Dimensión y aprendizaje

Elvira Mayordomo, Vinodchandran N. Variyam

Febrero 2004

Workshop MOISES

¿Qué es?

- **Trabajo en fase de realización**
- **Resultados negativos de aprendizaje usando técnicas de medida**

Hoy

- **El modelo de aprendizaje de Littlestone**
- **Un resultado sobre el tamaño de lo que se puede aprender**
- **Consecuencia:**

  si existen generadores con “exponential hardness” no se puede aprender P/poly con muuuuuchas preguntas

El modelo de Littlestone

- Se trata de aprender $T \subseteq \{0,1\}^n$
- El algoritmo va recibiendo una serie de casos x_1, x_2, ... de $\{0,1\}^n$
- Para cada uno el algoritmo responde a si está en T
- Después recibe la respuesta correcta

El modelo de Littlestone

- **“Online mistake-bound model”**
- **Se trata de acotar**
  - **el número máximo de errores**
  - **el tiempo para responder al caso x_i en función de n y de i**

Dimensión

- **Vamos a ver el tamaño de estas clases de lenguajes:**

X_a={ T | $\forall n$ $T^{=n}$ **se puede aprender en tiempo polinómico con menos de** $a2^n$ **errores** }

Dimensión

Teorema

X_a $(a \leq 1/2)$ tiene dimensión $H(a)$ en E

$H(a) = -a \log a - (1-a) \log(1-a)$

$E = DTIME(2^{O(n)})$

Dimensión

- **Si $A \subseteq B$ entonces $\dim(A) \leq \dim(B)$**
- **$\dim(E|E)=1$**
- **Si $\dim(A|E)<1$ entonces $\mu(A|E)=0$**

Dimensión

X_a (a<1/2) tiene dimensión H(a) en E

H(a)= -a log a –(1-a) log(1-a) <1

- X_a medida 0 en E

¿Se puede aprender P/poly?

P/poly son los lenguajes que se pueden reconocer con circuitos de tamaño polinómico

Parece que no tiene medida 0 en E ...

¿Se puede aprender P/poly?

[ReganSC] Si existen generadores pseudoaleatorios con "exponential hardness" entonces P/poly no tiene medida 0 en E

Bajo la misma hipótesis no se puede aprender P/poly con menos de $(1-\varepsilon)2^{n-1}$ errores en el modelo de Littlestone

¿Mejora lo anterior?

Desde Valiant 84 se sabe que si existen "generadores pseudoaleatorios débiles" entonces P/poly no es PAC-aprendible

¿Qué es?

- generadores pseudoaleatorios débiles
- generadores pseudoaleatorios con exponential hardness
- aprendizaje PAC

Generadores

- G: $\{0,1\}^m \rightarrow \{0,1\}^{2m}$
- con exponential hardness: los circuitos de tamaño $2^{n\varepsilon}$ (para algún ε) no distinguen entre:
  - elegir x $\in\{0,1\}^m$ aleatoriamente, calcular G(x)
  - elegir y $\in\{0,1\}^{2m}$ aleatoriamente

Generadores

- **G: $\{0,1\}^m \rightarrow \{0,1\}^{2m}$**
- **débiles:**
  - **los circuitos de tamaño polinómico no distinguen entre G(x), y**
  - **G es calculable con circuitos de tamaño polinómico**

Aprendizaje PAC

Se trata de aprender $T \subseteq \{0,1\}^n$

El algoritmo alumno va recibiendo pares

$<x, ¿x \in T?>$ según una cierta distribución sobre $x \in \{0,1\}^n$

En tiempo polinómico el alumno sabe quién es T con un pequeño error

(piensa que es T' con $Pr(|T \Delta T'| < \varepsilon) > 1 - \delta$)

Generadores

- **G: $\{0,1\}^m \rightarrow \{0,1\}^{2m}$**
- **débiles:**
  - **los circuitos de tamaño polinómico no distinguen entre G(x), y**
  - **G es calculable con circuitos de tamaño polinómico**

Resultados anteriores

Desde Valiant 84 se sabe que si existen “generadores pseudoaleatorios débiles” entonces P/poly no es PAC-aprendible

Resultados anteriores

- **No está claro cómo pasar de aprender con el modelo de Littlestone con muchos errores a aprender con PAC**

Se puede pasar

- **de Equivalence queries a PAC**
- **de Equivalence queries a Littlestone**
- **¿de Littlestone $(1-\varepsilon)2^{n-1}$ a Equivalence queries?**

Resultados anteriores

- **El resultado de Valiant parece depender de que el PSRG se calcule en P/poly**

La idea es que si puedo aprender P/poly puedo aprender a distinguir G(x) de y

Resultados anteriores

- **No está claro cómo escalar el resultado de Valiant para conseguir número de preguntas $(1-\varepsilon)\ 2^{n-1}$**

No funciona pedir que el generador tenga una cierta hardness

Nuestro resultado

- **Si existen generadores pseudoaleatorios con exponential hardness entonces no se puede aprender P/poly con menos de**

  $(1-\varepsilon)\ 2^{n-1}$ errores en el modelo de Littlestone

Hipótesis razonable [Razborov Rudich]