Probabilistic Mereological TTR and the Mass/Count Distinction

Peter Sutton and Hana Filip

Heinrich-Heine-Universität Düsseldorf, 40225 Düsseldorf, Germany,
peter.r.sutton.@icloud.com; hana.filip@gmail.com

Abstract. We propose that countability of nouns is not a matter of just a bipartite mass vs. count distinction. Instead we distinguish four classes of nouns. We propose that these divisions can be modelled in a mereological enrichment of prob-TTR (Cooper et al. 2014), via patterns that emerge in probabilistic type judgements associated with them.

Keywords: lattice theoretic semantics, Type Theory with Records, Bayesian Semantics

1 Introduction

The goals of this paper are threefold. (1) We argue that four semantic classes can be used to predict cross and intralinguistic variation in whether nouns are encoded as count or mass. These semantic classes are defined in terms of two kinds of sources: vagueness and overlap. (2) We enrich prob-TTR (Cooper et al. 2014, a probabilistic variant of TTR, Cooper 2012) with mereological relations to give probM-TTR. We adopt prob-TTR since its probabilistic basis is perfectly suited to the representation of vagueness. We enrich it so as to be able to express when two (mereological) entities overlap. (3) We derive our four classes by showing how vagueness and overlap interact.

2 Sources of Countability

We propose that countability in nouns can be accounted for by combining three factors. Two relate to vagueness, and one to overlap.

The first factor is individuation: what counts as a single quantity (an individual) of a given noun. Individuation can be vague. For some nouns (e.g. *mud*, *blood*) it is vague what a single quantity or individual portion is. There is no generally accepted way to partition the domain of such nouns into individuals. For other nouns (e.g. *cat*, *boy*, *lentil*, *rice*, *kitchenware*, *furniture*) there is little or no vagueness in this respect. There are generally accepted ways to partition these domains (e.g. into single cats, single lentils, single grains of rice). Vagueness of individuation can affect countability. If you can find no one clear way to partition the domain into individuals, there is no clear answer to how many individuals there are. Individuation features in many accounts of the mass/count distinction ((Grimm 2012), (Rothstein 2010), (Landman 2011) among many others). Vagueness of individuation has been suggested as a factor by Rothstein (2010).

The second factor is quantity: how much of a given amount of stuff is sufficient to classify it as falling under a noun. This can be vague too. For some nouns (*mud*, *blood*, *rice*, *lentil*) a single speck of mud or a single grain of rice is not always sufficient to count as mud or rice (you can truly say "we don't have any rice (for dinner)" when there are only a few grains left in the packet). However, it is vague just how much is enough in a given context. For other nouns (*cat*, *boy*, *kitchenware*, *furniture*), there is little or no vagueness in this respect. Vagueness in quantity can affect countability. If there is no clear minimal quantity of N that counts as N (in a context), then there are no clear minimal parts of N to count.

These two vagueness factors are highly interrelated, but they are separable. Nouns such as *lentil* and *rice* are not vague in the first respect, but are vague in the second. This type of quantity vagueness is suggested as the source of the mass/count distinction in Chierchia (2010).

The third factor is overlap. For some nouns (*furniture*, *kitchenware*), the things classified as individuals/single quantities may overlap. For example, a pestle and mortar can count as a single item of kitchenware, but individually, a pestle and a mortar can each count as single items of kitchenware. For other nouns (*cat*, *boy*, *lentil*, *rice*) there is no overlap in this respect. Overlap can affect countability, because it leads to double counting. Is a pestle and mortar one, two or even three items of kitchenware? The overlap of individual entities (things that count as one) is proposed as the source of the mass/count distinction in Landman (2011).

The first vagueness factor and the overlap factor can interact. If it is vague how to partition a domain into individuals, then it will not be possible to determine what the disjoint non-overlapping individuals are. In some sense, therefore, individuation-vague nouns (*mud*, *blood*) are also overlapping. Hence, nouns can be overlapping as well as vague with respect to both individuation and quantity $[+V, +O]$ (*mud*, *blood*), quantity vague but non-overlapping $[+V, -O]$ (*rice*, *lentil*), overlapping but non-vague $[-V, +O]$ (*kitchenware*, *furniture*), and neither vague nor overlapping $[-V, -O]$ (*cat*, *boy*). As Table 1 helps to show, these four classes display striking cross- and intralinguistic patterns when it comes to the encoding of their members as count $[+C]$ and mass $[-C]$.

	$[+O]$	$[-O]$
$[+V]$	mud$_{[-C]}$, blood$_{[-C]}$	rice$_{[-C]}$, lentils$_{[+C]}$= lešta$_{[-C],Bulgarian}$
$[-V]$	furniture$_{[-C]}$= huonekalu$_{[+C],Finnish}$	cat$_{[+C]}$, boy$_{[+C]}$

Table 1. Feature Matrix for $[\pm V], [\pm O], [\pm C]$

Nouns in the upper left quadrant of Table 1 are typically *substances* and are encoded universally as mass, while nouns in the lower right quadrant are prototypical count nouns, and both types tend not to display cross- or intralinguistic variation in the encoding as count or mass. Nouns in the upper right quadrant are typically *granular*. Such nouns will display count/mass variation e.g. $rice_{[-C]}$, $lentil_{[+C]}$, $lešta_{[-C],\text{lentil (Bulgarian)}}$. Nouns in the lower left quadrant are typically aggregates in the sense of Payne and Huddleston (2002) and are often superordinate categories. Such nouns display cross- or intralinguistic count/mass variation e.g. $furniture_{[-C]}$, $meubel_{[+C],\text{furniture (Dutch)}}$, $jalkineet_{[+C],\text{PL,footwear (Finnish)}}$.

This leads us to formulating the following Hypothesis: The presence of individuation vagueness entails uncountability ($[-C]$). The absence of vagueness and overlap entails countability ($[+C]$). The presence of quantity vagueness, without overlap allows grammaticization of a given noun concept as $[+C]$ or $[-C]$. The presence of overlap without quantity vagueness allows grammaticization of a given noun concept as $[+C]$ or $[-C]$.

As we shall go on to argue, this variability may ultimately be viewed as a reflection of the differences in perspective on what appears to be the same slice of the real world. Most importantly, no single source account, such as the vagueness-only account of Chierchia (2010) or the overlap-only account of Landman (2011), will be sufficient to account for all the count/mass data as represented in Table 1.

3 Probabilistic Mereological TTR

3.1 Individuation and Individuation Vagueness

The basic schema for a record type involving a nominal predicate will be the record type in (1) where P would be replaced with *cat*, *rice*, *mud* etc. The type $^*Ind$ is the type of *stuff* (a type involving substances, individuals and any sums thereof):

$$\begin{bmatrix} x & : {}^*Ind \\ c_p & : P(x) \end{bmatrix} \tag{1}$$

Following Krifka (1995), we view the application of nominal predicates as being affected by both the *qualitative* and *quantitative* factors. In addition to whatever observable and intrinsic properties that might usually be associated with the qualitative aspects of a noun's denotation, we assume properties such as boundedness, contiguity, size, and topology. For every type as in (1) we assume a dependent type including $P_{qual}(x)$ which packages together all of these properties such as in (2):

$$\begin{bmatrix} x & : {}^*Ind \\ c_p & : P_{qual}(x) \end{bmatrix} \tag{2}$$

The quantitative aspect of applying nominal predicates will be characterised as a function $f_{pqual} : (RecType \rightarrow \mathbb{N})$. This is a quantising function which maps Record types, such as (2), to a natural number, where the natural number represents a quantity of P. So, it takes the type of things with rice qualities, boundedness, contiguity, size, and topology, for example, and maps this to a number. (Where 1 would indicate a single rice grain, and, 2 a sum of two rice grains etc.)[1] Finally, we also introduce a condition that the output of the function is some particular value. All of these components together yield a schema for a record type in (3), with an example in (4).

$$\begin{bmatrix} c_{ppt} : \begin{bmatrix} x & : {}^{*}Ind \\ c_{p_{ql}} & : P_{Qual}(x) \end{bmatrix} \\ f_{p_{qt}} :(\begin{bmatrix} x & : {}^{*}Ind \\ c_{p_{ql}} & : P_{Qual}(x) \end{bmatrix} \rightarrow \mathbb{N}) \\ i \quad : \mathbb{N} \\ c_{p_{qt}} : f_{p_{qt}}(c_{ppt}) = i \end{bmatrix} \quad (3) \qquad \begin{bmatrix} c_{ppt} \quad : \begin{bmatrix} x & : {}^{*}Ind \\ c_{rice_{ql}} & : rice_{Qual}(x) \end{bmatrix} \\ f_{rice_{qt}} :(\begin{bmatrix} x & : {}^{*}Ind \\ c_{p_{ql}} & : rice_{Qual}(x) \end{bmatrix} \rightarrow \mathbb{N}) \\ c_{rice_{qt}} : f_{rice_{qt}}(c_{ppt}) = 1 \end{bmatrix} \quad (4)$$

The type in (4), with quantity value of 1, represents the type of single grains of rice. For this and other such special cases with quantity values of 1, we introduce a notational convention. Let Ind_{cat}, Ind_{rice} etc. be the type of single cats, (grains) of rice etc. We then abbreviate (4) as (5):

$$[\, x : Ind_{rice} \,] \tag{5}$$

In prob-TTR, judgements are of the form $p(a : T) = k$ where $k \in [0, 1]$. Within Cooper et el's learning model, $p_{A,\mathfrak{J}}(r : T)$ is the probability that an agent, A, assigns to r being of type T with respect to $\mathfrak{J}$ (her experience and learning data set).

Individuation vagueness (which affects substance nouns like *mud*) can be represented as the low probability of anything in an agent's experience being categorised as a single quantity. For example, with *mud*, nothing one has experienced counts as a single mud quantity. This is because there is no principled way to divide the stuff with the right mud qualities into parts. Assuming a *threshold probability* θ, below which an agent will not make a type judgement, this means that for all $r : Rec$ in $A, \mathfrak{J}$:

$$p(r : [\, x : Ind_{mud} \,]) < \theta \tag{6}$$

Nothing the agent has experienced is judged to be a single quantity of mud because for every mud experience, any part of that mud is as good a candidate to be a single unit of mud as the whole is. Intuitively, this is precisely because mud lacks the relevant boundedness, contiguity, size, and topology properties to allow individuation.

3.2 Quantity Vagueness

Even for those nouns that are *not* individuation vague, another form of vagueness (first described by Chierchia 2010) can still remain. In any given situation, if some small amount of P is not enough to count as P, but some larger amount is, it can be vague what the least amount to count as P is. For example, one lentil or rice grain left on a plate does not mean that you have not eaten (all)

[1] The quantising function may be more course grained for higher values, just as perception is. For example, for *rice* high quantity values of could represent some range of numbers of grains.

your rice/lentils. Other nouns are not, or at least, are not usually vague in this way. A single chair is usually enough to count as furniture. A single cat is usually enough to count as a cat. For individuation vague nouns (which have no clear single units), quantity vagueness can also be present: A tiny fleck of mud left on your boot does not mean that you have not cleaned (all) the mud off your boots, but it can be vague how much mud would.

This form of vagueness can be represented as gradience in a conditional probability distribution. For example, $p(x$ is rice $\mid x$ is some quantity i of rice) will be low for small quantities, higher for higher quantities. We get a sorites-like slope, where, even if there is a threshold probability for assertion, an agent may not be sure whether or not she is at the threshold. In probM-TTR, this will be represented as the probability distribution generated by a record being of a type as in (1), given that it is of a type as in (3) (with i as a variable). For *rice*, this would be represented in (7)

$$p_{A,\mathfrak{J}}(r : \begin{bmatrix} x & : {}^{*}Ind \\ c_{rice} & : rice(x) \end{bmatrix} \mid r : \begin{bmatrix} c_{ppt} & : \begin{bmatrix} x & : {}^{*}Ind \\ c_{rice_{ql}} & : rice_{Qual}(x) \end{bmatrix} \\ f_{rice_{quant}} & :(\begin{bmatrix} x & : {}^{*}Ind \\ c_{p_{ql}} & : rice_{Qual}(x) \end{bmatrix} \rightarrow \mathbb{N}) \\ i & : \mathbb{N} \\ c_{rice_{qt}} & : f_{rice_{quant}}(c_{ppt}) = i \end{bmatrix}) = k \quad (7)$$

So, the probability k that something is *rice*, given that it is i quantised portions of rice (roughly, i number of grains, although see footnote 1), varies with the value of i. There will be little or no such vagueness for nouns such as *cat*, *furniture*, etc. This is represented as an above the threshold probability of something being a cat, given it is a single cat (8), and an above the threshold probability of something being furniture, given that it is a single (item of) furniture (9).

$$p_{A,\mathfrak{J}}(r : \begin{bmatrix} x & : {}^{*}Ind \\ c_{cat} & : cat(x) \end{bmatrix} \mid r : \begin{bmatrix} x : Ind_{cat} \end{bmatrix}) > \theta \quad (8)$$

$$p_{A,\mathfrak{J}}(r : \begin{bmatrix} x & : {}^{*}Ind \\ c_{furn.} & : furniture(x) \end{bmatrix} \mid r : \begin{bmatrix} x : Ind_{furn.} \end{bmatrix}) > \theta \quad (9)$$

The difference between nouns like *cat*, and nouns like *furniture* will be that single cats (entities of type $[x : Ind_{cat}]$) do not overlap, but single items of furniture (entities of type $[x : Ind_{furn}]$) may well do.

3.3 Overlap

We will capture overlap and non-overlap in terms of *Disjointedness*: There is overlap within a type if the type is not disjoint. A type T is disjoint iff for all $a : T$ and $b : T$, if $a \neq b$, then $a \cap b = \varnothing$. The type of single items of furniture ($[x : Ind_{furn}]$) is not disjoint, since all of the following are of this type: $dressing\text{-}table$, $mirror$, $dressing\text{-}table \cup mirror$.

Interpreted probabilistically, we will again use the probability threshold for type judgements, θ:

Definition 1. T IS PROBABILISTICALLY DISJOINT RELATIVE TO θ

$T : Disj_\theta$ iff there is at least some a such that $p(a : T) \geq \theta$, and for all a, b such that $p(a : T) \geq \theta$ and $p(b : T) \geq \theta$, if $a \neq b$ then $a \cap b = \varnothing$

So T is disjoint relative to threshold θ iff T is not empty from the perspective of sufficient certainty of what is of type T, and no two things which are sufficiently probably of type T overlap.

4 Countability in probM-TTR

Within the formal representations of nouns we have developed here, the denotations of nouns admit of two different *perspectives*: The "zoomed out" perspective of what counts, with sufficient certainty, as the smallest amount of stuff which is of some type (as determined by conditional probability distributions like in (7) for *rice*); and the "zoomed in" perspective of what counts as single, *but also disjoint* entities of a noun (disjoint entities of the type Ind_p). From each perspective, countability may be prevented by either vagueness, which prevents the identification of what one should count, or overlap, which leads to double counting and so incompatible counting results. We will now derive our four classes of nouns via whether countability is prevented from one of these two perspectives, from both, or from neither.

Prototypical count nouns: For nouns such as *cat*, the zoomed in and the zoomed out perspectives coincide. Zooming out, the smallest quantities that clearly satisfy the predicate *cat* are single individual cats, as the probability of something being a cat, given that it is of type Ind_{cat}, is very high/above the threshold. Zooming in, we see that this type Ind_{cat} is also disjoint. Either way, we are left with just individual cats, and so only one possible counting result. The fact that we get the same result from either perspective can explain why non-vague, non-overlapping nouns are encoded universally as count.

Vague, non-overlapping nouns: For nouns such as *rice* and *lentils*, the zoomed in and the zoomed out perspectives do not coincide. Zooming out, it is (quantity) vague what the smallest entities are that satisfy the predicates *rice* and *lentils*. Single grains/lentils frequently do not count as satisfying these predicates, as the probability of something being rice/lentils, given that it is of type Ind_{rice}/Ind_{lentil} is very low/below the threshold. Large quantities do count frequently enough, but the cut-off point is vague. From this zoomed out perspective we cannot focus on a reasonable counting base and so countability is blocked. Zooming in, the types Ind_{rice} and Ind_{lentil} are anyway disjoint (single grains/single lentils do not overlap). Furthermore, Ind_{rice} and Ind_{lentil} are not (particularly) vague. So counting can occur at this level.

The difference in the result between perspectives can explain why we see mass/count variation in such nouns. Depending on whether their denotations are viewed from a zoomed in or a zoomed out perspective, we can either count or fail to be able to count what is there. On this understanding of the mass/count distinction, English, for example, conceptualises *rice* from a zoomed out perspective, and *lentils* from a zoomed in perspective, hence the difference in count/mass encoding. Bulgarian conceptualises *lešta* (lentil) from a zoomed out perspective, hence *lešta* is encoded as mass.

Non-vague, overlapping nouns: For nouns such as *furniture* and *kitchenware*, the zoomed-in and the zoomed-out perspectives also do not coincide, but not for the same reasons as with nouns such as *rice* and *lentils*. Zooming out, there is no quantity vagueness for furniture or kitchenware. The smallest quantities that clearly satisfy the predicates $furniture$, $kitchenware$ are items of kitchenware and items of furniture respectively (the probability of something being a kitchenware/furniture, given that it is of type $Ind_{furn.}/Ind_{kitch.}$ is very high/above the threshold). However, we cannot count entities of the types $Ind_{furn.}$ and $Ind_{kitch.}$, because they are not disjoint. This overlap leads to incompatible counting results. For example $\{pestle, mortar, pestle \cap mortar\}$ are all of type $Ind_{kitch.}$, but it is not, from this perspective decided whether this set constitutes one, two or even three items of kitchenware. Zooming in, we are forced, relative to a context, to try to find a disjoint subset of the set $Ind_{kitch.}$ and thereby ignore/remove the overlap. The type Ind_{kitch} is not vague and so it is clear what is of this type. Taking the above example, this allows two possible contexts. In context 1, we take the disjoint subset that leaves $\{pestle, mortar\}$ in which case there are two things. In context 2, we take the disjoint subset that leaves $\{pestle \cup mortar\}$ in which case there is only one thing.

The different outcome of the two perspectives can explain why we see mass/count variation in such nouns. In English, for example, both *kitchenware* and *furniture* are viewed from a zoomed out perspective and so are encoded as mass. In Finnish $huonekalu_{[+C],\mathrm{furniture}}$ is viewed from a zoomed in perspective and is encoded as count.

Vague and overlapping nouns: For nouns such as *mud* and *blood*, the zoomed in and the zoomed out perspectives *do* coincide. Zooming out, it is (quantity) vague what the smallest entities are that satisfy the predicates *mud* and *blood*. Tiny amounts of mud and blood most of the time do not count as satisfying these predicates, but in contrast to nouns such as *rice* and *lentils*, there is also individuation vagueness here. Within the denotations of *mud* and *blood*, we are not even sure what is of the types Ind_{mud} or Ind_{blood}. From this zoomed out perspective we cannot focus on a reasonable counting base and so countability is blocked. Zooming in does not get us much further. Because of individuation vagueness, we are not sufficiently sure that anything is of type Ind_{mud} or Ind_{blood} etc. That means that types such as Ind_{mud} and Ind_{blood} are not probabilistically disjoint because nothing has a high enough probability of being of these types. That means that we can not discern a disjoint subset within this set. Countability is prevented from a zoomed-in perspective too. From either perspective, there is no counting result for individuation vague nouns. This can explain why substance nouns always get encoded as mass.

5 Conclusion

Depending upon the ways we can conceptualise some referent of a noun, counting either gets a free pass, faces a conceptually avoidable stumbling block (vagueness or overlap), or is stopped in its tracks (vagueness and overlap). We have proposed that in the first case, nouns will be universally encoded as count. In the second, depending on how the stumbling block was treated within and between languages, we expect mass/count encoding to vary. In the third case, we expect universal mass encoding. If correct, we improve upon Landman (2011) and Chierchia (2010) by accounting for cross- and intralinguistic variation, while also motivating the stubborn resistance that prototypical mass nouns show to counting.

References

Chierchia, G.: Mass nouns, vagueness and semantic variation. Synthese 174, 99–149 (2010)

Cooper, R.: Type theory and semantics in flux. In: Kempson, R., Fernando, T., Asher, N. (Eds.), Philosophy of Linguistics, Handbook of the Philosophy of Science. Elsevier, 271–323 (2012)

Cooper, R., Dobnik, S., Lappin, S., Larsson, S.: A probabilistic rich type theory for semantic interpretation. Proceedings of the EACL 2014 Workshop on Type Theory and Natural Language Semantics (2014)

Grimm, S.: Number and Individuation. PhD Dissertation, Stanford University (2012)

Krifka, M.: Common nouns: A contrastive analysis of English and Chinese. In: Carlson, G., Pelletier, F. J. (Eds.), The Generic Book. Chicago UP, 398–411 (1995)

Landman, F.: Count nouns – mass nouns – neat nouns – mess nouns. In: The Baltic International Yearbook of Cognition: Vol. 6. pp. 1–67 (2011)

Payne, J., and Huddleston, R.: Nouns and Noun Phrases: In: Pullum, G. and Huddleston, R. (Eds.) The Cambridge grammar of the English language. Cambridge UP, 323–524 (2002)

Rothstein, S.: Counting and the mass/count distinction. Journal of Semantics 27 (3), 343–397 (2010)