This course presupposes familiarity with elementary set theory, propositional logic, and predicate logic, as in Allwood et al., Logic in Linguistics, Chs 1-5, although there will be some brush-up assignments.

The goal of the course is to introduce some basic mathematical and logical tools, and the kind of insights they offer into empirical and theoretical linguistic issues. It is followed by Semantics II.

Calendar:

Week 1 Compositionality and some of its consequences
Week 2 Presuppositions and implicatures; Determiners as relations between sets
Week 3 Competing accounts of the definiteness effect

Week 4 Operations in partially ordered sets
Week 5 Generalized quantifiers
Week 6 Polarity: monotonicity, focus, implicatures

Week 7 Free and bound variables, quantifiers
Week 8 Starting to build a grammar; models, assignments
Week 9 Categorial grammar and type theory
Week 10 Lambdas

Week 11 Empirical aspects: How to obtain reliable judgments?
Week 12 Scope a la Montague and May
Week 13 Different quantifiers, different scopes
Week 14 Student presentations

Readings:
- a selection of articles
- class handouts, to be used as lecture notes

Requirements:
- weekly readings and written assignments
- a squib or paper – small group projects are very welcome if every member makes a well-defined contribution

If more time is needed to discuss issues that are either difficult or particularly interesting to some members of the class, meetings may on occasion start at 9:30 or 10:00am. To make time for classroom discussion, meetings may regularly start at 10:45am.
Compositionality and Some Consequences

The compositionality principle is a fundamental methodological assumption concerning the relationship between syntax and semantics. It says, "The meaning of a complex expression is a function of the meanings of its parts and the way they are put together." The principle is historically attributed to Gottlob Frege (although it turns out he never stated it explicitly) and has been realized most influentially in pioneering work by Richard Montague (in UG and PTQ).

The principle can be read in more than one way. Minimally, it is a commonsense requirement that entails at least the following:

(a) The sentences of natural language are not all idioms: we do not learn their meanings as unanalyzed wholes but, rather, compute them using some kind of effective procedure. Therefore, it is not sufficient to propose global interpretations for sentences, without specifying how these interpretations are arrived at. (NB Philosophers often engage in such global speculations about meaning.)

(b) Since syntax gives us the units and puts them together, semantics must take syntax seriously. Conversely, interpretability may choose between alternative syntactic analyses.

(c) The meaning of a complex expression cannot depend on anything else than the meanings of constituents and of the operations that built it; e.g., it cannot depend on further context. If some aspects of interpretation do depend on further context, they do not belong to "semantics proper"; they may belong to pragmatics. E.g. Could you pass the salt? interpreted as 'Please pass the salt' is not a matter of compositional semantics.

(d) In view of the existence of indexical expressions (I, you, this, now, yesterday, etc.) context independence must be recast as follows: a complex expression cannot be more context dependent than its constituents are.

The assumption of compositionality in this broad sense underlies practically all the substantial empirical questions that have been asked about semantics in the past thirty years.

Conceived this broadly, a compositional grammar can be devised in a number of ways. For instance, we could have an algorithm that takes the final stage of a syntactic derivation as its input and assigns it a meaning using a variety of templates. This is the interpretation procedure traditionally implied by generative syntacticians; see Heim and Kratzer (1998) for an implementation. Linguists who spell out compositional analyses often follow Montague's specific bottom-up way of implementing the principle (and deviate from it in well-defined ways if they find it necessary). Given the bottom-up derivations of the Minimalist Program, this may be more generally applicable now than 30 years ago, when it was proposed.
Montague's version

Montague assigned (i) an explicit interpretation to every word (morpheme), (ii) explicated exactly how each rule operates on the interpretation of the input expressions, and (iii) required that the interpretation of the complex depend solely on the meaning, but never the form, of the constituents (or on anything else).

This idea can be captured formally quite precisely. An algebra is a set closed under some operations. The algebra of syntax consists of the set of all expressions of the language (lexical and complex) plus the set of its syntactic operations (by which the complex expressions are obtained from the lexical ones). The algebra of semantics consists of the set of all meanings of the language (lexical and complex) plus the set of its semantic operations (by which the complex meanings are obtained from the lexical ones). What meanings are does not matter here. The two algebras have to be similar in the sense that they contain the same number of operations that can be paired up so the members of the pair have the same arity. (iii) is ensured by the requirement that the function that associates meanings with expressions be a homomorphism. Let E be an expression, M a meaning, Syn a one-place syntactic rule, and Sem a matching one-place semantic rule. Then, the meaning assignment function h is a homomorphism iff h(Syn(E))=Sem(h(E)). In our diagram: the result of applying Syn to E, i.e. Syn(E) is E*; the result of h assigning a meaning to E, i.e. h(E) is M. Because h is a homomorphism, h(E*) coincides with Sem(M), i.e. both lead to the same M*.

Montague's approach is often referred to as "rule-by-rule" interpretation. Whenever a subexpression is assembled, its meaning is immediately specified.

Consider some general issues that arise in connection with compositionality.

Subsentential expressions

One consequence of Montague's strategy is that we must be able to assign meanings to all subparts of complex expressions. Let us illustrate some of the issues that come up here. In deriving John walk[s] it might be easy to say that the interpretation of John is the individual John, the interpretation of walk is the set of walkers, and the interpretation of the concatenation of subject and predicate is to assert that the former is an element of the latter. But what shall we do with Every man walk[s], Few men walk, etc? We might say that every man denotes the set of men, and Every man walks asserts that this set is a
subset of the set of walkers. Note that we are forced to interpret the subject-predicate relation
differently in John walks and Every man walks and may thus miss a significant generalization. Worse is
the case of Few men walk, where there is no comparable unique individual or set that we could assign
to the subject. Montague proposed a general solution by appealing to the theory of generalized
quantifiers and the lambda calculus in assigning interpretations to expressions and to operations.

What is "meaning"?

The assumption of compositionality has consequences for the definition of meaning. Above we were
assuming that the interpretations of John, walk, etc. are just the things they refer to (=denote) in a
particular moment (=in one model/world), and the interpretation of John walks is its corresponding
truth value.

The question is whether this is an (even remotely) adequate notion of meaning. Leibniz's principle says
that identicals are interchangeable salva veritate (=preserving truth). Concretely, if two expressions
have identical meanings, they must be interchangeable in any context. This is the guiding principle in
deciding what kind of a beast meaning should be. Meaning is defined instrumentally as the kind of
thing that allows us to conform to Leibniz's principle.

Reference turns out to be inadequate. For instance, imagine that our world is such that all and only
those who walk catch a flu, i.e., walk and catch a flu refer to the same set. But I want to walk may be
true while I want to catch a flu is not. Similarly, John walks and John catches a flu may both be true in
our world but Necessarily, John walks and Necessarily, John catches a flu can at the same time differ in
truth value. Thus, extensional semantics (the one that does not go beyond reference) is insufficient.

Drawing from fundamental results by Frege and Carnap, Montague proposed to use intensions as
meanings. Intensional logic relies on a class of possible worlds, rather than just one world; the intension
of an expression is a function that specifies, for each possible world, what that expression denotes in
that world. Two expressions are synonymous if they have the same intension, i.e. the same denotation
in every possible world. (Notation, to be remembered: For any expression α, \(^\alpha\) denotes its intension. If
β denotes an intension, ^β is its value in the world under consideration, i.e. its extension. ^α = α, but
^β ≠ β.)

^Walk and ^catch a flu will clearly be distinct. It is easy to imagine a possible world where not all and
only walkers catch a flu; the two intension functions are not the same.

NB A proposition is the set of worlds in which a given sentence is true; equivalently, it is a function
from worlds to truth values (i.e. the characteristic function of that set).

The equation of meaning with intension is an important but insufficient step. It turns out that two
logically equivalent expressions (that are true in the same set of possible worlds) need not be
interchangeable salva veritate. Two famous contexts where they fail are perception reports and
propositional attitudes. E.g. John walks entails John walks or Einstein dances, but Mary saw
John walk does not entail Mary saw John walk or Einstein dance. Or, any two tautologies are logically equivalent, but substituting an arbitrary mathematical truth for the complement clause in John is aware that two plus two equals four does not preserve the truth of the whole. One solution to this problem involves partiality: models are no longer supposed to represent largest sets of consistent propositions.

The methodology can be summarized as: Define "meaning" so as to be able to observe compositionality.

Caveat: while it is important to see that the meaning of an expression cannot be equated with its reference, for many problems addressed in this course (and for many problems addressed in current literature), extensional semantics is sufficient.

Cross-linguistic variation

Generative grammar has demonstrated beyond reasonable doubt that natural languages, while superficially wildly different, exhibit very detailed and thoroughgoing structural similarities; in other words, that "universal grammar" is not merely a wishful thought. Therefore no theory incapable of accounting for the unity behind the superficial variation stands a chance to be an even remotely valid theory of natural language.

Cross-linguistic variation in syntax is to some extent paralleled by cross-linguistic variation in interpretation. Two simple examples. (i) Given the right predicate, bare plurals in English and German may have an existential or a generic reading:

(1) Professors are sick.
    ‘There are professors stricken with illness’
    ‘Professors in general are disgusting’

But it is well-known that in many other languages, Romance languages among them, one or both interpretations may be unavailable.

(ii) The interaction of negation with disjunction and conjunction in English and German straightforwardly bears out the de Morgan laws:

(2) John didn’t study flute or accounting.
    ‘neither’         cf. not(p or q) = not p and not q
(3) John didn’t study flute and accounting.
    ‘not both’        cf. not(p and q) = not p or not q

In many other languages, Russian, Italian, Japanese, and Hungarian among them, the above interpretations are absent. The literal counterparts of (2) mean exclusively 'One or the other he didn’t study’ and the literal counterparts of (3) mean exclusively 'He studied neither one’.

The weakest understanding of compositionality is that we just need some effective procedure that delivers the correct interpretations for the constructions of the individual languages. Given the variation observed above, this will not suffice. What is needed is a compositional analysis that also accounts for exactly how languages differ. Without that, human languages will appear to be incommensurable.
Semantics versus pragmatics

Some of the simplest “logical words” are the names of natural numbers. So what does two mean?

(4) How many children do you have? I have two children.
(5) If you have two children, you are eligible for this benefit.

One example suggests that two means ‘exactly two’; the other suggests ‘at least two’. Is two ambiguous? Is its meaning dependent on the context? If yes, does this argue against compositionality? A rigorous investigation of the semantics/pragmatics boundary helps answer such questions.

Read: (a) Semantics I in Fromkin 2000.
(b) Gamut book Vol.2. Ch.1, The origins of intensional logic.
(c) Chapters 2 and 3 of Allwood et al.

Recommended:

General background:

Barwise, Jon and John Perry (1983), Situations and Attitudes. MIT Press.
Carnap, Rudolf (1947), Meaning and Necessity. Chicago UP.
Frege, Gottlob (1982), On sense and reference. In Geach and Black, eds., Translations from the

Specific references:

In (1)-(3), we assume a universe \( U = \{a,b,c,d\} \), \( A = \{a,b,c\} \), \( B = \{b,c\} \), and \( \emptyset \) is the empty set.

(1) (a) What are the elements of \( A \cap B \)?

(b) What are the elements of \( A \cup B \)?

(c) What are the elements of \( A \cap \emptyset \)?

(d) What are the elements of \( A \cup \emptyset \)?

(e) Complete the following: If a set \( X \) is smaller than or equal to \( Y \) (=is a subset of), then their union is the set ______ and their intersection is the set ______.

(2) (a) What are (all) the sets \( X \) such that \( A \cap X = A \)?

(b) What are (all) the sets \( X \) such that \( A \cup X = A \)?

(3) (a) Is \( \emptyset \) an element of \( A \)? ______ Explain why.

(b) Is \( \emptyset \) a subset of \( A \)? ______ Explain why.

(4) (a) Let \( K \) be the set \( \{k,m,n\} \). Give the power set of \( K \).

(b) Let \( L \) be the set \( \{k,m, \emptyset \} \). Give the power set of \( L \).

(c) Let \( N \) be the set \( \{k,m,\{ \emptyset \}\} \). Give the power set of \( N \).
(5) Venn-diagrams can be used to determine whether two sequences of set theoretic operations lead to the same result. For instance:

Does it make a difference whether (i) we first intersect $A$ and $B$ and then union the result with $C$, or (ii) we union $B$ and $C$ and then intersect the result with $A$?

$$(A \cap B) \cup C \ =?\ A \cap (B \cup C)$$

Using the same technique, determine the following.

(a) $(A \cap B) \cap C \ =?\ A \cap (B \cap C)$

(b) $(A \cap B) \cup C \ =?\ (A \cup C) \cap (B \cup C)$
Below you find the definitions of three functions in terms of ordered pairs:

\[ f = \{<a,b>, <b,c>, <d,b>\} \]
\[ g = \{<b,a>, <c,b>, <d,b>\} \]
\[ h = \{<b,c>, <d,b>, <a,b>\} \]

Are these functions identical? Justify your answer (i) in your own words, and (ii) by representing functions graphically (a separate diagram for each function). Use arrows to indicate how the function maps each element of the domain to an element of the co-domain.

<table>
<thead>
<tr>
<th>domain</th>
<th>co-domain (range)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td>c</td>
<td>c</td>
</tr>
<tr>
<td>d</td>
<td>d</td>
</tr>
</tbody>
</table>

(7) Make a one-letter change in the definition of \( f \) so that the result is no longer a function.

(8) Using Venn-diagrams, determine whether the following equivalences are valid or not.
(a) \(- (A \cup B) = (-A) \cup (-B)\)
(b) \(- (A \cap B) = (-A) \cup (-B)\)

(9) That \( p \) logically entails \( q \) means that whenever \( p \) is true, \( q \) is also true. Let \( P \) be the set of worlds in which \( p \) is true (=the truth set of \( p \)), and similarly for \( Q \). Which of the following Venn-diagrams will represent their relationship correctly?

(a) \( p \Rightarrow q \) means \( P \subseteq Q \) or (b) \( p \Rightarrow q \) means \( Q \subseteq P \)
Plan for this unit: (a) Discuss some particular semantic and pragmatic components of meaning that are not part of the assertion. (b) Discuss some basic properties of determiners. (c) Apply (a) to certain problems in (b).

Semantics: Presuppositions

A sentence S logically presupposes a sentence T if S cannot be meaningful (true or false) if T is false.

Test: A sentence S logically presupposes a sentence T if and only if both S and its natural full negation entail T.

Bill regretted offending Peter. No, he didn't / That's not true.
Both S and not-S presuppose that Bill offended Peter.

The king of France is bald. No, he isn't / That's not true.
Both S and not-S presuppose that France has a king.

A sentence S logically asserts a sentence T if and only if S entails T, but S does not presuppose T.

S1: Bill regretted offending Peter
   entails: Bill offended Peter
   Bill regretted that...
   No, he didn't (regret offending Peter)
   entails: Bill offended Peter
   Bill did not regret that...
S1 PRESUPPOSES that Bill offended Peter and ASSERTS that he regretted it.

S2: Bill contemplated offending Peter
   entails: Bill contemplated offending Peter
   No, he didn't (contemplate offending Peter)
   entails: Bill did not contemplate offending Peter
S2 neither presupposes nor asserts that Bill offended Peter; it ASSERTS that he contemplated it.

S3: Bill denied offending Peter
   entails: Bill was accused of offending Peter
   Bill said that he did not offend Peter
   No, he didn't (deny offending Peter)
   entails: Bill was accused of offending Peter
   Bill did not say that he did not offend Peter
S3 neither presupposes nor asserts that Bill offended Peter; it PRESUPPOSES that Bill was accused of offending Peter and ASSERTS that he denied it.
Two sentences may have the same entailments, but a different presupposition—assertion division:
S1: The doctor who fainted was sick.
   Preupposes: exactly one doctor fainted
S2: Exactly one doctor fainted and any doctor who fainted was sick.
   Asserts: exactly one doctor fainted

Some further examples of presuppositional expressions (see a list in Keenan):

Factive predicates
   That Fred left surprised me.
   It is a scandal that Fred left.
   Fred resents Zelda's infidelity.
   Fred noticed that the money was missing.
   Fred knew that the Earth was round.

Make things more precise: Is remember factive?
   Bill remembered that he offended Peter.
   Bill remembered offending Peter.
   Bill remembered to offend Peter.

Responsive predicates
   Bill agreed that the film was good.
   Bill accepted that the Earth was square.
   Bill confirmed that his car was missing.

Aspectual predicates
   Bill stopped smoking.
   Bill quit drinking.
   Bill continued walking.
   Bill resumed speaking.

Selectional restrictions of verbs
   That Fred was there surprised Mary (presupposes: dir. obj. is a thinking being)
   Oil was dripping (presupposes: subj. is a liquid)

"Definite" noun phrases
   Bill called.
   Bill called Fred's sister.
   Bill believed what the doctor said.
   How he escaped is a mystery.

Iteratives
   Fred called again.
   Fred ate another turnip.

Non-restrictive relative clauses
My brother, who lost his ticket, called me at night.

**Pragmatics: Presuppositions**

As in Stalnaker and much current work – part of the common ground.

**Pragmatics: Gricean implicatures**

Conventional implicature:

“If I say (smugly) *He is an Englishman; he is, therefore, brave*, I have certainly committed myself, by virtue of the meaning of my words, to its being the case that his being brave is a consequence of (follows from) his being an Englishman... I do not want to say that my utterance of this sentence would be, strictly speaking, false should the consequence in question fail to hold.” (Grice 1989, p.25)

For Potts (2003), a conventional implicature is, roughly, a logical presupposition that is not part of the common ground, e.g. the content of a non-restrictive relative or the contribution of *even*:

*Even John likes this film.*

Asserts: John likes this film

Presence of *even* conventionally implicates: Someone besides John likes this film. John is the least likely to like this film.

**Co-operative Principles (aka Conversational/Gricean Maxims):**

Quantity: Be just as informative as required
Quality: Try to make your contribution one that is true
Relation: Be relevant
Manner: Be perspicuous

**Conversational implicature:** Arises from a violation (flouting) of some Co-operative Principle; may well be situation dependent.

**Implicature arising from Quantity violation:**

Is he a good doctor? -- His office is nice, and his assistant is pretty.

**Implicature arising from Quality violation:**

Irony: (in pouring rain) What a great weather!

Metaphor: You are the cream in my coffee

Meiosis: (of man breaking up all furniture) He is a little intoxicated

Hyperbole: Every nice girl loves a sailor

**Implicature arising from Relation violation:**

Mrs. X is an old bag. -- Would you like some coffee?

Is she in today? -- Today is Thursday.

**Particularized conversational implicatures** exploit the special features of the situation.
Generalized conversational implicatures do not pertain to a particular situation; they live off of the form of the expression. Cancellable; non-detachable; initially not part of the expression’s meaning; not part of “what is said”.

“X went into a house yesterday and found a tortoise inside... My hearer would normally be surprised if some time later I revealed that the house was X’s own.”

The Gricean reasoning: If the speaker knows that the house was X’s own and this has some relevance for the conversation, the maxim of Quantity would require the use of his/her house. The hearer assumes that the speaker is cooperative, therefore the use of a house leads the hearer to believe that it was not X’s house.

Generalized conversational implicatures have been used very extensively in the semantics literature to relieve compositional semantics of the burden of accounting for certain aspects of interpretation. More recently it has been argued that some of these bits of interpretation (in scalar implicatures) should be sent back to the semantics or at least to the recursive component of grammar: Landman, Chierchia, Kratzer, Fox.

A crucial question going beyond Grice: Are generalized conversational implicatures present in all linguistic environments?

Read:
Keenan: Two kinds of presupposition in natural language
Grice: Logic and conversation

Recommended (from classics to work in progress):
Chierchia, 2003, Scalar implicatures, polarity phenomena, and the syntax/semantics interface
Fox, 2006, Free-choice disjunctions and the theory of scalar implicatures
Gazdar, 1979: Pragmatics: Implicature, Presupposition, and Logical Form
Heim, 1983, On the projection problem for presuppositions
Horn, 1984, Toward a new taxonomy of pragmatic inference
Kratzer, 2003, Scalar implicatures: Are there any
Landman, 2001, Events and Plurality
Potts, 2003, The Logic of Conventional Implicature
Schlenker, 2006, Transparency: An incremental theory of presupposition projection
Stalnaker, 1973, Presuppositions
Stalnaker, 1978, Assertion
Propositional Logic Assignment: Allwood et al. Ch. 4. (due Thursday)

Please internalize the tautologies in (8), pp. 55.

(1) The following equivalence can be easily explained using set theory. t [true] stands for 'tautology,' f [false] for 'contradiction':

(a) \( t \land p \equiv p \)
the conjunction of a tautology \( t \) and a sentence \( p \) that may be either true or false has the same truth value as \( p \)

Let the set of worlds in which \( p \) is true be \( P \). Since \( t \) is a tautology and a tautology is true in every world, \( P \subseteq T \). We know that whenever \( A \subseteq B \), \( A \cap B = A \). Hence, \( P \cap T = P \).

Translate (b)-(c)-(d) into English and explain them with reference to truth sets and set theory, as above:

(b) \( p \lor t \equiv t \)
(c) \( p \land f \equiv f \)
(d) \( p \lor f \equiv p \)

(2) Let \# be a propositional connective whose truth table is as follows:

<table>
<thead>
<tr>
<th>( p )</th>
<th>#</th>
<th>( q )</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>T</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>F</td>
</tr>
</tbody>
</table>

(a) Is \# identical to one of the standard connectives discussed in the book? If yes, which?
(b) Draw a Venn diagram and shade the areas in which \( p \# q \) is true.
(c) Express \# using nothing but negation + conjunction or negation + disjunction. That is, \( p \# q = ??? \)

(3) Determine whether the following are tautologies, using truth tables.

(a) \( (p \rightarrow q) \rightarrow (p \lor q) \)
(b) \( (p \land q) \equiv \neg (p \rightarrow (\neg q \lor \neg q)) \)

(4) Optional (but very much encouraged): Show how \( \rightarrow, \lor, \land \) can be expressed using just Sheffer's stroke \( | \) (which means 'not both'). Hint: notice that \( p \vert p \) is equivalent to \( \neg p \) (as below) and go on to get two (different) propositional variables into the picture.

| \( p \) | \( | \) | \( p \) |
|---|---|---|
| t | f | t |
| f | t | f |
Determiners

- Recall the compositionality issue: What do subsentential expressions denote?
  Each D denotes a relation between NP-denotations and VP-denotations.

\[ NP \cap \neg VP \quad \neg NP \cap VP \quad NP \cap VP \]

\[
\begin{align*}
\text{every}(\text{man})(\text{walks}) = \text{true} & \iff \text{man} \subseteq \text{walk} \\
\text{no}(\text{man})(\text{walks}) = \text{true} & \iff \text{man} \cap \text{walk} = \emptyset \\
\text{at least one}(\text{man})(\text{walks}) = \text{true} & \iff \text{man} \cap \text{walk} \neq \emptyset \quad \text{i.e.} \quad |\text{man} \cap \text{walk}| \geq 1 \\
\text{most-of-the}(\text{men})(\text{walk}) = \text{true} & \iff |\text{man} \cap \text{walk}| > |\text{man} \cap \neg \text{walk}| \\
\end{align*}
\]

- Do natural language determiners refer to arbitrary relations between their two arguments? Does their definition make reference to arbitrary areas of the diagram?

\[ \text{INTersectivity:} \quad \text{the truth of } D(NP)(VP) \text{ depends only on } NP \cap VP \quad \text{e.g. at least one} \]
\[ \text{CO-INTersectivity:} \quad \text{the truth of } D(NP)(VP) \text{ depends only on } NP \cap \neg VP \quad \text{e.g. every} \]
\[ \text{CONServativity:} \quad D(NP)(VP) = D(NP)(NP \cap VP) \quad \text{all nat.lg. Ds} \]
\[ \text{Domain independence:} \quad \text{the truth of } D(NP)(VP) \text{ does not depend on } \neg NP \cap VP \quad \text{all nat. lg. Ds} \]
\[ \text{RESTrictedness:} \quad \text{CONS} + \text{domain independence} \quad \text{all nat.lg. Ds} \]

Restrictedness means that natural language determiners can be seen as quantifying over just a designated subset of the universe, that denoted by their NP: a nice match between syntax and semantics.

Questions:
- Are there determiners that must work this way, i.e. for which quantifying over the whole universe would be plainly wrong?
- What do the and two mean?
The need for restricted quantification: the case of \textit{most}

Let $\text{Mx}[f(x)]$ be true iff more than 50\% of universe has property $f$. Now compare:

<table>
<thead>
<tr>
<th>World #1: the universe contains just</th>
<th>World #2: the universe contains just</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 dogs, of which 8 bark</td>
<td>3 dogs, which do not bark</td>
</tr>
<tr>
<td>20 cats, which do not bark</td>
<td>20 cats, which do not bark</td>
</tr>
</tbody>
</table>

true or false in #1: in #2:

- \textit{Most of the dogs bark} \hspace{2cm} ____ \hspace{2cm} ____
- $\text{Mx}[\text{dog}(x) \rightarrow \text{bark}(x)]$ \hspace{2cm} ____ \hspace{2cm} ____
- $\text{Mx}[\text{dog}(x) \& \text{bark}(x)]$ \hspace{2cm} ____ \hspace{2cm} ____
- $\text{Mx}[\text{dog}(x) \lor \text{bark}(x)]$ \hspace{2cm} ____ \hspace{2cm} ____

The discrepancy between the results for the English sentence and the attempted predicate logical representations is due to the fact that \textit{most} is proportional, and it talks about percentages of the dogs, not of the entities in the universe. Its meaning is not definable in predicate logic in the strict sense, where variables take their values from the entire universe. Its truth conditions can be adequately captured if we shrink the universe so the quantifier looks at dogs only. Conservativity says that all natural language quantifiers work well with such a shrunked universe, although only the proportional ones require it.

What does \textit{two} mean?

Some of the simplest “logical words” are the names of natural numbers. So what does \textit{two} mean?

(1) How many children do you have? I have two children.
(2) If you have two children, you are eligible for this benefit.

(1) seems to be most readily interpreted as ‘I have exactly two children’, and (2) as ‘If you have two or more children,…’. So is (2) ambiguous between ‘exactly two’ and ‘at least two’? Problems with the ambiguity hypothesis: (a) Why aren’t both readings equally available in (1) and (2)? (b) If ‘exactly two’ is a real meaning, (3) should be able to mean ‘I have either zero or one or three or five … children’.

(3) I don’t have two children.

But that is not a reading of (3); for that reading we need contrastive focus on \textit{two}:

(4) I don’t have TWO children, I have ONE/THREE.

The (Neo-)Gricean approach: the lexical interpretation is ‘at least two’. This is strengthened to ‘two and not more’ by the conversational maxims. If it is relevant how many children the person has AND the speaker is well-informed and cooperative, then she should make the strongest possible claim. If she had three children, she should have said so. So it is a generalized conversational implicature that any stronger proposition is understood NOT to be true, i.e. (1) implicates that there
are two children and not more. Then, (2) will have an ‘at least’ interpretation, because having more children may well be irrelevant for eligibility for the benefit.

But the irrelevance part does not very easily carry over to (3). In fact, something more systematic can be said about (2) and its kin. There is a whole set of contexts that systematically lack the ‘and no more’ implicature (in the absence of focus). In addition to (2) and (3), consider:

(5) Do you have two children? (Yes, I have even three.)
(6) Few people have two children. (Most people have at most one.)

The contexts in which the ‘and not more’ implicature is absent can be characterized semantically as downward entailing (other terms: monotonically decreasing, implication reversing):

Let $A \leq B$. If $A$ and $B$ are propositions, this amounts to $A \rightarrow B$.

OP is downward entailing iff $\text{OP}(B) \leq \text{OP}(A)$.

E.g. truth set (I have at least three children) $\leq$ truth set (I have at least two children)
‘whenever I have at least three, I also have at least two’
truth set (I do not have at least two ch.) $\leq$ truth set (I do not have at least three ch.)
‘whenever I fail to have at least two, I also fail to have at least three’

Why is the ‘and not more’ implicature (‘exactly’ interpretation) absent in downward entailing contexts? Negation reverses the entailment scale. The 3-children claim is stronger than the 2-children claim in the positive cases but weaker than the 2-children claim in the negative cases. An implicature (here ‘and not more’) is added when adding it excludes a stronger claim, and this only obtains in the positive cases.
Assignment:

Written (please type, but feel free to draw circles, trees, etc. by hand):

All of these require some prose. Please give thoughtful answers, but they need not be very long.

(1) Examine the inclusive and exclusive interpretations of or in natural language, based on the discussion in Allwood et al. Summarize the facts. Propose a unified analysis using (neo-)Gricean reasoning.

(2) Propose a reasonable interpretation for the determiner the. If you find it controversial, defend it.

(3) Exx. 7.3 and 7.13 from Fromkin, ed. In addition to answering the questions as formulated in the exercise, discuss the moral for the syntax of extraction. [These exercises are based on Cattell 1978, On the source of interrogative adverbs, in Language.]

(4) Exx 7.5 and 7.7 -- When answering questions about conservativity, many people find it easiest to look at the Venn diagrams and ask, Does the definition of this determiner allow me to ignore what individuals are in VP ∩ ~NP? But of course you can also follow the book's lead at the bottom of p. 382 ("... entails and is entailed by..."), that gives the same result.

Readings, required:

Grice: Logic and conversation
Keenan: Two kinds of presupposition in natural language
Keenan, The semantics of determiners – for the time being, only up to section 4 (p.48)

Next week we (start) discuss(ing) the definiteness effect, based on Milsark, Keenan, and Zucchi. It will speed up the discussion of the raw data if you read Milsark in advance:

Milsark, Toward an explanation of certain peculiarities in the existential construction in English

Logic brush-up:

Please review propositional logic in Allwood et al. If you are not familiar with indirect reasoning, you can skip it. However, the tautologies that the book chooses to present in that context are important irrespective of indirect reasoning, so understand and, hm, try to memorize them. -- We may meet at 9:45am on Tuesday.

If you want to do more exercises in propositional logic, please use Ch 2 of Gamut I. In addition to the excellent text, this book contains detailed solutions to the exercises. A good way of self-teaching is to do one problem, check your solution against the book’s, make sure you understand why it differs from yours if it does (Did you make a mistake? Were there two good but distinct solutions? Is your solution logically equivalent to that of the book?), and only go to the next problem if everything is clear.
(1) I give you the predicate logic formula for some simple sentences. Your task is to build more complex ones by elaborating on that input, step by step.

**John walks:**

John walks and talks:

John or Mary talks:

John saw Mary:

**Something barks:**

Something is a dog and it barks =

Something that is a dog barks =

Some dog barks:

Some dog barks or growls:

Some dog does not bark:

No dog barks:

John saw some dog:

**Everything changes:**

Everything changes if it is a rule =

Everything that is a rule changes =

Every rule changes:

Every rule changes or gets eliminated:

Not every rule changes:

No rule changes:

Every dog saw John:

Please do exx. 1 and 2 from p. 93. The back of the book has the solutions. Please check your solutions against the book’s and if there is divergence, understand what the book is doing. Mark the problem if you can’t figure out why the book has what it does. – Doing these counts as part of the assignment.
Competing explanations for the definiteness effect, Part I

What D's may occur in _there_-contexts, and why?

Milsark’s terminology and answer: Ds that occur in _there_-contexts are "weak"; those that do not are "strong". Weak D’s are cardinality markers; strong Ds are quantifiers. _There_ is an existential quantifier. Its application to a quantified DP results in double quantification: *.

Problems with Milsark’s answer:
Does not quite mesh with what Ds e.g. cross-sentential anaphora classifies as quantificational.
Language-specific: requires the presence of an expletive.
Explanation does not use the fact that the meaning is specifically existential.

Keenan's answer: Weak D's are essentially the intersective ones. [Keenan calls them existential D’s.]

A(n), some, one, two, three, ..., no, more/less than one, exactly one, cardinal few, cardinal many are intersective. _The, every, most, proportional few, proportional many_ are not intersective.

Why is intersectivity relevant? Two steps.

(i) Intersective determiners are sortally reducible, that is, they can be expressed using quantification over the whole universe:

\[ \text{three(cat)}(\text{in the yard}) = \text{three(entity)}(\text{cat } \cap \text{in the yard}) \]

(ii) In the simple cases we are considering, intersectives are also symmetrical, hence

\[ \text{three(entity)}(\text{cat } \cap \text{in the yard}) = \text{three(cat } \cap \text{in the yard)(entity)} \]

[[entity]] is the same as [[exist]] – to exist is the same as to be an entity in the universe of discourse.

Thus, if an intersective determiner occurs in the _there_-sentence, the sentence makes an assertion of existence, as desired. (_There_ itself is meaningless). If we have a non-intersective determiner, the _there_-sentence can at best be put to some other use.

Some problems with Keenan's answer:
Does not explain why _There is every cat in the yard_ is actually felt to be ungrammatical, only that it cannot be used to assert existence.
Does not naturally predict the following contrast, since the two dets have the same (asserted) meanings, given that both are tautological:

1. ok  There are _either zero or more than zero_ cats in the yard
2. * There are _all or else not all_ cats in the yard
• A note on list readings

Context: Who could help?
(1) *There’s every friend of mine / *There are most friends of mine
(2) There’s Susan / the guy next door / There are my friends

In contrast to other strong DPs, definites have so-called list readings in there-contexts. These do not assert existence but, instead, draw attention to someone/something.

Interestingly, list readings are absent from a number of contexts (this inventory is my own):

(3) *Is there the guy next door?
(4) *There isn’t the guy next door.
(5) *There’s rarely the guy next door. [cf. There’s always/usually the guy next door]
(6) *If there’s the guy next door, you needn’t worry.
(7) *It doesn’t seem that there’s the guy next door.

It may be that these are bad simply because they don’t happen to be appropriate for drawing attention to something. But there seems to be a regularity here: these are downward entailing contexts. Downward entailment is linguistically relevant for many reasons. We have seen one already: it prevents the emergence of conversational implicatures. What might be its significance here?

Assignment:

The OR problem from last time. If you did it already but without reading the TWO story, please think it over again. Develop a precise, step-by-step argument of what you want the “dictionary meaning” of OR to be and how you account for the range of possible uses of OR (and for what may not be possible).

Read “Many quantifiers” by Partee; summarize her general argument regarding the ambiguity of MANY and FEW and her specific argument concerning the definiteness effect.

Read Keenan, The semantics of determiners, pp.54-61 (from Constraints... to Definite...).

Leave appropriate time for these readings. They are short and clear but you have to proceed carefully from sentence to sentence.

Read Zucchi, The ingredients of the definiteness effect, pp.33-57, in preparation for next week’s discussion; we will not cover the Formal implementation part in this class.
Zucchi’s answer: Go back to the +/- definite distinction, but in a new non-trivial sense.

Definiteness according to Heim 1982: FC + DCC
  FC: Familiarity Condition: When a definite is uttered, its discourse referent is familiar.
  DCC: Descriptive Content Condition: Definites presuppose that the NP-set is non-empty.

Weak Ds don’t presuppose NP-set to be non-empty; strong Ds do (note: this amounts to interpreting some Ds differently from Keenan). (1) and (2) are distinguished, because all is presuppositional.

The “pure semantic question”: Why cannot presuppositional DPs occur in there-context?

First try: A sentence should assert something that it does not presuppose. But:
(3) ok The maid who is singing is singing.

Second try: Specifically there-sentences are felicitous only when the context does not entail either that NP-set is empty or that it is not empty. (The sentence makes a genuinely new assertion regarding existence.) If D is presuppositional, the context needs to entail that NP-set is nonempty.

The “compositionality question”: How are syntactic components used by the semantics?
(i) How does the meaning of the CODA come into play when it is not part of postverbal NP?

Note: people who know you at the meeting is not a NP:
(4) The people who know you are bound to protest, and there will be some people who know you at the meeting.
Nonemptiness of [[people who know you]] is entailed by first clause; still, second clause is acceptable. The reason must be that the nonemptiness of [[people who know you]]∩[[at the meeting]] is relevant, and that, luckily, is not entailed by context. Thus:

Third try: There-sentences are felicitous only when the context does not entail that [[NP]]∩[[coda]] is empty or that it is not empty. -- But:
(5) * There is every cat in the yard.
Every cat presupposes that cats exist, not that cats that are in the yard exist; * is not accounted for.

So, interpret coda not as predicated of DP but as domain restriction on every, cf.:
  Every cat is sleepy ‘every (cat ∩ in contextual domain) is sleepy’
Now coda is presupposed part of NP, whether or not it is syntactically part of NP.

(ii) How come (6) is okay, although first clause entails the existence of mistakes?
(6) There are some mistakes, indeed, there are five mistakes.
Do not treat numeral as determiner, let it be a cardinality adjective:
\[[five mistakes]\] = \{x: x is an individual made up of 5 atoms that are mistakes\}.
Then second clause is okay, since first does not entail that there are such 5-atom individuals.

**How do Keenan’s and Zucchi’s proposals compare?**

Keenan: weak = intersective; strong = not intersective
Zucchi: weak = not presuppositional, maybe a cardinality adjective; strong = presuppositional

These are very related properties:
(a) Among one-place natural language determiners, intersective = symmetrical. If presuppositionality is acknowledged, presuppositional D’s cannot be symmetrical (NP and Pred are not interchangeable if there is a presupposition pertaining to just NP); hence if D is presuppositional, it is never intersective.
(b) Only intersective D’s can reincarnate as cardinality adjectives.

So Keenan’s and Zucchi’s explications of “weak” and “strong” can be related to each other through their contents and they pick out fairly similar, though not identical, sets of data (at least for the simple determiners we considered). Still, Keenan’s and Zucchi’s definitions differ in that they rely on possibly different interpretations of the individual determiners (are presuppositions acknowledged?), they may package meanings differently (where does the coda’s content belong?) and, most importantly, they serve as bases for rather different explanations of the definiteness effect.

*  

**Rejoinder: Keenan again** (pp. 195-213 of Keenan 2003)

Some of the problems with Zucchi’s proposal:
Does not characterize weak DPs directly (pace learning from positive evidence).
*All* and both not equally presuppositional.
Predicts presuppositional (7)-(9) to be out, while they are not:
(7) There are(n’t) just two mistakes in your paper.
(8) There were no students but John at the party.
(9) There are(n’t) only/mostly freshmen in that course.
   only(A)(B) = all(B)(A)
   mostly(A)(B) = most(B)(A)

New proposal: D’s that occur in there-sentences limit their universe to the coda set (are conservative in their second argument, CONS-2).
D is CONS-2 iff D(A)(B) = D(A \lor B)(B)
There are mostly freshmen in the course = Most of those in-the-c are freshmen-in-the-c.
Note: D is INT iff both CONS-1 and CONS-2. D may be CONS-2 but not CONS-1.

More precisely: Weak DPs are (Boolean compounds) of DP’s built from lexical CONS-2 D’s.
Some of the definiteness effect literature:


Buring (1996), A weak theory of strong readings. Proc. of Semantics and Linguistic Theory (SALT) VI.

Comorovski (1995), Partitives and the definiteness effect, in Bach et al. Quantification in Natural Languages.

Diesing (1992), Indefinites. MIT Press.

Keenan (1987), A semantic definition of "indefinite NP". In Reuland--ter Meulen, eds.
[summarized in his handbook chapter and in his 2003 NALS paper].


Re OR:
from Chierchia, Scalar Implicatures, Polarity Phenomena, and the Syntax/Pragmatics Interface, 2001

.../

In the rest of this introduction, in order to set our baseline, I summarize (a version of) the NeoGricean stand on implicatures. The phenomenon is well known. The truth conditional content of a sentence like (1a) is taken to be (1b). Yet, such a sentence typically conversationally implicates (1c). Similarly for the sentences in (2)

(1) a. John is singing or screaming
   b. singing(j) ∨ screaming(j)
   c. ¬ (singing(j) ∧ screaming(j))

(2) a. Some students did well
    b. ∃x [student(x) ∧ did well(x)] (Some and possibly all students did well)
    c. ¬ ∀x [student(x) → did well(x)] (Not every student did well)

Implicatures of this sort arise whenever expressions that may be viewed as part taking in an informational scale are involved. For example, the (positive) quantifiers can be thought as being ordered along a scale of informativeness as follows:

(3) The (positive) quantifier scale: some < many < most < every

The reason why this is so is that, for example, a sentence like (4a) asymmetrically entails (and hence is informationally stronger than) a sentence like (4b):¹

(4) a. every man smokes
    b. some man smokes

More generally, whenever a determiner D occurs to the right of a determiner $D^+$ in the scale in (3), a sentence of the form “D N V” will entail a sentence of the form “$D^+ N V$”. Similarly, and and or can be thought of as part of an informational scale, as $p$ and $q$ (asymmetrically) entails $p$ or $q$, and hence the former is inherently more informative than the latter. The centrality of the notion of scale for implicatures has been motivated especially in Horn’s work (see references above); what can constitute a natural scale is somewhat controversial and need not concern us here.² Typical scales, besides < or, and> and the positive quantifiers in (3), are the following:

¹ The relevant entailment holds whenever the set of men is non-empty. Or equivalently, we will be assuming a presuppositional view of determiners like “every”, according to which “every N” is taken to be uninterpretable is there are no Ns is the domain of quantification. See, e.g., de Jong and Verkuyl (1985)
² On scales, besides the works cited in the text, see, e.g. Gazdar (1979), Hirschberg (1985). Heim (class lectures, Fall 99) individuates in monotonicity a necessary condition for being part of a scale. For arguments that monotonicity is not also a sufficient condition, see Landman (1998).
(5) Examples of Horn scales:
   a. Negative quantifiers: not all < few < none
   b. Predicates: cute < beautiful < stupendous
discrete < good < excellent
   ...  
c. Numerals: 1 < ... n < ....
d. the modals: possibly < necessarily
   may < must

where:
\[ \leq \] (“ is informationally weaker than ”) =df (asymmetrically) entails  

SIs derive from the systematic exploitation of Grice’s conversational maxims (especially, relevance and quantity). The way in which they come about may be schematically illustrated through an example (inspired most directly to Landman (1998)). Consider:

(6) a. Who is in that room?
b. John or Bill
c. John and Bill

Suppose a hearer gets (6b) as an answer to question (6a); s/he will then typically come to conclude that the answer in (6b) implicates that (6c) does not hold (i.e. that John and Bill are not both in the room) in the following (idealized) way:

(7) i. The speaker said (6b) and not (6c), which would have been also relevant
   ii. (6c) entails (6b) [ or and and are part of a scale)]
   iii. If the speaker had the info that (6c), she would have said so [quantity]
   iv. The speaker has no evidence that (6c) holds
   v. The speaker is well informed
Therefore:
   vi. It is unlikely/not the case that (6c) holds

/.../

---

3 Two caveats are in order. First, the notion of entailment is to be understood as generalized in the usual way to all types (that “end in t” – see Partee and Rooth 1983). Second, entailment must be understood as being relativized to contextually shared knowledge in the sense of Stalnaker (1978). See Heim (1984) and von Fintel (1999) for relevant elaborations of Stalnaker’s view.
Operations in partially ordered sets, assignment

- Reading
  Szabolcsi, Background notions in lattice theory and generalized quantifiers. In Ways of Scope Taking, section 1.1.
  Davey-Priestley, Maps between ordered sets (the excerpts attached to this handout)

- Written

  (1) Problems 58 and 59 from Background notions... The end of the chapter has detailed solutions. Please write down your own solutions, then check them against the book; if they differ, please try to figure out why the book claims what it does. If you are unsure, please alert me.

  (2) Take the set $E=\{a,b,c\}$ and the relation $R_n$ on it, as defined below. Is $<E,R_n>$ a poset? Why?

  $R_1=\{<a,c>, <b,c>, <c,a>, <b,b>\}$
  $R_2=\{<a,b>, <b,c>, <a,c>, <b,b>\}$
  $R_3=\{<a,b>, <b,c>, <a,a>, <b,b>, <c,c>\}$
  $R_4=\{<a,b>, <b,c>, <a,c>, <a,a>, <b,b>, <c,c>\}$

  (3) Davey—Priestley say that in Figure 1.3 map 5 is order preserving but not an order-embedding, whereas map 6 is an order-embedding. Demonstrate why, by writing out the correspondences that show both the positive and the negative claims to be true.

  (4) Davey—Priestley say that map 5 in Figure 1.3 preserves (order and) joins, but not all meets. Show why this is the case, by writing out what the joins and the meets of each pair of elements are in the domain set and in the range set and pointing out where divergences occur.

  (5) Exx 5 and 11 from Gamut I.

More on Boolean algebras

- I try to avoid taking the examples from math, but some of you may be interested to find, just as M. Jourdain found out that he had always been speaking in prose, that you have been doing lattice-theoretic meets and joins ever since you learned to do greatest common factors and least common multiples. You may want to think through exactly how these are the same; perhaps also think about whether you’d find very comparable applications to complements.

- We showed what it means that propositional logical conjunction, disjunction, and negation are just “domain specific” realizations of the general lattice theoretic meet, join, and complement operations, respectively. Let’s recap how this went. First, the name of this game is to look at truth values, rather than truth sets (if we looked at truth sets, we would really be talking about how the set theoretic operations are “domain specific” realizations of the lattice theoretic ones). In classical, non-modal propositional logic each propositional variable can be assigned just one of two values: T or F; no finer distinctions can be made. So there are only these two objects to consider. Our poset will be

\[
\text{PropL} = \langle \{T,F\}, \geq \rangle
\]

where the relation \( \geq \) is defined as follows:

\[
\begin{align*}
T & \geq T \\
F & \geq F \\
T & \geq F
\end{align*}
\]

The following Hasse diagram represents PropL:

![](image)

We check that PropL is a Boolean algebra by inspection, i.e. using the lattice theoretic definitions of meets, joins, and complements, we convince ourselves that these exist and do not lead out of PropL. It is important to realize that PropL being a Boolean algebra is not a stipulation (that’s how we want it to be); it is a matter of fact.

\[
\begin{align*}
T \text{ meet } T &= T \\
T \text{ join } T &= T \\
\text{ compl}(T) &= F \\
T \text{ meet } F &= F \\
T \text{ join } F &= T \\
\text{ compl}(F) &= T \\
F \text{ meet } F &= F \\
F \text{ join } F &= F \\
F \text{ meet } T &= F \\
F \text{ join } T &= T
\end{align*}
\]

We furthermore convince ourselves that we have just drawn up the well-known truth tables of propositional logical conjunction, disjunction, and negation.

- We also said that set theoretic intersection, union, and complement are “domain specific” realizations of the general lattice theoretic meet, join, and complement operations, respectively. To see this, we have to choose a poset whose elements themselves are sets. Unlike truth-values, sets are numerous, so the domain of the poset can be chosen in many ways (though not arbitrarily, as we’ll see later). The ordering relation will be the subset relation. Let us look at this:

\[
A = \langle \{\{a\}, \{b\}, \{a,b\}, \emptyset\}, \subseteq \rangle
\]

You can easily convince yourselves that A is a Boolean algebra, and moreover that for all \( x,y \in A \),

- \( x \text{ meet } y \) is the same as \( x \cap y \),
- \( x \text{ join } y \) is the same as \( x \cup y \), and
- \( \text{ complement}(x) \) is the same as \( \neg x \).
First read the rest of the Backgrounds chapter. If you wish, you may skip the discussion of (22)-(24) and pp. 16-17, which we did not discuss. Ignore formulas with $\lambda$-operators; they are mentioned only so that people who use the lambda notation know how it fits in. You can also skip the problems and solution at the end of the chapter. For the exercises, have clean copies of the Boolean algebra at hand.

Do not start on the exercises before you have read the whole text at least once. But the exercises will also help with digesting the text, so the best way is to go back and forth.

(1) In our universe \{a,b,c,d\}, or indeed in any finite universe, existential and the universal quantification can be expressed in terms of disjunctions and conjunctions, as follows:

$$\exists x [\text{mad}(x)] = \text{mad}(a) \lor \text{mad}(b) \lor \text{mad}(c) \lor \text{mad}(d)$$

‘there is a thing that is mad: it is either a or b or c or d, or more than one of these’

$$\forall x [\text{mad}(x)] = \text{mad}(a) \land \text{mad}(b) \land \text{mad}(c) \land \text{mad}(d)$$

‘every thing is mad, that is, a is and b is and c is and d is’

Likewise, $$\exists x [\text{man}(x) \land \text{mad}(x)] = \text{mad}(a) \lor \text{mad}(b) \lor \text{mad}(c)$$, and similarly for $$\forall$$.

These are important, because now we know that the two quantifiers inherit their properties from the Boolean operations in terms of which they are defined.

Now let us see how this works when we use \textit{generalized quantifiers}. Demonstrate how we can trace every man and some (=at least one) man back to the individual men. What will you have to union and intersect to do these? Perform these operations by shading the relevant areas of the diagram and check whether the union/intersection coincides with the generalized quantifiers every man and some man.

(2) The Backgrounds chapter identifies the GQ-level correlate of conservativity. If a noun phrase is composed of a D and a NP, then $$D(NP)(VP) = D(NP)(NP \cap VP)$$. The GQ-level correlate of the restrictor set is the smallest set that the GQ lives on. We say that GQ lives on the set A iff the following holds: $$VP \in GQ = (A \cap VP) \in GQ$$.

$$D(NP)(VP) = D(NP)(NP \cap VP)$$ and GQ is built from D and NP, then

$$GQ(VP) = GQ(NP \cap VP)$$

$$VP \in GQ = (A \cap VP) \in GQ$$ so NP is a set that GQ lives on

From now on, we can use “restrictor set” and “smallest live-on set” interchangeably.

We can now define the notion of a witness set of a GQ. This is useful, because it makes possible to make various intuitive observations that the notion of an element of the GQ does not.

W is a witness set of some GQ iff W is an element of that GQ and W is a subset of the smallest set the GQ lives on.
Encircle the set of witnesses of \([\text{every man}]\), \([\text{exactly two men}]\), \([\text{less than three men}]\) and compare them with the GQs themselves. (Note that a GQ is nothing other than the set of its own elements.)

(a) Are the monotonicity properties of the GQs preserved when we switch to the sets of their witnesses? Explain this with reference to the encircled areas of the Boolean algebra. E.g. If GQ is monotonically increasing, then the set of its elements .... In contrast OR Likewise, the set of its witnesses .... Etc.

(b) Use observations about witness sets to characterize indefinites versus principal filter denoters. Assume a universe that is “big enough” to bear out the differences.

(3)Take the DP every man but Bill; define the GQ it denotes using our usual language of set theory, and encircle the GQ in our Boolean algebra (with b=Bill). Then, explain how the GQ it denotes can be built from the denotations of its subexpressions. In doing so it will be useful to consult the diagram just drawn. What interpretation do you assign to but?

[After you are done, recommended reading is the first part of Moltmann 1996 in L&P.]
Monotonicity and NPI-licensing

- Monotonicity ($\text{MON}^\uparrow$, $\text{MON}^\downarrow$, $\neg\text{MON}$):

Take a functional perspective on determiners and DPs. Determiner-denotation is a function from NP-denotations to DP-denotations. DP-denotation is a function from Pred-denotations to sentence denotations (truth values).

Let $A \subseteq B$. Note: $p \Rightarrow q$ is $P \subseteq Q$

A function $f$ is mon. increasing iff it preserves the subset relation in its domain, i.e. $f(A) \subseteq f(B)$. A function $f$ is mon. decreasing iff it reverses the subset relation, in its domain, i.e. $f(B) \subseteq f(A)$. A function $f$ is non-monotonic iff neither.

With increasing functions, whatever is true in a small universe remains true if we enlarge the universe. With decreasing functions, whatever is true is a large universe remains true if we shrink it. With non-monotonic functions, we cannot make any such inference.

Some increasing DPs: John, the men, every man, three men, more than five men
   every man walks $\Rightarrow$ every man moves

Some decreasing DPs: neither John nor Mary, no man, few men, fewer than five men
   no man moves $\Rightarrow$ no man walks

Some non-monotonic DPs: only John, exactly five men
   Exactly five men walk/move $=/=\Rightarrow$ and $<=/=\Rightarrow$ Exactly five men move/walk

Some Ds that are increasing in their first, NP-argument: three, more than five
   three men walk $\Rightarrow$ three humans walk

Some Ds that are increasing in their first, NP-argument: every, neither, no, few, fewer than five
   every human walks $\Rightarrow$ every man walks

Some non-monotonic Ds: exactly five
   Exactly five men/humans walk $=/=\Rightarrow$ and $<=/=\Rightarrow$ Exactly five humans/men walk

Note: increasing DP $=/=\Rightarrow$ increasing determiner! every is the famous example: decreasing in its NP argument but increasing in its Pred argument, i.e. $\downarrow\text{MON}^\uparrow$:
   every human walks $\Rightarrow$ every man walks
   every man walks $\Rightarrow$ every man moves
Negative Polarity Item (NPI) licensing: from morpho-syntax to semantics

* John wanted any more wine.
ok John didn’t want any more wine. ok If you drink any more wine, you get drunk.
ok No one wanted any more wine. ok He hardly drank any more wine.
ok Few men wanted any more wine. ok Did he drink any more wine?

Klima 1964: NPI must be within the scope of an n-word (roughly).
Ladusaw 1980: NPI must be within the scope of a monotonically decreasing operator.

- Confirmation from every (decreasing in NP but increasing in Pred argument, i.e. ↓MON↑):
ok Every man who has ever drunk any wine knows that.
every human walks ⇒ every man walks
* Every man has ever drunk any wine.
every man walks ⇒ every man moves

Caveat: many NPIs have so-called free choice (FC) counterparts: anyone, anything, anywhere.
Use any more N, a single N, ever, yet, at all, much, all that Adj, budge an inch, etc. to test NPI-claims.

- Some qualifications needed:

NPI must be in the immediate scope of decreasing operator (Kroch 1979, Linebarger 1980, Honcoop 1998):
I don’t think that John/*everyone has ever been here.
What we need is decreasingness in some extended sense, to handle problem cases (von Fintel 1999):
Only John has ever been here.
I regret/am surprised that John has ever been here.
Many NPIs need a stronger licenser (Zwarts 1981, Giannakidou 1998, Szabolcsi 2004):
No one/*Less than five boys have been here yet.

- Why does NPI have to be within the scope of a decreasing operator?

Kadmon and Landman 1993: any is a domain-widener that comes with the requirement that the proposition containing it must make a stronger claim than the one with plain a(n).
*I saw any dog is bad because it is not stronger (is in fact weaker) than I saw a dog.
Lahiri’s 1998 answer, based on Hindi: the NPI ek/koii bhii means “even one/some thing”. Because “one/some” is a minimal amount, its implicatures contradict those arising from even. The contradiction can only be resolved in a decreasing (scale reversing) environment.
The problem with *I saw anything is assimilated to the problem with (Do you have a kid?) * I have even one!
Caveat: These explanations only account for subsets of NPIs: those that are focused minimal amounts and want just decreasing licensors. See Chierchia 2001 and Section 11 of Szabolcsi 2004.
Assignment:

Complete Keenan’s “Semantics of Dets”.

True or false? If false, give at least one counterexample. “An NPI is…” is meant to be true only if there are no exceptions.

[A] An NPI is licensed within the scope of an expression E if and only if E is monotone decreasing.
[B] An NPI is licensed within the scope of an expression E if E is monotone decreasing.
[C] An NPI is licensed within the scope of an expression E only if E is monotone decreasing.
[D] The NPI ever is licensed within the scope of an expression E if and only if E is monotone decreasing.
[E] The NPI ever is licensed within the scope of an expression E if E is monotone decreasing.
[F] The NPI ever is licensed within the scope of an expression E only if E is monotone decreasing.

Recommended readings:


Polarity From Different Perspectives, http://www.nyu.edu/gsas/dept/lingu/events/polarity/schedule.html
Why does NPI have to be within the scope of a decreasing operator?
Zooming in on Lahiri 1998 and Rooth 1992

- **Lahiri**: Hindi NPI *ek/koii bhii* means “even one/some thing”. Because “one/some” is a minimal amount, its implicatures contradict those arising from *even*. The contradiction can only be resolved in a decreasing (scale reversing) environment.

Conventional implicatures (presuppositions) of *even* (Karttunen and Peters 1979):

Even JOHN saw this
- asserts: John saw this
- implicates: (a) Someone other than John saw this. (existential implicature)
  (b) John is the least likely to have seen this. (scalar implicature)

Rooth 1992, 1995: The focusing of a constituent introduces a set of alternatives into the discourse. The alternatives induced by focusing John in *JOHN saw this* form a set of propositions C, such that

\[ C = \{ \neg John saw this, \neg Mary saw this, \ldots \} \]

Operators that associate with focus (*not, even, only, always(?), usually(?), etc.*) rely on this set C. E.g.,

Even JOHN saw this
- asserts: John saw this (=A)
- implicates: (a) there is a true proposition among the focus-induced alternatives C of A that is not identical to A.
  (b) Among the focus-induced alternatives C, A is the least likely.

Put these together with the NPI, a minimal amount expression in focus and modified by ‘even’:

\[ \neg \text{koi bhii} \text{ aayaa} \]

‘Anyone came’

- asserts: At least one entity came
- alternatives: \{\neg At least 1 entity came, \neg At least 2 entities came, \ldots\}
- implicates: (a) there is a true proposition of the form *At least n entities came*, other than *At least 1 entity came*, and
  (b) of all the alternatives, *At least 1 entity came* is the least likely.

*To be \( \geq 1 \)* is the weakest predicate, true of everything. Likelihood is understood in terms of entailment:

If \( p \) entails \( q \), \( q \) is more likely than \( p \) (because \( P \subseteq Q \)).

For all \( n \), *At least n entities came* entails *At least 1 entity came*, therefore *At least 1 entity came* is the most likely. This contradicts the requirement of *even* that it be the least likely. I.e. the existential implicature (a) contradicts the scalar implicature (b).
What if the sentence has a decreasing operator, e.g. negation, such that even scopes above it?

ok koi bhii nahiiN aayaa
one even not came  'It is not the case that anyone came'

asserts: there isn’t at least one entity that came
alternatives: {^There isn’t at least 1 entity that came, ^There aren’t at least 2 entities ..., ...}
implicates: (a) there is a true proposition of the form There aren’t at least n entities that came, other than There is not at least 1 entity that came,
(b) of the alternatives, There isn’t at least 1 entity that came is the least likely.

Now we have no contradiction. For all n, not(at least 1 came) entails not(at least n came), thus not(at least 1 came) is the least likely. This is what (b) wants.

• Rooth on the association of only and always with focus:

Mary only [vp introduced BILL to Sue].

Alternatives C induced by focus on Bill:
C = {^introduced Bill to Sue, ^introduced Mary to Sue, ^introduced Ken to Sue, ...}

C will serve as the restriction of only, and the property [[introduce Bill to Sue]] as its scope.
only(C)(VP1)(mary) : ∀P[(P ∈ C & P(mary)) → P=VP1]
’For every property P, where P comes from the set C and holds of Mary, P is [[VP1]]’
where [[VP1]] is the ordinary semantic value of VP1 and C is presupposed to be a subset of the focus semantic value of VP1, notated as [[VP1]]f (subset b/c further constrained by contextual factors)

Similarly, Mary always introduced OLD MEN to Sue  ’Always, when Mary introduces someone to Sue, Mary introduces an old man to Sue’. This latter claim is debated by Beaver and Clark 2003.

Recommended:
Beaver & Clark 2003, Always and only: why not all focus sensitive operators are alike. NALS 11.
Starting to Build a Grammar

First try: Predicate logic (Allwood et al. 5.4-5.7, Gamut I, Ch 3 [skip substitution method]):

Syntax:
Lexicon: Cat dp = \{John\}  \hspace{1cm} \text{read: The category dp contains just John}
  Cat pred = \{walks\}  \hspace{1cm} \text{The category pred contains just \textit{walks}}
Rule: If \(\alpha\) \(\in\) Cat dp and \(\beta\) \(\in\) Cat pred, then \(\alpha\beta\) \(\in\) Cat s. \hspace{1cm} \text{A dp followed by a pred make a s}

Semantics:
A model M is \(<D, I>\), where D is a set of individuals (the universe of discourse) and I is an interpretation function from elements of the lexicon to elements of the model. V is the function that specifies how complex expressions are evaluated.

Let \(D = \{j, m, f\}\) and let
\[
\begin{align*}
[\text{John}]^M & = I(\text{John}) = j  \\
[\text{walks}]^M & = I(\text{walks}) = \{m, j\}
\end{align*}
\]
\(V^M(\alpha\beta) = 1 \text{ iff } [\alpha]^M \in [\beta]^M\). The value of \(\alpha\beta\) (which, as we know from above, is a sentence) is 1 iff \(\alpha\)'s denotation is an element of \(\beta\)'s denotation.

How about Everyone walks? Since we are doing predicate logic, we \textbf{enrich} the above as follows:

Syntax:
Lexicon: Cat dp = \{John, x, y, z, ...\}
  Cat pred = \{walks\}
Rules: If \(\alpha\) \(\in\) Cat dp and \(\beta\) \(\in\) Cat pred, then \(\alpha\beta\) \(\in\) Cat s.
  If \(\phi\) \(\in\) Cat s, \(\forall x[\phi]\), \(\exists x[\phi]\) \(\in\) Cat s.

Semantics:
A model M is \(<D, I>\), where D is a set of individuals (the universe of discourse) and I is an interpretation function from elements of the lexicon to elements of the universe. \(g\) is an assignment of values to variables. \(g[x/d]\) is that assignment which differs minimally from \(g\) in that it assigns the individual \(d\) to the variable \(x\). V is the function that specifies how complex expressions are evaluated.

Let \(D = \{j, m, f\}\). \hspace{1cm} \text{Let } g = \{<x, m>, <y, j>, <z, f>\}.
\[
\begin{align*}
[\text{John}]^{M_g} & = I(\text{John}) = j  \\
[x]^{M_g} & = g(x) = m  \\
[\text{walks}]^{M_g} & = I(\text{walks}) = \{m, j\}
\end{align*}
\]
\(V^{M_g}(\alpha\beta) = 1 \text{ iff } [\alpha]^{M_g} \in [\beta]^{M_g}\).
\[ V^M_g(\forall x[\phi]) = 1 \text{ iff for every } d \in D, \ V^{M^x_d}[\phi] = 1. \]
\[ V^M_g(\exists x[\phi]) = 1 \text{ iff for some } d \in D, \ V^{M^x_d}[\phi] = 1. \]

Predicate logic treats quantifiers syntagmatically, i.e. as diacritics. It doesn’t assign them to a category. It just tells you how they occur with expressions that belong to some category. A compositional semantics of English cannot be built like this.

Reminders:

Free versus bound variables:

- \( \text{big}(x) \neq \text{big}(y) \)  
  Why?

Pointing with free variables:

This is big 'the entity I am pointing to (with my left hand) is big'  
= \( \text{big}(x) \)

That is big 'the entity I am pointing to (with my right hand) is big'  
= \( \text{big}(y) \)

The free variables \( x \) and \( y \) may accidentally point to the same object, but in general, you cannot assume that they do so. Free variables have their own individual identity.

- \( \forall x[\text{big}(x)] \equiv \forall y[\text{big}(y)] \)  
  Why?

Making general statements with quantifiers using bound variables:

Everything is big 'Whatever I point to, that entity is big'  
= \( \forall x[\text{big}(x)] \) or \( \forall y[\text{big}(y)] \)

Something is big 'I can point to an entity that is big'  
= \( \exists x[\text{big}(x)] \) or \( \exists y[\text{big}(y)] \)

A bound variable is a mere placeholder. The sentence says nothing specifically about \( x \) or \( y \).

Duals:

- \( \exists \) and \( \forall \) are duals.  
  Why?

Imagine a universe with 3 elements: \( a, b, c \). Then:

\( \exists x[\text{big}(x)] = \text{big}(a) \lor \text{big}(b) \lor \text{big}(c) \)
\( \forall x[\text{big}(x)] = \text{big}(a) \land \text{big}(b) \land \text{big}(c) \)

The propositional connectives \( \lor \) and \( \land \) are duals -- see the de Morgan laws. Then \( \exists \), which is disjunction over the universe and \( \forall \), which is conjunction over the universe are necessarily duals, too.

\( \neg \forall x[\text{big}(x)] = \exists x \neg [\text{big}(x)] \)  
and  
\( \neg \exists x[\text{big}(x)] = \forall x \neg [\text{big}(x)] \)

because

\( \neg (\text{big}(a) \land \text{big}(b) \land \text{big}(c)) = \neg \text{big}(a) \lor \neg \text{big}(b) \lor \neg \text{big}(c) \)  
and  
\( \neg (\text{big}(a) \lor \text{big}(b) \lor \text{big}(c)) = \neg \text{big}(a) \land \neg \text{big}(b) \land \neg \text{big}(c) \)
Assignment (due Nov.2)

Read Gamut 1, Ch 3, focusing on the discussion of models and assignments. SKIP the substitution method section. – Allwood et al. is good too, but their terminology is different from what Gamut and the handouts use, and the differences may just be confusing, so I recommend Gamut.

(1) Gamut, ch. 3, ex. 8: Work out 7bi, iii, and v using the assignment method. Please compare the results with the book’s solutions, and alert me if you aren’t sure about the discrepancies. (In previous Gamut assignments I wasn’t always sure if all of you did this comparison... In this case it is especially important that you consult the book’s solutions along the way, because that’s how you’ll learn how to write these things out. Don’t just cast a glimpse at the solutions after you are done with all three cases.)

(2) Which of these is true in all models? Briefly explain why.
   a. \( \forall x [f(x) \& g(y)] = \forall z [f(z) \& g(y)] \)
   b. \( \forall x [f(x)] \& g(x) = \forall y [f(y)] \& g(x) \)
   c. \( \forall x [f(x)] \& g(x) = \forall y [f(y)] \& g(y) \)
   d. \( \forall x [f(x)] \& g(x) = \forall y [f(y)] \& g(y) \)
   e. \( g(x) \& \forall y [f(y)] = g(x) \& \forall x [f(x)] \)

(3) Which of these is true in all models? Why?
   a. \( \forall x \neg [f(x)] = \neg \exists x [f(x)] \)
   b. \( \forall x \neg [f(x)] = \exists x \neg [f(x)] \)
   c. \( \forall y [f(y)] = \neg \exists y \neg [f(y)] \)
   d. \( \neg \forall x [f(x)] = \neg \exists x \neg [f(x)] \)
   e. \( \neg \forall x [f(x)] = \neg \exists x [f(x)] \)
   f. \( \neg \forall y [f(y)] = \neg \exists y \neg [f(y)] \)
   h. \( \exists x \neg [a(x) \lor b(x)] = \exists x [\neg a(x) \lor \neg b(x)] \)

(4) The Gamut sentences you already translated into predicate logic included (ones like) these:
   a. If John owns a dog, he has not shown it to Bill.
   b. If you see anyone, you should give no letter to her.
   c. If someone is noisy, everybody is annoyed at him.
   d. If Pedro owns donkeys, he beats them.
   e. Everyone who owns a car parks it on the street.

These sentences all have a peculiarity: the translations that Gamut proposes capture the truth conditions but do not have existential quantifiers for the indefinites a dog, anyone, someone, etc. At first blush, there are even two distinct translations using existential quantifiers for these indefinites that could be used. But they would not be truth-conditionally correct. (a) Explain what the problem is in each case. (b) Can you link some of these issues to the syntax of these sentences? Are there other sentences involving indefinites, not cited above, that would raise similar problems? (c) If you can, please comment on the fact that in the specific examples above it is universal quantification that can be used to circumvent the problems in (b). You may think about truth conditions, compositionality, perhaps other things. – Several recent theories in semantics originated with this type of sentences (dynamic semantics, E-type pronouns, etc.). I am not expecting you to write dissertations as part of this assignment. But please identify the nature of the problem very clearly.
Predicate logic treats quantifiers syntagmatically, i.e. as diacritics. It doesn’t assign them to a category. It just tells you how they occur with expressions that belong to some category. A compositional semantics of English cannot be built like this. In the second try, we remedy this and also introduce a recursive structure into categories and into the universe of the model. For example, we do not use a primitive category “pred”. It will be replaced by e\t, which is the category of things that look for an expression of category e on the left and together with it form an expression of category t. Similarly for predicate denotations in the model: they will be interpreted as functions that apply to the denotations of expressions of category e (individual entities) and return the value of expressions of category t (True or False).

**Categorial grammar has these two main ideas:**

1. Syntactic categories are nothing but explicit encodings of what combinations are possible. E.g. to say that *walks* belongs to category e\t is to say that it can combine with a member of category e that occurs to its left (its specifier) and the result will be a member of category t. On this approach *walks* cannot belong to the same category as *knows*, since the first takes one and the second two arguments.

2. Syntactic categories are mapped to semantic types in a trivially simple way. An expression of category e\t is a function of type <e,t>, one that applies to an entity-type thing and gives a truthvalue-type thing (true or false) as a value. This makes a strong commitment to syntax and semantics being parallel.

**Syntax:**

Categories: e, t ∈ Cat;
if A and B ∈ Cat, A/B and B\A ∈ Cat;
nothing else is in Cat.

Lexicon: Cat e = {John1, Mary1}
Cat t = ∅
Cat e\t = {walks}
Cat ((e\t)/e) = {knows}
Cat (t/(e\t)) = {John2, Mary2, everyone, something}

Rules: If α ∈ Cat A/B and β ∈ Cat B, then αβ ∈ Cat A. [A/B looks for a B on the right]
If β ∈ Cat B and α ∈ Cat B\A, then βα ∈ Cat A. [B\A looks for a B on the left]

**Semantics:**

The universe is now typed. Types are “categories of things in the model”. The tangible members of the universe of discourse are individuals, but functions of all sorts can be defined based on individuals and truth values. E.g., functions of type <e,e> (from entities to entities) or <e,t> (from entities to truth
values. Using categories and types is very natural in linguistics, since words do not combine
arbitrarily: they combine with particular kind of other words to yield particular kind of phrases. In logic,
the use of types is one way to get out of the Russell paradox.

A model M is &lt;D,I&gt;, where D is a set of individuals. I is an interpretation function from expressions to
things in the universe. (As before.)

The Dom(ain) of a type with respect to a universe D is the set of things in D that belong to that type:
entities, or truth values, or various kinds of functions.

Types: e,t ∈ Type;

Dom_e,D = D,
Dom_t,D = {0,1}

if B, A ∈ Type, &lt;B,A&gt; ∈ Type;

Dom&lt;B,A&gt;,D = {f: Dom_B,D → Dom_A,D}

nothing else is in Type.

Syntactic categories are related to semantic types by the following simple map f. Expressions of
category e are mapped to things of category e(ntity). Expressions of category t are mapped to things of
category (r)uthvalue). Function categories are mapped to appropriate functions. Syntactic directionality
becomes irrelevant.

f(e)=e
f(t)=t
f(A/B) = f(B\A) = &lt;f(B),f(A)&gt;

Let D = {j, m, f} and let I work as follows:

[[John1]]M = j
[[Mary1]]M = m
[[walks]]M = that function from individuals to truthvalues which assigns 1 to an individual iff it
walks, viz. the characteristic function of the set of walkers (in M, let it be {m,j})
[[knows]]M = that function from individuals to [a function from individuals to truthvalues] which
assigns 1 to a pair of individuals if one knows the other
[[John2]]M = that function from sets of individuals to truthvalues which assigns 1 to a set P iff j ∈ P,
viz. the characteristic function of the set of those sets that contain j
[[Mary2]]M = that function from sets of individuals to truthvalues which assigns 1 to a set P iff m ∈ P,
viz. the characteristic function of the set of those sets that contain m
[[everyone]]M = that function from sets of individuals to truthvalues which assigns 1 to a set P iff
human ⊆ P, viz. the char. function of the set of those sets that contain everyone
[[something]]M = that function from sets of individuals to truthvalues which assigns 1 to a set P iff
thing ∩ P ≠ Ø, viz. the char. function of the set of those sets that contain something

Merge is interpreted as functional application; who applies to whom depends on who is the function:

[[αβ]]M = [[α]]M([[β]]M) or [[β]]M([[α]]M)
**Everyone** and **something** now belong to a proper category and they receive exactly the kind of interpretations that the theory of generalized quantifiers assigned to them. Notice that we are preserving all the results that we had when studying GQs. We are not revising anything; we are merely building a grammar that accommodates those results.

When a function $\alpha$ is a characteristic function of some set, the application of $\alpha$ to some argument $\beta$ is equivalent to saying that $\beta \in \alpha$. That is, in our model $[[\text{walk}]](\text{[[John]]}) = j \in \{m,j\}$. Using this shortcut:

\[
\begin{align*}
\text{John1 walks, t} & \quad & \text{everyone walks, t} \\
\text{\hspace{1cm} j} & \in \{m,j\} & \{\text{\hspace{1cm} m,j} \} & \in \{\text{\hspace{1cm} P: human } \subseteq \text{P} \}
\end{align*}
\]

\[
\begin{align*}
\text{John1, e} & \quad & \text{everyone, t(e} & \text{t)} \\
\text{\hspace{1cm} \{m,j\}} & \quad & \text{\hspace{1cm} \{m,j\}} & \text{\hspace{1cm} \{m,j\}}
\end{align*}
\]

Alternatively, **John walks** might be built in analogy to **Everyone walks**, using **John2**. **John walks** will come out interpreted the same both ways, in one derivation with **John1** as the argument and **walks** the function, and in the other, with **John2** as the function and **walks** the argument.

\[
\begin{align*}
\text{John2 walks, t} & \quad & \text{everyone, t(e} & \text{t)} \\
\text{\hspace{1cm} \{m,j\} & \in \{\text{\hspace{1cm} P: j} & \in \text{P} \} & \in \{m,j\} & \in \{m,j\} = 1 \\
\text{\hspace{1cm} j} & \in \{m,j\} & \{\text{\hspace{1cm} P: human } \subseteq \text{P} \} & \{m,j\}
\end{align*}
\]

We have just implemented the claim that all noun phrases can be said to denote generalized quantifiers (cf. Keenan). **Everyone** denotes the set of properties that every individual entity has (shares), and **John** denotes the set of properties that the individual entity called John has. The denotation of **John** can also be “simplified” to the individual entity called John, but the denotation of **everyone** cannot be “simplified” to any individual entity. Similarly for **few kids**, **more than five kinds**, etc. So generalized quantifiers are the “least common multiple” for noun phrase denotations.

**John1** is useful in explaining the possibility of coreference between the name and a singular pronoun. **John2** is useful in explaining the possibility of coordinations like **John and every plumber**.

In our grammar, the interpretations of syntactic expressions are given directly, without invoking a level of logical form. E.g. **everyone** is directly correlated with a function such that... It is of utmost importance to see that this can be done. At the same time, it is very convenient to take advantage of a simple and transparent notation like $\{\text{P: human } \subseteq \text{P} \}$ when doing our calculations. Therefore, we might want to translate English expressions first into some convenient formal language, letting its interpretation serve as the interpretation for English. This is what Montague did in UG and PTQ. But the set theoretic notation we have employed so far is only manageable in some cases. What we need is a comparably simple and flexible all-purpose functional notation. The **lambda calculus** offers one. This is what we introduce next.
Appendix: Some further specifics

(a) We said that the determiner *every* denotes a relation between two sets of individuals. But then we said that those two sets are not on a par: the np-set is "closer" in that it serves as a restriction. Taken together with the fact that *every* first combines with np in syntax, it is useful to look at *every* as denoting a two-place function: first argument a np-denotation, second argument a predicate denotation, value 0 or 1. How about *every boy*? By "Currying" the two-place function, we get that *every* denotes a one-place function from np-denotations to one-place functions from predicate denotations to truth values. *Every boy* then denotes a one-place function from predicate denotations to truth values, and *Every boy walks* a truth value.

(b) Like *walks*, *boy* is interpreted as a function of type <e,t> (the characteristic function of the set of boys), but *boy* does not combine directly with subjects in English syntax. This can be fixed by adding an "inert" category, call it t^e^, also mapped to type <e,t>, which is input only to syntactic rules where the category t^e^ is an argument of some other function. There will be no rule that allows t^e^ to directly combine with e either from the left or from the right.

(c) Categories other than t and e are mapped to function types. We learn this from the definition of the denotation domains of the types: Dom_{A,B,D} = {f: Dom_{A,D} → Dom_{B,D}}, viz. <A,B> is the type of functions from A-type things to B-type things (relative to the universe of discourse D). Although the category-type mapping is conceptually trivial, the notation is not. You need to be aware of alternative notations to read the literature.

In Montague’s PTQ, the category of *walks* is notated as t/e, and the (extensional) type of *walks* as <e,t>. Montague’s notation is not meant to encode left-right in the categories, and in mapping categories to types you simply invert the order: (value-cat / argument-cat) maps to <argument-type, value-type>. In subsequent work the notation of types remains the same, but the notation of categories varies a lot.

In categories, e is sometimes replaced by n or np and t by s, viz. *walks* is (s/n) or (s/np). e and t continue to be used for types.

Lambek and his followers use a directional notation where *walks* is e\t and *knows* is (e\t)/e, to encode that the direct object is expected from the right and the subject from the left. This notation is mathematically sound, given the multiplication analogy value-cat ≈ numerator and argument-cat ≈ denominator and e.e\t=t and t/e.e=t. To perform the category-type mapping, you need to find the ultimate value category above the main slash and the corresponding argument category under the main slash, etc. This handout uses the Lambek notation to facilitate your reading of that literature.

Another notation is used by Mark Steedman and in my own earlier papers. Here only the leaning of the slash, but not the placement of the argument category, indicates whether the argument is expected from the left or the right. Thus, Lambeck’s ((subj\t)/obj) is Steedman’s ((t/subj)/obj). This has a perceptual advantage with multi-layered categories: it is easy to see at once who is the argument and who is the value. But in all cases, the type is written as <object, <subject, t>>.
Recommended reading regarding the indefinites problem you looked at in the last assignment is Gamut Vol. 2, ch. 7.4 (Discourse Representation Theory).

**Assignment (due Nov. 9)**

Read Ch 4 of Gamut II (The theory of types and categorial grammar) UP TO categorial grammar (pp. 75-92). [The categorial grammar part of this chapter is really not good. It is outdated, and I think their (16) and (21) have typos or are otherwise confusing. It's better if you skip it, since it'd take more time to clear up confusions than what it's worth. -- You'll read 4.4 on lambdas next week.] Notice that the book continues to give a syncategorematic treatment of quantification here, because it formalizes a logical language. We do not do that in our toy English grammar but this does not mean that what the book does is wrong in its own right.

Please start reading the project related stuff.

- Do ex 1 and check the back for solutions, as usual.

- Devise derivations, in the style of the handout for *John walks* and *Every boy walks*, for the following examples. You will have to reason out the categories of some of the words – do that in the backward engineering fashion we used to figure out the category of *everyone* in class. Work on the examples in the order given, use your old results in the new example (without repeating the detailed reasoning that led to it), and explain your new decisions at each step. Bear in mind that category assignment must make semantic sense (type theoretical, argument structure, etc.) ... whenever you want to assign something to category e, please ask yourself if that expression denotes an entity.

  In the examples with *and* and *or*, use the flat, simplified structure to stand for the more correct binary branching one, with $\alpha$ the category of the conjuncts/disjuncts:

  \[
  \begin{array}{c}
  \text{John and Mary} \\
  \alpha \\
  \alpha (\alpha/\alpha/\alpha) \\
  \alpha
  \end{array}
  \]

  Your main task is to determine what the categories are. **Please do not assume any empty elements in the sentence! You must work with just the words that you see.**

  [1] Kim sneezed. (take inflected verbs to be primitives)
An all-purpose function notation: Lambdas

Lambdas are a practical syntactic device for (i) extending the use of sentential connectives to subsentential cases, (ii) assigning explicit interpretations to arbitrary coherent parts of sentences, as required by compositionality, and (iii) defining arbitrary functional operations, e.g. type lifting.

(i) run and sing
John and Bill

(ii) every man

(iii) type lifting

Re (i). Step 1: To make the use of the sentential connective & legitimate, pad out * run' & sing' with variables: run'(x) & sing'(x). This amounts to saying that if run' and sing' had the argument x, their conjunction would be run'(x) & sing'(x). Step 2: But they do not in fact have that argument (run and sing is not a sentence). The assumption of x must be "withdrawn". This is indicated by the prefix λx.

Re (ii). Step 1: Every man walks would be represented as ∀x[man'(x) → run'(x)]. Everything in this formula, save for run', is the contribution of every man. Hence, to represent every man, get rid of run' by replacing it with a predicate variable P. Step 2: To withdraw the assumption of P, prefix λP.

Definitions:

Syntax: If α is a well-formed expression and x a variable, λx[α] is a well-formed expression.

Semantics: λx[α] denotes a function. When this function is applied to some b of the same type as x, the function value is computed by replacing every occurrence of x in α with b. This replacement process is called lambda conversion.

E.g, λx[x^2] denotes that function which assigns each number its square: λx[x^2](3)=3^2.

Two important qualifications

One: Only those x's in α can be replaced that are bound by the initial lambda operator. In λx[f(x) & ∀x[h(x)]], only the x in f(x) can be replaced; the one in h(x) is bound by the universal. Note: λx[f(x) & ∀x[h(x)]] ≠ λx[f(x) & ∀y[h(y)]]. To prevent mis-applications, it may be useful to reletter the "homonymous" variables that are bound by an operator other than the lambda (you don’t have to, if you are sure you are careful enough – in contrast to Case Two, below).

Two: If the argument to which the lambda-defined function applies is described by an expression that is, or contains, a free variable, care must be taken to ensure that this variable remains free in the course of computing the function value. Suppose λx[∀y[f(x) → h(y)]] is applied to the argument y, a free variable. Let the current assignment g of values to variables have g(y)=bill. Then, λx[∀y[f(x) → h(y)]](y) must be the same as λx[∀y[f(x) → h(y)]](bill), i.e. ∀y[f(bill) → h(y)]]. If we had mechanically replaced x by y, we would have gotten ∀y[f(y) → h(y)], which is an entirely different thing. To prevent misapplications, we must reletter those bound variables in α that happen to be "homonymous" with the free variable in the argument. That is, λx[∀y[f(x) → h(y)]] is not applied to y. It is first relettered as λx[∀z[f(z) → h(z)]]. Note that a free variable can never be relettered.
Since the lambda operator is an all-purpose device for defining functions, any functions, there are no restrictions on what its domain and co-domain might be. Above, we assumed for simplicity that $x$ was an individual variable, but in fact it might be a variable over any domain. Likewise, the $\alpha$ in $\lambda x[\alpha]$ may be anything: (a) a truth value, (b) a function, (c) an individual that varies with $x$, (d) a fixed object; etc.

(a) $\lambda x[\text{run}'(x)]$ is the characteristic function of the set of runners.
\[
\lambda x[\text{run}'(x)] = \text{run}' \quad \text{because for every argument } b, \quad \lambda x[\text{run}'(x)](b) = \text{run}'(b)
\]
A characteristic function is a function from some $D$ to $\{0,1\}$. It "characterizes" some subset $C$ of $D$ by assigning 1 to $d$'s that are also in $C$, and 0 to those that are not.

(b) $\lambda x[\lambda y[\text{employ}'(x)(y)]]$ is a function from potential employees $x$ to VP-denotations, viz., functions from potential employers $y$ to $\{0,1\}$. Not a characteristic function.

(c) $\lambda x[\text{mother-of}'(x)]$ is a function from individuals to their mothers.

(d) $\lambda x[\bullet]$ is a constant function that maps everything to $\bullet$.

**Vital conventions on representing argument order:**

The notation $f(a)(b)$ is short for $(f(a))(b)$. The function $f$ is first applied to $a$ and then to $b$. (This is called left-associativity.) The order of the lambda-prefixes represents the inviolable order of how the function can be applied to arguments:

\[
\lambda x \ [ \lambda y \ [f(x)(y)] \] \ (a \ b) = \lambda y[f(a)(y)](b) = f(a)(b) \quad \text{and never } f(b)(a)!
\]

Naturally, explicit bracketing may indicate that things are otherwise:

\[
\lambda x \ [ \lambda y[f(x)(y)](a) \ ] \ (b) = \lambda y[f(b)(y)](a) = f(b)(a)
\]

Different ways of writing the same thing:

\[
\lambda x[\lambda y[f(x)(y)]] = \lambda x.\lambda y[f(x)(y)] = \lambda x.\lambda y.f(x)(y) = \lambda x y.f(y,x)
\]

Heim--Kratzer follow the "dot convention" to indicate the scope of the lambda operator. Moreover, they always explicitly indicate the domain of the function:

\[
\lambda x : x \in D. f(x) \quad \text{or, for short, } \lambda x \in D. f(x)
\]

**How to compute the type of a lambda-expression:**

\[
\begin{array}{c}
\lambda x_\epsilon \ [ \lambda P_{\epsilon,t} \ [ P_{\epsilon,t}(x_\epsilon) ] ]_t \\
\downarrow \downarrow \downarrow \\
\epsilon, \ldots, <\epsilon, t>, t >
\end{array}
\]

\[
< \text{arg1}, < \text{arg2}, \text{value} > >
\]
Optional: Can a function be applied to any arbitrary function, including itself?

Some functions applied to themselves yield a perfectly well-behaved and useful new function. But some other self-applications allow us to replicate the Russell paradox. The set theoretic version of the Russell paradox goes as follows:

Assume that for every property P, there is a set \( \{x : Px\} \).

Now ask the reasonable question: What sets are elements of themselves?

Both \( \{x : x=x\} \in \{x : x=x\} \) and \( \emptyset \notin \emptyset \) would make good sense.

But let \( R \) be \( \{x : x \notin x\} \), i.e. the set of those things that are not elements of themselves.

Is \( R \) an element of \( R \)? If yes, ..., if no, ... Paradox.

\( R \) is an element of itself iff it is not an element of itself.

The functional version replicates this (from Curry—Feys 1958):

Let \( N \) be negation. Let us define a function \( Y \) such that \( Y(Y) \) is \( N(Y(Y)) \).

That is, applying \( Y \) to itself is the same as the negation of applying \( Y \) to itself. \( Y \) is called the paradoxical combinator. How to define a \( Y \) that behaves in this paradoxical way?

Let \( Y \) be \( W(B(N)) \), with \( W(h)(z) = h(z)(z) \) and \( B(f)(g)(x) = f(g(x)) \). This will do, because \( Y(Y) = W(B(N))(Y) = (W(B(N))(Y))(Y) = N(Y(Y)) \).

Written with lambdas, let \( W = \lambda h \lambda z[h(z)(z)] \) and \( B = \lambda f \lambda g \lambda x[f(g(x))] \). Then

\[
Y(Y) = W(B(N))(Y) = \lambda h \lambda z[h(z)(z)](\lambda f \lambda g \lambda x[f(g(x))])(N)(Y) = \\
\lambda h \lambda z[h(z)(z)](\lambda g \lambda x[N(g(x))])(Y) = \\
\lambda g \lambda x[N(g(x))](Y)(Y) = N(Y(Y))
\]

To exclude this paradox, things in the model are placed at different levels (sorted into types). The element-of relation is not defined for things of the same level/type. Similarly for functional application.

Using types is doing thought-policing: asking some questions in banned. But type-free systems are both mathematically and philosophically important, plus they have their own good uses in defining new operations/programs, for which purposes they are retained and paradoxes are dealt with otherwise. On the other hand, it is natural to use typing in connection with natural language, since syntactic categories delimit the abilities of expressions to combine: *determiner_verb, etc.

In a typed system, self-application is excluded by the broader constraint that a function is unable to apply to things that have the same type (i.e. not only to things identical to it). Suppose we want to have a well-formed \( G(G) \). Suppose \( G \) is of type \(<a,b>\). In order for a function to apply to such a \( G \), the function ought to have the type \(<<a,b>,c>\); but \( G \)'s type does not have that format. Or, conversely, in order for \( G \) to apply to an argument, the argument ought to have the type \( a \); but \( G \)'s type does not have that format. It is not possible to define a type of functions that take functions of the same type as argument.
Using Lambdas
Reading: Gamut II, Ch. 4.4

(1) Is any of these (i) ill-formed, (ii) false? Pay attention to free and bound variables; you can assume that types match properly.

(a) \( \lambda x[f(x)] (b) = f(b) \)
(b) \( \lambda x[f(x)] = f \)
(c) \( \lambda a\forall x[a(x)] (\lambda y[f(y)(z)]) = \forall x[f(x)(z)] \)
(d) \( \lambda a\forall x[a(x)] (\lambda y[f(y)(x)]) = \forall x[f(x)(x)] \)
(e) \( \lambda a\lambda x[a(x)] (\lambda y[f(y)(z)]) = \lambda x[f(x)(z)] \)
(f) \( \lambda a\lambda x[a(x)] (\lambda y[f(y)(x)]) = \lambda x[f(x)(x)] \)

(2) Reduce the following to their simplest form by performing the lambda conversions (i). You can assume that types match properly. Say in each case what a linguistic application might be. Explain the application as in the sample, saying (ii) what linguistic expressions the unreduced \( \lambda \)-function and its arguments might each correspond to, and (iii) what the result of reduction (\( \lambda \)-conversion) would correspond to and, if applicable, (iv) why what we have seen is good for us linguists. -- Your explanations need not be this verbose, but please give all of (i)-(ii)-(iii)-(iv), not only the result of the reduction.

(a) Sample solution:
   (i) \( \lambda P[P(m) & P(o)](g) = g(m) & g(o) \)
   (ii) \( \lambda P[P(m) & P(o)] \) might be the translation of Mickey and Oscar;
        \( g \) might be the translation of giggle.
   (iii) The result can be a formalization of `Mickey giggles and Oscar giggles'.
   (iv) We are able to assign an interpretation to conjoined proper names (which cannot be
        conjoined by & directly) that combines with a single instance of the predicate and the
        result is semantically correct in that the predicate gets distributed over the conjuncts.

(b) \( \lambda A \lambda B \lambda R[A(R) & B(R)] (\lambda P[P(m)]) (\lambda P[P(o)]) \)
   (i) \( \lambda A \lambda B \lambda R[A(R) & B(R)] (\lambda P[P(m)]) (\lambda P[P(o)]) = \)
   (ii) \( \lambda P[P(m)] \) might be the translation of ...
        \( \lambda A \lambda B \lambda R[A(R) & B(R)] \) might be the translation of ...
   (iii) The result can be a formalization of ...
   (iv) ...

(c) \( \lambda A \lambda B \lambda R[A(R) & B(R)] (\lambda P[P(m)]) (\lambda P \forall x[F(x) \rightarrow P(x)]) \)
(d) \( \lambda C \lambda D \forall x[C(x) \rightarrow D(x)] (f) (h) \)
(3) What would be the \( \lambda \)-expression formalizing the contribution of **not** (i) as a sentential operator, (ii) as a verb phrase operator?

(4) The following functions differ only in argument order: ACT: \( \lambda x \lambda y[f(x)(y)] \) and PASS: \( \lambda x \lambda y[f(y)(x)] \). Define a function that applies to ACT and turns it into PASS.

(5) Montague (1972 = PTQ) proposed a translation of “the transitive verb be”, which Partee (1987) reincarnated as the type shifter BE. It looks like this:

\[
{\text{be}}: \lambda \phi \lambda x[\phi(\lambda y[x=y])] \quad \text{where } \phi \text{ is of type } \langle\langle e,t,t\rangle>.
\]

To find out what it is doing, use it the way Montague did, to derive *John is a man*. Supply the \( \lambda \)-expressions for all 3 items and perform the conversions. Then explain the operation of be.

\[
{\text{be(a man)}}(\text{John}) =
\]
A fishing expedition is not field work: 
Do you agree that there’s something wrong with “He became no Einstein”? 

Field work should be a well-designed experiment. Some traps to avoid: 

(1) Make sure the speaker is judging the sentence and the intonation pattern you had in mind. If at all possible, have the speaker pronounce the sentence, or repeat the sentence for you. You’ll be surprised how often the speaker is judging something other than what you want. 
John and Mary loves icecream – the speaker may accept it but repeat it back with love. 
I didn’t study math and flute – you’ll never find out that the ‘not both’ reading needs stress on and unless you have the speaker pronounce it for you. 
I gave him it – okay with emphasis on it but don’t expect the speaker to figure that out. 

(2) Don’t expect the speaker to rise above the traps of lexical semantics, pragmatics, world knowledge, etc. that you think are irrelevant. 
I gave Mary herself – doesn’t make much sense, use showed. 
What do you wonder who to talk about to? – make sure a person is being talked to. 
I kiss a table – the speaker will focus on trying to imagine why you might want to say such a thing, not whether it is grammatical. 
Every hunter killed a deer – don’t ask for object wide scope. 
I have at most five brothers – will only be natural if the situation involves genuine uncertainty. 

(3) Work hard to create “good examples” but make sure that in doing so you are not creating a separate effect (a new empirical phenomenon). 
He has left yet is unacceptable on its own, but acceptable within I don’t think that he has left yet. 
Mary drinks a beer is unacceptable on its own, but acceptable within When she gets home / In the evenings Mary drinks a beer. 
There were five kids intelligent is unacceptable, but changing the predicate to sick makes it perfect. 

(4) Don’t ask metalinguistic questions: 
What does this sentence mean? 
Is this sentence ambiguous? 
Give me all the meanings of this sentence. 
Is “a man” specific or generic in this sentence? 
It is not the speaker’s job to give you an analysis. If you have a speaker who is a linguist, it is okay to discuss the data with them, but this is professional consultation, not field work.
(5) Don’t betray your expectations.  
_This sentence is unacceptable, isn’t it?_

(6) It is okay to use linguists as speakers – though most of the time you have to be as careful with them as with non-linguists. But they may be able to alert you to a glitch in your data, which is especially helpful if you are not a speaker yourself and you do not have an opportunity to conduct very extensive field work.

(7) Make sure your social status, less than perfect knowledge of the language, etc. don’t influence the judgment.  
_Can I say “....”?_ – field worker often gets Yes, but if (s)he asks _Would you say “....”?_ (s)he gets a straight No.

(8) Avoid satiating / tiring out your speaker. Don’t prime speaker, mix the critical examples with irrelevant ones (=fillers), don’t repeat the same pattern too much.

(9) Beware of the comparison set. Whether you get an OK or ? or * judgment depends greatly on what other sentences or readings the speaker is comparing your datum with (because you asked for a comparison, or simply because the data were presented in a sequence – your fillers may be responsible for the result). Read Clifton—Fanselow—Frasier in LI 37/1 (2006).

Some good practices:

(10) Establish a base line by having the speaker judge related but uncontroversial data. If your work relates to that of others, replicate their judgments before changing the conditions or questions. Read Feynman on cargo cult science, [http://wwwcdf.pd.infn.it/~loreti/science.html](http://wwwcdf.pd.infn.it/~loreti/science.html)

(11) Form a clear hypothesis about what your phenomenon is and exactly what causes it. Devise examples that don’t sound like IQ tests and are pragmatically credible. Make sure the test sentences test just what you are interested in. Present them in such a way that speaker can give you a natural reaction, not a linguistic analysis.

[a] Clarify the exact truth conditions. Create models in which the sentence is true and ones in which it is false. Make the models as simple as possible and make sure they are not contaminated by unspoken assumptions.

[b] If the sentence is ambiguous, check whether one reading entails the other. If there is an entailment, it is difficult to tell apart which reading you obtain, so you have to use special techniques, like “let the inverse reading shine” or the “pronominal anaphora test”.

[c] Psycholinguists who are good at behavioral experiments design tasks where the speaker performs an entirely natural linguistic or non-liguistic activity, rather than a metalinguistic one. Read papers from Anthony Sanford’s lab for good examples; also “What is a linguistic fact?” by Labov.

(12) When your fellow linguists ask you for judgments, be prepared to be patient and careful. If they are trying to do even some of the above things right, the operation will rarely take a minute. Ask to talk to them a bit later if you are not ready to spend some time on their project at the moment.
Quantification: The Logical Syntax of Generalized Quantifiers and Scope

At the outset, we asked what the interpretations of noun phrases like every man and few men might be. We know that the interpretation of Every man walks can be explicated as

\[(1) \ \forall x[\text{man}(x) \rightarrow \text{walk}(x)]\]

but the contribution of every man is not a contiguous subpart of (1). Using backwards engineering, we conclude that everything except walk in (1) is the contribution of every man. The lambda calculus then allows us to abstract away from walk and capture just what we need:

\[(2) \ \lambda P\forall x[\text{man}(x) \rightarrow P(x)]\]

The denotation of the expression in (2), the set of properties shared by every man, is a generalized quantifier (GQ). All non-pronominal noun phrases can be correctly described as denoting generalized quantifiers (or as semantic objects derived from those). Some, like John, can also be attributed simpler denotations (individuals of type e), but most noun phrases cannot.

John as an expression of category t/(e\text{t}) denotes a generalized quantifier, i.e. the set of properties John has (=the set of those sets that contain the individual John):

\[(3) \ \lambda P[J] = \{\text{is human, is angry at Mary, is identical to John, } \ldots\}\]

Every man denotes the set of properties that are shared by every man:

\[(4) \ \lambda P\forall x[\text{man}(x) \rightarrow P(x)] = \{\text{is human, is angry at Mary, is identical to John or Bill or Joe }, \ldots\}\]

Some man denotes the set of those properties that are exhibited by at least one man:

\[(5) \ \lambda P\exists x[\text{man}(x) \land P(x)] = \{\text{is human, is tall, is angry at Mary, walks, does not walk, is identical to John or Bill or Joe or Frank, } \ldots\}\]

NB In contrast to (3) and (4), (5) may contain “contradictory" properties. While John or every man cannot both be tall and not tall, there may be some man who is tall and some (other) man who is not.

Everything we say about the formalization of quantification and scope below is built on the notion that noun phrases denote generalized quantifiers.

To start with, notice that not all properties that an individual has need to be as simple as the ones listed above. Equally good is the property of being a dog that bit every man, or the property of being a person whom some dog bit. This observation is the foundation of the treatment of quantifier scope.
Take a dog bit every man. In present terms, the subject a dog taking scope over the object every man amounts to saying that the set of properties that at least one dog has includes the property of having bit every man. This property is written as \( \lambda z \forall x[M(x) \rightarrow B(z,x)] \).

Conversely, the object every man taking scope over the subject a dog amounts to saying that the set of properties shared by every man includes the property of being an individual such that there is some dog that bit that individual. This property is written as \( \lambda z \exists y[D(y) \& B(y,z)] \).

Finally, the subject scoping over negation in Some dog doesn’t growl amounts to saying that the set of properties that at least one dog has includes being an individual who does not growl, \( \lambda y \neg [G(y)] \).

(6) \( \lambda Q \exists y[D(y) \& Q(y)](\lambda z \forall x[M(x) \rightarrow B(z,x)]) \)  \( \text{subject existential > object universal} \)

(7) \( \lambda Q \forall x[M(x) \rightarrow Q(x)](\lambda z \exists y[D(y) \& B(y,z)]) \)  \( \text{object universal > subject existential} \)

(8) \( \lambda Q \exists x[D(x) \& Q(x)](\lambda y \neg [G(y)]) \)  \( \text{subject existential > negated predicate} \)

These observations suggest a way to derive these readings in a grammar. We effectively defined the semantic contributions of the following constituents. Generalization: the highest operator has the widest scope. (Notice how this meshes with the idea that the scope of an operator is its c-command domain!)

(9) some dog bit every man  \( S \rightarrow O \)

(10) some dog bit every man  \( O \rightarrow S \)

(11) some dog doesn’t growl  \( S \rightarrow \text{Neg} \)

It is of utmost importance to see that this is the common core of all Fregean treatments of quantification (Fregean in the sense that we introduce one quantifier at a time). In all such theories the introduction of a quantifier is interpreted as saying that a simple or complex property is an element of the GQ denoted by the quantifier. The definition of a complex property may involve another quantifier.

This exposition relies on the idea that not only bit every man but also some dog bit are syntactically and semantically coherent segments. We have independent justification for this from extraction facts. The formation and interpretation of such segments is possible with or without invoking traces (the latter is an option in a more powerful version of categorial grammar).
One way to execute this idea is (an extensional version of) what Montague (1972) did. We start out by applying the verb to placeholders (variables of type e) as arguments and build a sentence. Montague uses indexed pronouns as placeholders; I will use indexed empty categories ec. Properties are then formed from this sentence by lambda-binding the placeholders one by one. A quantifier can be introduced as soon as a property is formed.

In syntax, the quantifier will replace the placeholder. Because all placeholders are e-type variables, the result is that the higher (later) a quantifier is introduced, the wider its scope: the other quantifiers are already buried in the definition of the property that it combined with. (Strictly speaking, the involvement of a placeholder is only necessary for non-subject arguments, but I uniformize the treatment to make the differences between S>O and O>S minimal.)

Technically, Montague's PTQ collapses the two steps of lambda-binding a free variable and applying a generalized quantifier to the result into a single rule of quantifying-in. -- Heim and Kratzer (1998) generate structures much like Montague’s. However, they make lambda abstraction a reflex of movement, i.e. of the movement of the index on the variable to be bound. Thus, in the literature following H&K you may see λ2 or simply 2 in the position where the lambda calculus would have λx or λx2.

The two readings of A dog bit every man in the style of Montague's PTQ, extensionalized.

To be read from bottom up, as in a syntax with Merge:

(S>O)  |, A dog bit every man]
      |   ∃z[dog'(z) & ∀y[man'(y) → bit'(y)(z)]]
|<e,t,t> a dog
λP∃z[dog'(z) & P(z)]  λx2 ∀y[man'(y) → bit'(y)(x2)]
  |  [t, ec2 bit every man]
  |  ∀y[man'(y) → bit'(y)(x2)]
[<e,t,t,t> every man]
λQ ∀y[man'(y) → Q(y)]  λx1[bit'(x1)(x2)]
  |  |  [t, ec2 bit ec1]
  |  |  bit'(x1)(x2)
  |  [e ec2]
    x2
[<e,t,t> bit]
λoλs[bit'(o)(s)]  [e ec1]
x1

apply a dog to this

λ-bind subject x2
result is of type <e,t>

apply every man to this

λ-bind object x1
result is of type <e,t>

build sentence w/ 2 free vbls
(O>S)    \[\lambda y[\text{man}'(y) \to \exists z[\text{dog}'(z) & \text{bit}'(y)(z)]]\]
|    \[\lambda x_1 \exists z[\text{dog}'(z) & \text{bit}'(x_1)(z)]\]
|    \[\lambda Q \forall y[\text{man}'(y) \to Q(y)]\]
|    \[\lambda P \exists z[\text{dog}'(z) & P(z)]\]

\[\lambda x_2[\text{bit}'(x)(x_2)]\]

\[\lambda x_2[\text{bit}'(x)(x_2)]\]

\[\lambda x_2[\text{bit}'(x)(x_2)]\]

May 1977 and then 1985 introduces quantifier phrases in argument positions and then adjoins them to VP or S (IP). This adjunction rule is called Quantifier Raising (QR). The higher a quantifier is adjoined, the wider scope it takes, by c-command. Thus the resulting LF structures will be very similar to Montague’s.

\((\text{Subj}>\text{Obj})\) \[\lambda x_1 \exists z[\text{dog}'(z) & \text{bit}'(x_1)(z)]\]

\((\text{Obj}>\text{Subj})\) \[\lambda Q \forall y[\text{man}'(y) \to Q(y)]\]

If one wants to provide an interpretation for May’s syntax, one could just build his output structures in the manner of Montague. In this sense May 1977 is equivalent to Montague 1972, extensionalized. Thus, reversing the actual chronology, you might look upon Montague’s grammar as one that directly builds the representations that May’s QR produces with movement. Montague’s \[\lambda x_1 \exists z[\text{dog}'(z) & \text{bit}'(x_1)(z)]\] corresponds to May’s \[\lambda x_2[\text{bit}'(x)(x_2)]\], and the adjoined quantifiers obviously correspond. But while May’s grammar doesn’t spell out how those quantifiers get related to particular argument places within S, Montague’s does, with the mediation of forming properties by lambda-abstraction and applying the quantifiers to them. Crucially, a compositional semantics will not introduce a quantifier in one position, interpret it there, then move it and recompute the interpretation. One builds a compositional semantics for the desired output (the representation), not for the derivation, if traces are assumed in syntax.

Montague 1972 (PTQ) differs from the above substantially in that it is an intensional theory. Intensional transitive verbs like seek, imagine, owe, etc. cannot take an e-type argument for the direct object: they must apply to the intension of the generalized quantifier denoted by the direct object. Thus seek will be of type \(<s,GQ>, <e,t>\) (using GQ to abbreviate a type). Furthermore Montague “generalizes to the worst case” and treats all verbs uniformly. Extensional readings are obtained via meaning postulates; in the course of applying the postulate we shift to seek*, which is of type \(<e,<e,t>>\). These complications are immaterial to us at present.

Assignment
Required readings:

- My handout in full.

Recommended:

- Heim—Kratzer, Chapters 7-8 are highly recommended, because you’ll often encounter their technology in the literature.

Written:
Take didn’t to be a syntactically primitive adjunct of semantic type \(<<e,t>,<e,t>>, with interpretation \(\lambda P \lambda x[\neg P(x)]\). Using this, build the following sentences on the indicated interpretations. Your task is to assign the correct syntax and semantics to John or Sue and to determine how its interaction with negation can be handled using the Montagovian quantifier scoping apparatus introduced above.

Start out with a brief statement that explains how the interpretations specified below are matters of scope: what is scoping over what in each case when the sentence has the given reading. Go on to build the derivations only when you see these clearly.

(a) Mary didn't visit John or Sue ["she visited neither"]
(b) Mary didn't visit John or Sue ["I'm not sure which she didn’t visit"]
(c) Every boy visited John or Sue ["some visited John, some Sue, some perhaps both"]
(d) Every boy visited John or Sue ["I am not sure which was visited by every boy"]
x, y, z, v, w are of type e
P, Q, R are of type <e,t>
K, N, M are of type <<e,t>,t>

Please abbreviate the words with their initials. Please write out each conversion step, on a separate line. Please write by hand. – Please bring the solutions to the pre-class meeting on Tuesday. We’ll correct them together; you will not hand your solutions in.

\[ \lambda x \lambda y \lambda z [P(x)(y) \& Q(x)(z) \& R(x)(v)] (y)(v) = \]

__________________________________________________________________________
__________________________________________________________________________
__________________________________________________________________________
__________________________________________________________________________
__________________________________________________________________________

the type of \( \lambda x \lambda y \lambda z [P(x)(y) \& Q(x)(z) \& R(x)(v)] \) is ______________________

\[ \lambda y \exists x [\text{sphere}(x)](z) = \]
__________________________________________________________________________
__________________________________________________________________________

the type of \( \lambda y \exists x [\text{sphere}(x)] \) is ____________________

\[ \lambda N [\exists x [\text{rectangle}(x) \& N(\lambda y[x=y])] \& R(\forall x[\text{square}(x) \rightarrow R(x)]) = \]
__________________________________________________________________________
__________________________________________________________________________
__________________________________________________________________________
__________________________________________________________________________
__________________________________________________________________________
__________________________________________________________________________

the type of \( \lambda N [\exists x [\text{rectangle}(x) \& N(\lambda y[x=y])] \) is ______________________
\[\lambda R \exists x [\text{triangle}(x) \land R(x) \land \lambda P \forall y [\text{angle}(x)(y) \rightarrow P(y)] (\lambda z [\text{obtuse}(z)])] = \]

\[\lambda P \exists y [\text{number}(y) \land P(y)] (\lambda x \forall z [\text{prime}(z) \rightarrow \text{divide}(z)(x)]) = \]

\[\lambda R \lambda M \lambda y [M(R) \rightarrow R(y)] (\text{sneeze}) (\lambda P [P(\text{kim}) \land P(x)]) (x) = \]

the type of \[\lambda R \lambda M \lambda y [M(R) \rightarrow R(y)]\] is _____________________
\[ \lambda K \lambda M [K(\lambda x[M(\lambda y[touch(x)(y)])]) \ (\lambda P \exists v[\text{circle}(v) \ & \ P(v)]) \ (\lambda Q \forall y[\text{line}(y) \to Q(y)]) = \]

\[ \lambda P \exists y[\text{dog}(y) \ & \ P(y)] \ (\lambda x \exists y[\text{cat}(y) \ & \ \text{hate}(x)(y)]) = \]

the type of \((\lambda x \exists y[\text{cat}(y) \ & \ \text{hate}(x)(y)])\) is ______________________
OPTIONAL
Equivalent derivations without appealing to traces or syncategorematic lambda-binding

The derivations reviewed are existing options in PTQ but not the only options. Subject noun phrases may be easily entered directly, even if they are quantificational. Can direct object quantifiers be also introduced without invoking the quantifying-in machinery of traces/variables and syncategorematic lambda-binding?

Higher order variables and multiple verb types: Hendriks (1993), Studied Flexibility. PhD, UvA.

(1) It is true that every man kisses a woman.
(2) Het is waar dat iedere man een vrouw kust.

In both (1) and (2) the object can scope over the subject. In (1) it is possible to claim that on the O>S reading the constituent structure is as follows:

(3) ... every man kisses a woman

\[

t/(e\text{'t}) \quad (e\text{'t})/e \quad (t/e)\lambda \forall y[\text{man}(y) \rightarrow \text{kisses}(x)(y)]
\]

\[
\exists x[\text{woman}(x) \& \forall y[\text{man}(y) \rightarrow \text{kisses}(x)(y)]]
\]

This analysis cannot be replicated for Dutch. Given that the object een vrouw intervenes between iedere man and kust, the latter two cannot be claimed to form a surface constituent. The question is, does this force a Montague/May type of treatment of scope assignment, which creates a new constituent structure at some more abstract level?

Hendriks shows that a flexible type system can be designed in which the most conservative assumptions about constituent structure can be maintained and, nevertheless, all imaginable scope relations can be obtained. Flexibility means that lexical items (most importantly, verbs) do not have one fixed semantic type. They start out with a basic type from which an infinite set of new types can be obtained via three particular type transitions (Argument Raising, Value Raising, Argument Lowering).

Hendriks' proposal is novel in that it locates all the information relevant to quantifier scope in the interpretation of the verb. Here is a preliminary idea of how the proposal works using the problematic Dutch example above.

(4) a. iedere man: basic translation: \(\lambda P \forall x[\text{man}(x) \rightarrow P(x)]\)
   een vrouw: basic translation: \(\lambda P \exists x[\text{woman}(x) \& P(x)]\)
   syntactic category: dp, basic semantic type: \(<<e,t>,t>\)

   b. kust: basic translation: \(\lambda x \lambda y[\text{loves}(x)(y)]\)
   syntactic category: dp'/(dp's), basic semantic type: \(<e,<e,t>>\)

The syntactic derivation is as simple as the categories predict:
On the semantic side, the Argument Raising rule will endow kust with a semantic type that enables it to take the two noun phrases of type \(<e,t>,t>\) as arguments in the surface order -- but the fact that such a type can be derived in two different ways will make it possible for the noun phrases to be interpreted in two different scopal orders. The scopal order of the two noun phrases depends on which of the two argument slots was raised first. If the type of the object argument is raised first, and the type of the subject second, we get the skeleton of the subject wide scope reading, and if the type of the subject argument is raised first, we get the skeleton of the object wide scope reading. S and O are variables over generalized quantifier denotations.

The interpretation of kust on the S>O reading:

\[ \lambda O \lambda S[S(\lambda x[O(\lambda y[kust'(x)(y)])])] \]

The interpretation of kust on the O>S reading:

\[ \lambda O \lambda S[O(\lambda x[S(\lambda y[kust'(x)(y)])])] \]

Observe how Hendriks's method will assign the same interpretations to sentences as the Montague/May method; only the syntax (linguistic and logical) is different. In fact, each interpretation of the verb encodes a full derivation tree that was discussed under “Montague extensionalized”.

Another important application of higher order variables (without multiple verb types) is semantic reconstruction, as in Cresti 1995. Because Cresti is concerned with the more complicated case of how many-questions, below is an illustration from Sauerland—Elbourne 2002 and Leu 2003:

(5) A northern team is likely to make it to the finals.
   (i) ‘there is a northern team that is likely to make it to the finals’
   (ii) ‘it is likely that some northern team or other will make it to the finals’

(6) (i) a_northern_team’(\(\lambda x[\text{likely}'(\text{make_it_to_the_finals}'(x)])\])
   (ii) \(\lambda Q[\text{likely}'(Q(\lambda x[\text{make_it_to_the_finals}'(x)])](\text{a_northern_team}')\]

In (6ii) the existential quantifier of a northern team will remain outside the scope of likely, whereas in (6i) the whole QP denotation lands inside its scope, since it replaces Q under likely. Make it to the finals has a subject of type e, and a northern team is interpreted the same GQ in both cases.

**Summary of the two methods**

To summarize, we have seen two proposals for the logico-syntactic treatment of quantifier scope. On the first, quantificational noun phrases (QPs) always denote generalized quantifiers, and they are typically introduced into the sentence via binding a placeholder variable. On the second, QPs denote generalized quantifiers, but neither syntactic variable binding, nor flexible constituency is necessary; the scopal order is encoded in the interpretation of the verb that takes the QPs as arguments. The semantic results of the two methods are identical. It is useful to know that these possibilities exist, because one may choose whichever fits one's theory best.

Whether the theories we reviewed make the correct empirical predictions about what scope ambiguities exist is a separate matter. They do not – as demonstrated in Szabolcsi, ed. 1997, different quantifier types have different scope-taking abilities, whereas these theories offer a uniform treatment.
Existential vs distributive scope

(15) Some fireman or other thought that every building was unsafe
can’t mean: firemen vary with buildings

(16) Every fireman thought that two buildings were unsafe
possible reading: there are two buildings such that every fireman thought they were unsafe

(17) Some fireman imagined that every violinist had one arm
possible reading: a fireman imagined of every actual violinist that he/she had one arm
possible reading: a fireman thought up an all-one-armed-violinists world

(18) Some fireman or other imagined that two buildings were unsafe
can’t mean: firemen vary with buildings

(19) If three relatives of mine die, I will inherit a house
(i) if the # of my dead relatives reaches 3, I inherit one house
(ii) there are three relatives of mine such that if they all die, I inherit one house
(iii)* there are three relatives of mine such that for each, if he/she dies, I inherit a house

The existential scope of both indefinites and universals is unbounded and island-free (thus probably not obtained via movement), whereas their distributive scope is clause-bounded.

(20) Every fireman thought that two or more buildings were unsafe.
?? there are two or more buildings such that every fireman thought they were unsafe

(21) Every fireman thought that less than three buildings were unsafe.
?? there are less than three buildings such that every fireman thought they were unsafe

Quantifiers that do not take clause-internal inverse scope do not have clause-external existential scope.

Reinhart 1997: Indefinites like two buildings are interpreted using existential quantification over choice functions; distributivity is a property of the local predicate.

(22) f is a choice function CH iff S(f(S)), where S is a set.

(23) ∃f[CH(f) & every fireman thought that f(two(buildings)) were unsafe]
‘there is a choice function such that every fireman thought that the value this function returns when it is applied to a set of pairs of buildings was unsafe’

Notice that the set of pairs of buildings is the set of witness sets of [[two buildings]].

Beghelli, Ben-Shalom & Szabolcsi 1997, Variation, distributivity, and … In Ways of Scope Taking.
Beghelli & Stowell 1997, Distributivity and negation… In Ways of Scope Taking.
Reinhart 1997: Division of labor between QR and choice functions. L&P.
Hi,

Here are the works I recall mentioning in the ultimate and the penultimate lectures:

Fox 2004(?), On Logical Form. In Hendrick, ed. [on his home page] (for reconstruction and trace conversion)

Heim--Kratzer 1998, Semantics in Generative Grammar. (moving the index of the variable as the syntactic operation corresponding to lambda binding).

Curry--Feys 1958, Combinatory Logic Vol 1. Don't try to read this. For an introduction to the main ideas see e.g. my "Bound variables in syntax: Are there any?" at http://homepages.nyu.edu/~as109/papers.html and various works by Steedman, e.g. his The Syntactic Process, MIT Press.

Jacobson 1999, Variable-free semantics, Linguistics and Philosophy, for a treatment of pronouns and traces as identity maps (lambda x. x).


Szabolcsi 2000, The syntax of scope http://homepages.nyu.edu/~as109/papers.html (has a longish section on "Different quantifiers, different scopes")

Barker 2006, Direct Compositionality on demand at http://ling.ucsd.edu/~7Ebarker/Research/index.html (this is where I said that for the time being you should just read the prose).

Anna