Proof Checking with a Natural Language Interface

Peter Koepke, University of Bonn, Germany

Logical Methods in the Humanities Workshop

Stanford, April 12, 2006

Contents

- The Gödel completeness theorem
- Automatic proof checking: *Automath*
- Automatic proof checking: MIZAR
- MIZAR proof of the Gödel completeness theorem
- NAtural language PROof CHEcking: NAPROCHE
- Linguistics of mathematical language
- TeX-quality interface: $\TeX_{MACS}$
- Examples and demonstration

The Gödel Completeness Theorem

Theorem 1: Every valid formula of the special function calculus is provable.

Satz 1: Jede allgemeingültige Formel des engeren Funktionenkalküls ist beweisbar.
(Kurt Gödel, *Die Vollständigkeit der Axiome des logischen Funktionenkalküls*, 1930)

A first-order sequent calculus

Antecedens	$\dfrac{\Phi\varphi}{\Psi\varphi}$	$(\Phi \subseteq \Psi)$
Premises	$\dfrac{}{\Phi\varphi}$	$(\varphi \in \Phi)$
Cases	$\dfrac{\Phi\varphi\psi \quad \Phi\neg\varphi\psi}{\Phi\psi}$	
Contradiction	$\dfrac{\Phi\neg\varphi\psi \quad \Phi\neg\varphi\neg\psi}{\Phi\varphi}$	
$\vee$-Introduction I	$\dfrac{\Phi\varphi\chi \quad \Phi\psi\chi}{\Phi(\varphi\vee\psi)\chi}$	
$\vee$-Introduction II	$\dfrac{\Phi\varphi}{\Phi(\varphi\vee\psi)}$	
$\vee$-Introduction III	$\dfrac{\Phi\varphi}{\Phi(\psi\vee\varphi)}$	
Equality	$\dfrac{}{t\equiv t}$	$(t \in T^S)$
$\exists$-Introduction I	$\dfrac{\Phi\varphi\frac{t}{x}}{\Phi\exists x\varphi}$	
$\exists$-Introduction II	$\dfrac{\Phi\varphi\frac{y}{x}\psi}{\Phi\exists x\varphi\psi}$	$(y \notin \mathrm{free}(\Phi \cup \{\exists x\varphi, \psi\}))$
Substitution	$\dfrac{\Phi\varphi\frac{t}{x}}{\Phi t\equiv t'\varphi\frac{t'}{x}}$	

Example: fragment from group theory

25.$\Phi \vdash (v\dot{\circ}\dot{e})\dot{\circ}w = x\dot{\circ}w\frac{v\dot{\circ}\dot{e}}{x}$ (AR)

26.$\Phi \vdash \underline{(v\dot{\circ}\dot{e})\dot{\circ}w \doteq v\dot{\circ}w}$ ((=) angewendet auf 24.,25.)

27.$\Phi \vdash \underline{v\dot{\circ}w \doteq \dot{e}}$ (AR)

Wir fassen die Kette der unterstrichenen Gleichungen zusammen:

28.$\Phi \vdash v\dot{\circ}u \doteq (v\dot{\circ}u)\dot{\circ}(v\dot{\circ}w)$ (Transitivität von $\doteq$ angewendet auf 3.,6.)

29.$\Phi \vdash v\dot{\circ}u \doteq v\dot{\circ}(u\dot{\circ}(v\dot{\circ}w))$ (Transitivität von $\doteq$ angewendet auf 28.,10.)

30.$\Phi \vdash v\dot{\circ}u \doteq v\dot{\circ}((u\dot{\circ}v)\dot{\circ}w)$ (Transitivität von $\doteq$ angewendet auf 29.,16.)

31.$\Phi \vdash v\dot{\circ}u \doteq v\dot{\circ}(\dot{e}\dot{\circ}w)$ (Transitivität von $\doteq$ angewendet auf 30.,19.)

32.$\Phi \vdash v\dot{\circ}u \doteq (v\dot{\circ}\dot{e})\dot{\circ}w$ (Transitivität von $\doteq$ angewendet auf 31.,23.)

33.$\Phi \vdash v\dot{\circ}u \doteq v\dot{\circ}w$ (Transitivität von $\doteq$ angewendet auf 32.,26.)

34.$\Phi \vdash v\dot{\circ}u \doteq \dot{e}$ (Transitivität von $\doteq$ angewendet auf 33.,27.)

Wir eliminieren die Annahme „$v\dot{\circ}w \doteq \dot{e}$“: 34. bedeutet

34.$\Phi_{\mathrm{Gr}}, u\dot{\circ}v \doteq \dot{e}, v\dot{\circ}w \doteq \dot{e} \vdash v\dot{\circ}u \doteq \dot{e}$

35.$\Phi_{\mathrm{Gr}}, u\dot{\circ}v \doteq \dot{e}, \dot{\neg}\, v\dot{\circ}u \doteq \dot{e} \vdash \dot{\neg}\, v\dot{\circ}w \doteq \dot{e}$ (Kontraposition angewendet auf 34.)

36.$\Phi_{\mathrm{Gr}}, u\dot{\circ}v \doteq \dot{e}, \dot{\neg}\, v\dot{\circ}u \doteq \dot{e} \vdash \dot{\forall}y\, \dot{\neg}\, v\dot{\circ}y \doteq \dot{e}$ (($\forall$ 2) angewendet auf 35.)

37.$\Phi_{\mathrm{Gr}}, u\dot{\circ}v \doteq \dot{e}, \dot{\neg}\, v\dot{\circ}u \doteq \dot{e} \vdash \dot{\forall}x\, \dot{\neg}\dot{\forall}y\, \dot{\neg}\, x\dot{\circ}y \doteq \dot{e}$ ((AR); beachte $\dot{\neg}\dot{\forall}y\, \dot{\neg}\, x\dot{\circ}y \doteq \dot{e} \;=\; \dot{\exists}y\; x\dot{\circ}y \doteq \dot{e}$)

38.$\Phi_{\mathrm{Gr}}, u\dot{\circ}v \doteq \dot{e}, \dot{\neg}\, v\dot{\circ}u \doteq \dot{e} \vdash \dot{\neg}\dot{\forall}y\, \dot{\neg}\, v\dot{\circ}y \doteq \dot{e}$ (($\forall$ 1) angewendet auf 37.)

39.$\Phi_{\mathrm{Gr}}, u\dot{\circ}v \doteq \dot{e} \vdash v\dot{\circ}u \doteq \dot{e}$ (Prinzip des Widerspruchsbeweises angewendet auf 36.,38.)

Wir übernehmen abschließend „$u\dot{\circ}v \doteq \dot{e}$“ in die Folgerung. Nach 16.20 folgt aus 39.:

40.$\Phi_{\mathrm{Gr}} \vdash \dot{(}u\dot{\circ}v \doteq \dot{e} \dot{\to} v\dot{\circ}u \doteq \dot{e}\dot{)}$

41.$\Phi_{\mathrm{Gr}} \vdash \dot{\forall}v\, \dot{(}u\dot{\circ}v \doteq \dot{e} \dot{\to} v\dot{\circ}u \doteq \dot{e}\dot{)}$ (($\forall$ 2) angewendet auf 40.)

42.$\Phi_{\mathrm{Gr}} \vdash \dot{\forall}u\dot{\forall}v\, \dot{(}u\dot{\circ}v \doteq \dot{e} \dot{\to} v\dot{\circ}u \doteq \dot{e}\dot{)}$ (($\forall$ 2) angewendet auf 41.)

- A formal proof in the sequent calculus can be recursively checked for correctness: *Proof checking*.

- In principle, proofs of valid formulas can be found by searching through the set of all proofs: *Automatic proving*.

- Can the format of formal proofs be brought together with the informal presentation of proofs in mathematical practice? *"Natural proof checking"*?

N.G. de Bruijn (*1918)

First automatic proof checker:

Automath (~1967)

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

hen. 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 vernecular”

Natural deduction style

Automatic proof checker

Large mathematical library

Journal
Formalized Mathematics

www.mizar.org

From: Robert Solovay, *Fibonacci Numbers*
(MML-Identifier: `FIB_NUM`), *JFM* 14, 2002:

```
...

Lm11: sqrt 5 < 3

proof 3^2 = 3 * 3 by SQUARE_1:def 3
         .= 9;
then sqrt 9 = 3 by SQUARE_1:89;
hence thesis by SQUARE_1:95;
end;

...
```

References to: Andrzej Trybulec, Czeslaw Bylinski:
Some Properties of Real Numbers Operations:
min, max, square, and square root
(MML identifier: SQUARE_1), *JFM* 1, 1989

P. Braselmann and K., A formal proof of Gödel's completeness theorem, a series of 7 articles in:
Formalized Mathematics 13 (2005), 5-53

corresponding to the MIZAR articles

1. `SUBSTUT1.MIZ`: Definition of substitution
2. `SUBSTUT2.MIZ`: Technical facts about substitutions
3. `SUBLEMMA.MIZ`: The substitution lemma
4. `CALCUL_1.MIZ`: Sequent calculus; correctness
5. `CALCUL_2.MIZ`: Technical facts about the sequent calculus
6. `HENMODEL.MIZ`: Consistency; construction of Henkin-model
7. `GOEDELCP.MIZ`: Proof of the Gödel Completeness Theorem

From *Formalized Mathematics*

snb(f).

(31) If snb(C_1) is finite, then there exists C_2 such that $C_1 \subseteq C_2$ and C_2 has examples.

(32) If $X \vdash p$ and $X \subseteq Y$, then $Y \vdash p$.

(33) If C_1 has examples, then there exists C_2 such that $C_1 \subseteq C_2$ and C_2 is negation faithful and has examples.

In the sequel J_2 is a Henkin interpretation of C_3.

Next we state the proposition

(34) If snb(C_1) is finite, then there exist C_3, J_2 such that $J_2, \mathrm{valH} \models C_1$.

3. Gödel's Completeness Theorem

One can prove the following proposition

(35) If snb(X) is finite and $X \models p$, then $X \vdash p$.

References

[1] Grzegorz Bancerek. Connectives and subformulae of the first order language. *Formalized Mathematics*, 1(**3**):451–458, 1990.

[2] Grzegorz Bancerek. The fundamental properties of natural numbers. *Formalized Mathematics*, 1(**1**):41–46, 1990.

[3] Grzegorz Bancerek. The ordinal numbers. *Formalized Mathematics*, 1(**1**):91–96, 1990.

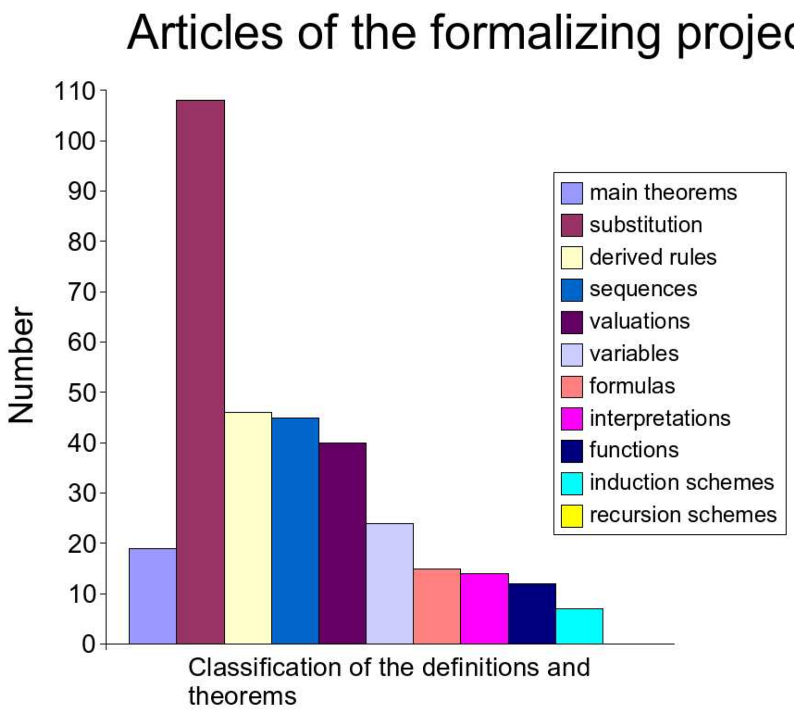
[4] Grzegorz Bancerek. Countable sets and Hessenberg's theorem. *Formalized Mathematics*, 2(**1**):65–69, 1991.

Original MIZAR in `GOEDELCP.MIZ`:

```
begin :: Goedel's Completeness Theorem,
:: Ebb et al, Chapter V, Completeness Theorem 4.1
theorem
  still_not-bound_in X is finite & X |= p implies X |- p
  proof
    assume A1: still_not-bound_in X is finite & X |= p;
    now assume not X |- p; then
      reconsider CX = X \/ {'not' p} as Consistent Subset of CQC-WFF
        by HENMODEL:9;
A2:  for A,J,v holds not J,v |= CX
```

```
    proof
      let A,J,v;
      now assume A3: J,v |= X \/ {'not' p};
        now let q such that A4: q in X;
          X c= X \/ {'not' p} by XBOOLE_1:7;
          hence J,v |= q by A3,A4,CALCUL_1:def 11;
        end; then
A5:     J,v |= X by CALCUL_1:def 11;
        now let q such that A6: q in {'not' p};
          {'not' p} c= X \/ {'not' p} by XBOOLE_1:7;
          hence J,v |= q by A3,A6,CALCUL_1:def 11;
        end; then
A7:     J,v |= {'not' p} by CALCUL_1:def 11;
        'not' p in {'not' p} by TARSKI:def 1; then
        J,v |= 'not' p by A7,CALCUL_1:def 11; then
        J,v |= X & not J,v |= p by A5,VALUAT_1:28;
        hence contradiction by A1,CALCUL_1:def 12;
      end;
      hence not J,v |= CX;
    end;
    still_not-bound_in 'not' p is finite by CQC_SIM1:20; then
    still_not-bound_in {'not' p} is finite by Th26; then
    still_not-bound_in X \/
    still_not-bound_in {'not' p} is finite by A1,FINSET_1:14; then
    still_not-bound_in CX is finite by Th27; then
    consider CZ,JH1 such that A8: JH1,valH |= CX by Th34;
    thus contradiction by A2,A8;
  end;
  hence thesis;
end;
```

Statistics:


Articles of the formalizing project
Number
0
10
20
30
40
50
60
70
80
90
100
110
Classification of the definitions and theorems
main theorems
substitution
derived rules
sequences
valuations
variables
formulas
interpretations
functions
induction schemes
recursion schemes

The project NAPROCHE: NAtural language PROof CHEcker

- natural language constructs, gramatically correct and varied;
- input through TeX quality WYSIWYG editor
- interactive proof checking

... From a linguistic perspective, the Language of Mathematics is distinguished by the fact that its core mathematical meaning can be fully captured by an intelligent translation into first-order predicate logic. ...

The ... project NAPROCHE aims at constructing a system which accepts a controlled but rich subset of ordinary mathematical language including TeX-style typeset formulas and transforms them into formal statements. We adapt linguistic techniques to allow for common grammatical constructs and to extract mathematically relevant implicit information about hypotheses and conclusions. Combined with proof checking software we obtain NAtural language PROof CHEckers which are prototypically used ... to teach mathematical proving.

Layers of a NAPROCHE system:

WYSIWIG mathematical text

↕ $\text{T}_\text{E}\text{X}_\text{MACS}$

TeX-style internal format with editing information

↕ Tokenizer

Tokenized format

↕ NLP (natural language processing)

First-order logic format

↕ Proof checker

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

NLP: direct translations between natural language and first order logic:

For every $y \in \mathbb{Z}$ holds $1|y$.

$$\updownarrow$$

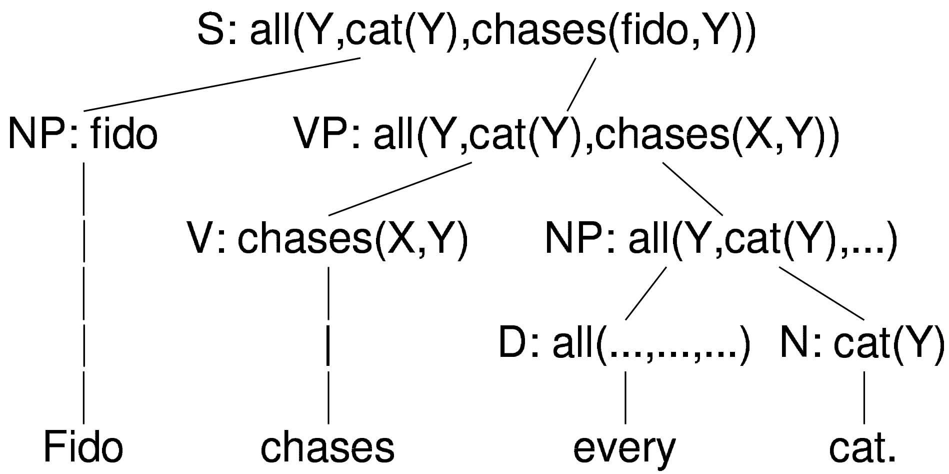
$$\forall y \in \mathbb{Z} \, 1|y$$

$$\updownarrow$$

all(Y,integer(Y),divides(1,Y))

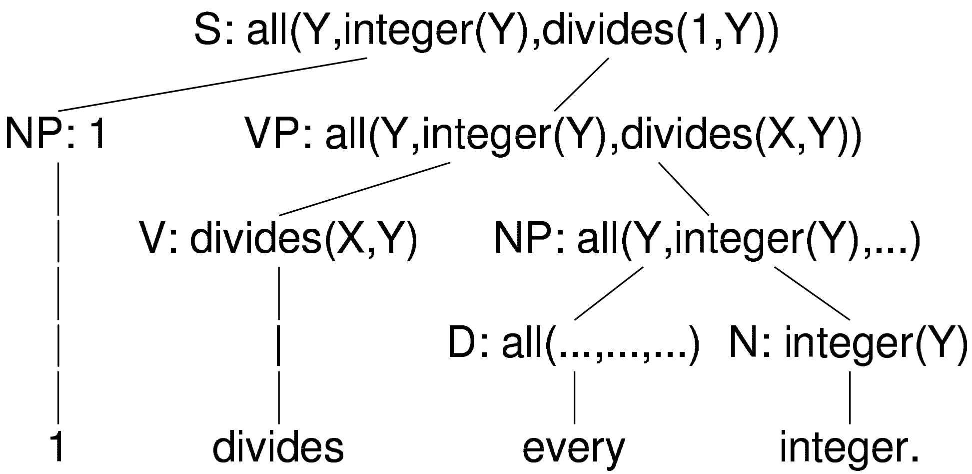
NLP: Semantics of simple natural language

"Fido chases every cat"



NLP: Semantics of simple mathematical language

“1 divides every integer.”



i.e., $\forall y \in \mathbb{Z}\ 1|y$

Further linguistic issues in mathematical texts

- ensure grammatical correctness by grammars
- mixture of text and mathematical formulas:
  pass the formulas unchanged through the NLP layer
- resolution of anaphors: *let X be a set of integers and let m be its maximal element*. Use standard NLP methods
- identification of mathematical *keywords* structuring a text:
  Proof, qed, define, ...
- handling of ellipses: $1, 2, ..., n$

The mathematical WYSIWYG editor TeXmacs

- www.texmacs.org, GNU General Public License, under development
- TeX/LaTeX-like file format and instant on-screen rendering using the TeX font system and TeX typesetting algorithms α
- on-screen editing
- uses `scheme` as extension language

Theorem. $(\neg\varphi \vee \psi) \rightarrow (\varphi \rightarrow \psi)$.
Proof.
Let $(\neg\varphi \vee \psi)$.
Let $\neg\varphi$. Let φ. Contradiction. ψ. Thus $\varphi \rightarrow \psi$. Thus $\neg\varphi \rightarrow (\varphi \rightarrow \psi)$.
Let ψ. Let φ. ψ. Thus $\varphi \rightarrow \psi$. Thus $\psi \rightarrow (\varphi \rightarrow \psi)$.
$\varphi \rightarrow \psi$. Thus $(\neg\varphi \vee \psi) \rightarrow (\varphi \rightarrow \psi)$.
Qed.

Internal representation (.tm file)

```
<TeXmacs|1.0.6>
<style|generic>
<\body>
  Example:
  <\quotation>
    Theorem. <with|mode|math|(\<neg\>\<varphi\>\<vee\>\<psi\>)\<rightarrow\>
              (\<varphi\>\<rightarrow\>\<psi\>)>.\
    Proof.
    Let <with|mode|math|(\<neg\>\<varphi\>\<vee\>\<psi\>)>.
    Let <with|mode|math|\<neg\>\<varphi\>>. Let <with|mode|math|\<varphi\>>.
    Contradiction. <with|mode|math|\<psi\>>. Thus
    <with|mode|math|\<varphi\>\<rightarrow\>\<psi\>>. Thus
    <with|mode|math|\<neg\>\<varphi\>\<rightarrow\>(\<varphi\>\<rightarrow\>\<psi\>)>.
    Let <with|mode|math|\<psi\>>. Let <with|mode|math|\<varphi\>>.
    <with|mode|math|\<psi\>>. Thus <with|mode|math|\<varphi\>\<rightarrow\>\<psi\>>.
    Thus <with|mode|math|\<psi\>\<rightarrow\>(\<varphi\>\<rightarrow\>\<psi\>)>.
    <with|mode|math|\<varphi\>\<rightarrow\>\<psi\>>. Thus
<with|mode|math|(\<neg\>\<varphi\>\<vee\>\<psi\>)\<rightarrow\>
```

```
(\<varphi\>\<rightarrow\>\<psi\>)>.
    Qed.
  </quotation>
</body>
```