**Lab 3 -- Questions & Instructions (Sep 9, 2021)**

Each question weighs 10 points, and you are required to attempt all questions…

Answers need to be succinct and in your own words… Verbosity is undesirable…

Total points = 150 (+ Bonus: 30 points)

Your submission will be a singular pdf file <Naming convention Roll\_No.\_FirstName\_Lab3.pdf> - containing theoretical answers, and code with outputs and observations

Deadline for submission: Sep 12, 2021.

1. Consider the following partial table of an ordered library catalogue:

|  |  |  |  |
| --- | --- | --- | --- |
| **Author\_ID** | **Book\_ID** | **Author\_Name** | **Book** |
| Da\_001 | Da001\_Sel | Damasio | Self Comes to Mind |
| To\_015 | To015\_Fel | Tolkien | Fellowship of the Rings\_Lord of the Rings |
| Mi\_009 | Mi009\_Emo | Minsky | Emotion Machine |
| Mi\_009 | Mi009\_Soc | Minsky | Society of Mind |
| Ra\_001 | Ra001\_Pha | Ramachandran | Phantoms in the Brain |
| Ro\_015 | Ro015\_Fan | Rowling | Fantastic Beasts and Where to Find Them |
| Ro\_015 | Ro015\_Gob | Rowling | Goblet of Fire\_Harry Potter |
| Ro\_015 | Ro015\_Pri | Rowling | Prisoner of Azkaban\_Harry Potter |
| Sa\_001 | Sa001\_Wha | Safina | What Animals Think |
| Wo\_015 | Wo015\_Wod | Wodehouse | Wodehouse at the Wicket |

**Note:**

* It is known that most user queries involve the Author\_Name and/or Book\_Name
* Typing the first three characters <xxx> of a search value for either of the attributes, on the library exploration portal, is sufficient to retrieve a set of records that begin with <xxx>.

1. For the given table: What are the potential candidate keys, and which one would you use as the primary key and why? **(10 points)**
2. How would you use extendible hashing with bucket size 4 to design an effective access strategy for the given scenario? **(10 points)**
3. Do you think a larger bucket size would be more effective? Experiment with at least one smaller and at least one larger bucket size to justify your claim. **(10 + 10 = 20 points)**
4. How would you use linear hashing with a bucket size of 4 to design an effective strategy for the given scenario? **(10 points)**
5. Do you think a larger bucket size would be more effective? Experiment with at least 1 smaller and at least one larger bucket size to justify your claim. **(10 + 10 = 20 points)**
6. For a new record: <Sa\_001, Sa001\_Voy, Safina, Voyage of the Turtle>: **(10 + 10 + 10 = 30 points)**
   1. Trace the insertion into the extendible hash bucket of Q.a
   2. Trace the insertion into the linear hash bucket of Q.c
   3. Write a short 150-200 words note on your observations between the above two insertions and subsequent time of retrieval for the record
7. For the record <Ro\_015, Ro015\_Phi, Rowling, Philosopher’s Stone\_Harry Potter>: **(10 + 10 + 10 = 30 points)**
   1. Trace the insertion into the extendible hash bucket of Q.e
   2. Trace the insertion into the liner hash bucket of Q.e
   3. Write a short 150-200 words note on your observations between the above two insertions and subsequent time of retrieval for the record
8. Can you think of a better way to encode the author\_id and book\_id - with respect to the format followed in the given table? Why do you think your code is better? **(10 + 10 = 20 points)**

**\*\*\*\***

**Bonus: Experiment resolution of collisions and retrieval of the record of Q.g. for the extendible hash bucket of Q.g using linked lists, B-trees of order 4, and B+- trees of order 4. (30 points)**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_