Carnegie Mellon University
15-826 - Multimedia Databases and Data Mining
C. Faloutsos - Fall 20XX

SAMPLE FINAL EXAMINATION

- **3 hours** duration
- **No aids allowed** - only 2 pages of your notes, and calculator.
- You should have **4 non-empty pages, including this cover**
- Graded out of **100** points (*actually, this sample has less than 100 points*)
- Numbers in square brackets indicate points.

LAST NAME	
First Name	
andrew login	

Contents

Q1 DFT [10 pts]

Recall that the n-point DFT of a sequence $\vec{x} = [x_i]$ $(i = 1, \ldots, n)$ is defined as the sequence $[X_f]$, with

$$X_f = 1/\sqrt{n} \sum_{i=0}^{n-1} x_i \exp(-j2\pi fi/n) \quad f = 0, 1, \ldots, n-1 \tag{1}$$

where $j = \sqrt{-1}$ is the imaginary unit.

Consider a real-valued time sequence of n=32 time ticks ($t = 0, 1, \ldots, 31$). It has only one non-zero value in its amplitude spectrum (or two, if we count its mirror). Specifically, we have $|X_3|$=2.

1. ([**1 pts**]) How many sequences x_t $(t = 0, \ldots, 31)$ can lead to such a spectrum?
2. ([**3 pts**]) Could the sequence $x_t = 2\cos(3\pi t/32)$ give that spectrum? Justify briefly.
3. ([**6 pts**]) Give the formula x_t that generates a sequence with the above spectrum

Q2 DWT [10 pts]

Consider time sequences like $\vec{x} = [x_i]$ $(i = 1, \ldots, n)$ with n=32 time-ticks. Let n_{nz} be the number of non-zero values it has.

1. ([**4 pts**]) Give such a sequence, such that
   - it has as many zeros as possible (i.e., minimum n_{nz})
   - it has only 2 non-zero Haar wavelet coefficients
2. ([**6 pts**]) Is it possible to find a time sequence with only 1 non-zero value in the time domain, and 2 non-zero values in the (Haar) DWT domain? Explain briefly.

Q3 IFS [15 pts]

Consider Barnsley's method of Iterated Function Systems (IFS). The transformations of Figure 1 lead to the Cantor dust on the diagonal of the unit square, where all its points are disconnected.

1. ([**4 pts**]) Is it possible to have a fractal set, with fractal dimension $D > 1$, and with all its points disconnected?
2. ([**5 pts**]) If yes, describe or draw the corresponding transformations, and give the resulting fractal dimension, up to 2 decimal points. If not, explain briefly.

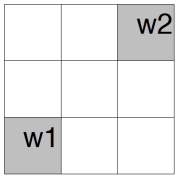

Figure 1: Barnsley's method: IFS transformations for the Cantor dust.

Q4 Graphs - Kronecker [10 pts]

Consider the two matries **A** and **B**:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \text{ and } \mathbf{B} = \begin{bmatrix} 10 & 20 & 30 \end{bmatrix}$$

1. **[3 pts]** What are the dimensions of the Kronecker product matrix **C**=**A** $\otimes$ **B**? A, posssibly correct answer, could be: 3 rows, 4 columns.
2. **[3 pts]** What is the value of the top left entry $c_{1,1}$ of **C**?
3. **[4 pts]** What is the value of the bottom right entry $c_{maxRow,maxCol}$?

Q5 Text [20 pts]

Consider the string editing distance and let the cost for insertion and deletion be $C_i = C_d = 1$; let the cost for substitution of any pair of differing letters be $C_s = 0.5$. Consider the strings ‘‘TOM’’ and ‘‘STUMP’’.

	-	T	O	M
-	0	1	2	3
S	1			
T	2			
U	3			
M	4			
P	5			

Table 1: Cost matrix for string editing distance

1. ([**15 pts**]) Give the $cost(i, j)$ matrix as in the course foils. Also see Table 1, which is partially filled for your convenience.
2. ([**3 pts**]) Show the lowest-cost path.
3. ([**2 pts**]) What is the string editing cost for the two input strings?

Q6 Other questions [35 pts]

And 4-5 more questions, on SVD, pageRank, power laws, forecasting (and, even 1 or 2, on z-ordering, signature files, etc), adding up to 100 points.

The exam questions will be similar to the quiz questions on the web:

`www.cs.cmu.edu/~christos/courses/826-resources/REVIEW-QUESTIONS`

END OF EXAM QUESTIONS _____GOOD LUCK