

FI - Entrance exam - Computer Science

Jméno a příjmení - pište do okénka	Číslo přihlášky	Číslo zadání
		1

The computer science test consists of 30 questions, where you choose one of the possible answers a, b, c, d, e. Just one answer is correct. Each correctly answered question is valued by one point, an incorrectly answered question is valued -0.25. You get zero points for multiple selected answers or no answer.

Algorithms and Data Structures

1 Let $f(n) \in \mathcal{O}(n)$ and $g(n) \in \mathcal{O}(n^2)$. Which of the following statements is true?

- A $f(n) \in \mathcal{O}(g(n))$
- B $\frac{g(n)}{f(n)} \in \mathcal{O}(1)$
- C $g(n) \in \mathcal{O}(f(n))$
- *D $f(n)g(n) \in \mathcal{O}(n^3)$
- E $\frac{f(n)}{g(n)} \in \mathcal{O}(1)$

2 Which of the following statements is always true for every binary search tree (BST)?

- A The depth of the tree is logarithmic with respect to the number of all nodes.
- B The minimal key is stored in one of the leaves.
- C The median key is stored in the root.
- D Every internal node (i.e. not a leaf) contains exactly two children.
- *E The node with the maximal key does not have a right child.

3 Consider a directed graph. We run a depth-first search on this graph and the search algorithm assigns two numbers to each vertex v : $v.d$ is the discovery time of v , and $v.f$ is the finishing time of v . Which one of these statements is true in general?

- *A None of the other statements is true in general.
- B If there is a path from vertex u to vertex v in the graph then $u.f > v.f$.
- C If there is a path from vertex u to vertex v in the graph then $u.d < v.d$.
- D If there is a path from vertex u to vertex v in the graph then $u.d > v.d$.
- E If there is a path from vertex u to vertex v in the graph then $u.f < v.f$.

4 Consider a hash table with linear probing and the hash function $h(x) = (5 \cdot x + 3) \bmod 7$. We start with an empty hash table. In what order do we have to insert the values 1, 2, 5, 7, 9, 12 so that the contents of the resulting hash table look as follows: 2, 12, 1, 5, 7, empty field, 9?

- A 2, 9, 5, 12, 1, 7
- B 9, 2, 12, 5, 1, 7
- *C 9, 2, 12, 1, 5, 7
- D 9, 2, 5, 12, 1, 7
- E 2, 9, 12, 5, 1, 7

- 5** Which one of these statements is true?
- A** The worst-case time complexity of inserting an element into a hash table is in $\mathcal{O}(1)$.
 - B** We say that an algorithm is partially correct, if it produces a correct output for at least some of the inputs.
 - C** The B-tree is a special case of the binary search tree.
 - D** The logarithm and square root functions have the same asymptotic growth.
 - *E** There exists an $\mathcal{O}(n)$ algorithm that creates a binary heap from an unsorted array of elements.

Database Systems

- 6** Consider the following relational algebra expression:
 $documents \leftarrow \sigma_{created > '2020-01-01'}(documents) \cup \sigma_{fileSize > 0}(documents)$
 Choose the SQL expression that will do the exact same thing.
- A** UPDATE documents SET created = '2020-01-01' WHERE fileSize > 0;
 - B** INSERT INTO documents SELECT * FROM documents WHERE created > '2020-01-01' UNION SELECT * FROM documents WHERE fileSize > 0;
 - C** INSERT INTO documents VALUES (SELECT * FROM documents WHERE created > '2020-01-01' OR fileSize > 0);
 - D** DELETE FROM documents WHERE documents IN (SELECT * FROM documents WHERE created > '2020-01-01' UNION SELECT * FROM documents WHERE fileSize > 0);
 - *E** DELETE FROM documents WHERE created <= '2020-01-01' AND fileSize <= 0;

- 7** Transform the following E-R diagram to the relational model



Select the answer that specifies the exact relations that the resulting relational model should have. Note that the primary keys are underlined.

- *A** $array(array_id, name), disk(disk_id, capacity), contains(array_id, disk_id)$
- B** $array(array_id, name), disk(disk_id, capacity), contains(array_id, disk_id)$
- C** $array(array_id, name), disk(disk_id, capacity), contains(array_id, disk_id)$
- D** $array(array_id, name), disk(disk_id, capacity)$
- E** $array(array_id, name), disk(disk_id, capacity), contains(array_id, disk_id)$

- 8** Let us have a relation $menu(day_of_week, meal_number, meal_name, meal_price)$ and the following set of functional dependencies:

$day_of_week, meal_number \rightarrow meal_name$
 $meal_name \rightarrow meal_price$

We know that the relation is in the first normal form (1NF).

Select the statement that is true.

- A** Relation menu is in the third normal form (3NF) but not in the Boyce-Codd normal form (BCNF).
- *B** Relation menu is in the second normal form (2NF) but not in the third normal form (3NF).
- C** Relation menu is in the third normal form (3NF) but not in the fourth normal form (4NF).
- D** Relation menu is in the Boyce-Codd normal form (BCNF) but not in the third normal form (3NF).
- E** Relation menu is in the first normal form (1NF) but not in the second normal form (2NF).

- 9** Consider a simplified database for storing cook-book recipes with relations *recipe*(*id*, *name*, *description*) and *ingredient*(*recipe_id*, *placing_order*, *name*, *gram_weight*). The *recipe_id* attribute references the primary key of the *recipe* relation (foreign key). Choose the SQL expression that will display the name of each recipe that contains a word "cake" and the total sum of the gram weight of all its ingredients. For example, the result will contain tuple ("Birthday cake", 200) for a recipe named "Birthday cake" with two ingredients: 150g flour and 50g sugar.
- A** `SELECT name, gram_weight FROM ingredient
WHERE recipe_id IN (SELECT id FROM recipe WHERE LOWER(name) LIKE '%cake%');`
- B** `SELECT name, total_gram_weight FROM recipe WHERE LOWER(name) LIKE '%cake%'
COMPUTE total_gram_weight AS SUM(gram_weight) FROM ingredient WHERE recipe_id
= recipe.id;`
- C** `SELECT recipe.name, SUM(gram_weight) FROM recipe INNER JOIN ingredient ON
(recipe_id = recipe.id)
WHERE LOWER(recipe.name) LIKE '%cake%';`
- *D** `SELECT recipe.name, SUM(gram_weight) FROM recipe, ingredient
WHERE recipe_id = recipe.id AND LOWER(recipe.name) LIKE '%cake%' GROUP BY
recipe.id, recipe.name;`
- E** `SELECT name, SUM(gram_weight) FROM ingredient
WHERE recipe_id IN (SELECT id FROM recipe WHERE LOWER(name) LIKE '%cake%')
GROUP BY recipe_id, name;`
-

- 10** Consider an e-shop database that contains a relation *product*(*id*, *name*, *description*, *price*, *discount*). The relation is filled with a large number of products with prices ranging from 1 to 1 000 000 and discount values ranging from 0 to 0.99. Approximately 1 % of the products have a non-zero discount. The execution performance of the SQL command `SELECT name, price, discount FROM product WHERE price BETWEEN 10000 AND 20000 AND discount < 0.1;` can be best improved by:
- A** Creating a B+tree index on the attribute discount.
- B** Creating an extendible-hashing index on the attribute discount and another extendible-hashing index on the attribute price.
- *C** Creating a B+tree index on the attribute price.
- D** Creating a linear-hashing index on the attribute discount.
- E** Creating a linear-hashing index on the attribute price.
-

Computer Networks

- 11** The mechanism for slowing down the IPv4 address space exhaustion is:
- A** an autonomous system that divides Internet to smaller networks hierarchically.
- B** the Network Address Translation applied on each input and output domain router.
- C** the network renumbering to the finer classful addressing and the transition to IPv6.
- D** the use of Gigabit networks and highly-efficient connecting elements.
- *E** the classless addressing with the CIDR notation and the Network Address Translation.
-

- 12** Assume the physical layer of the ISO/OSI model. What is the primary function of it? What standards are relevant and how does it define addressing?
- A** the transmission of bytes of passed packets between a sender and a receiver; using RFC 1771, RFC 2918 and RFC 3649 standards; but it does not define any addressing.
 - B** the transmission of packets (the content of passed frames) between a sender and a receiver; using RS-232-C, CCITT V.24, CCITT X.21, IEEE 802.x standards; and addressing is by the physical (MAC) addresses.
 - C** the transmission of bits of passed frames among routers; using RFC 1058, RFC 1723 and RFC 2081 standards; but it does not define any addressing.
 - D** the transmission of bits of passed frames among a sender and receivers (multi-cast); using RFC 3649 standard; and a combination of MAC and IPv6 addresses is used for addressing.
 - *E** the transmission of bits of passed frames between a sender and a receiver; using RS-232-C, CCITT V.24, CCITT X.21, IEEE 802.x standards; but it does not define any addressing.
-

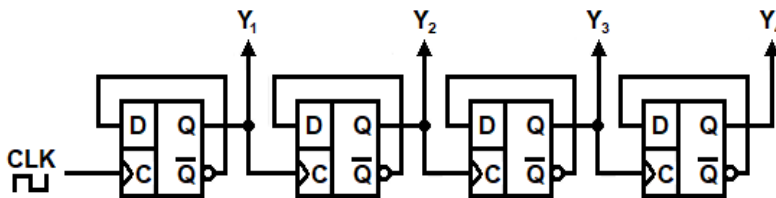
- 13** What are the advantages of networks with cycles, and how are the cycling packets eliminated?
- A** Robustness. The Backward Learning Algorithm prevents packet cycling and ensures the shortest packet delivery in all situations.
 - *B** Robustness. The Distributed Spanning Tree Algorithm prevents the creation of cycles, or the parameter TTL limits the number of hops of a packet.
 - C** Robustness is negligible due to large routing tables. When a routing cycle is detected, the corresponding interval in the routing table is split.
 - D** No advantages. Network topology must be hierarchical. When a cycling packet is detected by TTL (hop count), the link causing the cycle in routing is dropped.
 - E** Robustness. Hierarchical Routing defines a tree of routers that eliminates cycles in routing. When a link is down, a backup tree is used.
-

- 14** Which option is an example of the network protocol that includes a handshake, e.g., the three-way handshake, and describes why the handshake is important and what the handshake is responsible for within the selected protocol?
- *A** A complex transport protocol, e.g., TCP. Connection parameters are negotiated during the handshake. None or negative response during the handshake means the receiver is not able to receive any data.
 - B** TCP and UDP. The receiver sends a negative response during the handshake to terminate the connection. Otherwise, the sender starts sending data with the congestion window set to the maximum value.
 - C** UDP but only in combination with RTP and RTCP for the handshake. It sets the initial size of the congestion window to prevent congestion and allows the real-time communication.
 - D** All routing protocols when routing tables are transferred by UDP. The security of transfer is negotiated during the handshake.
 - E** The Multi Protocol Label Switching (MPLS). The handshake assigns a new label to the communicating nodes, and it ensures the labels are unique within the MPLS cloud.
-

- 15** Which option gives an example of media with collision domain on the Data Link Layer and how does it solve the collisions?
- A** In the local-area networks based on a bus or a ring topology and realized by a wired medium, the collisions are solved by a specific routing protocol based on the Distributed Spanning Tree Algorithm.
 - B** In the wide-area networks realized by optical fibers, the collisions are solved by the routing protocol BGP version 4.
 - *C** In the local-area networks based on a bus or a ring topology and connected by bridges, the collisions do not propagate through bridges, and a Medium Access Protocol (MAC) protocol is used.
 - D** In the wide-area networks realized by optical fibers, the collisions are solved by retransmissions in the intervals defined by system administrators.
 - E** In the local-area networks realized by a wireless medium, the collisions are solved by the routing protocols RIP and OSPF.

Computer systems

- 16** Consider the following logic circuit:



This circuit serves as:

- A** Shift register
- B** Parallel register
- C** Multiplexer
- *D** Frequency divider
- E** Frequency multiplier

- 17** The number 255 in the **base-8** number system can be expressed in the unpacked (zoned) binary-coded decimal (**BCD**) code as:
- A** 00000001 00000111 11010011
 - B** 00010111 00111101
 - *C** 11110001 11110111 11000011
 - D** 00010111 00111100
 - E** 11110001 11110111 11010011

- 18** Consider a hard disk drive with **four** active surfaces. This hard disk drive uses Zone Bit Recording (ZBR) and is divided into **five** zones. These zones have the following sectors count per track: Zone 0: 1024, Zone 1: 900, Zone 2: 800, Zone 3: 720, Zone 4: 652. Each zone contains 2048 tracks and each sector stores 4096 bytes. What is the capacity of this hard disk drive? (Note: 1 GB = 2^{30} B)
- *A** 128 GB
 - B** 256 GB
 - C** 32 GB
 - D** 64 GB
 - E** 192 GB

- 19** Consider a microprocessor which uses the paging mode for translating a 32-bit linear address into a 32-bit physical address. The linear address is split into two parts. The first part is of 10 most significant bits and serves as a pointer into the page directory to select a 10-bit page base address. The second part is of 22 least significant bits and is an offset. Which of the following statements is true?
- A The physical address space is 4 GB. The page size is 4 MB. The page directory can have up to 512 entries.
 - B The physical address space is 4 GB. The page size is 4 kB. The page directory can have up to 1024 entries.
 - C The physical address space is 8 GB. The page size is 2 MB. The page directory can have up to 1024 entries.
 - D The physical address space is 2 GB. The page size is 4 kB. The page directory can have up to 2048 entries.
 - *E The physical address space is 4 GB. The page size is 4 MB. The page directory can have up to 1024 entries.

- 20** Consider the processes P_1 , P_2 , P_3 , and P_4 and their time of arrival and duration given in the following table.

Process	Arrival time	Duration
P_1	0.0 s	7 s
P_2	2.0 s	4 s
P_3	4.0 s	1 s
P_4	5.0 s	4 s

What is the **average waiting time** if we use 1 CPU (i.e., 1 core) and the **preemptive Shortest Job First** scheduling algorithm?

- A 4 s
- *B 3 s
- C 5 s
- D 2 s
- E 7 s

Programming

21

```

read(n)
sum = 0
k = n
XXX {
    sum = sum + (n mod 10)
    n = n div 10
}
print sum

```

Let `mod` be the modulo operator and let `div` be the integer division operator. Assume that the input n is a positive integer. Which of the following should we place instead of "XXX" in the above code, so that it computes the sum of the decimal digits of n and is the best solution out of the provided ones?

- A `for i = 0 to k`
- B `for i = 1 to 10`
- C `while n == k`
- *D `while n > 0`
- E `while n < sum`

```

22 function fun1(unsigned integer n)
begin
    sum = 0
    for i = 1 to n do
        sum = <EXPR>
    end for
    return sum
end

function fun2(unsigned integer n)
begin
    sum = 0
    i = <INIT>
    while <COND> do
        sum = sum + i
        i = i + 1
    end while
    return sum
end

```

Assume an arbitrarily precise evaluation of expressions. Which of the following substitutions has to be used so that the functions fun1 and fun2 return a **different** result for at least one valid input (unsigned integer)?

- *A <EXPR> = sum + i - 1; <INIT> = 0; <COND> = i <= n
- B <EXPR> = sum + i; <INIT> = 1; <COND> = i <= n
- C <EXPR> = 3 * n; <INIT> = n - 1; <COND> = i <= n + 1
- D <EXPR> = sum + 1; <INIT> = n; <COND> = i <= n
- E <EXPR> = sum + i - 1; <INIT> = 0; <COND> = i < n

23 Which of the following three statements I, II, and III are true (in common languages such as C++, Java, C#)? Choose the option that contains exactly all the true statements (and none of the false ones).

- I. Local variables of functions are allocated on the heap.
- II. Function calls are implemented using the stack.
- III. If an exception is caught (in a catch block), it can be re-thrown (using throw).

- A I, II
- *B II, III
- C I, III
- D III
- E I, II, III

24 Which statement is generally true in common OOP languages such as C++, Java, C#?

- A If a class B inherits from a class A, the instances of B can access all attributes (member variables) of A.
- B The difference between static and non-static methods (member functions) is that the static methods may access static attributes (member variables) of a class.
- *C If a class B inherits from a class A (via public inheritance), every instance of B is considered to also be an instance of A.
- D The notions "class" and "instance" mean the same thing.
- E If the late binding (virtual method calls) is used, the actual method to be called is decided by the compiler at the compile time.

25 Which statement is **false**?

- A** In purely functional languages, functions cannot have any side effects.
 - B** A recursive function can always be rewritten iteratively.
 - C** The lazy evaluation strategy in functional programming allows working with infinite data structures.
 - D** A tail-recursive function can always be rewritten iteratively.
 - *E** When using the call-by-reference, the change of a parameter value inside a function can not be observed outside the function.
-

Software engineering

26 As a software engineer you shall choose the best fitting notation to describe the purchase process within an eshop application. Will you choose UML Use Case diagram, UML Activity diagram or UML State diagram and why?

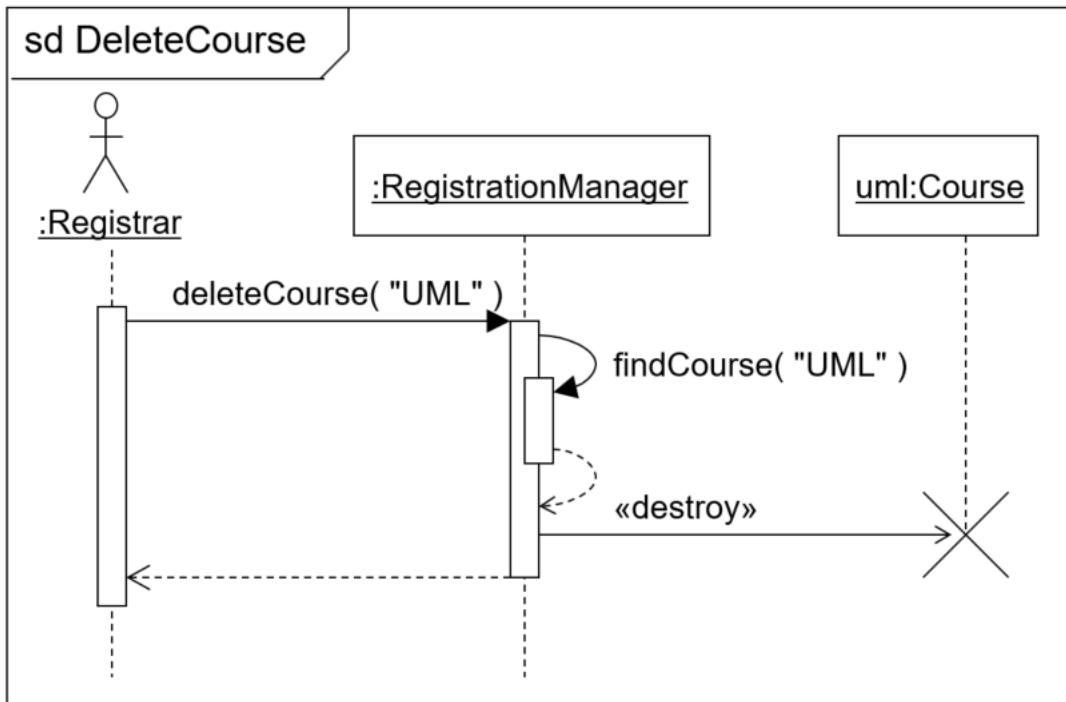
- A** I will choose UML State diagram because it is best suited for high-level process descriptions, which is what is needed in case of the purchase process.
 - B** I will choose UML State diagram because purchase is a complex object with many different states and there is just one actor involved.
 - *C** I will choose UML Activity diagram because it is well suited for the description of processes involving different system components and actors.
 - D** I will choose UML Use Case diagram because it is best suited for high-level process descriptions, which is what is needed in case of the purchase process.
 - E** I will choose UML Activity diagram because it offers the best mechanism to describe data transformations happening along the purchase process.
-

27 Match the descriptions of architectural patterns with their correct name:

- I. Architecture X decomposes the system into layers interacting over defined interfaces, which enables the architect to replace a layer with a newer implementation of it.
- II. Architecture Y composes the system of mutually independent components realizing certain functionality used by other components.
- III. Architecture Z is often used in web applications to separate the application logic (including data) from the presentation layer and the coordination between the two via a controller.

- A** X = pipe and filter architecture, Y = component architecture, Z = layered architecture
 - B** X = pipe and filter architecture, Y = service-oriented architecture, Z = layered architecture
 - *C** X = layered architecture, Y = service-oriented architecture, Z = model-view-controller architecture
 - D** X = layered architecture, Y = client-server architecture, Z = model-view-controller architecture
 - E** X = layered architecture, Y = service-oriented architecture, Z = component architecture
-

28 Which one of the following descriptions best corresponds to the UML sequence diagram in the following figure?



- A Within the deleteCourse() method execution, the RegistrationManager class performs two actions in parallel, namely a call of the findCourse() method and deletion of the uml object of the Course type.
- *B Within the deleteCourse() method execution, an instance of the RegistrationManager class calls the findCourse() method on itself, and then triggers the deletion of the uml object of the Course type.
- C Within the modelled scenario, two methods are executed in a sequence, namely: deleteCourse(), findCourse(). That by itself invalidates the uml object of the Course class, which is deleted in effect of that.
- D Within the deleteCourse() method execution, an instance of the RegistrationManager class calls the findCourse() method on itself, within which (inside the findCourse() execution) the uml object of the Course type is deleted.
- E Within the modelled scenario, these three methods are performed one after the other, namely: deleteCourse(), findCourse(), destroy().

29 Consider the following statements about automated and manual testing:

- I. The newly developed feature, which is still inaccurately defined and changing, is cheaper to test manually.
- II. Regression testing is cheaper when automated, than done manually.
- III. Automated testing is always cheaper and so it shall be always preferred over manual testing.
- IV. Performance and stress testing is cheaper when done manually, than automated.

Which of those statements are true?

- A Only II. and III., not I. and IV.
- B Only I. and IV., not II. and III.
- *C Only I. and II., not III. and IV.
- D Neither of them.
- E Only II., III. and IV., not I.

- 30** Refactoring is an activity that might have multiple effects. Which one of the following effects should **not** result from refactoring?
- A** Decomposition of complex classes to multiple simpler ones.
 - B** Compliance with S.O.L.I.D. principles.
 - *C** Implementation of new user-defined functions.
 - D** Change of system architecture.
 - E** Increase in system understandability.
-