

کد کنترل

334

F

آزمون (نیمه متهم گز) ورود به دوره های دکتری - سال ۱۴۰۱

دفترچه شماره (۱)

صبح جمعه ۱۴۰۰/۱۲/۶



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

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

رشته مهندسی فناوری اطلاعات (کد ۲۳۵۸)

جدول مواد امتحانی، تعداد، شماره سؤال ها و زمان پاسخ گویی

زمان پاسخ گویی	تا شماره	از شماره	تعداد سؤال	مواد امتحانی
۱۵۰ دقیقه	۴۵	۱	۴۵	مجموعه دروس تخصصی: - حل مسئله (ساختمان داده ها و طراحی الگوریتم ها) - شبکه (شبکه های کامپیوتری و امنیت شبکه) - مهندسی اطلاعات (پایگاه داده پیشرفته، بازیابی پیشرفته اطلاعات و داده کاوی) - خدمات فناوری اطلاعات (تجارت الکترونیک و آموزش الکترونیک)

استفاده از ماشین حساب مجاز نیست.

این آزمون نمره منفی دارد.

حق چاپ، تکثیر و انتشار سؤال ها به هر روش (الکترونیکی و...) پس از برگزاری آزمون، برای تمامی اشخاص حقیقی و حقوقی تنها با مجوز این سازمان مجاز می باشد و با متخلفان برابر مقررات رفتار می شود.

* متقاضی گرامی، وارد نکردن مشخصات و امضا در کادر زیر، به منزله غیبت و حضور نداشتن در جلسه آزمون است.

اینجانب با شماره داوطلبی با آگاهی کامل، یکسان بودن شماره صندلی خود را
با شماره داوطلبی مندرج در بالای کارت ورود به جلسه، بالای پاسخنامه و دفترچه سؤال ها، نوع و کد کنترل درج شده
بر روی دفترچه سؤال ها و پایین پاسخنامه ام را تأیید می نمایم.

امضا:

۱- اگر بخواهیم داده ساختار صف را با استفاده از پشته پیاده سازی کنیم، طوری که عملیات های پایه ای صف در زمان

سرشکن $O(1)$ قابل انجام باشد، کدام مورد درست است؟

(۱) با سه پشته می توان این کار را انجام داد و سه پشته برای این کار لازم است.

(۲) با دو پشته می توان این کار را انجام داد و دو پشته برای این کار لازم است.

(۳) با یک پشته می توان این کار را انجام داد.

(۴) این کار امکان پذیر نیست.

۲- فرض کنید m آرایه مرتب داریم که در مجموع آرایه ها شامل n عدد هستند. می خواهیم از هر کدام از آرایه ها یک

عدد را انتخاب کنیم، به طوری که اختلاف بیشینه و کمینه اعداد انتخاب شده کمترین مقدار ممکن شود. این کار در

چه زمانی ممکن است؟ (بهترین گزینه را انتخاب کنید.)

(۲) $O(n \log m)$

(۱) $O(m + \log n)$

(۴) $O(nm)$

(۳) $O(n + m)$

۳- نمایش پیش ترتیب (Preorder) و میان ترتیب (Inorder) یک درخت دودویی مرتب با n رأس داده شده است.

کدام مورد درست است؟

(۱) درخت متناظر با این لیست رئوس، منحصر به فرد است.

(۲) دقیقاً دو درخت با این ترتیب ملاقات رئوس وجود دارد.

(۳) دقیقاً $\frac{n}{2}$ درخت با این ترتیب ملاقات رئوس وجود دارد.

(۴) هیچ یک از موارد درست نیست.

۴- فرض کنید یک هرم بیشینه شامل حداکثر n عدد داده شده است. جستجوی یک مقدار در این هرم بیشینه در کدام

کلاس پیچیدگی است؟ (بهترین گزینه را انتخاب کنید.)

(۴) $O(\sqrt{n})$

(۳) $O(n)$

(۲) $O(\log(n))$

(۱) $O(\frac{n}{\log(n)})$

۵- کدام مورد در خصوص دو گزاره (الف) و (ب) به ترتیب درست است؟

الف - اگر $f(n) \in O(2^{2^n})$ ، آنگاه $f(n) \in \Omega(2^n)$.

ب - اگر $T(n) = 2T(\frac{n}{2}) + O(n^2)$ ، آنگاه $T(n) = \Theta(n^2)$.

(۲) درست - نادرست

(۱) درست - درست

(۴) نادرست - نادرست

(۳) نادرست - درست

۶- رشته ای به طول اولیه $m \geq 1$ در یک صفحه از نرم افزار حروف چین داریم. دو عمل زیر را می توانیم به ترتیب دلخواه روی این رشته انجام دهیم.

• کپی: کل رشته موجود در صفحه را در حافظه ذخیره کن.

• پیست: رشته ذخیره شده در حافظه را به انتهای رشته موجود در صفحه اضافه کن.

به طور مثال، اگر رشته اولیه ab باشد، پس از انجام یک عمل کپی و یک عمل پیست، رشته موجود در صفحه به صورت $abab$ در خواهد آمد. اگر $dp[n]$ نشان دهنده طول بزرگ ترین رشته قابل ایجاد با n بار استفاده از عمل کپی یا پیست (به ترتیب دلخواه) باشد، کدام رابطه بازگشتی زیر برای محاسبه $dp[n]$ به ازای $n \geq 5$ درست است؟ (فرض کنید در ابتدا حافظه خالی است).

$$dp[n] = 2dp[n-1] \quad (۲)$$

$$dp[n] = 2dp[n-2] \quad (۱)$$

$$dp[n] = \max(dp[n-1], 2dp[n-2]) \quad (۴) \quad dp[n] = \max(2dp[n-2], 2dp[n-3]) \quad (۳)$$

۷- محاسبه عنصر یا بیشترین تکرار در یک آرایه n عضوی دلخواه در چه زمانی ممکن است؟ (بهترین گزینه را انتخاب کنید).

$$O(n) \quad (۴) \quad O(n^2) \quad (۳) \quad O(n\sqrt{n}) \quad (۲) \quad O(n \log(n)) \quad (۱)$$

۸- کدام یک از دو گزاره (الف) و (ب) در خصوص الگوریتم هافمن به ترتیب درست است؟

الف - اگر فراوانی یک حرف در یک متن بیشتر از $1/4$ باشد، آنگاه درخت هافمن متن لزوماً شامل کد یک بیتی است.

ب - اگر فراوانی هر حرف در یک متن کمتر از $1/3$ باشد، آنگاه درخت هافمن متن لزوماً شامل کد یک بیتی نیست.

(۱) نادرست - درست (۲) درست - درست (۳) درست - نادرست (۴) نادرست - نادرست

۹- فرض کنید G یک گراف کامل وزن دار با n رأس است. درخت پوشای کمینه G در چه زمانی قابل محاسبه است؟ (بهترین گزینه را انتخاب کنید).

$$\Theta(n \log n) \quad (۲) \quad \Theta(n) \quad (۱)$$

$$\Theta(n^2 \log n) \quad (۳) \quad \Theta(n^2) \quad (۴)$$

۱۰- آرایه مرتب A شامل n عدد به صورت اکیداً صعودی داده شده است. یک نفر این آرایه را به اندازه k واحد شیفت دوری داده و نتیجه را به صورت یک آرایه B به ما داده است. هدف پیدا کردن مقدار k است. در چه زمانی می توان مقدار k را با داشتن آرایه B محاسبه کرد؟ (بهترین گزینه را انتخاب کنید).

$$O(\sqrt{n}) \quad (۲) \quad O(n) \quad (۱)$$

$$O(\log^2 n) \quad (۴) \quad O(\log n) \quad (۳)$$

۱۱- به ازای اعداد صحیح مثبت n و m و a ، در چه مرتبه زمانی می توان $a^n \bmod m$ را محاسبه کرد؟ فرض کنید

عملیات های ضرب و جمع به پیمانه m در زمان $O(1)$ قابل انجام هستند. (بهترین گزینه را انتخاب کنید).

$$O(m) \quad (۴) \quad O(n) \quad (۳) \quad O(\sqrt{n}) \quad (۲) \quad O(\log n) \quad (۱)$$

۱۲- کدام گزینه در خصوص دو گزاره (الف) و (ب) به ترتیب درست است؟

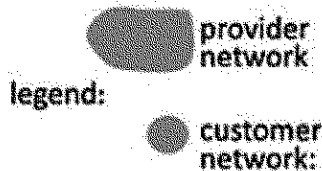
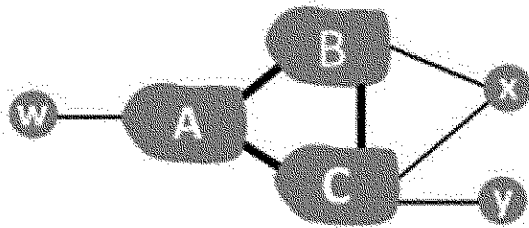
الف - به ازای هر گراف G ، ترتیبی از یال های گراف وجود دارد که با تنها یک دور ریلکس کردن یال ها به آن ترتیب در اجرای الگوریتم بلمن - فورد، کوتاه ترین مسیر از رأس s به تمام رأس های دیگر محاسبه می شوند.

ب - در مرتب سازی ادغامی هر عنصر با $O(\log n)$ عنصر دیگر مقایسه می شود.

(۱) درست - درست (۲) درست - نادرست

(۳) نادرست - درست (۴) نادرست - نادرست

- ۱۳- فرض کنید یک ارائه دهنده شبکه بخواهد ترافیک مربوط به مشتریان خود را منتقل کند. با توجه به شکل زیر، شبکه مسیر Cy را به کدام موجودیت ها آگهی می کند؟



(۱) فقط به X

(۲) فقط به X و A

(۳) فقط به A و B و X

(۴) به تمام موجودیت ها

- ۱۴- دو کلمه ۱۶ بیتی 11110101 11010011 و 10110011 01000100 را در نظر بگیرید. کدام گزینه برای واریسی مجموع اینترنت این دو کلمه درست است؟

(۲) 01010110 11100111

(۱) 10101001 00010111

(۴) 01011110 11000101

(۳) 01010110 11101000

- ۱۵- بهینه سازی باز ارسال سریع TCP را که در شکل زیر به تصویر کشیده شده در نظر بگیرید. در نظر داشته باشید که فرستنده مطمئن نیست که سگمنت با شماره ترتیب 100 در واقع تلف شده است. آیا یک فرستنده می تواند ACK تکراری برای یک سگمنتی که در واقع تلف نشده است، دریافت کند؟ با توجه به این مطلب کدام یک از عبارات درست است؟ (در مدل کانال در نظر گرفته شده بسته ها می توانند تلف شوند ولی دچار خرابی نخواهند شد). الف) اگر کانال بتواند ترتیب بسته ها را عوض کند، امکان ACK تکراری سه گانه حتی اگر بسته ای تلف نشده باشد، وجود دارد.

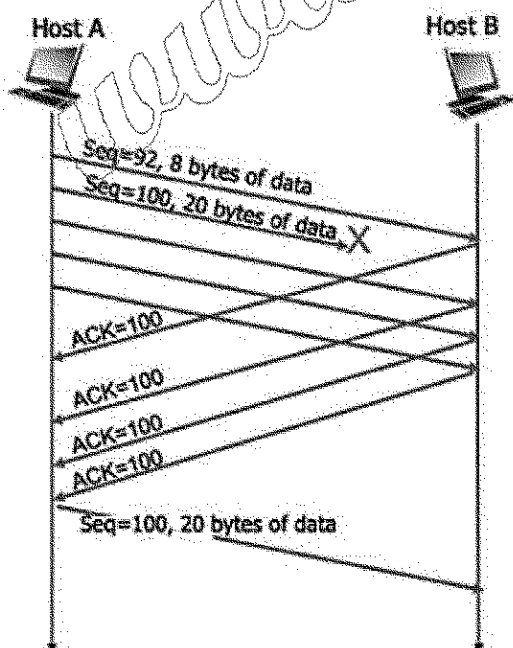
- ب) اگر کانال نتواند ترتیب بسته ها را عوض کند، ACK تکراری سه گانه در فرستنده نشانه آن است که حتماً یک بسته تلف شده است.

(۱) هر دو عبارت

(۲) عبارت الف

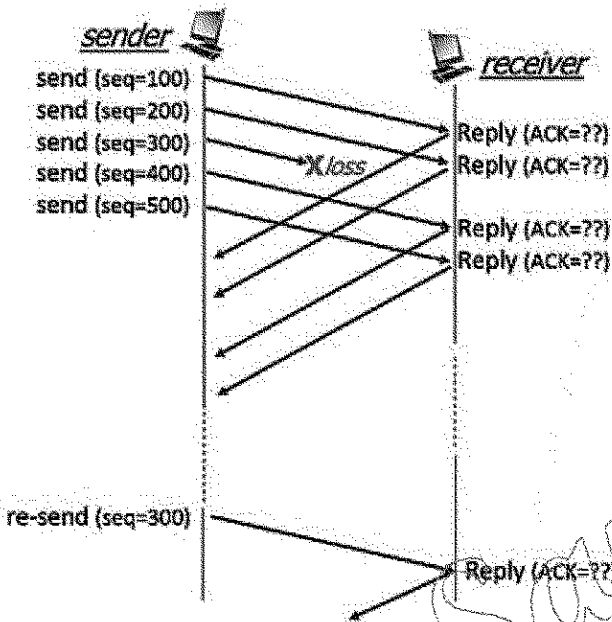
(۳) عبارت ب

(۴) علاوه بر اطلاعات فوق، فرستنده نیازمند اطلاعات اضافی است.



- ۱۶- با توجه به شکل زیر، فرستنده TCP سگمنت هایی که قسمت محموله آن ها 100 بایتی است را ارسال می کند. فرستنده TCP پنج قطعه با شماره های به ترتیب 100، 200، 300، 400 و 500 ارسال می کند. فرض کنید سگمنت با شماره ترتیب 300 تلف شده باشد. گیرنده TCP بسته های دریافتی را بافر می کند. عبارات زیر را در نظر بگیرید:
- * پس از دریافت سگمنت 100 گیرنده ACK با مقدار A را ارسال می کند.
 - * پس از دریافت سگمنت 200 گیرنده ACK با مقدار B را ارسال می کند.
 - * پس از دریافت سگمنت 500 گیرنده ACK با مقدار C را ارسال می کند.
 - * پس از دریافت بسته باز ارسالی 300 گیرنده ACK با مقدار D را ارسال می کند.
 - * گیرنده TCP پاسخی با ACK به مقدار E را ارسال نمی کند.

با توجه به موارد فوق کدام گزینه درست است؟



- (۱) A=100, B=200, C=400, D=500, E=600
 (۲) A=200, B=300, C=300, D=400, E=600
 (۳) A=100, B=200, C=200, D=500, E=300
 (۴) A=200, B=300, C=300, D=600, E=400

- ۱۷- کدام یک از موارد جزو مؤلفه های اساسی پروتکل IPv4 است؟

- (۱) قالب دیتاگرام IPv4 - پروتکل ها و الگوریتم های مسیریابی (مثل OSPF و BGP) - قواعد مربوط به نشانی دهی IPv4 - ICMP (پروتکل پیام کنترل اینترنت)
 (۲) قالب دیتاگرام IPv4 - قواعد اداره کردن بسته ها در روترها (مثلاً تکه سازی و دوباره سرهم کردن) - قواعد مربوط به نشانی دهی IPv4
 (۳) قالب دیتاگرام IPv4 - پروتکل ها و الگوریتم های مسیریابی (مثل OSPF و BGP) - قواعد مربوط به نشانی دهی IPv4
 (۴) قالب دیتاگرام IPv4 - پروتکل های کنترلر SDN - قواعد مربوط به نشانی دهی IPv4

- ۱۸- فرض کنید در روش امضای دیجیتال ElGamal، امضاکننده (آلیس) از k یکسان برای امضای دو پیام مختلف استفاده می کند. در این صورت کدام یک از گزینه ها نادرست است؟ (k عدد تصادفی است که توسط آلیس برای امضای پیام انتخاب می شود).

- (۱) مهاجم قادر است k را به دست آورد.
 (۲) حمله Existential forgery امکان پذیر می شود.
 (۳) مهاجم قادر است کلید خصوصی آلیس را به دست آورد.
 (۴) مهاجم قادر است امضای آلیس را روی یک پیام دیگر جعل کند.

۱۹- کدام یک از سرویس های امنیتی توسط پروتکل AH (Authentication Header) تامین نمی شود؟

- (۱) جلوگیری از حمله پیام های تکرار
- (۲) احراز اصالت پیام
- (۳) کنترل دسترسی
- (۴) محرمانگی

۲۰- در الگوریتم توافق کلید Diffie-Hellman، کدام یک از گزاره ها در مورد حمله Man in the Middle (مردی در میانه) درست است؟

- (۱) حتی اگر هر دو طرف گواهی دیجیتالی داشته باشند (بدین معنی که سهم های کلیدشان، امضای شخص ثالث قابل اعتماد را داشته باشد)، حمله مردی در میانه قابل انجام است، زیرا این الگوریتم ذاتاً در برابر این حمله امن نیست.
- (۲) اگر یکی از دو طرف گواهی دیجیتالی داشته باشد (بدین معنی که سهم کلیدش در الگوریتم، امضای شخص ثالث قابل اعتماد را داشته باشد)، حمله مردی در میانه کاملاً غیر قابل انجام می شود.
- (۳) حمله مردی در میانه تنها در لحظه تبادل سهم های کلید دو طرف روی کانال قابل انجام است و پس از آن امکان اوارفتن کلید و انجام حمله وجود ندارد.
- (۴) انجام حمله مردی در میانه حتی پس از پایان مرحله تبادل سهم های کلید دو طرف هم ممکن است و امکان بازیابی کلید توسط دشمن وجود دارد.

۲۱- در الگوریتم رمزنگاری DES، سه واحد رمزنگاری به صورت سری پشت هم قرار داده می شوند. رابطه چنین رمزکننده ای به صورت $C = E(K_3, D(K_2, E(K_1, P)))$ است که P همان Plain Text و C همان Cipher Text بوده و در آن اپراتور $E(K, \circ)$ به معنی رمز کردن یا کلید K و $D(K, \circ)$ به معنای رمزگشایی با کلید K است. رابطه رمزگشای DES نیز به طور مشابه قابل نوشتن است. طول کلید مؤثر در الگوریتم DES چقدر است و دلیل استفاده از واحد میانی رمزگشا (D) در میان دو رمزکننده (E) در طراحی این رمزکننده چیست؟

- (۱) طول کلید مؤثر ۱۶۸ بیت است. استفاده از D در مرحله میانی باعث امنیت بیشتر می شود و شکستن الگوریتم را مشکل تر می کند.
- (۲) طول کلید مؤثر ۱۶۸ بیت است. استفاده از D در مرحله میانی به کاربران DES امکان رمزگشایی پیام هایی که از سیستم های قدیمی تر DES آمده است را با تنظیم کلیدهای مناسب، می دهد.
- (۳) طول کلید مؤثر ۱۱۲ بیت است، زیرا کلید K_1 و K_3 باید حتماً یکسان انتخاب شوند. استفاده از D در مرحله میانی باعث امنیت بیشتر می شود و شکستن الگوریتم را مشکل تر می کند.
- (۴) طول کلید مؤثر ۱۱۲ بیت است، زیرا کلید K_1 و K_3 باید حتماً یکسان انتخاب شوند. استفاده از D در مرحله میانی به کاربران DES امکان رمزگشایی پیام هایی که از سیستم های قدیمی تر DES آمده است را با تنظیم کلیدهای مناسب، می دهد.

۲۲- در یک سیستم RSA، کلید خصوصی $d = 7$ و مقدار (تابع فی اولر پیمانه) $\phi(n) = 72$ است. اگر $p = 7$ یکی از اعداد اول سازنده پیمانه باشد، مقادیر کلید عمومی (e) و پیمانه (n) کدام هستند؟

- (۱) $e = 31, n = 91$
- (۲) $e = 17, n = 91$
- (۳) $e = 31, n = 77$
- (۴) $e = 17, n = 77$

- ۲۳- در پایگاه های داده توزیع شده و مبحث پروتکل نهایی (Commit) تراکنش در دو فاز، اگر سایت هماهنگ کننده در هنگام اجرای این پروتکل برای تراکنش T_i دچار خرابی شود، سایر سایت های فعال (مشارکت کننده) باید در مورد سرنوشت T_i تصمیم بگیرند. چند مورد از عبارت زیر در این خصوص درست است؟
- اگر یک سایت فعال حاوی یک رکورد $\langle \text{commit } T_i \rangle$ در فایل گزارش (Log) خود باشد، T_i باید تأیید نهایی شود.
 - اگر یک سایت فعال حاوی یک رکورد $\langle \text{abort } T_i \rangle$ در فایل گزارش خود باشد، T_i باید لغو شود.
 - اگر حتی یک سایت فعال حاوی یک رکورد $\langle \text{ready } T_i \rangle$ در گزارش خود نباشد، سایت هماهنگ کننده مربوطه نمی تواند تصمیم به تأیید T_i گرفته باشد، بنابراین T_i باید لغو شود.
 - در حالت غیر از موارد فوق، تمام سایت های فعال دارای یک رکورد $\langle \text{ready } T_i \rangle$ در گزارش های خود بوده، اما هیچ رکورد کنترلی دیگری در مورد T_i وجود ندارد. بنابراین از آنجایی که هماهنگ کننده هم دچار خرابی شده است، T_i باید لغو شود.

۴ (۴)

۳ (۳)

۲ (۲)

۱ (۱)

- ۲۴- وظیفه تأمین خاصیت دوام (Durability) بر عهده کدام زیرسیستم است؟

(۱) Backup Subsystem

(۲) Recovery Subsystem

(۳) Concurrency Control Subsystem

- (۴) این خاصیت مرتبط با هیچ زیرسیستمی نیست و مرتبط با رسانه ذخیره سازی، یعنی حافظه های دائمی مانند هارد دیسک است.
- ۲۵- نوع دیگری از پروتکل درخت (Tree Protocol) وجود دارد که نام آن پروتکل جنگل (Forest Protocol) است. پایگاه داده در پروتکل جنگل به شکل یک جنگل از درخت های ریشه دار است. تمامی قوانین این دو پروتکل یکسان هستند. به جز یک قانون که هر تراکنش T_i ، اولین قفل بر روی هر یک از درخت ها را می تواند بر روی هر گره دلخواه اعمال کند. اکنون طبق پروتکل جنگل کدام مورد درست است؟
- (۱) این پروتکل گرسنگی (Starvation) ندارد ولی عدم وجود بن بست (Deadlock) را تضمین می کند.
- (۲) این پروتکل توالی پذیری نمایی (View Serializable) و همچنین توالی پذیری نتیجه ای را تضمین نمی کند.
- (۳) این پروتکل قابلیت بازیابی (Recoverable) ندارد ولی توالی پذیری تعارضی (Conflict Serialization) را تضمین می کند.

- (۴) این پروتکل توالی پذیری تعارضی (Conflict Serialization) و همچنین طرد تسلسلی (Cascadeless) را تضمین می کند.

- ۲۶- یک مدیر کنترل همروندی مبتنی بر مهر زمان سختگیرانه (Strict Timestamp Ordering) را در نظر بگیرید. در زیر دنباله ای از رویدادها شامل رویدادهای شروع، که در آن St_i یعنی تراکنش T_i شروع می شود و co_i یعنی تراکنش T_i کامیت می شود. این دنباله نشان دهنده زمان واقعی است و زمان بند مبتنی بر مهر زمان به تراکنش ها، مهر زمان را به ترتیب شروع شان تخصیص می دهد. مشخص کنید که برای آخرین دستور چه اتفاقی می افتد؟ (ترتیب دستورات دنباله از چپ به راست است).

$St_1; St_2; St_3; r_1(A); w_1(A); r_2(A)$

(۲) صرف نظر می شود. (ignored)

(۱) بازگشت داده می شود. (rolled back)

(۴) اجرا می شود. (accepted)

(۳) به تأخیر می افتد. (delayed)

۲۷- رابطه $R(A, B, C, D)$ را در نظر بگیرید که در آن A کلید اصلی رابطه است. فرض کنید R_1 ، R_2 و R_3 سه قطعه (fragmentation) از رابطه R باشند که به صورت زیر تعریف شده اند. کدام یک از معیارهای صحت (Correctness) نقض شده است؟

$$R_1 = \pi_{AB} \sigma_{A \geq 2}(R)$$

$$R_2 = \sigma_{A < 2} \pi_{AB}(R)$$

$$R_3 = \pi_{CD}(R)$$

(۲) جدایی (Disjointness)

(۱) بازسازی (Reconstruction)

(۴) هم پوشانی (Overlap)

(۳) کامل بودن (Completeness)

۲۸- تعداد تکرار کلمات در اسناد D_1 و D_2 به شرح جدول زیر است که در آن T_1 ها ترم های موجود در اسناد هستند. عبارت $T_1 T_2 T_3$ به عنوان پرس و جوی ورودی دریافت می شود. در صورتی که از مدل زبانی یونیگرم برای بازیابی استفاده شود، امتیاز query likelihood برای دو سند D_1 و D_2 به ترتیب (از راست به چپ) چند خواهد شد؟

	T_1	T_2	T_3	T_4	T_5	T_6
D_1	۵	۵	۵	۳	۲	۱۰
D_2	۱	۷	۵	۴	۳	۵

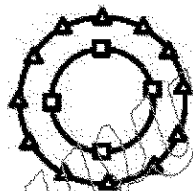
(۲) ۰/۱۶ و ۰

(۱) ۰/۰۸ و ۰/۰۳

(۴) ۰/۰۱۶ و ۰

(۳) ۰/۰۸ و ۰/۰۳

۲۹- در یک سیستم بازیابی، سندها را با بردارهای دوبعدی نمایش می دهند. برای پرس و جوی P مجموعه ای به صورت زیر گزارش شده است که در آن اسناد مرتبط (مربع) و غیر مرتبط (مثلث) روی دو دایره متحدالمرکز با فاصله مساوی قرار دارند. الگوریتم فیدبک روکیو (Rocchio's feedback algorithm) با ضرایب اهمیت یکسان اسناد مرتبط و غیر مرتبط، پرس و جوی جدید Q را از روی P می سازد. درباره این دو پرس و جو کدام گزینه درست است؟



- (۱) به دلیل توزیع متوازن اسناد پاسخ، برای Q دقیقاً همان مجموعه جواب P بازیابی می شود.
- (۲) به دلیل توزیع هم مرکز اسناد، برای Q اسناد مرتبط کاملاً متفاوتی بازیابی می شود.
- (۳) به دلیل تعداد بیشتر اسناد غیر مرتبط، برای Q اسناد مرتبط دقیق تری بازیابی می شود.
- (۴) به دلیل تراکم بیشتر اسناد مرتبط، برای Q اسناد مرتبط بیشتری بازیابی می شود.

۳۰- فرمول هموارسازی Absolute discounting(AD) و Dirichlet Prior(DP) و Jelinek-Mercer(JM) برای محاسبه احتمال وقوع ترم w در سند d را در نظر بگیرید. فرض کنید $c(w,d)$ تعداد تکرار w در d ، $|d|_u$ تعداد ترم های یکتای d ، $p(w|C)$ احتمال وقوع w در مجموعه مرجع C و $|d|$ طول سند d ، δ پارامتر AD، μ پارامتر DP و λ پارامتر JM هستند. حال فرض کنید فرمول کلی ما برای هموارسازی به صورت زیر است. در این صورت مقدار X در روش AD، DP و JM به ترتیب کدام است؟

$$P(w|d) = \begin{cases} p_{\text{seen}}(w|d) & \text{اگر } w \text{ در } d \text{ وجود دارد:} \\ \lambda p(w|c) & \text{در غیر این صورت:} \end{cases}$$

$$\lambda \text{ و } \frac{|d| + \mu}{\mu} \text{ و } \frac{|d|_u}{\delta |d|} \quad (۱)$$

$$\lambda \text{ و } \frac{\mu}{|d| + \mu} \text{ و } \frac{\delta |d|_u}{|d|} \quad (۲)$$

$$\frac{\lambda}{|d|} \text{ و } \frac{|d| + \mu}{\mu} \text{ و } \frac{|d|_u}{|d|} \quad (۳)$$

$$\frac{\lambda}{|d|} \text{ و } \frac{\mu}{|d| + \mu} \text{ و } \frac{\delta |d|_u}{\delta |d|} \quad (۴)$$

۳۱- چه تعداد از گزاره های زیر درباره الگوریتم LSI (Latent Semantic Indexing) درست است؟

- با کاهش تعداد ابعاد، فراخوانی (Recall) بهبود می یابد.
- ابعاد به دست آمده از این الگوریتم شبیه خوشه های معنایی کلمات هستند.
- هزینه محاسباتی تجزیه ماتریس یکی از چالش های اصلی این الگوریتم است.
- الگوریتم LSI به صورت اتوماتیک تعداد ابعاد بهینه برای نمایش اسناد را می یابد.
- به کمک مقادیر منفی در ابعاد حاصله، می توان پرس وجوهای منفی را هم پاسخ داد.
- کلمات کم بسامد یا دارای خطای املائی سبب ایجاد مقادیر منفی در ابعاد نهایی اسناد هستند.

۵ (۱) ۴ (۲) ۳ (۳) ۲ (۴)

۳۲- تعداد خوشه بندی های ممکن سلسله مراتبی N نمونه براساس شباهت دوبه دوی آنها با تعداد کدامیک از مجموعه های زیر هم رتبه است؟

- (۱) تعداد درخت های دودویی دارای N گره
- (۲) تعداد جایگشت های غیر تکراری N شی متمایز
- (۳) تعداد زیرمجموعه های یک مجموعه دارای N عضو
- (۴) تعداد درخت های پوشای کمینه گرافی با N رأس و $M > N$ یال

۳۲- یک frequent itemset به نام L و یک زیرمجموعه غیر تهی از آن به نام S را در نظر بگیرید. همچنین فرض کنید که S_1 زیرمجموعه ای از S است. حال دو قانون زیر را در نظر بگیرید. کدام جمله درست است؟

- $S \rightarrow (L - S)$
 $S_1 \rightarrow (L - S_1)$

- (۱) مقدار Support قانون $S \rightarrow (L - S)$ بزرگتر از قانون $S_1 \rightarrow (L - S_1)$ است.
- (۲) مقدار Confidence قانون $S_1 \rightarrow (L - S_1)$ بزرگتر مساوی قانون $S \rightarrow (L - S)$ است.
- (۳) مقدار Confidence قانون $S \rightarrow (L - S)$ بزرگتر مساوی قانون $S_1 \rightarrow (L - S_1)$ است.
- (۴) نمی توان دقیق اظهار نظر کرد.

۳۴- در مورد گزاره های زیر کدام گزینه درست است؟

گزاره A: مجموعه داده X دارای ۱۰۰ بعد است. با استفاده از PCA ابعاد آن به ۲۰ بعد کاهش می یابد. سپس مجدداً روی این مجموعه جدید الگوریتم PCA اعمال شده و ابعاد آن به ۳ بعد کاهش می یابد.

گزاره B: مجموعه داده X دارای ۱۰۰ بعد است. با استفاده از PCA ابعاد آن به ۳ بعد کاهش می یابد.

- (۱) لزوماً ابعاد به دست آمده در هر گزاره مشابه نیستند.
- (۲) ابعاد به دست آمده در هر دو گزاره مشابه هم هستند.
- (۳) ابعاد به دست آمده در هر دو گزاره مشابه نیستند.
- (۴) به دلیل حذف همبستگی ویژگی ها در مرحله اول، مرحله دوم گزاره A انجام پذیر نیست.

۳۵- اگر برای تشخیص ناهنجاری از الگوریتم Isolation Forest استفاده شود، داده های ناهنجار معمولاً چه جایگاهی در درخت های تصمیم به دست آمده پیدا می کنند؟

- (۱) به دلیل تفاوت عمده حداقل یکی از ویژگی های آنها با داده های عادی، در مسیرهای نزدیک به ریشه طبقه بندی می شوند.
- (۲) به دلیل شباهت زیاد نمونه های ناهنجار به همدیگر، در یک زیردرخت جداگانه طبقه بندی می شوند.
- (۳) به دلیل شباهت زیاد آنها به نمونه های عادی، در همه جای درخت طبقه بندی می شوند.
- (۴) به دلیل تعداد کم آنها در مسیرهای طولانی و دور از ریشه طبقه بندی می شوند.

۳۶- تجارت الکترونیکی در دوره گذار از محیط های سنتی به محیط های مبتنی بر L.T. متبحر به کدام مورد می شود؟

- (۱) افزایش قدرت چانه زنی مشتریان
- (۲) کاهش تهدید کالای جایگزین
- (۳) افزایش قدرت کانال های توزیع کالا
- (۴) افزایش موانع ورود تازه وارد ها به صنعت

۳۷- کدام گزینه، زیرمجموعه تبلیغات مستقیم از طریق موتورهای جستجو به شمار می آید؟

- (۱) تبلیغات کلیک - تبلیغات ارگانیک
- (۲) تبلیغات پولی - تبلیغات ارگانیک
- (۳) تبلیغات کلمات کلیدی - تبلیغات کلیک
- (۴) افزایش رتبه وبسایت (SEO) - تبلیغات کلمات کلیدی

۳۸- نسبت مشتریانی که واقعاً از یک وبسایت خرید می کنند به کل کاربران بازدیدکننده از آن وبسایت، چه نام دارد؟

- (۱) جذابیت
- (۲) نرخ کلیک
- (۳) نرخ همسان
- (۴) نرخ تبدیل

۳۹- کدام گزاره در خصوص سرمایه‌پذیری جمعی (crowd-funding) درست است؟

- (۱) سرمایه‌پذیری جمعی محدود به رمز ارزها است.
 - (۲) سرمایه‌پذیری جمعی نوع خاصی از سرمایه‌گذاری فرشته (angel) است.
 - (۳) در سرمایه‌گذاری جمعی تمامی قوانین رایج و متعارف سرمایه‌گذاری در بازار سرمایه رعایت می‌شود.
 - (۴) ریسک‌های این نوع سرمایه‌گذاری نسبت به دیگر سرمایه‌گذاری‌ها نظیر بازار بورس برای سرمایه‌گذاران بالاتر است.
- ۴۰- کدام دیجیتال گنجانده شده در یک تصویر دیجیتال یا فایل صوتی به‌طوری که قابل رؤیت نباشد، چه نامیده می‌شود؟
- (۱) علامت تجاری یا تریدمارک
 - (۲) پنهان‌نگاری یا استگانوگرافی
 - (۳) واترمارک
 - (۴) رمزنگاری

۴۱- کدام مورد تعریف دقیق‌تری از ابزارهای تعاملی ارائه می‌دهد؟

- (۱) ابزارهایی که به دانشجویان امکان می‌دهد تا با یک یا چند دانشجوی دیگر تعامل برقرار کنند و به‌صورت مشارکتی بیاموزند.
 - (۲) ابزارهایی که برای کاربردهایی نظیر ویدئو کنفرانس و استریم و پخش زنده مورد استفاده قرار می‌گیرند.
 - (۳) ابزارهایی که برای برقراری ارتباط بین استاد و دانشجویان به‌طور فعال مورد استفاده قرار می‌گیرد.
 - (۴) فرایندهایی که برای برقراری ارتباط بین دانشجویان در یک کلاس مورد استفاده قرار می‌گیرد.
- ۴۲- کدام گزاره در خصوص یادگیری مشارکتی درست است؟

- (۱) یادگیری مشارکتی می‌تواند هم حالت هم‌زمان (سنکرون) و هم غیرهم‌زمان (غیرسنکرون) باشد.
- (۲) یادگیری مشارکتی از اولین نسل سیستم‌های یادگیری الکترونیکی مد نظر قرار گرفت.
- (۳) استفاده از شبکه‌های اجتماعی، آسیب جدی یادگیری مشارکتی به‌شمار می‌آید.
- (۴) قوت یادگیری مشارکتی بیشتر در آموزش مسائل پیچیده‌تر آشکار می‌شود.

۴۳- کدام مورد درباره آموزش مجازی و یادگیری آنلاین درست نیست؟

- (۱) آموزش مجازی، سرعت و زمان یادگیری را با یادگیرنده تطبیق می‌دهد.
 - (۲) آموزش مجازی دسترسی به دوره‌های درسی بیشتری را برای یادگیرنده فراهم می‌کند.
 - (۳) در آموزش مجازی، فضای بهتری برای بحث و تبادل نظر میان معلم و یادگیرنده برقرار است.
 - (۴) در آموزش مجازی معلم مجبور نیست که برای جلب توجه دانشجو تلاش کند و فقط باید محتوای خوبی تهیه کند.
- ۴۴- آموزش مجازی سنکرون یا هم‌زمان به چه معنا است؟

- (۱) دانشجویان با به‌کارگیری سیستم مدیریت یادگیری به محتوای درسی دسترسی پیدا می‌کنند و در مورد آن نظر می‌دهند.
- (۲) دانشجویان با به‌کارگیری سیستم مدیریت یادگیری یا هر سامانه دیگری از طریق شرکت در کلاس مجازی، وبینار و یا تلفن به یادگیری می‌پردازند.
- (۳) دانشجویان از طریق تالارهای گفتگو درباره مطالب درسی به بحث و تبادل نظر می‌پردازند.
- (۴) دانشجویان در زمان و مکان دلخواه خود به یادگیری می‌پردازند.

۴۵- کلاس معکوس یا Flipped-Classroom توسط کدام عبارت بهتر تبیین و تشریح می‌شود؟

- (۱) معلم مطالب را از قبل آماده کرده و در اختیار دانشجو قرار می‌دهد. در کلاس دانشجو با آمادگی مطالب را با هم کلاسی‌ها مباحثه کرده و نقش معلم هماهنگی است.
- (۲) معلم به روش سنتی مطالب را آماده می‌کند ولی به دانشجویان تکلیف داده می‌شود که آنها هم مطالب را دوباره تهیه کنند.
- (۳) تنها تکلیف دانشجویان مشارکت در بحث است و خارج از کلاس کاری ندارند.
- (۴) تمرکز در این روش بر یادگیرنده است و معلم نقش خاصی ندارد.