ساختمان‌داده‌ها و طراحی الگوریتم‌ها

نیم‌سال دوم ۰۲-۰۳

مدرس: آبام- کاظمی

دانشکده‌ی مهندسی کامپیوتر

تمرین سری پنجم | درهم‌سازی | زمان آزمون: ۲۷ خرداد

۱ مسائل اصلی تمرین

مسئله‌ی ۱. درهم‌سازی زنجیره‌ای

فرض کنید برای درهم‌سازی n کلید متمایز از روش درهم‌سازی زنجیره‌ای استفاده شده است. تابع درهم‌سازی ساده و یکنوا و اندازه آرایه m می‌باشد. امید ریاضی تعداد برخوردها (تعداد جفت کلیدهایی که به یک خانه نگاشت می‌شوند) برحسب m و n چه پیچدگی محاسباتی دارد؟

مسئله‌ی ۲. ترتیب اعداد

فرض کنید از روش آدرس‌دهی باز با استفاده از وارسی خطی برای درهم‌سازی استفاده شده است. اندازه جدول درهم‌سازی ۱۰ و تابع درهم ساز $h(x) = x^۲ \bmod ۱۰$ است. با فرض آنکه ورودی‌ها به ترتیب ۱۹, ۱, ۳۰, ۴۹, ۵۲, ۳۷, ۱۲۳, ۵۳, ۱۷, ۵ (از چپ به راست) باشد، اعداد به چه ترتیب در جدول ذخیره می‌شوند.

مسئله‌ی ۳. رفع تصادم

وضعیت فعلی یک جدول درهم‌ساز در زیر آمده است. فرض کنید برای رفع مشکل تصادم از روش وارسی خطی استفاده شده است. با در نظر گرفتن فرض یکنواختی تابع درهم‌ساز، کلید بعدی با چه احتمالی در خانه‌ی دوم قرار می‌گیرد؟ (خانه‌های جدول از چپ به راست از ۱ تا ۱۸ شماره‌گذاری شده‌اند.)

$$H[۱..۱۸] = \{۶, -, ۱, -, ۳, -, ۱۴, -, ۹, ۲, -, ۱۱, -, -, -, ۰, ۴, ۵\}$$

مسئله‌ی ۴. وضعیت جدول

فرض کنید از آدرس‌دهی باز و وارسی خطی برای درهم‌سازی استفاده شده و تابع درهم سازی $i^۲$ به پیمانه ۷ است. بعد از دریافت همه اعداد ۰, . . . , ۶ می‌دانیم نحوه‌ی قرار گیری اعداد در جدول درهم‌ساز به صورت زیر است:

$$A[۰, .., ۶] = ۰, ۶, ۴, ۳, ۱, ۵, ۲$$

به ازای چند جایگشت ورودی وضعیت جدول درهم‌ساز به شکل بالا خواهد بود.

مسئله‌ی ۵. احتمال نگاشت

جدول درهم‌سازی ۱۰ خانه‌ای و تابع درهم‌ساز $h(x) = ۳x + ۵ \mod ۱۰$ را در نظر بگیرید. کدام گزینه درست است. توضیح دهید.

۱. احتمال آن که ورودی $x = ۴$ به خانه ۷ نگاشت شود برابر ۱/۱۰ است.

۲. احتمال آن که ورودی $x = ۴$ به خانه ۷ نگاشت شود برابر ۰ است.

۳. احتمال آن که ورودی $x = ۴$ به خانه ۷ نگاشت شود برابر ۱ است.

۴. احتمال آن که دو ورودی مختلف به یک خانه نگاشت شوند برابر ۱/۱۰ است.

مسئله‌ی ۶. تعداد صفرها

فرض کنید $H : \{۱, \cdots, n\} \to \{۱, \cdots, n\}$ یک تابع درهم‌ساز یکنواخت باشد. برای ورودی x، عدد z را برابر تعداد صفرهای سمت راست $H(x)$ قرار می‌دهیم. برای عدد $۰ \leqslant c \leqslant ۱$، احتمال $z \geqslant c \log n$ از چه مرتبه‌ای است؟ فرض کنید c ثابت است.

مسئله‌ی ۷. روش جستجو

فرض کنید می‌خواهیم درون یک جدول درهم‌سازی با خانه‌های $\{۰, ۱, ..., m - ۱\}$ دنبال عنصر داده‌شده k بگردیم. همچنین تابع درهم‌سازی $h : \mathcal{I} \to \{۰, ۱, ..., m - ۱\}$ که $\mathcal{I}$ نمایش‌دهنده‌ی فضای عناصر می‌باشد را در اختیار داریم. روش جست‌وجوی ما در زیر آمده است:
۱) مقدار $i \leftarrow h(k)$ را محاسبه کن و $j \leftarrow ۰$ قرار بده.
۲) در درایه‌ی i به دنبال k بگرد. اگر آن را یافتی یا اگر درایه خالی بود جست‌وجو را متوقف کن.
۳) $j \leftarrow (j + ۱)\ mod\ m$ و $i \leftarrow (i + j)\ mod\ m$ قرار بده و به مرحله‌ی ۲ برگرد.
فرض کنید m توانی از ۲ است.

۱. نشان دهید که این روش یک روش از روش کلی وارسی درجه ۲ است. این کار را با تعیین مقدارهای مناسب برای ثابت‌های $c_۱$ و $c_۲$ انجام دهید.

۲. ثابت کنید این الگوریتم در بدترین حالت هر درایه از جدول را وارسی می‌کند.

۲ مسائل اضافه‌تر برای علاقه‌مندان

مسئله‌ی ۸. توپ و سطل

نشان دهید زمانی که n توپ با احتمال یکسان و به صورت مستقل درون n سطل پرتاب شوند، احتمال اینکه بیشترین تعداد توپ در یک سطل از $۳ \ln n / \ln \ln n$ تجاوز کند، حداکثر $۱/n$ خواهد بود، برای n به حد کافی بزرگ.

مسئله‌ی ۹. سطل‌های بیشتر

نشان دهید که به ازای هر $۰ < \varepsilon$، اگر n توپ با احتمال یکسان و مستقلا در $n^{۲+\varepsilon}$ سطل پرتاب شوند، با احتمال $۱ - ۱/n^{\varepsilon}$ هیچ سطلی بیش از دو توپ ندارد.

مسئله‌ی ۱۰. Universal Hashing

نشان دهید که اگر یک خانواده $\mathcal{H}$ از توابع درهم‌سازی universal – ۲ باشد، آنگاه universal است.

مسئله‌ی ۱۱. tupleهای n تایی

فرض کنید U یک universe از tupleهای مانند $x = \langle x_۰, x_۱, \cdots, x_{n-۱} \rangle$ و برگرفته شده از

$$\mathbb{Z}_p = \{۰, ۱, \cdots, p-۱\}$$

باشد، که در آن p عددی اول است. برای هر tuple مانند $a = \langle a_۰, a_۱, \cdots, a_{n-۱} \rangle \in U$، تابع درهم‌سازی h_a به این صورت تعریف می‌شود:

$$h_a(x) = \Big(\sum_{j=۰}^{n-۱} a_j x_j \Big) \mod p$$

می‌دانیم که خانواده توابع درهم‌سازی $\mathcal{H} = \{h_a\}$ یک خانواده از توابع universal است. نشان دهید $\mathcal{H}$ یک خانواده ۲-universal نیست.

مسئله‌ی ۱۲. همچنان tupleهای n تایی

فرض کنید در توابع درهم‌سازی سوال قبل، تغییر زیر را اعمال کردیم:

$$h_a(x) = \Big(\sum_{j=۰}^{n-۱} a_j x_j + b \Big) \mod p$$

که در آن، $b \in \mathbb{Z}_p$ است. نشان دهید که خانواده جدید توابع درهم‌سازی ۲-universal است.