Roll No.												

PRESIDENCY UNIVERSITY

BENGALURU

End - Term Examinations – MAY/ JUNE 2025

Date: 03-06-2025 **Time:** 09:30 am – 12:30 pm

School: SOE	**Program:** B. Tech-ECE	
Course Code : ECE3028	**Course Name:** Speech Signal Processing DEV	
Semester: VI	**Max Marks**: 100	**Weightage**: 50%

CO - Levels	CO1	CO2	CO3	CO4	CO5
Marks	14	34	26	26	

Instructions:

(i) Read all questions carefully and answer accordingly.

(ii) Do not write anything on the question paper other than roll number.

Part A

Answer ALL the Questions. Each question carries 2marks. **10Q x 2M=20M**

1.	Feature extraction transforms raw signals into useful numerical representations. Define speech signal?	2 Marks	L1	CO1
2.	The message information has a number of different representations during the process of speech production. Describe continuant and non-continuant sound.	2 Marks	L1	CO1
3.	In linear predictive spectrum analysis, the excitation effects are removed by focusing on the low-time autocorrelation coefficients. Write any four examples of time-domain parametric representations of the speech signal.	2 Marks	L1	CO2
4.	Speech coding is the process of compressing speech signals to reduce the amount of data required for storage or transmission, while preserving intelligibility and quality. Define short time zero crossing rate.	2 Marks	L1	CO2
5.	Speech coders enable a broad range of applications including narrowband and broadband wired telephony, cellular communications, voice over internet protocol (VoIP) and so on. Summarize two ways of interpreting STFT.	2 Marks	L1	CO3
6.	Text-to-speech synthesis systems are an essential component of modern human–machine communications systems. Draw the block diagram of filter summation method of synthesis.	2 Marks	L1	CO3
7.	Pattern recognition applications often occur in conjunction with other digital speech processing applications. Outline the conditions for filter banks in DFT view.	2 Marks	L1	CO3

8.	Speech coding at bit rates on the order of 8 Kbps enables normal voice conversations in cell phones. Define Cepstrum.	2 Marks	L1	CO4
9.	The goal of language translation systems is to convert spoken words in one language to spoken words in another language so as to facilitate natural language voice dialogues between people speaking different languages. Why linear vector spaces are needed in speech signal processing?	2 Marks	L1	CO4
10.	Speech enhancement, is to remove or suppress noise or echo or reverberation picked up by a microphone along with the desired speech signal. What is homomorphic filter?	2 Marks	L1	CO4

Part B

Answer the Questions. Total Marks 80M

11.	**a.**	The goal of speech enhancement systems is to make the speech more intelligible and more natural. Draw the schematic model of auditory mechanism and describe it.	**4 Marks**	**L2**	**CO1**
	b.	Speech can be represented phonetically by a finite set of symbols called the phonemes of the language, the number of which depends upon the language and the refinement of the analysis. Explain the applications of speech signal processing.	**3 Marks**	**L2**	**CO1**
	c.	Voiced sounds are produced when the vocal tract tube is excited by pulses of air pressure resulting from quasi-periodic opening and closing of the glottal orifice. Classify the three distinct sound processing sections of human ear.	**3 Marks**	**L2**	**CO1**
Or					
12.	**a.**	The sounds created in the vocal tract are shaped in the frequency domain by the frequency response of the vocal tract. Explain in detail of phonemes in American English.	**4 Marks**	**L2**	**CO1**
	b.	Samples of a speech signal are assumed to be the output of the time-varying linear system. Describe the mechanism of speech production through vocal cords with neat diagram.	**3 Marks**	**L2**	**CO1**
	c.	The acoustic wave is transmitted from the outer ear to the inner ear where the ear drum and bone structures convert the sound wave to mechanical vibrations. Extrapolate about speech properties and speech waveforms.	**3 Marks**	**L2**	**CO1**

13.	**a.**	Loudness is a perceptual quality that is related to the physical property of sound pressure level. Why short time analysis is needed?	**4 Marks**	**L2**	**CO2**
	b.	"Threshold of audibility" is the sound pressure level that is required for a sound of a given frequency to be just audible. Locate any 3 differences of rectangular and Hamming window.	**3 Marks**	**L2**	**CO2**
	c.	Most musical sounds as well as voiced speech sounds have a periodic structure when viewed over short time intervals. Estimate why parametric representation is needed?	**3 Marks**	**L3**	**CO2**
Or					
14.	**a.**	Pitch is a subjective attribute of sound that is related to the fundamental frequency of the sound. Determine the short time magnitude.	**4 Marks**	**L2**	**CO2**
	b.	Voiced speech is quasi-periodic, but contains many frequencies. Summarize about short time energy.	**3 Marks**	**L2**	**CO2**

	c.	Pitch period is used for the fundamental period of the voiced speech signal. Explain how automatic gain control is done using short time energy.	**3 Marks**	**L3**	**CO2**

15.	**a.**	Masking is the phenomenon of critical band auditory analysis in terms of vibrations of the basilar membrane. Enumerate the steps for DFT in speech signal.	**4 Marks**	**L2**	**CO3**
	b.	Pure tones can mask other pure tones, and noise can mask pure tones as well. Interpret the DTFT in speech signal processing.	**3 Marks**	**L2**	**CO3**
	c.	Because of the slowly varying nature of the speech signal, it is common to process speech in blocks ("frames). Explain the concept of sampling rates in time and frequency.	**3 Marks**	**L2**	**CO3**
Or					
16.	**a.**	Two basic short-time analysis functions useful for speech signals are the short-time energy and the short-time zero-crossing rate. Draw the block diagram of discrete-time model of speech production for voiced and unvoiced speech sounds.	**4 Marks**	**L2**	**CO3**
	b.	During the unvoiced interval, the zero-crossing rate is relatively high compared to the zero-crossing rate in the voiced interval. Explain the discrete-time model of speech production for voiced and unvoiced speech sounds.	**3 Marks**	**L3**	**CO3**
	c.	The autocorrelation function is often used as a means of detecting periodicity in signals. Estimate the supportive equations for the speech production for voiced and unvoiced speech sounds.	**3 Marks**	**L2**	**CO3**

17.	**a.**	STFT is a sequence of discrete-time Fourier transforms of windowed signal segments. Write about total sampling rate.	**4 Marks**	**L2**	**CO3**
	b.	Sound spectrogram has been a basic tool for gaining understanding of how the sounds of speech are produced and how phonetic information is encoded in the speech signal. Illustrate the impulse response of the composite system of filter banks.	**3 Marks**	**L2**	**CO3**
	c.	The window length is on the order of the length of a pitch period of the waveform during voiced intervals. Interpret the three important equation for filter bank analysis.	**3 Marks**	**L3**	**CO3**
Or					
18.	**a.**	If the analysis window is short, the spectrogram is called a wide-band spectrogram. Explain with neat diagram of analysis and synthesis operations for short-time spectrum analysis.	**4 Marks**	**L2**	**CO3**
	b.	Multiplying by the analysis window in the time-domain results in convolution of the Fourier transform of the window with the impulses in the spectrum of the periodic signal. Draw the composite frequency response for N = 6 equally spaced ideal filters.	**3 Marks**	**L3**	**CO3**
	c.	The FFT algorithm can be used to implement both analysis and synthesis. Predict the output of analysis/synthesis system.	**3 Marks**	**L2**	**CO3**

19.	**a.**	Short-time Fourier analysis and synthesis is generally formulated with equally spaced channels with equal bandwidths. Explain pitch period estimation using parallel processing approach with the neat block diagram.	**10 Marks**	**L3**	**CO2**

	b.	The complex cepstrum operator transforms convolution into addition. Generalize the model for speech production and synthesis.	**5 Marks**	**L3**	**CO2**
	c.	Masking occurs when one sound makes a second superimposed sound inaudible. Defend about short time analysis of speech.	**5 Marks**	**L3**	**CO2**
		Or			
20.	**a.**	The independent variable of the cepstrum and complex cepstrum is nominally time. Describe in detail about zero crossing rate with neat diagrams.	**10 Marks**	**L3**	**CO2**
	b.	The cepstrum was first applied in speech processing to determine the excitation parameters for the discrete-time speech model. Explain how zero crossing is affected by DC offset.	**5 Marks**	**L3**	**CO2**
	c.	The essence of the pitch detection algorithm proposed by Noll is to compute a sequence of short-time cepstra. With neat diagram and necessary equations, represent the short time zero crossing.	**5 Marks**	**L3**	**CO2**

.

21.	**a.**	Presence of a strong peak implies voiced speech, and the quefrency location of the peak gives the estimate of the pitch period. Explain about Cepstrum with necessary equations and its terminologies.	**10 Marks**	**L3**	**CO4**
	b.	Linear predictive analysis is one of the most powerful and widely used speech analysis techniques. Write in detail about homomorphic systems for convolution.	**5 Marks**	**L3**	**CO4**
	c.	The basic approach is to find a set of predictor coefficients that will minimize the mean-squared prediction error over a short segment of the speech waveform. Estimate homomorphic systems for convolution in representation of DTFT with neat block diagram.	**5 Marks**	**L3**	**CO4**
		Or			
22.	**a.**	The most widely used method of linear predictive analysis is called the autocorrelation method. Explain homomorphic analysis for voiced signal.	**10 Marks**	**L3**	**CO4**
	b.	The Levinson–Durbin formulation provides one more piece of useful information about the PARCOR coefficients. Analyze homomorphic analysis for unvoiced signal.	**5 Marks**	**L3**	**CO4**
	c.	Masking is widely employed in digital representations of speech signals by “hiding” errors in the representation in areas where the threshold of hearing is elevated by strong frequency components in the signal. Draw the time domain representation of glottal pulse and zero pole diagram.	**5 Marks**	**L3**	**CO4**