Chapter 17. Speech Recognition and Understanding Systems

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

Lecture Notes on Artificial Intelligence, Spring 2012

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

Biointelligence Laboratory
School of Computer Science and Engineering
Seoul National Univertisy

http://bi.snu.ac.kr

Contents

Overview of Chapter 17

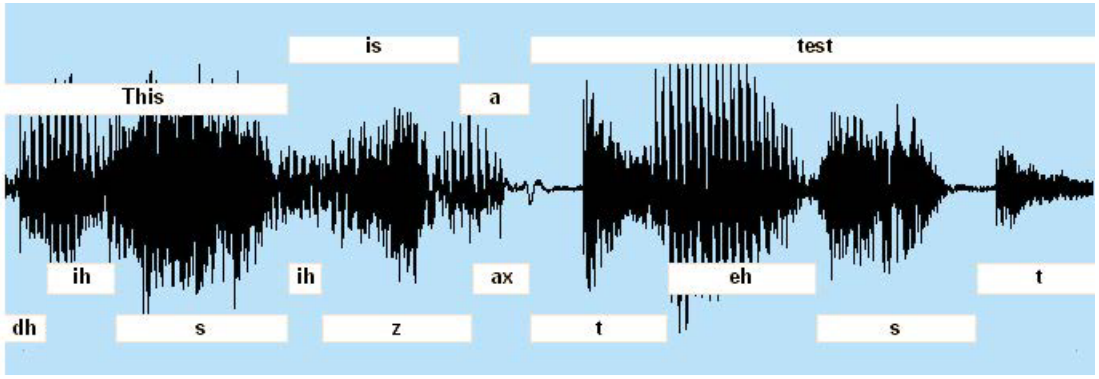
- Speech processing is divided between speech recognition and speech understanding
- The speech understanding study group
  - Work at BBN
  - Work at CMU
  - Supported by DARPA research progrm
- Subsequent work in speech recognition

Chapter 17. Speech Recognition and Understanding Systems

17.1 Speech Processing

Speech Precessing

- 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		**Vowels**	
[p]	**p**at	[iy]	li**y**
[t]	**t**om	[ih]	m**i**ss
[k]	**c**at	[ey]	l**a**zy
[b]	**b**oy	[eh]	m**e**ss
[d]	**d**ip	[ae]	**a**fter
[g]	**g**arment	[aa]	p**o**p
[m]	**m**at	[ao]	**o**rchestra
[n]	**n**ut	[uh]	w**oo**d
[ng]	si**ng**	[ow]	l**o**tus
[f]	**f**ive	[uw]	t**u**lip
[v]	do**v**e	[uh]	b**u**tter
[th]	**th**istle	[er]	b**ir**d
[dh]	fea**th**er	[ay]	**i**tem
[s]	**s**at	[aw]	fl**ow**er
[z]	ha**z**e	[oy]	t**oi**l
[sh]	sma**sh**	[y uw]	**few**
[zh]	ambro**s**ia	[ax]	r**u**ffian
[ch]	**ch**ic	[ix]	l**i**p
[jh]	pa**g**e	[axr]	leath**er**
[l]	**l**ick	[ux]	d**u**de
[w]	ki**w**i		
[r]	pa**r**se		
[y]	**y**ew		
[h]	**h**orse		
[q]	uh-oh (glottal stop)		
[dx]	bu**tt**er		
[nx]	wi**n**ter		
[el]	thistl**e**		

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

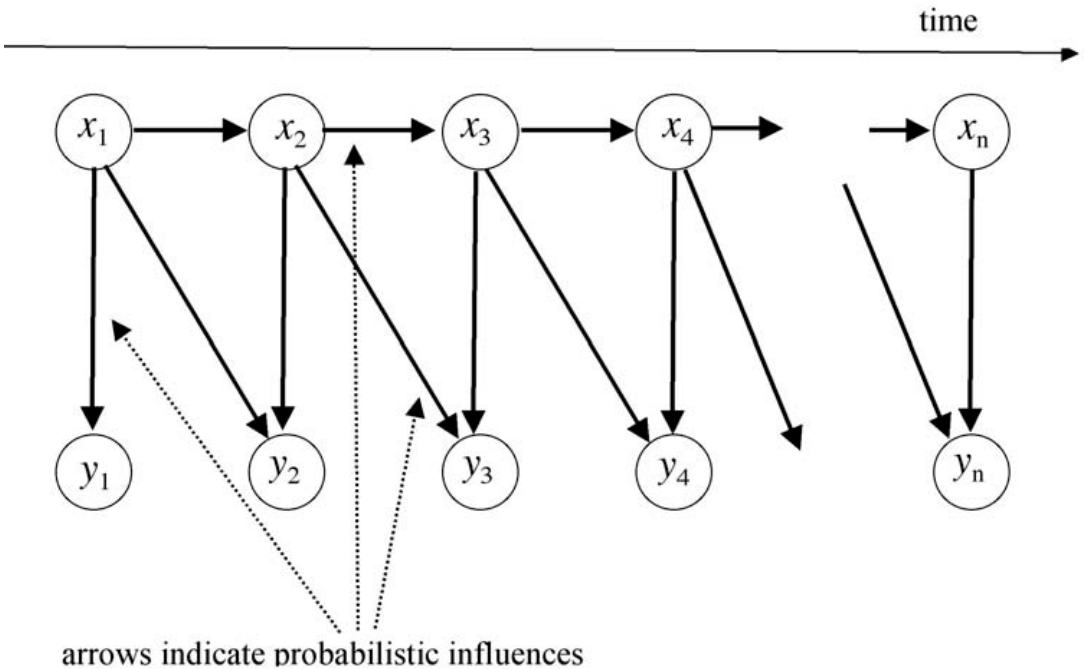
- HMM approach in DRAGON was ultimately adopted by all the leading speech recognition companies
- DARPA began funding speech recognition work again as part of its Strategic Computing program in 1984
- Dragon introduced Dragon NaturallySpeaking, a speech recognition program for personal computers

Chapter 17. Speech Recognition and Understanding Systems

Appendix

17.3.2 Work at CMU

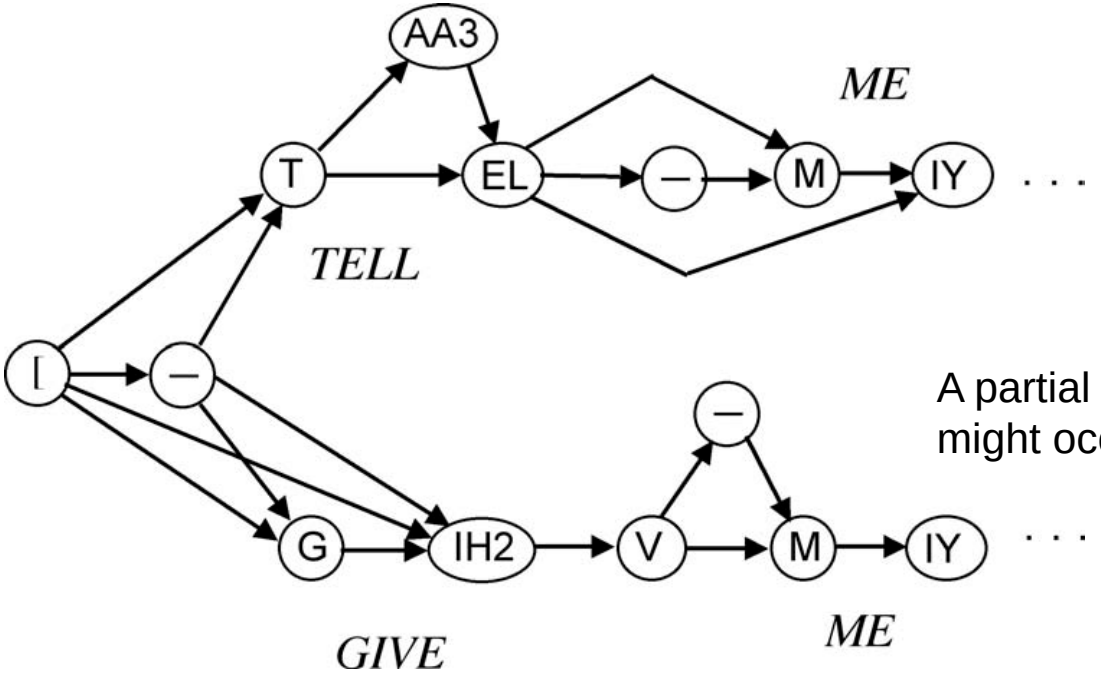
- Dragon
  - Designed to understand sentences about chess moves by James K. Baker
  - First examples of the use of Hidden Markov Models in AI.



Two hierarchical levels in speech generation.

17.3.2 Work at CMU

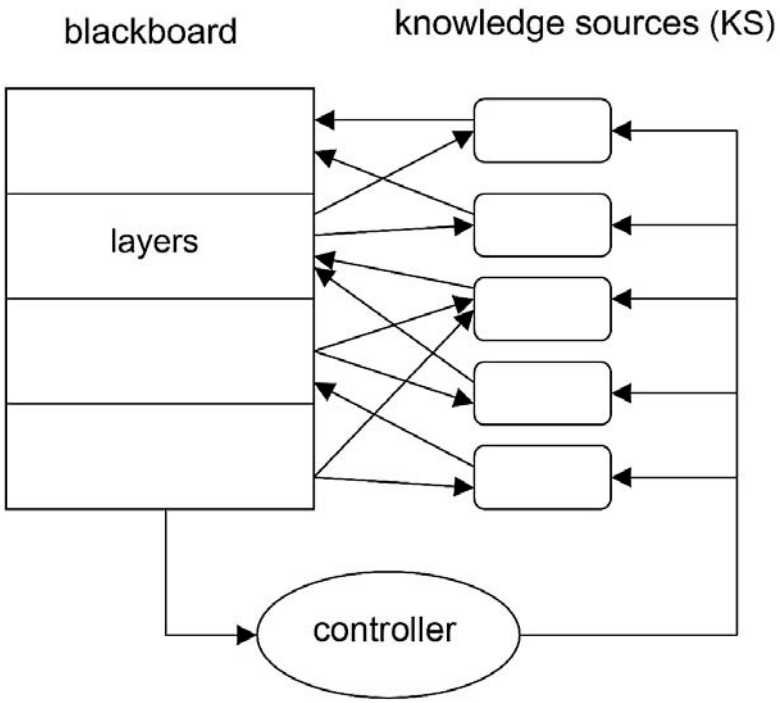
- HARPY
  - Bruce T. Lowerre designed and implemented the system
  - Understand spoken sentences and answer questions about, and to retrieve documents from, a database containing abstracts of AI papers



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