Combinatorial Gems

Po-Shen Loh

June 2009

Although this lecture does not contain many offical Olympiad problems, the arguments which are used are all common elements of Olympiad problem solving. Some of these perhaps could have made nice Olympiad problems, if not for the fact that they were already well-known due to their importance in higher mathematics.

Most of this is sourced from the excellent combinatorics course developed by Benny Sudakov for Princeton.

1 Warm-up

1. (Erdős-Szekeres.) Prove that every sequence of n^2 distinct numbers contains a subsequence of length n which is monotone (i.e. either always increasing or always decreasing).

   Solution: For each of the n^2 indices in the sequence, associate the ordered pair (x, y) where x is the length of the longest increasing subsequence ending at x, and y is the length of the longest decreasing one. All ordered pairs must obviously be distinct. But if they only take values with $x, y \in \{1, \ldots, n-1\}$, then there are not enough for the total n^2 ordered pairs. Thus n appears somewhere, and we are done.

2. (Trivial approximation.) For $x \in \mathbb{R}$ and $n \in \mathbb{Z}^+$, there is a rational number $\frac{p}{q}$, with $1 \leq q \leq n$, such that
$$\left| x - \frac{p}{q} \right| < \frac{1}{n}.$$

   Solution: Pick $q = n$ and let p range over the integers.

3. (Dirichlet approximation.) Same as above, except with
$$\left| x - \frac{p}{q} \right| < \frac{1}{nq}.$$

   Solution: Consider the numbers $\{ax\}$, where this is the fractional part, and a runs from 0 to n. Consider the n buckets $[0, \frac{1}{n})$, $[\frac{1}{n}, \frac{2}{n})$, ... By pigeonhole, some bucket contains both $\{ax\}$ and $\{a'x\}$, say with $\{ax\} < \{a'x\}$. This means there is some integer p such that $0 < a'x - ax - p < \frac{1}{n}$. Let $q = a' - a$ and divide through by it, and we are done.

2 Arranging in order

One useful tool, both in Olympiads and in real mathematics, is to consider an extremal arrangement of the objects in the problem. Remarkably, a concept known as the *median order* has recently been discovered to be very useful in the study of directed graphs.

Try these problems.

1. Prove that every tournament (complete graph with all edges directed) has a Hamiltonian path, i.e., a directed path which visits every vertex exactly once.

   Solution: Take a median order.

2. A vertex v in a directed graph is a *king* if every other vertex can be reached from it via directed paths of length at most 2. That is, $V = \{v\} \cup N^+(v) \cup N^{++}(v)$. Show that every tournament has a king.

   Solution: (From Havet and Thomassé, 2000.) Take the first vertex v_1 in a median order, and consider v_k. If already $\overrightarrow{v_1 v_k}$, then done. Otherwise, by the feedback property, at least half of the edges from v_1 to $v_2 \dots v_k$ point forward. These give at least $\frac{k-1}{2}$ landing points in $v_2 \dots v_{k-1}$, because we assumed that v_1 does not go to v_k. Similarly, there are at least $\frac{k-1}{2}$ vertices in $v_2 \dots v_{k-1}$ which are origination points for edges directed to v_k. Yet there are only $k-2$ vertices in this window, so by pigeonhole some vertex is both an origination point and a landing point, giving a directed path of length 2.

3. (Seymour's Second Neighborhood conjecture for tournaments.) Show that every tournament contains a vertex v such that $|N^{++}(v)| \geq |N^+(v)|$. Here, $N^+(v)$ is the set of all vertices which are out-neighbors of v, and $N^{++}(v)$ is the set of all vertices which are out-neighbors of out-neighbors of v, but *not* out-neighbors of v or v itself.

   Solution: Prove the following slightly stronger statement by induction. Take a median order, and let v be the final vertex. Call a vertex u in $N^-(v)$ *good* if there is a vertex t preceding u in the median order, such that $\overrightarrow{vt}$ and $\overrightarrow{tu}$. We will show that $N^+(v)$ is at most the number of good vertices.

   Let b be the first bad vertex in the median order. So the direction of the edge to v is $\overleftarrow{bv}$. By induction, between b and v the number of $N^+(v)$ is at most the number of good vertices, so it suffices to show the same for the vertices before b. But every u with $\overleftarrow{uv}$ forces $\overleftarrow{ub}$ because b is bad. So the number of $N^+(v)$ before b is at most the number of $\overleftarrow{ub}$, which is at most half of those vertices by the feedback property applied to b.

   Yet b was the first bad vertex, so all non-$N^+(v)$ vertices before b are automatically good. This implies that the number of good vertices there is at least the number of $N^+(v)$ vertices, so we are done.

3 Brouwer's fixed point theorem

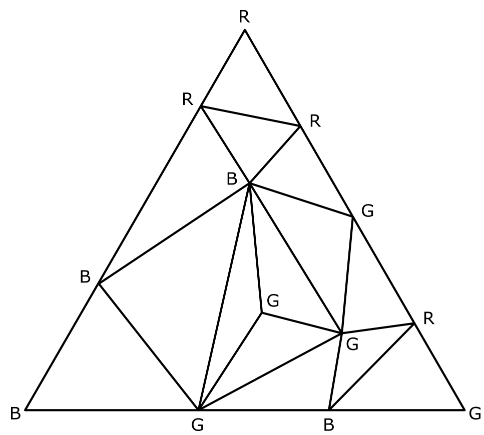
Any continuous map $f : [0,1] \to [0,1]$ has a fixed point $f(x) = x$. The map does **not** need to be injective or surjective. This is relatively easy to prove, because it is trivial unless $f(0) > 0$ and $f(1) < 1$. But then we are drawing a continuous curve between $(0, f(0))$ and $(1, f(1))$, which are on opposites sides of the line $f(x) = x$, so there must be a crossing.

The result for higher dimensions is less trivial. Surprisingly, the following purely combinatorial theorem is enough to imply it. We state it for 2 dimensions:

Theorem. *(Sperner's Lemma.) Consider any triangulation of the 2-dimensional simplex, i.e., a triangle. Suppose that all of the vertices have been colored from 3 colors, with the following properties.*

- *The 3 vertices of the large outermost triangle all receive distinct colors.*
- *Along any outer edge, the only colors which appear are the colors of the far endpoints of that outer edge.*

Then, somewhere in the triangulation, there is an elementary triangle with vertices of all 3 colors.



Solution: Induction on 2 (the dimension), with the stronger statement that the number of rainbow simplices is odd. Consider the 1-dimensional statement: a line segment subdivided, such that the outer endpoints receive distinct colors, but only those two colors are used internally. Then the number of rainbow elementary line segments is odd by obvious parity.

Now for dimension 2. Suppose the colors are R, B, G. Observe that if an elementary triangle is rainbow, then of its three sides, exactly 1 (odd) is BR. However, every not-rainbow elementary triangle contains either 0 or 2 (both even) BR sides. So, the parity of the total number of BR sides contained in elementary triangles is equal to the parity of the number of elementary triangles. It suffices to show that the former is odd.

Look on the elementary segments in the triangulation. Every BR segment which is on the outer edge contributes exactly 1 (odd) to this total, while every non-outer-perimeter BR segment contributes exactly 2 (even). We proved from the 1-dimensional case that the first quantity is also odd, so we are done.

Now we proceed to prove the Fixed Point theorem. Let f be a continuous function from the triangle (plus interior) to itself, and suppose for contradiction that there is no fixed point. First we show that every point in the triangle can be given a color. Since the vector $f(x) - x$ is nonzero, its angle can be used to determine the color. Specifically, if it is between 12 o'clock and 4 o'clock, color it blue. If it is between 4 o'clock and 8 o'clock, color it red. If it is between 8 o'clock and 12 o'clock, color it green.

This clearly would satisfy the Sperner conditions, since the outer vertices are the correct colors, and the outer edges only use the colors of their far endpoints. Now consider a sequence of successively finer triangulations. Color the vertices based on the rule we had set above.

In each triangulation, we find a rainbow elementary triangle, and let x_i be its center. Since the triangle is a compact set (closed and bounded), (x_i) has a convergent subsequence $z_i \to z$. Yet $f(x) - x$ is a continuous function, so there is always one of the three colors such that $f(x) - x$ is separated away from the region corresponding to that color for all x within some ϵ distance from z. But the triangles become successively smaller, and z_i eventually falls within $\epsilon/2$ distance of z, so eventually we see all three colors appear within ϵ distance of z, contradiction.

4 Ramsey theory and applications

Let s and t be positive integers. We define the *Ramsey Number* $R(s, t)$ to be the minimum integer n for which **every** red-blue coloring of the edges of K_n contains either a completely red K_s or a completely blue K_t. Ramsey's Theorem states that $R(s, t)$ is always finite, and this generalizes to more than 2 colors as well. The interesting question in this field is to find upper and lower bounds for these numbers, as well as for quantities defined in a similar spirit. One simple bound is the following.

Theorem. *(Erdős-Szekeres.) The Ramsey number $R(s, t)$ is at most $\binom{s+t-2}{s-1}$.*

Now try these.

1. Let $R(s_1, s_2, \ldots, s_k)$ be the k-color analogue of the Ramsey number. Prove that for any given (s_i), this is always finite.

   Solution: One could repeat the argument used to prove the Erdős-Szekeres bound, but here is a quick and dirty way to establish finiteness. It suffices to show that $R(s, \ldots, s)$ is always finite. Split the k colors into two equal groups, and call them color-group 1 and color-group 2. By 2-color Ramsey, there is a large complete graph in one of the two color-groups, but then the number of colors has fallen by half, so we can recurse.

2. (Schur's Theorem.) For any $k \geq 2$, there is $n > 3$ such that for any k-coloring of $\{1, \ldots, n\}$, there are three integers x, y, z of the same color such that $x + y = z$.

   Solution: We can take n to be the k-color Ramsey number $R(3, \ldots, 3)$, and actually even 1 less will do. Now consider the complete graph with vertex set $\{0, \ldots, n\}$, and color the edge ij with the color of $|i - j|$ according to the coloring of the integers. There is a monochromatic triangle, i.e. in the coloring of $[n]$ some $|i - j|$, $|j - k|$, and $|k - i|$ all got the same color. But two of these sum to the third!

3. (IMO 1978/6.) Prove that in any coloring of the integers $\{1, \ldots, 1978\}$ with 6 colors, there are integers x, y, z, all of the same color, satisfying $x + y = z$.

   Solution: By Schur's Theorem above, it suffices to show that 1978 beats the 6-color Ramsey number for triangles. Let r_k be the k-color Ramsey number for triangles. The Erdős-Szekeres argument gives $r_k \leq 2 + k \cdot (r_{k-1} - 1)$, where we have used the fact that $R(3,3,3,3,2) = R(3,3,3,3)$, for example. Let $s_k = r_k - 1$. We then have the recursion $s_k < ks_{k-1}$, so $s_6 < 6 \cdot 5 \cdot 4 \cdot 3 \cdot s_2$, and $s_2 = R(3,3) - 1 = 5$. This gives $s_6 < 1800$, which clearly suffices.

4. (Fermat's Last Theorem is false in $\mathbb{Z}_p$.) For every $m \geq 1$, there is p_0 such that for any prime $p \geq p_0$, the congruence $x^m + y^m \equiv z^m \pmod{p}$ has a solution.

   Solution: Let g be a generator of the cyclic group $\mathbb{Z}_p^\times$, also known as a primitive root. Color the elements of $\mathbb{Z}_p^\times$ with m colors, where g^{km+j} gets color j. Assume that p is large enough, so that Schur's Theorem implies that there are $x + y = z$ all of the same color. This means $x = g^{am+j}$, $y = g^{bm+j}$, $z = g^{cm+j}$. But then $(g^a)^m + (g^b)^m = (g^c)^m$, giving the desired solution.

5. (Hypergraph version.) A graph is a set V of vertices, together with a collection of edges, which are subsets of V with 2 elements each. An r-uniform hypergraph is a set V of vertices, together with a collection of r-edges, which are subsets of V with r elements each.

   Let the hypergraph Ramsey number $R^{(r)}(s,t)$ be the minimum n such that the following holds. Let each of the $\binom{n}{r}$ total r-edges on n vertices be colored either red or blue. Then there is a set of s vertices such that all $\binom{s}{r}$ of its r-edges are red, or a set of t vertices such that all $\binom{t}{r}$ of its r-edges are blue. Prove that $R^{(r)}(s,t)$ is finite.

   Solution: Induct on r, s, t, and use the fact that we already proved everything for $r = 2$. Let $X = R^{(r)}(s-1,t)$, and let $Y = R^{(r)}(s,t-1)$. Let $n = R^{(r-1)}(X,Y) + 1$, and let a coloring of the r-edges be given.

   It suffices to show that there is a red $K_s^{(r)}$ or a blue $K_t^{(r)}$. Consider an arbitrary vertex v, and construct an auxiliary $(r-1)$-uniform hypergraph on $n-1$ vertices. The vertices will correspond to $V \setminus \{v\}$, and the way we color an $(r-1)$-edge E is to copy the color of $E \cup \{v\}$ in the originally given coloring. By induction, in this auxiliary hypergraph we either find a red $K_X^{(r-1)}$ or a blue $K_Y^{(r-1)}$.

   Suppose it is the former (the latter follows similarly), and look on those X vertices we have now found. Since X was also a Ramsey number, in the original graph either there is already a blue $K_t^{(r)}$ there (so we are done), or there is only a red $K_{s-1}^{(r)}$. But from our information on the auxiliary graph, we also know that every $r-1$ of the vertices here make a red r-edge with v, so adding v will make a red $K_s^{(r)}$.

6. (Happy ending problem.) Given any 5 distinct points in the plane, no 3 collinear, show that some 4 are in *convex position*, i.e., forming the vertices of a convex quadrilateral.

   Solution: If the convex hull has size 4 or more, already done. Otherwise, convex hull is a triangle ABC and there are 2 internal points. Consider the line ℓ determined by those internal points. Since no 3 collinear, it divides $\{A, B, C\}$ into two groups, one of which has size 2. Use those 2, plus the 2 internal points.

7. (Erdős-Szekeres.) For every integer n, there is some finite N such that the following holds. Given any N distinct points in the plane, no 3 collinear, some n are in convex position.

   Remark. It is conjectured that $N = 1 + 2^{n-2}$ suffices for all $n \geq 3$, and known that $N \geq 1 + 2^{n-2}$ is required. The best known upper bound is of order $4^n/\sqrt{n}$.

   Solution: If a set S of points has the property that every 4-subset is in convex position, then all of S is in convex position. To see this, suppose there was some point P strictly inside the convex hull of S. Triangulate the convex hull using diagonals, and P will be strictly inside one of the triangles, say ABC. Then $PABC$ is concave, contradiction.

   Now suppose we have $R^{(4)}(5, n)$ many points. For every 4-set of points, color the corresponding 4-edge red if they are not in convex position, blue otherwise. Our hypergraph Ramsey bound implies that there must either be 5 vertices with all 4-edges red, or n vertices with all 4-edges blue.

   But the 5 vertices with all 4-edges red contradicts the Happy Ending Problem, so we must have the latter. By our opening remark, in fact the n points are all in convex position.

5 List coloring

Recall that the chromatic number of a graph is the minimum t such that every vertex can be assigned one of t colors without any two adjacent vertices receiving the same color. But what if we do not require that every vertex's color choices are exactly the same? Specifically, consider the situation where every vertex has a separate list of t distinct colors, but these may come from a much larger pool. We then ask for a proper coloring from these lists, which means that every vertex picks one of its colors without having any two adjacent vertices pick the same color. The *list chromatic number* of a graph G is defined to be the minimum t such that no matter how we assign color lists of size t to each vertex, it is always possible to properly color G from these lists.

At first glance, it may appear that the list chromatic number should be smaller than the chromatic number, because it seems that spreading out the colors across lists should make it easier to avoid conflicts. But this intuition is very very wrong.

1. (List chromatic number can be arbitrarily large.) For any k, there is a graph G with chromatic number 2, but list chromatic number greater than k.

   Remark. The construction I have in mind has about 2^{2k} vertices.

   Solution: Complete bipartite graph of equal sides. Pool of $2k - 1$ total colors. On each side, there is exactly one vertex with color list corresponding to each subset of k colors from the $2k - 1$ pool. So, $\binom{2k-1}{k}$ vertices per side, giving roughly 2^{2k} vertices in total. But the maximum number of different colors NOT selected by the LHS is at most $k - 1$. Indeed, if there were k colors missed, then there was some LHS vertex with exactly that as its color list, contradiction. Therefore, the LHS used a total of at least k colors.

   Similar argument for RHS, so the RHS used a total of at least k colors. But complete bipartite graph, so LHS and RHS need disjoint sets of colors. This requires at least $2k$ total colors, contradiction.

2. (Alon-Spencer, Exercise 2.9.) Let G be a bipartite graph with n vertices, with lists of size strictly greater than $\log_2 n$ assigned to each vertex. Prove there is a proper coloring from these lists.

Solution: For each color, randomly choose whether it will be permitted on the LHS or on the RHS. If every vertex still has a permitted color choice left in its list, then there is a proper list coloring. The probability that a vertex v fails is exactly 2^{-c_v}, where c_v is the number of colors in v's list. But $c_v > \log_2 n$, so the failure probability of v is strictly less than $\frac{1}{n}$.

By union bound over all n vertices v, the probability of failure is strictly below 1, so there is some partitioning of the colors which succeeds.