

Speaker Attitudes in Text Planning

Christine DEFRISE
IRIDIA
Université Libre de Bruxelles

Sergei NIRENBURG
Center for Machine Translation
Carnegie Mellon University

ABSTRACT. Natural language generation needs an input language whose expressive power is sufficient for generating texts with the level of quality desired by various NLP applications. In our generator, DIOGENES (e.g., Nirenburg et al., 1989), we use the text meaning representation language TAMERLAN (Nirenburg and Defrise, 1989 and *forthcoming*). Expressions in this language are used as input by the DIOGENES text planner to produce text plan expressions in the text plan language, TPL, that in their turn serve as input to syntactic realization. In this paper we describe the treatment of one of the several types of knowledge encoded in TAMERLAN, namely, speaker attitudes. We also illustrate how these input components are used in producing text plans.

1. Introduction

Our reasons for introducing attitudes as an explicit part of the representation of the meaning of a natural language clause are manifold. In what follows we will review three (partially interconnected) reasons. Representing attitudes a) helps reasoning about speaker goals, b) highlights the argumentative structure of a discourse and c) provides a convenient vehicle for representing modal meanings, including negation.

Almost all spoken and written discourse involves the participants' opinions, so much so that producing a perfectly 'objective' text is an almost impossible task. Within the set of possible goals relating to generating text, the introduction (explicit or implicit, lexicalized or not) of the producer's opinions and points of view serves two goals:

- modifying the consumer's model of the producer by stating facts (including opinions) about self which are not in principle observable by the consumer
- modifying the consumer's opinions by stating producer's opinions about facts of the world (the latter can in principle be observed by the consumer)

The above distinctions only become visible if one decides to represent attitudes overtly. Once this decision is made, it becomes clear that it brings about better description possibilities for additional linguistic phenomena, such as the argu-

mentative structure of discourse. It has been observed (e.g., Anscombe and Ducrot, 1983) that texts have a well-defined *argumentative structure* which reflects the producer's current goals and influences such processes as the ordering of text components and lexical selection in generation. The argumentative structure of a text is realized (or, in text understanding, detected) through linguistic means such as the use of scalar adverbs ('only', 'even', 'almost', 'hardly', etc.), connectives ('but', 'since'), adjectives ('unbearable', 'fascinating', etc.). Sets of such lexical items may have to be considered equivalent from a purely semantic point of view, but different in a facet of their pragmatic effect known as *argumentative orientation*. For example, to illustrate the interplay between semantic content and argumentative orientation (i.e. the producer's attitude towards an event), contrast (1) and (2), which have opposite truth conditions, but the same pragmatic value — from both (1) and (2) the consumer will infer that the producer regards Burma as an inefficient sleuth. In this example it is sufficient to retain pragmatic information concerning the producer's judgment of Burma while the semantic differences (induced by the use of "few" versus "none at all") can be disregarded. However, in other contexts the semantics will matter much more — consider, for instance, (3) for which there can be no paraphrase with "no clues at all."

- (1) Nestor Burma found few clues. Nobody was surprised.
- (2) Nestor Burma found no clues at all. Nobody was surprised.
- (3) Nestor Burma found few clues. But it was still better than having none at all.

The difference between (4) and (5), whose truth conditions are similar, is purely argumentative (or attitudinal) — (4) expresses a positive (optimistic!) attitude, (5) the opposite point of view. This example shows how crucial the extraction of the argumentative structure is, since it is the only clue for the inacceptability of (6).

- (4) Nestor has a little money.

(5) Nestor has little money.

(6) ?Nestor has little money. He wouldn't mind spending some on chocolate.

Finally, we use the attitude markers as a means of expressing modality. Traditionally, formal semanticists have extended first order logic to modal logic in order to account for modals. This places the modals at a purely semantic level, and does not allow for a distinction between what is observable for both producer and consumer, and what is not — such as opinions, beliefs, etc. We consider that expressions like 'perhaps,' 'possibly,' 'it is almost certain that' are clues as to what the producer's beliefs and attitudes are towards facts of the world and help the consumer modify or update his model of the producer. It is for the above reasons that we decided to include a detailed specification of producer attitudes into the input specification for generation.

1.1. Attitudes in TAMERLAN

TAMERLAN is a frame-based representation language for representing text meanings. In our approach, treatment of meaning is agent-centered, that is, all the processes of (and the knowledge used for) understanding, representation and realization of meaning are described with respect to the model of an intelligent agent. This agent includes a model of the world, a model of language and a model of language understanding and generation.¹ The world model includes knowledge (beliefs) about other agents in the world, including interlocutors. In understanding language communication (text or dialog), an intelligent agent extracts and represents a) text meaning; b) the active set of producer (speaker, author) goals and plans that led to the production of this text and c) a set of active parameters of the speech situation, including spatiotemporal characteristics, knowledge about participants and a set of pragmatic factors similar to Hovy's (1988) rhetorical goals. These three items form what we call the *supermeaning* of a text.

To represent text meaning proper, TAMERLAN uses the following basic entity types: *clause*, *relation* and *attitude*. TAMERLAN clauses delimit the propositional and pragmatic content of target language utterances. Relations represent links among events, objects, or textual objects (e.g., sentences, paragraphs, etc.). A definition and detailed description of TAMERLAN is given in Nirenburg and Defrise (forthcoming).

Attitudes are represented in TAMERLAN as a quintuple

$$attitude_i = type_i, value_i, attributed - to_i, scope_i, time_i,$$

where $type_i$ is the type of the attitude; $value_i$ is the value of the attitude, represented as a point or an interval on a

¹A more detailed description of our approach to agent-centered processing see in Nirenburg et al., 1986.

{0,1} scale; *attributed - to_i* points at the intelligent agent this attitude is attributed to; *scope_i* takes as its value that part of the meaning representation to which the attitude is held; and finally *time_i* represents the time at which this attitude is held.

In somewhat greater detail, the definition of the fragment of TAMERLAN dealing with attitudes is as follows.

| | |
|-------------------|--|
| <attitude> | ::= <attitude-type> <attitude-value> <scope> <attributed-to> <attitude-time> |
| <attitude-type> | ::= epistemic deontic volition expectation evaluative saliency |
| <attitude-value> | ::= <numerical-value> |
| <numerical-value> | ::= <point>* <semi-interval>* <interval>* |
| <semi-interval> | ::= > <point> < <point> |
| <interval> | ::= <point> <point> |
| <point> | ::= n, 0 <= n <= 1 |
| <scope> | ::= any Tamerlan expression or set of such |
| <attributed-to> | ::= any instance of the ontological type 'intelligent-agent' |
| <attitude-time> | ::= since <time> until <time> since <time> until <time> |
| <time> | ::= <absolute-time> <time-variable> |
| <absolute-time> | ::= <month>-<date>-<year>- <hours>: <minutes>:<seconds>. <second-fractions> |
| <time-variable> | ::= time_<integer> |

The taxonomy of attitude types is an enhancement of Reichman's treatment of "context spaces" (1985: 56). We use the terminology (if not exactly the spirit) of her distinction among the epistemic, evaluative and deontic *issue-type context spaces*. Context space is Reichman's term for a discourse segment. The issue context space corresponds to our attitude component, while the non-issue context space provides a shallow taxonomy for discourse segment types (Reichman defines *comment*, *narrative support*, and *nonnarrative support* as the non-issue type values). It will be discussed and

illustrated in the next section.

Ontological types are concepts in the intelligent agent's ontology and domain model. The organization of the ontology used in the DIOGENES project see, e.g., in Nirenburg and Levin (1989). Instances of ontological types are actual meanings, including those comprising a TAMERLAN text. Some instances are "remembered instances" (e.g., *John Kennedy, The Washington Post* etc.) and are stored in the agent's episodic memory. The absolute time at (or since or until) which an attitude has been held is shown, for instance, as 05-12-90-13:45:11.56. Relative (or unknown) times are locally represented as variables and treated with the help of temporal relations in TAMERLAN.

The **attributed-to** component of the attitude simply binds the attitude to a particular cognitive agent (which may be the producer of the utterance or some other known or unknown agent), who endorses the responsibility of the content of the utterance. This is important for understanding reported speech, and more generally the polyphony phenomena, in the sense of Ducrot (1984). Ducrot's theory of polyphony, an approach to extended reported speech treatment, provides a framework for dealing with the interpretation of a number of semantic and pragmatic phenomena, e.g. the difference in meaning and use between 'since' and 'because', certain particularities of negative sentences, etc.

The **scope** of the attitude representation pinpoints the entity to which this attitude is expressed. The values of the scope can be an entire clause, a part of it or even another attitude value, with its scope. In understanding the text the text consumer notes the attitudes of the producer toward the content. The attitudes can be expressed toward events (7), objects (8), properties (9) or other attitudes (10).

- (7) The train, unfortunately, left at 5 p.m.
- (8) This book is interesting.
- (9) The meeting was reprehensibly unproductive.
- (10) Unfortunately, I ought to leave.

McKeown and Elhadad (1989) also treat argumentative scales and attitudinals in a generation environment. They, however, consider these phenomena as part of syntax, thus avoiding the need to add a special pragmatic component to their system. This decision is appropriate from the point of view of minimizing the changes in an existing generator due to the inclusion of attitude information. However, if compatibility were not an issue, we believe that introducing a separate component is a more appropriate choice.

2. Attitude Types

The following example illustrates lexical realizations of the epistemic attitude (grouped by approximate *attitude-value*).

- 1 Paul left. I know for sure Paul left. I believe without doubt that Paul left. It is true that Paul left.
- 0.9 Paul must have left. Most probably, Paul left.
- 0.8 Paul may have left. I'm prepared to believe that Paul left. Perhaps Paul left. I'm almost sure Paul left.
- 0.6 It is possible that Paul left. I would think Paul left. Chances are Paul left.
- 0.5 I don't know whether Paul left (or not).
- 0.3 It is unlikely that Paul left. I doubt whether Paul left.
- 0 Paul didn't leave. It is impossible for Paul to have left. I don't know that Paul left. I don't believe (at all) that Paul left. It is not true that Paul left. I know that Paul didn't leave. I believe (without a doubt) that Paul didn't leave.

In our representation we do not distinguish what is from what the agent knows, believes or is certain about. "Objective" reality, thus, doesn't exist in the system. Facts and events belong to the "projected reality" (Jackendoff's term), i.e., reality as perceived by an intelligent agent. The fact that something is or is not, happened or did not happen, bears the mark of the agent's perception. Hence the epistemic attitude. Degrees of knowledge are identified with degrees of belief and degrees of certainty. If an agent knows something, he is certain about it and believes it. "Paul left" = "I (the text producer) believe that Paul left" = "I know that Paul left."

Similarly, we feel that if someone says "Paul didn't leave," it really means (to the text consumer who interprets it) "The producer doesn't believe at all that Paul left" = "The producer doesn't know that Paul left" = "It is impossible for Paul to have left" = "The producer doesn't believe that Paul left" = "It's not true that Paul left." Negation can be understood as an attitude towards the event "Paul left." Hence our decision to collapse the parity of sentence with the epistemic attitudes of the agent. Seeing negation as the realization of an agent's attitude has further advantages. Some uses of negation (the "polemic" use, in denials) as in the following dialog:

- A: Paul came to the party yesterday.
- B: He didn't come. <I saw him downtown with his girlfriend. At the time of the party, he was ... >

demand an analysis that take into account more than parity, contrasting explicitly different agent's attitudes towards the same event (this is similar to Ducrot's (1983) "polyphony"). we can provide a good representation of the above dialog using the "attributed-to" slot of an epistemic attitude frame. This representation will include the representation of the

meaning of the clause “Paul came to the party yesterday” in a TAMERLAN clause, say, clause_1, and two epistemic attitude frames, as follows:

```
(attitude_1
 (type epistemic)
 (value 1)
 (attributed-to A)
 (scope clause_1))
```

```
(attitude_2
 (type epistemic)
 (value 0)
 (attributed-to B)
 (scope clause_1))
```

In generating spoken text, the fact that the representation contains opposite epistemic attitudes with similar scopes will be realized through marked intonation. In contrast, a text featuring a simple negation (not a denial of a previous assertion, but a simple negative assertion) will not be represented using two opposite-value epistemic attitudes with similar scope.

Furthermore, representing parity as an attitude gives rise to “formulas” that elegantly translate certain semantic relations between sentences. For instance the synonymy of the natural language sentences “The book is not interesting” and “The book is uninteresting” is translated in terms of attitudes as

```
(attitude_3
 (type epistemic)
 (value 0)
 (attributed-to A)
 (scope (clause_2 attitude_4)))
```

```
(attitude_4
 (type evaluative)
 (value 1)
 (attributed-to A)
 (scope clause_2))
```

and

```
(attitude_5
 (type epistemic)
 (value 1)
 (attributed-to A)
 (scope (clause_2 attitude_6)))
```

```
(attitude_6
 (type evaluative)
 (value 0)
 (attributed-to A)
 (scope clause_2))
```

respectively (clause_2 represents the meaning glossed as “this book,” because the entire sentences only express the attitude toward the book). Therefore,

(epistemic 0 (evaluative 1)) = (epistemic 1 (evaluative 0))

The equality will be valid only if the “attributed-to” slots of the relevant attitudes have the same fillers. The above means that negation is generally understood as having a “lowering effect” — something not interesting is *less* than interesting. When the condition about the “attributed-to” fillers is not fulfilled, negation must be understood as *polemical*, and in this case the meaning of “the book is not interesting” could, in fact, be as in “the book is not interesting; it is fascinating.” (Once again, in speech a marked intonation will be used.)

The realization of the deontic attitude can be illustrated as follows:

1 I must go. I have to go
0.8-0.2 I ought to go. I'd better go. I should go. You may go.
0 I needn't go.

Some illustrations of the realization of the volition attitude:

1 I wish ... I want to... I will... I will gladly...
0.8-0.2 I hesitate to... It may be a good idea to... I'm reluctant to...
0 I'm unwilling to... I refuse to... I don't want...

Some lexical realizations of the expectation attitude:

1 Not surprisingly... As expected... Of course...
Needless to say...
0.8-0.2 Even (as in “Even Paul left”)
0 Surprisingly... It couldn't be expected...

The last two attitudes, evaluative and saliency can have in their scope not only clauses, relations or attitudes like the previous ones, but also objects and properties.² It is therefore difficult to give a limited and exhaustive set of examples of realizations.

The evaluative scale goes, like the others, from 1 to 0. The endpoints are interpreted as roughly “the best” (“very good”) and “the worst” (“very bad”). Depending on the scope, realizations will greatly vary and will include no lexical realization at all. If the scope is an event, adverbs like *fortunately* and *unfortunately* will be used. If the scope is the physical appearance of a person, the endpoints of the scale of evaluative attitude will be realized as “attractive” and “ugly,” etc.

The saliency attitude plays an important role in selecting the syntactic structure of the target sentences and in the lexical selection. Thus, it will influence the order of elements in a

²The evaluative attitude to “the book” in the example above would, in fact, belong to this latter class. Its scope is a clause only because there is no predication in the sentence other than the attitudinal one. In a sentence like “John read an interesting book” the attitude is clearly toward an object instance.

conjunction; it will be realized syntactically through topicalization ("It is Paul who ...") and lexically through connective expressions such as *last but not least* or *most importantly*.

3. Text Plan Representation

In a nutshell, the flow of data in DIOGENES can be described as follows. The first processing component in DIOGENES is its text planner which, taking into account the input "supermeaning" produces a *text plan*, a structure containing information about the order and boundaries of target language sentences; the decisions of reference realization and lexical selection (for both open and most closed-class lexical items). At the next stage, a set of semantics-to-syntax *mapping rules* are used to produce a set of target-language syntactic structures (we are using the f-structures of LFG — see, e.g., Nirenburg and Levin, 1989). Finally, a syntactic realizer produces a target language text from the set of f-structures.

The text plan language we use in DIOGENES includes the following types of constructs — the plan-sentence, the plan-clause, two kinds of plan-roles and the plan modifier. The frames for these constructs are as follows:

```
(S_#
  (type plan-sentence)
  (subtype <TYPE>)
  (clauses (<C_#>*)
)

(C_#
  (type plan-clause)
  (head <word-sense>)
  (realization {ellipsis | pro | lexical})
  (features <feature-value>* )
  (topic {passive | cleft | passive-cleft
          | active})
  (role <R_#>)*
  (modifiers <MOD_#>* <R_#>*)
)

(R_#
  (type plan-role)
  (head <word-sense>)
  (realization {ellipsis | pro | lexical})
  (features <feature-value>* )
  (role <R_#>)*
  (modifiers <MOD_#>* <C_#>* <R_#>*)
)

(R_#
  (type plan-role)
  (head $SETS)
  (elements <R_#><R_#>*)
  (type {CONJ | DISJ})
  (realization {ellipsis | pro | lexical})
  (features <feature-value>* )
)

(MOD_#
```

```
(type plan-modifier)
(head <word-sense>)
(realization {ellipsis | pro | lexical})
(features <feature-value>* )
(modifiers <MOD_#>*)
)
```

Types of plan sentences at present include simple, compound-conjunctive and compound-disjunctive. The realization property has 3 possible values — *lexical*, *ellipsis* and *pro*. *topic* is used to mark the topicalized/focused elements in the clause; this property is also used to specify that the clause will be active or passive and whether it will feature an active or passive cleft construction.

modifiers is a slot in which all the modifiers of a given plan concept are listed.

The text planner will to determine which of the thematic roles in the input are to be realized as arguments and which, as modifiers.

3.1. Text Planning Rules for Attitudes.

Text planning rules in DIOGENES deal with a variety of phenomena. Some are devoted to text structure proper — the number and order of sentences and clauses to express the meanings of input; clause dependency structures, etc. Others deal with treatment of reference — pronominalization, ellipsis, etc. Still others take care of lexical selection, determine tense and mood features of the target text, etc. In this section we illustrate text planning rules devoted to realization of attitudes.

Rule A1 deals with an attitude of the evaluative type; rules A2 through A4 with attitudes of the epistemic type.

```
A1. IF (and (= clause_i.attitude.type
             evaluative)
        (= clause_i.attitude.value
          (< 0.3))
        (= clause_i.attitude.scope
          clause_i.proposition))
      THEN (add-unit-filler C_i
            'attitude 'unfortunately)

A2. IF (and (= clause_i.attitude.type
             epistemic)
        (= clause_i.attitude.value 1)
        (= clause_i.attitude.scope
          clause_i.proposition))
      THEN (add-unit-facet-filler C_i
            'features 'mood 'declarative)

A3. IF (and (= clause_i.attitude.type
             epistemic)
        (= clause_i.attitude.value 0)
        (= clause_i.attitude.scope
          clause_i.proposition))
      THEN (add-unit-facet-filler C_i
```

```

        `features `mood `negative)
A4. IF (and (= clause_i.attitude.type
            epistemic)
        (= clause_i.attitude.value 0.5)
        (= clause_i.attitude.scope
         clause_i.proposition))
    THEN (add-unit-filler C_i
         `attitude `perhaps)

```

Attitudes get realized either lexically, through the inclusion of a lexical unit or through grammatical features. In the sample rules, the *if* clauses check the values in TAMERLAN and, depending on the actual match, either add features to the text plan or add a lexical realization for the attitudinal meaning (as in Rule A4).

4. Status and Future Work

In the DIOGENES project we adopt the methodological attitude of developing the generator functionalities in a breadth-first fashion. That is to say that, unlike many other projects, we do not tend to describe exhaustively a specific linguistic phenomenon (e.g., negation, anaphora, aspect, scope of quantifiers) or type of processing (e.g., text planning, lexical selection, syntactic realization) before proceeding to the next one (this approach can be considered depth-first). We prefer to go for a complete functioning system which contains all (or, in practice, most) of the above components and covers all (or most) of the above phenomena. It is clear that, at the beginning, the treatment of each (or most) of these components is incomplete, and not every phenomenon is described in sufficient detail. However, this methodology allows us to benefit from a complete experimentation environment and an open-ended architecture that facilitates the addition of knowledge to the system and its testing and debugging. At present we have a working prototype text planning and generation system with narrow coverage. Our current work is devoted to expanding the knowledge needed for achieving a deeper level of analysis of each of the linguistic phenomena covered in the system.

Acknowledgement

The first author was supported by the Belgian National Incentive Program for fundamental research in Artificial Intelligence initiated by the Belgian state — Prime Minister's Office — Science Policy Programming.

Bibliography.

- Anscombe, J.-C. and O. Ducrot. 1983. *L'argumentation dans la langue*. Brussels: Mardaga.
- Defrise, C. and S. Nirenburg (in preparation). Aspects of Text

Meaning. Center for Machine Translation, Carnegie Mellon University.

Ducrot, O. 1984. Polyphonie. In *Lalies*, 4.

Hovy, E. 1987. *Generating Natural Language under Pragmatic Constraints*. Yale University Ph.D. Dissertation.

McKeown, K. and M. Elhadad. 1989. *A Comparison of Surface Language Generators: A Case Study in Choice of Connectives*. MS. Columbia University.

Nirenburg, S. and V. Raskin. 1987. The Subworld Concept Lexicon and the Lexicon Management System. *Computational Linguistics*, Volume 13, Issue 3-4.

Nirenburg, S., E. Nyberg, R. McCardell, S. Huffman, E. Kenschaft and I. Nirenburg. 1988. Diogenes-88. Technical Report CMU-CMT-88-107. Carnegie-Mellon University. June.

Nirenburg, S. and L. Levin. 1989. Knowledge Representation Support. *Machine Translation*, 4, pp. 25 - 52.

Nirenburg, S., E. Nyberg and C. Defrise. 1989. Text Planning with Opportunistic Control. Technical Report CMU-CMT-88-113. Carnegie-Mellon University. June.

Reichman, R. 1985. *Getting Computers to Talk Like You and Me*. Cambridge, MA: MIT Press.