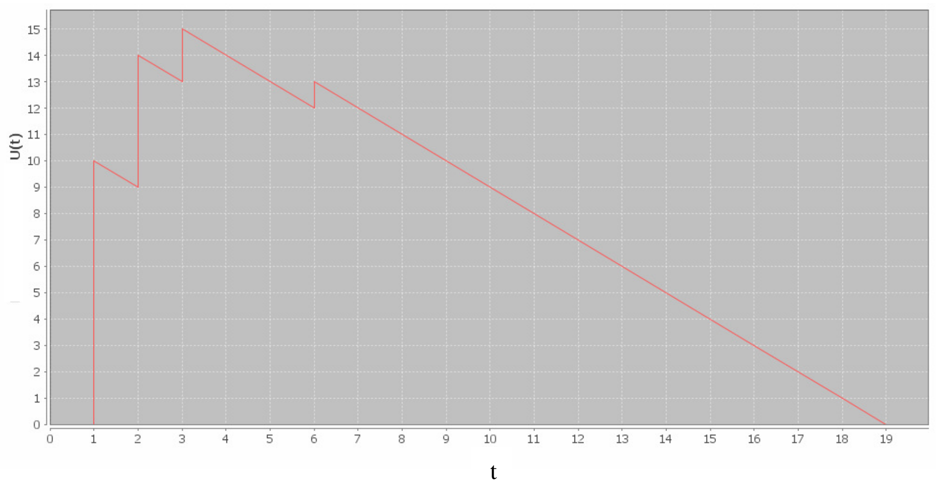
Operational Research / Queuing Systems Set 1

Problem 1

Plot the queuing processes $N(t)$ = *queue length* and $U(t)$ = *unfinished work* for an arrival stream specified by (t_n^+) = (0.5 s, 2 s, 5 s, 6.5 s) and (b_n) =(3 s.u., 2 s.u., 2 s.u., 3 s.u.), and for two cases:

(a) there are two processors, each serving requests at the speed of v = 1 s.u./s (no grading, processor-bound and time-sharing service mode are assumed),
(b) there is one processor serving requests at the speed of v = 2 s.u./s (time-sharing service mode is assumed).

Below is an example plot of $U(t)$. What is the value of S? What is the slope of the downward parts of the plot? What would be the slope in case b)? Retrieve the request arrivals stream.



Draw the plots of $U(t)$ for cases a) and b). Explain how they were arrived at. Give an analytic formula.

Explain the no grading, processor-bound and time-sharing service mode assumptions. Discuss the form the plots would take without them.

Discuss the fairness of the comparison of cases a) and b).

Name possible comparison criteria of cases a) and b). Is case a) superior? What are the merits of case b)?

Generalize the problem in possible directions. *Can our solution be automated?

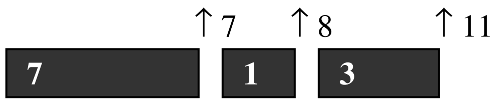
Problem 2

Compare mean system delays in a single-processor queuing system under FIFO and RR with service quantum 2 s (partial use of assigned quantum causes earlier commencement of the next quantum). Processor speed is 1 s.u./s. Three requests, X, Y and Z, of sizes 7 s.u., 1 s.u. and 3 s.u., respectively, arrive simultaneously and queue up in the order (a) XYZ, (b) YZX, (c) XZY.

Explain the workings of both queuing disciplines and premises to apply them in real-world queuing systems.

For RR, what other possibilities exist of handling partial use of assigned quantum?

In analyzing the XYZ scenario for FIFO, the diagram depicted below may be helpful. What diagram obtains for RR?



Analyze in the same way the scenarios b) i c) (perhaps also the other possible), separately for FIFO and RR, and calculate for each the mean system delay of a request.

What general properties of FIFO and RR can be deduced from the comparison of the calculated mean system delays in the respective scenarios? Discuss the value of such an analysis in the context of queuing systems.

Operational Research / Queuing Systems Set 2

Problem 1

A single-processor infinite-buffer queuing system with processor speed $v = 1$ s.u./s serves a "dense" arrival stream of requests creating a 75% offered load. The total service demand in a one-second observation period is a random variable with standard deviation $\sigma = 0.1$ s. What are the chances that the processor can spare half of the second to deal with other (e.g., system) tasks without a backlog of requests forming at the end of the observation period?

Let a random variable $\underline{X}$ represent the total service demand in a one-second observation period. What moments doe the probability distribution of $\underline{X}$ have? What argument yields the shape of this probability distribution?

Recall the relevant probability theorem and the form of $P(a \leq \underline{X} \leq b)$ that it implies. What values of a and b are of interest in our problem? Give a graphical illustration.

Recall the definition of the Laplace function. What properties of this function are useful when reading its tabulated numerical values (which can be found in textbooks or online)? Look up the necessary values in the tables.

Find the required probability and compare with that of forming a backlog at the end of the observation period.

Problem 2

A queuing system serves on average 800 transactions per second, each transaction on average requiring 5000 elementary operations to complete. An arriving transaction is immediately assigned a processor whose speed is 4,000,000 elementary operations/s. Find the mean number of transactions in system.

How many processors must there be in the system to support the above described operation? Think of a realistic application.

How long on average is the queue of transaction waiting to be processed?

Give the values of the mean request circulation (throughput) and lifetime in the system. What relationship does one now find helpful?

What does the obtained result say about the required number of processors?

Problem 3

A single-processor queuing system with a finite buffer of capacity Q works under offered load $r > 1$. In such a system, the loss fraction L never drops below a certain level – what?

How can one explain on operational grounds the fact that unbounded expansion of the buffer size does not prevent a nonzero loss fraction?

What happens to the processor idle time when Q tends to infinity? How can Little's law be applied to account for the above limit condition?

Operational Research / Queuing Systems Set 3

Problem 1

A queuing system of capacity Q handles telemetric reports generated by J identical terminals.

Find J_{max}, the maximum number of terminals that can be connected to the processor, and Q_{opt}, optimum buffer capacity (defined as the maximum number of accommodated reports) under the following assumptions:

- each terminal generates on average 20 reports per minute,
- a report contains on average 1800 records of data,
- the system uses a single processor capable of handling 12000 records per second,
- multiple access of the terminals is enabled by a common finite buffer,
- tolerable mean system delay of a report is 1.8 s, and
- tolerable loss fraction due to buffer overflow is 4%.

The table below shows, for various Q and offered load r, the mean system delay of a report, normalized to the mean report processing time (in boldface), and report loss fraction due to buffer overflow.

Q =	20		21		22		23		24		25	
r =												
0.1	**1.11**	0	**1.11**	0	**1.11**	0	**1.11**	0	**1.11**	0	**1.11**	0
0.2	**1.25**	0	**1.25**	0	**1.25**	0	**1.25**	0	**1.25**	0	**1.25**	0
0.3	**1.43**	0	**1.43**	0	**1.43**	0	**1.43**	0	**1.43**	0	**1.43**	0
0.4	**1.67**	0	**1.67**	0	**1.67**	0	**1.67**	0	**1.67**	0	**1.67**	0
0.5	**2**	0	**2**	0	**2**	0	**2**	0	**2**	0	**2**	0
0.6	**2.5**	0	**2.5**	0	**2.5**	0	**2.5**	0	**2.5**	0	**2.5**	0
0.7	**3.32**	0	**3.32**	0	**3.32**	0	**3.33**	0	**3.33**	0	**3.33**	0
0.8	**4.77**	0	**4.8**	0	**4.84**	0	**4.86**	0	**4.89**	0	**4.91**	0
0.9	**7.23**	0.01	**7.42**	0.01	**7.6**	0.01	**7.76**	0.01	**7.92**	0.01	**8.07**	0.01
1	**10.5**	0.05	**11**	0.05	**11.5**	0.04	**12**	0.04	**12.5**	0.04	**13**	0.04
1.1	**13.5**	0.11	**14.3**	0.1	**15.1**	0.1	**15.9**	0.1	**16.7**	0.1	**17.5**	0.1
1.2	**15.5**	0.17	**16.5**	0.17	**17.4**	0.17	**18.4**	0.17	**19.3**	0.17	**20.3**	0.17
1.3	**16.8**	0.23	**17.8**	0.23	**18.7**	0.23	**19.7**	0.23	**20.7**	0.23	**21.7**	0.23
1.4	**17.5**	0.29	**18.5**	0.29	**19.5**	0.29	**20.5**	0.29	**21.5**	0.29	**22.5**	0.29
1.5	**18**	0.33	**19**	0.33	**20**	0.33	**21**	0.33	**22**	0.33	**23**	0.33

Name the ingredients of a generic queuing system design problem and argue that all of them occur in the formulated problem.

What is the maximum allowable offered load?

What exactly do the boldface numbers in the table represent? How can the tolerable mean system delay be expressed in these terms?

Explain how the last two constraints of the problem influence the feasible values of Q.

Suggest some practical steps to raise J_{max}.

Problem 2

Each of 50 terminals connected to a common transceiver generates a request after a think time of average duration $^2/_3$ s. In 80% cases it is a message of average length 1000 bytes, and in 20% cases a control data report of average length 160 bytes. The transceiver works at 1 Mb/s in half-duplex; the average proportion of time it is switched to receive mode is 75% (during that time it is unavailable to the terminals). What is the resulting loss fraction?

Interestingly, we attempt to determine the loss fraction not knowing the buffer size (or even not explicitly assuming its presence) – how can one explain this?

The input data and specified mode of operation translate into the values of a_{m}, b_{m}, and the processor speed and idle time – what are they?

What analytic tool of queuing systems theory permits to link all the above quantities together?

Judge the obtained result and suggest some practical steps to improve it.

Problem 3

In a single-processor queuing system with buffer capacity $Q = 2$ in statistical equilibrium and under offered load $r = 0.75$, we have $p_0 \geq p_1 \geq p_2$. Find the range of possible p_1 values.

Assume that for the considered system, $p_k^+ \equiv p_k$, which in general is not true (on what assumptions is it true?), in this way we are able to link the queue state probabilities with the loss fraction.

What relationship between these probabilities now arises by virtue of the flow conservation equation?

Another relationship is yielded by the normalization constraint, hence we have two relationships (equations) and three unknowns – is it enough to solve our problem?

Instead of the numbers p_k consider the numbers $1 - p_k$, what are the relationships between them?