

# Exercises and hand-in

Advanced database technology

March 20, 2006

We consider three of the multi-dimensional index structures covered in GUW:

- The grid file.
- The kd-tree.
- The R-tree.

These structures are often used in practice as they are simple and typically work well. The aim of this exercise is to investigate in what situations (i.e., for what kind of data) these index structures behave poorly. For any data set, there are many possible choices for the index (which grid lines to use for grid files, and which rectangles to use for R-trees). We seek data that is bad *no matter* how these choices are made. To simplify, we will consider the 2-dimensional case, and static data structures (no updates).

Consider nearest neighbor queries for a point set that is distributed *almost* exactly along a circle.

- How will each of the index structures represent such a set?
- How will they cope with a query for the nearest neighbor of a point almost exactly at the center of the circle?<sup>1</sup>

**The following exercises are to be handed in at the latest by April 3**

1. Exercise 18.2.5 from GUW. [Hint: Use a third transaction to force the order of the two transactions (in any serial schedule) required in (iii).]
2. Exercise 14.3.3 from GUW.

---

<sup>1</sup>A much better solution in this case is to use the planar point location data structure you saw in the lecture. This is left as an exercise.