INDIAN STATISTICAL INSTITUTE
THIRD SEMESTER EXAMINATION, 2015/17 SESSION
PAPER - 13: INFORMATION STORAGE, RETRIEVAL AND DBMS
(MASTER IN LIBRARY INFORMATION SCIENCE)
21st November 2016 (10:00-13:00) (3 Hours)

This Question paper consists of 2 pages. **Attempt Questions and/or Sub-Questions sufficient enough to score 100 marks.** Please print all your answers in the answer Booklet provided. Scientific Calculator is allowed.

QUESTION 1.
Write briefly about
  i) Advanced information storage devices **[4 marks]**
  ii) IR in Big Data Analytics and Library Information Science **[4 marks]**
  iii) Role of RDBMS and E-R Diagrams in text information retrieval **[4 marks]**
  iv) Morphological distances **[4 marks]**
  v) Computation of Information dimension using multifractal spectra **[4 marks]**

QUESTION 2.
A. Show a system (a schematic diagram) for summarization of multiple documents **[4 marks]**
B. Write the formula with clear explanation of notations involved in K-Mixture model that is popular in the calculation of $P_i(k)$, probability of word (w_i) that appears k times in a document. **[8 marks]**

QUESTION 3.
What is spatial autocorrelation? Compute the Moran's Index for a spatial field of your choice. Choose at least five values in the spatial field. Show all steps involved in estimating Moran's I. **[15 marks]**

QUESTION 4.
What is Mahalanobis Distance? Compute it for the data, created of your choice, containing two variables (x and y) belonging to two groups (X and Y). Write your conclusions about the relevance of Mahalanobis Distance in the context of information retrieval. **[15 marks]**

QUESTION 5.
A. Write briefly about (i) Tonal similarity, and (ii) Time and pitch **[4 marks]**
B. Explain Hit-Or-Miss Transformation (HMT), and Blur Hit-Or-Miss Transformation (BHMT) and their use in Music Information Retrieval (MIR). **[8 marks]**

QUESTION 6.
Delimiter space of the first page of any document published in a standard technical journal possesses geometrically complex structure. Explain how rectangular granulometries with full set of equations could be used to quantitatively characterize such a delimiter space. **[8 marks]**

QUESTION 7.
Write a full set of equations involved in
A. Skeletonization by Influence Zone (SKIZ), and Weighted Skeletonization by Influence Zone (WSKIZ) transformations for three points (X^1, X^2, X^3) spread over discrete space, where $(X^1 \cap X^2 \cap X^3) = \phi$ **[8 marks]**
B. Morphological median computation between the (a) sets X and Y, where the $(X \cap Y) \neq \phi$, and (b) the two spatial functions f^1 and f^2, where $f^1 \wedge f^2 \neq f^1$ and $f^1 \wedge f^2 \neq f^2$. **[8 marks]**

QUESTION 8.
In Modified Gravity Model, additional required parameters are asymmetric distances, location significance indices. Write full set of equations to classify the three disjointed spatial zones namely X, Y, and Z of a spatial system S to compute variable-specific force of attraction between all possible combinations of the zones using the modified gravity model. Provide interaction matrices in support of the equations. **[10 marks]**

QUESTION 9.

A. Figure 1 is a four-bit greyscale image, $f(x,y)$, of size 7 x 7 pixels, depicting greyscale values of a picture in a document. These greyscale values range between 0 and 15 (4-bit image).

3	5	9	8	9	6	6
11	9	7	13	5	6	7
10	9	7	12	9	8	7
6	8	12	9	4	8	6
2	3	5	12	9	11	12
3	6	9	6	10	12	16
6	8	12	9	11	9	7

Figure 1

Compute $\sum_{x,y} f(x,y)$, and plot a histogram for the above 7 X 7 multi-valued image **[6 marks]**

B. What is thresholding? Convert the data shown in Figure 1 into binary form by choosing a median value as the threshold value.. **[6 marks]**

QUESTION 10.

Write a simple morphology-based algorithm to compute the ranks for pair of images. Following are the three spatial fields f^1, f^2, and f^3 of each of size 5X5.

208	25	41	37	168
231	72	248	108	10
33	140	245	234	217
233	245	124	203	239
162	247	205	245	174

(a)

71	128	192	245	215
174	245	66	140	65
168	87	130	36	208
42	150	179	39	63
31	58	228	66	237

(b)

44	140	47	237	79
166	76	94	198	130
187	190	160	125	131
166	49	199	112	209
115	176	21	114	203

(c)

The dilation $d(f^i, f^j)$ and erosion $e(f^i, f^j)$ distances interaction matrices for the above three spatial fields are given below:

	$d(f^i,f^j)$				$e(f^i,f^j)$		
	f^1	f^2	f^3		f^1	f^2	f^3
f^1	1	5	5	f^1	1	4	3
f^2	5	1	2	f^2	4	1	3
f^3	5	2	1	f^3	3	3	1

Compute the ranks for all possible pair of spatial fields. **[10 marks]**

END OF PAPER