Hybrid Segmentation Ver 2.0

- *Aswin Shanmugam S*

DON Lab, IIT Madras

Compilation Steps: (Tested only on 64 bit linux machines)

1. Set 3 environmental variables,

   export NISTINC=/home/aswin/hybrid_segmentation/front-end-dsp/nist/include
   export NISTLIB=/home/aswin/hybrid_segmentation/front-end-dsp/nist/lib
   export DSPLIB=/home/aswin/hybrid_segmentation/front-end-dsp/src

   Replace "/home/aswin/" with the directory where hybrid_segmentation was copied

2. Install NIST. Goto the directory hybrid_segmentation/front-end-dsp/nist/. Type "sh src/scripts/install.sh" and choose "10" (linux) as the option. (I have inluded nist's SPHERE libraries only for 64bit linux.)
3. Goto the directory hybrid_segmentation/front-end-dsp/ and give "make -B"
4. Goto the directory hybrid_segmentation/front-end-dsp/Segmentation and give "make -B"
5. Goto the directory hybrid_segmentation/ and give "make -B"

Requirements: HTK, ch_wave, tcsh, perl

Performing segmentation for different data:

Inputs required:

1 Wavefiles at 16KHz sampling rate in the directory "**wav_16KHz/**"

2 Give the list of *Affricates*, *Fricatives*, *Nasals*, *SemiVowels*, *SibilantFricatives*, *Silence*, *StopConsonants* and *Vowels* in the directory "**Phonelist_Description/**"

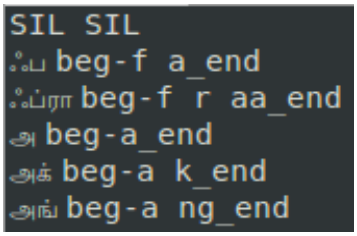
   Note: It is recommended to include only the unvoiced stop consonants in the *StopConsonants* list and leave the voiced stop consonants.

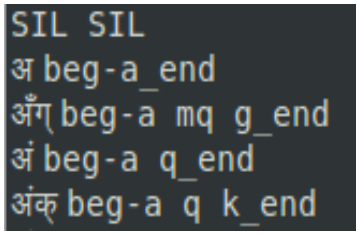
3 "**hmm/prompt-lab/**" should contain the syllable level transcription in *festival lab format* with random time-stamps (three column format with '#' in the first line. Only the transcription in the third column is important, the other two can be random. 'prompt-lab' generated from festival can be used)
   A screenshot of a sample 'promp lab' from Tamil is shown below:

```
#
0.1100 100 SIL
0.2200 100 என்
0.3300 100 பெ
0.4400 100 யர்
```

   Note: Silences should be marked only as 'SIL' and there shouldn't be two successinve SIL's in 'prompt lab'. Also silences in the beginning and end of an utterance is compulsory.

4 “**hmm/syldict**” should contain the list of unique syllables from the dataset and its phonetic transcription with begin, end context. [eg. (அல்ட் beg-a l tx_end), (அ beg-a_end), (चस्प् beg-c a s p_end)]
Screenshots of a portion of “syldict” from Tamil and Hindi are shown below:





Command for Execution:
sh run.sh

Note: It is better to divide the script “run.sh” into smaller chunks and then execute.

Output Directories:
After successful execution, final syllable and phone lab files will be present in “output_lab_syllable” and “output_lab_phone” respectively.

Note: If correct calculation of likelihood scores are needed, the phonelist description should also be updated in two scripts “hmm/scripts/cal_likelihood_category_hybrid.pl” and “hmm/scripts/cal_likelihood_category.pl”