

**Palo Alto Research Center**

**PLANS AND SITUATED ACTIONS:**

**The problem of human-machine  
communication**

**Lucy A. Suchman**

**XEROX**

# **PLANS AND SITUATED ACTIONS:**

## **The problem of human-machine communication**

**Lucy A. Suchman**

**February 1985**

**ISL-6**

**Corporate Accession P85-00005**

Copyright Lucy A. Suchman 1985. All rights reserved.

**XEROX**

**Xerox Corporation  
Palo Alto Research Centers  
3333 Coyote Hill Road  
Palo Alto, California 94304**



## **Abstract**

This thesis considers two alternative views of purposeful action and shared understanding. The first, adopted by researchers in Cognitive Science, views the organization and significance of action as derived from plans, which are prerequisite to and prescribe action at whatever level of detail one might imagine. Mutual intelligibility on this view is a matter of the recognizability of plans, due to common conventions for the expression of intent, and common knowledge about typical situations and appropriate actions. The second view, drawn from recent work in social science, treats plans as derivative from situated action. Situated action as such comprises necessarily *ad hoc* responses to the actions of others and to the contingencies of particular situations. Rather than depend upon the reliable recognition of intent, successful interaction consists in the collaborative production of intelligibility through mutual access to situation resources, and through the detection, repair or exploitation of differences in understanding.

As common sense formulations designed to accommodate the unforeseeable contingencies of situated action, plans are inherently vague. Researchers interested in machine intelligence attempt to remedy the vagueness of plans, to make them the basis for artifacts intended to embody intelligent behavior, including the ability to interact with their human users. The idea that computational artifacts might interact with their users is supported by their reactive, linguistic, and internally opaque properties. Those properties suggest the possibility that computers might 'explain themselves,' thereby providing a solution to the problem of conveying the designer's purposes to the user, and a means of establishing the intelligence of the artifact itself.

I examine the problem of human-machine communication through a case study of people using a machine designed on the planning model, and intended to be intelligent and interactive. A conversation analysis of "interactions" between users and the machine reveals that the machine's insensitivity to particular circumstances is a central design resource, and a fundamental limitation. I conclude that problems in Cognitive Science's theorizing about purposeful action as a basis for machine intelligence are due to the project of substituting plans for actions, and representations of the situation of action, for action's actual circumstances.



## Acknowledgements

The greatest contribution to this project over the last several years has been the combination of time, resources, freedom to work, faith that something good would come of it, and intellectual support provided by John Seely Brown, Manager of the Intelligent Systems Laboratory at Xerox Palo Alto Research Center. He and other colleagues have nourished my slowly emerging appreciation for the 'anthropologically strange' community that is Cognitive Science and its related disciplines. In particular I have benefited from discussions with Danny Bobrow, Johan deKleer, Sarah Douglas, Richard Fikes, Austin Henderson, David Levy, Tom Moran, Brian Smith, Kurt Vanlehn, and Terry Winograd. Recent conversations with Stan Rosenschein on 'situated automata' have opened yet another window for me onto the problem of how this community proposes that we understand action. Needless to say, while I am deeply grateful for their contributions, none of them is responsible for the results.

In my own field of anthropology, I have enjoyed the intellectual and personal friendship of Brigitte Jordan, whose creative energy and respectful sensibilities toward her own work and life are an example for mine. I am deeply grateful to Doug Macbeth, with whom I shared my discovery and exploration of the field of social studies and its possibilities. While he would doubtless argue innumerable points in the pages that follow, his influence is there. Mike Lynch and Steve Woolgar both provided thoughtful responses to early drafts. Particularly in its early stages, this project benefited greatly from lively discussions at the Interaction Analysis Lab at Michigan State University, where Fred Erickson, Rich Frankel, Brigitte Jordan, Willett Kempton, Bill Rittenberg, Ron Simons and others helped me first to penetrate the thickness of a video analysis. Jeanette Blomberg, and more recently Julian Orr are anthropological colleagues in the PARC community, and I have enjoyed their company. I am grateful to Hubert Dreyfus and John Gumperz, as members of my thesis committee, for their substantive and stylistic contributions, and for their enthusiasm for the project. My thesis advisor, Gerald Berreman, is through his own career an example of what ethical scholarship can be. While finding my work increasingly foreign and exotic, he has remained an unflagging supporter.

Finally, of course, I thank my friends new and old, in particular Mimi Montgomery who over the last several years, along with this thesis, has been my most constant companion.



## **Table of Contents**

PREFACE	1
CHAPTER 1: INTRODUCTION	3
CHAPTER 2: INTERACTIVE ARTIFACTS	7
CHAPTER 3: PLANS	21
CHAPTER 4: SITUATED ACTIONS	35
CHAPTER 5: COMMUNICATIVE RESOURCES	47
CHAPTER 6. CASE AND METHODS	65
CHAPTER 7: HUMAN-MACHINE COMMUNICATION	77
CHAPTER 8: CONCLUSION	123
REFERENCES	125
APPENDIX	133



## Preface

Thomas Gladwin (1964) has written a brilliant article contrasting the method by which the Trukese navigate the open sea, with that by which Europeans navigate. He points out that the European navigator begins with a plan—a course—which he has charted according to certain universal principles, and he carries out his voyage by relating his every move to that plan. His effort throughout his voyage is directed to remaining ‘on course.’ If unexpected events occur, he must first alter the plan, then respond accordingly. The Trukese navigator begins with an objective rather than a plan. He sets off toward the objective and responds to conditions as they arise in an *ad hoc* fashion. He utilizes information provided by the wind, the waves, the tide and current, the fauna, the stars, the clouds, the sound of the water on the side of the boat, and he steers accordingly. His effort is directed to doing whatever is necessary to reach the objective. If asked, he can point to his objective at any moment, but he cannot describe his course (Gerald Berreman 1966).

The subject of this thesis is the two alternative views of human intelligence and directed action represented by the Trukese and the European navigators. The European navigator embodies the prevailing scientific model of purposeful action, for reasons that are implicit in the final sentence of the quote above. That is to say, the Trukese navigator is hard pressed to tell us how he actually steers his course, while the comparable account for the European seems to be ready-at-hand, in the form of the very plan that is taken to guide his actions. While the objective of the Trukese navigator is clear from the outset, his projected course is necessarily vague, insofar as his actual course is contingent on unique circumstances that he cannot anticipate in advance. The plan of the European, in contrast, is derived from universal principles of navigation, and is essentially independent of the exigencies of his particular situation.

The image of the European navigator, deeply entrenched in the Western human sciences as *the* correct model of the purposeful actor, is now in the process of being reified in the form of new, computational artifacts. In this thesis I examine one such artifact, as a way of investigating the strengths and limitations of the general view that the design embodies. The properties of the plan make it attractive for the purpose of constructing a computational model of purposeful action, to the extent that for those fields devoted to what is now called Cognitive Science, the analysis and synthesis of plans effectively constitutes the study of action. The contention of this thesis, however, is that as students of human action we ignore the Trukese navigator at our peril, because while an account of how the European navigates may be ready-at-hand, the essential nature of situated action, however planned or unplanned, is Trukese. It behooves us, therefore, to study, and to begin to find ways to describe, the Trukese system.

There is an injunction in social studies of science to eschew interest in the validity of the products of science, in favor of an interest in their production. While I generally agree with this injunction, my investigation of one of the prevailing models of human action in Cognitive Science is admittedly and unabashedly interested. That is to say, I take it that there *is* a reality of human action beyond either the Cognitive Scientist’s models or my own accounts, to which both are trying to do justice. In that sense, I am not just examining the Cognitive Science model with the

dispassion of the uncommitted anthropologist of science, I am examining it in light of an alternative account of human action to which I am committed, and which I attempt to clarify in the process.

## 1. Introduction

The famous anthropological absorption with the (to us) exotic ... is, thus, essentially a device for displacing the dulling sense of familiarity with which the mysteriousness of our own ability to relate perceptively to one another is concealed from us (Clifford Geertz 1973, p. 14).

The problem of shared understanding, or mutual intelligibility, has defined the field of social studies for the past 100 years. On the one hand, the interpretation of action has been the social scientist's task; to come up with accounts of the significance of human actions is, after all, the principal charge of ethnographic anthropology. On the other hand, to understand the mutual intelligibility of action as a mundane, practical accomplishment of members-of-the-society is, in large measure, the social scientist's problem or subject matter. An account of that accomplishment would constitute an account of the foundation of social order.

While studies of mutual intelligibility have been concerned exclusively with human action, we now have a new technology which has brought with it the idea that rather than just using machines, we interact with them. In particular, the notion of "human-machine interaction" pervades technical and popular discussion of computers, their design and use. Amidst ongoing debate over specific problems in the design and use of interactive machines, however, no question is raised regarding the bases for the idea of human-machine interaction itself. At the same time, recent developments in the social sciences regarding the foundations of human interaction have had remarkably little influence on the discussion of interactive machines.

Every human tool relies upon, and reifies in material form, some underlying conception of the activity that it is designed to support. As a consequence, one way to view the artifact is as a test on the limits of the underlying conception. The motivation for this research was to examine the conception of purposeful action, and consequently of interaction, directing the design of interactive machines. I take it that interaction, or communication—I'll use the two interchangeably—turns on the extent to which my words and actions and yours are mutually intelligible. Beginning with this general characterization, I investigate the basis for beginning to speak of interaction, or mutual intelligibility, between humans and machines.

My central concern in the investigation is a new manifestation of an old problem in the study of mutual intelligibility; that is, the relation between observable behavior, and the processes—not available to direct observation—that make behavior meaningful. For psychological studies, the crucial processes are essentially cognitive, located inside the head of the actor; i.e. the formation and effect of beliefs, desires, intentions and the like. For social studies, the crucial processes are essentially interactional and circumstantial, located in the relationships among actors, and between actors and their embedding situations. In either case, what defines the problem of meaningful action is the observation that behavior is inherently subject to indefinitely many ascriptions of meaning or intent, while meaning and intent are inherently expressible through an indefinite number of possible behaviors. Whether the final arbiter of meaning is taken to be private psychological processes, or socially constructed criteria of accountability to the public world, the

question to be resolved—what defines the essence of an intentional action—is the same.

One new manifestation of this old problem is a technological one, however, arising from the idea of constructing artificial devices—human artifacts—that would behave purposefully and intelligently. A problem for this project, discussed at length in the next chapter, is to know when the project has succeeded. That is, if one builds a device whose behavior is indistinguishable from that of an intelligent actor, is that device intelligent? If not—if the process that generates the behavior, and not just the behavior produced, matters—then in what sense must the processes be the same? If the criterion for sameness of processes is the identity of input and output behaviors, then we're back where we started. But if not, then we have no behavioral test.

This paradox is rendered somewhat moot by the problems that arise in constructing a device that even *appears* to behave in ways that are purposeful or intelligent. It may well turn out that the resistance of meaningful action to simulation in the absence of any deep understanding will defend us against false impressions of success. In any case, my purpose here is to clarify some of the existing troubles in the project of constructing interactive machines, as a way of contributing to our understanding of interaction. To ground the investigation, I use a particular machine as a case study. The aim of the case study is not to criticize the particular design, but to view the machine as reifying certain premises about purposeful action. The task is to articulate those premises, to see how they succeed as a basis for human-machine communication and how they fail, and to explore the implications of their success and failure for both the design of human-machine communication, and the problem of purposeful action and shared understanding in general.

The machine studied is something of a hybrid of old and new technologies; a large photocopier, controlled by a computer-based system intended to act as an artificially intelligent ‘expert’ in the copier’s use. This ‘expert help system’ embodies a conception of human action shared by designers, the behavioral sciences and our commonsense. Briefly, the conception is that the significance of what people say and do is best understood as the reflection of their underlying plans. Applied to communication, this view holds that the coherence of action is individual, and is given in advance, and identifies the problem for conversants as the transmission and recognition of their respective plans.

In this thesis I argue that artifacts built on a planning model of human action confuse *plans* and *situated action*. The behavioral science and commonsense that supports the design of interactive artifacts treats a plan as something located in the actor’s head, which directs his or her behavior. In contrast, this study adopts a view of plans just as formulations of antecedent conditions and consequences of action, which account for action in a plausible way. Stated in advance, plans are necessarily vague, insofar as they are designed to accommodate the unforeseeable contingencies of actual situations of action. Reconstructed retrospectively, plans systematically ignore the necessary *ad hocness* of situated action in favor of an account of the action as in accord with the plan. As ways of talking about action, plans *per se* neither determine the actual course of situated action, nor adequately reconstruct it. While for purposes of practical action this fact is uninteresting, for purposes of a *science* of practical action it is crucial. Specifically, if we are

interested in situated action itself, we need to look at how it is that actors use the circumstances that a particular occasion provides—including, but crucially not reducible to, formulations like plans—to provide for their action's developing purpose and intelligibility.



## 2. Interactive Artifacts

Marginal objects, objects with no clear place, play important roles. On the lines between categories, they draw attention to how we have drawn the lines. Sometimes in doing so they incite us to reaffirm the lines, sometimes to call them into question, stimulating different distinctions (Sherry Turkle 1984).

In *The Second Self* (1984), Sherry Turkle describes the computer as an "evocative object," one that raises new questions regarding our common sense of the distinction between artifacts and intelligent others. Her studies include an examination of the impact of computer-based artifacts on children's conceptions of the difference between 'alive' *versus* 'not alive,' and 'machine' *versus* 'person.' In dealing with the questions that computer-based objects evoke, children make clear that the differentiation of physical and psychological entities, which as adults we largely take for granted, is the end product of a process of establishing the relationship between the observable behavior of a thing and its underlying nature. Children have a tendency, for example, to attribute aliveness to physical objects on the basis of behavior like autonomous motion, or reactivity, while reserving humanity for entities evidencing such things as emotion, speech, and apparent thought or purposefulness. Turkle's observation with respect to computational artifacts, however, is that children ascribe to them an 'almost aliveness,' and a psychology, while maintaining their distinctness from human beings: a view that, as Turkle points out, is remarkable among other things for its correspondence to the views held by those who are the artifacts' designers.<sup>1</sup>

The point of departure for this research is a particular aspect of the phenomenon that Turkle identifies; namely, the apparent challenge that computational artifacts pose to the longstanding distinction between the physical and the social, in the special sense of those things that one designs, builds, and uses, on the one hand, *versus* those things with which one communicates, on the other. While this distinction has been relatively non-problematic to date, now for the first time the term interaction—in a sense previously reserved for describing a uniquely interpersonal activity—seems appropriately to characterize what goes on between people and certain machines as well.<sup>2</sup> Interaction between people and machines necessarily implies mutual intelligibility, or shared understanding. What motivates this inquiry, therefore, is not only the recent question of how there

1. See especially pp. 62-63; Turkle also finds some cause for alarm in the fact that for children the distinction of machine and person seems to turn centrally on a separation of thought from feeling, i.e. computers exhibit the former, but lack the latter. This view, she points out, includes a kind of dissociation of intellect and emotion, and consequent trivialization of both, that characterizes the attitudes of many in the field of Artificial Intelligence.

2. I recognize that the term interaction actually originates with the physical sciences, to describe a mutual or reciprocal action or influence. As used here, however, I mean to cite the common sense assigned the term by social science; specifically, communication between persons. The migration of the term from the physical sciences to the social, and now back to some ground that stands between them, ties in intriguing ways to a general blurring of the distinction between physical and social in modern science, and to the general question of whether machines are actually becoming more like people or whether, in fact, people are coming to define themselves more as machines. There is clearly a mutual influence at work. See Dreyfus 1979, chapter 9.

could be mutual intelligibility between people and machines, but the prior question of how we account for the shared understanding, or mutual intelligibility, that we experience as people in our interactions with others whose essential 'sameness' is not in question. An answer to the more recent question, theoretically at least, presupposes an answer to the earlier one.

This chapter locates the origins of the idea of human-machine interaction in some distinctive properties of computational artifacts, and in the emergence of disciplines dedicated to making those artifacts "intelligent." I begin with a brief discussion of Cognitive Science, the interdisciplinary field devoted to simulating cognitive processes, and its relation to the project of creating intelligent artifacts. Along with the theoretical interest in intelligent artifacts, the computer's properties have inspired a practical effort at engineering interaction between people and machines. I argue that the description of computational artifacts as interactive is supported by their *reactive*, *linguistic*, and internally *opaque* properties. With those properties in mind, I consider the double sense in which researchers are interested in artifacts that "explain themselves:" on the one hand, as a solution to the longstanding problem of conveying the artifact's *intended purpose* to the user, through its design and attendant instructions and, on the other hand, as a means of establishing the intelligence, or *rational accountability*, of the artifact itself.

## 2.1 Automata and Cognitive Science

Historically the idea of *automata*—the possibility of constructing physical devices that are self-regulating in ways that we commonly associate with living, or animate beings—has been closely tied to the simulation of animal forms. McCorduck (1979) points out that human-like automata have been constructed since Hellenic times; statues that moved, gestured, spoke and generally were imbued by observers—even those well aware of the internal mechanisms that powered them—with everything from minds to souls.<sup>3</sup> In the 14th century in Western Europe, learned men were commonly believed to construct "talking heads" made of brass, considered as both the source of their wisdom, and its manifestation. More prosaically, Jacques de Vaucanson in the 18th century designed a series of renowned mechanical statues, the most famous being a duck, the inner workings of which produced a variety of simple outward behaviors.

At the same time, however, Julien de la Mettrie published *Man, A Machine*, in which he argued that the vitality characteristic of human beings was the result neither of something immanent in their material substance, nor of any immaterial force, but of the *organization* of their physical structure. Cognitive Scientists today maintain the basic premise of de la Mettrie with respect to mind, contending that mind is best viewed as neither substance nor as insubstantial, but as an abstractable structure implementable in any number of possible physical substrates. Two ideas are

---

3. See McCorduck 1979, Chapter 1; Churchland 1984, Chapter 6. For an extensive history of automata, see Cohen 1966.

central to this contention. First, there is the insight, due to Turing, that given the right form of input and the right mechanism for its manipulation, a machine with a finite number of states can produce an infinite range of behaviors. And second, there is the claim that intelligence is only incidentally embodied in the neurophysiology of the human brain—that what is essential about intelligence can be abstracted from that particular, albeit highly successful, substrate and embodied in an unknown range of alternative forms. Both ideas decouple reasoning and intelligence from things uniquely human, and open the way for the construction of intelligent artifacts.<sup>4</sup>

The preoccupation of Cognitive Science with mind in this abstract sense is in part a reaction to behaviorist psychology, and a movement to restore ‘meaning’ to psychological explanation (cf. Stich 1983, chpt. 1). Specifically, the commitment to a cognitivist account of human action stands in opposition to the stimulus-response brand of environmental determinism proposed by behaviorists like Watson and Skinner. The cognitivist strategy is to interject a mental operation between environmental stimulus and behavioral response; in essence, to relocate the causes of action from the environment that impinges upon the actor, to processes—abstractable as computation—in the actor’s head. The first premise of Cognitive Science, therefore, is that people—or ‘cognizers’ of any sort—act on the basis of symbolic ‘representations;’ a kind of cognitive code, instantiated physically in the brain, on which operations are performed to produce mental states like ‘the belief that *p*,’ which in turn produce behavior consistent with those states. The relation of environmental stimuli to those mental states, on the one hand, and of mental states to behavior, on the other, remains deeply problematic and widely debated within the field (see, for example, Fodor 1983; Pylyshyn 1984; Stich 1984). The agreement among all participants in Cognitive Science and its affiliated disciplines, however, is that cognition is not just potentially *like* computation, it literally *is* computational. There is no reason, in principle, why there should not be a computational account of mind, therefore, and there is no *a priori* reason to draw a principled boundary between people, taken as ‘information-processors’ or ‘symbol manipulators’ or, in George Miller’s phrase, ‘informavores’ (Pylyshyn 1984, p. x), and certain computing machines.

The practical implementation of the view that intelligence is the manipulation of symbols, and the most successful present day automata, are the so-called “expert systems” that structure and process large amounts of well-formulated data, and the industrial robots that perform routine, repetitive assembly and control tasks. Expert systems—essentially sophisticated programs that manipulate data structures to accord with rules of inference that experts are understood to use—have minimal sensory-motor, or “peripheral” access to the world in which they are embedded, input being most commonly through a keyboard, by a human operator. Industrial robots—highly specialized, computer-controlled devices designed to perform autonomously a single repetitive physical task—have relatively more developed sensory-motor apparatus than do expert systems, but the success of robotics is still confined to specialized activities, under controlled conditions. In both

4. See Turkle 1984 chapter 7; and McCorduck 1979 chapter 5. Turkle’s description of the present academic AI culture at MIT is particularly insightful.

cases, the systems can handle large amounts of encoded information, and syntactic relationships of great sophistication and complexity, in highly circumscribed domains. But when it comes either to direct interaction with the embedding world, or to the exercise of practical, everyday reasoning about the significance of events in the world, there is wide agreement that the state-of-the-art in "intelligent" machines has yet to attain the general abilities of the normal two-year old child.

## 2.2 The idea of human-computer interaction

In spite of the persisting limits on machine intelligence, the use of an intentional vocabulary is already well-established in both technical and popular discussion of computers. In part, the attribution of purpose to computer-based artifacts derives from the simple fact that each action by the user effects an immediate machine *reaction* (cf. Turkle 1984, chpt. 8). Technically and historically this immediacy stands in contrast to earlier forms of computing, specifically "batch processing," where user commands were queued, and executed without any intermediate feedback. A combination of progress in integrated circuit technology that enormously extended the storage capacity and speed of computers, new transmission devices, new languages for programming, and the expansion of computer applications out into the public world of non-programmers all contributed to the development of more direct "interaction" between user and computer. By some definitions (e.g. Oberquelle *et al* 1983, p. 313), the criterion for "interactive computing" is just that real-time control over the computing process is placed in the hands of the user, through immediate processing, and through the availability of interrupt facilities whereby the user can override and modify the operations in progress. This reactivity, combined with the fact that, like any machine, the computer's reactions are not random but by design, suggest the character of the computer as a purposeful, and, by association, as a social object.

A more profound basis for the relative 'sociability' of computer-based artifacts, however, is the fact that the means for controlling computing machines and the behavior that results are increasingly *linguistic*, rather than mechanistic. That is to say, machine operation becomes less a matter of pushing buttons or pulling levers with some physical result, and more a matter of specifying operations and assessing their effects through the use of a common language.<sup>5</sup> With or without machine intelligence, this fact has contributed to the tendency of designers, in describing what goes on between people and machines, to employ terms borrowed from the description of human interaction—dialogue, conversation, and so forth—terms that carry a largely unarticulated

5. The popular fantasy of the "talking machine" notwithstanding, the crucial element that invites a view of computers as interactive is language, not speech. While strictly speaking buttons and keys remain the principal input devices in computing, this is relatively trivial. The synthesis of speech by computers may well add to *our* inclination to ascribe understanding to them, but will not, in itself, contribute substantively to *their* sensibility. On the other hand, simulation of natural language *understanding*, even when the language is written rather than spoken, is proving to be a profoundly difficult problem that is inseparable from the problem of simulating intelligence as such.

collection of intuitions about properties common to human communication and the use of computer-based machines.

While for the most part the vocabulary of human interaction has been taken over by researchers in human-machine communication with little deliberation, several recent journal articles attempt to clarify similarities and differences between computer use and human conversation. Perhaps the most thoughtful and comprehensive of these is Hayes and Reddy (1983). They identify the central difference between existing "natural language processing" systems and human communication as a question of "robustness," or the ability to respond to unanticipated circumstances, and to detect and remedy troubles in communication:

[T]he ability to interact gracefully depends on a number of relatively independent skills: skills involved in parsing elliptical, fragmented, and otherwise ungrammatical input; in ensuring that communication is robust (ensuring that the intended meaning has been conveyed); in explaining abilities and limitations, actions and the motives behind them; in keeping track of the focus of attention of a dialogue; in identifying things from descriptions, even if ambiguous or unsatisfiable; and in describing things in terms appropriate for the context. While none of these components of graceful interaction has been entirely neglected in the literature, no single current system comes close to having most of the abilities and behaviours we describe, and many are not possessed by any current systems (*ibid.*, p. 232).

Hayes and Reddy believe, however, that "[e]ven though there are currently no truly gracefully interacting systems, none of our proposed components of graceful interaction appears individually to be much beyond the current state of the art, at least for suitably restricted domains of discourse." They then review the state of the art, including systems like LIFER (Hendrix 1977) and SCHOLAR (Carbonell 1971), which display sensitivity to the user's expectations regarding acknowledgement of input; systems that resolve ambiguity in English input from the user through questions (Hayes 1981); systems like the GUS system (Bobrow *et al* 1977) which represent limited knowledge of the domain that the interaction is "about;" work on the maintenance of a common focus over the course of the interaction (Grosz 1978; Sidner 1979); and Hayes and Reddy's own work on an automated explanation facility in a simple service domain (1983).

Two caveats on Hayes and Reddy's requirements for a gracefully interacting system—both of which, to their credit, they freely admit—are worth noting. First, they view the abilities cited as necessary but not sufficient for human interaction, the claim for the list being simply that "it provides a good working basis from which to build gracefully interacting systems" (*ibid.*, p. 233). And not surprisingly, the abilities that they cite are precisely a list of problems currently under consideration in research on human-machine communication. There is, in other words, no independent assessment of how the problems on which researchers work relate to the nature and organization of human communication as such. Second, research on those problems that have been identified is confined to highly circumscribed domains. The consequence of working from a partial list of abilities in limited domains is that practical inroads in human-machine communication can be furthered, while the basic question of what human interaction comprises (of which their list of abilities is an admittedly *ad hoc* selection), and why research in human-machine interaction has

proceeded only in limited domains (beyond strictly methodological convenience), are deferred.

Moreover, while Hayes and Reddy take the position that "it is very important for a gracefully interacting system to conduct a dialogue in as human-like a way as possible" (*ibid.*, p. 233), this assertion is a point of controversy in the research community. On the one hand, there is an argument to the effect that one should acknowledge, and even exploit the fact that people bring to computer use a tremendous range of skills and expectations from human interaction. Within research on human-computer communication, for example, projects are underway to allow people to enter commands into computers using "natural language" (i.e. languages like English, in contrast to programming languages). On the other hand, even Hayes and Reddy admit that:

the aim of being as human-like as possible must be tempered by the limited potential for comprehension of any foreseeable computer system. Until a solution is found to the problems of organizing and using the range of world knowledge possessed by a human, practical systems will only be able to comprehend a small amount of input, typically within a specific domain of expertise. Graceful interaction must, therefore, supplement its simulation of human conversational ability with strategies to deal naturally and gracefully with input that is not fully understood, and, if possible, to steer a conversation back to the system's home ground (*ibid.*, p. 233).

Hayes and Reddy would make these recovery strategies transparent to the user, but they also point to the "habitability" problem identified by Watt (1968) with respect to language, i.e. the tendency of human users to assume that a computer system has sophisticated linguistic abilities after it has displayed elementary ones. This tendency is not surprising, given the fact that our only precedent for language using entities to date has been other human beings. The misconceptions that ensue lead some, for example Fitter (1979), to argue that English or other "natural" languages are in fact not natural for purposes of human-computer communication:

for the purpose of man-computer communication, a *natural language* is one that makes explicit the knowledge and processes for which the man and computer share a common understanding ... it becomes the responsibility of the systems designer to provide a language structure which will make apparent to the user the procedures on which it is based and will not lead him to expect from the computer unrealistic powers of inference (*ibid.*, p. 340).

As soon as computational artifacts demonstrate *some* evidence of recognizably human abilities, in other words, we are inclined to endow them with the rest. In view of the tendency to ascribe full intelligence on the basis of partial evidence, one can argue that designers would do best to make available to the user just how the system is *not* like a participant in interaction.<sup>6</sup> In this spirit, Nickerson (1976) argues that

The model that seems appropriate for this view of person-computer interaction is that of an individual making use of a sophisticated tool and not that of one person conversing

---

6. Nickerson (1976) points out that there are some ways in which a computer is not like another person which lend a certain advantage to the user, e.g. interruptions can be done without regard for offense, responses can be delayed as long as is necessary.

with another. The term "user" is, of course, often used to denote the human component in a person-computer interaction, as it has been in this paper. It is, to my taste, preferable to the term "partner," not only because it seems more descriptive of the nature of the relationships that existing systems permit, and that future systems are likely to, but because it implies an asymmetry with respect to goals and objectives that "partner" does not. "User" is not a term that one would normally apply to a participant in a conversation (*ibid.*, p. 111).

The argument that processes should be revealed to the user, however, is potentially counter to the promotion of an intentional vocabulary in speaking about computer-based devices. Quite apart from either reactivity or language, it is precisely the complexity and opacity of the computer's inner workings that invites description in intentional terms (cf. Dennett 1978, chapter 1). Despite design philosophies like that embodied in the WEST system (Burton and Brown 1979), which includes a so-called "glass box" that reveals a part of the underlying mechanism to the user, the computer generally is a 'black box' for most users. This is the case not only because users lack technical knowledge of its internal workings but because, even for those who possess such knowledge, there is an "irreducibility" to the computer as an object that is unique among human artifacts (Turkle 1984, p. 272). The overall behavior of the computer is not describable with reference to any of the simple local events that it comprises; it is precisely the behavior of a myriad of those events in combination that constitutes the overall machine. To refer to the behavior of the machine, then, one must speak of "its" functionality. And once reified as an entity, the inclination to ascribe actions to the entity rather than to the parts is irresistible. For one thing, intentional explanations relieve us of the burden of actually understanding the mechanism, insofar as one need only assume that the design is rational in order to call upon the full power of common sense psychology and have, ready-at-hand, a basis for anticipating and construing its behavior. At the same time, precisely because the mechanism is in fact unknown, and insofar as underspecification is taken to be characteristic of human beings (as evidenced by the fact that we are inclined to view something that is fully specified as less than human), the personification of the machine is reinforced by the ways in which its inner workings are a mystery, and its behavior at times surprises us. Insofar as the machine is somewhat predictable, in sum, and yet is also both internally opaque and liable to unanticipated behavior, we are more likely to view ourselves as engaged in interaction with it than as just performing operations upon it, or using it as a tool to perform operations upon the world (cf. MacKay 1962).

### 2.3 Self-explanatory artifacts

At the same time that computational artifacts introduce new complexity and opacity into our encounters with machines, our reliance on computer-based technology, and its proliferation throughout the society increases. One result is the somewhat paradoxical objective that increasingly complex technology should be usable with decreasing amounts of training. This objective lends

renewed urgency to the problem of "user interface" design.<sup>7</sup> Designers have long held the view that ideally a device should be self-explanatory, that is, decipherable solely from information provided on or through the device itself. With a computer-based artifact, however, the notion of a self-explanatory machine becomes ambiguous, in a way that reflects a broader equivocation within the design community about just what is meant by a machine's intelligibility, and by human-machine communication.<sup>8</sup> Generally, in assessing whether an artifact is or could be self-explanatory, designers mean just the extent to which someone examining the artifact is able to reconstruct *the designer's intentions* regarding its use. This idea—that a self-explanatory artifact is one whose intended purpose is discoverable by the user—is presumably as old as the design and use of tools as such. With respect to computer-based artifacts, however, the notion of a self-explanatory artifact suggests that the artifact might actually *explain itself* in something more like the sense that a human being does. In this second sense the idea is that the artifact should not only be intelligible to the user as a tool, but that it should be intelligent, i.e. able to understand the actions of the user, and to provide for the rationality of its own.

In the remainder of this chapter, I look at these two senses of a self-explanatory machine and at the relationship between them. The first sense—that a tool should be decipherable by its user—reflects the fact that artifacts are constructed by designers, for a purpose, and that the user of a tool needs to know something of that design intent. The difficulty of reconstructing the artifact's intended use from the design alone has led to attempts over the years to supplement tools with instructions for their use. Now, computational tools seem to offer unique capabilities for the provision of instruction to their users. The distance is not far from the idea that instructions could be presented more effectively using the power of computation, to the idea that computer-based artifacts could actually instruct; i.e. could interact with people in a way that approximates the behavior of an intelligent human expert or coach. And this second idea, that the artifact could actually interact instructively with the user, ties the practical problem of instruction to the theoretical problem of building an intelligent, interactive machine.

### *2.3.1 The computer as an artifact designed for a purpose*

The designer of any artifact that is a tool must communicate the artifact's intended use and, in some cases, the rationale for its behavior, to the user. There is a strong sense, therefore, in which

7. In design parlance, the term "user interface" refers both to the physical place at which the user issues commands to a device, finds reports of its state, or obtains the products of its operation, and the procedures by which that occurs.

8. For example, in a recent article on human-machine communication in the International Journal of Man-Machine Studies (Oberquelle *et al.* 1983), the authors initially refer to computers as a medium for human communication: "Computers just play a special role as one element in a highly complex communication network *with several human agents*...[e.g. designer, implementer, user]" Further down on the same page they write: "The problems of today's computer use mainly result from difficulties in the communication *between the human and the machine*" (*Ibid.*, p. 309, emphasis added).

the problem of designing and using artifacts is precisely a problem of communication. On one premise this simply means that the artifact conveys the intentions of its designer more or less directly—that the intended use is, or at least should be, self-evident from the design. The problem with such a premise, however (as archaeologists well know), is that while the attribution of *some* design intent is a requirement for an artifact's intelligibility, the artifact's design *per se* does not unequivocally convey either its actual, or its intended use. While this problem in the interpretation of artifacts can be alleviated, it can never fully be resolved, and it defines the essential problem that the novice user of the artifact confronts. Insofar as the goal of design is that the artifact should be self-evident, therefore, the problem of deciphering an artifact defines the problem of the designer as well.

As with any communication, instructions for the use of a tool are guided generally by the maxim that utterances should be designed for their recipients. The extent to which the maxim is observed in instruction is limited in the first instance by the resources that the medium of communication affords. Face-to-face human interaction is the paradigm case of a system for communication that, because it is organized for maximum context-sensitivity, supports a response designed for just this recipient, on just this occasion. Face-to-face instruction brings that context-sensitivity to bear on problems of skill acquisition. The gifted coach, for example, draws on powers of language and observation, and uses the situation of instruction, in order to specialize instruction for the individual student. Where written instruction relies upon generalizations about its recipient and the occasion of its use, the coach draws pedagogical strength from exploitation of the unique details of particular situations.<sup>9</sup>

A consequence of the human coach's method is that his or her skills must be deployed anew each time. An instruction manual, in contrast, has the advantage of being durable, re-usable, and replicable. In part, the strength of written text is that, in direct contrast to the pointed commentary of the coach, text allows the *disassociation* of the occasion of an instruction's production from the occasion of its use. For the same reason, however, text affords relatively poor resources for recipient design. The promise of interactive computer systems, in these terms, is a technology that can move instructional design away from the written manual in the direction of the human coach, and the resources afforded by face-to-face interaction. Efforts at building self-explicating machines in their more sophisticated forms now adopt the metaphor of the machine as an expert in its own use, and the user as a novice, or student. The system studied here, called a "context-dependent

---

9. Face-to-face interaction is in most cases a necessary, but is of course never a sufficient, condition for successful human coaching. Coombs and Alty (1984) provide an interesting discussion of the failings of interactions between *human* advisors and new computer users. At the same time, they point out that the characteristics of the advisory sessions that new users found unsatisfactory show marked similarities to human interactions with most rule-based computer "help" systems, e.g. that the advisors provide only the recommended solutions to reported problems, while failing either to elicit the view of the user, or to articulate any of their own rationale. Satisfactory sessions, in contrast, were characterized by what initially appeared to be less structure and less economy, but which on further investigation was revealed as "well-motivated despite surface appearances, the objective not being strict problem-solving as we had assumed, but problem-solving through mutual understanding. This required sensitivity to different structural factors" (pp. 24-25).

"expert help system," is such an effort.<sup>10</sup> A basic aim of the system is that rather than providing a compendium of information and leaving decisions of relevance to the user, information should be occasioned by and fitted to the user's inquiries. In order to provide not just a set of instructions, but an occasioned response, the system designer must now define not only the information, but how the system should recognize the situation for which that information is appropriate. Crucially, the relevant situations are constituted by the user's actions. Consequently, the system must in some sense be able to find their significance.

Among the most interesting attempts to design a computer-based "coach" is Burton and Brown's WEST system (1979). The philosophy underlying WEST includes several observations to the effect that the skill of a human coach lies as much in isn't said as it does in what is said. The human coach does not disrupt the student's engagement in an activity in order to ask questions, but instead diagnoses a student's strengths and weaknesses through observation. And once the diagnosis is made, the coach interjects advice and instruction selectively, in ways designed to maximize learning through discovery and experience. In that spirit, the WEST system attempts to infer the student's knowledge of the domain—in this case a computer game called "How the West was Won," designed to teach the use of basic arithmetic expressions—by observing the student's behavior.<sup>11</sup>

While the project of identifying a student's problems directly from his or her behavior proved considerably more difficult than expected, the objectives for the WEST coach were accomplished in the prototype system to an impressive degree. Because in the case of learning to play WEST the student's actions take the form of input to the computer (entries on a keyboard) and therefore leave an accessible trace, and a context for those actions (the current state of, and history of consecutive moves across, the "board") is defined by the system, each student turn can be compared against calculations of the move that a hypothetical expert player would make given the same conditions. Each expert move, in turn, requires a stipulated set of associated skills. Evidence that a particular skill is lacking, accumulated across some number of moves, identifies that skill as a candidate for coaching. The coach then interjects offers of advice to the student at opportune moments in the course of the play, where what constitutes an opportune moment for interjection is determined

10. The system, described in Chapter 6, was designed by Richard Fikes at the Xerox Palo Alto Research Center.

11. The student is presented with a graphic display of a game board made up of 70 squares (representing the Western frontier), a pair of icons (representing the two players — user and computer), and three spinners. A player's task in each turn is to combine the three numbers that the spinners provide, using the basic operations, to produce a value that becomes the number of spaces the icon is moved along the board. To add an element of strategy, squares on the board are more and less desirable — for example, "towns" occur every ten spaces, and landing on one advances you to the next. The object is to be the first player to land on exactly 70.

Early observation of students playing the game revealed that they were not gaining the full benefit of the arithmetic practice, in that they tended to settle on a method for combining numbers (for example, multiply the first two numbers and add the third), and to repeat that same method at each turn. Recognizing that this might reflect either a weakness in the student's proficiency at constructing expressions, a failure to grasp the strategy of the game, or both, Brown and Burton saw the potential usefulness of a "coach" that could guide the student to an expanded repertoire of skills and a better understanding of the domain. For a description of a similarly motivated "advisory" system for the programming language PROLOG, see Coombs and Alty, 1984.

according to a set of rules of thumb regarding good tutorial strategy (for example, always coach by offering the student an alternate move that both demonstrates the relevant skill and accomplishes obviously superior results; never coach on two turns in a row, no matter what, and so forth.)

### 2.3.2 *The computer as an artifact having purposes*

While the computer-based coach can be understood as a logical development in the longstanding problem of instruction, the requirement that it be interactive introduces a second sense of self-explanatory machine which is more recent, and is uniquely tied to the advent of computing. The new idea is that the intelligibility of artifacts could be not just a matter of the availability to the user of *the designer's* intentions for the artifact, but of the intentions of *the artifact* per se. That is to say, the designer's objective now is to imbue the machine with the grounds for behaving in ways that are accountably rational; i.e. reasonable or intelligible to others, including, in the case of interaction, ways that are responsive to the other's actions.

In 1950, A. M. Turing proposed a now-famous—and still controversial—test for machine intelligence based on a view of intelligence as accountable rationality. Turing argued that if a machine could be made to respond to questions in such a way that a person asking the questions could not distinguish between the machine and another human being, the machine would have to be described as intelligent.<sup>12</sup> Turing expressly dismissed the possible objection that although the machine might succeed in the game, it could succeed through means that bear no resemblance to human thought. Turing's contention was precisely that success at performing the game, regardless of mechanism, is sufficient evidence for intelligence (1950, p. 435). The Turing test thereby became the canonical form of the argument that if two information-processors, subject to the same input stimuli, produce indistinguishable output behavior then, regardless of the identity of their internal operations, one processor is essentially equivalent to the other.

The lines of the controversy raised by the Turing test were drawn over a family of programs developed by Joseph Weizenbaum in the 1960's under the name ELIZA.<sup>13</sup> and designed to support "natural language conversation" with a computer (1983, p.23).<sup>14</sup> Anecdotal reports of occasions on

12. To implement his test, Turing chose a game called the "imitation game." The game was initially conceived as a test of the ability of an interrogator to distinguish which of two respondents was a man and which a woman. To eliminate the evidence of "physical embodiment," the interaction was to be conducted remotely, via teleprinter. Thus Turing's notion that the game could easily be adapted to a test of machine intelligence, by substituting the machine for one of the two human respondents.

13. Of the name ELIZA, Weizenbaum writes "Its name was chosen to emphasize that it may be incrementally improved by its users, since its language abilities may be continually improved by a 'teacher'. Like the Eliza of Pygmalion fame, it can be made to appear even more civilized, the relation of appearance to reality, however, remaining in the domain of the playwright" (*Ibid*, p. 23).

14. "Natural language understanding" is a principal area of AI research. (See Chapter 3, section 2.) Interestingly, the idea that language ability is *the* mark of intelligence is found also in the notion of "competent member" of the society as used by the sociologists Garfinkel and Sacks: "We do not use the term ('member') to refer to a person. It refers instead to mastery of natural language" (1970, p. 342).

which people approached the teletype to one of the ELIZA programs and, believing it to be connected to a colleague, engaged in some amount of "interaction" without detecting the true nature of their respondent, led many to believe that Weizenbaum's program had passed a simple form of the Turing test. In contrast to Turing, however, Weizenbaum himself denied the intelligence of the program—not on the basis of its interactional success, but on the basis of the underlying mechanism—in a paper that discussed the program's reliance on "a mere collection of procedures" (*Ibid*, p. 23):

The gross procedure of the program is quite simple; the text [written by the human participant] is read and inspected for the presence of a *keyword*. If such a word is found, the sentence is transformed according to a *rule* associated with the keyword, if not a content-free remark or, under certain conditions, an earlier transformation is retrieved. The text so computed or retrieved is then printed out (*Ibid*, p. 24, original emphasis).

In spite of Weizenbaum's disclaimers with respect to their intelligence, the ELIZA programs are still cited as instances of successful interaction between human and machine. The grounds for their success are clearest in DOCTOR, one of the ELIZA programs whose 'script' equipped it to respond to the human user as if the computer were a Rogerian therapist and the user a patient. The DOCTOR program exploited the maxim that shared premises can remain unspoken, i.e. the less we say in conversation, the more what *is* said is assumed to be self-evident in its meaning and implications (cf. Coulter, 1979, chpt. 5). Conversely, the very fact that a comment is made without elaboration implies that such shared background assumptions exist. The more elaboration or justification is provided, the less the appearance of transparency or self-evidence. The less elaboration, the more the recipient will take it that the meaning of what is provided *should be* findable without problem or explanation.<sup>15</sup> In the case of DOCTOR, computer-generated responses that might otherwise seem odd were rationalized by users on the grounds that there must be some psychiatric intent behind them, not immediately obvious to the "patient," but sensible nonetheless:

If, for example, one were to tell a psychiatrist 'I went for a long boat ride' and he responded 'Tell me about boats', one would not assume that he knew nothing about boats, but that he had some purpose in so directing the subsequent conversation. It is important to note that this assumption is one made by the speaker. Whether it is realistic or not is an altogether different question. In any case, it has a crucial psychological utility in that it serves the speaker to maintain his sense of being heard and understood. The speaker further defends his impression (which even in real life may be illusory) by attributing to his conversational partner all sorts of background knowledge, insights and reasoning ability. But again, these are the *speaker's* contribution to the conversation. They manifest themselves inferentially in the *interpretations* he makes of the offered response (Weizenbaum, *Ibid*, p. 26).

---

15. Put another way, the design of the DOCTOR program exploited the natural inclination of people to deploy the "documentary method of interpretation" in finding the sense of actions that are in some way problematic, but which they assume to be purposeful or meaningful (Garfinkel, 1967, p.78). Very simply, the "documentary method" refers to the observation that people take appearances as evidence for, or the document of, an ascribed underlying reality, while taking the reality so ascribed as a resource for the interpretation of the appearance (see Chapter 4).

In explicating the ELIZA programs, Weizenbaum was primarily concerned with the inclination of human users to find sense in the computer's output, and to ascribe to it an understanding, and therefore an authority, unwarranted by the actual mechanism.<sup>16</sup> Weizenbaum was concerned with the problem of what counts as intelligence, while being little concerned with the problem of what counts as conversation. While unmasking the intelligence of his program, Weizenbaum continued to describe it as "a program which makes natural language conversation with a computer possible" (*Ibid.* p. 23). Nevertheless, as part of his disclaimer regarding its intelligence, Weizenbaum points to a crucial shortcoming in the ELIZA strategy with respect to conversation:

ELIZA in its use so far has had as one of its principal objectives the *concealment* of its lack of understanding. But to encourage its conversational partner to offer inputs from which it can select remedial information, it must *reveal* its misunderstanding. A switch of objectives from the concealment to the revelation of misunderstanding is seen as a precondition to making an ELIZA-like program the basis for an effective natural language man-machine communication system (*Ibid.* p. 27).

More recently, the inevitability of troubles in communication, and the importance of their remedy to the accomplishment of "graceful interaction," has been re-introduced into the human-machine communication effort by Hayes and Reddy (1983). They observe that:

During the course of a conversation, it is not uncommon for people to misunderstand or fail to understand each other. Such failures in communication do not usually cause the conversation to break down; rather, the participants are able to resolve the difficulty, usually by a short clarifying sub-dialogue, and continue with the conversation from where they left off. Current computer systems are unable to take part in such clarifying dialogues, or resolve communication difficulties in any other way. As a result, when such difficulties occur, a computer dialogue system is unable to keep up its end of the conversation, and a complete breakdown is likely to result; this fragility lies in stark and unfavourable contrast to the robustness of human dialogue (*ibid.*, p. 234).

Hayes and Reddy go on to recommend steps toward a remedy for the fragility of human-computer communication, based on the incorporation, from human communication, of conventions for the detection and repair of misunderstanding. They acknowledge, however, that their

---

16. In this regard it is interesting to note that a great debate ensued surrounding the status of the DOCTOR program as a psychotherapeutic tool. That debate took on a humorous tone when Weizenbaum submitted a letter to the *Forum* of the Association for Computing Machinery, an excerpt of which follows:

Below is a listing of a PL/1 program that causes a typewriter console to imitate the verbal behavior of an autistic patient. The 'doctor' types his interrogatories on the console. It responds *exactly* as does an autistic patient — that is, not at all. I have validated this model following the procedure first used in commercial advertising by Carter's Little Liver Pills ('Seven New York doctors say ...') and later used so brilliantly by Dr. K.M. Colby in his simulation of paranoia [a reference to Colby, K.M. et al. 'Turing-like indistinguishability tests for the validation of a computer simulation of paranoid processes.' *Artificial Intelligence 3* (1972), 199-221.]; I gave N psychiatrists access to my program and asked each to say from what mental disorder it suffered. M psychiatrists (M<N) said the (expletive deleted) program was autistic. (The methodological assumption here is that if two processes have identical input/output behaviors, then one constitutes an explanation of the other.)

The program has the advantage that it can be implemented on a plain typewriter not connected to a computer at all (1974, p. 543).

recommendations are unlikely to be sufficient for successful communication in other than the simplest encounters, e.g. automated directory assistance, or reservation systems. The question of why this should be so—of the nature of the limits on human-machine communication, and the nature and extent of robustness in human interaction—is the subject of the following chapters.

### 3. Plans\*

Once the European navigator has developed his operating plan and has available the appropriate technical resources, the implementation and monitoring of his navigation can be accomplished with a minimum of thought. He has simply to perform almost mechanically the steps dictated by his training and by his initial planning synthesis (Gladwin, 1964, p. 175).

To the extent that communication consists in the mutual intelligibility of our actions, any account of communication presupposes, whether explicitly or implicitly, an account of the coherence and intelligibility of action as such. This chapter and the next discuss two alternative views of action. The first, adopted by most researchers in artificial intelligence, locates the organization and significance of human action in underlying plans.<sup>17</sup> On this view, plans are prerequisite to and prescribe action, at whatever level of detail one might imagine. Mutual intelligibility is a matter of the reciprocal recognizability of our plans, due to common conventions for the expression of intent, and shared knowledge about typical situations and appropriate actions. The second approach, in contrast, argues that while the course of action can always be reconstructed in terms of prior intent, conventional rules and common knowledge, the prescriptive significance of intent, rules and knowledge for situated action is inherently vague. The coherence of situated action is tied in essential ways not to *a priori* prescriptions, but to the action's particular circumstances. A consequence of action's situatedness is that communication must incorporate both a sensitivity to its circumstances, and built-in resources for the remedy of troubles in understanding that inevitably arise.

This chapter reviews the planning model of purposeful action and shared understanding. As a basis for efforts to achieve mutual intelligibility between people and machines, the planning approach draws on three related conceptions: (i) the planning model itself, which takes the significance of action to be derived from plans, and identifies the problem for interaction as their recognition and coordination, (ii) speech act theory, which accounts for the recognition of plans or intentions by proposing conventional rules for their expression, and (iii) the idea of shared background knowledge, as the common resource that stands behind individual action and gives it social meaning. Each of these conceptions attempts to address general problems in human communication—the relation of observable behavior to intent, the correspondence of intended and interpreted meaning, and the stability of meaning assignments across situations—in ways that are relevant to particular problems in "human-machine interaction."

\* Note to the reader: This chapter is intended principally as background to the thesis, developed in Chapters 4 and 5, that action is essentially situated. The reader familiar with the Planning Model might either skip Chapter 3, or read Chapters 4 and 5 first and then return, with those Chapters in mind, to reconsider the familiar view.

17. It should be noted that this view of purposeful action is as old as the (at least Occidental) hills — it is the basis for traditional philosophies of rational action and behavioral sciences. It is hardly surprising, therefore, that it should be embraced by the fields concerned with "intelligent" artifacts, particularly Cognitive Science and Information-Processing Psychology (cf. Dreyfus, forthcoming).

### 3.1 The planning model

The planning model treats a plan as an attempt to prescribe the sequence of actions that will accomplish some preconceived end. The model posits that action is a form of problem solving, where the actor's problem is to find a path from some initial state to a desired goal state, given certain conditions along the way:<sup>18</sup>

In problem-solving systems, actions are described by prerequisites (i.e. what must be true to enable the action), effects (what must be true after the action has occurred), and decomposition (how the action is performed, which is typically a sequence of subactions) (Allen, 1984, p. 126).

Actions are described, at whatever level of detail, by their preconditions and their consequences, and every action can be located relative to a goal. Goals define the actor's relationship to the situation of action, the situation is just those conditions that obstruct or advance the actor's progress toward his or her goals. Advance planning is inversely related to prior knowledge of the environment of action, and of the conditions that the environment is likely to present. Unanticipated conditions will require re-planning. In every case, whether constructed entirely in advance, or completed and modified during the action's course, the plan is prerequisite to the action.

#### 3.1.1 *Plan generation and execution monitoring*

One of the earliest attempts to implement the planning model in a machine occurred as part of a project at Stanford Research Institute, beginning in the mid 1960's, to build a robot that could navigate autonomously through a series of rooms, avoiding obstacles and moving specified objects from one room to another. The robot, named by its designers Shakey, was controlled by a problem solving program called STRIPS (1971), which employed a means-ends analysis to determine the robot's path. The STRIPS program examined the stated goal, and then determined a subset of "operators," or actions available to the robot that would produce that state. The preconditions of those actions in turn identified particular subgoal states, which could be examined in the same way. The system thus worked backward from the goal until a plan was defined from the initial state to the goal state, made up of actions that the robot could perform. Subsequent work on problem solving and plan generation consisted in large part of refinements to the basic means-ends strategy, toward the end of achieving greater efficiency by constraining the search through possible solution paths.<sup>19</sup>

---

18. See Newell and Simon, 1972 for the seminal formulation of this view.

19. For a review of subsequent work, see Sacerdoti, 1977, chapter 3.

Beyond the analysis of ends and means, by which a plan is constructed, artificial intelligence researchers have also had to address problems of "failure and surprise" (Nilsson, 1973) in the execution of their planning programs, due to the practical exigencies of action in an unpredictable environment. The objective that Shakey should actually be able to move autonomously through a real (albeit somewhat impoverished) world added a new class of problems to those faced by mathematical or game-playing programs operating in an abstract, formal domain:

for a problem-solver in a formal domain is essentially done when it has constructed a plan for a solution; nothing can go wrong. A robot in the real world, however, must consider the execution of the plan as a major part of every task. Unexpected occurrences are not unusual, so that the use of sensory feedback and corrective action are crucial (Raphael, cited in McCorduck, 1979, p. 224).

In Shakey's case, execution of the plan generated by the STRIPS program was monitored by a program called PLANEX. The PLANEX program monitored not the actual moves of the robot, however, but the execution of the plan. The program simply assumed, in other words, that the execution of the plan meant that the robot had taken the corresponding action in the real world. The program also made the assumption that every time the robot moved there was some normally distributed margin of error, that would be added to a "model of the world," or representation of the robot's location. When the cumulative error in the representation got large enough, the plan monitor initiated another part of the program that triggered a camera which could, in turn, take a reading of Shakey's location in the actual world.

Another level of uncertainty was accommodated by execution-monitoring with Sacerdoti's system NOAH (an acronym for Nets of Action Hierarchies), designed in part to contribute to a larger project, the Computer-Based Consultant, also at the Stanford Research Institute:

NOAH is an integrated problem solving and execution monitoring system. Its major goal is to provide a framework for storing expertise about the actions of a particular task domain, and to impart that expertise to a human in the cooperative achievement of nontrivial tasks (1977, p. 2).

The output of Sacerdoti's planning program is a "procedural net," or hierarchy of partially ordered actions, which becomes in turn the input to the execution-monitoring portion of the system. The execution monitor takes the topmost action in the hierarchy, provides the user with an instruction, and then queries the user regarding the action's completion. A principal objective of the innovations that Sacerdoti introduced for the representation of procedures in NOAH was to extend execution monitoring to include tracking and assessment of *the user's* actions in response to the instructions generated:

The system will monitor the apprentice's work to ensure that the operation is proceeding normally. When the system becomes aware of an unexpected event, it will alter instructions to the apprentice to deal effectively with the new situation (ibid. p. 3).

A positive response from the user to the system's query regarding the action was taken to mean that the user understood the instruction, and had successfully carried it out, while a negative response was taken as a request for a more detailed instruction. The system allowed as well for a "motivation response," i.e. a query from the user as to why a certain task needed to be done, to which it responded by listing tasks to which the current task was related, and for an "error response," i.e. an indication from the user that the current instruction could not be carried out.

Just as the accumulation of error in the PLANEX program required feedback from the world in order to re-establish the robot's location, the error response from the user in Sacerdoti's system required that NOAH somehow repair its representation of the user's situation:

The PLANEX system faced this problem by continuously checking the plan's kernel against its world model, to ensure that the execution was on the right track. PLANEX presumed that an adequate mechanism existed for accurately updating the world model. This was almost the case, since there were only a small number of actions that the robot vehicle could take, and the model of each action contained information about the uncertainty it would introduce in the world model. When uncertainties reached a threshold, the vision subsystem was used to restore the accuracy of the world model.

For the domain of the Computer-based Consultant, or even for a richer robot domain, this approach will prove inadequate ... NOAH cannot treat the world model as a given. It must initiate interactions with the user at appropriate points to ensure that it is accurately monitoring the course of the execution ...

[W]hen a serious error is discovered (requiring the system to be more thorough in its efforts to determine the state of the world), the system must determine what portions of its world model differ from the actual situation (Sacerdoti 1977, p. 71-72).

### 3.1.2 *Interaction and plan recognition*

Adherents of the planning model in artificial intelligence research have taken observations like Sacerdoti's, to the effect that mutual intelligibility requires *interaction*, as an injunction to extend the planning model from a single individual to two or more individuals acting in concert. The situation in which Shakey moved consisted of walls and boxes (albeit boxes that could be moved unexpectedly by a human hand). In the real world, of course, the situations of our actions are constituted not only by physical constraints, whether natural or artificial in origin, but by the actions of others. Generally, the planning model attempts to bring the fact that we inhabit a profoundly social world under the jurisdiction of the individual actor, by attaching to the others in the actor's world sufficient description, and granting to the actor sufficient knowledge, that he or she is able to respond to the actions of others as just another set of environmental conditions. The problem of social interaction, consequently, becomes an extension of the problem of the individual actor. The basic view of a single, goal-directed agent, acting in response to an environment of conditions is complicated—the conditions now include the actions of other agents—but intact. The problem for interaction, on this view, is to recognize the actions of others as the expression of their underlying plans.

The complement to plan generation and execution in artificial intelligence research, therefore, is plan recognition, or the attribution of plans to others based on observation of their actions. The starting premise for a theory of plan recognition is that an observer takes some sequence of actions as 'input,' and then forms hypotheses about the plans that could motivate and explain those actions. One approach to the problem of action understanding in AI research has been an effort to characterize the relation between plans and observable behavior. Allen (1984) is concerned, for example, with the fact that there is an uncertain relationship between an action's intended effect, and the method by which that effect is achieved—a problem he illustrates with the example of turning on a light:

There are few physical activities that are a necessary part of performing the action of turning on a light. Depending on the context, vastly different patterns of behavior can be classified as the same action. For example, turning on a light usually involves flipping a light switch, but in some circumstances it may involve tightening the light bulb (in the basement) or hitting the wall (in an old house). Although we have knowledge about how the action can be performed, this does *not* define what the action is. The key defining characteristic of turning on the light seems to be that the agent is performing some activity which will cause the light, which was off when the action started, to become on when the action ends. An important side effect of this definition is that we could recognize an observed pattern of activity as 'turning on the light' even if we had never seen or thought about that pattern previously (*ibid.*, p. 126).

Allen's point is two-fold: First, the 'same' action as a matter of intended effect can be achieved in any number of ways, where the ways are contingent on circumstance rather than on definitional properties of the action. And second, while an action can be accounted for *post hoc* with reference to its intended effect, an action's course cannot be predicted from knowledge of the actor's prior intent, nor can the course be inferred from observation of the outcome. Allen identifies the indeterminate relationship of intended effect to method as a problem for planning or plan recognition systems: a problem that he attempts to resolve by constructing a logical language for action descriptions that handles the distinction between what he calls the "causal definition" of an action (i.e. the pre and post conditions that must hold in order to say that the action has occurred, independent of any method), and the action's characterization in terms of a particular method or procedure for its accomplishment.<sup>20</sup>

While Allen's approach to the problem of plan recognition is an attempt to reconstruct logically our vocabulary of purposeful action, a few more psychologically oriented researchers in artificial intelligence actually have undertaken experiments designed to reveal the process by which people bring the actions of others under the jurisdiction of an ascribed plan. Schmidt *et al* (1978) observe, for example, that plan attribution seems to require certain transformations of the real-time

---

20. Another distinction that Allen attempts to capture is one viewed as crucial to the success of the planning model, *viz* the observation that while some components of an action are sequentially ordered in a necessary way (i.e. one is prerequisite to the other), other components, while necessary to the action, have no necessary sequential relation to each other. The incorporation of the latter into the structure of plans was viewed as a substantial breakthrough in early planning research (see Sacerdoti, 1975).

organization of the action described.<sup>21</sup> They report that throughout the process of plan attribution the problem to be solved by the subject remains "ill-formed," by which they mean that at any given time neither the range of possible plans that the other might be carrying out, nor the criteria for assessing just what plan is actually in effect, are clearly defined (*ibid.*, p. 80). Nonetheless, they report that their subjects are able to posit an underlying plan. The strategy appears to be to adopt tentatively a single hypothesis about the other's plan, rather than entertain all or even some number of logical possibilities simultaneously. The preferred hypothesis regarding the other's plan then affects what actions are noted and recalled in the subject's accounts of the action, and the temporal order of events is restructured into logical 'in order to' or 'because' relationships, such that relations among actions are not restricted to consecutive events in time. At the same time, the current hypothesis is always subject to elaboration or revision in light of subsequent events, to the extent that subjects are often required to suspend judgment on a given hypothesis, and to adopt a "Wait and See" strategy. Wherever possible, actions that violate the structure of an attributed plan are "explained away" before the plan itself is reconsidered. Schmidt *et al* conclude that all of these observations "support the generalization that action understanding is simply a process of plan recognition" (*ibid.*, p. 50).

### 3.1.3 *The status of plans*

Assessment of the planning model is complicated by equivocation in the literature on planning between plans as a conceptual framework for the analysis and simulation of action, and plans as a psychological mechanism for its actual production. When researchers describe human action in terms of plans, the discussion generally fineses the question of just how the formulations provided by the researcher are purported to relate to the actor's intent. The claim is at least that people analyze *each other's* actions into goals and plans in order to understand each other. But the suggestion that the plan is "recognized" implies that it has an existence prior to and independent of the attribution: that it actually determines the action.

The identification of the plan with the actor's intent is explicit in the writing of philosophers of action supportive of artificial intelligence research, like Margaret Boden who writes:

Unless an intention is thought of as an action-plan that can draw upon background knowledge and utilize it in the guidance of behavior one cannot understand how intentions function in real life (1973, p. 28).

Intentions in other words, are realized as plans-for-action that directly guide behavior. A logical extension of Boden's view, particularly given an interest in rendering it more 'precise' (i.e. computable), is the view that plans *are* actually prescriptions or instructions for action. This view is

---

21. The empiricism of their study is unusual in artificial intelligence research, where work generally proceeds on the basis of imagination and introspection.

exemplified by Miller, Galanter and Pribram (1960) who define an intention as "the uncompleted parts of a Plan whose execution has already begun" (*ibid.*, p. 61). With respect to the plan itself:

Any complete description of behavior should be adequate to serve as a set of instructions, that is, it should have the characteristics of a plan that could guide the action described. When we speak of a plan ... the term will refer to a *hierarchy* of instructions ... *A plan is any hierarchical process in the organism that can control the order in which a sequence of operations is to be performed.*

A Plan is, for an organism, essentially the same as a program for a computer ... we regard a computer program that simulates certain features of an organism's behavior as a theory about the organismic Plan that generated the behavior.

Moreover, we shall also use the term "Plan" to designate a rough sketch of some course of action ... as well as the completely detailed specification of every detailed operation ... We shall say that a creature is executing a particular Plan when in fact that Plan is controlling the sequence of operations he is carrying out. (*ibid.*, p. 17, original emphasis).

With Miller *et al.*, the view that purposeful action is planned is reconstructed as a psychological 'theory,' compatible with the interest in a mechanistic, computationally tractable account of intelligent action. By improving upon or "completing" our common sense descriptions of the structure of action, the structure is now represented not only as a plausible sequence, but as an hierarchical plan. The plan reduces, moreover, to a detailed set of instructions that actually serves as the program that controls the action. At this point, the plan as stipulated becomes substitutable for the action, insofar as the action is viewed as derivative from the plan. And once this substitution is done, the theory is self-sustaining: the problem of action is assumed to be solved by the planning model, and the task that remains is the model's refinement.

While attribution of the plan to the actor resolves the question of the plan's status, however, it introduces new problems with respect to what we actually mean by "purposeful action." First, if any plan of action can be analyzed at any level of detail, what level of description represents that which we would want to call purposeful action? If every level, there is no reason in principle to distinguish, for example, between deliberate action and involuntary response, as the latter can always be ascribed to a process of planning unavailable to the actor.<sup>22</sup> Second, if plans are synonymous with purposeful action, how do we account, on the one hand, for a prior intent to act which may never be realized, and on the other, for an intentional action for which we would ordinarily say no plan was formed ahead of time.<sup>23</sup> Finally, while the planning model would have a statement of intent reflect an actual set of instructions for action, even casual observation indicates that our statements of intent generally do not address the question of situated action at any level of

22. In fact, this is just what Boden would have us do. On her account, action can be reduced to basic units for which "no further procedural analysis could conceivably be given." Those units compose "complex procedural schemata or action-plans", which in turn produce "complex intentional effects." Psychological processes at the level of intention, in other words, are reducible ultimately to bodily operations.

23. Davis (cited in Allen, 1984) gives the example of a person driving, who brakes when a small child runs in front of the car. See also Searle's distinction (1980) between "prior intentions" and "intentions-in-action."

detail. In fact, because the relation of the intent to accomplish some goal to the actual course of situated action is enormously contingent, a statement of intent generally says little about the action that follows. It is precisely because our plans are inherently vague—because we can state our intentions without having to describe the actual course that our actions will take—that an intentional vocabulary is so useful for our everyday affairs.

The confusion in the planning literature over the status of plans mirrors the fact that in our everyday action descriptions we do not normally distinguish between accounts of action provided before and after the fact, and action's actual course. As common sense constructs, plans are a constituent of practical action, but they are constituent as an *artifact* of our reasoning about action, not as the generative *mechanism* of action. Our imagined projections and our retrospective reconstructions are the principal means by which we catch hold of situated action and reason about it, while situated action itself, in contrast, is essentially transparent to us as actors.<sup>24</sup> The planning model, however, takes over our commonsense preoccupation with the anticipation of action, and the review of its outcomes, and attempts to systematize that reasoning as a model for action itself, while ignoring the actual stuff, the situated action, which is the reasoning's object.

### 3.2 Speech acts

A growing number of research efforts devoted to machine intelligence have their objective, for both theoretical and practical reasons, human-machine communication using English, or "natural language" (for example, Brady and Berwick 1983; Bruce 1981; Joshi *et al.* 1981). Researchers in "natural language understanding" have embraced Austin's observation (1962) that language is a form of action, as a way of subsuming communication to the planning model. If language is a form of action, it follows that language understanding, like the interpretation of action generally, involves an analysis of speaker's utterances in terms of the plans those utterances serve:

Let us start with an intuitive description of what we think occurs when one agent A asks a question of another agent B which B then answers. A has some *goal*; s/he creates a plan (*plan construction*) that involves asking B a question whose answer will provide some information needed in order to achieve the goal. A then executes this plan, asking B the question. B interprets the question, and attempts to infer A's plan (*plan inference*) (Allen 1983, p. 110).

As with the interpretation of action, plans are the substrate on which the interpretation of natural language utterances rests, insofar as "[h]uman language behavior is part of a coherent plan of action directed toward satisfying a speaker's goals" (Appelt nd., p. 1). We understand language, and action more generally, when we successfully infer the other's goals, and understand how the other's action furthers them. The appropriateness of a response turns on that analysis, from which, in turn:

---

24. One result of the transparency of situated action is that we have little vocabulary with which to talk about it, though Chapters 4 and 5 attempt to present some recent efforts in social science. For a treatment of the philosophical vocabulary proposed by Heidegger, see Dreyfus, forthcoming.

[t]he hearer then adopts new goals (e.g., to respond to a request, to clarify the previous speaker's utterance or goal), and plans his own utterances to achieve those. A conversation ensues. (Cohen nd, p. 24).

Given such an account of conversation, the research problem with respect to language understanding is essentially the same as that of the planning model more generally, that is, to characterize actions in terms of their preconditions and effects, and to formulate a set of inference rules for mapping between actions and underlying plans. Among researchers in the natural language area of artificial intelligence research, Searle's speech act theory (1969) is taken to offer some initial guidelines for computational models of communication:

We hypothesize that people maintain, as part of their models of the world, symbolic descriptions of the world models of other people. Our plan-based approach will regard speech acts as operators whose effects are primarily on the models that speakers and hearers maintain of each other (Cohen and Perrault 1979, p. 179).

Searle's conditions of satisfaction for the successful performance of speech acts are read as the speech act's "preconditions," while its illocutionary force is the desired "effect:"

[u]tterances are produced by actions (*speech acts*) that are executed in order to have some effect on the hearer. This effect typically involves modifying the hearer's beliefs or goals. A speech act, like any other action, may be observed by the hearer and may allow the hearer to infer what the speaker's plan is (Allen 1983, p. 108).

In describing utterances by their preconditions and effects, speech acts seem to provide at least the framework within which actual mechanisms for engineering interaction between people and machines might emerge. But while Searle's "conditions of satisfaction" state conventions governing the illocutionary force of certain classes of utterances, he argues against the possibility of a rule-based semantics for construing the significance of any particular utterance. While the maxims that speech act theory proposes—for example, the felicity condition for a directive is that S wants H to do A—tell us something about the general conditions of satisfaction for a directive, they tell us nothing further about the significance of any given directive *in situ*. With respect to the latter, Gumperz (1982b) offers the following example from an exchange between two secretaries in a small office:

- A: Are you going to be here for ten minutes?  
B: Go ahead and take your break. Take longer if you want.  
A: I'll just be outside on the porch. Call me if you need me.  
B: OK. Don't worry.  
(*ibid.*, p. x)

Gumperz points out that B's response to A's question clearly indicates that B interprets the question as an indirect request that B stay in the office while A takes a break, and by her reply, A confirms that interpretation. B's interpretation accords with a categorization of A's question as an indirect speech act (Searle, 1979), and with Grice's discussion of implicature (1975); i.e. B assumes that A is cooperating, and that her question must be relevant, therefore B searches her mind for some

possible context or interpretive frame that would make sense of the question, and comes up with the break. But, Gumperz points out, *this analysis begs the question of how B arrives at the right inference:*

What is it about the situation that leads her to think A is talking about taking a break? A common sociolinguistic procedure in such cases is to attempt to formulate discourse rules such as the following: 'If a secretary in an office around break time asks a co-worker a question seeking information about the co-worker's plans for the period usually allotted for breaks, interpret it as a request to take her break.' Such rules are difficult to formulate and in any case are neither sufficiently general to cover a wide enough range of situations nor specific enough to predict responses. An alternative approach is to consider the pragmatics of questioning and to argue that questioning is semantically related to requesting, and that there are a number of contexts in which questions can be interpreted as requests. While such semantic processes clearly channel conversational inference, there is nothing in this type of explanation that refers to taking a break (1982b, p. 326).

The problem that Gumperz identifies here clearly applies equally to attempts to account for inferences such as B's by arguing that she "recognizes" A's plan to take a break. Clearly she does: the outstanding question is how. While we can always construct a *post hoc* account that explains her interpretation in terms of knowledge of typical situations and motives, with speech act theory as with the planning model it remains the case that neither typifications of intent nor general rules for its expression and interpretation are sufficient to account for the mutual intelligibility of our situated action. In the final analysis, attempts to construct a taxonomy of intentions and rules for their recognition seem to beg the question of situated interpretation, rather than answering it.

### 3.3 Background knowledge

Gumperz' example demonstrates that a problem for any account of human action is that an action's significance seems to lie as much in what it presupposes and implies about its situation as in any explicit or observable behavior as such. Even the notion of observable behavior becomes problematic in this respect, insofar as what we do, and what we understand others to be doing, is so thoroughly informed by assumptions about the action's significance. In the interpretation of purposeful action, it is hard to know where the observation leaves off, and where the interpretation begins. In recognition of the fact that human behavior is a figure defined by its ground, behavioral science has largely turned from the observation of behavior to explication of the background that seems to lend behavior its sense.

For Cognitive Science, the background of action is not the world as such, but knowledge about the world, and researchers agree that representation of knowledge about the world is a principal limiting factor on progress in machine intelligence. The prevailing method in knowledge representation has been to categorize the world into domains of knowledge (e.g. areas of specialization like medicine, along one dimension; or propositions about physical phenomena like liquids, along another), and then to enumerate facts about the domain and relationships between

them. Having carved out domains of specialized knowledge, the catch-all for anything not clearly assignable is "common sense," which then can be spoken of as if it were yet another domain of knowledge, albeit one that is unbounded and indefinitely large. While substantial progress has been made in selected areas of specialized knowledge, however, the domain of common sense knowledge remains stubbornly unwieldy.

One approach to bounding common sense knowledge, exemplified by the work of Schank and Abelson (1977), is to classify the everyday world as types of situations, and assign to each its own body of specialized knowledge. The claim is that our knowledge of the everyday world is organized by a "predetermined, stereotyped sequence of actions that define a well-known situation" or script (*ibid.*, p. 422). Needless to say:

Scripts are extremely numerous. There is a restaurant script, a birthday party script, a football game script, a classroom script, and so on (*ibid.*, p. 423).

Every situation, in other words, has its plan made up of ordered action sequences, each action producing the conditions that enable the next action to occur. Of course the normative order of these action sequences can be thrown off-course—by any one of what Schank and Abelson term "distractions," "obstacles," or "errors." Distractions, about which they have little to say, comprise the interruption of one script by another, while:

[a]n obstacle to the normal sequence occurs when someone or something prevents a normal action from occurring or some enabling condition for the action is absent. An error occurs when the action is completed in an inappropriate manner, so that the normal consequences of the action do not come about (*ibid.*, p. 426).

Not only does the typical script proceed according to a normal sequence of actions, but each script has its typical obstacles and errors that, like the script itself, are stored in memory along with their remedies, and retrieved and applied as needed. So while plans associate intentions with action sequences, scripts associate action sequences with typical situations.

In practice, however, the stipulation of relevant background knowledge for typical situations always takes the form of a partial list, albeit one offered as if the author could complete the list, given the requisite time and space:

If one intends to buy bread, for instance, the knowledge of which bakers are open and which are shut on that day of the week will enter into the generation of one's plan of action in a definite way; one's knowledge of local topography (and perhaps of map-reading) will guide one's locomotion to the selected shop; one's knowledge of linguistic grammar and of the reciprocal roles of shopkeeper and customer will be needed to generate that part of the action-plan concerned with speaking to the baker, and one's financial competence will guide and monitor the exchange of coins over the shop counter (Boden 1973, p. 28).

Like Boden's informal story of the business of buying bread, attempts in artificial intelligence research to axiomatize common sense knowledge rely upon an appeal to intuition that shows little sign of yielding to scientific methods. The difficulty is not just that every action presupposes a

large quantity of background knowledge; though it would pose practical problems, such a difficulty would be tractable eventually. The problem is that just because 'implicit knowledge' can in principle be enumerated indefinitely, deciding in practice about the enumeration of background knowledge remains a stubbornly *ad hoc* procedure, for which researchers have not succeeded in constructing rules that do not depend, in their turn, on some deeper *ad hoc* procedures.

With respect to communication, the image evoked by "shared knowledge" is a potentially enumerable body of implicit assumptions, that stands behind every explicit action or utterance, and from which participants in interaction selectively draw in understanding each other's actions. This implies that what does actually get said on any occasion must reflect the application of a principle of communicative economy, which recommends roughly that to the extent that either the premises or rationale of an action can be assumed to be shared, they can be left unspoken. This means, in turn, that speakers must have procedures for deciding the extent of the listener's knowledge, and the commensurate requirements for explication. The listener, likewise, must make inferences regarding the speaker's assumptions about shared knowledge, on the basis of what he or she chooses explicitly to say. What is unspoken and relevant to what is said, in other words, is assumed to reside in the speaker's and listener's common stock of background knowledge, the existence of which is proven by the fact that an account of what is said always requires reference to further facts that, though unspoken, are clearly relevant.

This image of communication is challenged, however, by the results of an exercise assigned by Garfinkel to his students (1972). Garfinkel's aim was to press the common sense notion that background knowledge is a body of things thought but unsaid, that stands behind behavior and makes it intelligible. Students were asked to report a simple conversation by writing on the left hand side of a piece of paper what was said, and on the right hand side what it was that they and their partners actually understood was being talked about. Garfinkel reports that when he made the assignment:

many students asked how much I wanted them to write. As I progressively imposed accuracy, clarity, and distinctness, the task became increasingly laborious. Finally, when I required that they assume I would know what they had actually talked about only from reading literally what they wrote literally, they gave up with the complaint that the task was impossible (*ibid.*, p. 317).

The request was that the students provide a complete description of what was communicated, in one particular conversation, as a matter of the participants' shared knowledge. The students' dilemma was not simply that they were being asked to write "everything" that was said, where that consisted of some bounded, albeit vast, content. It was rather that the task of enumerating what was talked about itself extended what was talked about, providing a continually receding horizon of understandings to be accounted for. The assignment, it turned out, was not to describe some existing content, but to generate it. As such, it was an endless task. The students' failure suggests not that they gave up too soon, but that what they were assigned to do was not what the participants in the conversation themselves did in order to achieve shared understanding.

While the notion of 'background assumptions' connotes an actual collection of things that are 'there' in the mind of the speaker—a body of knowledge that motivates a particular action or linguistic expression, and makes it interpretable—there is reason to question the view that background assumptions are part of the actor's mental state prior to action:

As I dash out the door of my office, for example, I do not consciously entertain the belief that the floor continues on the other side, but if you stop me and ask me whether, when I charged confidently through the door, I believed that the floor continued on the other side, I would have to respond that indeed, I did (Dreyfus, 1982, p. 25).

A background assumption, in other words, is generated by the activity of accounting for an action, when the sense of the action is called into question, but there is no particular reason to believe that the assumption actually characterizes the actor's mental state prior to the act. In this respect, the 'world taken for granted' denotes not a mental state, but something outside of our heads that, precisely because it is non-problematically there, we do not need to think about. By the same token, in whatever ways we do find action to be problematical, the world is there to be consulted should we choose to do so. Similarly, we can assume the intelligibility of our actions, and as long as the others with whom we interact present no evidence of failing to understand us, we do not need to explain ourselves, yet the grounds and significance of our actions can be explicated endlessly. The situation of action is thus an inexhaustibly rich resource, and the enormous problems of specification that arise in cognitive science's theorizing about intelligible action have less to do with action, than with the project of substituting definite procedures for vague plans, and representations of the situation of action, for action's actual circumstances.



#### 4. Situated actions

This total process [of Trukese navigation] goes forward without reference to any explicit principles and without any planning, unless the intention to proceed to a particular island can be considered a plan. It is nonverbal and does not follow a coherent set of logical steps. As such it does not represent what we tend to value in our culture as 'intelligent' behavior (Gladwin 1964, p. 175).

This chapter turns to recent efforts, within anthropology and sociology, to challenge traditional assumptions regarding purposeful action and shared understanding. A point of departure for the challenge is the idea that common sense notions like plans are not faulted versions of scientific models of action, but rather are resources for people's practical deliberations about action. As projective and retrospective accounts of action, plans are themselves located in the larger context of some practical activity. As common sense notions about the structure of that activity, plans are part of the subject matter to be investigated in a study of purposeful action, not something to be improved upon, or transformed into proper scientific theories.

The premise that common sense accounts of action are properly part of the subject matter of social studies is due to a recent branch of sociology named *ethnomethodology*. This chapter describes the inversion of traditional social theory recommended by ethnomethodology, and the implications of that inversion for the problem of purposeful action and shared understanding. To designate the alternative that ethnomethodology suggests—more a reformulation of the problem of purposeful action, and a research programme, than an accomplished theory—I use the term *situated action*. That term underscores the fact that the course of action depends in essential ways upon the action's circumstances. Rather than attempting to abstract action from its circumstances and reconstruct it as a rational plan, the approach is to study how people use their circumstances to achieve intelligent action. Rather than build a theory of action out of a theory of plans, the aim is to investigate how people produce and find evidence for plans in the course of situated action. More generally, rather than subsume the details of action under the study of plans, plans are subsumed by the larger problem of situated action.

The view of action that ethnomethodology recommends is neither behavioristic in any narrow sense of that term, nor mentalistic. It is not behavioristic in that it assumes that the significance of action is not reducible to uninterpreted physical movements. Nor is it mentalistic, however, in that the significance of action is taken to be based, in ways that are fundamental rather than secondary or epiphenomenal, in the social and the material world. The basic premise is twofold; first, that what traditional behavioral sciences take to be cognitive phenomena have an essential relationship to a publicly available, collaboratively organized world of artifacts and actions, and second, that the significance of artifacts and actions, and the methods by which their significance is conveyed, have an essential relationship to their particular, concrete circumstances.

A view of purposeful action and shared understanding as situated is outlined in this chapter under five propositions: (i) plans are derivative from actions *in situ*; (ii) in the course of situated action, deliberation arises when otherwise transparent activity becomes in some way problematic;

(iii) the practical objectivity of the situations of our action is achieved rather than given; (iv) a central resource for achieving the objectivity of situations is language, which stands in a generally indexical relationship to the circumstances that it presupposes, produces and describes; (v) as a consequence of the indexicality of language, mutual intelligibility is achieved on each occasion of interaction, through recourse to situation particulars, rather than being discharged once and for all by a stable body of shared meanings.

#### 4.1 Plans are derivative from actions

The pragmatist philosopher and social psychologist George Herbert Mead argues (1934) that directed action is best viewed as two integrally, but problematically, related kinds of activity. One kind of activity is an essentially situated and *ad hoc* improvisation—the part of us, so to speak, that actually acts. The other kind of activity is derivative from the first, and includes our representations of action in the form of future plans and retrospective accounts. Plans and accounts are distinguished from action *per se* by the fact that to represent our actions, we must in some way make an object of them. Consequently, our descriptions of our actions as purposeful come always before or after the fact, in the form of envisioned projections, and recollected reconstructions.

Mead's treatment of the relation of deliberation and reflection to action is one of the more controversial, and in some ways incoherent, pieces of his theory. But his premise of a disjunction between our actions and our grasp of them at least raises the question for social science of the relationship between projected or reconstructed courses of action, and actions *in situ*. Most accounts of directed action have taken this relationship to be a simple causal one, at least in a logical sense (see Chapter 3). That is to say, given a desired outcome, the actor is assumed to make a choice among alternative courses of action, based upon the anticipated consequences of each with respect to that outcome. Accounts of actions taken, by the same token, are just a report on the choices made. The student of purposeful action on this view need know only the predisposition of the actor and the alternative courses that are available in order to predict the action's course. The action's course is just the playing out of these antecedent factors, knowable in advance of, and standing in a determinate relationship to, the action itself.

Such an account may appear, at first blush, to hold reasonably well in the case of certain deliberative, instrumental tasks. But surely we want to recognize as purposeful action more than just such tasks and, as surely, we want a single account that will hold across the range of purposeful actions. On the one hand, then, we have embodied skills like walking across the room, or driving a car (cf. Dreyfus and Dreyfus, forthcoming) which, while they might well be characterized as instrumental—in both cases we act in order to go from point A to point B—are clearly not deliberative. Skilled activities like driving proceed in a way that is only derivatively and summarily characterizable in terms of procedures or rules—and such rules as do get formulated are only used when the activity needs for some reason to be explicated, as for instruction, or at times of

breakdown when otherwise transparent ways of proceeding need to be inspected or revised (see 4.2 below). On the other hand, there are other mundane activities—going to work, for example—that, while more deliberate perhaps than an embodied practice like driving are nonetheless, like driving, so contingent on unique circumstances on any given day that we generally don't anticipate alternative courses of action, or their consequences, until *some* course of action is already underway. It is frequently only on acting in a present situation that its possible future states become clear. And we often do not know ahead of time, or at least not with any specificity, what future state we even desire to bring about; only after we encounter some state of affairs that we find to be desirable do we identify that state as the goal toward which our previous actions, in retrospect, were directed "all along" or "after all" (Garfinkel 1967, pp. 98-99). Of course, we can always perform a *post hoc* analysis of situated action that will make it appear to have followed a rational plan. But that fact says more about the nature of our analyses than it does about our situated actions. To return to Mead's point, rather than direct situated action, rationality anticipates action before the fact, and reconstructs it afterwards.

#### 4.2 Deliberation and the Unready-to-hand

While we can always construct rational accounts of situated action before and after the fact, when action is proceeding smoothly it is essentially transparent to us. Similarly, when we use what Heidegger terms equipment that is "ready-to-hand," the equipment "has a tendency to 'disappear':"

Consider the example (used by Wittgenstein and Merleau-Ponty) of the blind man's cane. We can hand the man the cane and ask him to tell us what properties it has. After hefting and feeling it, he can tell us that it is light, smooth, about three feet long, and so on; it is present-at-hand for him. But when the man starts to use the cane (when he grasps it in that special mode of understanding that Heidegger calls 'manipulation') he loses his awareness of the cane itself; he is aware only of the curb (or whatever object the cane touches); or, if all is going well, he is not even aware of that. Thus it is that equipment that is ready-to-hand is invisible just when it is most genuinely appropriated (Dreyfus forthcoming, chpt. 6, p. 8).

In contrast, the "unready-to-hand," in Heidegger's phrase, comprises occasions wherein equipment that is involved in some practical activity becomes unwieldy, temporarily broken, or unavailable. At such times, inspection and practical problem solving occurs, aimed at repairing or eliminating the disturbance in order to "get going again." In such times of disturbance, our use of equipment becomes "explicitly manifest as a goal-oriented activity," and we may then try to formulate procedures or rules:

In deliberation one stops and considers what is going on and plans what to do, all in a context of involved activity. Here one finds the sort of reasoning the tradition formalized in the practical syllogism;

The scheme peculiar to [deliberating] is the 'if-then'; if this or that, for instance, is to be produced, put to use, or averted, then some ways and means, circumstances, or opportunities will be needed (Heidegger, *Being and Time*, p. 410) (*ibid.*, chpt. 6, p. 15)

Another kind of breakdown in the "ready-to-hand," that arises when equipment to be used is unfamiliar, is discussed in Chapter 6 in relation to the 'expert help system' and the problem of instructing the novice user of a machine. The important point here is just that the rules and procedures that come into play when we deal with the "unready-to-hand" are not self-contained or foundational, but contingent on and derivative from the situated action that the rules and procedures represent. The representations involved in managing problems in the use of equipment *presuppose* the very transparent practices that the problem renders noticeable or remarkable. Situated action, in other words, is not made explicit by rules and procedures. Rather, when situated action becomes in some way problematic, rules and procedures are explicated for purposes of deliberation and the action, which is otherwise neither rule-based nor procedural, is then made accountable to them.

#### 4.3 The practical objectivity of situations

If we look at the world commonsensically, the environment of our actions is a succession of situations that we 'walk in to,' and to which we respond. Communication, correspondingly, is an exchange of information, based on our observations, about the facts of the situations in which we find ourselves. Whatever agreement we have about what is factual, on this realist view, comes simply from our common ability to see things as they are. As I noted in Chapter 3, advocates of the planning model of purposeful action are interested not only in reifying this common sense account of the situation for the individual actor, but in bringing concerted action under the same account by treating the actions of others as just so many more conditions of the actor's environment.

In a move quite compatible with the interests of the planning model, the traditional sociological approach has been to posit, and then attempt to describe, an objective world of social facts, or received norms, to which our attitudes and actions are a response. Recognizing the human environment to be constituted crucially by others, traditional sociology identifies a further set of environmental conditions, beyond the material, to which human behavior is responsive, i.e. the sanctions of institutionalized group life. Emile Durkheim's famous maxim that "the objective reality of social facts is sociology's fundamental principle" (1938) has been the methodological premise of social studies since early in this century. Human action, the argument goes, cannot be adequately explained without reference to these 'social facts,' which are to be treated as antecedent, external and coercive *vis a vis* the individual actor.

In 1954, the sociologist Herbert Blumer published a critique of traditional sociology titled "What Is Wrong with Social Theory?" (see Blumer 1969, pp. 140-152). Blumer's critique argues that the social world is constituted by the local production of meaningful action, and that as such the social world has never been taken seriously by social scientists. Instead, Blumer says, investigations done by social scientists have looked at meaningful action as the playing out of various determining factors, all antecedent and external to the action itself. Whether those factors

are brought to the occasion in the form of individual predispositions, or are present in the situation as pre-existing environmental conditions or received social norms, the action itself is treated as epiphenomenal. As a consequence, Blumer argues, we have a social science that is *about* meaningful human action, but not a science *of* it.

For the foundations of a science of action, Blumer turns to Mead. Mead offers a metaphysics of action that is deeply sociological—Blumer points out that a central contribution of Mead's work is his challenge to traditional assumptions regarding the origins of the common sense world, and of purposeful action;

His treatment took the form of showing that human group life was the essential condition for the emergence of consciousness, the mind, a world of objects, human beings as organisms possessing selves, and human conduct in the form of constructed acts. He reversed the traditional assumptions underlying philosophical, psychological, and sociological thought to the effect that human beings possess minds and consciousness as original 'givens,' that they live in worlds of pre-existing and self-constituted objects, and that group life consists of the association of such reacting human organisms (*ibid.*, p. 61).

Mead's "reversal," in putting human interaction before the objectivity of the common sense world, should not be read as an argument either for idealism or relativism; Mead does not deny the existence of intrinsic properties of the environment in which we act. What Mead is working toward is not a characterization of the natural world *simpliciter*, but of the natural world *under interpretation*, or the world as construed by us through language. The latter is precisely what we mean by the *social* world and, on Mead's view, interaction is a condition for that world, while that world is a condition for intentional action.

By adopting Durkheim's maxim, and assuming the individual's responsiveness to received social facts, social scientists hoped to gain respectability under the view that human responses to the facts of the social world should be studyable by the same methods as are appropriate to studies of other organisms reacting to the natural world. A principal aim of normative sociology was to shift the focus of attention in studies of human behavior from the psychology of the individual to the conventions of the social group. But at the same time that normative sociology directs attention to the community or group, it maintains an image of the individual member that is rooted in behaviorist psychology and natural science—an image that has been dubbed by Garfinkel the "cultural dope."

By 'cultural dope' I refer to the man-in-the-sociologist's-society who produces the stable features of the society by acting in compliance with preestablished and legitimate alternatives of action that the common culture provides (1967, p. 68).

Insofar as the alternatives of action that the culture provides are seen to be non-problematic and constraining on the individual, *their* enumeration is taken to constitute an account of the situation of human action. The social facts—i.e. what actions typically come to—are used as a point of departure for retrospective theorizing about the "necessary character of the pathways whereby the end result is assembled" (*ibid.*, p. 68).

More recently, ethnomethodology has turned Durkheim's maxim on its head, with more profound theoretical and methodological consequences. Briefly, the view of ethnomethodology is that what traditional sociology captures is precisely our *common sense* view of the social world (see Sacks 1963; Garfinkel 1967; and Garfinkel and Sacks 1970). Following Durkheim, the argument goes, social studies have taken over common sense realism, and attempted to build a science of the social world by improving upon it. Social scientific theories, under this attempt, are considered to be scientific insofar as they remedy problems in, and preferably quantify, the intuitions of everyday, practical sociological reasoning.

In contrast, ethnomethodology grants common sense sociological reasoning a fundamentally different status than that of a defective approximation of an adequate scientific theory. Rather than being *resources* for social science to improve upon, the 'all things being equal' typifications of practical reasoning are to be taken as social science's *topic*. The notion that we act in response to an objectively given social world is replaced, as a methodological premise, by the assumption that our everyday social practices render the world publicly available and mutually intelligible.<sup>26</sup> The source of mutual intelligibility is not a received conceptual scheme, or a set of coercive rules or norms, but those common foundational practices that produce the typifications of which schemes and rules are made. The task of social studies then is to describe the practices, not to enumerate their product in the form of a catalogue of common sense beliefs about the social world. Ethnomethodology's interest, in other words, is in how it is that the mutual intelligibility and objectivity of the social world is achieved. It locates that achievement, moreover, in our interactions. So our common sense of the social world is not the precondition for our interaction, but its product. Similarly, the objective reality of social facts is not the fundamental *principle* of social studies, but social studies' fundamental *phenomenon*.

#### 4.4 The general indexicality of language

Our shared understanding of situations is due in great measure to language, "the typifying medium *par excellence*" (Schutz 1962, p. 14). The efficiency of language is due to the fact that, on the one hand, expressions have assigned to them conventional meanings, which hold on any occasion of their use, and on the other hand, the significance of a linguistic expression on some actual occasion lies in its relationship to circumstances that are indicated by, but not actually

---

26. Thus the interest in *ethnomethods*. It should be clear from this discussion that the methodology of interest to ethnomethodologists is not their own, but that deployed by members of the society in coming to know, and making sense out of, the everyday world. The outstanding question for social science under this view is not whether social facts are objectively grounded, but through what methods. The sense of methods here is not as a matter of techniques for the analyst, but as the deeply interactive relationship between, practical actions, communicative practices, shared understandings, and publicly available situations.

captured in, the expression itself. Language takes its significance from the embedding world, in other words, even while it transforms the world into something that can be thought of and talked about.

Expressions that rely upon their situation for significance are commonly called *indexical*, after the 'indices' of Charles Peirce (Burks 1949), the exemplary indexicals being first and second person pronouns, tense, and specific time and place adverbs like "here" and "now." In the strict sense exemplified by these commonly recognized indexical expressions, the distinction of conventional or literal meaning, and situated significance, breaks down. That is to say, these expressions are distinguished by the fact that while one can state formal procedures for finding the expression's meaning, or rules for its use, the expression's meaning can only be specified with reference to some actual circumstances (cf. Bates 1976, chpt. 1).

Among philosophers and linguists, the term "indexicality" typically is used to distinguish those classes of expressions whose meaning is conditional on the situation of their use in this way from those such as, for example, noun phrases that refer to a class of objects, whose meaning is claimed to be specifiable in objective, or context-independent terms. But the *communicative* significance of a linguistic expression is always contingent on the circumstances of its use. A statement not of what the language means in relation to any context, but of what the language-user means in relation to some particular context, would require a description of the context or situation of the utterance itself. And every utterance's situation comprises an indefinite range of possibly relevant features. Our practical solution to this "problem" is not to enumerate some subset of the relevant circumstances—we generally never mention our circumstances as such at all—but to 'wave our hand' at the situation or to gloss over it, as if we always included in our utterance an implicit *ceterus paribus* clause, and closed with an implicit *etcetera* clause. One consequence of this practice is that we always "mean more than we can say in just so many words":<sup>27</sup>

[S]peakers can ... do the immense work that they do with natural language, even though over the course of their talk it is not known and is never, not even 'in the end,' available for saying in so many words just what they are talking about. Emphatically, that does not mean that speakers do not know what they are talking about, *but instead that they know what they are talking about in that way* ... (Garfinkel and Sacks 1970, p. 342-4).

A central tenet of ethnomethodology, in other words, is that the significance of an expression always exceeds the meaning of what actually gets said, and the interpretation of an expression, accordingly, always turns not only on its literal or dictionary meaning, nor on that plus some body of unspoken assumptions, but on the situation of its use.<sup>27</sup> In this sense deictic expressions, time and place

---

27. By situation of use, I mean either the circumstances of an expression's production or of its interpretation. Of course, for any expression, including those traditionally termed indexical, we can find a meaning for the expression with reference to an imagined scenario. But it is equally the case that for a so-called non-indexical expression, such as definite description, reference to the circumstances of the description's use is required to understand its *communicative* significance, i.e. what about the world is this expression meant to convey here and now. In any event, I want to

adverbs, and pronouns are just particularly clear illustrations of the general fact that all situated language, including the most abstract or universal, stands in an essentially indexical relationship to the embedding world. The relation of efficient linguistic formulations to particular situations parallels the relation of plans—which I take to be essentially prospective and retrospective formulations of action—to situated action. As formulations of action, plans are subject to the constraint that:

[h]owever extensive or explicit what a speaker says may be, it does not by its extensiveness or explicitness pose a task of deciding the correspondence between what he says and what he means that is resolved by citing his talk verbatim (Garfinkel and Sacks 1970, p. 342-4).

The problem of communicating instructions for action, in particular certain of its seemingly intractable difficulties, becomes clearer with this view of language in mind.<sup>28</sup> Like all action descriptions, instructions rely upon an implicit 'etcetera' clause in order to be called complete, and the irremediable incompleteness of action descriptions means that the significance of an instruction with respect to situated action does not inhere in the instruction, but must be found by the instruction-follower. Far from replacing the *ad hoc* methods used to establish the significance of everyday talk and action, therefore, the interpretation of instructions is thoroughly reliant on those same methods:

To treat instructions as though ad hoc features in their use was a nuisance, or to treat their presence as grounds for complaining about the incompleteness of instructions, is very much like complaining that if the walls of a building were gotten out of the way, one could see better what was keeping the roof up (Garfinkel 1967, p. 22).

The project of instruction-writing is ill-conceived, in other words, if its goal is the production of exhaustive action descriptions, that can guarantee a particular interpretation. What 'keeps the roof up' in the case of instructions for action is not only the instructions *per se*, but their interpretation *in situ*. And the latter has all of the properties of *ad hoc*ery and uncertainty that characterizes every occasion of the situated use of language.

Our situated use of language, in sum, and consequently language's significance, presupposes and implies an horizon of things that are never actually mentioned—what Schutz referred to as the "world taken for granted" (1962, p. 74). Philosophers have been preoccupied with this fact about language as a matter of the truth conditionality of propositions, the problem being that the truth conditions of an assertion are always relative to a background, and the background does not form part of the semantic content of the sentence as such (Searle 1979). The same problems that have

---

distinguish hypothetical scenarios, or idealized descriptions of situations from the actual situations in which language use (including the construction of hypothetical scenarios and idealized descriptions) invariably occurs.

28. This problem is the topic of chapters 6 and 7.

plagued philosophers of language as a matter of principle are now practical problems for Cognitive Science. As I pointed out in Chapter 3, the view that mutual intelligibility rests on a stock of shared assumptions has been taken over by researchers in Cognitive Science, in the hope that an enumeration of the knowledge assumed by particular words or actions could be implemented as data structures in the machine, which would then 'understand' those words and actions. Actual attempts to include the background assumptions of a statement as part of its semantic content, however, run up against the fact that there is no fixed set of assumptions that underlies a given statement. As a consequence, the elaboration of background assumptions is fundamentally *ad hoc* and arbitrary, and each elaboration of assumptions introduces further assumptions to be elaborated, *ad infinitum*.

#### 4.5 The mutual intelligibility of action

To account for the foundations of mutual intelligibility and social order, traditional social science posits a system of known-in-common behavioral conventions or 'norms.' What we share, on this view, is agreement on the appropriate relation of actions to situations. We walk into a situation, in other words, identify its features, and match our actions to it. On any given occasion, then, the concrete situation must be recognizable as an instance of a class of typical situations, and the behavior of the actor must be recognizable as an instance of a class of appropriate actions. With respect to communication, this implies that:

the different participants must define situations and actions in essentially the same way, since otherwise rules could not operate to produce coherent interaction over time. Within the normative paradigm, this cognitive agreement is provided by the assumption that the actors share a system of culturally-established symbols and meanings. Disparate definitions of situations and actions do occur, of course, but these are handled as conflicting subcultural traditions or idiosyncratic deviations from the culturally established cognitive consensus (Wilson 1970, p. 699).

In contrast, Garfinkel proposes that the stability of the social world is not the consequence of a "cognitive consensus," or measured amount of shared agreement, but of our common application of a very general method—the documentary method of interpretation—in unique and specific circumstances. Through that method, appearances (e.g. utterances, actions) are taken as evidence or "documents" of a presupposed underlying pattern (e.g. intent), which in turn fills in the sense of the appearance (1967, chpt. 3). The documentary method describes an ability—the ascription of intent on the basis of evidence, and the interpretation of evidence on the basis of ascribed intent—that is as identifying of rationality as the ability to act rationally itself. At the same time, the documentary method is not reducible to necessary and sufficient conditions, either behavioral or contextual, for the identification of intent:

[T]here is no way, logically, to recognise or describe any given case as a case of some

intelligent performance ... independently of the context in which an agent is engaged in a line of conduct, and there are no recognition algorithms for contextual particulars conjoined to behavioural descriptions such that any given form of 'cognitive conduct' might be precisely defined over an explicit set of (necessary and sufficient) observational data (Coulter 1983, p. 162-3).

Given the lack of universal rules for the interpretation of action, the programme of ethnomethodology is to investigate and describe the use of the documentary method in actual situations. Studies indicate, on the one hand, the generality of the method and, on the other, the extent to which special constraints on its use characterize specialized domains of practical activity; e.g. natural science, courts of law, the practice of medicine (see Chapter 5). By 'indice', Peirce meant not only that the sign relies for its significance on the event or object that it designates, but also that the sign is actually a constituent of the referent. Indexical—i.e. situated—language generally is not only anchored in, but in large measure constitutes, the situation of its use. Ethnomethodology generalizes this constitutive function of language to action and communication, in the proposition that the purposefulness of action is recognizable in virtue of the methodic, skillful and therefore taken-for-granted practices whereby we render an action accountably rational:

In unknown ways, that accomplishment is for members [of the society] a commonplace phenomenon. And in the ways that the accomplishment is commonplace it is for our interests an awesome phenomenon, for in its unknown ways it consists (1) of members' use of concerted everyday activities as methods with which to recognize and demonstrate the isolatable, typical, uniform, potential repetition, connected appearance, consistency, equivalence, substitutability, directionality, anonymously describable, planful—in short, the rational properties of ... expressions and ... actions. (2) The phenomenon consists, too, of the analyzability of actions-in-context given that not only does no concept of context-in-general exist, but every use of "context" without exception is itself essentially indexical (Garfinkel 1967, p. 10).

In positing the reflexivity of purposeful action and the methods by which we convey and construe action's purposes, ethnomethodology does not intend to reduce meaningful action to method. The intent is rather to identify the mutual intelligibility of action as *the* problem of social order. With that problem in mind, ethnomethodological studies investigate situated action and practical reasoning about action in general, and in particular domains.<sup>29</sup>

In a contrived situation that, though designed independently and not with them in mind, closely parallels both the "Turing test" and encounters with Weizenbaum's ELIZA programs (see Chapter 2 for a discussion of both), Garfinkel set out to test the documentary method for finding an underlying intent in observable behavior. Students were asked to direct questions concerning

---

29. For example, the work of coroners at the Los Angeles Suicide Prevention Center (Garfinkel 1967, pp. 11-18), the deliberations of juries (*ibid.*, chpt. 4) and courtroom practices of attorneys (Atkinson and Drew 1979), the work of clinic staff in selecting patients for out-patient psychiatric treatment (Garfinkel 1967, chpt. 7), the work of physicians interviewing patients for purposes of diagnosis (Beckman and Frankel 1983), the work of scientists discovering an optical pulsar (Garfinkel, *et al.* 1981).

their personal problems to a person whom they knew to be a student counsellor, seated in another room. They were restricted to questions that could take yes/no answers, and the answers were then given by the counsellor on a random basis. For the students, the counsellor's answers were motivated by the questions. That is to say, by taking each answer as evidence for 'what the counsellor had in mind,' the students were able to find a deliberate pattern in the exchange that explicated the significance and relevance of each new response as an answer to their question. Specifically, the yes/no utterances were found to document 'advice' from the counsellor, intended to help in the solution of the student's problem. For example, students assigned to the counsellor, as the advice 'behind' the answer, the thought formulated in the student's question:

when a subject asked 'Should I come to school every night after supper to do my studying?' and the experimenter said 'My answer is no,' the subject in his comments said, 'He said I shouldn't come to school and study' (Garfinkel 1967, p. 92).

In cases where an answer seemed directly to contradict what had come before, students either attributed the apparent contradiction to a change of mind on the part of the counsellor, as the result of having learned more between the two replies, or to some agenda on the part of the counsellor that lent the reply a deeper significance than its first, apparently inconsistent interpretation would suggest. In other cases, the interpretation of previous answers was revised in light of the current one, or an interpretation of the question was found, and attributed to the counsellor, that rationalized what would otherwise appear to be an inappropriate answer. Generally:

The underlying pattern was elaborated and compounded over the series of exchanges and was accommodated to each present 'answer' so as to maintain the 'course of advice,' to elaborate what had 'really been advised' previously, and to motivate the new possibilities as emerging features of the problem (*ibid.*, p. 90).

Garfinkel's results with arbitrary responses make the success of Weizenbaum's DOCTOR program easier to understand, and lend support to Weizenbaum's hypothesis that the intelligence of 'interactions' with the DOCTOR program is due to the work of the human participant—specifically, to methods for interpreting the system's behavior as evidence for some underlying intent—rather than to intent located in the machine. The larger implications of the documentary method, however, touch on the status of an 'underlying' reality of psychological and social fact in general, prior to situated action and interpretation:

It is not unusual for professional sociologists to think of their ... procedures as processes of 'seeing through' appearances to an underlying reality; of brushing past actual appearances to 'grasp the invariant.' Where our subjects are concerned, their processes are not appropriately imagined as 'seeing through,' but consist instead of coming to terms with a situation in which factual knowledge of social structures—factual in the sense of warranted grounds of further inferences and actions—must be assembled and made available for potential use despite the fact that the situations it purports to describe are, in any calculable sense, unknown; in their actual and intended logical structures are essentially vague; and are modified, elaborated, extended, if not indeed created, by the fact and matter of being addressed. (Garfinkel 1967, p. 96).

#### 4.6 Plans as efficient formulations

I pointed out earlier (Chapter 3) that, with some equivocation, AI researchers take plans as models for action. The weak form of such a claim, if we adopt a loose definition of the term "model," is just that plans are to actions as something that represents or stands for something else, in order that projective or imaginative operations can be performed. Strictly speaking, however, a model is a representation that is actually *homologous* to, or structurally like, the thing that it represents. A model of the DNA molecule, for example, actually posits that the structure of the model captures the constituent parts of the molecule and their relationships. Similarly, under a strict definition, a model of action as the enactment of a plan claims that the structure of the plan captures the constituent parts of action and their relationships—as an hierarchical arrangement of goals, preconditions and consequences.

While the plans constructed by Cognitive Science may in this strict sense model our vocabulary, or idealizations of action, the relationship between idealizations like plans and *situated* action is less like a model than it is like a map. That is to say, plans share with models the function of supporting projections and reconstructions of action. But rather than abstracting action in the strict sense of a constructing a homologue of the action's structure, plans are a simplification or sketch of action. Like maps, and like linguistic formulations of action generally, the utility of a plan rests on a particular kind of relationship, constructed at the time of its use. Specifically, the usefulness of a plan requires that the actor construct a correspondence between the plan, and his or her actions under some actual circumstances.

The mapmaker assumes that in using the map to navigate the terrain, the traveler will have recourse not only to the map, but to the terrain itself. For that reason, the map does not have to represent all of the terrain's relevant detail. Similarly, the efficiency of a plan rests not just on the plan's description of preconditions, actions and consequences, but on our ability to look to our circumstances in order to see on any particular occasion, literally or figuratively, what is being talked about. While the map's usefulness turns on the ability of the traveler to exploit the systematic relationship of map to terrain, one can not, just by looking at the map, reconstruct that relationship. Given that fact, plans and goals do not provide the solution to the question of what determines rational action, they simply re-state the problem. To characterize purposeful action as in accord with plans and goals is just to say again that it is purposeful and that *somehow*, in a way not addressed by the characterization itself, we constrain and direct our actions according to the significance that we assign to a particular context. How we do that is the outstanding problem. The dependency of significance on a particular context, every particular context's open-endedness, and the essential *ad hocness* of contextual elaboration is *sine qua non* for practical affairs, but a perplexity for a science of human action. And, to anticipate the analysis of Chapter 7, it is an intractable problem for projects that rest on providing in advance for the significance of canonical descriptions—like instructions—for situated action.

## 5. Communicative Resources

Thus the whole framework of conversational constraints ... can become something to honor, to invert, or to disregard, depending as the mood strikes. (Erving Goffman 1975, p. 311).

Communicative action occurs in particular moments of actual time, in particular relationships of simultaneity and sequence. These relationships in time, taken together, constitute a regular rhythmic pattern. This regularity in time and timing seems to play an essential, constitutive role in the social organization of interaction ... Whereas there is no metronome playing while people talk, their talking itself serves as a metronome (Erickson and Shultz 1982, p. 72).

We are environments for each other (Ray McDermott 1976, p. 27).

The argument of the preceding chapter was that we never definitively determine *the* intent behind an action, in that descriptions at the level of intent just are not designed to pick out relations of strict causality, or even, in any strict sense, of correspondence to action. Instead, intentional descriptions are typifications over an indefinite range of possible actions and situations. Of course, the attribution of intent is generally non-problematic, even transparent, for members of the society who, from their practical perspective, and for their practical purposes, are engaged in the everyday business of interpreting each others' actions.<sup>30</sup> Intentional descriptions not only suffice to classify purposeful behavior but, given the unique and fleeting circumstances of situated action, and the need to represent it efficiently, seem ideally suited to the task. For studies of purposeful action, however, there is a methodological consequence of recognizing the inherent contingency of action; namely, we can shift our focus from explaining away uncertainty in the interpretation of action, to identifying the resources by which the inevitable uncertainty is managed, 'for all practical purposes.' The central tenet of *social* studies of action is that the resources for interpretation—the sources of action's intelligibility—are not only cognitive, but interactional. While acknowledging the role of conventional meanings and individual predispositions in mutual intelligibility, therefore, this chapter focuses on the neglected other side of shared understanding; namely the local, interactional work that produces intelligibility, *in situ*. The starting premise is that the significance of action is an essentially collaborative achievement. Rather than depend upon reliable recognition of intent, mutual intelligibility turns on the availability of communicative resources to detect, remedy, and at times even exploit the inevitable uncertainty of action's significance.

30. Which is not to say that disputes over the meaning of an action don't arise, in which case the uncertainty of intentional attributions is very much a practical problem. But in such cases it is the 'right' interpretation of the action, not the fact of its inherent uncertainty, that is of interest to participants. Gumperz points out (1982) that "participants in a conversation need not agree on the specifics of what is intended. People frequently walk away from an encounter feeling that it has been highly successful only to find later that they disagree on what was actually said" (p. 326). Note that some actions are specifically exempted from the question of the actor's intent, i.e. certain crimes, such as rape, are defined by the perpetrator's behavior regardless of his avowed intent, and others are identified according to arbitrary conventions, e.g., "We call that behavior chopping wood" (Heap 1980, p. 93).

In order to underscore the breadth and subtlety of the resources available for shared understanding, and the precision of their use, this chapter focuses on the richest form of human communication, face-to-face interaction.<sup>31</sup> The methodological premise is that face-to-face interaction incorporates the broadest range of possible resources for communication, with other forms of interaction being characterizable in terms of particular resource limitations or additional constraints. Sections 5.1-5.3 describe the organization of the most unrestricted form of face-to-face interaction, everyday conversation. In the final section of this chapter, I consider some modifications to everyday conversation that have developed for specialized purposes in institutional settings and, in chapter 6, some additional constraints introduced by restrictions on the mutual access of participants to each other and to a common situation. Finally, human-machine communication is analyzed, in chapter 7, as an extreme form of resource-limited interaction.

### 5.1 Conversation as 'ensemble' work

The prevailing view of conversation is that speakers and listeners, pursuing some common topic according to individual predispositions and agendas, engage in an alternating sequence of actions and reactions. For students of human cognition and of language, conversation generally has been treated as epiphenomenal with respect to the central concerns of their fields. Cognitively, conversation is just the meeting ground of individual psychologies, while linguistically, it is the noisy, real-world occasion for the exercise of basic language abilities. On either view, the additional constraints imposed by situated language are a complication that obscures the underlying structure of cognitive or linguistic competence. As a consequence, linguists generally have not used actual speech for the analysis of linguistic competence, on the assumption that the phrasal breaks, restarts, hesitations and the like found in actual speech represent such a defective performance, that the data are of no use. And in analyzing idealized utterances, linguists have focused exclusively on the speaker's side in the communicative process (Streeck 1980). When one takes situated language as the subject matter, however, the definition of the field must necessarily shift to communication under concrete circumstances. And when one moves back far enough from the utterances of the speaker to bring the listener into view as well, it appears that much in the actual construction of situated language that has been taken to reflect problems of speaker performance, instead reflects speaker competence in responding to cues provided by the listener (Goodwin 1981, pp. 12-13).

Analyses of face-to-face communication indicate, then, that conversation is not so much an alternating series of actions and reactions between individuals, as it is a joint action accomplished through the participants' continuous engagement in speaking *and* listening (cf. Schegloff 1968, 1982;

31. In the discussion that follows I consider only a small subset of these resources, e.g. I do not include the wealth of prosodic and gestural actions described by students of interaction. The rationale for neglecting those resources here, and in the analysis of Chapter 7, is that the case of human-machine interaction is so limited (the system has no language, for example), that the basic resources, let alone the expressive subtleties, of human interaction are in question.

Goodwin 1981, p. 5). In contrast to the prevailing preoccupation of linguists and of discourse analysts with speaking, where the listener is generally taken for granted or extraneous, conversational analysis approaches the action of listening as consequential to the extent that:

Analysis shows that the listener's failure to act at the right time in the right way literally prevents the speaker from finishing what he was trying to say—at least from finishing it in the way he was previously saying it. The speaker, in continuing to speak socially (i.e. in taking account in speaking of what the other is doing in listening), makes *accountable* the listener's violations of expectations for appropriate listening behavior (Erickson and Shultz 1982, p. 118-19, original emphasis).

In the same way that the listener attends to the speaker's words and actions in order to understand them, in other words, the speaker takes the behavior of the listener as evidence for the listener's response. Schegloff offers the example of the lecturer:

Anyone who has lectured to a class knows that the (often silent) reactions of the audience—the wrinkling of brows at some point in its course, a few smiles or chuckles or nods, or their absence—can have marked consequences for the talk which follows: whether, for example, the just preceding point is reviewed, elaborated, put more simply, etc., or whether the talk moves quickly on to the next point, and perhaps to a more subtle point than was previously planned (Schegoff 1982, p. 72).

The local resources, or *contextualization cues* (Gumperz 1982a), by which people produce the mutual intelligibility of their interaction consist in the systematic organization of speech prosody (*ibid.*), body position and gesture (Birdwhistell 1970; Erickson 1982; Scheflen 1974), gaze (C. Goodwin 1981, M. Goodwin, 1980), and the precision of collaboratively accomplished timing (Erickson and Shultz 1982).<sup>32</sup> The richness of both simultaneous and sequential coordination:

suggests that conversational inference is best seen not as a simple unitary evaluation of intent but as involving a complex series of judgments, including relational or contextual assessments on how items of information are to be integrated into what we know and into the event at hand ... (Gumperz 1982b, p. 328-9).

As with any skill, in ordinary conversation these 'judgments' are made with such proficiency that they are largely transparent, though at times of breakdown they may become contestable (see Gumperz and Tannen 1979). Viewed as highly skilled performance, the organization of conversation appears to be closer to what in playing music is called 'ensemble' work (Erickson and Shultz 1982, p. 71) than it is to the common notion of speaker stimulus and listener response.

---

32. For example, Erickson and Shultz suggest that what may be disturbing about certain speaker hesitations in conversation is not so much the interruption of talk *per se*, but the fact that when talk stops and starts in temporally unpredictable ways, it is difficult for listeners to coordinate their listening actions (Erickson and Shultz 1982, p. 114).

## 5.2 Conversational organization

One reason to begin a consideration of interaction with the organization of conversation is that studies of everyday conversation (e.g. Sacks, Schegloff & Jefferson 1978), and more recently studies in specific institutional settings where the type, distribution, and content of turns at talk are constrained in characteristic ways (see 5.4), indicate that all of the various forms of talk (e.g. interviews, cross-examinations, lectures, formal debates, and so on) can be viewed as modifications to conversation's structure. As the basic system for situated communication, conversation is characterized by (i) an organization designed to support local, 'endogenous' control over the development of topics or activities, and to maximize accommodation of unforeseeable circumstances that arise, and (ii) resources for locating and remedying the communication's troubles as part of its fundamental organization.

### 5.2.1 Local control

Taking ordinary conversation as their subject matter, Sacks, Schegloff and Jefferson (1978) set out to identify the structural mechanisms by which this most 'unstructured' of human activities is accomplished in a systematic and orderly way. Two problems for any interaction are the distribution of access to 'the floor' and, closely related, control over the development of the topic or activity at hand. In contrast to mechanisms that administer an *a priori*, externally imposed agenda (for example, the format for a debate), the organization of conversation maximizes local control over both the distribution of turns, and the direction of subject matter. That is to say, who talks, and what gets talked about is decided *in situ*, by the participants in the conversation, over their collaborative construction of the conversation's course.

That turn-taking is a collaborative achievement, rather than a simple alternation of intrinsically bounded segments of talk, is evident in the common occurrence in actual conversation of simultaneous talk, of joint production of a single sentence, and of silence. The observations that somehow one speaker only takes the floor when two begin together, that a listener may finish the speaker's turn without it constituting an interruption, and that any participant in a conversation, including the last to speak, may begin a new turn out of silence, raise theoretical questions about the proper definition of a turn's boundaries, and the process by which turn transitions are organized (cf. Goodwin 1981, p. 2). In answer to such questions, Sacks *et al* (1978) have delineated a set of conventions or normative rules by which turn-taking is accomplished.<sup>33</sup> The set of rules provides that for every place in the course of an utterance that is a projectable completion point, or potential turn-transition place, one of the following occurs:

---

33. By normative, I mean that these are 'rules' only in the sense that they describe common practices observed by analysts of conversation. Speakers and listeners do not "know" these rules in the sense that they would or could formulate them in so many words. Rather, it can be seen by an observer, having these rules in mind, that they describe the practices by which people in conversation achieve the orderly distribution of turns.

- a) the current speaker selects a next speaker, e.g. by directing a question or other sequentially implicative utterance at a particular recipient,
- b) another participant self-selects, by being the first to start speaking,
- c) the current speaker continues.

Options a-c are not simply alternatives, but an ordered set. That is to say, at each place where a transition might be effected, the rule set applies as a matter of "if not a, then b is an option," "if not b, then c is an option," recursively, until a change in speakers occurs. This does not mean, of course, that participants wait to see which rule applies and then respond appropriately; the options are theirs to exercise, not the workings of some external mechanism. Moreover, the rules are not necessarily, but contingently, ordered; which rule will prevail in any given case is the product of the work done at a possible turn-transition place by the participants. The contingency of the rules means that the exercise of each is constrained by the presence of others in the set, independently of their actually being employed on any given occasion. That is to say, lower priority options constrain higher priority options, e.g. for the (b) option to be exercised given the presence of (c), it will need to be employed before (c) is invoked, at which time priority returns to (a).

Due to the ordering of the rule set, and the fact that candidate transition places are determined by projectable completion points of the current turn, the system for turn-taking constrains both the way that current speakers develop their talk, and the way that others listen. Generally, the speaker has reason to let the listener know where he or she is in the turn's course. This is sometimes done explicitly, as in introductions on the order of "I want to say a couple of things," or the use of 'story prefaces' (cf. Sacks 1974), which announce the onset of an extended, multisentence turn with its own distinctive shape. More frequently, however, turn transitions are managed interactionally, and in contrast with attempts to enumerate a set of discrete "turn-yielding signals" (Duncan 1974), the rule-set identified by Sacks *et al* views each possible completion point of a current turn as a candidate transition place. Possible completion and appropriate transitions are identified not according to some special cues, but with reference to the same range of syntactic, semantic and non-linguistic resources by which the listener decides the significance of what the speaker is saying.

Because turn-constructional units (words, phrases, sentences, etc.) specify minimal turn sizes but otherwise allow for expansion—any unit from a particle to an extended exposition can constitute a turn—and because what happens at the continually recurring transition points is contingent on the interpretations and actions of others, the turn is essentially interactionally determined:

By virtue of its character, it is misconceived to treat turns as units characterized by a division of labor in which the speaker determines the unit and its boundaries, and other parties to the conversation have as their task the recognition of them. Rather, the turn is a unit whose constitution and boundaries involve such a distribution of tasks (as we have noted): That a speaker can talk in such a way as to permit projection of possible completion to be made from his talk, and to allow others to use its transition places to start talk, to pass up talk, to affect directions of talk, and so on, and that their starting to talk, if properly placed, can determine where he ought to stop talk. (Sacks *et al*, p. 42)

The speaker, in other words, does not define the turn unilaterally: turn completion is as much a function of the listener's inclination to respond as it is a matter of the speaker's readiness to yield, though insofar as the speaker controls the floor, he or she holds some advantage. In order to preclude the exercise of option (b), for example, before having had a say, the current speaker can postpone completion, for example, by withholding a point until after the supporting arguments have been made. Alternatively, by *passing* on option (b) at a possible transition place, listeners invite the speaker to continue, turning what could be a transition into a pause in the same speaker's turn. Or listeners may, on finding either in the speaker's exercise of option (a), or failure to exercise option (c), that a turn is completed, *then* look back over the turn to try to find what was said in order to respond to it.

The interactional structure of turn-taking presents some distinctive problems for the definition and categorization of units in conversational analysis. For example, one might argue reasonably that silence should be classified differently according to whether it occurs within the turn of a single speaker (a pause), or between turns of different speakers (a gap) (Goodwin 1981, p. 18). The problem that arises for analysts is exemplified, however, in a case like the following:<sup>34</sup>

John: Well I, I took this course.  
 (0.5)  
 Ann: In h ow to quit?  
 [  
 John: which I really recommend.  
 (Goodwin 1981, p. 18).

The ambiguous status of the silence in this example as either a pause or a gap is not so much an analytic problem as it is an inherent property of situated talk. That is to say, the silence is treated by Ann as a gap, by John as a pause, such that "the same silence yields alternative classifications at different moments in time from the perspective of different participants" (*ibid.*, p. 19).<sup>35</sup> No single classification of the silence will do, as its status is inextricably tied to an event developing over time, and is subject to transformation. From Ann's point of view, at the point where she begins to speak, John's turn appears to be complete. John's extension of the turn, however, makes the silence into a pause, and Ann's turn into an interruption that begins in the midst, rather than at the completion of his utterance. The status of what constitutes John's "turn" in this exchange, and therefore the status of the silence, is *essentially* ambiguous, in other words, in a way that will not be remedied by any exercise of the analyst. And in fact, attempts to remedy the ambiguity must do damage to the phenomenon, which is precisely that boundaries of a turn are mutable, and that the structure of

---

34. In this and later examples, transcripts are presented with whatever notation and punctuation was used in the original source. Generally, [']s indicate overlap, numbers in parentheses represent elapsed times.

35. See section 5.3 for a discussion of how such competing definitions are routinely negotiated.

conversation is achieved, by speakers and hearers, in this locally developing, contingent way.

As a consequence of its interactional nature, the turn is not the kind of object that can be first defined, and then examined for how it is passed back and forth between speakers. Instead, intrinsic structural elements of the turn are contingent on the process by which control changes hands between participants in conversation, as is the structure of the conversation produced. The point is not just that speakers can extend the length of their turns by the addition of further units of speech, but that through that essentially transparent mechanism they are able to change the emerging *meaning* of their talk within a turn to fit the actions of their listener (cf. Goodwin 1981, p. 11). The localness of the constraints on speakers' constructions of turns-at-talk, and the turn's contingency on other speakers, makes conversation maximally sensitive and adaptable to particular participants, and to unforeseen circumstances of the developing interaction. The turn-taking system for conversation demonstrates how a system for communication that accommodates any participants, under any circumstances, may be systematic and orderly, while it must be essentially *ad hoc*.

### 5.2.2 *Sequential organization and coherence*

In addition to providing a mechanism for control over the distribution of turns, the turn-taking system bears a direct relation to the control of inferences about the conversation's content. In general, a coherent conversation is one in which each thing said can be heard as relevant to what has come before. Most locally, this means that the relevance of a turn is conditional on that which immediately precedes it:

By conditional relevance of one item on another we mean; given the first, the second is expectable; upon its occurrence it can be seen to be a second item to the first; upon its nonoccurrence it can be seen to be officially absent (Schegloff 1972, p. 364).

Two utterances that stand in a relationship of conditional relevance of one on the other, in this local sense, constitute an *adjacency pair* in Schegloff and Sacks' terminology (1973), though conditional relevance is not limited to literal adjacency (cf. Levinson 1983, p. 304). The first part of an adjacency pair both sets up an expectation with respect to what should come next, and directs the way in which what does come next is heard (Schegloff 1972, p. 367). By the same token, the absence of an expected second part is a notable absence, and therefore takes on significance as well. In this way silences, for example, can be meaningful—most obviously, a silence following an utterance that implicates a response will be 'heard' as belonging to the recipient of the utterance, and as a failure to respond. Similarly, a turn that holds the place of the second part of an adjacency pair, but cannot be made relevant to the first, will be seen as a *non sequitur*, or as incoherent.

The conditional relevance of adjacency pairs is an instance of what we might call, following Durkheim, a 'social fact'—the implicativeness of the first part of an adjacency pair is external and constraining with respect to the second—but in a particular way. The constraint is not just a matter

of the coherence or intelligibility of the second part, but of the accountability of the respondent for inferences that the first part warrants. For example, in the case of interactions opened by a summons, such as calling someone's name, or ringing their doorbell:

A member of the society may not 'naively choose' not to answer a summons. The culture provides that a variety of 'strong inferences' can be drawn from the fact of the official absence of an answer, and any member who does not answer does so at the peril of one of those inferences being made ... although members can, indeed 'choose' not to answer a summons, they cannot do so naively, i.e. they know that if the inference of physical or interactional absence cannot be made, then some other inference will, e.g. they are cold shoulder, insulting, etc. (Schegloff 1972, p. 367-68).

The summons-answer pair is an example of a sequence in which the first part implicates a particular type of response in the immediate next turn, to the extent that if no response occurs the first speaker is justified in, for example, repeating the summons. And on eliminating the possibility that the summons has not been heard, and determining that it will in fact not be answered, the summoner is justified in making further inferences regarding the recipient's availability, interest, and the like.<sup>36</sup> In this sense, we are not so much constrained by the rules of conversation, as we are "caught in a web of inferences" (Levinson 1983, p. 321, note 16). That is to say, the rules of conversation are neither strictly optional (their breach does have consequences) nor obligatory (they may be breached without a necessary loss of coherence); they are, however, relentlessly meaningful.

While conditional relevance is a constraint on inference, it is a weak constraint in the sense that it does not prescribe *what* counts as a response to a given action, only that whatever is done next *will be viewed* as a response. In fact, the range of actions in a second part position that can be heard as a response is extended in virtue of the expectation that adjacency sets up, rather than constrained. That is to say, an action that is not in any explicit way tied to the action that it follows will nevertheless be interpreted as a response, in virtue of its position:

- A: Are you coming?
  - B: I gotta work.
- (Goffman 1975, p. 260)

The position of B's utterance as a response means that we look for its relevance to A's question. So in this case, B's statement can be heard as a negative reply, just as B's question can be heard as an affirmative reply in the following:

36. Schegloff points to the telephone as an example of a technology that embodies the organization of the summons-answer pair:

In this mechanical age it may be of interest to note that the very construction and operation of the mechanical ring is built on these principles. If each ring of the phone be considered a summons, then the phone is built to ring, wait for an answer, if none occurs, to ring again, wait for an answer, ring again, etc. And indeed, some persons, polite even when interacting with a machine, will not interrupt a phone, but wait for the completion of a ring before picking up the receiver (1972, p. 365).

- A: Have you got coffee to go?  
 B: Milk and sugar?  
 (Merritt 1976)

The sequential implicature exemplified by adjacency pairs is not literally conditional on adjacency, but instead allows for multiple levels of embedded sequences aimed at clarification and elaboration. The result is that answers to later questions can precede answers to earlier ones without a loss of coherence:

- B: ... I ordered some paint from you uh a couple  
 of weeks ago some vermillion  
 A: Yuh  
 B: And I wanted to order some more the name's Boyd ((Request 1))  
 A: Yes // how many tubes would you like sir ((Question 1))  
 B: An-  
 B: U hm (.) what's the price now eh with V.A.T.  
 do you know eh ((Question 2))  
 A: Er I'll just work that out for you= ((Hold))  
 B: =Thanks ((Accept))  
 (10.0)  
 A: Three pound nineteen a tube sir ((Answer 2))  
 B: Three nineteen is it= ((Question 3))  
 A: =Yeah ((Answer 3))  
 B: E::h (1.0) yes u hm ((dental click))  
 ((in parenthetical tone))  
 e:h jus-justa think, that's what three nineteen  
 That's for the large tube isn't it ((Question 4))  
 A: Well yeah it's the thirty seven c.c.s ((Answer 4))  
 B: Er, hh I'll tell you what I'll just eh eh ring you  
 back I have to work out how many I'll need. ((Account for  
 no Answer 1))  
 Sorry I did- wasn't sure of the price you see  
 A: Okay  
 (Levinson 1983, p. 305)

The local system of adjacency pair organization, in its canonical form operating over two turns, can by the accumulation of first parts (e.g. requests, questions) project an extended sequence of expected seconds like that of the last example, i.e: (R1(Q1(Q2(Q3(Q4—A4)A3)A2)A1). As Levinson points out with respect to this example, B's final obligation is to account for his failure to provide an answer to Q1, demonstrating his orientation to the 'social fact' that an answer is called for. That failure, in turn, effectively constitutes B's withdrawal of R1, freeing A of the obligation to respond to that original request:

What the notion of conditional relevance makes clear is that what binds the parts of adjacency pairs together is not a formation rule of the sort that would specify that a question must receive an answer if it is to count as a well-formed discourse, but the setting up of specific expectations which have to be attended to. Hence the non-

occurrences of an R[equest] 1 and an A[nswer] 1 in [the example] do not result in an incoherent discourse because their absences are systematically provided for (*ibid.*, p. 306).

The overall coherence of a conversation, in sum, is accomplished through the development and elaboration of a local coherence operating in the first instance across just two turns, current and next. The resiliency of embedding, however, is such that the backward 'reach' of relevance extends beyond the immediately preceding turn:

- C: (telephone rings)
  - A: Hello.
  - C: Is this the Y?
  - A: You have the wrong number.
  - C: Is this KI five, double four, double o?
  - A: Double four, double six.
  - C: Oh, I am sorry.
- (Goffman 1975, p. 285)

In this case the apology is only intelligible if we view the entire telephone call as its object, not just the utterance of A that it immediately follows. Similarly, to use another example of Goffman's (*ibid.*, p. 286), the applause at the end of a play is a response not to the delivery of the final line, or the drop of the curtain, but to the entire play. The relevance of an action, in other words, is conditional on any identifiable prior action or event, however far that may extend for the participants (i.e. it may be a lifetime, say, for mother and child), insofar as the previous action can be tied to the current action's immediate, local environment. As a consequence, conditional relevance does not allow us to predict from an action to a response, but only to project that what comes next will be a response and, retrospectively, to take that status as a cue to how what comes next should be heard. The interpretation of action, in this sense, relies upon the liberal application of *post hoc, ergo propter hoc*.

### 5.3 Locating and remedying communicative trouble

Communication takes place in real environments, under real "performance" requirements on actual individuals, and is vulnerable therefore to internal and external troubles that may arise at any time, from a misunderstanding, to a clap of thunder (cf. Schegloff 1982). Our communication succeeds in the face of such disturbances not because we predict reliably what will happen and thereby avoid problems, or even because we encounter problems that we have anticipated in advance, but because we work, *in situ*, to identify and remedy the inevitable troubles that arise:

It is a major feature of a rational organization for behavior that accommodates real-worldly interests, and is not susceptible of external enforcement, that it incorporates resources and procedures for repair of its troubles into its fundamental organization (Sacks, Schegloff & Jefferson 1978, p. 39).

The resources for detecting and remedying problems in communication, in other words, are the same resources that support communication that is trouble free. With respect to control, for example, the contingency of conversational options for keeping and taking the floor—specifically, the fact that transitions should be done at possible turn completion points and not before, and that at each possible completion point the speaker may extend his or her turn—means that gaps and overlaps can and do occur. The extent to which conversationalists accomplish speaker transitions with a minimum of gap or overlap is the product not only of the "accurate" projection of completion points, but of the repair of routine troubles. The following is a simple example of a familiar kind of conversational repair work:

- C: .hhh aa::of course under the circumstances Dee  
I would never:: again permit im tuh see im.  
D: Yeah  
(0.7)  
C: tlk. Be:cuz *he*—  
[  
D: Wul did'e ever git—ma:rried'r anything?  
C: Hu: :h?  
[  
D: Did yee ever git—ma:rried?  
C: .hhh I have no *idea*.
- (cited in Atkinson and Drew 1979, p. 40)

In addition to negotiating the transfer of control, participants in conversation must be alert to the possibility of substantive troubles of interpretation. Schegloff (1982) points out that tokens such as 'uh huh,' commonly viewed as a signal from the listener that encourages the speaker to continue, operate as they do not simply because there is a semantic convention to the effect that such tokens claim or signal understanding, but rather because through such tokens listeners pass up, at possible transition places, the opportunity to initiate repair on the preceding talk. The same option that provides for ordinary turn transitions, in other words, affords the recipient of an utterance the occasion to assert that he or she has some trouble in understanding, or to request some clarification.

The work of repair includes calling the other's attention to the occurrence of some troublesome item, remedying it, and resuming the original line of action in which the troublesome item is embedded. Jefferson identifies two kinds of trouble flag, a questioning repeat, and a non-specific interrogative (e.g. "What?" or "Huh?"). The two are different in that the repeat simultaneously flags the fact that there is some troublesome item in the prior talk, and *locates* it for the first speaker:

- Steven: One, two, three ((pause)) four, five, six, ((pause)) eleven, eight, nine, ten.  
Susan: "Eleven"?—eight, nine, ten?  
(ibid., p. 295)

An interrogative request for clarification, in contrast, leaves it to the first speaker both to locate the item that produced the request, and to remedy it. In that case, the location of the item and the

*remedy* may be done simultaneously, in the first speaker's reply:

- A: If Percy goes with—Nixon I'd sure like that.  
 B: Who?  
 A: Percy.  
 (ibid., p. 296)

In both cases, the adjacency of the trouble flag to the troublesome item is obviously a resource for the latter's identification.<sup>37</sup> On the other hand, listeners generally do not interrupt a speaker to flag some trouble, but rather wait for the next turn transition place, or point of completion. By permitting the speaker to complete the utterance in which the trouble is heard, the listener is warranted in assuming that there is no unsolicited remedy forthcoming, and the complaint becomes a legitimate one (ibid., p. 298).

A side sequence initiated by an assertion of misunderstanding or request for clarification sets up an exchange that the first speaker did not necessarily anticipate, but to which he or she is obliged to respond. That is to say, a failure on the part of the speaker to provide clarification in response to an explicit request is a "noticeable absence;" i.e. is seen as specifically *not providing clarification*, as opposed to just doing something else. The 'failure to respond' then becomes something about which complaints can be made, or inferences may be drawn (Atkinson and Drew 1979, p.57).

In responding to a request for clarification, the sequential implicativeness of the troublesome utterance is temporarily suspended in favor of finding a remedy for the recipient's problem. Routinely in face-to-face conversation, the 'adjacency' relation or continuity between utterance and response, and the coherence of the interaction, are sustained across such embedded side sequences. This is true even when the request for clarification results in complete reformulation of the initial utterance. That is to say, while the response may ultimately address the reformulation, not the original utterance, it will still be heard as a response to the original:

- (AA, CA, 2)  
 M: What=so what did you *do* did you have people—did  
     Morag (.) come (.) down with the car ag ain ( ) or what  
     [  
 A:       When last year  
 M: Mmm  
     how did you man age to *shift* it back and forward  
     [  
 A:       Last year I don't know ho:w I  
     managed it I got it a:ll in (0.8) two suitcases...  
 (cited in Atkinson and Drew 1979, p. 239).

37. It is worth noting in this case that while the "Who?" is in fact ambiguous, speaker *A* appears to have no trouble identifying its referent. It is hard to account for this in any way other than in virtue of *A* and *B*'s common knowledge of politics, i.e. that it is more likely that "Percy" would be a troublesome item in this context than that "Nixon" would. Such an analysis cannot be more than conjecture, however.

In this case it is just because A's "When last year" cannot be heard as a reply to M's question, that it is heard as an embedded request for clarification. By the same token, the fact that a reply to M's question is deferred makes A's response to the reformulation about 'managing it' relevant to the original question about 'Morag and the car.'

Turn transition places provide recurring opportunities for the listener to initiate some repair or request for clarification from the speaker. Alternatively, clarification may be offered by the speaker not because the recipient of an utterance asks for it, but because the speaker finds evidence for some misapprehension in the recipient's response:

Dr: Hev'y ever had palpitations  
 (0.6)  
 Pt: Noh. M' feet ain't painin' me but they swell sometime  
 Dr: No I— it's when yer heart starts beating really  
       fast an y'feel like y'can't catch yer breath.  
 Pt: No. uh-uh. I never had that  
 (Frankel 1984, p. 155)

While the patient here produces a response that is, formally, an answer to the doctor's question, the answer reveals a lexical problem that the doctor detects and remedies. The problem is of course unknown to the patient until the remedy is offered; if the doctor had failed to detect the misunderstanding, or had decided to let it pass and had consequently not offered the correction, the conversation to all appearances could have continued on as if there were no trouble. In this case the problem is evidently, from the doctor's point of view, 'worth bothering about,' in myriad other cases it is not. That is to say, given the lack of specific criteria for shared understanding in most cases, a crucial part of interactional competence is the ability to judge whether some evidence that the recipient has misunderstood the speaker's words warrants the work required for repair (Jefferson 1972). The decision whether to challenge a troublesome item, or to let it pass involves, in part, a weighing of the relative work involved in the item's clarification *versus* the foreseeable dangers of letting it go by. The risks of the latter are exemplified by the 'garden path' situation, where speakers fail to identify some communicative trouble at the point where it occurs, and only discover at some later point in the interaction that there *has been* some misunderstanding (cf. Jordan and Fuller 1975, Gumperz and Tannen 1979).<sup>38</sup> At the point of discovery, the coherence of the interaction over some indefinite number of past turns may be called into question, and the source of the trouble may be difficult or impossible to reconstruct. In contrast to the routine problems and remedies that characterize local repair in conversation, such a situation may come close to communicative failure, i.e. it may require abandoning the current line of talk, or beginning anew.

---

38. An instance of this in the case of human-machine communication is discussed in 7.5.2.

## 5.4 Specialized forms of interaction

A distinguishing feature of ordinary conversation is the local, *in situ* management of the distribution of turns, of their size, and of what gets done in them, those things being accomplished in the course of each current speaker's turn. There are, of course, numerous institutionalized settings that prescribe the organization and subject matter of interaction. Interactional organization is institutionalized along two dimensions that are of particular relevance to problems discussed in Chapter 7; (1) the pre-allocation of *types of turns*, i.e. who speaks when, and what form their participation takes and (2) the prescription of the substantive content and direction of the interaction, or the *agenda*.

### 5.4.1 Pre-allocation of turn types

Analysis of encounters between physicians and patients (Frankel 1984) and of the examination of witnesses in the courtroom (Atkinson and Drew 1979) reveal a turn-taking system that is pre-allocated in terms of both the types of turns, and the distribution of those types between the participants. While there is no explicit formulation of a rule for the organization of talk in medical encounters, for example, Frankel reports that physicians' utterances almost always take the form of questions (99% of the time), while patients' take the form of answers. And in the courtroom, by definition, the examiner has the sole right to ask questions, while the examined is obliged to answer. In the courtroom, the convention that only two parties participate holds in spite of the number of persons present, and with the exception only of certain prescribed methods for "interruption" from the other counsel which, in virtue of the prescription of when and how interruption is to be done, itself becomes a technical matter in the courtroom setting. In both medical and legal settings, the effect of the pre-allocation of turn types is to deliver control of the proceedings from the 'client' or layperson, back to the 'expert' or specialist. At the close of each question-answer sequence, control is relinquished to whomever would start another sequence and the expert, having rights to the role of questioner, is repeatedly the one to re-take control.

While the constraints on medical or legal interaction can be seen as institutional, and in that sense as external to any particular occasion, it is nonetheless the case that in every actual instance the constraints are realized locally and collaboratively. Insofar as the interaction is locally managed, turn-transitions are subject to the same problems that arise in everyday conversation, and are remedied *via* the same methods, as in the following example of an overlap that W remedies by repeating her answer to C's question:

[C=counsel, W=witness)

C: An (.) about how long did you say you  
ta:lked before (this was )

[

W: I don't remember

C: (started ta kiss (h) a)=

W: =I don't remember. (from Atkinson and Drew 1979, p. 67)

At the same time, the fact that procedural constraints on turn transitions are managed locally, even in these settings, means that general conventions of conversational turn-taking can be exploited to further the special purposes of the participants. Because of the fact that pauses in conversation, for example, will be ascribed significance insofar as they are seen to belong to a selected next speaker, a pause following an examination sequence can be used by the examining counsel to "comment" on the response to the jury, as in the following examination in a rape case cited by Atkinson and Drew:

- C: You were out in the woods with the defendant at this point isn't that so  
 (1.0)  
 W: Yeah  
 (7.0)  
 C: And the defendant (.) took (.) the car (1.0) and backed it (1.0) into some trees didn't  
 (0.5)  
 W: Mm hm  
 [  
 C: underneath some trees.  
 (ibid., p. 241)

In this case, the pre-allocated order of turns produces the inferentially implicative character of the 7 second pause—specifically, that the pause belongs to the counsel—and ensures that no other speaker will use the pause as an opportunity to take over the floor. The pause is used by the counsel in an unspoken turn that insinuates further 'information' into the message that the jury receives from the witness's answer. In the medical encounter, similarly, the physician can use a silence as an unspoken turn—in the following example, in order to avoid having to deliver bad news through disagreement:

- Pt: This— chemotherapy (0.2) it won't have any lasting effects on havin' kids will it?  
 (2.2)  
 Pt: It will?  
 Dr: I'm afraid so.  
 (Frankel 1984, p. 153)

Finally, although respective turns of physician and patient, or counsel and witness, are constrained to be either questions or answers, these are minimal characterizations, and provide no instruction for how, or what, specific utterances can be put into such a format. In the courtroom, for example, rules of evidence apply—relevance to the case at hand, status of the evidence as hearsay, the use of leading questions and the like—where the application of those rules is situated and problematic, and is itself part of the technical business of the proceedings. And the format of questions and answers in the courtroom accommodates a range of activities including accusations,

challenges, justifications, denials and the like. Those activities are not prescribed in the way that the question-answer format is, and what counts as a question or an answer is itself liable to challenge. As a consequence, rules for courtroom interaction, like those for everyday conversation, constitute a resource for social order, not a recipe, nor an explanation.

#### 5.4.2 *Agendas*

Various settings, of course, do comprise prescriptions not only about forms of talk, but also about the substantive direction and purposes of the interaction:

in several different types of speech-exchange situations, there can be occasions in which participation is constructed by a speaker in continuing response to interactional contingencies and opportunities from moment to moment, and occasions in which a participant has a preformed notion, and sometimes a prespecified text, of what is to be said, and plows ahead with it in substantial (though rarely total) disregard for what is transpiring in the course of his talking (Schegloff 1982, p. 72).

A major concern for participants in such settings is the distribution of knowledge about the 'agenda' (Beckman and Frankel 1983). The communicative task of novice and expert in a given setting is to coordinate their actions in a way that accommodates their asymmetrical relationship to the interaction's institutionalized purposes. At the same time, it is precisely the difference in their respective familiarities *vis a vis* the setting's protocols and purposes that in large measure distinguishes the 'expert' or specialist from the 'novice' or layperson (cf. Erickson and Shultz 1982, p. 4).

The work of Beckman and Frankel (1983) on physician's methods for eliciting a patient's 'chief complaint' is illustrative. They point out that the medical literature has generally viewed the agenda for medical interviews as the patient's, in the sense that it is the patient who comes to the physician with a complaint, and who is the source of the information required for the complaint's diagnosis. Given this view, a commonly cited problem for physicians is the experience of discovering, at the point where the physician is about to conclude the office visit or at least the history-taking segment of the interview, that the patient has 'hidden' some information that is relevant to a chief complaint. In contrast, by inverting the common view, Beckman and Frankel identify the relevant agenda in medical encounters as the physician's, and further locate the source of the 'hidden agenda' problem in ways that the *physician's* actions, in the opening sequence of the clinical encounter, serve systematically to foreclose a complete report of symptoms by the patient.<sup>39</sup>

The point of Beckman and Frankel's observation relevant for present purposes—a point that I return to in Chapter 6—is their insight that analysts of the medical interview have been misconceiving the essential problem for the interaction. Specifically, the problem is not that the

---

39. Specifically, they cite the physician's tendency, given any mention of symptoms by the patient, to engage in early hypothesis testing; "once hypothesis testing has begun, it is difficult for the patient to get a word in edgewise without deviating from conventional rules of discourse which relate types of speech acts to one another, in this case the relevance of an answer to the question that preceded it" (*ibid.*, p. 9).

patient 'hides' the agenda, but that the patient, as a novice in this setting, does not understand the institutional purposes of the interaction, i.e. the identification of a 'chief complaint,' or the physician's strategy for achieving those purposes. The patient's task is misconceived, therefore, if it is viewed as either carrying out the plan of the interview, or as failing to do so. The point is rather that the patient *does not know the plan*, and is therefore only able to cooperate to the extent that being responsive to the physician's actions, locally, constitutes cooperation in realizing the plan. To the extent that the patient's cooperation is contingent on the physician's actions, the success of the interview is as well.

The actual production of an agenda, through local interactional work, is evident in the following excerpt from a career counseling interview, reported by Erickson and Shultz (1982, p. 77-78):

- C: Well, let's start from scratch. What did you get in your English 100 last semester?  
S: A 'C.'  
C: Biology 101?  
S: 'A.'  
C: Reading 100?  
S: 'B.'  
C: Med tech... 'B'? (medical technology)  
S: 'B.'  
C: Gym?  
S: 'A.'  
C: Was that a full credit hour? What was it?  
S: It was a wrestling..two periods.  
C: Wrestling. (He writes this on the record card, then shifts postural position and looks up from the record at the student.) Ok, this semester... English 101?  
S: (Changes facial expression, but no nod or "mhm" in response to the question.)  
C: That's what you've got now...  
S: (Nods.)  
C: Biology 102? Soc Sci 101. (The counselor is looking down.)  
S: I.I.. I don't have Biology 102. I have, mm, 112.  
C: (The counselor corrects the record card.) Soc Sci 101?  
S: (Nods.) Mhm.  
C: Math 95.  
S: (Nods.)  
C: Med Tech 112.  
S: (Nods.)  
C: Gym.  
S: (Nods.)

In Erickson and Shultz's analysis, this interview comprises two adjacent routines, by which the counselor establishes the student's academic status first as a matter of courses completed, and then as a matter of courses underway. The problem negotiated by counselor and student is that the counselor's behavior is superficially the same across both routines, so that the juncture or transition between them, which requires a change in the task of the student, is initially missed by the student.

The student's failure to respond to the query "English 101?" demonstrates the problem to the counselor, who then offers a remedy.

While the organization of this and any interaction can be analyzed, *post hoc*, into a hierarchical structure of topics and subtopics, or routines and subroutines, the coherence that the structure represents is actually achieved in real-time, as a local, collaborative, sequential accomplishment. This view stands in marked contrast to the assumptions of students of discourse to the effect that the actual enactment of interaction is the behavioral realization of a plan. Instead, every instance of coherent interaction is an essentially local production, accomplished collaboratively in real time, rather than "born naturally whole out of the speaker's forehead, the delivery of a cognitive plan" (Schegloff 1982, p. 73):

Good analysis retains a sense of the actual as an achievement from among possibilities; it retains a lively sense of the contingency of real things. It is worth an alert, therefore, that too easy a notion of 'discourse' can lose us that.

If certain stable forms appear to emerge or recur in talk, they should be understood as an orderliness wrested by the participants from interactional contingency, rather than as automatic products of standardized plans. Form, one might say, is also the distillate of action and interaction, not only its blueprint. If that is so, then the description of forms of behavior, forms of discourse... included, has to include interaction among their constitutive domains, and not just as the stage on which scripts written in the mind are played out (*ibid.*, p. 89).

The organization of face-to-face interaction is the paradigm case of a system that has evolved in the service of orderly, concerted action over an indefinite range of essentially unpredictable circumstances. What is notable about that system is the extent to which mastery of its constraints *localizes* and thereby *leaves open* questions of control and direction, while providing built-in mechanisms for recovery from trouble and error. The constraints on interaction in this sense are not determinants of, but are rather "production resources" (Erickson 1982) for, shared understanding. The limits on available resources for accomplishing a shared agenda in a case of 'interaction' between people and machines, and for detecting and remedying the troubles that that task poses, is the subject of Chapter 7.

## 6. The Case and Methods

In this age, in which social critics complain about the replacement of men by machines, this small corner of the social world has not been uninvaded. It is possible, nowadays, to hear the phone you are calling picked up and hear a human voice answer, but nevertheless not be talking to a human. However small its measure of consolation, we may note that even machines such as the automatic answering device are constructed on social, and not only mechanical, principles. The machine's magnetic voice will not only answer the caller's ring, but will also inform him when its ears will be available to receive his message, and warn him both to wait for the beep and confine his interests to fifteen seconds (Shegloff 1972, p. 374.)

Chapter 7 describes people's first encounters with a machine called an 'expert help system;' a computer-based system attached to a large and relatively complex photocopier, and intended to instruct the user of the copier in its operation. The system's identification as an 'expert help system' both locates it in the general category of 'expert systems,' and indicates that a function of this system is to provide procedural instructions to the user. The idea behind 'expert' systems in general is that expertise consists in a body of knowledge about a particular domain (in this case, about how to operate the copier), and rules for its use (namely, to generate a plan for the user). The 'knowledge' of the system comprises a set of rules about copying jobs and procedures, encoded as data structures in the programming language LISP, that control the presentation of instructions to the user on a video display. The design objective is that the system should provide timely and relevant information to the user regarding the operation of the copier. The information should be presented not as a compendium, but in a step-wise order wherein each next instruction is invoked by the user's successful enactment of the last. To apply its 'knowledge' of the copier and provide the user with appropriate instruction, therefore, the system must somehow recognize the action of the user to which it should respond. It is this problem in particular, the problem of the system's recognition of the user's actions, that this study explores.

### 6.1 The 'Expert Help System'

In contrast to relatively unrestricted occasions of human interaction like ordinary conversation, certain properties of the event of using the 'expert help system' provide grounds for imagining that one might safely predict, in some detail, just how the event will go:

- The interaction is instrumental;
- The possible goals of the interaction are defined by the machine's functionality;
- The structure of the interaction is procedural, constituted by a sequence of actions whose order is partially enforced;
- The criteria of adequacy for each action can be specified.

In as much as, in machine operation, the user's purposes are constrained by the machine's functionality, and her actions by its design, it seems reasonable to suppose that the user's purposes should serve as a sufficient context for the interpretation of her actions. The strategy that the design adopts is to project the course of the user's actions as the enactment of a plan for doing the job, and then use the presumed plan as the relevant context for the action's interpretation.<sup>40</sup> Through the user's response to a series of questions about the state of her originals and the desired copies, her purposes are identified with a job specification, the specification (represented in the system as a data structure with variable fields) invokes an associated plan, and the enactment of the plan is prescribed by the system as a step-wise procedure.

Having mapped the user's purposes to a job specification, and the job specification to a plan, the plan is then effectively ascribed to the user. The rationale for this move is that the plan is conveyed to the user in the form of instructions for a step-wise procedure, the user is following the instructions and consequently, one can assume, is following the plan that the instructions describe. Under that assumption, the effects of certain actions by the user are mapped to a place in the system's plan, and that mapping is used to locate an appropriate next instruction. The actions by the user that effect changes in the machine's state comprise some physical actions on the machine (putting documents into document trays, opening and closing machine covers and the like), and directives to the system in the form of selections of text on a video display. The hope of the designer is that the effects of these actions by the user can be mapped reliably to a location in the system's plan, and that the location in the plan will determine an appropriate system response. The relevant sense of 'interaction' in this case, therefore, is very simply that the provision of instruction is not only fitted to the user's purposes, but occasioned by her actions.

The design assumes that it is the correspondence of the system's plan to the user's purposes that enables the interaction. In contrast, Chapter 7 demonstrates that user and system each has a fundamentally different relationship to the design plan. While the plan directly *determines* the system's behavior, the user is required to *find* the plan, as the prescriptive and descriptive significance of a series of procedural instructions. While the instructions, and the procedure that they describe, are the object of the user's work, they do not reconstruct the work's course, nor do they determine its outcome.

---

40. See chapter 3. Analysts of the intention-action relationship are troubled by the 'diffuse' and 'tacit' nature of intentions in many situations, and the consequent problem of determining just what is the actor's 'true' intent. This seems less of a problem with goal-directed activities, where the goal, as defined by the analyst, can simply be taken *a priori* as the intent of the actor. The argument of this thesis, of course, is that the relief from the problem of determining intent that task-oriented interaction seems to offer is only a temporary palliative to AI's problem; the real solution must lie in an alternative understanding of the nature of intentions and their relation to actions—one that views the problem of identifying intent as an essentially contingent, practical, and interactional rather than theoretical problem.

## 6.2 The problem of following instructions

The practical problem that the 'expert help system' was designed to solve turns on the nature of the work of following instructions, and on the relation of that work to the work of communicating instructions. The general task in following instructions is to bring canonical descriptions of objects and actions to bear on the actual objects and embodied actions that the instructions describe (Lynch, Livingston and Garfinkel 1983). Studies of instruction in cognitive and social science alike have focused, on the one hand, on the problem of providing adequate instructions and, on the other, on the problem of finding the practical significance of procedural instructions for situated action.

Social studies concerned with the production and interpretation of instructions have concentrated on the irremediable incompleteness of instructions (Garfinkel 1967, chpt. 1), and the nature of the work required in order to 'carry them out.' The problem of the instruction-follower is viewed as one of turning essentially incomplete descriptions of objects and actions into practical activities with predictable outcomes (Zimmerman 1970; Amerine and Bilnes, 1979). A general observation from these studies is that while instructions rely upon the recipient's ability to do the implicit work of anchoring descriptions to concrete objects and actions, that work remains largely unexamined by either instruction-writer or recipient, particularly when the work goes smoothly.

In a study of instruction-following as practical action, Amerine and Bilnes (1979) point out that instructions serve not only as prescriptions for what to do, but also as resources for retrospective accounts of what has already happened:

Successfully following instructions can be described as constructing a course of action such that, having done this course of action, the instructions will serve as a descriptive account of what has been done (*ibid.*, p. 5).

More than the 'correct' execution of an instruction, in other words, successful instruction-following is a matter of constructing a course of action that is accountable to the general description that the instruction provides. The work of constructing that course is neither exhaustively enumerated in the description, nor completely captured by a retrospective account of what was done. Instructions serve as a resource for describing what was done not only because they guide the course of action, but also because they filter out of the retrospective account of the action, or treat as 'noise,' everything that was actually done that the instructions fail to mention:

if the experiment is 'successful,' if it achieves its projected outcome, the instructions can serve as an account of 'what was done,' although in the actual performance a great deal more is necessarily done than can be comprised in the instructions (*ibid.*, p. 3).

The credibility of instructions rests on the premise not only that they describe what action to take, but that if they are followed correctly, the action will produce a predictable outcome. An unexpected outcome, accordingly, indicates trouble and warrants some remedy. As long as instructions are viewed as authoritative, the preference in remedying a faulted outcome is to account for the failure in outcome without discrediting the instruction. An obvious solution is to locate the

trouble somewhere in the instruction's 'execution.' In assessing the course of the work for troubles in execution, questions inevitably arise concerning the relation of the many actions that were taken that are not specified by the instructions, to the faulted outcome. Previously insignificant details may appear crucial, or the meaning of the instructions may be transformed in such a way that they are found not to have been followed after all. Amerine and Bilnes give an example, drawn from science experiments in a third grade classroom, of the kind of problem inherent in reasoning inductively about the relation between courses of action and outcomes:

To expedite carrying out this lesson two similar and functionally equivalent pans of water were placed on a table in the center of the room and the students were called on by pairs to try the exercise. Toward that end, when, as related above, this activity had become particularly competitive, one of the children approached a pan but was urged by classmates to use the other one because it was 'luckier.' We are not sure how this notion came about, although in a pair of trials closely preceding this comment the student using the 'unlucky' pan had failed, while the child using the other one had succeeded. At any rate, the student followed this advice and the experiment was successful. Both of the following two children rushed for the 'lucky' pan, though the loser settled for the 'unlucky' one (and succeeded nonetheless). In the case of the next pair, the second child waited for the first to finish using the 'lucky' pan, and then also used it. The 'unlucky' pan remained unused thereafter ... In neither case are such observations *by nature* illogical or irrelevant ... But in these science experiments our understanding of the relationship between the practical course of action and its outcome seems to leave no place for 'luck'... Therefore such factors become 'noise' (*ibid.*, p. 9-10).

The ability to discriminate between relevant information and 'noise' in a given domain of action, by invoking both precepts and practice, is a part of what we recognize as expertise. The point of the "lucky pan" example is that the process by which that ability is acquired is a fundamentally inductive and *ad hoc* one, regardless of the degree to which rules of action are encoded and prescribed. In the final analysis, no amount of prescription, however precise or elaborate, can relieve situated action "of the burden of finding a way through an unscheduled future while making a convincing case for what is 'somehow' extracted from that future" (Lynch *et al* 1983, p. 233). The latter is the problem of accountably rational, situated action, however adequate the instructions for that action may be.

### 6.3 Previous studies on communication of instruction

An appreciation for what is required in instruction-following makes it easier to understand the problem that the communication of instructions attempts to resolve; namely, the troubles inherent in turning an instruction into an action. Motivated by the project of designing instructional computer systems, researchers in artificial intelligence have looked at instruction as a question of communicative resources available to expert and novice. One of the earliest such projects, the Computer-Based Consultant project begun at Stanford Research Institute in the 1970's, continues to direct research on task-oriented communication through 'natural language' on what has become the canonical problem of assembling a simple mechanical device. The goal of the original project was:

to produce a computer system that could fill the role of an expert in the cooperative execution of complex tasks with a relatively inexperienced human apprentice. The system was to use rich channels of communication, including natural language and eventually speech. The main function of the consultant was to aid the apprentice in the diagnosis of faulty electromechanical equipment, and the formulation of plans for the assembly, disassembly, and repair of the equipment (Sacerdoti 1977, p. 3).

Using an air compressor as the assembly task, a corpus of dialogues was collected in situations "simulating direct interaction between a person and a computer" (Grosz 1979, p.3); i.e. where two people—one acting as expert, and the other acting as apprentice—had only limited visual access to each other. Grosz's analysis of the corpus turns on the premise that the requirement for successful communication between expert and novice, despite their limited visual access, is a common "focus" on the task at hand. Common focus allows the exploitation of language (e.g. definite descriptions like "the screw"), materials ("the screw visible on the hub of the flywheel"), and local history ("the screw loosened previously") in instruction, while language, materials and history are used, in turn, to maintain the common focus. Grosz's concern with the interactional, linguistic and material bases of successful instruction set the stage for a series of subsequent analyses in which, by varying the resources available for communication, researchers began to explore the relationship of various non-linguistic resources to the use of language (for a 'taxonomy,' see Rubin 1980). In an analysis of transcripts of instructors communicating with an apprentice through face-to-face, telephone, audio-taped and written media about the assembly of a toy water pump, Burke (1982) finds that the most obvious difference, that between speaking and writing, is actually less crucial than the difference between interactive (e.g. a keyboard) and non-interactive (e.g. audio-taped) instruction. The restrictions generally associated with written instruction, in other words, derive not from the writing so much as from the absence of interaction, while the effectiveness of verbal instruction derives less from the speech than from the interaction that is generally associated with it.

Ochs (1979) relates the distinction of interactive and non-interactive communication to degrees of planning, arguing that in the case of interaction, speakers plan only at a general level—the concretization of the plan being contingent and emergent—whereas non-interactive discourse can be entirely planned in advance. Burke's thesis confirms that in the non-interactive modalities of instruction on the assembly task there is a tendency to "overelaborate" descriptions, in an apparent attempt to anticipate possible troubles, and to compensate for the lack of opportunities for their *in situ* clarification. In face-to-face instruction, in contrast, Burke find that instructors initially provide minimal descriptions, and then monitor the apprentice's actions for evidence of the description's adequacy or inadequacy. By telephone, where visual access was unavailable, but where the interaction remained, the resources for monitoring the actions of the apprentice changed from visual to verbal (e.g. affirmations, repeats and transformed repeats of the instructions by the apprentice), but the monitoring again guided the description.

Cohen's (nd.) analysis of transcripts of instructor and apprentice communicating by telephone or keyboard, on the same assembly task, also emphasizes the ability of instructors to adjust the level of their descriptions in response to the demonstrated understanding or misunderstanding by the apprentice. He concludes that the principle difference between spoken and written interactive media is that experts in spoken instruction more often explicitly request that the novice identify an object, and often question the novice on his or her success, while experts using keyboards subsume reference to objects into instructions for action unless some prior referential miscommunication has occurred (*ibid.*, p. 21). Spoken interaction between expert and novice, in that sense, is more finely calibrated than written, though insofar as both are interactive, both support the collaborative construction of a "useful description" of the objects and actions in question, through practical analyses of the communication's success at each turn.

#### 6.4 The basic 'interaction'

While not incorporating 'natural language understanding' on the part of the computer, the aim of the 'expert help system' analyzed in Chapter 7 is to use the power of the computer in order to combine the portability of non-interactive instructions, with the timeliness, relevance and effectiveness of interaction (see Chapter 2). Each display presented to the user by the system (numbered 1-n for purposes of analysis) either describes the machine's behavior, or provides the user with some next instructions. In the latter case, the final instruction of each display prescribes an action whose effect is detectable by the system, thereby triggering a change to the next display.

##### MACHINE PRESENTS DISPLAY 1

User reads  
instruction,  
interprets referents  
and action  
descriptions

##### USER TAKES ACTION A

Design assumes  
Action A  
means that  
user has  
understood  
Display 1

##### MACHINE PRESENTS DISPLAY 2

FIGURE 1: THE BASIC STRUCTURE OF AN INTERACTION

Through the device of display changes keyed to actions by the user, the design accomplishes a simple form of occasioned response, in spite of the fact that only a partial trace of the user's behavior is available to the system. Among those user actions that are *not* available to the system is

the actual work of locating referents and interpreting action descriptions; the system has access only to the *product* of that work. Moreover, within the instruction provided by a given display are embedded instructions for actions whose effects are not detectable by the system. To anticipate our discussion of troubles that arise, if one of *these* instructions is misconstrued, the error will go by unnoticed. Since the implication of a next display is that prior actions have been noted, and that they have been found adequate, the appearance of Display 2 will confirm the correctness not only of Action A narrowly defined, but of all of the actions prescribed by Display 1.

To compensate for the machine's limited access to the user's actions, the design relies upon a partial enforcement of the order of user actions within the procedural sequence. This strategy works fairly well, insofar as a particular effect produced by the user (such as closing a cover on the copier), can be taken to imply that a certain condition obtains (a document has been placed in the machine for copying) which, in turn, implies a machine response (the initiation of the printing process). In this sense, the order of 'turns,' and what is to be accomplished in each, is predetermined. The system's 'recognition' of turn-transition places is essentially *reactive*, i.e. there is a determinate relationship between certain uninterpreted actions by the user, read as changes to the state of the machine, and the machine's transition to a next display. By establishing a determinate relationship between detectable user actions and machine responses, the design unilaterally administers control over the interaction, but in a way that is conditional on the user.

At the same time that the system controls the interaction, the design avoids certain problems that arise when instructions are provided consecutively, in a strictly invariant order. Every procedure is represented in the system as a series of Steps, each of which has an associated precondition (the effect of a prior action by user or machine), and an associated machine response (display of instructions and/or setting of machine state). Rather than proceeding through these instructions consecutively, the system begins processing at the *last* step of the procedure and checks to see whether that step has been done. If not, the preconditions are checked and, if they are all satisfied, the step is executed. Each precondition carries with it a reference to the earlier step in the procedure that will satisfy that precondition, so that if an unmet precondition is found the system will return to the earlier step, and proceed from there. If, therefore, a procedure is repeated, but in the second instance certain conditions hold over from the first, the system will not display instructions for the actions that have already been taken. Beginning with the final step, it will work backwards through the procedure just to the point where an unmet precondition is found, and will provide the instruction from that point on. Similarly, if the user takes an action that undoes a condition satisfied earlier, the system will encounter that state again at the next check. This technique produces appropriate instructions not because the system knows that this time through differs from the last, but just because, regardless of how they come about, certain detectable conditions (e.g. a document is in the machine) are linked unequivocally to an appropriate response (e.g. initiating the printing process). Chapter 7 examines how this design strategy works and how, for the very same reason that it works in some instances, in other instances troubles arise.

#### 6.4 Methodology

The study was directed by two methodological commitments, one general and the other particular to the problem at hand.

Generally, the study began with a commitment to an empirical approach, along with the conviction that situated action cannot be captured empirically through either examples constructed by the researcher, paper and pencil observations, or interview reports. Analyses of contrived examples, observations or interviews all rest upon accounts of circumstances that are either imagined or recollected. One objective in studying situated action is to consider just those fleeting circumstances that our interpretations of action systematically rely upon, but which our accounts of action routinely ignore. A second objective is to make the relation between interpretations of action and action's circumstances our subject matter. Both objectives are clearly lost if we use reports of action as our data.<sup>41</sup>

Another approach to the analysis of instructions might be to look at the textual cogency of the instructions themselves. An example offered by Searle (1979) illustrates the problem with such a strategy:<sup>42</sup>

Suppose a man goes to the supermarket with a shopping list given him by his wife on which are written the words 'beans, butter, bacon, and bread.' Suppose as he goes around with his shopping cart selecting these items, he is followed by a detective who writes down everything he takes. As they emerge from the store both the shopper and detective will have identical lists. But the function of the two lists will be quite different. In the case of the shopper's list, the purpose of the list is, so to speak, to get the world to match the words; the man is supposed to make his actions fit the list. In the case of the detective, the purpose of the list is to make the words match the world; the man is supposed to make the list fit the actions of the shopper. This can be further demonstrated by observing the role of 'mistake' in the two cases. If the detective goes home and suddenly realizes that the man bought pork chops instead of bacon, he can simply erase the word 'bacon' and write 'pork chops.' But if the shopper gets home and his wife points out that he has bought pork chops when he should have bought bacon he cannot correct the mistake by erasing 'bacon' from the list and writing 'pork chops.' (*ibid.*, p.4)

41. This is not to say that paper and pencil observations do not have their place. The precursor to the current study (see Suchman 1982) began with a period of approximately 20 hours of observation of new users of the same machine—minus the 'expert help system,' but equipped with written instructions—in actual office settings. That earlier study was undertaken in response to an unelaborated report, from those who supported the machine and its users "in the field," of user complaints that the machine was "too complicated." Given the relative simplicity of even the most complex photocopier, this complaint on face value was puzzling, particularly to the machine's designers. The combination of the vagueness of the complaint as reported, and the bewilderment of the designers, intrigued both me and certain of my co-workers at the research center, and we set about to try to ascertain what the "complexity" was really about. That led to the observations in "real" offices, which convinced me that indeed the machine was "too complicated" for the novice user who had no previous training; i.e. people trying to use the machine were visibly confused. The methodological problem at that point was that I, as an observer of their troubles, was equally confused. From the observations, therefore, I learned two important lessons. First, that there was indeed a problem. And second, that to understand the problem would require the use of an adequate, i.e. a videotaped, record.

42. Searle credits this example to Elizabeth Anscombe (1957). The point that Searle is interested in concerns the notion of 'direction of fit' between words and the world.

The subject of the present analysis, the user of the 'expert help system,' is in the position of the shopper with respect to the instructions that the system provides; that is, she must make her actions match the words. But in what sense? Like the instructions, a shopping list may be consulted to decide what to do next or to know when the shopping is done, may be cited after the fact to explain why things were done the way they were, and so forth. But also like the instructions, the list does not actually describe the practical activity of shopping (how to find things, which aisles to go down in what order, how to decide between competing brands, etc.); it simply says how that activity is to turn out.

Just as the list of the shopping's outcomes does not actually describe the organization of the activity of shopping, an analysis of instructions will not yield an analysis of the activity of carrying them out. In fact, contrary to the case in the story, there's no reason to believe that if a person has a set of instructions for operating a machine, and we generate a description of the activity of operating a machine from watching the person, that the description we generate should look anything like the instructions. In fact, if our description of the situated activity does mirror the structure of the instructions, there is reason to believe that something is amiss.

Unlike the detective in the story who is supposed to generate a list, our problem as students of situated action is more akin to the problem of a detective who is just sent out and told to report back on what going to the grocery store is all about and how it is done. What that description should look like—what its terms should be, what its structure should be, what of all that goes on it should report—is an open methodological question. If, in order to put some constraints on the description, we set out with a template that asks for a list just of what the actions come to, then what counts as "an action" is prescribed ahead of time as "its outcome," and the list format prescribes the structure of the description. Only that part of the activity that fills in the template will be recorded. The action's structure, in other words, will be decided in advance, and the method employed by the scientist will ensure that that structure is what is found.

One further issue that the story touches on is the problem of validity. The story says the detective might "suddenly realize" that there is some error in his description. But how might he actually realize that? If we just look for a discrepancy between the shopper's list and the detective's, what we find might reflect either an error in the shopper's activity, (it doesn't match the list) or in the description (it doesn't match the activity). In order to evaluate which, we must have a) independent access to the shopper's list, to compare against the activity; and b) a record of the activity. That is to say, two essential, methodological resources are a) the comparison of our own interpretations with those of our subjects, and b) a record that is not contingent on either.

However adequate the record, of course, the empiricism of social studies is not a positivist one because we cannot, by definition, provide a literal description of our phenomenon:<sup>43</sup>

---

43. Galaty (1981) makes a useful distinction between "data sources," as the business of the social world independent of the anthropologist's interest in it; "data," as the anthropologically processed information that appears in the form of, e.g. transcripts; and "analytic objects," conceptualized as events, troubles, and the like (note 2, p. 91). The point is that for the social scientist, the data is interpreted already at its source.

Any description of a phenomenon is based on perceived features that the phenomenon displays to the observer. A literal description, then, amounts to asserting that on the basis of those features the phenomenon has some clearly designated property, or what is logically the same thing, belongs to some particular, well-defined class of phenomena. (Wilson 1970, p. 72)

In order for a description to be literal, in other words, the class of phenomena of which the described is an instance must be definable in terms of sufficient conditions for counting some instance as a member of the class. For situated action, that would require classification of action not only as the relation of intent to behavior, but as the relation of both to mitigating circumstances—a classification which, I argued in chapter 4, is functionally and criterially different from that applied to intentional descriptions of actions and situations. Moreover, the social scientist's description is yet another order of remove from a literal description if the subject of the description is not only the intent of some actor, but the interpretations of that actor's intent by others on the scene. Judgments of correctness and veridicality for literal description are replaced in social studies by judgments of adequacy or verisimilitude for interpretive descriptions (Heap 1980, p. 104), the latter resting on criteria of evidence and warranted inference rather than conditions of truth.

The problems that social science struggles with in defining its methodology are the same outstanding problems that constitute its subject matter, i.e. the uncertain relation between accounts of the significance of action, and the observations and inferences on which those accounts must be based. That is to say, there is no privileged analytic stance for the social scientist, that exempts him or her from the problem of the practical observability and objectivity of the social world. The only advantage that accrues to the researcher—a substantial one, it turns out—is recourse to a record of the action and its circumstances, independent of the researcher's analysis. The availability of the audiovisual technology to provide such a record, for repeated inspection by the researcher and by colleagues, avoids both the reliance on unexplicated resources that characterizes traditional ethnographic accounts (where, given the fleeting nature of the events that the ethnographer describes, the only data available for inspection by others is the ethnographer's description), and explications of method that rely on introspection, or on reconstructed accounts of the analysis' production.

This study proceeded, therefore, in a setting where video technology could be used in a sort of uncontrolled experimentation. On the one hand, the situation was constructed so as to make certain issues studyable, specifically the work of using the machine with the assistance of the 'expert help system.' The construction consisted in the selection of tasks observed to pose problems for new users in "the real world." On the other hand, once given those tasks, the subjects were left entirely on their own. In the analysis, by the same token, the goal was to construct a characterization of the "interaction" that ensued, rather than to apply a predetermined coding scheme. Both predetermined coding schemes and controlled experiments presuppose a characterization of the phenomenon studied, varying only certain parameters to test the characterization. Application of that methodology to this study would be at least premature. The point of departure for the study

was the assumption that we lack a description of the structure of situated action. And because the hunch is that the structure lies in a relation between action and its circumstances that we have yet to uncover, we don't want to *presuppose* what are the relevant conditions, or their relationship to the structure of the action. We need to begin, therefore, at the beginning, with observations that capture as much of the phenomenon, and presuppose as little, as possible.

The consequence of this commitment to examining the circumstances of action is that we need to begin with a record of actual events, which is not pre-judged as to its analytic interest either in advance or in the making. The data for this study, accordingly, are a corpus of videotapes of first-time users of the 'expert help system.' First-time users were chosen on the grounds that the system was intended by its designers to be self-explanatory, or usable by people with no previous introduction to the machine. More generally, the troubles encountered by first-time users of a system are valuable in that they disclose work required to understand the system's behavior that, for various reasons, is masked by the proficient user.<sup>44</sup>

In some cases two people, neither of whom had ever used the system before, worked together in pairs. While this was not a stipulation for participation in the study, it was encouraged and found to be a useful arrangement. Two people asked to collaborate in using a relatively simple machine like a photocopier are faced with the problem of doing together what either could do alone. In the interest of the collaboration, each makes available to the other what she believes to be going on: what the task is, how it is to be accomplished, what has already been done and what remains, rationales for this way of proceeding over that, and so forth. Through the ways in which each collaborator works to provide her sense of what is going on to the other, she provides that sense to the researcher as well. An artifact of such a collaboration, therefore, is a naturally generated protocol.<sup>45</sup>

A second methodological commitment arose from the particular problem of looking at human-machine communication, and directed the analysis itself. The aim of the analysis was to find the sense of 'shared understanding' in human-machine communication. More particularly, I wanted to compare the user's and the system's respective views of the situation, over a sequence of events. In working to organize the transcripts of the videotapes, therefore, I arrived at the following simple framework:

---

44. This is the value of studying interactional troubles generally (cf. Gumperz 1982b, p. 308), and distinguishes my analysis from the usual "operability tests;" i.e. I am not interested in 'correct' or 'erroneous' moves by the user, but rather by studying what things look like when they are unfamiliar, I hope to understand better what is involved in their mastery.

45. Brown, Rubenstein and Burton (1976) argue persuasively for the use of teams to generate protocols, where the discussions and arguments that unfold are treated as evidence for the individual reasoning of the participants. The actions of the team members can also be viewed as organized by the task of collaboration itself, however, although in the interest of looking at the interaction of *both users* with the *machine*, I have deliberately avoided taking that view here. It is worth noting, in this regard, that analyses of "discourse" undertaken in the interest of building interactive AI systems generally tend to view communication as the coincidence of individual reasoning processes, rather than as an activity with a distinctive character arising from the collaboration itself.

THE USER		THE MACHINE	
I	II	III	IV
Actions not available to the machine	Actions available to the machine	Effects available to the user	Rationale

FIGURE 2: THE ANALYTIC FRAMEWORK

The framework revealed two initial facts about the relationship of user and system. First, it showed that the coherence of the user's actions was largely unavailable to the system, and something of why that was the case. Beginning with the observation that what the user was trying to do was, somehow, available to me as the researcher, I could ask how that was so. The richest source of information for the researcher, as a fully-fledged, "intelligent" observer, is the verbal protocol, recorded in Column I. In reading the instructions aloud, the user locates the problem that she is working on. Her questions about the instructions identify the problem more particularly, and further talk provides her interpretations of the machine's behavior, and clarifies her actions in response.

A second, but equally crucial resource is visual access to the user's actions. Of all of her actions, one could clearly see the very small subset, recorded in Column II, that were actually detected by the system. From the system's "point of view," correspondingly, one could see how it was that those traces of the user's actions available to the system—the user's behavior seen, as it were, through a key-hole—were mapped onto the system's plan, under the design assumption that, for example, button *x* pushed at this particular point in the procedure must mean that the user is doing *y*.

The framework proved invaluable for taking seriously the idea that user and machine were interacting. By treating the center two columns as the mutually available, behavioral 'interface,' one could compare and contrast them with the outer columns, as the respective interpretations of the user and the design. This comparison located precisely the points of confusion, as well the points of true intersection or 'shared understanding.' Both are discussed at length in the next chapter.

## 7. Human-machine communication

Interaction is always a *tentative* process, a process of continuously testing the conception one has of...the other. (Turner 1962, p. 23)

In Chapter 4, I outlined the view that the significance of actions, and their intelligibility, resides neither in what is strictly observable about behavior, nor in a prior mental state of the actor, but in an interactionally constructed relationship between observable behavior, circumstances and intent. Rather than enumerate an *a priori* system of shared rules for meaningful behavior, Chapter 5 described resources for constructing shared understanding, collaboratively and *in situ*. Face-to-face interaction was presented as the most powerful and highly developed system for accomplishing mutual intelligibility, exploiting a range of linguistic, observational and inferential resources.

Given this view of the source of action's intelligibility, the situation of action can be defined as the full range of resources that the actor has available to convey the significance of his or her own actions, and to interpret the actions of others. Taking that preliminary definition of the situation as a point of departure, my interest in this chapter is to consider 'communication' between a person and a machine in terms of the nature of their respective situations. For purposes of the analysis, and without ascribing intent in any way, I will assume that the machine is behaving according to the resources of 'its' situation, the user according to the resources of hers. The aim of the analysis then is to view the organization of human-machine communication, including its troubles, in terms of constraints posed by asymmetries in the respective situation resources of human and machine.

For the case considered here, we can assume that the situation of the user comprises preconceptions about the nature of the machine and the operations required to use it, combined with evidence found in and through the actual course of its use. The evidence found is both planned and fortuitous, consisting in information that the designer provides about the machine, and in the machine's observable behavior. The situation of the 'expert help system,' in contrast, comprises a plan for the use of the machine written by the designer and implemented as the program that determines the machine's behavior, and sensors that register changes to the machine's state, including some changes produced by the user's actions. The design plan defines what constitutes "intelligible action" by the user insofar as the machine is concerned, and determines what stands as an appropriate machine "response." The intersection of the situations of user and machine is the locus both for successful exploitation of mutually available resources, and for problems of understanding that arise out of the fundamental asymmetry of their two situations.

### 7.1 The conditional relevance of the machine's response

The general problem that the designer of an 'interactive' system must somehow contend with is how to ensure that the machine responds appropriately to the user's actions. As in human communication, an appropriate response requires an adequate interpretation of the action's significance. The adequacy of a given interpretation is judged indirectly, by the response that the

other makes to actions taken under that interpretation, and by the usefulness of the interpretation in understanding the others' further actions. This highly contingent process is precisely what we call interaction.

For purposes of this analysis, we can begin by considering two propositions about the basis for human interaction, each of which has a somewhat different implication for the project of designing an interactive machine:

*Proposition A:* A relevant response *anticipates* the other's actions.

*Proposition B:* A relevant response is *occasioned* by the other's actions.

Proposition *A* suggests that an interactive interface should be based on a model of the user that supports the prediction of actions, the specification of recognition criteria for the actions predicted, and the prescription of an appropriate response. Proposition *B* suggests that an interactive interface should maximize sensitivity to actions actually taken, by minimizing predetermined sequences of machine behavior. The former suggestion is constrained by limitations on the designer's ability to predict the user's actions, the latter by limitations on the system's access to the actions taken, its ability to draw relevant inferences about their significance, and its ability to construct a relevant response.

The design strategy in the system examined here is to try to provide the *effect* of *B*, through the *use of A*. That is to say, the designer predicts some intent on the part of the user, of the form "use the machine to accomplish outcome *x*." Tied to that statement of intent are a set of instructions that prescribe the actions to be taken, at a level of generality designed to ensure their relevance to any user, whatever the details of her actual situation. Ideally, the instructions tell the user what features of her actual situation are relevant for the machine's operation. By finding and/or producing those features, the user anchors the instructions to her particular circumstances.

This chapter looks at some of the consequences of taking a statement of intent and a presumed plan as grounds for the interpretation of situated action. To anticipate, the result of that strategy is an insensitivity to actual circumstances that is both the system's central resource, and its fundamental problem. I look first at the system's resources for construing the actions of the user; namely, *plans and states*. I then consider the problems posed for the designer by the user's principal resource, organized under the general rubric of *situated inquiries*, and by the user's ability to find the relevance of the system's response to those inquiries. Finally, I look at two cases of communicative trouble, the *false alarm* and the *garden path*. Chapter 8 concludes with implications of the analysis for a general account of mutual intelligibility, and for the particular requirements on the design of artifacts that would interact with their users.

## 7.2 The system's situation: Plans and detectable states

I have said that the situation of the 'expert help system' comprises a program that controls its behavior, and sensors that register certain changes to its state effected by actions of the user. Initially, the user's response to a series of questions is taken as a statement of intent, which determines selection of an appropriate plan. The plan is then presented to the user in the form of a step-wise set of procedural instructions. The designer assumes that in following the procedural instructions, the user effectively is engaged in carrying out the plan.

The design premise is further that as the user takes actions prescribed by the instructions, the actions will change the state of the machine in predetermined ways. By treating those changes to the machine's state as traces of the user's actions, the designer can specify how the user's action is to be recognized by the system, and how the system is to respond. The strategy of tying certain machine states to particular machine "responses" enables the appearance of instructions occasioned by the user's actions, as in the following example of a successful 'interaction':<sup>46</sup>

(Numbers in brackets identify tape and location of the sequence in a transcript. "Quotes" indicate that the subject is reading instructions from the display. See appendix for machine diagram and displays.)

I [22:7-68, 189-196]

(S's are proceeding from the display that establishes their goal as making two-sided copies of a bound document. Two-sided copying requires an unbound document, so they must begin by making a master unbound copy of their document, using the "Bound Document Aid," or BDA.)

	THE USER'S ACTIONS	THE MACHINE'S BEHAVIOR	
	I Not available to the machine	II Available to the machine	III Available to the user
			IV Rationale
			DISPLAY 1
S1	"To access the BDA, pull the latch labelled Bound Document Aid":: [Both S's turn to machine]		Setting control panel
		DISPLAY 2	
			Instructions for copying a bound document: Raising the document handler.

46. One way of viewing the interaction of S1 and S2 in the following sequence is as the adept completion of what the design attempts. Specifically, S1 decomposes and re-presents the instructions provided by the system, such that they are fit precisely to S2's actions in carrying them out. S1 is able to do this because of her attunement to what S2 is doing, including S2's troubles.

THE USER'S ACTIONS		THE MACHINE'S BEHAVIOR	
I	II	III	IV
	Not available to the machine	Available to the machine	Available to the user
S1	[Points] Right there.		
S2	[Hands on latch]		
S1	"And lift up to the left." [looks to S2, who struggles with the latch]		
	"Lift up and to the left."		
S2	[Still struggling]		
S1	Okay::		
S2	Pu::ll, and lift up to the left. [Looks at picture] Oh, the whole thing		
S1	Yea.		
S2	lift up and to the left.		
	RAISES DOCUMENT HANDLER		

## DISPLAY 3

S1	"Place your original face down, [Passes journal to S2] on the glass, centered over the registration guide." [looks to machine] Got that? [pause] Want me to read it again?	Instructions for placing document and closing document cover.
S2	Um:: I'm just trying to figure out what a registration guide is, but I guess that's this, um:	
S1	[Looking over her shoulder] Yea:	
S2	centered over this line thingy here.	

	THE USER'S ACTIONS	THE MACHINE'S BEHAVIOR		
	I Not available to the machine	II Available to the machine	III Available to the user	IV Rationale
S1	Okay, let me read it again. "Place your original face down on the glass, centered over the registration guide, to position it for the copier lens." Okay?			
S2	'Kay.			
S1	Okay. "Slide the document cover left over your original, until it latches."			

[Portion omitted, in which they first mis-locate, then locate, the document cover]

#### CLOSES COVER

DISPLAY 4

- |    |  |                                      |
|----|--|--------------------------------------|
| S2 | Okay, now,                             | Instructions<br>to start<br>printing |
| S1 | All right:<br>"Press the Start button" |                                      |

#### SELECTS START

The system presents to the user a series of displays, composed of text and drawings, that either describe the machine's behavior, or provide the user with some instructions for action. In the latter case, the final instruction of each display prescribes an action whose effect is detectable by the system, thereby initiating the processing that produces the next display. Below is the procedure from Sequence I, as specified by the designer to the program that controls the display of instructions to the user:

Step 1: Set Panel

[DISPLAY 1]

Step 2: Tell User "To access the BDA ... Raise the RDH"

[DISPLAY 2]

Step 3: Tell User "Place original face down ... Slide document cover left"

[DISPLAY 3]

Step 4: Make Ready.

Step 5: Tell User "Press Start". Requirements:

Panel Set (If not, try Step 1)

RDH raised (if not, try Step 2)  
 Document cover closed (If not, try Step 3)  
 Ready State (If not, try Step 4)

[DISPLAY 4]

Step 6: Complete printing Step (Sets CopiesMade) Requirements:  
 Printing State (If not, try Step 5)

The "Requirements" represent those features of the system's situation—i.e. of the system's own state—that are resources for determining an appropriate next action or response. Rather than proceeding through the steps of the procedure consecutively, the system starts with the *last* step of the procedure, Step 6 in this case, and checks to see whether it is done. A step is done if a check of the machine's state confirms that the conditions represented by that step's requirements have been met. When a requirement is found that is not met, a further set of specifications, tied to that requirement, send the system back to an earlier step in the procedural sequence. The system then displays the instructions tied to that earlier step to the user, until another change in state begins the same process again. Each time the user takes an action that changes the machine's state, in other words, the system compares the resulting state with the end state, returns to the first unfinished step in the sequence, and presents the user with the instructions for that step.

This device of working backward through the procedure is designed to avoid the presentation of redundant instructions. In II, having discovered that their original is larger than standard paper, the S's decide to re-do the job. They return to the job specification display to select the reduction feature, and then direct the machine to proceed:

II [22:223-255]

(Again S's are making two-sided copies of a bound document, this time with reduction.  
*The document is still on the copier glass, the document cover is closed.*)

THE USER'S ACTIONS		THE MACHINE'S BEHAVIOR	
I	II	III	IV
Not available to the machine	Available to the machine	Available to the user	Rationale

DISPLAY 1

- |    |  |               |
|----|--|---------------|
| S2 | It's supposed to—<br>it'll tell "Start",<br>in a minute.             | Setting panel |
| S1 | Oh. It will?   |               |
| S2 | Well it did:<br>in the past.<br>[pause]<br>A little start: box will: |               |

THE USER'S ACTIONS		THE MACHINE'S BEHAVIOR	
I	II	III	IV
	Not available to the machine	Available to the machine	Available to the user
S2	There it goes.		DISPLAY 4
S1	"Press the Start button"		Ready to print
		SELECTS START	
	Okay.		STARTS

On this occasion the system bypasses the instructions to raise the document handler, place the document on the glass, and close the document cover, all of which are irrelevant in that the actions they prescribe have already been taken. The system is able to respond appropriately because a detectable machine state (the closed document cover) can be linked by the designer to an *a priori* assumption about the user's intent with respect to a next action (ready to press start). As a result, the system can be engineered to provide the appropriate next instruction *in spite of* the fact that it does not actually have access to the history of the user's actions, or even to the presence or absence now of a document on the glass. The result is that while S1 predicts the system's behavior—specifically, that it will provide them with a "Start button"—on her recollection of an occasion (sequence I) on which the system actually behaved somewhat differently, her prediction holds. That is, just because on this occasion a relevant feature of the user's situation is accessible to the system, and its behavior *changes* accordingly, it appears to behave in the 'same' way. In human interaction, this graceful accomodation to changing circumstance is expected, and largely taken-for-granted. The success of the system's accomodation in this instance is evident in the accomodation's transparency to the users.

On other occasions, however, the inference from a machine state to an *a priori* assumption about the user's situation, on which the success of Sequence II rests, leads to trouble. I have said that given a statement of the user's goal (derived from the selections made on DISPLAY 0), the system initiates a plan, and then tracks the user's actions by mapping state changes to a step-wise procedure bound to that plan. In the following case, the S's have completed the unbound master copy of their document, and have gone on to attempt to make their two-sided copies. The order of pages in the copies are found to be faulted (a fault not available to the system, which has no access to the actual markings on the page), so they try again. As in II, for the S's this is a second attempt to accomplish the same job, while for the machine it is just another instance of the procedure. On this occasion, however, that discrepancy turns out to matter:

## III [22:582-608]

(Again making two-sided copies from a bound document, but this time having already completed their unbound master copy.)

	THE USER'S ACTIONS		THE MACHINE'S BEHAVIOR	
I	II	III	IV	
	Not available to the machine	Available to the machine	Available to the user	Rationale
S2	Okay, and then it'll tell us,  okay, and:: It's got to come up with the little Start thing soon. [pause]		DISPLAY 1	Setting panel
	Okay, we've done all that. We've made our bound copies. [pause]		DISPLAY 2	Instructions for copying a bound document: Raising the document handler.
S1	It'll go on though, I think. Won't it?			
S2	I think it's gonna continue on, after it realizes that we've done all that.			

In sequence II, the system's ignorance of the relation between one time through and the last did not matter, just because a check of the current state of the machine caused the appropriate behavior. Or, more accurately, the "current state" of the machine could be read as a local, technical matter independent of the embedding course of events. Here, however, a check of the current state belies the users' intent. To appreciate what they are doing now requires that one recognize the relation between this time through and the last, and the system does not. So while both users and system are, in some sense, doing the job again, there are two different senses of what, at this particular point, it means to do so. As far as the users are concerned, they are still trying to make two-sided copies of a bound document, so they leave their job description as such. For the machine, however, the description of their intent this time, having made their master copy, is two-sided copying from an *unbound* document. The result is that what they tell the machine they are doing is not what they intend to do, and what they intend to do is not available from the current state of the world as the machine is able to see it.

They find evidence of this trouble in an extended silence, which suggests that the system is not going to proceed (see also 7.4.2):

IV [22:742-825]

(Continued from III)

THE USER'S ACTIONS		THE MACHINE'S BEHAVIOR	
I	II	III	IV
	Not available to the machine	Available to the machine	Available to the user
S1	(8.0) Then again, maybe we need to change the task description.		DISPLAY 2
S2	What do you think?	Selects "Change"	DISPLAY 0
S1	No.		User may want to change job description.
S2	Okay, "Proceed."	Selects "Proceed"	DISPLAY 1
S1	Maybe I should just lift it up and put it =		DISPLAY 2
S2	How do we skip this then?		Making two-sided copies from a bound document.
S1	=down again. Maybe it'll think we're done.		Raising the document handler
S2	(laughs) Oh, Jean.		
S1	There.	Opens BDA	DISPLAY 3
	Okay, we've done what we're supposed to do.	Closes BDA	Instructions for placing document
	Now let's put this down. Let's see if that makes a difference. (looks back to display)		
S1	(laughs) It did something.		DISPLAY 2
			Instructions for copying a bound document

THE USER'S ACTIONS		THE MACHINE'S BEHAVIOR	
I	II	III	IV
	Not available to the machine	Available to the machine	Available to the user
S2	(inaudible) Good grief.		Rationale
S1	Oh, it's still telling us we need to do a bound document. And we don't need to do the bound document because we've done that. You know, maybe we ought to go back to the beginning, and erase that thing about the bound document.		
S2	Okay, that's a good idea.	Selects "Change"	
S1	Then say, "Is it bound?" just put no.	DISPLAY 0	User may want to change job description
S2	Not anymore.	Selects "No"	
S1	And then everything else is constant, isn't it? It's on standard paper: [		
S2	so we'll proceed.	Selects "Proceed"	
S1	So let's just proceed.		New job; two-sided from unbound document.

What the S's discover here is that, from the system's "point of view," their situation is determined by their statement of intent. The significance of a given action can only be assessed with reference to the action's location in a developing course of events, and the statement of their intent and presumed plan is meant to provide that reference. Statements of intent, however, are inevitably embedded in larger purposes, and embed smaller ones. While their initial statement still accurately describes their global purpose, it belies their local one. Nor in this instance is their current situation—the result of their previous actions—reflected in the system's current state. Their

current situation is available only through a history of which the system has no record, or through their reports and assertions about their situation, to which the system has no access.<sup>47</sup>

### 7.3 The user's resource: The situated inquiry

The premise of a self-explanatory machine is that its users will discover its intended use through information found in and on the machine itself. In physical design, the designer anticipates certain questions such that, in the event, an answer is there ready-at-hand. So, for example, the user's question "Where do I grab?" is answered by a handle fitted to the action of grabbing. In the traditional instruction manual, some further classes of inquiry are anticipated, and answers provided. The step-wise instruction set addresses the question "What do I do next?", and the diagram, "Where?" In all cases, the questions anticipated and answered must be those that any user of the system might ask, and the occasion for both questions and answers is found by the user.

For the novice engaged in a procedural task, the guiding inquiry is some form of the question "What next?" The question is an essentially indexical one, relying for its significance on the embedding situation. In the case at hand, the system effectively checks its own state to anticipate the user's question, and then presents the next outstanding requirement of the selected plan in response. This design strategy assumes that the job specification represents the user's intent, that the intent so represented determines the appropriate plan, and that user and system are engaged in carrying out the procedure for that plan.

#### 7.3.1 "Meta" inquiries

The design premise, in other words, is that the embedding situation for the question "What next" is just the procedure, and that the question is just a request for the next step. As long as the premise holds, the presentation of a next instruction constitutes an appropriate response (see, for example, sequence I). The design premise fails, however, in cases where the question "What next" is not a matter of proceeding with the current plan, but of its abandonment or repair:

---

47. Their attempt to make their situation accessible to the system by exploiting its insensitivity to their actual circumstances, and "faking" the required action fails, but the failure is a failure in performance, not in principle. Specifically, if they had opened and closed the *document cover*, rather than only the Bound Document Aid, they would have achieved the desired effect.

V [203:809-845]

(S's are making 5 two-sided copies of a bound document. They first must make a single, unbound master copy of their original.)

	THE USER'S ACTIONS	THE MACHINE'S BEHAVIOR	
	I	II	III
	Not available to the machine	Available to the machine	Available to the user
S1	"Instructions. Slide the document cover to the right."		DISPLAY 5
S2	(noting output)  Okay, it gave us one copy here.		
S1	Okay, "Slide the document cover right to remove the original."		
S2	We're supposed to have 5 copies and we only got one.		
S1	(looks to output) Oh. (looks to display) We only got one?		
S2	Yea.		
	(long pause)		
S1	What do we do then?		
	(long pause, both study display)		

This sequence is discussed at length in 7.5.1. For the moment, the observation is simply that the question "What do we do then" is not, in this instance, a simple request for a *next* in the sense of a next step in the procedure, but rather is a request for a remedy. The situation of the inquiry (indicated anaphorically by the "then," viz. "given that we were supposed to have 5 copies and we only got one") is not the procedure itself, but the conflict between the apparent outcome of the procedure (a single copy), and their stated intent (five copies). That situation, while clearly described by S2, is unavailable in the current state of the machine, which shows no evidence of

their trouble.<sup>48</sup> As a consequence of the fact that the situation of their inquiry is not that which the design anticipates, and is not otherwise accessible to the system, the answer that the system offers—do the next step in this procedure—is inappropriate.

Even in a case where the designer anticipates an inquiry having to do with the procedure itself rather than the next action, the situation of the request may be problematic:

VI [210:237-304]

(S's are making two-sided copies of a bound document. In response to the instruction to close the "document cover" (DISPLAY 3), they have mistakenly closed the entire "Bound Document Aid" or BDA instead, and as a consequence have returned to the previous instruction to open the BDA (DISPLAY 2).)

	THE USER'S ACTIONS	THE MACHINE'S BEHAVIOR		
	I	II	III	IV
	Not available to the machine	Available to the machine	Available to the user	Rationale
S1	"Pull the latch labelled—" "We did that." "Raise—" "We did that." (studying display) Okay. Okay.		DISPLAY 2	Instructions for copying a bound document
S2	"Lift up on the latch," "We did that."			
S1	Now let's change::			
S2	"Change task description?"			
S1	Yes.			
S2	Selects "Change"		DISPLAY 0	User may want to change job specs
	"Describe the document to be copied—" "Oh, we already did." "No, we don't want to do that."			
S1	Maybe we have to do it to copy that. (next page)			

48. That is, the current state of the machine indicates just that a copy has been made, the design rationale being that they have copied the first page of an *unbound master copy* of their bound document, and are ready to go on to the second page.

	THE USER'S ACTIONS		THE MACHINE'S BEHAVIOR	
I	II	III	IV	
	Not available to the machine	Available to the machine	Available to the user	Rationale
S2	(looks around machine (laugh) I don't know.			
S1	Well::			
S2	"Help" (laugh) "Select the question you would like help with."	Selects "Help"		User needs clarification of DISPLAY 0
S1	I guess we still do have to=			
S2	We still ha—			
S1	=answer this.			
S2	Oh, okay. Alright.			
S1	Okay.			
S2	We sti— but we did all that, didn't we?			
S1	Well, maybe not for this page.			

Their selection of "Change task description," in the context of a loop between DISPLAY 2 and DISPLAY 3, and their subsequent surprise at the re-appearance of DISPLAY 0 in response, suggests that the intent of their action was not to return to the job specification, but to find a *next* instruction. The fundamental ambiguity between any next instruction as either a continuation, or as the initiator of a repair, is discussed at length in 7.4. Our interest here is in the situation of the request for "help" that follows the return to DISPLAY 0. Specifically, the request for "Help" is a question *about that return* to DISPLAY 0, and the larger problem of the loop in which it is embedded. The design, however, takes the situation of the request to be a local one, *viz.* as having to do with interpreting the contents of DISPLAY 0 itself.

### 7.3.2 The request for clarification

Tied to the guiding inquiry "What next" is a set of subordinate questions about prescribed actions—questions that look for clarification of the forms "How," "Where" or "To what," and "Why."<sup>49</sup> The system's responsiveness to requests for elaboration turns again on the adequacy of the designer's prediction:

VII [202:13-33]

(S's are making two-sided copies of a bound document. They first must make a single, unbound master copy using the "Bound Document Aid," or BDA.)

	THE USER'S ACTIONS	THE MACHINE'S BEHAVIOR	
	I	II	III
	Not available to the machine	Available to the machine	Available to the user
			Rationale
S1	"You need to use the Bound Document Aid to make an unbound copy of your original." Where is—  Oh, here it is.	DISPLAY 1  DISPLAY 2	DISPLAY 1  Overview  Instructions for copying a bound document: picture of the BDA.

S1's question is actually interrupted by the change to DISPLAY 2, which anticipates that very question. In this instance, it happens that the display change is timed to the mechanism that sets the machine's control panel, rather than being conditional on any action of the user's. Ironically, just because on this occasion the system's behavior is determined by the internal processing of the system, *rather than* by the user's actions, it appears that the system's behavior is occasioned by the user's question.

The fact that the question anticipated turns out to be the user's question in this instance marks the success of the design. In the following sequence, however, the designer's prediction fails;

49. From the standpoint of the actor concerned with a procedural *next*, the other two logically possible queries, *viz.* "By whom" and "When," are already answered by the embedding situation. Though see Sequences XVIII and XXIV below.

## VIII [202:116-133]

(S's are making two-sided copies from a bound document. They have placed their document on the document glass.)

	THE USER'S ACTIONS	THE MACHINE'S BEHAVIOR	
	I	II	III
	Not available to the machine	Available to the machine	Available to the user
DISPLAY 3			
S1	Okay, wait a minute. "Slide the document cover left over your original until it latches." (looks to machine)		Copying a bound document: Closing the document cover
S2	(grasps BDA)		
S1	The document cover— (leans over to look in BDA)		
S2	Oh. (pulls on document feeder belt, which gives a little) No, no, no. (indicating entire BDA) <i>This would be the document cover, isn't it?</i>		
S1	"To provide an eyeshield for the copier (inaudible)."		

In this case, the designer anticipates a question regarding the *motivation* for the action, while the user's problem is with the action's *object*. In another instance, the question *what is the object* is anticipated, while the question actually asked concerns *how to do the action*:

IX [221:504-532]

(S's are making two-sided copies of an unbound document)

	THE USER'S ACTIONS		THE MACHINE'S BEHAVIOR	
I	II	III	IV	Rationale
	Not available to the machine	Available to the machine	Available to the user	DISPLAY 10
S1	"Place the copies: on the top paper tray."  [portion omitted in which they locate the tray]			Beginning second pass of two- sided copies
S1	Okay.			
S2	But, (turning back to display) <i>How do you do that?</i>			
S1	(looking at diagram) <i>"The top paper tray is to the right of the output tray."</i> Place copies in the top paper tray," (turning to tray) Oh, you just lift it up. (does)			

The answer to S2's inquiry about how to do the action in this case is found not in the instruction, which *locates* the object, but in the object itself. In another instance, similarly, a problem in interpreting an instruction is solved through a picture of the object on which the action is to be performed:

X [220:22-42]

(S's are making two-sided copies of a bound document.)

THE USER'S ACTIONS		THE MACHINE'S BEHAVIOR	
I	II	III	IV
Not available to the machine	Available to the machine	Available to the user	Rationale

DISPLAY 2

- S1 "To access the BDA,  
pull the latch labelled  
Bound Document Aid":  
[Both S's turn to  
machine]
- S1 [Points] Right there.
- S2 [Hands on latch]
- S1 "And lift up to the  
left."  
[looks to S2, who  
struggles with the  
latch]  
"Lift up and to the  
left."
- S2 [Still struggling]
- S1 Okay:
- S2 Pu::ll, and lift up  
to the left.  
[Looks at picture]  
Oh, the whole thing  
[  
S1 Yea.  
S2 lift up and to the  
left.
- Opens BDA

When the object that S2 first takes to be implicated in the action description "lift up and to the left" resists her attempts to perform the action described, and the description suggests no other interpretation of the *action*, she finds in the picture a different *object*. That re-interpretation of the object, in its turn, revises the significance of the action description. A conflict between the action on an object described by an instruction, and the action required by the object itself, can be a resource for the identification of trouble in the interpretation of an instruction, and its resolution:

XI [202:487-506]

(S's have mistaken the entire "Bound Document Aid" for the "document cover", and are caught in a loop between DISPLAY 3 and DISPLAY 2 (see Sequence VI))

	THE USER'S ACTIONS I	THE MACHINE'S BEHAVIOR II	III	IV
	Not available to the machine	Available to the machine	Available to the user	Rationale
DISPLAY 3				
S2	Okay. "Slide the document cover— left over your original, until it latches. (turns to machine) You know it says "slide"— this (finds document cover). Okay.			
S1	Ohh.			
S2	(laughs) Ohh, isn't that hilarious? Okay.			
S1	Okay.	Closes cover		
DISPLAY 4				
	It was something else that was supposed to go over that.			

In general, the relationship of instructions to the actions and objects they describe is reciprocal, rather than directional (cf. Burke, 1982). That is to say, while instructions answer questions about objects and actions, they also pose problems of interpretation that are solved in and through the same objects and actions that they reference:<sup>50</sup>

50. Burke's pump assembly task provides an interesting case in that to some extent, the necessary information for the assembly task is discoverable in the materials themselves, specifically the 'fit and stay' bindings of one component of the pump to another. At the same time, Burke noted a difference in confidence between those students who had linguistic instruction and those who did not, the former using the instructions, on the one hand, and the task actions and materials, on the other, as mutually informative, such that:

Both the instructions and the task actions are treated by the apprentice as problems to be solved. But each is used as a resource to solve the other as a problem (*ibid.*, p. 178).

XII [210:139-162]

(S's are making two-sided copies of a bound document.)

THE USER'S ACTIONS		THE MACHINE'S BEHAVIOR	
I	II	III	IV
Not available to the machine	Available to the machine	Available to the user	Rationale

DISPLAY 2

S2 "To access the BDA, pull the latch labelled Bound Document Aid."

(both turn to machine.)

S1 (takes hold of latch.)

S2 Pull it down:  
just push it down.

S1 (does, BDA starts to open)

S2 (startled)  
Oh, alright.

Opens BDA

DISPLAY 3

This is what you do.

S1 Is this what you do?  
Oh my gosh.

In this case, rather than the interpretation of the instruction being *prerequisite* to the action's execution the action, *after* the fact, clarifies what the instruction means.

Given the requests for clarification that are potential responses to any directive, one can easily predict that any one or more of them might occur, but not with any certainty which. The design of the 'expert help system' attempts to deal with the problem exhaustively, and frequently succeeds. Questions of "How," "Where," and "Why" are answered by a diagram and supplementary description, provided with each next instruction. In all of these instances, the user brings the descriptions that the system provides to bear on deciphering the material circumstances of her situation, and brings those circumstances to bear on her interpretation of the descriptions. The user exploits the meaning of object and action descriptions to pick out their referents, in other words, and uses the objects and actions picked out as resources for finding the significance of the description.

## 7.4 Conditional relevance of response

I have described how the responsiveness of the system is limited to those occasions where the user's actions effect some change in the machine's state, that ties them to the requirements of the underlying design plan. In principle, the design plan serves as the measure of what constitutes an adequate and appropriate action by the user; namely, one that satisfies the current procedural requirement. The requirements that the system imposes, in this procrustean sense, serve as prescriptions for successful use of the machine. The success assumes, however, that the user interprets the instructions and the system's responses in the way that the designer intended.

In the interest of conveying the intent of the design to the user, and in doing so through something like interaction, the designer implicitly exploits certain communicative conventions. Most generally, designer and user share the expectation that the relevance of each utterance is conditional on the last; that given an action by one party that calls for a response, the other's next action *will be* a response. The expectation does not ensure that any next action in fact will or must be a response to the last, but it does mean that wherever possible, the user will look for an interpretation of the next action that makes it so.

The user's expectation, in other words, is that each system response conveys, either implicitly or explicitly, an assessment of the last action she has taken and a recommendation for what to do next. More specifically, given some instruction to which the user responds with an action, the user has the following expectations with respect to the system's response:

- i) The system's response should be a new instruction, which stands as implicit confirmation of the adequacy of the user's previous action,
- ii) If the system does not respond, the user's previous action is somehow incomplete.
- iii) If the system's response is to repeat the instruction, the repetition implies that the user's previous action should be repeated (i.e. that the procedure is recursive) OR that there is some trouble in the previous action that should be repaired.

### 7.4.1 A new instruction confirms the previous action

We have a general expectation, in carrying out a step-wise procedure, that completion of one action allows progress to a new instruction, and a next action. The inverse of the expectation that a completed action indicates readiness for a new instruction is the fact that the appearance of a new instruction is taken, at least initially, as confirmation of the previous action:

## XIV [22:83-95]

(S's are making two-sided copies of a bound document. They first must make a single, unbound master copy using the "Bound Document Aid," or BDA.)

	THE USER'S ACTIONS	THE MACHINE'S BEHAVIOR	
I	II	III	IV
	Not available to the machine	Available to the machine	Available to the user
DISPLAY 3			
S2	Okay. "Slide the document cover: left over your original, until it latches."		Instructions for closing document cover
S1	(moves hand to BDA)		
S2	(Turns to machine) "Slide the document cover." (looks back to diagram) that's this (BDA). Right?		
S1	(Starts to close) We— it said left, though. (looks to display)		
S2	"To close the Document Cover, grasp the cover,		
S1	slide it firmly to the left."	CLOSES BDA	
DISPLAY 2			
	(You must) have done that.		Instructions for raising the document handler

Evidence for the adequacy of the action in this case is found in the fact that it generates a response, which is assumed to be a next instruction. The apparent change to a new instruction confirms the action in spite of the fact that the action description, "Slide the document cover," does not actually seem to fit the action taken. The action taken in fact is not closing the *document cover*, which is located inside the Bound Document Aid, but instead closing the Bound Document Aid itself. The assumption that DISPLAY 2 must be a *next* to DISPLAY 3, however, masks the fact that they are entering into a loop between those two displays (see Sequence VI).

#### 7.4.2 No response indicates that the previous action is incomplete

In conversation, silences are more than just the absence of talk; they are generally owned by one party or another, and they invariably acquire significance (see chapter 5). The significance of silence lies in its relationship to the talk that it follows and, retrospectively, the talk that it can be seen to precede. In particular, the convention that certain utterance types (questions and answers being the canonical example) are sequentially implicative of the appropriate *next* produces "noticeable absences" when the *next* is not forthcoming. An extended silence following a question, for example, will be seen as a non-response. In the case of the 'expert help system,' there is no response until the user completes the action prescribed by the final instruction of a given display. This design constraint, combined with the user's expectation from human interaction regarding sequential implicature and silence, means that the unresponsiveness of the system carries information. Specifically, when an action that is intended to satisfy a final instruction fails to elicit a response, the user takes the unresponsiveness as evidence for trouble in her performance of the action:

XV [203:1473-1488]

(S's are making two-sided copies using the "Recirculating Document Handler" or RDH.)

	THE USER'S ACTIONS		THE MACHINE'S BEHAVIOR	
	I	II	III	IV
	Not available to the machine	Available to the machine	Available to the user	Rationale
			DISPLAY 10	Copies complete
S1	Okay, "Remove the copies from the output tray." (Takes documents from document handler) Okay. Now:  (15.0)  (turns to output) Oh, (looks back to display)			
S2	The output tray:			
S1	This is the output tray.			
S2	(points to picture) That's the output tray, okay.			

In this instance, what the S's initially see as a pause turns, in virtue of its length, into a non-response. The non-response, in turn, carries information with respect to their last action. Specifically, the non-response indicates that this is still, in effect, their 'turn,' that the last action was not, somehow, the action prescribed by this instruction. The evidence that the non-response provides—that there is some problem in the action taken—initiates a re-inspection of the instruction, a re-identification of the instruction's object, and the action's repair.

#### *7.4.3 Repetition is ambiguous between iteration and repair*

There are two conditions on which the system may repeat a prior instruction:

- a) The action taken in response to the instruction *should be repeated* — the common case, for example, in a procedure that is *iterative*;
- b) The action taken in response to the instruction *is in error* in just such a way as to return the system to a state prior to the instruction; in effect, to undo a previous action. This produces a *loop*.

In human interaction, (b) does not occur. On the other hand, in human interaction repetition is used in a way that does not occur between user and system, namely to indicate that:

- c) The action taken in response to the instruction in some way fails to satisfy the intent of the instruction, and needs to be remedied.

Consistent with the observation that users import expectations from human interaction to construe the system's responses, users failed to recognize the occurrence of (b), and instead read all cases of repetition as either (a), or as (c).

**Repetition as iteration.** In procedural instructions, there are occasions on which the repeat of an instruction is to be taken at face value, as an explicit directive to do the previous action again:

## XVI [22:325-351]

(S's are making two-sided copies of a bound document. They have copied the first page of the document, using the Bound Document Aid.)

THE USER'S ACTIONS		THE MACHINE'S BEHAVIOR	
I	II	III	IV
Not available to the machine	Available to the machine	Available to the user	Rationale
DISPLAY 6			
S2	"If more pages are to be copied, then place the next page face down on the glass."		Iterative procedure for using the BDA
S1	Just keep it up until we're finished with the, with the, uh:		
S2	Oh, well how do you— she was— she said on both sides, right?		
S1	Well that's after we finish getting this (indicating document). We're just getting the originals to stick up here (Recirculating Document Handler).		
S2	Oh, you're right, you're right.		

Purposeful action is characterized by the fact that its projected outcome is a resource for assessing the action's course, rather than being just the action's consequence. In the case of composite procedures, this function of the projected outcome of an action is complicated. For one thing, success at a composite procedure depends crucially on reliable ways of discriminating between the procedure's outcome, and its intermediate states. Particularly for the novice, the expectation that a component procedure (in this case, making the unbound master copy of the document) will produce the *finished* product leads to confusion like that of S2 in the example above, and to more complex misunderstandings such as the following:

## XVII [203:1321-1343]

(S's are making five two-sided copies of a bound document. They have completed the *master* copy using the "Bound Document Aid." Unaware of the composite structure of the procedure, and seeking to explain the fact that this procedure has produced only one copy, they have adopted the hypothesis that the remaining four copies are produced automatically, by the machine, and they are waiting for them to appear.)

	THE USER'S ACTIONS		THE MACHINE'S BEHAVIOR	
I	II	III	IV	
	Not available to the machine	Available to the machine	Available to the user	Rationale
DISPLAY 6				
S2	"Place the next page face down on the glass. Slide the document cover (inaudible). Lower the RDH until it latches. (pause) Okay. (11.0) So we start over for five? It doesn't do it (inaudible)?			Iterative procedure for using the BDA: when RDH is lowered, user is ready to go on to make multiple copies.
S1	I guess we just have to do it five times, and then it'll: (pause) Do what it says, I guess.			

Coming to what they take to be the end of the iterative procedure described in DISPLAY 6, the S's hypothesize that their part in the procedure is done, that the next 'turn' is the system's. That hypothesis is challenged by the system's inaction (a silence of 11.0 seconds), which suggests some incompleteness in their own action, and something further for them to do. Their problem, then, is to find a *next*; some action, prescribed by the instructions, that is outstanding.<sup>51</sup> Given that problem, and its situation in the inherently ambiguous context of a procedure that is recursive (copying each page once until the end of the document, in order that the document can be copied five times), one possible solution is to see the persistence of this instruction as a *repeat* rather than a non-response, and therefore as a directive to do the procedure again.

In another case the option "Change task description," intended by the designer to enable a repair, but noticed in the context of the search for a *next*, suggests iteration where the designer did not intend it:

51. The outstanding instruction, in fact, is "Lower the RDH until it latches." For the designer, that is the precondition for continuing on to the second pass of the procedure. For the users, however, under the assumption that the procedure is effectively done, that could only be a sort of coda (like putting away the bowls once one has baked a cake), with no direct consequence for the outcome.

## XVIII [21:218-229]

(S's are in a loop between DISPLAY 3 and DISPLAY 2.)

	THE USER'S ACTIONS I	THE MACHINE'S BEHAVIOR II	THE MACHINE'S BEHAVIOR III	IV Rationale
	Not available to the machine	Available to the machine	Available to the user	
S1	"Pull the latch labelled—" We did that. "Raise—" We did that. (studying display) Okay. Okay.		DISPLAY 2	Instructions for copying a bound document: Raising the document handler.
S2	"Lift up on the latch." We did that.			
S1	Now let's change::			
S2	"Change task description?"			
S1	Yes.			
S2	Selects "Change"		DISPLAY 0	User may want to change job specs
	"Describe the document to be copied—" Oh, we already did: No, we don't want to do that.			
S1	Maybe we have to do it to copy that. (next page)			
S2	(looks around machine) (laugh) I don't know.			

If the objective of the S's in selecting "Change task description" at this point is to find a *next*, one way that they can make the system's response a relevant one is to interpret the return to DISPLAY 0 iteratively, as telling them to specify their job again. The possibility, if not plausibility, of that interpretation arises from the fact that the difference between going "backward" to something already done in a procedure, and going "forward" to repeat the action, is inherently problematical. The difference does not lie in any features of the instruction or action itself, but just in whether the instruction's re-appearance at a given time is read as a misunderstanding, or as

intended by the design.<sup>52</sup>

Finally, the novice user may *expect* recursion in what is by design a one-pass procedure:

### XIX [20:28-30]

(S's are making four one-sided copies of an unbound document.)

THE USER'S ACTIONS		THE MACHINE'S BEHAVIOR	
I	II	III	IV
Not available to the machine	Available to the machine	Available to the user	Rationale
		DISPLAY 7	Instructions for copying an unbound document: Load all pages in RDH.
	S PUTS SINGLE PAGE IN DOCUMENT HANDLER		DISPLAY 8
S1	Press the Start button. Where's the Start button? [looks around machine, then to display]		Ready to Print
S2	[points to display] Start? Right there it is.		
S1	There, okay.	SELECTS START	
		STARTS	Document is being copied
		DELIVERS COPIES	

52. For more on how this problem develops, see Sequence VI.

THE USER'S ACTIONS		THE MACHINE'S BEHAVIOR	
I	II	III	IV
Not available to the machine	Available to the machine	Available to the user	Rationale
S1 So it made four of the first?			Job complete

DISPLAY 9

[looks at display]

Removing  
documents  
from the  
document  
handler.

Okay.

S1 [Takes first page  
out of document handler]REMOVES  
ORIGINAL

DISPLAY 10

Removing  
the copies.[holding second page  
over the document  
handler, looks to  
display]

Does it say to put it in yet?

[Puts second page  
into document handler]REPLACES  
ORIGINAL

DISPLAY 9

"Remove the original—"  
Okay, I've re- I've moved  
the original. And put in  
the second copy.

S1's action of removing the first page of the document and replacing it with a second assumes that this procedure is iterative, *viz.* copy each page one-at-a-time, until done. While taken as a *next*, her action restores a state that from the system's "point of view" appears identical to the state *before* the action was taken—a document in the document handler—thereby cancelling the action's effect. For the S's, logically, the *last* page has been removed from the document handler, and putting the *next* page in is pre-requisite to going on; for the system there is just *a* document in the document handler, and its *removal* is required to go on.

Seen as an instruction to undo their last action, the instruction to "remove the original" would stand as evidence of trouble. But by paraphrasing "remove" as "move the first page to make a place for the second," S1 makes this response relevant by turning it into a *next*, iterative instruction,

and therefore a confirmation of her last action.<sup>53</sup>

**Repetition as repair.** The inclination to see each next instruction as a new instruction (see 7.4.1) means that a repetition might not initially even be recognized as such. Recall that this was the case in Sequence XIV:

XIV [22:83-95]

THE USER'S ACTIONS		THE MACHINE'S BEHAVIOR	
I	II	III	IV
Not available to the machine	Available to the machine	Available to the user	Rationale

DISPLAY 3

- S2 Okay.  
"Slide the document  
cover: left over your  
original, until it latches."
- S1 (moves hand to BDA)
- S2 (Turns to machine)  
"Slide the document  
cover"  
(looks back to diagram)  
that's this (BDA).  
Right?
- S1 (Starts to close)  
We— it said left,  
though.  
(looks to display)
- S2 "To close the  
Document Cover,  
grasp the cover,
- S1 slide it firmly  
to the left."
- CLOSES BDA
- DISPLAY 2
- (You must)  
have done that.

53. For discussion of this sequence as a "garden path," see 7.5

	THE USER'S ACTIONS		THE MACHINE'S BEHAVIOR	
	I	II	III	IV
	Not available to the machine	Available to the machine	Available to the user	Rationale
S2	Well, I didn't slide anything firmly to the left, I just banged it down. (laugh)			
S1	Okay.			

In fact, this is another instance of the loop described for Sequence VI. Specifically, mislocation of the object referred to as the "document cover" leads the S's to close the entire Bound Document Aid, an action that returns the system to its initial state and causes it to re-display the first instruction, namely, to open the BDA.<sup>54</sup> The design rationale that produces this system response is simple; i) the user must use the BDA to copy bound documents, ii) in order to use the BDA, it must be opened, iii) if the BDA is closed, the user should be presented with instructions for opening it. However, rather than taking the return to the previous instruction as evidence for some problem in their last action, the S's see it as a *next* instruction, and as confirmation.

The inclination to mistake a return to a previous instruction for a *next* can be appreciated by considering the anomalous character of this particular problem in terms of any parallels in human interaction. While repetition of the first part of an adjacency pair is justified in cases where there is no response, when a response does occur it terminates the sequence and provides for the relevance of a next. Insofar as the user believes her action constitutes a response to the current instruction, then, she has every reason to view the system's next turn as a *next*. The closest situation that one finds in human interaction to the loop in human-machine communication occurs when a response to a sequentially implicative utterance—the answer to a summons, for example—is not recognized as such;

As noted, upon the completion of the SA [summons-answer] sequence, the original summoner cannot summon again. The operation of this terminating rule, however, depends upon the clear recognition that an A has occurred. This recognition normally is untroubled. However, trouble sometimes occurs by virtue of the fact that some lexical items, e.g., 'Hello', may be used both as summonses and as answers. Under some circumstances it may be impossible to tell whether such a term has been used as summons or as answer. Thus, for example, when acoustic difficulties arise in a telephone conversation, both parties may attempt to confirm their mutual availability to one another. Each one may then employ the term 'hello' as a summons to the other. For each of them, however, it may be unclear whether what he hears in the earpiece is an answer to his check, or the other's summons for him to answer. One may, under such circumstances, hear a conversation in which a sequence of some length is constituted by

54. Fortunately, the action that the BDA suggests, just because it returns the machine to a previous state, is the only action other than that which the design intends to which the system would respond at all at this point.

nothing but alternatively and simultaneously offered 'hellos'. Such 'verbal dodging' is typically resolved by the use, by one party, of an item on which a second is conditionally relevant, where that second is unambiguously a second part of a two-part sequence. Most typically this is a question, and the question 'Can you hear me?' or one of its common lexical variants, regularly occurs (Schegloff, 1972, p. 366).

Recognized as such, a return to a previous instruction that *cannot* be construed as recursive is evidence for trouble. Take another instance of the same misunderstanding:

**XX [21:191-217]**

THE USER'S ACTIONS		THE MACHINE'S BEHAVIOR	
I	II	III	IV
Not available to the machine	Available to the machine	Available to the user	Rationale

DISPLAY 3

S2 "Slide the document  
cover over your  
original until it latches."

S1 (hand on BDA)

S2 Just push it down.

S1 Closes BDA

DISPLAY 2

Okay, here we go.  
(turns to display)  
"Pull the latch la—"  
Oh, we already  
did that.

(pause. They study display)

S1 Okay.  
S2 Okay.

(7 seconds)

Now what do we do?

In human interaction, when the response to an action is incoherent or inappropriate, the producer of the action has recourse to two alternative interpretations. She can treat the troublesome response as the product of an error on the listener's part (not hearing or mishearing, not understanding or misunderstanding), or as intended. If the troublesome response is seen as the product of some failure of hearing or understanding, the repair may be just to repeat the original action (cf. Coulter 1974, p. 30). Unless the trouble is one of hearing, however, we rarely repeat a directive verbatim if there appears to be some problem of understanding the first time around.

Instead, we try some reformulation, or elaboration. If one formulation fails to convey our intended meaning, we try another. Frequently, it is not simply that we try an alternative formulation of what we intended before, but that what we intend is conditional on the others' response. In that sense, our own intentions are clarified for us by the response of the other.

In every case, to the extent that we are heard to be repeating ourselves, the repeat is heard as an attempt to correct some problem in understanding the first time around (cf. Jordan and Fuller, 1974). Seen in this light, as a repair-initiator, repetition initiates a review of the repeated instruction:

XXI [21:218-229]

(Continued from XX)

	THE USER'S ACTIONS		THE MACHINE'S BEHAVIOR	
I	II	III	IV	
	Not available to the machine	Available to the machine	Available to the user	Rationale
S1	"Pull the latch labelled—" We did that. "Raise—" We did that. (studying display) Okay. Okay.			DISPLAY 2
S2	"Lift up on the latch." We did that.			

In this case, a review of the instruction confirms that the actions it prescribes have been done. The two alternative responses to the repeat, in that case, are either to assert that the action is complete, or to do it again. In face-to-face interaction these alternatives appear to be ordered; that is, we first assert that we have heard a prior utterance and responded to it and then, if the assertion does not suffice, we provide a demonstration. The discovery by users that assertions never suffice in the case of communication with the system, that the system has access only to demonstrations or actions, is part of the acquisition of proficiency in its use.

Actually re-doing an action frequently uncovers problems of understanding, not just because the same terrain is considered again, but because, considered again, the terrain is seen differently:

XXII [22:206-271]

(Another case of the loop between DISPLAY 2 and DISPLAY 3.)

	THE USERS ACTIONS	THE MACHINE'S BEHAVIOR	
I	II	III	IV
DISPLAY 2			
S2	Not available to the machine	Available to the machine	Available to the user
DISPLAY 3			
S1	Yea.		Rationale
S2	"And lift up to the left", <i>do it again.</i>		
S1		Opens to BDA	
S2	"Pull the latch labelled Bound," "Raise the RDH." "To access the BDA, pull the latch labelled Bound Document Aid." Okay, (gesture to latch, then back to display) we did.		
S1	There. (to display) I'm lifting up to the left.		
S2	Okay. "Place your original face down on the glass, centered over the registration=		
S1	"guide", Okay.		
S2	=guide."		
S1	Did that.		
S2	"Slide the document co—"		
S1	"cover left over=		
S2	Wait a minute.		
S1	=your originals,"		
S2	Well:		
S2	Here's the document glass, (indicating BDA) is that what they mean?		
S1	(looking at BDA) Document cover.		

	THE USERS ACTIONS	THE MACHINE'S BEHAVIOR	
	I	II	III
			IV
			Rationale
	Not available to the machine	Available to the machine	Available to the user
S2	"To close the document cover, grasp the cover and slide it firmly to the left."		
S1	(finding it) Oh, here's the document cover!	Closes Doc Cover	
S2	Oh, Jean, good girl!		
S1	There's the document— (Both turn back to display)		DISPLAY 4
S1	Okay, now: [	Selects "Start"	
S2	All right: "Press: the Start button" Jean, you're doin' great.		
	(Both look to BDA)		Machine starts
S1	Oh, I see, [		
S2	Alright.		
S1	we don't have to close this big thing.		
S2	No, we were— we were lookin' at the wrong thing. We were closing the bound document aid, instead of the:		
S1	instead of the document cover.		

When a review fails to reveal any new actions, however, one reasonable inference is that the next action must be the other's:

XXIII [202:147-170]

(Again, the loop between DISPLAY 2 and DISPLAY 3.)

THE USER'S ACTIONS		THE MACHINE'S BEHAVIOR	
I	II	III	IV
Not available to the machine	Available to the machine	Available to the user	Rationale
DISPLAY 2			
S1 (inaudible, rereads instructions) Okay, are we ready? "Pull the latch labelled bound— to release." and then you release::: the, uh, RDH (inaudible.) Okay, are we ready?			
(pause)			
Oh, it's supposed to do it by itself.			
(pause)			

S1's review indicates that the actions prescribed by the instructions have all been completed; the sense of her "Ready" here is as in "ready to go." There appears to be nothing further for them to do. Since the logical next is for the machine to copy the document, S1 concludes that it must do so without further action on their part. Concluding that it is the system's "turn" offers an alternative to the original interpretation of the repeat, as an indication that their action is somehow incomplete. If the system is in fact responding to their last action, that both confirms the action's adequacy, and accounts for the system's failure to provide a next instruction.

The length of time that passes with no apparent activity, however, casts doubt on that conclusion, as the system's silence takes on the character of a non-response. If the system is not responding, there must be some further action for *them* to take. They again attempt to read the repetition as a directive to repair some problem in the action as it was done the first time through:

XXIV [202:364-382]

THE USER'S ACTIONS		THE MACHINE'S BEHAVIOR	
I	II	III	IV
Not available to the machine	Available to the machine	Available to the user	Rationale

## DISPLAY 2

- S1 "Pull the latch labelled bound copy aid to release the RDH"
- S2 (points)  
This is the RDH.  
This (latch) is the release.
- S1 *But why does it want it to release it?*  
(to display)  
"Release (inaudible)  
to enable placement of  
the bound document on  
the glass," so we  
don't have that on  
the glass like it's  
supposed to be.

S1's "Why" here is a situated one; that is, she is not asking in general about the rationale for this instruction, but in particular about its intent now, given their particular history and present circumstances. While the answer provided is intended to justify the instruction on *any* occasion, she attributes to it a significance particular to *this* occasion. Because their inquiry is situated in their particular circumstances, the answer is taken as an answer to that situated inquiry. Specifically, S1 reads the "to enable" clause as relevant to the directive that they release the RDH *again*, to allow a repair of some fault in the document's placement.<sup>55</sup> Under this interpretation of the design, the directive to re-place the document would be conveyed by re-presenting this instruction to the user until the document is placed correctly. This interpretation not only accounts for the loop in which they've found themselves, but also suggests the way out of it.

55. This attributes to the system substantially greater sensitivity than it has, *viz.* the ability to tell how the document is sitting on the glass, and to notice that it is faulted in some way.

## 7.5 Communicative trouble

This section describes two forms of communicative trouble between user and system: the *false alarm* and the *garden path*. In the first case, a misconception on the user's part produces evidence of an error in her actions where none exists; in the second, a misconception on the user's part produces an error in her action, the presence of which is masked. In both cases, the user's trouble is unavailable to the system.

### 7.5.1 The false alarm

I noted earlier (section 7.4.3) that purposeful action is characterized by the fact that projected outcomes of action are a resource for constructing the action's course. In particular, the effects of actions taken are compared against expected outcomes, in order to judge the action's adequacy. Expectations with respect to the effect of actions taken are discovered in the breach:

XXV [203:809-845]

(S's are making two-sided copies of a bound document. They have copied the first page.)

	THE USER'S ACTIONS	THE MACHINE'S BEHAVIOR		
	I	II	III	IV
	Not available to the machine	Available to the machine	Available to the user	Rationale
DISPLAY 5				
S1	"Instructions. Slide the document cover to the right."			Copying a bound document: Opening the document cover
S2	(noting output) Okay, it gave us one copy here.			
S1	Okay, "Slide the document cover right to remove the original."			
S2	We're supposed to have 5 copies and we only got one.			
S1	(looks to output) Oh. (looks to display) We only got one?			
S2	Yea.			

THE USER'S ACTIONS		THE MACHINE'S BEHAVIOR	
I	II	III	IV
Not available to the machine	Available to the machine	Available to the user	Rationale

(long pause)

S1      What do we do then?

(long pause,  
both study display)

The system here offers the users two, competing pieces of evidence regarding the adequacy of their last action. The display offers a next instruction, which makes sense as a confirmation of their previous action (see 7.4.1). The output, however, indicates that the action has failed, in which case the next instruction is irrelevant. From the system's "point of view," however, there is no problem. And just because the system detects no problem here, it offers no prescription for a remedy. The result is an interactional impasse, where the question "What do we do then?" finds no answer. Or rather, the answer that the system provides makes sense only if what they intend to do is to continue making a single copy from a bound document.<sup>56</sup>

While from the point of view of the design that is precisely what they want to do, that intent is not a feature of *their* situation. Their situation, meanwhile—that they intended to produce five copies of the document, and have produced only one—is unavailable to the system. The consequence is that they ascribe a (spurious) misunderstanding of their intent to the machine, while the machine fails to detect the (genuine) misunderstanding on their part with respect to the structure of the procedure. The result is their effort to repair a line of action that is in no way faulted.

---

56. The problem here is one of intermediate states vs. outcomes. The procedure for two-sided copying requires use of the recirculating document handler, but use of that mechanism requires an unbound original. As a consequence of that mechanism constraint, the first pass of the procedure for copying a bound document is directed at producing one, unbound master copy of the document. This requires a procedure of copying each page, using the bound document aid, until there is one complete set to be put into the RDH. While an overview of the procedure was presented in DISPLAY 1, designed to correct the assumption that this first pass would produce the desired outcome directly, the overview evidently did not do so. S2's statement of the problem demonstrates their continuing confusion, but the fact that the confusion is unavailable to the system precludes what would be a second chance for the misconception's repair.

### 7.5.2 Garden path

To the extent that discrepant assumptions between users and designers produce evidence of misunderstanding, there is at least some hope that the trouble might be located and resolved. In 7.4 we looked at two events taken by users as evidence of trouble; namely, the non-response, and the repeat. As in Sequence XXV, false expectations with respect to an action's effect may lead the user to find evidence for trouble in her performance where, in design terms, none exists. Because in such cases the problem lies in the user's expectations rather than her actions, and because the evidence for her expectations that the user provides is unavailable to the system, the problem itself is unavailable to the system.

While the user is uncertain of her action in such cases, the action she takes is in fact the action that the design prescribes. Deeper problems arise when the user takes an action other than that prescribed by the design, but one that satisfies the procedural requirement. As a result of the ambiguity of the action's effect, the incorrect action is actually "mistaken" by the system for some other, correct action, from which it is indistinguishable by the system's sensors. As in XXV, the problem in such cases is inaccessible to the system. But whereas in XXV the misconception leads the user to find evidence of trouble where, by design, none exists, in these other cases trouble is *masked* by the fact that the user sees the action as non-problematic, and by the fact that because the action appears non-problematic to the system as well, the system's response appears to the user to confirm the action.

Take the following example:

XXVI [20:28-30]

(S's are making four copies of an unbound document.)

THE USER'S ACTIONS		THE MACHINE'S BEHAVIOR	
I	II	III	IV
Not available to the machine	Available to the machine	Available to the user	Rationale
		DISPLAY 7	Loading the originals
S1	PUTS SINGLE PAGE IN DOCUMENT HANDLER	DISPLAY 8	
Press the Start button".			Ready to Print
Where's the Start button?			
[looks around]			

THE USER'S ACTIONS		THE MACHINE'S BEHAVIOR	
I	II	III	IV
	Not available to the machine  machine, then to display]	Available to the machine	Available to the user
S2	[points to display] Start? Right there it is.		Rationale
S1	There, okay.	SELECTS START	STARTS DELIVERS COPIES
S1	So it made four of the first?		Job complete
		DISPLAY 9	
	[looks at display] Okay.		Removing documents from the document handler.

From the system's "point of view," this sequence produces no evidence of trouble. DISPLAY 7 instructs the S's to place their documents in the automatic document handler, the system's sensors "see" them do so, DISPLAY 8 instructs them to press start, they do, and the machine produces four copies of their document.

To a human observer with any knowledge of this machine, however, S1's question "So it made four of the first?" indicates a misunderstanding. Specifically, her question conveys the information that this in fact is not a single page document, but the first page of several. And in contrast to previous machines that require the placement of pages on the glass one-at-a-time, copying an unbound document of multiple pages with this machine requires loading the pages all-at-once. The problem here is not simply a failure of anticipation on the designer's part. On the contrary, in anticipation of this very situation the instruction for loading documents explicitly states that *all* of the pages should be placed in the document handler. There is no evidence, however, that the instruction is consulted by these users.<sup>57</sup>

---

57. A basic premise of instructions is that they explicate some problem of action: if there is no problem, there is logically no need for instruction. We can infer from the users' failure to consult the instructions at this point that they have a preconception about what to do, based on past experience. Such preconceptions probably account in large part for the common complaint from designers that people "ignore" instructions; they ignore them because they believe that they already know how to proceed.

Given the fact of the users' misconception, the further problem arises when the faulted action goes by unnoticed at the point where it occurs. It does so because what is available to the system is only the action's effect, and that effect satisfies the requirements for the next instruction. As an assertion in the form of a question, S1's statement not only formulates her view of the system's last operation, but requests confirmation of that formulation. Interactionally, her statement provides an occasion for the discovery of the misunderstanding. She even looks to the display for a response. The information provided there is efficient enough, however—it simply says, "The copies have been made"—to support her assertion, rather than challenging it. As a consequence, the misunderstanding displayed in S1's question is unavailable to the system, while the efficiency of the system's response masks the trouble for the user.

S1's action of placing the document in the document handler appears, in other words, to be a perfectly adequate response to DISPLAY 7. The system treats the action as satisfying the directive to place all of their documents in the document handler (where "all" in this case comprises one), and therefore provides a next instruction, while they take the appearance of the next instruction as confirmation that their last action, placing *the first page* of their document in the document feeder, satisfied the design intent.<sup>58</sup> The start-up of the machine, with no complaint about their prior action, reflects the fact that the directive to "Start" has two different, but compatible interpretations. For the users, the significance of the directive is "make 4 copies of page 1," while for the system it is just "make 4 copies of the document in the document handler." There is nothing in either DISPLAY 9 or DISPLAY 10 to indicate the discrepancy. Each is efficient enough to be read under either interpretation.

So at the point where the machine starts to print, S1 is making four copies of page 1 of her document, while the machine is just making four copies of the document in the document handler. This seems, on the face of it, a minor discrepancy. If the machine copies the document, why should it matter that it fails to appreciate more finely the document's status as one in a set of three?

The problem lies in the consequences of this continuing misunderstanding for the next exchange:

---

58. The fact that in this instance one could easily imagine a test, viz. there must be >1 document, doesn't alter the basic point: i) as observers we learn about their intent through their talk, and ii) a test is essentially an alternative to (i). Different versions of this same problem can be expected to be more and less amenable to testing.

XXVII [20:32-35]

THE USER'S ACTIONS		THE MACHINE'S BEHAVIOR	
I	II	III	IV
Not available to the machine	Available to the machine	Available to the user	Rationale
DISPLAY 9			
S1      [Takes first page out of document handler]	REMOVES ORIGINAL	DISPLAY 10	Removing originals from the document handler.
[holding second page over the document handler, looks to display] Does it say to put it in yet? [Puts second page into document handler]	REPLACES ORIGINAL	DISPLAY 9	Removing the copies.
			Removing originals.

The strength of S1's conception of what is going on (repeating the procedure for each page) provides her with a logical next action (loading her second page into the document handler) in advance of any instruction. The instruction is looked to for confirmation of her action, rather than for direction. Her certainty is evident in the terms of her question; the indexicals *it* with respect to the system as "next speaker," and to the second page as the object of the instruction, the *in* with respect to the location of the action, and the *yet* with respect to the time of the action, all imply a shared situation that makes the business of anchoring each indexical term non-problematic. That the instruction will appear, and what it will say is not in question, only when.

While S is going on to the next run of the procedure, however, the system is still engaged in the completion of the last. What remains are the *removal* of originals and copies from their respective trays:

XXVIII [20:38-42]

	THE USER'S ACTIONS	THE MACHINE'S BEHAVIOR	
I	II	III	IV
	Not available to the machine	Available to the machine	Available to the user
S1	"Remove the original—" Okay, I've re- I've moved the original. And put in the second copy.	DISPLAY 9	Removing originals.

The "misunderstanding" between users and system at this point turns on just what the document in the document handler is, and how it got there. For S1, a first page has been replaced by a second, a necessary step for the next pass of what she takes to be a recursive procedure. For the system, there just is *a* document in the document handler, and its removal is required for the procedure's completion. The result is an impasse wherein both user and system are "waiting for each other," on the assumption that their own turn is complete, that their next action waits on an action by the other.

The instruction to "Place all of your originals in the RDH face up" must be designed for any user who might come along, on any occasion. The designer assumes that on some actual occasion, the instruction, in particular the relative quantifier *all*, will be anchored by the particular user to a particular document with a definite number of pages. Under the assumption that the user will do that anchoring, the system just takes the evidence that *something* has been put into the RDH as an appropriate response, and takes whatever is put there as satisfying the description. On the one hand, this means that the system can provide the relevant instruction in spite of the fact that it does not have access to the particular identities of this user, or this document. On the other hand, the system's insensitivity to particulars of this user's situation is the limiting factor on its ability to assess the significance of her actions.

## 7.6 Summary

This analysis has tied the particular problem of designing a machine that responds appropriately to the actions of a user, to the general problem of deciding the significance of purposeful action. The ascriptions of intent that make purposeful action intelligible, and define a relevant response, are the result of inferences based on linguistic, non-linguistic and circumstantial evidence. I have argued that one way to characterize machines is by the severe constraints on their access to the evidential resources on which human communication of intent routinely relies. In the particular case considered here, the designer of the 'expert help system' attempts to circumvent those constraints through prediction of the user's actions, and detection of the effects of actions taken. When the actual course of action that the user constructs proceeds in the way that the design anticipates, effects of the user's actions can be mapped to the projected plan, and the system can be engineered to provide an appropriate response.

The new user of a system, however, is engaged in ongoing, situated inquiries regarding an appropriate next action. While the instructions of the 'expert help system' are designed in anticipation of the user's inquiries, problems arise from the user's ability to move easily between a simple request for a next action, 'meta' inquiries about the appropriateness of the procedure itself, and embedded requests for clarification of the actions described within a procedure. In reading the machine's response to her situated inquiries and taking the actions prescribed, the user imports certain expectations from human communication; specifically, that a new instruction in response to an action effectively confirms the adequacy of that action, while a nonresponse is evidence that the action is incomplete. In the case of repeated instructions, an ambiguity arises between interpreting the repetition as a straightforward directive to repeat the action, or as a directive for its repair. A further problem arises when the action that the user takes in response to an instruction is in error in just such a way as to return the system to a state prior to that instruction. Because this trouble does not arise in human interaction, new users initially fail to recognize the occurrence of such a loop.

Due to the constraints on the machine's access to the situation of the user's inquiry, breaches in understanding that for face-to-face interaction would be trivial in terms of detection and repair, become 'fatal' for human-machine communication (cf. Jordan and Fuller 1974). In particular, misconceptions with regard to the structure of the procedure lead users to take intermediate states of the procedure as faulted outcomes. Because the intermediate state is non-problematic from the system's point of view, the system offers no remedy. The result is an interactional impasse, with the user finding evidence of trouble in her actions where none in fact exists. In the case of the garden path, in contrast, the user takes an action that is in some way faulted, which nonetheless satisfies the requirements of the design under a different but compatible interpretation. As a result, the faulted action goes by unnoticed at the point where it occurs. At the point where the trouble is discovered by the user, its source is difficult or impossible to reconstruct.



## 8. Conclusion

*The scientist's task is not to duplicate phenomena but to make them accessible to the intellect.* In contemporary Western science this can mean only one thing: The scientist must substitute for the 'real thing' a system built on principles which he can understand. The 'ultimate reality' is approachable in its manifest entirety by neither science nor revelation, neither by poetry nor mystic illumination. There is no limit to the questions which man can ask and no limit therefore to what in principle can be revealed. The scientist's task is a never-ending one of unfolding a description which relates both to the phenomena (i.e., the evidence of his senses) and to his capacity to intellectually grasp the description (i.e., to his rational capacities) (Pylyshyn 1974, p. 65, original emphasis).

Researchers interested in machine intelligence and in human-machine communication have embraced the traditional philosophical and scientific view that purposeful action is planned action. Yet even casual observation of purposeful action indicates that, as common sense formulations of intent, plans are inherently vague. To the Cognitive Scientist, this vagueness is a fault to be remedied, insofar as in any event a plan is the prerequisite for purposeful action, and the details of action are derivative from the completion and modification of the plan. The task of the scientist who would model situated action, therefore, is to improve upon, or render more 'precise' and axiomatic, the plan. For situated action, however, the vagueness of plans is not a fault but, to the contrary, is ideally suited to the fact that the detail of intent and action must be contingent on the circumstantial and interactional particulars of actual situations.

One aim of this study has been to suggest the mutual relevance of two fields of endeavor—research in machine intelligence, and human-machine communication, on the one hand, and social studies of situated action and interaction, on the other—that today are largely unaware of each other. Just as the project of building intelligent artifacts has been enlisted in the service of a theory of mind, the attempt to build interactive artifacts, taken seriously, presses our understanding of human interaction. At the least, the attempt to simulate interaction challenges those of us committed to social studies to strengthen our characterizations of what interaction is. In this study, I emphasize three observations about human interaction that might serve as the basis for a strong characterization. First, the mutual intelligibility that we achieve in our everyday interactions—sometimes with apparent effortlessness, sometimes with obvious travail—is always the product of *in situ*, collaborative work. Second, the face-to-face communication that supports that work is designed to maximize sensitivity to situation particulars, and includes resources for detecting and remedying troubles in understanding as part of its fundamental organization. And third, every occasion of human communication is embedded in, and makes use of, a taken for granted but mutually accessible world.

Everything about our current communicative practices assumes such an embedding world. Communication is not primarily a symbolic manipulation that happens to go on in real-wordly settings, but a real-world activity in which we make use of language to delineate the relevancies of our environment. Our environment, in this sense, is not the material or biological world *simpliciter*, but the world under interpretation; i.e. the social world. If we build an account of action, or

language-use, without simultaneously building an understanding of language's relationship to the world, there is a sense in which we quite literally do not know what we are talking about.

The designer of an interactive artifact must contend with the relation of action to the world in two, related respects. First there is the fact that the significance of action, including but not limited to linguistic utterances, relies upon its situation. The user of a machine deploys a full complement of observational, linguistic, and inferential skills in order to construe—and misconstrue—the machine's behavior according to the behavior's particular circumstances. At the same time, the machine is mapping a restricted set of effects, left by the user's actions, onto a prescribed template of possible meanings. As a consequence of the asymmetrical access of user and machine to the situation of action, the ordinary collaborative resources of human interaction are unavailable. The real project of the designer of an interactive artifact, accordingly, is to engineer *alternatives* to interaction's situated properties.

The most plausible alternative—and the second way that designers must contend with the action-world relation—is to substitute a description of the situation of action, for access to the actual situation. The user is left to contend with the disparity between the hypothetical sequence of events on which the design is based, and the action's actual course. The discrepancy of description and situated action is not, in itself, the problem for human-machine communication. The problem is, rather, that designers tend to *identify* representations of action, like plans, with situated action. As a consequence, the question of the correspondence of idealized action descriptions to situated action is solved *a priori*, rather than defined as a problem for research. And only if that question is addressed directly can we hope to build our theories about human action on solid ground, and bring them usefully to bear on the design of new technology.

## References

[Allen, 1983]

Allen, James Recognizing Intentions from Natural Language Utterances. Chpt. 2 in *Computational Models of Discourse*, M. Brady and R. Berwick (eds.), Cambridge: MIT Press, 1983.

[Allen, 1984]

Allen, James Towards a general theory of action and time. *Artificial Intelligence* 23:123-154, 1984.

[Amerine & Bilmes, 1979]

Amerine, Ronald and Jack Bilmes Following Instructions. Unpublished manuscript, 1979.

[Anscombe, 1957]

Anscombe, G. E. M. *Intentions*. Blackwell, 1957.

[Appelt, nd]

Appelt, D. Planning english referring expressions. Unpublished manuscript, Palo Alto, CA: Artificial Intelligence Center, Stanford Research Institute, nd.

[Atkinson & Drew, 1979]

Atkinson, J. M. and P. Drew *Order in Court: The organization of verbal interaction in judicial settings*. Atlantic Highlands, N.J.: Humanities Press, 1979.

[Austin, 1962]

Austin, J. L. *How to do Things with Words*. Oxford: Clarendon Press, 1962.

[Bates, 1976]

Bates, Elizabeth *Language and Context: The Acquisition of Pragmatics*. New York: Academic Press, 1976.

[Beckman & Frankel, 1983]

Beckman, H. and R. Frankel Who hides the agenda: the impact of physician behavior on the collection of data. *Presented to the Fourth Annual SREPCIM Task Force on Interviewing*, Washington, D.C., April, 1983. (address reprint requests to Howard Beckman, M.D., POD 5C, University Health Center, 4201 St. Antoine, Detroit, MI, 48201.)

[Birdwhistell, 1970]

Birdwhistell, Ray *Kinesics and Context: Essays on Body Motion Communication*. Philadelphia: University of Pennsylvania, 1970.

[Blumer, 1969]

Blumer, Herbert *Symbolic Interactionism*. Englewood Cliffs, N.J.: Prentice-Hall, 1969.

[Bobrow, et al, 1977]

Bobrow, D. G., Kaplan, R. M., Kay M., Norman, D. A., Thompson H. & Winograd, T. GUS: a frame-driven dialogue system. *Artificial Intelligence*, 8, 155-173, 1977.

[Boden, 1973]

Boden, M. The structure of intentions. *Journal of Theory of Social Behavior* 3:23-46, 1973.

[Brady & Berwick, 1983]

Brady, M. and R. Berwick (eds.) *Computational Models of Discourse*. Cambridge: MIT Press, 1983.

## [Brown, et al, 1976]

Brown, J.S., Rubenstein and Burton Reactive Learning Environment for Comptuer Assisted Electronics Instruction. BBN Report 3314, Bolt Beranek and Newman, Inc., Cambridge, Mass., 1976.

## [Bruce, 1981]

Bruce, Bertram Natural Communication Between Person and Computer. In Lehnert and Ringle (eds.), *Strategies for Natural Language Processing*, 1981.

## [Burke, 1982]

Burke, Julie An Analysis of Intelligibility in a Practical Activity. Unpublished Ph.D. dissertation, University of Illinois at Urbana-Champaign, 1982.

## [Burks, 1949]

Burks, Arthur Icon, Index and Symbol. *Journal of Philosophy and Phenomenology Research*, 673-689, 1949.

## [Burton &amp; Brown, 1979]

Burton, Richard and John Seely Brown An investigation of computer coaching for informal learning activities. *International Journal of Man-Machine Studies* 11: 5-24, 1979.

## [Carbonell, 1971]

Carbonell, J. R. Mixed-initiative man-computer dialogues. *Technical Report 1970*, Bolt Beranek, and Newman, Inc., Cambridge, MA, 1971.

## [Churchland, 1984]

Churchland, Paul *Matter and Consciousness*. Cambridge, MA: MIT Press, 1984.

## [Cohen, 1966]

Cohen, J. *Human Robots in Myth and Science*. London: Allen and Unwin, 1966.

## [Cohen, nd]

Cohen, Philip Pragmatics, Speaker-Reference, and the Modality of Communication. Unpublished manuscript, Laboratory for Artificial Intelligence, Fairchild Camera and Instrument Corp., Palo Alto, Ca, nd.

## [Cohen &amp; Perrault, 1979]

Cohen, Philip and C. R. Perrault Elements of a Plan-Based Theory of Speech Acts. *Cognitive Science* 3, 177-212, 1979.

## [Coombs &amp; Alty, 1984]

Coombs, M. and J. Alty Expert systems: an alternative paradigm. *International Journal of Man-Machine Studies*, 20: 21-43, 1984.

## [Coulter, 1979]

Coulter, Jeff *The Social Construction of Mind*. Totowa, N.J.: Rowman and Littlefield, 1979.

## [Coulter, 1983]

Coulter, Jeff *Rethinking Cognitive Theory*. New York: St. Martin's Press, 1983.

## [Dennett, 1978]

Dennett, Daniel *Brainstorms*. Cambridge, MA: MIT Press, 1978.

[Dreyfus, 1979]

Dreyfus, Hubert *What Computers Can't Do: The limits of artificial intelligence* (revised edition). New York: Harper & Row, 1979.

[Dreyfus, 1982.]

Dreyfus, Hubert (ed.) *Husserl Intentionality and Cognitive Science*. Cambridge: MIT Press, 1982.

[Dreyfus, Forthcoming]

Dreyfus, Hubert *Being-in-the-World: A Commentary on Heidegger's Being and Time*, Division I. Forthcoming.

[Dreyfus & Dreyfus, Forthcoming]

Dreyfus, Hubert and Stuart Dreyfus *Putting Computers in Their Place: The Power of Intuitive Expertise in Management and Education*. Forthcoming.

[Duncan, 1974]

Duncan, S. Jr. On the structure of speaker-auditor interaction during speaking turns. *Language in Society*, 3, 161-180, 1974.

[Durkheim, 1938]

Durkheim, Emile *The Rules of Sociological Method*. New York: Free Press, 1938.

[Erickson, 1982]

Erickson, Frederick Money Tree, Lasagna Bush, Salt and Peper: Social Construction of Topical Cohesion in a Conversation Among Italian-Americans. In D. Tannen (ed.) *Georgetown University Round Table on Language and Linguistics: Analyzing Discourse: Text and Talk*. Washington, D.C.: Georgetown University Press, 1982.

[Erickson & Shultz, 1982]

Erickson, F. and J. Shultz *The Counselor as Gatekeeper*. New York: Academic Press, 1982.

[Fikes & Nilsson, 1971]

Fikes, Richard and Nils Nilsson STRIPS: A new approach to the application of theorem proving to problem solving. *Artificial Intelligence* 2: 189-205, 1971.

[Fitter, 1979]

Fitter, M. Towards more "natural" interactive systems. *International Journal of Man-Machine Studies*, 11: 339-349, 1979.

[Fodor, 1983]

Fodor, Jerry *The Modularity of Mind*. Cambridge, MA: MIT Press, 1983.

[Frankel, 1984]

Frankel, Richard From Sentence to Sequence: Understanding the Medical Encounter Through Microinteractional Analysis. *Discourse Processes* 7, 135-170, 1984.

[Galaty, 1981]

Galaty, John Models and Metaphors: On the semiotic explanation of segmentary systems. In L. Holy and M Stuchlik (eds.) *The Structure of Folk Models*. New York: Academic Press, 1981.

[Garfinkel, 1967]

Garfinkel, Harold *Studies in Ethnomethodology*. Prentice-Hall, 1967.

[Garfinkel, 1972]

Garfinkel, Harold Remarks on Ethnomethodology. In *Directions in Sociolinguistics: The Ethnography of Communication*. J. Gumperz and D. Hymes (eds.), New York: Holt, Rinehart and Winston, 1972.

[Garfinkel & Sacks, 1970]

Garfinkel, Harold and Harvey Sacks On formal structures of practical actions. In J. McKinney and E. Tiryakian (eds.) *Theoretical Sociology*. New York: Appleton-Century-Crofts, pp. 337-366, 1970.

[Garfinkel, et al, 1981]

Garfinkel, Harold, Michael Lynch and Eric Livingston The Work of a Discovering Science Construed with Materials from the Optically Discovered Pulsar. *Philosophy of the Social Sciences*, 11:131-158, 1981.

[Geertz, 1973]

Geertz, Clifford *The Interpretation of Cultures*. New York: Basic Books, 1973.

[Gladwin, 1964]

Gladwin, Thomas Culture and logical process. In W. Goodenough (ed.) *Explorations in cultural anthropology: essays presented to George Peter Murdock*. New York: McGraw-Hill, 1964.

[Goffman, 1975]

Goffman, Erving Replies and responses. *Language in Society*, 5, 257-313, 1975.

[Goodwin, 1981]

Goodwin, Charles *Conversational Organization: Interaction between speakers and hearers*. New York: Academic Press, 1981.

[Goodwin, 1980]

Goodwin, Marjorie Processes of Mutual Monitoring Implicated in the Production of Description Sequences. *Sociological Inquiry*, Vol. 50: 303-317, 1980.

[Grice, 1975]

Grice, H. P. Logic and Conversation. In P. Cole and J. Morgan (eds.), *Syntax and Semantics*, Vol. 3: *Speech Acts*. New York: Academic Press, 1975.

[Grosz, 1981]

Grosz, Barbara Focusing and Description in Natural Language Dialogues. *SRI International Technical Note* 185, April, 1979. (published in A. Joshi, B. Webber and I. Sag (eds.) *Elements of Discourse Understanding*. Cambridge: Cambridge University Press, 1981.)

[Gumperz, 1982a]

Gumperz, John *Discourse Strategies*. Cambridge: Cambridge University Press, 1982a.

[Gumperz, 1982b]

Gumperz, John The linguistic bases of communicative competence. In D. Tannen (ed.) *Georgetown University Round Table on Language and Linguistics: Analyzing Discourse: Text and Talk*. Washington, D.C.: Georgetown University Press, 1982b.

[Gumperz & Tannen, 1979]

Gumperz, J. and D. Tannen Individual and Social Differences in Language Use. In *Individual Differences in Language Ability and Language Behavior*. New York: Academic Press, 1979.

## [Hayes, 1981]

Hayes, Philip A construction specific approach to focused interaction in flexible parsing. *Proceedings of 19th Annual Meeting of the Association for Computational Linguistics*, Stanford University, pp.149-152, 1981.

## [Hayes &amp; Reddy, 1983]

Hayes, Philip and D. Raj Reddy Steps toward graceful interaction in spoken and written man-machine communication. *International Journal of Man-Machine Studies* 19: 231-284, 1983.

## [Heap, 1980]

Heap, James Description in Ethnomethodology. *Human Studies* 3, 87-106. 1980.

## [Hendrix, 1977]

Hendrix, G. G. Human engineering for applied natural language processing. *Proceedings Fifth International Joint Conference on Artificial Intelligence*, MIT, pp. 183-191, 1977.

## [Jefferson, 1972]

Jefferson, Gail Side Sequences. In Sudnow, D. (ed.) *Studies in Social Interaction*. New York: Free Press, 1972.

## [Jordan &amp; Fuller, 1974]

Jordan, Brigitte and Nancy Fuller On the Non-Fatal Nature of Trouble: Sense-Making and Trouble-Managing in *Lingua Franca* Talk. *Semiotica* 13:1, 11-31, 1974.

## [Joshi, et al, 1981]

Joshi, A., B. Webber and I. Sag (eds.) *Elements of Discourse Understanding*. Cambridge: Cambridge University Press, 1981.

## [Levinson, 1983]

Levinson, Steven *Pragmatics*. Cambridge: Cambridge University Press, 1983.

## [Lynch &amp; Garfinkel, 1983]

Lynch, M., E. Livingston and H. Garfinkel Temporal Order in Laboratory Work. In K. Knorr and M. Mulkay (eds.) *Science Observed*. London: Sage, 1983.

## [MacKay, 1962]

MacKay, D. M. The Use of Behavioral Language to Refer to Mechanical Processes. *British Journal of Philosophical Science*, XIII, 89-103, 1962.

## [McCorduck, 1979]

McCorduck, Pamela *Machines Who Think*. San Francisco: W.H. Freeman, 1979.

## [McDermott, 1976]

McDermott, Ray *Kids Make Sense*: An ethnographic account of the interactional management of success and failure in one first-grade classroom. Stanford University dissertation, 1976.

## [Mead, 1934]

Mead, G. H. *Mind, Self, and Society*. Chicago: University of Chicago Press, 1934.

## [Merritt, 1976]

Merritt, Marilyn Resources for saying in service encounters. Unpublished Ph.D. dissertation, Department of Linguistics, University of Pennsylvania, 1976.

[Miller, et al, 1960]

Miller, G., E. Galanter and K. Pribram *Plans and the Structure of Behavior*. New York: Holt, Rinehart and Winston, 1960.

[Newell & Simon, 1973]

Newell, Allen and Herbert Simon *Human Problem Solving*. Englewood Cliffs, N.J.: Prentice-Hall, 1973.

[Nickerson, 1976]

Nickerson, R. On conversational interaction with computers. In S. Treu (Ed.) *Proceedings of ACM/SIGGRAPH workshop*. Pittsburgh, Pa., October 14-15, 1976.

[Nilsson, 1973]

Nilsson, Nils A Hierarchical Robot Planning and Execution System. SRI Artificial Intelligence Center, *Technical Note 76*, Stanford Research Institute, Menlo Park, CA, April 1973.

[Oberquelle, et al, 1983]

Oberquelle, H., I. Kupka and S. Maass A view of human-machine communication and cooperation. *International Journal of Man-Machine Studies* 19, 309-333, 1983.

[Ochs, 1979]

Ochs, Elinor Planned and Unplanned Discourse In T. Givon (ed.) *Syntax and Semantics: Vol. 12, Discourse and Syntax*. New York: Academic Press, 1979.

[Pylyshyn, 1974]

Pylyshyn, Zenon Minds, Machines and phenomenology: Some reflections on Dreyfus' 'What Computers Can't Do.' *Cognition* 3(1): 57-77, 1974.

[Pylyshyn, 1984]

Pylyshyn, Zenon *Computation and Cognition*. Cambridge, MA: MIT Press, 1984.

[Rubin, 1980]

Rubin, Andee A theoretical taxonomy of the differences between oral and written language. In R. Spiro, B. Bruce and W. Brewer (eds.) *Theoretical Issues in Reading Comprehension*. Hillsdale, N.J.: Erlbaum, 1980.

[Sacerdoti, 1975]

Sacerdoti, Earl The nonlinear nature of plans. *Proceedings 4th International Joint Conference on Artificial Intelligence*, Tbilisi, USSR, 1975.

[Sacerdoti, 1977]

Sacerdoti, Earl *A Structure for Plans and Behavior*. New York: Elsevier, 1977.

[Sacks, 1963]

Sacks, Harvey Sociological Description. *Berkeley Journal of Sociology*, Vol. 8, 1963.

[Sacks, 1974]

Sacks, Harvey An analysis of the course of a joke's telling in conversation. In R. Bauman and J. Scherzer (eds.) *Explorations in the Ethnography of Speaking*. Cambridge: Cambridge University Press, 1974.

[Sacks, et al, 1978]

Sacks, H., E. Schegloff and G. Jefferson A simplest systematics for the organization of turn-taking in conversation. In J. Schenkein (ed.) *Studies in the Organization of Conversational Interaction*. New York: Academic Press, 1978.

## [Schank &amp; Abelson, 1977]

Schank, R. and R. Abelson Scripts, plans and knowledge. In P. Johnson-Laird and P. Wason (eds.), *Thinking: readings in cognitive science*. Cambridge University Press, 1977.

## [Scheflen, 1974]

Scheflen, A. E. *How Behavior Means*. Garden City, N.Y.: Anchor Press, 1974.

## [Schegloff, 1972]

Schegloff, Emanuel Sequencing in conversational openings. In J. Gumperz and D. Hymes (eds.) *Directions in Sociolinguistics: The Ethnography of Communciation*. New York: Academic Press, 1972.

## [Schegloff, 1982]

Schegloff, Emanuel Discourse as an interactional achievement: some uses of guh huh' and other things that come between sentences. In D. Tannen (ed.) *Georgetown University Round Table on Language and Linguistics: Analyzing Discourse: Text and Talk*. Washington, D.C.: Georgetown University Press, 1982.

## [Schmidt, 1975]

Schmidt, C. F. Understanding Human Action. *Conference on theoretical issues in natural language processing*, 1975.

## [Schmidt, et al, 1978]

Schmidt, C. F., N. Sridharan and J. Goodson The plan recognition problem. *Artificial Intelligence*, 11:45-83, 1978.

## [Schutz, 1962]

Schutz, Alfred *Collected Papers I: The Problem of Social Reality*. The Hague: Martinus Nijhoff, 1962.

## [Searle, 1969]

Searle, John *Speech Acts: An Essay in the Philosophy of Language*. Cambridge University Press, 1969.

## [Searle, 1979]

Searle, John *Expression and Meaning*. Cambridge University Press, 1979.

## [Searle, 1980]

Searle, John The Intentionality of Intention and Action. *Cognitive Science* 4, 47-70, 1980.

## [Sidner, 1979]

Sidner, C. L. Towards a computational theory of definite anaphora comprehension in English discourse. *Technical Report TR-537*, MIT AI Laboratory, Cambridge, MA, 1979.

## [Stich, 1983]

Stich, Stephen *From Folk Psychology to Cognitive Science*. Cambridge, MA: MIT Press, 1983.

## [Streeck, 1980]

Streeck, Jurgen Speech Acts in Interaction: A critique of Searle. *Discourse Processes* 3, 133-154, 1980.

## [Suchman, 1982]

Suchman, Lucy Toward a sociology of human-machine interaction: Pragmatics of instruction-following. CIS Working Paper, Palo Alto, CA: Xerox Palo Alto Research Center, November 1982.

[Turing, 1950]

Turing, A.M. Computing Machinery and Intelligence. *Mind*, Vol. LIX, No. 236: 433-461. October, 1950.

[Turkle, 1984]

Turkle, Sherry *The Second Self*. New York: Simon and Schuster, 1984.

[Turner, 1962]

Turner, Roy Words, Utterances and Activities. In *Ethnomethodology: Selected Readings*. Harmondsworth: Penguin, 1962.

[Watt, 1968]

Watt, W. C. Habitability. *American Documentation*, 19 (3) 338-351, 1968.

[Weizenbaum, 1966]

Weizenbaum, J. ELIZA: A computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 25th Anniversary Issue, Volume 26, No. 1:23-27, January 1983. (reprinted from *Communications of the ACM*, Vol. 29, No. 1:36-45, January, 1966.)

[Wilson, 1970]

Wilson, Thomas Conceptions of interaction and forms of sociological explanation. *American Sociological Review*, 1970.

[Zimmerman, 1970]

Zimmerman, Don The Practicalities of Rule Use. In J. Douglas (ed.) *Understanding Everyday Life*. Chicago: Aldine, 1970.

# Describe the document to be copied:

Is it a bound document? Yes  No

Copy both sides of each sheet? Yes  No

Is it on standard size (8.5" x 11") paper? Yes  No

Is it on standard thickness paper? Yes  No

Quality of original: darker than normal  normal  lighter than normal

1		
1	2	3
4	5	6
7	8	9
8	Clear	

About how many images are to be copied?

Number of copies:

1		
1	2	3
4	5	6
7	8	9
8	Clear	

## Comments

The currently described job will take about 1 minute.

Use standard paper? Yes  No

Staple each copy? Yes  No

Put images on both sides? Yes  No

Reduce size of images?  35% smaller  26% smaller  2% smaller

**PROCEED**

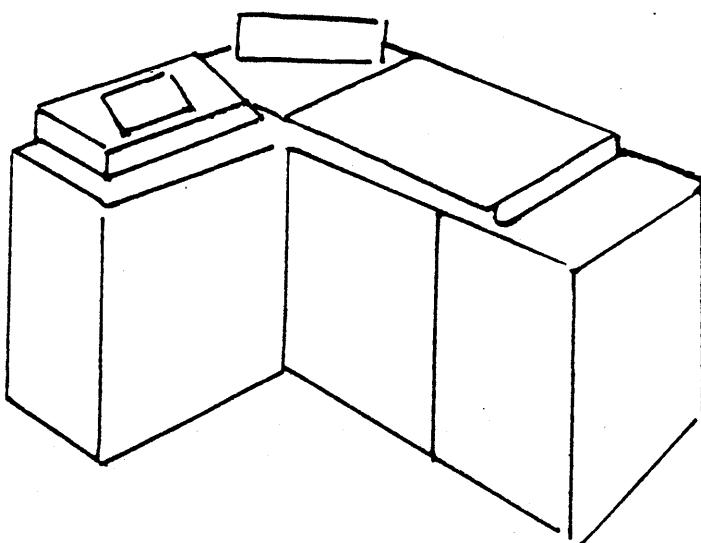
**HELP**

DISPLAY 0



**OVERVIEW:**

You need to use the Bound Document Aid (BDA) to make an unbound copy of your original. That copy can then be put into the Recirculating Document Handler (RDH) to make your collated two-sided copies.

**THE MACHINE****INSTRUCTIONS:**

Please wait.

**Change  
Task Description**

**Help**

**DISPLAY 1**

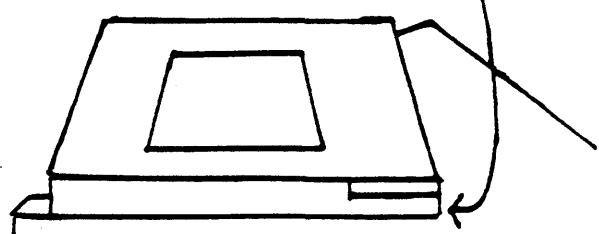


**OVERVIEW:**

You need to use the Bound Document Aid (BDA) to make an unbound copy of your original. That copy can then be put into the Recirculating Document Handler (RDH) to make your collated two-sided copies.

**How to access  
the BDA:**

To access the BDA, pull the latch labelled bound document aid,

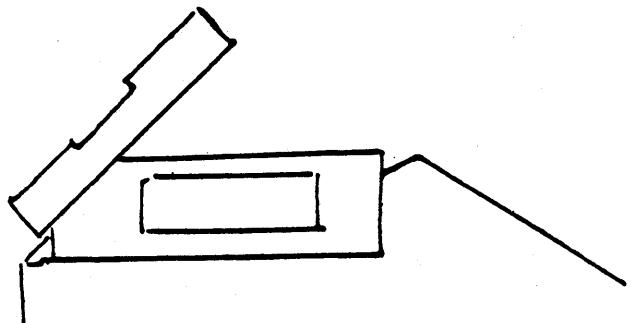


And lift up and to the left.

**INSTRUCTION:**

Pull the latch labelled bound document aid.  
(To release the RDH.)

Raise the RDH.  
(To enable placement of the bound document on the glass.)



Change Task  
Description

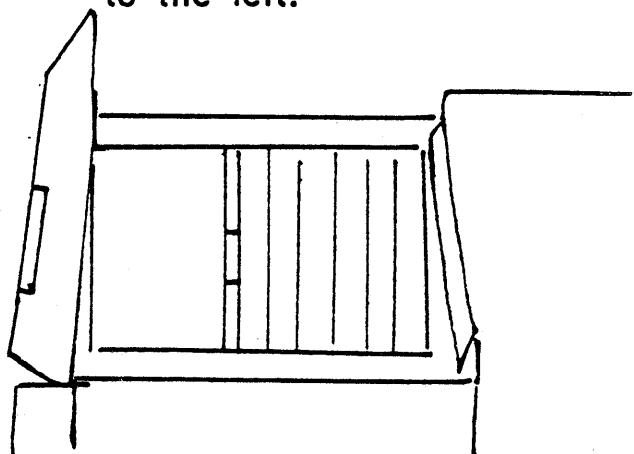
Help

DISPLAY 2



**OVERVIEW:****How to close  
the document cover:**

To close the document cover, grasp the cover, and slide it firmly to the left.

**INSTRUCTION:**

Place your original face down on the glass, centered over the registration guide.  
(To position it for the copier lens)

Slide the document cover left over your original until it latches.  
(To provide an eye shield from the copier lights)

**DISPLAY 3**

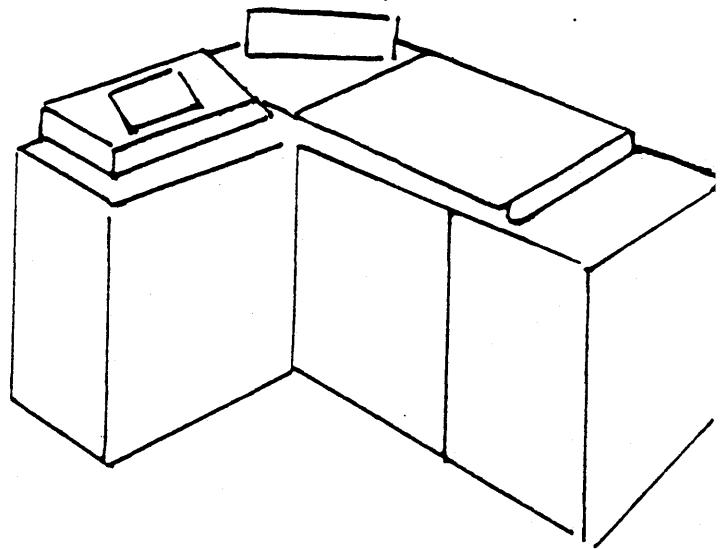


OVERVIEW:THE COPIERASSUMPTIONS:

The first page to be copied is on the glass.

INSTRUCTION:

Press the Start button  
(to produce a copy  
in the output tray)



**START**

**DISPLAY 4**

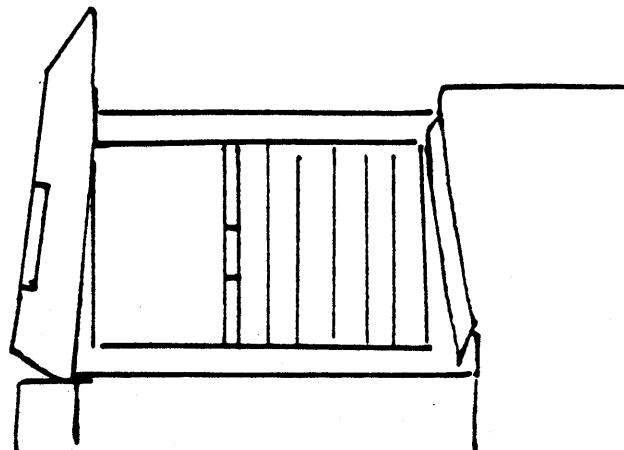


**ASSUMPTIONS:**

The copy of your original on the glass has been made.

**How to open the document cover:**

To open the document cover, grasp the cover, and slide it all the way to the right.

**INSTRUCTION:**

Slide the document cover right.  
(To remove the original)

Change task  
Description

Help

**DISPLAY 5**



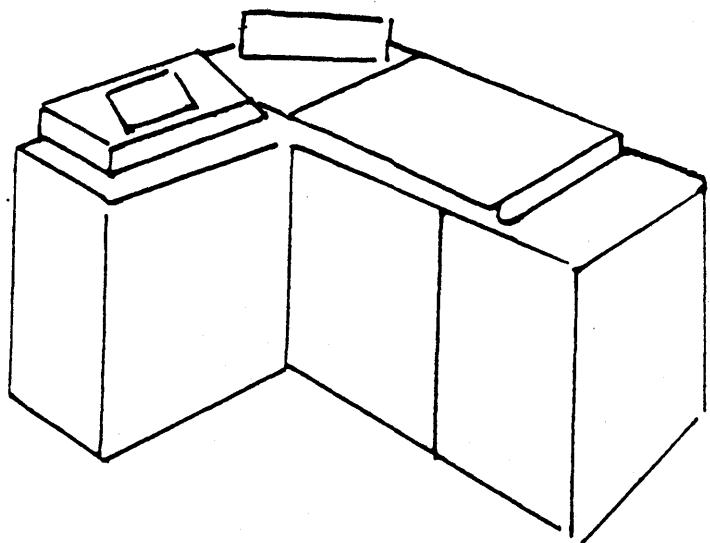
**THE MACHINE****INSTRUCTIONS:**

Remove the original from  
the glass.

If more pages are to be  
copied, then:

Place the next page  
face down on the glass.  
Slide the document  
cover left until it  
latches.

Otherwise, lower the RDH  
until it latches.



**Change  
Task Description**

**Help**

**DISPLAY 6**

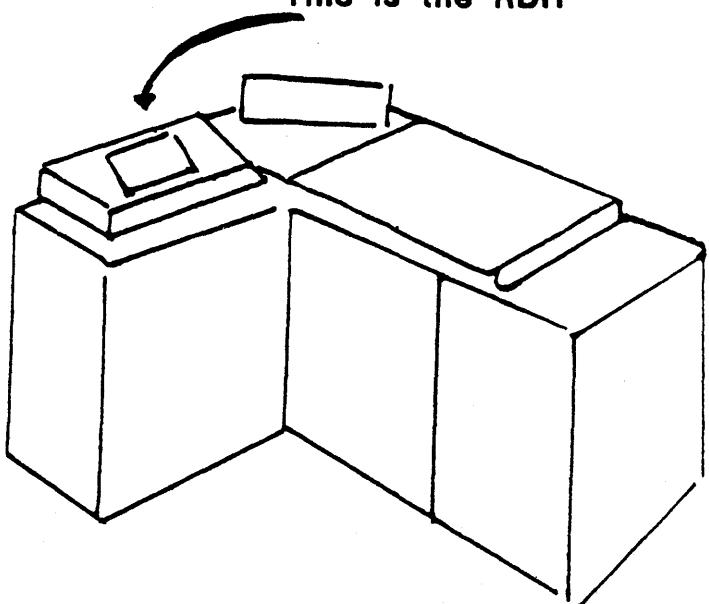


**OVERVIEW:**  
You can use the  
Recirculating Document  
Handler (RDH) to make  
your copies.

**INSTRUCTIONS:**

Place all of your originals  
in the RDH,  
first page on top.  
(so that the RDH can  
automatically feed each  
sheet into the copier.)

**THE MACHINE**  
This is the RDH



**Change  
Task Description**

**Help**

**DISPLAY 7**



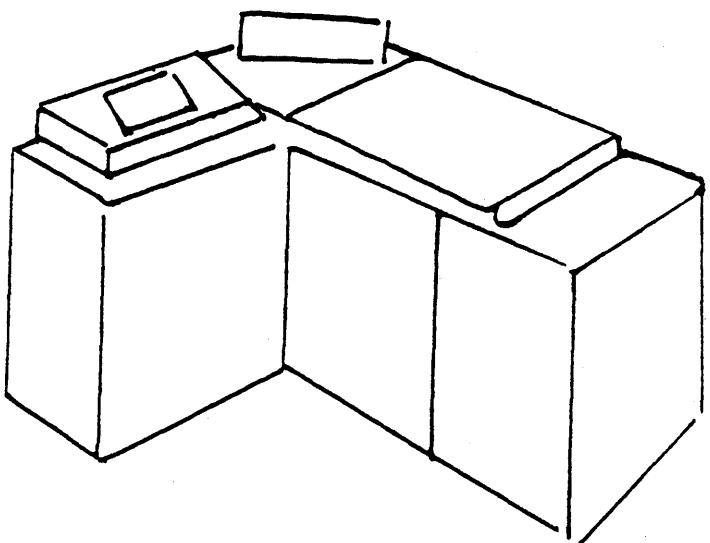
**OVERVIEW:**  
You can use the  
Recirculating Document  
Handler (RDH) to make  
your copies.

**ASSUMPTIONS:**  
The document to be  
copied is in the RDH.

**INSTRUCTIONS:**

Press the Start button.  
(to produce 4 copies  
in the output tray.)

**THE MACHINE**



Change Task  
Description

Start

Help

**DISPLAY 8**

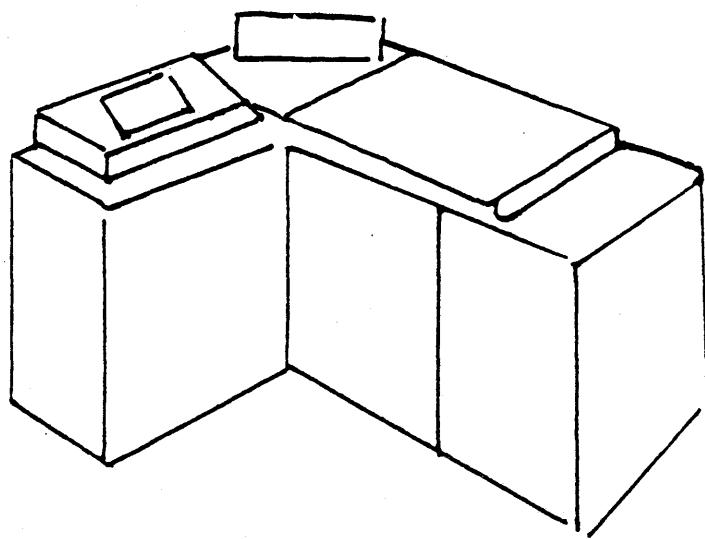


**THE MACHINE**

**ASSUMPTIONS:**  
The copies have  
been made.

**INSTRUCTIONS:**

Remove the originals  
from the RDH.



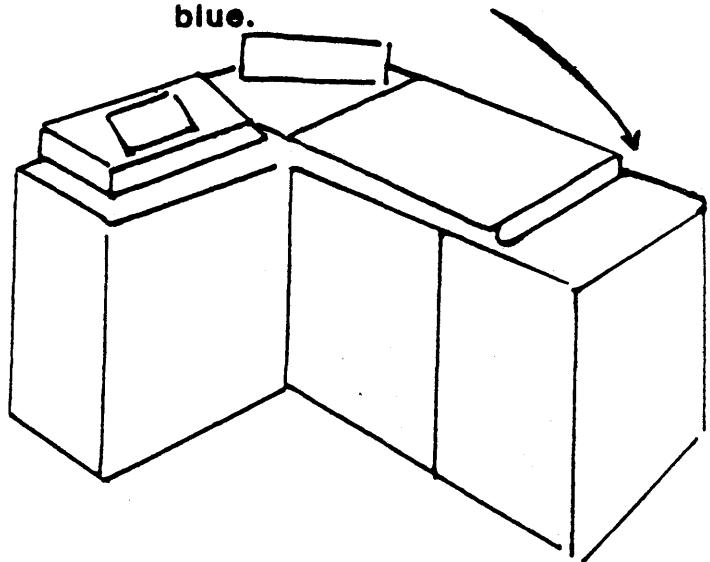
**Change Task  
Description**

**Help**

**DISPLAY 9**



**THE MACHINE**  
The output tray is  
where the copies  
come out. It is  
located on the right  
side and is colored  
blue.



**ASSUMPTIONS:**  
The copies have  
been made.

**INSTRUCTIONS:**

Remove the copies  
from the output tray.

Change Task  
Description

Help

**DISPLAY 10**

