Mathematical investigations

Carl Joshua Quines

November 28, 2019

Here's a short handout on finding ideas for mathematical investigations, based on a workshop I gave at my high school about a year ago. I am in no way an authority on this. Send any suggestions or complaints to cj@cjquines.com.

1 The average sequence problem

A mathematical investigation is not a single math problem, like the ones we are used to. There's not a single answer, or a single solution, or even a single proof. When investigating something, we're not answering a question that someone else gave us. We're asking our own questions, and try to find answers to them.

Here's a problem, which I like to call the **average sequence problem**.

Problem. A sequence of five non-negative real numbers begins and ends with 0. It has the property that each term, other than the first and last, is equal to one more than the average of the term before and after. What is this sequence?

We can represent the numbers algebraically using variables. Let's say the sequence is $0, x, y, z, 0$, for some numbers x, y, and z. So we can get a system of equations: $x = \frac{1}{2}y + 1$, $y = \frac{1}{2}(x + z) + 1$, and $z = \frac{1}{2}y + 1$. From here, we can solve for x, y, and z to find that the sequence is $0, 3, 4, 3, 0$.

This isn't an investigation on its own. But then we might wonder, what if the sequence had more numbers? Try to solve it in a similar way. For example, if it had three numbers, it's just $0, 1, 0$. If it had four numbers, we can write $0, x, y, 0$, and then we can write $x = \frac{1}{2}y + 1$, $y = \frac{1}{2}x + 1$, and then we can solve this to get $0, 2, 2, 0$.

What if we had six numbers? We could try to write $0, w, x, y, z, 0$ and solve the system of four equations. But we could also observe that all of our previous sequences are the same both forwards and backwards, and we might guess that it's of the form $0, x, y, y, x, 0$ instead. When we solve the system of equations, we should get $0, 4, 6, 6, 4, 0$. (Do the algebra yourself to check that this works!)

And we can try to collect more data, for more and more sequences:

3. $0, 1, 0$.
4. $0, 2, 2, 0$.
5. $0, 3, 4, 3, 0$.
6. $0, 4, 6, 6, 4, 0$.
7. $0, 5, 8, 9, 8, 5, 0$.

Already we begin to notice some patterns. Surprisingly, all the terms in the sequence are integers, and the sequence is the same both forwards and backwards.

Lined up like this, we can look down the second column, and see that it's the numbers $1, 2, 3, 4, 5, \ldots$, and we might guess that this continues for longer lengths of sequences. Similarly, the third column has $2, 4, 6, 8, \ldots$, and the fourth column has $3, 6, 9, \ldots$, so we might guess that this pattern continues for the next sequence. So it's like a multiplication table, except we're looking at its diagonals.

When the sequence has eight numbers, then, we would guess that the sequence is $0, 6, 10, 12, 12, 10, 6, 0$. And amazingly, this turns out to be true! So, if the number of terms is n, we can try to write a formula for the kth term, based on our guesses. And then we can try to prove this algebraically.

We could then be interested in other sequences. For example, what if instead of each number being *one more than* the average of its neighbors, we can ask for each number being *one less than* the average? What kind of sequence would we get then? What if we change *the term before and after* to *the two terms before*, or something else?

Mathematical investigation isn't *just* about answering questions. It's about asking our own questions, it's about experimenting with data and looking for patterns, it's about making guesses and seeing if they're true. It shouldn't feel like you're answering a problem. It should feel like you're **exploring** it.

2 Mathematical investigations

Here's how I would describe a mathematical investigation, which I'll shorten to MI. It consists of three parts:

1. Finding a good problem and asking good questions.
2. Experimenting, exploring, making observations, making guesses, and proving things about this problem.
3. Writing up your results and presenting your work to others.

We can differentiate it from an SIP in several ways. For example, an SIP involves the scientific method: there are hypotheses and experiments. An MI is different, in that you experiment, come up with guesses,[1] and then try to prove them. An SIP usually involves lab work. An MI usually needs involves reading more related literature.

When you read about MIs online, a lot of them focus on the second step. Circle in a Box is the first resource that comes to mind: there are several problems, building up on each other, asking about observations and guesses.

The second step is similar to solving a math problem: the problem is already given to you, and you can apply your problem-solving skills in order to solve it. So to be better in the second step is to become better in solving problems. There are many good resources for this, so I will not focus on that here.

Because of that, this handout will be dedicated to the first part: finding good problems, and asking good questions. I think is the hardest part of doing an MI. What separates a good problem from a bad problem, and how do we come up with good problems?

3 Good problems

I think a good problem should have three different qualities. It is *new*, it is *doable*, and it is *interesting*. Optionally, it can also *have applications*, which is usually a sign of a good problem, but isn't always necessary. Let's talk about each one.

A problem is good if it's *new*. While it's still very possible to do an MI with a problem that someone else has done before, it's often more exciting if the exploration is *new* in some sense. This doesn't have to mean that no one else has done it before. There's probably someone else, for example, who's seen the average sequence problem.

[1] Mathematicians like to call these guesses *conjectures*, and scientists like to call them *hypotheses*, but I think *guesses* is a better word.

I think the criteria for saying something is new is if you can't quickly find it online, and no one you've talked to has seen it before, then it's new. For example, I once investigated if there was a generalization for the sums of squares and sums of cubes formula. I did find one, but this was not a new problem at all.

A problem should be *doable*, because otherwise there's no point to doing it. There should be just the right balance between how approachble it is and how hard it is. If you solve the problem in a few minutes, then it's too easy. One thing you can do is look for a different problem, or make the problem harder, like we did for the average sequence problem.

If you don't make progress after an hour or two, it's too hard, and the best thing to do is probably try a different problem altogether. If it's similar to an unsolved problem, it's also probably too hard. The right balance should be that, every hour you work on it, it should feel like you're discovering something or making progress.

A problem should also be *interesting*. This one kind of overlaps with the previous too problems. For example, the problem "Which numbers can be expressed in exactly one way as sums of consecutive counting numbers?" is more interesting than, say, "Which numbers can be expressed in exactly seven ways as sums of consecutive counting numbers?"[2] This one is more subjective, but if you're not interested in a problem, then it's probably not interesting enough.

4 Where problems come from

Good problems can come from lots of different sources. For one, problems can come from *the real world*. Consider traffic, for example. How can we model traffic mathematically? We can model cars per lane, giving each of them a position, lane, and tracking their velocities, and giving them a probability that they switch into another lane. And then we can run simulations or make predictions and see how much they match the real world. This is what Yi does in *A probability-based model of traffic flow*. How else can we model traffic mathematically? Does this match up with what we expect traffic to look like? Does this help us make recommendations with what to do with traffic?

Another source is *other problems*. One of my friends, Wayne, did an investigation called *Counting Sudoku variants*. It's well-known that there are a certain number of valid Sudoku puzzles, but no work has been done to extend this to other kinds of Sudoku variants, like Sudo-Kurves. Wayne was also really interested in puzzles, so it was a good fit for him too. It's also a nice, doable problem. The paper is pretty approachable too.

Problems can also come from *contest problems*. Our average sequence problem from earlier is inspired by a problem on the PMO 2017 Areas, except we generalized it and asked more questions about it. Another of my friends, Kaan, did an investigation called *On denesting radicals*. It's an investigation of problems like, simplify $\sqrt{3+2\sqrt{2}}$, and which radicals can be simplified in certain ways.

A final source of problems is *other people*. Probably the easiest way to find a problem to work on, although it may not be a good problem, is to ask someone else! Looking at existing papers and reading their recommendations is always a good start. This has the risk that you might end up with a problem that is too hard, or not new. The solution is to modify the problem so that it *does* become a good problem.

[2]This example is taken from Benson et al's *Ways to Think About Mathematics*, which has a good chapter on MI.

5 Modifying problems

A common situation in MI is *almost* having a good problem, except it isn't new, or it's too easy, or it's too hard. But that doesn't mean you should throw away the problem completely. Often, it's possible to modify the problem to get a new one which could be more suitable.

One technique is to make problems *more specific*. For example, *Friedman numbers* are numbers expressible with their own digits, like $347 = 7^3 + 4$. It's not known which numbers are Friedman numbers, and this is a hard problem in general. We can make it more specific, by reducing operations (like, say, allowing only multiplication and addition), or considering only certain numbers (like powers of five). These help make the problem easier, and from there we can try to make it more general again.

Another good trick, especially for problems inspired from the real world, is to look at things from a Philippine context. This is an easy way to make a problem new, since there isn't a lot of work done about Philippine things. *Sungka* is a good example; it has been done before, but maybe consider problems along similar lines.

The opposite of this is making a problem *more general*. This is what we did for the average sequence problem, which was initially too easy, but we can make harder and more interesting by considering it more generally. You can also generalize other things; I read an MI by Aiylam called *Modified Stern–Brocot sequences* which generalizes this cool thing called the *Stern–Brocot tree*, which you should look up if you don't know about it.

If making a problem more specific is going down, and making a problem more general is going up, then another trick is going sideways. That is, making a *variant* on the problem. For Crowdmath, this was essentially what we did. We studied a game called *cops and robbers*, and asked ourselves, what if the cop wasn't restricted to moving a certain way? This lead me and my partner, Espen, to write *Variations of the cop and robber game on graphs.*

The general technique is "describe the problem, take a phrase, and then change it to something else." As another example, consider the game of *nim*. Here's a description.

Problem. There are several piles of stones, and on each turn, you choose a pile and remove a positive number of stones from that pile. The last player who plays wins. Who wins?

We can change "who wins" to "how many ways". More generally, we can change "is there" to "how many". We can change "how many" to "at most what" or "at least what". We can change "find all" to "find one" or "find some". There are lots of different ways to change a phrase!

A final technique is *combining problems.* One of the write-ups I've read combined the ideas of *Collatz sequence* and *random number generation* in something called *Novel application of Collatz-like sequences to cryptographically secure random number generation.* Michael and I also worked on a project where we combined *metric dimension* and *planar graphs* to make *Bounds on metric dimension for families of planar graphs.* So this technique works particularly well: pick one problem, find a related problem, then combine them.

6 Issues

If good problems are new, doable, and interesting, then a bad problem is either done previously, too easy or too hard, or not interesting. By modifying problems strategically, we can fix these issues.

Making a problem interesting is probably the hardest one to fix by modifying it. Probably the easiest way to fix this is to just look for a different problem. Michael and I took two weeks of reading up a *lot* of graph theory in order to decide that we wanted to do a project on metric dimension and planar graphs. It can take a lot of time to find a problem that's interesting enough for you to work on.

Changing the difficulty of a problem is often done by making it more specific or more general. By making a problem more specific, you make it easier. By making a problem more general, you make it harder. Another way to change the difficulty is to change the way things are counted. Instead of counting "how many Xs are there?" you can consider "what's a good upper bound on the number of Xs there are?" or "what's a good lower bound on the number of Xs there are?" To make it harder, you can go from "is there an X?" to "how many Xs are there?"

How do you make a problem new? First, are you *sure* it's not new? Sometimes a problem can feel like it's been done before, because the idea may seem so common, when it actually hasn't. If you're sure that it's not new, then you can make a variant on the problem, or give it a Philippine context, or come up with an application.

Here are some good examples. This first one is taken from Crowdmath 2017, and is known as the *broken stick problem*:

Problem. You have a segment of length one. You choose two points on this segment at random. They divide the segment into three smaller segments. What is the probability that the three smaller segments can be the sides of a triangle?

Some good ways to modify it: what is the expected area of the triangle? What if we require the second point chosen to be on the right of the first? If we pick n points, what is the expected number of triangles? What if we pick points on a circle? What if you turn it into a game?

Here's a well-known game called *chomp*:

Problem. Consider an $m \times n$ rectangular grid of chocolate. Two players alternate moves. In a move, a player picks one cell and removes it from the grid, along with all cells below it and to its right. The person who makes the last move loses. Who wins?

What if we change it to three or more dimensions? What if we have several grids of chocolate? The original problem is too hard, so what if $m = n$ or $m = 2$?[3]

Finally, here's the game of *chopsticks*:

Problem. Two players each extend a finger on each hand. In a turn, one player taps their hand against a hand of the other player, adding the number of fingers to the other player's hand. If after this, the number of fingers is more than five, that hand is dead, and taken out of play. If both of a player's hands are dead, that player loses. Who wins? What is the optimal strategy?

We can change "more than five" to "exactly five", and make addition carry over, which is a well-known variant. Or allow players to split their hand if one hand is dead and the other is alive. Or we can consider what happens if there are more fingers or more hands.[4]

7 Inspiration

It's helpful to know probability and programming for doing applied mathematics. Doing projects in algebra is hard. Geometry is doable if you know a lot of it, and number theory is accessible if you

[3] As a note, the case $m = 3$ was solved in 2002 by a high school student, who won the Siemens Competition for their work.

[4] The original game here is solved; an optimal strategy is online. I don't think any of the proposals I have here have been done before, though.

do a lot of contest math. I think combinatorics is probably the most approachable field to do MI in: graph theory and game theory are great places to start.

Good sources to find inspiration are:

- Tim Gowers wrote a good post on `https://bit.ly/362hiTR` with lots of interesting problems that I think are good starting points.
- *MIT PRIMES*, a math research program for high school students. A lot of them require a lot of background, but there are a few that don't; make sure to read several of them to see which ones. You can view slides from this year's presentations on `https://goo.gl/jQ9hCS`.
- *CROWDMATH*, an online collaborative math research project open to high school students. You can view topics on `https://goo.gl/WiDtpR`: choose a project, click Problems, and make sure to scroll down to check all of them.
- *Intel ISEF*, the largest science fair for high school students. You can search abstracts on `https://goo.gl/bb3XiJ`. You can search for each year's winning entries: there are several winning entries under the mathematics category each year, and you can look up the abstracts.