Priority Queue

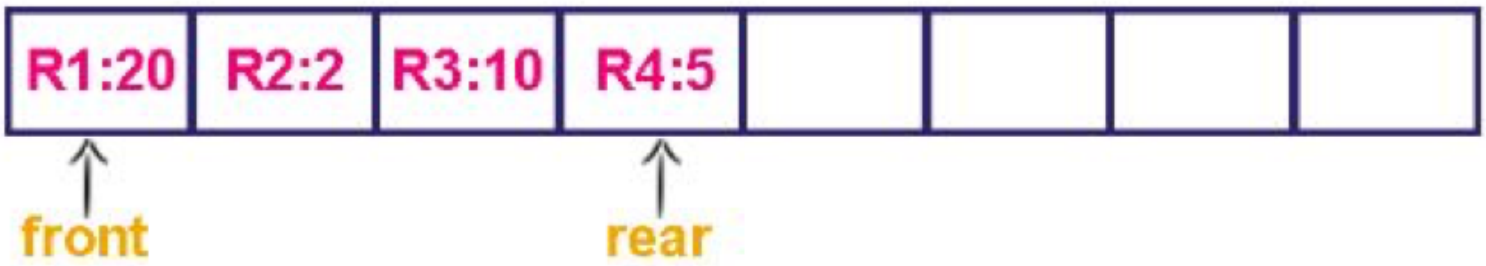
IIITS

Queue - Limitation

In normal queue data structure, insertion is performed at the end of the queue and deletion is performed based on the FIFO principle. This queue implementation may not be suitable for all situations.

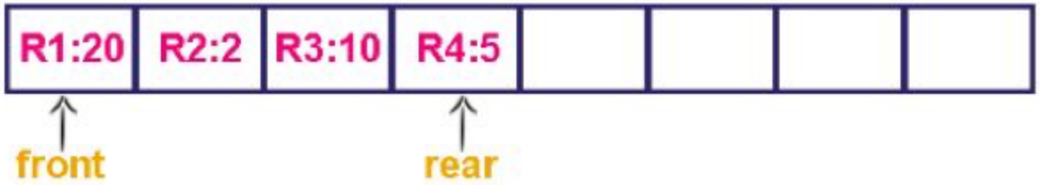
Example

Consider a networking application where server has to respond for requests from multiple clients using queue data structure. Assume four requests arrived to the queue in the order of R1 requires 20 units of time, R2 requires 2 units of time, R3 requires 10 units of time and R4 requires 5 units of time. Queue is as follows.



R1:20	R2:2	R3:10	R4:5				
front			rear				

Example - continue



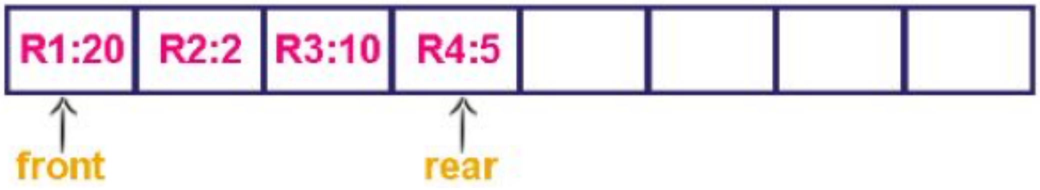
Now, check waiting time for each request to be complete.

R1 : 20 units of time
R2 : 22 units of time (R2 must wait till R1 complete - 20 units and R2 itself requires 2 units. Total 22 units)
R3 : 32 units of time (R3 must wait till R2 complete - 22 units and R3 itself requires 10 units. Total 32 units)
R4 : 37 units of time (R4 must wait till R3 complete - 35 units and R4 itself requires 5 units. Total 37 units)

Here, average waiting time for all requests (R1, R2, R3 and R4) is (20+22+32+37)/4 ≈ 27 units of time.

Can we make it better?

Example - continue



Now, consider another way of serving these requests. If we serve according to their required amount of time. That means, first we serve R2 which has minimum time required (2) then serve R4 which has second minimum time required (5) then serve R3 which has third minimum time required (10) and finnaly R1 which has maximum time required (20).

Now, check waiting time for each request to be complete.

R2 : 2 units of time
R4 : 7 units of time (R4 must wait till R2 complete 2 units and R4 itself requires 5 units. Total 7 units)
R3 : 17 units of time (R3 must wait till R4 complete 7 units and R3 itself requires 10 units. Total 17 units)
R1 : 37 units of time (R1 must wait till R3 complete 17 units and R1 itself requires 20 units. Total 37 units)

Here, average waiting time for all requests (R1, R2, R3 and R4) is (2+7+17+37)/4 ≈ 15 units of time.

Priority Queue - Implementation

1. Array - Max / Min
2. Array - Max/Min Circular
3. Array - Max/Min [Key - Priority] pair 2D/Two array/s
4. Linked List [Key - Priority] pair
5. Using Heap