Generic Sentences as Quantifications over Samples

Manfred Krifka
Leibniz-Zentrum Allgemeine Sprachwissenschaft (ZAS)
Institut für deutsche Sprache und Linguistik & Berlin School of Mind & Brain
Humboldt-Universität zu Berlin

Workshop "The Generic Notebook"
Humboldt-Universität zu Berlin
June 2, 2017

1. Types of Genericiy

(1) As a reminder, there are two types of Genericity, according to The Generic Book:

a. Statements about kinds:	*The rat reached Hawaii in 300 A.D.*	not treated
b. Characterizing sentences:	(*Hungry) rats are clever.*	treated
c. Both:	*The rat is a clever animal.* *Rats are clever animals.*	(treated)

2. The quantificational approach

(2) Krifka e.a. 1995: Characterizing sentences involve a generic quantifier, a dyadic operator with a restrictor and a nuclear scope that allows for exceptions.

(3) Dyadic structure (Carlson 1988), modulated by focus (Krifka 1995):
a. *Typhoons are dángerous.*
GEN(λx[typhoon(x)], λx[dangerous(x)])
b. *Typhoons arise in thís part of the Pacific.*
GEN(λx[typhoon(x)], λx∃s[s is a situation in this part of the Pacific, x arises in s])
c. *Typhóóns arise in this part of the Pacific.*
GEN(λs[s is a situation in this part of the pacific], λs∃x[typhoon(x) ∧ x arises in x])

(4) Exceptional quantification:

a. *Turtles live a long live*.	(even though >95% die young)
b. *#Children born in Rainbow lake are right-handed.*	(even though 100% are)
c. *Boys don't cry*.	(even though ≈ 0% do not cry).
d. *Mail from Antarctica goes in box Z.*	(even though there is none)

(5) What is exceptional quantification?

a. Quantification over relevant entities.	Why not then: *Turtles die young.*
b. Quantification over prototypical entities.	i. *Ducks lay speckled eggs.* ii. *Ducks have colorful feathers.*
c. Quantification over stereotypical entities.	i. *Lions have a mane.* ii. *Snakes are slimy.*

(6) Modal quantification over a set of accessible worlds that behave according to the rules.
a. GEN(A)(B): GEN i∈R (A(i), B(i)), or ∀i∈R GEN(A(i), B(i))
a. *#Children born in Rainbow lake are right-handed.* – could easily be different.
b. Explains *Boys don't cry.* – this is how boys should behave, deontic statement.
c. Explains *Mail from Antarctica goes in box Z.* – for those indices where there is such mail.
d. But: *Ducks lay speckled eggs.* – only the females do even in all normal worlds.

(7) Definitional generics as involving fixing of the language itself (Krifka 2013):
a. *A straight line is the shortest connection between two points.*
b. *A boy / Boys do not cry.*

3. The probability approach of Cohen 1999

(8) Cohen (1999, ff.): Characterizing / frequency statements (*usually, often...*) express probability judgements, they are interpreted as statements about relative hypothetical frequency.

(9) With P(B|A) as the probability of A, given B:
a. always(B, A) iff P(B|A) = 1
b. usually(B, A) iff P(B|A) > 0,5
c. sometimes(B, A) iff P(B|A) > 0
d. never(B,A) iff P(B|A) = 0

(10) *Birds (usually) can fly.* P(λx[x is a bird], λx[x can fly]) > 0,5

(11) But: #*Children born in Rainbow Lake are (usually) right-handed.* (even if all of them are).

(12) Probability as limit frequency, relating to ideas of Rudolf von Mises (1919, 1931):
– form a sequence S = x_0 x_1 x_2 … x_n of elements of A, compute the number of x_i with B(x_i) / n
– let S grow to infinity by choosing more x_i at random,
– the limit of this process is P(B|A)

(13) Rule (12) works for finite and infinite cases,
but is **not necessary** for finite cases, where we can just compute: P(B|A) = #B / #A;
hence the evaluation procedure in (9)/(12) is blocked if A is finite.

(14) We can restrict it for infinite cases by working with sets S, not sequences (no repetition), and require that the size of S can grow indefinitely:
Take samples S ⊆ A and compute r = #(B ∩ S) / #(S),
r approaches P(B|A) as S grows by adding elements of A at random.

(15) The part “at random” is crucial, and ways to determine it are controversial.
a. Otherwise *Natural numbers are usually / always divisible by seven* comes out as true, if we happen to select the infinite sequence 7, 14, 21, 28, …
b. Even if we formulate (12) as to involve **every** such infinite sequence, they need not converge to one particular ratio r.

(16) Prediction of restriction to infinite sets A:
Characterizing sentences and frequency statements express law-like generalizations, which are by their nature applicable to a potentially infinite number of cases.

(17) **Proposal**: Modeling of generics as a relation involving infinite number of word/time indices:
a. *Birds (usually) can fly*: P(λi∈Rλx[x is a bird in i] | λi∈Rλx[x can fly in i]) > 0.5,
b. checking of pairs ⟨i, x⟩, where i is an accessible world, x is an entity;
R: set of worlds where the laws of the world of evaluation hold; a potentially infinite set.

(18) Instances do not have to occur in the real word:
Mail from Antarctica goes in box Z.

(19) Prediction: Randomness leads to a homogeneity requirement
a. #*Israelis live on the coastal plain.*
b. #*People have black hair.*
c. #*People are over three years old.*
d. #*Bees are sterile.*
e. #*Primary school teachers are female*.
f. #*Books are paperback.*

(20) Cohen: Choosing at random requires a salient partition (Time, Space, Age groups etc.) where the generic probabilities should hold.
a. Space partition: Does not hold for Israelis in Jerusalem.
b. Space partitions: Does not hold for people in Northern Europe.
c. Age partition: Does not hold for babies.
d. Type partition: Does not hold for queens, drones.
e. Sex partition: Does not hold for males.
e. Subject partition: Does not hold for text books, etc.

(21) Note that some cases in (19) are already ruled out by intensional sets, e.g. (19)(a)
$P(\lambda i \in R \lambda x[\text{x is an Israeli in i}], \lambda i \in R \lambda x[\text{x lives in the coastal plane i}]) > 0.5$,
not likely true, as this might be different in different possible worlds i in R

4. Why are generics "simple"?

(22) Why are characterizing sentences "simpler" than "true" quantificational sentences, why are they acquired early, why do they probably occur in all languages (including Pirahã)? (cf. Leslie e.a. 2011 – but see Lazaridou-Chatzigoga e.a. 2015)

(23) a. Generic quantification starts out with a simple verification process:
Birds can fly. – check bird b_1 that you happen to find, can it fly? – yes;
– check bird b_2, can it fly? – yes. …
– ...
b. The final truth conditions arise as the limit of this process.

(24) If one just tallies the positive and negative cases, one has to do just a size comparison, no counting necessary, just size comparison (cf. Solt 2016 on *most* and *more than half*):
a. Encountered birds that can fly: …...
b. Encountered birds that cannot fly: …......

(25) Can be modeled by a simple **double pushdown automaton** (van Benthem 1986, for *most*)
a) Whenever you encounter a bird that can fly:
Put a "." on the positive stack; if positive stack is empty: remove a "." from negative stack
b) Whenever you encounter a bird that does not fly:
Put a "." on the negative stack; if negative stack is empty: remove a "." from positive stack.
c) Automaton for *every* and *no* is even easier
(only positive or negative stack, once it falls under 1, sentence is false).
d) No counting necessary, even speakers of languages with no number words can do it.

(26) Truth conditions of a generic sentence can be approximated, different from other quantifiers.

(27) Truth conditions of generic sentence correspond to **cognitive procedure of finding generalizations**, hence they appear so natural.

(28) The information generic sentences convey is not about the whole restrictor set, but rather about **what one should expect when encountering elements** of the restrictor set.

(29) The probability approach can be seen as a kind of **quantification over samples**; the probability-theoretic approach and the quantificational approach can be combined.

(30) They assume (presuppose?) a **potential infinity** of entities in the restrictor, hence they are suitable for expressing general rules.

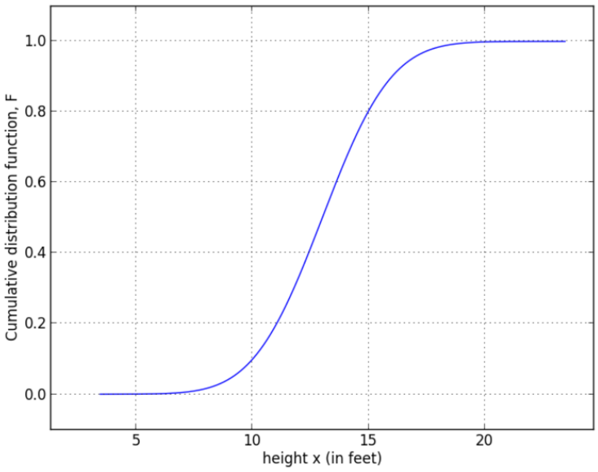
(31) In order to be **useful**, the procedure of arriving at / checking truth conditions of generics must be such that **samples are chosen at random**,
otherwise the truth value would depend on the way how one arrived at it,
which makes it of little value for guiding one's behavior, for interpersonal communication.

(32) Still, **non-random choices** can explain why generics often are interpreted in a **stereotypical** way, often express **prejudices**.

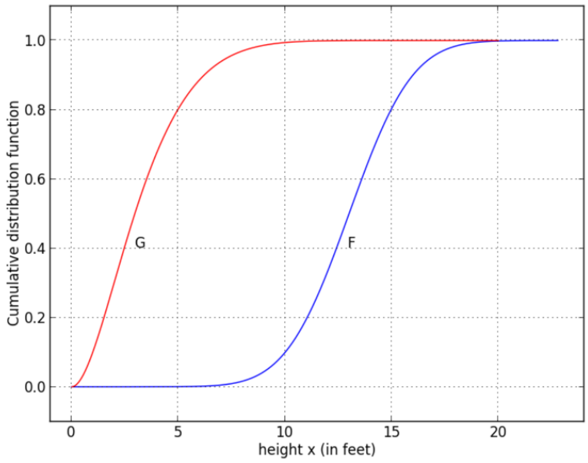
5. The probability approach of Deo & Madiman 2015

(33) Deo & Madiman 2015:
Characterizing sentences as claims about the probability distribution
or random elements along a scale distribution;
what is the probability that a random element from A exhibits a property up to value x?

(34) Represented by cumulative distribution functions (CDF's),
e.g. tallness of giraffes



CDF for height of giraffes

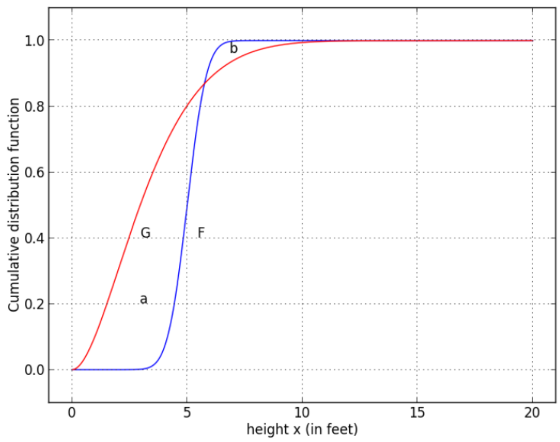


CDF for height of giraffes (F) and mammals (G)

(35) Generic comparisons: *Giraffes are taller than elephants.*
CDF for height of giraffes dominates CDF for height of elephants,
even though some elephants may be taller than some giraffes.

(36) Generics based on implicit comparison classes:
Giraffes are tall.
a. implicit comparison class, e.g. mammals
b. CDF of giraffes (F)
dominates CDF of mammals (G)
even though there might be some non-giraffe
mammals taller than some mammals.

(37) Comparison of CDFs can be subtle:
Horses are tall.
CDF F (horses) dominates CDF G (mammals),
as area b < area a



(38) Relation to comparison class (cf. Krifka 1995
for "distinguishing property" interpretation)
a. *Frenchmen eat horse meat.*
b. *Dutchmen are good sailors.*
The CDF of Frenchmen that eat horse meat is compared with the CDF of people in general.

(39) Remarks on the Deo & Madiman 2015 approach:
a. CDFs involve a rather complex process of lining up all (encountered) individuals following the relevant dimension, but this involves a potentially infinite set of individuals
b. CDFs are plausible representations as a result of generic quantification, but perhaps regular (Gaussian) distributions are more realistic for that.

6. Applicability restrictions, role of focusation

(40) a. *Ducks lay speckled eggs.* only ♀
b. *Ducks have colorful feathers.* only ♂

(41) a. *#Ducks are female.*
b. *#Ducks are male.*

(42) Proposal (Cohen 1999):
a. *lay speckled eggs* has alternatives: other means of giving birth; hence can be applied only to female entities.
b. *have colorful feathers* denotes a sign of sexual dimorphism restricted to one sex; hence can be applied only to entities of one sex.

(43) In quantificational account:
lay speckled eggs: λx: ♀(x) [x lays speckled eggs] or λx: x gives birth [x lays speckled eggs]
a. When applied to male object y: not defined; does not count when computing quantification,
b. GEN(λiλx[x is a duck in i], λiλx: ♀(x) [x lays speckled eggs in i]) for x: male, restrictor is not applicable, is not a falsifying case
c. GEN(λiλx[x is a duck in i], λiλx: Z(x) [x has colorful feathers in i]); Z ∈ {♀, ♂}

(44) In computing probabilities:
P(λiλx[x is a duck in i] | λiλx: ♀(x) [x lays speckled eggs]) > 0.5,
where for computing P(B|A), an x in B to which A cannot be applied does not count.

(45) Role of focus (Krifka 1995, Rooth 1985 for quantifiers, 1995 for characterizing sentences):
a. *Ducks* [*lay speckled EGGS*]$_F$
b. *Ducks lay* [*SPECKLED*]$_F$ *eggs.*
c. focusation introduces alternatives, predicate is restricted to alternatives

(46) In quantificational account:
a. GEN(λiλx[x is a duck in i],
λiλx: ⋃ {λiλx[F(x)] | F∈ ALT(lay speckled eggs)]} [x lay speckled eggs in i])
b. GEN(λiλx[x is a duck in i],
λiλx ⋃ {λiλx [x lay F eggs in i] | F∈ ALT(speckled)} [x lay speckled x in i]}

7. Quantifying over atoms and sums

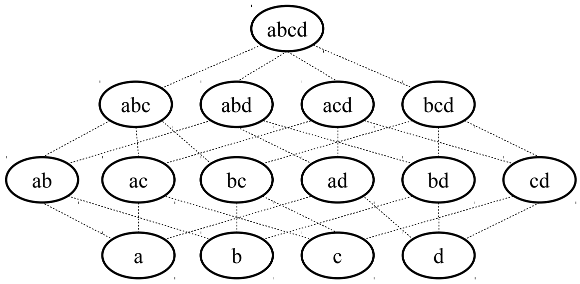
(47) Observation:
You have to take malaria medicine with you. There are many mosquitos, and...
a. *a mosquito carries malaria.* (allows for few exceptions)
b. *mosquitos carry malaria.* (allows for many exceptions)

(48) Existing proposal: Cohen 2001: bare plurals descriptive, indefinite singulars normative
but: *A lion is dangerous. / Lions are dangerous.* – both sentences are descriptive.

(49) Proposal: The number distinction is relevant;
a. singular count nouns apply to atomic individuals,
b. plural count nouns apply to sums
c. mass nouns apply to arbitrary sums as well.

(50) We need this anyway (Gerstner-Link & Krifka 1993, Krifka e.a. 1995)
a. *Lions gather near acacia trees when they are tired.*
b. **A lion gathers near acacia trees when it is tired.*

(51) From n atomic objects $2^n - 1$ objects and $2^n - 1 - n$ proper sum objects can be formed, i.e. there are many more sum objects than atomic objects.



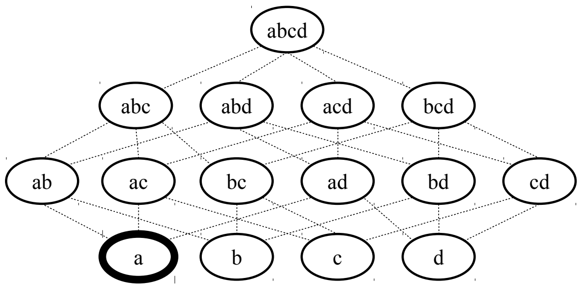
example: 4 atomic objects,
$2^4 - 1 = 15$ objects,
$2^4 - 1 - 4 = 11$ proper sum objects

(52) If S is a countably infinite set ($\aleph_0$), then 2^n, the cardinality of the power set of S, is greater than $\aleph_0$, i.e. it is not countably infinite ($\aleph_1$, following continuum hypothesis).

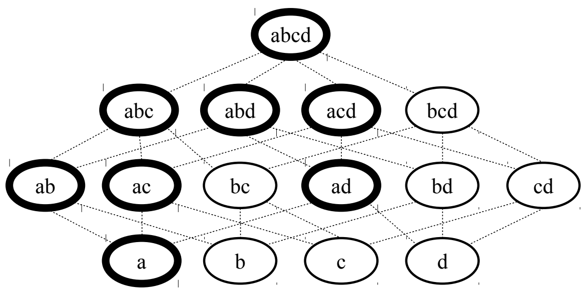
(53) (47)(a) quantifies over atomic mosquitos,
true following the probability account
if P(λiλx[x is a mosquito in i] | λiλx[x carries malaria in i]) > 0.5,
if only 1% of the mosquitos do, this is false.

(54) (47)(b) quantifies over atomic or sum individuals of mosquitos,
true if P(λiλx[x are mosquitos in i] | λiλx[x carry malaria in i]) > 0.5,
if only 1% of the mosquitos do, this may be true!

(55) Assume that there are 4 atomic mosquitos, only one carries malaria;
atomic individuals and sum individuals that **contain** objects that carry malaria:



1 object of 15: P = 0.0167



8 objects of 15: P = 0.5333

(56) Observe:
a. if $\#(\text{pow}(A)) = 2^{\#A}$, then $\#\{A' \mid A' \subseteq \text{pow}(A) \wedge b \in A\} = 2^{\#A}/2$, for all $b \in A$.
b. the sum lattice of sums of elements in A corresponds to $\text{pow}(A) - \{\emptyset\}$,
with $\#(\text{pow}(A) - \{\emptyset\}) = 2^{\#A} - 1$
c. $\#\{A' \mid A' \subseteq \text{pow}(A) \wedge b \in A\} / \#(\text{pow}(A) - \{\emptyset\})$
$= (2^{\#A}/2) / (2^{\#A} - 1)$
$= 2^{\#A}/(2^{\#A} - 1) * ½$
$> ½$
d. That is, if B contains one atomic element b and all the sums of b with elements in A, the ratio #(B) / #(A) is greater than ½, approaching ½ with increasing sets A.
f. If B contains more than just one atomic element, #(B) / #(A) is far greater.

(57) Existential interpretation in non-generic sentences (Krifka 1996):
a. *Be careful – the mosquitos in this jar carry malaria*! (o.k. if only a few do).
b. Notice that the negation of the sentence would not be true:
The mosquitos in this jar do not carry malaria.

(58) Depends on verbal predicate:
a. *The children are sick.* – at least some of the children are sick.
b. *The children are healthy.* – all the children are healthy. (negation of (a)).

(59) a. *My fingers are dirty.* – at least one is.
b. *My fingers are clean.* – all are

(60) a. *I had to return to the house because I left the windows open.* – at least one.
b. *I found that the windows were closed.* – all are.

(61) Distinction between interpretations:
a. existential interpretation: x is sick / dirty / open iff $\exists y[y<x \wedge y$ is sick/dirty/open],
sick / *dirty / open* express presence of sickness / dirtiness / openness
b. universal interpretation: x is clean / healthy / closed iff x is not sick / dirty / open
i.e. iff $\neg\exists y[y<x \wedge y$ is sick/dirty/open]

(62) a. *Mosquitos do not carry malaria.* false characterizing sentences; require that no mosquito
b. *Mosquitos are malaria-free.* (perhaps with few exceptions) carry malaria.

(63) Cumulative interpretation rule for cumulative predicates, and resulting distributivity
a. Cumulativity: If P(x) and P(y), then P(x+y), where x+y: the sum of x and y
b. If predicate P is applicable to x by cumulativity: distributivity, $P(x) \wedge y<x \rightarrow P(y)$

(64) But we often do not assume cumulativity:
a. Collective interpretations: *The apples have a weight of 4 kilograms.*
b. Cum-grano-salis interpretations, Link 1983: *The children built the raft.*
c. Generalization from parts to sums: *The animals in this cage are dangerous.*

(65) Application to other cases, e.g.
a. i. *A Dutchman is a good sailor.* (false)
ii. *Dutchmen are good sailors.* (possibly true)
b. i. *An Israeli lives in the coastal plane.* (odd or false)
ii. *Israelis live in the coastal plane.* (odd or true)
Note that (a), (b) are indistinguishable with the Deo & Madiman account.

(66) Perhaps also:
a. *Ducks lay speckled eggs.* judged true (groups likely contain females)
b. *A duck lays speckled eggs.* less likely judged true (about ½ of single ducks are males.)

(67) Application to definite generics, involving sum individuals as specimen.
a. *The Anopheles mosquito carries malaria.*
b. P(λiλx[x realizes the genus Anopheles in i] | λiλx[x carry malaria in i]) > 0.5,
where x ranges over atomic / sum individuals.

8. References

Benthem, Johan van. 1986. Semantic Automata. In: Benthem, Johan van, (ed), *Essays in logical semantics*. 151-167.

Carlson, Greg N. 1989. The semantic composition of English generic sentences. In: Chierchia, Gennaro, Barbara Partee & Raymond Turner, (eds), *Properties, types and meaning*. Kluwer, 167-192.

Cohen, Ariel. 1999. Generics, frequency adverbs and probability. *Linguistics and Philosophy* 22: 221-253.

Cohen, Ariel. 2001. On the generic use of indefinite singulars. *Journal of Semantics* 18: 183-209.

Gerstner-Link, Claudia & Manfred Krifka. 1993. Genericity. In: Jacobs, Joachim et al., (eds), *Handbuch der Syntax*. Berlin: Walter de Gruyter, 966-978.

Krifka, Manfred et al. 1995. Genericity: an introduction. In: Carlson, Greg N. & F. J. Pelletier, (eds), *The generic book.* Chicago, London: The University of Chicago Press, 1-124.

Krifka, Manfred. 1995. Focus and the interpretation of generic sentences. In: Carlson, Gregory N. & Francis Jeffry Pelletier, (eds), *The Generic Book*. Chicago: Chicago University Press, 238-264.

Krifka, Manfred. 1996. Pragmatic strengthening in donkey sentences and plural predications. *SALT* 6, Rutgers University: CLC Publications, Morrill Hall, Cornell University, 136-153.

Krifka, Manfred. 2013. Definitional generics. In: Mari, Alda, Claire Beyssade & Fabio Del Prete, (eds), *Genericity*. Oxford: 372-389.

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

Leslie, Sarah-Jane, Sangee Khemlani & Sam Glucksberg. 2011. All ducks lay eggs: the generic overgeneralization effect. *Journal of Memory and Language* 65: 15-31.

Lazaridou-Chatzigoga, Dimitra, Katsos, Napoleon, and Stockall, Linnaea. 2015. Genericity is easy? Formal and experimental perspectives. In Hansen, N. and Borg, E. (Eds.), *RATIO* 28(4), Special Issue: Investigating Meaning: Experimental Approaches, 470-494.

Solt, Stephanie. 2016. On measurement and quantification: The case of most and more than half. *Language* 92: 65-100.

von Mises, Richard. 1919. Grundlagen der Wahrscheinlichkeitsrechnung. *Mathematische Zeitschrift*