ANAPHORA IN MATHEMATICS AND IN COMPUTER SCIENCES EDUCATION*

Viviane Durand-Guerrier
Institut Montpelliérain Alexander Grothendieck, CNRS, Univ. Montpellier France
viviane.durand-guerrier@umontpellier.fr

Faiza Chellougui
Université de Carthage, Tunisie
chellouguifaiza@yahoo.fr

Simon Modeste
Institut Montpelliérain Alexander Grothendieck, CNRS, Univ. Montpellier France
simon.modeste@umontpellier.fr

Abstract: In this communication, we will discuss the difficulties met by students when dealing in mathematics with statement involving an Anaphora, and we will explore what happened in Informatics.

1 Introduction

Anaphora has been widely discussed in formal semantics. A classical example is known as "donkey-sentence". The "Donkeys Sentences" rewarded their name to a famous example due to Kamp (1981): how to represent in predicate calculus the sentence: "Every farmer who owns a donkey beats it" (1). According to Corblin (2002), "$\forall$X $\forall$Y [(farmer (X) donkey(Y) owns(X,Y) $\Rightarrow$ beats (X,Y)]" « is generally admitted as a correct representation of sentence (1) in first order logic" (p.91). Nevertheless, he argues that this is in contradiction with standard correspondence so that "a N" corresponds to $\exists$x, while "all N" correspond to "$\forall$x". For him, following Kamp (1981), the donkey sentences show that it is necessary to modify the language use for representing such sentences; the Discourse Representative Theory has been build in this purpose. What we learn as mathematics educators with "donkey sentences" is, that something that is usual in mathematics (i.e. considering that "a N" might be use to design a generic element, seems) to be unsuitable for "formal semanticians".

2 Anaphora in mathematics: an educational perspective

Although one could think that such anaphora are not supposed to appear in mathematical activity, we have shown in our research in mathematics education that they are likely to appear in a number of case. In mathematical practise, Anaphora is related with instability of letters' logical status: indeed, it is common that in mathematical texts, including texts addressed to students, the same letter is used with different logical status: bound variable free variable, generic element, constant. We illustrate this point by an example/

An example of students' difficulties related to Anaphora in calculus

The following item is the fourth one of a questionnaire comprising six that was submitted to 273 fresh university students (for a presentation of the questionnaire see Durand-Guerrier 2003, 16-20). As an introduction to the questionnaire, we indicated in introduction of the questionnaire that the given sentences were true statements to be used for answering the

* Presented at the ESSLLI 2017 workshop QUAD: Quantifiers and Determiners (Retoré, Steedman)

questions. The sentence for this item (*our translation from French*) is a classical theorem of analysis introduced at the transition between secondary and tertiary level.

The statement and some elements of a priori analysis

A4) (u_n) is the name of a sequence of real numbers determined by a recursive law of the type $u_{n+1} = f(u_n)$, where f is a continuous function on the set of real numbers. Then the statement below is true: *If the sequence (u_n) converges to the real number L, then L is a solution for equation (E): "$f(x) = x$".* Questions; What can be said about the convergence of the sequence (u_n) if: a) The equation (E) has no solution? b) The equation (E) has at least a solution? What can be said about eventual solutions for the equation (E) if: c) The sequence (u_n) converges? d) The sequence (u_n) does not converge?

The statement at stake is an informal written mathematical statement which, in a naive formalisation, could be considered as a sentence of type "if p, then q". However, this apparent simplicity is far from embedding all the information needed to interpret adequately this sentence. For this purpose, it will be necessary to unpack the logic of the statement (Selden & Selden, 1995). In order to provide such formalisation, we introduce three binary relations (two-places predicate). The first one, S(x, y) is interpreted by "x is a sequence, y is a continuous function, and for all n, $x_{n+1} = y(x_n)$ " ; the second one, R(x, z) is interpreted by "x is a sequence, z is a real number and x converges to z" ; the third one , T(y, z) is interpreted by "y is a continuous function, z is real number and z is a solution for equation $y(z)$ = z". Considering these predicates and getting rid of the quantification matters, a formalisation for sentence 4 could be " R(u, L) => T(f, L)" (1). At this step, if we intend to introduce quantifiers, we have to pay attention to logical status for letters u, f and L in the sentence 4, which is given as a true statement, that indeed it is, and consequently in formula (1).

Concerning u and f, they are clearly introduced as generic element, with u depending on f; this is a rather common way of doing at university: it prevents students from writing several quantifiers, and it is supposed to make easier the understanding of the sentence. So, we focus now our interest on letter L, for which the situation is more intricate. Indeed, 1. L has not been introduced; 2. L appears in a definite expression: the real number L ; 3. Sentence 4 is a conditional statement in which L appears both in antecedent and consequent. ; 4. u converges *if and only if* there exits L so that R(u, L). Regarding 1, 2 and 3 leads to contradictory conclusions: indeed, speaking of the real number L may suppose that L represents a constant, a particular real number well defined, the existence and the uniqueness of which are warranted, while not introducing L may indicate that L is a mute letter, a bounded variable in the scope of a universal quantifier, due to the common practise of implicit quantification of conditional statement (Durand-Guerrier 2003). Regarding 3 inclines to wonder if L has the same status in antecedent and consequent, which is generally the case in mathematics, and if yes, which status. Regarding 4 seems indicate that, in antecedent, L is a bounded variable in the scope of an existential quantifier. As a consequence, while introducing quantifiers, several possibilities appear according to the chosen interpretation for letter L. In the first interpretation, it is possible to consider that the letter u is useless, and so to get rid of. A second interpretation consists in considering that it is possible to change le logical status of letter L between the antecedent and the consequent of the conditional statements. A third interpretation consists in considering that L is a generic element.

We have shown (Durand-Guerrier, 2003) that these differences in logical status of letter in mathematical texts are source of difficulties for fresh university students. This was particularly clear through the results of the question A4 resented above: for the items A4-b and A4-c, less than 20% gave a correct answer, while nearly 80% gave a correct answer to items A4.a (application of modus tollens) and A4.c (application of the modus ponens) (Durand-Guerrier 2003, 21).

Some students' answers (our translation from French)

In the questionnaire, we asked students to justify carefully their answers; not all of them did it for this difficult question, nevertheless some students trying to explain why it was not possible to give an answer in item A4b and A4d give us precious indications. We focus here on explanations involving the letter L.

In several copies, L seems to be considered as a given element; it is the case in copy n°83:

> *"A4a: if (E) has no solution, then sequence u does not converge to L, but it might converge to another value L'.*
> *A4b: if (E) has at least one solution, then this solution is L and the sequence converges to L. In that case, lim u = L.*
> *A4c: If the sequence u is convergent, there exists at least a solution and this solution is L."*

In this last example, the existence of a solution seems to insure for the student the uniqueness, even when using the expression " at least".

Opposite, in copy n°17, the uniqueness in (E) is given as a necessary condition to convergence for *u*, while L seems once more to be given:

"*A4b: if the equation (E) as only one solution, it is L and the sequence u converges ; if the equation (E) has several solutions, then the sequence u does not converge, for it is only possible to converge in more than one point."*

In copy n°68, the uniqueness of the limit seems to be moved to equation:

> *"A4b: if the equation (E) has a solution, then u converge; if the equation (E) has more than one solution, then u diverge."*

For some students, it seems that sentence 4 asserts that "sequence *u* converge to L", and hence that they try to solve the contradiction with the premise in item A4a " (E) has no solution"; it appears clearly in copy n°258 :

> *A4a: the sequence u converges to L ; but without reaching value L.*

We can see in these examples, on the one hand, that students met difficulties to deal with letter L, and on the other hand, that mathematical knowledge are strongly involved in the answers, specially for items for which there is a lack of inference. The fact that the uniqueness is warranted for the limit of the sequence when it exists, but is not required for the equation, creates a dissymmetry hidden by the use of the letter L. Even if it could be considered that the choice of giving the statement introducing the letter L is not relevant, and introduces confusion, the donkey sentences remind us that anyway, the problem is serious and due partly to the structure of the sentence (e.g. Sansu 1997, Corblin 2002).

3 Is there Anaphora in computer sciences?

In informatics, the use of formal languages prevents *a priori* from the difficulties of interpretation (as the interpretation is given by the construction of the formal language) and we could think that Anaphora do not appear. Nevertheless, anaphora may appear in class when algorithms are given first in natural language; in such cases, the translation of the algorithm into a program requires to choose an interpretation. Relying on our results in mathematics education, we intend to investigate how students interpret instructions that look like Donkey Sentences, as (in a graph): "for every vertex having a loop, remove it". This instructions can be formalised (with P(l,v) is the predicate "the edge l is a loop of the vertex v") as : "for all vertices v, if there exists l such that P(l,v), then remove l", while a correct formalisation should be : "for all vertices v, for all edge l, if P(l,v) then remove l". In this respect, this sentence presents similarities with the donkey sentences, but this does not means that students deal in the same way with such sentences as they do in mathematics class. In the frame of a research project on quantification issues initiated in 2017 in Montpellier, we intend to explore the following didactical questions: are there differences in the treatment and interpretation of “for all” quantifier in mathematics and the “for every” in computer science? How much different are the treatment and interpretation of “if... then” conditional statement in mathematics and the “if... then... else...” conditional instruction in computer sciences? Do students deal the same way with statements and instructions regarding formal languages and natural language? How to interpret and treat specific instructions from algorithmic (as the while loop) in a didactical perspective? We are interested in discussing these questions during the workshop.

References

Corblin, F. (2002). *Représentation du discours et sémantique formelle*. Paris, Presses Universitaires de France.

Durand-Guerrier,V. (2003). Which notion of implication is the right one ? From logical considerations to a didactic perspective. *Educational Studies in Mathematics*. 53, 5-34

Kamp , H. (1981) A theory of truth and semantic representation, Groenendjik & al. (eds) *Formal methods in the study of language, Amsterdam, Mathematical Centre Tracks* 135, p.277-322.

Sandu, G. (1997) On the Theory of Anaphora: Dynamic Predicate Logic vs. Game-Theoretical Semantics, *Linguistics and Philosophy*, Vol. 20, No. 2 (Apr., 1997), pp. 147-174

Selden, J. and Selden, A. : 1995, 'Unpacking the logic of mathematical statements', in Educational Studies in Mathematics, 29, 123-151

-