

DT2112

Introduction to the Automatic Speech Recognition Exercise

Giampiero Salvi

KTH/CSC/TMH giampi@kth.se

VT 2019

Objective

Get familiarity with practical problems in Automatic Speech Recognition

You will **not** need to:

- ▶ implement ASR algorithms
- ▶ perform complex calculations

...but if interested all the source code is available

The Task

Build and test a “four digit” speech recogniser:

Example: “one seven four three”

Language: any language of your choice

Speaking mode: continuous speech

Speaking style: read speech

Speakers: speaker dependent

Vocabulary: small (10 words)

Robustness: low noise, close microphone

The Task

1. define a pronunciation dictionary
2. record training and test data
3. train monophone, single-mixture phone models
4. evaluate the recognition performance on the test data and on-line
5. answer the questions in the exercise protocol

Work in groups

- ▶ you have to test cross-speaker recognition
- ▶ form groups of maximum 3 people (min 2)
- ▶ borrow a USB headset for the recordings

The Environment

- ▶ the exercise is performed on Linux (Ubuntu)
- ▶ preferably use the CSC Ubuntu machines
- ▶ familiarity with Unix shell helps
- ▶ ... but there are detailed instructions and we are willing to help

You can try to perform the exercise on your own computers, but we can not help with the practical problems you may encounter

The Tools

you will get:

- ▶ a document with detailed instructions and the protocol to fill
- ▶ a package with all the relevant scripts and files
- ▶ the Hidden Markov Model Toolkit (HTK) for ASR related algorithms

To Pass

you have to:

- ▶ perform the steps described in the lab
- ▶ answer the questions described in the protocol
- ▶ hand in the answers through Canvas
(only one submission per group)

Practical Issues: Broken Audio

- ▶ click on loudspeaker icon top right of the screen
- ▶ choose “Sound settings...”
- ▶ Check the devices under “Play sound through”

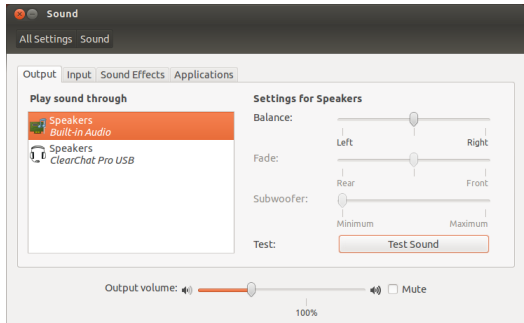
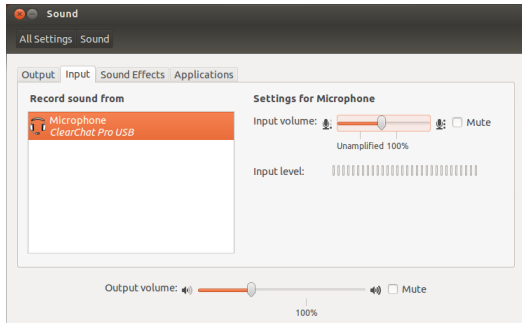


If the box is empty (broken audio settings):

1. logout from your session
2. press Ctrl+Alt+F1 (takes you to a text based login)
3. login with your usual user name and password
4. run the commands:

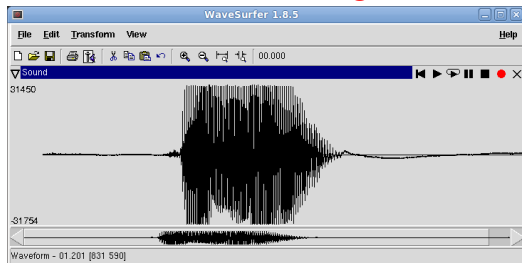
```
rm -rf .pulse .config/pulse
pulseaudio --kill
exit
```
5. press Ctrl+Alt+F7 to go back to the graphical login
6. login and the audio should work.

Practical Issues: Audio Settings



Practical Issues: Recording Levels

Bad recording!



Good recording:

