Making better PDF

Hartmut Henkel

EuroTEX 2006, Debrecen, Friday 7 July 2006

Overview

Introduction

Why look into a PDF file?

- Get some background: Learn what PDF is.
- Get a rough idea how it's created by pdfTeX.
- This might help to solve problems around PDF.
- Maybe you get interest in writing own PDF code.
- Maybe you might even want to tweak pdfTeX.

This might ultimately lead to "better PDF".

PDF tour

The LaTeX example input file:

foo.tex

```
\documentclass[12pt,a4paper]{article}
\usepackage[margin=1in]{geometry}
\usepackage{graphicx}
\pdfcompresslevel=0
\pagestyle{empty}
\begin{document}
\noindent
Euro\TeX\ 2006 \bfseries Ahoi!
\bigskip\par\noindent
\includegraphics[width=\textwidth]{campushotel1.jpg}
\end{document}
```

PDF tour

The PDF output looks like this A4 sheet:

EuroTEX 2006 **Ahoi!**

PDF tour — global structure of the PDF file

Header

```
%PDF-1.4
```

Body

```
... many objects: page descriptions, fonts, images, etc. ...
11 0 obj << /Type /Pages /Count 1 /Kids [3 0 R] >> endobj
14 0 obj << /Type /Catalog /Pages 11 0 R >> endobj
15 0 obj << /Producer (pdfeTeX-1.40.0)
/CreationDate (D:20060701211003+02'00') ... >> endobj
```

Cross reference table

(object numbers → byte offsets)

```
xref
0 16
0000000000 65535 f
0000000411 00000 n
0000073302 00000 n
0000000299 00000 n
0000000015 00000 n
0000079926 00000 n
0000076450 00000 n
0000079788 00000 n
0000076047 00000 n
0000073416 00000 n
0000075907 00000 n
0000080332 00000 n
0000076255 00000 n
0000080141 00000 n
0000080390 00000 n
0000080441 00000 n
```

Trailer

```
trailer
<< /Size 16
/Root 14 0 R
/Info 15 0 R ... >>
startxref
80710
%%EOF
```

PDF tour — how a PDF reader scans the PDF file

```
%PDF-1.4
```

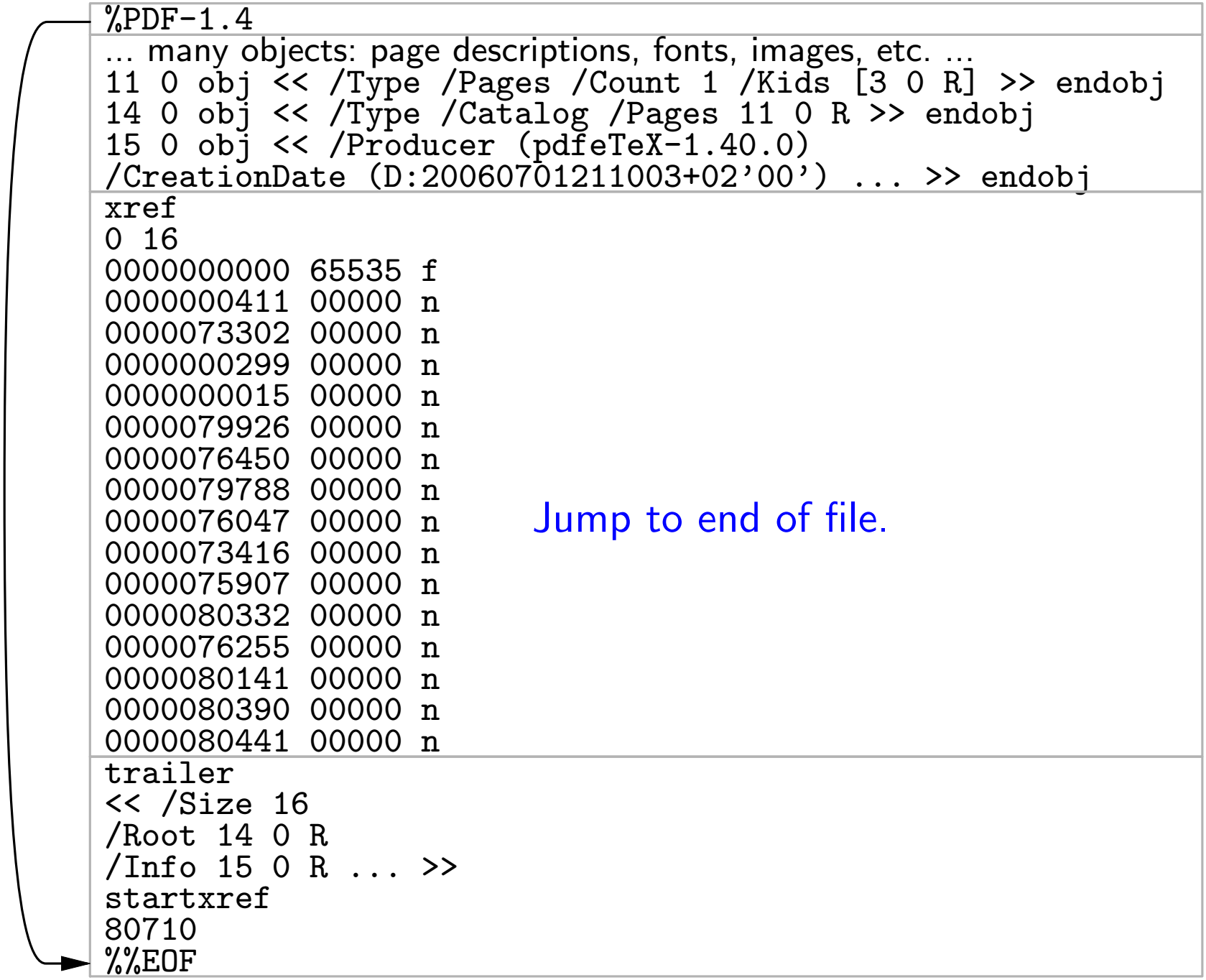
```
... many objects: page descriptions, fonts, images, etc. ...
11 0 obj << /Type /Pages /Count 1 /Kids [3 0 R] >> endobj
14 0 obj << /Type /Catalog /Pages 11 0 R >> endobj
15 0 obj << /Producer (pdfeTeX-1.40.0)
/CreationDate (D:20060701211003+02'00') ... >> endobj
```

```
xref
0 16
0000000000 65535 f
0000000411 00000 n
0000073302 00000 n
0000000299 00000 n
0000000015 00000 n
0000079926 00000 n
0000076450 00000 n
0000079788 00000 n
0000076047 00000 n
0000073416 00000 n
0000075907 00000 n
0000080332 00000 n
0000076255 00000 n
0000080141 00000 n
0000080390 00000 n
0000080441 00000 n
```

Start: Is it PDF? Which version?

```
trailer
<< /Size 16
/Root 14 0 R
/Info 15 0 R ... >>
startxref
80710
%%EOF
```

PDF tour — how a PDF reader scans the PDF file



PDF tour — how a PDF reader scans the PDF file

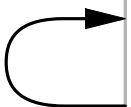
```
%PDF-1.4
```

```
... many objects: page descriptions, fonts, images, etc. ...
11 0 obj << /Type /Pages /Count 1 /Kids [3 0 R] >> endobj
14 0 obj << /Type /Catalog /Pages 11 0 R >> endobj
15 0 obj << /Producer (pdfeTeX-1.40.0)
/CreationDate (D:20060701211003+02'00') ... >> endobj
```

```
xref
0 16
0000000000 65535 f
0000000411 00000 n
0000073302 00000 n
0000000299 00000 n
0000000015 00000 n
0000079926 00000 n
0000076450 00000 n
0000079788 00000 n
0000076047 00000 n
0000073416 00000 n
0000075907 00000 n
0000080332 00000 n
0000076255 00000 n
0000080141 00000 n
0000080390 00000 n
0000080441 00000 n
```

Search backward for `startxref`.

```
trailer
<< /Size 16
/Root 14 0 R
/Info 15 0 R ... >>
startxref
80710
%%EOF
```



PDF tour — how a PDF reader scans the PDF file

```
%PDF-1.4
```

```
... many objects: page descriptions, fonts, images, etc. ...
11 0 obj << /Type /Pages /Count 1 /Kids [3 0 R] >> endobj
14 0 obj << /Type /Catalog /Pages 11 0 R >> endobj
15 0 obj << /Producer (pdfeTeX-1.40.0)
/CreationDate (D:20060701211003+02'00') ... >> endobj
```

```
xref
0 16
0000000000 65535 f
0000000411 00000 n
0000073302 00000 n
0000000299 00000 n
0000000015 00000 n
0000079926 00000 n
0000076450 00000 n
0000079788 00000 n
0000076047 00000 n
0000073416 00000 n
0000075907 00000 n
0000080332 00000 n
0000076255 00000 n
0000080141 00000 n
0000080390 00000 n
0000080441 00000 n
```

Get byte offset for `xref`.

(all byte offsets from begin of file)

```
trailer
<< /Size 16
/Root 14 0 R
/Info 15 0 R ... >>
startxref
80710
%%EOF
```

PDF tour — how a PDF reader scans the PDF file

```
%PDF-1.4
... many objects: page descriptions, fonts, images, etc. ...
11 0 obj << /Type /Pages /Count 1 /Kids [3 0 R] >> endobj
14 0 obj << /Type /Catalog /Pages 11 0 R >> endobj
15 0 obj << /Producer (pdfeTeX-1.40.0)
/CreationDate (D:20060701211003+02'00') ... >> endobj
xref
0 16
0000000000 65535 f
 1 0000000411 00000 n
 2 0000073302 00000 n
 3 0000000299 00000 n
 4 0000000015 00000 n
 5 0000079926 00000 n
 6 0000076450 00000 n
 7 0000079788 00000 n
 8 0000076047 00000 n
 9 0000073416 00000 n
10 0000075907 00000 n
11 0000080332 00000 n
12 0000076255 00000 n
13 0000080141 00000 n
14 0000080390 00000 n
15 0000080441 00000 n
trailer
<< /Size 16
/Root 14 0 R
/Info 15 0 R ... >>
startxref
80710
%%EOF
```

Get cross reference table by offset.

(left: object numbers)

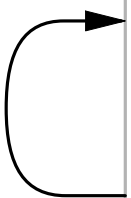
PDF tour — how a PDF reader scans the PDF file

```
%PDF-1.4
```

```
... many objects: page descriptions, fonts, images, etc. ...
11 0 obj << /Type /Pages /Count 1 /Kids [3 0 R] >> endobj
14 0 obj << /Type /Catalog /Pages 11 0 R >> endobj
15 0 obj << /Producer (pdfeTeX-1.40.0)
/CreationDate (D:20060701211003+02'00') ... >> endobj
```

```
xref
0 16
0000000000 65535 f
0000000411 00000 n
0000073302 00000 n
0000000299 00000 n
0000000015 00000 n
0000079926 00000 n
0000076450 00000 n
0000079788 00000 n
0000076047 00000 n
0000073416 00000 n
0000075907 00000 n
0000080332 00000 n
0000076255 00000 n
0000080141 00000 n
0000080390 00000 n
0000080441 00000 n
```

Search backward for trailer.



```
trailer
<< /Size 16
/Root 14 0 R
/Info 15 0 R ... >>
startxref
80710
%%EOF
```

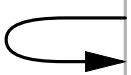
PDF tour — how a PDF reader scans the PDF file

```
%PDF-1.4
```

```
... many objects: page descriptions, fonts, images, etc. ...
11 0 obj << /Type /Pages /Count 1 /Kids [3 0 R] >> endobj
14 0 obj << /Type /Catalog /Pages 11 0 R >> endobj
15 0 obj << /Producer (pdfeTeX-1.40.0)
/CreationDate (D:20060701211003+02'00') ... >> endobj
```

```
xref
0 16
0000000000 65535 f
0000000411 00000 n
0000073302 00000 n
0000000299 00000 n
0000000015 00000 n
0000079926 00000 n
0000076450 00000 n
0000079788 00000 n
0000076047 00000 n
0000073416 00000 n
0000075907 00000 n
0000080332 00000 n
0000076255 00000 n
0000080141 00000 n
0000080390 00000 n
0000080441 00000 n
```

Read trailer dictionary.



```
trailer
<< /Size 16
/Root 14 0 R
/Info 15 0 R ... >>
startxref
80710
%%EOF
```

PDF tour — how a PDF reader scans the PDF file

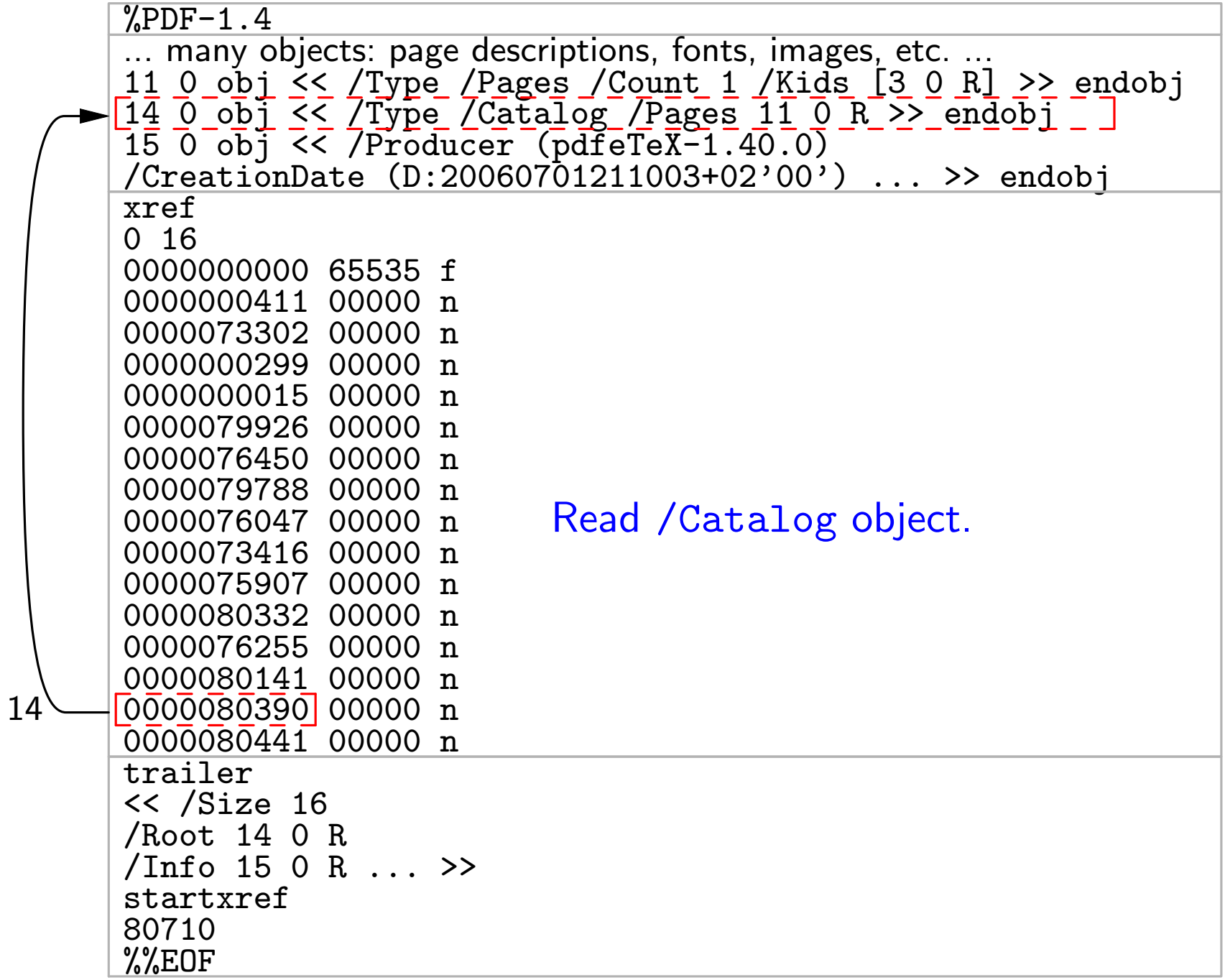
```
%PDF-1.4
... many objects: page descriptions, fonts, images, etc. ...
11 0 obj << /Type /Pages /Count 1 /Kids [3 0 R] >> endobj
14 0 obj << /Type /Catalog /Pages 11 0 R >> endobj
15 0 obj << /Producer (pdfeTeX-1.40.0)
/CreationDate (D:20060701211003+02'00') ... >> endobj
xref
0 16
0000000000 65535 f
0000000411 00000 n
0000073302 00000 n
0000000299 00000 n
0000000015 00000 n
0000079926 00000 n
0000076450 00000 n
0000079788 00000 n
0000076047 00000 n
0000073416 00000 n
0000075907 00000 n
0000080332 00000 n
0000076255 00000 n
0000080141 00000 n
0000080390 00000 n
0000080441 00000 n
trailer
<< /Size 16
/Root 14 0 R
/Info 15 0 R ... >>
startxref
80710
%%EOF
```

14

Lookup offset of /Catalog object.

(/Root ≡ /Catalog)

PDF tour — how a PDF reader scans the PDF file



PDF tour — how a PDF reader scans the PDF file

```
%PDF-1.4
... many objects: page descriptions, fonts, images, etc. ...
11 0 obj << /Type /Pages /Count 1 /Kids [3 0 R] >> endobj
14 0 obj << /Type /Catalog /Pages 11 0 R >> endobj
15 0 obj << /Producer (pdfeTeX-1.40.0)
/CreationDate (D:20060701211003+02'00') ... >> endobj
xref
0 16
0000000000 65535 f
0000000411 00000 n
0000073302 00000 n
0000000299 00000 n
0000000015 00000 n
0000079926 00000 n
0000076450 00000 n
0000079788 00000 n
0000076047 00000 n
0000073416 00000 n
0000075907 00000 n
0000080332 00000 n
0000076255 00000 n
0000080141 00000 n
0000080390 00000 n
0000080441 00000 n
trailer
<< /Size 16
/Root 14 0 R
/Info 15 0 R ... >>
startxref
80710
%%EOF
```

11

Lookup offset of /Pages object.

PDF tour — how a PDF reader scans the PDF file

```
%PDF-1.4
... many objects: page descriptions, fonts, images, etc. ...
11 0 obj << /Type /Pages /Count 1 /Kids [3 0 R] >> endobj
14 0 obj << /Type /Catalog /Pages 11 0 R >> endobj
15 0 obj << /Producer (pdfeTeX-1.40.0)
/CreationDate (D:20060701211003+02'00') ... >> endobj
xref
0 16
0000000000 65535 f
0000000411 00000 n
0000073302 00000 n
0000000299 00000 n
0000000015 00000 n
0000079926 00000 n
0000076450 00000 n
0000079788 00000 n
0000076047 00000 n
0000073416 00000 n
0000075907 00000 n
0000080332 00000 n
0000076255 00000 n
0000080141 00000 n
0000080390 00000 n
0000080441 00000 n
trailer
<< /Size 16
/Root 14 0 R
/Info 15 0 R ... >>
startxref
80710
%%EOF
```

11

Read /Pages object.

(gives access to all /Page objects)

PDF tour — gathering stuff for building a page

```
4 0 obj <<
/Length 227
>>
stream
BT
/F15 11.9552 Tf 72 757.935 Td[(EuroT)]TJ
31.376 -2.574 Td[(E)]TJ 6.501 2.574 Td[(X)-326(2006)]TJ
/F16 11.9552 Tf 39.992 0 Td[(Ahoi!)]TJ
ET
1 0 0 1 72 507.702 cm
 q 0.54967 0 0 0.54967 0 0 cm
q
821 0 0 427 0 0 cm
/Im1 Do
Q
 Q
endstream
endobj
```

Jump to first /Page object.
(there is only one in our file)

```
3 0 obj <<
/Type /Page /Contents 4 0 R /Resources 2 0 R
/MediaBox [0 0 595.276 841.89]
/Parent 11 0 R
>> endobj
2 0 obj <<
/Font << /F15 7 0 R /F16 10 0 R >>
/XObject << /Im1 1 0 R >> /ProcSet [ /PDF /Text /ImageC ]
>> endobj
11 0 obj <<
/Type /Pages /Count 1 /Kids [3 0 R]
>> endobj
```

PDF tour — gathering stuff for building a page

```
4 0 obj <<
/Length 227
>>
stream
BT
/F15 11.9552 Tf 72 757.935 Td[(EuroT)]TJ
31.376 -2.574 Td[(E)]TJ 6.501 2.574 Td[(X)-326(2006)]TJ
/F16 11.9552 Tf 39.992 0 Td[(Ahoi!)]TJ
ET
1 0 0 1 72 507.702 cm
 q 0.54967 0 0 0.54967 0 0 cm
q
821 0 0 427 0 0 cm
/Im1 Do
Q
 Q
endstream
endobj
```

Get all /Resources for this page.

```
3 0 obj <<
/Type /Page /Contents 4 0 R /Resources 2 0 R
/MediaBox [0 0 595.276 841.89]
/Parent 11 0 R
>> endobj
2 0 obj <<
/Font << /F15 7 0 R /F16 10 0 R >>
/XObject << /Im1 1 0 R >> /ProcSet [ /PDF /Text /ImageC ]
>> endobj
11 0 obj <<
/Type /Pages /Count 1 /Kids [3 0 R]
>> endobj
```

PDF tour — gathering stuff for building a page

```
4 0 obj <<
/Length 227
>>
stream
BT
/F15 11.9552 Tf 72 757.935 Td[(EuroT)]TJ
31.376 -2.574 Td[(E)]TJ 6.501 2.574 Td[(X)-326(2006)]TJ
/F16 11.9552 Tf 39.992 0 Td[(Ahoi!)]TJ
ET
1 0 0 1 72 507.702 cm
 q 0.54967 0 0 0.54967 0 0 cm
q
821 0 0 427 0 0 cm
/Im1 Do
Q
 Q
endstream
endobj
```

There are two fonts needed. . .

```
3 0 obj <<
/Type /Page /Contents 4 0 R /Resources 2 0 R
/MediaBox [0 0 595.276 841.89]
/Parent 11 0 R
>> endobj
2 0 obj <<
/Font << /F15 7 0 R /F16 10 0 R >>
/XObject << /Im1 1 0 R >> /ProcSet [ /PDF /Text /ImageC ]
>> endobj
11 0 obj <<
/Type /Pages /Count 1 /Kids [3 0 R]
>> endobj
```

PDF tour — gathering stuff for building a page

```
4 0 obj <<
/Length 227
>>
stream
BT
/F15 11.9552 Tf 72 757.935 Td[(EuroT)]TJ
31.376 -2.574 Td[(E)]TJ 6.501 2.574 Td[(X)-326(2006)]TJ
/F16 11.9552 Tf 39.992 0 Td[(Ahoi!)]TJ
ET
1 0 0 1 72 507.702 cm
 q 0.54967 0 0 0.54967 0 0 cm
q
821 0 0 427 0 0 cm
/Im1 Do
Q
 Q
endstream
endobj
```

. . . and there is one image needed.

(which is formally an /XObject)

```
3 0 obj <<
/Type /Page /Contents 4 0 R /Resources 2 0 R
/MediaBox [0 0 595.276 841.89]
/Parent 11 0 R
>> endobj
```

```
2 0 obj <<
/Font << /F15 7 0 R /F16 10 0 R >>
/XObject << /Im1 1 0 R >> /ProcSet [ /PDF /Text /ImageC ]
>> endobj
```

```
11 0 obj <<
/Type /Pages /Count 1 /Kids [3 0 R]
>> endobj
```

PDF tour — gathering stuff for building a page

```
4 0 obj <<
/Length 227
>>
stream
BT
/F15 11.9552 Tf 72 757.935 Td[(EuroT)]TJ
31.376 -2.574 Td[(E)]TJ 6.501 2.574 Td[(X)-326(2006)]TJ
/F16 11.9552 Tf 39.992 0 Td[(Ahoi!)]TJ
ET
1 0 0 1 72 507.702 cm
 q 0.54967 0 0 0.54967 0 0 cm
q
821 0 0 427 0 0 cm
/Im1 Do
Q
 Q
endstream
endobj
```

```
3 0 obj <<
/Type /Page /Contents 4 0 R /Resources 2 0 R
/MediaBox [0 0 595.276 841.89]
/Parent 11 0 R
>> endobj
```

```
2 0 obj <<
/Font << /F15 7 0 R /F16 10 0 R >>
/XObject << /Im1 1 0 R >> /ProcSet [ /PDF /Text /ImageC ]
>> endobj
```

```
11 0 obj <<
/Type /Pages /Count 1 /Kids [3 0 R]
>> endobj
```

Fonts and images are referenced by (arbitrary) internal names.

PDF tour — gathering stuff for building a page

```
4 0 obj <<
/Length 227
>>
stream
BT
/F15 11.9552 Tf 72 757.935 Td[(EuroT)]TJ
31.376 -2.574 Td[(E)]TJ 6.501 2.574 Td[(X)-326(2006)]TJ
/F16 11.9552 Tf 39.992 0 Td[(Ahoi!)]TJ
ET
1 0 0 1 72 507.702 cm
 q 0.54967 0 0 0.54967 0 0 cm
q
821 0 0 427 0 0 cm
/Im1 Do
Q
 Q
endstream
endobj
```

The actual page contents is defined by a stream object.

```
3 0 obj <<
/Type /Page /Contents 4 0 R /Resources 2 0 R
/MediaBox [0 0 595.276 841.89]
/Parent 11 0 R
>> endobj
```

```
2 0 obj <<
/Font << /F15 7 0 R /F16 10 0 R >>
/XObject << /Im1 1 0 R >> /ProcSet [ /PDF /Text /ImageC ]
>> endobj
```

```
11 0 obj <<
/Type /Pages /Count 1 /Kids [3 0 R]
>> endobj
```

PDF tour — reading the page stream

```
4 0 obj <<
/Length 227
>>
stream
BT
/F15 11.9552 Tf 72 757.935 Td[(EuroT)]TJ
31.376 -2.574 Td[(E)]TJ 6.501 2.574 Td[(X)-326(2006)]TJ
/F16 11.9552 Tf 39.992 0 Td[(Ahoi!)]TJ
ET
1 0 0 1 72 507.702 cm
 q 0.54967 0 0 0.54967 0 0 cm
q
821 0 0 427 0 0 cm
/Im1 Do
Q
 Q
endstream
endobj
```

A stream object has a dictionary.

(/Length; often filter for decompression)

```
3 0 obj <<
/Type /Page /Contents 4 0 R /Resources 2 0 R
/MediaBox [0 0 595.276 841.89]
/Parent 11 0 R
>> endobj
```

```
2 0 obj <<
/Font << /F15 7 0 R /F16 10 0 R >>
/XObject << /Im1 1 0 R >> /ProcSet [ /PDF /Text /ImageC ]
>> endobj
```

```
11 0 obj <<
/Type /Pages /Count 1 /Kids [3 0 R]
>> endobj
```

PDF tour — reading the page stream

```
4 0 obj <<
/Length 227
>>
stream
BT
/F15 11.9552 Tf 72 757.935 Td[(EuroT)]TJ
31.376 -2.574 Td[(E)]TJ 6.501 2.574 Td[(X)-326(2006)]TJ
/F16 11.9552 Tf 39.992 0 Td[(Ahoi!)]TJ
ET
1 0 0 1 72 507.702 cm
 q 0.54967 0 0 0.54967 0 0 cm
q
821 0 0 427 0 0 cm
/Im1 Do
Q
 Q
endstream
endobj
```

The actual stream data are bracketed by keywords.

```
3 0 obj <<
/Type /Page /Contents 4 0 R /Resources 2 0 R
/MediaBox [0 0 595.276 841.89]
/Parent 11 0 R
>> endobj
```

```
2 0 obj <<
/Font << /F15 7 0 R /F16 10 0 R >>
/XObject << /Im1 1 0 R >> /ProcSet [ /PDF /Text /ImageC ]
>> endobj
```

```
11 0 obj <<
/Type /Pages /Count 1 /Kids [3 0 R]
>> endobj
```

PDF tour — reading the page stream

```
4 0 obj <<
/Length 227
>>
stream
BT
/F15 11.9552 Tf 72 757.935 Td[(EuroT)]TJ
31.376 -2.574 Td[(E)]TJ 6.501 2.574 Td[(X)-326(2006)]TJ
/F16 11.9552 Tf 39.992 0 Td[(Ahoi!)]TJ
ET
1 0 0 1 72 507.702 cm
 q 0.54967 0 0 0.54967 0 0 cm
q
821 0 0 427 0 0 cm
/Im1 Do
Q
 Q
endstream
endobj
```

The page stream is made up from various operators with parameters.

```
3 0 obj <<
/Type /Page /Contents 4 0 R /Resources 2 0 R
/MediaBox [0 0 595.276 841.89]
/Parent 11 0 R
>> endobj
```

```
2 0 obj <<
/Font << /F15 7 0 R /F16 10 0 R >>
/XObject << /Im1 1 0 R >> /ProcSet [ /PDF /Text /ImageC ]
>> endobj
```

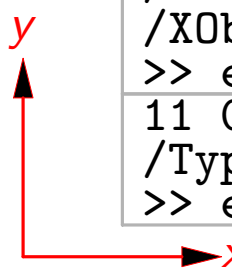
```
11 0 obj <<
/Type /Pages /Count 1 /Kids [3 0 R]
>> endobj
```

PDF tour — reading the page stream

```
4 0 obj <<
/Length 227
>>
stream
BT
/F15 11.9552 Tf 72 757.935 Td[(EuroT)]TJ
31.376 -2.574 Td[(E)]TJ 6.501 2.574 Td[(X)-326(2006)]TJ
/F16 11.9552 Tf 39.992 0 Td[(Ahoi!)]TJ
ET
1 0 0 1 72 507.702 cm
 q 0.54967 0 0 0.54967 0 0 cm
q
821 0 0 427 0 0 cm
/Im1 Do
Q
 Q
endstream
endobj
```

The origin is initially set to the lower left media corner.

```
3 0 obj <<
/Type /Page /Contents 4 0 R /Resources 2 0 R
/MediaBox [0 0 595.276 841.89]
/Parent 11 0 R
>> endobj
2 0 obj <<
/Font << /F15 7 0 R /F16 10 0 R >>
/XObject << /Im1 1 0 R >> /ProcSet [ /PDF /Text /ImageC ]
>> endobj
11 0 obj <<
/Type /Pages /Count 1 /Kids [3 0 R]
>> endobj
```



PDF tour — reading the page stream

```
4 0 obj <<
/Length 227
>>
stream
BT
/F15 11.9552 Tf 72 757.935 Td[(EuroT)]TJ
31.376 -2.574 Td[(E)]TJ 6.501 2.574 Td[(X)-326(2006)]TJ
/F16 11.9552 Tf 39.992 0 Td[(Ahoi!)]TJ
ET
1 0 0 1 72 507.702 cm
 q 0.54967 0 0 0.54967 0 0 cm
q
821 0 0 427 0 0 cm
/Im1 Do
Q
 Q
endstream
endobj
```

There are text sections.

(here happens the ‘typesetting’)

```
3 0 obj <<
/Type /Page /Contents 4 0 R /Resources 2 0 R
/MediaBox [0 0 595.276 841.89]
/Parent 11 0 R
>> endobj
```

```
2 0 obj <<
/Font << /F15 7 0 R /F16 10 0 R >>
/XObject << /Im1 1 0 R >> /ProcSet [ /PDF /Text /ImageC ]
>> endobj
```

```
11 0 obj <<
/Type /Pages /Count 1 /Kids [3 0 R]
>> endobj
```

PDF tour — reading the page stream

```
4 0 obj <<
/Length 227
>>
stream
BT
/F15 11.9552 Tf 72 757.935 Td[(EuroT)]TJ
31.376 -2.574 Td[(E)]TJ 6.501 2.574 Td[(X)-326(2006)]TJ
/F16 11.9552 Tf 39.992 0 Td[(Ahoi!)]TJ
ET
1 0 0 1 72 507.702 cm
 q 0.54967 0 0 0.54967 0 0 cm
q
821 0 0 427 0 0 cm
/Im1 Do
Q
 Q
endstream
endobj
```

There are coordinate transforms
(this is a movement)

```
3 0 obj <<
/Type /Page /Contents 4 0 R /Resources 2 0 R
/MediaBox [0 0 595.276 841.89]
/Parent 11 0 R
>> endobj
```

```
2 0 obj <<
/Font << /F15 7 0 R /F16 10 0 R >>
/XObject << /Im1 1 0 R >> /ProcSet [ /PDF /Text /ImageC ]
>> endobj
```

```
11 0 obj <<
/Type /Pages /Count 1 /Kids [3 0 R]
>> endobj
```

PDF tour — reading the page stream

```
4 0 obj <<
/Length 227
>>
stream
BT
/F15 11.9552 Tf 72 757.935 Td[(EuroT)]TJ
31.376 -2.574 Td[(E)]TJ 6.501 2.574 Td[(X)-326(2006)]TJ
/F16 11.9552 Tf 39.992 0 Td[(Ahoi!)]TJ
ET
1 0 0 1 72 507.702 cm
 q 0.54967 0 0 0.54967 0 0 cm
q
821 0 0 427 0 0 cm
/Im1 Do
Q
 Q
endstream
endobj
```

There are coordinate transforms
(these are scalings)

```
3 0 obj <<
/Type /Page /Contents 4 0 R /Resources 2 0 R
/MediaBox [0 0 595.276 841.89]
/Parent 11 0 R
>> endobj
```

```
2 0 obj <<
/Font << /F15 7 0 R /F16 10 0 R >>
/XObject << /Im1 1 0 R >> /ProcSet [ /PDF /Text /ImageC ]
>> endobj
```

```
11 0 obj <<
/Type /Pages /Count 1 /Kids [3 0 R]
>> endobj
```

PDF tour — reading the page stream

```
4 0 obj <<
/Length 227
>>
stream
BT
/F15 11.9552 Tf 72 757.935 Td[(EuroT)]TJ
31.376 -2.574 Td[(E)]TJ 6.501 2.574 Td[(X)-326(2006)]TJ
/F16 11.9552 Tf 39.992 0 Td[(Ahoi!)]TJ
ET
1 0 0 1 72 507.702 cm
 q 0.54967 0 0 0.54967 0 0 cm
q
821 0 0 427 0 0 cm
/Im1 Do
Q
 Q
endstream
endobj
```

Grouping limits operator scopes.
(e. g. scaling of /Im1)

```
3 0 obj <<
/Type /Page /Contents 4 0 R /Resources 2 0 R
/MediaBox [0 0 595.276 841.89]
/Parent 11 0 R
>> endobj
```

```
2 0 obj <<
/Font << /F15 7 0 R /F16 10 0 R >>
/XObject << /Im1 1 0 R >> /ProcSet [ /PDF /Text /ImageC ]
>> endobj
```

```
11 0 obj <<
/Type /Pages /Count 1 /Kids [3 0 R]
>> endobj
```

PDF tour — reading the page stream

```
4 0 obj <<
/Length 227
>>
stream
BT
/F15 11.9552 Tf 72 757.935 Td[(EuroT)]TJ
31.376 -2.574 Td[(E)]TJ 6.501 2.574 Td[(X)-326(2006)]TJ
/F16 11.9552 Tf 39.992 0 Td[(Ahoi!)]TJ
ET
1 0 0 1 72 507.702 cm
q 0.54967 0 0 0.54967 0 0 cm
q
821 0 0 427 0 0 cm
/Im1 Do
Q
Q
endstream
endobj
```

Grouping limits operator scopes.
(e. g. scaling of /Im1)

```
3 0 obj <<
/Type /Page /Contents 4 0 R /Resources 2 0 R
/MediaBox [0 0 595.276 841.89]
/Parent 11 0 R
>> endobj
```

```
2 0 obj <<
/Font << /F15 7 0 R /F16 10 0 R >>
/XObject << /Im1 1 0 R >> /ProcSet [ /PDF /Text /ImageC ]
>> endobj
```

```
11 0 obj <<
/Type /Pages /Count 1 /Kids [3 0 R]
>> endobj
```

PDF tour — reading the page stream

```
4 0 obj <<
/Length 227
>>
stream
BT
/F15 11.9552 Tf 72 757.935 Td[(EuroT)]TJ
31.376 -2.574 Td[(E)]TJ 6.501 2.574 Td[(X)-326(2006)]TJ
/F16 11.9552 Tf 39.992 0 Td[(Ahoi!)]TJ
ET
1 0 0 1 72 507.702 cm
 q 0.54967 0 0 0.54967 0 0 cm
q
821 0 0 427 0 0 cm
/Im1 Do
Q
 Q
endstream
endobj
```

Begin and end text also groups regarding transforms.

```
3 0 obj <<
/Type /Page /Contents 4 0 R /Resources 2 0 R
/MediaBox [0 0 595.276 841.89]
/Parent 11 0 R
>> endobj
```

```
2 0 obj <<
/Font << /F15 7 0 R /F16 10 0 R >>
/XObject << /Im1 1 0 R >> /ProcSet [ /PDF /Text /ImageC ]
>> endobj
```

```
11 0 obj <<
/Type /Pages /Count 1 /Kids [3 0 R]
>> endobj
```

PDF tour — putting text on the page

```
4 0 obj <<
/Length 227
>>
stream
BT
/F15 11.9552 Tf 72 757.935 Td[(EuroT)]TJ
31.376 12.574 Td[(E)]TJ 6.501 2.574 Td[(X)-326(2006)]TJ
/F16 11.9552 Tf 39.992 0 Td[(Ahoi!)]TJ
ET
1 0 0 1 72 507.702 cm
 q 0.54967 0 0 0.54967 0 0 cm
q
821 0 0 427 0 0 cm
/Im1 Do
Q
 Q
endstream
endobj
```

Select a font and its size.
(operator `Tf`)

```
3 0 obj <<
/Type /Page /Contents 4 0 R /Resources 2 0 R
/MediaBox [0 0 595.276 841.89]
/Parent 11 0 R
>> endobj
```

```
2 0 obj <<
/Font << /F15 7 0 R /F16 10 0 R >>
/XObject << /Im1 1 0 R >> /ProcSet [ /PDF /Text /ImageC ]
>> endobj
```

```
11 0 obj <<
/Type /Pages /Count 1 /Kids [3 0 R]
>> endobj
```

PDF tour — putting text on the page

```
4 0 obj <<
/Length 227
>>
stream
BT
/F15 11.9552 Tf 72 757.935 Td[(EuroT)]TJ
31.376 -2.574 Td[(E)]TJ 6.501 2.574 Td[(X)-326(2006)]TJ
/F16 11.9552 Tf 39.992 0 Td[(Ahoi!)]TJ
ET
1 0 0 1 72 507.702 cm
 q 0.54967 0 0 0.54967 0 0 cm
q
821 0 0 427 0 0 cm
/Im1 Do
Q
 Q
endstream
endobj
```

Move relative to current point.
(operator Td)

```
3 0 obj <<
/Type /Page /Contents 4 0 R /Resources 2 0 R
/MediaBox [0 0 595.276 841.89]
/Parent 11 0 R
>> endobj
```

```
2 0 obj <<
/Font << /F15 7 0 R /F16 10 0 R >>
/XObject << /Im1 1 0 R >> /ProcSet [ /PDF /Text /ImageC ]
>> endobj
```

```
11 0 obj <<
/Type /Pages /Count 1 /Kids [3 0 R]
>> endobj
```

PDF tour — putting text on the page

```
4 0 obj <<
/Length 227
>>
stream
BT
/F15 11.9552 Tf 72 757.935 Td[(EuroT)]TJ
31.376 -2.574 Td[(E)]TJ 6.501 2.574 Td[(X)-326(2006)]TJ
/F16 11.9552 Tf 39.992 0 Td[(Ahoi!)]TJ
ET
1 0 0 1 72 507.702 cm
 q 0.54967 0 0 0.54967 0 0 cm
q
821 0 0 427 0 0 cm
/Im1 Do
Q
 Q
endstream
endobj
```

Output string of glyphs.
(operator TJ)

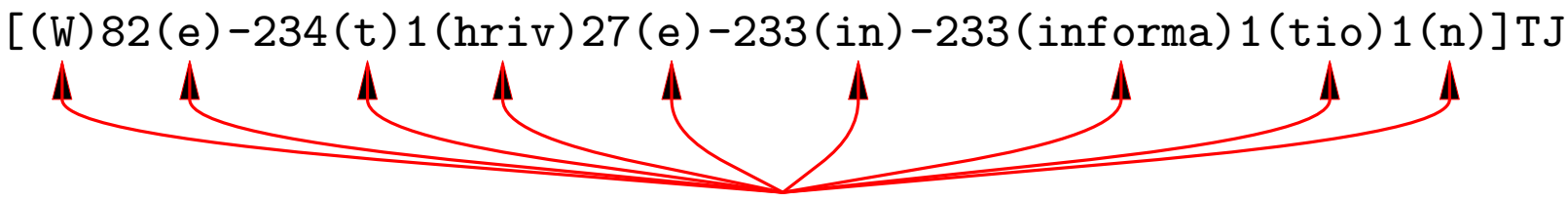
```
3 0 obj <<
/Type /Page /Contents 4 0 R /Resources 2 0 R
/MediaBox [0 0 595.276 841.89]
/Parent 11 0 R
>> endobj
```

```
2 0 obj <<
/Font << /F15 7 0 R /F16 10 0 R >>
/XObject << /Im1 1 0 R >> /ProcSet [ /PDF /Text /ImageC ]
>> endobj
```

```
11 0 obj <<
/Type /Pages /Count 1 /Kids [3 0 R]
>> endobj
```

PDF tour — The glyph positioning operator TJ

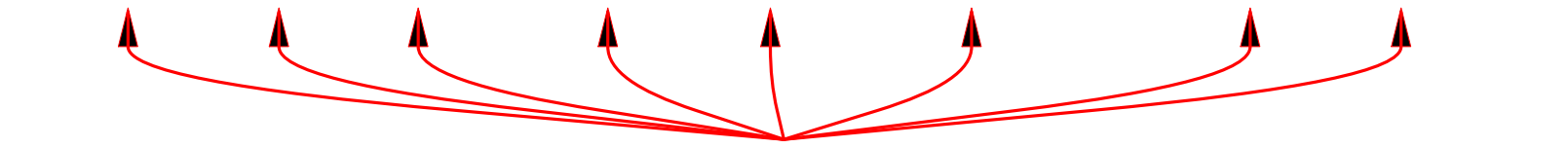
This takes an array as operand, glyphs alternating with movements:



The glyphs in brackets (= implicit movements).

PDF tour — The glyph positioning operator TJ

This takes an array as operand, glyphs alternating with movements:

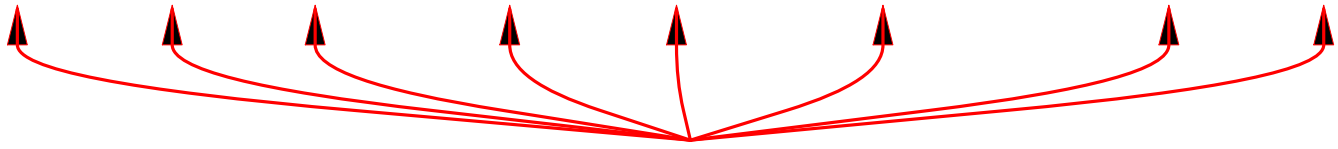


Explicit movements (units of fontsize/1000; move left = pos.).

PDF tour — The glyph positioning operator TJ

This takes an array as operand, glyphs alternating with movements:

```
[(W)82(e)-234(t)1(hriv)27(e)-233(in)-233(informa)1(tio)1(n)]TJ
```



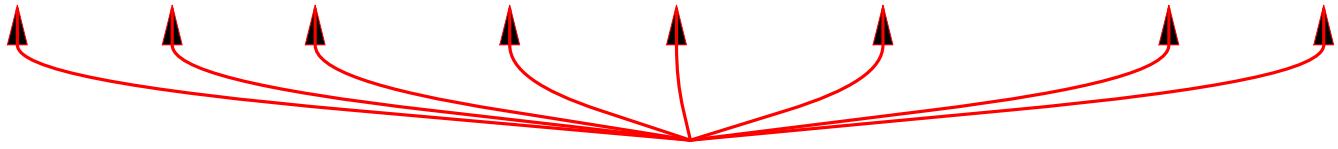
Explicit movements (units of fontsize/1000; move left = pos.).

- The implicit movements per glyph (glyph width) are taken from the `/Widths` array in the font object.
- These `/Width` entries are from the TFM file.

PDF tour — The glyph positioning operator TJ

This takes an array as operand, glyphs alternating with movements:

```
[(W)82(e)-234(t)1(hriv)27(e)-233(in)-233(informa)1(tio)1(n)]TJ
```



Explicit movements (units of fontsize/1000; move left = pos.).

- The implicit movements per glyph (glyph width) are taken from the `/Widths` array in the font object.
- These `/Width` entries are from the TFM file.
- Example with `cmr12` font on next slide...

Glyph positioning as of pdfTEX-1.30.6

```
BT
/F15 11.955 Tf 72 757.935 Td[(W)82(e)-234(t)1(hriv)
27(e)-233(in)-233(informa)1(tio)1(n|)-1(|thic)27(k)
-232(w)27(orlds)-233(b)-27(ec)-1(a)1(use)-234(o)
1(f)-233(our)-233(ma)1(rv)27(elo)1(us)-234(a)
1(nd)-233(ev)27(eryda)28(y)-233(capa)1(cit)27(y)
]TJ 0 -14.446 Td[(to)-296(select,)-303(edit,)-302(single)
-296(out,)-302(structur)1(e)-1(,)-302(hig)1(hligh)
27(t,)-302(gr)1(oup,)-302(pair,)-302(merg)1(e)-1(,)
-302(ha)1(rmonize,)-303(syn)28(thes)-1(i)1(z)-1(e,)
]TJ 0 -14.446 Td[(fo)-27(cus,)-453(org)1(anize,)-453(c)
-1(o)1(ndense)-1(,)-453(r)1(e)-1(duce,)-453(b)-27(oil)
-427(do)27(wn,)-453(c)27(ho)-27(o)1(s)-1(e,)-453(cat)
1(e)-1(g)1(orize,)-453(cat)1(alog)1(,)-453(classif)-1(y)
more lines...
ET
```

The glyph /Widths array.

```
6 0 obj <<
/Type /Font /Subtype /Type1
/Encoding 8 0 R /FirstChar 11 /LastChar 124
/Widths 9 0 R
/BaseFont /HNUPFH+CMR12 /FontDescriptor 4 0 R
>> endobj
9 0 obj
[571 544 544 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 272 326 272 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
1006 0 0 0 0 0 0 0 0 0 490 544 435 544 435 299 490 544
272 0 517 272 816 544 490 544 0 381 386 381 544 517 707
517 517 435 0 979 ]
endobj
```

Glyph positioning as of pdfTEX-1.30.6

```
BT
/F15 11.955 Tf 72 757.935 Td[(W)82(e)-234(t)1(hriv)
27(e)-233(in)-233(informa)1(tio)1(n|)-1(|thic)27(k)
-232(w)27(orlds)-233(b)-27(ec)-1(a)1(use)-234(o)
1(f)-233(our)-233(ma)1(rv)27(elo)1(us)-234(a)
1(nd)-233(ev)27(eryda)28(y)-233(capa)1(cit)27(y)
]TJ 0 -14.446 Td[(to)-296(select,)-303(edit,)-302(single)
-296(out,)-302(structur)1(e)-1(,)-302(hig)1(hligh)
27(t,)-302(gr)1(oup,)-302(pair,)-302(merg)1(e)-1(,)
-302(ha)1(rmonize,)-303(syn)28(thes)-1(i)1(z)-1(e,)
]TJ 0 -14.446 Td[(fo)-27(cus,)-453(org)1(anize,)-453(c)
-1(o)1(ndense)-1(,)-453(r)1(e)-1(duce,)-453(b)-27(oil)
-427(do)27(wn,)-453(c)27(ho)-27(o)1(s)-1(e,)-453(cat)
1(e)-1(g)1(orize,)-453(cat)1(alog)1(,)-453(classif)-1(y)
more lines...
ET
```

Width of the letter ‘a’ in cmr12.

```
6 0 obj <<
/Type /Font /Subtype /Type1
/Encoding 8 0 R /FirstChar 11 /LastChar 124
/Widths 9 0 R
/BaseFont /HNUPFH+CMR12 /FontDescriptor 4 0 R
>> endobj
```

```
9 0 obj
[571 544 544 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 272 326 272 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
1006 0 0 0 0 0 0 0 0 0 490 544 435 544 435 299 490 544
272 0 517 272 816 544 490 544 0 381 386 381 544 517 707
517 517 435 0 979 ]
endobj
```

Glyph positioning as of pdfTeX-1.30.6

```
BT
/F15 11.955 Tf 72 757.935 Td[(W)82(e)-234(t)1(hriv)
27(e)-233(in)-233(informa)1(tio)1(n|)-1(|thic)27(k)
-232(w)27(orlds)-233(b)-27(ec)-1(a)1(use)-234(o)
1(f)-233(our)-233(ma)1(rv)27(elo)1(us)-234(a)
1(nd)-233(ev)27(eryda)28(y)-233(capa)1(cit)27(y)
]TJ 0 -14.446 Td[(to)-296(select,)-303(edit,)-302(single)
-296(out,)-302(structur)1(e)-1(,)-302(hig)1(hligh)
27(t,)-302(gr)1(oup,)-302(pair,)-302(merg)1(e)-1(,)
-302(ha)1(rmonize,)-303(syn)28(thes)-1(i)1(z)-1(e,)
]TJ 0 -14.446 Td[(fo)-27(cus,)-453(org)1(anize,)-453(c)
-1(o)1(ndense)-1(,)-453(r)1(e)-1(duce,)-453(b)-27(oil)
-427(do)27(wn,)-453(c)27(ho)-27(o)1(s)-1(e,)-453(cat)
1(e)-1(g)1(orize,)-453(cat)1(alog)1(,)-453(classif)-1(y)
more lines...
ET
```

Look: These are no TeX kerns!

```
6 0 obj <<
/Type /Font /Subtype /Type1
/Encoding 8 0 R /FirstChar 11 /LastChar 124
/Widths 9 0 R
/BaseFont /HNUPFH+CMR12 /FontDescriptor 4 0 R
>> endobj
```

```
9 0 obj
[571 544 544 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 272 326 272 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
1006 0 0 0 0 0 0 0 0 0 490 544 435 544 435 299 490 544
272 0 517 272 816 544 490 544 0 381 386 381 544 517 707
517 517 435 0 979 ]
endobj
```

Glyph positioning as of pdfTEX-1.30.6

Observation:

- Many tiny correcting movements ± 1 between strings of glyphs.
- These happen particularly for CM and LM fonts...

Glyph positioning as of pdfTEX-1.30.6

Observation:

- Many tiny correcting movements ± 1 between strings of glyphs.
- These happen particularly for CM and LM fonts...
- ...but not for the 35 Adobe standard Type 1 fonts!

Glyph positioning as of pdfTeX-1.30.6

Observation:

- Many tiny correcting movements ± 1 between strings of glyphs.
- These happen particularly for CM and LM fonts...
- ...but not for the 35 Adobe standard Type 1 fonts!
- Same also with older pdfTeX versions.

Glyph positioning as of pdfTeX-1.30.6

Observation:

- Many tiny correcting movements ± 1 between strings of glyphs.
- These happen particularly for CM and LM fonts...
- ...but not for the 35 Adobe standard Type 1 fonts!
- Same also with older pdfTeX versions.
- These movements with CM/LM fonts look strange.
- Why are they there?

Glyph positioning as of pdfTEX-1.30.6

The PDF reader's view:

- The PDF reader positions stuff on the page only from the info in the PDF file.

Glyph positioning as of pdfTeX-1.30.6

The PDF reader's view:

- The PDF reader positions stuff on the page only from the info in the PDF file.
- Movements are incremental.
- The PDF file gives movements as decimal real numbers.
- They might be *rounded* to some precision, e. g. by pdfTeX.
- The PDF reader takes these numbers as exact.

Glyph positioning as of pdfTeX-1.30.6

The PDF reader's view:

- The PDF reader positions stuff on the page only from the info in the PDF file.
- Movements are incremental.
- The PDF file gives movements as decimal real numbers.
- They might be *rounded* to some precision, e. g. by pdfTeX.
- The PDF reader takes these numbers as exact.

How far the PDF reader moves forward after placement of a glyph:

- The `/Widths` array tells this for any used glyph.

Glyph positioning as of pdfTeX-1.30.6

pdfTeX's view:

- pdfTeX internally keeps positions on a 'scaled point' raster.
- These TeX positions are exact, no rounding.
- No accumulation of rounding errors in the TeX coordinate system.

Glyph positioning as of pdfTeX-1.30.6

pdfTeX's view:

- pdfTeX internally keeps positions on a 'scaled point' raster.
- These TeX positions are exact, no rounding.
- No accumulation of rounding errors in the TeX coordinate system.
- But pdfTeX *has to* round values when writing to PDF file.

Glyph positioning as of pdfTeX-1.30.6

pdfTeX's view:

- pdfTeX internally keeps positions on a 'scaled point' raster.
- These TeX positions are exact, no rounding.
- No accumulation of rounding errors in the TeX coordinate system.
- But pdfTeX *has to* round values when writing to PDF file.

pdfTeX's strategy against accumulation of positional rounding errors of the PDF reader:

- pdfTeX constantly keeps track of *two* positions:
  1. The position in the TeX coordinate system.

Glyph positioning as of pdfTeX-1.30.6

pdfTeX's view:

- pdfTeX internally keeps positions on a 'scaled point' raster.
- These TeX positions are exact, no rounding.
- No accumulation of rounding errors in the TeX coordinate system.
- But pdfTeX *has to* round values when writing to PDF file.

pdfTeX's strategy against accumulation of positional rounding errors of the PDF reader:

- pdfTeX constantly keeps track of *two* positions:
  1. The position in the TeX coordinate system.
  2. The position where the PDF reader *thinks* it is on the page.

Glyph positioning as of pdfTeX-1.30.6

pdfTeX's view:

- pdfTeX internally keeps positions on a 'scaled point' raster.
- These TeX positions are exact, no rounding.
- No accumulation of rounding errors in the TeX coordinate system.
- But pdfTeX *has to* round values when writing to PDF file.

pdfTeX's strategy against accumulation of positional rounding errors of the PDF reader:

- pdfTeX constantly keeps track of *two* positions:
  1. The position in the TeX coordinate system.
  2. The position where the PDF reader *thinks* it is on the page.
- pdfTeX then can correct a rounding error in the next movement command.

Glyph positioning as of pdfTEX-1.30.6

pdfTEX's algorithm to prevent accumulation of position rounding errors:

- Output glyph into TJ array. (a

Glyph positioning as of pdfTEX-1.30.6

pdfTEX's algorithm to prevent accumulation of position rounding errors:

- Output glyph into TJ array. (a
- Update TEX position by glyph width from TFM file.
- Update PDF position by glyph width from the `/Widths` array.

Glyph positioning as of pdfTeX-1.30.6

pdfTeX's algorithm to prevent accumulation of position rounding errors:

- Output glyph into TJ array. `(a`
- Update TeX position by glyph width from TFM file.
- Update PDF position by glyph width from the `/Widths` array.
- Calculate position error between PDF and TeX positions.
- If (error $>$ 1/2000 fontsize) then
  - Output correcting movement into TJ array. `(a)-1`

Glyph positioning as of pdfTeX-1.30.6

pdfTeX's algorithm to prevent accumulation of position rounding errors:

- Output glyph into TJ array. `(a`
- Update TeX position by glyph width from TFM file.
- Update PDF position by glyph width from the `/Widths` array.
- Calculate position error between PDF and TeX positions.
- If (error $>$ 1/2000 fontsize) then
  - Output correcting movement into TJ array. `(a)-1`
  - Update PDF position accordingly.

Glyph positioning as of pdfTEX-1.30.6

Still: Why are there these tiny ± 1 corrections in the TJ matrix for CM and LM fonts?

Glyph positioning as of pdfTEX-1.30.6

Still: Why are there these tiny ± 1 corrections in the TJ matrix for CM and LM fonts? Reason:

- CM fonts are not designed on a 1/1000 fontsize *raster*.
- CM predates PostScript fonts.

Glyph positioning as of pdfTEX-1.30.6

Still: Why are there these tiny ± 1 corrections in the TJ matrix for CM and LM fonts? Reason:

- CM fonts are not designed on a 1/1000 fontsize *raster*.
- CM predates PostScript fonts.
- Glyph 'a' in `cmr12.tfm` has width 0.489578 $\times$ fontsize
- In the `/Widths` array there is 490, not 489.578!
- pdfTEX-1.30.6 rounds to *integer* `/Width` array values.

Glyph positioning as of pdfTEX-1.30.6

Still: Why are there these tiny ± 1 corrections in the TJ matrix for CM and LM fonts? Reason:

- CM fonts are not designed on a 1/1000 fontsize *raster*.
- CM predates PostScript fonts.
- Glyph 'a' in `cmr12.tfm` has width 0.489578 $\times$ fontsize
- In the `/Widths` array there is 490, not 489.578!
- pdfTEX-1.30.6 rounds to *integer* `/Width` array values.
- After three 'a' the accumulated rounding error is $> 1/1000$ fontsize!
- That might be the reason for these `)1(` and `)-1(` corrections.

Glyph positioning as of pdfTEX-1.30.6

Still: Why are there these tiny ± 1 corrections in the TJ matrix for CM and LM fonts? Reason:

- CM fonts are not designed on a 1/1000 fontsize *raster*.
- CM predates PostScript fonts.
- Glyph 'a' in `cmr12.tfm` has width 0.489578 $\times$ fontsize
- In the `/Widths` array there is 490, not 489.578!
- pdfTEX-1.30.6 rounds to *integer* `/Width` array values.
- After three 'a' the accumulated rounding error is $> 1/1000$ fontsize!
- That might be the reason for these `)1(` and `)-1(` corrections.
- Most standard PostScript fonts are designed on a 1/1000 fontsize raster, therefore integer `/Widths` array values are *exact*.

Glyph positioning as of pdfTEX-1.30.6

A rather straight-forward solution:

- Add one digit after the decimal point for the `/Widths` entries.
- Then the rounding error accumulates ten times slower.
- The `)1(` and `)-1(` corrections will happen about ten times less often — which is just ok.

Glyph positioning as of pdfTEX-1.30.6

A rather straight-forward solution:

- Add one digit after the decimal point for the `/Widths` entries.
- Then the rounding error accumulates ten times slower.
- The `)1(` and `)-1(` corrections will happen about ten times less often — which is just ok.

Side effects:

- Tidier page stream.
- Up to 3 % smaller PDF file when using CM or LM fonts.

Result on next slide. . .

Glyph positioning as of pdfTEX-1.40.0

```
BT
/F15 11.9552 Tf 72 757.935 Td[(W)82(e)-233(thriv)27(e)
-233(in)-233(information||thic)27(k)-233(w)28(orlds)
-233(b)-27(ecause)-233(of)-233(our)-233(marv)27(elous)
-233(and)-233(ev)27(eryda)28(y)-233(capacit)27(y)
]TJ 0 -14.446 Td[(to)-296(select,)-303(ed)1(it,)
-303(single)-296(out,)-302(structure,)-302(highligh)
27(t,)-302(group,)-303(pair,)-302(merge,)-302(harmonize,)
-302(syn)27(thesize,)]TJ 0 -14.446 Td[(fo)-27(cus,)
-453(organize,)-453(condense,)-453(reduce,)-453(b)
-27(oil)-428(do)28(wn,)-453(c)27(ho)-27(ose,)
-453(categorize,)-453(catalog,)-453(classify)82(,)
-453(list,)]TJ 0 -14.446 Td[(abstract,)-405(scan,)-405(lo)
more lines...
ET
```

Width of the letter ‘a’ in cmr12.

```
6 0 obj <<
/Type /Font /Subtype /Type1
/FirstChar 11 /LastChar 124
/Widths 8 0 R
/BaseFont /MDVKCM+CMR12 /FontDescriptor 4 0 R
>> endobj
```

```
8 0 obj
[571.2 544 544 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 272 326.4 272 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 1006 0 0 0 0 0 0 0 0 0 [489.6] 544 435.2 544
435.2 299.2 489.6 544 272 0 516.8 272 816 544 489.6
544 0 380.8 386.2 380.8 544 516.8 707.2 516.8 516.8
435.2 0 979.2 ]
endobj
```

Conclusion

- PDF is no inscrutable data format.
- PDF files generated by pdfTeX are rather readable with a standard text editor when `\pdfcompresslevel=0`.
- An example was presented on a tiny improvement of the PDF file quality.

Conclusion

- PDF is no inscrutable data format.
- PDF files generated by pdfTEX are rather readable with a standard text editor when `\pdfcompresslevel=0`.
- An example was presented on a tiny improvement of the PDF file quality.

Thank you very much for your attention.