پاسخ سؤالات کوئیز اول

مسئله‌ی ۱. پاسخ سؤال اول

در اینجا پاسخ با جزئیات کامل برای بررسی دقیق هر دو الگوریتم آورده شده است. پاسخ‌های درستی که کوتاه‌تر باشند، نیز نمره‌ی کامل دریافت خواهند کرد. همچنین برای الگوریتم اول، به دست آوردن کران پایین کفایت می‌کند و به دست آوردن تتا در این پاسخنامه نیز به جهت تحلیل دقیق این الگوریتم است و برای مطالعه‌ی بیشتر شما دانشجویان این درس است.

الگوریتم اول: با توجه به خطی بودن رابطه‌ی امیدریاضی، میانگین تعداد مقایسه‌ها در هر دورِ حلقه‌ی بیرونی را حساب کرده و در انتها آن‌ها را با یکدیگر جمع می‌کنیم. پس در ابتدا میانگین تعداد مقایسه‌ها را در دورِ اول حلقه‌ی بیرونی یا به عبارتی زمانی که i برابر یک است را حساب می‌کنیم. رابطه امید ریاضی هست:

$$E[C_۱] = \Sigma_{i=۱}^{n-۱} i p(C_۱ = i)$$

که در آن، $C_۱$ تعداد مقایسه‌ها است و همچنین بیشترین مقدار آن می‌تواند $n - ۱$ باشد (زمانی که اگر کوچکترین عنصر در ابتدا دنباله آمده باشد). حال فرض کنید اولین نابجایی در مقایسه‌ی kام اتفاق بیفتد. می‌خواهیم احتمال آن، یعنی $p(C_۱ = k)$ را حساب کنیم. این‌کار را با استفاده از حاصل تقسیم تعداد جایگشت‌های مطلوب بر تعداد کل جایگشت‌ها انجام می‌دهیم. اگر نابجایی اول در مقایسه‌ی kام اتفاق بیفتد، یعنی عضوی که در حال مقایسه شدن با بقیه است (عنصری که توسط حلقه‌ی بیرونی مشخص می‌شود و در اینجا عضو ابتدای دنباله است) در $k - ۱$ مقایسه‌ی قبلی، کوچک‌تر بوده است. این یعنی عنصر $k + ۱$ام کوچکترین عنصر در بین $k + ۱$ عنصر ابتدایی و عنصر اول، دومین کوچکترین عنصر در بین $k + ۱$ عنصر ابتدایی است. اگر چنین شرایطی برقرار باشد می‌دانیم که حتما در مقایسه‌ی kام به اولین نابجایی (با توجه به الگوریتم) خواهیم رسید. با این حساب با مشخص شدن این دو جایگاه در بین $k + ۱$ عنصر ابتدایی، تعداد جایگشت‌ها را می‌شماریم. که برابر با عبارت زیر است:

$$\frac{\binom{n}{k+۱}(k-۱)!(n-k-۱)!}{n!} = \frac{۱}{k(k+۱)}$$

حال رابطه‌ی امیدریاضی را بازنویسی می‌کنیم.

$$E[C_۱] = \Sigma_{i=۱}^{n-۱} \frac{i}{i(i+۱)} = \Sigma_{i=۱}^{n-۱} \frac{۱}{i+۱} \sim log(n)$$

حال به سراغ تعداد دور دوم حلقه‌ی بیرونی می‌رویم. زمانی که این دور شروع می‌شود، می‌دانیم که در دور اول هیچ نابجایی‌ای وجود نداشته. پس تا به الان $n - ۱$ مقایسه‌انجام داده‌ایم. همچنین با این اوصاف، متوجه شده‌ایم کوچک‌ترین عضو دنباله در ابتدا آمده است. رابطه‌ی امیدریاضی برای این دور از حلقه‌ی بیرونی را می‌نویسیم:

$$E[C_۲] = \Sigma_{i=n}^{۲n-۳} ip(C_۲ = i) = \Sigma_{i=۱}^{n-۲}(n-۱+i)p(C_۲ = n-۱+i)$$

حال مسئله معادلِ این می‌شود که یک عضو از دنباله کم شده است با این تفاوت که به احتمال $\frac{۱}{n}$ این عضو در ابتدای دنباله می‌آید. لذا می‌توانیم از رابطه‌ی قبلی و احتمال جایگشت‌های مطلوب استفاده نماییم. پس داریم:

$$\Sigma_{i=۱}^{n-۲} \frac{n-۱+i}{i(i+۱)n}$$

حال سعی می‌کنیم برای این کسر یک کران پیدا کنیم:

$$E[C_۲] = \Sigma_{i=۱}^{n-۲} \frac{n-۱+i}{i(i+۱)n} \leqslant \frac{۱}{n}\Sigma_{i=۱}^{n-۲} \frac{۲n}{i(i+۱)} = ۲\Sigma_{i=۱}^{n-۲} \frac{۱}{i(i+۱)} = ۲\Sigma_{i=۱}^{n-۲} \frac{۱}{i} - \frac{۱}{i+۱} \leqslant ۲$$

همین منطق را برای دورِ سوم حلقه‌ی بیرونی به کار می‌بریم مجددا:

$$E[C_۳] = \Sigma_{i=۱}^{n-۳} \frac{(۲n-۳+i)}{i(i+۱)n(n-۱)} \leqslant \frac{۱}{n(n-۱)}\Sigma_{i=۱}^{n-۳} \frac{۳n}{i(i+۱)} = \frac{۳}{n-۱}\Sigma_{i=۱}^{n-۳} \frac{۱}{i} - \frac{۱}{i+۱} \leqslant \frac{۳}{n-۱}$$

به همین طریق برای دور چهارم خواهیم داشت:

$$E[C_۴] \leqslant \frac{۴}{(n-۱)(n-۲)}$$

و همین‌طور باقی دورها. حال سعی می‌کنیم مجموع این میانگین‌ها در هر دور را حساب کنیم تا میانگین تعداد مقایسه‌های کل به دست آید.

$$۲ + \frac{۳}{n-۱} + \frac{۴}{(n-۱)(n-۲)} + \frac{۵}{(n-۱)(n-۲)(n-۳)} + ...$$

هر جمله از این دنباله، کوچکتر می‌شود. حدس می‌زنیم از جمله‌ی دوم به بعد تمامی جملات از $\frac{۴}{n}$ کوچکتر باشند. آن را بررسی می‌کنیم و برای اینکار، با توجه به نزولی بودنِ جملات، صرفا کافی‌ست آن را با جمله دوم مقایسه کنیم:

$$\frac{۳}{n-۱} ? \frac{۴}{n} \rightarrow \frac{۳}{۴} ? \frac{n-۱}{n}$$

که برای $n \geqslant ۴$ کسر سمت راست بزرگتر خواهد بود. می‌دانیم تعداد جملات حداکثر برابر با تعداد حلقه‌های بیرونی است. این تعداد کمتر از n است پس حاصل‌جمع از $\frac{۴}{n} \times n$ کمتر است. حاصل‌جمع کران‌های به دست آمده برای دور دوم و اول نیز $۲ + logn$ بود پس با این حساب:

$$E[C] \leqslant log(n) + ۶$$

همچنین در دور اول حلقه‌ی بیرونی مشاهده کردیم که تعداد $log(n)$ مقایسه در حالت متوسط لازم است. پس اردر زمانی این الگوریتم هست:

$$E[C] = \theta(log(n))$$

در بهترین حالت در مقایسه‌ی اول می‌توان به نابجایی رسید و در بدترین حالت باید حداقل $\frac{n(n-۱)}{۲}$ مقایسه انجام دهیم. پیچیدگی بدترین حالت $O(n^۲)$ و بهترین حالت $O(۱)$ است.

الگوریتم دوم: در این الگوریتم در بهترین حالت در همان مقایسه‌ی اول و بدترین حالت باید کل دو حلقه را پیمایش کنیم؛ یعنی $\frac{n(n-۱)}{۲}$ مقایسه باید انجام دهیم. یک نکته‌ی مهم در رابطه با این الگوریتم این است که اگر در دورِ اول حلقه‌ی بیرونی نابجایی پیدا نکنیم، یعنی آرایه هیچ‌گاه نزولی نبوده و با پیمایش عضو به عضو در آن، صعود مشاهده کرده‌ایم و در نتیجه نابجایی وجود نداشته. در این صورت آرایه مرتب است و بدون نابجایی خواهد بود و باقی دورهای حلقه‌ی بیرونی عملا اضافه هستند و بی‌جهت بدترین حالت خود را به میزان زیادی افزایش می‌دهیم با این کار. پیچیدگی بدترین حالت $O(n^۲)$ و بهترین حالت $O(۱)$ است.

حال برای محاسبه‌ی امیدریاضی در دورِ اول حلقه‌ی بیرونی، مجددا مانند ایده‌ی الگوریتم قبل، تعداد جایگشت‌های مطلوب را بر تعداد کل جایگشت‌ها تقسیم کرده و احتمالِ پیدا کردن اولین نابجایی در مقایسه‌ی kام را پیدا کرده و سپس امیدریاضی را محاسبه می‌کنیم. می‌دانیم اگر در مقایسه‌ی kام اولین نابجایی باشد، به این معنی است که k عنصر اول به ترتیب صعودی هستند و عنصر kام از عنصر بعدی بزرگتر است. این معادل با این است که $k+۱$ عنصر اول، به گونه‌ای باشند که k عنصر اول صعودی مرتب باشند و عنصر آخر بزرگترین عنصر نباشد (در این صورت همواره بزرگترین عنصر در جایگاه kام قرار گرفته و از عنصر بعدی قطعا بزرگتر خواهد بود). پس اگر تعداد کل جایگشت‌ها را $(k+۱)!$ در نظر بگیریم، تعداد جایگشت‌ها مطلوب ما برابر با k خواهد بود. حال امیدریاضی را حساب می‌کنیم:

$$E[C_۱] = \Sigma_{i=۱}^{n-۱} \frac{i^۲}{(i+۱)!} \sim e - ۱$$

حال در دور دوم حلقه‌ی بیرونی تعداد مقایسه‌ها از n تا $۲n-۳$ خواهد بود ولی احتمال رسیدن به دور دوم برابر با احتمال دنباله‌ی مرتب است که هست : $n!$. میانگین تعداد مقایسه‌ها از دور دوم به بعد حداکثر $\frac{(\binom{n}{۲})^۲}{n!}$ است چرا که $\binom{n}{۲}$ مقایسه داریم در کل و طول سیگما نیز به همین میزان است. که این عبارت با بزرگ شدن n کوچکتر شده و به صفر میل می‌کند. پس می‌توان کرانِ e و یا $e-۱$ را برای کران مطمئن حالت میانگین این الگوریتم در نظر گرفت. و در نتیجه حالت متوسط از پیچیدگی $O(۱)$ است.

مسئله‌ی ۲. پاسخ سؤال دوم

الف) تعریف $log^*(n)$ به صورت زیر است:

$$log^*(n) = min\{i \geqslant ۰ : log^{(i)} n \leqslant ۱\}$$

پس عبارت $log^*(log\, n)$ به صورت زیر خواهد بود:

$$log^*(n) = min\{i' \geqslant ۰ : log^{(i')}(log\, n) \leqslant ۱\}$$

با توجه به دو عبارت بالا و همان گونه که در تمرین گفته شد نهایتا رابطه زیر بدست می آید:

$$log^*(n) = log^*(log\, n) + ۱ \Longrightarrow log^*(n) = \theta(log^*(log\, n))$$

طرف دوم با توجه به اطلاعات ما بدیهتا برقرار است، اما برای اثبات کافی است $n_۰$ و $c_۱$ و $c_۲$ ای تعریف کنیم که به ازای هر $n \geqslant n_۰$ رابطه زیر برقرار باشد:

$$c_۱ log^*(log\, n) \leqslant log^*(n) \leqslant c_۲ log^*(log\, n)$$

که به ازای $n_۰ = ۱۰$ و $c_۱ = ۱$ و $c_۲ = ۱۰$ رابطه برقرار می باشد؛ پس عبارت صورت سوال درست است.

ب) در روش اول ابتدا آرایه خود را به دو قسمت مساوی تقسیم می کنیم. حال به صورت بازگشتی عناصر نیمه اکثریت را در این دو بازه پیدا می کنیم.اما نکته مهم این است که در هربار نصف کردن آرایه و پیدا کردن عنصر نیمه اکثریت، ممکن است دو عنصر با خاصیت فوق وجود داشته باشند و هردو بیشتر از $\frac{n}{۴}$ در بازه های خود وجود داشته باشند. از طرفی بدیهی است که اگر عنصری بخواهد عنصر نیمه اکثریت کل آرایه باشد باید حداقل در یکی از دو بازه عنصر نیمه اکثریت باشد چون در غیر این صورت در هرکدام از بازه ها تعداد آن از $\frac{n}{۴} \times \frac{۱}{۲}$ کمتر می شود و در نتیجه در مجموع دو بازه تعداد آن در کل آرایه کمتر از $\frac{n}{۴}$ طول آن است. پس نمی تواند عنصر نیمه اکثریت کل باشد! پس نهایتا ما حداکثر ۴ کاندید برای این مهم داریم و پس از اجرای الگوریتم بازگشتیمان یک بار آرایه را پیمایش کرده و تعداد این ۴ عنصر را پیدا می کنیم و اگر تعداد هرکدام از آن ها از $\frac{n}{۴}$ بیشتر بود آنگاه عنصر نیمه اکثریت آرایه A است. رابطه بازگشتی نهایی ما نیز به صورت زیر می باشد:

$$T(n) = ۲T(\frac{n}{۲}) + O(n)$$

نهایتا با توجه به قضیه اصلی داریم: $T(n) = \theta(nlog\, n)$

برای روش دوم نیز کافی است مانند روش اول عمل کنیم با این تفاوت که آرایه A را به سه قسمت مساوی به طول $\frac{n}{۳}$ تقسیم کنیم و به همان روش تقسیم و غلبه عنصر مورد نظر را در صورت وجود پیدا کنیم.