ENTAILMENTS AND IMPLICATURES OF CUMULATIVE SENTENCES*

JUNRI SHIMADA
Keio University

Abstract

According to Krifka (1999), due to scalar implicature, sentences that express cumulative quantification are usually understood with a meaning stronger than their actual truth conditions, typically with an "exactly" interpretation. To evaluate this pragmatic account accurately, the present paper studies the entailment relation between such sentences in detail. It turns out that what Krifka claims to be empirical facts are either unpredicted or incorrect in many cases. Instead, a whole bunch of interesting but controversial scalar implicatures are predicted, calling for future investigation.

1 Introduction

Scha (1981) has observed that sentences containing two noun phrases can have an interesting reading which he calls *cumulative*. According to Scha, the cumulative reading of sentence (1), for instance, has the truth conditions in (2), where $|\ |$ denotes the cardinality of a set:

(1) Three boys kissed seven girls.

(2) $|\mathsf{boy} \cap \{x \mid \exists y[\mathsf{girl}(y) \wedge \mathsf{kiss}(x,y)]\}| = 3 \wedge |\mathsf{girl} \cap \{y \mid \exists x[\mathsf{boy}(x) \wedge \mathsf{kiss}(x,y)]\}| = 7$

As can be seen from (2), the two noun phrases participating in a cumulative reading play symmetrical semantic roles, and there is no sense in which one takes scope over the other. This is the defining characteristic of a cumulative reading. Since this paper focuses on cumulative

*I would like to thank Eric McCready, Makoto Kanazawa, Orin Percus, Uli Sauerland, Raj Singh and Chris Tancredi for sharing intuitions and comments, and the reviewer for carefully reading this paper and providing various helpful suggestions for improvement. This work has been presented at the Semantic Research Group meeting held at Keio University on February 24, 2014 and at the Workshop on Meaning, Quantification and Context held at Nihon University on July 28, 2014. As for the latter occasion, I would like to thank Takashi Iida for organizing it and having me in the program.

The Art and Craft of Semantics: A Festschrift for Irene Heim, vol. 2, MITWPL 71, pp. 123–153

readings, for the sake of simplicity, I will refer to sentences with cumulative readings as *cumulative sentences*.

More recently, Krifka (1999) has proposed that an analysis like (2) does not represent the real truth conditions of a cumulative sentence, but something that arises as a result of scalar implicature.[1] The present paper is inspired by this work and examines some of its claims.[2] Thus, this paper is not about compositional semantics, but a model-theoretic investigation into entailments and scalar implicatures of cumulative sentences. It will become clear that the standard Gricean account does not predict meanings like (2) as implicatures of cumulative sentences.

The organization of this paper is as follows. In Section 2, we review Krifka's (1999) accounts of how cumulative sentences are analyzed and how scalar implicatures affect their interpretations. From Section 3 through Section 5, we investigate in detail what entailment relations hold between cumulative sentences. Section 6 returns to Krifka's examples and Section 7 discusses what implicatures are predicted and actually arise for cumulative sentences. Section 8 concludes the paper.

2 Review of Krifka (1999)

Since Scha's (1981) work, cumulative readings have drawn much attention in the study of natural language semantics, partly because the very successful and influential theory of Generalized Quantifiers (Barwise and Cooper 1981) does not seem capable of dealing with them. Under GQ theory, determiners with a numeral greater than 1 are essentially distributive, so if a sentence has two NPs with such a determiner, one of them necessarily takes scope over the other. As a consequence, cumulative readings, where noun phrases are apparently scopeless, cannot be generated.

One popular solution to this puzzle has been advocated by Krifka (1989, 1992, 1999), which adopts Link's (1983) lattice-theoretic ontology. On Link's ontology, individuals form a lattice structure with respect to the "part-of" relation, which we will denote with $\sqsubseteq$. For any set X of individuals, there is a unique smallest individual that every member x of X is part of, and this is called the join of X, which we will represent with $\sqcup X$. For instance, the individual $\sqcup\{\mathsf{john}, \mathsf{bill}\}$, or its notational equivalent $\mathsf{john} \sqcup \mathsf{bill}$, denotes the join of the two individuals john and bill, and we have $\mathsf{john} \sqsubseteq \mathsf{john} \sqcup \mathsf{bill}$ and $\mathsf{bill} \sqsubseteq \mathsf{john} \sqcup \mathsf{bill}$. An individual that has no other (non-bottom-element) individual as a part is called an atom. While the denotation of a singular count noun is a set of atoms, the denotation of the corresponding plural form is its closure under join. For instance, the denotation of *boy* is the set of boy-atoms. If John, Bill and Chris are the only boys in the world, $[\![\text{boy}]\!] = \mathsf{boy} = \{\mathsf{john}, \mathsf{bill}, \mathsf{chris}\}$. Since $[\![\text{boys}]\!]$ is its closure under join, $[\![\text{boys}]\!] = \mathsf{boys} = \{\mathsf{john}, \mathsf{bill}, \mathsf{chris}, \mathsf{john} \sqcup \mathsf{bill}, \mathsf{bill} \sqcup \mathsf{chris}, \mathsf{chris} \sqcup \mathsf{john}, \mathsf{john} \sqcup \mathsf{bill} \sqcup \mathsf{chris}\}$.[3]

In order to understand Krifka's (1999) proposal, let us first review the classic account of the scalar implicature of the numeral interpretation and how the analysis is rendered on the lattice-theoretic ontology.

[1] Landman (2000) also makes essentially the same proposal, although Landman's theory has a different technical implementation from Krifka's and makes use of a maximalization operation.

[2] Krifka's (1999) main focus seems to be on how to treat determiners like *at least three*, *at most seven* etc., but this topic is not discussed in the present paper.

[3] With Link's $*$ operator, one may write $\mathsf{boys} = *\mathsf{boy}$.

When someone utters sentence (3) to report what happened at the party, the hearer will usually understand that John did not kiss more than seven girls at the party.

(3) John kissed seven girls.

However, even if John kissed eight or more girls, it will not contradict what is asserted by (3); if John kissed eight girls, then he certainly kissed seven girls. The hearer's supposition that John kissed only seven girls has often been taken to be a case of scalar implicature, which would arise due to Grice's (1975) Maxim of Quantity in the following way (cf. Horn 1972, Levinson 1983). Instead of (3), the speaker could have uttered another sentence of the form *John kissed n girls* where n represents a natural number other than 7. All sentences of this form where $n > 7$ entail (3). Therefore, if the speaker knew that the sentence *John kissed n girls* was true for some $n > 7$, then that sentence would have been more informative than (3), so uttering (3) should have violated Grice's Maxim of Quantity. Assuming that the speaker conformed to Grice's Maxims, the hearer can then reason that there is no $n > 7$ such that the speaker knew that the sentence *John kissed n girls* was true, and might even strengthen this inference and come to suppose that there is no $n > 7$ such that the sentence *John kissed n girls* is true.

The standard GQ analysis takes scalar implicature into account and gives the following truth conditions to sentence (3):

(4) $|\{x \mid \mathsf{girl}(x) \wedge \mathsf{kiss}(\mathsf{john}, x)\}| \geq 7$

Thus, the literal meaning of (3) is 'John kissed seven or more girls', but due to the scalar implicature, the hearer feels as if its truth conditions were those where the "≥" sign in (4) had been replaced with the "=" sign. Adopting Link's ontology, the same truth conditions as (4) are rendered as in (5), where # is the function that gives the number of atoms that are part of a given (possibly plural) individual.

(5) $\exists X[\#(X) = 7 \wedge \mathsf{girls}(X) \wedge \mathsf{kiss}(\mathsf{john}, X)]$

Observe that in (5), unlike in (4), we have the "=" sign rather than the "≥" sign. This apparent symbolic difference notwithstanding, (5) is equivalent to (4) in substance. (5) only asserts the existence of a collection of exactly seven girls who were kissed by John, and this is consistent with there being eight or more girls who were kissed by John. On the other hand, (4) uses the "≥" sign because it talks about the collection of all the girls who were kissed by John. Because of the "=" sign in (5), one might perhaps regard (5) as expressing the meaning of (3) in a more intuitive and straightforward way.

Let us now turn to cumulative readings. The key notion that Krifka (1999), building on earlier work (Krifka 1989, 1992), entertains is *cumulativity* of predicates, which can be defined as follows:

(6) a. A one-place predicate P is cumulative iff $\forall x_1 \forall x_2[(P(x_1) \wedge P(x_2)) \rightarrow P(x_1 \sqcup x_2)]$.
b. A two-place predicate P is cumulative iff $\forall x_1 \forall x_2 \forall y_1 \forall y_2[(P(x_1, y_1) \wedge P(x_2, y_2)) \rightarrow P(x_1 \sqcup x_2, y_1 \sqcup y_2)]$.

We have already seen that plural nouns denote (one-place) cumulative predicates on Link's ontology. Krifka's proposal is that verbal predicates, such as the two-place predicate kiss, are also generally cumulative. Then, given the analysis in (5), it seems natural to posit the following truth conditions for the cumulative reading of (1):

(7) $\exists X \exists Y[\#(X)=3 \wedge \mathsf{boys}(X) \wedge \#(Y)=7 \wedge \mathsf{girls}(Y) \wedge \mathsf{kiss}(X,Y)]$

This is Krifka's solution to the conundrum about cumulative readings that GQ theory was confronted with. Evidently, it succeeds in giving a very intuitive account of cumulative sentences, and providing a compositional semantics for them now seems to be a piece of cake.

This solution, however, has led to a new puzzle. It turns out that the truth conditions in (7) are not equivalent to Scha's truth conditions in (2). Just as (5) is consistent with there being eight or more girls who were kissed by John, (7) is consistent with there being four or more boys who kissed some girl or another and there being eight or more girls who were kissed by some boy or another.[4] On the other hand, (2) asserts that there are only three boys who kissed girls and that there are only seven girls who were kissed by boys. Krifka (1999) proposes that scalar implicature makes up for this discrepancy, just as in the classic case of (3). According to Krifka, when the sentence *n boys kissed m girls* is uttered, the Maxim of Quantity has the effect that the hearer assumes that n and m are the greatest numbers that make a true sentence.[5] Thus, the hearer would feel as if (1) meant (2), even though the genuine truth conditions are (7). In support of this view, Krifka provides the following examples:

(8) a. Three boys ate seven apples, perhaps even eight apples.
    b. ?Three boys, perhaps even four boys, ate seven apples.

According to Krifka, *three boys ate seven apples* should implicate that it is not the case that three boys ate eight apples. Therefore, *perhaps even eight apples* in (8a) can be understood as canceling this induced scalar implicature. The fact that (8a) sounds natural shows that the "exactly" interpretation is not the genuine truth conditions of *three boys ate seven apples*. Similarly, *three boys ate seven apples* should also implicate that it is not the case that four boys ate seven apples. So (8b) should be as good as (8a) for a parallel reason. Krifka finds (8b) to be degraded, however, and suggests that although the effect of a scalar implicature is clear with the object NP, it is less so with the subject NP.[6] In Section 6.1, we will return to (8) and account for the contrast, but not in terms of some subject-object asymmetry as Krifka suggests.

Let us now look at (9), another very interesting example from Krifka (1999).

(9) In Guatemala, 3% of the population owns 70% of the land.

Krifka says that (9) is interpreted like the following:

---

[4]Incidentally, the following truth conditions, which look analogous to (4), do not express the same thing as (7).

(i) $|\mathsf{boy} \cap \{x \mid \exists y[\mathsf{girl}(y) \wedge \mathsf{kiss}(x,y)]\}| \geq 3 \wedge |\mathsf{girl} \cap \{y \mid \exists x[\mathsf{boy}(x) \wedge \mathsf{kiss}(x,y)]\}| \geq 7$

If there are seven boys and seven girls, and each boy kissed only one girl and each girl was kissed by only one boy, then (7) is false because no three boys kissed seven girls, but (i) comes out true. Krifka (1999) gives something like (ii) (≈ Krfika's (11) on p. 262) as the truth conditions for the cumulative reading of *three apples ate seven apples*, but this must simply be a mistake, and he should have meant something like (iii) (≈ Krifka's (31a) on p. 270).

(ii) $|\mathsf{boy} \cap \{x \mid \exists y[\mathsf{apple}(y) \wedge \mathsf{eat}(x,y)]\}| \geq 3 \wedge |\mathsf{apple} \cap \{y \mid \exists x[\mathsf{boy}(x) \wedge \mathsf{eat}(x,y)]\}| \geq 7$

(iii) $\exists X \exists Y[\#(X)=3 \wedge \mathsf{boys}(X) \wedge \#(Y)=7 \wedge \mathsf{apples}(Y) \wedge \mathsf{eat}(X,Y)]$

[5]However, Krifka notes that *n boys ate m apples* does not always entail *n′ boys ate m′ apples* when $n' \leq n$ and $m' \leq m$. This is indeed the important point that I will elaborate on in the following section.

[6]As to the difference in acceptability between (8a) and (8b), Krifka (1999) suggests that it is "presumably because the subject NP is most likely interpreted as the topic" (p. 4) in the version found on the author's website, but this part has apparently been deleted in the published version in Turner (1999).

(10) In Guatemala, at most 3% of the population owns at least 70% of the land.

To see this point, first consider the sentences in (11) and (12).

(11) a. In Guatemala, 2% of the population owns 70% of the land.
b. In Guatemala, 3% of the population owns 70% of the land.
c. In Guatemala, 4% of the population owns 70% of the land.

(12) a. In Guatemala, 3% of the population owns 60% of the land.
b. In Guatemala, 3% of the population owns 70% of the land.
c. In Guatemala, 3% of the population owns 80% of the land.

In each set, the numeral in one of the two NPs increases from the (a) sentence to the (c) sentence, while the numeral in the other NP is kept constant. (11b) and (12b) are identical to (9). Assuming that 4 or more % of the Guatemalan population owns some land of Guatemala, we see that (11a) entails (11b), which in turn entails (11c), but (11b) does not entail (11a), and nor does (11c) entail (11b). For (12), the entailment pattern is the other way around. That is, (12c) entails (12b), which in turn entails (12a), but entailments in the reverse direction do not go through (even if we assume that 80 or more % of the land is owned by Guatemalan people). Next, consider the following sets of data, where the subject numeral is headed by *at most* and the object numeral by *at least*:

(13) a. In Guatemala, at most 2% of the population owns at least 70% of the land.
b. In Guatemala, at most 3% of the population owns at least 70% of the land.
c. In Guatemala, at most 4% of the population owns at least 70% of the land.

(14) a. In Guatemala, at most 3% of the population owns at least 60% of the land.
b. In Guatemala, at most 3% of the population owns at least 70% of the land.
c. In Guatemala, at most 3% of the population owns at least 80% of the land.

Here, we see that exactly the same entailment patterns are found as in (11)–(12). That is, we have (13a) $\models$ (13b) $\models$ (13c) and (14c) $\models$ (14b) $\models$ (14a), and no entailments in the reverse directions hold. It is then understandable that (9) is interpreted like (10).

If we followed Krifka's account for cases like (1), we would predict that the truth conditions of (9) are something like (15) and that the sentence induces the implicature that only 3% of the population owns land, and that only 70% of the land is owned by people.

(15) $\exists X \exists Y [\textsf{\%-of-the-population}(X) = 3 \wedge \textsf{population}(X) \wedge \textsf{\%-of-the-land}(Y) = 70 \wedge \textsf{land}(Y) \wedge \textsf{own}(X,Y)]$

However, this prediction is clearly inconsistent with the intuition that (9) is interpreted like (10). As Krifka notes, if it is the case that all Guatemalan people have some land of Guatemala and every part of Guatemala's land is possessed by some Guatemalan citizen, the theory predicts that the sentence *in Guatemala, 100% of the population owns 100% of the land* should be used. Paradoxically, such a sentence seems pretty uninformative, despite the fact that Krifka's account is based on Grice's Maxim of Quantity.

Krifka's explanation for (9) is as follows. Sentences like (9) describe a bias in a statistical distribution by selecting a small set among one dimension that is related to a large set of the other dimension, and in so doing, one should try to decrease the first set and increase the second. In the case of (9), 3% of the population has been chosen as the small set and 70% of the land as the contrasting large set. In terms of informativity, (9) entails, but is not entailed by, *in Guatemala, n% of the population owns m% of the land* where $n \geq 3$, $m \leq 70$ and either $n \neq 3$ or $m \neq 70$. So (9) is in fact informative. Essentially, Krikfa attributes the difference in implicature between (1) and (9) to the difference in utterance context. That is, in an utterance context where a biased distribution is discussed, the entailment relation for cumulative sentences can be quite different than in an utterance context where the existence of certain groups of individuals is simply asserted. We will return to these Guatemala sentences in Section 6.2.

I take it that Krifka's semantic analysis of cumulative sentences is essentially correct, in that the truth conditions of (1) and (9) should look like (7) and (15) respectively.[7] I also agree that with scalar implicature, we may interpret these sentences with somewhat stronger truth conditions. However, in order to evaluate Krifka's pragmatic accounts accurately, we need to know precisely how scalar implicature may strengthen cumulative sentences, and for that, it is imperative to elucidate which cumulative sentence entails which other cumulative sentence, and how the utterance context affects the entailment relation. We will therefore investigate these issues from Section 3 through Section 5. In doing so, we make two kinds of distinction. The first concerns whether the nouns involved in the cumulative sentence have a count denotation (e.g. *boys*) or a mass denotation (e.g. *land*). For a cumulative sentence with two NPs, three combinatorial possibilities exist: two count nouns, one count noun and one mass noun, and two mass nouns. As it turns out, these different combinations lead to different entailment relations. The second kind of distinction is about whether the utterance context is one where the existence of certain groups of individuals is simply asserted, which we will call an *existential context*, or one where a (possibly biased) distribution is discussed, which we will call a *distribution-describing context*. The following table points which section investigates which case.

	existential context	distribution-describing context
two count nouns	Section 3.1	Section 3.2
one count noun + one mass noun	Section 4.2	Section 4.3
two mass nouns	Section 5	

3 Entailments of cumulative sentences with two count nouns

This section investigates entailments of cumulative sentences with two count nouns. Let N_1 and N_2 be the denotations of two count nouns and let V be the denotation of a two-place verb. For all $n, m \in \mathbb{N}$, we set

$$nR^{\text{cc}}m \equiv \exists X \exists Y[\#(X) = n \wedge N_1(X) \wedge \#(Y) = m \wedge N_2(Y) \wedge V(X,Y)].$$

That is, $nR^{\text{cc}}m$ expresses the truth conditions of a relevant cumulative sentence. We assume that N_1, N_2 and V are cumulative (i.e., closed under join, see (6)). We will also assume that N_1, N_2 and V are *divisive* in the following sense:

[7] Robaldo (2011) proposes different truth conditions for cumulative sentences.

(16) a. A one-place predicate P is divisive
iff for all X such that $P(X)$ holds, $P(x)$ holds for every (non-bottom-element individual[8]) $x \sqsubseteq X$.

b. A two-place predicate P is divisive
iff for all X,Y such that $P(X,Y)$ holds, for every (non-bottom-element individual) $x \sqsubseteq X$ there is some (non-bottom-element individual) $y \sqsubseteq Y$ such that $P(x,y)$ holds, and for every (non-bottom-element individual) $y \sqsubseteq Y$ there is some (non-bottom-element individual) $x \sqsubseteq X$ such that $P(x,y)$ holds.

V is not divisive when it allows collective interpretation. For example, suppose the following sentence is true:

(17) John, Bill, Mary and Ann lifted the piano.

This will be translated as $\mathsf{lift}(\mathsf{john} \sqcup \mathsf{bill} \sqcup \mathsf{mary} \sqcup \mathsf{ann}, \mathsf{the\text{-}piano})$. If lift were divisive, we would deduce from this that $\mathsf{lift}(\mathsf{john}, \mathsf{the\text{-}piano})$ held, since $\mathsf{john} \sqsubseteq \mathsf{john} \sqcup \mathsf{bill} \sqcup \mathsf{mary} \sqcup \mathsf{ann}$, but this is apparently wrong. (17) can be true without John having lifted the piano alone, if John and Bill lifted the piano in joint effort (i.e., $\mathsf{lift}(\mathsf{john} \sqcup \mathsf{bill}, \mathsf{the\text{-}piano})$) and Mary and Ann did the same (i.e., $\mathsf{lift}(\mathsf{mary} \sqcup \mathsf{ann}, \mathsf{the\text{-}piano})$). This shows that lift is not divisive. What is discussed below concerns verbs that do not allow collective interpretation and are hence divisive, such as *kiss*.

3.1 Existential context

We first examine existential contexts, where a cumulative sentence is used to simply assert the existence of certain groups of individuals. This amounts to inquiring for which combinations of n, m, n' and m', $nR^{\text{cc}}m$ entails $n'R^{\text{cc}}m'$. The two propositions below turn out to be fundamental. In what follows, for any $X \in D_e$, we write $\text{AT}(X)$ to denote $\{x \mid x \text{ is an atom} \wedge x \sqsubseteq X\}$. We have $\#(X) = |\text{AT}(X)|$.

Proposition 1. *For all $n, m, l \in \mathbb{N}$ such that $n \leq l < m$,*

$$nR^{\text{cc}}m \models nR^{\text{cc}}l,$$
$$mR^{\text{cc}}n \models lR^{\text{cc}}n.$$

Proof. Let $n \leq l < m$. Suppose $nR^{\text{cc}}m$ holds, i.e., there are individuals A and B such that

$$\#(A) = n \wedge N_1(A) \wedge \#(B) = m \wedge N_2(B) \wedge V(A,B).$$

Since V is divisive, for each $a \in \text{AT}(A)$, we can pick a $b \in \text{AT}(B)$ such that $V(a,b)$. Let us write this b by $f(a)$. That is, f is a function from $\text{AT}(A)$ into $\text{AT}(B)$ such that $V(a, f(a))$. We write $f(\text{AT}(A))$ to denote the image of $\text{AT}(A)$ by f. Then, $f(\text{AT}(A)) \subseteq \text{AT}(B)$ and $|f(\text{AT}(A))| \leq |\text{AT}(A)| = \#(A) = n$. Since $|\text{AT}(B)| = \#(B) = m$, $\text{AT}(B) \backslash f(\text{AT}(A))$ has at least $m - n$ elements. Since $m - l \leq m - n$, we can pick $m - l$ distinct elements $b_1, \ldots, b_{m-l}$ out of $\text{AT}(B) \backslash f(\text{AT}(A))$. Put $B' = \bigsqcup(\text{AT}(B) \backslash \{b_1, \ldots, b_{m-l}\})$. Then, $\#(B') = |\text{AT}(B) \backslash \{b_1, \ldots, b_{m-l}\}| = m - (m - l) = l$. Also, $N_2(B')$ holds by the divisivity of N_2, since $N_2(B)$ holds and $B' \sqsubseteq B$. Lastly, we can verify that

[8] Krifka's ontology excludes the existence of a bottom element. In the present paper, however, we will be assuming a bottom-element individual ($\perp$). See Section 4.

$V(A,B')$ holds as follows. First, for every $a \in \text{AT}(A)$, there is a $b \in \text{AT}(B')$ such that $V(a,b)$, because $V(a,f(a))$ holds and $f(a) \sqsubseteq \bigsqcup f(\text{AT}(A)) \sqsubseteq B'$ (since $f(\text{AT}(A)) \subseteq \text{AT}(B) \backslash \{b_1,\ldots,b_{m-l}\}$). Second, for every $b \in \text{AT}(B')$, there is an $a \in \text{AT}(A)$ such that $V(a,b)$, because $V(A,B)$ holds and $\text{AT}(B') \subseteq \text{AT}(B)$. We have thus established

$$\#(A) = n \wedge N_1(A) \wedge \#(B') = l \wedge N_2(B') \wedge V(A,B'),$$

which implies $nR^{\text{cc}}l$. This completes the proof of $nR^{\text{cc}}m \models nR^{\text{cc}}l$. One can prove $mR^{\text{cc}}n \models lR^{\text{cc}}n$ similarly. □

Proposition 2. *For all $n,m,l \in \mathbb{N}$ such that $1 \leq n-1 \leq l$ and $(m-l+1)n > m$,*

$$nR^{\text{cc}}m \models (n-1)R^{\text{cc}}l,$$
$$mR^{\text{cc}}n \models lR^{\text{cc}}(n-1).$$

Proof. By reductio ad absurdum. Let $1 \leq n-1 \leq l$ and $(m-l+1)n > m$. This actually means $m \geq l$, because otherwise we would have $(m-l+1)n \leq 0$, contrary to the assumption that $(m-l+1)n > m > 0$. Now, assume that $nR^{\text{cc}}m$ holds but $(n-1)R^{\text{cc}}l$ does not. Since $nR^{\text{cc}}m$ holds, there are individuals A and B such that

$$\#(A) = n \wedge N_1(A) \wedge \#(B) = m \wedge N_2(B) \wedge V(A,B).$$

For every $a \in \text{AT}(A)$, put

$$f(a) = \bigsqcup(\text{AT}(A) \backslash \{a\}), \quad g(a) = \bigsqcup\{y \sqsubseteq B \mid \exists x \sqsubseteq f(a)[V(x,y)]\}.$$

It is easy to verify that $f(a) \sqsubseteq A$, $g(a) \sqsubseteq B$, $\#(f(a)) = n-1$, and $N_1(f(a))$, $N_2(g(a))$ and $V(f(a),g(a))$ hold, and hence $(n-1)R^{\text{cc}}\#(g(a))$. If it were the case that $\#(g(a)) \geq l$, by Proposition 1, $(n-1)R^{\text{cc}}l$ should hold, contrary to the assumption. So $\#(g(a)) < l$, and hence $m - \#(g(a)) \geq m-l+1$. This means that there are at least $m-l+1$ atom-individuals $y \in \text{AT}(B)$ such that $y \not\sqsubseteq g(a)$ and $V(a,y)$. Because this is the case for each of the n atom-individuals $a \in \text{AT}(A)$ and because $(m-l+1)n > m = \#(B)$ by assumption, the pigeonhole principle tells us that there are some $a_1,a_2 \in \text{AT}(A)$ and $b \in \text{AT}(B)$ such that $a_1 \neq a_2$, $b \not\sqsubseteq g(a_1)$, $V(a_1,b)$, $b \not\sqsubseteq g(a_2)$ and $V(a_2,b)$. Then,

$$\begin{aligned}
g(a_1) \sqcup g(a_2) &= \Big(\bigsqcup\{y \sqsubseteq B \mid \exists x \sqsubseteq f(a_1)[V(x,y)]\}\Big) \sqcup \Big(\bigsqcup\{y \sqsubseteq B \mid \exists x \sqsubseteq f(a_2)[V(x,y)]\}\Big) \\
&= \bigsqcup(\{y \sqsubseteq B \mid \exists x \sqsubseteq f(a_1)[V(x,y)]\} \cup \{y \sqsubseteq B \mid \exists x \sqsubseteq f(a_2)[V(x,y)]\}) \\
&\sqsupseteq \bigsqcup(\{y \sqsubseteq B \mid \exists x \in \text{AT}(f(a_1))[V(x,y)]\} \cup \{y \sqsubseteq B \mid \exists x \in \text{AT}(f(a_2))[V(x,y)]\}) \\
&= \bigsqcup\{y \sqsubseteq B \mid \exists x \in \text{AT}(f(a_1)) \cup \text{AT}(f(a_2))[V(x,y)]\} \\
&= \bigsqcup\{y \sqsubseteq B \mid \exists x \in (\text{AT}(A)\backslash\{a_1\}) \cup (\text{AT}(A)\backslash\{a_2\})[V(x,y)]\} \\
&= \bigsqcup\{y \sqsubseteq B \mid \exists x \in \text{AT}(A)[V(x,y)]\} \\
&= B.
\end{aligned}$$

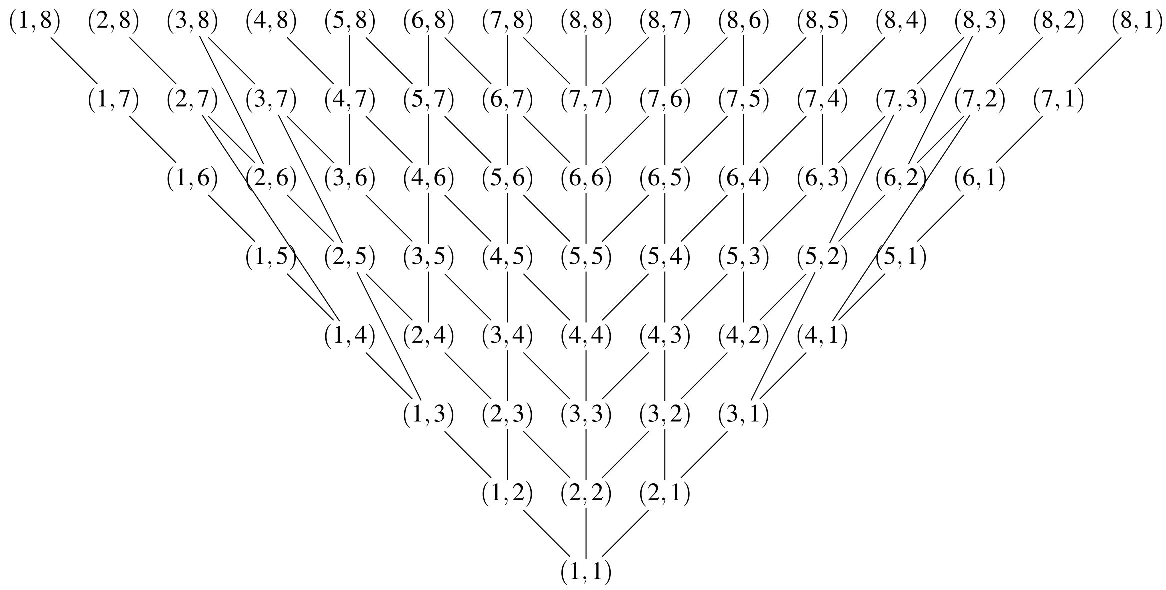

Diagram 1

Since $b \sqsubseteq B$, we obtain $b \sqsubseteq g(a_1) \sqcup g(a_2)$. Because b is an atom, this implies that we have either $b \sqsubseteq g(a_1)$ or $b \sqsubseteq g(a_2)$,[9] but this contradicts $b \not\sqsubseteq g(a_1)$ and $b \not\sqsubseteq g(a_2)$. This completes the proof of $nR^{\mathrm{cc}}m \models (n-1)R^{\mathrm{cc}}l$. One can prove $mR^{\mathrm{cc}}n \models lR^{\mathrm{cc}}(n-1)$ similarly. □

Corollary 3. *For all $n \in \mathbb{N}$ such that $n \geq 2$,*

$$nR^{\mathrm{cc}}n \models (n-1)R^{\mathrm{cc}}(n-1).$$

Proof. Let $n \geq 2$. Since $(n-(n-1)+1)n = 2n > n$, the claim follows from Proposition 2. □

Makoto Kanazawa proved that Propositions 1 and 2 exhaust all valid entailments between cumulative sentences whose truth conditions are of the form $nR^{\mathrm{cc}}m$, and we report this in Kanazawa and Shimada (2014). Diagram 1 visualizes the partial order induced by this entailment relation for $n, m \leq 8$. Here, node (n,m) represents the cumulative sentence with the truth conditions $nR^{\mathrm{cc}}m$. When two nodes are connected by a line, the one above entails the one below. Thus, for instance, we can read from the diagram that *three boys kissed five girls* entails *three boys kissed four girls*, but does not entail *two boys kissed five girls*.

3.2 Distribution-describing context

We now examine entailments in distribution-describing contexts. In an existential context, whenever *n boys kissed m girls* entails *n′ boys kissed m′ girls*, we have $n' \leq n$ and $m' \leq m$, because we can never infer from *n boys kissed m girls* that more than *n* boys kissed girls or more than *m* girls

[9] This follows from the tacit assumption that D_e is a distributive lattice. Since $b \sqsubseteq g(a_1) \sqcup g(a_2)$, we have $b = b \sqcap (g(a_1) \sqcup g(a_2)) = (b \sqcap g(a_1)) \sqcup (b \sqcap g(a_2))$ (see (19) and (20d) below). Because b is an atom, this shows that each of $b \sqcap g(a_1)$ and $b \sqcap g(a_2)$ equals either $\bot$ (bottom element) or b, but they cannot both be $\bot$ at the same time. Hence, either $b \sqcap g(a_1) = b$ or $b \sqcap g(a_2) = b$, that is, either $b \sqsubseteq g(a_1)$ or $b \sqsubseteq g(a_2)$.

were kissed by boys. By contrast, when the same sentence is uttered in a distribution-describing context, the background information ensures that more then n boys kissed girls or more than m girls were kissed by boys, and the sentence is simply meant to give some information on the distribution of those individuals. Therefore, the entailment may go through even if $n' > n$ or $m' > m$.

Proposition 4. *For all $n,m,n',m' \in \mathbb{N}$ such that $n < n'$ and $m < m'$,*

$$nR^{\text{cc}}m \wedge \exists X \exists Y[\#(X) = n' \wedge N_1(X) \wedge N_2(Y) \wedge V(X,Y)] \\ \wedge \exists X \exists Y[N_1(X) \wedge \#(Y) = m' \wedge N_2(Y) \wedge V(X,Y)] \models (n+1)R^{\text{cc}}(m+1).$$

Proof. Let $n < n'$ and $m < m'$. Suppose that $nR^{\text{cc}}m$, $\exists X \exists Y[\#(X) = n' \wedge N_1(X) \wedge N_2(Y) \wedge V(X,Y)]$ and $\exists X \exists Y[N_1(X) \wedge \#(Y) = m' \wedge N_2(Y) \wedge V(X,Y)]$ all hold. That is, there are individuals A and B such that

$$\#(A) = n \wedge N_1(A) \wedge \#(B) = m \wedge N_2(B) \wedge V(A,B),$$

and there are individuals C and D such that

$$\#(C) = n' \wedge N_1(C) \wedge N_2(D) \wedge V(C,D), \qquad (\bigstar)$$

and there are individuals E and F such that

$$N_1(E) \wedge \#(F) = m' \wedge N_2(F) \wedge V(E,F). \qquad (\bigstar\bigstar)$$

Since $\#(A) = n < n' = \#(C)$, there is at least one individual c such that $c \in \text{AT}(C)$ but $c \notin \text{AT}(A)$. Since $N_1(C)$ holds and $c \sqsubseteq C$, $N_1(c)$ holds by the divisivity of N_1. Likewise, since $\#(B) = m < m' = \#(F)$, there is at least one individual f such that $f \in \text{AT}(F)$ but $f \notin \text{AT}(B)$. $N_2(f)$ holds by the divisivity of N_2. Since $N_1(A)$ and $N_1(c)$ both hold, $N_1(A \sqcup c)$ holds by the cumulativity of N_1. Likewise, $N_2(B \sqcup f)$ holds. Since $c \notin \text{AT}(A)$, $\#(A \sqcup c) = \#(A) + 1 = n + 1$. Likewise, $\#(B \sqcup f) = \#(B) + 1 = m + 1$. We show that $(n+1)R^{\text{cc}}(m+1)$ holds by examining the following two cases.

Case 1: there is some $b \in \text{AT}(B)$ such that $V(c,b)$ and there is some $a \in \text{AT}(A)$ such that $V(a,f)$. In this case, by the cumulativity of V, we have $V(A \sqcup c \sqcup a, B \sqcup b \sqcup f)$, but since $a \sqsubseteq A$ and $b \sqsubseteq B$, this means $V(A \sqcup c, B \sqcup f)$. Thus,

$$\#(A \sqcup c) = n+1 \wedge N_1(A \sqcup c) \wedge \#(B \sqcup f) = m+1 \wedge N_2(B \sqcup f) \wedge V(A \sqcup c, B \sqcup f),$$

implying $(n+1)R^{\text{cc}}(m+1)$.

Case 2: either (i) there is no $y \in \text{AT}(B)$ such that $V(c,y)$ or (ii) there is no $x \in \text{AT}(A)$ such that $V(x,f)$. In the case of (i), by ($\bigstar$) and by the divisivity of N_2 and V, there is some individual $d \in \text{AT}(D)$ such that $N_2(d)$ and $V(c,d)$. By assumption, $d \notin \text{AT}(B)$, so $\#(B \sqcup d) = m + 1$. Also, by the cumulativity of N_2 and V, we have $N_2(B \sqcup d)$ and $V(A \sqcup c, B \sqcup d)$. Hence

$$\#(A \sqcup c) = n+1 \wedge N_1(A \sqcup c) \wedge \#(B \sqcup d) = m+1 \wedge N_2(B \sqcup d) \wedge V(A \sqcup c, B \sqcup d),$$

implying $(n+1)R^{\text{cc}}(m+1)$. In case (ii), by ($\bigstar\bigstar$) and by the divisivity of N_1 and V, there is some individual $e \in \text{AT}(E)$ such that $N_1(e)$ and $V(e,f)$, and $e \notin \text{AT}(A)$. We can similarly verify that

$$\#(A \sqcup e) = n+1 \wedge N_1(A \sqcup e) \wedge \#(B \sqcup f) = m+1 \wedge N_2(B \sqcup f) \wedge V(A \sqcup e, B \sqcup f),$$

which implies $(n+1)R^{\text{cc}}(m+1)$. □

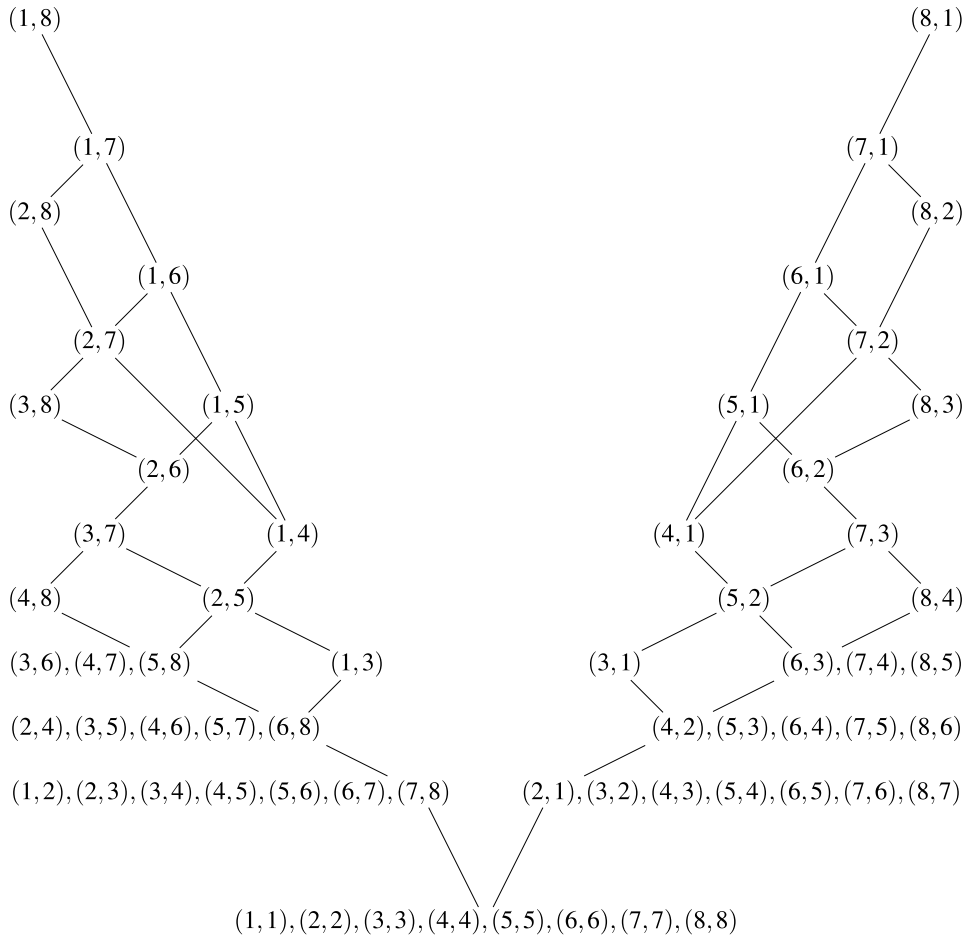

Diagram 2

Suppose that someone utters the sentence *six boys kissed nine girls* in a context where the people engaged in the conversation already know that ninety boys kissed girls and sixty girls were kissed by boys. By Proposition 4, we then know that *seven boys kissed ten girls* is true, and in turn, so is *eight boys kissed eleven girls*, and so on.

Corollary 5. *For all $n,m,n',m' \in \mathbb{N}$ such that $n < n'$ and $m < m'$,*

$$nR^{cc}m \wedge n'R^{cc}m' \models (n+1)R(m+1).$$

Proof. Since $n'R^{cc}m' \models \exists X \exists Y[\#(X) = n' \wedge N_1(X) \wedge N_2(Y) \wedge V(X,Y)]$ and $n'R^{cc}m' \models \exists X \exists Y[N_1(X) \wedge \#(Y) = m' \wedge N_2(Y) \wedge V(X,Y)]$, the claim follows from Proposition 4. □

We can now draw entailment diagrams for distribution-describing contexts by combining the result in the previous and the present subsections. Diagram 2 visualizes the partial order induced by the entailment relation for cumulative sentences with the truth conditions $nR^{cc}m$ for $n,m \leq 8$, when $8R^{cc}8$ is given as background information.[10] Thus, one may regard node (n,m)

[10] As Makoto Kanazawa (p.c.) points out, the term "distribution-describing" may be somewhat of a misnomer for referring to a context where the speaker utters *n boys kissed m girls* with $n \leq n'$ and $m \leq m'$ when *n' boys kissed m'*

as representing $nR^{\mathrm{cc}}m \wedge 8R^{\mathrm{cc}}8$. In many cases, multiple sentences which are represented by separate nodes in Diagram 1 have collapsed into a single node. For example, for all n,m such that $n \leq m \leq 8$, Corollary 3 gives $mR^{\mathrm{cc}}m \models nR^{\mathrm{cc}}n$, but now Corollary 5 gives $nRn \wedge 8R^{\mathrm{cc}}8 \models mR^{\mathrm{cc}}m$, and consequently, $(1,1)$, $(2,2)$, ..., $(8,8)$ are represented by a single node.

4 Entailments of cumulative sentences with one count noun and one mass noun

In this section, we investigate entailments of cumulative sentences with one count noun and one mass noun such as (18).

(18) Three boys drank 7 liters of beer.

As can be seen in this example, we use a measure phrase such as *7 liters* to talk about the size of a mass individual. A natural idea that suggests itself is to assume that units of measurement such as *liter* express a *measure* for individuals, which is a function from individuals to nonnegative real numbers. Now, let N_1 be the denotation of a count noun, N_2 the denotation of a mass noun, μ a measure, and V the denotation of a two-place verb, and for all $n \in \mathbb{N}$ and $s \in \mathbb{R}$, we set

$$nR^{\mathrm{cm}}s \equiv \exists X \exists Y[\#(X) = n \wedge N_1(X) \wedge \mu(Y) = s \wedge N_2(Y) \wedge V(X,Y)].$$

Then, the truth conditions of a cumulative sentence with one count noun and one mass noun are of the form $nR^{\mathrm{cm}}s$ with some $n \in \mathbb{N}$ and $s \in \mathbb{R}$. Again, we will assume that N_1, N_2 and V are all cumulative and divisive. Our task is to examine for which $n,n' \in \mathbb{N}$ and $s,s' \in \mathbb{R}$, $nR^{\mathrm{cm}}s$ entails $n'R^{\mathrm{cm}}s'$. To that end, the following subsection makes ontological assumptions for our semantic theory and clarifies what we mean by measures for individuals and what basic properties they have. The reader who is not interested in mathematical details may skip the proofs.

4.1 Preliminaries

Measures for individuals have been discussed in the literature, including Krifka (1990). On Krifka's ontology, the domain D_e of individuals constitutes a complete join-semilattice with no bottom element. However, as we will see shortly, it is convenient to have a bottom element for a simplified treatment of measurement of individuals. Let us therefore assume that D_e has a bottom

girls is given as background information. If *n boys kissed m girls* is really meant to describe the distribution of the relation between the n' boys and the m' girls, then those n boys must be a subset of those n' boys and those m girls a subset of those m' girls. However, the sentence obviously does not guarantee such an interpretation; for that purpose, one would have to use the sentence *n of the n' boys kissed m of the m' girls*. Note, however, that from $nR^{\mathrm{cc}}m$ and $n'R^{\mathrm{cc}}m'$, it follows that there are individuals A and B such that $\#(A) = n \wedge N_1(A) \wedge \#(B) = m \wedge N_2(B) \wedge V(A,B)$ and that there are individuals A' and B' such that $\#(A') = n' \wedge N_1(A') \wedge \#(B') = m' \wedge N_2(B) \wedge V(A',B')$, so by cumulativity, we obtain $N_1(A \sqcup A')$, $N_2(B \sqcup B')$ and $V(A \sqcup A', B \sqcup B')$, and we have $\#(A \sqcup A') \leq n + n'$ and $\#(B \sqcup B') \leq m + m'$. Therefore, for some $n'' \geq n$ and $m'' \geq m$, it is true that n'' boys kissed m'' girls, and those n boys talked about in the sentence *n boys kissed m girls* is a subset of these n'' boys, and those m girls a subset of these m'' girls. So, we end up considering something like Digram 2 after all. It is true that without giving specific numbers for n'' and m'', we cannot determine the exact diagram to look at, and it may very well be the case that the participants in the conversation do not know these numbers. However, my purpose here is to demonstrate how additional background information can affect the entailment relation, and Diagram 2 suffices for that. The same remark applies to the discussion of "distribution-describing" contexts in the rest of this paper.

element, which means that D_e becomes a complete Boolean lattice (i.e. complete Boolean algebra). In addition to the join operation $\sqcup$, let us denote the meet operation with $\sqcap$, complementation with $\neg$, and the top and the bottom elements with $\top$ and $\bot$ respectively. Since D_e is complete, for every family $\{x_i\}_{i\in I}$ of individuals, $\bigsqcup\{x_i\}_{i\in I}$ and $\bigsqcap\{x_i\}_{i\in I}$ exist, for which we may also write $\bigsqcup_{i\in I} x_i$ and $\bigsqcap_{i\in I} x_i$ respectively. The "part-of" relation $\sqsubseteq$ and the join and the meet operations are connected via the following equivalence:

(19) $x \sqsubseteq y$ iff $x \sqcup y = y$ iff $x \sqcap y = x$

Furthermore, the following laws hold for all $x, y, z \in D_e$:

(20) a. [Idempotency] $x \sqcup x = x, \quad x \sqcap x = x.$
b. [Associativity] $x \sqcup (y \sqcup z) = (x \sqcup y) \sqcup z, \quad x \sqcap (y \sqcap z) = (x \sqcap y) \sqcap z.$
c. [Commutativity] $x \sqcup y = y \sqcup x, \quad x \sqcap y = y \sqcap x.$
d. [Distributivity] $x \sqcup (y \sqcap z) = (x \sqcap y) \sqcup (z \sqcap z), \quad x \sqcup (y \sqcap z) = (x \sqcup y) \sqcap (x \sqcup z).$
e. [Absorption] $x \sqcup (x \sqcap y) = x, \quad x \sqcup (x \sqcap y) = x.$
f. [Identity] $x \sqcup \bot = x, \quad x \sqcap \top = x.$
g. [Annihilation] $x \sqcap \bot = \bot, \quad x \sqcup \top = \top.$
h. [Complement] $x \sqcup \neg x = \top, \quad x \sqcap \neg x = \bot.$
i. [Double complement] $\neg(\neg x) = x.$
j. [De Morgan's laws] $\neg(x \sqcup y) = \neg x \sqcap \neg y, \quad \neg(x \sqcap y) = \neg x \sqcup \neg y.$

Lemma 6. *For all $x, y \in D_e$, we have*

(i) $x \sqcup (\neg x \sqcap y) = x \sqcup y,$
(ii) $x \sqcap (\neg x \sqcap y) = \bot.$

Proof. (i) By distributivity, complement, commutativity and identity laws, $x \sqcup (\neg x \sqcap y) = (x \sqcup \neg x) \sqcap (x \sqcup y) = \top \sqcap (x \sqcup y) = (x \sqcup y) \sqcap \top = x \sqcup y$. (ii) By associativity, complement, commutativity and annihilation laws, $x \sqcap (\neg x \sqcap y) = (x \sqcap \neg x) \sqcap y = \bot \sqcap y = y \sqcap \bot = \bot$. □

Now that our ontology has the bottom element $\bot$, measures should be defined for $\bot$ as well. Below is our definition of measure.

(21) μ is a measure iff μ is a function from D_e into $\mathbb{R} \cup \{\infty\}$ that satisfies the following three conditions:
i. [Nonnegativity] For every $x \in D_e$, $\mu(x) \geq 0$.
ii. $\mu(\bot) = 0$.
iii. [Countable additivity] If $\{x_n\}_{n\in\mathbb{N}}$ is a family of pairwise non-overlapping individuals (i.e. for all $i, j \in \mathbb{N}$, $i \neq j \rightarrow x_i \sqcap x_j = \bot$), then $\mu\left(\bigsqcup\{x_n\}_{n\in\mathbb{N}}\right) = \sum_{n=1}^{\infty} \mu(x_n)$.

If we assume $\mu(\bot) < \infty$, the second condition can actually be derived from the third. Since $\bot \sqcup \bot = \bot$ and $\bot \sqcap \bot = \bot$, the additivity of μ implies $\mu(\bot) = \mu(\bot \sqcup \bot) = \mu(\bot) + \mu(\bot)$, so $\mu(\bot) < \infty$ implies $\mu(\bot) = 0$.[11] For more on Boolean algebras and measures on them, the reader is referred to Givant and Halmos (2009).

[11] $\mu(x) = 0$ does not necessarily imply $x = \bot$. We could have an alternative definition of measure where the nonnegativity condition is replaced with the following positivity condition:

Lemma 7. *Let μ be a measure. For all $x, y \in D_e$ such that $x \sqsubseteq y$,*

(i) $\mu(\neg x \sqcap y) = \mu(y) - \mu(x)$,
(ii) $\mu(x) \leq \mu(y)$.

Proof. Let $x \sqsubseteq y$, i.e., $x \sqcup y = y$ (by (19)). By Lemma 6(i)(ii) and by the additivity of μ, we have $\mu(y) = \mu(x \sqcup y) = \mu(x \sqcup (\neg x \sqcap y)) = \mu(x) + \mu(\neg x \sqcap y)$, and hence $\mu(\neg x \sqcap y) = \mu(y) - \mu(x)$, proving (i). Since $\mu(\neg x \sqcap y) \geq 0$ by the nonnegativiey of μ, (ii) follows from (i). □

Lemma 8. *Let μ be a measure. If $\{x_n\}_{n\in\mathbb{N}}$ is a monotone-increasing sequence of individuals (i.e., $x_1 \sqsubseteq x_2 \sqsubseteq \dots$) and $\lim_{n\to\infty} \mu(x_n) < \infty$, then $\mu\left(\bigsqcup\{x_n\}_{n\in\mathbb{N}}\right) = \lim_{n\to\infty} \mu(x_n)$.*

Proof. Define a sequence $\{y_n\}_{n\in\mathbb{N}}$ of individuals inductively by $y_1 = x_1$ and $y_n = \neg x_{n-1} \sqcap x_n$ for $n > 1$. Let $i, j \in \mathbb{N}$ and $i < j$. Since $y_i \sqsubseteq x_i \sqsubseteq x_{j-1}$, we have $y_i \sqcap y_j \sqsubseteq x_{j-1} \sqcap y_j = x_{j-1} \sqcap (\neg x_{j-1} \sqcap x_j) = \bot$ by Lemma 6(ii). This shows that $\{y_n\}_{n\in\mathbb{N}}$ is a family of pairwise non-overlapping individuals. We can furthermore show that $\bigsqcup_{n\in\mathbb{N}} x_n = \bigsqcup_{n\in\mathbb{N}} y_n$ as follows. It is obvious that $\bigsqcup_{n\in\mathbb{N}} y_n \sqsubseteq \bigsqcup_{n\in\mathbb{N}} x_n$ since $y_n \sqsubseteq x_n$ for every $n \in \mathbb{N}$. To see that $\bigsqcup_{n\in\mathbb{N}} x_n \sqsubseteq \bigsqcup_{n\in\mathbb{N}} y_n$ holds as well, observe that $x_n = \bigsqcup_{k=1}^{n} y_k$ holds for all $n \in \mathbb{N}$, which can be easily shown by induction on n. Therefore, by the countable additivity of μ and by Lemma 7(i),

$$\mu\left(\bigsqcup\{x_n\}_{n\in\mathbb{N}}\right) = \mu\left(\bigsqcup\{y_n\}_{n\in\mathbb{N}}\right) = \sum_{n=1}^{\infty} \mu(y_n) = \mu(x_1) + \sum_{n=1}^{\infty} (\mu(\neg x_{n-1} \sqcap x_n))$$
$$= \mu(x_1) + \sum_{n=1}^{\infty} (\mu(x_n) - \mu(x_{n-1})) = \lim_{n\to\infty} \mu(x_n).$$ □

Lemma 9. *Let μ be a measure.*

(i) *For all $x, y \in D_e$, $\mu(x \sqcup y) \leq \mu(x) + \mu(y)$.*
(ii) *For every family $\{x_i\}_{i\in I}$ of individuals with I at most countable, $\mu\left(\bigsqcup\{x_i\}_{i\in I}\right) \leq \sum_{i\in I} \mu(x_i)$.*

Proof. Since $(\neg x \sqcap y) \sqcap y = \neg x \sqcap (y \sqcap y) = \neg x \sqcap y$, we have $\neg x \sqcap y \sqsubseteq y$ (by (19)). Hence, by Lemma 6 and Lemma 7(ii), $\mu(x \sqcup y) = \mu(x \sqcup (\neg x \sqcap y)) = \mu(x) + \mu(\neg x \sqcap y) \leq \mu(x) + \mu(y)$, proving (i). (ii) easily follows from (i). □

One popular view on mass nouns is that their denotations have no atoms (Landman 1991, pp. 312–313). On this view, whenever an individual x is beer, x has a proper part which is also beer. Since we are talking about measuring individuals, a natural adaptation of this idea will be (22). We will refer to this as the *nonatomicity of mass measurement*:[12]

i′. [Positivity] For every $x \in D_e$ such that $x \neq \bot$, $\mu(x) > 0$.

On this alternative definition, $\mu(x) = 0$ would imply $x = \bot$.

[12]If we assumed positivity instead of nonnegativity for measures (see footnote 11), (22) would follow from the nonatomicity assumption for mass denotations, which can be stated as follows:

(i) Nonatomicity of mass denotations
Let N be the denotation of a mass noun. For all $x \in D_e$ such that $N(x)$ holds, x is not an atom (i.e. there is some individual y such that $y \sqsubseteq x$, $y \neq \bot$ and $y \neq x$).

(22) Nonatomicity of mass measurement
Let N be the denotation of a mass noun and μ be a measure. For all $x \in D_e$ such that $N(x)$ and $\mu(x) > 0$, there is some individual y such that $y \sqsubseteq x$, $y \neq x$ and $0 < \mu(y) < \mu(x)$.

Lemma 10. *Let N be the denotation of a mass noun and μ be a measure. For all $x, y \in D_e$, if $N(y)$ holds and $x \sqsubseteq y$ and $\mu(x) < \mu(y)$, then there is some individual z such that $x \sqsubseteq z \sqsubseteq y$ and $\mu(x) < \mu(z) < \mu(y)$.*

Proof. Suppose that $N(y)$ holds and $x \sqsubseteq y$ and $\mu(x) < \mu(y)$. By Lemma 7(i), $\mu(\neg x \sqcap y) = \mu(y) - \mu(x) > 0$, so $\neg x \sqcap y \neq \bot$. Then, since $\neg x \sqcap y \sqsubseteq y$, by the divisivity of N, we have $N(\neg x \sqcap y)$. By the nonatomicity of mass measurement, there is an individual w such that $w \sqsubseteq \neg x \sqcap y$ and $0 < \mu(w) < \mu(\neg x \sqcap y)$. Put $z = x \sqcup w$. We show that z is a desired individual. First, since $w \sqsubseteq \neg x \sqcap y \sqsubseteq y$, we have $w \sqcup y = y$, and hence $z \sqcup y = (x \sqcup w) \sqcup y = x \sqcup (w \sqcup y) = x \sqcup y = y$ (due to $x \sqsubseteq y$). This shows $z \sqsubseteq y$. Since $x \sqsubseteq x \sqcup w = z$, we have $x \sqsubseteq z \sqsubseteq y$. It remains to show that $\mu(x) < \mu(z) < \mu(y)$. Since $w \sqsubseteq \neg x \sqcap y$, we have $x \sqcap w \sqsubseteq x \sqcap (\neg x \sqcap y) = \bot$ by Lemma 6(ii), so $x \sqcap w = \bot$. Therefore, by the additivity of μ,

$$\mu(z) = \mu(x \sqcup w) = \mu(x) + \mu(w).$$

Since $\mu(w) > 0$, this shows $\mu(x) < \mu(z)$. Furthermore, since $\mu(w) < \mu(\neg x \sqcap y)$, by Lemma 7(i),

$$\mu(z) = \mu(x) + \mu(w) < \mu(x) + \mu(\neg x \sqcap y) = \mu(x) + (\mu(y) - \mu(x)) = \mu(y). \qquad \square$$

Given Lemma 10, we can see that when the domain of a measure μ is restricted to the denotation of a mass noun, μ is a *nonatomic measure* in the following sense:

(23) A measure μ is called nonatomic iff there is no element x such that
   i. $\mu(x) > 0$ and
   ii. for all y, if $y \sqsubseteq x$ and $y \neq x$, then $\mu(y) = 0$.

It is known that nonatomic measures take a continuum of values. The following lemma is a straightforward adaptation of this result to the current setting.[13] It plays a fundamental role in the discussion of entailments where mass nouns are involved.

Lemma 11 (uses the axiom of choice)**.** *Let N be the denotation of a mass noun and μ be a measure. For all $x \in D_e$, if $N(x)$ holds and $\mu(x) = r$, then for all $s \in \mathbb{R}$ such that $0 \leq s \leq r$, there is an individual y such that $y \sqsubseteq x$ and $\mu(y) = s$.*

Proof. Suppose that $N(x)$ holds and $\mu(x) = r$. We consider the set M of monotone-increasing partial functions ϕ from $[0, r]$ into individuals such that $\phi(r) = x$, and $\phi(p)$, whenever it is defined, gives an individual whose measurement by μ is p. That is,

$$M = \{\phi \mid \phi \text{ is a partial function from } [0,r] \text{ into } D_e \text{ such that } \phi(r) = x \wedge \\ \forall t_1, t_2 \in \operatorname{dom}\phi\, [t_1 \leq t_2 \rightarrow \phi(t_1) \sqsubseteq \phi(t_2)] \wedge \forall t \in \operatorname{dom}\phi\, [\mu(\phi(t)) = t]\}.$$

[13] Sierpiński (1922) was the first to prove that nonatomic measures take a continuum of values. Although Sierpiński did not do so, it now seems standard to use Zorn's lemma (and hence the axiom of choice) to prove this kind of result. Sierpiński's theorem was proved for measures on subsets of a Euclidean space, which has an intrinsic metric. In our case, the domain of the measure is the set of individuals, where no intrinsic metric is given. The use of Zorn's lemma therefore seems essential in our case.

The elements of M, which are partial functions, are partially ordered by inclusion of their graphs. If C is a chain (i.e. totally ordered subset) in M, we can easily see that $\bigcup C \in M$, and thus $\bigcup C$ gives an upper bound of C in M. Therefore, by Zorn's lemma, M has a maximal element ϕ_0. Below, we show that $\operatorname{dom}\phi_0 = [0,r]$. It will then follow that for any given s such that $0 \le s \le r$, we have $\phi_0(s) \sqsubseteq \phi_0(r) = x$ and $\mu(\phi_0(s)) = s$, proving the claim. It is obvious that $0 \in \operatorname{dom}\phi_0$, because if $0 \notin \operatorname{dom}\phi_0$, then $\phi_0 \cup (0,\bot)$ would be an element of M properly containing ϕ_0, in contradiction to the maximality of ϕ_0. We thus need only show that if $0 < s < r$ then $s \in \operatorname{dom}\phi_0$.

Suppose to the contrary that there is a real number s such that $0 < s < r$ and $s \notin \operatorname{dom}\phi_0$. Let $p_0 = \sup\{t \in \operatorname{dom}\phi_0 \mid t < s\}$ and $q_0 = \inf\{t \in \operatorname{dom}\phi_0 \mid s < t\}$. We first show $p_0, q_0 \in \operatorname{dom}\phi_0$.

First, we show that $p_0 \in \operatorname{dom}\phi_0$. Let $\{p_n\}_{n\in\mathbb{N}}$ be a monotone-increasing sequence in $\operatorname{dom}\phi_0$ converging to p_0, i.e., $p_1 \le p_2 \le \dots$ and $\lim_{n\to\infty} p_n = p_0$. Let $a = \bigsqcup\{\phi_0(p_n)\}_{n\in\mathbb{N}}$. Since both $\{p_n\}_{n\in\mathbb{N}}$ and ϕ_0 are monotone-increasing, so is $\{\phi_0(p_n)\}_{n\in\mathbb{N}}$. Hence, by Lemma 8,

$$\mu(a) = \mu\left(\bigsqcup\{\phi_0(p_n)\}_{n\in\mathbb{N}}\right) = \lim_{n\to\infty}\mu(\phi_0(p_n)) = \lim_{n\to\infty} p_n = p_0.$$

We can then show that $\phi_0 \cup (p_0,a)$ is an element of M. First, if $t \in \operatorname{dom}\phi_0$ and $t < p_0$, then $t < p_m$ for some $m \in \mathbb{N}$, and hence by the monotonicity of ϕ_0, we have $\phi_0(t) \sqsubseteq \phi_0(p_m) \sqsubseteq a$. Also, if $t \in \operatorname{dom}\phi_0$ and $p_0 < t$, then for all $n \in \mathbb{N}$, $p_n < t$ and hence $\phi_0(p_n) \sqsubseteq \phi_0(t)$ by the monotonicity of ϕ_0, and hence $a \sqsubseteq \phi_0(t)$. This shows that $\phi_0 \cup (p_0,a)$ is monotone-increasing. Also, $\mu((\phi_0 \cup (p_0,a))(p_0)) = \mu(a) = p_0$. Thus, $\phi_0 \cup (p_0,a) \in M$, and since ϕ_0 is a maximal element of M, we in fact have $\phi_0 \cup (p_0,a) = \phi_0$. Hence $p_0 \in \operatorname{dom}\phi_0$.

Next, we show that $q_0 \in \operatorname{dom}\phi_0$. Let $\{q_n\}_{n\in\mathbb{N}}$ be a monotone-decreasing sequence in $\operatorname{dom}\phi_0$ such that $\lim_{n\to\infty} q_n = q_0$ and let $b = \bigsqcap\{\phi_0(q_n)\}_{n\in\mathbb{N}}$. Let $y_n = \neg\phi_0(q_n) \sqcap x$. For each i, we have $q_{i+1} \le q_i$ and hence $\phi(q_{i+1}) \sqsubseteq \phi(q_i)$, that is, $\phi(q_{i+1}) \sqcup \phi(q_i) = \phi(q_i)$. Therefore, by associativity, commutativity, De Morgan and idempotency laws, we have

$$\begin{aligned} y_i \sqcap y_{i+1} &= (\neg\phi_0(q_i) \sqcap x) \sqcap (\neg\phi_0(q_{i+1}) \sqcap x) = (\neg\phi_0(q_{i+1}) \sqcap \neg\phi_0(q_i)) \sqcap (x \sqcap x) \\ &= \neg(\phi_0(q_{i+1}) \sqcup \phi_0(q_i)) \sqcap x = \neg\phi_0(q_i) \sqcap x = y_i, \end{aligned}$$

and hence $y_i \sqsubseteq y_{i+1}$. Thus, $\{y_n\}_{n\in\mathbb{N}}$ is monotone-increasing. For each n, since $q_n \le r$, we have $\phi(q_n) \sqsubseteq \phi(r) = x$. Therefore, by Lemma 8 and by Lemma 7(i),

$$\begin{aligned} \mu\left(\bigsqcup\{y_n\}_{n\in\mathbb{N}}\right) &= \lim_{n\to\infty}\mu(y_n) = \lim_{n\to\infty}\mu(\neg\phi_0(q_n) \sqcap x) \\ &= \lim_{n\to\infty}(\mu(x) - \mu(\phi_0(q_n))) = \lim_{n\to\infty}(r - q_n) = r - q_0. \end{aligned}$$

By distributivity, De Morgan and double complement laws, we have

$$\begin{aligned} \neg\left(\bigsqcup\{y_n\}_{n\in\mathbb{N}}\right) \sqcap x &= \neg\left(\bigsqcup\{(\neg\phi_0(q_n) \sqcap x)\}_{n\in\mathbb{N}}\right) \sqcap x = \neg\left(\bigsqcup\{\neg\phi_0(q_n)\}_{n\in\mathbb{N}} \sqcap x\right) \sqcap x \\ &= \neg\left(\neg\bigsqcap\{\phi_0(q_n)\}_{n\in\mathbb{N}} \sqcap x\right) \sqcap x = \neg(\neg b \sqcap x) \sqcap x = (\neg\neg b \sqcup x) \sqcap x \\ &= (b \sqcup \neg x) \sqcap x = (b \sqcap x) \sqcup (\neg x \sqcap x) = (b \sqcap x) \sqcup \bot = b \sqcap x = b. \end{aligned}$$

The last equality is due to $b \sqsubseteq x$. Since $y_n \sqsubseteq x$ for all $n \in \mathbb{N}$, we have $\bigsqcup\{y_n\}_{n\in\mathbb{N}} \sqsubseteq x$. Hence, by Lemma 7(i),

$$\mu(b) = \mu\left(\neg\left(\bigsqcup\{y_n\}_{n\in\mathbb{N}}\right) \sqcap x\right) = \mu(x) - \mu\left(\bigsqcup\{y_n\}_{n\in\mathbb{N}}\right) = r - (r - q_0) = q_0.$$

By an argument parallel to the one made in the previous paragraph, we can then show that $\phi_0 \cup (q_0, b) \in M$, and the maximality of ϕ_0 implies that $q_0 \in \operatorname{dom}\phi_0$.

Now, since $s \notin \operatorname{dom}\phi_0$, $p_0, q_0 \in \operatorname{dom}\phi_0$ and $p_0 \leq s \leq q_0$, we in fact have $p_0 < s < q_0$ and hence $\mu(\phi_0(p_0)) < \mu(\phi_0(q_0))$. By the definitions of p_0 and q_0, no number between p_0 and q_0 should be in $\operatorname{dom}\phi_0$. However, since $\phi_0(q_0) \sqsubseteq x$ and $\phi_0(q_0) \neq \bot$, we have $N(\phi_0(q_0))$ by the divisivity of N, and by Lemma 10, there must be an individual c such that $\phi_0(p_0) \sqsubseteq c \sqsubseteq \phi_0(q_0)$ and $\mu(\phi_0(p_0)) < \mu(c) < \mu(\phi_0(q_0))$. This contradicts the maximality of ϕ_0. □

With Lemma 11, we can see, for instance, that *John drank 2 liters of beer* entails *John drank 1 liter of beer*, which in turn entails *John drank 0.6 liters of beer*.

4.2 Existential context

Entailments of cumulative sentences with one count noun and one mass noun in existential contexts are covered by the following two propositions.

Proposition 12. *For all $n \in \mathbb{N}$ and for all $s, t \in \mathbb{R}$ such that $s > t > 0$,*

$$nR^{\mathrm{cm}}s \models nR^{\mathrm{cm}}t.$$

Proof. Let $s > t > 0$. Suppose that $nR^{\mathrm{cm}}s$ holds, i.e., there are individuals A and B such that

$$\#(A) = n \wedge N_1(A) \wedge \mu(B) = s \wedge N_2(B) \wedge V(A, B).$$

Since V is divisive, for each $a \in \mathrm{AT}(A)$, there is some $B' \sqsubseteq B$ such that $V(a, B')$. Let us write this B' by $f(a)$. That is, f is a function from $\mathrm{AT}(A)$ into $\{y \mid y \sqsubseteq B\}$ such that $V(a, f(a))$. Since $s > t > 0$, we have $t/s < 1$. Then, Lemma 11 tells us that for each $a \in \mathrm{AT}(A)$, there is an individual B'' such that $B'' \sqsubseteq f(a)$ and $\mu(B'') = (t/s)(1/n)\mu(f(a))$. Write this B'' by $g(a)$. Thus, for each $a \in \mathrm{AT}(A)$, $g(a) \sqsubseteq f(a)$ and $\mu(g(a)) = (t/s)(1/n)\mu(f(a))$. By the divisivity of N_2 and V, $N_2(g(a))$ and $V(a, g(a))$ hold. Let $G = \bigsqcup\{g(a) \mid a \in \mathrm{AT}(A)\}$. Then, $G \sqsubseteq B$, and by the cumulativity of V, we have $V(A, G)$. Furthermore, by Lemma 9(ii) and Lemma 7(ii),

$$\begin{aligned}
\mu(G) &= \mu\left(\bigsqcup\{g(a) \mid a \in \mathrm{AT}(A)\}\right) \\
&\leq \sum_{a \in \mathrm{AT}(A)} \mu(g(a)) = (t/s)(1/n) \sum_{a \in \mathrm{AT}(A)} \mu(f(a)) \\
&\leq (t/s)(1/n)\left(n \cdot \max_{a \in \mathrm{AT}(A)} \mu(f(a))\right) = (t/s) \max_{a \in \mathrm{AT}(A)} \mu(f(a)) \\
&\leq (t/s)\mu(B) = (t/s)s = t.
\end{aligned}$$

Now, if $\mu(G) = t$, then we have

$$\#(A) = n \wedge N_1(A) \wedge \mu(G) = t \wedge N_2(G) \wedge V(A, G),$$

which implies $nR^{\mathrm{cm}}t$. So suppose $\mu(G) < t$. Since $G \sqsubseteq B$, by Lemma 7(i),

$$\mu(\neg G \sqcap B) = \mu(B) - \mu(G) = s - \mu(G) > t - \mu(G) > 0.$$

Therefore, by Lemma 11, there is an individual H such that $H \sqsubseteq \neg G \sqcap B$ and $\mu(H) = t - \mu(G)$. Put $I = G \sqcup H$. Since $G \sqcap H \sqsubseteq G \sqcap (\neg G \sqcap B) = \bot$ by Lemma 6(ii), we have $G \sqcap H = \bot$, and hence by the additivity of μ, $\mu(I) = \mu(G \sqcup H) = \mu(G) + \mu(H) = \mu(G) + (t - \mu(G)) = t$. Since $I \sqsubseteq B$, by divisivity of N_2, we have $N_2(I)$. Also, since $H \sqsubseteq B$, by the divisitvity of V, we have $V(A', H)$ for some $A' \sqsubseteq A$, and by the cumulativity of V, we have $V(A \sqcup A', G \sqcup H)$, that is, $V(A, I)$. Hence

$$\#(A) = n \wedge N_1(A) \wedge \mu(I) = t \wedge N_2(H) \wedge V(A, I),$$

implying $nR^{\mathrm{cm}}t$. □

Proposition 13. *For all $n \in \mathbb{N}$ such that $n \geq 2$ and for all $s \in \mathbb{R}$ such that $s > 0$,*

$$nR^{\mathrm{cm}}s \models (n-1)R^{\mathrm{cm}} \left(\frac{n-1}{n} \cdot s \right).$$

Proof. By reductio ad absurdum. Let $n \geq 2$ and $s > 0$. Assume that $nR^{\mathrm{cc}}s$ holds but $(n-1)R^{\mathrm{cm}} \left(\frac{n-1}{n} \cdot s \right)$ does not. Since $nR^{\mathrm{cm}}s$ holds, there are individuals A and B such that

$$\#(A) = n \wedge N_1(A) \wedge \mu(B) = s \wedge N_2(B) \wedge V(A, B).$$

For every $a \in \mathrm{AT}(A)$, put

$$f(a) = \bigsqcup(\mathrm{AT}(A) \backslash \{a\}), \quad g(a) = \bigsqcup\{y \sqsubseteq B \mid \exists x \sqsubseteq f(a)[V(x, y)]\}.$$

It is easy to see that $f(a) \sqsubseteq A$, $g(a) \sqsubseteq B$, $\#(f(a)) = n-1$, and $N_1(f(a))$, $N_2(g(a))$ and $V(f(a), g(a))$ hold, and hence $(n-1)R^{\mathrm{cm}}\mu(g(a))$ holds. If it were the case that $\mu(g(a)) \geq \frac{n-1}{n} \cdot s$, then by Proposition 12, $(n-1)R^{\mathrm{cc}}(\frac{n-1}{n} \cdot s)$ should hold, contrary to the assumption. So $\mu(g(a)) < \frac{n-1}{n} \cdot s$. Since $g(a) \sqsubseteq B$, by Lemma 7(i),

$$\mu(\neg g(a) \sqcap B) = \mu(B) - \mu(g(a)) > s - \frac{n-1}{n} \cdot s = \frac{s}{n}.$$

Now, if $\{\neg g(a) \sqcap B \mid a \in \mathrm{AT}(A)\}$ were a family of pairwise non-overlapping individuals, it would follow from the additivity of μ that

$$\mu(\bigsqcup\{\neg g(a) \sqcap B \mid a \in \mathrm{AT}(A)\}) = \sum_{a \in \mathrm{AT}(A)} \mu(\neg g(a) \sqcap B) > n \cdot \frac{s}{n} = s.$$

However, since $\bigsqcup\{\neg g(a) \sqcap B \mid a \in \mathrm{AT}(A)\} \sqsubseteq B$, by Lemma 7(ii), we have

$$\mu(\bigsqcup\{\neg g(a) \sqcap B \sqsubseteq B \mid a \in \mathrm{AT}(A)\}) \leq \mu(B) = s,$$

so this is a contradiction. This shows that $\{\neg g(a) \sqcap B \mid a \in \mathrm{AT}(A)\}$ is not a family of pairwise non-overlapping individuals, so there exist some $a_1, a_2 \in \mathrm{AT}(A)$, $a_1 \neq a_2$ such that

$$(\neg g(a_1) \sqcap B) \sqcap (\neg g(a_2) \sqcap B) \neq \bot.$$

Since $V(A, B)$ holds and $(\neg g(a_1) \sqcap B) \sqcap (\neg g(a_2) \sqcap B) \sqsubseteq B$, by the divisivity of V, there is some $A' \sqsubseteq A$, $A' \neq \bot$ such that $V(A', (\neg g(a) \sqcap \neg g(a_1) \sqcap B) \sqcap (\neg g(a_2) \sqcap B))$. Let $a' \in \mathrm{AT}(A')$. Then, by the divisivity of V, there is some $C \sqsubseteq (\neg g(a_1) \sqcap B) \sqcap (\neg g(a_2) \sqcap B)$, $C \neq \bot$ such that $V(a', C)$. Now, if $a' \neq a_1$, then $a' \sqsubseteq f(a_1)$ and it therefore follows from the definition of g that $C \sqsubseteq g(a_1)$. Then, since $C \sqsubseteq (\neg g(a_1) \sqcap B) \sqcap (\neg g(a_2) \sqcap B) \sqsubseteq \neg g(a_1) \sqcap B$, we have $C \sqsubseteq g(a_1) \sqcap (\neg g(a_1) \sqcap B) = \bot$ by Lemma 6(ii), contradicting $C \neq \bot$. This shows that $a' = a_1$. Likewise, we can show that $a' = a_2$. We thus obtain $a' = a_1 = a_2$, contradicting $a_1 \neq a_2$. □

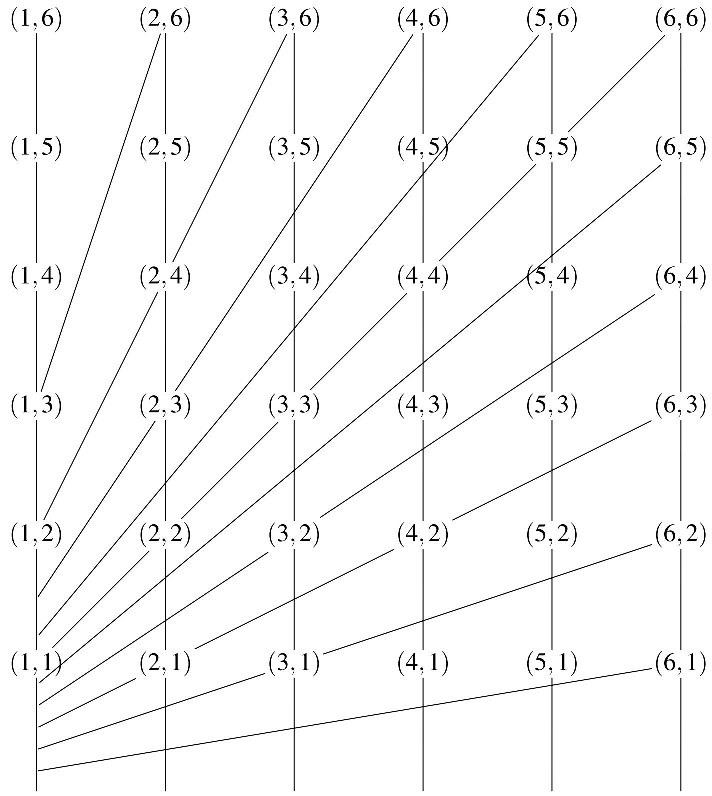

Diagram 3

It would be helpful to draw a diagram for the entailment relation between cumulative sentences whose truth conditions are of the form $nR^{\text{cm}}s$, just as we did in Section 3. However, now that s ranges over real numbers, we would need infinitely many nodes to represent those sentences, and hence, it is not really possible to draw such a diagram. Nevertheless, I have tried to visualize the relation, albeit schematically, in Diagram 3 for $1 \leq n \leq 6$ and $0 < s \leq 6$. Here, the vertical lines come from Proposition 12, while the oblique ones come from Proposition 13. Only the nodes with whole number coordinates have been explicitly marked, but the reader should know that on each vertical line lie infinitely many nodes. For instance, nothing is marked at the intersection of the third vertical line and the oblique line starting from node $(4,6)$, but this point represents $(3,4.5)$. Also, there should really be infinitely many oblique lines present in the diagram, but drawing them all would render the digram pitch-black and unintelligible, so only a small number of them are represented. We can read from the diagram, for instance, that *three boys drank 5 liters of beer* entails *two boys drank 3 liter of beer*, which in turn entails *two boys drank 2.7 liters of beer*.

4.3 Distribution-describing context

Let us now consider entailments in distribution-describing contexts. We want to know, for example, what is entailed by the sentence *n boys drank s liters of beer* when the background information ensures that more then n boys drank beer or more than s liters of beer was drunk by boys. As the following proposition shows, when it is given that more then n boys drank beer, we obtain some valid entailments. By contrast, it is not difficult to see that the information that more than s liters of beer was drunk by boys does not lead to new entailments.

Proposition 14. *For all $n, n' \in \mathbb{N}$ such that $n < n'$ and for all $s \in \mathbb{R}$,*

$$nR^{\mathrm{cm}}s \wedge \exists X \exists Y[\#(X) = n' \wedge N_1(X) \wedge N_2(Y) \wedge V(X,Y)] \models (n+1)R^{\mathrm{cm}}s.$$

Proof. Suppose that both $nR^{\mathrm{cm}}s$ and $\exists X \exists Y[\#(X) = n' \wedge N_1(X) \wedge N_2(Y) \wedge V(X,Y)]$ hold. That is, there are individuals A and B such that

$$\#(A) = n \wedge N_1(A) \wedge \mu(B) = s \wedge N_2(B) \wedge V(A,B),$$

and there are individuals C and D such that

$$\#(C) = n' \wedge N_1(A) \wedge N_2(D) \wedge V(C,D).$$

Since $\#(A) = n < n' = \#(C)$, there is at least one individual c such that $c \in \mathrm{AT}(C)$ but $c \notin \mathrm{AT}(A)$. $N_1(c)$ holds by the divisivity of N_1, and $\#(A \sqcup c) = n+1$. Since $V(C,D)$ holds, by the divisivity of V, there is some individual $D' \sqsubseteq D$, $D' \neq \bot$ such that $V(c,D')$. By the divisivity of N_2, we have $N_2(D')$. We show that $(n+1)R^{\mathrm{cm}}s$ holds by examining the following two cases.

Case 1: $B \sqcap D' \neq \bot$. Since $B \sqcap D' \sqsubseteq D'$, by divisivity, we have $N_2(B \sqcap D')$ and $V(c, B \sqcap D')$. Then, by the cumulativity of V, we have $V(A \sqcup c, B \sqcup (B \sqcap D'))$, but since $B \sqcup (B \sqcap D') = B$ by absorption law, we in fact have $V(A \sqcup c, B)$. Hence

$$\#(A \sqcup c) = n+1 \wedge N_1(A \sqcup c) \wedge \mu(B) = s \wedge N_2(B) \wedge V(A \sqcup c, B),$$

implying $(n+1)R^{\mathrm{cm}}s$.

Case 2: $B \sqcap D' = \bot$. If $\mu(D') = 0$,[14] by the additivity of μ, we have $\mu(B \sqcup D') = \mu(B) + \mu(D') = s + 0 = s$, and hence

$$\#(A \sqcup c) = n+1 \wedge N_1(A \sqcup c) \wedge \mu(B \sqcup D') = s \wedge N_2(B \sqcup D') \wedge V(A \sqcup c, B \sqcup D'),$$

implying $(n+1)R^{\mathrm{cm}}s$. So suppose $\mu(D') > 0$. By Lemma 11, there is an individual E such that $E \sqsubseteq D'$ and $0 < \mu(E) < \min(\mu(D'), s)$. Since $E \sqsubseteq D'$, by divisivity, we have $N_2(E)$ and $V(c,E)$. Since $0 < s - \mu(E) < s$, by Proposition 12, $nR^{\mathrm{cm}}s \models nR^{\mathrm{cm}}(s - \mu(E))$, and more precisely, following the proof of Proposition 12, we see that there is an individual $I \sqsubseteq B$ such that $\mu(I) = s - \mu(E)$ and $N_2(I)$ and $V(A,I)$ hold. Put $J = I \sqcup E$. By cumulativity, $N_2(J)$ and $V(A \sqcup c, J)$ hold. Since $I \sqsubseteq B$ and $E \sqsubseteq D'$, we have $I \sqcap E \sqsubseteq B \sqcap D' = \bot$ and hence $I \sqcap E = \bot$. Therefore, by the additivity of μ,

$$\mu(J) = \mu(I \sqcup E) = \mu(I) + \mu(E) = (s - \mu(E)) + \mu(E) = s.$$

Hence

$$\#(A \sqcup c) = n+1 \wedge N_1(A \sqcup c) \wedge \mu(J) = s \wedge N_2(J) \wedge V(A \sqcup c, J),$$

implying $(n+1)R^{\mathrm{cm}}s$. □

[14] On the alternative theory where positivity is required of measures (see footnote 11), $D' \neq \bot$ would already imply $\mu(D') \neq 0$, so the possibility of $\mu(D') = 0$ would not need consideration.

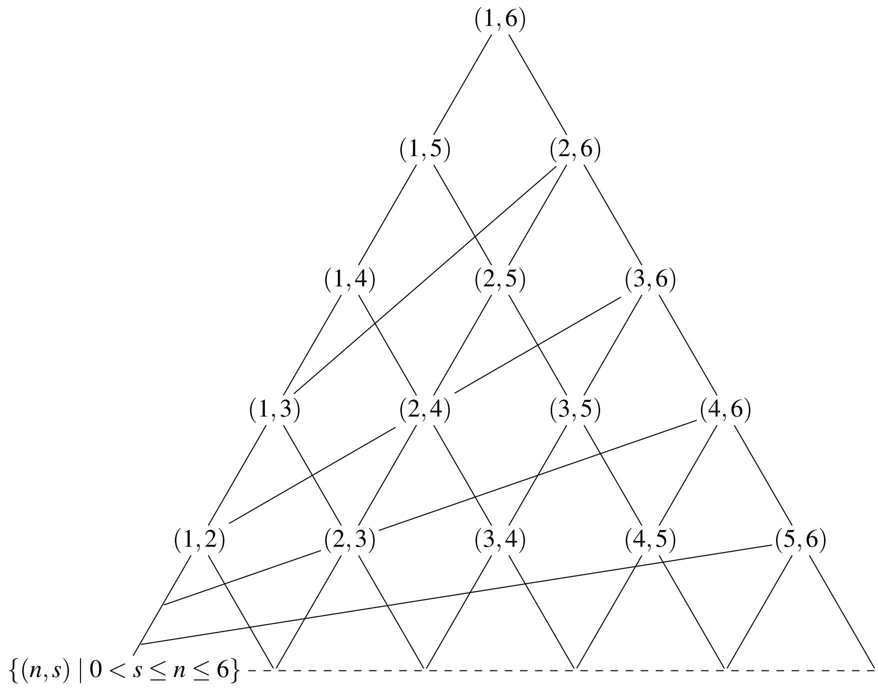

Diagram 4

In Diagram 4, by adding the effect of Proposition 14 to Diagram 3, I have tried to schematically visualize the entailment relation between sentences whose truth conditions are of the form $nR^{\mathrm{cm}}s$ for $1 \leq n \leq 6$ and $0 < s \leq 6$ in a context where it is given that $6R^{\mathrm{cm}}6$ holds. Again, infinitely many nodes and lines have been omitted. Note that all the sentences whose truth conditions are $nR^{\mathrm{cm}}s$ where $0 < s \leq n \leq 6$ have collapsed into the single node at the bottom left corner here.

5 Entailments of cumulative sentences with two mass nouns

Finally, let us consider cumulative sentences where both the nouns involved are mass nouns, i.e., sentences whose truth conditions are of the following form:

$$sR^{\mathrm{mm}}t \equiv \exists X \exists Y[\mu_1(X) = s \wedge N_1(X) \wedge \mu_2(Y) = t \wedge N_2(Y) \wedge V(X,Y)].$$

There do not seem to be many verbs that allow this kind of sentence, but (24) is an example in point.

(24) 16 grams of gas occupied 22.4 liters of volume.

(24) does not seem to entail or be entailed by any sentence of the form *s grams of gas occupied t liters of volume* where $s \neq 16$ or $t \neq 22.4$, be it in an existential context or in a distribution-describing context. Thus, we simply do not have any nontrivial entailments for cumulative sentences of this type.[15]

[15]Depending on the verb, we sometimes obtain some valid entailments. For instance, it seems to me that on the most natural interpretation, (i) does not entail or is entailed by *s tons of snow covered t square meters of land* where $s \neq 100$ or $t \neq 1000$, since I would normally think of separating the fallen snow vertically in my mind.

6 Discussion of Kirkfa's examples

Let us now return to Krifka's (1999) examples.

6.1 Three boys, perhaps even four boys, ate seven apples

According to Krifka, a scalar implicature is observed clearly with the object as in (8a), but not so with the subject as in (8b).

(8) a. Three boys ate seven apples, perhaps even eight apples.
b. ?Three boys, perhaps even four boys, ate seven apples.

Here, we are dealing with cumulative sentences with two count nouns in an existential context. So, a relevant diagram to look at is Diagram 1, where we see that $3R^{cc}8 \models 3R^{cc}7$ and $3R^{cc}7 \not\models 3R^{cc}8$. In other words, *three boys ate eight apples* is more informative than *three boys ate seven apples*. Then, the classic theory of scalar implicature predicts that the utterance of the latter sentence implicates the negation of the former. This coincides with Krifka's prediction and can explain the goodness of (8a); (8a) is expressly canceling this implicature. On the other hand, we have $4R^{cc}7 \not\models 3R^{cc}7$, as shown by model $\mathfrak{M}_1$ in Figure 1.[16] So *four boys ate seven apples* is not more informative than *three boys ate seven apples*. In the entailment diagram, the nodes that immediately dominate $(3,7)$ are $(3,8)$ and $(4,9)$.[17] Hence, what the classic theory predicts is that the utterance of $3R^{cc}7$ should implicate the negation of $3R^{cc}8$ and the negation of $4R^{cc}9$. Since $3R^{cc}7 \wedge \neg 3R^{cc}8 \wedge \neg 4R^{cc}9 \not\models \neg 4R^{cc}7$ as shown by model $\mathfrak{M}_2$ in Figure 1,[18] it is not predicted that we obtain $\neg 4R^{cc}7$ as a scalar implicature, contrary to Krifka's prediction.

It is then only natural that (8b) is judged to be not very acceptable, since it seems as if it were trying to cancel an implicature that were nonexistent. I should note, though, that quite a few people seem to find (8a) and (8b) to be equally good and see no contrast. My suggestion is that those who find (8b) to be degraded are either aware that $4R^{cc}7 \not\models 3R^{cc}7$ or uncertain whether $4R^{cc}7 \models 3R^{cc}7$. Then, in order to account for the difference in acceptability between (8a) and (8b), we do not need

(i) 100 tons of snow covered 1000 square meters of land.

However, as Makoto Kanazawa (p.c.) points out, if we separate the snow horizontally in our mind and focus on some thinner layer of snow, and if we deem such a layer alone, disregarding the rest of the snow, to be covering the land, we may say that for any s such that $0 < s < 100$, (i) entails *s tons of snow covered 1000 square meters of land*, though there might be some lower limit (since perhaps a snow layer cannot be so thin to be seen through in order to be deemed to be covering land).

As another example, Eric McCready (p.c.) points out that (ii) seems to entail *200 grams of sugar dissolves in t grams of water* for any $t > 100$. The trick here is that (ii) is interpreted as a description of a generic property of sugar, and almost reads like "200 grams of sugar can dissolve in 100 grams of water." This interpretation is obviously triggered by the present tense. As Uli Sauerland (p.c.) points out, when the sentence is changed into the past tense as in (iii), it does not entail *200 grams of sugar dissolved in t grams of water* for any $t > 100$.

(ii) 200 grams of sugar dissolves in 100 grams of water.

(iii) 200 grams of sugar dissolved in 100 grams of water.

[16] In $\mathfrak{M}_1$, it is true that four boys kissed seven girls, so $\mathfrak{M}_1 \models 4R^{cc}7$. However, we cannot find (plural) individuals X and Y such that $\#(X) = 3 \wedge \mathsf{boys}(X) \wedge \#(Y) = 7 \wedge \mathsf{girls}(Y) \wedge \mathsf{kiss}(X,Y)$. Hence $\mathfrak{M}_1 \not\models 3R^{cc}7$.

[17] $(4,9)$ is not included in Diagram 1, but one can easily confirm that $4R^{cc}9 \models 3R^{cc}7$ with Proposition 2.

[18] If we disregard b_4 and g_8, we can see that $\mathfrak{M}_2 \models 3R^{cc}7$. We can also see that $\mathfrak{M}_2 \models \neg 3R^{cc}8$ and $\mathfrak{M}_2 \models \neg 4R^{cc}9$. Yet, if we disregard g_7, we can see that $\mathfrak{M}_2 \models 4R^{cc}7$. Hence $3R^{cc}7 \wedge \neg 3R^{cc}8 \wedge \neg 4R^{cc}9 \not\models \neg 4R^{cc}7$.

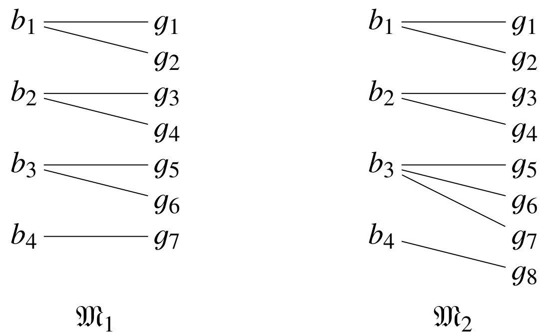

Figure 1: Models $\mathfrak{M}_1$ (left) and $\mathfrak{M}_2$ (right). *b*'s and *g*'s represent boys and girls respectively, and the line between a boy and a girl indicates that the boy kissed the girl.

to resort to some difference between the subject and the object as Krifka suggests, since it seems to be simply due to the different combinations of numbers. Because Diagram 1 is symmetrical, this point can be verified by switching the numbers. Since the same apple cannot be eaten by more than one person, let us first change the verb and the object noun:

(25) a. Three boys kissed seven girls, perhaps even eight girls.
b. ?Three boys, perhaps even four boys, kissed seven girls.

Again, not everyone will find a contrast here. Now, (26) gives the examples where the numbers in the subject and the object have been switched, and I believe that for those who find a contrast between (25a) and (25b), there will be the same degree of contrast between (26a) and (26b).

(26) a. ?Seven boys kissed three girls, perhaps even four girls.
b. Seven boys, perhaps even eight boys, kissed three girls.

Here, the oddness of (26a) can be attributed to the fact that $7R^{\mathrm{cc}}4 \not\models 7R^{\mathrm{cc}}3$.

A question remains, however, as to why some people find (8b) acceptable. I can think of two possible reasons.

First, even though $4R^{\mathrm{cc}}7 \not\models 3R^{\mathrm{cc}}7$, people might not be aware of this fact. Because of the fact that *four boys kissed Mary* entails *three boys kissed Mary*, by analogy, people might get the mistaken impression that *four boys kissed seven girls* entails *three boys kissed seven girls*. In that case, (8b) will be (wrongly) understood as expressly canceling a scalar implicature, and hence taken to be as good as (8a). This is not an implausible story. While truth-value judgement is something we can do out of linguistic intuition, telling whether a sentence entails another does not belong to our linguistic capacity but to our logical capacity, and people make logical errors all the time. In fact, many logic puzzles present a set of sentences and ask what consequence follows or doesn't follow. These puzzles challenge our logical capacity. If we had direct linguistic intuition about which sentence entails which other, they would cease to be puzzles.

Second, *perhaps even …* is not necessarily a device for canceling an implicature. Consider the following:

(27) Aliens have already visited Earth, and perhaps even abducted some of us.

Since the utterance of *aliens have visited Earth* alone clearly does not implicate that it is not the case that aliens have abducted some humans, the function of *perhaps even abducted some of us* is not to cancel some implicature, but to add some more information. Nevertheless, one might think that what follows *perhaps even* should entail what precedes it. Surely, in order to be able to abduct earthlings, aliens need to be physically present on/around Earth, so their abducting us entails their visiting us. This seems to explain the naturalness of (27). It is indeed true that when random phrases with no entailment relation are combined, we generally obtain quite queer sentences, as illustrated by the following:

(28) ?John has a lisp, and perhaps even a lot of body hair.

Hairiness does not entail a lisp, so this lack of entailment seems to explain the oddness of the example. However, in some cases, a combination with no entailment relation does seem to work. Since there are many ugly single men, bachelorhood surely does not entail handsomeness. Even so, (29) sounds like a suitable sentence to utter when in search of husband material.

(29) John is handsome, and perhaps even single.

It is beyond the scope of this paper to develop a semantic/pragmatic theory of *perhaps even* ..., but what we see is that it is possible to use *perhaps even* ... not in order to cancel an implicature but in order to add some relevant information, and what comes after *perhaps even* does not need to entail what comes before it. Then, even if one is aware that $4R^{cc}7 \not\models 3R^{cc}7$, one might still use (8b) to add some relevant, logically independent information.

6.2 Guatemala sentences

Let us now discuss the Guatemala sentences. The interesting observation was that (9) is interpreted like (10). Krifka's explanation is that in the case of (9), the utterance context is naturally taken to be a description-describing context, where the speaker wants to convey information on the skewness of the distribution of people and land, and therefore, if n is a small number and m a large number, *in Guatemala, n% of the population owns m% of the land* entails, but is not entailed by, *in Guatemala, n'% of the population owns m'% of the land* where $n' \geq n$, $m' \leq m$ and either $n' \neq n$ or $m' \neq m$.

This explanation may sound plausible at first glance, but turns out not quite right. Suppose 100% of the Guatemalan population owns some Guatemalan land and 100% of the Guatemalan land is owned by Guatemalan citizens, and consider (30a–c), which have been derived from (11a–c) by switching the numbers between the subject and the object.

(30) a. In Guatemala, 70% of the population owns 2% of the land.
b. In Guatemala, 70% of the population owns 3% of the land.
c. In Guatemala, 70% of the population owns 4% of the land.

Since we have (11a) $\models$ (11b) $\models$ (11c), if Krifka's explanation were on the right track, we should have the same entailment pattern, namely (30a) $\models$ (30b) $\models$ (30c). However, this is incorrect, and what we have is in fact (30c) $\models$ (30b) $\models$ (30a). Likewise, consider (31a–c), which have been derived from (12a–c) in the same manner.

(31) a. In Guatemala, 60% of the population owns 3% of the land.
b. In Guatemala, 70% of the population owns 3% of the land.
c. In Guatemala, 80% of the population owns 3% of the land.

Since we have (12c) $\models$ (12b) $\models$ (12a), Krifka would predict (31c) $\models$ (31b) $\models$ (31a). Again, this is wrong, and what we have is (31a) $\models$ (31b) $\models$ (31c). It then seems that the interesting interpretation of (9) has nothing to do with the skewness of the distribution, i.e., the fact that the subject numeral is small and the object numeral is large.

Could this phenomenon be caused by some asymmetry between the subject and the nonsubject? To answer this question, consider (32)–(35), which have been obtained by rewriting (11), (12), (30) and (31) respectively, using the verb *belong to* instead of *own* in order to exchange the roles of the subject and the nonsubject.

(32) a. In Guatemala, 70% of the land belongs to 2% of the population.
b. In Guatemala, 70% of the land belongs to 3% of the population.
c. In Guatemala, 70% of the land belongs to 4% of the population.

(33) a. In Guatemala, 60% of the land belongs to 3% of the population.
b. In Guatemala, 70% of the land belongs to 3% of the population.
c. In Guatemala, 80% of the land belongs to 3% of the population.

(34) a. In Guatemala, 2% of the land belongs to 70% of the population.
b. In Guatemala, 3% of the land belongs to 70% of the population.
c. In Guatemala, 4% of the land belongs to 70% of the population.

(35) a. In Guatemala, 3% of the land belongs to 60% of the population.
b. In Guatemala, 3% of the land belongs to 70% of the population.
c. In Guatemala, 3% of the land belongs to 80% of the population.

The valid entailments we find are (32a) $\models$ (32b) $\models$ (32c), (33c) $\models$ (33b) $\models$ (33a), (34c) $\models$ (34b) $\models$ (34a) and (35a) $\models$ (35b) $\models$ (35c). Thus, the entailment patterns are exactly the same as in the original (11), (12), (30) and (31). This shows that what we are looking at has nothing to do with some special property of the subject.

To summarize what we observe for the Guatemala sentences, then, the number associated with *population* is consistently given an 'at most' interpretation while the number associated with *land* is consistently given an 'at least' interpretation, and neither the combination of numbers nor the subject/nonsubject difference affects this.

I suggest that the source of the interesting interpretation of the Guatemala sentences has do with the types of the nouns involved. In the preceding sections, we have seen that cumulative sentences have different entailment relations, depending not only on whether the context is existential or distribution-describing, but also on whether the two nouns involved are count or mass. For cumulative sentences with two count nouns in distribution-describing contexts, Krifka's explanation is quite right, as shown by the following proposition.

Proposition 15. *For all $n, n', N, m, m', M \in \mathbb{N}$ such that $n \leq m$, $n' \leq m'$, $n \leq n' \leq N$, $m' \leq m \leq M$,*

$$nR^{cc}m \wedge NR^{cc}M \models n'R^{cc}m',$$
$$mR^{cc}n \wedge MR^{cc}N \models m'R^{cc}n'.$$

Proof. Let n, n', N, m, m', M satisfy the given conditions. Since $m' - n' \geq 0$ and $n' - n \geq 0$, we have $n \leq n + (m' - n') = m' - (n' - n) \leq m' \leq m$. It then follows from Proposition 1 that

$$nR^{cc}m \models nR^{cc}(m' - (n' - n)).$$

Also, by using Corollary 5 $n' - n$ times, we obtain $nR^{cc}(m' - (n' - n)) \wedge NR^{cc}M \models (n + (n' - n))R^{cc}(m' - (n' - n) + (n' - n))$, that is,

$$nR^{cc}(m' - (n' - n)) \wedge NR^{cc}M \models n'R^{cc}m'.$$

From the above two entailments, we obtain $nR^{cc}m \wedge NR^{cc}M \models n'R^{cc}m'$. One can prove $mR^{cc}n \wedge MR^{cc}N \models m'R^{cc}n'$ similarly. □

$NR^{cc}M$ in Proposition 15 is to be understood as background information in a distribution-describing context. For some concrete examples, imagine that the sentences in (36)–(39) are uttered in a context where it is given that one hundred boys kissed one hundred girls.

(36) a. Two boys kissed seventy girls.
b. Three boys kissed seventy girls.
c. Four boys kissed seventy girls.

(37) a. Three boys kissed sixty girls.
b. Three boys kissed seventy girls.
c. Three boys kissed eighty girls.

(38) a. Seventy boys kissed two girls.
b. Seventy boys kissed three girls.
c. Seventy boys kissed four girls.

(39) a. Sixty boys kissed three girls.
b. Seventy boys kissed three girls.
c. Eighty boys kissed three girls.

(36)–(39) have the same combinations of numbers as (11), (12), (30) and (31) respectively. By Proposition 15, we have (36a) $\models$ (36b) $\models$ (36c), (37c) $\models$ (37b) $\models$ (37a), (38a) $\models$ (38b) $\models$ (38c) and (39c) $\models$ (39b) $\models$ (39a). Thus, while (11) and (30) show opposite entailment patterns, the corresponding (36) and (38) share the same entailment pattern. Similarly, (37) and (39) have the same entailment pattern unlike (12) and (31).

Returning to the Guatemala sentences, we see that the two nouns involved are *(% of the) population* and *(% of the) land*, and these are mass nouns. We would then make the wrong prediction that no entailments hold between Guatemala sentences with different combinations of numbers, since there are in general no nontrivial entailments for cumulative sentences with two mass nouns, as we briefly saw in Section 5. This puzzle dissolves once we reflect on the meanings of *population* and *land*. The population of Guatemala consists of Guatemalan people, each one of whom is not divisible. Hence, each Guatemalan citizen becomes an atom in the denotation of *(% of the) population*. Thus, while *(% of the) land* may have a genuine mass denotation, *(% of the) population* practically has a count denotation. Consequently, the Guatemala sentences are interpreted like cumulative sentences with one count noun and one mass noun in distribution-describing contexts. Then, what we should consider is something like Diagram 4, and it will be immediate that all the entailments found in (11), (12), (30) and (31) are exactly as expected. For (32)–(35), where the positions of the two nouns have been switched, we need only picture the mirror image of the diagram. Again, the entailments found in (32)–(35) are exactly as expected.

7 Scalar implicatures of cumulative sentences

According to Krifka (1999) and Landman (2000), *n boys kissed m girls* generally implicates that only n boys kissed girls and only m girls were kissed by boys. We can now tell that this is indeed predicted by the classic Gricean theory when $n = m$. To see this, suppose there are numbers $n' \geq n$ and $m' \geq m$ such that either $n' \neq n$ or $m' \neq m$ and $n'R^{cc}m'$. Without loss of generality, assume $n' \leq m'$. It then follows from Proposition 1 that we have $n'R^{cc}m' \models n'R^{cc}n'$. Also, by using Corollary 3 $n' - n$ times, we obtain $n'R^{cc}n' \models nR^{cc}n$. Together, we obtain $n'R^{cc}m' \models nR^{cc}n$. Thus, *$n'$ boys kissed m' girls* is more informative than *n boys kissed n girls*. Hence, the use of the latter sentence should implicate the denial of the former for any n' and m' satisfying the given conditions, and this amounts to implicating that only n boys kissed girls and only n girls were kissed by boys.

When $n \neq m$, however, the classic Gricean theory does not predict *n boys kissed m girls* to implicate that only n boys kissed girls and only m girls were kissed by boys. The question is, then, do we really get such implicatures? Due to scalar implicature, numerically-quantified existential sentences are often understood to mean what is obtained by adding *exactly* to the original sentence. For instance, (40b) is the resulting interpretation for (40a) (= (3)), as we saw in Section 2.

(40) a. John kissed seven girls.
b. John kissed exactly seven girls.

Now, if one is to mechanically manipulate cumulative sentences in the same fashion, one might like to imagine (41b) to be the pragmatically-strengthened meaning of (41a) (= (1)).

(41) a. Three boys kissed seven girls.
b. Exactly three boys kissed exactly seven girls.

In essence, this is what many researchers on cumulative sentences, including Krifka and Landman, seem to assume, though the technical details of how to derive such a strengthened meaning may differ from theory to theory. Regarding the interpretation of the "exactly" sentence (41b), many (Krifka 1999, Landman 2000, Winter 2000, Ferreira 2007, Brasoveanu 2010) write to the effect that it has the truth conditions in (2) and therefore will be false in model $\mathfrak{M}_2$. I would like to cast doubt on this apparent consensus, however. While it is clear that the "exactly" sentence (40b) is false in a situation where John kissed eight girls, it is not clear to me whether (41b) is true or false in model $\mathfrak{M}_2$. Just what exactly is *exactly* doing in (41b)? In any event, regardless of whether (41a) is perceived to mean (41b) or not, it seems irrational to me to suppose that (2) gives the pragmatically-strengthened meaning of (41a). According to Krifka and Landman, in order to describe $\mathfrak{M}_2$, one would be forced to use the sentence *four boys kissed eight girls*, but why? That would make it impossible for one to communicate the fact that three boys kissed seven girls in $\mathfrak{M}_2$, because this is not entailed by the sentence *four boys kissed eight girls*. I would like to suggest that when people feel like *n boys kissed m girls* implicates that only n boys kissed girls and only m girls were kissed by boys, they might be operating under the mistaken impression that *n boys kissed m girls* is entailed by *n' boys kissed m' girls* whenever $n' \geq n$ and $m' \geq m$.

Let us then see what scalar implicatures are predicted when $n \neq m$. We have already seen an example in Section 6.1 with (8a). In much the same way, (41a) is predicted to implicate (42a) (due to Proposition 1) and (42b) (due to Proposition 2).

(42) a. It is not the case that three boys kissed eight girls.

b. It is not the case that four boys kissed nine girls.

As we have seen already, the fact that some people find a contrast between (8a) and (8b) suggests that they may well be aware of the entailment $3R^{cc}8 \models 3R^{cc}7$ and consequently obtain (42a) as an implicature for (41a), though the evidence is not conclusive. It is quite likely, however, that for many other people, (41a) does not implicate (42a), and (8a) is good not because of the existence of the entailment $3R^{cc}8 \models 3R^{cc}7$ but for some other reasons (see Section 6.1). The situation with (42b), on the other hand, seems rather different. I simply do not think (41a) has (42b) as an implicature. I suppose that the problem with these predictions lies in the fact that even though Propositions 1 and 2 are proved in an elementary manner, their truth is not immediate to the normal person who is casually holding a conversation, unless perhaps he or she is someone like John von Neumann. In fact, when I bombarded people with questions like "Suppose seven boys kissed three girls. Then, certainly, six boys kissed three girls. Right or wrong?" they seemed very perplexed, which indicates that Proposition 1 is already not so obvious, and I believe that the entailments given by Proposition 2 are much more surprising for most people. If one is not aware of the relevant entailments, one is of course not expected to get these implicatures.[19]

So far, we have only considered the classic Gricean view, where the utterance of a sentence is taken to implicate the negations of the logically stronger alternative sentences. More recently, Fox (2007a) has proposed that the utterance of a sentence implicates the negations of all the *non-weaker* alternatives that are *innocently excludable*. Given a sentence p to be asserted and a set A of alternative sentences, the set of non-weaker alternatives, $\mathrm{NW}(p,A)$, is defined as in (43a), and the concept of "innocently excludable" is defined as in (43b).[20]

(43) a. $\mathrm{NW}(p,A) = \{q \in A \mid p \not\models q\}$.

b. q is innocently excludable iff $\neg\exists q' \in \mathrm{NW}(p,A)[(p \wedge \neg q) \models q']$.

In our case, we have $p \equiv 3R^{cc}7$ and $A = \{nR^{cc}m \mid n,m \in \mathbb{N}\}$. Then, since $3R^{cc}7 \not\models 1R^{cc}4$, we have $1R^{cc}4 \in \mathrm{NW}(p,A)$. Furthermore, it is easy to see that $1R^{cc}4$ is innocently excludable. Fox's theory therefore predicts *three boys kissed seven girls* to implicate that it is not the case that one boy kissed four girls.[21] It is quite clear that we do not get such an implicature, so at least for cumulative sentences, Fox's theory does not seem to work.

8 Conclusion

This paper has investigated what entailments hold between cumulative sentences with two numerically-quantified NPs. It showed that the entailment relation varies depending on whether the utterance context is existential or distribution-describing, and also on whether the two nouns participating in the cumulation are count or mass.

It is worth noting that the fact that the count/mass distinction affects the entailment relation for existential sentences has been brought to light by virtue of looking at cumulative sentences. For comparison, consider existential sentences with one numerically-quantified NP. For n a natural number, let

$$P^{c}n \equiv \exists X[\#(X) = n \wedge N(X) \wedge V(X)].$$

[19] This may not be the whole story, since even though I am now aware of the entailment $4R^{cc}9 \models 3R^{cc}7$, I still do not take (41a) to implicate (42b).

[20] See Chemla (2009) for a different formulation.

[21] I would like to thank Raj Singh for reminding me of Fox's theory and showing what it can predict.

$P^{\mathrm{c}}n$ represents the truth conditions of an existential sentence with one count noun such as *John kissed n girls*. For s a positive real number, let

$$P^{\mathrm{m}}s \equiv \exists X[\mu(X) = s \wedge N(X) \wedge V(X)].$$

$P^{\mathrm{m}}s$ represents the truth conditions of an existential sentence with one mass noun such as *John drank s liters of beer*. Assuming cumulativity and divisivity for predicates, it is easy to see that $P^{\mathrm{c}}n \models P^{\mathrm{c}}m$ iff $n \geq m$. With Lemma 11, it is also easy to see that $P^{\mathrm{m}}s \models P^{\mathrm{m}}t$ iff $s \geq t$. Consequently, the entailment relation for existential sentences with one count noun is a subrelation of the entailment relation for existential sentences with one mass noun, i.e.,

$$\{(n,m) \mid P^{\mathrm{c}}n \models P^{\mathrm{c}}m\} \subset \{(s,t) \mid P^{\mathrm{m}}s \models P^{\mathrm{m}}t\}.$$

By contrast, the entailment relation for existential cumulative sentences with two count nouns is not a subrelation of the entailment relations for existential cumulative sentences involving mass nouns. First, we have

$$\{((n,m),(n',m')) \mid nR^{\mathrm{cc}}m \models n'R^{\mathrm{cc}}m'\} \not\subset \{((n,s),(n',s')) \mid nR^{\mathrm{cm}}s \models n'R^{\mathrm{cm}}s'\},$$

because we have, for instance, $3R^{\mathrm{cc}}5 \models 2R^{\mathrm{cc}}4$ and $3R^{\mathrm{cm}}5 \not\models 2R^{\mathrm{cm}}4$. To see why $3R^{\mathrm{cm}}5 \not\models 2R^{\mathrm{cm}}4$, imagine a situation where there are three boys each of whom drank exactly $5/3$ liters of beer. Then, no matter which two boys we pick, it is not the case that the two boys drank 4 liter of beer. Also, we have

$$\{((n,m),(n',m')) \mid nR^{\mathrm{cc}}m \models n'R^{\mathrm{cc}}m'\} \not\subset \{((s,t),(s',t')) \mid sR^{\mathrm{mm}}t \models s'R^{\mathrm{mm}}t'\},$$

since $\{((s,t),(s',t')) \mid sR^{\mathrm{mm}}t \models s'R^{\mathrm{mm}}t'\} = \{((s,t),(s',t')) \mid s = s' \wedge t = t'\}$.

The reader might recall Fox and Hackl's (2006) following hypothesis:

(44) The Universal Density of Measurements
Measurement scales needed for natural language semantics are always dense.

It is reasonable to assume density of measurement for mass entities such as weights, so if John weighs 120 lbs, it is also true that John weighs $119 + \varepsilon$ lbs for some ε such that $0 < \varepsilon < 1$. Fox and Hackl's idea is to extend this to count nouns as well, so if John has 3 children, it is also true that John has $2 + \varepsilon$ children for some ε such that $0 < \varepsilon < 1$. Fox and Hackl themselves admit that this is quite an unintuitive move, but it allows them to account for the fact that *John has more than three children* does not implicate that John has exactly four children.[22] In order to fill the gap between theory and intuition, Fox and Hackl devise a complex system of division of labor

[22]Fox and Hackl's theory cannot account for the fact that (i) does not implicate that John has exactly three children.

(i) John has three or more children.

Fox (2007b) attributes the lack of scalar implicature in (i) to the presence of *or*, but such an account seems dubious. Japanese has the bound suffix *ijo*, whose meaning is 'more than or equal to'. Thus, (ii) (where TOP = topic, NOM = nominative, CL = classifier) means 'John has three or more children'. Crucially, even though (ii) does not contain a word for *or*, it clearly does not implicate that John has exactly three children.

(ii) John-wa kodomo-ga san-nin-ijo iru.
John-TOP children-NOM three-CL-more.than.or.equal.to. exist

and interaction between syntax, semantics and pragmatics. In light of the discussion in the present paper, it is now fairly obvious that Fox and Hackl's approach is hard to maintain, as it exploits the fact that the entailment relation for existential sentences with one count noun can be embedded in the entailment relation for existential sentences with one mass noun via the identity map, which is actually an exceptional situation. As made clear in the previous paragraph, the entailment relations for existential sentences with count nouns are generally not embeddable via the identity map in the entailment relations for corresponding existential sentences involving mass nouns.

Regarding scalar implicatures of cumulative sentences, as things stand now, it is not clear whether such things really exist, and if so, what exactly they are. It is therefore desired that empirical data be collected on how people interpret cumulative sentences. Whatever the empirical facts are, by investigating the entailment relations between cumulative sentences, I hope to have laid the groundwork for discussing the pragmatics of cumulative sentences. I also hope to have made the point that the story about scalar implicatures for cumulative sentences cannot be as simple as in the case of sentences with only one quantified NP. Specifically, I have shown that the standard theories of scalar implicature make interesting, but apparently controversial predictions. This seems to reveal a conceptually problematic aspect of these theories, namely, that they seem to take it for granted that humans can immediately see the entailment relation between alternative sentences so that they can derive the pragmatic interpretation of a sentence. Having examined the entailments between cumulative sentences, some of which struck me as quite surprising, I cannot help but find it somewhat baffling that those theories seem to make that assumption, when the average Joe can hardly solve the logic puzzles in Raymond Smullyan's puzzle books.[23]

In this paper, we have only looked at cumulative sentences involving only two nouns, but it should also be interesting to investigate what entailments and scalar implicatures arise for cumulative sentences with three or more nouns participating in the cumulation. In Kanazawa and Shimada (2014), we report some partial results on the entailment relation between cumulative sentences with three count nouns.

References

Barwise, Jon, and Robin Cooper. 1981. Generalized quantifiers and natural language. *Linguistics and Philosophy* 4:159–219.

Brasoveanu, Adrian. 2010. Modified numerals as post-suppositions. In *Logic, language and meaning: 17th Amsterdam colloquium, Amsterdam, the Netherlands, December 2009, revised selected papers*, ed. Maria Aloni, Harald Bastiaanse, Tikitu de Jager, and Katrin Schulz, 203–212. Berlin, Germany: Springer.

Chemla, Emmanuel. 2009. Universal implicatures and free choice effects: experimental data. *Semantics and Pragmatics* 2:1–33.

[23]The reviewer suggests that what matters is that the implicature system be able to compute logical relationships between sentences, and the human need not be aware of these relationships. In response to this, I would like to note that first-order logic is undecidable and natural language can obviously express every proposition that first-order logic can, which implies that there is no *general* computational procedure for deciding the set of stronger or non-weaker alternatives to a given sentence. This is of course far from sufficient to justify my viewpoint, but some "semantic" consideration by the human seems to be required in order to derive a pragmatically-strengthened meaning of a sentence, and I take that to be fairly natural given that pragmatics is about how humans use sentences.

Ferreira, Marcelo. 2007. Scope splitting and cumulativity. In *Proceedings of the ESSLLI workshop on quantifier modification*, ed. Rick Nouwen and Jakub Dotlačil.

Fox, Danny. 2007a. Free choice and the theory of scalar implicatures. In *Presupposition and implicature in compositional semantics*, ed. Uli Sauerland and Penka Stateva, 71–120. New York: Palgrave Macmillan.

Fox, Danny. 2007b. Too many alternatives: Density, symmetry and other predicaments. In *Proceedings of SALT 17*, ed. T. Friedman and M. Gibson, 89–111. Ithaca, NY: Cornell University.

Fox, Danny, and Martin Hackl. 2006. The universal density of measurement. *Linguistics and Philosophy* 29:537–586.

Givant, Steven, and Paul Halmos. 2009. *Introduction to Boolean algebras*. New York: Springer.

Grice, Paul. 1975. Logic and conversation. In *Syntax and semantics, vol. 3: Speech acts*, ed. P. Cole and J. Morgan, 41–58. New York: Academic Press.

Horn, Laurence R. 1972. On the semantic properties of logical operators in English. Doctoral Dissertation, UCLA.

Kanazawa, Makoto, and Junri Shimada. 2014. Toward a logic of cumulative quantification. In *Joint proceedings of the second workshop on Natural Language and Computer Science (NLCS'14) & 1st international workshop on Natural Language Services for Reasoners (NLSR 2014)*, ed. Valeria de Paiva, Walther Neuper, Pedro Quaresma, Christian Retoré, Lawrence S. Moss, and Jordi Saludes, 111–124. Centre for Informatics and Systems, University of Coimbra.

Krifka, Manfred. 1989. Nominal reference, temporal constitution and quantification in event semantics. In *Semantics and contextual expressions*, ed. R. Bartsch, J. van Benthem, and P. van Emde Boas, 75–115. Dordrecht: Foris.

Krifka, Manfred. 1990. Four thousand ships passed through the lock: Object-induced measure functions on events. *Linguistics and Philosophy* 13:487–520.

Krifka, Manfred. 1992. Definite NPs aren't quantifiers. *Linguistic Inquiry* 23:156–163.

Krifka, Manfred. 1999. At least some determiners aren't determiners. In *The semantics/pragmatics interface from different points of view*, ed. K. Turner, 257–291. Oxford: Elsevier.

Landman, Fred. 1991. *Structures for semantics*. Dordrecht: Kluwer.

Landman, Fred. 2000. *Events and plurality: The Jerusalem lectures*. Dordrecht: Kluwer.

Levinson, Stephen. 1983. *Pragmatics*. Cambridge University Press.

Link, Gödehard. 1983. The logical analysis of plurals and mass terms: A lattice-theoretical approach. In *Meaning, use and the interpretation of language*, ed. R. Bäuerle, C. Schwarze, and A. von Stechow, 303–323. Berlin, New York: de Gruyter.

Robaldo, Livio. 2011. Distributivity, collectivity, and cumulativity in terms of (in)dependence and maximality. *Journal of Logic, Language and Information* 20:233–271.

Scha, Remko J. H. 1981. Distributive, collective, and cumulative quantification. In *Formal methods in the study of language, part 2*, ed. J. A. G. Groenendijk, T. M. V. Janssen, and M. B. J. Stokhof, 483–512. Amsterdam: Mathematisch Centrum.

Sierpiński, Wacław. 1922. Sur les fonctions d'ensemble additives et continues. *Fundamenta Mathematicae* 3:240–246.

Winter, Yoad. 2000. Distributivity and dependency. *Natural Language Semantics* 8:27–69.