NAME

unroff-html-man – back-end to translate manual pages to HTML 2.0

SYNOPSIS

unroff [**–fhtml**] [**–man**] [*file* | *option...*]

OVERVIEW

When called with the **–fhtml** and **–man** options, the troff translator *unroff* loads the back-end for converting UNIX manual pages to the Hypertext Markup Language (HTML) version 2.0.

Please read **unroff**(1) first for an overview of the Scheme-based, programmable troff translator and for a description of the generic options that exist in addition to **–f** and **–m**. The translation of basic troff requests, special characters, escape sequences, etc. as well as the HTML-specific options are described in **unroff-html**(1). For information about extending and programming *unroff* also refer to the *Unroff Programmer's Manual*.

OPTIONS

The **–man** extension provides one new keyword/value option in addition to those listed in **unroff**(1) and **unroff-html**(1):

do-signature (boolean)

If set to 1, a signature is appended to each output file. The signature is composed of a horizontal rule and a one-line message consisting of version information and date and time. The default value of this option is 1.

DESCRIPTION

unroff reads and parses its input files (each containing a UNIX manual page); the HTML output is written to a separate output file for each input file. The name of an output file is obtained by appending the suffix ".html" to the name of the corresponding input file. Any **document** option is ignored if input files are named in the command line. As usual, the special file name '**–**' can be used to interpolate standard input.

If no file name is given in the command line, a manual page is read from standard input and sent to standard output, unless the **document** option is given, in which case the HTML output is written to the specified file (with ".html" appended). Example: this call to *unroff* translates two manual pages and creates two corresponding output files, **cc.1.html** and **send.2.html**:

```
unroff –fhtml –man /usr/man/man1/cc.1 /usr/man/man2/send.2
```

The following **–man** macros are recognized and translated (in addition to any user-defined macros):

```
.TH   .SH   .SS   .I    .B    .SB   .SM
.BI   .BR   .IB   .IR   .RB   .RI   .TP
.IP   .HP   .RS   .RE   .LP   .PP   .P
```

In addition, the following Sun-specific macros are silently ignored (.TX generates an informational message containing its argument):

```
.TX   .IX   .DT   .PD   .UC
```

The following predefined troff strings are recognized (*S expands to the empty string):

```
\*R   \*S   \*(lq   \*(rq
```

The title of each HTML document generated is obtained by calling the primitive *substitute* (as explained in the Programmer's Manual) with the value of the option **title** and the first and second arguments passed to the initial call to **.TH**. Thus, the specifiers "%1%" and "%2%" can be used in the option to interpolate the command (or whatever is documented in the manual page) and the section number. If **title** has not been specified, the string "Manual page for %1%(%2%)" is taken. As generating the HTML title element is deferred until the call to **.TH**, any macros or other troff requests that produce output must not be used before the initial **.TH**.

HTML header elements <h2> and <h3> are created for **.SH** and **.SS** requests. The markup created for the initial NAME section differs in that the contents of the section (usually a single line) is itself placed inside a header element.

The font switching macros are based on changes to the fonts ‘R’, ‘I’, and ‘B’, as explained under FONTS in **unroff-html**(1). Of course, this fails if the fonts (which are mounted on startup) are unmounted by explicit **.fp** requests. As HTML is lacking the concept of text size, the macro **.SB** is just an alias for **.B**, and **.SM** simply echoes its arguments.

The translation rules for **.TP** and **.IP** employ a heuristic to determine whether to generate a definition list or an unordered list: if the first in a sequence of tagged/indented paragraph macros is called with a tag consisting of the special character \(bu, a definition list is begun, otherwise an unordered list. Subsequent invocations cause the list style to change if appropriate. Use of tagged paragraphs inside non-filled (pre-formatted) text violates the HTML definition and should be avoided. A warning message is printed in this and other questionable situations.

As hanging tags cannot be realized with HTML 2.0, a kludge is used for the **.HP** (hanging paragraph) macro: the macro starts a definition list (as does the ordinary **.TP** macro), and everything up to the next request that causes a break is placed inside the definition tag. This method obviously fails if no break occurs in subsequent lines, but it works for the common, idiomatic use of hanging paragraphs in manual pages.

SEE ALSO

unroff(1), **unroff-html**(1), **troff**(1), **man**(5 or 7).

Unroff Programmer’s Manual.

http://www.informatik.uni-bremen.de/˜net/unroff

Berners-Lee, Connolly, et al., HyperText Markup Language Specification—2.0, Internet Draft, Internet Engineering Task Force.