Iconic Interfaces in Intelligent AAC Systems

Bruce R. Baker, Semantic Compaction Systems
Eric H. Nyberg 3rd, Carnegie Mellon University/Semantic Compaction Systems
Kathleen F. McCoy, ASEL, University of Delaware/A.I. DuPont Institute
Patrick W. Demasco, ASEL, University of Delaware/A.I. DuPont Institute

Abstract

The problems in developing a system with the complete linguistic ability of a human speaker are overwhelming given the current state of technology and linguistic theory. However, some of the issues involved can be finessed by providing the system with certain kinds of linguistic information. Using multi-meaning sequenced icons as an input medium allows the system to get much information out of few actuations. This paper discusses the use of such an interface in an intelligent AAC system.

Multi-Meaning Sequenced Icons

The ideal interface represents input selections transparently and has a relatively small number of input choices or keys. An interface intended for individuals who are cognitively or physically impaired must generate language with a low number of actuations. Letters have serious problems as input media because they require 6+ actuations per content word. AAC operators usually select one key every 5 to 8 seconds. Congenitally speech impaired people often have weak reading and spelling skills. As an alternative to letters, single meaning pictures have serious limitations, because hundreds of pictures are required to instantiate even the simplest vocabulary (with only one lexical item per picture).

A patented technique called *semantic compaction* approaches the problem of sentence generation by interpreting icons to have different meanings in different contexts (Baker, 82), (Baker, 84). This technique exploits the natural polysemy inherent in illustrations of real-life objects. For example, an apple is not only an apple, it is also red, a fruit, and round. When used in a sentence describing a favorite color, an apple icon can be interpreted as indicating the color red. In general, this approach allows operators to access a much larger number of concepts with the same number of input keys when multi-meaning icons are used in place of letters, words or single meaning pictures.

For both cognitively impaired and cognitively intact individuals, semantic compaction has been thoroughly explored using simple transducer programs to retrieve pre-stored language items when particular sequences of icons are actuated. For example, on a system designed for a cognitively intact operator, an average of 2 actuations is required to retrieve a word (3 or 4 actuations for a template sentence). Such a system supports a substantial (greater than 50%) reduction in the physical effort required for communication, using an input layout with fewer than 100 keys. The cognitive requirements are substantially reduced as well. This is a major improvement, but it is not enough. If the sentence to be generated has 8 words (not an unreasonable sentence length), then this type of system will still require 16 actuations to create it. If the operator needs 8 seconds for each actuation (which is the case for a large population of operators), then an 8 word sentence will still take roughly 128 seconds, or over two minutes.

Symbol Parsing

Existing AAC devices allow the operator to actuate certain choices or sequences of choices on an input device to retrieve pre-stored language items (single words or template sentences). While this is certainly useful, it does not help the cognitively impaired operator who does not have intact syntactic knowledge. The system will output only the words that are retrieved by the operator, in exactly the order that they are accessed, whether or not the resulting utterance is meaningful. This places most of the cognitive burden of communication on the operator, who must provide syntactic, stylistic and pragmatic information to produce an utterance. The pre-stored language items that are retrieved by the operator are generated by the device in the order that they are retrieved, even if that order does not produce a meaningful sentence. Systems that accept sequences of symbols or icons and produce words or sentences in the same order can be called *symbol parsing* or *transduction* systems.

Intelligent Word Parsing

It is becoming possible to develop more intelligent systems which take advantage of artificial intelligence and linguistic techniques to reduce the syntactic load and other cognitive burdens placed on the operator (Demasco et al., 89). The ideal system could require only the content words of an intended utterance and yet produce a syntactically and semantically well formed sentence. Such an intelligent parser has several tasks. Consider the input: *John go store yesterday.* The system must have semantic information about the individual words to determine that *go* is a verb, *John* is an agent of the action (since *go* must take an animate agent), *store* is the object (since physical locations are things that can be gone to), etc. (McCoy et al., 89a).

Once the semantic roles have been deduced, a natural language generation system could be used to fill in necessary syntactic information (e.g., determiners, prepositions, necessary verb inflections) to produce a sentence such as: *John went to the store yesterday.*(McCoy et al., 89b).

An intelligent *word parser* could lift much of the cognitive load from the operator since s/he would no longer be concerned with syntactic information. The system would be constrained to produce semantically and syntactically well-formed sentences.

Of course, the production of such utterances is not without a cost. The system requires large amounts of knowledge and powerful inferencing techniques in order to produce the utterances.

Icons and Intelligent Parsing

By combining a scaled-down version of an intelligent parser with multi-meaning sequenced icons, it is possible to make a usable, intelligent AAC system a reality. A well-chosen geographic layout can provide the system with syntactic and semantic information based on the locations of the icons that are selected, thus reducing on the knowledge and inferencing required by the intelligent parser.

While this layout strategy builds on the basic idea behind the Fitzgerald key (Musselwhite & St. Louis, 82), it differs in several important ways.

- The symbols in the proposed system are multi-meaning.
- The user need not select all of the words in the sentence to be generated -- the word parse is responsible for automatically adding function words that are required for purely syntactic reasons.
- The arrangement is actually helping the system (although it is likely to have benefits for language-impaired operators as well).

In representing a large vocabulary by combining icons together, the icons in three 8-icon columns can combine to produce 300 agents. Similarly, the same 24 icons can combine to produce 300 patients if the icons are situated in a different location to distinguish them from the agent icons. Two columns of 8 icons representing modifiers or function words can be placed at crucial parts of the board. The central notion behind this layout is that the icons can sequence to access a large number of unique strings. It is the multi-meaning nature of icons and their ability to form sequences that indicate a single notion which give them their combinatorial expressive power. If we were using single meaning pictures, we would only be able to represent 24 agents and 24 patients if we used the same layout. It is important to note that the usefulness of an intelligent system will be defeated if it makes use of a burdensome interface that places a high physical and/or cognitive load on the operator while placing severe restrictions on the depth of his/her vocabulary. For this reason, multi-meaning icons seem to be the most appropriate interface media for intelligent AAC systems.

Conclusion

We have described a method that synthesizes iconic interfaces with intelligent parsing. The result is a system which combines the advantages of both with minimal constraints. By keeping the vocabulary at approximately 1000 words, it is possible to implement this system on current generation hardware. The result would be an intelligent communication system useful to a large number of people with disabilities.

Acknowledgments

This work is partially supported by Grant #H133E80015 from the National Institute on Disability and Rehabilitation Research. Additional support has been provided by the Nemours Foundation.

References

B. Baker. Minspeak. Byte, 186ff, September, 1982.

B. Baker. Semantic compaction for sub-sentence vocabulary units compared to other encoding and prediction systems. In *Proceedings of the 10th Conference on Rehabilitation Technology*, pages 118-120, RESNA, San Jose, CA, 1984.

P. Demasco, K. McCoy, Y. Gong, C. Pennington, and C. Rowe. Towards more intelligent AAC interfaces: the use of natural language processing. In *Proceedings of the 12th Annual Conference*, pages 141-142, RESNA, New Orleans, Louisiana, June 1989.

K. McCoy, P. Demasco, Y. Gong, C. Pennington, and C. Rowe.A Semantic Parser for Understanding Ill-Formed Input. In *Proceedings of the 12th Annual Conference*, pages 145-146, RESNA, New Orleans, Louisiana, June 1989.

K. McCoy, P. Demasco,Y. Gong, C. Pennington, and C. Rowe. Toward A Communication Device Which Generates Sentences. In *Proceedings of the 12th Annual Conference*, pages 141-142, RESNA, New Orleans, Louisiana, June 1989.

C. Musselwhite and K. St. Louis, *Communication Programming for the Severely Handicapped: Vocal and Non-Vocal Strategies*, College Hill Press, San Diego, CA, 1982.

Contact

Bruce Baker
Semantic Compaction Systems
801 McNeilly Road
Pittsburgh, PA 15226