

Package for creation of Czech spontaneous speech recogniser

This package contains sample data and scripts for creation of Czech spontaneous speech recogniser. Within the training on limited amount of data, simple language model is created as well as models of Czech monophones and triphones.

CREATED AT:

SpeechLab at FEE CTU Prague,
Technicka 2, 166 27 Prague 6, Czech republic.
<http://noel.feld.cvut.cz/speechlab>

PACKAGE CONTENT:

LM: Language modelling part

- `texts/` - sample data for the training of language model
- `cut_for_OOV.pl` - words for training limitation
- `lm_train.bash` - training of the LM

AM: Acoustic modelling part

- `scripts/` - bash and perl scripts for acoustic model creation
- `config/` - sample configuration scripts
- `data/` - sample dataset for AM training
- `lists/` - lists of words and phones in the training dataset
- `misc/` - miscellaneous settings and lists
- `models/` - place for your acoustic models

INSTALATION:

- just untar the package
- be sure to have HTK Tools from <http://htk.eng.cam.ac.uk/> in your PATH

RUNNING:

LM: Enter the LM directory and run `lm_train.bash`.

Final LM can be found in `LM/lm_oov/tg1_cut0`.

Possible settings in `lm_train.bash`:

- `LM_size`: n-gram model degree <3>
- `cutoff`: cutoff degree for n-grams <0>
- `OOV_threshold`: number of words for LM training <60000>

AM: Enter the AM/scripts directory and follow these steps:

- 1) Run `parametrize.sh` - parametrizes the training data for use with ac. modelling.
- 2) Run `train_monophones.sh` - performs several steps for trainig monophone models.
Final models can be found in `AM/models/hmm9`.
- 3) Run `recognise_monophone.sh` - recognises the testing records (Czech digits) using monophones (95% accuracy).
- 4) Run `triphones_from_monophones.sh` - trains triphone models.
Final models can be found in `AM/models/hmm50`
Warning messages occure within the training due to the lack of training material.
- 5) Run `recognise_triphone.sh` - recognises the testing records (Czech digits) using triphones (100% accuracy).
- 6) Run `recognise_LVCSR.sh` - recognises the testing records (generous text) using triphones (12.77% accuracy).