PDF information and navigation elements with hyperref, pdfTeX, and thumbpdf

Heiko Oberdiek
oberdiek@ruf.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`

General Document Information

Application: search engines.

`hyperref` **options:**

Title:	`pdftitle`
Subject:	`pdfsubject`
Author:	`pdfauthor`
Keywords:	`pdfkeywords`
Creator:	`pdfcreator`
Producer:	`pdfproducer`

Times:

Created:	/CreationDate
Modificated:	/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 = {pdfeTeX-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 TEX

Eyes

- Reading input lines
- Catcodes are set

Mouth

- Forming tokens
- Expansion of macros

Stomach

- Assignments
- Typesetting

Differences between mouth and stomach

	mouth	**stomach**
"zero":	`\empty`	`\relax`
`{...}`:	parameter	group
variable:	read and check	assignments
fonts:	–	different fonts and encodings ligatures
Commands:	`\string`, `\number` `\if`, `\the`	`\def`, `\hbox`, `$` `\begingroup`, `\special`

TEX into PDF strings

- Reusing the argument of section commands (TEX 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-TEX: `\ss = \char"19` ⇒ ß

LATEX 2ε: `\fontencoding{OT1}\selectfont\char"19` ⇒ ß

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

`\ss = \OT1-cmd \ss \OT1\ss` ⇒ ß

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: ® = \textregistered, ˜ = \textasciitilde

- Traditional TEX and LATEX names.
  Examples: Æ = \AE, ß = \ss

- Accented characters.
  Examples: Ü = \"U, ç = \c c, Å = \r A (⇐ \AA)

- Escape octal sequences are the result: \textmu ⇒ \265

PU encoding (Unicode)

☹ Characters having catcodes 11 (letter) and 12 (other):
Checking each token, high memory and time consumption.

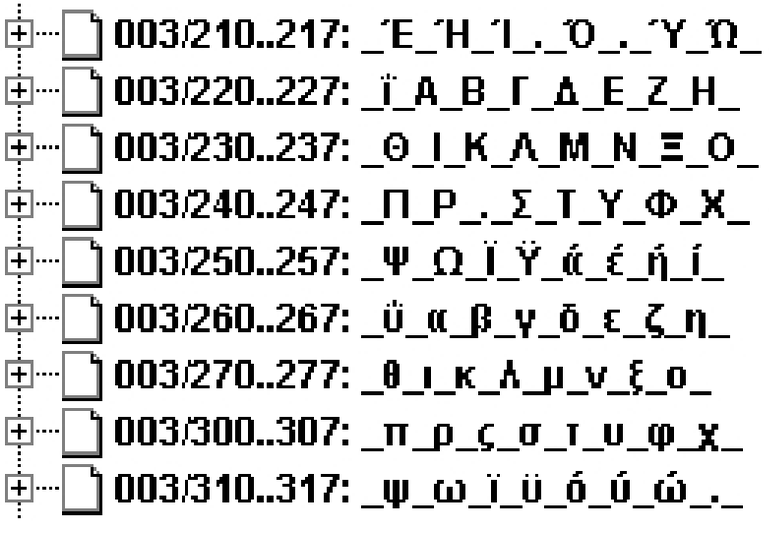
☹ Ligatures: same as above.

☹ Direct commands:
Easier to implement, but each command must be redefined each time.

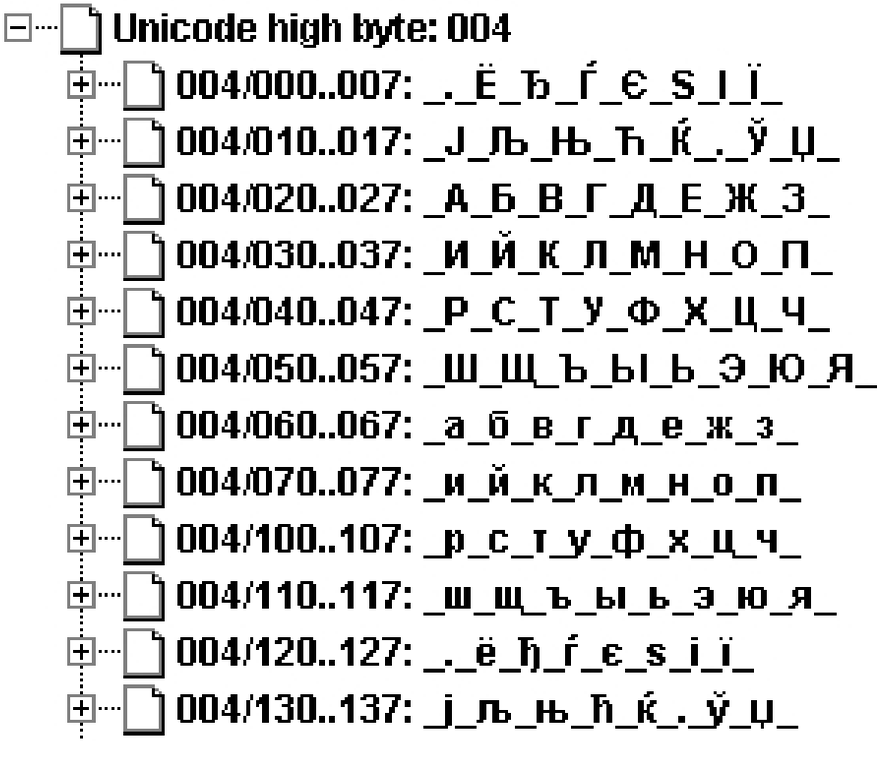
☺ Commands of LaTeX 2_ε's fontencoding mechanism:
Smallest memory consumption and fastest method.
This method is used by package `hyperref`.

Examples of Unicode bookmarks

Greek:



Cyrillic:



Full-screen Back Next

\pdfstringdef converts TeX 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.

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.

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 ancestores, not grandparents.
- If an intermediate level is omitted, the leaf starts leftmost:

	table of contents	bookmarks
`\part{I}`	Part I	main entry
`\chapter{5}`	→Chapter I.5	→subentry
`\subsection{1}`	→→→Subsection I.5.0.1	main entry

Creating bookmarks with `\pdfbookmark`

- The bookmarks of section commands have an level number:

documentclass	`\part`	`\chapter`	`\section`	`\subsection`	...
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` choices one of its arguments:
  a TEX 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}{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 pdfTEX.

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.

Link to external file

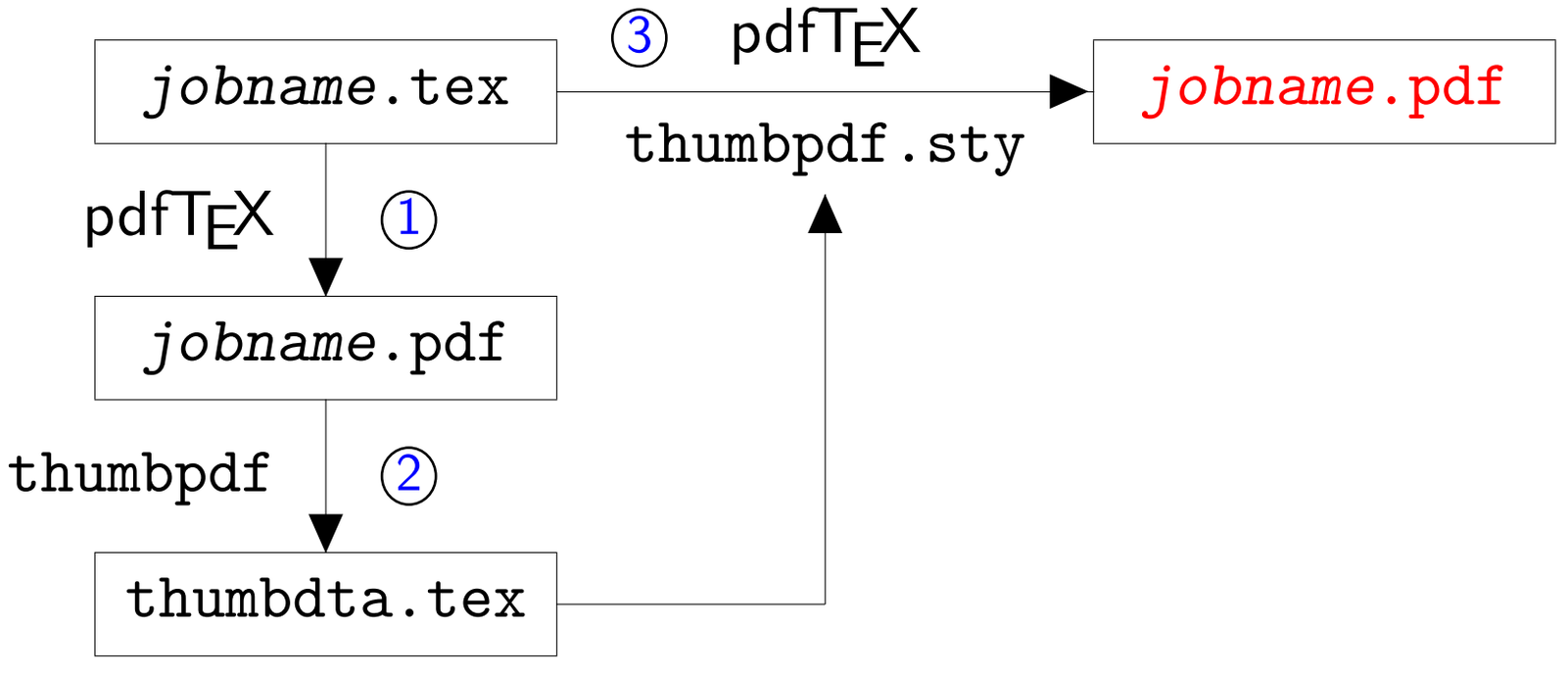
The PDF specification contains the possible actions on page 96.
At last a bookmark that points to that page:

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

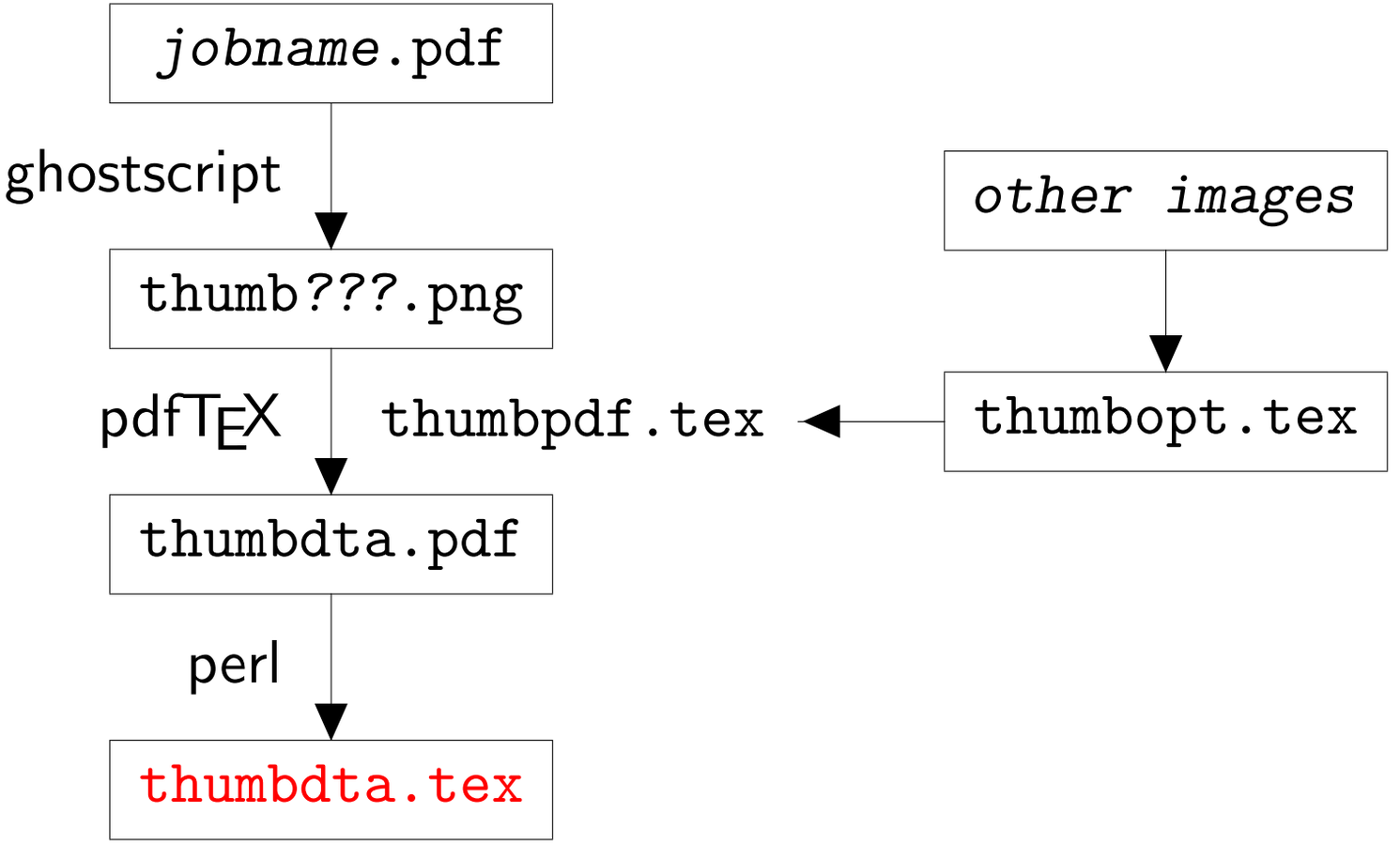
Thumbnails

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

Adding thumbnails with `thumbpdf`



How thumbpdf works



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.

`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}`

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.ese-metz.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/

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