bewijzen in de computer

Freek Wiedijk
Katholieke Universiteit Nijmegen

Nationale Wiskunde Dagen
Noordwijkerhout
2004 02 06, 16:15

principia mathematica

wiskunde in volledig detail in een formele taal

Leibniz, eind 17de eeuw

uitvinder van differentiëren en integreren

bouwer van één van de eerste rekenmachines: $\times$, $\div$, $\sqrt{}$

lingua characteristica universalis (universele taal)
calculus ratiocinator

Boole, The Calculus of Logic, 1848

Frege, Begriffsschrift, 1879

Peano, tijdschrift Rivista di Matematica, 1891-1906

Russell & Whitehead, Principia Mathematica

drie delen: 1910, 1912, 1913

automath

computer $\longrightarrow$ bewijzen in volledig detail uitwerken wordt praktisch

eerste ter wereld: **de Bruijn,** Eindhoven

‘de automath’ = computer om wiskunde mee te controleren

automath = taal om wiskunde in op te schrijven

1968: eerste ideeën

jaren zeventig: groot onderzoeksproject

Jutting, 1977: vertaling van een heel wiskundeboek

dik pak computertekst

controle op correctheid: paar uur computertijd

(tegenwoordig: een halve seconde)

wiskunde in de computer

berekenen en bewijzen

---

de twee activiteiten van de wiskundige

- **berekenen**

  het gaat om het antwoord: wat?

- **bewijzen**

  het gaat om het begrip: waarom?

is bewijzen nog wel van deze tijd?

middelbaar onderwijs: bewijzen wordt nauwelijks meer onderwezen

- er is geen grootste priemgetal
- decimalen van
  $\sqrt{2} = 1.41421356237309504880168872420969807856967187537694807317\ldots$
  herhalen niet
- er zijn meer reële getallen dan natuurlijke getallen
- er zijn evenveel punten in een lijn als punten in het vlak

universitair onderzoek: computer-experimenten steeds belangrijker

. . . nemen de plaats van bewijzen over?

bewijzen spelen tegenwoordig een rol in de **informatica**
kritische hardware/software mag geen 'bugs' hebben

wiskunde in de computer

- **numeriek**

  getallen: computer $\longrightarrow$ mens

- **computer algebra**

  formules: computer $\longrightarrow$ mens

- **stellingenbewijzers**

  bewijzen: computer $\longrightarrow$ mens

- **bewijscheckers**

  bewijzen: mens $\longrightarrow$ computer

  (computer kijkt alleen of het klopt)

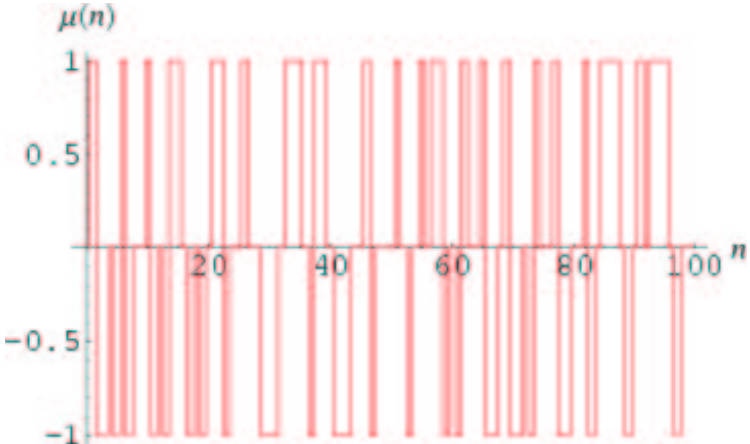
numeriek

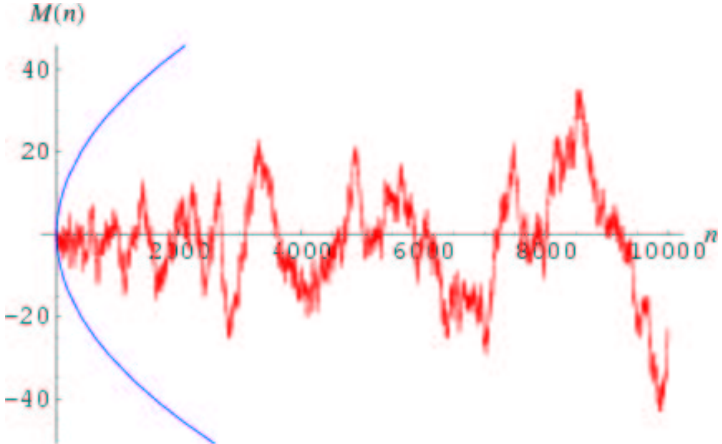
het vermoeden van Mertens

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





het vermoeden van Mertens (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...

computer algebra

symbolische integratie

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

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

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

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

het algoritme van Risch

integraal uitdrukken in 'elementaire functies' $\sqrt{}$, $\ln$, e^x, $\sin$, $\arcsin$, ...

– zegt of het kan

– geeft het antwoord als het kan

mathXpert

populaire computer algebra systemen

- maple
- mathematica

computer algebra voor het onderwijs

- **mathXpert**
  Beeson, 1997

  niet heel krachtig
  wel heel duidelijk

  – formules zien er uit als in de leerboekjes
  – niet alleen de antwoorden maar ook de berekeningen

automatische stellingenbewijzers

het vermoeden van Robbins

---

computers

. . . kunnen binnenkort beter schaken dan mensen

. . . kunnen binnenkort beter bewijzen dan mensen?

Robbins, 1933: is iedere Robbins algebra een Boole algebra?

eqp, 1996: **ja !**

acht dagen computertijd

één van de weinige bewijzen het eerst gevonden door een computer

niet erg conceptueel: probeert gewoon heel veel mogelijkheden

interessant onderzoek, maar momenteel niet relevant voor de wiskunde

bewijscheckers

de pentium bug

pentium: intel processor

midden 1994: processor maakt rekenfouten
FDIV bug, niet genoeg precisie

$\frac{5505001}{294911} = 18.66665197\ldots$ (wiskunde)

$\frac{5505001}{294911} = 18.66600093\ldots$ (pentium)

pentium processor met bug wordt gratis vervangen
schatting schade: 475 miljoen dollar

Harrison: werkt voor intel
bewijst full time met de **HOL** bewijschecker micro programma's correct

de metro en de B methode

1998: metro zonder bestuurder

météor = métro est-ouest rapide

lijn 14 van de metro in Parijs

madeleine $\leftrightarrow$ bibliothèque nationale

systeem voor de besturing

sacem = système d'aide à la conduite, à l'exploitation et à la maintenance

geprogrammeerd in de programmeertaal ada

veiligheid van sacem bewezen met de **B methode**

commerciële bewijschecker

het vermoeden van Kepler

Kepler in strena sue de nive sexangula, 1661:

is de **sinaasappelstapeling** de efficientste manier om bollen op te stapelen?

Hales, 1998: **ja !**

bewijs: bevat groot en onbegrijpelijk programma
bekijkt heel veel gevallen
3 gigabytes programma's/data + paar maanden computertijd

referees zeggen 99% zeker te zijn dat het klopt

flyspeck project
'formal proof of Kepler'

mizar

bewijschecker uit Polen

Trybulec

- 1973: eerste ideeën
- vandaag: grootste wiskunde bewijscheck project ter wereld

anderhalf miljoen regels gecodeerde wiskunde
achthonderd artikelen

gebaseerd op de axiomatische verzamelingenleer
wiskundige & leesbare bewijstaal

hyperproof

bewijschecken in de informatica

bewijschecken in de wiskunde

bewijschecken in het onderwijs

  eerste orde predicatenlogica

Barwise & Etchemendy, 1994: **hyperproof**

eenvoudige redeneringen over een eenvoudige blokkenwereld

draait momenteel alleen op apple macintosh computers

redeneren in een diagram

voorbeeld van een checkbaar bewijs

Pythagoreïsche tripels

---

een oplossing van

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

met a, b en c geen gemeenschappelijke delers, is altijd van de vorm

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

voorbeelden

$$\begin{array}{llcrcrcr}
m = 2 & n = 1 & \rightarrow & 3^2 & + & 4^2 & = & 5^2 \\
m = 3 & n = 2 & \rightarrow & 5^2 & + & 12^2 & = & 13^2 \\
m = 4 & n = 1 & \rightarrow & 15^2 & + & 8^2 & = & 17^2 \\
m = 4 & n = 3 & \rightarrow & 7^2 & + & 24^2 & = & 25^2
\end{array}$$

mizar bewijs

```
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 m^2 = (c + a)/2 & n^2 = (c - a)/2;
 take m,n;
```

mizar bewijs (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;
```

dit is bijna het echte mizar bewijs, maar te kort door de bocht

- stappen zijn te groot
- verwijzingen tussen stappen ontbreken

het volledige bewijs (fragment)

```
 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;
```

waarom proofchecken?

wordt wiskunde nu eindelijk echt **wis**kunde?

- bugs in de checker, kapotte hardware
  (in de praktijk geen probleem)

de Bruijn kriterium: eenvoudige checker

- definiëren je definities wel wat je denkt dat ze definiëren?

QED

1994, anonieme groep wiskundigen & informatici:

laten we alle wiskunde in de computer stoppen !

erg interessant **QED manifesto**
antwoorden op twaalf mogelijke tegenwerpingen

tot nog toe nog een utopie

het leukste computerspel ter wereld

wiskunde: spannend

programmeren: spannend

bewijzen in de computer: het beste van twee werelden

je hoeft niet alleen op je begrip te vertrouwen $\longrightarrow$ de computer helpt !
je weet zeker dat wat je doet helemaal goed is $\longrightarrow$ geen 'bugs' !

een stelling die nog niet helemaal bewezen is

$=$

een 'level' van een computerspel dat nog niet is uitgespeeld