به نام خدا



دانشگاه صنعتی امیرکبیر

(پلی‌تکنیک تهران)

درس پردازش داده‌های حجیم

استاد حقیرچهرقانی

تمرین دوم

علیرضا مازوچی

۴۰۰۱۳۱۰۷۵

بخش اول: سوالات تشریحی

سوال ۱

در جدول زیر فاصله دو‌به‌دوی هر جفت کلمه آورده شده است:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | hen | she | he | then | when | مجموع |
| hen | 0 | 2 | 1 | 1 | 1 | ۵ |
| she | ۲ | ۰ | ۱ | ۳ | ۳ | ۹ |
| he | ۱ | ۱ | ۰ | ۲ | ۲ | ۶ |
| then | ۱ | ۳ | ۲ | ۰ | ۲ | ۸ |
| when | ۱ | ۳ | ۲ | ۲ | ۰ | ۸ |

الف) متناسب با مجموع فواصل کلمه hen مرکز این خوشه خواهد بود.

ب) کلمه she با فاصله ۲ تا کلمه hen دارای بیشترین فاصله تا خوشه است.

ج) بیشترین فاصله‌ای که بین جفت کلمات در جدول فواصل وجود دارد فاصله ۳ است؛ پس انسجام خوشه برابر با ۳ خواهد بود.

سوال ۲

1. غلط؛ اولین بلاک شامل دو عدد یک است. اولین بلاک قطعا یک عدد یک دارد.
2. غلط؛ چهارمین بلاک با صفر شروع شده است. بلاک‌ها با یک شروع می‌شوند.
3. غلط؛ دومین بلاک با صفر شروع شده است. بلاک‌ها با یک شروع می‌شوند.
4. صحیح

سوال ۳

الف) 256

فرض کنید داخل پنجره 511 عدد یک باشد و مابقی صفر باشد. اگر از هر نوع بلاک دقیقا یک عدد داشته باشیم. بلاک‌های ۱-تایی، ۲-تایی، ۴-تایی تا 256-تایی خواهیم داشت که مجموع تعداد یک‌های آن برابر با خواهد بود. پس امکان دارد بلاک 256 تایی ظاهر شود.

من ادعا می‌کنم امکان ندارد بلاک‌های بزرگتر از 256 ظاهر شود و برای اثبات از برهان خلف استفاده می‌کنم. فرض کنید حداقل یک بلاک بزرگ‌تر از ۲۵۶ وجود داشته باشد. با توجه به آنکه از هر نوع بلاک ۱-تایی تا بزرگترین بلاک باید یک یا دو تای آن را داشته باشیم و از آنجایی که وجود دارد بلاکی که بزرگتر از ۲۵۶ باشد و اندازه بلاک‌ها باید توان ۲ باشد پس حداقل یک بلاک ۱-تایی، حداقل یک بلاک ۲-تایی،... حداقل یک بلاک ۲۵۶-تایی و حداقل یک بلاک ۵۱۲-تایی خواهیم داشت. با این حساب حداقل عدد خواهیم داشت که طبیعتا در پنجره هزارتایی جا نمی‌شود. به تناقض می‌خوریم و حکم ثابت می‌شود.

با مثال و اثبات ارائه‌شده می‌توان نتیجه گرفت که بزرگترین بلاکی که امکان ظاهر شدن دارد ۲۵۶ است.

ب) 256

امکان ندارد بزرگ‌ترین بلاک ۱۲۸-تایی باشد. برای اثبات از برهان خلف کمک می‌گیرم. فرض کنید امکان‌پذیر باشد. پس در این شرایط حداکثر دو بلاک ۱-تایی، حداکثر دو بلاک ۲-تایی،... و حداکثر ۲ بلاک ۱۲۸-تایی خواهیم داشت که تعداد یک‌های موجود در بلاک‌ها در این حالت حداکثر برابر با خواهد بود. پس به تناقض می‌خوریم و حکم اثبات می‌شود.

باتوجه به این اثبات و نتیجه قسمت قبل، اندازه بزرگترین بلاک در این حالت دقیقا برابر با ۲۵۶ است. به طور جانبی می‌توان نتیجه گرفت که بزرگترین بلاک حداقل 256-تایی است.

ج) در قسمت الف یک حالت کلی بیان شد که برای این قسمت قابل استناد است. برای دقیق‌تر شدن مثال می‌توانید فرض کنید تمام ۵۱۱ عدد یک همگی در ابتدا آمده‌اند و سپس ۵۱۱ صفر آمده است. ترتیب بلاک‌بندی هم مطابق با آنچه که گفته شد در نظر گرفته می‌شود.