

PLANS AND SITUATED ACTIONS:

on problems of human-machine
communication

by A. Hutchins

MIT PRESS

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 accomodate the unforeseeable contingences 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

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 accomodate 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

could be mutual intelligibility between people and machines, but the prior question of how we

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.⁴

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 these mental states on the one hand and of mental states to behavior on the other remains

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

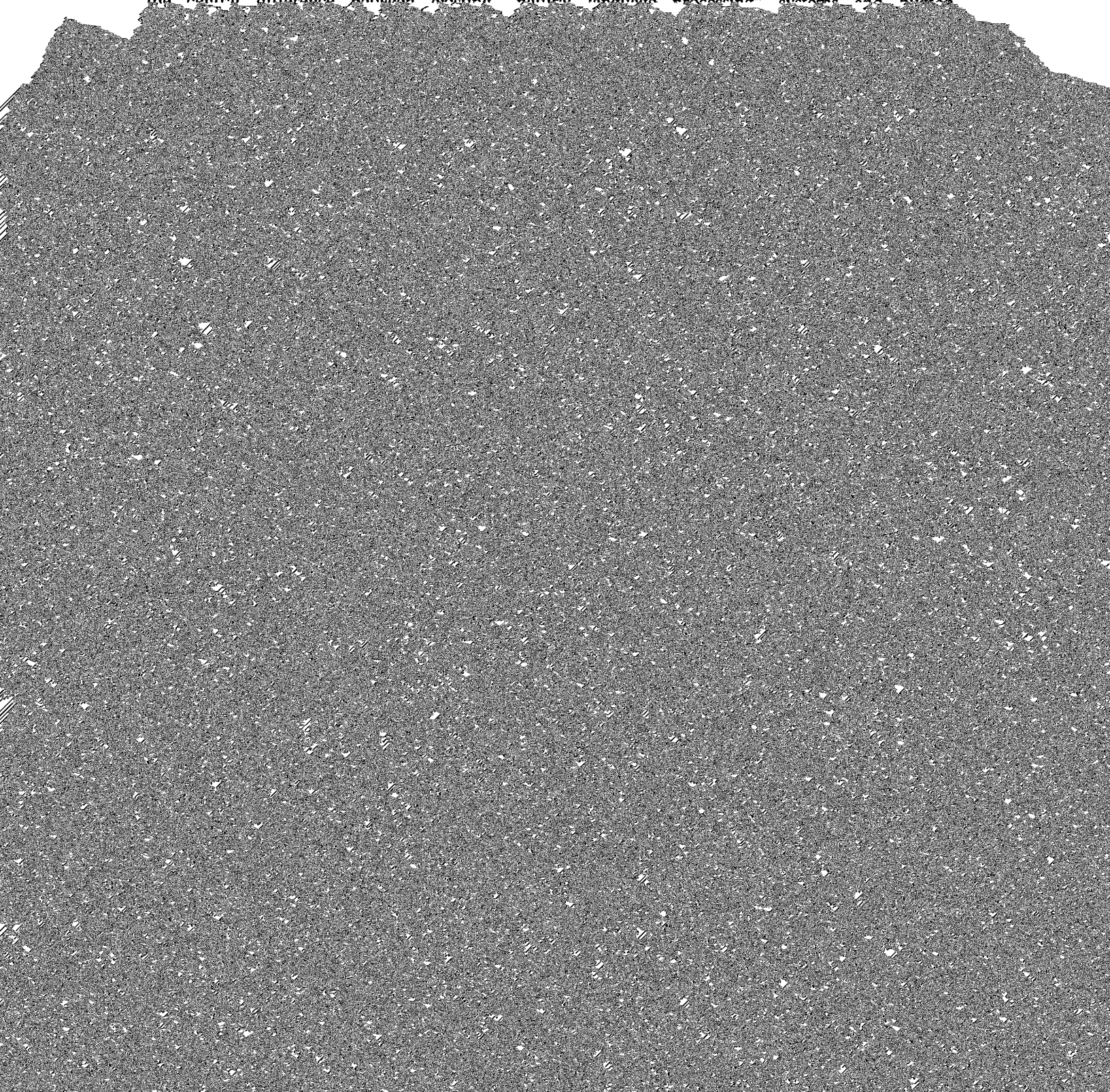
In spite of the persisting limits on machine intelligence, the use of an intentional vocabulary is already well-established in both technical and popular discussion of computers. In part, the attribution of purpose to computer-based artifacts derives from the simple fact that each action by the user 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.⁵ 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



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 is however the responsibility of the system designer to provide a

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 (Turtle 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.⁷ 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 is meant by a machine's intelligibility, and by human-machine communication.⁸ 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.⁹

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.¹⁰ 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.¹¹

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

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.¹⁵ 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

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).

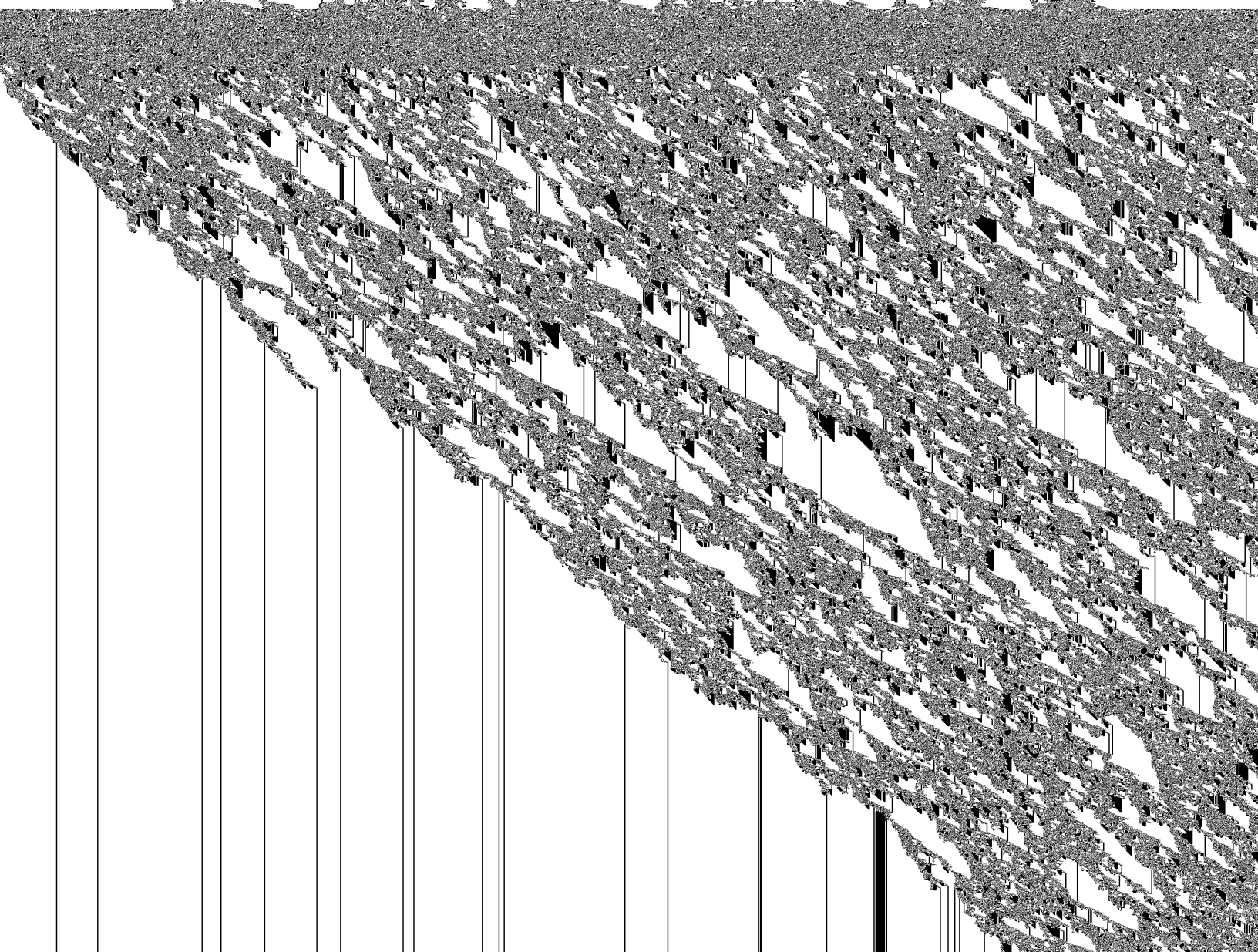
To the extent that communication consists in the mutual intelligibility of our actions, any account of communication presupposes, whether explicitly or implicitly, an account of the coherence and intelligibility of action as such. This chapter and the next discuss two alternative views of action. The first, adopted by most researchers in artificial intelligence, locates the organization and significance of human action in underlying plans.¹⁷ 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

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:¹⁸

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*



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

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).

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

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).

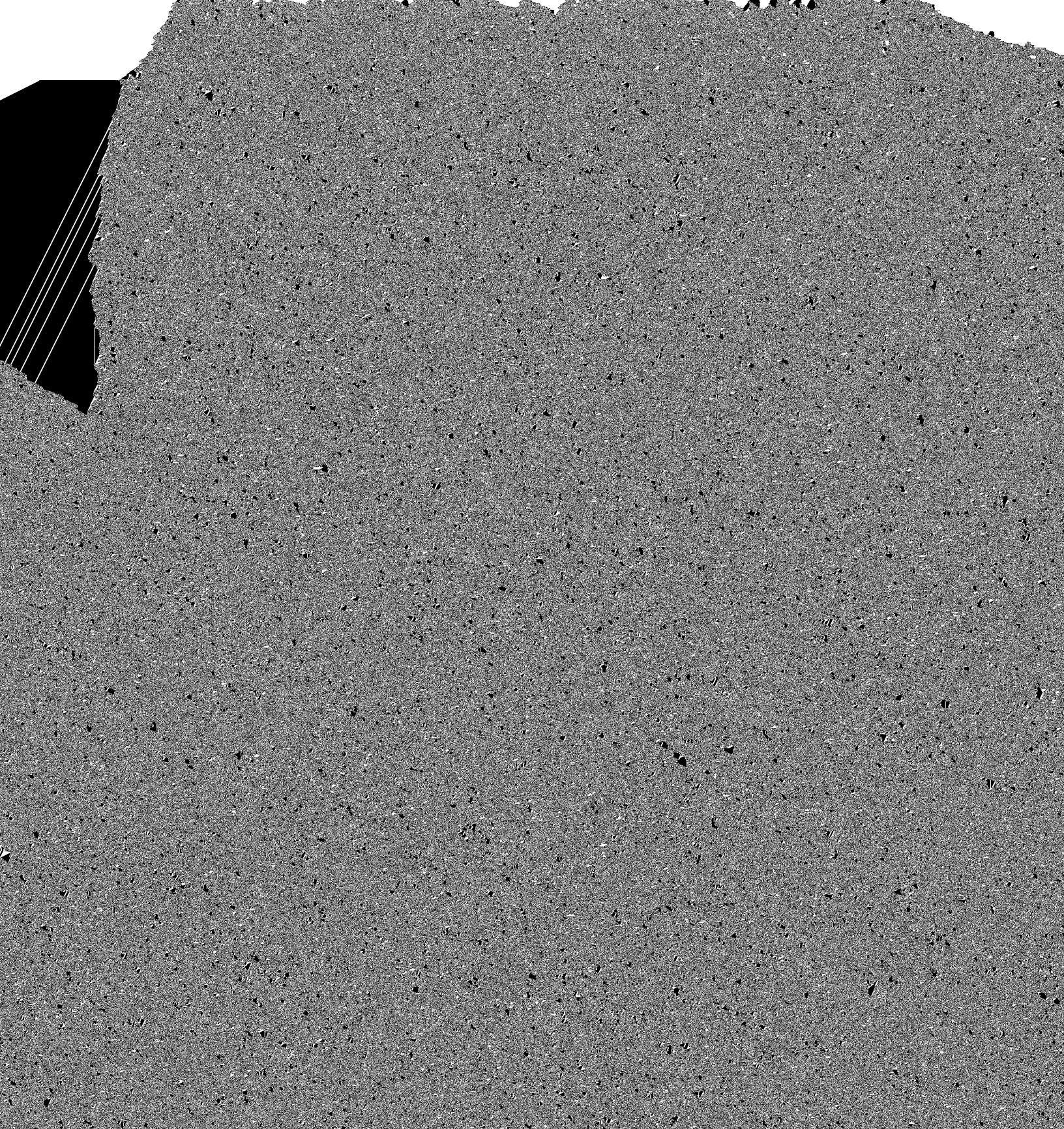
With Miller *et al*, the view that purposeful action is planned is reconstructed as a psychological 'theory,' compatible with the interest in a mechanistic, computationally tractable account of intelligent action. By improving upon or "completing" our common sense descriptions of the structure of action, 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

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

[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



possible context or interpretive frame that would make sense of the question, and comes up with the break. But Gempert points out *this analysis bears the question of how R arrives at the right*

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

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 as account of what is said always requires reference to further facts that cannot

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

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

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.

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

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 *ceteris 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:"

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 assumptive and retrospective.

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 consensual agreement, but of the common application of

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;

... natural sciences, courts of law, the practice of medicine (see Chapter 5). But ...

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:

4.6 Plans as efficient formulations

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.³⁰ 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.³¹ 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 (Schegloff 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).³² 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

- 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

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:³⁴

- 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).³⁵ 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

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 shouldering, 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.³⁶ 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 vermilion
A: Yuh
B: And I wanted to order some more the name's Boyd
A: Yes // how many tubes would you like sir
B: An-
B: U:hm (.) what's the price now eh with V.A.T.

((Request 1))

((Question 1))

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 accomodates 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: .hhhh 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: .hhhh 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

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.³⁷ 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

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

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 ca:r (1.0) and backed
it (1.0) into some trees didn't e
(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 accomodates 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.³⁹

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 11?

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"

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 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.⁴⁰ 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

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

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

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.⁴¹

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:⁴²

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

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:⁴³

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.

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.⁴⁴

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.⁴⁵

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

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

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':⁴⁶

(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	II	III	IV
Not available to the machine	Available to the machine	Available to the user	Rationale
		DISPLAY 1	Setting control panel
		DISPLAY 2	
S1	"To access the BDA, pull the latch labelled Bound Document Aid."		Instructions for copying a bound

	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	[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	II	III	IV
Not available to the machine	Available to the machine	Available to the user	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 to start printing
S1	[All right:: "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

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	There it goes.	DISPLAY 4	Ready to print
S1	"Press the Start button"		
	SELECTS START	STARTS	
	Okay.		

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

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::	DISPLAY 1	Setting panel
	It's got to come up with the little Start thing soon.		
	[pause]		
	Okay, we've done all that. We've made our bound copies.	DISPLAY 2	Instructions for copying 2 bound

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	Rationale
		Available to the user	
S1	(8.0) Then again, maybe we need to change the task description.	DISPLAY 2	Instructions for copying a bound document
S2	What do you think?	Selects "Change"	
		DISPLAY 0	User may want to change job description.
S1	No.		
S2	Okay, "Proceed."	Selects "Proceed"	
		DISPLAY 1	Making two-sided copies from a bound document.
S1	Maybe I should just lift it up and put it =	DISPLAY 2	Raising the document handler
S2	[		
S1	How do we skip this then?		
S1	=down again.		
	Maybe it'll think we're done.		
S2	(laughs) Oh, Jean.		
S1	There.	Opens BDA	
	Okay, we've done what we're supposed to do.	DISPLAY 3	Instructions for placing document
		Closes BDA	
	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

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 Not available to the machine	II Available to the machine	III Available to the user	IV Rationale
S1	"Instructions. Slide the document cover to the right."		DISPLAY 5	Instructions for copying a bound document: removing the document from the glass.
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	<i>What do we do then?</i> (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.⁴⁸ 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

THE USER'S ACTIONS		THE MACHINE'S BEHAVIOR	
	I Not available to the machine	II Available to the machine	III Available to the user
S2	(looks around machine (laugh) I don't know.		
S1	Well::		
S2	"Help" (laugh) "Select the question you would like help with."	Selects "Help"	
S1	I guess we still do have to =		User needs clarification of DISPLAY 0
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."⁴⁹ 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	IV
Not available to the machine	Available to the machine	Available to the user	Rationale
		DISPLAY 1	Overview
S1	"You need to use the Bound Document Aid to make an unbound copy of your original." Where is— Oh, here it is.	DISPLAY 2	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 Not available to the machine		III Available to the user	IV Rationale
		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	<i>"To provide an eyeshield for the copier (inaudible)."</i>		

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
Not available to the machine	Available to the machine	Available to the user	Rationale
		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)		

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"::		Instructions for copying a bound document: Raising the document handler.
	[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			
S2	lift up and to the left.		
	Yea.		
	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		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. (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:⁵⁰

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)		
		DISPLAY 3	
S2	(startled) Oh, alright.		
	Opens BDA		
	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

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	Rationale
		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	
	(You must) have done that.	DISPLAY 2	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

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*

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—		

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

XVIII [21:218-229]

(S's are in a loop between DISPLAY 3 and 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: 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-52

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	

	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] Okay.			Removing documents from the document handler.
S1	[Takes first page out of document handler]			
		REMOVES ORIGINAL		
			DISPLAY 10	
	[holding second page over the document handler, looks to display] Does it say to put it in yet? [Puts second page into document handler]			Removing the copies.
		REPLACES ORIGINAL		
	"Remove the original—" Okay, I've re- I've moved the original. And put in the second copy.		DISPLAY 9	

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.⁵³

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"		

	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.⁵⁴ 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

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		
	Okay, here we go. (turns to display) "Pull the latch la—" Oh, we already did that. (pause. They study display)	DISPLAY 2	
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)

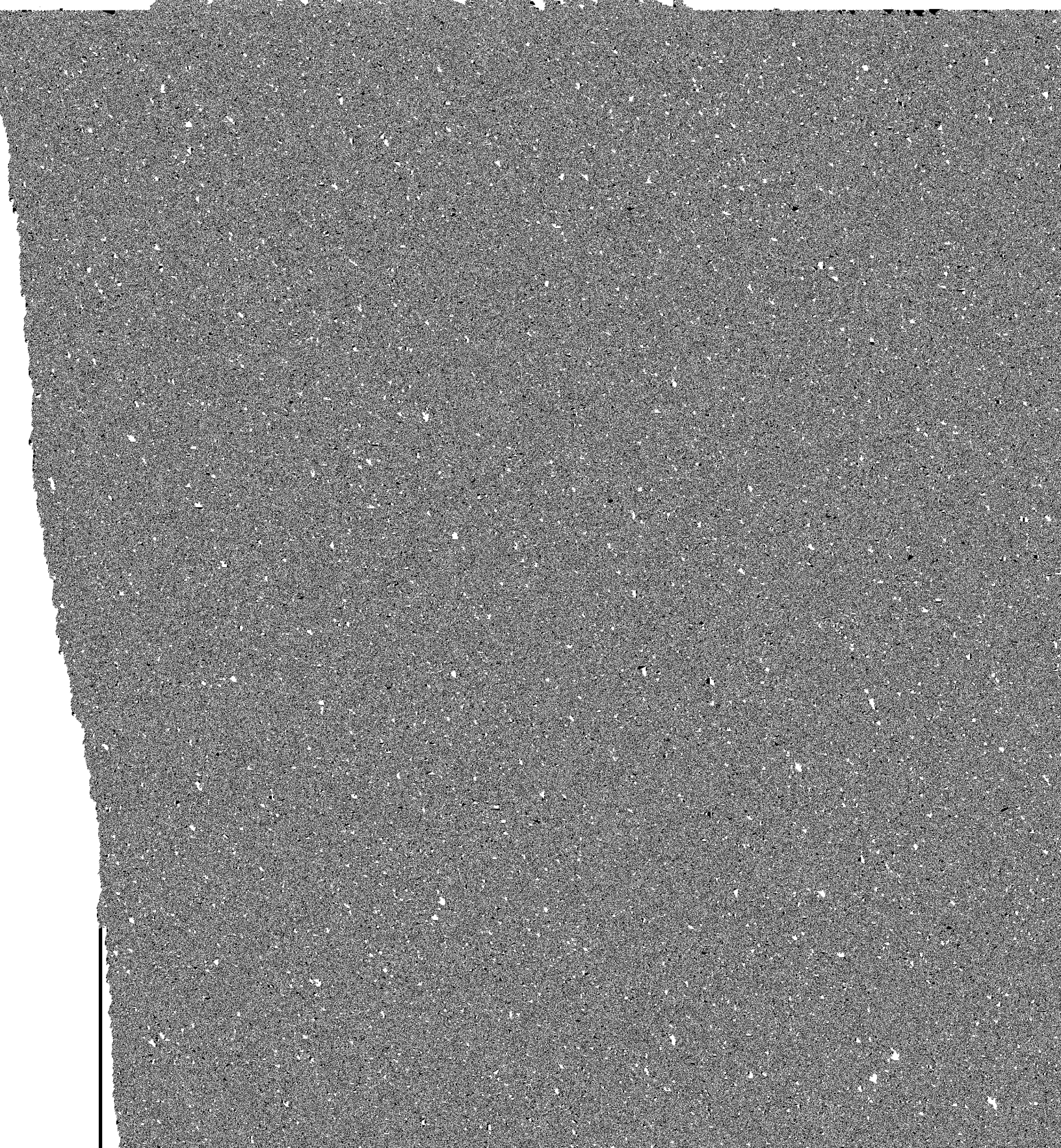
THE USER'S ACTIONS		THE MACHINE'S BEHAVIOR	
I	II	III	IV
Not available to the machine	Available to the machine	Available to the user	Rationale
S1	"Pull the latch labelled—" We did that. "Raise—" We did that. (studying display) Okay. Okay.	DISPLAY 2	
S2	"Lift up on the latch," We did that.		

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 1 and DISPLAY 2)



THE USERS ACTIONS		THE MACHINE'S BEHAVIOR	
I	II	III	IV
	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."		Rationale
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:		
S2	[All right: "Press: the Start button" Jean, you're doin' great.		
	(Both look to BDA)	Selects "Start"	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

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	<i>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.</i>		

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.⁵⁵ 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.⁵⁶

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". Where's the Start button? [looks around			Ready to Print

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	[points to display] Start? Right there it is.		
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.⁵⁷

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.⁵⁸ 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	Removing originals from the document handler.
S1	[Takes first page out of document handler]		
	REMOVES ORIGINAL	DISPLAY 10	
	[holding second page over the document handler, looks to		Removing the copies.

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	Rationale
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

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

[Brown, et al, 1976]

Brown, J.S., Rubenstein and Burton Reactive Learning Environment for Computer 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 & 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 & Perrault, 1979]

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

[Coombs & 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*

[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 & 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 & 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]

[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*

- [Schank & 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.
- [Scheffen, 1974]
Scheffen, 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 Communication*. New York: Academic Press, 1972.
- [Schegloff, 1982]
Schegloff, Emanuel Discourse as an interactional achievement: some uses of 'uh 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.

[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

About how many images are to be copied?

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

Describe the desired copies:

Number of copies:

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

Use standard paper? **Yes** No

Staple each copy? Yes **No**

Put images on both sides? Yes **No**

Reduce size of images? **No** 35% smaller 25% smaller 2% smaller

PROCEED

HELP

Comments

The currently described job will take about 1 minute.

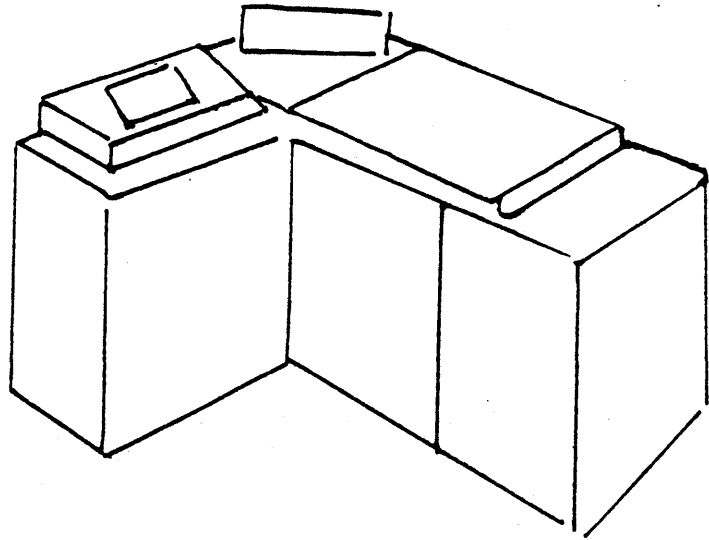
DISPLAY 0

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.

INSTRUCTIONS:

Please wait.

THE MACHINE

**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.

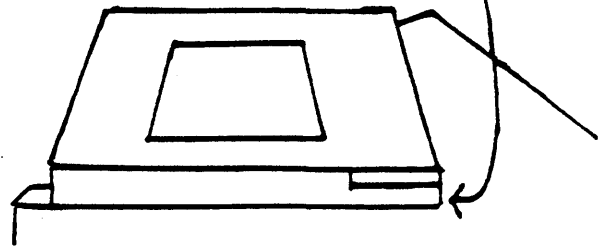
INSTRUCTION:

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

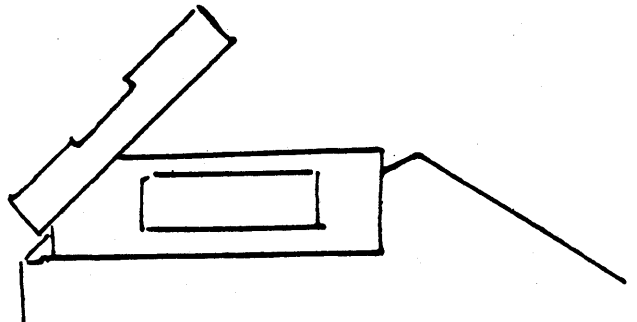
Raise the RDH.
(To enable placement of the bound document on the glass.)

How to access the BDA:

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



And lift up and to the left.



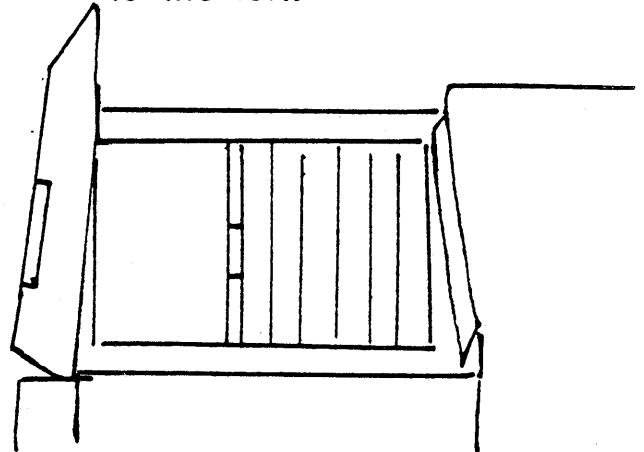
Change Task
Description

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)

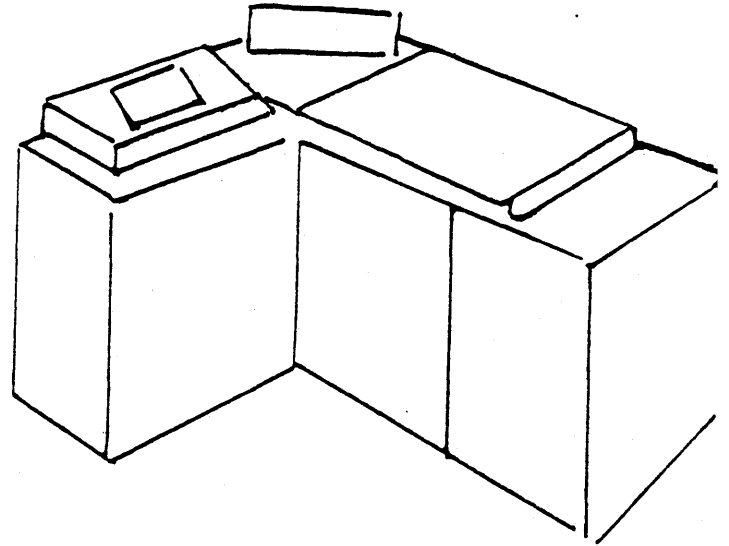
DISPLAY 3

OVERVIEW:**THE COPIER****ASSUMPTIONS:**

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:

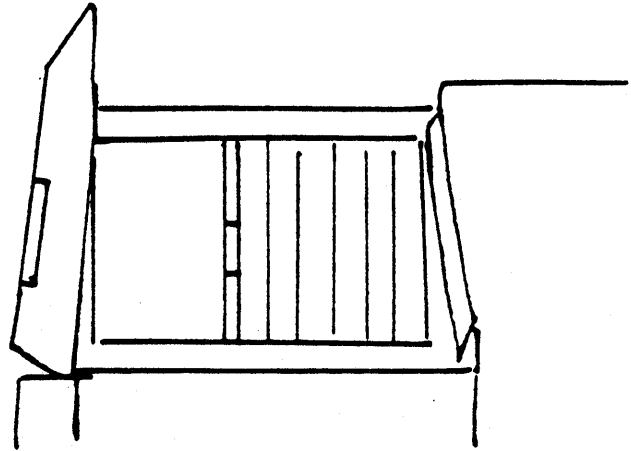
The copy of your original on the glass has been made.

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

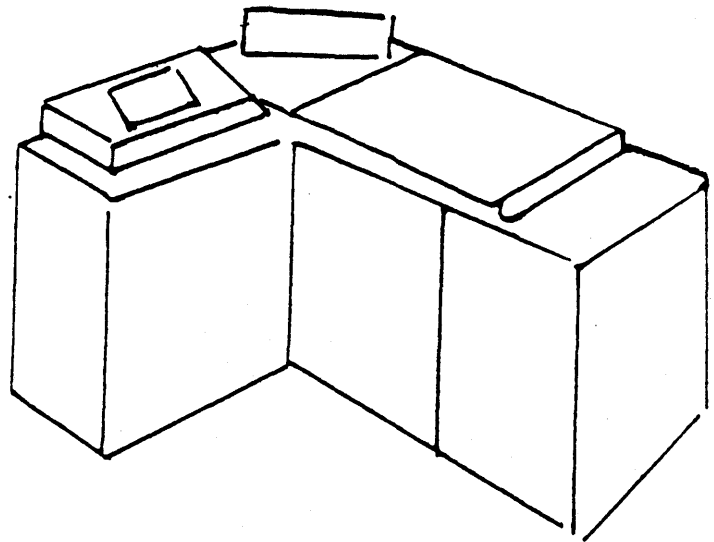
THE MACHINE**INSTRUCTIONS:**

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

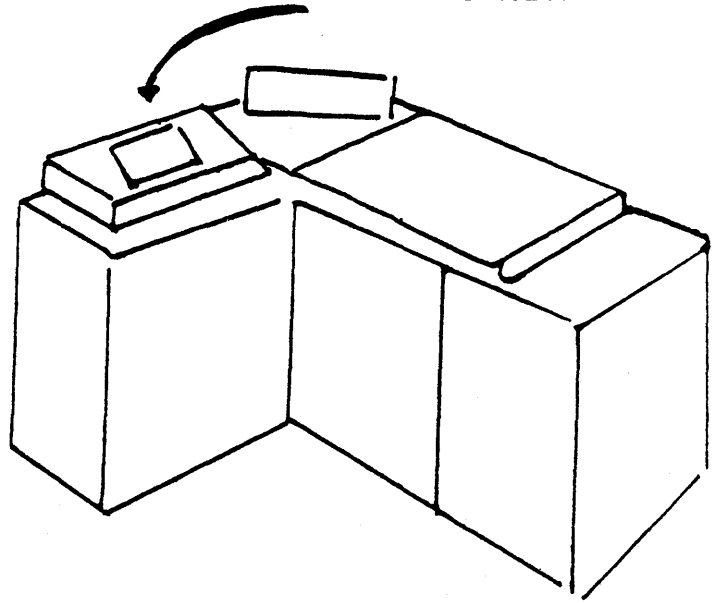
OVERVIEW:

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:

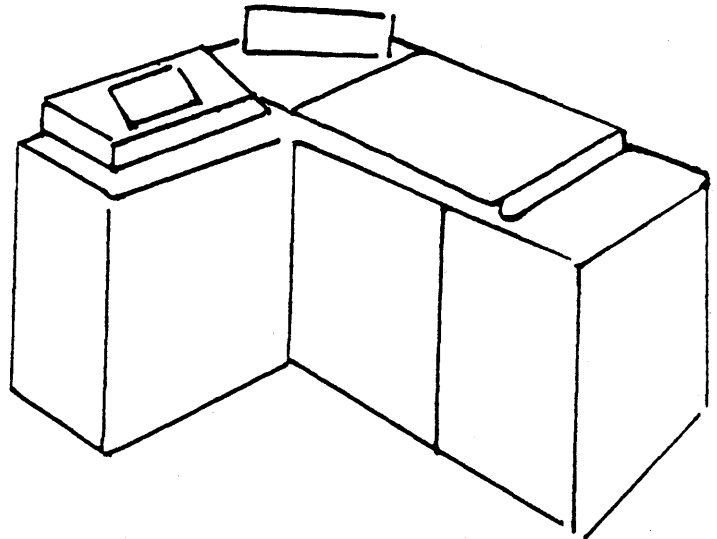
You can use the
Recirculating Document
Handler (RDH) to make
your copies.

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

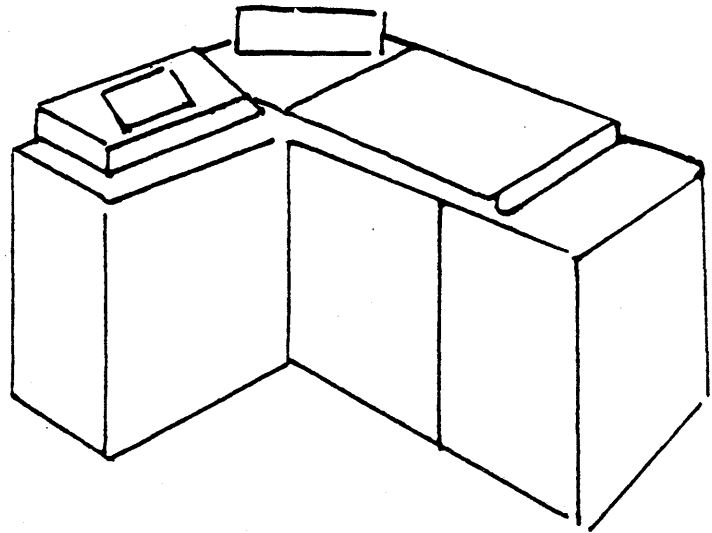
DISPLAY 8

THE MACHINE

ASSUMPTIONS:
The copies have
been made.

INSTRUCTIONS:

Remove the originals
from the RDH.



**Change Task
Description**

Help

DISPLAY 9

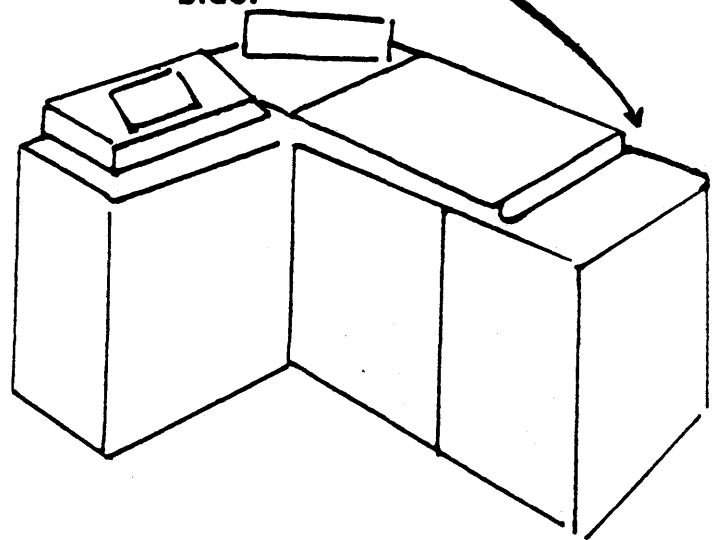
THE MACHINE

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

