

Re: Sorting of Chinese song titles in Karaoke books

Source: <http://sci.tech-archive.net/Archive/sci.lang/2005-03/3273.html>

From: Lee Sau Dan (danlee_at_informatik.uni-freiburg.de)

Date: 03/27/05

Date: 27 Mar 2005 09:52:56 +0800

>>>> "Charles" == Charles Belov <docorbit@invalid.invalid> writes:

Charles> I was at a multilingual karaoke place in the San
Charles> Francisco Bay Area, and decided to take a look at the
Charles> Chinese song book. I know just about 300 characters of
Charles> Chinese, not enough to even be dangerous, but I noticed a
Charles> few things about the sorting sequence of the titles.

Charles> – The primary sort is on the number of words in the song
Charles> title.

Charles> – Within that primary sort:

Charles> 1) Titles beginning with an English word come first 2)
Charles> Titles beginning with a Chinese numeral come next in
Charles> numerical order from first word to last word, that is

Charles> 11 1x 21 22 2x where 1 above is actually the Chinese
Charles> character for the number 1—they print the Chinese
Charles> character in the book—x is a non-numeric Chinese word

Charles> 3) Everything else comes last.

Did you observe any patterns under category (3)? Do characters with fewer strokes seem to come first?

I suspect that's a result of sorting based on the Big5 encoding.

Charles> I note in the Unicode CJK Unified Ideograph code block,
Charles> the Chinese character for 1 does indeed come first, but
Charles> it is not immediately followed by the Chinese character
Charles> for 2.

If not an encoding-based (big5 is more common in implementations than Unicode!) sorting, then it could be stroke count based sorting.

That's why 1 comes before 2.

Charles> I'm guessing there is some standard way of sorting the
Charles> "everything else," and my query to anyone who uses Chinese
Charles> karaoke song lists is "What is that sequence?"

Ordering by stroke count is the traditional standard. I think this
has been used for thousands of years. Name lists are usually
footnoted "names are ordered by stroke count" to avoid arguments that
the names that go first indicate more important or respected people.
Sometimes, lazier people will just leave the list 'randomly' ordered
and put a note "names are not ordered".

--

Lee Sau Dan

§õ|u'°

~{@nJX6X~}

E-mail: danlee@informatik.uni-freiburg.de

Home page: <http://www.informatik.uni-freiburg.de/~danlee>