CHAPTER 5

Being Constructive in Doing Mathematics

Robin Havea, Sione Paea
The University of the South Pacific

Abstract

The traditional method of doing mathematics is primarily based on classical logic. By doing mathematics constructively, we mean doing mathematics using intuitionistic logic which can be seen as a generalisation of classical logic. Carefully selected examples are used to demonstrate the notion of constructivity in mathematics. The emphasis lies in the importance of the computational content of mathematics.

Introduction

Teaching mathematics at any level is not always an easy exercise. Depending on the subject matter taught and the background of the students in a class, the challenge could be higher than anticipated. It is an ongoing challenge and throughout the years the Pacific Island countries have been continuously invested in improving the skills and expertise of mathematics teachers through long-term and short-term trainings at tertiary institutes in the region and abroad. Research on mathematics education in the Pacific is well documented including recent work done by Begg, Bakalevu, and Havea (2018).

The purpose of this paper is to highlight and introduce the readers (especially mathematics teachers and interested students) to an existing alternative approach to doing mathematics. To be specific, this so-called alternative approach is what we refer to as doing mathematics with a 'twist' in the line of reasoning and is tied to what is widely-known in the realm of mathematics as constructive mathematics which is a very active and highly specialised research field. Constructive mathematics proper can be very technical and may require the sophisticated machinery of mathematical logic to unpack the subtleness and depth of how it works, but we shall keep all technicalities to a minimum and concentrate on presenting accessible demonstrations by means of examples borrowed from senior secondary and undergraduate mathematics. It is our intention that this article will serve as a simple and brief guideline to teachers and students of mathematics so they can easily identify when an argument in mathematics is constructive (or non-constructive), which is not always easy for a non-specialist.

It should be clearly pointed out at the outset that this paper is not intended to be a piece of propaganda nor a suggestion to discredit and abolish the existing approach of how we teach and do mathematics. Furthermore, we give a slightly philosophical view of how one thinks and approaches doing mathematics in an "intuitive" manner.

The authors were brought up as students in the Tonga education system and had taught at secondary schools in Tonga. Based on years of experience through research and constantly questioning the status quo, the authors decided to share their opinions and it is our hope that this article will stimulate teachers and tertiary educators alike to be critical of how mathematics works, because it is a very vibrant discipline but that depends on how one looks at it. Some may think that mathematics is nothing other than routine and textbook discipline and there is no need to be philosophical about it simply because everything is either black or white without any grey area. Needless to say, it is always a healthy approach to explore alternatives as that would open new grounds be it in the epistemological or ontological levels of doing mathematics.

In the next section, we will give a quick tour of constructivism in mathematics. This is a very broad area and we cannot explain every single detail in few paragraphs or pages. As such, where necessary we may direct the reader to relevant sources while concentrating on demonstrations and using carefully chosen examples. Readers may find that having a decent working knowledge of undergraduate mathematics and (but not necessarily) elementary classical logic would be an advantage when reading this paper. However, for readers who are interested in foundation of constructive mathematics, see the works of Beeson (1985), Bishop (1967), Bishop and Bridges (1985), and Troelstra and van Dalen (1988).

What is constructive mathematics?

We try to give a short and less technical answer to the question posed as the title of this section. The literature is rich with information on a full-fledged discussion on what constructivity is all about, particularly in the context of mathematics. There is a widespread interest in constructivism in mathematics shared amongst mathematicians, mathematical logicians, and theoretical computer scientists. The authors recommend the works of Bridges and Dediu (1997) and Bridges and Mines (1984) as very accessible introductory sources of information regarding constructive mathematics.

The terms "traditional mathematics" or "classical mathematics" refer to the usual way of how we do mathematics. In particular, this is the usual mathematics that is based on classical logic. As a refresher, let us look at a particular example where classical logic is used and it is manifested Abstract Algebra as a principle which says that if the product of real numbers a and b is zero, then at least one of them must be zero; that is, if $ab = 0$, then either $a = 0$ or $b = 0$. It is this very reasoning that allows us to solve many quadratic equations of the form $ax^2 + bx + c = 0$, where *a,b,c* are real numbers with $a \neq 0$. In particular, take for instance the product $(x + 2)(x - 3) = 0$, and courtesy of classical logic, we deduce that either $x + 2 = 0$ or $x - 3 = 0$giving us the solution that $x = -2,3$. But what is constructive mathematics? The answer to this question is very extensive and can be very technical. However, we shall adopt the Richman approach that *constructive mathematics is just doing mathematics using intuitionistic logic* (Richman, 1990). Under the umbrella of constructive mathematics, there are three varieties: Brouwer's intuitionistic mathematics, Markov's Russian constructivism, and Bishop's constructive mathematics (Bridges and Richman, 1987). These three varieties have subtle differences but they all share in common the strict interpretation of mathematical existence.

Existence means Computability

So, what does "computability" mean? It simply means that if you want to show that a mathematical object exists (mathematical objects are those that we use in our mathematics including numbers, functions, matrices, continuous functions, differentiable functions, to name a few), it means that you should be able to compute or construct that mathematical object. In other words, if you claim that the object in question exists, then you should be able to provide a routine or an algorithm that anyone (including a machine that is directed by instructions in codes) can follow and systematically find or at least approximate that object to whatever precision required. Anyone that is familiar with writing a computer programme may have a deeper insight right away because writing a piece of code is actually giving instructions to the machine to follow in order to successfully complete the required task. It is a very intuitive way of computing and establishing the existence of the mathematical object in question. As far as logic is concerned, intuitionistic logic allows you to avoid non-constructive decisions and gives you an opportunity to be very "honest" with your mathematics; that is, if you claim that an object exists, then you should be able to construct it. On the contrary, classical logic allows you to make certain moves in your reasoning where you could prove that a mathematical object exists without even showing how to compute it, which is an issue that is very central to computing.

Most of us were taught and brought up learning (and eventually teaching) mathematics based on classical logic even if we weren't aware or told about it. There are several basic principles in classical logic that our mathematical reasoning relies on. One of the most notable principle is the Law of Excluded Middle (LEM) which simply states that any mathematical statement P is either true or false; in logical notation we write

$$P \text{ or } \neg P,$$

where ¬P stands for "not P'. LEM is a tautology which means that it is always true no matter what meaning we associate with the propositional variable P. This is because we can't expect a statement to be half-true or half-false; there is no grey area or anything in between true and false. You can only have one or the other; you cannot have both which is recalled in the following truth-value table.

P	**¬*P***	***P* or ¬*P***
True	False	True
False	True	True

(Note that a statement that involves the disjunction "or" is true when at least one of the disjuncts is true.) Although it is trivially acceptable in classical logic, LEM cannot be proved using intuitionistic logic and is, therefore, highly regarded as non-constructive. As such, any mathematical statement that is equivalent to or implied by LEM is considered highly non-constructive and, hence, not acceptable in constructive mathematics. But why does LEM allow one to be non-constructive? LEM allows you to "cheat" when you argue that an object x exists. Suppose we want to prove that x exists. An application of LEM allows us to argue that because LEM asserts that

"x exists" or "x does not exist",

we only have two alternatives to worry about. So instead of showing (directly) that "x exists" holds, we (indirectly) show that if we could rule out "x does not exist" that is enough to establish that the other alternative "x exists" must be true! How would you rule out "x does not exist"? We assume that "x does not exist" holds, and based on that assumption we ended up with a contradiction and, hence, we reject "x does not exist" simply because it is contradictory. Therefore, we conclude that the other alternative, "x exists", must be the case. For more information on LEM, see the works of Bridges and Richman (1987) and Havea (2005).

Our next Example, 1, is a variation of a well-known classical theorem and we have rephrased it to demonstrate the power and application of LEM. Notice that it is an existential statement because it purports the non-existence of two integers.

Example 1. Consider the following statement which is trivially true in classical mathematics.

There are no integers p and q, with $q \neq 0$, such that $\sqrt{2} = \frac{p}{q}$.

Let us see how we translate this so that we could apply LEM. The statement clearly claims that there are no integers p and q such that the conclusion followed. The other alternative is that there are integers p and q such that the conclusion followed. In short, we have

"There *are no* integers p and q …" **or** "There *are* integers p and q…"

If we could rule out the alternative "There are integers p and q ...", then, by courtesy of LEM, we have to prove that the other alternative, "There are no integers p and q …" is true! We argue as follows. Suppose that there are integers p and q, with $q \neq 0$, such that $\sqrt{2} = \frac{p}{q}$, and we further assume that the rational expression $\frac{p}{q}$ is in its lowest and simplest form; that is, the numbers p and q have no common factor other than 1 which means that the greatest common divisor is 1 and we write $\gcd(p, q) = 1$. Then

$$\sqrt{2} = \frac{p}{q} \Leftrightarrow 2 = \frac{p^2}{q^2} \Leftrightarrow p^2 = 2q^2 \Leftrightarrow p^2 \text{ is even} \Leftrightarrow p \text{ is even} \Leftrightarrow p = 2k,$$

for some integer k. Furthermore,

$$p^2 = 2q^2 \Leftrightarrow (2k)^2 = 2q^2 \Leftrightarrow q^2 = 2k^2 \Leftrightarrow q^2 \text{ is even} \Leftrightarrow q \text{ is even} \Leftrightarrow q = 2l,$$

for some integer l. Hence,

$\sqrt{2} = \frac{p}{q} = \frac{2k}{2l} \Rightarrow \gcd(p, q) \neq 1,$ a contradiction!

Here we ended up with a contradiction because of the assumption that the existence of p and q such that $\sqrt{2} = \frac{p}{q}$. Therefore, we conclude that there are no integers p and q such that $\sqrt{2} = \frac{p}{q}$.

Apart from the finer details of the argument in the preceding example, the point to notice is the general form of the argument which is allowed by LEM. What we have done in the example is rule out one alternative, so concluding that it must be the other alternative that is the case.

As mentioned earlier, there are other principles that are constructively unacceptable because they allow us to make certain moves and reasoning in our mathematics which are highly non-constructive; these are statements that have to do with mathematical existence. To be specific, because of the strict interpretation of "existence" as "computable", we need to be more elaborative and precise about what we assume and expect to get at the end. Some classical non-constructive principles could be converted into constructive principles by adding (or deleting) some assumptions to (or from) the classical versions. The reader is invited to see the works of Bridges and Dediu (1997), Bridges and Richman (1987), and Havea (2005) for more extensive discussions of a considerable number of well-known non-constructive principles in classical mathematics.

We list a few and commonly well-known principles below.

Axiom of Choice (AC). If A and B are sets and S is a nonempty subset of $A \times B$ such that for each $a \in A$ there exists $b \in B$ with $(a, b) \in A \times B$, then there exists a function $f: A \to B$ (called the *choice function*) such that $(a, f(a)) \in S$ for all $a \in A$.

There is an interesting relationship between AC and LEM whereby Goodman and Myhill (1978) showed that AC implies LEM. To be specific, under the assumption that AC is true, one could deduce that LEM is also true, and because LEM is non-constructive, hence, AC is also non-constructive.

Recall that a binary sequence (a_n) is simply a sequence that contains 0s and 1s.

Limited Principle of Omniscience (LPO). If (a_n) is a binary sequence, then either $a_n = 0$ for all n or else there exists n such that $a_n = 1$.

Weak LPO (WLPO). For any binary sequence (a_n), either $a_n = 0$ for each n, or else it is impossible that $a_n = 0$ for all n.

Lesser LPO (LLPO). If (a_n) is a binary sequence containing at most one term equal to 1, then either $a_{2n} = 0$ for each n, or else $a_{2n+1} = 0$ for all n.

There is a clear indication that Brouwer was very suspicious of the constructive status of the above omniscience principles although he used different names for LPO and LLPO (Bishop, 1970). For more detailed discussion of these principles, the reader is advised to see the work of Bridges and Richman (1987).

Examples and Demonstrations

In this section we look at some carefully chosen examples to demonstrate the constructive and non-constructive challenges that we encounter even in some very well-known theorems. It should be pointed out that when we say that a theorem is non-constructive it does not mean that we completely reject such a theorem outright but, rather, we look and apply or add necessary conditions so that we have a constructive version of that theorem. When doing so, we are also interested in checking to see what is the best we can hope for in a constructive setting by means of using Brouwerian examples; see the works of Bridges and Richman (1987) for detailed discussion and the role of Brouwerian examples in constructive mathematics.

Example 2. This example is due to Bishop (Bishop, 1972; Goodman and Myhill, 1972) showing how LEM is used to prove the well-known classical theorem:

There exist irrational numbers r and s such that r^s is rational.

We argue as follows. Consider the real number $\sqrt{2}^{\sqrt{2}}$. By LEM, either $\sqrt{2}^{\sqrt{2}}$ is rational or $\sqrt{2}^{\sqrt{2}}$ is irrational (using the fact that any real number is either rational or irrational). In the former case, if $\sqrt{2}^{\sqrt{2}}$ is rational, then we simply take $r = s = \sqrt{2}$ and we are done! In the latter case, if $\sqrt{2}^{\sqrt{2}}$ is irrational, then we take $r = \sqrt{2}^{\sqrt{2}}$ and $s = \sqrt{2}$ in which case both r and s are irrational and, hence, $r^s = \left(\sqrt{2}^{\sqrt{2}}\right)^{\sqrt{2}} = \sqrt{2}^2 = 2$ which is rational!

Carefully studying this proof reveals that under LEM, we were able to prove the statement without even showing how to find the irrational numbers r and s. It is very clever, neat, and classically an acceptable argument. However, if we are interested in the numerical content of the statement, then this proof is not helpful at all. A constructive proof of the statement would enable us to compute the two irrational numbers or even approximate them to any precision that pleases us. Thus, from a constructive standpoint, the proof is non-constructive. Why is it non-constructive? Simply because we appealed to LEM and it led us to such conclusion.

Example 3. In this example, we consider the classical *Intermediate Value Theorem*:

If $f: [a, b] \to \mathbb{R}$ is a continuous mapping such that $f(a) < 0$ and $f(b) > 0$, then there exists $c \in (a, b)$ such that $f(c) = 0$.

One way of proving this is using the interval-halving technique, more commonly known to a secondary school student as the *bisection method*. Following a similar presentation given by Bridges and Vîtă (2006), without loss of generality, suppose that the interval in question is (a,b)=(0,1). We proceed in the following manner. Consider $f\left(\frac{1}{2}\right)$:

if $f\left(\frac{1}{2}\right) = 0$, then we take $c = \frac{1}{2}$ and stop the process

if $f\left(\frac{1}{2}\right) > 0$, then f satisfies the hypotheses of the theorem with $a = 0$ and $b = \frac{1}{2}$

if $f\left(\frac{1}{2}\right) < 0$, then f satisfies the hypotheses of the theorem with $a = \frac{1}{2}$ and $b = 1$.

In either of the last two cases above, we are guaranteed two things: either the process terminates and produces the required result, or it continues forever, thereby producing a descending sequence of compact intervals whose unique point of intersection is the required zero. What you should notice in the process is that this is a purely algorithmic proof! It gives you a step-by-step procedure of how to locate or at least approximate the root c.

There is an interesting phenomenon that a typical computer programmer may have noticed or find a bit confusing when implementing the algorithm in the preceding example. We demonstrate this phenomenon in our next example.

Example 4. In this demonstration, we showcase how a computer may get confused and register an incorrect answer based on the bisection method discussed in Example 3. Again, we follow and use the argument used by Bridges and Vîtă (2006) in the following way. Suppose we are implementing the algorithm on a machine that works with 50-bit precision. Consider the following cubic function defined on the closed interval [0,1]:

$$f(x) = \left(x - \frac{3}{4}\right)\left(x - \frac{1}{2}\right)^2 - 2^{-51}.$$

Using MATHEMATICA, it is easy to see that $f(x)$ satisfies the hypotheses of the Intermediate Value Theorem,

$$f(0) = -\frac{3}{16} - 2^{-51} = -\frac{422212465065985}{2251799813685248} = -0.1875 < 0,$$

$$f(1) = \frac{1}{16} - 2^{-51} = \frac{140737488355327}{2251799813685248} = 0.0625 > 0.$$

Carrying out the interval-halving technique leads to f having a zero between 0 and 1. Now, let us look at the evaluation at the midpoint, where $x = \frac{1}{2}$:

$$f\left(\frac{1}{2}\right) = -2^{-51} = -\frac{1}{2251799813685248} = -4.44089\times10^{-16}.$$

Since our computer's floating-point representation of $f\left(\frac{1}{2}\right)$ is 0 (Floating-point numbers are numbers that involve floating decimal points and are mostly used when dealing with very small and large magnitudes. In engineering and most technical calculations, we use floating points to represent non-integer numbers with a certain fixed number of decimal points. This is very useful when talking about both small and large magnitudes; for example, with a fixed number of decimal points, we can speak of the diameter of a single hair, or the distance between two galaxies in the universe), we are faced with the problem of underflow where the machine registers a value (like, for example, 10^{-10}) that is very close to 0 as simply 0 which is not correct. However, the only (real) zero (or x-intercept) of f is actually $\frac{3}{4}$ which is quite a distance away from $\frac{1}{2}$.

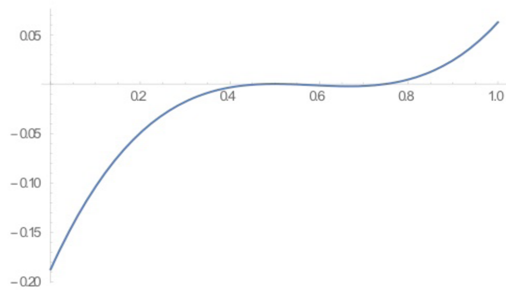

Figure 1. Graph of $f(x) = \left(x - \frac{3}{4}\right)\left(x - \frac{1}{2}\right)^2 - 2^{-51}$

The graph of f in Figure 1 shows the critical places where the machine mistakenly "thought" the zero might be which is 0.5 but actually it is at 0.75.

The preceding example is a demonstration of how mathematical existence is dealt with in a realistic and practical manner. The Intermediate Value Theorem guarantees the existence of a zero under certain favourable conditions but finding it using the bisection method can be problematic depending on the type of function that we are working with.

Our next example is another demonstration of how classical logic is used in a very appealing manner. The theorem is well-known and commonly taught in a typical undergraduate course in discrete mathematics.

Example 5. The following theorem was proved by Euclid using a very clever contradiction argument.

There are infinitely many primes.

The statement is about the existence of an infinite set of primes. So, we proceed by assuming that there is a "finite" set $\{p_1, p_2, \ldots, p_n\}$ of primes. We define the integer

$$p = (p_1 \times p_2 \times \cdots \times p_n) + 1.$$

Clearly p is greater than the smallest prime 2. Thus p has prime factors and note that p itself may be prime. Since the primes p_k, for $k = 1, \ldots, n$, are not factors (i.e. divisors) of p, whatever the prime factors of p are must be distinct from each p_k. So, we have here another prime that is not in the set $\{p_1, p_2, \ldots, p_n\}$ of primes; that is, a contradiction. Therefore, there are infinitely many primes.

Looking at Euclid's proof from a constructive point of view, there are at least two ways in which it can be criticised. Following the ideas of Bridges and Vîtă (2006), we argue as follows.

1. We were able to construct a new prime out of an already known finite number of primes. This is perfectly fine and algorithmic but the unnecessary contradiction makes the computational side of the argument; that is, emphasis is on the derivation of the contradiction making the algorithmic process less significant.

2. We witness an application of some form of LEM which is very subtle. To be specific, the argument rests on the negativity of the statement about "infinite sets" which assumes that

> A set is infinite if and only if it is contradictory that it be finite.

It is worth pointing out that there is positivity in Euclid's proof as far as constructivity is concerned and is associated with and hinted at by the observation that one should be able to construct a new prime out of already known primes. Generally, it emphasises the possibility that if we start with a finite subset B of A, then we can compute an element of A that is distinct from each element of B; in the preceding example, take A to be the set of all primes and B the finite subset of primes.

Apart from abstract analysis, there is a wide range of examples and demonstrations of how constructive mathematics is carried out over the real number line. Interested readers having a background in classical real analysis are advised to look into the comprehensive work of Bridges (1994).

Conclusion

Constructive mathematics is honest mathematics! If you claim that an object exists, then you should be able to demonstrate how to actually construct, or compute, that object. It is all about the strict interpretation of mathematical existence as simply constructability. Mathematical existence in the classical sense can be seen as being "ideal" or even "virtual" whereas in constructive mathematics it is more "realistic". In a more practical context, if you claim that an object exists, then you should be able to provide an algorithm or set of instructions where anyone (or even a programmable machine) can follow and find (or construct) the object in question to whatever precision you please.

So, why do we need to do mathematics constructively? It all depends on what you want to do. If you are interested in the computational content of your mathematics, then constructive mathematics, or doing mathematics using intuitionistic logic, provides a suitable platform and framework. We learn and teach mathematics primarily based on classical logic and at times we tend to neglect the very heart of doing mathematics which has to do with being able to compute mathematical objects. Further, we should be able to avoid certain decisions that would lead us to non-constructive moves but that can only be dictated by the very logical principles that we use. In particular, intuitionistic logic provides a better alternative as far as computability is concerned.

Anyone with a slight interest in the foundation of mathematics would welcome the varieties and different approaches in mathematics. It would be completely misleading for the authors to present as a case where constructive mathematics is the answer to everything and that we must abandon the classical approach – no, not at all. We believe that it is equally important and relevant for teachers and lecturers of mathematics to have at least an appreciation of the many approaches to doing mathematics. Of course, the traditional way of doing mathematics using classical logic will be the common approach in all aspects of teaching and doing mathematics but there are certain limitations when it comes to computational content. Having said that, in order to appreciate constructive mathematics, it is very important to have a full appreciation of classical mathematics.

We end this note with a challenge to all teachers and lovers of mathematics – How can we teach our students to think algorithmically?

References

Beeson, M. J. (1985). Foundation of Constructive Mathematics. Ergebnisse der Mathematik und ihrer Grenzgebiete 6. Berlin: Springer-Verlag.

Begg, A., Bakalevu, S., & Havea, R. (2018). Mathematics Education in the South Pacific In J. Mack & B. Vogeli (Eds.), Mathematics and its teaching in the Pacific-Asia Region Series on Mathematics Education 15, World Scientific, 2018.

Bishop, E. (1967). Foundations of Constructive Analysis. McGraw-Hill, New York.

Bishop, E. (1970). Mathematics as a numerical language. In 1970 Intuitionism and Proof Theory (Proc. Conf., Buffalo, N.Y., 1968) (pp. 53-71). North-Holland, Amsterdam.

Bishop, E. (1972). Aspects of constructivism. Las Cruses, New Mexico: New Mexico State University.

Bishop, E., & Bridges, D. (1985). Constructive Analysis. Grundlehren der Mathematischen Wissenschaften 279. Berlin: Springer-Verlag.

Bridges, D. (1994). A constructive look at the real number line. In P. Ehrlich (Ed.), Real Numbers, Generalizations of the Reals, and Theories of Continua (special issue of Synthèse (pp. 29-92). Amsterdam: Kluwer Academic Publishers.

Bridges, D., & Dediu, L. (1997). Paradise lost or paradise regained? EATCS Bulletin, 63, 141-155.

Bridges, D., & Mines, R. (1984). What is constructive mathematics? Math. Intelligencer 6(4), 403-410.

Bridges, D., & Richman, F. (1987). Varieties of Constructive Mathematics. Cambridge: Cambridge University Press.

Bridges, D. S., & Vîţă, L. S. (2006). Techniques of Constructive Mathematics. Universitext: Springer.

Goodman, N. D., & Myhill, J. (1972). The formalization of Bishop's constructive mathematics. In Toposes, Algebraic Geometry and Logic (Conference at Dalhousie University, Halifax, 1971), Lecture Notes in Mathematics 274 (pp. 83-96). Berlin: Springer-Verlag.

Goodman, N. D., & Myhill, J. (1978). Choice implies excluded middle. Zeit. Math. Logik und Grundlagen der Math, 24, 461.

Havea, R. (2005). Mathematics without the law of the excluded middle. In I. C. Campbell & E. Coxon (Eds.), Polynesian Paradox: Essays in Honour of Prof. 'I. Futa Helu. Suva: Institute of Pacific Studies, University of the South Pacific.

Richman, F. (1990). Intuitionism as generalization. Philosophia Mathematica s2-5, 124–128.

Troelstra, A. S., & van Dalen, D. (1988). Constructivism in Mathematics:

An Introduction (Vol. 2). Elsevier, Amsterdam.