

PDF information and navigation elements with hyperref, pdf $\text{T}_{\text{E}}\text{X}$, and thumbpdf

Heiko Oberdiek

oberdiek@ruf.uni-freiburg.de

16th September 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 ISOLatin1
Unicode: 16bit, AcrobatReader version ≥ 4

Digestion of T_EX

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

mouth

“zero” : `\empty`
{...}: parameter
variable: read and check
fonts: –
Commands: `\string, \number`
`\if, \the`

stomach

`\relax`
group
assignments
different fonts and encodings
ligatures
`\def, \hbox, $`
`\begingroup, \special`

T_EX into PDF strings

- Reusing the argument of section commands (T_EX 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- T_EX : `\ss = \char"19` \Rightarrow β

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

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

`\ss = \OT1-cmd \ss \OT1\ss` \Rightarrow β

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`, \textasciitilde = `\textasciitilde`
- Traditional T_EX and L^AT_EX 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_ _ñ_ _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_ _İ_
- +... 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_EX** 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.:
 - \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:	-1	0	1	2	...
article:	0		1	2	...

- 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_EX 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{ TeX string }{PDF string}`

Example:

```
\section{\texorpdfstring{H$_2$O$_2$}{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 TEX .

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

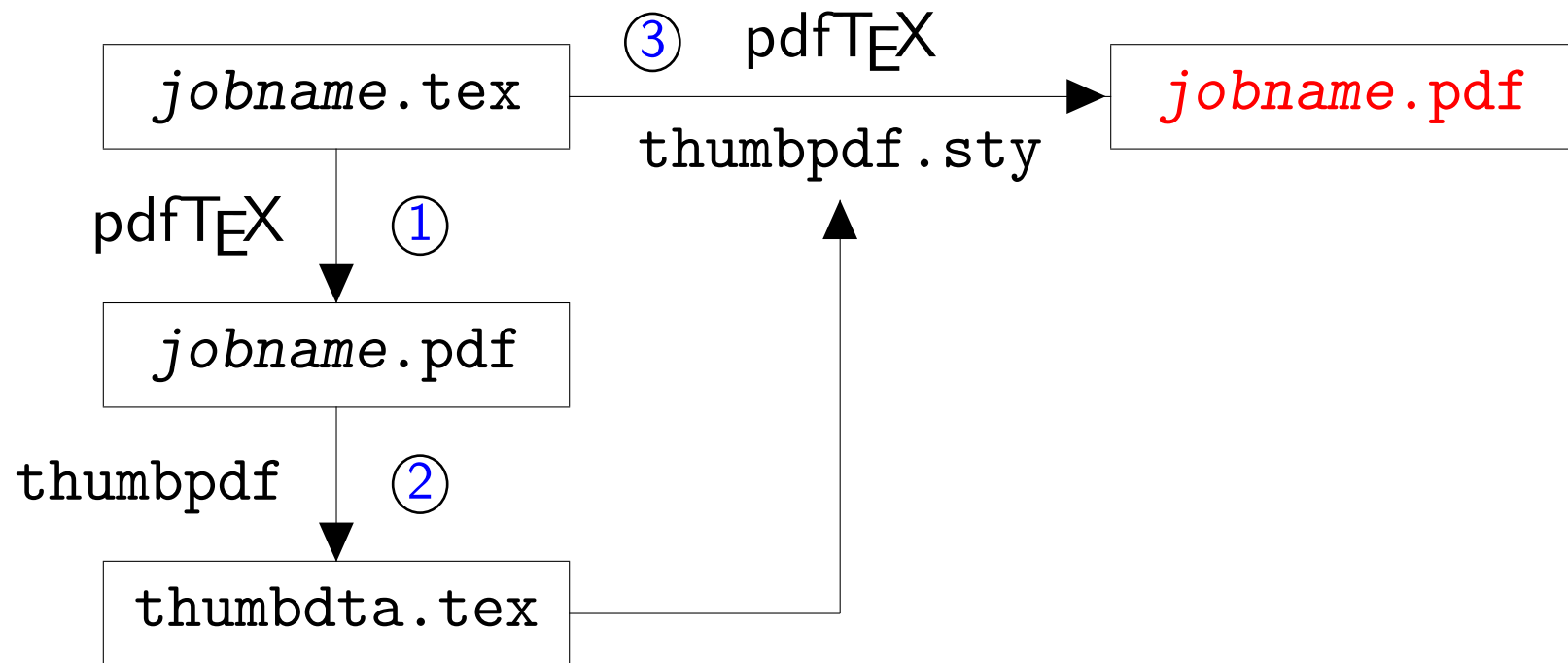
Thumbnails

Bookmarks

Back

Next

Adding thumbnails with thumbpdf



Full-screen

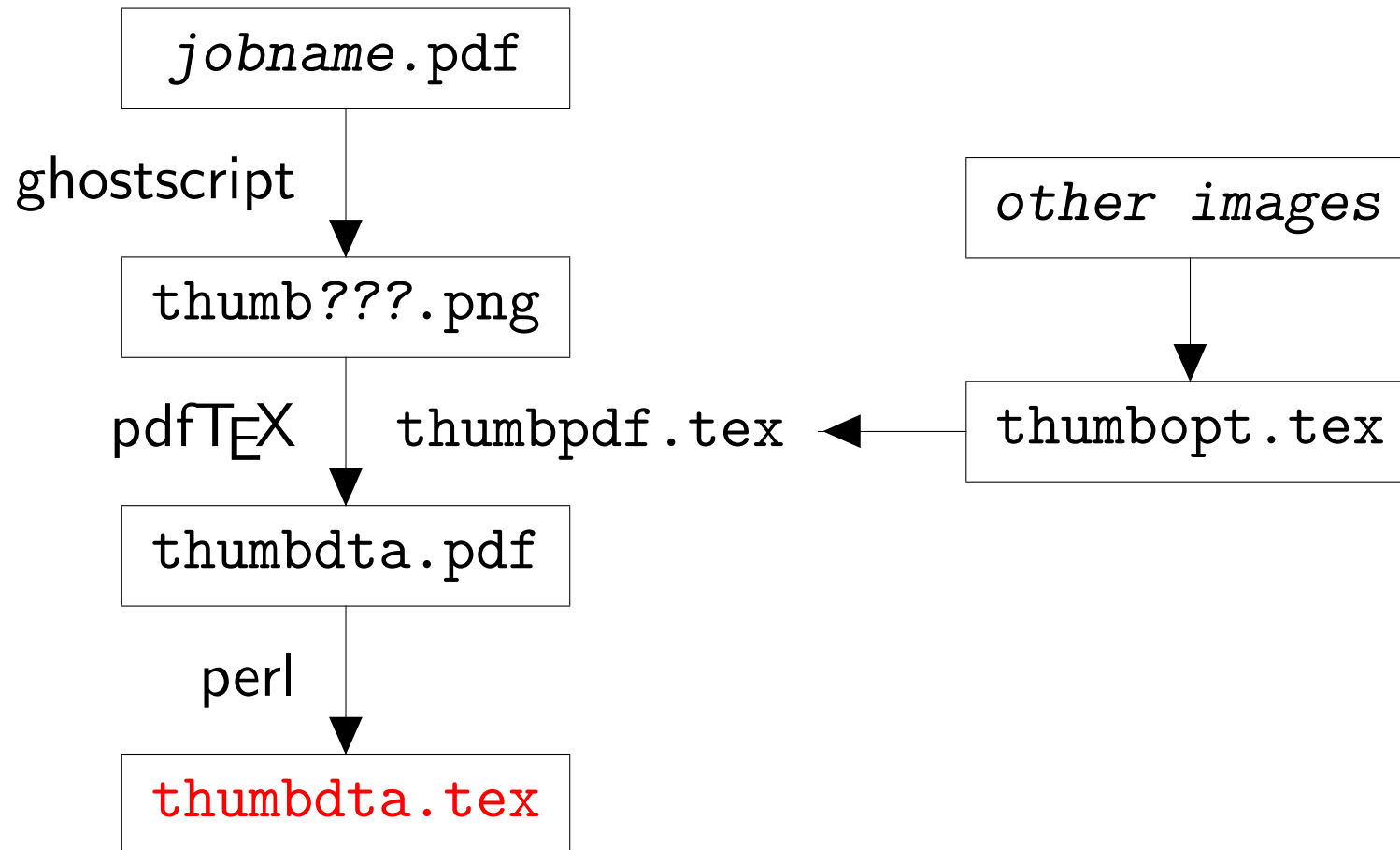
Thumbnails

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 `thumbdta.tex`.
- `\thisthumb` sets the thumbnail for the current page.

`thumbopt.tex`

`jobname.tex`

`\thumb{one}`

→

`\thisthumb{one}`

`\thumb{one.png}`

→

`\thisthumb{one.png}`

`\thumb[two]{one}`

→

`\thisthumb{two}`

`\thumb[one]{./extras/one.png}`

→

`\thisthumb{one}`

Full-screen

Thumbnails

Bookmarks

Back

Next

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://ctan.tug.org/tex-archive/macros/latex/contrib/supported/hyperref/>
Test versions:
<http://www.tug.org/applications/hyperref/hyperref.zip>
- hypbmsec (Heiko Oberdiek): <ftp://ctan.tug.org/tex-archive/macros/latex/contrib/supported/oberdiek/>
- thumbpdf (Heiko Oberdiek):
<ftp://ctan.tug.org/tex-archive/macros/pdftex/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\@firstoftwo  
  \else  
    \expandafter\@secondoftwo  
  \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