WIS-CS-74-205
COMPUTER SCIENCES DEPARTMENT
The University of Wisconsin
1210 West Dayton Street
Madison, Wisconsin 53706

January 21, 1974

# ENGLISH AS A BASIS FOR COMMAND LANGUAGES FOR MACHINES AND SOME PROBLEMS OF SPATIAL REFERENCE

by

Norman Sondheimer

Computer Sciences Technical Report #205

July 1974

# ENGLISH AS A BASIS FOR COMMAND LANGUAGES FOR MACHINES AND SOME PROBLEMS OF SPATIAL REFERENCE*

Norman Sondheimer


## ABSTRACT

Current research on natural language speech-understanding provides encouragement for the development of systems for the vocal control of machines. However, the designer of a natural language based command language still faces difficulties posed by certain areas of semantics that have not been well studied. One such area is that of spatial reference, which is the way people refer to objects and actions in space. This paper looks at some problems raised by terms used to make spatial references. In particular, the semantics of English locative prepositions and prepositional adverbs and a few related terms are analyzed. In light of this analysis, suggestions are made pertaining to the form that the spatial references that these terms convey could take in English-based command languages that require unambiguous reference.


---

*This paper was delivered in part as "English as a Command Language for Machines and the Semantics of 'Left' and 'Right'" at the Milwaukee Symposium on Automatic Control, March, 1974.

TABLE OF CONTENTS

## I. INTRODUCTION

### Speech-Understanding and Machine Control

In the United States, there has been in recent years considerable interest in developing systems that allow man-machine interaction in spoken English. At one session of a recent artificial intelligence conference, efforts were reported underway to apply speech-understanding to a system which accepts chess moves (Reddy et al., 1973), a question-answering system for geological data (Woods and Makhoul, 1973), and one to interact with a simulated hand-eye robot (Walker, 1973). The inspiration for these three projects and much of the other recent efforts can be found in the report of the Newell Study Group on Speech-Understanding Systems (Newell et al., 1971). The report evaluated the state-of-the-art and suggested a program of research for the near future.

The Newell report does not suggest and these projects are not attempting, to develop systems that can comprehend completely unrestricted, naturally occurring English. Rather they have very limited goals. For example, they are only aiming for a vocabulary of a thousand words. So, strictly speaking, they are developing systems based on the English language. Even with these limitations, the completed systems will be able to operate in sophisticated domains and useful situations. The state-of-the-art is such that many other practical applications are possible.

One interesting application area is that of "machine control". This can be defined as the guidance of a machine by a human controller as it mechanically operates in its workspace. A speech-understanding system should supply to a machine control environment the ability to give that guidance in a command language that is derived from a natural language, in which case English would be a basis for command languages for machines, as the title states.

There are many applications where such a capability would be beneficial. For example, a computerized hand of the type developed by various artificial intelligence research groups, e.g., see Dobrotin and Scheinmann (1973), might with vocal control be capable of performing the basic operations in radioactivity laboratories. An industrial hand like the Unimation UNIMATE should benefit from vocal supervisory control as it performs its material transfer functions. In fact, plans exist to develop such a capability as part of an advanced automation project at Stanford Research Institute (Rosen et al., 1973). Vocal control could be applied to the home to allow a bedridden invalid to manipulate his environment. With vocal control attached to mobile warehousing equipment several machines could be supervised by one human. These examples would be especially strong if the vocal supervisory system could be interfaced with algorithmic or heuristic routines to allow the human to assist the machine by advising it or modifying its course of action.

What makes machine control based on a natural language desirable is its naturalness. The fact that a controller will be using forms from his native language to guide the machine should a) simplify his training and b) give him a spontaneous way to control the machine. The latter opinion is based on the assumption that a controller will find it easier to express himself quickly in his native language. Further, manual control of machines is often complex. For example, the UNIMATE has one control for each of its five or six degrees of freedom. Yet it performs operations that have common verbal descriptions, for example "putting" and "taking". If the understanding system could do the conversion from the language forms to the instructions corresponding to the individual degrees of freedom, not only would the control be more natural but it should also be effected more quickly. The arguments above could also be made for nonvocal natural language machine control. Vocal control has even more arguments in its favor. With it the controller is free to move about with the flow of the work and use different perspectives when deciding on appropriate commands. When dealing with complicated machines a vocal supervisory system could provide a "third hand" when the human's two are occupied. Finally, if the controller is a paraplegic or quadraplegic, vocal control could in a sense replace the nonfunctioning hands.

Besides being desirable, natural language machine control appears, on the whole, to be feasible. It would seem that a set of utterances which are a subset of the English language should be discoverable that a) gives complete control of a machine to a human controller and b) is simple enough linguistically to be recognized by state-of-the-art systems. In fact, short phrases and a prescribed vocabulary might even be sufficiently powerful (Rosen, 1972, page 53). Examples of the kind of utterance required include forms like "a little higher", "to the right", "stop at once", "six inches, forward", "cancel last command", and "repeat last step". A language on this order should easily fit within the bounds that the Newell report sets.

Of course developing a command language will not be a trivial task since the language will have to be carefully designed. Not just any one thousand words will be accepted. Not just any grammar will fit a parsing scheme. Not every reference related to the control of a machine will be understandable.

This last topic of reference or meaning is one of the most difficult because less is understood about the semantics and pragmatics of language. To be sure there is work being done here, e.g., see the references in Simmons (1965) and (1970). There has even been excellent work done on meaning in machine control environments, e.g., Coles (1969) and Winograd (1972). But there are still areas where it is not well understood how people convey meaning, e.g., anaphora, presupposition and quantification. The designer of a system should consider the literature available on these subjects carefully.

There are, however, problem areas that have not been considered in the literature. One of the most important of these for the designer of a machine control system is that of "spatial reference". This can roughly be defined as the way people refer to objects and activities in space. Since machine control systems operate in physical space an understanding of the semantics of spatial reference is essential to a designer.

Included among the various types of spatial references are those made to directions, orientations and spatial relations. The typical terms that express this last kind of reference are locative prepositions and prepositional adverbs, e.g., "up", "over", "behind" and "in front of". Examples of the variety of problems that arise with these words include figuring out what direction is intended by the order "move back" or deciding what relation to use to define the side of the square the contact is to be on so that the instruction "put the circle on top of the square" can be carried out. One of the goals of this paper is to assist the designer of a natural language machine control system by analyzing the meanings of these terms.

What is done in the paper is to study some of the elements that go into understanding these references. A wide range of examples are given in hopes of showing the complete range of ambiguities that arise when there is confusion about the identity of some of the elements or about their characteristics. This analysis is found in Sections II thru IV.

Note that other words share concepts with the locative prepositions and prepositional adverbs. One example being words that refer to spatially defined parts of objects, e.g., "left", "top", "front" and "upper". Also verbs like "advance", "drop" and "follow" show the same phenomena. Some of these terms will figure in our analysis. In fact the discussion starts off with "left" and "right".

There are a variety of ways a designer of a machine control system can make use of an analysis of this type in developing an English-based language. One way he can use it is to identify unambiguous forms. He can also use it to identify how to restrict meaning so that phrases can be used unambiguously. These types of results for spatial reference are given in Section V.

## II. FRAME OF REFERENCE

### "Left" and "Right"

As was stated the analysis of the problems of understanding spatial references will begin with a look at the problems of understanding "left" and "right" and following such apparently simply orders as "eyes right" and "move to the left". Consider the old gag that has a new recruit lined up with his platoon. The first time his drill sergeant says "eyes right", the recruit finds himself looking straight at the man to his left: he looked the wrong way. On one level this is humorous because the recruit does not seem to know his left from his right. In truth, something like this experience is common to many people. For example, there are those situations where they have found themselves asked by someone walking towards them in a hallway to "please, move to the left" and then found the other person moved the same absolute direction they did.

What is the problem here? One possible explanation of the recruit's problem can be found if we consider the recruit's and the sergeant's physical placement; namely, that they are facting each other. In this situation one's right side is across from the other's left and his left is across from the other's right. We can think of them as possessing "left/right axes" that are parallel but opposite. When the sergeant uses "right", the recruit does not know if he is referring to the right of the recruit's axis or the right of his own axis. He picks the sergeant's axis. Turning so that he satisfies that interpretation he ends up doing what the sergeant considers "eyes left".[2] Of course, the rules governing military drills are generally well known so the recruit's response appears humorous to people. However there are no rules governing requests for movement in hallways. Since the physical situation is essentially the same, left opposite right, the speaker can either be using his left or his addressee's left and the addressee can assume either his own left or the speaker's left. In other words, there is with "left" and "right" the possibility of interpretations according to the speaker's or addressee's axis. When these axes differ the choice of which frame to use becomes important since the resultant reference can differ. This difficulty can be called the "frame of reference" problem.[3]

More than just the speaker and addressee have axes. Obviously, other people do. This too must be considered in establishing a frame of reference. Consider where you would look if when you were watching television, you saw a golfer on a tee through a camera behind the green and the announcer said "he hit the ball to the left". You could look for the ball on the side to your left or you could attempt to determine where the announcer is and use his left/right axis. But since the golfer has his own left/right axis you could also pick it. A survey of 36 people showed about 40% taking this option.

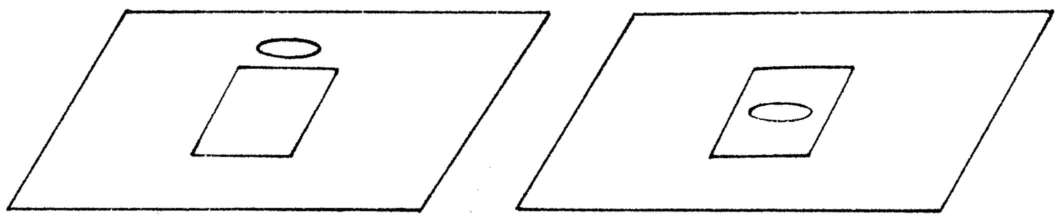
A more subtle source of frames are the objects in the environment. I have two cabinets on either side of my kitchen sink. I once told a friend that a certain pot was in "the cabinet to the left of the sink". My friend's first choice for the reference of that phrase was the cabinet that I would say is "to the right" of the sink. The explanation she gave for her choice was based on the consideration that if she were the sink looking out from the wall the cabinet she picked would have been on her left. I assumed she would use her own axis as she stood in front of the sink. Instead she had presumed that the sink has its own axis and used it. Hence speakers, addressee's, other people and objects can be the source of the left/right axis that serves as the frame of reference of an observation.

"Up", "Down" and Like Terms

The same problems that arise with "left" and "right" arise with terms connected with the two other human axes: the top/bottom and the front/back. The difficulties in establishing these references are distributed in importance somewhat differently. For example, with the top/bottom axis there are more often differences between the object's axis and the general human determination of vertical than there is between the different human observer's vertical.

Associated with the top/bottom axis are terms such as "on top of", "under", "over", "above", "below", "beneath", "underneath", "up", "down", and, of course, "top" and "bottom". The major element in observer's interpretation of these terms is the force of gravity. Because of gravity, everyone has essentially the same vertical. It is a reasonable approximation to assume that speaker, addressee and other living beings in the conversation use the same top/bottom axis. Objects, however, and people who are referred to as objects and not as observers have their own axes. For example, if a book were lying open on a table a reference to the "top" of a page would normally be interpreted according to a top/bottom axis that is perpendicular to gravity. Likewise a cut could be said to be on the "bottom" of a person's foot whether the person was standing or sitting with his foot upon a desk.

Difficulties arise when an object's axis differs from that of gravity. Consider the request to place two pieces of paper on a bulletin board which has been removed from the wall and is now lying on a table. "Put the circle on top of the square" can be interpreted according to the inherent top/bottom axis of the board (Figure 1a) or it could be established by gravity (Figure 1b). If a cereal box were lying on its side an order to "stamp the price on top of the box" would be similarly ambiguous.



a. b.

Figure 1

Although the important differences are between objects and observers, research reported in Rock (1974) has pointed out that there is in certain cases a need to consider the "egocentric coordinate reference system" and hence differences between observers' top/bottom axes. Consider for example a circle drawn on the ground. If an observer was standing next to it, he would say its top was the point farthest from him. If he was lying next to it, he would say the top was the point farthest from his feet, closest to or farthest above his head. In other words, what he considers the top of the circle is the highest point of its retinal image according to his top/bottom axis. Since the speaker's and the addressee's retinal images can vary, the frame of reference problem also appears here. But this effect is only noticed when no other object axis intrudes and the images are perpendicular to gravity, so gravity is still the primary approximation to all observers' top/bottom axes.

"Front", "Back" and Like Terms

The frame of reference problem also occurs with the front/back axis terms. These include "in front of", "in back of", "behind", "forward", "backward", "ahead", "advance", "front" and "back". Problems most commonly arise when objects in the environment can have axes that can be used as frames. For example, "John is sitting in front of Jane" is a valid description of the fact that John is forward of Jane on Jane's front/back axis. But if they were sitting in different rows of a movie theater it could also be a valid description of the fact that John is in a row which is ahead of the row Jane is in according to the theater's front/back axis. It is important to know which axis is the frame of reference, since Jane could turn around to face someone in the row behind her and hence John could be "in back of" on one axis and "in front of" on another. We have the same problem when a conversant's axis differs from the environment's. A classroom with its podium and blackboards has an identifiable front/back axis. If a person in the classroom is not facing its front, when he is given the order to "move back" he can either move to the back of the room or "back away" from what he is facing.

Natural examples of differences between a speaker's and an addressee's axis affecting the interpretation of an utterance are more difficult to imagine. One possible one can arise when two people are intently bird watching through binoculars. If one says "look straight ahead", the addressee, not being certain that the speaker knows which way he is looking, is left open to decide whether he is asked to look where the speaker is looking or to look in the direction he himself is facing.

One final aspect of the frame of reference problem can be shown by referring to the problems of giving people travel directions. You might tell a friend to "get onto Elm Drive and then take a right onto Fig Road". If your friend gets onto Elm Drive on the opposite side of Fig Road from where you expect him to,the turn he will make onto Fig Road will be the opposite of the one you intended. The problem arises because your reference to right uses a noncurrent frame of reference. The same holds true for such expressions as "you will see it on your left", "it was behind where the building used to be" and "imagine you were standing there, then it would be behind you".

To summarize, the source of the axis that defines many spatial references can be the speaker, the addressee, and animate and inanimate objects. These axes can be established by sources current to the observation or even noncurrent ones. Deciding which axis is being used is called the frame of reference problem.

## III. CONVENTIONS

### Conflicting Conventions

In the last section living and nonliving objects were described as having "fronts", "backs", "tops", "bottoms", "lefts" and "rights". Since these are linguistic descriptions of physical properties and only humans possess language, all nonhuman beings and things must be considered as being assigned these descriptions by human speakers. In a sense, man only knows his own left from his own right and has to figure out for everything else what its right and left must be. Unfortunately, people do not always come up with the same assignment for every object all the time. Since they do not, a speaker's and an addressee's understanding of a spatial reference can differ even when the frame of reference is known.

This problem can be demonstrated by considering some general ways that humans make assignments of axes to nonhumans. There appear to be at least two major sets of conventions or heuristics that are applied here. One set is essentially anthropomorphic conventions that are based on the identification of human characteristics in a being or object assign to them the axes as they would be assigned to humans. The other set is familiarity conventions based on how the human associates himself with the object through use or approach. These different conventions can be based on intrinsic or extrinsic properties of objects. There may be other conventions or better ways to describe their organization. However, this analysis is sufficient for purposes of showing how humans can come to different conclusions when assigning axes to objects.

Anthropomorphic Conventions

Anthropomorphosis is the ascribing of human attributes to beings or objects nonhuman. In a sense any assignment of axes is of this nature since all interpretations of spatial references are in light of the triple human axes. However, the anthropomorphic conventions discussed here are those based on the identification of other human characteristics which in turn cause the assignment of the axes. These characteristics include facial features, a standard orientation vis-à-vis gravity and a predilection to movement in one direction with respect to the object's body. These properties are easiest to identify in other mammals. They have the basic elements of a face on one side of their body and generally move with that side in the fore. Both these characteristics suggest frontness. They also have a standard orientation to gravity which suggests a vertical axis. Humans take these suggestions and assign front/back, top/bottom and, derivatively, left/right axes as they would be assigned in a human. The fact that people speak of an animal's "left ear" or "right side" serves to point out that people are assigning their properties where they are not inherent since animals do not naturally make a distinction between left and right, see Corbales and Beale (1971).

Where these and other properties are identifiable in other animals axes are assigned accordingly. Plants have their standard orientation to gravity and are hence assigned top/bottom axes. Man-made objects are subject to these processes. The fact that cars generally move a certain way suggests frontness and hence the front/back axis. Some boxes such as cereal boxes have more prominent graphics on one side. This suggests that this side is the "front". Objects that rest most easily on one side have a standard vertical orientation and hence are given a "top" and "bottom".

The anthropomorphic conventions are widely applicable and very powerful. The sink example in the previous section showed one unexpected application of them. With the wide variety of beings in nature conflicts are likely to occur. For example consider the crayfish. It is a small animal that has eyes, claws and mouth that suggest that its front is to one side. However, it moves most naturally in the opposite direction. In a situation on a picnic where you are attempting to capture one the instruction "get in front of the crayfish" is ambiguous. More problems arise with man-made objects. A mobile home has a front established by the side through which it is entered. But there are also people who identify another front which is the side which is in the fore when the home is towed. Conflicts also arise between these and other conventions.

Familiarity Conventions

When people spatially relate themselves to an object in one normal or predominant way, they tend to assign axes to it according to that relation. The relation could, for example, be based on viewing the object, wearing it or operating it. The "top" of a page or painting is the edge that would normally be highest in people's retinal image when they view them. The "left" pocket of a shirt is the one on the left side of the human when he is wearing the shirt. The "left" glove is worn on the left hand. The "left" burner on a stove is the one that would be to the human's left when he operates the stove. Containers have their top/bottom axes defined by the side from which the contents are dispensed, e.g., the "top" of a cereal box is the side people are supposed to open.

The front/back axis conventions show two different distributions. When a human makes a close association between himself and an object, it tends to pick up his front/back axis arrangement directly. This can happen when the predominant spatial relation is such that the human is contained in the object or it is attached to him. This happens with many objects, e.g., chairs, lecture halls, pants, shirts, glasses, telescopes, trumpets and so on. When the relation is looser the side the human faces in his normal position is the "front" and the side opposite the "back". The looseness is seen in the way humans view or use objects such as paintings, televisions, stoves, dressers, pianoes and many others.

When these conventions can be variously applied they can give conflicting results. Because humans are spatially associated with it in two different ways, some people define the "left" side of the traditional theatrical stage by what would be the audience's left while others use what would be the actor's left when he faced the audience. Some ice cream cartons have two sides which can be opened. One is made to be lifted and the other to come apart. Both are chosen by different people when they are asked to identify the "top" of the carton. Bureaucrats and their petitioners relate to a desk in two different ways. The bureaucrat thinks the side he sits on, the one with the drawers, is the "front". The petitioners think they face the "front" of the desk and the bureaucrat sits "behind" it. Of course, if a petitioner is also an office worker his judgment might differ.

It was mentioned above that the anthropomorphic conventions can conflict with these. For example, the "left" drawer of a dresser can be, and usually is, defined by what would be a human's left/right axis when he was opening and closing the drawers. However, some people actually anthropomorphize the dresser and set up an opposite left/right axis. This conflict is always possible whenever the front/back axis is defined by a loose connection between human and object since anthropomorphic conventions assign the left/right axis so that it always keeps the human relation with the front/back axis and here the two are assigned independently.

Use of these Conventions Based on Extrinsic Properties

Up to this point the examples of the applications of these conventions have been restricted to judgments on essential properties of objects. They can also be seen to apply to objects' unessential properties. For example, a plain symmetrical table such as a bridge table has no intrinsic "front" or "back". If such a table were pushed up against a wall, people would feel free to refer to an object as being on the "front" or "back" of the table. Rocks have no inherent front/back axis. If a rock was partially buried in a hillside, people could refer to someone who is buried "behind" it or standing "in front of" it. This sensitivity to extrinsic properties brings conflicts. If a truck were rolling down a hill backwards, many people would say that someone was "in back of" it if the person was in its path. This judgment would be indirectly based on the truck's intrinsic motional properties. Based on the truck's present motion, many other people would say that it had already rolled past what was "in back of" it and that the person was "in front of" it.

The most common convention-based problem arises because of extrinsic properties of objects derived from the position with respect to an observer. Only the interpretation of terms connected with the front/back axis is affected. If a photographer was heard to say to his subject "stand in front of the tree", since the trees have no intrinsic front/back axis it would be clear that he was telling his subject to stand on the side of the tree he is facing. This is a situational application of the familiarity convention that humans use to establish the front/back axes of objects with which they are loosely related. Here the relation that defines the axis is simply the present position of the observer vis-à-vis the object. The side he is nearest becomes the front and the farther side the back.

Apparently any being or thing is subject to this convention. This means that when an object has an inherent front ambiguities can arise. If a photographer says "stand in front of the car", he could want the subject to stand between him and the car or to stand next to the side with the grill and headlights.

Not the speaker alone need be the establisher of the axis. "Can you see the table or am I standing in front of it?" show that the addressee's relation to an object can establish the front. "When you approach the gate you will see the house behind the trees" is an example of a non-current position being used to establish the axis. So almost the entire range of frame of reference problems could show up when this one convention for assigning an axis to an object is applied.

## IV. SOME OTHER PROBLEMS

### Beyond the Axial System: A Path Structure

The discussion in the last two sections was based on the analogy of a triple set of axes. The intended impression was that the relations being discussed could be handled by thinking of straight lines either projecting from an object or floating in space. This analogy breaks down in many cases. Consider the example of one person being "in front of" another in a line. It is a common occurrence for lines at popular movies to bend around corners. In this case one person could be "in front of" another but not straight in front. Consider a highway intersection where either a turn is possible or the road could be followed further. Imagine that the road curves after the intersection such that a driver coming out of the intersection must turn his steering wheel to stay on the road. In this situation the instruction to the driver to "go straight ahead" should not be interpreted as implying that a straight line of movement is requested. The same phenomenon is evident in the use of some terms generally associated with the top/bottom axis. For example "Cairo is above Memphis" refers to their relationship on the Mississippi and only very indirectly to gravitational relations. Likewise "St. Louis is ten miles down the road" refers to measurements taken along the highway. Returning to the front/back axis, one unique example presents itself with the terms "back" and "backward". On a roadway the instruction "back up" can be interpreted, depending on the situation, in the same way as "move in the direction opposite of the way you are facing", "go backward following the road", or interestingly "go back along the way you went forward". In other words "back" and "backwards" have a special interpretation that involves the object's previous movement.

All these examples require that a new framework be developed in contraposition to the axial system. Since most of the examples that motivate it involve travel, an appropriate title for this system is a Path Structure. A machine control system will have to be aware when this structure rather than the Axial one is being used.

Notice that only one ordering, instead of three, need be postulated for the path structure. Some heuristics that establish the ordering on this axis are flow for river, movement to a goal for lines and highways, and history of travel for the "back" and "backwards" cases. Conflicts can occur in cases such as when an inexperienced river traveler going against the current applies his nonriver experience to say that something that is upstream is "down the river".

With this structure it is possible to describe the meaning of the terms "before", "after", "past" and one sense of "beyond". "Get off before the bridge" means to get off the road at some point less far along the path, which is defined in terms of the road and the direction of movement, than the bridge. "Go past the bridge" means to get to any point which is farther along the path than the bridge. There is a frame of reference problem in a different sense here. To talk of a point on a journey and relate other points to it requires some way of establishing the direction of approach to that point. This can be done by identifying a point as a frame of reference. For example the frame is clear in such sentences as "from here, the turn off is before the bridge" and "go from the crossroads till past the bridge". The frame could be the speaker's, addressee's, some third party's or an object's position. There can be conflicts when the frame is not specified in the dialogue or in the environment.

Some Special Problems

There are some important ambiguities independent of both the axial and path system that must be mentioned in this section to complete the discussion of spatial reference. First, "over", "up", "down" and many other terms appear as particles of verbs in situations where they are best considered as having idiomatic meaning involving spatial reference. "Watch over", for instance, suggests standing above someone but could hold even if this physical situation was not the case. "Stand up" and "sit down" suggest a change in height but would hold even if there was none. "Get up" refers as much to waking as to changing height.

"Over" is involved with "under" in another important class of meanings. Both words can refer to nonvertical relations. "Tape paper over the windows", "there is paint all over your face" and "he is wearing a shirt under his coat" refer to a sense of covering that is independent of the top/bottom axis. The same sense can be seen in "the furniture is under the sheet" and "the bridge goes over the water".

Finally, "over" has yet another meaning. It can be involved with a meaning showing a relation between two objects on either side of a third or one object crossing another, e.g., "he lives over the hill from here" or "climb over the wall". Similar senses involve "across", e.g., "put it across the room from the door", "through", e.g., "he went through the door", and "beyond", e.g., "from here, the hospital is beyond the library". "Across" and "beyond" involve a frame of reference problem like that mentioned for "before", "after", "past" and the other sense of "beyond". "Put it across the room" leaves open whether the position desired is across from the speaker, addressee or some prominent feature. "He lives beyond the river" could be from where the speaker or addressee are (consider a phone conversation or letter) or some third person or prominent object or place.

## V. PHRASES FOR ENGLISH-BASED SYSTEMS THAT REQUIRED UNAMBIGUOUS REFERENCE

In the last three sections a variety of problems were analyzed that were involved with establishing the intended reference of various spatial terms. The existence of these problems will have to be considered by anyone designing an English-based machine control system since he must somehow allow for spatial reference. How he must handle these problems depends to a great extent on the performance requirements of the system he is designing. There are a number of paths he could take.

One way of handling these problems that would be useful in many situations is to employ a heuristic method. In this way a designer could be allowed to include ambiguous references in his language. He could then have his system set up a space of all possible interpretations for each reference it receives. The elements of this space will be based on the differing frames, conventions and other phenomena that apply to the reference. The system could then apply heuristics to search the space for the intended reference. The search could be guided by various kinds of knowledge, both preprogrammed and learned, to help it arrive at its conclusion.

Unfortunately, with a heuristic method it is impossible to guarantee that its conclusion will be the correct one. Since many spatial references are subject to the Frame of Reference and Conflicting Convention problems this becomes a serious problem. Essentially, a heuristic method will often result in only a "best guess" solution. Unfortunately, the penalty for guessing wrong in some applications makes a heuristic method unacceptable there, e.g., in a radioactivity laboratory. Of course, in these situations the system could ask the human for disambiguating information.

Reliance on this device can undesirably affect the timeliness of the control. So a designer might be required to develop languages where he can be sure that the system will always arrive at one interpretation which will be the intended one.

In situations where a designer must assure communication does succeed, he essentially has to assure that spatial references have the same unambiguous meaning to a controller and to his system. To accomplish this he can identify the unambiguous phrases that make spatial references and include only them in his language. He can add to these the phrases that he knows are unambiguous in his context.

If the designer finds that he can not paraphrase all the references he needs to make in unambiguous terms, he must turn to other means. Basically he must identify terms whose meanings he can restrain to one sense. Then he must train controllers so that when they use these terms they intend these specific meanings. If the meanings he picks for the terms are among the various alternative meanings the terms inherently possess, then he is essentially restricting semantics as it was assumed he would restrict syntax. He must be careful in picking the meanings that he wishes to legislate. The best choice will be the one that minimizes controller training with respect to the other choices. Also a controller must consider how much of this type of legislation he will rely on.

A carefully designed English-based language of this type should be nearly as natural as the same subset of English without semantic restrictions. It can be expected to be much easier to train a controller in the use of it than something artificial such as reference through a grid and vector scheme. The language should also be easier for him to use.[4]

The rest of the paper follows through with this analysis in a general way for some of the spatial reference terms. In particular since most of the forms discussed have few unambiguous usages, it provides general suggestions on alternative means of expressing the references or on feasible restrictions of the terms' semantics.

Phrases for the Left/Right Axis

A designer knows in general that the interpretation of the words "left" and "right" are open to many different frames and conventions. Fortunately he can assume an unambiguous reference in a speaker's reference to his own axis in the first person possessive, e.g., "my left arm" and "to my left". Since every form in the language is a command given by one entity to another, it is unlikely that a language would be based solely on these forms. For example "move to my right" is a different kind of usage from "move right". One orders movement to a point, the other movement in a certain direction. It would be better to allow references to the machine's own axis to control its movement and refer to its parts. These references are clear in the second person possessive forms, e.g., "move to your right" or "close your right clamp". This, however, opens the problem of which convention specifies the arrangement of the machine's axis, which arises because the nonliving machine is receptive to many different conventions. Here then is an example of where it is necessary to train the controller to use one set convention. Since the machine takes the role of addressee in the conversation it would be reasonable to establish its left/right axis with respect to its front/back axis as a human's would be arranged. As to controller training, a designer could in fact do something as simple as painting the words "left" and "right" on the appropriate side of the machine and train the controller by pointing out the notation.

Continuing with the possessive form, if a designer wishes to design a system having two controllers, it is possible to use the possessive form "his" as long as the speaker knows that he is understood to be using the other controller's personal axis. The problem becomes appreciably more difficult if a designer tries to continue further with "its" or "the dresser's" or "of the dresser" or like forms. Training is necessary for every object and there is not always a natural choice for the orientation of the axes. Unless there are a limited number of objects in the activity space, a designer's best plan is to avoid referring to the object's frame and put everything in terms of "my" and "your", e.g., "the side of the box to my right" or "the drawer to your left". Again this can only work if the object's axis is parallel to either the controller's or the machine's axis. If a designer can not depend on this, he can force language that linguistically puts the machine in the right place by allowing the use of a noncurrent frame. This can be done by appending such phrases as "when you face...", where the dots are replaced by some clearly defined point and the meaning of "face" is explained. Examples of this "placing" language are "the drawer on your left when you face the drawers", "the side of the room to your right when you face the blackboards" or "the cabinet to your left when you face the sink". A designer must then make the system intelligent enough to model these situations. This suggests a solution to the travel directions and similar problems. If he feels a need for such references, a designer can force slightly extended placing language, e.g., "turn to what would be your left if you were on Elm Street coming from Newton Street". This last suggestion is rather awkward. When faced with the lack of current frames, phrases using "nearest" and "towards" could be used when there is an object to relate to, e.g., "the drawer nearest the window" and "turn towards the church".

Phrases for Other Terms

The phrases necessary to avoid ambiguity and requiring the least training for terms involved with the front/back axis are slightly different. "In front of me" is as clear as "my left" and "in front of you" is similar to "your left" in that it requires training. Again "in front of" anything else should be avoided. But differing from the left/right axis treatment, phrases such as "in front of you as you face..." can not be used if it is necessary to avoid the conflicting convention problems since it does not specify how far you are from an object and you could be between the two objects you are trying to relate. Instead these uses of "in front of" can be replaced by "between". For example, "the box in front of the table", using the table's front as defined by the speaker's view of it, could be said as "the box between the table and myself".

This last substitution seems to change the meanings that arise if more than one box is between the speaker and the table. Since the nearest box to the table is meant, the phrase "and nearest the table" could be added. With the type of order shown by "put the pen in front of the box" where placement should be as close to the box as possible, a designer could use "next to" with "between", e.g., "put the pen next to the box and between the box and myself". If distance measures must be used, e.g., "three feet in front of the box", the measure can be applied to a phrase with "away from", e.g., "between ... and three feet away from the box". Other fixed points can take the place of "myself" and "me". An example would be "the box between the table and the lamp".

"Ahead of" can be treated like "in front of". "Behind" and "in back of" can be treated much the same way. However, with an object as in "the box behind the table" the box can not be related to the speaker and the table unless a phrase such as "by the side of the table farthest from me" is used. Likewise, the noun and adjectival uses of "front" and "back" can be avoided with "nearest" and "farthest", e.g., "the side nearest you" or "the side farthest from door". Because of the problem with such usages as "move the box forward", "forward", "backward", and the directional usage of "ahead" would best be replaced by "in the direction ... facing" where "I am" or "you are" replaces the dots or by "towards" followed by some fixed point as shown before. So except in a few cases the use of front/back axes terms seems best avoided.

Avoidance of the ambiguities in the top/bottom axis terms appears more difficult. Expressions like "above me" leave open an interpretation based on gravity or the longitudinal axis of the body with the speaker referring to his body as an object. Modifying terms have other problems, e.g., "higher" could be gravitationally determined or supplied by environmental frames and "upper" could be supplied by any of the possible sources.

Here if a designer could guarantee that all observers are upright then he could allow such phrases as "above me". He may also try to guarantee that gravity and environmental axes coincide, but he probably can not be sure that objects will be upright or essentially horizontal items will not have top/bottom terms applied to them on the basis of their retinal image. In these cases, the most promising approach again appears to be to relate the references to fixed points with expressions like "put the circle on the surface of the square nearest the ceiling", "put the circle next to the edge of the square farthest from me" and "look towards the ceiling".

Where the application requires the identification of paths as opposed to axes, it appears necessary to train a controller to identify uses of the path system. Phrases identifying the path can then be added to the kinds of solutions proposed for front/back terms so the system can be aware of the usage. Three examples are "he is between John and the box office in the line", "he lives between Memphis and the delta on the river" and "go along the road in the direction you are now going". Here the phrases "in the line", "on the river" and "along the road" establish the path and the phrases beginning with "between" and "in the direction" show the frame and orientation. The two path meanings of "back up" can be handled by "go back along the road" and "go back the way you came". Adding "to ..." should serve to show the goal of the movement that is intended. Phrases to identify the path and the identification of the frame where necessary can disambiguate the class of usages involved with "past", "after", "before", and "beyond", e.g., "from Minneapolis, Cairo is the last port on the river before St. Louis".

Finally the class of idioms must be identified as such and either avoided or the controller purposefully trained and the machine prepared to accept them. The cover sense of "over" and "under" can be avoided by the use of the verb "cover" or "cross", e.g., "cover the window with paper" and "the bridge crosses the river". Finally, the class of "over", "across", "beyond" and "through" can stand as is as long as the frame of reference is mentioned.

ACKNOWLEDGEMENT

The thoughtful guidance of Dr. Richard Venezky is very gratefully acknowledged, as is Jerry Shelton's careful reading of the manuscript.

## FOOTNOTES

[1]If you do not have this experience in your past, the probable reason it is lacking is the phenomenon of pointing. When people use terms like "left" and "right" they often accompany them with a physical gesture that helps to specify the reference. In the hallway situation a gesture with the hand and fingers or the eyes and head would show you where you should move. Similarly the sight of the other person starting to move in one direction would help you to deduce that the reference was to another.

If some mechanical pointing device were available, it might be included in a system as an "analogue control" to complement the language which is a "symbolic control", see Ferrell and Sheridan (1967). However, the language itself remains the primary tool for machine control. The problems that are identified here are inherent in the language and a speech-understanding system will doubtless have to confront them.

[2]There are problems with the use of this axes analogy. Some are discussed in Section V. Others include the fact that "turn left" cannot be meant the same way as "turn 90 degrees counterclockwise". If someone turns 89 degrees he will probably feel that he has turned "left". A true understanding system will need a formalism based on something like fuzzy logic (Zadeh, 1973). Also, relational phrases like "to the left of" need more than just the direct application of the axis system to show their meaning. For example in Figure 2, the box labelled I is to the left of the box labelled II. However, the left/right by which the relationship is established cannot be drawn from the center of II to the center of I. Instead, the two boxes must be projected onto the axis and then their projections can be compared.

This also leads into the realm of fuzziness since box III is formally to the left of box II according to the definition but one would like to say the relationship is less strong than the one between II and I. Nevertheless, the axis analogy appears to be sufficient to analyze the problems discussed in Sections II and III and no further analysis is made of its shortcomings here.

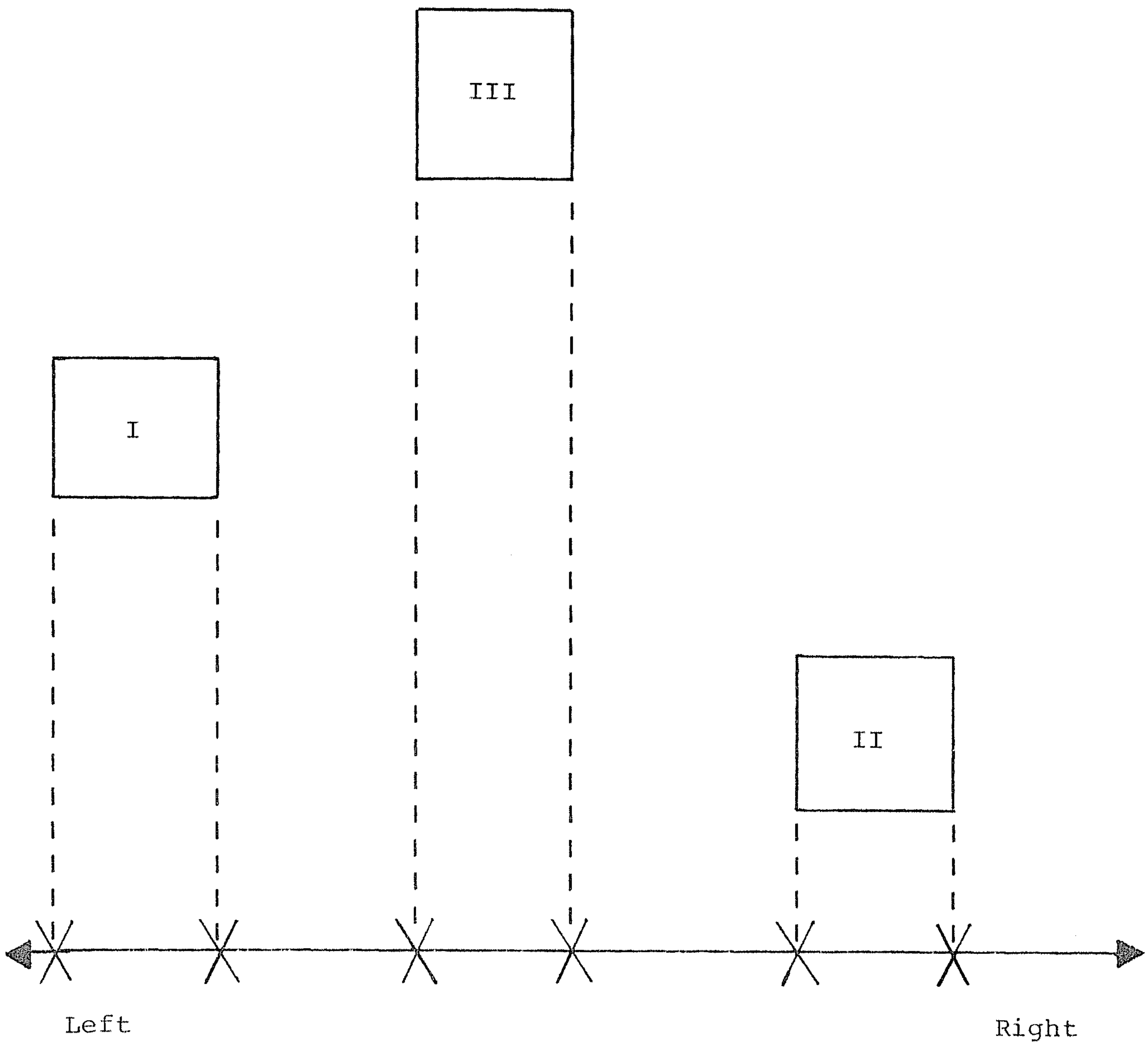



Figure 2

[3]In the linguistic literature this has been called the problem of place deixis, see especially Fillmore (1966) and Lyons (1968). That terminology is avoided here since the problems discussed in the remainder of Section II go beyond those thought of as being involved with deictic reference.

[4]Even languages that abandon the syntax of English appeal to English concepts. The well-known numerical control language APT allows the user to make liberal use of them in his references to machine tool movement. Through the commands GOLFT, GORGT, GOFWD, GOBACK, GOUP and GODOWN the programmer has access to a triple axis system. These axes are carefully defined in a natural way with the forward (FWD) direction understood as the direction in which the cutter has been moving, see Figure 3. Languages for more complex environments than the machine tool one, such as industrial robots' environments, will want to allow references to other relations. The analysis in this paper should also be helpful in their design.

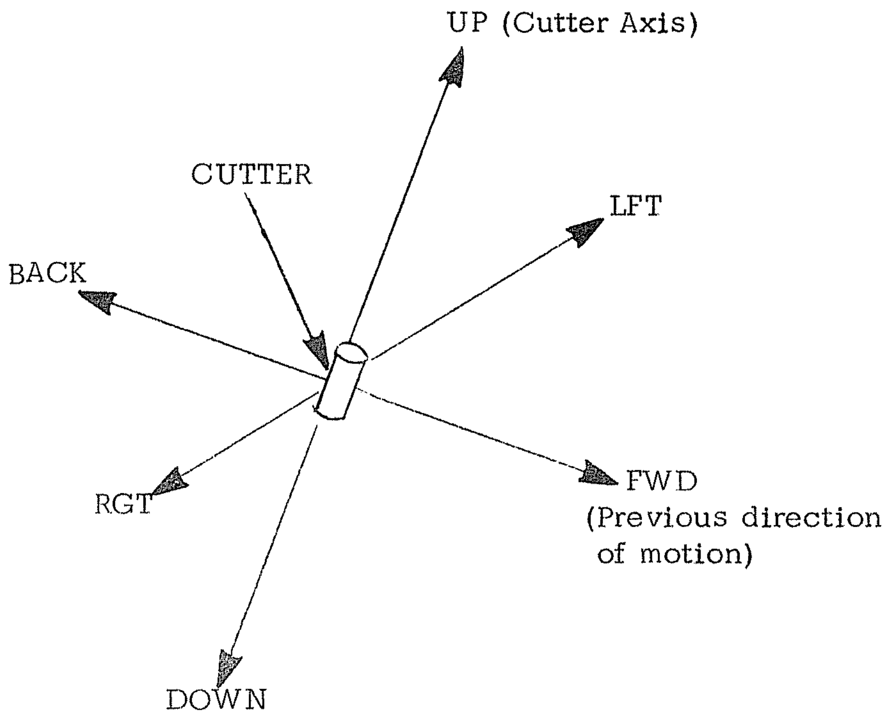

UP (Cutter Axis)
CUTTER
LFT
BACK
FWD
(Previous direction
of motion)
RGT
DOWN

Figure 3

## REFERENCES


1. Coles, L. S., "Talking with a Robot in English". Proceedings of the First International Joint Conference on Artificial Intelligence. Washington, D.C., 1969, pages 587-596.

2. Corbalis, M. C. and I. L. Beale, "On Telling Left from Right". Scientific American 224, 3 (March, 1971), pages 96-104.

3. Dobrotin, B. M. and V. D. Scheinman, "Design of a Computer Controlled Manipulator for Robot Research". Proceedings of the Third International Joint Conference on Artificial Intelligence. Stanford, 1973, pages 291-297.

4. Ferrell, W. R. and T. B. Sheridan, "Supervisory Control of Remote Manipulation". IEEE Spectrum 4, 10 (October, 1967), pages 81-88.

5. Fillmore, C. J., "Deictic Categories in the Semantics of 'Come'". Foundations of Language 2 (1966), pages 219-227.

6. Lyons, J., Introduction to Theoretical Linguistics. Cambridge at the University Press, 1968.

7. Newell, A., J. Barnett, J. Forgie, C. Green, D. Klatt, J. C. R. Licklider, J. Munson, R. Reddy and W. Woods, Speech-Understanding Systems: Final Report of a Study Group. Computer Science Department, Carnegie-Mellon University, Pittsburgh, May 1971.

8. Reddy, D. R., L. D. Erman, R. D.Fennell and R. B. Neely, "The Hearsay Speach Understanding System: An Example of the Recognition Process". Proceedings of the Third International Joint Conference on Artificial Intelligence. Stanford, 1973, pages 185-193.

9. Rock, I., "The Perception of Disoriented Figures". Scientific American 230, 1 (January, 1974), pages 78-85.

10. Rosen, C. A., "Robots, Productivity and Quality". Proceedings of the 1972 ACM National Conference. Boston, 1972, pages 47-57.

11. Rosen, C. A. et. al., Exploratory Research in Advanced Automation. First Semi-Annual Report, Grant GI-38100X, SRI project 2591, Stanford Research Institute, Menlo Park, California, December, 1973.

12. Simmons, R. F., "Answering English Questions by Computer: A Survey". Communications of the ACM 8, 1 (January, 1965), pages 53-70.

13. Simmons, R. F., "Natural Language Question-Answering Systems: 1969". Communications of the ACM 3, 1 (January, 1970), pages 15-30.

14. Walker, D. E., "Speech Understanding Through Syntactic and Semantic Analysis". Proceedings of the Third International Joint Conference on Artificial Intelligence. Stanford, 1973, pages 208-215.

15. Winograd, T., Understanding Natural Language. Academic Press, New York, 1972.

16. Woods, W. A. and J. Makhoul, "Mechanical Inference Problems in Continuous Speech Understanding". Proceedings of the Third International Joint Conference on Artificial Intelligence. Stanford, 1973, pages 200-207.

17. Zadeh, L. A., "Outline of a New Approach to the Analysis of Complex Systems and Decision Process". IEEE Transactions on Systems, Man, and Cybernetics, SMC-3, 1 (January, 1973), pages 28-44.