Kaldi Codes

Ankit Grover

September 13, 2022

1. Print the and inspect Finite State Transducer.

```
fstprint --isymbols=lang/phones.txt --osymbols=lang/words.txt lang/L.fst
```

2. Draw the Monophone HMM model diagram(DT).

```
draw-tree data/lang/phones.txt exp/mono0a/tree |\
dot -Tps-Gsize=8,10.5 | ps2pdf - ~/tree.pdf;
cd ~;
evince tree.pdf ; # Evince is a PDF viewer
```

3. Print the GMM-HMM Model topology.

```
 gmm-copy --binary=false final.mdl - | head -n 20
```

4. Print the GMM-HMM/Nnet Model information.

```
gmm-info 40.mdl
nnet-am-info final.mdl
```

5. Print the Nnet2/3 Model topology information.

```
nnet-am-copy --binary=false final.mdl - | head
```

6. Show Forced Alignments

```
show-alignments phones.txt 40.mdl ark:1.ali | head
```

7. View HMM model transitions .

```
show-transitions data/lang/phones.txt exp/mon0a/final.mdl
```

8. Create Lexicon FST

```
python utils/lang/make_lexicon_fst.py  --sil-prob=0.5
--sil-disambig='#1'  --sil-phone="SIL" data/local/dict/lexiconp.txt |\

fstcompile --isymbols=data/lang/phones.txt \ --osymbols=data/lang/words.txt
--keep_isymbols=false --keep_osymbols=false |\

fstaddselfloops  data/lang/phones/wdisambig_phones.int lang/phones/wdisambig_words.int |\

fstarcsort --sort_type=olabel > L_disambig_py.fst
```

9. Draw the Lexicon FST diagram

```
fstdraw --isymbols=data/lang/phones.txt --osymbols=data/lang/words.txt L_disambig_py.fst |\
dot -Tpdf >Lpy.pdf
```

10. Map indices of transcript to text

```
tra=" ark:utils/sym2int.pl --map-oov $oov -f 2- $lang/words.txt $sdata/JOB/text| "

#tra is used as a form of mapped transcript for alignment, gop, other tasks
```

11. Strip accent , stress markers, pos-deps and list of pure phone IDs

```
../gop_speechocean762/s5/local/remove_phone_markers.pl phones.txt
new_phones.txt phone-to-pure-phone-mapping.int
```

12. View the .ark/.scp files via file format conversion.

```
head -n 10 | feats.scp | tail -1 | copy-feats scp:- ark,t:-
```

13. View the feature dimensions

```
feat-to-dim scp:data/feats.scp - | less
# OR
feat-to-dim ark:feats.ark - | less
```

14. Copy an integer vector file in a Kaldi.

```
copy-int-vector --binary=false vector_file - | head
```

15. Convert alignments to word level ctm. Do not ask me what I'm doing here.

```
linear-to-nbest ark:exp/mono0a/ali.1  'ark:utils/sym2int.pl -f 2-
data/lang/words.txt data/train_yesno/split1/1/text|'  '' '' 'ark:-' |\

lattice-align-words-lexicon "data/lang/phones/align_lexicon.int" "exp/mono0a/final.mdl"
"ark:-" "ark:-"  |\

nbest-to-ctm ark:- - | utils/int2sym.pl -f 5- data/lang/words.txt
```

16. Convert alignments to word alignments(when pos dep phones are present).

```
linear-to-nbest ark:exp/mono0a/ali.1  'ark:utils/sym2int.pl -f 2- data/lang/words.txt
data/train_yesno/split1/1/text|'  '' '' 'ark:-' |\

lattice-align-words "data/lang/phones/word_boundary.int"
"exp/mono0a/final.mdl" "ark:-" "ark:-"  |\

nbest-to-ctm ark:- - | utils/int2sym.pl -f 5- data/lang/words.txt
```

17. Convert alignments to phone ctm.

```
linear-to-nbest ark:exp/mono0a/ali.1  'ark:utils/sym2int.pl -f 2- data/lang/words.txt
data/train_yesno/split1/1/text|'  '' '' 'ark:-' |\

lattice-align-words-lexicon "data/lang/phones/align_lexicon.int" "exp/mono0a/final.mdl"
"ark:-" "ark:-" |\

lattice-to-phone-lattice "exp/mono0a/final.mdl"
"ark:-" "ark:-" |\

nbest-to-ctm ark:- - | utils/int2sym.pl -f 5- data/lang/phones/phones.txt
```

18. Get phone-level posteriors from post.ark (nnet-am) outputs

```
post-to-tacc --binary=false $dir/final.mdl "ark:ali-to-post $dir/ali|" 1.tacc

prob-to-post ark:wavlib/lab/posteriors/1272-0000_posterior.ark ark:- |\

post-to-pdf-post wavlib/exp/nnet2_online/nnet2_er/final.mdl ark:- ark,t:- |\

post-to-phone-post --transition-id-counts=1272-0000.tacc \
wavlib/exp/nnet2_online/nnet2_er/final.mdl ark:- ark,t:phone_temp.ark
```

Feature extraction

19. MFCC feature extraction

```
compute-mfcc-feats --config=conf/mfcc.conf \
scp:exp/make_mfcc/train/wav1.scp ark:/data/mfcc/raw_mfcc_train.1.ark
```

20. Cepstral Mean and Variance Normalization sufficient statistics

```
compute-cmvn-stats --spk2utt=ark:data/train.1k/spk2utt  scp:data/train.1k/feats.scp \
ark:exp/mono/cmvn.ark
```

21. I-Vector Feature Extraction

```
ivector-extract-online2 --config=conf/ivector_extractor.conf ark:data/train.1k/spk2utt \
scp:data/train.1k/feat.scp "ark,t:exp/nnet2_online/ivectors.1.ark"

# OR

gmm-global-get-post 1.dubm '$feats' ark:- | \
ivector-extract-online --ivector-period=10 final.ie '$feats' ark,s,cs:- ark,t:ivectors.1.ark
```