# *Assignment #2*

1. [**Relational Algebra**] **(20)** Consider the following database schema:

   **Movies** (Title, Director, Actor);

   **Location** (Theater, Address, Phone number);

   **Schedule** (Theater, Title, Time).

   Express the following queries in relational algebra (select σ, project Π , Cartesian product X, join (theta-join))

   Q1: which theaters feature "Zootopia"?

   Q2: List the names and address of theaters featuring a film directed by Steven Spielberg.

   Q3: What are the address and phone number of the Le Champo theater?

   Q4: List pairs of actors that acted in the same movie. (* you want to use renaming on Movies and join the Movies with its copy Movie').

2. **[Join Operators] (30)** This sets of questions test the understanding of basic database search operators. Consider a join $\bowtie_{R.A=S.B}$. We ignore the cost of output the result, and measure the cost with the number of I/Os.

   Given the information about relations to be joined below:

   Relation *S* contains 20,000 tuples and has 10 tuples per block. Relation *R* contains 100,000 tuples and has 10 tuples per block. Attribute *B* is the primary key of *S*. In total, 52 blocks are available in memory. Assume neither relation has any index.

   a (10) Describe a block nested join algorithm, Give the cost of joining *R* and *S* with a block nested loops join.

b (10) Describe a sort-merge join algorithm. Give the cost of joining $R$ and $S$ with a sort-merge join.

c (10) Describe a hash-join algorithm. Give the cost of joining $R$ and $S$ with a hash join.

3. **[XML] (30)**

a. (10) Consider the attached data for a simple Construction Project Management company. Create XML documents (.xml files) to store the data. Provide the XML documents as separate files.

b. (10) Consider the XML documents defined in (a) above for the Construction Project Management dataset. Give a Document Type Definition (DTD) representation for each of the XML documents. Provide the DTD representations as .dtd files. How do you encode keys? Foreign keys?

c. (10) Consider the XML documents defined in (a) above for the Construction Project Management dataset. Give an XML schema representation for each of the XML documents. Provide the XML schema representation as .xsd files. How do you encode keys? Foreign keys?

4. **[JSON] (20)**

a. (10) Consider again the Construction Project Management company dataset attached with this assignment. Assuming you don't have any relational schema. How would you model the data to store it as JSON documents. Use the provided data to create the JSON documents. Provide the JSON documents as .json files.

b. (10) Consider the JSON data model and JSON documents defined in (a) for the Construction Project Management dataset. Give a JSON Schema that can be used to validate the JSON documents with sample data. Provide the JSON Schema as .json files.