HINTING BY PARAPHRASING IN AN INSTRUCTION SYSTEM

Vladimir PERICLIEV
Svjatoslav BRAJNOV
Irina NENOVA

Department of Mathematical Linguistics
Institute of Mathematics with Computing Centre, bl.8
Bulgarian Academy of Sciences
1113 Sofia, Bulgaria

Abstract

Previous work has emphasized the need for paraphrases as means of ensuring a feedback with a system. In this paper, we discuss how a paraphrase may be used as a heuristic device, viz. as a hint. We describe an experimental instruction system in mathematics incorporating this feature. The system accepts a restricted class of algebraic story problems, formulated in non-stylized Bulgarian language, and is capable of solving them and providing one or more "hinting" paraphrases, that is, paraphrases alleviating their formalisation (=translation into equations).

1. INTRODUCTION

Previous work has emphasized the need for paraphrases as means of ensuring a feedback with the system. For example, question-answering systems, before responding, paraphrase the requests formulated in natural (Kaplan 1979, McKeown 1983) or a formal language (de Roeck, Lowden 1986) in order that the user ascertain that his/her question has really been correctly understood. This step is necessary to avoid (possibly) costly searches in the data base for requests that have never been made. An additional reason is that sometimes the format of the retrieved information also cannot clear up a potential misunderstanding (Thomas, Gould 1975).

However, there are other applications, different from feedback, to which a paraphrastic facility may profitably contribute. In the paper, we discuss how a paraphrase may be used as a heuristic device, viz. as providing a hint, in an instruction system in mathematics.

The paper is organized in the following way. Sect. 2 is a brief overview of the instruction system incorporating this feature. Sects. 3 and 4 describe respectively some general requirements to a hinting paraphrase and the measures we have taken to satisfy them in the system. Sect. 5 discusses an example, and Sect. 6, some implementation details.

2. SYSTEM OVERVIEW

We have designed an experimental instruction system in mathematics. The system operates in a limited domain: it is capable of solving a restricted class of story problems in algebra for Secondary Schools in Bulgaria (the so called "number problems"). The system accepts non-stylized story problems in Bulgarian as they can be found in mathematical textbooks or are spontaneously formulated by the user. It solves the problem, and is capable of providing either of the following 3 options:

(a) Result (resultant number(s) are displayed).

(b) Equations (the equation(s) to which problems translate are displayed).

(c) Paraphrases (one or more "hinting" paraphrases are displayed, together with the text of the original problem).

All the three options serve as (different degrees of) hinting needed in case the users (Secondary School pupils) have problems with finding a solution. Furtheron, we focus on problems concerned with the hinting paraphrastic facility of the system.

3. GENERAL REQUIREMENTS TO A HINTING PARAPHRASE

The profit of using a paraphrase, or a "reformulation", of a problem as a heuristic tool has been emphasized by researchers in heuristics, pedagogy and psychology of education. Nevertheless, such a possibility is usually beyond the scope of instruction systems (Sleeman, Brown 1982, Weischedel et al 1978, Pulman 1984).

The question still remains as to what can count as a hinting pararhrase (HP) (obviously , not <u>any</u> paraphrase can serve this purpose equally well). Basing ourselves on research in mathematical pedagogy and psycholinguistics (since conceptual and linguistic structures in this early age are known to be strongly interdependant), we derived the following general requirements to a HP:

1. The HP should simplify the original problem (OP) as regards the users of the system (by this we mean simplification of OP in both conceptual and linguistic aspects with respect to the task assigned, viz. to fomalise (=translate into equations) the OP.

2. The HP should be different from the OP (this requirement is self-evident).

3. The HP should keep close to the OP from a conceptual and linguistic viewpoints (this is to ensure that the users conceive the "sameness" of HP and OP).

Since the most important task of the HP is to simplify the translation of verbally formulated problems into equations (solving these equations being as a rule unproblematic for children), we took the following general solution regarding an "appropriate" HP: An "appropriate" HP to a problem is the one that can be, somewhat metaphorically expressed, literally translated into the respective equations of this problem. Obviously, this would, to the greatest extent possible, simplify the OP (in the sense in which in the translation from one NL to another, the easiest to perform is the literal translation). This decision is further supported by the fact that pupils usually translate to themselves the OP into intermediary language which is most close to the equations derivable from this problem.

4. CONSTRAINTS ON "APPROPRIATE" HPs

From what is stated above, a number of specific constraints on the content and form of the HP can be derived. We briefly mention them below in connection with two of the major decisions that have to be made in a generation process: first, making a decisions as to the discourse structure of the HP (i.e. determining what and when to say, or an ordered message to be conveyed), and, secondly, making a decision as to the verbal formulation of the discourse structure of the HP (i.e. determining how to express this information in Bulgarian, what syntactic structures to use, what lexemes, etc.).

At the first stage, we should gain in conceptual, and, at the second stage, in linguistic simplification, thus approximating the requirement as to the literalness we have imposed.

4.1. Discourse structure

In the light of our aims, it is clear that the discourse structure of the HPs should be standardized, or canonized. This means that we need not be concerned (like most scholars working on discourse organization, e.g. Mann 1984, McKewon 1985) with describing types of discourse structures of actual texts in the domain of interest, but rather with prescribing a discourse pattern that satisfies the discourse goal.

Each of the texts in our domain, story problems in algebra for Secondary Schools, is known to be characterized by unknown(s) (i.e. what is looked for in the problem), and condition(s) (i.e. the equation(s), relating the unknown(s), or variables, to the given(s), or constants, in the problem). Some problems also involve auxiliary unknown(s) (i.e. further unknown(s), often mentioned in the problem formulation somewhat misleadingly (e.g. "...Another number is given.."), which have to be manipulated, but are not themselves part of the solution).

The discourse structure of the HPs, therefore, will have to reflect the basic conceptual constituents of the problems:

1. the unknown(s)

2. the auxiliary unknown(s) <optionally>

3. the condition(s),

in that particular order.

It may be noted that a lot of problems, as they are formulated in mathematics textbooks, do not actually satisfy this discourse schema: the unknowns are interspersed in the text, the unknown(s) and auxiliary unknown(s) are not explicitly discriminated, the conditions precede (auxiliary) unknown(s), etc.

For instance, a typical problem to be found in a textbook may begin as follows: "The sum of two numbers is 8..." Clearly, starting the problem formulation by a condition, instead of with declaring first the unknown(s), is misleading. Thus, notice that this problem may have quite different continuations, among which

...The first number is 2.
Which is the second?

in which we have just one unknown, or

...Their product is 12.
What are these numbers?

a version in which there are two unknowns.

The resolution of this local ambiguity requires additional intellectual effort on the part of the pupil, re-reading, etc., circumstances which our HPs should evade.

In addition to describing the major conceptual constituents of the problems, in the canonical discourse structure of the HPs, the conditions of problems themselves, usually compound propositions, should be broken down into parts. The ordering of these propositions should reflect the hierarchical structure of the conditions, starting with those at the lowest level of the hierarchy, proceeding with those at the next higher level, and so on.

4.2. Verbal formulation

The verbal formulation of the HPs includes the choice of syntactic and lexical structures

4.2.1. Syntactic structure

Only restricted classes of syntactic structure should be used in the HP. As a rule self-embedded, or recursive, constructions that appear in actual texts ("...the sum of the sum and difference..."), ellipsis and "confusing" anaphorical constructions should be avoided.

4.2.2. Lexical structure

The lexemes in the HP should directly, explicitly name operations, numbers and variables. Only one relation should be admissible in the HPs, viz. the relation "equal", but not the confusing for pupils "is greater than/with", etc.

In both syntactic and lexical choices, wherever possible, the HPs should stick to the verbal expressions in the OPs.

It will be clear from what is said how the HPs conform to the requirements of Sect. 3.

5. AN EXAMPLE

Below we give an example, worked by the system, which illustrates some of the simplifications of the OP achieved in the HP (for convenience, the OP and the HP are translated into English).

The OP is:

(1) If the sum of one number with another

(2) which is with 4 smaller than it

(3) is multiplied by 2,

(4) you will find the product of the second number with the number 5.

(5) Find the first number.

One of the HPs of this OP will look something like:

(1) A number is looked for.

(2) Another number is given.

(3) Add the two numbers.

(4) If you multiply the sum obtained with 2,

(5) you will find the product of the second number with the number 5.

(6) If from the first number you subtract 4,

(7) you will obtain the second number.

In comparison with the OP, the HP explicates the two numbers of the problem that will be further manipulated: first, the unknown, and, then, the auxiliary unknown. In clause (1) of the OP the operation of addition is implicitly given by its result ("the sum"), whereas in clause (3) of the HP the same operation is elaborated by an explicit mentioning of the particular arithmetical operation of addition . The imbedded relative clause (2) of the OP is expressed separately from the main sentence in the HP ((6) and (7) of the HP). This provides a possibility, reading the condition of the problem from left to right, to write down, sequentially and independently, the different equations. In the paraphrase of the relative clause (2) of the OP, the relation "is smaller than", known to be confusing for small children, is replaced by its corresponding operation "subtraction", and the pronominal reference (expressed in the English test with "it") is avoided. Notice also that (4) from the OP and (5) from the HP are phrased in the same way (thus preserving partial sameness of the OP and the HP).

The reorganised text of the HP can be seen to significantly simplify the OP (which will be particularly true for Secondary School children).

6. IMPLEMENTATION

Below we briefly describe some aspects of the implementation design.

The system comprises 3 modules:

(i) Analyser

(ii) Solver

(iii) Paraphraser.

The Analyser is a "traditional" semantic grammar, using hierarchically organised templates. The Solver solves the equations obtained as a result of the parsing phase (if the system is in a "Result" mode).

The generation process goes through two major phases. The paraphrastic facility of the system has two components, responsible for the tasks at these phases: the Canonizer, and the Generator.

In the first phase, the Canonizer constructs the discourse structure, or the canonical form, of the HP. The process includes the representation of the discourse structure into a sequence of elementary propositions, instantiated by the result derived by the Analysis module. This sequence begins with the proposition describing the unknown(s), and, optionally, propositions for auxiliary unknown(s). In the sequence follow the propositions describing conditions (=equations).

For example, as a result of the analysis of the OP, mentioned in Sect. 5, the following sequence is obtained:

```
equal(*(+(X,Y),2),*(Y,5))
equal(-(X,4),Y)
unknown(X)
```

The Canonizer shifts the last proposition unknown(X) at the begining of the sequence of propositions and adds yet another proposition auxiliary_unknown(Y). As a result

```
unknown(X)
auxliary_unknown(Y)
equal(*(+(X,Y),2),*(Y,5))
equal(-(X,4),Y)
```

is obtained.

Each compound proposition of the latter type is substituted with an equivalent <u>propositional expression</u>. In order to achieve this, all constituent propositions are substituted by variables, after which the simple proposition obtained is unified with the compound proposition.

In the above case, from the unification of the two compound propositions "equal" with the simple proposition equal(α,β), we obtain:

equal(*(+(X,Y),2),*(Y,5))=
=equal(α,β).{*(Γ,2)/α,*(Y,5)/β}.{+(X,Y)/Γ}
equal(-(X,4),Y)=equal(α,β).{-(X,4)/α,Y/β},

where the expressions in braces are substitutions.

The propositional expression thus describes the <u>process</u> of obtaining the compound proposition in question from simple propositions.

After the substitution of each compound proposition of the equivalent propositional expression, the following canonical representation obtains:

```
unknown(X)
auxiliary_unknown(Y)
equal(α,β).{*(Γ,2)/α,*(Y,5)/β}.{+(X,Y
)/Γ}
equal(-(X,4),Y)=equal(α,β).{-
(X,4)/α,Y/β}
```

The canonical representation used is easily seen to have certain advantages. On the one hand, it explicates all computations necessary for construction of the system of equations, and, on the other hand, it defines a <u>plan for verbalization</u>, to be used by the Generator, in which, first of all, the simple propositions are verbalized, then their verbalizations are used in the verbalization of the compound propositions at the next higher level of hierarchy, and so on. The text to be obtained following such a plan of verbalization can be literally translated into a system of equations by virtue of the fact that the text itself is generated in inverse order - from simple to compound propositions.

In the second phase of the process of generation of the HPs, the canonical form of the HPs is translated into Bulgarian text by the Generator. The Generator itself is a <u>unification grammar of templates</u> (the templates used for generation).

Each template describes a syntactic construction by means of particular wordforms, lexical classes and variables. Some of the templates are used to propagate anaphorical relations (definite NPs, or pronominal references).

As already mentioned, the Generator follows the plan for verbalization defined by the canonical representation. A set of <u>selection rules</u> governs the choice of particular templates, <u>before</u> unification begins. In case of alternatives as to the choice of a template, the Generator consults the derivational history of the analysis, which is kept in a special register, and selects the template, and the concrete verbal formulation, used in the OP (this ensuring partial "sameness" of HPs and OPs).

The system is implemented in PROLOG-2 and runs on IBM ATs and compatibles.

7. CONCLUSION

In the paper, we tried to show how a paraphrase can be used as a hinting tool in an instruction system in mathematics, and described a system incorporating this feature. In the current implementation, the system may give reasonably good paraphrases of the original problem, but still there is a lot to be desired, even abstracting from any real application for educational purposes. It is a rather difficult thing to make the "right" compromise between the simplification needed in such tasks and a nice verbal phrasing of the problems. We shall continue the work on the refinement of the system and on developing an explanation facility.

REFERENCES

Kaplan S., Cooperative Responses from a Portable Natural Language Data Base Query, Ph.D. Dissertation, University of Pennsylvania, 1979.

Mann W., "Discourse structures for text generation", COLING84, Stanford, 1984.

McKeown K., "Paraphrasing questions using given and new information", AJCL, 9, 1, 1983.

McKeown K., Text Generation: Using Discourse Strategies and Focus Constraints to Generate Natural Language Text, Cambridge University Press, 1985.

Pulman S., "Limited domain systems for language teaching", In: COLING84, Stanford, 1984.

de Roeck A., B. Lowden, "Generating English paraphrases from formal relational calculus expressions", In: COLING86, Bonn, 1986.

Sleeman D., J. Brown (eds), Intelligent Tutoring Systems, N.Y., Academic Press, 1982.

Thomas J., J. Gould, "A psychological study of query by example", In: Proceedings of NCC, 44, 1975.

Weischedel R. et al, "An AI approach to language instruction", Artificial Intelligence, 10, 3, 1978.