

Chapter 17. Speech Recognition and Understanding Systems

The Quest for Artificial Intelligence, Nilsson, N. J., 2009.

Lecture Notes on Artificial Intelligence, Spring 2016

Summarized by Jang, Ha-Young and Lee, Chung-Yeon

Biointelligence Laboratory
School of Computer Science and Engineering
Seoul National University

Contents

17.1 Speech Processing

Speech Processing

17.2 The Speech Understanding Study Group

The Speech Understanding Study Group

17.3 The DARPA Speech Understanding Research Program

17.3.1 Work at BBN

17.3.2 Work at CMU

17.3.3 Summary and Impact of the SUR Program

17.4 Subsequent Work in Speech Recognition

Subsequent Work in Speech Recognition

Appendix

Overview of Chapter 17

- Speech processing is divided between speech recognition and speech understanding
- The speech understanding study group
 - Work at BBN
 - Work at CMU
 - Supported by DARPA research program
- Subsequent work in speech recognition

Chapter 17. Speech Recognition and Understanding Systems

17.1 Speech Processing

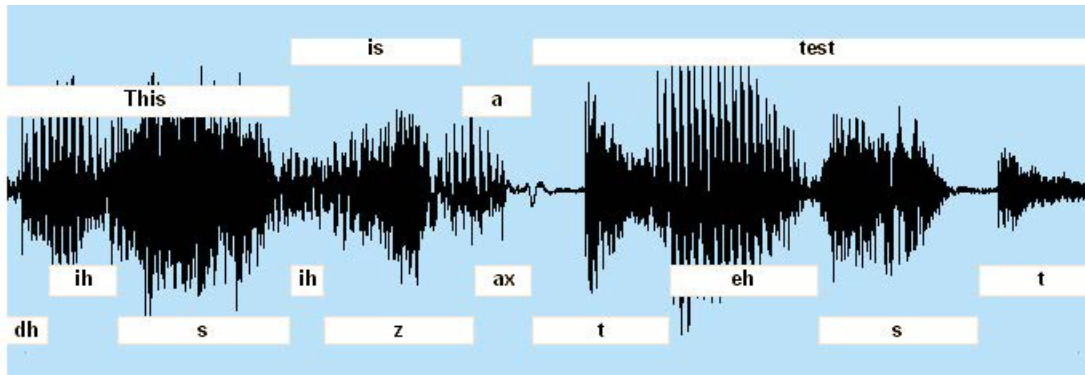
Speech Preprocessing

■ Speech recognition

- Process of converting an acoustic stream of speech input into a text representation of its component.

■ Speech understanding

- Understanding what is spoken.



A speech waveform

Symbol Example Sound Symbol Example Sound

Consonants

[p]	pat
[t]	tom
[k]	cat
[b]	boy
[d]	dip
[g]	garment
[m]	mat
[n]	nut
[ŋ]	sing
[f]	five
[v]	dove
[θ]	thistle
[ð]	feather
[s]	sat
[z]	haze
[ʃ]	smash
[ʒ]	ambrosia
[ç]	chic
[j]	page
[l]	lick
[w]	kiwi
[r]	parse
[y]	yew
[h]	horse
[q]	uh-oh (glottal stop)
[dx]	butter
[nx]	winter
[el]	thistle

Vowels

[iy]	lily
[ih]	miss
[ey]	lazy
[eh]	mess
[ae]	after
[aa]	pop
[ao]	orchestra
[uh]	wood
[ow]	lotus
[uw]	tulip
[uh]	butter
[er]	bird
[ay]	item
[aw]	flower
[oy]	toil
[y uw]	few
[ax]	ruffian
[ix]	lip
[axr]	leather
[ux]	dude

Consonants and vowels in the ARPAbet phonetic alphabet

Chapter 17. Speech Recognition and Understanding Systems

17.2 The Speech Understanding Study Group

The Speech Understanding Study Group

- **Feasibility study on a system that can recognize speech**
 - Larry Roberts in DARPA and Cordell Green in U.S. Army in early 1970
- **Meeting on speech processing**
 - Carnegie Mellon University at the end of March 1970
 - Form a 'study group' to make recommendations concerning the launching of DARPA supported project in speech understanding.
- **First meeting of the study group**
 - BBN on May 26 and 27, 1970
- **Final meeting of the study group**
 - SDC on July 26-28, 1970

Chapter 17. Speech Recognition and Understanding Systems

17.3 The DARPA Speech Understanding Research Program

17.3.1 Work at BBN




■ SPEECHLIS

- Answer spoken questions about the moon rocks database

■ HWIM

- Travel budget manager's automated assistant
- Respond to spoken questions

17.3.2 Work at CMU

- **Dragon** 
 - Designed to understand sentences about chess moves by James K. Baker
 - First examples of the use of Hidden Markov Models in AI.
- **HARPY** 
 - Bruce T. Lowerre designed and implemented the system
 - Understand spoken sentences and answer questions about, and to retrieve documents from, a database containing abstracts of AI papers
- **HEARSAY-II** 
 - Understand spoken sentences and answer questions about, and to retrieve documents from, a database containing abstracts of AI papers
 - Blackboard architecture

17.3.3 Summary and Impact of the SUR Program

- More thorough search of potential solutions
- More thorough built-in knowledge of transition phenomena between adjacent words
- More thorough testing, tuning, and debugging

Chapter 17. Speech Recognition and Understanding Systems

17.4 Subsequent Work in Speech Recognition

Subsequent Work in Speech Recognition

- HMM approach in DRAGON was ultimately adopted by all the leading speech recognition companies
- DARPA began funding speech recognition work again as part of its Strategic Computing program in 1984
- Dragon introduced Dragon NaturallySpeaking, a speech recognition program for personal computers

Chapter 17. Speech Recognition and Understanding Systems

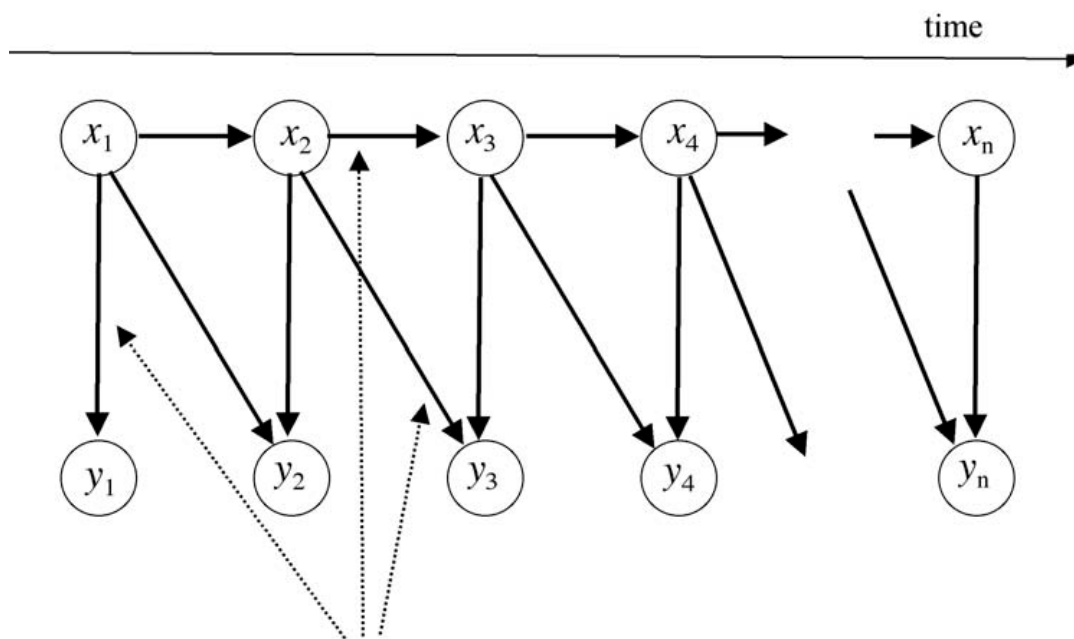
Appendix



17.3.2 Work at CMU

■ Dragon

- Designed to understand sentences about chess moves by James K. Baker
- First examples of the use of Hidden Markov Models in AI.



arrows indicate probabilistic influences

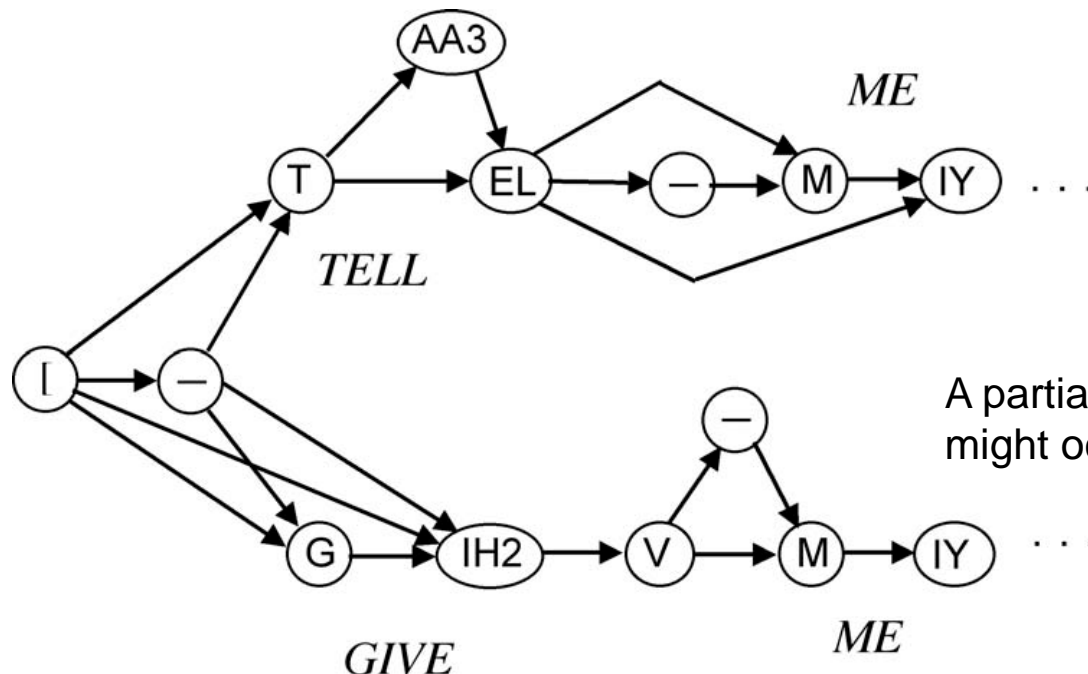
Two hierarchical levels in speech generation.



17.3.2 Work at CMU

■ HARPY

- Bruce T. Lowerre designed and implemented the system
- Understand spoken sentences and answer questions about, and to retrieve documents from, a database containing abstracts of AI papers



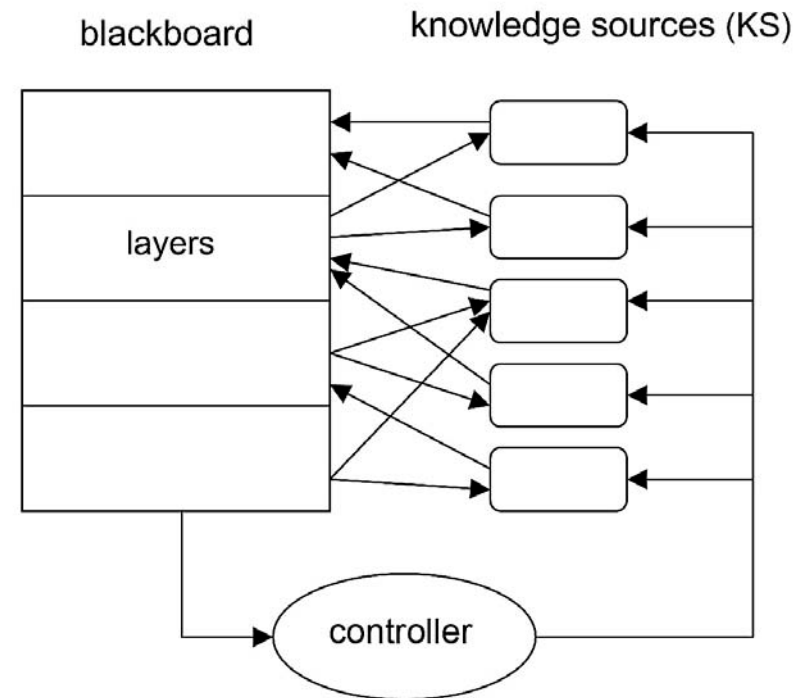
A partial network of the phones that might occur in a spoken sentence

17.3.2 Work at CMU



■ HEARSAY-II

- Understand spoken sentences and answer questions about, and to retrieve documents from, a database containing abstracts of AI papers
- Blackboard architecture



The Blackboard architecture