NAME

html-norm − normalize HTML tags

SYNOPSIS

html-norm [*sgmlnorm-options*] [*in-html-file* or <*in-html-file*] >*out-html-file*

DESCRIPTION

html-norm is a convenient interface to **sgmlnorm**(1), a tag normalizer for SGML files. It conceals the details of providing the HTML grammar files needed by the normalizer.

It is a good idea to run **html-norm** on any HTML files that you plan to filter with **html-pretty**(1) to ensure correct indenting.

OPTIONS

All command-line options are passed on directly to **sgmlnorm**(1); none are interpreted by **html-norm** itself. For convenience, the **sgmlnorm**(1) option descriptions are repeated here. **html-norm** supplies the **−c** *catalog-file* and **−k** *c* options.

−c *catalog-file*
: Provide the full path name of the SGML catalog file, which is needed to relate the document type name in the initial SGML
<!DOCTYPE document-type-name PUBLIC "...">
declaration to a grammar file that defines the syntax of the SGML tags in the remainder of the file.

  This option is always required, because the current version of **sgmlnorm** has no built-in default catalog location.

  At some UNIX sites, a suitable option will be **−c** */usr/local/lib/html-check/lib/catalog*

−d *file-directory-name*
: Add a file directory name to the default search path used to locate files specified in system identifiers in
<!DOCTYPE document-type-name SYSTEM "filename">
declarations. Multiple **−d** options are allowed.

−e
: Describe open entities in error messages. Error messages always include the position of the most recently opened external entity.

−k cm
: This option requires either **c**, or **m**, or both, in either order. The letter **c** requests output of any embedded SGML comments; otherwise, comments are normally discarded. The letter **m** requests output of marked sections. [It is unclear what the **m** option is supposed to do.]

−w *warning_type*
: Control warnings and errors. Multiple **−w** options are allowed. The following values of *warning_type* enable warnings:

  mixed
  : Warn about mixed content models that do not allow #pcdata anywhere.

  sgmldecl
  : Warn about various dubious constructions in the SGML declaration.

  should
  : Warn about various recommendations made in ISO 8879 that the document does not comply with. (Recommendations are expressed with ‘‘should’’, as distinct from requirements which are usually expressed with ‘‘shall’’.)

  default
  : Warn about defaulted references.

  duplicate
  : Warn about duplicate entity declarations.

  undefined
  : Warn about undefined elements: elements used in the DTD but not defined.

unclosed Warn about unclosed start and end-tags.

empty Warn about empty start and end-tags.

net Warn about net-enabling start-tags and null end-tags.

min-tag Warn about minimized start and end-tags. Equivalent to combination of **unclosed**, **empty** and **net** warnings.

unused-map Warn about unused short reference maps: maps that are declared with a short reference mapping declaration but never used in a short reference use declaration in the DTD.

unused-param Warn about parameter entities that are defined but not used in a DTD.

all Warn about conditions that should usually be avoided (in the opinion of the author). Equivalent to: **mixed**, **should**, **default**, **undefined**, **sgmldecl**, **unused-map**, **unused-param**, **empty** and **unclosed**.

A warning can be disabled by using its name prefixed with **no-**. Thus **-wall -wno-duplicate** will enable all warnings except those about duplicate entity declarations.

The following values for *warning_type* disable errors:

no-idref
Do not give an error for an ID reference value which no element has as its ID. The effect will be as if each attribute declared as an ID reference value had been declared as a name.

no-significant
Do not give an error when a character that is not a significant character in the reference concrete syntax occurs in a literal in the SGML declaration. This may be useful in conjunction with certain buggy test suites.

SEE ALSO

html-check(1), **html-ncheck**(1), **html-pretty**(1), **html-spam**(1), **htmlchek**(1), **nsgmls**(1), **sgmlnorm**(1), **sgmls**(1), **spam**(1), **spent**(1),

AUTHOR

Nelson H. F. Beebe, Ph.D.
Center for Scientific Computing
Department of Mathematics
University of Utah
Salt Lake City, UT 84112
Tel: +1 801 581 5254
FAX: +1 801 581 4148
Email: <beebe@math.utah.edu>
WWW URL: http://www.math.utah.edu/~beebe