Android Pocket Sphinx Implementation in Eclipse(ADT) and Using your own Models(Training done by Sphinx Train):

*First Step:(importing an existing project)*

Firstly import an existing tutorial project from below website to your Eclipse (ADT) or Android studio.(Ref: https://github.com/HelloSpoon/PocketSphinx\_Sample).

Some bugs may come while implementing ,kindly try to fix them.

Observe the written code over there and try to find the order of the calling of different callback methods used in the code.

Check the folder structure of the imported project. The important folders that you need a look are

a)Assets (where the models , grammar , dictionary required for recognition exists)

b)src (where MainActivity.java file exists ie, the code we have to modify as per our requirement)

Info about important call back methods in code:

When the audio chunck of data is processed “onPartialResult” method is called.

When the recognizer is stopped [after recognizer.stop() ] “onResult” method is called.

To get the continuous speech recognition FireRecognition() method in the code is to be called again after one recognition phase.

Integrating your models to the project

Delete old models from assets and copy your model files into that assets directory.

Also delete the old check sum files (.md5) and create the new .md5 files for the respective ones.

Also open assets.lst file in assets dir to check whether the paths are correct

Don’t forget to use your model dictionary file

Change the grammar file as per your recognition requirement .

Second Step(Models Creation and its issues)

I am using Ubuntu(Linux) for Models creation, it is easy for compilation and bug fixing on comparision to compilation process in windows OS.

I followed the tutorial from the below link: [www.speech.cs.cmu.edu/sphinx/**tutorial**.html](http://www.speech.cs.cmu.edu/sphinx/tutorial.html)

Download Sphinxbase , sphinxtrain from above link.

I will show how to create models for digits(0-9).

Firstly you need the recording of every digits such that all recording could sum over to atleast few minutes of length.

Create a folder with as your database name(I created it as “digits”)

Create two subfolders “etc” and “wav”.

Keep all your recording in the “wav” folder

The “etc” folder should be like:

* your\_db.dic - *Phonetic dictionary*
* your\_db.phone - *Phoneset file*
* your\_db.lm.DMP - *Language model*
* your\_db.filler - *List of fillers*
* your\_db\_train.fileids - *List of files for training*
* your\_db\_train.transcription - *Transcription for training*

creating .dic,.lm files(using cmusphinx lm tool)

firstly create a txt file with all words typed line by line and save it as digits.corpus file

go to the link: [www.speech.cs.cmu.edu/**tools**/**lmtool**.html](http://www.speech.cs.cmu.edu/tools/lmtool.html)

upload the corpus file and compile it in above link then you will get some files take the .dic and .lm files from those and rename them to digits.

Creating .dmp file from .lm file:

Go to sphinxbase dir and compile and make binaries using commands

./configure

make

then to convert .lm to .dmp use the below command.

Command: sphinx\_lm\_convert -i digits.lm -o digits.dmp

.sph format creation

Copy this file from the link to ur system

( <http://www.cs.cmu.edu/~./mharvill/RATS/software_releases/131008_SSB/go_wav2sph.pl> )

*In terminal go to folder location of this file use the following command*

**Perl go\_wav2sph.pl <input\_file\_list> <dest\_dir> <temp\_dir> <log\_dir>**

Input\_file\_list = txt file containing all the paths of .wav files

Create some dest,temp,log dirs and also mention their paths in the command

Finally all the respective .sph files are created in dest folder.

Converting mixture\_weights to sendump:

use the following command in linux:

First go to sphinx train folder,go to the folder that contains mk\_s2sendump.

./mk\_s2sendump

-pocketsphinx yes

-moddeffn path of this file in ur system/mdef

-mixwfn path of this file in ur system /mixture\_weights

-sendumpfn path of this file in ur system /sendump

Points to remember while creating .fileids and .transription files

After specifying the paths of all wav files don’t end the file with any new lines. Follow the same thing

For the transcript file also

Caution while creating .filler file

Any .filler file should be like this, the order should be like this else errors will be placed

<s> SIL

</s> SIL

<sil> SIL

Caution with .dic,.phone,.transcript(matching problem)

All the words uttered in .wav files to be written in capital letters only in transcripton file.

Phones that are created in .dic file should only be specified in .phone file

SIL phone should be included in .phone file

Sometimes it may give warning regarding the ONE in .dic file(phone HH not used in transcription).