

Spring, 2008 CIS 455 / 555

Internet and Web Systems

Final Examination

120 minutes, 8 questions, 120 pts

Try to allocate no more than 1 minute per point, and be concise with your answers.

Your name: _____

1. For the PageRank algorithm (improved version), (a) explain its intuition in terms of the authoritativeness; (b) explain it with respect to the “random surfer” model; (c) specify how it is computed in matrix form (including how the matrices are created). (20pts)

2. (a) BRIEFLY explain what strictness means and its importance in isolation. (5pts)

(b) Provide an example in which locks are used, but not acquired and released according to the two-phase convention, which results in a lack of isolation. (5pts)

- 3. Explain how a master (coordinator) node in two-phase commit goes through the recovery process (consulting its logs) if it crashes in the middle of a vote. (10pts.)**

- 4. Outline the steps for taking a keyword query of the form “k1 k2” and returning answers in the document vector model based on TF/IDF. (20pts.)**

5. **Write the map and reduce functions that, given a distributed list of documents, count all documents with names starting with “c,” which include at least ten keywords starting with “k”. (10pts.)**

- 6. Explain the roles of the principals (nodes) and the authentication steps in the Kerberos protocol. (20pts.)**

7. **Provide pseudocode showing how one would use Pastry to GET a value by key. Describe the failure cases and how Pastry handles them (if at all). (10pts.)**

8. Given an XML version of an inverted index over English documents, `inverted.xml`, of the form:

```
<index>
  <entry><word> ... </word>
    <foundIn>...</foundIn>*
  </entry>*
</index>
```

where *s mean 0 or more elements, and a table of French word translations, `french.xml`, of the form:

```
<english-french>
  <word><english>...</english><french>...</french></word>*
</english-french>
```

Write an XQuery that takes the word “université” and returns the set of corresponding entries. (20pts)