LIUM Notes

**Resources**

Santa Barbara Corpus

http://www.linguistics.ucsb.edu/research/santa-barbara-corpus#SBC001

Java 1.8 appears to have some problems with JDK 1.8 (Check below for project with new build script)

https://github.com/ahmetaa/lium-diarization

Build command: ant -d stand-alone-jar (requires that ant be installed)

**Lium File Restrictions**

Audio can be in Sphere format or Wave format (16kHz / 16bit PCM mono) [Auto-detected by extension]

\*.seg is the output file containing the segmentation

Script to convert audio to proper Wave format (16kHz/16bit PCM mono)

ffmpeg -i test1.wav -acodec pcm\_s16le -ac 1 -ar 16000 test2.wav

Tool Files

Input:

Diarization file (\*.seg? File)

File containing acoustic vector or an audio file.

Diarization parameters (segmentation files)

Input

--sInputMask=<path> (Diarization file)

Can be

* absolute path: /home/myseg.seg
* relative path from the current directory: seg/myseg.seg
* a path where the %s is substituted: seg/%s.seg (%s is substituted by the show name param)

--sInputFormat

Output

-sOutputMask

-sOutputFormat

Feature Parameters (sound files)

Input:

-fInputMask

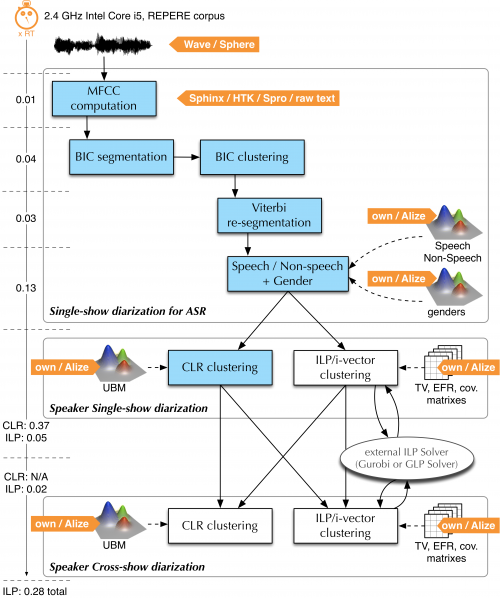
* absolute path: /home/myfile.mfcc
* relative path: file/myfile.wav
* %s is substituted: ./file/%s.sph (like ./file/myshow.sph)

-fInputDesc=type[:deltatype] (e.g. ="audio16kHz2sphinx,1:3:2:0:0:0,13,1:1:300:4")

Output:

-fOutputMask

-fOutputDesc



fr.lium.spkDiarization.

Compute the MFCC (tools.Wave2FeatureSet)

* Convert .wav to .mfcc in test\_out/t001/%s.mfcc

Check the MFCC (programs.MsegInit)

* Performs safety checks on generated file

GLR based segmentation, make small segments (programs.Mseg)

* Segmentation detection softare that finds instantaneous change point corresponding to segment boundaries.

Segmentation: linear clustering (programs.MClust)

* Hierarchical agglomerative clustering. Using Gaussians only.

hierarchical clustering (programs.MClust)

initialize GMM (programs.MtrainInit)

* Initialize the GMMs

EM computation (programs.MtrainEM)

* Train of a GMM using the EM algorithm

Viterbi decoding (programs.MDecode)

* Basic Viterbi decoder using a set of GMMs

Speech/Music/Silence segmentation (programs.MDecode)

filter spk segmentation according pms segmentation (tools.Sfilter)

Set gender and bandwith (programs.Mscore)

* A program that computes the likelihood scores given a set of GMMs

ILP Clustering (programs.ivector.ILPClustering)

Example of

**Wave2FeatureSet**

--fInputMask=./test\_wav/t001-1.wav

–fInputDesc=audio16kHz2sphinx,1:1:0:0:0:0,13,0:0:0

--fOutputMask=././test\_out/t001-1/%s.mfcc

--fOutputDesc=sphinx,1:1:0:0:0:0,13,0:0:0

**MSegInit**

--fInputMask=././test\_out/t001-1/%s.mfcc

--fInputDesc=sphinx,1:1:0:0:0:0,13,0:0:0

--sInputMask=

--sOutputMask=././test\_out/t001-1/%s.i.seg

**MSeg**

--fInputMask=././test\_out/t001-1/%s.mfcc

--fInputDesc=sphinx,1:1:0:0:0:0,13,0:0:0

--sInputMask=././test\_out/t001-1/%s.i.seg

--sOutputMask=././test\_out/t001-1/%s.s.seg

**MClust**

--fInputMask=././test\_out/t001-1/%s.mfcc

--fInputDesc=sphinx,1:1:0:0:0:0,13,0:0:0

--sInputMask=././test\_out/t001-1/%s.s.seg

--sOutputMask=././test\_out/t001-1/%s.l.seg

**MTrainInit**

--fInputMask=././test\_out/t001-1/%s.mfcc

--fInputDesc=sphinx,1:1:0:0:0:0,13,0:0:0

--sInputMask=././test\_out/t001-1/%s.h.3.seg

--tOutputMask=././test\_out/t001-1/%s.init.gmms

**MTrainEM**

--fInputMask=././test\_out/t001-1/%s.mfcc

--fInputDesc=sphinx,1:1:0:0:0:0,13,0:0:0

--sInputMask=././test\_out/t001-1/%s.h.3.seg

--tOutputMask=././test\_out/t001-1/%s.gmms

--tInputMask=././test\_out/t001-1/%s.init.gmms

**MDecode**

--fInputMask=././test\_out/t001-1/%s.mfcc

--fInputDesc=sphinx,1:1:0:0:0:0,13,0:0:0

--sInputMask=././test\_out/t001-1/%s.h.3.seg

--sOutputMask=././test\_out/t001-1/%s.d.3.seg

--tInputMask=./test\_out/t001-1/%s.gmms

**MDecode**

--fInputDesc=sphinx,1:3:2:0:0:0,13,0:0:0:0

--fInputMask=././test\_out/t001-1/%s.mfcc

--sInputMask=././test\_out/t001-1/%s.i.seg

--sOutputMask=././test\_out/t001-1/t001-1.pms.seg

--tInputMask=./models/sms.gmms

**SFilter**

--fInputDesc=sphinx,1:3:2:0:0:0,13,0:0:0:0

--fInputMask=././test\_out/t001-1/%s.mfcc

--sFilterMask=././test\_out/t001-1/t001-1.pms.seg

--sInputMask=././test\_out/t001-1/%s.d.3.seg

--sOutputMask=././test\_out/t001-1/t001-1.flt.3.seg

**MScore**

--fInputDesc=sphinx,1:3:2:0:0:0,13,1:1:0:0

--fInputMask=././test\_out/t001-1/%s.mfcc

--sInputMask=././test\_out/t001-1/t001-1.flt.3.seg

--sOutputMask=././test\_out/t001-1/t001-1.g.3.seg

--tInputMask=./models/gender.gmms

echo initialize GMM

java -Xmx$mem -classpath "$LOCALCLASSPATH" fr.lium.spkDiarization.programs.MTrainInit --help --nbComp=8 --kind=DIAG --fInputMask=$features --fInputDesc=$fDesc --sInputMask=./$datadir/%s.h.$h.seg --tOutputMask=./$datadir/%s.init.gmms $show

echo EM computation

java -Xmx$mem -classpath "$LOCALCLASSPATH" fr.lium.spkDiarization.programs.MTrainEM --help --nbComp=8 --kind=DIAG --fInputMask=$features --fInputDesc=$fDesc --sInputMask=./$datadir/%s.h.$h.seg --tOutputMask=./$datadir/%s.gmms --tInputMask=./$datadir/%s.init.gmms $show