Quiz 10 Solution
NoSQL Presentation

1. Write a relational algebra expression to find the libnum of libraries that house a copy of all books written by Anton Chekhov.

$$-(\pi_{\text{libnum, booknum}}(Co)) / (\pi_{\text{booknum}}(\sigma_{\text{first = 'Anton' and last = 'Chekhov'}} (W \bowtie A)))$$

2. Suppose the Copy table contains 50,000 rows and there are 100 rows per page. Also suppose there are 50 main memory page buffers that can be used for sorting.

a. How many secondary memory accesses (reads and writes) will be done in order to sort the Copy table? Use the external sorting algorithm (i.e. partial sort then k-way merging) described in the lecture on external sorting.

b. How many merge phases are required to complete the sort?

Total Pages: 50000/100 = 500 pages
Partial Sort: Do the following 10 times; this creates 10 sorted sequences
           Read 50 pages, sort, write 50 sorted pages

       (10*50) reads + (10*50) writes = 1000 secondary memory accesses

Merge Phase: Merge 10 sorted sequences into a single sorted sequence
        Since this only requires 11 buffers only **one merge phase is required**
        500 reads + 500 writes = 1000 secondary memory accesses

Partial Sort + One Merge Phase = 1000+1000 = 2000 secondary memory accesses

Directly using the formula: $2*500*\text{ceiling}(\log_{49} 500) = 1000 * 2 = 2000$ secondary memory accesses

Extra Credit NoSQL Presentation
Due 11:59 PM Friday May 7

- 20-30 minute video + slides

- Upload one zip file that contains the video and pdf of the slides to Canvas

- Undergraduates can work in pairs; graduate students must work alone

- This is an optional assignment