Hyphenation Exception Log

Barbara Beeton

This is the periodic update of the list of words that \TeX fails to hyphenate properly. The full list last appeared in \textit{TUGboat} 16, no. 1, starting on page 12, with updates in \textit{TUGboat} 22, no. 1/2, pages 31–32, and \textit{TUGboat} 23, no. 3/4, pages 247–248. This installment contains only exceptions identified since the last update.

In the list below, the first column gives results from \TeX’s \texttt{\showhyphens{...}}; entries in the second column are suitable for inclusion in a \texttt{\hyphenation{...}} list.

In most instances, inflected forms are not shown for nouns and verbs; note that all forms must be specified in a \texttt{\hyphenation{...}} list if they occur in your document. See the section below, “Converting this list into a list of hyphenation exceptions”.

Like the full list, this update has been subdivided into two parts: English words, and names and non-English words (including transliterations from Cyrillic and other non-Latin scripts) that occur in English texts.

Thanks to all who have submitted entries to the list. Since some suggestions have demonstrated a lack of familiarity with the rules of the hyphenation algorithm, here is a short reminder of the relevant idiosyncrasies. Hyphens will not be inserted before the number of letters specified by \texttt{\lefthyphenmin}, nor after the number of letters specified by \texttt{\righthyphenmin}. For U.S. English, \texttt{\lefthyphenmin=2} and \texttt{\righthyphenmin=3}; thus no word shorter than five letters will be hyphenated. (For the details, see \textit{The \TeXbook}, page 454.) This particular rule is violated in some of the words listed; however, if a word is hyphenated correctly by \TeX except for “missing” hyphens at the beginning or end, it has not been included here.
Some other permissible hyphens have been omitted for reasons of style or clarity. While this is at least partly a matter of personal taste, an author should think of the reader when deciding whether or not to permit just one more break-point in some obscure or confusing word. There really are times when a bit of rewriting is preferable.

One other warning: Some words can be more than one part of speech, depending on context, and have different hyphenations; for example, ‘analyses’ can be either a verb or a plural noun. If such a word appears in this list, hyphens are shown only for the portions of the word that would be hyphenated the same regardless of usage. These words are marked with a ‘*’; additional hyphenation points, if needed in your document, should be inserted with discretionary hyphens.

The reference used to check these hyphenations is Webster’s Third New International Dictionary, Unabridged.

Hyphenation for languages other than English

Patterns now exist for many languages other than English, including languages using accented alphabets. CTAN holds an extensive collection of patterns in tex-archive/language/hyphenation and its subdirectories.

Converting this list into a list of hyphenation exceptions

Werner Lemberg has created a script that will convert this article into a real \hyphenation block that can be incorporated into a document either directly or by inputting a file. Many inflected forms will be included automatically, some evident in the printed version, but many included silently.

The script, hyphenex.sh, runs under Unix. It is posted on CTAN, along with its output for the current set of exceptions, in tex-archive/info/digests/tugboat/

Names and non-English words used in English text

AT-Pase
ATP-ase
Delaware
Del-a-ware
Du-ane
Duane
FreeBSD
Free-BSD
Hadamard
Ha-da-mard
Har-alam-bous
Har-a-lam-bous
Jack-owski
Jac-kow-ski
Ma-cOS
Mac-OS
Math-SciNet
Math-Sci-Net
NetBSD
Net-BSD
OpenBSD
Open-BSD
OpenOf-fice
Open-Office
PfaEdit
Pfa-Edit
Richard
Ric-hard
Southall
South-all
Tomaszewski
To-ma-szew-ski
VMware
VM-ware
Werner
Wer-ner
WinEdt
Win-Edt

The List — English words

crancase
crossover
cuf-flink(s)
cus-tomiz-able
cus-tomize(s,d)
dichloromethane
ethane
flu-o-ro-car-bon
geother-mal
grandu-n cle
hy-drother-mal
key-note
manslaugh-ter
methane
ni-tromethane
of-fline
of-fload(s,ed)
palette
pipeline-ing
prewrap(ped)
pro-grammable
promi-nent
promis-cu-ous
promis-sory
promise
prowess
qu-a-sitriv-ial
rephrase(s,d)
rewrap(ped)
subex-pres-sion
taffrail
tri-ethy-lamine
vi-sual(ly)

• Barbara Beeton

bnb@ams.org