PHIL 236.00
Professor Douglas Marshall
Final Paper

Final Paper – A Discussion about Explanatory Proofs

During the course of studying mathematical philosophy, it's important to not allow a single proof to be sufficiently demonstrative. In fact, we've discussed this in class – as mathematicians we don't just create proofs in one category but include proofs from number theory, topology and probability. It is so important to aim higher than demonstrating that a fact *is*; we want to attain real knowledge by demonstrating the *why*. This is not a radical view of mathematics – as a part of the standard view understandability is our goal. So in this essay, I will describe different metrics of a proof's explanatory degree. Furthermore, I will discuss how these metrics are critical.

Take a look at the Goldbach Conjecture – the original statement describes that in the set of integers **Z**, every integer greater than or equal to 2 is the sum of three primes.[1] Primes of course are the irreducible integers. While computer-generated proofs exist they do not explain the essence of the conjecture, and so it remains a conjecture. The purpose of an explanatory proof is to justify a statement (make it believable) and also to provide historical context for a deeper understanding of the theorem. This is why proof-analysis creates progression.

Mathematical proofs are different from scientific demonstrations, and this difference is fundamentally fragile. Where Aristotle says that "all Fs are Gs" isn't true if there aren't any Fs, this illuminates the question. Constructivists would say that there are Fs because mathematical objects are created. But because of this, facts about mathematical objects are up to human

---

[1] Weisstein, Goldbach Conjecture. MathWorld.

interpretation and the vagaries of hidden lemmas. In an explanatory proof, certain things must be eliminated including vagueness for the sake of explicit clarity. Their values lies not in their correctness but in the clear line of thought; the element of explanatory proof is that of conviction.

From Lakatos, Definition *P* tells us objects are polyhedra if and only if they are a part of the system of polygons for which the equation V – E + F = 2 holds true.[2] This is the clearest example of the definitions discussed by the Teacher which were created by various students in their largely vain attempt. The picture-frame, the urchin, star-polyhedron and the simple cylinder can be accepted as polyhedra due to their vague definition. The principal problem I see in Lakatos is the difficulty of constructing an explanatory proof that the Euler characteristic is 2. Innumerable proofs are presented via concept-stretching, monster-barring and namely the illumination of hidden lemmas which bar global counterexamples from being valid refutations. In terms of explanatory value, monster-barring theorems are not improvements on the naïve conjecture as Gamma says.[3] This is mainly because they continue to work within the framework of the refuter. They don't explain any more why the theorem is true because they simply narrow already existing concepts and generate complexity.

Furthermore, Omega concludes that freaks and monsters which violate our expectations of what a polyhedron should be are actually good – he says that one may get to a relative saturation point of a particular thread of deductive guessing but then find a revolutionary idea with more explanatory power and depth.[4] What he means by this is that the method of proofs

---

[2] Lakatos, *Proofs and Refutations,* p. 17

[3] Ibid., p. 91

[4] Lakatos, p. 102

and refutations is very helpful to making proofs explanatory. It means that proofs are never over and have infinite possibilities in terms of increasing depth, explanation and rigor.

Certainly, the mathematical community has placed a great emphasis on guarding proofs from counterexamples. So by showing that no proof is truly safe from refutation, Lakatos implies that they are all safe; in fact, criticism has a beneficial impact. And Lakatos' book allows us to see how important it is to make explanatory justifications for any conjecture before simply publishing it on some inductive basis which is not based on a representative sample.

At this point I think it would be helpful to define some traditional metrics of what makes a proof explanatory for our analysis. In the 1970s, Steiner suggested that a proof is explanatory if and only if it makes reference to a characterizing property of an entity which is an integral part of the theorem.[5] We also have Kitcher who says that explanatory proofs are measured by context-less unification. Specifically, Kitcher's definition is that proofs can only be explanatory if they are part of a small collection of argumentative techniques in the framework of deductive reasoning.[6] Additionally, Dawson describes proofs as informal arguments which (to paraphrase, if they are explanatory) will not only demonstrate that a mathematical statement is true but also answer the question of why.[7]

Knowing how to do a proof is not the same as attaining knowledge of truth or facts, a fact to which Sellar alludes. If this were so then as he says, the statement that ducks know how to swim would be metaphorically the same as the statement that they know that water supports

---

[5] Ernest, *The Philosophy of Mathematics Education Today,* p. 85

[6] Ibid., p. 85

[7] Ibid., p. 88

them.[8] Similarly, Eric Watkins tells us that the sensory impression of a red square is itself neither red nor square.[9] And what Watkins basically says in his article *Kant, Sellars and the Myth of the Given* is that being able to manipulate symbolic representations of mathematical objects is not necessarily the same as the acquisition of real knowledge. Furthermore, he says that sensations from external world actually constitute a constraint on our cognition.[10] What this implies to me is that just as sensations in the external world are our key input for deduced knowledge about the world around us (based on attaching linguistic concepts to the sensory inputs which we receive), in mathematics we have representations of mathematical objects in linguistic form which are meaningless without our assigning them value. So the deductivist approach may not be completely explanatory because of the fact that the way in which we do math puts some constraints on our acquisition of knowledge. Merely being able to write the mathematical language is not the same as using it in order to transmit real insights.

Now as Kitcher alludes to, the most explanatory proofs make the least amount of assumptions and in making the methods of mathematical proof simple they are able to explain the most theorems in the least number of ways. So for Kitcher, the common universal element between explanatory proofs is the fact that they incorporate the least number of hidden lemmas. I believe that the nature of being explanatory is subjective; this means that we can measure whether or not proofs are explanatory but we have to do so in a heuristic way which takes each of these metrics into account.

---

[8] Sellars, *Philosophy and the Scientific Image of Man,* p. 2

[9] Watkins, *Kant, Sellars, and the Myth of the Given,* p. 320

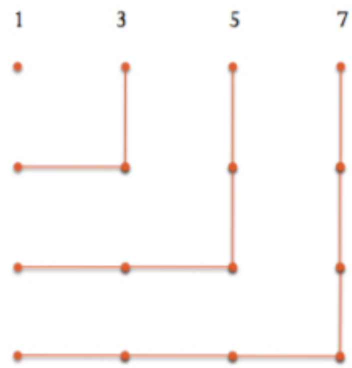
[10] Ibid., p. 323

I'd like to provide an example of two different mathematical proofs by induction, the first of which is non-explanatory and the second of which is. In doing so I hope to distinguish between algebraic induction (which is not explanatory) and geometric induction (which is explanatory).

The statement we are proving is that the sum of the first n odd positive integers is n^2. We can prove this by induction in two ways.[11] First, the base case is that the first odd positive integer is 1 and this is equal to 1^2. Then by the inductive hypothesis we can assume that the sum of the first n odd positive integers is n^2 and then look at the n+2 case. Here, we have that the sum of the first n + 1 odd positive integers is $n^2 + (2(n+1) – 1)$. Clearly, we see that this is equal to $n^2 + 2n + 2 – 1 = n^2 + 2n + 1 = (n + 1)^2$ and therefore the proof should be complete by the principle of mathematical induction. We can argue that this is not explanatory by the metrics outlined earlier. By Kitcher's method of unifying patterns of argument and making the least number of assumptions, it is explanatory to the extent that it invokes the accepted method of mathematical induction with the least number of lemmas. However, it is clear that it doesn't really reference a characterizing property of positive odd numbers other than the fact that the n^{th} odd number is equal to $2n – 1$. Furthermore, by Dawson's metric it doesn't conceptualize why the sum of the first n odd numbers is equal to n^2; it simply relies on algebraic manipulations and the principle of mathematical induction which is valid to prove that this is true but isn't explanatory. In general, proof by contradiction allows us to indirectly infer theorems' validity yet (like counterexamples) does not really constitute a proof. Examples by themselves do not prove statements.

---

[11] The two proofs I describe are from Cariani's *Mathematical Induction and Explanatory Value in Mathematics,* p. 9-10



There exists a geometric proof on the other hand, which is pretty self-explanatory and is generalizable. It makes use of geometric induction which is clear from the diagram. And by omitting the step of converting thoughts into words it allows us to immediately grasp the concept and thus is the more explanatory out of the two.

I would concede that while inductive proofs may have an explanatory aspect called into question, direct proofs are generally much more explanatory. This is because they are generalizable, another criterion which I would judge to be a sufficient metric of being explanatory. The proof I present to you now is simply based on definitional arguments, yet does not include any of the monster-adjustment which Lakatos describes.[12] Consider the proof that if R and S are rings with identity then the direct product R x S (call it T) is also a ring with identity. Right away we can say that R and S are rings with identity ⇒ R and S contain identities for the operation ·, identities which we can denote by 1. Without loss of generality we can let (r, s) ∈ T be arbitrary. Then (1, 1) · (r, s) = (1 · r, 1 · s) = (r, s) because 1 is the identity for R and for S. Then this becomes (r · 1, s · 1) = (r, s) · (1, 1). And because (1, 1) · (r, s) = (r, s) = (r, s) · (1, 1) and we were given any arbitrary (r, s) in T ⇒ for all a ∈ T, (1, 1) · a = a = a · (1, 1) ⇒ (1, 1) is the identity for T under the operation · ⇒ T is a ring with identity, we have established incontrovertibly that the direct product of two rings with identity is also a ring with identity. This proof relies exclusively on properties of rings with identities, and so it appeals to a characterizing property while explaining precisely why the property is true without excessive use of previous theorems. In fact, it relies exclusively on the definitions rather than previously

[12] Shahriari, *Algebra in Action: A Course in Groups, Rings, and Fields,* Exercise 15.2.3 (c), p. 322

proved results so it should prove to be an example of a highly explanatory proof within the realm of group theory and its accepted definitions.

Also, I'd like to add some more requisites and lack thereof for a proof to be explanatory. As Resnik says in his article *Proof as a Source of Truth*, there's a clear distinction between the psychological and logical powers of a proof.[13] In order for a proof to be explanatory it need not necessarily be logically correct, and as Resnik says proofs cannot be separated from their context. Their validity depends very much on the historical and social element. Now I would like to expand upon Resnik's mention of a proof that for each finite collection of primes there is a greater prime.[14] His proof that there is no greatest prime is quite sufficient, however as he writes the directness of a proof is correlated positively with its explanatory value. The proof goes something like this. Say we are given an exhaustive collection of finite primes $p_1, p_2, \ldots, p_k$ such that $p_1 < p_2 < \ldots < p_k$. Now, let's first talk about the old proof. The old proof says that $E = (p_1 \times p_2 \times \ldots \times p_k) + 1$ is not divisible by primes, and yet we can divide E until we get prime factors which divide E with no remainder yet aren't in the set of all primes (a contradiction).[15] Proofs by contradiction are definitely convincing but they don't refer to a characterizing property of the prime numbers. Namely, let's look at E and notice that because E is not evenly divisible by any of the finite prime numbers described its only divisors are going to be 1 and itself. So in the set of integers E is going to be a prime number. Furthermore, it's greater than the other primes so we have described that there is always a greater prime than $p_1, p_2, \ldots, p_k$. This proof is

---

[13] Resnik, *Proof as a Source of Truth,* p. 15

[14] Ibid., p. 15

[15] Ibid., p. 14

more explanatory than the initial one because it refers to the characterizing quality of prime numbers (that they are divisible only 1 and by themselves) instead of the fact that all integers can be factored into prime factors. Instead of arriving at a contradiction (which shows a refutation of some other principle which follows from the original principle) it more directly demonstrates that there is always a greater prime, and in this fashion we can proceed inductively (adding E to the prime numbers discovered) to show that the set of all prime numbers is infinite. Our method may not find all prime numbers (as we discussed in class) but it definitely will generate an infinite set of prime numbers. Both proofs are equally valid but we prefer the more extensive one because it explains things more directly.

And as Kuhn says in his article *The Nature and Necessity of Scientific Revolutions,* the discovery of life on the moon for example would be totally destructive of the existing paradigm which has established until this point that life cannot exist in those conditions.[16] But whether or not our discovery of life occurs on the moon or elsewhere, it still expands our knowledge base in totality. This is generalizable to all sciences including math. In mathematics, any discovery (whether or not it causes an irreversible paradigm shift) causes our explanatory power to increase and develop.

As Aristotle says in Chapter 9 of *Posterior Analytics* it is pretty clear that there are limitations to Bryson's proof of the squaring of the circle.[17] From my perspective, I would say there are limited things which we can infer about the circle from its relationship to different squares because they are fundamentally different in terms of their essence. We can certainly

---

[16] Kuhn, *The Nature and Necessity of Scientific Revolutions,* p. 89

[17] Aristotle, *Posterior Analytics,* p. 13

make comparisons, but they are not equivalent. Thus, it would be very hard for Bryson to infer anything beyond the general range of the area of a circle by comparing circles and squares.

I'd like to quickly discuss some other proofs and assess the extent to which they can answer the question of why the theorem is true. A week ago, we discussed the uses of the pigeonhole principle which itself has many different kinds of proofs. The Pigeonhole Principle, just as a reminder, states that if we place more than n pigeons in n pigeonholes then there exists a hole with more than one pigeon. Of course, this property doesn't apply to pigeons only, and this seems fairly intuitive in all senses. I want to argue that certain kinds of induction which do not rely on the inductive principle (that is, proofs by induction which are constructive rather than circular in logic) can be very explanatory. For instance, we can look at a single pigeonhole and insert a pigeon. If we add any more it is clear that there will be more than one pigeon contained in the single pigeonhole. And for every additional pigeonhole we add we will see that adding more than one pigeon will result in the same condition. And now we have already populated each pigeonhole and as a result there is no more room. This kind of induction is quite explanatory because it demonstrates that given n pigeonholes each of which contains a maximum of one pigeon, we're never going to have a number of pigeons which is greater than n.

Another proof involves trying to construct an injective function from a set of n + 1 pigeons to a set of n pigeonholes. We can arrive at a contradiction that two pigeons must be the same, otherwise the function is not injective. This proof is less explanatory in my opinion because it depends on prior (not innate in the sense of Kant) knowledge of the injective property. So I conclude that there is no universal measure which can determine the explanatory qualities of a proof – each proof has its own subjective (to historical context as well as to the interpretation of the reader) degree.

Similarly, we can prove that a minimally spanning tree T has $|V_T| - 1 = |E_T|$ where V is the set of vertices and E is the set of edges in the tree by contradiction (any less edges disconnects the graph and any more creates a cycle, making T not spanning or not minimal). However, this is clearly not as explanatory as a more direct proof by inductive construction would be. Although it appeals to the definition of a tree, it also incorporates other concepts like connectedness and cycles which are required to complete the proof. As we discussed in class, the simplest proofs are often better at explaining.

Consider the following explanation that a minimally spanning tree is constructed by choosing a single starting vertex. At this stage we have $|V_T| = 1$, $|E_T| = 0$. Each time we add an additional vertex we only add one more edge, thus maintaining the difference $|V_T| - 1 = |E_T|$. I would conclude that direct proofs (and direct forms of induction) and proofs by contrapositive are generally the most explanatory, because proofs by algebraic induction and contrapositive can be convincing but require us to accept other established facts about mathematical logic. Proofs by cases tend to establish theorems neither with certainty nor with any explanation as to why the theorem holds. An inability to find a counter-example to the four-color theorem based on a computerized search is certainly convincing but statistically invalid because there are many different types of graphs with which we are not familiar.

In addition, I agree with Popper's notion that there is no upper bound on the power of explanation (that is, theorems can always be explained better).[18] I also agree with Popper on the

[18] Popper, *The Logic of Scientific Discovery,* p. 452

idea that deduction presupposes the property of being explanatory.[19] It is clear from reading Lakatos that heuristic proof-analysis is critical; the logical positivism about which he writes in the Introduction to *Proofs and Refutations* is truly detrimental because it is clear that the explanatory value of a theorem cannot be pinpointed or deduced incontrovertibly by rigid rules of mathematical logic; it exists outside the realm of deduction but is critical to proof-analysis, the study of proofs (all proofs by nature tend to be formalist, and those which are truly non-deductive are not really proofs per the logical positivists' dogma) and the mathematical philosophy because math is discovered. We may believe certain statements which as of now are unverified yet seem plausible, and the discussion of these conjectures is indispensable.

[19] Ibid., p. 40

References:

Aristotle and Barnes, J (editor). (1975). *Aristotle: Posterior Analytics*. Second Edition. Oxford: Clarendon Press.

Cariani, F. (2016). Mathematical Induction and Explanatory Value in Mathematics.

Ernest, P. (2018). *The Philosophy of Mathematics Education Today*. First edition. Springer International Publishing AG.

Kuhn, T. (1998). The Nature and Necessity of Scientific Revolutions. *Philosophy of Science; The Central Issues*, (1), p. 86-101.

Popper, K. (2005). *The Logic of Scientific Discovery*. Second edition. London and New York: Taylor & Francis e-Library.

Resnik, M. and Detlefsen, M (editor). (1992). Proof as a Source of Truth. *Proof and Knowledge in Mathematics*, p. 6-32.

Sellars, W. (1963). *Philosophy and the Scientific Image of Man*. First edition. London: Routledge & Kegan Paul Ltd, p. 1-40.

Shahriari, S. (2017). *Algebra in Action: A Course in Groups, Rings, and Fields*. 1st edition. Rhode Island: American Mathematical Society, p. 322.

Watkins, E. (2012). Kant, Sellars, and the Myth of the Given. *The Philosophical Forum, Inc*, p. 311-326.

Weisstein, E. (2019). *Goldbach Conjecture*. From MathWorld - A Wolfram Web Resource. <https://mathworld.wolfram.com/GoldbachConjecture.html>