YAWP 0.5.4 MANUAL

Yet Another Word Processor

2022-04-04

Carlo Alessandro Verre

carlo.alessandro.verre@gmail.com

I sound my barbaric yawp over the roofs of the world

Walt Whitman

CONTENTS

1. INTRODUCTION

1.1. WHAT IS YAWP?

The name "yawp" here means Yet Another Word Processor, and yawp is an automatic word processor for plain text files, with PDF output. If you really need all the features of a full-fledged WYSIWYG word processor like LibreOffice Writer, yawp is not for you. But if you just want to create a simple quick-and-dirty no-frills document, with yawp you can:

- edit a text file by your favorite editor
- run yawp in order to:
    - backup read format and rewrite the text file
    - export the text file in a PDF file
    - open the PDF file for check or print
- possibly go back to the editor and update the text file, or finish
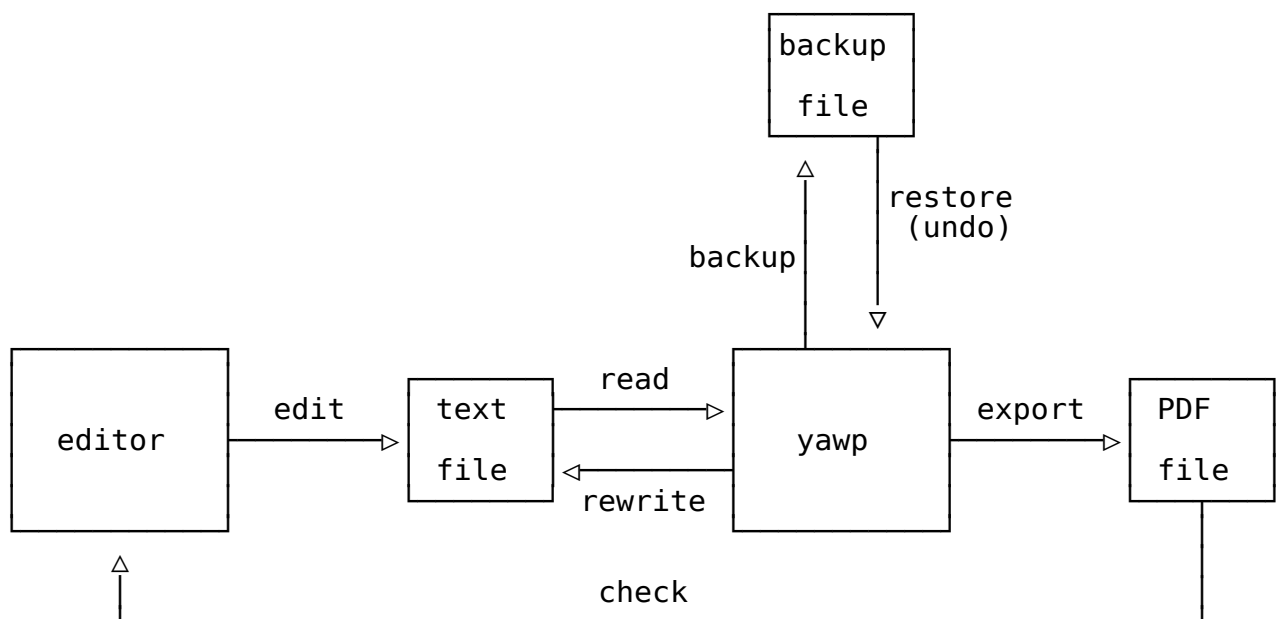
Main features are:

- yawp processes in place a single text file, hereinafter referred to simply as the "file"
- yawp before processing makes a timestamped backup of the file, allowing undo operation (see '2. Usage Modes')
- yawp processing is driven by the text in the file and by arguments only, not by commands or tags embedded in text
- yawp justifies (see '3. Justification') text at left and right in:
    - unindented paragraphs
    - dot-marked indented paragraphs (as this one)
- yawp accepts unjustified pictures (as schemas, tables and code examples) freely intermixed with text
- yawp adopts an ad hoc policy for Python files, formatting the docstrings but not the Python code (see '4. Python Files')
- yawp performs multi-level chapter renumbering (see '5.1. Numbered Chapters')
- yawp inserts an automatic contents chapter in the file (see '5.2. Contents Chapter')
- yawp recognizes relevant subjects (quoted by '"') and inserts an automatic index chapter in the file (see 5.3. Index Chapter)
- yawp cuts the file in pages, by automatic insertion of two-lines page headers (see '6. Paging')
- yawp also has some limited graphic feature, you can sketch pictures with segments (by '`') and arrowheads (by '^'), yawp redraws them by proper graphic characters (as in the picture above, see '7. Graphics')

- yawp exports the resulting lines in PDF format, with control over
  character size and page layout, and opens for you the generated PDF
  file, allowing preview and printing (see '8. Pdf Exporting')
- yawp corrects errors made by CUPS-PDF about font size and page
  margins (see '9. Corrections')
- yawp is "stable", namely if after a yawp execution you run yawp again
  on the same file with the same arguments then the file content
  doesn't change (except date and time in page headers, see '7.
  Paging')
- as a beta release, yawp 0.5.4. contains debug functionalities not
  aimed at the end user, they will disappear in some future release
  (see '10.1. Debugging')

Believe or not, anything has been kept as simple as possible.

As an example, this documentation you're reading has been created as
yawp.pdf from yawp.txt by typing:

    | $ yawp -v -w 75 -F -E 'Yawp 0.5.4 Manual' yawp.txt

Other examples are scattered below.

1.2. INSTALLATION

CUPS-PDF provides a PDF writer backend to CUPS, and yawp needs it to export
the file in PDF format. For example, if your Linux belongs to the Debian
family, type:

    | $ sudo apt-get -y update
    | $ sudo apt-get -y install printer-driver-cups-pdf

If you don't have pip, type:

    | $ sudo apt-get -y install python3-pip

If you type at terminal:

    | $ pip3 install --upgrade yawp

this command will:

- install current version of yawp if not present
- upgrade yawp to the current version if already installed

If you see a message of this type:

    | WARNING: The script yawp is installed in ... which is not on PATH.

don't worry, a reboot should fix the problem.

1.3. MESSAGES

During execution yawp can write three kinds of messages:

- an "information message" says what's going on (but only if -v is on)
- a "warning message" (starting with 'WARNING:') says what may be
  wrong, and processing continues
- an "error message" (starting with 'ERROR:') says what's wrong, file
  backup and rewriting don't take place and yawp execution is
  terminated.

All three types of messages are written on stderr, in order to avoid
interference with -p option, which writes the file on stdout.

When applicable, warning and error messages are preceded by position and
content of the offending line in the file.

2. USAGE MODES

General behaviour is controlled by the following arguments:

- "-h, --help": show a short help message and exit
- "-H, --manual": open this yawp-generated "yawp Manual" in PDF format and exit
- "-V, --version": show program's version number and exit
- "-v, --verbose": display information messages on stderr

Two arguments, -U and -N, define three usage modes. If neither -U nor -N is on, yawp runs in "standard mode":

- the file is read, eliminating:
    - page header lines
    - trailing blanks in lines
    - "horizontal tab" '\t' in lines, each replaced by four blanks
- the file content is formatted, see chapters 3. 4. 5. and 6.
- if -g is on, pictures are redrawn, see '7. Graphics' chapter
- the file is backed up in a timestamped copy
- the file is rewritten

- "-N, --no-format": leave the file unchanged

If -N is on, yawp runs in "no-format mode":

- the file is read, eliminating:
    - trailing blanks in lines
    - "horizontal tab" '\t' in lines, each replaced by four blanks
- if -g is on, pictures are redrawn, see '7. Graphics' chapter
- the file is backed up in a timestamped copy
- the file is rewritten

Note that the input page headers are eliminated in standard mode, while they are kept in no-format mode.

A text file is a sequence of lines, separated by line terminators. Both in standard and in no-format mode, file reading is in "universal newlines" mode, so three line terminators are accepted:

- "line feed" = '\n' (Unix-Linux-Apple)
- "carriage return" + "line feed" = '\r\n' (Microsoft-Windows)
- "carriage return" = '\r' (Apple old MacOS pre-OSX)

When the file is rewritten, the line terminator is always '\n'.

- "-U, --undo": restore the file from its previous version

If -U is on, yawp runs in "undo mode":

- the file is recovered from the most recent backup
- -g argument has no effect

BEWARE: undo operation can't be undone, the file goes back to the penultimate version and the last one is lost.

To turn on both -U and -N is not allowed.

In all three usage modes, the file is possibly written and/or exported, depending on -p and -P arguments:

- "-p, --print-file": at end write file on stdout

- "-P, --file-pdf": file name of exported PDF file ('0' = no PDF export, default: '%P/%f.pdf', see '8. Pdf Exporting' chapter)

3. JUSTIFICATION

Let's distinguish in four categories the input file lines:

- a line is an "empty line" if it contains no characters, all trailing blanks in input lines are stripped away, hence every input line containing only blanks becomes an empty line
- otherwise a line is a "dot line" if the first nonblank character is a "decimal point" '.' or a "black small circle" '•' followed by a blank (but on output such a '.' is always replaced by '•')
- otherwise a line is an "indented line" if it starts with a blank
- otherwise a line is an "unindented line"

Header lines inserted on output by yawp (see 7. Paging) are preliminarily canceled from input.

The formatting algorithm, driven by the input lines, oscillates between two states:

- "picture state", where input lines are directly written out as they are
- "text state", where input lines are accumulated into a paragraph buffer for further justification and writing at paragraph end

The picture state is the initial state. In this state, if the input is:
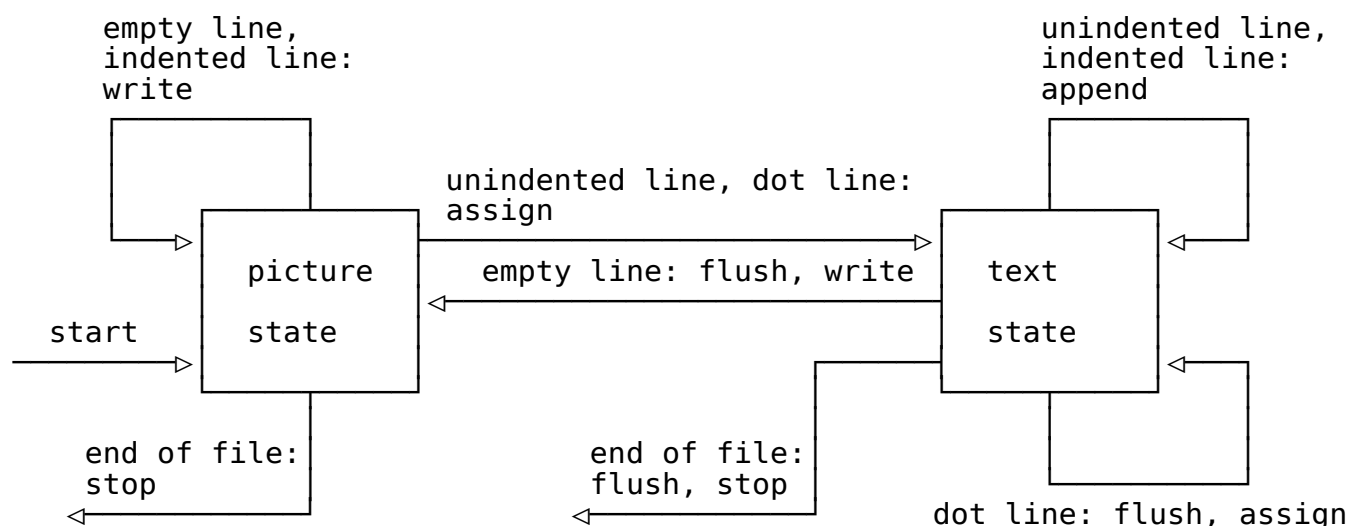
- an empty line or an indented line: the line is written out as is
- an unindented line: text state is entered, an "unindented paragraph" begins, the line is shrunk and assigned to the paragraph buffer, paragraph left indentation is set to zero
- a dot line: text state is entered, an "indented paragraph" begins, the line is shrunk and assigned to the paragraph buffer, paragraph left indentation is set to the position of initial dot character plus two
- end of input file: processing is terminated

Here to "shrink" a string means:

- to strip away all leading and trailing blanks
- and reduce each group of consecutive blanks to a single blank

When we are in text state, if the input line is:

- an empty line: the paragraph buffer is flushed (justified, written out and emptied), state goes back to picture state, the empty line is written out
- an indented or unindented line: the line is shrunk and appended to the paragraph buffer
- a dot line: paragraph buffer is flushed, a new paragraph is started, the line is shrunk and assigned to the paragraph buffer, paragraph left indentation is set to the position of initial dot plus two
- end of input file: paragraph buffer is flushed, processing is ended

----

```
      empty line,                                      unindented line,
      indented line:                                   indented line:
      write                                            append

          ┌─────────────┐                      ┌─────────────┐
          │             │                      │             │
          │  ┌───────────┴──┐  unindented line, dot line:  ┌──┴───────────┐  │
          └─▷│              │  assign             │              │◁─┘
             │   picture    ├────────────────────────────▷│     text     │
             │              │  empty line: flush, write    │              │
   start     │    state     │◁────────────────────────────┤    state     │◁─┐
   ──────────▷│              │                             │              │  │
             └──┬───────────┘                             └──────────┬───┘◁─┘
                │                                                     │
      end of file:                    end of file:                   │
      stop                            flush, stop                     │
                ◁                             ◁                       │
                                                          dot line: flush, assign
```

Actions associated to transitions are:

 • write: the input line is immediately written out unchanged as a
   "picture line"
 • assign: the input line is shrunk and assigned to the paragraph buffer
 • append: the input line is shrunk and appended to the paragraph buffer
 • flush: the paragraph buffer is flushed, namely is justified, written
   out as "text lines", and emptied
 • stop: formatting is finished

Max length of text lines can be controlled by:

 • "-w, --chars-per-line": line width in chars per line (default: 0 =
   automatic)

-w and -W (character width, see '8. PDF exporting') can be zero hence
"automatic", namely:

 • if -w is automatic and -W is not, -w is computed from -W
 • if -W is automatic and -w is not, -W is computed from -w
 • if both -w and -W are automatic,
     • -w is deduced from max line length in file (page headers are not
       considered)
     • -W is computed from -w

As an example, if you type:

   | $ echo BANNER >banner.txt

and impose only one character per line:

```
$ yawp -v -N -w 1 -p banner.txt
Correct: -L 43.325pt = 0.602in = 15.284mm = 1.528cm
Correct: -R 40.57pt = 0.563in = 14.312mm = 1.431cm
Correct: -T 53.814pt = 0.747in = 18.985mm = 1.898cm
Correct: -B 15.991pt = 0.222in = 5.641mm = 0.564cm
Read: yawp <-- '~/banner.txt'
    0 header lines, max 0 chars per line, 1 page
    1 body line, max 6 chars per line
    1 total line, max 6 chars per line
Compute: -W 481.89pt = 6.693in = 170.0mm = 17.0cm
Correct: -W 507.252pt = 7.045in = 178.947mm = 17.895cm
Compute: char height 803.15pt = 11.155in = 283.333mm = 28.333cm
Correct: char height 854.414pt = 11.867in = 301.418mm = 30.142cm
Backup: '~/banner.txt' --> '~/banner-2022.04.03-12.35.31.txt'
Rewrite: yawp --> '~/banner.txt'
    0 header lines, max 0 chars per line, 1 page
    1 body line, max 6 chars per line
    1 total line, max 6 chars per line
Print: '~/banner.txt' --> stdout
BANNER
Export: '~/banner.txt' --> '~/banner.pdf'
```

banner.txt is not altered because the -N argument, but the resulting
banner.pdf contains the string 'BANNER', a large character per page.

Line justification in text is controlled by -l argument:

  • "-l, --left-only": justify text lines at left only (default: at left
    and right)

After justification:

  • groups of consecutive empty lines between text paragraphs are reduced
    to a single empty line
  • groups of consecutive empty lines in contact with a picture are left
    unchanged

Now we can define some "best practice" to follow when writing the file.

Unindented paragraphs should be:

  • preceded by an empty line
  • started by an unindented line
  • continued by indented or unindented lines
  • ended by an empty line (better) or by a dot line

Indented paragraphs should be:

  • preceded by a line of any kind
  • initiated by a dot line
  • continued by indented or unindented lines
  • ended by an empty line or by another dot line

Pictures should be:

  • preceded by an empty line
  • initiated and continued by indented lines only
  • ended by an empty line (better) or by an unindented line or by a dot
    line

----
4. PYTHON FILES

Python files deserve a special treatment. If the textfile filename ends
with '.py' extension, then we suppose the file is a Python source, hence we
are interested to format docstrings and not Python code. So the formatting
function is alternatively turned on and off by switch lines. A "switch
line" is a line containing a "'''" string.

Note that yawp never formats switch lines, formatting takes place from the
line after the start switch line until the line before the next stop switch
line.

So your Python file must follow some simple rules:

- docstrings to be formatted must start and ended by "'''" and not
  '"""'
- long strings not to be formatted must start and end with '"""' and
  not "'''"
- a "'''" inside a string can be coded for instance as '\'\'\'', so
  containing line will not be identified as a switch line

An error in switch lines could format and destroy your Python code. A
preliminary check prints an error message and stops execution before file
formatting if the total number of switch lines is odd. This should
intercept 90% of errors, anyway after yawp processing check the result and
if needed go back to previous version by -U. An example:

```
$ cat pycode.py
''' Start switch line is not formatted.
This is a one-line unindented paragraph.

This is a first multiline unindented paragraph.
This is a first multiline unindented paragraph.
This is a first multiline unindented paragraph.

    This is a picture.
        This is a picture.
    This is a picture.

    . This is a first multi line indented paragraph.
This is a first multiline indented paragraph.
This is a first multiline indented paragraph.
        . This is another multi line indented paragraph.
This is another multiline indented paragraph.
This is another multiline indented paragraph.
''' # Stop switch line is not formatted.

def double(x): # Python code is not formatted.
 ''' Start switch line is not formatted.
This is another multiline unindented paragraph.
This is another multiline unindented paragraph.
This is another multiline unindented paragraph.
''' # Stop switch line is not formatted.
    return x + x # Python code is not formatted.
```

```
$ yawp -v -w 45 -p -P 0 pycode.py
Correct: -L 43.325pt = 0.602in = 15.284mm = 1.528cm
Correct: -R 40.57pt = 0.563in = 14.312mm = 1.431cm
Correct: -T 53.814pt = 0.747in = 18.985mm = 1.898cm
Correct: -B 15.991pt = 0.222in = 5.641mm = 0.564cm
Read: yawp <-- '~/pycode.py'
    0 header lines, max 0 chars per line, 1 page
    28 body lines, max 56 chars per line
    28 total lines, max 56 chars per line
Compute: -W 10.709pt = 0.149in = 3.778mm = 0.378cm
Correct: -W 11.272pt = 0.157in = 3.977mm = 0.398cm
Compute: char height 17.848pt = 0.248in = 6.296mm = 0.63cm
Correct: char height 18.987pt = 0.264in = 6.698mm = 0.67cm
Backup: '~/pycode.py' --> '~/pycode-2022.04.02-17.47.05.py'
Rewrite: yawp --> '~/pycode.py'
    0 header lines, max 0 chars per line, 1 page
    33 body lines, max 48 chars per line
    33 total lines, max 48 chars per line
Print: '~/pycode.py' --> stdout
#!/usr/bin/python3
''' Start switch line is not formatted.
This is a one-line unindented paragraph.

This   is   a   first   multiline   unindented
paragraph.   This   is   a   first   multiline
unindented   paragraph.   This   is   a  first
multiline unindented paragraph.

    This is a picture.
        This is a picture.
    This is a picture.

    • This  is  a  first  multi line indented
      paragraph.  This  is  a first multiline
      indented  paragraph.  This  is  a first
      multiline indented paragraph.
        • This is another multi line indented
          paragraph.   This   is   another
          multiline  indented paragraph. This
          is   another   multiline   indented
          paragraph.
''' # Stop switch line is not formatted.

def double(x): # Python code is not formatted.
 ''' Start switch line is not formatted.
This    is   another   multiline   unindented
paragraph.    This   is   another   multiline
unindented   paragraph.   This   is   another
multiline unindented paragraph.
''' # Stop switch line is not formatted.
    return x + x # Python code is not formatted.
```

5. CHAPTERS

The file can be partitioned in "chapters" by "chapter lines". A chapter can
be:

- the "nameless chapter", the first one, from first line until first
  chapter line, it could be the cover of the document
- a "numbered chapter", started by a "numbered chapter line", how many
  do you want
- the "contents chapter", started by a "contents line", no more than
  one
- the "index chapter", started by an "index line", no more than one

Numbered chapters, contents chapter and index chapter may appear in any
order.

5.1. NUMBERED CHAPTERS

Start of a "numbered chapter" is recognized by a numbered chapter line. A
line is a "numbered chapter line" if:

- is the first line or is preceded by an empty line
- is the last line or is followed by an empty line
- does not start with blank
- contains:
    - one or more "int-dot couples", each made by
        - one or more decimal digits, between '0' and '9'
        - a "decimal point" '.'
    - a "blank" ' '
    - a chapter title

The "level" of a numbered chapter line is the count of int-dot couples in
its prefix, examples:

- 12345. a level-1 numbered chapter line
- 1.345. a level-2 numbered chapter line
- 0.0.0. a level-3 numbered chapter line

numbered chapter lines must follow two quite obvious sequence rules:

- first numbered chapter line in file must be a level-1 numbered
  chapter line
- each other numbered chapter line can have a level between 1 and the
  level of the previous chapter line plus 1, but no more

Numbered chapter title is:

- shrunk and uppercased in text
- shrunk and titlecased when inserted:
    - into contents chapter
    - into page headers by '%c' (see '6. Paging')

To "titlecase" a string means to uppercase the first character of each word
in it and to lowercase all the others.

Numbers in input don't matter, yawp replaces them by the right ones, only
the level matters. So you are free to create destroy or swap chapters as
you like, they will be renumbered accordingly. Example:

```
$ cat chapters.txt
0. AAA aaa

32.33. BBB bbb

0.0. CCC ccc

0. DDD ddd

3.14. EEE eee

0.0.0. FFF fff

0.0. GGG ggg
```

```
$ yawp -v -p -P 0 chapters.txt
Correct: -L 43.325pt = 0.602in = 15.284mm = 1.528cm
Correct: -R 40.57pt = 0.563in = 14.312mm = 1.431cm
Correct: -T 53.814pt = 0.747in = 18.985mm = 1.898cm
Correct: -B 15.991pt = 0.222in = 5.641mm = 0.564cm
Read: yawp <-- '~/chapters.txt'
    0 header lines, max 0 chars per line, 1 page
    13 body lines, max 14 chars per line
    13 total lines, max 14 chars per line
Compute: -w 14
Compute: -W 34.421pt = 0.478in = 12.143mm = 1.214cm
Correct: -W 36.232pt = 0.503in = 12.782mm = 1.278cm
Compute: char height 57.368pt = 0.797in = 20.238mm = 2.024cm
Correct: char height 61.03pt = 0.848in = 21.53mm = 2.153cm
Backup: '~/chapters.txt' --> '~/chapters-2022.04.02-17.51.24.txt'
Rewrite: yawp --> '~/chapters.txt'
    0 header lines, max 0 chars per line, 1 page
    13 body lines, max 14 chars per line
    13 total lines, max 14 chars per line
Print: '~/chapters.txt' --> stdout
1. AAA AAA

1.1. BBB BBB

1.2. CCC CCC

2. DDD DDD

2.1. EEE EEE

2.1.1. FFF FFF

2.2. GGG GGG
```

5.2. CONTENTS CHAPTER

For "contents chapter" we mean the list of chapters, possibly with the number of the page they begin on. Namely, the contents chapter will list all numbered chapters and the index chapter, but will not list the nameless chapter or the contents chapter itself.

Contents line starting the contents chapter is defined by -c argument:

- "-c, --contents-title": title of contents chapter (default: 'contents'), example:

```
$ yawp -c 'list of chapters' ...
```

A line is a "contents line" if:

- is the first line or is preceded by an empty line

- is the last line or is followed by an empty line
- does not start with blank
- its shrunk and uppercased value is equal to the shrunk and uppercased
  value of the -c argument

If the file contains a contents line:

- lines until next chapter line (or until end of file) are supposed to
  be the old contents chapter and are deleted
- a new contents chapter is inserted instead

The file may or may not contain a contents chapter, but no more than one.
The contents chapter may appear everywhere, on top, on bottom or in the
middle of file, and after or before the index chapter.

Contents chapter title is:

- shrunk and uppercased in text
- shrunk and titlecased when inserted into page headers by '%c' (see
  '6. Paging')

BEWARE: an error in the next chapter line could erase a piece of your file,
so after yawp processing check the result and if needed go back to previous
version by -U.

Example:

```
$ cat contents.txt
Contents

    (any
    old
    content
    of
    the
    contents
    chapter
    will
    be
    deleted)

1. AAA AAA

1.1. BBB BBB

1.2. CCC CCC

2. DDD DDD

2.1. EEE EEE

2.1.1. FFF FFF

2.2. GGG GGG
```

```
$ yawp -v -p -P 0 contents.txt
Correct: -L 43.325pt = 0.602in = 15.284mm = 1.528cm
Correct: -R 40.57pt = 0.563in = 14.312mm = 1.431cm
Correct: -T 53.814pt = 0.747in = 18.985mm = 1.898cm
Correct: -B 15.991pt = 0.222in = 5.641mm = 0.564cm
Read: yawp <-- '~/contents.txt'
    0 header lines, max 0 chars per line, 1 page
    26 body lines, max 14 chars per line
    26 total lines, max 14 chars per line
Compute: -w 14
Compute: -W 34.421pt = 0.478in = 12.143mm = 1.214cm
Correct: -W 36.232pt = 0.503in = 12.782mm = 1.278cm
Compute: char height 57.368pt = 0.797in = 20.238mm = 2.024cm
Correct: char height 61.03pt = 0.848in = 21.53mm = 2.153cm
Backup: '~/contents.txt' --> '~/contents-2022.04.04-12.38.57.txt'
Rewrite: yawp --> '~/contents.txt'
    0 header lines, max 0 chars per line, 1 page
    23 body lines, max 20 chars per line
    23 total lines, max 20 chars per line
Print: '~/contents.txt' --> stdout
CONTENTS

    • 1.      Aaa Aaa
    • 1.1.    Bbb Bbb
    • 1.2.    Ccc Ccc
    • 2.      Ddd Ddd
    • 2.1.    Eee Eee
    • 2.1.1. Fff Fff
    • 2.2.    Ggg Ggg

1. AAA AAA

1.1. BBB BBB

1.2. CCC CCC

2. DDD DDD

2.1. EEE EEE

2.1.1. FFF FFF

2.2. GGG GGG
```

5.3. INDEX CHAPTER

For "index chapter" we mean an alphabetical list of subjects, possibly
showing which pages a subject is on. For "subject" we mean any string
preceded and followed by a "double quote" '"'. But double quotes preceded
or followed by a "single quote" "'" or by another double quote '"' are not
taken into account, in order to allow things like '"' in text.

Subject length is controlled by:

   • "-m, --max-subject": max subject length in index chapter (default:
     36)

this should intercept unpaired double quotes.

Index line starting the index chapter is defined by -i argument:

   • "-i, --index-title": title of index chapter (default: 'index'),
     example:

```
$ yawp -i 'list of subjects' ...
```

A line is an "index line" if:

- is the first line or is preceded by an empty line
- is the last line or is followed by an empty line
- does not start with blank
- its shrunk and uppercased value is equal to the shrunk and uppercased
  value of the -i argument

If the file contains an index line:

- lines until next chapter line (or until end of file) are supposed to
  be the old index chapter and are deleted
- a new index chapter is inserted instead

The file may or may not contain an index chapter, but no more than one. The
index chapter may appear everywhere, on top, on bottom or in the middle of
file, and after or before the contents chapter.

Index chapter title is:

- shrunk and uppercased in text
- shrunk and titlecased when inserted:
    - into contents chapter
    - into page headers by '%c' (see '6. Paging')

BEWARE: an error in the next chapter line could erase a piece of your file,
so after yawp processing check the result and if needed go back to previous
version by -U.

For an example, see the index chapter at end of this manual.

‾‾
6. PAGING

Normally the file is not splitted in pages by page headers. Insertion of
page headers is controlled by -f and -F arguments:

   • "-f, --form-feed" : insert a page header on full page

If -f is turned on, a page header is inserted into the file when:

   • current line doesn't fit into current page
   • current picture doesn't fit into current page

If a figure is too long to enter into the current page, a page eject is
forced. But if the figure is higher than one page, page eject doesn't take
place.

   • "-F, --form-feed-chap": insert a page header on full page and before
     contents index and level-one chapters

If -F is turned on (or both -f anf -F are), a page header is inserted into
the file when:

   • current line doesn't fit into current page
   • current picture doesn't fit into current page
   • current line is a contents line or an index line or a level-one
     numbered chapter line

So each page, except the first one, is prefixed by a page header. Each page
header is made up of two lines:

   • a line starting with a "form feed" '\f', containing data such as file
     name or page number
   • a dashed separation line of "macron" '‾' characters

So a line in the output can be considered as:

   • a "header line" if it's the first or the second line of a page header
   • or a "body line" otherwise

When the file is initially read, if -U and -N are off, all header lines are
eliminated. So, if you want to eliminate page headers from your file, just
run yawp without -f or -F.

If -f and -F are off and contents chapter and/or index chapter are
requested, contents chapter and/or index chapter will still appear, but
without page numbers.

Putting page headers in a Python file doesn't make sense. So if the file is
a Python file then -f and -F are automatically turned off.

The content of first header line is controlled by -e -E -o -O and -a
arguments:

   • "-e, --even-left": even page headers, left (default: '%n/%N')
   • "-E, --even-right": even page headers, right (default: '%f.%e
     %Y-%m-%d %H:%M:%S')
   • "-o, --odd-left": odd page headers left (default: '%c')
   • "-O, --odd-right":odd page headers, right (default: '%n/%N')

```
+-------+-------+
|-e   -E|-o   -O|
|-------+-------|
|       |       |
|  even |  odd  |
|  page |  page |
|       |       |
+-------+-------+
```

Each "percent variable" is evaluated as follows.

VAR	VALUE
'%P'	file "long path", with no ending '/'
'%p'	file "short path", with no ending '/'
'%f'	file name, with no extension
'%e'	file extension, with no separator '.'
'%Y'	current year, 4 digits
'%m'	current month, 2 digits
'%d'	current day, 2 digits
'%H'	current hour, 2 digits
'%M'	current minute, 2 digits
'%S'	current second, 2 digits
'%n'	current page number
'%N'	total number of pages
'%c'	current contents, index or level-one chapter
'%%'	'%'

No other percent variable is allowed.

If for instance the file is '/home/xxxx/yyy/zzz.www.txt' and the user is
'xxxx' then we get:

- '%P' --> '/home/xxxx/yyy'
- '%p' --> '~/yyy'
- '%f' --> 'zzz.www'
- '%e' --> 'txt'
- '%P/%f.%e' --> '/home/xxxx/yyy/zzz.www.txt'
- '%p/%f.%e' --> '~/yyy/zzz.www.txt'

If you don't need front-back printing, set:

- "-a, --all-pages-E-e": all page headers contain -E at left and -e at
  right

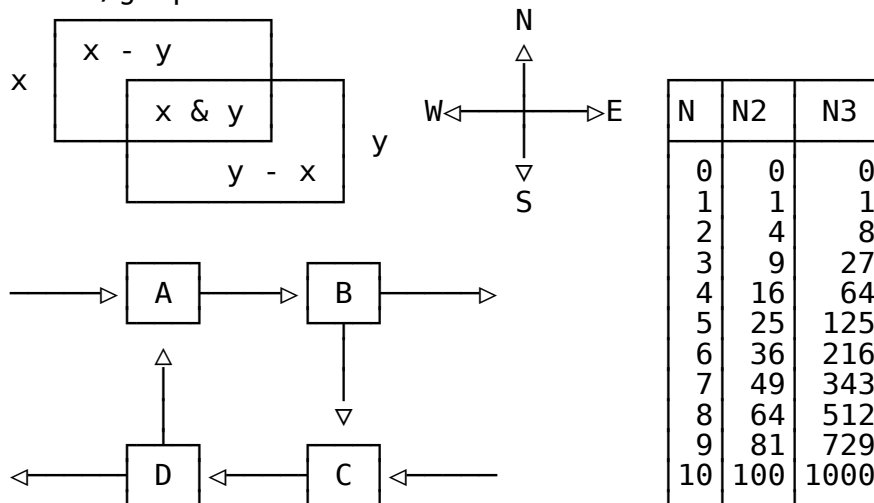and so you will get the same header format on even and odd pages. Note the
swap between -e and -E.

```
+-------+-------+
|-E   -e|-E   -e|
|-------+-------|
|       |       |
|  even |  odd  |
|  page |  page |
|       |       |
+-------+-------+
```

7. GRAPHICS

You can sketch pictures with boxes arrows and tables by two "draw characters":

- "back quote" '`' to draw horizontal and vertical "segments"
- "circumflex accent" '^' to mark an "arrowhead"

If you turn on the argument:

- "-g, --graphics": redraw '`'-segments and '^'-arrowheads

and -U is off then yawp will redraw the pictures by replacing the draw characters with suitable graphic ones, depending on the other four characters around (left right above and below). Standalone draw characters are not replaced.

This feature works:

- in "standard mode" (-N and -U off): in picture lines only, but not in text lines
- in "no-format mode" (-N on): in all lines of the file
- in "undo mode" (-U on): nowhere

An example:

```
$ cat graphics.txt
    ```````````                    N
   `  x - y      `                 ^
 x `  ``````````````               `        ``````````````
   `   `  x & y `   `      W^```````^E   `N `N2 ` N3 `
   ``````````````  `    y              ``````````````
   `   ` y - x `   `   ^   ` 0` 0`   0`
   ``````````````   S   ` 1` 1`   1`
                                          ` 2` 4`   8`
   `````   `````   ` 3` 9`  27`
 ``````^ A ``````^ B ```````^   ` 4` 16`  64`
     `````   `````   ` 5` 25` 125`
       ^   `                         ` 6` 36` 216`
       `         `   ` 7` 49` 343`
       `         ^                         ` 8` 64` 512`
     `````   `````   ` 9` 81` 729`
 ^``````` D `^````` C `^`````   `10`100`1000`
     ``````   ``````   ``````````````
```

```
$ yawp -v -g -p -P 0 graphics.txt
Correct: -L 43.325pt = 0.602in = 15.284mm = 1.528cm
Correct: -R 40.57pt = 0.563in = 14.312mm = 1.431cm
Correct: -T 53.814pt = 0.747in = 18.985mm = 1.898cm
Correct: -B 15.991pt = 0.222in = 5.641mm = 0.564cm
Read: yawp <-- '~/graphics.txt'
    0 header lines, max 0 chars per line, 1 page
    17 body lines, max 53 chars per line
    17 total lines, max 53 chars per line
Compute: -w 53
Compute: -W 9.092pt = 0.126in = 3.208mm = 0.321cm
Correct: -W 9.571pt = 0.133in = 3.376mm = 0.338cm
Compute: char height 15.154pt = 0.21in = 5.346mm = 0.535cm
Correct: char height 16.121pt = 0.224in = 5.687mm = 0.569cm
Backup: '~/graphics.txt' --> '~/graphics-2022.04.02-18.07.09.txt'
Rewrite: yawp --> '~/graphics.txt'
    0 header lines, max 0 chars per line, 1 page
    17 body lines, max 53 chars per line
    17 total lines, max 53 chars per line
Print: '~/graphics.txt' --> stdout
```

```
                                    N
       x - y                        △
  x                         W◁───────────▷E
       x & y                        ▽
                y                   S
       y - x
```

N	N2	N3
0	0	0
1	1	1
2	4	8
3	9	27
4	16	64
5	25	125
6	36	216
7	49	343
8	64	512
9	81	729
10	100	1000

```
      ──────▷ A ─▷ B ──────▷
              △      │
              │      ▽
      ◁────── D ◁─ C ◁──────
```

Argument -g works with -N too, so you can redraw picture files without the
need for a blank at beginning of the lines, and without yawp trying to
format the file. We can say that when -N is turned on, the whole file is a
single picture. As another example, we want to print a chessboard.

```
$ cat chessboard.txt
`````````````````
`  `X` `X` `X` `X`
`````````````````
`X` `X` `X` `X` `
`````````````````
`  `X` `X` `X` `X`
`````````````````
`X` `X` `X` `X` `
`````````````````
`  `X` `X` `X` `X`
`````````````````
`X` `X` `X` `X` `
`````````````````
`  `X` `X` `X` `X`
`````````````````
`X` `X` `X` `X` `
`````````````````
```

```
$ yawp -v -N -g -p -A 1 -L 1cm -T 1cm chessboard.txt
WARNING: -L 1cm < 2cm, you may get unexpected results
Correct: -L 13.737pt = 0.191in = 4.846mm = 0.485cm
WARNING: -R 1cm < 2cm, you may get unexpected results
Correct: -R 10.456pt = 0.145in = 3.689mm = 0.369cm
WARNING: -T 1cm < 2cm, you may get unexpected results
Correct: -T 23.699pt = 0.329in = 8.36mm = 0.836cm
WARNING: -B 1cm < 2cm, you may get unexpected results
Correct: -B 0.0pt = 0.0in = 0.0mm = 0.0cm
Read: yawp <-- '~/chessboard.txt'
    0 header lines, max 0 chars per line, 1 page
    17 body lines, max 17 chars per line
    17 total lines, max 17 chars per line
Compute: -w 17
Compute: -W 31.681pt = 0.44in = 11.176mm = 1.118cm
Correct: -W 33.349pt = 0.463in = 11.765mm = 1.176cm
Compute: char height 31.681pt = 0.44in = 11.176mm = 1.118cm
Correct: char height 33.704pt = 0.468in = 11.89mm = 1.189cm
Backup: '~/chessboard.txt' --> '~/chessboard-2022.04.02-18.09.55.txt'
Rewrite: yawp --> '~/chessboard.txt'
    0 header lines, max 0 chars per line, 1 page
    17 body lines, max 17 chars per line
    17 total lines, max 17 chars per line
Print: '~/chessboard.txt' --> stdout
```

```
┌─┬─┬─┬─┬─┬─┬─┬─┐
│ │X│ │X│ │X│ │X│
├─┼─┼─┼─┼─┼─┼─┼─┤
│X│ │X│ │X│ │X│ │
├─┼─┼─┼─┼─┼─┼─┼─┤
│ │X│ │X│ │X│ │X│
├─┼─┼─┼─┼─┼─┼─┼─┤
│X│ │X│ │X│ │X│ │
├─┼─┼─┼─┼─┼─┼─┼─┤
│ │X│ │X│ │X│ │X│
├─┼─┼─┼─┼─┼─┼─┼─┤
│X│ │X│ │X│ │X│ │
├─┼─┼─┼─┼─┼─┼─┼─┤
│ │X│ │X│ │X│ │X│
├─┼─┼─┼─┼─┼─┼─┼─┤
│X│ │X│ │X│ │X│ │
└─┴─┴─┴─┴─┴─┴─┴─┘
```

```
Export: '~/chessboard.txt' --> '~/chessboard.pdf'
```

Some remarks:

- '-N' prevents formatting the file as text, but the backup is still performed
- defaults '-w 0' and '-W 0' make the picture take up all available space between left and right margins
- '-A 1' transforms rectangles into squares in PDF file, so by printing it on an A4 sheet you get a square chessboard of about 18x18cm, 7x7in.

----

8. PDF EXPORTING

Normally the file is exported to a PDF file, which is then opened by
calling the system default PDF browser (as atril or evince). If you don't
want this to happen, set '-P 0'.

- "-P, --file-pdf": at end export and open PDF file ('0' = no PDF
  export, default: '%P/%f.pdf')

Each "percent variable" is evaluated as explained in '6. Paging', but '%n'
'%N' and '%c' are not applicable and not allowed. '%P' and '%p' are both
allowed but give the same result. The resulting file name must end with
'.pdf'.

Character size is defined by -W and -A arguments:

- "-W, --char-width": character width (pt/in/mm/cm, default: '0' =
  automatic)

Value is an integer or float literal followed by a lowercase suffix:

- 'pt' for points (1 inch = 72 points)
- 'in' for inches
- 'mm' for millimeters
- 'cm' for centimeters

so for instance these are all equivalent, giving a value of one inch:

- -W 72pt
- -W 1.0in
- -W 25.4mm
- -W 2.54cm

Only a zero value (as '0' or '0.0') can lack the suffix, because of course
0pt = 0in = 0mm = 0cm, but a zero value for -W means "automatic".

Both -W and -w (line width in chars per line, see '3. Justification') can
be zero hence automatic, namely:

- if -w is automatic and -W is not, -w is computed from -W
- if -W is automatic and -w is not, -W is computed from -w
- if both -w and -W are automatic,
  - -w is deduced from max line length in file (page headers are not
    considered)
  - -W is computed from -w

Given the character width by -W, the character height is derived from -A:

- "-A, --char-aspect": character aspect ratio = char width / char
  height ('1' = square grid, default: '3/5')

-A is a ratio, so can be:

- an integer or float literal (as '1' or '0.6')
- two integer or float literals, separated by a "slash" '/' (as '3/5'
  or '3/5.0')

Dimensions of the paper sheet, expressed in points inches centimeters or
millimeters (as explained for -W before) are defined by -S:

- "-S, --paper-size": portrait paper size (width x height, pt/in/mm/cm,
  default: 'A4' = '210x297mm')

Format is portrait, in other words the width can't be greater than the
height.

Values can be names too, see the following table of allowed names. These names are case-insensitive ('A4' and 'a4' are equivalent), while the 'x' and the suffix in the width x height value must be lowercase.

NAME	VALUE
HALF LETTER	5.5x8.5in
LETTER	8.5x11.0in
LEGAL	8.5x14.0in
JUNIOR LEGAL	5.0x8.0in
LEDGER	11.0x17.0in
TABLOID	11.0x17.0in
A0	841x1189mm
A1	594x841mm
A2	420x594mm
A3	297x420mm
A4	210x297mm
A5	148x210mm
A6	105x148mm
A7	74x105mm
A8	52x74mm
A9	37x52mm
A10	26x37mm
B0	1000x1414mm
B1	707x1000mm
B1+	720x1020mm
B2	500x707mm
B2+	520x720mm
B3	353x500mm
B4	250x353mm
B5	176x250mm
B6	125x176mm
B7	88x125mm
B8	62x88mm
B9	44x62mm
B10	31x44mm

Paper width and height are exchanged with each other by -Z:

  • "-Z, --landscape": turn page by 90 degrees (default: portrait)

Print quality is defined by -Q, but reducing the print quality doesn't make much sense, it only slightly reduces the length of the exported PDF file:

  • "-Q, --print-quality": print quality ('0' '1' or '2', default: '2')

Unprintable margins around the paper sheet are controlled by:

  • "-L, --left-margin": left margin (pt/in/mm/cm, default: '2cm')
  • "-R, --right-margin": right margin (pt/in/mm/cm, default: '-L')
  • "-T, --top-margin": top margin (pt/in/mm/cm, default: '2cm')
  • "-B, --bottom-margin": bottom margin (pt/in/mm/cm, default: '-T')

Margins are expressed in points inches centimeters or millimeters as explained above for -W.

If -R has the default value '-L' then it's forced to the value of -L.

If -B has the default value '-T' then it's forced to the value of -T.

BEWARE: margins of less than 2cm are allowed but not guaranteed, and may give unexpected results.

9. CORRECTIONS

Export of the PDF file is performed by CUPS via the "lp" Unix command.
Geometry is controlled by yawp passing to lp various arguments:

- -o cpi=N (number of characters per inch, default=10)
- -o lpi=N (number of lines per inch, default=6)
- -o page-left=N (left page margin, value in points)
- -o page-right=N (right page margin, value in points)
- -o page-top=N (top page margin, value in points)
- -o page-bottom=N (bottom page margin, value in points)

These options are undocumented in the lp man page, but you can find them
for instance in:

    https://www.computerhope.com/Unix/ulp.htm

Unfortunately, by printing the PDF file on paper the above options are not
respected, but are affected by not negligible errors. With "lpr" it's the
same.

A hardwired mechanism in yawp tries to correct such errors. This device has
been tuned for the default paper size 'A4' = '210x297mm', both portrait and
landscape, on a HP Officejet 2620 printer. With another format or another
printer you may get different results. In a future release this may be made
parametric thru a configuration file.

If -v is on, such corrections are signaled by information messages starting
with 'Correct:'.

10. AFTERWORDS

10.1. DEBUGGING

This is a beta release, so it contains some functionality aimed to the debug and not to the end user. They may disappear in a future release.

If you turn on:

  • "-s, --echo-shell": debug, display invoked Unix commands

then Unix commands invoked during execution are displayed on stderr, along with their output.

If you turn on:

  • "-k, --calibration": debug, don't correct character size and page margins

then no correction on character size and page margins is performed. This is useful in setting up the correction device seen in previous chapter.

10.2. CREDITS

A somewhat analogous (but very different) program is the Unix command "fmt". We are not aware of any other programs of this kind. For details, type:

  | $ man fmt

Yawp has been developed under Xubuntu 21.10:

    http://www.xubuntu.org

by Python 3.9.7 and IDLE, the Python's integrated editor:

    https://www.python.org

and published on Pypi by flit 3.7.1:

    https://pypi.org/project/flit

10.3. HISTORY

  • version 0.5.4
    • just some other fix to the Yawp Manual
    • beta release

  • version 0.5.3
    • just a fix to the installation instructions
    • beta release

  • version 0.5.2
    • correction of page margins, new algorithm
    • page margins, no more bounded by 2cm..8cm interval (but below 2cm you may get unexpected results)
    • -g argument, now works with -N too (but not with -U)
    • information messages about corrected values, added
    • index subjects, taken not only from text but from pictures too
    • debugging argument "-b, --dump-buffer", eliminated
    • Yawp Manual, updated
    • beta release

  • version 0.5.1
    • completely redefined and rewritten
    • beta release

- version 0.4.2
  - obsolete and deprecated

- version 0.4.1
  - first version on pypi.org
  - obsolete and deprecated

10.4. ACRONYMS

ACRONYM	MEANING
CUPS	Common Unix Printing System
HP	Hewlett-Packard
IDLE	Integrated Development and Learning Environment
PDF	Page Description Format
PyPI	Python Package Index
WYSIWYG	What You See Is What You Get
YAWP	Yet Another Word Processor

10.5. CHARACTERS

CHARACTER NAME	CHAR	HEX	DEC
back quote	'`'	0060	96
black small circle	'•'	2022	8226
blank	' '	0020	32
box drawings light down and horizontal	'┬'	252c	9516
box drawings light down and left	'┐'	2510	9488
box drawings light down and right	'┌'	250c	9484
box drawings light horizontal	'─'	2500	9472
box drawings light up and horizontal	'┴'	2534	9524
box drawings light up and left	'┘'	2518	9496
box drawings light up and right	'└'	2514	9492
box drawings light vertical	'│'	2502	9474
box drawings light vertical and horizontal	'┼'	253c	9532
box drawings light vertical and left	'┤'	2524	9508
box drawings light vertical and right	'├'	251c	9500
carriage return	'\r'	000d	13
circumflex accent	'^'	005e	94
decimal point	'.'	002e	46
digit nine	'9'	0039	57
digit zero	'0'	0030	48
double quote	'"'	0022	34
form feed	'\f'	000c	12
horizontal tab	'\t'	000b	11
line feed	'\n'	000a	10
macron	'¯'	00af	175
percent sign	'%'	0025	37
single quote	"'"	0027	39
slash	'/'	002f	47
white down-pointing triangle	'▽'	25bd	9661
white left-pointing triangle	'◁'	25c1	9665
white right-pointing triangle	'▷'	25b7	9655
white up-pointing triangle	'△'	25b3	9651

11. ARGUMENTS

General arguments:

- "-h, --help": show a short help message and exit
- "-H, --manual": open this yawp-generated "yawp Manual" in PDF format and exit
- "-V, --version": show program's version number and exit
- "-v, --verbose": display information messages on stderr
- "-N, --no-format": don't format the file
- "-U, --undo": restore the file from its previous version
- "-g, --graphics": redraw '`'-segments and '^'-arrowheads
- "-p, --print-file": at end print file on stdout

Formatting arguments:

- "-w, --chars-per-line": line width in chars per line (default: '0' = automatic)
- "-l, --left-only": justify text lines at left only (default: at left and right)
- "-c, --contents-title": title of contents chapter (default: 'contents')
- "-i, --index-title": title of index chapter (default: 'index')
- "-m, --max-subject": max index subject length (default: '36')

Paging arguments:

- "-f, --form-feed": insert a page header on full page
- "-F, --form-feed-chap": insert a page header on full page and before contents index and level-one chapters
- "-e, --even-left": even page headers, left (default: '%n/%N')
- "-E, --even-right": even page headers, right (default: '%f.%e %Y-%m-%d %H:%M:%S')
- "-o, --odd-left": odd page headers, left (default: '%c')
- "-O, --odd-right": odd page headers, right (default: '%n/%N')
- "-a, --all-pages-E-e": all page headers contain -E at left and -e at right

PDF arguments:

- "-P, --file-pdf": file name of exported PDF file ('0' = no PDF export, default: '%P/%f.pdf')
- "-W, --char-width": character width (pt/in/mm/cm, default: '0' = automatic)
- "-A, --char-aspect": character aspect ratio = char width / char height ('1' = square grid, default: '3/5')
- "-S, --paper-size": portrait paper size (width x height, pt/in/mm/cm, default: 'A4' = '210x297mm'
- "-Z, --landscape": turn page by 90 degrees (default: portrait)
- "-Q, --print-quality": print quality ('0' '1' or '2', default: '2')
- "-L, --left-margin": left margin (pt/in/mm/cm, default: '2cm')
- "-R, --right-margin": right margin (pt/in/mm/cm, default: '-L')
- "-T, --top-margin": top margin (pt/in/mm/cm, default: '2cm')
- "-B, --bottom-margin": bottom margin (pt/in/mm/cm, default: '-T')

Debugging arguments:

- "-s, --echo-shell": debug, display invoked Unix commands
- "-k, --calibration": debug, don't correct character size and page margins

Positional argument:

- "file": text file to be processed

No numeric arguments can have a negative value.

INDEX