GATE And Tech

Daily Quiz Questions For GATE CSE 2025

GATE And Tech
(Lakshman Patel)

January 21, 2025

1 Discrete Mathematics (30 Questions)

1. Consider the following expressions:

   (i) False

   (ii) Q

   (iii) True

   (iv) $\neg P \vee Q$

   (v) $Q \vee P$

   The number of expressions given above that are logically implied by $Q \wedge (Q \Rightarrow P)$ is ________. (Numerical Answer Type)

   Solution:

   Correct Answer: 4

2. Consider the following statement:

   If I get a perfect score on this test, then I am a genius.

   Which of the following is/are logically not equivalent to the statement above? (Multiple Select Question)

   A. If I am a genius, then I will get a perfect score on this test.

   B. If I am not a genius, then I won't get a perfect score on this test.

   C. Being a genius is sufficient for my getting a perfect score on this test.

   D. Being a genius is necessary for my getting a perfect score on this test.

   Solution: A is the converse which is not equivalent. B is the contrapositive so it is equivalent. C is also the converse so it is not equivalent. D is equivalent. So only B and D are equivalent.

   Correct Answer: A; C

3. Let

   - m = "Juan is a math major,"
   - c = "Juan is a computer science major,"

- g = "Juan's girlfriend is a literature major,"
- h = "Juan's girlfriend has read Hamlet," and
- t = "Juan's girlfriend has read The Tempest."

Which of the following expresses the statement "Juan is a computer science major and a math major, but his girlfriend is a literature major who hasn't read both The Tempest and Hamlet."

A. $c \wedge m \wedge (g \vee (\sim h \vee \sim t))$

B. $c \wedge m \wedge g \wedge (\sim h \wedge \sim t)$

C. $c \wedge m \wedge g \wedge (\sim h \vee \sim t)$

D. $c \wedge m \wedge (g \vee (\sim h \wedge \sim t))$

Solution:

Correct Answer: C

4. The function $((p \vee (r \vee q)) \wedge \sim (\sim q \wedge \sim r)$ is equal to the function

A. $q \vee r$

B. $((p \vee r) \vee q)) \wedge (p \vee r)$

C. $(p \vee q) \wedge \sim (p \vee r)$

D. $(p \wedge r) \vee (p \wedge q)$

Solution:

Correct Answer: A

5. The truth table for $(p \vee q) \vee (p \wedge r)$ is the same as the truth table for

A. $(p \vee q) \wedge (p \vee r)$

B. $(p \vee q) \wedge r$

C. $(p \vee q) \wedge (p \wedge r)$

D. $p \vee q$

Solution:

Correct Answer: D

6. The Boolean function $[\sim (\sim p \wedge q) \wedge \sim (\sim p \wedge \sim q)] \vee (p \wedge r)$ is equal to the Boolean function

A. q

B. $p \wedge r$

C. $p \vee q$

D. p

Solution:

Correct Answer: D

7. Which of the following functions is the constant 1 function?

A. $\sim p \vee (p \wedge q)$

B. $(p \wedge q) \vee (\sim p \vee (p \wedge \sim q))$

C. $(p \wedge \sim q) \wedge (\sim p \vee q)$

D. $((\sim p \wedge q) \wedge (q \wedge r)) \wedge \sim q$

Solution:

Correct Answer: B

8. Consider the statement, "Either $-2 \leq x \leq -1$ or $1 \leq x \leq 2$." The negation of this statement is

A. $x < -2$ or $2 < x$ or $-1 < x < 1$

B. $x < -2$ or $2 < x$

C. $-1 < x < 1$

D. $x \leq -2$ or $2 \leq x$ or $-1 < x < 1$

Solution:

Correct Answer: A

9. The truth table for a Boolean expression is specified by the correspondence $(P, Q, R) \rightarrow S$ where $(0,0,0) \rightarrow 0, (0,0,1) \rightarrow 1, (0,1,0) \rightarrow 0, (0,1,1) \rightarrow 1, (1,0,0) \rightarrow 0, (1,0,1) \rightarrow 0, (1,1,0) \rightarrow 0, (1,1,1) \rightarrow 1$. A Boolean expression having this truth table is

A. $[(\sim P \wedge \sim Q) \vee Q] \vee R$

B. $[(\sim P \wedge \sim Q) \wedge Q] \wedge R$

C. $[(\sim P \wedge \sim Q) \vee \sim Q] \wedge R$

D. $[(\sim P \wedge \sim Q) \vee Q] \wedge R$

Solution:

Correct Answer: D

10. Which of the following statements is FALSE:

A. $(P \wedge Q) \vee (\sim P \wedge Q) \vee (P \wedge \sim Q)$ is equal to $\sim Q \wedge \sim P$

B. $(P \wedge Q) \vee (\sim P \wedge Q) \vee (P \wedge \sim Q)$ is equal to $Q \vee (P \wedge \sim Q)$

C. $(P \wedge Q) \vee (\sim P \wedge Q) \vee (P \wedge \sim Q)$ is equal to $[(P \vee \sim P) \wedge Q] \vee (P \wedge \sim Q)$

D. $(P \wedge Q) \vee (\sim P \wedge Q) \vee (P \wedge \sim Q)$ is equal to $P \vee (Q \wedge \sim P)$

Solution:

Correct Answer: A

11. For each of the following sentences in first-order logic, is a valid statement?:

A. $P(A) \Rightarrow \forall x P(x)$

B. $P(A) \Rightarrow \forall x \neg P(x)$

C. $P(A) \Rightarrow \exists x P(x)$

D. $P(A) \Rightarrow \exists x \neg P(x)$

Solution:

A. Satisfiable but not valid.

B. Satisfiable but not valid.

C. Valid.

D. Satisfiable but not valid.

Correct Answer: C

12. Which of the following statements is/are correct?

S1: If $(\neg\alpha \vee \beta)$ is valid, then $\alpha \models \beta$.

S2: Is $\neg\exists x \neg P(x) \Rightarrow \forall x P(x)$ valid?

A. S1

B. S2

C. Both

D. None

Solution:

S1: $(\neg\alpha \vee \beta)$ means that in all models (truth assignments to the symbols in α and β, which are sentences like $(A \wedge B)$ or whatever), $(\neg\alpha \vee \beta)$ is true. That is, no matter what I plug in (true or false) for the symbols in this sentence, I'll get true for the whole thing. Now, we decompose...

By the defined semantics of implication, we know that in all model, either α is false or β is true (1). Also, sentences are either true or false given a model. So, for any model M in that set of all models, either α is false or true. If it's true, then we know that β is true by (1) above. We have shown that for all models where α is true, β is true. This is the definition of $\alpha \models \beta$, so we're done.

S2: This prove is a bit complex, so I'm going to write it rather formally. This degree of formality is neither necessary nor particularly desired, but it's instructive. The above sentence in FOL is valid. To prove it, we want to show that for any model (a set of objects and an assignment of true or false to all $P(x)$ for any object x), this sentence is true. We'll do so by appealing only to the definitions of the semantics logical connectives and quantifiers.

- For any model M, either $\neg\exists x \neg P(x)$ is true or false. All sentences are either true or false given a model. This is true of any sentence and model in FOL.
- If $\neg\exists x \neg P(x)$ is true given M, then $\exists x \neg P(x)$ is false.
- If $\exists x \neg P(x)$ is false, then there are no objects in M such that $\neg P(x)$ is true.
- Thus, $P(x)$ must be true for all objects in M.
- But this is the definition for the semantics of $\forall x P(x)$, so we have proved that for any model in which $\neg\exists x \neg P(x)$ is true, $\forall x P(x)$ is true.
- Therefore, $\neg\exists x \neg P(x) \Rightarrow \forall x P(x)$ must be true in all models and therefore valid.

Correct Answer: C

13. We use the dictionary:

M	domain of men
s	Sharon
$N(x)$	x is nice
$L(x,y)$	x loves y

Which of the following formulas corresponds to the sentence: There is a nice man who loves Sharon.

A. $\exists x \in M[N(x) \land L(x,s)]$

B. $\exists x \in M[N(x) \to L(x,s)]$

C. $\forall x \in M[N(x) \land L(x,s)]$

D. $\forall x \in M[N(x) \to L(x,s)]$

Solution:

The second answer makes no sense because this formula does not state that there actually exists a nice man. The third answer makes no sense because this formula states that all men are nice and all men love Sharon.

The fourth answer makes no sense because this formula states that all nice men love Sharon.

Correct Answer: A

14. Which of the following formulas does not express that there is at most one element of D that has property $P(x)$?

A. $\forall x_1, x_2 \in D[P(x_1) \land P(x_2) \to x_1 = x_2]$

B. $\forall x_1, x_2 \in D[x_1 \neq x_2 \to \neg P(x_1) \lor \neg P(x_2)]$

C. $\neg\exists x_1, x_2 \in D[x_1 \neq x_2 \land P(x_1) \land P(x_2)]$

D. $\neg\exists x_1, x_2 \in D[P(x_1) \land x_1 \neq x_2 \to \neg P(x_2)]$

Solution:

The combination of the $\exists$ and the $\to$, it is hard to explain what this formally really means. It is easier to state why the other three are correct:

- The first one states that if there are two elements having property P, then they must be the same.
- The second one states that if there are two different elements, then they cannot both have property P.

- The third one states that there are no two different elements that both have property P.

So in all cases, there cannot be two different ones, but it is possible that there exists no element with property P. Hence there is at most one element having property P.

Correct Answer: D

15. Consider the model (men, women, parent_of), and an interpretation in this model:

M	domain of men
W	domain of women
$P(x,y)$	y is a parent of x

You may assume that both parents of all people in $(M \cup W)$ are also in $(M \cup W)$.

Which of the following formulas is not true in this model?

A. $\forall x \in (M \cup W) \exists y_1 \in M \exists y_2 \in W\,[P(y_1,x) \wedge P(y_2,x)]$

B. $\forall x \in W \exists x', x'' \in W\,[P(x,x') \wedge P(x,x'')]$

C. $\forall x \in W \exists x', x'' \in W\,[P(x,x') \wedge P(x',x'')]$

D. $\forall x \in (M \cup W) \exists y \in (M \cup W)[P(x,y) \vee P(y,x)]$

Solution:

This states that every person is the parent of (at least) one man and one woman, which is certainly not true in this model.

The second formula seems to state that every woman has two mothers, however, these 'two' mothers may be the same. So it actually states that every woman has at least one mother.

The third formula states that every woman has a mother and a grandmother.

The fourth formula states that for all persons there is a person who is either his or her parent or his or her child. The first part of this disjunction holds. The second part does not.

Correct Answer: A

16. For which $n, m \geq 1$ is $K_{n,m}$ a tree?

A. For $n = 1$ and any m, or $m = 1$ and any n.

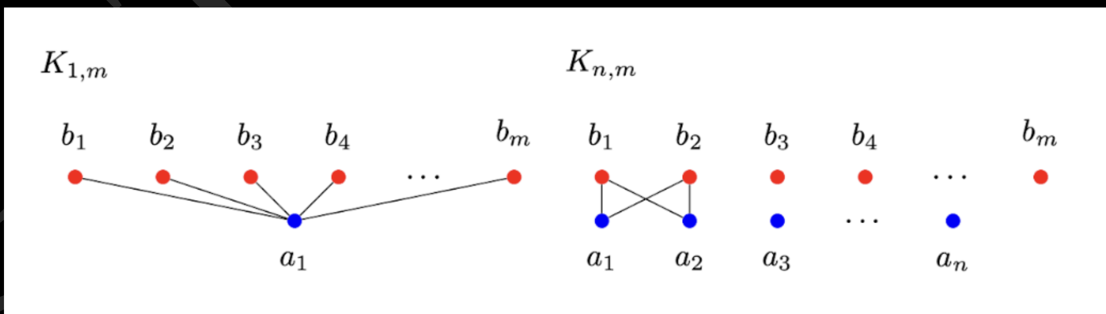
B. For $n = 1$ and $m = 1$.

C. For $n = m$.

D. This is never a tree.

Solution:

These are the (partial) graphs $K_{1,m}$ and $K_{n,m}$ where $n \geq 2$:



It is clear that $K_{1,m}$ is connected and has no cycles, so it is a tree. However, although most of the edges in $K_{n,m}$ are not drawn, we see a cycle $a_1 \to b_2 \to a_2 \to b_1 \to a_1$, so it cannot be a tree.

For symmetry reasons, the results also follow for $K_{n,1}$ and $K_{n,m}$ with $m \geq 2$.

Correct Answer: A

17. We have a proof by induction that shows that a predicate $P(n)$ holds for all n starting at zero. This proof follows the standard scheme, in which the base case just is about $P(0)$. What from this proof is used to establish that $P(3)$ holds?

A. The proof of the base case, and the induction steps for $k = 0, k = 1$ and $k = 2$.

B. The proof of the base case, and the induction steps for $k = 0, k = 1$, $k = 2$ and $k = 3$.

C. The base case and the induction step for $k = 2$.

D. The base case and the induction step for $k = 3$.

Solution:

We need the base case to prove $P(0)$. Then, we use the induction step for $k = 0$ to prove $P(1)$ from $P(0)$. Next, we use the induction step for $k = 1$ to prove $P(2)$ from $P(1)$. Finally, we use the induction step for $k = 2$ to prove $P(3)$ from $P(2)$. It is clear that none of the other options can be correct at the same time.

Correct Answer: A

18. We want to count the number of ways that one can divide nine distinguishable objects into four non-distinguishable (possibly empty) groups. What can we best use for this?

A. Binomial coefficients.

B. Stirling numbers of the first kind.

C. Stirling numbers of the second kind.

D. Bell numbers.

Solution:

The Stirling number of the second kind $\left\{ \begin{matrix} n \\ m \end{matrix} \right\}$ is the number of ways to divide n distinguishable objects over m indistinguishable groups. In this case, we want to divide nine distinguishable elements over four non-distinguishable groups, which are possibly empty.

Therefore, we have to count the number of ways to divide nine distinguishable elements over one group, two groups, three groups, and four groups. So the answer here would be

$$\left\{ \begin{matrix} 9 \\ 1 \end{matrix} \right\} + \left\{ \begin{matrix} 9 \\ 2 \end{matrix} \right\} + \left\{ \begin{matrix} 9 \\ 3 \end{matrix} \right\} + \left\{ \begin{matrix} 9 \\ 4 \end{matrix} \right\} = 1 + 255 + 3025 + 7770 = 11051$$

Correct Answer: C

19. We want to count the number of ways that one can select a non-empty selection of at most four objects out of nine distinguishable objects. What can we best use for this?

A. Binomial coefficients.

B. Stirling numbers of the first kind.

C. Stirling numbers of the second kind.

D. Bell numbers.

Solution:

The binomial coefficient $\binom{n}{m}$ is the number of ways to select m objects from n distinguishable objects.

In this case, we want to select at most four out of nine distinguishable objects. So the answer here would be

$$\binom{9}{1}+\binom{9}{2}+\binom{9}{3}+\binom{9}{4}=9+36+84+126=254$$

Correct Answer: A

20. The complete bipartite graphs $K_{m,n}$ with $m,n \geq 2$ have a Hamiltonian circuit if and only if ...

A. $m = n$.

B. m and n are both even.

C. m is even, n is even, or both.

D. None of the above.

Solution:

In a Hamiltonian circuit, one has to visit all vertices exactly once and finish in the starting vertex.

Recall that $K_{m,n}$ is a complete bipartite graph with m red vertices and n blue vertices, where there are edges between all combinations of red and blue vertices, but there are never edges between vertices of the same color.

So if we have a Hamiltonian circuit, we may assume that we start in a red vertex, followed by a blue vertex, followed by a red vertex, and so on, ending in the same red vertex that we started from.

So this means that $m+n$ must be even, but also that $m=n$ since the vertices in the circuit change color every step.

Hence if such a Hamiltonian circuit exists, then $m=n$ must hold.

And if we assume that $m=n$, we have m red vertices and m blue vertices. So we may give them the names $r_1, r_2, \ldots, r_m$ and $b_1, b_2, \ldots, b_m$. And we can construct the following Hamiltonian circuit:

$$r_1 \to b_1 \to r_2 \to b_2 \to r_3 \to \cdots b_{m-1} \to r_m \to b_m \to r_1$$

The second answer makes no sense because if $m=2$ and $n=4$ you cannot put them all in a cycle where the color changes each step. The third answer makes no sense because having one of m and n or even both of them to be even is not strong enough: $(m=2$ and $n=3), (m=2$ and $n=4), (m=3$ and $n=2)$, and $(m=4$ and $n=2)$ are all examples where you cannot put all vertices in a cycle where the color changes each step.

The fourth answer makes no sense because one of the answers above is correct.

Correct Answer: A

21. The sum of the numbers in any row of Pascal's triangle is ...

A. A power of two.

B. The factorial of a number.

C. A Bell number.

D. None of the above.

Solution:

On the $(n+1)^{\text{th}}$ row of Pascal's triangle, we have the binomial coefficients

$$\binom{n}{0}, \binom{n}{1}, \binom{n}{2}, \ldots, \binom{n}{n}$$

Recall that Newton's Binomial Theorem gives

$$(1+x)^n = \binom{n}{0} + \binom{n}{1}x + \binom{n}{2}x^2 + \cdots + \binom{n}{n}x^n$$

Now if we substitute $x = 1$ we get

$$(1+1)^n = \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n}$$

And hence the sum of the $(n+1)^{\text{th}}$ a row is $(1+1)^n = 2^n$, a power of two.

The second answer makes no sense because on the third row of the triangle, the sum is $1+2+1 = 4$, which is not a factorial: $1, 2, 6, 24, \ldots$ You would get the factorials in the triangle for the Stirling numbers of the first kind.

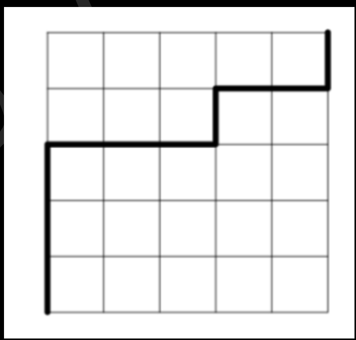
The third answer makes no sense because on the third row of the triangle, the sum is $1+2+1 = 4$, which is not a Bell number: $1, 2, 5, 15, \ldots$ You would get the Bell numbers in the triangle for the Stirling numbers of the second kind.

The fourth answer makes no sense because one of the answers above is correct.

Correct Answer: A

22. Let be given a square grid with $n \times n$ squares, where $n \geq 1$. Now consider all paths from the lower-left corner to the upper-right corner, in which each step is either up or to the right. Such a path always consists of $2n$ steps, of which n are horizontal and n are vertical.

Here is an example of such a path in the grid for $n = 5$:



Now how many paths like this are there?

A. $\binom{2n}{n}$

B. 2^n

C. 2^{2n}

D. None of the above.

Solution:

As every path has $2n$ steps, where n of them are horizontal and n of them are vertical, a path is determined by indicating which of the $2n$ steps are horizontal and which are vertical. So we can create such a path by putting the numbers $1, 2, \ldots, 2n$ in a bag, and picking n numbers out of this bag and let these numbers represent the vertical steps. So the path in the example is represented by the numbers 1−2−3−7−10. Note that the order of the numbers chosen does not matter. From the construction, it follows that the number of ways this can be done is $\binom{2n}{n}$.

Correct Answer: A

23. An interschool basketball tournament is being held at the Olympic sports complex. There are multiple basketball courts. Matches are scheduled in parallel, with staggered timings, to ensure that spectators always have some match or other available to watch. Each match requires a team of referees and linesmen. Two matches that overlap require disjoint teams of referees and linesmen. The tournament organizers would like to determine how many teams of referees and linesmen they need to mobilize to effectively conduct the tournament. To determine this, which graph theoretic problem do the organizers have to solve?

A. Find a minimal colouring.

B. Find a minimal spanning tree.

C. Find a minimal cut.

D. Find a minimal vertex cover.

Solution:

Matches are nodes. There is an edge between two matches if they overlap. Each team of referees and linesmen is a colour. A valid colouring assigns disjoint teams to overlapping matches. A minimal colouring determines the smallest number of such teams that are needed.

Correct Answer: A

24. The power set $\mathcal{P}((A \times B) \cup (B \times A))$ has the same number of elements as the power set $\mathcal{P}((A \times B) \cup (A \times B))$ if and only if

A. $A = \emptyset$ or $B = \emptyset$

B. $B = \emptyset$ or $A = B$

C. $A = \emptyset$ or $B = \emptyset$ or $A = B$

D. $A = \emptyset$ or $B = \emptyset$ or $A \cap B = \emptyset$

Solution:

Correct Answer: C

25. For sets A and B, let $f : A \to B$ and $g : B \to A$ be functions such that $f(g(x)) = x$ for each x.

Which of the following is/are not correct? (Mark all the appropriate choices)

A. The function f must be one-to-one.

B. The function f must be onto.

C. The function g must be one-to-one.

D. The function g must be onto.

Solution:

If $g(x_1) = g(x_2)$, then $x_1 = f(g(x_1)) = f(g(x_2)) = x_2$, so g is one-to-one. Also, f is onto because each $x \in B$ is in the image of f, namely $x = f(g(x))$. The other two statements are false, e.g. by constructing an example in which A is a larger finite set than B.

Correct Answer: A;D

26. A broken calculator has all its 10-digit keys and two operation keys intact. Let us call these operation keys A and B. When the calculator displays a number n pressing A changes the display to $n+1$. When the calculator displays a number n pressing B changes the display to $2n$. For example, if the number 3 is displayed then the key strokes ABBA changes the display in the following steps $3 \to 4 \to 8 \to 16 \to 17$. If 1 is on the display what is the least number of keystrokes needed to get 260 on the display?

A. 8

B. 9

C. 7

D. 18

Solution:

9, there are exactly two sequences, for example, $BBBBBBABB$.

Correct Answer: B

27. How many elements are in the following set?

$$\{(A,B) \mid A \subseteq B \subseteq \{1,2,3,\ldots,n\}\}$$

A. 2^{n-1}

B. 3^n

C. 2^{n+1}

D. 2^{2n}

Solution:

For each pair (A,B), we can define a function $f_{A,B} : \{1,\ldots,n\} \to \{a,b,c\}$ as follows:

$$f_{A,B}(i) = \begin{cases} a & \text{if } i \in A \backslash B \\ b & \text{if } i \in B \backslash A \\ c & \text{if } i \in A \cap B \quad (i \text{ is common to } A \text{ and } B) \end{cases}$$

For each function $f : \{1,\ldots,n\} \to \{a,b,c\}$, we can define a pair $\left(A_f, B_f\right)$ as follows:

$$A_f = \{i \mid f(i) = a \text{ or } f(i) = c\} \quad B_f = \{i \mid f(i) = b \text{ or } f(i) = c\}$$

The total number of functions from $\{1,\ldots,n\}$ to $\{a,b,c\}$ is 3^n.

Correct Answer: B

28. A graph is k-regular if all the vertices have degree exactly k. What is the minimum number of vertices in a k-regular graph that has no 3-length cycles?

A. k

B. $k+1$

C. $2k$

D. $2k+1$

Solution:

Consider any k-regular graph with $< 2k$ vertices, and two vertices u and v with an edge between them. Other than v, u has $k-1$ neighbours. Other than u, v has $k-1$ neighbours. But the number of vertices other than u and v is strictly less than $2k-2 = 2(k-1)$. Thus there is a vertex w which is a common neighbour to both u and v, and $u-v-w$ is a cycle of length 3.

The complete bipartite graph on $2k$ vertices, $K_{k,k}$, is k-regular and has no odd cycles. So $2k$ is the minimum number of vertices needed for a k-regular graph to not have cycles of length 3.

Correct Answer: C

29. A candy factory uses 5 fruit flavours $\{A,B,C,D,E\}$. Each candy is made using one or more of these flavours. The taste of a candy depends on which flavours are included.

A kid may prefer some combination of flavours more than others. For example, the kid may prefer the combination $\{A,B,C\}$ over $\{B,C\}$. The preference order of a kid is a total ordering of all the combinations, the ones occurring earlier being preferred more.

Suppose you want to throw a party and do not want more than one kid with the same preference order. What is the maximum number of kids that can attend such a party?

A. $\left(2^{(5!)}\right)-1$

B. $\left(\left(2^5\right)-1\right)!$

C. $((5*4)/2)-1$

D. $5*5$

Solution:

Each nonempty subset of the set of 5 fruit flavours will result in a distinct taste. So there are $F = \left(2^5\right)-1$ tastes. These can be arranged in $F!$ ways, which is the number of distinct preference orders. Inviting more than these many kids will result in at least two kids having the same preference order, by the pigeonhole principle. Hence (B) is the answer.

Correct Answer: B

30. Consider the following statements about finite simple graphs G :

(i) If each vertex of a graph G has degree at least 2 then G contains a cycle as a subgraph.

(ii) If the number of edges of a graph G is at least as large as the number of its vertices, then G contains a cycle as a subgraph.

Which of the above two statements holds for all graphs?

A. (i) only

B. (ii) only

C. both (i) and (ii)

D. neither of them

Solution:

(i) Let $P = \langle v_1, v_2, \ldots, v_t \rangle$ be a maximal path in G. Since v_t has degree at least 2 in G, there is an edge $\{v_t, x\}$ in G which is different from the edge incident on v_t in the path P. Since P is a maximal path the vertex x must belong to P. So $x = v_i$ for some $1 \le i < (t-1)$. The path $\left\langle v_i, v_{(i+1)}, \ldots, v_t \right\rangle$ together with the edge $\{v_t, v_i\}$ forms a cycle in G.

(ii) Suppose not, and let H be a graph with the smallest number of vertices which contradicts the statement. That is, H has at least as many edges as vertices, and does not contain a cycle. From the solution to part (a) above we get that there must be a vertex v in H whose degree is at most one. Deleting v from H results in a graph H' with (i) one fewer vertex, and (ii) at most one fewer edge, so H' also satisfies the premise of the statement. But then H' does not have any cycle, because H did not have any by assumption. Thus H' is a graph with fewer vertices than H for which the statement holds, which contradicts our assumption about H. Hence proved.

Correct Answer: C

2 Engineering Mathematics (30 Questions)

31. Given a real number x, define $g(x) = x^2 e^x$ if $x \geq 0$ and $g(x) = xe^{-x}$ if $x < 0$. Which of the following is/are correct? (Multiple Select Answers)

A. The function g is continuous everywhere.

B. The function g is differentiable everywhere.

C. The function g is one-to-one.

D. The range of g is the set of all real numbers.

Solution:

Correct Answer: A; C;D

32. Consider the polynomial $p(x) = (x + a_1)(x + a_2)\cdots(x + a_{10})$ where a_i is a real number for each $i = 1, \ldots, 10$. Suppose all of the eleven coefficients of $p(x)$ are positive. Which of the following is/are correct? (Multiple Select Answers)

A. The polynomial $p(x)$ must have a global minimum.

B. Each a_i must be positive.

C. All real roots of $p'(x)$ must be negative.

D. All roots of $p'(x)$ must be real.

Solution:

All are true. (A) The degree is even, so $p(x)$ goes to $+\infty$ as $x \to \pm\infty$. So $p(x)$ must attain a global minimum somewhere by continuity. (B) The roots of $p(x)$ are $-a_i$. By positivity of coefficients of $p(x)$, no nonnegative number is a root of $p(x)$. Thus all $-a_i$ are negative, so all $a_i > 0$. (C+ D) All 10 roots of $p(x)$ are real and negative. There is a root of $p'(x)$ between consecutive roots of $p(x)$ (this is valid even in case of multiple roots). So all 9 roots of $p'(x)$ are real and negative as well. For negativity, one can also note that all coefficients of $p'(x)$ are positive and apply the logic in (B) to $p'(x)$.

Correct Answer: A;B;C;D

33. Consider polynomials of the form $f(x) = x^3 + ax^2 + bx + c$ where a, b, c are integers. Name the three (possibly non-real) roots of $f(x)$ to be p, q, r. Which of the following is/are incorrect? (Multiple Select Answers)

A. If $f(1) = 2021$, then $f(x) = (x-1)\left(x^2 + sx + t\right) + 2021$ where s, t must be integers.

B. There is such a polynomial $f(x)$ with $c = 2021$ and $p = 2$.

C. There is such a polynomial $f(x)$ with $r = \frac{1}{2}$.

D. The value of $p^2 + q^2 + r^2$ does not depend on the value of c.

Solution:

(A) is true by the remainder theorem. Long division automatically gives integers s, t. Uniqueness of quotient and remainder for polynomial long division means those are the only values of s, t that work. (B) is false by substituting $x = 2$ into $f(x)$ and noting that $c = 2021$ forces $f(2)$ to be odd, in particular nonzero. To see that (C) is false, substitute $x = \frac{1}{2}$ into $f(x)$, multiply by 8 to clear denominators, and see that the leading term makes the integer $8f\left(\frac{1}{2}\right)$ odd. So $f\left(\frac{1}{2}\right)$ is nonzero. (General version of (B) and (C) that one gets by the similar reasoning: suppose a polynomial $p(x)$ with integer coefficients has a rational root $\frac{r}{s}$ written in lowest form. Then the leading coefficient of $p(x)$ is divisible by s and the constant term of $p(x)$ is divisible by r. Often used special case: for a polynomial $p(x) = x^n +$ lower terms with integer coefficients, any rational root must be an integer.) For (D) note that $p^2 + q^2 + r^2 = (p + q + r)^2 - 2(pq + pr + qr) = (-a)^2 - 2b$ does not depend on c.

Correct Answer: B; C

34. A stationary point of a function f is a real number r such that $f'(r) = 0$. A polynomial need not have a stationary point (e.g. x^3+x has none). Consider a polynomial $p(x)$. Which of the following is/are correct? (Multiple Select Answers)

A. If $p(x)$ is of degree 2022, then $p(x)$ must have at least one stationary point.

B. If the number of distinct real roots of $p(x)$ is 2021, then $p(x)$ must have at least 2020 stationary points.

C. If the number of distinct real roots of $p(x)$ is 2021, then $p(x)$ can have at most 2020 stationary points.

D. If r is a stationary point of $p(x)$ AND $p''(r) = 0$, then the point $(r, p(r))$ is neither a local maximum nor a local minimum point on the graph of $p(x)$.

Solution:

(A) $p'(x)$ is a polynomial of degree 2021, which is odd, so it has a root by intermediate value theorem by looking at behaviour as $x \to \pm\infty$. (B) The graph of $p(x)$ has to turn between any two consecutive zeros, giving a stationary point, in fact, a local max/min (C) The graph of $p(x)$ can turn more than once between zeros, or turn outside extreme zeros or have stationary points that are not maxima or minima. (D) is false, e.g., $p(x) = x^4$.

Correct Answer: A; B

35. Given three distinct positive constants a, b, c we want to solve the simultaneous equations

$$ax + by = \sqrt{2}$$
$$bx + cy = \sqrt{3}$$

Which of the following is/are correct? (Multiple Select Answers)

A. There exists a combination of values for a, b, c such that the above system has infinitely many solutions (x, y).

B. There exists a combination of values for a, b, c such that the above system has exactly one solution (x, y).

C. Suppose that for a combination of values for a, b, c, the above system has NO solution. Then $2b < a + c$.

D. Suppose $2b < a + c$. Then the above system has NO solution.

Solution:

Each of the given equations defines a line in the XY plane. (A) One can arrange both lines to be identical by having each equation a scalar multiple of the other, e.g., $a = 1, b = \frac{\sqrt{3}}{\sqrt{2}}, c = \frac{3}{2}$. (B) There is a unique solution when the two lines are distinct and not parallel. (C) The two lines are given to be parallel. So slopes are equal, i.e., $b^2 = ac$. Thus b is the geometric mean of a and c, so $b <$ the arithmetic mean $\frac{a+c}{2}$. (Recall that a, b, c are distinct and positive.) (D) is absurd. Just ensure $b^2 \neq ac$.

Correct Answer: A; B; C

36.

$$f(x) = \frac{x}{x + \sin x} \quad \text{and} \quad g(x) = \frac{x^4 + x^6}{e^x - 1 - x^2}.$$

Which of the following is/are incorrect? (Multiple Select Answers)

A. Limit as $x \to 0$ of $f(x)$ is $\frac{1}{2}$.

B. Limit as $x \to \infty$ of $f(x)$ does not exist.

C. Limit as $x \to \infty$ of $g(x)$ is finite.

D. Limit as $x \to 0$ of $g(x)$ is 720.

Solution:

Calculate (A) and (C) using L'Hôpital's rule. (Or in (A) use that $\sin x$ behaves like x near 0 and in (C) the limit is 0 because e^x dominates any polynomial for large x.) In (B) the limit is 1 as $f(x)$ is sandwiched between $\frac{x}{x\pm1}$, both of which $\to 1$. L'Hôpital's rule is not applicable as the expression one gets after attempting it does not have a limit as $x \to \infty$, so L'Hôpital's rule does not tell us anything. In (D) the limit is 0 by L'Hôpital's rule used correctly. Only one application is enough.

Correct Answer: B;D

37. Define the right derivative of a function f at $x = a$ to be the following limit if it exists.

$$\lim_{h\to 0^+} \frac{f(a+h)-f(a)}{h}, \text{ where } h \to 0^+ \text{means}$$

h approaches 0 only through positive values.

Which of the following is/are correct? (Multiple Select Answers)

A. If f is differentiable at $x = a$ then f has a right derivative at $x = a$.

B. $f(x) = |x|$ has a right derivative at $x = 0$.

C. If f has a right derivative at $x = a$ then f is continuous at $x = a$.

D. If f is continuous at $x = a$ then f has a right derivative at $x = a$.

Solution:

A. True. Obvious from the definition of the derivative.

B. True. The right derivative is 1.

C. False. Consider the floor function at integer values.

D. False. Take $x \sin \frac{1}{x}$ made continuous at 0 .

Correct Answer: A; B

38. Consider the real matrix

$$A = \begin{pmatrix} \lambda & 2 \\ 3 & 5 \end{pmatrix}.$$

Assume that -1 is an eigenvalue of A. Which of the following is/are true? (Multiple Select Answers)

A. The other eigenvalue is in $\mathbb{C}\backslash\mathbb{R}$.

B. $A + I_2$ is singular.

C. $\lambda = 1$.

D. Trace of A is 5.

Solution:

Correct Answer: B; D

39. Let $A \in M_2(\mathbb{R})$ be a nonzero matrix. Which of the following is/are true? (Multiple Select Answers)

A. If $A^2 = 0$, then $(I_2 - A)^5 = 0$.

B. If $A^2 = 0$, then $(I_2 - A)$ is invertible.

C. If $A^3 = 0$, then $A^2 = 0$.

D. If $A^2 = A^3 \neq 0$, then A is invertible.

Solution:

Correct Answer: B; C

40. If $A = \begin{vmatrix} a & b \\ b & a \end{vmatrix}$ and $A^2 = \begin{vmatrix} \alpha & \beta \\ \beta & \alpha \end{vmatrix}$, then which of the following is/are true? (Multiple Select Answers)

A. $\alpha = a^2 + b^2$

B. $\beta = 2ab$

C. $\alpha = a^2 - b^2$

D. $\beta = -2ab$

Solution:

Correct Answer: A; B

41. If P is an invertible matrix and $A = PBP^{-1}$, then which of the following statements is/are necessarily true? (Multiple Select Answers)

A. $B = P^{-1}AP$

B. $|A| = |B|$

C. A is invertible if and only if B is invertible

C. $B^T = A^T$.

Solution:

Correct Answer: A; B; C

42. A $n \times n$ matrix A is said to be symmetric if $A^T = A$. Suppose A is an arbitrary 2×2 matrix. Then which of the following matrices is/are symmetric (here $\mathbf{0}$ denotes the 2×2 matrix consisting of zeroes): (Multiple Select Answers)

A. $A^T A$

B. $\begin{bmatrix} \mathbf{0} & A^T \\ A & \mathbf{0} \end{bmatrix}$

C. AA^T

D. $\begin{bmatrix} A & \mathbf{0} \\ \mathbf{0} & A^T \end{bmatrix}$

Solution:

Correct Answer: A; B; C

43. The sum of the diagonal elements of a matrix A is called the trace of A and is denoted by $\mathrm{tr}(A)$. Which of the following statements about the trace are true? (Multiple Select Answers)

A. $\mathrm{tr}(A+B) = \mathrm{tr}(A) + \mathrm{tr}(B)$;

B. $\mathrm{tr}(2A) = 2\,\mathrm{tr}(A)$;

C. $\mathrm{tr}\left(A^T\right) = \mathrm{tr}(A)$;

D. $\mathrm{tr}\left(A^{-1}\right) = \mathrm{tr}(A)$.

Solution:

Correct Answer: A; B; C

44. An upper triangular matrix is a square matrix with all entries below the diagonal being zero. Suppose A and B are upper triangular matrices. Which of the following statements are true? (Multiple Select Answers)

A. The matrix $A + B$ is upper triangular.

B. The matrix A^T is upper triangular.

C. The matrix A^{-1} is upper triangular.

D. The matrix AB is upper triangular.

Solution:

Correct Answer: A; C; D

45. Consider the matrices

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 15 \\ 0 & 0 & 0 & 13 & 14 \\ 0 & 0 & 10 & 11 & 12 \\ 0 & 6 & 7 & 8 & 9 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 5 & 4 & 3 & 2 & 1 \\ 9 & 8 & 7 & 6 & 0 \\ 12 & 11 & 10 & 0 & 0 \\ 14 & 13 & 0 & 0 & 0 \\ 15 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Which of the following hold true? (Multiple Select Answers)

A. $|A| = |B|$

B. $\text{trace}(A) = \text{trace}(B)$

C. $|A| = -|B|$

D. $\text{trace}(AB) = \text{trace}(BA)$

Solution:

Correct Answer: A; B; D

46. Suppose that A is an $n \times n$ matrix with $n = 10$ and b is an $n \times 1$ vector. Suppose that the equation $Ax = b$ for an $n \times 1$ vector does not admit any solution. Which of the following conclusions can be drawn from the given information? (Multiple Select Answers)

A. A^{-1} does not exist.

B. The equation $A^T x = b$ also does not admit any solution.

C. $|A| = 0$.

D. Suppose c is another $n \times 1$ vector such that $Ax = c$ also does not admit a solution. Then the vector c is a constant multiple of the vector b.

Solution:

Correct Answer: A; C

47. Let $A = \left(\left(a_{ij}\right)\right)$ be a 7×7 matrix with $a_{i,i+1} = 1$ for $1 \leq i \leq 6, a_{7,1} = 1$ and all the other elements of the matrix are zero. Which of the following statements are true? (Multiple Select Answers)

A. $|A| = 1$

B. $\text{trace}(A) = 0$

C. $A^{-1} = A$

D. $A^7 = I$, where I is the identity matrix

Solution:

Correct Answer: A; B; D

48. Let A and B be events such that $P(A) = 0.4, P(B) = 0.5$ and $P(A \cup B) = 0.7$. Which of the following are true? (For sets $A, B, A\Delta B = (A^c \cap B) \cup (A \cap B^c)$). (Multiple Select Answers)

A. A and B are mutually exclusive

B. A and B are independent

C. $P(A\Delta B) = 0.1$

D. $P\left(A^c \cup B^c\right) = 0.8$

Solution:

Correct Answer: B; D

49. Which of the following statement(s) is/are true for an arbitrary $n \times n$ matrix A?

A. Exchanging two rows of A does not change its determinant.

B. Exchanging two rows of A does not change its trace.

C. Replacing each diagonal element of A with a 1 does not change its determinant.

D. Exchanging two columns of A negates its determinant.

Solution:

Correct Answer: D

50. Let A, B be $n \times n$ matrices. Which of the following properties of A and B is/are preserved under matrix multiplication? (Multiple Select Answers)

A. Being upper triangular.

B. All diagonal elements being zero.

C. Being diagonal.

D. Being symmetric.

Solution:

Correct Answer: A; C

51. A continuous random variable whose probability density function is given, for some $\lambda > 0$, by

$$\begin{aligned} f(x) &= \lambda e^{-\lambda x}, \quad && \text{if } x \geq 0 \\ &= 0, && \text{if } x < 0 \end{aligned}$$

is said to be an exponential random variable (or, exponentially distributed) with parameter λ.

Suppose that the number of miles that a car can run before its battery wears out is exponentially distributed with an average value of 10,000 miles. Suppose the car has already run for 3000 miles, and a person now desires to take an additional 5000-mile trip, what is the probability that he or she will be able to complete the trip without having to replace the car battery?

A. 0.308

B. 0.604

C. 0.908

D. 0.412

Solution:

The correct answer is option (B). Though one can calculate the answer in the usual way one computes conditional probability, the calculation is greatly simplified if one remembers that exponentially distributed random variables are memoryless. Refer the following screenshot taken from the book of "A First Course in Probability (5th edition)" by Sheldon M. Ross (page no. - 217) -

We say that a nonnegative random variable X is memoryless if

$$P\{X > s + t \mid X > t\} = P\{X > s\} \quad \text{for all } s, t \geq 0$$

If we think of X as being the lifetime of some instrument, Equation (5.1) states that the probability that the instrument survives for at least $s + t$ hours, given that it has survived t hours, is the same as the initial probability that it survives for at least s hours. In other words, if the instrument is alive at age t, the distribution of the remaining amount of time that it survives is the same as the original lifetime distribution (that is, it is as if the instrument does not remember that it has already been in use for a time t). The condition (5.1) is equivalent to

$$\frac{P\{X > s + t, X > t\}}{P\{X > t\}} = P\{X > s\}$$

or

$$P\{X > s + t\} = P\{X > s\} P\{X > t\}$$

Since Equation (5.2) is satisfied when X is exponentially distributed (for $e^{-\lambda(s+t)} = e^{-\lambda s} e^{-\lambda t}$), it follows that exponentially distributed random variables are memoryless.

Therefore, the answer to our original question is : $P\{\text{ remaining lifetime} > 5\} = 1 - F(5) = e^{-5\lambda} = e^{-1/2} \approx 0.604$, where $F(x)$ is the cumulative distribution function (cdf) for the given random variable.

Correct Answer: B

52. For a non-zero real number a, the inverse of $J = \begin{pmatrix} a & 1 & 0 \\ 0 & a & 1 \\ 0 & 0 & a \end{pmatrix}$ is ______.

A. $\begin{pmatrix} a^{-1} & a^{-2} & a^{-3} \\ 0 & a^{-1} & a^{-2} \\ 0 & 0 & a^{-1} \end{pmatrix}$

B. $\begin{pmatrix} a^{-1} & -a^{-2} & a^{-3} \\ 0 & a^{-1} & -a^{-2} \\ 0 & 0 & a^{-1} \end{pmatrix}$

C. $\begin{pmatrix} a^{-1} & 1 & 0 \\ 0 & a^{-1} & 1 \\ 0 & 0 & a^{-1} \end{pmatrix}$

D. $\begin{pmatrix} a^{-1} & a^{-2} & 0 \\ 0 & a^{-1} & a^{-2} \\ 0 & 0 & a^{-1} \end{pmatrix}$

Solution:

Correct Answer: B

53. Let X be a Binomial(n,p) random variable with the probability mass function

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, 2, \ldots, n.$$

Let T be a random variable defined as:

$$T = \begin{cases} 1 & \text{if } X = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Let $\mathbb{E}(T)$ and $\mathbb{V}(T)$ denote the expectation and the variance of a random variable T. Which of the following statements is/are true? (Multiple Select Answers)

A. $\mathbb{E}\left(T^2\right) = np(1-p)^{n-1}$ and $\mathbb{V}(T) = np(1-p)^{n-1}\left\{\left(1 - \left(np(1-p)^{n-1}\right)\right)\right\}$

B. $\mathbb{E}(T) = np$ and $\mathbb{V}(T) = np(1-p)$

C. $\mathbb{E}(T) = np(1-p)^{n-1}$ and $\mathbb{E}\left(T^2\right) = np(1-p)^{n-1}$

D. $\mathbb{E}(T) = p$ and $\mathbb{V}(T) = p(1-p)$

Solution:

Correct Answer: A; C

54. Which of the following statements is/are true? (Multiple Select Answers)

A. Let A be a $n \times m$ matrix with $n < m$. Then there is a nonzero solution y with $Ay = 0$ only if A has full row rank.

B. Let A be a $n \times m$ matrix with $n > m$. There is a nonzero solution y with $Ay = 0$.

C. The row rank of an $n \times m$ matrix is equal to its column rank only when $n = m$.

D. Let A be an $n \times n$ matrix. Suppose $A = BC$, where B has size $n \times r$ and C has size $r \times n$. The rank of A is less than or equal to r.

Solution:

Correct Answer: D

55. Let A be a 20×11 matrix with real entries. After performing some row operations on A, we get a matrix B which has 12 nonzero rows. Which of the following is/ are not true? (Multiple Select Answers)

A. Rank of A is 12 .

B. Rank of A and B are not related.

C. If v is a vector such that $Av = 0$ then Bv is also 0 .

D. Rank of B is at most 11 .

Solution:

A. False. E.g.: Suppose A had 12 identical nonzero rows (and the rest all zeroes) to start with, and we swapped two rows to get B, which also has exactly 12 nonzero rows. The rank of A is 1 .

B. False. Follows from the definition; "row or column operations do not change the rank".

C. True. Same reason as above; the null space is unchanged.

D. True. B has 11 columns so its rank cannot be more.

Correct Answer: C;D

56. For each month in the year (i.e., January, February, March,...), let us assume the probability that a person's birthday falls in that particular month is exactly 1/12, and let us assume that this is independent for different persons. What is the smallest value of the natural number n such that, among n independently chosen persons the probability that there is a pair of them born in the same month is at least 1/2 ?

A. 3

B. 4

C. 5

D. 6

Solution:

Correct Answer: C

57. Winnie-the-Pooh and Oski are competing for honey! When Oski finds honey in Berkeley Hills, there is less honey left for Winnie. The probability that Winnie finds honey given that Oski has found honey is 0.2 . The probability that Winnie finds honey given that Oski has NOT found honey is 0.7. Oski can find honey with probability 0.4. Paddington Bear sees that Winnie has successfully found honey today and wants to calculate the probability that Oski also found honey today. What is the probability that Oski found honey today given that Winnie also found honey today?

A. 0.016

B. 0.18

C. 0.16

D. 0.018

Solution:

Let O be the event Oski finds honey and let W be the event Winnie finds honey.

$$\begin{aligned} P(O \mid W) &= \frac{P(W \mid O)P(O)}{P(W)} \\ &= \frac{P(W \mid O)P(O)}{P(W \mid O)P(O) + P(W \mid O^c)P(O^c)} \\ &= \frac{0.2 \times 0.4}{0.2 \times 0.4 + 0.7 \times 0.6} \\ &= 0.16 \end{aligned}$$

Correct Answer: C

58. Oski and Joe Bruins are playing a game. Joe Bruins draws a random variable X_J distributed as Uniform $[0,1]$. Oski, trying to outsmart Joe, first looks at Joe's random variable X_J and then draws a random variable X_O distributed Uniformly in $\left[X_J - 0.5, X_J + 0.5\right]$. The bear with the higher score wins. Kevin, the bear from Zootopia, looks at the point obtained by the winner, given by $X_W = \max\left\{X_O, X_J\right\}$ and the points obtained by the loser $X_L = \min\left\{X_O, X_J\right\}$. The value of $\mathbb{E}[X_W]$ and $\mathbb{E}[X_L]$ is ______.

A. $\mathbb{E}[X_W] = \frac{2}{3}$ and $\mathbb{E}[X_L] = \frac{1}{3}$

B. $\mathbb{E}[X_W] = \frac{3}{5}$ and $\mathbb{E}[X_L] = \frac{2}{5}$

C. $\mathbb{E}[X_W] = \frac{5}{9}$ and $\mathbb{E}[X_L] = \frac{4}{9}$

D. $\mathbb{E}[X_W] = \frac{5}{8}$ and $\mathbb{E}[X_L] = \frac{3}{8}$

Solution:

Note that $X_O = X_J + U$ where U is Uniform $[-0.5, 0.5]$ and independent of X_J.

$$\begin{aligned}
\mathbb{E}[X_W] &= \mathbb{E}\left[X_W \mid X_J < X_O\right] P\left(X_J < X_O\right) + \mathbb{E}\left[X_W \mid X_J \geq X_O\right] P\left(X_J \geq X_O\right) \\
&= \mathbb{E}\left[X_J + U \mid U > 0\right] P(U > 0) + \mathbb{E}\left[X_J \mid U \leq 0\right] P(U \leq 0) \\
&= \left(\mathbb{E}\left[X_J \mid U > 0\right] + \mathbb{E}[U \mid U > 0]\right) P(U > 0) + \mathbb{E}\left[X_J \mid U \leq 0\right] P(U \leq 0) \\
&= \left(\mathbb{E}\left[X_J\right] + \mathbb{E}[U \mid U > 0]\right) P(U > 0) + \mathbb{E}\left[X_J\right] P(U \leq 0) \\
&= \left(\frac{1}{2} + \frac{1}{4}\right)\frac{1}{2} + \frac{1}{2} \times \frac{1}{2} \\
&= \frac{5}{8}.
\end{aligned}$$

Now, $\mathbb{E}[X_W] + \mathbb{E}[X_L] = \mathbb{E}\left[X_O + X_J\right] = 2\mathbb{E}\left[X_J\right] = 1$. Therefore, $\mathbb{E}[X_L] = \frac{3}{8}$.

Correct Answer: D

59. Three events A, B, C are mutually independent if and only if

A. $P[A \mid B \cap C] = P(A)$ and $P(B \cap C) = P(B)P(C)$;

B. $P(A \cap B \cap C) = P(A)P(B)P(C)$

C. $P[A \cap B \mid C] = P(C)$ and $P[C \mid A \cap B] = P(A \cap B)$;

D. None of the above.

Solution:

Correct Answer: D

60. A multiple choice exam has 4 choices for each question. A student has studied enough so that the probability they will know the answer to a question is 0.5, the probability that they will be able to eliminate one choice is 0.25, otherwise all 4 choices seem equally plausible. If they know the answer they will get the question right. If not they have to guess from the 3 or 4 choices.

As the teacher you want the test to measure what the student knows. If the student answers a question correctly what's the probability they knew the answer?

A. 0.774

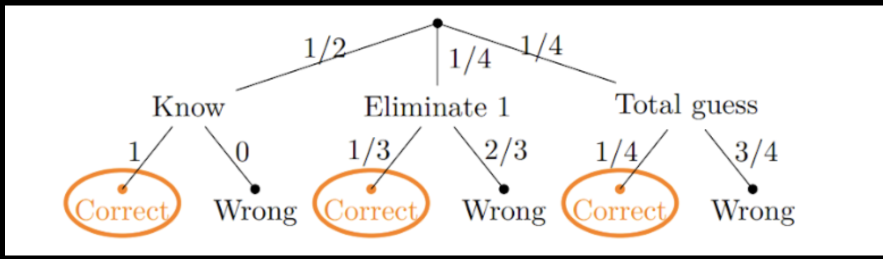
B. 0.884

C. 0.664

D. 0.777

Solution:

We show the probabilities in a tree:



For a given problem let C be the event the student gets the problem correct and K the event the student knows the answer. The question asks for $P(K \mid C)$. We'll compute this using Bayes' rule:

$$P(K \mid C) = \frac{P(C \mid K)P(K)}{P(C)} = \frac{1 \cdot 1/2}{1/2 + 1/12 + 1/16} = \frac{24}{31} \approx 0.774 = 77.4\%$$

Correct Answer: A

3 C Programming and Data Structures (20 Questions)

61. Based on the following program.

```
procedure mystery (A : array [1..100] of int)
    int i,j,position,tmp;
    begin
        for j := 1 to 100 do
            position := j;
            for i := j to 100 do
                if (A[i] > A[position]) then
                    position := i;
                endfor
            tmp := A[j];
            A[j] := A[position];
            A[position] := tmp;
        endfor
    end
```

When the procedure terminates, the array A has been:

A. Reversed

B. Left unaltered

C. Sorted in descending order

D. Sorted in ascending order

Solution:

At iteration j of the outer loop, the inner loop determines the index position between j and 100 such that A[position] is maximum, and A[position] is interchanged with $A[j]$. Thus at the end, position $A[j]$ holds the j^{th} largest value in the array.

Correct Answer: C

62. Based on the following program.

```
procedure mystery (A : array [1..100] of int)
    int i,j,position,tmp;
    begin
        for j := 1 to 100 do
            position := j;
            for i := j to 100 do
                if (A[i] > A[position]) then
                    position := i;
                endfor
            tmp := A[j]
            A[j] := A[position];
            A[position] := tmp;
        endfor
    end
```

The number of times the test A[i] > A[position] is executed is:

A. 100

B. 5050

C. 10000

D. Depends on contents of A

Solution:

In iteration j, there are $100 - j + 1$ comparisons made. So in all, there are

$$100 + 99 + \cdots + 2 + 1 = 5050$$

comparisons.

Correct Answer: B

63. Consider the code below, defining the function mystery:

```
mystery(a,b){
    if (a<0 or b < 0) return 0;
    else if (a == 0) return b+1;
    else if (b == 0) return mystery(a-1,1);
    else return mystery(a-1, mystery(a,b-1));
}
```

If mystery$(3,2)$ is α and mystery$(3,3)$ is β. Then the value of $\alpha - \beta = ?$ (Numerical Answer Type)

Solution:

mystery$(3,0)$ = mystery$(2,1) = 2 + 3 = 5$.

mystery $(3,1)$ = mystery $(2,$ mystery $(3,0))$ = mystery $(2,5) = 2 \cdot 5 + 3 = 13$.

mystery $(3,2)$ = mystery $(2,$ myster $y(3,1))$ = mystery $(2,13) = 2 \cdot 13 + 3 = 29$.

mystery $(3,3)$ = mystery $(2,$ mystery $(3,2))$ = mystery $(2,29) = 2 \cdot 29 + 3 = 61$.

$\therefore$ The value of $\alpha - \beta = 29 - 61 = -32$.

Correct Answer: -32

64. In the code fragment on the right, start and end are integer values and prime(x) is a function that returns true if x is a prime number and false otherwise.

```
i := 0 ; j := 0 ; k -:= 0;
for (m := start; m <= end; m := m+1)
    k :=k + m;
    if (prime(m)){
        i := i + m;
    }else{
        j:= j + m;
    }
}
```

At the end of the loop:

A. $k < i + j$

B. $k = i + j$

C. $k > i + j$

D. Depends on the start and end

Solution:

In each iteration, the value added to k is also added to exactly one of i and j.

Correct Answer: B

65. Suppose we are working with a programming language that supports automatic garbage collection. This means that:

A. Uninitialized variables are assigned null values.

B. Unreferenced dynamically allocated memory is added back to free space.

C. Unreachable if-then-else branches are pruned.

D. Expressions where array indices exceed array bounds are flagged.

Solution:

Unreferenced dynamically allocated memory is added back to free.

Correct Answer: B

66. Consider the code below, defining the function A :

```
A(m, n, p) {
    if (p == 0) return m+n;
    else if ( n == 0 && p == 1) return 0;
    else if (n == 0 && p == 2) return 1;
    else if (n == 0) return m;
    else return A(m, A(m,n-1,p), p-1);
}
```

The value of $A(m, n, 1)$ as a function of m and n is ______. (Multiple Selects Answers)

A. mn

B. $mn - 1$

C. $n + m(n - 1)$

D. $m + m(n - 1)$

Solution:

Evaluating $A(m,n,1)$ for small values of n suggests that $A(m,n,1) = mn$. We verify it using induction. The base case is clear: $A(m,0,1) = 0 = m \cdot 0$. For the induction step, assuming that $A(m,n-1,1) = m(n-1)$, we have

$$A(m,n,1) = A(m,A(m,n-1,1),0) = A(m,m(n-1),0) = m + m(n-1) = mn.$$

Correct Answer: A;D

67. How many times is the comparison $i \geq n$ performed in the following program?

```
int i=85, n=5;
main() {
    while (i >= n) {
        i=i-1;
        n=n+1;
    }
}
```

A. 40

B. 41

C. 42

D. 43

Solution:

The value of $i - n$ is 80 initially. We run the loop as long as $i - n \geq 0$ and in each iteration, $i - n$ decreases by 2. Just before the k^{th} time the comparison is performed (for $k \geq 1$), the value of $i - n$ is $80 - 2k + 2$.

Hence just before the forty-first comparison, the value of $i - n$ is 0. After the forty-first comparison, the loop is executed one last time. We need to make the comparison once more to exit the loop.

Thus the correct answer is 42.

Correct Answer: C

68. Spaghetti code is a phrase used to refer to unstructured and difficult-to-maintain source code. Such code has a complex and tangled control structure, resulting in a program flow that is conceptually like a bowl of spaghetti, twisted and tangled. The overuse of which of the following programming constructs (despite other possible constructs) results in the creation of a spaghetti code?

A. Polymorphism

B. Nested for-loop constructs

C. Repeat until loop constructs

D. Goto statements

Solution:

The correct answer is option (D). Using Goto statements is generally considered to be a bad programming practice, as it encourages arbitrary flow control. As per Wikipedia, code that overuses GOTO statements rather than structured programming constructs, resulting in convoluted and unmaintainable programs, is often called spaghetti code. Such code has a complex and tangled control structure, resulting in a program flow that is conceptually like a bowl of spaghetti, twisted and tangled.

Correct Answer: C

69. Which of the following relationships holds in general between the scope of a variable and the lifetime of a variable (in a language like C or Java)?

A. The scope of a variable is contained in the lifetime of the variable.

B. The scope of a variable is the same as the lifetime of the variable.

C. The lifetime of a variable is disjoint from the scope of the variable.

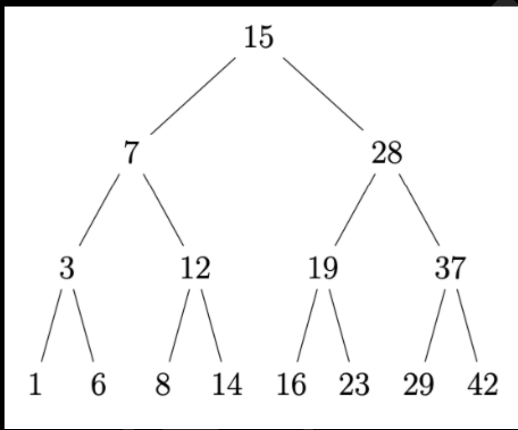
D. None of the above.

Solution:

The scope of a variable is contained in the lifetime of the variable.

Correct Answer: A

70. Suppose we constructed the binary search tree shown at the right by starting with an empty tree and inserting one element at a time from an input sequence, without any rotations or other manipulations. Which of the following assertions about the order of elements in the input sequence cannot be true?



A. 8 came after 3 and 19 came after 29.

B. 7 came before 8 and 23 came after 37.

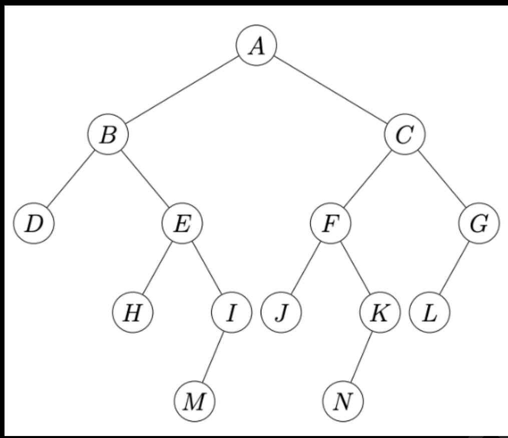
C. 1 came after 12 and 29 came before 42.

D. 3 came before 14 and 16 came before 28.

Solution:

28 is an ancestor of 16, so 16 must have come after 28. In the tree, only ancestor-descendant relationships matter in determining the order in which elements arrive. An ancestor must always come before any of its descendants. Incomparable elements could come in any order.

71. Suppose that the figure to the right is a binary search tree. The letters indicate the names of the nodes, not the values that are stored. What is the predecessor node, in terms of value, of the root node A?



A. D

B. H

C. I

D. M

Solution:

The predecessor of A is the rightmost node in the left subtree of A.

Correct Answer: C

72. When a user submits a query, a search engine does the following. For every webpage that has been visited by the search engine, it computes a score indicating how relevant that page is to the query. Finally, it reports the pages with the top k scores on the screen, for a number k specified by the user. A good data structure for accumulating the scores and ranking them is:

A. a queue

B. a heap

C. a stack

D. a binary search tree

Solution:

Let n be the number of pages visited by the search engine at the time a query is submitted. Assume that it takes constant time to compute the relevance score for each page w.r.t. a query. Then it takes $O(n)$ time to compute the relevance scores, a further $O(n)$ time to build a heap of n relevance scores, and $O(k \cdot \log n)$ time for k delete-max operations to return the top k scores.

Correct Answer: A

73. In the code fragment below, start and end are integer values and prime(x) is a function that returns True if **x** is a prime number and False otherwise.

```
i = 0;
j = 0;
k = 0;
for m = start to end {
    if prime(m) == True {
        i = i + 1;
        // Statement 1
    }else{
        j = j - 1 ;
        // Statement 2
    }
}
```

We wish to maintain the invariant $k == i - j$ after each iteration of the for loop. What should we insert at Statement 1 and Statement 2?

A. Statement 1: $k = k + 1$ and Statement $2 : k = k + 1$

B. Statement 1: $k = k + 1$ and Statement $2 : k = k - 1$

C. Statement $1 : k = k - 1$ and Statement $2 : k = k + 1$

D. Statement 1: $k = k - 1$ and Statement $2 : k = k - 1$

Solution:

$i + (-j)$ is the number of the times the loop is iterated, so increment k in both branches of the if.

Correct Answer: A

74. Consider the following program. Assume that x and y are integers.

```
f(x, y)
{
    if (y != 0)
        return (x * f(x, y-1));
    else
        return 1;
}
```

What is $f(6,3)$?

A. 243

B. 729

C. 125

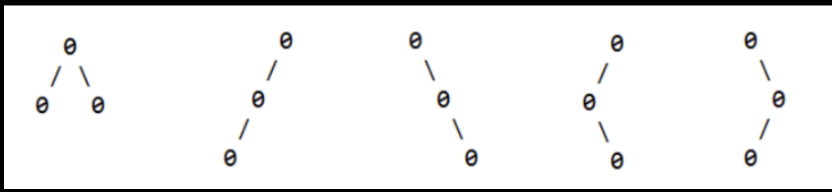
D. 216

Solution:

This code computes x^y.**f**$(6,3)$ will compute 6^3 which is 216.

Correct Answer: D

75. A binary tree starts with a single root node at the top of the tree. Each node can have either a left child or a right child, or both, or neither. The children of a node are drawn below it, connected by edges. Here are the five possible binary trees with three nodes.



Note that the directions left and right of the children matter. In the second tree, the root has a left child that has a left child, while, in the fourth tree, the root has a left child that has a right child, and so on. How many different binary trees can be constructed with four nodes?

A. 3

B. 5

C. 14

D. 30

Solution:

$T(n) = (2n)!/(n+1)!n!$ binary trees can be constructed with n unlabeled nodes. So, $T(4) = 8!/5!4! = 14$.

Correct Answer: C

76. Which of the following is/are not correct? (Multiple Select Answers)

A. Given two heaps with n elements each, it is possible to construct a single heap comprising all $2n$ elements in $O(n)$ time.

B. Building a heap with n elements can always be done in $O(n \log n)$ time.

C. We can always find the maximum in a min-heap in $O(\log n)$ time.

D. In a heap of depth d, there must be at least 2^d elements. (Assume the depth of the first element (or root) is zero).

Solution:

A. TRUE. Simply traverse each heap and read off all $2n$ elements into a new array. Then, make the array into a heap in $O(n)$ time by calling MAXHEAPIFY for $i = 2n$ down to 1.

B. TRUE. In fact, we can build a heap in $O(n)$ time by putting the elements in an array and then calling MAX-HEAPIFY for $i = n$ down to 1.

C. FALSE. The maximum element in a min-heap can be anywhere in the bottom level of the heap. There are up to $n/2$ elements in the bottom level, so finding the maximum can take up to $O(n)$ time.

D. TRUE. The minimum number of elements in a heap of depth d is one more than the maximum number of elements in a heap of depth $(d-1)$. Since a level at depth d in a binary heap can have up to 2^d elements, the number of elements at depth $(d-1)$ is $\sum_{i=0}^{d-1} 2^i = 2^d - 1$. So the minimum number of elements in a heap of depth d is $\left(2^d - 1\right) + 1 = 2^d$.

Correct Answer: C

77. Which of the following is/are not correct? (Multiple Select Answers)

A. Any two (possibly unbalanced) BSTs containing n elements each can be merged into a single balanced BST in $O(n)$ time.

B. Inserting an element into a binary search tree of size n always takes $O(\log n)$ time.

C. To delete the i^{th} node in a min heap, you can exchange the last node with the i^{th} node, then do the min-heapify on the i^{th} node, and then shrink the heap size to be one less the original size.

D. Any n-node unbalanced tree can be balanced using $O(\log n)$ rotations.

Solution:

A. TRUE. Use in-order traversal of the two BSTs to create two sorted lists of length n in $O(n)$ time, merge them into a single sorted list of length $2n$ in $O(n)$ time, and then create a balanced BST from the sorted lists in $O(n)$ time.

B. FALSE. Inserting an element into a binary search tree takes $O(h)$ time, where h is the height of the tree. If the tree is not balanced, h may be much larger than $\log n$(as large as $n-1$).

C. FALSE. The last node may be smaller than the i^{th} node's parent; min-heapify won't fix that.

D. FALSE. The worst-case unbalanced tree is a list, and balancing it requires $\Omega(n)$ rotations.

Correct Answer: B; C;D

78. Which of the following is/are correct? (Multiple Select Answers)

A. In a BST, we can find the next smallest element to a given element in $O(1)$ time.

B. In an AVL tree, during the insert operation there are at most two rotations needed.

C. In a min-heap, the next largest element of any element can be found in $O(\log n)$ time.

D. AVL trees can be used to implement an optimal comparison-based sorting algorithm.

Solution:

A. False. Finding the next smallest element, the predecessor, may require traveling down the height of the tree, making the running time $O(h)$.

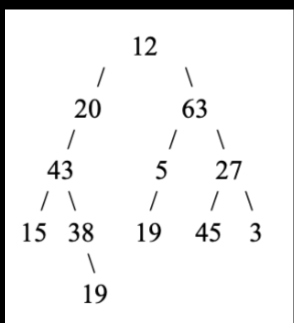
B. True. The AVL property is restored on every operation. Therefore, inserting another item will require at most two rotations to restore the balance.

C. False. A min-heap cannot provide the next largest element in $O(\log n)$ time. To find the next largest element, we need to do a linear, $O(n)$, search through the heap's array.

D. True. AVL trees can be used to sort N numbers in $O(N \log N)$ time, by inserting all the numbers in the tree, and iteratively calling NEXT-LARGEST N times.

Correct Answer: B;D

79. Which of the following is/are correct? (Multiple Select Answers)



A. Pre-Order: $12, 20, 43, 15, 38, 19, 63, \mathbf{5}, 19, 27, 45, 3$

B. In-Order: $15, 43, 38, 19, 20, 12, 19, 5, 63, 45, 27, 3$

C. Post-Order: $15, 19, 38, 43, 20, 19, \mathbf{5}, 45, 3, 27, 63, 12$

D. None of the above

Solution:

Correct Answer: A;B;C

80. You have n lists, each consisting of m integers sorted in ascending order. Merging these lists into a single sorted list will take time:

A. $O(nm \log m)$

B. $O(mn \log n)$

C. $O(m+n)$

D. $O(mn)$

Solution:

We can merge two sorted lists of size k and ℓ in time $O(k+\ell)$. We begin by merging the lists in pairs to generate $\frac{n}{2}$ lists of size $2m$ each, in total time $O(mn)$. If we repeat this, we get $\frac{n}{4}$ lists of size $4m$ each, again in total time $O(mn)$. Thus, in $O(\log n)$ rounds, we converge to a single sorted list. Each round takes time $O(mn)$, so the total time is $O(mn \log n)$.

Another strategy to achieve complexity $O(mn \log n)$ is to build a min-heap of size n, consisting the first element in each of the n lists. We then extract the minimum element from the heap to the output and replace it in the heap with the next element from the list from which the minimum came. Thus, it takes time $O(\log n)$ to generate each element in the output and there are $O(mn)$ output elements in all, so the overall time is $O(mn \log n)$.

On the other hand, if we can apply min() to n elements at a time, we can merge all n lists is parallel. Let the n lists be $\ell_1, \ell_2, \ldots, \ell_n$. We maintain indices $i_1, i_2, \ldots, i_n$ into these n lists, initially set to 1. At each stage, we add $j = \min(\ell_1[i_1], \ell_2[i_2], \ldots, \ell_n[i_n])$ to the sorted output list and increment the corresponding index. Overall, we add mn elements to the sorted output, and each element is generated in constant time, so the time taken is $O(mn)$.

Correct Answer: B

4 Algorithms (20 Questions)

81. Suppose that for some given array A, the quicksort algorithm initially invoked with input A, always partitions its input in the ratio $1:100$ in each of its recursive invocations. What is the tightest upper bound on the overall time complexity required to sort this input array A using the given quicksort algorithm?

A. $O(n \log n)$

B. $O(n)$

C. $O\left(n^2\right)$

D. None of the above

Solution:

The correct answer is option (A). Option (C) is clearly wrong as it violates the lower bound on the time required to sort any arbitrary array of n numbers. As for the remaining, as long as the input array is being partitioned in a constant proportion in each the recursive invocation of the quicksort algorithm, the overall time complexity required by the same remains $O(n \log n)$. Refer the following screenshot taken from the book of "Introduction to Algorithms (3rd edition)" by Cormen et al. (page no. 176) -

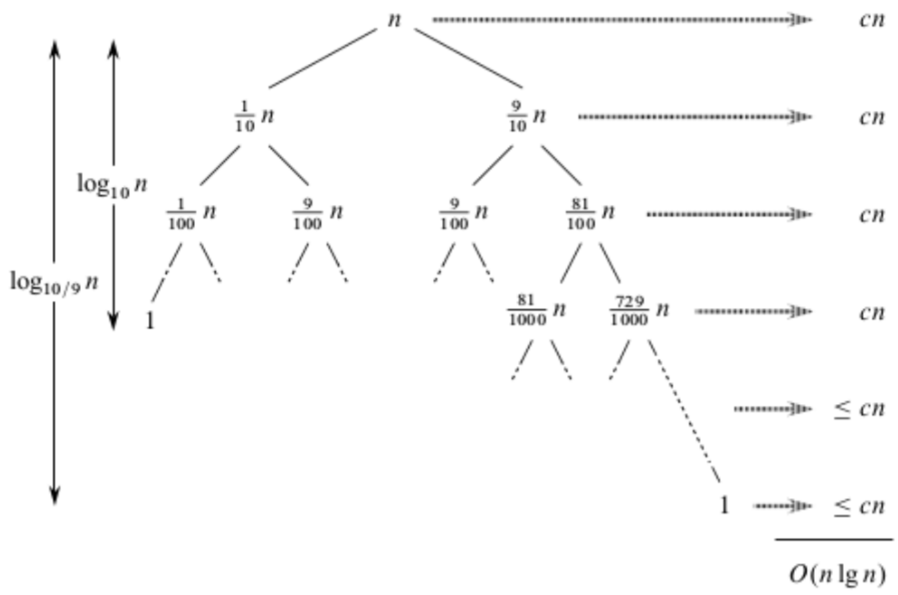


Figure 7.4 A recursion tree for QUICKSORT in which PARTITION always produces a 9-to-1 split, yielding a running time of $O(n \lg n)$. Nodes show subproblem sizes, with per-level costs on the right. The per-level costs include the constant c implicit in the $\Theta(n)$ term.

Suppose, for example, that the partitioning algorithm always produces a 9-to-1 proportional split, which at first blush seems quite unbalanced. We then obtain the recurrence

$$T(n) = T(9n/10) + T(n/10) + cn ,$$

on the running time of quicksort, where we have explicitly included the constant c hidden in the $\Theta(n)$ term. Figure 7.4 shows the recursion tree for this recurrence. Notice that every level of the tree has cost cn, until the recursion reaches a boundary condition at depth $\log_{10} n = \Theta(\lg n)$, and then the levels have cost at most cn. The recursion terminates at depth $\log_{10/9} n = \Theta(\lg n)$. The total cost of quicksort is therefore $O(n \lg n)$. Thus, with a 9-to-1 proportional split at every level of recursion, which intuitively seems quite unbalanced, quicksort runs in $O(n \lg n)$ time—asymptotically the same as if the split were right down the middle. Indeed, even a 99-to-1 split yields an $O(n \lg n)$ running time. In fact, any split of *constant* proportionality yields a recursion tree of depth $\Theta(\lg n)$, where the cost at each level is $O(n)$. The running time is therefore $O(n \lg n)$ whenever the split has constant proportionality.

Correct Answer: A

82. Given two functions f and g such that $f(n)$ is $O(g(n))$, which of the following statements involving f and g are true?

I. $\log_2 f(n)$ is $O(\log_2 g(n))$

II. $2^{f(n)}$ is $O\left(2^{g(n)}\right)$

III. $f(n)^2$ is $O\left(g(n)^2\right)$

A. I, II, and III

B. I, II, but not III

C. I, III, but not II

D. III, but neither I nor II

Solution:

The correct answer is option (D). This is an exercise problem from the book of "Algorithm Design (1st edition)" by Kleinberg and Tardos (page no. 68). The following explanation for the solution to the question has been taken from

https://github.com/mathiasuy/Soluciones-Klenberg/blob/master/chapter_02/problem%20(5).pdf

(i) This is false in general, since it could be that $g(n) = 1$ for all n, $f(n) = 2$ for all n, and then $\log_2 g(n) = 0$, whence we cannot write $\log_2 f(n) \le c \log_2 g(n)$.

On the other hand, if we simply require $g(n) \ge 2$ for all n beyond some n_1, then the statement holds. Since $f(n) \le cg(n)$ for all $n \ge n_0$, we have $\log_2 f(n) \le \log_2 g(n) + \log_2 c \le (\log_2 c)(\log_2 g(n))$ once $n \ge \max(n_0, n_1)$.

(ii) This is false: take $f(n) = 2n$ and $g(n) = n$. Then $2^{f(n)} = 4^n$, while $2^{g(n)} = 2^n$.

(iii) This is true. Since $f(n) \le cg(n)$ for all $n \ge n_0$, we have $(f(n))^2 \le c^2(g(n))^2$ for all $n \ge n_0$.

Correct Answer: D

83. Two of the two key ingredients that an optimization problem must have in order for dynamic programming to apply are

I. Optimal substructure - a problem is said to exhibit optimal substructure if an optimal solution to the problem contains within it optimal solutions to subproblems, and

II. Overlapping subproblems - when a recursive algorithm revisits the same problem repeatedly, we say that the optimization problem has overlapping subproblems.

Now, consider the following two problems -

I. Unweighted shortest simple path - Find a simple path from u to v consisting of the fewest edges.

II. Unweighted longest simple path - Find a simple path from u to v consisting of the most edges.

Choose the correct option from the following options.

A. The unweighted shortest simple path problem exhibits the optimal substructure property.

B. The unweighted longest simple path problem exhibits the optimal substructure property.

C. The unweighted shortest simple path problem DOES NOT exhibit the optimal substructure property.

D. The unweighted longest simple path problem DOES NOT exhibit the optimal substructure property.

Solution:

The correct answer is option (A). See the following screenshot from the book of "Introduction to Algorithms (3rd edition)" by Cormen et al. (page no. - 382-383).

The unweighted shortest-path problem exhibits optimal substructure, as follows. Suppose that $u \neq v$, so that the problem is nontrivial. Then, any path p from u to v must contain an intermediate vertex, say w. (Note that w may be u or v.) Thus, we can decompose the path $u \overset{p}{\leadsto} v$ into subpaths $u \overset{p_1}{\leadsto} w \overset{p_2}{\leadsto} v$. Clearly, the number of edges in p equals the number of edges in p_1 plus the number of edges in p_2. We claim that if p is an optimal (i.e., shortest) path from u to v, then p_1 must be a shortest path from u to w. Why? We use a "cut-and-paste" argument: if there were another path, say p_1', from u to w with fewer edges than p_1, then we could cut out p_1 and paste in p_1' to produce a path $u \overset{p_1'}{\leadsto} w \overset{p_2}{\leadsto} v$ with fewer edges than p, thus contradicting p's optimality. Symmetrically, p_2 must be a shortest path from w to v. Thus, we can find a shortest path from u to v by considering all intermediate vertices w, finding a shortest path from u to w and a shortest path from w to v, and choosing an intermediate vertex w that yields the overall shortest path. In Section 25.2, we use a variant of this observation of optimal substructure to find a shortest path between every pair of vertices on a weighted, directed graph.

You might be tempted to assume that the problem of finding an unweighted longest simple path exhibits optimal substructure as well. After all, if we decompose a longest simple path $u \overset{p}{\leadsto} v$ into subpaths $u \overset{p_1}{\leadsto} w \overset{p_2}{\leadsto} v$, then mustn't p_1 be a longest simple path from u to w, and mustn't p_2 be a longest simple path from w to v? The answer is no! Figure 15.6 supplies an example. Consider the path $q \to r \to t$, which is a longest simple path from q to t. Is $q \to r$ a longest simple path from q to r? No, for the path $q \to s \to t \to r$ is a simple path that is longer. Is $r \to t$ a longest simple path from r to t? No again, for the path $r \to q \to s \to t$ is a simple path that is longer.

This example shows that for longest simple paths, not only does the problem lack optimal substructure, but we cannot necessarily assemble a "legal" solution to the problem from solutions to subproblems. If we combine the longest simple paths $q \to s \to t \to r$ and $r \to q \to s \to t$, we get the path $q \to s \to t \to r \to q \to s \to t$, which is not simple. Indeed, the problem of finding an unweighted longest simple path does not appear to have any sort of optimal substructure. No efficient dynamic-programming algorithm for this problem has ever been found. In fact, this problem is NP-complete, which—as we shall see in Chapter 34—means that we are unlikely to find a way to solve it in polynomial time.

Correct Answer: A

84. Given a graph $G = (V, E)$, suppose a depth-first traversal is performed on the graph G. Assume that each of the vertex $v \in V(G)$ has two attributes associated with it as follows -

I. v.d (discovery time), which records the time when the vertex was first discovered during the traversal, and

II. v.f (finishing time), which records the time when the vertex was explored for the last time during the traversal.

A depth-first forest is the graph G, with the same vertex set V, but with only those edges in $E(G)$ present which were used to perform the depth-first traversal (note that it is a forest, as there can be more than one connected component present in G). Then, for any pair of vertices $u, v \in V(G)$, which of the following conditions hold?

A. The intervals $[u.d, u.f]$ and $[v.d, v.f]$□ are entirely disjoint, and neither u nor v is a descendant of the other in the depth-first forest.

B. The interval $[u.d, u.f]$□ is contained entirely within the interval $[v.d, v.f]$, and u is a descendant of v in a depth-first tree.

C. The interval $[v.d, v.f]$□ is contained entirely within the interval $[u.d, u.f]$, and v is a descendant of u in a depth-first tree.

D. Either I, or II, or III.

Solution:

The correct answer is option (D). This is the Parenthesis Theorem mentioned in the book of "Introduction to Algorithms (3rd edition)" by Cormen et al. (page no. - 606-608), as it elucidates the parenthesis structure of the discovery and finishing times in a depth-first traversal, i.e., if we represent the discovery of vertex u with a left parenthesis " $(u$ " and represent its finishing by a right parenthesis " $u)$", then the history of discoveries and finishings makes a well-formed expression in the sense that the parentheses are properly nested.

Below given is the screenshot of the theorem from the book -

Theorem 22.7 (Parenthesis theorem)

In any depth-first search of a (directed or undirected) graph $G = (V, E)$, for any two vertices u and v, exactly one of the following three conditions holds:

- the intervals $[u.d, u.f]$ and $[v.d, v.f]$ are entirely disjoint, and neither u nor v is a descendant of the other in the depth-first forest,
- the interval $[u.d, u.f]$ is contained entirely within the interval $[v.d, v.f]$, and u is a descendant of v in a depth-first tree, or
- the interval $[v.d, v.f]$ is contained entirely within the interval $[u.d, u.f]$, and v is a descendant of u in a depth-first tree.

And, in case you are interested, the following is a formal proof of the theorem -

The procedure DFS below records when it discovers vertex u in the attribute $u.d$ and when it finishes vertex u in the attribute $u.f$. These timestamps are integers between 1 and $2|V|$, since there is one discovery event and one finishing event for each of the $|V|$ vertices. For every vertex u,

$$u.d < u.f . \tag{22.2}$$

Proof We begin with the case in which $u.d < v.d$. We consider two subcases, according to whether $v.d < u.f$ or not. The first subcase occurs when $v.d < u.f$, so v was discovered while u was still gray, which implies that v is a descendant of u. Moreover, since v was discovered more recently than u, all of its outgoing edges are explored, and v is finished, before the search returns to and finishes u. In this case, therefore, the interval $[v.d, v.f]$ is entirely contained within the interval $[u.d, u.f]$. In the other subcase, $u.f < v.d$, and by inequality (22.2), $u.d < u.f < v.d < v.f$; thus the intervals $[u.d, u.f]$ and $[v.d, v.f]$ are disjoint. Because the intervals are disjoint, neither vertex was discovered while the other was gray, and so neither vertex is a descendant of the other.

The case in which $v.d < u.d$ is similar, with the roles of u and v reversed in the above argument. ■

Correct Answer: D

85. Which of the following is/are correct? (Multiple Select Answers)

A. With all equal-sized intervals, a greedy algorithm based on the earliest start time will always select the maximum number of compatible intervals.

B. The problem of weighted interval scheduling can be solved in $O(n \log n)$ time using dynamic programming.

C. If we divide an array into groups of 3, find the median of each group, recursively find the median of those medians, partition, and recurse, then we can obtain a linear-time median-finding algorithm.

D. If we used the obvious $\Theta\left(n^2\right)$ merge algorithm in the divide-and-conquer convex-hull algorithm, the overall time complexity would be $\Theta\left(n^2 \log n\right)$.

Solution:

A. True. The algorithm is equivalent to the earliest finish time algorithm.

B. True. The algorithm was covered in recitation.

C. False. $T(n) = T(n/3) + T(2n/3) + O(n)$ does not solve to $T(n) = O(n)$. The array has to be broken up into groups of at least 5 to obtain a linear time algorithm.

D. False. The time complexity would satisfy the recurrence $T(n) = 2T(n/2) + \Theta\left(n^2\right)$, which solves to $\Theta\left(n^2\right)$ by the Master Theorem.

Correct Answer: A; B

86. Which of the following is/are correct? (Multiple Select Answers)

A. The quicksort algorithm that uses linear-time median finding to run in worst-case $O(n \log n)$ time requires $\Theta(n)$ auxiliary space.

B. $2^{2n} \in \Theta(2^n)$.

C. If $T(n) = \frac{9}{4}T\left(\frac{2}{3}n\right) + n^2$ and $T(1) = \Theta(1)$, then $T(n) = O\left(n^2\right)$.

D. Performing an $O(1)$ amortized operation n times on an initially empty data structure takes worst-case $O(n)$ time.

Solution:

A. False. It can be implemented with $O(\log n)$ auxiliary space.

B. False. This statement is equivalent to saying $k^2 \in O(k)$ for $k = 2^n$. Constants in exponents matter asymptotically!

C. False. This is an example of Case II of Master Theorem, since $a = \frac{9}{4}, b = \frac{3}{2}, f(n) = n^2$ and $n^2 = \Theta\left(n^{\log_{3/2} 9/4} \log^0 n\right)$. Thus, the recurrence evaluates to $T(n) = \Theta\left(n^2 \log n\right)$, which is not $O\left(n^2\right)$.

D. True. This is the definition of amortization.

Correct Answer: D

87. Which of the following is/are correct? (Multiple Select Answers)

A. Given an array A containing n comparable items, sort A using merge sort. While sorting, each item in A is compared with $O(\log n)$ other items of A.

B. Given a directed graph $G = (V, E)$, run a breadth-first search from a vertex $s \in V$. While processing a vertex u, if some $v \in \text{Adj}^+(u)$ has already been processed, then G contains a directed cycle.

C. Run Bellman-Ford on a weighted graph $G = (V, E, w)$ from a vertex $s \in V$. If there is a witness $v \in V$, i.e., $\delta_{|V|}(s, v) < \delta_{|V|-1}(s, v)$, then v is on a negative weight cycle of G.

D. Floyd-Warshall and Johnson's Algorithm solve all-pairs shortest paths in the same asymptotic running time when applied to weighted complete graphs, i.e., graphs where every vertex has an edge to every other vertex.

Solution:

A. False. As a counter-example, during the final merge step between two sorted halves of the array, each of size $\Theta(n)$, a single item from one array may get compared to all the items from the other list.

B. False. BFS can't be used to find directed cycles. A counterexample is $V = \{s, a, b, t\}$ and $E = \{(s, t), (s, a), (a, b), (b, t)\}$. Running BFS from s will first process vertices in levels $\{s\}$, then $\{a, t\}$, then $\{b\}$. When processing vertex b, vertex $t \in \text{Adj}^+(b)$ has already been processed, yet G is a DAG.

C. False. A witness is only guaranteed to be reachable from a negative weight cycle; it may not actually be on a negative weight cycle.

D. True. A complete graph is dense, i.e., $|E| = \Theta\left(|V|^2\right)$, so Johnson's algorithm runs in $O\left(|V|^2 \log |V| + |V||E|\right) = O\left(|V|^3\right)$ time, which is the same as Floyd-Warshall.

Correct Answer: D

88. Which of the following is/are correct? (Multiple Select Answers)

A. If there is an algorithm to solve 0-1 Knapsack in polynomial time, then there is also an algorithm to solve Subset Sum in polynomial time.

B. The solution to the recurrence $T(n) = 2^n T(n-1)$ is $T(n) = \Theta\left((\sqrt{2})^{n^2+n}\right)$. (Assume $T(n) = 1$ for n smaller than some constant c).

C. The solution to the recurrence $T(n) = T(n/6) + T(7n/9) + O(n)$ is $O(n)$. (Assume $T(n) = 1$ for n smaller than some constant c).

D. In a simple, undirected, connected, weighted graph with at least three vertices and unique edge weights, the heaviest edge in the graph is in no minimum spanning tree.

Solution:

A. True. Subset Sum is the special case of 0-1 Knapsack. Specifically, one can (in linear time) convert an instance (A, T) of Subset Sum into an equivalent instance of 0-1 Knapsack, with an item i for each integer $a_i \in A$ having size $s_i = a_i$ and value $v_i = a_i$, needing to fill a knapsack of size T; and then solve the instance via the polynomial-time algorithm for 0-1 Knapsack.

B. True. Let $T(0) = 1$. Then $T(n) = 2^n \cdot 2^{n-1} \cdot 2^{n-2} \ldots 2^1 = 2^{(n+(n-1)+(n-2)+\ldots+1)}$. Because $\sum_{i=1}^{n} i = n(n+1)/2$, we therefore have $T(n) = 2^{(n^2+n)/2} = (\sqrt{2})^{n^2+n}$. Some students also correctly solved the problem by using the substitution method. Some students made the mistake of multiplying the exponents instead of adding them. Also, it has to be noted that $2^{n^2/2} \neq \Theta\left(2^{n^2}\right)$.

C. True. Using the substitution method:

$$\begin{aligned} T(n) &\leq cn/6 + 7cn/9 + an \\ &\leq 17cn/18 + an \\ &\leq cn - (cn/18 - an) \end{aligned}$$

This holds if $c/18 - a \geq 0$, so it holds for any constant c such that $c \geq 18a$. Full credit was also given for solutions that uses a recursion tree, noting that the total work at level i is $(17/18)^i n$, which onverges to $O(n)$.

D. False. If the heaviest edge in the graph is the only edge connecting some vertex to the rest of the graph, then it must be in every minimum spanning tree.

Correct Answer: A; B; C

89. Which of the following is/are not correct? (Multiple Select Answers)

A. The weighted task scheduling problem with weights in the set $\{1, 2\}$ can be solved optimally by the same greedy algorithm used for the unweighted case.

B. Two polynomials p, q of degree at most $n-1$ are given by their coefficients, and a number x is given. Then one can compute the multiplication $p(x) \cdot q(x)$ in time $O(\log n)$.

C. Suppose we are given an array A of n distinct elements, and we want to find $n/2$ elements in the array whose median is also the median of A. Any algorithm that does this must take $\Omega(n \log n)$ time.

D. There is a density $0 < \rho < 1$ such that the asymptotic running time of the FloydWarshall algorithm on graphs $G = (V, E)$ where $|E| = \rho |V|^2$ is better than that of Johnson's algorithm.

Solution:

A. False. The algorithm will fail given the set of tasks (given in the form $((s_i, f_i), w_i)$: $\{((0,1),1), ((0,2),2)\}$.

B. False. We need at least $\Theta(n)$ time to evaluate each polynomial on x and to multiply the results. Some students argued incorrectly that it must take $O(n \log n)$ using FFT, but FFT overkills because it computes all coefficients, not just one.

C. False. It's possible to do this in linear time using SELECT: first, find the median of A in $\Theta(n)$ time, and then partition A around its median. Then we can take $n/4$ elements from either side to get a total of $n/2$ elements in A whose median is also the median of A.

D. False. The asymptotic running time of Floyd-Warshall is $O\left(V^3\right)$, which is at best the same asymptotic running time of Johnson's (which runs in $O(VE + V \log V)$ time), since $E = O\left(V^2\right)$

Correct Answer: A; B; C; D

90. Which of the following is/are correct? (Multiple Select Answers)

A. Suppose that every operation on a data structure runs in $O(1)$ amortized time. Then the running time for performing a sequence of n operations on an initially empty data structure is $O(n)$ in the worst case.

B. If there is a randomized algorithm that solves a decision problem in time t and outputs the correct answer with probability 0.5, then there is a randomized algorithm for the problem that runs in time $\Theta(t)$ and outputs the correct answer with probability at least 0.99.

C. Let $\Sigma = \{a, b, c, \ldots, z\}$ be a 26-letter alphabet, and let $s \in \Sigma^n$ and $p \in \Sigma^m$ be strings of length n and $m < n$ respectively. Then there is a $\Theta(n)$-time algorithm to check whether p is a substring of s.

D. The recurrence

$$T(n) = \sqrt{n}T(\sqrt{n}) + n$$

has solution $T(n) = O(n \lg\lg n)$.

Solution:

A. True: $\sum_{i=1}^{n} \widehat{c_i} \geq \sum_{i=1}^{n} c_i$.

B. False. Every decision problem has an algorithm that produces the correct answer with a probability 0.5 just by flipping a coin to determine the answer.

C. True. E.g., using suffix trees.

D. True. First, prove a base case. For $4 \leq n \leq 16$, let $T(n) = O(1) \leq k$ for some $k > 0$. For some $c \geq k/4$, we have that $T(n) \leq cn \lg\lg n$. Now, prove the inductive case. Assume $T(n) \leq cn \lg\lg n$ for some $c > 0$. Then:

$$T(n) = \sqrt{n}T(\sqrt{n}) + n \leq \sqrt{n}c\sqrt{n}\lg\lg\sqrt{n} + n$$

We now find $c > 0$ so that

$$\begin{aligned}
\sqrt{n}c\sqrt{n}\lg\lg\sqrt{n} + n &\leq cn\lg\lg n \\
nc(\lg\lg n - \lg 2) + n &\leq cn\lg\lg n \\
-c\lg 2 + 1 &\leq 0 \\
1 &\leq c
\end{aligned}$$

Hence, the inductive case holds for any $c \geq 1$. Setting $c = \max\{k/4, 1\}$, we have $T(n) \leq cn\lg\lg n$ for all $4 \leq n$.

Correct Answer: A; C;D

91. Which of the following is/are correct? (Multiple Select Answers)

A. If we use a max-queue instead of a min-queue in Kruskal's MST algorithm, it will return the spanning tree of maximum total cost (instead of returning the spanning tree of minimum total cost). (Assume the input is a weighted connected undirected graph.)

B. A randomized algorithm for a decision problem with one-sided-error and correctness probability 1/3 (that is, if the answer is YES, it will always output YES, while if the answer is NO, it will output NO with probability 1/3) can always be amplified to a correctness probability of 99%.

C. Suppose that you have two deterministic online algorithms, A_1 and A_2, with a competitive ratios c_1 and c_2 respectively. Consider the randomized algorithm A^* that flips a fair coin once at the beginning; if the coin comes up heads, it runs A_1 from then on; if the coin comes up tails, it runs A_2 from then on. Then the expected competitive ratio of A^* is at least $\min\{c_1, c_2\}$.

D. Suppose that a randomized algorithm A has expected running time $\Theta\left(n^2\right)$ on any input of size n. Then it is possible for some execution of A to take $\Omega\left(3^n\right)$ time.

Solution:

A. True. The proof is essentially the same as for the usual Kruskal's algorithm. Alternatively, this is equivalent to negating all the edge weights and running Kruskal's algorithm.

B. True. Since the error is one-sided, it in fact suffices for the correctness probability to be any constant > 0. We can then repeat it, say, k times, and output NO if we ever see a NO, and YES otherwise. Then, if the correct answer is YES, all k repetitions of our algorithm will output YES, so our final answer is also YES, and if the correct answer is NO, each of our k repetitions has a $1/3$ chance of returning NO, in which case our final answer is, correctly, NO, with probability $1-(2/3)^k$, so $k = \log_{3/2} 100$ repetitions suffice.

C. False. Suppose the problem is to guess which of two cups holds the bean. Algorithm A_1 checks left cup then right cup and has competitive ratio $c_1 = 2$. Algorithm A_2 checks right cup then left cup and has competitive ratio $c_2 = 2$. The randomized algorithm has competitive ratio $c = 0.5 \cdot 2 + 0.5 \cdot 1 = 1.5 < \min\{c_1, c_2\}$

D. True. Imagine a scenario where the runtime of A is $\Theta(3^n)$ with probability $n^2 3^{-n}$, and $\Theta\left(n^2\right)$ otherwise. It is apparent that the expected run time of A is $\Theta\left(n^2\right)$. But, with non-zero probability, some execution may take $\Omega(3^n)$ time.

Correct Answer: A;B;D

92. Let $G = (V, E)$ be a connected undirected graph with edge-weight function $w : E \to \mathbb{R}$. Let $w_{\min}$ and $w_{\max}$ denote the minimum and maximum weights, respectively, of the edges in the graph.

Which of the following is/are correct? (Multiple Select Answers)

A. If the graph G has more than $|V|-1$ edges and there is a unique edge having the largest weight $w_{\max}$, then this edge cannot be part of any minimum spanning tree.

B. Any edge e with weight w_{min}, must be part of some MST.

C. If G has a cycle and there is unique edge e which has the minimum weight on this cycle, then e must be part of every MST.

D. If the edge e is not part of any MST of G, then it must be the maximum weight edge on some cycle in G.

Solution:

A. False. This heavy edge could be the only edge connecting some vertex to a graph and thus must be included in the MST.

B. True. In some ordering of the edges by increasing weight, e will be the first, thus included in the tree constructed by the Kruskal's algorithm. Any tree constructed by the Kruskal's algorithm is an MST.

C. False. Consider the graph of a tetrahedron where the three edges of the base triangle have weights $2, 3, 4$ and the three edges connecting the peak to the base all have weight 1 . Then the MST is composed of those latter three edges and does not include the edge with weight 2 in the bottom cycle.

D. True. Take any MST T. Since e is not part of it, e must complete some cycle in T. e must be the heaviest edge on that cycle. Otherwise, a smaller weight tree could be constructed by swapping the heavier edge on the cycle with e and thus T cannot be MST.

Correct Answer: B;D

93. Which of the following is/are correct? (Multiple Select Answers)

A. If a problem has an algorithm that is correct for 2/3 fraction of the inputs, then it also has an algorithm that is correct for 99.9% of the inputs.

B. A perfect hash table that is already built requires 2 hash functions, one for the first-level hash, and another for the second-level hash.

C. If $T(n) = T(n/2) + T(\sqrt{n}) + n$, then $T(n) = \Theta(n)$.

D. For all asymptotically positive $f(n)$, $f(n) + o(f(n)) = \Theta(f(n))$.

Solution:

A. False. If an algorithm is correct for only 2/3 of the inputs, then it is not necessarily possible to amplify the results to make it correct for those inputs. (Consider, for example, a problem that is easy on 2/3 of the inputs but hard/undecidable for the remaining 1/3 of the inputs.) An algorithm must be correct for 2/3 of the possible sequences of random choices for amplification to apply.

B. False. In the implementation of perfect hashing which we discussed in recitation, the second-level hash table requires a different hash function for every bucket. Note that if we relax the requirement for taking only $O(n)$ space and allow the hash table to instead take $O\left(n^2\right)$ space, then we can use just one hash function.

C. True. The master method does not apply directly. But $\sqrt{n}$ is much smaller than $n/2$, therefore ignore the lower order term and guess that the answer is $T(n) = \Theta(n)$. Check by substitution.

D. True. Clearly, $f(n) + o(f(n))$ is $\Omega(f(n))$. Let $g(n) \in o(f(n))$. For any $c > 0, g(n) \leq c(f(n))$ for all $n \geq n_0$ for some n_0. Hence, $g(n) = O(f(n))$, whence $f(n) + o(f(n)) = O(f(n))$. Thus, $f(n) + o(f(n)) = \Theta(f(n))$.

Correct Answer: C;D

94. Which of the following is/are correct? (Multiple Select Answers)

A. A greedy algorithm for a problem can never give an optimal solution on all inputs.

B. Suppose we have computed a minimum spanning tree of a graph and its weight. If we make a new graph by doubling the weight of every edge in the original graph, we still need $\Omega(E)$ time to compute the cost of the MST of the new graph.

C. A maximum spanning tree (i.e., a spanning tree that maximizes the sum of the weights) can be constructed in $O(E \log V)$ time.

D. Let $G = (V, E)$ be a connected, undirected graph with edge-weight function $w : E$ to reals. Let $u \in V$ be an arbitrary vertex, and let $(u, v) \in E$ be the least-weight edge incident on u; that is, $w(u, v) = \min\{w(u, v') : (u, v') \in E\}$.$(u, v)$ belongs to some minimum spanning tree of G.

Solution:

A. False. Prim's algorithm for minimum spanning trees is a counter-example: it greedily picks edges to cross cuts, but it gives an optimal solution on all inputs.

B. False. Consider sorting the edges by weight. Doubling the weights does not change the sorted order. But this means that Kruskal's and Prim's algorithm will do the same thing, so the MST is unchanged. Therefore the weight of the new tree is simply twice the weight of the old tree and can be computed in constant time if the original MST and weight are known.

C. True. Negate each cost and find a minimum cost spanning tree, by any algorithm shown in the book or in class. (Prim's or Kruskal's) These algorithms make no assumption that the costs are positive, and finding the minimum of the negated costs clearly finds the maximum of the original costs.

D. True. Consider any MST. Assume it doesn't contain (u, v), or else we're done. Since it is a spanning tree, there must be a path from u to v. Let (u, x) be the first edge on this path. Delete (u, x) from the tree and add (u, v). Since (u, v) is a least weight-edge from u, we did not increase the cost, so as long as we still have a spanning tree, we still have a minimum spanning tree. Do we still have a spanning tree? We broke the tree in two by deleting an edge. So did we reconnect it, or add an edge in a useless place? Well, there is only one path between a given two nodes in a tree, so u and v were separated when we deleted e. Thus our added edge does in fact give us back a spanning tree.

Correct Answer: C;D

95. Which of the following is/are correct? (Multiple Select Answers)

A. If the running time of an algorithm satisfies the recurrence

$$T(n) = T(0.1n) + T(0.2n) + T(0.7n) + cn,$$

for some positive constant c, and $T(1) = 1$, then $T(n) = O(n)$.

B. Suppose we are given a connected undirected graph with m edges, such that all edge weights are distinct, except for two edges, which have the same weight. Then this graph can have at most two MSTs.

C. Suppose you are given a collection of sets $S_1, \ldots, S_n$ such that their union is equal to $\{1, 2, \ldots, n\}$, as well as non-negative numbers $w_1, \ldots, w_n$. Your goal is to find a subcollection $i_1, \ldots, i_t$ such that the union of $S_{i_1}, \ldots, S_{i_t}$ is $\{1, 2, \ldots, n\}$, and $w_{i_1} + \ldots + w_{i_t}$ is minimized. Then the greedy algorithm given in class can be modified to find a solution that is within a factor of $\log n$ from the optimal solution.

D. If $T(n) = T(n-1) + O(n)$, then $T(n)$ is at most $O\left(n^2\right)$.

Solution:

A. False. Every level of the recursion tree contributes the same, so the running time of this algorithm is proportional to $n \log n$.

B. True, we have shown in homework and class, that every MST corresponds to a sorting of the edges according to weights. There are only two possible sorted orders.

C. True. The greedy algorithm would pick the set that covers the most elements per weight. If the optimal solution has weight k, then in each step some set must cover elements at the rate of $1/k$. This leads to a solution of cost at most $k \log n$.

D. True. The solution is $O(1)(1 + 2 + \ldots + n) = O\left(n^2\right)$.

Correct Answer: B; C; D

96. Professor Fiorina uses the following algorithm for merging k sorted lists, each having n/k elements. She takes the first list and merges it with the second list using a linear-time algorithm for merging two sorted lists, such as the merging algorithm used in merge sort. Then, she merges the resulting list of $2n/k$ elements with the third list, merges the list of $3n/k$ elements that results with the fourth list, and so forth, until she ends up with a single sorted list of all n elements.

What is the worst-case running time of the professor's algorithm in terms of n and k?

A. $\theta(\sqrt{n}\, k)$

B. $\theta(\sqrt{nk})$

C. $\theta(n^2 k)$

D. $\theta(nk)$

Solution:

Merging the first two lists,each of n/k elements, takes $2n/k$ time. Merging the resulting $2n/k$ elements with the third list of n/k elements takes $3n/k$ time, and so on. Thus for a total of k list, we have:

$$\begin{aligned}
\text{Time } &= \frac{2n}{k} + \frac{3n}{k} + \cdots + \frac{kn}{k} \\
&= \sum_{i=2}^{k} \frac{in}{k} \\
&= \frac{n}{k} \sum_{i=2}^{k} i \\
&= \frac{n}{k} \frac{(k+2)(k-1)}{2} \\
&= \theta(nk)
\end{aligned}$$

Correct Answer: D

97. Which of the following is/are correct? (Multiple Select Answers)

A. A constant $n_0 \geq 1$ exists such that for any $n \geq n_0$, there is an array of n elements such that insertion sort runs faster than merge sort on that input.

B. Given a number a and a positive integer n, the value $a^{2^n} = a^{(2^n)}$ can be computed in $O(n \lg n)$ time by repeated squaring. (Assume that multiplying two numbers takes constant time.)

C. An adversary can force randomized quicksort to run in $\Omega\left(n^2\right)$ time by providing as input an already sorted or reverse-sorted array of size n.

D. Bucket sort is a suitable auxiliary sort for radix sort.

Solution:

A. True. If the n elements are already in sorted order, the running time for insertion sort is $O(n)$, whereas that of merge sort is $O(nlgn)$.

B. True.

$$a^{2^n} = a^{2^{n-1}} a^{2^{n-1}}$$

Letting $T(n) = a^{2^n}$, we have:

$$\begin{aligned} T(n) &= 2T(n-1) + \Theta(1) \\ &= \Theta(n) \\ &= O(n \lg n) \end{aligned}$$

Note: it is important to understand that $\Theta(n) = O(n \lg n)$.

C. False. The running time is not dependent on the input that the adversary provides. This is because randomized quicksort will randomly choose a pivot to partition, independent of the input.

D. True. An auxilliary sort for radix sort must be stable. Bucket sort is stable and hence suitable for use as an auxilliary sort for radix sort.

Correct Answer: A;B;D

98. Which of the following is/are correct? (Multiple Select Answers)

A. Let $\mathcal{H}$ be a class of universal hash functions for a hash table of size $m = n^3$. Then, if we use a random $h \in \mathcal{H}$ to hash n keys into the table, the expected number of collisions is at most $1/n$.

B. Consider two implementations of a hash table with m slots storing n keys, where $n < m$. Let $T_c(m,n)$ be the expected time for an unsuccessful search in the table if collisions are resolved by chaining, using the assumption of simple uniform hashing. Let $T_o(m,n)$ be the expected time for an unsuccessful search in the table if collisions are resolved by open addressing, using the assumption of uniform hashing. Then, we have $T_c(m,n) = \Theta(T_o(m,n))$.

C. If the Huffman code of some character is ' 0 ' or ' 1 ' then this character's frequency in the code is at least 50%.

D. In a graph G = (V, E), if the shortest distance between vertices A and B is a single edge connecting them, this edge is guaranteed to be in some minimum spanning tree of G.

Solution:

A. True. Number of ways to choose 2 keys out of n keys is $\frac{n(n-1)}{2}$ Therefore, the expected number of collisions,

$$\begin{aligned} \text{E[number of collisions]} &= \frac{n(n-1)}{2}\frac{1}{m} \\ &= \frac{n(n-1)}{2}\frac{1}{n^3} \\ &\leq \frac{1}{n} \end{aligned}$$

B. False.

$$T_c(m,n) = \Theta\left(1+\frac{n}{m}\right)$$
$$T_o(m,n) = \Theta\left(\frac{1}{1-\frac{n}{m}}\right)$$

In the worst case, in the chaining version, we need to search through all the values in one particular slot. On the other hand, in the open-addressing version, we need to search through every slot. Hence $T_c(m,n) \neq \Theta(T_o(m,n))$.

C. False.

Example 1: 'a' has freq. 1/100, 'b' has freq. 99/100, an optimal coding is 'a' gets 0 ' and 'b' gets ' 1 '.

Example 2; 'a' 'b' and 'c' each has freq. 1/3. an optimal coding: 'a'-'0', 'b'-'10', c-'11'.

D. False. Consider a cycle of 4 vertices a-b-c-d-a and $w(a,b) = w(b,c) = w(c,d) = 1, w(d,a) = 2$. The shortest path connecting a and d is the edge (a, d), the single MST is a-b-c-d.

Correct Answer: A

99. Which of the following is/are correct? (Multiple Select Answers)

A. Every comparison-based sort uses at most $O(n\log n)$ comparisons in the worst case.

B. RADIX-SORT is stable if its auxiliary sorting routine is stable.

C. Randomized-Select can be forced to run in $\Omega(n\log n)$ time by choosing a bad input array.

D. It is possible to compute the smallest $\sqrt{n}$ elements of an n-element array, in sorted order, in $O(n)$ time.

Solution:

A. False. INSERTION-SORT, for example, uses $\Theta\left(n^2\right) \neq O(n\log n)$ comparisons in the worst-case (a reverse-sorted array). The statement would be true if it read "... at least $\Omega(n\log n)$ comparisons in the worst case."

B. True. If two numbers are equal, then they have the same digits. Each intermediate sort is stable, so the two equal numbers never change relative positions.

C. False. Randomized-Select runs in expected $O(n)$ time; the only way it can take longer is if its random choices of pivots are unlucky. The input array cannot force these unlucky choices.

D. True. We can SELECT the $\sqrt{n}$ th smallest element and partition around it, then sort those $\sqrt{n}$ elements in $O(n)$ time. Alternately, we can build a minheap in $O(n)$ time and call EXTRACT-Min $\sqrt{n}$ times, for a total runtime of $O(n+\sqrt{n}\log n) = O(n)$.

Some incorrect solutions amounted to: "we must do $\sqrt{n}$ order statistic queries, each of which take $O(n)$ time, for a total running time of $O(n\sqrt{n})$." However, this argument does not preclude us from coming up with a more clever algorithm (like the one above) that is more efficient. In fact, a similar argument would "prove" that sorting must take $\Omega\left(n^2\right)$ time (despite the existence of MERGESORT etc.), because we must do n order statistic queries!

Correct Answer: B;D

100. Which of the following is/are correct? (Multiple Select Answers)

A. Suppose that we have a hash table with $2n$ slots, with collisions resolved by chaining, and suppose that $n/2$ keys are inserted into the table. Each key is equally likely to be hashed into each slot (simple uniform hashing). Then the expected number of keys for each slot is 1/4.

B. For every two functions $f(n)$ and $g(n)$, either $f(n) = O(g(n))$ or $g(n) = O(f(n))$.

C. Suppose we use HEAPSORT instead of INSERTION-SORT as a subroutine of BUCKET-SORT to sort n elements. Then BUCKET-SORT still runs in average case linear time, but its worst-case running time is now $O(n \log n)$.

D. If memory is limited, one would prefer to sort using HEAPSORT instead of MERGESORT.

Solution:

A. True. Define X_j (for $j = 1, \ldots, n/2$) to be the indicator which is 1 if element j hashes to slot i, and 0 otherwise. Then $E\left[X_j\right] = \Pr\left[X_j = 1\right] = 1/2n$. Then the expected number of elements in slot i is $E\left[\sum_{j=1}^{n/2} X_j\right] = \sum_{j=1}^{n/2} E\left[X_j\right] = n/4n = 1/4$ by linearity of expectation.

B. False. Let $f(n) = \sin n$ and $g(n) = \cos n$; then neither case holds. Another example is $f(n) = \sqrt{n}$ and $g(n) = n^{\sin n}$. Finally, one could let $f(n)$ and $g(n)$ be any strictly-negative functions; by a technical condition of the definition, $f(n)$ must be at least 0 to be $O(g(n))$.

Many incorrect answers argued that one of the statements $f(n) \leq cg(n), f(n) = cg(n)$, or $g(n) \leq cf(n)$ must be true. This is correct for any particular value of n, but it doesn't mean that the same statement is true for all sufficiently large values of n, which is the condition needed in the definition of big- O.

C. True. Even if all the elements land in the same bucket (the worst-case input), HEAPSORT sorts them in $O(n \log n)$ time.

D. True. MERGESORT is not in-place, which means it requires an auxiliary array as big as the input. HEAPSORT is in-place, which means it only uses $O(1)$ auxiliary space.

Correct Answer: A; C; D

5 Databases (20 Questions)

101. For relational algebra expressions, which of the following options is correct?

I. $(R \bowtie S) \bowtie T = (T \bowtie R) \bowtie S$

II. $\sigma_{\mathrm{A}}(\sigma_{\mathrm{B}}(\mathrm{R})) = \sigma_{\mathrm{B}}(\sigma_{\mathrm{A}}(\mathrm{R}))$

III. $\Pi_A(\Pi_B(R)) = \Pi_B(\Pi_A(R))$

IV. $R \bowtie (S \cap T) = (R \bowtie S) \cap (R \bowtie T)$

V. $\sigma_A(R \bowtie S) = \sigma_A(R) \bowtie S$

A. I, II, IV

B. I, II, III, IV

C. II, IV, V

D. All of them

Solution:

III. Fields in expression A might depend on B, so swapping the order won't work.

V. This only works if A references just fields in relation R.

Correct Answer: A

102. Which of these statements about serializable schedules is true?

A. Every serializable schedule is recoverable.

B. Every serializable schedule contains no conflicting actions.

C. Every 2PL schedule is serializable.

D. None of the above.

Solution:

Correct Answer: C

103. Which of these statements about recoverable schedules is true?

A. Every recoverable schedule is serializable.

B. In a recoverable schedule, if a transaction T commits, then any other transaction that T read from must also have committed.

C. In a recoverable schedule, no transaction will ever be aborted because a transaction that it read from has aborted.

D. None of the above.

Solution:

Correct Answer: B

104. In which of the following situations is optimistic concurrency control with validation likely to perform better than locking with 2PL?

A. A high-contention workload where all the transactions need to update a single record.

B. A read-mostly workload, where most transactions just read a small number of data items, and a few transactions write data items.

C. A distributed database where all the transactions need to read and write objects on multiple servers.

D. All of the above.

Solution:

Correct Answer: B

105. Which of these statements about deadlock are true? (Multiple Select Answers)

A. If all transactions use two-phase locking, they cannot deadlock.

B. Once two transactions deadlock, one of them must be aborted to maintain correctness.

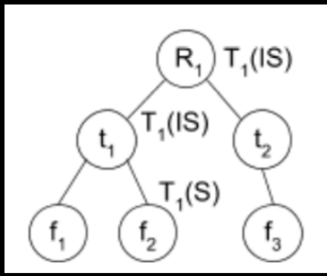
C. Systems that support update locks (S,X and U modes) cannot deadlock.

D. Validation-based concurrency control schemes cannot deadlock.

Solution:

Correct Answer: B;D

106. In the database below with hierarchical locking, transaction T_1 holds some locks. Which of the following transactions can still acquire the locks they need to run? (assume that only one of these candidate transactions would run at a time.) (Multiple Select Answers)



A. T_2 : write t_1.

B. T_3 : write t_2.

C. T_4 : read f_2 and write f_1.

D. T_5 : insert a child node under t_2.

Solution:

Correct Answer: B;C;D

107. You are running validation on transaction T_2. The only other transaction is T_1, which has already validated but has not finished. T_1 has a read set of {A, B} and write set {B, C}. Which read and write sets of T_2 will allow it to validate? (Multiple Select Answers)

A. RS = {D}, WS = {A}

B. RS = {E}, WS = {D, F}

C. RS = {B, A}, WS = {D}

D. RS = {A}, WS = {C}

Solution:

Correct Answer: A:B

108. For each statement below, which of the following options is incorrect?

i. Using finer locks can always be more efficient than coarser locks.

ii. Two-phase locking is guaranteed to produce a schedule whose precedence graph is acyclic.

iii. Two-phase locking ensures recoverability.

iv. Strict two-phase locking is guaranteed to produce a conflict serializable schedule.

v. A system providing exclusive access on each object (locking before accessing and unlocking afterwards) guarantees serializability.

A. i, iii, iv

B. i, iii, v

C. i, ii, iii, v

D. All of them

Solution:

Correct Answer: B

109. Consider a table People with attributes SALARY and NAME. Given the relational algebra expression Π_{NAME} ($\sigma_{\text{SALARY} > 10,000}$ ($\Pi_{\text{NAME}} \cup$ SALARY (People)), which of the following are valid rewrites? (Multiple Select Answers)

A. Π_{NAME} ($\sigma_{\text{SALARY} > 10,000}$(People))

B. $\sigma_{\text{SALARY} > 10,000}$ ($\Pi_{\text{NAME}} \cup$ SALARY (People)

C. $\sigma_{\text{SALARY}>10,000}$ (Π_{NAME} ($\Pi_{\text{NAME} \cup \text{SALARY}}$ (People))

D. Π_{NAME} ($\sigma_{\text{SALARY}>10,000}$ (Π_{NAME} (Π_{SALARY}(People)))

Solution:

Correct Answer: A

110. Suppose that we want to join a table R with fields $a_1, \ldots, a_n$ against n other tables $S_1, \ldots, S_n$, where each attribute a_i in R matches a key in the corresponding table S_i. This is called a "star join". How many valid left-deep join plans exist for this query?

A. $\Theta(n)$

B. $\Theta(2^n)$

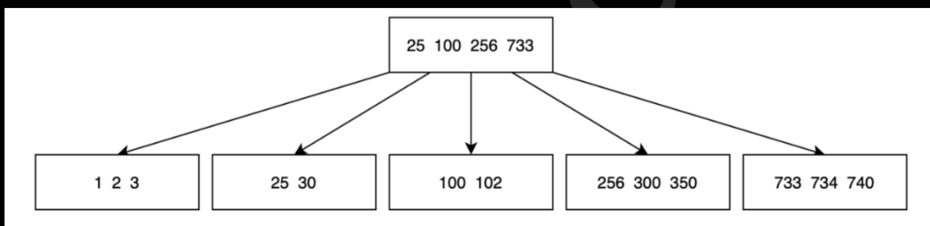
C. $\Theta(n!)$

D. $\Theta(n^n)$

Solution:

Correct Answer: C

111. Considering the following B+ tree:



The order n of the B+ tree is 4 . Follow the same rules taught in class of the maximum and minimum pointers or keys of the tree.

Insert elements in the following order: INSERT(1000), INSERT(301), INSERT(735)

During all the operations, there are x leaf overflows and y non-leaf overflows.

After all the operations, the root node contains z keys.

The value of x+y-z is equal to ________. (Numerical Answer Type)

Solution:

The first and second insertion add one key in the leftmost and second-leftmost leaf node. No overflow happens till now. The third insertion tries to insert into the leftmost leaf node once again. The number of keys of that node will be $5 > 4$, and one leaf overflow happens. The current root node should insert a key and a pointer to this newly split leaf node, but the keys would be $5 > 4$, so one non-leaf overflow happens. A new root with 1 key and 2 pointers will be created.

So, $x = 1, y = 1, z = 1$

$\therefore$ The value of $x + y - z = 1 + 1 - 1 = 1$.

Correct Answer: 1

112. Assume there are M records for B+ tree to index. In the first case, there are N separate exact search queries on this index. In the second case, there is a range query on this index with results spanning N leaf nodes. How many pointers in the B+ tree will these two cases traverse separately to retrieve the results?

A. $\Theta(\mathrm{MN})$

B. $\Theta(N \log \mathrm{M})$

C. $\Theta(N + \log M)$

D. $\Theta(M + \log N)$

The first case: _______ , the second case: _______. (Fill in one of A, B, C, or D for each case).

Solution:

The first case: B, the second case: C.

113. Consider a relation $R(A, B, C, D)$ with the following set of dependencies:

$$A \to \spadesuit, \quad CD \to A, \quad A \to CD$$

Unfortunately, we don't know what $\spadesuit$ is-it could be any nonempty subset of R's attributes. However, we do know that the given set of dependencies form a minimal basis; that is, every dependency in this set is nontrivial and cannot be inferred from the others in the set.

Which one of the following must be true regardless of what is inside $\spadesuit$?

I. R is in BCNF

II. R is not in BCNF

III. R is in 3NF

IV. R is not in 3NF

A. II and IV

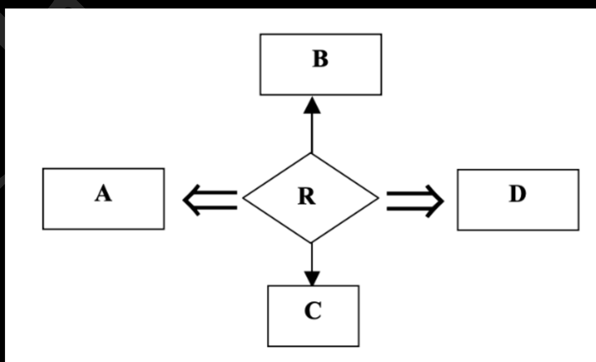
B. II

C. II and III

D. I and III

Solution:

114. Consider the following ER diagram.



The key of an entity set **A**, is a1. Similarly, the keys for entity sets **B**, **C**, and **D** are b1, c1, and d1 respectively. If we translate the relationship set **R** into a relation R(a1, b1, c1, d1), what are all the candidate keys of the relation R?

A. {a1, b1}, {a1, c1}, {a1, d1}, {b1, c1}, {b1, d1}, and {c1, d1}

B. {a1}, {b1}, {c1}, and {d1}

C. { a1, b1, c1, d1 }

D. {a1, b1, c1}, {a1, b1, d1}, {a1, c1, d1}, {b1, c1, d1}

Solution:

Correct Answer: D

115. Refer to a relation $S(A,B,C,D,E)$ with the FD's:

$$AB \to C, \quad B \to D, \quad DE \to A$$

$B \to D$ is a BCNF violation for S.

Suppose we decide to decompose S into $S_1(B,D)$ and $S_2(A,B,C,E)$.

Which of the following statements are true?

I. $\{AB \to C\}$ is a minimal basis for the FD's that hold in S_2

II. $AB \to C$ is a BCNF violation for S_2

III. S_2 should be decomposed further into $S_3(A,B,C)$ and $S_4(C,E)$

A. II only

B. I and II only

C. II and III only

D. I, II, and III

Solution:

Correct Answer: D

116. Refer to a relation $S(A,B,C,D,E)$ with the FD's:

$$AB \to C, \quad B \to D, \quad DE \to A$$

Which of the following statements are true?

I. Instead of decomposing S using $B \to D$, we could decompose S using $DE \to A$ first.

II. It does not matter whether we start with $B \to D$ first or $DE \to A$ first: at the end of the BNCF decomposition algorithm we will get the same set of relations.

A. I only

B. II only

C. Both I and II

D. Neither I nor II

Solution:

117. Refer to a relation $T(A,B,C,D,E)$ with the following FD's:

$$A \longrightarrow BC, \quad CD \longrightarrow E, \quad \clubsuit \longrightarrow D$$

Unfortunately, we don't know what $\clubsuit$ is - it could be any nonempty subset of T 's attributes. (In particular, $\clubsuit$ might even contain D itself, which would make $\clubsuit \longrightarrow D$ a trivial dependency.)

Which of the following must be true regardless of what is inside $\clubsuit$?

A. Every key of T contains A

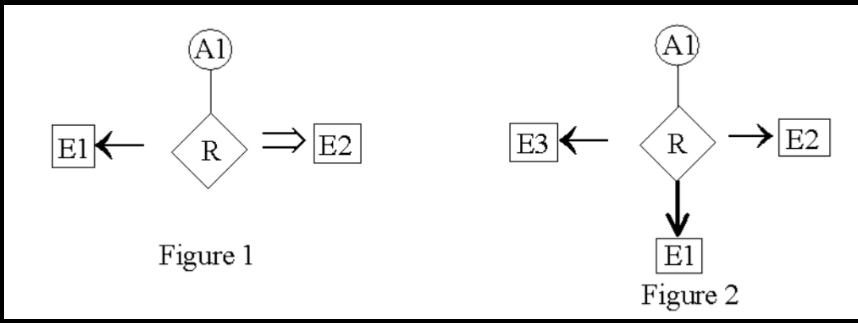
B. No key of T contains A

C. Some key of T contains A while some other key does not

D. None of the above

Solution:

118. Which of the following statements are true?



1. (Refer to figure 1) Consider a one-to-one binary relationship R between entity set E1 and E2, in which E2 participation is total. It is possible to represent the same information by migrating attribute A1 of relationship R to entity set E1 but not to entity set E2.

2. (Refer to figure 2) Consider a one-to-one-to-one ternary relationship R between entity set E1, E 2 , and E3. It is possible to represent the same information by migrating attribute A1 of relationship R to either of entity sets E1, E2, or E3.

A. only 1 is true

B. only 2 is true

C. both 1 and 2 are true

D. neither 1 nor 2 are true

Solution:

Correct Answer: D

119. Suppose we have two relations $R(\underline{A},B)$ and $S(\underline{A},B)$ with the same schema. The only key of R is $\{A\}$; the only key of S is $\{A\}$ as well.

Let relation $T(A,B)$ be the set union of R and S, i.e., $T = R \cup S$. What are the keys of T?

A. $\{A\}$

B. $\{B\}$

C. $\{A\}$ and $\{B\}$

D. $\{A,B\}$

Solution:

120. Refer to the following database schema:

```
Person(SSN, employerSymbol, salary)
Holding (SSN, symbol, numShares)
```

A person is uniquely identified by a social security number (SSN). A company is uniquely identified by its stock ticker symbol. Each person is employed by exactly one company, but may hold any number of different stocks.

Suppose we wish to find the average salary of the persons who own more than 100 shares of Microsoft (MSFT) or more than 100 shares of Yahoo! (YHOO). Which of the following queries will correctly compute the desired average?

I.

```
SELECT AVG (salary)
FROM Person
WHERE SSN IN (SELECT SSN FROM Holding
             WHERE (symbol = 'MSFT' OR symbol = 'YHOO')
             AND numShares > 100);
```

II.

```
SELECT AVG (salary)
FROM Person, Holding
WHERE Person.SSN = Holding.SSN
AND ((symbol = 'MSFT' AND numShares > 100) OR
    (symbol = 'YHOO' AND numShares > 100));
```

A. I only

B. II only

C. Both I and II

D. Neither I nor II

Solution:

6 Operating System (20 Questions)

121. Which of the following is/are correct? (Multiple Select Answers)

A. A user-level process cannot modify its own page table entries.

B. The scheduler is the part of an Operating System that determines the priority of each process.

C. Shortest Remaining Time First is the best preemptive scheduling algorithm that can be implemented in an Operating System.

D. The working set model is used to compute the average number of frames a job will need in order to run smoothly without causing thrashing.

Solution:

A. TRUE. If a user-level process was allowed to modify its own page table entries, then it could access physical memory being used by other processes or the OS kernel. Kernel mode is required to modify page table entries.

B. FALSE. The scheduler schedules processes based on user-specified priorities.

C. FALSE. SRTF cannot be implemented because it requires knowledge of the future.

D. FALSE. The working set model is used to compute the minimum (total) number of frames a job will need in order to run smoothly without causing thrashing.

Correct Answer: A

122. Recall the various deadlock detection and prevention algorithms we've discussed in this course, and consider the following snapshot of a system with five processes (P1, P2, P3, P4, P5) and four resources (R1, R2, R3, R4). There are no current outstanding queued unsatisfied requests.

Currently Available Resources

R1	R2	R3	R4
2	1	2	0

	Current Allocation				Max Need				Still Needs			
Process	**R1**	**R2**	**R3**	**R4**	**R1**	**R2**	**R3**	**R4**	**R1**	**R2**	**R3**	**R4**
P1	0	0	1	2	0	0	3	2	0	0	2	0
P2	2	0	0	0	2	7	5	0	0	7	5	0
P3	0	0	3	4	6	6	5	6	6	6	2	2
P4	2	3	5	4	4	3	5	6	2	0	0	2
P5	0	3	3	2	0	6	5	2	0	3	2	0

The system currently is not deadlocked, what is the process execution order?

A. P1, P2, P5, P4, P3.

B. P2, P4, P5, P1, P3.

C. P1, P4, P5, P2, P3.

D. P1, P2, P3, P4, P5.

Solution:

Using the Banker's algorithm, the system is not deadlocked and will not become deadlocked. The process finishing order is: P1, P4, P5, P2, P3.

Correct Answer: C

123. Consider a memory system with a cache access time of 10 ns and a memory access time of 110 ns - assume the memory access time includes the time to check the cache. If the effective access time is 10% greater than the cache access time, what is the hit ratio H? (up to two-digit of decimal) (Multiple Select Answers)

A. 99/100

B. 100/99

C. 0.99

D. 0.89

Solution:

$\text{Effective Access Time } = H^* T_{\text{cache}} + (1-H)^* T_{\text{memory}}$

$1.1 * T_{\text{cache}} = H^* T_{\text{cache}} + (1-H)^* T_{\text{memory}}$

$1.1 \times 10 = H^* 10 + (1-H)110$

$11 = H^* 10 + 110 - 110^* H$

$-99 = -100^* H$

$H = 99/100$

Correct Answer: A; C

124. Suppose that we build a disk subsystem to handle a high rate of I/O by coupling many disks together. Properties of this system are as follows:

- Uses 10 GB disks that rotate at 10,000 RPM, have a data transfer rate of 10MBytes/s (for each disk), and have an 8 ms average seek time, 32 KByte block size.
- Has a SCSI interface with a 2 ms controller command time.
- Is limited only by the disks (assume that no other factors affect performance).
- Has a total of 20 disks.

Each disk can handle only one request at a time, but each disk in the system can be handling a different request. The data is not striped (all I/O for each request has to go to one disk).

What is the average service time to retrieve a single disk block from a random location on a single disk, assuming no queuing time (i.e. the unloaded request time)? Hint: There are four terms for this service time!

A. 16.28 ms

B. 16.38 ms

C. 16.18 ms

D. 16.20 ms

Solution:

$\text{Time}_{\text{service}} = \text{Time}_{\text{controller}} + \text{Time}_{\text{seek}} + \text{Time}_{\text{rotational}} + \text{Time}_{\text{transfer}}$

$\text{Time}_{\text{contoller}} = 2 \text{ ms}$

$\text{Time}_{\text{seek}} = 8 \text{ ms}$

$\text{Time}_{\text{rotational}} = 1/2 \text{ Time}_{\text{rotation}} = 1/2 \times [(60 \text{ s/M})/(10000\text{RPM})] = 0.003 \text{ s} = 3 \text{ ms}$

$\text{Time}_{\text{transfer}} = (32 \times 1024 \text{ bytes })/\left(10 \times 10^6 \text{ bytes / sec }\right) = 0.0032768 \text{ s} = 3.2768 \text{ ms}$

$\text{Time}_{\text{service}} = 2 + 8 + 3 + 3.28 \text{ ms} = 16.28 \text{ ms}$

Correct Answer: A

125. Consider the following two threads, to be run concurrently in a shared memory (all variables are shared between the two threads):

Thread A	Thread B

for $(i = 0; \quad i < 5; \quad i++)$ {
 $\quad x = x + 1;$
 }
	for $(j = 0; \quad j < 5; \quad j++)$ {
	 $\quad x = x + 2;$
	 }

Assume a single-processor system, that load and store are atomic, that x is initialized to 0 before either thread starts, and that x must be loaded into a register before being incremented (and stored back to memory afterwards).

What is/are the final value of x after both threads have completed? (Multiple Select Answers)

A. $x \geq 0$

B. $x \leq 15$

C. $x > 0$

D. $x < 15$

Solution:

Each $x = x + 1$ statement can either do nothing (if erased by Thread B) or increase x by 1. Each $x = x + 2$ statement can either do nothing (if erased by Thread A) or increase x by 2. Since there are 5 of each type, and since x starts at $0, x$ is ≥ 0 and $x \leq (5 \times 1) + (5 \times 2) = 15$.

Correct Answer: A; B

126. Which of the following is/are correct? (Multiple Select Answers)

A. The size of an inverted page table grows with the size of the virtual address space that it is supporting.

B. Threads within the same process can share data with one another by passing pointers to objects on their stacks.

C. Resource cycles always lead to deadlock.

D. Anything that can be done with a monitor can also be done with semaphores.

Solution:

A. False: Since an inverted page table holds a hash on entries, it only needs to grow with the size of the physical address space. Entries not found in the inverted page table can be assumed to be invalid.

B. True: Since the threads in the same process share an address space, they can all access the same memory (including each other's stacks).

C. False: No. If there are multiple equivalent resources, then a cycle could exist that wasn't a deadlock: The reason is that some thread that wasn't a part of the cycle could release a resource needed by a thread in the cycle, thereby breaking the cycle.

D. True: Since one can construct monitors using semaphores, one could implement any monitor-based algorithm with semaphores by simply replacing the monitor implementation (i.e. locks and condition variables) with one using semaphores.

Correct Answer: B; D

127. Which of the following is/are correct? (Multiple Select Answers)

A. Each thread has its own stack.

B. Starvation implies deadlock.

C. Shortest Run Time First (SRTF) is the "optimal" scheduling algorithm, but it is generally not implemented directly, due to excessive context switching overhead.

D. Unlike paging, segmentation doesn't prevent processes from accessing physical memory not allocated to them.

Solution:

A. Yes, the thread is the unit of execution, and it uses the stack to store the return addresses and arguments passed to functions/methods.

B. FALSE: A system my exit from starvation, but not from deadlock (without external intervention).

C. It is not implemented because it is impossible to predict the future.

D. Segmentation avoids access to unallocated memory by checking that an address is within segment limits.

Correct Answer:

128. Which of the following is/are correct? (Multiple Select Answers)

A. A thread that wants to signal other threads upon completion of a critical section should do so with a Condition Variable after releasing the lock.

B. The offset within the page of a data element is the same in virtual memory page and physical memory page, regardless of whether a single level or multi-level page table is used.

C. With multi-level page tables, the TLB holds translations for each of the level.

D. The size of page tables is always much smaller than the size of physical memory.

Solution:

Correct Answer: B

129. Which of the following is true about File System operation? (Multiple Select Answers)

A. The directory structure can be searched on Open and need not be examined on Read or Write.

B. File Index structures are optimized to handle small and large files in the same manner.

C. SCAN has shorter average seek distance than C-SCAN because it services requests while the head is moving in either direction.

D. If the transaction log is stored on disk, requests must be written in order, as with FIFO scheduling.

Solution:

Correct Answer: A; D

130. Which of the following is/are correct? (Multiple Select Answers)

A. A direct mapped cache can sometimes have a higher hit rate than a fully associative cache with an LRU replacement policy (on the same reference pattern).

B. If the Banker's algorithm finds that it's safe to allocate a resource to an existing thread, then all threads will eventually complete.

C. Even though most processors use a physical address to address the cache, it is possible to overlap the TLB lookup and cache access in hardware.

D. You can always reduce the number of page faults by increasing the amount of memory available to a process.

Solution:

A. Consider a cache of N cache lines with an $N+1$ sequential access pattern whose addresses stride across cache lines (i.e. with 32-byte cache lines, address $0, 32, 64, \ldots 32$ N, $0, 32, 64, \ldots. 32N, \ldots.$). The LRU cache would miss on every access (since the $N+1^{st}$ entry is never in the cache); the direct-mapped cache would miss on only 2 out of N + 1 accesses, with only 0 and 32 N conflicting.

B. Threads fail to complete for many reasons other than cycles in the resource graph. For instance, a thread could go into an infinite loop independent of the Banker's algorithm.

C. Not all bits in a Virtual address are translated during mapping to Physical address. Particularly, the page offset stays the same. If the cache index fits entirely in the page offset, then the SRAM lookup of the cache can happen in parallel with the TLB lookup, with the Tag check happening after the TLB access finishes. [For instance, with $4K$ pages and a 16-way, 64 K cache, the cache index would come entirely from the offset. For larger caches, those bits of the index that fit inside the offset could start the cache lookup, with the remainder happening after the TLB lookup completes.]

D. This result depends on the replacement policy. A FIFO page replacement policy exhibits Belady's anomaly, in which certain access patterns experience increased page faults with increased memory.

Correct Answer: A; C

131. Which of the following is/are correct? (Multiple Select Answers)

A. The Shortest Remaining Time First (SRTF) algorithm is the best preemptive scheduling algorithm that can be implemented in an Operating System.

B. Anything that can be done with a monitor can also be done with semaphores.

C. Page Tables have an important advantage over simple Segmentation Tables (i.e. using base and bound) for virtual address translation in that they eliminate external fragmentation in the physical memory allocator.

D. In the Fast File System (FFS) without a buffer cache, the number of disk accesses to retrieve the first byte of a file is always less than the number of disk accesses to retrieve the last byte of a file.

Solution:

A. Since SRTF relies on knowledge of the future, it cannot be implemented. It is merely a strawman against which to compare practical scheduling algorithms.

B. Since it is possible to create both locks and condition variables out of semaphores, once can simply build monitors out of semaphores, then use these monitors to do whatever we wantthus demonstrating the equivalence.

C. Since page tables operate on pages, which are of fixed size, the physical memory allocator can use any physical page for any part of a virtual address space; thus, there is no externally wasted physical memory. In a simple segmentation scheme, however, the chunks of physical memory vary widely in size (based on the size of the segments); thus, the physical memory allocator is forced to find variable sized chunks - leading to external fragmentation.

D. For small files that can be described entirely with direct pointers, the first and last byte takes the same number of disk accesses, namely two (one for inode, one for datablock).

Correct Answer: B; C

132. Which of the following is/are correct? (Multiple Select Answers)

A. The number of allocated kernel stacks is always equal to the number of allocated userspace stacks.

B. When fork returns a negative integer, an error has occurred that must be dealt with in both the parent and child process.

C. Switching between two threads within the same process is generally more efficient than switching between two threads belonging to different processes.

D. Locks can prevent a thread from being preempted.

Solution:

A. If a process is running a user-level threading library, then there can be more stacks in userspace (one per user-level thread) than in the kernel (just on kernel thread for the process). Additionally, there are kernel threads that perform tasks entirely within the operating system and do not require a userspace stack.

B. When fork returns a negative integer, an error has occurred, but no child process is actually created.

C. Switching between two threads always requires changing the execution context (registers, stack pointer, program counter, etc.). For two threads belonging to different processes, the OS must also change the current address space, which incurs an additional cost.

D. Locks do not prevent a thread from being preempted. Threads can be interrupted during a critical section. Locks only guarantee that the critical section is only entered by one thread at a time.

Correct Answer: C

133. Consider a machine that requires all addresses by mapped through its 128 entry TLB. What is the maximum amount of memory that can be accessed, without incurring any TLB faults, if the page size is 4 KB?

A. 128 KB

B. 128 MB

C. 512 KB

D. 4 GB

Solution:

Correct Answer: C

134. What is the total number of processes at the end of the execution of the following program? Assume there is one process in the beginning that starts running at main. Also, assume that all system calls succeed. (Numerical Answer Type)

```
main() {
    fork();
    fork();
    fork();
}
```

Solution:

fork#	#processes
1	2
2	4
3	8

First fork is called by the first process, it generates one more process (or doubles). Both these processes call second fork and doubles the number of processes to 4… and so on.

Correct Answer: 8

135. Which of the following are stored in a process control block (PCB) of a process? (Multiple Select Answer)

A. Stack pointer

B. Program counter

C. Register values

D. Process data files

Solution:

While the PCB of a process stores file descriptors of files in use by the process, the PCB does not store the files themselves

Correct Answer: A; B; C

136. Which of the following happens after a fork() call is invoked? (Multiple Select Answer)

A. The child process is created and is an identical clone of the parent.

B. The return value of fork() is set to 0 for the child process.

C. The return value of fork() is set to 0 for the parent process.

D. The return value of fork() is set to the child's pid for the child process.

E. The return value of fork() is set to the child's pid for the parent process

Solution:

Which of the following happens after a fork() call is invoked? (multiple choice question with multiple correct answers) A . Incorrect. The child process is created and is an identical clone of the parent. On fork(), the child's memory is a nearly an exact copy of the parent's address space, and both have the same PC value. However, the return value of fork() is different, hence they are not identical.

B. Correct. The return value of fork() is set to 0 for the child process. To differentiate the parent and child, the return value of fork() is used. The return value will be 0 if we are in the child process and the OS-assigned PID value within the parent process.

C. Incorrect. The return value of fork( ) is set to 0 for the parent process. To differentiate the parent and child, the return value of fork() is used. The return value will be 0 if we are in the child process and the OS-assigned PID value within the parent process.

D. Incorrect. The return value of fork() is set to the child's pid for the child process.

To differentiate the parent and child, the return value of fork() is used. The return value will be 0 if we are in the child process and the OS-assigned PID value within the parent process.

E. Correct. The return value of fork() is set to the child's pid for the parent process To differentiate the parent and child, the return value of fork() is used. The return value will be 0 if we are in the child process and the OS-assigned PID value within the parent process.

Correct Answer: B; E

137. Consider the following program:

```c
#include <stdlib.h>
#include <stdio.h>
int sum(int varX, int varY) {
    char *cArray = NULL;
    cArray = (char*) malloc (1024 * sizeof(char));
    static int funcCounter = 0;
    funcCounter++;
    free(cArray);
    return (funcCounter + varX + varY);
}
int main (int argc, char** argv) {
    int x = 0, y = 0;
    x = sum(1,2);
    y = sum (3,4);
    printf("%d %d\n", x, y);
    return 0;
}
```

Where is the cArray stored in memory? (Multiple Select Answers)

A. Heap segment

B. Stack segment

C. Data segment

D. None of the above

Solution:

A. Correct. Heap segment

B. Incorrect. Stack segment

C. Incorrect. Data segment

All memory allocated via malloc() calls go into the heap, even if they are called within functions.

Correct Answer: A

138. Which of the following statements is/are true about a context switch? (Multiple Select Answers)

A. A context switch from one process to another will happen every time a process moves from user mode to kernel mode

B. For preemptive schedulers, a trap of any kind always leads to a context switch

C. A context switch will always occur when a process has made a blocking system call, irrespective of whether the scheduler is preemptive or not

D. For non-preemptive schedulers, a process that is ready/willing to run will not be context switched out

Solution:

Correct Answer: C; D

139. Assume we have a demand-paged memory. The page table is held in registers. It takes 8 milliseconds to service a page fault if an empty page is available or the replaced page is not modified and 20 milliseconds if the replaced page is modified. Memory access time is 100 nanoseconds. Assume that the page to be replaced is modified 70 percent of the time. What is the maximum acceptable page-fault rate for an effective access time of no more than 200 nanoseconds?

A. 6.1×10^{-6}

B. 6.1×10^{-3}

C. 6.1×10^{-9}

D. 6.1×10^{6}

Solution:

Let's used variable x for the page-fault rate. The average time consumption for pages that are available on the main memory is,

$$(1-x) \times 100 \cdot 10^{-9}$$

The average time consumption for page faults if empty pages are available or the replaced pages are not modified is,

$$x \times (1-70\%) \times 8 \cdot 10^{-3}$$

The average time consumption for page faults if the replaced pages are modified is,

$$x \times 70\% \times 20 \cdot 10^{-3}$$

Therefore the total average access time is no more than 200 nanoseconds indicates,

$$\begin{aligned}
&(1-x) \times 100 \cdot 10^{-9} + x \times (1-70\%) \times 8 \cdot 10^{-3} + x \times 70\% \times 20 \cdot 10^{-3} \leq 200 \times 10^{-9} \\
&\Rightarrow -x + 140000x + 24000x \leq 1 \\
&\Rightarrow 163999x \leq 1 \\
&\Rightarrow x \leq 6.1 \times 10^{-6}.
\end{aligned}$$

The maximum acceptable page-fault rate for an effective access time of no more than 200 nanoseconds is approximately 6.1×10^{-6}.

Correct Answer: A

140. Which of the following is/are correct? (Multiple Select Answers)

A. With a single linear page table (and no other support), fetching and executing an instruction that performs an add of a constant value to a register will involve exactly two memory references.

B. Paging approaches suffer from external fragmentation, which grows as the size of a page grows.

C. A disadvantage of the SCAN and C-SCAN scheduling algorithms are that they ignore the influence of rotation time on positioning cost.

D. A TLB miss is usually faster to handle than a page miss.

Solution:

A. True; single linear page table implies one extra lookup per address translation; fetching the instruction requires one address translation.

B. False; paging has fixed-sized pages and thus suffers from internal fragmentation.

C. True; SCAN and C-SCAN just schedule based on the track (or cylinder) number.

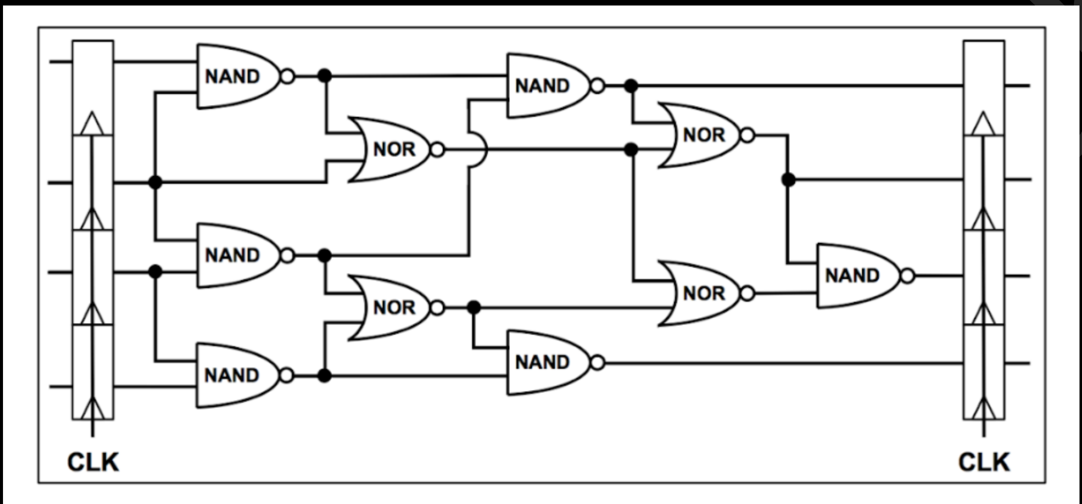
D. True; missing in the TLB just requires accessing RAM to walk the page tables; handling a page miss requires fetching a page from disk (milliseconds).

Correct Answer: A; C; D

7 Digital Logic (20 Questions)

141. For the following circuit, the timing characteristics of the components are summarized below.

- Flip-flop: clock-to-Q maximum delay (propagation delay) t_{pcq} = 30ps, clock-to-Q minimum delay (contamination delay) t_{ccq} = 20ps, setup time t_{setup} = 20ps, hold time t_{hold} = 20ps.
- Logic gate (each NAND, NOR): propagation delay t_{pd} = 25ps, contamination delay t_{cd} = 20ps.



Suppose that there is no clock skew. What is the maximum clock frequency of this circuit?

A. 7.67 GHz

B. 8.67 GHz

C. 6.67 GHz

D. 6.76 GHz

Solution:

$$t_{pcq} + t_{pd(MAX)} + t_{setup} \leq T_c$$
$$30ps + 4 \times 25ps + 20ps \leq T_c$$
$$150ps \leq T_c$$
$$f \leq \frac{1s}{T_c} = \frac{1s}{150ps} \approx 6.67\text{GHz}$$

The maximum operating frequency is about 6.67 GHz.

Correct Answer: C

142. A system has one output F and four inputs where the first two inputs A and B represent one 2-bit unsigned binary number, and the second two inputs C and D represent another 2-bit unsigned binary number. F is to be 1 if and only if the sum of the two numbers is odd.

The minimal POS form of F :

A. Has 2 sum terms

B. Has 3 sum terms

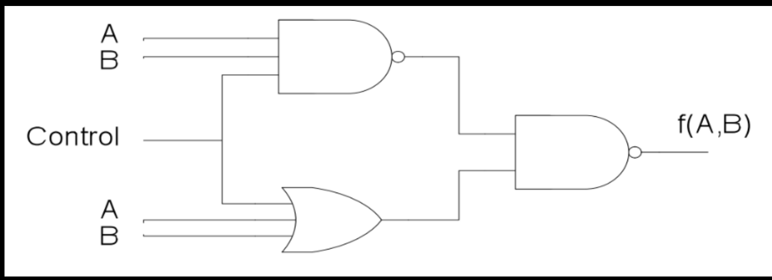
C. Has 4 sum terms

D. None of the above

Solution:

Correct Answer: A

143. What statement is correct for $f(A,B)$ in the following circuit?



A. $f(A,B) = A + B$ when Control $= 0$

B. $f(A,B) = \overline{\overline{A} \cdot \overline{B} \cdot (A+B)}$ when Control $= 1$

C. $f(A,B) = \overline{A} + \overline{B}$ when Control $= 1$

D. $f(A,B) = \overline{A} \cdot \overline{B}$ when Control $= 0$

Solution:

Correct Answer: D

144. What are the essential prime implicants of $f(A,B,C,D) = \Pi(0,1,6,9,11,12,14)$

A. $A'B'C, A'CD, B'CD'$

B. A'BC', AB'D', BD

C. A'BC', A'CD, BD

D. None of the above

Solution:

Correct Answer: B

145. Which of the following is/are correct? (Multiple Select Answers)

A. An n-bit number in bias notation and an n-bit two's complement number always represent the same range of numbers.

B. An n-bit two's complement number can represent exactly twice the number of unique values as an n-bit unsigned number.

C. For any number representation system, you must have at least $2n + 1$ bits to represent the value 2^{2n}.

D. The binary equivalent of the unsigned number $(23.75)_{10}$ is: $(10111.11)_2$

Solution:

A. Suppose the bias is 0, then the range of bias notation is the same as the range of unsigned.

B. Two's complement represents the same number of unique values as unsigned.

C. Using bias notation, you can represent this number with 1 bit and a bias of 2^{2n}.

Correct Answer: D

146. Consider a read-only memory (ROM) that performs binary to 3-digit binary-coded decimal (BCD) conversion as follows. The binary number is applied on the 10 input address lines, $\mathbf{A_9A_8A_7A_6A_5A_4A_3A_2A_1A_0}$, and the equivalent 3-digit BCD number is read on the 12 output data lines, $D_{11}D_{10}D_9D_8D_7D_6D_5D_4D_3D_2D_1D_0$, after the memory access time (note, the most significant BCD digit will appear on the four most significant bits on the output and the least significant BCD digit will appear on the four lest significant bits on the output). Consequently, when applying the input address $\mathbf{A_9A_8A_7A_6A_5A_4A_3A_2A_1A_0} = 1100000000$, after the memory access time the output data will become:

A. $D_{11}D_{10}D_9D_8D_7D_6D_5D_4D_3D_2D_1D_0 = 000100010010$

B. $D_{11}D_{10}D_9D_8D_7D_6D_5D_4D_3D_2D_1D_0 = 001001010110$

C. $D_{11}D_{10}D_9D_8D_7D_6D_5D_4D_3D_2D_1D_0 = 011101101000$

D. None of the above

Solution:

Correct Answer: C

147. Consider a comparator circuit between two unsigned 2-bit numbers. There is one output F and four inputs (i.e., two bits for each number, $\mathbf{A}_1\mathbf{A}_0$ and B_1B_0). F will be activated to a 1 when $\mathbf{A}_1\mathbf{A}_0$ is greater than or equal to B_1B_0 and it will be 0 otherwise. Then, function F is:

A. $\mathbf{F}(\mathbf{A}_1, \mathbf{A}_0, \mathbf{B}_1, \mathbf{B}_0) = \prod \mathbf{M}(1, 2, 3, 6, 7, 11)$

B. $\mathbf{F}(\mathbf{A}_1, \mathbf{A}_0, \mathbf{B}_1, \mathbf{B}_0) = \prod \mathbf{M}(1, 2, 3, 6, 7, 11, 15)$

C. $\mathbf{F}(\mathbf{A}_1, \mathbf{A}_0, \mathbf{B}_1, \mathbf{B}_0) = \prod \mathbf{M}(0, 1, 2, 3, 5, 6, 7, 10, 11, 15)$

D. None of the above

Solution:

Correct Answer: A

148. A flip-flop has a 3-ns delay from the time the clock edge occurs to the time the output is complemented. What is the maximum frequency that a 10-bit binary ripple counter that uses this flipflop can operate with reliability?

A. 10 MHz

B. 20 MHz

C. 25 MHz

D. 33 MHz

Solution:

Correct Answer: D

149. The single precision IEEE binary floating point representation of a number is

$$01000000011000000000000000000000.$$

What is the decimal value of this number?

A. 1.5

B. 3.5

C. 2.22

D. None of the above

Solution:

Correct Answer: B

150. Using a synchronous 3-bit up counter with parallel load capability ($\mathbf{Q_2}$ and $\mathbf{D_2}$ are the most significant bits) the circuit shown in Figure generates the following periodic sequence:

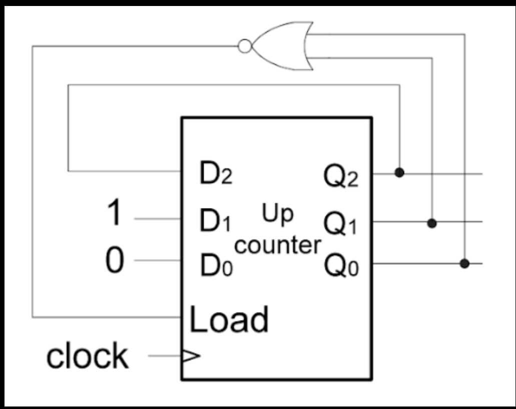


Figure: Circuit for this question (it is assumed that after power up the state is 000).

A. $0, 2, 3, 4, 6, 7$

B. $0, 2, 3, 4, 6$

C. $0, 2, 3, 4$

D. $0, 2, 4$

Solution:

Correct Answer: A

151. Which of the following equations is correct? (Multiple Select Answer)

A. $x \oplus y = (x \uparrow y') \uparrow (x' \uparrow y)$

B. $x \oplus y = (x \uparrow (x \uparrow y)) \uparrow ((x \uparrow y) \uparrow y)$

C. $x \oplus y = (x \downarrow y) \downarrow (x' \downarrow y')$

D. $x \oplus y = (x \downarrow y') \uparrow (x' \downarrow y)$

Solution:

Correct Answer: A; B; C

152.

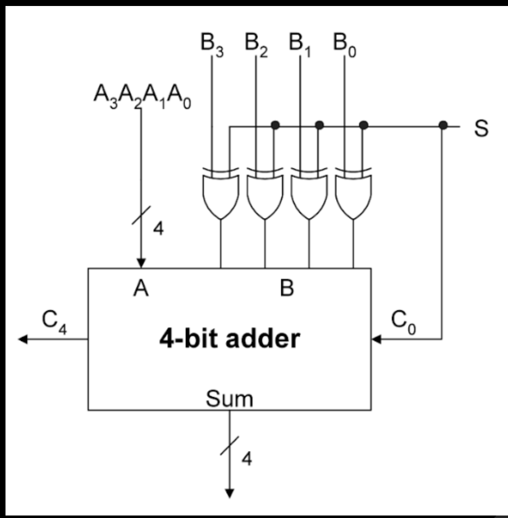


Figure: Circuit for this question. Note, c_0 is carry in and c_4 is carry out for the 4-bit adder.

Reminder: This adder and subtractor unit operates on 2's complement numbers and the S input signal determines whether an addition or subtraction will occur.

In the Figure , if $A_3A_2A_1A_0 = 0111, B_3B_2B_1B_0 = 1000$ and $S = 0$ then the output is:

A. Sum = 0000 and $C_4 = 0$

B. Sum = 0000 and $C_4 = 1$

C. Sum = 1111 and $C_4 = 0$

D. None of the above

Solution:

Correct Answer: C

153. The function $f = X_1 \oplus X_2 \oplus X_3 \ldots \oplus X_n$ defined over n variable has how many minterms:

A. $n/2$

B. $n^2/2$

C. 2^{n-1}

D. None of the above

Solution:

Correct Answer: C

154. How many 4-to-16 decoders are needed to implement an 8 -to-256 decoder? Note, all the decoders have an enable input and all the inputs and outputs are un-inverted.

A. 2

B. 8

C. 9

D. 17

Solution:

Correct Answer: D

155. The programmable array logic (PAL) implementation of function $\mathbf{G}(\mathbf{A}, \mathbf{B}, \mathbf{C})$ shown in the Figure is:

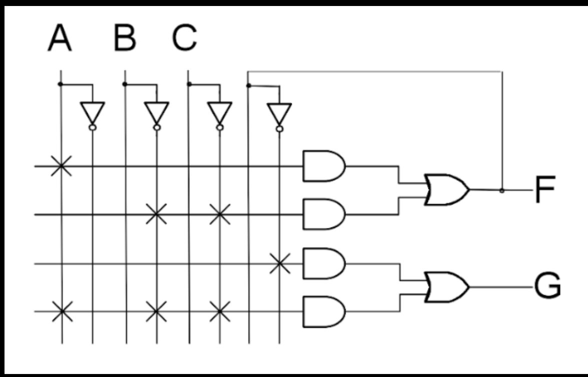


Figure: Circuit for this question. An X indicates that the wires are connected, i.e., the vertical signal is an input to the corresponding horizontal AND gate.

A. $\mathbf{G}(\mathbf{A}, \mathbf{B}, \mathbf{C}) = \Sigma\mathbf{m}(1,2,3,4,5,6)$

B. $\mathbf{G}(\mathbf{A}, \mathbf{B}, \mathbf{C}) = \Sigma\mathbf{m}(1,2,3,4,5)$

C. $\mathbf{G}(\mathbf{A}, \mathbf{B}, \mathbf{C}) = \Sigma\mathbf{m}(1,2,3,4)$

D. $\mathbf{G}(\mathbf{A}, \mathbf{B}, \mathbf{C}) = \Sigma\mathbf{m}(1,2,3)$

Solution:

Correct Answer: C

156. A four-variable logic function $F(A, B, C, D)$ equals to 1 if the number of input signals equal to 1 is greater than or equal to the number of input signals equal to 0. Function $\mathbf{F}(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathrm{D})$ can be written as:

A. $\mathbf{F}(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}) = \Sigma\mathbf{m}(3,5,6,7,11,12,13,14,15)$

B. $\mathbf{F(A,\ B,\ C,\ D)} = \Sigma\mathbf{m}(3,5,6,7,13,14,15)$

C. $\mathbf{F}(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}) = \prod\mathbf{M}(0,1,2,4,8,10)$

D. $\mathbf{F}(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}) = \Pi\mathbf{M}(0,1,2,4,8)$

Solution:

Correct Answer: D

157. Consider a combinational logic system that determines if a 4-bit binary quantity A, B, C, D in the range of 0000(0) through 1100 (12 in base 10) is divisible by the decimal numbers six. That is, the function is true if the input can be divided by six with no remainder. Assume that the binary patterns 1101 (13) through 1111 (15) are "don't cares." Tread the values as unsigned. Minimized Sum of Products form of this function is:

A. $A'B'C'D' + A'BCD'$

B. $A'B'C'D' + BCD'$

C. $AB + BCD' + A'B'C'D'$

D. None of the above

Solution:

Correct Answer: C

158. The periodic sequence $\mathbf{Q}_1\mathbf{Q}_0$ generated by the circuit shown in the Figure is:

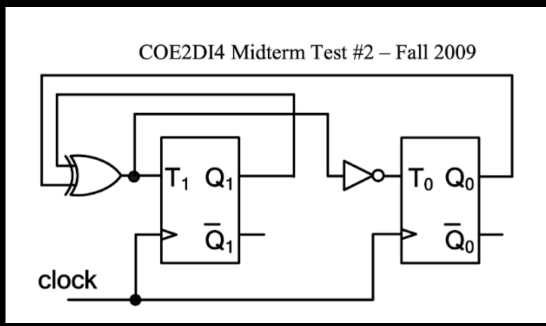


Figure: Circuit for this question (it is assumed that after power up the state is **00**).

A. $00, 01, 10, 11$

B. $00, 01, 11, 10$

C. $00, 10, 11, 01$

D. $00, 10, 01, 11$

Solution:

Correct Answer: B

159. The programmable array logic (PAL) implementation of function F shown in the Figure is:

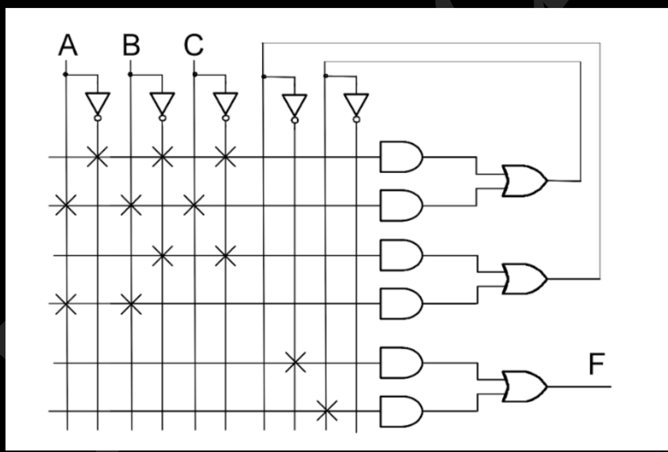


Figure: Circuit for this question. An X indicates that the wires are connected, i.e., the vertical signal is an input to the corresponding horizontal AND gate.

A. $F(A, B, C) = \Sigma\mathbf{m}(0, 1, 2, 3, 5, 7)$

B. $F(A, B, C) = \Sigma\mathbf{m}(0, 1, 2, 5, 7)$

C. $F(A, B, C) = \Sigma\mathbf{m}(0, 1, 2, 7)$

D. $F(A, B, C) = \Sigma\mathbf{m}(0, 1, 7)$

Solution:

Correct Answer: A

160.

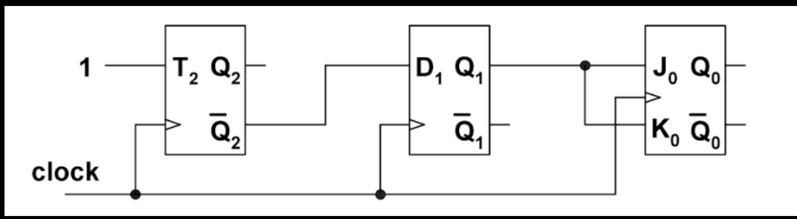


Figure: Circuit for this question (it is assumed that after power up the state is 000).

The periodic sequence $\mathbf{Q_2Q_1Q_0}$ generated by the circuit shown in the Figure is:

A. $000, 110, 001, 111$

B. $000, 101, 010, 111$

C. $000, 100, 001, 111$

D. None of the above

Solution:

Correct Answer: A

8 Computer Organization and Architecture (20 Questions)

161. Which of the following types of memory access methods, for which can you accurately predict the amount of time it will take to receive the data after requesting it? (Multiple Select Answers)

A. Sequential

B. Direct

C. Random

D. Associative

Solution:

The access time for sequential depends on how far away the data is from the current head position (e.g., tape). The access time for direct depends both on how far away the head is from the track and how far away the head is from the sector (e.g., hard drive). Both of these are not predictable.

Random access uses a decoder to instantly access the location regardless of the current state of the memory (e.g., SRAM). Associative uses a standard "compare circuit" which uniquely and quickly identifies the location of the data (e.g., cache). Both of these are consistent and predictable.

Correct Answer: C; D

162. Assume a memory access to main memory on a cache "miss" takes 30 ns and a memory access to the cache on a cache "hit" takes 3 ns . If 80% of the processor's memory requests result in a cache "hit", what is the average memory access time?

A. 8.4 nS

B. 33 nS

C. 24.6 nS

D. 27.0 nS

Solution:

Since we know that not all of the memory accesses are going to the cache, the average access time must be greater than the cache access time of 3 ns . This eliminates answers 'e' and ' f '. Similarly, not all accesses are going to the main memory, so the average must be less than 30 ns eliminating ' b ' as an answer. For the exact answer, you need to see that 80% of the time, the access will be 3 ns while the rest of the time (20%) the access time will be 30 ns . You should be able to see then that the average is going to be closer to 3 ns than 30 ns meaning the answer is a. You could also see this with the following equation.

$$\begin{aligned}(0.8 \times 3 \text{ ns}) + (0.2 \times 30 \text{ ns}) &= 2.4 \text{ ns} + 6 \text{ ns} \\ &= 8.4 \text{ ns}\end{aligned}$$

Correct Answer: A

163. An ISA includes three classes of machine instructions: type I taking 1 clock cycle, type II taking 2 clock cycles and type III taking 4 clock cycles. The execution of a particular program, running on a machine with a clock rate of 2.1 GHz, will require 25% type I instructions, 40% type II instructions, and 35% type III instructions.

What is the average CPI for this program? (Numerical Answer Type)

Solution:

$$\text{Average CPI } = .25 * 1 + .40 * 2 + .35 * 4 = .25 + .8 + 1.4 = 2.45$$

Correct Answer: 2.45

164. The clock rate for Machine A is 2.4 GHz, and the clock rate for machine B is 3.0 GHz. For a particular program, the average CPI on machine A is 1.2. For the same program, the average CPI on machine B is 2.0.

From those facts, what can be established about the relative speeds of the two machines, with respect to this program?

A. Machine A is 4/3 times as fast as Machine B.

B. Machine A is 3/2 times as fast as Machine B.

C. Machine A is 2/3 times as fast as Machine B.

D. Machine A is 1/2 times as fast as Machine B.

Solution:

One approach is to calculate the execution time for a typical program, having, say, I machinelevel instructions:

$$\begin{aligned}\text{Time }_A &= \text{ InstructionCount } \times CPI_A \times \text{ CycleTime }_A \\ &= I \times 1.2 \times \frac{1}{2.4 \times 10^9} = \frac{1}{2} I \times 10^{-9} \\ \text{Time }_B &= \text{ InstructionCount } \times \text{CPI}_B \times \text{ CycleTime }_B \\ &= I \times 2.0 \times \frac{1}{3.0 \times 10^9} = \frac{2}{3} I \times 10^{-9}\end{aligned}$$

So, a typical program will execute (slightly) more quickly on Machine A. More precisely, the relative performance would be:

$$\frac{\text{Tim}_B}{\text{Tim}_A} = \frac{\frac{2}{3} I \times 10^{-9}}{\frac{1}{2} I \times 10^{-9}} = \frac{4}{3}$$

So, Machine A is 4/3 times as fast as Machine B.

Correct Answer: A

165. Consider two different machines. The first has a single-cycle datapath (i.e., a single-stage, non-pipelined machine) with a cycle time of 15 ns. The second is a pipelined machine with 5 pipeline stages and a cycle time of 3ns.

What is the speedup of the pipelined machine versus the single cycle machine if the pipeline stalls 1 cycle for 25% of the instructions? (Numerical Answer Type)

Solution:

New $CPI = 1 + .25 \times 1 = 1.25$

Since the number of instructions is the same, the speedup is $\frac{CPI_{old} \times \text{cycle time}_{\text{old}}}{CPI_{\text{new}} \times \text{cycle time}_{\text{new}}} = \frac{1 \times 15\text{ ns}}{1.25 \times 3\text{ ns}} = 4$.

Correct Answer: 4

166. Consider two different 5-stage pipeline machines (IF ID EX MEM WB). The first machine resolves branches in the ID stage, uses one branch delay slot, and can fill 80% of the delay slots with useful instructions. The second machine resolves branches in the EX stage (i.e., it determines whether the branch is taken and the target address of a taken branch in the EX stage) and uses a predict-not-taken scheme. Assume that the cycle times of the machines are identical. Given that 35% of the instructions are branches, 25% of branches are taken, and that stalls are due to branches alone, which machine is faster?

A. The first machine is faster

B. The second machine is faster

C. Both machines are equally faster

D. None of the above

Solution:

For the first machine, a cycle is wasted every time a delay slot can't be filled i.e. for 20% of the branches. Thus, $CPI = 1 + 0.35 \times .20 \times 1 = 1.07$.

For the second machine, two cycles are wasted due to the unknown target address for 25% of the branches. Thus, $CPI = 1 + .35 \times .25 \times 2 = 1.175$.

Therefore, the first machine is faster.

Correct Answer: A

167. Virtual memory problem: Assume a computer has on-chip and off-chip caches, main memory, and virtual memory. Assume the following hit rates and access times: on-chip cache 95%, 1 ns, off-chip cache 99%, 10 ns, main memory: X%, 50 ns, virtual memory: 100%, 2,500,000 ns. Notice that the on-chip access time is 1 ns. We do now want our effective access time to increase much beyond 1 ns. Assume that an acceptance effective access time is 1.6 ns. What should X be (the percentage of page faults) to ensure that EAT is no worse than 1.6 ns?

A. 94.994%

B. 99.994%

C. 92.994%

D. 90.994%

Solution:

$\text{EAT} = 1\text{ ns} + .05 * (10\text{ ns} + .01 * (50\text{ ns} + (1 - X) * 2{,}500{,}000\text{ ns}))$.

Since we want EAT to be no more than 1.6 ns, we solve for X with $1.6\text{ ns} = 1\text{ ns} + .05*(10\text{ ns} + .01 * (50\text{ ns} + (1-X)*2{,}500{,}000\text{ ns}))$. $X = 1-(((((1.25\text{ ns}-1\text{ ns})/.05)-10\text{ ns})/.01)-50\text{ ns})/2{,}500{,}000)$. $X = 0.99994 = 99.994\%$.

Our miss rate for virtual memory must be no worse than .006%.

Correct Answer: B

168. Alice, after realizing that she's lost in the class, panics and decides to buy a laser printer in order to print out all the course notes. Suppose that she bought a laser printer that produces up to 45 pages per minute, where each page consists of 5000 characters. The manual for the laser printer states that the system uses interrupt-driven I/O by raising an interrupt for every character.

If each interrupt takes 50 microseconds to process, how much CPU time will be spent processing interrupts (in %)? (Numerical Answer Type)

Solution:

Percentage of CPU time spent: 18.75%

Assume that you print 45 pages. Then, the ratio to find is: $\frac{\text{time}_{\text{ interrupts}}}{\text{time}_{\text{ print 45 pages}}}$.

$$\text{time}_{\text{ interrupts}} = 45\text{ pages} * \frac{5000\text{ characters}}{1\text{ page}} * \frac{50\text{ microsece}}{1\text{ character}} * \frac{1\text{ second}}{10^6\text{ microsec}} * \frac{1\text{ minute}}{60\text{ seconds}} = 0.1875\text{ minutes}.$$

It is obvious that $\text{time}_{\text{ print 45 pages}} = 1$ minute.

Therefore, the percentage of CPU time spent is 18.75%

Correct Answer: 18.75

169. Consider the following MIPS assembly language code:

I1: ORI $s0, $0, 5

I2: ADDI $s1, $0, 10

I3: ADD $s1, $s0, $s1

I4: LW $s0, -4($s1)

I5: ADD $s0, $s0, $s0

I6: SW $s0, -4($s1)

Complete the following table showing the timing of the above code on the 5-stage pipeline given (IF, ID, EX, MEM, WB) supporting forwarding and pipeline stall. Draw an arrow showing forwarding between the stage that provides the data and the stage that receives the data. Show all stall cycles (draw an *X* in the box to represent a stall cycle). Determine the number of clock cycles to execute this code.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
I1: ORI	IF	ID	EX	-	WB										
I2: ADDI															
I3: ADD															
I4: LW															
I5: ADD															
I6: SW															

Total number of clock cycles to execute the above code = (Numerical Answer Type)

Solution:

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
I1: ORI	**IF**	**ID**	**EX**	-	**WB**										
I2: ADDI		**IF**	**ID**	**EX**	-	**WB**									
I3: ADD			**IF**	**ID**	**EX**	-	**WB**								
I4: LW				**IF**	**ID**	**EX**	**M**	**WB**							
I5: ADD					**IF**	**X**	**ID**	**EX**	-	**WB**					
I6: SW							**IF**	**ID**	**EX**	**M**	-				

Total number of clock cycles to execute the above code = **10**

Correct Answer: 10

170. Gollum is a capable programmer who works for Google. He wrote a program that takes one web page as the input and outputs "yes" if the web page contains "my precious" in its text. He wants to run this program to process one million web pages. He can choose between two machines. These two machines are identical except that one machine uses an out-of-order superscalar CPU and the other uses a 2-way SMT CPU.

The OoO superscalar machine can only run one thread and the 2-way SMT machine can run two threads simultaneously. However, due to the added complexity, the single thread performance for the 2-way SMT machine is worse than that of the OoO superscalar machine. On average, the superscalar machine is 1.2 times faster than the SMT machine when running one thread. However, with two threads, the SMT machine is 1.5 times faster than the superscalar machine.

Which machine should Gollum use in order to process the one million web pages he has as quickly as possible?

A. The OoO Superscalar machine because it has better single-thread performance.

B. The 2-way SMT machine because it can utilize two threads to achieve 1.5 times the performance of the OoO machine.

C. The OoO Superscalar machine because it can run one thread faster than both threads on the SMT machine.

D. Both machines will perform equally well; the choice doesn't matter.

Solution:

Suppose that P instructions/cycle is the performance of the SMT when it runs 1 thread (the unit could be different, but instructions/cycle is a good one to use here). Then the superscalar's performance is $1.2 \times P$.

If we run 2 threads, either in sequence or at the same time, at best the superscalar's performance will be the same, so the best-case performance is still $1.2 \times P$ (often times worse though). However, the SMT can run 2 threads at the same time and will improve the performance when it does that. The performance will be $1.5 \times 1.2 \times P$.

With one million web pages to process, Gollum has plenty of threads to run. Thus, it's better if Gollum picks the SMT and runs two threads at a time to process one million web pages.

Correct Answer: B

171. Which kind of dependence can cause data hazards in a single-core, pipelined, in-order processor? (Mark all the appropriate choices)

A. read-after-write dependence

B. write-after-read dependence

C. write-after-write dependence

D. read-after-read dependence

Solution:

Correct Answer: A

172. Which of the following may lead to a reduction in cache capacity misses? (Mark all the appropriate choices)

A. Switching from write-back to write-through.

B. Increasing cache size.

C. Increasing associativity (for a given capacity).

D. Decreasing block size (for a given capacity).

Solution:

Correct Answer: B; D

173. An L2 cache with 64-byte lines, 4 ways, and 4096 sets, for a machine with 32-bit virtual addresses and 40-bit physical addresses.

What will be the total capacity of the cache?

A. 16 KB

B. 256 KB

C. 1 MB

D. 32 MB

Solution:

Correct Answer: C

174. Which of the following is the best justification for using the middle bits of an address as the set index into a cache rather than the most significant bits?

A. Indexing with the most significant bits would necessitate a smaller cache than is possible with middle-bit indexing, resulting in generally worse cache performance.

B. It is impossible to design a system that uses the most significant bits of an address as the set index.

C. The process of determining whether a cache access will result in a hit or a miss is faster using middle-bit indexing.

D. A program with good spatial locality is likely to make more efficient use of the cache with middle-bit indexing than with high-bit indexing.

Solution:

Correct Answer: D

175. You are developing a new enhancement that provides a 2.5 x speedup to certain kinds of instructions. What percentage of a program, as measured by its original execution time, must consist of these instructions if you want to gain an overall speedup of 10%? (Mark all the appropriate choices)

A. 15.2%

B. 0.125

C. 12.5%

D. 0.152

Solution:

Let f be the required fraction.

$$\frac{T_{\text{old}}}{T_{\text{new}}} = 1.1$$

$$\frac{T_{\text{old}}}{\left((1-f)+\frac{f}{2.5}\right)\times T_{\text{old}}} = 1.1$$

$$\frac{1}{1-\frac{3f}{5}} = 1.1$$

$$\frac{3f}{5} = 1 - \frac{1}{1.1}$$

$$f = 0.152 \text{ or } 15.2\%$$

Correct Answer: A; D

176. Two enhancements are proposed: one that can enhance 40% of execution time with a speedup of 1.5, and another that can enhance 25% of execution time with some greater speedup value. Only one of these two can be implemented. How much of a speedup is necessary in the second enhancement to give a better enhancement than the first?

A. 2.14

B. 4.12

C. 1.42

D. 2.41

Solution:

$$T_{new1} \geq T_{new2}$$
$$0.6 + \frac{0.4}{1.5} \geq 0.75 + \frac{0.25}{s}$$
$$\frac{7}{60} \geq \frac{0.25}{s}$$
$$s \geq \frac{15}{7} \text{ or } 2.14$$

Correct Answer: A

177. Several researchers have suggested that adding a register-memory addressing mode to a load-store machine might be useful. The idea is to replace sequences of

```
ld x1, 0(x8)
add x2, x2, x1
```

with

```
add x2, 0(x8)
```

Assume that the new instruction will cause the clock cycle time to increase by 5% and will not affect the CPI. Also, assume loads constitute 25.1% of all instructions. What percentage of the loads must be eliminated for the machine with the new instruction to have at least the same performance? (Mark all the appropriate choices)

A. 19%

B. 0.91

C. 0.19

D. 91%

Solution:

Let L be the fraction of loads that are eliminated. This means that $0.251 \times L$ of all instructions are eliminated.

$$\text{CPU time}_{\text{old}} = \#\text{ of instructions } \times \text{ CPI } \times \text{ cycle time}$$
$$\text{CPU time}_{\text{new}} = ((1 - 0.251 \times L) \times \#\text{ of instructions }) \times \text{CPI} \times ((1 + .05) \times \text{ cycle time })$$
$$\text{CPU time}_{\text{new}} \leq \text{CPU time}_{\text{old}}$$
$$(1 - 0.251 \times L) \times 1 \times 1.05 \leq 1$$
$$0.251 \times L \geq 1 - \frac{1}{1.05}$$
$$L \geq 0.19 \text{ or } 19\%$$

Correct Answer: A; C

178. Assume:

- A processor has a direct mapped cache
- Data words are 8 bits long (i.e. 1 byte)
- Data addresses are to the word
- A physical address is 20 bits long
- The tag is 11 bits
- Each block holds 16 bytes of data

How many blocks are in this cache?

A. 16

B. 15

C. 32

D. 31

Solution:

Given that the physical address is 20 bits long, and the tag is 11 bits, there are 9 bits left over for the index and offset

We can determine the number of bits of offset as the problem states that:

- Data is word addressable and words are 8 bits long
- Each block holds 16 bytes

As there are 8 bits / byte, each block holds 16 words, thus 4 bits of offset are needed.

This means that there are 5 bits left for the index. Thus, there are 2^5 or 32 blocks in the cache.

Correct Answer: C

179. Consider a 16-way set-associative cache

- Data words are 64 bits long
- Words are addressed to the half-word
- The cache holds 2 Mbytes of data
- Each block holds 16 data words
- Physical addresses are 64 bits long

How many bits of tag, index, and offset are needed to support references to this cache?

A. Tag: 48 bits; Index: 10 bits; Offset: 5 bits

B. Tag: 49 bits; Index: 8 bits; Offset: 5 bits

C. Tag: 49 bits; Index: 10 bits; Offset: 4 bits

D. Tag: 49 bits; Index: 10 bits; Offset: 5 bits

Solution:

We can calculate the number of bits for the offset first:

- There are 16 data words per block which implies that at least 4 bits are needed
- Because data is addressable to the 1/2 word, an additional bit of offset is needed
- Thus, the offset is **5** bits

To calculate the index, we need to use the information given regarding the total capacity of the cache:

- 2 MB is equal to 2^{21} total bytes.
- We can use this information to determine the total number of blocks in the cache...
    - 2^{21} bytes **x** (1 block / 16 words) **x** (1 word / 64 bits) **x**(8 bits /1 byte) = 2^{14} blocks
- Now, there are 16 (or 2^4) blocks / set
    - Therefore there are 2^{14} blocks **x** (1 set /2^4 blocks) = 2^{10} or 1024 sets
- Thus, 10 bits of index are needed

Finally, the remaining bits form the tag:

- $64 - 5 - 10 = 49$
- Thus, there are 49 bits of tag

To summarize: Tag: 49 bits; Index: 10 bits; Offset: 5 bits

Correct Answer: D

180. Which of the following changes to a cache design would decrease the number of set index bits it has? (Mark all the appropriate choices)

A. doubling the associativity without changing the block size or total cache size

B. doubling the block size without changing the associatity or total cache size

C. changing from a random replacement policy to an LRU replacement policy

D. doubling the total cache size and the associativity without changing the block size

Solution:

Correct Answer: A; B

9 Theory of Computation (20 Questions)

181. Let be given a language L. Which of the following languages is not necessarily equal to the others?

A. LL^*

B. $LL^* \cup \{\lambda\}$

C. L^*L^*

D. L^*

Solution:

Correct Answer: A

182. Let be given a language L. Which of the following languages is not necessarily equal to the others?

A. L

B. $\left(L^R\right)^R$

C. $L \cup \left(L^R\right)^R$

D. $\left(\left(L^R\right)^R\right)^R$

Solution:

Correct Answer: D

183. Which of the following regular expressions describes a language different from the others?

A. $(a \cup b)^*$

B. $(a^* b^*)^*$

C. $(b^* a^*)^*$

D. $(a^* \cup b^*)$

Solution:

Note that

$$\mathcal{L}((a \cup b)^*) = \mathcal{L}((a^* b^*)^*) = \mathcal{L}((b^* a^*)^*) = \{a, b\}^*$$

and that

$$\mathcal{L}(a^* \cup b^*) = \{a\}^* \cup \{b\}^*$$

So in particular $ab \notin \{a\}^* \cup \{b\}^*$, but $ab \in \{a, b\}^*$.

Correct Answer: D

184. Which of the following regular expressions describes a language different from the others?

A. $a^* b^*$

B. $(\lambda \cup aa)^* (\lambda \cup bb)^*$

C. $(a \cup aa)^* (b \cup bb)^*$

D. $(\lambda \cup a)^* (\lambda \cup b)^*$

Solution:

Note that

$$\mathcal{L}(a^* b^*) = \mathcal{L}((a \cup aa)^* (b \cup bb)^*) = \mathcal{L}((\lambda \cup a)^* (\lambda \cup b)^*) = \{a^n b^m \mid n, m \in \mathbb{N}\}$$

and that

$$\mathcal{L}((\lambda \cup aa)^* (\lambda \cup bb)^*) = \left\{a^{2n} b^{2m} \mid n, m \in \mathbb{N}\right\}$$

So in particular $ab \notin \left\{a^{2n} b^{2m} \mid n, m \in \mathbb{N}\right\}$, but $ab \in \{a^n b^m \mid n, m \in \mathbb{N}\}$.

Correct Answer: B

185. Consider the following context-free grammar for a fragment of English:

$$\begin{aligned} S &\to N \text{ walks } \mid N \text{ loves } N \mid S \text{ and } S \\ N &\to \text{ the } AM \\ A &\to \text{ tall } \mid \text{ small } \mid \lambda \\ M &\to \text{ man } \mid \text{ woman} \end{aligned}$$

Which of the following statements about the language produced by this grammar is not true?

A. The shortest sentences in the language have three words.

B. The sentences in this language can be arbitrarily long.

C. This language contains an infinite sentence.

D. For each $n \geq 3$ this language contains a sentence of exactly n words.

Solution:

Correct Answer: C

186. Consider the following context-free grammar for a fragment of a programming language, with alphabet $\Sigma = \{: , = , ; , 0 , 1 , +, *, (,), \mathrm{x}, \mathrm{y}\}$:

$$S \rightarrow V := E \mid S ; S$$
$$E \rightarrow 0 \mid 1 \mid V \mid E + E \mid E * E \mid (E)$$
$$V \rightarrow \mathrm{x} \mid \mathrm{y}$$

Which of the following statements about the language produced by this grammar is not true?

A. The shortest programs in this language have four symbols.

B. The programs in this language can be arbitrarily long.

C. This language contains an infinite program.

D. This language does not contain a program of five symbols.

Solution:

Correct Answer: C

187. Which of the following requirements does not necessarily hold for an invariant of a context-free grammar G?

A. An invariant is a predicate on words from $(\Sigma \cup V)^*$.

B. An invariant holds for the word $S \in (\Sigma \cup V)^*$.

C. If an invariant holds for a word, and one symbol in the word is replaced according to a rule from the grammar, the invariant still holds.

D. The invariant does not hold for the word of which we want to show that it is not in the language $\mathcal{L}(G)$.

Solution:

If we want to prove that some word is not in the language generated by a grammar, we do want to have that the invariant doesn't hold for this specific word. However, this is not part of the definition of an invariant.

The other three options are consequences of the definition of an invariant.

Correct Answer: D

188. Which of the following requirements does not necessarily hold for an invariant of a context-free grammar G?

A. The invariant holds for all words in $\mathcal{L}(G)$.

B. The invariant holds for all words $S, w_1, w_2, \ldots$ in any production $S \to w_1 \to w_2 \to \ldots$ of the language.

C. There is a word in $(\Sigma \cup V)^*$ for which the invariant holds.

D. $\mathcal{L}(G)$ consists of the words in Σ^* that satisfy the invariant.

Solution:

The most trivial example of an invariant is $P(w) :=$ true. It is clear that this is indeed an invariant. However, this invariant also holds for the words in Σ^* which are not in $\mathcal{L}(G)$.

The other three options are consequences of the definition of an invariant.

Correct Answer: D

189. What is the minimum number of states for a DFA that accepts the language

$$\mathcal{L}(\lambda \cup aba^*)$$

A. two

B. three

C. less than two

D. more than three

Solution:

Correct Answer: D

190. There exist DFAs with 2020 states that accept the word $a^{2021}b^{2021}$. Does such a DFA always accept a word $a^n b^{2021}$ as well, for some $n > 2021$?

A. Yes, because while processing the a's in $a^{2021}b^{2021}$, there has to be a state that occurs twice, which means there is a loop while processing the a's.

B. No, because it only accepts words of the form $a^n b^n$.

C. No, because the language $\{a^n b^n \mid n \in \mathbb{N}\}$ is not regular, so it is not accepted by a DFA.

D. You cannot know this, this is the case for some of these automata, but not for all.

Solution:

This is basically the proof of the pumping lemma for regular languages, which is not really discussed in this course, but the explanation given above should be convincing.

The second answer makes no sense because we don't know anything about the words being accepted, besides that $a^{2021}b^{2021}$ is accepted.

The third answer makes no sense because this claim is not about this language.

The fourth answer makes no sense because we do know this: the first answer is correct.

Correct Answer: A

191. Consider the NFA:

$$M := \langle \Sigma, Q, q_0, F, \delta \rangle$$

with:

$$\begin{aligned}
\Sigma &= \{a, b\} \\
Q &= \{q_0, q_1\} \\
F &= \{q_0\} \\
\delta(q_0, a) &= \{q_0, q_1\} \\
\delta(q_0, b) &= \varnothing \\
\delta(q_0, \lambda) &= \varnothing \\
\delta(q_1, a) &= \varnothing \\
\delta(q_1, b) &= \{q_1\} \\
\delta(q_1, \lambda) &= \{q_0\}
\end{aligned}$$

What is true?

A. $aab \in L(M)$ and $baa \in L(M)$

B. $aab \in L(M)$ and baa $\notin L(M)$

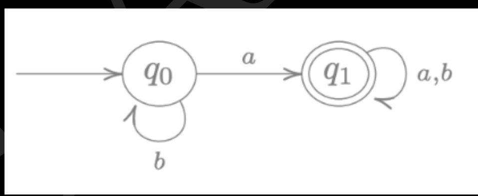
C. $aab \notin L(M)$ and $baa \in L(M)$

D. $aab \notin L(M)$ and $baa \notin L(M)$

Solution:

Correct Answer: B

192. Consider the deterministic finite automaton M_{11} :



Which of the following regular expressions does not describe the language of this automaton?

A. $b^*a(a \cup b)^*$

B. $(a \cup b)^*a(a \cup b)^*$

C. $(a^*b^*)^*\, ab^*$

D. All of the above describe the language of the automaton.

Solution:

Correct Answer: D

193. Does the following equality hold?

$$\mathcal{L}(a^*b^*) = \mathcal{L}((ab)^*)$$

A. Yes.

B. No, but $\mathcal{L}(a^*b^*) \subset \mathcal{L}((ab)^*)$.

C. No, but $\mathcal{L}((ab)^*) \subset \mathcal{L}(a^*b^*)$.

D. No, and neither language is a subset of the other.

Solution:

Note that $a \in \mathcal{L}(a^*b^*)$, but $a \notin \mathcal{L}((ab)^*)$. So $\mathcal{L}(a^*b^*) \subset \mathcal{L}((ab)^*)$ does not hold.

Note also that $abab \in \mathcal{L}((ab)^*)$, but $abab \notin \mathcal{L}(a^*b^*)$. So the statement $\mathcal{L}((ab)^*) \subset \mathcal{L}(a^*b^*)$ also does not hold.

Correct Answer: D

194. Consider the context-free grammar G_{11} :

$$S \to aA \mid \lambda$$
$$A \to bS$$

Someone wants to show that

$$aba \notin \mathcal{L}(G_{11})$$

using the property

$$P(w) := w \text{ has the same number of } a\text{'s and } b\text{'s}$$

in the hope that it is an invariant for this grammar. Does that work?

A. Yes, as the property holds for all elements in $\mathcal{L}(G_{11})$.

B. Yes, $S \to aA \to abS$, so in each production the number of a's and b's stays the same.

C. No, this is not an invariant, as the production steps $aAa \to abSa \to aba$ show.

D. No, this is not an invariant, as the production step $S \to aA$ shows.

Solution:

Note that $P(S)$ holds as S has zero a's and zero b's. However, from S we can produce aA in a single step. But $P(aA)$ does not hold as aA has one a and zero b's. So the invariance is broken by the rule $S \to aA$.

The third answer makes no sense because we should only look at single step productions and besides that, the property should hold before doing this step, but in this case, the property doesn't hold before any of the two production steps.

The first and the second answer make no sense because the answer is 'no'.

Correct Answer: D

195. Let $M = \langle \Sigma, Q, q_0, F, \delta \rangle$ be a non-deterministic finite automaton. Which statement is correct for each $q_i \in Q$ and $x \in \Sigma$?

A. $\delta(q_i, x) \in \Sigma$

B. $\delta(q_i, x) \subseteq \Sigma$

C. $\delta(q_i, x) \in Q$

D. $\delta(q_i, x) \subseteq Q$

Solution:

Recall that the transition function δ gives a set of states in an NFA. Hence it must be one of the options with the set of states Q and not one of the options with the alphabet Σ.

In addition, because the result of δ is a set, the relation with Q has to be about being a subset ($\subseteq$) and not about being an element ($\in$).

Correct Answer: D

196. Let be given an arbitrary deterministic finite automaton M. We want to make another deterministic finite automaton M', such that

$$\mathcal{L}(M') = \overline{\mathcal{L}(M)}$$

(so M' should recognize the complement of the language that M recognizes.) How can we do this?

A. M' is like M, but we change all the final states to a non-final state, and all non-final states to a final state.

B. M' is like M, but we reverse all transitions.

C. M' is like M, but we add a sink state.

D. M' is like M, but with the sink state removed.

Solution:

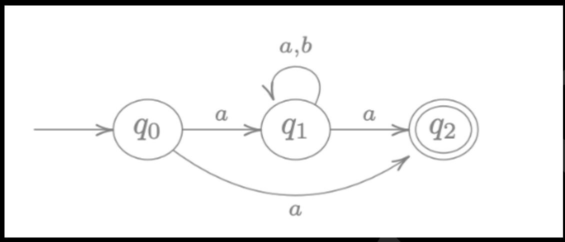
The language $\overline{\mathcal{L}(M)}$ is the complement of the language $\mathcal{L}(M)$. And for a word w it holds that $w \in \overline{\mathcal{L}(M)}$ if and only $w \notin \mathcal{L}(M)$.

However, in terms of a DFA M, it holds that $w \in \mathcal{L}(M)$ if and only if M ends in a final state after reading w.

So if we combine these two things we get that $w \in \overline{\mathcal{L}(M)}$ if and only if w ends in a non-final state in M. So if we create M' by swapping the final and non-final states, we get that $w \in \overline{\mathcal{L}(M)}$ if and only if w ends in a final state in M', which means that $\overline{\mathcal{L}(M)} = \mathcal{L}(M')$

Correct Answer: A

197. Let be given the following non-deterministic finite automaton M_{15}:



This recognizes the words that both start and end with the symbol a. How many states has a deterministic finite automaton that recognizes the same language as M_{15} and has a minimal number of states?

A. Less than three states.

B. Exactly three states.

C. Exactly four states.

D. More than four states.

Solution:

Correct Answer: C

198. Which of the following is/are correct? (Multiple Select Answers)

In all parts of this question, the alphabet Σ is $\{0,1\}$.

A. If L is a regular language and F is a finite language (i.e., a language with a finite number of words), then $L \cup F$ must be a regular language.

B. If L is a regular language, then $\left\{ww^R : w \in L\right\}$ must be a regular language. (Here, w^R denotes the reverse of string w.)

C. Regular expressions that do not contain the star operator can represent only finite languages.

D. Define $EVEN(w)$, for a finite string w, to be the string consisting of the symbols of w in even-numbered positions. For example, $EVEN(1011010) = 011$. If L is a regular language, then $\{EVEN(w) : w \in L\}$ must be regular.

Solution:

A. True; all finite languages are regular languages and regular languages are closed under union.

B. False; we can show this language in not regular using techniques similar to Example 1.40 on page 81 of Sipser. A common mistake is confusing the language above with LL^R.

C. True; the star operator in regular expressions is the equivalent of a loop in DFAs. If a deterministic finite automata with n states does not contain a loop, then at most it can recognize strings of length less than n. The set of such strings is finite.

D. True; given a DFA $M = (Q, \Sigma, \delta, q_0, F)$ that recognizes L, we can build an NFA $M' = (Q, \Sigma, \delta', q_0, F)$ that recognizes $L' = \{EVEN(w) : w \in L\}$. Intuitively, the transition function δ' is set so that from every state A there are outgoing transitions to any state that is two hops (i.e., reachable in M by reading two input characters) away from A, where the new transition symbol in M' is the second of the two symbols in M.

Correct Answer: A; C; D

199. Which of the following statements is incorrect?

S1: For every pair of regular expressions R and S, the languages denoted by $R(SR)^*$ and $(RS)^*R$ are the same.

S2: If L_1 and L_2 are languages such that L_2, L_1L_2, and L_2L_1 are all regular, then L_1 must be regular.

A. S1 only

B. S1 and S2 both

C. Neither S1 nor S2

D. S2 only

Solution:

S1: True; intuitively, one can see that minimally both expressions are R and that when the star operator is exercised they both expand to expressions of alternating Rs and Ss, where both expressions begin and end with R. It is possible to prove that they describe equivalent sets using induction.

S2: False; consider $L_1 = \left\{0^{2^i} : i > 0\right\}$ and $L_2 = \{0\}^*$.

Correct Answer: D

200. The following language L over the alphabet $\{a, b, c\}$ is:

$L = \{wcx : w, x \in \{a, b\}^*$ and the number of $a's$ in w is equal to the number of $b's$ in $x\}$.

A. Regular

B. Not regular

C. Context-free

D. None of the above

Solution:

Assume to the contrary that L is regular. Let p be the pumping length given by the pumping lemma. Let s be the string $a^p c b^p$. Because s is a member of L and s has length more than p, the pumping lemma guarantees that s can be split into three pieces, $s = xyz$, satisfying the three conditions of the lemma. We show that this is impossible.

Since, by the pumping lemma, we know that $|xy| \leq p$ and $|y| > 0$, then for this s, we see that y is made up of one or more a 's that all come before the ' c ' symbol. If we pump down, by letting $i = 0$, then the resulting string is $s' = xy^0z$, where the number of a 's before ' c ' is at least one less than the number of b 's after it. This new string s' is not in L. Thus, since s can not be pumped, L is not a regular language.

Correct Answer: B

10 Compiler Design (20 Questions)

201. Which of the following is/are correct? (Multiple Select Answers)

A. An $LR(k)$ grammar is a context-free grammar where the handle in a right sentential form can be identified with a lookahead of at most k input. We shall only consider $k = 0, 1$.

B. An $LR(0)$ parser can take shift-reduce decisions entirely on the basis of the states of $LR(0)$ automaton of the grammar. Consider the following grammar with the augmented start symbol and the production rule.

C. An $LR(0)$ item is called complete if the ' $\bullet$ ' is at the right end of the production, $A \rightarrow \alpha\bullet$. This indicates that the DFA has already 'seen' a handle and it is on the top of the stack.

D. A grammar G is of type $LR(0)$ if the DFA of its viable prefixes has the following properties:

- no state has both complete and incomplete items,
- no state has two complete items.

Solution:

Correct Answer: A;B;C;D

202. Which of the following statements is correct in the LR(0) parser?

I. A state with a unique complete item $A \rightarrow \alpha\bullet$, indicates a reduction of the handle α by the rule $A \rightarrow \alpha$.

II. A state with incomplete items indicates shift actions.

A. I only

B. II only

C. Both of them

D. None of them

Solution:

Correct Answer: C

203. Which of the following is/are correct? (Multiple Select Answers)

A. Number of states of the $LR(1)$ automaton are more than that of $LR(0)$ automaton.

B. Merging of two $LR(1)$ states with the same $LR(0)$ item cannot give rise to a new shift/reduce conflict.

C. Two states of an LALR parser cannot have the same set of $LR(0)$ items.

D. The number of states of an $LR(0)$ and an LALR(1) automaton are the same.

Solution:

Correct Answer: A;B;C;D

204. Consider the CFG G with $N = \{S, X\}, T = \{a, b, d\}$, start symbol S and productions

$$S \to Xb, X \to aXd \mid \varepsilon$$

Which of the following is correct?

A. First$(Xb) = \{a, b\}$

B. Follow$(X) = \{b, d\}$

C. The grammar is $LL(1)$

D. All of the above

Solution:

A. First$(Xb) = \{a, b\}$.

B. Follow$(X) = \{b, d\}$.

C. The grammar is LL(1) since First $(aXd) = \{a\}$ and Follow$(X) = \{b, d\}$ are disjoint.

Correct Answer: D

205. For the following problem, consider the following context-free grammar:

$$\begin{aligned}
&S \to A\$ \\
&A \to xAx \\
&A \to C \\
&B \to yBy \\
&B \to C \\
&C \to zBz \\
&C \to wAw \\
&C \to \lambda
\end{aligned}$$

What are the terminals and non-terminals of this grammar?

A. Terminals: $\{x, y, z, w\}$, non-terminals: $\{S, A, B, C\}$

B. Terminals: $\{x, y, z, w, \$\}$, non-terminals: $\{S, A, B, C\}$

C. Terminals: $\{x, y, z, w, \$, \lambda\}$, non-terminals: $\{S, A, B, C\}$

D. None of the above

Solution:

Terminals: $\{x, y, z, w, \$\}$, non-terminals: $\{S, A, B, C\}$

Correct Answer: B

206. For the following problem, consider the following context-free grammar:

$$S \rightarrow A\Phi$$
$$A \rightarrow xAx$$
$$A \rightarrow C$$
$$B \rightarrow yBy$$
$$B \rightarrow C$$
$$C \rightarrow zBz$$
$$C \rightarrow wAw$$
$$C \rightarrow \lambda$$

What is/are the first and follow sets for each of the non-terminals of the grammar? (Multiple Select Answers)

A. First(S) = $\{x, z, w, \$\}$, Follow ($S$) = $\{\}$

B. First(A) = $\{x, z, w, \lambda\}$, Follow (A) = $\{\$, x, w\}$

C. First(B) = $\{y, z, w, \lambda\}$, Follow (B) = $\{y, z\}$

D. First(C) = $\{z, w, \lambda\}$, Follow (C) = $\{\$, x, y, z, w\}$

Solution:

First(S) = $\{x, z, w, \$\}$
First(A) = $\{x, z, w, \lambda\}$
First(B) = $\{y, z, w, \lambda\}$
First(C) = $\{z, w, \lambda\}$

Follow (S) = $\{\}$
Follow (A) = $\{\$, x, w\}$
Follow (B) = $\{y, z\}$
Follow (C) = $\{\$, x, y, z, w\}$

Correct Answer: A; B; C; D

207. Given the following grammar

$S \rightarrow$ a S'
$S' \rightarrow$ b S' | ϵ

Which of the following is/are correct? (Multiple Select Answers)

A. FIRST(S) = $\{a\}$

B. FIRST (S') = $\{b, \epsilon\}$

C. FOLLOW(S)={$}

D. FOLLOW(S')={$}

Solution:

FIRST(S) = $\{a\}$

FIRST (S') = $\{b, \epsilon\}$

FOLLOW(S)={$}

FOLLOW(S')={$}

Correct Answer: A;B;C;D

208. Consider the following grammar:

$$\begin{aligned} S &\rightarrow bAb \mid bBa \\ A &\rightarrow aS \mid CB \\ B &\rightarrow b \mid BC \\ C &\rightarrow c \mid cC \end{aligned}$$

If the grammar is LL(1), write the answer as 1; otherwise, write 0. (Numerical Answer Type)

Solution:

There were actually three reasons, why the given grammar is not LL(1):

1. The grammar is not left-factored (non-terminals S and C).

2. The grammar is left-recursive ($B \rightarrow BC$).

3. The grammar is ambiguous (the string $bbcca$ has two different left-most derivations):

$\underline{\text{S}} \rightarrow \text{bBa} \rightarrow \text{bBCa} \rightarrow \text{bb}\underline{\text{Ca}} \rightarrow \text{bbcCa} \rightarrow bbcca,$

$\underline{\text{S}} \rightarrow \text{b}\underline{Ba} \rightarrow \text{b}\underline{\text{BCa}} \rightarrow \text{b}\underline{\text{BCCa}} \rightarrow \text{bb}\underline{\text{CCa}} \rightarrow \text{bbc}\underline{\text{Ca}} \rightarrow bbcca$

Correct Answer: 0

209. S $\rightarrow$ bT

T $\rightarrow$ Ab | Ba

$A \rightarrow aS$ | CB

B $\rightarrow$ bD

C $\rightarrow$ cD

D $\rightarrow \epsilon$ | cD

The first and follow sets for non-terminal is ________.

A. First(T) = $\{a,b,c\}$, Follow(B) = $\{a,b\}$

B. First(D) = $\{c\}$, Follow(C) = $\{b\}$

C. First(S) = $\{b\}$, Follow(T) = $\{a,b,\$\}$

D. None of the above

Solution:

Non-Terminal	First	Follow
S	$\{b\}$	$\{b,\$\}$
T	$\{a,b,c\}$	$\{b,\$\}$
A	$\{a,c\}$	$\{b\}$
B	$\{b\}$	$\{a,b\}$
C	$\{c\}$	$\{b\}$
D	$\{c,\epsilon\}$	$\{a,b\}$

Correct Answer: A

210. Given the following two CFGs over the alphabet $\Sigma = \{a,b,c\}$,

$$\begin{aligned} A &\rightarrow aaC \mid C \\ B &\rightarrow bB \mid a \\ C &\rightarrow c \mid B \end{aligned}$$

Which of the following is/are incorrect? (Multiple Select Answers)

A. The CFG is SLR(1).

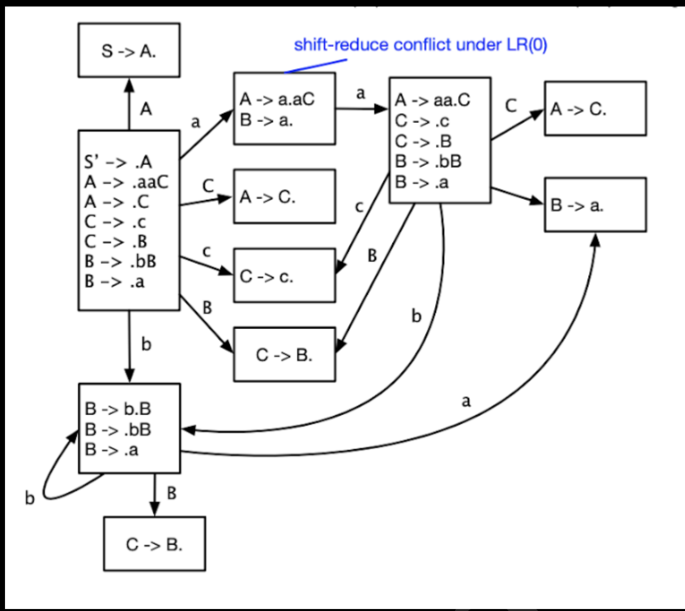
B. The CFG is LR(0).

C. The CFG is not SLR(1).

D. The CFG is not LR(0)

Solution:

The CFG is SLR(1). It is not in LR(0) because there is a shift-reduce conflict, which can be resolved in SLR(1) since Follow $(B) = \{\$\}$ does not contain $'\ a\ '$.



Correct Answer: B; C

211. Consider the following CFG over the language $\Sigma = \{a, b, (,), -, 0, 1\}$:

$$S \to aR(V \mid b(Q$$
$$Q \to bV1 \mid RQ \mid 0) -$$
$$V \to 0V \mid) -$$
$$R \to aR \mid 1$$

If the given grammar is LL(1), then write 1; otherwise, write 0. (Numerical Answer Type)

Solution:

The LL(1) parsing table:

	a	b	(	)	{	0	1
S	$aR(V$	$b(Q$					
Q	RQ	$bV1$				$0)-$	RQ
V				$)-$		$0V$	
R	aR						1

Yes, it's LL(1). Each cell in the parsing table has at most one item.

Correct Answer: 1

212. Consider the following grammar over the alphabet a, b, c :

$$S \rightarrow Xa$$
$$X \rightarrow bX$$
$$X \rightarrow Y$$
$$Y \rightarrow Zc$$
$$Z \rightarrow bZ$$
$$Z \rightarrow \epsilon$$

Which of the following is/are correct? (Multiple Select Answers)

A. LL(1)

B. Not LL(1)

C. LR(0)

D. Not LR(0)

Solution:

The easiest way to show that the grammar is not LL(1) is to show that it is ambiguous. For instance, consider the string "bca"

$$S \rightarrow Xa \rightarrow bXa \rightarrow bYa \rightarrow bZca \rightarrow bca$$
$$S \rightarrow Xa \rightarrow Ya \rightarrow Zca \rightarrow bZca \rightarrow bca$$

These two leftmost derivations form two different parse trees, and therefore the grammar is ambiguous.

An LR(0) grammar is one where:

There are no conflicts in an LR(0) parsing table. Specifically, no shift/reduce or reduce/reduce conflicts.

LR(0) parsers are more lenient with left recursion than LL(1), but each state of the LR(0) automaton must be able to make a clear decision based only on the current symbol.

This grammar, due to the ϵ-productions and the multiple possibilities for starting a string with b (via $X \rightarrow bX$ or $Y \rightarrow Zc$), is likely to introduce reduce/reduce conflicts in an LR(0) parser.

Therefore, this grammar is Not LR(0).

Correct Answer: B; D

213. Consider the following flex specification that your friend has written:

c	{ printf("1"); }
$ac+b*$	{ printf("2"); }
cC	{ printf("3"); }
ab*	{ printf("4"); }
a	{ printf("5"); }
ac * b+	{ printf("6"); }

Which rules can never be executed? (Multiple Select Answers)

(Note that the analyzer outputs the token that matches the longest possible prefix).

A. Rule 2

B. Rule 4

C. Rule 5

D. Rule 6

Solution:

Rule 5 (a) can never be executed, because it is included in rule 4 ($ab*$).

Rule 6 ($ac*b+$) can never be executed, because when there is no c (i.e. $ab+$) it is included in rule 4 ($ab*$), and when there are one or more c's (i.e. $ac+b+$) it is included in rule 2 (ac + b*).

Correct Answer: C; D

214. The following problem describe a deterministic (i.e., DFA)LR(0) parsing automaton. Show your grammar and fill in the parsing automaton with transitions and each state labeled with its set of LR(0) items. You do not need to analyze the automaton to determine whether the grammar is LR(0) or SLR(1). The grammar and the automaton constitute a complete answer to each sub-problem.

Assume that the first step of the automaton construction is to add a new production $S' \to S$ to the grammar, as described in class. This production should be included in your grammars, in your automatons, and in your counts.

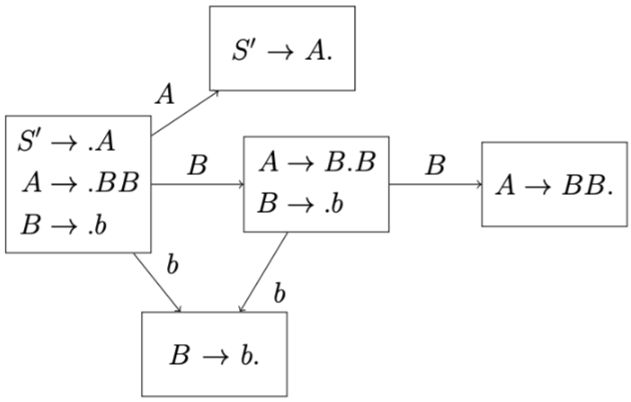
Give the simplest possible grammar (fewest productions and fewest terminals) that result in a parsing automaton satisfying the description.

An automaton (and corresponding CFG) with a minimal number of states is α, without loops, where one state has two incoming transitions.

The unit digit of $\alpha^{2025^{2025^{2025!}}}$ is ________. (Numerical Answer Type)

Solution:

$$A \to BB$$
$$B \to b$$



(Note: We accepted any solution with at most 7 states as close enough to minimal.)

Correct Answer: 5

215. Which of the following sentences regarding Viable prefixes is/are correct? (Multiple Select Answers)

A. Viable prefixes is the set of prefixes of right-sentential forms that do not extend past the end of the right-most handle.

B. Viable prefixes can be recognized using a DFA.

C. Viable prefixes is the set of prefixes of right-sentential forms that can appear on the stack of a shift-reduce parser.

D. For any context-free grammar, the set of viable prefixes is a regular language.

Solution:

Correct Answer: A; B; C; D

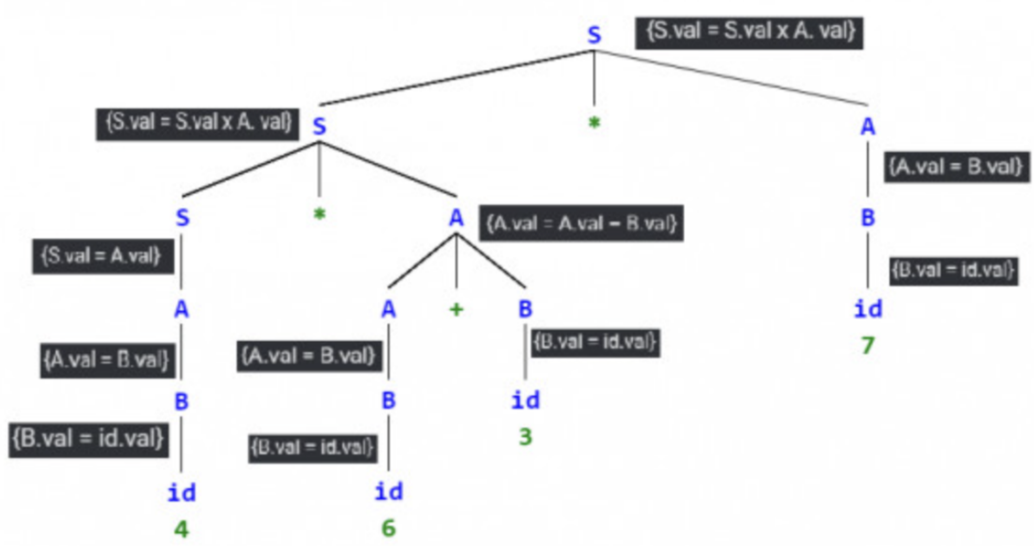
216. Consider the following grammar and their syntax-directed translation (SDT) rules.

$$
\begin{array}{ll}
S \rightarrow S * A & \{S.val = S.val \times A.\text{ val }\} \\
S \rightarrow A & \{S.val = A.val\} \\
A \rightarrow A + B & \{A.val = A.val - B.val\} \\
B \rightarrow (S)' & \{B.val = 2\} \\
A \rightarrow B & \{\text{ A.val } = \text{ B.val }\} \\
B \rightarrow \text{ id} & \{\text{ B.val } = \text{ id.val }\}
\end{array}
$$

What is the evaluation value of the expression $4 * 6 + 3 * 7$? (Numerical Answer Type)

Solution:

The syntax tree:



The expression will be evaluated as : **(4* (6 – 3))* 7 = 84**

Correct Answer: 84

217. Which one of the following statements is TRUE?

A. SLR parser has more states than LALR parser.

B. LALR parser has more states than Canonical LR.

C. Canonical LR has fewer states than SLR parser.

D. Both SLR and LALR parsers have the same number of states.

Solution:

If number of states LR(0) = n_1, number of states SLR(1) = n_2, number of states LALR(1) = n_3, number of states CLR(1) = n_4 then, $n_1 = n_2 = n_3 \leq n_4$.

Correct Answer: D

218. A context-free grammar G is ambiguous iff

A. G is not in Chomsky's normal form.

B. some string $w \in L(G)$ has at least two different derivations.

C. some string $w \in L(G)$ has at least two different parse trees.

D. every string $w \in L$ (G) has at least two different parse trees.

Solution:

Having two derivations for some string w doesn't make a CFG ambiguous. For example, $S \rightarrow AB; A \rightarrow a; B \rightarrow b$

For string $w = ab$, we have two derivations, one leftmost and one rightmost.

Correct Answer: C

219. Given a grammar $S1 \rightarrow Sc, S \rightarrow SA \mid A, A \rightarrow aSb \mid ab$, there is a rightmost derivation $S1 \Rightarrow Sc \Rightarrow SAc \Rightarrow SaSbc$ Thus, $SaSbc$ is a right sentential form, and its handle is _______.

A. SaS

B. bc

C. Sbc

D. aSb

Solution:

For an Answer visit: https://gateoverflow.in/58011/ugc-net-cse-june-2013-part-2-question-22?show=409708#a409708

Correct Answer: D

220. Which of the following is/are incorrect? (Multiple Select Answers)

A. Every lexical token can also be specified using context-free grammar.

B. Predictive parsing is less powerful than LL parsing with backtracking.

C. The number of entries in the state stack used by the LR parsing routine is independent of the size of the input string.

D. The complexity of the LR parsing algorithm is less than that of the CYK algorithm.

Solution:

- **A. Every lexical token can also be specified using context-free grammar.**
  - **Correct.** Lexical tokens can be described by context-free grammars, though they are typically specified using regular expressions.
- **B. Predictive parsing is less powerful than LL parsing with backtracking.**
  - **Incorrect.** Predictive parsing (LL(1)) is limited, while LL parsing with backtracking can handle a wider range of grammars, making it more powerful.
- **C. The number of entries in the state stack used by the LR parsing routine is independent of the size of the input string.**
  - **Correct.** The state stack size depends on the grammar and the number of states, not on the length of the input string.
- **D. The complexity of the LR parsing algorithm is less than that of the CYK algorithm.**
  - **Incorrect.** LR parsing has a time complexity of $O(n)$, while CYK has $O(n^3)$, making LR parsing generally more efficient.

Correct Answer: B; D

11 Computer Networks (20 Questions)

221. Assuming circuit switching, if the network has allocated exactly enough bandwidth to handle the application's peak bandwidth P, and the application has average bandwidth A, the level of utilization of that application's circuit is given by (where / denotes division):

A. AP

B. P/A

C. $(P-A)/A$

D. $(P-A)/P$

Solution:

Correct Answer: A

222. Consider an IP packet (without options) with total length 1500 bytes. The packet is split into fragments by a network that can only handle IP packets of up to 500 bytes. Which set of IP packet lengths could describe the set of fragments?

A. $500, 500, 500$

B. $400, 400, 400, 300$

C. $420, 420, 420, 300$

D. None of the above

Solution:

Correct Answer: C

223. The IP packet header includes a time-to-live field that is decremented by each router along the path. Why is the time-to-live field necessary?

A. The TTL field is decremented at each router along the path, and routers drop packets with a TTL of 0, so the TTL field prevents packets from looping indefinitely if they are stuck in a forwarding loop.

B. The TTL field allows hosts to determine the appropriate window size.

C. The TTL field is necessary for scalability. Routers store each packet for the amount of time given by the TTL field; without the TTL field, routers would need to store packets forever, which would consume infinite space.

D. The TTL field is not necessary, which is why it was removed in IPv6.

Solution:

Correct Answer: A

224. Consider two packets that are sent back to back (i.e., one right after another) along the path A-B-C-D, where A, B, C, and D are store-and-forward routers. Assume the capacity of link (A, B) is 10Mbps, the capacity of (B, C) is 1 Mbps, and the capacity of (C, D) is 100 Mbps. The propagation delay along each link is 10 ms. Assume the size of the first packet is 1000 bits, and that the size of the second packet is 500 bits.

What is the inter-arrival time between the two packets at node D ? The inter-arrival time is equal to the arrival time of the second packet minus the arrival time of the first packet at node D.

Notes: Assume there is no cross-traffic; the two packets are the only ones in the network. The arrival time of a packet is the time when the last bit of the packet was received. With a store-and-forward router, a packet is forwarded only after the last bit of the packet was received.

A. 0.459 ms

B. 0.495 ms

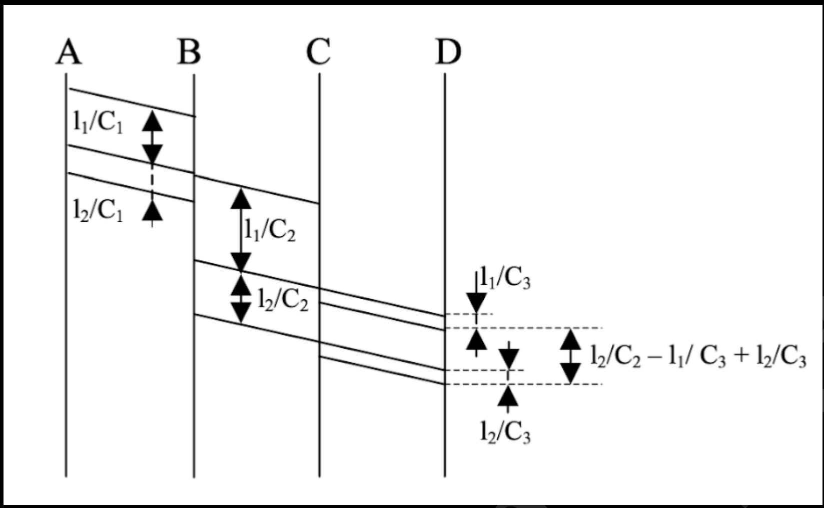
C. 0.594 ms

D. 0.954 ms

Solution:

See the time diagram below bellow ($l_1 = 1000$ b, $l_2 = 500$ b, $C_1 = 10$Mbps, $C_2 = 1$Mbps, $C_3 = 100$Mbps)

The inter-arrival time between the two packets at D is $1_2/C_2 - 1_1/C_3 + 1_2/C_3 = 500$ b$/10^6$Mbps$-$ 1000 b$/10^8$Mbps $+ 500$ b$/10^8$Mbps $= 0.495$ ms



Correct Answer: B

225. You are given the following network prefixes:

171.64.121.0/24	171.64.180.0/24
171.64.18.0/24	171.64.181.0/24
171.64.19.0/24	171.64.187.0/24

You are to aggregate the prefixes together into the smallest possible number of shorter prefixes. What is the smallest number of prefixes, and what are they?

A. 1 prefix; 171.64.179.0/21

B. 2 prefixes; 171.64.180.0/20, 171.64.18.0/21

C. 4 prefixes; 171.64.180.0/23, 171.64.187.0/24, 171.64.121.0/24, 171.64.18.0/23

D. 3 prefixes 171.64.180.0/20, 171.64.187.0/24, 171.64.18.0/23

Solution:

Correct Answer: C

226. Which of the following is/are true? (Multiple Select Answers)

A. 171.64.128/16 cannot be a prefix because it is a Class B address.

B. If a routing table contains prefixes 31.75/16 (for which packets are sent to port 1) and 31.75.93.128/25 (for which packets are sent to port 2) then an arriving packet with IP address 31.75.93.129 will be sent to port 2.

C. A routing table can correctly contain the two prefixes 128.50.128/17 and 128.50.128/18 simultaneously.

D. If a routing table is organized in order of increasing prefix length, then a routing decision may be performed by finding the last matching prefix.

Solution:

Correct Answer: B; C; D

227. You notice that the packets you send to your friend always follow the same path and take a minimum of 100 ms to reach her, with a variation in delay from $100-150$ ms, and an average of 125 ms . You send an average of 100 new packets per second, and none of them are dropped. What is a good estimate of the average number of your packets buffered in routers along the path?

A. 125

B. 2.5

C. 12.5

D. There isn't enough information in the question to make a good estimate.

Solution:

The average queuing delay is 25 ms.

By Little's law, The average number of packets in the queue = $100 * 25$ m = 2.5

Correct Answer: B

228. Why is the retransmission timeout (RTO) approximately equal to the round-trip-time (RTT)? (Multiple Select Answers)

A. A timeout much smaller than the estimated RTT will lead to unnecessary retransmissions.

B. A timeout that is roughly equal to the RTT gives enough time to the receiver application to process data.

C. A timeout that is roughly equal to the RTT gives the sender enough time to receive ACKs from the receiver.

D. A timeout much larger than the estimated RTT may result in poor network utilization.

Solution:

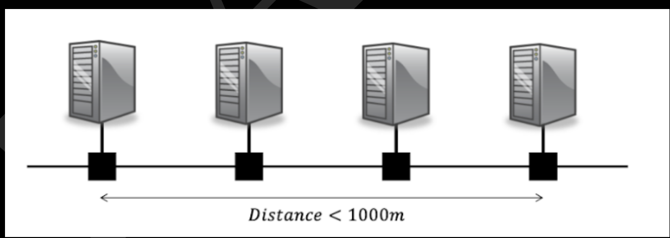
Option A because timeout shorter than RTT means all the packets will be retransmitted before the acknowledgement arrives even if they are successfully delivered to the destination.

Option C because confirmation regarding a successful delivery of a packet arrives in RTT seconds to the destination.

Option D because waiting too long before transmitting a lost packet may cause throughput of the flow degrade as it would prevent the congestion window to proceed.

Correct Answer: A; C; D

229.



A 1000 m long 1 Gb/s Ethernet network uses CSMA/CD to control access to a shared copper cable as shown in the figure. The Ethernet specification requires that if a collision occurs, it must detect the collision before it finishes transmitting a packet. What is the size of the minimum packet in this network? (Assume the speed of propagation is 2×10^8 m/s.)

A. 10^4 bits

B. 1250 bytes

C. 10^5 bits

D. 1012 bytes

Solution:

The RTT on the cable is 2000 m/$2{\times}10^8 = 10^{-5}$ s. Therefore $P/10^9 \geq 10^{-5}$ which means the packet size must be greater than 10^4 bits or 1250 bytes.

Correct Answer: A; B

230. Consider an Ethernet network with three hosts, Host A, Host B, and Host C as shown in Figure 1. No machine is configured as an IP router, and there is no IP router on this network. Assume that the IP addresses and subnet masks are as shown in the figure.

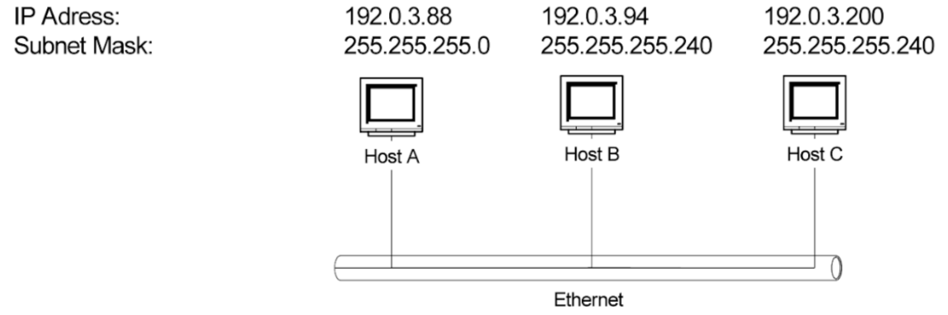


Figure 1: Network Topology

For each of the following IP datagram transmissions, which of the transmissions will be successful?

A. Host C sends an IP datagram to Host A

B. Host A sends an IP datagram to Host C

C. Host B sends an IP datagram to Host A

D. Host B sends an IP datagram to Host C

Solution:

A. Host C sends an IP datagram to Host A does not work

B. Host A sends an IP datagram to Host C works

C. Host B sends an IP datagram to Host A works

D. Host B sends an IP datagram to Host C works does not work

Correct Answer: B; C

231. An IP router (which uses Longest Prefix Matching to make forwarding decisions) has the following table, which maps classful prefixes to outgoing interfaces:

Prefix	Interface
200.10.192.0/24	3
200.10.128.0/24	1
200.10.160.0/24	2
200.10.224.0/24	3

A packet with destination IP 200.10.187.11 will be forwarded to interface ________. (Numerical Answer Type)

Solution:

Correct Answer: 2

232. One advantage of a path vector protocol over distance vector and link state is that

A. it is guaranteed to converge to the shortest paths

B. it enables the implementation of policies that may not correspond to simple metrics

C. it enables load balancing across parallel paths

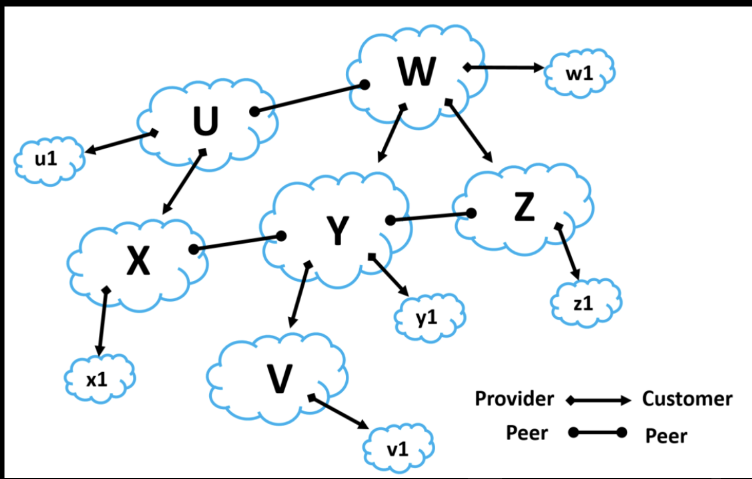
D. None of the above

Solution:

Correct Answer: B

233. The Commander is devising an invasion strategy based on the enemy network's communication patterns. You have mapped out the enemy's network below, and in addition, you discover that they, just like Earth, also use standard BGP routing policies. In particular, they also follow the Gao-Rexford model, which includes valley-free routing.

Which paths are possible? (Multiple Select Answers)



A. x1 → X → U → W → Y → V → v1

B. v1 → V → Y → Z → z1

C. v1 → V → Y → Z → W → w1

D. **u**1 → U → W → Y → y1

Solution:

Option A is not valid because X prefers to route through Y (a peer) instead of U (a provider), so this route would never be selected. This means $x1$ would never even know about the route x1 – X – U – W – Y – V – v1.

Correct Answer: B; D

234. Two computers exchange messages over a link and agree to use a CRC with the generator $x^3 + 1$ to detect errors. One computer wishes to send the message 10011001000010011101. What remainder should it append to the message?

A. $R = 000$

B. $R = 0100$

C. $R = 100$

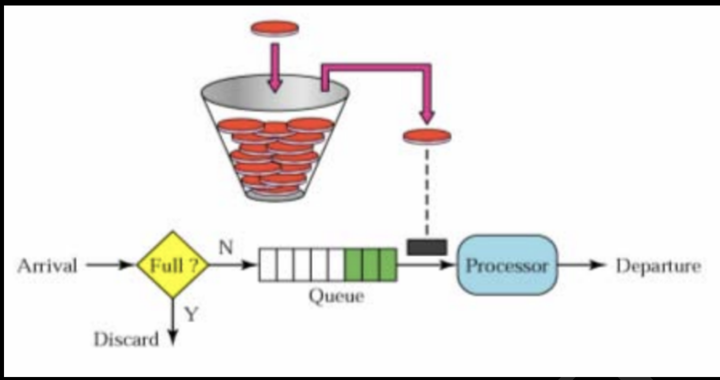
D. None of the above.

Solution:

Correct Answer: C

235. Host A has to 'inject' 30 Mbits of data into a network via a token bucket regulator. The token bucket has a capacity of 15 Mbits and is filled with tokens at the rate of 5 Mbps . Data is buffered if it arrives at the regulator when there are no tokens in the bucket.

How long does it take, in total, for the 30 Mbits of data to enter the network, assuming that the host sends at a peak rate of 20 Mbps and the token bucket is initially full?



A. 7 seconds

B. 5 seconds

C. 3 seconds

D. 2 seconds

Solution:

The host can send all 30 Mbits to the token bucket at the rate of 20 Mbps.

The data arrives to the token bucket in two phases:

1st second: first burst of size 20 Mbps arrives
2nd second: burst of size 10 Mbps arrives

The data leaves the token bucket (i.e. enters the network) in three phases:

1st second: burst of 20 Mbits, using the full bucket (15 Mbits) and the first next token that arrives
2nd second: burst of 5 Mbits (using the 2nd token)
3rd second: burst of 5 Mbits (using the 3rd token)

Thus, it takes (at most) 3 seconds for the data to enter the network.

Correct Answer: C

236. Which of the following statements is/are true about physical and link-layer protocols?

A. Manchester encoding requires that sender and receiver have clocks that run at approximately the same rate, so they can differentiate between the encoding of a single " 0 " and than of multiple " 0 "s.

B. When many hosts seek to actively communicate, token-ring schemes can achieve higher total goodput on a shared LAN than Ethernet.

C. An Ethernet adapter passes every non-corrupt frame that it receives up to the network layer.

D. Ethernet switches (as compared to hubs) eliminate the need for broadcasting packets to all hosts.

Solution:

Correct Answer: B

237. Which of the following link protocols best describes Ethernet?

A. frequency-division multiple access (FDMA)

B. time-division multiple access (TDMA)

C. token-passing

D. carrier sense multiple access (CSMA)

Solution:

Correct Answer: D

238. Which statement is TRUE with regard to the source and destination MAC addresses of a packet as it traverses from one network to another across an internetwork environment?

A. These addresses will remain the same for the entire route.

B. The destination MAC, and only the destination MAC, will change on each network hop.

C. The source MAC, and only the source MAC, will change on each network hop.

D. Both source and destination MAC addresses will change on each network hop.

Solution:

Correct Answer: D

239. TCP and UDP use the 16-bit Internet Checksum for computing the value of the checksum field in each segment. In this question you are asked to use the same method but for simplicity we restrict to 8-bit words. Suppose you have the following three 8-bit words:

$$11101011$$
$$01010101$$
$$10110111$$

What is the value of the 8-bit Internet Checksum (i.e., the 1 s complement of the 8 -bit sum) of these words?

A. 00000111

B. 00001000

C. 00001010

D. None of the above

Solution:

Correct Answer: A

240. Consider the Go-Back-N protocol with a sender window size of $N = 4$ and a sequence number range of 1024. Suppose at time t the next in-order packet that the receiver is expecting has sequence number 228. Assume that the medium does not reorder messages. Which of the following is a possible pair of values for the sender's `send_base` and `nextseqnum` pointers at time t?

A. `send_base` = 224; `nextseqnum` = 226

B. `send_base` = 225; `nextseqnum` = 228

C. `send_base` = 227; `nextseqnum` = 232

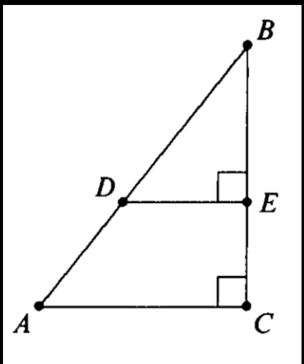
D. `send_base` = 229; `nextseqnum` = 233

Solution:

Correct Answer: B

12 General Aptitude (35 Questions)

241. In triangle $ABC, \angle C = 90°, AC = 6$ and $BC = 8$. Points D and E are on $\overline{AB}$ and $\overline{BC}$, respectively, and $\angle BED = 90°$. If $DE = 4$, then $BD =$



A. 5

B. 16/3

C. 20/3

D. 15/2

Solution:

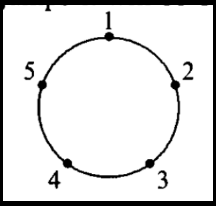
$\triangle ABC$ is a right triangle, the Pythagorean Theorem implies that $BA = 10$. Since $\triangle DBE \sim \triangle ABC$,

$$\frac{BD}{BA} = \frac{DE}{AC}. \quad \text{So} \quad BD = \frac{DE}{AC}(BA) = \frac{4}{6}(10) = \frac{20}{3}$$

OR Since $\sin B = DE/BD$, we have $BD = DE/\sin B$. Moreover, $BA = 10$ by the Pythagorean Theorem, so $\sin B = AC/BA = 3/5$. Hence $BD = 4 \div 3/5 = 20/3$.

Correct Answer: C

242. Five points on a circle are numbered 1, 2, 3, 4, and 5 in clockwise order. A bug jumps in a clockwise direction from one point to another around the circle; if it is on an odd-numbered point, it moves one point, and if it is on an even-numbered point, it moves two points. If the bug begins on point 5, after 1995 jumps it will be on point



A. 1

B. 2

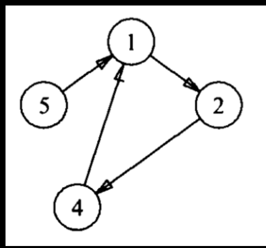
C. 3

D. 4

Solution:

With the first jump, the bug moves to point 1, with the second to 2, with the third to 4, and with the fourth it returns to 1. Thereafter, every third jump it returns to 1.

Thus, after $n > 0$ jumps, the bug will be on 1,2, or 4, depending on whether n is of the form $3k+1$, $3k+2$,or $3k$, respectively. Since $1995 = 3(665)$, the bug will be on point 4 after the 1995 jumps.



Correct Answer: D

243. There exist positive integers A, B, and C, with no common factor greater than 1 , such that

$$A \log_{200} 5 + B \log_{200} 2 = C$$

What is $A+B+C$?

A. 6

B. 7

C. 8

D. 9

Solution:

Note that $C = A\log_{200} 5 + B\log_{200} 2 = \log_{200} 5^A + \log_{200} 2^B = \log_{200}\left(5^A \cdot 2^B\right)$, so $200^C = 5^A \cdot 2^B$. Therefore, $5^A \cdot 2^B = 200^C = \left(5^2 \cdot 2^3\right)^C = 5^{2C} 2^{3C}$.

By uniqueness of prime factorization, $A = 2C$ and $B = 3C$. Letting $C = 1$ we get $A = 2, B = 3$ and $A + B + C = 6$. The triplet $(A, B, C) = (2, 3, 1)$ is the only solution with no common factor greater than 1.

Note. The uniqueness is guaranteed by the Fundamental Theorem of Arithmetic.

Correct Answer: A

244. Consider the triangular array of numbers with $0, 1, 2, 3, \ldots$ along the sides and interior numbers obtained by adding the two adjacent numbers in the previous row. Rows 1 through 6 are shown.

0
1 1
2 2 2
3 4 4 3
4 7 8 7 4
5 11 15 15 11 5

Let $f(n)$ denote the sum of the numbers in row n. What is the remainder when $f(100)$ is divided by 100?

A. 12

B. 30

C. 50

D. 74

Solution:

Calculating the first five values of f,

$$f(1) = 0, \quad f(2) = 2, \quad f(3) = 6, \quad f(4) = 14, \quad f(5) = 30,$$

we are led to the conjecture that $f(n) = 2^n - 2$. We prove this by induction: Observe that each of the interior numbers in row n is used twice and each of the end numbers is used once as a term in computing the interior terms of row $n+1$; i.e.,

$$f(n+1) = [2f(n) - 2(n-1)] + 2n = 2f(n) + 2$$

so if $f(n) = 2^n - 2$, then $f(n+1) = 2f(n) + 2 = 2(2^n - 2) + 2 = 2^{n+1} - 2$. Therefore, we seek the remainder when $f(100) = 2^{100} - 2$ is divided by 100 . Use the fact that 76^2 has a remainder 76 when divided by 100. We find

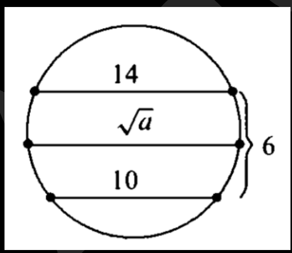
$$\begin{aligned} 2^{10} &= 100K + 24, \\ 2^{20} &= 100L + 76, \\ 2^{40} &= 100M + 76, \\ 2^{80} &= 100N + 76, \\ 2^{100} &= 100Q + 76, \end{aligned}$$

for positive integers K, L, M, N, Q, so $f(100) = 2^{100} - 2$ has remainder 74 when divided by 100.

Correct Answer: D

Query: What other positive integers N have the property that N^2 has remainder N when divided by 100?

245. Two parallel chords in a circle have lengths 10 and 14, and the distance between them is 6. The chord parallel to these chords and midway between them is of length $\sqrt{a}$ where a is _______.



A. 184

B. 156

C. 168

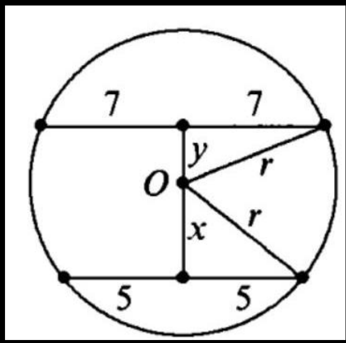
D. 176

Solution:

Let x be the distance from the center O of the circle to the chord of length 10 and let y be the distance from O to the chord of length 14. Let r be the radius. Then,

$$\begin{aligned} & x^2 + 25 = r^2, \\ & y^2 + 49 = r^2, \\ \text{so} \quad & x^2 + 25 = y^2 + 49. \end{aligned}$$

Therefore, $\quad x^2 - y^2 = (x-y)(x+y) = 24.$



If the chords are on the same side of the center of the circle, $x - y = 6$. If they are on opposite sides, $x + y = 6$. But $x - y = 6$ implies that $x + y = 4$, which is impossible. Hence $x + y = 6$ and $x - y = 4$. Solve these equations simultaneously to get $x = 5$ and $y = 1$. Thus, $r^2 = 50$, and the chord parallel to the given chords and
midway between them is two units from the center. If the chord is of length $2d$, then $d^2 + 4 = 50, d^2 = 46$, and $a = (2d)^2 = 184$.

OR

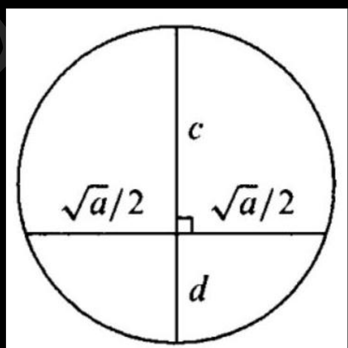
The diameter perpendicular to the chords is divided by the chord of length $\sqrt{a}$ into segments with lengths c and d as shown. Then

$$cd = \left(\frac{\sqrt{a}}{2}\right)^2 = \frac{a}{4}$$

Treat the chords 3 units above and 3 units below similarly:

$$(c-3)(d+3) = \left(\frac{14}{2}\right)^2$$
$$(c+3)(d-3) = \left(\frac{10}{2}\right)^2$$

Adding the last two equations, we get $2cd - 18 = 49 + 25 = 74$. Thus, $2cd = 92$ so $a = 4cd = 184$.



Correct Answer: A

246. Sunny runs at a steady rate, and Moonbeam runs m times as fast, where m is a number greater than 1. If Moonbeam gives Sunny a head start of h meters, how many meters must Moonbeam run to overtake Sunny?

A. $\dfrac{h+m}{m-1}$

B. $\dfrac{h}{h+m}$

C. $\dfrac{h}{m-1}$

D. $\dfrac{\mathrm{hm}}{\mathrm{m}-1}$

Solution:

Let x be the number of meters that Moonbeam runs to overtake Sunny, and let r and mr be the rates of Sunny and Moonbeam, respectively. Because Sunny runs $x-h$ meters in the same time that Moonbeam runs x meters, it follows that $x - h/r = x/mr$. Solving for x, we get $x = hm/m - 1$.

Correct Answer: D

247. A circle of radius 2 has center at $(2,0)$. A circle of radius 1 has center at $(5,0)$. A line is tangent to the two circles at points in the first quadrant. Which of the following is closest to the y-intercept of the line?

A. $\sqrt{2}/4$

B. $8/3$

C. $1+\sqrt{3}$

D. $2\sqrt{2}$

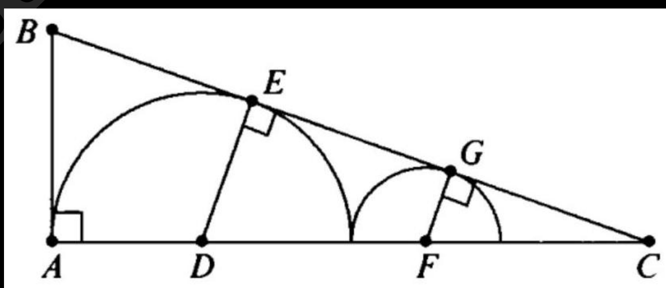
Solution:

Let D and F denote the centers of the circles. Let C and B be the points where the x-axis and y-axis intersect the tangent line, respectively. Let E and G denote the points of tangency as shown. We know that $AD = DE = 2, DF = 3$, and $FG = 1$. Let $FC = u$ and $AB = y$. Triangles FGC and DEC are similar, so

$$\frac{u}{1} = \frac{u+3}{2}$$

which yields $u = 3$. Hence, $GC = \sqrt{8}$. Also, triangles BAC and FGC are similar, which yields

$$\frac{y}{1} = \frac{BA}{FG} = \frac{AC}{GC} = \frac{8}{\sqrt{8}} = \sqrt{8} = 2\sqrt{2}$$



Correct Answer: D

248. The number of geese in a flock increases so that the difference between the populations in year $n+2$ and year n is directly proportional to the population in year $n+1$. If the populations in the years 1994, 1995, and 1997 were 39,60, and 123, respectively, then the population in 1996 was ____.

A. 81

B. 84

C. 87

D. 90

Solution:

Let x be the number of geese in 1996, and let k be the constant of proportionality. Then $x - 39 = 60k$ and $123 - 60 = kx$. Solve the second equation for k, and use that value to solve for x in the first equation, obtaining $x - 39 = 60 \cdot 63/x$. Thus $x^2 - 39x - 3780 = 0$. Factoring yields $(x - 84)(x + 45) = 0$. Since x is positive, it follows that $x = 84$.

Correct Answer: B

249. How many ordered triples of integers (a, b, c) satisfy

$$|a+b|+c=19 \text{ and } ab+|c|=97\ ?$$

A. 0

B. 6

C. 10

D. 12

Solution:

If $c \geq 0$, then $ab - |a+b| = 78$, so $(a-1)(b-1) = 79$ or $(a+1)(b+1) = 79$.

Since 79 is prime, $\{a, b\}$ is $\{2, 80\}, \{-78, 0\}$, $\{0, 78\}$, or $\{-80, -2\}$.

Hence, $|a+b| = 78$ or $|a+b| = 82$, and from the first equation in the hypothesis, it follows that $c < 0$, a contradiction.

On the other hand, if $c < 0$, then $ab + |a+b| = 116$, so $(a+1)(b+1) = 117$ or $(a-1)(b-1) = 117$.

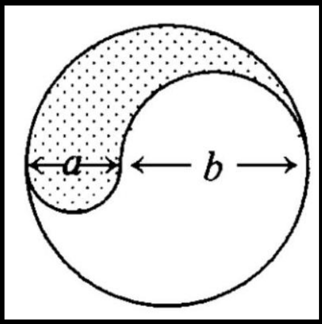
Since $117 = 3^2 \cdot 13$, we distinguish the following cases:

$$\begin{gathered}
\{a, b\} = \{0, 116\} \text{ yields } c = -97 \\
\{a, b\} = \{2, 38\} \text{ yields } c = -21; \\
\{a, b\} = \{8, 12\} \text{ yields } c = -1; \\
\{a, b\} = \{-116, 0\} \text{ yields } c = -97 \\
\{a, b\} = \{-38, -2\} \text{ yields } c = -21; \\
\{a, b\} = \{-12, -8\} \text{ yields } c = -1
\end{gathered}$$

Since a and b are interchangeable, each of these cases leads to two solutions, for a total of 12.

Correct Answer: D

250. The figure shown is the union of a circle and two semicircles of diameters a and b, all of whose centers are collinear. The ratio of the area of the shaded region to that of the unshaded region is ________.



A. $\sqrt{\frac{a}{b}}$ B. $\frac{a}{b}$

C. $\frac{a^2}{b^2}$ D. $\frac{a^2+2ab}{b^2+2ab}$

Solution:

The area of the shaded region is

$$\frac{\pi}{2}\left(\left(\frac{a+b}{2}\right)^2+\left(\frac{a}{2}\right)^2-\left(\frac{b}{2}\right)^2\right)=\frac{\pi}{2}\frac{a+b}{2}\left(\frac{a+b}{2}+\frac{a-b}{2}\right)$$
$$=\frac{\pi(a+b)a}{4}$$

and the area of the unshaded region is

$$\frac{\pi}{2}\left(\left(\frac{a+b}{2}\right)^2-\left(\frac{a}{2}\right)^2+\left(\frac{b}{2}\right)^2\right)=\frac{\pi}{2}\frac{a+b}{2}\left(\frac{a+b}{2}+\frac{b-a}{2}\right)$$
$$=\frac{\pi(a+b)b}{4}$$

Their ratio is a/b.

Correct Answer: B

251. Call a 7-digit telephone number $d_1\ d_2\ d_3-d_4\ d_5\ d_6\ d_7$ memorable if the prefix sequence $d_1d_2d_3$ is exactly the same as either of the sequences $d_4\ d_5\ d_6$ or $d_5\ d_6\ d_7$ (possibly both). Assuming that each d_i can be any of the ten decimal digits $0,1,2,\ldots 9$, the number of different memorable telephone numbers is ______.

A. 19,810

B. 19,910

C. 19,990

D. 20,000

Solution:

There are 10,000 ways to write the last four digits $d_4\ d_5\ d_6\ d_7$, and among these there are $10000-10=9990$ for which not all the digits are the same. For each of these, there are exactly two ways to adjoin the three digits $d_1d_2d_3$ to obtain a memorable number. There are ten memorable numbers for which the last four digits are the same, for a total of $2\cdot 9990+10=19990$.

OR

Let A denote the set of telephone numbers for which $d_1\ d_2\ d_3$ and $d_4\ d_5\ d_6$ are identical and B the set for which $d_1\ d_2\ d_3$ is the same as $d_5\ d_6\ d_7$. A number $d_1\ d_2\ d_3-d_4\ d_5\ d_6\ d_7$ belongs to

$A \cap B$ if and only if $d_1 = d_4 = d_5 = d_2 = d_6 = d_3 = d_7$. Hence, $n(A \cap B) = 10$. Thus, by the Inclusion-Exclusion Principle, $n(A \cup B) =$

$$n(A) + n(B) - n(A \cap B) = 10^3 \cdot 1 \cdot 10 + 10^3 \cdot 10 \cdot 1 - 10 = 19990$$

Correct Answer: C

252. Which one of the following statements is false?

A. All equilateral triangles are congruent to each other.

B. All equilateral triangles are convex.

C. All equilateral triangles are equiangular.

D. All equilateral triangles are regular polygons.

E. All equilateral triangles are similar to each other.

Solution:

(A) Triangles with side lengths of $1, 1, 1$ and $2, 2, 2$ are equilateral and not congruent, so (A) is false. Statement (B) is true since all triangles are convex. Statements (C) and (E) are true since each interior angle of an equilateral triangle measures $60°$. Furthermore, all three sides of an equilateral triangle have the same length, so (D) is also true.

Correct Answer: A

253. The marked price of a book was 30% less than the suggested retail price. Alice purchased the book for half the marked price at a Fiftieth Anniversary sale. What percent of the suggested retail price did Alice pay?

A. 25%

B. 30%

C. 35%

D. 60%

Solution:

If the suggested retail price was P, then the marked price was $0.7P$. Half of this is $0.35P$, so Alice paid 35% of the suggested retail price.

Correct Answer: C

254. The number of all ordered pairs (a, b) of positive integers for which

$$\frac{1}{a} + \frac{1}{b} = \frac{3}{2018}$$

is ______. (Numerical Answer Type)

Solution:

The given equation is equivalent to

$$2018(a + b) = 3ab$$

It possible, but not so pleasant, to work with this equation. It's much easy to work with the equivalent equation

$$(3a - 2018)(3b - 2018) = 2018^2$$

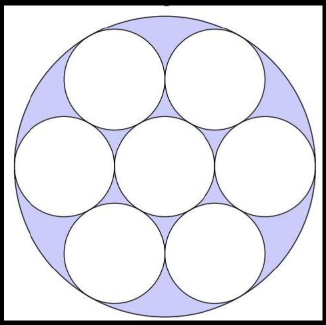
Each of the factors on the left is congruent to $1 (\bmod 3)$. There are 6 positive factors of $2018^2 = 2^2 \cdot 1009^2$ that are congruent to $1 (\bmod 3) : 1, 2^2, 1009, 2^2 \cdot 1009, 1009^2, 2^2 \cdot 1009^2$.

These lead to the 6 possible pairs: $(a, b) = (673, 1358114), (674, 340033), (1009, 2018), (2018, 1009), (340033, 674)$, and $(1358114, 673)$.

As for negative factors, the ones that are congruent to $1 (\bmod 3)$ are $-2, -2 \cdot 1009, -2.\ 1009^2$. However, all of these lead to pairs where $a \le 0$ or $b \le 0$.

Correct Answer: 9

255. Each of the small circles in the figure has radius one. The innermost circle is tangent to the six circles that surround it, and each of those circles is tangent to the large circle and to its small-circle neighbors.



The area of the shaded region is _______.

A. π

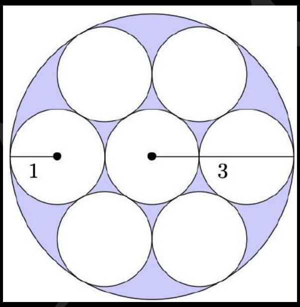
B. 1.5π

C. 2π

D. 3π

Solution:

The outer larger circle has radius 3 as shown in the figure below



and so we find that

(the area of the shaded region) = (the area of the larger circle) - (the areas of the 7 smaller circles)

$$= \pi \cdot 3^2 - 7 \cdot \left(\pi \cdot 1^2\right)$$
$$= 2\pi.$$

Correct Answer: C

256. The remainder of $3^{21} + 5^{21}$ divided by 64 is _______.

A. 40

B. 20

C. 12

D. 10

Solution:

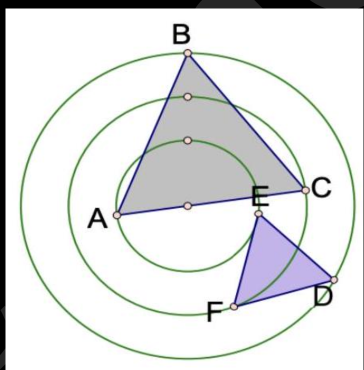
After apply in the binomial theorem we work in modulo 64 arithmetic

$$
\begin{aligned}
3^{21}+5^{21} &= (4-1)^{21}+(4+1)^{21} \\
&= \left[\sum_{k=0}^{21}\binom{21}{k}4^k\cdot(-1)^{21-k}\right]+\left[\sum_{k=0}^{21}\binom{21}{k}4^k\cdot 1^{21-k}\right] \\
&= \sum_{k=0}^{21}\binom{21}{k}4^k\cdot\left[(-1)^{21-k}+1^{21-k}\right] \\
&\equiv \sum_{k=0}^{2}\binom{21}{k}4^k\cdot\left[(-1)^{21-k}+1^{21-k}\right] \quad \left(\text{because } 4^k\equiv 0 \mod 64 \text{ for } k\geq 3\right) \\
&\equiv \binom{21}{0}4^0\cdot\left[(-1)^{21}+1^{21}\right]+\binom{21}{1}4^1\cdot\left[(-1)^{20}+1^{20}\right]+\binom{21}{2}4^2\cdot\left[(-1)^{19}+1^{19}\right] \\
&\equiv \binom{21}{1}4^1\cdot\left[(-1)^{20}+1^{20}\right] \\
&\equiv 21\cdot 4\cdot 2\equiv 40 \mod 64.
\end{aligned}
$$

Correct Answer: A

257. Given three concentric circles of radii a, b and c, there exists an equilateral triangle of side lengths equal to x and with vertices on these circles, as in the adjacent figure, if and only if

$$
\begin{gathered}
x^4+a^4+b^4+c^4-x^2a^2-x^2b^2= \\
x^2c^2+a^2b^2+b^2c^2+c^2a^2.
\end{gathered}
$$



If $a=3, b=5$ and $c=7$, there are two solutions for x. One is $x=8$ and the other is given by the equation $x^2=d$. What is the value of d?

A. 17

B. 18

C. 19

D. 20

Solution:

(Method I) Substituting the given values of a, b and c in (1) we get $0=1216-83x^2+x^4$ and after the given factor is used we get $0=\left(x^2-64\right)\left(x^2-19\right)$.

(Method II) The equation is biquadratic and so it factors as $\left(x^2-64\right)\left(x^2-d\right)=0$ this means $64+d=a^2+b^2+c^2=9+25+49$ and so $d=83-64=19$.

Correct Answer: C

258. Suppose $a > 0, b > 0, c > 0$ and $abc = 1$. What is the minimum value of

$$(a+b)^3 + (b+c)^3 + (c+a)^3?$$

A. 8

B. 16

C. 24

D. 32

Solution:

Using the Arithmetic-Geometric Mean Inequality we have that

$$\frac{(a+b)^3 + (b+c)^3 + (c+a)^3}{3} \geq \sqrt[3]{(a+b)^3(b+c)^3(c+a)^3} = (a+b)(b+c)(c+a)$$

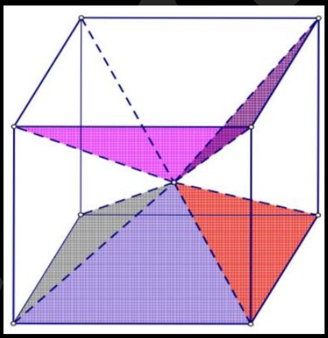
We also have that

$$(a+b)(b+c)(c+a) \geq 2\sqrt{ab} \cdot 2\sqrt{bc} \cdot 2\sqrt{ca} = 8$$

This implies that the minimum value of the original expression is 24 and it occurs when $a = b = c = 1$.

Correct Answer: C

259. A cube is dissected into six pyramids by connecting a given point in the interior of the cube with each vertex of the cube. The volumes of five of these pyramids are $1, 4, 9, 10$ and 13. What is the volume of the sixth pyramid?



A. 4

B. 5

C. 6

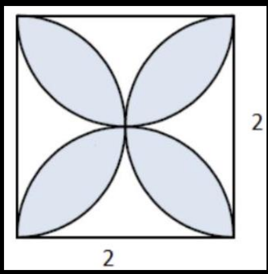
D. 7

Solution:

The sum of the volumes of opposite pyramids is the same since it is $1/3$ of the volume of the cube. This means that the sum of opposite pyramids is 14, and the volume of the 6th pyramid is 5.

Correct Answer: B

260. The figure shows a square and four semicircles generated with each side of the square as a diameter. If the side length of the square is 2, what is the area of the shaded region?



A. $2\pi - 4$

B. $3\pi - 4$

C. $4\pi - 4$

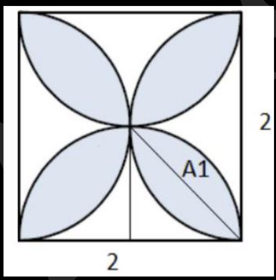
D. None of the above

Solution:

The area of A_1 can be obtained by deducting the area of the triangle from the area of the quarter circle with radius 1.

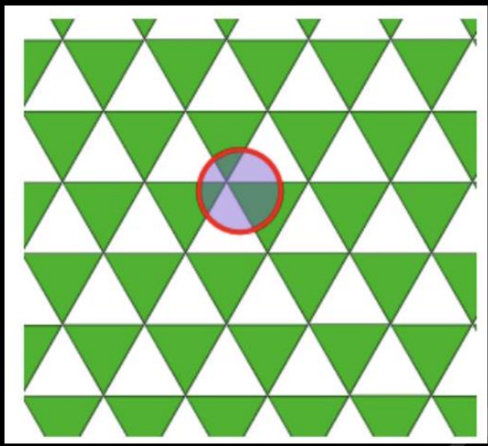
We have $A_1 = \frac{\pi}{4} - \frac{1}{2}$.

All shaded area $= 8 \cdot A_1 = 2\pi - 4$.



Correct Answer: A

261. The figure below shows a tessellation of the plane with equilateral triangles of side length 1. In the experiment of tossing a coin of diameter 1, what is the probability that the coin does not overlap with any triangle vertex?

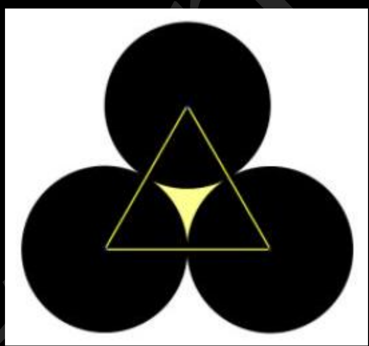


A. $\dfrac{8\sqrt{3}-4\pi}{7\sqrt{3}}$ B. $\dfrac{8\sqrt{3}-4\pi}{8\sqrt{3}}$

C. $\dfrac{8\sqrt{3}-4\pi}{9\sqrt{3}}$ D. $\dfrac{8\sqrt{3}-4\pi}{10\sqrt{3}}$

Solution:

Consider one of the equilateral triangles:



The only way the coin will not overlap with any triangle vertex will happen if the coin lands in the yellow region shown in the diagram above. Therefore, it suffices to determine the area of this yellow region:

The three sectors are each of 60-degree central angle, summing to a 180-degree central angle, i.e., a semicircle. Therefore, the area of the yellow region equals = the area of the equilateral triangle of side length 1 - the area of a semicircle of radius 1/2.

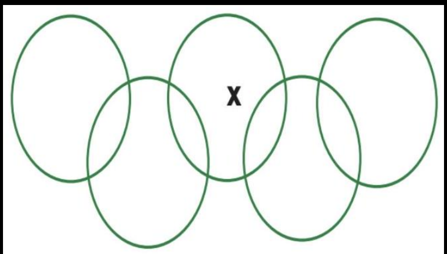
$$= 1^2 \cdot \frac{\sqrt{3}}{4} - \frac{\pi \cdot (1/2)^2}{2} = \frac{\sqrt{3}}{4} - \frac{\pi}{8}$$

Dividing this value by the area of the equilateral triangle will give us the desired probability:

$$P = \frac{\frac{\sqrt{3}}{4} - \frac{\pi}{8}}{\frac{\sqrt{3}}{4}} = 1 - \frac{\pi}{8} \cdot \frac{4}{\sqrt{3}} = \frac{8\sqrt{3}-4\pi}{8\sqrt{3}}$$

Correct Answer: B

262. The five intersecting circles in the figure determine nine regions. In each of these regions, you write one of the numbers from 1 to 9 such that each number is written exactly once and the sum of the numbers inside each circle is 11. Which number must be written in the region marked with an **X**?



A. 7

B. 4

C. 6

D. 3

Solution:

Note that numbers 8 and 9 can only be written inside the regions most to the left or right. Also, note that the numbers $1, 2, 3$, and 4 can only be written in the small regions determined by exactly two intersecting circles.

Using these two observations it follows that only 6 can be written in the region marked by **X**.

Correct Answer: C

263. If the space shuttle is flying with a speed of $17,321$mi/h at an elevation of 126 mi above the equator and the equatorial radius of the Earth is 3963 mi, then how long (to the nearest minute) does it take to complete one revolution?

A. 93 minutes

B. 115 minutes

C. 89 minutes

D. 48 minutes

Solution:

The length of one revolution is $L = 2\pi R$, where R is the radius of the circle described by the shuttle.

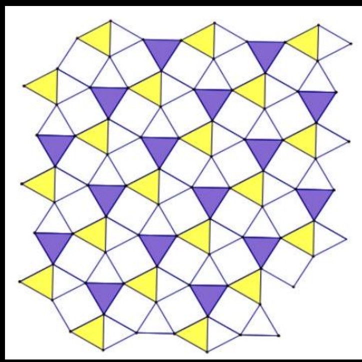
Since the radius of the Earth is 3963 miles and the elevation of the shuttle is 126 miles then $R = 3963 + 126 = 4089$ miles.

Thus $L = 2\pi 4089 \approx 25692$ miles.

The time it takes the shuttle to complete one revolution is approximately $\frac{25692}{17321} \approx 89$ minutes.

Correct Answer: C

264. Consider the tiling of the plane with squares and equilateral triangles as in the figure. If the pattern continues on a very big area, what is the approximate ratio between the number of squares and the number of equilateral triangles used?



A. $\frac{1}{3}$ B. 1

C. $\frac{2}{3}$ D. $\frac{1}{2}$

Solution:

Count the number of equilateral triangles and squares that lie inside a n by n square with the vertices on the sides or the midpoints of the sides of the equilateral triangles (could take n to be the length of a triangle side plus the length of two heights).

If a triangle is not entirely inside the n by n square then combine it with another one with the same property so that you form an equilateral triangle. Because the plane can be covered with such squares and the tiling is identical for every such n by n square we get that the quotient is $\frac{1}{2}$.

Correct Answer: D

265. The statement **"If it is snowing then schools are closed"** is logically equivalent to which of the following?

A. "If it is snowing then schools are not closed."

B. "If it is not snowing then schools are closed."

C. "If it is not snowing then schools are not closed."

D. "If schools are closed then it is snowing."

E. "If schools are not closed then it is not snowing."

Solution:

The statement "If schools are not closed then it is not snowing." is the counterpositive of the statement "If it is snowing then schools are closed" so they are logically equivalent.

Correct Answer: E

266. You have 100 tiles, numbered 0 to 99. Take any set of three tiles. If the number on one of the tiles is the sum of the other two numbers, call the set "good". Otherwise, call the set "bad". How many good sets of three tiles are there?

A. 160

B. 1600

C. 1225

D. 2401

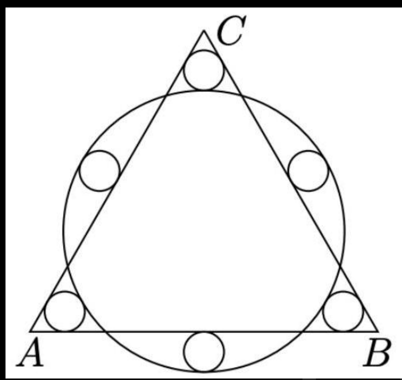
Solution:

Students need to notice a pattern and invent some useful counting techniques in order to solve this problem. For example, they may reason as follows. If 3 is the largest number, there is one good set, $\{1, 2, 3\}$. If 4 is the largest, there is one good set, $\{1, 3, 4\}$. For 5 and 6 there are two, $\{1, 4, 5\}, \{2, 3, 5\}$ and $\{1, 5, 6\}, \{2, 4, 6\}$, respectively. For 7 and 8 there are three each, for 9 and 10 there are four each, and so on. In this way, when we reach 97 there are 48 good sets, for 98 there are also 48, and for 99 there are 49. It remains to compute the sum

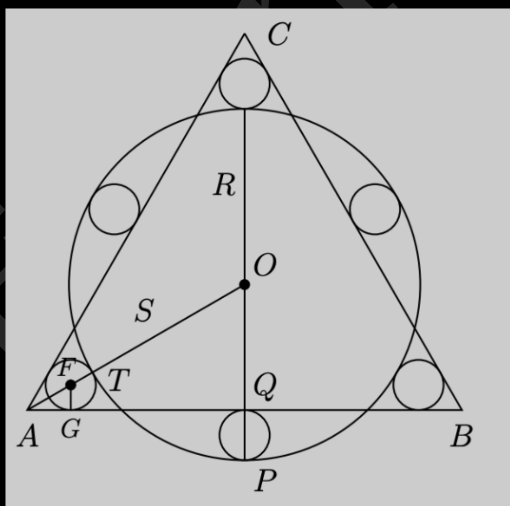
$$\begin{aligned} & 1+1+2+2+3+3+\cdots+47+47+48+48+49 \\ & \quad = 49+(1+48)+(1+48)+(2+47)+(2+47) \\ & \quad +\cdots+(24+25)+(24+25) = 49\times 49 = 2401 \end{aligned}$$

Correct Answer: D

267. Six small circles each have radius of one unit. Each is tangent to a larger circle and to one or two sides of an equilateral triangle $\triangle ABC$, arranged symmetrically, as shown. The radius of the larger circle is ______. (Numerical Answer Type)



Solution:



Let R be the radius of the big circle, O its centre, let $S = OA$, and let P be the point on the big circle directly below O in the diagram.

Let Q be the point of intersection of AB and OP and let T be the point of intersection of OA and the big circle.

Because $\triangle AOQ$ is $30-60-90$, $OQ = \frac{1}{2}S$.

Let F be the center of the small circle between A and O and let G be where that circle meets AB. Then

$$\begin{aligned} & FG = 1 \\ & FT = 1 \\ & AF = 2 \text{ (again by } 30-60-90 \text{ triangle)} \\ \therefore\ & AT = 3. \end{aligned}$$

So we have (considering OA)$S = R + 3$ and (considering OB) $R = \frac{S}{2} + 2$.

Therefore $R = \frac{R+3}{2} + 2$, so $2R = R + 3 + 4 = R + 7$.

So, $R = 7$.

Correct Answer: 7

268. Answer the question based on the information given below.

Literacy rates (in percentages) among various classes in India.

Years	Scheduled Castes (C)	Scheduled Tribes (T)	Other Categories (O)
1961	10.27	8.54	28.30
1971	14.67	11.30	34.45
1981	21.38	16.35	43.57
1991	37.41	29.60	52.21
2001	54.69	47.10	64.80

Note: These three classes account for the whole of India's population.

If none of the given classes constituted more than 50% or less than 25% of India's population in any of the years, in how many of the given years was the number of literates in India at least one-third of the total population?

A. 4

B. 3

C. 2

D. 1

Solution:

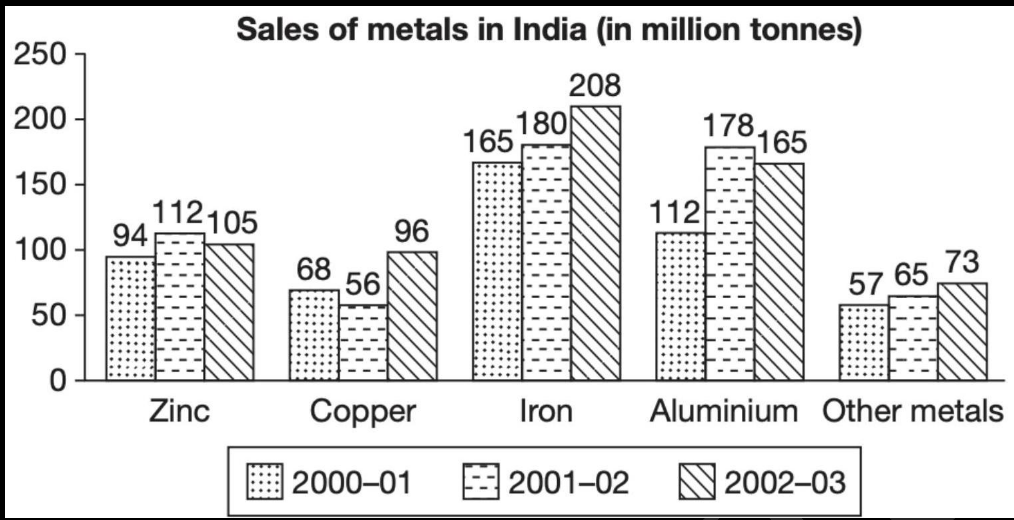
For 1961 and 1971, the number of literates would definitely be less than one-third. While for 1991 and 2001 it would definitely be more than one-third.
$\therefore$ We must check only for 1981.
Considering the best case, the literacy rate in India would be only

$$\begin{aligned}
&= \frac{43.57 \times 50 + 21.38 \times 25 + 16.35 \times 25}{50 + 25 + 25} \\
&= \frac{2178.5 + 534.5 + 408.75}{100} = \frac{3121.75}{100} \\
&= \frac{3121.75}{100} = 31 \times 22\%, \text{ which is less than one-third.}
\end{aligned}$$

$\therefore$ It is at least one-third for only two years.

Correct Answer: C

269. Answer the question based on the information given below.



The sales target for the sales of metals in the year 2002-2003 was 20% more than that of the actual sales in the year 2000-2001. What is the approximate percentage deficit or surplus achieved in the actual sales in the year 2002-2003?

A. 8% deficit B. 9% surplus C. 30% surplus D. 120% deficit

Solution:

Total sales in 2002-03

$$= 105 + 96 + 208 + 165 + 73 = 647 \rightarrow \text{(i)}$$

Total sales in 2000-01 $= 94 + 68 + 165 + 112 + 57 = 496$

Let $647 \cong 650$ and $496 \cong 500$

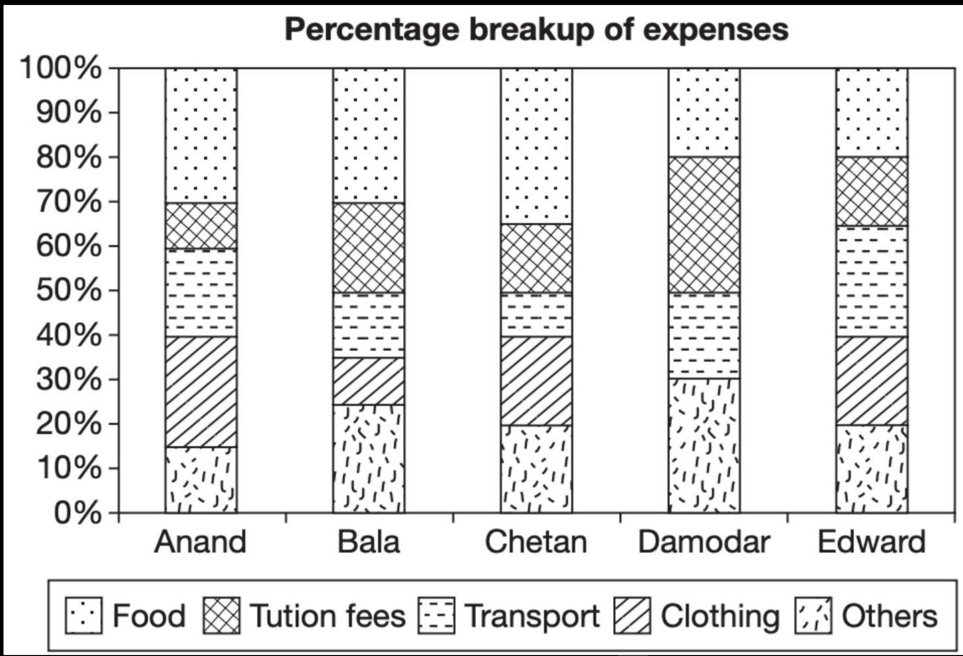
Sales target in 2002-02 $= 1.2 \times 500 = 600$ m tons

Hence, there is surplus in the sales target.

$\therefore$ Required percentage $= \dfrac{650 - 600}{600} \times 100 \cong 9\%$

Correct Answer: B

270. The question are based on the following stack bar.



If Anand's expenses under each of the five heads is not less than that of Chetan's, then the total expenses of Anand is at least how many times that of Chetan's?

A. 1.2 B. 1.5 C. 1.8 D. 2.0

Solution:

10% of Anand's expenses (on clothing) is at least equal to 15% of Chetan's expenses.

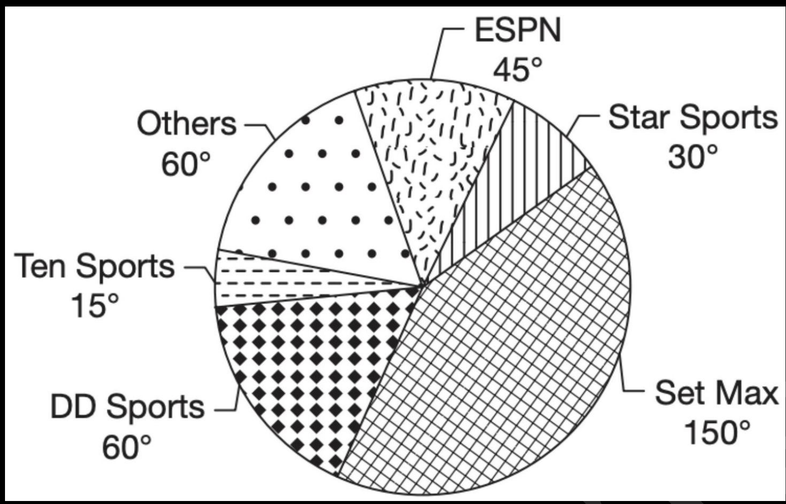
$\therefore$ 10% of A $\geq$ 15% of C

$\therefore \frac{A}{C} \geq 1.5$

$\therefore$ Total expenses of Anand is at least 1.5 times that of Chetan.

Correct Answer: B

271. The question is based on the following pie chart which shows the viewership of different sports channels in the month of February 2003 in India. There is no overlap in viewership of channels.



If the viewership of DD Sports for the first half of February is half that of the second half of February, then what is the ratio of viewership of DD Sports for the second half of February to that of ESPN for the whole month?

A. 8 : 9 B. 4 : 3 C. 2 : 3 D. 9 : 8

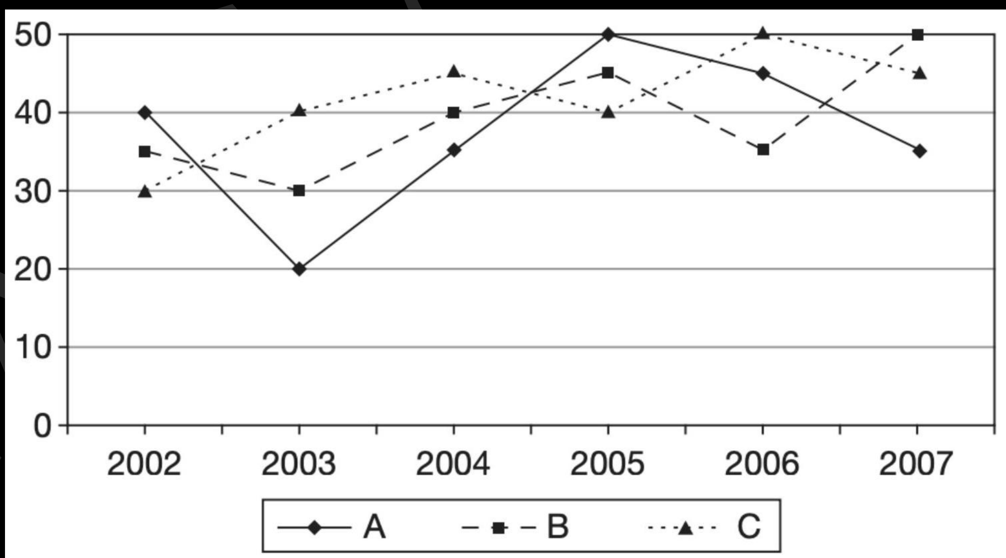
Solution:

Viewership of DD Sports during second half of February $= 60 \times \frac{2}{3} = 40°$

$\therefore$ Required ratio = 40° : 45° = 8 : 9

Correct Answer: A

272. The question is based on the following graph which shows the profit percentage earned on three products A, B, and C over the years 2002 through 2007.



Profit = Selling Price – Cost price

$$\text{Profit percentage} = \frac{\text{Profit}}{\text{Cost price}} \times 100$$

Selling price of product B in 2004 and 2005 were equal. What was the ratio of its cost prices in 2004 to that of 2005?

A. 29 : 28 B. 25 : 27 C. 36 : 35 D. 33 : 25

Solution:

Let the S.P. of B in 2004 and 2005 be 100.

Therefore,

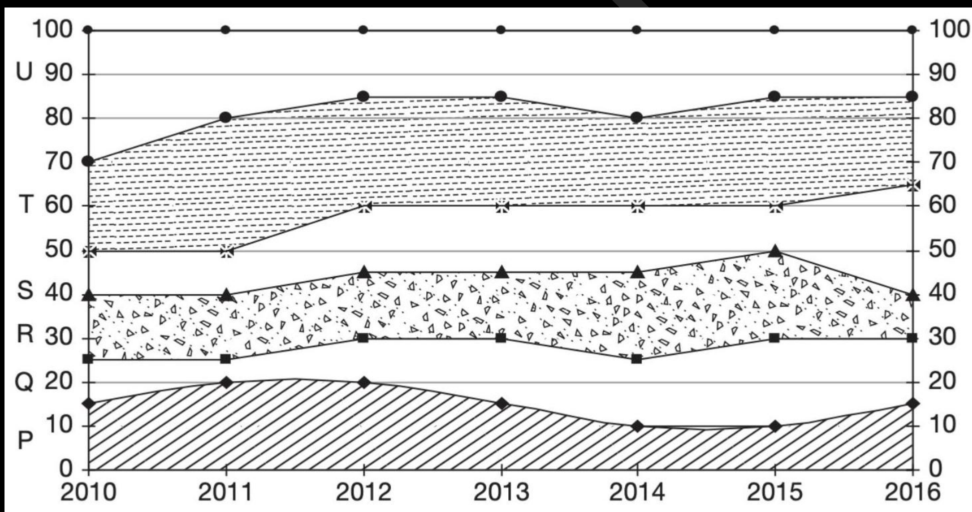
$$\text{C.P in 2004} = \frac{160}{1.4}$$
$$\text{C.P in 2005} = \frac{100}{1.45}$$

$$\therefore \text{ Required ratio } = \frac{100}{1.4} : \frac{100}{1.45}$$

$$= 1.45 : 1.4 = 29 : 28$$

Correct Answer: A

273. Answer the question on the basis of the information given below. A survey was done regarding the number of mobile phone subscribers of six companies in a country over the period 2010 to 2016. The graph gives the percentage break up of the number of subscribers of these six companies P, Q, R, S, T and U. It is observed that, each year, the total number of subscribers of these six companies increased by 25% over the previous year.



What is the percentage increase in the number of subscribers of company P from 2011 to 2016?

A. 110% B. 121% C. 129% D. 137%

Solution:

Let the total subscribers in 2011 be 100 .

The total subscribers in 2016 would be 305.

The subscribers of company P in 2011 = 20

Subscribers of company P in 2016 would be $\frac{15}{100} \times 305 = 45.75$

The percentage increase in subscribers = $\frac{25.75}{20} \times 100 == 128.75\%$

Correct Answer: C

274. Answer the question based on the information given below.

In Indian Public School (IPS), 80% of the students who appeared for the Class X Board exams in 2014 passed the exam. Among these who passed the board exams, 60% joined Indian Intermediate college (IIC) for their Class XI. The students of IPS who joined IIC opted for Science, Commerce and Humanities streams in the ratio 3 : 4 : 5. Therefore, 60% of students in Science stream, 40% of students in Commerce stream and 50% of students in Humanities stream of Class XI in IIC happen to be students of IPS who passed in 2014. The total number of students in all the three streams of Class XI in IIC is 800. Each student opts for only one stream.

How many students failed in the board exams in IPS in 2014?

A. 160

B. 120

C. 200

D. 240

Solution:

Let the total number of students who appeared for the board exams in IPS be N.

0.8 N passed the board exams while 0.2 N failed. Of the 0.8 N who passed, 60% or 0.48 N joined IIC.

The 0.48 N students opted for Science, Commerce and Humanities in the ratio of 3 : 4 : 5.

$\therefore$ 0.12 N from IPS opted for Science, 0.16 N for IPS opted for Commerce and 0.2 N from IPS opted for Humanities. 60% of Science stream = 0.12 N

$\Rightarrow$ Total Science stream = 0.2 N

Similarly, total Commerce stream = 0.4 N

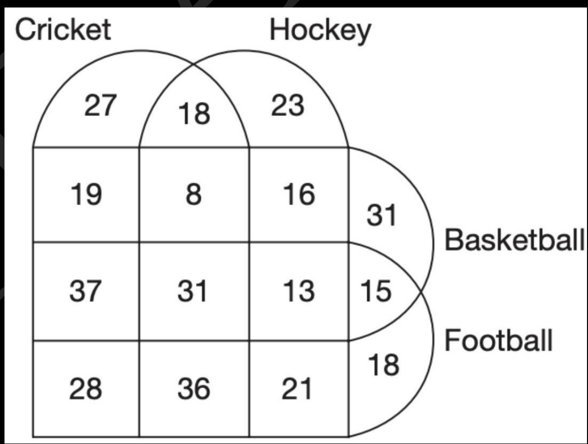
Total Humanities stream = 0.4 N

Given, Science + Commerce + Humanities = 0.2 N + 0.4 N + 0.4 N = N = 800

Students who failed = 0.2 N = 160

Correct Answer: A

275. Answer the question based on the information given below. The given figure provides the details of the number of students in a school who play any of the four games, such as cricket, basketball, football, and hockey.



The total number of students in the class is 500.

How many of the students play at most one game?

A. 99 B. 60 C. 159 D. 258

Solution:

The number of students playing at least one game = 341.

$\therefore$ The number of students playing none of the games = 159

The number of students playing exactly one game = 27 + 23 + 31 + 18 = 99

$\therefore$ The number of students playing at most one game = 159 + 99 = 258

Correct Answer: D

– Love From GATE And Tech –

– Love From Lakshman Patel –

Total 275 Questions