Meaning in linguistic theory:

Formal semantics and Montague Grammar

Patrick D. Elliott https://patrickdelliott.com

September 21, 2021

MIT

Goals of formal semantics

- *What do we know, when we know a language?*
- Given a potentially unbounded number of sentences, human language processors have stable judgments about when one sentence *entails* another.
- *The philosophers sneezed.*
  $\Rightarrow$ *Each of the philosophers sneezed.*
- *The philosophers lifted a piano.*
  $\not\Rightarrow$ *Each of the philosophers lifted a different piano.*

What is the unconscious knowledge that undergirds this capacity?

A central goal of formal semantics:
Account for this productive human capacity by engineering a device able to compute the meaning of an unbounded number of sentences.

A methodological gambit:
logicians and programming language theorists have developed sophisticated tools for recursively computing the meanings of sentences of artificial languages…let's treat natural language as this kind of thing, and see how far we can get.

The logicality of language

Human language processors have fast and unconscious judgements about sentence acceptability.

Let's consider the case of so-called Negative Polarity Items (NPIS).

(1) a. John didn't see any luxans.

b. ✗ John saw any luxans.

(2) a. Aeryn isn't happy at all.

b. ✗ Aeryn is happy at all.

NPIs cont.

Hypothesis 1
NPIS only occur in sentences with negation.

But...

(3) a. Aeryn doubts that John saw any luxans.

b. ✗ Aeryn thinks that John saw any luxans.

Hypothesis 2
NPIS only occur in sentences with a "negative" expression.

But...

(4) a. Everybody with any sense stayed home.

b. Nobody with any sense stayed home.

(4) a. ✗ Everbody who stayed home had any sense.

b. Nobody who stayed home had any sense.

Environments can be distinguished on the basis of their *logical* properties.

Let's say that a predicate P is **more specific** than a predicate Q if P applies to a subset of the entities that Q does. E.g., *Texan* is more specific than *American*.

(5) John is Texan $\overset{\Rightarrow}{\not\Leftarrow}$ John is American.

A simple positive sentence is an **upward entailing** environment, since it licenses inferences from *more specific* to *less specific*.

Some basic logic cont.

In a simple negative sentence, this pattern is flipped.

(6) John isn't American $\overset{\Rightarrow}{\underset{\not\Leftarrow}{}}$ John isn't Texan.

Negation creates a **downward entailing** environment; inferences are licensed from *less specific* to *more specific*.

The Ladusaw-Fauconnier generalization

NPIS are licensed in **downward entailing** environments (Fauconnier 1975, Ladusaw 1979).

As we've seen, this correctly distinguishes between simple positive vs. negative sentences.

But it's much more powerful than this...

Doubt vs. *think*

Recall:

(7) Aeryn { doubts | ✗ thinks } that John saw any luxans.

(8) Aeryn doubts that John is American
$\Rightarrow$ Aeryn doubts that John is Texan.
$\not\Leftarrow$

(9) Aeryn thinks that John is Texan
$\Rightarrow$ Aeryn thinks that John is American.
$\not\Leftarrow$

Recall: *nobody* licenses NPIS indiscriminately, whereas *everybody* licenses NPIS only in a certain environment:

(10) *Everybody* [with any sense] (restrictor) [stayed at home] (scope)

(11) Everybody [who is American] stayed at home.
$\Rightarrow$ / $\not\Leftarrow$ Everybody [who is Texan] stayed at home.

(12) Everybody who stayed at home [is Texan].
$\Rightarrow$ / $\not\Leftarrow$ Everybody who stayed at home [is American].

Extensions

The Ladusaw-Fauconnier generalization extends to a broad range of other cases too, such as the antecedent vs. consequent of a conditional:

(13) a. If you have any problems,
you were probably being lazy.

b. ✗ If you were being lazy,
then you probably have any problems.

You can verify this for yourself — the generalization is remarkably productive, and his been the subject of numerous refinements (see Chierchia 2013 for an overview).

- One of the factors underlying our judgments of sentence acceptability is the **logical/inferential properties** of linguistic environments.
- We can tentatively conclude that human language processors are also *human inference machines* — we compute abstract logical properties on-the-fly.
- Importantly, these judgments are *fast* and *automatic*. Our ability to compute the logical properties of linguistic environments doesn't necessarily rely on a laborious, conscious reasoning process.

The starring role of quantifiers

A widespread view: natural language is messy, imperfect, and *illogical*.

> *"The actual unique reference made, if any, is a matter of the particular use in the particular context; ...Neither Aristotelian nor Russellian rules give the exact logic of any expression of ordinary language;* ***for ordinary language has no exact logic.****"*

— Strawson 1950 (emphasis mine)

- A vivid example of the (apparent) difficulty of using logical tools for analyzing meaning:
- *Every semanticist uses Emacs.*
- We might like to think of this sentence as consisting of a sequence of words, and we might have some idea of the kind of concepts they convey.
- A first-order representation suitable for an inference task:
- $\forall x[\mathbf{semanticist}(x) \rightarrow \mathbf{uses}(x, \mathbf{Emacs})]$
    - What is the meaning of *every semanticist*?
    - What is the meaning of *every*?

The (pre-)history of formal semantics

- A number of techniques needed to be developed/rediscovered in order to surmount this problem:
  - Compositionality via *function-argument application* (Gottlob Frege 19th century work on mathematical logic).
  - Model theory and truth-conditional semantics (Alfred Tarski in the 20th century)
- This culminated in the foundational insight of contemporary formal semantics — Richard Montague's development of a logic that mirrors the structure of natural language, with an analysis of quantification as his crowning achievement.

Natural vs. formal languages

"There is in my opinion no important theoretical difference between natural languages and the artificial languages of logicians; indeed I consider it possible to comprehend the syntax and semantics of both kinds of languages with a single natural and mathematically precise theory."
— Montague 1970

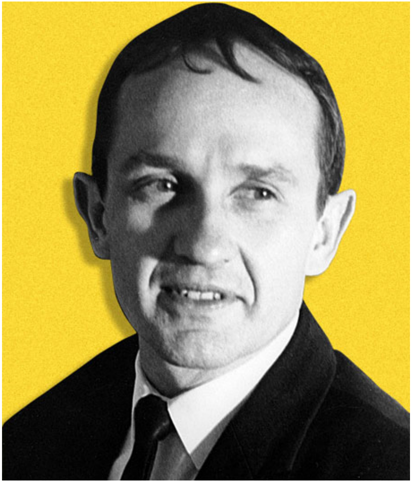


Figure 1: from
`https://richardmontague.com/`

Montague built on Chomsky's (1957) insight that the structure of natural language is amenable to a formal characterization.

$$
\begin{aligned}
S &\rightarrow \text{NP VP} \\
\text{NP} &\rightarrow \text{Det N} \\
\text{Det} &\rightarrow \text{Every} \\
\text{N} &\rightarrow \text{semanticist} \\
&\ldots
\end{aligned}
$$

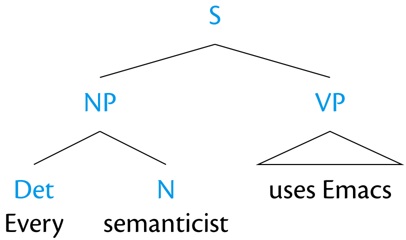
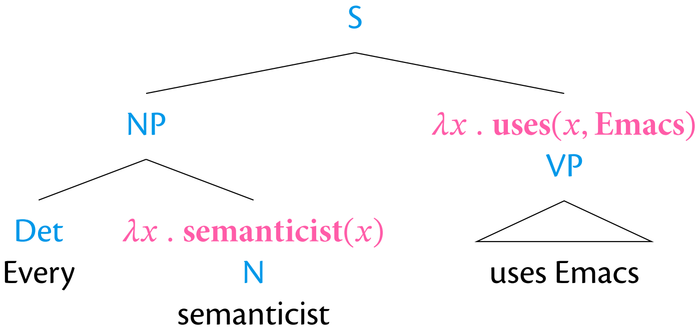
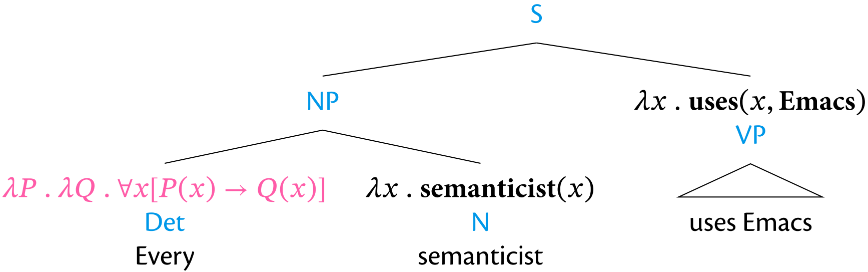


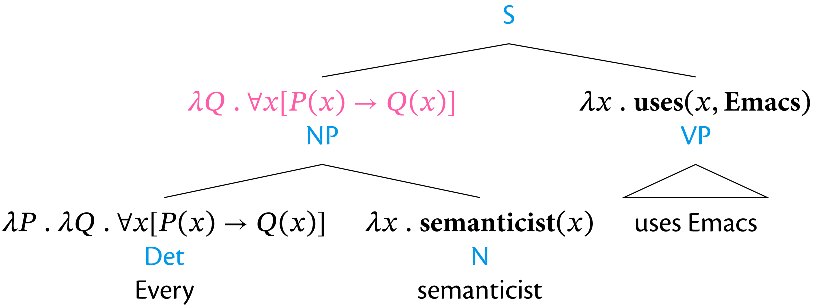
Figure 2: A quantificational sentence.


S
NP
λx . uses(x, Emacs)
VP
Det
λx . semanticist(x)
Every
N
semanticist
uses Emacs

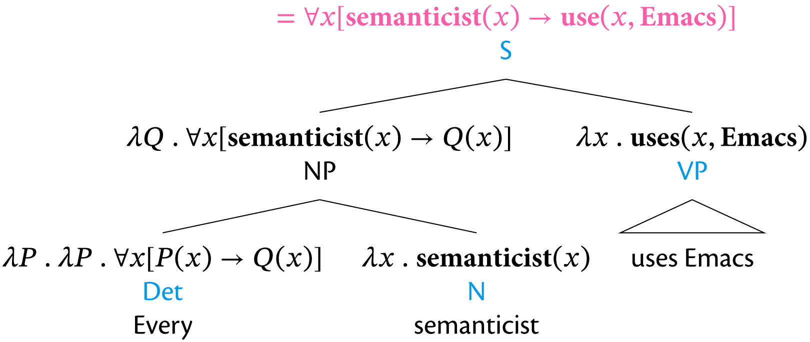
Generalized quantifiers iii



Generalized quantifiers iv



Generalized quantifiers v



A remarkable discovery of formal semantics: the meaning of *any noun phrase* can be characterized as a generalized quantifier.

$$\lambda Q\,.\,\forall x[\mathbf{semanticist}(x) \rightarrow Q(x)]$$
$$\approx \{\, Q \mid \mathbf{semanticist} \subseteq Q \,\}$$

(The function conveyed by *every linguist* characterized the set of sets containing the semanticists)

More GQs

- $[\![\text{some semanticist}]\!] = \{ Q \mid \textbf{semanticist} \cap Q \neq \varnothing \}$
- $[\![\text{between 3 and 5 dogs}]\!] = \{ Q \mid 3 \leq |\textbf{dog} \cap Q| \leq 5 \}$
- $[\![\text{no cat}]\!] = \{ Q \mid \textbf{cat} \cap Q = \varnothing \}$
- $[\![\text{Josie}]\!] = \{ Q \mid \textbf{Josie} \in Q \}$

> *Montague's interpretation of the sentence [...] is the same as Russell's; the big difference is that Montague derives the interpretation compositionally;* ***the semantic structure is homomorphic to the syntactic structure****.*

— Partee 2008 (emphasis mine)

Some open questions

- How can the logical treatment of GQs be combined with probabilistic information; arguably necessary for giving a semantics for *vague* predicates (Goodman & Lassiter 2015, Emerson 2020).
- *Every tall man is annoying.*
- The extension of *tall man* is subject to contextual uncertainty — how does this interact with the austere logic of GQs?

Formal semantics today

Montague's program revolutionized the study of an extremely rich set of empirical phenomena. Some examples:

- Plurality.
- Modal verbs and sentential embedded.
- Comparatives and gradable adjectives.
- Tense and aspect.
- Questions and imperatives.

The logical approach has revealed aspects of natural language understanding which were not even previously recognized as puzzling or surprising.

Beyond the classical toolbox

- The frontiers of formal semantics have significantly expanded over the last couple of decades.
- Linguists have discovered a rich typology of entailments which display rich interactions with the logical properties of the environments they appear in — treating inferences as a homogeneous class is not sufficient.

(14) a. Paul is smoking again. ⇒ *Paul used to smoke*

b. Paul isn't smoking again. ⇒ *Paul used to smoke*

- Sherry will talk about *presuppositions* in more detail in the next section.

Sometimes a sentence can affect the context in a way which seemingly goes beyond their inferential properties.

(15) Andreea is married. # He's waiting outside.

(16) Andreea has a husband. He's waiting outside.

In my recent research (Elliott 2020), I've combined insights from dynamic semantics (Kamp 1981, Heim 1982) and programming language theory (Mcbride & Paterson 2008) to reconcile Montague's approach to meaning with patterns of anaphora resolution.

Future prospects

- Something formal semantics doesn't have much to say about: lexical meaning. Probabilistic/vector-based approaches provide a rich, computationally sophisticated toolbox.
- A big challenge: integrating insights from probabilistic and formal approaches to meaning.
- A small but burgeoning literature on this. More work still to be done (see e.g., Coecke, Sadrzadeh & Clark 2010).

Acknowledgments

I’m grateful to Daniel Harris for guiding me towards Barbara Partee’s wonderful paper *The Starring Role of Quantifiers in the History of Formal Semantics* (2008) — this presentation loosely follows the same narrative.

- Chierchia, Gennaro. 2013. *Logic in grammar - polarity, free choice, and intervention.* (Oxford Studies in Semantics and Pragmatics 2). Oxford: Oxford University Press. 468 pp.
- Chomsky, Noam. 1957. *Syntactic Structures.* Mouton.
- Coecke, Bob, Mehrnoosh Sadrzadeh & Stephen Clark. 2010. *Mathematical Foundations for a Compositional Distributional Model of Meaning.* `http://arxiv.org/abs/1003.4394` (19 September, 2021).
- Elliott, Patrick D. 2020. Towards a principled logic of anaphora. lingbuzz/005562. MIT. Submitted to Semantics & Pragmatics.

- Emerson, Guy. 2020. *Linguists Who Use Probabilistic Models Love Them: Quantification in Functional Distributional Semantics.* http://arxiv.org/abs/2006.03002 (19 September, 2021).
- Fauconnier, G. 1975. Polarity and the Scale Principle. *undefined.*
- Goodman, Noah D. & Daniel Lassiter. 2015. Probabilistic semantics and pragmatics: Uncertainty in language and thought. In.

- Heim, Irene. 1982. *The semantics of definite and indefinite noun phrases*. University of Massachusetts - Amherst dissertation.
- Kamp, Hans. 1981. A theory of truth and semantic representation. In Paul Portner & Barbara H. Partee (eds.), *Formal semantics: The essential readings*, 189–222. Blackwell.
- Ladusaw, William A. 1979. *Polarity sensitivity as inherent scope relations*. University of Texas at Austin dissertation.
- Mcbride, Conor & Ross Paterson. 2008. Applicative programming with effects. *Journal of Functional Programming* 18(1).

Montague, Richard. 1970. Universal grammar. *Theoria* 36(3). 373–398.

Strawson, P. F. 1950. On Referring. *Mind* 59(235). 320–344.