**CS 63018 & CS 73018 Probabilistic Data Management**

**Homework 3**

**Instructor:** Xiang Lian

**Due Date:** Please refer to the course website

1. Please list two (2) types of probabilistic queries over uncertain data. Give the formal problem definitions and discuss its real applications. **[20 points]**

2. Discuss the challenges of answering probabilistic queries over uncertain data. **[10 points]**

3. Please refer to the reading materials about the index on the course website and list two (2) types of indexes over multi-dimensional spatial objects. **[20 points]**

4. Please read the lecture slides of Chapter 6 (object-level pruning for join on uncertain data streams (USJ)), and explain what is the pruning condition of the object-level pruning (as shown in the example of Figure 1). Please prove that this object-level pruning is correct. **[20 points]**



**Figure 1.** Object-Level Pruning

5. Please read the abstract, introduction, and problem definition of the following paper. Then, please write a summary of this paper (including motivation, challenges, and problem definition). **[30 points]**

Xiang Lian and Lei Chen. Efficient Processing of Probabilistic Reverse Nearest Neighbor Queries over Uncertain Data. In *Very Large Data Bases Journal* (VLDBJ), 18(3), pages 787-808, 2009. <https://dl.acm.org/citation.cfm?id=1553332>

**Bonus Question** Describe the probabilistic pruning method for probabilistic reverse nearest neighbor in Question (5), and prove its correctness. [extra 20 points]

**Submission**

Submit an electronic copy of your homework solution to the [Blackboard](https://learn.kent.edu/).

18