1 What Is Logic?

1 Arguments

Symbolic logic is usually described as the study of the difference between *valid* and *invalid* arguments, so we begin with an explanation of this terminology. An argument is a piece of discourse which contains some *premises*, a *conclusion*, and perhaps also some reasoning in which an attempt is made to *derive* the conclusion *from* the premises. The premises are those statements which, for the purposes of the argument, are being accepted as true. The conclusion is the statement whose truth the argument aims to demonstrate, given acceptance of the premises. Hence this 'logical' sense of 'argument' is quite different from its sense in 'Smith and Jones are having an argument', though an argument in the logical sense might be reconstructible from what Smith or Jones says.

For present purposes, we take an argument to consist just in a listing of the premises, and then the conclusion. A *valid* argument is an argument whose conclusion *follows* from its premises, and correspondingly, an invalid argument is an argument whose conclusion *does not follow* from its premises. Here are two very simple arguments, one valid, the other invalid, which illustrate the difference between the conclusion following from the premises and its not following:

A: (1) If our currency loses value then our trade deficit will narrow.
(2) Our currency will lose value.
(3) ∴ Our trade deficit will narrow.[1]

B: (1) If our currency loses value then our trade deficit will narrow.
(2) Our trade deficit will narrow.
(3) ∴ Our currency will lose value.

In argument A, the truth of the conclusion (3) is guaranteed by the truth of the two premises (1) and (2): *if* (1) and (2) are true, then (3) will be true as well. So (3) follows from (1) and (2). But it is not difficult to see that in argument B, the truth of (1) and (2) does not guarantee the truth of (3). (1) of B says that a depre-

[1] The symbol '∴' abbreviates 'therefore'.

ciating currency will lead to a narrowing trade deficit but leaves it open whether there are other conditions which would also lead to a narrowing trade deficit, for example, the institution of protectionist trade policies. So the truth of (2) in B does not justify concluding that our currency will lose value: if there *are* other conditions which would lead to a narrowing trade deficit, it may be one of them which accounts for the truth of (2), without our currency having lost value at all. This means that in B, (3) does not follow from (1) and (2).

This discussion of A and B brings out two important points:

- In discerning the difference between A and B we did not rely on any special knowledge about economics. The fact that the conclusion of A follows from its two premises is something that should be evident even to the reader who has no views about how the trade deficit could be narrowed: all that is required is an understanding of the English words 'if…then…' and an ability to recognize the recurrence of the sentences 'our currency loses value' and 'our trade deficit will narrow'.[2] Equally, even if you cannot think of any alternative method of producing a narrowing of the trade deficit, you should still be able to see that nothing in the premises of B rules out the existence of such a method, and that is enough to tell you that the conclusion of B does not follow from its premises.

- In assessing A and B we did not try to settle whether the premises are in fact true, or whether the conclusion is in fact true. Whether or not the conclusion of an argument follows from its premises is a conditional matter, having to do with whether, *if* the premises are true, this by itself guarantees the truth of the conclusion. If the premises of an argument are in fact true and the conclusion in fact false, then of course the premises do not guarantee the truth of the conclusion. But if, say, the premises and the conclusion are all in fact false, the argument may still be valid, and if they are all true, the argument may still be invalid. For example, A is a valid argument, but it is quite conceivable that its premises and conclusion are all false, and B is an invalid argument, though it is equally conceivable that its premises and conclusion are all true.

We can label these two points *topic neutrality* and *independence of actual truth or falsity*. The techniques of logical evaluation which we apply to arguments do not depend on an argument's topic or on whether its premises and conclusion are actually true or false. Their focus lies elsewhere, because the property of an argument which determines whether it is valid or invalid has nothing to do with its topic or with whether its premises and conclusion are

[2] In A and B the sentence 'our currency loses value' does not strictly speaking itself recur, since for grammatical reasons we have to change the tense from future to present. Tense differences are one kind of difference between sentences which are often unimportant from the point of view of sentential logic (others will emerge as we go along). In such cases, we count sentences which differ only as regards tense as having the same content, and we use the same sentence letter for them.

actually true or false. To see what the relevant property of an argument is, compare A and B with the following two arguments:

C: (1) If the Continuum Hypothesis is true, the Zermelo-Frankel universe is constructible.
(2) The Continuum Hypothesis is true.
(3) ∴ The Zermelo-Frankel universe is constructible.

D: (1) If the Continuum Hypothesis is true, the Zermelo-Frankel universe is constructible.
(2) The Zermelo-Frankel universe is constructible.
(3) ∴ The Continuum Hypothesis is true.

In C the conclusion follows from the premises (even though (3) in C is false) while in D it does not. And a comparison with A and B shows that the reason C is valid is the same as the reason A is, and the reason D is invalid is the same as the reason B is. The reader is not expected to know so much as what the premises and conclusion of C and D mean. But in considering D, it should nevertheless be evident that there may be some other condition which would confer constructibility (whatever that is) on the Zermelo-Frankel universe (whatever that is) besides the truth of the Continuum Hypothesis (whatever that is).[3] Hence (1) and (2) do not warrant drawing the conclusion (3).

2 Logical form and validity

If A and C are valid for the same reason, and B and D are invalid for the same reason, this suggests that what makes A and C valid is something they have in common, and what makes B and D invalid is something they have in common. What is common to A and C is their pattern. Each argument has as its first premise an 'if…then…' sentence; in each argument, the sentence following the 'if' in premise 1 is itself premise 2 and the sentence following the 'then' in premise 1 is itself the conclusion. Similarly, B and D have the same pattern, importantly different from that of A and C: the first premise of each is an 'if…then…' sentence, the second premise is the sentence following the 'then' in the first premise, and the conclusion is the sentence following the 'if' in the first premise. We can exhibit the pattern common to A and C by abstracting from the particular English sentences which occur in them. In place of sentences, we use letters 'P', 'Q', 'R' and so on. Then the common pattern of A and C can be written out as:

[3] An excellent source for information on the Continuum Hypothesis and the Zermelo-Frankel universe (and on many other topics not discussed in much detail in this book) is Partee *et al;* see Part A for set theory.

E: If P then Q
P
∴ Q

Similarly, the common pattern of B and D is given by

F: If P then Q
Q
∴ P

When an argument instantiates a certain pattern, that pattern is called the *logical form* of the argument, and it is to its logical form that we look to settle whether an argument is valid or invalid. However, there is a complication. In this book we shall successively present three systems of logic, each of which extends the previous system, beginning with *sentential* logic, the system of Part I. In sentential logic, the basic units are letters like 'P' and 'Q' above, which stand for complete sentences, like the sentences 'the Continuum Hypothesis is true' and 'our trade deficit will narrow' from which the arguments of the previous section were constructed (hence the name 'sentential' logic). Note that we do not allow the use of letters to stand for complete sentences which contain other complete sentences as constituents, connected with expressions like 'if…then…'; for example, we would never represent (1) in D by a single letter.

With a more powerful system of logic, in which the basic units are smaller than complete sentences, we can often give a more discerning description of the form of an English argument than is possible in sentential logic. Thus it is a misnomer to speak of 'the' logical form of an argument; the form is relative to the system of logic under discussion. Hence E and F give what we call the *sentential logical forms* of A through D. Notions like validity and invalidity for English arguments are also relativized to the system of logic, and so at the moment we are discussing just validity in sentential logic, or *sentential validity*. For an argument to be sententially valid is for it to have a valid sentential form. A sentential form is a form whose basic units are letters like the 'P' and 'Q' of E and F. A *valid* sentential form is one such that in any English argument with that form, the conclusion follows from the premises. For the moment, we will assume that with each English argument we can associate just one sentential form, so that it is permissible to speak of 'the' sentential logical form of an English argument; later, we shall see that this is debatable. In addition, we shall often omit the qualifier 'sentential' in Part I of this book, on the understanding that by the logical form of an English argument we mean, in Part I, the form ascribed to the argument in sentential logic by the translation procedures we are going to develop.

A little reflection tells us that any English argument which fits the pattern E, that is, which has the logical form E, will be a sententially valid argument, regardless of its subject matter and regardless of whether its premises and conclusion are true or false. Similarly, any English argument which has the logical form F will be a sententially invalid argument: the first premise is consistent with there being other conditions besides P which would bring about the

truth of Q, so despite the second premise informing us of the truth of Q, there is no warrant for inferring the truth of P.[4]

In making these judgements, we use a relativized notion of validity and a relativized notion of invalidity. However, there is an asymmetry between these notions. If an English argument is sententially valid, then it is valid absolutely; even if we can construct a more finely grained form for it in the logical systems of Parts II and III of this book, those forms will still pronounce the argument valid. This is because the methods of sentential logic remain available in more powerful systems, so if an argument's conclusion has been shown by these methods to follow from its premises, it will still follow in more powerful systems. Thus a positive verdict on an English argument at the level of sentential logic cannot be overturned later. But a negative verdict *can* be overturned: if an argument is sententially *in*valid, it can still turn out to have a valid form in a more advanced system, with more powerful methods for demonstrating that the argument's conclusion follows from its premises (think of telescopes with different magnifications—what can be seen with the less powerful telescope can be seen with the more powerful, but not necessarily vice-versa). The curious reader who would like examples of sententially invalid arguments which are valid in stronger systems will find some at the start of Part II. However, the examples of sententially invalid arguments which are to be found in Part I will not be like those later examples; as well as being sententially invalid, they will be invalid in stronger systems of logic as well, and thus invalid absolutely. In particular, there will be no future redemption for arguments B and D.

Here are two more examples of English arguments, one valid, the other invalid, and so simple as hardly to merit the designation 'argument'.

G: (1) Nixon was president and Agnew was vice president.
(2) ∴ Nixon was president.

H: (1) Either Nixon was president or Agnew was.
(2) ∴ Nixon was president.

In G the conclusion follows from the premise since the truth of the premise guarantees the truth of the conclusion. The sentential logical form of G is

I: P and Q
∴ P

and it is clear that any English argument with this form, be it about politics, economics, mathematics or sport, will be a sententially valid argument and therefore valid absolutely. In H the conclusion does not follow from the premise: if our only information is that one or the other of Nixon or Agnew was president, then we have no warrant for concluding that it was Nixon. The form of H is

[4] A note to the instructor: no argument whose conclusion is a tautology could have the form F since we do not allow the use of a sentence-letter for an English sentence which contains sentential connectives.

J: Either P or Q
∴ P

and again it is clear that any English argument will be sententially invalid if our procedures for determining logical form ascribe the form J to it: if our only information is that one or the other of two basic alternatives is the case, we are not warranted in drawing any conclusion about which it is. Of course, it would be simple to turn H into a valid argument by adding an extra premise:

K: (1) Either Nixon was president or Agnew was.
(2) It was not Agnew.
(3) ∴ Nixon was president.

The sentential form of K is

L: Either P or Q
It is not the case that Q
∴ P

and any English argument with this form will be an absolutely valid argument, since if one of two alternatives holds (premise 1) and we know which one does not hold (premise 2), it must be the other which does hold.

We said above that logic is the study of the difference between valid and invalid arguments. We have now seen that whether or not an argument is valid depends on its form, so we can revise our definition of logic to say that it is the study of the difference between valid and invalid argument-forms. An important point to note is that we are interested in the difference between forms that are *in fact* valid and forms that are *in fact* invalid. This is not quite the same as the difference between what *seems* valid to us and what *seems* invalid, even for the simplest sentential forms. The latter distinction is investigated in cognitive psychology, and there is plenty of experimental evidence that the judgments people make about validity and invalidity do not always reflect the facts. For example, experimental subjects sometimes resist arguments with the form I. It seems that in judging the 'acceptability' of a piece of reasoning, we often use criteria in addition to validity; for instance, we prefer not to draw conclusions which involve a loss of information from the premises, such as occurs in G. This preference has a certain practical utility, but it is irrelevant to the logical issue of whether the conclusion of the argument follows from the premises, and it is only this issue with which we are here concerned. (For more information about psychological aspects of deductive reasoning, see Johnson-Laird and Byrne.)

3 Complications

The examples we have looked at so far have been very simple, and their intuitive statuses as valid or invalid more or less evident to inspection. When we look at more realistic examples of English arguments, there are at least two

ways in which they may be harder to assess than any of the foregoing. One problem is that it may be difficult to determine what the argument *is*, for in ordinary discourse, we rarely make our premises completely explicit, list them in order and clearly demarcate our conclusions. Here is an example of a real argument which leaves the reader a lot of work to do. It occurred at an American Bar Association debate on a motion calling for legislation barring discrimination against homosexuals in employment and housing. The *New York Times* reported one delegate speaking against the motion as making the following argument:

> This is not to me a question of unjust discrimination, because it's not to punish someone for what they are, but for what they do. They don't have to be homosexuals.

Perhaps we can reconstruct the intended argument as follows:

M: (1) If a type of behavior is freely chosen then it is permissible to discriminate against people who behave in that way.
(2) Homosexual behavior is freely chosen.
(3) ∴ It is permissible to discriminate against homosexuals.

Reconstructed in this manner, the argument is clearly valid, so if its premises are true then its conclusion must be true as well. But it is not certain that M captures the speaker's intention—does 'They don't have to be homosexuals' mean that sexual orientation is a matter of choice or just that *manifestation* of sexual orientation is under voluntary control?—and second-guessing what someone may have intended is not really a part of logic: logic is about what makes arguments in general valid or invalid, and psychological speculation about which particular argument a speaker may have been expressing on a given occasion is out of place. So we will not be concerned with this kind of interpretative difficulty in what follows but instead will confine ourselves to arguments which are expressed much more explicitly. The example, however, does help to emphasize the distinction between the questions (a) does the conclusion of the argument follow from its premises? and (b) are the premises true? If the answer to (a) is 'yes', then the argument is valid. If the answer to (b) is also 'yes', then the conclusion must be true as well. Of course, the answer to (b) in the case of M is highly controversial.[5]

It would be useful to have a term for arguments to which the answer to *both* (a) and (b) is 'yes'. Such an argument, one whose form is valid *and* whose premises are true, is said to be *sound*. So we have the following definitions:

> A *valid* argument is an argument with a valid form; concomitantly, an invalid argument is one that does not have a valid form.

[5] Premise 1 could justify religious discrimination. Premise 2 appears to ignore the likely contribution to an individual's sexual orientation of factors beyond his or her control. Aspiring lawyers concerned that law school may be too intellectually demanding should take heart from this example.

> A *valid argument-form* is a form which guarantees that the conclusion of an argument with that form is true if its premises are (refer to E, I and L in §2).
>
> A *sound* argument is a valid argument with true premises.

Hence the conclusion of a sound argument is guaranteed to be true. These definitions are not limited to sentential logic. They capture the intuitive idea of validity which all our systems of logic aim to articulate.

The second way in which arguments may be harder to assess than were our earlier examples is that the complexity of the argument may be much greater. Here is an example which we will meet again:

> If the safe was opened, it must have been opened by Smith, with the assistance of Brown or Robinson. None of these three could have been involved unless he was absent from the meeting. But we know that either Smith or Brown was present at the meeting. So since the safe was opened, it must have been Robinson who helped open it.

Some readers may be able to discern that this is a valid argument, though they may feel unsure of their judgement. Is there any methodical procedure we could follow to determine without question that the argument is valid? Since validity is fundamentally a property of logical forms, we could begin by writing out the sentential form of the argument (this is an exercise in Chapter 2). However, the result will be much more complicated than such forms as I or J or L above. What do we do with a complicated form once we have it?

There are in fact various tests which can be applied to an argument-form to find out if it is valid or invalid, and in later chapters we will develop the techniques on which tests of this sort are based. However, the result of a test on the argument-form gives us information about the English argument whose form we say it is only if we are *right* that the form in question is that of the argument. To begin with, therefore, we need to acquire facility in the representation of arguments, and of English sentences in general, in the kind of symbolic notation to which the argument-forms we have already exhibited approximate.

4 A note to the reader

What can the reader expect to learn from the study of logic? What 'use' is it? Grasp of logic plays a central role in reasoning capacities, and humans seem to be endowed with these to a greater degree than other animals, at least on the average. On the other hand, a sense of humor is also fairly characteristic of humans—again, on the average—so the rationale for studying logic is not just that logic is a distinguishing feature of human nature. The main reason is that the distinction between valid and invalid arguments gives rise to a rich theoretical structure in a way that the distinction between funny and unfunny jokes

does not. And rich theoretical structures are worth studying for their own sake.

There is in addition a more practical aspect of the subject, though one that has only emerged in the last few decades, concerning the role of logic in computation. It is not for nothing that the main component board of the modern personal computer is called a 'logic board'. On the software side, the most prominent application of the principles developed and studied in this book is in a kind of program called an 'expert system'. An expert system contains a database of facts about some domain linked to an inference engine which allows the program to draw conclusions if it is given information about a situation within its domain as premises. Expert systems are used, for example, in medical diagnosis, oil and gas exploration, and financial decision making, in all of which they are generally more reliable than people. The deductive systems studied in this book are abstract inference engines of the sort implemented by such programs.

A natural question to ask about the subjective impact of studying logic is whether working through a book like this will make the reader a more logical person. Some systematic research has been done in this area, with results that are not entirely comforting to logic teachers. On the anecdotal side, however, a student once told the author that after a month of his symbolic logic course, she had found herself able to rout her boyfriend every time they had an argument. She added that she had then lost the boyfriend. Fortunately, she seemed quite pleased with this outcome.

5 Summary

- Logic is the study of the distinction between valid and invalid arguments.
- A valid argument is one whose conclusion follows from its premises.
- Whether or not an argument's conclusion follows from its premises depends on the form of the argument.
- The form we discern in an argument depends on which logical system we have at our disposal.
- We begin with the simplest system of logic, known as sentential logic.
- In sentential logic the forms we discern are called sentential logical forms.
- To test an English argument for sentential validity, we first write out its sentential form and then test that form for validity.
- An argument which is sententially valid is valid absolutely, but a sententially invalid argument may prove to be valid in the framework of a more powerful system of logic.
- While studying this book, readers should be careful not to intimidate their friends unintentionally.