

## پروژه دوم درس ساختمان داده ها

# روش های جستجو

سررسید تحویل : ۱۳۸۶/۴/۱۳

## ۱ صورت مسئله

متنی در اختیار داریم که دارای تعدادی غلط املائی است و می خواهیم این غلط ها را با توجه به لغت نامه ای که در اختیار داریم پیدا کنیم . شما باید برنامه ای بنویسید که غلط های املائی یک متن را با توجه به لغت نامه داده شده پیدا کرده و گزارش دهد.

## ۲ ساختار ورودی

لغت نامه در فایل dic.dat ذخیره شده است . سطر اول dic.dat شامل یک عدد طبیعی بین ۱ تا ۱۰۰۰۰۰۰ است که بیانگر تعداد لغات، لغت نامه می باشد . در ادامه هر سطر شامل یک لغت با حداقل ۱ و حداکثر ۳۰ حرف می باشد. متنی که باید غلط گیری شود داخل فایل text.in قرار دارد. هر سطر این فایل شامل حداقل ۱ و حداکثر ۱۰ کلمه است که با یک یا چند فاصله ی خالی از هم جدا شده اند و طول هر سطر حداکثر ۲۵۶ حرف است. هر کلمه فقط از حروف کوچک (a تا z) تشکیل شده است . متن داده شده حداکثر از ۱۰۰۰۰۰۰ سطر تشکیل شده است.

### ورودی نمونه

فایل dic.dat	فایل text.in
6 this is a test examination english	ths is a tst

## ۳ ساختار خروجی

در هر سطر فایل خروجی (error.out) باید یک غلط گزارش داده شود. این گزارش شامل شماره سطر ، شماره کلمه در آن سطر و خود کلمه می باشد. در ضمن غلط ها باید به ترتیب ( از بالای متن به پایین ) گزارش شوند.

### خروجی نمونه

فایل error.out
1 1 ths 2 1 tst

## ۴ نحوه نمره دهی

هدف از انجام این پروژه، آشنایی با الگوریتم های جستجوی کارا می باشد. به همین دلیل، نمره دهی بر اساس سرعت اجرای برنامه شما می باشد، به این صورت که یک تست بزرگ<sup>۱</sup> (!) به برنامه شما داده می شود و نمره شما بر اساس زمانی که به طول می انجامد تا برنامه شما خروجی را تولید کند، محاسبه می شود. در این رابطه به نکات زیر توجه فرمائید:

- اگر خروجی برنامه شما نادرست باشد، هیچ نمره ای به شما تعلق نمی گیرد.
- شما می توانید برنامه را با هر الگوریتم دلخواهی بنویسید. فقط به یاد داشته باشید که هر چه الگوریتم (تان) سریعتر باشد، نمره بیشتری کسب می کنید.
- برنامه شما می بایست به یکی از زبانهای C یا C++ باشد.

نحوه نمره دهی به این صورت می باشد که بعد از مرتب کردن (بر اساس زمان اجرای برنامه) برنامه هایی که خروجی صحیح تولید کرده اند، اگر برنامه شما، برنامه  $A-i$  ام در این لیست باشد، نمره شما از رابطه  $20 + 80 \times \log_n^{n+1-i}$  محاسبه می شود، که در آن  $n$  تعداد برنامه هایی است که خروجی صحیح تولید کرده اند. این رابطه تضمین می کند که در صورت صحیح بودن خروجی برنامه شما حداقل ۲۰٪ نمره سوال را کسب خواهید کرد!

## ۵ ملاحظات

- شما بایستی برنامه خود را حداکثر تا تاریخ سررسید به آدرس [ce.ds.40254@gmail.com](mailto:ce.ds.40254@gmail.com) ایمیل کنید. موضوع نامه ارسالی بایستی به صورت prj2\_STID باشد که STID شماره دانشجویی شما می باشد. فایل فشرده ارسالی بایستی شامل موارد زیر باشد:
  - فایل برنامه شما با نام prj2.cpp یا prj2.c
  - فایلی با نام prj2.pdf که حاوی توضیحاتی درمورد الگوریتم شما و نحوه پیاده سازی آن می باشد.
- با توجه به اینکه برنامه های شما به صورت خودکار تصحیح می شود، به فرمت ورودی ها و خروجی ها دقت بیشتری کنید.
- به سریعترین برنامه ها (سه برنامه اول) علاوه بر نمره پروژه، نمره تشویقی نیز تعلق می گیرد.
- به دلیل اینکه، کل زمان اجرای برنامه محاسبه می شود. در عملیات با فایل ها دقت بیشتری کنید.
- به برنامه های مشابه تا ۱۵۰٪ نمره منفی تعلق می گیرد.
- هرگونه سوال و ابهام را می توانید در بخش گفتگو صفحه ابرمتنی درس مطرح کنید.

---

<sup>۱</sup> تستی با حداکثر مقادیر ذکر شده در صورت مسئله.