**CSCE 4013/5013 Cloud Computing and Security**

**Quiz #5 (20 points)**

**Name:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_, ID:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Problem #1 (10 pts)**

1. Yes / No Do you participate in this quiz?

**Problem #2 (10 pts)**

1. (10 points) We use “pairs” approach to calculate the relative frequency of word pairs in two files.

|  |  |
| --- | --- |
| File 0 | File 1 |
| A B A B C A D  A C B C B A | B C B A C D  A D B A D C |

1. In-mapper combining is implemented.
2. There is only one reducer.
3. A word pair is a pair of words that are right next to each other.

Please list the input to the reducer.

Solution:

|  |  |
| --- | --- |
| **Key** | **Value** |
| (A, \*) | [7, 5] |
| (A, B) | [4, 2] |
| (A, C) | [2, 1] |
| (A, D) | [1, 2] |
| (B, \*) | [8, 5] |
| (B, A) | [4, 2] |
| (B, C) | [4, 2] |
| (B, D) | [1] |
| (C, \*) | [6, 5] |
| (C, A) | [2, 1] |
| (C, B) | [4, 2] |
| (C, D) | [2] |
| (D, \*) | [1, 5] |
| (D, A) | [1, 2] |
| (D, B) | [1] |
| (D, C) | [2] |