186/200

Name: Richa Saraswat
ID#: 112997679

**CS 4513 - Database Management - Fall 2014 - Dr. Le Gruenwald**
**MIDTERM EXAM**
**(Time and Date: 1:30 PM – 2:45 PM, 10/22/2014; Maximum Score: 200 points; Closed books, notes and electronic devices)**

**1) (75 points):**
An organization needs a database to hold its current order information. An initial analysis of the requirements produced the below facts:

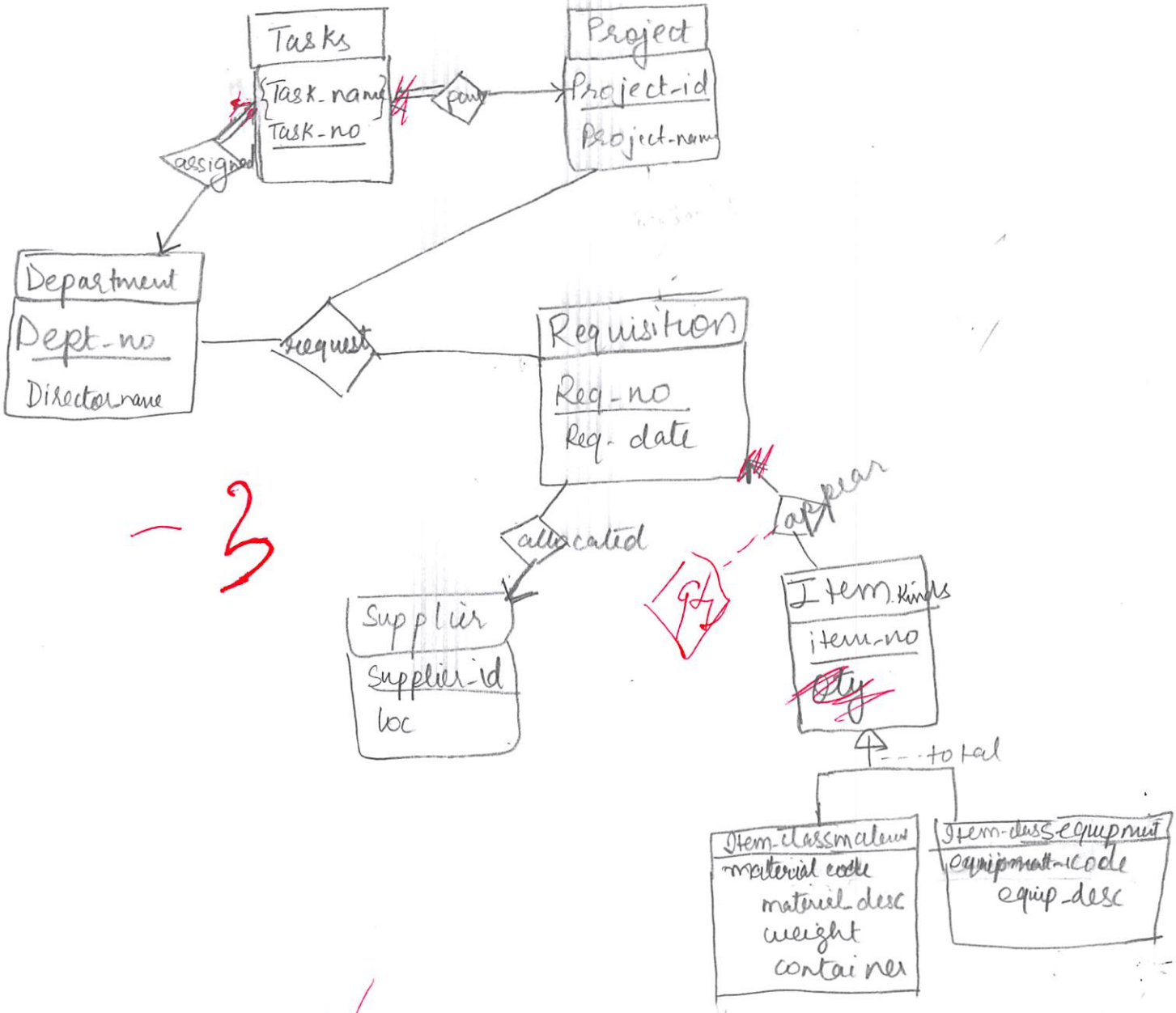
- Tasks are assigned task-names and unique task-numbers within the organization. Each such task is part of one project, and each task is assigned to one department. Each project has a unique project-id and a project-name. Each department is identified by its dept-no. The organization also records the director-name of each department.
- Requisitions are made of items. Each requisition is assigned a unique requisition-no and a requisition-date. A requisition contains the requesting department and the project for which the request is made.
- Each requisition is for the items for one project. The requisition is allocated to only one supplier who has a unique supplier-id and a location.
- Item kinds are identified by item-no; more than one item kind can appear in the same requisition.
- Any number of any item kinds may be requested in one requisition; qty is the number of a particular item kind in a requisition.
- Items are classified as being either item-class material or item-class equipment. Items of class material are allocated a material-code, and items of class equipment are allocated an equipment-code.
- Each material-code is described by a material-description, weight, and container. Each equipment-code is described by an equipment-description only.
  As an example, item-no = 77 is of class material and is given a material-code = 93, which in turn identifies the item as being cement, 200 lbs, paper-bag.

**Answer the following questions for problem 1:**
**1.1)** Construct an ER diagram that represents the database; **underline all primary keys. No additional attributes are allowed.**

**1.2)** How would your design change if the analysis of requirements also includes the following:
Some assignments of tasks to departments will result in a foreign contract agreement. Each foreign contract agreement has a cost, a unique agreement number, and a name that consists of a country name and city name.
To answer part (b), draw only the part of the ER diagram that reflects the changes (if any).

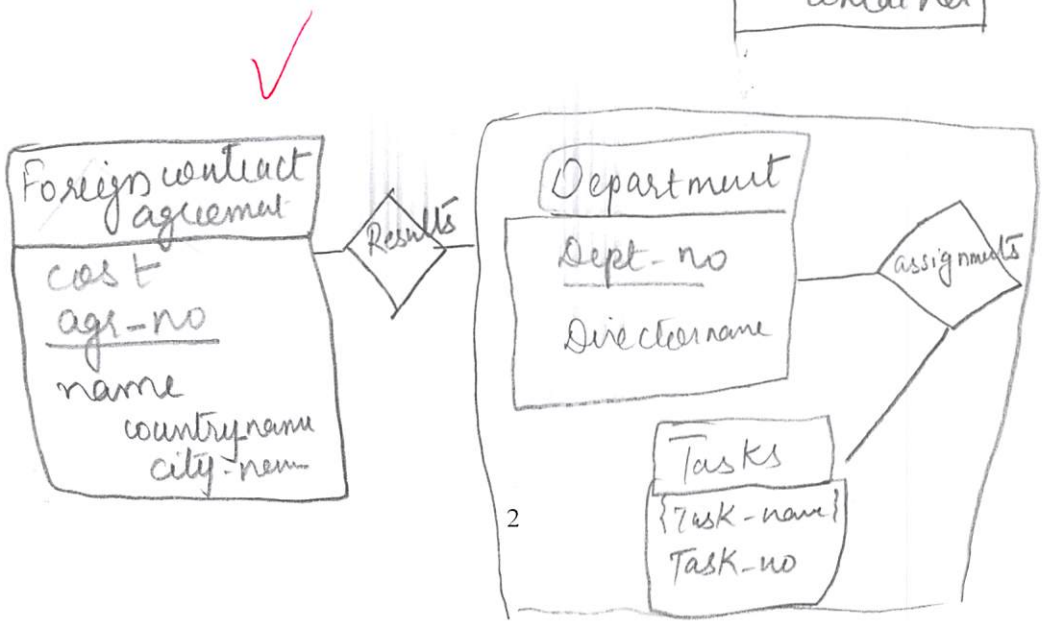
**(Workspace for problem 1)**

1.1)




-3

1.2)



**2) (15 points):** What are the causes of **bucket skew** in a hash file organization?

In hash file organization the new record is inserted at the end if there is no empty space available before. If the allocated bucket size is full, then bucket overflow is caused. The reasons of bucket skew are:

- if the total number of records stored ($n_r$) to the frequency of the record is less than total bucket-size allocated
- There is insufficient buckets allocated for all records

-6

**3) (20 points):** Johnson wants to store information (names, addresses, descriptions of embarrassing moments, etc.) about his employees. Not surprisingly, the volume of data compels him to buy a database management system. To save money, he wants to buy one with the fewest possible features, and he plans to run it as a stand-alone application on his PC. Of course, Johnson does not plan to share his list with anyone. The following questions (3.1) and (3.2) are about which DBMS features Johnson should buy. For each of the questions, circle the MOST CORRECT answer (if you circle more than one answers for a question, you will get a zero score for that question):

**3.1)** Johnson should buy the Concurrency Control feature

a) True

b) False

**3.2)** Johnson should buy the Query Processing feature

a) True

b) False

**4) (10 points):** Which of the following is among the ethical responsibilities of Website sponsors/operators? Circle the MOST CORRECT answer (if you circle more than one answers, you will get a zero score for this question):

a) They must take reasonable care to ensure the information they provide is accurate;

b) A site should have a mechanism to review content and filter out or remove dangerous material;

c) A site should make clear which information is supplied by users and what has, or has not, been verified;

d) Operators of sites that display material based on rankings or votes should anticipate manipulation and prepare to protect against it;

e) All of the above.

Student course section grade

**5) (30 points):** Given the following relational schemas:

Student(Name, StudentID, Classification, Major)
Course(CourseName, CourseNumber, CreditHours, Department)
Prerequisite(CourseNumber, PrerequisiteNumber)
Section(SectionIdentifier, CourseNumber, Semester, Year, Professor)
GradeReport(StudentID, SectionIdentifier, Grade)

Write an expression in SQL for each of the following queries:

**5.1)** Retrieve the names and majors of all straight A students (students who have a grade of A in all their courses).

(Select G1.SectionIdentifier = S.SectionIdentifier and S.course-no = C.course-no and Grade = "A")

-3 select name, major from student, gradeReport G1
where student.StudentID in (select distinct G1.studentID from
gradeReport "T" where T.studentID = G1.studentID and T.grade = "A")

**5.2)** Increase the number of credit hours by 1 for every course offered by the "History" department.

```
update course
    set creditHours = creditHours + 1
        where department = "History";
```

✓

**6) (50 points)**: given the following **sequential data file with the search key being project's name (each project has a unique id and a unique name):**

| Project (id, | name, | budget, | months, | description) |
|---|---|---|---|---|
| 10 | Cipher | 10,000,000 | 12 | terrorism data |
| 30 | Intro | 5,000,000 | 4 | data mining |
| 20 | Miles | 10,000,000 | 8 | transportation |
| 50 | Repair | 10,000 | 1 | Repair offices |
| 70 | Roads | 2,000,000 | 7 | New roads for Cleveland |
| 40 | Water | 5,000 | 4 | Water connection |
| 60 | World Trade | 1,000,000 | 5 | World trade agreement |

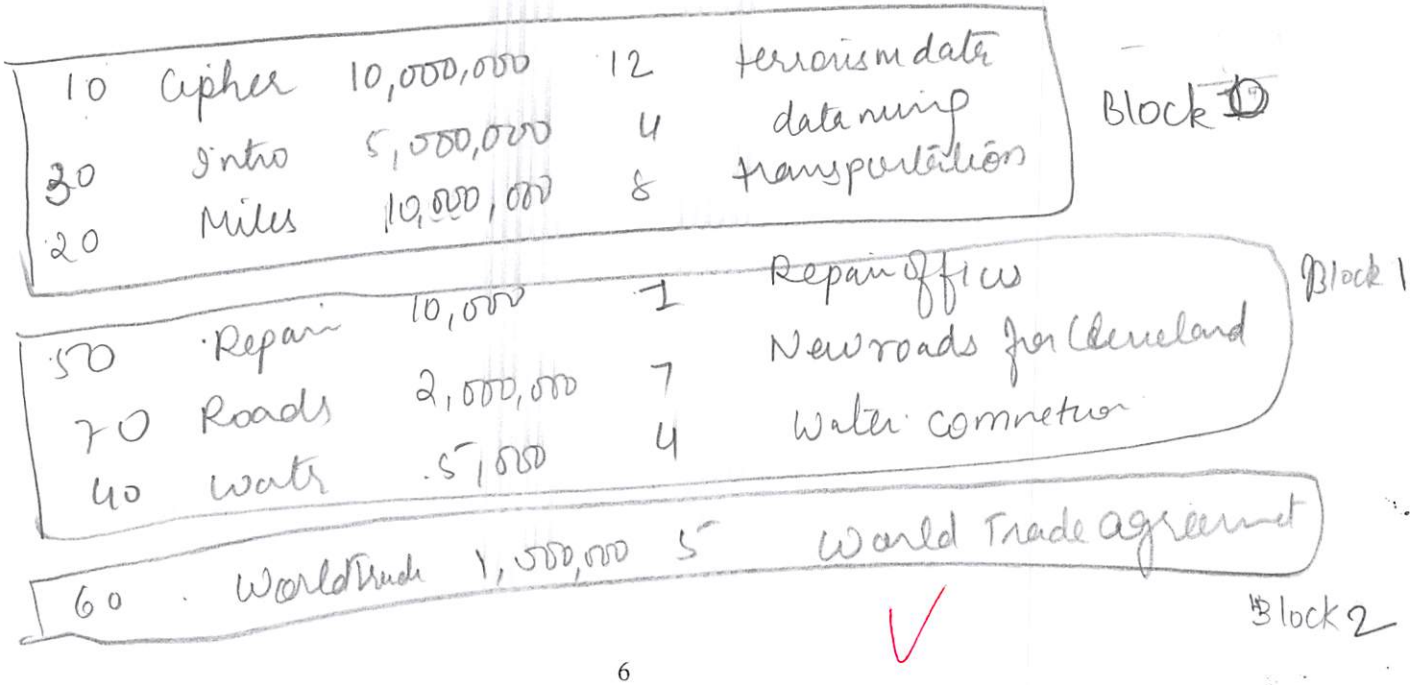
Assume that the following attribute sizes are used in the above data file: id: 4 bytes, name: 200 bytes, budget: 8 bytes, months: 4 bytes, and description: 250 bytes. Also assume that each disk block can store up to 1500 bytes, and no record spans two blocks, i.e. if a record does not fit into a particular disk block, it will be stored in the next disk block that has enough space.

Answer the following questions:

**6.1)** Show a primary index for the above data file using the implementation of indexed- sequential files discussed in class.

**6.2)** Show a B tree index on project's id with order 4 using the definition and insertion procedure of the B tree discussed in class. Show the tree after each insertion. Each tree node must have its complete information.

6.1) The above data can be stored as.



**(Additional Workspace for Problem 6)**

The primary index can be limited as

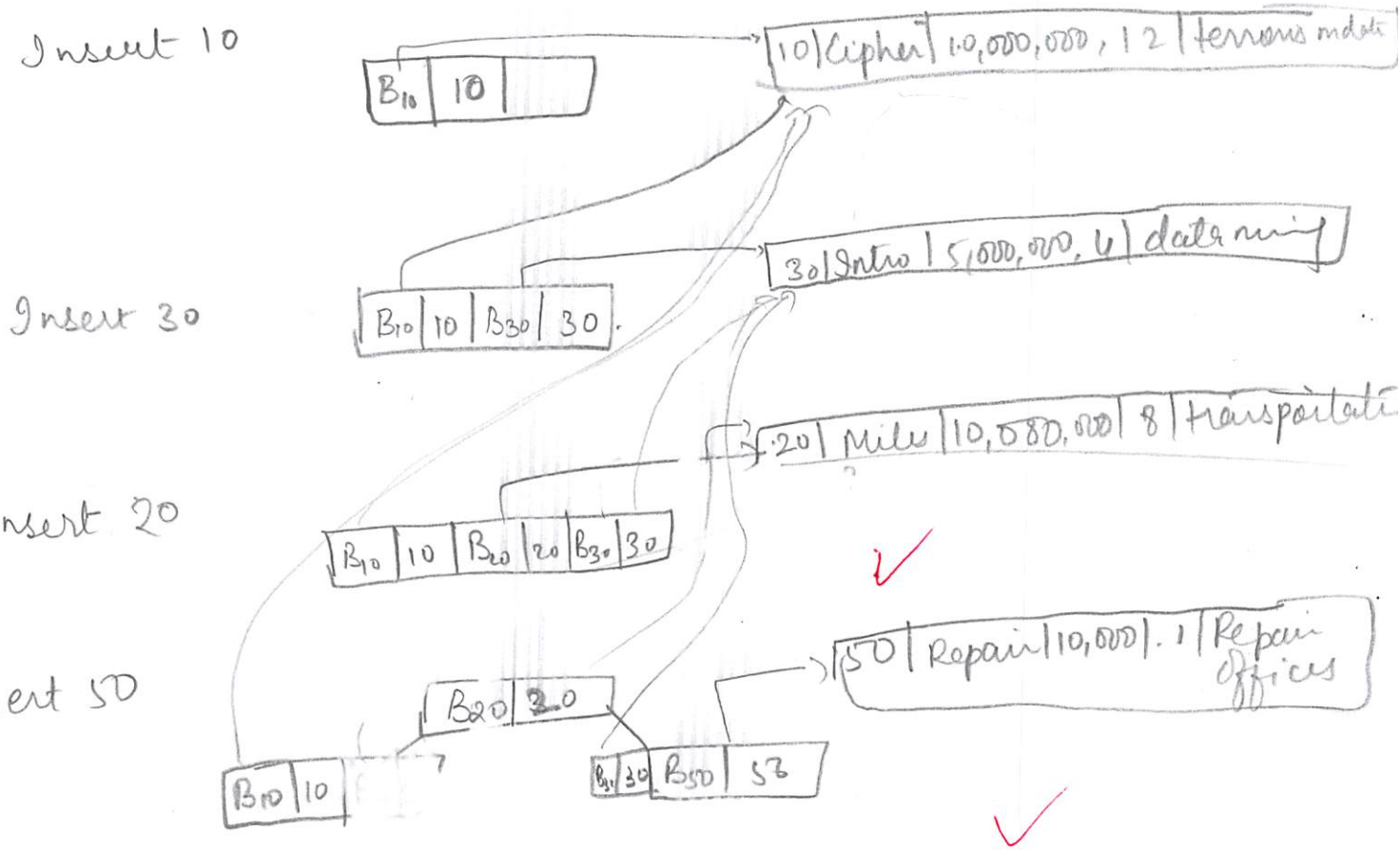
| Search key | address |
|---|---|
| Cipher | Block 0 |
| Repair | Block 1 |
| WorldTrade | Block 2 |

✓

6.2 id 10, 30, 20, 50, 70, 40, 60 — 2

Here $max = n-1 = 4-1 = 3$

$min = \lceil \frac{n}{2} \rceil - 1 = \lceil \frac{4}{2} \rceil - 1 = 1$




✓ ✓

**(Additional Workspace for Problem 6)**