

Re: Chinese character & pinyin frequency analysis

Source: <http://sci.tech-archive.net/Archive/sci.lang/2007-10/msg00373.html>

- *From:* Ruud Harmsen <realemailonsite@xxxxxxxxxxxxxxxxxxxxx>
 - *Date:* Wed, 10 Oct 2007 19:33:34 +0200
-

Wed, 10 Oct 2007 09:13:37 -0700: hxy <huangxinyu714@xxxxxxxxx>: in sci.lang:

Nice page. I have a question. The code for Chinese characters in your page does not look like unicode. Your code is `&#` plus 5 digits. I remember that uncode for Chinese characters is `&#x` plus 4 digits, at least I did that way with the unicode chart, and it works. What type of code is it in your page?

I suppose `&#` plus 5 digits is decimal, and `&#x` plus 4 digits is hexadecimal. The latter is easier because the Unicode codepages are also in hexadecimal. Otherwise it makes no difference and it just a matter of personal preference.

Some say decimal is supported by more browsers than hexaecimal, but that's now probably a thing of the past.

--

Ruud Harmsen, <http://rudhar.com>

.