CARNEGIE MELLON UNIVERSITY
DEPARTMENT OF COMPUTER SCIENCE
15-826 MULTIMEDIA AND DATA MINING
C. FALOUTSOS, FALL 2024

Homework 2
Due: hard copy, in class, at **2:00pm, on Fri 09/20/2024**

VERY IMPORTANT:
- Deposit **hard copy** of your answers, in class.
- **Type** the full info on **each** page: your **name**, **Andrew ID**, **course#**, **Homework#**, **Question#** on each of the pages.

Reminders:
- *Plagiarism*: Homework is to be completed *individually.*
- *Typeset* your answers. You may use the pdf of the handout, to type/circle your answers. Illegible handwriting may get zero points.
- *Late homeworks*: Please follow the instructions here

For your information:
- Graded out of **100** points; **2** questions total
- Rough time estimate: *2-3 hours*
- Weight: 3% of total course grade.

Revision : 2024/09/13 00:20

Question	Points	Score
Z-/Hilbert ordering	80	
Graph patterns	20	
Total:	100	

Question 1: Z-/Hilbert ordering . . . . [80 points]

Consider a $2^n \times 2^n$ grid, and the z-curve on it. As usually, its first step is *vertical*, that is:

- the (0,0) cell has decimal z-value = 0
- the (0,1) cell is next, with decimal z-value = 1

Figure 1 shows the first step (arrow) of such a z-curve, on an 8×8 grid (which obviously has ranges (0,7) × (0,7)).
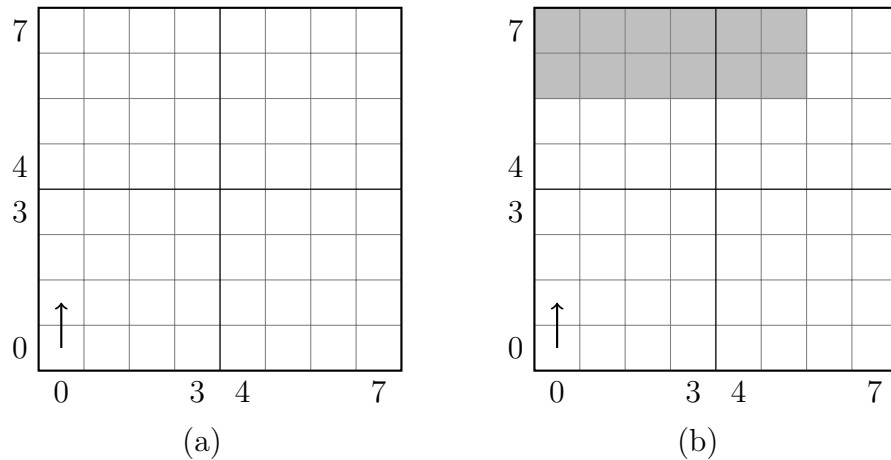

Figure 1:   Grids, for z-values

(a) [**5 points**]   What is the z-value of the (7,7) cell, in *decimal*?

(a) _____

(b) [**5 points**]   Again in decimal, what is the z-value of the (3,3) cell?

(b) _____

(c) [**15 points**]   *Hilbert:* Again in decimal, what is the Hilbert value of the (3,3) cell? (consider the Hilbert curve that has a vertical first step).

(c) _____

(d) [**15 points**]   Which cell(s) correspond to the *binary* z-value of 11-**-**? (We put 'dashes' for readability.) Shade your answer on Figure 1(a), or give the coordinates as $(x1, x2) \times (y1, y2)$.

(d) _____

(e) [**20 points**]   Give the z-value(s) in *binary*, for the gray-shaded region of Figure 1(b). Use "don't care" characters ("*"), as much as feasible.

(e) _____

(f) [**20 points**] Consider the regions $r1$ and $r2$, with binary z-values $z1$=01-**-**
and $z2$= 11-**-**. Circle your answer:

      A. They overlap partially (= non-zero common area, non-zero additional areas)

      B. $r1$ contains $r2$

      C. $r2$ contains $r1$

      D. They are disjoint (= zero common area), but they share a line segment

      E. They are disjoint, but share a point

      F. The are disjoint, apart from each other

Question 2: Graph patterns .... [20 points]

Consider graph with $N = 10^6$ nodes. We are told that it is a real graph, past its gelling point (that is, it has a giant-connected-component).

(a) [**10 points**] **Number of edges:** Which of the following is the best estimate for the number of edges E:

       A. $E \approx N^{0.5}$

       B. $E \approx N$

       C. $E \approx N^{1.5}$

       D. $E \approx N^2$

       E. $E \approx N^{2.5}$

       F. $E \approx N^3$

(b) [**10 points**] **Diameter**: Which of the following is the best estimate for the ('effective') diameter D: (Reminder: 'effective' diameter is the 90-percentile of all the pair-wise distances between nodes).

       A. $D \approx 0$

       B. $D \approx 1$

       C. $D \approx 2$

       D. $D \approx 6$

       E. $D \approx \log_2(N) \approx 20$

       F. $D \approx \sqrt{N} \approx 1000$