



تمرین کامپیوتری سوم



سیستم‌های عامل - پاییز ۱۳۹۹

گزارش کار

دانشکده مهندسی برق و کامپیوتر

نام و نام خانوادگی:

استاد:

فاطمه سیددباغی

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 شده است

پردازش 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 اوّل) چرا برای پیاده‌سازی یک برنامه بصورت چندریسه‌ای، بهتر است ابتدا این برنامه بصورت سری

پیاده‌سازی شود؟

برای بهتر مشخص شدن قسمتی‌هایی از برنامه که امکان پیاده‌سازی آن‌ها به صورت موازی وجود دارد و پیاده‌سازی با این روش باعث تسریع برنامه میشود.

سوال دوم) با بررسی زمان اجرای بخش‌های مختلف برنامه، $topstoH^3$ های برنامه را مشخص کنید.

1) store_train_data که فایل train را خوانده و در یک وکتور دو بعدی ذخیره میکند. 2) normalize_data که حلقه ای روی تمام خانه های موجود در train_data دارد و آن‌ها را نرمالایز میکند. 3) predict_price که حلقه ای روی train_data دارد و مقدار خانه های موجود در دیتا را در weight ضرب میکند. 4) calc_accuracy که حلقه ای روی تمامی داده های هدف دارد و با مقایسه با مقدار واقعا درستی نتیجه را حدس میزند.

جدول اوّل) 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 اوّل |
|----------|-----------|------------|-------------|-----------|-----------|------------|
| 0.043627 | 0.045748 | 0.047292 | 0.042787 | 0.040324 | 0.041579 | 0.044034 |

³ توابعی که در برنامه‌تان بیشترین زمان اجرا 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؟

تغییر تعداد ریشه ها و انتخاب تعداد ریشه ها متناسب با تعداد هسته های موجود.

سوال چهارم) در هنگام پیاده‌سازی این بخش، به چه چالش‌هایی برخورد کردید و بیان کنید که به چه صورت

آنها را رفع کردید.

علاوه بر مورد اشاره شده در سوال سه هنگام دسترسی به متغیرهای گلوبال احتمال برخورد با مشکل بود که با تعریف کردن متغیرهای

گلوبال بصورت وکتوری به سبب تعداد ریشه ها از مشکل‌هایی که ممکن بود در حافظه پیش بیاید جلوگیری کردیم (هر ریشه با توجه

به tid خود به خانه ی مخصوص به خود دسترسی داشت) و نتیجه‌های نهایی را در main با انجام محاسبات روی متغیرهای گلوبال

پس از اتمام ریشه ها انجام دادیم.

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.

هنگامی که چند ریشه بخوانند با هم تغییری در قسمتی از حافظه ایجاد کنند. قفل تضمین میکند که هنگامی که یک ریشه در یک

بخش بحرانی از کد است ریشه دیگر وارد آن بخش نشود تا زمانی که قفل توسط ریشه ی اول آزاد شود.

⁴ Lock

⁵ Performance

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 |
|----------|-----------|------------|-------------|-----------|-----------|-----------|
| 0.024404 | 0.021246 | 0.023022 | 0.024634 | 0.027466 | 0.022714 | 0.027345 |

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 |
|-------------|--------------------------|------------------------|
| 1.787698 | 0.024404 | 0.043627 |