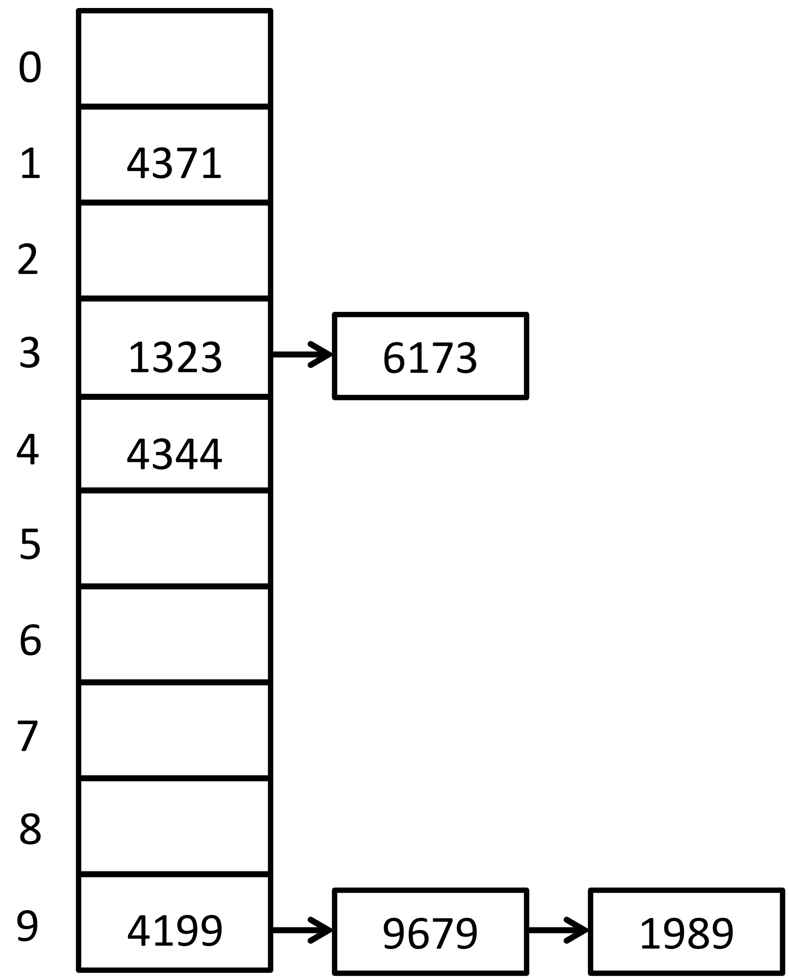
Solutions for Written Assignment 4 CSCI 2100A 2017 Spring

Exercise 4.1 (50 points)

(1)



(2)

0	9679
1	4371
2	1989
3	1323
4	6173
5	4344
6	
7	
8	
9	4199

(3)

0	9679
1	4371
2	
3	1323
4	6173
5	4344
6	
7	
8	1989
9	4199

(4)

0	
1	4371
2	
3	1323
4	6173
5	9679
6	
7	4344
8	
9	4199

(5)

Any reasonable answers are acceptable.

For example:

- The removal operations of open hashing are simple. The performance degrades more gracefully than the open addressing ones, especially when the load factor is high. However, it inherits the disadvantages of linked lists. When storing small records, the overhead of the linked list can be significant and the traversing a linked list may take a long time.
- Linear probing has greater locality and can provide good performance when the load factor is not high. However, it may

cause primary clustering (a tendency for one collision to cause more nearby collisions) when the load factor is high or the distribution of input is not uniform.

- Quadratic probing can avoid the problem of primary clustering, but it may still cause the problem of secondary clustering. Its performance degrades even more drastically than that of linear probing, as the load factor becomes higher.
- Double Hashing can relieve the problem of secondary clustering, but the performance is greatly affected by the choice of the second hashing function and requires more computation.

Exercise 4.15 (10 points)

O(N^2).

Exercise 6.1 (15 points)

(1) 34 pairs (the inversion pairs are omitted).

(2) $n(n-1)/2$ possible pairs.

(3) No. It depends on the sorting algorithm that we use. For insertion sort, it takes more time to sort A than B. However, for merge sort, it takes the same time to sort A and B.

Exercise 6.4 (15 points)

(1) O($NlogN$).

(2) O($NlogN$).

(3) O($NlogN$).

Exercise 6.6 (10 points)

[3	1	4	1	5	9	3	6	5	3	5]
[3	1	4	1	5	3	6	5	3	5]	9
[3	1	4	1	3	3]	5	[6	5	5]	9
[3	1	3	1	3]	4	5	[6	5	5]	9
[1	1]	3	[3	3]	4	5	[6	5	5]	9
1	1	3	3	3	4	5	5	5	6	9