Sponsored by ACM/SIGCHI
with support of AAAI
and Lockheed AI Center

Architectures for Intelligent Interfaces: Elements and Prototypes

Monterey, California
March 29–April 1, 1988

Chaired by Joseph W. Sullivan and Sherman W. Tyler

Workshop Committee:

Jock Mackinlay
Robert Neches
Edwina Rissland
Norm Sondheimer

Intelligent Multi-Media Interface Technology*

Jeannette G. Neal and Stuart C. Shapiro

Calspan-UB Research Center (CUBRC)
4455 Genesee Street
Buffalo, NY 14225

Abstract

Cognitive work load is a significant factor for users of sophisticated computer systems in stressful situations such as that of military mission planning. The human-computer interface to a sophisticated system should ideally be simple and natural to use, require a minimal learning period, present output in a form that best conveys information to the user, and reduce cognitive load for the user. In pursuit of this ideal, the Intelligent Multi-Media Interface project is devoted to the development of interface technology that integrates speech, natural language text, graphics, and pointing gestures for human-computer dialogues. These dialogues are modelled on the manner in which two people are able to converse at a graphics device like a chalkboard using multiple media and modalities. The objective of the project is to develop interface technology that uses the media/modalities intelligently in a flexible, context-sensitive, and highly integrated manner. As part of the project, a knowledge-based interface system is being developed as a prototype proof-of-concept for this intelligent multi-media interface technology. The application domain being used to drive the research is a military tactical air control domain.

*This research is supported by the Defense Advanced Research Projects Agency and monitored by the Rome Air Development Center under Contract No. F30603-87-C-0136.

1. Introduction

The introduction of improved and advanced processing capabilities into Air Force Command and Control (C^2) systems is proceeding at an ever-increasing rate. It is essential that the human/machine interfaces resulting from these developments not be limiting factors which degrade the performance of the system. Many times this issue is either overlooked or handled much like a retrofit after the fact. Much of the general guidance for the introduction of decision support capabilities has been based on the overall simplified statement that the goal is to allocate information processing and decision functions between man and machine in a way which optimizes the use of their respective strengths and compensates for their respective weaknesses. Many aspects of C^2 operations are mandated to be human decision processes (e.g., nuclear enablement, fire-control, target designation). In many cases, the processes involved and the mechanization of the interface has resulted in far less than optimal performance. Considerable R&D has been initiated to support elements of this critical human/machine interface, but attention has only recently been focused on the development of intelligent multi-media interface technology.

This paper discusses an intelligent multi-media interface system, called CUBRICON (the CUBRc Intelligent CONversationalist), being developed as part of the Intelligent Multi-Media Interfaces (IMMI) project funded by DARPA and monitored by RADC. The IMMI project is devoted to the application of artificial intelligence methodology to the development of human/computer interface technology that integrates speech, natural language text, and graphics for interactive dialogues between human and computer. Using as our model two people talking at a chalkboard, we are developing the notion of a single language that incorporates speech, text, graphics, and gestures as various "syntactic" options for the conveying of information and requests. The objective is to simplify operator communication with sophisticated computer systems. The computer system accepts coordinated graphic/verbal input and composes multi-media output to best convey information to the user.

2. System Design

The CUBRICON system design is based upon a *unified view of language*. Language is a means of communication, whether verbal, visual, tactile, or gestural. Human beings communicate with each other via written and spoken natural language, frequently supplemented by pictures, diagrams, pointing to objects, and other gestures. The CUBRICON system design provides for the use of a unified multi-media language, defined by an integrated grammar, consisting of textual, graphic, and combined text/graphic symbols. Input and output streams are treated as compound streams with components corresponding to different media. This approach is intended to imitate, to a certain extent, the ability of humans to simultaneously accept input from different sensory devices (such as eyes and ears), and to simultaneously produce output in different media (such as voice, pointing motions, and drawings). The CUBRICON system includes (a) language parsing and generation to accommodate the compound streams, (b) knowledge representation and inferencing to provide reasoning ability, and (c) automated knowledge-based medium selection and formulation of responses.

The goals for this system design include the development of a clearly defined methodology of knowledge-based multi-media communication that provides the computer system with the following new functionality in interactive interface technology:

1. acceptance and understanding of multi-media input such that references to entities in a natural language sentence can be accompanied by coordinated simultaneous "pointing" to the respective entities on a graphics display; the ability to use a simultaneous pointing reference and natural language reference to disambiguate one another when appropriate.
2. automated system selection of appropriate output media/modalities for expressing information to the user with the selection performed in a flexible context-sensitive manner;
3. automated system generation of relevant output to the user in coordinated multi-media.

3. Related Research

Research and development of artificial intelligence (AI) systems for man/machine interfaces have focused on natural language (NL) text, speech, and graphics primarily in isolation, rather than in integrated interfaces. The major parsing systems include the Conceptual Analyzer at Yale University [Schank & Riesbeck81]; systems based on the augmented transition network (ATN) methodology such as LUNAR [Woods72], LIFER [Hendrix77], the RUS Parser [Bobrow78] and PSI-KLONE [Bobrow80; Sondheimer84]; the PROLOG-based parsers of Dahl [Dahl81, Dahl83], McCord [McCord85], and Warren & Pereira [Warren82]; the HPSG System [Proudian85] developed at Hewlett-Packard; frame-based parsing [Hayes84]; and the multi-strategy parsing system of Carbonell & Hayes [Carbonell87]. Major work on natural language generation is reported in [McKeown86], [Shapiro82], [Mann84], and [McDonald86].

Significant speech understanding systems were developed under DARPA sponsorship. These systems included HWIM at BBN [Woods76, Woods83], the Hearsay II [Erman81] and HARPY [Lowerre76] systems at CMU, and the Speech Understanding System at SRI [Walker78]. Ongoing Speech Understanding research is exemplified by follow-on work with HWIM at BBN [Woods83] and the development of the SPEREXSYS system at AFIT [Routh84].

Graphical display for input/output has reached a high level of sophistication in systems such as TACTICIAN from Titan Systems, Inc., the Steamer System [Stevens83], and design tools such as CAD/CAM systems.

Computer-based multi-media communication has only recently received significant attention. This research includes recent work in multimedia electronic document systems at Brown University [Feiner82], an experimental multimedia mail system at ISI [Katz84], a multimedia message system at SRI [Aceves84], and the Diamond message system at BBN [Thomas85]. Hypertext [Conklin87] provides for multi-modal multi-dimensional document representation and access.

Intelligent interactive human-computer communication via multi-media language (i.e., simultaneous natural language and graphics) has not yet been developed. Work is just beginning on intelligence in interfaces [Neches86]. Unified multi-media language (i.e., natural language and graphics) for interactive interfaces, is also just beginning to be studied. For example, the TEMPLAR system [Press86] being developed at TRW for RADC and DARPA will include natural language as an input medium with the ability to accept a pointing reference that it translates into a natural language expression and inserts into the language input stream.

The combination of intelligence with the use of a unified multi-media language for interactive interface technology distinguishes the CUBRC Intelligent Multi-Media Interfaces project from other research in interface technology. The goal for the CUBRC Interfaces project is to develop new functionality in interface technology: system acceptance of multi-media input and use of simultaneous multiple media inputs to disambiguate each other; automated selection of the media for output in a flexible context-sensitive manner; and automated composition of the output message/utterance in coordinated multi-media language. This methodology is discussed in more detail in subsequent sections of this paper.

4. System Overview

The CUBRC team has designed an integrated user interface system based on the approach described briefly in Section 2. Figure 1 provides an overview of the software system and hardware I/O devices.

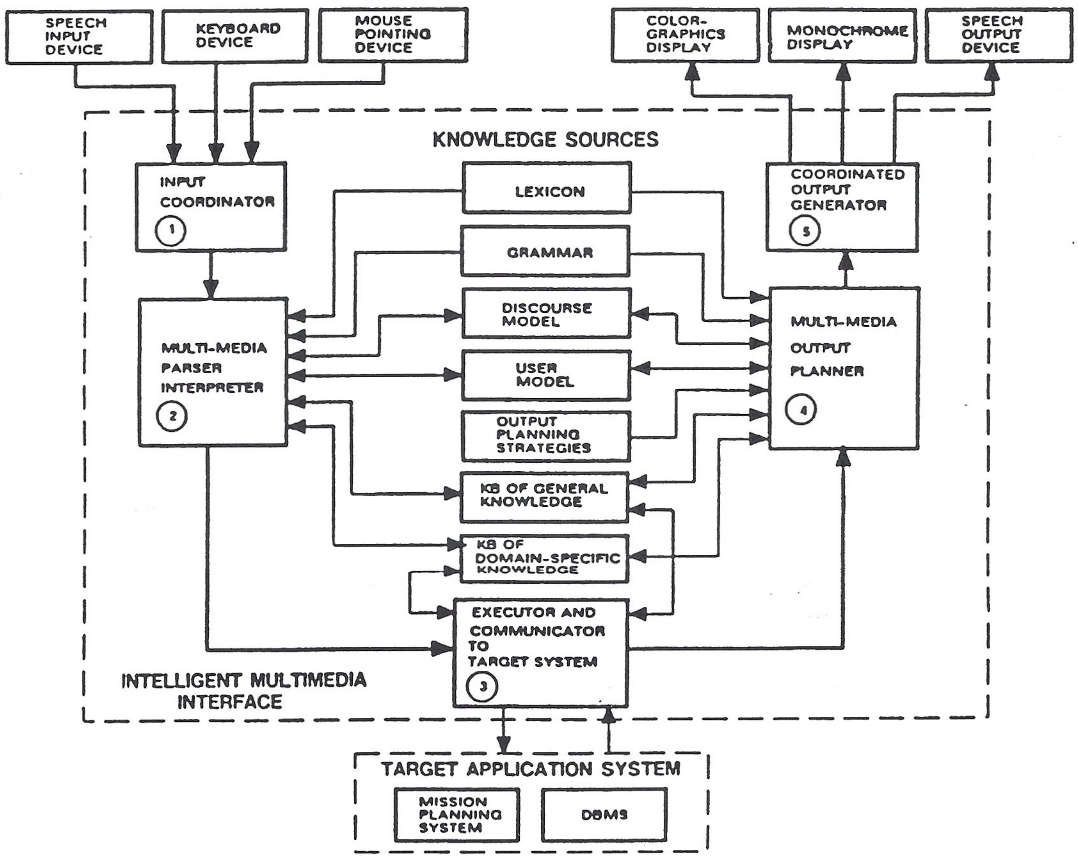


Figure 1. System Overview

The design provides for input to be accepted from three input devices: speech input device, keyboard, and mouse device pointing to objects on a graphics display. Output is produced for three output devices: color-graphics display, monochrome display, and speech output device. The primary path that the input data follows is indicated by the modules that are numbered in the figure: (1) Input Coordinator, (2) Multi-media Parser Interpreter, (3) Executor/Communicator to Target System, (4) Multi-media Output Planner, and (5) the Coordinated Output Generator. The Input

Coordinator module accepts input from the three input devices and fuses the input streams into a single compound stream, maintaining the temporal order of tokens in the original streams. The Multi-media Parser/Interpreter is an augmented transition network (ATN) that has been extended to accept the compound stream produced by the Input Coordinator and produce an interpretation of this compound stream. Appropriate action is then taken by the Executor module. This action may be a command to the mission planning system, a database query, or an action that entails participation of the interface system only. An expression of the results of the action is then planned by the Multi-media Output Planner for communication to the user. The output planner is an extended and generalized ATN that produces a multi-media output stream representation with components targeted for different devices (e.g., color-graphics display, speech device, monochrome display). This output representation is translated into visual/auditory output by the Coordinated Output Generator module. This module is responsible for producing the multi-media output in a coordinated manner in real time (e.g., the Planner module can specify that a certain icon on the color-graphics display must be highlighted when the entity represented by the icon is mentioned in the simultaneous natural language output).

The CUBRICON system includes several knowledge sources to be used during processing. The knowledge sources include: (1) a lexicon, (2) a grammar defining the language used by the system for multi-media input and output, (3) a discourse model, (4) a user model, (5) a knowledge base of output planning strategies to govern the composition of multi-media responses to the user, (6) a knowledge base of information about generally shared world knowledge, and (7) a knowledge base of information about the specific task domain of tactical air control. These knowledge sources are used for both understanding input to the system and planning/generating output from the system. They will be discussed in more detail in subsequent sections.

The prototype CUBRICON system is being developed using an extended and generalized ATN and the SNePS semantic network processing system [Shapiro79a; Shapiro81; Shapiro86]. SNePS is a fully intensional propositional semantic network and has been used for a variety of purposes and applications [Maida85; Shapiro86; Neal86, Neal87]. SNePS provides: (a) a flexible knowledge representation facility in the semantic network formalism; (b) representation of rules in the network in a declarative form so they can be reasoned about like any other data; (c) a bi-directional inference subsystem [Shapiro82] which focuses attention towards the active processes and cuts down the fan-out of pure forward or backward chaining; (d) a simulated multi-processing control structure [McKay80]; (e) special non-standard connectives [Shapiro79b] to model human reasoning processes; (f) existential, universal, and numerical quantifiers [Shapiro79c].

5. KNOWLEDGE SOURCES FOR MULTI-MEDIA LANGUAGE PROCESSING

The knowledge sources included in the CUBRICON system design for use in multi-media language understanding and production are a lexicon, grammar, discourse model, user model, a knowledge base of output planning strategies to govern the composition of multi-media responses to the user, a knowledge base of information about generally shared world knowledge, and a knowledge base of information about the task domain of tactical air control.

5.1 The Lexicon And Grammar

A lexicon is the collection of all morphemes, tokens, or signals that carry meaning in a given language. The CUBRICON system's lexicon consists of words, graphic figures, and pointing signals. The grammar defines how the morphemes, tokens, and signals of the lexicon can combine to form legal composite language structures. An example of a multi-modal language structure that is legal according to the CUBRICON grammar is a noun phrase. A noun phrase consists of the typical linguistic syntax (e.g., determiner followed by zero or more modifiers followed by a noun) accompanied by zero or more pointing signals (pointing to objects on the graphics display.) The lexicon and grammar together define the multi-modal language used by the system.

5.2 The Discourse Model

Continuity and relevance are key factors in discourse. Without these factors, people find discourse disconcerting and unnatural. The attentional discourse focus space representation [Grosz78, Grosz86; Sidner83; Grosz85] is a key knowledge structure that supports continuity and relevance in dialogue. CUBRICON tracks the attentional discourse focus space of the dialogue carried out in multi-media language and maintains a representation of the focus space in a list structure.

CUBRICON is based on the premise that graphics are an integral part of its language along with natural language and other forms of text and pointing. The CUBRICON system treats objects presented on the graphics display as having been intentionally "expressed" or "mentioned". All objects on the graphics display are "in focus" and the CUBRICON system maintains a list of all these objects as an ancillary focus space representation. This ancillary graphic focus space is also called a virtual display. CUBRICON also maintains a history of these virtual displays. This ancillary graphic focus space representation is used in a manner that is analogous to the use of the main dialogue focus space list.

The dialogue attentional focus space representation is used for determining the interpretation of anaphoric references [Sidner83] and definite descriptive references [Grosz81] uttered via natural language. In the CUBRICON system, the main dialogue focus list is consulted in determining the referent of a pronoun. In the case of a definite reference, if an appropriate referent is not found in the main dialogue focus list, then CUBRICON consults the ancillary graphic focus space representation. The motivation for this is the fact that when a person expresses a definite reference such as "the airbase" with just one such object in view (as on a graphics display) and none have been discussed, then the person most likely refers to the one in view even though he knows about several others.

We have not addressed the problem of understanding the input phrase "the airbase" when more than one is in view and none have been discussed in the human-computer dialogue. The phrase is ambiguous but a reasonable approach would be to interpret the phrase as referencing the one in view which is most relevant to the user's task.

5.3 The User Model

Many aspects of a user are highly relevant to interface technology. These aspects include level of expertise in the current task, perspective based on his role, his value system, degree and nature of impairedness due to fatigue or illness, and preferences concerning mode of communication. Carberry [Carberry87] provides a brief summary of current research on user modeling. To address all of these aspects of user modelling is, of course, beyond the scope of this project. The aspect of the user that is most relevant at this stage of CUBRICON's development is the importance rating that the user attaches to the different entity types that are relevant to a given task. We will call this the user's entity rating system.

The current CUBRICON system includes a representation of the user's entity rating system as a function of the task being addressed by the user. For a given task in the process of being carried out by the user, the entity rating system representation includes a numerical importance rating (on a scale from zero to one) assigned to each of the entity types used in the application task domain. The numerical rating assigned to a given entity type represents the degree of importance of the entity to the user. The CUBRICON design provides for the entity rating system representation to change automatically under program control when the user's task changes. In the current CUBRICON implementation the entity rating system representation is static and must be changed "by hand". Automated modification of the user entity rating system by the system as a function of the task has not been implemented yet, but is under development.

The user's entity rating system plays an important role in composing responses to the user. (1) The entity rating system representation is used in determining what information is relevant in answering questions or responding to commands from the user. (2) The entity rating system is used in selecting ancillary information to enhance or embellish the main concept being expressed and to prevent the user from making false inferences that he might otherwise make. (3) The entity rating system is also used in organizing the form in which information is presented.

As an example of (1) above, if the user instructs the system to "Display the Inner Fulda Region", CUBRICON uses the user entity rating system representation to determine what objects within the Region should be displayed. If the user is a military mission planner, then displaying all the country cottages in the region, for example, is irrelevant. The objects to display are those that are relevant to the job of the mission planner. Thus the objects that the system selects from its data base for display are airbases, missile sites, targets, etc. Section 8 discusses examples of the use of this entity rating system representation in interactive dialogue between a user and the CUBRICON system.

5.4 Knowledge Bases: General And Domain-Specific

The CUBRICON system includes knowledge bases containing general and domain-specific information. General information includes world knowledge applicable across different task domains, while domain-specific information is applicable to the particular task domain of the target information system being used as a "back end" for the interface system. Crucial information included in the knowledge bases is information concerning the presentation or expression of the entities/concepts known to the system. The knowledge base includes information on how to express an entity in its unified verbal/graphic language. This information includes the words and symbols used to express an entity, which symbols are appropriate under which conditions, and when particular colors are to be used.

6. MULTI-MEDIA LANGUAGE UNDERSTANDING

The CUBRICON system design provides for the acceptance of language consisting of coordinated "simultaneous" natural language and pointing (via a mouse device) to objects on a graphics display. The IMMI research program includes the development of a formal grammar defining the syntax of the multi-media language. The grammar is being implemented in the form of an extended and generalized ATN. The traditional ATN which takes a linear textual input stream has been modified so that it takes a multi-media input stream with components from the different input devices. Input from the devices is accepted via a polling mechanism, in order to provide the coordination information required for simulated parallelism.

The current implementation of the CUBRICON system accepts input consisting of (1) natural language without pointing, (2) pointing without natural language (e.g., to select from a menu or select an icon from a display), or (3) combined coordinated natural language with pointing to objects on a graphics display. When pointing is used in conjunction with natural language, then each noun phrase can consist of zero or more words of text along with zero or more pointing references to icons on the display (there must be at least one point or one word). The pointing input that is a component of a noun phrase can occur anywhere within the noun phrase: first token(s), within the natural language text of the noun phrase, or as the last token(s).

The use of pointing with natural language forms a very efficient means of expressing a definite reference. This enables a person to use a demonstrative pronoun as a determiner in a noun phrase and simultaneously point to an entity on the graphics display to form a succinct reference. Thus a person would be able to say "this SAM system" and point to an entity on the display to disambiguate which of several SAM systems is meant. The alternative using natural language only would be to say something like "the SAM system at 10.35 degrees longitude and 49.75 degrees of latitude" or "the SAM system just outside of Kleinburg". The combined natural language and pointing reference is clearly efficient since the natural language reference can be very short in comparison with the language-only reference and the cognitive load is reduced for the user since he does not need to generate a definite reference for the object with sufficient modification to specify it for the system.

The following sentences illustrate the types of input accepted by the CUBRICON system. These inputs presuppose that a map is displayed on the color-graphics screen with mouse-sensitive icons representing various entities. Each "<point>" represents a mouse pointing reference to an icon on the graphics display. The system's responses to such input will be discussed in Section 7 where output planning and generation is discussed.

INTERROGATIVE:

"Where is the 43rd Soviet Tank Battalion?"

"What is the mobility of this <point> SAM?"

"Is this <point> the base for these Troop Battalions <point>$_1$ <point>$_2$ <point>$_3$?"

IMPERATIVE:

"Display the Bruenn Region."

"Display the status of this <point> airbase."

Pointing references share an interesting and sometimes frustrating characteristic with natural language references: they can be ambiguous. This occurs in situations where the extents of icons on the display overlap. Figure 2 shows overlapping extents for a port facility and a population center. If the user points to the overlapping area, the point is ambiguous, having two interpretations. However, a natural language reference spoken simultaneously with the point can disambiguate the pointing reference. For example, if the phrase "this port facility" is spoken with the point to the overlapping extents of Figure 2, then the compound reference is no longer ambiguous. Our CUBRICON system correctly handles this type of situation. The natural language phrase specifies the class to which the referenced entity belongs and is used to select from the interpretations of the ambiguous pointing reference.

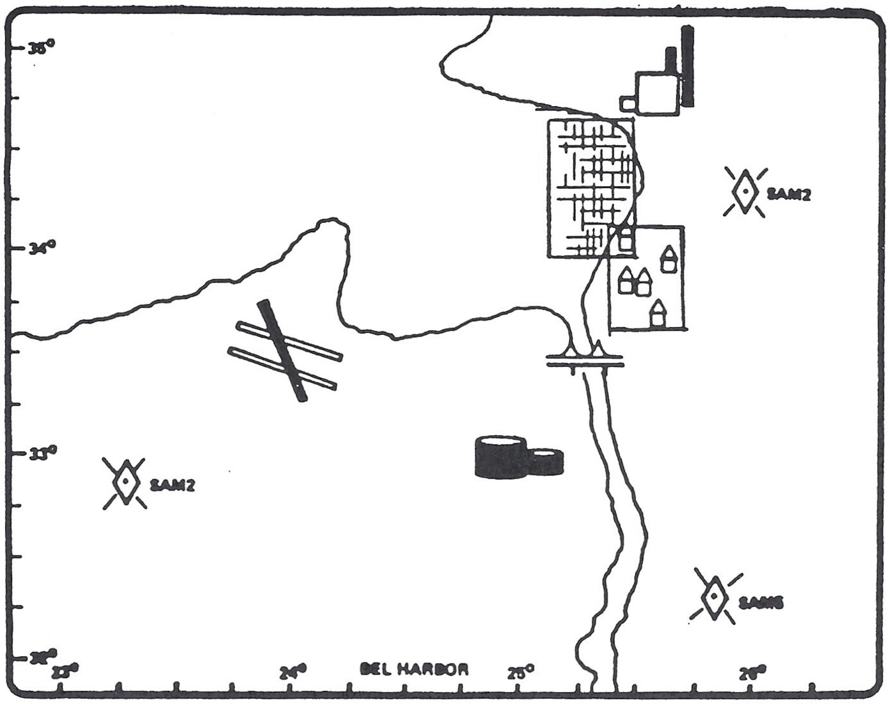


Figure 2. Rectangular Icon Extents that Overlap

Some systems use default techniques to handle ambiguous pointing. These techniques include: (1) a point returns the entity represented by the "top" or "foremost" icon where the system has a data structure it uses to remember the order in which icons are "painted" on the display (i.e., which are further in the background and which are foremost in the foreground); (2) the icons or entities are assigned weights representing importance and the icon with the largest weight is selected as the interpretation of an ambiguous point; or (3) the icon whose "center" is closest to the location pointed at is selected. Combinations of the above techniques can also be used. A serious disadvantage of these point-interpretation techniques is that it is difficult, if not impossible, for certain icons to be selected via a point reference.

The CUBRICON system handles the synergistic and efficient combination of simultaneous pointing and natural language that people very commonly and naturally use. We have deliberately not used any of the techniques described in the preceding paragraph, but instead provide the system with the ability to use coordinated natural language and pointing for mutual disambiguation. Section 8 discusses examples from a user-computer dialogue that illustrate the methodology discussed in this section.

7. PLANNING THE MULTI–MEDIA RESPONSE

The CUBRICON system design provides for output in the form of coordinated speech, text displayed on a CRT, and pictoral information on a color–graphics display. The system uses the ATN methodology for generation as well as understanding. ATNs have been used to generate text by Simmons & Slocum [Simmons72], in the XCALIBUR system [Carbonell85], and by Shapiro [Shapiro75, Shapiro82]. In this project, this ATN methodology is being generalized to generate language in the form of text, graphics, or a text/graphics combination from the semantic network knowledge bases.

A multi–media language provides a wider range of choices for the formulation of an expression for a given concept than a single–media language such as natural language used in isolation. There are alternative ways of expressing a given concept in a strictly textual language. For example, Ronald Reagan can be referred to as "Ronald Reagan", as "the President of the United States", or as "the Commander in Chief of the Armed Forces of the U.S.". If language is extended to include graphics as well, then Mr. Reagan could be referred to by pointing to his picture, by creating/displaying his image as a reference, or by a combination of text and graphics. Thus a reference can be composed of a proper noun, a noun phrase with modifiers, a graphic form with features of color, size, location on a display, etc. Furthermore, the graphic features "modifying" a form or shape can communicate meaning just as the adjectives and relative clauses do that modify a head noun of a noun phrase.

The Multi–Media Output Planner composes the response that is to be produced to the user by the Output Generator in coordinated multi–modalities. The Output Planner determines the media and modalities for expressing the response information to the user, but then must determine whether the resources are available in order to do so. If they are not, then the Planner must take appropriate action to modify the state of the resources, modify the information to be expressed, and/or select different modalities for expressing the information before the composition of the output can be accomplished.

The CUBRICON design includes the use of the following output modalities in isolation or in combinations: color–graphic/pictorial displays, tables, histograms, written natural language prose, spoken natural language, and "Fill in the blank" forms. This list does not exhaust the possibilities, of course, but provides a good varied selection with which to "prove our concept" and upon which to build.

Selection of the most appropriate modalities for expressing information in the CUBRICON system is based on the nature and characteristics of the information. Our system design is based on the premise that graphic/pictorial presentation is always desirable. The following is a brief summary of the selection criteria.

1. Color–graphics: Selected whenever the CUBRICON system knows how to represent the information pictorially.
2. Table: Selected when the values of common attribute(s) of several entities must be expressed.
3. Histogram: Selected when a quantitative attribute of several entities must be displayed in a comparative form.
4. Forms: A predefined form is selected when the task engaged in by the user requires the form.
5. Natural language prose: Selected for the expression of a proposition, relation, event, or combination thereof. The knowledge structures being expressed are heterogeneous.

Natural language can be presented in either spoken or written form. The following summarizes the selection criteria for spoken versus written language:

Spoken natural language is selected for:

- Dialogue descriptions to assist the user in comprehending the presented information. These include explanations of graphic displays or display changes and verbal highlighting of objects on the displays (e.g., "The enemy airbases are highlighted in red").
- Warnings to alert the user of important events that have or are about to take place (e.g., new critical information comes into the application system database and the system notifies user: "The XXX airbase has been damaged by enemy shellfire").
- Informing the user about the system's activity (e.g., "I'm still working" when the user must wait for output from the system).
- Short expressions of relatively non-technical information that can be remembered when presented serially (e.g., a "yes"/"no" answer to a user's question). G.A. Miller [Miller56] provides us with the guideline that short term memory can only hold approximately seven elements.

Written natural language is selected for longer technical responses that strain the user's short term memory if speech were used.

Most frequently, multiple modalities are desirable to express a body of information to the user. For example, to inform the user about the movements of a certain tank battalion, a desirable presentation might be an explanation delivered in combined spoken speech and coordinated drawing on a graphic map display showing movements of the battalion, as well as a printed textual summary with ancillary information on the monochrome display. The multiple modalities should be selected to complement and enhance one another. Andriole [Andriole86] has used "graphic equivalence" effectively using dual displays or split screens to present the same material in different forms to aid user comprehension and problem solving performance. We are not restricting the system to presenting the *same* material in different forms, but, instead, our design provides for related material or different aspects of a given event or concept to be presented in different forms/modalities (as appropriate based on the nature and characteristics of the information). We are also not restricted to graphic display presentations.

Our top level output planning process is summarized below. This planning process presupposes that the relevant information has been obtained to respond to the user.

1. For each information item or cluster, determine the modality in which it should ideally be expressed. Graphic/pictorial presentation is always desirable. Natural language can always be used, as a last resort if no other modality is available.
2. Determine whether the resources are available to express the information as desired. Resources: (1) Color graphics display: Are the items to be expressed graphically already on the color display (e.g., objects of interest in a geographical domain may already be displayed on a map)? If so, no additions are necessary. If not, is there room to add them in their "natural" position? (e.g., can the desired objects be inserted in the area already on the graphics display without changing the area shown or does the displayed area need to be extended or changed totally?) (2) Monochrome display: Similar to the color graphics display. (3) Speech output device: Always available.
3. If the desired resources are not available, modify the state of the resources. The desired resources would be "not available" in at least two cases: (1) the physical device is not

functional (e.g., needs repair) or (2) the device (e.g., a display) already contains critical information that cannot be disrupted nor covered by a window.

Graphics display: If not all the items to be expressed graphically are on the graphics display, then the system must compose a new display. Borrowing terminology from the geographical situation, the possible cases are:

a. "Zoom out" with intelligent addition of required and relevant ancillary objects to fill in the new area and possible deletion of previously present objects to prevent overcrowding of the display.
b. "Zoom in" with intelligent addition of required and relevant objects to create an intelligible display.
c. "Pan" to a different area with intelligent addition and deletion of objects.
d. Combination of the above.
e. Display a different disjoint area. (i) Completely replace display with new "area" or (ii) Open a window on the display to show new information.

4. If the desired resources are still not available to accommodate the information to be expressed, try modifying the information to be expressed: trim the amount of information by filtering on the basis of relevance with regard to user model, task model, and/or discourse model.
5. If the information can still not be expressed in the given modality due to insufficient resources for the selected modality, then select another modality and go back to step 2.
6. Compose the output, having resolved resource constraints.

The methodology discussed above for selecting output modalities and solving resource constraints has not been implemented yet, but its development is underway. Some of the lower level output composition routines for natural language, map graphics, and tables have been implemented and are undergoing further development and refinement. The next section discusses working examples from a dialogue with the system. The examples illustrate some of the functionality of the system in its current state of implementation.

8. EXAMPLE DIALOGUE

In this section we present short sample dialogues to illustrate the functionality and processing discussed in the previous sections. The dialogues are concerned with mission planning and situation assessment in a tactical air control domain.

Consider the following initial user input to the CUBRICON system:

"Display the Inner Fulda region."

The planning and composition of output for the user is dependent upon the nature of the information, the discourse context, and user model. Since this example command is the initial input, the process is simpler than otherwise. The CUBRICON system knows that a region can be represented graphically and therefore chooses graphics as the primary modality for display. Regions are represented in the CUBRICON knowledge base with an associated boundary. The boundary is retrieved by the system and the main roads, major cities, waterways, and national borders are displayed on the color graphics display. These items are displayed by use of the MAP Display System [Hilton87] developed by Lt. Mike Hilton at RADC. The CUBRICON system then searches its knowledge base for task-specific objects within the region that should be displayed. The selection of these objects is based on the user model discussed in Section 5. This user model consists of the user's entity rating system: a task-dependent assignment of importance ratings to the objects in the task domain. Objects are selected for display that are above a certain criticality threshold for importance. The system does not display all entities that it knows about in the region, but only those that are above the criticality threshold for the particular task in which the user is engaged (a sub-task of planning an air-strike mission). Thus the system decides to display all airbases, surface-to-air missile (SAM) sites, fuel storage sites, critical factories and plants, but not objects such as food production plants or minor industry that are not germane to mission planning. The resulting color map display is shown in Figure 3.

Based on the information provided by the user model, the system knows the important attributes of each object. Since these attributes are not displayed or communicated via the map display but have been selected as relevant, the system must determine a modality for presenting this information. Since this information consists of lists of objects with different values of common attributes, the system determines that a table presentation is appropriate. The system composes a table showing the most important attributes of the most important objects. The number of attributes and the number of objects is constrained by available space. The resultant table is displayed on the monochrome display and is shown in Figure 3. As stated in Section 7, the top level decision-making process for determining the appropriate modality for an information item or cluster has not been implemented yet, but is under development. The map composition and table composition processes are implemented and undergoing refinement and enhancement.

As discussed in Section 6, the CUBRICON system accepts natural language input with coordinated pointing to objects on the graphics display. We now provide some examples of this multi-media input. The user can input:

"What is this <point>?"

where the <point> is a point to an icon on the color graphics display using the mouse device. The system replies:

"It is a munitions factory."

To illustrate the system's ability to disambiguate a pointing reference, the user inputs:

"What is the mobility of this <point> SAM?"

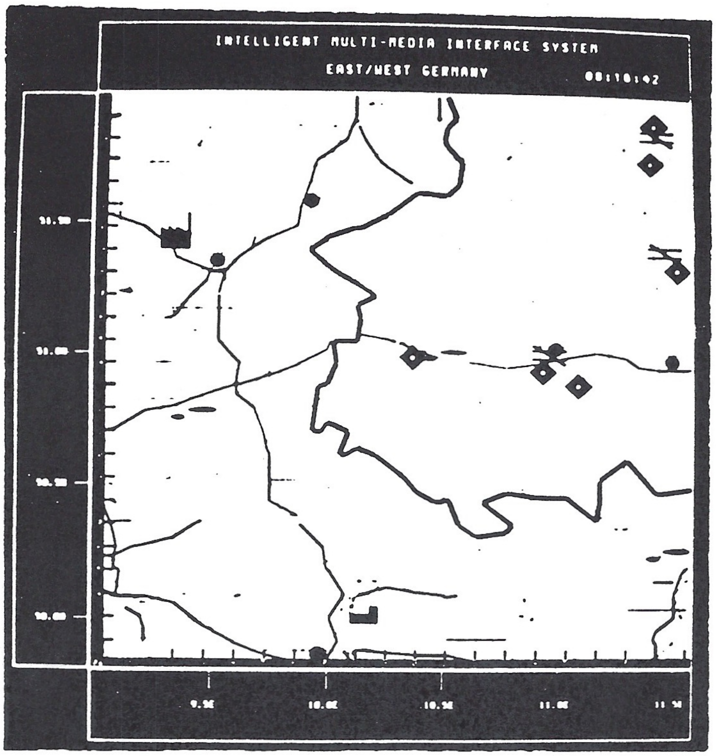
where the <point> is a mouse click on the area where the extents of two icons overlap as indicated by the arrow in Figure 3. As a result, the mouse handler returns a representation of each of the two objects. In determining the referent of the noun phrase "this <point> SAM", CUBRICON determines that the descriptor "SAM" applies to only one of the two objects returned by the mouse handler. More specifically, "SAM" identifies a super-category of the one object. The other object is eliminated since it is not in any sub-category of SAMs. The system then retrieves the particular SAM's mobility from the knowledge base and responds:

"The mobility of this SA-3 is high."

In order to further illustrate the CUBRICON system's modality selection and output composition process, consider this next user input:

"Show the Dresden airbase." (or "Where is the Dresden airbase?")

As discussed in Section 7, whenever possible the CUBRICON system prefers to present information to the user graphically with ancillary information presented simultaneously in an another modality. Since CUBRICON knows how to display an airbase graphically (it has an icon associated with the class in the knowledge base), and since each particular airbase in the knowledge base has an



Inner Fulda Region

Item	Name	Longitude	Latitude	Disposition
fuel storage tank	heine	11.450E	50.230N	friendly
fuel storage tank	rhein	9.400E	50.750N	friendly
fuel storage tank	haina	10.480E	50.980N	enemy
munition factory	fritz munitions	10.152E	49.991N	friendly
steel plant	zarogoza steel	9.344E	51.421N	friendly
fighter base	erfurt	10.960E	51.000N	enemy
fighter base	cochsted	11.430E	51.860N	enemy
fighter base	allstedt	11.460E	51.400N	enemy
SA-6	--	10.933E	51.133N	enemy
SA-3	--	11.400E	51.750N	enemy
SA-3	--	10.366E	50.983N	enemy
SA-3	--	11.516E	51.333N	enemy
SA-2	--	10.933E	50.933N	enemy
SA-2	--	11.083E	50.883N	enemy
SA-2	--	11.416E	51.900N	enemy

Figure 3. The Displays Composed by the System

associated geographical location, then the system will display the airbase on the color-graphics map with additional information displayed in another modality. If the Dresden airbase were already displayed on the color map display, then the system would choose to blink the particular airbase icon as its way of pointing to the object and accompany this pointing action with a textual response. If the Dresden airbase could be added to the current map, it would do so and blink the airbase icon as mentioned above. However, the Dresden airbase is outside of the region shown in the map display currently on the color CRT. Therefore the system must now decide how to show the airbase. What map should be displayed?

In composing a new map on which to display the Dresden airbase, the system has some choices. These choices include: open a window on the color graphics display showing the area around the Dresden airbase, replace the old map on the CRT with a new area around the Dresden airbase, or compose a new map including both the old map and the region around the Dresden airbase.

An important guideline to which the CUBRICON system tries to adhere is to maintain the context of the user/computer dialogue. With regard to the graphic displays, this means that the system tries to retain the most recently discussed or mentioned objects on the displays so as to maintain continuity in the dialogue. The discourse focus space representations, discussed in Section 5 are the key knowledge sources in this process. The system composes a new map containing the objects that are on the old map as well as the Dresden airbase. The algorithm that the system uses to determine the boundary for a new map of this type is to determine the smallest rectangle that encloses the old objects on the current map as well as the new objects to be displayed and then add a small "border" area around all sides. This essentially extends the area shown to include both the old and new objects.

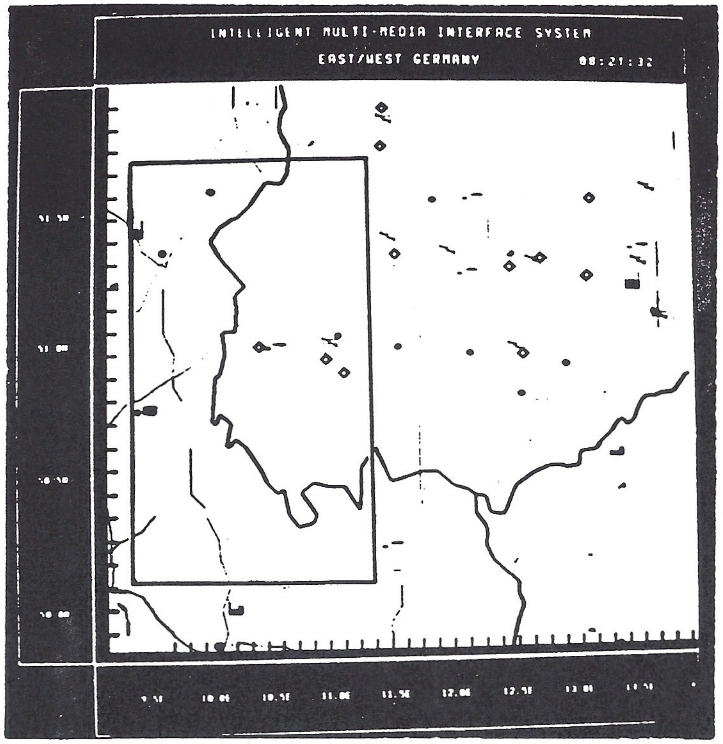
A key feature of the CUBRICON system is that it is able to display any rectangular region within a "master map" boundary containing the East-West Germany region. That is, the system can select any degree of longitude for the eastern or western boundary of the map to be displayed and any degree of latitude for the northern or southern boundary (within the confines of the master map). For any such map, as indicated previously, the non-domain-specific items such as roads, major cities, waterways, and national boundaries are displayed by Hilton's MAP Display System.

Another important guideline to which the CUBRICON system adheres is to maintain consistency throughout a display so as to prevent the user from making false inferences about what is or is not located within the region. In the case of our map display, this means that there should be consistency in the types of objects shown across the entire map. If SAMs are displayed in the old region, then they should be displayed in the newly added map area. Similarly for other types of objects. If this is not done, then the user would probably infer that there were no SAMs in the new area since he sees none on the display in the new area, when in reality there are SAMs in the new area. Figure 4 shows the new map display composed by the CUBRICON system in response to the user's input "Display the Dresden airbase." The rectangular outline within the map is used to indicate the previously displayed area. This provides graphic context: the new entities in the context of the previously displayed area. Guided by the consistency principle, the system also modifies the tabular presentation that is on the monochrome display. Both displays are also shown in Figure 4.

In our demo script, the user then asks:

"Where is the Lindsey airbase?"

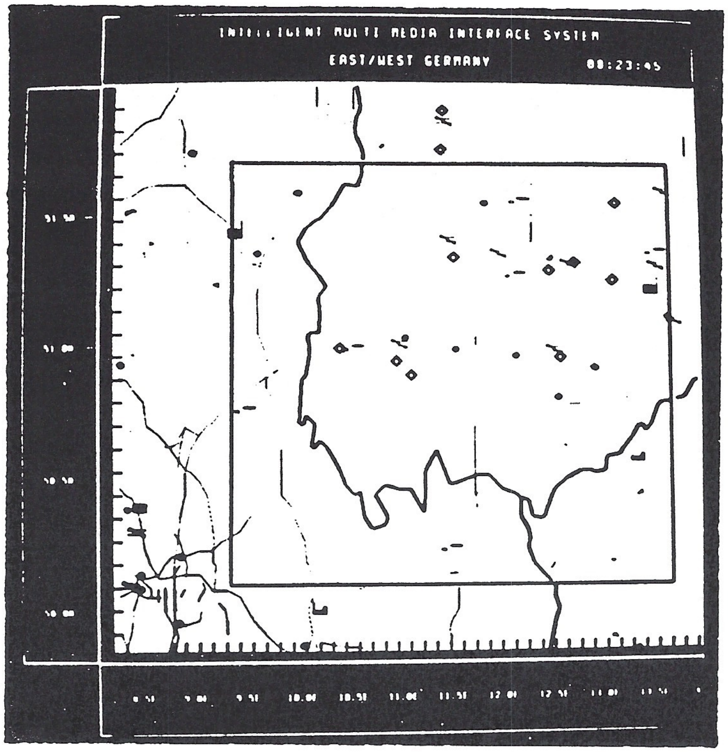
The CUBRICON system responds with the new map and table shown in Figure 5.



Expanded Region

Item	Name	Longitude	Latitude	Disposition
fuel storage tank	zollschen	12.100E	51.260N	enemy
fuel storage tank	welzow	12.150E	51.550N	enemy
fuel storage tank	bautzen	13.500E	51.360N	enemy
munition factory	schwartz munitions	13.320E	50.561N	enemy
steel plant	gallschutz steel	13.480E	51.211N	enemy
fighter base	merseberg	11.960E	51.350N	enemy
fighter base	grossenhain	13.530E	51.310N	enemy
fighter base	finsterwalde	13.610E	51.600N	enemy
fighter base	dresden	13.700E	51.100N	enemy
fighter base	brandis	12.660E	51.310N	enemy
fighter base	attenberg	12.510E	50.980N	enemy
SA-6	--	12.433E	51.350N	enemy
SA-6	--	13.683E	51.216N	enemy
SA-4	--	12.955E	51.473N	enemy
SA-4	--	12.016E	51.633N	enemy
SA-4	--	11.916E	51.450N	enemy
SA-3	--	13.100E	51.250N	enemy
SA-3	--	12.566E	50.950N	enemy
SA-2	--	12.466E	51.283N	enemy
SA-2	--	13.133E	51.550N	enemy
SA-2	--	12.716E	51.316N	enemy
fighter base	dresden	13.700E	51.100N	enemy
SA-2	--	11.416E	51.900N	enemy
SA-2	--	11.083E	50.883N	enemy
SA-2	--	10.933E	50.933N	enemy
SA-3	--	11.516E	51.333N	enemy
SA-3	--	10.366E	50.983N	enemy
SA-3	--	11.400E	51.750N	enemy
SA-6	--	10.933E	51.133N	enemy
fighter base	allstedt	11.460E	51.400N	enemy
fighter base	cochsted	11.430E	51.860N	enemy
fighter base	erfurt	10.960E	51.000N	enemy
steel plant	zarogoza steel	9.344E	51.421N	friendly
munition factory	fritz munitions	10.152E	49.991N	friendly
fuel storage tank	haina	10.480E	50.980N	enemy
fuel storage tank	rhein	9.400E	50.750N	friendly
fuel storage tank	heine	11.450E	50.230N	friendly

Figure 4. Map and Table Maintaining Context and Consistency



Expanded Region

Item	Name	Longitude	Latitude	Disposition
fuel storage tank	bliss	8.340E	50.390N	friendly
fighter base	lindsey	8.330E	50.100N	friendly
fighter base	rhein main	8.390E	50.300N	friendly
fuel storage tank	heine	11.450E	50.230N	friendly
fuel storage tank	rhein	9.400E	50.750N	friendly
fuel storage tank	haine	10.480E	50.980N	enemy
munition factory	fritz munitions	10.152E	49.991N	friendly
steel plant	zarogoza steel	9.344E	51.421N	friendly
fighter base	erfurt	10.960E	51.000N	enemy
fighter base	cochsted	11.430E	51.860N	enemy
fighter base	allstedt	11.460E	51.400N	enemy
SA-6	--	10.933E	51.133N	enemy
SA-3	--	11.400E	51.750N	enemy
SA-3	--	10.366E	50.983N	enemy
SA-3	--	11.516E	51.333N	enemy
SA-2	--	10.933E	50.933N	enemy
SA-2	--	11.083E	50.883N	enemy
SA-2	--	11.416E	51.900N	enemy
fighter base	dresden	13.700E	51.100N	enemy
SA-2	--	12.716E	51.316N	enemy
SA-2	--	13.133E	51.550N	enemy
SA-2	--	12.466E	51.283N	enemy
SA-3	--	12.566E	50.950N	enemy
SA-3	--	13.100E	51.250N	enemy
SA-4	--	11.916E	51.450N	enemy
SA-4	--	12.016E	51.633N	enemy
SA-4	--	12.955E	51.473N	enemy
SA-6	--	13.683E	51.216N	enemy
SA-6	--	12.433E	51.350N	enemy
fighter base	altenberg	12.510E	50.980N	enemy
fighter base	brandis	12.660E	51.310N	enemy
fighter base	dresden	13.700E	51.100N	enemy
fighter base	finsterwalde	13.610E	51.600N	enemy
fighter base	grossenhain	13.530E	51.310N	enemy
fighter base	merseberg	11.960E	51.350N	enemy
steel plant	gallschutz steel	13.480E	51.211N	enemy
munition factory	schwertz munitions	13.320E	50.561N	enemy
fuel storage tank	bautzen	13.500E	51.360N	enemy
fuel storage tank	welzow	12.150E	51.550N	enemy
fuel storage tank	zollschen	12.100E	51.260N	enemy

Figure 5. Newly Composed Map and Table Displays

In this section we have discussed examples to (a) illustrate some of the key functionality of the CUBRICON system, (b) show how some of the important CUBRICON knowledge sources are used, and (c) show how the important guidelines are applied by the system. The functionality that we illustrated includes the ability of the system to accept and understand a multi-media noun phrase that consists of a textual noun phrase accompanied by a pointing reference; use natural language to disambiguate the accompanying pointing reference; select the modalities for presentation of information to the user; select relevant information to present to the user; compose the multi-media presentations, particularly map graphics and tables. The CUBRICON knowledge sources that were used in the examples of this section include the nature and characteristics of the information to be expressed, the discourse model, user model, and task model. The guidelines discussed in the context of the examples were the maintain-context guideline and the consistency guideline.

9. CURRENT STATUS AND FUTURE DIRECTION

The CUBRICON project is fairly young and, as a result, our system implementation is not complete. The system includes a modest grammar for multi-media language understanding and generation. The ability to handle natural language with coordinated pointing references has been implemented. Similarly, the ability to use natural language to disambiguate a coordinated pointing reference has been implemented. The knowledge sources discussed in Section 5 have been implemented with the qualifications mentioned in Section 5. The system's process for selecting the media/modalities for generation of responses to the user is under development and has not yet been implemented. The system's process of intelligently composing new map regions and selecting relevant objects to display has been implemented.

The CUBRICON system is being implemented on a Symbolics Lisp Machine with both color and monochrome displays. A speech recognition device is now being added to the hardware suite and a speech production device will subsequently be added.

The CUBRICON team is continuing its development of the processes discussed in this paper. These include the system's process of (1) selecting the appropriate media/modalities for expressing responses to the user; and (2) determining whether the desired resources are available and, if not, either modifying the state of the resources or the information to be expressed or selecting different modalities for expressing the information. The team is also continuing development of the system's ability to compose additional forms of multi-media output.

10. SUMMARY

This paper has discussed the architecture of the CUBRICON system and its components, key functionality of the system, the important knowledge sources used by the system, multi-media language understanding, the system's process for composing multi-modal output, examples to illustrate some of the system's key functionality and processing methodology, and the current status and future direction of the CUBRICON project.

Key functionality of the system includes the ability to handle simultaneous input in natural language and graphics pointing, with language to disambiguate pointing and conversely, automated selection by the interface system of media for output generation, and automated composition of multi-media responses to be conveyed to the user.

The knowledge sources used by the CUBRICON system are the nature and characteristics of the information, discourse attentional focus space model, user model, task model, guidelines for enhancing human comprehension, and the constraints or limitations of the hardware input and output devices.

The system's processing of multi-media input was discussed with emphasis on the system's ability to use the natural language to disambiguate pointing to objects on a graphics display and conversely.

The system's output composition process includes selection of appropriate modalities and media, determination of whether resources are available, subsequent modification of resources or information to be expressed (if necessary), modification of selected output modalities/media (if necessary), and composition of the output.

Examples were presented to illustrate some of the key functionality of the CUBRICON system, to show how the CUBRICON knowledge sources are used, and to show how two important guidelines for enhancing human comprehension are applied by the system.

This paper discussed the Intelligent Multi-Media Interfaces Project. This project is devoted to the development of interface technology that integrates speech, natural language text, graphics, and pointing gestures for human-computer dialogues. The objective of the project is to develop interface technology that uses the media/modalities intelligently in a flexible, context-sensitive, and highly integrated manner.

11. ACKNOWLEDGEMENTS

The other members of the CUBRC team are: K. Bettinger, S. Byoun, Z. Dobes, D. Funke, M. Summers, C. Thielman.

The MAP Display System was developed by Lt. Mike Hilton (RADC), who graciously allowed us to use it.

12. REFERENCES

[Aceves84] Aceves, J.J.G., Poggio, A. & Eliott, D. 1984. Research into multimedia message system architecture, Final Report. *Proj. No. 5363*, SRL International, Menlo Park, CA.

[Andriole86] Andriole, S.J. 1986. Graphic equivalence, graphic explanations, and embedded process modeling for enhanced process modeling for enhanced user-system interaction. *IEEE Trans. on Systems, Man, and Cybernetics*, Vol. 16, No. 6.

[Bobrow78] Bobrow, R.J. 1978). The RUS system, BBN Report No. 3878.

[Bobrow80] Bobrow, R.J. & Webber, B. 1980. Knowledge representation for syntactic/semantic processing, *Proc. AAAI-80*, pages 316-323.

[Carberry87] Carberry, S. 1987. First international workshop on user modeling, *AI Magazine*, Vol.8, No.3, pages 71-74.

[Carbonell85] Carbonell, J.G., Boggs, W.M., Mauldin, M.L., & Anick, P.G. 1985. An integrated natural language interface, in S. Andriole (ed.) *Applications in Artificial Intelligence* Petrocelli Books, Inc., pages 227–245.

[Carbonell87] Carbonell, J.G. & Hayes, P.J. 1987. Robust parsing using multiple construction-specific strategies, in *Natural Language Parsing Systems*, L. Bolc (ed.), Springer–Verlag Pub., pages 1–32.

[Conklin87] Conklin, J. 1987. Hypertext: An introduction and survey, *Computer*, Vol. 20, No. 9, pages 17–41.

[Dahl81] Dahl, V. 1981. Translating spanish into logic through logic, *AJCL*, Vol.7, No.3, pages 149–164.

[Dahl83] Dahl, V. & McCord, M.C. 1983. Treating coordination in logic grammars, *American Journal for Computational Linguistics*, Vol. 9, No. 2, pages 69–91.

[Erman81] Erman, L.D., Hayes–Roth, F., Lesser, V.R., & Reddy, D.R. 1981. The Hearsay–II speech–understanding system: Integrating knowledge to resolve uncertainty, in B. Webber & N. Nilsson (eds), *Readings in Artificial Intelligence*, pages 349–389.

[Feiner82] Feiner, S., Nagy, S. & van Dam, A. 1982. An experimental system for creating & presenting interactive graphical documents, *Trans. Graphics*, Vol. 1, No. 1, pages 59–72.

[Grice75] Grice, H.P. 1975. Logic and conversation, in P. Cole & J.L. Morgan (eds.), *Syntax and Semantics, Vol. 3: Speech Acts*, Academic Press, pages 41–48.

[Grosz78] Grosz, B. J. 1978. Discourse analysis, in D. Walker (ed.), *Understanding Spoken Language*, Elsevier North–Holland, New York, pages 229–345.

[Grosz81] Grosz, B.J. 1981. Focusing and description in natural language dialogues, *Elements of Discourse Understanding*, A.Joshi, B. Webber, & I.Sag (eds.), Cambridge Univ. Press, pages 84–105.

[Grosz85] Grosz, B.J. & Sidner, C.L. 1985. Discourse structure and the proper treatment of interruptions, *Proc. of IJCAI*, pages 832–839.

[Grosz86] Grosz, B.J. 1986. The representation and use of focus in a system for understanding dialogs," in *Readings in Naatural Language Processing*, B.J. Grosz, K.S. Jones, B.L. Webber (eds.), Morgan Kaufmann Pulishers, pages 353–362.

[Hayes84] Hayes, P.J. 1984. Entity oriented parsing, *Proceedings of COLING84*, Stanford University, pages 212–217.

[Hendrix77] Hendrix, G.G. 1977. The LIFER manual, Technical Note 138, Artificial Intelligence Center, SRI International.

[Hilton87] Hilton, M.L. 1987. Design and implementation of the MAP DISPLAY SYSTEM, RADC REPORT.

[KATZ84] Katz, A. 1984. An experimental internetwork multimedia mail system, *Proc. IFIP 6.5 Working Conf. Computer Message Services*, Nottingham, England.

[Lowerre76] Lowerre, B.T. 1976. The HARPY speech recognition system, Ph.D. Thesis, C.S. Dept., Carnegie–Mellon University.

[Maida85] Maida, A.S. & Shapiro, S.C. 1985. Intensional concepts in propositional semantic networks," in *Readings in Knowledge Representation*, R.J. Brachman & H.J. Levesque (eds.), Morgan Kaufmann Pub., pages 169–190.

[Mann84] Mann, W.C. 1984. Discourse structures for text generation, *Proceedings of COLING84*, Stanford University, pages 367–375.

[McCord85] McCord, M.C. 1985. Modular logic grammars, *Proc. ACL*, pages 104–117.

[McDonald86] McDonald, D. 1986. Description directed control: its implications for natural language generation, in *Readings of Natural Language Processing*, B.J. Grosz, K.S. Jones, B.L. Webber (eds.), Morgan Kaufmann Pulishers, pages 519–537.

[McKay80] McKay, D.P. & Shapiro, S.C. 1980. MULTI – A LISP based multiprocessing system, *Conference Record of the 1980 LISP Conference*, Stanford Univ., pages 29–37.

[McKeown86] McKeown, K. 1986. Discourse strategies for generating natural–language text, in *Readings in Natural Language Processing*, B.J. Grosz, K.S. Jones, B.L. Webber (eds.), Morgan Kaufmann Pulishers, pages 479–499.

[Miller56] Miller, G.A. 1956. The magical number seven plus or minus two, *Psychological Review*, 63, pages 81–97.

[Neal86] Neal, J.G. & Shapiro, S.C. 1986. Knowledge representation for reasoning about language, in *The Role of Language in Problem Solving*, J.C. Boudreauz, B.W. Hamill, & R. Jernigan (eds.), Springer–Verlag Pub., pages 27–47.

[Neal87] Neal, J.G. & Shapiro, S.C. 1987. Knowledge based parsing, in *Natural Language Parsing Systems*, L. Bolc (ed.), Springer-Verlag Pub., pages 49-92.

[Neches86] Neches, R. & Kaczmarek, T. 1986. *AAAI-86 Workshop on Intelligence in Interfaces*, USC/Information Sciences Institute, August, 1986.

[Press86] Press, B. 1986. The U.S. Air Force TEMPLAR project status and outlook, *Western Conf, on Knowledge-Based Engineering and Expert Systems*, Anaheim, CA, pages 42-48.

[Proudian85] Proudian, D. & Pollard, C. 1985. Parsing head-driven phrase structure grammar, *Proc. ACL*, pages 167-171.

[Routh84] Routh, R.L. & Milne, R.W. 1984. Continuous speech recognition using natural language constraints, *Proc NAECON-84*, pages 916-23.

[Schank81] Schank, R.C. & Riesbeck, C. 1981. *Inside Computer Understanding*, Lawrence Erlbaum Associates.

[Shapiro75] Shapiro, S.C. 1975. Generation as parsing from a network into a linear string, *AJCL*, Microfiche 33, pages 45-62.

[Shapiro79a] Shapiro, S.C. 1979a. The SNePS semantic network processing system, in Findler, ed., *Associative Networks - The Representation and Use of Knowledge by Computers*, Academic Press, New York, pages 179a-203.

[Shapiro79b] Shapiro, S.C. 1979b. Using non-standard connectives & quantifiers for representing deduction rules in a semantic network, invited paper presented at Current Aspects of AI Research, a seminar held at the Electrotechnical Laboratory, Tokyo.

[Shapiro79c] Shapiro, S.C. 1979c. Numerical quantifiers and their use in reasoning with negative information, *Proc. IJCAI*, pages 791-796.

[Shapiro81] Shapiro, S.C. & the SNePS Implementation Group 1981. *SNePS User's Manual*, C.S. Dept., SUNY at Buffalo, N.Y.

[Shapiro82a] Shapiro, S.C. 1982. Generalized augmented transition network grammars for generation from semantic networks, *AJCL*, Vol. 8, No. 1, pages 12-25.

[Shapiro82b] Shapiro, S.C. & Martens, J., & McKay, D. 1982. Bi-directional inference, *Proc. of the Cognitive Science Society*, pages 90-93.

[Shapiro82c] Shapiro, S.C. & Neal, J.G. 1982. A knowledge engineering approach to natural language understanding, *Proc. ACL*, pages 136–144.

[Shapiro86] Shapiro, S.C. & Rapaport, W. 1986. SNePS considered as a fully intensional propositional semantic network, *Proceedings of AAAI-86*, pages 278–283; in G. McCalla & N. Cercone (eds.) *Knowledge Representation*, Springer–Verlag Pub.

[Sidner83] Sidner, C.L. 1983. Focusing in the comprehension of definite anaphora, in M. Brady & R.C. Berwick (eds.), *Computational Models of Discourse*, The MIT Press, pages 267–330.

[Simmons72] Simmons, R. & Slocum, J. 1972. Generating English discourse from semantic networks, *CACM*, 15:10, pages 891–905.

[Sondheimer84] Sondheimer, N.K., Weischedel, R.M. & Bobrow, R.J. 1984. Semantic interpretation using KL–ONE, *Proc. of ACL-84*, pages 101–107.

[Stevens83] Stevens, A., Roberts, B., & Stead, L. 1983. The use of a sophisticated graphics interface in computer–assisted instruction, *IEEE Computer Graphics and Applications*, pages 25–31.

[Thomas85] Thomas, R.H., Forsdick, H.C., Crowley, T.R., Schaaf, R.W., Tomlinson, R.S., Travers, V.M., & Robertson, G.G. 1985. Diamond: a multimedia message system built on a distributed architecture, *IEEE Computer*, December, pages 65–78.

[Walker78] Walker, D.E. 1978. *Understanding Spoken Language*, Elsevier North–Holland.

[Warren82] Warren, D.H.D. & Pereira, F.C.N. 1982. An efficient easily adaptable system for interpreting natural language queries, *AJCL*, Vol. 8, No. 2, pages 97–119.

[Woods72] Woods, W.A., Kaplan, R. & Nash–Webber, B. 1972. The lunar sciences natural language information system: Final report, BBN Rep. No. 2378, Bolt Beranek & Newman, Inc.

[Woods76] Woods, W.A., et al. 1976. *Speech Understanding Systems: Final Report*, BBN Report No. 3438.

[Woods83] Woods, W.A. 1983. Language processing for speech understanding, BBN Report No. 5376.