Generalized Quantifiers

Marco Degano

Structures for Semantics 2023
11 May 2023

Plan

1. Introduction

2. Semantic Universals

3. Monotonicity

4. The Algebraic Interpretation

5. Polyadic Quantification

Literature

Mandatory:

- Gamut 7.2

Further recommended reading:

- Barwise and Cooper. *Generalized quantifiers and natural language.* Linguist Philos, 1981.
- Westerståhl. *Generalized Quantifiers*, SEP.
- Keenan, *The semantics of determiners*, Handbook of Contemporary Semantic Theory, Blackwell, 1996.
- Peters & Westerståhl. Quantifiers in Language and Logic, OUP, 2008
- Keenan & Westerståhl. *Generalized Quantifiers in Linguistics and Logic*. Handbook of Logic and Language 2nd ed., 2011
- Westerståhl, *Generalized quantifiers*. The Cambridge Handbook of Formal Semantics, 2016
- Szymanik. *Quantifiers and Cognition*, Springer 2016.
- Keenan & Paperno (eds.), Handbook of Quantifiers in Natural Language, Springer 2017.

Outline

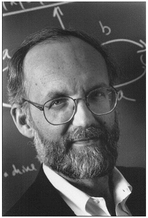
1. Introduction
2. Semantic Universals
3. Monotonicity
4. The Algebraic Interpretation
5. Polyadic Quantification

Natural Language Quantifiers

Quantification is quite common in ordinary language:

- **All** dogs bark.
- **Some** cats are black.
- **Most** people enjoy music.
- **Exactly two** cars were stolen last night.

Barwise & Copper (1981): seminal study on the applications of GQs to natural language.



Jon Barwise (1942 - 2000)

Robin Cooper (1947 -)

Generalized Quantifiers

The term **generalized quantifiers** was introduced by Mostowski (1957) to study mathematically interesting quantifiers not definable in terms of the first-order $\exists$ or $\forall$, like *finitely many* or *most* (originally defined as classes of models closed under isomorphisms).

Andrzej Mostowski (1913 - 1975)

What do you think motivated Isomorphism Closure?

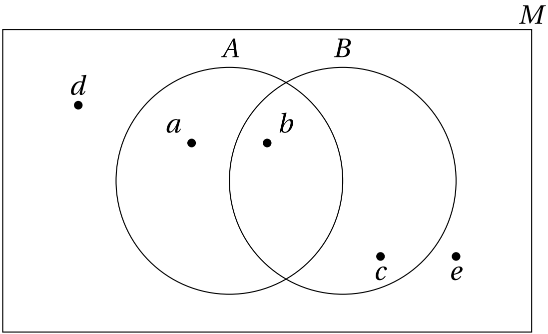
Quantifiers as second-order relations

Given a model M, we can treat a generalized quantifier as a second order relation over subsets of the universe.

$every_M[A,B]$ iff $A^M \subseteq B^M$

$some_M[A,B]$ iff $A^M \cap B^M \neq \emptyset$

$exactly\ two_M[A,B]$ iff $|A^M \cap B^M| = 2$



Types of Quantifiers

A quantifier Q is of type $\langle n_1, \ldots, n_k \rangle$ if it applies to k formulas and binds n_i variables in the ith formula.

- Terms: *John*, *every cat*, ... $\langle 1 \rangle$
- One place determiners: *every*, *most*, ... $\langle 1, 1 \rangle$
- Two place determiners: *more ... than ...* $\langle \langle 1, 1 \rangle, 1 \rangle$
- Reciprocals:

(1) The boys like *each other*. $\langle 1, 2 \rangle$

- *Different* and *same*:

(2) a. Every boy in my class reads a different book. $\langle \langle 1, 1 \rangle, 2 \rangle$
b. Every student answered the same question. $\langle \langle 1, 1 \rangle, 2 \rangle$

Montague Grammar vs GQT for $\langle 1 \rangle$

How did Montague treated terms?

John $\mapsto \lambda P P(j)$
every man $\mapsto \lambda P \forall x (M(x) \to P(x))$

The **syntax** of these logical translations directly reflect their semantic interpretation.

Richard Montague (1930 - 1971)

In GQT, the translation process is trivial and the efforts are concentrated on the **model-theoretic** part. Take a model $M = (D, [\ ])$:

$[\text{John}] = \{X \subseteq D \mid j \in X\}$

$[\text{every man}] = \{X \subseteq D \mid [\text{man}] \subseteq X\}$

Relational vs Functional for Determiners $\langle 1, 1\rangle$

(3) Every man sleeps

Functional:
(every[man])([sleep]) $f \in \left[D_{(e,t)} \rightarrow \left[D_{(e,t)} \rightarrow \{0,1\}\right]\right]$

Relational:
every (man, sleep) $R \subseteq D_{(e,t)} \times D_{(e,t)}$

The two views are equivalent:

Let $f \in [B \rightarrow [A \rightarrow \{0,1\}]]$. Then

$$R_f = \{\langle a, b\rangle \mid a \in A \ \& \ b \in B \ \& \ (f(b))(a) = 1\}$$

Let $R \subseteq A \times B$. Then

$f_R \in [B \rightarrow [A \rightarrow \{0,1\}]]$ is such that for all $b \in B \ \& \ a \in A$:
$\left(f_R(b)\right)(a) = 1$ iff $\langle a, b\rangle \in R$

Some Examples of type $\langle 1,1 \rangle$

$every(A,B) = 1$ iff $A \subseteq B$

$some(A,B) = 1$ iff $A \cap B \neq \emptyset$

$no(A,B) = 1$ iff $A \cap B = \emptyset$

$most(A,B) = 1$ iff $|A \cap B| > |A - B|$

$less\ than\ five(A,B) = 1$ iff $|A \cap B| < 5$

$all\ but\ two(A,B) = 1$ iff $|A - B| = 2$

Generalized Quantifiers

A **monadic** quantifier is of type $\langle 1, \ldots, 1 \rangle$. It is a relation over subsets of M.

A **polyadic** quantifier is of type $\langle n_1, \ldots, n_k \rangle$, where $n_i > 1$ for at least one i. It is a relation between k relations over M, where the i'th relation is n-ary.

Similar definitions can be given for the functional view.

Outline

1. Introduction
2. Semantic Universals
3. Monotonicity
4. The Algebraic Interpretation
5. Polyadic Quantification

Logic and Language

Universal cross-linguistic regularities in the domain of quantifiers?

Relationship between class of natural languages and class of logically possible languages?

QUESTION 1: Are there constraints on which functions (or relations) can be denoted by natural language determiners?

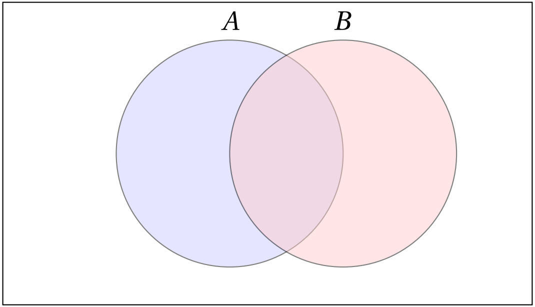
QUESTION 2: Do lexical (=syntactically simple) Dets satisfy stronger constraints on their possible denotations than syntactically complex one?

If $card(M) = n$, how many determiners can we define? (How many functions $(et, (et, t))$?)
2^{4^n}. With just two objects ($n = 2$), we get 65 536.

Isomorphism closure (ISOM) - Topic Neutrality

A determiner D is topic-neutral iff for all $A, B \subseteq M$:

If $(M, A, B) \cong (M', A', B')$, then $D_M(A, B) \leftrightarrow D'_M(A', B')$



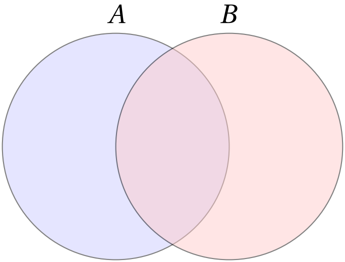
A topic-neutral determiner is a determiner which is not sensitive to the properties of, and the relations between, the elements in the domain and in the sets which it relates.

Most well-known determiners are topic-neutral, e.g. *every*. Possessive determiners like *John's* are not. Why?

Extensionality (EXT)

EXT: A determiner D satisfies extension iff for all $A, B \subseteq M$:

if $M \subseteq M'$, then $D_M(A,B) \Leftrightarrow D_{M'}(A,B)$



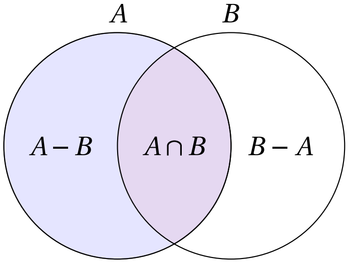
$D_M(A,B)$ does not change its meaning if the universe becomes bigger without affecting the size of A and B. The set $M - (A \cup B)$ is irrelevant.

Generalization: All natural language determiners satisfy EXT.

Conservativity (CONS)

CONS: A determiner D is conservative iff for all $A, B \subseteq M$:

$$D_M(A,B) \Leftrightarrow D_M(A, A \cap B)$$



To verify $D_M(A,B)$ it is sufficient to look at the interpretation of $A - B$ and of $A \cap B$. The set $B - A$ is irrelevant.

Conservativity Test

Conservativity test: Det N VP ⇔ Det N are N that VP.

(4) a. All men smile. ⇔
b. All men are men that smile.

(5) a. Some men smile. ⇔
b. Some men are men that smile.

(6) a. Most men smile. ⇔
b. Most men are men that smile.

Conservativity Generalization

Is CONS a weak or a strong condition? How many quantifiers satisfy CONS?

(Keenan and Stavi 1986):

For $|D_e| = n$,
$\left|D_{\langle\langle e,t\rangle,\langle\langle e,t\rangle t\rangle\rangle}\right| = 2^{4^n}$
$|CONS| = 2^{3^n}$

For $|D_e| = 2$,
$\left|D_{\langle\langle e,t\rangle,\langle\langle e,t\rangle t\rangle\rangle}\right| = 65536$
$|CONS| = 512$

Generalization: (With at most a few exceptions) Natural language determiners are conservative.

Can you think of a non-conservative determiner?

$Only_M(A,B) \Leftrightarrow B \subseteq A$

Only men smile $\not\Leftrightarrow$ Only men are men that smile.

The case of *Only*

Is *only* a determiner? What do you think?

Only is not a determiner. Some evidence:

(7) a. Only/*every/*some/*the/*most John cries.
b. John only/*every/*some/*the/*most cries.
c. John sleeps only/*every/*some/*the/*most with his teddy bear.

Only is a determiner. Some evidence:

(8) All and only boys cry.

Trees

By EXT, $M - (A \cup B)$ is irrelevant.

By CONS, $B - A$ is irrelevant.

By ISOM only the cardinality of $A - B$ and $A \cap B$ count.

Determiners satisfying these three constraints can be perspicuously represented by means of a **tree of numbers** (van Benthem 1983, 1984).

Johan van Benthem
(1949 -)

Triangle Representation

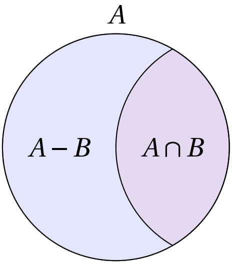
(0, 0)

(1, 0) (0, 1)

(2, 0) (1, 1) (0, 2)

(3, 0) (2, 1) (1, 2) (0, 3)

… … … … …



Pairs $(|A - B|, |A \cap B|)$
Row$_0$: $card(A) = 0$,
Row$_1$: $card(A) = 1$, …

+
- +
- - +
- - - +

Every

-
- +
- + +
- + + +

Some

-
- -
- - +
- - + -

Exactly two

Outline

1. Introduction

2. Semantic Universals

3. Monotonicity

4. The Algebraic Interpretation

5. Polyadic Quantification

Monotonicity

Monotonicity is an important and well-known phenomenon in many contexts.

In general, a function f relative to an ordering $\leq_1$ of the arguments and an ordering $\leq_2$ of the values is **increasing** iff $x \leq_1 y$, then $f(x) \leq_2 f(y)$.

For GQs, we take $\leq_1$ to be **set-inclusion** and $\leq_2$ to be **entailment.**

Type $\langle 1 \rangle$ quantifiers:

Q_M is monotone increasing iff
if $X \subseteq Y \subseteq M$, then $Q_M(X) \Rightarrow Q_M(Y)$

Q_M is monotone decreasing iff
if $X \subseteq Y \subseteq M$, then $Q_M(Y) \Rightarrow Q_M(X)$

Every man	increasing	*No man*	decreasing
Exactly two	neither		

Monotonicity

For Determiners D of type $\langle 1,1\rangle$:

MON↑: A determiner D is **right monotone increasing** iff $B \subseteq B'$ and $D(A)(B)$ then $D(A)\left(B'\right)$

MON↓. A determiner D is **right monotone decreasing** iff $B \subseteq B'$ and $D(A)\left(B'\right)$ then $D(A)(B)$

↑MON. A determiner D is **left monotone increasing** iff $A \subseteq A'$ and $D(A)(B)$ then $D\left(A'\right)(B)$

↓MON. A determiner D is **left monotone decreasing** iff $A \subseteq A'$ and $D\left(A'\right)(B)$ then $D(A)(B)$

$\downarrow Every \uparrow$ $\qquad \uparrow Some \uparrow$ $\qquad Most \uparrow$ $\qquad Exactly\ two$

Monotonicity & Universals

Barwise & Cooper (1981): The simple determiners of any natural language express right monotone quantifiers or conjunctions of monotone quantifiers (e.g, two men := at most two men AND at least two men.).

It also usually holds that left monotone determiners (increasing or decreasing) are also right monotone (increasing or decreasing).

Monotonicity and Reasoning

Monotonicity patterns could be exploited to model reasoning patterns in natural language.

If D is increasing and $X \subseteq Y$, then $D(A, X)$ entails $D(A, Y)$.

(9) a. Every boy sings beautifully.
b. Every boy sings.

If D is decreasing and $Y \subseteq X$, then $D(A, X)$ entails $D(A, Y)$.

(10) a. No boy sings.
b. No boy sings beautifully.

Natural Logic

$Most\uparrow$ has arguably a not trivial semantics, but inferences based on monotonicity patterns are immediate:

(11) a. Most Americans who know a foreign language speak it at home.
b. $\rightsquigarrow$ Most Americans who know a foreign language speak it at home or at work.

The field of *natural logic* (Moss 2015 for an overview) examines different inference patterns based on their computational complexity and closeness to natural language.

(12) a. More than two-thirds of the students passed the exam.
b. At least one-third of the students are athletes.
c. $\rightsquigarrow$ Some student who is an athlete passed the exam.
(Example from Keenan 2005)

The inference pattern $Q(A,B)$ & $Q^{\mathrm{d}}(A,C) \Rightarrow \mathrm{some}(A, B \cap C)$ $[Q^{d} := \neg Q \neg]$ is valid only for $MON\uparrow$ quantifiers.

Universals & Cognition

Why almost all natural language determiners are closed under isomorphism, they satisfy EXT and CONS and they are monotonic?

Barwise & Copper already proposed that these quantifiers are in a sense *simpler* to verify.

This led to an influential research agenda exploring the relationship between logic, complexity and learnability of GQs (e.g., Szymanik 2016 and related work).

Jakub Szymanik

Negative Polarity Items

The English adverb ‘ever’ requires a negative environment and it is thus an example of a Negative Polarity Item (NPI):

(13) Sentential Negation
- a. John hasn’t ever been to Paris.
- b. #John has ever been to Paris.

(14) ‘No’ vs ‘Some’
- a. No student here has ever been to Paris.
- b. #Some student here has ever been to Paris.

(15) ‘At most’ vs ‘At least’
- a. At most five students here have ever been to Paris.
- b. #At least five students here have ever been to Paris.

(16) ‘Every’ vs ‘Some’
- a. Every student who has ever been to Paris speaks French.
- b. #Some student who has ever been to Paris speaks French.

Ladusaw-Fauconnier Generalization

Klima (1964): NPIs must be licensed by a negative expression.

Ladusaw-Fauconnier Generalization: Negative polarity items occur within arguments of monotonic decreasing functions [if $x \leq_1 y$, then $f(y) \leq_2 f(x)$], but not within arguments of monotonic increasing functions.

In the case of **negation**, we take $\leq_1$ and $\leq_2$ as entailment:

$$p \text{ entails } q, \text{ then } \neg q \text{ entails } \neg p$$

(17) Sentential Negation
 a. John hasn't ever been to Paris.
 b. #John has ever been to Paris.

GQs and NPIs

As discussed, GQs can also be described in terms of monotonicity:

(18) ‘No’↓ vs ‘Some’↑
  a. No student here has ever been to Paris.
  b. #Some student here has ever been to Paris.

(19) ‘At most↓’ vs ‘At least↑’
  a. At most five students here have ever been to Paris.
  b. #At least five students here have ever been to Paris.

(20) ‘↓Every’ vs ‘↑Some’
  a. Every student who has ever been to Paris speaks French.
  b. #Some student who has ever been to Paris speaks French.

Monotonicity

What do you think are the merits and limits of this analysis?

It showcases the relevance of formal semantics modelling in natural language.

Not fully explanatory (why monotonicity?)

It requires further amendments for other items (e.g., anti-additivity for *yet*).

It does not explain the occurrence of NPI in questions:

(21) Have you ever been to Paris?

Outline

1. Introduction

2. Semantic Universals

3. Monotonicity

4. The Algebraic Interpretation

5. Polyadic Quantification

The Algebraic Interpretation

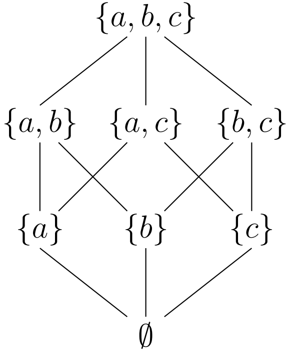
It is natural to think of GQs as having a Boolean structure.

Given a domain D, we can think of the denotation of GQs as sets of properties (i.e., subsets of $\wp(D)$).

$Alice(A) = 1 \iff a \in A$

$[\![Alice]\!] = \{X | a \in X\}$

$[\![Every\ woman]\!] = \{X | W \subseteq X\}$



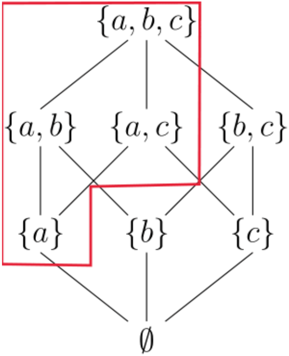
$\langle \wp(D), \subseteq \rangle$ with $D = \{a, b, c\}$

Example: Ultrafilters

The denotation of **Alice** is an ultrafilter on the powerset lattice $\langle P(D), \subseteq \rangle$

$[\![Alice]\!] = \{X | a \in X\}$

Filter:
It is non-empty.
It is upward-closed.
It is closed under binary intersection.
Ultrafilter:
For any set $B \in \wp(D)$, exactly one of $Alice(B)$ or $Alice(\overline{B})$ will hold.



$\langle \wp(D), \subseteq \rangle$ with $D = \{a, b, c\}$

Outline

1. Introduction

2. Semantic Universals

3. Monotonicity

4. The Algebraic Interpretation

5. Polyadic Quantification

Polyadic Quantifiers

(22) Most critics reviewed two books. $\langle 1,1,2 \rangle$

$Q(A,B,R) \Leftrightarrow |A \cap \{a \in A : |B \cap R_a| = 2\}| > |A - \{a \in A : |B \cap R_a| = 2\}|$

$R_a = \{b : Rab\}$

(22) can be formalized as a quantifier $Q(A,B,R)$, where A is the set of critics, B the set of books and R the binary relation *reviewed*. The type of the quantifier is thus $\langle 1,1,2 \rangle$.

$Q(A,B,R)$ can be defined by means of two type $\langle 1,1 \rangle$ quantifiers:

$Q(A,B,R) \Leftrightarrow \textbf{most}(A, \{a : \textbf{two}(B, \{b : Rab\})\})$

This operation is called iteration ($most \cdot two$), and can be generalized as follows:

$$(Q \cdot Q')[A,B,R] \Longleftrightarrow \mathrm{Q}\left[A, \{a : \mathrm{Q}'[B, R_a]\}\right],$$

where $R_a = \{b : Rab\}$.

Frege Boundary

Frege Boundary: All polyadic quantification in natural language is iterated monadic quantification.

(23) Twenty students attended ten courses.
Cumulative: Each of the 20 students attended at least one course and each of the 10 courses was attended by at least one student.

How to capture the cumulative reading of (23)?

With a cumulation operator definable by means of iteration and the existential quantifier:

$$\mathrm{Cum}\left(\mathrm{Q},\mathrm{Q}'\right)[A,B,R] \Longleftrightarrow$$
$$(\mathrm{Q}\cdot Some)[A,B,R] \wedge (\mathrm{Q}'\cdot Some)[B,A,R^{-1}]$$

Irreducible Polyadic Quantifiers

However, not all natural language quantifiers can be expressed by means of iteration.

(24) Most people are grateful to firemen who rescue them.

(24) has a reading where the quantifier applies to the pairs composed of a person and a fireman who rescues that person.

This reading can be captured by **lifting** the type of $most$. This operation is also known as **resumption.**

$$\text{Res}^k(most)_M(R,G) \iff |R \cap G| > |R - G| \qquad \langle k, k\rangle$$

$$\text{Res}^2(most)_M(R,G) \iff |R \cap G| > |R - G| \qquad \langle 2, 2\rangle$$

$R(a,b) \Leftrightarrow a$ rescued b $\qquad G(a,b) \Leftrightarrow a$ is grateful to b

Resumption is not definable from any finite number of monadic quantifiers (Hella, Väänänen, & Westerståhl 1997).

Frege Boundary

Similar lifting strategies can be applied to branching quantifiers and reciprocals (Peters & Westerståhl 2006, ch. 10).

As a result, while the Frege Boundary does probably not hold, it can still be argued that complex polyadic quantifiers can be formed from monadic quantifiers by means of dedicated (lifting) operations.

For a discussion on polyadic quantification and complexity, see Szymanik (2016, Part III).

THANK YOU!