Lecture 14: Guest Lecture by Noga Alon: Nonconstructive Proofs in Combinatorics

Lecturer: *Sanjeev Arora* Scribe:*Don Sheehy*

We are interested in looking at nonconstructive existence proofs in combinatorics where no efficient algorithm is known for solving the corresponding search problem. This lecture presents three examples illustrating topological methods, algebraic methods, and probabilistic methods.

1 Topological Methods

For the first example we consider a problem of spliting up a necklace (a string of beads).

THEOREM 1
Let N be an opened necklace with ka_i beads of type i, where $1 \leq i \leq t$ and t is the number of types of beads. It is possible to cut N in at most $(k-1)t$ places and split the resulting intervals into k collections, each with exactly a_i beads of type i.

Figure 1 gives an example for $k = 2$ and $t = 3$. We see that $(k-1)t = (2-1)3 = 3$ cuts suffice. In fact, the Example in the figure can be done with just 2 cuts.

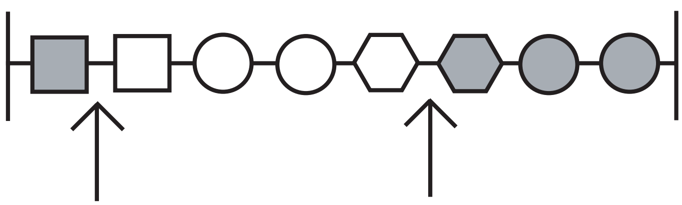

Figure 1:

There are situations where all $(k-1)t$ cuts from the theorem are necessary. For instance, if the beads are grouped by type on the string as in Figure 2 then each of the t bead types requires $k-1$ cuts. So, we see that the bound from the theorem is tight.

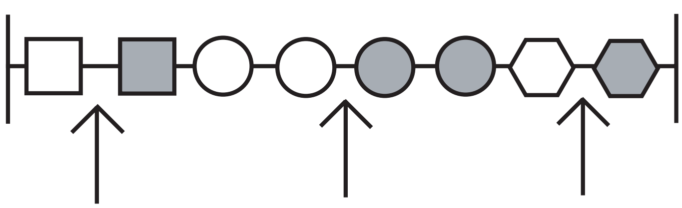

Figure 2:

The problem of finding these cuts is known to be NP-Hard even for the case when $k = 2$. We will use the Borsuk-Ulam Theorem to prove the Theorem for the $k = 2$ case. The version of the

Borsuk-Ulam Theorem we will use states that for all continuous, antipodal maps $f : S^t \to \mathbf{R}^t$, there is a point x in S^t such that $f(x) = 0$.

PROOF:(of Theorem for $k = 2$) If N has n beads, split the unit interval $[0, 1]$ into n equal segments $X_1, \ldots, X_n$ and "color" X_j with color i if and only if jth bead is of type i. Given $x = (x_1, \ldots, x_{t+1}) \in S^t$ partition the unit interval into subintervals $I_1, \ldots, I_{t+1}$ of length x_i^2. This is possible because for all points $x \in S^t, \sum_{i=1}^{t+1} x_i^2 = 1$. (For our purposes, we don't need to worry about the boundaries between intervals).

This defines t continuous measures $\mu_1, \ldots \mu_t$ on $[0, 1]$, where if $Y \subseteq [0, 1]$ then $\mu_i(Y)$ is the fraction of I_i that contains Y.

We now define the map $f : S^t \to \mathbf{R}^t$ by $f(x_1, \ldots, x_{t+1}) = (f_1, \ldots, f_t)$ where

$$f_i(x) = \sum_{i=1}^{t+1} SIGN(x_i)\mu_j(I_i) \qquad (1)$$

Note that f is continuous and that $f(x) = -f(-x)$. So, the Borsuk-Ulam Theorem implies that there exists some x such that $f(x) = 0$. We can use this point to define the two collections of intervals.

$$I_+ = \bigcup_{j:x_j>0} I_j$$
$$I_- = \bigcup_{j:x_j<0} I_j$$

We see that $\mu_j(I_+) = \mu_j(I_-)$ for all j as desired.

The final step is to clear up any fractional beads. This is always possible by shifting the cuts by a small amount.

□

The Algorithmic Question: Given a necklace N with t types of beads and $2a_i$ beads of type i, find (efficiently) t cuts as needed.

2 Algebraic Methods

THEOREM 2
If $G = (V, E)$ is a graph with maximum vertex degree 5 and average vertex degree > 4, then G contains a 3-regular subgraph.

The Algorithmic Question: Given such a graph G, find the 3-regular subgraph.

LEMMA 3
Let P be a multilinear function over a field F. So,

$$P(x_1, \ldots, x_n) = \sum_{U \subseteq \{1,\ldots,n\}} a_U \prod_{i \in U} x_i$$

If $P(x_1, \ldots, x_n) = 0$ for all $x \in \{0, 1\}^n$ then $P = 0$ (i.e. $a_U = 0$ for all U).

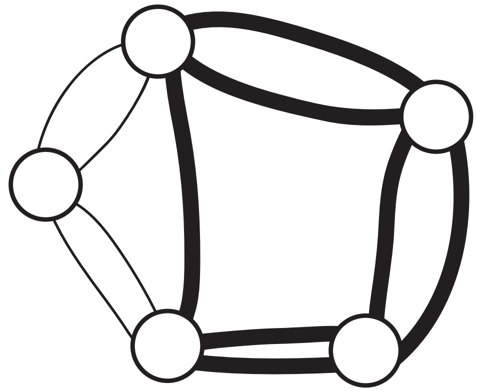

Figure 3: A graph satisfying the conditions of the Theorem. A 3-regular subgraph is indicated with bold lines.

Proof:(of Lemma) The proof will be by induction on n.

Base case: If $n = 1$ then $P(x) = ax + b$ for some integer coefficients a and b. We see that $P(0) = 0a + b = 0$ so it must be that $b = 0$. This implies that $P(x) = ax$, but $P(1) = 0$ so $a = 0$ as well.

Induction Step: Let $P_, P_2$ be polynomials so that

$$P(x_1, \ldots, x_n) = P_1(x_1, \ldots, x_{n-1})x_n + P_2(x_1, \ldots, x_{n-1}).$$

Take $x_n = 0$ and the induction hypothesis to get that $P_2 = 0$. Take $x_n = 1$ with the induction hypothesis to get that $P_1 = 0$. □

Proof:(of Theorem) We have $G(V, E)$ where the maximum vertex degree is 5. The average vertex degree is greater than 4 so $|E| \geq 2m + 1$. Suppose that G does not contain a 3-regular subgraph. For each edge $e \in E$, define a variable $x^{(e)}$. For each vertex/edge pair $(v \in V, e \in E)$, define a variable $a_v^{(e)}$ where

$$a_v^{(e)} = \begin{cases} 1 & v \in e \\ 0 & otherwise \end{cases} \tag{2}$$

Define the following polynomial over $GF(3)$

$$Q(x^{(e)}; e \in E) = \prod_{v \in V} \left[1 - \left(\sum_{e \in E} a_v^{(e)} x^{(e)} \right)^2 \right] \tag{3}$$

Intuitively, a choice of $x^{(e)} \in \{0, 1\}^{|E|}$ corresponds to selecting a subset of edges to make a subgraph. The sum squared on the RHS is 0 if $deg(v) \equiv 0 \pmod 3$, and it is 1 otherwise. So, $Q(x^{(e)}) \neq 0$ if and only if $x^{(e)}$ corresponds to a 3-regular subgraph. By assumption, no such 3-regular subgraph exists, so $Q(x^{(e)}) = 0$ for all $x^{(e)} \in \{0, 1\}^{|E|} \setminus \{\overline{0}\}$. In order to use Lemma 3, we want to tweak Q to make it multilinear. So, we define a new polynomial.

$$P(x^{(e)}; e \in E) = Q - \prod_{e \in E}(1 - x^{(e)}) \pmod{(x^{(e)^2} - x^{(e)})} \tag{4}$$

So now, P is multilinear and $P(x^{(e)}) = 0$ for all $x^{(e)} \in \{0,1\}^{|E|}$. So, $P \equiv 0$ by the Lemma. However, it can be checked that the coefficients of P are not all 0, a contradiction.

□

3 Probabilistic Methods

Theorem 4

If $G(V,E)$ is a digraph with maximum outdeg Δ and minimum indegree δ and if $e(\Delta\delta+1)(1-\frac{1}{k}) < 1$ then G has a directed simple cycle of length $0 \pmod{k}$. [In particular, if all in and outdegrees are between 20 and 30, there is a cycle of length $0 \pmod{3}$.

The Algorithmic Question: Given such a digraph and the integer k, find a directed simple cycle of length $0 \pmod{k}$.

Lemma 5

(Lovasz Local Lemma) Let $A_1, \ldots, A_n$ be events in an arbitrary probability space such that each A_i is mutually independent of all others but at most $d \geq 2$ and for all i, $\mathbf{Pr}(A_i) \leq p$. If $ep(d+1) < 1$ (here, e is the constant $2.71\ldots$) then

$$\mathbf{Pr}\left(\bigwedge_{i=1}^{n} \overline{A_i}\right) > 0. \tag{5}$$

Proof:(of Theorem) We may assume all indegrees are exactly δ. Let f be a random mapping $V \to \mathbf{Z}_k$ (uniform, indepent). For each vertex $v \in V$ define an event A_v as follows.

$$A_v = \{\nexists u : (u,v) \in E \text{ and } f(u) \equiv f(v) + 1 \pmod{k}\} \tag{6}$$

Intuitively, A_v is the event that none of the vertices u with edges directed into v have $f(u) \equiv f(v) + 1 \pmod{k}$. The probability of A_v is not hard to compute.

$$\mathbf{Pr}(A_v) = \left(1 - \frac{1}{k}\right)^{\delta} \tag{7}$$

Each A_v is mutually independent of all A_u but except those for which

$$\big(\{u\} \cup N^-(u)\big) \cap N^-(v) = \emptyset$$

Here, N^- denotes the predecessor set (i.e. $N^-(v) = \{u : (u,v) \in E\}$). The number of u such that A_u is not mutually independent of A_v is at most $\delta + \delta(\Delta - 1) = \delta\Delta$. So, by the Lovasz Local Lemma (Lemma 5) there exists some function f so that for all v there exists u such that $(u,v) \in E$ and $f(u) \equiv f(v) + 1 \pmod{k}$. This function f will give us the desired cycle.

take $v_0, v_1 \in V$ so that $(v_1, v_0) \in E$ and $f(v_0) \equiv f(v_1) + 1 \pmod{k}$.

Similarly define $v_0, v_1, \ldots$ so that $(v_{i+1}, v_i) \in E$ and $f(v_{i+1}) \equiv f(v_i) + 1 \pmod{k}$

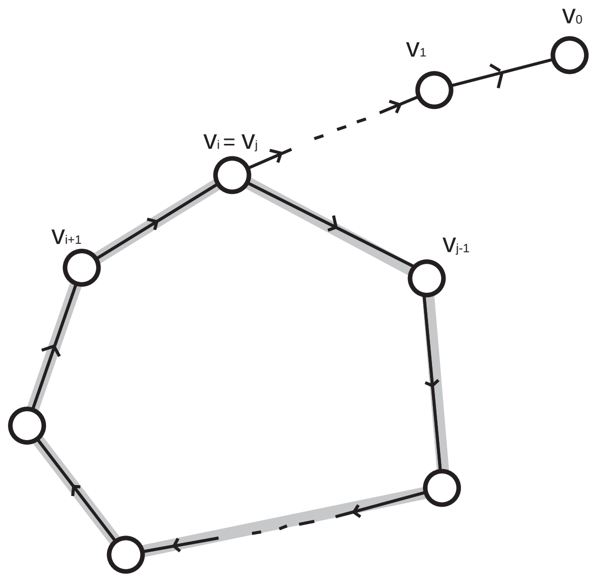

Figure 4:

At some point, we must get back to a vertex that we have already seen. Let j be the minimum so that there exists some $i < j$ such that $v_j = v_i$. Now we take the simple directed cycle $v_i v_{i+1} v_{i+2} \cdots v_j = v_i$, and its length is $f(v_i) - f(v_j) \equiv 0 \pmod{k}$ as desired.

$\square$

References

[1] N. Alon, J. H. Spencer, and Paul Erdös. *The Probabilistic Method*. John Wiley and Sons Inc, 1991.