Ein Assistenzsystem für die Mathematik

Christoph Benzmüller

FR Informatik, Universität des Saarlandes

Tag der offenen Tür, Saarbrücken, 5. Juli 2003

Inhalt des Vortrag

Vision eine leistungsfägigen und integrierten mathematischen Assistenzsystems

Prototyp entwickelt an der Uni des Saarlandes:

Ωmega

(AG Prof. Siekmann)

Demonstration durch: M. Pollet und A. Fiedler

12:30 Uhr, Foyer

Mathematische Assistenzsysteme

Assistenzsystem für die Mathematik:

- Beweisen mathematischer Aussagen
- Mathematische Berechnungen
- Verwaltung mathematischen Wissens in Datenbanken
- Multi-modale Interaktion mit dem Mathematiker:
  Graphische Repräsentation, Hypertext, Dialog
- Lehren mathematischer Inhalte
- Exploration neuen mathematischen Wissens
- Verifikation mathematischer Texte

$\Rightarrow$ Ressourcenbündelung erforderlich

Beweisen mathematischer Aussagen

Frege, Russel, Hilbert Prädikatenkalkül und Typentheorie als formale Basis für die Mathematik

$$\forall x, y, z.(x + (y + z)) = ((x + y) + z)$$

Gentzen Kalkül des Natürlichen Schließens (ND)

ND-Regeln (Bsp.)

$$\frac{\mathbf{A} \Rightarrow \mathbf{B} \quad \mathbf{A}}{\mathbf{B}}\ mp$$

$$\frac{\mathbf{A} \quad \mathbf{B}}{\mathbf{A} \wedge \mathbf{B}}\ \wedge I \qquad \frac{\mathbf{A} \wedge \mathbf{B}}{\mathbf{A}}\ \wedge E_l \qquad \frac{\mathbf{A} \wedge \mathbf{B}}{\mathbf{B}}\ \wedge E_r$$

$$\frac{\begin{array}{c}[\mathbf{A}]_1 \\ \vdots \\ \mathbf{B}\end{array}}{\mathbf{A} \Rightarrow \mathbf{B}}\ \Rightarrow I^1$$

... usw. ...

ND-Beweis für $(\mathbf{A} \wedge \mathbf{B}) \Rightarrow (\mathbf{B} \wedge (\mathbf{C} \vee \mathbf{A}))$

$$\dfrac{\dfrac{\dfrac{[\mathbf{A} \wedge \mathbf{B}]_1}{\mathbf{B}}\ \wedge E_r \qquad \dfrac{\dfrac{[\mathbf{A} \wedge \mathbf{B}]_1}{\mathbf{A}}\ \wedge E_l}{\mathbf{C} \vee \mathbf{A}}\ \vee I_r}{\mathbf{B} \wedge (\mathbf{C} \vee \mathbf{A})}\ \wedge I}{(\mathbf{A} \wedge \mathbf{B}) \Rightarrow (\mathbf{B} \wedge (\mathbf{C} \vee \mathbf{A}))}\ \Rightarrow I^1$$

Beweisen mathematischer Aussagen

Robinson (1965): Resolutionskalkül als Grundlage zur Automatisierung

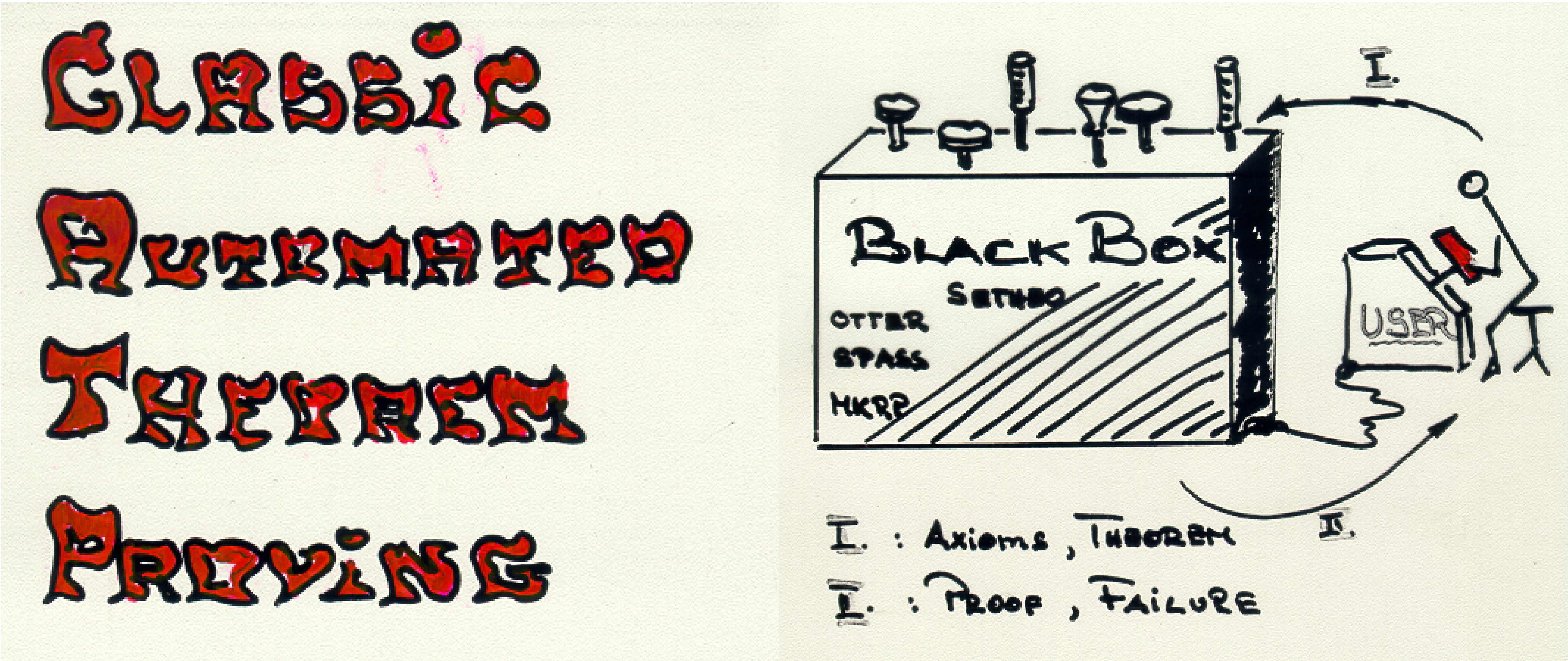


Bild: Jörg Siekmann

Beweisen mathematischer Aussagen

Erfolge: Robbins Lemma wurde erstmals bewiesen mit Maschine (EQP) in 1997

Beispielbeweis durch OTTER für: $\sqrt{2}$ ist irrational

Problemeingabe

```
%Here's an input file that gets a proof quickly.
%Note that he has a cancellation rule for multiplication.

set(auto).
set(ur_res).
assign(max_distinct_vars, 1).
list(usable).
x = x.
m(1,x) = x.                          %identity
m(x,1) = x.
m(x,m(y,z)) = m(m(x,y),z).           %associativity
m(x,y) = m(y,x).                     %commutativity
m(x,y) != m(x,z) | y = z.            %cancellation
-d(x,y) | m(x,f(x,y)) = y.           %this and next line define divides
m(x,z) != y | d(x,y).
-d(2,m(x,y)) | d(2,x) | d(2,y).      % 2 is prime  (with 12)
m(a,a) = m(2,m(b,b)).                % a/b = sqrt(2)
-d(x,a) | -d(x,b) | x = 1.           % a/b is in lowest terms
2 != 1.                              % I almost forgot this!
end_of_list.
```

Beweisausgabe

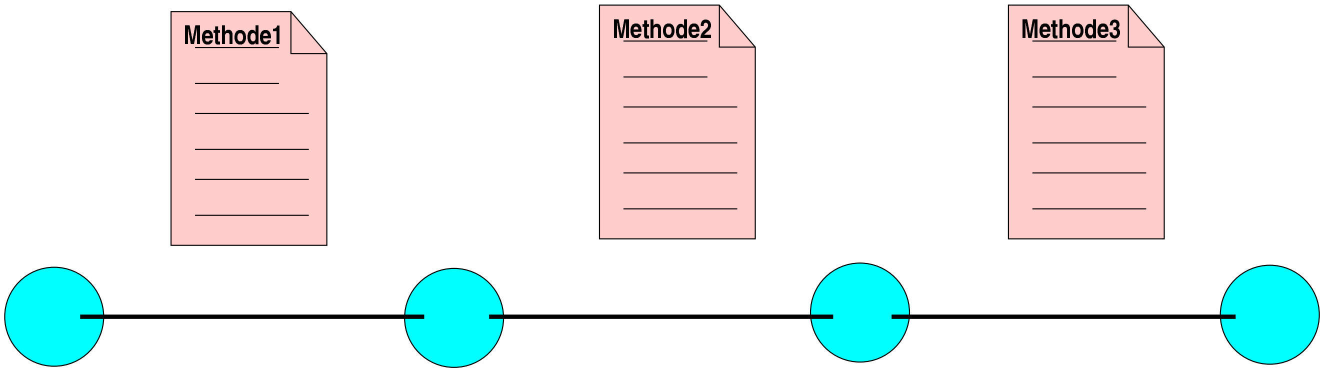
```
---------------- PROOF ----------------

1 [] m(x,y)!=m(x,z)|y=z.
2 [] -d(x,y)|m(x,f(x,y))=y.
3 [] m(x,y)!=z|d(x,z).
4 [] -d(2,m(x,y))|d(2,x)|d(2,y).
5 [] -d(x,a)| -d(x,b)|x=1.
6 [] 2!=1.
7 [factor,4,2,3] -d(2,m(x,x))|d(2,x).
13 [] m(x,m(y,z))=m(m(x,y),z).
14 [copy,13,flip.1] m(m(x,y),z)=m(x,m(y,z)).
16 [] m(x,y)=m(y,x).
17 [] m(a,a)=m(2,m(b,b)).
18 [copy,17,flip.1] m(2,m(b,b))=m(a,a).
30 [hyper,18,3] d(2,m(a,a)).
39 [para_from,18.1.1,1.1.1] m(a,a)!=m(2,x)|m(b,b)=x.
42 [hyper,30,7] d(2,a).
46 [hyper,42,2] m(2,f(2,a))=a.
48 [ur,42,5,6] -d(2,b).
50 [ur,48,7] -d(2,m(b,b)).
59 [ur,50,3] m(2,x)!=m(b,b).
60 [copy,59,flip.1] m(b,b)!=m(2,x).
145 [para_from,46.1.1,14.1.1.1,flip.1] m(2,m(f(2,a),x))=m(a,x).
189 [ur,60,39] m(a,a)!=m(2,m(2,x)).
190 [copy,189,flip.1] m(2,m(2,x))!=m(a,a).
1261 [para_into,145.1.1.2,16.1.1] m(2,m(x,f(2,a)))=m(a,x).
1272 [para_from,145.1.1,190.1.1.2] m(2,m(a,x))!=m(a,a).
1273 [binary,1272.1,1261.1] $F.

------------ end of proof -------------
```

Beweisen mathematischer Aussagen:

Beweisplanen: Domänenspezifisches Schließen auf abstrakterer Ebene



Beispiele für Beweismethoden:

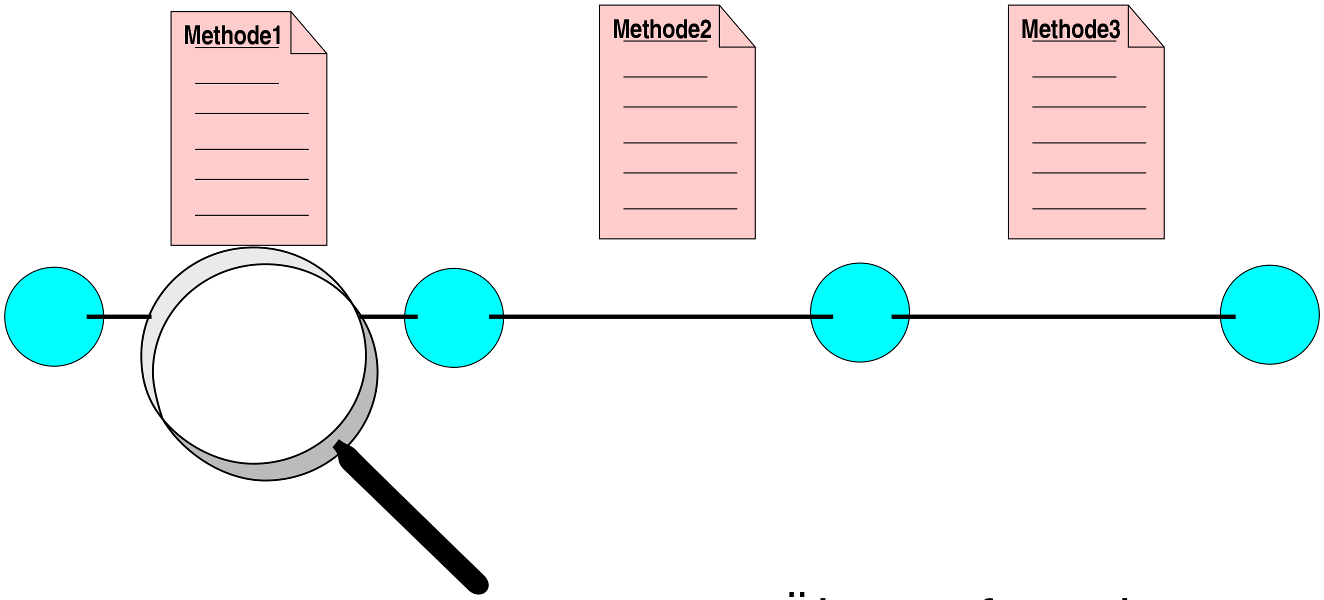
- Diagonalisierungsprinzip
- Induktionsbeweis

\+ heuristische Steuerung

Klassische Automatische Beweiser:

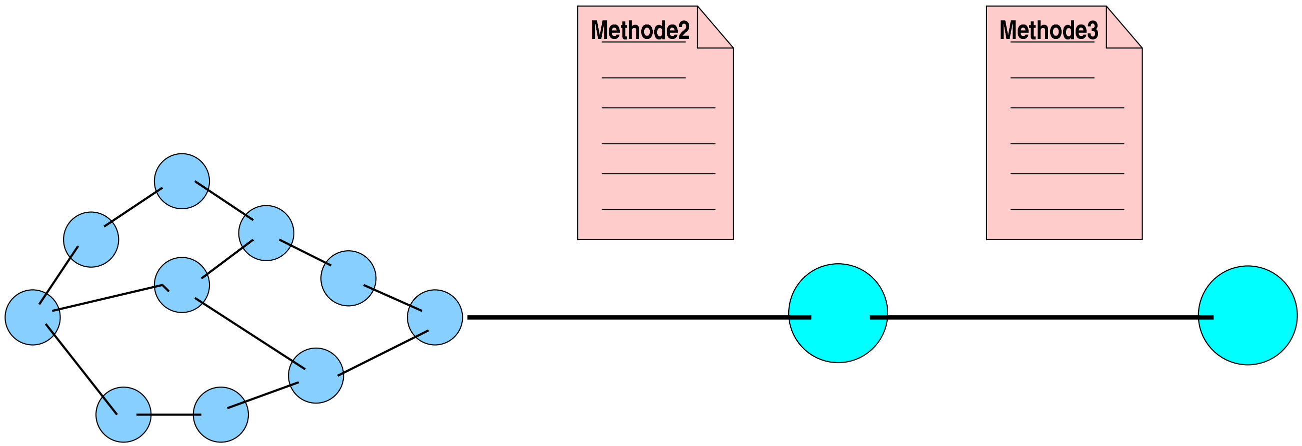
- Integration in Beweismethoden
- und in Steuerungsheuristiken

Beweisen mathematischer Aussagen:



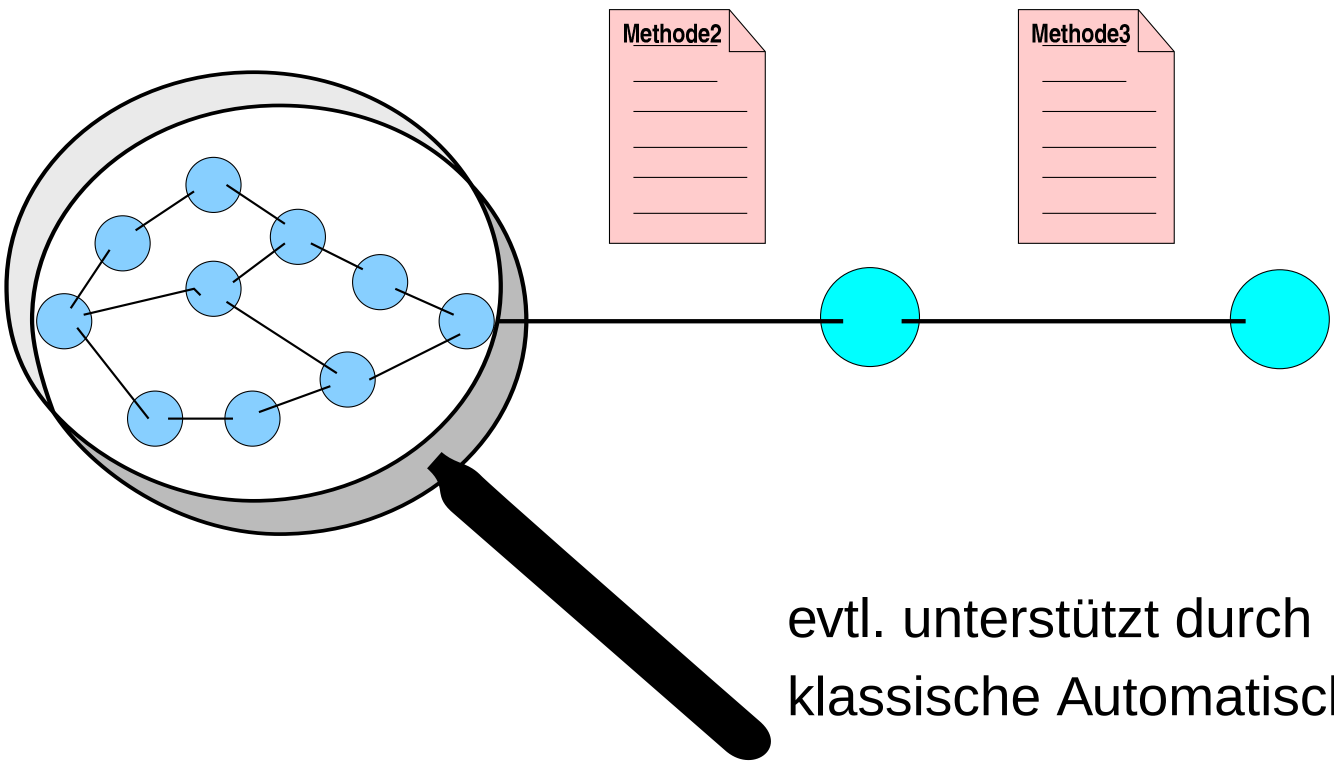
Überprüfung der Korrektheit durch . . .

Beweisen mathematischer Aussagen:



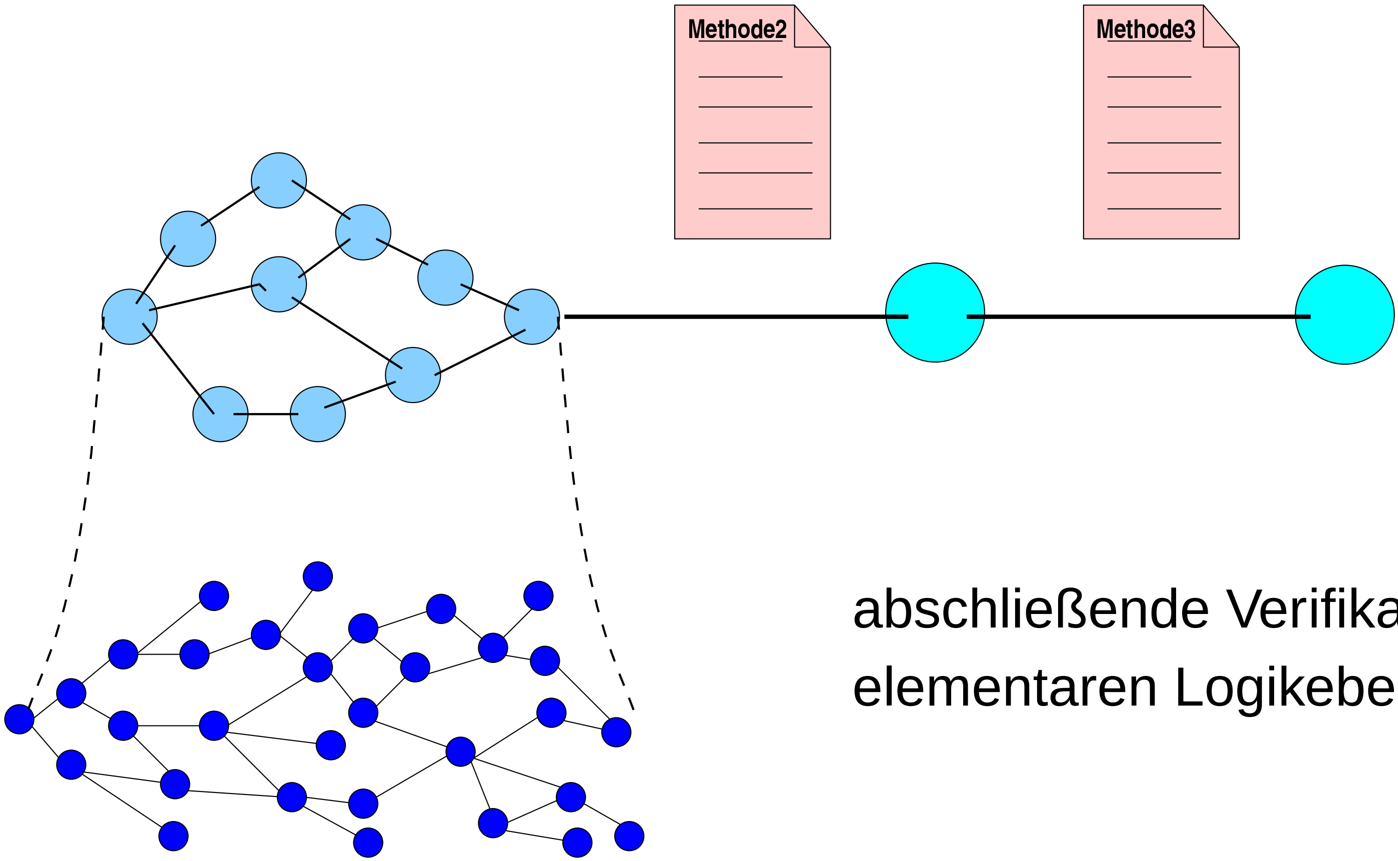
Beweisverfeinerung (Expansion)
über mehrere Ebenen
...

Beweisen mathematischer Aussagen:



evtl. unterstützt durch
klassische Automatische Beweiser . . .

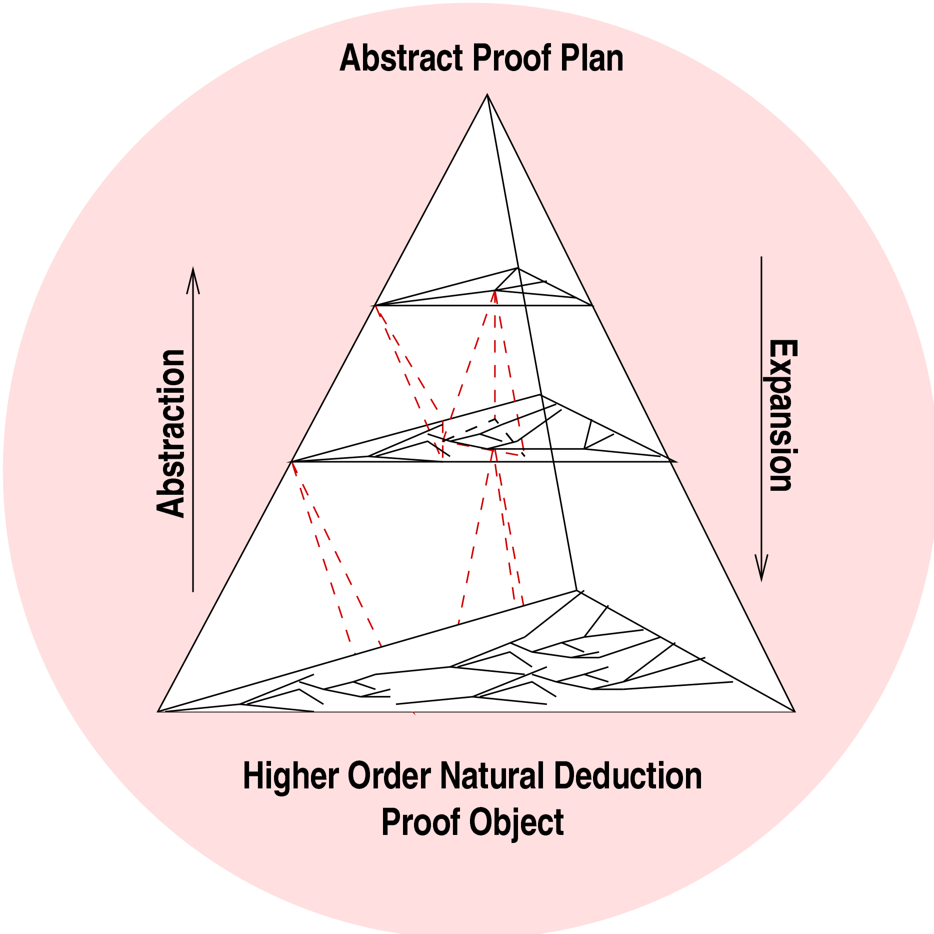
Beweisen mathematischer Aussagen:



abschließende Verifikation auf der elementaren Logikebene

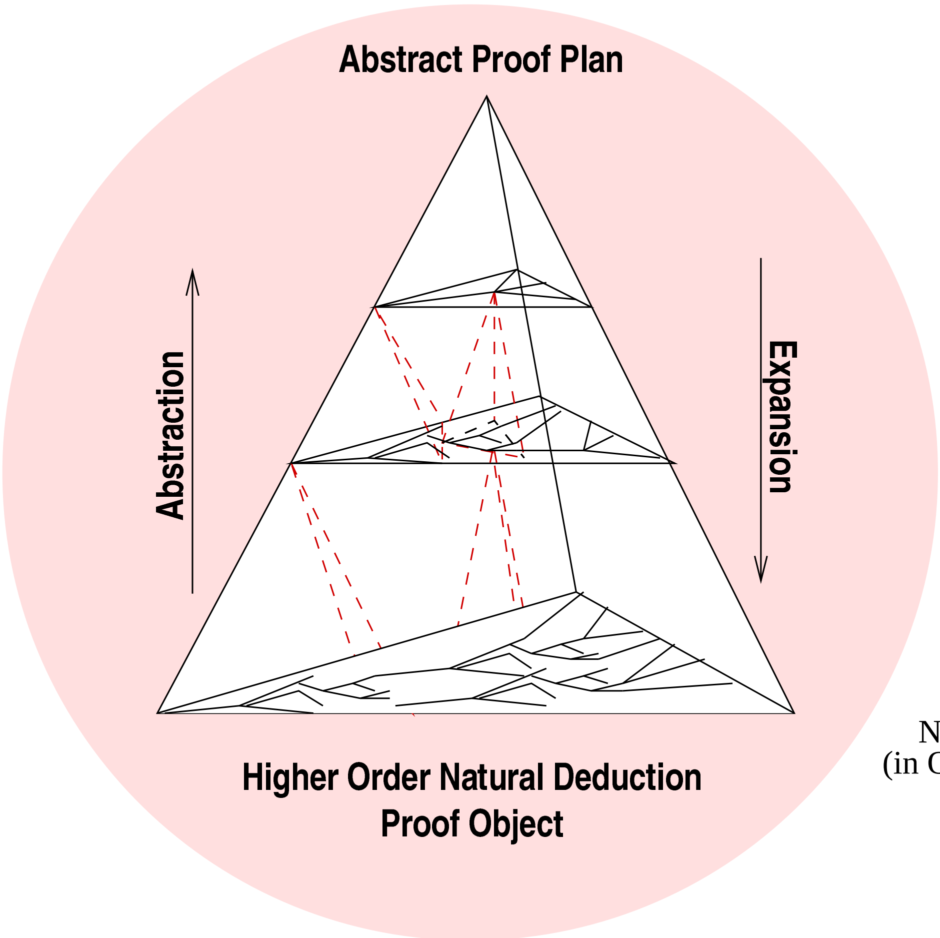
Beweisen mathematischer Aussagen:

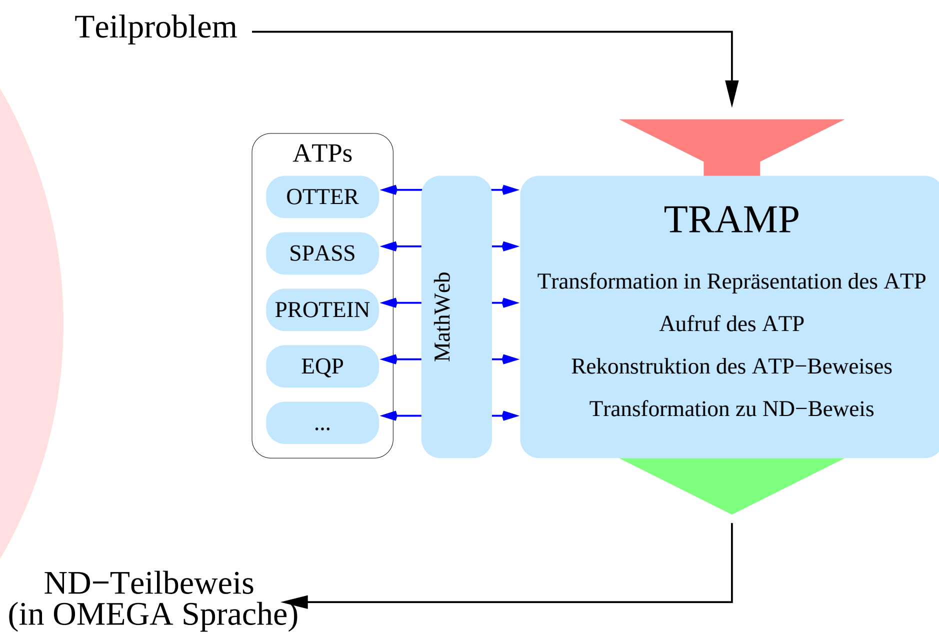
ΩMEGA Beweisobjekt



Beweisen mathematischer Aussagen:

Überbrückung des Kommunikationsproblems

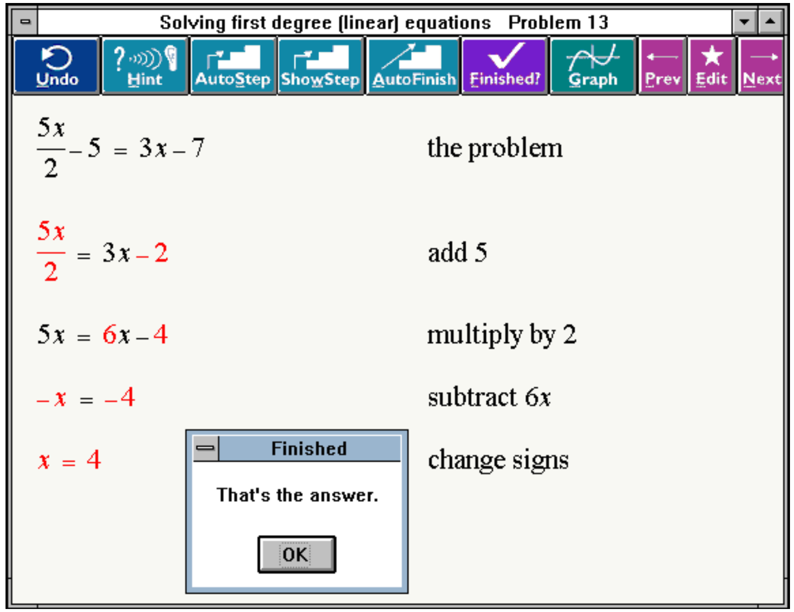




Computer Algebra Systeme

Erste Rechenmaschine: Abakus (ab ca. 500 v. Chr.)

Wilhelm Schickard's Rechenmaschine
(1592 - 1635)



MathPert System (Michael Beeson)

Heutige Systeme: Derive, MAPLE, MathCad, Mathematica, Reduce, . . .

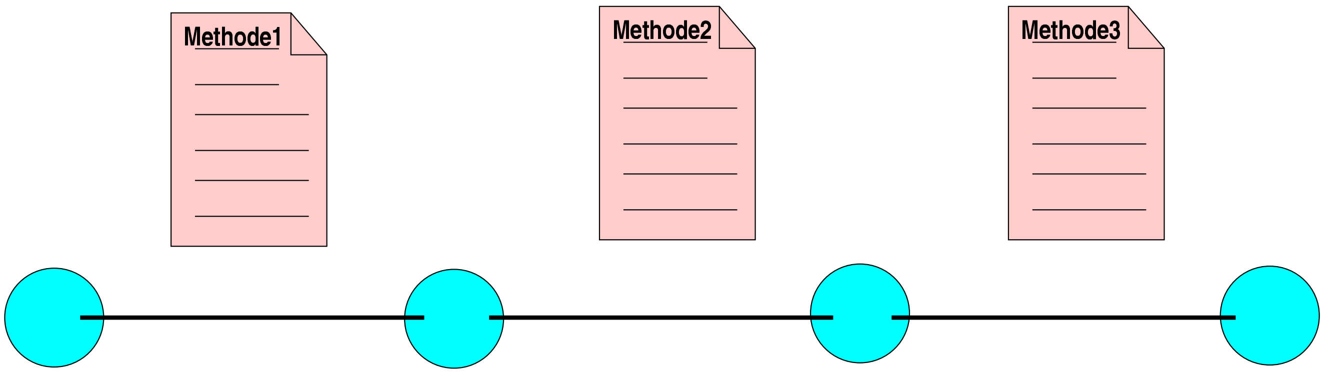
Computer Algebra Systeme

Komplementäre Schwächen und Stärken von Beweisern und Computer Algebra Systemen

- Beweiser: Schwächen bei der Symbolischen Berechnung
  - Berechnung als Beweissuche
  - logische Repräsentationen schlecht für Berechnung
- Computer Algebra Systeme: Schwächen beim Symbolisches Schließen; eingeschränkte Tauglichkeit als Beweiser
  - Algorithmen abstrahieren von Nebenbedingungen
  - Bsp.: $1 = \frac{x-2}{x-2}$ gilt nur falls $x \neq 2$

$\Rightarrow$ Integration von Beweisern und CAS erstrebenswert

Computer Algebra Systeme:

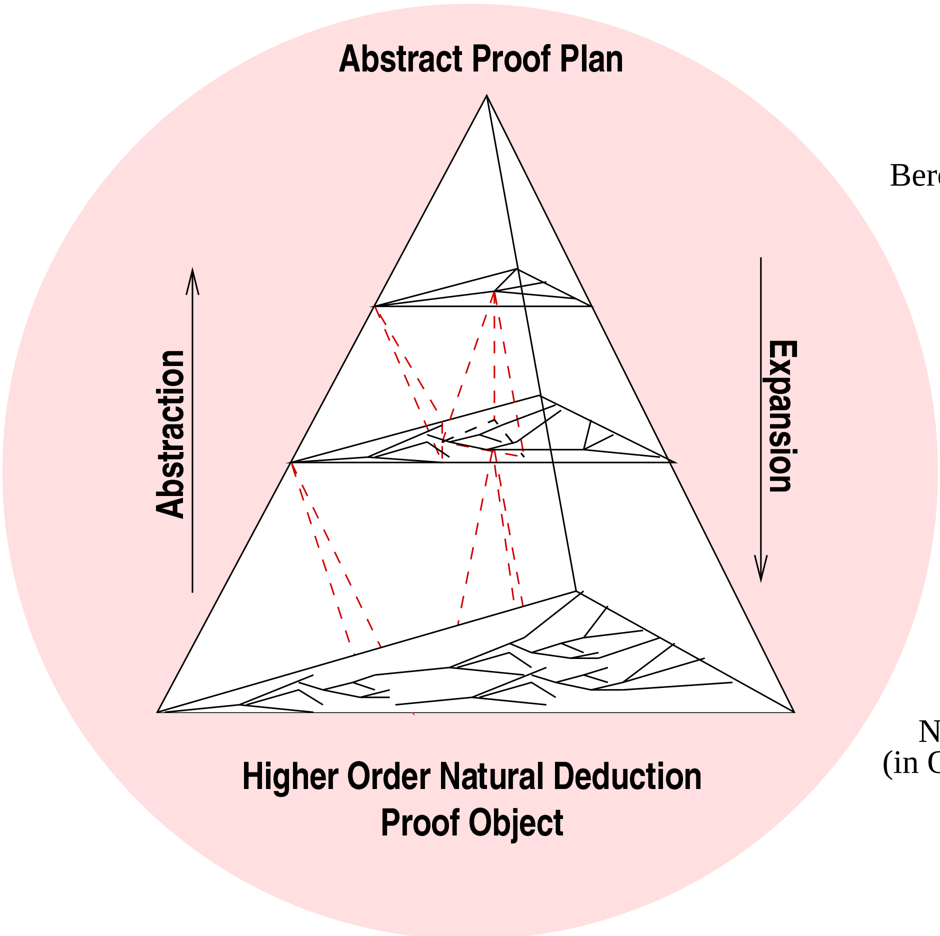


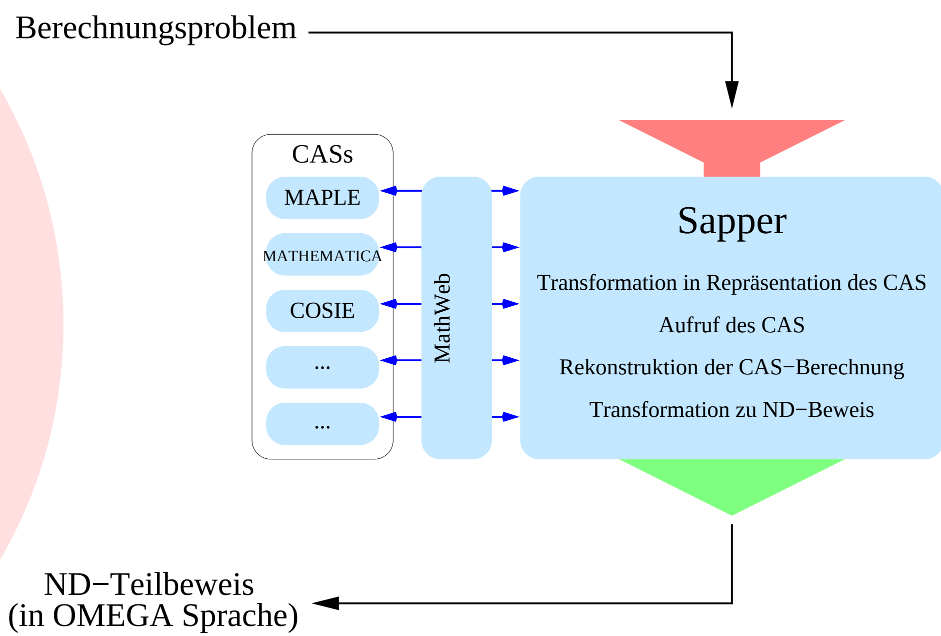
Computer Algebra Systeme:

- Integration in Beweismethoden
- und in Steuerungsheuristiken

Computer Algebra Systeme:

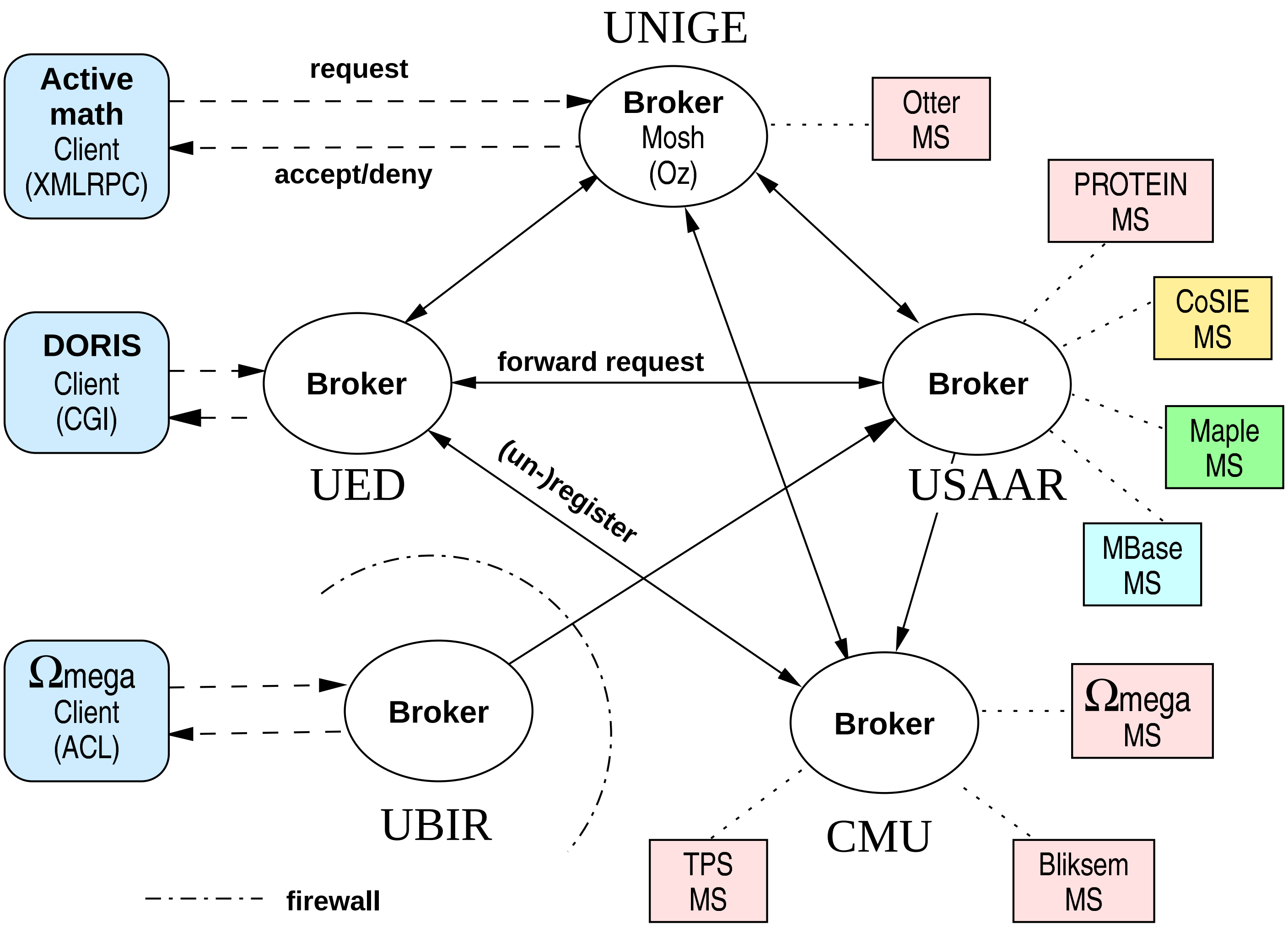
Überbrückung des Kommunikationsproblems





Mathematisches Semantisches Web:

MATHWEB-sb: Ein Netzwerk mathematischer Service Systeme im Internet



Mathematische Wissensbanken

- Ontologie mathematischer Theorien: Mengen, Relationen, Funktionen, . . . , Gruppen, . . . , natürliche Zahlen, . . . , relle Zahlen, . . .
- Theorie: Definitionen, Axiome, Lemmata, Theoreme, Beweise, . . .
- komplexe Vererbungshierarchie gemäß Ontologie

Eindrucksvolle mathematische Wissensbank: MIZAR (`www.mizar.org`)

Journal of Formalized Mathematics, Volume 15, 2003

Table of contents

. . .

4. On the Hausdorff Distance Between Compact Subsets *by Adam Grabowski*

5. Chains on a Grating in Euclidean Space *by Freek Wiedijk*

6. Bessel's Inequality *by Hiroshi Yamazaki, Yasunari Shidama, and . . .*

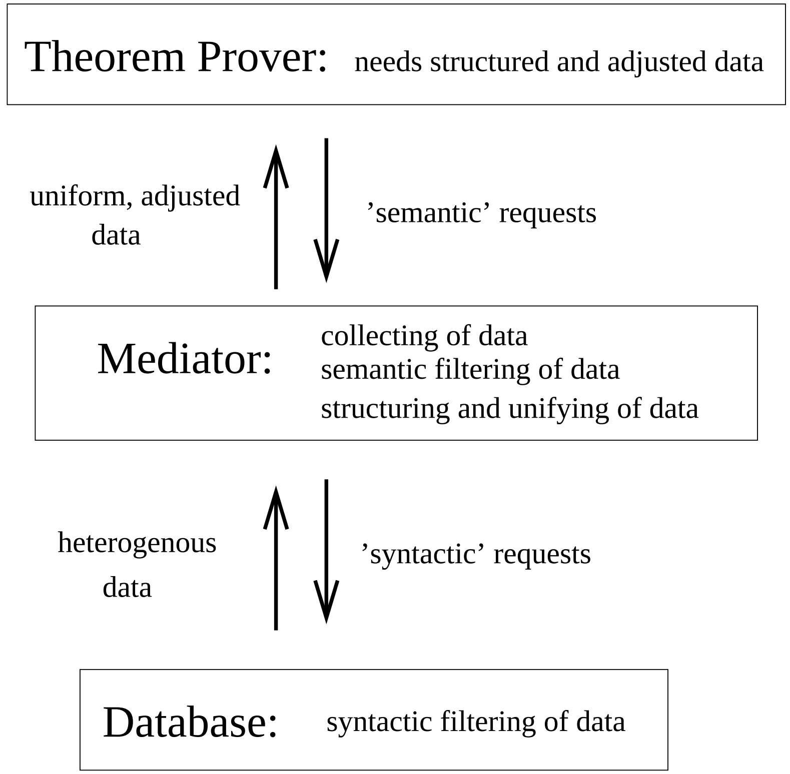
. . .

Mathematische Wissensbanken:

MBASE

- grosse mathematische Wissensbank
- aus ΩMEGA Projekt hervorgegangen
- Import von MIZAR Daten nach MBASE möglich

Kooperation mit
M. Kohlhase, CMU, USA



Interaktion mit dem Mathematiker:

Warum überhaupt Benutzerinteraktion?

- auf längere Zeit nicht eliminierbar
- wichtig für Ausbildung und Lehre

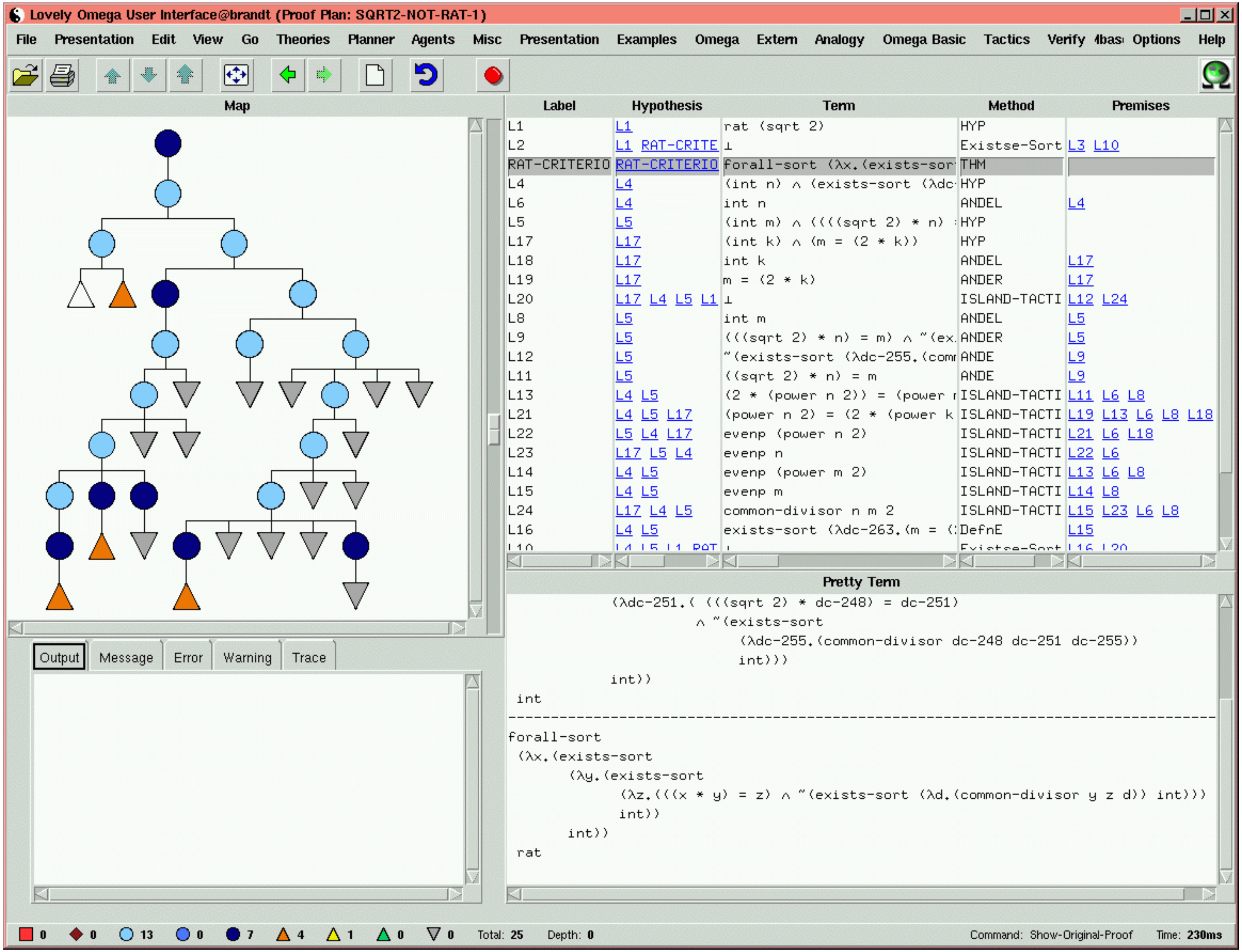
Idealerweise Kommunikation mathematischer Inhalte durch

- Textuelle Repräsentationen
- Graphische Repräsentationen: Beweisgraphen, Diagramme, . . .
- Maus, Hypertext
- Natürlichsprachige Kommunikation
- . . .

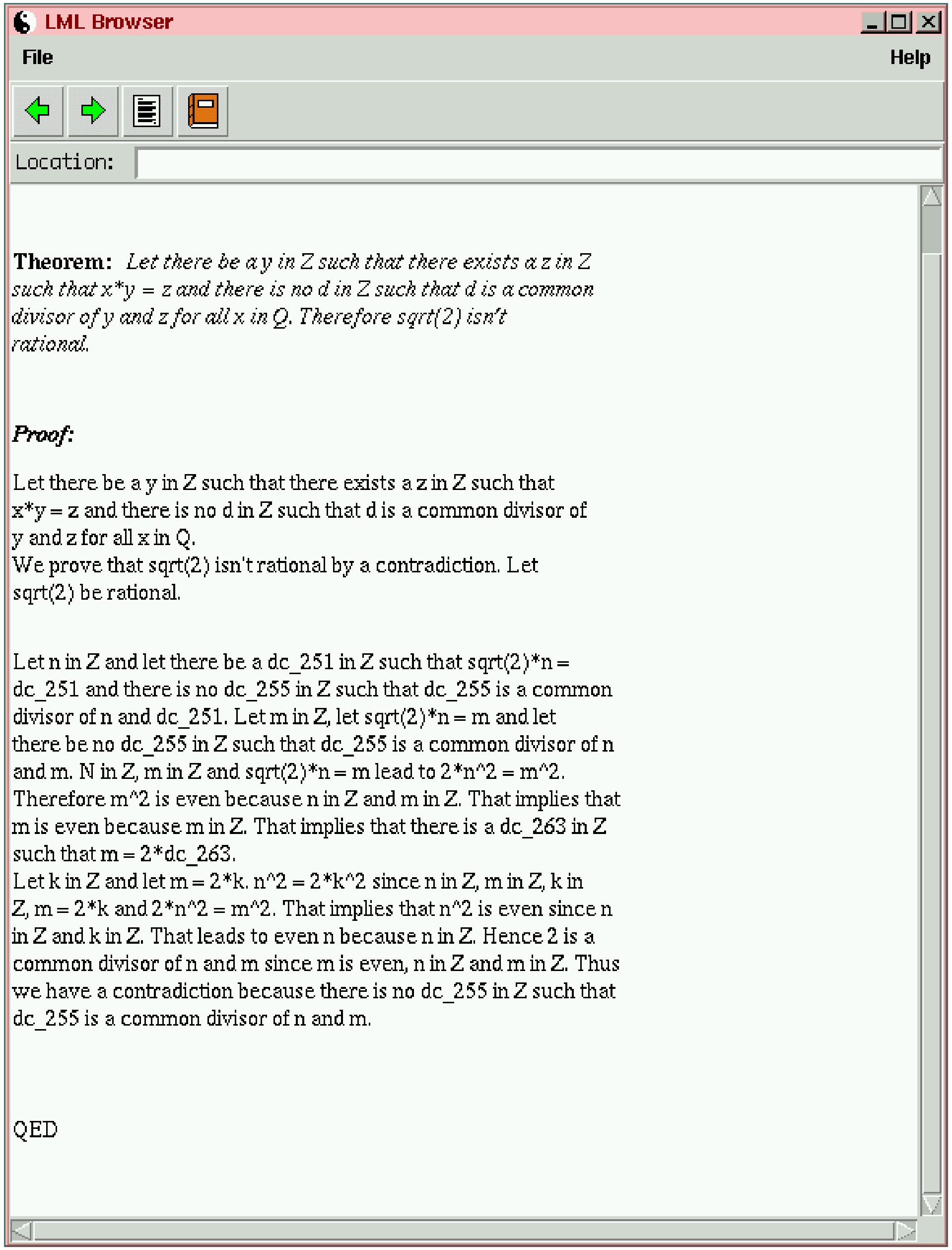
Wichtig auch

- Pro-aktives versus passives mathematisches Assistenzsystem

Interaktion mit dem Mathematiker:



Interaktion mit dem Mathematiker:



Theorem: *Let there be a y in Z such that there exists a z in Z such that x*y = z and there is no d in Z such that d is a common divisor of y and z for all x in Q. Therefore sqrt(2) isn't rational.*

Proof:

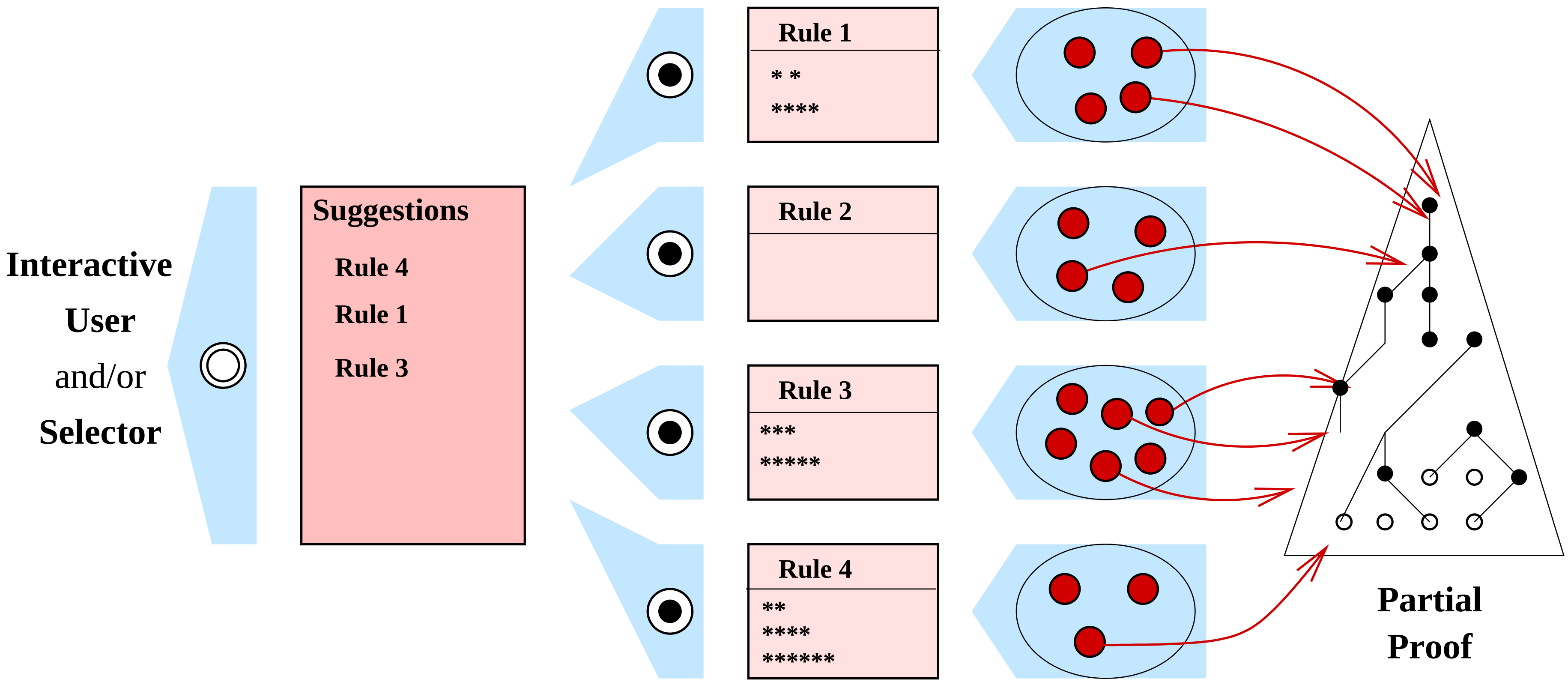
Let there be a y in Z such that there exists a z in Z such that x*y = z and there is no d in Z such that d is a common divisor of y and z for all x in Q.
We prove that sqrt(2) isn't rational by a contradiction. Let sqrt(2) be rational.

Let n in Z and let there be a dc_251 in Z such that sqrt(2)*n = dc_251 and there is no dc_255 in Z such that dc_255 is a common divisor of n and dc_251. Let m in Z, let sqrt(2)*n = m and let there be no dc_255 in Z such that dc_255 is a common divisor of n and m. N in Z, m in Z and sqrt(2)*n = m lead to 2*n^2 = m^2. Therefore m^2 is even because n in Z and m in Z. That implies that m is even because m in Z. That implies that there is a dc_263 in Z such that m = 2*dc_263.
Let k in Z and let m = 2*k. n^2 = 2*k^2 since n in Z, m in Z, k in Z, m = 2*k and 2*n^2 = m^2. That implies that n^2 is even since n in Z and k in Z. That leads to even n because n in Z. Hence 2 is a common divisor of n and m since m is even, n in Z and m in Z. Thus we have a contradiction because there is no dc_255 in Z such that dc_255 is a common divisor of n and m.

QED

Interaktion mit dem Mathematiker:

Dynamische Generierung von Vorschlägen durch Pro-aktive Agenten



Lehren mathematischer Inhalte:

Behauptung: Die Mathematikausbildung wird sich durch den Einsatz mathematischer Lernumgebungen und mathematischer Assistenzsysteme entscheidend verändern.

$\Rightarrow$ Siehe Vortrag von Erica Melis um 14:30

Beispiel der Verwendung von ΩMEGA:

- zur interaktiven Bearbeitung von Beispielaufgaben in der Mathematik-Lernumgebung ACTIVEMATH
- zur unterstütztenden Steuerung eines natürlichsprachigen tutoriellen Dialogs

Lehren mathematischer Inhalte:

Tutor-1:

Bitte zeigen Sie : $\overline{(A \cup B) \cap (C \cup D)} = (\overline{A} \cap \overline{B}) \cup (\overline{C} \cap \overline{D})$

Student-1:

(`correct`) nach deMorgan-Regel-2 ist $\overline{(A \cup B) \cap (C \cup D)} = \overline{A \cup B} \cup \overline{C \cup D}$

Tutor-2:

Das ist richtig.

Student-2:

(`correct`) $\overline{A \cup B}$ ist laut DeMorgan-1 $\overline{A} \cap \overline{B}$

Tutor-3:

Das stimmt auch.

Student-3:

(`correct`) und $\overline{C \cup D}$ ist ebenfalls laut DeMorgan-1 $\overline{C \cap D}$

Tutor-4: Auch das stimmt.

...

Exploration neuen mathematischen Wissens

Können Maschinen neue mathematischen Beweise finden?

Antwort bereits geliefert Ja

Frage nun:

Können Maschinen neue mathematische Strukturen entdecken?

Antwort (eingeschränktes) Ja

- Beispiel: HR System (Simon Colton, Imperial College, London) hat neue Integer-Sequenzen entdeckt für Encyclopedia of Integer Sequences

Ziel vorerst:

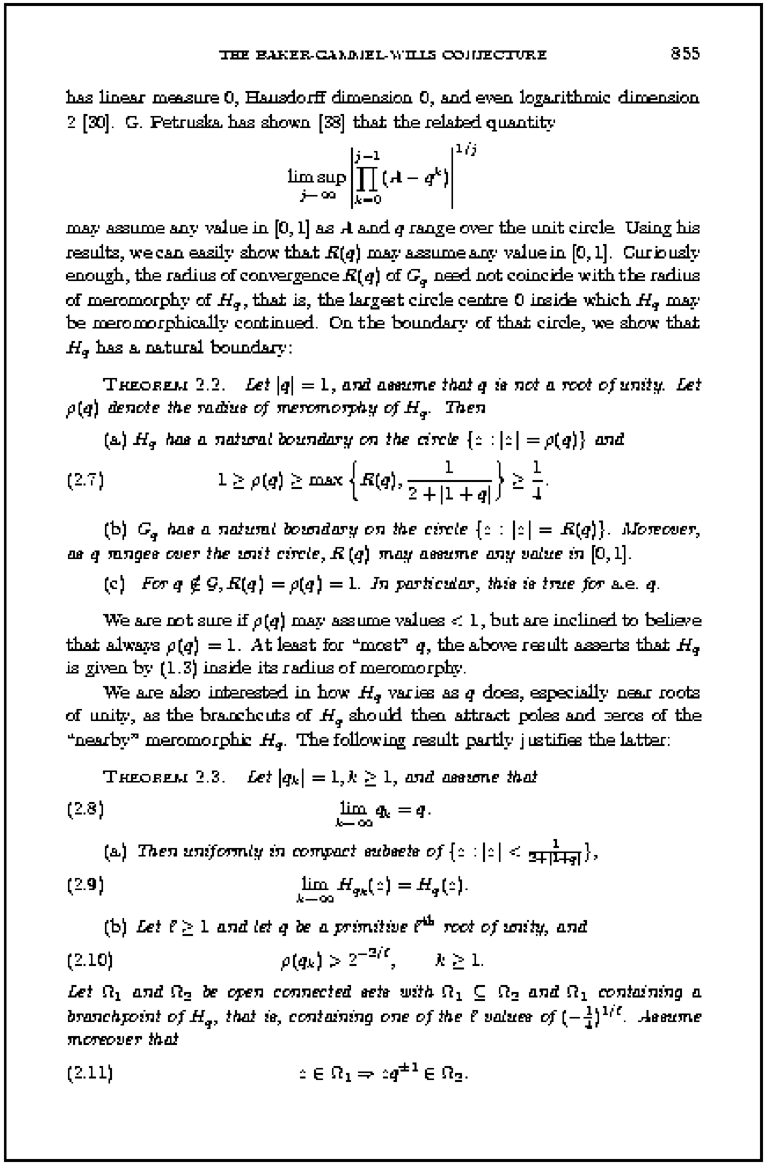
Unterstützung des Mathematikers bei Exploration

Verifikation mathematischer Publikationen

Ziel:

- Überprüfung der Korrektheit mathematischer Publikationen durch mathematische Assistenzsysteme

Erste Verlage/Journale denken bereits über machinenüberprüfbare Beweise nach . . .



THE BAKER-GAMMEL-WILLS CONJECTURE 855

has linear measure 0, Hausdorff dimension 0, and even logarithmic dimension 2 [30]. G. Petruska has shown [38] that the related quantity

$$\limsup_{j\to\infty}\left|\prod_{k=0}^{j-1}(A-q^k)\right|^{1/j}$$

may assume any value in $[0,1]$ as A and q range over the unit circle. Using his results, we can easily show that $R(q)$ may assume any value in $[0,1]$. Curiously enough, the radius of convergence $R(q)$ of G_q need not coincide with the radius of meromorphy of H_q, that is, the largest circle centre 0 inside which H_q may be meromorphically continued. On the boundary of that circle, we show that H_q has a natural boundary:

THEOREM 2.2. *Let $|q| = 1$, and assume that q is not a root of unity. Let $\rho(q)$ denote the radius of meromorphy of H_q. Then*

(a) *H_q has a natural boundary on the circle $\{z : |z| = \rho(q)\}$ and*

$$1 \geq \rho(q) \geq \max\left\{R(q), \frac{1}{2+|1+q|}\right\} \geq \frac{1}{4}. \tag{2.7}$$

(b) *G_q has a natural boundary on the circle $\{z : |z| = R(q)\}$. Moreover, as q ranges over the unit circle, $R(q)$ may assume any value in $[0,1]$.*

(c) *For $q \notin \mathcal{G}$, $R(q) = \rho(q) = 1$. In particular, this is true for a.e. q.*

We are not sure if $\rho(q)$ may assume values < 1, but are inclined to believe that always $\rho(q) = 1$. At least for "most" q, the above result asserts that H_q is given by (1.3) inside its radius of meromorphy.

We are also interested in how H_q varies as q does, especially near roots of unity, as the branchcuts of H_q should then attract poles and zeros of the "nearby" meromorphic H_q. The following result partly justifies the latter:

THEOREM 2.3. *Let $|q_k| = 1, k \geq 1$, and assume that*

$$\lim_{k\to\infty} q_k = q. \tag{2.8}$$

(a) *Then uniformly in compact subsets of $\{z : |z| < \frac{1}{2+|1+q|}\}$,*

$$\lim_{k\to\infty} H_{q_k}(z) = H_q(z). \tag{2.9}$$

(b) *Let $\ell \geq 1$ and let q be a primitive ℓ^{th} root of unity, and*

$$\rho(q_k) > 2^{-2/\ell}, \qquad k \geq 1. \tag{2.10}$$

Let Ω_1 and Ω_2 be open connected sets with $\Omega_1 \subseteq \Omega_2$ and Ω_1 containing a branchpoint of H_q, that is, containing one of the ℓ values of $(-\frac{1}{4})^{1/\ell}$. Assume moreover that

$$z \in \Omega_1 \Rightarrow zq^{\pm 1} \in \Omega_2. \tag{2.11}$$

Zusammenfassung

- spannendes, ambitioniertes und multi-disziplinares Forschungsfeld
- ΩMEGA-Team eines der weltweit größten Teams auf diesem Gebiet
- Ressourcenbündelung durch Kooperationen erforderlich
  - An UdS: DFKI, SFB 378, Computerlinguistik (Prof. Pinkal)
  - EU Netzwerke:
    CALCULEMUS (ΩMEGA-Team ist Coordinator), MKMNet
  - Carnegie Mellon University, USA
  - The University of Edinburgh, Scotland
  - The University of Birmingham, England
  - Cornell University, USA
  - . . . viele weitere Kooperations-Partner . . .

Demonstration: Direkt nach Vortrag

Theorem: $\sqrt{2}$ is irrational.

Proof: (by contradiction)
Assume $\sqrt{2}$ is rational, that is, there exist natural numbers m, n with no common divisor such that $\sqrt{2} = m/n$. Then $n\sqrt{2} = m$, and thus $2n^2 = m^2$. Hence m^2 is even and, since odd numbers square to odds, m is even; say $m = 2k$. Then $2n^2 = (2k)^2 = 4k^2$, that is, $n^2 = 2k^2$. Thus, n^2 is even too, and so is n. That means that both n and m are even, contradicting the fact that they do not have a common divisor.

Demonstration durch: Martin Pollet und Armin Fiedler

12:30 Uhr, Foyer

Demonstration: Direkt nach Vortrag

Proof:
Let 2 be a common divisor of x and y if x is even and y is even for all $y \in \mathbb{Z}$ for all $x \in \mathbb{Z}$. Let x be even if and only if x^2 is even for all $x \in \mathbb{Z}$. Let there be a $y \in \mathbb{Z}$ such that there exists a $z \in \mathbb{Z}$ such that $x \cdot y = z$ and there is no $d \in \mathbb{Z}$ such that d is a common divisor of y and z for all $x \in \mathbb{Q}$.
We prove that $\sqrt{2}$ isn't rational by a contradiction. Let $\sqrt{2}$ be rational.
Let $n \in \mathbb{Z}$ and let there be a $dc_{269} \in \mathbb{Z}$ such that $\sqrt{2} \cdot n = dc_{269}$ and there doesn't exist a $dc_{273} \in \mathbb{Z}$ such that dc_{273} is a common divisor of n and dc_{269}.
Let $m \in \mathbb{Z}$, let $\sqrt{2} \cdot n = m$ and let there be no $dc_{279} \in \mathbb{Z}$ such that dc_{279} is a common divisor of n and m.
We prove that $m^2 = 2 \cdot n^2$ in order to prove that there is a $dc_{287} \in \mathbb{Z}$ such that $m^2 = 2 \cdot dc_{287}$. $m^2 = 2 \cdot n^2$ because $\sqrt{2} \cdot n = m$.
Hence m^2 is even. Hence m is even since $m \in \mathbb{Z}$. Thus there exists a $dc_{343} \in \mathbb{Z}$ such that $m = 2 \cdot dc_{343}$.
Let $k \in \mathbb{Z}$ and let $m = 2 \cdot k$. $2 \in \mathbb{Z}$.
We prove that $n^2 = 2 \cdot k^2$ in order to prove that there is a $dc_{353} \in \mathbb{Z}$ such that $n^2 = 2 \cdot dc_{353}$. $n^2 = 2 \cdot k^2$ since $m^2 = 2 \cdot n^2$ and $m = 2 \cdot k$.
That implies that n^2 is even. That leads to even n because $n \in \mathbb{Z}$. That leads to a contradiction because $m \in \mathbb{Z}$, $n \in \mathbb{Z}$, there is no $dc_{279} \in \mathbb{Z}$ such that dc_{279} is a common divisor of n and m, m is even and $2 \in \mathbb{Z}$.

■