Computer Science and Data Science: second-round sample tasks

Module 1. Simple test questions with one correct answer (2 points for the correct answer)
Task 1
Positive integer N gives the remainder of 11 when divided by 17. Find positive integer M not exceeding 17 such that the product NM gives remainder of 1 when divided by 17.
a)   14
b)   17
c)   1
d)   13
Answer – a
Maximum 2 points.

Task 2
In a complete graph K6 on vertices v1, v2, … v6 the weights of the seven edges v1v2, v2v3, …, v5v6, v6v1, and v1v4 are equal to 1, and all the other edges have weight 2. How many minimal spanning trees does such graph have?
a) 15
b) 54
c) 1
d) 7
Answer – a
Maximum 2 points.

Task 3
What the Test Plan describes:
1) Testing procedure
2) Schedule for the implementation of test cases
3) Criteria for the end of testing
4) All together
Answer – 4
Maximum 2 points.

Task 4
What is the average estimate of the complexity of implementing tests in accordance with the plan compared to the total development costs:
1) 10%
2) 20%
3) 40%
4) 70%
Answer – 3
Maximum 2 points.

Task 5
Choose the correct option: 1 Petabyte is equal to:

ONE CLICK TO OPEN ALL DOORS

od.globaluni.ru

A) 8192 bits.
B) 8e+12 bits.
C) 8e+15 bits.
D) 9e+15 bits.
Answer – C
Maximum 2 points.

Task 6

If the entire text consists of k characters, then with a volumetric approach, the size of the information that contains I is determined by the formula: $I = k \cdot i$, where i is the information weight of one character in the used alphabet.

How much information is contained in the text if it consists of 315 characters, and one character is a character from the alphabet which power is 256:
A) 2520.
B) 80640.
C) 15280.
Answer – A
Maximum 2 points.

Task 7

What is Authorization?
1. granting a certain person or group of persons the rights to perform certain actions, as well as the process of verifying these rights when trying to perform these actions.
2. procedure for checking the legality of the user
3. the process of entering identification data
4. part of the information system
Answer – 1
Maximum 2 points.

Task 8

What is Authentication?
1. the act of proving an assertion
2. procedure for checking access rights
3. procedure for matching a user with records in the database
4. procedure for checking the presence of a user
Answer – 1
Maximum 2 points.

Task 9

What is OLAP?
A) a set of concepts, principles and requirements that underlie in the software products that make it easier for analysts to access data
B) means for data processing and analysis

C) the software technology for data storage
D) a model of human-computer interaction
Answer – A
Maximum 2 points.

Task 10
1.2. List the disadvantages of Gantt chart data analysis.
A) inflexibility, dependence, overflow
B) inflexibility, independence, overflow
C) flexibility, dependence, overflow
D) flexibility, independence, overflow
Answer – A
Maximum 2 points.

Module 2. Mid-level tasks (7 points for the correct solution)
Task 11
Let S = {a, b, c, d, e, f , g}. Determine which of the following are partitions of S:
a) P1 = [{a, c, e}, {b}, {d, g}]
b) P2 = [{a, e, g}, {c, d}, {b, e, f }],
c) P3 = [{a, b, e, g}, {c}, {d, f }],
d) P4 = [{a, b, c, d, e, f , g}].
Answer – c, d
2 correct answers - 7 points
1 correct answer - 4 points
Each wrong answer - minus 4 points
Maximum 7 points.

Task 12
What types of testing are used in software design?
1) Module testing
2) System testing
3) Periodic testing
4) Integration testing
5) Regression testing
6) Thematic testing
Answer – 1, 2, 4, 5
4 correct answers - 7 points
3 correct answers - 5 points
2 correct answers - 3 points
1 correct answer - 1 point
Each wrong answer - minus 4 points
Maximum 7 points.

Task 13
2.3 A single-channel (mono) sound recording is made with a sampling frequency of 16 kHz and 24-bit resolution. The recording lasts 1 minute, its results are written to a file, and data compression is not performed. Which of the following values is closest to the resulting file size?
A) 0.2 Megabyte.

B) 3 Megabyte.
Answer – B
1 correct answer - 7 points
1 wrong answer - minus 7 points
Maximum 7 points.

Task 14
Key principles of the CIA triad
1. Confidentiality
2. Integrity
3. Availability
4. Continuous access
5. Infinity access
6. Aptitude for save
Answer – 1, 2, 3
All 3 correct answers selected - 7 points
Only 2 correct answers selected - 5 points
Only 1 correct answer selected - 2 points
For choosing one wrong answer minus 3 points
For choosing two wrong answers minus 6 points
For choosing three wrong answers minus 7 points
Maximum 7 points.

Task 15
What is the purpose of Data Storage (DS)?
A) to ensure the collection, storage and quick access to the key information
B) for the preparation of all kinds of reports
C) for the operational analytical processing and Data Mining
D) to view video files;
E) to listen to audio files;
F) for the development of the database.
Answer – a, b, c
All 3 correct answers selected - 7 points
Only 2 correct answers selected - 5 points
Only 1 correct answer selected - 2 points
For choosing one wrong answer minus 3 points
For choosing two wrong answers minus 6 points
For choosing three wrong answers minus 7 points
Maximum 7 points.

Module 3. Tasks of a difficult level with a detailed answer (15 points for the right answer)
Task 16
In the $(x_1 \circ x_2 \circ x_3) \wedge (x_4 \circ x_5 \circ x_6) \wedge \ldots \wedge (x_{31} \circ x_{32} \circ x_{33})$ each all the symbols $\circ$ are independently and uniformly changed to one of $\{\vee, \wedge\}$, and each of the variables $x_1, \ldots, x_{33}$ is assigned randomly a value of 0 (false) or 1 (true). What is the probability that the resulting expression evaluates at 1?

Answer: $\left(\frac{9}{16}\right)^{11}$.

ONE CLICK TO OPEN ALL DOORS

od.globaluni.ru

Solution:

The answer is a fraction to a power. The degree is eleven because we have 11 expressions like $(x_n \circ x_{n+1} \circ x_{n+2})$ joined by conjunction. The denominator of the fraction is the number of options in the expression $(x_n \circ x_{n+1} \circ x_{n+2})$ minus 32 (since 3 variables capable of taking two values, and two operations capable of taking two values, 8*4=32).

The numerator of the fraction is the number of variants of the expression $(x_n \circ x_{n+1} \circ x_{n+2})$, which takes the value 1 - 18. We reduce the fraction 18/32, we get 9/16.

Evaluation criterion:

for finding the degree - 5 points,

for finding the numerator - 5 points

for finding the denominator - 5 points.

Maximum 15 points.

Task 17

There is a sample containing 30 numerical values of some sign of a random variable X:

19	25	22	16	22	14	17	19	18	20
22	26	24	18	16	19	22	14	18	14
25	17	18	14	20	18	24	25	16	18

Formulate the definitions of the mode, median and calculate their values.

Answer: Mo = 18; Me = 18,5.

Solution:

The mode is most often encountered meaning. (**Definition of the mode - 3 points**).

Let us write down the distribution row (all values in ascending order): 14, 14, 14, 14, 16, 16, 16, 17, 17, 18, 18, 18, 18, 18, 18, 19, 19, 19, 20, 20, 22,22, 22, 22, 24, 24, 25, 25, 25, 26 (**Writing down distribution row - 2 points**).

Mo = 18. (**The correct answer is 3 points**)

The median is the number in the middle in an ascending order of the numbers (if their number is odd) or the half-sum of the numbers in the middle places in an ordered set of these numbers (if their number is even). (**Definition of the concept of median - 3 points**).

The original series has an even number, so the median is equal to half the sum of the numbers in the middle places in the ordered set of these numbers.

14, 14, 14, 14, 16, 16, 16, 17, 17, 18, 18, 18, 18, 18, 18, 19, 19, 19, 20, 20, 22,22, 22, 22, 24, 24, 25, 25, 25, 26.

$$M_e = \frac{18 + 19}{2} = 18{,}5.$$

(**The correct answer is 4 points**)

Maximum 15 points.

Task 18

You have received a datagram from some server:

e0eaeb3be718a85fa241a340a85db450d53fd53dc60aa6

The structure of the datagram is as follows:

2 bytes	2 bytes	2 bytes	...	2 bytes

ONE CLICK TO OPEN ALL DOORS

od.globaluni.ru

Sender port	Receiver port	Datagram length	Data (payload)	Checksum

The length of a datagram is the sum of all its fields, including the length field. All network traffic is encrypted using a simple XOR cipher with a cyclically repeating 2- byte key. The datagram header, including the destination and sender ports, is also encrypted.

Determine what data the server has sent if the port is 2021.

Supplementary materials: see hex codes of ASCII symbols at

https://www.cisco.com/c/en/us/td/docs/ios/12_4/cfg_fund/command/reference/cfnapph.html

Answer: OPENDOORS_2022!

Solution:

Knowing the datagram structure (1) and that it was sent by a server (2), we can deduce that e0 ea is the encrypted representation of the sender port, i.e. 2021. A simple XOR cipher is a bitwise XOR operation performed on every character using a given key. Therefore, an XOR operation on the first bytes of the datagram and the two bytes corresponding to the port number will help to obtain the key. To this end, we convert the number 2021 into hex to get 07e5 and calculate e0ea XOR 07e5 = e70f (or 1110011100001111). e70f is the secret key.

Next, we sequentially use the key to decrypt the datagram data. First, we determine the size of the data by decrypting the size of the datagram e7 18 XOR e7 0f = 17 (or 23 in decimal.) Thus, the length of the payload is 23 - 8 = 15, and it starts at the 7th byte a8 5f a2 41 a3 40 a8 5d b4 50 d5 3f d5 3d c6. By performing an XOR operation on this sequence with the e7 0f key, we obtain ASCII character codes in hex. Each ASCII character is encoded in one byte. The decryption result is OPENDOORS_2022!

Correctly translated the server port into hexadecimal number system - 3 points

Correctly identified the secret key - 5 points

Correctly determined the size of the message - 3 points

Correctly deciphered the entire message - 4 points.

Maximum 15 points.