Speech Recognition technology from the ICSI Realization Group

Dan Ellis <dpwe@icsi.berkeley.edu> 1998feb26

Background

ICSI: International Computer Science Institute
UCB-affiliated independent research lab,
~75% directly-funded by European governments/industry

Realization Group: http://www.icsi.berkeley.edu/real/

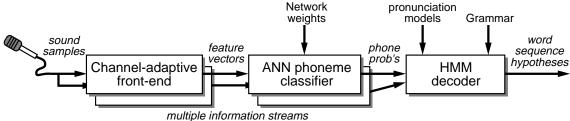People: Prof. Morgan, Prof. Steven Greenberg, 5 international post-docs,
12 UCB grad students, 2 technical staff

Mission: • Improve basic speech recognition, esp. robustness • Hardware...

Speech recognition

"Hybrid Artificial Neural Network/Hidden Markov Model system (ANN/HMM)" (IEEE SPmag May'95)
- low-complexity alternative to Gaussian mixture models, invented at ICSI.

Training: for • ANN weights (also pronunciations, grammar)
with • transcribed utterances (1000s or more) • takes 10s of hours

Accuracy: Numbers (39 words, continuous, telephone): 5-8% word error rate (WER)
Switchboard (unconstrained informal phone conversations): ? (ARPA best ~ 35%)

(see also: http://www.icsi.berkeley.edu/~dpwe/isrintro/)

Technology

• 3 programs (front-end, classifier, decoder), Unix commands, communicate via pipes

• **speed:** recognition ~ real time on Sun Ultra-1 (decoder is limit & can be 'pruned')

• **porting:** Autoconf packages for any? Unix; nothing exotic

+ Graphical User Interface tools: Tcl extensions (for Unix)