CS466 Final exam

Mingjun Zha

TOTAL POINTS

114 / 120

QUESTION 1

Problem 1 20 pts

1.1 2 / 2

✓ **- 0 pts Correct**

- 2 pts Incorrect

1.2 2 / 2

✓ **- 0 pts Correct**

- 2 pts incorrect

1.3 2 / 2

✓ **- 0 pts Correct**

- 2 pts Incorrect

1.4 2 / 2

✓ **- 0 pts Correct**

- 2 pts Incorrect

1.5 2 / 2

✓ **- 0 pts Correct**

- 2 pts Incorrect

1.6 2 / 2

✓ **- 0 pts Correct**

- 2 pts Incorrect

1.7 2 / 2

✓ **- 0 pts Correct**

- 2 pts Incorrect

1.8 2 / 2

✓ **- 0 pts Correct**

- 2 pts Incorrect

1.9 2 / 2

✓ **- 0 pts Correct**

- 2 pts Incorrect

1.10 2 / 2

✓ **- 0 pts Correct**

- 2 pts Incorrect

QUESTION 2

2 Problem 2 19 / 20

- 0 pts Correct

- 20 pts No answer provided

✓ **- 1 pts minor calculation mistake**

- 15 pts Wrong table structure

- 10 pts Correct table structure but wrong prob calculation and wrong inferred states

- 5 pts incorrect sequence of states and final probability

- 1 pts Minor backtracing mistake

QUESTION 3

3 Problem 3 20 / 20

✓ **- 0 pts Correct**

- 5 pts missed one or two labels

- 10 pts missed multiple labels

- 20 pts Did not provide an answer

QUESTION 4

4 Problem 4 20 / 20

✓ **- 0 pts Correct**

- 4 pts Cannot find second mapping

- 3 pts not separable second mapping

QUESTION 5

5 Problem 5 20 / 20

✓ **- 0 pts Correct**

- 1 pts Cluster center incorrect / minor mistake /

suspected calculation mistake

- **3 pts** Missing distance calculations for the first iteration
- **3 pts** Missing distance calculations for the second iteration
- **2 pts** Need to explicitly show cluster assignments (not just plot the points without labeling them)
- **1 pts** Minor mistake
- **8 pts** Second iteration not included
- **1 pts** Updated cluster centers not explicitly shown
- **4 pts** Incorrect new cluster centers

QUESTION 6

6 Problem 6 **15 / 20**

- **0 pts** Correct
- ✓ **- 5 pts** **Ignores multi-edges / somewhat incorrect construction**
- **8 pts** Constructs overlap graph (or anything using 4mers) instead
- **2 pts** Does not label multiedges
- **6 pts** Incorrect construction
- **3 pts** Does not find Euler path despite graph being correct
- **2 pts** Needs to explicitly write out reconstructed string
- **2 pts** Found Euler path does not correspond to the Graph

Problem 1

1. True

2. joint : $P(X, Y)$
   marginal : $P(X)$

3. classification is predicting the discrete task.
   Regression is the task to predict continous quantity.

4. True

5. False

6. False

7. True

8. False

9. True.

10. True.

1.1 **2 / 2**

✓ **- 0 pts** **Correct**

- 2 pts Incorrect

Problem 1

1. True

2. joint : $P(X, Y)$
   marginal : $P(X)$

3. classification is predicting the discrete task.
   Regression is the task to predict continous quantity.

4. True

5. False

6. False

7. True

8. False

9. True.

10. True.

1.2 **2 / 2**

✓ **- 0 pts** **Correct**

- 2 pts incorrect

Problem 1

1. True

2. joint : $P(X, Y)$
   marginal : $P(X)$

3. classification is predicting the discrete task.
   Regression is the task to predict continous quantity.

4. True

5. False

6. False

7. True

8. False

9. True.

10. True.

1.3 **2 / 2**

✓ **- 0 pts** **Correct**

- 2 pts Incorrect

Problem 1

1. True

2. joint : $P(X, Y)$
   marginal : $P(X)$

3. classification is predicting the discrete task.
   Regression is the task to predict continous quantity.

4. True

5. False

6. False

7. True

8. False

9. True.

10. True.

1.4 2 / 2

✓ **- 0 pts** Correct

- 2 pts Incorrect

Problem 1

1. True

2. joint : $P(X, Y)$
   marginal : $P(X)$

3. classification is predicting the discrete task.
   Regression is the task to predict continous quantity.

4. True

5. False

6. False

7. True

8. False

9. True.

10. True.

1.5 **2 / 2**

✓ **- 0 pts** **Correct**

- 2 pts Incorrect

Problem 1

1. True

2. joint : $P(X, Y)$
   marginal : $P(X)$

3. classification is predicting the discrete task.
   Regression is the task to predict continous quantity.

4. True

5. False

6. False

7. True

8. False

9. True.

10. True.

1.6 **2 / 2**

✓ **- 0 pts** **Correct**

- 2 pts Incorrect

Problem 1

1. True

2. joint : $P(X, Y)$
   marginal : $P(X)$

3. classification is predicting the discrete task.
   Regression is the task to predict continous quantity.

4. True

5. False

6. False

7. True

8. False

9. True.

10. True.

1.7 2 / 2

✓ **- 0 pts** **Correct**

- 2 pts Incorrect

Problem 1

1. True

2. joint : $P(X, Y)$
   marginal : $P(X)$

3. classification is predicting the discrete task.
   Regression is the task to predict continous quantity.

4. True

5. False

6. False

7. True

8. False

9. True.

10. True.

1.8 **2 / 2**

✓ **- 0 pts** **Correct**

- 2 pts Incorrect

Problem 1

1. True

2. joint : $P(X, Y)$
   marginal : $P(X)$

3. classification is predicting the discrete task.
   Regression is the task to predict continous quantity.

4. True

5. False

6. False

7. True

8. False

9. True.

10. True.

1.9 **2 / 2**

✓ **- 0 pts** **Correct**

- 2 pts Incorrect

Problem 1

1. True

2. joint : $P(X, Y)$
   marginal : $P(X)$

3. classification is predicting the discrete task.
   Regression is the task to predict continous quantity.

4. True

5. False

6. False

7. True

8. False

9. True.

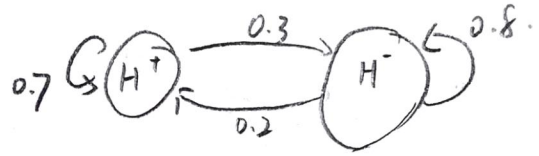
10. True.

1.10 **2 / 2**

✓ **- 0 pts** **Correct**

- 2 pts Incorrect

Problem 2.



i) NARNRDCCRN

	N	A	R	N	R	D	C	C	R	N
H^+	0	Max{(0), (0.2×0.1×0.2)} = 0.004	Max{(0.004×0.7×0.2), (0.048×0.2×0.2)} = 0.00192	Max{(0), 0.00768×0.2×0)} = 0	Max{(0), 0.00245×0.2×0.2)} = 9.8×10^{-5}	Max{(9.8×10⁻⁴×0.7×0.5), (3.9×10⁻⁴×0.2×0.5)} = 3.93×10^{-5}	Max{(5.5×10⁻⁶, 0)} = 5.5×10^{-6}	Max{7.71×10⁻⁷, 0.47×10⁻⁷} = 7.7×10^{-7}	Max{1.08×10⁻⁷, 7×10⁻⁹} = 1.08×10^{-7}	Max{(0), 501)} = 0
H^-	0.4×0.5=0.2	Max{(0), (0.2×0.8×0.3)} = 0.048	Max{(0.004×0.3×0.2), (0.048×0.8×0.2)} = 0.00768	Max{(0.00192×0.3×0.4), 0.00768×0.8×0.8)} = 0.000245 76	Max{(0), 10.00245 76×0.8×0.2)} = 3.9×10^{-4}	0	Max{1.18×10⁻⁶, 0} = 1.18×10^{-6}	Max{1.65×10⁻⁷, 0.94×10⁻⁷} = 1.65×10^{-7}	Max{4.6×10⁻⁸, 2.6×10⁻⁸} = 4.6×10^{-8}	Max{(1.3×10⁻⁸, 1.5×10⁻⁸)} = 1.5×10^{-8}

ii) The probability would be 1.5×10^{-8}.

iii) $H^-H^-H^-H^-H^-H^+H^+H^+H^+H^-$ → the most likely hidden sequence

2 Problem 2 19 / 20

- **0 pts** Correct
- **20 pts** No answer provided

✓ **- 1 pts minor calculation mistake**

- **15 pts** Wrong table structure
- **10 pts** Correct table structure but wrong prob calculation and wrong inferred states
- **5 pts** incorrect sequence of states and final probability
- **1 pts** Minor backtracing mistake

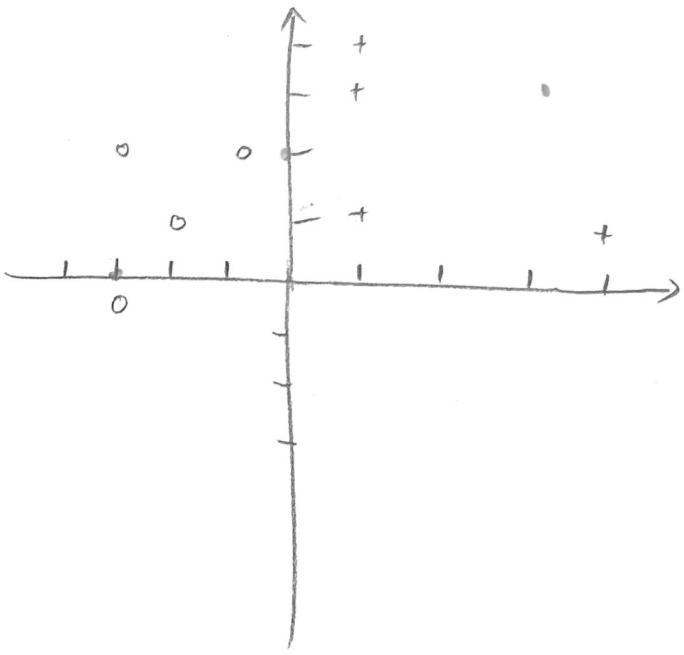
Problem 3

1)

x	y	label
-3	0	-
3	3	+
0	2	-

2)

x	y	Label
-3	0	-
3	3	+
0	2	+



3 Problem 3 20 / 20

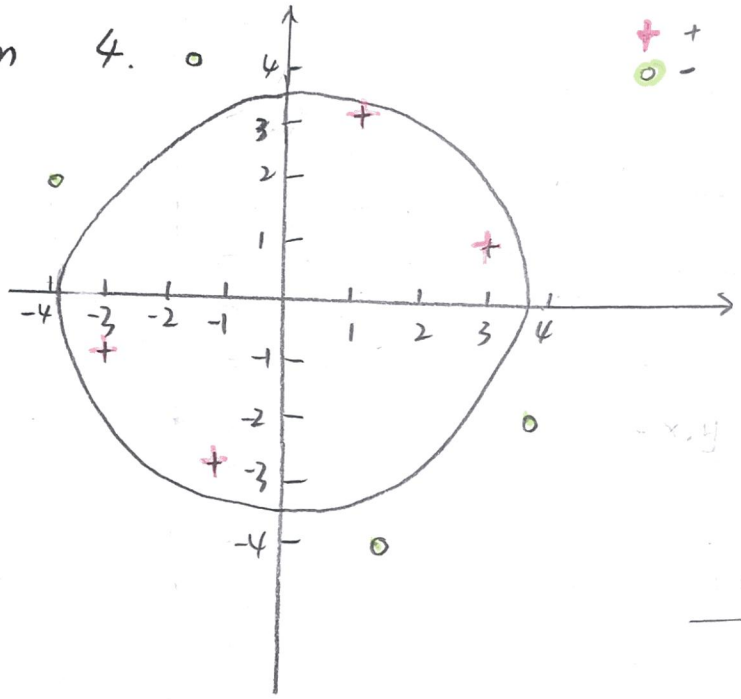
✓ **- 0 pts Correct**

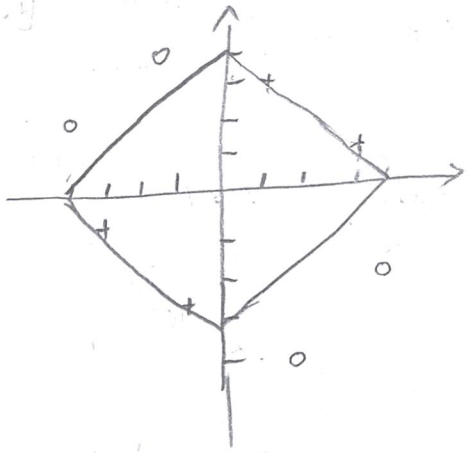
- 5 pts missed one or two labels

- 10 pts missed multiple labels

- 20 pts Did not provide an answer

Problem 4.





1: $(x, y) \to (x, y, x^2+y^2)$

2: $(x, y) \to (x, y, \frac{1}{2}xy)$

4 Problem 4 20 / 20

✓ - 0 pts Correct

- 4 pts Cannot find second mapping
- 3 pts not separable second mapping

Problem 5

$k = 2$ $\quad d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$

P_1: (0, 0)

P_2: (2, 2)

	P_1	P_2
(1, 0.5)	1.12	1.80
(2.5, 3)	3.91	1.12
(2, 1)	2.24	1
(3, 2)	3.61	1
(3.5, 2)	4.03	1.5
(-0.5, 0)	0.5	3.2
(-0.5, 1)	1.12	2.69
(-1, 0.5)	1.12	3.35
(1, -1)	1.41	3.16
(0.5, -1)	1.12	3.35

P_1: (1, 0.5), (-0.5, 0), (-0.5, 1), (-1, 0.5), (1, -1), (0.5, -1)

P_2: (2.5, 3), (2, 1), (3, 2), (3.5, 2)

↓

We repeated previous step, find P_1 & P_2 depending on mean.

P_1: $((1 - 0.5 - 0.5 - 1 + 1 + 0.5)/6,\ (0.5 + 0 + 1 + 0.5 - 1 - 1)/6) = (0.083, 0)$

P_2: $((2.5 + 2 + 3 + 3.5)/4,\ (3 + 1 + 2 + 2)/4) = (2.75, 2)$

We use the new P_1 and P_2, to find each distance.

	(0.083, 0) P_1	(2.75, 2) P_2
(1, 0.5)	1.04	2.30
(2.5, 3)	3.85	1.03
(2, 1)	2.16	1.25
(3, 2)	3.54	0.25
(3.5, 2)	3.96	0.75
(-0.5, 0)	0.583	3.82
(-0.5, 1)	1.16	3.4
(-1, 0.5)	1.19	4.04
(1, -1)	1.36	3.47
(0.5, -1)	1.08	3.75

P_1: (1, 0.5), (-0.5, 0), (-0.5, 1), (-1, 0.5), (1, -1), (0.5, -1)

P_2: (2.5, 3), (2, 1), (3, 2), (3.5, 2)

The predication is same as previous one, this algorithm is converge.

5 Problem 5 20 / 20

✓ **- 0 pts Correct**

- 1 pts Cluster center incorrect / minor mistake / suspected calculation mistake

- 3 pts Missing distance calculations for the first iteration

- 3 pts Missing distance calculations for the second iteration

- 2 pts Need to explicitly show cluster assignments (not just plot the points without labeling them)

- 1 pts Minor mistake

- 8 pts Second iteration not included

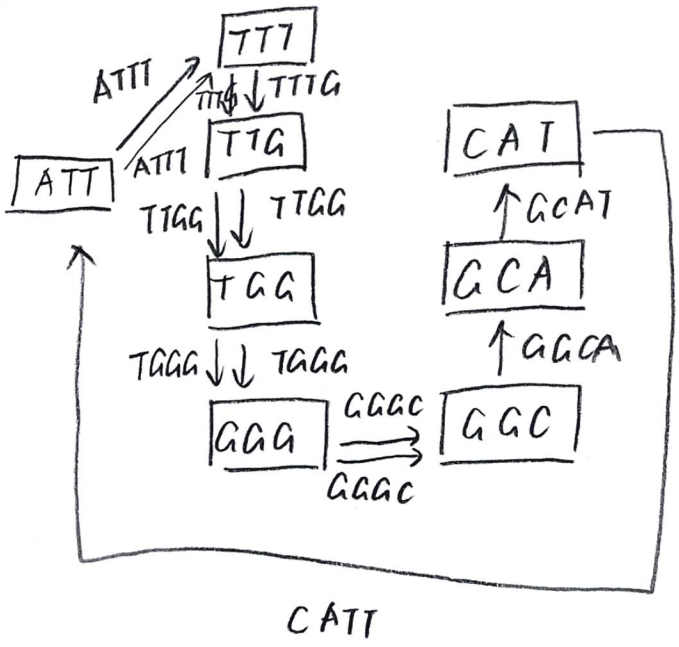
- 1 pts Updated cluster centers not explicitly shown

- 4 pts Incorrect new cluster centers

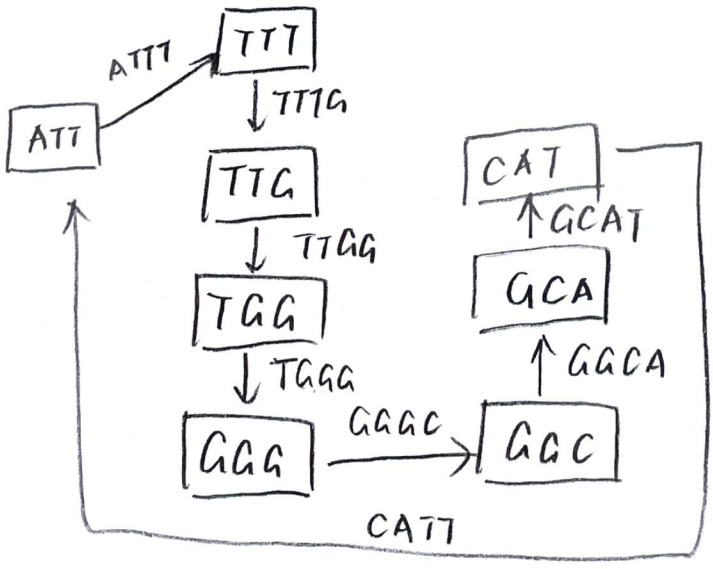
Problem 6 : String: ATTTGGGCATTTGGGC

a) ATTT
TTTG
TTGG
TGGG
GGGC
GGCA
GCAT
CATT

b)



c)



string : ATTTGGGCATT

d) No, they are not identical. Eulerian walk visits each edge exactly once. and keep them balanced The original one have so many duplicated nodes.

6 Problem 6 15 / 20

- **0 pts** Correct

✓ **- 5 pts** **Ignores multi-edges / somewhat incorrect construction**

- **8 pts** Constructs overlap graph (or anything using 4mers) instead
- **2 pts** Does not label multiedges
- **6 pts** Incorrect construction
- **3 pts** Does not find Euler path despite graph being correct
- **2 pts** Needs to explicitly write out reconstructed string
- **2 pts** Found Euler path does not correspond to the Graph