

Relacionando el muestreo (casi) uniforme con la existencia de un FPRAS

IIC3810

Marcelo Arenas y Luis Alberto Croquevielle

La noción de p -relación

Será conveniente ver las funciones de $\#P$ como relaciones.

Definición

Una relación $R \subseteq \Sigma^ \times \Sigma^*$ es una p -relación si:*

- ▶ *Existe un polinomio q tal que si $(x, y) \in R$, entonces $|y| \leq q(|x|)$*
- ▶ *$R \in PTIME$, vale decir, existe un algoritmo de tiempo polinomial que, dado $(x, y) \in \Sigma^* \times \Sigma^*$, verifica si $(x, y) \in R$*

Cada p -relación representa una función en $\#P$

Dada una relación $R \subseteq \Sigma^* \times \Sigma^*$, defina la función $f_R : \Sigma^* \rightarrow \mathbb{N}$ como:

$$f_R(x) = \begin{cases} |\{y \mid (x, y) \in R\}| & \text{si } \{y \mid (x, y) \in R\} \text{ es finito} \\ 0 & \text{en otro caso} \end{cases}$$

Proposición

Si R es una p -relación, entonces $f_R \in \#P$

Ejercicio

Demuestre la proposición.

Cada función en $\#P$ puede ser representada como una p -relación

Sea $f : \Sigma^* \rightarrow \mathbb{N}$ una función en $\#P$

- ▶ Existe una MT no determinista M tal que para todo $x \in \Sigma^*$ se tiene que $f(x) = \text{accept}_M(x)$

Cada ejecución de M se puede codificar usando el alfabeto Σ

Utilizando las codificaciones de las ejecuciones de M definimos:

$$R_f = \{(x, y) \in \Sigma^* \times \Sigma^* \mid y \text{ codifica una ejecución de } M \\ \text{con entrada } x \text{ que termina en un estado final}\}$$

Cada función en $\#P$ puede ser representada como una p -relación

Proposición

Si f está en $\#P$, entonces R_f es una p -relación

Demostración: Como la MT no determinista M en la transparencia anterior es de tiempo polinomial, para una entrada x se puede:

- ▶ Codificar cualquier ejecución de M que acepta usando un string de largo polinomial en $|x|$
- ▶ Verificar si una ejecución termina en estado final (simulando el funcionamiento de M) en tiempo polinomial



Funciones en $\#P$ y p -relaciones

Por lo tanto, de ahora en adelante trabajamos con p -relaciones.

Estudiaremos los problemas de conteo y de generación uniforme asociados a p -relaciones, sabiendo que los resultados se extienden de manera inmediata a funciones en $\#P$

Un generador uniforme para una p -relación

Dada una p -relación $R \subseteq \Sigma^* \times \Sigma^*$, sea:

$$N_R(x) = |\{y \in \Sigma^* \mid (x, y) \in R\}|$$

Además, suponga que $\perp$ es un símbolo reservado que no es usado en Σ

Un algoritmo aleatorizado $\mathcal{G} : \Sigma^* \rightarrow \Sigma^* \cup \{\perp\}$ es un generador uniforme para R si para todo $x, y \in \Sigma^*$:

- ▶ si $(x, y) \notin R$, entonces $\Pr(\mathcal{G}(x) = y) = 0$
- ▶ si $(x, y) \in R$, entonces $\Pr(\mathcal{G}(x) = y) = \frac{1}{N_R(x)}$

Un generador casi uniforme para una p -relación

Las herramientas que veremos más adelante no nos permitirán obtener generadores uniformes, sino que una versión más débil

- Consideramos nuevamente una p -relación $R \subseteq \Sigma^* \times \Sigma^*$

Definición

Un algoritmo aleatorizado $\mathcal{G} : \Sigma^* \times (0, 1) \rightarrow \Sigma^* \cup \{\perp\}$ es un generador casi uniforme para R si para todo $x, y \in \Sigma^*$ y $\varepsilon \in (0, 1)$:

- si $(x, y) \notin R$, entonces $\Pr(\mathcal{G}(x, \varepsilon) = y) = 0$
- si $N_R(x) > 0$, entonces $\Pr(\mathcal{G}(x, \varepsilon) = \perp) = 0$
- si $(x, y) \in R$, entonces:

$$(1 - \varepsilon) \cdot \frac{1}{N_R(x)} \leq \Pr(\mathcal{G}(x, \varepsilon) = y) \leq (1 + \varepsilon) \cdot \frac{1}{N_R(x)}$$

Un esquema de generación casi uniforme

Definición

Dada una p -relación $R \subseteq \Sigma^* \times \Sigma^*$, un algoritmo aleatorizado $\mathcal{G} : \Sigma^* \times (0, 1) \rightarrow \Sigma^* \cup \{\perp\}$ es un **fully polynomial almost uniform generator (FPAUG)** para R si

1. $\mathcal{G}$ es un generador casi uniforme para R
2. Existe un polinomio $q(u, v)$ tal que para todo $x \in \Sigma^*$ y $\varepsilon \in (0, 1)$, el número de pasos ejecutados por $\mathcal{G}(x, \varepsilon)$ es menor o igual a $q(|x|, \frac{1}{\varepsilon})$

Una definición de FPRAS para relaciones

Definición

Dada una p -relación $R \subseteq \Sigma^* \times \Sigma^*$, un algoritmo aleatorizado $\mathcal{A} : \Sigma^* \times (0, 1) \rightarrow \mathbb{N}$ es un **fully polynomial randomized approximation scheme (FPRAS)** para R si existe un polinomio $q(u, v)$ tal que para cada $x \in \Sigma^*$ y $\varepsilon \in (0, 1)$:

1. El número de pasos ejecutados por $\mathcal{A}(x, \varepsilon)$ es menor o igual a $q(|x|, \frac{1}{\varepsilon})$
2. $\Pr(|\mathcal{A}(x, \varepsilon) - N_R(x)| \leq \varepsilon \cdot N_R(x)) \geq \frac{3}{4}$

Observación

Dado $f \in \#P$ representado como R_f , se puede demostrar que esta definición de FPRAS es equivalente a la vista en el capítulo anterior.

Un comentario sobre las definiciones anteriores

La noción de algoritmo aleatorizado se formaliza usando MT probabilísticas.

- ▶ Estas máquinas funcionan con cintas de bits, por lo que las probabilidades resultantes son de la forma $\frac{n}{2^k}$

Por lo tanto, al describir un algoritmo aleatorizado, en teoría no podemos decir algo como “la probabilidad de error del algoritmo es $\frac{1}{3}$ ”

Algunos comentarios sobre las definiciones anteriores

Tratar de tener algoritmos aleatorizados con probabilidades arbitrarias no entrega intuiciones nuevas

- Y hace mucho más técnicas y complicadas las demostraciones

Supuesto

Todas las probabilidades que vamos a considerar (por ejemplo, la probabilidad $\frac{1}{N_R(x)}$) son de la forma $\frac{n}{2^k}$

La relación entre FPAUG y FPRAS

Pasaremos ahora a enunciar y demostrar que la existencia de un FPAUG implica la existencia de un FPRAS.

- ▶ Este resultado es válido para una amplia clase de relaciones
- ▶ Esto nos va a permitir utilizar una gran cantidad de herramientas desarrolladas para el muestreo de variables aleatorias en la construcción de FPRAS

Primero debemos formalizar la noción de p -relación auto-reducible, la cual es necesaria al demostrar la relación entre FPAUG y FPRAS.

Relaciones auto-reducibles

Intuitivamente, un problema se dice auto-reducible si es que se puede solucionar mediante la resolución de instancias más simples del mismo problema.

Ejemplo

Sea φ una fórmula proposicional con variables $x_1, \dots, x_n$

La notación $\varphi[\frac{x_i}{v}]$ indica que la variable x_i es reemplazada por $v \in \{0, 1\}$

- ▶ Si $v = 0$ reemplazamos x_i por el operador 0-ario $\perp$, y si $v = 1$ reemplazamos x_i por el operador 0-ario $\top$
- ▶ $\varphi[\frac{x_i}{v}]$ tiene una variable menos que φ

Determinar si φ es satisfacible se reduce a determinar si $\varphi[\frac{x_1}{0}]$ o $\varphi[\frac{x_1}{1}]$ es satisfacible

- ▶ Así, una instancia de SAT se reduce a instancias más simples de SAT

Definición

Una relación $R \subseteq \Sigma^* \times \Sigma^*$ es auto-reducible si:

1. Existe una función $g : \Sigma^* \rightarrow \mathbb{N}$ tal que g es computable en tiempo polinomial y para cada $(x, y) \in R$ se tiene que $|y| = g(x)$
2. Existen funciones $\psi : \Sigma^* \times \Sigma^* \rightarrow \Sigma^*$ y $\sigma : \Sigma^* \rightarrow \mathbb{N}$ tales que:
 - ▶ ψ y σ son computables en tiempo polinomial
 - ▶ $\sigma(x) \in O(\log(|x|))$
 - ▶ $\forall x \in \Sigma^* : \text{si } g(x) > 0, \text{ entonces } 0 < \sigma(x) \leq g(x)$
 - ▶ $\forall x, w \in \Sigma^* : |\psi(x, w)| \leq |x|$
 - ▶ $\forall x, y \in \Sigma^* \text{ con } y = a_1 \cdots a_n :$
 $(x, y) \in R \text{ si y sólo si } (\psi(x, a_1 \cdots a_{\sigma(x)}), a_{\sigma(x)+1} \cdots a_n) \in R$

Ejercicios

Demuestre que las siguientes relaciones son auto-reducibles:

1. $R_{\text{SAT}} = \{(\varphi, \sigma) \mid \varphi \text{ es una fórmula proposicional y } \sigma \text{ es una valuación tal que } \sigma(\varphi) = 1\}$
2. $R_{\text{IS}} = \{(G, S) \mid G \text{ es un grafo y } S \text{ es un conjunto independiente de } G\}$

Una relación natural no auto-reducible

Dado un grafo $G = (N, A)$, decimos que $S \subseteq N$ es un conjunto independiente **maximal** de G si:

1. S es un conjunto independiente de G
2. Para todo conjunto independiente S' de G , no se cumple que $S \subsetneq S'$

Considere la relación:

$$R_{\text{MIS}} = \{(G, S) \mid G \text{ es un grafo y } S \text{ es un conjunto independiente maximal de } G\}$$

¿Es R_{MIS} auto-reducible?

- ¿Cómo se puede demostrar que no lo es?

Una propiedad de las relaciones auto-reducibles

Para demostrar que R_{MIS} no es auto-reducible identificamos una propiedad de las relaciones auto-reducibles que no es cumplida por R_{MIS}

- Bajo una suposición de complejidad

Dado un alfabeto Σ , suponga dado un orden lineal en Σ

- Este orden lineal induce un orden lexicográfico $\leq$ en Σ^*

Definición

Dada una relación $R \subseteq \Sigma^* \times \Sigma^*$:

$$\text{Exists}(R) = \{x \mid \exists y : (x, y) \in R\}$$

$$\text{Min}(R) = \{(x, y) \mid x \in \text{Exists}(R) \wedge y = \arg \min_{\leq} \{z \mid (x, z) \in R\}\}$$

$$\text{Max}(R) = \{(x, y) \mid x \in \text{Exists}(R) \wedge y = \arg \max_{\leq} \{z \mid (x, z) \in R\}\}$$

Una propiedad de las relaciones auto-reducibles

Teorema

Si R es una p -relación auto-reducible tal que $\text{Exists}(R) \in PTIME$, entonces $\text{Min}(R) \in PTIME$ y $\text{Max}(R) \in PTIME$

Ejercicio

Demuestre el teorema

R_{MIS} no es auto-reducible

Proposición

Si R_{MIS} es auto-reducible, entonces $\text{PTIME} = \text{NP}$

Ejercicio

Demuestre las siguientes propiedades:

1. R_{MIS} es una p -relación y $\text{Exists}(R_{\text{MIS}}) \in \text{PTIME}$
2. $\text{Min}(R_{\text{MIS}})$ es co-NP-completo

A partir de estas propiedades y del teorema demuestre la proposición.

Solucionando el ejercicio: $\overline{\text{Min}(R_{\text{MIS}})}$ es NP-hard

Vamos a mostrar que $\text{CNF-SAT} \leq_m^p \overline{\text{Min}(R_{\text{MIS}})}$

- Dada una formula proposicional φ en CNF, mostramos como construir en tiempo polinomial un grafo $G = (N, A)$ y un conjunto $S \subseteq N$ tales que:

$$\varphi \text{ es satisfacible} \quad \text{si y sólo si} \quad (G, S) \in \overline{\text{Min}(R_{\text{MIS}})}$$

Debemos representar S como un string sobre un alfabeto Σ fijo.

- Esto es necesario porque debemos tener un orden lexicográfico sobre los conjuntos independientes maximales de G

En estricto rigor también G debería ser representado como un string sobre Σ

- Aunque esto no es fundamental para la demostración

Solucionando el ejercicio: $\overline{\text{Min}(R_{\text{MIS}})}$ es NP-hard

Vamos a dar la idea de la demostración con un ejemplo.

- ▶ Dejamos como un ejercicio el generalizar esta idea a cualquier fórmula proposicional en CNF

Suponga que $\varphi = C_1 \wedge C_2$, donde $C_1 = (r \vee t)$ y $C_2 = (t \vee \neg s \vee \neg t \vee \neg u)$

- ▶ Consideramos $\Sigma = \{0, 1\}$ en la reducción

Los nodos del grafo G

El conjunto N de nodos de G es definido como:

$$N = \{C_1, C_2, r, s, t, u, \neg r, \neg s, \neg t, \neg u, \star\}$$

donde $\star$ es un símbolo que no es mencionado en φ

Representando un conjunto independiente de G

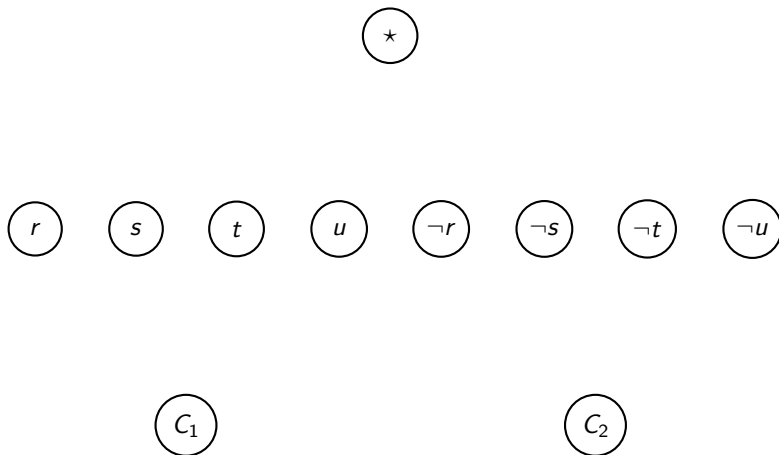
Para representar un conjunto $S \subseteq N$ usamos un string $w \in \{0,1\}^*$ de largo 11

- ▶ El primer bit de w es 1 si $C_1 \in S$, y 0 en caso contrario. El segundo bit de w es 1 si $C_2 \in S$, y 0 en caso contrario
- ▶ El tercer bit de w es 1 si $\star \in S$, y 0 en caso contrario
- ▶ Los siguientes bits de w son contruidos de la misma forma para los nodos $r, s, t, u, \neg r, \neg s, \neg t, \neg u$

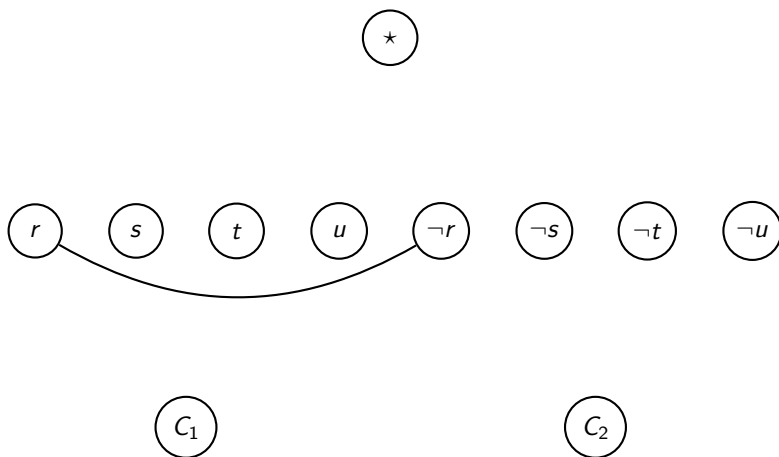
Ejemplo

El conjunto $S = \{C_2, \star, u, \neg s, \neg t\}$ es representado por el string 01100010110

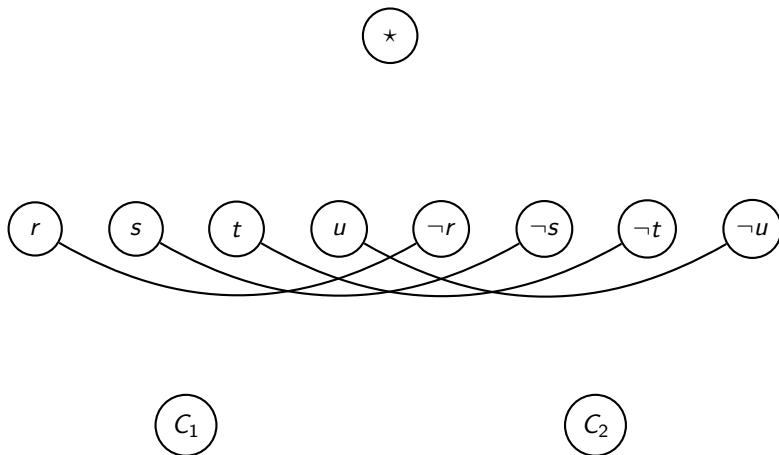
Los arcos del grafo G



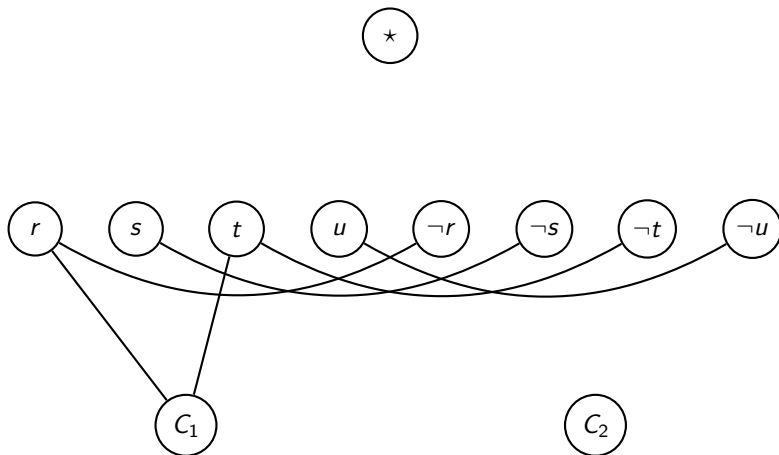
Los arcos del grafo G



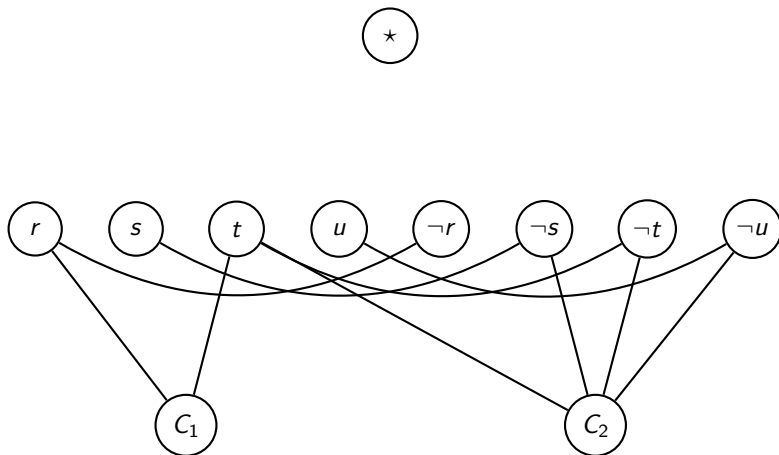
Los arcos del grafo G



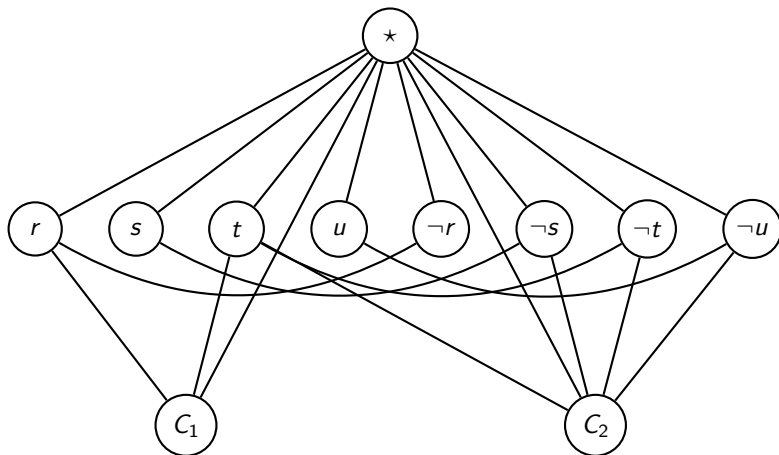
Los arcos del grafo G



Los arcos del grafo G



Los arcos del grafo G



La equivalencia entre los problemas

$S = \{\star\}$ es un conjunto independiente maximal de G

- ▶ Es representado por el string 00100000000

Tenemos que φ es satisfacible si y sólo si $(G, S) \in \overline{\text{Min}(R_{\text{MIS}})}$

- ▶ ¿Por qué se cumple esto en general?

$S' = \{r, \neg s\}$ también es un conjunto independiente maximal de G

- ▶ S' representa a una valuación σ que satisface φ : $\sigma(r) = 1$ y $\sigma(s) = 0$
- ▶ S' es representado por el string 00010000100
- ▶ Tenemos que $(G, S) \in \overline{\text{Min}(R_{\text{MIS}})}$ puesto que 00010000100 es menor que 00100000000 en orden lexicográfico

Ejercicios

1. Para $\varphi = (r \vee t) \vee (t \vee \neg s \vee \neg t \vee \neg u)$, encuentre S'' tal que $(G, S'') \in \text{Min}(R_{\text{MIS}})$
2. Generalice la construcción mostrada para cualquier fórmula proposicional φ en CNF
 - ▶ Si φ menciona m cláusulas y n variables proposicionales, entonces el grafo G debe tener $m + 2 \cdot n + 1$ nodos
 - ▶ Como en el ejemplo, se debe tener que $S = \{\star\}$
3. Para la construcción realizada en 2, demuestre que φ es satisfacible si y sólo si $(G, S) \in \overline{\text{Min}}(R_{\text{MIS}})$

Teorema (Jerrum, Valiant & Vazirani)

Sea R una p -relación auto-reducible. Si existe un FPAUG para R , entonces existe un FPRAS para R .

La idea de la demostración: primer ejemplo

Vamos a mostrar algunas de las ideas fundamentales de la demostración considerando $\#SAT$

- ▶ Recuerde que $R_{SAT} = \{(\varphi, \sigma) \mid \varphi \text{ es una fórmula proposicional y } \sigma \text{ es una valuación tal que } \sigma(\varphi) = 1\}$ es una p-relación auto-reducible

Suponemos que tenemos un generador uniforme para R_{SAT}

- ▶ Si este generador funciona en tiempo polinomial entonces vamos a obtener un FPRAS para $\#SAT$
- ▶ En la demostración del teorema la hipótesis será que existe un FPAUG para la relación que estemos considerando

La idea de la demostración: #SAT

Sea $\mathcal{G}$ un generador uniforme para R_{SAT} que funciona en tiempo polinomial.

Para cada fórmula proposicional φ tenemos:

- ▶ si $\sigma(\varphi) = 0$, entonces $\Pr(\mathcal{G}(\varphi) = \sigma) = 0$
- ▶ si $\sigma(\varphi) = 1$, entonces $\Pr(\mathcal{G}(\varphi) = \sigma) = \frac{1}{\# \text{SAT}(\varphi)}$

La idea de la demostración: #SAT

Sea φ un fórmula proposicional y $\{x_1, \dots, x_n\}$ el conjunto de variables mencionadas en φ

Podemos utilizar $\mathcal{G}$ para generar una valuación σ tal que $\sigma(\varphi) = 1$

- Si $\mathcal{G}(\varphi) = \perp$, entonces sabemos que φ no es satisfacible y $\#SAT(\varphi) = 0$

Suponemos que $\sigma(x_i) = v_i$ para cada $i \in \{1, \dots, n\}$

La idea de la demostración: #SAT

Tenemos que $\#SAT(\varphi[\frac{x_1}{v_1}, \dots, \frac{x_n}{v_n}]) = 1$

- ▶ $\varphi[\frac{x_1}{v_1}, \dots, \frac{x_n}{v_n}]$ es obtenida desde φ reemplazando cada variable x_i por el valor v_i
 - ▶ Si $v_i = 0$ reemplazamos x_i por el operador 0-ario $\perp$, y si $v_i = 1$ reemplazamos x_i por el operador 0-ario $\top$
 - ▶ $\varphi[\frac{x_1}{v_1}, \dots, \frac{x_n}{v_n}]$ no tiene variables

Así, tenemos que:

$$\#SAT(\varphi) = \frac{\#SAT(\varphi)}{\#SAT(\varphi[\frac{x_1}{v_1}, \dots, \frac{x_n}{v_n}])} \cdot \#SAT(\varphi[\frac{x_1}{v_1}, \dots, \frac{x_n}{v_n}])$$

La idea de la demostración: #SAT

Por lo tanto, tenemos que:

$$\#SAT(\varphi) = \frac{1}{\rho}$$

donde:

$$\rho = \frac{\#SAT(\varphi[\frac{x_1}{v_1}, \dots, \frac{x_n}{v_n}])}{\#SAT(\varphi)}$$

Podemos entonces estimar $\#SAT(\varphi)$ utilizando una estimación para ρ

- Utilizamos $\mathcal{G}$ para estimar ρ

Estimando ρ

Sea X una variable aleatoria que toma valor 1 para σ , y toma valor 0 para cada valuación σ' tal que $\sigma'(\varphi) = 1$ y $\sigma' \neq \sigma$

- ▶ En particular, tenemos que $X \sim \mathbf{Ber}(\rho)$

Tenemos que $\mathbf{E}[X] = \rho$, por lo que podemos estimar ρ a través del muestreo de X

- ▶ Realizamos $t \geq 1$ nuestras independientes $X_1, \dots, X_t$ de X , y utilizamos como estimador el promedio $\bar{X} = \frac{1}{t} \sum_{i=1}^t X_i$
 - ▶ Puesto que $\mathbf{E}[\bar{X}] = \rho$ y tiene una menor varianza

Estimando ρ

Para estimar ρ utilizamos el siguiente algoritmo:

```
 $f_{av} := 0$   
for  $j := 1$  to  $t$  do  
     $\sigma' := \mathcal{G}(\varphi)$   
    if  $\sigma' = \sigma$  then  
         $f_{av} := f_{av} + 1$   
return  $\frac{f_{av}}{t}$ 
```

¿Qué tan buena es la estimación de ρ ? ¿Cuántas muestras t debemos realizar para tener una buena estimación de ρ ?

Estimando ρ

Usando la desigualdad de Chebyshev obtenemos:

$$\begin{aligned}\Pr(|\bar{X} - \mathbf{E}[\bar{X}]| \geq \varepsilon \cdot \mathbf{E}[\bar{X}]) &\leq \frac{\mathbf{Var}[\bar{X}]}{\varepsilon^2 \cdot \mathbf{E}[\bar{X}]^2} \\ &= \frac{\mathbf{Var}[X]}{t \cdot \varepsilon^2 \cdot \mathbf{E}[X]^2} \\ &= \frac{\rho \cdot (1 - \rho)}{t \cdot \varepsilon^2 \cdot \mathbf{E}[X]^2} \\ &\leq \frac{1}{t \cdot \varepsilon^2 \cdot \mathbf{E}[X]^2}\end{aligned}$$

Por lo tanto, si $\frac{4}{\varepsilon^2 \cdot \mathbf{E}[X]^2} \leq t$, entonces:

$$\Pr(|\bar{X} - \mathbf{E}[\bar{X}]| \geq \varepsilon \cdot \mathbf{E}[\bar{X}]) \leq \frac{1}{4}$$

Estimando ρ

Para obtener una buena estimación de ρ realizamos entonces $t = \lceil \frac{4}{\varepsilon^2 \cdot \mathbf{E}[X]^2} \rceil$ muestras

- ▶ ¿Cuán grande es t ?

Dado que $\mathbf{E}[X] = \rho = \frac{1}{\#\text{SAT}(\varphi)}$, obtenemos:

$$t = \left\lceil \frac{4 \cdot (\#\text{SAT}(\varphi))^2}{\varepsilon^2} \right\rceil$$

Por lo tanto, t puede ser exponencial en el tamaño de φ

- ▶ ¿Puede ser solucionado este problema utilizando una valuación distinta de σ que satisfaga φ ?

Reduciendo el número de muestras

Consideramos nuevamente la valuación σ que satisface φ

- Recuerde que $\sigma(x_i) = v_i$ para cada $i \in \{1, \dots, n\}$

Tenemos que:

$$\begin{aligned} \#SAT(\varphi) = & \frac{\#SAT(\varphi)}{\#SAT(\varphi[\frac{x_1}{v_1}])} \cdot \frac{\#SAT(\varphi[\frac{x_1}{v_1}])}{\#SAT(\varphi[\frac{x_1}{v_1}, \frac{x_2}{v_2}])} \cdot \frac{\#SAT(\varphi[\frac{x_1}{v_1}, \frac{x_2}{v_2}])}{\#SAT(\varphi[\frac{x_1}{v_1}, \frac{x_2}{v_2}, \frac{x_3}{v_3}])} \cdot \dots \\ & \cdot \frac{\#SAT(\varphi[\frac{x_1}{v_1}, \dots, \frac{x_{n-1}}{v_{n-1}}])}{\#SAT(\varphi[\frac{x_1}{v_1}, \dots, \frac{x_n}{v_n}])} \cdot \#SAT(\varphi[\frac{x_1}{v_1}, \dots, \frac{x_n}{v_n}]) \end{aligned}$$

Reduciendo el número de muestras

Por lo tanto:

$$\begin{aligned} \#SAT(\varphi) = & \frac{1}{\left(\frac{\#SAT(\varphi[\frac{x_1}{v_1}])}{\#SAT(\varphi)} \right)} \cdot \frac{1}{\left(\frac{\#SAT(\varphi[\frac{x_1}{v_1}, \frac{x_2}{v_2}])}{\#SAT(\varphi[\frac{x_1}{v_1}])} \right)} \cdot \\ & \frac{1}{\left(\frac{\#SAT(\varphi[\frac{x_1}{v_1}, \frac{x_2}{v_2}, \frac{x_3}{v_3}])}{\#SAT(\varphi[\frac{x_1}{v_1}, \frac{x_2}{v_2}])} \right)} \cdot \dots \cdot \frac{1}{\left(\frac{\#SAT(\varphi[\frac{x_1}{v_1}, \dots, \frac{x_n}{v_n}])}{\#SAT(\varphi[\frac{x_1}{v_1}, \dots, \frac{x_{n-1}}{v_{n-1}}])} \right)} \end{aligned}$$

Definiendo $\rho_i = \frac{\#SAT(\varphi[\frac{x_1}{v_1}, \dots, \frac{x_i}{v_i}])}{\#SAT(\varphi[\frac{x_1}{v_1}, \dots, \frac{x_{i-1}}{v_{i-1}}])}$, obtenemos:

$$\#SAT(\varphi) = \prod_{i=1}^n \frac{1}{\rho_i}$$

Estimando ρ_1

Sea X una variable aleatoria tal que para toda valuación σ' que satisface φ :

$$X(\sigma') = \begin{cases} 1 & \text{si } \sigma'(x_1) = v_1 \\ 0 & \text{en caso contrario} \end{cases}$$

Tenemos que $X \sim \mathbf{Ber}(\rho_1)$

Realizamos $t \geq 1$ nuestras independientes $X_1, \dots, X_t$ de X , y utilizamos como estimador el promedio $\bar{X} = \frac{1}{t} \sum_{i=1}^t X_i$

Estimando ρ_1

Para estimar ρ_1 utilizamos el siguiente algoritmo:

```
 $f_{av} := 0$   
for  $j := 1$  to  $t$  do  
     $\sigma' := \mathcal{G}(\varphi)$   
    if  $\sigma'(x_1) = v_1$  then  
         $f_{av} := f_{av} + 1$   
return  $\frac{f_{av}}{t}$ 
```

¿Qué tan buena es la estimación de ρ_1 ? ¿Solucionamos el problema que teníamos con el enfoque anterior?

Estimando ρ_1

Usando nuevamente la desigualdad de Chebyshev obtenemos:

$$\Pr(|\bar{X} - \mathbf{E}[\bar{X}]| \geq \varepsilon \cdot \mathbf{E}[\bar{X}]) \leq \frac{1}{t \cdot \varepsilon^2 \cdot \mathbf{E}[X]^2}$$

Entonces realizamos $t = \lceil \frac{4}{\varepsilon^2 \cdot \mathbf{E}[X]^2} \rceil$ muestras

- ▶ Dado que $\mathbf{E}[X] = \rho_1 = \frac{\#\text{SAT}(\varphi[\frac{x_1}{v_1}])}{\#\text{SAT}(\varphi)}$, obtenemos:

$$t = \left\lceil \frac{4 \cdot (\#\text{SAT}(\varphi))^2}{\varepsilon^2 \cdot (\#\text{SAT}(\varphi[\frac{x_1}{v_1}]))^2} \right\rceil$$

Por lo tanto, nuevamente t puede ser exponencial en el tamaño de φ

- ▶ Por ejemplo, si $\#\text{SAT}(\varphi[\frac{x_1}{v_1}]) = 1$ y $\#\text{SAT}(\varphi) = 2^{n-1} + 1$

Reduciendo el valor de t

¿Qué salió mal? El problema de elegir σ antes de realizar la estimación de ρ_1 es que el valor $\mathbf{E}[X]$ puede ser muy pequeño, por lo que el valor $\frac{1}{\mathbf{E}[X]^2}$ puede ser muy grande.

- ▶ En el caso anterior podíamos tener que $\mathbf{E}[X] = \frac{1}{2^{n-1}+1}$, por lo que $\frac{1}{\mathbf{E}[X]^2} = (2^{n-1} + 1)^2$

Para evitar este problema, tenemos que elegir el valor v_1 por el que vamos a reemplazar la variable x_1 **después** de realizar las t muestras.

Reduciendo el valor de t

Sea:

$$\alpha = \frac{\#\text{SAT}(\varphi[\frac{x_1}{0}])}{\#\text{SAT}(\varphi)} \quad \text{y} \quad \beta = \frac{\#\text{SAT}(\varphi[\frac{x_1}{1}])}{\#\text{SAT}(\varphi)}$$

Nótese que $\alpha + \beta = 1$

Sean Y, Z variables aleatorias tales que para toda valuación σ que satisface φ :

$$Y(\sigma) = \begin{cases} 1 & \text{si } \sigma(x_1) = 0 \\ 0 & \text{si } \sigma(x_1) = 1 \end{cases}$$
$$Z(\sigma) = \begin{cases} 1 & \text{si } \sigma(x_1) = 1 \\ 0 & \text{si } \sigma(x_1) = 0 \end{cases}$$

Nótese que $Y \sim \mathbf{Ber}(\alpha)$, $Z \sim \mathbf{Ber}(\beta)$ y $Z = 1 - Y$

Reduciendo el valor de t

Consideramos los estimadores $\overline{Y}$ y $\overline{Z}$ que calculamos con el siguiente algoritmo:

```
 $fav_Y := 0$   
 $fav_Z := 0$   
for  $j := 1$  to  $t$  do  
   $\sigma := \mathcal{G}(\varphi)$   
  if  $\sigma(x_1) = 0$   
    then  $fav_Y := fav_Y + 1$   
    else  $fav_Z := fav_Z + 1$   
return  $(\frac{fav_Y}{t}, \frac{fav_Z}{t})$ 
```

Reduciendo el valor de t

Finalmente reemplazamos x_1 por:

$$v_1 = \begin{cases} 0 & \text{si } \bar{Y} \geq \bar{Z} \\ 1 & \text{si } \bar{Y} < \bar{Z} \end{cases}$$

Lo cual corresponde a utilizar el siguiente estimador:

$$X = \max\{\bar{Y}, \bar{Z}\}$$

Vale decir, reemplazamos x_1 por el valor v_1 que esperamos que aparezca un mayor número de veces en las valuaciones que satisfacen φ

Reduciendo el valor de t

¿Solucionamos el problema con el valor de $\mathbf{E}[X]$? ¡Sí!

Dado que $\overline{Y} \leq X$ y $\overline{Z} \leq X$, tenemos que $\mathbf{E}[\overline{Y}] \leq \mathbf{E}[X]$ y $\mathbf{E}[\overline{Z}] \leq \mathbf{E}[X]$

► Por lo tanto $\alpha \leq \mathbf{E}[X]$ y $\beta \leq \mathbf{E}[X]$

Tenemos que $\alpha + \beta \leq 2 \cdot \mathbf{E}[X]$, de lo cual concluimos $\frac{1}{2} \leq \mathbf{E}[X]$

Concluimos entonces que:

$$\frac{1}{\mathbf{E}[X]^2} \leq 4$$

¿Qué más debemos hacer?

Debemos realizar procedimientos de estimación similares para $\rho_2, \dots, \rho_n$

- ▶ En todos ellos usamos el mismo valor de muestras t

Además, debemos calcular cómo los errores en las estimaciones de $\rho_1, \dots, \rho_n$ se componen para obtener un error para la estimación de

$$\prod_{i=1}^n \rho_i$$

¿Qué más debemos hacer?

Finalmente, a partir de la estimación de $\prod_{i=1}^n \rho_i$, obtenemos una estimación de:

$$\#SAT(\varphi) = \prod_{i=1}^n \frac{1}{\rho_i} = \frac{1}{\prod_{i=1}^n \rho_i}$$

Esto nos da como resultado un FPRAS para $\#SAT$.

Vamos a mostrar en un segundo ejemplo como las estimaciones $\rho_1, \dots, \rho_n$ son utilizadas para generar un FPRAS para un problema de conteo a partir de un generador uniforme de soluciones.

La idea de la demostración: segundo ejemplo

En este segundo ejemplo vamos a explicar cómo componer los errores en las estimaciones locales para obtener una cota superior en el error total de la estimación.

Y también queremos considerar una forma alternativa de auto-reducción.

- Hay muchas formas de definir una noción de relación auto-reducible

La idea de la demostración: segundo ejemplo

Utilizamos $\cdot$ para denotar el producto interior usual en $\mathbb{R}^n$ ($n \geq 1$)

Considere el siguiente problema:

$$\text{KS} = \{(\vec{a}, b) \mid \vec{a} \in \mathbb{N}^n \text{ para } n \geq 1, \\ b \in \mathbb{Z} \text{ y existe } \vec{x} \in \{0, 1\}^n \text{ tal que } \vec{a} \cdot \vec{x} \leq b\}$$

KS es una versión simplificada del problema de la mochila, de hecho tenemos que $\text{KS} \in \text{PTIME}$

KS como una relación y el problema de conteo asociado

Ejercicio

Podemos representar KS como la siguiente relación:

$$R_{\text{KS}} = \{((\vec{a}, b), \vec{x}) \mid \vec{a} \in \mathbb{N}^n \text{ y } \vec{x} \in \{0, 1\}^n \text{ para } n \geq 1, b \in \mathbb{Z} \text{ y } \vec{a} \cdot \vec{x} \leq b\}$$

Demuestre que R_{KS} es auto-reducible.

Definimos la función de conteo $\#KS$ como $\#KS(\vec{a}, b) = N_{R_{\text{KS}}}((\vec{a}, b))$

- Suponiendo que $\vec{a} \in \mathbb{N}^n$, tenemos que $\#KS(\vec{a}, b)$ es el número de vectores $\vec{x} \in \{0, 1\}^n$ tales que $\vec{a} \cdot \vec{x} \leq b$

La idea de la demostración para KS

Sea $(\vec{a}, b)$ una entrada de $\#KS$

- ▶ Suponemos que $\vec{a} = (a_1, \dots, a_n)$ con $0 < a_1 \leq \dots \leq a_n$ y $b \geq 0$
 - ▶ ¿Por qué podemos suponer esto?

Definimos $b_0 = 0$ y para cada $i \in \{1, \dots, n\}$:

$$b_i = \min \left\{ \sum_{j=1}^i a_j, b \right\}$$

Es importante notar que:

$$\begin{aligned} \#KS(\vec{a}, b_0) &= 1 \\ \#KS(\vec{a}, b_i) &\leq \#KS(\vec{a}, b_{i+1}) \quad \text{para todo } i \in \{0, \dots, n-1\} \\ \#KS(\vec{a}, b_n) &= \#KS(\vec{a}, b) \end{aligned}$$

La idea de la demostración para KS

De la misma forma que para #SAT, la demostración se basa en la igualdad:

$$\begin{aligned} \#KS(\vec{a}, b) = \#KS(\vec{a}, b_n) &= \frac{\#KS(\vec{a}, b_n)}{\#KS(\vec{a}, b_{n-1})} \cdot \frac{\#KS(\vec{a}, b_{n-1})}{\#KS(\vec{a}, b_{n-2})} \cdot \dots \\ &\quad \frac{\#KS(\vec{a}, b_1)}{\#KS(\vec{a}, b_0)} \cdot \#KS(\vec{a}, b_0) \end{aligned}$$

Para cada $i \in \{1, \dots, n\}$ definimos:

$$\rho_i = \frac{\#KS(\vec{a}, b_{i-1})}{\#KS(\vec{a}, b_i)}$$

La idea de la demostración para KS

Tenemos que $0 < \rho_i \leq 1$ para cada $i \in \{1, \dots, n\}$

Considerando que $\#KS(\vec{a}, b_0) = 1$, concluimos que:

$$\frac{1}{\#KS(\vec{a}, b)} = \prod_{i=1}^n \rho_i$$

Por lo tanto, si logramos tener buenas estimaciones de cada ρ_i podemos obtener una buena estimación de $\frac{1}{\#KS(\vec{a}, b)}$

► Y de esta forma de $\#KS(\vec{a}, b)$

Estimando ρ_i

Sea X_i una variable aleatoria tal que para toda valuación $\vec{x} \in \{0, 1\}^n$ que satisface $\vec{a} \cdot \vec{x} \leq b_i$:

$$X_i(\vec{x}) = \begin{cases} 1 & \text{si } \vec{a} \cdot \vec{x} \leq b_{i-1} \\ 0 & \text{en caso contrario} \end{cases}$$

Tenemos que $X_i \sim \mathbf{Ber}(\rho_i)$

Realizamos $t \geq 1$ muestras independientes $Y_{i,1}, \dots, Y_{i,t}$ de X_i , y utilizamos como estimador el promedio $\overline{Y}_i = \frac{1}{t} \cdot \sum_{j=1}^t Y_{i,j}$

- Recuerde que $\mathbf{E}[\overline{Y}_i] = \rho_i$ y tiene una menor varianza

¿Pero como podemos muestrear X_i ?

En este punto necesitamos suponer que tenemos un generador uniforme para R_{KS}

- ▶ Si este generador funciona en tiempo polinomial entonces vamos a obtener un FPRAS para $\#KS$
- ▶ En la demostración del teorema la hipótesis será que existe un FPAUG para la relación que estemos considerando

Sea $\mathcal{G}$ el generador uniforme que necesitamos. Para cada entrada $(\vec{c}, d)$ de KS con $\vec{c} \in \mathbb{N}^m$, y cada vector $\vec{y} \in \mathbb{N}^m$ tenemos:

- ▶ si $\vec{c} \cdot \vec{y} > d$, entonces $\Pr(\mathcal{G}(\vec{c}, d) = \vec{y}) = 0$
- ▶ si $\vec{c} \cdot \vec{y} \leq d$, entonces $\Pr(\mathcal{G}(\vec{c}, d) = \vec{y}) = \frac{1}{\#KS(\vec{c}, d)}$

Estimando ρ_i

Para estimar $\rho_i = \frac{\#KS(\vec{a}, b_{i-1})}{\#KS(\vec{a}, b_i)}$ utilizamos el siguiente algoritmo:

```
 $f_{av} := 0$   
for  $j := 1$  to  $t$  do  
     $\vec{v} := \mathcal{G}(\vec{a}, b_i)$   
    if  $\vec{a} \cdot \vec{v} \leq b_{i-1}$  then  
         $f_{av} := f_{av} + 1$   
return  $\frac{f_{av}}{t}$ 
```

¿Qué tan buena es la estimación de ρ_i ? ¿Que tan buena es la estimación de $\frac{1}{\#KS(\vec{a}, b)}$ dadas las estimaciones de $\rho_1, \dots, \rho_n$?

- Tenemos que acotar la probabilidad de error

La probabilidad de error

$$\text{Sea } Z = \prod_{i=1}^n \overline{Y_i}$$

$$\text{Y sea } \varepsilon \in (0, 1)$$

Para obtener un FPRAS para $\#KS$, primero tenemos que acotar superiormente la siguiente probabilidad:

$$\Pr\left(\left|Z - \frac{1}{\#KS(\vec{a}, b)}\right| \geq \varepsilon \cdot \frac{1}{\#KS(\vec{a}, b)}\right)$$

La probabilidad de error

Dado que $\overline{Y}_i$ es independiente de $\overline{Y}_j$ para $i \neq j$; tenemos que:

$$\mathbf{E}[Z] = \mathbf{E}\left[\prod_{i=1}^n \overline{Y}_i\right] = \prod_{i=1}^n \mathbf{E}[\overline{Y}_i] = \prod_{i=1}^n \rho_i = \frac{1}{\#\text{KS}(\vec{a}, b)}$$

Usando entonces la desigualdad de Chebyshev obtenemos:

$$\begin{aligned} \Pr\left(\left|Z - \frac{1}{\#\text{KS}(\vec{a}, b)}\right| \geq \varepsilon \cdot \frac{1}{\#\text{KS}(\vec{a}, b)}\right) &= \Pr(|Z - \mathbf{E}[Z]| \geq \varepsilon \cdot \mathbf{E}[Z]) \\ &\leq \frac{\mathbf{Var}[Z]}{\varepsilon^2 \cdot \mathbf{E}[Z]^2} \end{aligned}$$

La probabilidad de error

Pero tenemos que:

$$\begin{aligned}\frac{\text{Var}[Z]}{\varepsilon^2 \cdot \mathbf{E}[Z]^2} &= \frac{1}{\varepsilon^2} \cdot \frac{\text{Var}[Z]}{\mathbf{E}[Z]^2} \\&= \frac{1}{\varepsilon^2} \cdot \left(\frac{\mathbf{E}[Z^2] - \mathbf{E}[Z]^2}{\mathbf{E}[Z]^2} \right) \\&= \frac{1}{\varepsilon^2} \cdot \left(\frac{\mathbf{E}[Z^2]}{\mathbf{E}[Z]^2} - 1 \right) \\&= \frac{1}{\varepsilon^2} \cdot \left(\frac{\mathbf{E}[(\prod_{i=1}^n \overline{Y}_i)^2]}{(\prod_{i=1}^n \mathbf{E}[Y_i])^2} - 1 \right) \\&= \frac{1}{\varepsilon^2} \cdot \left(\frac{\mathbf{E}[\prod_{i=1}^n \overline{Y}_i^2]}{\prod_{i=1}^n \mathbf{E}[Y_i]^2} - 1 \right) \\&= \frac{1}{\varepsilon^2} \cdot \left(\frac{\prod_{i=1}^n \mathbf{E}[\overline{Y}_i^2]}{\prod_{i=1}^n \mathbf{E}[Y_i]^2} - 1 \right)\end{aligned}$$

La probabilidad de error

$$\begin{aligned} &= \frac{1}{\varepsilon^2} \cdot \left(\prod_{i=1}^n \frac{\mathbf{E}[\overline{Y}_i^2]}{\mathbf{E}[\overline{Y}_i]^2} - 1 \right) \\ &= \frac{1}{\varepsilon^2} \cdot \left(\prod_{i=1}^n \frac{(\mathbf{Var}[\overline{Y}_i] + \mathbf{E}[\overline{Y}_i]^2)}{\mathbf{E}[\overline{Y}_i]^2} - 1 \right) \\ &= \frac{1}{\varepsilon^2} \cdot \left(\prod_{i=1}^n \left[1 + \frac{\mathbf{Var}[\overline{Y}_i]}{\mathbf{E}[\overline{Y}_i]^2} \right] - 1 \right) \\ &= \frac{1}{\varepsilon^2} \cdot \left(\prod_{i=1}^n \left[1 + \frac{\mathbf{Var}[\overline{Y}_i]}{\rho_i^2} \right] - 1 \right) \end{aligned}$$

Por lo tanto necesitamos acotar superiormente $\frac{\mathbf{Var}[\overline{Y}_i]}{\rho_i^2}$

Acotando superiormente $\frac{\text{Var}[\bar{Y}_i]}{\rho_i^2}$

Tenemos que:

$$\begin{aligned}\frac{\text{Var}[\bar{Y}_i]}{\rho_i^2} &= \frac{\text{Var}\left[\frac{1}{t} \sum_{j=1}^t Y_{i,j}\right]}{\rho_i^2} \\&= \frac{\frac{1}{t^2} \cdot \text{Var}\left[\sum_{j=1}^t Y_{i,j}\right]}{\rho_i^2} \\&= \frac{\frac{1}{t^2} \cdot \sum_{j=1}^t \text{Var}[Y_{i,j}]}{\rho_i^2} \\&= \frac{\frac{1}{t^2} \cdot \sum_{j=1}^t \rho_i \cdot (1 - \rho_i)}{\rho_i^2} \\&= \frac{\frac{1}{t^2} \cdot t \cdot \rho_i \cdot (1 - \rho_i)}{\rho_i^2} \\&= \frac{1}{t} \cdot \left(\frac{1}{\rho_i} - 1\right)\end{aligned}$$

Lema

Para cada $i \in \{1, \dots, n\}$ se tiene que:

$$\#KS(\vec{a}, b_{i-1}) \leq \#KS(\vec{a}, b_i) \leq (n+1) \cdot \#KS(\vec{a}, b_{i-1})$$

Ejercicios

1. Sea $\vec{x} \in \{0, 1\}^n$ tal que $\vec{x} = (x_1, \dots, x_n)$ y $\vec{a} \cdot \vec{x} \leq b_i$. Demuestre que si $\vec{a} \cdot \vec{x} > b_{i-1}$, entonces existe una posición $j \in \{1, \dots, n\}$ tal que $x_j = 1$ y para el vector $\vec{y} = (x_1, \dots, x_{j-1}, 0, x_{j+1}, \dots, x_n)$ se tiene que $\vec{a} \cdot \vec{y} \leq b_{i-1}$.
2. Demuestre que el lema es consecuencia de la propiedad anterior.

Acotando superiormente $\frac{\text{Var}[\overline{Y}_i]}{\rho_i^2}$

Del lema concluimos que para cada $i \in \{1, \dots, n\}$:

$$\frac{1}{\rho_i} = \frac{\#\text{KS}(\vec{a}, b_i)}{\#\text{KS}(\vec{a}, b_{i-1})} \leq (n+1)$$

Por lo tanto:

$$\frac{\text{Var}[\overline{Y}_i]}{\rho_i^2} = \frac{1}{t} \cdot \left(\frac{1}{\rho_i} - 1 \right) \leq \frac{n}{t}$$

Acotando superiormente $\frac{\text{Var}[Z]}{\varepsilon^2 \cdot \mathbf{E}[Z]^2}$

De los cálculos anteriores concluimos que:

$$\begin{aligned}\frac{\text{Var}[Z]}{\varepsilon^2 \cdot \mathbf{E}[Z]^2} &= \frac{1}{\varepsilon^2} \cdot \left(\prod_{i=1}^n \left[1 + \frac{\text{Var}[\bar{Y}_i]}{\rho_i^2} \right] - 1 \right) \\ &\leq \frac{1}{\varepsilon^2} \cdot \left(\prod_{i=1}^n \left[1 + \frac{n}{t} \right] - 1 \right) \\ &= \frac{1}{\varepsilon^2} \cdot \left(\left[1 + \frac{n}{t} \right]^n - 1 \right)\end{aligned}$$

Escogemos t de manera de hacer pequeña la cota superior para $\frac{\text{Var}[Z]}{\varepsilon^2 \cdot \mathbf{E}[Z]^2}$

- ▶ Vamos a escoger t de manera que sea polinomial en n y $\frac{1}{\varepsilon}$, dado que n es menor que el tamaño de la entrada ($\vec{b}, a$)
- ▶ Esperamos que n y $\frac{1}{\varepsilon^2}$ sean valores grandes, por lo que t debe disminuir el impacto de estos valores
 - ▶ Tomamos $t = c \cdot n^2 \cdot \varepsilon^{-2}$, donde c es una constante

El valor de t y una cota superior para $\frac{\mathbf{Var}[Z]}{\varepsilon^2 \cdot \mathbf{E}[Z]^2}$

Tomamos $t = 5 \cdot n^2 \cdot \varepsilon^{-2}$

Tenemos que:

$$\begin{aligned} \frac{\mathbf{Var}[Z]}{\varepsilon^2 \cdot \mathbf{E}[Z]^2} &\leq \frac{1}{\varepsilon^2} \cdot \left(\left[1 + \frac{n}{t} \right]^n - 1 \right) \\ &= \frac{1}{\varepsilon^2} \cdot \left(\left[1 + \frac{\varepsilon^2}{5 \cdot n} \right]^n - 1 \right) \end{aligned}$$

El valor de t y una cota superior para $\frac{\text{Var}[Z]}{\varepsilon^2 \cdot \mathbf{E}[Z]^2}$

Además, tenemos que:

$$\begin{aligned} \left[1 + \frac{\varepsilon^2}{5 \cdot n}\right]^n &= \sum_{i=0}^n \binom{n}{i} \left(\frac{\varepsilon^2}{5 \cdot n}\right)^i \\ &= \sum_{i=0}^n \frac{n!}{i! \cdot (n-i)! \cdot n^i} \cdot \left(\frac{\varepsilon^2}{5}\right)^i \\ &\leq \sum_{i=0}^n \frac{1}{i!} \cdot \left(\frac{\varepsilon^2}{5}\right)^i \\ &< \sum_{i=0}^{\infty} \frac{\left(\frac{\varepsilon^2}{5}\right)^i}{i!} \\ &= e^{\frac{\varepsilon^2}{5}} \end{aligned}$$

El valor de t y una cota superior para $\frac{\text{Var}[Z]}{\varepsilon^2 \cdot \mathbf{E}[Z]^2}$

Lema

$$e^{\frac{\varepsilon^2}{5}} \leq \frac{\varepsilon^2}{4} + 1$$

Ejercicio

Demuestre el lema considerando que $0 < \frac{\varepsilon^2}{4} < \frac{1}{4}$ y el intervalo donde la función $f(x) = e^{\frac{4}{5} \cdot x} - x - 1$ es negativa.

Concluimos que:

$$\left[1 + \frac{\varepsilon^2}{5 \cdot n}\right]^n \leq e^{\frac{\varepsilon^2}{5}} \leq \frac{\varepsilon^2}{4} + 1$$

El valor de t y una cota superior para $\frac{\text{Var}[Z]}{\varepsilon^2 \cdot \mathbf{E}[Z]^2}$

Finalmente obtenemos que:

$$\begin{aligned}\frac{\text{Var}[Z]}{\varepsilon^2 \cdot \mathbf{E}[Z]^2} &\leq \frac{1}{\varepsilon^2} \cdot \left(\left[1 + \frac{\varepsilon^2}{5 \cdot n} \right]^n - 1 \right) \\ &\leq \frac{1}{\varepsilon^2} \cdot \left(\frac{\varepsilon^2}{4} + 1 - 1 \right) \\ &= \frac{1}{\varepsilon^2} \cdot \left(\frac{\varepsilon^2}{4} \right) \\ &= \frac{1}{4}\end{aligned}$$

Por lo tanto, para todo $0 < \varepsilon < 1$ obtenemos:

$$\Pr\left(\left|Z - \frac{1}{\#\text{KS}(\vec{a}, b)}\right| \geq \varepsilon \cdot \frac{1}{\#\text{KS}(\vec{a}, b)}\right) \leq \frac{\text{Var}[Z]}{\varepsilon^2 \cdot \mathbf{E}[Z]^2} \leq \frac{1}{4} \quad (\dagger)$$

Obteniendo un FPRAS para $\#KS$

Dado $0 < \delta < 1$, para terminar debemos demostrar que:

$$\Pr(|Z^{-1} - \#KS(\vec{a}, b)| \leq \delta \cdot \#KS(\vec{a}, b)) \geq \frac{3}{4}$$

De esta forma Z^{-1} nos da un FPRAS para $\#KS$

Considerando $\varepsilon = \frac{\delta}{2}$ en (b) obtenemos:

$$\Pr\left(\left(1 - \frac{\delta}{2}\right) \cdot \frac{1}{\#KS(\vec{a}, b)} \leq Z \leq \left(1 + \frac{\delta}{2}\right) \cdot \frac{1}{\#KS(\vec{a}, b)}\right) \geq \frac{3}{4}$$

Dado que $0 < \frac{\delta}{2} < 1$, entonces tenemos que:

$$\Pr\left(\frac{1}{1 + \frac{\delta}{2}} \cdot \#KS(\vec{a}, b) \leq Z^{-1} \leq \frac{1}{1 - \frac{\delta}{2}} \cdot \#KS(\vec{a}, b)\right) \geq \frac{3}{4}$$

Obteniendo un FPRAS para #KS

Dado que $(1 - \delta) \leq \frac{1}{1 + \frac{\delta}{2}}$ y $\frac{1}{1 - \frac{\delta}{2}} \leq (1 + \delta)$, concluimos que:

$$\Pr((1 - \delta) \cdot \#KS(\vec{a}, b) \leq Z^{-1} \leq (1 + \delta) \cdot \#KS(\vec{a}, b)) \geq$$

$$\Pr\left(\frac{1}{1 + \frac{\delta}{2}} \cdot \#KS(\vec{a}, b) \leq Z^{-1} \leq \frac{1}{1 - \frac{\delta}{2}} \cdot \#KS(\vec{a}, b)\right) \geq \frac{3}{4}$$

Por lo tanto Z^{-1} nos da un FPRAS para #KS

- El número de pasos ejecutados por el algoritmo es polinomial en el tamaño de la entrada $(\vec{a}, b)$ y $\frac{1}{\delta}$ si suponemos que $\mathcal{G}$ funciona en tiempo polinomial, puesto que $\mathcal{G}$ es invocado $n \cdot t$ veces y:

$$n \cdot t = n \cdot 5 \cdot n^2 \cdot \epsilon^{-2} = 5 \cdot n^3 \cdot \left(\frac{\delta}{2}\right)^{-2} = \frac{20 \cdot n^3}{\delta^2}$$

La demostración general

Vamos a extender las ideas utilizadas para $\#SAT$ y $\#KS$ al caso general.

Vale decir, dada una p -relación R auto-reducible, vamos a demostrar que si existe un FPAUG para R , entonces existe un FPRAS para R

Algunos supuestos para la relación R

Suponemos que $R \subseteq \Sigma^* \times \Sigma^*$, y dados $x, w \in \Sigma^*$ definimos:

$$\text{Ext}_R(x, w) = \{y \in \Sigma^* \mid (x, y) \in R \text{ y existe } z \in \Sigma^* \text{ tal que } y = wz\}$$

Sean $g : \Sigma^* \rightarrow \mathbb{N}$, $\psi : \Sigma^* \times \Sigma^* \rightarrow \Sigma^*$ y $\sigma : \Sigma^* \rightarrow \mathbb{N}$ funciones que muestran que R es auto-reducible.

- De acuerdo a la definición vista en las transparencias anteriores

Además, sea $\mathcal{G} : \Sigma^* \times (0, 1) \rightarrow \Sigma^* \cup \{\perp\}$ un FPAUG para R

Finalmente, sean $c, d \in \mathbb{R}^+$ tales que $|\Sigma|^{\sigma(x)} \leq |x|^c + d$ para todo $x \in \Sigma^*$

- Sabemos que existe porque $\sigma(x) \in O(\log(|x|))$

Un esquema de aproximación para R

```
EAR( $x, \varepsilon$ )  
  if  $\mathcal{G}(x, \varepsilon) = \perp$  then return 0  
  else  
     $N := 1$   
     $m := g(x)$   
     $t := \lceil 180 \cdot (|x|^c + d)^3 \cdot m^3 \cdot \varepsilon^{-2} \rceil$   
    while  $g(x) > 0$  do  
      for  $j := 1$  to  $t$  do  
         $y_j := \mathcal{G}\left(x, \frac{\varepsilon}{5m}\right)$   
        Sea  $w \in \Sigma^{\sigma(x)}$  el prefijo de largo  $\sigma(x)$  más común en  $\{y_1, \dots, y_t\}$   
         $\alpha := \frac{|\{j \in \{1, \dots, t\} \mid y_j \in \text{Ext}_R(x, w)\}|}{t}$   
         $x := \psi(x, w)$   
         $N := \frac{1}{\alpha} \cdot N$       /* se tiene que  $\alpha > 0$  */  
  return  $N$ 
```


EAR es un FPRAS para R

Vamos a demostrar que **EAR** es un FPRAS para R

- ▶ Sean $x \in \Sigma^*$ y $\varepsilon \in (0, 1)$ una entrada de **EAR**

Si $N_R(x) = 0$, tenemos que **EAR**(x, ε) retorna el resultado correcto 0 dado que $\mathcal{G}$ es un FPAUG para R

- ▶ En el resto de la demostración suponemos que $N_R(x) > 0$

De la misma forma, **EAR**(x, ε) retorna el resultado correcto si $g(x) = 0$

- ▶ ¿Por qué?

En el resto de la demostración suponemos que $g(x) > 0$

EAR es un FPRAS para R

Tenemos que el valor de la función g disminuye en cada iteración

- ▶ ¿Por qué?

Sea s la cantidad total de iteraciones realizadas por el algoritmo.

- ▶ Tenemos que $s \leq g(x) = m$

EAR es un FPRAS para R

EAR funciona en tiempo polinomial en $|x|$ y $\frac{1}{\varepsilon}$

- ▶ Dado que R es una p -relación, sabemos que existe un polinomio fijo $p(u)$ tal que $m = g(x) \leq p(|x|)$
- ▶ Además, sabemos que m puede ser calculado en tiempo polinomial dada la definición de relación auto-reducible
- ▶ Finalmente, tenemos que $\mathcal{G}$ es un FPAUG para R , y **EAR** realiza a lo más $m \cdot \lceil 180 \cdot (|x|^c + d)^3 \cdot m^3 \cdot \varepsilon^{-2} \rceil$ llamadas a la función $\mathcal{G}$

Nos queda entonces por demostrar:

$$\Pr\left((1 - \varepsilon) \cdot N_R(x) \leq \mathbf{EAR}(x, \varepsilon) \leq (1 + \varepsilon) \cdot N_R(x)\right) \geq \frac{3}{4}$$

La propiedad central

Dado $i \in \{1, \dots, s\}$, para la iteración i del algoritmo sean:

- ▶ x_i, N_i los valores de las variables x y N al principio de la iteración
 - ▶ Tenemos que $x_1 = x$ y $N_1 = 1$
- ▶ w_i, α_i los valores de las variables w y α calculados en la iteración

Propiedad de aproximación

Para cada $i \in \{1, \dots, s\}$:

$$\frac{\alpha_i}{1 + \frac{\varepsilon}{2m}} \leq \frac{|\text{Ext}_R(x_i, w_i)|}{N_R(x_i)} \leq \left(1 + \frac{\varepsilon}{2m}\right) \cdot \alpha_i$$

Usando la propiedad de aproximación

Tenemos que:

$$\begin{aligned} \Pr\left((1 - \varepsilon) \cdot N_R(x) \leq \mathbf{EAR}(x, \varepsilon) \leq (1 + \varepsilon) \cdot N_R(x)\right) = \\ \Pr\left((1 - \varepsilon) \cdot N_R(x) \leq \mathbf{EAR}(x, \varepsilon) \leq (1 + \varepsilon) \cdot N_R(x) \mid \text{propiedad de aproximación}\right) \cdot \\ \Pr\left(\text{propiedad de aproximación}\right) + \\ \Pr\left((1 - \varepsilon) \cdot N_R(x) \leq \mathbf{EAR}(x, \varepsilon) \leq (1 + \varepsilon) \cdot N_R(x) \mid \overline{\text{propiedad de aproximación}}\right) \cdot \\ \Pr\left(\overline{\text{propiedad de aproximación}}\right) \geq \\ \Pr\left((1 - \varepsilon) \cdot N_R(x) \leq \mathbf{EAR}(x, \varepsilon) \leq (1 + \varepsilon) \cdot N_R(x) \mid \text{propiedad de aproximación}\right) \cdot \\ \Pr\left(\text{propiedad de aproximación}\right) \end{aligned}$$

Usando la propiedad de aproximación

Por lo tanto, para demostrar que **EAR** es un FPRAS para R basta acotar inferiormente la siguiente probabilidad:

$$\Pr\left((1 - \varepsilon) \cdot N_R(x) \leq \mathbf{EAR}(x, \varepsilon) \leq (1 + \varepsilon) \cdot N_R(x) \mid \text{propiedad de aproximación}\right) \cdot \Pr(\text{propiedad de aproximación})$$

Primero vamos a demostrar que la propiedad de aproximación es suficiente para tener una buena aproximación de $N_R(x)$

- Después de esto vamos a acotar inferiormente la probabilidad de que la propiedad de aproximación se cumpla

La propiedad de aproximación es suficiente

Definimos x_{s+1} y N_{s+1} como los valores de las variables x y N al final de la iteración s del algoritmo.

- ▶ Tenemos que $g(x_{s+1}) = 0$, $N_R(x_{s+1}) = 1$ y el valor retornado por el algoritmo es N_{s+1}
 - ▶ ¿Por qué?

Dado $i \in \{1, \dots, s\}$, tenemos que:

$$\begin{aligned}x_{i+1} &= \psi(x_i, w_i) \\ N_R(x_{i+1}) &= |\text{Ext}_R(x_i, w_i)| \\ N_{i+1} &= \frac{1}{\alpha_i} \cdot N_i\end{aligned}$$

La propiedad de aproximación es suficiente

Además, tenemos que:

$$\begin{aligned}\frac{\alpha_i}{\left(1 + \frac{\varepsilon}{2m}\right)} &\leq \frac{|\text{Ext}_R(x_i, w_i)|}{N_R(x_i)} \leq \left(1 + \frac{\varepsilon}{2m}\right) \cdot \alpha_i \\ \Rightarrow \frac{N_R(x_i)}{\left(1 + \frac{\varepsilon}{2m}\right)} &\leq \frac{|\text{Ext}_R(x_i, w_i)|}{\alpha_i} \leq \left(1 + \frac{\varepsilon}{2m}\right) \cdot N_R(x_i) \\ \Rightarrow \frac{N_R(x_i) \cdot N_i}{\left(1 + \frac{\varepsilon}{2m}\right)} &\leq |\text{Ext}_R(x_i, w_i)| \cdot \frac{N_i}{\alpha_i} \leq \left(1 + \frac{\varepsilon}{2m}\right) \cdot N_R(x_i) \cdot N_i \\ \Rightarrow \frac{N_R(x_i) \cdot N_i}{\left(1 + \frac{\varepsilon}{2m}\right)} &\leq N_R(x_{i+1}) \cdot N_{i+1} \leq \left(1 + \frac{\varepsilon}{2m}\right) \cdot N_R(x_i) \cdot N_i\end{aligned}$$

Así, la cantidad $N_R(x_i) \cdot N_i$ es *casi* una invariante del ciclo **while**

La propiedad de aproximación es suficiente

Suponiendo que la propiedad de aproximación es cierta concluimos que:

$$\frac{N_R(x_1) \cdot N_1}{\left(1 + \frac{\varepsilon}{2m}\right)^s} \leq N_R(x_{s+1}) \cdot N_{s+1} \leq \left(1 + \frac{\varepsilon}{2m}\right)^s \cdot N_R(x_1) \cdot N_1$$

Por lo tanto, dado que $s \leq m$ tenemos que:

$$\frac{N_R(x_1) \cdot N_1}{\left(1 + \frac{\varepsilon}{2m}\right)^m} \leq N_R(x_{s+1}) \cdot N_{s+1} \leq \left(1 + \frac{\varepsilon}{2m}\right)^m \cdot N_R(x_1) \cdot N_1$$

Algunas propiedades de $N_R(x_i) \cdot N_i$

Dado que $x_1 = x$ y $N_1 = 1$, tenemos que $N_R(x) = N_R(x_1) \cdot N_1$

- Recuerde que queremos calcular $N_R(x)$

Dado que $N_R(x_{s+1}) = 1$, tenemos que $N_R(x_{s+1}) \cdot N_{s+1} = N_{s+1}$

- Además, sabemos que **EAR**(x, ε) = N_{s+1}

La propiedad de aproximación es suficiente: conclusión

Juntando los resultados anteriores obtenemos:

$$\frac{N_R(x)}{\left(1 + \frac{\varepsilon}{2m}\right)^m} \leq \mathbf{EAR}(x, \varepsilon) \leq \left(1 + \frac{\varepsilon}{2m}\right)^m \cdot N_R(x)$$

Dado que $\varepsilon \in (0, 1)$, tal como el caso de #KS obtenemos:

$$\left(1 + \frac{\varepsilon}{2m}\right)^m \leq e^{\frac{\varepsilon}{2}} \leq \varepsilon + 1$$

Ejercicio

Demuestre que $e^{\frac{\varepsilon}{2}} \leq \varepsilon + 1$ considerando que $0 < \varepsilon < 1$ y el intervalo donde la función $f(x) = e^{\frac{x}{2}} - x - 1$ es negativa.

La propiedad de aproximación es suficiente: conclusión

Así, suponiendo que la propiedad de aproximación se cumple obtenemos:

$$\frac{N_R(x)}{(1 + \varepsilon)} \leq \mathbf{EAR}(x, \varepsilon) \leq (1 + \varepsilon) \cdot N_R(x)$$

Sabemos que $(1 - \varepsilon) \leq \frac{1}{1 + \varepsilon}$, puesto que $\varepsilon > 0$

Suponiendo que la propiedad de aproximación se cumple, obtenemos entonces:

$$(1 - \varepsilon) \cdot N_R(x) \leq \mathbf{EAR}(x, \varepsilon) \leq (1 + \varepsilon) \cdot N_R(x)$$

La propiedad de aproximación es suficiente: conclusión

Vale decir, hemos demostrado que:

$$\Pr\left((1 - \varepsilon) \cdot N_R(x) \leq \mathbf{EAR}(x, \varepsilon) \leq (1 + \varepsilon) \cdot N_R(x) \mid \begin{array}{c} \text{propiedad de aproximación} \end{array} \right) = 1$$

De esta forma, para terminar la demostración tenemos que demostrar:

$$\Pr\left(\text{propiedad de aproximación} \right) \geq \frac{3}{4}$$

Acotando inferiormente $\Pr$ (propiedad de aproximación)

Fije $i \in \{1, \dots, s\}$

- ▶ i corresponde a una iteración de **EAR**

Para cada $u \in \Sigma^{\sigma(x_i)}$ definimos la variable aleatoria:

$$X_u = \frac{|\{j \in \{1, \dots, t\} \mid \exists z \in \Sigma^* : y_j = uz\}|}{t}$$

X_u es un promedio de t variables aleatorias que toman valor 0 ó 1, las cuales denotamos como $X_{j,u}$ para $j \in \{1, \dots, t\}$

- ▶ $X_{j,u}(y_j) = 1$ si $y_j = uz$ para algún $z \in \Sigma^*$, y $X_{j,u}(y_j) = 0$ en otro caso

Vale decir, tenemos que $X_u = \frac{1}{t} \sum_{j=1}^t X_{j,u}$

Acotando inferiormente $\Pr$ (propiedad de aproximación)

Tenemos que:

$$\mathbf{Var}[X_u] = \mathbf{Var}\left[\frac{1}{t} \sum_{j=1}^t X_{j,u}\right] = \frac{1}{t^2} \sum_{j=1}^t \mathbf{Var}[X_{j,u}] \leq \frac{1}{t^2} \sum_{j=1}^t 1 = \frac{1}{t}$$

Dado que $t = \lceil 180 \cdot (|x|^c + d)^3 \cdot m^3 \cdot \varepsilon^{-2} \rceil$, por la desigualdad de Chebyshev concluimos que:

$$\begin{aligned} \Pr\left(|X_u - \mathbf{E}[X_u]| \geq \frac{\varepsilon}{6 \cdot (|x|^c + d) \cdot m}\right) &\leq \frac{36 \cdot (|x|^c + d)^2 \cdot m^2 \cdot \mathbf{Var}[X_u]}{\varepsilon^2} \\ &\leq \frac{36 \cdot (|x|^c + d)^2 \cdot m^2}{\varepsilon^2 \cdot t} \\ &\leq \frac{36 \cdot (|x|^c + d)^2 \cdot m^2}{\varepsilon^2 \cdot 180 \cdot (|x|^c + d)^3 \cdot m^3 \cdot \varepsilon^{-2}} \\ &= \frac{1}{5 \cdot (|x|^c + d) \cdot m} \end{aligned}$$

Definiendo α_i en términos de las variables aleatorias X_u

El valor de cada variable aleatoria X_u es una función de las variables $y_1, \dots, y_t$

- Podemos entonces hablar de $X_u(y_1, \dots, y_t)$

De la misma forma, el valor de la variable aleatoria α_i es una función de $y_1, \dots, y_t$, y podemos hablar de $\alpha_i(y_1, \dots, y_t)$

Definiendo α_i en términos de las variables aleatorias X_u

Suponga que v es el valor de w_i . Si los valores de las variables $y_1, \dots, y_t$ en la iteración i son $a_1, \dots, a_t$, respectivamente, entonces tenemos que:

$$\alpha_i(a_1, \dots, a_t) = X_v(a_1, \dots, a_t)$$

Además, en general tenemos que:

$$\alpha_i(y_1, \dots, y_t) = \max_{u \in \Sigma^{\sigma(x_i)}} X_u(y_1, \dots, y_t)$$

Es importante notar que **no** podemos concluir que:

$$\alpha_i(y_1, \dots, y_t) = X_v(y_1, \dots, y_t),$$

dado que v es un string **fijo** calculado en la iteración i

No podemos concluir que $\alpha_i(y_1, \dots, y_t) = X_v(y_1, \dots, y_t)$

Suponga que ejecutamos el algoritmo nuevamente obteniendo v' como el valor de w_i

- ▶ Además, suponga que los valores de las variables $y_1, \dots, y_t$ en la iteración i de esta nueva ejecución son $b_1, \dots, b_t$, respectivamente.

Sabemos que $\alpha_i(b_1, \dots, b_t) = X_{v'}(b_1, \dots, b_t)$

Pero si $v \neq v'$ y $X_v(b_1, \dots, b_t) < X_{v'}(b_1, \dots, b_t)$, entonces tenemos que:

$$\alpha_i(b_1, \dots, b_t) \neq X_v(b_1, \dots, b_t)$$

Acotando superiormente $\Pr(|\alpha_i - \mathbf{E}[X_{w_i}]| \geq \delta)$

Queremos entender cuán cerca está α_i de $\mathbf{E}[X_{w_i}]$

- Vale decir, queremos acotar superiormente $\Pr(|\alpha_i - \mathbf{E}[X_{w_i}]| \geq \delta)$

Dado que $\alpha_i = \max_{u \in \Sigma^{\sigma(x_i)}} X_u$, no podemos utilizar la desigualdad de Chebyshev para acotar superiormente $\Pr(|\alpha_i - \mathbf{E}[X_{w_i}]| \geq \delta)$

- ¿Por qué?

Acotando superiormente $\Pr(|\alpha_i - \mathbf{E}[X_{w_i}]| \geq \delta)$

Si $|\alpha_i - \mathbf{E}[X_{w_i}]| \geq \delta$ en la iteración i , entonces se debe tener que $|X_{w_i} - \mathbf{E}[X_{w_i}]| \geq \delta$

- Y si esperamos que esto se cumpla en la iteración i de cada ejecución de **EAR**, entonces en cada ejecución debe ser posible encontrar $u \in \Sigma^{\sigma(x_i)}$ tal que $|X_u - \mathbf{E}[X_u]| \geq \delta$

Por lo tanto, tenemos que:

$$\Pr(|\alpha_i - \mathbf{E}[X_{w_i}]| \geq \delta) \leq \Pr\left(\bigvee_{u \in \Sigma^{\sigma(x_i)}} |X_u - \mathbf{E}[X_u]| \geq \delta\right)$$

Acotando superiormente $\Pr(|\alpha_i - \mathbf{E}[X_{w_i}]| \geq \frac{\varepsilon}{6 \cdot (|x|^c + d) \cdot m})$

Utilizando la conclusión de la transparencia anterior:

$$\begin{aligned}
 \Pr\left(|\alpha_i - \mathbf{E}[X_{w_i}]| \geq \frac{\varepsilon}{6 \cdot (|x|^c + d) \cdot m}\right) &\leq \\
 \Pr\left(\bigvee_{u \in \Sigma^{\sigma(x_i)}} \left[|X_u - \mathbf{E}[X_u]| \geq \frac{\varepsilon}{6 \cdot (|x|^c + d) \cdot m}\right]\right) &\leq \\
 \sum_{u \in \Sigma^{\sigma(x_i)}} \Pr\left(|X_u - \mathbf{E}[X_u]| \geq \frac{\varepsilon}{6 \cdot (|x|^c + d) \cdot m}\right) &\leq \\
 \sum_{u \in \Sigma^{\sigma(x_i)}} \frac{1}{5 \cdot (|x|^c + d) \cdot m} &= \\
 \frac{1}{5 \cdot (|x|^c + d) \cdot m} \cdot |\Sigma|^{\sigma(x_i)} &\leq \\
 \frac{1}{5 \cdot (|x|^c + d) \cdot m} \cdot (|x|^c + d) &= \frac{1}{5m} \cdot \left(\frac{|x_i|^c + d}{|x|^c + d}\right) \leq \frac{1}{5m}
 \end{aligned}$$

¿Por qué **EAR** elige el prefijo de largo $\sigma(x)$ más común en $\{y_1, \dots, y_t\}$?

Como w_i se elige como el prefijo de largo $\sigma(x_i)$ más común en $\{y_1, \dots, y_t\}$, sabemos que:

$$\begin{aligned}\alpha_i &= \frac{|\{j \in \{1, \dots, t\} \mid y_j \in \text{Ext}_R(x_i, w_i)\}|}{t} \geq \frac{t}{|\Sigma|^{\sigma(x_i)}} \cdot \frac{1}{t} = \\ &\frac{1}{|\Sigma|^{\sigma(x_i)}} \geq \frac{1}{|x|^c + d} \geq \frac{1}{|x|^c + d}\end{aligned}$$

Concluimos entonces que:

$$\begin{aligned}\Pr\left(|\alpha_i - \mathbf{E}[X_{w_i}]| \leq \alpha_i \cdot \frac{\varepsilon}{6m}\right) &\geq \Pr\left(|\alpha_i - \mathbf{E}[X_{w_i}]| \leq \frac{\varepsilon}{6 \cdot (|x|^c + d) \cdot m}\right) \\ &\geq 1 - \frac{1}{5m}\end{aligned}$$

Dos desigualdades útiles

Tenemos que $1 + \frac{\varepsilon}{5m} \geq 1 + \frac{\varepsilon}{6m}$

Por otra parte, para $\varepsilon \in (0, 1)$ también se tiene que:

$$\frac{1}{1 + \frac{\varepsilon}{5m}} \leq 1 - \frac{\varepsilon}{6m},$$

puesto que $m \geq 1$ y:

$$\begin{aligned} \frac{1}{1 + \frac{\varepsilon}{5m}} \leq 1 - \frac{\varepsilon}{6m} &\Leftrightarrow \frac{5m}{5m + \varepsilon} \leq \frac{6m - \varepsilon}{6m} \\ &\Leftrightarrow 30 \cdot m^2 \leq 30 \cdot m^2 + 6 \cdot m \cdot \varepsilon - 5 \cdot m \cdot \varepsilon - \varepsilon^2 \\ &\Leftrightarrow \varepsilon^2 \leq m \cdot \varepsilon \\ &\Leftrightarrow \varepsilon \leq m \end{aligned}$$

Una cota inferior para $\Pr(|\alpha_i - \mathbf{E}[X_{w_i}]| \leq \alpha_i \cdot \frac{\varepsilon}{6m})$

Usando todas las desigualdades anteriores concluimos que:

$$\begin{aligned} \Pr\left(\frac{1}{\left(1 + \frac{\varepsilon}{5m}\right)} \cdot \alpha_i \leq \mathbf{E}[X_{w_i}] \leq \left(1 + \frac{\varepsilon}{5m}\right) \cdot \alpha_i\right) &\geq \\ \Pr\left(\left(1 - \frac{\varepsilon}{6m}\right) \cdot \alpha_i \leq \mathbf{E}[X_{w_i}] \leq \left(1 + \frac{\varepsilon}{6m}\right) \cdot \alpha_i\right) &= \\ \Pr\left(|\alpha_i - \mathbf{E}[X_{w_i}]| \leq \alpha_i \cdot \frac{\varepsilon}{6m}\right) &\geq \\ 1 - \frac{1}{5m} \end{aligned}$$

Acotando $\mathbf{E}[X_{w_i}]$

Por definición de X_{w_i} , se tiene que:

$$\begin{aligned}\mathbf{E}[X_{w_i}] &= \mathbf{E}\left[\frac{1}{t} \sum_{j=1}^t X_{j,w_i}\right] \\&= \frac{1}{t} \sum_{j=1}^t \mathbf{E}[X_{j,w_i}] \\&= \frac{1}{t} \sum_{j=1}^t \left(1 \cdot \mathbf{Pr}(X_{j,w_i} = 1) + 0 \cdot \mathbf{Pr}(X_{j,w_i} = 0)\right) \\&= \frac{1}{t} \sum_{j=1}^t \sum_{y \in \text{Ext}_R(x_j, w_i)} \mathbf{Pr}\left(\mathcal{G}\left(x_j, \frac{\varepsilon}{5m}\right) = y\right) \\&= \sum_{y \in \text{Ext}_R(x_i, w_i)} \mathbf{Pr}\left(\mathcal{G}\left(x_i, \frac{\varepsilon}{5m}\right) = y\right)\end{aligned}$$

Acotando $\mathbf{E}[X_{w_i}]$

Como $\mathcal{G}$ es un FPAUG para R , tenemos que:

$$\left(1 - \frac{\varepsilon}{5m}\right) \cdot \frac{1}{N_R(x_i)} \leq \Pr\left(\mathcal{G}\left(x_i, \frac{\varepsilon}{5m}\right) = y\right) \leq \left(1 + \frac{\varepsilon}{5m}\right) \cdot \frac{1}{N_R(x_i)}$$

Por lo tanto:

$$\sum_{y \in \text{Ext}_R(x_i, w_i)} \left(1 - \frac{\varepsilon}{5m}\right) \cdot \frac{1}{N_R(x_i)} \leq \mathbf{E}[X_{w_i}] \leq \sum_{y \in \text{Ext}_R(x_i, w_i)} \left(1 + \frac{\varepsilon}{5m}\right) \cdot \frac{1}{N_R(x_i)}$$

De lo cual concluimos que:

$$\left(1 - \frac{\varepsilon}{5m}\right) \cdot \frac{|\text{Ext}_R(x_i, w_i)|}{N_R(x_i)} \leq \mathbf{E}[X_{w_i}] \leq \left(1 + \frac{\varepsilon}{5m}\right) \cdot \frac{|\text{Ext}_R(x_i, w_i)|}{N_R(x_i)}$$

Acotando $\mathbf{E}[X_{w_i}]$

Dado que:

$$\begin{aligned} \left(1 + \frac{\varepsilon}{5m}\right) &\leq \left(1 + \frac{\varepsilon}{4m}\right) \\ \frac{1}{1 + \frac{\varepsilon}{4m}} &\leq \left(1 - \frac{\varepsilon}{5m}\right) \end{aligned}$$

Concluimos que:

$$\frac{1}{1 + \frac{\varepsilon}{4m}} \cdot \frac{|\text{Ext}_R(x_i, w_i)|}{N_R(x_i)} \leq \mathbf{E}[X_{w_i}] \leq \left(1 + \frac{\varepsilon}{4m}\right) \cdot \frac{|\text{Ext}_R(x_i, w_i)|}{N_R(x_i)}$$

Otra desigualdad útil

Para $\varepsilon \in (0, 1)$, se tiene que:

$$1 + \frac{\varepsilon}{2m} \geq \left(1 + \frac{\varepsilon}{5m}\right) \cdot \left(1 + \frac{\varepsilon}{4m}\right)$$

Puesto que $\varepsilon \in (0, 1)$ y:

$$\begin{aligned} 1 + \frac{\varepsilon}{2m} \geq \left(1 + \frac{\varepsilon}{5m}\right) \cdot \left(1 + \frac{\varepsilon}{4m}\right) &\Leftrightarrow 1 + \frac{\varepsilon}{2m} \geq 1 + \frac{\varepsilon}{4m} + \frac{\varepsilon}{5m} + \frac{\varepsilon^2}{20m} \\ &\Leftrightarrow \frac{\varepsilon}{2m} \geq \frac{\varepsilon}{4m} + \frac{\varepsilon}{5m} + \frac{\varepsilon^2}{20m} \\ &\Leftrightarrow 10\varepsilon \geq 5\varepsilon + 4\varepsilon + \varepsilon^2 \\ &\Leftrightarrow \varepsilon \geq \varepsilon^2 \\ &\Leftrightarrow 1 \geq \varepsilon \end{aligned}$$

$$\text{Acotando inferiormente } \Pr\left(\frac{\alpha_i}{(1+\frac{\varepsilon}{2m})} \leq \frac{|\text{Ext}_R(x_i, w_i)|}{N_R(x_i)} \leq (1 + \frac{\varepsilon}{2m}) \cdot \alpha_i\right)$$

Usando las desigualdades anteriores, tenemos que:

$$\begin{aligned} \frac{1}{(1 + \frac{\varepsilon}{5m})} \cdot \alpha_i \leq \mathbf{E}[X_{w_i}] &\Rightarrow \frac{1}{(1 + \frac{\varepsilon}{5m})} \cdot \alpha_i \leq \left(1 + \frac{\varepsilon}{4m}\right) \cdot \frac{|\text{Ext}_R(x_i, w_i)|}{N_R(x_i)} \\ &\Rightarrow \frac{1}{(1 + \frac{\varepsilon}{5m}) \cdot (1 + \frac{\varepsilon}{4m})} \cdot \alpha_i \leq \frac{|\text{Ext}_R(x_i, w_i)|}{N_R(x_i)} \\ &\Rightarrow \frac{1}{(1 + \frac{\varepsilon}{2m})} \cdot \alpha_i \leq \frac{|\text{Ext}_R(x_i, w_i)|}{N_R(x_i)} \end{aligned}$$

De la misma forma obtenemos:

$$\mathbf{E}[X_{w_i}] \leq \left(1 + \frac{\varepsilon}{5m}\right) \cdot \alpha_i \Rightarrow \frac{|\text{Ext}_R(x_i, w_i)|}{N_R(x_i)} \leq \left(1 + \frac{\varepsilon}{2m}\right) \cdot \alpha_i$$

$$\text{Acotando inferiormente } \Pr\left(\frac{\alpha_i}{(1+\frac{\varepsilon}{2m})} \leq \frac{|\text{Ext}_R(x_i, w_i)|}{N_R(x_i)} \leq \left(1 + \frac{\varepsilon}{2m}\right) \cdot \alpha_i\right)$$

Juntando todo lo anterior, finalmente concluimos que:

$$\Pr\left(\frac{\alpha_i}{(1+\frac{\varepsilon}{2m})} \leq \frac{|\text{Ext}_R(x_i, w_i)|}{N_R(x_i)} \leq \left(1 + \frac{\varepsilon}{2m}\right) \cdot \alpha_i\right) \geq$$

$$\Pr\left(\frac{\alpha_i}{(1+\frac{\varepsilon}{5m})} \leq \mathbf{E}[X_{w_i}] \leq \left(1 + \frac{\varepsilon}{5m}\right) \cdot \alpha_i\right) \geq 1 - \frac{1}{5m}$$

Acotando inferiormente $\Pr(\text{propiedad de aproximación})$: el paso final

Recuerde que nuestro objetivo es demostrar que:

$$\Pr(\text{propiedad de aproximación}) \geq \frac{3}{4}$$

Tenemos que:

$$\Pr(\text{propiedad de aproximación}) =$$

$$\begin{aligned} & \Pr\left(\bigwedge_{i=1}^s \left[\frac{\alpha_i}{\left(1 + \frac{\varepsilon}{2m}\right)} \leq \frac{|\text{Ext}_R(x_i, w_i)|}{N_R(x_i)} \leq \left(1 + \frac{\varepsilon}{2m}\right) \cdot \alpha_i \right]\right) = \\ & \prod_{i=1}^s \Pr\left(\frac{\alpha_i}{\left(1 + \frac{\varepsilon}{2m}\right)} \leq \frac{|\text{Ext}_R(x_i, w_i)|}{N_R(x_i)} \leq \left(1 + \frac{\varepsilon}{2m}\right) \cdot \alpha_i\right) \geq \\ & \prod_{i=1}^s \left(1 - \frac{1}{5m}\right) = \left(1 - \frac{1}{5m}\right)^s \geq \left(1 - \frac{1}{5m}\right)^m \end{aligned}$$

Acotando inferiormente $\Pr$ (propiedad de aproximación): el paso final

Dado que $m \geq 1$, tenemos que $\left(1 - \frac{1}{5m}\right)^m \geq \frac{4}{5} > \frac{3}{4}$

Concluimos finalmente que:

$$\Pr(\text{propiedad de aproximación}) \geq \left(1 - \frac{1}{5m}\right)^m \geq \frac{3}{4}$$

Esto termina la demostración del teorema. □