Quantifiers and discourse referents

Bart Geurts
University of Nijmegen

Quantified expressions like "most kangaroos" or "at most three wombats" interact with anaphoric pronouns in various ways, two of which are especially relevant to my present concerns. First, quantified NPs enable donkey anaphora, as in the emblematic:

(1) Every farmer who owns a donkey beats *it*.

According to the standard analysis, the indefinite "a donkey" introduces a discourse referent (DR), which the quantifier is instrumental in making accessible to the pronoun, and the main issue is to explain how precisely this goes. Secondly, quantifiers introduce DRs for pronouns to refer back to:

(2) Last week, most (of the) kangaroos were sick and some of *them* stayed indoors.

There are at least two ways of interpreting this sentence, depending on whether the plural pronoun refers back to the set of all kangaroos in the local zoo (say) or the subset of kangaroos that were sick. Either way, the pronoun picks up a DR associated with the quantifier, and the problem is to explain how quantifiers make availabe DRs for subsequent reference. This is the problem I will primarily be concerned with, but the solution I will outline has ramifications for donkey anaphora, which will be taken up at the end.

Let's assume that the truth-conditional meaning of quantified sentences is adequately captured by means of generalised quantifiers, i.e. relations between sets. Hence, the first conjunct in (2) says that the set of kangaroos that were sick was larger than the set of kangaroos that weren't. For ease of reference, here are some of the standard definitions (with due apologies for mishandling meta-variables):

(3) a. all(A,B) iff $A \subseteq B$
b. most(A,B) iff $|A \cap B| > |A \backslash B|$
c. some(A,B) iff $A \cap B \neq \varnothing$
d. no(A,B) iff $A \cap B = \varnothing$

Since the first conjunct of (2) is of the form "most(A,B)", the pronoun in the second conjunct either picks up a DR representing the A-set or one representing $A \cap B$. It seems likely, therefore, that there is some connection between the truth-conditional meaning of "most" and the DRs it makes available. The question I would like to consider is how, in general, the set of DRs associated with a quantifier Q relates to Q's truth-conditional meaning.

From a purely semantic point of view, it would be pleasing if the set of DRs associated with Q would be determined in a principled way by the sets Q takes as arguments. This would be the case if the family of sets repre-

Proceedings of the 6th International Conference of Cognitive Science, Seoul, July 27-29, 2008.

sented by DRs was, e.g., just $\{A,B\}$ or the closure of $\{A,B\}$ under union, intersection, and complementation, i.e. $\{\varnothing$, A, B, $A \cap B$, $A \cup B$, $A \backslash B$, $B \backslash A\}$ (call this set $\{A,B\}^*$). However, a moment's reflection suffices to see that neither $\{A,B\}$ nor $\{A,B\}^*$ will do. We already know that a quantifying sentence Q(A,B) may yield DRs representing A and $A \cap B$. As far as I can see, there is only one further set that is generally represented by DRs, viz. $A \backslash B$:

(4) Few professors showed up at the party. *They* were too busy writing research proposals.

Again, there are two ways of reading this, depending on whether the pronoun is taken to refer to the set of professors in the local university (say) or the set of professors who failed to show up at the party; the second reading requires a DR representing $A \backslash B$.

I submit that, between them, examples (2) and (4) exhaust the available options, and therefore propose the following generalisation:

(5) The DRs introduced by Q(A,B) represent all and only the non-empty members of $\{A, A \cap B, A \backslash B\}$.

To illustrate, for "most(A,B)" this yields DRs for A, $A \cap B$, and $A \backslash B$, while for "all(A,B)" (5) predicts that there will only be a DR (or DRs) representing A, as in this case $A \cap B = A$ and $A \backslash B = \varnothing$.

Prima facie, the generalisation in (5) may seem too coarse, because it doesn't distinguish between upward entailing and downward entailing quantifiers. We have seen that upward entailing "most" makes it possible for subsequent pronouns to refer to A and $A \cap B$, and that downward entailing "few" enables anaphora to A and $A \backslash B$. The following example shows that "few" may license anaphora to $A \cap B$, as well:

(6) Few professors showed up at the party, but *they* had a good time.

Here the pronoun may refer to all professors (A), the professors who failed to show up at the party ($A \backslash B$), or the ones who didn't ($A \cap B$). It's a bit harder to come up with good examples of upward entailing quantifiers that license anaphora to the complement set $A \backslash B$, but here's one. Every morning, eunuch E reports to his boss, caliph C, on the situation in C's harem:

(7) E: Your Highness, I'm extremely happy to report that almost all your wives are still in the palace.
C: What! Where are *they*?

True, there are differences between quantifiers which aren't reflected in the generalisation in (5). I believe, however, that such differences as can be observed are plausibly described in terms of salience and therefore don't require a more sophisticated version of the proposed generalisation.

Assuming, then, that (5) is correct as it stands, why does it hold? Usually, it is seen as a fact of life that quantifiers introduce the DRs they do. This is unsatisfactory. Granted, it may be an accident that something along the lines of (5) holds, but our working hypothesis should be that that it isn't. So let us consider how the question can be addressed. Note, to begin with, that a purely semantic answer doesn't seem to be feasible. Though is highly likely that the set of DRs associated with a quantifiying sentence is contingent on the sentence's truth-conditional content, pragmatic constraints appear to be involved, as well. Two such constraints readily come to mind. First, unless set theory is the topic of discourse, the empty set is not represented by a DR: even if it may be useful to establish that the set of honest lawyers is empty, there will be no need to discuss this *set* any further. Secondly, even if Q is a symmetric quantifier, a statement of the form Q(A,B) never is: it is about the A-set, and therefore any DRs associated with this form must represent A itself or a subset of A. Putting together the foregoing considerations, I propose the following:

(8) Q(A,B) only introduces DRs for those $X \in \{A,B\}^*$ such that $X \neq \varnothing$ and $X \subseteq A$.

It will not be hard to see that this accounts for the generalisation in (5).

Thus far, we have taken an extensional stance on the meanings of quantifier expressions, supposing as we did that such expressions are to be interpreted as relations between sets of individuals. This approach has its limits, which become relevant when we turn to examples like the following:

(9) [a] There are no kangaroos in Luxembourg. [b] ?*They* are sick.

Suppose (9a) is construed as no({x: kangaroo(x)}, {x: inLuxembourg(x)}). Then we predict that this sentence introduces a DR representing the set of kangaroos in the domain of discourse. But the ill-formedness of (9b) indicate that this would be wrong.

The problem, I would like to suggest, is that not every occurrence of an indefinite can be analysed as a generalised quantifier, and that in the context of an existential sentence like (9a), indefinites never admit of generalised-quantifier construals. Rather, (9a) means something along the lines of, "nowhere in Luxembourg is the kangaroo property instantiated", which doesn't imply the existence of kangaroos. On this analysis, (9a) merely requires the existence of the kangaroo property, thereby engendering a generic DR, which can be retrieved by means of a pronoun:

(10) There are no kangaroos in Luxembourg, but in Belgium *they* are quite common.

Note, however, that such generic take-ups are possible on general-quantifier construals, too.

Let us now revert to the extensional stance, and develop a bit further the principle proposed in (8). (11b) is a DRT-style representation of (11a) that implements this principle:

(11) a. Most (of the) kangaroos were sick.
b. [X, Y, Z: most(X,Y),
$X \subseteq \hat{x}[: K(x)]$, (*i*)
$Y = X \cap \hat{x}[: S(x)]$, (*ii*)
$Z = X \setminus Y$] (*iii*)

Here, lowercase and uppercase letters from the end of the alphabet act as individual and plural DRs, respectively, "most" is interpreted as in (3b), $\hat{x}\phi$ denotes the set of individuals matching description ϕ, and all set-theoretic notation has its standard meaning. ($11b_i$) introduces a DR that represents the domain of the quantifier "most", whose value is a subset of the set of kangaroos; I assume that this domain is restricted by the context. ($11b_{ii}$) and ($11b_{iii}$) introduce DRs representing, respectively, the set of kangaroos that were sick and the set of kangaroos that weren't sick.

Having shown how a particular way of representing quantified statements can be motivated by semantic and pragmatic considerations, it remains to be seen how the resulting analysis fares with donkey anaphora. Here is how (1) is represented in the format I have argued for (omitting the DR for A\B, which won't be needed):

(12) $[X, Y\!: \text{every}(X,Y),$
$\quad X \subseteq \hat{x}[y\!: F(x), D(y), O(x,y)],$
$\quad Y = X \cap \hat{x}[\underline{z}\!: B(x,z)]\ ]$

The problem with this is that the DR representing the donkey pronoun, marked here by underlining, doesn't have access to its intended referent, i.e. y. What to do? In DRT and other dynamic theories of interpretation, this problem is solved by defining the meaning of a quantifier in such a way that any DRs introduced in its restrictor are accessible from its (nuclear) scope. However, this means that donkey anaphora in quantified sentences is licensed by fiat (and the same goes for conditionals). Fortunately, there is a more principled way of solving the problem. To see how, note that in (12) $Y \subseteq X$ and $X \subseteq \hat{x}[y\!: F(x), D(y), O(x,y)]$. Therefore, it must be the case that $Y \subseteq \hat{x}[y\!: F(x), D(y), O(x,y)]$, too, and it logically follows from (12) that:

(13) $[X, Y\!: \text{every}(X,Y),$
$\quad X \subseteq \hat{x}[y\!: F(x), D(y), O(x,y)],$
$\quad Y = X \cap \hat{x}[y,\underline{z}\!: F(x), D(y), O(x,y), B(x,z)]\ ]$

Now, in this representation, y is accessible to z, so we can equate the two:

(14) $[X, Y\!: \text{every}(X,Y),$
$\quad X \subseteq \hat{x}[y\!: F(x), D(y), O(x,y)],$
$\quad Y = X \cap \hat{x}[y\!: F(x), D(y), O(x,y), B(x,y)]\ ]$

In prose: X is a (contextually restricted) set of farmers who own donkeys; Y is that subset of X all of whose members are farmers who own and beat a donkey; and X is a subset of Y. All of which is to say that every farmer who owns a donkey is a farmer who owns and beats a donkey.

What I'm suggesting is that donkey anaphora in quantified sentences is mediated by *bridging inferences*. If this is right, donkey pronouns are on a par with the italicised pronouns in the following examples:

(15) a. When the doorbell rang I thought *it* was Vernon.
b. What's that shadow creeping up the wall? Could *it* be a burglar?

To be sure, there is a difference, too: while the inferences needed to account for (15a,b) are, perhaps, defeasible (though their probability must be very close to 1), the inference I need for my analysis is logically valid. But clearly this difference doesn't jeopardise the proposed account.