Query Optimization  
7. Exercise, Summer 2009

Due 2009-06-10.

1. Implement a program that measures the physical characteristics (in particular bandwidth and seek time) of a hard disk. Measure the values for your own hard disk.

2. Given these values, determine the break even point where a scan becomes cheaper than an index lookup.

3. We assume the query optimizer has underestimated the costs of a scan by a factor $\alpha$ ($\alpha > 1$) and overestimated the costs of an index lookup by a factor $\beta$ ($\beta < 1$). What is the maximum error caused by this misestimation?

4. Assume the following kinds of SQL queries are executed frequently. Which index structures would you propose?

   (a) `select e.room`  
       `from employee`  
       `where e.id=17`

   (b) `select m.plot`  
       `from movies m`  
       `where m.title="Metropolis" and m.year=1927`

   (c) `select e.name`  
       `from employee e, departments d`  
       `where e.did=d.id and d.building="123" and e.gender="F"`