

Planning Reference Choices for Argumentative Texts

Xiaorong Huang*

Techne Knowledge Systems

439 University Avenue

Toronto, Ontario M5S 3G4

Canada

xh@FormalSystems.ca

Abstract

This paper deals with the reference choices involved in the generation of argumentative text. Since a natural segmentation of discourse into attentional spaces is needed to carry out this task, this paper first proposes an architecture for natural language generation that combines *hierarchical planning* and *focus-guided navigation*, a work in its own right. While hierarchical planning spans out an attentional hierarchy of the discourse produced, local navigation fills details into the primitive discourse spaces. The usefulness of this architecture actually goes beyond the particular domain of application for which it is developed.

A piece of argumentative text such as the proof of a mathematical theorem conveys a sequence of derivations. For each step of derivation, the premises derived in the previous context and the inference method (such as the application of a particular theorem or definition) must be made clear. Although not restricted to nominal phrases, our reference decisions are similar to those concerning nominal subsequent referring expressions. Based on the work of Reichmann, this paper presents a discourse theory that handles reference choices by taking into account both textual distance as well as the attentional hierarchy.

1 Introduction

This paper describes how reference decisions are made in *PROVERB*, a system that verbalizes machine-found natural deduction (ND) proofs. A piece of argumentative text such as the proof of a mathematical theorem can be viewed as a sequence

*Much of this research was carried out while the author was at Dept. of CS, Univ. of the Saarland, supported by DFG (German Research Council). This paper was written while the author was a visitor at Dept. of CS, Univ. of Toronto, using facilities supported by a grant from the Natural Sciences and Engineering Research Council of Canada.

of derivations. Each such derivation is realized in *PROVERB* by a *proof communicative act* (PCA), following the viewpoint that language utterances are actions. PCAs involve referring phrases that should help a reader to unambiguously identify an object of a certain type from a pool of candidates. Concretely, such references must be made for previously derived conclusions used as premises and for the inference method used in the current step.

As an example, let us look at the PCA with the name *Derive* below:

(Derive Derived-Formula: $u * 1_U = u$
Reasons: ($unit(1_U, U, *)$, $u \in U$)
Method: Def-Semigroup*unit)

Here, the slot *Derived-Formula* is filled by a new conclusion which this PCA aims to convey. It can be inferred by applying the filler of *Method* to the filler of *Reasons* as premises. There are alternative ways of referring to both the *Reasons* and the *Method*. Depending on the discourse history, the following are two of the possible verbalizations:

1. (inference method omitted):

“Since 1_u is the unit element of U , and u is an element of U , $u * 1_U = u$.”

2. (reasons omitted):

“According to the definition of unit element, $u * 1_U = u$.”

An explicit reference to a premise or an inference method is not restricted to a nominal phrase, as opposed to many of the treatments of subsequent references found in the literature. Despite this difference, the choices to be made here have much in common with the choices of subsequent references discussed in more general frameworks (Reichman, 1985; Grosz and Sidner, 1986; Dale, 1992): they depend on the *availability* of the object to be referred to in the context and are sensitive to the segmentation of a context into an *attentional hierarchy*. Therefore, we have first to devise an architecture for natural language generation that facilitates a natural and effective segmentation of discourse. The

basic idea is to distinguish between language production activities that effect the global shift of attention, and language production activities that involve only local attentional movement. Concretely, *PROVERB* uses an architecture that models text generation as a combination of *hierarchical planning* and *focus-guided navigation*. Following (Grosz and Sidner, 1986) we further assume that every posting of a new task by the hierarchical planning mechanism creates new attentional spaces. Based on this segmentation, *PROVERB* makes reference choices according to a discourse theory adapted from Reichman (Reichman, 1985; Huang, 1990).

2 The System *PROVERB*

PROVERB is a text planner that verbalizes natural deduction (ND) style proofs (Gentzen, 1935). Several similar attempts can be found in previous work. The system EXPOUND (Chester, 1976) is an example of *direct translation*: Although a sophisticated linearization is applied on the input ND proofs, the steps are then translated locally in a template-driven way. ND proofs were tested as inputs to an early version of MUMBLE (McDonald, 1983); the main aim, however, was to show the feasibility of the architecture. A more recent attempt can be found in THINKER (Edgar and Pelletier, 1993), which implements several interesting but isolated proof presentation strategies. *PROVERB* however can be seen as the first serious attempt for a comprehensive system that produces adequate argumentative texts from ND style proofs. Figure 1 shows the architecture of *PROVERB* (Huang, 1994a; HuangFiedler, 1997): the macroplanner produces a sequence of PCAs, the DRCC (Derive Reference Choices Component) module of the microplanner enriches the PCAs with reference choices. The TSG (Text Structure Generator) module subsequently produces the text structures as the output of the microplanner. Finally, text structures are realized by TAG-GEN (Kilger and Finkler, 1995), our realization component. In this paper, we concentrate only on the macroplanner and the DRCC component.

2.1 Architecture of the Macroplanner

Most current text planners adopt a hierarchical planning approach (Hovy, 1988; Moore and Paris, 1989; Dale, 1992; Reithinger, 1991). Nevertheless there is psychological evidence that language has an unplanned, spontaneous aspect as well (Ochs, 1979). Based on this observation, Sibun (Sibun, 1990) implemented a system for generating descriptions of objects with a strong domain structure, such as houses, chips, and families. Her system produces text using a technique she called *local organization*. While a hierarchical planner recursively breaks generation tasks into subtasks, local organization navigates the domain-object following the local focus of

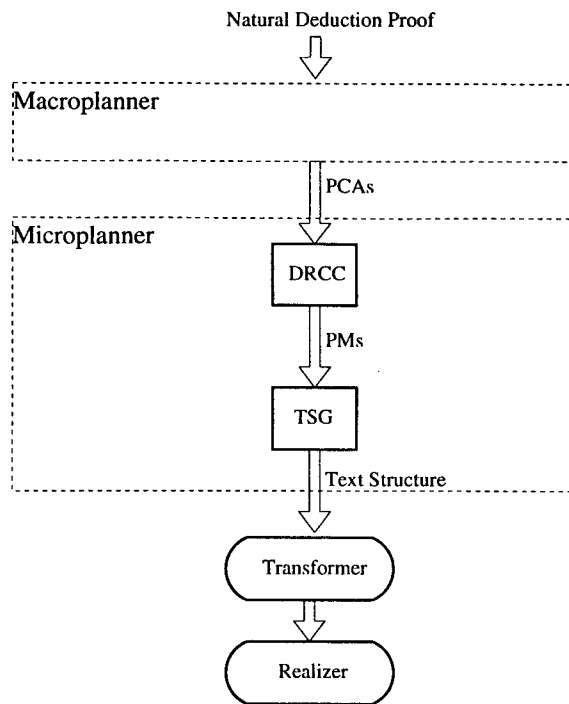


Figure 1: Architecture of *PROVERB*

attention.

PROVERB combines both of these approaches in a uniform planning framework (Huang, 1994a). The *hierarchical planning* splits the task of presenting a particular proof into subtasks of presenting subproofs. While the overall planning mechanism follows the RST-based planning approach (Hovy, 1988; Moore and Paris, 1989; Reithinger, 1991), the planning operators more resemble the schemata in schema-based planning (McKeown, 1985; Paris, 1988) since presentation patterns associated with specific proof patterns normally contain multiple RST-relations. *PROVERB*'s hierarchical planning is driven by proof patterns that entail or suggest established ways of presentation. For *trivial* proofs that demonstrate no characteristic patterns, however, this technology will fail. *PROVERB navigates* such relatively small parts of a proof and chooses the next conclusion to be presented under the guidance of a local focus mechanism.

While most existing systems follow one of the two approaches exclusively, *PROVERB* uses them as complementary techniques in an integrated framework. In this way, our architecture provides a clear way of factoring out domain-dependent presentation knowledge from more general NLG techniques. While *PROVERB*'s hierarchical planning operators encodes accepted format for mathematical text, its local navigation embodies more generic principles of

language production.

The two kinds of *planning operators* are treated accordingly. Since hierarchical planning operators embody explicit communicative norms, they are given a higher priority. Only when none of them is applicable, will a local navigation operator be chosen.

2.2 Proof Communicative Acts

PCAs are the primitive actions planned by the macroplanner of *PROVERB*. Like speech acts, they can be defined in terms of the communicative goals they fulfill as well as their possible verbalizations. The simplest one conveying the derivation of a new intermediate conclusion is illustrated in the introduction. There are also PCAs that convey a partial plan for further presentation and thereby update the reader's global attentional structure. For instance, the PCA

(Begin-Cases Goal: *Formula*
Assumptions: (*A B*))

creates two attentional spaces with *A* and *B* as the assumptions, and *Formula* as the goal by producing the verbalization:

"To prove *Formula*, let us consider the two cases by assuming *A* and *B*."

2.3 Hierarchical Planning

Hierarchical planning operators represent communicative norms concerning how a proof is to be presented can be split into *subproofs*, how the subproofs can be mapped onto some linear order, and how primitive subproofs should be conveyed by PCAs. Let us look at one such operator, which handles proof by case analysis. The corresponding schema of such a proof tree¹ is shown in Figure 2, where

$$\frac{\frac{\dots, \frac{F}{\vdots}, \frac{G}{\vdots}}{\vdots} \quad \frac{F}{\vdots} \quad \frac{G}{\vdots}}{\frac{?L_4 : F \vee G \quad ?L_2 : Q \quad ?L_3 : Q}{?L_1 : \Delta \vdash Q}} \text{CASE}$$

Figure 2: Proof Schema Case

the subproof rooted by $?L_4$ leads to $F \vee G$, while subproofs rooted by $?L_2$ and $?L_3$ are the two cases proving Q by assuming F or G , respectively. The applicability encodes the two scenarios of case analysis, where we do not go into details. In both circumstances this operator first presents the part leading to $F \vee G$, and then proceeds with the two cases. It also inserts certain PCAs to mediate between parts

¹We adopt for proof tree the notation of Gentzen. Each bar represents a step of derivation, where the formula beneath the bar is derived from the premises above the bar. For the convenience of discussion, some formulae are given an identifying label, such as $?L_1$.

of proofs. This procedure is captured by the planning operator below.

Case-Implicit

- Applicability Condition: $((\text{task } ?L_1) \vee (\text{local-focus } ?L_4)) \wedge (\text{not-conveyed } (?L_2 ?L_3))$
- Acts:
 1. if $?L_4$ has not been conveyed, then present $?L_4$ (subgoal 1)
 2. a PCA with the verbalization: "First, let us consider the first case by assuming F ."
 3. present $?L_2$ (subgoal 2)
 4. a PCA with the verbalization: "Next, we consider the second case by assuming G ."
 5. present $?L_3$ (subgoal 3)
 6. mark $?L_1$ as conveyed
- features: (hierarchical-planning compulsory implicit)

2.4 Planning as Navigation

The *local navigation* operators simulate the unplanned part of proof presentation. Instead of splitting presentation goals into subgoals, they follow the local derivation relation to find a proof step to be presented next.

2.4.1 The Local Focus

The node to be presented next is suggested by the mechanism of *local focus*. In *PROVERB*, our local focus is the last derived step, while focal centers are semantic objects mentioned in the local focus. Although logically any proof node that uses the local focus as a premise could be chosen for the next step, usually the one with the greatest semantic overlap with the *focal centers* is preferred. In other words, if one has proved a property about some semantic objects, one will tend to continue to talk about these particular objects, before turning to new objects. Let us examine the situation when the proof below is awaiting presentation.

$$\frac{\frac{[1] : P(a, b) \quad [1] : P(a, b), [3] : S(c)}{[2] : Q(a, b)}, \quad [4] : R(b, c)}{[5] : Q(a, b) \wedge R(b, c)}$$

Assume that node [1] is the local focus, $\{a, b\}$ is the set of focal centers, [3] is a previously presented node and node [5] is the root of the proof to be presented. [2] is chosen as the next node to be presented, since it does not introduce any new semantic object and its overlap with the focal centers $\{a, b\}$ is larger than the overlap of [4] with the focal centers $\{b\}$.

For local focus mechanisms used in another domain of application, readers are referred to (McKeown, 1985).

3 The Attentional Hierarchy

The distinction between hierarchical planning and local navigation leads to a very natural segmentation

NNo	S;D	Formula	Reason
7.	7;	$\vdash \text{group}(F, *) \wedge \text{subgroup}(U, F, *) \wedge \text{unit}(F, 1, *) \wedge \text{unit}(U, 1_U, *)$	(Hyp)
8.	7;	$\vdash U \subset F$	(Def-subgroup 7)
9.	7;	$\vdash 1_U \in U$	(Def-unit 7)
10.	7;	$\vdash \exists_x x \in U$	(\exists 9)
11.	;11	$\vdash u \in U$	(Hyp)
12.	7;11	$\vdash u * 1_U = u$	(Def-unit 7 11)
13.	7;11	$\vdash u \in F$	(Def-subset 8 11)
14.	7;11	$\vdash 1_U \in F$	(Def-subset 8 9)
15.	7;11	$\vdash \text{semigroup}(F, *)$	(Def-group 7)
16.	7;11	$\vdash \text{solution}(u, u, 1_U, F, *)$	(Def-solution 12 13 14 15)
17.	7;11	$\vdash u * 1 = u$	(Def-unit 7 13)
18.	7;11	$\vdash 1 \in F$	(Def-unit 7)
19.	7;11	$\vdash \text{solution}(u, u, 1, F, *)$	(Def-solution 13 17 18 15)
20.	7;11	$\vdash 1 = 1_U$	(Th-solution 17 16 19)
21.	7;	$\vdash 1 = 1_U$	(Choice 10 20)
22.	;	$\vdash \text{group}(F, *) \wedge \text{subgroup}(U, F, *) \wedge \text{unit}(F, 1, *) \wedge \text{unit}(U, 1_U, *) \Rightarrow 1 = 1_U$	(Ded 7 21)

Figure 3: Abstracted Proof about Unit Element of Subgroups

of a discourse into an *attentional hierarchy*, since following the theory of Grosz and Sidner (Grosz and Sidner, 1986), there is a one-to-one correspondence between the intentional hierarchy and the attentional hierarchy. In this section, we illustrate the attentional hierarchy with the help of an example, which will be used to discuss reference choices later.

The input proof in Figure 3 is an ND style proof for the following theorem²:

Theorem:

Let F be a group and U a subgroup of F . If 1 and 1_U are unit elements of F and U respectively, then $1 = 1_U$.

The definitions of *semigroup*, *group*, and *unit* are obvious. $\text{solution}(a, b, c, F, *)$ stands for “ c is a solution of the equation $a * x = b$ in F .” Each line in the proof is of the form:

Label $\Delta \vdash$ *Conclusion* (*Justification reasons*)

where *Justification* is either an ND inference rule, a definition or theorem, which justifies the derivation of the *Conclusion* using as premises the formulas in the lines given as *reasons*. Δ can be ignored for our purpose.

We assume a reader will build up a (partial) proof tree as his model of the ongoing discourse. The corresponding discourse model after the completion of the presentation of the proof in Figure 3 is a proof tree shown in Figure 4. Note that the bars in Gentzen’s notion (Figure 2) are replaced by links for clarity. The numbers associated with nodes are the corresponding line numbers in Figure 4. Children of nodes are given in the order they have been presented. The circles denote nodes which are first

²The first 6 lines are definitions and theorems used in this proof, which are omitted.

derived at this place, and nodes in the form of small boxes are copies of some previously derived nodes (circled nodes), which are used as premises again. For nodes in a box, a referring expression must have been generated in the text. The big boxes represent *attentional spaces* (previously called *proof units* by the author), created during the presentation process.

The naturalness of this segmentation is largely due to the naturalness of the hierarchical planning operators. For example, attentional space U2 has two subordinate spaces U3 and U4. This reflects a natural shift of attention between a subproof that derives a formula of the pattern $\exists_x P(x)$ (node 10, $\exists_x x \in U$), and the subproof that proceeds after assuming a new constant u satisfying P (node 11, $u \in U$). When *PROVERB* opens a new attentional space, the reader will be given information to post an open goal and the corresponding premises. Elementary attentional spaces are often composed of multiple PCAs produced by consecutive navigation steps, such as U5 and U6. It is interesting to note that elementary attentional space cannot contain PCAs that are produced by consecutive planning operators in a pure hierarchical planning framework.

Adapting the theory of Reichman for our purpose (Reichman, 1985), we assume that each attentional space may have one of the following status:

- an attentional space is said to be *open* if its root is still an open goal.
- The *active* attentional space is the innermost attentional space that contains the local focus.
- The *controlling* attentional space is the innermost proof unit that contains the active attentional space.
- *precontrol* attentional spaces are attentional spaces that contain the controlling attentional space.

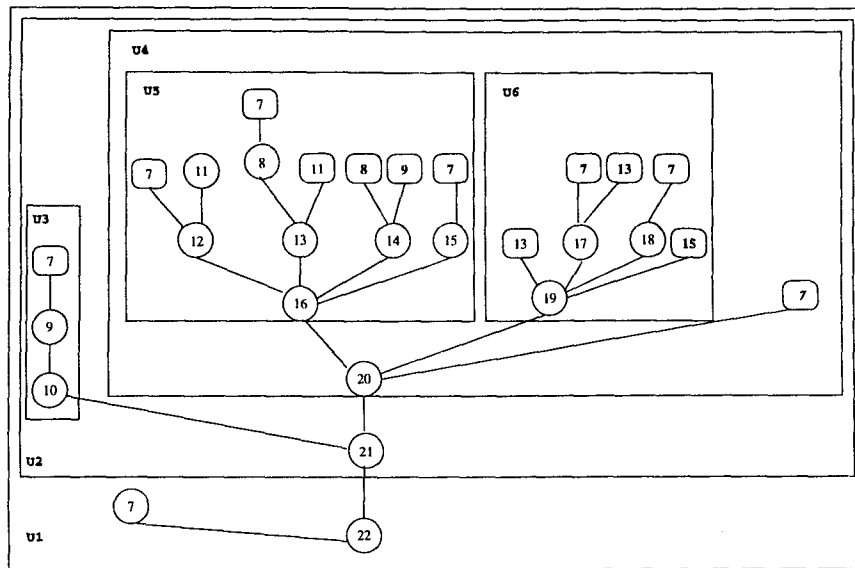


Figure 4: Proof Tree as Discourse Model

- *Closed spaces* are attentional spaces without open goals.

4 A Classification of Reference Forms

A referring expression should help a reader to identify an object from a pool of candidates. This section presents a classification of the possible forms with which mathematicians refer to conclusions previously proved (called reasons) or to methods of inference available in a domain.

4.1 Reference Forms for Reasons

Three reference forms have been identified by the author for *reasons* in naturally occurring proofs (Huang, 1990):

1. The *omit* form: where a reason is not mentioned at all.
2. The *explicit* form: where a reason is literally repeated.
3. The *implicit* form: By an implicit form we mean that although a reason is not verbalized directly, a hint is given in the verbalization of either the inference method, or of the conclusion. For instance, in the verbalization below

“Since u is an element in U , $u * 1_U = u$ by the definition of unit.”

the first reason of the PCA in Section 1, “since 1_U is the unit element of U ” is hinted at by the inference method which reads “by the definition of unit”.

Although omit and implicit forms lead to the same surface structure, the existence of an implicit hint in the other part of the verbalization affects a reader’s understanding.

4.2 Reference Forms for Methods

PROVERB must select referring expressions for methods of inference in PCAs as well. Below are the three reference forms identified by the author, which are analogous to the corresponding cases for reasons:

1. the *explicit* form: this is the case where a writer may decide to indicate explicitly which inference rule he is using. For instance, explicit translations of a definition may have the pattern: “*by the definition of unit element*”, or “*by the uniqueness of solution*.” ND rules have usually standard verbalizations.
2. the *omit* form: in this case a word such as “thus” or “therefore” will be used.
3. The *implicit* form: Similar to the implicit form for the expression of reasons, an implicit hint to a domain-specific inference method can be given either in the verbalization of the reasons, or in that of the conclusion.

5 Reference Choices in *PROVERB*

5.1 Referring to Reasons

Because reasons are intermediate conclusions proved previously in context, their reference choices have much in common with the problem of choosing *anaphoric* referring expressions in general. To account for this phenomenon, concepts like activat-

edness, foregroundness and consciousness have been introduced. More recently, the shift of focus has been further investigated in the light of a structured flow of discourse (Reichman, 1985; Grosz and Sidner, 1986; Dale, 1992). The issue of salience is also studied in a broader framework in (Pattabhiraman and Cercone, 1993). Apart from salience, it is also shown that referring expressions are strongly influenced by other aspects of human preference. For example, easily perceivable attributes and basic-level attributes values are preferred (Dale and Haddock, 1991; Dale, 1992; Reiter and Dale, 1992).

In all discourse-based theories, the update of the focus status is tightly coupled to the *factoring* of the flux of text into segments. With the segmentation problem settled in section 3, the DRCC module makes reference choices following a discourse theory adapted from Reichman (Reichman, 1985). Based on empirical data, Reichman argues that the choice of referring expressions is constrained both by the *status of the discourse space* and by the object's *level of focus* within this space. In her theory, there are seven status assignments a discourse space may have. Within a discourse space, four levels of focus can be assigned to individual objects: *high*, *medium*, *low*, or *zero*, since there are four major ways of referring to an object using English, namely, by using a pronoun, by name, by a description, or implicitly.

Our theory uses the notions of *structural closeness* and *textual closeness*, and takes both of these factors into account for argumentative discourse.

5.1.1 Structural Closeness

The *structural closeness* of a reason reflects the foreground and background character of the innermost attentional space containing it. Reasons that may still remain in the focus of attention at the current point from the structural perspective are considered as *structurally close*. Otherwise they are considered as *structurally distant*. If a reason, for instance, is last mentioned or proved in the active attentional space (the subproof which a reader is supposed to concentrate on), it is likely that this reason still remains in his focus of attention. In contrast, if a reason is in a closed subproof, but is not its conclusion, it is likely that the reason has already been moved out of the reader's focus of attention. Although finer differentiation may be needed, our theory only distinguishes between reasons residing in attentional spaces that are *structurally close* or *structurally distant*. DRCC assigns the *structural status* by applying the following rules.

1. Reasons in the *active* attentional space are structurally close.
2. Reasons in the *controlling* attentional space are structurally close.
3. Reasons in *closed* attentional spaces:
 - (a) reasons that are the root of a closed attentional space immediate subordinate to the active at-

tentional space are structurally close.

- (b) Other reasons in a closed attentional space are structurally distant.

4. Reasons in *precontrol* attentional spaces are structurally distant.

Note that the rules are specified with respect to the innermost proof unit containing a proof node. Rule 3 means that only the conclusions of closed subordinated subproofs still remain in the reader's focus of attention.

5.1.2 Textual Closeness

The *textual closeness* is used as a measure of the level of focus of an individual reason. In general, the level of focus of an object is established when it is activated, and decreases with the flow of discourse. In Reichman's theory, although four levels of focus can be established upon activation, only one is used in the formulation of the four reference rules. In other words, it suffices to track the status *high* alone. Therefore, we use only two values to denote the level of focus of individual intermediate conclusions, which is calculated from *textual distance* between the last mentioning of a reason and the current sentence where the reason is referred to.

5.1.3 Reference Rules

We assume that each intermediate conclusion is put into high focus when it is presented as a newly derived conclusion or cited as a reason supporting the derivation of another intermediate result. This level of focus decreases, either when a attentional space is moved out of the foreground of discussion, or with the increase of textual distance. The DRCC component of *PROVERB* models this behavior with the following four *reference rules*.

Referring Expressions for Reasons

1. If a reason is both structurally and textually close, it will be omitted.
2. If a reason is structurally close but textually distant, first try to find an implicit form; if impossible, use an explicit form.
3. If a reason is structurally distant but textually close, first try to find an implicit form; if impossible, omit it.
4. An explicit form will be used for reasons that are both structurally and textually far.

Note that the result of applying rule 2 and rule 3 depends on the availability of an implicit form, which often interacts with the verbalization of the rest of a PCA, in particular with that of the inference method. Since the reference choice for methods is handled independent of the discourse segmentation (Huang, 1996), however, it is not discussed in this paper.

Fourteen PCAs are generated by the macroplanner of *PROVERB* for our example in Figure 3. The

microplanner and the realizer of *PROVERB* finally produces:

Proof:

Let F be a group, U be a subgroup of F , 1 and 1_U be unit elements of F and U , respectively. According to the definition of unit element, $1_U \in U$. Therefore there is an X , $X \in U$.

Now suppose that u is such an X . According to the definition of unit element, $u * 1_U = u$. Since U is a subgroup of F , $U \subset F$. Therefore $1_U \in F$. Similarly $u \in F$, since $u \in U$. Since F is a group, F is a semigroup. Because $u * 1_U = u$, 1_U is a solution of the equation $u * X = u$.

Since 1 is a unit element of F , $u * 1 = u$. Since 1 is a unit element of F , $1 \in F$. Because $u \in F$, 1 is a solution of the equation $u * X = u$. Since F is a group, $1_U = 1$ by the uniqueness of solution.

Some explanations are in order. *PROVERB*'s microplanner cuts the entire text into three paragraphs, basically mirroring the larger attentional spaces U3, U5 and U6 in Figure 4. Since nodes 22 and 21 are omitted in this verbalization, node 20 (the last sentence) is merged into the paragraph for U6.

Let's examine the reference choices in the second last sentence:

Because $u \in F$, 1 is a solution of the equation $u * X = u$.

which is actually line 19 in Figure 3 and node 19 in Figure 4. Among the four reason nodes 13, 17, 18, 15, only node 13 is explicitly mentioned, since it is in a closed attentional space (U5) and is mentioned five sentences ago. Node 17 and 18 are in the current space (U6) and was activated only one or two sentence ago, they are therefore omitted. Node 15 is also omitted although also in the same closed space U5, but it was mentioned one sentence after node 13 and is considered as near concerning textual distance.

6 Conclusion

This paper describes the way in which *PROVERB* refers to previously derived results while verbalizing machine-found proofs. By distinguishing between hierarchical planning and focus-guided navigation, *PROVERB* achieves a natural segmentation of context into an attentional hierarchy. Based on this segmentation, *PROVERB* makes reference decisions according to a discourse theory adapted from Reichman for this special application.

PROVERB works in a fully automatic way. The output texts are close to detailed proofs in textbooks and are basically accepted by the community

of automated reasoning. With the increasing size of proofs which *PROVERB* is getting as input, investigation is needed both for longer proofs as well as for more concise styles.

Although developed for a specific application, we believe the main rationales behind of our system architecture are useful for natural language generation in general. Concerning segmentation of discourse, a natural segmentation can be easily achieved if we could distinguish between language generation activities affecting global structure of attention and those only moving the local focus. We believe a global attentional hierarchy plays a crucial role in choosing reference expressions beyond this particular domain of application. Furthermore, it turned out to be also important for other generation decisions, such as paragraph scoping and layout. Finally, the combination of hierarchical planning with local navigation needs more research as a topic in its own right. For many applications, these two techniques are a complementary pair.

Acknowledgment

Sincere thanks are due to all three anonymous reviewers of ACL/EACL'97, who provided valuable comments and constructive suggestions. I would like to thank Graeme Hirst as well, who carefully read the final version of this paper.

References

- Chester, Daniel. 1976. The translation of formal proofs into English. *Artificial Intelligence*, 7:178-216.
- Dale, Robert. 1992. *Generating Referring Expressions*. ACL-MIT Press Series in Natural Language Processing. MIT Press.
- Dale, Robert and Nicholas Haddock. 1991. Content determination in the generation of referring expressions. *Computational Intelligence*, 7(4).
- Edgar, Andrew and Francis Jeffrey Pelletier. 1993. Natural language explanation of natural deduction proofs. In *Proc. of the first Conference of the Pacific Association for Computational Linguistics*, Vancouver, Canada. Centre for Systems Science, Simon Fraser University.
- Gentzen, Gerhard. 1935. Untersuchungen über das logische Schließen I. *Math. Zeitschrift*, 39:176-210.
- Grosz, Barbara J. and Candace L. Sidner. 1986. Attention, intentions, and the structure of discourse. *Computational Linguistics*, 12(3):175-204.
- Hovy, Eduard H. 1988. *Generating Natural Language under Pragmatic Constraints*. Lawrence Erlbaum Associates, Hillsdale, New Jersey.
- Huang, Xiaorong. 1990. Reference choices in mathematical proofs. In L. C. Aiello, editor, *Proc. of*

- 9th European Conference on Artificial Intelligence*, pages 720–725. Pitman Publishing.
- Huang, Xiaorong. 1994. Planning argumentative texts. In *Proc. of COLING-94*, pages 329–333, Kyoto, Japan.
- Huang, Xiaorong. 1996. *Human Oriented Proof Presentation: A Reconstructive Approach*. Infix, Sankt Augustin.
- Huang, Xiaorong and Armin Fiedler 1997. Proof Verbalization as an Application of NLG. In *Proc. of IJCAI-97*, Nagoya, Japan, forthcoming.
- Kilger, Anne and Wolfgang Finkler. 1995. Incremental generation for real-time applications. Research Report RR-95-11, DFKI, Saarbrücken, Germany.
- McDonald, David D. 1983. Natural language generation as a computational problem. In *Brady and Berwick: Computational Models of Discourse*. MIT Press.
- McKeown, Kathleen. 1985. *Text Generation*. Cambridge University Press, Cambridge, UK.
- Moore, Johanna and Cécile Paris. 1989. Planning text for advisory dialogues. In *Proc. 27th Annual Meeting of the Association for Computational Linguistics*, pages 203–211, Vancouver, British Columbia.
- Ochs, Elinor. 1979. Planned and unplanned discourse. *Syntax and Semantics*, 12:51–80.
- Paris, Cécile. 1988. Tailoring object descriptions to a user's level of expertise. *Computational Linguistics*, 14:64–78.
- Pattabhiraman, T. and Nick Cercone. 1993. Decision-theoretic salience interactions in language generation. In Ruzena Bajcsy, editor, *Proc. of IJCAI-93*, volume 2, pages 1246–1252, Chambéry, France. Morgan Kaufmann.
- Reichman, Rachel. 1985. *Getting Computers to Talk Like You and Me. Discourse Context, Focus, and Semantics*. MIT Press.
- Reiter, Ehud and Robert Dale. 1992. A fast algorithm for the generation of referring expressions. In *Proc. of COLING-92*, volume 1, pages 232–238.
- Reithinger, Norbert. 1991. *Eine parallele Architektur zur inkrementellen Generierung multimodaler Dialogbeiträge*. Ph.D. thesis, Universität des Saarlandes. Also available as book, Infix, Sankt Augustin, 1991.
- Sibun, Penelope. 1990. The local organization of text. In K. McKeown, J. Moore, and S. Nirenburg, editors, *Proc. of the fifth international natural language generation workshop*, pages 120–127, Dawson, Pennsylvania.