Integrating Text Planning and Production in Generation

Eduard H. Hovy

Yale University
Artificial Intelligence Project
2158 Yale Station
New Haven, CT 06520

Abstract

While the task of language generation seems to separate quite naturally into the two aspects of language generation (text planning and text production), it is necessary to have the planning and the production interact at generator decision points in such a way that the former need not contain explicit syntactic knowledge, and that the latter need not contain explicit goal-related information. This paper describes the decision points, the types of plans that are used in making the decisions, and a process that performs the task. These ideas are embodied in a program.[1]

1 Introduction: the Problem

Our current understanding of language generation includes text planning and text production. In generation work of a decade ago (Simmons & Slocum 72; Goldman 75), no text planning phase ever appeared. In the last few years, much work has been done developing text planners. The issue of interaction between planning and production phases was addressed in various ways. This paper suggests a better way to achieve the necessary interaction.

In the simplest systems, planners make only very high-level decisions, such as selecting appropriate speech acts (Cohen 78; Jacobs 85), and play no further role in text expansion.

In systems with more elaborate plans, the text can be planned out in considerable detail before actual production is started. In this approach, there is a one-way flow of information from the planner to the generator (McDonald (personal communication); McDonald 80; Appelt 81). The production process requires this information whenever it must decide how to expand a generator instruction into a series of more detailed instructions. If the decision criteria are based purely on syntactic and rhetorical grounds (using notions such as sentence focus and stress); it is comparatively easy to build them into a pre-expansion planner, since they are relatively simple and only impinge on expansion at a small number of points (such as subject choice and clause content).

However, when you want the decisions to take into account pragmatic considerations such as speaker intentions, conversational setting, and hearer characteristics, it is much more difficult to plan all the decisions before commencing actual expansion. For example, suppose the generator wants to create in the hearer sympathy for a 65-year old beggar. In the sentence "the [*say-age* AGE-INSTANCE-23] woman is homeless", *say-age* should return "old" or even "ancient" rather than "65-year old". For the planner to precompute this decision, it will have to compute all the decisions (via *say-sentence* and *say-subject*, etc.), such as selecting a subject, a head noun, and adjectives, before it will be in a position firstly to realize that AGE is to be said as an adjective, and secondly to determine what the options are in this case. In order to do this computation, the planner will have to have access to information which one would like to claim is properly the exclusive concern of expansion, such as syntactic and lexical knowledge. (For instance, Appelt's planner contains grammatical knowledge spread throughout. Appelt alludes to the problems that this causes in (Appelt 81).) Furthermore, if the planner is going to do this computation down to the level of individual words, it may as well do the generation simultaneously. One can try to get around this problem by having the planner assemble a set of injunctions upon which the expansion can base its decisions. To do so, the injunctions would have to span the space of possible locutions arising from the representation; assembling them would be a very large task.

A better solution is to perform planning only when necessitated by the expansion. This approach is characterized by a two-way communication at decision points. As the example in section 3 shows, five decision points enable a generator to produce flexible yet good text. These decision points are: topic choice; sentence content; sentence organization; clause organization; and word choice.

[1]This work was supported in part by the Advanced Research Projects Agency monitored by the Office of Naval Research under contract N00014-82-K-0149.

There are good reasons why this approach is desirable:

- **modularity:** the planner need not precompute all generator decisions, but must simply be able to furnish answers when required by the generator. That means it need not contain knowledge about the generation process itself.
- **parsimony:** it is unnecessary to make more decisions than exactly and only those the generator requires.
- **opportunism:** it should be possible opportunistically to take advantage of possibilities of locution that arise due to syntax (of which the planner should have no knowledge) and which the production mechanism cannot decide by itself. This can be done only if the planner can make decisions during the actual production.

2 Plans

The generator described here tries to satisfy multiple pragmatic goals. It can be given a goal to affect the hearer in each of the following areas: his knowledge; beliefs (affective values attached to knowledge); future behaviour; emotional state (aspects: hearer's own; toward topic; toward conversation); relation wrt the speaker (aspects: social; emotional). It requires a set of hearer characteristics, and is able to reason about relatedness of concepts to those in the model. Unfortunately it is impossible to describe the goal structure and plan selection process here. (Related work was done in (Cohen 78) and (Hermann & Laucht 78).)

Some of the plans this program contains are sets of injunctions (among others, DESCRIBE-OBJECT, DESCRIBE-CAUSE/HISTORY, CORRECT-MISCONCEPTION, CONVINCE); they are similar to McKeown's (McKeown 82) generational scripts (schemata). These plans are used by a depth-first network traversal planner to make topic-related decisions and instructions. For example, the CONVINCE plan serves the goal to get the hearer to attach a certain affective connotation to some concept. To do this, it directs the story representation traversal by selecting for consideration and, if further criteria are met, for inclusion in the text, concepts that help to support the connotation. Analysis of various written arguments (taken from communist newspapers, pro- and anti- labour strike leaflets, etc.) indicates that the CONVINCE plan contains, at least, the following suggestions about a topic:

- consider the topic if speaker and hearer agree over its connotation
- consider the topic if it opposes a concept accepted by both
- consider concepts the topic is a subgoal to
- consider concepts the topic is an instance of
- consider the topic's results
- consider the current state of affairs relating to the topic
- minimize difference between speaker's desire/interpretation and reality, if it is small
- minimize difference between hearer's desire/interpretation and reality, if it is small
- find someone hearer respects who agrees with speaker's interpretation

This plan is described further in the next section.

Other plans are suggestions for achieving a goal. For example, the relation plan *make the hearer dominant* contains, amongst others, the instructions:

- topic choice: don't change the topic; follow his lead (i.e., don't select unrelated topics)
- sentence content: don't include in a sentence clauses about concepts he doesn't know
- sentence organization: focus on what he deems important
- sentence organization: make questions rather than assertions; ask his opinion
- word choice: use words he knows

whereas the increase-knowledge plan *teach the hearer* contains:

- topic choice: change the topic when required
- sentence content: don't include in a sentence concepts he doesn't know, unless they are the topic
- sentence organization: make short, simple sentences
- word choice: use words he knows

Clearly, there is a potential conflict in the two plans given: one calls for never changing the topic and the other explicitly calls for the opposite. For planners such as NOAH (Sacerdoti 77) this poses a serious problem. However, this is what gives generation its spice! People holding contradictory goals and plans can speak; a generator must be able to hold conflicting plans and merge their instructions into one sensible generator instruction or decision.

The integrated method of generation proposed here is well suited to manage such contradictory demands.

Rather than being mandatory instructions, an activated plan's steps become suggestions on suggestion lists. Since the guidance which the expansion process requires from the planner always takes the form of selecting one of a number of options, a suggestion is some criterion for evaluating a list of alternatives and preferring one (or more). For example, if the speaker's goal is to teach the hearer, one of his plan-suggestions is to use words that the hearer knows. Given a list of possible words, the suggestion will select only such word(s). If none exist, no suggestion is made, and the decision is made by whatever other criteria apply. Thus, when the expansion has to make a word-choice decision, a query containing the syntactically obtained alternatives is sent to the appropriate suggestion list for pragmatic and rhetorical evaluation.

Merging conflicting preferences can be achieved in various ways. Some alternatives are: most popular option; result of the suggestion serving least recently served goal; using some fixed ranking of goals. Woolf's program (Woolf & McDonald 84) contains "meta-rules" to guide the planner through its network of plans (in effect, a planner planner!); similar criteria can be defined in this scheme under which one of the active suggestions will be preferred over other, conflicting ones. My program currently uses the (simplistic) first alternative.

3 An Example

Tracing through an example will make clear the decisions and the planning criteria required to produce text. The example is part of a story denoting one of the primary elections in the Carter-Kennedy Presidential nomination contest in 1980. OUTCOME-4 is one of the results of the election. A similar result, OUTCOME-5, denotes Kennedy's getting 218 delegates. In the conversation, the hearer has been defined to be antipathetic toward Carter and sympathetic toward Kennedy, while the program has the opposite sympathies. (This is accomplished by including the relevant concepts among the interlocutors' sympathy and antipathy characteristics.) The program's goal is to convince the hearer that, even though Kennedy won, Carter didn't do badly. This goal activates the CONVINCE plan.

```
(OUTCOME-4
  (INSTANCE-OF : VOTE-OUTCOME)
  (VOTING : ELECTION-12)
  (ACTOR : CARTER-15)
  (AMOUNT-GOT : 215)
  (UNIT : DELEGATE)
  (CONNOTATION-VALUE : FAIL)
  (DIFF-AMOUNT : 3))
```

where VOTE-OUTCOME, CARTER-15, ELECTION-12, etc., are concepts further defined in a property inheritance network.

3.1 Topic Choice and Change

OUTCOME-4 was nominated for consideration by the CONVINCE plan's fifth rule during scrutiny of the representation of the election. When OUTCOME-4 itself is considered by the topic choice suggestions, the seventh rule fires and builds a *say-sentence* instruction to make a sentence that stresses the small difference between the actual outcome, OUTCOME-4, (to which the hearer is sympathetic) and the speaker's desired outcome (a hypothetical outcome, also represented, to which the hearer is antipathetic). The stress information *min* will enable *say-stress* to generate an appropriate adverb or adjective. In more detail, rule 7 contains:

```
IF the speaker is antipathetic to the topic
   AND there is a hypothetical equivalent
       to the topic
   AND the speaker is sympathetic to the
       equivalent
   AND the difference between these two is
       not large
THEN make a sentence expressing the
     difference, with stress: min
```

3.2 Sentence Content

Since both OUTCOME-4 and OUTCOME-5 are selected by the CONVINCE plan, and since their sentence instructions are both sent to the expander only when a satisfactory end to the line of argumentation has been found, the expansion has the options of making two sentences, or of combining them into a relational sentence such as:

"While Kennedy got 218 delegates, Carter got 215"

"Carter only got 3 delegates fewer than Kennedy"

The decision is made by the planner's sentence content suggestions. If most suggestions call for long or complex sentences, the relation will be said; otherwise, two sentences will be formed. Typically, suggestions to make long or complex sentences are activated by the goals to confuse or bore the hearer, or to make him feel inferior. Similar criteria decide whether or not to include the clause "in the election on Feb 20".

3.3 Sentence Organization

In a typical sentence, you can usually select almost any clause to be the sentence subject. McKeown describes rules concerning focus or stress in order to make subject choices which resulted in natural, flowing text. Additional criteria you can take into account are pragmatic; for example, if it is your intention to anger the hearer, and you know he does not like some aspect of the topic, you may select that aspect as the subject, to give it prominence.

The example contains at least three possible sentence subjects (the ACTOR, AMOUNT-GOT, DIFF-AMOUNT aspects):

"Carter lost the election"
"215 delegates were won by Carter"
"3 delegates was the margin by which Carter lost"

3.4 Clause Organization

Within a clause, the generator has to decide which aspects of the representation to say and how to order them. For example, when making a noun group, it must select the head noun and then decide whether to describe it in full, only give unsaid information, or give an abbreviated version. It then has to select and order the modifiers, both pre- and post-nominal (and some modifiers can appear in both positions), before it can return a form from which the eventual noun group will be built.

Some possibilities are ruled out by text flow rules (one doesn't say "the Georgian male 65-year old Jimmy Carter, the President"); other decisions can be based upon goal-related criteria, and therefore can form another point of interaction between planner and generator. For example, if your goal is to calm the hearer, you should not explicitly include aspects about an object that you know he disapproves of.

3.5 Word Choice

This decision must be made in any representation system rich enough to associate more than one word with a concept. Like (Goldman 75), the program uses discrimination nets attached to the representational primitives; here the discriminations depend both on features of the particular instance of the concept, and on the pragmatic issues mentioned (for example, interaction of word affect with the hearer's sympathies: saying "terrorist" to an IRA soldier may get you shot; saying "freedom fighter" certainly will not!).

Selecting a verb has implications for sentence content and organization. For example, when the subject has been said, *say-predicate* has to select a verb and build up *say*-function instructions for the rest of the sentence. The discrimination net for the concept VOTE-OUTCOME contains, among others, the verbs "win", "lose", and "get". While "win" is inappropriate, either of the others can be said:

"Carter got 215 delegates"
"Carter lost the election"

The two verbs are passed to the word-choice suggestions of the planner, which (in this case) prefer the former option, since the CONVINCE plan's sugggestions call for preferring words with connotations that match the speaker's interpretations. (Since OUTCOME-4 is the failure of an event which the program was sympathetic to, and since "lose" has negative connotations which would oppose the sympathy, it is rejected.)

4 Conclusion

The final text, generated from the whole representation, is:
KENNEDY ONLY GOT A SMALL NUMBER OF DELEGATES IN THE ELECTION ON 20 FEBRUARY. CARTER JUST LOST BY A SMALL NUMBER OF DELEGATES. HE HAS SEVERAL DELEGATES MORE THAN KENNEDY IN TOTAL.

5 Acknowledgements

Thanks to Lawrence Birnbaum and Rod McGuire.

6 References

1. Appelt, D.E., *Planning Natural Language Utterances to Satisfy Multiple Goals*, Ph.D. dissertation, Stanford, 1981.

2. Cohen, P.R., *On Knowing What to Say: Planning Speech Acts*, Ph.D. dissertation, University of Toronto, 1978.

3. Goldman, N.M., *Conceptual Generation*, in R.C. Schank (ed.), **Conceptual Information Processing**, Elsevier North-Holland, 1975.

4. Herrmann, T. & Laucht, M., *Planung von Aeusserungen alsSelektion von Komponenten Implikativer Propositionsstrukturen*, Universitaet Mannheim Technical Report no. 3, 1978.

5. Jacobs, P.S., *PHRED: a Generator for Natural Language Interfaces*, University of California (Berkeley) Technical Report, 1985.

6. McDonald, D., *Natural Language Production as a Process of Decision Making under Constraint*, Ph.D. dissertation, MIT, 1980.

7. McKeown, K.R., *Generating Natural Language Text in Response to Questions about Database Queries*, Ph.D. dissertation, University Of Pennsylvania, 1982.

8. Sacerdoti, E., **A Structure for Plans and Behavior**, Elsevier North-Holland, 1977.

9. Simmons, R.F. & Slocum, J., *Generating English Discourse from Semantic Networks*, Communications of the ACM, Vol 15, No 10, 1972.

10. Woolf, B. & McDonald, D.D., *Context-Dependent Transitions in Tutoring Discourse*, AAAI Proceedings, 1984.