@RayanAzmoon

کد کنترل

334

E

334E

نام:

نام خانوادگی:

محل امضا:

«اگر دانشگاه اصلاح شود مملکت اصلاح می‌شود.»
امام خمینی (ره)

جمهوری اسلامی ایران
وزارت علوم، تحقیقات و فنّاوری
سازمان سنجش آموزش کشور

صبح جمعه
۱۳۹۶/۱۲/۴
دفترچهٔ شماره (۱)

آزمون ورودی دورهٔ دکتری (نیمه‌متمرکز) ـ سال ۱۳۹۷

رشتهٔ مهندسی فناوری اطلاعات (کد ۲۳۵۸)

تعداد سؤال: ۴۵

مدت پاسخگویی: ۱۵۰ دقیقه

عنوان مواد امتحانی، تعداد و شمارهٔ سؤالات

ردیف	مواد امتحانی	تعداد سؤال	از شمارهٔ	تا شمارهٔ
۱	مجموعه دروس تخصصی: حل مسئله (ساختمان داده‌ها و طراحی الگوریتم‌ها) ـ شبکه (شبکه‌های کامپیوتری و امنیت شبکه) ـ مهندسی اطلاعات (پایگاه داده پیشرفته، بازیابی پیشرفته اطلاعات و داده‌کاوی) ـ خدمات فناوری اطلاعات (تجارت الکترونیک و آموزش الکترونیکی)	۴۵	۱	۴۵

این آزمون نمرهٔ منفی دارد.

استفاده از ماشین‌حساب مجاز نیست.

حق چاپ، تکثیر و انتشار سؤالات به هر روش (الکترونیکی و ...) پس از برگزاری آزمون، برای تمامی اشخاص حقیقی و حقوقی تنها با مجوز این سازمان مجاز می‌باشد و با متخلفین برابر مقررات رفتار می‌شود.

* داوطلب گرامی، عدم درج مشخصات و امضا در مندرجات جدول ذیل، به‌منزلهٔ عدم حضور شما در جلسهٔ آزمون است.

اینجانب ..... با شمارهٔ داوطلبی ..... در جلسهٔ این آزمون شرکت می‌نمایم.

امضا:

۱- الگوریتمی را در نظر بگیرید که ورودی $a_1,...,a_n$ شامل n عدد مجزا را به‌ترتیب داده شده می‌خواند و هنگام خواندن a_i مقدار متغیر x را به احتمال $1/i$ برابر a_i قرار می‌دهد. الگوریتم در پایان مقدار x را به‌عنوان خروجی گزارش می‌کند. با چه احتمالی خروجی الگوریتم برابر a_i است؟

۱) می‌تواند هر مقداری کوچکتر یا مساوی $1/i$ باشد.

۲) می‌تواند هر مقداری در بازه $[1/n, 1/i]$ باشد.

۳) $1/i$

۴) $1/n$

۲- در گراف جهت‌دار G وزن یال‌ها را یک قرار می‌دهیم و شار بیشینه از رأس s به t را محاسبه می‌کنیم. مقدار جریان بیشینه برابر کدام مورد است؟

۱) تعداد مسیرهای بین s و t

۲) تعداد کوتاه‌ترین مسیرهای بین s و t

۳) تعداد مسیرهای مجزای یالی بین s و t

۴) تعداد مسیرهای مجزای رأسی بین s و t

۳- چندتا از گزاره‌های زیر درست است؟

- اگر مسئله تصمیم‌گیری X در ان‌پی باشد، مسئله تصمیم‌گیری not X نیز در ان‌پی است.
- هر مسئله ان‌پی ـ سخت به یک مسئله‌ی ان‌پی ـ کامل قابل کاهش است.
- تمام مسائل ان‌پی ـ کامل به تمام مسائل ان‌پی ـ سخت قابل کاهش‌اند.

۱) ۰

۲) ۱

۳) ۲

۴) ۳

۴- فرض کنید گراف جهت‌دار G شامل n رأس و m یال، روابط دوستی بین n فرد را مدل می‌کند. به عبارت دقیق‌تر از u به v یال وجود دارد، اگر شخص u شخص v را بشناسد. در این شبکه هرگاه شخص از یک خبر مطلع شود آن را به اطلاع همهٔ دوستان خود می‌رساند. می‌خواهیم یک خبر مشخص را به اطلاع همه برسانیم. می‌خواهیم کمترین تعداد افرادی را پیدا کنیم که با مطلع شدن آن‌ها، همه از خبر فوق مطلع شوند. بهترین الگوریتم برای محاسبه این مجموعه افراد از چه مرتبه‌ای است؟

۱) $O(m+n)$

۲) $O(\min(n,m))$

۳) $O(n\log n+m)$

۴) این مسئله ان‌پی ـ سخت است.

۵- در جست و جوی سطح اول (BFS)، به هر رأس یک بازهٔ زمانی نسبت می‌دهیم، طوری که زمان قراردادن رأس در صف شروع بازه و زمان برداشتن آن از صف انتهای بازه باشد. کدام مورد درست است؟

۱) ترتیب طول بازه‌های بین رئوس، معادل با ترتیب فاصله‌ی آنان از رأس شروع است.

۲) اگر یک رأس از نوادگان رأس دیگر در درخت BFS باشد، آنگاه بازه‌ی آن‌ها اشتراک ندارد.

۳) اگر بازه‌ی دو رأس اشتراک نداشته باشند، آن‌گاه یکی از آن‌ها نوه‌ی دیگری در درخت BFS است.

۴) اگر شروع بازه‌ی b بعد از شروع بازه‌ی a باشد، فاصلهٔ رأس شروع تا b بیشتر از فاصلهٔ رأس شروع تا a است.

۶- فرض کنید $T=(V,E)$ یک درخت شامل n رأس باشد. می‌خواهیم $C\subseteq V$ با حداقل تعداد رأس پیدا کنیم، طوری‌که برای هر یال در E، حداقل یکی از دو سر آن یال در C باشد. الگوریتم حریصانه زیر را در نظر بگیرید. C را در ابتدا تهی قرار می‌دهیم. در هر مرحله رأس با درجه بیشینه در T را در C قرار می‌دهیم (درصورتی‌که بیش از یک رأس با درجه بیشینه بود، یکی از آن‌ها را به دلخواه انتخاب می‌کنیم) و تمام یال‌های مجاور آن را از T حذف می‌کنیم. این کار را تا زمانی‌که T تهی از یال شود ادامه می‌دهیم. کمترین مقدار n که برای آن، الگوریتم حریصانه فوق درست کار نمی‌کند، کدام است؟

۱) ۳

۲) ۴

۳) ۵

۴) برای هر n الگوریتم حریصانه جواب بهینه را برمی‌گرداند.

۷- گراف $G=(V,E)$ با وزن‌های مثبت را در نظر بگیرید. $d(u,v,k)$ را برابر طول کوتاه‌ترین مسیر از u به v در نظر بگیرید که حداکثر k یال داشته باشد. چندتا از رابطه‌های بازگشتی زیر برای $k>1$ درست می‌باشند؟
($w=(u,v)$ وزن یال (u,v) را نشان می‌دهد و وزن یالی که وجود نداشته باشد بی‌نهایت در نظر گرفته می‌شود.)

$$d(u,v,k)=\min_{x\in V}(d(u,v,k-1),d(u,x,k-1)+w(x,v))$$

$$d(u,v,k)=\min_{x\in V}(d(u,x,k-1)+w(x,v))$$

$$d(u,v,k)=\min_{x\in V}(d(u,v,k-1),d(u,x,\lceil k/2\rceil)+d(x,v,\lfloor k/2\rfloor))$$

۱) ۳

۲) ۲

۳) ۱

۴) ۰

۸- در یک گراف همبند با n رأس و m یال، وزن یال‌ها ۱ یا ۲ هستند. بهترین الگوریتم برای یافتن درخت پوشای کمینه این گراف دارای چه زمان اجرایی است؟

۱) $O(n\log n+m)$

۲) $O(n\log n)$

۳) $O(n)$

۴) $O(m)$

@RayanAzmoon

۹- سعید و وحید بازی زیر را انجام می‌دهند. ابتدا دو بازیکن روی یک عدد n توافق می‌کنند. بازی با $x=2$ آغاز می‌شود. بازیکن اول با انداختن سکه انتخاب شده و بازی را آغاز می‌کند. از این به بعد هرکدام از بازیکنان به دلخواه x را به توان۲ یا ۳ می‌رساند. مثلاً اگر سعید بازی را آغاز کرده و ۳ را انتخاب کند، x برابر با ۸ خواهد شد. بعد از حرکت هر کدام از بازیکنان که $x > n$ شود، بازی متوقف شده و بازیکن آخر به‌عنوان برنده بازی انتخاب می‌شود. بازی‌ها به‌طور متوسط در چند مرحله به پایان می‌رسند؟

۱) $\Theta(\log \log n)$

۲) $\Theta(\log^2 n)$

۳) $\Theta(\log n)$

۴) قابل محاسبه نیست.

۱۰- می‌خواهیم ماتریس‌های $M_1(5\times10)$ و $M_2(10\times3)$ و $M_3(3\times12)$ و $M_4(12\times5)$ را با همین ترتیب در هم ضرب کنیم. این کار با حداقل چند ضرب عددی قابل انجام است؟

۱) ۳۶۰

۲) ۴۰۵

۳) ۵۸۰

۴) ۶۳۰

۱۱- کدام جملات در خصوص فرایند دریافت IP از طریق پروتکل DHCP درست است؟ (فرض کنید آدرس IP سرور DHCP برابر 80.66.180.1 باشد و آدرس IP که قرار است تخصیص داده شود 80.66.180.20 باشد.)

الف. آدرس IP مقصد بسته‌های discover و offer برابر 255.255.255.255 است.

ب. آدرس IP مقصد بسته‌های request و ack برابر 255.255.255.255 است.

پ. آدرس IP مقصد بسته request برابر 80.66.180.1 و آدرس مقصد بسته ack برابر 80.66.180.20 است.

ت. آدرس IP مقصد بسته request برابر 255.255.255.255 و آدرس مقصد بسته ack برابر 80.66.180.20 است.

ث. آدرس IP مبدا بسته ack برابر 80.66.180.1 و آدرس مقصد بسته ack برابر 80.66.180.20 است.

۱) الف و پ ۲) ت و ث

۳) الف و ب ۴) الف و ت و ث

۱۲- می‌خواهیم یک فایل ۱۵ مگابایتی را ارسال کنیم. مبدأ و مقصد از طریق سه لینک و دو سوئیچ به یکدیگر متصل هستند. زمان ارسال از طریق هر کدام از روش‌های circuit switch(cs)، virtual circuit switch(vc) و datagram packet switch(ps) به‌ترتیب از راست به چپ چند میلی‌ثانیه است؟

(اندازه بسته‌های داده را ۱/۵ مگابایت و اندازه بسته‌های کنترلی برای برقراری و قطع تماس را در صورت نیاز ۱۵ کیلوبایت در نظر بگیرید. در ضمن تأخیر انتشار در هر لینک ۱۰ میلی‌ثانیه، و پهنای باند هر لینک 15×10^6 بیت بر ثانیه می باشند.)

۱) ۸۱۴۴ ، ۸۱۶۲ ، ۹۰۳۰ ۲) ۸۱۶۰ ، ۹۷۹۲ ، ۹۶۳۰

۳) ۸۱۴۴ ، ۸۱۹۲ ، ۸۰۳۰ ۴) ۸۱۳۰ ، ۹۷۶۲ ، ۹۶۰۰

۱۳- در TCP رنو در هر کدام از حالات زیر اندازه پنجره ازدحام (**cwnd**) برابر کدام مورد است؟

الف. دریافت یک **ack** تکراری در حالت **fast recovery**

ب. مشاهده سه **ack** تکراری در حالت **slow start**

پ. مشاهده یک **ack** جدید در حالت **congestion avoidance**

۱) الف. cwnd+MSS، ب. cwnd/2 + 3*MSS، پ. cwnd+MSS*MSS/cwnd

۲) الف. cwnd، ب. cwnd+3*MSS، پ. cwnd+MSS/cwnd

۳) الف. cwnd+MSS/cwnd، ب. ssthresh، پ. cwnd+MSS

۴) الف. cwnd، ب. cwnd/2، پ. cwnd+MSS

۱۴- فرض کنید دو گره در قالب یک نشست **voice-over-IP** از طریق یک مسیر که شامل ۴ مسیریاب است، به همدیگر پیغام ارسال می‌کنند. پهنای باند تمام لینک‌ها **۱ Mbps** است و فاصله مکانی دو گره از یکدیگر **۳۰۰۰ km** است. تمام بسته‌های ارسالی شامل داده اصلی و **ACK**، ۱۵۰۰ بایت حجم دارند. فرض کنید که سرعت انتشار برابر $۲\times۱۰^{۸}\ \frac{m}{s}$ است و اندازه طول صف در هر کدام از مسیریاب‌ها حداکثر ۵ بسته است و زمانی که فرستنده هر بسته‌ای را که برای اولین بار ارسال می‌کند،آن بسته ممکن است با احتمال ۰/۵ تلف شود، اما ارسال دوم هر بسته حتماً موفقیت‌آمیز است. بافر **playback** باید چند کیلوبایت باشد تا هیچ وقت خالی نشود و تجربه کاربر دچار مشکل نشود؟

(فرض کنید $1Mbps = 10^6 bit/s$ ،$1KB = 1024B$ و **Retransmission Timeout =RTT** باشد.)

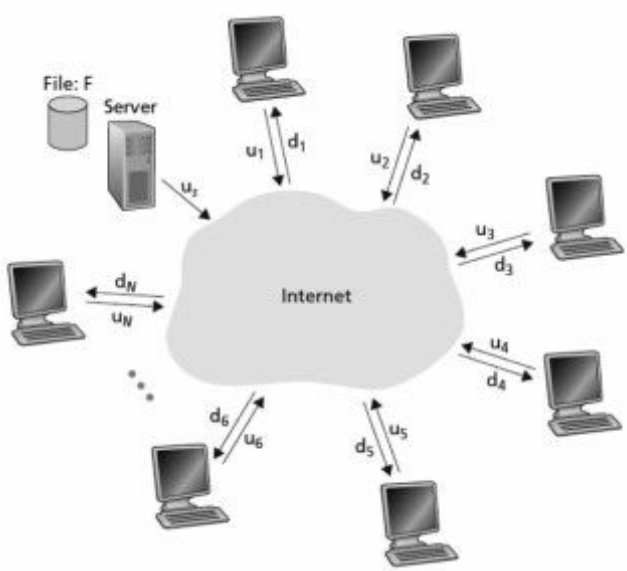
۱) ۸۷۰

۲) ۱۰۶

۳) ۶۷

۴) ۲۹

۱۵- شبکه **Peer-to-peer** شکل زیر را در نظر بگیرید. **N** گره علاقه دارند یک فایل با حجم **F** بایت که در ابتدا بر روی سرور قرار دارد را دریافت کنند. درصورتی که فرض شود اطلاعات به‌صورت بایتی منتقل می‌شود، کمترین زمان برای دریافت اطلاعات توسط کلیه گره‌ها کدام است؟ (فرض کنید سرعت دانلود تمام گره‌ها بسیار بالاست.)



۱) F / u_s

۲) $NF / (u_s + u_1 + ... + u_N)$

۳) $F / \min(u_s, (u_1 + ... + u_N)/N)$

۴) $F / \min(u_s, (u_s + u_1 + ... + u_N)/N)$

۱۶- درصورتی‌که E تابع رمزنگاری، Sign تابع امضاء دیجیتال، K_A^+ و K_B^+ به‌ترتیب کلید عمومی A و B و K_A^- و K_B^- به‌ترتیب کلید خصوصی A و B بوده و M متن آشکار باشد، کدام مورد برای حفظ محرمانگی در زمان ارسال پست الکترونیکی و پردازش کم مناسب است؟

۱) ارسال $E_{K_B^+}(M)$ از فرستنده A به گیرنده B

۲) ارسال $Sign_{K_A^-}(M)$ و $E_{K_B^+}(M)$ از فرستنده A به گیرنده B

۳) ارسال $(K_B^+)Sign_{K_A^-}$ و $E_{K_B^+}(M)$ از فرستنده A به گیرنده B

۴) ارسال $(E_{K_B^+}(M))Sign_{K_A^-}$ و $E_{K_B^+}(M)$ از فرستنده A به گیرنده B

۱۷- در صورتی‌که امکان جعل آدرس فرستنده در بسته IP وجود نداشته باشد، کدام حمله هنوز بسیار خطرناک است؟

۱) DOS ۲) smurf

۳) DDOS ۴) هرسه حمله

۱۸- کدام عبارت(ها)، درست است؟

الف) IDS هایی که براساس امضا کار می‌کنند، سرعت تشخیص خوبی دارند ولی خطای منفی زیادی دارند.

ب) IDS هایی که براساس تشخیص ناهنجاری کار می‌کنند، اگر همراه با یک IDS مبتنی بر امضاء استفاده شوند، خطای منفی زیادی دارند.

ج) در صورت استفاده از HIDS در همه سرورها نیازی به استفاده از NIDS در آن شبکه نیست.

۱) الف ۲) ب

۳) ج ۴) الف و ب

۱۹- کدام‌یک از پروتکل‌های امنیتی زیر، سرویس محرمانگی پیام را ارائه نمی‌دهند؟

۱) TLS ۲) AH

۳) ESP ۴) PGP

۲۰- علت استفاده از توابع چکیده‌ساز همراه با امضای دیجیتال، کدام است؟

۱) افزودن محرمانگی به سرویس احراز هویت ارائه شده توسط امضای دیجیتال

۲) مقاومت بیشتر در برابر حملات جعل امضا

۳) جلوگیری از حمله مردمیانی

۴) سربار محاسباتی کمتر

۲۱- کدام مورد درخصوص طرح‌های توالی‌پذیر، درست است؟

۱) هر طرح توالی‌پذیر نتیجه‌ای، توالی‌پذیر تعارضی و توالی‌پذیر نمایی نیز است.

۲) هر طرح توالی‌پذیر تعارضی، توالی‌پذیر نمایی و توالی‌پذیر نتیجه‌ای نیز است.

۳) هر طرح توالی‌پذیر نمایی، توالی‌پذیر تعارضی و توالی‌پذیر نتیجه‌ای نیز است.

۴) هر طرح توالی‌پذیر نمایی، توالی‌پذیر تعارضی نیز است ولی توالی‌پذیر نتیجه‌ای نیست.

۲۲- در صورتی که در طرح زیر از پروتکل چندنسخه‌سازی مبتنی بر زمان‌مهر استفاده شود، کدام مورد درست است؟

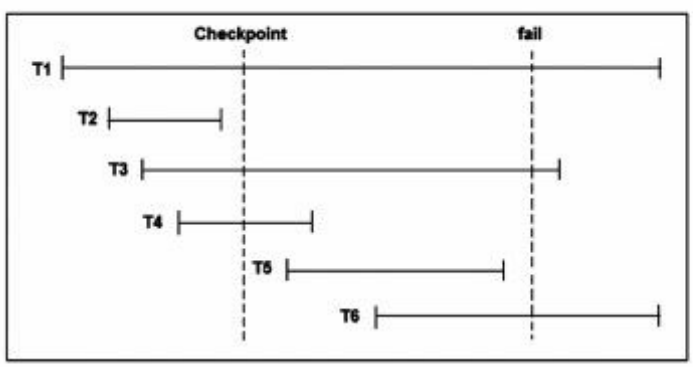
T۱	T۲	T۳
$TS(T۱)=۱۰۰$	$TS(T۱)=۲۰۰$	$TS(T۳)=۳۰۰$
	Read(Z)	
	Read (Y)	
	Write(Y)	
		Read(Y)
Read(X)		Read(Z)
Write(X)		
		Write(Y)
		Write(Z)
	Read(X)	
Read(Y)		
Write(Y)		
	Write(X)	

۱) این طرح توالی‌پذیر نیست و در پایان اجرای آن از X سه نسخه، Y سه نسخه و Z دو نسخه وجود دارد.
۲) این طرح توالی‌پذیر نیست و در پایان اجرای آن از X دو نسخه، Y دو نسخه و Z یک نسخه وجود دارد.
۳) این طرح توالی‌پذیر است و در پایان اجرای آن از X سه نسخه، Y سه نسخه و Z دو نسخه وجود دارد.
۴) این طرح توالی‌پذیر است و در پایان اجرای آن از X دو نسخه، Y دو نسخه و Z یک نسخه وجود دارد.

۲۳- کدام مورد درخصوص پروتکل‌های قفل‌گذاری، درست نیست؟

۱) در پروتکل قصد قفل‌گذاری، قفل IX با خودش سازگار است.
۲) پروتکل قفل‌گذاری جنگلی، توالی‌پذیری را تضمین نمی‌کند.
۳) پروتکل درختی، عدم وجود طرد تسلسلی و ترمیم‌پذیر بودن طرح را تضمین می‌کند.
۴) عملیات قفل‌برداری در پروتکل درختی می‌تواند نسبت به پروتکل 2PL زودتر انجام شود.

۲۴- در طرح اجرای زیر چناچه نقطه وارسی از نوع commit consistent checkpoint باشد و از immediate modification استفاده شده باشد، کدام مورد درست است؟



۱) فقط تراکنش‌های T1، T3 و T6 باید undo شوند.
۲) تراکنش‌های T1 و T3 باید undo و تراکنش‌های T4 و T5 باید redo شوند.
۳) تراکنش‌های T1 و T3 باید undo و تراکنش‌های T2 و T4 باید redo شوند.
۴) تراکنش‌های T1، T3 و T6 باید undo و تراکنش‌های T4 و T5 باید redo شوند.

۲۵- کدام اعمال، موجب انتقال **Log Record** ها از حافظه اصلی به حافظه پایدار (**Stable Storage**) نمی‌شود؟

۱) Commit شدن تراکنش
۲) انتقال بلوک داده تغییریافته به دیسک
۳) انجام Checkpoint
۴) تغییر یک بلوک داده در حافظه اصلی

۲۶- در طرح اجرای $R_1(B), W_1(A), R_2(A), R_3(A),\ldots, W_1(B), W_3(B)$ ، چه دستوری در جای خالی باید قرار بگیرد تا این طرح اجرا، توالی‌پذیر تعارضی نباشد، ولی توالی‌پذیر نمایی باشد؟

۱) $W_2(B)$ یا $R_2(B)$
۲) $W_2(A)$ یا $R_2(A)$
۳) $W_2(B)$
۴) $W_2(A)$

۲۷- یک پرس‌وجو را در نظر بگیرید که در کل چهار سند مرتبط با آن در مجموعه اسناد موجود است. سیستم بازیابی اطلاعات در پاسخ به این پرس‌وجو لیست مرتب شده‌ای از ۱۰ سند را برگردانده است. این ۱۰ سند از نظر ارتباط با پرس‌وجو قضاوت شده‌اند و وضعیت مرتبط بودن آن‌ها در جدول مقابل آمده است. "+" در ستون «ارتباط» به این معنی است که سند بازیابی شده مرتبط با پرس‌وجو است و "-" به معنی نامرتبط بودن سند به پرس‌وجو می‌باشد. **Average Precision** این پرس‌وجو، کدام است؟

رتبه	ارتباط
۱	-
۲	+
۳	-
۴	+
۵	+
۶	-
۷	-
۸	-
۹	-
۱۰	-

۱) $\frac{۲}{۵}$
۲) $\frac{۸}{۱۵}$
۳) $\frac{۳}{۱۰}$
۴) $\frac{۳}{۴}$

۲۸- در مدل فضای برداری (**vector space model**)، کدام مورد در مقایسه با سایر موردها انتخاب مناسبی برای بسامد ترم (**term frequency – TF**) نیست؟ (**c(w, d)** تعداد تکرار ترم **w** در سند **d** است.)

۱) $1+\log(1+\log(c(w,d)))$
۲) $0.5+0.5\frac{c(w,d)}{\max_{w'\in d} c(w',d)}$
۳) $1+\log(c(w,d))$
۴) $2^{c(w,d)}-1$

۲۹- روش بازخورد **Rocchio** در مدل فضای برداری (**vector space model**) در کدام شرایط خوب عمل نمی‌کند؟

۱) برای پرس‌وجویی که تعداد سندهای مرتبط آن زیاد است.
۲) برای پرس‌وجویی که تعداد سندهای نامرتبط آن زیاد است.
۳) برای پرس‌وجویی که پرس‌وجو بسیار نزدیک به مرکز مجموعه سندهای مرتبط به پرس‌وجو است.
۴) برای پرس‌وجویی که سندهای مرتبط در فضا حداقل دو زیرخوشه با مرکزهای مختلف تشکیل می‌دهند.

۳۰- در روش هموارسازی **Absolute Discounting**، از فرمول زیر برای هموارسازی احتمال وقوع ترم **w** در سند **d** استفاده می‌کند، که در آن **c(w,d)** تعداد تکرار ترم **w** در سند **d**، δ پارامتر هموارسازی، $|d|_u$ تعداد ترم‌های یکتای سند **d**، **p(w|C)** احتمال وقوع ترم **w** در مجموعه مرجع **C** و $|d|$ طول سند **d** است. مجموعه مقادیر مجاز برای پارامتر δ، کدام است؟

$$p(w|d)=\frac{\max\left(c(w,d)-\delta,0\right)+\delta|d|_u\,p(w|C)}{|d|}$$

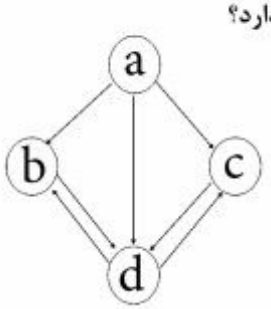
۱) $\delta\in[-1,1]$

۲) $\delta\in[0,1]$

۳) $\delta\geq|d|_u$

۴) $\delta\geq 0$

۳۱- چهار صفحه **d**، **c**، **b**، **a** و پیوندهای بین آن‌ها را به شکل گراف زیر در نظر بگیرید. الگوریتم **HITS** (**Hypertext-Induced Topic Search**) را بر روی این گراف اجرا می‌کنیم و برای هر صفحه دو امتیاز **hub** و **authority** محاسبه می‌کنیم. کدام صفحه بالاترین امتیاز **hub** در الگوریتم **HITS** را دارد؟



۱) a

۲) b

۳) c

۴) d

۳۲- مجموعه تراکنش‌های زیر را در نظر بگیرید، می‌خواهیم با استفاده از الگوریتم **Apriori**، با شرط **minimum support= ۳**، **itemset**های مکرر را پیدا کنیم. الگوریتم **Apriori** برای پیداکردن همه **itemset**های مکرر این مجموعه تراکنش‌ها، چند بار کل داده‌ها را اسکن می‌کند؟

TID	Items
۱	A, B
۲	A, C, D, E
۳	B, C, E, F, G
۴	A, B, C, E
۵	A, B, C, F

۱) ۱

۲) ۲

۳) ۳

۴) ۴

۳۳- چهار شئ **A, B, C, D** را در نظر بگیرید. می‌خواهیم با استفاده از الگوریتم **Agglomerative Hierarchical Clustering** این چهار شئ را خوشه‌بندی کنیم. اطلاعات شباهت اشیاء در جدول زیر آمده است. برای محاسبه شباهت گروهی اشیاء از روش **single-link** استفاده می‌کنیم. میزان شباهت بین دو خوشه {A , B} و {C , D} با استفاده از روش **single-link** کدام است؟

	A	B	C	D
A	۱	۰/۸۵	۰/۷	۰/۳
B	۰/۸۵	۱	۰/۸	۰/۶
C	۰/۷	۰/۸	۱	۰/۹
D	۰/۳	۰/۶	۰/۹	۱

۱) ۰/۲

۲) ۰/۳

۳) ۰/۶

۴) ۰/۸

۳۴- یک مکعب داده‌ای (data cube) پنج بعدی با معیار اندازه‌گیری count را در نظر بگیرید که در آن شکل کلی سلول‌ها به صورت count : (a۱,a۲,a۳,a۴,a۵) است.این مکعب شامل دو سلول به صورت پایه ۲۰ : (A,B,C,D,E) و ۱۰ : (A,F,C,G,H) است. با فرض اینکه A≠F ، B≠G و D≠H و E≠H، می‌خواهیم iceburg cube با شرط count ≥ 20 را محاسبه کنیم. تعداد سلول‌های این iceburg cube (که شرایط مدنظر count ≥ 20 را دارند) کدام است؟

۱) ۱۶

۲) ۳۲

۳) ۶۰

۴) ۶۴

۳۵- برای ارزیابی کیفیت یک روش خوشه‌بندی از معیار خارجی خلوص (purity) استفاده می‌کنیم. فرض کنید نتیجه خوشه‌بندی داده‌ها توسط خبره، دسته‌های T۱،T۲،T۳،T۴ است. از طرف دیگر داده‌ها توسط روشی که می‌خواهیم ارزیابی کنیم خوشه‌بندی شده‌اند که نتیجه خوشه‌های C۱ ، C۲ ، C۳ ، C۴ است. توزیع داده‌ها در دسته‌ها و خوشه‌ها در جدول زیر آمده است. c(Ci , Tj) تعداد اقلامی از دسته Tj را نشان می‌دهد که توسط روش خوشه‌بندی در خوشه Ci قرار گرفته‌اند. با توجه به اطلاعات جدول، معیار purity برای این خوشه‌بندی کدام است؟

	T۱	T۲	T۳	T۴
C۱	۰	۱۵	۳۰	۵
C۲	۱۰	۱۰	۵	۰
C۳	۰	۵	۰	۰
C۴	۰	۰	۵	۱۵

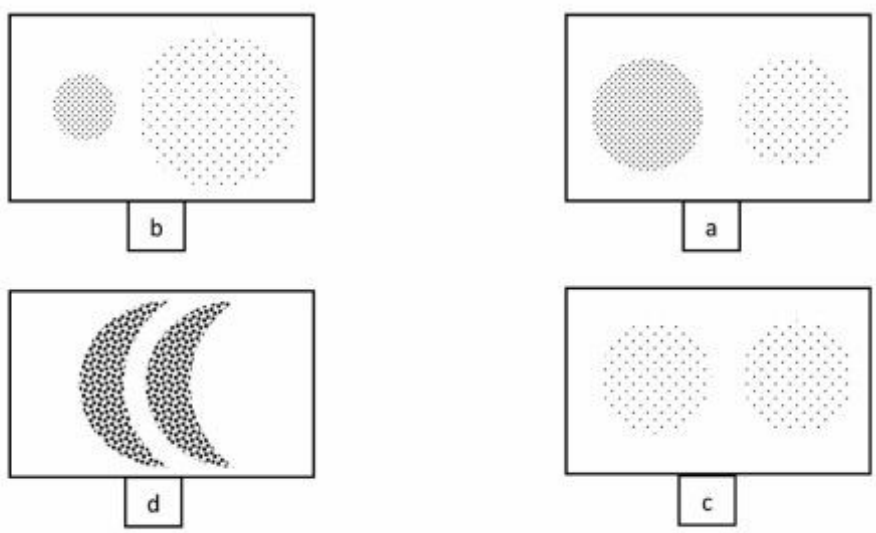
۱) ۰/۶

۲) ۰/۷

۳) ۰/۶۸

۴) ۰/۷۵

۳۶- چهار شکل a ، b، c، d را در نظر بگیرید، که هر کدام نشان دهنده مجموعه نقاطی است که قرار است خوشه‌بندی شوند. در مورد هر شکل می‌خواهیم داده‌ها را به دو خوشه تقسیم کنیم. در کدام شکل(ها) استفاده از الگوریتم k-means برای خوشه‌بندی مناسب است؟



۱) فقط c ۲) a, c ۳) a, b, c ۴) a, b, c, d

۳۷- تجارت الکترونیک جزئی (Partial EC)، به کدام مثال از کسب‌وکار و یا تجارب الکترونیکی اطلاق می‌شود؟

۱) سامانه اینترنتی تبلیغات و جستجوی اتومبیل

۲) کسب‌وکار مبتنی بر انجام تبلیغات دیجیتال درون بازی‌های تلفن همراه

۳) سامانه انتقال واحد مجازی بیت‌کوین از یک بانک یا حساب مجازی به بانک یا حساب مجازی دیگر

۴) کسب‌وکاری که مبتنی بر اینترنت و وب سایت سفارش پخت و پز غذای خانگی دریافت می‌کند و غذای سفارش شده را با پیک موتوری تحویل سفارش دهنده بدهد.

۳۸- آسیب درک شده (perceived damage) از رقابت بین عمده فروشان یک محصول و یا خدمات و فروش مستقیم برخط همان محصول توسط تولید کننده، باعث کدام مورد می‌شود؟

۱) تعارض در کانال‌های توزیع فروش الکترونیکی (channel conflict) و عدم ادامه همکاری

۲) حذف واسطه‌ها در up stream زنجیره تأمین برای تولید آن کالا یا خدمات

۳) رشد نوآوری در فرایندها و استفاده از روش‌های D2C توسط توزیع‌کنندگان

۴) عدم پذیرش تجارت الکترونیک B2B توسط مصرف‌کنندگان

۳۹- در مدیریت زنجیره تامین برای جلوگیری از Bullwhip effect، کدام روش استفاده می‌شود؟

۱) استفاده گسترده از RFID برای ردیابی جریان کالا و بهبود جریان شفاف اطلاعات به طول زنجیره تامین

۲) ارائه تخفیف در قیمت برای خریدهای انبوه توسط مشتریان نهایی

۳) ایجاد انبار در مقاطع مختلف زنجیره تامین

۴) حذف واسطه‌ها

۴۰- کدام مورد با چالش‌های به‌کارگیری تحلیل جریان (clickstream analysis)، یک کاربر سایت تجارت الکترونیکی مرتبط نیست؟

۱) بهبود کاربردپذیری تجربه کاربری

۲) پیچیدگی محاسباتی مرتبط با تخمین احتمال خرید کاربر از سایت تجارت الکترونیکی

۳) پردازش داده‌های حجیم برای استخراج داده (data mining) جهت ایجاد پروفایل برای کاربر

۴) نگرانی کاربر از جانب حفظ محرمانگی داده‌های او با توجه به وجود امکان تشخیص هویت کاربر

۴۱- کدام عامل باعث می‌شود تا یک ابزار یادگیری برخط به عنوان ابزار غیرهمزمان (asynchronous)، دسته بندی شود؟

۱) عدم امکان حضور و مشارکت هم‌زمان یادگیرندگان و استادان در یک نشست

۲) عدم امکان حضور و مشارکت هم‌زمان یادگیرندگان و استادان در چند نشست

۳) وجود تراکنش‌های توزیع شده سازگار (consistent)

۴) وجود تاخیر در شبکه اینترنت حین ارسال منابع درسی

۴۲- سناریوی زیر چه سطحی از یادگیری را براساس طبقه‌بندی بلوم می‌سنجد؟

«یادگیرنده یک صفحه html را مشاهده می‌کند که در آن مفاهیم میانگین و میانه تعریف شده است. سپس با کلیک روی گزینهٔ «صفحهٔ بعد» به صفحه‌ای هدایت می‌شود که در آن ۱۴ عدد در یک ستون نوشته شده و از کاربر خواسته شده است که میانگین و میانهٔ این اعداد را حساب نموده و در دو جای مشخص شده وارد نماید.»

۱) کاربرد ۲) خلاقیت ۳) تحلیل ۴) دانش

۴۳- برای ارزیابی تاثیر فناوری بر یادگیری برخط، از کدام معیار (metrics) استفاده می‌شود؟

۱) تعداد منابع و موارد بازدید شده و یا اتمام یافته در یک بازه آموزشی

۲) دقت سامانه‌های توصیه‌گر (recommender systems)

۳) کیفیت سرویس (QoS) شبکه‌های آموزشی

۴) کاربردپذیری پورتال‌ها

۴۴- کدام مورد از چالش‌های یادگیری اینترنتی به صورت synchronous می‌باشد و با استفاده از منابع آموزشی چندرسانه‌ای است؟

۱) یافتن محتوای آموزشی مورد نیاز

۲) کنترل دسترسی و حفظ یکپارچگی نمرات

۳) غیرهمگون بودن سطح اطلاعات و پیشینه آموزشی افراد یادگیرنده در یک کلاس مجازی

۴) به اشتراک گذاشتن دانش با سایر افراد و تشکیل گروه‌های آموزشی (group formation)

۴۵- روش‌های توصیه مبتنی بر محتوا (content-based recommendation) برای تخمین متون و منابع آموزشی مرتبط و مورد نیاز کاربر، به چه اطلاعاتی نیاز دارند؟

۱) خصوصیات محتوا (content features) و تاریخچه استفاده سایر یادگیرندگان از محتوای مد نظر

۲) پروفایل کاربر و خصوصیات محتوا (content features)

۳) مدل‌های دانش و منطق فازی

۴) چارت امتحانات