What You Infer Might Hurt You – A Guiding Principle for a Discourse Planner

Ingrid Zukerman
Department of Computer Science
Monash University
Clayton, VICTORIA 3168, AUSTRALIA
ingrid@bruce.cs.monash.edu.au

Abstract

Most Natural Language Generation systems developed to date assume that a user will learn only what is explicitly stated in the discourse. This assumption leads to the generation of discourse that states explicitly all the information to be conveyed, and that does not address further inferences from the discourse. The content planning mechanism described in this paper addresses these problems by taking into consideration the inferences the user is likely to make from the presented information. These inferences are modeled by means of inference rules, which are applied in a prescriptive manner to generate discourse that conveys the intended information, and in a predictive mode to draw further conclusions from the presented information. In addition, our mechanism minimizes the generated discourse by presenting only information the user does not know or about which s/he has misconceptions. The domain of our implementation is the explanation of concepts in high school algebra.

Introduction

It has been widely accepted that much of what is intentionally conveyed during language use is not explicitly expressed [Grice 1978]. The recognition of this fact has significantly influenced research in Natural Language Understanding, e.g., [Norvig 1989], and in Plan Recognition, e.g., [Allen and Perrault 1980], where extensive inferences are drawn from a piece of discourse. A few Natural Language Generation (NLG) systems have addressed particular types of inferences which can be made from statements issued by a system [Joshi et al. 1984, Reiter 1990, Zukerman 1990, Cawsey 1991]. However, most NLG systems developed to date, e.g., [Appelt 1982, McKeown 1985, Paris 1988, Hovy 1988, Moore and Swartout 1989, Cawsey 1990], assume that a user[1] will learn only what is explicitly stated in the discourse.

This assumption may result in the generation of discourse that on one hand is *overly explicit*, i.e., includes information that could have been easily inferred by the hearer from the presented information, and on the other hand *fails to address possible indirect inferences drawn by the user*, i.e., is not complemented by information that contradicts possible erroneous inferences. For example, a possible erroneous indirect inference from the statement "wallabies are like kangaroos" is that wallabies are the same size as kangaroos. A system that is sensitive to a user's inferences should complement this statement with the disclaimer "but smaller."

In this paper, we present a content planning mechanism which addresses these problems. Our mechanism generates *Rhetorical Devices (RDs)*, such as Descriptions, Instantiations and Similes. The generation of RDs is performed by applying inference rules which relate RDs to beliefs. These rules are applied in two different ways during the discourse planning process: *forward reasoning* and *backward reasoning*.

Forward reasoning reasons from the RDs to their possible effects. This process accounts for the generation of the disclaimer for the above Simile between wallabies and kangaroos, by first applying a similarity-based inference rule which conjectures that the hearer will transfer what s/he knows about kangaroos to wallabies, and then blocking the transference of features which are not correct with respect to wallabies. This reasoning mechanism was used in [Zukerman 1990] for the generation of Contradictions and Revisions of a user's possible inferences.

Backward reasoning reasons from a communicative goal to the RDs that may be used to achieve it. For instance, the concept of a stack may be conveyed to a student by means of a Definition, an Analogy (say to a stack of plates in a cafeteria), an Example, or a combination of these RDs. This reasoning mechanism has been widely used in NLG systems, e.g., [Appelt 1982, Hovy 1988, Moore and Swartout 1989, Cawsey 1990]. However, these systems do not represent explicitly the inferential process that allows a hearer to deduce a belief from an RD. In our system, this process

[1] The terms hearer/user/addressee are used interchangeably in this paper.

is modeled by the inference rules.

In the following section, we describe briefly the rules of inference used by our mechanism. In the remainder of this paper, we describe the tasks performed by the content planner.

The Rules of Inference

Our mechanism takes into consideration three types of inferences: (1) direct inferences, (2) indirect inferences, and (3) uniqueness implicatures.

Direct inferences reproduce directly the content of the discourse. The likelihood that a hearer will understand a statement by means of a direct inference is influenced by the complexity and abstractness of the information in the statement and by the addressee's ability to understand abstract explanations, such as stand-alone descriptions or definitions. The abstract-understand inference rule assesses this likelihood.

Indirect inferences produce inferences that are further removed from what was said. These inferences are not always sound. The indirect inference rules considered at present in our model are based on the ones described in [Zukerman 1990], namely: generalization, specialization, similarity and applicability. However, they draw inferences from RDs, rather than from already acquired beliefs. The first three rules reflect student behaviour observed in [Matz 1982]. The generalization rule was also postulated in [Van Lehn 1983, Sleeman 1984]. Applicability is a simple deductive reasoning rule. It states that if we can apply the first set of steps of a procedure to an object, then we can apply the entire procedure to this object. The likelihood of acquiring a message through indirect inferences from an RD depends on the hearer's confidence in the corresponding inference rules and on the strength of the beliefs which participate in the inference process. For instance, in the above wallabies-kangaroos example, the likelihood that the hearer will infer the desired features of wallabies from the Simile "wallabies are like kangaroos" depends on his/her knowledge about kangaroos and on his/her confidence in the similarity inference rule. Finally, in the current system, the application of indirect inference rules emulates a behaviour observed in [Sleeman 1984], whereby good students retain more correct conclusions than incorrect ones, while the opposite happens for mediocre students.

Given a proposition $P(O)$, **uniqueness implicatures** license the inference that P is true *only* with respect to O. For example, upon hearing the statement "Joe has one leg," most people will probably infer that Joe has one leg *only* [Hirschberg 1985][2]. Although the occurrence of uniqueness implicatures is mainly influenced by the wording of the discourse, their impact on a hearer's understanding of the discourse is affected by the hearer's beliefs and by the manner in which the hearer processes the incoming information. That is, people usually expect to add the incoming information to their knowledge pool [Zukerman 1991b]. However, if other information is already in place, a conflict ensues due to the uniqueness implicature resulting from the normal wording of the discourse. For instance, if a speaker says "Bracket Simplification applies to Like Algebraic Expressions," the uniqueness implicature will license the inference that Bracket Simplification applies *only* to Like Algebraic Expressions. Now, if the user believes that Bracket Simplification applies to Numbers, the uniqueness implicature will cause a conflict with this belief.

The domain of our inference rules is RDs, and their range is beliefs. That is, their format is: *Inference(RD)* $\xrightarrow{pr}$ *Belief*, where *pr* is the probability that when applied to the RD in the antecedent, the rule will produce the belief in the consequent. For example, if the RD is an Instantiation and the rule is generalization, then *pr* is the probability that the hearer will generalize the intended belief from the Instantiation. Thus, our rules allow our system to conjecture the effect of an RD on a hearer's beliefs, and act accordingly, i.e., omit information that may be inferred from this RD, and add information that addresses possible incorrect inferences from the RD.

Operation of the Content Planner

Our content planner receives as input a **concept** to be conveyed to the hearer (e.g., Distributive Law), a list of **aspects** that must be conveyed about this concept (e.g., operation and domain), and a **communicative goal** which states the degree to which these aspects must be known (e.g., know well). The output of the content planner is a set of RDs, where each RD is composed of a rhetorical action, such as Assert and Instantiate, applied to a proposition.

In order to convey the intended aspects of a concept, our mechanism first determines the information to be presented, and then proposes RDs to convey this information. However, it is possible that the hearer does not understand the concepts mentioned in a particular RD well enough to understand this RD. Therefore, the generation process is repeated with new communicative goals and aspects with respect to the concepts mentioned in the proposed RDs, in order to add information about these concepts, if necessary. Other discourse planning tasks, such as organizing the generated messages, and selecting a set of RDs among a number of candidate sets of RDs which convey the intended information, are the subject of future research.

Throughout this section, we will use the following sample input to illustrate the operation of the content planner: (Bracket-Simplification, {*domain,operation*}, KNOW). In this input, the communicative goal is for the hearer to know the domain and operation of the Bracket Simplification procedure.

[2] Uniqueness implicatures differ from the implicatures discussed in [Reiter 1990], since they pertain to propositions, rather than to concepts.

Table 1: Propositions Relevant to the Sample Aspects	
Aspect	**Domain Predicate**
domain	[Bracket-Simplification apply-to Like-Algebraic-Expressions] [Bracket-Simplification apply-to Numbers]
operation	[Bracket-Simplification use-1 $\pm$] [Bracket-Simplification use-2 $\times$]

Deciding which Information to Present

In this step, our system produces a list of propositions that must be conveyed in order to satisfy a given communicative goal with respect to specified aspects of a given concept. Our system caters for Grice's Maxim of Quantity [Grice 1975] in that the generated propositions contain only information that the user does not know or about which the user has misconceptions. This feature is particularly useful in situations such as the ones described in [Sleeman 1984], where the user knows most of the steps in a procedure, and needs to be instructed only with respect to a few of them.

Our system first retrieves from a knowledge base the propositions relevant to the given aspects. Next, based on consultation with a model of the hearer's beliefs [Zukerman 1992], the propositions already known by the user are filtered out, and propositions which address misconceptions held by the user are added. Propositions that are weakly believed by the hearer are presented, but they must be prefixed with a Meta Comment which credits the hearer with the belief in question [Zukerman 1991b], e.g., "*As you probably know*, Bracket Simplification applies to Numbers."

For instance, in order to satisfy the aspects in our sample input, the first step determines that the propositions in Table 1 must be known by the hearer[3]. Now, consider a situation where the hearer has the following beliefs with respect to the aspects in question:

[Bracket-Simplification apply-to Algebraic-Expressions] [Bracket-Simplification apply-to Numbers] [Bracket-Simplification use-2 $\times$]

In this case, the propositions [Bracket-Simplification use-2 $\times$] and [Bracket-Simplification apply-to Numbers] are omitted from the propositions to be conveyed, and the negation of the wrongly believed proposition [Bracket-Simplification apply-to Algebraic-Expressions] is added. This process results in the propositions in Table 2.

Proposing Rhetorical Devices

In this step, the content planner proposes RDs to convey the set of propositions produced in the previous step. To this effect, it takes into consideration inferences a hearer is likely to perform based on these RDs. Our procedure is based on the tenet that while processing a piece of discourse in an interactive setting, a hearer will draw immediate inferences from the discourse, but will perform further reaching inferences after the entire discourse has been processed. Thus, in order to address these immediate inferences, our algorithm draws one round of inferences from a proposed RD. Each inference rule that is applicable to this RD may be instantiated more than once during a round of inferences. This process is carried out by the procedure *Propose-RDs*.

Propose-RDs(propositions-to-be-conveyed)

1. Select a set of propositions that pertain to a particular aspect to be conveyed.
2. Apply inference rules in *backward reasoning* mode in order to propose a set of RDs which convey these propositions. Each RD in this set constitutes a different alternative for conveying the propositions in question.
3. For each alternative RD in the set of RDs, apply inference rules in *forward reasoning* mode in order to draw the inferences that can de made from this RD.
   (a) Update the list of propositions to be conveyed as follows:
      i. If an inference is correct and it corresponds to one of the propositions to be conveyed, then if the inference is strong enough, the proposition no longer has to be said, and it is deleted from the list of propositions to be conveyed. Otherwise, the inference has had some effect on the proposition to be conveyed, but this effect is not sufficient to determine that the proposition will be believed by the addressee. (A correct inference that does not correspond to one of the propositions to be conveyed has no effect on the discourse[4].)
      ii. If an inference is incorrect, then if it does not correspond to any of the propositions to be conveyed, its negation is added to the list of propositions to be conveyed. (If it corresponds to a

[3]The relationships use-1 and use-2 indicate the temporal ordering of a mathematical operation.

[4][Zukerman 1990] describes a mechanism which produces discourse that addresses such inferences if they are weak.

Table 2: Propositions to be Conveyed	
Aspect	Domain Predicate
domain	[Bracket-Simplification apply-to Like-Algebraic-Expressions] ¬[Bracket-Simplification (always) apply-to Algebraic-Expressions]
operation	[Bracket-Simplification use-1 ±]

proposition to be conveyed, it is already being addressed.)

(b) Update the model of the hearer with the above inferences.

(c) If the updated list of propositions to be conveyed is not empty, then add the RDs produced by Propose-RDs(updated-propositions-to-be-conveyed) to the RD proposed in this alternative.

To illustrate the workings of this algorithm, let us return to our Bracket Simplification example. For our discussion, we assume that the addressee is able to understand abstract explanations, i.e., the *pr* of the rule *Abstract-understand(Assertion)* $\xrightarrow{pr}$ *Belief* is quite high. Now, the aspects to be conveyed with respect to Bracket Simplification are domain and operation. In the current implementation, we select operation first, since the inferences from the RDs generated to convey this aspect tend to affect other propositions to be conveyed. Next, we apply rules of inference in backward reasoning mode to generate RDs that can convey the proposition [Bracket-Simplification use-1 ±]. This step yields the RDs {Assertion} and {Assertion + Instantiation}, where both RDs have a sufficiently high probability of conveying the intended proposition. In both alternatives, the relationship use-1 in the Assertion is conveyed by a descriptor such as "before multiplying" which identifies the position of the ± operation in the Bracket Simplification procedure.

Let us first consider the alternative initiated by {Assertion}. In this case, the application of the inference rules in forward reasoning mode does not affect any of the other propositions to be conveyed. Hence, we update the model of the hearer to reflect the fact that s/he has been informed of the first step of Bracket Simplification, and re-activate our algorithm with respect to the propositions in the aspect domain.

During the backward reasoning step, our mechanism determines that the proposition [Bracket-Simplification apply-to Like-Algebraic-Expressions] may also be conveyed either by an Assertion or by an Assertion accompanied by an Instantiation. In both cases, during the forward reasoning stage, the following inferences may be drawn from the Assertion: (1) a similarity-based inference based on the hearer's belief that Numbers are similar to Like Algebraic Expressions; (2) a generalization based on the belief that Like Algebraic Expressions are a subset of Algebraic Expressions; and (3) a uniqueness implicature. The similarity-based inference corroborates the hearer's correct belief that Bracket Simplification applies to Numbers; the generalization corroborates his/her incorrect belief in the applicability of Bracket Simplification to Algebraic Expressions; and the uniqueness implicature concludes that Bracket Simplification applies *only* to Like Algebraic Expressions, and hence not to Numbers or to Algebraic Expressions.

The uniqueness implicature, which conflicts with the similarity-based inference and with the user's belief that Bracket Simplification applies to Numbers, may be prevented by prefixing the proposed Assertion with information that corroborates the user's belief, e.g., "*In addition to Numbers*, Bracket Simplification applies to Like Algebraic Expressions." At first glance, it appears that information that was omitted in the filtering process (see preceding section) is now being reinstated. However, the generation of this preamble links the new information to an existing belief held by the hearer, rather than informing the hearer that Bracket Simplification applies to Numbers.

The generalization, which conflicts with the uniqueness implicature and corroborates the hearer's erroneous belief that Bracket Simplification applies to Algebraic Expressions, is already being addressed by the second domain proposition in Table 2. Hence, nothing needs to be added to the list of propositions to be conveyed. However, the fact that the generalization can be inferred from the proposed Assertion supports the generation of an expectation violation Meta Comment [Zukerman 1991b], such as "but" or "however," which links this Assertion with the RD(s) that will be generated to convey the second domain proposition.

The generation of RDs for the second domain proposition in Table 2 is performed similarly. This yields the output in Table 3 for the alternative where an Assertion was generated for the first and third proposition in Table 2, and a Negation for the second proposition. Our current implementation produces the names of the RDs and the propositional representation. The English text has been added for illustrative purposes.

We conclude this discussion by considering briefly the alternative headed by {Assertion + Instantiation} of the proposition [Bracket-Simplification use-1 ±]. This alternative will result in a discourse which is markedly different from the one in Table 3, if the proposition is instantiated with respect to a Like Algebraic Expression, such as $2(2x + 3x)$, and thereafter, in the forward inference step, the generalization inference rule

Table 3: Sample RDs Produced during Content Planning	
Mention	[Bracket-Simplification apply-to Numbers] "In addition to Numbers,
Assert	[Bracket-Simplification apply-to Like-Algebraic-Expressions] Bracket Simplification applies to Like Algebraic Expressions,
Negate	[Bracket-Simplification (always)apply-to Algebraic-Expressions] but it does not always apply to Algebraic Expressions.
Assert	[Bracket-Simplification use-1 $\pm$] Before multiplying, we add or subtract the terms inside the brackets."

produces the inference [Bracket-Simplification apply-to Like-Algebraic-Expressions] from this Instantiation. In this case, this proposition will be deleted from the list of propositions to be conveyed.

Conveying the Concepts in an RD

At this point in the content planning process, we have a number of candidate sets of RDs, where each set conveys the specified aspects of the intended concept. For each of these sets, we now have to ascertain that the hearer understands the concepts mentioned in its RDs well enough to understand these RDs. To this effect, for each of these concepts, the content planner performs the following actions: (1) it determines the aspects of the concept which are relevant to the understanding of the proposition which contains the concept, (2) it determines a communicative goal for these aspects, and (3) it regresses to generate RDs that accomplish this communicative goal with respect to the selected aspects of the concept. This process generalizes the mechanism described in [Zukerman 1991a].

The determination of the aspects the hearer must know about a concept in order to understand a proposition which contains this concept is based on the main predicate of the proposition and on the role of the concept with respect to this predicate. For example, in order to understand the Assertion [Bracket-Simplification apply-to Like-Algebraic-Expressions] proposed above, the hearer must know what Like-Algebraic-Expressions are and what they look like. Hence, the system returns the aspects *membership-class* and *structure*.

The determination of a communicative goal with respect to the selected aspects of a concept is based on the relevance of this concept to the original communicative goal. That is, the more relevant the concept is to this communicative goal, the better it should be known by the addressee, and vice versa. This consideration is implemented by lowering the expertise requirements with respect to a concept as the recursion becomes deeper. In this manner, we preclude the elaboration of concepts which are far removed from the main concept to be conveyed, while at the same time, ensuring a minimal level of competence with respect to these concepts.

Conclusion

The content planning mechanism presented in this paper generates RDs by taking into consideration the inferences a hearer is likely to draw from the presented information. To this effect, our mechanism applies inference rules both in backward and in forward reasoning mode. Although these inference rules are generally applicable, the conditions for the application of the direct and indirect inference rules and for the acquisition of the conclusions they draw vary for different types of users. Uniqueness implicatures, on the other hand, are influenced by expectations which are common to all users, in addition to the wording of the discourse.

Our mechanism minimizes the generated discourse by presenting only information that the user does not know or about which s/he has a misconception, and by omitting information which the hearer is likely to infer from the presented information. The inference mechanism that supports the latter capability also enables our mechanism to address possible incorrect inferences from the discourse. To perform these tasks, our mechanism requires a model of a user's beliefs and skills, and of his/her possible inferences. The former may be acquired with the help of a diagnostic system, such as the ones described in [Sleeman 1982, Burton 1982], and the latter is based on research by [Matz 1982, Van Lehn 1983, Sleeman 1984] about mathematical inferences commonly drawn by students.

A prototype of our content planning mechanism is in advanced stages of implementation. The implementation of the generation of Assertions, Negations and Instantiations in the framework of the algorithm Propose-RDs has been completed. The generation of Analogies and Similes is currently being implemented. Once the system is fully operational, it will be evaluated by presenting the texts generated for different types of students to the corresponding target audiences. The response of the students to these texts will be compared with their response to texts from algebra textbooks and texts produced by the traditional NLG approach.

Finally, an interesting line of investigation for further work consists of activating the system in a reflective mode after a session with the user has been completed. In this mode, the system would draw further reaching inferences from the generated discourse. Typically, these inferences would interact with each other, thereby requiring a processing mechanism that combines the inferences until the beliefs in the user model reach quiescence. The result of this process would then be the starting point of the next interaction with the user.

References

Allen, J.F. and Perrault, C.R. (1980), Analyzing Intention in Utterances. In *Artificial Intelligence* 15, pp. 143-178.

Appelt, D.E. (1982), Planning Natural Language Utterances to Satisfy Multiple Goals. Technical Note 259, SRI International, March 1982.

Burton, R.R. (1982), Diagnosing Bugs in a Simple Procedural Skill. In *Intelligent Tutoring Systems*, Sleeman, D. and Brown, J.S. (Eds.), London: Academic Press, pp. 157-183.

Cawsey, A. (1990), Generating Explanatory Discourse. In *Current Research in Natural Language Generation*, Dale, R., Mellish, C. and Zock, M. (Eds.), Academic Press, pp. 75-102.

Cawsey, A. (1991), Using Plausible Inference Rules in Description Planning. In *Proceedings of the Fifth Conference of the European Chapter of the ACL*.

Grice, H.P. (1975), Logic and Conversation. In *Syntax and Semantics, Volume 3: Speech Acts*, Cole, P.J. and Morgan, J.L. (Eds.), Academic Press, pp. 41-58.

Grice, H.P. (1978), Further Notes on Logic and Conversation. In *Syntax and Semantics, Volume 9: Pragmatics*, Cole, P.J. and Morgan, J.L. (Eds.), Academic Press.

Hirschberg, J.B. (1985), A Theory of Scalar Implicature. Doctoral Dissertation, The Moore School of Electrical Engineering, University of Pennsylvania, Philadelphia, Pennsylvania.

Hovy, E.H. (1988), Planning Coherent Multisentential Text. In *Proceedings of the Twenty-Sixth Annual Meeting of the Association for Computational Linguistics*, Buffalo, New York, pp. 163-169.

Joshi, A., Webber, B.L., and Weischedel, R.M. (1984), Living Up to Expectations: Computing Expert Responses. In *AAAI-84, Proceedings of the National Conference on Artificial Intelligence*, Austin, Texas, pp. 169-175.

Matz, M. (1982), Towards a Process Model for High School Algebra Errors. In *Intelligent Tutoring Systems*, Sleeman, D. and Brown, J.S. (Eds.), London: Academic Press, pp. 25-50.

McKeown K.R. (1985), Discourse Strategies for Generating Natural Language Text. In *Artificial Intelligence* 27, pp. 1-41.

Moore, J.D. and Swartout, W. (1989), A Reactive Approach to explanation. In *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence*, Detroit, Michigan, pp. 1504-1510.

Norvig, P. (1989), Marker Passing as a Weak Method for Text Inferencing. In *Cognitive Science* 13, pp. 569-620.

Paris, C.L. (1988), Tailoring Object Descriptions to a User's Level of Expertise. In *Computational Linguistics* 14(3), pp. 64-78.

Reiter, E. (1990), Generating Descriptions that Exploit a User's Domain Knowledge. In *Current Research in Natural Language Generation*, Dale, R., Mellish, C. and Zock, M. (Eds.), Academic Press, pp. 257-285.

Sleeman, D. (1982), Inferring (mal) rules from pupil's protocols. In *ECAI-82, Proceedings of the European Artificial Intelligence Conference*, pp. 160-164.

Sleeman, D. (1984), Mis-Generalization: An Explanation of Observed Mal-rules. In *Proceedings of the Sixth Annual Conference of the Cognitive Science Society*, Boulder, Colorado, pp. 51-56.

Van Lehn, K. (1983), Human Procedural Skill Acquisition: Theory, Model and Psychological Validation. In *AAAI-83, Proceedings of the National Conference on Artificial Intelligence*, Washington, D.C., pp. 420-423.

Zukerman, I. (1990), A Predictive Approach for the Generation of Rhetorical Devices. In *Computational Intelligence* 6(1), pp. 25-40.

Zukerman, I. (1991a), Avoiding Mis-communication in Concept Explanations. In *Proceedings of the Thirteenth Annual Conference of the Cognitive Science Society*, Chicago, Illinois, pp. 406-411.

Zukerman, I. (1991b), Using Meta Comments to Generate Fluent Text in a Technical Domain. In *Computational Intelligence* 7(4), Special Issue on Natural Language Generation, pp. 276-295.

Zukerman, I. (1992), Content Planning based on a Model of a User's Beliefs and Inferences. In *Proceedings of the Third International Workshop on User Modeling*, Dagstuhl, Germany.