

1 Unique Elements

In the last class, we saw that the expected number of unique elements a set S can have with respect to a randomly chosen hashing function belonging to H is $N/2$, where H is a 2-Universal hash family U to T , $|S| = N$, and $|T| = 2N$.

1.1 Notations

U - Universal Set

S - an arbitrary subset of U

N - size of the set S i.e. $|S|$

T - hash table to store the elements of S

H - 2-Universal family of hash functions from U to T

X - A random variable defined as the number of unique elements in S with respect to $h \in H$, when h is chosen uniformly at random from H .

$Y = N - X$

In the last class, we saw that when $T \geq 2N$,

$$E[X] \geq N/2 \tag{1.1}$$

Claim:

$$\Pr[X \geq N/3] \geq 1/4 \tag{1.2}$$

when $|T| \geq 2N$.

Proof:

$$\begin{aligned} E[Y] &\leq N - \frac{N}{2} \\ &\leq \frac{N}{2} \end{aligned}$$

$$\Pr[X \geq \frac{N}{3}] = 1 - \Pr[X < \frac{N}{3}]$$

$$\begin{aligned} \Pr[X < \frac{N}{3}] &= \Pr[Y \geq \frac{2N}{3}] \\ &< \frac{E[Y]}{\frac{2N}{3}} \quad (\text{by Markov's inequality}) \\ &< \frac{\frac{N}{2}}{\frac{2N}{3}} \\ &< \frac{3}{4} \end{aligned}$$

Therefore

$$\Pr[X \geq N/3] \geq \frac{1}{4}$$

2 Unique-SAT

Definition: Here, we are given a formula $\phi(x_1, x_2, x_3, \dots, x_n)$ and a promise that ϕ has exactly one satisfying assignment or no satisfying assignment. Can we find the satisfying assignment if it exists or say that it is not satisfiable if none exists?

2.1 Valiant-Vazirani Lemma

Theorem:

There is a randomized algorithm A such that for every boolean formula ϕ ,

$$\begin{aligned} \phi \in \text{SAT} &\Rightarrow \Pr[A(\phi) \text{ has exactly one satisfying assignment}] \geq 1/48n \\ \phi \notin \text{SAT} &\Rightarrow A(\phi) \notin \text{SAT} \end{aligned}$$

Proof:

Let

$L = \{ \langle \phi, 1^n, 1^k, h, \alpha \rangle \mid \phi \text{ has } n \text{ variables, } 1 \leq k \leq n, h \in H \text{ is a hashing function from } \Sigma^n \rightarrow \Sigma^{k+2}, \alpha \in \Sigma^{k+2}, \exists a \text{ st } \phi(a) = 1 \text{ and } h(a) = \alpha \}$

Here h is specified by a tuple $\langle a, b \rangle$ (Recall that $h_{ab}(x)$ is the first $k+2$ bits of $ax + b$). Clearly, $L \in \text{NP}$. Hence, there exists a function f that reduces L to SAT.

Consider the following randomized algorithm A .

- 1 Input $\phi(x_1, x_2, \dots, x_n)$
- 2 Randomly pick $k \in \{1, 2, \dots, n\}$
- 3 Randomly pick $h \in H$ [family of 2-Universal hash functions from $\Sigma^n \rightarrow \Sigma^{k+2}$]
- 4 Randomly pick $\alpha \in \Sigma^{k+2}$
- 5 Output $f(\langle \phi, 1^n, 1^k, h, \alpha \rangle)$

Now

$$\begin{aligned} \phi \notin \text{SAT} &\implies \langle \phi, 1^n, 1^k, h, \alpha \rangle \notin L \\ &\implies f(\langle \phi, 1^n, 1^k, h, \alpha \rangle) \notin \text{SAT} \end{aligned}$$

Suppose $\phi \in \text{SAT}$. Then there exists a constant k such that the number of assignments that satisfy ϕ lie between 2^k and 2^{k+1} . In Step 2, the correct value of k is chosen with probability $1/n$. From now we assume that this event has happened.

If we randomly pick h , by earlier Claim at least $\frac{|S|}{3}$ elements from T have exactly one inverse in S with respect to h with probability $\geq \frac{1}{4}$. From now we assume that this even happened.

So the probability of a randomly chosen element in S being unique (wrt h) is at least $\frac{1}{12}$. Since, $2^k \leq |S| \leq 2^{k+1}$ and $|T| = 2^{k+2}$, we have $2|S| \leq |T| \leq 4|S|$.

Thus, given that k and h are good random choices, the probability of a randomly chosen $\alpha \in T$ has exactly one inverse in $S(\text{wrt } h) \geq \frac{1}{4} \cdot \frac{1}{12}$

Hence, the probability that $\langle \phi, 1^n, 1^k, h, \alpha \rangle$ has exactly one witness at least $\frac{1}{48n}$

Now, we have the following corollary.

Corollary. *If UniqueSAT is in P, then SAT is in BPP.*

Let B be an algorithm that solves UniqueSAT in polynomial time. Given a formula ϕ , run $A(\phi)$ to obtain a new formula ψ . If B accepts ψ , then accept, else reject.

If $\phi \notin \text{SAT}$, then $\psi \notin \text{SAT}$ and B rejects ψ and so the above algorithm rejects ϕ . If $\phi \in \text{SAT}$, then with probability at least $1/48n$, ψ has exactly one satisfying assignment. Thus if $\phi \in \text{SAT}$, then the above algorithm accepts with probability at least $1/48n$. Since this algorithm has one sided error, the success probability can be amplified to $1 - 1/2^{p(n)}$ for some polynomial p .

3 Random Walks

Consider a line segment divided into n partitions with 0 marked as the beginning of the segment and n marked as the end of the segment with positions 1 to $n - 1$ being in between the two positions in a sequential order. Consider a particle at position 0 which moves right by 1 position and keeps moving either left or right with probability $\frac{1}{2}$ till it reaches position n .

i.e. $\forall j, 0 < j < n$

$\Pr[\text{Particle moves to position } j + 1 \text{ from } j] = \frac{1}{2}$

$\Pr[\text{Particle moves to position } j - 1 \text{ from } j] = \frac{1}{2}$

We are interested in finding the expected number of moves to reach n starting at position 0.

There are several possible sequences of moves that will take the particle from 0 to n . For example, keep moving right till it reaches n , or $n + 2$ steps to right and 1 step to left, $n + 4$ steps to right and 2 steps left etc. This would give an expression that is somewhat similar to

$$\sum_{k \geq 0} C_k^{n+2k} \cdot \left(\frac{1}{2}\right)^{n+2k} [n + 2k]$$

We can compute the expectation easily using (again) linearity of expectation. Let X_j be a random variable defined as the number of steps to reach n starting from j .

$$E[X_j] = \frac{1}{2}[1 + E[X_{j+1}]] + \frac{1}{2}[1 + E[X_{j-1}]]$$

Let h_j denote $E[X_j]$. Then

$$h_j = \frac{h_{j-1}}{2} + \frac{h_{j+1}}{2} + 1$$

Observe that $h_n = 0$ and $h_0 = 1 + h_1$.

Claim: $h_j = h_{j+1} + 2j + 1$

Proof:

Let $f_j = h_j - h_{j-1}$. Since, $h_0 = 1 + h_1$, we have $f_1 = h_1 - h_0 = -1$. Since $h_j = \frac{h_{j-1}}{2} + \frac{h_{j+1}}{2} + 1$, $2h_j = h_{j-1} + h_{j+1} + 2$. Thus

$$h_{j+1} - h_j = h_j - h_{j-1} - 2$$

$$\begin{aligned} f_{j+1} &= f_j - 2 \\ &= -(2j + 1) \end{aligned}$$

Therefore

$$h_{j+1} - h_j = -(2j + 1)$$

$$h_{j+1} - h_j = -(2j + 1)$$

$$h_j = h_{j+1} + 2j + 1$$

Hence the proof.

With this result, we have

$$\begin{aligned} h_0 &= h_1 + 2 \cdot 0 + 1 \\ &= h_1 + 1 \\ &= h_n + \sum_{j=1 \dots n} (2j - 1) \\ &= n^2 \end{aligned}$$

Thus the expected number of steps to reach n starting at 0 is n^2 .