

نام درس: ذخیره و بازیابی اطلاعات

رشته تحصیلی: گرایش: علوم کامپیوتر - ریاضی

کد درس: ۲۶۳۱۶۷ - ۲۶۱۲۵۰

تعداد سؤال: ۳۰ نیمی تکمیلی -- تشریحی ۸

زمان امتحان: تستی و تکمیلی ۷۵ دقیقه تشریحی ۹۰ دقیقه

تعداد کل صفحات: ۵

۱- در کدام گزینه خصوصیات حافظه در معنای عام آن آمده است؟

الف- نشانی پذیری، ظرفیت، نرخ انتقال، دستیابی پذیری، خواندن و نوشتن

ب- تعداد نوک خواندن/نوشتن، زمان دستیابی، نرخ انتقال و ظرفیت

ج- ظرفیت، زمان دستیابی، نرخ انتقال و تعداد رویه ها

د- ظرفیت، دستیابی پذیری، خواندن و نوشتن

۲- منظور از پک (Pack) چیست؟

الف- دیسک های چند صفحه ای

ب- نوارهای دو رویه ای

ج- دیسک های تک صفحه ای

د- نوارهای دورویه ای صوتی

۳- منظور از زمان استوانه جویی (Seek time) چیست؟

الف- زمانی است که سپری می شود تا آغاز داده مورد نظر در اثر دوران دیسک به زیر نوک R/W برسد.

ب- زمانی است که سپری می شود تا نوک R/W به استوانه مورد نظر برسد.

ج- زمانی است که سپری می شود تا نوک R/W به استوانه مورد نظر و به ابتدای داده مورد نظر در اثر دوران دیسک برسد.

د- زمانی است که صرف طی کردن کل استوانه های دیسک در یک سازمان دهی مجدد می شود.

۴- منظور از رکورد منطقی چیست؟

الف- رکورد در محیط ذخیره سازی

ب- رکورد در سطح برنامه کاربر

ج- رکورد در سطح انتزاعی

د- هریک از موارد فوق می تواند صحیح باشد.

۵- کدام یک از موارد زیر از دلایل متغیر شدن طول رکورد نمی باشد؟

الف- طول فیلدها متغیر باشد.

ب- تعداد صفات خاصه برای نمونه های مختلف موجودیت متفاوت باشد.

ج- پدیده فایل توزیع شده وجود داشته باشد.

د- پدیده فقره اطلاع تکرار شونده وجود داشته باشد.

۶- مفهوم فایل در معنای عام خود دارای کدام دسته از ویژگی های زیر است؟

الف- پایایی (Resistance)، اشتراکی بودن (Shared) و زمان دسترسی (Access time)

ب- ساختار ذخیره سازی، پایایی و اندازه بزرگ

ج- اندازه بزرگ، پایایی و اشتراکی بودن

د- زمان دسترسی، کلید داخلی (Internal key) و کلید خارجی (External key)

۷- کمترین مقدار داده که در یک عمل خواندن/نوشتن (I/O) توسط سیستم فایل بین درون و برون ماشین مبادله می شود، چیست؟

الف- رکورد

ب- شیار

ج- باکت

د- بلاک

۸- اگر بیش از یک رکورد همجوار، هم طول باشند و یا طول رکوردها مضربی از یک عدد باشد کدام یک از تکنیک های تعیین

محدوده رکورد در بلاک می تواند صرفه جویی بیشتری در حافظه مصرفی داشته باشد؟

الف- درج نشانگر پایان رکورد

ب- درج طول در بخش غیر داده ای رکورد

ج- ایجاد جدول مکان نما

د- ایجاد جدول طول ها

نام درس: ذخیره و بازیابی اطلاعات

رشته تحصیلی: گرایش: علوم کامپیوتر - ریاضی

کد درس: ۲۶۳۱۶۷ - ۲۶۱۲۵۰

تعداد سؤال: نسی ۳۰ تکمیلی -- تشریحی ۸

زمان امتحان: تستی و تکمیلی ۷۵ دقیقه تشریحی ۹۰ دقیقه

تعداد کل صفحات: ۵

۹- کدامیک از موارد زیر به عنوان معایب استفاده از چگالی لود اولیه می باشد؟

الف- ایجاد حافظه هرز و سبکبار شدن (Underload) بلاک ها.

ب- پرهزینه شدن عملیات درج و ایجاد حافظه هرز.

ج- پرهزینه شدن عملیات شیفت درون بلاکی و درج رکورد بهنگام سازی شده.

د- هیچکدام.

۱۰- در مدل بافرینگ ساده (Single Buffering)، در چه صورت همروندی عملیات CPU و عملیات پردازشگر I/O می تواند تا حدی امکان پذیر باشد؟

الف- اسلوب انتقالی و فایل بلاک بندی نشده باشد.

ب- اسلوب مکان نمایی و فایل بلاک بندی شده باشد.

۱۱- هرگاه سیستم آدرس حاوی رکورد ذخیره شده بر روی دیسک را به نحوی داشته باشد می تواند آنرا مستقیم بخواند، در این حالت مدت زمان خواندن مستقیم یک بلاک چه اندازه است؟

الف- b_{tt} ب- $s+r+b_{tt}$ ج- $s+2r+b_{tt}$ د- $s+r+2b_{tt}$

۱۲- با فرض آنکه $B=2000$ byte ، $T_f=10$ و زمان یک دور کامل دیسک برابر 0.02 میلی ثانیه باشد با در نظر گرفتن این شرط که زمان خواندن بلاک، بزرگتر از زمان پردازش آن است، نرخ انتقال واقعی، در مدل بافرینگ ساده و درهم خوانی برابر با چند byte/sec است؟

الف- 500 ب- 1000 ج- 2000 د- 4000

۱۳- یکی از تکنیکهای کاهش زمان درنگ دوران، پراکنده خوانی می باشد در این صورت متوسط زمان درنگ دوران برابر است با:

الف- دوبرابر زمان انتقال بلاک

ب- زمان انتقال بلاک

ج- نصف زمان انتقال بلاک

د- یک چهارم زمان انتقال بلاک

۱۴- کدامیک از موارد زیر از تکنیک های کاهش افزونگی طبیعی نمی باشد؟

الف- استفاده از کدها فم

ب- استفاده از تکنیک ماتریس بیتی

ج- استفاده از تکنیک ثبت تراکنش ها

د- استفاده از تکنیک برش از آغاز

۱۵- چنانچه بین رکورد فعلی و رکورد بعدی هیچ ارتباط ساختاری وجود نداشته باشد آنگاه همواره...

الف- $T_N = \frac{1}{2} T_F$ ب- $T_N = T_F$

ج- $T_N = 2T_F$ د- بازیابی رکورد بعدی اساسا ناممکن است.

۱۶- در فایل با ساختار پایل (برهم)، T_I (زمان مورد نیاز برای درج یک رکورد) برابر است با:

الف- $s+r+2b_{tt}$ ب- $s+r+b_{tt}$ ج- $s+2r+b_{tt}$ د- $s+3r+b_{tt}$

۱۷- معایب شاخص بندی در ساختارهای شاخص دار بویژه در حالت چند شاخص کدام است؟

الف- مصرف حافظه برای ایجاد شاخص (ها) و فزونکاری (Overhead) در عملیات ذخیره سازی

ب- فزونکاری در عملیات ذخیره سازی و غیر متراکم شدن فایل

ج- مصرف حافظه برای ایجاد شاخص (ها) و گم شدن کلید اصلی

د- هر سه گزینه صحیح است.

نام درس: ذخیره و بازیابی اطلاعات

رشته تحصیلی: گرایش: علوم کامپیوتر - ریاضی

کد درس: ۲۶۳۱۶۷ - ۲۶۱۲۵۰

تعداد سؤال: نسی ۳۰ تکمیلی -- تشریحی ۸

زمان امتحان: تستی و تکمیلی ۷۵ دقیقه تشریحی ۹۰ دقیقه

تعداد کل صفحات: ۵

۱۸- تکنیک های درج سرریزی در فایل ترتیبی شاخص دار کدامند؟

الف- درج در اولین بلاک جادار ناحیه سرریزی و درج با جابجایی

ب- درج در آخرین بلاک جادار ناحیه سرریزی و درج بدون جابجایی

ج- درج در ناحیه تصادم

د- درج با نامه نگاری ۱۰ تایی به باکت سطح بالا

۱۹- معایب فایل ترتیبی شاخص دار در کدام گزینه آمده است؟

الف- عدم تقارن، پویا بودن شاخص و مسئله درج سرریزی ها

ب- عدم تقارن، ایستا بودن شاخص و مسئله درج سرریزی ها

ج- مسئله درج سرریزی ها و ایستا بودن شاخص

د- عدم تقارن و پویا بودن شاخص

۲۰- وقتی بر روی تمامی صفات خاصه فایل، شاخص داشته باشیم اصطلاحاً به فایل چه می گویند؟

الف- کاملاً اندیسی (Indexed) ب- کاملاً وارون (Inverted) ج- برهم (پایل) د- هیچکدام

۲۱- کدام یک از موارد زیر برای فایل مستقیم (Direct) درست است؟

الف- فایل مستقیم درهم است و احتمال تصادف در این فایل وجود ندارد.

ب- پردازش سریال فایل مستقیم ممکن است.

ج- اگر رکورد در آدرس طبیعی خود باشد بایک عمل I/O واکنشی می شود.

د- همه موارد فوق صحیح است.

۲۲- در فایل مستقیم اگر تعداد رکوردها برابر 1000 و کلیدها اعداد 6 رقمی مثبت باشند فاکتور لود برابر است با: (بدون در نظر

گرفتن تابع مبدل)

الف- $\frac{1}{20000}$ ب- $\frac{1}{10000}$ ج- $\frac{1}{1000}$ د- $\frac{1}{100}$

۲۳- در مورد توابع مبدل کدام گزینه صحیح است؟

الف- توابع دارای قطعیت (Deterministic) متعلق به توابع احتمالاتی (Probabilistic) است.

ب- توابع درهم ساز (Hashing) متعلق به توابع قائل به نظم است.

ج- توابع احتمالاتی متعلق به توابع دارای قطعیت است.

د- توابع درهم ساز متعلق به توابع احتمالاتی است.

۲۴- در ساختار فایل مستقیم برای بازیابی رکورد بعدی کدام گزینه صحیح است؟

الف- $T_N = 2T_F$ ب- $T_N = \frac{1}{2}T_F$ ج- $T_N = \frac{1}{3}T_F$ د- $T_N = T_F$

۲۵- در فایل مستقیم گسترش پذیر (Extendable Direct)، حداکثر دفعات دستیابی لازم به دیسک برای واکنشی یک رکورد چه

تعداد خواهد بود؟

الف- ۲ ب- ۱ ج- ۳ د- ۵

نام درس: ذخیره و بازیابی اطلاعات

رشته تحصیلی: گرایش: علوم کامپیوتر - ریاضی

کد درس: ۲۶۳۱۶۷ - ۲۶۱۲۵۰

تعداد سؤال: ۳۰ نیمی تکمیلی -- تشریحی ۸

زمان امتحان: تستی و تکمیلی ۷۵ دقیقه تشریحی ۹۰ دقیقه

تعداد کل صفحات: ۵

۲۶- برای ساختار درختی جستجوی دودویی با عمق n ، متوسط تعداد دستیابی مستقیم برای واکشی رکورد برابر است با:

الف- $\frac{n+1}{2}$ ب- $\frac{n-1}{2}$ ج- $\frac{n}{2}$ د- n

۲۷- در عملیات حذف یک رکورد از درخت جستجوی دودویی اگر گره حذف شدنی دارای دو فرزند باشد، پس از حذف گره مربوطه، رکورد بعدی آن طبق نظم، باید جایگزین آن رکورد شود، رکورد بعدی طبق نظم کدام است؟

- الف- رکورد سمت چپ ترین در زیر درخت سمت چپ گره حذف شدنی است.
ب- رکورد سمت راست ترین در زیر درخت سمت چپ گره حذف شدنی است.
ج- رکورد سمت راست ترین در زیر درخت سمت راست ترین گره حذف شدنی است.
د- الف و ج صحیح است.

۲۸- ساختاری مبتنی بر درخت داریم که بخش داده‌ای فایل در بخش شاخص آن ادغام شده است این ساختار چه نام دارد؟

- الف- $B - tree$ ب- فایل درختی ج- $B^+ - tree$ د- درخت نخ کشی شده

۲۹- تفاوت آشکار ساختار درخت $k-d$ با ساختار درخت جستجو در چیست؟

- الف- عمق درخت ب- سطح درخت
ج- یکسان نبودن فیلد کلید در سطوح مختلف د- یکسان بودن فیلد کلید در سطوح مختلف

۳۰- اگر نشانوند کلمات با طول دلخواه و متغیر باشد (مثلا لغات دیکشنری) کدام یک از ساختارهای زیر مناسب تر است؟

- الف- درخت نخ کشی شده ب- $B - tree$ ج- Trie د- فایل درختی

سؤالات تشریحی

نکات مهم:

نکته اول: سوالات تشریحی در دو بخش دسته بندی شده است. نکته دوم: بایستی در هر بخش فقط به دو سوال پاسخ دهید.

بخش اول: از چهار سوال زیر فقط به دو سوال پاسخ دهید. (هر سوال ۱/۵ نمره دارد)

۱- هرم مربوط به سلسله مراتب حافظه ها را رسم کرده و از نظر سرعت دستیابی، قیمت و حجم ذخیره سازی مقایسه کلی کنید.

۲- معماری سیستم فایل را بطور کامل رسم کنید.

۳- برای فایل ساختار پایل (برهم) هریک از موارد زیر را محاسبه کنید.

الف- زمان واکشی یک رکورد (T_F) ب- بازیابی رکورد بعدی (T_N)

نام درس: ذخیره و بازیابی اطلاعات

رشته تحصیلی: گرایش: علوم کامپیوتر - ریاضی

کد درس: ۲۶۳۱۶۷ - ۲۶۱۲۵۰

تعداد سؤال: نسی ۳۰ تکمیلی -- تشریحی ۸

زمان امتحان: تستی و تکمیلی ۷۵ دقیقه تشریحی ۹۰ دقیقه

تعداد کل صفحات: ۵

۴- زمان های لازم برای خواندن تمام فایل ($T_{X_{seq}}$, $T_{X_{ser}}$) و سازماندهی مجدد فایل (T_Y) را برای فایل ترتیبی محاسبه کنید.

بخش دوم: از چهار سوال زیر فقط به دو سوال پاسخ دهید. (هر سوال 1/75 نمره دارد)

۵- تکنیک بلاکبندی رکوردهای با طول ثابت را رسم کرده و میزان حافظه هرز به ازاء یک رکورد (W_R) را در رسانه ذخیره سازی نوار بدست آورید.

۶- دو تکنیک تداخل بلاک ها (Interleaving) و تغییر مکان نقطه آغاز شیارها (Track Staggering) را برای کاهش زمان درنگ دوران بطور کامل شرح دهید.

۷- از توابع درهم ساز زیر، هریک را به همراه مثالی شرح دهید.

الف- انتخاب ارقام میانی مربع کلید (Mid Square) ب- تقسیم (Division)

۸- عملیاتهای درج زیر را مرحله به مرحله به صورت متوالی برای یک ساختار B_2 (B-tree مرتبه ۲) رسم کنید.

مرحله ۱- درخت تهی مرحله ۲- درج رکورد با کلید ۴

مرحله ۳- درج رکورد با کلیدهای ۷، ۲۲ و ۲۱ مرحله ۴- درج رکورد با کلید ۳۵

مرحله ۵- درج رکورد با کلیدهای ۱۲، ۲۳، ۳۹ و ۱۶