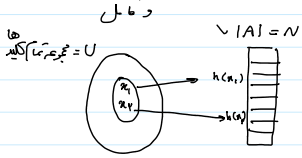
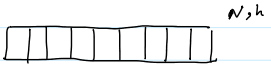
جلسه پانزدهم

13:18 Sunday, November 22, 2020

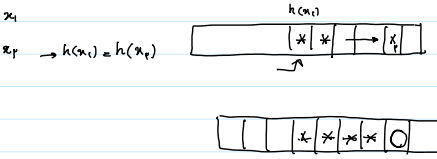
موضوع: ادامه درهم سازی، روش وارسی مربعی و هش دوگانه، هش جهانی و کامل





تصادم: x_1 و x_2 $\quad x_2 \neq x_1 \quad h(x_1) = h(x_2)$

آدرس دهی باز: * وارسی خطی:



* روش وارسی مربعی یا Quadratic probing:

$h(x)$ ، $h(x)+1^2$ ، $h(x)+2^2$ ، $h(x)+3^2$ ، ...

مثال: $N=31$ $\quad h(x)=23$ $\quad \ldots - 1 - 2^2 - 3^2 - \ldots$

$23 - 24 - 27 - \ldots - 30 - 32$

مشکل روش وارسی مربعی: secondary clustering : در این حالت ممکن است تعداد خانه‌های پُر شده بسیار کم باشد.

$h(x)=1$ $\quad 1\ 2\ 4\ 9\ 16\ 25\ \ldots$

$N=8$

[قضیه]: اگر N اول باشد، حداقل $\lceil \frac{N}{2} \rceil$ از خانه‌ها بررسی می‌شوند. لذا اگر $\frac{n}{N} \leq \frac{1}{2}$ باشد حتما خانه خالی پیدا می‌شود.

اثبات: $N \times (h(x)+j^2) = N \times (h(x)+i^2)$: فرض

$N \times (j^2 - i^2) = 0 \quad \rightarrow \quad N \times (j-i)(j+i) = 0$

$N \mid (j+i)$ $\quad$ $0 < j+i < 2N$

$j-i$ یا $j+i$ از N بیشتر است

□

قضیه [بدون اثبات] متوسط تعداد عمل در جستجوی ناموفق $\frac{1}{1-\alpha}$

" " " موفق: $\frac{1}{\alpha} \log \left[\frac{1}{1-\alpha}\right]$

$n = \alpha N \rightarrow \alpha \leq \frac{1}{2}$

* هش دوگانه Double Hashing: $h_1(x)$ و $h_2(x)$

$h_1(x)$ ، $h_1(x)+h_2(x)$ ، $h_1(x)+2h_2(x)$ ، ...

مثلا $h_2(x) = x \% 5$

طول گام‌ها $h_2(x)$ چه خاصیت‌هایی دارد؟

* هیچ‌گاه صفر نشود.

* بخش قابل توجهی از جدول وارسی شود

$|U| \gg N$ و n ... هش درهم سازی خوبی هم در نظر بگیریم با ورود کاربران

$|U| \gg N$ و n ← هر تابع درهم سازی خوبی هم درنظر بگیریم، بازورودی‌ای ای وجود دارد که به ازای آن اتفاقات بدی رخ می‌دهد

Adversary :

راه حل : تابع هش را به صورت تصادفی انتخاب کنیم .

تعریف : فرض کنید H یک مجموعه (خانواده) از توابع درهم سازی باشد. در این صورت H یک خانواده جهانی universal است اگر

* $\forall x,y \in U \;,\; \Pr_{h \in H}[h(x)=h(y)] \leq \frac{1}{N}$ ← اندازه برابر

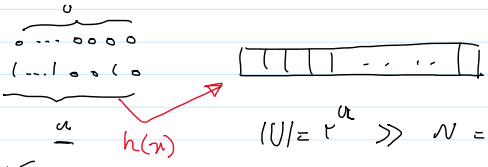
* در واقع تعداد توابعی از H که $h(x)=h(y)$ است کمتر از $\frac{|H|}{N}$ باشد، $\frac{|H|}{N} \times \frac{1}{|H|} \geq \frac{1}{N}$

قضیه : فرض کنید H یک خانواده جهانی است و h یک تابع تصادفی از H است. در این صورت به ازای n عمل درج، متوسط تعداد تصادم‌ها برابر با $\frac{n}{N}$ برای یک کلید است.

$x \neq y$ ، احتمال تصادم $\frac{1}{N}$ ← یک عنصر به طور متوسط با $\frac{n}{N}$ تصادم می‌کند .

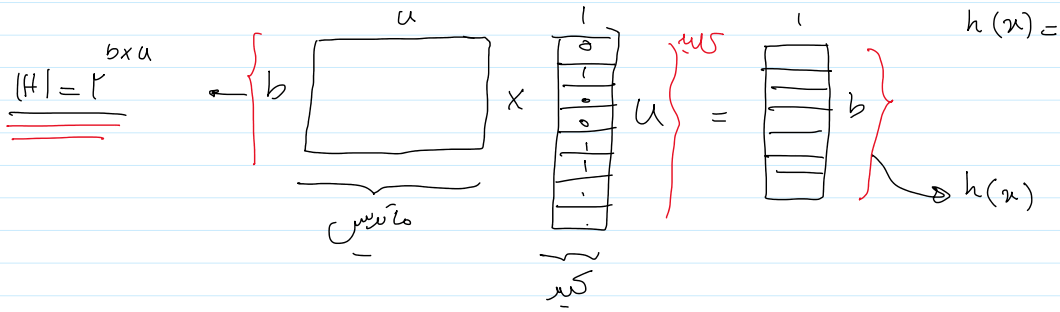
سوال : آیا چنین خانواده ای وجود دارد؟

* مثال : فرض کنید U شامل تمام اعداد u بیتی است. فرض کنید N هم برابر با 2^b است ← مبنای 2



$b \ll u$

مجموعه H (خانواده توابع درهم‌سازی) : تمام ماتریس‌ها با اندازه $b \times u$ و با مقدارهای 0 و 1 ، به ازای یک $h \in H$



ضرب : ضرب

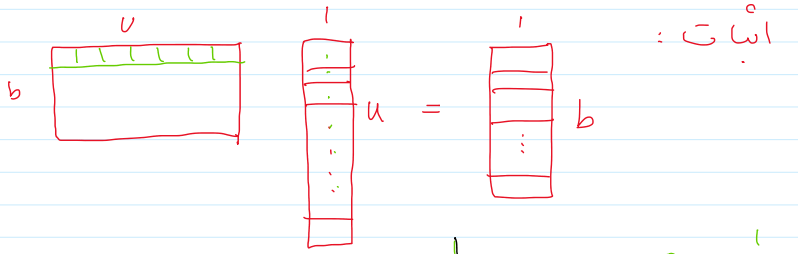
جمع : XOR

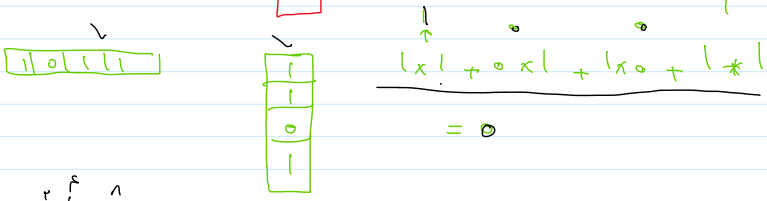
قضیه : فرض کنید h به صورت تصادفی انتخاب شده است ($h \in H$) در این صورت

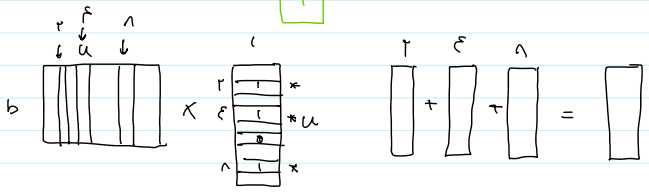
$$x \neq y \quad \Pr[h(x)=h(y)] = \frac{1}{2^b} = \frac{1}{N}$$

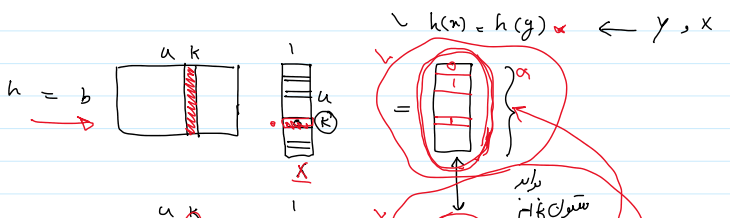
اثبات :

اثبات :









ستون k را کنار بگذاریم

$\Rightarrow$ ستون k ام یکبار جابجا می‌شود. $\longleftarrow \frac{1}{2b} = \frac{1}{n}$