

NAME**ptx** — permute index**SYNOPSIS**

```

ptx [-f] [-r|-A] [-R] [-t|-w width] [-o only-words] [-i ignore-words]
      [-b separators] [-F trunc-flag] [-g field-gap] [file]...
ptx [-O|-T] [-f] [-r|-A] [-t|-w width] [-o only-words] [-i ignore-words]
      [-b separators] [-F trunc-flag] [-g field-gap] [-M macro] [file]...
ptx [-G] [-f] [-r|-A] [-t|-w width] [-o only-words] [-i ignore-words]
      [-b separators] [-F trunc-flag] [-g field-gap] [-M macro] [file [into]]

```

DESCRIPTION

Tokenises each line in *files* (standard input stream if "-", the default), duplicates them such that each copy has a different token (keyword) rotated to the front, sorts them, then lays them out for use in a permuted index to the standard output stream. As an example:

```

ptx - permute index
is tokenised into
ptx - permute index
- permute index ptx
permute index ptx -
index ptx - permute
and sorted according to the locale's collation sequence as
- permute index ptx
index ptx - permute
permute index ptx -
ptx - permute index
then laid out as

```

```

      ptx - permute index
ptx - permute index
      ptx - permute index
      ptx - permute index

```

If an input line doesn't fit on one **78**-column-wide-by-default line, it's truncated (noted by *trunc-flag* ("/" by default)) and, if possible, overflow is continued on the opposite side; for example, after limiting to **25** columns:

```

      ptx - permute /      ← right bit truncated
- permute index ptx      ← left bit continued on the right
      ptx - permute /      ← right bit truncated
permute/ ptx -           ← left bit truncated

```

a minimum *field-gap* is maintained (default **2**).

Abstractly, there are thus four sections: overflow-from-right, bit before keyword, keyword and bit after keyword, overflow-from-left (at most one overflow is used at any time), read, in order, as

```

                                overflow-from-left
      bit before keyword  keyword and bit after keyword
overflow-from-right
but squished into one line. The default output mode produces this format directly, but -OG instead yield
troff(1) code:
      .xx "overflow-from-right" "bit before keyword" "keyword and bit after keyword" "overflow-from-left"
(xx is the default macro) and -T yields tex(1) code:
      \xx{overflow-from-right}{bit before keyword}{keyword and bit after keyword}{overflow-from-left}

```

With **-r**, the first token on each line is removed, and taken to be a chapter reference; these are output on a completely separate column to the left (right with **-R**) of the index, or, with **-OTG**, as the fifth argument. For example, with **-r** if the input were instead

```

ptx(1) ptx - permute index
it would be laid out as
ptx(1)          ptx - permute index
ptx(1) ptx - permute index
ptx(1)          ptx - permute index
ptx(1)          ptx - permute index
or

```

```

      ptx      - permute index      ptx(1)
ptx - permute  index                ptx(1)
      ptx -    permute index        ptx(1)
      ptx -    permute index        ptx(1)

```

-A also adds a chapter reference, but in the form "*file:line-number*" (though *file* is empty if reading the standard input stream). As a special bonus, without **-OTG** and **-R**, ":" is appended.

OPTIONS

-O, --format=roff Produce troff(1) macros. "''" and '\ ' are escaped.

-G, --traditional **-O** but also take the second argument as *to*, and behave as-if invoked as *> to*.

-O, --format=tex Produce tex(1) macros. '\', '{', '}', '#', '\$', '%', and '&' are escaped.

-f, --ignore-case Sort without regard to case.

-r, --references Delete the first token in each line and produce it as the fifth macro argument or as another column.

-A, --auto-reference Produce "*file:line-number*" it as the fifth macro argument or as another column.

-w, --width=width Lay out text to fit a *width*-column wide page. Defaults to **72**.

-t, --typeset-mode **-w 100**

-o, --only-file=only-words Tokenise file *only-words*. When splitting an input line, only select words that match one of the resulting tokens as the keywords. A line with no keywords is removed. Always matched case-insensitively.

-i, --ignore-file=ignore-words Tokenise file *ignore-words*. When splitting an input line, do not pick words that that match one of the resulting tokens as the keywords. Traditionally this is the *eign* file from the troff distribution, containing words like "a", "I", "the", "and", &c. Always matched case-insensitively.

-b, --break-file=separators When tokenising input files, break on any of the characters in *separators*, in addition to the space, the tab, and the new-line. Always matched case-sensitively.

-F, --flag-truncation=truncatedg If a line is truncated, the cut-off bit is replaced with *trunc-flag*. Defaults to *"I"*.

-g, --gap-size=field-gap Allocate *field-gap* columns between each of the four columns. Defaults to **2** (though note that this is likely to be like **3** in practice, as any spaces between tokens are also reproduced).

-R, --right-side-refs Without **-OTG**, put the chapter references from **-rA** on the right.

-M, --macro-name=macro With **-OTG**, call *macro* with four (by default) or five (**-rA**) arguments. Defaults to **xx**.

EXIT STATUS

1 if a *file*, *only-words*, *ignore-words*, or *separators* couldn't be opened or read, or if *into* couldn't be created or written.

SEE ALSO

wcwidth(3)

STANDARDS

Compatible with the GNU system (sans its **-SW** flags). With **-G**, largely compatible with Version 7 AT&T UNIX. Some implementations automatically pre-load a fixed list of *ignore-words*; you should find one for English in a file called `eign` somewhere in your `troff(1)` distribution.

HISTORY

The first edition of the UNIX Programmer's Manual has a permuted **INDEX**, but doesn't describe a **ptx**.

The second edition, in section VI (then "User-maintained programs"), describes a `ptx`:

```
NAME      ptx  --  permuted index
SYNOPSIS  ptx1 input temp1
          sort temp1 temp2
          ptx2 temp2 output
```

the list of *ignore-words* is fixed at "a", "and", "as", "is", "for", "of", "on", "or", "the", "to", and "up".

Version 3 AT&T UNIX fuses the **SYNOPSIS** into

```
ptx input output
```

Version 4 AT&T UNIX sees a **SYNOPSIS** of

```
ptx [ -t ] input [ output ]
```

where **-t** "causes `ptx` to prepare its output for the phototypesetter". "an" is also ignored now.

Version 6 AT&T UNIX removes `ptx` from the manual but still carries it. This being the first userland-source-available UNIX we know that: lines are tokenised are broken down at spaces and tabs. These are all welded together into one space. Keywords starting with the left parenthesis ('(') are explicitly sorted last. The tilde ('~') is used as in-band signalling to separate the front- and back-end of each line for sorting, it featuring in the input disrupts reconstruction. The output columns are separated by a two-column gap, overflow to the left (or right) is indicated with "...", and spills only occur from the left and without additional spacing. **-t** changes the target width from **72** to **100**, as present-day, The default output mode is "visual", as present-day, for `nroff(1)` use, and **-t** also enables

```
.xx "bit before keyword" "keyword and bit after keywordoverflow-from-left"
```

output. Most of these semantics are unlikely to have meaningfully changed so far.

Version 7 AT&T UNIX `ptx(1)` is new and synthesised as

```
ptx [ option ] ... [ input [ output ] ]
```

and is effectively present-day, semantically *and* seeing most of the modern usage: *option* is any of **-frtwoibg**. The only noted output format is

```
.xx "tail" "before keyword" "after keyword" "head"
```

with the modern layout and "/" *trunc-flag* and **-t** is "Prepare the output for the phototypesetter; the default line length is 100 characters." but all it does is **-w 100** if **-w** wasn't already given. "''" is escaped. The default *field-gap* is **3**.

-io override each other. *only-words* and *ignore-words* take one word per line. They are always matched case-insensitively. If neither is given, `/usr/lib/eign` is used as *ignore-words*.

AT&T System III UNIX finally documents the tilde ('~') being poisonous.

AT&T System V Release 1 UNIX promotes `mptx(5)` ("the macro package for formatting a permuted index"), which is new, and produces a result similar to this implementation's default output with **-R**. The gap is 1em (2 characters), which doesn't match the unchanged default *field-gap*.

ptx disappears from AT&T System V Release 3 UNIX.

The BSD ships Version 7 AT&T UNIX **ptx**.