

موعد تحویل: ۳ آبان ۱۳۹۳

تغییر شکل رشته‌های DNA

در داخل هسته‌ی تمام سلول‌های زنده، مولکولی به نام DNA وجود دارد که رشته‌ای است حاوی اطلاعاتی در مورد سلول. این رشته در واقع مجموعه‌ای از کدها است که سلول پس از انجام عملیاتی مانند نسخه‌برداری^۱ و تبدیل^۲ روی آن‌ها، می‌تواند کارکرد درست و حیاتی‌اش را مشخص کند.

این رشته‌ها، از واحدهای کوچک‌تری به نام اسیدهای نوکلئوتید تشکیل شده‌اند. این اسیدهای نوکلئوتید مانند حروف الفبا هستند که کدها را ایجاد می‌کنند. اسیدهای نوکلئوتیدی که در رشته‌ی DNA وجود دارند، چهار نوع هستند: آدنین، تیمین، سیتوزین و گوانین که به اختصار آن‌ها را به ترتیب با A، T، C و G نمایش می‌دهیم.

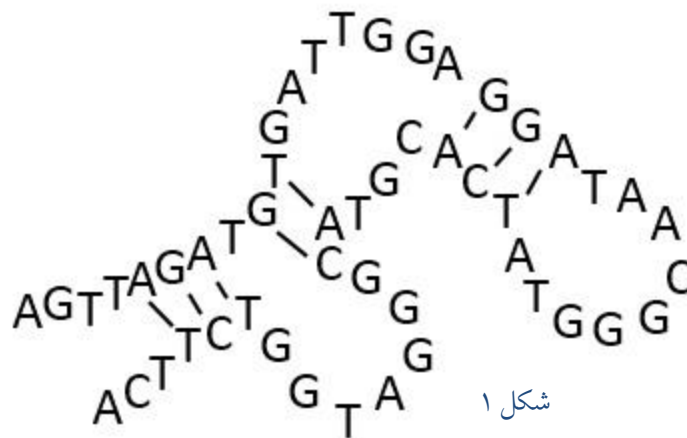
به این ترتیب قسمتی از یک رشته‌ی DNA می‌تواند مانند رشته‌ی زیر باشد:

.....ACCTGCAATGCCTCGAATCT.....

از ویژگی‌های این اسیدهای نوکلئوتید آن است که می‌توانند به صورت دو به دو (A با T و C با G) پیوندهای محکمی تشکیل دهند:

A---T C---G

توجه داشته‌باشید که زوج‌های دیگر اسیدهای نوکلئوتید با یک‌دیگر پیوندی قوی ایجاد نمی‌کنند و فقط همین دو زوج در این تمرین مد نظر ما هستند.



^۱ Transcription

^۲ Translation

با در نظر داشتن ویژگی بیان شده، قسمت‌های مختلف یک رشته از اسیدهای نوکلئوتید می‌توانند با یک‌دیگر پیوند ایجاد کنند و به این ترتیب شکل یک رشته می‌تواند مانند آنچه در شکل ۱ مشاهده می‌کنید، تغییر کند. توجه کنید که تمام پیوندهای ممکن در این شکل نشان داده نشده‌اند.

از ویژگی‌های دیگر این پیوندها آن است که به‌صرف^۳ هستند؛ اگر یک رشته به صورت ACTG داشته‌باشیم و A با T پیوند ایجاد کند دیگر امکان ایجاد پیوند بین C و G وجود ندارد. در حالی‌که در رشته‌ی ACGT، اگر A با T پیوند ایجاد کند، هم‌زمان C با G نیز می‌تواند پیوند ایجاد کند. به دو اسید نوکلئوتید که در رشته‌ی DNA با هم پیوند برقرار می‌کنند، یک جفت^۴ می‌گویند.



جفت غیرقابل قبول



جفت قابل قبول

محققان علم زیست‌شناسی با مطالعه‌ی جفت‌هایی که ممکن است در یک رشته ایجاد شود، سعی می‌کنند تا شکل احتمالی یک رشته از اسیدهای نوکلئوتید مانند DNA را تعیین کنند.

در این تمرین می‌خواهیم یک برنامه‌ی بازگشتی بنویسیم که با داشتن رشته‌ای از اسیدهای نوکلئوتید، بیشترین تعداد جفت‌های ممکن که به صورت هم‌زمان می‌توانند در یک رشته وجود داشته‌باشند را پیدا کند.

راهنمایی: بهتر است ابتدا یک تابع با امضای `int count_max_pairs(string s, int i, int j)` بنویسید که تعداد جفت‌های موجود در زیررشته‌ی `i` تا `j` رشته‌ی `s` را محاسبه می‌کند. برای محاسبه‌ی این تعداد نیز با توجه به قانون جفت‌های قابل قبول، از دیگر زیررشته‌های موجود در زیررشته‌ی `i` تا `j` استفاده کنید.

^۳ Aligned

^۴ Base pairs

ورودی استاندارد

ورودی شامل چندین مورد آزمون است. در خط اول ورودی عدد n که تعداد موردهای آزمون است، داده می‌شود. در هر یک از n خط بعدی یک رشته شامل حروف A، T، C و G که رشته‌ای از اسیدهای نوکلئوتید است، داده می‌شود.

خروجی استاندارد

برنامه‌ی شما باید بیشترین تعداد جفت‌های ممکن برای هر رشته‌ی ورودی را، در یک خط چاپ کند.

ورودی و خروجی نمونه

خروجی نمونه	ورودی نمونه
2	2
5	ACTGCG
	TTGCCCGAAGC

نحوه‌ی تحویل

فایل(های) برنامه‌ی خود را در یک پوشه به نام A4-SID قرار داده و آن را با فرمت zip آرشیو کنید و در نهایت فایلی با نام A4-SID.zip را در سایت درس آپلود کنید. (SID پنج رقم آخر شماره‌ی دانشجویی شماست. به عنوان مثال اگر شماره‌ی دانشجویی شما ۸۱۰۱۹۲۱۲۳ است، نام فایل شما باید A4-92123.zip باشد.)

بارم‌بندی

دقت کنید برای برنامه‌ای که به صورت بازگشتی نوشته نشود، نمره‌ی صفر در نظر گرفته می‌شود.

ساختار درست برنامه	تعریف درست توابع بازگشتی	کارکرد کلی برنامه	جمع کل
۱۰ نمره	۳۰ نمره	۶۰ نمره	۱۰۰ نمره

دقت کنید

- برنامه‌ی شما باید در سیستم عامل لینوکس نوشته و با مترجم g++ ترجمه شود.
- مثالی که در این تمرین به عنوان ورودی نمونه آورده شده است، مثال ساده و کوچکی است. سعی کنید با ورودی‌های مختلف برنامه‌ی خود را تست کنید.
- به فرمت و نام فایل‌های خود دقت کنید. در صورتی که هر یک از موارد گفته شده رعایت نشود، نمره‌ی صفر برای شما در نظر گرفته می‌شود.
- در صورت کشف تقلب در کل و یا قسمتی از تمرین، برای هر دو طرف نمره‌ی ۱۰۰ - منظور خواهد شد.