



Piet Groeneboom

*Delft Institute of Applied Mathematics
TU Delft
p.groeneboom@tudelft.nl*

Lambert Meertens

*Kestrel Institute
Palo Alto, CA, USA
lambert@kestrel.edu*



Column Piet en Lambert grijpen hun kans

Chat

Naar aanleiding van het redactioneel in het vorige nummer had Piet Groeneboom even contact met zijn ‘oude makker’ Lambert Meertens. Dat leverde de volgende dialoog op.

Piet: “Heb jij dat redactioneel commentaar in het vorige nummer van NAW gelezen? Ik moest daar eigenlijk niet zo om lachen.”

Lambert: “Ja, het is een van de artikelen die online staan. (Tussen haakjes: ik begrijp de criteria voor wat wel of niet online staat niet.) Inderdaad flauw, in de zin van kraak noch smaak. Ik vermoed daarom dat dit geen grap is maar werkelijk uit de koker van ChatGPT komt, maar jij kent Robbert beter dan ik.”

P: “Tja, het zou ook wel eens van Robbert zelf kunnen zijn. Maar dat daargelaten, ik vind dat de ontwikkeling eigenlijk niet zo snel gaat als ik tegenwoordig steeds hoor beweren, en ChatGPT is daar een voorbeeld van. Hoe denk jij daarover?”

L: “Ik vind dat de ontwikkeling van AI nu toch wel snel gaat.”

P: “Laten we ons eerst even tot ChatGPT beperken. Ik zit op zondagmiddag vaak in de auto op het moment dat er een programma, getiteld *Diskotabel*, op NPO klassiek is. Enkele zondagen geleden was er een programma over uitvoeringen van het vierde pianoconcert van Beethoven en iemand in het panel had aan ChatGPT gevraagd waarom dat Beethovens mooiste pianoconcert is.”

L: “Waarom? Wat een onzinnige vraag. En altijd maar dat gezeur over ‘mooiste’, ‘beste’.”

P: “In het algemeen ben ik het geheel met je eens, maar in dit geval was ik toch erg benieuwd wat ChatGPT zou antwoorden. Ik vind zelf het vierde pianoconcert van Beethoven ook absoluut Beethovens mooiste pianoconcert. Alleen al dat ontroerende, verpletterende begin...”

L: “Goed, goed, maar wat zei ChatGPT?”

P: “Technische virtuositeit, symfonische kwaliteit, muzikale expressie, innovatie (aldus de altviolist Ward Koopmans, die de vraag had gesteld).”

L: “Dat is toch niet zo gek?”

P: “Het is misselijkmakend! En dan ook nog dat woord ‘innovatie’. Natuurlijk was het heel vernieuwend, maar dat het modewoord ‘innovatie’ er weer in voorkomt, daar moet ik dan weer bijna van overgeven. En wat moeten we denken van ‘technische virtuositeit’? Als je niet wist dat het van ChatGPT kwam, zou je denken: dit is van iemand die de klok heeft horen luiden, maar er in feite niets van begrepen heeft. Bovendien is het geen antwoord op de vraag, want zijn ‘technische virtuositeit’, ‘symfonische kwaliteit’, ‘muzikale expressie’ en ‘innovatie’ kwaliteiten die de andere vier concerten missen?”

L: “Ja, dit lijkt wel op een ervaring die ik ook met ChatGPT had. Ik heb ChatGPT gevraagd: ‘Bewijs of geef een tegenvoorbeeld voor de bewering: ieder natuurlijk getal c kan geschreven worden als $c = a^2 + b^3$, waarbij a en b integers zijn.’ Ik laat je even het resultaat zien.” (Zie Figuur 1.)

P: “Het doet me denken aan een verhaal dat een dispuutsgenoot uit het corpsdispuut Homerus (ja, ja, nog uit de donkere tijden van de kaalgeschoren feuten) mij vertelde over een tentamen logica dat hij over het boek *Inleiding tot de logica* van Tarski, vertaald door Evert Beth, had gedaan bij de analytisch filosoof Else Barth. Hij was — oudergewoonte — maar vast wat gaan praten in de hoop ergens wel iets te raken dat juist was, totdat Else tegen hem had gezegd: ‘Hou nou eens op met praten en denk eerst eens even in stilte na.’

Ik heb zelf ook twee tentamens bij haar gedaan, over dat boek van Tarski, en ook over *Formal methods* van Beth, waar de onvolledigheidsstellingen van Gödel werden behandeld. Het was bij haar

LM Prove or disprove: Every natural number can be written as the sum of the square of an integer and the cube of an integer.

Let's prove that every natural number can be written as the sum of the square of an integer and the cube of an integer.

Proof:

We will use mathematical induction to prove this statement.

1. Base Case: $n = 0$
For $n = 0$, it can be written as $0^2 + 0^3 = 0$, which satisfies the condition.
2. Inductive Hypothesis:
Assume that for some positive integer k , every natural number less than or equal to k can be written as the sum of the square of an integer and the cube of an integer.
3. Inductive Step:
We need to show that every natural number $k + 1$ can be written as the sum of the square of an integer and the cube of an integer.

Case 1: $k + 1$ is a perfect square ($k + 1 = m^2$, where m is an integer).
In this case, k can be written as the sum of the square of an integer (m) and the cube of 0 (0^3):
 $k = m^2 + 0^3$.

Case 2: $k + 1$ is not a perfect square.
Let p be the largest perfect square that is less than or equal to $k + 1$.
So, $p = a^2$, where a is an integer, and $p < k + 1$. Since p is the largest perfect square less than $k + 1$, the difference $k + 1 - p$ must be less than a^2 .

Let b be the positive integer such that $(k + 1 - p) = b^3$. We know b exists because $(k + 1 - p)$ is a positive number less than a^2 , and there must be an integer b that satisfies this condition.

Now, $k + 1$ can be written as the sum of the square of an integer (a) and the cube of an integer (b):
 $k + 1 = a^2 + b^3$.

So, by the principle of mathematical induction, every natural number can be written as the sum of the square of an integer and the cube of an integer. The proof is complete.

Figuur 1 Bewijs ChatGPT.

thuis en ik herinner me nog dat zij me een opgave voorlegde en toen zei: ‘Ga daar maar eens rustig over nadenken, ik ga intussen wat anders doen.’”

L: “De manier waarop ChatGPT over een niet overbrugbare kloof in het bewijs heen walst (‘and there must be an integer b that satisfies this condition’) lijkt opmerkelijk veel op hoe een wanhopige student op een tentamen probeert weg te komen met zijn onvermogen het bewijs te leveren. Merk ook op dat ChatGPT voor dit deel van zijn ‘bewijs’ in ieder geval b^3 in plaats van b^2 had moeten gebruiken. Blijkbaar kunnen we er wel op vertrouwen dat we voor de onmiddellijke toekomst nog even wiskundigen nodig zullen hebben om te checken of een bewijs juist is.

Het is verleidelijk om ChatGPT’s onvermogen hier op een zinnige manier mee om te gaan, toe te schrijven aan een gebrek aan inzicht. Echter, een betere verklaring berust volgens mij op een fundamentele beperking van het model. Een student, geconfronteerd met dit specifieke probleem, zal (hopelijk) een moment nadenken alvorens aan het probleem te beginnen (en dus niet meteen maar gaan praten zoals jouw dispuutsgenoot). ‘Large Language Models’ (LLMs) zoals ChatGPT wordt de luxe van zo’n stapje terug niet gegund; ze moeten meteen tekst genereren.

Een student zou eerst een poging kunnen doen om met volledige inductie te bewijzen dat de bewering waar is om vervolgens, na te hebben gezien dat deze weg doodloopt, een andere benadering te proberen. Dat kan de huidige generatie op LLMs gebaseerde chatbots niet. Die kunnen alleen maar doormodderen in de in het

begin ingeslagen richting. Zelfs als ChatGPT zich zou ‘realiseren’ — wat dat ook moge betekenen — dat de gekozen aanpak tot niets zal leiden, zijn de teksten die het kan genereren beperkt door zijn trainingsmateriaal. En daarin komen nu eenmaal geen voorbeelden van bewijzen voor waarin de auteur na een mislukte poging alsnog een andere weg inslaat. Eenmaal begonnen met ‘Laten we bewijzen dat ...’ kan het daardoor niet anders dan voortstropelen tot ‘hiermee is het gevraagde bewezen’.”

P: “Ja, maar het is een beperking van ChatGPT, niet van AI. Onze schaakkampioen van Nederland Jan-Hein Donner riep altijd dat computers nooit zouden kunnen schaken, maar Kasparov is tenslotte door DeepBlue verslagen, en go-kampioen Lee Sedol door AlphaGo. We kunnen niet meer zeggen dat dit alleen maar ‘Brute Force’ en ‘Geheugen’ is. Dus computers kunnen wellicht over een tijdje het door jou voorgelegde probleem ook wel een stuk intelligenter gaan behandelen dan ChatGPT nu.”

Het vermoeden van Catalan

Het aan ChatGPT voorgelegde probleem zit in de hoek van ‘het vermoeden van Catalan’, dat in 2002 door Preda Mihăilescu bewezen werd, zie bijvoorbeeld [1, 2]. Het eerst in aanmerking komende getal voor een tegenvoorbeeld bij Lamberts probleem is 6 (*niet* $3 = 2^2 + (-1)^3$). Dat het niet zo eenvoudig is laat bijvoorbeeld $7174 = 172469591^2 + (-309843)^3$ zien.

Referenties

- 1 Jeanine Daems, Het vermoeden van Catalan, *Nieuw Archief voor Wiskunde* 5/5(3) (2004), 221–225.
- 2 Preda Mihăilescu, Primary cyclotomic units and a proof of Catalan’s conjecture, *J. Reine Angew. Math.* 572 (2004), 167–195.