

Possible Worlds and Imagination in Informal Logic

John Nolt
University of Tennessee

Logic's main task is to evaluate strength of reasoning. By "strength of reasoning," I mean the degree of logical connection between premises and conclusion, regardless of the premises' truth. Of course we can't tell whether an argument ought to persuade us unless we know whether its premises are true. But truth of premises is in general not a logical matter. Logic is concerned, rather, with the question, "If the premises were true, how likely would the conclusions' truth be? This likelihood is strength of reasoning.

The simplest measure of strength of reasoning is the all-or-nothing distinction: valid/invalid. One common way of making this distinction is to hold that a valid argument is one with a valid form, and an invalid argument is one with an invalid form. What counts as a valid or invalid form is then decided by some formal system—usually Aristotelian syllogistic or the predicate calculus. But form is a poor guide to validity in natural language. I am not referring here to the fact that natural language arguments may have several logical forms, though such formal ambiguity is certainly trouble enough. Rather, my concern is that validity often has nothing at all to do with form as defined by the standard logical calculi. Consider the following inferences:

- The infection is spreading rapidly.
∴ The infection is spreading.
- A calf was born.
∴ Something happened.
- The vase is blue.
∴ The vase is colored.

The validity of each is due to aspects of meaning not captured by the standard logical operators.

Advocates of formal approaches will reply, of course, that each inference **becomes** formally valid as soon as we add certain "hidden premises." But what if we insist on asking about these inferences themselves, just as they stand? The answer can only be that **as they stand** they must be judged invalid by the formal criterion. But they are not invalid. In each case, if the stated premises (and these alone) are true, then the conclusion must (in the strongest sense of "must") be true. Each argument **as it stands** is just as strong as any formally valid argument. Thus the formal criterion of validity is simply wrong as a measure of strength for arguments in ordinary language.¹

It may be that some more advanced notion of logical form, perhaps imported from some rich categorial grammar, could provide a more adequate formal criterion of validity. But such a notion is not currently available, and even if it were it would be too complicated to teach to students who take only one or two courses in logic. My concern is with what we ought to teach just such students. And my first conclusion is that we should not teach them to apply the formal criterion of validity to arguments in ordinary language. It is not accurate outside of a limited range of formalizable arguments.

The formal criterion is also inadequate in a practical sense. Consider the well-known and obviously valid argument:

- There are more people than hairs on any single person's head.
- No one is bald.
- ∴ At least two people's heads have the same number of hairs.

By suitable translation, this argument becomes formally valid (at least in second order logic). But the translation and the theory behind it are too advanced for beginning students. The formal criterion is useless to them here. Nevertheless, they are perfectly capable of seeing that the argument is valid. In cases like this, imagination and intuition are far more efficient, and only slightly less reliable, than formalism. Moreover, examples of this sort are exceedingly common in practical reasoning. Virtually all the arguments our students encounter in their science and mathematics classes, for example, are, like this one, arguments they will never be able to formalize.

If we, as logic teachers, are to give short-term students any help in judging the strength of these arguments, it will not be by teaching them to discern their logical form. Instead, it will be by sharpening the imagination they already display in evaluating these arguments intuitively. For that they need to understand validity, not in terms of the inaccurate criterion of logical form, but as a matter of conceptual necessity. They need to regard an argument as valid just in case the simultaneous truth of its premises and falsity of its conclusion cannot coherently be conceived.

I want to emphasize that this approach is in no way **opposed** to formal methods. Indeed, formal methods, such as truth tables or Venn diagrams, are extremely useful in helping

students to see why it is inconceivable for premises to be true and conclusions false with respect to certain limited classes of arguments. We should teach these methods, but we should also give students the means to evaluate arguments to which they are inapplicable, either for practical or for theoretical reasons. And for this we need the conceptual, not the formal, criterion of validity.

Admittedly the conceptual criterion is more vague than the formal one. But I have argued that **for our students' purposes** is both more accurate and more widely applicable.

Now I think that the most useful way to formulate the conceptual criterion is in terms of possible worlds. The criterion, once again, is that an argument is valid if and only if the simultaneous truth of its premises and falsity of its conclusion cannot be coherently conceived. The connection with possible worlds is obvious and familiar. For a situation can be coherently conceived just in case a world in which it occurs is possible (in just the sense of "possible" that is relevant here). That is, possible worlds and coherently conceivable worlds are the same thing. Thus, appealing directly to our students' intuitions, we can say that an argument is valid just in case there is no possible (i.e., coherently conceivable) world in which its premises are true and its conclusion is false. The test for validity thus becomes a thought-experiment, an attempt to conceive certain kinds of worlds. The task of the logic teacher is to develop the students' ability to perform this test accurately.

Formulating the criterion of validity in terms of possible worlds enables us to link it systematically with other logical concepts which are also expressible in these terms: contingency, inconsistency, necessity, incompatibility, entailment, logical equivalence, and so on. This allows us, if we wish, to give students a broad and systematic understanding of these concepts at an early stage in their studies.² But the most important advantages of the possible worlds approach emerge when we turn from the valid/invalid distinction to more sensitive measures of argument strength. This, of course, is the realm of induction.

There are two salient facts about induction: (1) many, if not most, of the arguments students encounter in practice are inductive, and (2) these vary widely in form, and most bear little resemblance to the standard textbook examples. Some of the more innovative logic texts, most notably Steven N. Thomas's **Practical Reasoning in Natural Language**,³ address this problem in the way I am advocating—i.e., informally by developing the students' intuitions. Thomas suggests that we estimate strength of reasoning (he calls it "degree of validity") along a scale that ranges from nil to deductively valid. He provides many sample arguments together with his estimates of their strength. But he gives us no standard procedure for making such estimates, and the result is a very large and worrisome element of subjectivity.

Thomas's method also suffers from a more subtle defect; it fails to separate strength of reasoning from collateral evidence. Consider the statistical syllogism:

75% of bankers drive Cadillacs.
Alphonso is a banker.
∴ Alphonso drives a Cadillac.

This is a fairly strong argument. In fact, there is a sense in which it is precisely 75% as strong as the valid argument:

All bankers drive Cadillacs.
Alphonso is a banker.
∴ Alphonso drives a Cadillac.

But suppose we know that Alphonso is broke and that virtually no one who is broke drives a Cadillac. Then what should we say about the statistical syllogism? The answer, of course, is that we should still say that its reasoning is 75% as strong as the reasoning of the valid inference. This collateral evidence is not part of the statistical syllogism; adding it produces a different argument. What's wrong with the statistical syllogism under these conditions is neither that its premises are false nor that its reasoning is weak, but rather that it fails to tell the whole story. It violates what Carnap and others have called **the requirement of total evidence**.⁴

Now the more subtle problem with Thomas's approach is that it fails to distinguish violations of the requirement of total evidence from weakness in reasoning. Confronted with the statistical syllogism just mentioned, the student doesn't know whether to classify the inference as strong (because of the strength of the reasoning) or weak (because of the contrary collateral evidence) or somewhere in between. The subjectivity of Thomas's approach is multiplied by such confusions. We can give students little help with the great majority of nondeductive arguments unless we can produce a more objective, more systematic and better articulated evaluation procedure.

The possible worlds approach can provide just that. The fundamental idea is to measure strength of reasoning by the proportion of conceivable worlds in which the conclusion is true among worlds in which the premises are true. If this proportion is 100%, the argument is valid. If it is 0%, the premises imply the negation of the conclusion. And there are infinitely many degrees of strength between these two extremes. Thus strength of reasoning is judged along a continuum, on which validity is a limiting case of inductive strength.

But is such a measure feasible—in theory or in practice? Let's take theory first. The most obvious objection is that since the number of conceivable worlds is infinite, proportions or percentages such as I have been describing are simply unavailable. There's no such thing as a percentage of an infinite collection. In the most obvious sense that is true. But it may not be true if we regard the percentage as a frequency. The set of all natural numbers is infinite. Yet in their natural order, the frequency of those evenly divisible by 10 is 10%. It is a familiar fact of probability theory that many sequences, both orderly and random, are amenable to such frequency measures. Do similar measures exist for the domain of possible (i.e. conceivable) worlds? The first thing to note, of course, is that, unlike the natural numbers, the domain of possible worlds is not a naturally well-ordered sequence. It seems probable that certain sequences of members of this domain will have measurable limiting frequencies for certain properties, and others will not. And simple reordering of a sequence may change such frequencies drastically. Perhaps there is some way of choosing sequences which in some sense "typify" the distribution of a property among worlds, but it's not clear how this could be done, or even what it would mean, since we have no precise characterization of the domain of conceivable worlds.

Thus from a theoretical viewpoint, the prospect of rigorously defining a concept of proportion for possible worlds is bleak. But the picture brightens considerably when we recall that we are not seeking a formal or mathematically rigorous theory, but only an intuitive procedure for evaluating informal arguments. Thomas's method is the best we have now. If we can improve on it, we'll still have accomplished something useful.

Now despite the theoretical difficulties inherent in the notion of proportions of possible worlds, I think we have or can develop a fairly objective intuitive grasp of this concept. It works like this: first we imagine the range of conceivable worlds in which the premises of an argument are true, with all the wild variations that logical or conceptual possibility permits. Then we imagine ourselves choosing worlds from this range "at random." The frequency among these choices of worlds in which the conclusion is true is the measure of the argument's strength. The reason these choices should be made "at random" is that we want them, insofar as possible, to typify the class as a whole. If we intentionally or unconsciously "bias" our choices, they may yield any frequency whatsoever.

A simple example will illustrate this point. Consider the fallacious appeal to popularity:

Everyone thinks that Neil Armstrong walked on the moon.
∴ Neil Armstrong walked on the moon.

To estimate the strength of this reasoning, we first imagine the range of conceivable worlds in which its premise is true. These include worlds in which everyone is right and Armstrong did walk on the moon, as well as worlds in which everyone is wrong and he didn't. Then we estimate the frequency among these worlds of worlds in which he did make the moonwalk. It is easy to prejudice this estimate by only considering worlds too much like the actual world. If we think, for example, only of worlds in which people's beliefs about such matters are generally accurate, in which NASA and the news media which report the moonwalk are truthful, and so on, we'll probably conclude that the conclusion is true in most of these worlds. But the proper procedure is to make a "random" selection from among **all conceivable** worlds in which the premise is true. These include worlds in which people never believe the truth, or in which NASA and the press fabricated the moonwalk, or in which there was a moonwalk, but by an imposter, not by Armstrong, or in which both the moon and Neil Armstrong are myths—the variations are endless. It is clear that in a "fair" selection from among these worlds, the proportion of those in which Armstrong really did walk on the moon is not high. Thus we conclude, as we should, that the argument's reasoning is weak.

I wish to make no claim here either of rigor or of theoretical adequacy. I have not defined the relevant notions of frequency and randomness. This evaluation procedure is intuitive and imprecise. Nevertheless, I think it is useful, and I want to offer five pragmatic reasons for adopting it. First, in my two years of experience teaching it to undergraduates, it has seemed to produce a reasonable degree of agreement about argument strength. Second, the estimates of strength that it yields are fairly accurate (i.e., it judges valid arguments as valid, strongly inductive arguments as strong, and traditional fallacies as weak). Third, given precisely formulated statistical arguments, it yields precise numerical percentages as measures of reasoning strength—percentages which, moreover, are right. Fourth, unlike Thomas's method, it makes a clear distinction between strength of reasoning and collateral evidence. And, finally, it provides a unified conceptual framework within which the student can readily locate and organize, not only the "all-or-nothing" logical concepts discussed earlier (validity, inconsistency, necessity, etc.), but also those that admit of degrees (various kinds of probability, inductive strength, strength of premises, strength of conclusion, and so on). These can all be characterized in terms of possible worlds.

My first two claims are empirical: that the method yields relatively consistent answers, and that these answers are

generally accurate. I have no scientific data on this—only my own classroom observations. So I'll just register these claims with the hope that others will put them to the test, and move on.

My third claim, that the method yields precise and accurate results for precisely stated statistical arguments, can be illustrated in a simple way by successive consideration of the following three inferences:

- (1) Alphonso, Beula, Cynthia and Douglas are the only bankers.
Exactly three bankers drive Cadillacs.
∴ Alphonso drives a Cadillac.
- (2) There are exactly four bankers.
Exactly three of them drive Cadillacs.
Alphonso is a banker.
∴ Alphonso drives a Cadillac.
- (3) (Exactly) 75% of bankers drive Cadillacs.
Alphonso is a banker.
∴ Alphonso drives a Cadillac.

The third of these is the statistical syllogism discussed earlier, whose strength of reasoning, as we noted, is 75% (or 0.75 on a scale from 0 to 1). It is easily seen that this is the appropriate figure for the other two as well. The procedure I am advocating yields just this figure for all three.

The procedure, once again, is this: imagine the range of possible worlds in which all the premises are true. Choose worlds from this range "at random" and estimate the frequency among them of worlds in which the conclusion is true. That frequency is strength of reasoning.

Let's apply this procedure to the first inference. Imagine the range of worlds in which only Alphonso, Beula, Cynthia and Douglas are bankers and exactly three of them drive Cadillacs. This range of worlds falls naturally into four classes, according to which of the four does **not** drive a Cadillac. Thus one of these classes consists of all these worlds in which Alphonso does not drive a Cadillac, one consists of all in which Beula does not drive a Cadillac, and so on. Now these classes are precisely isomorphic, in the sense that there are just as many ways of conceiving that the one who does not drive the Cadillac is Alphonso as there are ways of conceiving that it is Beula or Cynthia or Douglas. Thus it seems intuitively obvious that if we choose "at random" from among the union of these classes, in the long run we will find that 75% of our choices are worlds in which Alphonso drives a Cadillac, i.e., in which the conclusion is true. The strength of reasoning is therefore 75%.

For the other two cases, the principle is the same, but there is more detail to consider. The worlds in which the premises of (2) are true fall into classes according to the identities of the three bankers other than Alphonso. All the worlds in which the same three individuals are the other bankers belong to a single class. There will be infinitely many of these classes, since we can imagine different bankers **ad infinitum**. But each class consists of worlds structured exactly like the worlds of the previous example. That is, in all the worlds of each class, the same four individuals are the only bankers, and all but one of them drive a Cadillac. The identity of the one who doesn't drive a Cadillac, however, varies from world to world within a given class.

Because of this structural similarity, we can repeat the reasoning of the last example with respect to each class of worlds. We divide each class into four subclasses, so that the same banker is without a Cadillac in all the worlds of each

subclass. Thus, as in the last example, Alphonso drives a Cadillac in 75% of the worlds of each class. But since each of the infinitely many classes is isomorphic to all the others (in an intuitively obvious sense), "random" choice of worlds from their union ought in the long run to yield a sequence in which the frequency of worlds in which Alphonso drives a Cadillac is, once again, 75%.

Arguments like (3) add still another level of complexity, but nothing new in principle. Here the worlds in which the premises are true fall into classes, based first upon the number of bankers they contain. This number is always a multiple of four, since otherwise it is impossible for exactly 75% of the bankers to drive Cadillacs. (You can't, for example, take 75% of three bankers.) Now the simplest of these classes is the one containing worlds in which there are exactly four bankers, three of whom drive Cadillacs. This is precisely the same class of worlds we looked at in the last example. So by our previous reasoning, the frequency among them of worlds in which Alphonso drives a Cadillac is 75%. Similar reasoning yields the same conclusion for each of the other classes (worlds containing eight bankers, worlds containing twelve, and so on). Thus, once again, it seems clear that in a sequence "random" choices from the union, we'll obtain a frequency of 75% for worlds in which the conclusion is true. Once again, this figure is the strength of the argument's reasoning.

These examples illustrate the method of possible worlds in application to an extremely simple series of statistical arguments. The method also works for more sophisticated statistical arguments, but I will not try to show that here.⁵

Instead, I'll turn to my fourth claim—that the method distinguishes clearly between strength of reasoning and collateral evidence. This is already implicit in the discussion of the previous three examples. In each case we looked at the class of **all conceivable** worlds in which the premises are true. It is the requirement that we consider this entire class without "bias" that prevents us from illegitimately importing background knowledge into our evaluation. Consider the third inference once again. It doesn't matter if in the actual world Alfonso is broke and doesn't drive anything. Because we are looking at **all conceivable** worlds in which the premises are true, not just the actual world (or worlds similar to it), our frequency estimates are unaffected by this collateral evidence. For among worlds in which the premises are true, we must consider all worlds in which this collateral information is false, as well as all those in which it is true. The strength of reasoning remains 75%, regardless of what is true or false and regardless of what we know or fail to know in the actual world.

Collateral evidence can enter into our evaluation only by being explicitly included among the premises of the argument. But that makes a new argument. The strength of the original remains constant. This eliminates much of the confusion inherent in evaluation techniques which fail to separate strength of reasoning from collateral evidence.

My fifth and final claim is that the framework of possible worlds provides a unified scheme for defining and inter-

relating, not only the usual all-or-nothing logical concepts, but also those that admit of degrees. The strength of a **statement**, for example, can be defined as the proportion among all possible worlds of possible worlds in which it is false. The weakest statements thus are necessary truths (false in no worlds), and the strongest are inconsistencies (false in all worlds). Contingent statements are arrayed in between. With this definition, it is easy to explain in terms of possible worlds how and why strength of reasoning tends to vary in direct proportion to the strength of the premises and in inverse proportion to the strength of the conclusion. It also becomes relatively easy to explain to students why in those limiting cases when premises are inconsistent (i.e., maximally strong) or the conclusion is necessary (i.e., maximally weak), the reasoning is maximally strong (i.e., valid). This is but one example of the sorts of conceptual links which can be forged within the framework of possible worlds. We can also distinguish and interrelate several important notions of probability, and a variety of other concepts of varying utility and importance. But there is not room for these details in a brief presentation, such as I have attempted to give here.⁶

Notes

1. The hard-core advocate of formal approaches will not be satisfied by this reply. He or she will insist that in declaring these inferences valid, we are at least implicitly appealing to hidden assumptions, and that they cannot be regarded as valid in isolation. But there is no more justification for holding that the necessary truth "All blue things are colored" is needed to complete the third argument, for example, than there is for holding that the necessary truth $((A \supset B) \& A) \supset B$ is necessary to complete instances of **modus ponens**. And Lewis Carroll showed us what was wrong with the latter idea. (See "What the Tortoise Said to Achilles," *Mind* 4, 1895, pp. 278-280.)
2. As an example of what can be done along these lines, see Raymond Bradley and Norman Swartz, **Possible Worlds: An Introduction to Logic and its Philosophy**, Indianapolis and Cambridge, Hackett Publishing Co., 1979. Swartz and Bradley also hint at the possibility of extending the possible worlds concept to induction (see pp. 379-81), as I advocate in what follows.
3. Englewood Cliffs, N.J., Prentice-Hall, 1981.
4. See Rudolf Carnap, **Logical Foundations of Probability**, Chicago and London, University of Chicago Press, 1962, pp. 211-15.
5. A number of examples are given in my **Informal Logic: Possible Worlds and Imagination**, New York, McGraw-Hill Book Company, 1983, Chapter 6.
6. These details are explained in **Informal Logic** (see footnote 5). ●

John Nolt, Department of Philosophy, University of Tennessee, Room 801 McClung Tower, Knoxville, Tennessee, 37996