

LaTeX und HTML, Teil 2

„Gleichzeitiges“ Publizieren in LaTeX *und* HTML

Günter Partosch, HRZ Gießen

E-Mail: `Guenter.Partosch@hrz.uni-giessen.de`

Inhalt des Vortrags

- Problemstellung
- Ausgangsinformationen
- Konverter, Beschreibungssprachen und andere Dinge
- Testbeispiel
- Zu untersuchende Elemente im Testbeispiel
- Mögliche Zuordnung einfacher HTML-Elemente zu LaTeX-Strukturen und umgekehrt
- Konvertierung LaTeX \rightarrow HTML
 - Umsetzung „per Hand“
 - Konvertierung mit LaTeX2HTML
 - Konvertierung mit Hyperlatex
- Konvertierung HTML \rightarrow LaTeX
 - `html2latex.sed`
- Beschreibungssprachen
 - YODL
 - Hyperlatex als Beschreibungssprache

Problemstellung

- Viele Autoren müssen sowohl in Papierform (also z.B. mit LaTeX) als auch im World-Wide Web (also mit HTML) publizieren.
- Es ist sehr aufwendig, unzuverlässig und sehr fehlerträchtig, auf Dauer zwei Erscheinungsformen einer Veröffentlichung zu pflegen.
- Konvertierung „per Hand“ ist nur in Ausnahmefällen möglich.
- Am geeignetsten erscheinen Verfahren und Vorgehensweisen, die automatisch und ohne großen Aufwand die Ausgangsdateien in die gewünschte Darstellung umwandeln.

Ausgangsinformationen

Als Ausgangspunkte wurden die folgenden Dokumente im WWW benutzt:

Allgemein:

„*The LaTeX-PC Converter FAQ in English*“ (<http://www.kfa-juelich.de/isr/1/texcnven.txt>)

Konvertierung von HTML nach LaTeX:

- „*Other tools*“ in „*Converting from HTML*“
(<http://www.w3.org/pub/WWW/Tools/html2things.html>)
- Abschnitt 23.11 „*Comment convertir du HTML en LaTeX ?*“ in „*FAQ LaTeX francaise V2.1*“
(ftp://ftp.dante.de/tex-archive/help/LaTeX-FAQ-francaise/FAQ_LaTeX_francaise_V2.1)

Konvertierung von LaTeX nach HTML:

- „*LaTeX etc.*“ in „*Word Processor filters etc*“
(http://www.w3.org/pub/WWW/Tools/Word_proc_filters.html)
- Abschnitt 23.12 „*Comment convertir du LaTeX en HTML ?*“ in „*FAQ LaTeX francaise V2.1*“
(ftp://ftp.dante.de/tex-archive/help/LaTeX-FAQ-francaise/FAQ_LaTeX_francaise_V2.1)

Konverter, Beschreibungssprachen und andere Dinge

axtex	Einbettung von LaTeX in HTML
BETA Format	Beschreibungssprache
h2l	HTML→LaTeX
html2latex	HTML→LaTeX
html2latex.sed	HTML→LaTeX
html2tex	HTML→LaTeX
htmlatex.pl	HTML→LaTeX
Hyperlatex	Beschreibungssprache, LaTeX→HTML
HyperTeX	Darstellung von Hypertext-DVI-Dateien
IDVI	Darstellung von Hypertext-DVI-Dateien mittels Java-Applets
JAM	Beschreibungssprache
l2x	LaTeX→HTML
LaTeX2HTML	LaTeX→HTML
LaTeX2Hyp	LaTeX→HTML
LTX2X	LaTeX→HTML
math2html	LaTeX→HTML
TeX2RTF	LaTeX→HTML
typehtml	LaTeX als HTML-Parser
vulcanize	LaTeX→HTML
YODL	Beschreibungssprache

- keine Binär- oder Quellprogramme erhältlich: axtex, htmlatex.pl, vulcanize
- Installation nicht erfolgreich oder Installation aus anderen Gründen nicht vorgenommen: html2tex, HyperTeX, IDVI, l2x
- Die übrigen Produkte wurden genauer getestet.

Testbeispiel

```
\documentclass[a4paper]{article}
% LaTeX2e-Datei zum Testen verschiedener Konverter
\usepackage{german}
\title{Test f"ur Konverter}
\author{G"unter Partosch}
\thanks{Hochschulrechenzentrum Gie"sen}}
\date{\today}

\newcommand{\HTML}{HTML~3.2}

\begin{document}
\maketitle

\section{Erster Abschnitt}

In diesem Abschnitt werden einige
Textauszeichnungen f"ur {\HTML} getestet,
wie z.B.

\begin{itemize}
\item Das ist \texttt{Typewriter}.
\item Das ist \textsc{SmallCaps}.
\item Das ist \textit{Italic} und
      \textit{\textbf{Italic boldface}}.
\end{itemize}

\begin{description}
\item[Sansserif] Das ist \textsf{Sansserif}.
Und noch ein bi"schen anderer Text, so da"s
eine neue Zeile begonnen wird.
\item[boldface] Das ist \textbf{boldface}.
Und noch ein bi"schen anderer Text, so da"s
eine neue Zeile begonnen wird.
\end{description}

\noindent und Umlaute ("a"o"u"A"O"U"s) und

\begin{enumerate}
\item Das ist \emph{hervorgehobener} Text.
\item Das ist \verb|verbatim| Text.
\end{enumerate}
```

```
\section{Zweiter Abschnitt}
```

```
\subsection{Tabelle}
```

Hier wird eine kleine Tabelle eingefügt:

```
\begin{tabular}{||l|c|r||}  
\hline  
abc & x & 120.55 \\  
defg & -- & 11.37\\  
hijklmn & xxx & 12.95\\  
\hline  
\end{tabular}
```

```
\subsection{verbatim}
```

Ein kleiner verbatim gesetzter Text:

```
\begin{verbatim}  
jhj hj h jh jh jh jj h  
  k kl k lo klo      kk k k ol klo kl  
klj kj j  jkl k  
\end{verbatim}  
  
\end{document}
```

Zu untersuchende Elemente im Testbeispiel

- a) Dokumentklasse `article`, Dokumentoption `a4paper`
- b) Kommentar
- c) LaTeX2e-Package `german`
- d) Dokumenttitel aus `\title`, `\author`, `\thanks` und `\date`; Ausgabe des Titels durch `\maketitle`
- e) neues Kommando durch `\newcommand`; Aufruf dieses neuen Kommandos
- f) Abschnittsbefehle `\section`, `\subsection`
- g) Aufzählungen mit `itemize`-, `enumerate`- und `description`-Umgebung
- h) Font-Befehle `\texttt`, `\textsc`, `\textit`, `\textbf`, `\textsf`; auch Schachtelungen
- i) deutsche Umlaute und Eszet
- j) Verbatim-Textteile mit `\verb`- und `verbatim`-Umgebung
- k) Tabelle mit `tabular`-Umgebung

Mögliche Zuordnung einfacher HTML-Elemente zu LaTeX-Strukturen (und umgekehrt)

\	\backslash
{	\{
}	\}
%	\%
\$	\\$
#	\#
_	_
~	\~
^	\^
<	<
>	>
&	&
 ... 	\textbf{ ... }
<BIG> ... </BIG>	{\large ... }
<BLOCKQUOTE> ... </BLOCKQUOTE>	\begin{quote} ... \end{quote}
<CENTER> ... </CENTER>	\begin{center} ... \end{center}
<CITE> ... </CITE>	\textit{ ... }
<CODE> ... </CODE>	\texttt{ ... }
<DFN> ... </DFN>	\textit{ ... }

<DL> ... <DT> ... <DD> ... </DL>	\begin{description} ... \item[...] \end{description}
 ... 	\emph{ ... }
<Hi> ... </Hi>	\section, \subsection, ... bzw. \chapter, \section,...
<HR> ... </HR>	\hrule
<I> ... </I>	\textit{ ... }
<KBD> ... </KBD>	\texttt{ ... }
<LISTING> ... </LISTING>	\begin{verbatim} ... \end{verbatim}
<MENU> ... </MENU>	\begin{itemize} ... \end{itemize}
 	\begin{enumerate} ... \item \end{enumerate}
<P> ... </P>	\par ... \par
<PRE> ... </PRE>	\begin{alltt} ... \end{alltt}
<SAMP> ... </SAMP>	\verb ...
<SMALL> ... </SMALL>	{\small ... }
<STRIKE> ... </STRIKE>	
 ... 	\textbf{ ... }
... _{...}	\$... _{ ... }\$

... ^{...}	\$... ^{ ... }\$
<TT> ... </TT>	\texttt{ ... }
<U> ... </U>	\underline{ ... }
 	\begin{itemize} ... \item \end{itemize}
<VAR> ... </VAR>	\textsl{ ... }
<XMP> ... </XMP>	\begin{verbatim} ... \end{verbatim}

Umsetzung „per Hand“

```
<HTML>
<HEAD>
<TITLE>Test f&uuml;r Konverter</TITLE>
</HEAD>
```

```
<BODY>
```

```
<H1>Erster Abschnitt</H1>
```

In diesem Abschnitt werden einige
Textauszeichnungen für
HTML 3.2 getestet, wie z.B.

```
<UL>
<LI>Das ist <TT>Typewriter</TT>.
<LI>Das ist SMALLCAPS.
<LI>Das ist <I>Italic</I> und
      <I><B>Italic boldface</B></I>.
</UL>
```

```
<DL>
<DT>Sansserif
<DD>Das ist Sansserif. Und noch ein
bi&szlig;chen anderer Text, so da&szlig;
eine neue Zeile begonnen wird.
<DT>boldface
<DD>Das ist <B>boldface</B>. Und noch ein
bi&szlig;chen anderer Text, so da&szlig;
eine neue Zeile begonnen wird.
</DL>
```

und Umlaute
(ä ö ü Ä Ö Ü ß) und

```
<OL>
<LI>Das ist <EM>hervorgehobener</EM> Text.
<LI>Das ist <SAMP>verbatim</SAMP> Text.
</OL>
```

```
<H1>Zweiter Abschnitt</H1>
```

```
<H2>Tabelle</H2>
```

Hier wird eine kleine Tabelle
eingefügt:

```
<TABLE BORDER>
<TR><TD>abc <TD> x <TD> 120.55
<TR><TD>defg <TD> -- <TD> 11.37
<TR><TD>hijklmn <TD> xxx <TD> 12.95
</TABLE>
```

```
<H2>verbatim</H2>
```

Ein kleiner verbatim gesetzter Text:

```
<LISTING>
jhj hj h jh jh jj h
 k kl k lo klo      kk k k ol klo kl
klj kj j  jkl k
</LISTING>
```

```
</BODY>
</HTML>
```

LaTeX2HTML

- **Beschreibung:** LaTeX2HTML ist ein Konverter, der ein LaTeX-Dokument in mehrere miteinander verbundene HTML-Dateien konvertiert. Er ist in der Lage, auch kompliziertere und komplexere LaTeX-Strukturen (wie Mathematik, Verzeichnisse, Numerierungen, Querverweise, Bilder und Tabellen) korrekt zu bearbeiten. Ggf. werden dazu auch entsprechende GIF-Bilder erzeugt, die über `` in die HTML-Dateien eingebunden werden. Jede erzeugte HTML-Datei wird zum leichteren Navigieren mit entsprechenden „Knöpfen“ versehen.
- **Erste Informationen:** Nikos Drakos: „*All About LaTeX2HTML*“ (<http://cbl.leeds.ac.uk/nikos/tex2html/doc/latex2html/latex2html.html>)
- **Dokumentation:**
 - a) Nikos Drakos: „*The LaTeX2HTML Translator*“ (<ftp://ftp.rzg.mpg.de/pub/soft/latex2html/manual/manual.html>),
 - b) Nikos Drakos: „*The LaTeX2HTML Translator*“ (<ftp://ftp.rzg.mpg.de/pub/soft/latex2html/manual/manual.ps>),
 - c) Michael Goossens und Janne Saarela: „*From LaTeX to HTML and Back at CERN*“ (<http://wwwcn.cern.ch/asdoc/WWW/publications/l2hen/l2hen.html>)
- **Programmautor:** Nikos Drakos (<http://cbl.leeds.ac.uk/nikos/personal.html> und <mailto:nikos@mpn.com>)
- **Version:** 96.2 (=96.1 Rev. f); Dokumentation vom 07.06.1996
- **Quellen:**
 - a) <ftp://ftp.rzg.mpg.de/pub/soft/latex2html/sources/latex2html-96.1.revh.tar.gz> (Rev. h);
 - b) verschiedene spezielle Patch-Dateien in <ftp://ftp.rzg.mpg.de/pub/soft/latex2html/patches/>

- **Binär-Programme:** nicht verfügbar
- **Betriebssystem/Plattform:** UNIX
- **Benötigte Hilfsprogramme:** Perl 4.0 (Patch-Level mindestens 36) oder Perl 5.0, DBM oder NDBM (UNIX Database Management System), [dvipsk (mindestens 5.516)], [gs (mindestens 2.6.1)], LaTeX2e, [netpbm- oder pbmplus-Library], [math2html]
- **Anpaßbarkeit:**
 - a) Voreinstellungen abänderbar in `latex2html.config` und `.latex2html-init`
 - b) *vorsichtige* Änderungen der Benutzerbefehle in den Style-Dateien (`html.sty`, `htmllist.sty`, `url.sty` usw.) möglich
 - c) sinnvolle Änderungen in verschiedenen Perl-Dateien (z.B. in `german.perl` und `texdef.perl`) denkbar; Kenntnisse in Perl5 und von der Arbeitsweise des Programms sind erforderlich
 - d) sonstige Änderungen nur mit großem Aufwand (gute Perl5-Kenntnisse!)
- **Methode:** Konvertierung von LaTeX nach HTML
- **Bemerkung:** Es gibt eine Diskussionsliste zu LaTeX2HTML (`mailto:latex2html@mcs.anl.gov`). Um sich anzumelden, schicken Sie einen E-Mail-Brief mit „`subscribe latex2html`“ an `mailto:majordomo@mcs.anl.gov`.

Einige Programm-Eigenschaften

- behandelt (fast) alle LaTeX-Befehle korrekt
- Querverweise, Fußnoten, Inhaltsverzeichnisse, Abbildungsverzeichnisse, Tabellenverzeichnisse werden automatisch in Hypertext-Links übersetzt.
- versucht, auch unvollständige oder syntaktisch falsche LaTeX-Dokumente zu übersetzen
- bei der Übersetzung zu ignorierende LaTeX-Befehle können spezifiziert werden (in der Datei `.latex2html-init`)
- ebenso LaTeX-Befehle, die von LaTeX verarbeitet werden sollen
- Nicht-Übersetzbares (math. Formeln, Bilder, aufwendige Tabellen) werden in GIF-Bilder umgewandelt und an der korrekten Stelle automatisch eingefügt.
- eigene Befehle und Umgebungen werden korrekt behandelt
- unterstützt automatisch bestimmte Style-Dateien (`alltt`, `german`, `french`, `graphics`, `graphicx`, `epsfig`, `color` u.a)
- generiert automatisch konfigurierbare Navigations-„Knöpfe“
- Hypertext-Erweiterungen verfügbar durch das Package `html.sty`
- Links zu externen Zielen durch `\htmladdnormallink{text}{URL}`, `\htmladding{URL}` und `\htmladdnormallinkfoot{text}{URL}`
- bedingte Übersetzung durch `\begin{latexonly} ... \end{latexonly}`, `\latex{...}` bzw. `\begin{htmlonly} ... \end{htmlonly}`, `\html{...}` bzw. `\latexhtml{...}{...}`
- „Roh“-HTML durch `\begin{rawhtml} ... \end{rawhtml}`
- Querverweise mit `\hyperref{anker}{vortext}{nachtext}{label}` und `\label{label}`
- symbolische Referenzen zwischen veränderlichen Dokumenten mittels `\externallabels{...}{...}` und `\external{...}`
- Inline-Bilder durch `\htmlimage{...}` (und zahlreichen Optionen) steuerbar
- Änderungen zwischen verschiedenen Versionen des Dokuments in der LaTeX- und HTML-Version (`changebar.sty`) sichtbar

Aufruf von LaTeX2HTML:

`latex2html [optionen] datei`

Einige Optionen für LaTeX2HTML (von insgesamt über 40 Optionen):

- `-split zahl`
Aufteilung des HTML-Dokuments bis zur Stufe *zahl*
- `-link zahl`
Verweise auf untergeordnete Teile bis zur Tiefe *zahl*
- `-external_image`
an Stelle von Inline-Bildern externe Bilddateien benutzen
- `-t titel`
Titel des HTML-Stamm-Dokuments
- `-dir verzeichnis`
Ausgabe der zu generierenden HTML-Dateien im Verzeichnis *verzeichnis*
- `-prefix präfix`
Präfix, der den Namen aller zu produzierenden `.gif`-, `.pl`- und `.html`-Dateien vorangestellt wird
- `-address autor-adresse`
Jede HTML-Datei wird mit *autor-adresse* gekennzeichnet.
- `-no_images`
keine Inline-Bilder erzeugen
- `-images_only`
übrig gebliebene Inline-Bilder erzeugen
- `-show_section_numbers`
Abschnittsnummern generieren
- `-html_version version`
HTML-Version (2.0, 2.1, 2.2, 3.0, 3.1)
- `-h`
Auflistung aller Programmooptionen

Protokoll des Aufrufs (LaTeX-Datei nicht speziell vorbereitet)

```
word{g029}3367:>latex2html convtest
```

```
This is LaTeX2HTML Version 96.1-f (May 31, 1996) by  
Nikos Drakos,  
Computer Based Learning Unit, University of Leeds.
```

```
OPENING /tmp_mnt/home/hrza/g029/convtest.tex
```

```
Loading /opt/tex/latex2html/styles/texdefs.perl...
```

```
Loading /opt/tex/latex2html/styles/german.perl...
```

```
german style interface for LaTeX2HTML, 4.12.95
```

```
Reading ...
```

```
Processing macros .....+.
```

```
Translating
```

```
...0/5.....1/5.....2/5....3/5.....4  
/5....5/5.....
```

```
Doing footnotes ...
```

```
Doing section links .....
```

Es werden die folgenden Dateien im Unterverzeichnis convtest generiert:

```
convtest.css  
convtest.html  
footnode.html  
node1.html  
node2.html  
node3.html  
node4.html  
node5.html
```

Hyperlatex

- **Beschreibung:** Das Paket Hyperlatex ermöglicht es, eine Datei so vorzubereiten, daß sie von dem Hyperlatex-Konverter nach HTML gewandelt oder von LaTeX bearbeitet werden kann. Jede so erzeugte HTML-Datei kann zum Zwecke des leichteren Navigierens mit entsprechenden Menüs und „Knöpfen“ versehen werden.
- **Erste Informationen:** siehe „Dokumentation“
- **Dokumentation:**
 - a) Otfried Schwarzkopf: „*The Hyperlatex Markup Language*“
(hyperlatex.tex {hyperlatex.ind, hyperlatex.sty} in
ftp://ftp.dante.de/
tex-archive/support/hyperlatex/
Hyperlatex-2.0.5.tar.gz)
 - b) Otfried Schwarzkopf: „*The Hyperlatex Markup Language*“
(http://atoms.physics.wm.edu/~schwitr/hyperlatex/)
 - c) Otfried Schwarzkopf: *The Hyperlatex Story*; TUGboat 16, No. 2, p. 159 ff
- **Programmautor:** Otfried Schwarzkopf
(mailto:otfried@postech.ac.kr)
- **Version:** 2.0.6 (Januar 1997)
- **Quellen:**
 - a) Emacs-Makrodateien Hyperlatex-2.0.5/hyperlatex.el und
Hyperlatex-2.0.5/tabbing.el in ftp://ftp.dante.de/
tex-archive/support/hyperlatex/
Hyperlatex-2.0.5.tar.gz
 - b) LaTeX-Package hyperlatex.sty im gleichen Archiv
 - c) Hyperlatex-Dateien (siteinit.hlx usw.)
- **Binär-Programme:** nicht erhältlich
- **Betriebssystem/Plattform:** UNIX
- **Benötigte Hilfsprogramme:** gunzip, tar, LaTeX, emacs, csh, [ps2gif],
[dvipsk]

- **Anpaßbarkeit:**
 - a) Übersetzungsvorgang nach HTML in den `.el`-Dateien modifizierbar (gute LISP-Kenntnisse erforderlich)
 - b) Eingabesprache in `hyperlatex.sty` anpaßbar (gute LaTeX2e-Kenntnisse erforderlich)
 - c) Zusätzlich können die `.hlx`-Dateien, die vom Hyperlatex-Konverter gelesen werden, modifiziert werden.
- **Methode:** Hyperlatex wirkt hier wie eine vorgeschaltete Dokumentenbeschreibungssprache; Extraktion und Umwandlung durch einen speziellen Filter
- **Aufruf:**
`hyperlatex [-gif | -html | -dvi] datei`

mit
`-gif` nur GIF-Bilder erzeugen (u.a. LaTeX, ps2gif aufrufen)
`-html` HTML-Datei erzeugen
`-dvi` DVI-Datei erzeugen (LaTeX aufrufen)

Hyperlatex als Konverter

- Um Hyperlatex-Minimal-Funktionen zu bekommen:
`\usepackage{hyperlatex}`
- Bei der HTML-Generierung durch Hyperlatex werden die meisten LaTeX-Kommandos ausgewertet oder zumindest korrekt ignoriert.
- Wenn gar nichts oder nur sehr wenig geht: `\usepackage{emulate}`
- Eines der wenigen nicht korrekt behandelten Kommandos ist `\textsf` → entsprechende Anweisung (`\W\newcommand{\textsf}`) in der Präambel oder in der Datei `~/hyperlatex/emulate.hlx`.
- Einige LaTeX-Befehle haben für das HTML-Dokument eine leicht abgewandelte Bedeutung.
- Umlaute und Eszet umsetzen (z.B. mit `\"a`, Eszet mit `\ss{ }`)

Aufruf ohne Vorbereitung

```
word{g029}3096:hyperlatex>hyperlatex convtest
Hyperlatex formatting
/home/hrza/g029/hyperlatex/convtest.tex...
Parsing ...
Package "siteinit" inserted
Package "init" not found...
Package "article" inserted
Hyperlatex version 2.0.6
Package "german" not found...
Title of work is "Untitled"
Using filename "./convtest.html"
Parsing: Erster Abschnitt ...
>> Error: (error "Unknown command: textsf")
>> point at
>> {Sansserif}.
>> Und noch ein bi"schen anderer Text, so da"s
>> eine neue Zeile begonnen wird.
>> \item[bold
Hint: Try running Latex, it may give a better error
message.
```

Nach ersten Änderungen

```
word{g029}3099:hyperlatex>hyperlatex convtest
Hyperlatex formatting
/home/hrza/g029/hyperlatex/convtest.tex...
Parsing ...
Package "siteinit" inserted
Package "init" not found...
Package "article" inserted
Hyperlatex version 2.0.6
Package "german" not found...
Title of work is "Untitled"
Using filename "./convtest.html"
Parsing: Erster Abschnitt ...
Parsing: Zweiter Abschnitt ...
Parsing: Tabelle ...
Parsing: verbatim ...
Parsing: Footnotes ...
Formatting ...
Package "siteinit" inserted
Package "init" not found...
Package "article" inserted
Hyperlatex version 2.0.6
Package "german" not found...
Title of work is "Untitled"
Using filename "./convtest.html"
Formatting: Erster Abschnitt ...
Wrote /home/hrza/g029/hyperlatex/convtest.html
Formatting: Zweiter Abschnitt ...
Wrote /home/hrza/g029/hyperlatex/convtest_1.html
Formatting: Tabelle ...
Wrote /home/hrza/g029/hyperlatex/convtest_2.html
Formatting: verbatim ...
Wrote /home/hrza/g029/hyperlatex/convtest_3.html
Formatting: Footnotes ...
Wrote /home/hrza/g029/hyperlatex/convtest_4.html
Wrote /home/hrza/g029/hyperlatex/convtest_5.html
Hyperlatex formatting done.
```

html2latex.sed

- **Beschreibung:** `html2latex.sed` ist ein Sed-Programm zum Umsetzen von HTML-Konstrukten in LaTeX-Konstrukte.
- **Erste Informationen:** nicht verfügbar
- **Dokumentation:** nicht verfügbar
- **Programmautor:** nicht bekannt
- **Version:** nicht bekannt
- **Quellen:**
`http://www.w3.org/pub/WWW/Tools/TeX/html2latex.sed`
- **Binär-Programme:** nicht verfügbar
- **Betriebssystem/Plattform:** läuft unter jedem Betriebssystem, für das sed verfügbar ist
- **Benötigte Hilfsprogramme:** sed
- **Anpaßbarkeit:** Zuordnung der HTML-Elemente zu LaTeX-Konstrukten kann leicht in `html2latex.sed` geändert werden (sed-Kenntnisse erforderlich)
- **Methode:** Konvertierung von HTML nach LaTeX

Aufruf

```
sed -f html2latex.sed eingabe >ausgabe
```

Sed-Datei

(nach einigen Korrekturen; nicht die ursprüngliche Datei)

```
li\
\\documentclass{article}
li\
\\usepackage{german}
# Too bad there's no way to make sed ignore case!
/<[Xx][Mm][Pp]>/,/<.[Xx][Mm][Pp]>/b lit
/<.[Xx][Mm][Pp]>/b lit
/<[Ll][Ii][Ss][Tt][Ii][Nn][Gg]>/,/<.[Ll][Ii][Ss][Tt][Ii][Nn][Gg]>/b lit
/<.[Ll][Ii][Ss][Tt][Ii][Nn][Gg]>/b lit
/<[Pp][Rr][Ee]>/,/<.[Pp][Rr][Ee]>/b pre
/<.[Pp][Rr][Ee]>/b pre
# Stuff to ignore
s?<[Ii][Ss][Ii][Nn][Dd][Ee][Xx]>??
s?</[Aa][Dd][Dd][Rr][Ee][Ss][Ss]>??g
s?<[Nn][Ee][Xx][Tt][Ii][Dd][^>]*>??g
s?<[Hh][Tt][Mm][Ll]>??g
s?</[Hh][Tt][Mm][Ll]>??g
s?<[Hh][Ee][Aa][Dd]>??g
s?</[Hh][Ee][Aa][Dd]>??g
# translate HTML entities:
s?&szlig;?"s?g
s?&auml;?"a?g
s?&ouml;?"o?g
s?&uuml;?"u?g
s?&Auml;?"A?g
s?&Ouml;?"O?g
s?&Uuml;?"U?g
s?&nbsp;?<XYZ>?g
# character set translations for LaTeX special chars
s?&gt;.?>?g
s?&lt;.?<?g
s?\\?\\backslash ?g
s?{?\\{?g
s?}?\\}?g
s?%?\\%?g
s?\\$?\\$?g
s?&?\\&?g
s?#?\\#?g
s?_?\\_?g
```



```

s?~?\\~?g
s?\\^?\\^?g
s?<XYZ>?~?g
# Paragraph borders
s?<[Pp]>?\\par ?g
s?</[Pp]>??g
# Headings
s?<[Tt][Ii][Tt][Ll][Ee]>?\\title{?g
s?</[Tt][Ii][Tt][Ll][Ee]>?}?g
s?<[Hh]n>?\\part{?g
s?</[Hh]n>?}?g
s?<[Hh]1>?\\section*{?g
s?</[Hh][0-9]>?}?g
s?<[Hh]2>?\\subsection*{?g
s?<[Hh]3>?\\subsubsection*{?g
s?<[Hh]4>?\\subsubsection*{?g
s?<[Hh]5>?\\paragraph{?g
s?<[Hh]6>?\\subparagraph{?g
s?<[Bb][Oo][Dd][Yy]>?\\begin{document}?g
s?</[Bb][Oo][Dd][Yy]>?\\end{document}?g
# UL is itemize; and other markups
s?<[Uu][Ll]>?\\begin{itemize}?g
s?</[Uu][Ll]>?\\end{itemize}?g
s?<[Oo][Ll]>?\\begin{enumerate}?g
s?</[Oo][Ll]>?\\end{enumerate}?g
s?<[Ll][Ii]>?\\item ?g
s?<[Ee][Mm]>?\\emph{?g
s?</[Ee][Mm]>?}?g
s?<[Ss][Aa][Mm][Pp]>?\\verb|?g
s?</[Ss][Aa][Mm][Pp]>?|?g
# DL is description
s?<[Dd][Ll]>?\\begin{description}?g
s?</[Dd][Ll]>?\\end{description}?g
# closing delimiter for DT is first < or end of line
which ever comes first NO
#s?<[Dd][Tt]>\\([<]*\\)<?\\item[\\1]<?g
#s?<[Dd][Tt]>\\([<]*\\)$?\\item[\\1]?g
#s?<[Dd][Dd]>??g
s?<[Dd][Tt]>?\\item[<?g
s?<[Dd][Dd]>?]?g
# Other common SGML markup. this is ad-hoc
s?<sec[ab]>??
s?</sec[ab]>??g
# Italics
s?<it>\\([<]*\\)</it>?{\\it \\1 }?g
s?<[Ii]>?\\textit{?g
s?</[Ii]>?}?g
s?<[Bb]>?\\textbf{?g

```

```

s?</[Bb]>?}?g
s?<[Tt][Tt]>?\\texttt{?g
s?</[Tt][Tt]>?}?g
# Get rid of Anchors
:pre
s?<[Aa][^>]*>??g
s?</[Aa]>??g
# This is a subroutine in sed, in case you are not a
sed guru
: lit
s?<[Xx][Mm][Pp]>?\\begin{verbatim}?g
s?</[Xx][Mm][Pp]>?\\end{verbatim}?
s?<[Ll][Ii][Ss][Tt][Ii][Nn][Gg]>?\\begin{verbatim}?g
s?</[Ll][Ii][Ss][Tt][Ii][Nn][Gg]>?\\end{verbatim}?

```

Ergebnisdatei (LaTeX)

```
\documentclass{article}
\usepackage{german}

\title{Test f"ur Konverter}

\begin{document}

\section*{Erster Abschnitt}

In diesem Abschnitt werden einige
Textauszeichnungen f"ur
HTML~3.2 getestet, wie z.B.

\begin{itemize}
\item Das ist \texttt{Typewriter}.
\item Das ist SMALLCAPS.
\item Das ist \textit{Italic} und
      \textit{\textbf{Italic boldface}}.
\end{itemize}

\begin{description}
\item[<Sansserif
]Das ist Sansserif. Und noch ein
bi"schen anderer Text, so da"s
eine neue Zeile begonnen wird.
\item[<boldface
]Das ist \textbf{boldface}. Und noch ein
bi"schen anderer Text, so da"s
eine neue Zeile begonnen wird.
\end{description}

und Umlaute
("a"o"u"A"O"U"s) und

\begin{enumerate}
\item Das ist \emph{hervorgehobener} Text.
\item Das ist \verb|verbatim| Text.
\end{enumerate}

\section*{Zweiter Abschnitt}
```

```
\subsection*{Tabelle}
```

Hier wird eine kleine Tabelle
eingefügt:

```
<TABLE BORDER>
<TR><TD>abc <TD> x <TD> 120.55
<TR><TD>defg <TD> -- <TD> 11.37
<TR><TD>hijklmn <TD> xxx <TD> 12.95
</TABLE>
```

```
\subsection*{verbatim}
```

Ein kleiner verbatim gesetzter Text:

```
\begin{verbatim}
jhj hj h jh jh jh jj h
 k kl k lo klo      kk k k ol klo kl
klj kj j  jkl k
\end{verbatim}
```

```
\end{document}
```

html2latex

- **Beschreibung:** html2latex basiert auf dem NCSA-HTML-Parser (Mosaic) und konvertiert HTML-Auszeichnungen in LaTeX-Auszeichnungen.
- **Erste Informationen:** siehe „Dokumentation“
- **Dokumentation:** „*html2latex - convert HTML markup to LaTeX markup*“ (Datei `html2latex.html` in <http://www.vuw.ac.nz/non-local/software/html2latex-0.9c.tar.Z>)
- **Programmautor:** Nathan Torkington
(mailto:Nathan.Torkington@vuw.ac.nz)
- **Version:** 0.9c
- **Quellen:** <http://www.vuw.ac.nz/non-local/software/html2latex-0.9c.tar.Z>
- **Binär-Programme:** Es gab (gibt?) eine veraltete MS-DOS-Version in <ftp://ftp.dante.de/tex-archive/support/html2latex/html2ltx.zip>.
- **Betriebssystem/Plattform:** UNIX; Umstellung für andere Plattformen möglich
- **Benötigte Hilfsprogramme:** uncompress, tar, make, cc (oder gcc)
- **Anpaßbarkeit:** Zuordnung der HTML-Elemente zu LaTeX-Konstrukten kann in `html2latex.c` geändert werden (gute C-Kenntnisse erforderlich!); relativ starr
- **Methode:** Konvertierung von HTML nach LaTeX
- **Aufruf:**
`html2latex [optionen] [html-datei(en)]`

mit den *optionen*:

- | | |
|--------------------|--|
| -n | Abschnitte numerieren |
| -p | Seiten nach Titelseite und Inhaltsverzeichnis einfügen |
| -c | Inhaltsverzeichnis generieren |
| -s | auf Standardausgabe schreiben |
| -t <i>titel</i> | Titelseite mit <i>titel</i> generieren |
| -a <i>autor</i> | Titelseite mit <i>autor</i> generieren |
| -h <i>kopf</i> | Text <i>kopf</i> unmittelbar nach <code>\begin{document}</code> einfügen |
| -f <i>fuß</i> | Text <i>fuß</i> unmittelbar vor <code>\end{document}</code> einfügen |
| -o <i>optionen</i> | Optionen für <code>\documentstyle</code> |

Beispiel für einen Aufruf von `html2latex`

```
html2latex -t 'Test fuer Konverter' \  
-a 'Guenter.Partosch@hrz.uni-giessen.de' \  
-c -n -o '[german]{article}' \  
convtest.html
```

Ergebnis (LaTeX-Datei)

```
% This file was converted from HTML to LaTeX with
Nathan Torkington's
% html2latex program
% Version 0.9c
\documentstyle[german]{article}
\begin{document}
\title{Test fuer Konverter}
\author{Guenter.Partosch@hrz.uni-giessen.de}
\maketitle
\tableofcontents
\section{Erster Abschnitt}
In diesem Abschnitt werden einige
Textauszeichnungen f\uuml;r
HTML\&nbsp;3.2 getestet, wie z.B.
\begin{itemize}\item Das ist {\tt Typewriter}.
\item Das ist SMALLCAPS.
\item Das ist {\it Italic} und
        {\it {\bf Italic boldface}}.
\end{itemize}\begin{description}\item[Sansserif]
Das ist Sansserif. Und noch ein
bi\&szlig;chen anderer Text, so da\&szlig;
eine neue Zeile begonnen wird.
\item[boldface]
Das ist {\bf boldface}. Und noch ein
bi\&szlig;chen anderer Text, so da\&szlig;
eine neue Zeile begonnen wird.
\end{description}
und Umlaute
(\&auml;\&ouml;\&uuml;\&Auml;\&Ouml;\&Uuml;\&szlig;)
und
\begin{enumerate}\item Das ist {\it hervorgehobener}
Text.
\item Das ist {\tt verbatim} Text.
\end{enumerate}\section{Zweiter
Abschnitt}\subsection{Tabelle}
Hier wird eine kleine Tabelle
eingef\uuml;gt:
abc  x  120.55
defg --  11.37
hijklmn xxx 12.95
\subsection{verbatim}
Ein kleiner verbatim gesetzter Text:
```

```
\begin{verbatim}jhj hj h jh jh jh jj h
k kl k lo klo      kk k k ol klo kl
klj kj j  jkl k
\end{verbatim}\end{document}
```

Warum zusätzliche Dokumentenbeschreibungssprachen?

Warum Dokumentenbeschreibungssprachen?

Warum nicht Konvertierung LaTeX \leftrightarrow HTML?

- Zitat: „...conversion is not a solution to HTML authoring. A well written HTML document must differ from a printed copy. These differences cannot be recognized mechanically. Converted LaTeX can never be as readable as a document written in HTML...“
- Warum nicht SGML?
- Eine Beschreibungssprache ist auf jeden Fall sinnvoll, denn die logische Auszeichnung, nicht deren Ausgestaltung und Formatierung sollte im Vordergrund stehen.
- Es gibt zwar Konvertierer von SGML nach LaTeX und HTML, aber SGML als Beschreibungssprache ist zu aufwendig und zu schwierig zu erlernen.
- Fast alle Dokumente können mit einer deutlich einfacheren Beschreibungssprache beschrieben werden.
- Beschreibungssprachen der hier vorgestellten Art beschränken sich auf solche Sprachelemente, die sich leicht benutzen und abbilden lassen. Zu diesem Zweck werden die Sprachelemente auch entsprechend gestaltet.
- Eine solche Beschreibungssprache ist deutlich einfacher als SGML zu erlernen oder zu benutzen.

YODL

- **Beschreibung:** YODL (*Yet Oneother Document Language*) ist eine an C angelehnte Dokumenten-Beschreibungssprache, die mit Hilfe von Konvertern nach LaTeX, HTML und einigen anderen Formaten übertragen werden kann. YODL ist eine Makrosprache, bei der die Makroaufrufe in den zu verarbeitenden Eingabetext eingestreut werden.
- **Erste Informationen:**
 - a) Karel Kubat: „YODL Yet Oneother Document Language“ (<http://www.w3.org/pub/WWW/Tools/YODL.html>),
 - b) <ftp://ftp.icce.rug.nl/pub/unix/yodl.lsm>
 - c) Karel Kubat: <ftp://ftp.icce.rug.nl/pub/unix/yodl.README> im YODL-Format
- **Dokumentation:** Karel Kubat: „YODL VI.15“ <http://www.icce.rug.nl/docs/programs/yodl/yodl.html>
- **Programmautor:** Karel Kubat (<mailto:karel@icce.rug.nl>)
- **Version:** 1.15 (11.02.1997)
- **Quellen:**
 - a) <ftp://ftp.icce.rug.nl/pub/unix/yodl-1.15.tar.gz> (UNIX),
 - b) <ftp://ftp.icce.rug.nl/pub/unix/yodl4linux.tar.gz> (Linux)
- **Betriebssystem/Plattform:** UNIX; kann teilweise auf andere Betriebssysteme portiert werden
- **Benötigte Hilfsprogramme:** gzip, tar, make, gcc, bison, flex, install, dvips, LaTeX, groff, sed
- **Anpaßbarkeit:** eigene Makros können in der zu verarbeitenden YODL-Eingabedatei bzw. in speziellen Makrodateien vereinbart werden; sonstige Änderungen sehr aufwendig (Makefile und diverse .c-Dateien)
- **Methode:** vorgeschaltete Beschreibungssprache, die mit Hilfe entsprechender Filter (u.a. nach LaTeX bzw. HTML) übersetzt werden kann
- **Aufruf:** `yodl2html yodl-datei` bzw. `yodl2tex yodl-datei`

YODL als Beschreibungssprache

YODL verfügt u.a. über die folgenden Makro-Befehle:

`article(title) (author) (date)`: Beginn eines report-Dokuments; analog gibt es `book(title) (author) (date)` und `report(title) (author) (date)`

`abstract(text)`: Zusammenfassung in einem article- oder report-Dokument

`part(titel)`: Beginn eines neuen Teils in einem book-Dokument

`chapter(titel)`: Beginn eines Kapitels in book- oder report-Dokumenten

`sect(titel)`: Beginn eines Abschnitts; analog gibt es `subsect(titel)`, `subsubsect(titel)` und `subsubsubsect(titel)`

`npart(titel)`: part ohne Numerierung und Inhaltsverzeichniseintrag; analog gibt es `nchapter(titel)`, `nsect(titel)`, `nsubsect(titel)`, `nsubsubsect(titel)` und `nsubsubsubsect(titel)`

`bf(text)`: *text* wird fett formatiert

`em(text)`: *text* wird hervorgehoben

`tt(text)`: *text* wird in einer äquidistanten Schrift ausgegeben

`center(text)`: *text* wird zentriert

`description(list)`: description-Liste

`dit(itemname)`: Eintrag *itemname* in einer description-Liste

`enumerate(list)`: Aufzählungsliste

`eit()`: Eintrag in einer Aufzählungsliste

`itemize(list)`: itemize-Liste

`it()`: Eintrag in einer itemize-Liste

`quote(text)`: *text* wird als Zitat ausgegeben.

`affiliation(institution)`: Ausgabe von *institution* bei der Autor-Angabe

`clearpage()`: Beginn einer neuen Seite; im HTML-Modus Ausgabe einer Querlinie

`email(adresse)`: Im HTML-Modus wird `<HREF="mailto:adresse">` ausgegeben.

`figure(datei) (unterschrift) (label)`: Ausgabe des Bildes in der Datei *datei* mit dem beschreibenden Text *unterschrift* und dem Verweisziel *label*; YODL unterstellt die Dateinamensendung `.gif` im HTML-Modus bzw. `.ps` im LaTeX-Modus

`footnote(text)`: gibt *text* als Fußnote aus (ggf. in *text* runden Klammern)

`includefile(datei)`: fügt den Inhalt von *datei* ein

`label(label)`: vereinbart *label* als Ziel eines Verweises

`link(text) (label)`: in HTML-Dokumenten: Ausgabe des anklickbaren Textes *text*, der auf *label* zeigt

`nl()`: Zeilenwechsel

`ref(label)`: Verweis auf *label*

`url(anker) (ziel)`: in LaTeX-Dokumenten: Ausgabe von *anker*; in HTML-Dokumenten: Ausgabe des anklickbaren Ankers *anker*, der auf *ziel* zeigt

`verb(text)`: gibt *text* im Verbatim-Modus aus; keine weiteren Zeichenumsetzung oder Expandierung durch YODL

`verbatiminclude(datei)`: gibt den Inhalt der Datei *datei* im Verbatim-Modus aus; wird von YODL nicht weiterverarbeitet

`latexcommand(cmd)`: Ausgabe von *cmd* im LaTeX-Modus; wird nicht weiter von YODL untersucht

`latexdocumentstyle(style)`: fügt im LaTeX-Modus *style* in der Anweisung `\documentstyle` ein

`latexoptions(optionen)`: im LaTeX-Modus: fügt *optionen* in die `\documentstyle`-Anweisung ein

`latexlayoutcmds(latex-kommandos)`: *latex-kommandos* werden im LaTeX-Modus unmittelbar hinter `\documentstyle` ausgegeben.

`standardlayout()`: im LaTeX-Modus: kein Einzug am Absatzbeginn, dafür aber etwas größere Abstände zwischen zwei Absätzen

`whenlatex (text)` : *text* wird nur im LaTeX-Modus ausgegeben; *text* wird ggf. von YODL weiter expandiert

`htmlbodyopt (attribut) (wert)` : versieht das HTML-Element `<BODY>` mit der zusätzlichen Angabe *attribut=wert*

`htmlcommand (cmd)` : Ausgabe des HTML-Textes *cmd* im HTML-Modus; *cmd* wird von YODL nicht weiterverarbeitet

`htmlnewfile ()` : im HTML-Modus Beginn einer neuen HTML-Datei

`htmltag (tagname) (start)` : Ausgabe des HTML-Elements *tagname* (spitze Klammern werden eingefügt); Ist *start=0*, so wird *tagname* ein / vorangestellt.

`plainhtml (titel)` : Beginn eines einfachen HTML-Dokuments (ähnlich `article`)

`whenhtml (text)` : *text* wird nur im HTML-Modus ausgegeben; *text* wird ggf. von YODL weiter expandiert.

In YODL können eigene Makros (z.B. in der Datei `.../yodl/lib/shared.yo`) vereinbart werden:

```
DEFINEMACRO (i) (1) (\
  whenlatex (latexcommand ({\it } ARG1+latexcommand ())) \
  whenhtml (htmlcommand (<I> ARG1+htmlcommand (</I>))) )
```

```
DEFINEMACRO (sf) (1) (\
  whenlatex (latexcommand ({\sc } ARG1+latexcommand ())) \
  whenhtml (htmlcommand (<!-- --> ARG1+htmlcommand \
    (<!-- -->))) )
```

```
DEFINEMACRO (sc) (1) (\
  whenlatex (latexcommand ({\sf } ARG1+latexcommand ())) \
  whenhtml (htmlcommand (<!-- --> ARG1+htmlcommand \
    (<!-- -->))) )
```

YODL-Eingabedatei

```
affiliation(Hochschulrechenzentrum Gie\ssen)
```

```
mailto(Guenter.Partosch@hrz.uni-giessen.de)
```

```
article(Test f\"ur Konverter)  
      (G\"unter Partosch)  
      (20.02.1997)
```

```
sect(Erster Abschnitt)
```

In diesem Abschnitt werden einige Textauszeichnungen f\"ur HTML getestet, wie z.B.

```
itemize(  
  it() Das ist tt(Typewriter).  
  it() Das ist sc(SmallCaps).  
  it() Das ist i(Italic) und  
    i(bf(Italic boldface)).  
)
```

```

description(
  dit(Sansserif) Das ist sf(Sansserif). Und noch ein bi\sschen
    anderer Text, so da\ss eine neue Zeile begonnen
    wird.
  dit(boldface) Das ist bf(boldface). Und noch ein bi\sschen
    anderer Text, so da\ss eine neue Zeile begonnen
    wird.
)

```

und Umlaute (\ "a\ "o\ "u\ "A\ "O\ "U\ss) und

```

enumerate(
  eit() Das ist em(hervorgehobener) Text.
  eit() Das ist verb(verbatim) Text.
)

```

sect(Zweiter Abschnitt)

subsect(Tabelle)

Hier wird eine kleine Tabelle eingef\ "ugt:

```

\begin{tabular}{||l|c|r||}
\hline
abc & x & 120.55 \\
defg & -- & 11.37 \\
hijklmn & xxx & 12.95 \\
\hline
\end{tabular}

```

```
subsect(verbatim)
```

Ein kleiner verbatim gesetzter Text:

```
verb(  
jhj hj h jh jh jh jj h  
  k kl k lo klo      kk k k ol klo kl  
klj kj j  jkl k  
)
```

YODL-Ergebnisdatei (LaTeX)

```
\documentstyle{article}
\hfuzz=4pt % don't fuss when less than 4pt
\input{/home/hrz/g029/yodl/lib/yodl/xlatin1.tex}
\setlength{\parskip}{3mm} % height between par
\setlength{\parindent}{0mm} % no par indent

\title{Test f\"{u}r Konverter}
\author{G\"{u}nter Partosch\ \ \ Hochschulrechenzentrum Gie\ss{}en}
\date{20.02.1997}
\begin{document}
\maketitle

\tableofcontents
\clearpage

\section{Erster Abschnitt}
```

In diesem Abschnitt werden einige Textauszeichnungen f\"{u}r HTML getestet, wie z.B.


```

\begin{itemize}
  \item Das ist {\tt Typewriter}.
  \item Das ist {\sf SmallCaps}.
  \item Das ist {\it Italic} und
        {\it {\bf Italic boldface}}.
\end{itemize}

\begin{description}
  \item[Sansserif] Das ist {\sc Sansserif}. Und noch ein bi\ss{}chen
                    anderer Text, so da\ss{} eine neue Zeile begonnen
                    wird.
  \item[boldface] Das ist {\bf boldface}. Und noch ein bi\ss{}chen
                    anderer Text, so da\ss{} eine neue Zeile begonnen
                    wird.
\end{description}

und Umlaute ({\a}{\o}{\u}{\A}{\O}{\U}\ss{}) und

\begin{enumerate}
  \item Das ist {\em hervorgehobener} Text.
  \item Das ist \begin{verbatim}
verbatim
\end{verbatim}
Text.
\end{enumerate}

\section{Zweiter Abschnitt}

\subsection{Tabelle}

```

Hier wird eine kleine Tabelle eingefügt:

```
{\tt
$\backslash$begin{\tt\{ }tabular{\tt\{ }\{\tt$\mid$\}\{\tt$\mid$\}l{\tt$\mid$\}c{\tt
$\mid$\}r{\tt$\mid$\}\{\tt$\mid$\}\{\tt\{ }\}
{\tt $\backslash$}hline
abc {\tt\&} x {\tt\&} 120.55 {\tt $\backslash$}defg {\tt\&} -- {\tt\&} 11.37{\tt
$\backslash$}hijklmn {\tt\&} xxx {\tt\&} 12.95{\tt $\backslash$}\{\tt
$\backslash$}hline
{\tt $\backslash$}end{\tt\{ }tabular{\tt\{ }\}
```

```
\subsection{verbatim}
```

Ein kleiner verbatim gesetzter Text:

```
\begin{verbatim}
```

```
jhj hj h jh jh jh jj h
 k kl k lo klo      kk k k ol klo kl
klj kj j   jkl k
```

```
\end{verbatim}
```

```
\end{document}
```

YODL-Ergebnisdatei (HTML, Ausschnitt)

<p>In diesem Abschnitt werden einige Textauszeichnungen für HTML getestet, wie z.B.

<p>

- Das ist <code>Typewriter</code>.
- Das ist <!-- -->SmallCaps<!-- -->.
- Das ist <I>Italic</I> und
 <I>Italic boldface</I>.

<p>

 Sansserif Das ist <!-- -->Sansserif<!-- -->. Und noch ein
bißchen

 anderer Text, so daß eine neue Zeile begonnen
 wird.

 boldface Das ist boldface. Und noch ein
bißchen

 anderer Text, so daß eine neue Zeile begonnen
 wird.


```
<p>und Umlaute (&auml;&ouml;&uuml;&Auml;&Ouml;&Uuml;&szlig;) und  
<p><ol>  
  <li> Das ist <em>hervorgehobener</em> Text.  
  <li> Das ist <listing>  
verbatim  
</listing>  
Text.  
</ol>
```

Hyperlatex als Beschreibungssprache

- `\htmladdress{adresse}`
- `\htmltitle{titel}`
- `\htmldirectory{verzeichnis}`
- `\htmlname{basisname}`
- `\htmllevel{level}` mit `level=html2, html32`
- `\htmldepth{tiefe}`
- `\htmlmenu{tiefe}`
- `\htmlautomenu{tiefe}`
- **Wie bisher:** `\title{...}`, `\author{...}`, `\date{...}`, `\maketitle`, `\thanks`, abstract-Umgebung
- **Gliederung:** `\chapter{...}`, `\section{...}`
auch `\htmheading{...}`
- **Umgebungen:** `quote`, `quotation`, `verse`, `center`, `itemize`, `enumerate`, `description`
- **Zusätzliche Umgebungen:** `blockquote`, `menu`, `example`;
brauchbar `\setlength{\exampleindent}{einzug}`
- **Vereinbarung neuer Befehle oder Umgebungen möglich:**
`\newcommand`, `\newenvironment`, ...
- **Font-Befehle:** `\textbf{...}`, `\textit{...}`, `\textsc{...}`,
`\texttt{...}`, `\underline`, `\bf`, `\it`, `\tt`, `\em`, `\emph{...}`
- **Zusätzliche Font-Befehle:** `\cit`, `\code`, `\dfn`, `\file`, `\kbd`, `\samp`,
`\strong`, `\var`
- **Font-Größen:** `\large`, `\normalfont`, ...
brauchbar auch `\html{basefont size=x} x=1,..., 7`
- **Bedingte Übersetzung:**

<code>\texonly{...}</code>	<code>\T</code>	<code>\begin{iftex}</code>
...		
<code>\end{iftex}</code>		
<code>\htmlonly{...}</code>	<code>\W</code>	<code>\begin{ifhtml}</code>
...		
<code>\end{html}</code>		

`\texorhtml{...}{...}`
- **„Rohes“ HTML:**
`\htmlsym{...}`
`\html{tag}` und `\html*{tag}`
`\begin{rawhtml}... \endhtml{rawhtml}`

Beispiel für `\htmlsym`:

```
\T \newcommand{\onequarter}{\$1/4\$}
```

```
\W \newcommand{\onequarter}{\htmlsym{##188}}
```

- **Querverweise:** `\label{label}`, `\link{anker}{label}`
`\begin{Label}{label}... \end{Label}`
- **Externe Querverweise (URLs):** `\xlink{anker}{URL}`,
`\xlink{anker}[gedruckt]{URL}`
- **Bilder:** `\htmlimage[...]{datei}`
- **GIF-Bilder:**
`\begin{gif}[tags][auflösung][font-auflösung]{name}`
...TeX-Material ...
`\end{gif}`
- **Querlinien:** `\htmlrule[...]`
- **Tabellen:** einfache tabular-Gebilde werden korrekt übertragen
Trick für Tabellenränder: `\htmlattributes*{TABLE}{BORDER}`
unmittelbar vor `\begin{tabular}{...}`

Hyperlatex als Beschreibungssprache; geänderte Eingabe

```
\documentclass[a4paper]{article}
\usepackage{hyperlatex}
\usepackage{emulate}
\usepackage{german}

% LaTeX2e-Datei zum Testen verschiedener Konverter
\title{Test f{"u}r Konverter}
\author{G{"u}nter Partosch}
\thanks{Hochschulrechenzentrum Gie{\ss}en}}
\date{\today}

\newcommand{\HTML}{HTML~3.2}
\htmladdress{\xlink{\html{CODE}}%
Guenter.Partosch@hrz.uni-giessen.de\html{/CODE}}%
{mailto:Guenter.Partosch@hrz.uni-
giessen.de}}
\htmltitle{Test fuer Konverter}
\htmlattributes*{BODY}{COLOR="#FF0000"}

\begin{document}
\maketitle

\section{Erster Abschnitt}

In diesem Abschnitt werden einige
Textauszeichnungen f{"u}r {\HTML} getestet,
wie z.B.

\begin{itemize}
\item Das ist \texttt{Typewriter}.
\item Das ist \textsc{SmallCaps}.
\item Das ist \textit{Italic} und
\textit{\textbf{Italic boldface}}.
\end{itemize}
...
```

Hyperlatex als Beschreibungssprache; Aufrufprotokoll

```
word{g029}3135:hyperlatex>hyperlatex convtest
Hyperlatex formatting
/home/hrza/g029/hyperlatex/convtest.tex...
Parsing ...
Package "siteinit" inserted
Package "init" not found...
Package "article" inserted
Hyperlatex version 2.0.5
Package "german" not found...
Title of work is "Test fuer Konverter"
Using filename "./convtest.html"
Parsing: Erster Abschnitt ...
Parsing: Zweiter Abschnitt ...
Parsing: Tabelle ...
Parsing: verbatim ...
Parsing: Footnotes ...
Formatting ...
Package "siteinit" inserted
Package "init" not found...
Package "article" inserted
Hyperlatex version 2.0.5
Package "german" not found...
Title of work is "Test fuer Konverter"
Using filename "./convtest.html"
Formatting: Erster Abschnitt ...
Wrote /home/hrza/g029/hyperlatex/convtest.html
Formatting: Zweiter Abschnitt ...
Wrote /home/hrza/g029/hyperlatex/convtest_1.html
Formatting: Tabelle ...
Wrote /home/hrza/g029/hyperlatex/convtest_2.html
Formatting: verbatim ...
Wrote /home/hrza/g029/hyperlatex/convtest_3.html
Formatting: Footnotes ...
Wrote /home/hrza/g029/hyperlatex/convtest_4.html
Wrote /home/hrza/g029/hyperlatex/convtest_5.html
Hyperlatex formatting done.
```


Ergebnisdatei CONVTEST . HTML

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 3.2 Draft//EN">
<HTML>
<!-- HTML file produced from file: convtest.tex --
  -- using Hyperlatex v 2.0.6 (c) Otfried Schwarzkopf -->
<HEAD>
<TITLE>Test fuer Konverter</TITLE>
</HEAD><BODY COLOR="#FF0000">

    <H1 ALIGN=CENTER>Test fuer Konverter</H1>
    <H2 ALIGN=CENTER>Guenter Partosch<A HREF="convtest_5.html#2">(2)</A></H2>
    <H3 ALIGN=CENTER>February 22, 1997</H3>
    <MENU>
<LI><A HREF="convtest_1.html">Erster Abschnitt</A>
<LI><A HREF="convtest_2.html">Zweiter Abschnitt</A>
<LI><A HREF="convtest_5.html">Footnotes</A>
</MENU>
<HR ><ADDRESS><A HREF="mailto:Guenter.Partosch@hrz.uni-
giessen.de"><CODE>Guenter.Partosch@hrz.uni-giessen.de</CODE></A></ADDRESS><BR>
</BODY></HTML>
```

Ergebnisdatei CONVTEST_1.HTML

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 3.2 Draft//EN">
<HTML>
<!-- HTML file produced from file: convtest.tex --
  -- using Hyperlatex v 2.0.6 (c) Otfried Schwarzkopf -->
<HEAD>
<TITLE>Test fuer Konverter -- Erster Abschnitt</TITLE>
</HEAD><BODY>
<IMG ALT="" SRC="previous.xbm">
  <A HREF="convtest.html"><IMG ALT="" SRC="up.xbm"></A>
  <A HREF="convtest_2.html"><IMG ALT="" SRC="next.xbm"></A><BR><B>Zurück zum
Anfang </B><A HREF="convtest.html">Top</A><BR><B>Vorwärts zu Abschnitt </B><A
HREF="convtest_2.html">Zweiter Abschnitt</A><BR><HR >
<H1>Erster Abschnitt</H1>
<P>In diesem Abschnitt werden einige
Textauszeichnungen für HTML 3.2 getestet,
wie z.B.
<UL><LI>Das ist <TT>Typewriter</TT>.
<LI>Das ist SMALLCAPS.
<LI>Das ist <I>Italic</I> und
      <I><B>Italic boldface</B></I>.
</UL>
```

```

<DL><DT><B>Sansserif</B><DD> Das ist Sansserif.
Und noch ein bi&szlig;chen anderer Text, so da&szlig;
eine neue Zeile begonnen wird.
<DT><B>boldface</B><DD> Das ist <B>boldface</B>.
Und noch ein bi&szlig;chen anderer Text, so da&szlig;
eine neue Zeile begonnen wird.
</DL>
<P>und Umlaute (&auml;&ouml;&uuml;&Auml;&Ouml;&Uuml;&szlig;) und
<OL><LI>Das ist <EM>hervorgehobener</EM> Text.
<LI>Das ist <CODE>verbatim</CODE> Text.
</OL>
<HR ><ADDRESS><A HREF="mailto:Guenter.Partosch@hrz.uni-
giessen.de"><CODE>Guenter.Partosch@hrz.uni-giessen.de</CODE></A></ADDRESS><BR>
<IMG ALT="" SRC="previous.xbm">
    <A HREF="convtest.html"><IMG ALT="" SRC="up.xbm"></A>
    <A HREF="convtest_2.html"><IMG ALT="" SRC="next.xbm"></A></BODY></HTML>

```

Ergebnisdatei CONVTEST_2.HTML

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 3.2 Draft//EN">
<HTML>
<!-- HTML file produced from file: convtest.tex --
  -- using Hyperlatex v 2.0.6 (c) Otfried Schwarzkopf -->
<HEAD>
<TITLE>Test fuer Konverter -- Zweiter Abschnitt</TITLE>
</HEAD><BODY>
<A HREF="convtest_1.html"><IMG ALT="" SRC="previous.xbm"></A>
  <A HREF="convtest.html"><IMG ALT="" SRC="up.xbm"></A>
  <A HREF="convtest_5.html"><IMG ALT="" SRC="next.xbm"></A><BR><B>Zurück zum
Abschnitt </B><A HREF="convtest_1.html">Erster Abschnitt</A><BR><B>Zurück zum
Anfang </B><A HREF="convtest.html">Top</A><BR><B>Vorw&auml;rts zu Abschnitt </B><A
HREF="convtest_5.html">Footnotes</A><BR><HR >
<H1>Zweiter Abschnitt</H1>
<MENU>
<LI><A HREF="convtest_3.html">Tabelle</A>
<LI><A HREF="convtest_4.html">verbatim</A>
</MENU>
<HR ><ADDRESS><A HREF="mailto:Guenter.Partosch@hrz.uni-
giessen.de"><CODE>Guenter.Partosch@hrz.uni-giessen.de</CODE></A></ADDRESS><BR>
<A HREF="convtest_1.html"><IMG ALT="" SRC="previous.xbm"></A>
  <A HREF="convtest.html"><IMG ALT="" SRC="up.xbm"></A>
  <A HREF="convtest_5.html"><IMG ALT="" SRC="next.xbm"></A></BODY></HTML>
```

Ergebnisdatei CONVTEST_3.HTML

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 3.2 Draft//EN">
<HTML>
<!-- HTML file produced from file: convtest.tex --
-- using Hyperlatex v 2.0.6 (c) Otfried Schwarzkopf -->
<HEAD>
<TITLE>Test fuer Konverter -- Tabelle</TITLE>
</HEAD><BODY>
<IMG ALT="" SRC="previous.xbm">
  <A HREF="convtest_2.html"><IMG ALT="" SRC="up.xbm"></A>
  <A HREF="convtest_4.html"><IMG ALT="" SRC="next.xbm"></A><BR><B>Zur&uuml;ck zum
Anfang </B><A HREF="convtest_2.html">Zweiter Abschnitt</A><BR><B>Vorw&auml;rts zu
Abschnitt </B><A HREF="convtest_4.html">verbatim</A><BR><HR >
<H2>Tabelle</H2>
<P>Hier wird eine kleine Tabelle eingef&uuml;gt:
<TABLE><TR><TD COLSPAN="1" ALIGN="LEFT">

abc </TD><TD COLSPAN="1" ALIGN="CENTER"> x </TD><TD COLSPAN="1" ALIGN="RIGHT"> 120.55
</TD></TR>
<TR><TD COLSPAN="1" ALIGN="LEFT">
defg </TD><TD COLSPAN="1" ALIGN="CENTER"> - </TD><TD COLSPAN="1" ALIGN="RIGHT">
11.37</TD></TR>
<TR><TD COLSPAN="1" ALIGN="LEFT">
hijklmn </TD><TD COLSPAN="1" ALIGN="CENTER"> xxx </TD><TD COLSPAN="1" ALIGN="RIGHT">
12.95</TD></TR>
<TR><TD COLSPAN="1" ALIGN="LEFT">
</TD></TR></TABLE>
```

```
<HR ><ADDRESS><A HREF="mailto:Guenter.Partosch@hrz.uni-  
giessen.de"><CODE>Guenter.Partosch@hrz.uni-giessen.de</CODE></A></ADDRESS><BR>  
<IMG ALT="" SRC="previous.xbm">  
  <A HREF="convtest_2.html"><IMG ALT="" SRC="up.xbm"></A>  
  <A HREF="convtest_4.html"><IMG ALT="" SRC="next.xbm"></A></BODY></HTML>
```

Ergebnisdatei CONVTEST_4.HTML

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 3.2 Draft//EN">
<HTML>
<!-- HTML file produced from file: convtest.tex --
  -- using Hyperlatex v 2.0.6 (c) Otfried Schwarzkopf -->
<HEAD>
<TITLE>Test fuer Konverter -- verbatim</TITLE>
</HEAD><BODY>
<A HREF="convtest_3.html"><IMG ALT="" SRC="previous.xbm"></A>
  <A HREF="convtest_2.html"><IMG ALT="" SRC="up.xbm"></A>
  <IMG ALT="" SRC="next.xbm"><BR><B>Zur&uuml;ck zum Abschnitt </B><A
HREF="convtest_3.html">Tabelle</A><BR><B>Zur&uuml;ck zum Anfang </B><A
HREF="convtest_2.html">Zweiter Abschnitt</A><BR><HR >
<H2>verbatim</H2>
<P>Ein kleiner verbatim gesetzter Text:
<PRE>
jhj hj h jh jh jj h
 k kl k lo klo      kk k k ol klo kl
klj kj j  jkl k
</PRE>
<HR ><ADDRESS><A HREF="mailto:Guenter.Partosch@hrz.uni-
giessen.de"><CODE>Guenter.Partosch@hrz.uni-giessen.de</CODE></A></ADDRESS><BR>
<A HREF="convtest_3.html"><IMG ALT="" SRC="previous.xbm"></A>
  <A HREF="convtest_2.html"><IMG ALT="" SRC="up.xbm"></A>
  <IMG ALT="" SRC="next.xbm"></BODY></HTML>
```

Ergebnisdatei CONVTEST_5.HTML

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 3.2 Draft//EN">
<HTML>
<!-- HTML file produced from file: convtest.tex --
-- using Hyperlatex v 2.0.6 (c) Otfried Schwarzkopf -->
<HEAD>
<TITLE>Test fuer Konverter -- Footnotes</TITLE>
</HEAD><BODY>
<A HREF="convtest_2.html"><IMG ALT="" SRC="previous.xbm"></A>
  <A HREF="convtest.html"><IMG ALT="" SRC="up.xbm"></A>
  <IMG ALT="" SRC="next.xbm"><BR><B>Zur&uuml;ck zum Abschnitt </B><A
HREF="convtest_2.html">Zweiter Abschnitt</A><BR><B>Zur&uuml;ck zum Anfang </B><A
HREF="convtest.html">Top</A><BR><HR >
<H1>Footnotes</H1>
<DL>
<A NAME="1">&#160;</A><DT><B>(1)</B><DD>Hochschulrechenzentrum Gie&szlig;en

<A NAME="2">&#160;</A><DT><B>(2)</B><DD>Hochschulrechenzentrum Gie&szlig;en

</DL>
<HR ><ADDRESS><A HREF="mailto:Guenter.Partosch@hrz.uni-
giessen.de"><CODE>Guenter.Partosch@hrz.uni-giessen.de</CODE></A></ADDRESS><BR>
<A HREF="convtest_2.html"><IMG ALT="" SRC="previous.xbm"></A>
  <A HREF="convtest.html"><IMG ALT="" SRC="up.xbm"></A>
  <IMG ALT="" SRC="next.xbm"></BODY></HTML>
```