CSE 562: systems project (due December 9, 2005)

Use Java for project implementation.
Submit all work in electronic form.

Part A: Indexing

You will implement a hashed file (H-file) in main memory, using linear hashing (textbook, p. 368). The following operations need to be supported:

- create an H-file with a given record type;
- save an H-file to disk;
- restore an H-file from disk;
- insert a record into an H-file;
- retrieve a record from an H-file given a key value;
- retrieve all records from an H-file, one by one;
- display the contents of an H-file.

Assumptions:

1. every record type consisting of variable-length character string fields should be supported;
2. each record fits in one block;
3. one field is designated as the key;
4. file blocks are implemented as independently allocated main memory fragments of size 512 bytes;
5. if the extra credit problem of Part B is not attempted, it is enough to implement a single H-file. Its record type will be determined by the parameters of the create file operation.

Part B: query processing

You are given a subset of SQL2 consisting of:

- relation definitions of the form:

  `CREATE TABLE` *Table*(
  A_1 `VARCHAR2`(n_1)
  ...
  A_k `VARCHAR2`(n_k))

Additionally, one attribute is designated as `PRIMARY KEY`.

- `DROP TABLE` *Table*.
- SQL2 single-tuple `INSERT` commands.
- the `@File` command to execute an .sql file consisting of the SQL commands listed here.
- queries of the following form:

  `SELECT` *List of Attributes*
  `FROM` *Table*
  `WHERE` *Condition*.

  The `WHERE` *Condition* is a conjunction of atomic equality conditions of the form $A = c$ or $A = B$ where A and B are attributes and c a constant.

You are supposed to implement:

1. a parser for the input
2. a query evaluator for the above subset of SQL2 that performs one of the following does:
   - selection using hashing (if the query contains an equality condition on the key);
   - selection using scanning (in other cases).

The selection should be followed by a final projection.

Extra credit

You may get up to 5% of the final grade for implementing:

- multiple H-files and multiple range variables, and
- hash-join and join conditions.