

Estructura de datos avanzados (CS3102) - Teo. 1 - Lab.

1.01 - 2022-2 Final

Miguel Angel Alvarado Ramos

TOTAL POINTS

17.5 / 28

QUESTION 1

1 Grid File 2.5 / 2.5

✓ + 2.5 pts *Correcto!*

+ 0 pts ![FQ1.png](/files/3beed28b-b407-4e83-

8f21-135663415409)

+ 1.5 pts ![FQ1.png](/files/3beed28b-b407-4e83-

8f21-135663415409)

QUESTION 2

2 Hamming/Levenshtein 2.5 / 2.5

✓ + 2.5 pts ***Correcto!***

+ 0 pts ***Distancia Hamming:*** \$\$\$

Distancia Levenshtein: \$\$\$

QUESTION 3

3 PM2 QuadTree 2.5 / 2.5

✓ + 2.5 pts *Correcto!*

+ 2 pts ![Q5P4.png](/files/9fef8717-8115-4754-bcc8-aa977fc05b42)

+ 1 pts ![Q5P4.png](/files/9fef8717-8115-4754-bcc8-aa977fc05b42)

+ 0 pts ![Q5P4.png](/files/9fef8717-8115-4754-bcc8-aa977fc05b42)

QUESTION 4

4 Well Separated Pair Descomposition

2.5 / 2.5

✓ + 2.5 pts *Correcto!*

+ 1.5 pts ![FQ4.png](/files/62189b0c-5d23-4b58-9e68-2ac8e693dc8c)

Fair split tree:

![FQ4.jpg](/files/86da95df-c6ff-4528-bea9-02e6ef740cae)

Well separated pairs:

- $\{d, e, f\}, \{g, h, i, j, k\}$

- $\{a, b, c\}, \{g, h, i, j, k\}$

- $\{a, b, c\}, \{d, e, f\}$

- $\{g, h, i\}, \{j, k\}$

+ 0 pts ![FQ4.png](/files/62189b0c-5d23-4b58-9e68-2ac8e693dc8c)

Fair split tree:

![FQ4.jpg](/files/86da95df-c6ff-4528-bea9-02e6ef740cae)

Well separated pairs:

- $\{d, e, f\}, \{g, h, i, j, k\}$

- $\{a, b, c\}, \{g, h, i, j, k\}$

- $\{a, b, c\}, \{d, e, f\}$

- $\{g, h, i\}, \{j, k\}$

+ 1 pts ![FQ4.png](/files/62189b0c-5d23-4b58-

9e68-2ac8e693dc8c)

****Fair split tree:****

![FQ4.jpg](/files/86da95df-c6ff-4528-bea9-02e6ef740cae)

****Well separated pairs:****

- $\{d,e,f\}, \{g,h,i,j,k\}$

- $\{a,b,c\}, \{g,h,i,j,k\}$

- $\{a,b,c\}, \{d,e,f\}$

- $\{g,h,i\}, \{j,k\}$

+ 2 pts ![FQ4.png](/files/62189b0c-5d23-4b58-9e68-2ac8e693dc8c)

****Fair split tree:****

![FQ4.jpg](/files/86da95df-c6ff-4528-bea9-02e6ef740cae)

****Well separated pairs:****

- $\{d,e,f\}, \{g,h,i,j,k\}$

- $\{a,b,c\}, \{g,h,i,j,k\}$

- $\{a,b,c\}, \{d,e,f\}$

- $\{g,h,i\}, \{j,k\}$

QUESTION 5

5 M-Tree 2 / 2.5

+ 2.5 pts Correcto!

+ 0 pts ![FQ5.png](/files/5ed5c3e6-4b41-4286-b811-e89a72370116)

****Árbol:****

![FQ5.2.png](/files/0bd5a559-dbf6-4b47-99d3-c1ae8b39af11)

✓ + 2 pts ![FQ5.png](/files/5ed5c3e6-4b41-4286-

b811-e89a72370116)

****Árbol:****

![FQ5.2.png](/files/0bd5a559-dbf6-4b47-99d3-c1ae8b39af11)

+ 1 pts ![FQ5.png](/files/5ed5c3e6-4b41-4286-b811-e89a72370116)

****Árbol:****

![FQ5.2.png](/files/0bd5a559-dbf6-4b47-99d3-c1ae8b39af11)

💬 No indicaste los pivots. Los nodos tienen capacidad 3.

QUESTION 6

6 Hull of Foci 2.5 / 2.5

✓ + 2.5 pts Correcto!

+ 0 pts ****Primer focus:**** Punto más lejano a h

Entonces f_1 : a

****Segundo focus:**** Punto más lejano a f_2

Entonces f_2 : k

Definimos: $V_k(p) = \sum_{n=0}^k |d(f_n, p) - \text{edge}|$

****Tercer focus:****

$V_2(b) = 10.29 + 3.36 = 13.66$

$V_2(c) = 6.22 + 4.71 = 10.94$

$V_2(f) = 3.95 + 7.48 = 11.43$

$V_2(g) = 7.69 + 6.28 = 13.97$

$V_2(h) = 4.87 + 8.76 = 13.63$

$V_2(i) = 2.72 + 10.64 = 13.36$

$$V_2(j) = 4.57 + 8.93 = 13.50$$

El valor mínimo se obtiene con c . Entonces

$$f_3: c$$

Cuarto focus:

$$V_3(b) = 10.29 + 3.36 + 10.64 = 24.30$$

$$V_3(c) = 3.95 + 7.48 + 11.60 = 23.04$$

$$V_3(f) = 7.69 + 6.28 + 6.70 = 20.68$$

$$V_3(h) = 4.87 + 8.76 + 10.64 = 24.27$$

$$V_3(i) = 2.72 + 10.64 + 8.76 = 22.13$$

$$V_3(j) = 4.57 + 8.93 + 5.55 = 19.05$$

El valor mínimo se obtiene con j . Entonces

$$f_4: j$$

Entonces, el conjunto foci es: $\{a, k, c, j\}$

+ 1.5 pts **Primer focus:** Punto más lejano a

$$h$$

$$f_1: a$$

Segundo focus: Punto más lejano a f_2

$$f_2: k$$

Definimos: $V_k(p) = \sum_{n=0}^k |d(f_n, p) - \text{edge}|$

Tercer focus:

$$V_2(b) = 10.29 + 3.36 = 13.66$$

$$V_2(c) = 6.22 + 4.71 = 10.94$$

$$V_2(f) = 3.95 + 7.48 = 11.43$$

$$V_2(g) = 7.69 + 6.28 = 13.97$$

$$V_2(h) = 4.87 + 8.76 = 13.63$$

$$V_2(i) = 2.72 + 10.64 = 13.36$$

$$V_2(j) = 4.57 + 8.93 = 13.50$$

El valor mínimo se obtiene con c . Entonces

$$f_3: c$$

Cuarto focus:

$$V_3(b) = 10.29 + 3.36 + 10.64 = 24.30$$

$$V_3(c) = 3.95 + 7.48 + 11.60 = 23.04$$

$$V_3(f) = 7.69 + 6.28 + 6.70 = 20.68$$

$$V_3(h) = 4.87 + 8.76 + 10.64 = 24.27$$

$$V_3(i) = 2.72 + 10.64 + 8.76 = 22.13$$

$$V_3(j) = 4.57 + 8.93 + 5.55 = 19.05$$

El valor mínimo se obtiene con j . Entonces

$$f_4: j$$

Entonces, el conjunto foci es: $\{a, k, c, j\}$

QUESTION 7

7 Circunferencia mínima 1 / 2

+ 2 pts Correcto!

+ 0 pts - Elija un punto aleatorio. Supongamos, sin perder generalidad, que el punto elegido es

$$p_3$$

- Seleccione el punto más lejano, haciendo $p_x = \max_{p \neq p_3} d(p, p_3)$. Asuma, por ejemplo, $p_x = p_1$.

- Seleccione el punto más lejano al punto previo $p_y = \max_{p \neq p_1} d(p, p_3)$. Asuma, por ejemplo, $p_y = p_2$.

- Calcule el centro $c = (p_1 + p_2) / 2$.

- Calcule el radio $r = d(c, p_1)$.

- **Si** $d(c, p_3) < r$: `return c, r`

- **Sino**:

$$D = 2 \cdot [a_1 \cdot (b_2 - c_2) + b_1 \cdot (c_2 - a_2) + c_1 \cdot (a_2 - b_2)]$$

$$c_x = \frac{1}{D} [(a_1^2 + a_2^2) \cdot (b_2 - c_2) + (b_1^2 + b_2^2) \cdot (c_2 - a_2) + (c_1^2 + c_2^2) \cdot (a_2 - b_2)]$$

```

]$$
$$c_y=\frac{1}{D}[(a_1^2+a_2^2).(c_1-b_1) +
(b_1^2+b_2^2).(a_1-c_1) + (c_1^2+c_2^2).(b_1-a_1)
]$$
Defina el centro $$c=(c_x,c_y)$$$. Calcule el radio
$$r=(c,p_1)$$$.
`return c,r`

```

✓ + 1 pts - Elija un punto aleatorio. Supongamos, sin perder generalidad, que el punto elegido es p_3 .

- Seleccione el punto más lejano, haciendo $p_x = \max_{p \neq p_3} d(p, p_3)$. Asuma, por ejemplo, $p_x = p_1$.

- Seleccione el punto más lejano al punto previo $p_y = \max_{p \neq p_1} d(p, p_3)$. Asuma, por ejemplo, $p_y = p_2$.

- Calcule el centro $c=(p_1+p_2)/2$.

- Calcule el radio $r=d(c, p_1)$.

- ****Si**** $d(c, p_3) < r$: `return c,r`

- ****Sino****:

```

$$D=2.[ a_1.(b_2-c_2) + b_1.(c_2-a_2) + c_1.(a_2-b_2)
]$$

```

```

$$c_x=\frac{1}{D}[(a_1^2+a_2^2).(b_2-c_2) +
(b_1^2+b_2^2).(c_2-a_2) + (c_1^2+c_2^2).(a_2-b_2)
]$$

```

```

$$c_y=\frac{1}{D}[(a_1^2+a_2^2).(c_1-b_1) +
(b_1^2+b_2^2).(a_1-c_1) + (c_1^2+c_2^2).(b_1-a_1)
]$$

```

```

Defina el centro $$c=(c_x,c_y)$$$. Calcule el radio
$$r=(c,p_1)$$$.
`return c,r`

```

QUESTION 8

8 Graphs 0 / 3

+ 3 pts Correcto!

✓ + 0 pts [!FQ8.jpg\]\(/files/0e8b1840-3d28-4ecb-a9f2-e5d9863663e3\)](#)

+ 0.5 pts Por el intento

[!FQ8.jpg\]\(/files/0e8b1840-3d28-4ecb-a9f2-e5d9863663e3\)](#)

QUESTION 9

9 MinHash 0 / 3

+ 3 pts Correcto!

✓ + 0 pts [Click here to replace this description.](#)

QUESTION 10

10 Gabriel Graph 0 / 1

+ 1 pts Correcto!

✓ + 0 pts ****No****. Un grafo es nevegable si el tiempo de búsqueda de cualquier punto es poli-logarítmico. Las redes Small World logran este desempeño debido a las conexiones a largo alcance. El grado de Gabriel no posee esta propiedad.

QUESTION 11

11 Smal-world 0.5 / 1

+ 1 pts Correcto!

✓ + 0.5 pts Es un grafo en el que la distancia entre dos nodos cualquiera crece con el logaritmo del número de datos.

También se puede definir como un grafo con alto coeficiente de agrupamiento global, pero bajo grado promedio de los nodos.

+ 0 pts Es un grafo en el que la distancia entre dos nodos cualquiera crece con el logaritmo del número de datos.

También se puede definir como un grafo con alto coeficiente de agrupamiento global, pero bajo

grado promedio de los nodos.

- ☞ Realmente no entiendo lo que pusiste, pero propones algunos conceptos claves.

QUESTION 12

12 Grafo plano 0 / 0.5

+ 0.5 pts Correcto!

✓ + 0 pts **Respuesta: c)**

Los modelos de Small World no generan grafos planos debido a las conexiones a larga distancia.

QUESTION 13

13 Shared memory 0 / 0.5

+ 0.5 pts Correcto!

✓ + 0 pts **Respuesta: a)**

![Q6P6.png](/files/79dceec9-48fe-4839-8687-11c716415f21)

QUESTION 14

14 Sector tree 0.5 / 0.5

✓ + 0.5 pts Correcto!

+ 0 pts **Respuesta: b)**

Sector Tree no divide los radios, solo segmentos de arco.

QUESTION 15

15 Grid File: fusión 0.5 / 0.5

✓ + 0.5 pts Correcto!

+ 0 pts **Respuesta: d)**

Los tres son criterios para fusionar celdas en Grid File.

QUESTION 16

16 Minkowski distance 0.5 / 0.5

+ 0.5 pts Correcto!

✓ + 0.5 pts **Respuesta: $\pi_1 = \pi_{\infty} > \pi_2$ **

![minkowski.png](/files/992a1d74-5f05-4de8-8c6e-377129176ea9)

Si el radio es r , entonces:

- $\pi_1 = \frac{4 * (|1-0| + |0-1|)}{2} = 4$

- $\pi_2 = 3.1416 \dots$

- $\pi_{\infty} = \frac{4 * (\max(1,1) + \max(1,1))}{2} = 4$

QUESTION 17

17 OS-Tree 0 / 0.5

+ 0.5 pts Correcto

✓ + 0 pts **Respuesta: d)**

La estructura que emplea puntos adicionales es Probably Correct OS-Tree.

Profesor: Victor Flores Benites

Apellidos: Alvarado Remy

Nombres: Moisés Angel

Sección: _____

Fecha: 28/11/22

Nota:

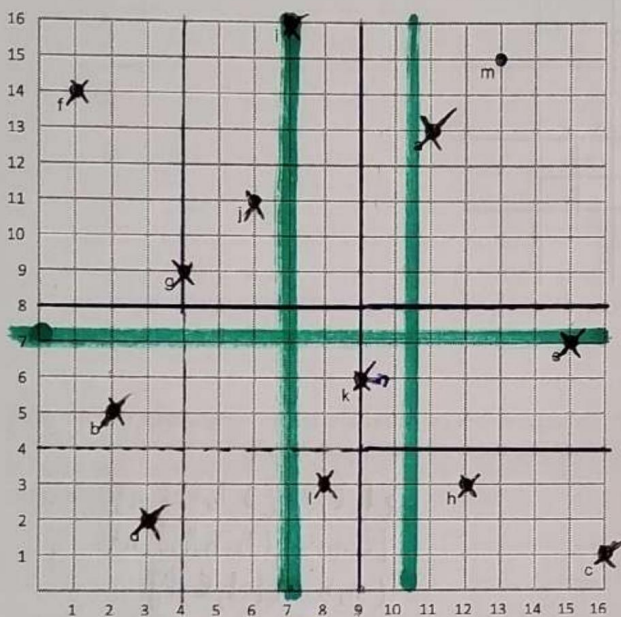
Indicaciones:

La Duración es de **120 minutos**.

La evaluación consta de **12 preguntas**.

Pregunta 1 (2.5 puntos)

Inserte los siguientes puntos, en orden alfabético, a un Grid File. Considere capacidad del bucket $d = 3$.



Resuelto en el blog de Notas.

Pregunta 2 (2.5 puntos)

Calcule las distancias Hamming y Levenshtein de las palabras: **AFGANISTAN** y **UZBEKISTAN**

	A	F	G	A	N	I	S	T	A	N
U	1	2	3	4	5	6	7	8	9	10
Z	2	1	2	3	4	5	6	7	8	9
B	3	2	3	4	5	6	7	8	9	10
E	4	3	4	5	6	7	8	9	10	11
K	5	4	5	6	7	8	9	10	11	12
I	6	5	6	7	8	9	10	11	12	13
S	7	6	7	8	9	10	11	12	13	14
T	8	7	8	9	10	11	12	13	14	15
A	9	8	9	10	11	12	13	14	15	16
N	10	9	10	11	12	13	14	15	16	17

AFGANISTAN

UZBEKISTAN

1 1 1 1 1 0 0 0 0 0

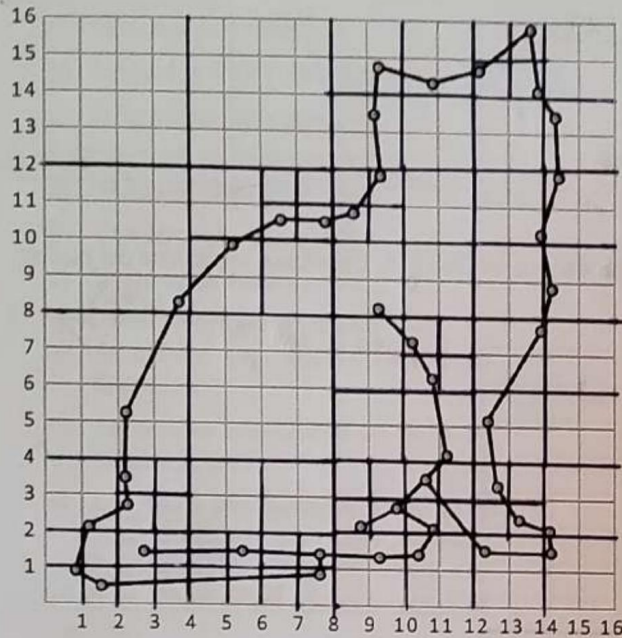
5

Distancia Hamming (0.5 pts): 5

Distancia Levenshtein (2 pts): 5

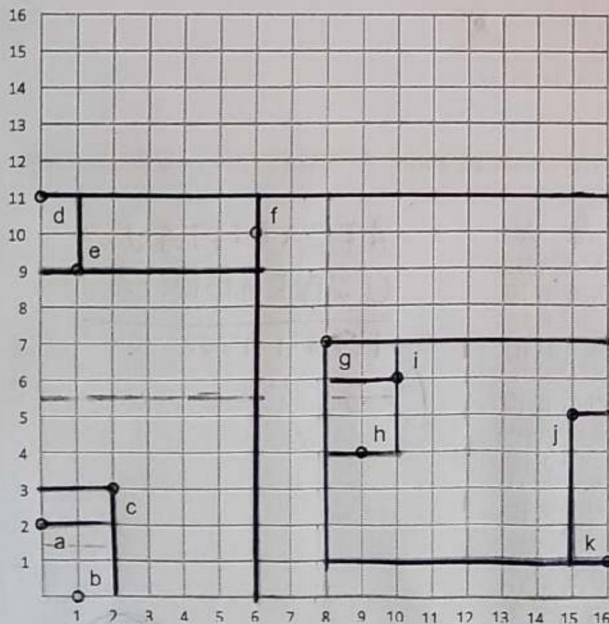
Pregunta 3 (2.5 puntos)

Inserte los siguientes puntos y aristas a un PM₂ Quadtree.

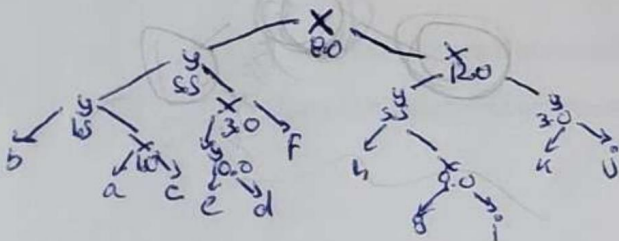


Pregunta 4 (2.5 puntos)

Estime los Well Separated Pairs (WSP) de los siguientes puntos. De como respuesta el Fair split tree, y el conjunto de WSP (solo ejecute el algoritmo en el root y sus hijos).



$\{d, e, f\}, \{g, h, i, j, k\}$
 $\{a, b, c\}, \{g, h, i, j, k\}$
 $\{a, b, c, d, e, f\}$
 $\{g, h, i, j, k\}$

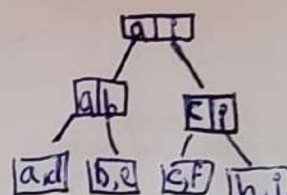
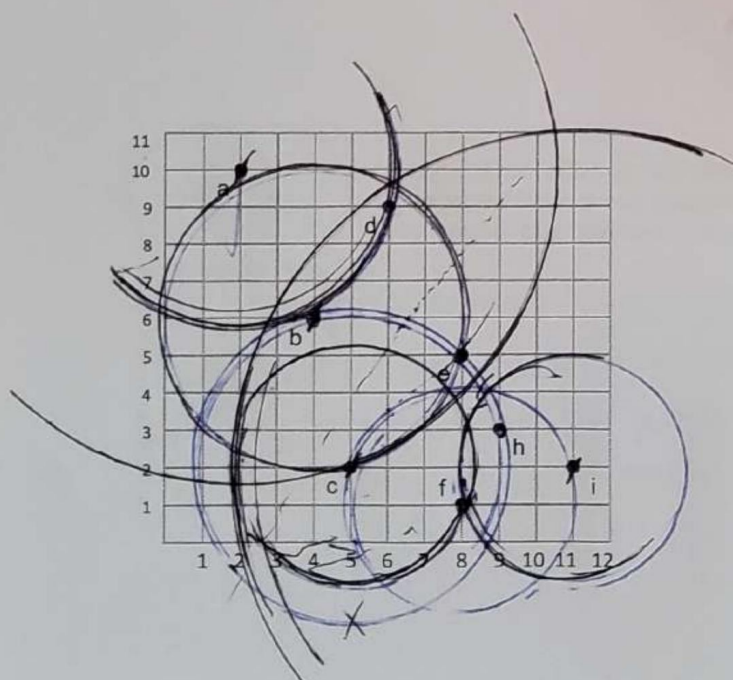


Emplee la siguiente tabla de distancias en las **preguntas 5 y 6**.

	a	b	c	d	e	f	g	h	i	j	k
a	0.00	8.49	8.19	2.61	2.39	5.56	7.45	5.62	5.07	7.02	10.71
b	8.49	0.00	10.45	11.93	10.98	12.20	5.15	9.21	5.20	2.18	8.81
c	8.19	10.45	0.00	14.27	4.07	11.92	7.56	6.51	2.81	0.81	11.24
d	2.61	11.93	14.27	0.00	5.23	13.91	6.35	4.70	2.21	2.88	10.65
e	2.39	10.98	4.07	5.23	0.00	15.43	11.82	11.42	12.88	14.58	4.26
f	5.56	12.20	11.92	13.91	15.43	0.00	14.91	5.86	14.51	2.98	4.07
g	7.45	5.15	7.56	6.35	11.82	14.91	0.00	6.82	15.00	14.32	14.84
h	5.62	9.21	6.51	4.70	11.42	5.86	6.82	0.00	16.86	3.92	16.82
i	5.07	5.20	2.81	2.21	12.88	14.51	15.00	16.86	0.00	4.67	19.18
j	7.02	2.18	0.81	2.88	14.58	2.98	14.32	3.92	4.67	0.00	22.70
k	10.71	8.81	11.24	10.65	4.26	4.07	14.84	16.82	19.18	22.70	0.00

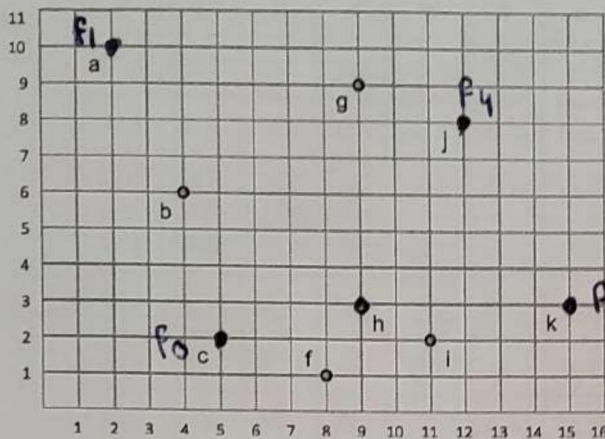
Pregunta 5 (2.5 puntos)

Inserte los siguientes puntos, en orden alfabético, en un M-Tree ($m = 2$, $M = 3$). De como respuesta el árbol resultante (indique los pivots) y las regiones generadas.



Pregunta 6 (2.5 puntos)

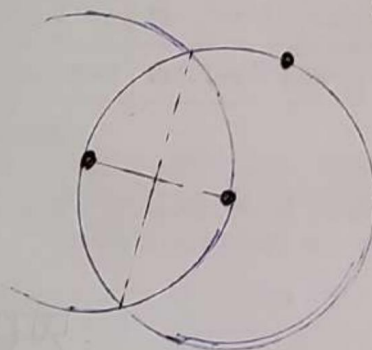
Estime un conjunto Foci de 4 puntos empleando el algoritmo Hull of Foci. De como respuesta la lista de puntos. Asuma h como la semilla del algoritmo.



$F_1: 14,76 - 4,76 = 10,00$
 $F_2: 14,76 - 8,54 = 6,22$
 $F_3: 14,76 - 10,82 = 3,94$
 $F_4: 14,76 - 7,07 = 7,69$
 $F_5: 14,76 - 9,90 = 4,86$
 $F_6: 14,76 - 12,04 = 2,72$
 $F_7: 14,76 - 10,20 = 4,56$
 $F_8: 14,76 - 11,46 = 3,30$
 $F_9: 14,76 - 10,05 = 4,71$
 $F_{10}: 14,76 - 7,28 = 7,48$
 $F_{11}: 14,76 - 3,49 = 11,27$
 $F_{12}: 14,76 - 6,00 = 8,76$
 $F_{13}: 14,76 - 4,12 = 10,64$
 $F_{14}: 14,76 - 5,83 = 8,93$
 $F_{15}: 13,65 + 14,76 = 28,41$
 $F_{16}: 10,42 + 14,76 = 25,18$
 $F_{17}: 13,96 + 14,76 = 28,72$
 $F_{18}: 11,36 + 14,76 = 26,12$
 $F_{19}: 11,36 + 14,76 = 26,12$
 $F_{20}: 13,49 + 14,76 = 28,25$

Pregunta 7 (2 puntos)

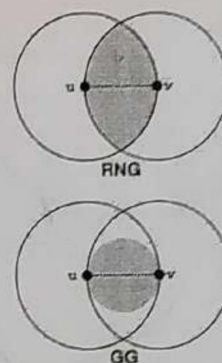
Proponga un algoritmo que estime la circunferencia mínima que contenga tres puntos $p_1: (a_1, a_2)$, $p_2: (b_1, b_2)$, $p_3: (c_1, c_2)$.



calcular el punto de la mediatriz de p_1 y p_2 . Luego calcular la mediatriz de p_1 y p_3 . Entonces m_2 es el centro del círculo mínimo que contiene a p_1, p_2 y p_3 . con radio máximo de m_2 a alguno de los puntos.

Pregunta 8 (3 puntos)

Sean los puntos dos puntos u y v , muestre que si están conectados en Relative neighborhood graph (RNG), se garantiza que la región sombreada (figura superior) no contiene ningún punto. Demuestre lo mismo para Gabriel graph (GG, figura inferior).



Pregunta 9 (3 puntos)

Al aplicar MinHash a las cadenas de caracteres T_1 y T_2 se obtienen las *signatures* s_1 y s_2 , los cuales son números enteros en el rango $[0, n - 1]$, donde n es la longitud del diccionario. Demuestre las siguientes propiedades:

1. $\text{MinHash}(T_1 + T_2) \leq \text{MinHash}(T_1) + \text{MinHash}(T_2)$
2. $\text{MinHash}(kT_1) \leq \text{MinHash}(T_1)$, donde $k \in \mathbb{Z}^+$

Pregunta 10 (1 punto)

¿Gabriel graph es navegable? Justifique.

Si porque se puede llegar a cualquier punto a otro.

Pregunta 11 (1 punto)

¿Cómo se define una red Small-world?

es aquella que esta muy dispersa en los puntos pero tiene una alta conectividad entre los puntos. es decir para llegar de un punto a otro el camino que se recorre es mínimo.

Pregunta 12 (3 puntos)

En las siguientes preguntas, marque la respuesta correcta:

- ¿Cuál de los siguientes modelos genera grafos planos?
 - ☒ a) Modelo de Kleinberg
 - b) Modelo de Watts Strogatz
 - c) Ninguna de las anteriores.
- El alcance del *shared memory* es:
 - a) *Threads* del mismo *block*.
 - ☒ b) *Threads* del mismo *grid*.
 - c) Ninguna de las anteriores.
- No es una propiedad de Sector Tree:
 - a) Se definen segmentos radiales con ángulos $\pi/2^k$.
 - ☒ b) Se definen segmentos de radio de igual longitud.
 - c) Los arcos almacenados en el árbol son definidos por una ecuación lineal sobre el ángulo del sector θ .
- No es un criterio para fusionar celdas en Grid File:
 - a) Cambio de granularidad.
 - b) La cantidad de archivos se reduce.
 - c) Eliminar claves inactivas.
 - ☒ d) Son criterios todas las anteriores.
- Dada la distancia Minkowski L_p , se define π_p como la relación entre el perímetro y el diámetro. Entonces:
 - ☒ a) $\pi_1 < \pi_2 < \pi_\infty$
 - b) $\pi_1 > \pi_2 > \pi_\infty$
 - c) $\pi_1 = \pi_2 = \pi_\infty$
- Sobre OS-Tree, es falso que:
 - a) OS-Tree poco eficiente para $d \geq 3$.
 - ☒ b) Probabilidad de fallo de Probably Correct OS-tree es aproximadamente igual a la relación entre el número de datos $|S|$ y el número de datos de entrenamiento $|T|$. Solo válido para $|T|$ grande.
 - c) Probably Correct OS-tree es útil una cantidad media de dimensiones (10-30).
 - d) OS-Tree emplea puntos adicionales (distintos a los datos) para determinar los hiperplanos paralelos que limitan las coberturas.