AD-A017 681

DIALOGUE-BASED RESEARCH IN MAN-MACHINE COMMUNICATION

William C. Mann

University of Southern California

Prepared for:

Office of Naval Research
Advanced Research Projects Agency

November 1975

ARPA ORDER NO. 2930
NR 154-374

ISI/RR-75-41
November 1975

William C. Mann

Dialogue-Based Research in Man-Machine Communication

INFORMATION SCIENCES INSTITUTE
UNIVERSITY OF SOUTHERN CALIFORNIA
4676 Admiralty Way/ Marina del Rey/California 90291
(213) 822-1511

Preparation of this paper was supported by the Office of Naval Research, Personnel and Training Research Programs, Code 458, under Contract N00014-75-C-0710, NR 154-374, under terms of ARPA Order Number 2930.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ISI/RR-75-41	2. GOVT ACCESSION NO. -----	3. RECIPIENT'S CATALOG NUMBER ------------
4. TITLE (and Subtitle) DIALOGUE-BASED RESEARCH IN MAN-MACHINE COMMUNICATION		5. TYPE OF REPORT & PERIOD COVERED RESEARCH 6. PERFORMING ORG. REPORT NUMBER ----------
7. AUTHOR(s) WILLIAM C. MANN		8. CONTRACT OR GRANT NUMBER(s) N00014-75-C-0710
9. PERFORMING ORGANIZATION NAME AND ADDRESS USC/INFORMATION SCIENCES INSTITUTE 4676 ADMIRALTY WAY MARINA DEL REY, CALIFORNIA 90291		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61153N RR042-06-01 RR042-06 NR154-374
11. CONTROLLING OFFICE NAME AND ADDRESS HUMAN RESOURCES RESEARCH OFFICE ADVANCED RESEARCH PROJECTS AGENCY 1400 WILSON BLVD., ARLINGTON, VA 22209		12. REPORT DATE OCTOBER 15, 1975 13. NUMBER OF PAGES ~~30~~ 35
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) PERSONNEL AND TRAINING RESEARCH PROGRAMS OFFICE OF NAVAL RESEARCH - CODE 458 800 N. QUINCY STREET ARLINGTON, VA 22217		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE ---------

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

IN PRESS, *INTERNATIONAL JOURNAL OF MAN-MACHINE STUDIES*

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Artificial Intelligence, Cognitive Psychology, Complexity, Computer, Dialogue, Evaluation, Linguistic, Model, Performance, Process, Research Methodology, Theory.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

(OVER)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. ABSTRACT

This paper first surveys current knowledge of human communication from a point of view which seeks to find or develop knowledge that will be useful to computer system designers. The relevant scientific knowledge is found to be fragmentary and hard for designers to use.

Next, the problem of complexity is explored. Building a useful knowledge of human communication is an extremely complex task. Controlling this complexity and its effects, without giving up usefulness, is seen as the central problem in designing a research approach.

Finally, a new research methodology is presented. It contains some innovations that help control the complexity of the task, and others that make the results useful to designers. The methodology is unique in that

- It is based on case analysis rather than functional system design.
- The results are in the form of individual computer algorithms (much smaller than systems).
- The algorithms are transferable into useful (nonresearch) systems.

This research is an integral part of a larger set of research objectives to substantially improve man/machine communication--particularly for the growing level of on-line, interactive use of computers by the Department of Defense and the military departments.

DIALOGUE-BASED RESEARCH IN MAN-MACHINE COMMUNICATION

BY

William C. Mann

ISI/RR-75-41

November 1975

USC/Information Sciences Institute
4676 Admiralty Way
Marina del Rey, California 90291

The research reported herein is supported by the Personnel and Training Research Programs of the Office of Naval Research (Code 458), Arlington VA 22217, Contract Authority Identification Number NR 154-374, Contract N00014-75-C-0710, under ARPA Order Number 2930 from the Human Resources Research Office of the Advanced Research Projects Agency. The report is approved for public release; distribution unlimited. Reproduction in whole or in part is permitted for any purpose of the United States Government. Opinions expressed herein are those of the author and do not necessarily reflect policy of any agency of the United States Government.

I.

ABSTRACT

This paper first surveys current knowledge of human communication from a point of view which seeks to find or develop knowledge that will be useful to computer system designers. The relevant scientific knowledge is found to be fragmentary and hard for designers to use.

Next, the problem of complexity is explored. Building a useful knowledge of human communication is an extremely complex task. Controlling this complexity and its effects, without giving up usefulness, is seen as the central problem in designing a research approach.

Finally, a new research methodology is presented. It contains some innovations that help control the complexity of the task, and others that make the results useful to designers. The methodology is unique in that

- It is based on case analysis rather than functional system design.
- The results are in the form of individual computer algorithms (much smaller than systems).
- The algorithms are transferable into useful (nonresearch) systems.

This research is an integral part of a larger set of research objectives to substantially improve man/machine communication--particularly for the growing level of on-line, interactive use of computers by the Department of Defense and the military departments.

This research was supported by the Advanced Research Projects Agency of the Department of Defense and was monitored by ONR under Contract No. N00014-75-C-0710. A nearly identical version has been submitted for publication.

INTRODUCTION

When people communicate with machines they do so by specializing and extending their ability to communicate with each other. To design systems that can communicate with people, we need to know how people communicate.

This paper first surveys current knowledge of human communication from a point of view which seeks to find or develop knowledge that will be useful to system designers. The relevant scientific knowledge is found to be fragmentary and hard for designers to use.

Next, the problem of complexity is explored. Building a useful knowledge of human communication is an extremely complex task. Controlling this complexity and its effects, without giving up usefulness, is seen as the central problem in designing a research approach.

Finally, a new research methodology is presented. It contains some innovations that help control the complexity of the task, and others that make the results useful to designers.

The methodology is unique in that a variety of features of the approach that limit the complexity of algorithm development are described.

The development of these algorithms produces new knowledge about people and how they communicate. Using the algorithms in systems makes those systems more flexible and more skilled in communicating with people, because they incorporate methods that people already use, recognize and understand.

- It is based on case analysis rather than functional system design.
- The results are in the form of individual computer algorithms (much smaller than systems).
- The algorithms are transferable into useful (nonresearch) systems.

PART I: NEEDS AND SOURCES OF KNOWLEDGE OF MAN-MACHINE COMMUNICATION

The Technical Nature of Man-machine Communication

When people start to learn to use computers, they build on their knowledge of human communication. They are told about "commands", "statements" and "error messages". They are told to "tell the computer" things, and "it will tell you" others. Their own ability to understand and engage in two-party communication is the primary basis on which their skills in computer use are built.

The specialization from human dialogue to computer use has always been a drastic one. The languages, participants' goals, interpretation rules, means of control of computer interfaces are unfamiliar to most people, and often even hard to conceive.*

Man-machine interfaces today exploit very little of people's communication abilities, so their power to serve people's needs turns out to be strongly limited as well. The communication problem is a bottleneck, preventing people from getting easy access to computer methods and information that could otherwise be extremely useful to them. One direction of major improvement is therefore to identify particular expressive and receptive abilities of people, and then make machines appropriately responsive. But there is a catch. A special understanding of how people communicate is needed.

This paper describes a research methodology which is explicitly designed for developing the necessary understanding.

USEFUL MODELS OF COMMUNICATION

Consider the information needs of the designer of a system which includes man-machine communication.

The designer of a machine which is intended for man-machine communication is necessarily designing only half a system. The machine, and perhaps the medium, are subject to design, but man is not - he is only adaptive, not arbitrarily variable. He comes equipped with richly complex symbol manipulation abilities, which the computer system must engage effectively. ***So the designer needs a model of that other half of his system.***

---

* The work described here is intended especially to apply to users of interactive computer systems who are not programmers and who are not performing stereotyped tasks. A major share of the trouble comes from excessive reliance on command forms (Mann 1975).

What kind of a model of the human communicator would be helpful? It should be:

1. Complete enough to know what specific human capabilities will be exercised by a proposed system.

2. Explicit enough to predict human performance.

3. Detailed enough to represent the human performance limits realistically.

4. Detailed enough to support a tradeoff comparison of human versus machine performance of parts of the task.

5. Explicit enough about human methods so that they can be imitated in system designs.

Completeness, explicitness and detail can be found in varying degrees, of course. The models of symbol-manipulating man that are currently in existence are almost entirely inadequate for each one of the designers' needs suggested above. (Examination of current systems shows that designers' intuitions are also an inadequate source.)

There is no reason to expect that more adequate models will be easy to come by. Human symbolic behavior is known to be complex, unstable, and difficult to characterize. But it is extremely worthwhile to seek the knowledge, for several reasons:

- To reduce the frustration of using systems poorly matched to human capabilities and goals.

- To reduce the costs of error and labor associated with current systems.

- To increase our understanding of man.

On a very practical scale, it appears that even modest improvements in our current models of man would have significant direct benefits to designers. Many kinds of information tasks can only be performed to people's satisfaction by people. Satisfactory explicit methods are unknown. Having even approximate models of people's methods would allow designers to imitate and improve on them in many cases.

We have chosen processes as the appropriate medium to express the kind of models we are seeking. Practically, this means that the models will be expressed and tested as computer programs.

CURRENT SOURCES OF MODELS

Although no integrated and detailed account exists of how people communicate, there is an abundance of relevant information which can contribute to developing such an account. Some of the existing work in Artificial Intelligence, Linguistics, Cognitive Psychology, and the Philosophy of Language is particularly relevant. A selection of these is reviewed below, followed by our approach to an integrated theory.

Artificial Intelligence

Much of the recent work on processing of natural language has been undertaken as research in Artificial Intelligence. There is a sequence of developments on programmed question answering, another on tools for linguists, another on design of special purpose subsets of natural language for particular uses, another on teaching machines, another on solving algebraic word-problems, another on getting a program to obey natural language commands.

Most of these efforts are pursued as contributions to a theory of problem solving: "Here is some text. Use it to select an appropriate action," where appropriateness is to conform to human judgment about the text. The problem is a selection problem (or a synthesis problem). The program is to derive enough information from the text and its context to act.

Different programs involved human interaction in different degrees, from no involvement (for many) to provision of an interface for ongoing dialogue (as in Winograd's blocks-world program (Winograd 1973)). The current level of development of these programs does not reveal much about how the person who uses them functions as a communicator. The programs are interesting as models of receivers. Usually no claims are made for their correspondence to functional descriptions of people.

Developing appropriate knowledge representations will be a major factor in determining the generality of our dialogue models. We have the benefit of a long sequence of knowledge representation work. This background is currently represented in part by work at SRI (especially Strips and QA4 and their descendants) (Fikes and Nilsson, 1971; Rulifson 1972) and MIT (especially the Planner family) (Hewitt, Bishop and Steiger, 1973). (The recent work on Automatic Programming at ISI uses a knowledge representation which is in the same general style.) There is a widely scattered generation of systems bearing the label "semantic net," descended from relational systems (Raphael 1968; Quillian 1968, 1969; Elliott 1965.) Recent work has applied semantic net representation to a variety of information tasks (Lindsay and Norman, 1972; Rumelhart, Lindsay and Norman 1972; Norman and Rumelhart, 1975; Simmons 1973; Schank 1971, 1973; Hendrix 1973a, 1973b; Collins and Passafiume, 1974; Collins, Passafiume, Gould and Carbonell, 1973; Moore 1971; Moore and Newell, 1973; Srinivasan 1973). The Sophie system is particularly significant in combining a dialogue capability in restricted English with a variety of models of an electronic circuit being repaired (Brown, Burton and Bell, 1974). The Slate system combines relational representation with methods for retrieving, combining and interpreting larger units of information in a model of human short-term memory (Mann 1974).

Linguistics

The concerns of linguistics encompass several loosely related groups of phenomena. For our purposes, we will consider only those which are approximately invariant under the differences between speech and writing. The others, which deal specifically with auditory and articulatory phenomena are less germane to the general topic of communication.

Linguistic theory is a theory of symbols and compositions of symbols. Strings of symbols are characterized as grammatical, meaningful, equivalent, and so forth. Categories of strings, such as the categories of sentences and non-sentences, are distinguished by exhibiting formal systems in which they have differing derivational properties. In one of the mainstreams of current work, the theory is regarded as expressing the "competence of the native speaker," without any commitment to correspondence between theoretical entities and operations performed by actual speakers (Chomsky, 1965).

The "native speaker" is not speaking to a listener, and his competence remains fixed. No state changes are represented, and thus language is not treated as a vehicle of communication. Whatever consequences there are of the communication use of language are not directly representable as such in this kind of theory.

An example below illustrates on the one hand the difference in focus between this style and our concerns, and on the other hand the need for an explicitly communication-oriented theory.

> ...At this point discourse may become relevant in a way that it hasn't usually been taken to be for syntax. If you say to me "you're staring at me" and I reply "no, I'm not", are we to say that the deep structure of my sentence includes a transformed version of your sentence? It certainly wouldn't include "you're staring at me"; if it included any thing, it would be "I'm staring at you", but even that hardly seems plausible. And if the deep structure of my sentence does not include some transformed version of your sentence, then what is my deletion based on? It might be worth working on that question as a source for new ideas about deletions in general -- (Partee 1971).

Notice that the driving question is that of finding the right underlying form to use in the justification of the claim that "no I'm not" is grammatical. There is no consideration of what the transmitted content of "you're staring at me" might have been. Yet this knowledge seems to be just what is vital for answering Partee's question about the proper underlying form for representing the speaker's competence.

For communication modeling, a theory of speaker's competence is particularly helpful in suggesting what underlying units are used in communicating, and how the units are related.

Some linguists are working more directly on the kind of language that actually occurs (performance rather than competence). George Lakoff presents grammatical theories which he regards as subject (in principle) to refutation from psychological experiments. A significant grammar of the English of scientific publications has been developed at NYU (Sager 1972, 1975). Grimes, and many others, are exploring structural units larger than sentences (Grimes 1972). Several are working on theories composed of processes (Woods 1971; Kaplan 1971; Kay 1971).

These developments provide a rich source of ideas about the necessary components of a linguistically adequate theory of communication.

Cognitive and Experimental Psychology

The Cognitive Psychology movement has its roots primarily in Experimental Psychology. It aims to produce theories of complex human activities, especially involving memory, perception and language behavior. The units of analysis are usually complex objects such as sentences, stories or scenes, and the experiments may deal with either individuals (e.g., Newell and Simon 1972) or groups (e.g., Friend 1973, Colby 1973; Abelson 1973). The work proceeds from a view of man as an information processor, appropriately represented by theories containing information processes.

The current state of Cognitive Psychology is supportive of our goals for theory of communication in a number of ways. It is coping with elements such as commonly occur in communication, and is discovering structural properties which affect their communicative use (e.g., Chase and Clark, 1972), (Anderson and Ortony 1975). Retelling a story resembles ordinary communicative behavior in many ways that performance on a paired-associate memory test with nonsense syllables does not. Representation problems and the consequences of using particular representations are being dealt with explicitly (Simon, 1975). In particular there are production-system representation methods that have been through many stages of refinement (Waterman, 1971, 1973). Sufficient theories to account for particular complex behaviors are being developed, and an understanding of the evaluation of such theories is emerging. The field is demonstrating ways of coping with rich complexity in an empirical way.

Philosophy of Language

There is a line of development in philosophy which is currently addressing some of the central problems of communication, but from a different perspective. The development runs through early Logical Positivism, through the later work of Wittgenstein and work of J. L. Austin to recent work by John R. Searle, P. Grice and now many others (Wittgenstein 1953; Austin 1962; Searle 1969; Grice 1975; Cole and Morgan 1975). They are developing a theory of the relationships that occur between speakers and their words. The notion of an illocutionary act, and the theory of "speech acts" are important parts of this development. The theory identifies some aspects of the words which are important to understanding the words in dialogue, including, for example, characterizing the different kinds of commitment to the spoken words that speakers may express.

PART 2: COMPLEXITY

Complexity of the Theory

Our goal is to build a theory of human communication that specifies how it is accomplished. Approaching this goal we face a serious problem of coping with the complexity of the phenomena. The diversity of the contributing fields cited above is one evidence of this complexity. The complexity of a single isolated act of communication arises from a number of sources:

The complexity of personhood
Individual differences
The medium
The tasks being attempted by the communicators
The complexity of their available languages and conventions
The side effects of communicating
The extent and variety of relevant world-knowledge

The difficulty of dealing with these complexities is compounded and magnified by the weakness of currently available theoretical formalisms.

For theories of physical events, it is a practical necessity to describe at a variety of levels. For example, sterilization of a surface with boiling water can be described in terms of:

Epidemiology
Cell physiology
Organic chemistry
Atomic physics

each of which can be further divided yielding coherent descriptions at several levels. These descriptions are primarily in a relationship of composition; the objects of a more detailed descriptive level are taken as the building blocks of the adjacent less detailed one.

Likewise, human communication can be described at a variety of levels, some of them related by composition relations. Our goal is not involved in all of the complexity of communication because we are not trying to address all of the possible levels at once. But this is still a complex matter. In order to meet the designer's needs outlined above, a "sufficient" theory is needed -- that is, one that specifies how a representative communicating agent acts. Such a theory is equal in complexity to the task of building a robot that acts that way.

The difficulty is not the volume of information involved, but its diversity. Having computers a million times faster or larger would not allow us to build that robot, since we would not know what to put into the programs. The limitation comes from the state of the theory itself.

Controlling Complexity In The Research Activity

Complexity in the theory leads directly to complexity in the activity in which the theory is developed. The research cannot be less complex, since it takes effort to establish each independent point. In fact, elaborate effort is often required to establish single hypotheses. How can the complexity of the theory-building effort be controlled and prevented from causing the effort to fail?

Just as we reviewed the state of some communication sciences, we could review the state of the art of complexity limitation in science. But this is far too great a task. A large fraction of the techniques of science -- instruments, laboratories, controlled comparisons, formalisms of expression -- serve to limit complexity. Closer to home, there are several methods in common use:

Strong restriction of the world of knowledge (the micro-world method) The restriction comes in two flavors: selective attention to real-world knowledge, exemplified by the Scholar, Sophie systems and by recent work of Malhatra (Brown et al. 1974; Collins et al. 1973, 1974; Carbonell 1970; Grignetti and Warnock, 1973; Malhotra 1975) or a wholly synthetic world of objects and concerns, exemplified by the Blocks World (Winograd 1973).

Selection of simple problems as the subjects of communication, such as Cryptarithmetic (Newell et al. 1972)

Limitation of the scope of the theory, for example, to the syntactic competence of native speakers of a language. This is a strong restriction, in that the nature of the competence is assumed not to change as a result of the act of speaking or hearing. State changes of the language users are not represented, and so any theory of communication which includes such state changes is outside of the scope.

Failure to reconcile theoretical fragments, represented either by an informal (descriptive) coordination of formal results, or none whatever. Serious problems of consistency of the fragments remain. Collections of fragments of theory do not fulfill the notion of sufficiency described above.

Use of synthetic data. It is common in Linguistics and Philosophy to deal with constructed examples that highlight particular points. This is a legitimate and fruitful approach, but it tends to leave large collections of phenomena unexamined.

> "593. A main cause of philosophical disease -- a one-sided diet: one nourishes one's thinking with only one kind of example." (Wittgenstein 1953)

It also tends to leave parts of the theory unreconciled, since the process of creating clear examples eliminates complicating interactions.

These ways of reducing complexity of research all share a common risk: That necessary information will be ruled out of the investigation, either by being excluded from the data or excluded from the scope of possible results. (For example, it may be necessary to have data on context or a notion of communication in order to develop a sufficient theory of meaning.) Because of this risk we should seek and prefer controls on complexity that do not restrict the scope of associated data or theory.

PART 3: RESEARCH METHODOLOGY

A desire for sufficient theory, and for strong controls on the immediate complexity of the research, have guided the design of the work described below. We believe that it addresses the concerns raised above in an interesting and powerful way.

Overview

The complexity of building new theoretical hypotheses is limited by requiring that a new hypothesis only account for a single case. Hypotheses (expressed as processes) are validated and/or progressively generalized by requiring that they account for additional single cases. The size of a case is controllable by the way we select it.

Incremental progress toward a sufficient theory is achieved by developing it in parts, each part being in the form of a process, and the overall theory in the form of a set of cooperating processes. The processes are made explicit by stating them in a suitable process formalism, a programming language. The processes are tested as parts of a computer program, and their mutual consistency, compatibility and coverage are made explicit and tested by combining them in this way.

Use of actual human communication helps retain comprehensiveness and contact with a full range of communication phenomena. We choose to work on dialogue, partly because it is the starting point for learning man-machine communication, and partly because we want to avoid the complexity of side effects (e.g., A speaks to B for the benefit of C).

The complexity of knowledge required is limited to that actually involved in particular cases being studied, rather than spanning a knowledge domain. The required knowledge is identified by examining the cases, since the studies work from transcripts. Certain central phenomena are selected to lead into the development and gauge its progress.

Dialogue Process Modeling

This modeling effort involves a series (cycle) of experiments with human dialogue, having results which are processes (represented as computer programs) which collectively are able to follow and make sense of dialogue transcripts. The effort addresses a limited range of descriptive levels.

What do we plan to exclude? All of the long-term communication effects, from the multi-century development of language up to the personal long-term effects, including the gross developments of language acquisition, are outside our scope. At the detailed levels, articulatory or phonological or auditory phenomena are also outside the scope. The smallest unit of analysis will ordinarily be the word or morpheme. The goal is to understand immediate effects in communication with words.

The effort addresses a limited range of described levels. The cycle of experiments works as follows:

1. In each single experiment, processes are developed which can cope with the transcript of a single dialogue. These processes are a kind of empirical hypotheses.

2. The results of multiple experiments are examined to identify those processes which reoccur, which therefore successfully cope with phenomena from several dialogues. These are the verified processes which we expect to successfully transfer into future communication systems. In multiple experiments a kind of refinement and verification of hypotheses occurs, the hypotheses being stated as processes and debugged as computer programs.

DATA SOURCES

The data are dialogues. For stable results, we need to cope with a diversity of kinds of dialogues, including diversity of source, selecting on the basis of experience. A priori, the following are sources of readily available data with interesting characteristics:

- We have in hand a wide selection of Operator-User dialogues from the TENEX time sharing system, including many appeals for help or advice.

- NASA has provided us with spacecraft-to-ground transcripts and voice tapes of the Apollo 13 mission (the one with the blowout). These provide another interesting kind of task-oriented dialogue.

- Radio dial-up talk shows. These provide well-bounded context and discussion between strangers.

- We construct some experimental situations ourselves.

We use only media in which the participants are separated (telephone, linked typewriters, computer mediated exchange, etc.) because we need to be able to represent the symbolic content of the exchanges fully and easily. Facial expressions, gestures and the like are beyond our abilities to encode well enough to make them useful to us. (Also, these media are most like the common man-machine media.)

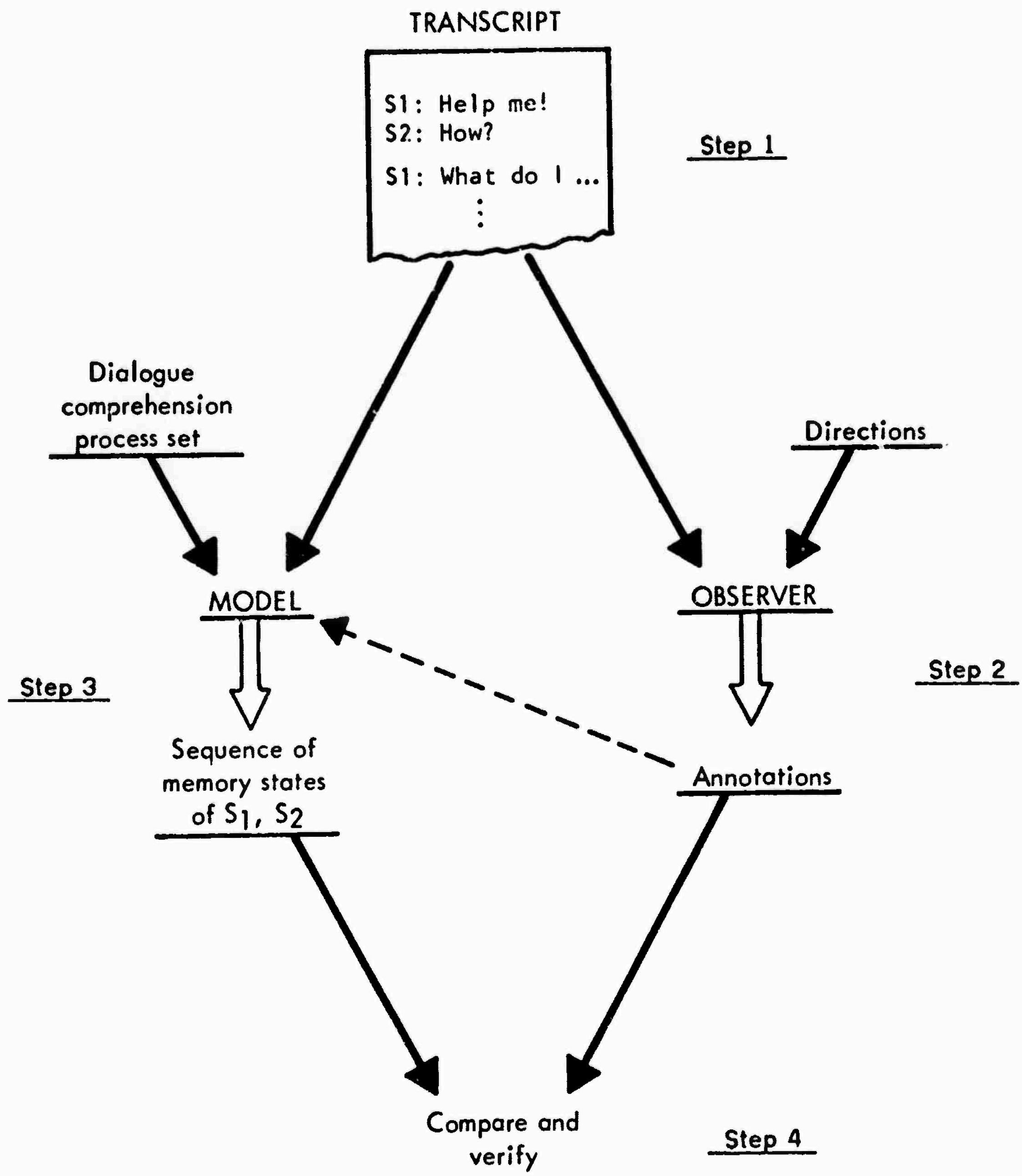

Figure 1 - An Experiment with a Single Dialogue

SINGLE EXPERIMENT

Dialogue process modeling is a series of experiments which develop communicating processes. The single experiment consists of four parts:

1. Capture human dialogues.
2. Identify important phenomena in the dialogue.
3. Create processes that can follow the dialogue.
4. Evaluate the processes relative to the phenomena identified in item 2.

Each of these is explained in detail below. Figure 1 illustrates the experiment.

Experiment Step 1

Step 1 is to capture the dialogue. A transcript of the dialogue, suitable for both computer-reading and human reading, is produced. The remainder of the experiment deals with this transcript, in particular with reconciling a human reading and a computer reading of it.

This step (and all prior steps) also includes some selection of transcript material for further analysis, since dialogue is abundant and easy to create, and there will always be far more than we can analyze. This is one of several controls we have on the difficulty of our task. Another is the length of the exchange that we select.

To help follow the explanation, consider the following example, which is an actual operator-user dialogue, from the ISI TENEX computing facility, cleaned up for readability but essentially unaltered in content:

The dialogue below, between parties O (a computer system operator) and L (one of the system users), was mediated by the system. O and L were at separate locations, typing. RUNOFF is a commonly used text-formatting program. Spelling, punctuation and visual format have been edited for readability and privacy.

****

L:
How do I get RUNOFF to work, I keep executing it but it just grabs
my input file and then says done but gives me no output? Go ahead.
O:
The output comes out on the line printer
L:
Throw it away but can I get it to go to a file? Go ahead.
O:
Confirm your commands with a comma and you'll be queried for files,
etc. Go ahead.
L:
Thanx mucho
BREAK

****

Experiment Step 2

Of course, a dialogue is generally a unique, one-of-a-kind event. It contains systematic features that make communication possible, together with a lot of unexplainable, idiosyncratic detail. Understanding the regularities that govern a dialogue requires describing them in common terms. The first stage of description in common terms involves deriving systematic data from dialogue. For this work, we use a person we call the Observer.

Step 2 in the single-experiment figure shows the Observer creating a commentary on the dialogue, based on a set of categories of observation described below.

The Observer is asked to assert only those phenomena for which he has high confidence that his interpretation would be widely agreed upon, say, by 95% of a group of people as competent as himself in communication. We do so in order to keep the attention of the modelers on a set of phenomena for which we expect a consistent set of technical accounts. There is presently no profit in dealing with marginal cases. Our situation is like that of the linguists interested in grammaticality

"Even though few reliable operational procedures have been developed, the theoretical (that is, grammatical) investigation of the knowledge of the native speaker can proceed perfectly well. The critical problem for grammatical theory today is not a paucity of evidence but rather the inadequacy of present theories of language to account for masses of evidence that are hardly open to serious question." (Chomsky 1965)

The Observation Process

The role of the Observer is not to tell why the dialogue went as it did, nor what methods the participants used to create their parts, since these are theoretical questions. Such questions belong to the accounting for the observations rather than to the observational process. Figure 2 below illustrates the Observer's activity, annotating transcripts of the dialogue according to instructions and his own understanding and judgment.

The Observer of communication activity is a kind of instrument. His judgments on the condition of the communication constitute the properties of the communication that need to be explained. They are analogous to the judgments of grammaticality in linguistics, which attach the * to some strings and not to others.

The Observer is a kind of reducer of the data, an inscrutable perceiver and filter, whose presence is acceptable and necessary because his comments somehow represent his whole community of communicators, and therefore represent underlying regularities which arise from the conventional communication methods of that community. Thus in this methodology the observer must be a person: he cannot be replaced by a computer program.

The Observer is able to answer sufficiently specific questions about the dialogue with high reliability and repeatability, and is normally able to state correctly the communication effects of parts of the dialogue.* Of course, the reliability and repeatability of Observers must be verified experimentally.

In order to have the observations in an experiment reflect a single coherent point of view, each dialogue is modeled relative to the comments of a single observer rather than with pooled, possibly inconsistent, observations.

Part of the work of the project is to develop this framework of observation into an easily used tool. We intend that the theory rest on "obvious" cases, and that the observation process be easy to understand and use without extensive training.

---

*The Observer is doing a kind of encoding of the interaction. There is a line of development of group interaction coding methods in social psychology, including for example Bales, 1951. We expect these developments to be of some help, but our demands (for example, relative to content) are quite different.

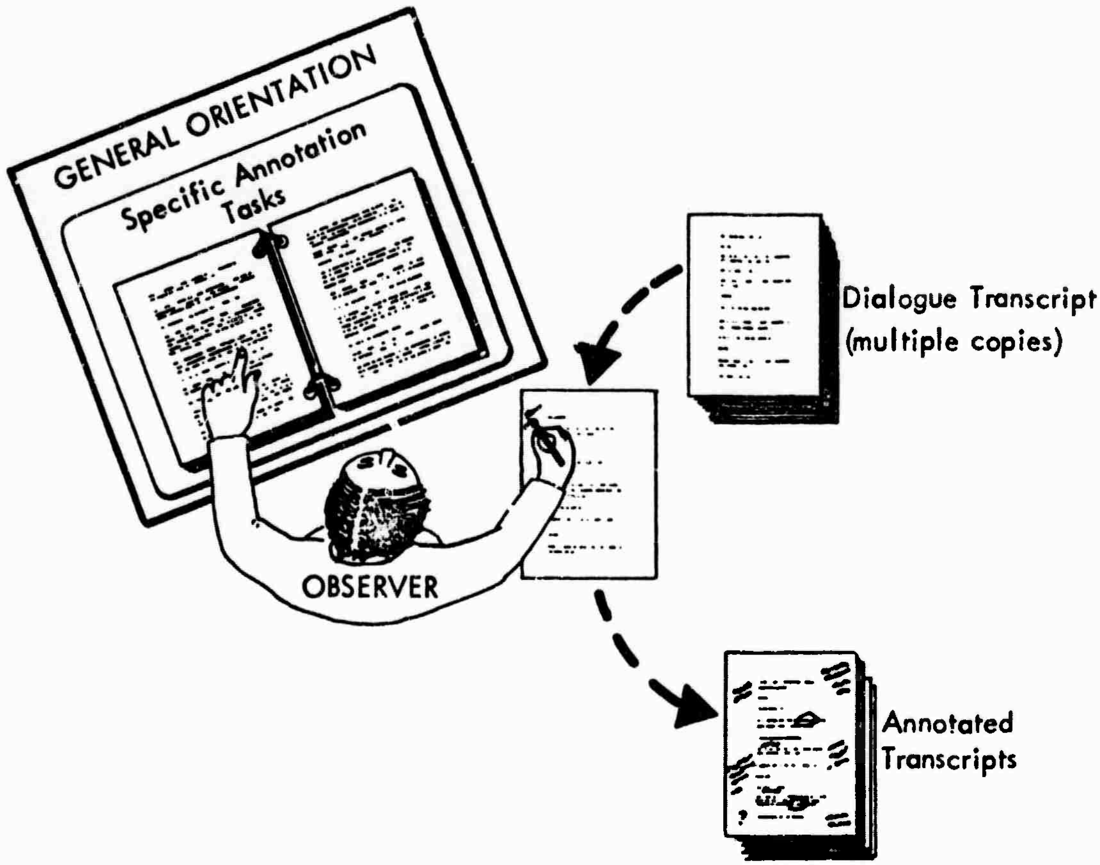

Figure 2 - The Observer's Task

The Observer is trained to seek several prescribed categories of phenomena, which are described below. The function of each of these categories of observation is to influence the attention of the model-builders toward some aspect of ordinary, successful human communication. By attending to a suitable number of cases, the model builders are to derive processes whose structure reflects the regularities of at least those cases, and hopefully much more. Since each of these models must account for multiple categories of observations, interactions and mutual dependencies among people's methods can be reflected in the processes.

Some parts of the dialogues will not be addressed directly by any observation, but processing them will nevertheless be necessary as a prerequisite to successful processing of the rest. Processes may be involved that only contribute indirectly to accounting for observations. (For the categories described below, parsing processes are such.) These processes can be developed *and their execution contexts and consequences identified. Therefore they can be verified just as if they were directly accounting for varieties of observation.*

So the particular observation categories chosen are influential, but ultimately arbitrary. They are not primitives, and they do not necessarily combine to give a complete picture of human communication. Experience may suggest that one or all be replaced.

There is actually very little risk that the present formulation will fail to get at the key communication phenomena. If there were some basic sort of process whose effects were not represented directly in the scope of the observations, it would be modeled anyway if it had major consequences in the scope of the observations, which is just what we would expect of any such basic process.

Categories of Observation

The Observer is trained to identify seven more-or-less independent kinds of phenomena:

1. Repeated Reference
2. Requests
3. Expression of Comprehension
4. Topic Structure
5. Correction Actions
6. Similar Expressions - (two kinds)*

---

*All of the work of the Observer is described in detail, with definitions, instructions and examples, in a recent report ISI/RR-75-33, (Mann, Moore, Levin and Carlisle 1975).

Repeated Reference deals with what is mentioned or referred to. It includes the actions of nouns, pronouns and verbs in English. We deal with two kinds: ordinary reference and text reference. In order to get reliable observations, we ask the Observer to identify multiple mentions of the same thing, rather than requiring him to spell out just what is referred to.

For example, we would expect the Observer to note that "RUNOFF" and the first 2 occurrences of "it" refer to the same thing in the example dialogue above.

Requests deals with all kinds of real and apparent behavior-seeking in dialogue. The Observer makes detailed annotations about 5 subcategories:

1. Questions -- request immediate, verbal response.
2. Orders -- request immediate, nonverbal response.
3. Directives -- request certain behavior in the future.
4. Rhetoricals -- look like Requests but are not.
5. Prohibitives -- request *to not do* something.

The Observer annotates how the Request occurs, and also how in most cases how the partner responds. In the example, "can I get it to go to a file?" is one of the Questions the Observer would find, and he would note that eventually it was answered.

Expression of Comprehension deals with the direct and indirect evidences given by one partner to the other that he has comprehended something which has gone before. In the example, O's way of speaking of "output" indicated that he comprehended what L said.

Topic Structure deals with the way topics of conversation are begun, acknowledged, carried on, suspended and dropped. The short example above has only one topic, which persists over the whole dialogue.

Correction actions are parts of the dialogue in which one party cancels or changes some previous part of the dialogue that has already been understood by each party (often in two different ways.) There are no correction actions in the example above.

The idea of Similar Expressions is related to the idea of paraphrase, but is looser. The Observer scores alternative ways of saying what was said, judging whether they would be suitable substitutes under some imaginable circumstances. The intent here is to force the model to represent underlying functional equivalences rather than making it responsive only to surface similarities between expressions in communication. The two kinds of Similar Expressions arise because under one condition, the Observer is not shown the text preceding the fragment in question, and in the other condition he sees the entire preceding dialogue.

In the example above, O might as well have said "RUNOFF sends its output to the printer" in the second turn. In some contexts, "Get it away from you by throwing it" might function as well as "Throw it away" , but in this context it would not. These are representative findings from the observation of Similar Expressions.

In every category, the Observer is instructed to assert only on cases that he feels are clear and easy to decide. We want to avoid putting much attention on marginal cases.

Experiment Step 3

The next major step (Step 3 in the single-experiment figure) is to create a process model (a new computer program for this specific dialogue) that can cope with the dialogue transcript. Informally, this means that the model must ***maintain a simulated knowledge state for each participant*** that is adequate for supporting continuing interpretation of the dialogue. The adequacy of this continuing interpretation is assessed by the Fidelity evaluation and Recurrence evaluation methods described in later sections.

The program must follow the dialogue in enough detail so that references of pronouns and noun phrases are identified, requests and how they are resolved are identified, corrections are correctly carried out, topic flow is correctly assimilated and so forth.

The model is ***not*** attempting to simulate the Observer. It is simulating ***the receptive acts of the participants*** in the dialogue. The model takes as input the transcript of the dialogue, with the text produced by each party identified. It has a portion of its memory allocated as a simulated memory for each of the participants. A participant's memory indicates provision for general knowledge, what is relatively static, and current awareness, which changes with each turn. The model's task is simply to keep these memories up to date.

This step is the crucial step toward general communication processes. Here, all of the issues of "How can we cope with this kind of thing?" are raised in a context in which there is enough information on the specific use to suggest solutions. Like any serious design task, the creation of a proven model is an obscure tangle of specifications, conjectures, revisions and tests. An idealized sequence of steps is shown below.

****

STEPS IN MODELING A DIALOGUE

1. Write approximate memory states that would be adequate as the starting point, ending point, and as various intermediates, based on the dialogue transcript and the Observer's commentary.

2. Represent the knowledge which is relied upon in the dialogue.

3. Write processes that will collectively move through the memory states if correctly invoked.

4. Create invocation conditions on the processes which will cause them to get invoked at only and all the right times.

ITERATE WITH SUCCESSIVE APPROXIMATIONS

****

We do not wring any dialogue dry. The modeling is pursued to some level of detail at which the generality of the process being built is in serious doubt.*

The simulated memories of the participants will be structured to represent different kinds of knowledge and knowledge status. Current awareness will be distinguished from other available knowledge. Knowledge which a participant believes is not known to the other participant will be distinguished from knowledge which he regards as shared. Linguistic knowledge, facts about the world, and knowledge about the environment of the dialogue will be distinguishable. Each simulated participant will have in memory a model of the other party and of himself. Each party's goals and abilities will be represented. Of course, the depth of representation will depend in each case on the demands of the dialogue.

---

* This is another control on immediate task difficulty.

The subprocesses of a process model will include at least the following:

1. English Parser
2. Inference Processes
3. Object Completion Processes
4. Processes Which Perform Corrections
5. Hypothesis Comparison Processes
6. Evidence Propagation Methods
7. Time-knowledge Manager
8. Awareness Manager (selective forgetter)
9. Processes Which Use and Modify Own Goals
10. Processes Which Maintain Knowledge of Partner's Goals

For all of these parts, there are available ideas in current and past work cited herein.

We have a strong advantage here over most programming activity in that all of the performative demands on the program to be produced are explicit and available when the program is written. The input (the dialogue transcript) and the standard of evaluation (the Observer's commentary) are open to inspection. (Hindsight is clearer than foresight.)

People's ability to anticipate the contingencies of program input is rather limited. This is one reason that debugging is the major cost component of commercial programming. Having the input at hand is much more efficient.

Also, we can make use of methods that are known to be inadequate for the general case. We can explore their function for specific successful cases, and either generalize after gaining experience or find a set of methods that jointly cover the requirement.

Another advantage is that each dialogue is a fresh problem. There is not a backlog of design commitments, habits, and development cost which must be accommodated. From dialogue to dialogue, we can keep what works and drop the rest. The amount of work carried forward from one dialogue to the next depends on their underlying similarity. This is another controllable feature, since we select dialogues and dialogue sources as we go along. From the point of view of scientific method, we are being driven by the data rather than by anticipations of what the data might be, or by anticipations in the form of system specifications. We expect that this practice will keep the focus away from peripheral and artificial problems.

Experiment Step 4

Step 4 of the experiment is called Fidelity Evaluation. It is a comparative evaluation of the process model relative to the Observer's commentary. The purposes of the evaluation are:

- To identify program states and actions which are in agreement with the Observer's comments.

- To identify Observer's comments for which the corresponding process states and actions are either absent or disagree with the observations.
- To identify the directly contributing processes whose states and actions are in agreement with the Observer's comments.

The comparison is between the observations of a dialogue and the trace of the model for that dialogue. (The trace is a sequential record of all of the actions of the model during the interpretation of the dialogue.) For each observation which the Observer asserted, the trace is examined to see whether there is a corresponding action or partial state of the model. If so, then the observation is counted as having been successfully modeled. The time (in the trace) at which the observation was fulfilled is identified.

For each successfully modeled observation, the trace is examined to identify the directly contributing processes. These are usually those which are active (in the stack) at the time of fulfillment of the observation, together with the processes which produced the conditions, knowledge or data which controlled the branching of the model's control flow immediately preceding the time of fulfillment.

Multiple-Experiment Comparisons

<u>The Need for Comparison</u>

A successful accounting for a single dialogue yields a set of processes of essentially unknown value. The value of the processes, both as a scientific account and as components of future systems, depends on their generality, i.e., their effectiveness on different communications, different environments and communication goals. The high-value processes are those which have high coverage of the scope of human communication (illustrated in Figure 3 below). There is a spectrum of values from the totally ad-hoc process which can cope with only one dialogue, to the totally general process which copes with every case in its domain. The next step, Multi-experiment comparison, is intended to identify processes which, in our sample of experiments, have been found repeatedly effective.

We identify our high value processes as those which are reasonable on a variety of dialogues. A diversity of dialogue sources, media and goals in our experiments is necessary in order to get a good measure of value for our processes. This diversity also serves the system designer who will use the processes, since it gives him the evidence that the processes are reliable and transferable.

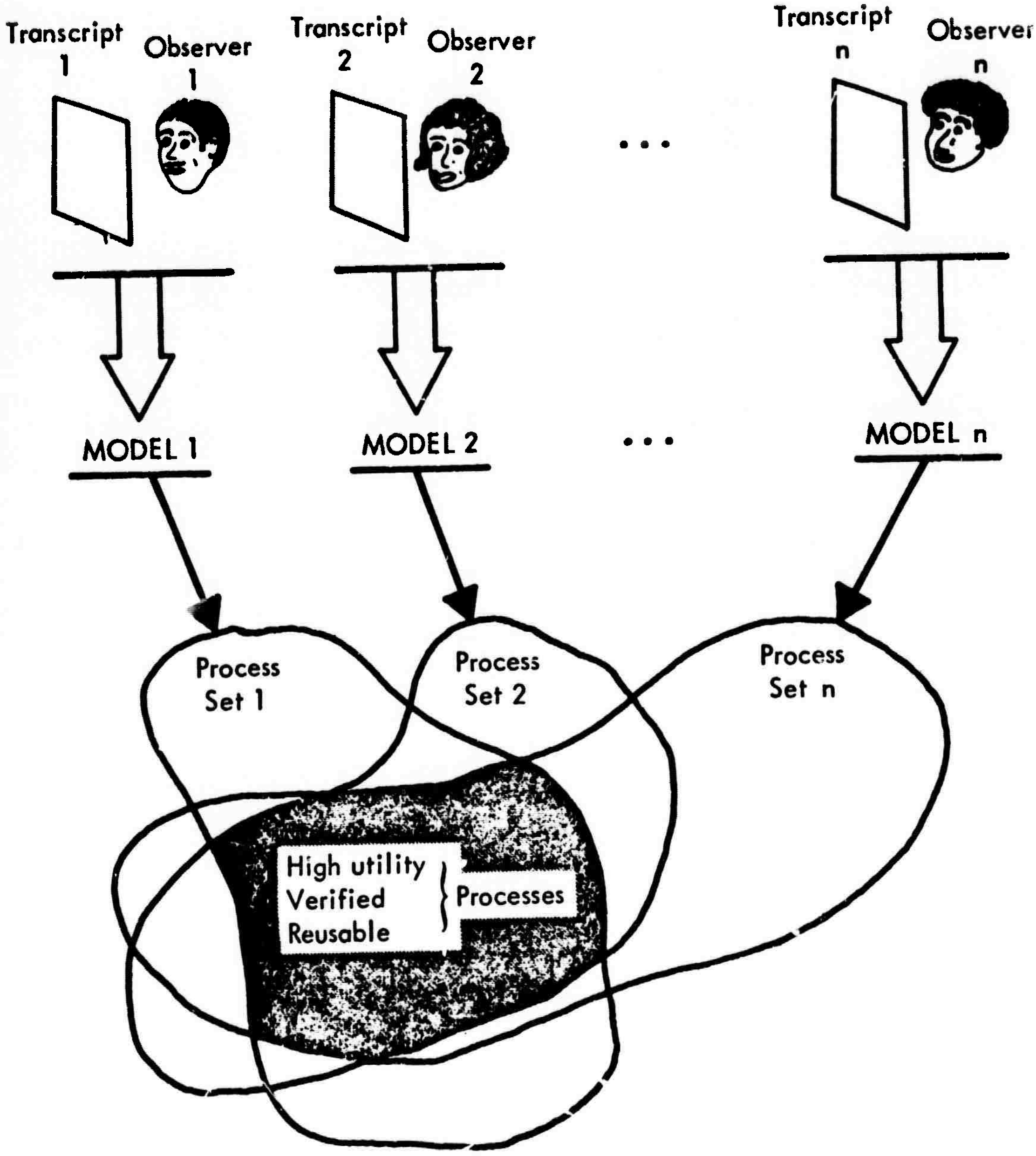

Figure 3 - Evaluating Multiple Experiments

Recurrence Evaluation Method

The Recurrence Evaluation method is quite simple. At the end of each experiment, we identify those processes that have:

- been used unaltered from previous experiments, or could be retrofitted into previous experiment;
- been identified as contributing to successful accounts of the Observer's comments.

These are the only processes that are considered to have demonstrated value.

How Recurrence Evaluation Drives Algorithm Development

The comparison in Recurrence Evaluation has a pervasive effect on the whole modeling process. Any processes that are dependent on features of a particular dialogue, or that work in a known-unreliable way, will not show demonstrated value at this step. ***The goal in modeling is to maximize the product of this comparative step.*** This leads the modelers to include general features in processes whenever it is sufficiently convenient to do so, to seek the unifying principles, and to represent them in models. It avoids the problem of creating many independent, incommensurate and finally useless single-case models.

Deriving Valuable Results

There are several approaches to the products of this methodology, depending on what is regarded as valuable. The activity can be supported from various viewpoints, of which we will consider two briefly:

1. Scientific Knowledge
2. Application in Computer Systems

For both of these, the primary results are processes rather than whole systems, since the individual process is a decisively more convenient unit of investigation or application than the enclosing system in which it occurs.

Scientific Values

We have discussed several branches of science above which express their theories in terms of discrete symbolic processes like those we develop. In others (e.g., the medical theory of communication disorders and brain damage) such theories are possible but seldom found. In any of these, the process we develop can be considered as theories of phenomena of that branch. There is very widespread interest in science in understanding human symbolic activity.

The evidence for the effectiveness and generality of a process is explicit and easily examined -- it is the dialogues and observations which the process has contributed to accounting for, and the model traces which exhibit the manner of contribution. The empirical approach of this method is intended to make its results directly relevant.

Beyond this, we anticipate that this approach will eventually have an integrating effort, providing common representations across discipline boundaries and coordinating knowledge developed in the separate disciplines. For this purpose, the general framework and its processes are both of interest.

Application in Computer Systems

One reason for designing the methodology so that it produces processes is that processes are primary components of computer systems. We plan relatively direct transfer of convenient processes into working system environments. The earliest applications are planned as modifications of existing systems. (For example, a process for correcting one's previous statements might be moved from a dialogue model into an existing text editing program so that the editor could understand a method that people use for correcting themselves, applying it to previously entered text.

This "Method of Embedding" provides a very direct transfer of the technical results into applications, which was one of the goals described above. Algorithms are selected informally for embedding.

Of course, the process can be designed into new systems as well. Again, the transfer is very direct, especially relative to research that produces only factors for the designers' consideration.

Experience with This Methodology

The research effort using this methodology has been under way less than a year at this writing. Results so far are extremely encouraging, but inconclusive. One major progress report is available, (Mann et al. 1975) and more are forthcoming. By the time this paper is published we expect much more to be known.

Summary

The approach described above includes several innovations in man-machine communication research. ***Human dialogue*** is used as the primary data source. Communicating processes are derived from communication transcripts by a ***case analysis method.*** The method develops ***explicit processes*** for use in systems which include man-machine communication.

REFERENCES

Abelson, Robert P., "The Structure of Belief Systems," ***Computer Models of Thought and Language***, W. H. Freeman and Company, San Francisco, 1973, pp. 287-340.

Anderson, Richard C. and Andrew Ortony, "On Putting Apples into Bottles -- A Problem of Polysemy," ***Cognitive Psychology***, Vol. 7, No. 2, pp. 167-180, April 1975

Austin, J. L., ***How to do Things with Words***, The William James Lectures delivered at Harvard University in 1955, Harvard University Press, Cambridge, Massachusetts, 1962.

Bales, R. F., ***Interaction Process Analysis***, Addison-Wesley Press, Inc., 1951.

Brown, John Seely, Richard R. Burton, and Alan G. Bell, "SOPHIE: A Sophisticated Instructional Environment for Teaching Electronic Troubleshooting (An Example of AI in CAI)", Bolt, Beranek and Newman, Inc., Report 2790, (A.I. Report 12), March 1, 1974, pp. 24-25, 96.

Carbonell, J. R., ***Mixed-Initiative Man-Computer Instructional Dialogues***, Bolt, Beranek and Newman, Inc., May 31, 1970.

Charniak, E., ***Toward a Model of Children's Story Comprehensions***, Massachusetts Institute of Technology, December 1972.

Chase, W. G. and Clark, H. H. "Mental Operations in the Comparison of Sentences and Pictures." in L. W. Gregg (ed.), ***Cognition in Learning and Memory***, John Wiley & Sons, Inc., New York, 1972.

Chomsky, Noam, ***Aspects of the Theory of Syntax***, The M.I.T. Press, 1965, pp. 19-20.

Colby, Kenneth M., "Simulations of Belief Systems," ***Computer Models of Thought and Language***, W. H. Freeman and Company, San Francisco, 1973, pp. 251-286.

Cole, Peter and Jerry L. Morgan, (ed.), ***Syntax and Semantics***, Vol. 3, Academic Press, New York, 1975.

Collins, A., E. H. Warnock, and J. J. Passafiume, "Analysis and Synthesis of Tutorial Dialogues", Report 2789, Bolt, Beranek and Newman, Inc., March 1974.

Collins, A. M., J. J. Passafiume, L. Gould, and J. G. Carbonell ***Improving Interaction Capabilities in Computer-Assisted Instruction***, Bolt, Beranek and Newman, Inc., BBN Report 2631, August 31, 1973.

Elliott, Roger W., "A Model for a Fact Retrieval System", The University of Texas Computation Center, TNN-42, May 1965.

Fikes, Richard E. and Nils J. Nilsson, "STRIPS: A New Approach to the Application of Theorem Proving to Problem Solving," *Artificial Intelligence*, II, 1971, pp. 189-208.

Friend, Ken, "An Information Processing Approach to Small Group Interaction in a Coalition Formation Game," Psychology Department, Carnegie-Mellon University, 1973.

Grice, H. Paul, "Logic and Conversation", *Syntax and Semantics*, Vol. 3, Academic Press, New York, 1975, pp. 41-58.

Grimes, Joseph E., "The Thread of Discourse", Cornell University, 1972, Ithaca, N.Y., Department of Modern Languages and Linguistics,

Grignetti, M. C. and E. H. Warnock, "Mixed-Initiative Information System for Computer-Aided Training and Decision Making", Bolt, Beranek and Newman, Inc., Report 1971, September 15, 1973.

Hendrix, Gary G., Craig W. Thompson, and Jonathan Slocum, "Language Processing Via Canonical Verbs and Semantic Models," The Department of Computer Sciences and CAI Laboratory, The University of Texas, Austin, Texas. Appearing in *Advance Papers of the Third International Joint Conference on Artificial Intelligence*, August 20-23, 1973, Stanford University, Stanford, California, pp. 262-269.

Hendrix, Gary G., "Question Answering Via Canonical Verbs and Semantic Models: A Model of Textual Meaning," University of Texas at Austin, Computer Science Technical Report NL12, January 1973.

Hewitt, C., P. Bishop, and R. Steiger, "A Universal Modular ACTOR Formalism for Artificial Intelligence," *Advance Papers of the Third International Joint Conference on Artificial Intelligence*, August 20-23, 1973, Stanford University, Stanford, California, pp. 235-245.

Kaplan, Ronald M., "A General Syntactic Processor," *Courant Computer Science Symposium 8: Natural Language Processing*, New York, December 20-21, Algorithmics Press, Inc., 1971, pp. 193-242.

Kay, Martin, "The MIND System," *Courant Computer Science Symposium 8: Natural Language Processing*, December 20-21, 1971, Algorithmics Press, Inc., New York, pp. 155-188.

Lindsay, P. H., and D. A. Norman, *Human Information Processing, An Introduction to Psychology*, Academic Press, 1972.

Malhotra, A., "Design Criteria for a Knowledge-Based English Language System for Management," Ph.D. Thesis, Massachusetts Institute of Technology, February 1975.

Mann, William C., "A Memory-Process Model of Symbolic Assimilation," Ph.D. Thesis, Department of Computer Science, Carnegie-Mellon University, Pittsburgh, Pennsylvania, April 1974.

Mann, William C., James A. Moore, James A. Levin, James H. Carlisle, ***Observation Methods for Human Dialogue***, Information Sciences Institute, ISI/RR-75-33, Marina del Rey, California, June 1975.

Moore, James and A. Newell, "How Can Merlin Understand?", Department of Computer Science, Carnegie-Mellon University, November 15, 1973.

Moore, James, "The Design and Evaluation of a Knowledge Net for Merlin," Ph.D. Thesis, Department of Computer Science, Carnegie-Mellon University, 1971.

Newell, Allen, and Herbert A. Simon, ***Human Problem Solving***, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1972.

Norman, D. A., and D. E. Rumelhart, ***Explorations in Cognition***, W. H. Freeman and Company, San Francisco, 1975.

Partee, Barbara H., "Linguistic Metatheory," in William Orr Dingwall, ed., ***A Survey of Linguistic Science***, Linguistics Program, University of Maryland, September 1971, pp. 674-675.

Quillian, M. R., "Semantic Memory," in M. L. Minsky, (ed.), ***Semantic Information Processing***, M.I.T. Press, Cambridge, Mass., 1968.

Quillian, M. R., "The Teachable Language Comprehender: a Simulation Program and Theory of Language," ***Communications of the ACM***, 12, pp. 459-476, 1969.

Raphael, B., "SIR: A Computer Program for Semantic Information Retrieval," in M. Minsky, ed., ***Semantic Information Processing***, M.I.T. Press, Cambridge, Mass , 1968.

Rulifson, J. F., ***QA4: A Procedural Calculus for Intuitive Reasoning***, Ph.D.EThesis, Computer Science Department, Stanford University, November 1972.

Rumelhart, D. E., P. H. Lindsay, and D. A. Norman, "A Process Model for Long-term Memory," in E. Tulving and W. Donaldson, eds., ***Organization of Memory***, Academic Press, 1972.

Sager, Naomi, "Sublanguage Grammers [sic] in Science Information Processing," ***Journal of the American Society for Information Science***, pp. 10-16, January-February 1975.

Sager, Naomi, "Syntactic Formatting of Science Information," ***Proceedings of Fall Joint Computer Conference***, AFIPS Press, Montvale, N.J., 1972, pp. 791-800.

Schank, Roger C., "Identification of Conceptualizations Underlying Natural Language," *Computer Models of Thought and Language*, W. H. Freeman and Company, San Francisco, 1973, pp. 187-248.

Schank, Roger, "The Conceptual Analysis of Natural Language," *Courant Computer Science Symposium 8: Natural Language Processing*, New York, Algorithmics Press, Inc., December 20-21, 1971, pp. 291-310.

Searle, John R., "Indirect Speech Acts", *Syntax and Semantics*, Vol. 3, Academic Press, New York, pp. 59-82, 1975.

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

Simmons, R. F., "Semantic Networks: Their Computation and Use for Understanding English Sentences," *Computer Models of Thought and Language*, W. H. Freeman and Company, San Francisco, 1973, pp. 63-113.

Simon, Herbert A., "The Functional Equivalence of Problem Solving Skills," *Cognitive Psychology*, Vol. 7, No. 2, April 1975, pp. 268-288,

Slocum, J., "Question Answering Via Canonical Verbs and Semantic Models: Generating English from the Model," University of Texas at Austin, Computer Science Technical Report NLl3, January 1973.

Srinivasan, C. V., "The Architecture of Coherent Information System: A General Problem Solving System," Department of Computer Science, Rutgers University, New Brunswick, New Jersey. Also, appearing in *Advance Papers of The Third International Joint Conference on Artificial Intelligence*, August 20-23, 1973, Stanford University, Stanford, California.

Waterman, D. A., and A. Newell, "PAS-II: An Interactive Task-Free Version of An Automatic Protocol Analysis System", Department of Computer Science, Carnegie-Mellon University, June, 1973.

Waterman, D. A., and A. Newell, "Protocol Analysis as a Task for Artificial Intelligence", Department of Computer Science, Carnegie-Mellon University, May 24, 1971.

Wilks, Yorick, "The Stanford Machine Translation Project," *Courant Computer Science Symposium 8: Natural Language Processing*, New York, December 20-21, 1972, Algorithmics Press, Inc., 1971, pp. 243-290.

Winograd, Terry, "A Procedural Model of Language Understanding," *Computer Models of Thought and Language*, W. H. Freeman and Company, San Francisco, California, 1973, pp. 152-186.

Wittgenstein, Ludwig, *Philosophical Investigations*, 3rd ed., Macmillan Company, 1958.

Woods, W. A., "An Experimental Parsing System for Transition Network Grammars," *Courant Computer Science Symposium 8: Natural Language Processing*, December 20-21, 1971, Algorithmics Press, Inc., New York, pp. 111-154.