Expressing quantifier scope in French generation

Pierre-Joseph GAILLY *

Computer Science Department, University of Liège, B4000 Liège, Belgium

Abstract

In this paper we propose a new method to express quantification and especially quantifier scope in French generation. Our approach is based on two points: the identification of the sentence components between which quantifier scope can indeed be expressed and a mechanism to reinforce the expression of quantifier scope. This approach is being integrated in a written French generator, called Hermès, which will become the generator of a portable natural language interface.

1 Introduction

Natural language interfaces are being applied to various applications; most of them, especially data base and expert system interfaces, require the production of precise texts. One of the key problems of precision is the correct expression of quantifier type and scope.

Our research has led us to identify a new **practical** method to express quantification in a precise and natural way and to implement our approach in the portable written French generator called Hermès (/Gailly 87a/ and /Gailly 87b/) to be integrated in a portable natural language interface (/Binot, Gailly & Ribbens 86/). The generator's goal is to produce French sentences from a message (a representation of their meaning) expressed in Formula, a representation language based on an extension of first order predicate calculus. The main point of our approach is the identification of the importance of properly choosing the determiners in noun phrases in conjunction with the selection of the surface structure of the sentences produced. However, we must stress the fact that we are not trying to provide a theoretical and comprehensive study of how quantification is expressed in natural language (such as in the work of /Cooper 83/).

In the rest of this paper we will first discuss the difference between the problems raised by quantification in parsing and generation, then discuss the present approaches to the problem and finally present our own approach.

2 The difference between understanding and generating quantifier scope

In understanding, the problem is, starting from the input sentence, to identify the quantifier scope intended by the user. In presence of an ambiguous sentence, the system will use all available information (i.e. syntactic structure, word choice and order, semantic and pragmatic knowledge, context and dialog history) to find the most plausible quantifier order. Consider:

Ex. 1 *Un homme meurt tous les soirs.*
  (A man dies all evenings.[1])

Ex. 2 *Chaque soir un homme meurt.*
  (Every evening a man dies.)

Although the first example is poorly expressed, a clever understanding system should comprehend both sentences in the same way, as human beings do, because the interpretation "*The same man dies every evening.*" is nonsensical. If the verb "*meurt*" ("*dies*") is replaced by "*mange chez Paul*" ("*eats at Paul's place*"), the two sentences have different meanings and the system should understand them in different ways.

In generation, the goal is to express the quantifier scope imposed by the input message correctly and as clearly as possible in order to avoid ambiguity.

We should not rely on the user's domain knowledge but provide him with clear clues about quantifier scope. Therefore, we should never produce unclear sentences such as example 1.

3 Existing approaches to the problem

Few authors considered the problem of expressing quantifier scope correctly and in a natural way. Surveys of the field (e.g. /Mann 82/ or /McKeown 86/) or even workshops proceedings (e.g. /Appelt 85/) do not mention the problem. At present, the problem seems to be solved in two ways:

- By indenting the text produced, (especially in database query reformulation), in order to show the structure of the text (scope of quantifiers and propositional connectives) graphically. Though this method can prove useful in small scale applications (e.g. /De Roeck & Lowden 86/ and some commercial systems), in fact it avoids the problem.

- By using word order[2] as suggested by /Chester 76/ and /Grishman 86/. The surface structure is chosen in such a way that the variable bound by the most dominating quantifier is generated as the leftmost noun phrase and so on.

The "word order" method is not satisfactory. Though word order is kept unchanged in the following examples, different scopes are clearly indicated by choosing adequate determiners:

Ex. 3 *Chaque chien poursuit un chat.*
  (Each dog is chasing a cat.)

Ex. 4 *Tous les chiens poursuivent un même chat.*
  (All dogs are chasing a same cat.)

As far as French is concerned, /Colmerauer 77/ proposed a more extensive treatment of the subject for parsing. His approach is based on the syntactic function of noun phrases and word order. Again, this method is not satisfactory: determiner selection allows to express different scopes, keeping word order and syntactic functions unchanged:

Ex. 5 *J'ai rencontré l'armateur de chaque navire.*
  (I met the owner of each ship.)

Ex. 6 *J'ai rencontré l'armateur de tous les navires.*
  (I met the owner of all ships.)

4 Our approach

Our approach is based on an integrated choice of both the surface structure and the determiners associated with noun phrases. We believe that scope expression is ruled by the following facts:

- A direct[3] dominance relationship cannot be expressed between any pair of noun phrases occurring in the sentence but only between noun phrases involved in specific syntactic structures.

- To each of these specific structures can be attached a **weak** default scope rule providing a partial and default ordering of the different noun phrases involved in the structure.

- Adequate determiner selection (through a mechanism we call scope expression reinforcement) can either override the default ordering given by scope rules or turn the partial ordering into a total one if necessary.

---

[1]Though we have tried to provide a faithful translation of the examples, we have kept the French sentence structure in the English version; therefore the translation can sometimes prove poor.

[2]This rule was proposed for English. Examples will show that it is not applicable, neither in French nor in English.

[3]That is the scope relationship between two consecutive quantifiers.

In the next section we will introduce the principle of scope expression reinforcement. The limits of scope expression will be considered in section 6. Section 7 will describe the default scope rules and provide examples of the use of these rules as well as of reinforcement.

5 Scope expression reinforcement by determiner selection

Let us consider some simple ways of expressing universal and existential quantifiers.

Several determiners express universal quantification. The definite article *"le"* or *"les"* (*"the"*) weakly express universal quantification. *"Chaque"* (*"each"* or sometimes *"every"*)[4] and *"tous les"* (*"all the"*) both strongly indicate that all entities specified by the noun phrase participate in the semantic relationship with the element (e.g. the verb) they depend from. But *"chaque"* adds a distributive feature which indicates that there is a state or action for each entity described in the noun phrase and therefore has the effect of giving a dominant position to the corresponding universal quantifier. Using *"tous les"* adds a collective[5] feature and does not give this domination effect. This leaves the possibility for another quantifier to "take" the dominance.

Other determiners express existential quantification, such as *"un"* (*"a"*), *"des"*[6] and *"un même"* (*"a same"* or *"one"*)[7] etc. *"Un"* does not give any domination effect; obviously, *"un même"* indicates that the entity described in the noun phrase and participating to the semantic relationship is the same for all states or actions described. This gives dominance to the corresponding existential quantifier.

Let us now consider some examples and evaluate their quality as generated text:

Ex. 7 *Chaque homme est dans une pièce.*
(Each man is in a room.)

Ex. 8 *Tous les hommes sont dans une même pièce.*
(All men are in a same room.)

Ex. 9 *Tous les hommes sont dans une pièce.*
(All men are in a room.)

Ex. 10 *Chaque homme est dans une même pièce.*
(Each man is in a same room.)

Obviously, examples 7 and 8 are clear and unambiguous; they provide distinct and adequate expressions for the two possible quantifier scopes. Example 9 is poor. Most readers understand it as having the same meaning as example 7 but some could understand it as example 8. Such a sentence is potentially ambiguous and generating such a sentence should be avoided. Example 10 is also poor because the dominance indications given by *"chaque"* and *"un même"* are contradictory.

These examples have shown that the imposed quantifier scope can be expressed correctly and clearly by reinforcing the expression of the dominating quantifier.

6 Limits of scope expression

We consider that scope can only be expressed between noun phrases which are linked by a direct syntactic relationship. In French, quantifier dominance can only be expressed between:

- the different elements directly linked to the same verb (subject, direct or indirect object, agent and other complements),

- the different elements directly linked to the head of a noun phrase, that is between:
  - the head and its complements,
  - the head and the elements directly linked to the main verb of a relative clause,
  - the head and the complements of an adjective,
- a main clause and a conjunctive subordinate clause.

Our study has shown that these hypotheses on syntactic structures can provide a criterion applicable to the structure of the input message. This criterion helps us to detect situations where dominance cannot be expressed properly in one sentence and to decide, in the process of generation, to split the message into several parts which can be expressed precisely.

7 Default scope rules and examples

Default scope rules indicate how the reader perceives quantifier scope whenever no reinforcement is used. We will propose a default scope rule for each of the syntactic structures given in section 6.

7.1 Simple sentences

The quantification associated to noun phrases will respect the following default quantification priority scheme:
1. non essential[8] complements placed before the verb,
2. subject,
3. non essential complements placed between the verb and the essential complements,
4. essential complements,
5. non essential complements placed after the essential ones.

This rule bears some resemblance with the "left to right order" rule or with Colmerauer's proposal but in fact it is based more on semantic considerations than syntactic ones. The usual word position in a French sentence (i.e. subject, verb, essential complements, non essential complements) reflects the semantic proximity of these elements with respect to the verb: a nearer position corresponds to a greater proximity (points 2, 4 and 5 of the above rule). We consider that placing some non essential complement in an unusual position (points 1 and 3) raises its importance with a corresponding raise in its scope priority.

Let us also point out that this rule only provides a partial ordering. For example, if a verb has two essential complements, the rule does not order them.

In /Gailly 87b/ we have given an extensive list of examples showing how complex[9] quantification structures can be expressed in a natural way. We will just quote a few complex examples taken from the classical suppliers-parts database and requiring reinforcement:

Ex. 11 $\forall$ *supplier* $\exists$ *article* $\exists$ *date* $\forall$ *customer*
Donnez-moi la liste des fournisseurs qui livrent un même article à tous les clients à une même date.
(Give me the list of all suppliers which deliver a same article to all customers at a same date.)

The dominance of *"article"* and *"date"* has been reinforced by the use of *"un même"*.

Ex. 12 $\forall$ *date* $\exists$ *article* $\forall$ *supplier* $\exists$ *customer*
Donnez-moi la liste des dates où un même article a été livré par chaque fournisseur à au moins un client.
(Give me the list of dates when a same article was delivered by each supplier to at least one customer.)

The dominance of *"article"* has been expressed by selecting a subject position and by using *"un même"*; the relative dominance of *"supplier"* has been expressed by using *"chaque"*.

---

[4]In French, there is only one word *"chaque"* to express the meaning of both English words *"each"* and *"every"*.

[5]Though the collective feature added by *"tous les"* is weaker then the distributive one added by *"chaque"*.

[6]This is the plural form of the indefinite article; it has no equivalent in English.

[7]In French, we cannot choose between the indefinite article *"a"* and the numeral *"one"* as they are both expressed by *"un"*.

[8]Essential complements are those verb complements whose omission would make the sentence nonsensical. For example the verb *"aller"* (*"to go"*), in the meaning implying some movement, requires the specification of the target location. This semantic notion seems more adequate than Colmerauer's proposal of "near complements" which is based on the syntactic function of the complement.

[9]Involving alternating existential and universal quantifiers. Commuting quantifiers obviously do not raise any problem.

183

7.2 Noun phrases

We have studied the cases of noun's complements, relative clauses and complements of an adjective. The head of the noun phrase has default dominance on the subordinated noun phrases occurring inside the noun phrase.

The scope relationship between the quantifiers associated with the noun phrases subordinated to the head can be treated by default scope rules similar to the one given for simple sentences. These rules assume that the head dominates and that there is no "interleaving" between the quantifier scope inside the noun phrase and the quantifier scope in the sentence (or clause) containing the noun phrase.

We also considered the problem of expressing the dominance of the quantification associated with one of the subordinated noun phrases over the quantification associated with the head. This involves an interaction between the quantifiers at the sentence level and those at the noun phrase level. This expression is not always possible (because the default dominance of the head appears to be too strong). This problem can only be treated satisfactorily in the case of a *"complément déterminatif"* (equivalent to the use of a genitive or a *"of"* construct in English) and of relative clauses. Reinforcement is necessary for the quantifier associated with the subordinated noun phrase. The following two examples illustrate this last point.

Ex. 13 $\exists$ *play* $\forall$ *student* (Default scope rule)
J'ai lu la pièce que tous les élèves ont écrite.
(I read the play all the students wrote.)

Ex. 14 $\forall$ *student* $\exists$ *play* (Universal quantifier reinforced)
J'ai lu la pièce que chaque élève a écrite.
(I read the play each student wrote.)

7.3 Conjunctive subclauses

We believe that it is not possible to provide a general and reliable default scope rule assigning a priority scheme to the different noun phrases appearing in the main clause and in the conjunctive subclause(s). Anyhow, two cases can be treated satisfactorily:

1. if one of the clauses does not contain any quantification or if all the quantified noun phrases involved appear in the same clause (either explicitly or repeated in the form of a pronoun) then the simple sentence approach can be applied.

2. if the sentence contains only two quantifiers, then reinforcement of the dominating quantifier can be used.

8 Conclusions

We have shown how quantification scope can be expressed in a natural and precise way in French generation. We have insisted on two important aspects: selecting sentence structure and determiners. We have also shown some limitations of natural language expression of quantifier scope. In some cases input message restructuring can solve the problem but in others we reach the intrinsic limits of natural language.

The methodology we propose is, at the time of writing, being implemented in the French generator Hermès. Further detail on our method (including a discussion of the interaction of quantifier scope with negation) can be found in /Gailly 88/.

As further conclusions we will now consider two possible extensions to our approach and a possible application to understanding:

- Application to other languages. We have not studied this problem in detail but we are convinced that correct quantifier scope expression for other languages will have to consider both the influence of sentence structure and of determiner choice. The set of default scope rules will have to be adapted to account for the syntax of the language. The reinforcement mechanism will have to incorporate the determiners available in the particular language.

- Application to domains where a precise expression is not required. In our approach we provide obvious indications of quantifier scope; this leads to a precise style compatible with the target applications. In other domains, a weaker expression might be required, that is reinforcement should not be used whenever the quantifier scope is obvious from the context. Our method could still be used but the reinforcement mechanism would only be invoked whenever the lack of reinforcement would produce an ambiguous sentence. This kind of decision would be taken by an expert module using pragmatic knowledge as well as a complete domain and discourse model.

- Use of our heuristics in scoping algorithms for parsing. The algorithm proposed by /Hobbs & Shieber 87/, for example, produces all possible scopings consistent with the logical structure of English. As the authors suggest, this algorithm can be extended in order to produce scopings in decreasing order of plausibility: our default scope rules could be used as syntactic and order rules. Our observation of the dominating effect of some determiners could also be used as lexical rules but the authors consider that lexical rules are difficult to integrate in their algorithm.

9 Acknowledgments

References

D. E. Appelt 85 "Proceedings of the 1984 international workshop on language generation.", C.S.L.I. Report, 1985.

J.-l. Binot, P.-J. Gailly & D. T. Ribbens 86 "Eléments d'une interface portable et robuste pour le français écrit", Proc. of the "Huitièmes Journées Francophones de l'Informatique", Grenoble, Jan. 1986.

D. Chester 76 "The translation of formal proofs into English", Artificial Intelligence, Vol 7, pp 261-278, 1976

A. Colmerauer 77 "Un sous-ensemble intéressant du français", Rapport 363, Groupe d'intelligence artificielle, Université d'Aix-Marseilles II, 1977

A. Cooper 83 "Quantification and syntactic theory", D. Reidel Publishing Company, 1983.

A. N. De Roeck & B. G. T. Lowden 86 "Generating English paraphrases from formal relational calculus expressions", Proc. of COLING 86, pp.581-583, 1986.

P.-J. Gailly 87a "Une première et succincte présentation du système de génération du français Hermès", University of Liège, Computer Science Department, Internal report, Jan. 1987.

P.-J. Gailly 87b "Le système de génération du français Hermès, Objectifs, Problèmes et Principes", University of Liège, Computer Science Department, Working report, July 1987.

P.-J. Gailly 88 "Hermès: Un générateur portable du français écrit" University of Liège, Computer Science Department, In preparation.

R. Grishman 86 "Computational linguistics", Cambridge University Press, 1986

J. R. Hobbs & S. M. Shieber 87 "An algorithm for generating quantifier scopes", Computational Linguistics, Vol.13, pp.47-63, 1987

K. R. McKeown 86 "Language generation: applications, issues and approaches", Proc. of I.E.E.E., Vol.74, No.7, July 1986

W. C. Mann 82 "Text generation, Panel report", A.J.C.L., Vol.8, No.2, April 1982