1 Review: Categorial grammar

- The basic categories $S$, $NP$, $N$, $PP$, $CP$ are in $Cat$.
- For any $A, B \in Cat$, the derived categories $A/LB$ and $A/RB$ are in $Cat$.
- Nothing else is in $Cat$.

Combinatorial rules:

**Right application**: If $\alpha = ([\alpha], A/RB, [\alpha])$ and $\beta = ([\beta], B, [\beta])$, then there is a $\gamma = ([\alpha-\beta], A, f : f(w) = [\alpha](w)([\beta](w)))$.

**Left application**: If $\alpha = ([\alpha], A, [\alpha])$ and $\beta = ([\beta], B/LA, [\beta])$, then there is a $\gamma = ([\alpha-\beta], B, f : f(w) = [\beta](w)([\alpha](w)))$.

2 Coordination

To make sense of coordination we need metavariables over syntactic categories. We'll use capital letters $X, Y, Z$ for this purpose.

- $(X/LX)/RX$: and, or

This means that, for any category $X \in Cat$, and or may have category $(X/LX)/RX$. This means that coordination structures have an asymmetric, binary-branching structure, as in:

```
NP
  /\NP
  NP/\NP
NP/\NP
   /\NP
   NP/\NP
   (NP/\NP)/RN
Barcode (NP/\NP)/RN
and NP
(NP/\NP)/RN laughed S/LNP
S/LNP
S/LNP
((S/LNP)/L(S/LNP))/R(S/LNP)
(S/LNP)/L(S/LNP)
S/LNP
(S/LNP)/L(S/LNP)
or S/LNP
S/LNP ran
```
3 Features

Our CG will capture facts about agreement, etc. again in an HPSG-like manner: agreement means selection of a head with a particular (e.g.) gender/number, case assignment is selection of an NP with a certain case feature. So we might have a schema like this:

- S/NP: run, walk, jog, laugh, ...
- **Lexical Rule**: S/NP $\rightarrow$ S$_L$NP[NOM]

or, when we’re not trying to be parsimonious, we might just write down lexical entries like

- S$_L$NP[SG,NOM]: runs, walks, jogs, laughs, ...

Presumably number agreement should be introduced through a lexical rule which simultaneously adjusts phonology and syntactic category. (Specifically?) Argument-marking prepositions, as in *give the book to Mary*: hopefully there are only a few such items. To handle them, we introduce a special feature for each, e.g. [TO]. *give* will then select for a second object which is a PP with the [TO] feature.

- ([to], PP[TO]/$_R$NP, I), where I is the identity function
- ([give], ((S$_L$NP)/$_R$PP[TO])/$_R$NP, [give])

4 First-order (predicate) logic

4.1 Syntax

(1) The primitive vocabulary of $L_{FO}$ is the union of the following (disjoint) sets:

a. A countable set of **individual constants**, $\{c_i \mid i \in \mathbb{N}\}$;

b. A countable set of **individual variables**, $\{x_i \mid i \in \mathbb{N}\}$;

c. For each $n > 0$, a countable set of $n$-place predicates $\{P_i \mid i \in \mathbb{N}\}$;

d. A set of connectives $\{\neg, \land, \lor, \rightarrow, \leftrightarrow\}$.

We’ll generally write $P_{(i)}$ or $Q_{(i)}$ for one-place predicates, and $R_{(i)}$ or $S_{(i)}$ for two- or three-place predicates, letting the syntax disambiguate.

A bit more annoying but necessary terminology:

- A **TERM** is anything that is either a constant or a variable.

- An $n$-place predicate is also said to be of **ARITY** $n$.

Rules for building up compound expressions:

(2) a. If $d$ is a term and $P$ is a one-place predicate, then $P(d)$ is in $L_{FO}$ (‘is a well-formed formula’).

b. If $d, d'$ are terms and $P$ is a two-place predicate, then $P(d, d')$ is in $L_{FO}$.
c. (etc. for predicates of higher arity)
d. If $\phi \in \mathcal{L}_{FO}$ then $\exists x_i \phi$ is in $\mathcal{L}_{FO}$ for every $i \in \mathbb{N}$.
e. If $\phi \in \mathcal{L}_{FO}$ then $\forall x_i \phi$ is in $\mathcal{L}_{FO}$ for every $i \in \mathbb{N}$.
f. If $\phi, \psi \in \mathcal{L}_{FO}$ then $\neg \phi, \phi \land \psi, \phi \lor \psi, \phi \rightarrow \psi, \phi \leftrightarrow \psi$ are in $\mathcal{L}_{FO}$.
g. Nothing else is in $\mathcal{L}_{FO}$.

Again, we could have only had primitive connectives $\neg$ and $\land$, or $\neg$ and $\lor$, or various other combinations, and introduced the others as abbreviations in the metalanguage (e.g., $\phi \rightarrow \psi \equiv \neg(\phi \land \neg \psi)$). This is a design choice; nothing hinges on it.

4.2 Semantics

This time we’ll have to introduce separate interpretation and evaluation functions, in order to handle free variables.

A MODEL for $\mathcal{L}_{FO}$ is a pair $\mathcal{M}_{FO} = (D, I)$, where

- $D$ is a non-empty set of objects (the domain);
- $I$ is a function whose domain is the union of the constants and the predicates, obeying the following constraints:
  - If $\alpha$ is a constant then $I(\alpha) \in D$.
  - If $\alpha$ is a predicate of arity $n$ then $I(\alpha)$ is a $n$-place relation over $D$.

An $n$-place relation is just a set of $n$-tuples of objects, each of which is in $D$. For instance, if like is a two-place predicate then its denotation should be a set of pairs of individuals, each of which likes the other (according to the model in question): $\{(\text{Bob, Mary}), (\text{Mary, Mary}), (\text{Sue, Mary}), ...\}$

Intuitively, a variable is BOUND iff it is in the scope of a quantifier which controls its interpretation, in the sense we’ll define in a moment. A variable is FREE iff it is not bound. For example, $x$ is bound and $y$ is free in $\forall x[f(x, y)]$.

An ASSIGNMENT function $g$ is a function from variables to $D$.

The evaluation function is always relativized to an assignment function, thus: $[\cdot]^{\mathcal{M}_{FO}:g}$:

- If $\alpha$ is a constant or a predicate then $[\alpha]^{\mathcal{M}_{FO}:g} = I(\alpha)$. (i.e., ignore the assignment function when dealing with things that we know don’t contain variables.)
- If $\alpha$ is a variable then $[\alpha]^{\mathcal{M}_{FO}:g} = g(\alpha)$.
- If $\alpha$ is a term and $\gamma$ is a 1-place predicate then $[\gamma(\alpha)]^{\mathcal{M}_{FO}} = 1$ iff $[\alpha]^{\mathcal{M}_{FO}:g} \in [\gamma]^{\mathcal{M}_{FO}:g}$.
- If $(\alpha, \beta, ...)$ is an $n$-tuple of terms and $\gamma$ is an $n$-place predicate then $[\gamma(\alpha, \beta, ...)]^{\mathcal{M}_{FO}:g} = 1$ iff $(\alpha, \beta, ...) \in [\gamma]^{\mathcal{M}_{FO}:g}$.

Note that the definitions simply look to $g$ for variables and to $I$ for everything else. After that, we don’t care whether an expression is a variable or not, except in the definitions for quantifiers.
The connectives are defined exactly as in propositional logic.

Key clauses for quantified formulae follow. In effect, we’ll interpret quantified statements by playing around with the assignment function. By \( g[x := d] \) we mean the function \( g' \) that is identical to \( g \) in every way except that \( g(x) = d \).

- \( [\forall x \phi]^{M_{FO}} = 1 \) iff, for all \( d \in D \), \( [\phi]^{M_{FO},g[x := d]} = 1 \).
- \( [\exists x \phi]^{M_{FO}} = 1 \) iff, for some \( d \in D \), \( [\phi]^{M_{FO},g[x := d]} = 1 \).

Gloss: \( \forall x \phi \) is true iff \( \phi \) would be true no matter what object we assigned to the variable \( x \). \( \exists x \phi \) is true relative to \( g \) iff \( \phi \) is true under some way of assigning an object to the variable \( x \).

Note that the rules do not forbid vacuous quantification, i.e. quantifying over a variable which does not occur free within the scope of the quantifier.

**Practice:**

- Build up an interpretation step-by-step for the sentence \( \forall x_3 L(x_3, c_1) \) and assign it a truth-value, assuming that \( D = \{ \text{Lady Gaga, Dan Lassiter} \} \), \( L \) is the love relation, and \( I(c_1) = \text{Lady Gaga} \).

- Consider the formula \( \forall x_2 P(c_2) \). What is it equivalent to, and why? Spell out the reason in detail, referring to the use of assignments in the interpretation of universal quantification.

An equivalent way to think about the interpretation of \( \forall \) and \( \exists \) is that we are quantifying (in the metalanguage) over possible assignment functions:

- \( [\forall x \phi]^{M_{FO},g} = 1 \) iff, for all possible \( g' \), \( [\phi]^{M_{FO},g'} = 1 \).
- \( [\exists x \phi]^{M_{FO},g} = 1 \) iff, for some possible \( g' \), \( [\forall x \phi]^{M_{FO},g'} = 1 \). (i.e., it’s possible to find some way of assigning a value to the variable \( x \), such that \( \phi \) is true.)

### 4.3 First-order entailment, etc.

\( \phi \models \psi \) iff, for all \( M_{FO} \) and \( g \) such that \( [\phi]^{M_{FO},g} = 1 \), \( [\psi]^{M_{FO},g} = 1 \).

- Equivalently: The set of pairs \( (M_{FO}, g) \) such that \( [\phi]^{M_{FO},g} = 1 \) is a subset of the set of pairs \( (M_{FO}, g) \) such that \( [\psi]^{M_{FO},g} = 1 \).

\( \phi \) is valid/a tautology iff \( [\phi]^{M_{FO},g} = 1 \) for all \( M_{FO} \) and all \( g \).
\( \phi \) is a contradiction iff there are no \( M_{FO} \) and \( g \) such that \( [\phi]^{M_{FO},g} = 1 \).

**References**