Mathematics as Information Selection by Brian Peterson

Imagine an infinite library of books and parchments, divided into two sections.

The volumes in section one each contain marks and strings of marks. A volume could have nothing but one mark in it, or a series of marks. There is no upper limit on how long any given string of marks may be, which is why the library is necessarily infinite in size. There is also no upper limit on the variety and combinations of scribbles which could occur in these volumes.

Section two of the infinite library is also infinite, and also contains volumes of various scribbles. However, the scribbles in section two are all in English, and each volume in section two specifically cites a volume in section one. The English sentences in section two describe the grammar and rules of application for the scribbles occurring in the volume cited.

Now, for each volume in section one, there is an infinite set of volumes in section two dictating the ways in which one can interpret and play with the scribbles occurring there. The rules in English might define the scribbles as sentences in a possible or real language. The rules might also define the scribbles as axioms of an imaginary mathematical system, or a theorem. Regardless of the interpretation we give to the scribbles, there is no requirement that they make sense to us in any way. For instance, the rules may be so random that we wouldn't want to call the scribbles a language or axiom under the current interpretation. There might also be middle cases where we wouldn't know what we wanted to say.

(A side note: as a matter of logical necessity, every volume in section two is also in section one, and every single one of THOSE has an infinite number of volumes in section two which give it a different interpretation. Makes your head spin.)

Here is an example of how this library could work. One parchment in section one contains a scribble that looks like "+". Of the infinite volumes giving an interpretation to that scribble, one such volume corresponds to our rules for addition, making the scribble a "plus" sign. But yet another corresponds to another possible rule, called "quus". The rule for quus is identical to plus, except when one of the numbers in the formula is higher than 265, in which case the answer is always 5. And so on—there is no limit to the number of rules one could imagine "+" to mean.

A moment's contemplation should reveal that of all the volumes in section two, the vast, vast majority of them are like quus rather than plus. Actually, that's being too generous: quus is too close to being a useful rule for our purposes. It's much easier to imagine a crazy set of interpretations which we couldn't possibly imagine a use for, under any circumstances. But the crazy rules outnumber the sensible or sometimes sensible rules by staggering magnitudes—of course.

The corollary of this is that most of the interpretations in section two cannot be used for communication, cannot be used to make predictions, cannot be used to analyze the stock market, and certainly can't be used to model reality. Others—again, a staggering number of them, but still a much, much smaller set—can be used imperfectly for any of these tasks. Quus is handy, for instance, so long as the numbers you use are all less than 265.

A still smaller set—possibly finite, but still a cosmic number—can be used very accurately for a variety of tasks. Perhaps we would be motivated to say that some of the rules in this set work "perfectly" to make predictions or model reality—whatever *that* is supposed to mean. But compared to the entirety of section two, these volumes are less than a grain of sand on a huge beach.

Now, what kinds of conclusions should we draw from this?

The main moral I wish to point out is derived from something Richard Dawkins once said. To paraphrase, there are infinitely more ways to be dead than alive. In the context of his point, he was referring to the various possible combinations of genes on a chromosome, the vast majority of which lead to phenotypes that would die in the womb or even fail to form a body. Even the set of genotypes which would create a creature that survived birth but not live long is large compared to the set of genotypes leading to a creature which would also successfully reproduce itself.

Natural selection works because of the tautology that replicators which successfully make more of themselves will multiply. And the reason why one genotype works better than an alternative is in the details: what kinds of interactions occur within a body and within an environmental niche, and how do they eventually lead to reproduction?

As we stroll through section two of our Borgesian library, we will find that most of the volumes there aren't very interesting to us. Why? Because they don't constitute an interesting pattern and can't be used for anything, at least very well. They are like a randomly assembled string of genes which will certainly interact with the environment according to the laws of chemistry, but fail to even grow a body. Evolution built animals to recognize patterns, because genomes coding for pattern-recognizing phenotypes have had better success replicating than genomes which don't. The universe, while probably not being deterministic in a strict sense, still has a great degree of reliability, still has patterns in it, or otherwise there could be nothing like life or consciousness. And the tiny, tiny set of volumes in section two of the library which can be used as tools to detect that reliability are what we call mathematics.

Thus, my answer to the question, "Why does math work?" is: because if it didn't, we wouldn't call it math. There is basically a selection bias at work here. You know what a selection bias is: if you are trying to do a scientifically solid survey, you have to make sure that the people you select to study aren't selected on the basis of the very thing you are trying to measure, otherwise your results will be filled with "false positives". Example: you do a study which aims to see if the children raised by aggressive and violent parents tend to be aggressive and violent themselves. If you restrict your survey to children raised by their biological parents, you are guilty of selection bias, which will fail to account for the possibility that aggressive and violent tendencies may come not from the environment, but from the inherited genes.

Something very similar is at work when one notices the effectiveness of mathematics, notices the patterns it produces, and then asks, in wonderment, how this could be. When one does this, one is forgetting to account for the much larger set of rules and symbols which aren't effective, don't produce discoverable patterns, and wouldn't normally even be thought of *for that very reason*. Our minds gravitate towards the rules we call mathematics precisely because they are the only ones which have the very qualities the thread wants explained. And they have those qualities because, out of the infinite number of possible axioms and rule sets, it is literally guaranteed that there exists a logically possible universe to which they can have some kind of useful application. We only pay attention to the ones which appear to help us in this one, the only universe we care about.