

Die T_EXnische Komödie

dante

Deutschsprachige
Anwendervereinigung T_EX e.V.

20. Jahrgang Heft 1/2008 Januar 2008

1/2008

Impressum

»Die T_EXnische Komödie« ist die Mitgliedszeitschrift von DANTE e.V. Der Bezugspreis ist im Mitgliedsbeitrag enthalten. Namentlich gekennzeichnete Beiträge geben die Meinung der Schreibenden wieder. Reproduktion oder Nutzung der erschienenen Beiträge durch konventionelle, elektronische oder beliebige andere Verfahren ist nur im nicht-kommerziellen Rahmen gestattet. Verwendungen in größerem Umfang bitte zur Information bei DANTE e.V. melden.

Beiträge sollten in Standard-L^AT_EX-Quellcode unter Verwendung der Dokumentenklasse dtk erstellt und per E-Mail oder Datenträger (CD) an untenstehende Adresse der Redaktion geschickt werden. Sind spezielle Makros, L^AT_EX-Pakete oder Schriften dafür nötig, so müssen auch diese komplett mitgeliefert werden. Außerdem müssen sie auf Anfrage Interessierten zugänglich gemacht werden.

Diese Ausgabe wurde mit pdfT_EX 3.141592-1.40.5-2.2 (Web2C 7.5.6) erstellt. Als Standard-Schriften kamen die Type-1-Fonts Latin-Modern und LuxiMono zum Einsatz.

Erscheinungsweise: vierteljährlich

Erscheinungsort: Heidelberg

Auflage: 2700

Herausgeber: DANTE, Deutschsprachige Anwendervereinigung T_EX e.V.
Postfach 10 18 40
69008 Heidelberg

E-Mail: dante@dante.de
dtkred@dante.de (Redaktion)

Druck: Konrad Tritsch Print und digitale Medien GmbH
Johannes-Gutenberg-Str. 1-3, 97199 Ochsenfurt-Hohe Stadt

Redaktion: Herbert Voß (verantwortlicher Redakteur)

Mitarbeit : Rudolf Herrmann Lutz Ihlenburg Bernd Raichle
Christine Römer Volker RW Schaa Hilmar Schlegel
Uwe Ziegenhagen

Redaktionsschluss für Heft 2/2008: 15. April 2008

ISSN 1434-5897

Editorial

Liebe Leserinnen und Leser,

beim letzten Mal habe ich bereits darauf hingewiesen, dass wir mit dieser Ausgabe in das 20. Erscheinungsjahr gehen. Dass in diesem langen Zeitraum alle Ausgaben immer pünktlich erscheinen konnten, lag nicht nur an der Arbeit der Redaktion, sondern auch an den vielen Autoren, die mit ihren Beiträgen für den Inhalt unserer Vereinszeitschrift sorgen. Es gab und gibt immer wieder schwierige Zeiten, in denen zu wenig Manuskripte vorliegen, weshalb wiederholt die Bitte an alle Mitglieder ergeht, uns auch sehr kurze Beiträge aus dem Alltag eines T_EX-Entwicklers oder T_EX-Anwenders zukommen zu lassen.

Diese Kurzausgabe beschäftigt sich vorrangig mit dem Thema »Trennregeln und Trennmuster«, welches für viele T_EX-Anwender nur bedingt von Interesse ist, da diese Thematik selten Anlass für Diskussionen bietet. Dies liegt aber nicht in der Natur der Sache, sondern an dem von T_EX verwendeten Trennalgorithmus und der verwendeten Trennmuster- und Ausnahmeregeldatei. Dadurch sind Trennfehler sehr selten geworden, mit dem Ergebnis, dass das Interesse der T_EX-Anwender für den Trennalgorithmus gering ist. Durch die Einführung der neuen deutschen Rechtschreibung hat sich aber auch im Bereich der Trennregeln einiges geändert. Dies war Anlass, einmal grundsätzlich über das Thema nachzudenken.

Aufgrund der Kurzausgabe dieser Komödie finden Sie die Liste der neuen CTAN-Pakete erst wieder in der nächsten Ausgabe.

Mit T_EXnischen Grüßen,

Ihr Herbert Voß

Hinter der Bühne

Vereinsinternes

Grußwort

Liebe Mitglieder,

wir hoffen, dass Sie das neue Jahr 2008 gut begonnen haben und dass all Ihre Wünsche in Erfüllung gehen mögen.

Dies ist eine besondere Ausgabe von »Die $\text{T}_{\text{E}}\text{X}$ nische Komödie«. Nachdem der Versand der gesammelten Konferenzbeiträge der Euro $\text{T}_{\text{E}}\text{X}$ 2005 viel Anklang fand, haben wir beschlossen, die *Proceedings* der vergangenen EuroBach $\text{T}_{\text{E}}\text{X}$ 2007 ebenfalls an alle Mitglieder zu verschicken. Der Proceedings-Band gilt als normale Ausgabe unserer Zeitschrift, mit einem separaten vereinspezifischen Teil, den sie gerade in Händen halten.

Hier finden Sie noch einmal die Einladung zur kommenden $\text{T}_{\text{E}}\text{X}$ -Tagung DANTE 2008 in Jena. Wir hoffen, dass viele Mitglieder an der ersten DANTE-Tagung in Thüringen teilnehmen werden. Insbesondere würden wir uns über eine zahlreiche Teilnahme an der Mitgliederversammlung freuen, in der neben den üblichen Berichten über Vorstandsarbeit, Finanzen und Kassenprüfung der Vorstand für die nächste Amtsperiode gewählt wird.

Falls Sie kombiniertes Mitglied von DANTE e.V. und der internationalen $\text{T}_{\text{E}}\text{X}$ Users Group TUG sind, beachten Sie bitte, dass der Konferenzband ebenfalls eine Ausgabe des TUGboat ersetzt und Sie für das Jahr 2008 daher nur zwei Ausgaben des TUGboat erhalten.

Mit freundlichem Gruß,

Klaus Höppner Volker RW Schaa
Vorsitzender Stellvertretender Vorsitzender

T_EX-Theatertage

DANTE 2008 – Einladung zur Mitgliederversammlung

Christine Römer, Klaus Höppner

Liebe Mitglieder von DANTE e.V.,

die T_EX-Tagung DANTE 2008 findet vom 5. bis zum 7. März 2008 an der Friedrich-Schiller-Universität Jena statt. Veranstalter sind DANTE e.V. und das Institut für Germanistische Sprachwissenschaft der Friedrich-Schiller-Universität Jena.

Am Mittwoch sind Tutorien geplant, Donnerstag und Freitag sind für Vorträge und die 38. Mitgliederversammlung von DANTE e.V. vorgesehen. Die Tagesordnung der Mitgliederversammlung am Donnerstag, den 6. März 2008 um 9.00 Uhr in der

Friedrich-Schiller-Universität Jena
Multimediazentrum I, Hörsaal E 028
Ernst-Abbe-Platz 8
07743 Jena

lautet:

1. Begrüßung und Tagesordnung
2. Bericht des Vorstands
3. Finanzbericht
4. Bericht der Kassenprüfer
5. Entlastung des Vorstands
6. Wahl des Vorstands
7. Wahl von Kassenprüfern
8. Verschiedenes

Ihre Stimmunterlagen erhalten Sie direkt vor Ort, um vorherige Anmeldung wird gebeten. Eine Übertragung des Stimmrechts ist im Rahmen des § 13 (4) der Vereinssatzung möglich.

Wie üblich sind auch Nichtmitglieder als Gäste willkommen.

Firmen und Institutionen, die ihre Produkte präsentieren oder die Tagung finanziell unterstützen wollen, werden gebeten, sich frühzeitig per E-Mail an dante2008@dante.de zu wenden.

Unter <http://www.dante.de/dante2008/> finden Sie die Homepage der Tagung mit allen weiteren Informationen. Hier finden Sie auch einige Hotels, in denen ein Zimmerkontingent für Teilnehmer der Tagung zu vergünstigten Preisen bereit steht. Wenn Sie eines dieser Zimmer buchen wollen, geben Sie bei der Buchung bitte das Code-Wort »DANTE« an. Bitte beachten Sie, dass die rabattierten Zimmerkontingente zeitlich befristet sind und danach wieder die normalen Zimmerpreise der Hotels gelten!

Mit Fragen, Wünschen und Anregungen wenden Sie sich bitte an

DANTE e.V.
Stichwort: DANTE 2008
Postfach 10 18 40
69008 Heidelberg
E-Mail: dante2008@dante.de

Mit freundlichen Grüßen,

Christine Römer (Friedrich-Schiller-Universität Jena)
Klaus Höppner (DANTE e.V.)

Bretter, die die Welt bedeuten

Einige Fragen zum Beitrag »Hyphenation Exception Log für deutsche Trennmuster, Version 1«

Stephan Hennig

Dieser Beitrag versucht einige Fragen aufzuwerfen, die die Beiträge von Werner Lemberg zum Thema Trennmuster und Ausnahmelisten in »Die T_EXnische Komödie« 2/2003 und 2/2005 bisher unbeantwortet ließen.

Rückblick

In »Die T_EXnische Komödie« 2/2003 rief Werner Lemberg dazu auf, Fehler in den Trennmustern für die deutsche Rechtschreibung zu sammeln und diese in einer öffentlichen Ausnahmeliste zusammenzufassen.[4] In der Ausgabe 2/2005 beschrieb er Methoden und Skripte, mit denen die Suche nach Fehlern in den Trennmustern für die traditionelle und neue deutsche Rechtschreibung vereinfacht werden kann und stellte eine erste Version der Ausnahmeliste für die traditionelle Rechtschreibung vor.[5]

Beide Beiträge sind interessant, jedoch bleiben meiner Meinung nach eine Reihe von Problemen unbeantwortet. Ich stelle meine Fragen hier in der Hoffnung, dass sie von allgemeinem Interesse sind und die eine oder andere Seite Antworten darauf geben kann. Auch wenn Werner Lembergs Beitrag als Aufhänger dient, richten sich die folgenden Fragen nicht nur an ihn, sondern an alle Mitglieder.

Fragenkatalog

1. Aus welchem Grund wurde eine Ausnahmeliste für die traditionelle Rechtschreibung vorgestellt und nicht für die neue?

2. Gibt es Erkenntnisse zur Qualität der Trennmuster für die deutsche Rechtschreibung, sowohl die traditionelle als auch die neue? Wie gut oder schlecht sind die existierenden Trennmuster?
3. Gibt es einen Bedarf für Neuberechnete Trennmuster?
4. Ist die Arbeit an Ausnahmelisten zur Zeit sinnvoll?
5. Welche Berührungspunkte gibt es mit anderen freien Programmen zur Textbearbeitung?
6. Wie kann eine Infrastruktur aussehen, die verteiltes Arbeiten an zu schaffenden Listen korrekter Worttrennungen für die traditionelle und neue Rechtschreibung ermöglicht?
7. Was sind gute Wortquellen?
8. Wie funktioniert die Worttrennung in T_EX?
9. Wie werden Trennmuster erstellt?
10. Welchen Einfluss haben ungewöhnliche Wörter auf die Trennmuster?
11. In welcher Form kann und möchte DANTE e.V. unterstützend bei der Pflege deutscher Trennmuster und Ausnahmelisten wirken?
12. Ist es zweckmäßig auch Trennmuster und Ausnahmelisten nur mit Haupttrennstellen zu erstellen?

Anmerkungen zu den Fragen

Es folgen Anmerkungen (keine Antworten) zu den aufgeworfenen Fragen:

Zu 1) Aus welchem Grund wurde eine Ausnahmeliste für die traditionelle Rechtschreibung vorgestellt und nicht für die neue?

Zur vorgestellten Ausnahmeliste für die traditionelle Rechtschreibung bemerkt Werner Lemberg in [5]:

Entfernt man alle Zeilen, die [...] »-st« enthalten, kann man die folgenden Korrekturen auch mit den neuen deutschen Trennmustern benützen ...

Dieses Vorgehen birgt Probleme. Zum Beispiel wird das Wort »Abendstern« mit `dehypht.tex`¹ korrekt »Abend-stern« getrennt. Mit `dehypfn.tex` wird es dagegen falsch »Abends-tern« getrennt. Zum einen wäre der Eintrag »Abendstern« in einer gemeinsamen Ausnahmeliste redundant, falls die traditionelle Rechtschreibung verwendet wird. Zum anderen würde er gestrichen, sobald

¹Die Datei `dehypht.tex` enthält die Trennmuster für die traditionelle, `dehypfn.tex` diejenigen für die neue deutsche Rechtschreibung. Die Trennmuster werden beim Laden der Pakete `german`, `ngerman` oder `babel` mit den entsprechenden Optionen aktiviert.

nach der oben angegebenen Regel die Ausnahmeliste für die neue Rechtschreibung abgeleitet wird. Diese fehlerhafte Trennung ist so nicht korrigierbar. Aus technischer Sicht wäre es daher vorzuziehen, zwei getrennte Ausnahmelisten zu führen. Frage 1 könnte daher auch anders formuliert werden: Welche Ausnahmelisten (und Trennmuster) sollten gepflegt werden? Sprechen nichttechnische Gründe dafür, nur die einen oder die anderen zu pflegen?

Zu 2) Gibt es Erkenntnisse zur Qualität der Trennmuster für die deutsche Rechtschreibung, sowohl die traditionelle als auch die neue? Wie gut oder schlecht sind die existierenden Trennmuster?

Werner Lemberg hat in [5] einige Beobachtungen zu den Trennmustern für die traditionelle Rechtschreibung zusammengestellt. Meine Meinung zu den existierenden Trennmustern ist: Sie sind akzeptabel. Das heißt, L^AT_EX trennt nicht auffallend besser oder schlechter als andere Anwendungen. Eine manuelle Kontrolle der Trennungen ist in jedem Fall notwendig, einige Korrekturen sind es meistens auch. Systematische Betrachtungen zu den Trennmustern sind mir allerdings nicht bekannt. Existieren solche?

Zu 3) Gibt es einen Bedarf für Neuberechnete Trennmuster?

Bei der Beantwortung dieser Frage spielen meiner Meinung nach mehrere Punkte eine Rolle:

1. Ausnahmelisten sind unbestritten sinnvoll, da in der deutschen Sprache beliebige Wortzusammensetzungen und somit Buchstabenkombinationen gebildet werden können. Dadurch verlieren die Buchstabenmuster, an denen sich der Trennalgorithmus orientiert, an Signifikanz und fehlerhafte Worttrennungen sind unvermeidlich. Die Trennalgorithmen sollten im Idealfall jedoch auf Komposita beschränkt bleiben. Werner Lembergs Ausnahmeliste `dehyphtex.tex` (Stand: 28. 1. 2006) enthält 2090 fehlerhafte und fehlende Trennungen *einfacher* Wörter. Ab welcher Zahl besteht ein Bedarf für neue Trennmuster?
2. In [5] berichtet Werner Lemberg, dass die originale Wortliste »verschollen« ist, aus der die Trennmuster für die traditionelle Rechtschreibung erzeugt wurden. Die Trennmuster können daher nicht jederzeit wieder neu erstellt werden. Für eine freie Software wie T_EX ist dies kein wünschenswerter Zustand.
3. Die Trennmuster für die neue Rechtschreibung `dehypfn.tex` wurden (aus diesem Grund) nicht aus einer Wortliste berechnet. Walter Schmidt hat

in Handarbeit die Datei `dehyphn.tex` an die modifizierte s-t-Trennregel, vermehrte Doppel-s-Schreibung und weitere Änderungen angepasst. So ist auch die unterschiedliche Trennung von »Abendstern« zu erklären. Allerdings wird auch das einfache Wort »Fassade« mit `dehypht.tex` korrekt getrennt, mit `dehyphn.tex` (Stand: 7. 5. 2001, Revisionlevel 31) jedoch die erste Trennstelle nicht erkannt. Können solche Fehler durch Neuberechnete Trennmuster für die neue Rechtschreibung vermieden oder verringert werden?

Zu 4) Ist die Arbeit an Ausnahmelisten zur Zeit sinnvoll?

Falls Frage 3 positiv beantwortet werden kann, stellt sich die Frage, welchen Nutzen Ausnahmelisten zum jetzigen Zeitpunkt bringen. Durch mühsame Arbeit können damit zwar einige fehlerhafte Trennungen korrigiert werden, systematische Fehler in den Trennmustern jedoch nicht. Dies gilt umso mehr, da die originale Wortliste nicht mehr zur Verfügung steht und deren Qualität daher schlecht eingeschätzt werden kann.

Wäre es möglicherweise sinnvoller, sich zunächst auf das Erstellen neuer Trennmuster zu konzentrieren? Dabei sollten die bekannten Trennfehler einfacher Wörter selbstverständlich in die Erstellung einbezogen werden. Allerdings wird für das Erstellen neuer Trennmuster weit mehr benötigt als eine Liste einiger derzeit falsch getrennter Wörter. Nämlich eine qualitativ möglichst hochwertige Liste getrennter deutscher Wörter. Zwei Qualitätsmerkmale der Liste sollten nach Werner Lemberg der Listenumfang an einfachen und zusammengesetzten Wörtern und die Fehlerfreiheit sein.[5]

Im Gegensatz zu früher ist es wohl nicht so dringend notwendig, die Trennmuster klein zu halten [...] sollte eher Augenmerk auf wirklich fehlerfreie Trennung einer möglichst großen Anzahl von einfachen und zusammengesetzten Wörtern gelegt werden.

Zu den Fehlern in der vorliegenden Wortliste urteilt Werner Lemberg in [5]:

Mit großer Wahrscheinlichkeit habe ich etliche Einträge übersehen oder falsch behandelt.

Sofern in die *Ausnahmeliste* nur fehlerfreie Einträge aufgenommen werden, sind unentdeckte Trennfehler unproblematisch, da sie keinen zusätzlichen Schaden anrichten. Sollen jedoch *Trennmuster* erzeugt werden, sind Fehler in der zugrundeliegenden Wortliste unbedingt zu vermeiden. Der nächste Schritt wäre meiner Meinung nach, auf Werner Lembergs Arbeit aufzubauen und

eine zweite, dritte, vierte und x-te Kontrolle der existierenden Liste(n) durch »andere Augen« durchzuführen. Daran können und sollten sich möglichst viele beteiligen. Die dazu erforderliche Infrastruktur wird in Frage 6 thematisiert.

Zu 5) Welche Berührungspunkte gibt es mit anderen freien Programmen zur Textbearbeitung?

Neben $\text{T}_{\text{E}}\text{X}$ gibt es eine ganze Reihe von freien Programmen zum Verfassen und Bearbeiten von Texten. Wie werden die Probleme der Rechtschreibkontrolle und Worttrennung dort gelöst? Können die existierenden Wortlisten für die Nutzung mit $\text{T}_{\text{E}}\text{X}$ aufbereitet werden? Sind dort Trennungen enthalten? Gibt es lizenzrechtliche Probleme? Beispielhaft seien Programme wie `OpenOffice`, `Aspell` oder `SiSiSi` erwähnt.

Zu 8) Wie funktioniert die Worttrennung in $\text{T}_{\text{E}}\text{X}$?

Die folgenden beiden Fragen betreffen die grundlegende Arbeitsweise des Programms $\text{T}_{\text{E}}\text{X}$: Absatzumbruch und Worttrennungen sind in $\text{T}_{\text{E}}\text{X}$ eng miteinander verwoben. Der Schwerpunkt dieser Frage soll darauf liegen, wie aus den Trennmustern mögliche Trennstellen ermittelt werden.

Zu 9) Wie werden Trennmuster erstellt?

Hier geht es um das umgekehrte Problem, wie werden aus Listen mit Worttrennungen Trennmuster abgeleitet?

Die letzten beiden Fragen finde ich auch abseits der hier behandelten Problematik interessant. Ich wundere mich, wie wenige Beiträge in »Die $\text{T}_{\text{E}}\text{X}$ nische Komödie« die Grundlagen von $\text{T}_{\text{E}}\text{X}$ und $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$ beleuchten. Gehört »The $\text{T}_{\text{E}}\text{X}$ book« [3] zum Allgemeinwissen eines jeden $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$ -Anwenders? Ich kann von mir leider nicht behaupten, dieses Buch vollständig gelesen und verstanden zu haben.

Zu 10) Welchen Einfluss haben ungewöhnliche Wörter auf die Trennmuster?

Gibt die Antwort auf Frage 9 Hinweise darauf, ob es Wörter gibt, die *nicht* in Trennmuster aufgenommen werden sollten, da sie deren Qualität (Signifikanz und Trennschärfe bestimmter Buchstabenmuster) verringern oder die Trennmuster deutlich vergrößern würden? Welche Wörter sind dies? Komposita, wie zum Beispiel »Stabs-chef«, mit einer Trennung innerhalb der sonst untrennbaren Buchstabenkombination »sch« wurden bereits erwähnt. Gibt es Wörter, die grundsätzlich besser durch Ausnahmelisten abgedeckt werden (Fremdwörter, Eigennamen, geografische Bezeichnungen)?

Zu 11) In welcher Form kann und möchte DANTE e.V. unterstützend bei der Pflege deutscher Trennmuster und Ausnahmelisten wirken?

Ansprechpartner zu den Trennmusterdateien sind zur Zeit Walter Schmidt für `dehyphn.tex` und Bernd Raichle bzw. »DANTE e.V., Koordinator `german.sty`« für `dehypht.tex`. Für die Ausnahmeliste `dehyphtex.tex` ist es Werner Lemberg. Ist DANTE e.V. in die Pflege institutionell eingebunden? Kann DANTE e.V. die Arbeit an Wortlisten unterstützen, zum Beispiel durch Bereitstellung von technischer Infrastruktur oder Projektmitteln? Da ich mir wenige Aufgaben vorstellen kann, die so unmittelbar im Fokus von DANTE e.V. stehen sollten wie die langfristige Pflege von deutschen Wortlisten, Trennmustern und Ausnahmelisten, erscheint mir eine Unterstützung der derzeit damit betrauten Personen durch den Verein sinnvoll (sofern dies nicht bereits der Fall ist).

Zu 12) Ist es zweckmäßig auch Trennmuster und Ausnahmelisten nur mit Haupttrennstellen zu erstellen?

Werner Lemberg hat in [4] verschiedene Ansätze erwähnt, die sich dem Problem der Unterscheidung zwischen Haupt- und Nebentrennstellen widmen. Die Ansätze von Clasen und Sojka benötigen dafür Trennmuster, die gewichtete oder nur Haupttrennstellen enthalten.[8, 1]

Auch ohne die beschriebenen Erweiterungen des Trennalgorithmus können Trennmuster sinnvoll sein, die nur Haupttrennstellen enthalten. Verwendet man in einem Dokument zum Beispiel eine ausreichende Zeilenlänge, so kann durch Trennungen ausschließlich an Haupttrennstellen die Lesbarkeit des Textes erhöht werden. Würden diese zusätzlichen Trennmuster als weitere Sprachen in Babel integriert, könnte man als Hauptsprache für ein Dokument zum Beispiel `ngerman-ht` wählen und nur für kritische Textstellen, für welche aus irgendeinem Grund mit diesen Trennmustern kein befriedigender Absatzumbruch gefunden werden kann, mit einem Sprachwechsel vorübergehend die herkömmlichen, umfangreicheren Trennmuster `ngerman` aktivieren.

Die Pflege dieser Trennmuster würde wohl weitgehend unabhängig von der der herkömmlichen Trennmuster erfolgen müssen. Zum Beispiel würde die Ausnahmeliste wahrscheinlich erheblich mehr Komposita enthalten müssen und diese mit anderen Trennstellen. Damit würde sich die Zahl der zu verwaltenden Dateien verdoppeln. Besteht ein Bedarf an Trennmustern und Ausnahmelisten nur mit Haupttrennstellen?

Anmerkungen zu Werner Lembergs Skripten

Abschließend möchte ich Änderungen an zwei der von Werner Lemberg vorgestellten Skripten vorschlagen: `strippunct.sed` und `prepare-wordlist.sh`. Außerdem wird eine neue Version des Skripts `log2words.sed` vorgestellt, welche auch von Werner Lemberg stammt.

Kurz zur Funktionsweise des bisherigen Skripts `strippunct.sed` (vgl. Listing 5 in [5]): In Zeile 1 wird eine Auswahl an Sonderzeichen entfernt, jedoch bleiben einige nichtalphabetische Zeichen erhalten. In den Zeilen 2 und 3 werden Wörter entfernt, die einen Bindestrich oder ein (falsches) Apostroph im Wortinnern, am Wortanfang oder -ende enthalten. Zeile 4 ersetzt schließlich Leerzeichen durch Zeilentrennzeichen, damit alle verbleibenden Wörter in eine neue Zeile geschrieben werden.

Eine Schwäche des bisherigen Skripts ist, dass nicht alle Sonderzeichen entfernt werden. Beispielsweise verbleiben mit Werner Lembergs Skript in der aus dem Freedict-Wörterbuch (siehe letzter Abschnitt) extrahierten Wortliste viele geklammerte Wörter. Mir fällt jedoch kein Grund ein, aus dem in der aufbereiteten Wortliste irgendwelche nichtalphabetischen Zeichen auftreten sollten. Das vorgeschlagene Skript erfüllt die folgenden Anforderungen:

1. Es werden alle Wörter entfernt, die im Wortinnern mindestens ein *beliebiges* nichtalphabetisches Zeichen enthalten.
2. Es werden wie bisher Wörter entfernt, die mit mindestens einem Bindestrich, (falschen) Apostroph und/oder einer Ziffern beginnen oder enden.
3. Es werden schließlich sämtliche verbliebenen nichtalphabetischen Zeichen entfernt, insbesondere auch Klammern und Ziffern.
4. Neben Leerzeichen gelten auch Tabulatoren als Worttrenner. Aufeinanderfolgende Worttrenner werden durch ein Zeilentrennzeichen ersetzt.

Anforderung 1 erfasst wie bisher den »Wochenend-Einkaufszettel«, zusätzlich aber auch Abkürzungen, Formeln oder Gebilde wie »a. a. O.«, »H₂O« oder »Z*****t«. Durch die Anforderung 2 werden Wortrümpfe wie »Bildungs-« oder »3fach« entfernt. Die Zahl der durch Anforderungen 1 und 2 entfernten *brauchbaren* Wörter sollte gering sein, etwa durch Schrägstrich getrennte Wörter, wie »schwarz/weiß«. Hauptsächlich dürfte es sich um Komposita handeln.

Ausdruck	Beschreibung
<code>[:alpha:]</code>	Posix-Zeichenklasse: alle alphabetischen Zeichen der Landessprache einschließlich akzentuierte Buchstaben
<code>[:space:]</code>	Posix-Zeichenklasse: Leerzeichen, Tabulator, Zeilen- und Seitentrenner
<code>[:digit:]</code>	Posix-Zeichenklasse: Ziffern
<code>[:punct:]</code>	Posix-Zeichenklasse: Interpunktionszeichen
<code>[[:alpha:]]</code>	passt auf ein beliebiges alphabetisches Zeichen
<code>[^[:alpha:]]</code>	passt auf ein beliebiges nichtalphabetisches Zeichen
<code>*</code>	Quantor: »nicht oder beliebig oft«
<code>\{1,\}</code>	Quantor: »mindestens einmal«

Tabelle 1: Einige reguläre Ausdrücke.

Im Vergleich zum bisherigen Skript `strippunct.sed` wurde auch die Reihenfolge der Aktionen geändert. Da Anforderung 3 auch die in Anforderungen 1 und 2 verwendeten Interpunktionszeichen einschließt, wurde diese Aktion nach hinten verschoben. Die bereits erwähnten Beispiele würden sonst zu früh zu »WochenendEinkaufszettel«, »aaO« usw. schrumpfen und könnten nicht mehr entfernt werden.

Außerdem werden im vorgeschlagenen Skript statt expliziter Zeichenmengen weitgehend Posix-Zeichenklassen verwendet (vgl. Tabelle 1). Grundsätzlich würde der Suchausdruck `[[:punct:][:digit:]]` Anforderung 3 genügen. Ich habe jedoch den theoretisch allgemeineren Ausdruck `[^[:alpha:][:space:]]` verwendet, der auf alles außer alphabetische Zeichen und *Whitespace* passt. Der Quantor `\{1,\}` entspricht übrigens dem `+` in erweiterten regulären Ausdrücken. Dieser Ausdruck wurde hier jedoch umschrieben, da `+` in Posix-konformen Programmen mitunter nicht zur Verfügung steht.

Das überarbeitete Skript `strippunct.sed` ist in Listing 1 zu sehen. Der Suchausdruck in Zeile 1 könnte zwar auch etwas kürzer formuliert werden, ich habe allerdings diesen Suchausdruck gewählt, da er symmetrisch ist und in dieser Form leichter verständlich sein sollte. Die Zeilen 2 bis 4 sind stark an das ursprüngliche Skript angelehnt.²

²Zeile 2 erledigt mehr als Anforderung 2. Es werden sämtliche Bindestriche, Apostrophe und Ziffern entfernt. Wichtig ist jedoch, dass nicht nur diese Zeichen entfernt werden, sondern gegebenenfalls auch unmittelbar führende und folgende Wortteile.

Listing 1: `strippunct.sed`, Version 2.0.

```

1 s/[[:alpha:]]^[[:space:]]*^[[:alpha:]][[:space:]]^[[:space:]]*^[[:alpha:]]
  :]]//g
2 s/[[:alpha:]]*[[[:digit:]]' '-\{1,\}][[:alpha:]]*//g
3 s/^[[:alpha:]][[:space:]]//g
4 s/[[:space:]]\{1,\}/\n/g

```

Im Skript `prepare-wordlist.sh` wird in Zeile 6 das Programm `comm` mit einer nicht überall verfügbaren Option `-i` verwendet, um zwei Dateien ohne Beachtung von Groß- und Kleinschreibung auf nicht übereinstimmende Zeilen zu prüfen. Dieselbe Funktionalität kann auch mit `grep` erreicht werden. Dazu ist Zeile 6 des Skripts mit der aus Listing 2 zu ersetzen.

Listing 2: Ein Ersatz für `comm -i`.

```

6 | grep -Fixvf words.txt -

```

Die von `grep` zu suchenden Muster werden aus der angegebenen Datei `words.txt` gelesen, der Liste bereits geprüfter Wörter (letzte Option `f` mit folgendem Dateinamen). Es findet ein einfacher Stringvergleich statt (Option `F`), Groß- und Kleinschreibung werden ignoriert (Option `i`), die Muster müssen zeilenweise passen (Option `x`) und es werden *die* Zeilen der zu durchsuchenden Datei ausgegeben, für die *keine* übereinstimmenden Muster in `words.txt` gefunden wurden (Option `v`).

Einige zusätzliche Hinweise: a) Die Filterung mit `grep` läuft langsamer ab als mit `comm` (bei großen Dateien erheblich), da nicht ausgenutzt werden kann, dass die zu vergleichenden Dateien sortiert sind. Dafür sollte dieser Aufruf überall funktionieren. b) Mir ist aufgefallen, dass der Aufruf aus Listing 2 unter Umständen nicht korrekt arbeitet, wenn die Datei `words.txt` leer ist.³ Daher sollte zunächst zur Datei `words.txt` ein beliebiger Eintrag hinzugefügt werden, zum Beispiel das Wort »die«. c) Außerdem ist in einer gemischten Umgebung von Unixshell und Windowskommandozeile auf die richtige (gleiche) Behandlung von Zeilenenden zu achten. Der letzte Hinweis betrifft nicht nur `grep`.

In Absprache mit Werner Lemberg stelle ich hier auch eine neue Version des Skripts `log2words.sed` vor, welche er mir schickte, während dieser Beitrag entstand. Das in Listing 3 zu sehende Skript, kommt besser mit überlangen Wörtern zurecht.

³Zum Beispiel mit GNU `grep` 2.4.2 unter MinGW.

Listing 3: log2words.sed, Version 2.0.

```

1 1,/Underfull/ d
2 /^Underfull/ d
3 /^\\hbox/ d
4 /^$/ d
5 \\[\\] \\T1/cmr/m/n/10 | {
6   s|[\\] \\T1/cmr/m/n/10 ||
7   : loop
8   N
9   s|\\n\\(.*\\)$|\\1|
10  t loop
11  s|\\n||
12  s|[\\]||
13  s/ÿ/ß/g
14 }
15 /(\./,$ d

```

Ausblick

Obwohl Werner Lemberg auf seine Beiträge kein allzu großes Echo bekam, hoffe und vermute ich, dass die hier gestellten Fragen, die auch grundlegende Themen betreffen, für einen größeren Kreis von Interesse sind. Außerdem denke ich, dass die Pflege von Wortlisten als Grundlage von Trennmustern und Ausnahmelisten ein Langzeitprojekt sein sollte und der Mitarbeit vieler interessierter Helfer und Kontrolleure bedarf – zumindest bis die Wortlisten einen gewissen Umfang erreicht haben. Ich halte »Die T_EXnische Komödie« aus diesen Gründen für ein geeignetes Diskussionsforum rund um dieses Thema und freue mich, wenn an dieser Stelle die eine oder andere der gestellten Fragen näher beleuchtet wird oder weitere Fragen aufgeworfen werden.

Zur Zeit findet die Diskussion des Projekts »Freie Wortlisten und Trennmuster für die deutsche Sprache« in der neu eingerichteten Google-Gruppe `trennmuster-opensource` statt.^[2] Wer sich an der Diskussion beteiligen möchte, benötigt ein Zugangskonto bei Google, die Diskussion kann online aber auch ohne Konto verfolgt werden.⁴ Der E-Mailverkehr vor der Gruppeneinrichtung ist im Dateibereich der Gruppe zu finden (Datei `mbox-vorgeschichte.zip`), außerdem auch eine kurze Projektbeschreibung, die hier fehlenden Abschnitte

⁴In naher Zukunft ist geplant auf einen anderen Softwareprojekthost umzuziehen.

meines Entwurfs, sowie neue Trennmuster, die aus Werners aktueller, erheblich erweiterter Wortliste abgeleitet wurden.

Das in Frage 7 thematisierte Problem der Wortlistenbeschaffung kann dank Werner Lemberg und Georg Verweyen als (fast) gelöst betrachtet werden. Zur Zeit verfügt das Projekt über drei Listen mit Umfängen zwischen vierhunderttausend und vier Millionen Wörtern. Eine weitere Liste mit einem repräsentativen Wortschatz aus dem Kernkorpus des Projekts »Digitales Wörterbuch der Deutschen Sprache« steht in Aussicht. Werner Lembergs Liste steht in einem öffentlich zugänglichen Git-Repositoryum und kann mittels

```
git clone git://repo.or.cz/wortliste.git
```

bezogen werden. Genauere Angaben zu den Listen können der Projektbeschreibung entnommen werden. Die Hauptaufgabe besteht nun »lediglich« darin, diese Listen oder Teile davon auf Rechtschreibung und korrekte Trennung zu kontrollieren, damit sie die Grundlage für neue Trennmuster bilden können. Mathias Nater hat bereits einen Entwurf für eine Eingabemaske und die Datenbankbindung fertiggestellt.[6, 7] Er ist für jede Hilfe bei der Arbeit dankbar. Interessierte sollten sich an die Gruppe wenden.[2]

Literatur

- [1] Matthias Clasen: *Proposals for extensions to T_EX*; CTAN:systems/tex-extensions/clasen.
- [2] Google-Gruppe: *Trennmuster-Opensource*.
- [3] Donald E. Knuth: *The T_EXbook*; Bd. A von *Computers and Typesetting*; Addison-Wesley; Reading, MA, USA; 1986.
- [4] Werner Lemberg: *Hyphenation Exception Log für deutsche Trennmuster*; *Die T_EXnische Komödie*; 15(2), S. 28–31; 2003.
- [5] Werner Lemberg: *Hyphenation Exception Log für deutsche Trennmuster*; *Die T_EXnische Komödie*; 17(2), S. 24–51; 2005.
- [6] Mathias Nater: <http://www.mnn.ch/hyph/silbentrennung1.html>.
- [7] Mathias Nater: <http://www.mnn.ch/opendehyph/index.php>.
- [8] Petr Sojka: *Notes on Compound Word Hyphenation in T_EX*; *TUGboat*; 16(3), S. 290–296; 1995.

Von fremden Bühnen

Verleihung der Goethe-Plakette an Hermann Zapf

Volker RW Schaa



Am 11. Dezember 2007 wurde unser Ehrenmitglied Prof. Dr. h.c. Hermann Zapf in der Wiesbadener Staatskanzlei mit der Goethe-Plakette ausgezeichnet. Die Goethe-Plakette ist die höchste Auszeichnung des Hessischen Ministeriums für Wissenschaft und Kunst (HMWK) und wird seit 1949 in unregelmäßigen Abständen verliehen. Sie wird Persönlichkeiten zugesprochen, die durch ihr Lebenswerk in besonderer Weise zur kulturellen Entwicklung des Landes Hessen beigetragen haben.

Im Anhang sind Auszüge aus der offiziellen Pressemitteilung des Hessischen Ministeriums für Wissenschaft und Kunst, dem Bericht von hr-online und aus dem Wiesbadener Kurier, einer regionalen Zeitung, zu finden.

Presseerklärung des HMWK[1]

Goethe Plakette des Landes Hessen an Prof. Dr. h.c. Hermann Zapf, Darmstadt

Der hessische Minister für Wissenschaft und Kunst überreichte Hermann Zapf am 11. Dezember 2007 in der Staatskanzlei in Wiesbaden die Goethe-Plakette, die höchste kulturelle Auszeichnung des Landes.

Zu den bisherigen Preisträgern gehören Max Horkheimer, Paul Hindemith, Carl Orff, Th. W. Adorno, Siegfried Unseld, Sir Georg Solti, Marcel Reich-Ranicki, etc.

Als Buch- und Schriftkünstler gestaltete Hermann Zapf seit 1938 über 200 Alphabete für Bleisatz, Photosatz und digitale Wiedergabe. Seit Anfang der 70er Jahre beschäftigte er sich bereits mit typographischen Computerprogrammen, zu einer Zeit, als man solche Ideen in Deutschland noch nicht ernst nahm. Der »Macintosh« von Apple kam erst acht Jahre später auf den Markt.



1977 übertrug man ihm die weltweit erste Professur für »Typographic Computer Programs« am Rochester Institute of Technology in Rochester/New York (1977–1987).

1991 bekam er auf sein »hz-Programm, einem System zur Verbesserung der Satzqualität in Büchern und Zeitschriften« ein Europa-Patent. Hermann Zapf wurde 1974 der

Gutenberg-Preis der Stadt Mainz überreicht. Die Royal Society of Arts in London ernannte ihn 1985 zum »Honorary Designer for Industry«. Unter weiteren Ehrungen erhielt er 1969 den Frederic W. Goudy Award des R.I.T. in Rochester, den Euro Design Award, Oostende 1994 und 1973 die Johann Heinrich Merck Ehrung der Stadt Darmstadt.

Pressemitteilung Nr. 230/2007 des HMWK[2]

Staatssekretär Prof. Dr. Lorz würdigt das Werk dreier vorbildlicher Kulturschaffender.

Wiesbaden — Der Kirchenmusiker und Musikpädagoge Martin Lutz, der Jazzjournalist und Hörfunkmoderator Heinz Werner Wunderlich und der Buch- und Schriftkünstler Prof. Dr. h.c. Hermann Zapf sind vom hessischen Minister für Wissenschaft und Kunst, Udo Corts, mit der Goethe-Plakette des Landes ausgezeichnet worden. Staatssekretär Prof. Dr. Ralph Alexander Lorz würdigte die drei Geehrten heute bei der Überreichung der Plaketten in der Wiesbadener Staatskanzlei als vorbildliche Kulturschaffende, die auf ganz unterschiedlichen Gebieten Beispielhaftes leisteten: »Ihr Wirken und ihr Engagement ist ein Beleg für kulturelle Vielfalt in Hessen und zeigt sehr

überzeugend, wie bereichernd und völkerverbindend kulturelles Wirken sein kann.« [...]»



Hermann Zapf, 89, arbeitete als Schriftgrafiker, Kalligraph und Schriftdesigner und entwarf noch vor dem Zweiten Weltkrieg seine erste Type »Gilgengart«. Er entwarf zahlreiche Schriften, darunter auch die allen Computernutzern bekannten »Zapf Dingbats«. Schon in den frühen sechziger Jahren des 20. Jahrhunderts beschäftigte er sich mit der Kombination von Typographie und Computerprogrammen. 1976 erhielt er den weltweit ersten Lehrstuhl für »Typographic Computer Programs« am amerikanischen Rochester Institute for Technology. 1994 bekam er

ein Europa-Patent für das von ihm entwickelte »hz-Programm« zur Verbesserung der Satzqualität in Büchern und Zeitschriften. Hermann Zapf wurde unter anderem mit dem Gutenberg-Preis der Stadt Mainz (1974) und dem Euro Design Award (1994) ausgezeichnet. 2003 wurde ihm die Ehrendoktorwürde (Fine Arts) der University of Illinois verliehen. [...]

hr-online[3]

[...] Schriftenerfinder

Hermann Zapf, 89, arbeitete als Schriftgrafiker, Kalligraph und Schriftdesigner und entwarf noch vor dem Zweiten Weltkrieg seine erste Type »Gilgengart«. Er entwarf zahlreiche Schriften, darunter auch die allen Computer-Nutzern bekannten »Zapf Dingbats«. Schon in den frühen sechziger Jahren des 20. Jahrhunderts beschäftigte er sich mit der Kombination von Typographie und Computerprogrammen. 1976 erhielt er den weltweit ersten Lehrstuhl für »Typographic Computer Programs« am amerikanischen Rochester Institute for Technology. 1994 bekam er ein Europa-Patent für das von ihm entwickelte »hz-Programm« zur Verbesserung der Satzqualität in Büchern und Zeitschriften. Hermann Zapf wurde unter anderem mit dem Gutenberg-Preis der Stadt Mainz (1974) und dem Euro Design Award (1994) ausgezeichnet. 2003 wurde ihm die Ehrendoktorwürde (Fine Arts) der University of Illinois verliehen. [...]

Wiesbadener Kurier[4]

Verleihung der Goethe-Plaketten in der Hessischen Staatskanzlei

[...] Einer stillen, aber im Alltag nicht weniger präsenten Kunst ist der dritte Preisträger verpflichtet: Hermann Zapf, 1918 geborener Buch- und Schriftkünstler, hat zahlreiche Schriften erfunden und schon in den 60er Jahren die Computertechnik für die Typografie nutzbar gemacht. Staatssekretär Ralph Alexander Lorz, der die Goethe-Plaketten stellvertretend für Kunstminister Udo Corts verlieh, hielt es »für höchste Zeit, dass nun auch das Land Hessen« den Mainzer Gutenberg-Preisträger Zapf »für seine enorme künstlerische Leistung« auszeichnet. [...]



Literatur

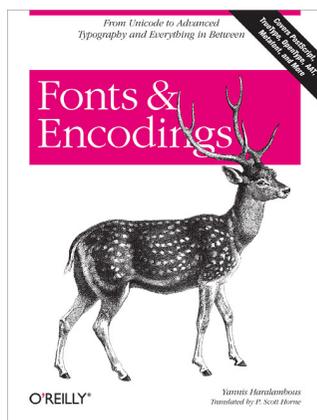
- [1] Presseerklärung des HMWK, 11. Dezember 2007, Original von Hermann Zapf, *Goethe-Plakette des Landes Hessen an Prof. h.c. Hermann Zapf, Darmstadt*
- [2] Pressemitteilung Nr. 230/2007, 11. Dezember 2007, Melanie Krombach, Referat M2, Hessisches Ministerium für Wissenschaft und Kunst *Goethe-Plakette für Martin Lutz, Heinz Werner Wunderlich, Hermann Zapf*
- [3] hr-online, 11. Dezember 2007, Goethe-Plakette, *Martin Lutz, Heinz Werner Wunderlich und Hermann Zapf geehrt*
http://www.hr-online.de/website/rubriken/kultur/index.jsp?rubrik=5676&key=standard_document_33542800
- [4] Wiesbadener Kurier, 12. Dezember 2007, Volker Milch, »Wie passt das alles in ein Leben?«, *Verleihung der Goethe-Plaketten in der hessischen Staatskanzlei*
http://www.wiesbadener-kurier.de/region/objekt.php3?artikel_id=3086446

Rezensionen

Yannis Haralambous: *Fonts & Encodings*

Ulrik Vieth

Yannis Haralambous ist ein alter Bekannter in der internationalen \TeX -Szene, nicht nur als Mitbegründer des Omega-Projekts, sondern auch durch seine vielfältigen Beiträge bei der Entwicklung exotischer Schriften. Es erscheint daher äußerst passend, dass Yannis es unternommen hat, ein Buch über das Thema Fonts und Codierungen zu schreiben. Die Originalausgabe erschien bereits 2004 in französischer Sprache unter dem Titel *Fontes & Codages*. Nun liegt seit einigen Monaten endlich auch die lang erwartete englische Übersetzung unter dem Titel *Fonts & Encodings* vor.



Inhalt

Mit einem Umfang von über 1000 Seiten beeindruckt das Werk nicht nur durch die Breite der behandelten Themen, sondern auch durch die Tiefe der behandelten Details. In einigen Fällen verwendet der Autor bewusst Dutzende Seiten auf die Dokumentation der Details diverser Font-Formate, die bisher nirgendwo anders in dieser Vollständigkeit beschrieben wurden.

Der Inhalt gliedert sich in einen Hauptteil mit 14 Kapiteln und 600 Seiten sowie einen Anhang mit weiteren 7 Kapiteln und weiteren knapp 400 Seiten. Einige Kapitel sind eher encoding-spezifisch, andere eher font-spezifisch, ein weiterer Teil liegt in der Grauzone zwischen diesen Extremen. Einige Kapitel orientieren sich durchaus an den Bedürfnissen der Endbenutzer, die Fonts

einfach nur installieren und nutzen wollen, andere dagegen sind hauptsächlich für Font-Designer oder Entwickler von Font-Software von Interesse.

Der erste Teil des Buches beschäftigt sich mit encoding-spezifischen Themen. Kapitel 1 beschreibt die historische Entwicklung der Codierungen, angefangen von 7-Bit ASCII über die diversen 8-Bit ISO Latin Codierungen bis hin zu ost-asiatischen 16-Bit Codierungen. Die Kapitel 2–4 beschreiben dann den Unicode-Standard, beginnend mit einem Überblick der darin enthaltenen Symbole und Zeichen, bis zu komplexen technischen Implementierungsdetails. Kapitel 5 beschließt diesen Teil mit der Vorstellung einiger nützlicher Tools zur Eingabe von Unicode-Zeichen auf verschiedenen Systemplattformen.

Der zweite Teil beschäftigt sich mit der Font-Installation auf verschiedenen System-Plattformen und bewegt sich dabei in der Grauzone zwischen Fonts und Encodings. Die Kapitel 6–8 behandeln dieses Thema jeweils für die Macintosh-, Windows- und Unix/X11-Plattform. Während die Beschreibung der Macintosh-Plattform recht ausführlich auf die Unterschiede von MacOS 9 und MacOS X eingeht, ist die Beschreibung der Unix/X11-Plattform leider etwas knapp gehalten. Hier hätte man sich durchaus einige Ergänzungen wünschen können, beispielweise zum Font-Management in modernen Linux-Desktop-Umgebungen wie KDE oder Gnome.

Die folgenden zwei Kapitel widmen sich dann der Plattform-unabhängigen Nutzung von Fonts in \TeX /Omega sowie in Web-Browsern. Kapitel 9 beginnt mit der Beschreibung der Font-Auswahl in \LaTeX /NFSS2, gefolgt von einer detaillierten Anleitung zur Font-Installation für *dvips*. Im weiteren Verlauf folgen dann eine Reihe von Beispielen zur Erstellung virtueller Fonts mit *fontinst* für zunehmend komplexere Anwendungsfälle. Kapitel 10 beschließt diesen Teil mit Hinweisen zur Nutzung von Fonts in Web-Browsern mit Hilfe von (X)HTML/CSS oder alternativ mit SVG.

Der letzte Teil beschäftigt sich mit den Font-spezifischen Themen. Kapitel 11 behandelt die Klassifikation lateinischer Schriftarten und liefert gleichzeitig einen reichhaltig illustrierten Überblick über die historische Entwicklung der wichtigsten Schriftfamilien. Die Kapitel 12–13 widmen sich dann der Entwicklung, Bearbeitung und Optimierung von PostScript-, TrueType- und OpenType-Fonts mit *FontLab* oder *FontForge*. Kapitel 14 beschreibt schließlich die Nutzung der erweiterten typographischen Möglichkeiten von OpenType- und AAT-Fonts.

Der Anhang des Buches widmet sich dann der ausführlichen Beschreibung der diversen Font-Formate. Angefangen von Bitmap-Fonts und $\text{T}_{\text{E}}\text{X}$ -Fonts über PostScript- und TrueType- bis hin zu OpenType- und AAT-Fonts werden in Anhang A–E alle wichtigen Font-Formate beschrieben. Anhang F beschäftigt sich letztendlich mit METAFONT und den daraus abgeleiteten Systemen wie METAPOST , MetaFog und MetaType-1 .

Fazit

Bei einem Buch dieses Umfangs ist es verständlich, dass von der Erstellung des Manuskripts bis zur Fertigstellung einer Übersetzung mehrere Jahre vergehen können, was leider zu Lasten der Aktualität geht. Man muss daher wohl davon ausgehen, dass die im Jahre 2007 frisch erschienene englische Übersetzung im wesentlichen nur auf dem Stand von 2003 beruht.

Bei den meisten Kapiteln mit Referenz-Charakter ist ein solcher Mangel an Aktualität nicht weiter tragisch, da die Beschreibungen der Codierungen und Font-Formate dauerhafte Gültigkeit behalten. Dagegen ist es jedoch sehr bedauerlich, dass gerade im Kapitel über $\text{T}_{\text{E}}\text{X}$ einige sehr wichtige Entwicklungen der letzten Jahre völlig verpasst oder übersehen wurden.

So beschreibt der Autor in Kapitel 9 zwar ausführlich die Font-Installation für $\text{T}_{\text{E}}\text{X}/\text{Omega}$ und dvips , jedoch wird $\text{pdfT}_{\text{E}}\text{X}$ in diesem Zusammenhang kein einziges Mal erwähnt, obwohl die meisten Details in modernen $\text{T}_{\text{E}}\text{X}$ -Live-Installationen für beide Systeme gleichermaßen anwendbar sind.

Hàn Thê Thành wird lediglich einmal als Autor vietnamesischer Fonts zitiert, dagegen finden seine Leistungen zur Implementierung der mikrotypographischen Erweiterungen in $\text{pdfT}_{\text{E}}\text{X}$ keine Beachtung. Dies ist um so verwunderlicher, als dass der Autor auch noch ein Beispiel behandelt, wo der Effekt von *margin kerning*, der in $\text{pdfT}_{\text{E}}\text{X}$ direkt verfügbar wäre, recht umständlich mit Hilfe von virtuellen Fonts in *fontinst* simuliert wird.

Als Erweiterungen der Computer-Modern-Fonts wird das Paket CM-Super empfohlen, während die wesentlich erfolgreichereren $\text{Latin-Modern-Fonts}$ nur beiläufig als Beispiel für eine MetaType-1 -Anwendung erwähnt werden.

Schließlich verweist der Autor zur Nutzung von OpenType- und AAT-Fonts in $\text{T}_{\text{E}}\text{X}$ (verständlicherweise) auf seine eigenen Entwicklungs- und Forschungsvorhaben im Rahmen des Omega2 -Projekts, während der inzwischen recht erfolgreiche Newcomer $\text{X}_{\text{F}}\text{T}_{\text{E}}\text{X}$ völlig unerwähnt bleibt.

Fairerweise muss man zugestehen, dass einige dieser Entwicklungen bei Entstehung des Manuskripts im Jahre 2003 nicht vorhersehbar waren. Dennoch wäre es sicher möglich gewesen, in die englische Fassung des Jahres 2007 einige Ergänzungen und Überarbeitungen einzubringen, statt nur eine Übersetzung des nicht mehr ganz aktuellen Standes abzuliefern.

So wurde hier leider eine Gelegenheit verpasst, die das Buch noch wertvoller und nützlicher für T_EX-User gemacht hätte, die sich für die aktuellen Entwicklungen der Font-Technologie interessieren.

Trotz dieser Mängel bleibt das Buch in weiten Teilen eine wertvolle Referenz für fortgeschrittene T_EX-User und Software-Entwickler, die sich ernsthaft mit Font-Technologien und Codierungen beschäftigen wollen. Es gibt wohl kein anderes Buch, das die gesamte Thematik in einer ähnlichen Breite und Tiefe abdeckt. Stattdessen müsste man Dutzende von Referenzen aus verschiedenen Quellen zusammensuchen und hätte immer noch Lücken zu füllen.

Insgesamt gesehen ist das Werk sicherlich empfehlenswert. Dennoch wären Ergänzungen und Überarbeitungen bestimmter Kapitel sehr wünschenswert.

Zum Abschluss sei schließlich noch eine kleine aber nette Kuriosität erwähnt: Wie viele andere moderne Sachbücher verfügt auch dieses Werk nicht nur über ein Sachregister, sondern auch über ein eigenes Personenregister. Dabei muss man allerdings feststellen, dass der Autor hier offenbar recht großzügig vorgegangen ist. So finden sich gegenwärtige Autoren von Font-Software hier nicht nur in der Gesellschaft der größten Schriftgestalter der Geschichte, sondern auch in der Nähe von Personen der Zeitgeschichte (z. B. Lenin, Hitler, Mao) sowie literarischer Figuren (z. B. Sherlock Holmes, James Bond, James T. Kirk), die lediglich beiläufig erwähnt werden.

Während sich der Autor dieser Buchbesprechung für die Aufnahme in diese recht einzigartige Gesellschaft bedankt, möchte er dennoch bezweifeln, ob es wirklich zweckmäßig ist, literarische Figuren auf die gleiche Art wie Personen der Gegenwart in dasselbe Register aufzunehmen.

Yannis Haralambous, Fonts & Encodings, O'Reilly Media Inc., 1016 + xx pages, first edition: September 2007, ISBN: 978-0-596-10242-5, US \$ 59.99

Leserbriefe

Hermann Zapf

Liebe Freunde,

im November-Heft 2007 der Zeitschrift »Die $\text{T}_{\text{E}}\text{X}$ nische Komödie« sind auf den Seiten 27–40 unter dem Titel »Das Schriftpaket `initials`« verschiedene Alphabete abgedruckt. Sie stammen alle aus der längst verstaubten Mottenkiste von Alphabeten aus der Zeit der Jahrhundertwende des 20. Jahrhunderts.

Es gibt unter den Mitgliedern von »Dante« genügend junge Talente, die zeitgemäße Initialen entwerfen können und sich bestimmt über eine solche Gelegenheit freuen würden, etwas speziell zu entwerfen.

Solche furchtbaren Initialen schaden nur unserer guten Sache und erweisen unseren Mitgliedern keinen guten Dienst. Ich habe immer die Sorgfalt und die Details in den Schriften von Dante bewundert, aber diese Alphabete haben mich erschreckt.

Wer ist für die Schriftauswahl verantwortlich, um diese Monster wieder aus dem Schriftangebot herauszunehmen?

(Kopien an Volker Schaa, Hans Hagen und die Schriftleitung von »Dante«).

Mit herzlichen Grüßen

Hermann

Spielplan

Termine

2008

25. 1. – 26. 1. **Asian T_EX Conference 2008 (Asia T_EX'08)**
Kongju National University, South Korea
<http://conf.ktug.kr/2008/>, Email: AsiaTeX08@ktug.or.kr
5. 3. – 7. 3. **DANTE 2008**
Frühjahrstagung und 38. Mitgliederversammlung
Friedrich-Schiller-Universität Jena
<http://www.dante.de/dante2008/>
30. 4. – 4. 5. **16. BachoT_EX-Konferenz BachoT_EX 2008**
Bachotek, Nähe Brodnica, Polen
<http://www.gust.org.pl/BachoTeX/2008>
28. 5. – 31. 5. **Linuxtag 2008**
Messe Berlin
Weitere Informationen bei Karlheinz Geyer
geyerk.fv.tu@nds.tu-darmstadt.de
<http://www.linuxtag.org/2008/>
21. 7. – 24. 7. **TUG 2008**
T_EX's 30th birthday
Cork, Irland
<http://tug2008.ucc.ie/>
20. 8. – 25. 8. **2nd ConT_EXt user meeting**
Bohinj, Slovenia
<http://meeting.contextgarden.net/2008/>

Stammtische

In verschiedenen Städten im Einzugsbereich von DANTE e.V. finden regelmäßig Treffen von T_EX-Anwendern statt, die für jeden offen sind. Im WWW gibt es aktuelle Informationen unter <http://www.dante.de/events/stammtische/>.

Aachen

Torsten Bronger
bronger@physik.rwth-aachen.de
Gaststätte Knossos, Templergraben 28
Zweiter Donnerstag im Monat, 19.00 Uhr

Berlin

Rolf Niepraschk
Tel.: 0 30/3 48 13 16
Rolf.Niepraschk@gmx.de
Humboldt Universität Berlin,
Fachbereich Wirtschaftswissenschaften
Spandauer Straße 1
Zweiter Donnerstag im Monat, 19.00 Uhr

Bremen

Axel Sommerfeldt
Tel.: 04 21/3 96 40 57
Axel.Sommerfeldt@arcor.de
Wechselnder Ort
Erster Donnerstag im Monat, 18.30 Uhr

Darmstadt

Karlheinz Geyer
geyerk.fv.tu@nds.tu-darmstadt.de
www.da-tex.org
Restaurant Poseidon,
Rheinstraße 41, 64283 Darmstadt
Erster Freitag im Monat, ab 19.30 Uhr

Dresden

Carsten Vogel
lego@wh10.tu-dresden.de
Studentenwohnheim, Borsbergstraße 34,
Dresden, Ortsteil Striesen
ca. alle 8 Wochen, Mittwoch, 19.00 Uhr

Düsseldorf

Georg Verweyen
Georg.Verweyen@web.de
Bistro/Café Zicke
Böckerstr. 5 a (Ecke Bergerallee)
40213 Düsseldorf
Zweiter Mittwoch in ungeraden Monaten,
20.00 Uhr

Erlangen

Walter Schmidt, Peter Seitz
w.a.schmidt@gmx.net
Gaststätte »Deutsches Haus«
Luitpoldstraße 25
3. Dienstag im Monat, 19.00 Uhr

Freiburg

Heiko Oberdiek
Tel.: 07 61/4 34 05
oberdiek@uni-freiburg.de
Wechselnder Ort
Dritter Donnerstag im Monat, 19.30 Uhr

Hamburg

Lothar Fröhling
lothar@thefroehlings.de
Zum Schwarzenberg
Schwarzenbergstr. 80 – 21073 HH
letzter Dienstag im Monat, 19.30 Uhr

Hannover

Mark Heisterkamp
heisterkamp@rrzn.uni-hannover.de
Seminarraum RRZN
Schloßwender Straße 5
Zweiter Donnerstag im Monat, 18.30 Uhr

Heidelberg

Luzia Dietsche
Tel.: 0 62 21/54 45 27
luzia.dietsche@urz.uni-heidelberg.de
»Restaurant Tomato, der Turm«
Alte Glockengießerei 9
Letzter Mittwoch im Monat, 19.30 Uhr

Karlsruhe

Klaus Braune
Tel.: 07 21/6 08 40 31
braune@rz.uni-karlsruhe.de
Universität Karlsruhe, Rechenzentrum
Zirkel 2, 3. OG, Raum 316
Erster Donnerstag im Monat, 19.30 Uhr

Köln

Helmut Siegert
 Institut für Kristallographie
 Zölpicher Straße 49b
 Letzter Mittwoch im Monat, 19.30 Uhr

München

Michael Niedermaier
<http://www.uwe-siart.de/typografie/stammtisch.html>
 Momentan wechselnder Ort
 Erste Woche des Monats an wechselnden
 Tagen, 19.00 Uhr

Stuttgart

Bernd Raichle
bernd.raichle@gmx.de
 Bar e Ristorante »Valle«
 Geschwister-Scholl-Str. 3
 Zweiter Dienstag im Monat, 19.30 Uhr

Trier

Martin Sievers
stammtisch-trier@texberatung.de
 Fetzenkneipe (Haus Fetzenreich)

Sichelstraße 36 (beim Sieh-Um-Dich)
 54290 Trier
 Dritter Montag des Monats, 20.15 Uhr

Ulm

Adelheid Grob
adelan@heidi.in-ulm.de
<http://latex.in-ulm.de>
 Gaststätte »Peppers Ulm«
 Deinselsgasse 8
 Erster Donnerstag im Monat, 19.30 Uhr

Wuppertal

Andreas Schrell
 Tel.: 02193/53 10 93
as@schrell.de
 Restaurant Croatia »Haus Johannisberg«
 Südstraße 10
 an der Schwimmooper Wuppertal-Elberfeld
 Zweiter Donnerstag im Monat, 19.30 Uhr

Würzburg

Bastian Hepp
dante@sning.de
 nach Vereinbarung

Adressen

DANTE, Deutschsprachige Anwendervereinigung T_EX e.V.
Postfach 10 18 40
69008 Heidelberg

Tel.: 0 62 21/2 97 66 (Mo., Mi.–Fr., 10.00–12.00 Uhr)
Fax: 0 62 21/16 79 06
E-Mail: dante@dante.de

Konto: Volksbank Rhein-Neckar eG
BLZ 670 900 00
Kontonummer 2 310 007
IBAN DE67 6709 0000 0002 3100 07
SWIFT-BIC GENODE61MA2

Präsidium

Präsident:	Klaus Höppner	president@dante.de
Vizepräsident:	Volker RW Schaa	vice-president@dante.de
Schatzmeister:	Tobias Sterzl	treasurer@dante.de
Schriftführer:	Manfred Lotz	secretary@dante.de
Beisitzer:	Günter Partosch	
	Bernd Raichle	
	Herbert Voß	

Server

ftp: [ftp.dante.de](ftp://ftp.dante.de)
WWW: <http://www.dante.de/>

Autoren/Organisatoren

Stephan Hennig Peter-Rosegger-Weg 4 16540 Hohen Neuendorf	[7]	Volker RW Schaa siehe Seite 30	[4, 18]
Klaus Höppner siehe Seite 30	[4, 5]	Ulrik Vieth Vaihinger Straße 69 70567 Stuttgart	[22]
Christine Römer Friedrich-Schiller-Universität Jena Institut für Germanistische Sprachwissenschaft Fürstengraben 30 07740 Jena Christine.Roemer@uni-jena.de	[5]	Herbert Voß siehe Seite 30	[3]
		Hermann Zapf Seitersweg 35, 64287 Darmstadt	[26]

Die T_EXnische Komödie

20. Jahrgang Heft 1/2008 Januar 2008

Impressum

Editorial

4 Grußwort

T_EX-Theatertage

5 DANTE 2008 – Einladung zur Mitgliederversammlung

Bretter, die die Welt bedeuten

7 Einige Fragen zum Beitrag »Hyphenation Exception Log für deutsche Trennmuster, Version 1«

Von fremden Bühnen

18 Verleihung der Goethe-Plakette an Hermann Zapf

Rezensionen

22 Yannis Haralambous: *Fonts & Encodings*

Leserbriefe

26 Hermann Zapf

Spielplan

27 Termine

28 Stammtische

Adressen

31 Autoren/Organisatoren