Process Analysis

ft. Games and Understanding

Evan Chen

February 6, 2020
DCY-PROCESS, OTIS*

§1 Lecture notes

§1.1 Discussion

Rigid problems are especially good candidates for solutions that involve lots of soft methods, because they involve a fixed, concrete structure. (See https://usamo.wordpress.com/2019/05/03/hard-and-soft-techniques/ for an explanation of "soft method".)

This lecture features problems for which the same type of "deep understanding" is important, but this time the problems do not necessarily have a concrete flavor: most frequently, because it involves a game or process which has such a large state space that it is much harder to imagine concretely.

You might be used to thinking about processes in terms of invariants or monovariants, but I think that's really just a special case of trying to understand the process with soft methods, with invariants or monovariants being a formal technique used after the fact. (Religiously trying to use monovariants leads to not solving USA TST 2017/4, one of the walkthroughs.) Instead, it's much more important to just try to explore the problem, rather than trying to search specifically for a magical invariant or monovariant that may not even exist (and even if it does exist, could have been more motivated by exploration).

Here are some heuristics:

- The particular task you are asked to prove often feels superficial compared to what you actually figure out.
- Thus upon solving the problem you feel like you have a pretty good idea what's going on, and why the problem "should" be true.
- There may or may not be a concrete structure you are studying (if there is, the problem is quite likely "rigid"). However in both cases the problem has some *complexity*, and understanding it well is the point of the problem.

Note: this is not quite the same as the IOI unit. The IOI unit features algorithms as well, but the flavor of problems there is more constructive, e.g. "determine an algorithm that achieves so-and-so goal". (Whereas here, problems are often of the form "show that the process always terminates", etc.) In some sense, rigid is to free as process is to IOI.

§1.2 Examples

Example 1.1 (Shortlist 2016 A2)

Find the smallest constant $C > 0$ for which the following statement holds: among any five distinct positive real numbers, one can label four different ones as p, q, r, s such that

$$\left|\frac{p}{q} - \frac{r}{s}\right| \le C.$$

Walkthrough.

(a) We will first solve the four-variable version. Given four real numbers, describe the permutation (a, b, c, d) of them which minimizes $\left|\frac{a}{b} - \frac{c}{d}\right|$.

(b) Arrange the numbers $a < b < c < d < e$. According to (a), there are only five expressions (rather than $5 \cdot 4 \cdot 3 \cdot 2 = 120$) which we need to consider. What are they?

(c) One should find that only five distinct fractions arise. Arrange them in a circle and use this to solve the problem.

The main part of this problem is part (a): the idea that if/once you are stuck, it's immensely helpful to consider the case of four variables first. Conversely, it is very easy to get stuck on this problem or even figure out the correct answer if one never considers (a). I think if you get stuck on the problem for hours and then realize you never attempted the $n = 4$ case seriously, then you need to be much more willing to take a step back.

Example 1.2 (USA TST 2017/4)

You are cheating at a trivia contest. For each question, you can peek at each of the $n > 1$ other contestant's guesses before writing your own. For each question, after all guesses are submitted, the emcee announces the correct answer. A correct guess is worth 0 points. An incorrect guess is worth -2 points for other contestants, but only -1 point for you, because you hacked the scoring system. After announcing the correct answer, the emcee proceeds to read out the next question. Show that if you are leading by 2^{n-1} points at any time, then you can surely win first place.

Walkthrough. Hint: *not* monovariants. (Also, to clarify, questions can have more than two possible answers.)

(a) Show that we can WLOG assume that you answer all questions incorrectly, by discarding any rounds where you get the answer correct.

(b) By shifting the scores by 1, re-frame the problem in such a way your score never changes. This should make the problem easier to think about.

(c) Assume for now that in each round, only two distinct answers are given. Thus each round corresponds to a partition of $\{1, \dots, n\}$ into two halves. How many possible rounds are there? If two identical rounds occur, can you cancel them out?

(d) Solve the problem in the special case described in (c).

(e) Extend the solution to the general case. (This requires some care. It is tempting to try and argue that the assumption in (c) can be made without loss of generality, but as far as I know this is not actually true. Instead, keep track of a history of sets of contestants answering correctly. See the remarks in the solution for further details.)

(f) Optional: can you improve the bound to $2^{n-2} + 1$?

The right mentality for the problem is to really think about the process, which is what leads to the observations such as those in (a) - (c), rather than immediately trying to contrive a monovariant (which was a common failure mode for the problem).

Example 1.3 (RMM 2018/3)

Ann and Bob play a game on the unit edges of an infinite square lattice, making moves in turn. Ann makes the first move. A move consists of orienting any unit edge of the grid that has not been oriented before. If at some stage, some oriented edges form an oriented cycle, Bob wins. Does Bob have a strategy that guarantees him to win?

Walkthrough.

(a) Optionally, show that on a triangular grid, Bob can win in at most five turns. So it'll be necessary to use the square grid in a substantial way.

(b) Play the game for a bit, and convince yourself that there is unlikely to be a bounded strategy for Bob.

(c) Solve the problem in the case where Ann moves *second* rather than first. (Hint: pair up the edges of the grid.)

(d) Modify your strategy for (c) to work when Ann moves first, thus solving the original problem.

One main lesson from the problem might be (c). There are two motivations for doing this different problem: it is more natural for Ann this way, and the problem condition (avoiding directed cycles) has nothing to do with the order of the players, i.e. there is no natural reason that the cycle-creator should move second versus first.

I think the main reason I was able to solve the problem was that (c) gave me a foothold to approach the full problem.

§2 Practice problems

Instructions: Solve [40♣]. If you have time, solve [60♣]. Problems with red weights are mandatory.

> Great work! Because this message is prerecorded, any observations related to your performance are speculation on our part. Please disregard any undeserved compliments.
>
> The Announcer in *Portal 2*

[3♣] **Problem 1** (USAMO 2000/3)**.** A game of solitaire is played with R red cards, W white cards, and B blue cards. A player plays all the cards one at a time. With each play he accumulates a penalty. If he plays a blue card, then he is charged a penalty which is the number of white cards still in his hand. If he plays a white card, then he is charged a penalty which is twice the number of red cards still in his hand. If he plays a red card, then he is charged a penalty which is three times the number of blue cards still in his hand.

Find, as a function of R, W, and B, the minimal total penalty a player can amass and the number of ways in which this minimum can be achieved.

[2♣] **Problem 2** (TSTST 2014)**.** Let $\leftarrow$ denote the left arrow key on a standard keyboard. If one opens a text editor and types the keys "ab$\leftarrow$ cd $\leftarrow\leftarrow$ e $\leftarrow\leftarrow$ f", the result is "faecdb". We say that a string B is *reachable* from a string A if it is possible to insert some amount of $\leftarrow$'s in A, such that typing the resulting characters produces B. So, our example shows that "faecdb" is reachable from "abcdef".

Prove that for any two strings A and B, A is reachable from B if and only if B is reachable from A.

[3♣] **Problem 3** (Shortlist 2016 A3)**.** Find all positive integers n such that the following statement holds: Suppose real numbers a_1, a_2, $\dots$, a_n, b_1, b_2, $\dots$, b_n satisfy $|a_k|+|b_k|=1$ for all $k=1,\dots,n$. Then there exists ε_1, ε_2, $\dots$, ε_n, each of which is either -1 or 1, such that

$$\left|\sum_{i=1}^{n}\varepsilon_i a_i\right| + \left|\sum_{i=1}^{n}\varepsilon_i b_i\right| \le 1.$$

[3♣] **Problem 4** (USA TST 2015/3)**.** A physicist encounters 2015 atoms called usamons. Each usamon either has one electron or zero electrons, and the physicist can't tell the difference. The physicist's only tool is a diode. The physicist may connect the diode from any usamon A to any other usamon B. (This connection is directed.) When she does so, if usamon A has an electron and usamon B does not, then the electron jumps from A to B. In any other case, nothing happens. In addition, the physicist cannot tell whether an electron jumps during any given step. The physicist's goal is to isolate two usamons that she is 100% sure are currently in the same state. Is there any series of diode usage that makes this possible?

[3♣] **Problem 5** (USAMO 1994/2)**.** The vertices of a 99-gon are initially colored so that consecutive vertices are red, blue, red, blue, $\dots$, red, blue, yellow. We make a sequence of modifications in the coloring, changing the color of one vertex at a time to one of the three given colors (red, blue, yellow), under the constraint that no two adjacent vertices may be the same color. By making a sequence of such modifications, is it possible to arrive at the coloring in which consecutive vertices are red, blue, red, blue, red, blue, $\dots$, red, yellow, blue?

[5♣] **Problem 6** (IMO 2016/6). There are $n \geq 2$ line segments in the plane such that every two segments cross and no three segments meet at a point. Geoff has to choose an endpoint of each segment and place a frog on it facing the other endpoint. Then he will clap his hands $n-1$ times. Every time he claps, each frog will immediately jump forward to the next intersection point on its segment. Frogs never change the direction of their jumps. Geoff wishes to place the frogs in such a way that no two of them will ever occupy the same intersection point at the same time.

(a) Prove that Geoff can always fulfill his wish if n is odd.

(b) Prove that Geoff can never fulfill his wish if n is even.

[3♣] **Problem 7** (EGMO 2018/3). The n contestants of EGMO are named C_1, C_2, $\dots$, C_n. After the competition, they queue in front of the restaurant according to the following rules.

- The Jury chooses the initial order of the contestants in the queue.
- Every minute, the Jury chooses an integer i with $1 \leq i \leq n$.
  - If contestant C_i has at least i other contestants in front of her, she pays one euro to the Jury and moves forward in the queue by exactly i positions.
  - If contestant C_i has fewer than i other contestants in front of her, the restaurant opens and the process ends.

For every n, prove that this process must terminate and determine the maximum number of euros that the Jury can collect by cunningly choosing the initial order and the sequence of moves.

[3♣] **Problem 8** (USAMO 2019/5). Let m and n be relatively prime positive integers. The numbers $\frac{m}{n}$ and $\frac{n}{m}$ are written on a blackboard. At any point, Evan may pick two of the numbers x and y written on the board and write either their arithmetic mean $\frac{1}{2}(x+y)$ or their harmonic mean $\frac{2xy}{x+y}$. For which (m, n) can Evan write 1 on the board in finitely many steps?

[5♣] **Problem 9** (Shortlist 2015 C4). Let n be a positive integer. Two players Frisk and Sans play a game in which they take turns choosing positive integers $k \leq n$. The rules of the game are:

(i) A player cannot choose a number that has been chosen by either player on any previous turn.

(ii) A player cannot choose a number consecutive to any of those the player has already chosen on any previous turn.

(iii) The game is a draw if all numbers have been chosen; otherwise the player who cannot choose a number anymore loses the game.

The protagonist Frisk takes the first turn. Please determine for each n the outcome of the game assuming optimal play.

[5♣] **Problem 10** (Shortlist 2013 N5). Fix an integer $k > 2$. Two players, called Ana and Banana, play the following game of numbers. Initially, some integer $n \geq k$ gets written on the blackboard. Then they take moves in turn, with Ana beginning. A player making a move erases the number m just written on the blackboard and replaces it by some number m' with $k \le m' < m$ that is coprime to m. The first player who cannot move anymore loses.

An integer $n \ge k$ is called good if Banana has a winning strategy when the initial number is n, and bad otherwise. Consider two integers $n, n' \ge k$ with the property that each prime number $p \le k$ divides n if and only if it divides n'. Prove that either both n and n' are good or both are bad.

[9♣] **Problem 11** (APMO 2019/4)**.** Consider a 2018×2019 board with integers in each unit square. In one operation, one can replace the integer in some cell of the board with the average of its orthogonal neighbors. Is it always possible to make all numbers equal?

[5♣] **Problem 12** (ELMO 2019/2)**.** Let $n \ge 3$ be a fixed positive integer. Evan has a convex n-gon in the plane and wishes to construct the centroid of its vertices. He has no standard ruler or compass, but he does have a device with which he can dissect the segment between two given points into m equal parts. For which m can Evan necessarily accomplish his task?

[5♣] **Problem 13** (Shortlist 2009 C5)**.** Five identical empty buckets of 2-liter capacity stand at the vertices of a regular pentagon. Cinderella and her wicked Stepmother go through a sequence of rounds. At the start of each round, the Stepmother distributes one liter of water arbitrarily over the five buckets. Then Cinderella chooses two neighboring buckets and empties them; the next round then begins. Can Cinderella always prevent the Stepmother from causing a bucket to overflow?

[9♣] **Problem 14** (IMO 2012/3)**.** The liar's guessing game is a game played between two players A and B. The rules of the game depend on two fixed positive integers k and n which are known to both players.

At the start of the game A chooses integers x and N with $1 \le x \le N$. Player A keeps x secret, and truthfully tells N to player B. Player B now tries to obtain information about x by asking player A questions as follows: each question consists of B specifying an arbitrary set S of positive integers (possibly one specified in some previous question), and asking A whether x belongs to S. Player B may ask as many questions as he wishes. After each question, player A must immediately answer it with yes or no, but is allowed to lie as many times as she wants; the only restriction is that, among any $k+1$ consecutive answers, at least one answer must be truthful.

After B has asked as many questions as he wants, he must specify a set X of at most n positive integers. If x belongs to X, then B wins; otherwise, he loses. Prove that:

(a) If $n \ge 2^k$, then B can guarantee a win.

(b) For all sufficiently large k, there exists an integer $n \ge (1.99)^k$ such that B cannot guarantee a win.

[5♣] **Problem 15** (IMO 2000/3)**.** Let $n \ge 2$ be a positive integer and λ a positive real number. Initially there are n fleas on a horizontal line, not all at the same point. We define a move as choosing two fleas at some points A and B, with A to the left of B, and letting the flea from A jump over the flea from B to the point C so that $\frac{BC}{AB} = \lambda$.

Determine all values of λ such that, for any point M on the line and for any initial position of the n fleas, there exists a sequence of moves that will take them all to the position right of M.

[5♣] **Problem 16** (Shortlist 2011 C5)**.** Let m be a positive integer, and consider a $m \times m$ checkerboard consisting of unit squares. At the center of some of these unit squares there is an ant. At time 0, each ant starts moving with speed 1 parallel to some edge of the checkerboard. When two ants moving in the opposite directions meet, they both turn 90° clockwise and continue moving with speed 1. When more than 2 ants meet, or when two ants moving in perpendicular directions meet, the ants continue moving in the same direction as before they met. When an ant reaches one of the edges of the checkerboard, it falls off and will not re-appear.

Considering all possible starting positions, determine the latest possible moment at which the last ant falls off the checkerboard, or prove that such a moment does not necessarily exist.

[9♣] **Problem 17** (ELMO 2019/3)**.** Let $n \geq 3$ be a positive integer. In a game, n players sit in a circle in that order. Initially, a deck of $3n$ cards labeled $\{1, \dots, 3n\}$ is shuffled and distributed among the players so that every player holds 3 cards in their hand. Then, every hour, each player simultaneously gives the smallest card in their hand to their left neighbor, and the largest card in their hand to their right neighbor. (Thus after each exchange, each player still has exactly 3 cards.)

Prove that each player's hand after the first $n-1$ exchanges is their same as their hand after the first $2n-1$ exchanges.

[3♣] **Problem 18** (ELMO SL 2019 C3)**.** In the game of Ring Mafia, there are 2019 counters arranged in a circle, 673 of these which are mafia, and the remaining 1346 which are town. Two players, Tony and Madeline, take turns with Tony going first. Tony does not know which counters are mafia but Madeline does.

On Tony's turn, he selects any subset of the counters (possibly the empty set) and removes all counters in that set. On Madeline's turn, she selects a town counter which is adjacent to a mafia counter and removes it. (Whenever counters are removed, the remaining counters are brought closer together without changing their order so that they still form a circle.) The game ends when either all mafia counters have been removed, or all town counters have been removed.

Is there a strategy for Tony that guarantees, no matter where the mafia counters are placed and what Madeline does, that at least one town counter remains at the end of the game?

[9♣] **Problem 19** (Goodstein theorem)**.** Let a_2 be any positive integer. We define the infinite sequence $a_2, a_3, \dots$ recursively as follows. If $a_n = 0$, then $a_{n+1} = 0$. Otherwise, we write a_n in base n, then write all exponents in base n, and so on until all numbers in the expression are at most n. Then we replace all instances of n by $n+1$ (including the exponents!), subtract 1, and set the result to a_{n+1}. For example, if $a_2 = 11$ we have

$$\begin{aligned} a_2 &= 2^3 + 2 + 1 = 2^{2+1} + 2 + 1 \\ a_3 &= 3^{3+1} + 3 + 1 - 1 = 3^{3+1} + 3 \\ a_4 &= 4^{4+1} + 4 - 1 = 4^{4+1} + 3 \\ a_5 &= 5^{5+1} + 3 - 1 = 5^{5+1} + 2 \end{aligned}$$

and so on. Prove that $a_N = 0$ for some integer $N > 2$.

> **Remark.** I personally think the Goodstein problem is okay, but I know a lot of people who *really* enjoy it.

[5♣] **Problem 20** (Kvant M2573, originally RMM 2019)**.** Two ants crawl along the edges of a convex polyhedron. Each ant's route ends at the vertex where it begins, with the ant never visiting any point more than once until the end of its route. On each face F of the polyhedron, write down the number of edges of F which are part of the first ant's route and the number of edges of F which are part of the second ant's route. Does there exist a polyhedron and some pair of routes as above, such that exactly one face bears a pair of non-equal numbers?

[1♣] **Mini Survey.** At the end of your submission, answer the following questions.

(a) About how many hours did the problem set take?

(b) Name any problems that stood out (e.g. especially nice, instructive, boring, or unusually easy/hard for its placement).

Any other thoughts are welcome too. Examples: suggestions for new problems to add, things I could explain better in the notes, overall difficulty or usefulness of the unit.

§3 Solutions to the walkthroughs

§3.1 Solution 1.1, Shortlist 2016 A2

The answer is $C = \frac{1}{2}$. For construction, take the five numbers ε, 1, $2-\varepsilon$, 2, $2+\varepsilon$ for arbitrarily small ε.

We now show selecting the five numbers is always possible. Assuming the numbers are $a < b < c < d < e$, consider the five fractions

$$a/c, \quad b/d, \quad c/e, \quad a/d, \quad b/e$$

and arrange them in a pentagon. These fractions are less than 1, so three of them must be on the same side of $\frac{1}{2}$; two are adjacent on the pentagon and give the desired.

> **Remark.** Finding the answer is a large part of the difficulty of this problem. The way I motivated is solving the sub-problem with four numbers $a < b < c < d$, the minimum fraction is $|\frac{a}{c} - \frac{b}{d}|$ (check this). Thus given five distinct numbers $a < b < c < d < e$, the five values in consideration are:
>
> - $a/c - b/d$
> - $a/c - b/e$
> - $a/d - b/e$
> - $a/d - c/e$
> - $b/d - c/e$
>
> One can then arrange these five fractions (less than 1) around the vertices of a pentagon, connecting fractions with no common term; trying to maximize the minimal difference naturally gives the optimal value of 1/2.

§3.2 Solution 1.2, USA TST 2017/4

We will prove the result with 2^{n-1} replaced even by $2^{n-2}+1$.

We first make the following reductions. First, change the weights to be $+1$, -1, 0 respectively (rather than 0, -2, -1); this clearly has no effect. Also, WLOG that all contestants except you initially have score zero (and that your score exceeds 2^{n-2}). WLOG ignore rounds in which all answers are the same. Finally, ignore rounds in which you get the correct answer, since that leaves you at least as well off as before — in other words, we'll assume your score is always fixed, but you can pick any group of people with the same answers and ensure they lose 1 point, while some other group gains 1 point.

The key observation is the following. Consider two rounds R_1 and R_2 such that:

- In round R_1, some set S of contestants gains a point.
- In round R_2, the set S of contestants all have the same answer.

Then, if we copy the answers of contestants in S during R_2, then the sum of the scorings in R_1 and R_2 cancel each other out. In other words we can then ignore R_1 and R_2 forever.

We thus consider the following strategy. We keep a list $\mathcal{L}$ of subsets of $\{1, \dots, n\}$, initially empty. Now do the following strategy:

- On a round, suppose there exists a set S of people with the same answer such that $S \in \mathcal{L}$. Then, copy the answer of S, causing them to lose a point. Delete S from $\mathcal{L}$. (Importantly, we do not add any new sets to $\mathcal{L}$.)

- Otherwise, copy any set T of contestants, selecting $|T| \geq n/2$ if possible. Let S be the set of contestants who answer correctly (if any), and add S to the list $\mathcal{L}$. Note that $|S| \leq n/2$, since S is disjoint from T.

By construction, $\mathcal{L}$ has no duplicate sets. So the score of any contestant c is bounded above by the number of times that c appears among sets in $\mathcal{L}$. The number of such sets is clearly at most $\frac{1}{2} \cdot 2^{n-1}$. So, if you lead by $2^{n-2}+1$ then you ensure victory. This completes the proof!

Remark. Several remarks are in order. First, we comment on the bound $2^{n-2}+1$ itself. The most natural solution using only the list idea gives an upper bound of $(2^n-2)+1$, which is the number of nonempty proper subsets of $\{1, \dots, n\}$. Then, there are two optimizations one can observe:

- In fact we can improve to the number of times any particular contestant c appears in some set, rather than the total number of sets.
- When adding new sets S to $\mathcal{L}$, one can ensure $|S| \leq n/2$.

Either observation alone improves the bound from 2^n-1 to 2^{n-1}, but both together give the bound $2^{n-2}+1$. Additionally, when n is odd the calculation of subsets actually gives $2^{n-2} - \frac{1}{2}\binom{n-1}{\frac{n-1}{2}} + 1$. This gives the best possible value at both $n=2$ and $n=3$. It seems likely some further improvements are possible, and the true bound is suspected to be polynomial in n.

Secondly, the solution is highly motivated by considering a true/false contest in which only two distinct answers are given per question. However, a very natural mistake (which graders assessed as a two-point deduction) is to try and prove that in fact one can "WLOG" we are in the two-question case. The proof of this requires substantially more care than expected. For instance, set $n=3$. If $\mathcal{L} = \{\{1\},\{2\},\{3\}\}$ then it becomes impossible to prevent a duplicate set from appearing in $\mathcal{L}$ if all contestants give distinct answers. One might attempt to fix this by instead adding to $\mathcal{L}$ the *complement* of the set T described above. The example $\mathcal{L} = \{\{1,2\},\{2,3\},\{3,1\}\}$ (followed again by a round with all distinct answers) shows that this proposed fix does not work either. This issue affects all variations of the above approach.

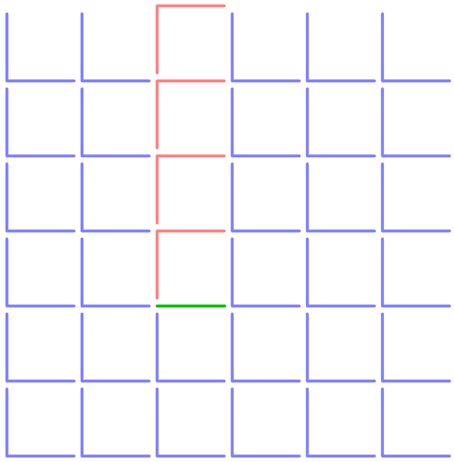
Remark. Here are some motivations for the solution:

1. The exponential bound 2^n suggests looking at subsets.
2. The $n=2$ case suggests the idea of "repeated rounds". (I think this $n=2$ case is actually really good.)
3. The "two distinct answers" case suggests looking at rounds as partitions (even though the WLOG does not work, at least not without further thought).
4. There's something weird about this problem: it's a finite bound over unbounded time. This is a hint to *not* worry excessively about the actual scores, which turn out to be almost irrelevant.

§3.3 Solution 1.3, RMM 2018/3

The answer is no; Ann can prevent him from forming a cycle.

Suppose Ann makes the first move at some horizontal edge e, say. Then we partition all edges other than e into L-pairs as shown in the following figure (with e in green):



Then, every time Bob plays on an edge in a pair, Ann plays the on the other edge in the pair, orienting it so that either both arrowheads point towards the center, or neither of them do.

We claim this strategy works. And indeed, no oriented cycle can "use a marked corner", and by construction this prevents the cycle from existing at all.

> **Remark.** If Bob goes first instead, then the simpler tiling of lower-left L's works fine, without the need for the green/red perturbation above. Solving this case makes the above solution more motivated.