

# Molecuul Muis Manuscript

## Verslag KNCV symposium

Gerard van Nes

18 oktober 1991

### Abstract

Op 18 oktober 1991 werd door de sectie Computertoepassingen van de KNCV (Koninklijke Nederlandse Chemische Vereniging) een symposium m.b.t. chemische tekstverwerking georganiseerd. Naast een algemene introductie over 'Electronisch Publiceren', kregen zowel de Apple Macintosh als MS-DOS geïnteresseerden (chemici) de huidige mogelijkheden van het verwerken van chemische teksten voorgeschoteld. Het pakket  $\text{T}_{\text{E}}\text{X}$  kwam in een laatste lezing naar voren. De duidelijk geslaagde dag werd bezocht door een kleine honderd deelnemers, inclusief een tiental leveranciers. Vele (leerzame) demonstraties, zowel tijdens de lezingen, als ook tussen de lezingensessies door, zorgden mede voor een duidelijk overzicht.

## 1 Algemeen

*Chemische* tekstverwerking is eigenlijk een wereld apart. Vooral de grote verscheidenheid aan disciplines binnen dit vakgebied zorgt ervoor dat de verwerking van chemisch tekstmateriaal minstens zo complex is als de verwerking van wiskundige teksten. Eenvoudige formules als  $\text{H}_2\text{O}$  bevinden zich in dit gebied aan de onderkant, terwijl het andere uiterste gekenmerkt wordt door ondermeer structuren van bijvoorbeeld biochemische verbindingen en complexe organische reactiemechanismen.

Gebrek aan een goed gedefinieerde gebruikersinterface is wel het grootste struikelpunt. Chemische formules kunnen enerzijds via een separaat, al dan niet 'dedicated', tekenpakket worden aangemaakt om vervolgens in teksten opgenomen te worden. Aan de andere kant leveren geïntegreerde pakketten de moeilijkheid *hoe* een gebruiker de formules in een tekst moet opgeven. Een vorm van standaardisatie is helaas nog ver te zoeken.

Toch gaven de afzonderlijke lezingen<sup>1</sup> een duidelijk beeld van de huidige stand van zaken, inclusief de mogelijkheden en onmogelijkheden. Alleen de lezing over  $\text{T}_{\text{E}}\text{X}$  bleek, vooral vanwege de persoonlijke voorkeuren van de spreker naast zijn gebrek aan (chemische)  $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}/\text{T}_{\text{E}}\text{X}$  tekstverwerkingskennis in het algemeen van een niet optimaal niveau.

## 2 Electronisch publiceren

Een uitstekende keuze van de organisatoren voor wat betreft onderwerp en spreker (P.W.J.M. Boumans) betrof de eerste bijdrage over 'Electronisch Publiceren'. Ondanks het feit dat de lezing nauwelijks iets met *chemische* tekstverwerking te doen had, kwamen toch zaken uitgebreid aan bod die direct te maken hadden met de pipeline 'interactief artikel schrijven' t/m 'het uiteindelijk uitgeven, inclusief de mogelijkheden erna'. Een nieuwe mogelijkheid van electronisch publiceren kwam ter sprake in de vorm van 'Spectrochimica Acta' (waarvan de spreker de 'Editor-in-Chief' is). Deze tijdschriftenreeks heeft nu een echte electronische lijn, t.w. 'Spectrochimica Acta Electronica', welke ingaat op een toekomstige behoefte: het uitgeven van, naast de hardcopy tekst, een diskette welke ondermeer de tekst, graphics, bijbehorende executable programma's, databestanden, parametersets en source code bevat.

Uitgebreid werd door de uiterst boeiende spreker gesproken over de plaats van 'markup'-talen, t.o.v. WYSIWYG en DTP systemen (welke laatste, aldus de spreker, binnen de wetenschappelijke wereld feitelijk al op zijn retour zijn). Ter sprake kwam tevens de plaats van SGML,  $\text{T}_{\text{E}}\text{X}$  en  $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$ , en de APEM (Auhor's Primary Electronic Manuscript) benadering.

De lezing was een korte uiteenzetting van hetgeen de spreker eerder in een 'standaard werk' had gepubliceerd [2]. Dit laatste is zeker het lezen waard!

<sup>1</sup>De teksten van alle lezingen waren opgenomen in de tijdens de dag beschikbare proceedings; prima verzorgd door de organisatoren!

### 3 Apple Macintosh omgeving

Tekstverwerking van chemische documenten binnen deze omgeving werd uiteengezet door J. Murre. De mogelijkheden lijken veel op die van de MS-DOS omgeving zoals in de volgende sectie ter sprake komt. Alleen is vanwege de volledige grafische interface van een Apple Macintosh (of te wel de clipboard mogelijkheden), *geen* geïntegreerd tekstverwerkings/teken pakket beschikbaar.

Los van het (algemene) tekstverwerkingsprogramma, zijn er veel *niet-chemische* tekenprogramma's beschikbaar. Voor het direct maken van chemische formules niet altijd geschikt (lees: dus ongeschikt).

Speciale chemische tekenprogramma's die ter sprake kwamen waren ChemDraw, Chemintosh II, ISIS/Draw, Organic Fonts (gebruikt een speciale font set; als font aan te roepen).

De mogelijkheden van deze producten zijn dikwijls zeer groot, locale intelligentie is aanwezig ten aanzien van 'niet-bestaande' formules, en de kwaliteit is zeer professioneel te noemen, naast de eenvoud in het gebruik, zeker in vergelijking met de hierna beschreven tekenpakketten voor MS-DOS systemen<sup>2</sup>. Niet vreemd eigenlijk voor een Macintosh.

Probleem is echter, en dat geldt ook bij de DOS systemen, dat er geen 'uitwisselbare' (lees: 'zichtbare') markup is gebruikt voor de chemische formules. Het heeft voordelen, doch zeker ook duidelijk veel nadelen!

Algemene gedachten: goede software, leuke hardware (zullen Macintosh bezitters zeker beamen<sup>3</sup>), prima voor locale documenten. Geen standaardisatie. Uitwisseling zeer beperkt, zo niet onmogelijk.

### 4 MS-DOS omgeving

Een uiteenzetting van chemische tekstverwerking binnen deze omgeving werd gegeven door de heren J.F.M. Bie & W.J.G. Schielen, beiden van het laboratorium Organische Chemie van de Katholieke Universiteit Nijmegen. Ook hier kwam weer de splitsing ter sprake van:

#### 1. Separate tekstverwerking met een puur chemisch tekenpakket.

Als tekstverwerking kan bijvoorbeeld WordPerfect, Word for Windows, Ventura Publisher, Pagemaker (of ook  $\text{\TeX}$ ) gebruikt worden. Het tekenpakket kan zijn 'geheel algemeen' (Autocad, Easycad, Harvard-Graphics, SlideWrite of Drawperfect). Niet direct geschikt dus voor chemische formules, veel tijd consumerend, onnauwkeurig, beperkt in gebruik etc. Niet fraai dus.

Als tekenmodule kan ook een speciaal chemisch te-

kenpakket genomen worden met als voorbeeld PLT, PsiGen, ChemWindow en Wimp-2001, de laatste twee onder MS-Windows. Alle kennen vanzelfsprekend een set van templates. Bekende eigenschappen zijn: het veel heen en weer switchen tussen tekstverwerking en tekenpakket, transport via HPGL, (E)PS, of CGM file, moeilijk om correcties aan te brengen, soms geen optimale (terug)koppeling etc. Min of meer omslachtig dus; maar het werkt en het biedt voor veel gevallen een oplossing.

#### 2. Integratie tekstverwerking en tekenpakket.

Eerst de eenvoudige programma's (zowel qua tekst als tekenfunctie). Chiwriter, Spellbinder Scientific, Total Word en T-3 kunnen bijvoorbeeld via een speciaal karakterset / 'chemisch font' (niet al te ingewikkelde) chemische formules weergeven. Wel zeer veel beperkingen, zowel wat betreft de chemische formules als zeker voor wat betreft de tekstverwerking zelf.

Chemtext gaat in mogelijkheden verder, daar het een intern chemisch tekenpakket in zich heeft. Zeer goede tekenmogelijkheden, echter op het gebied van tekstverwerking zijn er toch nog de nodige beperkingen.<sup>4</sup>

Conclusie DOS omgeving: Het blijft behelpen, zeker als men in eerste instantie hoge eisen stelt aan de tekstverwerking zelf. Algemene beperking is ook dat men niet met een *logische markup* werkt, hetgeen de uitwisseling van chemische documenten en het hergebruik binnen andere software modules zeker niet ten goede komt.

### 5 $\text{\TeX}$ presentatie

Wél het enige produkt met een '*visible markup*'. Kenmerkend dus voor onderhoudbaarheid en duidelijke logische structuur, waarbij men gebruik kan maken van de eigen favoriete editor op PC t/m supercomputer. De (logische) markup eigenschappen kwamen in de lezing van P.E.S. Wormer (Theoretische Chemie, Katholieke Universiteit Nijmegen) echter niet naar voren. Hetzelfde gold ook voor wat betreft de  $\text{\TeX}$  chemische formule mogelijkheid binnen gebieden als anorganische en organische chemie.

De spreker bleek een oude (één van de vroege) gebruikers van  $\text{\TeX}$  te zijn. Hij schroomde ook niet om naast enkele mogelijkheden, vooral de moeilijkheden bij het gebruik van  $\text{\TeX}$  'overdreven' naar voren te brengen. Veel niet terzake betreffende onderwerpen werden uiteengezet ('Knuth aanhangers zijn soms te vergelijken met sekteleiden'<sup>5</sup>).

$\text{\LaTeX}$  zelf had hij echter nooit gebruikt daar hij met  $\text{\TeX}$

<sup>2</sup>Dat was mijn persoonlijke indruk na diverse demonstraties gezien te hebben.

<sup>3</sup>Voor de duidelijkheid: ik bezit geen Apple Macintosh; ik heb er zelfs nooit mee gewerkt; helaas misschien.

<sup>4</sup>Sorry, ik als  $\text{\LaTeX}$  gebruiker ben wat gewend (of verwend??).

<sup>5</sup>Trouwens, en dit komt van mijzelf (GvN), ik las onlangs dat achter de WordPerfect bron de Mormonen uit Salt Lake City schuilgingen. Maar dit is vanzelfsprekend niet ter zake.

alles kon doen. De filosofie achter het  $\text{T}_{\text{E}}\text{X}/\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$  gebruik werd ook niet genoemd.<sup>6</sup> Daarnaast werden door de spreker onjuistheden weergegeven ('...Het nadeel van  $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$  op MS-DOS computers is dat de bibliotheek maar net in 640 Kb past...', etc.). Daar  $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$  zelf niet in het geheugen van zijn PC paste, had hij het ook nooit gebruikt (...). Ik begreep dat hij een gebruiker van  $\text{e}\text{T}_{\text{E}}\text{X}$  was. Een raadsel dus.

Puur chemische tekstverwerking werd slechts getoond met een (standaard) voorbeeld uit de  $\text{CHEMSTRUCT}$  macro [5] en het noemen van de  $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$ -macro's van Haas & O' Kane [3]. Enige eigen ervaring ontbrak echter bij hem.

Daarnaast kwam ook hier weer naar voren het genereren van chemische formules via een apart tekenpakket (AutoCAD werd daarbij genoemd, en een niet-chemisch voorbeeld ervan getoond...). Duidelijk was het dat, mede vanwege zijn chemische studierichting, de spreker nauwelijks tot geen kennis had van *chemische* tekstverwerking.

Jammer eigenlijk deze presentatie. Met een beetje verder zoeken door de organisatoren van dit symposium had (zeker onder de NTG leden) een betere (serieuze én beter bekend met de materie) en objectieve spreker gevonden kunnen worden.

## 6 Chemische $\text{T}_{\text{E}}\text{X}$ verwerking

Voor wat meer 'volledigheid' en ter verdere studie is het toch zinvol iets te zeggen over de huidige stand van zaken (voor zover mij bekend) op het gebied van  $\text{T}_{\text{E}}\text{X}$  en chemische tekstverwerking'.

Drie macro pakketten zijn in omloop:

1. Macro's van Roswitha Haas [3]  
Te vinden op enkele ftp sites (inclusief de RUU-server) en op de  $\text{T}_{\text{E}}\text{X}$ -SUN distributie tape. Uitgebreide documentatie is beschikbaar. Te gebruiken vanuit zowel  $\text{T}_{\text{E}}\text{X}$  als  $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$  (gebruikt picture-environment).
2. Macro's van Michael Ramek [5]  
Te vinden op enkele ftp sites. Documentatie geschreven door H. Partl. Ontworpen om gebruikt te worden vanuit de plain  $\text{T}_{\text{E}}\text{X}$  omgeving.
3. Macro's van Maurice Laugier [4]  
Te gebruiken in een  $\text{T}_{\text{E}}\text{X}$  omgeving.

Nog niet tegengekomen op een ftp site of een installatie tape (wie wel?). Aan te vragen via [louijean@frgren81](mailto:louijean@frgren81).

Elke set macro's kent zijn eigen goede en minder goede kanten. Ze zijn ontworpen vanuit een organisch chemisch gezichtspunt (tamelijk algemeen dus). Een vergelijking van de diverse macro sets vóór het uiteindelijke gebruik is dus nodig.

Ondanks de dikwijls goede kwaliteit, geldt echter als algemene beperking het gemis van een (gebruiksvriendelijke) gebruikersinterface<sup>7</sup>. Enige oefening en het maken van noodzakelijke aanvullende (eigen) macro's is dikwijls een vereiste om tot een goede oplossing te komen.

## 7 Conclusie

Een door de KNCV organisatoren uitstekend georganiseerde dag met veel informatie uitwisseling, boeiende demonstraties, een goede set van lezingen en een directe beschikbaarheid van de symposiumlezingen op papier [1]. Alleen jammer dat de  $\text{T}_{\text{E}}\text{X}$  lezing hiervan afweek.

Een gestandaardiseerde en geïntegreerde tekst- en chemisch 'teken'pakket is duidelijk missende. Wie voelt zich geroepen?

## References

- [1] *Molecuul Muis Manuscript*, KNCV Symposiareeks #3, 51 pagina's, (1991).
- [2] P.J.M. Boumans, *The Dissemination of the Results of Scientific Research in the Era of Electronic Media*, Spectrochimica Acta, Special Supplement, 11-45, (1989).
- [3] R.T. Haas & K.C. O'Kane, *Typesetting Chemical Structure Formulas with the Text Formatter  $\text{T}_{\text{E}}\text{X}/\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$* , Comput. Chem, 11.4, 251-271, (1987).
- [4] M. Laugier, *Composition des formules chimiques en  $\text{T}_{\text{E}}\text{X}$* , Cahiers GUTenberg #10,11, 209-216, (1991).
- [5] M. Ramek, *CHEMSTRUCT, Chemische Strukturformeln mit  $\text{T}_{\text{E}}\text{X}$* , Handbuch-Nummer X14, 1-18, (1988).

<sup>6</sup>Degenen die mij een beetje kennen weten dat ik gebruikers ook zoveel mogelijk aanraad om vooral  $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$  te gebruiken i.p.v.  $\text{T}_{\text{E}}\text{X}$ . Alleen als het niet anders kan moet m.i. naar  $\text{T}_{\text{E}}\text{X}$  worden uitgeweken, en dan nog bij voorkeur slechts via ( $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$  logische) macro's. Het gebruik van een logische markup staat bij mij nu eenmaal hoog in het vaandel.

<sup>7</sup>Wie staat te trappelen van ongeduld om daar wat aan te doen?