1. [15 points] Explain why you might have to search multiple paths during a lookup in an R-tree. Your explanation should include an example that illustrates your point. Also, explain why this doesn’t happen in B-trees.
2. [15 points] In the DBMin algorithm,

(a) [5 points] In the “clustered sequential” access pattern, would it be better to use LRU or MRU? Explain your answer.
(b) [10 points] Describe (briefly!) the difference between locking records, latching data structures, and pinning pages in a DBMS.
3. [15 points] In the Stonebraker and Hellerstein data model paper, they discuss a number of problems with the hierarchical data model. Explain one of these problems. You can use examples from that paper if it helps.
4. [10 points] In the “Dangers of Replication” paper, which replication schemes suffer from excessive deadlocks? Which suffer from “system delusion”? 
5. [15 points] In the Mariposa system, there is no centralized authority. So how does the “system” decide if it should replicate a file?
6. [15 points] In the BUCKY benchmark the object-relational implementation uses a pair of references to encode the student <-> department relationship. Compare and contrast this to the relational approach of using key-foreign key to do the same thing.
7. [15 points] Consider a selection of the form

```
SELECT *
FROM R
WHERE R.a = c1 and R.b = c2
```

Joe 764 wants to evaluate this query, and he is trying to decide whether to use bitmap indices or B-tree indices on R.a and R.b. Write equations that would help him decide which to use. You do not need to go into extreme detail here; I am looking for equations that use constants such as $RIDSize$, $RIDCompTime$, $WordLength$, $BitwiseAndTime$, $NumAValues$, $NumBValues$. You may not need all of these, and you can define others if you like.