

Analysis of Human Communication During Assembly Tasks

K. Suzanne Barber and Gerald J. Agin

CMU-RI-TR-86-13

**The Robotics Institute
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213**

June 1986

Copyright © 1986 Carnegie-Mellon University

Table of Contents

1. INTRODUCTION	1
2. BACKGROUND FOR ANALYSIS	3
3. DESCRIPTION OF EXPERIMENTS	5
3.1 IBM Proprinter Assembly	5
3.2 Insertion Experiments	7
4. SUMMARY OF FINDINGS	9
5. FOCUSING	11
5.1 Goal Specification	14
5.2 Investigation	15
5.3 Grouping	15
6. COMMUNICATION VALIDATORS	18
6.1 Clarifications	18
6.2 Verifications	19
7. REFERENCES	22
8. DESCRIPTORS	25
8.1 Characteristic Identifiers	25
8.2 Analogies	28
8.3 Functional and Possibility Descriptions	28
8.4 Result Descriptors	28
8.5 Perceptual Descriptors	29
8.6 Positional Descriptors	30
8.7 Orientation Descriptions	31
8.8 Insertion Descriptions	32
8.9 Negation	33
9. DIALOG STRUCTURE	34
9.1 Transition of Dialog from Task to Task	34
9.2 Transition of Dialog Within a Task	34
9.3 Transition of Dialog Within a Subtask	36
9.3.1 Parts Identification	36
9.3.2 Orientation Specification	36
9.3.3 Insertion Descriptions	37
10. IMPLICATIONS FOR A ROBOTIC SYSTEM	38
11. CONCLUSIONS	41

List of Figures

Figure 3-1: IBM Proprinter	5
Figure 3-2: Hardware for Insertion Experiments	8
Figure 9-1: Transitions of Dialog	35
Figure 9-2: Patterns of Orientation Specification	37
Figure 10-1: Overall Organization of Assembly	40

Abstract

This paper studies human-to-human interaction in an attempt to shed some light on the kinds of human-to-machine interaction that will be necessary for intelligent robot learning of assembly tasks. Experiments were performed in which an "expert" guided an "apprentice" through a complex assembly task using spoken language but no visual communication. An analysis of the dialog reveals that certain protocols and conventions facilitate communication, and that communication breaks down when these protocols are not observed. Five types of protocols were observed: focusing, validators, referencing, descriptors and dialog structure. The implications of these results for human-robot communication are discussed.

1. INTRODUCTION

Robots are potentially useful devices for automatic assembly because of their versatility and programmability. A single robot can accomplish a variety of tasks if it is programmed for each task. Currently, the only way to adapt a robot to a new task is by conventional programming techniques. It would be desirable to give robots the capability of learning by other means, similar to human operatives on an assembly line learning a new task. The ideal scenario would involve an intelligent system, equipped with vision and speech input and output that could learn by example. Intermediate scenarios would involve various forms of mixed-initiative interaction between the human "teacher" and the robot.

In order to move toward this goal, we are studying the semantics of human-to-human interaction in teaching and learning assembly tasks. The understanding we gain should be useful in building the first generation of machines that can learn by teaching rather than by programming.

This paper describes and analyzes several dialogs collected when one person communicated an assembly task to another. The information that is transferred between teacher and student corresponds to the information that is necessary for the robotic assembly system to have. Information that is not passed from teacher to student, but nonetheless is used in the assembly process either must be shared prior knowledge or must be deducible from the evidence at hand. It will be necessary either to give the robotic system the same deductive powers or knowledge base, or to provide additional instruction at teaching time. A better understanding of human to human communication will aid effective communication between human and robot.

The first set of experiments dealt with the entire assembly process. We examined dialog collected while one person, the expert(E), taught another person, the apprentice(A), to assemble an IBM Proprinter. The expert was taught the assembly task prior to the experiment. The apprentice had no prior knowledge of the task and was presented with a random arrangement of the necessary parts at the beginning of the experiment. The subjects were allowed to freely interact verbally but were separated by a partition so that no visual information was shared. If face-to-face teaching were allowed, means of communication would be used that are difficult to record and analyze.

The dialogs suggest humans employ a collection of strategies to identify, orient and insert each part. The dialog analysis provides us with the type of descriptors and instructions utilized in assembly. From the dialog structure, a means of organizing the techniques and strategies, descriptions and instructions has been explored.

A second set of experiments were conducted to investigate tactile information during an insertion task. This important information was, for the most part, not communicated verbally in the Proprinter assembly. Tactile information is essentially shared information or can be easily deduced from the environment so that humans utilize these perceptions without being instructed to do so. In these experiments, the subject was asked to perform a series of different insertions while describing how it felt. Information such as correct or incorrect insertion, the perception of different phases in the insertion, and the determination of completion were elicited.

Other researchers who have used similar techniques to study dialog are Cohen [1984] in a study comparing spoken and keyboard communication, Akin and Reddy [1977] in their research of image understanding utilizing the picture-puzzle paradigm, and Grosz [1982] in an analysis of different types of dialogs to characterize the language used when people communicate to solve a problem in one communication mode under different environmental restrictions. Our work differs from theirs in a number of respects. For our experiments, the only mode of communication was a completely interactive communication restricted only with respect to shared visual information. The dialogs of interest were those concerned with assembly tasks. In the picture-puzzle paradigm described in Akin and Reddy's work, the subject asked questions about a scene and the experimenter answered these questions using the actual photograph the subject was attempting to identify. The dialog generated in our experiments was a result of teaching by an expert and interaction by the apprentice as the learning process necessitated. Our interest in teaching assembly process directed our analysis to some specific questions. A sample of these questions are, "What is the assembly process as seen through the dialog structure?", "What teaching techniques were utilized to generate a good description or instruction?" and "What type of descriptors are used to identify parts, to describe insertions, and to relate tactile information and spatial orientations?"

2. BACKGROUND FOR ANALYSIS

A significant amount of research has investigated language with respect to a system's ability to engage a user in discourse. A large body of work has been concerned with the difference in spoken and written communication. Previous investigations have concentrated on syntactic differences (Hindle 1983; Kroch and Hindle 1982; Thompson 1980) with the goal of adapting parsing techniques to handle the syntax of spoken language. In the comparison of spoken and written discourse, written language is syntactically more integrated than spoken, employing nominalizations, participles, complements and relative clauses (Chafe 1982) while spoken language exhibits regular patterns of false starts and hesitations. Chapanis (1975) employed statistical measures such as the number of sentences, number of words and time required for problem solution to characterize the differences in language across different communication modalities. Cohen (1984) asserts that research is needed to compare the discourse structure of spoken and keyboard language to determine whether and how current techniques need to be adapted to the way people speak. Cohen maintains that the communication situation helps to determine the pragmatics of reference—what speakers intend hearers to do with referring expressions. Current theories propose that speaker's intentions underlying the use of indirect speech acts can be recognized as a by-product of a more general, independently motivated process of inferring a speaker's plans (Bruce 1983; Cohen and Perrault 1979; Cohen and Levesque 1980; Perrault and Allen 1980). Although these approaches provide information as to the superiority of one mode of communication over another and the effect on language a particular mode may exhibit, they do not address information such as the structure of the dialog and the actual components of the language which provide communication between humans.

A system also ought to be able to reason about the speaker's uses of descriptions—for identification, correcting previous misidentifications, attribution, etc. Akin and Reddy [1976] present an experimental paradigm for knowledge acquisition and illustrate different types of knowledge that seem to be useful in image understanding research. Three major aspects of knowledge are presented: primitive feature extraction operators, rewriting rules, and flow of control. A limited number of feature extraction operators were repeatedly used by the subjects to specify location, size, shape, quantity, color, texture, and patterns or various components found in scenes. Six types of rewriting rules were identified; assertions, negative assertion, context-free, conditional, generative and analytical inferences. Flow of control exhibited characteristics of a hypothesize and test paradigm capable of using imprecise, conflicting hypotheses in cooperation with others in multidimensional problem space.

Grosz' work [1982] is closer to the goals of our analysis. She analyzed dialog to determine a

person's language needs when using a computer system. Two types of dialogs were studied: task-oriented and question-answering dialogs. In each of the dialogs along with respective environmental restrictions, Grosz studied discourse-level phenomena: those features of utterances in dialogs that come from the utterances being part of a cohesive unit of discourse, the relation between dialog and task, structure of the dialog and the influence of an utterance on the utterance that follows.

The emphasis of the analysis presented in this paper addresses (1) the structure of dialog in relation to the process of assembly, which we found exemplified a relationship between dialog and task, (2) the teaching strategies and techniques for describing the assembly process, and (3) what those actual descriptors were. Via the study of human dialog in assembly task, an understanding of the dialog necessary in human-robot dialog for assembly tasks is the goal of the analysis for the experiments we investigated.

3. DESCRIPTION OF EXPERIMENTS

There were 2 different types of experiments conducted for the purposes of this analysis. The two types will be discussed separately.

3.1 IBM Proprinter Assembly

In this experiment, one subject, the expert(E), taught another subject, the apprentice(A), how to assemble the IBM proprinter pictured below.

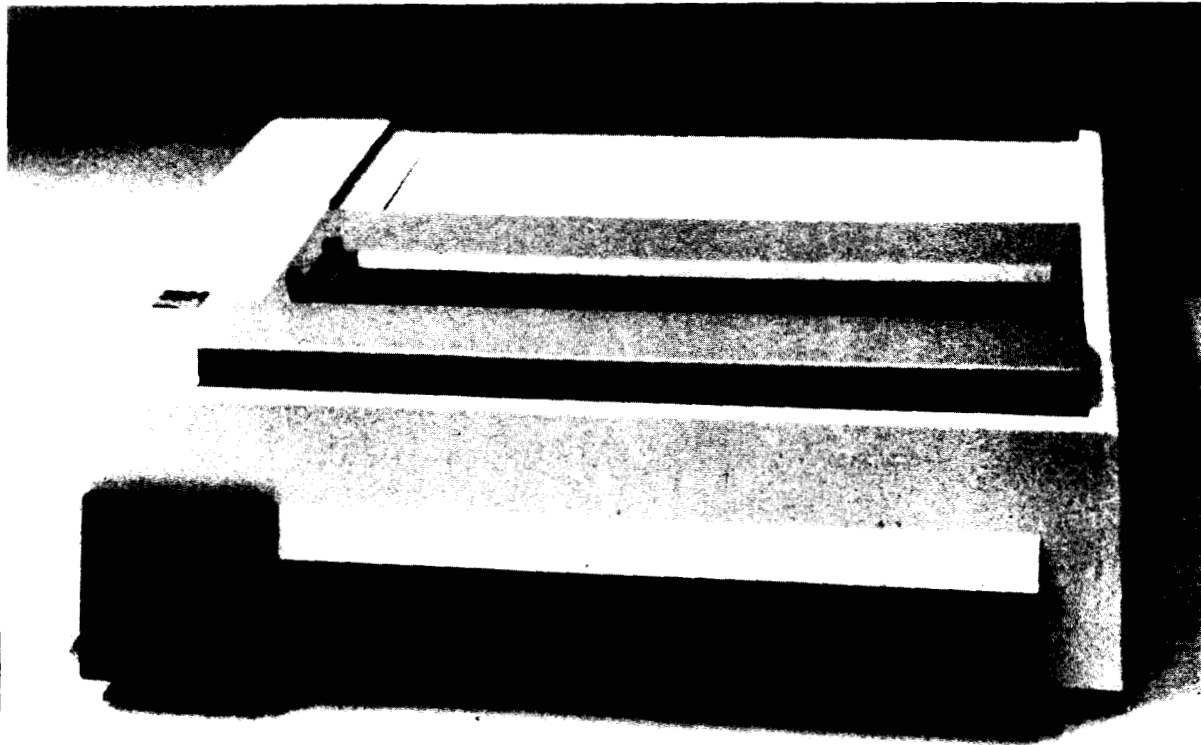


Figure 3-1: IBM Proprinter

This assembly involves 31 tasks. Each task includes identification, orientation and insertion of a part, except for the first part in the assembly which needed only to be identified and oriented. The entire assembly process can be accomplished without the aid of tools or fasteners. There are different types of parts which require a variety of insertion techniques such as peg-in-a-hole, twist to lock, slide in's and snap in's. Many of the rods inserted in the assembly were inserted into holes on the right panel and snapped in on the left panel. This type of insertion requires motions along several different axes during the insertion process. Present robot arms are incapable of accomplishing these

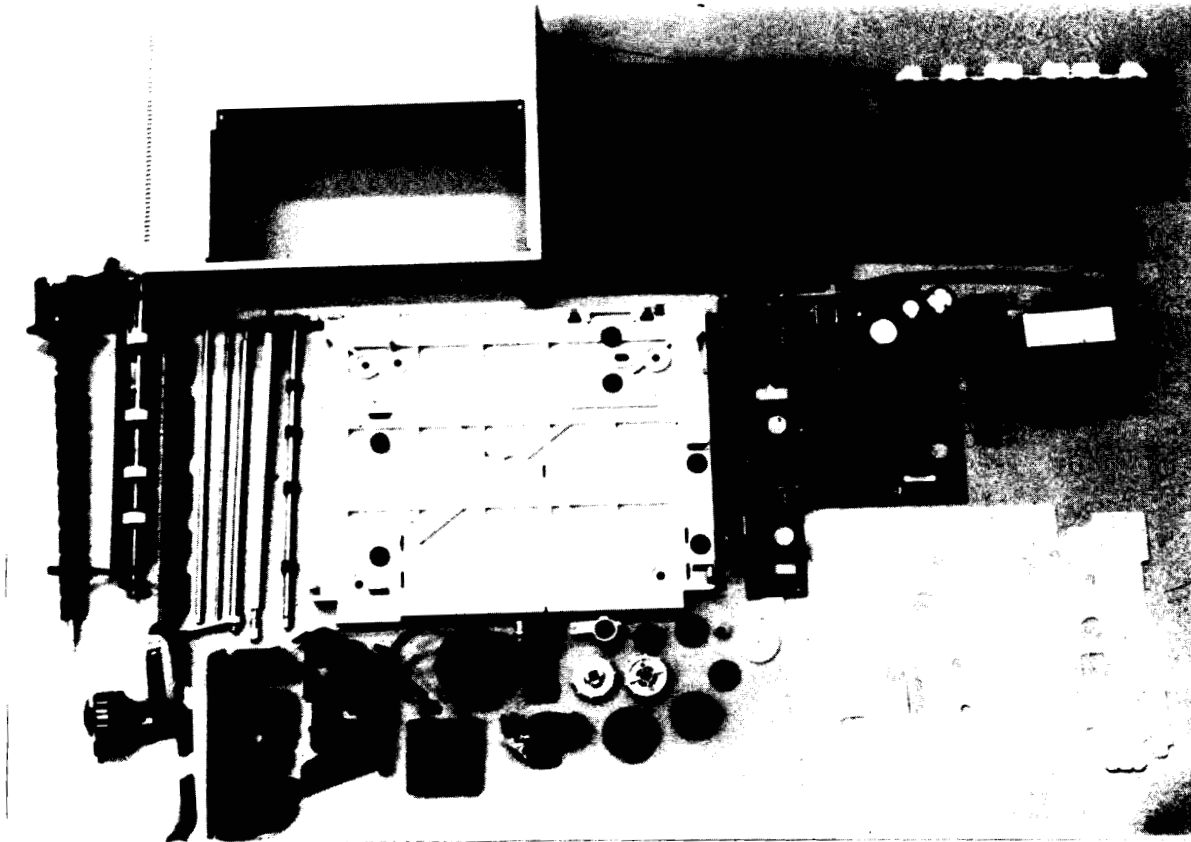


Figure 3-2: Hardware for Insertion Experiments

free associate how performing the insertion felt and looked, how the completed insertion felt and looked and whatever else came to mind.

The first experiment was performed by a Ph.D. student in English. Participants in the second and third experiments were electrical engineering undergraduates. All three subjects focused on issues such as the compliance of the object being inserted throughout different stages in the insertion, indications via the compliance as to necessary adjustments for correct insertion, or the compliance of the object inserted incorrectly.

types of tasks. Each insertion is unique and has only one possible resulting position. There are two gears which could possibly be inserted at an incorrect orientation but subsequent insertions would then be impossible.

Prior to the experiment, the expert was taught the assembly. The apprentice had no prior knowledge of the task. The subjects were separated by a partition in the same room so that no visual information was shared during the assembly. Verbal communication was not restricted; therefore, the subjects were allowed to interact freely. The entire experiment was videotaped for analysis at a later date.

The expert was given essentially no instructions except those concerning how to assemble the printer. Special care was used in these pre-experiment instructions. The experimenter who taught the expert was careful not to label or give names to parts, orientations or types of insertions. The expert was instructed to simply watch how to assemble the printer and was then asked to perform the task. This approach was used to insure the teaching performed by the expert exemplified the individual's own interpretation of the assembly process. The apprentice was presented with the parts of the printer arranged in a random fashion. The principal instructions and information given to the apprentice were that the parts in view were all needed in the assembly of the printer, and the assembly would be taught by the expert. The apprentice was also informed that questions could be asked at any time; thus, the line of communication should always be seen as two way.

Five dialogs were collected. Dialog #1 involved two electrical engineering graduate students. Because the apprentice had a very good mechanical sense, he was able to complete the assembly with very little guidance and very little useful dialog was generated. To correct this deficiency, the apprentice carefully chosen for dialog #2 was to be someone who would rely heavily upon the expert to perform the assembly. This indeed was the case. The expert in this experiment was a computer scientist and the apprentice was a graduate student in public policy. Some problems in this experiment occurred due to the expert's poor teaching capabilities and some inability to perceive difficulties encountered by the apprentice. Nonetheless, some important insights were obtained from this experiment. The third experiment included an undergraduate in architecture and a graduate student in history. There was a significant improvement in the obtainable information in this experiment compared to the previous two due to improved teaching capabilities on the part of the expert. Yet, it was apparent at this point in the experimentation that trial and error on the part of the apprentice was a problem. At the times when trial and error occurred, primarily in insertions, the apprentice was not reliant upon the expert for information. Thus, this information was not present in the dialog and our ability to identify the information used in these incidents was very restricted. To

resolve this problem in dialog # 4, the apprentice was told that a subtask such as a part identification, orientation or insertion was not to be attempted until instructed to do so by the expert. The expert was an electrical engineering graduate student and the apprentice was a public policy graduate student. The teaching methods and information used by the expert were very explicit in the dialog and the apprentice relied almost totally upon the the expert's teaching. The fifth dialog was another successful endeavor involving the apprentice from experiment #3 as the expert and another graduate student in public policy as the apprentice. Once the apprentice from experiment #3 had performed the experiment, he felt he had become an expert. This fact combined with obvious teaching capabilities made him a suitable choice. The same restrictions on trial and error applied in this experiment.

It is important to recognize that failure to communicate does not imply failure of the experiment. Analysis of the contrasts between successful and unsuccessful communication provides useful data. We must not only learn how to teach an assembly task, we must understand how not to teach in ways which had been predicted as unsuccessful.

3.2 Insertion Experiments

These experiments involved only one person at a time who performed a variety of unrelated "insertion" tasks. A faucet assembly and a toy assembly were included. Selected insertions from the IBM proprinter also provided some interesting insertions to investigate more closely. The assemblies are pictured below.

The insertion tasks included several types:

1. screw on's—caps screwed onto threaded pegs, threaded cylinders into holes
2. peg-in-a-hole—hexagonal, round and square pegs
3. snap in's—disk into holes, rod into semicircular destinations
4. slide on's—rods into crevices, connector inserted on prongs.

With respect to each of these sample insertion types, each varied greatly in terms of compliance at all stages of the insertion, in terms of initial approach to insertion and determination of completion.

The subject was shown how to perform the task before the actual experiment and was then allowed to "play around," investigating the parts and insertions. During the actual experiment, the subject was instructed to describe the insertions into a tape recorder. The subject was encouraged to simply

4. SUMMARY OF FINDINGS

As a result of the protocol analysis of both the proprinter assembly and the insertion dialogs, it is clear that communication between humans during assembly tasks is facilitated when the following protocols are employed.

Focusing is a preparatory description or instruction intended to provide an introduction to upcoming dialog. At the beginning of a subtask, goal specification is most frequently used as a focusing device. Within a subtask, attention to such points of interest as a part, group of parts, location, type of insertion or orientation is promoted by investigations, grouping or incremental focusing.

Communication validators work to insure that information is mutually understood between the apprentice and the expert. Clarifications and verifications are two forms of communication validators observed in the dialogs.

Referencing is the process of recalling common knowledge or previously learned information. Referencing reduces a need to reiterate descriptions and instructions.

Ten types of descriptors communicated the characteristics of parts, locations, insertions, and orientations:

1. Characteristic identifiers such as names, shape, weight, size, color, material and texture described parts and locations.
2. Positional descriptors such as back, front, corners, edges, the table, and the ceiling were extensively utilized.
3. Orientation descriptions specified the position of a section of the part, the position of the entire part or the resulting position. Descriptors such as horizontal or vertical were common.
4. Insertion descriptions included specifications of the location on the part to be inserted, the destination of insertion, the orientation for insertion and the technique of insertion.
5. Analogies provided descriptions by comparison to a common objects, common functions, common textures or knowledge previously learned.
6. Functional descriptions describe the function of a part or location.
7. Possibility descriptions describe the part to be inserted and allow the insertion destination be deduced from the possibilities.

8. Result descriptors provide an instruction and a result from which actual implementation is to be deduced.
9. Perceptual descriptors elicit the use of tactile and visual abilities.
10. Negation describes the characteristics not possessed by the part, location, insertion or orientation.

In the experiment of the Proprinter assembly studied, we found that in all of the dialogs a simple task hierarchy was apparent. This hierarchy was evidenced in the structure of the dialog. We label this hierarchy as follows.

ASSEMBLY GOAL

Task

subtask

The assembly goal is to complete the assembly of the IBM Proprinter. In order to accomplish this goal, tasks were performed for each part that were composed of subtasks. Three subtasks were consistently utilized to achieve a task—locating, orienting and the inserting. Distinct subdialogs were found in the analysis relating to these subtasks. By examining the structure of the dialog, an organization to teaching assembly tasks is provided.

In the discussions to follow, excerpts from the actual dialogs studied will provide examples.

5. FOCUSING

Focusing is a preparatory description or instruction intended to provide an introduction to the following dialog. It was used by the expert to direct the apprentice's attention to upcoming descriptions or instructions. Additional research in the representation and the use of focus in dialog understanding has been done by Grosz [1977]. Four subjects of focusing devices were prevalent in the experiments we studied—goal or intention, part specification, orientation specification, and locational definition. Goal or intention focusing specified the purpose of upcoming instructions or descriptions. Part specification established a general type of part to be investigated such as asking the apprentice to focus on all the gears in the assembly. Orientation specification directed attention to an orientation for insertion of the part in focus or an orientation so that additional parts could be inserted on the part in focus. Locational definition specified particular locations either on the parts or areas of insertion destinations. What is focused on, when focus of attention is established, and why focus shifts is primarily a function of the task hierarchy.

By focusing the attention of the apprentice, the expert accomplishes 3 goals. First, the question, "what happens next?" is answered. From the dialog studied, it appears possible that the apprentice uses the information provided to establish expectations of what perceptual tools such as visual or tactile will be required. Given the information of "what happens next," the apprentice expects a certain category of questions will be answered. For example, if the expert indicates that the next task will be an insertion, questions such as "which part will be inserted where and at what orientation?" are raised. It is now the expert's job to satisfy these questions.

E: Ok, pick up the larger of those 2 pieces we just identified. ok, and that will go on the left hand side

A: Where does it go?

How should it sit?

Where should the bottom go?

Second, focusing devices extract points of interest from the environment. Thus, the expert facilitates a means for concentrating attention on particulars such as an approximate physical location or group of parts. We will call this specified environment a *defined teaching frame*. A good example of defined teaching frame used in all five experiments involved defining at the beginning of the assembly the local destination of all future insertions as seen in the following dialog.

E: First, you have to look for is a large flat whitish, cremish color piece. It is probably the biggest piece you have on the table

And, it's got some small black circles all over it

A: Ok, I found it

E: So what we have to do now is orient it so the rest of the pieces in the experiment can be mounted on top of it, ok?

Therefore, unless otherwise indicated, the general area of focus for the destinations of insertions was the part identified above. These defined teaching frames were primarily specifications of approximate physical areas the expert wanted A to focus on visually. Subsequent instructions and descriptions were often relative within a defined teaching frame. Relatives such as big, small, near, front, etc. were common.

And last, focusing provides organization to the dialog. It should be recognized that task hierarchy does impose the principal structure to the dialog. Parts must be assembled in a particular order and, of course, a part must be located before it can be inserted. But, focusing labels this structure so that the transition of focus is explicitly provided by the expert. Thus, references and descriptions are simplified within a specified arena of focused attention. This is represented in the following dialog taken from one of the experiments.

E: Ok, the next piece to look for is large and rather tall...

(continues describing the part)

E: Ok, now you have the right one. We will have to orient it so ... do you see the red switch on it?

(the orientation is described with reference to the red switch)

E: Ok, now we've got it right. Now then if you will look back at the flat plastic that we are using for our base, I will describe where it goes.

(destination is described)

E: Now, pick it up and we are going to insert it by sliding it in correctly.

(orientation and technique for insertion are described)

E: Ok, push down to make it click in place.

A: It's in.

E: We are on to the next piece.

Explicit focusing strategies occurred frequently in the dialog studied. Yet, there were points in the assembly when a shift in attention was not explicitly directed, but was obvious to both the expert and

the apprentice. For example, after one part has been identified, oriented, and inserted, the task hierarchy indicates a new part is now to be identified. Indicators of focus shifts preempted by the task hierarchy were very simple such as the expert saying "next" or "now." As indicated before, the obvious shift occurred after an insertion and before a new part identification. Therefore, in these cases, the task hierarchy provides the primary means of focus.

The term "levels of focusing" explains that focus may shift to subtasks without forgetting previously focused information. For example, in most cases, when focus is established for a particular part, all other parts are unimportant and are really of no concern to the apprentice. Yet, when focus shifts to a location for an insertion destination from part identification, the part previously identified is not forgotten because it is that part which will be subsequently inserted. The previous sample of dialog also exhibits the different levels of focusing which occurs in assembly dialog. In the previous example, the part itself remains in the higher level of focus, the reference to the base prompts the apprentice that the description of the part's destination is forthcoming and the immediate focus shifts. The focus shifts back to the higher level when the destination is identified, verified by A and consequently, begins discussions of the part again. This higher level of focus is closed with the verification of the insertion by A and the signal by E, "next piece" to indicate a new task.

Problems frequently occur especially in the less successful experiments when shifts in focus were not evident to the apprentice. In the following example, the expert fails to shift the focus of intent as she moves from describing orientation to specifying the destination and technique for insertion.

E: Ok, position the typing

A: yes?

E: Ok, you want to fit into that... in that corner. See how it kinda fits in that corner there.

A: Should the typing be facing out?

E: Oh, yes, the typing should be facing out.

The fallacy of shifting intention without notice is evident here. The apprentice still wanted to talk about the typing even after the focus shift. The expert is forced to return to a discussion of the typing. The "Communication validator" (section) will describe techniques for providing effective closure to focus.

The following discussions will address the types of devices which established focus for the four subjects mentioned previously: goals or intentions, part identification, orientation, insertion.

5.1 Goal Specification

The most evident transition of focus is a shift from subtask to subtask frequently exemplified by a goal specification. The format of goal specifications varied somewhat but for all the experiments there were some very detectable patterns. For location and orientation subtasks, the goal specification was usually just a phrase such as "now, you have to orient it" or "the first thing to look for is probably the largest piece."

Since the part was already identified when orienting became an issue, a phrase was sufficient and occurred a significant percentage of the time. However, goal specification for the location of parts or destination could be more elaborate by including not only the instruction "to look," but "where to look" or the intended use of the part or destination. This additional information in the goal specification will be referred to as possible parameters. Possible parameters such as the area to visually focus upon, the green piece, and the function, holding the "whole piece," is given in the succeeding dialog.

E: Ok, I am going to describe another place now on the green board again where we are going to put this whole piece.

It should be noted again that phrases like "whole piece" can be used here because the part is at a higher level of focus; thus, it is clear to which piece reference is being made.

A brief overview was the most common form of goal specification for insertions. The following example exhibits 3 of the 4 most predominant contents of an insertion overview—part to be inserted, type of insertion, description of destination, the resulting position.

E: Ok, so now you are going to pick up the gray piece and fit it onto the board so that tab fits into the slot.

Goal specification serves basically two purposes :

- an introduction into new subtasks or a new focus within a higher level of focus,
- an explanation for the apprentice of the intent and purpose of the following instructions.

In the end, a focus is established if the goal specification is successful. Success is contingent upon recognizing the shift in focus and establishing the correct focus. Success is measured by the apprentice's receptiveness to the succeeding instructions and descriptions.

Additional focusing devices: investigation, grouping, stepwise identification, and gradual focus, were exhibited in these experiments primarily within a subtask instead of introducing a type of subtask. These devices were also under a higher level of focus.

5.2 Investigation

Investigations can be defined as an imperative or request by the expert to investigate and as a result find a characteristic identifier which defines an orientation of a part or location on a part already in focus. Investigation was a strategy used by the expert to focus the apprentice's attention by noting points of interests in the surrounding area or on the part already in focus. The increased focus due to investigation preempts and facilitates more specific descriptions and instructions.

Investigation is primarily used in the search of part characteristics for identification subtasks. Essentially, what the expert says during an investigation is "Let's go look at this part and investigate some characteristics about the part."

E: If you were looking at it .. There is one side of it that the metal ribbon is coming out of it? There's the piece where the metal ribbon is sorta lying on it? If you look on the other side of it you will see that there is a slot at the bottom that is a very skinny rectangle.

A: Yes.

Investigations occurring during insertion tasks were primarily for the location of destinations for insertions. The expert instructed the apprentice to investigate an area already mentioned at a higher level of focus. Thus, one will notice the extensive use of pronouns in investigation due to referrals of locations or parts mentioned previously. This increased focus assures the expert the correct location is in focus and specific, solo descriptions are more easily understood by the apprentice. The following example of locational investigation includes the investigation followed by the intended specific location description.

E: Alright, if you look on either side of it... I don't know how you are holding it but there is going to be 2 small circles about 3 inches apart which look like they might be able to fit over those spikes we were taking about before.

The orientation is determined in reference to the viewpoint of the apprentice. Initializing the focus for orientations usually involved establishing the point of view of the apprentice. Just as the approach angle for a robot is established before grasping an object, the approach angle of A needs to be established before investigation.

5.3 Grouping

Grouping of similiar parts, orientations and destinations was a strategy used by all experts. With respect to 4 gears which were parts in the printer assembly, every expert used grouping as a primary identifying technique. To define a group, a union of parameters must be defined. In the case of the

gears, the expert simply had to ask the apprentice to gather all four gears among the parts available. Explicitly, the union was the name "gears." Implicitly, the apprentice was asked to locate all four disks of various diameters with jagged edges and a hole of unknown inner diameter. But, it is not only common characteristics in parts that can form a group, locations formed unions for insertion destination groups in these experiments also.

Once a group is defined based upon these common denominators, descriptions can take place without regard to other parts on the assembly. These descriptions occur to distinguish the parts within the group in two ways. First, the parts can be contrasted. This involves focusing on a common characteristic such as the diameter of the gears and contrasting the size. Contrasting frequently uses relatives such as the larger, smaller or medium gear diameter rather than absolute dimensions. The use of relativity simplifies necessary descriptions within the focus of a group.

Extraction of a solo identifying characteristic is a second means of distinguishing parts in a group. Extraction was also used in reference to the gears. As seen below, the diameters are given more absolute dimensions and distinctive characteristics are mentioned.

E: So you see some plastic gears... around you see actually several! ...huh... one is about the size of a quarter and it is perfectly flat it doesn't have anything.. it is just a basic gear... There are actually 2 about the size of a quarter one has some extra things going on.

It is important to note that the common characteristics apply to all the parts in the assembly while the solo identifying characteristics and relative descriptions are referred to once the group is established and "make sense" only within focus of group.

Possibility grouping is essentially a grouping of all the possible ways a subtask could be performed. Grouping of possibilities frequently occurred with orientations because it is sometimes easier to suggest a member of plausible orientations and then negate the incorrect orientations.

E: The biggest piece of plastic that is sticking off this thing is pointing toward the right.

A: Right.

E: now, there are two ways it could be sticking out. One so that it is on the top half of the circle and one so that it is at the bottom half of the circle.

Incremental focusing works to establish a group of one. Incremental focusing strategies are usually in the form of utterances which exhibit a top-down approach. The following example initially provides a general identification and builds on that description by adding more specific characteristics.

And you have some gears

some plastic

some black plastic gear

As the description becomes more extensive the focus is intensified.

6. COMMUNICATION VALIDATORS

Communication validators attempt to insure that what is intended to be taught and what is learned are the same. Two communication validation strategies observed in the dialogs were clarification and verification. Clarifications address vagueness in teaching with spot checks or "add-on" information. Clarifications also address confusion in learning with questions concerning what is not understood. Verifications serve to establish that the intended goals of teaching are conveyed. Verification most frequently occurred at the conclusion of an instruction or description.

6.1 Clarifications

Clarifications are requests on the part of A to "clarify" the information provided by the expert. A clarification is usually initiated by A for two different reasons. The first reason is a result of the apprentice's frustrations with instructions or descriptions given by the expert. The apprentice responds by telling the expert what he or she perceives. This is a way of asking the expert to consider what A is aware of and utilize for that information in teaching.

The second reason originates from a need to insert additional information to either crystallize the instruction or description or to assist the expert's teaching. This additional information most often presents itself in a corrective clarification when A understands the teaching but wishes to refine the instruction or description.

E: They've got just little pieces sticking out of them in all sorts of shapes and sizes and you should find two of these. They also have 1 side that is sorta flat.

A: They're not exactly the same though.

E: Exactly the same, correct.

Simple "add-on" information is the third clarifier which is primarily provided by the apprentice. The principal purpose of "add-ons" is to provide more exactness to the description. The apprentice essentially performs a spot check.

E: You have 4 black gears.

A: Yes, they are different sizes.

E: Correct.

6.2 Verifications

Verifications, usually performed by the expert, establish that some action has been performed correctly. Verification is a very important element in the dialog facilitating closure to an instruction or description. Consequently, a "go ahead" signal is given which allows teaching to proceed. Without verifications, descriptions or instructions that are not understood by the apprentice are not addressed and further teaching is fruitless. The expert can not rely on the apprentice to always interject when uncertainties arise. When necessary verifications were overlooked in the experiments studied, the subsequent teaching was a waste and the expert was inevitably forced to return to the problem.

(The part to be inserted has been identified)

E: *Ok, now we've got it. This is going to go next to the capacitors on that green circuit board.*

(Orientation is described with no verification)

E: *You have to slide that into a piece of plastic which is sticking out the .. of the float white plastic piece that is on the table.*

E: *It look like a slot ... it looks like 2 metal pieces are going to fit into a slot.*

(The actual insertion with resulting position is describe with no verification)

A: *Can I ask you a question?*

E: *Sure.*

A: *Ok when I put it in or whatever... There's like a white... Is one piece of metal supposed to be... huh? you know on that white plastic board. This thing supposed to be on top of this little white thing?*

The apprentice had no idea where the actual point of insertion was. The entire section of dialog had to be repeated after the insertion located was finally verified.

The complexity of a verification's structure varies. Verifications can be as simple as the apprentice saying "Ok, I got it," or as complicated as the expert giving an overview of the entire subtask. Verifications are exemplified in basically five formats which are listed numerically below and accompanied by examples or discussions.

1) Inspections of the parts located, orientation established or position after an insertion

ex.1 Verification of parts located

(instructions and descriptions to locate 4 parts)

E: *Ok, you've got those 3 pieces and the metal rod*

ex.2 Orientation and position verification

E: The black gear is already on right?

The filler is going toward the gear side gear that is already on there

A: Then, the white piece?

E: yes, It is going to have a black piece, a white piece, a gray piece and then a white piece going from left to right

2) Inspections of how the part functions once in place after insertion

E: Ok, now it can rotate and the teeth will now stick through the hole

3) Reiterations of instructions—overviews

Overviews served as a very effective mode of completion verification by reiterating what had been accomplished and describing the end result. Though overviews by the apprentice were much less common than those performed by the expert, the effectiveness of reiterations by the apprentice was impressive. It provided the expert with a sense of the apprentice's interpretation of the subtask with respect to point of view, descriptive style, and what information provided by the expert was used by the apprentice. By accommodating the apprentice's perceptions, the expert's teaching was enhanced.

When a subtask is perceived as successful, the expert usually introduced an overview by the apprentice with something like "Describe to me." Repetition by the apprentice due to confusion is characterized by a formulation of what the expert has described and what is seen by the apprentice. It is often concluded by "is that what you mean?"

4) Tactile Verification simply verifies a subtask due to "how it feels." This verification was used primarily in insertion subtasks.

E: So it doesn't fit real well. So don't worry if it is a little loose.

5) Descriptive Verification consists of describing characteristic identifiers followed by verification request.

E: And you can tell they are the right ones because ...

the whole in the middle of them is square and will fit on the square rod.

Do you see those two?

A: Yes.

6) Bare affirmatives are the simplest of all the verifiers. Yet, this type of verifications can be dangerous. If there is not an explicit verification request proposed by the expert which indicate exactly what is being verified, A may verify something other than that intended. Or if the expert or the apprentice does not provide an overview especially of insertions, the risk of incorrect results increases. Verifications of simple "ok's" or "yes's" can assume different meanings. As noted by Grosz [1] and in this analysis as well, there are four different meanings to "O.K.".

- I heard you.
- I heard you and I understand.
- I heard you, and I am now doing what you said.
- I heard you and subsequently have completed the subtask. (Implied is "what's next?")

Differentiating meanings of these affirmatives is vital to the effectiveness of verifications.

7. REFERENCES

Referencing is the process of recalling knowledge common to both subjects or recalling previously learned information. Referencing makes use of an important link between what is presently in focus and information previously learned, or of common knowledge which may or may not be in focus. As evidenced in the experiments studied, there are six different types of references—pronouns, common knowledge, names, descriptions, order of assembly, and events.

A reference can be as simple as a pronoun that is referring to something at a higher level of focus. Pronoun referrals were most evident in orientation or insertion subtasks. When a part has been located, the part remains in focus as attempts are made to orient or insert the part. Pronoun references such as "it" are frequently used in these lower levels of focus since the higher level of focus eliminates the need to constantly re-identify or name the part.

Yet, there are times when the information to be recalled is not in focus. Thus, some explicit strategies must be employed to recover this information and bring it into focus as it pertains to the particular needs of the task. In these incidences, it is important to specify that a recall is being made and at times why the recall is necessary. This requirement is primarily based on a need to identify a focus shift so that it is clear to the expert and the apprentice to what the following discussions pertain. Request for recall can be performed by the expert or apprentice but most often is a teaching technique used by the expert. Prompts such as "do you remember" or "this insertion is done like the one we performed on the motor" were common types of recall requests by the expert.

Many common descriptions of a parts, orientations or insertions could be considered a request for recall. For example, references to color and shape such a gray, black, round or square are all common enough so that no explanation is necessary for the human subject. Thus, due knowledge shared by the subjects, these types of recalls did not require a focus shift due to such common usage.

Naming is referencing based entirely on the name. Recall based on naming assumes two forms. First, "assumed naming" occurs when the basic description of the part, location or orientation, or insertion is implicit within the name. These types of names include gears, switches, motors, holes, tabs, horizontal and flat. "Assumed naming" is a recall from a pool of common knowledge the expert assumes is shared by the apprentice. "Defined naming" recalls something which was previously identified and labeled with a name, in other words, "taught" or defined in the experiment.

Two characteristics important for successful "assumed naming" references were (1) the expert

must be certain the part, location, insertion or orientation can actually be assumed to be known by the apprentice and (2) the expert used those names as focusing devices and supplemented the assumed name with specifics.

"Assumed naming" is only a focusing device. The expert must realize that holes come in all shapes, motors assume many different forms, there may be more than one tab on a part, and so on. In many cases, using a name as an only means of reference may be inadequate. For example, the most common description of an extension was a "tab"; yet, it was necessary for the expert to specify location and shape to actually give relevance to the description.

E: Where you were looking before you saw those other 2 holes going down and all the way through the 2 very small ones... they are right in front of 2 things that I am calling tabs.

"Defined naming" proved extremely useful by facilitating the recall of information out of focus. "Defined named" locations were presented in absolute and relative forms. Relative forms require a reference to establish meaning unlike absolute forms whose meaning is independent of a reference frame. Once a part was in focus, sections of the part could be labeled with relative names such as the "back," "front," "bottom" and so on. Destinations were also labeled with relative names. Interestingly enough, "defined naming" was not used often. In place of exact names, descriptions were frequently used and with moderate success. References like "the bulky piece", the "big black plastic piece" and the "rod with the yellow rollers" are good examples. It should be noted that the ease of recall using descriptions increased as the assembly progressed and number of repeated references using these descriptions increased. Yet, if the part, location, or orientation was totally out of focus, the expert was required to give a summary description which mandated much more effort than an exact name for recall. Repeated descriptions became a problem when similar descriptions were used for different locations or part.

Order of assembly references recalled on information based upon when the information was mentioned in the assembly. These references are characterized by phrases like "the first part we identified," "the previous insertion."

References to events are essentially descriptions. A referenced event such as a previous part identification may provide recall of general characteristics of a location or part or may directly specify these characteristics. References to events can be used to note similarities between a task currently being performed and a previous task.

Or simply, the event might be more easily referred to by how it happened instead of what actually occurred. These event references were not effective when used as the only means of recall.

E: Remember the one we had trouble with...

Rarely did a type of reference stand alone, but combinations of the reference devices provided recall.

8. DESCRIPTORS

A descriptor is a word or phrase that refers to specific characteristics of a part, orientation, insertion, or location. The descriptions used in these experiments depend primarily on what is to be described and the surrounding environment. While the parts, locations, orientations, and insertions may be constant, the surrounding environment changes as the assembly progresses.

In the dialogs analyzed, ten distinct types of descriptors were observed: characteristic identifiers, positional, orientational, insertion, analogies, functional and possibility, result, perceptual, and negation. Following is a discussion of the descriptions with some examples from the dialogs. A set of descriptors necessary for task descriptions is also provided as given by the two sets of experiments studied.

8.1 Characteristic Identifiers

Characteristic identifiers such as name, shape, size, material, texture, or color provide descriptions of the attributes of part or location. For part description, characteristic identifiers were always used. Characteristic identifiers can be either what we will call general identifiers or solo identifiers. General identifiers specify characteristics that may be shared by a number of parts or locations. These identifiers primarily serve to provide focus rather than single out the part or location. Solo identifiers, on the other hand, describe characteristics which single out the part or location within the focus established by the general identifiers or within the entire collection of part and locations to be located. The difference between general and solo characteristics depends on the context. It should be realized that these seven identifiers rarely, if at any time, were used singly to describe a part or location. It was combinations of the descriptors which achieved identification. The frequency of each descriptor's occurrence was not tallied in the analysis. This is because different words can express similar concepts [7]. For example, pole and rod were used in two different experiments to describe the same part.

Teaching characteristic identifiers is usually presented in two forms—by question or request. A question essentially asks if the particular characteristic can be located. This approach insures a response or verification from the apprentice indicating that either more information is necessary or that the characteristic has been found. A request is just that—a request that the apprentice attempt an identification of some characteristic. In this case, the expert must request verification.

The identifiers most frequently used were names, color, material, texture, shape, and size.

Names. Names that provided characteristic identification worked on the assumption that the description was implicit in the name. Examples of names in common usage in the dialog studied are listed below.

ruler	bar coding	wiring	tabs
drill bit	switch	slots	circuit
knob	connector	spool	grillwork
typing	prongs	trough	teeth
valleys	printing	indentations	motor
board	pole	radiator	button
rollers	outcroppings	handle	part
cork	groove	openings	hooks
rod	bar	gears	wheels
grating	shovel	crevices	PC board

cylinder

Shape. Shape is an identifier that described parts and locations of destination. One basic criteria for the success of a shape description is establishing the orientation of visual viewpoint. Many shape descriptors are sensitive to part orientation. Failure to take this into account may lead to ambiguities.

E: Yes, tell me how you are holding it in your hand.

A: The white part on this one that I have in my hand is towards the back and the flat side of it oriented out to the right side .

E: The flat side is oriented to the right so there is a dark plastic piece that curves along the flat part and out toward you.

A correct orientation certainly comes into play when referring to symmetry. Symmetry is a very useful descriptor when talking about rods and gears.

Following is a list of the shape descriptors found in the dialogs studied.

flat	rectangular	long	short
thin	circular	semicircular	fat
shaped	cylindrical	square	skinny
spiky	narrow	points	taller
round			

Analogies were also drawn to compare the shape of named descriptors of common knowledge and the shape of the part or location in focus. One expert compared the shape of a paper tray in the printer to that of a snow shovel.

Not only was shape a common characteristic identifier, the lack of describable shape effectively provided information most often in the case of part descriptions. These terms were used when the expert realized that no shared knowledge existed to identify the part's shape and solo identifiers were difficult to describe. A list of these non-specific shape descriptors as observed in the dialog are given below.

horrible	strange	stuff
odd-shaped	thing	bizarre
funny	pieces	clippy thing
different		

Weight. Though, not used often, weight was invariably mentioned in the proprinter assembly with respect to a particular piece due to the fact it was extremely heavier than all the other parts.

Size. Size was described in either relative or absolute terms, most often relative. Size was always mentioned in reference to parts such as gears, cylinders, rods and most often in a relative sense. When referring to the size of a gear, rod or cylinder, references such as a small gear, "large rod" were all references to the diameter size. Size relatives were referenced to objects in focus to objects of common knowledge. Examples of absolute size include "about the size of a quarter," etc.

Relative sizes such as the following were employed in the dialog.

big	small	little	medium
large	biggest	little bity	

Color. Color was perhaps the most consistently used identifier for descriptions. Most uses of color was very ordinary as expected; yet, some were specified in relative terms such as dark, light, whitish. Color was primarily used as a general identifier, as numerous parts in the assembly were either white, black or gray. One of the parts had yellow wheels and another had a red switch. In these two cases, color identification of these parts always uniquely identified the part. Holes and slots were sometimes denoted as having color according to the surrounding material.

Material and Texture. Material and texture were usually general identifiers when addressing part's overall material or texture makeup. Deviations in material such as an attached piece of foam were invariably used as solo identifiers. Those mentioned in the dialog were plastic, foam, metal and aluminum.

8.2 Analogies

Analogies are descriptions relative to extracted information either from the common knowledge of the apprentice or learned information during the assembly. Four different types of analogies were observed in the experiments studied. The unknown parts, locations, orientations, or insertions were described to be

1. "like a common object,"
2. "like a common function,"
3. "like a common texture,"
4. "like a similar orientation, insertion or part which must be recalled."

The success was contingent upon the apprentice being familiar with the object upon which the analogy is drawn. Also, the expert must subsequently note what the unknown object has in common with the subject of the analogy.

8.3 Functional and Possibility Descriptions

It should be noted that texture, possibility and function descriptions did not always appear in the form of an analogy. The following dialog is an example of a function description.

E: Ok, we have two black pieces that work as framework for either side... one is larger than the other side.

A prime example of an explicit possibility description is given below.

E: Looking straight at it... There is a round circular space where it... that looks like you could possibly put a gear into... hold the ...the round piece in your hand.

E: There are also two tabs on the ruler that look like they should fit into it. They are separate from the rest of it and you should be able to fit the ruler piece right into that.

Texture and material as mentioned previously can act as a general or solo identifier.

8.4 Result Descriptors

Result descriptions were typically characterized by an instruction to perform an action with a specified expected result. The words "so that" and "until" frequently occurred in the dialogs studied. Result descriptions were used most often in orientation and insertion subtasks. These instructions deleted the explicit instruction of how to perform the subtask but simply asked the apprentice to

orient the part or insert the part with a resulting orientation, position, or tactile specification. The resulting orientation, position or tactile perception supplied the "how to" information to the apprentice.

ex.1 E: *Ok, so slide it downwards so that it clicks into place.*

ex.2 E: *And push it all the way backwards so that it now sticks straight up in the air.*

The "until" descriptors were observed in both the insertion and Proprinter experiments. In the Proprinter assembly, these types of descriptions were characteristic of insertion specification. These instructions described an action to be perform "until" some specified force or tactile indicator was perceived. Two examples from the insertion experiment are given below.

"Lower it until you can wiggle it back and forth but you can not turn it."

"When screwing it on, it doesn't take much pressure at all and screws on easily. Continue to screw it on until at some point you can feel it is getting harder and harder to screw on. Screw it on until it approaches a limit and stays there and it takes the same kind of pressure to keep it turning."

8.5 Perceptual Descriptors

Some descriptors by the expert called for explicit use of visual and tactile cues by the apprentice.

Tactile. Tactile descriptions were rarely provided explicitly. Words such as "click," "snap," or "slide" were descriptions of insertions from which the apprentice would ascertain how the insertion was to feel.

The following group of descriptors were found repeatedly in each of the five proprinter experiments and in the three insertion experiments when describing actual insertions

slides	clips	push
put	snap	pull
plug	fit	clamp
twist	dropped	rotate
push	press	mount
set	install	click
connect	place	slips
mesh		

Instructions such as "it should not move" after the insertion were clues to how the resulting position

should feel. We found that so many tactile perceptions are taken for granted that the expert was not required to explicitly provide the information. If a human understands what sector of the part to insert, at what orientation and into which destination, the rest is automatic. This was the reason for the second set of experiments.

In the insertion experiments, additional information was provided as to how the insertion felt and what kinds of indications existed to signal correct or incorrect insertion. The tactile information was supplied by noting the compliance of the part being inserted and the forces which were detected.

Compliance is the inverse of stiffness. Indicators such as "it fits snugly," "it's real wobbly" or "there is a lot of play" provided information to the compliance of the part. Changes in compliance throughout an insertion supplied progress information. Compliance was also the primary indicator of a correct or incorrect insertion.

Force detection was most often observed in determining completion. Instructions were given to "push," "pull" or "twist" until some amount of force was detected. Notation of what forces should be expected were also made. These notations were made to provide information as to the kinds of forces which would need to be supplied.

balization. Visualization essentially asks A to imagine the tasks before performance.

E: You can see that there is a little hole waiting for it ... If you were to push it in there... it is about a circular piece and then you could push it in and click it.

8.6 Positional Descriptors

There were three types of position descriptors: absolutes, extremes, and relatives. For the proprinter, all insertions were made on the baseboard or onto a part already in or going into this board. Therefore, a reference frame was attached to this general destination location.

Absolutes described positions or locations which were visible to the subject such as "on the table" and "toward the ceiling." Extremes are descriptors such as corners, edges, sides, and end. Relatives were numerous. Following is a short list.

center	top	right	upside down
far	above	below	upward
bottom	back	inside	left
around	middle	front	outer
next to	forward	away	inner

farther away

above

over

behind

Positional descriptors are defined in two different types of reference frames: global and regional. Global relatives are established by the environment as a whole. This environment includes the table on which the assembly takes place in relationship to the person performing the assembly, the ceiling and the floor. Therefore, relative directions such as "right" and "left" are obviously established. Also, in certain subdialogs, "back" and "front" have global definitions. "Back" was defined as being away from the person performing the assembly while "front" was defined as closest to the person. Examples of these global specifications of direction are most frequent in orientation subtasks. References to tabs or similar extensions "sticking out" either "upwards" or "downwards" were defined within the global reference frame.

Regional position relatives were observed most often in describing locations on parts or locations of insertion destinations. In these cases, two factors become very important. These are knowledge of the part's or destination's characteristics and an established reference frame. It is these two pieces of information to which regional position relatives referred.

Problems encountered by using relative positions such as up, down, top and so on were a result of ill-defined reference frames which is a mistake made by humans due to egocentricity.

8.7 Orientation Descriptions

Orientations were described in one of three ways: (1) a characteristic identifier was positioned in a reference frame, 2) (the orientation of the entire part was described or (3) the resulting position was specified.

When using characteristic identifiers, general identifiers for a particular part were used to recall a part not in focus or to verify the correct part was being oriented. Otherwise, it was solo identifiers such as various extensions, typing or labeling, holes, handles which were easily positioned in a reference frame that described orientations. Characteristic identifiers were used in much the same manner as in positioning. Extensions and other solo identifiers were oriented most often in the general reference frame. Extensions or handles were noted as "sticking out" and typing, labels, or holes were often "facing out" in some relative direction such as up, down, left, and right. These types of descriptors will be known as directional pointers.

Some relatives can be positional as well as directional. For example, the "back" of a part may indicate a location on the part while an extension may point to the "back" to define an orientation.

When describing the orientation of the entire part, two descriptors were most common—vertical and horizontal. In teaching humans as well as robots, words like horizontal and vertical should be defined in a global sense. By this I mean, defining "parallel to the table" as horizontal and "perpendicular to the table" as vertical. This eliminates the question often asked in this experiment, "Horizontal to what?"

Some orientations were specified according to the resulting position of an insertion. These orientation specifications were based upon the theory that given the resulting position after the insertion, the correct orientation to begin insertion could be deduced. Given that the apprentice's viewpoint was correct and the resulting position was described, the expert would subsequently ask "Do you have or can you find the resulting orientation X?" The reference frame in these incidents could be either the general reference frame or one established by the destination. This type of orientation specification occurred primarily when specifying orientations explicitly for insertions.

8.8 Insertion Descriptions

When attempting to specify an insertion, four components of information must be present. These are

1. Location on part that will be inserted
2. Orientation of part at insertion
3. Location or destination for insertion
4. Mode of insertion

Three approaches were used for locating the insertion destination:

1. with reference to surrounding locations or parts previously inserted;
2. with reference to resulting positions of the part to inserted. This works with parts like rods. If the final location of the rod is known, the location of insertion can be deduced;
3. type of insertion. When the apprentice looks at possible insertion locations, the type of insertion will eliminate some of the possible destinations. (For sliding insertion, it is appropriate to look for a slot.)

The insertion experiments were investigated to observe the necessary components of information for insertion technique descriptions. The subjects were encouraged to free association and openly express what kind of tactile information they were receiving that indicated the progress of an insertion. Therefore, no assertion as to the structure of dialog during the actual insertion can be made. Following is a list of the components found in the dialog.

1. Initial insertion instructions. Sample instructions are to drop, insert, hook. Approach angles for easier insertion are specified when appropriate.
2. Notations of compliance of initial insertion. At this point, the differences in correct and incorrect insertion as determined by compliance were usually mentioned.
3. Additional insertion instructions. Indications as to the force that one should expect to encounter and changes in compliance during insertion are made. Also, additional descriptions such as "screw on," "push," "click" and so on are made. It should be noted, though, that insertion descriptors can exist anywhere in the specification of an insertion. Points or locations where a specified force should be applied were most frequently made here.
4. Compliance and force to be expected for indication of completion.

8.9 Negation

A strategy that frequently proves useful when other strategies are too difficult is negation: by describing a number of possibilities and subsequently eliminating those which do not apply. The approach was a process of elimination. Negation served as an excellent teaching technique.

E: They are not rods, they're not...huh... they do not have large metal piece attached to them, they are strictly plastic and they've got little pieces sticking out of them.

Negation is a good method of employing the surrounding environment to assist descriptions. In two of the dialogs, the expert had difficulty describing one of the two motors in the assembly. As a solution, the expert described the other motor and then located the desired motor to be located as not having the characteristics mentioned.

9. DIALOG STRUCTURE

9.1 Transition of Dialog from Task to Task

In the experiments studied, we observed some regularities in the dialog structure. The transition of dialog between tasks was defined by the order of assembly.

Introductions to new tasks were always introductions to parts identification subtasks. Words such as "next" provided additional clues that a new task was forthcoming. By noting these transitions, a change of focus was also noted. Though the focus established for an entire task can not be distinctly defined, there are elements of interests within every task that must be satisfied. In other words, once the question of "What part was to be identified?" is answered, all other parts can be momentarily disregarded unless another part is to be inserted into the part identified. Nonetheless, a primary focus is established and subsequent subtasks such as orientations and insertions become pertinent to this identified part.

Because the insertion subtask was always the last subtask to be performed within a task, the closure of an insertion procedure was often synonymous with task closure. Closures almost invariably were performed with the use of verifications. Incidents did occur where no verification took place either in the insertion subtask or the task as a whole. These occurrences were often due to the simplicity of the task or the fact that a similar task had been previously performed. The effort or detail given to closure was frequently a function of the difficulty of the task, and consequently a function of the difficulty in providing descriptions and instructions.

Once a task closure takes place, it is a signal to the apprentice that a new part will become the focus of attention and previous information can be disregarded unless recall is requested.

9.2 Transition of Dialog Within a Task

The following flow chart best describes the usual transitions of dialog within a task observed in these experiments.

Identifications were always the first subtask in a task. The initial identification within a task provided a general focus. Thus, questions such as "What is being inserted?" and "What is being oriented?" were answered. Even if additional parts were identified such that a task existed within a task, references were still made to the part initially in focus. The best explanation of a task within a task is provided by an example.

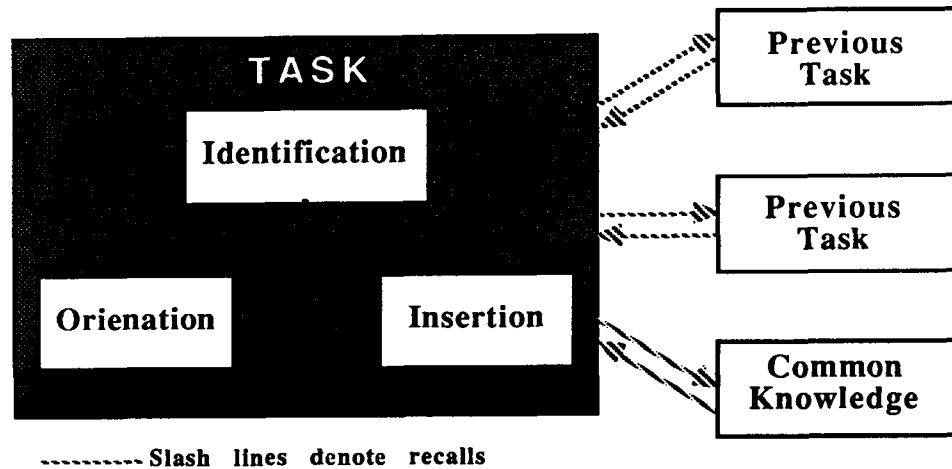


Figure 9-1: Transitions of Dialog

(The black frame in which the gears will be inserted has just been identified; it is noted as the piece of longest dimension.)

E: *Take the piece that has this longest dimension. That is the one we will be working with now.*

A: *Basically the bigger of the two?*

E: *Right, the bigger of the two*

E: *Alright, before we put this piece onto the board we have to put two pieces onto this piece. Now if you will look on your table there are several gears.*

A: *Yes.*

(the appropriate gears to be inserted onto the black frame are identified, and the gears are inserted into the black frame.)

E: *Ok, I am going to describe another place now on the green board again, where we are going to put this whole piece.*

(Finally, the black frame with gears is inserted.)

In Figure 9-1, the arrow connecting orientation and insertion is bi-directional because sometimes descriptions concerning orientation and insertions are interspersed together. This occurs when orientations are given specifically for insertions. An orientation was described successfully during an insertion subtask as long as it was clearly indicated that orientation instructions were being given to achieve later insertion goals.

Yet, equally or even more frequently, the orientation description occurred directly after the identification. This was primarily due to characteristic identifiers in immediate focus which were used in orientation specification.

Insertions were always the last subtask to be performed in a task.

9.3 Transition of Dialog Within a Subtask

9.3.1 Parts Identification

The approximate location of the part and the characteristic identifiers for description were the two essential components of part identification. In most cases, the initial descriptions of the part were general and followed by more specific descriptions which singled out the part.

9.3.2 Orientation Specification

The two most prominent patterns of orientation specification are shown in the flow chart below.

The specification of proper orientation has proven to be the most difficult aspect of teaching an assembly task. Orientations were usually established in the global reference frame spoken of in section . Exceptions sometimes occurred in orientation specifications for the purpose of insertion. In such cases, the orientation was described relative to the insertion destination or the viewpoint of the apprentice.

One of the three types of orientation descriptions followed a reference frame description: resulting position, characteristic identifier or entire part description. Orientation specified according to the resulting position of an insertion obviously occurred during insertion subtasks. Establishing an orientation via characteristic identifiers or the entire part usually occurred when no specific reason for establishing the orientation was given. In the most successful cases, verification followed the orientation description.

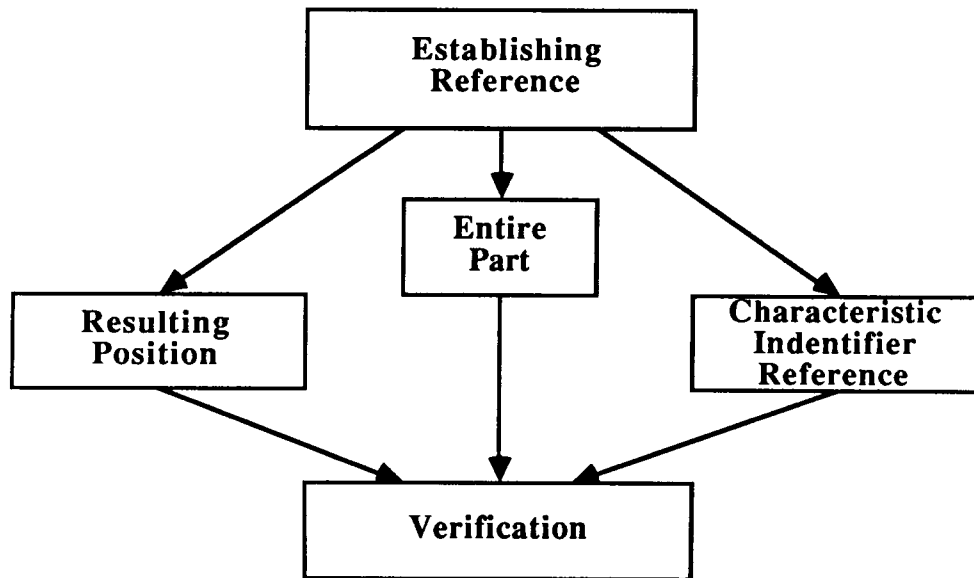


Figure 9-2: Patterns of Orientation Specification

9.3.3 Insertion Descriptions

Location on the part to be inserted and the destination for insertion were frequently mentioned initially in the insertion subtask. No clear pattern was observed that predicted which of the two would be described first or second. Orientation specification usually followed. The description of the technique of insertion was always the last of the four components described. Deviations to this structure were also observed yet, this structure occurred most often.

10. IMPLICATIONS FOR A ROBOTIC SYSTEM

The goal of this research was to determine the scope of information necessary to teach an assembly task. From the dialog analysis of humans conversing to achieve an assembly task, we have derived some important aspects of the assembly process that will assist in the understanding of the necessary elements in human to robot communication for the same purposes.

For the robotic system, it is necessary to determine which aspects of the assembly process are constant among different assembly problems, which aspects are a function of the situation and which aspects are a function of the environment. The constants can be hard-coded for the system. The aspects that vary from situation to situation need to be represented as parameters that the supervisory program can handle. The actual parameters appropriate for a given assembly step need to be deduced at teaching time, stored away, and passed back to the action subroutines at execution time. The aspects that are a function of the environment need to be dealt with by sensory feedback—this capability will also be hard-coded into the action routine. Each action routine will be an "expert" at its task, customizable by varying its input parameters, and able to deal with limited variations in the environment.

The constants of the assembly process are the information considered by both the expert and the apprentice as "common knowledge." Many of the movements, such as how to rotate, twist, push, pull or screw something on, were understood by both subjects. Gears and rods were examples of part descriptors that were mutually understood. Many of the characteristic identifiers were shared concepts. These include the common colors, shape, and materials.

From this analysis, we have also obtained some understanding of what needs to be parameterized for identification, orientation, and insertion. For part identification, we have observed seven types of descriptors which characterized a part. Orientations can be specified by two methods, each requiring different types of parameters. The orientation may be specified with respect to the entire part in which some axis or plane on the part is specified as vertical or horizontal to some axis or plane in the environment. The second method of orientation specification requires a characteristic and a direction in which that characteristic identifier points. Parameters to insertion tasks require four basic components: location on the part to be inserted, destination of insertion, orientation of part, and technique of insertion. Also, parameters for the actual insertion process, which include force limits, torques, insertion approaches and clearances, will need to be specified.

Though they are not complete in themselves, relative descriptions provided some very rich

language which could be used if parameters were employed. Comparative descriptors such as larger, smaller, smallest would require two parameters. What? is smaller than What?. Given a group of parts, parameters such as their similarities or singularities could be specified.

A choice among various methods of verifying the completed subtask should also be a parameter. If all the parameters are not specified the system should be able to query the teacher about such oversights. The system should also be capable of query if an identification, orientation or insertion is seen as impossible.

Many processes (insertion, for example) must not only rely on common knowledge and given parameters for different stages of insertions, but also use sensory feedback to allow for those aspect of the process which are a function of the environment. For insertion purposes, the sensory information will provide indications when parameters such as force and torque limits are satisfied.

The actual organization of the assembly process varied somewhat from one dialog to the other. Yet, there were some consistencies which allowed us to determine a rough structure. The final structure for organization can account for variances by including queries by the robotic system at appropriate times. The flow chart below exemplifies the organization of the assembly process as seen in the dialogs.

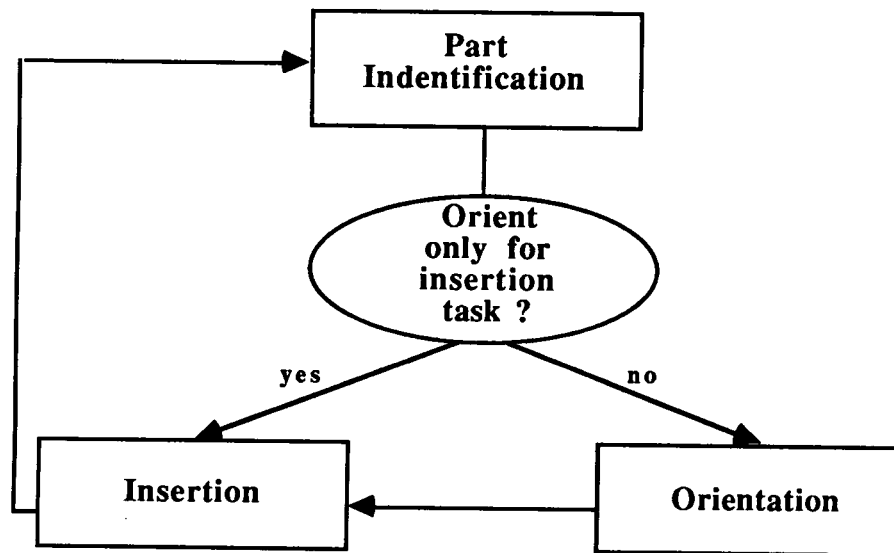
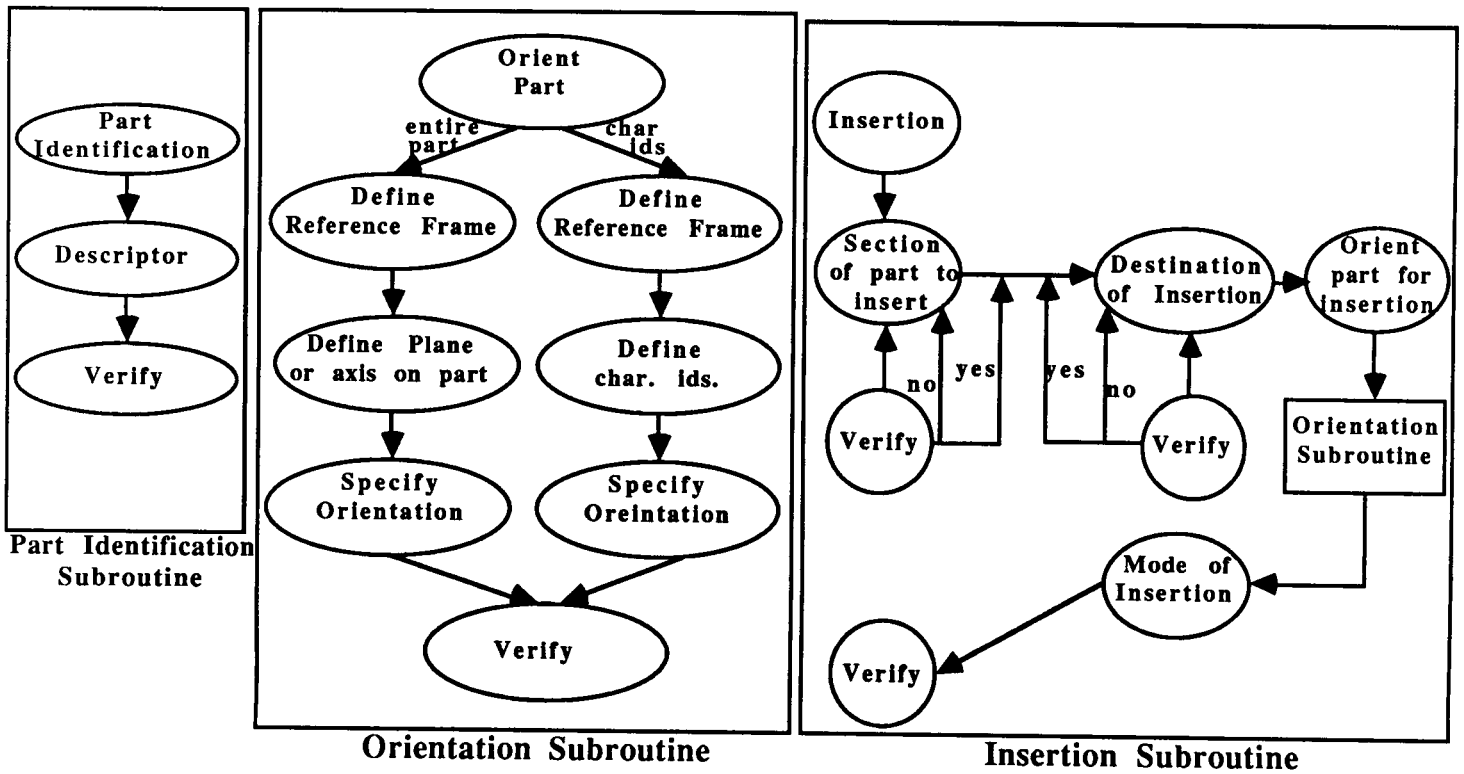


Figure 10-1: Overall Organization of Assembly

11. CONCLUSIONS

In an effort to better understand the assembly process, we examined the scope of discourse in dialogs in which an expert taught an apprentice an assembly task. From this dialog analysis, we were able to observe teaching methods for assembly, and the kind of information transmitted during an assembly task. Consequently, we feel that we have a better understanding of how to structure a dialog when a human teaches a robotic system an assembly task and the kind of information which will be important for teaching.

The dialog analysis used for this research served as a very important tool. It provided close examination of the multitude of components in the discourse for purposes of assembly. Examination of the variety of dialogs as a result of various modes of communication ignores the variances in the dialog but concentrates on similarities.

Further research of interest related to this research should be considered. First, the implications of focus in assembly tasks for human discourse and consequently for the human-robot system discourse are worth further study. Second, proposals to resolve the ambiguities of "relative descriptions" and formulate a means of expression for human-robot communication would explore a very potentially information-rich description type. And third, an extension of the insertion experiments can be examined because tactile information during insertion is still very difficult to obtain.

References

- [1] Akin, O. and Reddy, R. (1977): Knowledge Acquisition for Image Understanding Research. *Computer Graphics and Image Processing* 6, pp.307-334.
- [2] Bruce, B.C. (1983): Belief Systems and Language Understanding. *Trends in Linguistics, Studies and Monographs 19: computers in Language Research 2*, Walter de Gruyter and Co., New York, New York.
- [3] Chafe, W.L. (1982): Integration and Involvement in Speaking, Writing, and Oral Literature. In Tannen, D., Ed., *Spoken and Written Language: Exploring Orality and Literacy*. Ablex Publishing Corporation, Norwood, New Jersey.
- [4] Cohen, Phillip R. (1984): The Pragmatics of Referring and the Modality of Communication. *Computational Linguistics*, vol.10, no.2, pp.97-129.
- [5] Cohen, P.R. and Levesque, H.J. (1980): Speech Acts and the Recognition of Shared Plans. *Proceedings of the Third Biennial Conference*, Canadian Society for Computational Studies of Intelligence, Victoria, B.C., pp.263-271.
- [6] Cohen, P.R. and Perrault, C.R. (1979): Elements of a Plan-based Theory of Speech Acts. *Cognitive Science*, vol. 3, no.3, pp.177-212.
- [7] Grosz, B. (1982): "Sublanguage—studies of language in restricted semantic domains" Walter de Gruyter., New York, New York.
- [8] Grosz, B. (1977, July): The Representation and Use of Focus in Dialogue Understanding. Technical Report 151, Artificial Intelligence Center, SRI International.
- [9] Hindle, D. (1983): Deterministic Parsing of Syntactic Non-fluencies. *Proceedings of the 21st Annual Meeting of the Association for Computational Linguistics*, Cambridge, Massachusetts, 123-128.
- [10] Kroch, A.S. and Hindle, D. (1982): On the Linguistic Character of Non-standard Input. *Proceedings of the 20th Annual Meeting of the Association for Computational Linguistics*, Toronto, Canada, pp.161-163.
- [11] Perrault, C.R. and Allen J.F. (1980): A Plan-based Analysis of Indirect Speech Acts. *American Journal of Computational Linguistics*, vol.6, no.3, pp.167-182.
- [12] Thompson, B. (1980): Linguistic Analysis of Natural Language Communication with Computers. *Proceeding of COLING-80, Tokyo*, pp.190-201.