Algebraic Combinatorics: HW9

Dallin/Evan

5/17/2024

> **Problem 1** (Testing for Database consistency)**.** Two friends **Dallin** and **Evan** separately maintain copies of a large database of information. Every week they must compare their databases for consistency. Since transmission is expensive, they would like to find if there is any inconsistency between their databases by sending very little amount of information. Let their full databases be represented as the following $\{0,1\}$-vectors:
>
> $$\text{Dallin: } (a_0, a_1, \cdots, a_{n-1})$$
> $$\text{Evan: } (b_0, b_1, \cdots, b_{n-1})$$
>
> Design a randomized strategy that detects an inconsistency with high probability $\left(\text{say} \geq 1 - \frac{1}{n}\right)$ while transmitting only $O(\log n)$ bits.

Solution. Define the degree-n polynomials A and B with terms $a_i x^i$ and $b_i x^i$, respectively. For random testing, we let S be a set of m random elements from $\mathbb{R}$ such that $m \geq n^2$. We proceed by the folowing algorithm:

1. Dallin and Evan agree upon an $r \in S$ and transmit this value to each other.
2. Dallin evaluates $A(r) \pmod{2^{\log n}}$ and transmits this value to Evan. Taking the output $\pmod{2^{\log n}}$ ensures that the output is no more than $\log n$ bits while still being able to detect inconsistencies.
3. Evan similarly evaluates $B(r) \pmod{2^{\log n}}$ and transmits this value to Dallin.
4. Evan and Dallin have effectively evaluated $P(x) = A(x) - B(x)$ at $x = r$.

If $P(r) \neq 0$, the databases are inconsistent. Otherwise, by the Schwartz-Zippel lemma, the databases are consistent with a probability at most $\frac{n}{|S|}$ and are thus inconsistent with a probability $\geq 1 - \frac{n}{m} \geq 1 - \frac{1}{n}$. ■