section explores some of the statuses within our work communities that an artificial person might appropriately fill. The third section, "An Epistemology for Artificial Persons", takes up issues of real-world knowledge and logic; building on Ossorio's foundations it suggests a fundamental logical form that is intended in the future to form the basis for a general-purpose AI language. The last section, "Some Algorithms of Common Sense", offers some technically detailed means of handling real-world inference.

Descriptive Psychology since its inception in the early 1960s has intended to make substantive contributions to the enterprise popularly known as "artificial intelligence", or "AI" to its familiars. Peter Ossorio from the first insisted on the distinction between "person"—one whose

Advances in Descriptive Psychology, Volume 5, pages 81-103.
Editors: Anthony O. Putman and Keith E. Davis.

ISBN: 0-9625661-0-1.

history is paradigmatically that of deliberate action—and "human" (Ossorio, 1966), which he later formalized in the concept of "embodiment" (Ossorio, 1980/ 1982). His early work for the Air Force (Ossorio, 1964) led to computer-based "intelligence" that is still unsurpassed in the field; his seminal book *"What Actually Happens"* (Ossorio, 1971/1978) was written with at least one eye constantly on issues of computer implementation. A number of theses and dissertations by Ossorio's students dealt with issues of simulating human judgement (e.g., Mitchell, 1967; Putman, 1969), representing complex knowledge in computer-implementable form (e.g., Damon Tempey's work, published as part of Ossorio, 1971) or rigorously formalizing complex processes so that computer programs might use them (e.g., Busch, 1974; Jeffrey & Putman, 1983). The last ten years has seen a steady progression of concepts (Jeffrey, 1981; Putman and Jeffrey, 1985), formalisms (Putman, 1982) and functioning artificially intelligent programs. Looking at the historical record, one can see that, had all this activity taken place within the context of an academic computing science department, the Descriptive Psychology approach might be widely acknowledged as among the three or four primary schools of AI in the country.

This is not the case, of course; the Descriptive Psychology approach and its many contributions remain virtually unknown outside the Descriptive Psychology community. One reason for this state of affairs is the lack of a thorough-going explication of what the enterprise of AI looks like from the Descriptive Psychology viewpoint; lacking that, newcomers to this viewpoint have little basis for assessing the actual and potential contributions Descriptive Psychology can make. The current paper intends to be a step toward filling this deficit. This paper is not an overview or review of previously published or accomplished work; it presents essentially new material which builds on and links to the previously published works. It intends to provide a framework within which both past and future Descriptive Psychology work in AI can be seen and understood for what it is.

This paper addresses some fundamental questions: "What is AI? Why should we pursue it? How can it be done?" It divides naturally into four semi-autonomous sections. The first attempts to delineate the subject matter of AI; it offers specifications for what would qualify as an "artificial person". The second section explores some of the statuses within our work communities that an artificial person might appropriately fill. The third section, "An Epistemology for Artificial Persons", takes up issues of real-world knowledge and logic; building on Ossorio's foundations it suggests a fundamental logical form that is intended in the future to form the basis for a general-purpose AI language. The last section, "Some Algorithms of Common Sense", offers some technically detailed means of handling real-world inference.

SPECIFICATIONS FOR AN ARTIFICIAL PERSON

The link between the enterprise of "artificial intelligence" and the concept of a "person" is an interesting one. It seems clear that, since its inception, AI has typically defined its domain by various kinds of reference to persons and certain of their abilities. Exactly which abilities to include in the purview (and why exactly these and not some others) has been substantially less clear. As a result, the working AI community's "definition" of AI has traditionally been a moving target (Kurzweil, 1985), which today consists essentially of a relatively short list of discrete areas of research interest.

I would like to propose a somewhat fresh look at this link, and thereby at the question "What comprises the domain of AI?" The somewhat fresh aspect stems from the perspective brought to bear on the issue. My professional training is as a psychologist and mathematician. I have spent over twenty years practicing my craft as a Descriptive Psychologist (see Ossorio, 1971, Ossorio, 1971/1978; Davis, 1981; Davis and Bergner, 1983; Davis and Mitchell, 1982, 1985). I have been a professional computer programmer in various stints since 1967, and have been working for the past ten years primarily as an architect of languages and applications in AI (Jeffrey & Putman, 1983; Putman & Jeffrey, 1985). The perspective of Descriptive Psychology is the foundation of the remarks in this paper.

Consider the following assertion: the enterprise of Artificial Intelligence consists of attempting to duplicate, within the hardware/software configuration of a computer system, *all* of the characteristics of a human person, excepting solely those characteristics inextricably dependent on flesh-and-blood embodiment. To be less precise and more clear: AI is the enterprise of creating an artificial person. (By "artificial" I do not mean to imply "not real"; the term is used in its original and primary sense, "produced by human art".)

A few preliminary comments on this assertion: like any intellectual enterprise, AI at any given time will focus its attention on certain of these characteristics and ignore the rest—some questions are more interesting than others. Certain, perhaps many, person characteristics will at any given time be seen as trivial or irrelevant to "real AI"; further, we can predict with certainty that the class of "interesting" or "relevant" issues will continue to change with time. And deciding which characteristics of a person are "inextricably dependent on flesh-and-blood embodiment" will certainly generate some amount of controversy. This is all as it should be. Nonetheless, it seems useful to define the domain of AI in some way other than as a discrete list of current topics; after all, do we know of any other field in which, as soon as we know how to do something, it ceases to be part of the domain?

The primary purpose of specifying the domain of AI is to suggest a "scope of effort" for the field, and thereby suggest standards of adequacy for our theories and artifacts. To fulfill this purpose requires a substantially more detailed elaboration of the above "definition". The remainder of this section is devoted to a first step in that elaboration, which is meant to serve, as the title asserts, as specifications for an artificial person. To avoid overburdening an initial effort, most of the specifications are offered with little further elaboration; I have chosen simply to list the specifications and bunch elaborations together in a numbered schema at the end. Thus, for example, 1.1 is the first elaboration of the specification numbered 1, and so on.

Certain of the terms used in these specifications form a part of the technical vocabulary of Descriptive Psychology; that is, they are conceptually articulated as part of that discipline's body of work. In most cases, the reader will not be misled by assuming the ordinary English language understanding of a term. For the purposes of technical implementation, however, we require a considerably more precise articulation of these terms and their relations to each other, which Descriptive Psychology provides; that precise articulation has formed the basis for numerous working AI artifacts, including several expert systems (of size equivalent to 2700 production rules) and several systems (e.g., thyroid diagnosis, content-based document retrieval) based on judgement simulation which perform very favorably compared to other, related work. Interested readers are referred to, in particular, Ossorio (1971/1978), Ossorio (1980/1981), Putman & Jeffrey (1985), and Jeffrey & Putman (1983).

Specifications

An artificial person:

1. Has places (statuses) in communities.
2. Has motivations; can and will engage in purposive action to accomplish stated and/or desirable ends. These motivations are consonant with status and community.
3. Chooses from among available courses of action in accord with choice principles of its community.
4. Knows:
   a. its own status.
   b. what state of affairs it is attempting to bring about.

c. why it is attempting to bring about that state of affairs.
d. the significance of the behavior it is currently engaging in.
e. the social practice structure of the behavior it is engaging in.
f. what would count as success at its current behavior.
g. the important ways in which this behavior can go wrong, and what to do about it.
h. what its behavioral options are and what it needs to know to choose among them.
i. the current state of affairs relevant to its current actions.
j. the status of anyone with whom it interacts.
k. how to deal appropriately with someone in that status.
l. what actions it has taken in the past, what practices it has engaged in and with whom, and the results of these.
m. what incomplete courses of action, if any, remain from previous interactions.
n. what objects, processes, events and states of affairs are known to this community.

5. Knows how to:
   a. engage in all performances for which it is eligible.
   b. find any information it needs in order to act.
   c. recognize particular instances of the objects, processes, events and states of affairs known to this community.
   d. appraise behavior—including its own—in the light of community standards, and adjust its own behavior accordingly.
   e. construct and engage in a course of action to accomplish a stated goal.
   f. appraise and adjust its own knowledge.

6. Knows the language, concepts and locutions used in this community. Responds appropriately to verbal and written conversation in the language of the community, including questions, requests, and demands. Communicates in the language of the community.

Elaborations

Terms used in the above specifications that form part of the technical vocabulary of Descriptive Psychology include: status, community, motivation, action, course of action, choice principles, object, process, event, state of affairs, significance, social practice, performance, eligible, appraise, knowledge, know, know how, language, concept, locution.

1.1 A key insight of behavioral science is that all behavior occurs in the context of a specific community. We understand a given behavior by reference to that community and its standards. Every person—human or artificial—is a member of multiple communities and has a specific place (status) in each. A person's status in a given community connects to the permissions and restrictions on action within the community. Thus, to know what to do and how to do it, we must know what the person's status is within which community.

2.1 A person acts only when an opportunity is perceived to bring about a desired or valued state of affairs. But no state of affairs can be taken as per se desirable; it is adjudged desirable in the light of circumstances when viewed from the perspective of a particular status within a given community. Thus, what an artificial person is motivated to do depends on what status it is occupying in which community, and what its circumstances are.

2.1.1 For example, consider the state of affairs described by "Our organization spent five million dollars last year in excess of revenues". The financial VP of a publicly-traded corporation would perceive that state of affairs as highly undesirable, and would go to great lengths to avoid it. The marketing director of a start-up consumer products firm, charged with building brand recognition and market share, might well see this state of affairs as desirable and be motivated to pursue it. Status, community and circumstance combine to yield diametrically opposing motivations.

3.1 At any given time, a person must choose what to do from the various available courses of action. By "available", I mean simply that doing this particular thing would be in some way appropriate to the current situation. How appropriate a given course of action is—and in what way it is appropriate—depends on what Ossorio calls the "choice principles" of the community. (Ossorio, 1981/1983). A course of action can be preferred on at least the following grounds: (a) it best promotes one's self-interest; (b) it maximizes one's pleasure or decreases one's discomfort; (c) it is appropriate in the social or aesthetic sense; (d) it is the ethically right thing to do. Different statuses in different communities appropriately weigh these grounds differently when choosing behavior. An advertising marketing executive in the entertainment industry, for example, is expected to feel somewhat less constrained by the literally provable truth than is, for example, a bank's outside auditor. An artificial person acting in either role will fail if it fails to appreciate the difference such choice principles make.

3.2 Most AI theories and artifacts have a palpable bias in favor of the scientific version of the aesthetic choice standard: Choose the course of action which is logically fitting to the situation. Disputes about how to choose have been largely parochial—which procedures or algorithms best select the most fitting conclusion. This bias mirrors the traditional severely limited view of "intelligence" in the field of psychology (a view which, not coincidentally, has in recent years been substantially and effectively challenged; see especially Gardner, 1983 and Sternberg, 1985). So long as we build artificial persons with only severely limited agency, we can get away with this bias. But as we move increasingly toward building artificial persons whose agency can be compared to a human's, we will have to conform to more stringent standards. Consider this one: an artificial person must be able to choose behavior responsibly and appropriately in situations in which humans would have conflicting grounds for choosing, and must be able to demonstrate understanding of the conflict. The representation and utilization of choice principles seem necessary to adhere to this standard. The long tradition within philosophy and psychology of posing "moral dilemmas" can serve as a beginning reference point for constructing such conflictual test cases.

4d.1 One key aspect of the formal representation of behavior is its hierarchical structure. Regarding any behavior, it is both formally and procedurally sensible to ask either of two questions: "How do you do that?" and "What are you doing by doing that?" The first question leads to procedural elaboration as required; the second leads to elaboration of the broader pattern of which this action is a part. The latter is referred to within Descriptive Psychology as the Significance of the behavior. It is particularly important in understanding the meaning of a given behavior, and constructing means/ends explanations.

4e.1 Individual behaviors do not occur in isolation. They are part of some patterned sequence of behavior, typically involving more than one actor, in which one action can be seen as a response to prior actions in the sequence and as laying the groundwork for subsequent actions. This patterning of actions is referred to by the concept of social practices. The elaboration and description of social practices plays a major role in the creation of artificial persons (see Jeffrey & Putman, 1983; see also Schank & Abelson, 1977 and other AI practitioners whose concept of "scripts" closely parallels this).

4n.1 Any attempt to create an artificial person must quickly come to grips with the need to represent real-world phenomena. Most typically, that requirement has been translated into the need to represent objects of various sorts, as witness the well-known Stanford-derived KEE system. Ossorio (1971/1978) makes a cogent argument to the effect that the representation of real-world phenomena requires formal recourse to four major units, namely: object, process, event, and state of affairs. More importantly, he provided a detailed articulation of representational schema for each of these, along with "transition rules" explicitly stating the formal and conceptual relations among them. These have formed the basis for several AI artifacts, at least one language, and a procedural paradigm for software development. (Jeffrey & Putman, 1983; Ossorio, 1964; Putman, 1982; Putman & Jeffrey, 1985).

5c.1 The distinction drawn in these specifications between "know" and "know how"—knowledge and skill, if you prefer—parallels the AI tasks of representation and recognition. As machine vision practitioners have attested for years, it is one thing to represent an object in all its attributes, and quite another thing to be able to recognize an instance of that object when you see one. The two domains, and tasks, are certainly not independent of each other, but neither are they interchangeable.

5e.1 We have struggled for at least decades to pin down exactly what distinguishes a person from a clever mechanism. Turing's famous test defined the battleground, if you will, but by no means settled the issue; over time we have become more and more clever at catching the machine acting mechanical. I suspect that the crux of this matter for most of us lies in the person's ability to surprise us by acting in ways that are unexpected but still appropriate. Even intimate knowledge of an artificial person's programming should not enable us to predict with certainty what it will do, but mere randomness is just a trick; what we need is unpredictability in the service of effective action. What we can say with certainty is that persons can string together sometimes remarkably long and complex chains of action and interaction to accomplish their desired goals; further, they can make these chains up as they go along. An artificial person needs the same ability, and will surprise us at times by its appropriate use of it.

5f.1 Human knowledge is at best a tentative thing. What we know, and what we make of what we know, changes in the light of our own experience, and we do not require outside intervention to reprogram us. Neither should an artificial person.

6.1 Language has, appropriately I think, occupied a special place in the concerns of the AI community. Until very recently, many would argue that the human linguistic ability was our uniquely distinctive characteristic. As we learn that other species (certain cetaceans and primates, for example) seem to have some linguistic ability, our esteem for language has not gone down; rather, our esteem for those species has gone up. Peter Ossorio has made valuable and distinctive contributions to our knowledge of linguistic behavior; I believe his work in this area is of substantial interest to the AI community. Interested readers are referred especially to Ossorio (1964, 1969/1978) and Mitchell (1981).

JOB DESCRIPTIONS FOR ARTIFICIAL PERSONS

The specifications outlined in the first section of this paper were meant to address substantively the question "What is AI?" In this section we turn our attention to "why" questions, and attempt some answers to them.

What, after all, is the point of AI? Why are so many people spending so much time and effort trying to create artificial persons of various sorts? We in the AI community rarely asks ourselves these questions because the answers are self-evident to us. Like chess, if you play, you don't need to ask why; if you don't play, you wouldn't appreciate the answer anyway. The aesthetic satisfaction of AI work is complete in itself.

But in addition to the aesthetic, the enterprise of AI has its pragmatic concerns. Granted that artificial persons are fascinating and intellectually stimulating—what good are they? Where are they going to make a useful difference in the lives of humans? I submit that the relative lack of conceptual work on these pragmatic questions has to date been AI's main self-limiting weakness. We have not built many really useful artificial persons in part simply because we have had scant guidance regarding what would be useful to build.

One reason for this lack is the name of our enterprise itself. We are out to create artificial *intelligence*, after all, and everyone knows that intelligence per se is a good and valuable thing. You don't ask the owners of a gold mine what they intend to *do* with all that gold; they just mine the gold, and leave it to others to decide how to use it. Accordingly, much of AI's most "pragmatic" work has consisted of creating systems for building "experts", and assuming that others will decide how to use them.

Unfortunately, this point of view reflects the rather radically academic worldview of AI practitioners in general. As corporate personnel officers can attest, intelligence per se does not get many jobs done. What is needed is intelligence finely focused on the pragmatic concerns of specific tasks. We don't need someone who can prove esoteric theorems; we need someone who can keep an eye on accounts receivable and let the right people know when some accounts show trouble signs. In general, we don't need experts, either; what we do need are people—artificial or otherwise—who know how to support us in our enterprises by accomplishing some important, complex but often pretty mundane jobs. This section intends to offer some substantive help with these concerns by delineating several broad categories of artificial persons. Each category is "expert" in the sense of knowing substantially and in detail about some area of endeavor, but that is almost beside the point; the categories are defined not in terms of what or how much they know, but rather in terms of in what specific ways they fit into the existing work communities. In short, they are defined in terms of a brief "job description".

Guide

Guides help people with "the ropes to know and the ropes to skip", as the title of one business self-help book put it. Whenever we move into a field outside our own area of expertise, or into a new organizational environment, we are confronted with our ignorance of the vast body of implicit knowledge about how things work around here. A guide helps us navigate through these unfamiliar waters. The issues a guide can help on can range from fundamental and basic ("What is expected of someone in my position?" "Does one ever bypass the chain of command here, and if so, when and how?" "How is this particular analysis actually done?") to the nitty-gritty detail ("Where is the copier and who gets to use it?" "What am I supposed to do with my copy of Form P88M?" "I just got error message 1022; what does that mean?").

Regardless of the focus, however, good guides have certain common characteristics: (a) They in fact know what they are talking about; (b) They are easily available to us when we need them; (c) They are willing to share their knowledge with no further end in mind other than contributing to our success; (d) They do not make us feel stupid, inadequate or beholden for asking for help. Looking over these characteristics, it is not surprising that good guides are more often read about than encountered. Guide is an obvious and useful role for an artificial person; Joe Jeffrey built one that worked quite well at Bell Labs (Jeffrey & Putman, 1983).

Coach

I can make an excellent souffle if I have someone to talk me through the steps. Without such coaching, my souffles turn out remarkably like scrambled eggs. Most managers can do a good job of performance review with a little coaching; without coaching, they do what they can—and that's usually not very good. In virtually every endeavor, saving only those in which one is both proficient and recently practiced, coaching is needed to ensure good performance. And saving only those activities that require a flesh-and-blood body to demonstrate, artificial persons make excellent coaches. Talking one through the steps involved in doing something, digging into detail on a particularly tricky part, stopping to give "the big picture" or to clarify a term, helping to troubleshoot when one has gone wrong, a good coach gives one the confidence to succeed at endeavors one would otherwise not undertake or else perform poorly.

A coach (a) knows all the ways to succeed at this endeavor, (b) knows how to explain the relevant steps to people ranging from raw rookies to rusty experts, (c) focuses solely on giving you what you need to succeed. Clearly, there is overlap between the roles of guide and coach. The crucial difference lies not in expertise but in the form of interaction. You expect a guide to tell you about something and leave its application up to you; you expect a coach to be actually involved in your doing of the thing, talking you through it and offering tips from the sidelines. Artificial coaches in the areas of people management, marketing, and performance review have been built by the author in recent years.

Friday

Robinson Crusoe had his "Man Friday"; Mike Hammer had his "Girl Friday"; I suspect what most of us want from AI is a "Thing Friday"—an artificial person who can get things done for us. A "Friday" knows how to do all those little things we would rather not bother with ourselves—look things up, keep track of things, do the picky little calculations and make sure all the forms are filled out, remember to notify everybody about the meeting or the decision, make the reservations and remind us to pay the phone bill. The list of things we want a Friday to do is limitless, because a Friday is nothing more or less than a tireless and faithful assistant. It takes care of the scutwork and the details so we can get on with our jobs.

Being embodied in a software/hardware configuration, a Friday is particularly well suited for handling tasks involving data and records. For example, one recent project built a Friday to handle loan documentation for a bank. The Friday interviews loan clerks (via flexible forms and ordinary questions) about a particular loan, and then creates and prints all necessary documents required to secure the bank's claim on the collateral. This is an important task. Doing it properly can save a large bank literally millions of dollars annually, but doing it properly takes both a great deal of specific knowledge and a lot of time. This is an ideal task for a Friday. The loan documentation Friday knows what information is required to document which sorts of loans, so it can ask for just that information without asking for extraneous items. It is expert in the classic "expert system" sense as well, because it knows the law regarding what forms to file and with whom to secure a given type of collateral. And it knows how to draw up, print, and address the needed documents, taking over a massive clerical task and doing it right. The fact that it knows all these things is almost beside the point; the purpose of a Friday is to *do*, not just to know.

Analyst

Sometimes, however, what is really wanted is just an answer to our questions. "What are the Russians up to in the Baltic?" "Are we likely to find oil in this location?" "Which components should be put together for this customer's computer system?" "What disease does this patient have, and what is the best treatment?" "What are our best options for financing this acquisition?" These are a few of the questions addressed by currently operational expert systems. What they have in common is the role assigned to the artificial person—the role of analyst, responsible for thoroughly examining some particular state of affairs and making useful sense of it.

The standards of adequacy for an artificial analyst seem to be the same as for their human counterparts. The analyst is expected to take a thorough, meticulous look at all the relevant facts, to reason and extrapolate from those facts in a rigorous and defensible manner, to present its conclusions along with any important reservations in a clear and easily understood report, and to be ready to present facts and reasoning in support of its conclusions as required. Note that the analyst role explicitly assumes that the decision maker needs or wants no further help once the report is finished. If this is not the case, then some other role, such as Coach or Friday, is probably more appropriate than Analyst in designing the artificial person.

Librarian

Some of the original AI work done within Descriptive Psychology involved the creation of an artificial librarian (Ossorio, 1964). Virtually all enterprises these days generate and utilize more documents than can be kept track of by any human. As Ossorio realized some twenty years ago, what is needed is an artificial person who (a) knows the substance of every document in the library, (b) knows what an individual means by a particular request for information, and (c) can steer individuals to just those documents most relevant to their requests. To date, no one appears to have improved upon Ossorio's analysis or his basic "J-space" methods; indeed, most computerized databases use some combination of keywords and menus, which require substantial adaptation by the user to the conventions of the library, rather than vice-versa.

These "job descriptions" for artificial persons, while encapsulating a great deal of analysis and practical experience, are certainly not meant to be exhaustive. They are intended to help expand our conceptions of the roles artificial persons can play, while offering some useful initial categories to their designers. Part of the fascination of AI lies in watching entirely new roles being conceived and filled.

"What" and "why" having been addressed, let us now turn our attention to the question most dear to the true engineer's heart—"how?" The final two sections of this paper deal with "how" questions, beginning with the fundamental "how" of AI—how do we represent knowledge and knowing?

AN EPISTEMOLOGY FOR ARTIFICIAL PERSONS

"What is knowledge and knowing?" Philosophy 101 students for generations have encountered this question and have learned that answers to it comprise the field of epistemology . . . which hard-nosed engineering students for generations have placed among the three or four most useless fields of human endeavor. Hard-nosed variants of the question, however, come up when one takes seriously the task of creating an artificial person, namely: "How can we represent what there is to know? How can we know that something is actually the case, rather than merely something that is possible?" Perhaps it is stretching things a bit to call these epistemological questions; after all, for the engineer the "how" in those questions carries almost all the weight and interest. Nonetheless, they point to two critical domains within AI to which Descriptive Psychology can make very substantial contributions: knowledge representation, and forms of logic.

Ossorio virtually single-handedly laid the groundwork for the domain of knowledge representation in his book "*What Actually Happens*" (1971/1978), which not coincidentally is subtitled "The representation of real-world phenomena". His conceptual and technical elaborations of the basic reality constructs—object, process, event, state of affairs—form the basis of a complete technology for knowledge representation. (This should not be taken, however, as downplaying the very substantial work required to instantiate Ossorio's basic formats in computer-useful form.) Rather than reiterating that body of work, this paper assumes it as given and builds from it.

The second domain—forms of logic—is so fundamental that it can initially be difficult to see what one is getting at in talking about it. It addresses a concern that among flesh and blood persons is rarely discussed except by academic philosophers, but which is a down-and-dirty practical issue when trying to construct artificial persons. In a nutshell, the issue is: How can we represent facts about the world so that the logic of the relations among the facts—including importantly those facts which form the context for these facts—is also represented? To represent the fact that Roger Clemens pitched a one-hitter last night is one thing; it is quite another thing to represent that fact in such a way that we can also conclude that the Red Sox thereby clinched the pennant because the Yankees lost to the Tigers, and Clemens virtually assured his selection for the Cy Young award. Knowledge representation concerns itself with the adequate representation of facts; forms of logic concerns itself with what and how we can conclude from these facts.

AI in the past decade has begun to probe deeply into questions of logical form. The classic, sturdy "if-then" logical form, often expanded into "if-then-else" for completeness, has been widely used in actual applications. It has done very well for many applications, not so well for others; the limitations and critiques of this form are well known in the field. The "tree" structure instantiated in the PROLOG (PROgramming in LOGic) language was one step into further complexity. The past decade has seen an outpouring of theoretical work (and in some cases laboratory implementation) on such topics as default reasoning (e.g., Reiter, 1980; Glymour & Thomason, 1984), circumscription (e.g., Grosof, 1984), "common-sense" reasoning (e.g., McCarthy, 1986) and various approaches to non-monotonic logic (AAAI, 1984). Indeed, it is not overstating the case to say that virtually every LISP program written in AI labs in the 1980s either formally or informally created its own specialized "form of logic"; this is one of the language's more powerful and, its critics say, dangerous features in use.

The purpose of this section is not to critique these forms of logic. It is to offer an alternative formulation which has certain advantages: (a) It is complete, in that any logical relationship among facts can be represented without distortion within it; (b) It is powerful, in that surprisingly complex deductions and lines of inquiry can be derived from establishing surprisingly few facts; (c) It is intuitively clear, in that it is derived from a fundamental logic used by ordinary people in carrying out their day-to-day affairs.

Descriptive Psychologists will recognize this as simply paradigm case logic (Ossorio, 1980/1981) made technically explicit. Others have made similar points in rather different terms (Doyle's maintenance system and the general thrust of Hayes's ideas (Hayes, 1985) seem like reasonably close cousins); nonetheless, the conceptual connections to the rest of Descriptive Psychology imbedded in this formulation seem reason enough to offer it as, if not wholly "new", at least not outmoded and potentially quite useful. It is meant to guide the artificial person in its understanding of the world in which it finds itself; as such, it should most likely form the basis for an as-yet unspecified programming language for AI.

Fundamental Logical Form for Representing Knowledge
About the Real World:

SA1 unless SA2 in which case SA21 or SA22 or . . . SA2n.
unless SA3 in which case SA31 or . . . SA3m
...
unless SAi in which case SAi1 or . . . SAij.

where SA = state of affairs.

Unpacking that somewhat into English, it says: state of affairs SA1 is taken to be the case unless some other state of affairs (SA2, SA3, . . . SAi), which indicates otherwise, is known to be the case, in which case some other state of affairs (SA21 . . . SA2n) is taken to be the case. SA1 is the paradigm case.

One distinctive feature of this form is that a certain state of affairs (in this case, SA1) is taken to be the case unless we know something which indicates otherwise. This embodies paradigm case logic, and is a very powerful means of representing the "contextual" knowledge which makes real-world action possible. Humans often need to know very little about the specific situation at hand before complex actions and conclusions are warranted; this logical form makes available to artificial persons the same kind of scope and power.

Obviously, the concept of "state of affairs" is central to this logic. Ossorio details at great length the conceptual connections which articulate this concept (Ossorio, 1971/1978); it would be pointless to go over that ground again here. Instead, it seems useful to list some of the major "varieties" of states of affairs, and note that we need to make explicit allowance for representing each variety. What follows is meant to be an heuristic list, certainly not an exhaustive one; readers are explicitly invited to note needed additions.

Major varieties of SA:

1. Person P may/should/must engage in action A.
2. Proposition P is true (false).
3. Event E has occurred.
4. Variable X has value V.
5. A has relation R to B.
6. if SA1 then SA2.
7. SA1 and SA2 and . . . SAn.
8. SA1 or SA2 or . . . SAn.

It was previously mentioned that the above is an elaboration of paradigm case logic. As such, the question immediately arises, "What are the allowable variations on the paradigm case?" What "short forms" should be taken into account, and what more elaborate forms need to be recognized? Again, what follows is not exhaustive.

Allowable variations include:

1. Omit "in which case", leaving: SA1 unless SA2.
2. Omit "unless", leaving: SA1 in which case SA2.
3. Expand any SA with any allowable substitution, including the entire "SA unless . . ." form.
   For example:
   (if SA1 then SA2) unless (SA3 *or* SA4) in which case (SA5 AND SA6).

Relations (between objects, processes, events or states of affairs) are of special interest, precisely because so much of what we need to know about the world takes the form of, "Does this relation hold between these two things?" As SA variety 5 above suggests, to say that "A has relation R to B" is to identify a state of affairs as being the case. The actual relations which are of interest in the real world are seemingly endless; a partial listing follows of some for which we clearly must make explicit allowance.

Major varieties of R (relation) include:

1. is the result of
2. is caused by
3. is part of
4. is identical to
5. greater than, less than, equal to (arithmetic)
6. is a means to
7. can be taken to mean
8. is preferable to
9. is an alternative to
10. is an instance of
11. can be used in place of
12. contains as a constituent
13. precedes
14. follows
15. occurs simultaneously with
16. is compatible with
17. is an attribute of
18. has the attribute
19. is a part of X, as is also

I want to be the first to acknowledge that the above is only the barest sketch of the specifications of an AI language; very substantial work would be required to convert it into an implementable form. But such specification was not the purpose of this section. What I hoped to accomplish was simply to suggest a way of thinking about representing the world—the "fact unless fact in which case fact" logical form—and point to some obvious features of this way of thinking. The worth of this way of thinking can ultimately be tested only in actual implementations, which one hopes may be stimulated by this brief sketch.

SOME ALGORITHMS OF COMMON SENSE

In this last section, let us turn our attention to a very technical "how" issue: inference. Inference has been a core concern of the AI community since its founding. In some intuitively obvious sense, intelligence and inference are inextricably linked; indeed, one of our primary critical standards for a new AI artifact is the extent to which in operation it goes beyond the obvious given facts to deal with derivations.

Once we move beyond the intuitively obvious, however, inference becomes a somewhat difficult topic. We have made good progress with inference based on formal logic—the "All men are mortal. Plato is a man. Therefore, Plato is mortal." type of inference. We have run into difficulty handling "fuzzier" inferences, such as Aristotle's classic syllogism of practical reason: "I need a horse. There are horses for sale in the marketplace. Therefore, I should go down to the marketplace and buy a horse." The kind of everyday, "common sense" inference made by the average six-year-old child has, to date, overtaxed our ability to reproduce it in AI artifacts. The AI community is working on this class of problems and making progress (e.g., Shank & Abelson, 1977; Winograd, 1982); nonetheless, there appears to be substantial room remaining for new ideas and fresh approaches.

Consider the following restatement of the inference dilemma: The facts we need often are different from the facts we have. Inference becomes difficult when the needed facts cannot be formally derived from the facts at hand. How can we "know" what we don't know and cannot deduce?

This "dilemma" is actually a statement of the ordinary situation faced by any person—human or artificial—who is called upon to act in the real world. Not uncommonly, the knowledge we need in order to respond appropriately to current circumstances is not equivalent to the set of facts we have directly observed or established to be the case. Nonetheless, humans can act, and with a substantial degree of confidence, because of our knowledge of how the sort of situation we are facing ordinarily works—in short, because of our "common sense". If the basic task of AI is to create artificial persons, then we must give them that same common sense.

Representations of exactly "how the sort of situation we are facing ordinarily works"—what Ossorio calls "paradigm case formulations" (Ossorio, 1980/1981) and Shank calls "scripts" (Shank & Abelson, 1977)—are of necessity both detailed and complex. Every such representation, regardless of terminology and structure, can only tell us what is ordinarily the case; common sense comes in when we use what we know to be the case along with what is ordinarily the case to derive what we will take to be the case. To do this, we need some algorithms for common sense, which is the substantive topic of this section.

These algorithms (or rules, if you prefer) use language derived from Ossorio's conceptualization of real-world phenomena (Ossorio, 1971/1978). As with most of Ossorio's work, the terms are elaborated in substantial technical detail, but are carefully chosen so that the average reader will not be misled by assuming the ordinary English-language meaning. Accordingly, I have chosen simply to use the terms—object, process, event, state of affairs, relation—as a set of undefined primitives, like "point" and "line" in geometry, while inviting the interested reader to pursue the elaboration of these terms and their conceptual interconnections in Ossorio's "*What Actually Happens*".

Several of the algorithms speak of "reason to believe" and "reason enough". This reflects the perspective of Descriptive Psychology, which asserts that persons act on the basis of what they take to be the case. This includes both what they know (in the strict factual sense of what they have observed and/or deduced logically) and what they take to be the case because it is ordinarily part of the situation they have observed. Indeed, persons ordinarily do not distinguish between these two "types" of knowledge unless they have good reason to do so; they simply observe and act. Artificial persons, of course, must be more methodical about these matters than are most humans. Accordingly, it is necessary to make explicit the logic involved in such common sense inferences, which is rooted in paradigm case logic: We take it that things are the way they ordinarily are unless we have reason to believe otherwise.

Some algorithms for common-sense inference:

1. If object O1 exists, then any object O1n which in the paradigm case is O1's constituent also exists, unless:

    1. an alternate decomposition of O1 which does not contain O1n as a constituent is found to apply, or
    2. there is reason to believe O1n does not exist, in which case

a. a known alternative decomposition of O1 which does not contain O1n applies, or
b. there exists a decomposition of O1 which is unknown to us and does not contain O1n, or
c. O1 does not in fact exist.

2. If process P1 has occurred, then any subprocess P1n which is a paradigm case constituent of P1 has also occurred, unless:

   1. an alternative decomposition of P1 which does not contain P1n as a constituent is found to apply, or
   2. there is reason to believe P1n has not occurred, in which case

      a. a known alternative decomposition of P1 which does not contain P1n applies, or
      b. there exists a decomposition of P1 which is unknown to us and does not contain P1n, or
      c. P1 has not in fact occurred.

3. If state of affairs SA1 exists, then any object O1, process P1, event E1, or state of affairs SA1n which in the paradigm case is SA1's constituent also exists, unless:

   1. an alternate decomposition of SA1 which does not contain O1, P1, E1 or SA1n as a constituent is found to apply, or
   2. there is reason to believe O1, P1, E1, or SA1n does not exist, in which case
      a. a known alternative decomposition of SA1 which does not contain O1, P1, E1 or SA1n applies, or
      b. there exists a decomposition of SA1 which is unknown to us and does not contain O1, P1, E1 or SA1n, or
      c. SA1 does not in fact exist.

4. If object O1 exists, that is reason (but generally not reason enough) to believe that object O2, of which O1 is a paradigm case constituent, also exists.

Corollary:

4.1 If object O1, a paradigm case component of object O2, exists, that is reason (but generally not reason enough) to believe that object O3, which is also a paradigm case component of O2, exists. A similar rule applies to processes (5.1) and states of affairs (6.1).

5. If process P1 has occurred, that is reason (but generally not reason enough) to believe that process P2, of which P1 is a paradigm case constituent, has occurred or is occurring.
6. If object O1, process P1, event E1, or state of affairs SA1 exists, that is reason (but generally not reason enough) to believe that SA2, of which O1, P1, E1 or SA1 is a paradigm case constituent, exists.

Note: Rules 4, 5, and 6 suggest places in this inference scheme where "weights" or probability estimates could be both useful and in accord with common sense. How strong the "reason to believe" is depends on which component of which whole we have observed. The weighting is not generally a function of logical connection, but rather represents the cumulation of empirical experience with these types of objects, processes, etc.

7. If event E1 has occurred, then any process P1 which in the paradigm case is terminated by E1 has both occurred and is no longer occurring unless:

    1. an alternative process which is terminated by E1 is found to have occurred, or
    2. another event, E2, which begins P1 is found to have occurred after E1, in which case P1 has both occurred and is still occurring, or
    3. there is reason to believe P1 has not occurred, in which case

        a. some other process, P2, which is known to be terminated by E1 has occurred and is no longer occurring, or
        b. there exists some process P3 which is unknown to us which is terminated by E1 and which both has occurred and is no longer occurring, or
        c. E1 did not in fact occur, or,

    4. there is reason to believe P1 is still occurring, in which case

        a. another event, E2, which begins P1 has occurred after E1, in which case P1 has both occurred and is still occurring, or
        b. E1 did not in fact occur.

Note: By substituting "object" or "state of affairs" for the word "process" in Rule 7, and substituting "exist" for "occur", similar rules for relating events to objects and states of affairs can be written.

8. If two objects, O1 and O2, exist, both of which are constituents of object O3, that is stronger reason to believe that O3 exists than is the existence of either O1 or O2 alone. A similar rule applies to processes and states of affairs.

Corollary:

8.1 If two objects, O1 and O2, exist, both of which are constituents of object O3, that is reason (but generally not reason enough) to believe that the relation between O1 and O2 is R12, the relation that exists between O1 and O2 as constituents of O3. A similar rule applies to processes and states of affairs.

In conclusion, I would like to revisit the remarks at the beginning of this paper. Descriptive Psychology's approach to the enterprise of artificial intelligence is both distinctive and powerful. I have attempted to illustrate in this paper both the conceptual scope of the Descriptive approach, as well as some of the technical depth it contributes. Perhaps more importantly, I have attempted to give the reader an appreciation of the difference it makes to take this Descriptive approach. To those readers whose interest has been piqued by this exposition, I would like to echo the classic advice given in New York delis to hesitant patrons: "Try it! You'll like it!"

ACKNOWLEDGMENTS

Many people contributed to this paper in different ways. Joe Jeffrey and I worked together so closely for so long that his ideas and insights are inextricably woven into my own. I am especially grateful for Prof. Jeffrey's enthusiastic response to a three-page outline of this paper, without which enthusiasm I would never have begun writing. Paul Zeiger helped me sort through the computing science approach to AI, while reminding me that not everyone shares my belief in a behavioral approach. Keith Davis, Peter Ossorio and Tom Mitchell as always had helpful editorial suggestions on the manuscript. Finally, I would like to acknowledge the students and practitioners of Descriptive Psychology, who have pioneered an important new approach in the face of considerable difficulty.
Author's address: 1019 Baldwin Avenue, Ann Arbor, Michigan 48104.

REFERENCES

Busch, E. K. (1974). *The Boulder Police Department: A descriptive study of organizational structure-function and individual behavior*. Unpublished masters thesis, University of Colorado, Boulder.

Davis, K. E. (Ed.). (1981). *Advances in Descriptive Psychology* (Vol. 1). Greenwich, CT: JAI Press.

Davis, K. E. & Bergner, R, (Eds.). (1983). *Advances in Descriptive Psychology* (Vol. 3). Greenwich, CT: JAI Press.

Davis, K. E. & Mitchell, T. O. (Eds.). (1982). *Advances in Descriptive Psychology* (Vol. 2). Greenwich, CT: JAI Press.

Davis, K. E. & Mitchell, T. O. (Eds.). (1985). *Advances in Descriptive Psychology* (Vol. 4). Greenwich, CT: JAI Press.

Gardner, H. (1983). *Frames of mind*. New York: Basic Books.

Glymour, C. & Thomason, R. H. (1984). "Default reasoning and the logic of theory perturbation". In, *Non-monotonic reasoning workshop* (pp. 93-102). Menlo Park, CA: AAAI.

Grosof, B. W. (1984). Default reasoning as circumscription. In *Non-monotonic reasoning workshop* (pp. 115-124). Menlo Park, CA: AAAI.

Hayes, P. (1985). The second naive physics manifesto. In J. Hobbs & R. Moore (Eds.) *Formal theories of the common-sense world*. Norwood, NJ: Ablex.

Jeffrey, H. J. (1981). "A new paradigm for artificial intelligence". In K. E. Davis (Ed.), *Advances in Descriptive Psychology* (Vol. 1, pp. 177-194). Greenwich, CT: JAI Press.

Jeffrey, H. J., & Putman, A. O. (1983). The Mentor project: replicating the functions of an organization. In K. E. Davis & R. Bergner (Eds.), *Advances in Descriptive Psychology* (Vol. 3, pp. 243-269). Greenwich, CT: JAI Press.

Kurzweil, R. (1985, May-June). What is artificial intelligence, anyway? *American Scientist*, 73, 258-264.

McCarthy, J. (1986). Applications of circumscription to formalizing common-sense knowledge. *Artificial Intelligence*, 28, 89-116.

Mitchell, T. O. (1967). *Computer simulation of observational judgement of persons*. Unpublished manuscript.

Mitchell, T. O. (1981). On the interpretation of utterances. In K. E. Davis (Ed.), *Advances in Descriptive Psychology* (Vol. 1, pp. 155-176). Greenwich, CT: JAI Press.

Ossorio, P. G. (1964). *Classification space analysis* (RADC-TDR-64-287). Rome, NY: RADC Systems and Technology Division.

Ossorio, P. G. (1966). *Persons* (LRI Report No. 3). Los Angeles and Boulder: Linguistic Research Institute.

Ossorio. P. G. (1969/1978). *Meaning and Symbolism* (LRI Report No. 15). Whittier, CA and Boulder: Linguistic Research Institute, 1978. (Originally published in 1969 as LRI Report No. 10. Boulder: Linguistic Research Institute.)

Ossorio, P. G. (1971). *State of affairs system*. (RADC-TR-71-102). Rome, N.Y.: RADC Systems and Technology Division.

Ossorio, P. G. (1971/1978). *"What actually happens"*. Columbia, SC: University of South Carolina Press, 1978. (Originally published in an earlier version in 1971 as LRI Report No. 10a. Whittier, CA and Boulder: Linguistic Research Institute. Later issued as LRI Report No. 20).

Ossorio, P. G. (1980/1981). Representation, evaluation and research. In K. E. Davis (Ed.) *Advances in Descriptive Psychology* (Vol. 1, pp. 105-135). Greenwich, CT: JAI Press.

(Originally published in 1980 as LRI Report No. 24. Boulder: Linguistic Research Institute.)

Ossorio, P. G. (1982). Embodiment. In K. E. Davis & T. O. Mitchell (Eds.), *Advances in Descriptive Psychology* (Vol. 2, pp. 11-30). Greenwich, CT: JAI Press. (Originally published in 1980 as LRI Report No. 23. Boulder: Linguistic Research Institute.)

Ossorio, P. G. (1981/1983). A multicultural psychology. In K. E. Davis & R. Bergner (Eds.), *Advances in Descriptive Psychology* (Vol. 3, pp. 13-44). Greenwich, CT: JAI Press. (Originally published in 1981 as LRI Report No. 29. Boulder: Linguistic Research Institute.)

Putman, A. O. (1969). *Observer's judgement as a function of observer's abilities.* Unpublished manuscript.

Putman. A. O. (1982). DIAMOND: Specifications for a social practice language. Technical Report. Denver: Management Support Technology, Inc.

Putman, A. O., & Jeffrey, H. J. (1985). A new paradigm for software and its development. In K. E. Davis & R. Bergner (Eds.), *Advances in Descriptive Psychology* (Vol. 4, pp. 119–138). Greenwich, CT: JAI Press.

Reiter, R. (1980). A logic for default reasoning. *Artificial Intelligence*, 13, 81-132.

Shank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals and understanding.* Hillsdale, NJ: Erlbaum.

Sternberg, T. (1985). *Beyond IQ.* Cambridge: Cambridge University Press.

Winograd, T. (1982). *Language as a cognitive process: Vol. 1: Syntax.* New York: Addison-Wesley.