On Contest Creation

A reflection of my math contest creation experiences

Dennis Chen

August 7, 2021

This is a reflection of the contests I've created in the last two years. At the same time, it is meant to serve as a loose guide for anyone looking to create their own contest or refine their craft.

1 How to start

1.1 What's your purpose?

The mission of your contest isn't something you actively need to think about, but I think it's important to keep your contest purpose somewhere in your head because it informs your **test compilation values**.

Take the MAT, for instance. Most contests have primary goals of being a) educational and b) fun, and sometimes this comes at the expense of problem quality. MAT instead focuses on we strive first and foremost to make a beautiful test with only standout problems. I like to believe that our contest is also educational and fun, but it is undeniable our focus on **problem quality** supercedes most groups.[1]

Even though none of these goals are at odds with each other, the difficulty of putting together a test compounds exponentially with each goal you have. So if you're not meeting your goals, you either have to tighten your focus, spend more time, or put less problems on the test.

Two of your contest lines do not have to have the same philosophy. The MAT is produced like a work of art, meaning it is smaller and takes a lot longer to finish. The JMCs are meant for practice, so less emphasis is placed on egoism (making sure each individual problem is good) and moreso on cohesion (the test is good practice as a whole).

Your philosophy for a line of contests can also change between contests. For instance, the first MAT had a large emphasis on egoism. We wanted to put together the best test we could, and we made sure that the difficulty scale made sense on a **relative** scale. But next time we're going to put together a contest whose difficulty makes more sense on an **absolute** scale. This is especially because we're going to be moving locally, where the average contestant will not be as skilled.

1.2 Getting people

2 Time Management

3 Testsolving and final checks

You should get a number of people with different skill levels **that have had minimal contact with the test** involved in the test-solving process. Test-solving is a process that

[1] Actually, now that I think about it, MAT really is like Blue Lock. A contest full of standout problems and a team full of standout strikers. Perhaps I should use this analogy more often, at least in MAC.

means different things to different people, and no two groups use test-solving for the same purposes. Here are some common ones.

1. Validity checking – checking that all of the problems are correctly stated, the answers are what the contest-holders think they are, etc.

2. Quality checking – checking that the individual problems are good and the composition is good.

3. Difficulty checking – checking that the test is reasonably hard to be interesting to most of the target audience, and more commonly, that it is not too hard for the target audience. This is a lot harder, especially for less established contests, because they aren't exactly sure who their target audience is.

Whatever the case you should **never use testsolvers as a substitute for test-setters**. If you want help with the test or even individual problems, do not get testsolvers. Testsolvers are supposed to be the final check before you push a contest to production (i.e. print the papers).

3.1 Common mistakes

I recommend you read at the very least. If you work through the problems, you'll have a better idea of what I'm trying to communicate.

These mistakes are why you get people from outside to perform thorough **validity checks**. Some of these are so obvious that it almost seems impossible that an entire group of people missed it.

Here are some mistakes that MAC has made. The bolded parts are the added fixes. All problems mentioned will be spoiled, so if you want to try them first, do not read the rest of this subsection.

Problem 1 (JMC 2020/11) Jack has eight sticks of different lengths in the set $\{1, 2, 3, \ldots, 7, 8\}$. How many nonempty subsets of these eight sticks can Jack choose so the range of the lengths is at most 4 meters?

There is nothing wrong with the formulation of the problem. But to get our answer, all of us did casework on the range, which we wrongly assumed had to be between 1 and 4 (inclusive). It turned out that we forgot about sets with size 1, which had range 0. Thus the correct answer was never on the test! Interestingly enough, none of the people who submitted pointed this error out and we only realized a year later.

There's a number of takeaways to be had here, but I think the most important is you shouldn't solve and write up at the same time, particularly for casework problems. Sometimes the elegance or conciseness of a solution can disguise the fact that it doesn't really cover every case properly.

Problem 2 (MAT 2021/5) Let a, b, and c be **distinct** positive integers with $a + b + c = 10$. Then there exists a quadratic polynomial p satisfying $p(a) = bc$, $p(b) = ac$, and $p(c) = ab$. Find the maximum possible value of $p(10)$.

In short, the problem is solved by doing some algebraic manipulation and ending up with $p(10) = ab + bc + ca$. So

There is a procedure that contestants go through every time they solve a "find maximum" problem. First, they find the upper bound, and second, they verify it exists and makes sense. They also make sure to watch out for redundant information.

As a problem-writer, even if you are **confident** in your answer, you should still substitute your variables into the equation and make sure the result makes sense. For **integers** a, b, c, $ab + bc + ca$ is maximized when $a = 3$, $b = 3$, and $c = 4$ (or some permutation thereof).

But consider the common mistake made in AIME I 2020/11. If $a = b$, then there is a level of redundancy as the conditions reduce to $p(a) = ac$, $p(a) = ac$, and $p(c) = a^2$. Since two points do not uniquely determine a quadratic, $p(10)$ could have been anything. Some contestants did not notice our mistake and put 33, which they were awarded points for. This is why the "distinct" condition was necessary, and the only reason we didn't catch this mistake is because we never checked our construction **from the beginning**.

In both of these problems, our issues were not related to the main idea of the problem — rather, it was because we did not check details closely enough. This is likely because many of us got conditioned into evaluating the merit of a problem when thinking about it as we grow more familiar with our problem base. Outside perspectives are not painted with such bias, which is why it is important to include them.