REASONING IN AN ABSURD WORLD: DIFFICULTIES WITH PROOF BY CONTRADICTIION

Samuele Antonini  
Department of Mathematics  
University of Pavia, Italy

Maria Alessandra Mariotti  
Department of Mathematics  
University of Siena, Italy

The study presented in this report is part of a wide research project* concerning proofs by contradiction. Starting from the notion of mathematical theorem as the unity of statement, proof and theory, a structural analysis of proofs by contradiction has been carried out, producing a model to be used in the observation, analysis, and interpretation of cognitive and didactical issues related to this particular type of proof. In particular, the model highlights the complex relationship between the original statement to be proved and a new statement (the secondary statement) that is actually proved. Through the analysis of an exemplar protocol, this paper discusses on cognitive difficulties concerning the relationship between the reference theory and the proof of the secondary statement.

INTRODUCTION

The study presented in this report is part of a research project concerning difficulties that students encounter when are faced with proofs by contradiction, both at the high school and the university level. Although it was often observed that students spontaneously produce indirect arguments1, or, at least with a structure similar to that of proofs by contradiction (Freudenthal, 1973; Thompson, 1996; Reid & Dobbin, 1998; Antonini, 2003a, Antonini, 2003b), current literature agrees on the fact that students show much more difficulties with indirect than direct proofs. Different aspects have been highlighted, from different points of view. First of all, some authors remarked that this issue does not find an adequate attention, neither at the high school (Thompson, 1996) nor at the university level (Bernardi, 2002). Some difficulties were identified in the negation of a statement, commonly required in proofs by contradiction: because of its peculiarities, this process of negation presents a specific complexity in the mathematical domain (Thompson, 1996; Antonini, 2003a; Wu Yu et al., 2003). In a historic epistemological study, Barbin (1988) raised the issue of acceptability, pointing out that students’ attitudes towards this scheme of proof seem to echo ancient debates in the history of mathematics. According to Leron (1985) a specific difficulty seems to be related to the need of starting arguments with false assumptions: through these assumptions one enters a false, impossible world, and thinking in such an impossible world asks a highly demanding cognitive strain, which may explain the difficulties observed. Moreover, at the end of the proof, as soon as a contradiction is deduced, this world has to be rejected, so that students feel

* This research study was supported by the Italian Ministry of Education and Research (MIUR)- Prin 2003 #2003011072_003

1 Polya (1945) describes the role that proof by contradiction can assume in the production of conjectures.
deceived, and dissatisfied: they are faced with the unexpected destruction of the mathematical objects on which the proof was based (Leron, 1985). The research project (Antonini, 2003a), which this contribution is part of, is consistent with this direction of study and focuses on a cognitive and didactic analysis aimed to describe and interpret students’ difficulties in proof by contradiction within the broader context of proving activity in mathematics. In the following, the results of such analysis will be briefly outlined and subsequently employed to explain specific difficulties related to developing arguments from false assumptions.

**METHODOLOGY**

Consistently with its aim, the project developed two main research lines: empirical and theoretical. Collection of data was carried out through various means, interviews, questionnaires, recording and transcripts of classroom activities, and involved students at the high school (12th and 13th grade) and at the University level (Scientific Faculties such as Mathematics, Physic, Biology, …). In both cases, it was reasonable to assume that students were acquainted, and even familiar, with mathematical proof, as well as with proof by contradiction. For more details on the experimental design see (Antonini, 2003a).

**THEORETICAL FRAME**

The analysis of proof by contradiction started from the general notion of mathematical theorem introduced in (Mariotti et al., 1997; Mariotti, 2000). According to the ‘didactic’ definition formulated by the authors, a mathematical theorem is characterized by the system of relations between a statement, its proof, and a theory within which the proof make sense. In particular, the definition refers to “the existence of a reference theory as a system of shared principles and deduction rules […]” (Mariotti et al., 1997, p. 182). As far as deduction rules are concerned, a clear difference emerges between direct proofs and proofs by contradiction. In fact, direct proofs, based on deduction rules, lead back to well established argumentation schema; on the contrary, an intrinsic structural complexity of proofs by contradiction emerges showing specific aspects that can explain specific difficulties. The following analysis aims at describing such a complexity.

**STRUCTURAL ANALYSIS OF PROOF BY CONTRADICTION**

We consider the proof by contradiction of a given statement, that we call principal statement. Such a proof consists in the direct proof of another statement, that we call secondary statement. The move from one statement to the other is commonly introduced by expressions like “prove by contradiction” or “assume by contradiction” that signal to the reader the change in the type of argument that is going to be developed. For instance, we consider a proof by contradiction of the following statement (principal statement):

*let a and b two real numbers. If \( ab=0 \) then \( a=0 \) or \( b=0 \).*
Proof: assume by contradiction that \( ab = 0 \) and that \( a \neq 0 \) and \( b \neq 0 \). Since \( a \neq 0 \) and \( b \neq 0 \) one can divide both sides of the equality \( ab = 0 \) by \( a \) and by \( b \), obtaining \( 1 = 0 \).

Actually, this proof is a direct proof of the following statement (secondary statement):

let \( a \) and \( b \) two real numbers; if \( a \neq 0 \) and \( b \neq 0 \) and \( ab = 0 \) then \( 1 = 0 \)”.

The hypothesis of this statement, “\( a \neq 0 \) and \( b \neq 0 \) and \( ab = 0 \)”, is the negation of the principal statement (i.e. the conjunction of the hypothesis and the negation of the thesis of the statement to be proved) and its thesis is a false proposition, i.e. \( 1 = 0 \).

Thus, in order to prove the principal statement, that we indicate with \( E \), one provides the direct proof of the secondary statement, that we indicate with \( E^* \).

<table>
<thead>
<tr>
<th>Principal statement ( E )</th>
<th>Secondary statement ( E^* )</th>
</tr>
</thead>
<tbody>
<tr>
<td>a, b real numbers</td>
<td>a, b real numbers</td>
</tr>
<tr>
<td>If ( ab = 0 ) then ( a = 0 ) or ( b = 0 )</td>
<td>If ( a \neq 0 ) and ( b \neq 0 ) and ( ab = 0 ) then ( 1 = 0 )</td>
</tr>
</tbody>
</table>

**Table 1.** Principal and secondary statement in a proof by contradiction

From the point of view of logic, a proof by contradiction of the principal statement can be considered accomplished if the meta-statement \( E^* \rightarrow E \) is valid; in fact, in this case from \( E^* \) and \( E^* \rightarrow E \) it is possible to derive the validity of \( E \) by the well known “modus ponens” inference rule. But, the validity of the implication \( E^* \rightarrow E \) depends on the logic theory, i.e. the meta-theory, within which the assumed inference rules are stated. As it is commonly the case, i.e. in the classic logic theory, such a meta-theorem is valid, but it does not happen in other logic theories, such as the minimal or the intuitionistic logic.\(^2\)

This analysis, although very brief, clearly shows the complexity of the argumentative structure of a proof by contradiction, and in particular highlights key elements of this complexity, such as the secondary statement \( E^* \), its proof in respect to the reference mathematical theory, the meta-statement \( E^* \rightarrow E \), and its validity in respect to the assumed meta-theory.

We assume that specific difficulties may be related to each of these key elements and to their relationships.

In this paper we are going to address the problems related to the proof of the secondary statement \( E^* \); further research have been carried out and are still in progress, concerning the cognitive and didactic problems related to the validity of the meta-statement \( E^* \rightarrow E \), and more generally to the acceptability of the proof by contradiction on the whole (Antonini, 2003a, 2004).

\(^2\) For a definition in terms of rules of inference of the classic, minimal and intuitionistic logic, see Prawitz (1971).
FALSE HYPOTHESIS AND THE MATHEMATICAL THEORY OF REFERENCE

Consider statement $E^*$, and the mathematical theorem which has $E^*$ as its statement; according to the definition given above, such a theorem is defined by the triplet $(E^*, C, T)$, where $C$ is a direct proof of $E^*$ and $T$ is the mathematical theory within which this proof is constructed and validated. One of the main characteristic of this theorem concerns the fact that both the hypothesis and the thesis of the statement $E^*$ are constituted by false propositions; although from the logic point of view this fact does not present any particular problem, from the cognitive point of view this peculiarity may have serious consequences. In fact, conflicts may arise between the theoretical and the cognitive point of view. From the logical point of view, we observe:

- in spite of the falsity of both the hypothesis and the thesis, the statement $E^*$ is logically well formulated. Moreover, in spite of the falsity of its hypothesis (and just because of it, according to the truth tables) the implication $E^*$ results to be true;
- the proof $C$ constitutes a valid proof of the implication $E^*$. That means something more than the fact that $E^*$ is logically true, it means that it is possible to construct a deductive chain within a mathematical theory, and this despite the fact that both the hypothesis and the thesis are false;
- deduction in a mathematical theory is independent of the interpretation of the statements involved, that means that axioms and theorems of a mathematical theory can be applied to objects which are mathematically impossible, and for this reason, absurd: for instance, two real numbers $a$ and $b$ different from 0 and such that $ab=0$, the rational square root of 2, parallel lines that intersect each other, … .

For example, let us consider the previous theorem and analyse its proof according to our discussion.

(Principal) statement: let $a$ and $b$ two real numbers. If $ab=0$ then either $a=0$ or $b=0$.

Proof: assume by contradiction that $ab=0$ and $a≠0$ and $b≠0$. Since $a≠0$ and $b≠0$ both sides of the equality $ab=0$ can be divided, respectively, by $a$ and by $b$, obtaining $1=0$.

This is a direct proof of the secondary statement “let $a,b$ real numbers; if $a≠0$ and $b≠0$ and $ab=0$ then $1=0$”. The hypothesis of this statement is “$a≠0$ and $b≠0$ and $ab=0$”: it is false because do not exist two real numbers $a$ and $b$ such that $a≠0$ and $b≠0$ and $ab=0$; the thesis is “$1=0$”: it is false because $1≠0$. The implication expressed by the statement is true, because the falsity of the hypothesis.

The proof is within the real numbers theory (or more generally the mathematical fields theory) and is based on the following two axioms:

1. If a number is not zero, it has a multiplicative inverse;
2. If both sides of an equality are multiplied by the same number, the equality relation is maintained.

In this proof, these axioms are used to make some deduction on impossible mathematical objects. Axiom 1 is applied to the non-existing real numbers a and b such that a≠0 and b≠0 and ab=0; axiom 2 is applied to “ab=0”, an equality formulated with the two non-existing numbers.

In summary, while the truth of the secondary statement E* depends on the falsity of its antecedent, the validity of its proof C is based on the validity of a deductive chain within the mathematical theory T, that is applied to impossible mathematical objects.

From the didactical point of view, some authors, for instance Durand-Guerrier (2003), already pointed out students’ difficulties in evaluating or accepting the truth-value of an implication with a false antecedent; the following discussion will focus on specific difficulties originated by a proof validating an implication where both the antecedent and the consequent are false.

**ANALYSIS OF A CASE**

The protocol we are going to analyse clearly shows what kind of difficulties may emerge when a student is producing a proof by contradiction, and as a consequence of a ‘false’ assumption, he/she is required to manage the mathematical theory of reference in respect to impossible objects.

The subject of the interview, Maria, is an university student (the last year of the Faculty of Pharmacy) and as it is possible to notice, she is familiar with proof. In the following an excerpt of the interview is reported; in the transcript “I” indicates the interviewer, “M” indicates the student, and bold character indicates that the subject in some way emphasised her words.

**Excerpt 1 (from Maria’s interview)**

1. I: Could you try to prove by contradiction the following: “if ab=0 then a=0 or b=0”?
2. M: [...] well, assume that ab=0 with a different from 0 and b different from 0... I can divide by b... ab/b=0/b... that is a=0. I do not know whether this is a proof, because there might be many things that I haven’t seen.
3. M: moreover, so as ab=0 with a different from 0 and b different form 0, that is against my common beliefs (ita. contro le mie normali vedute) and I must pretend to be true, I do not know if I can consider that 0/b=0. I mean, I do not know what is true and what I pretend it is true.
4. I: let us say that one can use that 0/b=0.
5. M: it comes that a=0 and consequently ... we are back to reality. Then it is proved because ... also in the absurd world it may come a true thing; thus I cannot stay in the absurd world. The absurd world has its own rules, which are absurd, and if one does not respect them, comes back.
6. I: who does come back?
7. M: It is as if a, b and ab move from the real world to the absurd world, but the rules do not function on them, consequently they have to come back ...
8. M: But my problem is to understand which are the rules in the absurd world, are they the rules of the absurd world or those of the real world? This is the reason why I have problems to know if 0/b=0, I do not know whether it is true in the absurd world. […]

9. I: (The interviewer shows the proof by contradiction of “√2 is irrational”) what do you think about it?

10. M: in this case, I have no doubts, but why is it so? … perhaps, when I have accepted that the square root of 2 is a fraction I continued to stay in my world. I made the calculations as I usually do, I did not put myself problems like “in this world, a prime number is no more a prime number” or “a number is no more represented by the product of prime numbers”. The difference between this case and the case of the zero-product is in the fact that this is obvious whilst I can believe that the square root of 2 is a fraction, I can believe that it is true and I can go on as if it is true. In the case of the zero-product I cannot pretend that it is true, I cannot tell myself such a lie and believe it too!

Maria’s arguments are firmly based on what she believes it is “true”; she seems to refer to numbers relations in a numerical world which she knows and is familiar with (integers, rationals, or perhaps real numbers). On the contrary, when she has to construct arguments in what she calls an absurd world, the world where there exist two numbers, different from zero and such as their product is zero, Maria looses the control, because in this world she does not know any more what is true and what is false (8). In Maria’s opinion, when one assumes something false, everything can happen, it might even occur that 0/b≠0 (3).

Maria is not able to control the relationship between her argumentation and the mathematical theory within which such argumentation should make sense. The absurdity of the assumption, on which the deduction has to be based, upset the truth values of statements that for the subject are fundamental (3), so that she suspects that it might be possible the existence of a different theory, suitable for such an absurd situation. She opposes the “real world” and the “absurd world”, each referring to its own different rules (5;8).

Thus Maria fails to grasp a fundamental point, which constitutes a key element of the proof scheme by contradiction: the fact that the secondary statement is valid in the mathematical theory of reference. She focuses her attention on truth rather than on validity, and she looks for truth in a world where there are two non zero numbers, whose product is equal to zero. Our interpretation of Maria’s difficulties finds further support in her final remark; she states that she has no problems to accept the proof of the irrationality of √2, because in that case she finds easy to believe that √2 is truly a fraction; although absurd, the world where √2 is a fraction, is acceptable for Maria and in such a world, what are for Maria the fundamental truths are not upset (10).

CONCLUSIONS

The example discussed above clearly shows the crucial role played by the mathematical theory of reference in proving statements starting from false hypotheses. The example and its discussion confirms through empirical evidence
what argued by other researchers, for instance Leron (1985), about the cognitive strain needed by reasoning in a *false world*, i.e. reasoning on the base of false assumptions.

At the source of the difficulties with proof by contradiction there seems to be the fact that in the *absurd world* some of the fundamental properties are upset, so that they no more can be true. It seems that false hypotheses may produce a shortcut in the system of beliefs of a subject and induce impasse or doubts on the proving process: the subject loses the control on the deductive steps of the proof, because he/she does not know what is or is not true.

This attitude is consistent with the well spread opinion that a theorem expresses the link between properties that are true, or at least assumed true. Using the words of Berenice, a 13\textsuperscript{th} grade student of a scientific high school, during the interview:

**Excerpt 2 (from Berenice’s interview)**

1. I: could you tell me what is a theorem for you?
2. […]
3. B: generally speaking, a theorem is … is something … and … that … that is proved on the base of things that I know for sure that are true and from them I can start to prove for sure other things, I mean … more or less
4. […] I mean, I do … I must come to say that one certain thing is true, I assume that one certain thing is true, or I assume one certain thing and … through some steps that I know I can do because I know that they are … are correct, I come to say that that thing is such as it was assumed.

According to what expressed by Durand-Guerrier (2003), implication with false antecedent are not accepted or however are considered as false by students: in order to accept a proof it seems necessary to start from true, or at least potentially true assumptions. In the *real world*, as Maria says, or at least in a world that can be assumed real (as the world where $\sqrt{2}$ can be expressed by a fraction), the theory and the rules to be applied are not put into question. On the contrary, assuming false hypotheses can block the deductive process because it may ask to apply the mathematical theory to absurd situations.

The structural analysis of a proof by contradiction, presented above, highlighted the complex system of relationships between key elements, such as the principal, the secondary statement, and its proof, that were subsequently employed in the discussion of an exemplar protocol; according to our general assumption, specific difficulties were interpreted focusing on the proof of the secondary statement. Further investigation have been carried out and are still in progress in this same direction; for instance, besides the problems related to the proof of the secondary statement, we drew our attention on the passage from the principal to the secondary statement and vice-versa and the problems that this passage can present, besides the difficulty coming from effect of mathematical negation on the correct formulation of the secondary statement.
We claim that this method of research can be generalized: in fact, the structural analysis provides an effective model for generating specific research hypotheses concerning students’ difficulties related to different elements involved in a proof by contradiction.

REFERENCES

Antonini, S.(2003a), Dimostrare per assurdo: analisi cognitiva in una prospettiva didattica, Tesi di Dottorato, Dipartimento di Matematica, Università di Pisa.


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


Polya, G.(1945), How to solve it, Princeton University Press.


Thompson, D.R.(1996), Learning and Teaching Indirect Proof, The Mathematics Teacher v. 89(6), 474-82.