CHAPTER EIGHT: INTRODUCING PROPOSITIONAL LOGIC

1 Basic Semantic Notions

Historically, logic as an object of inquiry started out as an attempt to eludcidate and codify the laws of valid reasoning, that is, to get clear about and make explicit when conclusions follow from their premises. And linguistic semantics is, in large part, concerned with the relationship between the forms of sentences and what follows from them. So logic and semantics are inextricably bound up with each other. In this chapter we introduce basic concepts of logic with a view to semantic applications. (But as we'll see in due course, logic can also be brought to bear in the analysis of linguistic form, not only linguistic meaning.) As good a place as any for us to start is by informally introducing some basic semantic ideas. In very large part, contemporary linguistic semantics originates with Frege's 1892 paper "Über Sinn und Bedeutung (On Sense and Reference)". Among the many important and lasting contributions which Frege made to the foundations of semantics in this paper, we mention just three here: (1) sense vs. reference; (2) the idea that the reference of a declarative sentence is its truth value; and (3) the idea of *compositionality*.

The first and most obvious of Frege's main contributions to semantics was to draw a distinction between the *sense* of a linguistic utterance and its *reference* (also called its *denotation*).[1] We say that an utterance *expresses* its sense and *refers to*, or *denotes*, its reference. For present purposes, the sense of an utterance can be roughly identified with the nontechnical notion of *literal meaning*. (We will have to get clearer about what this amounts to in a later chapter.) Knowing what the sense of an utterance is is part of knowing the particular natural language the expression being uttered belongs to, just as knowing what the sound of the expression is is; as a linguist would put it, it is part of linguistic *competence*. But what the reference of an utterance is is *not* a fact about the language; rather, *the reference of an utterance is jointly determined by the sense of the utterance and contingent facts about the world* (in other words, how things happen to

[1]Frege's influence on linguistics is by no means limited to his 1892 paper. For example, his *Begriffsschrift* (1879) invented propositional logic and predicate logic. Even Frege's mistakes were productive: the discovery by Russell of the paradox that bears his name in Frege's *Grundgesetze der Arithmetik* was one of the central motivations for the development of type theory and its descendants, including typed lambda calculi, higher-order logic, and categorical logic.

be) that are external to the language.[2]

What kind of sense an utterance expresses depends (among other things) on the syntactic category of the expression being utttered. For example, common nouns are usually considered to express *properties*, e.g. *cat* expresses what might be called the property of felinity; proper nouns (e.g. *The Morning Star*, *Miss America*) express what are sometimes called *individual concepts*, and transitive verbs (e.g. *bite*, *love*) express what we might call *binary relationships*. For the moment, the precise inventory of *ontological categories* (kinds of things we assume to exist) is not crucial; what these are and how to model them mathematically is left to semantic theory (which we will come back to in due course). For now the point is just that different categories of expressions tend to be used to express different kinds of senses.

What kind of thing an utterance refers to (denotes) also depends on what kind of sense it expresses. For example, most mainstream semantic theories would say that the reference of *cat* is the set of all cats, which is a very different thing from the sense (the property of felinity). However, the reference is *determined by* the sense together with contingent facts about how the world happens to be (in this case, which things actually possess the property of felinity). (I'm not a cat, but I might have been, a possibility which I contemplate on occasion.) Similarly, the individual concept expressed by *Miss America* is quite a different thing from the individual it would most likely refer to in an utterance at the time I am writing, given the contingent fact of who actually won the Miss America Paegent most recently, namely Kirsten Haglund of Farmington Hills, Michigan.

What about the sense and reference of declarative sentences? Consider the sentence *I'm hungry.* If I utter it at 9:11pm Sunday November 3, 2008, the sense it expresses is what philosophers and semanticists call a *proposition*, namely the proposition that Carl Pollard is hungry at that particular time. That is quite a different proposition than the one expressed if President Elect Barack Obama utters the same sentence the following Thursday day at 12:30pm. This (and similar facts) is why we speak of the sense expressed by an *utterance* of an expression, not by the expression itself. In fact utterances of *different* sentences can express the same proposition (if

[2]Nonlinguist readers need to be alert to the distinction between linguistic expressions, which belong to the natural language in question, and *utterances* of expressions, which are specific events. As we'll see later, the sense of an expression can be highly dependent on various aspects of the *utterance context*, which necessitates distinguishing between the (context-independent) *meaning* of an expression and the (context-dependent) *sense* or *interpretation* of a particular utterance of that expression, but this is a subtlety we will ignore for the time being.

uttered in the same context):

(1) a. A motorcycle ran over my ferret.
  b. My ferret was run over by a motorcycle.
  c. My ferret, a motorcycle ran over.
  d. What ran over my ferret was a motorcycle.
  e. It was a motorycle that ran over my ferret.

Utterances of these sentences certainly differ in terms of what is emphasized and what is not, what is presented as new information and what as old, and various other facets of the utterance context (such factors fall within the purview of the linguistic subdiscipline of pragmatics); but spoken by a given one-ferret-owning speaker at a given time, all could express the same proposition. Indeed, utterances of sentences from *different languages* can express the same proposition:

(2) a. I'm hungry.
  b. J'ai faim. [French]
  c. Ich habe Hunger. [German]
  d. Wo du-zi e-le. [Chinese]

The second of Frege's two contributions to semantics we want to call attention to here is the idea that the reference of a declarative sentence utterance is its truth value (either true or false). This may seem an odd choice, but it is actually quite natural because after all, what else about declarative sentence utterances is jointly determined by the sense (in this case, the proposition expressed) and contingent fact (how the world is, or, in other words, which propositions are true and which are false)?

In this chapter we will focus on propositions, how to model them, and how to theorize about them, a subject which is usually called *propositional logic*. But later, we will need to consider not just senses and references of declarative sentences, but also of smaller phrases and even words; and then, we will have to grapple with the question of how language users determine the senses of sentences given that the lexicon specifies the senses of the words. At that point, we will lean heavily a version of the third of Frege's crucial contributions to semantics–the principle of *compositionality*—which (in the form we will adopt) holds that the sense of an utterance is a function of the senses of its immediate constituents.[3] Assuming for now that phrases are

[3]In Frege's original formulation, additionally the *reference* of an utterance was held to be a function of the *reference* of its immediate constituents, which—as we will see—leads to needless complications.

put together from words using context-free rules of the kind we considered in Chapter 6, another way to say this is that each phrase structure rule (PSR) not only combines expressions of certain categories to form expressions of another category, but additionally is associated with a specific function that takes the senses of the immediate constituents as arguments and returns the sense of the resulting expression as value.

2 Entailment

For any two propositions p and q, the most basic question we can ask about them is whether p *entails* q. To say that p entails q is to say that, independently of contingent fact—no matter how things happen to be—if p is true with things that way, then so is q. Sometimes people speak loosely of one utterence entailing another (or even more loosely of one sentence entailing another), but we will be at pains to reserve the term *entailment* for a relation between propositions, and *following from* for the corresponding relation between declarative sentence utterances. That is, for utterances u and v, we say v follows from u to mean that the proposition expressed by v is entailed by the proposition expressed by u. More generally, we say that a proposition p follows from a *set* P of propositions just in case, no matter how the world is, if all the propositions in P are true with the world that way, then so is p. Correspondingly, we say that an utterance (called the *conclusion*) follows from a set of utterances (called the *premisses*) just in case the proposition expressed by the conclusion is entailed by the set of propositions expressed by the premisses.

Now let's start to develop a theory about propositions. In due course we'll extend this theory to cover other kinds of senses (such as properties and individuals). And after that, the next step will be to develop a theory of the connection between the forms of utterances (how the expressions uttered are built up from their constituents using grammar rules) and the senses they express. (This connection is often referred to as the *syntax-semantics interface.*) And later still, we will have to consider the contribution that the context of utterance makes to its interpretation, as well as the the potential that utterances have to *change* various aspects of the context.

To begin, suppose we have a set P whose members are (our theoretical representations of) propositions. Then you should be able to persuade yourself after a little reflection that the entailment relation should be modelled by a preorder. (Why?) However, there is disagreement among philosophers and semanticists as to whether entailment is antisymmetric. To see why, consider propositions whose truth value is independent of how the world is. These come in two flavors, the *necessary truths*, which are true no matter what, and the *necessary falsehoods*, which are false no matter what. It's easy to see that the necessary truths are the tops in the entailment preorder and the necessary falsehoods are the bottoms. Propositions which are neither necessarily true nor necessarily false (i.e. neither tops nor bottoms with respect to entailment) are called *contingent*. Now consider two necessary truths, say the ones expressed by (utterances of) the two sentences *all donkeys are donkeys* and *all horses are horses*. Clearly these entail each other (any way things are in which one is true, namely any way whatsoever, is also a way the other is true), yet it is not at all clear that they are the same proposition! It is also possible for two not-clearly-identical propositions to entail each other, even if they are contingent, e.g. *I'm hungry* and *I'm hungry, and either France is shaped like a hexagon or it isn't*, or *Phil is a groundhog* and *Phil is a woodchuck*.

Mutually entailing propositions are called (*truth-conditionally) equivalent*. It's easy to see that equivalence is an equivalence relation on propositions. So whether entailment is antisymmetric comes down to whether truth-conditional equivalence is just identity of propositions. We will keep an open mind about this, by just assuming entailment is a preorder without committing as to whether it is an order.

3 Formulas: the Syntax of Standard Propositional Logic

.

Standard propositional logic (PL) is a particularly simple idealization of the syntax and semantics of a fragment of English[4] which is essentially concerned only with the "logic words" *it-is-not-the-case-that*, *and*, *or*, *implies*, and *iff*. PL ignores the distinction beween sentences and utterances. It also ignores the internal structure of "atomic" sentences (that is, ones that were not built up from smaller sentences using the logic words), which are simply represented using a fixed set of symbols $\{A, B, \ldots\}$ called **propositional letters**. (As we'll soon discuss, these are taken to represent arbitrarily chosen sentences that express contingent propositions.) Moreover, PL does not distinguish between distinct sentences that represent neccessary truths (or necessary falshoods).

The set of PL **formulas** is then defined recursively as follows:

[4] Or, to be more precise, of the "Mathese" dialect.

- (The length-one string corresponding to) each propositional letter is a formula (these formulas are called **atomic**);
- T and F are formulas;
- if ϕ and ψ are formulas, then so are $(\sim \phi)$, $(\phi \wedge \psi)$, $\phi \vee \psi$, $(\phi \rightarrow \psi)$, and $(\phi \leftrightarrow \psi)$;[5]
- nothing else is a formula.

4 Traditional PL Semantics

Here we briefly consider the standard way of assigning "senses" to PL formulas. In later chapters, we will look at alternatives that are better adapted to modelling the complexities of the syntax and semantics of natural language and the interface between them.

We start by defining a **(PL-)assignment** to be a function a from the set of atomic formulas to the set $2 = \{0, 1\}$. We think of 1 and 0 as (respectively) the truth values 'true' and 'false', and in this context that is what we will usually call them. Intuitively, we can think of an assignment as a complete description of one particular way things might be, or what some philosophers (following Carnap) have called a *state description*. The intuition is that if $a(A) = 1$ (respectively, 0) then a represents one of the ways things might be in which A is true (respectively, false).

It is easy to show that, for any assignment a, there is a unique function $\bar{a}$ from formulas to truth values satisfying the following conditions:

a. For each atomic formula A, $\bar{a}(A) = a(A)$;

b. $\bar{a}(T) = 1$ and $\bar{a}(F) = 0$;

c. for any formula ϕ, $\bar{a}(\sim \phi) = 1$ iff $\bar{a}(\phi) = 0$;

d. for any two formulas ϕ and ψ, $\bar{a}(\phi \wedge \psi) = 1$ iff $\bar{a}(\phi) = 1$ and $\bar{a}(\psi) = 1$;

e. for any two formulas ϕ and ψ, $\bar{a}(\phi \vee \psi) = 1$ iff $\bar{a}(\phi) = 1$ or $\bar{a}(\psi) = 1$;

f. for any two formulas ϕ and ψ, $\bar{a}(\phi \rightarrow \psi) = 1$ iff $\bar{a}(\phi) = 0$ or $\bar{a}(\psi) = 1$;

g. for any two formulas ϕ and ψ, $\bar{a}(\phi \leftrightarrow \psi) = 1$ iff $\bar{a}(\phi) = \bar{a}(\psi)$.

[5] But we usually omit outermost parentheses when no confusion will result.

A function, such as $\bar{a}$, that is obtained by extending an assignment in this way, is called a **valuation**. Equivalently, we can simply define a valuation to be a function v from formulas to 2 that satisfies the conditions (b) through (g) above; then it is easy to see that v is uniquely determined by the values it assigns to atomic formulas.

Now we're ready to explain the traditional semantics of PL. A valuation v is said to **satisfy** a formula ϕ, written $\models_v \phi$, if it assigns the value 1 (true) to it. Let V be the set of valuations, and let Prop be its power set. We call its members (sets of valuations) **propositions**. We now define the function sem from formulas to propositions as follows: if ϕ is a formula, then $\mathsf{sem}(\phi)$ is the set of valuations that satisfy it. But it should be obvious on a moment's reflection that Prop is in one-to-one correspondence with the set of functions from assignments to truth values, and therefore we can think of $\mathsf{sem}(\phi)$ as picking out all the various ways things might be in which ϕ is true.

We model entailment by the familiar subset inclusion relation on $\mathsf{Prop} = \wp(\mathsf{V})$, i.e. p entails q if every valuation in p is also in q. (Notice that with this way of modelling propositions and entailment, entailment is antisymmetric!)

Now for any two formulas ϕ and ψ, we say that ψ **follows from** ϕ, or equivalently, that ψ is a **(semantical) consequence** of ϕ, written $\phi \models \psi$, iff $\mathsf{sem}(\phi)$ entails $\mathsf{sem}(\psi)$. Equivalently: ψ follows from ϕ iff every valuation that satisfies ϕ also satisfies ψ.

More generally, if Γ is a set of formulas, then we say v **satisfies** Γ, written $\models_v \Gamma$, provided it satisfies every formula in Γ; and if additionally ϕ is a formula, we say ϕ **follows from** Γ (or equivalently, is a **(semantical) consequence** of Γ), written $\Gamma \models \phi$, iff every assignment that satisfies Γ also satisifes ϕ.

Some further traditional semantic terminology for PL runs as follows: a formula is called **satisfiable** if some valuation satisfies it, **unsatisfiable** otherwise. Another name for an unsatisfiable PL formula is a **contradiction**; a formula satified by *any* valuation is called **valid**, or a **tautology**. And two formulas which follow from each other (have each other as semantical consequences) are called **logically equivalent**. Notice that for a formula ϕ to be valid is the same as $\Gamma \models \phi$ where Γ is the empty set of formulas, i.e. $\emptyset \models \phi$. This in turn is easily seen to be equivalent to $T \models \phi$. This is usually written as simply $\models \phi$.

5 A Test for Validity: Truth Tables

Suppose Γ is a finite set of formulas (premisses) and ϕ is a formula (the conclusion). How can we tell whether the conclusion follows from the premisses? Well, if Γ is empty, we've already seen this is the same as asking whether ϕ is valid. Otherwise, suppose $\Gamma = \{\phi_0, \ldots, \phi_{n-1}\}$ for some $n > 0$. Then it is not hard to show that $\Gamma \models \phi$ iff

$$\models (\ldots(\phi_0 \wedge \phi_1)\ldots \wedge \phi_{n-1}) \rightarrow \phi$$

In other words, the problem of checking whether a conclusion follows from a finite set of premisses can always be reduced to checking whether some formula is valid.

But how do you do that? Well, for ϕ to be valid means for every valuation to map it to true; but there are only as many valuations as there are assignments, and when we consider assignments, clearly we really only have to consider the truth values they assign to the atomic formulas that actually occur as subformulas of ϕ. Thus, if ϕ contains exactly two different atomic formulas (not occurrences of atomic formulas), say A and B, then there are only four different assignments from $\{A, B\}$ to 2 that we have to worry about.

A **truth table** is just a table that displays, for a PL formula ϕ and the subformulas it was recursively constructed from, and for each assignment of truth values to the atomic formulas that occur in ϕ, what value the valuation that extends that assignment assigns to ϕ and to each of its subformulas. The following is a simple example:

```
    Truth Table for (A /\ B) -> A

-----------------------------------
| A | B ||  A /\ B | (A /\ B) -> A|
|___|___||_________|______________|
|---|---||---------|--------------|
| T | T ||    T    |      T       |
|---|---||---------|--------------|
| T | F ||    F    |      T       |
|---|---||---------|--------------|
| F | T ||    F    |      T       |
|---|---||---------|--------------|
| F | F ||    F    |      T       |
-----------------------------------
```

Note that, in the table, each row below the double horizontal line corresponds to an assignment of truth values to the atomic subformulas of the formula being checked (and therefore, to an equivalence class of valuations that agree with each other on these atomic formulas); the columns to the left of the double vertical line correspond to the atomic subformulas themselves; the rightmost column to the formula itself, and all the other columns to the nonatomic subformulas. In this case, we see that all the cells in the rightmost column below the horizontal double line are filled with T (true) rather than F (false), and can therefore conclude that $(A \wedge B) \rightarrow A$ is valid.