



if a word is found in several places in text, we have to delete it from input file and just note the first appearance of file

شرح پروژه

در این پروژه هدف ما این است که یک موتور جستجوی متن پیاده‌سازی کنیم.

موتور جستجو با گرفتن یک فایل و لیستی از کلمات به عنوان ورودی، فایل‌ی شامل نتایج جستجو را ایجاد می‌کند.

read big file /
read input /
create output /

جستجو در فایل‌های بزرگ زمان زیادی خواهد برد در نتیجه برای انجام این کار از چند نخ استفاده می‌کنیم. البته باید در نظر داشت که تعداد زیاد نخ‌ها می‌تواند تاثیر منفی گذاشته و سرعت ما را کاهش دهد؛ بنابراین برنامه باید حداکثر از ۴ نخ استفاده کند. در نظر داشته باشید که میزان متنی که هر نخ بررسی می‌کند باید با هم برابر باشد.

هر نخ بعد از پیدا کردن یکی از کلمات موردنظر باید آن را در فایل خروجی بنویسد؛ در نتیجه‌ی رقابت برای نوشتن، وضعیت *Race Condition* به وجود خواهد آمد که برای حل این مشکل می‌بایست برنامه را در دو حالت پیاده‌سازی کنید. یک حالت با استفاده از *Mutex lock* و حالت دیگر با استفاده از *Semaphore* انجام می‌شود تا مشکل همگام‌سازی حل گردد.

فایل خروجی شما باید شامل موارد زیر باشد:

- موقعیت هر کلمه در متن ورودی (شماره خط)
- شماره نخ‌ی که آن کلمه را پیدا کرده
- زمان پیدا شدن کلمه در متن
- زمان نوشتن خروجی در فایل

فایل تحویلی شما باید شامل:

- هر دو حالت پیاده‌سازی *Mutex Lock* و *Semaphore* به همراه مستندات (نمره اضافه) باشد.

شرح نمره اضافه

- در فایل‌ی جداگانه به نام مستندات (*PDF*) زمان اجرای برنامه‌ها در حالتی که برنامه با یک نخ اجرا می‌شود و دو حالت دیگر که برای حل مشکل همگام‌سازی پیاده‌سازی شده است را مقایسه کنید و دلیل تفاوت زمان اجرا در حالت‌های مختلف را شرح دهید.
- همچنین می‌توانید زمان اجرای برنامه را هنگامی که با تعداد زیادی نخ (مثلاً ۱۰ نخ) اجرا می‌شود را بررسی نموده و گزارش دهید.
- استفاده صحیح و تدریجی از *Git* نیز دارای نمره اضافه می‌باشد.

نکات

- ۱- پروژه قابل پیاده‌سازی با زبان‌های *C.CPP.Java* می‌باشد.
 - ۲- انجام و ارائه پروژه به صورت انفرادی است.
 - ۳- در صورت مشاهده شباهت مابین پروژه‌های تحویلی دانشجویان و اثبات تقلب، نمره داده‌شده **-۱۰۰** خواهد بود.
 - ۴- فایل تحویلی را مانند مثال زیر ارسال کنید:
- FirstName_LastName_StudentID_OS_PR2.zip

مهلت تحویل: جمعه ۱۳ خرداد ۱۴۰۱ - ساعت ۲۳:۵۹