

Introduction to Automatic Speech Recognition (ASR)

- Applications

- Challenges

- types of speeches

- types of recognition methods

- - .

Sound (signal)
to text

Speech is viewed as future computer
interface'

Primary application:

1. spoken input (used directly)

2. speech as data/knowledge
(both ips is speech)

1. Spoken I/P: - for dictation

- navigation
- transactions
- personal letters
- business correspondence
- stocks purchase
- Reservations in airlines, buses, -----
- transfer balances

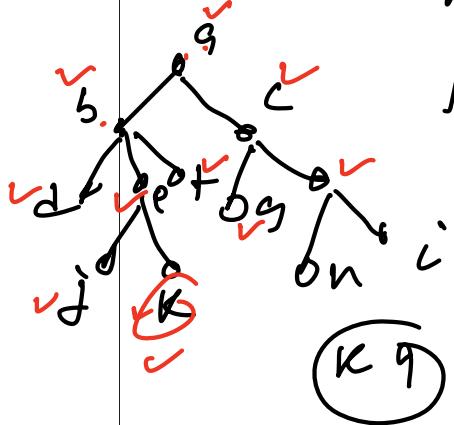
2. Data/Knowledge

- meeting capture
- knowledge management
- multi-media indexing

Others - voice dialing - IS + category
call routing ✓
data entry ✓
computer aided language learn-
ing ✓

Algorithms: ① probabilistic / stochastic

Bayes - HMM
probabilistic (hidden markov model)



② Graph search and
automata manipulations ✓

K q search

skew →

Can we classify speech?

- no. of speakers:
 - a) speaker dependent ASR
 - b) ... independent "

→ Cfg is trained to understand one speaker only. (dictation)
general purpose
- Nature of utterance
 - a. - IWR - Isolated word recognition
 - b. - Connected / continuous word recognition

Keyword spotting - search keywords
(GoogleSearch)

- Vocabulary size (10, or 100, or 1000s)
- spectrum band width:
 - telephone: 300 - 3400 Hz
(Narrow band)
 - wide-band (song)

— xx —

