1.4.1. A closer look at entailment

We will begin with a formal definition of the idea of entailment as a conditional guarantee of truth. When a conclusion φ is ***entailed*** by a set Γ of premises, we have a guarantee that φ is true provided that the members of Γ are all true. This is a strong guarantee for it holds, under the stated conditions, in all possible worlds. We can state this definition more formally in two equivalent ways.

$\Gamma \Rightarrow \varphi$	if and only if	there is no logically possible world in which φ is false while all members of Γ are true
	if and only if	φ is true in every logically possible world in which all members of Γ are true

It is worth emphasizing that these are not two different concepts of entailment, for the two statements to the right of *if and only if* say the same thing. Still, there is no redundancy because each of the two emphasizes different aspects of the concept. The second—which we will speak of as the ***positive form***—is closely tied to the motivation for the concept, to the reason why the concept is valuable. The first form—the ***negative form***—makes the content of the concept especially clear, and this form of definition will generally be the more useful when we try to prove things concerning entailment. The other deductive properties and relations we will consider can be given analogous pairs of definitions, a negative form ruling out certain patterns of truth values and another form stating a more positive generalization.

The pattern of truth values ruled out by entailment turns out to be one of the more cumbersome ones to state; and, since we will refer to it often, it will be useful to have special vocabulary for it. We will say that a possible world ***divides*** a set Γ from a set Δ when every member of Γ is true in the world and every member of Δ is false. This use of the word *divide* will prove helpful in a number of ways, but there is one respect in which it may be misleading. Other uses of *divide* point to a symmetric relation, when a is divided from b, b is divided from a. But the specific truth values that must be assigned to sets Γ and Δ for a world to

divide Γ from Δ make this relation between the two sets fundamentally asymmetric. To counteract the suggestion of symmetry you might think of Γ being divided from Δ by being "set above" it, thinking of truth as being "higher" than falsity. As with the premises of an argument, we will have no need, when considering this concept of division, to distinguish between a sentence and a set with that sentence as its only member, so we may regard one or the other terms of the relation of division as a sentence. Using this idea, we can state the negative form of the definition of entailment as follows:

$\Gamma \Rightarrow \varphi$	if and only if	there is no logically possible world that divides Γ from φ

We will say that a possible world divides an argument when it divides its premises from its conclusion, so we can say that an argument is valid when no possible world divides it.

The kind of possible world ruled out by the negative form of the definition must, of course, also have some relation to the positive form. The positive form is generalization concerning all possible worlds of a certain sort. When a generalization is false, it is because of ***counterexample***, something of sort about which we generalize that does not have the property we have said all such things have. A counterexample to the claim that all birds fly is a bird that does not fly. In the positive definition of entailment, the generalization is about all possible worlds in which the premises are all true and such worlds are said to all have the property that the conclusion is true in them. A counterexample to such a generalization is then a world in which the premises are all true but the conclusion is not. Thus a possible world that divides an argument is a counterexample to the claim that its premises entail its conclusion.

It is important to notice how little a claim of entailment says about the actual truth values of the premises and conclusion of an argument. We can distinguish four patterns of truth values that the premises and conclusion could exhibit. Of these, a claim that an argument is valid rules out only the one at the far right of Figure 1.4.1-1.

	Patterns admitted			ruled out
Premises	all **T**	not all **T**	not all **T**	all **T**
Conclusion	**T**	**T**	**F**	**F**

Fig. 1.4.1-1. Patterns of truth values admitted and ruled out by entailment.

So, knowing that an argument is valid tells us about actual truth values only that we do not find the conclusion actually false when the premises are all actually true. The real content of a claim of entailment lies not in what it tells us about the actual world but in the fact that it makes a claim about all possible worlds. The other three patterns all appear in the actual truth values of some valid arguments (though not all are possible for certain arguments).

To see examples of this, consider the case of an argument whose conclusion is among its premises—for example,

Indianapolis is the capital of Indiana
Springfield is the capital of Illinois

Indianapolis is the capital of Indiana

Such an argument is trivial but, because of this, it is obviously valid. Its conclusion certainly does no more than extract information from the premises; and, because it is one of the premises, there is certainly no possible world in which it is false while the premises are all true. Now the example above has true premises and a true conclusion, the first of the patterns in Figure 1.4.1-1. The other two patterns of truth values allowed for valid arguments can be produced by changing *Illinois* and *Indiana*, respectively, to *Ohio*.

Indianapolis is the capital of Indiana
Springfield is the capital of Ohio

Indianapolis is the capital of Indiana

Indianapolis is the capital of Ohio
Springfield is the capital of Illinois

Indianapolis is the capital of Ohio

That these two patterns of truth values should be possible is clear also from the idea of extracting information. Information can be extracted from a set of sentences even though they are not all true, and the information extracted in such a case might be either

true or false.

Of course, seeing one of these permitted patterns does not tell us that the argument is valid; no information that is limited to actual truth values can do that because validity concerns all possible worlds, not just the actual one. In particular, having true premises and a true conclusion does not make an argument valid; the following argument is not valid:

Indianapolis is the capital of Indiana

Springfield is the capital of Illinois

For, although the single premise and the conclusion are both true, there is a logical possibility of the capital of Illinois being different while that of Indiana is as it actually is, so there is a possible world that divides the premise from the conclusion.

Glen Helman 25 Aug 2005