Author: Kenneth Whistler
Title: Character Properties for avagrahas, etc.

Kent Karlsson submitted another errata report which has to be considered in terms of changes to existing character properties, rather than a fix to the BETA files for new characters.

See below for Kent's full report.

My assessment is:

1. Regarding the ellipses for Thai (U+0E2F) and Lao (U+0EAF), I believe these are conceived of something like "etc.", are shaped like letters, and are conceived of by Thais/Laos as letterlike. I think Lo is the deliberately chosen and correct general category for these, and should stay unchanged.

2. Regarding the avagrahas, Devanagari U+093D, etc., these are functionally "apostrophes" in Devanagari, used for the representation of the occasionally elided 'a' vowel in Sanskrit text. Rather than treating them as punctuation (Po), however, which would tend to break words at that position, I believe giving them the general category Lo is correct. This keeps the avagrahas within the scope of words for the purposes of word-breaking. Note that avagraha is traditionally transliterated with an apostrophe "'", but this is treated as a *letter* of the transliteration, comparable to the use of apostrophe for French liaison, or the use of apostrophe to transliterate a glottal stop, rather than as word-separating punctuation.

The inconsistency, however, is for the Khmer avagraha, U+17DC, used for Khmer script rendition of the Sanskrit avagraha. That should be consistent with the other avagrahas. So I believe the correct change to make is:

        U+17DC Po ---> Lo

--Ken

```
----- Begin Included Message -----

> -----Original Message-----
> From: Kent Karlsson [mailto:kentk@md.chalmers.se]
> Sent: Sunday, January 13, 2002 12:06 PM
> To: errata@unicode.org
> Subject: Two to six characters that aren't letters but have general
> category Lo
>
>
> These two characters work as ellipses, but have general category Lo:
>
> 0E2F;THAI CHARACTER PAIYANNOI;Lo;0;L;;;;;N;THAI PAI YAN NOI;paiyan
noi;;;
>     // ellipsis, abbreviation, not a letter;  Lo ---> Po
>
> 0EAF;LAO ELLIPSIS;Lo;0;L;;;;;N;;;;;
>     // ellipsis, (abbreviation), not a letter; Lo ---> Po
>
>
> The following ones are under the title "Sign" in the book, but has
> general category Lo:
>
> 093D;DEVANAGARI SIGN AVAGRAHA;Lo;0;L;;;;;N;;;;;
>     // letter or "sign"? (listed as "sign" on p. 402 of TUS3)
> Lo ---> Po?
>
> 0ABD;GUJARATI SIGN AVAGRAHA;Lo;0;L;;;;;N;;;;;
>     // letter or "sign"? (listed as "sign" on p. 411 of TUS3)
> Lo ---> Po?
>
> 0B3D;ORIYA SIGN AVAGRAHA;Lo;0;L;;;;;N;;;;;
>     // letter or "sign"? (listed as "sign" on p. 413 of TUS3)
> Lo ---> Po?
>
> 0E45;THAI CHARACTER LAKKHANGYAO;Lo;0;L;;;;;N;THAI LAK KHANG
YAO;lakkhang
> yao;;;
>     // letter or "sign"? (listed as "sign" on p. 431 of TUS3)
> Lo ---> Po?
>
>
> It appears that all of these should have general category Po instead.
>
> Note that this one already has category Po:
> 17DC;KHMER SIGN AVAKRAHASANYA;Po;0;L;;;;;N;;;;;
>

----- End Included Message -----
```