

Wat is T_EX?

Victor Eijkhout en Nico Poppelier

versie 0.8 – 20 december 1989

Samenvatting

In de huidige tijd van WYSIWYG ‘desktop publishing’ systemen lijkt het vreemd een systeem te propageren met honderden, vaak ingewikkelde, instructies. Wat het tekstopmaakstelsel T_EX desondanks aantrekkelijk maakt, is het feit dat het programmeerbaar is: de basisinstructies kunnen worden samengevoegd tot zeer gecompliceerde instructies die een eenmaal ontworpen opmaak geheel automatisch feilloos reproduceren. Hierdoor hoeft de gebruiker uitsluitend nog de structuur van het document aan te geven en niet de vormgeving, het uiterlijk, ervan. Met deze gedachte sluit T_EX nauw aan bij recente ontwikkelingen in de uitgeverwereld omtrent SGML.

1 Inleiding

T_EX¹ is een tekstopmaakstelsel met een wat merkwaardige ontstaansgeschiedenis. Het is ontwikkeld door Donald Knuth, een hoogleraar computerwetenschappen aan de universiteit van Stanford in Californië. Terwijl hij bezig was met het schrijven van zijn boekenreeks ‘The Art of Computer Programming’ kreeg hij ideeën voor een ‘programmeertaal voor tekstopmaak’. Tussen 1978 en 1984 groeide T_EX zo geleidelijk tot zijn huidige vorm.

Het is, toegegeven, een merkwaardig idee dat een softwareprodukt zich al meer dan zes jaar niet meer ontwikkeld heeft. Toch betekent dat niet dat T_EX verouderd is; veeleer was het zijn tijd ver vooruit toen het ontstond. In die jaren is overigens slechts een handjevol fouten in het programma gevonden.

Er zijn meer ongewone aspecten aan T_EX. Zo is er voor praktisch elk type computer wel een versie van T_EX, en vaak zelfs een gratis, een zogenaamde ‘public domain’, versie. De uitvoer van T_EX is al net zo universeel: voor bijna elke laserprinter en fotozetter is er wel een programma dat door T_EX opgemaakte teksten kan afdrukken, ongeacht op welke computer het opmaken is gebeurd.

Bij veel mensen die wel eens van T_EX gehoord hebben, heeft het idee post gevat dat het een moeilijk systeem is, alleen te bedienen door computerprogrammeurs. Dit misverstand is begrijpelijk gezien de ontstaansgeschiedenis van T_EX. Het is echter niet juist.

De waarheid heeft twee kanten. Enerzijds heeft T_EX ongewoon veel opdrachten, enkele honderden, wat meer is dan gewone zetsystemen hebben, en veel meer dan gewone programmeertalen. Ook is de structuur van T_EX gewoonweg lastig in vergelijking met normale zetsystemen en programmeertalen. Anderzijds kan iemand die bedreven is in T_EX de ‘primitieve’ opdrachten van T_EX samenvoegen tot nieuwe opdrachten op een hoger niveau, ‘macro’s’ genaamd, die zeer gecompliceerde dingen doen. Hiermee kan het bedienen van T_EX dus willekeurig eenvoudig worden gemaakt.

2 Wat kan T_EX zoal?

T_EX heeft in zijn simpelste vorm een aantal elementaire zetinstructies ingebouwd als ‘zet de volgende tekst cursief’, ‘geef tien punt witruimte’ en ‘gebruik vrije regelval’. Omdat T_EX geprogrammeerd kan worden, kunnen veel ingewikkelder instructies dan dit gerealiseerd worden.

Een voorbeeld: iemand zou een macro ‘paragraaf’ kunnen maken zodat een gebruiker intikt `\paragraaf` `Voordelen van TeX` en daarmee de volgende serie handelingen veroorzaakt:

- er komt een zekere hoeveelheid witruimte achter de voorafgaande tekst;
- het systeem houdt een ‘paragraafteller’ bij en drukt de waarde daarvan af, gevolgd door de tekst van de paragraafkop, in een zeker lettertype en -grootte; daarna wordt de paragraafteller met één vermeerderd;

1. Uitgesproken ‘tech’; dit is de hoofdlettervorm van $\tau\epsilon\chi$; de officiële schrijfwijze wanneer de gezakte ‘E’ niet mogelijk is, is TeX.

- de titel wordt, met het bijbehorende nummer en met het paginanummer, automatisch aan een door het programma gegenereerde inhoudsopgave toegevoegd;
- als er kopregels gewenst zijn, wordt de paragraaftitel daar automatisch in ingevuld;
- achter de titel komt wat witruimte, met die verstande dat T_EX nooit de pagina zal breken tussen de titel en de daaropvolgende tekst – al naar gelang wat van te voren besloten is kan de witruimte boven de titel wat samengedrukt worden, of kan de titel naar de volgende pagina verhuizen, waarbij de huidige pagina mogelijk opgerekt wordt om de onderste regel te laten aansluiten met andere pagina's;
- de eerste alinea na de titel springt niet in, maar alle daarop volgende wel.

Het valt zelfs zo te regelen dat de gebruiker intikt
`\paragraaf Voordelen van`
`TeX\referentie{voordelen}`
 en dat hij later kan schrijven
`... zoals betoogd in`
`paragraaf \refereer{voordelen}`
 waarbij T_EX dan automatisch het juiste nummer van de paragraaf invult.

Verschillende mensen over de wereld hebben al uitgebreide pakketten van dit soort 'hoog-niveau opdrachten' voor T_EX gemaakt. Het bekendste pakket – in T_EX termen: een *format* – is L^AT_EX, dat uit 1985 stamt en is ontworpen door Leslie Lamport, een medewerker van *Digital Equipment Corporation*. L^AT_EX is gebaseerd op het idee van beschrijving van de logische structuur van een document: de opdrachten die in de tekst staan geven informatie over de *functie* van onderdelen van de tekst. De vertaling van de logische structuur naar een visuele structuur is vastgelegd in zogenaamde 'documentstijlen'. De gebruiker moet aan het begin van een document eerst aangeven 'dit is een brief' of 'dit is een artikel', en afhankelijk hiervan staan hem opdrachten ter beschikking. Zo bevat de artikel-stijl wel de opdracht 'titel', maar de brief-stijl niet.

Over het algemeen geldt dat wat T_EX nog niet kan, hem geleerd kan worden. Hier is echter wel wat vaardigheid voor nodig. In het tijdschrift van de T_EX-gebruikersgroep verschijnen regelmatig oplossingen op allerlei zetproblemen. De eerste auteur van dit artikel heeft bijvoorbeeld 'variozetsel' en 'hangend inspringen' op zijn naam staan.

Los van deze toeters en bellen is het resultaat van T_EX op het gebied van gewone, 'platte' tekst vaak al beter dan dat van andere systemen. De reden hiervoor is het feit dat T_EX niet per regel werkt maar per alinea: pas als een complete alinea af is wordt bepaald waar de regels gebroken gaan worden. Aangezien T_EX ook een bijzonder goed woordafbrekingsstelsel heeft, leidt dit proces tot zeer egale alinea's. Overigens vallen de beslissingen van T_EX hoe de alinea gebroken moet worden op velerlei

wijze te beïnvloeden.

Op het gebied van woordafbreking heeft T_EX nog meer in huis: er zijn voorzieningen om het Duitse woord 'backen' te laten afbreken als 'bak-ken', en om de trema in 'onbeïnvloedbaar' te laten verdwijnen in 'onbe-invloedbaar'.

3 De commerciële status van T_EX

Het draagt bij aan de vreemde status van T_EX dat het in zeker opzicht een 'gratis' pakket is. Voor enkele, op universiteiten veel voorkomende, typen computers is tegen kostprijs een magneetband met T_EX te krijgen, en ook voor microcomputers als MS-DOS-computers, de Apple Macintosh en de Atari ST zijn gratis versies te krijgen. Aan deze gratis versies zijn meestal twee problemen verbonden: één waar ze te krijgen, en twee, hoe ze te installeren. Er zijn echter ook commerciële versies van T_EX, die vaak meer gebruiksgemak bieden of sneller zijn.

Een aanzienlijk deel van het zetwerk dat momenteel met T_EX gedaan wordt is duidelijk als zodanig te herkennen. Dit is grotendeels te wijten aan het 'gratis' karakter van T_EX. Het enige lettertype dat namelijk normaliter meegeleverd wordt, is een kopie van de Bodoni-achtige 'Modern 8A' van het bedrijf Monotype. Er is echter een bedrijf dat een groot aantal ITC-lettertypes voor T_EX levert, zoals Times Roman, Baskerville, Bembo en Gill Sans. Ook zijn er bedrijven die lettertypes voor niet-Westeuropese talen leveren, die kunnen worden gebruikt met T_EX.

Het zetwerk van het L^AT_EX-pakket is meestal ook eenvoudig te herkennen als de gebruiker de standaard meegeleverde documentstijlen heeft gebruikt, die nogal 'Amerikaans' zijn. L^AT_EX is echter zeer zeker niet synoniem met deze opmaak, zoals de beide auteurs van dit artikel bewezen hebben. De eerste auteur heeft, met advies van een typograaf, documentstijlen ontwikkeld die dezelfde opdrachten bevatten als de L^AT_EX documentstijlen ('artikel', 'rapport', 'boek' en 'brief'), maar met een meer Nederlandse opmaak; de tweede auteur heeft, in opdracht van een grote uitgever, documentstijlen ontwikkeld die exact de opmaak van reeds bestaande tijdschriften reproduceren.

4 Wie gebruikt T_EX?

Gezien de ontstaansgeschiedenis van T_EX is het weinig verbazend dat het eerste antwoord op de vraag 'Wie gebruikt T_EX?' luidt: wiskundigen, natuurkundigen, sterrenkundigen en informatici. Het wordt hen dan ook erg aantrekkelijk gemaakt T_EX te gebruiken, aangezien steeds meer tijdschriften artikelen in T_EX-vorm accepteren – en dat vaak belonen met versnelde publicatie.

Ook buiten de wetenschappelijke wereld begint T_EX de laatste tijd geaccepteerd te raken. Zo werkt de Amerikaanse ‘TV-Guide’ – wekelijks meer dan honderd verschillende edities, met een gezamenlijk omvang van ruim 10 000 pagina’s – al geruime tijd met T_EX. Dichter bij huis gebruikt het Energiebedrijf voor Groningen en Drente T_EX voor het drukken van handleidingen.

In beide gevallen garandeert de macro-faciliteit van T_EX dat, zonder dat de mensen die de teksten voor bijvoorbeeld handleidingen intoetsen een grafische opleiding gevolgd hebben, de uitvoer geheel uniform in opmaak is. Voor documenten zoals handleidingen, waarvan geregeld nieuwe versies verschijnen, is de bovengenoemde mogelijkheid om door middel van symbolische sleutels aan secties en andere genummerde onderdelen van de tekst te kunnen refereren natuurlijk bijzonder waardevol: het betekent dat verwijzingen altijd correct zijn, zonder dat ze ooit gecontroleerd hoeven worden.

5 T_EX in het gebruik

T_EX lijkt beduidend meer op klassieke zetsystemen dan op de moderne ‘desktop publishing’ programma’s. Eerst moet een tekst ingetoetst worden, besprekend met opdrachten, en pas nadat de computer deze invoer bewerkt heeft kan het resultaat op het scherm of op papier bewonderd worden.

Omdat T_EX niet echt visueel werkt, is het voor toepassingen als advertentie-opmaak minder geschikt. De grote kracht van T_EX ligt echter in de mogelijkheid een bepaalde, vooraf vastgelegde opmaak geheel te kunnen automatiseren. Enkele voorbeelden: een tijdschrift, een boekenreeks, handleidingen of brochures van een bedrijf, de concertprogramma’s van een impressariaat of een regelmatig opnieuw verschijnende prijslijst.

Terwijl tot nu toe een grafisch opgeleid persoon zich strikt aan bepaalde richtlijnen moest houden om een reeks publicaties een uniforme opmaak te geven, kan een competent T_EX-programmeur deze opmaak in een aantal eenvoudig te gebruiken macro’s verstoppen.

Ten eerste vermindert dit de kans op fouten aanzienlijk en ten tweede betekent het dat er niet meer continu een graficus nodig is. Zelfs voor gebruik binnen een grafisch bedrijf zijn er voordelen aan het gebruik van T_EX: voor een opmaak die meerdere malen gebruikt moet worden werkt het produktieverhogend, terwijl bij eenmalige documenten de auteur makkelijker correcties aan kan brengen.

De vraag waar iemand de kennis vandaan haalt om T_EX-macros te ontwikkelen, kan beantwoord worden al naar gelang de behoefte. Zo zijn er grote bedrijven die hun eigen ‘T_EXpert’ in huis hebben. Aan

de andere kant van het spectrum, waar bijvoorbeeld wetenschappers hun eigen artikelen in T_EX zetten, bestudeert men de boeken over T_EX en L^AT_EX, en leert men al doende.

Tussen deze uitersten zitten nog enkele mogelijkheden: er zijn cursussen in T_EX en L^AT_EX, al dan niet gegeven onder auspiciën van de internationale T_EX Users Group, en er zijn free-lance T_EX-programmeurs die hun diensten aanbieden voor het ontwikkelen van L^AT_EX documentstijlen of gespecialiseerde macros.

6 T_EX als ‘engine’

Voor veel programma’s, zoals databases, geldt dat het uitvoer verzorgende gedeelte redelijk schamel is. Per slot van rekening is een database-programma in de eerste plaats ontworpen om gegevens op te slaan, niet om deze af te drukken. De programmeerbaarheid van T_EX heeft tot gevolg dat het systeem bij uitstek geschikt is om als ‘print engine’ voor bijvoorbeeld een database te fungeren. Als de rapportgenerator van een database-programma, of een ander uitvoer genererend programma, in staat is T_EX-instructies tussen de gegevens te plaatsen, kan de kracht van T_EX benut worden om de uitvoer op elke gewenste manier af te drukken.

Hoewel wij mensen kennen die zelf hun database-programma zo ver gekregen hebben dat het T_EX macros in de uitvoer zet, wordt deze toepassing van T_EX nog vrij weinig door producenten van programmatuur zelf toegepast. Als voorbeeld kennen wij wiskundige pakketten waarmee gecompliceerde formules berekend kunnen worden; het pakket kan dan de T_EX-instructies voor een dergelijke formule opleveren. Een ander voorbeeld is het Amerikaanse computerbedrijf Digital Equipment Corporation, dat programmatuur gebruikt voor het schrijven en onderhouden van handleidingen, waar T_EX als afdruckmachine achter zit. De persoon die de handleidingen intoetst krijgt echter nooit iets van T_EX te zien.

Dergelijke toepassingen zullen in de toekomst waarschijnlijk meer gaan voorkomen, nu de uitgeverwereld de ISO-standaard SGML voor documentproductie lijkt te gaan aanvaarden.

7 T_EX en SGML

Het idee van gestructureerde documentopmaak, waarop onder andere, maar niet uitsluitend, L^AT_EX is gebaseerd, is in de meest verregaande vorm uitgewerkt in de door de standaardisatie-organisatie ISO vastgelegde taal SGML. SGML wordt gebruikt of zal gebruikt gaan worden door de Amerikaanse ministeries van Defensie en Economie, de Amerikaanse uitgever, de Europese Gemeenschap, het onderzoeksinstituut CERN in Genève en een aantal Europese

uitgevers, waaronder *Elsevier*. Het laat zich aanzien dat SGML in de komende decennia een steeds grotere rol zal gaan spelen en dat langzaam maar zeker op SGML gebaseerde systemen de huidige WYSIWYG-systemen zullen gaan overvleugelen of dat op SGML gebaseerde systemen met een WYSIWYG-karakter op de markt zullen worden gebracht.

Ook binnen de gemeenschap van $\text{T}_{\text{E}}\text{X}$ - en $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$ -gebruikers is er over SGML nagedacht. Het basisidee van SGML is gegeneraliseerde, dat wil zeggen structuurbeschrijvende, in plaats van uitvoergerichte, zogenaamde descriptieve, ‘markup’. In verschillende opzichten is dit idee al gerealiseerd in $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$, of realiseerbaar in $\text{T}_{\text{E}}\text{X}$. Doordat $\text{T}_{\text{E}}\text{X}$ volkomen programmeerbaar is, kunnen macro’s geschreven worden die bijna als directe vertaling van SGML-instructies kunnen fungeren. $\text{T}_{\text{E}}\text{X}$ lijkt hierdoor het perfecte ‘back end’ voor een op SGML gebaseerd tekstopmaakstelsel.

8 Besluit

Onze ervaring is dat $\text{T}_{\text{E}}\text{X}$ een flexibel en uitstekend bruikbaar systeem is. $\text{T}_{\text{E}}\text{X}$ heeft de volgende voordelen:

- $\text{T}_{\text{E}}\text{X}$ levert zetwerk van superieure kwaliteit.
- Het kan geprogrammeerd worden zodat een vooraf vastgelegde opmaak exact wordt gerepro-

duceerd.

- Het kan worden gebruikt voor publicaties in niet-Westeuropese schriften of met wiskundige formules.
- $\text{T}_{\text{E}}\text{X}$ werkt op elk type computer hetzelfde. Het is een *de facto* standaard en het is beschreven in een boek dat in iedere grote boekhandel verkrijgbaar is.
- Door het systeem zo in te richten dat het de mogelijkheid biedt tot structuurbeschrijvende ‘markup’, kunnen auteurs zich volledig toeleggen op de inhoud van hun publicaties en kunnen ze publicaties meer dan een keer, met verschillende opmaak gebruiken.
- Omdat de opmaak-instructies die in $\text{T}_{\text{E}}\text{X}$ -documenten worden gebruikt in letters worden uitgeschreven is het mogelijk volstrekt probleemloos dergelijke documenten uit te wisselen tussen computers over de hele wereld en dus om publicaties of delen daarvan op te sturen aan vakgenoten en zelfs aan uitgevers.

Hoewel $\text{T}_{\text{E}}\text{X}$ in de academische wereld reeds lang ingeburgerd is, moet het daarbuiten als een belofte voor de toekomst gezien worden. $\text{T}_{\text{E}}\text{X}$ is echter door zijn programmeerbaarheid zowel rechtstreeks, dat wil zeggen als opmaakprogramma, als indirect, als uitvoercomponent van een groter systeem, te gebruiken. De belofte is dus groot.