not arbitrary with respect to the conjecture under discussion, rather the criticism is that the proposed new meanings are arbitrary in general). Certainly avoidance of arbitrariness in the latter sense is desirable, but again Walton offers no response other than rejection.

11.3.1. Argument from Arbitrariness of a Verbal Classification.

Premise If an argument, Arg occurs in a context of dialogue that requires a non-arbitrary definition for a key property F that occurs in Arg, and F is defined in an arbitrary way in Arg, then Arg ought to be rejected as deficient.

Premise Arg occurs in a context of dialogue that requires a non-arbitrary definition for a key property F that occurs in Arg.

Premise Some property F that occurs in Arg is defined in a way that is arbitrary.

Conclusion Therefore, Arg ought to be rejected as deficient.

Critical Questions:

1. Does the context of dialogue in which Arg occurs require a non-arbitrary definition of F?
2. Is some property F that occurs in Arg defined in an arbitrary way?
3. Why is arbitrariness of definition a problem in the context of dialogue in which Arg was advanced? (cf. [54, p. 320])

11.4. Methodology in mathematics. The section on monster-barring in [23] is interspersed with heated debate on methodology: whether mathematicians should study typical, ordinary examples and generate interesting and useful theorems about these, or focus on boundary cases, studying mathematics in its "critical state, in fever, in passion" [23, p. 23]. The teacher concludes that monster-barring is *not* a valid method; indeed, it is presented as the least sophisticated method after the method of surrender. The main criticisms are that monster-barrers are anti-falsificationists who defend a conjecture at any cost, which makes the conjecture deteriorate into meaningless dogma, and that the method is *ad hoc*, since the border between monsters and counterexamples is done in fits and starts. "Using this method one can eliminate any counterexample to the original conjecture by a sometimes deft but always *ad hoc* redefinition of the polyhedron, of its defining terms, or the defining terms of its defining terms. We should somehow treat counterexamples with more respect, and not stubbornly exorcise them by dubbing them monsters" (*Teacher*, in [23, p. 23]). The Duhem-Quine thesis [46], that a scientific theory cannot be tested in isolation since a test of one theory always depends on other assumptions or hypotheses, is also relevant to this discussion. We cannot falsify a conjecture, rather we can show that a collection of assumptions, concepts, counterexample and conjecture is internally inconsistent. The choice then arises as to which of the collection we reject. Monster-barrers would argue that we should reject the counterexample and certain concept definitions, and retain the conjecture, whereas critics of monster-barring might argue that we should reject the conjecture under discussion.

12. Lakatos's conjecture-changing methods

12.1. Lakatos's exception-barring methods. Lakatos's exception-barring methods, piecemeal exclusion and strategic withdrawal, are presented early on in [23, pp. 24–30]; a sign that they are considered unsophisticated and inferior to later methods. However, they are worth considering in some detail, because their legacy can be seen throughout the book: his later methods of lemma incorporation are essentially proof-oriented versions of exception-barring. This is made explicit by Lakatos in his distinction between "primitive exception-barring", which makes no use of a proof, and lemma-incorporation:

> The best exception-barrers do a careful analysis of the prohibited area... in fact your method [the method of lemma-incorporation] is, in this respect, a limiting case of the exception-barring method [...].
>
> [23, p. 37]

The exception barring methods target propositions which are "hopefully false" (as opposed to ones which are true — which should be accepted, or "hopelessly false" — which should be rejected — p. 26, *ibid.*), *i.e.*, propositions which hold for most but not all examples considered. They consist in determining the domain of validity for a claim, and result in a modified version of a conjecture (or claim). This may be done by identifying specific counterexamples, generalising from these to form a class of exceptions and excluding this class in a Toulminian fashion (Toulmin's rebuttal seems to be written with exactly this situation in mind; stating the conditions under which a claim does not hold). A second approach is to "withdraw to safety", by generalising from some specific supporting examples and limiting the claim to only this class.

Both methods are criticised as resulting in a "chaotic position" (p. 25, *ibid.*), since we can neither know when all exceptions have been identified and safely excluded, nor even whether some may lie within our supposedly safe stronghold. The latter method is further criticised as even if the boundaries are drawn so narrowly as to be certain of the truth of our proposition, if there are further supporting examples which lie outwith the boundaries then we have not succeeded in our task: the object of the exercise is to draw the *right* boundaries — as wide as possible and no wider — rather the *safest* boundaries, which result in a dull conservatism and a failure to illuminate.

12.2. A logical representation of exception-barring. Logically, these methods can be expressed as follows: given conjecture $\forall x(Px \rightarrow Qx)$, a set of counter (or negative) examples Neg such that $\forall n \in Neg(Pn \wedge \neg Qn)$, and a set of positive examples Pos such that $\forall p \in Pos(Pp \wedge Qp)$;

(*i*) find a concept C_1 such that for all n, C_1n, and for all p, $\neg C_1p$, and modify the conjecture to $\forall x((Px \wedge \neg C_1x) \rightarrow Qx)$ (piecemeal exclusion), and

(*ii*) find a concept C_2 such that for all p, C_2p, and for all n, $\neg C_2n$, and modify the conjecture to $\forall x((Px \wedge C_2x) \rightarrow Qx)$ (strategic withdrawal).

These two methods are logically equivalent if $\forall x(Px \rightarrow (C_1x \vee C_2x))$. For instance, given the property of being an integer (P), the property of being a number with an even number of divisors (Q), and an initial conjecture: $\forall x(Px \rightarrow Qx)$, this could be modified by examining counterexamples (1,4,9,16, ...), generalising from them to the property of being a square number (C_1), and excluding this class from the domain to get $\forall x((Px \wedge \neg C_1x) \rightarrow Qx)$ (all integers except squares have an even number of divisors). Alternatively, the same conjecture could be formed by using strategic withdrawal to generalise from supporting examples (2,3,5,6,...) to the property of being a non-square number (C_2), resulting in the modified conjecture $\forall x((Px \wedge C_2x) \rightarrow Qx)$ (all non-square integers have an even number of divisors). This possibility for logical equivalence is not noted by Lakatos, maybe because none of his examples of exception-barring result in two concepts where one is the complement of the other. The fact that this is an equivalence theorem in mathematics, *i.e.*, $\forall x((Px \wedge C_2x) \leftrightarrow Qx)$ (x is a non-square integer if and only if it has an even number of divisors) also raises the question of how these methods apply to types of conjecture other than implications.

Again, this is not considered by Lakatos, and is also often overlooked in argumentation literature: all examples of Toulmin's claims are of the type Pa, and his warrants are generally implications $\forall x(Px \to Qx)$.[11]

12.3. A mediæval treatment of defeasibility. Syncategorematic terms were a preoccupation of many scholastic logicians from the twelfth century until the eclipse of traditional logic in the early modern era. These are terms which have no non-logical sense but which affect the logical form of a proposition. Reflection on these terms led to a doctrine of *exponible* propositions: propositions which could be analysed as conjunctions or disjunctions of simpler propositions. The scholastic analysis of *exceptive* syncategorematic terms, such as "besides", foreshadows the twentieth century debate over defeasibility which we have been discussing. Exceptive propositions were analysed as possessing at least two parts. For example, Walter Burley, writing in the early fourteenth century, states that:

> [E]ach exceptive has two exponents, an affirmative one and a negative one. For example, 'Every man besides Socrates runs' is expounded like this: 'Every man other than Socrates runs and Socrates does not run'. And 'No man besides Socrates runs' is expounded like this: 'No man other than Socrates runs and Socrates runs'. [7, p. 256]

(Some authors would add a third exponent 'Socrates is a man' to each pair [5, p. 235].) However, on most modern analyses, in each case only the first exponent is a logical consequence of the original statement. The second exponent would be seen as an implicature at best.

This historical detour may clarify the relationship between Toulmin's rebuttals and Pollock's rebutters and undercutters. Toulmin's rebuttals are introduced with an exceptive term, typically "unless". If this is understood as modifying the warrant in his layout (something on which there is less than complete consensus — see §12.5), then the warrant would be an exceptive proposition. For instance, the warrant in the Harry from Bermuda example would read 'anyone born in Bermuda will generally be British, unless his parents were aliens'. This would characteristically translate into first order logic as $\forall x((Bx \wedge \neg Ax) \to Sx)$, where Bx = 'x was born in Bermuda', Sx = 'x is a British subject', and Ax = 'x's parents were aliens'. However, on Burley's reading, the warrant would

[11] Gasteren [16] addresses this imbalance, focussing on equivalence conjectures, and showing how analysis of the form of a conjecture can guide the design of its proof.

translate as $\forall x((Bx \wedge \neg Ax) \rightarrow Sx) \wedge \forall x(Ax \rightarrow \neg Sx)$, which entails $\forall x(Bx \rightarrow (\neg Ax \leftrightarrow Sx))$. Hence, on the narrower, modern reading of exceptive propositions, Toulmin's rebuttals are undercutters, but on the looser, scholastic reading they conjoin an undercutter and a rebutter. Toulmin preserves this ambiguity by eschewing such explicit formalisation, but most of the examples of rebuttals in [52] and [53] provide reasons to reject the corresponding claim, and would thus support the scholastic reading. Nonetheless, Toulmin does introduce some examples of rebuttals most plausibly understood on the modern reading, perhaps including Harry's parents being aliens, since that need not have prevented him from acquiring British nationality by naturalisation.

We can also fruitfully consider Lakatos's exception-barring methods in terms of exceptives, since the proposition resulting from piecemeal exclusion will be exceptive, whether affirmative, 'all polyhedra except polyhedra with cavities are Eulerian', or equivalently, negative 'no polyhedra except those with cavities are non-Eulerian'. In the discussion in [23] Lakatos implies that strategic withdrawal results in a negative exceptive 'no polyhedra but convex polyhedra are Eulerian'. (Again, this is equivalent to an affirmative exceptive, 'all polyhedra except convex polyhedra are non-Eulerian'). The discussion on boundaries suggests that despite the representation of the conjecture as an implication, the goal is to find an equivalence conjecture, *i.e.*, 'all polyhedra are Eulerian if and only if convex', which suggests that Lakatos was tacitly following the scholastic reading of his exceptive proposition. This is only explicitly discussed in the context of strategic withdrawal: "Could you have withdrawn too radically, leaving lots of Eulerian polyhedra outside the walls?" (p. 28, *ibid.*). Indeed, when the piecemeal-excluded conjecture is modified to *for all polyhedra that have no cavities (like the pair of nested cubes) and tunnels (like the picture-frame)* $V - E + F = 2$ (p. 26 *ibid.*), there is an obvious "counterexample" of a pyramid with a tunnel in it, for which $V - E + F$ *is* equal to 2 (shown in figure 9), of which no mention is made in [23]. Thus, in contrast with strategic withdrawal, Lakatos's description of piecemeal exclusion seems to focus on the logical implication, and thereby the modern reading of the exceptive.

12.4. Lakatos and Walton. In terms of Walton's argumentation schemes, his *Argumentation Scheme for Argument from an Exceptional Case* and *Argumentation Scheme for Argument from Precedent*, shown below, are both pertinent to Lakatos's piecemeal exclusion.

Figure 9. A pyramid with a tunnel in it. $V - E + F = 13 - 20 + 9 = 2$, and thus it is a supporting example for Euler's conjecture, however it is barred by the piecemeal-exclusion move to exclude all polyhedra with tunnels.

12.4.1. Argument from an Exceptional Case.

Premise For all x, if the case of x is an exception, then the established rule does not apply to the case of x.
Premise The case of a is an exception.
Conclusion Therefore, a need not do A.

Critical Questions:

1. Is the case of a a recognized type of exception?
2. If it is not a recognized case, can evidence why the established rule does not apply to it be given?
3. If it is a borderline case, can comparable cases be cited?

(cf. [54, p. 344])

Here we have an exception utilising step.

12.4.2. Argument from Precedent.

Premise The existing rule says that for all x, if x has property F then x has property G.
Premise But in this case C, a has property F, but does not have property G.
Conclusion Therefore, the existing rule must be changed, qualified, or given up, or a new rule must be introduced to cover case C.

Critical Questions:

1. Does the existing rule really say that for all x, if x has property F then x has property G?
2. Is case C legitimate, or can it be explained away as not really in violation of the existing rule?

3. Is case C an already recognized type of exception that does not require any change in the existing rule? (cf. [54, p. 344])

And here an exception establishing step. One can perhaps see how (an adaptation of) these schemes could be deployed to set up a system of defeasible rules, procedures or definitions. But we are not yet that close to Lakatos.

12.5. Revisiting Toulmin's rebuttal. Although perhaps seemingly obvious, there is room for ambiguity regarding whether Toulmin's rebuttal is intended to be general or specific. As discussed in §8.2, all of Toulmin's examples in [52, Ch. 3] concern claims about specific facts. This suggests to us that Toulmin intended the data and the claim to be specific facts and the warrant the (general) bridge between the two. In subsequent work however, such as [1, 2, 3] (discussed in §8), other authors have shown that Toulmin's layout can also be used for general claims. Toulmin's focus on the specific raises the question of whether he intended the rebuttal to be specific or general (this is equivalent to asking whether the rebuttal rebuts the claim or the warrant): in his examples it is specific. Consider his discussion of Anne's hair colour [52, p. 117]: based on the datum that *Anne is one of Jack's sisters*, and warrant that *any sister of Jack's may be taken to have red hair* (which itself has the backing that *all his sisters have previously been observed to have red hair*) we may conclude that, unless the rebuttal that *Anne has dyed her hair/gone white/lost her hair* holds, subject to the qualifier *presumably*, *Anne now has red hair*. Here, Anne *is named* in the rebuttal: this is one specific counterexample. We may have expected a general rebuttal, such as *any sister of Jack's, who has dyed her hair/gone white/lost her hair*; thus repairing the general warrant to: *any sister of Jack's who has not dyed her hair/gone white/lost her hair may be taken to have red hair*. The situation is particularly interesting in Toulmin's general cases, in which he uses a pronoun that might be reasonably taken to refer to the general *or* the specific case. For example, in "we can presumably claim that Harry is British, since anyone born in Bermuda will generally be British . . . , unless **his** parents were aliens", does "his" refer to Harry or to anyone born in Bermuda?

We see an analogue in the method of piecemeal exclusion, although not noted by Lakatos. In cases where there are few counterexamples (or only one), it may be preferable to exclude these by name, rather than generalising to a class and excluding that. For instance, in the conjecture *all primes are odd*, given the counterexample 2 we could generalise to the (singleton) class of smallest primes and modify our conjecture to *all primes except for the smallest one are odd*. However, possibly to avoid the extra inference step, this theorem is usually expressed as *all primes except 2 are odd*. Goldbach's conjecture, that *every even number except 2 can be expressed as the sum of two primes* provides a second example, in which the obvious classification "smallest even number" is passed over for the simpler "2". Indeed, examples of this type in mathematics usually, if not always, involve the smallest member of the domain of a claim, such as the trivial group, the empty set, the singleton graph, etc. A theorem which held for all primes except 287, or all groups except for the real numbers under addition, would be curious in the extreme and would certainly merit further investigation. In [36] one of us called this method "counterexample-barring".

13. Lakatos's proof-changing methods

13.1. Three types of lemma-incorporation. Lemma-incorporation is triggered by a counterexample, and comes in three flavours, depending on the type of counterexample. A *global counterexample* is a counterexample to the main conjecture, and a *local counterexample* is a counterexample to one of the proof steps (Lakatos calls these *lemmas*). In argumentation terminology (compare, §4), global counterexamples are rebutters for *any* argument that P, and local counterexamples are undercutters for (*some* step of) *some* argument that P. That is, a rebutter for an argument that P is (or implies the existence of) a derivation of $\neg P$; an undercutter for an argument that P is (or implies the existence of) a defeating answer for one of the critical questions in the derivation of P. Lakatos considers counterexamples which are both global and local, or one and not the other. He suggests that the first step, when faced with a counterexample, is to determine which type it is. If it is both global and local, *i.e.*, there is a problem with both the argument and the conclusion, then one should use strategic withdrawal to modify the conjecture, but — crucially — the domain to which we withdraw must be generated by the problematic proof step S_i. That is, we create a concept "all X which satisfy proof step S_i" and then limit the claim to this concept (in the Euler example it is "simple polyhedra", that is, polyhedra for which step 1 of Cauchy's proof can be performed [p. 34 *ibid.*]). This is a combination of both rebuttal and undercutting, where both of Pollock's defeaters come into play. If the counterexample is local but not global, *i.e.*, the conclusion may still be correct but the reasons for believing it are flawed, then standard piecemeal exclusion is used on the problematic proof step. That is, objects which support and refute the problematic proof step S_i are examined, a concept found which is true of all the supporting examples and false for all the counterexamples and the claim made in S_i then limited to all objects for which this concept holds (in the Euler example it is "boundary triangles", that is, step 3 of Cauchy's proof is modified to *removal of any* boundary *triangle preserves the Euler characteristic* [p. 11 *ibid.*]). The global conjecture is left unchanged. This is an example of just undercutting, which is captured by Pollock's undercutting defeater. If the counterexample is global but not local, *i.e.*, there is a problem with the conclusion but no obvious flaw in the reasoning which led to the conclusion, then Lakatos suggests searching for a hidden assumption in the proof, then modifying the culprit proof step and the global conjecture by making the assumption an explicit condition. This is a case of simple rebuttal, *i.e.* rebuttal without undercutting. Logically, we would have a global conjecture $\forall x(Px \to Qx)$, a set of counter (or negative) examples Neg such that $\forall n \in Neg\ (Pn \wedge \neg Qn)$, and a set of positive examples Pos such that $\forall p \in Pos(Pp \wedge Qp)$, and a set of conjectures which constitute proof steps for the global conjecture:

$$\begin{gathered}\forall x(P_1 x \to Q_1 x)\\ \forall x(P_2 x \to Q_2 x)\\ \vdots\\ \forall x(P_n x \to Q_n x)\end{gathered}$$

For one of the proof steps it may be possible to find a concept C_3 such that for all n, $C_3 n$, and for all p, $\neg C_3 p$. In this case the proof step should be modified to $\forall x(P_i x \to (Q_i x \wedge C_3 x))$ (a third form of exception-barring?): we would then have a global and local counterexample, which can be dealt with as discussed above.

13.2. A new, fourth type, of lemma-incorporation? While Lakatos does deal with the surprising case of a counterexample which appears to be global but not local, he does not consider the fourth scenario, in which we have an object which is neither a global nor local counterexample but still seems surprising and is not what was intended by the original claim. In this case we would expect to find a hidden assumption in the *global* conjecture. For instance, consider the beaker in Figure 10(a). This is Eulerian, since it has zero vertices, two edges and four faces, and, if we accept that the cylinder is not a counterexample to any of the proof steps then this is not either. However, it is (probably) not what was meant by the initial claim. Thus, while not a counterexample in the traditional sense, the beaker would constitute a fourth strange object for Lakatos (the same argument holds for the (topologically equivalent) bowler hat, in Figure 10(b).[12] The modification in this case could involve finding a hidden assumption in the global conjecture, making it explicit and then performing exception-barring (for instance, the criterion that the polyhedron must be a closed system). An alternative, since Lakatos would presumably prefer the method to involve the proof, would be to perform a version of global-only lemma-incorporation, in which an assumption hidden in the proof is identified, made explicit, and incorporated into the global conjecture (for instance, the criterion that it is possible to remove only one face at a time).

(a) The beaker (b) The bowler hat

Figure 10. Both the beaker and the bowler hat are Eulerian, since $V - E + F = 0 - 2 + 4 = 2$. However, they are (presumably) not the sort of objects intended to be covered by Euler's conjecture.

13.3. Rebutting without a known undercutter in mathematics. Lakatos's second case study follows the development of Cauchy's proof [8] of the conjecture that 'the limit of any convergent series of continuous functions is itself continuous' [23, App. 1]. He shows how hidden lemma-

[12] Figure 10(a) is cropped from a photo by AndreyTTL; Figure 10(b) is a photo by Fozrocket.

incorporation and the third kind of counterexample was used to repair the faulty conjecture and proof. The counterexample, found by Fourier is:

$$\cos x - \frac{1}{3}\cos 3x + \frac{1}{5}\cos 5x - \dots$$

which converges to the step function. Lakatos credits Seidel with inventing the method of proofs and refutations, arguing that he discovered it at the same time as he discovered the proof-generated concept of uniform convergence, and that he was aware of the importance of his method [23, p. 136]. He quotes Seidel:

> Starting from the certainty just achieved, that the theorem is not universally valid, and hence that its proof must rest on some extra hidden assumption, one then subjects the proof to a more detailed analysis. It is not very difficult to discover the hidden hypothesis.
>
> [51, p. 383], on [23, p. 136]

Given that in this example it took twenty six years after the proof was published (and thirty five years after Fourier's series became known) to identify the hidden assumption of uniform convergence in the proof, we may question how easy it is to discover a hidden hypothesis. Lakatos, however, suggested that the main reason for such a long gap, and the willingness of mathematicians to ignore the contradiction, was a commitment on the part of mathematicians to Euclidean methodology: deductive argument was considered infallible and therefore there was no place for proof analysis.[13] Lakatos's example of the cylinder in his main case study appears to have been concocted by himself, as a plausible

[13] This example is complicated: Cauchy's claim is generally regarded as obviously false, and the clarification of what was wrong is usually taken to be part of the more rigorous formalisation of the calculus developed by Weierstrass, involving the invention of the concept of *uniform convergence* (the historical route is sketched in [19, pp. 213–17]). The episode was treated by Lakatos in two different ways. Cauchy claimed that the function defined by pointwise limits of continuous functions must be continuous [8]. In fact, what we take to be counterexamples were already known when Cauchy made his claim, as Lakatos points out in his earlier analysis of the evolution of the ideas involved [23, App. 1]. After discussion with Abraham Robinson, Lakatos then saw that there was an alternative analysis. Robinson was the founder of non-standard analysis, which found a way to rehabilitate talk of infinitesimals, for example, positive numbers greater than zero, but less than any "standard" real number (see [48], first edition 1966). Lakatos's alternative reading, presented in [24], is that Cauchy's proof was correct, but that his notion of (real) number was different from that adopted by mainstream analysis to this day.

nineteenth century addition to the discussion, to show how an important method may have applied to his Eulerian example. (While [23] is a rational reconstruction, much of what the "students" discuss has some historical correlation, as evidenced by the numerous footnotes. The cylinder, first introduced by Gamma [23, p. 22], has no associated historical footnote.[14]) The main discussion occurs on pp. 42–50, in which Alpha puts forward various 'implicit assumptions' in Cauchy's proof, which the cylinder undercuts. Other students, especially Gamma, argue that such assumptions have only just been invented, specifically for the purpose of being violated by the cylinder. In argumentation terms, Alpha argues that any rebutter *must* also be an undercutter, although it may undercut a premise which is initially hidden or missing: proof analysis will make implicit undercutting explicit. Gamma, on the other hand, argues that there *are* cases of mathematical counterexamples which are genuinely global but not local: that there can be rebutters without undercutters. The discussion culminates in the *Principle of Retransmission of Falsity* (p. 47); the criterion that a proof-analysis is valid and the corresponding mathematical theorem true if and only if there is no third type of counterexample, *i.e.*, if all global counterexamples must also be local (even if they locally violate a lemma which is not yet explicit in the proof).

Lakatos's identification of counterexamples which are global but not local shows that it is, in some way, possible in mathematics to rebut without a known undercutter (see ([1, p. 298] and [38, pp. 22–24] for previous discussion on this). However, his recommendation for dealing with such counterexamples suggests that he thought that it was possible to see such entities as undercutters as well as rebutters, with their role being to make a misleading proof more precise. It is an interesting question as to whether this situation, in which any rebutter *must be* an undercutter, either explicitly or implicitly, is specific to mathematics. Certainly, in other areas of thought, there are many examples which appear to rebut without undercutting. The British journalist Bruce Anderson provides a rather nice example of this:

---

[14] An additional clue that the cylinder may have been conceived by Lakatos is that when Gamma first appears [23, p. 8] it is to raise questions about Cauchy's proof of the Euler conjecture, to which Lakatos appends the following footnote: "The class is a rather advanced one. To Cauchy, Poinsot, and to many other excellent mathematicians in the nineteenth century these questions did not occur." Of course, it doesn't follow that *all* of Gamma's interjections are anachronistic, but it is interesting that he's introduced in this way.

> A generation ago, Reg Prentice was minister of education in a Labour government. He wrote a paper making the case for increased spending on nursery schools. The Chief Secretary was Joel Barnett, a tough little Mancunian accountant. 'Reg,' he said, 'brilliant paper. It was so moving; reminded me why I came into the Labour movement. Your arguments are unanswerable. And the answer's no.' [4]

We contend that Barnett's response would be untenable in mathematics (but perhaps only mathematics).

Our analysis of Lakatos's proof-changing methods from a Waltonian perspective might resolve the rebutting-without-undercutting problem (see §13 below).

13.4. Lakatos's method of proofs and refutations. Lakatos's method of *proofs and refutations* extends his *lemma-incorporation*, in which he suggests using the proof steps to find counterexamples (by looking for objects which would violate them). For any counterexamples found, one should determine whether they are local or global counterexamples, and then perform lemma-incorporation. This may correspond to using Walton's defeating questions to suggest premisses which may be mistaken or incomplete.

13.5. Extending Walton's schemes. In sections 13.5.3 to 13.5.5 we suggest extensions to Walton's schemes, which reflect Lakatos's proof-changing methods. In order to understand them, it is useful to suggest a new procedural scheme first (§13.5.1). We can then instantiate this to Cauchy's proof plan (§13.5.2). The procedural proof outlined by Cauchy may seem a rather curious beast, in that it starts off with one type of object (a polyhedron), performs various operations on the object (which change it from a polyhedron to a network, to a triangulated graph), shows that the resulting object, now of a different type, has a certain property, and concludes from this that the original object must have another property (Eulerianness). However, this is a common proof strategy in mathematics: arguably, this is *the* signature proof strategy of contemporary mathematics, perhaps beginning with Galois theory. Marquis [34] discusses this general approach, and gives a worked example considering K-theory:

> In algebraic K-theory, one starts with the category of rings and ring homomorphisms, then associates to each ring a commutative semigroup

> and to each ring homomorphism a semigroup homomorphism, then applies the above functor to end up in the category of abelian groups and group homomorphisms, and thus obtains information about some of the structural properties of these rings. [34, p. 257]

Another example of this type of reasoning is the application of Galois theory to demonstrate the impossibility of constructions in Euclidean geometry. Indeed, Galois's development of such techniques is arguably a critical turning point in mathematical method.

We can see these as examples of analogical reasoning in mathematics, in which there is a known, well-defined, bijective mapping between a source and target domain. This is a specific and important type of analogy: inferences made in a target domain can be used to infer certain knowledge about the source domain. An example of a known, well-defined, bijective mapping is a symmetrical transformation such as scaling, reflection, and rotation. For instance, analogical inferences about a kite-shaped quadrilateral with its apex pointing upwards and a kite-shaped quadrilateral with its apex pointing right, where the mapping between the two shapes is a rotation of 90°, will be rigourous. While analogical reasoning is ubiquitous, mathematics is probably a unique domain in that there exist known, well-defined, bijective mappings, and therefore analogical inference in this context is a rigourous form of reasoning: in other domains there is often a 'verification' stage after the analogical inference (clearly this can also occur in mathematical domains in which the mappings are not well-defined). As such, we can see our following scheme, *Suggested Argumentation Scheme for a Procedural Argument*, as a subcategory of Walton's *Argument from Analogy* [54, p. 315] mentioned in §9. Our scheme differs from Walton's schemes in a few regards. Most importantly, it is procedural. Secondly, the scheme contains multiple inferences: Walton's schemes in [54, Ch. 9] describe just one inference (although some of his examples in [54, Ch. 10] describe multiple inferences, such as *Argument from Guilt by Association*, in which premise 2 follows from premise 1). Thirdly, because of the second point, the order of the premisses matters. One possibility to bring our scheme closer to a Waltonian scheme would be to split it into several schemes: we present it as a single scheme for ease of understanding.

13.5.1. Suggested Argumentation Scheme for a Procedural Argument.
Premise 1 Take an arbitrary x such that Px, say m. Then:
 (i) $f_1 : P \to T_1$, $f_1(m) = m_1$ and $P_1 m_1$,

(ii) $f_2 : T_1 \to T_2$, $f_2(m_1) = m_2$ and $P_2 m_2$,
(iii) $f_3 : T_2 \to T_3$, $f_3(m_2) = m_3$ and $P_3 m_3$,
$\vdots$
(n) $f_n : T_{n-1} \to T_n$ and $f_n(m_{n-1}) = m_n$ and $P_n m_n$, for functions f_1 to f_n, types T_1 to T_n and properties P_1 to P_n.

Premise 2 Therefore (since m was arbitrary), there exist functions f_1 to f_n st $\forall x(Px \to \exists y \text{ st } (f_n(f_{n-1}(\ldots f_2(f_1(x)))))) = y) \wedge P_n y)$.

Premise 3 $(P_n m_n \to P_{n-1}(f_n^{-1}(m_n))) \wedge (P_{n-1}(m_{n-1})) \to P_{n-2}(f_{n-1}^{-1}(m_{n-1}))) \wedge \cdots \wedge (P_2 m_2 \to P_1(f_2^{-1}(m_2))) \wedge (P_1 m_1 \to E(f_1^{-1}(m_1)))$.

Conclusion $\forall x(Px \to Ex)$.

Critical Questions:

1. Is m really arbitrary? Does it have any properties that other objects with property P might not have?
2. Is the function f_i well-defined, for all i? Does it always take objects of type T_i and output objects of type T_{i+1}?
3. Does property P_n always hold for m_n?
4. What is the relationship between P_n and E? Can we conclude that because Pm and $P_n m_n$, that Em?

We instantiate this scheme in terms of Lakatos's description of Cauchy's proof sketch, rephrased from [23, pp. 7–8], as the following procedural scheme, with critical questions all taken from [23, p. 8] (clearly there would be further critical questions too). Note that the third function in step (iii) is applied multiple times.

13.5.2. Example Instantiation of the Argumentation Scheme for Procedural Argument.

Premise 1 Take the cube, which is an arbitrary polyhedron. Then:
(i) Remove a face from the cube. We can stretch the remaining surface flat on the blackboard, and the Euler characteristic of this surface will be that of the cube, minus one (as we removed a face).
(ii) Triangulate the connected network from step (i). The Euler characteristic from step (i) is preserved (since for any new edge which is added, one will always get a new face).
(iii) Drop the triangles one by one from the triangulated map from step (ii). Again, the Euler characteristic from step (ii) is preserved (since there are only two alternatives — the disappearance of one edge

and one face, or else of two edges, one vertex and a face). Apply this step repeatedly (n times) until a single triangle remains.

Premise 2 The triangle from step (*iii*) has an Euler characteristic of one (since $V - E + F = 3 - 3 + 1 = 1$).

Premise 3 Therefore for any polyhedron, there exist procedures (*i*), (*ii*) and (*iii*) which will turn the polyhedron into a single triangle, and that triangle has an Euler characteristic of one.

Premise 4 If we start with a triangle, which has an Euler characteristic of one, then there is a way in which we can add triangles one by one (n times), resulting in a triangulated map each time and preserving the Euler characteristic (we either add one edge and one face, or else of two edges, one vertex and a face). There is a way of removing edges to "untriangulate" the map. The Euler characteristic is preserved from the previous step (since for any new edge which is removed, one will always get one less face). There is now a way in which we can add a face and assemble the network into a polyhedron. Since there is one more face, the Euler characteristic is now two.

Conclusion All polyhedra have an Euler characteristic of two.

Critical Questions:

1. Does the cube have particular properties which allow us to perform this procedure, which other polyhedra do not share? For instance, can we perform step (*i*) on any polyhedron, with the property that the resulting surface can be stretched flat on a board? (Proposed by ALPHA [*ibid.*])
2. When we perform step (*ii*), does the property of preserving the Euler characteristic always hold (do we always get a new face for any new edge)? (Proposed by BETA [*ibid.*])
3. When we perform step (*iii*), does the property of preserving the Euler characteristic always hold (do we always remove one edge and one face, or two edges, one vertex and a face)? Are there always a countable number of repetitions of step (*iii*) which will result in a single triangle? (Proposed by GAMMA [*ibid.*])

We can now outline our suggested argumentation schemes for Lakatos's proof-changing methods. In all three schemes, initially only the first two premisses are noted: given these, Lakatos suggests doing the work involved to find the third premise and then the conclusion may be reached.

13.5.3. Suggested Argumentation Scheme for Global and Local Lemma-Incorporation.

Premise C_1 is a global counterexample to conjecture P, that is $C_1 \Rightarrow \neg P$.

Premise C_1 is a local counterexample to conjecture P, that is at least one of the premisses on which P depends, $Prem_i$ is itself a conclusion in an (instantiation of) a scheme which contains critical question(s) which must receive a defeating answer if C_1 holds.

Premise X is the concept of *objects for which* $Prem_i$ *holds.*

Conclusion Replace P by $X \Rightarrow P$.

Note that if P is the proposition $R \Rightarrow Q$, then the replacement is $(R \wedge X) \Rightarrow Q$, which is clearly in line with our Lakatosian example. As noted earlier (p. 43) this is a form of strategic withdrawal in which the concept to which one withdraws is provided by the notion of a problematic premise being satisfied.

13.5.4. Suggested Argumentation Scheme for Local and not Global Lemma-Incorporation.

Premise C_1 is a local counterexample to conjecture P, that is at least one of the (instantiations of) the schemes upon which the derivation of P depends, $Prem_i$, contains critical question(s) CQ_i which must receive a defeating answer if C_1 holds.

Premise C_1 is not a global counterexample to conjecture P, that is $C_1 \not\Rightarrow \neg P$.

Premise X is the concept of *objects for which* $Prem_i$ *fails*

Conclusion Replace $Prem_i$ by $\neg X \Rightarrow Prem_i$.

Again, note that if $Prem_i$ is the proposition $R \Rightarrow S$, then the replacement is $(R \wedge \neg X) \Rightarrow S$, which is in line with the Lakatosian example. We also noted earlier (p. 44) that this is a form of piecemeal exclusion, in which the concept which one excludes is designed to exactly cover the cases in which a problematic premise fails.

13.5.5. Suggested Argumentation Scheme for Global and not Local Lemma-Incorporation.

Premise C_1 is a global counterexample to conjecture P, that is $C_1 \Rightarrow \neg P$.

Premise C_1 is not a local counterexample to conjecture P, that is none of the (instantiations of) schemes upon which the derivation of P depends contains critical question(s) CQ_i which must receive a defeating answer if C_1 holds ($\forall i$ st $1 \leqslant i \leqslant n, P_i \Rightarrow C_1$, where n is the number of premisses).

Premise $Prem_i$ is a premise in the derivation of P which contains an unstated assumption, A_1 such that $(A_1 \wedge Prem_i) \Rightarrow \neg C_1$.

Conclusion Replace P by $A_1 \Rightarrow P$, and $Prem_i$ by $A_1 \wedge Prem_i$.

The absence of an undercutter follows from the incompleteness of the set of critical questions. Hence, in a narrow sense, we can have rebutting without undercutting. But the incompleteness can always (in principle) be remedied: there is always an A_1 for the proof analysis to find. Once it has been remedied, the undercutter has been found, since $C_1 \Rightarrow \neg A_1$ (by transposition), and thus the question 'Does A_1 hold?' must receive a defeating answer if C_1 holds.

The unstated assumption may well be a further property in one of the proof steps (from premise 1 in argumentation scheme in §13): this is the case in Lakatos's examples. That is, $f_i\colon T_{i-1} \to T_i$ and $f_n(m_{i-1}) = m_i$ and $P_i m_i$ *and* there is a further property P' such that $P'm_i$, for some $1 \leqslant i \leqslant n$. Initially this further property is not explicit, but it holds for all (known) positive examples and fails for all (known) global but not local counterexamples. The patch is then to make it explicit, by adding it to the relevant proof step. The presence of an unstated assumption in the derivation suggests that the set of critical questions is incomplete. (At some point the question 'Does A_1 hold?' should have been posed, since a negative answer would undercut the derivation.) There is nothing intrinsically wrong with this — it is *informal* logic after all — but it does indicate that the proof analysis is shoddy, which is of course Lakatos's point.

Closing remarks

Whether we are schemifying Toulmin or Toulminizing abduction, an analysis of where our five protagonists may agree and disagree suggests all kinds of fascinating new areas. Connections between the different theories have typically been neglected, to the extent that our most contemporary thinker, Walton, is the only author to discuss any of the others in the works we cite, and even he does not consider Lakatos. Finding and expanding such connections may well be a fruitful avenue of research, so that different descriptions of our elephant may, possibly, begin to converge. Additionally, we hope that our particular breed of interest, the mathematical elephant, will prove to be similar in many areas to other breeds, and both informal logic and the philosophy of mathematical practice will benefit. In any case, we shall enjoy enormously the attempt to describe it in such a way.

Acknowledgments. We are grateful to the Mathematical Reasoning Group at the University of Edinburgh for a lively debate on some of the ideas presented here. In particular, we are grateful to Alan Smaill and Markus Guhe for their thoughts and comments on earlier drafts of this work. We would also like to thank the organisers of the conference on *Argumentation as a cognitive process*, held at Nicolaus Copernicus University, Toruń, Poland on 13–15 May 2010, for all their work in bringing such an interesting community together. Alison Pease is supported by EPSRC grant EP/F035594/1.

References

[1] Aberdein, A., "The uses of argument in mathematics", *Argumentation* 19 (2005): 287–301.

[2] Aberdein, A., "Managing informal mathematical knowledge: Techniques from informal logic", pages 208–221, in: J. M. Borwein and W. M. Farmer (eds.), *MKM 2006, LNAI 4108*, Springer-Verlag, Berlin, 2006.

[3] Banegas, J. A., "L'argumentació en matemàtiques", pages 135–147 in: E. Casaban i Moya (ed.), *XIIè Congrés Valencià de Filosofia*, València, 1998. http://my.fit.edu/~aberdein/Alcolea.pdf

[4] Anderson, B., "Diary", *Spectator* 313 (9481), May 15th 2010. http://www.spectator.co.uk/politics/all/5996573/part_2/diary.thtml

[5] Bocheński, I. M., *A History of Formal Logic*, Chelsea Pub Co, New York, N.Y., 1970 [1956].

[6] Boden, M. A., *The Creative Mind: Myths and Mechanisms*, Weidenfield and Nicholson, London, 1990.

[7] Burley, W., *On the Purity of the Art of Logic: The Shorter and Longer Treatises*, Yale University Press, New Haven, CT, 2000. Translated by P. V. Spade.

[8] Cauchy, A. L., *Cours d'Analyse de l'École Royale Polytechnique*, de Bure, Paris, 1821.

[9] Colton, S., *Automated Theory Formation in Pure Mathematics*, Springer-Verlag, 2002.

[10] Colton, S., A. Bundy, and T. Walsh, "On the notion of interestingness in automated mathematical discovery", *International Journal of Human Computer Studies* 53, 3 (2000): 351–375.

[11] Conan Doyle, A., *The Memoirs of Sherlock Holmes*, Forgotten Books, 2008 [1894].

[12] Dauben, J. W., "Peirce's place in mathematics", *Historia Mathematica* 9, 3 (1982): 311–325.

[13] Dove, I. J., "On mathematical proofs and arguments: Johnson and Lakatos", pages 346–351 in: F. H. Van Eemeren and B. Garssen (eds.), *Proceedings of the Sixth Conference of the International Society for the Study of Argumentation*, volume 1, Sic Sat, Amsterdam, 2007.

[14] Dunham, W., *Euler: The Master of Us All*, The Mathematical Association of America, Washington, DC, 1999.

[15] Epstein, S. L., "Learning and discovery: One system's search for mathematical knowledge", *Computational Intelligence* 4, 1 (1988): 42–53.

[16] Gasteren, A. J. M., *On the shape of mathematical arguments*, Lecture notes in computer science, volume 445, Springer, Berlin, 1990.

[17] Gray, J., *The Hilbert Challenge*, Oxford University Press, Oxford, 2000.

[18] Grosholz, E. R., *Representation and Productive Ambiguity in Mathematics and the Sciences*, Oxford University Press, New York, NY, 2007.

[19] Hairer, E., and G. Wanner, *Analysis by Its History*, Undergraduate Texts in Mathematics, Springer, New York, NY, 2008.

[20] Hanson, N. R., "Is there a logic of scientific discovery?", pages 20–35 in: H. Feigl and G. Maxwell (eds.), *Current Issues in the Philosophy of Science*, Symposia of scientists and philosophers. Proceedings of Section L of the American Association for the Advancement of Science, 1959, Holt, Rinehart and Winston, New York, NY, 1961.

[21] Heath, T. L., *The Thirteen Books of Euclid's Elements*, Cambridge University Press, Cambridge, 2nd edition, 1925.

[22] Hitchcock, D., "Toulmin's warrants", pages 69–82 in: F. H. van Eemeren, J. Blair, C. Willard, and A. F. Snoeck-Henkemans (eds.), *Anyone Who Has a View. Theoretical Contributions to the Study of Argumentation*, Kluwer Academic Publishers, Dordrecht, 2003.

[23] Lakatos, I., *Proofs and Refutations*, Cambridge University Press, Cambridge, 1976.

[24] Lakatos, I., "Cauchy and the continuum: The significance of non-standard analysis for the history and philosophy of mathematics", pages 43–60 in: J. Worral and C. Currie (eds.), *Mathematics, science and epistemology*, Philosophical Papers, volume 2, Cambridge University Press, Cambridge, 1978.

[25] Lakatos, I., "Proofs and refutations (I)", *The British Journal for the Philosophy of Science* 14 (1963): 1–25.

[26] Lakatos, I., "Proofs and refutations (II)", *The British Journal for the Philosophy of Science* 14 (1963): 120–139.

[27] Lakatos, I., "Proofs and refutations (III)", *The British Journal for the Philosophy of Science* 14 (1963): 221–245.

[28] Lakatos, I., "Proofs and refutations (IV)", *The British Journal for the Philosophy of Science* 14 (1964): 296–342.

[29] Larvor, B., *Lakatos: An Introduction*, Routledge, London, 1998.

[30] Lenat, D., *AM: An Artificial Intelligence Approach to Discovery in Mathematics*, PhD thesis, Stanford University, 1976.

[31] Leng, M., "What's there to know? A fictionalist account of mathematical knowledge", pages 84–108 in: M. Leng, A. Paseau, and M. Potter (eds.), *Mathematical Knowledge*, Oxford University Press, Oxford, 2007.

[32] Macedo L., and A. Cardoso, "Creativity and surprise", in: G. Wiggins (ed.), *Proceedings of the AISB'01 Symposium on Artificial Intelligence and Creativity in Arts and Science*, 2001.

[33] Mancosu, P., "Mathematical explanation: Problems and prospects", *Topoi* 20, 1 (2001): 97–117.

[34] Marquis, J.-P., "Abstract mathematical tools and machines for mathematics", *Philosophia Mathematica* 5 (1997): 250–272.

[35] Muntersbjorn, M., "Construction, articulation, and explanation: Phases in the growth of mathematics", pages 19–42 in: B. Van Kerkhove, J.-P. van Bendegem, and J. De Vuyst (eds.), *Philosophical Perspectives on Mathematical Practice*, College Publications, London, 2010.

[36] Pease, A., *A Computational Model of Lakatos-style Reasoning*, PhD thesis, School of Informatics, University of Edinburgh, 2007. http://hdl.handle.net/1842/2113

[37] Pease, A., M. Guhe, and A. Smaill, "Analogy formulation and modification in geometry", pages 358–364 in: *Proceedings of the Second International Conference on Analogy*, 2009.

[38] Pease, A., A. Smaill, S. Colton, and J. Lee, "Bridging the gap between argumentation theory and the philosophy of mathematics", *Foundations of Science* 14, 1–2 (2009): 111–135.

[39] Peirce, C. S., *Collected Papers of Charles Sanders Peirce*, Eight Volumes, Harvard University Press, Cambridge, Mass, 1931–58.

[40] Peirce, C. S., *The New Elements of Mathematics*, Mouton, The Hague, 1976.

[41] Plato, *The Republic*, OUP, Oxford, 1993.

[42] Pollock, J., "The structure of epistemic justification", *American Philosophical Quarterly, monograph series* 4 (1970): 62–78.

[43] Pollock, J., *Contemporary Theories of Knowledge*, Rowman and Littlefield, Totowa, NJ, 1986.

[44] Pollock, J., *Cognitive Carpentry*, The MIT Press, Cambridge, MA., 1995.

[45] Polya, G., *Mathematics and plausible reasoning*, volume 1, Induction and analogy in mathematics, Princeton University Press, 1954.

[46] Quine, W. V. O., "Two dogmas of empiricism", *From a Logical Point of View*, Harvard University Press, 1953. First version of the paper, without any reference to Duhem, in *The Philosophical Review* 60 (1951): 20–53.

[47] Reed, C., and G. Rowe, "Translating Toulmin diagrams: Theory neutrality in argument representation", *Argumentation* 19, 3 (2005): 267–286.

[48] Robinson, A., *Non-standard Analysis*, Princeton University Press, Princeton, New Jersey, 1996. Revised Edition.

[49] Robinson, R., "Analysis in Greek geometry", *Mind* 45 (1936): 464–73.

[50] Sandifer, C. E., *The early mathematics of Leonhard Euler*, The Mathematical Association of America, 2007.

[51] Seidel, P. L., "Note über eine Eigenschaft der Reihen, welche Discontinuirliche Functionen Darstellen", *Abhandlungen der Mathmatisch-Physikalischen Klasse der Königlich Bayerischen Akademie der Wissenschaften* 5 (1847): 381–93.

[52] Toulmin, S., *The uses of argument*, CUP, Cambridge, 1958.

[53] Toulmin, S., R. Rieke, and A. Janik, *An Introduction to Reasoning*, Macmillan, London, 1979.

[54] Walton, D., C. Reed, and F. Macagno, *Argumentation Schemes*, Cambridge University Press, Cambridge, 2008.

[55] Walton, D. N., *Argument Schemes for Presumptive Reasoning*, Lawrence Erlbaum Associates, Mahwah, NJ, 1996.

[56] Walton D. N., and C. A. Reed, "Diagramming, argumentation schemes and critical questions", pages 195–211 in: F. H. van Eemeren, J. Anthony Blair, Ch. A. Willard, and A. Francisca Snoeck Henkemans (eds.), *Anyone Who Has a View: Theoretical Contributions to the Study of Argumentation*, Kluwer Academic Publishers, Dordrecht, 2003.

Alison Pease
Centre for Intelligent Systems and their Applications
Informatics Forum
University of Edinburgh
8 Crichton Street
Edinburgh, EH8 9AB
A.Pease@ed.ac.uk

Andrew Aberdein
Department of Humanities and Communication
Florida Institute of Technology
150 West University Blvd
Melbourne, Florida 32901-6975, U.S.A.
aberdein@fit.edu