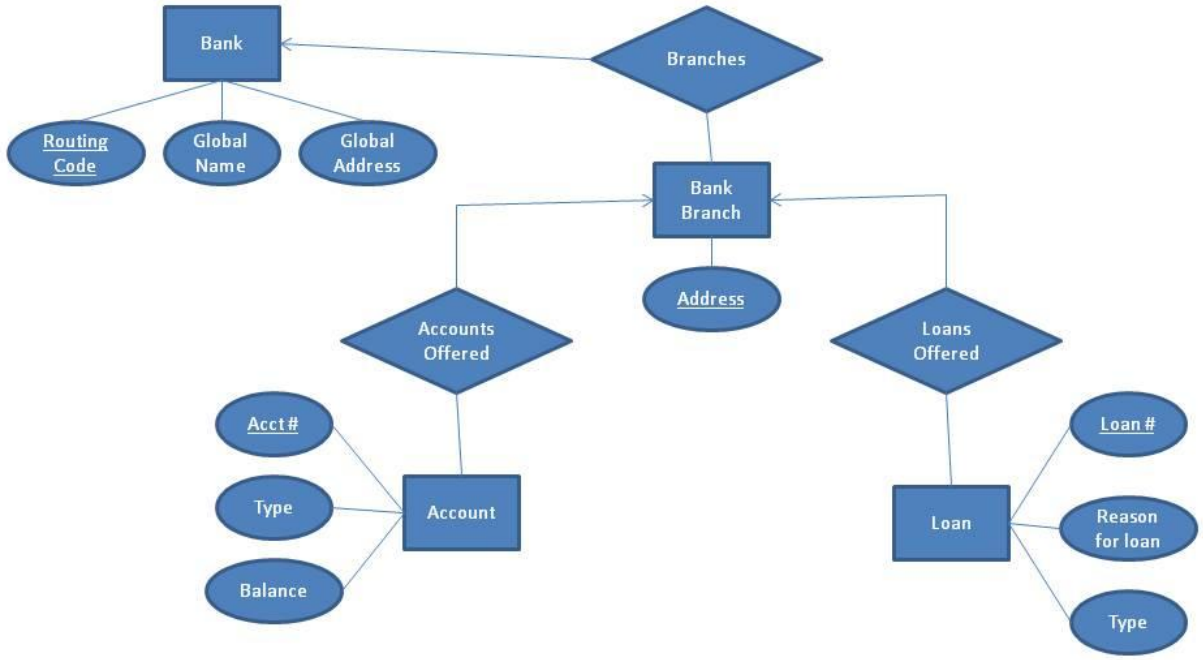
Ken Ford
11/3/17
CSCI 3287

Homework 2

1)



2)

a)

```
SELECT *
FROM R X, R Y
WHERE((X.A = Y.A AND X.B = Y.B) AND (X.C <> Y.C))
OR ((X.C = Y.C) AND (X.A <> Y.A OR X.B <> Y.B));
```

b)

```
SELECT *
FROM R X, R Y
WHERE(( X.A = Y.A AND X.B = Y.B AND X.C = Y.C)
AND (X.D <> Y.D OR X.E <> Y.E));
```

c)

```
SELECT *
FROM S X, S Y
WHERE((X.A = Y.A AND (X.B <> Y.B OR X.C <> Y.C))
OR (X.C = Y.C AND (X.A <> Y.A OR X.B <> Y.B));
```

3)

a)

Read in 160 pages in 4 page chunks = 40 page chunks

4 pages = 1 chunk = 1 IO

Therefore, 40 page chunks is equal to 40 IO for reading in the pages.

Pages must be written back out as well, where the cost of a write = 2(read)
160 pages written out x cost of 2 per page = 320 IO

In total then, we have 40 + 320 = 360 IO

b)

Cost formula = ceiling($(\log_{B-1}(N/B)+1)$)
= ceiling($\log_{19}(160/20) + 1$) = 2 passes

c)

With external merge sort, each time that we make a pass through the pages to sort the data, we end up merging the data. In our initial pass, this remains the case as the first pass of merging is done at the same time that the first sort pass was. With this in mind, the cost of the first pass of merging is what was calculated in part a, or 360 IO.

d)

The total cost of the external merge sort algorithm would be 360 + 360 = 720 IO, as each pass made through the data has an equal cost of 360 IO.