Week 1: Induction, Pigeonhole, and the Extremal Principle

Lecture notes · Thu Sep 24 / Fri Sep 25

Toolbox: three ways to prove something exists.

1. **Pigeonhole.** N objects in n boxes put at least $\lceil N/n \rceil$ in one box. Averaging form: if $y_1 + \cdots + y_n = s$, some $y_i \geq s/n$ and some $y_j \leq s/n$.
2. **The art is choosing the boxes:** residues mod n, odd parts, partial sums, fractional parts in short intervals, small cells of a region. Ask: *what do I want two objects to share?*
3. **Extremal principle.** A finite nonempty set of reals has a largest and smallest element; a nonempty set of positive integers has a least one. Pick the extreme object (longest, closest, largest area, least counterexample) and use that nothing beats it, usually by contradiction. "Sort the numbers" is this principle in disguise.
4. **Induction.** Ordinary, strong ($P(k)$ for all $k < n$ implies $P(n)$), or on a well-chosen parameter ($p+q$ for a fraction p/q, the largest element, the number of odd terms). It is also how "repeat this step" becomes a proof.
5. **Strengthen the hypothesis.** If the inductive step fails, prove something stronger: $\sum_{k \leq n} 1/k^2 < 2$ resists induction, but $\sum_{k \leq n} 1/k^2 \leq 2 - 1/n$ goes through in one line.

Example 1 (Pigeonhole with partial sums)**.** Among any n integers $a_1, \ldots, a_n$ some nonempty block of consecutive terms has sum divisible by n. *Let $s_0 = 0$ and $s_k = a_1 + \cdots + a_k$. The $n+1$ numbers $s_0, \ldots, s_n$ fall into the n boxes "residue mod n," so two of them, s_i and s_j with $i < j$, are congruent mod n. Then $n \mid s_j - s_i = a_{i+1} + \cdots + a_j$, a nonempty consecutive block. Including $s_0 = 0$ is what makes the count $n+1$.*

Example 2 (Extremal principle in a tournament)**.** In a round-robin tournament (every pair plays once, no ties) there is a player K such that every other player either lost to K or lost to someone who lost to K. *Choose K with the maximum number of wins (possible: finitely many players). If some Y beat K and also beat everyone K beat, then Y has at least (wins of K) $+1$ wins, contradicting maximality. So every Y lost to K or to somebody K beat. Nothing was constructed; we only used "no one has more wins than K."*

Example 3 (Strong induction, Putnam 2005 A1)**.** Every positive integer is a sum of one or more numbers of the form $2^r 3^s$ $(r, s \geq 0)$ in which no summand divides another. *Strong induction on n, with base case $n = 1 = 2^0 3^0$. Call a sum with the required property a* good representation*, and note that doubling every summand of a good representation gives a good representation of twice the number: multiplying all summands by 2 changes none of the divisibility relations among them.*

Even n. *Double a good representation of $n/2$.*

Odd $n > 1$. *Let 3^s be the largest power of 3 with $3^s \leq n$, so $3^s \leq n < 3^{s+1}$. If $n = 3^s$ we are done. Otherwise $n - 3^s$ is positive and even (odd minus odd), so $m = (n - 3^s)/2$ is a positive integer, and $n < 3^{s+1}$ gives $n - 3^s < 2 \cdot 3^s$, i.e. $m < 3^s$. By the induction hypothesis m has a good representation $m = \sum 2^{a_i} 3^{b_i}$. Double it and adjoin 3^s:*

$$n \;=\; 3^s + \sum 2^{a_i+1} 3^{b_i}.$$

Three checks show this is good. (i) The doubled summands do not divide one another, by the doubling remark. (ii) A doubled summand $2^{a_i+1} 3^{b_i}$ carries a positive power of 2, so it cannot divide 3^s. (iii)

3^s divides $2^{a_i+1}3^{b_i}$ only if $s \le b_i$; but $3^{b_i} \le 2^{a_i}3^{b_i} \le m < 3^s$ forces $b_i < s$. The step worked because we peeled off the largest *power of* 3 *(that is what makes $m < 3^s$), not because we looked at $n-1$.*

Pitfall. (1) Pigeonhole gives *two objects in one box*, not the two you wanted; make sure your conclusion follows for *any* two in a box (in Example 1, any $i < j$ works). (2) "Take the smallest counterexample" needs a well-ordered set: positive integers, yes; positive reals, no. (3) An extremal object exists only for finite sets or compact situations; say why it exists ("S is finite, so some triple has maximal area"). (4) State $P(n)$ exactly and check the base case; the classic lost point is an inductive step that silently assumes $n \ge 2$.

How to recognize this technique on the exam. Pigeonhole is the first guess when a problem says "among any N of these, some two…" or hands you more objects than natural classes; look for a count one more than a modulus, a number of intervals, or a number of odd numbers. The extremal principle is the first guess when something must exist in a finite configuration (points, tournaments, colorings) and no formula is in sight: pick the largest, closest, or longest and argue it already works or could be improved. Induction is the natural shape when a statement indexed by n contains a smaller case inside it; if that sub-case is not "$n-1$," expect strong induction or a nonstandard parameter.

Writing a solution that scores. State precisely what you will prove, including any reformulation ("We show that for all $n \ge 1$, …"). Write full sentences with every quantifier visible; a page of equations with no words is not a proof. Justify every step a skeptical reader could question, naming the tool as you use it ("by pigeonhole on the $n+1$ residues…"). Scores are 10 for a complete argument with cosmetic slips, 8–9 for a small gap, and almost always 0–2 otherwise: partial progress, an answer without proof, or the "easy half" rarely earns more than 1. Cross out what you abandon; never claim more than you have shown.

Suggested timing (15 minutes). 0–3: pigeonhole, Example 1, "choose the boxes." 3–7: extremal principle, Example 2, sorting heuristic. 7–12: induction, Example 3, choosing the parameter and strengthening the hypothesis. 12–15: the writing box and exam format (Sat Dec 5, four 90-minute sessions with breaks, 12 problems at 10 points, median near 0); then groups start on Problems 1–4.