Do students judge mathematical proofs to be valid reasoning?

Patricia Baggett, New Mexico State University, baggett@nmsu.edu, and Andrzej Ehrenfeucht, University of Colorado, andrzej@cs.colorado.edu

Psychonomic Society 2003

Abstract

Mathematical proofs use two-valued logic. This means that in any mathematical structure, every syntactically correct sentence is either true or false. The meaning of logical connectives (and, or, not, ...) is well defined and doesn't depend on the context in which a connective is used. Students with different mathematical backgrounds were asked to evaluate mathematical sentences. Most students:

1. also used other logical values (e. g. "neither true nor false") when given the opportunity;

2. assigned the values true and false differently than is done in mathematics; and

3. interpreted logical connectives differently depending on the context.

It is possible that students' difficulty with proofs is due to the fact that they do not see some mathematical proofs as being valid reasoning. The finding also calls into question NCTM's (2000, p. 56) statement, "A mathematical proof is a formal way of expressing particular kinds of reasoning and justification."

Introduction

1.

The system of two-valued logic that is used in mathematics was fully described at the beginning of the 20th century. The first complete description was given by Bertrand Russell and Alfred North Whitehead in Principia Mathematica (1910). The system is rule-based, and is often called "formal". Within set theory one can prove that the system is complete (completeness theorem) in the sense that if a statement cannot be proven from the assumptions, then there exists a mathematical structure (a counter-example) in which the assumptions are true and the statement is false.

No one claims that mathematicians use these rules when they reason about mathematics. We are rather sure that they don't. The only claim is that the system of rules of two-valued logic could have been, in principle, a tool for proving theorems.

This system was not universally accepted. The main challenge came from Brouwer, who designed another system called Intuitionism (see Heyting, 1934), and since then many other systems of multi-valued logic have been designed.

2.

From the beginning it has been clear that the use of logical connectives in two-valued logic (especially "if ... then") is different than in everyday life. Also we know that its use outside mathematics is severely restricted, due to the fact that all conclusions derived from inconsistent premises are essentially unreliable, and that showing the consistency of premises is mostly impossible.

3.

Students at all levels have difficulty with understanding and constructing proofs. Some generally accessible proofs, such as proofs of the Pythagorean theorem, are based on geometric constructions and are rather untypical. More typical proofs are simply sequences of statements written in mathematical jargon or as formulas, which are either quotes of axioms and already proven theorems or corollaries of statements that have already been proved. Such "verbal" proofs are mostly accessible only to specialists.

4.

We wanted to see if the difference between the use of logical operators such as "if .. then" or "for every" by students and in two-valued logic, which is consistent with their use in mathematical proofs, is really very large.

Subjects

We used four groups of subjects. The numbers in each group are given.

48 computer science majors (juniors) taking a course in the theory of algorithms.

36 education undergraduates and practicing teachers taking a low-level mathematics course.

19 education undergraduates and graduates, and practicing teachers taking a science course.

22 graduate students in computer science taking a course in the theory of computation

Task

Subjects were given a questionnaire with the following instructions:

This is not a test in mathematics or logic. Answer the questions using your knowledge of math and your common sense. The task is to judge the validity of the sentences presented below. Circle just one answer in each row.

The subjects were asked to choose one of five answers: true, false, both true and false, neither true nor false, I don't know.

There were eight questions. Each question had the form "If ... then ...", where both parts could be easily evaluated as true or false. The questions were:

(1) If 1 = 2 then 2 = 3.

(2) If 1 = 2 then 2 = 5.

(3) If 1 = 3 then 5 = 5.

(4) If 5 = 5 then 1 = 5.

(5) If 4 = 4 then 1 = 1.

(6) If every odd number is prime, then there is an even prime number.

(7) If every odd number is prime, then there is an odd prime number.

(8) If every even number is prime, then one is a prime number.

In two-valued logic all sentences except (4) are true. The patterns are:

	predecessor:	successor:	implication:
(1) If 1 = 2 then 2 = 3.	false	false	true
(2) If 1 = 2 then 2 = 5.	false	false	true
(3) If 1 = 3 then 5 = 5.	false	true	true
(4) If 5 = 5 then 1 = 5.	true	false	false
(5) If 4 = 4 then 1 = 1.	true	true	true
(6) If every odd number is prime, then there is an even prime number.	false	true	true
(7) If every odd number is prime, then there is an odd prime number.	false	true	true
(8) If every even number is prime, then one is a prime number.	false	false	true

Only one of the sentences is a logical tautology that doesn't depend on any knowledge of arithmetic. Sentence 7, "If every odd number is prime, then there is an odd prime number," follows the pattern: If for every x, P(x), then there is an x such that P(x). And this is valid for any property P.

References

Heyting, Arend (1934). *Intuitionism: An Introduction*. Amsterdam: North Holland.

NCTM (2000). *Principles and Standards for School Mathematics*. Reston, VA: NCTM.

Selden, A. & Selden, J. (2003). Validations of proofs considered as texts: Can undergraduates tell whether an argument proves a theorem? *Journal for Research in Mathematics Education* 34 (1), 4-36.

Whitehead, A. & Russell, B. (1910-13). *Principia Mathematica*. Cambridge: Cambridge University Press.

Results

The 125 subjects gave these answers:

sentence:	true:	false:	both:	neither:	I don't know:
(1) If 1=2 then 2=3.	36*	39	15	22	13
(2) If 1=2 then 2=5.	21*	59	7	13	15
(8) If every even number is prime, then one is a prime number.	30*	50	8	19	18
(3) If 1=3 then 5=5.	26*	50	17	19	13
(6) If every odd number is prime, then there is an even prime number.	30*	50	6	22	17
(7) If every odd number is prime, then there is an odd prime number.	96*	13	6	2	8
(5) If 4=4 then 1=1.	100*	4	5	10	6
(4) If 5=5 then 1=5.	4	72*	14	17	18

(Answers based on two-valued logic are marked with an asterisk.)

Conclusions

Subjects do not consider implications with false predecessors as being true (sentences 1, 2, 8, 3, and 6). Either they judge them as false, or they cannot assign a value of true or false to them. (For example, they may consider them meaningless. Some subjects actually said this after the quiz.)

Sentence (7) is an interesting exception. It is considered true by most subjects (77%). It is possible that the reason for this is that it is a logical tautology, but we would need more data to reach such a conclusion.

Most subjects concluded that (5) was true, as we expected.

It was rather surprising that only 72 subjects (58%) judged (4) as false. But this finding is in agreement with a finding of Selden & Selden (2003) that students have difficulty finding errors in incorrect reasoning.

More results

Only 8 subjects (6%) judged all sentences according to two-valued logic. Five of these were graduate students in computer science who were quite familiar with Boolean algebra.

Overall we did not see any consistency or patterns in subjects' answers. For example, if we compare the values assigned to the two first questions, (1) If 1 = 2 then 2 = 3; and (2) If 1 = 2 then 2 = 5, we have (T = true, F = false, O = other):

Answers given to questions 1 (If 1=2, then 2=3) and 2 (If 1=2, then 2=5)	1 2 T T	1 2 T F	1 2 F T	1 2 F F	1 2 T O	1 2 O T	1 2 F O	1 2 O F	1 2 OO
Number of subjects	20	15	0	37	1	1	2	7	42

Among all 125 subjects' answers, 83 (66%) were unique. Among the others, no pattern of answers, except the 8 answers in agreement with two-valued logic, was chosen by more than 4 subjects.

This lack of pattern strongly indicates that if we would ask subjects the same questions again, their answers would be not consistent with their previous answers.

Overall conclusion

It is well known that common concepts such as work or energy have specialized meanings in physics that are different from their common usage. Similarly the concepts of function or group have a technical meaning in mathematics. It seems that the same is true about logical connectives that are used in everyday situations and in mathematical proofs.

Acknowledgment.

We thank NASA (the National Aeronautics and Space Administration) for Sponsoring this work under the MASEF (Math and Science Educators For the Future program