Assignment 4 is 60 points.

Q 1-12: 4.5 points each

Q 13-14: 3 points each

Normalization

1. Code create statements for the 3 normalized tables candidate, contributor and contribution. Table candidate should have a primary key of cand_id. Contributor and contribution tables should have a surrogate key of int type defined as autoincrement. Contribution table should have columns for cand_id and contbr_id. Include your create table statement here.

```
create table candidate(
cand_id varchar(12) primary key,
cand_nm varchar(50)
);

create table contributor(
contbr_id int primary key auto_increment,
contbr_nm varchar(50),
contbr_city varchar(40),
contbr_st varchar(40),
contbr_zip varchar(20),
contbr_employer varchar(60),
contbr_occupation varchar(40)
);


create table contribution (
contb_id int primary key auto_increment,
cand_id varchar(12),
contbr_id int,
contb_receipt_amt decimal(8,2),
contb_receipt_dt varchar(20),
foreign key fcand_id (cand_id) references candidate(cand_id),
foreign key fcontbr_id (contbr_id) references contributor(contbr_id)
);
```

2. Code 3 insert statements using subselect (read “Inserting from a Query” page 185 in textbook) to select data from the campaign table and insert it into the normalized tables. You should have 22 rows in the candidate table, 14,174 rows in the contributor table, and 18,118 rows in the contribution table. Include your 3 insert statements here.

```
insert into candidate
select distinct cand_id, cand_nm
from campaign;

insert into contributor
select distinct null, contbr_nm, contbr_city, contbr_st, contbr_zip,
contbr_employer, contbr_occupation
from campaign;


insert into contribution
select null, cand_id, contbr.contbr_id, contb_receipt_amt,
contb_receipt_dt
from campaign c, contributor contbr
where c.contbr_nm = contbr.contbr_nm and
      c.contbr_city = contbr.contbr_city and
      c.contbr_st = contbr.contbr_st and
      c.contbr_zip = contbr.contbr_zip and
      c.contbr_employer = contbr.contbr_employer and
      c.contbr_occupation = contbr.contbr_occupation ;
```

NOTE: To find the value of contbr_id to insert into the contribution table, must join on a combination of name, city, state, zip, employer and occupation. Name alone does not uniqely identify the contributor.

3. Alter the contribution table to add foreign key constraints for columns cand_id and contrbr_id. Include your alter table statement here.

NOTE: If the foreign key constraints are not declared on the create table above, then declare them here.

```
alter table contribution add constraint foreign key (cand_id)
references candidate(cand_id);

alter table contribution add constraint foreign key (contbr_id)
references contributor(contbr_id);
```

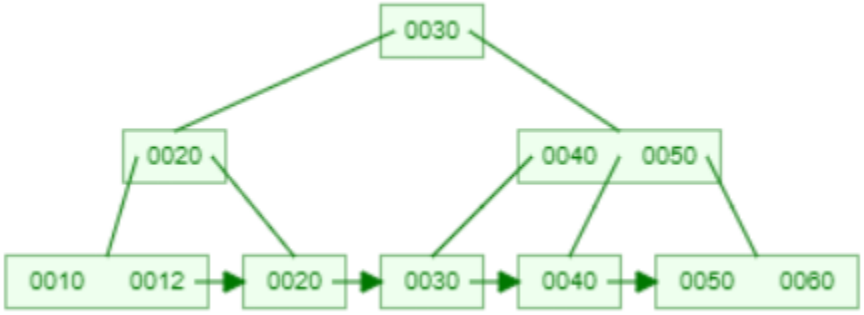
4. Create a view named “vcampaign” that is a join of the 3 normalized tables and has columns `cand_id, cand_nm, contbr_nm, contbr_city, contbr_st,`

```
   contbr_zip, contbr_employer, contbr_occupation,
   contb_receipt_amt, contb_receipt_dt

create view vcampaign as
select ca.cand_id, ca.cand_nm,
       cr.contbr_nm, cr.contbr_city, cr.contbr_st, cr.contbr_zip,
       cr.contbr_employer, cr.contbr_occupation,
       ct.contb_receipt_amt, ct.contb_receipt_dt
from contribution ct join candidate ca on ca.cand_id=ct.cand_id
    join contributor cr on ct.contbr_id=cr.contbr_id;
```

B+ Tree Visualization Exercises

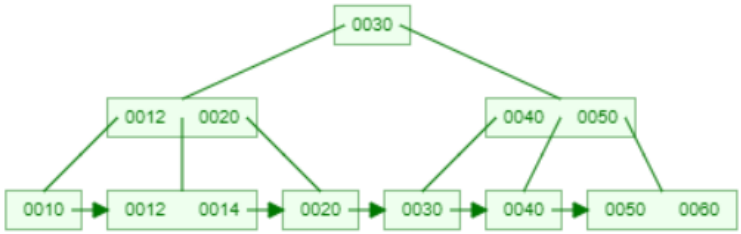
5. Do an insert of key value 12. Draw or embed a screenshot of the updated index.



6. How many nodes were either created or modified for the insert of 12?

   1 node (0010) was modified.

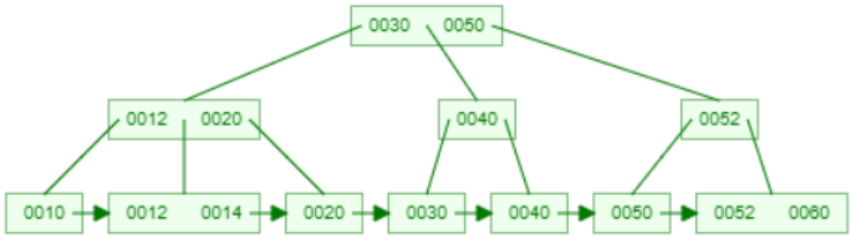
7. Now do an insert for a key value 14. Show an updated diagram.



8. How many nodes were either created or modified for an insert of 14?

Node (0010, 0012) was split to make room for 0014 and the parent node (0020) was modified to (0012, 0020). 1 new node and 2 changed nodes.

9. Do an insert of key value 52 and show an updated diagram.



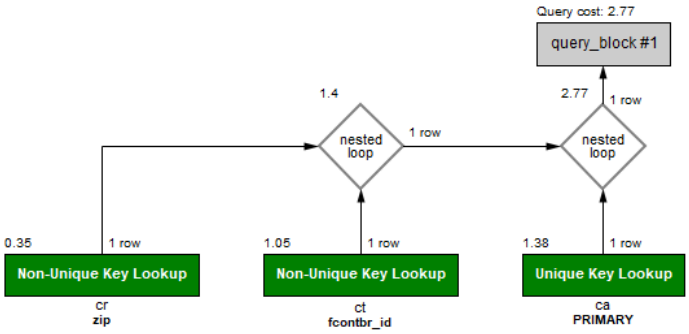
10. How many nodes were either created or modified for an insert of 52?

Leaf node (0050 0060) was split and the parent node (0040 0050) was also split. Root node was modified. 2 new nodes and 3 modified nodes.

Conclusion: insert, delete of a B tree index may involve several reads/writes.

Query Plan Exercises

11. Is the new index being used? Explain in your words the execution plan.



Yes, the zip index is used.

The index on contbr_zip is used to lookup rows for zip='93933' in contributor table. For each row returned, lookup on contribution table using index on foreign key contbr_id and join the rows. For each row lookup by cand_id primary index the row in candidate table.

Concurrency Exercises

Inconsistent Writes

12. Based on lecture material there are 2 ways to fix this problem. Pick one and test it out. How did you fix the problem?

```
Option 1:  set transaction isolation level SERIALIZABLE;

Option 2:  select * from duty for update;
```

Other Exercises

13. Consider this situation: you try to get cash at an ATM, but the ATM fails after updating your account and committing, but just before cash is dispensed. As a system designer, how do you cope with the situation that the money has been debited from the account and committed but the cash was unable to be dispensed? [hint: what do you think "compensating transaction" means? do a google search.]

When the cash cannot be dispensed, then a compensation transaction to deposit the amount back into the account is done.

14. Consider this situation: you try to buy an airline ticket at a web site. The transaction commits on the server, but crashes just before the message confirming the reservation is sent to the client. As a system designer, how would you cope with the situation of a reservation was made and committed in the database, but the confirmation message was never received by the client?

An email is also sent to the client with the confirmation number. In addition, the user can log back into the web site to verify that the reservation was made and the confirmation number.