Mathematical Proofs as Derivation-Indicators: Theory and Implementation

Peter Koepke, University of Bonn, Germany

Mathematical Institute

International Center for Philosophy NRW

Seminar on the Philosophy of Mathematics

Utrecht, November 3, 2009

Contents

- The Gödel completeness theorem
- Proofs and formal derivations
- Formal derivations from Euclid to MIZAR
- What is a mathematical proof?
- The derivation-indicator view of mathematical practice
- The Naproche project
- Implementation of derivation-indication
- Results

The Gödel completeness theorem

Über die Vollständigkeit des Logikkalküls
(1929)

1. Einleitung

Der Hauptgegenstand der folgenden Untersuchungen ist der Beweis der Vollständigkeit des in Russell, *Principia mathematica*, P. I, Nr. 1 und Nr. 10, und ähnlich in Hilbert–Ackermann, *Grundzüge der theoretischen Logik* (zitiert als H. A.), III, § 5, angegebenen Axiomensystems des sogenannten engeren Funktionenkalküls. Dabei soll "Vollständigkeit" bedeuten, daß jede im engeren Funktionenkalkül ausdrückbare allgemein giltige Formel (allgemein giltige Zählaussage nach Löwenheim) sich durch eine endliche Reihe formaler Schlüsse aus den Axiomen deduzieren lässt. Diese Behauptung lässt sich leicht als äquivalent erkennen mit der folgenden: Jedes widerspruchslose nur aus Zählaussagen bestehende Axiomensystem[1] hat eine Realisierung. (Widerspruchslos heißt dabei, daß durch endlich viele formale Schlüsse kein Widerspruch hergeleitet werden kann.)

(Doctoral Dissertation, Vienna 1929)

The Gödel completeness theorem

Every logically true mathematical statement has a formal derivation.

The Gödel completeness theorem

Every logically true mathematical statement has a formal derivation.

Every true mathematical statement has a formal derivation within some (foundational) axiom system.

The Gödel completeness theorem

Every logically true mathematical statement has a formal derivation.

Every true mathematical statement has a formal derivation within some (foundational) axiom system.

Every mathematical proof can be replaced by a formal derivation.

The Gödel completeness theorem

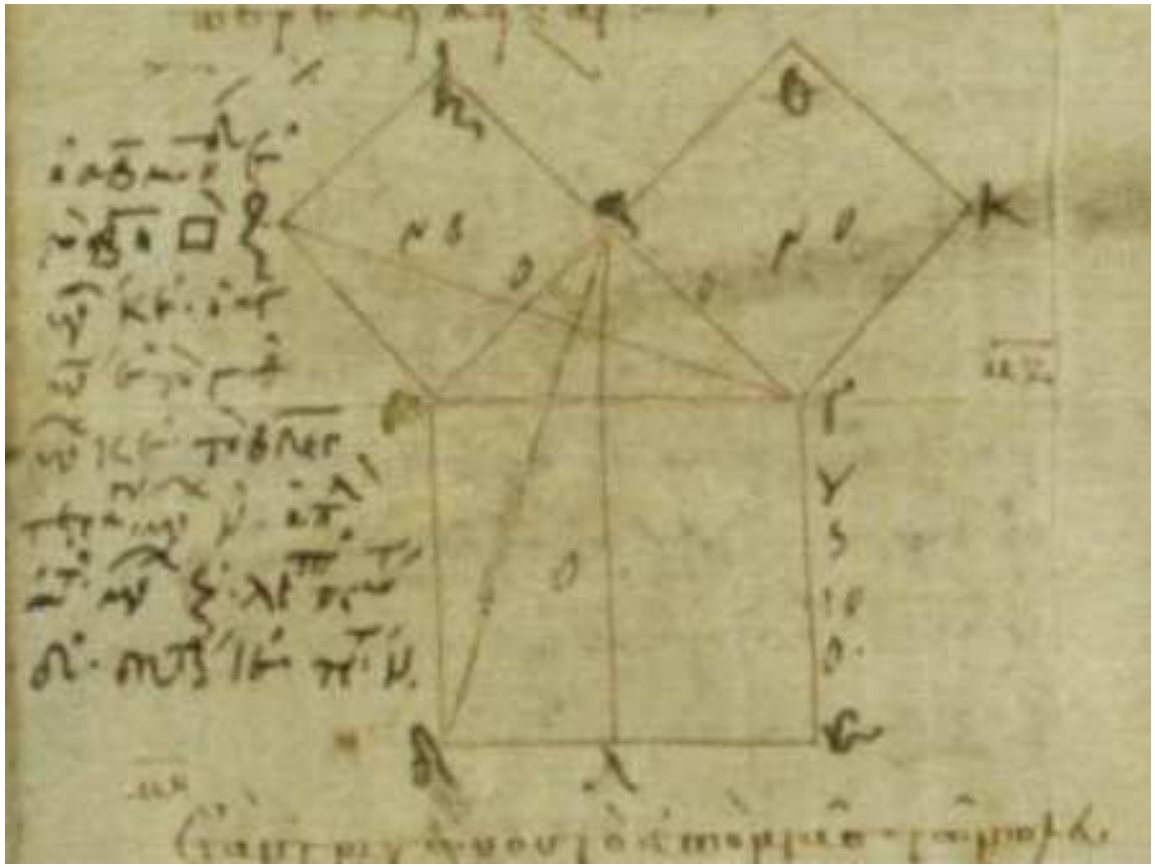
Every logically true mathematical statement has a formal derivation.

Every true mathematical statement has a formal derivation within some (foundational) axiom system.

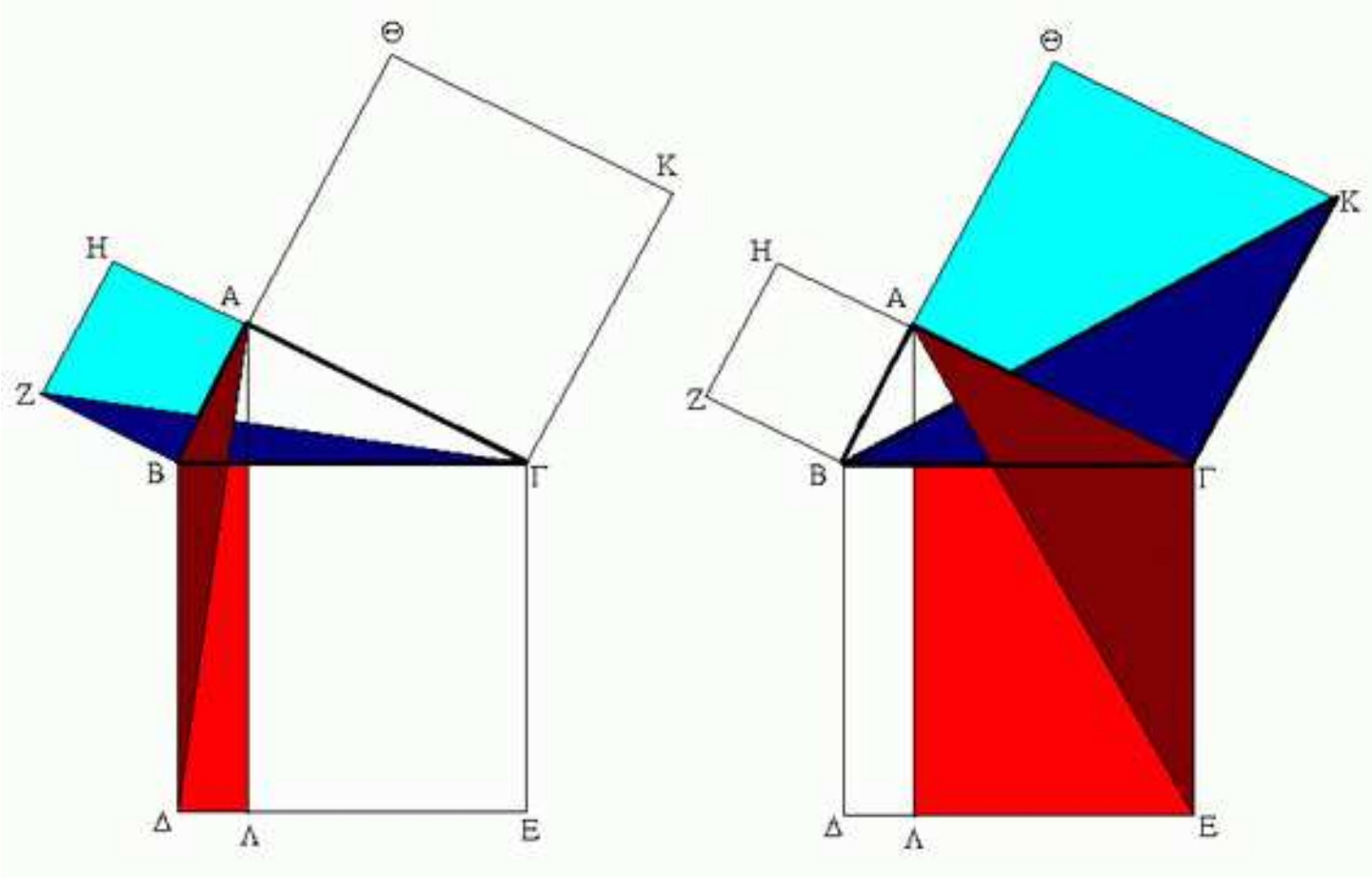
Every mathematical proof can be replaced by a formal derivation.

Mathematics can be in principle be carried out completely formal (Formal mathematics).

Pythagoras theorem



Pythagoras theorem

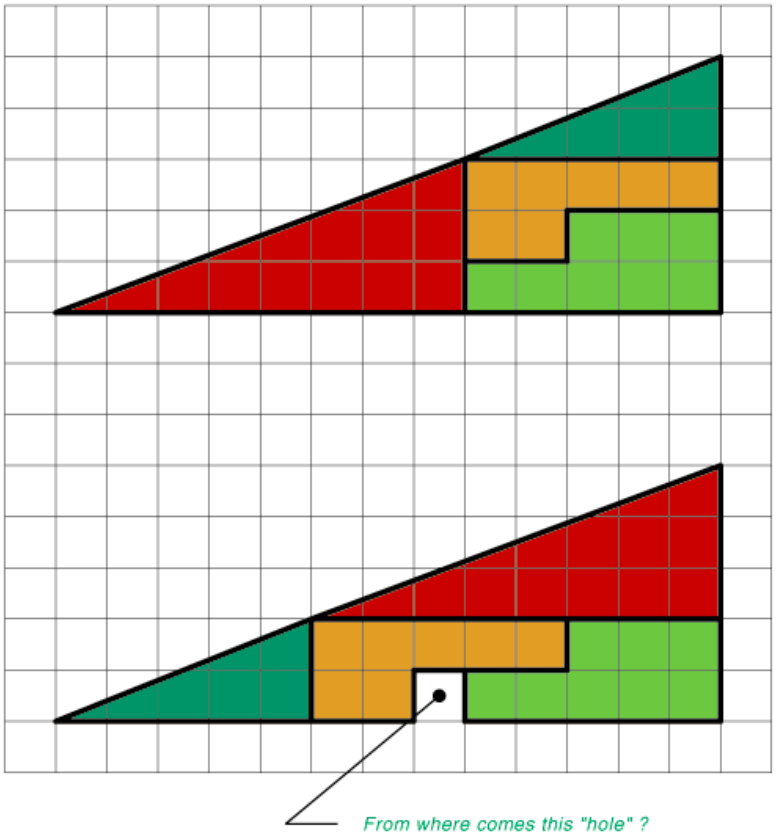


Graphical proofs?

Graphical proofs?

63=65?

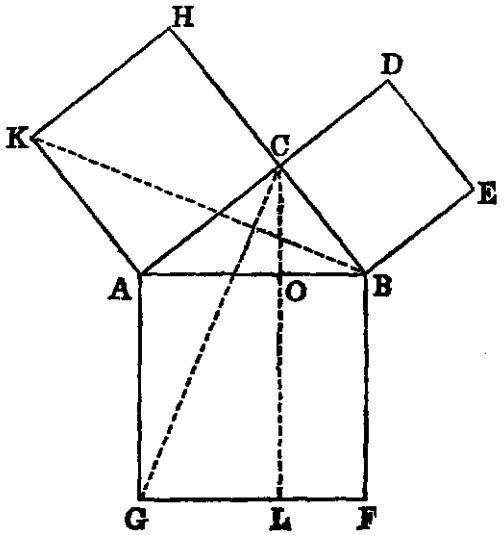
HOW CAN THIS BE TRUE ?



Below the four parts are moved around

The partitions are exactly the same, as those used above

Theorem. In a right-angled triangle (ABC) the square on the hypotenuse (AB) is equal to the sum of the squares on the other two sides (AC, BC).



Dem.---On the sides $A\,B$, $B\,C$, $C\,A$ describe squares [xlvi.]. Draw $C\,L$ parallel to $A\,G$. Join $C\,G$, $B\,K$. Then because the angle $A\,C\,B$ is right (hyp.), and $A\,C\,H$ is right, being the angle of a square, the sum of the angles $A\,C\,B$, $A\,C\,H$ is two right angles; therefore $B\,C$, $C\,H$ are in the same right line [xiv.]. In like manner $A\,C$, $C\,D$ are in the same right line. Again, because BAG is the angle of a square it is a right angle: in like manner $C\,A\,K$ is a right angle. Hence $B\,A\,G$ is equal to $C\,A\,K$: to each add $B\,A\,C$, and we get the angle $C\,A\,G$ equal to $K\,A\,B$. Again, since $B\,G$ and $C\,K$ are squares, $B\,A$ is equal to $A\,G$, and $C\,A$ to $A\,K$. Hence the two triangles $C\,A\,G$, $K\,A\,B$ have the sides $C\,A$, $A\,G$ in one respectively equal to the sides $K\,A$, $A\,B$ in the other, and the contained angles $C\,A\,G$, $K\,A\,B$ also equal. Therefore [iv.] the triangles are equal; but the parallelogram $A\,L$ is double of the triangle $C\,A\,G$ [xli.], because they are on the same base $A\,G$, and between the same parallels $A\,G$ and $C\,L$. In like manner the parallelogram $A\,H$ is double of the triangle $K\,A\,B$, because they are on the same base $A\,K$, and between the same parallels $A\,K$ and $B\,H$; and since doubles of equal things are equal (Axiom vi.), the parallelogram $A\,L$ is equal to $A\,H$. In like manner it can be proved that the parallelogram $B\,L$ is equal to $B\,D$. *Hence the whole square $A\,F$ is equal to the sum of the two squares $A\,H$ and $B\,D$.*

Dem.---On the sides $A\,B$, $B\,C$, $C\,A$ describe squares [xlvi.]. Draw $C\,L$ parallel to $A\,G$. Join $C\,G$, $B\,K$. Then because the angle $A\,C\,B$ is right (hyp.), and $A\,C\,H$ is right, being the angle of a square, the sum of the angles $A\,C\,B$, $A\,C\,H$ is two right angles; therefore $B\,C$, $C\,H$ are in the same right line [xiv.]. In like manner $A\,C$, $C\,D$ are in the same right line. Again, because BAG is the angle of a square it is a right angle: in like manner CAK is a right angle. Hence $B\,A\,G$ is equal to $C\,A\,K$: to each add $B\,A\,C$, and we get the angle $C\,A\,G$ equal to $K\,A\,B$. Again, since $B\,G$ and $C\,K$ are squares, $B\,A$ is equal to $A\,G$, and $C\,A$ to $A\,K$. Hence the two triangles $C\,A\,G$, $K\,A\,B$ have the sides $C\,A$, $A\,G$ in one respectively equal to the sides $K\,A$, $A\,B$ in the other, and the contained angles $C\,A\,G$, $K\,A\,B$ also equal. Therefore [iv.] the triangles are equal; but the parallelogram $A\,L$ is double of the triangle $C\,A\,G$ [xli.], because they are on the same base $A\,G$, and between the same parallels $A\,G$ and $C\,L$. In like manner the parallelogram $A\,H$ is double of the triangle $K\,A\,B$, because they are on the same base $A\,K$, and between the same parallels $A\,K$ and $B\,H$; and since doubles of equal things are equal (Axiom vi.), the parallelogram $A\,L$ is equal to $A\,H$. In like manner it can be proved that the parallelogram $B\,L$ is equal to $B\,D$. *Hence the whole square $A\,F$ is equal to the sum of the two squares $A\,H$ and $B\,D$.*

Dem.---On the sides $A\,B$, $B\,C$, $C\,A$ describe squares [xlvi.]. Draw $C\,L$ parallel to $A\,G$. Join $C\,G$, $B\,K$. Then because the angle $A\,C\,B$ is right (hyp.), and $A\,C\,H$ is right, being the angle of a square, the sum of the angles $A\,C\,B$, $A\,C\,H$ is two right angles; therefore $B\,C$, $C\,H$ are in the same right line [xiv.]. In like manner $A\,C$, $C\,D$ are in the same right line. Again, because BAG is the angle of a square it is a right angle: in like manner CAK is a right angle. Hence BAG is equal to CAK: to each add $B\,A\,C$, and we get the angle $C\,A\,G$ equal to $K\,A\,B$. Again, since $B\,G$ and $C\,K$ are squares, $B\,A$ is equal to $A\,G$, and $C\,A$ to $A\,K$. Hence the two triangles $C\,A\,G$, $K\,A\,B$ have the sides $C\,A$, $A\,G$ in one respectively equal to the sides $K\,A$, $A\,B$ in the other, and the contained angles $C\,A\,G$, $K\,A\,B$ also equal. Therefore [iv.] the triangles are equal; but the parallelogram $A\,L$ is double of the triangle $C\,A\,G$ [xli.], because they are on the same base $A\,G$, and between the same parallels $A\,G$ and $C\,L$. In like manner the parallelogram $A\,H$ is double of the triangle $K\,A\,B$, because they are on the same base $A\,K$, and between the same parallels $A\,K$ and $B\,H$; and since doubles of equal things are equal (Axiom vi.), the parallelogram $A\,L$ is equal to $A\,H$. In like manner it can be proved that the parallelogram $B\,L$ is equal to $B\,D$. *Hence the whole square $A\,F$ is equal to the sum of the two squares $A\,H$ and $B\,D$.*

Dem.---On the sides $A\,B$, $B\,C$, $C\,A$ describe squares [xlvi.]. Draw $C\,L$ parallel to $A\,G$. Join $C\,G$, $B\,K$. Then because the angle $A\,C\,B$ is right (hyp.), and $A\,C\,H$ is right, being the angle of a square, the sum of the angles $A\,C\,B$, $A\,C\,H$ is two right angles; therefore $B\,C$, $C\,H$ are in the same right line [xiv.]. In like manner $A\,C$, $C\,D$ are in the same right line. Again, because $B\,A\,G$ is the angle of a square it is a right angle: in like manner $C\,A\,K$ is a right angle. Hence $B\,A\,G$ is equal to $C\,A\,K$: to each add $B\,A\,C$, and we get the angle $C\,A\,G$ equal to $K\,A\,B$. Again, since $B\,G$ and $C\,K$ are squares, $B\,A$ is equal to $A\,G$, and $C\,A$ to $A\,K$. Hence the two triangles $C\,A\,G$, $K\,A\,B$ have the sides $C\,A$, $A\,G$ in one respectively equal to the sides $K\,A$, $A\,B$ in the other, and the contained angles $C\,A\,G$, $K\,A\,B$ also equal. Therefore [iv.] the triangles are equal; but the parallelogram $A\,L$ is double of the triangle $C\,A\,G$ [xli.], because they are on the same base $A\,G$, and between the same parallels $A\,G$ and $C\,L$. In like manner the parallelogram $A\,H$ is double of the triangle $K\,A\,B$, because they are on the same base $A\,K$, and between the same parallels $A\,K$ and $B\,H$; and since doubles of equal things are equal (Axiom vi.), the parallelogram $A\,L$ is equal to $A\,H$. In like manner it can be proved that the parallelogram $B\,L$ is equal to $B\,D$. *Hence the whole square $A\,F$ is equal to the sum of the two squares $A\,H$ and $B\,D$.*

1.	Φ_{Gr}	$\neg \circ v_0\, e \equiv v_0$		$\neg\exists v_0\, \neg \circ v_0\, e \equiv v_0$	VR
2.	Φ_{Gr}	$\neg \circ v_0\, e \equiv v_0$		$\neg \circ v_0\, e \equiv v_0$	VR
3.	Φ_{Gr}	$\neg \circ v_0\, e \equiv v_0$		$\exists v_0 \neg \circ v_0\, e \equiv v_0$	$\exists$S auf 2
4.	Φ_{Gr}			$\circ v_0\, e \equiv v_0$	WR auf 1,3
5.				$(v_2 \equiv \circ v_0\, e)\frac{\circ v_0\, e}{v_2}$	$(\equiv)$
6.		$\circ v_0\, e \equiv v_0$		$(v_2 \equiv \circ v_0\, e)\frac{v_0}{v_2}$	Sub auf 5
7.	Φ_{Gr}	$\circ v_0\, e \equiv v_0$		$v_0 \equiv \circ v_0\, e$	AR auf 6
8.	Φ_{Gr}			$v_0 \equiv \circ v_0\, e$	KS auf 4,7
9.	Φ_{Gr}	$v_0 \equiv e$		$v_0 \equiv e$	VR
10.	Φ_{Gr}	$v_0 \equiv e$		$(\neg \circ v_0\, e \equiv e \vee v_0 \equiv e)$	$\vee$S auf 9
11.	Φ_{Gr}	$\neg v_0 \equiv e$		$(\neg v_2 \equiv e)\frac{v_0}{v_2}$	VR
12.	Φ_{Gr}	$\neg v_0 \equiv e$	$v_0 \equiv \circ v_0\, e$	$(\neg v_2 \equiv e)\frac{\circ v_0\, e}{v_2}$	Sub auf 11
13.	Φ_{Gr}	$\neg v_0 \equiv e$	$v_0 \equiv \circ v_0\, e$	$\neg \circ v_0\, e \equiv e$	12
14.	Φ_{Gr}	$\neg v_0 \equiv e$		$v_0 \equiv \circ v_0\, e$	AR auf 8
15.	Φ_{Gr}	$\neg v_0 \equiv e$		$\neg \circ v_0\, e \equiv e$	KS auf 14
16.	Φ_{Gr}	$\neg v_0 \equiv e$		$(\neg \circ v_0\, e \equiv e \vee v_0 \equiv e)$	$\vee$S auf 15
17.	Φ_{Gr}			$(\neg \circ v_0\, e \equiv e \vee v_0 \equiv e)$	FU auf 10,16
18.	Φ_{Gr}	$\neg(\neg \circ v_0\, e \equiv e \vee v_0 \equiv e)$	$\neg\neg\exists v_0 \neg(\neg \circ v_0\, e \equiv e \vee v_0 \equiv e)$	$(\neg \circ v_0\, e \equiv e \vee v_0 \equiv e)$	AR auf 17
19.	Φ_{Gr}	$\neg(\neg \circ v_0\, e \equiv e \vee v_0 \equiv e)$	$\neg\neg\exists v_0 \neg(\neg \circ v_0\, e \equiv e \vee v_0 \equiv e)$	$\neg(\neg \circ v_0\, e \equiv e \vee v_0 \equiv e)$	VR
20.	Φ_{Gr}	$\neg(\neg \circ v_0\, e \equiv e \vee v_0 \equiv e)$		$\neg\exists v_0 \neg(\neg \circ v_0\, e \equiv e \vee v_0 \equiv e)$	WR auf 18,19
21.	Φ_{Gr}	$\exists v_0 \neg(\neg \circ v_0\, e \equiv e \vee v_0 \equiv e)$		$\neg\exists v_0 \neg(\neg \circ v_0\, e \equiv e \vee v_0 \equiv e)$	$\exists$A auf 20
22.	Φ_{Gr}	$\neg\exists v_0 \neg(\neg \circ v_0\, e \equiv e \vee v_0 \equiv e)$		$\neg\exists v_0 \neg(\neg \circ v_0\, e \equiv e \vee v_0 \equiv e)$	VR
23.	Φ_{Gr}			$\neg\exists v_0 \neg(\neg \circ v_0\, e \equiv e \vee v_0 \equiv e)$	FU auf 21,22

1.	Φ_{Gr}	$\neg \circ v_0 e \equiv v_0$	$\neg\exists v_0 \neg \circ v_0 e \equiv v_0$	VR
2.	Φ_{Gr}	$\neg \circ v_0 e \equiv v_0$	$\neg \circ v_0 e \equiv v_0$	VR
3.	Φ_{Gr}	$\neg \circ v_0 e \equiv v_0$	$\exists v_0 \neg \circ v_0 e \equiv v_0$	$\exists$S auf 2
4.	Φ_{Gr}		$\circ v_0 e \equiv v_0$	WR auf 1,3

1.	Φ_{Gr}	$\neg \circ v_0\, e \equiv v_0$	$\neg\exists v_0\, \neg \circ v_0\, e \equiv v_0$	VR
2.	Φ_{Gr}	$\neg \circ v_0\, e \equiv v_0$	$\neg \circ v_0\, e \equiv v_0$	VR
3.	Φ_{Gr}	$\neg \circ v_0\, e \equiv v_0$	$\exists v_0\, \neg \circ v_0\, e \equiv v_0$	$\exists$S auf 2
4.	Φ_{Gr}		$\circ v_0\, e \equiv v_0$	WR auf 1,3
		$\vdots$		

1.	Φ_{Gr}	$\neg \circ v_0\, e \equiv v_0$	$\neg\exists v_0\, \neg \circ v_0\, e \equiv v_0$	VR
2.	Φ_{Gr}	$\neg \circ v_0\, e \equiv v_0$	$\neg \circ v_0\, e \equiv v_0$	VR
3.	Φ_{Gr}	$\neg \circ v_0\, e \equiv v_0$	$\exists v_0\, \neg \circ v_0\, e \equiv v_0$	$\exists$S auf 2
4.	Φ_{Gr}		$\circ v_0\, e \equiv v_0$	WR auf 1,3
		$\vdots$		

1.	Φ_{Gr}	$\neg \circ v_0\, e \equiv v_0$	$\neg\exists v_0\, \neg \circ v_0\, e \equiv v_0$	VR
2.	Φ_{Gr}	$\neg \circ v_0\, e \equiv v_0$	$\neg \circ v_0\, e \equiv v_0$	VR
3.	Φ_{Gr}	$\neg \circ v_0\, e \equiv v_0$	$\exists v_0\, \neg \circ v_0\, e \equiv v_0$	$\exists$S auf 2
4.	Φ_{Gr}		$\circ v_0\, e \equiv v_0$	WR auf 1,3
		$\vdots$		

$\vdots$

21.	Φ_{Gr}	$\exists v_0 \neg(\neg \circ v_0\, e \equiv e \vee v_0 \equiv e)$	$\neg\exists v_0 \neg(\neg \circ v_0\, e \equiv e \vee v_0 \equiv e)$	$\exists$A auf 20
22.	Φ_{Gr}	$\neg\exists v_0 \neg(\neg \circ v_0\, e \equiv e \vee v_0 \equiv e)$	$\neg\exists v_0 \neg(\neg \circ v_0\, e \equiv e \vee v_0 \equiv e)$	VR
23.	Φ_{Gr}		$\neg\exists v_0 \neg(\neg \circ v_0\, e \equiv e \vee v_0 \equiv e)$	FU auf 21,22

Formal proofs - derivations

$$\frac{\Gamma \quad \varphi}{\Gamma \quad \psi \quad \varphi}$$

$$\frac{}{\Gamma \quad \varphi \quad \varphi}$$

$$\frac{\Gamma \quad \varphi \quad \psi}{\Gamma \quad \varphi \rightarrow \psi}$$

$$\frac{\Gamma \quad \varphi \qquad \Gamma \quad \varphi \rightarrow \psi}{\Gamma \quad \psi}$$

$$\frac{\Gamma \quad \varphi \qquad \Gamma \quad \neg\varphi}{\Gamma \quad \bot}$$

$$\frac{\Gamma \quad \neg\varphi \quad \bot}{\Gamma \quad \varphi}$$

$$\frac{\Gamma \quad \varphi\frac{y}{x}}{\Gamma \quad \forall x\varphi}\ ,$$

if $y \notin \mathrm{free}(\Gamma \cup \{\forall x\varphi\})$

$$\frac{\Gamma \quad \forall x\varphi}{\Gamma \quad \varphi\frac{t}{x}}$$

$$\frac{}{\Gamma \quad t \equiv t}$$

$$\frac{\Gamma \quad \varphi\frac{t}{x} \qquad \Gamma \quad t \equiv t'}{\Gamma \quad \varphi\frac{t'}{x}}$$

Formal proofs - derivations

- derivations are formed by repeated applications of (simple) syntactic rules

- whether a formal text is a derivation can (easily) be checked algorithmically

Formal proofs - derivations

N. Bourbaki:

If formalized mathematics were as simple as the game of chess, then once our chosen formalized language had been described there would remain only the task of writing out our proofs in this language, [...] But the matter is far from being as simple as that, and no great experience is necessary to perceive that such a project is absolutely unrealizable: the tiniest proof at the beginnings of the Theory of Sets would already require several hundreds of signs for its complete formalization. [...] formalized mathematics cannot in practice be written down in full, [...] We shall therefore very quickly abandon formalized mathematics, [...]

Formal proofs

Saunders Mac Lane:

As to precision, we have now stated an absolute standard of rigor: A mathematical proof is rigorous when it is (or could be) written out in the first-order predicate language $L(\in)$ as a sequence of inferences from the axioms ZFC, each inference made according to one of the stated rules. [...] When a proof is in doubt, its repair is usually a partial approximation to the fully formal version.

Computer-supported formal proofs

J. McCarthy:

Checking mathematical proofs is potentially one of the most interesting and useful applications of automatic computers. ... Proofs to be checked by computer may be briefer and easier to write than the informal proofs acceptable to mathematicians. This is because the computer can be asked to do much more work to check each step than a human is willing to do, and this permits longer and fewer steps.

McCarthy, J. "Computer Programs for Checking Mathematical Proofs," Proceedings of the Symposium in Pure Math, Recursive Function Theory,Volume V, pages 219-228, AMS, Providence, RI, 1962.

Automatic proof checker

Automath (~1967)

N.G. de Bruijn

From the Automath formalization of E. Landau, *Grundlagen der Analysis*, 1930 by L. S. van Benthem Jutting, 1979:

lich. Für die folgende spezielle Zahl ist aber ein kleiner lateinischer Buchstabe üblich auf Grund der

Definition 73:

$$i = [0,\ 1].$$

Satz 300:

$$ii = -1.$$

Beweis:

$$\begin{aligned} ii = [0,\ 1][0,\ 1] &= [0 \cdot 0 - 1 \cdot 1,\ 0 \cdot 1 + 1 \cdot 0] \\ &= [-1,\ 0] = -1. \end{aligned}$$

Satz 301: *Für reelle* u_1, u_2 *ist*

$$u_1 + u_2 i = [u_1,\ u_2].$$

```
ic:=pli(0,1rl):complex
+10300
t1:=tsis12a(0,1rl,0,1rl):is(ts(ic,ic),pli(mn"r"(ts"r"(0,0),ts"r"(1rl,1rl)),pl"r"(ts"r"(0,1rl),
ts"r"(1rl,0))))
t2:=tris(real,mn"r"(ts"r"(0,0),ts"r"(1rl,1rl)),m0"r"(ts"r"(1rl,1rl)),m0"r"(1rl),pl01(ts"r"(0,0),
m0"r"(ts"r"(1rl,1rl)),ts01(0,0,refis(real,0))),ism0"r"(ts"r"(1rl,1rl),1rl,satz195(1rl))):
is"r"(mn"r"(ts"r"(0,0),ts"r"(1rl,1rl)),m0"r"(1rl))
t3:=tris(real,pl"r"(ts"r"(0,1rl),ts"r"(1rl,0)),ts"r"(1rl,0),0,pl01(ts"r"(0,1rl),ts"r"(1rl,0),
ts01(0,1rl,refis(real,0))),ts02(1rl,0,refis(real,0))):is"r"(pl"r"(ts"r"(0,1rl),ts"r"(1rl,0)),0)
t4:=isrecx12(mn"r"(ts"r"(0,0),ts"r"(1rl,1rl)),m0"r"(1rl),pl"r"(ts"r"(0,1rl),
ts"r"(1rl,0)),0,t2,t3):is(pli(mn"r"(ts"r"(0,0),ts"r"(1rl,1rl)),
pl"r"(ts"r"(0,1rl),ts"r"(1rl,0))),cofrl(m0"r"(1rl)))
t5:=satz298j(1rl):is(cofrl(m0"r"(1rl)),m0(1c))
-10300
satz2300:=tr3is(cx,ts(ic,ic),pli(mn"r"(ts"r"(0,0),ts"r"(1rl,1rl)),
pl"r"(ts"r"(0,1rl),ts"r"(1rl,0))),cofrl(m0"r"(1rl)),m0(1c),t1".10300",t4".10300",t5".10300"):
is(ts(ic,ic),m0(1c))
```

The MIZAR system (1973 -) of Andrzej Trybulec

Language modeled after
"mathematical vernacular"

Natural deduction style

Automatic proof checker

Large mathematical library

Journal
Formalized Mathematics

www.mizar.org

MIZAR example: Proof of Pythagoras

```
theorem for p1,p2,p3 st p1<>p2 & p3<>p2 &
    (angle(p1,p2,p3)=PI/2 or angle(p1,p2,p3)=3/2*PI) holds
    (|.p1-p2.|^2+|.p3-p2.|^2=|.p1-p3.|^2
    proof let p1,p2,p3; assume A1: p1<>p2 & p3<>p2 &
    (angle(p1,p2,p3)=PI/2 or angle(p1,p2,p3)=3/2*PI);
  then A2: euc2cpx(p1)<> euc2cpx(p2) by Th6;
  A3: euc2cpx(p3)<> euc2cpx(p2) by A1,Th6;
  A4: euc2cpx(p1)-euc2cpx(p2)=euc2cpx(p1-p2) by Th19;
  A5: euc2cpx(p3)-euc2cpx(p2)=euc2cpx(p3-p2) by Th19;
  A6: euc2cpx(p1)-euc2cpx(p3)=euc2cpx(p1-p3) by Th19;
  A7: angle(p1,p2,p3)=angle(euc2cpx(p1),euc2cpx(p2),euc2cpx(p3))
by Def4;
  A8: |.euc2cpx(p1-p2).|=|.p1-p2.| by Th31;
  A9: |.euc2cpx(p3-p2).|=|.p3-p2.| by Th31;
   |.euc2cpx(p1-p3).|=|.p1-p3.| by Th31;
  hence thesis by A1,A2,A3,A4,A5,A6,A7,A8,A9,COMPLEX2:91;
end;
```

Derivations

- (Euclid)

- (Hilbert-style) calculi

- Automath

- MIZAR

What is a mathematical proof?

- description of the/some mathematical "reality"?
- argumentative text about the/some mathematical "reality"?
- argumentative text within some system of initial assumptions (axioms)?
- Wittgenstein: ..... ?
- abbreviation for some (long) formal derivation?
- recipe for building a formal derivation if required?
- a formal derivation in some very rich formal system (Montague: English as a formal language)?

Jody Azzouni: The derivation-indicator view of mathematical practice

ABSTRACT. A version of Formalism is vindicated: Ordinary mathematical proofs indicate (one or another) mechanically checkable derivation of theorems from the assumptions those ordinary mathematical proofs presuppose. The indicator view explains why mathematicians agree so readily on results established by proofs in ordinary language that are (palpably) not mechanically checkable. Mechanically checkable derivations in this way structure ordinary mathematical practice without its being the case that ordinary mathematical proofs can be 'reduced to' such derivations. In this way, one threat to formalist-style positions is removed: Platonic objects aren't needed to explain how mathematicians understand the import of ordinary mathematical proofs. (Philosophia Mathematica, 2004)

Derivation-indication

N. Bourbaki:

If formalized mathematics were as simple as the game of chess, ...

... there would remain only the task of writing out our proofs in this language, ...

Derivation-indication

Saunders Mac Lane:

As to precision, we have now stated an absolute standard of rigor: A mathematical proof is rigorous when it is (or could be) written out in the first-order predicate language $L(\in)$ as a sequence of inferences from the axioms ZFC, each inference made according to one of the stated rules. [...] When a proof is in doubt, its repair is usually a partial approximation to the fully formal version.

Derivation-indication

- Mathematicians agree that proofs can be written out in increasingly formal detail

- This leads to a fully formal derivation after some (long) finite time

- The indicator function lies mainly in the natural language parts of proofs

- Can one identify indicators by natural language processing?

- Derivations may be derivations performed by an Automatic Theorem Prover (ATP).

The Naproche project: Natural language proof checking

- studies the syntax and semantics of the language of proofs, emphasizing natural language and natural argumentation aspects
- models natural language proofs using computer-supported methods of formal linguistics and formal logic
- "reverse engineering" approach to derivation-indication
- joint work with Bernhard Schröder, linguistics; Bonn, Essen, Cologne; www.naproche.net
- development of a mathematical authoring system with a LaTeX-quality graphical interface

The Naproche project: Natural language proof checking

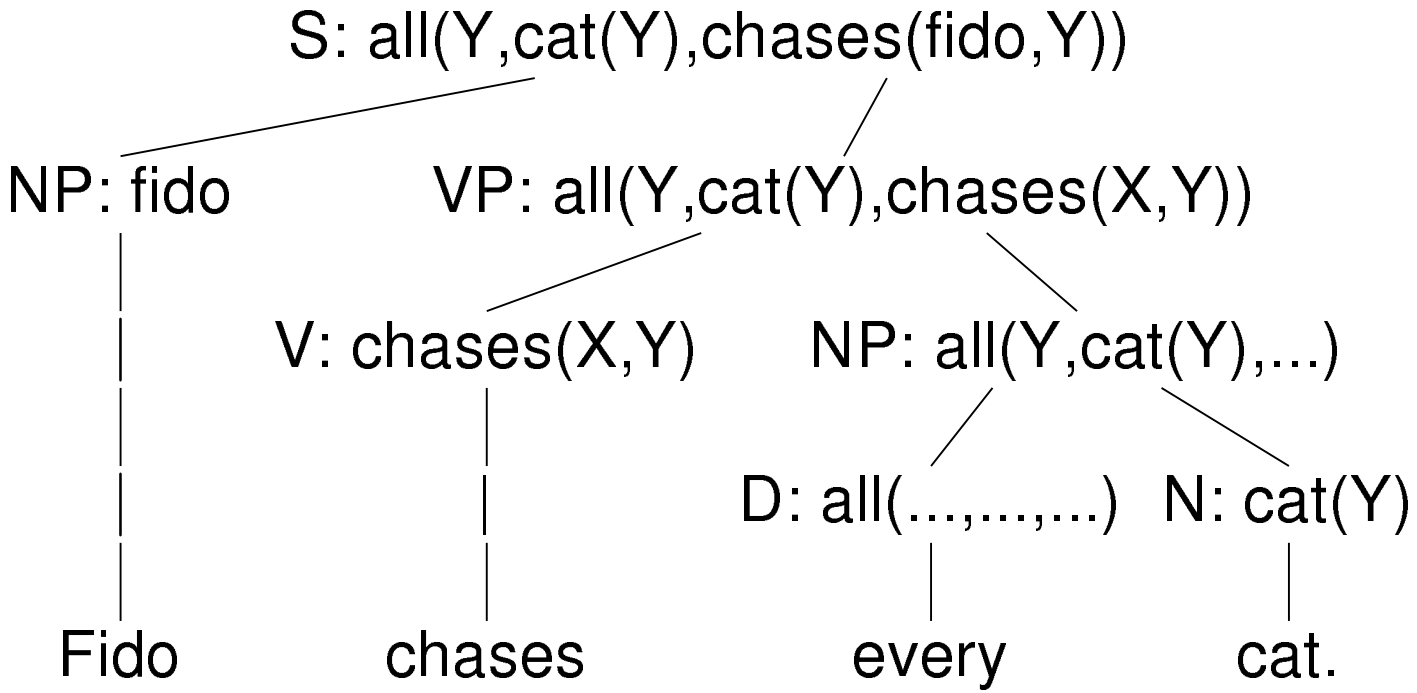
- To devise a strictly formal system for mathematics, implemented by computer, whose input language is an extensive part of the common mathematical language, and whose proof style is close to proof styles found in the mathematical literature.

Mathematical statements

“1 divides every integer.” $\longleftrightarrow$ “Fido chases every cat.”

Linguistic analysis

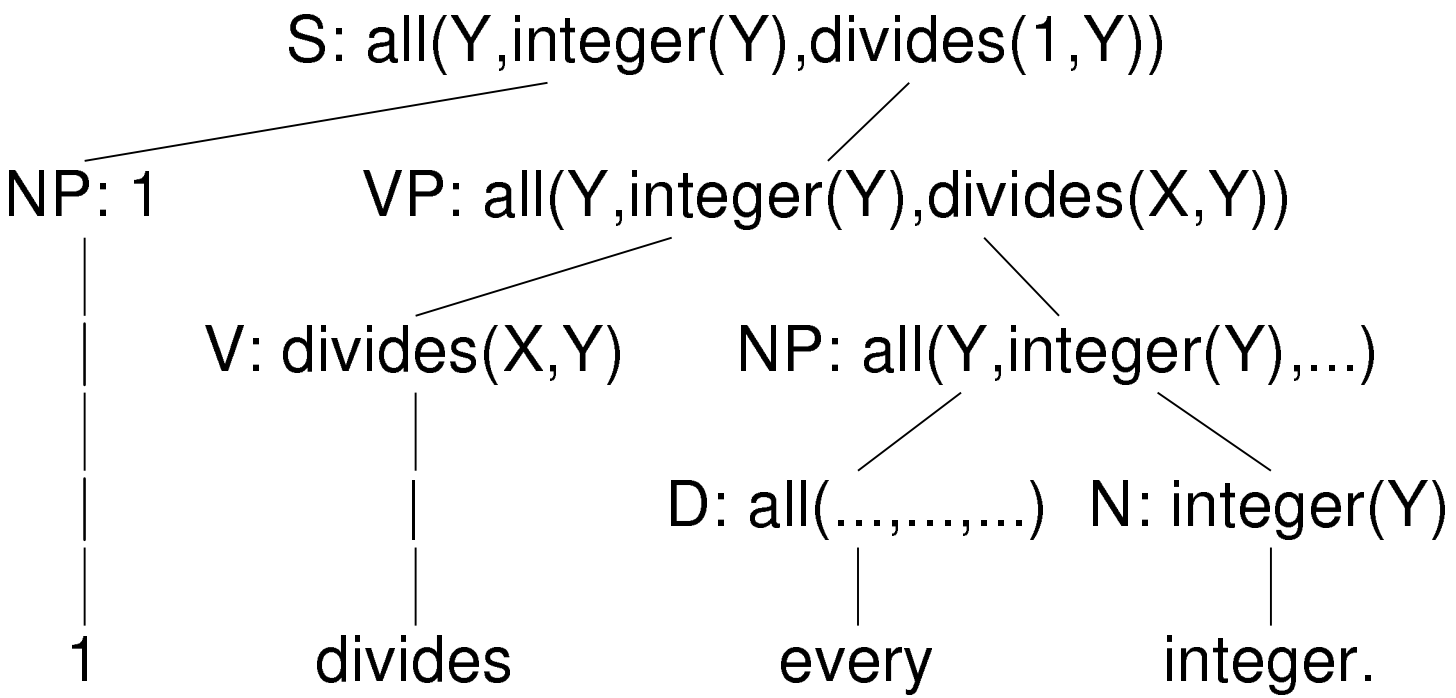
“Fido chases every cat.”



$\forall Y(\mathrm{cat}(Y) \rightarrow \mathrm{chases}(\mathrm{fido}, Y))$.

Linguistic analysis

“1 divides every integer.”



$\forall Y\,(\mathrm{integer}(Y) \rightarrow 1|Y).$

Layers of the Naproche system:

↓ Standard or web editor

TeX-style input text

↕ Natural language processing (NLP)

Proof representation structure (PRS)

↕ First-order translation

First-order logic format (TPTP)

↕ Proof checker or automatic theorem prover (ATP)

“Accepted”/“Not accepted”, with error messages

E. Landau, *Grundlagen der Analysis*, 1930: Theorem 30

Theorem 30 (Distributive Law):

$$x(y+z) = xy + xz.$$

Preliminary Remark: The formula

$$(y+z)x = yx + zx$$

which results from Theorem 30 and Theorem 29, and similar analogues later on, need not be specifically formulated as theorems, nor even be set down.

Proof: Fix x and y, and let $\mathfrak{M}$ be the set of all z for which the assertion holds true.

I) $x(y+1) = xy' = xy + x = xy + x \cdot 1$;

1 belongs to $\mathfrak{M}$.

II) If z belongs to $\mathfrak{M}$, then

$$x(y+z) = xy + xz,$$

hence

$$\begin{aligned} x(y+z') &= x((y+z)') = x(y+z) + x = (xy+xz) + x \\ &= xy + (xz + x) = xy + xz', \end{aligned}$$

so that z' belongs to $\mathfrak{M}$.

Therefore, the assertion always holds.

Theorem 30: For all x, y, z, $x * (y + z) = (x * y) + (x * z)$.
Proof:
Fix x, y. $x * (y + 1) = x * y' = x * y + x = (x * y) + (x * 1)$.
Now suppose $x * (y + z) = (x * y) + (x * z)$. Then $x * (y + z') = x * ((y + z)') =$
$(x * (y + z)) + x = ((x * y) + (x * z)) + x = (x * y) + ((x * z) + x) = (x * y) + (x * z')$.
Thus by induction, for all z $x * (y + z) = (x * y) + (x * z)$. Qed.

Components of the Naproche system: linguistic analysis

- standard analysis by a Prolog Definite Clause Grammar (DCG), the grammar defines a controlled natural language for mathematics (CNL), i.e. a formal subset of the common mathematical language

- translation into a formal semantics (without ambiguity)

Components of the Naproche system: linguistic analysis

- formal semantics: proof representation structures (PRS), extending discourse representation structures (DRS)

- DRS: tool for anaphor resolution (Let $\underline{x}$ be a set. It is ...) and for interpretation of natural language quantification (Every prime number is positive; a prime number is positive)

- PRS, moreover, represent global text structurings: Theorem / Proof, introductions and retractions of assumptions

Components of the Naproche system: Checking logical correctness

- translating the PRS conditions into some first-order format
- use TPTP-format (Thousands of Problems for Theorem Provers)
- generate relevant premises for every condition
- automatic theorem prover (ATP) used to prove every condition from its relevant premises
- proof is accepted if ATP can prove every condition
- feedback of success/error messages

Results

- The Naproche system allows natural reformulation of (simple) mathematical texts

- some example texts and parts of Landau, Foundations of Analysis have been reformulated and checked

Chapter 1 from Landau in Naproche

by Merlin Carl, Marcos Cramer, Daniel Khlwein

November 3, 2009

Abstract

This is a reformulation of the first chapter of Landau's *Grundlagen der Analysis* in the Controlled Natural Language of Naproche. Talk about sets is still avoided. One consequence of this is that Axiom 5 (the induction axiom) cannot be formulated; instead we use an induction proof method.

Axiom 3: For every x, $x' \neq 1$.

Axiom 4: If $x' = y'$, then $x = y$.

Theorem 1: If $x \neq y$ then $x' \neq y'$.
Proof:
Assume that $x \neq y$ and $x' = y'$. Then by axiom 4, $x = y$. Qed.

Theorem 2: For all x $x' \neq x$.
Proof:
By axiom 3, $1' \neq 1$. Suppose $x' \neq x$. Then by theorem 1, $(x')' \neq x'$. Thus by induction, for all x $x' \neq x$. Qed.

Theorem 3: If $x \neq 1$ then there is a u such that $x = u'$.
Proof:
If $1 \neq 1$ then there is a u such that $1 = u'$.
Assume $x' \neq 1$. If $u = x$ then $x' = u'$. So there is a u such that $x' = u'$.
Thus by induction, if $x \neq 1$ then there is a u such that $x = u'$. Qed.

Definition 1:
Define $+$ recursively:
$x + 1 = x'$.
$x + y' = (x + y)'$.

Theorem 5: For all x, y, z, $(x + y) + z = x + (y + z)$.
Proof:
Fix x, y.
$(x + y) + 1 = (x + y)' = x + y' = x + (y + 1)$.
Assume that $(x + y) + z = x + (y + z)$. Then $(x + y) + z' = ((x + y) + z)' = (x + (y + z))' = x + (y + z)' = x + (y + z')$. So $(x + y) + z' = x + (y + z')$.
Thus by induction, for all z, $(x + y) + z = x + (y + z)$. Qed.

Lemma 4a: For all y, $1 + y = y'$.
Proof:
By definition 1, $1 + 1 = 1'$.
Suppose $1 + y = y'$. Then by definition 1, $1 + y' = (1 + y)'$. So $1 + y' = (y')'$.
Thus by induction, for all y $1 + y = y'$. Qed.

Possible applications

- Natural language interfaces to formal mathematics
- Mathematical authoring and checking tools
- writing texts that are simultaneously acceptable by human readers and formal mathematics systems (“Logic for men and machines”)
- Tutorial applications: teaching how to prove

General issues

- Naproche attempts to implement parts of the derivation-indication approach to proofs
- natural language components serve as indicators
- there are natural(ly looking) proofs that are fully formal with respect to the Naproche system
- this defines a "fortified formalism", using linguistic methods and computer implementations, which allows to view some natural proofs as fully formal
- can a "fortified formalism" help to mediate between the "two streams" in the philosophy of mathematics (formalistic / naturalistic)?

Thank You!