

HALADÓ GÉPI TANULÁS 7

Kornai András

HGT 2022/10/20

- Csoportbeszámolók
- HMM/GMM bef.
- NLP (szemantika)

MAIN PARAMETERS

- Number of states *between 3 and 15*
- Dimension of feature vector *between 12 and 20*
- Number of mixtures *between 8 and 32*
- Diagonal vs. Full *diagonal*
- Number of channels *between 1 and 3*
- Topology/stepsize *Bakis (between 1 and 8 pixels)*

CONCLUSIONS

- Feature extraction still a black art
- Dimension reduction by 75% rule of thumb valid
- Think globally, act locally

WHAT CAN WE LEARN?

- Generic ML techniques are not powerful enough to compete with systems architected by people who understand the task
- Within the broad HMM space, human-architected systems converge to one another in many respects
- The models so optimized embody knowledge about the black box

CHAIN VS. DIRECT MODELING

In *chain modeling* we build a (sequence of) physical model(s), measure their parameters as directly as possible and fit the rest. Example: finite element modeling of tongue/oral cavity with elastic deformations, followed by explicit acoustic modeling.

In *direct modeling* we abstract away from the physical process, detail parameters are fit directly (trained), gross parameters are fit indirectly (optimizing #defines). Example: DARPA-style speech work.

Direct modeling provides only indirect evidence for structure.

End-to-end optimization often a big win.

NLP

- Early attempts go back to 1940s and 1950s: as soon as there were computers, people wanted to use them for IR and MT
- Chain modeling was abandoned in speech in the 1980s, other areas followed only very slowly
- The box is not so black! Major analysis stages classically involve
- Segmentation of input into words (very nontrivial for Chinese)
- Segmentation of word stream into sentences (hard e.g. for speech reco output)
- Morphological analysis of words, lexical lookup
- Syntactic analysis of sentences aka *parsing*
- Computing the meaning
- Text generation based on meaning, speech to text