



قدم به قدم، همراه دانشجو...

WWW.GhadamYar.Com

جامع ترین و به روز ترین پرتال دانشجویی کشور (پرتال دانش)
با ارائه خدمات رایگان، تحصیلی، آموزشی، رفاهی، شغلی و...
برای دانشجویان

- (۱) راهنمای ارتقاء تحصیلی. (کاردانی به کارشناسی، کارشناسی به ارشد و ارشد به دکتری)
- (۲) ارائه سوالات کنکور مقاطع مختلف سالهای گذشته، همراه پاسخ، به صورت رایگان
- (۳) معرفی روش‌های مقاله و پایان نامه نویسی و ارائه پکیج‌های آموزشی مربوطه
- (۴) معرفی منابع و کتب مرتبط با کنکورهای تحصیلی (کاردانی تا دکتری)
- (۵) معرفی آموزشگاه‌ها و مراکز مشاوره تحصیلی معتبر
- (۶) ارائه جزوایت و منابع رایگان مرتبط با رشته‌های تحصیلی
- (۷) راهنمای آزمون‌های حقوقی به همراه دفترچه سوالات سالهای گذشته (رایگان)
- (۸) راهنمای آزمون‌های نظام مهندسی به همراه دفترچه سوالات سالهای گذشته (رایگان)
- (۹) آخرین اخبار دانشجویی، در همه مقاطع، از خبرگزاری‌های پربازدید
- (۱۰) معرفی مراکز ورزشی، تفریحی و فروشگاه‌های دارای تخفیف دانشجویی
- (۱۱) معرفی همایش‌ها، کنفرانس‌ها و نمایشگاه‌های ویژه دانشجویی
- (۱۲) ارائه اطلاعات مربوط به بورسیه و تحصیل در خارج و معرفی شرکت‌های معتبر مربوطه
- (۱۳) معرفی مسائل و قوانین مربوط به سرگذری، معافیت تحصیلی و امریه
- (۱۴) ارائه خدمات خاص ویژه دانشجویان خارجی
- (۱۵) معرفی انواع بیمه‌های دانشجویی دارای تخفیف
- (۱۶) صفحه ویژه نقل و انتقالات دانشجویی
- (۱۷) صفحه ویژه ارائه شغل‌های پاره وقت، اخبار استخدامی
- (۱۸) معرفی خوابگاه‌های دانشجویی معتبر
- (۱۹) دانلود رایگان نرم افزار و اپلیکیشن‌های تحصیلی و...
- (۲۰) ارائه راهکارهای کارآفرینی، استارت آپ و...
- (۲۱) معرفی مراکز تایپ، ترجمه، پرینت، صحافی و ... به صورت آنلاین
- (۲۲) راهنمای خرید آنلاین ارزی و معرفی شرکت‌های مطرح (۲۳)



WWW.GhadamYar.Ir

۰۹۱۲ ۳۰ ۹۰ ۱۰۸

WWW.PortaleDanesh.com

باما همراه باشید...

WWW.GhadamYar.Org

۰۹۱۲ ۰۹ ۰۲ ۸۰۱

کد کنترل

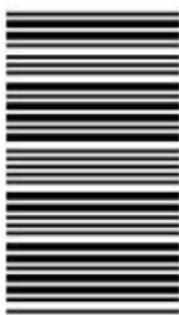
334

E

نام:

نام خانوادگی:

محل امضا:



334E

صبح جمعه
۱۳۹۶/۱۲/۴
دفترچه شماره (۱)



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

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

آزمون ورودی دوره دکتری (نیمه‌تمتر کز) - 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 کامپیوتری و امنیت شبکه — مهندسی اطلاعات (پایگاه داده پیشرفته، بازیابی پیشرفته اطلاعات و داده‌گاوی) — خدمات فناوری اطلاعات (تجارت الکترونیک و آموزش الکترونیکی)	۴۵	۱	۴۵

استفاده از ماشین حساب مجاز نیست.

این آزمون نمرة منفی دارد.

حق جانب تکرار و انتشار سوالات به هر روش (الکترونیک و...) بس از برگزاری آزمون، برای تمام اندیشنهای حقش و حقوقی تنها با مجوز این سازمان مجاز می‌باشد و با مخالفین برگزار غیرقانونی شود.

* 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 i مقدار متغیر x را به احتمال $1/n$ 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 $1/n$ باشد.

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

$1/i$ (۳)

$1/n$ (۴)

-۲ در گراف جهت‌دار 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

-۳ چندتا از گزاره‌های زیر درست است؟

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

* هر مسئله ان پی - سخت به یک مسئله ان پی - کامل قابل کاهش است.

* تمام مسائل ان پی - کامل به تمام مسائل ان پی - 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 n به v وجود دارد، اگر شخص v شخص w 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 بپیدا کنیم که با مطلع شدن آن‌ها، همه از خبر فوق مطلع شوند. بهترین

الگوریتم برای محاسبه این مجموعه افراد از چه مرتبه‌ای است؟

$O(\min(n, m))$ (۲)

$O(m + n)$ (۱)

(۴) این مسئله ان پی - 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 باشند، 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 (۱)

۴ (۲)

۵ (۳)

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

-۷ گراف $G = (V, E)$ با وزن‌های مشبّت 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 بازنگشته‌ی 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 (۱)

۲ (۲)

۳ (۳)

۴ (۴)

-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؟

$$O(n \log n + m) \quad (۱)$$

$$O(n \log n) \quad (۲)$$

$$O(n) \quad (۳)$$

$$O(m) \quad (۴)$$

-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 اندختن 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. 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 $n > X$ شود، بازی متوقف شده و بازیکن آخر به عنوان برنده بازی انتخاب می‌شود. بازی‌ها به طور متوسط در چند مرحله به پایان می‌رسند؟

$$\Theta(\log \log n) \quad (1)$$

$$\Theta(\log^2 n) \quad (2)$$

$$\Theta(\log n) \quad (3)$$

(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؟

$$360 \quad (1)$$

$$405 \quad (2)$$

$$580 \quad (3)$$

$$630 \quad (4)$$

-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 مبدأ 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 virtual circuit switch(vc)، circuit switch(cs) و datagram packet switch(ps) به ترتیب از راست به چپ چند میلی ثانیه 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 میلی ثانیه، و پهنای باند هر لینک 15×10^6 بیت بر ثانیه می‌باشد).

$$9030, 8162, 8144 \quad (1)$$

$$9630, 9792, 8160 \quad (2)$$

$$8030, 8192, 8144 \quad (3)$$

$$9600, 9762, 8130 \quad (4)$$

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. $cwnd + \text{MSS} * \text{MSS}/cwnd$, ب. $cwnd/2 + 3 * \text{MSS}$

۲) الف. $cwnd + 3 * \text{MSS}$, ب. $cwnd + \text{MSS}$

۳) الف. $ssthresh$, ب. $cwnd + \text{MSS}/cwnd$

۴) الف. $cwnd + \text{MSS}$, 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 بایت حجم دارند. فرض کنید که

سرعت انتشار برابر $2 \times 10^8 \frac{\text{m}}{\text{s}}$ است و اندازه طول صفر 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 $5/870$ 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 Retransmission Timeout = RTT و $1\text{KB} = 1024\text{B}$, $1\text{Mbps} = 10^6\text{bit/s}$ باشد.)

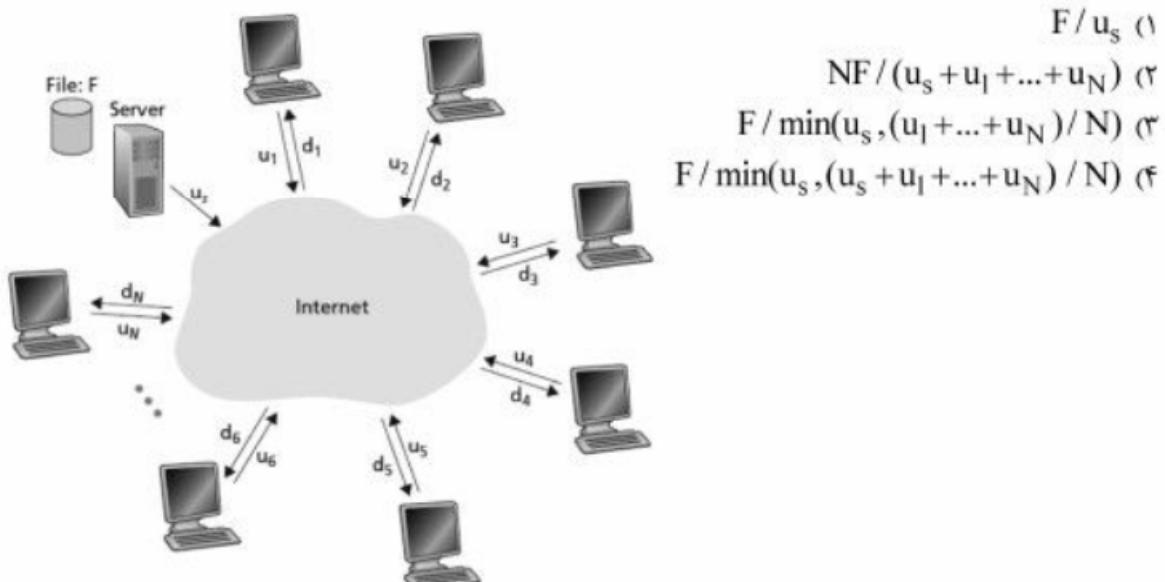
۱) ۸۷۰

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 به صورت باقیت منتقل می‌شود، کمترین زمان برای دریافت اطلاعات توسط کلیه گره‌ها کدام است؟ (فرض کنید سرعت دانلود تمام گره‌ها بسیار بالاست.)



- 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 محترمانگی در زمان ارسال پست الکترونیکی و پردازش کم مناسب است؟

(۱) ارسال $E_{K_B^+}(M)$ به گیرنده B

(۲) ارسال $E_{K_B^+}(M)$ و $\text{Sign}_{K_A^-}(M)$ به گیرنده B

(۳) ارسال $E_{K_B^+}(M)$ و $\text{Sign}_{K_A^-}(K_B^+)$ به گیرنده B

(۴) ارسال $E_{K_B^+}(M)$ و $\text{Sign}_{K_A^-}(E_{K_B^+}(M))$ به گیرنده B

- ۱۷ در صورتی که امکان جعل آدرس فرستنده در بسته IP وجود نداشته باشد، کدام حمله هنوز بسیار خطرناک است؟
 smurf (۱)
 DOS (۲)

(۳) هر سه حمله
 DDoS (۴)

- 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 سروورها 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 محترمانگی پیام را ارائه نمی‌دهند؟

(۱) TLS
 (۲) AH

(۳) PGP
 (۴) ESP

- ۲۰

عملت استفاده از توابع چکیده‌ساز همراه با امضا دیجیتال، کدام است؟

(۱) افزودن محترمانگی به سرویس احراز هویت ارائه شده توسط امضا دیجیتال

(۲) مقاومت بیشتر در برابر حملات جعل امضا

(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 چندنیسخه‌سازی مبتنی بر زمان مهر استفاده شود، کدام مورد درست است؟

T۱	T۲	T۳
$TS(T_1) = 100$	$TS(T_1) = 200$	$TS(T_2) = 300$
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 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 در خصوص پروتکل‌های قفل‌گذاری، درست نیست؟

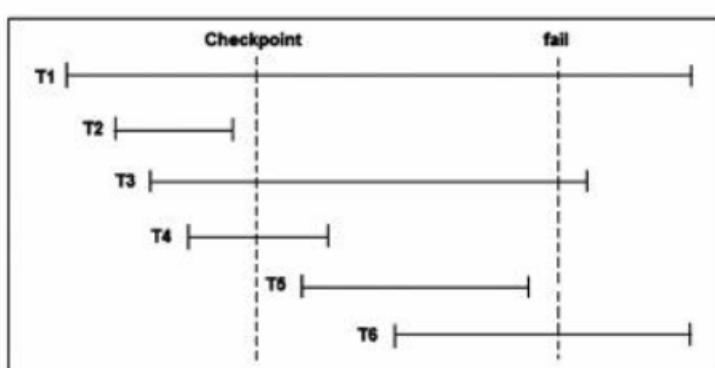
۱) در پروتکل قصد قفل‌گذاری، قفل IX با خودش سازگار است.

۲) پروتکل قفل‌گذاری جنگلی، توالی‌پذیری را تضمین نمی‌کند.

۳) پروتکل درختی، عدم وجود طرد 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 چنانچه نقطه وارسی از نوع commit consistent checkpoint باشد و از modification استفاده شده باشد، کدام مورد درست است؟



۱) فقط تراکنش‌های T1, T3 و T6 باید undo شوند.

۲) تراکنش‌های T1 و T3 undo و تراکنش‌های T4 و T5 redo شوند.

۳) تراکنش‌های T1 و T3 undo و تراکنش‌های T2 و T4 redo شوند.

۴) تراکنش‌های T1, T3 و T6 undo و تراکنش‌های T4 و T5 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) Checkpoint انجام
 (۴) تغییر یک بلوک داده در حافظه اصلی

- ۲۶ - در طرح اجرای (R₁(B), W₁(A), R₂(A), ..., W₂(B), R₃(B), ..., W₃(B)) ، چه دستوری در جای خالی باید قرار بگیرد تا این طرح اجرا، توالی‌پذیر تعارضی نباشد، ولی توالی‌پذیر نمایی باشد؟

- (۱) R_۲(B) یا W_۲(B)
 (۲) R_۲(A) یا W_۲(A)
 (۳) W_۲(B)
 (۴) W_۲(A)

- ۲۷ - یک پرس‌وجو را در نظر بگیرید که در کل چهار سند مرتبط با آن در مجموعه استناد 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 نامرتب بودن سند به پرس‌وجو می‌باشد. Average Precision این پرس‌وجو، کدام است؟

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

- (۱) $\frac{2}{5}$
 (۲) $\frac{8}{15}$
 (۳) $\frac{3}{10}$
 (۴) $\frac{3}{4}$

قدم به قدم، همراه دانشجو
WWW.GhadamYar.Ir

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

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

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

$$\text{TF-IDF}(w, d) = \text{TF}(w, d) \cdot \text{IDF}(w, d) \quad (۳)$$

- 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 فضای 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؟

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

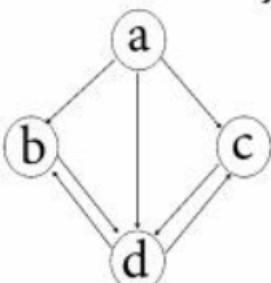
$$\delta \in [-1, 1] \quad (1)$$

$$\delta \in [0, 1] \quad (2)$$

$$\delta \geq |d|_u \quad (3)$$

$$\delta \geq 0 \quad (4)$$

- ۳۱ چهار صفحه a , b , c , d 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 $itemset_minimum\ support = ۲$

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 شی A, B, C, D را در نظر بگیرید. می‌خواهیم با استفاده از الگوریتم **Agglomerative Hierarchical Clustering** این چهار شی را خوشبندی 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	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 $(A, F, C, G, H) : 10$ است. با فرض اینکه $(A, B, C, D, E) : 20$, $E \neq H$ و $D \neq G$, $B \neq F$ 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؟

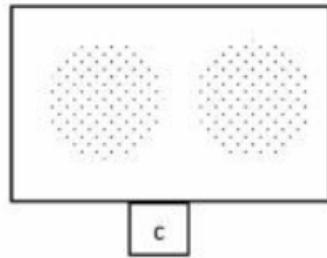
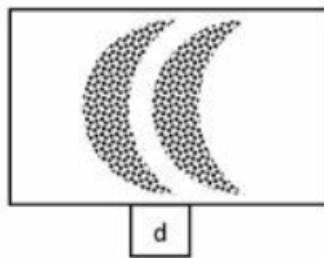
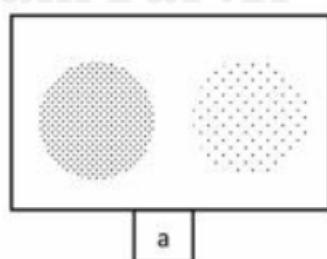
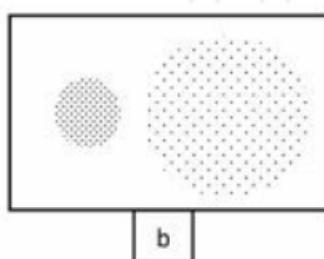
- ۱۶ (۱)
۲۲ (۲)
۶۰ (۳)
۶۴ (۴)

- 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. $c(T_j)$ تعداد اقلامی از دسته T_j را نشان می‌دهد که توسط روش خوشبندی در خوشبندی C_i 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

- ۰/۶ (۱)
۰/۷ (۲)
۰/۶۸ (۳)
۰/۷۵ (۴)

- 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؟



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 تحويل سفارش دهنده بدهد.

- ۳۸ - آسیب درک شده (perceived damage) از رقابت بین عمدۀ 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 نوادری 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 انبوه 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 کاربردپذیری تجربه کاربری
- (۲) پیچیدگی محاسباتی مرتبط با تخمین احتمال خرید کاربر از سایت تجارت الکترونیکی
- (۳) پردازش داده‌های حجمی برای استخراج داده (data mining) جهت ایجاد پروفایل برای کاربر
- (۴) نگرانی کاربر از جانب حفظ محمولگانی 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)، دسته‌بندی 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 گزینه «صفحة 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.»

- (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 مدنظر
- (۲) پروفایل کاربر و خصوصیات محتوا (content features)
- (۳) مدل‌های دانش و منطق فازی
- (۴) چارت امتحانات

قدم به قدم، همراه دانشجو...

WWW.GhadamYar.Ir



به اطلاع داوطلبان شرکت کننده در آزمون دکترای سال 1397 می‌رساند، این کلید اولیه غیر قابل استناد است و بس از دریافت نظرات داوطلبان و صاحب نظران، کلید تهابی سوالات تهیه و بر اساس آن کارنامه داوطلبان استخراج خواهد شد، در صورت تمایل هی توانید حداکثر تا تاریخ 15/12/1396 با مراجعته به سیستم پاسخگویی اینترنتی به نشانی request.sanjesh.org و تکمیل فرم اعتراض، به کلید سوالات آزمون دکترای سال 1397 اقدام نمایید.
لازم به ذکر است نظرات داوطلبان فقط از طریق اینترنت دریافت خواهد شد و به موارد ارسالی از طریق دیگر رسیدگی نخواهد شد.



عنوان دفترچه	نوع دفترچه	شماره باستخانمه	گروه امتحانی
مهندسی فناوری اطلاعات	E	1	فنی و مهندسی

شماره سوال	گزینه صحیح	شماره سوال	گزینه صحیح
1	4	31	1
2	3	32	3
3	2	33	4
4	1	34	2
5	2	35	1
6	3	36	2
7	2	37	4
8	4	38	1
9	1	39	1
10	2	40	4
11	3	41	1
12	2	42	1
13	1	43	1
14	2	44	1
15	4	45	2
16	1		
17	3		
18	1		
19	2		
20	4		
21	2		
22	1		
23	3		
24	3		
25	4		
26	3		
27	1		
28	4		
29	4		
30	2		

خروج