Kristopher Blair
Normalization Assignment

#1
Consider the following relations for an order-processing application database for Acme Products, Wile E. Coyote, owner. Determine whether the relations are in BCNF and, if not, decompose them.

```
Order(Order#, Order_date, Customer#, Total_amount)
Order_Item(Order#, Item#, Quantity_ordered, Price_each, Total_price,
Discount%)
```

Assume that each item has a different discount. The `Price_each` refers to one item, `Total_price` is the cost per item multiplied by the number of items purchased, `Order_date` is the date on which the order was placed, and the `Total_amount` is the amount of the order.

ANSWER:
The relations are not in BCNF. There is a partial dependency in the Order_Item relation.

- `Order(Order#, Order_date, Customer#, Total_amount)`
- `Order_Item(Order#, Item#, Quantity_ordered, Total_price)`
- `Item(Item#,Price_each,Discount)`

`This shows the correct BCNF relation for this data.`

#2
Determine whether the following relation is in BCNF and, if not, decompose it.

```
Rx(Doctor#, Patient#, Date, Diagnosis, Treatment_code, Charge)
```

A tuple describes a visit of a patient to a doctor along with a treatment code and a charge. Assume that each treatment code has a fixed charge (regardless of patient).

ANSWER:
The relation shown is not in BCNF. It should be as follows:

- Rx(Doctor#, Patient#, Date, Diagnosis, Treatment_code)
- Doc_Treatments(Doctor#, Treatment_code, Charge)

#3
Consider the relation `R(A, B, C, D, E, F, G, H, I, J)` and the set of functional dependencies `{AB→C, A→DE, B→F, F→GH, D→IJ}`. What is the key for `R`? Decompose `R` into 2NF and then 3NF.
Repeat the decomposition into 2NF and then 3NF starting with the original version of `R` and using the following set of functional dependencies instead: `{AB→C, BD→EF, AD→GH, A→I, H→J}`.

ANSWER:
The key for R is AB. Decomposing gives you the following:

- R(A, B, C)
- R_A(A, D, E)
- R_B(B, F, G, H)
- R_D(D, I, J)

Using the second set of functional dependencies, it decomposes to the following:

- R(A, B, C)
- R_BD(B, D, E, F)
- R_AD(A, D, G, H)
- R_A(A, I)
- R_H(H, J)

#4

Consider the relation `DISK_DRIVE(Serial_number, Manufacturer, Model, Batch, Capacity, Retailer)`. Each tuple in the relation `DISK_DRIVE` contains information about a disk drive with a unique serial number, made by a manufacturer, with a particular model number, released in a certain batch, which has a certain storage capacity, and is sold by a certain retailer. For example, the tuple

```
Disk_drive('1978619', 'Acme Drives', 'A2235X', '765324', 500, 'CompuMax')
```

specifies that Acme Drives made a 500GB disk drive with serial number 1978619 and model number A2235X; it was released in batch 765324 and is sold by CompuMax.

Translate each of the following into a FD:

1. The manufacturer and serial number uniquely identifies the drive.
2. A model number is registered by a manufacturer and therefore can not be used by another manufacturer.
3. All disk drives in a particular batch are the same model.
4. All disk drives of a certain model of a particular manufacturer have exactly the same capacity.

After stating the above FDs, decompose `DISK_DRIVE` into 3NF.

ANSWERS:

FDs

1. Serial_number, Manufacturer → Model, Batch, Capacity, Retailer
2. Manufacturer → Model
3. Batch → Model
4. Model → Capacity

Decomposition into 3NF:

- DISK_DRIVE(Serial_number, Manufacturer, Batch, Retailer)
- Manufacturer_Models(Manufacturer, Model)
- Batches(Batch, Model)
- Models(Model, Capacity)