**Homework Assignment 4**

Due: 11:59PM June 7, 2021

1. Read the relevant textbook sections and answer the questions below.

(a) Ans) trigger

(b) Ans) table functions

(c) Ans) transaction

(d) Ans) index

(e) According to the textbook description from Section 4.2, what is the main difference between views and named subqueries defined by WITH?

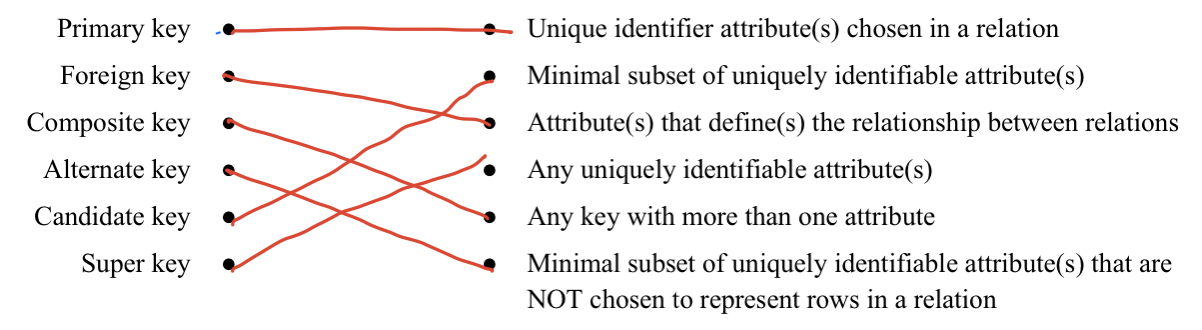
Ans) if view once created remain available until explicitly droped. But with clause is local query

(f) List the conditions that an SQL view is said to be updatable.

Ans) In FROM clause, when number of table is only one. SELECT clause contains only attribute names of the relation, and does not have any expressions, aggregates, or DISTINCT specification. The query does not have a GROUP BY or HAVING clause

2. Match each of the following key types to the corresponding definition.

Ans)



3. Answer the following questions that are from the textbook exercise problem sets. You may refer to the Internet as well as the textbook for assistance; however, your solution should contain your own ideas in your own language.

a) Show how to define a view tot\_credits(year, num\_credits), giving the total number of credits taken in each year.

Ans) CREAT VIEW tot\_credits as SELECT year, SUM(num\_creditts) AS num\_credits FROM course JOIN takes ON course.course\_id = takes.course\_id GROUP BY year;

b) For the view of Exercise 4.20, explain why the DBMS would not allow a tuple to be inserted into the database through this view.

Ans) Because there is GROUP BY clause, the insert is not allowed.

c) In the variable-length record representation, a null bitmap is used to indicate if an attribute has the null value.

i. Section 13.2.2 of the textbook explains what a null bitmap is.

Ans) It contains information about which attribute is null value.

ii. For variable-length fields, if the value is null, what would be stored in the offset and length fields?

Ans) nothing is stored.

iii. In some applications, tuples have a very large number of attributes, most of which are null. Can you modify the record representation such that the only overhead for a null attribute is the single bit in the null bitmap?

Ans) if there is null attribute, then just one bit is needed. Because value, offset, and length is not necessary.

d) Explain why the allocation of records to blocks affects database-system performance

significantly.

Ans) it can reduce the number of disk accesses for a given operation. Therefore, it significantly improves performance

e) Some databases use magnetic disks in a way that only sectors in outer tracks are used,

while sectors in inner tracks are left unused. What might be the benefits of doing so?

Ans) the data transfer rate will be greater on the outer tracks then inner tracks, and more importantly by using only outer track, the disk arm movement is minimized.

f) Operating systems try to ensure that consecutive blocks of a file are stored on consecutive disk blocks. Why is doing so very important with magnetic disks? If SSDs were used instead, is doing so still important, or is it irrelevant? Explain why.

Ans) The disk needs movement to read the information. Therefore, performance is much higher if it is in consecutive spaces. SSD has no moving part. Therefore, it does not matter.

g) The lost update anomaly is said to occur if a transaction Tj reads a data item, then another transaction Tk writes the data item (possibly based on a previous read), after which Tj writes the data item. The update performed by Tk has been lost, since the update done by Tj ignored the value written by Tk.

i. Give an example of a schedule shown the lost update anomaly.

Ans)

|  |  |
| --- | --- |
|  |  |
| Read(A)  Write(A) | Read(A)  Write(A) |

ii. Give an example schedule to show that the lost update anomaly is possible with the read committed isolation level.

Ans)

|  |  |
| --- | --- |
|  |  |
| Begin  Read(A)  Write(A)  Commit | Begin  Read(A)  Write(A)  Commit |

iii. Explain why the lost update anomaly is not possible with the repeatable read isolation level.

Ans) the lost update anomaly is not possible in repeatable read isolation level and repeatable read transaction only accesses that has been commited before it start.

4. Consider the following timelines where two transactions are intervening each other. The two vertical downward arrows represent the progression of time. The horizontal arrows represent the dataflow between transaction and storage.

Ans)

|  |  |  |  |
| --- | --- | --- | --- |
|  | REPEATABLE READ | READ COMMITED | READ UNCOMMITED |
| (a) | Null | Null | Null |
| (b) | Null | Null | Bob |
| (c) | Null | Null | Null |
| (d) | Jane | Jane | Jane |
| (e) | Jane | Jane | Mary |
| (f) | Jane | Mary | Amy |
| (g) | Jane | Mary | Mary |

5. Answer the following questions that are from the textbook exercise problem sets. You may refer to the Internet as well as the textbook for assistance; however, your solution should contain your own ideas in your own language.

a) When is it preferable to use a dense index rather than a sparse index? Explain your answer.

Ans) when index contains all of data. When fast search is needed. When memory overhead is not considered.

b) What is the difference between a clustering index and a secondary index?

Ans) A clustering index is sorting data according to the index. In a non-clustering index, the order of data is different from that of the index, and it is a dense index.

c) Indexes speed query processing, but it is usually a bad idea to create indexes on every attribute, and every combination of attributes, that are potentially search keys. Explain why.

Ans) The more the index contains many attributes, the more space the index record occupies. Also, the more clauses such as DML (UPDATE, INSERT, DELETE), the lower the performance.

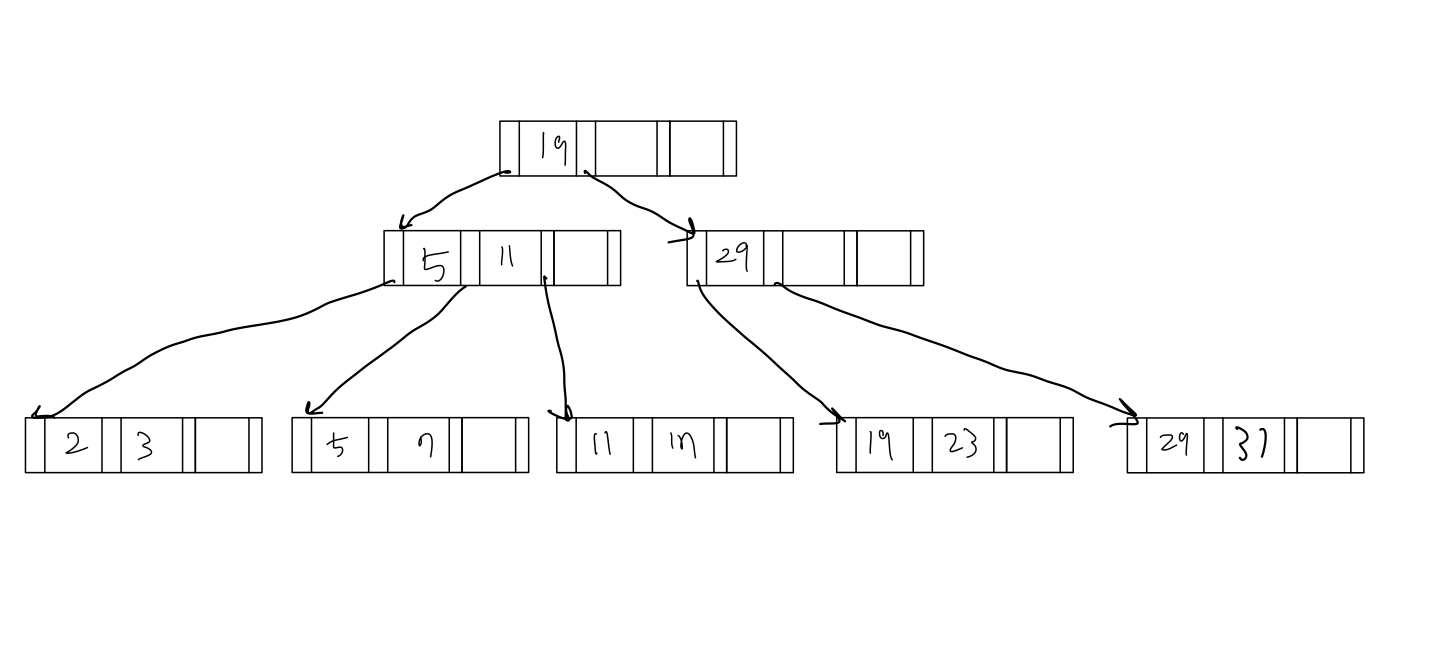
d) Construct a B+tree for the following set of key values:

(2, 3, 5, 7, 11, 17, 19, 23, 29, 31)

Assume that the tree is initially empty, and values are added in ascending order. Construct B+trees for the vases where the number of pointers that will fit in one node is as follows:

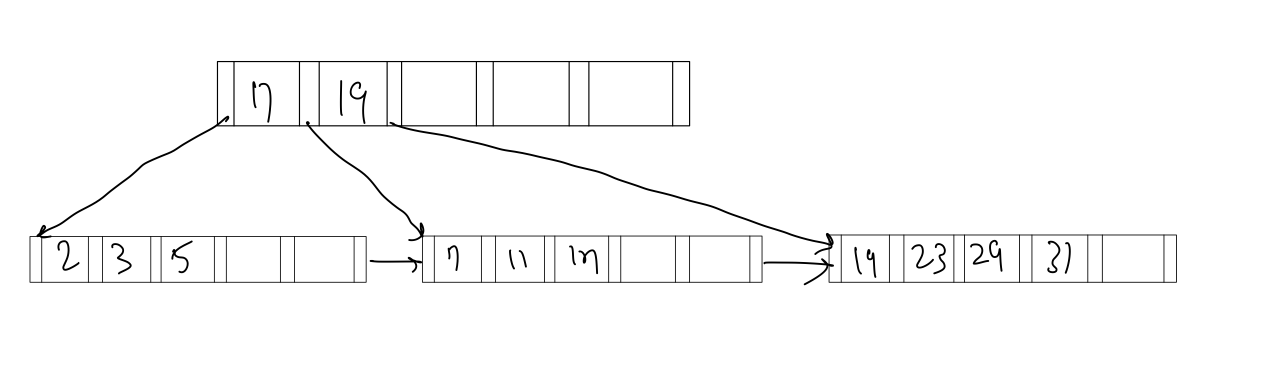
i. Four

Ans)



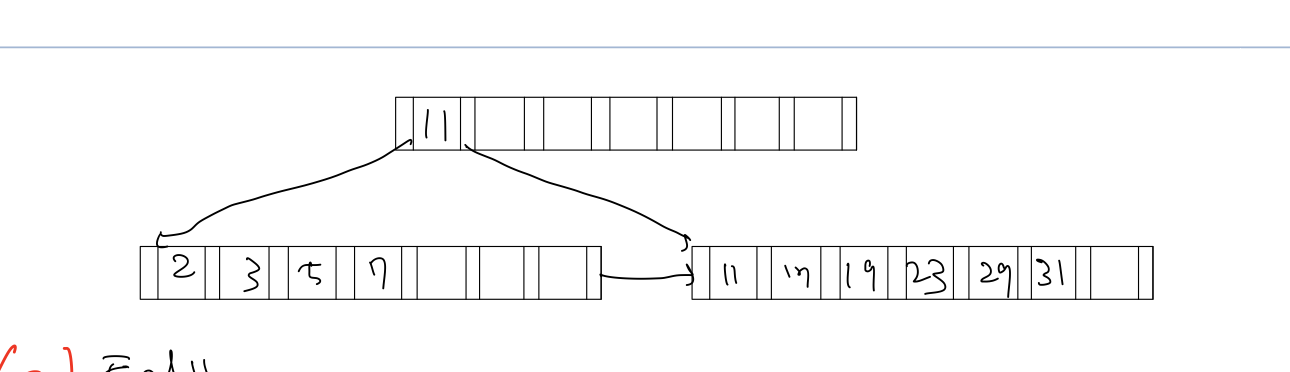
ii. Six

Ans)



` iii. Eight

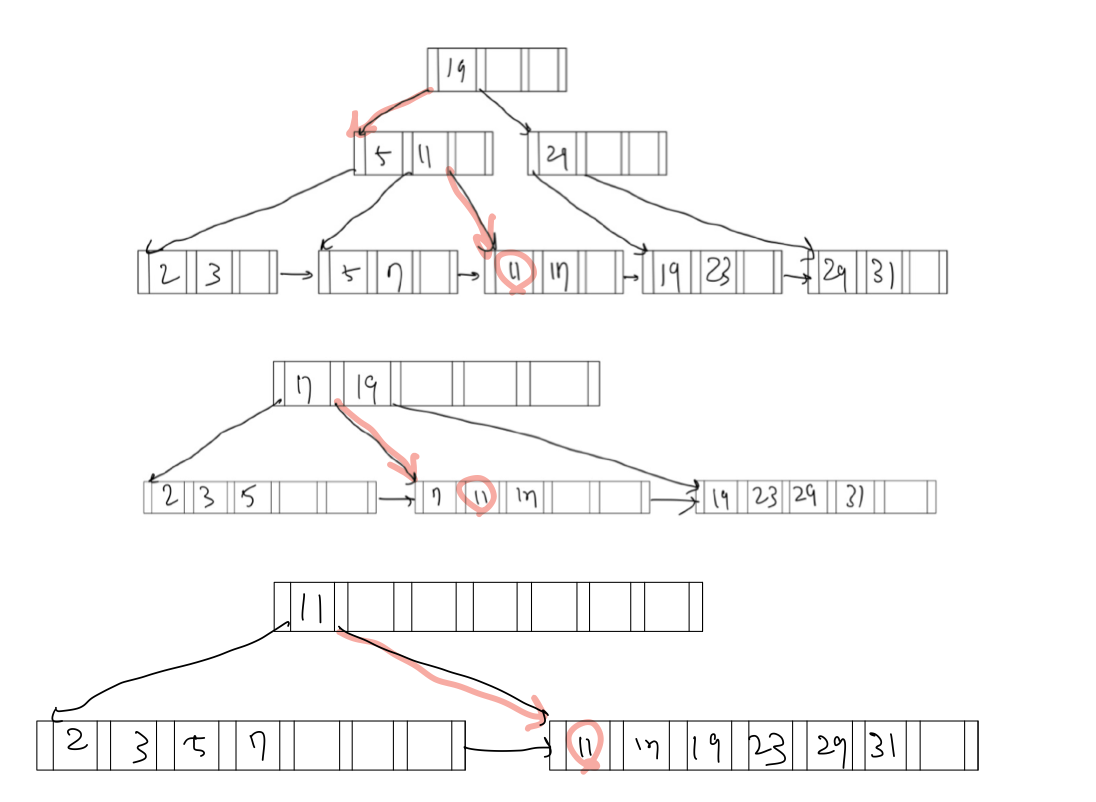
Ans)



e) For each B+tree of Exercise 14.3, who the steps involved in the following queries:

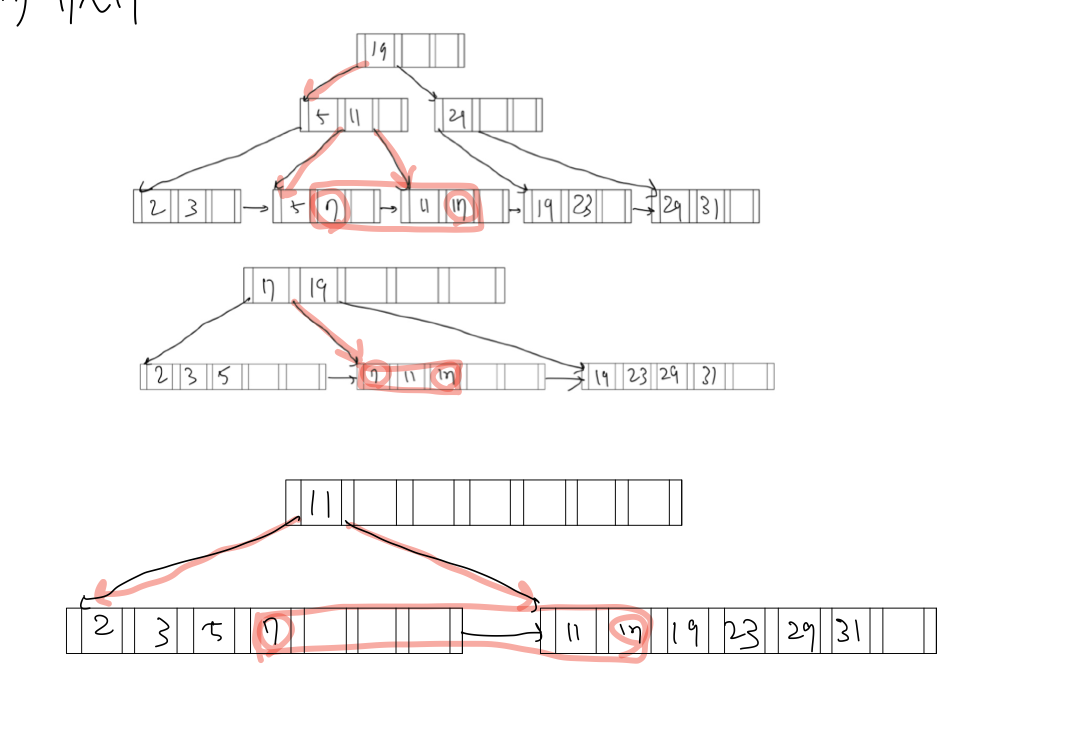
i. Find records with a search-key value of 11.

Ans)



ii. Find records with a search-key value between 7 and 17, inclusive.

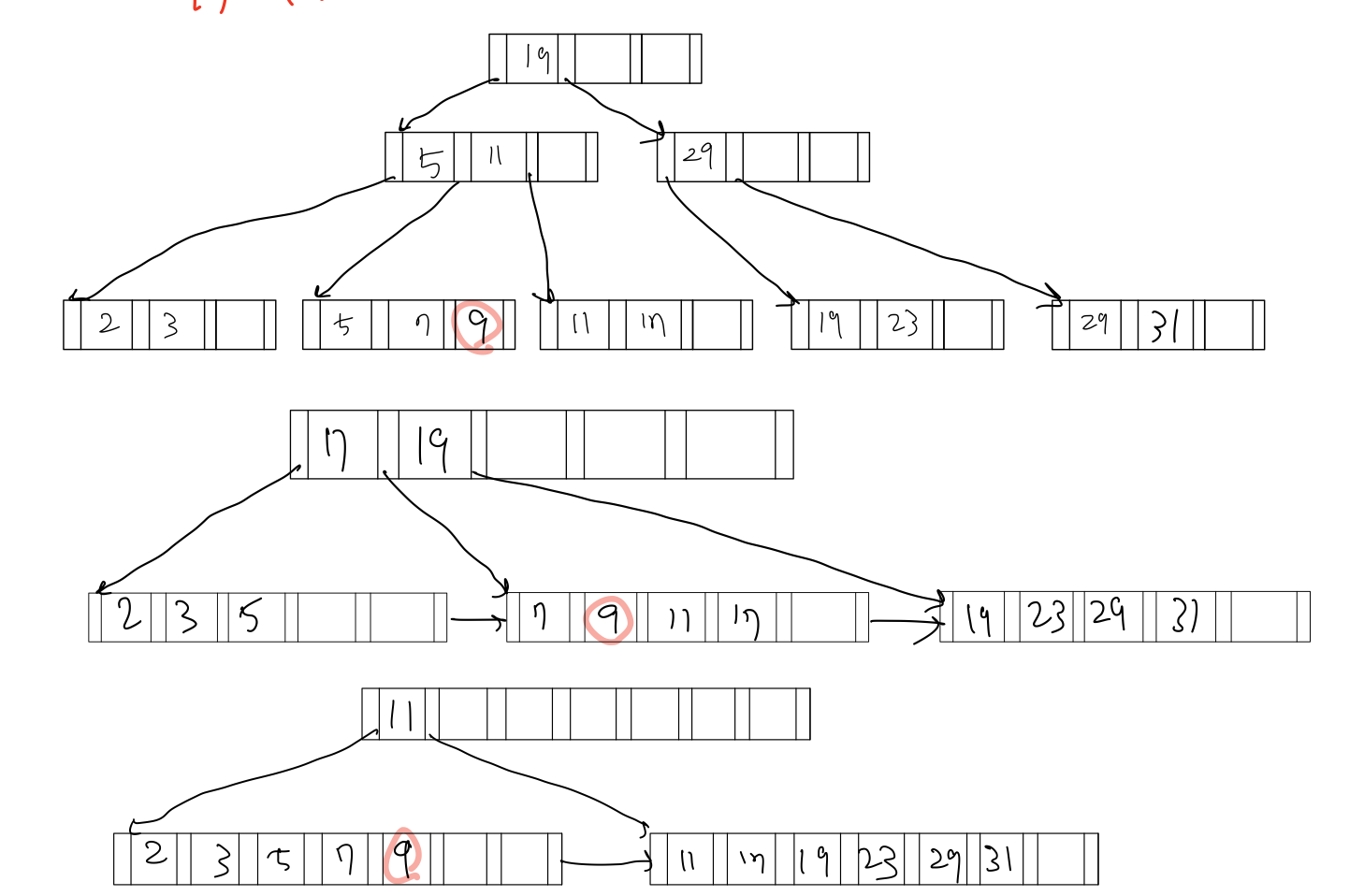
Ans)



f) For each B+tree of Exercise 14.3, show the form of the tree after each of the following series of operations:

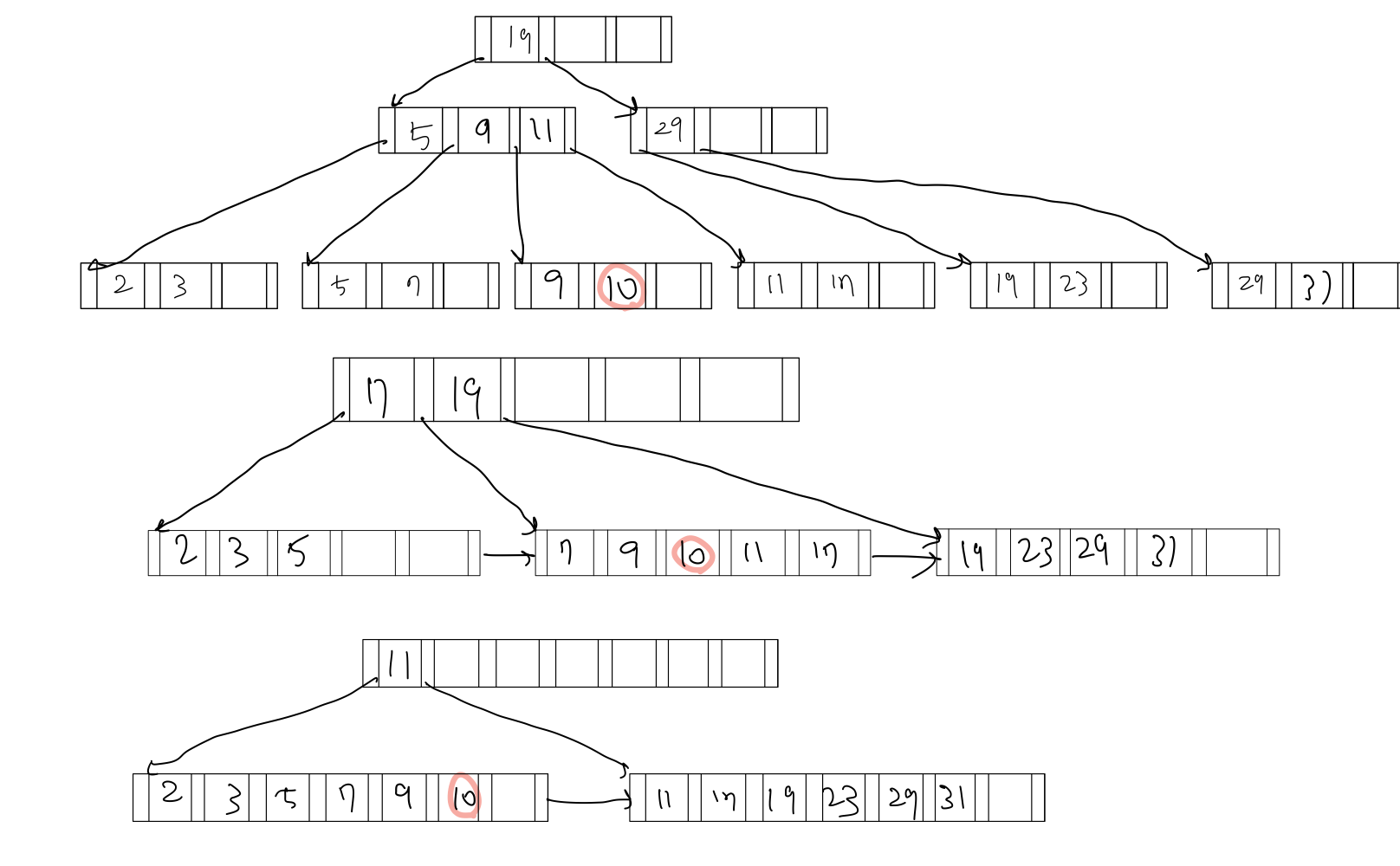
i. Insert 9

Ans)



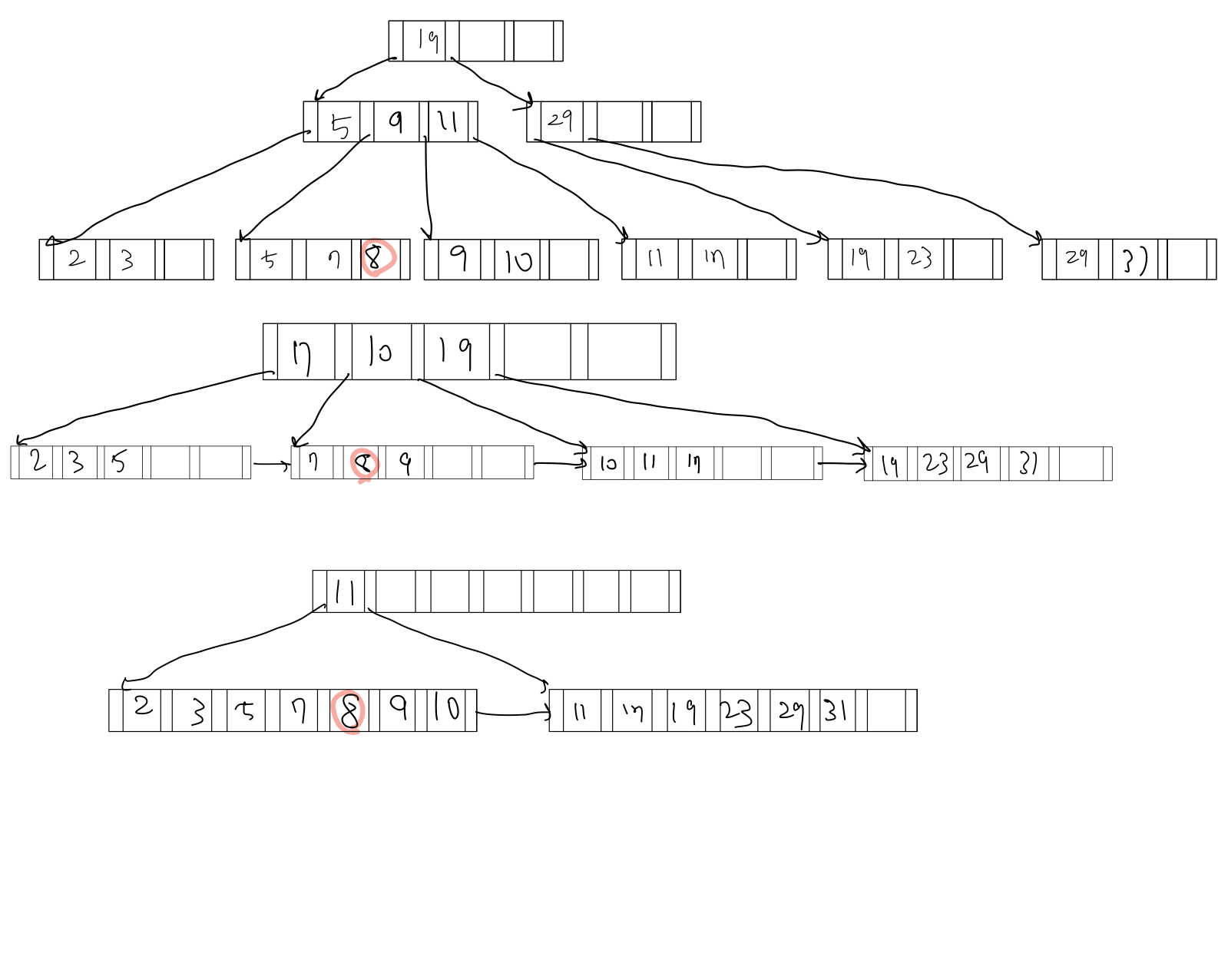
ii. Insert 10

Ans)



iii. Insert 8

Ans)



g) The leaf nodes of a B+tree file organization may lose sequentiality after a sequence of inserts. Explain why sequentiality may be lost.

Ans) Sequentiality may be lost if there are no more seats in the track on disk.