نام درس: ذخیره و بازیابی اطلاعات
رشته تحصیلی-گرایش: مهندسی کامپیوتر
کد درس: ۲۶۱۲۵۰

تعداد سؤال: تستی ۲۵ تکمیلی — تشریحی ۷
زمان امتحان: تستی و تکمیلی ۴۰ دقیقه تشریحی ۶۰ دقیقه
تعداد کل صفحات: ۴

۱. در همه طراحی‌های ساختار فایل، آنچه مورد توجه است:

الف. به حداقل رساندن دفعات دستیابی به فایل است.

ب. دستیابی به اطلاعات با یک مراجعه به دیسک است.

ج. حداقل رساندن دفعات دستیابی به دیسک و به حداکثر رساندن احتمال وجود اطلاعات مورد نظر برنامه کاربردی در حافظه است.

د. کاهش دفعات دستیابی به حداکثر سه بار می‌باشد.

۲. با فراخوانی fread، تعداد عناصر خوانده شده باز گردانده می‌شود، در این مورد اگر fread مقدار صفر را بازگرداند؟

الف. برنامه به انتهای فایل رسیده است. ب. برنامه به ابتدای فایل رسیده است.

ج. برنامه دارای خطای نحوی است. د. برنامه اداری خطای صرفی است.

۳. فرمان نام فایل tail چه عملی را انجام می‌دهد؟

الف. فایل نامبرده را حذف می‌کند. ب. ده خط آخر فایل متنی را چاپ می‌کند.

ج. حالت محافظت را در فایل نامبرده عوض می‌کند. د. محتویات فایل را نمایش می‌دهد.

۴. .....عبارت است از زمان لازم برای انتقال بازوی دستیابی (access arm) به سیلندر مناسب.

الف. زمان انتقال ب. تأخیر چرخش ج. زمان پیگرد د. زمان درنگ

۵. می‌خواهیم یک کپی پشتیبان، از فایلی به طول ۳ میلیون رکورد صد بایتی تهیه کنیم. اگر بخواهیم این فایل را روی یک نوار bpi ۶۲۵۰ با شکاف بین بلاکی ۳، اینچ ذخیره کنیم، چند اینچ نوار لازم است؟

الف. ۶۳۲۰۰۰ ب. ۲۶۳۳۳ ج. ۳۱۶۰۰۰ د. ۰/۰۱۶

۶. نقطه ضعف اصلی CD-ROM این است که .....

الف. ظرفیت ذخیره سازی در آن کم است. ب. جستجو در آن بسیار کند است.

ج. بهای آن زیاد است. د. دوام آن کم است.

۷. در حالت متوسط، جستجو دودویی تقریباً به .....مقایسه نیاز دارد.

الف. $(\log_2 n)$ ب. $[\log_2 n+1]$ ج. $([\log_2 n]+\frac{1}{2})$ د. $([\log_2 n]+2)$

۸. کدامیک از عبارات زیر درباره بلوک‌بندی رکوردها نادرست است؟

الف. بلوک‌بندی مرتبه عملکرد جستجوی ترتیبی را تغییر می‌دهد.

ب. بلوک‌سازی، اختلاف میان سرعت دستیابی در حافظه و زمان دستیابی در حافظه ثانویه را نشان می‌دهد.

ج. بلوک‌سازی تعداد مقایسه‌هایی را که باید در حافظه انجام شوند، تغییر نمی‌دهد.

د. با بلوک‌سازی در زمان صرفه‌جویی می‌شود زیرا مقدار جستجو کاهش می‌یابد.

۹. پایین‌ترین سطح سازمان‌دهی که معمولاً به فایلی اعمال می‌شود .....

الف. جریان رکوردها است. ب. جریان فیلدهاست.

ج. جریان بیتها است. د. جریان بایتها است.

۱۰. روال فشرده سازی pack در system V به طور معمول به کاهشی در حدود ۲۵ تا ۳۰ درصد در فایلهای متن دست پیدا می‌کند، که این مقدار در فایلهای دودویی به طور قابل توجهی کمتر است. زیرا .....

الف. در آنها پراکندگی کمتر است.
ب. در آنها پراکندگی یکنواخت‌تر است.
ج. در آنها طول رانش بیشتر است.
د. در آنها پراکندگی ثابت است.

۱۱. کدامیک از گزینه‌های زیر در مورد انقیاد صحیح است؟

الف. یک مزیت مهم جلو انداختن انقیاد آن است که این روش امن‌تر است.

ب. عیب انقیاد مستقیم در فایل آن است که سازمان‌دهی دوباره فایل داده‌ها باید منجر به اصلاح همه فایلهای شاخص انقیاد یافته شود.

ج. انقیاد در زمان ایجاد فایل، منجر به دستیابی کندتر می‌شود.

د. هنگامی که انقیاد طی اجرای برنامه رخ دهد، به آن انقیاد و محکم گفته می‌شود.

۱۲. دستور ..... در یونیکس دو فایل را که باید به ترتیب اسکی مرتب شوند، مقایسه می‌کند تا ببیند وجوه مشترک آنها در چیست؟

الف. diff
ب. cmp
ج. qsort
د. Comm.

۱۳. توصیف زیر در رابطه با کدام روش مرتب‌سازی می‌باشد؟

"یک الگوریتم مرتب‌سازی که به ویژه برای مرتب سازی فایلهای بزرگی مناسب است که در حافظه می‌گنجند زیرا اجزای آن می‌تواند با I/O همپوشانی کند."

الف. مرتب سازی هرمی
ب. مرتب سازی ادغامی
ج. مرتب سازی درجی
د. مرتب سازی انتخابی

۱۴. کدامیک از موارد زیر از خواص یک درخت B از مرتبه m محسوب نمی‌شوند؟

الف. هر صفحه حداکثر m فرزند دارد.
ب. ریشه حداقل دو فرزند دارد (مگر اینکه برگ باشد)
ج. برگها می‌توانند در دو سطح قرار گیرند.
د. هر صفحه به جز ریشه و برگها حداقل [m/۲] فرزند دارد.

۱۵. درخت‌های B و B^* و B^+ پیشوندی ساده در کدامیک از خصوصیات زیر مشترک نیستند؟

الف. همگی ساختارهای شاخص صفحه‌ای هستند.

ب. در هر سه روش درختهایی نگهداری می‌شود که ارتفاع آنها موازین است.

ج. درهمه موارد، درختها از بالا به پایین رشد می‌کنند و موازنه از طریق شکستن بلوک، ادغام و توزیع دوباره حفظ می‌شود.

د. هر سه روش را می‌توان به عنوان ساختارهای درختی مجازی پیاده‌سازی کرد.

۱۶. چگونه می‌توان بدون ایجاد تعداد زیادی از فایلهای کوچک، تعداد نامحدودی از لیست‌های متفاوت ایجاد کرد که هر یک طول متغییری داشته باشد؟

الف. استفاده از رکوردهایی با طول متغییر
ب. استفاده از رکوردهایی با طول ثابت
ج. استفاده از لیست‌های پیوندی
د. با استفاده از رکوردهایی با طول متغییر لیستهای پیوندی

۱۷. کدام روش موجب افزایش برخوردها می‌شود؟

الف. استفاده از باکت‌ها
ب. بخش کردن رکوردها
ج. استفاده از حافظه اضافی
د. قراردادن بیش از یک آدرس برای یک رکورد

نام درس: ذخیره و بازیابی اطلاعات
رشته تحصیلی-گرایش: مهندسی کامپیوتر
کد درس: ۲۶۱۲۵۰

تعداد سئوالات: تستی ۲۵ تکمیلی — تشریحی ۷
زمان امتحان: تستی و تکمیلی ۴۰ دقیقه تشریحی ۶۰ دقیقه
تعداد کل صفحات: ۴

۱۸. در رابطه با یک درخت B از مرتبه m کدامیک از گزینه‌های زیر نادرست است؟
الف. هر صفحه حداکثر m فرزند دارد. ب. تمام برگها در یک سطح قرار دارند
ج. هر صفحه به جز ریشه و برگها، حداقل ۲ − m فرزند دارد. د. ریشه حداقل دو فرزند دارد (مگر اینکه برگ باشد)

۱۹. کدامیک از گزینه‌های زیر در مورد مدل پردازش کلک تربیتی نادرست است؟
الف. این مدل در مسائلی که شامل اجرای عملیات مجموعه‌ای بر روی دو یا چند فایل ورودی و ایجاد یک یا چند فایل خروجی است، به کار می‌رود.
ب. مقدار دهی اولیه از قطعات اصلی این مدل است.
ج. در این مدل یک حلقه همزمان سازی اصلی به کار می‌رود و تا زمانی که رکوردها باقی مانده‌اند، این حلقه ادامه دارد.
د. کلیدهای خوانده شده از فایل بدون انجام مقایسه، انتخاب می‌شوند.

۲۰. در مرتب‌سازی، هرمی، هرم یک درخت دودویی است. کدامیک از ویژگیهای زیر در مورد این درخت نادرست است؟
الف. هر گره دارای کلیدی است که آن کلید بزرگتر یا مساوی کلید واقع در گره پدرش است.
ب. برگهای این درخت دودویی می‌تواند بیش از دو سطح داشته باشد.
ج. برای نگهداری این درخت می‌توان آرایه‌ای اختصاص داد که گره ریشه در آن اندیس ۱ و اندیس فرزندان چپ و راست گره i ، به ترتیب برابر 2i و 2i+1 باشند.
د. اندیس گرده پدر j برابر ⌊j/2⌋ است.

۲۱. کدامیک از عبارات زیر صحیح است؟
الف. الگوریتم ادغام موازنه شده دارای این مزیت است که نوشتن برنامه برای پیاده سازی آن آسان است، زیرا این الگوریتم هوشمند نیست.
ب. الگوریتم ادغام موازنه شده دارای این عیب است که نوشتن برنامه برای پیاده سازی آن مشکل است، زیرا این الگوریتم هوشمند نیست.
ج. الگوریتم ادغام موازنه شده دارای این عیب است که نوشتن برنامه برای پیاده سازی آن مشکل است، زیرا این الگوریتم هوشمند است.
د. الگوریتم ادغام موازنه شده دارای این مزیت است که نوشتن برنامه برای پیاده سازی آن آسان است، زیرا این الگوریتم هوشمند است.

۲۲. در کدامیک از تکنیک‌های پردازش رکوردهای سرریز، اندیس‌سازی را با درهم‌سازی ترکیب می‌کند؟
الف. درهم‌سازی دوگانه ب. سرریز فزاینده زنجیره‌ای ج. پیوند با ناحیه سرریز دیگر د. جدول‌های پراکندگی

۲۳. درهم‌سازی ..... باعث می‌شود رکوردهای سرریز در فاصله دوری از آدرسهای خانگی قرار بگیرند و رکوردها در سطح فایل پخش شوند.
الف. کامل ب. دوگانه ج. کمینه د. اندیس شده

۲۴. کدامیک از گزینه‌های زیر نادرست است؟

الف. صفحه صفحه کردن یک راه حل بالقوه برای مشکل ناکارا بودن استفاده از دیسک در درختهای جستجوی دودویی است

ب. تقسیم یک درخت به چندین صفحه، باعث کند شدن جستجو در حافظه جانبی می‌شود.

ج. مشکل اصلی استفاده درختهای صفحه‌ای فنوز هم استفاده مؤثر از دیسک است.

د. تعداد پیگردهای لازم برای نسخه صفحه‌ای یک درخت جستجوی دودویی کاملاً پر و موازنه شده $\log_{k+1}(N+1)$ است که N تعداد کلیدها و K تعداد کلیدهایی که در یک صفحه نگهداری می‌شوند را نشان می‌دهد.

۲۵. جستجوی ترتیبی عملی از مرتبه ..... است، حال آنکه دستیابی مستقیم، عملی از مرتبه ..... است.

الف. $O(n)، O(1)$ ب. $O(n)، O(n^2)$ ج. $O(1)، O(n)$ د. $O(n)، O(\log(n))$

سؤالات تشریحی

۱. اجزای ساختار I/O هسته (kernel) در یونیکس را با استفاده از یک شکل نمایش دهید. (۱/۵ نمره)

۲. ابزار dd در یونیکس برای چه کاربردهایی در نظر گرفته شده است؟ (ذکر ۵ مورد) (۱ نمره)

۳. جستجوی دودویی و جستجوی ترتیبی را با هم مقایسه کنید. (۱ نمره)

۴. روش ادغام موازنه شده را در مورد مرتب‌سازی فایلهای روی نوار شرح دهید. (۱/۵ نمره)

۵. مراحل ایجاد درخت B از مرتبه چهار مربوط به دنباله زیر را نمایش دهید. (۱/۵ نمره)

C S D T A M P I B W N G U R K E H O L J Y Q Z F X V

۶. دو مزیت مهم ساختار درخت B^+ را نسبت به ساختار B ذکر کنید. (۰/۵ نمره)

۷. ۵ روش درهم سازی را نام ببرید. (۱ نمره)