



تمرین کامپیوتری سوم



سیستم‌های عامل - پاییز ۱۳۹۹

گزارش کار

دانشکده مهندسی برق و کامپیوتر

نام و نام خانوادگی: محمد علی زارع

تاریخ:

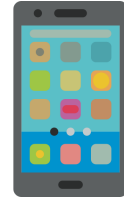
استاد:

۸۱۰۱۹۷۶۲۶

دکتر مهدی کارگهی

2	مقدمه
3	پیاده‌سازی سری
3	سوال اول
3	سوال دوم
3	جدول اول
3	پیاده‌سازی چندریسه‌ای
3	سوال سوم
4	سوال چهارم
4	سوال پنجم
4	جدول دوم

مقدمه



در این تمرین به تحلیل داده‌هایی که از مشخصات و قیمت فروش گوشی‌های موبایل جمع‌آوری شده‌است پرداخته شده است. در ابتدا برنامه اقدام به خواندن و تجزیه مجموعه داده¹ی ارائه شده می‌کند و آنها را در حافظه خود ذخیره می‌کند. پس از استخراج داده‌ها و ویژگی‌های آنها، برنامه اقدام به نرمال‌سازی² داده‌ها و در نهایت اقدام به تعیین طبقه قیمتی گوشی‌ها می‌کند. این تمرین به دو روش این مسئله پیاده‌سازی شده است که در ادامه گزارش، نتایج حاصل آمده است.



¹ Dataset

² Data Normalization

پیاده‌سازی سری

سوال اول

چرا برای پیاده‌سازی یک برنامه بصورت چندریسه‌ای، بهتر است ابتدا این برنامه بصورت سری پیاده‌سازی شود؟ تا بخش‌های قابل موازی‌سازی مشخص شود و تسک‌ها شناسایی شوند.

سوال دوم

با بررسی زمان اجرای بخش‌های مختلف برنامه، Hotspot³ های برنامه را مشخص کنید. خواندن از فایل، نرمالایز کردن، محاسبه اسکور برای پیشبینی

جدول اول

زمان‌های اجرای ۶ اجرای متوالی از برنامه و میانگین آن‌ها را بازای ورودی نمونه‌ای که در شرح تمرین آمده است، در جدول زیر بیاورید.

میانگین	اجرای ششم	اجرای پنجم	اجرای چهارم	اجرای سوم	اجرای دوم	اجرای اول
۰.۰۲۹	۰.۰۳۶	۰.۰۳۷	۰.۰۳۵	۰.۰۱۴	۰.۰۳۳	۰.۰۱۹

³ توابعی که در برنامه‌تان بیشترین زمان اجراها را به خود اختصاص می‌دهند.

پیاده‌سازی چندریسه‌ای

سوال سوم

اگر هنگام موازی‌سازی برنامه به زمان اجرای بیشتری نسبت به حالت سری برخورد کنید، چه رویکردهایی را برای کاهش زمان اجرا و استفاده حداکثری از موازی‌سازی پیش می‌گیرید؟

- تعداد ریشه‌ها را طوری تغییر می‌دهیم که سربار ساخت و مدیریت آن‌ها بیشتر از زمان صرفه‌جویی شده نشود.
- در تقسیم کار و تسک‌ها تجدید نظر می‌کنیم.
- بخش‌هایی که ریشه‌ها باید منتظر انجام کار ریشه دیگری بمانند (مثلا mutex lock ها) را طوری تغییر می‌دهیم که یا این وابستگی کاهش یابد یا در هنگام منتظر بودن کار مفید دیگری انجام شود.

سوال چهارم

در هنگام پیاده‌سازی این بخش، به چه چالش‌هایی برخورد کردید و بیان کنید که به چه صورت آن‌ها را رفع کردید.

- اشتراک گذاشتن داده‌ها میان ریشه‌ها
 - با استفاده از تعریف متغیر گلوبال که همه ریشه‌ها بتوانند از آن بخوانند.
- استفاده از نتیجه ریشه‌ای در ریشه دیگر. مثلا پیدا کردن مقدار مینی‌م و ماکسی‌م در همه داده‌ها برای نرمالایز کردن.
 - استفاده از join. به این شکل که جواب همه ریشه‌ها در تابع مین جمع‌آوری می‌شود و با محاسبه جواب کلی، دوباره ریشه‌ها را با استفاده از این جواب فعال می‌کنیم.
 - امکان استفاده از متغیری گلوبال و mutex lock نیز وجود دارد.

سوال پنجم

با توجه به تجربه‌ای که در پیاده‌سازی این تمرین بدست آوردید، به نظر شما در چه مواقعی از قفل⁴ در یک طراحی چندریسه‌ای ضروری است؟ تاثیر استفاده از قفل‌ها را بر روی کارایی⁵ سامانه بیان کنید.

⁴ Lock

⁵ Performance

هنگامی که چندین ریشه ممکن است یک داده را تغییر دهند و یا یک ریشه از یک داده که ریشه‌های دیگری آن را تغییر می‌دهند بخواهد استفاده کند.

اگر تعداد دفعات گرفتن و آزاد کردن قفل‌ها در ریشه‌ها زیاد باشد می‌تواند باعث کاهش کارایی شود چون ریشه‌ها باید منتظر آزاد شدن قفل توسط ریشه‌های دیگر شوند.

جدول دوم

زمان‌های اجرای ۶ اجرای متوالی از برنامه و میانگین آن‌ها را بازای ورودی نمونه‌ای که در شرح تمرین آمده است، در جدول زیر بیاورید.

میانگین	اجرای ششم	اجرای پنجم	اجرای چهارم	اجرای سوم	اجرای دوم	اجرای اول
۰.۰۱۵	۰.۰۲۰	۰.۰۱۴	۰.۰۹	۰.۰۱۴	۰.۰۲۳	۰.۰۱۳

میزان تسریع ($\frac{Serial\ Time}{Parallel\ Time}$) برنامه نسبت به حالت سری را در زیر بیاورید.

میزان تسریع	میانگین زمان اجرای موازی	میانگین زمان اجرای سری
۱.۹۳	۰.۰۱۵	۰.۰۲۹