

# PDF information and navigation elements with hyperref, pdf $\TeX$ , and thumbpdf

Heiko Oberdiek  
[oberdiek@uni-freiburg.de](mailto:oberdiek@uni-freiburg.de)

13rd October 1999

## Abstract

Additional possibilities for information and navigation through paper-less on-line documents that offer the PDF format.

Start

# Subjects

- General document information `hyperref`
- Bookmarks/outlines `hyperref`
- Thumbnails `thumbpdf`

Full-screen

Back

Next

# General Document Information

Application: search engines.

## hyperref options:

Title: [pdftitle](#)  
Subject: [pdfsubject](#)  
Author: [pdfauthor](#)  
Keywords: [pdfkeywords](#)  
Creator: [pdfcreator](#)  
Producer: [pdfproducer](#)

## Times:

Created: [/CreationDate](#)  
Modified: [/ModDate](#)

## Informations:

File name, file size  
PDF version  
Optimized(linearized)

# Setting the general document information

```
\usepackage{hyperref}
\hypersetup{%
  pdftitle      = {PDF information and navigation elements},
  pdfsubject    = {Slides for talk at EuroTeX'99 in Heidelberg},
  pdfkeywords   = {pdf, hyperref, bookmarks, thumbnails},
  pdfauthor     = {\textcopyright\ Heiko Oberdiek},
  pdfcreator    = {\LaTeX\ with package \flqq hyperref\frqq},
  pdfproducer   = {pdfTeX-0.\the\pdftexversion\pdftexrevision},
}
\pdfinfo{/CreationDate (D:19990909000000-01'00')}
```

# PDF strings

- Examples, where PDF strings are used:
  - Bookmark names, information dictionaries, text annotations
- PDF strings follow the PostScript rules:
  - Delimited by parentheses: `(This is a string)`
  - Escape sequences for white space, `\(, \), \\`;  
octal notation, e. g.: `'äöüß' = (\344\366\374\337)`
- **PDFDocEncoding**: 8bit, superset of ISO Latin1  
**Unicode**: 16bit, AcrobatReader version  $\geq 4$

# Digestion of T<sub>E</sub>X

## Eyes

- Reading input lines
- Catcodes are set

## Mouth

- Forming tokens
- Expansion of macros

## Stomach

- Assignments
- Typesetting

Full-screen

Back

Next

# Differences between mouth and stomach

	<b>mouth</b>	<b>stomach</b>
“zero”:	<code>\empty</code>	<code>\relax</code>
{...}:	parameter	group
variable:	read and check	assignments
fonts:	–	different fonts and encodings ligatures
Commands:	<code>\string, \number</code> <code>\if, \the</code>	<code>\def, \hbox, \$</code> <code>\begingroup, \special</code>

# T<sub>E</sub>X into PDF strings

- Reusing the argument of section commands (T<sub>E</sub>X string) for the bookmarks (PDF strings).
- Limitations because of missing stomach digestion:
  - No manipulation of boxes, no mathematics, no colors, . . .
  - No change of fonts or encodings, no ligatures.
  - No assignments (`\xspace` uses `\futurelet`).
  - Unexpandable commands appear verbatim.

## Font encoding mechanism

plain- $\text{T}_E\text{X}$ : `\ss = \char"19`  $\Rightarrow$   $\beta$

$\text{L}_A\text{T}_E\text{X } 2_\epsilon$ : `\fontencoding{OT1}\selectfont\char"19`  $\Rightarrow$   $\beta$

`\fontencoding {T1}\selectfont\char"19`  $\Rightarrow$   $\text{ı}$

`\ss = \OT1-cmd \ss \OT1\ss`  $\Rightarrow$   $\beta$

The font slot positions depend on the encoding:

`\OT1\ss = \char"19`

`\T1\ss = \char"FF`

`\?\ss = \UseTextSymbol{OT1}\ss`

`\OT1\ss = \csname OT1\string\ss\endcsname`

# PD1 encoding (PDFDocEncoding)

- Most names: `\text<glyph name>`  
Examples:  $\text{\textcircled{R}}$  = `\textregistered`,  $\text{\textasciitilde}$  = `\textasciitilde`
- Traditional T<sub>E</sub>X and L<sup>A</sup>T<sub>E</sub>X names.  
Examples:  $\text{\text{Æ}}$  = `\text{Æ}`,  $\text{\text{ß}}$  = `\text{ss}`
- Accented characters.  
Examples:  $\text{\text{Ü}}$  = `\text{"U}`,  $\text{\text{ç}}$  = `\text{c c}`,  $\text{\text{Å}}$  = `\text{r A}` ( $\leftarrow$  `\text{AA}`)
- Escape octal sequences are the result: `\textmu`  $\Rightarrow$  `\265`

## PU encoding (Unicode)

- ☹ Characters having catcodes 11 (letter) and 12 (other):  
Checking each token, high memory and time consumption.
- ☹ Ligatures: same as above.
- ☹ Direct commands:  
Easier to implement, but each command must be redefined each time.
- ☺ Commands of  $\text{\LaTeX} 2_{\epsilon}$ 's fontencoding mechanism:  
Smallest memory consumption and fastest method.  
This method is used by package `hyperref`.

# Examples of Unicode bookmarks

Greek:

- +... 003/210..217: \_E\_ \_H\_ \_I\_ \_O\_ \_Y\_ \_Ω\_
- +... 003/220..227: \_i\_ \_A\_ \_B\_ \_Γ\_ \_Δ\_ \_E\_ \_Z\_ \_H\_
- +... 003/230..237: \_Θ\_ \_I\_ \_K\_ \_Λ\_ \_M\_ \_N\_ \_Ξ\_ \_O\_
- +... 003/240..247: \_Π\_ \_Ρ\_ \_Σ\_ \_Τ\_ \_Υ\_ \_Φ\_ \_Χ\_
- +... 003/250..257: \_Ψ\_ \_Ω\_ \_i\_ \_ÿ\_ \_á\_ \_é\_ \_ή\_ \_í\_
- +... 003/260..267: \_ü\_ \_α\_ \_β\_ \_γ\_ \_δ\_ \_ε\_ \_ζ\_ \_η\_
- +... 003/270..277: \_θ\_ \_ι\_ \_κ\_ \_λ\_ \_μ\_ \_ν\_ \_ξ\_ \_ο\_
- +... 003/300..307: \_π\_ \_ρ\_ \_ς\_ \_σ\_ \_τ\_ \_υ\_ \_φ\_ \_χ\_
- +... 003/310..317: \_ψ\_ \_ω\_ \_i\_ \_ü\_ \_ó\_ \_ú\_ \_ώ\_ \_.

Cyrillic:

- ... Unicode high byte: 004
- +... 004/000..007: \_.\_ \_Ё\_ \_Ъ\_ \_Г\_ \_Є\_ \_S\_ \_I\_ \_i\_
- +... 004/010..017: \_J\_ \_Л\_ \_Ь\_ \_Ң\_ \_К\_ \_.\_ \_Ў\_ \_Ц\_
- +... 004/020..027: \_А\_ \_Б\_ \_В\_ \_Г\_ \_Д\_ \_Е\_ \_Ж\_ \_З\_
- +... 004/030..037: \_И\_ \_Й\_ \_К\_ \_Л\_ \_М\_ \_Н\_ \_О\_ \_П\_
- +... 004/040..047: \_Р\_ \_С\_ \_Т\_ \_У\_ \_Ф\_ \_Х\_ \_Ц\_ \_Ч\_
- +... 004/050..057: \_Ш\_ \_Щ\_ \_Ъ\_ \_Ы\_ \_Ь\_ \_Э\_ \_Ю\_ \_Я\_
- +... 004/060..067: \_а\_ \_б\_ \_в\_ \_г\_ \_д\_ \_е\_ \_ж\_ \_з\_
- +... 004/070..077: \_и\_ \_й\_ \_к\_ \_л\_ \_м\_ \_н\_ \_о\_ \_п\_
- +... 004/100..107: \_р\_ \_с\_ \_т\_ \_у\_ \_ф\_ \_х\_ \_ц\_ \_ч\_
- +... 004/110..117: \_ш\_ \_щ\_ \_ъ\_ \_ы\_ \_ь\_ \_э\_ \_ю\_ \_я\_
- +... 004/120..127: \_.\_ \_ё\_ \_ђ\_ \_ѓ\_ \_є\_ \_s\_ \_i\_ \_i\_
- +... 004/130..137: \_j\_ \_л\_ \_ь\_ \_ҥ\_ \_к\_ \_.\_ \_ў\_ \_ц\_

Full-screen

Back

Next

# `\pdfstringdef` converts **T<sub>E</sub>X** to PDF strings

Syntax: `\pdfstringdef{\command}{TEX string}`

1. Selecting the encoding PD1 or PU.
2. Redefinitions of many commands, so that they produce correct results or do not harm.
3. Expansion of the string.
4. Token checks and removal of forbidden tokens with comprehensive warnings.

# Redefinitions by `\pdfstringdef`

- Redefinitions to get correct results, e. g.:
  - Glyph commands: `\{`, `\%`, `\space`, `\dots`, . . .
  - Logos: `\TeX`, `\LaTeX`, `\MF`, . . .
- Many tokens are removed silently, e. g.:
  - $\text{\LaTeX}$  commands: `\label`, `\index`, `\textbf`, . . .
  - Stomach tokens: curly group braces, `\relax` (`\protect`), . . .
- The behaviour of `\xspace` is simulated.

# Bookmarks with package hyperref

- Places for options.
- Options for bookmarks.
- Creating bookmarks.
- Replacement methods.

Full-screen

Back

Next

# Places for hyperref options

1. Global: `\documentclass [ . . . ]` (e. g. driver)
2. Package: `\usepackage [ . . . ]`
3. Configuration file: `hyperref.cfg` with `\hypersetup`
4. After package has been loaded: `\hypersetup { . . . }`  
(e. g. PDF information options)

## Bookmark options

**bookmarks**: Make bookmarks (default: `true`).

**bookmarksnumbered**: Put section numbers in bookmarks (`false`).

**bookmarksopen**: Open up bookmark tree (default: `false`).

**bookmarksopenlevel**: Level, to which bookmarks are open.

**pdfpagemode**: How document starts when opened (default: `None`):

`None`: Neither outlines nor thumbnails are visible.

`UseOutlines`, `UseThumbs`: Outlines, thumbnails are visible.

`FullScreen`: Full-screen mode.

**unicode**: Bookmarks in Unicode (`false`). After package has been loaded, it switches between Unicode and PDFDocEncoding.

Full-screen

Back

Next

# Bookmarks by section commands

- Automatically addition of bookmarks by:
  - `\part`, `\chapter`, `\section`, `\subsection`, ...
  - `\addcontentsline`
- Help file `\jobname.out`:
  - It is written in the **first** run.
  - The bookmarks are set in the **second** run.
  - There is **no** “rerun” warning.

# Tree structure of bookmarks

- Subentries must be added to **direct** ancestors, **not** grandparents.
- If an intermediate level is omitted, the leaf starts **leftmost**:

	table of contents	bookmarks
<code>\part{I}</code>	Part I	main entry
<code>\chapter{5}</code>	→ Chapter 1.5	→ subentry
<code>\subsection{1}</code>	→ → → Subsection 1.5.0.1	main entry

# Creating bookmarks with `\pdfbookmark`

- The bookmarks of section commands have an **level** number:

documentclass	<code>\part</code>	<code>\chapter</code>	<code>\section</code>	<code>\subsection</code>	...
book/report:	<b>-1</b>	<b>0</b>	<b>1</b>	<b>2</b>	...
article:	<b>0</b>		<b>1</b>	<b>2</b>	...

- Syntax of `\pdfbookmark` (default for level: **0**):

```
\pdfbookmark[level]{bookmark text}{anchor name}
```

- `\currentpdfbookmark{bookmark}{anchor}% current level`
- `\subpdfbookmark{bookmark}{anchor}% current level + 1`

## Using anchors with `\pdfbookmark`

- An **anchor name** consists of the argument and the level:

```
\pdfbookmark[0]{Titlepage}{tit}% anchor: tit.0
```

- The bookmark can point to another **target (anchor)**:

```
\hypertarget{place.1}{}
```

Redirecting the bookmark to the previous defined **target**:

```
\begingroup
  \makeatletter
  \def\hyper@anchorstart #1\hyper@anchorend{}%
  \pdfbookmark[1]{Go to the place}{place}%
\endgroup
```

# Replacement methods

- `\texorpdfstring` chooses one of its arguments: a T<sub>E</sub>X or a PDF string.
- `\pdfstringdefPreHook` is a place for redefining commands. Additions are made by `\pdfstringdefDisableCommands`.
- Package `hypbmsec` extends the `\section` commands.

# `\texorpdfstring` method

Syntax: `\texorpdfstring{ $\text{TeX string}$ }{PDF string}`

Example:

```
\section{\texorpdfstring{H$_2$O}{Water}}
```

# Hook for private macro redefinitions

- `\pdfstringdefPreHook` is called, before expanding the string.
- `\pdfstringdefDisableCommands` adds user redefinitions to `\pdfstringdefPreHook`

```
\pdfstringdefDisableCommands{%  
  \renewcommand{\textcolor}[1]{#1}%  
  \renewcommand{\url}{\pdfstringdefwarn{\url}}%  
  \let\textcolor\@gobble  
}
```

- `\pdfstringdefwarn` prints a warning message.

# Package hypbmsec

- The syntax of `\section` commands is extended:
  - Second optional bookmark argument.
  - Bookmark in parentheses.
- Syntax:
  - `\section { toc/head = bookmark = text }`
  - `\section [ toc/head = bookmark ] { text }`
  - `\section [ toc/head ] [ bookmark ] { text }`
  - `\section ( bookmark ) { toc/head = text }`
  - `\section [ toc/head ] ( bookmark ) { text }`
  - `\section ( bookmark ) [ toc/head ] { text }`

# Additional features of PDF format

Hyperref does not support all possibilities of PDF format, e. g.:

- `/CreationDate` and `/ModDate`.
- Bookmarks with other functions.

The following examples uses commands of pdf $\text{T}\text{E}\text{X}$ .

Full-screen

Back

Next

# Bookmarks with other functions

- Menu functions of AcrobatReader, sound, video, . . .
- pdfTEX low level command:

`\pdfoutline action count n {text}`

- The absolute value of *n* is the count of the direct subentries.
- If *n* is negative, the subentries are closed.

Full-screen

Bookmarks

Page Only

Back

Next

## Link to external file

The PDF specification contains the possible actions on page 96.  
At last a bookmark that points to that page:

```
\pdfoutline
```

```
user {<</S /GoToR /F (pdfspec.pdf) /D [95 /Fit]>>}  
count 0 {Description of actions}
```

Full-screen

Bookmarks

Page Only

Back

Next

# Thumbnails

- Another method for navigation:
  - Choosing pages.
  - Selecting page areas.
- Contents:
  - Thumbnail sketches of the pages.
  - Other pictograms, symbols, . . .
  - Empty.

Full-screen

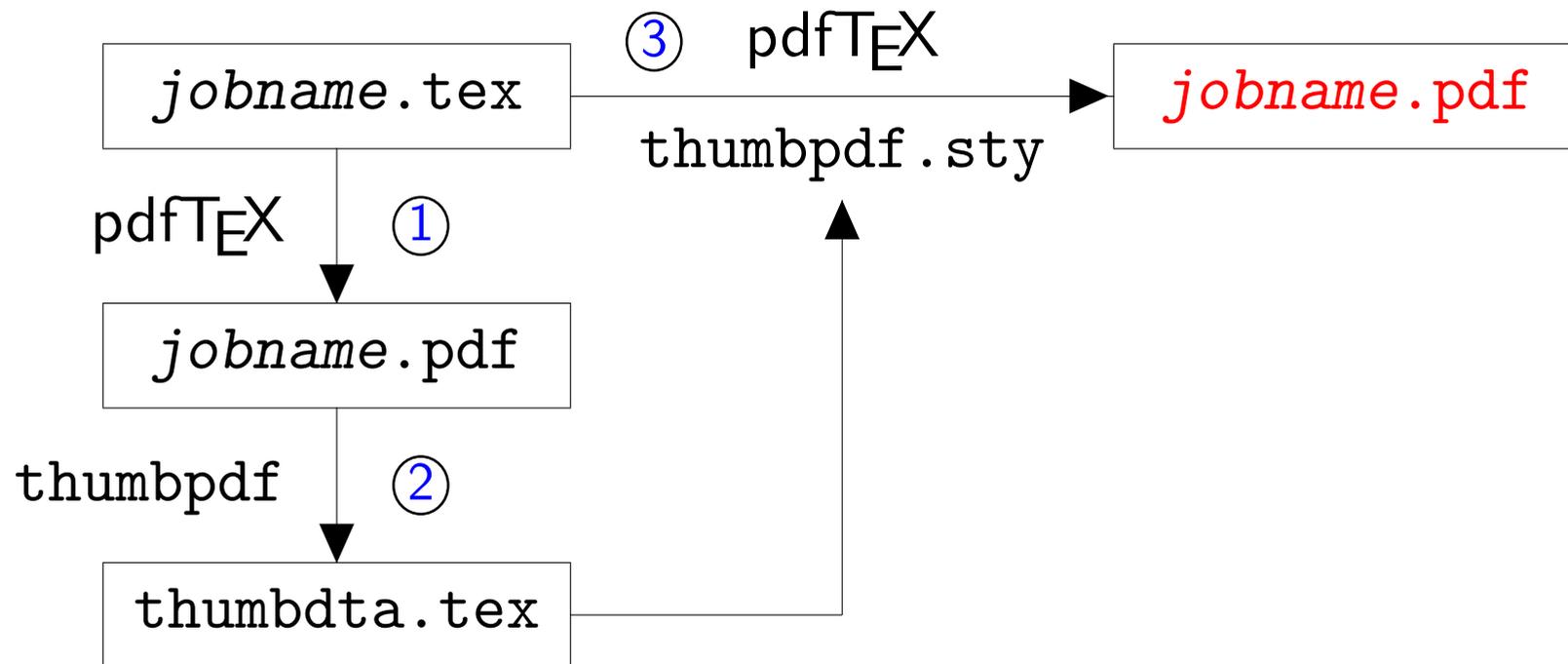
Thumbnails

Bookmarks

Back

Next

# Adding thumbnails with thumbpdf



Full-screen

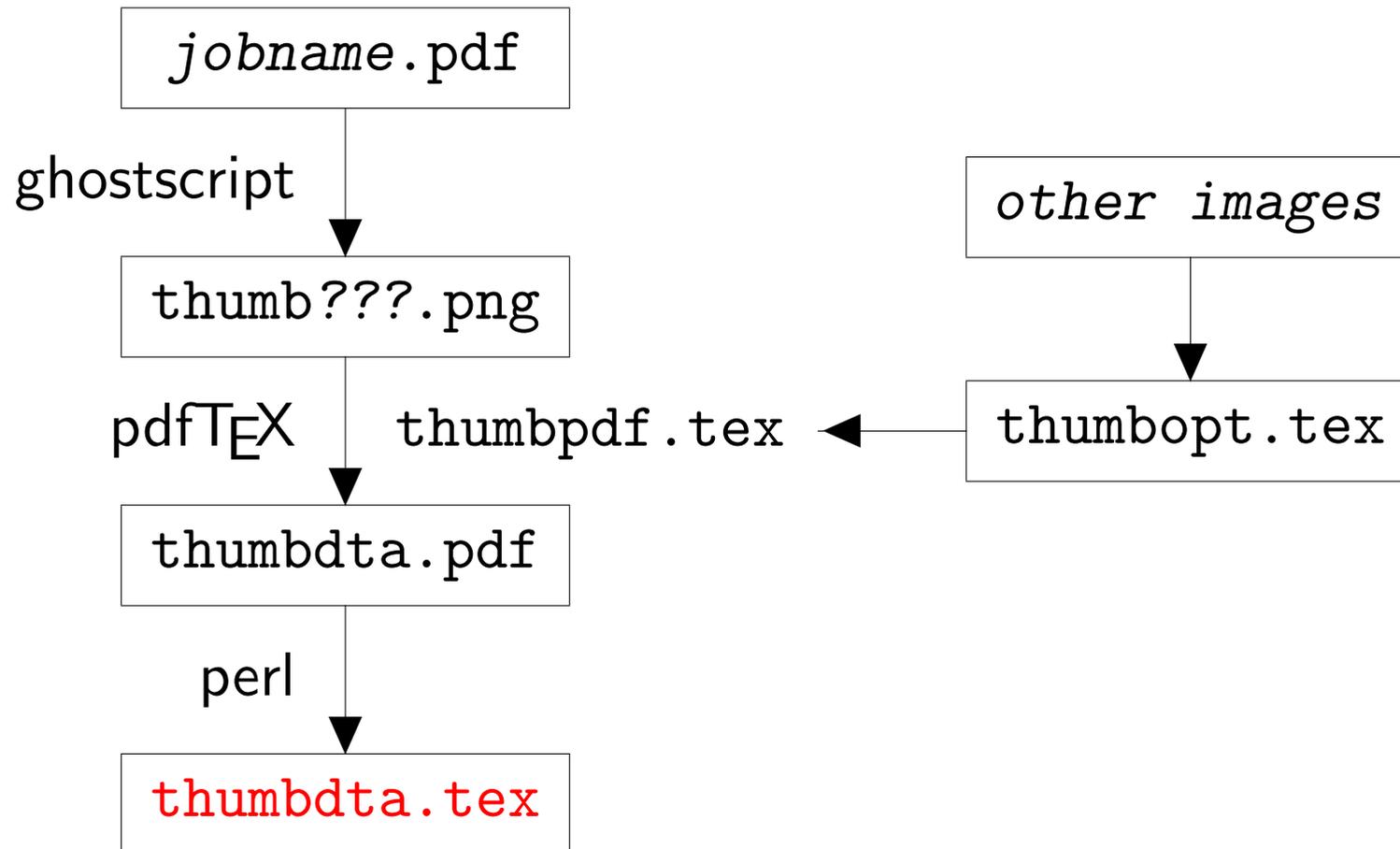
Thumbnails

Bookmarks

Back

Next

# How thumbpdf works



Full-screen

Thumbnails

Bookmarks

Back

Next

## Other images as thumbnails

- With `\thumb` other images are declared as thumbnails in the optional file `thumbopt.tex`.
- `\thisthumb` sets the thumbnail for the current page.

<code>thumbopt.tex</code>		<code>jobname.tex</code>
<code>\thumb{one}</code>	→	<code>\thisthumb{one}</code>
<code>\thumb{one.png}</code>	→	<code>\thisthumb{one.png}</code>
<code>\thumb[two]{one}</code>	→	<code>\thisthumb{two}</code>
<code>\thumb[one]{./extras/one.png}</code>	→	<code>\thisthumb{one}</code>

# Summary

- **General document information:**
  - Options in `\hypersetup` after `\usepackage{hyperref}`.
- **Bookmarks:**
  - Automatically added by `hyperref`.
  - **Oral expansion**, no stomach commands (exceptions).
  - Replacement methods: `\texorpdfstring`, `\pdfstringdefDisableCommands`, and `hypbmsec`.
  - Low level commands only for special effects.
- **Thumbnails:** Automatically added by `thumbpdf`.

# Links

- PDF specification (version 1.3):  
<http://partners.adobe.com/asn/developer/PDFS/TN/PDFSPEC.PDF>
- pdfTEX: (Hàn Thế Thành):  
<http://www.tug.org/applications/pdftex/>  
Win32: <ftp://ftp.esemetz.fr/pub/TeX/win32-beta/>
- hyperref (Sebastian Rahtz): ftp:  
<//ftp.ctan.org/tex-archive/macros/latex/contrib/hyperref/>  
Test versions:  
<http://www.tug.org/applications/hyperref/hyperref.zip>
- hypbmsec (Heiko Oberdiek): ftp:  
<//ftp.ctan.org/tex-archive/macros/latex/contrib/oberdiek/>
- thumbpdf (Heiko Oberdiek):  
<ftp://ftp.ctan.org/tex-archive/support/thumbpdf/>

Full-screen

Bookmarks

Quit

Back

Next

## Reusing of informations

Using of data that are available in `\maketitle`.

**Caution:** This only works before the first page is shipped out.

```
\newcommand{\org@maketitle}{}% LATEX-Check
\let\org@maketitle\maketitle
\def\maketitle{%
    \hypersetup{
        pdftitle={\@title},
        pdfauthor={\@author}
    }%
    \org@maketitle
}
```



## Tasks of encoding command `\OT1\ss`

- In a protected environment it expands to `\noexpand \ss`.
- Warning, if used in math mode.
- The character (`\OT1\ss = \char"19`) is set, if the currently active encoding matches.
- Else the value of the default encoding (`\?\ss`) is used (`\?\ss = \UseTextSymbol{OT1}\ss`).

## Bookmarks by `\addcontentsline`

- `\addcontentsline` also adds a bookmark entry:

```
\section*{Starred section}
\addcontentsline{toc}{section}{Starred section}
```

- For the link the last valid target (destination) is used.
- If there is a warning “`contentsline with no destination`”, a dummy target can be created:

```
\newcounter{dummy}
\begin{document}
  \refstepcounter{dummy}
  \addcontentsline{toc}{section}{Contents}
```

## Definition of `\texorpdfstring`

- `\pdfstringdef` sets a switch `\ifHy@pdfstring`.
- `\texorpdfstring` is full expandable:

```
\newcommand*{\texorpdfstring}{%  
  \ifHy@pdfstring  
    \expandafter\@secondoftwo  
  \else  
    \expandafter\@firstoftwo  
  \fi  
}
```



## Examples for `\texorpdfstring`

Syntax: `\texorpdfstring{TEX string}{PDF string}`

```
\section{Pythagoras:
```

```
  \texorpdfstring{$ a^2 + b^2 = c^2 $}{%  
    a\texttwosuperior\ + b\texttwosuperior\ =  
    c\texttwosuperior}%
```

```
}
```

```
\section{%
```

```
  \texorpdfstring{\textcolor{red}}{}{Red} Mars%
```

```
}
```

# Properties of hypbmsec

- Should work with packages that do not change the syntax of the `\section` commands. `hypbmsec` should be loaded **last**.
- Works **without** `hyperref` (bookmark argument is ignored).
- **Parameter delimiters** **inside** the optional parameter are protected by **curly braces**:

(... (... **{}**) ...) or [**{**... [**...**]...**}**]

# General document information

Example for setting `/CreationDate` and `/ModDate`:

```
\pdfinfo{/CreationDate (D:19990909000000-01'00')}
\begingroup
  \def\twodigits#1{\ifnum#1<10 0\fi\the#1}%
  \count0=\time \divide\count0 by 60
  \edef\x{\twodigits{\count0}}%
  \multiply\count0 by 60
  \count1=\time \advance\count1 by -\count0
  \edef\x{\x\twodigits{\count1}}%
  \edef\x{/ModDate (D:\the\year
    \twodigits\month \twodigits\day \x 00-01'00')}}%
\expandafter\endgroup
\expandafter\pdfinfo\expandafter{\x}%
```

## Example for “Named Actions”

```
\newcommand{\baction}[3][0]{%
  \begingroup
    \pdfstringdef\x{#3}%
    \pdfoutline
      user {<< /S /Named /N /#2 >>}
      count #1 {\x}%
  \endgroup
}
\baction[-3]{NOP}{Navigation}
\baction[2]{FullScreen}{Full-screen}
  \baction{PageOnly}{Page only}
  \baction{ShowThumbs}{Thumbnails}
\baction[6]{NOP}{Selecting pages}
  \baction{PrevPage}{Previous page}
...
```

Full-screen

Back

Quit