Interset: A Natural Language Interface for Teleoperated Robotic Assembly of the EASE Space Structure

Daniel K. Boorsma

Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Abstract

A teleoperated robot was used to assemble the EASE (Experimental Assembly of Structures in Extra-vehicular activity) space structure under neutral buoyancy conditions, simulating a telerobot performing structural assembly in the zero gravity of space. This previous work carried out at NASA and MIT used a manually controlled teleoperator as a test bed for system performance evaluations. From these results several Artificial Intelligence options were proposed. One of these was further developed into a real-time assembly planner written in Prolog.The interface for this system is effective in assembling EASE structures using windowed graphics and a set of networked menus. As the problem space becomes more complex and hence the set of control options increases, a natural language interface may prove to be beneficial to supplement the menu based control strategy. This strategy can be beneficial in situations such as: describing the local environment, maintaining a data base of task event histories, modifying a plan or a heuristic dynamically, summarizing a task in English, or operating in a novel situation.

Interset is conceptually a natural language extension to an assembly planner approach, it attempts to map a sentence into a "speech act network", the representation emphasizing spatial and topological relations about work pieces and key processes.There are four general system goals:1) provide a natural language process capable of verbal dialog about the telerobot and EASE structures, 2) maintain a database of control events with respect to assembly conditions over the lifetime of the construction project such as, the list of completed subtasks, 3) update the systems contents to reflect a changing world model, and 4) modify or test system parameters more rapidly such as avoid using a sensor and observe the resulting degraded system. Currently, the prototype is written in Scheme (lisp) and implemented using a VAXstation 2000, IBM RT, and MacIntosh II workstations. Interset consists of the following structures: a 'slightly' grammatical parser, a deductive assertion maintenance system, rule application control, event and time referencing via a filter, perspective stereo graphics, several dozen small knowledge bases including planning representations and rules.The program demonstrates an AI application of natural language and a semi-autonomous telerobot caring out the assembly of the EASE space structure.

Introduction

Interset's general theme is an isomorphism between single sentences about concrete objects mapped to generalized concrete object models. Sentences are viewed as speech acts in an action based "commitment" framework [14]. Interset is the name of a system that is written in Scheme [13] and runs on either a MacIntosh or a MicroVax. The sentence understanding capability uses a schema based approach. Sentences are parsed

and analyzed for "shape" before schema activations. As a result, space station truss assembly may prove be an applicable domain because of the need to carry out a variety of assembly information processing tasks about concrete objects. Specifically, the MIT EASE space structure (a tetrahedral modular element for construction of the space station truss network). In addition, provide an effective interface with the astronaut and the teleoperated robot.

Natural Language (NL) fits into a larger context of how to get the most effective Man Machine Interface (MMI) for a task. The computer is really a type of media for disseminating structural, parametric and control information. "Media Computations" are two concepts put together, MMI plus an AI system that understands the target processes and can display that understanding effectively [7]. Media computations can use: windows, keyboard, mouse, data glove, joy stick, EVA cuff controller, voice, heads-up display, etc. What's important is the interface understands enough to get the operator a "good" understanding of the key processes. Increases effectiveness, and carries out recurring operations. In an environment such as space operations there will be many complex systems requiring effective interface technologies.

In the space vehicle systems context: control, dynamics, solid mechanics, thermo-dynamics, fluid dynamics, and propulsion characteristics will continually undergo change. Even when the systems are functioning properly control and monitoring will be required expending resources to carry out these functions. Simplifying communications during telerobotic operations could provide a useful tool for managing this complexity. English input and output can be helpful to: reduce mistakes, improve operator attention, increase productivity and probably understand our own communications needs better and what is important when discussing space construction concepts. Essentially, the focus is to allow sentence and/or sentence fragments to be used for the construction operation queries. Also relieve the operator from performing repetitive cognitive operations that would normally require ongoing attention. Such as looking through lists or having to remember context dependent logistics. This approach will unfortunately have its detracting points.

First of all the task environment is still evolving this makes knowledge acquisition a moving target in that the representations in the system will have to change based on engineering changes elsewhere in the system. In operation ambiguous interpretation could cause serious accidents. While the system may or may not have influence on the process control functions of teleoperation, faulty logistical information could be just as hazardous. An argument stating that, "Operational telerobotics does not really need logistical support and that it's really quite straight forward", could be made. Another argument is natural language is not advanced enough yet to be used as a reliable technology. Better would be to use menus or specialized command languages. All of these have a certain amount of validity.

The Environment

There are a variety of resources in low earth orbit they are: shuttle and crew, pallet of spares and connectors, beam assembly teleoperator (BAT), previous construction, astronaut that can be EVA or IVA, assembly teleoperation work station (computer, console, etc.), telerobotic operator (mission specialist), subassembly construction area, final construction work area and perhaps an EVA work station [1]. The EASE structure that is assembled at neutral buoyancy is analogous in structure to space station design criteria. EASE Design and Development is part of a joint effort between Marshall Space

Flight Center and MIT Space Systems Laboratory the Principal Investigator is Dr. David Akin. The goals of the collaboration is to compare construction techniques between the Neutral Buoyancy Simulator (NBS) and Low Earth Orbit (LEO). Build a rigid stable structure that is simple and does not clutter the work environment and demonstrate a realistic subassembly for the final space station truss assembly.

The assembly consists of struts and interlocking connectors which are called clusters. Interlocking 6 struts and 4 clusters to form a tetrahedron. EASE structures can be attached together to form truss assembly. The truss assembly can be decomposed in a variety of ways (for example use the octahedron as the motif rather than the tetrahedron). Interset uses a combination of connection topology and a tetragonal space group (4-3-fold which is not an example of a space station truss assembly, that group is actually isometric where a = b = c and the planner angles are at 90 degrees) [4]. Construction is carried out at two primary sites the shuttle subassembly staging area and the truss assembly area. A subassembly is connected at the first staging area and is then transported to final assembly area. Where it is attached to the truss end. Subassemblies are constructed with the aid of a planner written in prolog. The teleoperator using a joy stick and commands the BAT to construct a tetrahedron. The BAT then moves the new assembly to the truss end where is attached. During this process English commands could be given to operate BAT or to request logistical information to improve operations in case of error, malfunction or anomaly.

A "Slightly" Grammatical Parser

Given that a sentence in natural language corresponds to an internal world model. It is possible to create a formal representation system. Then, For any relevant fact about the world there can be a corresponding structure in the representation. Systematic formal operations in representation structures can be devised to carry out valid reasoning. However considerable effort to make the schemas and knowledge representations is usually required. The reference to the adjective slightly in the name of the parser procedure addresses two limitations. One, having a perfect syntactic parse is probably not needed if the domain is fairly simple and there is potentially a large set of knowledge that can correctly interpret the intent of the sentence in the discourse context. Two, using and creating ungrammatical sentence fragments is actually the rule of most communication. In addition, periods of duress and quick decision have a tendency to increase the noise in the sentences (such as: extraneous cursing, miscuing of material, etc.) So an approach of this sort is really of ergonomic practicality.

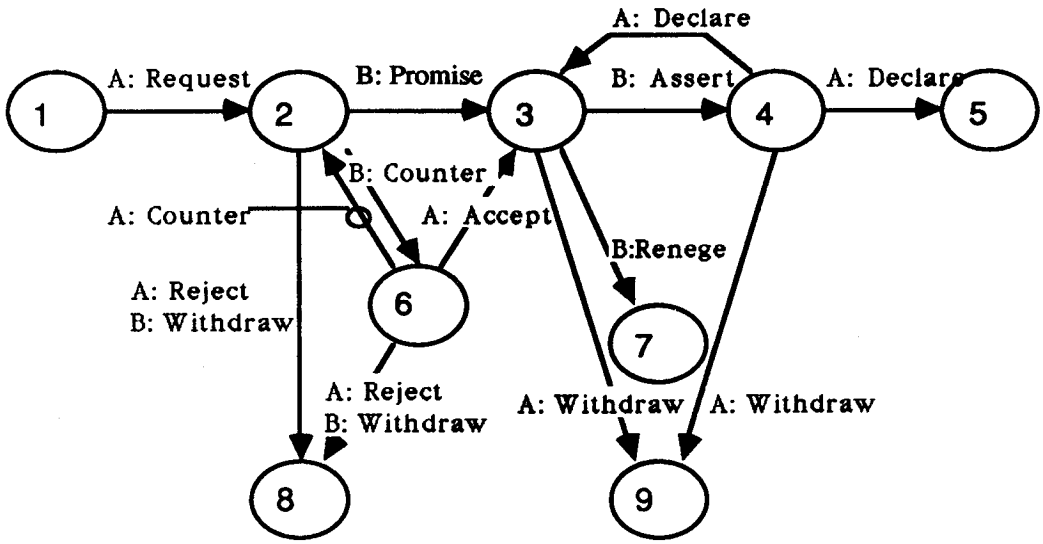
Natural language activity in Interset consists of a variety of functional areas: actors model, a task state space, Interset status information, assembly and construction model, a Filter for the present event space, speech act instantiation, execution control if an unambiguous sentence is "understood", schemas for semantic interpretation, deductive query section, lexicon with endings and meanings. Also, speech act networks, sentence shape, word shapes, A rule based parser and illocutionary act (input sentence).

Walking through the system operation first that occurs the illocutionary act. Words are found in the Lexicon. The parser uses syntactic rules using a head first scan. Syntactic analysis often times can be used to disambiguate word meaning [15]. Predicates are added at this time to label the sense of the word meaning based on the parse. These predicates are still in flux and will not be developed in this paper. The resulting parse tree represents the syntactic categories (noun, adjective, verb, etc.), word root with suffix,

infix, or prefix, and the word meaning senses. Word meaning senses is used to characterize the sentence in a general sense (the "shape" of the sentence).

Sentence shape can be used to make decisions about whether an appropriate schema exists and where in the action network the conversation probably is. A "Schema" of "Frame" contains expectation values, defaults programs, connections to other objects or schemas, and identifies itself as being of a general "type" of thing. We can use all of these types of information potentially to help resolve the what the sentence or sentence fragment might mean. Meaning corresponds to an appropriate representation or process in the data base or as often times referred to as the knowledge base. Sentence shape can be viewed as a vector pointing to the "space" of schemas [12]. If a schema is not found then a "neighbor" could be selected, if this is not appropriate then the space is marked to indicate that a sentence was indexed to this point but nothing was there. Use of learning algorithm could provide knowledge acquisition at this point if one was available to create additional meaning structure.

Dialogues about limited domains often follow similar patterns or chains of discourse (ie. the action network below). The method to generate these is to listen to conversations made by astronauts during the construction process to see what kinds of references are made, the recurrent chains, and the range of sentence types. Then building on the structure as needed when new situations appear they will be properly encoded as schematic information resulting in a robust language understanding capability.



Schema extended to conversation for action (nodes represent states in the conversation)

Winograd and Flores - Understanding Computers and Cognition.

The main property of the shape metric is categorical and/or based on number of nodes away rather than on a continuous means end analysis. The categories are based on the shape predicates. "Nearness" can help disambiguate fragment interms of context. The

temporary activation of a schema acts to preserve relevant data to be used for interpreting the next sentence or in generating a response. The "History" of activation can be used to generate explanations and activity logs.

Knowledge Base and Representations

Process entities, things that can operate upon the domain (such as the BAT). Are important for determining relative position and orientation w.r.t perspective (ie. up, down, left, right, etc.). Process entities also contain status information about their functional parameters (ie. control parameters, orientations and local diagnostics, sensor information, if available).

Task state space. Building the truss assembly is probably best thought of as a state space. Tasks will generally be completed based on a depth first search.

Interset status information. Interset like all systems will have many states that will be relevant to completing the set of construction tasks this information will be accessible to the teleoperator and the ground station as needed. Updating interaction preferences inregards to media computations will be controlled here also.

Assembly and Construction Model. Contains knowledge and rules for assembly. Contains the memory model for space groups, topologly, and sets.

Human Factors

- Interset can be used in a relative operation cycle between 5 seconds and hours. 5 Seconds is the lowest bounds because of typing and minimal computations to carry out the operations.

- Support logistics and information requests are the primary goals at present.

- If interacting with process control systems the BAT, control should be controlled at a fairly high level (ie. "Put that strut together with the cluster." or "Transport that subassembly to the truss work end.") to avoid unnecessary detail.

- Provide task cues.

- Allow for imprecise labels or forgotten terms.

- Viewing human operator w.r.t Single Channel Hypothesis [8] problems exist with modalities of control and coordination of tasks (ie. what level should tasks be abstracted to?)

- Optimally, acoustic input of words.

Conclusion

Interset maps English sentences to representations used to reason geometrically about assembling structures that are similar in nature to space station components. The speech act network is an important segment of schemas which are used to facilitate mapping sentences to the appropriate internal representations. This approach is a familiar one albeit requires considerable knowledge acquisition effort. The overall interface operates

in media computation paradigm. Human factors are of great importance in making the system useful. Ease of use and avoiding errors probably will prove to be important capabilities.

References:

[1] Aerospace Environmental Systems Conference #16,The Development of an EVA Universal Work Station, Miles Moffatt and Fred Abeles (Grumman Corp.), Society of Automotive Engineers, Inc., July 1986.

[2] Artificial Intelligence Applications in Teleoperated Robotic Assemble of the EASE Space Structure, Herbert E. M. Viggh, February 1988, MIT SSL#6-88.

[3] Chomsky's Universal Grammar an Introduction, V. J. Cook, Basil Blackwell Ltd. 1988.

[4] Contemporary Crystallography, Martin Buerger, McGraw-Hill Inc.1970, pp. 1-44.

[5] Counterexamples in Topology, Lynn Arthur Steen and J. Arthur Seebach, Jr., Springer-Verlag New York Inc.1978, pp. 1-30.

[6] The English Imperative, Eirlys Davies, Croom Helm Ltd.1986.

[7] The Media Lab, Stewart Brand, Penguin Books, 1987.

[8] Motor Control and Learning a Behavioral Emphasis, R. Schmidt, Human Kinetics Publishers Inc., 1982.

[9] Optimization Issues in the Design and Control of Large Space Structures, Edited by Manohar P. Kamat, American Society of Civil Engineers, 1985.

[10] Scheme Reference Manual (MIT), Scheme Release 7 Draft, October 18, 1988.

[11] Space Groups for Solid State Scientists, G. Burns, Academic Press 1978.

[12] The Society of the Mind, Marvin Minsky, Touchstone Book, 1985.

[13] The Structure and Interpretation of Computer Programs, Hal Abelson, Gerald Sussman and Julie Sussman, 1985.

[14] Understanding Computers and Cognition, Terry Winograd and Fernando Flores, Addison-Wesley Publishing Inc. 1987.

[15] UNITRAN A Principle-Based Approach to Machine Translation, Bonnie Jean Dorr, MIT AI Lab, December 1987.