Jason Bridges
Phil 340, winter 2002

How do we construct theories of meaning that satisfy the compositionality constraint and Convention T?

Recall Davidson's two constraints on the *form* of a theory of meaning for a natural language:

> **The compositionality constraint**: A theory of meaning of a natural language L must show how the meanings of sentences of L are determined by properties of the simple expressions composing the sentences, coupled with the order in which the expressions appear. More rigorously, such a theory must consist of the following parts:
>
> 1. A set of axioms that assign semantic properties to each of the simple expressions of L.
> 2. A set of axioms that specify how the meanings of complex expressions are determined on the basis of the semantic properties of the simple expressions composing them.
> 3. A set of theorems, implied by these axioms, that give the meanings of all possible sentences of L.
>
> **Convention T (Davidson's version).** The theorems of a theory of meaning for an object language L must take the form: **S is true iff p** (where 'S' is replaced by an expression in the metalanguage referring to a sentence in the object language, and 'p' is replaced by a sentence in the metalanguage that is true iff the object-language sentence is true.)

[*Note:* Davidson imposes a third constraint on theories of meaning, derived from his views about the *evidence* proper to such theories; we are deferring discussion of that constraint for now.]

How do we go about constructing theories that satisfy these two constraints?

First sample theory: language E1 (sentences)

Consider the artificial language E1. E1 consists of two expressions: "Bob is in China" and "Sally is clever", which are intended to have the same meaning as the comparable sentences in English. [Pretend, to avoid complications, that there is only one person with the English name "Bob", and only one with "Sally".]

Here's a theory of meaning for E1 (the metalanguage of the theory will, of course, be English):

> <u>Axiom 1</u>: "Bob is in China" is true iff Bob is in China.
> <u>Axiom 2</u>: "Sally is clever" is true iff Sally is clever.

This theory satisfies Convention T: just think of the axioms as theorems (which we can do, since every sentence implies itself). It also satisfies the compositionality constraint. But that constraint has no teeth here, for there are no complex expressions in the language. There are two simplex expressions, which can't be combined to form a complex expression.

Second sample theory: language E2 (truth-functional connectives)

E2 consists of E1 plus two further expressions, "or" and "not–", which are meant to function just like "or" and "it is not the case that" do in English. [*Note for logic enthusiasts:* "or" and "not–" are *truth-functional connectives*. A truth-functional connective is an expression that combines with whole sentences to produce a new sentence whose truth-value is a function of the truth-values of the original sentences.]

We now have a language that can produce an infinite number of sentences: "not–Sally is clever", "Bob is in China or Sally is clever", "Bob is in China or not–Sally is clever", "not–(Sally is clever) or (Sally is clever or not–not–Bob is in China)" and so on. [*Note:* The parentheses in the last case are needed to avoid ambiguities. For ease of exposition, we'll ignore this complication henceforth. In English various devices are used to avoid comparable ambiguities. Compare: "It's neither the case that Bob is in China nor that Sally is clever" and "Either it's not the case that Bob is in China, or Sally is clever."]

How do we state axioms that can generate truth-conditions for an infinite number of sentences? The solution is the mathematical technique of *recursion*. In the simplest form of recursion, we generate a sequence by defining each item in the sequence in terms of the immediately preceding item. Consider this recursive definition of exponent (or 'power') functions:

R1: $z^0 = 1$

R2: $z^{n+1} = z^n * z$

Here, we've defined z^n for every possible n, and we've done it in only two clauses. The crucial maneuver lies in R2, which gives us a recipe for determining z^{n+1} so long as we know z^n. Suppose we want to know what 2^3 is. The second clause tells us that $2^3 = 2^2 * 2$. But what's 2^2? Well, the second clause also tells us that $2^2 = 2^1 * 2$. Combining, we get that $2^3 = (2^1 * 2) * 2$. But what's 2^1? Well, the second clause also tells us that $2^1 = 2^0 * 2$. Thus $2^3 = ((2^0 * 2) * 2) * 2$. But what's 2^0? Well, the first clause tells us that it's 1. So $2^3 = ((1 * 2) * 2) * 2$. That's 8.

The same general technique can be applied in theories of meaning. Consider this theory of meaning for E2:

Axiom 1: "Bob is in China" is a sentence and is true iff Bob is in China.

Axiom 2: "Sally is clever" is a sentence and is true iff Sally is clever.

Axiom 3: For any sentence S, "not–" + S is a sentence and is true iff S is not true.

Axiom 4: For any sentences S and R, S + "or" + R is a sentence and is true iff S is true or R is true.

It can be proven that this theory generates correct T-sentences for every possible sentence of E2. Let's test it by determining what T-sentence it generates for the following sentence of E2: "Sally is clever or not–Bob is in China."

a) Axiom 2 tells us that "Sally is clever" is a sentence, and axioms 1 and 3 together tell us that "not–Bob is in China" is a sentence. Hence by axiom 4, "Sally is clever or not–Bob is in China" is true iff "Sally is clever" is true or "not-Bob is in China" is true.
b) By axiom 3, "not-Bob is in China" is true iff "Bob is in China" is not true.
c) By axiom 1, "Bob is in China" is not true iff Bob is not in China.
d) Putting together b) and c), "not-Bob is in China" is true iff Bob is not in China.
e) By axiom 2, "Sally is clever" is true iff Sally is clever.
f) Putting together a), d) and e), "Sally is clever or not–Bob is in China" is true iff Sally is clever or Bob is not in China.

This is just the result we wanted.

Third sample theory: language E3 (names and predicates)

But suppose we want a language with still more flexibility. We define a language E3, which is like E2 except that "Bob", "Sally", "is clever", and "is in China" are now to be understood as independent semantic units, free to combine with other expressions in the ways permitted of the comparable expressions in English. Thus, unlike E2, E3 will allow the sentences, "Sally is in China" and "Bob is clever".

How are we to understand the contributions of these expressions to determining the truth-condition of sentences?

Proper names: Consider first "Sally" and "Bob", which are *proper names*. Clearly, one job of a proper name is to stand for (equivalently, to refer to) a particular object. "Bob" stands for Bob; "Sally" refers to Sally.

Moreover, it seems plausible that its referring to a particular object is *all* that a proper name contributes to the determination of the truth conditions of sentences in which it appears.

The evidence for this hypothesis is the following fact: if "Bob is in China" is true, then any sentence of the form "x is in China", where "x" is replaced by an expression that refers to the same object as "Bob", *must* also be true. If Bob is also known as "Noodles", and it's true that Bob is in China, it must be true that Noodles is in China. This suggests that the contribution these two names make to their respective sentences is simply to refer to Bob.

We may state the idea as follows: Proper names are **referring expressions**, where a referring expression is an expression whose contribution to determining the truth-condition of a containing sentence is exhausted by the following property: the expression refers to (stands for) a particular object.

Thus we are lead to the following axioms for E3:

Axiom 1: "Bob" is a name, and refers to Bob.

Axiom 2: "Sally" is a name, and refers to Sally.

Note: "name" is not meant in these axioms to have any special meaning. It simply serves as a label. This is required for axioms 1 and 2 to interact with axioms 3 and 4 below in the desired way. Comparable axioms could be constructed without using such labels so long as 3 and 4 below were also modified, but it's a bit more perspicuous to do it this way.

Predicates: What about "is clever" and "is in China"? Can we view them as referring expressions as well? It is of fundamental importance for understanding how language works to see that the answer is no.

The problem is not with the idea that such expressions refer to things. We can say, if it pleases us, that "is clever" in some sense or other refers to an entity, call it cleverness. The problem is with the idea that the *sole* contribution of a predicate to determining truth-conditions is its referring to something. Such a view has the implication that a sentence like, "Bob is clever" amounts simply to a *list* of objects (in this case, two). It could then equally well be written, "Bob, cleverness". But "Bob, cleverness", precisely because it is just a list, does not *say* anything. It cannot be used to assert that something is *so*. Equivalently, it has no truth-condition.

This phenomenon—that the subject and predicate of a sentence unite to form a whole that is more than a list—is sometimes known as *the unity of the proposition*. It was one of the major preoccupations of early analytic philosophy (the period of Frege, Russell and early Wittgenstein. See, for example, Russell's *Principles of Mathematics*, §54.)

We need axioms for "is clever" and "is in China" that will acknowledge the unity of the proposition. One simple and perfectly legitimate way of doing so is to pass the buck into the metalanguage:

Axiom 3: Where N is a name, N + "is clever" is a sentence, and is true iff the object referred to by N is clever.

Axiom 4: Where N is a name, N + "is in China" is a sentence, and is true iff the object referred to by N is in China.

The remaining axioms of E3 are simply copies of axioms 3 and 4 from the theory of meaning for E2 (renumbered as 5 and 6):

Axiom 5: For any sentence S, "not–" + S is a sentence and is true iff S is not true.

Axiom 6: For any sentences S and R, S + "or" + R is a sentence and is true iff S is true or R is true.

Exercise: Derive the T-sentence for "Sally is in China or not–Bob is clever", showing all steps.

Fourth sample theory: language E4 (quantifiers and definite descriptions)

With E4, we add two quantifiers—"all" and "some"—and an operator for forming definite descriptions—"the".

The tackling of quantifiers within an axiomatic truth-conditional theory of meaning requires the introduction of technical apparatus that goes beyond the boundaries of this course. (For the historical source of the apparatus in question, see Tarski's discussion of the notion of *satisfaction* in "The Semantic Conception of Truth".)

Instead, we will discuss the relationship between quantifiers and definite descriptions in an informal way. It's an excellent illustration of the kind of issue that becomes of central importance on Davidson's view of what a theory of meaning ought to do.

A *quantifier* like "all" or "some" combines with other expressions to form a *quantifier phrase:* "All men", "Some of the people I used to date", "Most doughnuts", "Few of the members of the APA".

These phrases, just like proper names, fit into the subject-place, rather than the predicate-place, of sentences: "Most doughnuts are in China." But it should be obvious from the examples I have given that, unlike proper names, quantifier phrases cannot be regarded as referring expressions. "All human beings" does not contribute to the truth condition of "All human beings are mortal" by referring to some specific object. "All human beings are mortal" is true iff *everything* that is a human being is mortal. Or consider, "Some Texans can dance." This sentence may plausibly be said to be true iff *at least one* Texan can dance. But there is no *particular* Texan who must be able to dance in order for the sentence to be true.

Quantifier phrases embody a distinct semantic category, neither referring expressions nor predicates.

Now consider *definite descriptions*, formed in English by combining "the" with other expressions: "the inventor of the pencil", "the house atop Haunted Hill", "the smallest prime number".

Frege regarded definite descriptions as referring expressions. On the face of it, this is not implausible. The "The inventor of the pencil can dance", like "Some inventors of the pencil can dance", requires for its truth that *at least one* inventor of the pencil can dance. But unlike "Some inventors of the pencil can dance", it also requires that there be *no more than one* inventor of the pencil. (If the pencil were invented by a committee, then even if everyone on the committee could dance, it would not be true that *the* inventor of the pencil can dance.) The joint requirements of *at least one* and *no more than one* seemed to Frege to ensure that a definite description does the work of a referring expression, the work of *referring* to a *particular* object.

Since Russell's seminal work on this issue (see his "On Denoting"), Frege's view that definite descriptions are referring expressions has come to seem increasingly untenable. The issues are complex; here is one of them:

<u>Empty referring expressions vs. non-denoting descriptions</u>: When a purported referring expression, such as a proper name, has nothing that it refers to, it is said to be empty. What is the truth-condition of a sentence that contains such an expression? The answer would seem to be that *it has no truth condition at all.* Not that the sentence is always false, but that it is neither true nor false, and so cannot be used to say anything.

Suppose you think you overhear me speak of someone named "Flug". Later you remark to a friend, "Bridges likes Flug." In fact you misheard me; I was merely clearing my throat. There is no "Flug". What is the truth condition of your utterance? The right answer seems to be that it doesn't have one. What you said isn't false; it's contentless.

This result is reflected in our theory of meaning by the fact that the contribution of a referring expression to truth conditions is given by an axiom stating what that expression refers to. If there is nothing that "Flug" refers to, no such axiom can be truly stated. Hence no T-sentences can be derived for sentences containing "Flug".

Contrast the case of *non-denoting* definite descriptions, that is, descriptions that nothing in fact satisfies. Suppose you say, "I had breakfast with the present king of France today." The present king of France does not exist. If "the king of France" were a referring expression, we would have to conclude, parallel to the "Flug" case, that what you said had no truth condition. We would have to conclude that you were not simply wrong, but failed to so much as make an assertion at all.

But this seems strikingly implausible. Surely you said *something*; it's just that what you said was false. Note that on the view that definite descriptions are referring expressions, I wouldn't be saying something true if I respond, "You didn't have breakfast with the present king of France today." Note that on this view I can't even truly say, "The present king of France does not exist." But I just did so in the previous paragraph.

(For a contrasting view on the issue of non-denoting definite descriptions, see Strawson's "On Referring".)

(There is a further argument against viewing definite descriptions as referring expressions, having to do with differences in how the two kinds of expression behave in contexts involving *possibility*. We shall discuss this matter in depth when we read Kripke.)

Russell suggests, on the basis of grounds of this sort, that definite descriptions should be placed in the category of quantifier phrases, rather than of referring expressions. And he suggests an ingenious treatment of their form that shows how such an assimilation can be effected. The formal part is beyond the bounds of this course, but the idea can be stated informally: "The F is G" is true iff there exists some x such that the following three conditions hold: x is F, everything that is F is identical to x, and x is G. Post-Russellian logical advances have led to streamlined versions of this approach.