Computer and Data Science: Second-round Sample Tasks for the Open Doors Postgraduate Track

You will be asked to complete 35 tasks, including:
- 21 entry-level tasks, each correct answer worth 1 point;
- 10 intermediate-level tasks, each correctly answered task worth 3 points;
- 1 intermediate-level task worth 4 points;
- 3 advanced tasks (constructed response), each correctly completed task valued at 15 points.

Evaluation criteria and standard answers are provided for the advanced tasks requiring constructed responses.

Applied Mathematics

Task 1
Entry level (1 point)

In accordance with the POSIX Extended writing standard, the following regular expression is /^[abc]+[a-zA-Z]?8{3}/. Select the string that corresponds to it.
   a) aaFDF888
   b) aaF888
   c) xzyaaR888
   d) aaR8

Answer: b.

Task 2
Entry level (1 point)

If the graph is given by an adjacency matrix:
{0, 1, 1,0,1 1,0,0,1,0 1,0,0,1,1 0,1,1,0,1 1,0,1,1,0 },
What is the weight of the minimum spanning tree?
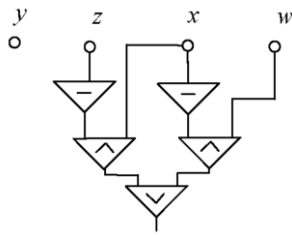   a) 2
   b) 3
   c) 4
   d) 1

Answer: c.

Task 3
Entry level (1 point)

Which Boolean function does this scheme of functional elements correspond to?

a) $xw \lor xz$
b) $\underline{x}y \lor w\underline{z}$
c) $\underline{x}w \lor x\underline{z}$
d) $\underline{x}w \lor x\underline{z}$

Answer: d.

Task 4
Intermediate level (3 points)

Which of the following functions is in disjunctive normal form (DNF)?
a) $f(x, y, z) = x \land x \oplus x \land y \land z$
b) $f(x, y, z) = (x \lor \underline{x}) \land (x \lor y \lor z) \land (y \lor x)$
c) $f(x, y, z) = x \land y \lor x \land \underline{y} \land z$
d) $f(x, y, z) = x \lor y \lor z$

Answer: c, d.

Task 5
Intermediate level (3 points)

For which sets A and B, is $|A \cup B| = 5$?
a) $A = \{1,2,3\}\ B = \{1,6,7,2\}$
b) $A = \{1,2,3\}\ B = \{1,6\}$
c) $A = \{1,2,3\}\ B = \{6,7\}$
d) $A = \{1,2,3\}\ B = \{1,2\}$

Answer: a, c.

Task 6
Advanced level (15 points)

Transform the function $F(x_1, x_2, x_3, x_4) = 1001\ 0010\ 1010\ 1100$, using the formula of disjunctive decomposition over a set of variables x_1, x_2. Provide the resulting functions for the variables x_3, x_4, using a set of operations from the basis: NOT, OR, AND.

Please note that the evaluation will consider your problem-solving process; providing only the final answer is not sufficient.

Solution:

Using the disjunctive decomposition formula for two variables, we obtain:

$$F(x_1, x_2, x_3, x_4) = V_{(a_1, a_2)} x_1^{a_1} \cdot x_2^{a_2} \cdot f_{1..4}(a_1, a_2, x_3, x_4)$$

$$= V_{(0,0)\,(0,1)\,(1,0)\,(1,1)} x_1^{a_1} \cdot x_2^{a_2} \cdot f(a_1, a_2, x_3, x_4)$$

$$= x_1^0 \cdot x_2^0 \cdot f_1(0,0,x_3,x_4) \vee x_1^0 \cdot x_2^1 \cdot f_2(0,1,x_3,x_4) \vee x_1^1 \cdot x_2^0 \cdot f_3(1,0,x_3,x_4) \vee x_1^1 \cdot x_2^1 \cdot f_4(1,1,x_3,x_4)$$

$$= \underline{x_1} \cdot \underline{x_2} \cdot f_1(0,0,x_3,x_4) \vee \underline{x_1} \cdot x_2 \cdot f_2(0,1,x_3,x_4) \vee x_1 \cdot \underline{x_2} \cdot f_3(1,0,x_3,x_4) \vee x_1 \cdot x_2 \cdot f_4(1,1,x_3,x_4)$$

(A correct entry of the disjunctive decomposition formula for the specified variables is worth 5 points.)

Let us determine the values of the functions f1, f2, f3, f4 on each set, using the values on similar sets of the original function F (you can use the truth table or the Karnaugh map for visual comparison, which are presented below). For example, in the first column, we alternately define

$$f_1(0000) = 1,$$
$$f_1(0001) = 0,$$
$$f_1(0010) = 0,$$

$f_1(0011) = 1$; we also determine the functions f_{2-4}.

The truth table of the function F

x_1	x_2	x_3	x_4	F
0	0	0	0	1
0	0	0	1	0
0	0	1	0	0
0	0	1	1	1
0	1	0	0	0
0	1	0	1	0
0	1	1	0	1
0	1	1	1	0
1	0	0	0	1
1	0	0	1	0
1	0	1	0	1
1	0	1	1	0
1	1	0	0	1
1	1	0	1	1
1	1	1	0	0
1	1	1	1	0

A Karnaugh map for F

$x_1 x_2$ / $x_3 x_4$	00	01	11	10
00	1	0	1	0
01	0	0	0	1
11	1	1	0	0
10	1	0	0	1

ONE CLICK TO OPEN ALL DOORS

x_3	x_4	$f_1(0,0,x_3,x_4)$	$f_2(0,1,x_3,x_4)$	$f_3(1,0,x_3,x_4)$ ∨	$f_4(1,1,x_3,x_4)$
0	0	1	0	1	1
0	1	0	0	0	1
1	0	0	1	1	0
1	1	1	0	0	0

(A correct comparison with the values of the original function is worth 5 points.)

All that remains is to compare each function with the set of Boolean algebra operations, obtaining:

$$F(x_1,..x_4) = \underline{x_1} \cdot \underline{x_2} \cdot (\underline{x_3} \vee x_4) \cdot (x_3 \vee \underline{x_4}) \vee \underline{x_1} \cdot x_2 \cdot (\underline{x_3} \vee x_4) \vee x_1 \cdot \underline{x_2} \cdot \underline{x_4} \vee x_1 \cdot x_2 \cdot \underline{x_3}$$

(A correctly written disjunctive decomposition formula for x_1 и x_2 with pre-defined functions for x_3 and x_4 is worth 5 points.)

Computer Science: Information Systems

Task 7
Entry level (1 point)

Which of the following is a telecommunications device?
    a) Network adapter
    b) Processor
    c) Printer
    d) Video card
Answer: a.

Task 8
Entry level (1 point)

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.

Task 9
Entry level (1 point)

Which knowledge representation model aligns with the understanding of long-term memory organization in humans?
    a) Production representation
    b) Semantic networks
    c) Frames
    d) Integral representation

ONE CLICK TO OPEN ALL DOORS

od.globaluni.ru

Answer: b.

Task 10
Intermediate level (3 points)

Which are the main stages of working with data in information systems?
   a) Extraction and transformation stage
   b) Machine learning stage
   c) Loading phase
   d) Data model development stage

Answer: a, c.

Task 11
Intermediate level (3 points)

Which of the following is the purpose of Data Storage (DS)?
   a) Ensuring the collection, storage and quick access to the key information
   b) Data visualization
   c) Operational analytical processing and data mining
   d) Viewing video files

Answer: a, c.

Computer Science: Artificial Intelligence

Task 12
Entry level (1 point)

What happens when we introduce more variables to a linear regression model?
   a) Both might increase or decrease depending on the variables introduced.
   b) The R-squared value may increase or decrease, but the adjusted R-squared always increases.
   c) Both R-squared and adjusted R-squared always increase.
   d) The R-squared value may increase or remain constant; the adjusted R-squared may increase or decrease.

Answer: d.

Task 13
Entry level (1 point)

Which of the following represents the upper and lower bounds of MCC in case of binary classification?
   a) (0,1]
   b) [1,0)

c) [-1,1]
d) (-1,1)
Answer: c.

Task 14
Entry level (1 point)

What happens during neural network training?
a) The number of layers and the number of neurons in each layer are adjusted.
b) The parameters of layers and weights are adjusted.
c) The length of the gradient descent step is adjusted.
d) The input object is configured.
**Answer: b.**

Task 15
Intermediate level (3 points)

Which TWO of the following clustering algorithms suffer from the problem of convergence at local optima?
a) K-Means++ algorithm
b) EM algorithm
c) Brute force algorithm
d) K-Means algorithm
Answer: b, d.

Task 16
Advanced level (15 points)

Consider a two-layer perceptron with 2 neurons in the hidden layer and 2 neurons in the output layer. The input vector has a length of 2.
describesis
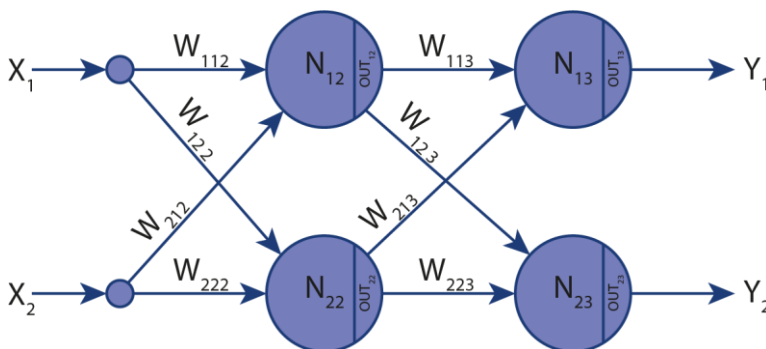
Each neuron is identified by the letter N_{jl}, where j describes the neuron number in the layer, and l is the layer number. Each neuron has an input, the weight of which is given by the value w_{ijl}, where i is the number of the neuron's input; j describes the neuron number in the layer; l is the layer number. The neurons in the input layer do not perform any computational operations; they solely

distribute the input signals. OUT_{jl} is the output value of the neuron. (X_1, X_2) is the input vector, i.e. the vector that is fed to the input of the network. (Y_1, Y_2) is the output vector that the neural network generates as a result of work. For the output layer $OUT_{jl} = Y_j$. The activation functions of each neuron are sigmoid (logistic function).

The perceptron is trained using the backpropagation algorithm without a training schedule. The algorithm uses the concept of error gradient for each neuron—∇E_{jl}. The rms error function is used as the error.

Weights of neurons in the hidden layer:

W_{112}	W_{212}	W_{122}	W_{222}
0.7599	0.1452	0.9608	0.6935

Output layer neuron weights:

W_{113}	W_{213}	W_{123}	W_{223}
-0.2808	-0.1997	-0.9667	-0.6111

The input vector:

X_1	X_2
-0.5851	-0.5826

For a given input vector, an output vector is expected (training vector):

D_1	D_2
0.2030	-0.8542

Assignment:
1. Write the neuron activation formula.
2. Write the formula for calculating the outputs of neuron N_{12}.
3. Write the formula for calculating the error gradient for the output layer neuron N_{23}.
4. Write the formula for calculating the error gradient for the hidden layer neuron N_{22}.
5. Write the formula for calculating the correction of the weights of neuron N_{13}.
6. Calculate the output of each neuron OUT_{jl}.
7. Calculate all error gradients ∇E_{jl}.
8. Calculate all new weights W_{ijl}.

Round all calculation results to four decimal places (10^{-4}).

Please note that the evaluation will consider your problem-solving process; providing only the final answer is not sufficient.

Solution:

1. Neuron activation formula.

$$\text{sig(x)} = \frac{1}{1 + e^{-x}}$$

ONE CLICK TO OPEN ALL DOORS

od.globaluni.ru

2. The formula for calculating the outputs of neuron N_{12}.

$$OUT_{12} = sig(x_1 * w_{112} + x_2 * w_{212})$$

3. The formula for calculating the error gradient for the output layer neuron N_{23}.

$$\nabla E_{23} = Y_2 * (1 - Y_2) * (Y_2 - D_2)$$

4. The formula for calculating the error gradient for the hidden layer neuron N_{22}.

$$\nabla E_{22} = OUT_{22} * (1 - OUT_{22}) * (\nabla E_{13} * w_{213} + \nabla E_{23} * w_{223})$$

5. The formula for calculating the correction of the weights of neuron N_{13}.

$$w_{113} = w_{113} - \nabla E_{13} * OUT_{12}$$
$$w_{213} = w_{213} - \nabla E_{13} * OUT_{22}$$

6. The outputs of each neuron OUT_{jl}.

$$OUT_{12} = sig(x_1 * w_{112} + x_2 * w_{212}) = sig(-0,44461749 - 0,08459352) = 0,3707$$

$$OUT_{22} = sig(x_1 * w_{122} + x_2 * w_{222}) = sig(-0,56216408 - 0,4040331) = 0,2756$$

$$OUT_{13} = sig(OUT_{12} * w_{113} + OUT_{22} * w_{213}) = sig(-0,10409256 - 0,05503732) = 0,4603$$

$$OUT_{23} = sig(OUT_{12} * w_{123} + OUT_{22} * w_{223}) = sig(-0,35835569 - 0,16841916) = 0,3713$$

7. Gradients ∇E_{jl}.

$$\nabla E_{13} = Y_1 * (1 - Y_1) * (Y_1 - D_1) = OUT_{13} * (1 - OUT_{13}) * (OUT_{13} - D_1)$$
$$= 0,4603 * (1 - 0,4603) * (0,4603 - 0,2030) = 0,0639$$

$$\nabla E_{23} = Y_2 * (1 - Y_2) * (Y_2 - D_2) = OUT_{23} * (1 - OUT_{23}) * (OUT_{23} - D_2)$$
$$= 0,3713 * (1 - 0,3713) * (0,3713 + 0,8542) = 0,2861$$

$$\nabla E_{12} = OUT_{12} * (1 - OUT_{12}) * (\nabla E_{13} * w_{113} + \nabla E_{23} * w_{123})$$
$$= 0,3707 * (1 - 0,3707) * (0,0639 * -0,2808 + 0,2861 * -0,9667)$$
$$= -0,0687$$

$$\nabla E_{22} = OUT_{22} * (1 - OUT_{22}) * (\nabla E_{13} * w_{213} + \nabla E_{23} * w_{223})$$
$$= 0,2756 * (1 - 0,2756) * (0,0639 * -0,1997 + 0,2861 * -0,6111)$$
$$= -0,0375$$

8. New weights W_{ijl}.

$$w_{112} = w_{112} - \nabla E_{12} * x_1 = 0,7599 + 0,0687 * -0,5851 = 0,7197$$

$$w_{212} = w_{212} - \nabla E_{12} * x_2 = 0,1452 + 0,0687 * -0,5826 = 0,1052$$

$$w_{122} = w_{122} - \nabla E_{22} * x_1 = 0,9608 + 0,0375 * -0,5851 = 0,9389$$

$$w_{222} = w_{222} - \nabla E_{22} * x_2 = 0{,}6935 + 0{,}0375 * -0{,}5826 = 0{,}6717$$

$$w_{113} = w_{113} - \nabla E_{13} * OUT_{12} = -0{,}2808 - 0{,}0639 * 0{,}3707 = -0{,}3045$$

$$w_{213} = w_{213} - \nabla E_{13} * OUT_{22} = -0{,}1997 - 0{,}0639 * 0{,}2756 = -0{,}2173$$

$$w_{123} = w_{123} - \nabla E_{23} * OUT_{12} = -0{,}9667 - 0{,}2861 * 0{,}3707 = -1{,}0727$$

$$w_{223} = w_{223} - \nabla E_{23} * OUT_{22} = -0{,}6111 - 0{,}2861 * 0{,}2756 = -0{,}6899$$

Answer:

1.

$$sig(x) = \frac{1}{1 + e^{-x}}$$

2.

$$OUT_{12} = sig(x_1 * w_{112} + x_2 * w_{212})$$

3.

$$\nabla E_{23} = Y_2 * (1 - Y_2) * (Y_2 - D_2)$$

4.

$$\nabla E_{22} = OUT_{22} * (1 - OUT_{22}) * (\nabla E_{13} * w_{213} + \nabla E_{23} * w_{223})$$

5.

$$w_{113} = w_{113} - \nabla E_{13} * OUT_{12}$$
$$w_{213} = w_{213} - \nabla E_{13} * OUT_{22}$$

6.

OUT_{12}	OUT_{22}	OUT_{13}	OUT_{23}
0.3707	0.2756	0.4603	0.3713

7.

∇E_{12}	∇E_{22}	∇E_{13}	∇E_{23}
-0.0687	-0.0375	0.0639	0.2861

8.

W_{112}	W_{212}	W_{122}	W_{222}	W_{113}	W_{213}	W_{123}	W_{223}
0.7197	0.1052	0.9389	0.6717	-0.3045	-0.2173	-1.0727	-0.6899

Assessment criteria:

1. Correctly recording the neuron activation formula, the formula for calculating the outputs of neuron N_{12}, the formula for calculating the gradient for the hidden layer neuron N_{22}, and the formula for calculating the correction of neuron N_{13} weights earns 3 points.

ONE CLICK TO OPEN ALL DOORS

od.globaluni.ru

2. Calculating the output of each OUT_{jl}. neuron and its error gradient is ∇E_{jl} earns 6 points.

3. Calculating all new W_{ijl} weights is worth 6 points.

Computer Science: Interdisciplinary Applications

Task 17
Entry level (1 point)

There is a public key cryptosystem with the following parameters: e_A is the public key of subscriber A; e_B is the public key of subscriber B; d_A is the private key of subscriber A; d_B is the private key of subscriber B. Subscriber A sends a message to subscriber B. With which key does subscriber A encrypt the message?

a) e_A
b) e_B
c) d_A
d) d_B

Answer: b.

Task 18
Entry level (1 point)

Consider an additive Fibonacci generator with the operation equation $X_n = (X_{n-1} + X_{n-2})$ mod 17. At a certain point, the generator first produced the number 9, and then the number 11. What number will appear next in the output?

a) 3
b) 5
c) 10
d) 13

Answer: a.

Task 19
Entry level (1 point)

What is the multiplicative inverse of 3 mod 17?

a) 1
b) 3
c) 6
d) 13

Answer: c.

Task 20
Intermediate level (4 points)

An organization's security officer suspects that there is an insider in the company, communicating with competitors through a steganographic channel. One of the employees became nervous when interviewed by the security officer. After intercepting the information flow from that employee's computer, the security officer realized that the employee in question was trying to send a distress signal. What is the name of the insider's accomplice?
The intercepted fragment in hexadecimal form is:

F4 F1 F4 FC F4 F9 F4 F3 F4 F5 F2 FC F4 F8 F4 F5 F4 FC F5 F0

Table 1. ASCII Table Fragment

hex	char		hex	char
2C	,		4D	M
2E	.		4E	N
41	A		4F	O
42	B		50	P
43	C		51	Q
44	D		52	R
45	E		53	S
46	F		54	T
47	G		55	U
48	H		56	V
49	I		57	W
4A	J		58	X
4B	K		59	Y
4C	L		5A	Z

Answer: ALICE

Task 21
Advanced level (15 points)

The round function of the block cipher takes a 16-bit value as input. The round function includes the following operations:
1) Performing an XOR between the original value and the key;
2) Splitting the resulting value into nibbles and replacing each nibble according to Table 2.

Table 2. Round Function Substitution Table

Input	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
Output	3	E	5	9	A	D	8	D	F	4	1	0	7	6	B	2

3) Performing a cyclic shift of the obtained value by 3 positions to the left.

What is the value of the lower (right) half of the encrypted message if the original message is 3D7B4AE3h and the upper (left) half of the encrypted message is 35FAh?

Please note that the evaluation will consider your problem-solving process; providing only the final answer is not sufficient.

Solution:

We begin by decrypting the upper part of the encrypted message.

Let us convert 35FAh to binary form:

$$35FAh = 0011\ 0101\ 1111\ 1010$$

and perform a cyclic shift to the right by 3 positions. Thus, we obtain:

$$0100\ 0110\ 1011\ 1111 = 46BFh$$

(A correctly performed shift is worth 3 points.)

Then we perform the inverse transformation, according to the table.

$$46BFh \rightarrow 9DE8h$$

(A correctly performed inverse substitution, according to the table, is worth 3 points.)

Then we convert the obtained value and the value of the upper half of the original message into binary form:

$9DE8h = 1001\ 1101\ 1110\ 1000$,

$37BDh = 0011\ 0111\ 1011\ 1101$,

and calculate their sum modulo 2 to obtain the key value.

$$1001\ 1101\ 1110\ 1000$$
$$xor$$
$$0011\ 0111\ 1011\ 1101$$
$$=$$
$$1010\ 1010\ 0101\ 0101$$

(A correctly inverted XOR operation is worth 3 points.)

(In total, correctly finding the key is worth $3 + 3 + 4 = 9$ points.)

Let us begin encrypting the lower half of the original text. First, we convert it into binary form:

$$4AE3h = 0100\ 1010\ 1110\ 0011$$

and add modulo 2 with the key

$$0100\ 1010\ 1110\ 0011$$
$$xor$$
$$1010\ 1010\ 0101\ 0101$$
$$=$$
$$1110\ 0000\ 1011\ 0110$$

(Correctly adding with the key is worth 2 points.)

Then we convert the obtained value into hexadecimal form:

$1110\ 0000\ 1011\ 0110 = E0B6h$

and perform the substitution according to the table:

$E0B6h \rightarrow B308h = 1011\ 0011\ 0000\ 1000$

(Correctly substituting according to the table is worth 2 points.)

All that remains is cyclically shifting the value to the left by 3 positions and obtaining the result

$1001\ 1000\ 0100\ 0101 = 9825h$

(A correctly performed shift earns 2 points.)

(In total, performing the round function with the known key is worth $2 + 2 + 2 = 6$ points)
Answer: 9825h

Computer Science: Software Engineering

Task 22
Entry level (1 point)

In which lifecycle model are all software requirements clearly defined before development begins?
- a) Spiral
- b) Waterfall
- c) Incremental
- d) Iterative

Answer: b.

Task 23
Entry level (1 point)

Which of the above is a methodology of "flexible" software development?
- a) V-Model
- b) Agile
- c) Waterfall
- d) All of the above

Answer: b.

Task 24
Entry level (1 point)

Which language is classified as declarative?
a) C, C++
b) Assembler, Pascal
c) PHP, Python
d) HTML, SQL

Answer: d.

Task 25
Intermediate level (3 points)

Which TWO stages does the software development life cycle include?
- a) Requirements analysis
- b) Testing
- c) Compilation
- d) Discussion

Answer: a, b.

<div align="center">

Task 26
Intermediate level (3 points)

</div>

Which TWO of the following actions does the user perform during the testing phase?

a) syntax debugging

b) test selection

c) defining the test method

d) determining the form for issuing results

Answer: b, c.

<div align="center">

Computer Science: Hardware and Architecture

Task 27
Entry level (1 point)

</div>

Which technologies are used to increase the reliability and speed of a hard disk drive?

a) increasing the capacity of the hard disk drive

b) reducing the number of revolutions of the hard disk drive

c) RAID

d) longitudinal recording to disk

Answer: c.

<div align="center">

Task 28
Entry level (1 point)

</div>

Which of the following statements are true about low-level formatting?

a) Low-level formatting is the layout of a disk into tracks and sectors.

b) Low-level formatting is the process of writing a file system structure to a disk that allows the disk to be used by the operating system to store programs and data.

c) Low-level formatting should be performed regularly.

d) Low-level formatting depends on the file system type.

Answer: a.

<div align="center">

Task 29
Entry level (1 point)

</div>

Which of the following statements are true about processor commands?

a) The command currently executed by the processor is located in the controller.

b) The instruction currently being executed by the processor is located in the processor status register.

c) Any command executed changes the value of the processor status register.

ONE CLICK TO OPEN ALL DOORS

od.globaluni.ru

d) A processor instruction has only one operand.

Answer: c.

Task 30
Intermediate level (3 points)

Which of the following statements align with the principles of Von Neumann architecture?

a) A program is written in memory as a sequential code.

b) A processor, having finished executing one instruction, moves on to the next.

c) The address of the next instruction is always the operand of the current instruction.

d) There is RAM for storing instructions and separate RAM for storing data.

e) Memory consists of numbered cells, any of which is accessible to the processor.

f) The binary-decimal number system is used to represent data.

Answer: a, b, e.

Task 31
Intermediate level (3 points)

Which of the following tasks does the Cache Controller perform?

a) providing quick access to intensively used data

b) preloading data into cache from RAM

c) loading data into RAM from the hard drive

d) lazy writing of data from cache memory to RAM

e) loading data into cache memory from a network resource

f) monitoring the execution of instructions by the processor

Answer: a, b, d.

Telecommunications

Task 32
Entry level (1 point)

Which information does the switch collect about connected devices based on the packet information?

a) The sender's IP address

b) The recipient's IP address

c) The sender's MAC address

d) The recipient's MAC address

Answer: c.

Task 33
Entry level (1 point)

Why is it necessary to specify the DNS server address?
>   a) to determine the MAC address from the IP address
>   b) to obtain the IP address of the resource
>   c) to determine the DNS name from the IP address
>   d) to obtain the DHCP server address

Answer: b.

Task 34
Entry level (1 point)

What is the ping utility used for?
>   a) To check whether the MAC address matches the IP address
>   b) To check the download speed of files
>   c) To test the presence of devices on the network
>   d) To create data backups

Answer: c.

Task 35
Intermediate level (3 points)

A router is a device that:
>   a) Connects networks.
>   b) Creates virtual networks.
>   c) Connects devices in a network.
>   d) Routes packets between networks.

Answer: a, d.

ONE CLICK TO OPEN ALL DOORS

od.globaluni.ru