

Michael van Hartskamp

Doelenstraat 18, 2282 MX Rijswijk

michael@vanhartskamp.com

\TeX -rubriek

De kracht van \TeX

Het professioneel drukken van een wiskundige tekst is een gecompliceerde en kostbare aangelegenheid. Bovendien is de oplage vaak klein, bijvoorbeeld wanneer het een proefschrift betreft. Vroeger werd zo'n tekst getypt, waarna er met de hand de formules in werden geschreven. Het programma troff betekende een behoorlijke vooruitgang en het is enige tijd populair geweest. Het programma \TeX heeft een grote omwenteling in het zetten van wiskundeteksten veroorzaakt. Michael van Hartskamp, gepromoveerd topoloog (Vrije Universiteit) met een vader uit de drukkerswereld, zal een rubriek over dit onderwerp verzorgen.

Het lijkt zo tegenstrijdig: het Nieuw Archief probeert zich te vernieuwen en ondertussen start men met een rubriek over een computerprogramma dat al in 1978 beschikbaar was en sinds 1989 niet wezenlijk meer veranderd werd. \TeX , want daar hebben we het over, heeft echter een ongeëvenaarde bijdrage geleverd aan het wiskundige schrijfwerk. Daarom staan we daar graag even bij stil en voorzien we waar mogelijk van nadere informatie om het werken met \TeX te vereenvoudigen of te verbeteren.

De vraag is natuurlijk hoe \TeX zo'n sterke positie heeft weten te verwerven, en hoe \TeX deze heeft kunnen behouden. Daar zullen we deze eerste keer naar kijken. Naast de sterke punten gaan we ook in op de zwakke punten van \TeX en wat de \TeX -wereld gedaan heeft ter verbetering.

Geschiedenis

Tot ver in de vorige eeuw werden boeken volgens de klassieke methode gedrukt. De zetter bekeek de tekst en maakte aan de hand hiervan een verdeling van de tekst over de regels. De regels vormden samen weer een bladzijde. Elke regel bestond uit letters die één voor één op de regel geschoven werden. De aldus gezette tekst werd vervolgens gebruikt als drukplaat. Het op deze wijze foutloos zetten van tekst was een tijdrovende bezigheid. Zeker het zetten van wiskundige teksten was specialistisch werk, en dus extra kostbaar.

Het sprak voor zich dat uitgevers en drukkers naar andere minder kostbare methoden zochten. Een nog altijd veelgebruikte methode is om het te drukken materiaal te fotograferen en van die foto's drukplaten te maken. De verantwoordelijkheid voor de juiste vormgeving verschoof hiermee naar degene die het papieren origineel aanleverde.

Het correct afleveren van wiskunde was daarmee nog steeds niet opgelost. Terwijl gewone tekst eenvoudig op schrijfmachines kon worden getikt, was voor het wiskundig formulewerk extra werk nodig. Soms werden symbolen met de hand geschreven in speciaal open gelaten ruimte, maar voor een professioneler resultaat was al gauw een hoop geknoei nodig met speciale wielen die andere tekens aan de schrijfmachine konden onttrekken. Hoewel dit werk veelal werd uitbesteed aan niet-wiskundig opgeleide secretaresses was dit proces nog altijd zeer moeizaam en tijdrovend.

Omdat steeds meer niet-specialisten zich bezighielden met de voor-

bereiding van het drukproces was de kwaliteit van het drukwerk lang niet meer zo hoog als vroeger in de tijd van de gespecialiseerde zettters. Aan één zo'n tegenvallend resultaat is \TeX te danken. De Amerikaanse wiskundige en computerwetenschapper Donald Knuth bekeek de nieuwe edities van zijn beoogd zeventien delig levenswerk 'The Art of Computer Programming' en was hierover zeer ontevreden. Hij wist dat het mogelijk was om met de computer beter zetterwerk af te leveren. Speciaal voor zijn levenswerk bedacht hij, samen met zijn promovendi, \TeX : een programma speciaal geschikt voor het zetten van wiskunde. Vele jaren later, in 1997, leverde Knuth de eerste in \TeX gezette herdrukken van zijn levenswerk af.

\TeX was bedoeld om papieren originelen af te leveren die gebruikt kunnen worden in het drukproces. De laatste jaren werd het origineel voor de drukplaat echter steeds vaker aangeleverd als uitdraai van een printer en dus was de volgende logische stap in de ontwikkeling van de drukpers om meteen vanuit de computer de drukplaten te genereren. Zonder veel moeite bleek \TeX in staat ook deze overgang te maken.

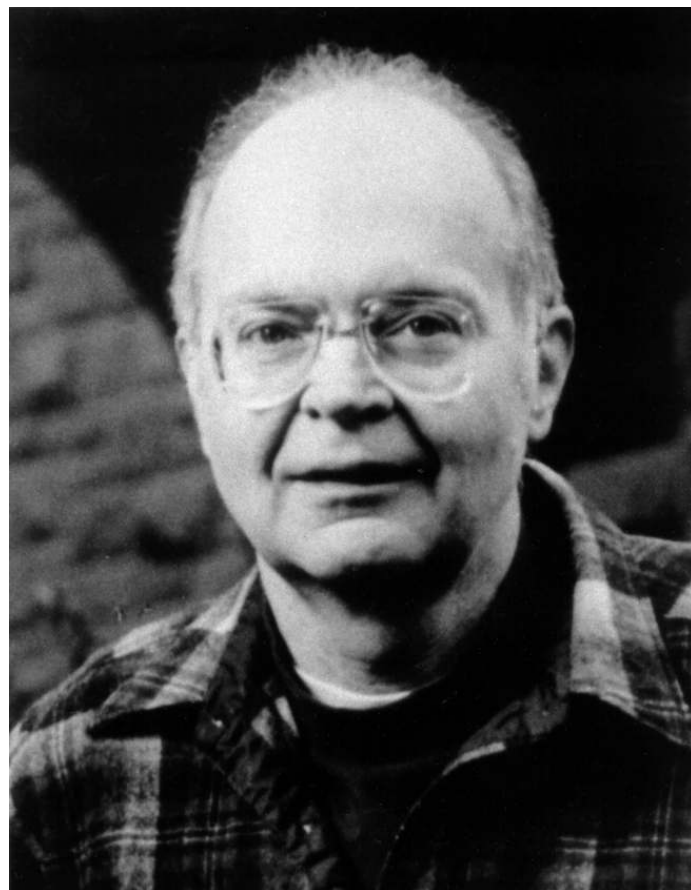
De ontwikkeling van \TeX door Knuth is een verhaal op zich. Met \TeX 78 werd in 1978 de eerste versie van \TeX getoond. Daarna is \TeX doorontwikkeld. Met het verschijnen van versie 3 nam Knuth een belangrijke stap. Hij wilde \TeX niet meer verder ontwikkelen. Alleen het weghalen van fouten in het programma zou nog tot veranderingen kunnen leiden. Bij iedere fout werd het versienummer aangevuld met een extra cijfer uit de decimale ontwikkeling van π . Het aantal gevonden fouten sinds versie 3 is vrij gering gebleven. Knuth durft het dan ook aan om op de volgende fout een beloning van \$327.68 te zetten. Een beloning die hij bij het vinden van een eerdere fout steeds heeft verdubbeld.

De eerste versie

De ontwikkeling van een programma zoals \TeX was voor Knuth ook vanuit onderzoeksperspectief interessant. Zo werd voor \TeX een algoritme ontwikkeld dat (engelse) woorden vrijwel altijd correct afbreekt. De uitdaging zit erin dat de computer geen lettergrepen onderscheiden kan en niet genoeg geheugen had voor een volledige woordenlijst. Op basis van de opeenvolging van specifieke letters worden voor een woord de mogelijke afbreekplaatsen gewaardeerd. De plaatsen met de hoogste score worden als toelaatbaar gekenmerkt. Een bepaalde collectie patronen zorgt voor de juiste gewichten.

Een ander nevenproduct van \TeX was het zogenaamde 'literate programming': \TeX werd zelfs herschreven in het daarvoor ontwikkelde pascaldialect Web. De programmeur werd geacht zijn programma te documenteren met \TeX -tekst. Zo ontstaat een helder begrijpelijk computerprogramma dat enerzijds door de computer kan worden geïnterpreteerd en anderzijds, dankzij \TeX , ook door de mens.

De belangrijkste reden van het succes van \TeX in de wiskundige wereld was ongetwijfeld het gebrek aan concurrentie. De standaard



Donald E. Knuth

Foto: Stanford Engineering School

tekstverwerkings programma's WordStar, Word en WordPerfect boden aanvankelijk geen enkele ondersteuning voor wiskundige formules. \TeX daarentegen was gebouwd met wiskunde in het achterhoofd en kon eigenlijk iedere wiskundige tekst vrijwel probleemloos aan.

Macropakketten

Toch werd \TeX zeker door de wiskundige gebruikers Spartaans gevonden, de informatici hadden daar wat minder moeite mee. Om iets eenvoudigs te maken was vaak al veel programmeerwerk in een onplezierige programmeertaal nodig. Van diverse kanten werd daaraan gewerkt: Leslie Lamport maakte het populaire \LaTeX , de AMS maakte $\text{Ams}\TeX$. \LaTeX zou uitgroeien tot het standaard macropakket. Ook het werk van de AMS werd uiteindelijk in \LaTeX geïntegreerd. Aan \LaTeX wordt nog altijd gewerkt. Versie 2.09 werd opgevolgd door een versie 2 ϵ . Deze laatste versie moet al jarenlang dienen als voorloper van \LaTeX 3. De toegenomen (vrij ongestructureerde) complexiteit van \LaTeX , maakt het pakket soms erg ondoorzichtig. Voor programmeurs en gebruikers blijft het doorgronden van \LaTeX een bijna ondoenlijke puzzel. De Nederlander Hans Hagen heeft de laatste jaren besteed aan het schrijven van een doorzichtiger macropakket met vergelijkbare functionaliteit als \LaTeX . Het resultaat is $\text{Con}\TeX$. Met behulp hiervan wordt het Nieuw Archief gezet. Door wiskundigen wordt het overigens nog maar mondjesmaat gebruikt.

De kracht van \TeX

\TeX 's mogelijkheden voor het (correct) vormgeven van wiskunde zijn nog altijd ongeëvenaard. \TeX kent vrijwel alle regels die de vroegere zettters ook in acht namen.

```
This is TeX, Version 3.14159 (Web2C 7.3.1)
**\relax

*Welkom in de wondere wereld van \TeX!

*\bye
[1]
Output written on texput.dvi (1 page, 260 bytes).
Transcript written on texput.log.
```

Figuur 1 Hoezo is \TeX onvriendelijk?

Door de mogelijkheid erin te programmeren is T_EX geschikt om het vervelende handwerk van het schrijven te automatiseren. T_EX kan al jaren hoofdstukken, paragrafen en literatuurverwijzingen automatisch nummeren en dat voorkomt veel werk voor de maker van inhoudsopgaven en indices, en ook een hoop ergernis voor de lezer die altijd net die ene onjuiste verwijzing treft die een handmatige nummering onherroepelijk oplevert.

T_EX is zeer sterk in het zetten van alinea's tekst. Waar de meeste tekstverwerkers per regel besluiten of een woord naar de volgende regel wordt verplaatst of niet, doet T_EX dit per alinea. Zo kan een suboptimale keuze in één regel leiden tot een veel mooier geheel en er ontstaat een uniformer tekstbeeld, dat daardoor prettiger leest. T_EX maakt zijn keuzes op basis van het kortste-pad-algoritme. In essentie komt het hier op neer: de steden voor het pad zijn de spaties tussen de woorden. Een pad tussen twee steden komt overeen met een regel tekst voor de alinea. De prijs van een pad wordt bepaald aan de hand van de hoeveelheid tekst en de wijze waarop deze kan worden gezet: ver uiteen staande tekst of gedrongen tekst is lelijk, en het bijbehorende pad wordt lang gemaakt, enzovoort. Ook tussen opeenvolgende regels is een relatie: gedrongen tekst gevolgd door veruiteen staande tekst leest niet prettig en het bijbehorende pad is dus ook relatief duur.

De kracht van T_EX is dat het meestal automatisch goed gaat. Afwijken van de regels wordt vaak erg moeilijk gemaakt. Als T_EX al eens een steekje laat vallen moet de auteur dus handmatig ingrijpen. Dat komt maar zelden voor en dus zullen veel auteurs de trucjes niet kennen of niet realiseren dat ze ze moeten gebruiken. Ook aan deze zaken zullen we aandacht besteden. We zullen voorgaande met een beroemd voorbeeld illustreren. Aan het eind van een zin (na een punt) laat T_EX meer witruimte dan tussen woorden. Het menselijk oog wordt hierdoor geholpen. T_EX kan echter geen verschil maken tussen een afkorting en het eind van de regel, d.w.z. dat achter de 'z' in de voorafgaande afkorting 'd.w.z.' er teveel witruimte komt omdat T_EX ten onrechte denkt dat de zin is afgelopen. Door T_EX te vertellen dat hier een gewone spatie hoort is het probleem verholpen. Dat laatste kan door achter de punt een T_EX-backslash te zetten gevolgd door een spatie. Hoewel de afkorting in de vorige zin de tweede keer tussen aanhalingstekens is geplaatst geldt ook daar hetzelfde.

Verder blijkt T_EX ook prima geschikt voor het elektronisch of op papier aanleveren van tekst aan de drukker. Wel blijkt dat het standaard T_EX-lettertype soms iets te dun is voor de moderne drukker. Ook voor het publiceren op het Internet is T_EX geschikt, maar dan is de omweg via PDF aan te raden.

Zwakke punten

De eerlijkheid gebiedt ook de zwakke punten te bekijken. Het gebruik van een ander lettertype was vroeger een heidens karwei. Nu is het een vrij eenvoudig op te lossen probleem en zijn duizenden commerciële lettertypes ook bruikbaar binnen T_EX. Voor vrijwel iedere wiskundige tekst zijn echter speciale symbolen onontbeerlijk en de mooiste afdruk wordt verkregen indien de symbolen bijpassend vormgegeven zijn. Helaas bevatten de commerciële lettertypes niet veel meer dan een plus- en een minteken. Er zijn mij slechts drie lettertypes bekend waarvoor de meest voorkomende wiskundige symbolen in bijpassende stijl beschikbaar zijn.

Het gebruik van afbeeldingen in tekst is iets waar T_EX eigenlijk nooit voor gemaakt is. Het levert dan ook tot op de dag van vandaag regelmatig problemen op. Wie een folder of een glossy magazine wil maken kan T_EX beter thuis laten.

Tot slot T_EX is nog steeds duidelijk gericht op technuten. De instal-

latie is inmiddels stukken eenvoudiger dan ze vroeger ooit was, maar T_EX blijft een niet-interactief programma dat meer een programmeerhulpmiddel is dan een tekstverwerker. Het communiceren met T_EX blijft programmeren. Er zijn wel enkele 'moderne' point-and-click interfaces: voor ervaren T_EX-ers zijn die een ergernis, maar ook voor de beginner is het behelpen.

De conclusie is duidelijk. T_EX heeft een lange geschiedenis en die is nog niet ten einde. T_EX's sterkste punt is het mooi weergeven van tekst en wiskunde. Voor een wiskundige is het plezierig dat niet alleen de inhoud maar ook de vorm prettig is en het lezen van zijn/haar werk vereenvoudigt, waardoor de toegankelijkheid verbetert. De zwakke kanten zijn dat T_EX als moeilijk en ontoegankelijk wordt ervaren en niet zo goed met afbeeldingen overweg kan.

Wordt vervolgd

In deze rubriek zal ik proberen om met name op deze laatste punten verlichting te brengen. De volgende keren zal deze rubriek in het teken staan van: T_EX thuis, T_EX vragen beantwoord krijgen en plaatjes. Maar ook leuke hulpmiddelen zoals het programma BibT_EX voor de referenties aan het eind van een artikel zal ik onder de loep nemen. ↩