



**28 ASSISTENTI - PROFILO TECNICO NEL CAMPO DELL'ICT**

Lettera C del bando del 20 novembre 2023

**Testo n. 2**

*Un quesito a scelta - tra due proposti dalla Commissione - su tre materie scelte dal candidato tra le quattro materie del programma*

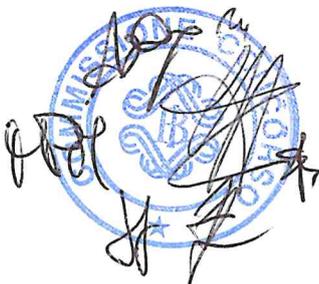
**BASI DI DATI E DATA SCIENCE**

**QUESITO N. 1**

- A. Si descrivano le principali forme normali di uno schema di una base di dati relazionale, indicandone vantaggi e svantaggi.
- B. Si intende realizzare una base di dati per gestire le vendite di un centro commerciale. Si consideri che:
- i negozi sono identificati da un codice e hanno un nome;
  - le categorie di prodotto sono identificate da un codice e hanno una descrizione;
  - i prodotti sono identificati da un codice, appartengono ad una categoria e hanno un nome;
  - ogni prodotto appartiene a una e una sola categoria;
  - ogni negozio vende tutti i prodotti di una o più categorie;
  - ogni categoria di prodotti può essere venduta da uno o più negozi;
  - ogni negozio attribuisce un prezzo di vendita a ciascun prodotto, che non varia nel tempo;
  - è mantenuto un registro delle vendite in cui per ciascuna vendita sono riportati: negozio, prodotto, quantità, data.

Sulla base di tali informazioni:

1. si rappresenti il diagramma Entità-Relazione (E-R);
  2. si fornisca un corrispondente schema logico relazionale che rispetta la forma normale di Boyce e Codd;
  3. si formuli un'interrogazione in linguaggio SQL che ordina i negozi in base al fatturato del mese di gennaio 2024.
- C. Si descriva come utilizzare gli indici per ottimizzare le prestazioni delle interrogazioni evidenziando vantaggi e rischi della loro applicazione.



## QUESITO N. 2

- A. Nell'ambito dell'attività di preparazione dei dati per l'addestramento di un algoritmo di *machine learning*, si discuta il tema della riduzione dimensionale.
- B. Un'azienda proprietaria di una catena di supermercati affida ai propri *data scientist* il compito di raggruppare i clienti in *cluster* omogenei sulla base degli acquisti effettuati nel 2023.

A tal fine, fornisce loro un *dataset* contenente il dettaglio degli scontrini emessi, di cui è riportato un estratto.

Tessera fedeltà	Data scontrino	Categoria prodotto	Quantità
A112124	2023-10-07	Biscotti	2
A112124	2023-12-15	Biscotti	4
A154545	2023-10-06	Pane	1
A154545	2023-11-22	Acqua	3
A154545	2023-10-12	Latte	2
A274741	2023-10-30	Latte	1

1. Si identifichi il problema di *machine learning* in cui ricade lo scenario;
  2. si descriva il principio di funzionamento di un algoritmo applicabile alla risoluzione del problema;
  3. si descrivano e si applichino allo scenario le attività di preparazione del dato per costruire il *dataset* di addestramento, nell'ipotesi che il modello da addestrare accetti solo *feature* numeriche.
- C. Si illustrino gli ambiti di applicazione, i vantaggi e gli svantaggi della tecnica di *feature normalization*.



## PROGETTAZIONE E SVILUPPO DEL SOFTWARE

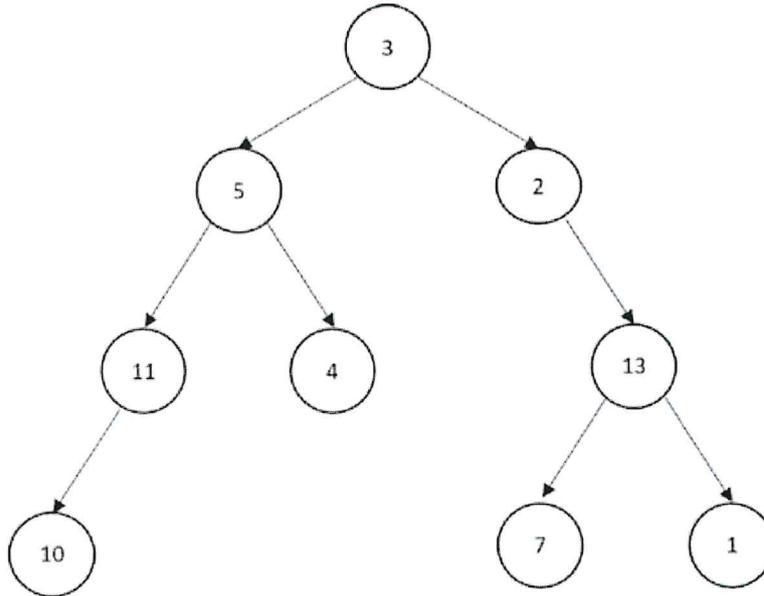
### QUESITO N. 3

- A. Si descrivano i concetti fondamentali della programmazione *object-based* e se ne esponano differenze e vantaggi rispetto al paradigma della programmazione procedurale.
- B. Si descrivano i concetti di ereditarietà e polimorfismo che i linguaggi di programmazione *object-oriented* supportano in aggiunta a quelli del paradigma *object-based*, illustrandone rischi e vantaggi.
- C. Due *client* (produttori) accodano messaggi in memoria che sono consumati in ordine di arrivo da un terzo *client* (consumatore). I messaggi prodotti dal primo *client* (ProducerA) sono composti da un identificativo univoco e da un testo in chiaro; quelli prodotti dal secondo *client* (ProducerB) sono composti da un identificativo univoco e da un testo cifrato. Il terzo *client* (Consumer) può consumare indifferentemente i messaggi provenienti da entrambi i produttori.
1. Si disegnino i diagrammi UML delle classi e di collaborazione;
  2. si scrivano in pseudocodice orientato agli oggetti i metodi di *enqueue* e *dequeue* dei messaggi, nell'ipotesi in cui la coda venga implementata tramite *buffer* circolare realizzato con *array*; si trascurino problemi di concorrenza e si sollevi un'eccezione in caso di scrittura su coda piena e di lettura da coda vuota.



#### QUESITO N. 4

- A. Si definisca la struttura dati albero binario e si illustri come può essere rappresentata.
- B. Si descrivano gli algoritmi di visita di un albero binario *in-order*, *pre-order* e *post-order* e li si applichino all'albero binario mostrato in figura per determinare le corrispondenti sequenze di visita dei nodi.



- C. Si estenda la definizione di albero binario per descrivere la struttura dati albero binario di ricerca (*Binary Search Tree*, BST).
- D. Si discuta la complessità temporale asintotica della ricerca di un nodo all'interno di un BST; si descriva quindi in cosa consiste il bilanciamento di un BST e in che modo tale operazione consenta di ridurre il costo della ricerca di un nodo.



## ARCHITETTURA DEI SISTEMI INFORMATICI E DELLE RETI

### QUESITO N. 5

- A. Dopo aver illustrato la differenza tra programma e processo, si descriva in quali sezioni è tipicamente organizzato il *layout* della memoria associata a un processo.
- B. Nel contesto della gestione dei processi:
1. si descriva la struttura dati *process control block* (PCB) e si illustri come tale struttura è utilizzata dal sistema operativo;
  2. si descrivano l'operazione di *context switch* e le principali fasi in cui si articola.
- C. Dopo aver brevemente introdotto il concetto di *system call*, si descriva la differenza tra le *system call* `fork()` ed `exec()` e si determini un possibile *output* dell'esecuzione del seguente programma in linguaggio C.

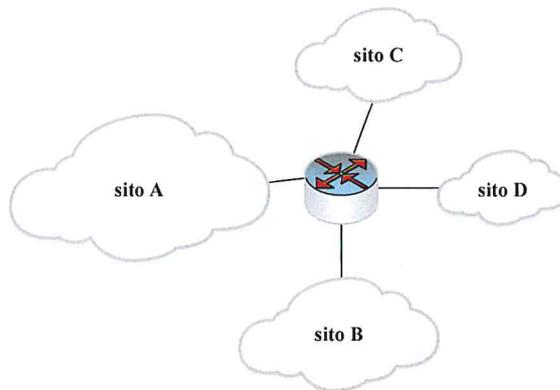
```
#include <stdio.h>
#include <unistd.h>
int main()
{
    pid_t pid;

    pid = fork();
    if (pid == 0) {
        execlp("/bin/echo", "echo", "ciao", NULL);
    } else {
        pid = fork();
        if (pid == 0) {
            printf("hello\n");
        }
    }
    printf("hola\n");
    return 0;
}
```



## QUESITO N. 6

- A. Si definisca, nell'ambito del protocollo IPv4, la struttura dello spazio di indirizzamento, facendo riferimento alla definizione di *subnet mask* e alla suddivisione in classi.
- B. Un'azienda possiede uno spazio di indirizzi IP di classe C (192.168.128.0/24) da utilizzare per la propria *Local Area Network* (LAN), costituita da 4 segmenti di rete (siti) interconnessi attraverso un *router*:
- il sito A dovrà consentire il collegamento di 30 *host*;
  - i siti B e C dovranno consentire il collegamento di 12 *host* ciascuno;
  - il sito D dovrà consentire il collegamento di 4 *host*.



1. Si definiscano le *subnet* da assegnare ai vari siti utilizzando, se possibile, l'indirizzamento IP a maschera fissa;
2. si descriva e si applichi il meccanismo del *Variable Length Subnet Mask* (VLSM) per ottimizzare l'indirizzamento dei 4 siti.



## SICUREZZA DEI SISTEMI INFORMATICI E DELLE RETI

### QUESITO N. 7

- A. Si descrivano le diverse tipologie di fattori di autenticazione, riportando almeno un esempio per ciascuna di esse. Si illustri quindi come è possibile combinare tali fattori per realizzare diversi livelli di autenticazione.
- B. Si consideri lo scenario in cui l'autenticazione degli utenti per l'accesso a un servizio web esposto su Internet è effettuata mediante inserimento in un form web di username (email) e password.
1. Nell'ambito degli attacchi volti alla compromissione di credenziali che potrebbero interessare tale servizio, si illustrino caratteristiche e differenze delle tecniche di attacco *online* e di attacco *offline*, soffermandosi su *password guessing*, *password spraying* e *rainbow table attack*;
  2. si discutano i presidi di sicurezza che è opportuno implementare per mitigare l'efficacia delle tecniche di attacco citate al punto precedente; si descriva, in particolare, come le tecniche di *hashing* e di *salting* possano essere utilizzate per realizzare una soluzione di memorizzazione delle credenziali sicura.
- C. Si descriva la vulnerabilità nota come *user enumeration*; si illustrino le tipologie di attacco che un attore malevolo potrebbe perpetrare sfruttando tale vulnerabilità e quali accorgimenti sia necessario adottare per mitigarne il rischio.



## **QUESITO N. 8**

- A. Si descriva il funzionamento dei cifrari a sostituzione e si illustri la tecnica di crittoanalisi più opportuna per romperne la cifratura.
- B. Si illustrino le caratteristiche del c.d. “cifrario perfetto”, evidenziando quali proprietà lo rendono immune da ogni tipo di attacco di crittoanalisi e quali sono i limiti nel suo reale utilizzo.
- C. Gli utenti A e B devono instaurare una comunicazione sicura sotto il profilo della confidenzialità su un canale non sicuro.
1. Si descriva un metodo che permetta ad A e B di raggiungere l’obiettivo in presenza di un attore malevolo in grado solo di ascoltare e registrare i messaggi;
  2. assumendo che l’attore malevolo possa inserirsi attivamente nella comunicazione tra A e B, sottraendo e inserendo messaggi, si determini se la soluzione proposta al punto precedente si possa considerare ugualmente valida.

## **PROVA IN LINGUA INGLESE**

2020 was the year of the “staycation”. A staycation is a holiday that someone spends in their own country or at home, rather than travelling abroad. In your opinion, what are the pros and cons of a staycation?

