

Habits of Highly Mathematical People

Jeremy Kun

July 27, 2016

<https://medium.com/@jeremykun/habits-of-highly-mathematical-people-b719df12d15e>

The most common question students have about mathematics is “when will I ever use this?” Many math teachers would probably struggle to give a coherent answer, beyond being very good at following precise directions. They will say “critical thinking” but not much else concrete. Meanwhile, the same teachers must, with a straight face, tell their students that the derivative of arccosine is important. (It goes beyond calculus, in case you were wondering.)

So here is my list. The concrete, unambiguous skills that students of mathematics, when properly taught, will practice and that will come in handy in their lives outside of mathematics. Some of these are technical, the techniques that mathematicians use every day to reason about complex, multi-faceted problems. Others are social, the kinds of emotional intelligence one needs to succeed in a field where you spend almost all of your time understanding nothing. All of them are studied in their purest form in mathematics. The ones I came up with are,

1. Discussing definitions
2. Coming up with counterexamples
3. Being wrong often and admitting it
4. Evaluating many possible consequences of a claim
5. Teasing apart the assumptions underlying an argument
6. Scaling the ladder of abstraction

Discussing Definitions

A primary skill that mathematicians develop is fluidity with definitions. There’s a lot more to this than it sounds at first. What I mean by this is that mathematicians obsess over the best and most useful meaning of every word they use. Mathematicians need logical precision because they work in the realm of things which can be definitively proven or disproven. And if something can be done “definitively,” it must necessarily be definable.

Let me start with a mathematical example first, one which has some relationship to real life, the word “random.” Randomness as a concept has plagued mathematics for much of its recent history because it’s difficult to nail down a *precise* definition of what it means for an event to be random. Statisticians deal with this conundrum by saying that *things* can’t be random, but rather *processes* can be random and you can define the probability of an event happening as a result of the process. That was a very brief overview, but it’s the foundation for pretty much all of statistics.

But it's not the only definition of randomness. Because we intuitively want to say, for example, that flipping a coin and getting 20 heads in a row is "less random" than getting HTHHTHHHTTTHTHHHTHHH. Mathematicians looked at the situation and decided the statistical definition of randomness is not enough, and invented a second definition called *Kolmogorov complexity*. Very roughly, an event is called "Kolmogorov random" if the shortest computer program that produces the event is as long as the description of the event. (This uses a purely mathematical definition of a "computer" that was invented before actual computers, think of Alan Turing) Colloquially, you can imagine that a Kolmogorov random event requires the description of the event itself to be written out, in full, in the source code of the computer program that produces it.

Kolmogorov complexity has grown into a fascinating part of mathematics and computer science, but it's not the end of the story. At the risk of going into too much detail, mathematicians discovered that Kolmogorov complexity could not be calculated for most events. So it's very difficult to apply it to non-theoretical problems. Mathematicians wanted a definition that could describe numbers that *look* random and are random enough for practical purposes, but in fact are highly non-random in the Kolmogorov sense. The result was the current definition of *cryptographically secure randomness*.

Still speaking loosely, randomness in the cryptographic sense means that no *efficient* computer program which tries to distinguish between pseudo-random and truly random events (in the statistical sense) has a significant advantage over a 50/50 guess. This is a guarantee that your numbers are random enough that your enemies won't be able to predict which ones you'll use next, because your enemies are limited to computations that take less time than their lifespan. This is the basis of modern cryptography, and after engineers took the charge, the resulting systems keep our internet communications secure and private.

Okay, so mathematicians have spent a lot of time thinking about definitions which, way down the line, affected how we use math in the real world. That's still not an argument in favor of teaching *everyone* mathematics. How can having everyone think about definitions help them in real life? Now it's time for some realistic examples. The first one comes from Keith Devlin, a mathematician and consultant who was asked to help some US defense agencies improve their intelligence analysis post 9/11. He describes his first presentation to a room full of defense contractors, in which he starts by discussing his thoughts on a precise definition of the word "context." I have chopped up his essay to highlight the important bits.

As I fired up my PowerPoint projection. . . I was sure the group would stop me half way through and ask me to get on the next plane back to San Francisco and not waste any more of their time.

In the event, I never got beyond the first content slide. But not because I was thrown out. Rather, the rest of the session was spent discussing what appeared on that one slide. . . [I was told afterwards,] "That one slide justified having you on the project."

So what had I done? Nothing really — from my perspective. My task was to find a way of analyzing how context influences data analysis and reasoning in highly complex domains involving military, political, and social contexts. I took the oh-so-obvious (to me) first step. I need to write down as precise a mathematical definition as possible of what a context is. It took me a couple of days. . . I can't say I was totally satisfied with it. . . but it was the best I could do, and it did at least give me a firm base on which to start to develop some rudimentary mathematical ideas.

The fairly large group of really smart academics, defense contractors, and senior DoD personnel spent the entire hour of my allotted time discussing that one definition. The discussion brought

out that all the different experts had a different conception of what a context is — a recipe for disaster.

What I had given them was, first, I asked the question “What is a context?” Since each person in the room besides me had a good working concept of context — different ones, as I just noted — they never thought to write down a formal definition. It was not part of what they did. And second, by presenting them with a formal definition, I gave them a common reference point from which they could compare and contrast their own notions. There we had the beginnings of disaster avoidance.

As a mathematician, Devlin did nothing unusual. In fact, the most common question a mathematician has when encountering a new topic is, “What exactly do you mean by that word?”

And while Devlin’s particular example is very specialized, a consultant for defense intelligence, his technique is universal. It’s one of the foundations of the wishy-washy term “critical thinking.” So now an average citizen who might discard the idea of mathematics is listening to the news and hears a politician say, “We have strong evidence of weapons of mass destruction in Iraq.” If they had a good mathematics education they will ask, “What exactly do you mean by *strong evidence* and *weapons of mass destruction*?” And, the crucial follow-up question, does the definition provided justify the proposed response, starting a war? If you don’t understand the definition you can’t make an informed voting decision. (Of course, if you watch the news for entertainment and to be part of a political tribe, the truth is irrelevant)

Everyone has to deal with new definitions, whether it’s a new definition of marriage and gender, or the legal definition of “intent,” “reasonable,” or “privacy.” A seasoned mathematician will readily notice that the government has no useful definition of “religion.” Being able to think critically about definitions is a foundation of informed discourse.

Typical mathematics students begin thinking hard about definitions early in their undergraduate career, and they develop this skill deeper and deeper through graduate school and a research career. A mathematician typically encounters new definitions daily, on both small and big scales, and fluency in discussing definitions is something that would benefit everyone.

Producing examples and counterexamples

Let’s practice our definitions in an informal setting. By “counterexample” I mean an example that shows something breaks or is wrong. For example, the number 5 is a counterexample to the claim that 10 is a prime number, because 5 evenly divides 10.

Mathematicians spend a lot of time coming up with examples and counterexamples to various claims. This point ties very closely to the previous about definitions in two ways.

1. Often, when coming up with a new definition, one has a set of examples and counterexamples that one wants the definition to adhere to. So examples and counterexamples help guide one to build good definitions.
2. When encountering a new existing definition, the first thing every mathematician does is write down examples and counterexamples to help them understand it better.

However, examples and counterexamples go beyond just thinking about definitions. They help one evaluate and make sense of claims. Anyone who has studied mathematics knows this pattern well, and it goes by the name of “conjecture and proof.”

The pattern is as follows. As you’re working on a problem, you study some mathematical object and you write down what you want to prove about that object. This is the conjecture, like an informed (or uninformed) guess about some pattern that governs your object of study. This is followed by the proof, where you try to prove or disprove the claim.

As a bad analogy, maybe you conjecture that the Earth is the center of the universe. Then you try to come up with examples of the object that satisfy the claim. In our solar system example maybe you make a toy model that shows how the Earth *could* be the center of the universe, if the universe were as simple as the toy. Or you could try to go make some measurements involving the sun and moon and come up with evidence that the claim is false, that actually the Earth revolves around the sun. The difference in mathematics is that the “evidence” is a counterexample and it’s only called such if it’s provable. “Evidence” in mathematics is often just a temporary placeholder until the truth is discovered, though for some high profile math problems mathematicians have found nothing but “evidence,” even after hundreds of years of study.

The analogy is also bad because this happens in mathematics on an almost microscopic level. When you’re deep in a project, you’re making new little conjectures every few minutes, and mostly disproving them because you later realize the conjectures were highly uninformed guesses. It’s a turbo-charged scientific process with hundreds of false hypotheses before you arrive at a nice result. The counterexamples you find along the way are like signposts. They guide your future intuition, and once they’re deeply ingrained in your head they help you accept or denounce more complicated conjectures and questions with relative ease.

Again, being able to generate interesting and useful examples and counterexamples is a pillar of useful discourse. If you’ve ever read the transcript of a supreme court hearing, such as this one about the legality of prisoners wearing a beard for religious reasons, you’ll see that most of the arguments are testing examples and counterexamples against previously established legal definitions of “reasonable” and “religion” and “intent.” This mindset has also had countless applications to physics, engineering, and computer science.

But a subtler part of this is that mathematicians, by virtue of having made so many wrong, stupid, and false conjectures over their careers, are the least likely to blindly accept claims based on a strong voice and cultural assumptions. If, as a collective modern society, we agree that people are too willing to believe others (say, politicians, media “experts,” financial talking heads), then studying mathematics is also a fantastic way to build a healthy sense of skepticism. This is just as useful for engineers as it is for plumbers, nurses, and garbage collectors.

Being wrong often and admitting it

Two mathematicians, Isabel and Griffin, are discussing a mathematical claim in front of a blackboard. Isabel thinks the claim is true, and she is vigorously arguing with Griffin, who thinks it is false. Ten minutes later they have completely switched sides, and now Isabel thinks it’s false while Griffin thinks it’s true.

I witness scenarios like this all the time, but only in the context of mathematics. The only reason it can happen is because both mathematicians, regardless of who is actually right, is not only willing to accept

they're wrong, but eager enough to radically switch sides when they see the potential for a flaw in their argument.

Sometimes I will be in a group of four or five people, all discussing a claim, and I'll be the only one who disagrees with the majority. If I provide a good argument, everyone immediately accepts they were wrong without remorse or bad feelings. More often I'm in the majority, being forced to retrace, revise, and polish my beliefs.

Having to do this so often—foster doubt, be wrong, admit it, and start over—distinguishes mathematical discourse even from much praised scientific discourse. There's no p-value hacking or lobbying agenda, and there's very little fame outside of the immediate group of people you're talking to. There's just the search for insight and truth. The mathematical habit is putting your personal pride or embarrassment aside for the sake of insight.

Evaluating many possible consequences of a claim

Scott Aaronson wrote a blog post about the JFK assassination and the conspiracy theories surrounding it. In it, he takes the claim “The JFK assassination was a conspiracy of CIA-size proportions” and evaluates it with straightforward arguments clearly inspired by a mathematical (and theoretical computer science-y) mindset. For example,

10. Almost all JFK conspiracy theories must be false, simply because they're mutually inconsistent. Once you realize that, and start judging the competing conspiracy theories by the standards you'd have to judge them by if at most one could be true, enlightenment may dawn as you find there's nothing in the way of just rejecting all of them.

Another:

12. If the conspiracy is so powerful, why didn't it do something more impressive than just assassinate JFK? Why didn't it rig the election to prevent JFK from becoming President in the first place? (In math, very often the way you discover a bug in your argument is by realizing that the argument gives you more than you originally intended — vastly, implausibly more. Yet every pro-conspiracy argument I've read seems to suffer from the same problem. For example, after successfully killing JFK, did the conspiracy simply disband? Or did it go on to mastermind other assassinations? If it didn't, why not? Isn't pulling the puppet-strings of the world sort of an ongoing proposition? What, if any, are the limits to this conspiracy's power?)

Indeed, exploring the limits of a claim is the mathematician's bread and butter. It's one of the simplest high-level tools one has for evaluating the validity of a claim before going through the details of the argument. Indeed, it can also be used as a litmus test for deciding which arguments are worthwhile to understand in detail.

Sometimes, the limits of an argument result in an even better and more elegant theorem that includes the origin claim. More often, you simply realize you were wrong. So this habit is a less formal variation on being wrong often, and coming up with counterexamples.

Teasing apart the assumptions underlying a claim

One perhaps regrettable feature of mathematics is that is fraught with ambiguity. We like to think of math as rigor incarnate. And I would even argue that's a reasonable idea, once the math has been built upon for a hundred or so years. But even so, the process of doing math—of learning existing ideas or inventing new ideas—is more about human to human communication than stone-cold rigor.

As such, when someone makes a claim in mathematics (out loud), they're usually phrasing it in a way that they hope will convey the core idea to another human as easily as possible. That usually means they're using words in ways one might not expect, especially if the conversation is between two mathematicians with a shared context and you're an outsider trying to understand.

When you face a situation like this in mathematics, you spend a lot of time going back to the basics. You ask questions like, "What do these words mean in this context?" and, "What obvious attempts have already been ruled out, and why?" More deeply, you'd ask, "Why are these particular open questions important?" and, "Where do they see this line of inquiry leading?"

These are the methods a mathematician uses to become informed on a topic. The unifying theme is to isolate each iota of confusion, each assumption underlying a belief or claim. It's decidedly different from the kinds of discourse you see in the world. For example, in the contentious 2016 US presidential election, who has tried to deeply understand Donald Trump's worldview? Most liberals hear, "I'm going to build a wall and make Mexico pay for it," and deride him as insane. Treating the claim mathematically, you would first try to understand where the claim came from. To whom is Trump appealing to? What alternatives to the immigration problem has he ruled out, and why? Why is immigration such an important issue to his supporters, and what assumptions on his part causes him to answer in this way? What is he tapping into that makes his bid successful?

I'm not taking a political stance either way, but rather pointing out that, if a mathematician is in a situation of complete bafflement, teasing apart underlying assumptions is part of the playbook. A large part of what has been pointed out as the "liberal media underestimating Trump" comes from not answering these kinds of questions. Instead they tweet the most misguided quotes they can find in the hopes of popping Trump supporters' filter bubbles. Which, if poll data is to be believed, is not very effective. . .

Scaling the ladder of abstraction

The last habit is a concept I'm borrowing from Bret Victor. It's the idea that when you're reasoning about a problem, there are many different resolutions at which you can think (he uses "rung"). In Victor's example, if you're designing a car-driving algorithm, you can study it at the finest resolution, where you are writing an algorithm and watching it behave in a single execution.

At a higher level, you can control different parameters of the algorithm (and time) with a slider, elevating one algorithm into a family of algorithms that can be tuned. And you can further generalize which parameters and behaviors are tunable to find a way to search through the space of all possible algorithms. As you go, you're looking for high-level patterns that can help you achieve your end goal, designing a great car-driving algorithm back down at the bottom rung (the coarsest resolution).

Mathematicians regularly scale this ladder, particularly in the later stages of graduate school when you need to learn to read a lot of research papers. Here, you don't have time to deeply understand every bit of notation and every claim in every paper (just those papers that are important enough to read in detail). Instead, you develop a ladder of abstraction: the lowest rung is individual definitions and theorems and examples, the next level up is the high-level flow of a paper, next up is a paper in relation to other papers and the broader mathematical context, and above that is the overarching trends in the field, what's considered important and fashionable, etc.

You might start at the lowest rung of the ladder, understanding some examples of a definition to get your bearings, then jump up to the main theorem of the paper and understand how it's perceived as a huge improvement over previous work. You might see they use some technique from the 50's in a field you're not familiar with, but you use that idea as a black box and try to understand the high-level proof of the main theorem, stepping down one rung. Then maybe you go to the open problem section to see what work is left to do, and if it seems enticing enough you can prepare yourself to do that by reading the rest of their paper in detail.

Indeed, mathematicians have to exercise this ladder-scaling muscle whenever they give a talk on their own work. There are many kinds of audience, each of which can appreciate the content of a mathematical idea at a different resolution. Some theorems are best explained as a competitive game in one context, an optimization problem in another, and an analogy to metallurgy in yet another.

Part of the big struggle of mathematics is synthesizing all of the information in all of these ladder rungs into a coherent world-view that you can personally scale up and down at will. Victor tries to make the gymnastic easier by designing a powerful user interface. Mathematicians practice it with whatever techniques they can get their hands on. Either approach, the end goal is valuable.

End Notes

I don't want to imply that developing the habits of highly mathematical people is unambiguously a good thing. In the real world, many of these habits are a double-edged sword. Anyone who has gone through an undergraduate math education has known a person (or been that person) to regularly point out that X statement is not precisely true in the very special case of Y that nobody intended to include as part of the discussion in the first place. It takes a lot of social maturity beyond the bare mathematical discourse to understand when this is appropriate and when it's just annoying.

Moreover, it usually takes a few years beyond introductory proofs to become comfortable with the "always being wrong" part. It can be a huge turn-off to many students who don't have good role models or other students to talk to who are at the same stage of understanding. Math really is an emotional roller coaster.

The point is that adhering too religiously to these principles in every situation will make people think you're an asshole or make you feel like a buffoon more than it will help you. It's knowing when it matters to hold to these principles that allows one to wield mathematical thinking skills like a chef's knife, safely and efficiently slicing up ideas and arguments into their essential forms.