

Current Issues in Natural Language Generation:

An Overview of the AAI Workshop on Text Planning and Realization

Eduard H. Hovy, David D. McDonald, and Sheryl R. Young

■ *Text planning is one of the most rapidly growing subfields of language generation. Until the 1988 AAI conference, no workshop has concentrated on text planning and its relationship to realization. This report is a summary of that workshop.*

Traditionally, systems that automatically generate natural language have been conceived as consisting of two principal components: a text planner and a realization grammar. Recent advances in the art, especially in the incorporation of generation systems into large computer applications, have prompted researchers to question this traditional categorization and the architectures used to embody generator systems. Although a series of workshops have been conducted on language generation as a whole, no workshop has concentrated on text planning and its relationship to realization. For this reason, a one-day workshop was organized for the Seventh National Conference on Artificial Intelligence.

Some months before the workshop, a set of questions was sent to a selected group of researchers in the field. The position papers produced in response are contained in the workshop proceedings. About 30 researchers attended the workshop, with 10 presenting their papers. A 15-minute discussion followed every two or three papers. Also included in the proceedings are the program and a list of participants and their addresses.

It is clear from the workshop that text planning is one of the most rapidly growing subfields of language generation, both in the number of researchers and the extent and complexity of the issues addressed. Perhaps for this reason, insufficient time has been available to develop a stable set of terms and paradigms. Many

supposedly stable terms and categorizations were called into question in papers and discussions, and a number of promising avenues for future research were outlined. For example, the historically common decomposition of language generation systems into two stages—planning and realization—was widely deemed more convenient than accurate: The components of a generator should be able to communicate at any level where their information is applicable. How the generation process should be modularized and which particular decisions each module should influence was a topic that dominated the discussions.

Purpose

In comparison to parsing, natural language generation got off to a slow start. However, after about 15 years of work, researchers in this area have developed a substantial body of knowledge and techniques. As nearly everyone in the field sees it, a generator is a back-end process initiated by an applications program (which can be a database system, an expert system, an intelligent computer-aided instruction tutor, and so on). The applications program is usually carrying on some sort of conversation with the user, using a parser and a discourse management system. The generator, which acts as the speaker, takes information from the application (some sort of interface to the application's internal representation is required) and, following a regime that varies from design to design, organizes the information, maps its elements to words and syntactic constructions, and produces the output text (typically, today, one-paragraph long). Extensive reviews of specific systems and approaches to generation can be found in McDonald (1987) and McKeown and Swartout (1987).

Most of the work in generation has focused on the syntactic issues of realization: how to represent rules of grammar, what types of grammars are computationally useful, and how to traverse input representations to produce language. The main issues have been fairly well mapped out, and as a result, we have a number of theoretical works and programs that serve as comparative standards and paradigm setters (McDonald 1983; Mann and Matthiessen 1983; Joshi 1986; Jacobs 1987).

In contrast, the accompanying organization and planning questions have received less attention. Yet without some form of text planning, generator programs are barely able to coherently put together two clauses. In the last 10 years, a number of text-planning theses have appeared (Appelt 1985; McKeown 1985; Hovy 1988; Paris 1987). Largely from this experience, we came to understand the sorts of tasks that a text planner has to perform: determining which elements to say, coherently structuring the input elements, building noun phrases (including pronominalization) from a set of attributes, determining the appropriate level of detail to use, controlling the slant

... text planning is one of the most rapidly growing subfields of language generation ...

and style of the text, and so on. Unfortunately, to date, no one has provided an encompassing model of text planning that coordinates these tasks in a single framework.

Few programs have attempted to put planning and realization together; the focus has almost always been on one aspect or the other. In addition, neither task by itself is sufficient. Systems that focused on planning tended to produce crude language or compromised on the flexibility of the possible output by using prebuilt phrases; systems that focused on realization were often extremely fluent but worked only one one or two kinds of examples. Inevitably, the result was that each side has made some pretty strong assumptions about the capabilities of

the other. In many cases, these assumptions go beyond what current generation technology is capable of. In addition, we have no idea about the boundary between planning and realization, for example, whether the most appropriate model is top-down hierarchical expansion, simple two-stage planning realization, interleaved processing, or shows no distinction at all.

To clarify these issues, it was high time that researchers in both planning and realization sat down together and discussed how the areas relate. Hopefully, this workshop was the beginning of such a discussion. The ultimate goal is the development of a set of terms that describe the points of interaction between planners and realizers and the kinds of interaction which can take place. Having such a set will facilitate future research and help workers in each camp make realistic estimates of the services provided by the other.

Developing such terms is a long-term task. To give some concrete direction to the discussion and bias it toward this issue, the following five questions were posed and distributed to participants some months before the workshop: (1) What is text planning, and what is/are the planning task(s)? (2) What kinds of information or guidance can a realization component legitimately expect from a text planner? (3) What kinds of information or guidance can a text planner legitimately expect from a realization component? (4) Where, if anywhere, is the boundary between planning and realization, and what is the nature of the interface? (5) How should a planner and a realizer interact—serially or interleaved?

Summary of Papers

The workshop was intended to be a forum for positions rather than descriptions of existing text-planning and realization modules. Workshop participants tended to address one of three principal issues: the knowledge important for generation, the architectures best equipped to manage the manifold influences on the task, and guidelines for how future generation research should be conducted.

. . . the historically common decomposition of language generation systems into two stages—planning and realization—was widely deemed more convenient than accurate . . .

Generator Knowledge Requirements

A common planning issue was the nature of the knowledge required to control realization and text planning. Hovy suggested that because these two aspects of generation each consist of a number of semi-independent tasks, they have to be monitored and controlled by a separate set of heuristics which cooperate in nonhierarchical fashion over arbitrarily long stretches of text. These heuristics and the heuristics controlling them give rise to the style of the text. Moore argued for the interleaving of planning and realization while retaining a record of the system's communicative goals and its plan tree; this record is of great use in understanding and controlling later interactions with the audience as the discourse continues. Jacobs noted that instead of having planning always perform the actual selection of alternatives, one can take planning to merely mean the predisposition toward certain locutions, lexical items, and so on, as mandated by the communicative context and goals. The ability to plan partial specifications at various levels provides the system builder with some useful flexibility.

Patten suggested using settings of register, dependent on stereotypical communicative settings and interactions, to preselect many of the choices required by realization. Sanford and Roach made the point that because language conveys more than simply factual information, this additional information, such as the speaker's stance toward the topic, should be included by controlling the generation process.

Making the distinction between the textual and tactical parts of generation (the former being content selection, the latter realization), Paris gave two examples of tactical constraints that might require replanning: the question of whether a complex entity will be realized as a single word or a longer piece and the

occasional inability to realize a relation under certain focus requirements. Zukerman discussed three phenomena that must be taken into account by any text planner in addition to the planner's concern about conveying primary information, namely, additional information to be conveyed, commentary about this information, and style.

Werner argued that lexical selection must take the possibilities anaphora into account and described one way of doing so. Maybury described a way of generating explanations using schemas and rules of conjunction. Nirenburg, McCardell, Nyberg, Huffman, and Kenschaff described a system for performing lexical selection within a machine-translation system in which the input language analyzer plays the role of text planner for the output language realizer. Rambow, Carcagno, and Kittredge provided a set of answers to the questions posed for the workshop, suggesting that text planning consists of two principal tasks: collecting or delimiting what to say and organizing this material into coherent paragraphs.

Issues of Architecture

Assuming that generation consists of a number of relatively separable subtasks, where each subtask helps either build the specification of the text (and ultimately the text itself) or constrain what the specification can contain, how should the flow of processing control occur? McDonald discussed a model in which each module performs its task and then sets preconditions for further processing on modules following it. Different preconditions and different modules selected to fulfill them result in different realizations for the same underlying communicative goals. Both Moore and Hovy suggested the interleaving of text planning and realization tasks under the guidance of control tasks, where text planning and realization tend to be concerned with language- and information-based issues, and control is concerned with how possible variations of presentation can be used to achieve overarching interpersonal and situation-determined communicative goals. Wong and Simmons discussed a blackboard model on

which sentence specifications and constraints can be written and amended as necessary. Most of the other papers assumed a sequential pipeline model.

Ward proposed a localist-connectionist approach in which the most highly activated node (a word) is generated in each cycle. The utterance of each word activates all the words, concepts, and syntactic types that can follow it. The output is not guaranteed to be grammatical, and the generator requires a large amount of finely tuned information in its network.

Guidelines for Future Research

Addressing the thorny issue of how we find suitable levels and languages for expressing the constraints on, and specifications of, the text being built, Meteer suggested considering how one ensures that the planned message can, in fact, be expressed by studying the expressive possibilities afforded by language under different circumstances

In his paper, Bateman suggested that the distinctions and categories uncovered by systemic-functional linguistics be taken as a guide for future research in natural language generation. This branch of linguistics has been active for many years and has developed a large conceptual framework and a rich set of terms and ideas about language as a social semiotic.

Conclusion

Two principal conclusions can be drawn from the papers and discussions. First and most striking was the participants' willingness to consider as the legitimate domain of inquiry a far broader range of issues than had been seriously considered before, issues such as contextual and interpersonal influences on the generation process and the different kinds of planning tasks and methods required to address the problem. Second (a result of the first), was the number of promising new avenues opened for research (and, of course, the concomitant lack of a clear picture of the whole process to the point where useful old notions, such as the division between planning and realization and the types of tasks performed by each, have lost some of their traditional fixity).

Much discussion centered on the fact that generation occurs in the

context of larger systems and that to be effective, generators must make active use of such systems' knowledge. We should develop guidelines for the design of knowledge repositories to facilitate both application and generation tasks. In addition, we should be aware research on spoken language systems parallels our own in the understanding that processing is enhanced when information from various sources is combined (speech recognition, for instance, is much enhanced by the top-down use of pragmatic knowledge).

It seemed clear that similar meetings should be held in the future to work out a set of terms and delineate frameworks of research. McKeown's suggestion for the next workshop is well taken: Given a single short paragraph, have each participant present an analysis of it so as to start from some common ground and converge on some common terms. In all, the workshop engendered a lot of discussion and was a most enjoyable event.

References

- Appelt, D. E. 1985. *Planning English Sentences*. Cambridge, England: Cambridge University Press.
- Hovy, E. H. 1988. *Generating Natural Language under Pragmatic Constraints*. Hillsdale, N.J.: Lawrence Erlbaum.
- Jacobs, P. S. 1987. Knowledge-Intensive Natural Language Generation. *Artificial Intelligence* 33(3): 325-378.
- Joshi, A. K. 1986. The Relevance of Tree Adjoining Grammar to Generation. In *Natural Language Generation: Recent Advances in Artificial Intelligence, Psychology, and Linguistics*, ed. G. Kempen, 63-70. Boston: Kluwer Academic.
- McDonald, D. D. 1987. Natural Language Generation. In *Encyclopedia of Artificial Intelligence*, ed. S. Shapiro, 642-654. New York: Wiley-Interscience.
- McDonald, D. D. 1983. Description-Directed Control: Its Implications for Natural Language Generation. *Computers and Mathematics* 9(1): 111-130. Also in Grosz, B.; Sparck Jones, K.; and Webber, B., eds. 1986. *Readings in Natural Language Processing*, 519-538. San Mateo, Calif.: Morgan Kaufmann.
- McKeown, K. R. 1985. *Text Generation: Using Discourse Strategies and Focus Constraints to Generate Natural Language Text*. Cambridge, England: Cambridge University Press.
- McKeown, K. R., and Swartout, W. R. 1987. Language Generation and Explanation. In *Annual Review of Computer Science* 2, 49-53. Palo Alto, Calif.: Annual Reviews. Also in Zock, M., and Sabah, G. 1988. *Advances in Natural Language Generation*, 1-54. London: Pinter.
- Mann, W. C., and Matthiessen, C. M. I. M. 1983. *Nigel: A Systemic Grammar for Text Generation*, Technical Report, RR-83-105, Information Sciences Inst., Univ. of Southern California.
- Paris, C. L. 1987. *The Use of Explicit User Models in Text Generation: Tailoring to a User's Level of Expertise*, Ph.D. diss., Columbia Univ.

Eduard H. Hovy is currently doing research in text planning as a member of the Penman Language Generation Project at the Information Sciences Institute, the University of Southern California, 4676 Admiralty Way, Marina del Rey, CA 90292-6695. He attained a bachelor of science in mathematics, computer science, and theoretical physics in 1977, a B.Sc. (Hons) in mathematics in 1979 from the Randse Afrikaanse Universiteit in Johannesburg, South Africa, and a Ph.D. in computer science (AI) from Yale University in 1987.

David D. McDonald is currently senior scientist and directs research and development efforts in text processing at MAD Intelligent Systems Inc., 55 Wheeler Street, Cambridge, MA 02138. McDonald received a bachelor of science in linguistics at the Massachusetts Institute of Technology (MIT) in 1972 and a Ph.D. in computer science from the MIT Artificial Intelligence Laboratory in 1980. He then joined the computer science faculty at the University of Massachusetts, Amherst, where he directed research in natural language generation, intelligent tutoring systems, and knowledge representation and was a member of the Atari Encyclopedia Project.

Sheryl R. Young is a research faculty member in the Computer Science Department at Carnegie-Mellon University, Pittsburgh, PA 15213. She attained a bachelor of science in mathematics and psychology at the University of Michigan in 1978. She graduated with a Ph.D. in cognitive science from the University of Colorado in 1984.