de kunst van het bewijzen

Freek Wiedijk & Herman Geuvers

Radboud Universiteit Nijmegen & Technische Universiteit Eindhoven

Vakantiecursus wiskunde 2008

Technische Universiteit Eindhoven

2008 08 23, 14 : 15

Centrum voor Wiskunde en Informatica, Amsterdam

2008 08 29, 18 : 30

lagere en hogere wiskunde

de twee soorten wiskunde

- **berekenen**

  bij berekenen gaat het om *de uitkomst*

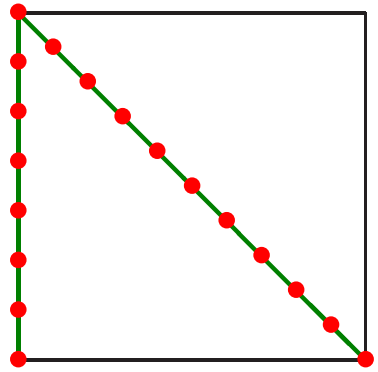
  wat?

- **bewijzen**

  bij bewijzen gaat het om *het begrip*

  waarom?

het diagonaal van een vierkant



- **berekenen:** wat is de verhouding tussen de zijde en de diagonaal?

  Pythagoras: $1 : \sqrt{2} = 1 : 1{,}4142135623\ldots$

- **bewijzen:** is dit als een verhouding van gehele getallen te schrijven?

  Pythagoras: niet mogelijk, *want* . . .

computerwiskunde

vier soorten

- **berekenen:** *getallen*

  numerieke wiskunde, visualisatie, experimentatie

- **berekenen:** *formules*

  computer algebra

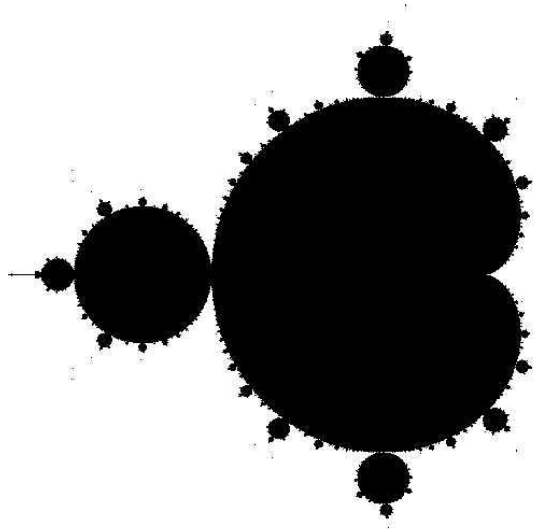
- **bewijzen:** *door de computer*

  automatische stellingenbewijzers

- **bewijzen:** *door de mens*, met behulp van de computer

  bewijsassistenten

numerieke wiskunde: de Mandelbrot-verzameling



$$z_0 = 0,\ z_1 = z_0^2 + c,\ z_2 = z_1^2 + c,\ \ldots$$

$$\{c \in \mathbb{C} \mid \text{de rij } z_i \text{ gaat niet naar } \infty\}$$

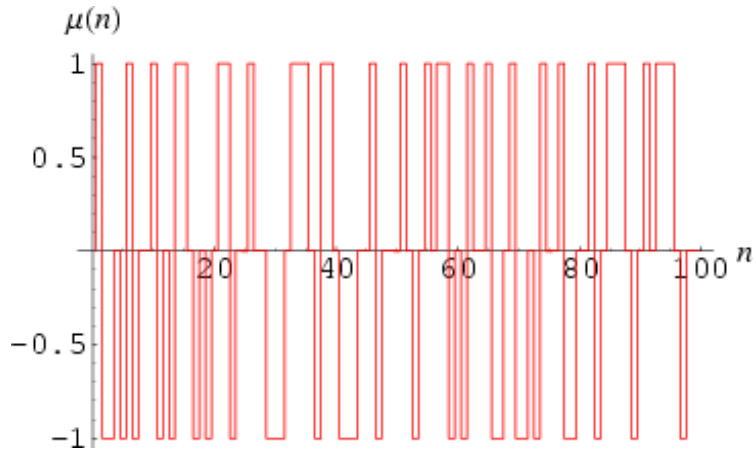
hoe belangrijk is het dat iedere pixel in dit plaatje correct is?

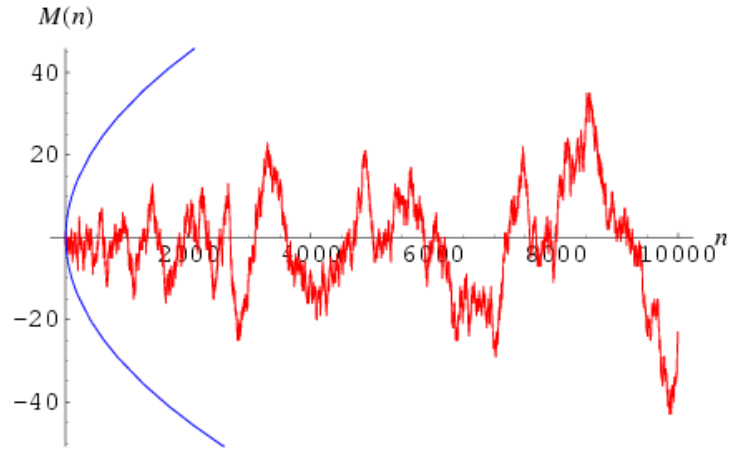
numerieke wiskunde: het Mertens-vermoeden

Möbius functie:

$$\mu(n) = \begin{cases} 0 & \text{als } n \text{ dubbele priemfactoren heeft} \\ 1 & \text{als } n \text{ een even aantal verschillende priemfactoren heeft} \\ -1 & \text{als } n \text{ een oneven aantal verschillende priemfactoren heeft} \end{cases}$$

Mertens, 1897: $\left|\sum_{k=1}^{n} \mu(n)\right| < \sqrt{n}$?





numerieke wiskunde: het Mertens-vermoeden (vervolg)

Odlyzko & te Riele, 1985: vermoeden van Mertens is niet waar !

50 uur computertijd

eerste n waar het mis gaat heeft tientallen cijfers

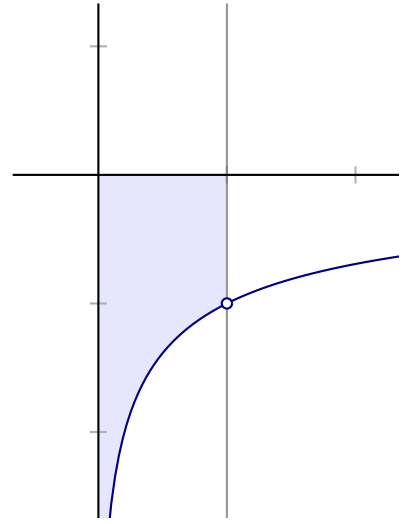
indirect bewijs !

2000 nulpunten van Riemann zeta functie op 100 decimalen nauwkeurig

14.13472514173469379045725198356247027078425711569924317568556746014996342980925676494901039317156101 27...
21.02203963877155499262847959389690277733434052490278175462952040358759858606889079971365851418015141 95...
25.01085758014568876321379099256282181865954967255799667249654200674509209844164427784023822455806244 07...
30.42487612585951321031189753058409132018156002371544018096214603699332938933327792029058429390208911 06...
32.93506158773918969066236896407490348881271560351703900928000344078481560863055100593884849613534872 45...
37.58617815882567125721776348070533282140559735083079321833300111362214908961853726473032910494582380 34...
40.91871901214749518739812691463325439572616596277727953616130366725328052872007128299600371988954687 55...
43.32707328091499951949612216540680578264566837183687144687889368552108832230505362645634937106319093 35...
48.00515088116715972794247274942751604168684400114442511777531251981409021641630828133033537230540099 77...
49.77383247767230218191678467856372405772317829967666210078195575043351161151573927873270750740093133 00...
52.97032147771446064414729660888099006382501788882122477990074814031756495030418805413758782709439929 88...
56.44624769706339480436775947670612755278226447171663184545096984395847528027450566690301131427485238 74...
59.34704400260235307965364867499221903109877280646666969812245175474680015269962981183810248707463354 84...
60.83177852460980984425990182452400380291009045121917825710134882480849366729492053843084167039434335 65...
65.11254404808160666087505425318370502934814929516672240596650108667534323266868538441677478443865947 14...
67.07981052949417371447882889652221677010714495174555887419666955169490121895619698353029397508583303 43...
69.54640171117397925292685752655473844301247420960251015732453999966338767227491041953334493317834035 63...
72.06715767448190758252210796982616839048090662145669708668330615148840737239960834836352533041217453 29...
75.70469069908393316832691676203034592281190353069740030164777530157419702770632360838403702183465279 80...
77.14484006887480537268266485630463701579603244923446104176523145315113916425371508940828869469973775 97...
79.33737502024936792276359287711622819061324674312003087843872049710154193267709097467745199461212410 90...

computer algebra: een integraal uitrekenen

Matlab, Mathematica, Maple, Magma



```
> Int(ln(x)/(1 - x), x = 0..1);
```

$$\int_0^1 \frac{\ln x}{1-x}\,dx$$

```
> value(%);
```

$$-\frac{\pi^2}{6}$$

```
> evalf(%);
```

$$-1.644934068$$

correcte uitkomst: $-1.64493406\,84822\,64\ldots$

TPTP

Thousands of Problems for Theorem Provers

7068 problemen

CASC

CADE ATP System Competition

Conference on Automated Deduction, Automated Theorem Prover

winnaar 2008: Vampire

169 van de 200 problemen opgelost

tweede 2008: E

164 van de 200 problemen opgelost

Otter $\approx$ Prover9

automatische stellingenbewijzers: het Robbins-vermoeden

is iedere **Robbins algebra** een Boolse algebra?

$$a \vee b = b \vee a$$

$$a \vee (b \vee c) = (a \vee b) \vee c$$

$$\neg(\neg(a \vee b) \vee \neg(a \vee \neg b)) = a$$

Bill McCune + EQP, 1996: inderdaad!

8 dagen computertijd

bewijs van 34 regels

automatische stellingenbewijzers kunnen tot nog toe:

- 'puzzels' waarin heel veel gevallen moeten worden onderzocht
- elementaire stapjes in een bewijs

bewijsassistenten

- **numerieke wiskunde** en **computer algebra:** geen bewijzen
- **automatische stellingenbewijzers:** geen interessante wiskunde
- **bewijsassistenten:** wél bewijzen & wél interessante wiskunde

  de prijs die moet worden betaald:
  gebruiker moet veel zelf doen

  bewijsassistenten = **interactieve stellingenbewijzers**
  samenspel van mens en computer

vier bewijsassistenten

100 leuke stellingen, 80 geformaliseerd

1.	The Irrationality of the Square Root of 2	≥ 17
2.	Fundamental Theorem of Algebra	4
3.	The Denumerability of the Rational Numbers	6
4.	Pythagorean Theorem	6
5.	Prime Number Theorem	2
6.	Gödel's Incompleteness Theorem	3
7.	Law of Quadratic Reciprocity	4
8.	The Impossibility of Trisecting the Angle and Doubling the Cube	1
9.	The Area of a Circle	1
10.	Euler's Generalization of Fermat's Little Theorem	4
11.	The Infinitude of Primes	6
12.	The Independence of the Parallel Postulate	0
13.	Polyhedron Formula	1

...

google: 100 theorems

de beste bewijsassistenten

vijf systemen serieus gebruikt voor wiskunde:

HOL	HOL Light	69
	ProofPower	42
	Isabelle	40
	Coq	39
	Mizar	45

HOL Light

in lange traditie:

LCF $\rightarrow$ HOL $\rightarrow$ HOL Light

Stanford, US $\rightarrow$ Cambridge, UK $\rightarrow$ Portland, US

John Harrison

bewijst floating point hardware correct bij Intel

formaliseert wiskunde in zijn vrije tijd

bijzonder elegant systeem

makkelijk uit te breiden

niet gebruikersvriendelijk

Isabelle

soort ‘opvolger’ van HOL

samenwerking tussen twee universiteiten:

Cambridge, UK

vooral: computerveiligheid

München, Duitsland

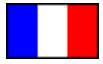
vooral: wiskunde

gebalanceerd systeem

mooie bewijstaal

krachtige automatisering

Coq



INRIA en Microsoft

Institut National de Recherche en Informatique et en Automatique

systeem met de indrukwekkendste formalisatie tot nu toe

systeem dat wij in Nijmegen gebruiken

geïntegreerde programmeertaal

$\approx$ Haskell

wiskundige expressief

intuïtionistisch

Luitzen Egbertus Jan Brouwer

pionier van de topologie

proefschrift, 1907

Over de grondslagen van de wiskunde

niet valide om over een reëel getal x te redeneren

'óf x is nul, óf x is ongelijk aan nul'

'uitgesloten derde'

tussenwaardestelling intuïtionistisch niet bewijsbaar

Mizar

Andrzej Trybulec

Białystok, Polen

ook: Nagano, Japan

wiskundigste van de bewijsassistenten

grootste bibliotheek met geformaliseerde wiskunde

2,1 miljoen regels code

gebruikersvriendelijk

soms moeilijk te begrijpen

geformaliseerde stellingen

de hoofdstelling van de algebra

Carl Friedrich Gauss, 1799

Mizar: Robert Milewski, 2000

HOL Light: John Harrison, 2000

Hellmut Kneser, 1940

Coq: Herman Geuvers e.a., 2000

$$x^2 = -1$$

$$x = i = \sqrt{-1}$$

welke polynomiale vergelijkingen kunnen we nu oplossen?

$$x^2 = i \ ?$$

de priemgetalstelling

Paul Erdös en Atle Selberg, 1949

Isabelle: Jeremy Avigad, 2004

Jacques Hadamard, Charles Jean de la Vallée-Poussin, 1896

HOL Light: John Harrison, 2008

aantal priemgetallen $\leq$ gegeven getal:

$$\begin{aligned} \pi(1000000000000) &= 37607912018 \\ \frac{1000000000000}{\ln(1000000000000)} &= 36191206825{,}27\ldots \\ \int_2^{1000000000000} \frac{dx}{\ln(x)} &= 37607950279{,}75\ldots \end{aligned}$$

wat is de verhouding in de limiet van het aantal getallen naar ∞?

de stelling over Jordan-krommen

Oswald Veblen, 1905

HOL Light: Tom Hales, 2005

Mizar: Artur Korniłowicz e.a., 2005

Jordan-krommen geen Jordan-krommen

in hoeveel stukken verdeelt een Jordan-kromme het platte vlak?

de eerste onvolledigheidsstelling

Kurt Gödel, 1931

Boyer-Moore prover: Natarajan Shankar, 1986

Coq: Russell O'Connor, 2003

HOL Light: John Harrison, 2005

de Gödel-zin:

'deze zin is niet bewijsbaar'

als deze zin waar is, zijn er onbewijsbare ware zinnen (onvolledigheid)
als deze zin onwaar is, zijn er bewijsbare onware zinnen (inconsistentie)

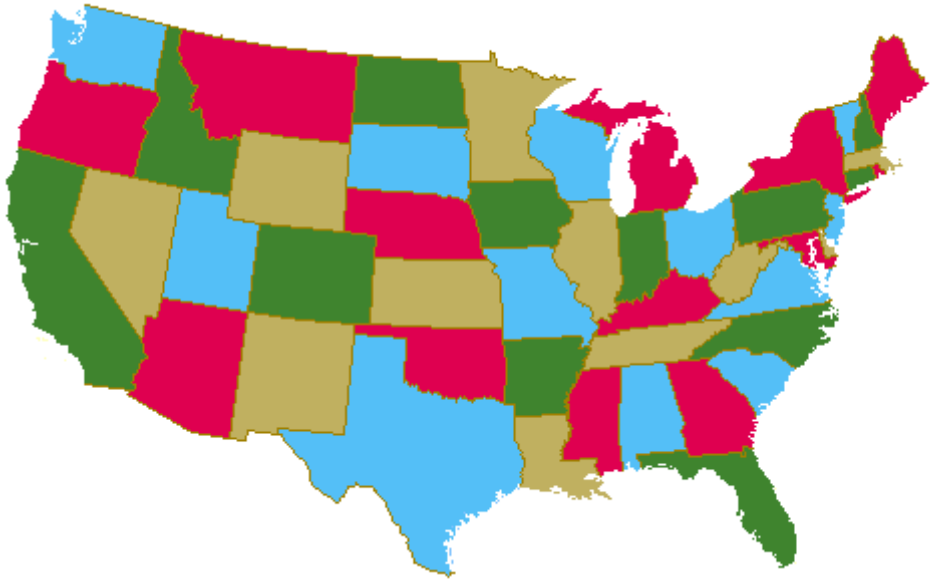
is er een bewijssysteem waarin waarheid en bewijsbaarheid samenvalt?

de vierkleurenstelling

Kenneth Appel en Wolfgang Haken, 1976

Neil Robertson e.a., 1996

Coq: Georges Gonthier, 2004



kan *iedere* kaart met maar vier verschillende kleuren worden gekleurd?

zelf formaliseren?

wetenschappelijk onderzoek

---

tijd voor het formaliseren van om één bladzijde uit een leerboek:
één volle werkweek $\approx$ 40 uur

lengte van een formalisatie van één bladzijde uit een leerboek:
een aantal computerschermen vol $\approx$ 200 regels

$\approx$ een kwartier per regel
moeilijk!

toch proberen Mizar te leren?

writing a Mizar article in nine easy steps
`http://www.cs.ru.nl/~freek/mizar/mizman.pdf`

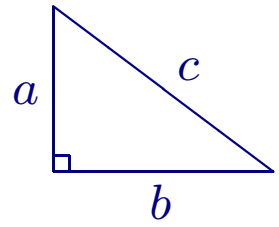
Een bolleboos riep laatst met zwier
gewapend met een vel A-vijf:
Er is geen allergrootst getal,
dat is wat ik bewijzen ga.
Stel, dat ik u nu zou bedriegen
en hier een potje stond te jokken,
dan ik zou zonder overdrijven
het grootste kunnen op gaan noemen.
Maar ben ik klaar, roept u gemeen:
'Vermeerder dat getal met twee!'
En zien we zeker en gewis
dat dit toch niet het grootste was.
En gaan we zo nog door een poos,
dan merkt u: dit is onbegrensd.
En daarmee heb ik q.e.d.
Ik ben hier diep gelukkig door.
'Zo gaan', zei hij voor hij bezwijmde,
'bewijzen uit het ongedichte'.

```
theorem
  not ex n st for m holds n >= m
proof

  assume not thesis;
  then consider n such that
A1: for m holds n >= m;

  set n' = n + 2;

  n' > n by XREAL_1:31;

  then not for m holds n >= m;

  hence contradiction by A1;

end;
```

Pythagoreïsche drietallen: een formule

een oplossing van

$$a^2 + b^2 = c^2$$



zonder gemeenschappelijke delers is altijd van de vorm:

$$a = m^2 - n^2 \qquad b = 2mn \qquad c = m^2 + n^2$$

voorbeelden

$$\begin{array}{llcrcrcr}
m = 2 & n = 1 & \rightarrow & 3^2 & + & 4^2 & = & 5^2 \\
m = 3 & n = 2 & \rightarrow & 5^2 & + & 12^2 & = & 13^2 \\
m = 4 & n = 1 & \rightarrow & 15^2 & + & 8^2 & = & 17^2 \\
m = 4 & n = 3 & \rightarrow & 7^2 & + & 24^2 & = & 25^2 \\
 & & \not\rightarrow & 9^2 & + & 12^2 & = & 15^2
\end{array}$$

Pythagoreïsche drietallen: 'formal proof sketch'

```
reserve a,b,c,m,n for Nat;

let a,b,c; assume a^2 + b^2 = c^2;
assume a,b are_relative_prime;
then a is odd or b is odd; assume a is odd;

ex m,n st a = m^2 - n^2 & b = 2*m*n & c = m^2 + n^2
proof
  b is even; c is odd;
X: (c + a)/2,(c - a)/2 are_relative_prime;
  ((c + a)/2)*((c - a)/2) = (c^2 - a^2)/4 .= (b/2)^2;
  then ((c + a)/2)*((c - a)/2) is square;
  then (c + a)/2 is square & (c - a)/2 is square by X;
  consider m,n such that (c + a)/2 = m^2 & (c - a)/2 = n^2;
  take m,n;
```

Pythagoreïsche drietallen: 'formal proof sketch' (vervolg)

```
  consider m,n such that m^2 = (c + a)/2 & n^2 = (c - a)/2;
  take m,n;
  thus a = (c + a)/2 - (c - a)/2 .= m^2 - n^2;
  b^2 = (c + a)*(c - a) .= 4*m^2*n^2 .= (2*m*n)^2;
  hence b = 2*m*n;
  thus c = (c + a)/2 + (c - a)/2 .= m^2 + n^2;
end;
```

syntactisch correct

semantisch correct

te kort door de bocht

stappen te groot

relaties tussen de stappen niet expliciet

Pythagoreïsche drietallen: fragment van de uitgewerkte formalisatie

```
 then
X: (c + a)/2,(c - a)/2 are_relative_prime by Lm3;
 ((c + a)/2)*((c - a)/2) = ((c + a)*(c - a))/(2*2)
   by REAL_1:35
  .= (c^2 - a^2)/4 by SQUARE_1:67
  .= (b^2)/(2*2) by H1,INT_1:3
  .= (b^2)/(2^2) by SQUARE_1:def 3
  .= (b/2)^2 by SQUARE_1:69;
 then ((c + a)/2)*((c - a)/2) is square by A1,Def1;
 then (c + a)/2 is square & (c - a)/2 is square by X,Lm4;
 then (ex m st m^2 = (c + a)/2) &
  (ex n st n^2 = (c - a)/2) by Def1;
 then consider m,n such that
A9: m^2 = (c + a)/2 & n^2 = (c - a)/2;
```

een zeer korte geschiedenis van de wiskunde

de drie revoluties

Euclid

bewijzen

Cauchy

rigoreus

de Bruijn

formeel