Prepared By Abhiyan Dev Rimal
https://github.com/abhirimal/be-software-notes

Chapter=2 (1/2)

Sound and Audio

1) What is MIDI? Explain the need of MIDI standard. Write some features of MIDI

Ans: MIDI stands for Musical Instrument Digital Interface.

→ It is a protocol that enables computer, synthesizer, keyboard and other musical devices to communicate with each other.

⇒

→ It is a standard that manufacturers of electronic musical instruments have agreed upon.

→ It is a set of specifications they use in building their instruments so that the instruments of different manufacturers can, without difficulty, communicate musical information between one another.

→ Some of the features of MIDI are:-

i) MIDI helps to record music.

ii) It doesnot take much storage space.

iii) Easy editing with MIDI recordings.

iv) Synthesizing of different sounds.

v) The MIDI data format is digital.

Scanned with CamScanner

12) What are the four MIDI reception modes? Explain.

Ans: To tune a MIDI device to one or more channels, the device must be set to one of the MIDI modes.

→ There are four reception modes:

i) Mode 1: Omni On/Poly
ii) Mode 2: Omni On/Mono
iii) Mode 3: Omni Off/Poly
iv) Mode 4: Omni Off/Mono

→ The first half of the mode name specifies how the MIDI device monitors the incoming MIDI channels.

→ If Omni is turned On, the MIDI device monitors all the MIDI channels and responds to all channel messages, no matter which channel they are transmitted on.

→ If Omni is turned off, the MIDI device responds only to channel messages sent on the channel the device is set to receive.

→ The second half of the mode name tells the MIDI device how to play notes coming in over the MIDI cable.

→ If the option poly is set, the device can play several notes at a time.

Scanned with CamScanner

-> If the mode is set to Mono, the device plays ~~notes~~ ~~like~~ one note at a time.

3) MIDI Devices (Not asked)

-> They are:-

i) Sound Generators

-> It synthesizes the sound.

-> It produces an audio signal that becomes sound when fed into a loud speaker.

ii) Microprocessor

-> It communicates with the keyboard.

-> It sends and receives MIDI messages.

iii) Keyboard

-> The keyboard affords the musician's direct control of the synthesizer.

iv) Memory -> stores settings

v) Master keyboard

↳ increases the quality of the synthesizer.

Scanned with CamScanner

IIII 4) Explain different types of MIDI messages.

Ans. MIDI messages transmit information between MIDI devices.

→ It also determine what kinds of musical events can be passed from device to device.

→ MIDI messages are divided into two different types:

1) Channel messages
   i) Channel voice messages
   ii) Channel mode messages

2) System Messages
   i) System real-time messages
   ii) System common messages
   iii) System exclusive messages

1) Channel messages

→ Channel messages go only to specified devices.

→ There are two types of channel messages.

a) Channel voice messages

→ It sends actual ~~messages~~ performance data between MIDI devices. describing keyword action, controller action etc.

→ They describe music by defining pitch, amplitude, etc.

Scanned with CamScanner

b) Channel mode messages

→ Channel mode message determine ~~what~~ the way that a receiving MIDI device respond to channel voice message.

→ It deals with how to play notes coming in over MIDI cabels.

2) System messages

→ System messages go to all devices in a MIDI system because no channel numbers are specified.

→ There are three types of system messages.

a) System-real time messages

→ They avoid delay.

→ To avoid delay these messages are sent in the middle of other messages, if necessary.

b) System common messages

→ They are commands that prepares sequencers and synthesizer to play songs.

c) System exclusive messages

→ It allows MIDI manufactures to create customized MIDI messages to send between their MIDI devices.

Scanned with CamScanner

111 5) List out the names and the functions of MIDI software.

Ans: Once a computer is connected to a MIDI system, a variety of MIDI applications can run on it.

-> The software applications generally fall into four major categories.

i) Music recording and performance applications

-> This category of applications provides functions such as recording of MIDI messages.

-> It also supports editing, and playing back the messages in performance.

ii) Musical notations and printing applications.

-> They allow writing music using traditional musical notation.

-> User can play and print music on paper for live performance or publication.

iii) Synthesizer patch editors and librarians

-> These programs allow information storage of different synthesizer patches in the computer's memory and disk drives.

-> It also allows editing of patches in the computer.

Scanned with CamScanner

iv) Music education applications

→ These software applications teaches different aspects of music using the computer monitor, keyboard and other controllers attached ~~to~~ of attached MIDI instruments.

Q) Discuss the relationship between MIDI and SMPTE timing standard.

Scanned with CamScanner

classmate
Date
Page

IIIII 6) Define the term speech

Ans: Speech can be defined as the communication or expression of thoughts in spoken words.

→ Simply, speech is the ability to express thoughts through vocal sounds.

→ Speech can be perceived, understood and generated by humans and by machines too.

→ The brain can recognize the very fine line between speech and noise.

II 7) How is speech generated? Explain the various steps involved in sound generation.

or

List the sound generation technique and explain them briefly.

Ans: Speech generation is the ability to generate understandable and close to natural speech.

→ Requirements
- speech must be understandable
- speech must sound natural

→ There are two methods of speech generation and they are:-

1) Time-dependent sound concatenation
2) Frequency-dependent sound concatenation

Scanned with CamScanner

IIII i) Time-dependent sound concatenation

Ans. Speech generation can be performed by sound concatenation in timely fashion.

→ Individual speech units are composed like building blocks, where the composition can occur at different levels.

→ The individual phones are understood as speech unit.

Info

Phone → it is the smallest speech unit.
→ they are generally vowels or consonants.
→ m of mat or b of bat

Allophones → It is a kind of phone that changes its sound based on how a word is spelled.
→ think of the letter 't' and what kind of sound it makes in the word 'bit' compared to 'tip', 'little', etc.

→ Originally phones are stored.

→ Example: To produce the word 'crumb'. The individual phones K, r, A, m is used.

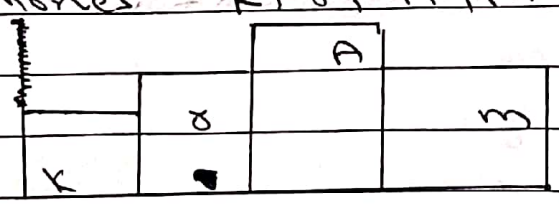

fig: Phone sound concatenation

Scanned with CamScanner

diphone -> pair of phones

-> But this technique is problematic since transitioning between phones while pronouncing make it sound unnatural.

-> So we can better improve transition by using ~~diphones~~ di-phones.

-> Two ~~words~~ ~~to~~ Two phones together can create a di-phone.

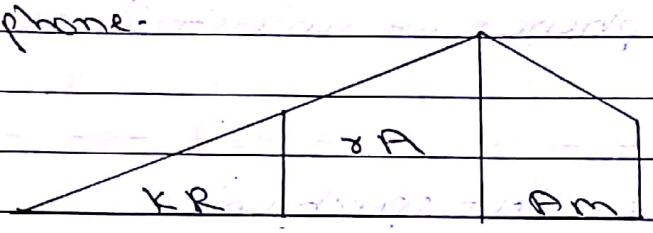

fig: Diphone sound concatenation.

-> To make the transition problem even easier, syllables can be created.

-> The best pronounciation of a word is achieved through storage of the whole word.

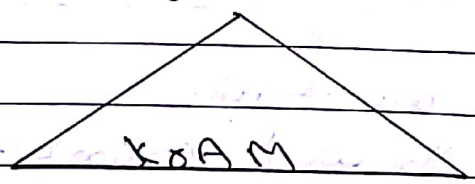

fig: Word Sound Concatenation

speech synthesis -> the computer generated simulation of human speech

12) Frequency-dependent sound concatenation

-> Speech generation can also be based on a frequency dependent sound concatenation.
-> This technique can be done through a formant synthesis.
-> Formants are frequency peaks in the spectrum which have a high degree of energy.
-> formant synthesis simulates the vocal tract through a filter.

-> Using speech synthesis, an existent text can be transformed into an acoustic signal.

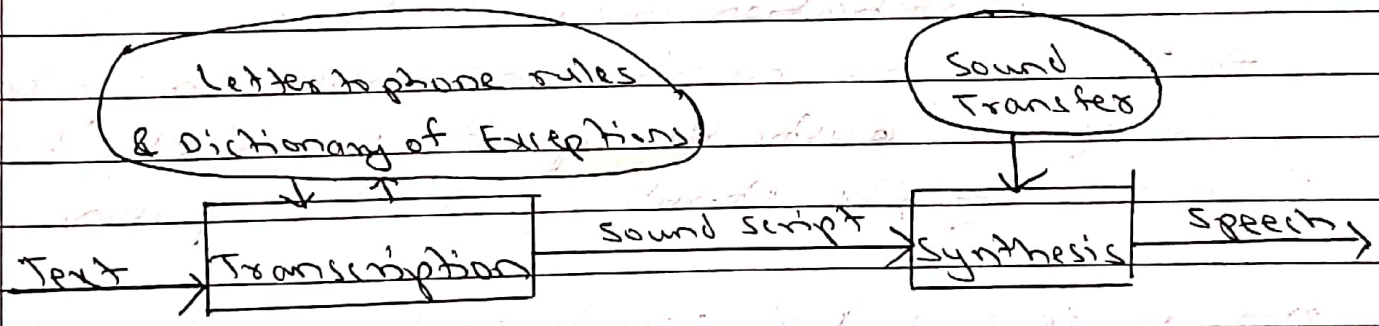

fig:- Components of speech synthesis system

-> In the first step, ~~the~~ transcription is performed, in which text is translated into sound script.
-> This process is done using letter to phone rules and dictionary of exceptions.

-> In the second step, the sound script is translated into a speech signal.

Scanned with CamScanner

8) Speech Analysis

Ans: Speech analysis is the process of analyzing speech to derive valuable insights.

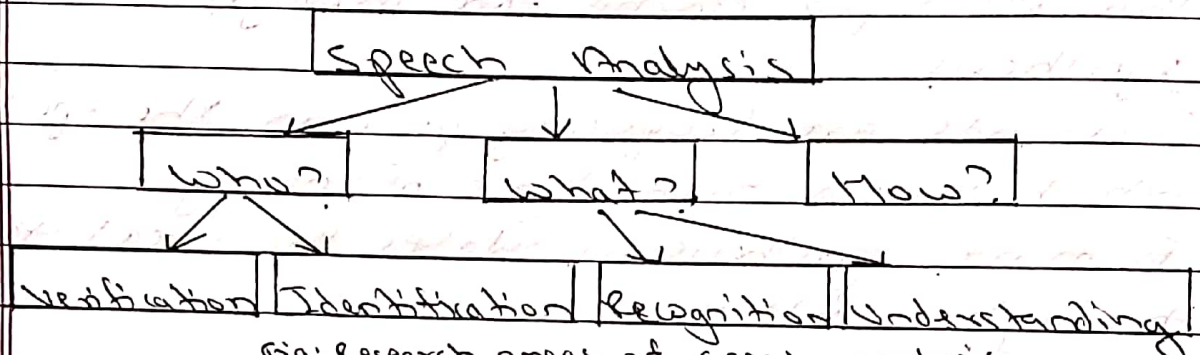

Fig: Research areas of speech analysis

1) → Speech analysis can serve to analyze who is speaking. i.e to recognize a speaker for his/her identification and verification.

→ The computer identifies and verifies the speaker using an acoustic fingerprint. An acoustic fingerprint is a digitally stored speech of a person.

2) → Another task of speech analysis is to analyze what has been said.
i.e to recognize and understand the speech itself.

3) → Another task is to analyze speech patterns with respect to how a certain statement was said.

→ Eg: A spoken sentence sounds different if a person is angry or calm.

→ An application of this research would be a lie detector.

Scanned with CamScanner

III a) Speech recognition and understanding.

or

Explain the components of speech recognition and understanding.

Ans:

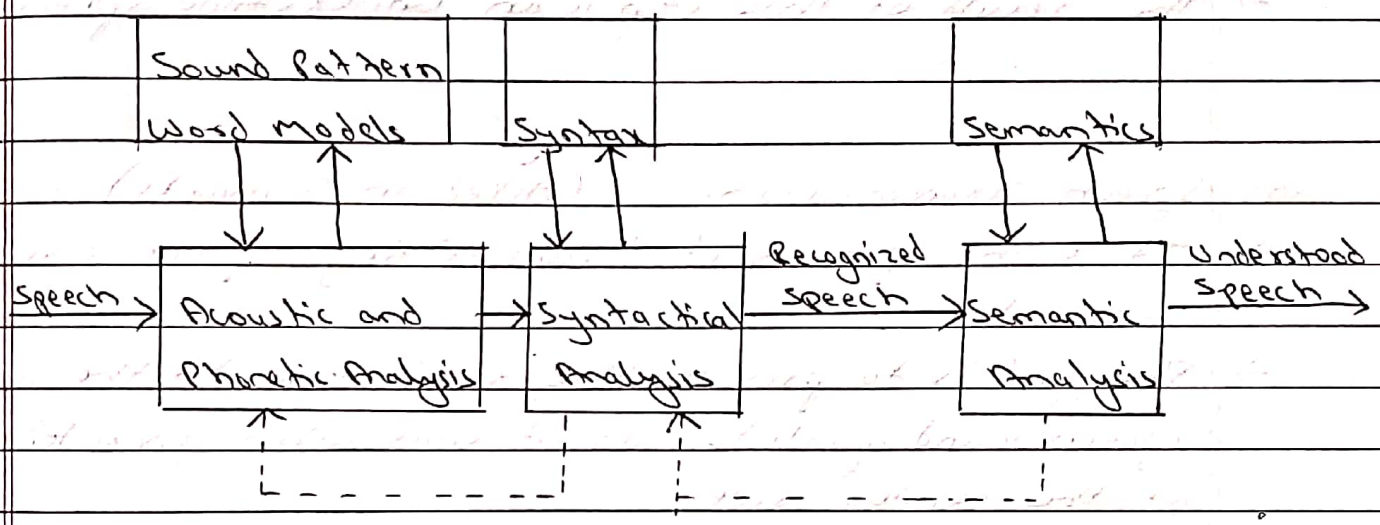

Fig:- Components of speech recognition and understanding

→ The system which provides speech recognition and understanding applies following principle:-

→ In the first step, the principle is applied to sound pattern and word model.

→ An acoustic and phonetic analysis is performed.

→ In the second step, certain speech units go through syntactical analysis.

→ In this step, the errors in the previous step can be recognized.

Scanned with CamScanner

semantics → the study of meanings.

→ The third step deals with the semantics of the previously recognized language.
→ Here the decision errors of the previous step can be recognized and corrected.
→ The result of this step is an understood speech.

10) Speech Transmission. (Asked once in 3 marks)
→ It is the process of transmitting speech over networks.
→ The goal is to provide the receiver with the same speech/sound quality as it was when generated at the sender side.

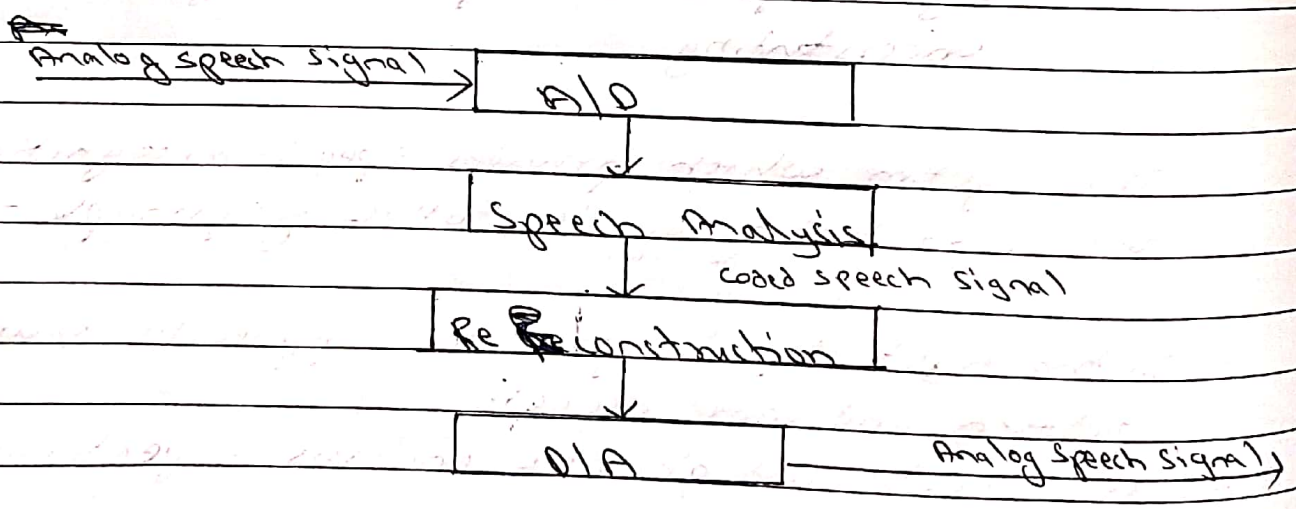

Fig:- Components of speech transmission system

A/D → audio to digital converter
D/A → digital to audio converter

Scanned with CamScanner

- First the speech is in analog form.
- Then it is converted to digital by A/D converter.
- Speech analysis is performed.
- Then the analyzed speech are reconstructed.
- Now the digital speech is converted to analog using D/A converter.
- Finally analog speech is received at receiver's end.

Scanned with CamScanner