

# Una segunda propiedad útil

## Lema

Sea  $X \subseteq \{0, 1\}^m$ , y suponga que se escoge con distribución uniforme y de manera independiente  $n + 1$  funciones de hash aleatorias

$$h_1, \dots, h_{n+1} : \{0, 1\}^m \rightarrow \{0, 1\}^n$$

1. Si  $|X| \leq 2^{n-1}$ , entonces:

$$\Pr(\exists x \in X \forall k \in \{1, \dots, n+1\} \exists y \in X : (y \neq x \wedge h_k(x) = h_k(y))) \leq \frac{1}{4}$$

2. Si  $|X| > (n+1)2^n$ , entonces:

$$\Pr(\exists x \in X \forall k \in \{1, \dots, n+1\} \exists y \in X : (y \neq x \wedge h_k(x) = h_k(y))) = 1$$

# Demostración de la parte 1 del lema

Por el lema anterior tenemos que para cada  $x, y \in X$  tal que  $x \neq y$ ,  $k \in \{1, \dots, n+1\}$  y  $j \in \{1, \dots, n\}$ :

$$\Pr_{h_k}(\pi_j(h_k(x)) = \pi_j(h_k(y))) = \frac{1}{2}$$

Dado que los elementos de la matriz Booleana que define a  $h_k$  son escogidos de manera independientes, concluimos que:

$$\begin{aligned} \Pr_{h_k}(h_k(x) = h_k(y)) &= \prod_{j=1}^n \Pr_{h_k}(\pi_j(h_k(x)) = \pi_j(h_k(y))) \\ &= 2^{-n} \end{aligned}$$

# Demostración de la parte 1 del lema

Así, dado  $x \in X$ :

$$\begin{aligned} \Pr_{h_k}(\exists y \in X : (y \neq x \wedge h_k(x) = h_k(y))) &\leq \sum_{y \in (X \setminus \{x\})} \Pr_{h_k}(h_k(x) = h_k(y)) \\ &\leq |X| \cdot 2^{-n} \\ &\leq 2^{n-1} \cdot 2^{-n} \\ &= \frac{1}{2} \end{aligned}$$

# Demostración de la parte 1 del lema

Dado que las funciones  $h_1, \dots, h_{n+1}$  son escogidas de manera independiente, concluimos que:

$$\begin{aligned} & \Pr(\forall k \in \{1, \dots, n+1\} \exists y \in X : (y \neq x \wedge h_k(x) = h_k(y))) \\ &= \prod_{k=1}^{n+1} \Pr_{h_k}(\exists y \in X : (y \neq x \wedge h_k(x) = h_k(y))) \\ &\leq \prod_{k=1}^{n+1} \frac{1}{2} \\ &= 2^{-n-1} \end{aligned}$$

# Demostración de la parte 1 del lema

Así, finalmente tenemos que:

$$\begin{aligned} & \Pr(\exists x \in X \forall k \in \{1, \dots, n+1\} \exists y \in X : (y \neq x \wedge h_k(x) = h_k(y))) \\ & \leq \sum_{x \in X} \Pr(\forall k \in \{1, \dots, n+1\} \exists y \in X : (y \neq x \wedge h_k(x) = h_k(y))) \\ & \leq \sum_{x \in X} 2^{-n-1} \\ & = |X| \cdot 2^{-n-1} \\ & \leq 2^{n-1} \cdot 2^{-n-1} \\ & = \frac{1}{4} \end{aligned}$$

# Demostración de la parte 2 del lema

Definimos una secuencia de conjuntos  $X_1, \dots, X_{n+1}$  tales  $X_i \subseteq X$  para cada  $i \in \{1, \dots, n+1\}$

Primero,  $X_1$  es definido como:

$$X_1 = \{x \in X \mid \exists y \in X : (y \neq x \wedge h_1(x) = h_1(y))\}$$

Dado  $i \in \{1, \dots, n\}$ , definimos  $X_{i+1}$  como:

$$X_{i+1} = \{x \in X_i \mid \exists y \in X : (y \neq x \wedge h_{i+1}(x) = h_{i+1}(y))\}$$

## Demostración de la parte 2 del lema

Si  $x \in X_{n+1}$ , entonces tenemos que la siguiente condición se satisface:

$$\forall k \in \{1, \dots, n+1\} \exists y \in X : (y \neq x \wedge h_k(x) = h_k(y))$$

Por lo tanto, si demostramos que  $|X_{n+1}| > 0$  concluimos que:

$$\Pr(\exists x \in X \forall k \in \{1, \dots, n+1\} \exists y \in X : (y \neq x \wedge h_k(x) = h_k(y))) = 1$$

# Demostración de la parte 2 del lema

Dado que el conjunto  $\{0, 1\}^n$  tiene  $2^n$  strings y  $h_1 : \{0, 1\}^m \rightarrow \{0, 1\}^n$ , si  $|X| > 2^n$  entonces existen elementos  $x, y \in X$  tales que:

$$x \neq y \quad y \quad h_1(x) = h_1(y)$$

Usando este argumento es posible concluir que  $|X_1| \geq |X| - 2^n$

► ¿Por qué?

# Demostración de la parte 2 del lema

Para cada  $i \in \{1, \dots, n\}$  podemos aplicar el mismo argumento obteniendo:

$$\begin{aligned} |X_{i+1}| &\geq |X_i| - 2^n \\ &\geq |X| - (i+1)2^i \end{aligned}$$

Concluimos que:

$$\begin{aligned} |X_{n+1}| &\geq |X| - (n+1)2^n \\ &> (n+1)2^n - (n+1)2^n \\ &= 0 \end{aligned}$$

Esto era lo que teníamos que demostrar.



# Tenemos los ingredientes necesarios ...

## Teorema

$$\overline{GRAPH-ISO} \in AM$$

## Corolario

$$GRAPH-ISO \in co-AM$$

# Tenemos los ingredientes necesarios ...

El teorema anterior se demuestra estableciendo el siguiente resultado:

## Teorema (Schöning)

*Existe una MT probabilística no determinista  $M$  tal que  $t_M(n)$  es  $O(n^k)$  y para cada par  $(G_1, G_2)$  de grafos:*

- ▶ *Si  $G_1$  no es isomorfo a  $G_2$ , entonces  $Pr_s(M(G_1, G_2, s) \text{ acepta}) = 1$*
- ▶ *Si  $G_1$  es isomorfo a  $G_2$ , entonces  $Pr_s(M(G_1, G_2, s) \text{ acepta}) \leq \frac{1}{4}$*

# Demostración del teorema: notación inicial

Antes de definir la MT probabilística no determinista  $M$ , vamos a discutir algunas nociones y herramientas útiles para la demostración.

Suponga que  $G_1$  y  $G_2$  son dos grafos con  $m > 0$  nodos cada uno.

- Recuerde que  $\text{num}(G_1, G_2)$  fue definido como:

$$\{(H, i, f) \mid H \text{ es un grafo isomorfo a } G_1 \text{ o } G_2, i \in \{1, 2\} \text{ y } f \in \text{Aut}(G_i)\}$$

Para utilizar las herramientas desarrolladas primero tenemos que representar cada  $(H, i, f) \in \text{num}(G_1, G_2)$  como un string en  $\{0, 1\}^\ell$

- ¿Cuál es el valor de  $\ell$ ?

# Demostración del teorema: notación inicial

¿Cuántos bits necesitamos para representar una tupla  $(H, i, f) \in \text{num}(G_1, G_2)$ ?

- ▶ Podemos representar  $H$  usando su matriz de adyacencia, para lo cual necesitamos  $m^2$  bits
- ▶ Para almacenar el valor de  $i$  necesitamos un bit
- ▶ Podemos almacenar la biyección  $f$  como una lista de  $m$  números  $a_1 \dots a_m$  tal que  $f(i) = a_i$ 
  - ▶ Dado que cada  $a_i \leq m$ , basta con utilizar  $1 + \lfloor \log_2(m) \rfloor$  bits para almacenar  $a_i$
  - ▶ Por lo tanto necesitamos  $m(1 + \lfloor \log_2(m) \rfloor)$  bits para almacenar la lista  $a_1 \dots a_m$

Suponemos entonces que  $\ell = m^2 + 1 + m(1 + \lfloor \log_2(m) \rfloor)$

# Demostración del teorema: notación inicial

Desde ahora en adelante consideramos a cada elemento de  $\text{num}(G_1, G_2)$  como un string de  $\ell$  bits

- ▶ Tenemos que  $\text{num}(G_1, G_2) \subseteq \{0, 1\}^\ell$

Defina  $X(G_1, G_2)$  como  $\text{num}(G_1, G_2)^m$

- ▶ Cada elemento de  $\text{num}(G_1, G_2)^m$  es de la forma  $w_1 w_2 \cdots w_m$ , donde para cada  $i \in \{1, \dots, m\}$  se tiene que  $w_i$  es un string en  $\text{num}(G_1, G_2)$

Tenemos que:

- ▶  $X(G_1, G_2) \subseteq \{0, 1\}^{\ell \cdot m}$
- ▶ Si  $G_1$  no es isomorfo a  $G_2$ , entonces  $|X(G_1, G_2)| \geq (4 \cdot m!)^m$
- ▶ Si  $G_1$  es isomorfo a  $G_2$ , entonces  $|X(G_1, G_2)| = (2 \cdot m!)^m$

# Demostración del teorema: notación inicial

Finalmente defina  $n = 1 + \lceil m \cdot \log_2(2 \cdot m!) \rceil$

Tenemos que:

$$\begin{aligned} 1 + \lceil m \cdot \log_2(2 \cdot m!) \rceil &= 1 + \lceil m \cdot (1 + \log_2(m!)) \rceil \\ &\leq 1 + \lceil m \cdot (1 + \log_2(m^m)) \rceil \\ &= 1 + \lceil m \cdot (1 + m \log_2(m)) \rceil \\ &\leq 1 + \lceil m \cdot (1 + m^2) \rceil \\ &= 1 + m + m^3 \end{aligned}$$

Concluimos que  $n + 1 = 2 + \lceil m \cdot \log_2(2 \cdot m!) \rceil < 2^{m-2}$  para todo  $m \geq 14$

► Vamos a utilizar esta propiedad en la siguiente transparencia

# Demostración del teorema: notación inicial

Suponga que  $G_1$  no es isomorfo a  $G_2$  y que  $m \geq 14$

► Tenemos que  $2^{m-2} > (n+1)$

Concluimos que  $|X(G_1, G_2)| > (n+1)2^n$  puesto que:

$$\begin{aligned} |X(G_1, G_2)| &\geq (4 \cdot m!)^m \\ &= 2^{\log_2((4 \cdot m!)^m)} \\ &= 2^{m \cdot \log_2(4 \cdot m!)} \\ &= 2^{m + m \cdot \log_2(2 \cdot m!)} \\ &= 2^{m-1 + (1 + m \cdot \log_2(2 \cdot m!))} \\ &\geq 2^{m-1 + \lceil m \cdot \log_2(2 \cdot m!) \rceil} \\ &= 2^{m-2+n} \\ &= 2^{m-2} \cdot 2^n \\ &> (n+1)2^n \end{aligned}$$

# Demostración del teorema: notación inicial

Si  $G_1$  es isomorfo a  $G_2$  tenemos que  $|X(G_1, G_2)| \leq 2^{n-1}$  puesto que:

$$\begin{aligned} |X(G_1, G_2)| &= (2 \cdot m!)^m \\ &= 2^{\log_2((2 \cdot m!)^m)} \\ &= 2^{m \cdot \log_2(2 \cdot m!)} \\ &\leq 2^{\lceil m \cdot \log_2(2 \cdot m!) \rceil} \\ &= 2^{1 + \lceil m \cdot \log_2(2 \cdot m!) \rceil - 1} \\ &= 2^{n-1} \end{aligned}$$

# Demostración del teorema: notación inicial

En la demostración vamos a considerar funciones de hash aleatorias

$$h : \{0, 1\}^{\ell \cdot m} \rightarrow \{0, 1\}^n$$

Estas funciones están dadas por matrices Booleanas  $A$  de  $n \times (\ell \cdot m)$

- Los elementos de  $A$  son escogidos con distribución uniforme y de manera independiente

Necesitamos  $(\ell \cdot m \cdot n)$  bits para representar  $A$

# Demostración del teorema: notación inicial

Necesitamos entonces  $(\ell \cdot m \cdot n)$  bits para representar una función de hash aleatoria  $h : \{0, 1\}^{\ell \cdot m} \rightarrow \{0, 1\}^n$

Vale decir, necesitamos la siguiente cantidad de bits para representar  $h$ :

$$[m^2 + 1 + m(1 + \lfloor \log_2(m) \rfloor)] \cdot m \cdot [1 + \lceil m \cdot \log_2(2 \cdot m!) \rceil]$$

El valor  $(\ell \cdot m \cdot n)$  es polinomial en  $m$ , de lo cual concluimos que es polinomial en el tamaño de  $(G_1, G_2)$

# Demostración del teorema: la definición de $M$

Recuerde que la MT probabilística no determinista  $M$  está tratando de verificar si dos grafos  $G_1$  y  $G_2$  **no** son isomorfos.

$M$  recibe como entrada una tupla de la forma  $(G_1, G_2, h_1, \dots, h_{n+1})$

- ▶ Cada  $h_i$  es una función de hash aleatoria de  $\{0, 1\}^{\ell \cdot m}$  en  $\{0, 1\}^n$ 
  - ▶  $m$  es el número de nodos de  $G_1$ ,  $\ell$  y  $n$  son definidos como fue mostrado en las transparencias anteriores
- ▶ La tupla  $(h_1, \dots, h_{n+1})$  corresponde al string de bits aleatorios que recibe  $M$

# Demostración del teorema: la definición de $M$

Con entrada  $(G_1, G_2, h_1, \dots, h_{n+1})$  la MT  $M$  realiza los siguientes pasos:

1. Si  $G_1$  y  $G_2$  no tienen el mismo número de nodos entonces retorne **sí**, si no vaya al paso 2
2. Sea  $m$  el número de nodos de  $G_1$  y  $G_2$
3. Si  $m < 14$  entonces vaya al paso 3.1, si no vaya al paso 4
  - 3.1 Construya todas las posibles biyecciones  
 $f : \{1, \dots, m\} \rightarrow \{1, \dots, m\}$
  - 3.2 Si alguna de estas biyecciones  $f$  es un isomorfismo de  $G_1$  en  $G_2$  entonces retorne **no**. En caso contrario retorne **sí**
4. Adivine  $(H_1, g_1, i_1, f_1, \dots, H_m, g_m, i_m, f_m)$  tal que para cada  $j \in \{1, \dots, m\}$ :  
 $g_j$  es un isomorfismo de  $H_j$  en  $G_1$  o  $G_2$ , y  $f_j \in \text{Aut}(G_{i_j})$
5. Sea  $x = (H_1, i_1, f_1, \dots, H_m, i_m, f_m)$

$$x \in X(G_1, G_2)$$

# Demostración del teorema: la definición de $M$

6. Para  $k = 1$  hasta  $n + 1$  haga lo siguiente:

6.1 Adivine  $(H_1, g_1, i_1, f_1, \dots, H_m, g_m, i_m, f_m)$  tal que para cada  $j \in \{1, \dots, m\}$ :  $g_j$  es un isomorfismo de  $H_j$  en  $G_1$  o  $G_2$  y  $f_j \in \text{Aut}(G_{i_j})$

6.2 Sea  $y = (H_1, i_1, f_1, \dots, H_m, i_m, f_m)$

$y \in X(G_1, G_2)$

6.3 Si  $x = y$  o  $h_k(x) \neq h_k(y)$ , entonces retorne **no**

7. Retorne **sí**

# Demostración del teorema: la probabilidad de error de $M$

La MT no determinista  $M$  funciona en tiempo polinomial.

Además,  $M$  acepta una entrada  $(G_1, G_2, h_1, \dots, h_{n+1})$  si y sólo si alguna de las siguientes condiciones se cumple:

- ▶  $G_1$  y  $G_2$  no tienen el mismo número de nodos
- ▶  $G_1$  y  $G_2$  tienen  $m < 14$  nodos cada uno y no son isomorfos
- ▶  $G_1$  y  $G_2$  tienen  $m \geq 14$  nodos cada uno y

$$\exists x \in X(G_1, G_2) \forall k \in \{1, \dots, n+1\}$$

$$\exists y \in X(G_1, G_2) : (x \neq y \wedge h_k(x) = h_k(y))$$

# Demostración del teorema: la probabilidad de error de $M$

Usamos las condiciones de aceptación de  $M$  para establecer la probabilidad de error de esta máquina de Turing

Vale decir, dados dos grafos  $G_1$  y  $G_2$  queremos calcular la probabilidad:

$$\Pr_{h_1, \dots, h_{n+1}}(M(G_1, G_2, h_1, \dots, h_{n+1}) \text{ acepte})$$

dependiendo de si  $G_1$  y  $G_2$  son o no son isomorfos.

# Demostración del teorema: la probabilidad de error de $M$

Suponemos primero que  $G_1$  y  $G_2$  **no** son isomorfos.

Si  $G_1$  y  $G_2$  no tienen el mismo número de nodos, o si tienen el mismo número de nodos  $m < 14$ , entonces:

$$\Pr_{h_1, \dots, h_{n+1}}(M(G_1, G_2, h_1, \dots, h_{n+1}) \text{ acepte}) = 1$$

De hecho en estos casos las funciones  $h_1, \dots, h_{n+1}$  no son tomadas en cuenta.

# Demostración del teorema: la probabilidad de error de $M$

Si  $G_1$  y  $G_2$  tienen el mismo número de nodos  $m \geq 14$ , entonces tenemos que  $|X(G_1, G_2)| > (n+1)2^n$

Concluimos por el último lema demostrado que para cualquier secuencia  $h_1, \dots, h_{n+1} : \{0, 1\}^{\ell \cdot m} \rightarrow \{0, 1\}^n$  de funciones de hash aleatorias:

$$\begin{aligned} \exists x \in X(G_1, G_2) \forall k \in \{1, \dots, n+1\} \\ \exists y \in X(G_1, G_2) : (x \neq y \wedge h_k(x) = h_k(y)) \end{aligned}$$

Por lo tanto en este caso también tenemos que:

$$\Pr_{h_1, \dots, h_{n+1}}(M(G_1, G_2, h_1, \dots, h_{n+1}) \text{ acepta}) = 1$$

# Demostración del teorema: la probabilidad de error de $M$

Finalmente, consideramos el caso en que  $G_1$  y  $G_2$  son grafos isomorfos.

- Suponemos que  $G_1$  y  $G_2$  tienen el mismo número de nodos  $m$

Si  $m < 14$  entonces la MT  $M$  no se puede equivocar al decidir si  $G_1$  es isomorfo a  $G_2$ , y tenemos que:

$$\Pr_{h_1, \dots, h_{n+1}}(M(G_1, G_2, h_1, \dots, h_{n+1}) \text{ acepte}) = 0$$

Suponemos entonces que  $m \geq 14$

# Demostración del teorema: la probabilidad de error de $M$

Tenemos que  $|X(G_1, G_2)| \leq 2^{n-1}$

Concluimos por el último lema demostrado:

$$\Pr_{h_1, \dots, h_{n+1}}(\exists x \in X(G_1, G_2) \forall k \in \{1, \dots, n+1\} \\ \exists y \in X(G_1, G_2) : (x \neq y \wedge h_k(x) = h_k(y))) \leq \frac{1}{4}$$

Por lo tanto:

$$\Pr_{h_1, \dots, h_{n+1}}(M(G_1, G_2, h_1, \dots, h_{n+1}) \text{ acepte}) \leq \frac{1}{4}$$

