

**CS585: Big Data Management**  
**(MongoDB Project 2)**  
**(Done in teams)**

**Total Points:** 100 (10 each question)

**Release Date:** 04/14/2017

**Due Date:** 04/29/2017 (midnight 11:59pm )

**References:** Lecture notes & MongoDB Manual

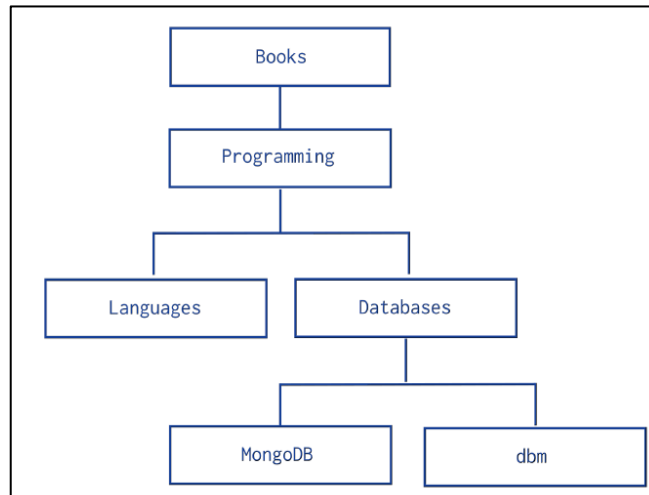


Figure 1: Tree Structure Relationships

### Question 1

1) Assume we model the records and relationships in Figure 1 using the Parent-Referencing model (Slide 49 in MongoDB-2). Write a query to report the ancestors of “MongoDB”. The output should be an array containing values `[{Name: “Databases”, Level: 1}, {Name: “Programming”, Level: 2}, {Name: “Books”, Level: 3}]`

*\* Note: “Level” is the distance from “MongoDB” node to the other node. It should be computed in your code*

2) Assume we model the records and relationships in Figure 1 using the Parent-Referencing model (Slide 49 in MongoDB-2). You are given only the root node, i.e., `_id = “Books”`, write a query that reports the height of the tree. (It should be 4 in our case).

3) Assume we model the records and relationships in Figure 1 using the Child-Referencing model (Slide 54 in MongoDB-2). Write a query to report the parent of “dbm”.

4) Assume we model the records and relationships in Figure 1 using the Child-Referencing model (Slide 54 in MongoDB-2). Write a query to report the descendants of “Books”. The output should be an array containing values `[“Programming”, “Languages”, “Databases”, “MongoDB”, “dbm”]`

5) Assume we model the records and relationships in Figure 1 using the Child-Referencing model (Slide 54 in MongoDB-2). Write a query to report the siblings “Databases”.

## Question 2

As you did in MongoDB Project 1, Create a collection named “test”, and insert into this collection the documents found in this link (10 documents): <http://docs.mongodb.org/manual/reference/bios-example-collection/>

- 1) Write an aggregation query that groups by the award name, i.e., the “award” field inside the “awards” array, and reports the count of each award. (Use Map-Reduce mechanism)
- 2) Write an aggregation query that groups by the birth year, i.e., the year within the “birth” field, and report an array of `_ids` for each birth year. (Use Aggregate mechanism)
- 3) Report the document with the smallest and largest `_ids`. You first need to find the values of the smallest and largest, and then report their documents.
- 4) Use the `$text` operator to search for and report all documents containing “Turing Award” as one sentence (not separate keywords).
- 5) Use the `$text` operator to search for and report all documents containing either “Turing” or “National Medal”.

### **Submission Mechanism**

Submit all your statements in a single text file electronically using blackboard system.