JZ Joost Zwagerman Essayprijs

Het gemiddelde woord

Daan Mulder

Hoe begint het gemiddelde prijswinnende essay? Op de website van de Joost Zwagerman Essayprijs staan de vorige acht winnaars onder elkaar opgesomd. Je hoeft alleen op hun namen te klikken en de essays openen zich, elk in nieuw tabblad. Ik kopieer de 8 eerste zinnen naar een Word-document. Ze tellen samen op tot 107 woorden, afgerond gemiddeld 13 per essay. Relatief kort: de gemiddelde Nederlandse zin, volgens *Onze Taal*, telt 15 à 20 woorden. Ik ben met mijn 6 woorden tellende eerste zin dus goed begonnen. Maar hoe nu verder? Hoe reken ik het gemiddelde uit van deze zinnen?

Ik zou me door de Russische wiskundige Andrej Markov kunnen laten inspireren. Hij analyseerde in 1913 de eerste twintigduizend van *Jevgeni Onegin*, Poesjkins roman in verzen, om de kans uit te rekenen dat een gemiddelde Russische letter een klinker is: 43 procent, aldus Markov. Vooral wilde hij bewijzen dat de kans op een klinker beïnvloed wordt door welke letter eraan voorafgaat: de kans op een klinker neemt na een medeklinker toe, en juist af na een klinker. Taal is geen willekeurige reeks symbolen; de symbolen zijn afhankelijk van elkaar, wat de tekst voorspelbaarder maakt.

Markovs methode werd later, toen de computer pen en papier kon vervangen, gebruikt om de kans uit te rekenen dat op een bepaalde letter een andere volgt (‘Hoe groot is de kans dat een na een T een E volgt?’) of, nog uitgebreider, dat op een combinatie van letters een andere combinatie volgt (‘Hoe groot is de kans dat na de letters ‘GED’ de letters ‘AAN’ volgen?’). Zodra je die kansen weet, kun je weer nieuwe zinnen maken, die aan dezelfde statistiek voldoen. Ik voer de computer de acht eerste zinnen. Die berekent voor elk tweetal letters

(inclusief de spatie) de kans dat daarop weer een bepaald tweetal letters volgt. Zo genereert het model: ‘LÉRIE HETTE AT SCHOOLPLEIN TIEND DE BIJ DE TRIOMPHET IK WOOL INE MEGASTEN ZIJN’. Het lijkt me sterk dat ik met deze zin een prijs zal winnen.

Ik probeer de voorspellende kracht van het model te vergroten door langere combinaties van letters te gebruiken, maar dan begint de computer hele zinsdelen van de oorspronkelijke zinnen te herhalen. Dat is logisch: de lettercombinatie ‘IK W’ komt maar één keer voor in de acht eerste zinnen, waardoor er in het model ook maar één vervolg mogelijk is: ‘ERK’, waarop ook slechts één optie kan volgen, ‘BIJ’, enzovoort, tot het model uiteindelijk de hele eerste zin van de winnaar van vorig jaar geplagieerd heeft. Om dit te voorkomen heb ik meer data nodig: ik voer alle acht de essays, meer dan 140.000 tekens, aan de computer. Die geeft me, door steeds combinaties van vier letters aan elkaar te koppelen, de zin ‘KRIOELEN ROND JE VAN DE DRINKBAKKEN VOORSPELLEND ZIJN ALS ONBEHALVE DE CROISSANTJES’, dat onmiskenbaar beter is dan de vorige zin, prachtig zelfs, maar nog altijd ongeschikt als eerste zin van een essay.

Wat me opvalt aan de acht eerste zinnen, is dat vier ervan een jaartal bevatten: 1930, 2002, 2005 en 2020. Vijf zinnen bevatten het woord ‘ik’, het vaakst voorkomende woord na ‘de’ en ‘het’. Deze ik-zinnen onthullen steeds iets essentieels over het ik: de ik-figuur blijkt bijvoorbeeld bij Rijkswaterstaat te werken, of als postbezorger bij PostNL. Deze onthulling heeft ook in elke zin iets omineus. Je weet als lezer gelijk dat dit nog niet het hele verhaal kan zijn, zoals ‘Op het erf waar ik vroeger woonde staat nu een megastal.’ (Jolanda van Benthem, 2019) Tegelijkertijd bevat elke zin een bijzonder en specifiek detail, zoals die megastal, of die betrekking bij PostNL. Er is zelfs één zin met al deze prijswinnende eigenschappen – het jaartal, het woord ‘ik’, de omineuze onthulling, het specifieke detail: ‘Mijn eerste courgette at ik in september 2005.’ (Anouk Kootstra, 2020)

Had ik ook een jaartal in mijn eerste zin moeten zetten? Moet ik die eerste zin niet toch herschrijven, en in de ik-vorm zetten? ‘Hoe win ik anno 2026 een essayprijs?’, is dat specifiek genoeg, of moet er nog een bijzonder detail bij? En, misschien wel het belangrijkste: wat moet ik over mezelf onthullen? Mag ik niet toch opnieuw beginnen?

***

Ik had mijn eerste mystieke ervaring in 2004 bij het bekijken van een Flash-video. YouTube bestond nog niet. In plaats daarvan ging ik als elfjarige naar gprime.net, een website waar video's in Flash-formaat op verzameld stonden: korte, knullig geanimeerde video's, die daardoor makkelijk en snel te downloaden waren. Je scrolde langs een lange lijst met de hoofdletterloze bestandsnamen, met daarachter de bestandsgrootte in kilobyte. Een favoriet was bijvoorbeeld 'magical trevor (745 kB)', een kort liedje met een getekende clip, over een tovenaar die een koe liet verdwijnen. Een ander filmpje ('9/11 pentagon strike conspiracy (3052 kB)') probeerde me ervan te overtuigen dat de aanslagen op 11 september een door de Amerikaanse regering georganiseerde *inside job* waren. Over deze kwestie heb ik destijds lang getwijfeld. Mijn ouders durfde ik er niet naar te vragen, uit angst dat ze me zouden verbieden deze website nog langer te bezoeken.

Het meest mysterieuze filmpje op deze website heette 'craziest (3401 kB)'. Het was een geanimeerd kort verhaal, geschreven door de Amerikaanse schrijfster Liz Dubelman. Dat het om fictie ging, was mij in eerste instantie absoluut niet duidelijk. Ik klikte de video aan en er verschenen Scrabble-letters in beeld. Ik hoorde een vrouwenstem vertellen dat Scrabble voor haar een religie was, en het Scrabblebord een orakel. Hierbij herschikten de letters in beeld zich tot het woord 'ORACLE'. Haar leven, zo vertelde ze, draaide om de 'triple-triple', de mogelijkheid om op het Scrabblebord in een van de hoeken het woord 'CRAZIEST' te leggen. De C en de"T" verdrievoudigen dan allebei de woordwaarde, terwijl de Z, op zichzelf al de maximale 10 punten waard in de Engelse variant van het spel, op een vakje met dubbele letterwaarde komt te liggen. Wie dit woord weet te leggen, scoort in één beurt 311 punten.

'*I played game after game. After each one, I would copy the board into a notebook so I could divine the true meaning. I have over 1,700 notebooks, and each tells a tale as meaningful as the Bible. The books are layered in words and numbers, but the message is clear. Life is patterns. … Scrabble will give you clues to the future, but very hard clues.*'

Als ik wil beschrijven wat me als kind aan dit filmpje zo aansprak, kom ik uit bij taal die (ex-)katholieken vaak gebruiken om de kerkdiensten uit hun jeugd te beschrijven: 'Ook al begreep ik het niet, toch vond ik het mooi. Het ging duidelijk over iets hogers.'

Mijn voorliefde voor deze video moet zelf een *clue to the future* zijn geweest: zeven jaar na het zien ervan ging ik wiskunde studeren. Een bijkomend voordeel van deze studie was dat mensen

om me heen me boeken bleven aanraden over de relatie tussen taal en wiskunde. Borges, natuurlijk, met zijn korte verhaal over een universum dat uit een enorme bibliotheek bestaat, van zeshoekige kamers, met daarin boekenkasten, met daarin boeken, met daarin willekeurige reeksen van letters. Als de bewoners van dit universum erachter komen dat hun bibliotheek alle mogelijke boeken bevat, leidt dit aanvankelijk tot een jubelstemming. Daarna daalt het besef in dat ze met geen mogelijkheid de betekenisvolle boeken kunnen scheiden van de miljarden en miljarden malen talrijkere onzin. 'Miljarden en miljarden' is hier nog een understatement: de kans op zelfs maar één coherent geschreven pagina is 0,000…1, waar op de puntjes nog honderden nullen ingevoegd moeten worden.

*'*Op deze bovenmatige hoop volgde, logischerwijs, een zeer diepe depressie. De zekerheid dat een of andere plank in een of andere zeshoek waardevolle boeken bevatte en dat die kostbare boeken ontoegankelijk waren, leek bijna onduldbaar. Een blasfemische sekte stelde voor het gezoek te staken en alle mensen letters en symbolen dooreen te laten gooien tot ze, dankzij een onwaarschijnlijke gift van het toeval, die canonieke boeken hadden gevormd.'

De ik-figuur uit het voorwoord van Willem Frederik Hermans' verhalenbundel *Paranoia* worstelt met een vergelijkbaar probleem. Hij beschrijft de weerzin die hij voelt bij het schrijven op een nieuw, blanco vel papier:

'(...) wat opzettelijk aan betekenis wordt vastgelegd, is een zó gering deel van alles wat ik zou dienen op te schrijven, dat het hoogstens met twee of drie haren uit het niet steekt waarin het verzinkt. (…) Er is maar één werkelijk woord: chaos. (…) Ik schrijf, hoewel ik weet dat men alleen één woord schrijven kan door er tienduizenden over te slaan.'

Als je stilstaat bij het aantal mogelijke geschreven woorden, is een lege pagina niet langer leeg, maar een onmetelijk verkeersknooppunt, waar op elke afslag weer een nieuwe afslag volgt, een treinstation met duizenden wissels, zo ver het oog reikt, een uitgestrekt paleis met steeds weer splitsende gangen. De ingebeelde aanblik van al deze mogelijkheden roept bij mij dezelfde gevoelens op als het bekijken, ergens ver buiten de stad, van een heldere sterrenhemel: nietigheid, ten eerste, en, ten tweede, de wens om in dit alles toch in Godsnaam een Systeem te ontwaren – een Systeem waar de personages van Dubelman en Borges vergeefs naar zoeken, en dat door Hermans werd afgezworen. Elke zin is een wonder in het licht van de talloze manieren waarop die had kunnen mislukken.

***

Het streven om een computer taal te laten voortbrengen, is lang stukgelopen op de ongelooflijke hoeveelheid zinnen die er denkbaar zijn. In de essaybundel *Rekenen op taal*, geschreven door Hugo Battus, een pseudoniem van de in 2014 overleden schrijver en computationeel taalkundige Hugo Brandt Corstius, legt hij uit waarom: simpele taalmodellen weten zich geen raad met correlaties in langere zinnen. Brandt Corstius geeft als voorbeeld het woord 'respectievelijk': als er twee, drie, of tien woorden vóór het woord 'respectievelijk' staan, moet het woord gevolgd worden door (respectievelijk!) twee, drie of tien andere woorden. Het simpele taalmodel dat Brandt Corstius in zijn essay bespreekt kan die complexiteit niet aan. Hetzelfde geldt voor het door Markov bedachte model waarmee ik eerder een goede eerste zin hoopte te genereren: zodra zo'n model het woord 'RESPECTIEVELIJK' heeft geschreven, gaat het door naar de volgende lettercombinatie, zonder te onthouden hoeveel woorden er daarna nog moeten volgen.

Je zou dit probleem kunnen proberen te omzeilen door het model met veel langere lettercombinaties te laten rekenen, zo lang dat ze een hele zin bevatten: hoe groot is de kans dat op de ene zin de andere volgt? Daarbij stuit je op weer een nieuw probleem, de zogeheten 'vloek van de dimensionaliteit': er zijn veel en veel te veel mogelijke zinnen. Het is haast niet te bevatten, maar zelfs als je alle ongrammaticale zinnen wegdenkt, zijn er zo ongelooflijk veel mogelijke zinnen, dat het gros ervan nog altijd niet gezegd is. Het is een prachtig gegeven (het is nog altijd mogelijk om de eerste te zijn die een bepaalde zin zegt!), maar het heeft één nadeel: het maakt het uitrekenen van 'een gemiddelde kans' dat op een bepaalde zin een andere volgt onmogelijk.

Een ander essay van Brandt Corstius bespreekt de geschiedenis van de vertaalmachine. Hij schreef het in 1973. Het aanvankelijke optimisme over de mogelijkheid van automatisch vertalen, dat in de jaren vijftig ontstond toen meer en meer onderzoeksinstituten de beschikking over een computer kregen, was toen alweer verdampt. Sterker nog, Brandt Corstius noemt de vertaalmachine een 'utopie'. Hij haalt de Israëlische taalfilosoof en wiskundige Yehoshua Bar-Hillel aan, die de voorbeeldzinnen '*the pen is in the box*' en '*the box is in the pen'* bedacht: in de ene zin betekent 'pen' gewoon 'pen', maar in de andere zin betekent 'pen' 'hok'. Als mensen

achterhalen we die betekenis doordat een doos niet in een pen kan staan, maar wel in een hok. Maar hoe moet een computer daar ooit achter komen?

‘Zijn argumenten [van Bar-Hillel] komen in het kort hier op neer dat voor formeel vertalen behalve een woordenboek ook een encyclopedie nodig is, en wel een encyclopedie waarin alle kennis over de wereld te vinden is. (…) Wij weten niet hoe we onze ervaringen zo opslaan dat we er later uit kunnen putten. We weten ook niet hoe we dat een computer moeten leren.’

Toen ik dit essay uit had, kon ik het niet laten om de zin ‘the box is in the pen’ aan een hedendaags AI-model voor te leggen. Die geeft de correcte vertaling, en antwoordt door te zeggen dat ‘dit voorbeeld vaak gebruikt wordt om te laten zien hoe lastig automatisch vertalen kan zijn’. Ook door mijzelf verzonnen voorbeelden (‘Er is een slot in een kamer’ en vice versa) worden correct vertaald. Wat Hugo Brandt Corstius vijftig jaar geleden nog een utopie noemde, is inmiddels werkelijkheid geworden. Of het daadwerkelijk een utopie is, valt nog te bezien.

***

Nu is dan toch het moment aangebroken om het over AI te hebben, of, specifieker, de grote taalmodellen, de LLM’s. De wiskundige in mij kan niet anders dan de taalmodellen prachtig vinden: tachtig jaar lang hebben we geprobeerd om een computer taal te laten voortbrengen, en ineens is het gelukt. Alle andere facetten van mijn persoonlijkheid verzetten zich ertegen: tegen de hielenlikkende miljardair-eigenaars, het totaal onverantwoorde gebruik van fossiele brandstoffen voor de benodigde datacentra, het inhalige onbetaalde gebruik van auteursrechtelijk beschermde teksten, de afzichtelijke eenvormigheid van de geproduceerde AI-teksten, enzovoort. Soms denk ik dat het idee van een taalmodel slechts als wiskundige abstractie verkend had moeten worden, zonder het ooit daadwerkelijk uit te voeren. Het probleem is dat we dan ook nooit hadden kunnen weten of de techniek werkt: niemand heeft daar een goede reden voor, sterker nog, niemand weet eigenlijk hoe het werkt. We weten alleen hoe we de omstandigheden kunnen creëren waarin het systeem zichzelf aanpast om de antwoorden te geven die het doet.

Een veelvoorkomende uitleg van dit soort systemen is dat ze het ‘gemiddelde’ of ‘meest waarschijnlijke’ woord kunnen bepalen. Zo is het taalmodel slechts een ‘stochastische papegaai’, die op basis van kansberekening een enorme hoeveelheid data napraat.

Cultuursocioloog Siri Beerends en filosoof Lisa Doeland schreven in *De Groene Amsterdammer* bijvoorbeeld dat we ‘met een veredelde rekenmachine te maken hebben die slechts het meest waarschijnlijke woord in een zin berekent’. Of, in een column van Peter Munneke in *NRC*: ‘De essentie van generatieve taalmodellen is dat ze altijd het meest waarschijnlijke vervolg kiezen. (...) Het probleem (...) vind ik precies de neiging van taalmodellen om zo gemiddeld mogelijke teksten te maken.’ Of lector IT Pascal Wiggers, in *Trouw*: ‘[Een AI-taalmodel is] niet meer dan een wiskundig model dat berekent wat het meest waarschijnlijke volgende woord in een gesprek is.’

Eerlijk gezegd begon het nadenken over dit essay omdat ik steeds las dat deze taalmodellen ‘het gemiddelde’ of ‘het meest waarschijnlijke’ woord bepalen. Nu ik er langer over heb nagedacht geloof ik eigenlijk niet meer dat ‘het gemiddelde woord’ bestaat: er is, statistisch gezien, geen waarschijnlijkste woord waarmee deze zin eindigt, omdat ik de eerste ben die deze zin schrijft. Dat we aanvoelen dat de vorige zin zal eindigen met het woord ‘schrijft’ is precies het taalgevoel waarvan we ons al zo lang afvragen hoe we het de computer kunnen laten nadoen.

De opmerking dat de computer een gemiddelde uitrekent is bedoeld als een nuttige correctie op het idee dat het taalmodel de wijsheid in pacht heeft. Deze formulering verhult alleen de keuzes die bij het opstellen van het taalmodel zijn gemaakt. Een gemiddelde suggereert objectiviteit, alsof er met de data een vastomlijnde berekening is uitgevoerd om dat gemiddelde te bepalen. Dat is niet zo. De ‘gemiddelde’ toon van de chatbot is bijvoorbeeld al een keuze: de chatbot spreekt zo overdreven enthousiast en niet-aanstootgevend omdat het bedrijf erachter vooral geen negatieve publiciteit wil krijgen.

Het lijkt me het beste om te blijven onthouden wat een taalmodel daadwerkelijk is: een enorme collectie computerprogrammaatjes, die door middel van een training op een automatische manier zijn aangepast om zo goed mogelijk te zeggen wat de baas wil horen. Die training bestaat voor een groot deel uit het gokken van woorden uit bestaande stukken tekst. Zoals het model van Markov op basis van de voorgaande paar letters een suggestie voor de volgende lettercombinatie doet, doet een LLM dat op basis van veel meer van de voorgaande tekst, waarbij die input ook nog eens op een complexe manier aan elkaar wordt geknoopt.

Bij een zin als ‘the pen is in the box’ is het ondoenlijk om voor elke combinatie van objecten de kans te bepalen dat object A in object B zit. Het taalmodel moet dus een omleg aanleggen: puur gedreven door de opdracht om beter het gewenste antwoord te kunnen gokken, zal het intern een programmaatje aanmaken, dat bij een woord kan zeggen hoe groot het object ongeveer is, waarmee weer bepaald kan worden dat ‘de pen is in de doos’ een meer gewenste zin is dan ‘de doos is in de pen’. Ondanks dat de AI de betekenis van deze woorden niet begrijpt, zal het model zo toch conclusies over de wereld gaan trekken die wij correct vinden, mits het model met coherente tekst getraind is. Sterker nog, zo blijkt het wiskundige bewijzen te kunnen leveren die nog niet eerder door iemand waren bedacht.

***

Omdat je op het einde van een essay toch terug hoort te grijpen op het begin, had ik dit essay willen eindigen met de vraag: ‘Hoe eindigt het gemiddelde essay?’ Bij nader inzien bestaat het gemiddelde essay niet. Ik hoef me over die vraag dus niet langer druk te maken.

De personages van Borges gaan uiteindelijk aan hun Bibliotheek ten onder: ‘De zekerheid dat alles is geschreven doet ons teniet of maakt ons tot schimmen.’ Te midden van een vloed aan AI-gegenereerde teksten is dat gevoel makkelijker voor te stellen dan ooit. Tegelijkertijd heb ik na het schrijven van dit essay een sprankje nieuwe hoop: zolang er tekst is die ik wil schrijven, kan ik in het paleis van mogelijke zinnen blijven dolen. Het monster zal ik er niet hoeven tegenkomen; de uitgestrektheid van de taal zal me beschermen.

Verantwoording

‘Craziest’, het korte verhaal van Liz Dubelman, is hier te vinden op YouTube:
https://www.youtube.com/watch?v=ORGMRi_nAgM

Zelf experimenteren met Markov-taalmodellen kan hier:
https://projects.haykranen.nl/markov/demo/

Voor Borges’ verhaal ‘De Bibliotheek van Babel’ heb ik gebruik gemaakt van de bundel *De Aleph en andere verhalen*, De Bezige Bij, 2020 (eerste druk: 1998), in de vertaling van Barber van de Pol. Het Spaanse origineel verscheen in 1941 in de bundel *El jardín de senderos que se bifurcan*.

Het voorwoord van de bundel *Paranoia* van Willem Frederik Hermans is te vinden in die bundel, Van Oorschot, 1999 (eerste druk: 1953).

De essaybundel van Hugo Brandt Corstius is te vinden als *Rekenen op taal*, van auteur Hugo Battus, Querido, 1983.