Hyphenation exception log

Barbara Beeton

This is a brief update of the list of words that TeX fails to hyphenate properly for U.S. English. The full list last appeared in *TUGboat* 16:1, starting on page 12, with periodic updates, most recently in *TUGboat* 39:1, p. 7.

A cumulative list appears in the CTAN collection, at `ctan.org/pkg/hyphenex`, or in TeX Live, in the file `tb0hyf` (`.tex` or `.pdf`). We are dispensing here with the details of how the list is organized; that information can be found in the previous update or on CTAN.

The heavy concentration of names in this update is due to the (un?)timely appearance of a 700-page book on a historical topic in the production queue at the AMS. Names often do not follow the pattern of ordinary English words, and are therefore not handled gracefully by the existing patterns. Worse still, they often appear in the first sentence of a paragraph, introducing the subject; and, once introduced, they are likely to occur again in the same work. Hence the desirability that they be dealt with once, in a local `\hyphenation` list.

British hyphenation

The patterns for British hyphenation — quite different from the U.S. patterns — were developed from a word list provided by the Oxford University Press based on their 1986 *Minidictionary of Spelling and Word Division*. As of 2014, the *New Oxford Spelling Dictionary* shows that many points of word division have changed. (The changes largely appear to be converging with U.S. practice.) The UK TeX Users Group has approached OUP to request access to the new word list, in order to update the patterns. There is apparently progress in this direction, at least in principle, although nothing definitive is known yet. Should the negotiations be successful, it will be reported here as soon as feasible.

The list — English words

`al-manac`	al-ma-nac
`al-manack`	al-ma-nack
`awarded`	awar-ded
`bankroll`	bank-roll
`courage`	cour-age
`coura-geous(ly)`	cou-ra-geous(-ly)
`dataflow`	data-flow
`dat-a-point`	data-point
`dataset`	data-set
`datatype`	data-type
`drosophila`	dro-soph-i-la
`ephemera`	ephem-era
`ephemeris(ides)`	ephem-eris(-i-des)
`equated`	equa-ted
`equidi-men-sional`	equi-di-men-sional
`method-ol-ogy(ies)`	meth-od-o-lo-gy(ies)
`method-olog-i-cal(ly)`	meth-od-o-logical(ly)
`nonar-chimedean`	non-ar-chi-me-dean
`philoso-pher`	phi-los-o-pher
`philoso-phies`	phi-los-o-phies
`plurisub-har-monic`	pluri-sub-har-monic
`polynya`	po-lyn-ya
`pol-y-semy`	poly-se-my
`potable`	po-ta-ble
`premise`	prem-ise
`prepo-si-tion`	prep-o-si-tion
`prepo-si-tional`	prep-o-si-tional
`semistable`	semi-sta-ble
`sto-plist`	stop-list
`transat-lantic`	trans-at-lan-tic
`transept`	tran-sept
`transpa-cific`	trans-pacific
`trapez-ium`	tra-pe-zium
`trape-zoid`	trap-e-zoid
`trape-zoidal`	trap-e-zoi-dal
`trigonom-e-try`	trig-o-nom-e-try
`trigono-met-ric`	trig-o-no-met-ric

Names and non-English words used in English text

`Alexan-dria`	Alex-an-dria
`An-dalu-cia(n)`	An-da-lu-cia(n)
`An-dalu-sia(n)`	An-da-lu-sia(n)
`Bowditch`	Bow-ditch
`Brinkmann`	Brink-mann
`Canada`	Can-a-da
`Cana-dian`	Ca-na-di-an
`Chicago`	Chi-ca-go
`Franklin`	Frank-lin
`Ge-oge-bra`	Geo-gebra
`Hamil-ton(ian)`	Ham-il-ton(-nian)
`Haskell`	Has-kell
`Hedrick`	Hed-rick
`Hegelian`	Hegel-ian
`Hilbert`	Hil-bert
`Jeremiah`	Je-re-miah
`Laplace`	La-place
`Lester`	Les-ter
`Morawetz`	Mora-wetz
`Mordell`	Mor-dell
`Nicholas`	Nich-o-las
`Northamp-ton`	North-amp-ton
`Por-tuguese`	Por-tu-guese
`Richard`	Rich-ard
`Southamp-ton`	South-amp-ton
`Sylvester`	Syl-ves-ter
`Wilczyn-ski`	Wil-czyn-ski
`Wolfskehl`	Wolfs-kehl
`Woodrow`	Wood-row

⋄ Barbara Beeton
  http://tug.org/TUGboat
  TUGboat (at) tug dot org