

کد کنترل



334

E

محل امضای:

نام: نام خانوادگی:

صبح جمعه
۱۳۹۶/۱۲/۴

دفترچه شماره (۱)

«اگر دانشگاه اصلاح شود مملکت اصلاح می‌شود.»
امام خمینی (ره)جمهوری اسلامی ایران
وزارت علوم، تحقیقات و فناوری
سازمان سنجش آموزش کشور

آزمون ورودی دوره دکتری (نیمه‌تمدد) - سال ۱۳۹۷

رشته مهندسی فناوری اطلاعات (کد ۲۳۵۸)

مدت پاسخگویی: ۱۵۰ دقیقه

تعداد سوال: ۴۵

عنوان مواد امتحانی، تعداد و شماره سوالات

ردیف	مواد امتحانی	تعداد سوال	از شماره	تا شماره
۱	مجموعه دروس تخصصی: حل مسئله (ساختمان داده‌ها و طراحی الگوریتم‌ها) - شبکه‌های کامپیوتري و امنیت شبکه) - مهندسی اطلاعات (پایگاه داده پیشرفته، بازیابی پیشرفته اطلاعات و داده کاوی) - خدمات فناوری اطلاعات (تجارت الکترونیک و آموزش الکترونیکی)	۴۵	۱	۴۵

استفاده از ماشین حساب مجاز نیست.

این آزمون نمره متفقی دارد.

حق جایه تکبر و انتشار سوالات به هر روش (الکترونیک و...) بس از برگزاری آزمون، برای تمام اشخاص حقیقی و حقوقی تنها با مجوز این سازمان مجاز می‌باشد و با منظمهن برای غفران و فثار می‌شود.

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

اینجانب با شماره داوطلبی در جلسه این آزمون شرکت می‌نمایم.

امضا:

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

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

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

(۳) $1/i$

(۴) $1/n$

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

(۱) تعداد مسیرهای بین s و t

(۲) تعداد کوتاه‌ترین مسیرهای بین s و t

(۳) تعداد مسیرهای مجازی بین s و t

(۴) تعداد مسیرهای مجازی رأسی بین s و t

چندتا از گزاره‌های زیر درست است؟

-۳

• اگر مسئله تصمیم‌گیری X در ان بی باشد، مسئله تصمیم‌گیری $\text{not } X$ نیز در ان بی است.

• هر مسئله ان بی - سخت به یک مسئله ان بی - کامل قابل کاهش است.

• تمام مسائل ان بی - کامل به تمام مسائل ان بی - سخت قابل کاهش‌اند.

(۱)

(۲)

(۳)

(۴)

-۴ فرض کنید گراف جهت دار G شامل n رأس و m یال، روابط دوستی بین n فرد را مدل می‌کند. به عبارت دقیق‌تر از n به v یال وجود دارد. اگر شخص n شخص v را بشناسد. در این شبکه هرگاه شخص از یک خبر مطلع شود آن را به اطلاع همه 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 مرتباً است؟

$O(\min(n, m))$ (۲)

$O(m + n)$ (۱)

(۴) این مسئله ان بی - سخت است.

$O(n \log n + m)$ (۳)

-۵ در جست و جوی سطح اول (BFS)، به هر رأس یک بازه زمانی نسبت می‌دهیم، طوری که زمان قراردادن رأس در صفحه شروع بازه و زمان برداشتن آن از صفحه 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 باشد، آن‌گاه یکی از آن‌ها نوهدی دیگری در درخت BFS است.

(۴) اگر شروع بازه‌ی b بعد از شروع بازه‌ی a باشد، فاصله رأس شروع تا b بیشتر از فاصله 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،

طوری که 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 (درصورتی که

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 زمانی که 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 حداقل k یال داشته باشد. چندتا از رابطه‌های بازگشته‌ی زیر برای $k > 1$ درست می‌باشند؟

(۱) وزن یال $w = (u, v)$ 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؟

$$O(n \log n + m) \quad (1)$$

$$O(n \log n) \quad (2)$$

$$O(n) \quad (3)$$

$$O(m) \quad (4)$$

- ۹ سعید و حیدر 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 هر کدام از بازیکنان به دلخواه X را به توان ۲ یا ۳ می‌رساند. مثلًا 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 به عنوان برنده بازی انتخاب می‌شود. بازی‌ها به طور 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 $M_1(5 \times 10)$ و $M_2(10 \times 3)$ و $M_3(3 \times 12)$ و $M_4(12 \times 5)$ 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 ۱۸۰.۶۶.۸۰.۸۰ باشد و آدرس IP که قرار است تخصیص داده شود ۲۰.۶۶.۸۰.۸۰ باشد).
- الف. آدرس IP مقصد بسته‌های discover و offer برابر ۲۵۵.۲۵۵.۲۵۵ است.
 ب. آدرس IP مقصد بسته‌های request و ack برابر ۲۵۵.۲۵۵.۲۵۵ است.
 پ. آدرس IP مقصد بسته request برابر ۸۰.۶۶.۱۸۰.۱ و آدرس مقصد بسته ack برابر ۸۰.۶۶.۱۸۰.۲۰ است.
 ت. آدرس IP مقصد بسته request برابر ۲۵۵.۲۵۵.۲۵۵ و آدرس مقصد بسته ack برابر ۲۰.۶۶.۱۸۰.۲۰ است.
 ث. آدرس IP مبدأ بسته ack برابر ۸۰.۶۶.۱۸۰.۱ و آدرس مقصد بسته ack برابر ۸۰.۶۶.۱۸۰.۲۰ است.
- (۱) الف و پ
 (۲) ت و ث
 (۳) الف و ت و ث
 (۴) الف و ب
- ۱۲ می‌خواهیم یک فایل ۱۵ مگابایتی را ارسال کنیم. مبدأ و مقصد از طریق سه لینک و دو سوییج به یکدیگر متصل هستند. زمان ارسال از طریق هر کدام از روش‌های virtual circuit switch(vc)، circuit switch(es) و datagram packet switch(ps) به ترتیب 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 15×10^9 بیت بر ثانیه می‌باشد).
- | | |
|----------------------|----------------------|
| (۱) ۹۰۳۰، ۸۱۶۲، ۸۱۴۴ | (۲) ۹۷۹۲، ۸۱۶۰، ۹۶۳۰ |
| (۳) ۸۰۳۰، ۸۱۹۲، ۸۱۴۴ | (۴) ۸۱۳۰، ۹۷۶۲، ۹۶۰۰ |

-۱۳ در TCP رنو در هر کدام از حالات زیر اندازه پنجه 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 جدید در حالت congestion avoidance

(۱) الف. $cwnd + MSS * MSS / cwnd$, ب. $cwnd / 2 + 3 * MSS$

(۲) الف. $cwnd + 3 * MSS / cwnd$, ب. $cwnd + MSS$

(۳) الف. $cwnd + MSS / ssthresh$, ب. $cwnd + MSS / cwnd$

(۴) الف. $cwnd + MSS / cwnd / 2$, ب. $cwnd / 2$

-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 بايت حجم دارند. فرض کنید که سرعت انتشار برابر $\frac{m}{s} 2 \times 10^8$ است و اندازه طول صف در هر کدام از مسیریاب‌ها حداقل ۵ بسته است و زمانی که فرستنده هر بسته‌ای را که برای اولین بار ارسال می‌کند، آن بسته ممکن است با احتمال $5/5$ تلف شود، اما ارسال دوم هر بسته حتماً موفقیت‌آمیز است. بافر playback باید چند کیلوبايت باشد تا هیچ وقت خالی نشود و تجربه کاربر دچار مشکل نشود؟

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

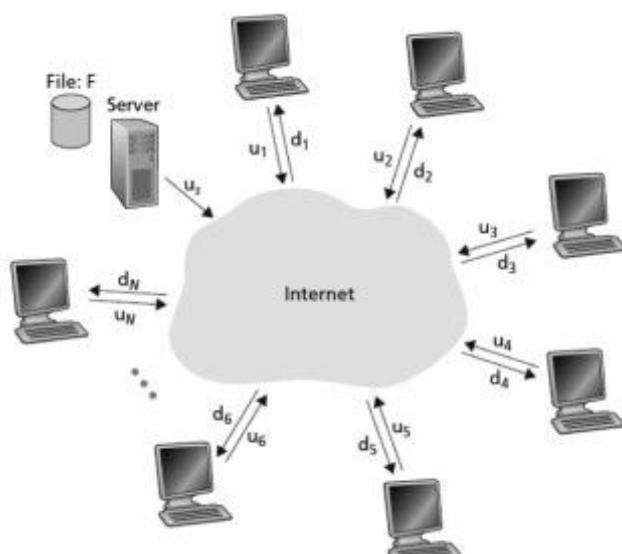
(۱) ۸۷۰

(۲) ۱۰۶

(۳) ۶۷

(۴) ۲۹

-۱۵ شبکه Peer-to-peer شکل زیر را در نظر بگیرید. N گره علاوه دارند یک فایل با حجم F بایت که در ابتدا بر روی سرور قرار دارد را دریافت کنند. درصورتی که فرض شود اطلاعات به صورت 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) $NF / (u_s + u_l + \dots + u_N)$

(۳) $F / \min(u_s, (u_l + \dots + u_N) / N)$

(۴) $F / \min(u_s, (u_s + u_l + \dots + u_N) / N)$

- ۱۶ در صورتی که E تابع رمزگاری، Sign تابع امضاء دیجیتال، K_A^+ و K_B^+ به ترتیب کلید عمومی A و B و K_A^- و K_B^- به ترتیب کلید خصوصی A و B بوده و M متن آشکار باشد، کدام مورد برای حفظ محbermanگی در زمان ارسال پست الکترونیکی و پردازش کم مناسب است؟
- (۱) ارسال $E_{K_B^+}(M)$ از فرستنده A به گیرنده B
 - (۲) ارسال $E_{K_B^+}(M)$ و $\text{Sign}_{K_A^-}(M)$ از فرستنده A به گیرنده B
 - (۳) ارسال $E_{K_B^+}(M)$ و $\text{Sign}_{K_A^-}(E_{K_B^+}(M))$ از فرستنده A به گیرنده B
- ۱۷ در صورتی که امکان جعل آدرس فرستنده در بسته IP وجود نداشته باشد، کدام حمله هنوز بسیار خطرناک است؟
- (۱) smurf
 - (۲) DOS
 - (۳) DDOS
 - (۴) هراسه حمله
- ۱۸ کدام عبارت(ها)، درست است؟
- الف) IDS هایی که براساس امضای 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 کدام یک از پروتکلهای امنیتی زیر، سرویس محbermanگی پیام را ارائه نمی‌دهند؟
- (۱) AH
 - (۲) TLS
 - (۳) PGP
 - (۴) ESP
- ۲۰ علت استفاده از توابع چکیده‌ساز همراه با امضای دیجیتال، کدام است؟
- (۱) افزودن محbermanگی به سرویس احراز هویت ارائه شده توسط امضای دیجیتال
 - (۲) مقاومت بیشتر در برابر حملات جعل امضای
 - (۳) جلوگیری از حمله مردمیانی
 - (۴) سربار محاسباتی کمتر
- ۲۱ کدام مورد در خصوص 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 زمان مهر استفاده شود، کدام مورد درست است؟

T ₁ TS(T ₁) = 100	T ₂ TS(T ₁) = 200	T ₃ TS(T ₂) = 300
Read(X) Write(X)	Read(Z) Read(Y) Write(Y)	Read(Y) Read(Z)
Read(Y) Write(Y)	Read(X)	Write(Y) Write(Z)
		Write(X)

(۱) این طرح توالی‌پذیر نیست و در پایان اجرای آن از X سه نسخه، Y 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 در خصوص پروتکل‌های قفل‌گذاری، درست نیست؟

(۱) در پروتکل قصد قفل‌گذاری، قفل IX با خودش سازگار است.

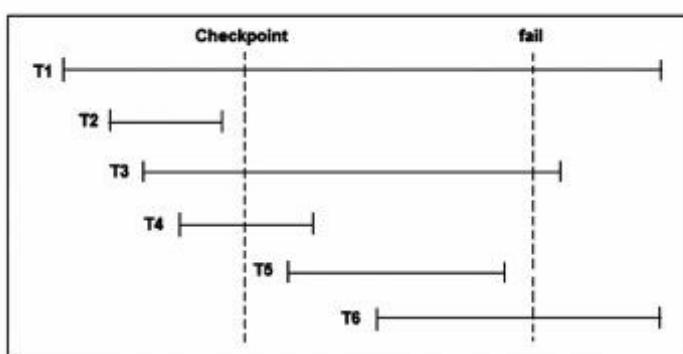
(۲) پروتکل قفل‌گذاری جنگلی، توالی‌پذیری را تضمین نمی‌کند.

(۳) پروتکل درختی، عدم وجود طرد تسلسی 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 چنانچه نقطه وارسی از نوع **commit consistent checkpoint** باشد و از **immediate modification**

استفاده شده باشد، کدام مورد درست است؟



(۱) فقط تراکنش‌های T1، T3، T5 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 و T3 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) انتقال بلوک داده تغییریافته به دیسک
- (۲) تغییر یک بلوک داده در حافظه اصلی
- (۳) انجام Commit
- (۴) Checkpoint

- ۲۶ - در طرح اجرای $R_1(B), R_2(A), R_3(A), \dots, W_1(B), W_2(B), \dots$ ، چه دستوری در جای خالی باید قرار بگیرد تا این طرح اجرا، توالی‌بذیر تعارضی نباشد، ولی توالی‌بذیر نمایی باشد؟

- (۱) $R_2(B)$ یا $W_2(B)$
- (۲) $R_2(A)$ یا $W_2(A)$
- (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 قضاؤت 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 نامرتب بودن سند به پرس‌وجو می‌باشد. Average Precision این پرس‌وجو، کدام است؟

ارتباط	رتبه	
-	۱	$\frac{2}{5}$ (۱)
+	۲	$\frac{8}{15}$ (۲)
-	۳	$\frac{3}{10}$ (۳)
+	۴	$\frac{2}{4}$ (۴)
+	۵	
-	۶	
-	۷	
-	۸	
-	۹	
-	۱۰	

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

$$\text{TF}(w, d) = \frac{c(w, d)}{\max_{w' \in d} c(w', d)} \quad (۱)$$

$$\text{TF}(w, d) = 1 + \log(1 + \log(c(w, d))) \quad (۲)$$

$$\text{TF}(w, d) = 1 + \log(c(w, d)) \quad (۳)$$

- 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 نامرتب 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؟

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

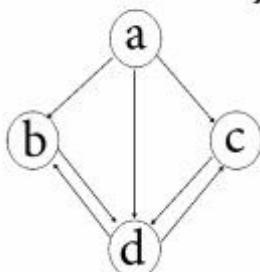
$\delta \in [-1, 1]$ (۱)

$\delta \in [0, 1]$ (۲)

$\delta \geq |d|_u$ (۳)

$\delta \geq 0$ (۴)

- ۳۱ چهار صفحه a ، b ، c 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؟



a (۱)

b (۲)

c (۳)

d (۴)

- ۳۲ مجموعه تراکنش‌های زیر را در نظر بگیرید، می‌خواهیم با استفاده از الگوریتم Apriori، با شرط itemset .minimum support=۳ مکرر را پیدا کنیم. الگوریتم Apriori برای پیداکردن همه itemset‌های 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 Clustering استفاده از اشیاء از روشن **single-link** استفاده می‌کنیم. میزان شباهت این دو خوشه $\{A, B\}$ و $\{C, D\}$ را با شباهت گروهی اشیاء از روشن **single-link** استفاده می‌کنیم. میزان شباهت بین دو خوشه $\{A, B\}$ و $\{C, D\}$ را با استفاده از روشن **single-link** کدام است؟

	A	B	C	D
A	1	0.85	0.7	0.3
B	0.85	1	0.8	0.6
C	0.7	0.8	1	0.9
D	0.3	0.6	0.9	1

۰/۲ (۱)

۰/۳ (۲)

۰/۶ (۳)

۰/۸ (۴)

- ۳۴- یک مکعب داده‌ای (data cube) پنج بعدی با معیار اندازه‌گیری count را در نظر بگیرید که در آن شکل کلی سلول‌ها به صورت $(a_1, a_2, a_3, a_4, a_5) : count$ است. این مکعب شامل دو سلول به صورت پایه است. با فرض اینکه $(A, F, C, G, H) : 10$ و $(A, B, C, D, E) : 20$ می‌خواهیم iceburg cube با شرط $count \geq 20$ را محاسبه کنیم. تعداد سلول‌های این iceburg cube (که شرایط مدنظر count ≥ 20 را دارند) کدام است؟

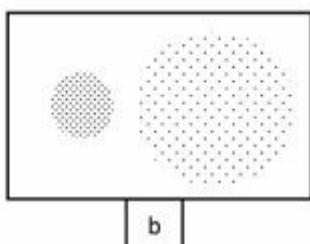
- (۱) ۱۶
 (۲) ۳۲
 (۳) ۶۰
 (۴) ۶۴

- ۳۵- برای ارزیابی کیفیت یک روش خوشبندی از معیار خارجی خلوص (purity) استفاده می‌کنیم. فرض کنید نتیجه خوشبندی داده‌ها توسط خبره، دسته‌های T_1, T_2, T_3, T_4 است. از طرف دیگر داده‌ها توسط روشی که می‌خواهیم ارزیابی کنیم خوشبندی شده‌اند که نتیجه خوشبندی C_1, C_2, C_3, C_4 است. توزیع داده‌ها در دسته‌ها و خوشبندی در جدول زیر آمده است. (c(C_i, T_j) تعداد اقلامی از دسته T_j را نشان می‌دهد که توسط روش خوشبندی در خوشبندی Ci قرار گرفته‌اند. با توجه به اطلاعات جدول، معیار purity برای این خوشبندی کدام است؟

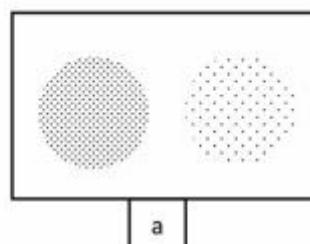
	T ₁	T ₂	T ₃	T ₄
C ₁	۰	۱۵	۳۰	۵
C ₂	۱۰	۱۰	۵	۰
C ₃	۰	۵	۰	۰
C ₄	۰	۰	۵	۱۵

- (۱) ۰/۶
 (۲) ۰/۷
 (۳) ۰/۶۸
 (۴) ۰/۷۵

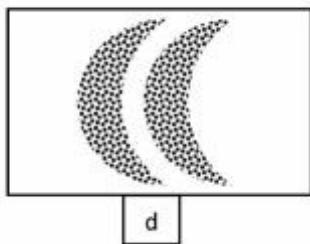
- ۳۶- چهار شکل a, b, c, d را در نظر بگیرید، که هر کدام نشان دهنده مجموعه نقاطی است که قرار است خوشبندی شوند. در مورد هر شکل می‌خواهیم داده‌ها را به دو خوشبندی تقسیم کنیم. در کدام شکل(ها) استفاده از الگوریتم k-means برای خوشبندی مناسب است؟



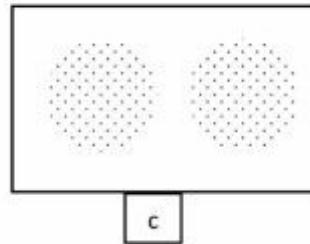
b



a



d



c

a, b, c, d (۴)

a, b, c (۳)

a, c (۲)

c فقط (۱)

-۳۷- تجارت الکترونیک جزئی (Partial EC)، به کدام مثال از کسبوکار و یا تجارت 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 عمدۀ فروشان یک محصول و یا خدمات و فروش مستقیم بر خط 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 توزیع کنندگان

(۴) عدم پذیرش تجارت الکترونیک B2B توسط مصرف‌کنندگان

-۳۹- در مدیریت زنجیره تأمین 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 تأمین

(۲) ارائه تخفیف در قیمت برای خریدهای اتباه توسط مشتریان نهایی

(۳) ایجاد انبار در مقاطع مختلف زنجیره تأمین

(۴) حذف واسطه‌ها

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

(۱) بهبود کاربرد پذیری تجربه کاربری

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

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

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

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

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

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

(۳) وجود تراکنش‌های توزیع شده سازگار (consistent)

(۴) وجود تأخیر در شبکه اینترنت حین ارسال منابع درسی

-۴۲- ستاربیوی 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 گزینه «صفحه بعد» به صفحه‌ای هدایت می‌شود که در آن ۱۴ عدد در یک ستون نوشته شده و از کاربر

خواسته شده است که میانگین و میانه این اعداد را حساب نموده و در دو جای مشخص شده وارد نماید.»

(۴) دانش

(۳) تحلیل

(۲) خلاقیت

(۱) کاربرد

۴۳- برای ارزیابی تأثیر فناوری بر یادگیری برشط، 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 تخمين متون و منابع آموزشی

مرتبه و مورد نیاز کاربر، به چه اطلاعاتی نیاز دارند؟

(۱) خصوصیات محتوا (content features) و تاریخچه استفاده سایر یادگیرنده‌گان از محتوای مدنظر

(۲) پروفایل کاربر و خصوصیات محتوا (content features)

(۳) مدل‌های دانش و منطق فازی

(۴) چارت امتحانات