CIS4301 Notes: Final Review

Ryan Roden-Corrent

April 22, 2014

1 HTML

- Forms
    - Submit form elements to a script
    - GET: parameters in url, should not modify DB
    - POST: parameters embedded in request, use for inserts on DB

2 URLS

- pass key value pairs to url after the **?** token
    - `some.php?key1=val1&key2=val2`
- Use **#** token to reference specific part of page
    - `some.php#somepageid`

3 JDBC

Example Questions
- What does `Class.forName("org.mysql.Driver")` do?
- What is a cursor?
    - Acts like an iterator to look at query results in succession

3.1 Prepared Statements

What are they, what are the benefits?

Faster already partially compiled

Safer knows what times to accept, protect against injection

4   Analyze and Explain

ANALYZE runs query and shows performance stats

EXPLAIN Provides info on query plan

5   K Means

Used to identify clusters in data, which can grant insight into distinct groups that the data represents.

1. Choose a K (number of clusters)

2. Determine centroids

3. Assign each point to one of the centroids

4. Average points for each centroid

5. Determine new centroids

6. Repeat until centroids no longer move

Can perform K Means multiple times with different values of K to determine which K value is most appropriate. Could use a mesurement like average ditance between points and centroids to estimate the correctness of a K value.

6   MapReduce

1. Map - Apply function to collection

2. Shuffle - Assign map results to reduce machines

3. Reduce - Apply function to a collection to reduce into accumulator

4. Combine - Merge reductions into final result