

Use the ER diagram for the Instrument Rental and Repair database (attached) to answer the following questions.

1. For each entity in the diagram, tell whether a record will have a fixed length or a variable length and state the reason why you believe that to be.

- INSTRUMENT will be variable length because ShopUsed is a multivalued attribute, and many of the other attributes such as InstrumentName, Comments, Model, Maker, InstrumentName, LoanLocation, and LocationName could all be variable length strings.
- RENTAL will be variable length because Pymt\_Made is a multivalued attribute, and Authorized\_By could be a variable length string if it is based on a name.
- STUDENT will be variable length since S\_Inst# is a multivalued attribute and Name is a composite attribute.
- REPAIR is a variable length because R-Desc will likely be a variable length string.
- REPAIR\_SHOP is a variable length because ContactName is a multi-valued attribute, and S\_Address is likely to be a variable length string.

2. Calculate the record size R in bytes in the INSTRUMENT file. You will have to state your assumption for the data type and size of each attribute.

INSTRUMENT				
Attribute	Fixed/Variable/Multivalued	Size (bytes)	Format	Assumption
InstrumentName	Variable	1 - 30	variable character	Between 1 and 30 characters. Shouldn't be empty.
Serial Number	Variable	1 - 10	XXXXXXXXXX	Assume a minimum of 1 character and a maximum of 10 character
Condition	Variable	1 - 30	variable character	Between 1 and 30 characters. Shouldn't be empty.
FundsFrom	Variable	1 - 30	variable character	Between 1 and 30 characters. Shouldn't be empty.
Size	Variable	14 - 23	I.IIXw.wwXd.dd up to IIII.IIXwww.wwXdddd.dd	Size of the control volume of the instrument in inches. Always larger than 0.01 inch on a side, never larger than 1000.00 inches on a side.
Last_Inspection_Date	Fixed	10	mm:dd:yyyy	Will be inspected when purchased at a minimum.
Comments	Variable	0 - 200	variable character	0 - 200 characters.
Model	Variable	0 - 30	variable character	Between 0 and 30 characters.
Maker	Variable	1 - 30	variable character	Between 1 and 30 characters. Shouldn't be empty.
ShopUsed	Multivalued	0 - 35	empty up to XXXXXX,XXXXXX,XXXXXX,XXXXXX, XXXXXX,XXXXXX	Shop IDs are assumed to be 6 characters. There will likely never be more than 5 shops that can repair a particular instrument. The instrument may never have been repaired.
InstrumentYearMade	Fixed	4	yyyy	Never should be empty.
LoanLocation	Variable	0 - 30	variable character	Between 1 and 30 characters.
LocationName	Variable	1 - 30	variable character	Between 1 and 30 characters. Shouldn't be empty.
PurchaseDate	Fixed	10	mm:dd:yyyy	
ReplacementValue	Variable	6 - 7	XXX.XX up to XXXX.XX	No rental instrument is worth less than \$100.00 or more than \$9999.99
InstrumentStatus	Fixed	1	boolean 0 = rented, 1 = available	Assume rented or available are the only options.

Total:	51 - 510
--------	----------

3. Assume that your disk has these parameters:

- Block size B = 512 bytes
- Interblock gap size G = 128 bytes
- Number of blocks per track = 20
- Number of tracks per surface = 400
- A disk pack consists of 12 double-sided disks.

Assume that the file INSTRUMENTs has r = 600 records. Suppose that only 75 percent of the INSTRUMENT records have a value for Model, 21 percent for LoanLocation and 66 percent for Comments. Each record has a 1 byte end-of-record marker. We will use spanned record organization where each block has a 5 byte pointer to the next block. (This space is not used for record storage.)

a. Calculate the average record length R in bytes.

INSTRUMENT					
Attribute	Fixed/Variable/Multivalued	Min Size (bytes)	Max Size (bytes)	% Assumptions	Average
InstrumentName	Variable	1	30		15.5
Serial Number	Variable	1	10		5.5
Condition	Variable	1	30		15.5
FundsFrom	Variable	1	30		15.5
Size	Variable	14	23		18.5
Last_Inspection_Date	Fixed	10	10		10
Comments	Variable	0	200	0.66	66
Model	Variable	0	30	0.75	11.25
Maker	Variable	1	30		15.5
ShopUsed	Multivalued	0	35		17.5
InstrumentYearMade	Fixed	4	4		4
LoanLocation	Variable	0	30	0.21	3.15
LocationName	Variable	1	30		15.5
PurchaseDate	Fixed	10	10		10
ReplacementValue	Variable	6	7		6.5
InstrumentStatus	Fixed	1	1		1
End of Record Character	Fixed	1	1		1
Average:					231.9

$$\therefore R = 232 \text{ bytes}$$

b. Calculate bfr and tell what bfr represents in this particular case.

$$bfr = \left\lfloor \frac{B}{R} \right\rfloor = \left\lfloor \frac{512 \text{ bytes} - 5 \text{ bytes}}{232 \text{ bytes}} \right\rfloor = 2 \quad (512 \text{ bytes/block} - 5 \text{ bytes for next block pointer})$$

The  $bfr$  represents the block size divided by the summation of the arithmetic means of the minimum and maximum sizes for all attributes in the table (since we don't currently know about the distribution of usage for each attribute). More simply, it's the number of average sized records that can fit in a block.

- c. Calculate the number of blocks  $b$  needed for the file.

$$b = \left\lceil \frac{r}{bfr} \right\rceil = \left\lceil \frac{600 \text{ records}}{2} \right\rceil = 300 \text{ blocks}$$

4. What is the primary purpose of a GOOD hashing algorithm?

The primary purpose of a good hash function is to avoid collisions, but it should also distribute records evenly, and minimize unused space.

5. Estimate the size of the INSTRUMENT file after 5 years. State your "real world" assumptions about how many records per year will be added and why.

Assuming that there are an average of 5 rental transactions happening every day which seems like a reasonable amount of rentals per day. Some days there may not be any rentals, some days there may be more than 5 rentals. That means there are  $5 \frac{\text{rentals}}{\text{day}} \cdot \left( 365 \frac{\text{days}}{\text{year}} + 0.25 \frac{\text{leap days}}{\text{year}} \right) \cdot 5 \text{ years} = 5 \cdot 365.25 \cdot 5 = 9131.25 \text{ rentals}$  over a 5 year period. This means that there are this many records that need to be saved to the file ( $r = 9131.25 \text{ records}$ ).

$$b = \left\lceil \frac{r}{bfr} \right\rceil = \left\lceil \frac{9132 \text{ records}}{2} \right\rceil = 4566 \text{ blocks}$$

$$4566 \text{ block} \cdot 512 \frac{\text{bytes}}{\text{block}} = 2.33779 \text{e}6 \text{ bytes} = 2.33779 \text{ gigabytes}$$