

# extractpdfmark

Extract page mode and named destinations as PDFmark from PDF

<https://github.com/trueroad/extractpdfmark>

<http://www.ctan.org/pkg/extractpdfmark>

TeX のようなシステムを使って PDF のドキュメントを作る場合、図としてたくさんの小さな PDF を用意して、メインの PDF に貼り付けるようなことが行われます。このとき、図の PDF では同じフォントを使っていることが多いと思います。

図の PDF にフォントのサブセットが埋め込まれていた場合、TeX システムはそのフォントをそのままメインの PDF に埋め込みます。結果として、メイン PDF には重複した同じフォントの異なるサブセットが埋め込まれます。サブセットが異なるため重複したフォントを削除することはできません。よってメイン PDF のファイルサイズは、かなり大きくなってしまいます。

一方、図の PDF にフォントをフルセットで埋め込んだ場合、TeX システムはフルセットのフォントをメインの PDF へ埋め込みます。メイン PDF にはたくさんの重複したフォントが埋め込まれますが、すべて同じフルセットのフォントです。そのため、Ghostscript を使えば重複を削除することができ、これによってメイン PDF のサイズを低減することができます。(注意: Ghostscript 9.17 から 9.21 は重複フォント削除のために

`-dPDFDontUseFontObjectNum`

コマンドラインオプションが必要です。Ghostscript 9.22 以降は重複フォント削除ができなくなってきたため、フルセット埋め込みの方法を採ることができません。この場合は、次に示すフォントを一切埋め込まない方法を採る必要があります。)

さらに、図の PDF に一切フォントを埋め込まなかった場合、TeX システムはフォントが欠けたメイン PDF を出力します。この場合、Ghostscript は必要なフォントを埋め込むことができるので、必要なディスク容量を劇的に減らすことができるようになります。

いずれの方法でも、Ghostscript へ TeX システムの出力したメイン PDF を入力し、最終的な PDF を出力することになります。しかし、残念ながら、この処理で Ghostscript は PDF のページモードやリンクの宛先名を残してくれません。そのため、最終的な PDF は意図したとおりに開かれなかったり、リモートからのリンクが機能しなくなったりしてしまいます。

[http://bugs.ghostscript.com/show\\_bug.cgi?id=696943](http://bugs.ghostscript.com/show_bug.cgi?id=696943)

[http://bugs.ghostscript.com/show\\_bug.cgi?id=695760](http://bugs.ghostscript.com/show_bug.cgi?id=695760)

## extractpdfmark

Extract PDFmark は PDF からページモードやリンクの宛先名を抽出し、PDFmark として保存することができます。このツールを使うと、ページモードやリンク先名を残したまま、サイズの小さい PDF を得ることができます。

Usage:

```
extractpdfmark TeX-Output.pdf > PDFmark.ps
rungs -q -dBATCH -dNOPAUSE -sDEVICE=pdfwrite \
      -sOutputFile=Final.pdf TeX-Output.pdf PDFmark.ps
```

Git repository:

<https://github.com/trueroad/extractpdfmark>

Licence:

Copyright (C) 2016-2019 Masamichi Hosoda

Extract PDFmark is free software: you can redistribute it and/or modify it under the terms of the GNU General Public License as published by the Free Software Foundation, either version 3 of the License, or (at your option) any later version.

Extract PDFmark is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY; without even the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU General Public License for more details.

You should have received a copy of the GNU General Public License along with Extract PDFmark.

If not, see [<http://www.gnu.org/licenses/>](http://www.gnu.org/licenses/).