

# Semantics and Pragmatics of NLP

## The Semantics of Discourse: Overview

Alex Lascarides

School of Informatics  
University of Edinburgh

# Outline

- 1 Shortcoming of FOL approaches to semantics
  - Anaphora across sentence boundaries
- 2 Changing the Approach: Discourse Representation Theory
  - A new way of constructing LF
  - A new way of interpreting LF

# Motivation for DRT

Pronouns:

(1) John owns a car. It is red.

wrong:  $\exists x(\text{CAR}(x) \wedge \text{OWN}(j, x)) \wedge \text{RED}(y)$

complex construction:  $\exists x(\text{CAR}(x) \wedge \text{OWN}(j, x) \wedge \text{RED}(x))$

Problems with:

(2) John doesn't own a car. ??It is red.

$\neg \exists x(\text{CAR}(x) \wedge \text{OWN}(j, x) \wedge \text{RED}(x))$

Not recording the right relationship between meaning and *context*.

# More Problems: Time

- (3) John entered the room. He sat down. He lit a cigarette.  
It was pitch dark.

Talking about Time: (sentences true or false *at a time*)

$M \models_t P\phi$  iff there is a time  $t' \prec t$  and  $M \models_{t'} \phi$

$M \models_t F\phi$  iff there is a time  $t' \succ t$  and  $M \models_{t'} \phi$

wrong:

$Ps_1 \wedge Ps_2$

wrong and complex construction:  $P(s_1 \wedge Fs_2)$

complex construction:  $P(Ps_1 \wedge s_2)$

And what about difference between events and states??

Not recording the right relationship between meaning and  
*context*

# More Problems: Presuppositions

Interferes with compositionality of LF construction:

- (4) John's son is bald.
- (5) If baldness is hereditary, then John's son is bald.
- (6) If John has a son, then John's son is bald.

These are all examples of *anaphora*.

## Representing Discourse: *Context Change Potential*

- When we utter *A woman snorts*, we don't simply make a claim about the world, we also *change the context in which subsequent utterances are interpreted*.
- **Anaphora:** semantics involving a relationship between what the anaphor denotes and an *antecedent* in that context.
  - For pronouns the relationship is =
- The structure of the context constrains what can, and cannot, be antecedents ((1) vs. (2)).

## Caching out these Ideas

*John owns a car. It is red*

- Start a new discourse with an empty box:


- expand this box with information from the first sentence:

x,y
john(x), car(y) own(x,y)

← **discourse referents:**

Things the discourse is about.

← **conditions:** relations and properties among discourse referents

Proper names are now conditions; so all NPs introduce a discourse referent and the Nbar introduces conditions on it.

# Processing the Second Sentence

## *John owns a car. It is red*

- Pronoun is an NP, and so like all NPs it introduces a new discourse referent  $z$ .
- the VP contributes  $\text{red}(z)$  (as before).
- Pronouns are special!  
They introduce an equality condition to a discourse referent (of same number and gender) in the context:

- So  $z=y$ :

$x, y, z$
$\text{john}(x), \text{car}(y)$ $\text{own}(x, y)$ $\text{red}(z), z=y$

# LF Construction has Changed!

Before:

**Compositionality:** the contribution to LF of an NL expression determined entirely by the contributions of its (syntactic) daughters.

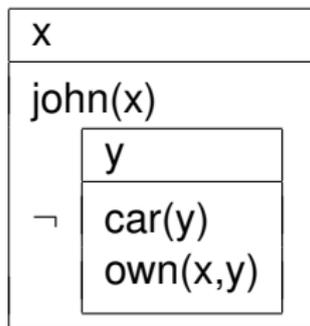
Now:

- Adding  $z=y$  is *not* compositional!
- Construction now depends on what's already in the box, and not just on syntax.
- This accurately reflects the fact that the meaning of a pronoun is dependent on context.

# Negation

*John doesn't own a car. It is red*

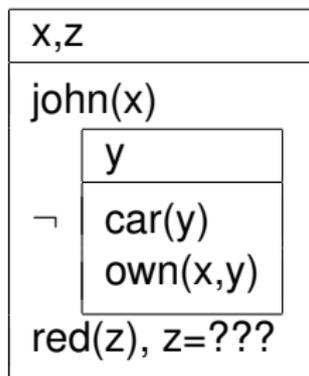
- Use  $\neg$  to indicate what's false:



- So we get boxes inside boxes!

# Structure Blocks Antecedents *John doesn't own a car.* *It is red*

- The antecedent discourse referent for a pronoun must be introduced in the same box or a 'bigger' box.
- *It is red* is *outside* the negation; *y* is *inaccessible* and pronoun is uninterpretable.



- Construction is dependent on *form*; not on interpretation.

## Important things we'll ignore for now

Selectional restrictions:

(7) John petted his cat. He purred affectionately.

Coherence:

(8) John can open Bill's safe. He knows the combination.

Will also gloss over grammatical constraints:

- \*John loves him<sub>john</sub>
- John buys a new car every year. It is/They are always red.
- John buys a new car every year. Last year it was/\*they were red.

Want to focus on the interaction between anaphora and logical structure: *not*, *if... then*, quantifiers etc.

# DRS And Predicate Logic

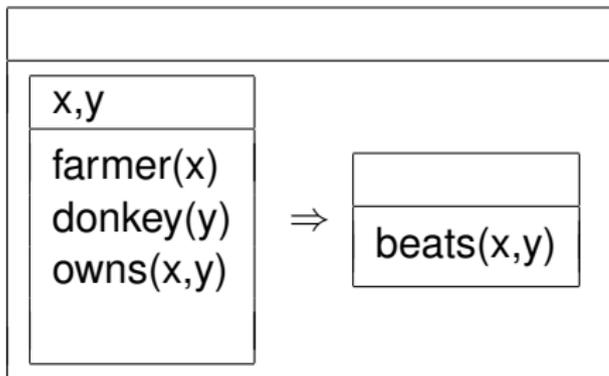
*Why use the funny box notation?*

Answers:

- One can translate certain DRS fragments into FOL with discourse referents being free variables.
- BUT:
  - If one did this during LF construction, then the hierarchical structure of DRSs would be lost, and this plays an important part in constraining how to insert new material.
  - It's more convenient to use the box notation.

# A Taster: DRSs can Get Complicated

(9) Every farmer who owns a donkey beats it.



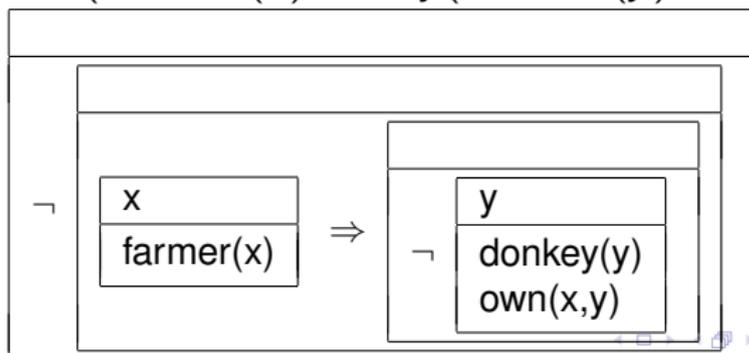
# Semantic Equivalence But Structural Differences

(10) A farmer owns a donkey.

$$\exists x \exists y (\text{FARMER}(x) \wedge \text{DONKEY}(y) \wedge \text{OWN}(x, y))$$

x,y
farmer(x)
donkey(y)
own(x,y)

(11) It's not the case that all farmers don't own a donkey.

$$\neg \forall x (\text{FARMER}(x) \rightarrow \neg \exists y (\text{DONKEY}(y) \wedge \text{OWN}(x, y)))$$


# DRS Languages

- DRSs can be nested and combined using  $\neg$ ,  $\vee$ ,  $\Rightarrow$ .
  - if  $K_1$  and  $K_2$  are DRSs,  
then  $\neg K_1$ ,  $K_1 \vee K_2$  and  $K_1 \Rightarrow K_2$  are *DRS conditions*.
- They also contain predicate symbols (e.g., woman, love), like FOL does.
  - woman(x) and love(x,y) are *atomic DRS conditions*
- DRS languages contain symbols  $x, y, \dots$ , they're called *discourse referents*, not variables.
- (Vanilla) DRS languages *don't* contain  $\forall$  or  $\exists$ .
  - Quantification is implicit, in the semantics of DRSs

# Informal Semantics: Boxes as Pictures

A DRS is *satisfied* in a model iff it is an accurate image of the information recorded inside the model.

*A woman snorts. She collapses.*

x y
woman(x) snort(x) collapse(y), x=y

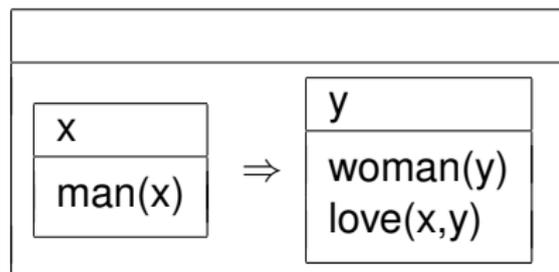
is satisfied iff it is possible to associate the discourse referents x and y with entities in the model such that:

- 1 the first entity is a woman and snorts
- 2 the second entity collapses and is equal to the first entity.

## More Informal Semantics: Complex Conditions

**Negated DRS:** satisfied iff it is *not* possible to find the picture inside the model.

**Disjunctive DRSs:** satisfied iff at least one of the pictures can be found in the model.



No matter which entities we use to verify the antecedent picture, we can verify the consequent picture (with those entities plus others).

# Accessibility

- It's a *geometrical* concept:  
configuration of DRSs; how they're nested.
- Discourse referents introduced in DRS  $K_1$  are accessible to (anaphoric) conditions in DRS  $K_2$  iff  $K_1$  *subordinates*  $K_2$  or  $K_1$  *equals*  $K_2$ .

# So what's Subordination then?

$K_1$  subordinates  $K_2$  iff:

- 1  $K_1$  contains the DRS condition  $\neg K_2$ ; or
- 2  $K_1$  contains the DRS condition  $K_2 \Rightarrow K$  or  $K \Rightarrow K_2$ ; or
- 3  $K_1$  contains the DRS condition  $K_2 \vee K$  or  $K \vee K_2$ ; or
- 4  $K$  contains the condition  $K_1 \Rightarrow K_2$ ; or
- 5  $K_1$  subordinates  $K$  and  $K$  subordinates  $K_2$  (transitive closure).

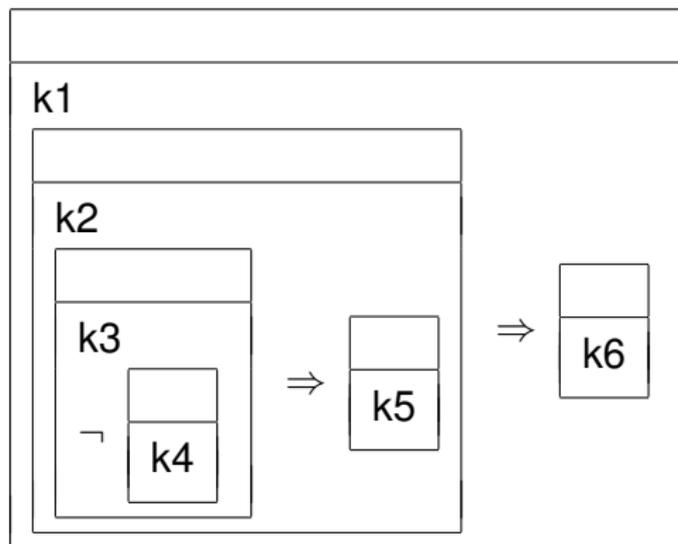
# Working out what's subordinate the simple way

To see if  $K_1$  subordinates  $K_2$

- 1 Start at  $K_2$ ;
- 2 If there's a DRS immediately to your left, move to that.
- 3 If not, but there's a DRS immediately up, move to that.
- 4 else, stop.

if  $K_1$  is on this path, then  $K_1$  subordinates  $K_2$

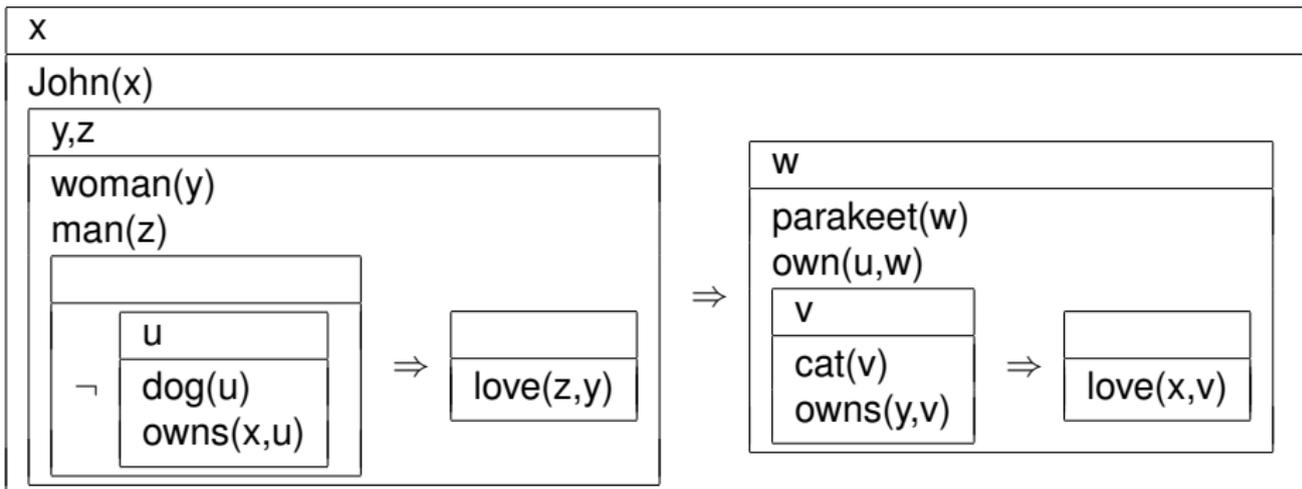
## Starting at k6...



Path is: k6, k2, k1

So discourse referents introduced in k5, k4 and k3 would be *inaccessible* to conditions in k6.

## Another Example



$x, y, z, w$  and  $v$  are accessible to  $love(x, v)$

$u$  isn't accessible to  $love(x, v)$

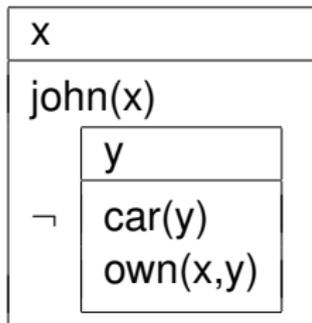
$x, y$  and  $z$  are accessible to  $love(z, y)$

$u, w$  and  $v$  are not accessible to  $love(z, y)$

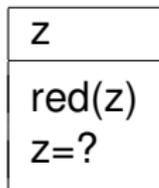
# An Inaccessible Pronoun

*John doesn't own a car. It is red*

(One of) the DRSs for *John doesn't own a car.*



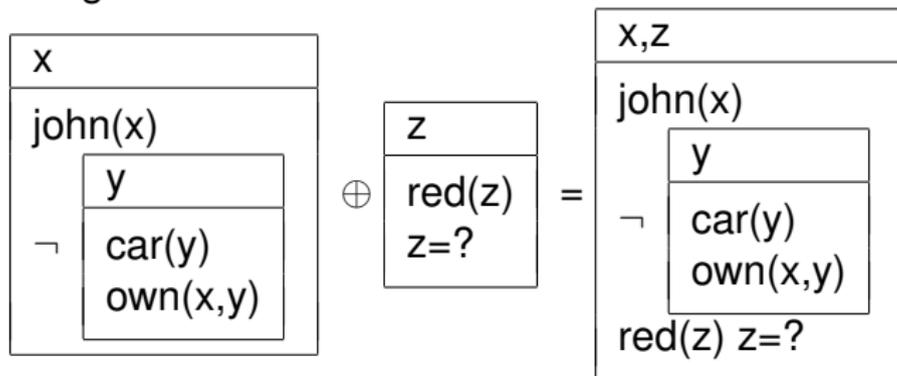
The DRS for *It is red:*



z=? : an instruction to resolve the equality with something accessible.

# The LF of the Discourse: Merge plus resolution

Merge:



Resolving  $z=?$ :

$x$  is accessible, but wrong gender...

## More Discourses that DRT 'gets right'

- (12) Mia ordered a five dollar shake. Vincent tasted it.
- (13) Mia didn't order a five dollar shake. ??Vincent tasted it.
- (14) Butch stole a chopper. It belonged to Zed.
- (15) Butch stole a chopper or a motor cycle. ??It belonged to Zed.
- (16) If a boxer loves a woman she is happy.
- (17) Every woman snorts. ??She collapses.

# Conclusion

- FOL as a semantic representation of NL discourse is problematic because when dealing with *anaphora*, either:
  - it gets the truth conditions wrong, and/or
  - LF construction would be really complicated
- DRSs potentially fare better, because they:
  - Offer a story about how things like negation and conditionals block things from being antecedents to anaphora
  - Through merging DRSs and making pronouns look for antecedents, LF construction may turn out OK too.
- But we will see how LF construction is done next time.