Discourse Representation Theory

Daniel Bonevac

April 24, 2012

Fourteenth-century logicians such as Burley and Buridan note that quantifiers not only express relations between terms but also introduce items in discourse that can be referred to anaphorically. The problem this raises is not one of truth conditions, as we have seen—it is not difficult to give first-order representations of sentences such as ‘an animal is running and it is a man’ or ‘Every farmer who owns a donkey beats it’—but rather of compositionality. The problem, in other words, is typically not a lack of appropriate truth conditions but a rule-governed way of generating them.

In Aristotelian logic, we can represent ‘an animal is running’ as ‘Some A is R,’ but we then have no way of indicating that the running animal is a man. Appending another categorical proposition will not help. In first-order logic, similarly, we can represent ‘an animal is running’ as \( \exists x (Ax \land Rx) \). Having done that, however, we have no way of adding another formula representing that the \( x \) in question is a man. What we have to do in both cases is extend the original representation, writing ‘Some A is R and M’ or \( \exists x (Ax \land Rx \land Mx) \). Since the anaphora occurs within the same sentence, we could perhaps write rules requiring this. But similar anaphora can occur across sentential boundaries. Buridan’s example could just as easily have been ‘an animal is running. It is a man.’ These could moreover be separated by intervening discourse: ‘an animal is running. People look on in surprise. It is uncommon to see such a sight in the square, amidst the crowds, in the midday heat. It is a man.’ This makes a strategy of waiting to close off the representation of the first sentence until all later anaphors have been collected implausible.

Sometimes, the problem is a lack of appropriate truth conditions. Consider this discourse:

Mary: “A man fell over the edge!”

John: “He didn’t fall; he jumped.”

A quantificational analysis leads to the formula \( \exists x (x \text{ is a man} \land x \text{ fell over the edge} \land x \text{ didn’t fall} \land x \text{ jumped}) \) for the discourse, which is contradictory. This seems appropriate enough, for the second assertion contradicts the first. But we have no way of telling whether the second assertion, considered alone, is true or false, for it receives no independent truth conditions. Representing just John’s assertion would yield \( \exists x (x \text{ didn’t fall} \land x \text{ jumped}) \), which is true if anything jumped and didn’t fall. Nothing there ties the assertion to the person spoken of by Mary.

The problem posed by so-called donkey sentences is thus quite general. Quantified noun phrases not only relate sets or relations on a universe of discourse but also make
subsequent anaphoric connection to certain things or sets possible.

Lauri Karttunen (1976) hypothesizes that quantified noun phrases introduce discourse referents, items to which later anaphoric elements can link. Hans Kamp (1981) and Irene Heim (1982) develop that hypothesis into a theory, which has become known as Discourse Representation Theory (DRT). An appearance of a noun phrase (hereafter, NP) that is indefinite, that is, of the form \( a(n) \ F \), establishes a discourse referent if and only if it justifies the occurrence of a coreferential pronoun or definite NP later in the text. For example, in example (1), the indefinite NP \( a \ man \) establishes a discourse referent. It justifies the coreferential pronoun \( he \) in the second sentence, in the sense that the second sentence cannot occur without the first in the absence of something else that would justify the occurrence of such a pronoun—a linguistic context providing other possibilities of anaphoric linkage, for example, or an act of demonstration such as pointing that would make the pronoun deictic rather than anaphoric.

Karttunen’s notion, unlike most traditional syntactic and semantic concepts, is multisentential. It applies readily to anaphora across sentential boundaries. It is also procedural. Karttunen defines not what a discourse referent is but what it takes to establish one. He thus analyzes indefinite NPs not in terms of what they stand for but in terms of what they do. Dynamic semantics extends this approach to language in general. It sees a sentence as a way of transforming one context into another.

NPs such as ‘an animal’ do not, however, always establish discourse referents in Karttunen’s sense of licensing further anaphora. In particular, when they occur within the scope of quantifiers or negations, they do not permit further anaphoric links:

\[
\text{Every farmer owns a donkey. *I feed it sometimes.}
\]
\[
\text{It is not true that an animal is running. *It’s a man.}
\]

An adequate theory based on Karttunen’s approach, then, should explain how and when NPs license anaphora. Discourse Representation Theory (DRT) does this in two steps. First, it analyzes indefinites and other NPs with \( \text{?mon}\) determiners—those fitting the pattern of Aristotle’s particular quantifiers—as introducing discourse referents in every case. Only sometimes, however, are those referents accessible to later anaphors. Second, it specifies formally a relation of accessibility that determines when anaphoric connections are possible. That relation depends crucially on the Discourse Representation Structure built from the discourse.

Kamp (1981) presents the essentials of Discourse Representation Theory. The general form of the theory is to process a discourse in two steps. First, it parses a sentence syntactically and applies an algorithm to construct a Discourse Representation Structure (DRS). Second, DRSs receive truth conditions by way of a model-theoretic semantics.

The intermediate level of DRSs, or semantic representations, is the key to the theory’s novelty. The DRS represents both context and content. Think of prior discourse, for example, as having built up an initial DRS. Then a sentence, processed by the DRS construction algorithm, transforms that initial DRS, acting as content of the previous discourse and context of the current sentence, into a new DRS, representing the content of the entire discourse and the context of further utterances. At each stage of the construction, the DRS determines the truth conditions of the discourse up to that point. But
it also constitutes a conceptual level that provides information that cannot be recovered from truth conditions alone.

In short, then, Discourse Representation Theory treats indefinites and other NPs with similar logical properties as referring expressions at the level of the DRS construction algorithm and as quantifiers at the level of the truth definition. Pronouns do not introduce independent elements into a DRS; they refer to items already there. They select their referents from sets of antecedently available entities. Deictic pronouns do so from the real world; anaphoric pronouns select from constituents of the representation—Karttunen’s discourse referents. The theory tries to specify sets of referential candidates, by specifying which entities are accessible to a given anaphor. Strategies for selecting referents from among the set of possible referents are complex, relying on semantic, pragmatic, and discourse factors. The theory itself does not spell out these strategies, but is easily supplemented with them.

The DRS construction algorithm provides rules for building or altering DRSs, given the syntactic parse of a sentence—specifically, an analyzed surface structure. To see how it works, let’s begin with ‘An animal is running.’ I will ignore tense, aspect, and other complications; the following DRS would be a component of a richer structure incorporating this additional information.

A DRS \( m = \langle U, C \rangle \) consists of a universe or domain and a set of conditions. The universe is a set of entities to be thought of as discourse referents; the conditions provide information about those referents. (For convenience, I shall write a DRS \( \langle \{x_1, ..., x_n\}, \{C_1, ..., C_m\}\rangle \) as \( \langle x_1, ..., x_n : C_1, ..., C_m \rangle \). One can thus think of a DRS as an information state or as a partial model, a representation of a situation or part of a world.

Indefinite NPs introduce discourse referents into the universe. They also introduce conditions—entries providing information about discourse referents—expressing the content of the description. ‘An animal,’ then, introduces a discourse referent and a condition saying that it is an animal. Names and personal pronouns also introduce discourse referents and, often, additional conditions.

Finally, verbs and adjectives introduce conditions. Applying the construction algorithm to An animal is running, in the null context, produces the following DRS: \( \langle u : \text{animal}(u), u \text{ is running} \rangle \). This DRS consists of a domain \( \{u\} \) and a set of conditions: \( \{\text{animal}(u), u \text{ is running}\} \), which express or simply are properties and relations among objects in that domain. It represents the content of ‘An animal is running’ and acts as a context for ‘It’s a man.’ The anaphoric pronoun \( it \) must refer to a discourse referent already introduced. We might, as soon as we reach it in the construction algorithm, search for that referent and use it in the conditions to be introduced. To separate the problem of searching among the possible referents for the actual one, however, it is more convenient to introduce another discourse referent, together with a condition identifying it with a previously introduced referent. So, we obtain a DRS of this form: \( \langle u, v : \text{animal}(u), u \text{ is running}, \text{man}(v), v = ? \rangle \). What are the possible referents of \( v \)? The theory states that, in the absence of an act of demonstration, all candidates are discourse referents previously introduced. Assuming the earlier DRS to be the entire context for the utterance, nothing but \( u \) is available. Identifying \( v \) with \( u \) then yields \( \langle u, v : \text{animal}(u), u \text{ is running}, \text{man}(v), v = u \rangle \).

We have now set up enough machinery in discourse representation theory to handle
simple cases of anaphoric connection. The theory correctly predicts the anaphoric properties of the sentences in Buridan’s very simple discourse. After processing the first sentence, we obtain the first DRS above; after processing the second, we obtain the second DRS, which embodies the information conveyed by both sentences.

This situation is quite general. The construction algorithm operates on sentence 1 in context 0, producing DRS 1. It then takes DRS 1 as context in processing sentence 2, yielding DRS 2, and so on. At each stage, the DRS produced embodies the information in all sentences processed up to that stage. Indeed, it must; otherwise it could not serve as context for the next sentence. The meaning of each sentence is not a truth condition that can be stated independently of previous discourse but a function taking discourse contexts into discourse contexts.

Discourse Representation Theory provides a semantics for DRSs and, thereby, for sentences and discourses. A discourse is true if and only if the DRS built from it by the construction algorithm is true. So, primarily, we will speak of DRSs as having truth values.

A DRS is a partial model. It is true in a model $\mathcal{M}$ if and only if it is a part of $\mathcal{M}$; there must be a way of embedding the DRS into $\mathcal{M}$. More formally, a DRS $m$ is true in a model $\mathcal{M}$ if and only if there is a homomorphic embedding of $m$ into $\mathcal{M}$. This means that there must be a function from the universe of $m$ into that of $\mathcal{M}$ preserving the properties and relations $m$ specifies. More formally still, there must be a function $f: U_m \rightarrow U_{\mathcal{M}}$ such that, if $\langle a_1, \ldots, a_n \rangle \in F_m(R)$, then $\langle f(a_1), \ldots, f(a_n) \rangle \in F_{\mathcal{M}}(R)$.

To see how the theory accounts for the quantificational force of indefinites, consider the DRS above, generated by the construction algorithm from the sentence ‘An animal is running.’ This DRS consists of a domain $\{u\}$ and a set of conditions $\{\text{animal}(u), u \text{ is running}\}$. The DRS is true in a model $\mathcal{M}$, according to our definition, if and only if it can be embedded into $\mathcal{M}$. This means that there must be a function $f$ from $\{u\}$ into $U_{\mathcal{M}}$ such that $f(u) \in F_{\mathcal{M}}(\text{animal})$ and $f(u) \in F_{\mathcal{M}}(\text{is running})$. Thus, the DRS, and the sentence that generated it, are true if and only if some animal is running, exactly as we would expect.

The DRS the construction algorithm generates from the entire discourse is true in a model $\mathcal{M}$ if and only if there is a function $f$ from $\{u, v\}$ into $U_{\mathcal{M}}$ such that $f(u) \in F_{\mathcal{M}}(\text{animal})$, $f(u) \in F_{\mathcal{M}}(\text{is running})$, $f(v) \in F_{\mathcal{M}}(\text{man})$, and $f(u) = f(v)$. The DRS is true, in other words, if there is an animal that is running and is also a man.

Note that the indefinite an animal has quantificational force here, in the sense that the truth conditions for the sentence might appropriately be represented by an existentially quantified formula, even though the indefinite did not introduce a quantifier into the DRS. It introduced nothing but a discourse referent and a condition. In Discourse Representation Theory, indefinite descriptions are referential terms, not existential quantifiers. There is nevertheless no simple answer to the question of what they denote. Their contribution to truth conditions depends on the role played by the clause containing the description, which depends, in turn, on the structure of the DRS. The same is true of any NP with a ↑mon↑ determiner. The theory accounts not only for the quantificational force of such NPs but also for their frequent success in establishing discourse referents, licensing further anaphoric connections.

Discourse Representation Theory thus explains the quantificational and anaphoric characteristics of the indefinite an animal at different levels of the theory. The indef-
inite introduces a discourse referent at the conceptual level of the DRS, which is then accessible to later pronouns. And the truth definition, specifying that the resulting DRS is true if and only if there is a way of embedding it in a model, determines that the discourse is true if and only if there is an animal, specifically a man, who is running.

Other kinds of determiners receive a different treatment. Let’s turn to Burley’s donkey sentence, ‘Every farmer who owns a donkey beats it.’ That is equivalent to ‘If a farmer owns a donkey, he beats it.’ We already know how to understand the antecedent; processing it yields the DRS \( \langle u, v : \text{farmer}(u), \text{donkey}(v), u \text{ owns } v \rangle \). A DRS for a conditional has the form \( m \Rightarrow m' \), where \( m \) and \( m' \) are DRSs for the antecedent and consequent. So, the DRS for this sentence becomes \( \langle u, v : \text{farmer}(u), \text{donkey}(v), u \text{ owns } v \rangle \Rightarrow \langle w, x : w \text{ beats } x, w = u, x = v \rangle \). This is the general strategy for ‘every,’ which introduces a conditional structure. Notice that the phrase ‘farmer who owns a donkey’ now corresponds to an identifiable part of the semantic representation; there is no reason to adhere to the Fregean principle that only in the context of a sentence does an expression have meaning, though it is of course true that only in the context of a discourse does a sentence have specific truth conditions.

The semantic condition for conditionals makes it clear why ‘every’ lives on its subject term. A conditional \( m \Rightarrow m' \) is true in a model \( \mathcal{M} \) if and only if every embedding of \( m \) into \( \mathcal{M} \) extends to an embedding of \( m' \) into \( \mathcal{M} \). This implies that the donkey sentence is true in \( \mathcal{M} \) if and only if every submodel of \( \mathcal{M} \) in which a farmer owns a donkey extends to one in which that farmer beats that donkey. But that is just to say that every farmer-owns-donkey pair in \( \mathcal{M} \) is also a farmer-beats-donkey pair, just as we would expect.

Discourse Representation Theory, by adopting a dynamic strategy in which the meaning of a sentence is a function from discourse contexts to discourse contexts—or, viewed differently, a function from representations to representations, or, from still another point of view, from partial models to partial models—respects surface structure, derives truth conditions compositionally in rule-governed ways, explains anaphoric connections within sentences and across sentence boundaries, and assigns appropriate truth conditions. It does so, however, by treating NPs with different determiners differently. Some, such as ‘a(n)’ and ‘some,’ introduce discourse referents; their quantificational force arises from the semantics, but receives no direct representation. Others, such as ‘every,’ introduce conditionals but also receive no direct representation. Yet others, such as ‘no,’ ‘never,’ and so on, introduce negations. And some, such as ‘many,’ ‘at least n,’ and ‘uncountably many,’ introduce plural discourse referents and conditions on them.