Lexical Semantics: Where Linguistics Meets the Real World

Alex Lascarides

Centre for Cognitive Science *and*
Human Communication Research Centre,
University of Edinburgh

Ann Copestake

CSLI,
Stanford University

June 24th, 1996

Overview

- Some approaches to lexical semantics
- How to model them in unification-based frameworks
- Role of pragmatics
- Formalizing interactions of lexical semantics and pragmatics

Basic assumptions

- Lexical semantics is needed for an account of linguistic phenomena (e.g. syntax, regular polysemy)
- Lexical knowledge is the interface between world knowledge and linguistic knowledge/processing
- Discourse / pragmatic processing interacts with lexical semantics
- Lexical semantic information can be modelled using a typed feature structure formalism, extended to handle defaults

First half outline

Approaches and representation:

1. Sense enumeration
2. Verb classes and alternations
3. Noun / verb interaction
4. Derivation and extended uses
5. Collocations etc

Introducing *bake*

(1) a. Kim is baking
b. The potatoes are baking
c. Kim is baking a cake
d. Kim is baking a cake for Sandy
e. Kim is baking Sandy a cake
f. The clay baked
g. The clay baked hard
h. Sandy was baking, sitting in the hot sun

Sense enumeration: a strawman

(2) a. $\langle \text{bake}, \text{TakesNP}, bake'_1 \rangle$
b. $\langle \text{bake}, \text{TakesNP}, bake'_2 \rangle$
c. $\langle \text{bake}, \text{TakesNP.NP}, bake'_3 \rangle$
d. $\langle \text{bake}, \text{TakesNP.NP.PP}, bake'_4 \rangle$
e. $\langle \text{bake}, \text{TakesNP.NP.NP}, bake'_5 \rangle$

Senses connected by *meaning postulates*, such as:

- $\forall x, y[bake'_3(x, y) \rightarrow bake'_1(x) \wedge bake'_2(y)]$
- $\forall x, y[bake'_4(x, y, z) \rightarrow bake'_3(x, y)]$

But:

- Doesn't capture generalisations (cf *cook*, *paint*)
- No distinction between homonyms (accidental polysemy) and related senses
- Meaning postulates are unrestricted

Lexical entry for transitive *bake*

$$
\begin{bmatrix}
\textbf{transitive-verb} \\
\text{ORTH} : \ \textbf{bake} \\
\text{SYN} : \begin{bmatrix}
\text{CAT} : \ \textbf{v} \\
\text{SUBJ} : \begin{bmatrix} \textbf{phrase} \\ \text{SYN:CAT} : \ \textbf{n} \\ \text{SEM:INDEX} : \ \boxed{\text{x}} \end{bmatrix} \\
\text{SUBCAT} : \left\langle \begin{bmatrix} \textbf{phrase} \\ \text{SYN:CAT} : \ \textbf{n} \\ \text{SEM:INDEX} : \ \boxed{\text{y}} \end{bmatrix} \right\rangle
\end{bmatrix} \\
\text{SEM} : \begin{bmatrix}
\text{INDEX} : \ \boxed{\text{e}} \\
\text{LISZT} : \left\langle \begin{bmatrix} \textbf{\_bake3\_rel} \\ \text{EVENT} : \ \boxed{\text{e}} \\ \text{ARG1} : \ \boxed{\text{x}} \\ \text{ARG2} : \ \boxed{\text{y}} \end{bmatrix} \right\rangle
\end{bmatrix}
\end{bmatrix}
$$

HPSGish syntax and MRS — neutral wrt various approaches to lexical semantics

bread (NP)

$$\begin{bmatrix} \textbf{phrase} \\ \text{ORTH}: \ \textbf{bread} \\ \text{SYN}: \begin{bmatrix} \text{CAT}: \ \textbf{n} \\ \text{SUBCAT}: \langle\rangle \end{bmatrix} \\ \text{SEM}: \begin{bmatrix} \text{INDEX}: \boxed{\text{w}} \\ \text{LISZT}: \left\langle \begin{bmatrix} \textbf{\_bread\_rel} \\ \text{INST}: \boxed{\text{w}} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

LISZTs are appended when signs are combined

$$\begin{bmatrix} \text{LISZT}: \left\langle \begin{bmatrix} \textbf{\_bake3\_rel} \\ \text{EVENT}: \boxed{\text{e}} \\ \text{ARG1}: \boxed{\text{x}} \\ \text{ARG2}: \boxed{\text{y}} \end{bmatrix}, \begin{bmatrix} \textbf{\_bread\_rel} \\ \text{INST}: \boxed{\text{y}} \end{bmatrix} \right\rangle \end{bmatrix}$$

Abbreviated description

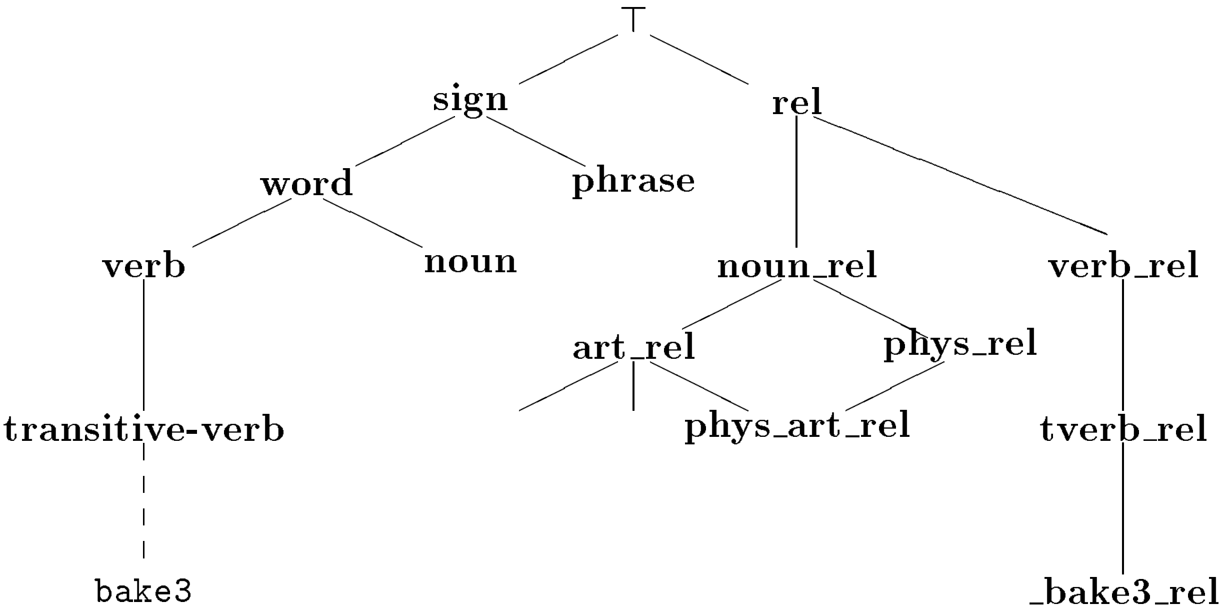
$$
\begin{bmatrix}
\textbf{transitive-verb} & \\
\text{ORTH}: & \textbf{bake} \\
\text{SYN}: & \begin{bmatrix} \text{CAT}: & \textbf{v} \\ \text{SUBJ}: & \textbf{NP}_{\boxed{\mathbf{x}}} \\ \text{SUBCAT}: & \langle \textbf{NP}_{\boxed{\mathbf{y}}} \rangle \end{bmatrix} \\
\text{SEM}: & \begin{bmatrix} \text{INDEX}: & \boxed{\text{e}} \\ \text{LISZT}: & \left\langle \begin{bmatrix} \textbf{\_bake3\_rel} & \\ \text{EVENT}: & \boxed{\text{e}} \\ \text{ARG1}: & \boxed{\text{x}} \\ \text{ARG2}: & \boxed{\text{y}} \end{bmatrix} \right\rangle \end{bmatrix}
\end{bmatrix}
$$

Type for simple transitive verbs

$$
\begin{bmatrix}
\textbf{transitive-verb} & \\
\text{ORTH}: & \textbf{string} \\
\text{SYN}: & \begin{bmatrix} \text{CAT}: \ \textbf{v} \\ \text{SUBJ}: \ \textbf{NP}_{\boxed{\mathbf{x}}} \\ \text{SUBCAT}: \langle \textbf{NP}_{\boxed{\mathbf{y}}} \rangle \end{bmatrix} \\
\text{SEM}: & \begin{bmatrix} \text{INDEX}: \ \boxed{\text{e}} \\ \text{LISZT}: \left\langle \begin{bmatrix} \textbf{rel} \\ \text{EVENT}: \ \boxed{\text{e}} \\ \text{ARG1}: \ \boxed{\text{x}} \\ \text{ARG2}: \ \boxed{\text{y}} \end{bmatrix} \right\rangle \end{bmatrix}
\end{bmatrix}
$$

```
bake3 < transitive-verb
[ ORTH = "bake"
  SEM : LISZT = < _bake3_rel > ]
```

Type hierarchies



- Information on higher types is inherited by lower types (and lexical entries, such as bake3)
- Multiple inheritance is possible
- Extension to formalism allows for defaults

Verb classes: Levin (1993)

Large-scale descriptive account of (some) English verbs, pushing the idea that syntactic behaviour is (partly) semantically determined

- alternations
- semantically coherent classes which exhibit the same alternations
- some extended senses (eg *whistle* as a movement verb)
- some data on morphology / conversion
- extensive bibliography

26 Verbs of Creation and Transformation

26.1 *Build verbs*

Class members include: bake, cook

(3) Material/Product Alternation (transitive):

a. Martha baked a loaf out of wholewheat flour

b. Martha baked some wholewheat flour into a loaf

(4) Unspecified Object Alternation:

a. Martha bakes bread

b. Martha bakes

(5) Benefactive alternation

a. Martha baked a loaf (out of wholewheat flour) for her aunt

b. Martha baked her aunt a loaf (out of wholewheat flour)

26.3 Verbs of Preparing

Class members include: bake, cook, fix, fry

(6) * Material/Product Alternation (transitive):

a. ? Donna fixed a sandwich from last night's leftovers

b. * Donna fixed last night's leftovers into a sandwich

(7) Benefactive alternation

a. Donna fixed a sandwich for me

b. Donna fixed me a sandwich

(8) * Causative alternations

a. Donna fixed a sandwich

b. * a sandwich fixed

45 Verbs of change of state

45.3 Cooking verbs

Class members include: bake, cook, fry

(9) Causative/Inchoative Alternation

a. Jennifer baked the potatoes
b. The potatoes baked

(10) Middle Alternation

a. Jennifer baked Idaho potatoes
b. Idaho potatoes bake beautifully

(11) Instrument Subject Alternation

a. Jennifer baked the potatoes in the oven
b. This oven bakes potatoes well

(12) Resultative Phrase

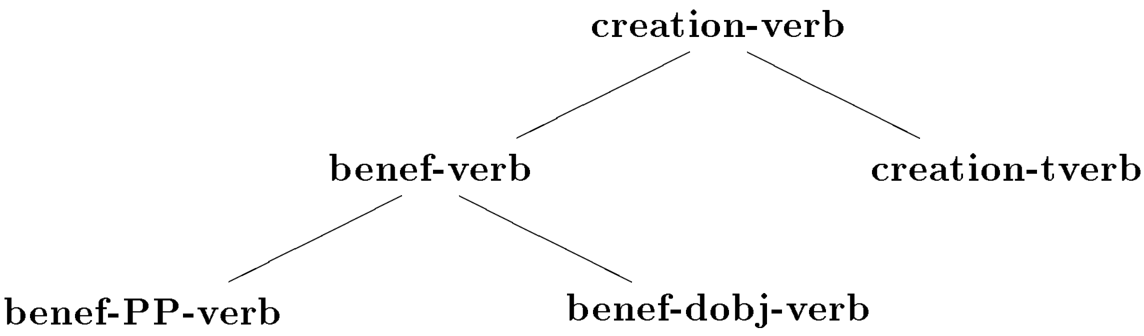
a. Jennifer baked the potatoes to a crisp

Representing alternations: Lexical rules

$$
\left[\begin{array}{l}
\textbf{creation-tverb} \\
\text{ORTH}: \boxed{1} \\
\text{SYN}: \left[\begin{array}{l} \text{CAT}: \textbf{v} \\ \text{SUBJ}: \textbf{NP}_{\boxed{\mathbf{x}}} \\ \text{SUBCAT}: \langle \textbf{NP}_{\boxed{\mathbf{y}}} \rangle \end{array}\right] \\
\text{SEM}: \left[\begin{array}{l} \text{INDEX}: \boxed{e} \\ \text{LISZT}: \left\langle \boxed{2} \left[\begin{array}{l} \textbf{rel} \\ \text{EVENT}: \boxed{e} \\ \text{ARG1}: \boxed{x} \\ \text{ARG2}: \boxed{y} \end{array}\right] \right\rangle \end{array}\right]
\end{array}\right]
\rightarrow
$$

$$
\left[\begin{array}{l}
\textbf{benef-PP-verb} \\
\text{ORTH}: \boxed{1} \\
\text{SYN}: \left[\begin{array}{l} \text{CAT}: \textbf{v} \\ \text{SUBJ}: \textbf{NP}_{\boxed{\mathbf{x}}} \\ \text{SUBCAT}: \langle \textbf{NP}_{\boxed{\mathbf{y}}}, \textbf{PP(for)}_{\boxed{\mathbf{z}}} \rangle \end{array}\right] \\
\text{SEM}: \left[\begin{array}{l} \text{INDEX}: \boxed{e} \\ \text{LISZT}: \left\langle \boxed{2} \left[\begin{array}{l} \textbf{rel} \\ \text{EVENT}: \boxed{e} \\ \text{ARG1}: \boxed{x} \\ \text{ARG2}: \boxed{y} \end{array}\right], \left[\begin{array}{l} \textbf{benef\_rel} \\ \text{EVENT}: \boxed{e} \\ \text{ARG1}: \boxed{z} \end{array}\right] \right\rangle \end{array}\right]
\end{array}\right]
$$

Representing alternations: Underspecification



$$
\begin{bmatrix}
\textbf{creation-verb} & \\
\text{ORTH}: & \textbf{string} \\
\text{SYN}: & \begin{bmatrix} \text{CAT}: \ \textbf{v} \\ \text{SUBJ}: \ \textbf{NP}_{\boxed{\textbf{x}}} \\ \text{SUBCAT}: \langle \textbf{NP}, \ldots \rangle \end{bmatrix} \\
\text{SEM}: & \begin{bmatrix} \text{INDEX}: \ \boxed{\text{e}} \\ \text{LISZT}: \left\langle \begin{bmatrix} \textbf{rel} \\ \text{EVENT}: \ \boxed{\text{e}} \\ \text{ARG1}: \ \boxed{\text{x}} \\ \text{ARG2}: \ \boxed{\text{y}} \end{bmatrix}, \ldots \right\rangle \end{bmatrix}
\end{bmatrix}
$$

Underspecification vs lexical rules

Underspecification

- is essential if word can be used generally: *brush*
- may avoid spurious ambiguity and may be more efficient computationally
- cannot model recursive operation (*reretie*, *great-great-grandfather*)
- requires that semantics be "additive":
  - $\lambda x[P(x)]$ specialized to $\lambda x[P(x) \wedge Q(x)]$
  - but not $\lambda x[P(x)]$ specialized to $\lambda y \exists x[P(x) \wedge R(x, y) \wedge Q(x)]$
  - or $\lambda x[P(x)]$ specialized to $\lambda w[Q(P)(w)]$
- can result in complex type hierarchy
- is unintuitive if there is asymmetry
- in general, is not suitable for extended senses

Construction Grammar

- Fillmore and Kay (in press), Goldberg (1995) on alternations
- deals with constructions such as *the Xer the Yer* (e.g. *the bigger the better*)
- alternative to standard approaches to alternations — e.g. dative construction
- recent approaches to lexical rules and phrasal schemata within HPSG have many similarities with this

Nonce usages

(13) electronic mail

(14) an email message

(15) an email

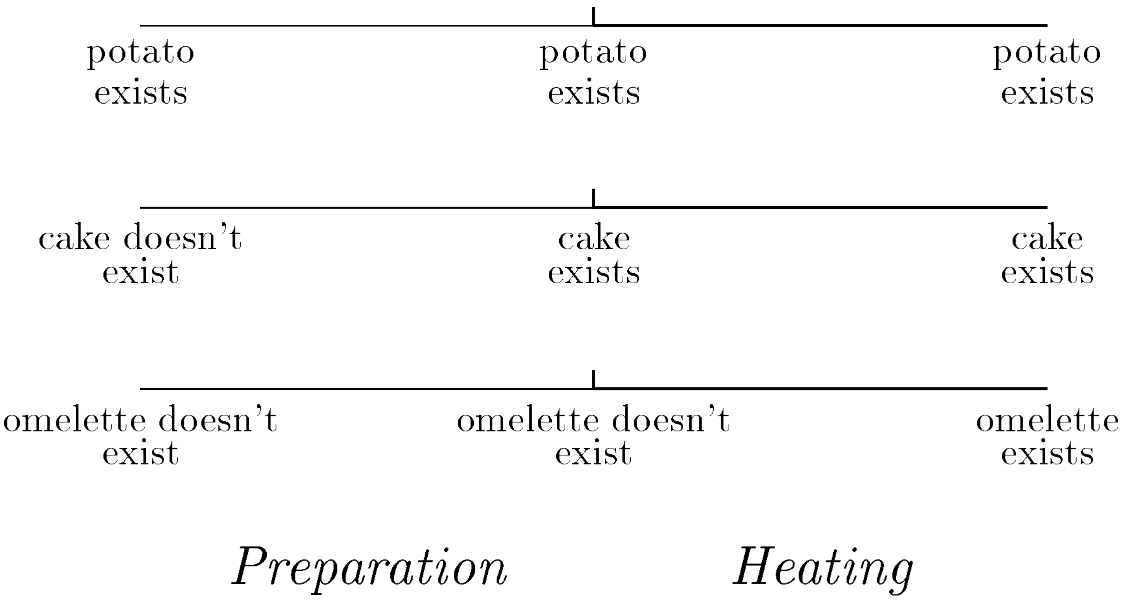
(16) I emailed him yesterday

(17) I emailed him the paper

- generative devices are essential to model this

Change of state vs. Creation

(18) a. John baked a potato for/?in 90 minutes.
b. John baked a cake ?for/in 90 minutes.
c. John fried an omelette ?for/in 3 minutes.



Co-composition: Pustejovsky (e.g. 1991, 1995)

- Rejects sense enumeration account of different meanings
- Single lexical entry for *bake*
- Noun influences phrase meaning: creation use only arises when noun specifies that *agentive* role is *bake*
- Phrase meaning arises from *co-composition* of meanings of noun and verb

Generative lexicon: Pustejovsky (1995)

Lexical meaning is captured in four levels of representation:

1. Argument Structure
2. Event Structure
3. Qualia Structure
4. Inheritance Structure

- Argument Structure: adicity, semantic type, syntax-semantics interface
- Event Structure: more complex than Vendlerian distinction (see also Dowty, Krifka, Moens and Steedman, Zaenen)
- Inheritance Structure: inheritance hierarchy (also *projective inheritance*)

Qualia structure

- Inspired by Moravcsik's (1975) interpretation of Aristotle's *aitiae*
- Nouns (and verbs) have associated qualia:

  Constitutive the relation between an object and its constituents or proper parts

  Formal Role that which distinguishes the object within a larger domain

  Telic purpose and function

  Agentive factors involved in origin or 'bringing about'

- Qualia allow multiple dimensions of meaning to be represented in a single entry / feature structure

Co-composition for *bake*: Pustejovsky (1995:123)

$$
\begin{bmatrix}
\textbf{bake} & \\
\text{EVENTSTR}: & \begin{bmatrix} \text{E1}: & \textbf{e}_1\textbf{:process} \\ \text{HEAD}: & \textbf{e}_1 \end{bmatrix} \\
\text{ARGSTR}: & \begin{bmatrix} \text{ARG1}: & \boxed{1} \begin{bmatrix} \textbf{animate\_ind} & \\ \text{FORMAL}: & \textbf{physobj} \end{bmatrix} \\ \text{ARG2}: & \boxed{2} \begin{bmatrix} \textbf{mass} & \\ \text{FORMAL}: & \textbf{physobj} \end{bmatrix} \end{bmatrix} \\
\text{QUALIA}: & \begin{bmatrix} \textbf{state\_change\_lcp} & \\ \text{AGENTIVE}: & \textbf{bake\_act}(\textbf{e}_1, \boxed{\textbf{1}}, \boxed{\textbf{2}}) \end{bmatrix}
\end{bmatrix}
$$

$$
\begin{bmatrix}
\textbf{cake} & \\
\text{ARGSTR}: & \begin{bmatrix} \text{ARG1}: & \textbf{x:food\_ind} \\ \text{D-ARG1}: & \textbf{y:mass} \end{bmatrix} \\
\text{QUALIA}: & \begin{bmatrix} \text{CONST}: & \textbf{y} \\ \text{FORMAL}: & \textbf{x} \\ \text{TELIC}: & \textbf{eat}(\textbf{e}_2,\textbf{z},\textbf{x}) \\ \text{AGENTIVE}: & \textbf{bake\_act}(\textbf{e}_1,\textbf{w},\textbf{y}) \end{bmatrix}
\end{bmatrix}
$$

bake a cake

$$
\left[\begin{array}{ll}
\multicolumn{2}{l}{\textbf{bake a cake}} \\
\text{EVENTSTR}: & \left[\begin{array}{l}
\text{E1}: \ \mathbf{e_1}\textbf{:process} \\
\text{E2}: \ \mathbf{e_2}\textbf{:state} \\
\text{RESTR}: \ <_\alpha \\
\text{HEAD}: \ \mathbf{e_1}
\end{array}\right] \\
\text{ARGSTR}: & \left[\begin{array}{ll}
\text{ARG1}: \boxed{1} & \left[\begin{array}{l}
\textbf{animate\_ind} \\
\text{FORMAL}: \ \textbf{physobj}
\end{array}\right] \\
\text{ARG2}: \boxed{2} & \left[\begin{array}{l}
\textbf{artifact} \\
\text{CONST}: \boxed{3} \\
\text{FORMAL}: \ \textbf{physobj}
\end{array}\right] \\
\text{D-ARG1}: \boxed{3} & \left[\begin{array}{l}
\textbf{material} \\
\text{FORMAL}: \ \textbf{mass}
\end{array}\right]
\end{array}\right] \\
\text{QUALIA}: & \left[\begin{array}{l}
\textbf{create-lcp} \\
\text{FORMAL}: \ \textbf{exist}(\mathbf{e_2}, \boxed{\mathbf{2}}) \\
\text{AGENTIVE}: \ \textbf{bake\_act}(\mathbf{e_1}, \boxed{\mathbf{1}}, \boxed{\mathbf{3}})
\end{array}\right]
\end{array}\right]
$$

Formalizing co-composition for *bake*

$$\begin{bmatrix} \textbf{process-or-creation} \\ \text{SEM}: \begin{bmatrix} \text{INDEX}: \boxed{\text{e}} \\ \text{LISZT}: \left\langle \begin{bmatrix} \textbf{rel} \\ \text{EV}: \boxed{\text{e}} \\ \text{ARG1}: \boxed{\text{x}} \\ \text{ARG2}: \boxed{\text{y}} \end{bmatrix}, \ldots \right\rangle \end{bmatrix} \end{bmatrix}$$

$$\begin{bmatrix} \textbf{process-tverb} \\ \text{SEM}: \begin{bmatrix} \text{INDEX}: \boxed{\text{e}}\ \textbf{process} \end{bmatrix} \end{bmatrix}$$

$$\begin{bmatrix} \textbf{creation-tverb} \\ \text{SYN}: \begin{bmatrix} \text{SUBCAT}: \langle \textbf{NP}_{\boxed{\textbf{y}}} \text{QUALIA:AGENTIVE} \boxed{\textbf{1}} \rangle \end{bmatrix} \\ \text{SEM}: \begin{bmatrix} \text{INDEX}: \boxed{\text{e}}\ \textbf{achievement} \\ \text{LISZT}: \boxed{1} \left\langle \begin{bmatrix} \textbf{rel} \\ \text{EV}: \boxed{\text{e}} \\ \text{ARG1}: \boxed{\text{x}} \\ \text{ARG2}: \boxed{\text{y}} \end{bmatrix}, \begin{bmatrix} \textbf{create\_rel} \\ \text{EV}: \boxed{\text{e}'} \\ \text{ARG1}: \boxed{\text{x}} \\ \text{ARG2}: \boxed{\text{y}} \\ \text{ARG3}: \boxed{\text{z}} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

Complications for co-composition for *bake*

- inchoative not possible with creation reading
- part-baked loaves etc
- Not all verbs behave this way:

  (19) steam the pudding

  (20) ?freeze the ice cubes
- *bake potatoes* isn't simply a process, there is some culmination (when they're cooked)
- Creation only with flour-based products:

  (21) ?Kim baked the gratin in 90 minutes
- Many nouns would have to have very underspecified qualia (e.g. *doll*)
- Many verbs can take on creation sense, in context.

  (22) Kim folded me a unicorn

Logical Metonymy

(23) a. John enjoyed reading the book.
b. John enjoyed the book.

The meaning is partly determined by the NP:

(24) John enjoyed the hamburger.

(25) John enjoyed the beer.

(26) ? John enjoyed the dictionary.

Control:

(27) John enjoyed eating your hamburger.

(28) John enjoyed your hamburger.

Similarly *begin*, *finish*. . .

Generalisation for *enjoy* etc

- *enjoy* and similar verbs always take events
- object-denoting argument is *coerced* to an event
- the type of event is determined by the qualia structure of the noun

For example, *book* has telic role **read**

(29) John enjoyed the book.

(30) John enjoyed reading the book.

Generalisation for *enjoy* (continued)

Example (31) is odd because *pebble* has no conventional purpose (contrast *like*, *want*):

(31) ?John enjoyed the pebble.

(32) is odd because *enjoy* needs an eventuality with some duration (state, process, accomplishment) and *dictionary* is usually *consulted* (achievement):

(32) ?John enjoyed the dictionary.

- Lexical generalisations have exceptions that are triggered by pragmatic knowledge, such as information about the world.

(33) The goat enjoyed the book.

(34) My goat eats anything.
He really enjoyed your book.

Alternative account with multiple lexical entries cannot account for coordination:

(35) John enjoys books, films and playing the guitar

Purely pragmatic account does not allow for effects of lexical choice.
Suppose I am using a copy of *War and Peace* as a doorstop:

(36) I didn't enjoy that book

(37) I didn't enjoy that

(38) *I didn't enjoy that doorstop

Presupposition:

(39) John enjoyed the book

(40) John didn't enjoy the book

Logical metonymy — continued

Analysis of LOB corpus (Briscoe et al 1990):

- Metonymic examples are common
- Explicit verb is relatively rarely used when the noun qualia would give the correct reading
- Verb is normally supplied in cases where qualia would give an incorrect reading unless immediate context is sufficiently informative

Coercion of NP does not allow for:

(41) Sam picked up and finished his beer

(42) Sam wrote but later regretted that article

absence of adverbs modifying metonymic predicate

(43) * Sam enjoyed the book aloud

Coercion and underspecification

$$\begin{bmatrix} \textbf{lex-enjoy-verb} \\ \text{SUBCAT} : \langle \textbf{v-or-np} \rangle \\ \text{SEM} : \ \textbf{tvsem} \end{bmatrix}$$

$$\begin{bmatrix} \textbf{coercing} \\ \text{SUBCAT} : \langle \textbf{NP} \rangle \end{bmatrix} \quad \begin{bmatrix} \textbf{non-coercing-NP} \\ \text{SUBCAT} : \langle \textbf{NP} \rangle \end{bmatrix} \quad \begin{bmatrix} \textbf{non-coercing} \\ \text{SUBCAT} : \langle \textbf{VPing} \rangle \end{bmatrix}$$

- For the coerced form, there is an extra conjunct in the semantics, which is instantiated from the noun's telic role.
- Value of telic role is defeasible (indicated by preceeding slash)

$$\begin{bmatrix} \textbf{literature} \\ \text{QUALIA TELIC} : \ \textbf{/read\_rel} \end{bmatrix}$$

- See later for details

Implicit arguments

Sentences in 44 are interpreted as having an implicit (narrow-scope) existential bakee. At least by default, this is bread, cakes, pastries etc.

(44) a. Kim bakes

b. Kim baked on Tuesday

c. Sixty potatoes were needed for the party. Kim was baking all afternoon.

Similarly, intransitive *drink* implies alcoholic beverage, *eat* implies a meal (Fillmore, 1986).

$$\begin{bmatrix} \text{SYN} : \begin{bmatrix} \text{SUBCAT} : \langle \textbf{NP}_{\boxed{\textbf{y}}} \, \text{SEM:LISZT}\langle / \textbf{cake-etc\_rel}\rangle\rangle \end{bmatrix} \\ \text{SEM} : \begin{bmatrix} \text{LISZT} : \left\langle \begin{bmatrix} \textbf{\_bake\_rel} \\ \text{EV} : \boxed{\text{e}} \\ \text{ARG1} : \boxed{\text{x}} \\ \text{ARG2} : \boxed{\text{y}} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

Derivation

---

baker n (from OED)

1. one who bakes; spec one whose business it is to make bread
2. (in US) a small portable tin oven in which baking is performed
3. an artifical salmon fly used in angling

Derivation (continued)

- *-er* suffixation allows both agentive and instrumental readings
- model this by two lexical rules (underspecification is highly implausible here)
- lexicalized forms may have additional or alternative information

```
baker < (bake + er-agentive)
[SEM:LISZT < [ rel      [ /prof_rel > ]
               ARG1:x     INST:x ]
               ARG2:y ],
```

- possibly could treat profession as default consequence of rule

Compounds

- ‘Regular’ compounds: *wickerwork chair*, *garbage man*
- Restrictions: ?*blacksmith hammer* / *blacksmith’s hammer*
- Strange relations: *lavender honey*
- Novel uses: *apple juice chair*
- Pragmatics can override conventional meaning:

  In the back street where I grew up, everybody was poor. We were so poor that we never went on holiday. Our only toys were the garbage cans. We never built sandcastles, only garbage men. (Bauer 1983:86)

Figurative uses

bake and other similar verbs have extended uses:

(45) How the London Clay bakes in the sun!

(46) ‘I’m going to bake’, thought Julia.

(47) Don’t roast [the bottles of wine] in front of the fire.

(48) I’ve an idea that the critics will roast it.

(49) Why on earth did your Superintendent go and grill the poor girl?

(50) He broil’d with impatience.

Possibly all main cooking verbs can be used to mean *get very hot*, but further uses aren’t predictable.

roast	**bake**
beef, pork, chicken, head of lamb, camel, goat, !cow, lizard, crab, *ham, *gammon	bread, cake, pies
potato, chesnuts, turnip, crab apples, ?apple	ham, fish, snake, barracuda steak
coffee beans	apple, potato
duke	Gratin Dauphinois, pork chops with apples
metal ore	clay, concrete, earth
	pottery, resistance films

Partial regularity: *roast* is used for cooking birds, mammals or joints of meat. But also nuts, beans, potatoes and not ham, gammon etc.

Synchronically treat as conventional.

An alternative approach to polysemy: Schütze (1995)

- gradient rather than discrete representations
- acquisition of ambiguous forms
- disambiguation modelled by proximity relations

Halfway point summary

- Gone from "billiard ball" model of the lexicon to structured, hierarchical model, with senses/usages interrelated
- Sketched formalisation using typed (default) feature structures
- Suggested interactions with pragmatics for logical metonymy, implicit arguments, derived forms, compounds, figurative uses

Treating Metonymy Pragmatically: Hobbs

- We could do all metonymy in pragmatics.
- Hobbs *et al.* do this in *Interpretation as Abduction*.
- Abduction is inference to the best explanation.

$$\begin{array}{l} p \rightarrow q \\ \underline{q \qquad\quad} \\ p \end{array}$$

- Abduction in NLP:
  - We must provide an explanation of why the sentence would be true.

The Algorithm

- To interpret a sentence:
  - Prove the logical form of the sentence, together with the constraints that predicates impose on their arguments, allowing for coercions, Merging redundancies where possible, Making assumptions where necessary.
- Prove means:
  Prove via predicate logic the logical form of the sentence that has been derived from syntax
- Merging redundancies entails getting the best explanation.
- Making assumptions is the abduction bit.

The Role Abduction in Interpreting Utterances

- S and H have
  - their own beliefs
  - mutual beliefs
- An utterance lies at the edge of mutual belief and S's beliefs, and is an attempt to expand the set of mutual beliefs.
  - The bits in mutual belief are *old information*
  - The bits outside mutual belief are *new information.*
  - The bits outside mutual belief will require abduction in order to prove them.

A Simple Example

(51) The Boston office called.

Three problems:

1. Determining the relation between *Boston* and *office*.
2. Determining the reference for *the Boston office*.
3. Resolving the metonymy to *Someone at the Boston office...*

Interpreting (51)

- We must prove the LF via abduction.

(51′) $(\exists x, y, z, e)(call(e, x) \wedge person(x) \wedge rel(x, y) \wedge$
$office(y) \wedge Boston(z) \wedge nn(z, y))$

- There's an event e of a person x calling.
- x may not be the explicit subject, but it must be related to it or coercible from it, represented by $rel(x, y)$.
- y is an office which bears some unspecified relation nn to Boston.
- Abduction must be used to find out why $nn(z, y)$ and $rel(x, y)$ are true.

Example Continued: The Mutual KB

- $Boston(B_1)$
  $office(O_1) \wedge in(O_1, B_1)$
  $person(J_1) \wedge work\text{-}for(J_1, O_1)$
- If y is in z, then y and z are in a possible compound relation:
  $\forall y \forall z(in(y, z) \rightarrow nn(y, z))$
- If x works for y, then y can be coerced from x:
  $\forall x \forall y(work\text{-}for(x, y) \rightarrow rel(x, y))$

Proving the Logical Form

- Everything in the LF can be proved from the KB except $call(e, x)$
- Abduction permits us to assume this, so we do and add it to the mutual belief set.
- $call(e, x)$ is the new information.
- We could have assumed $person(x)$, rather than proving it with $person(J_1)$.
  This would have given
  the less specific reading of (51) that someone called, rather than John called.

The Proof Graph

Logical Form:

$\boxed{call'(e,x)} \quad \wedge\, person(x) \;\wedge\; rel(x,y) \wedge office(y) \;\wedge\; Boston(z) \;\wedge\; nn(z,y)$

Knowledge Base:

$person(J_1)$

$work\text{-}for(x,y) \;\supset\; rel(x,y)$

$work\text{-}for(J_1, O_1)$

$office(O_1)$

$Boston(B_1)$

$in(y,z) \;\supset\; nn(z,y)$

$in(O_1, B_1)$

The Three Pragmatics Problems

They are all solved as a by-product:

- The implicit relation in the compound nominal *Boston Office* is *in*.
- *The Boston Office* is resolved to O_1.
- The metonymy has been expanded to: *John, who works for the Boston office, called.*

Problems: Interaction with grammar

Syntax

- *Boston office* in this context behaves syntactically as a group noun:
  - The Boston office were calling (BrE).
  - The Boston office called. They have a crisis.
  - The Boston office called. ? It has a crisis.

Restrictions on metonymy and compounding are not represented:

- Possible metonymic relationships are restricted.
  - The Boston office is built of brick.
  - * John is built of brick. (with the same meaning)

Problems Continued: Organizing information

Logical forms

- You must be really careful to get the logical forms right.
  - You must have $call(e, x)$ and $person(x)$ rather than $call(e, y)$.

Non-Modularity

- Lexical knowledge, world knowledge and language heuristics are all represented at the same level.
- This enables a uniform inference regime.
- But it is unwieldy.

Problems Continued: Choice

- The problem of *which* inferences to make is *the* problem in pragmatics.
  - Eg., should we assume $person(x)$, or prove it with $person(J_1)$?
- Hobbs solves this by assigning costs to predicates, and guiding assumptions so that they have least cost.
- But there are problems with this:
  - You can't assign weights in a systematic fashion.
  - The weights have to be assigned in a context free fashion. This means that some of the context-sensitive aspects of inference can't be modelled.

A Hybrid Approach: Lexical Representation

Use Persistent Typed Default Unification (PDU)

- Works on typed default FSs.
  - These explicitly mark what information is default and what isn't.
- The operation is order independent.
- If a default bit of information survives PDU, it's still explicitly marked as default.

Some Examples of PDU

Where $\mathbf{t}' \sqsubset \mathbf{t}$ and $\mathbf{a} \sqcap \mathbf{b} = \bot$:

$$\begin{bmatrix} \mathbf{t} \\ \mathrm{F}: \top \end{bmatrix} \overset{<>}{\sqcap} \begin{bmatrix} \mathbf{t}' \\ \mathrm{F}: \mathbf{b} \end{bmatrix} = \begin{bmatrix} \mathbf{t}' \\ \mathrm{F}: \mathbf{b} \end{bmatrix} \qquad \text{Normal Unification}$$

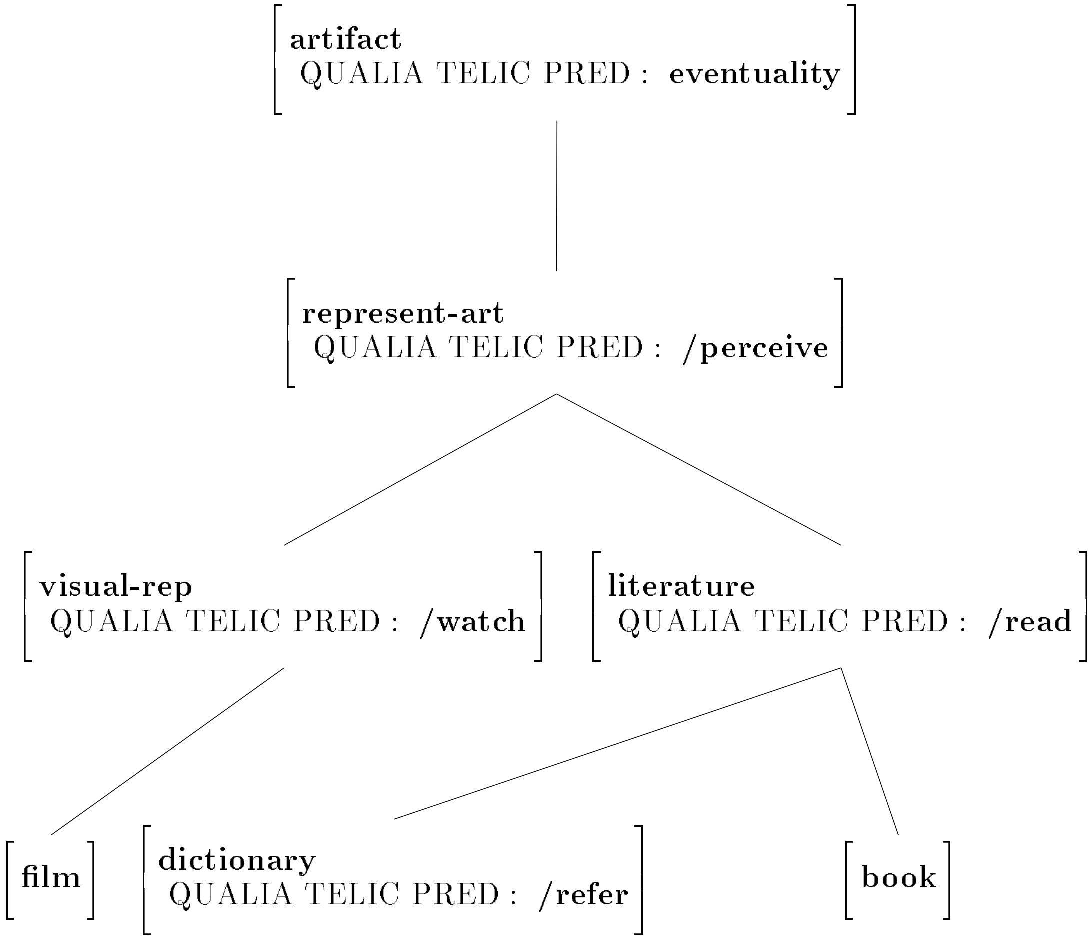
$$\begin{bmatrix} \mathbf{t} \\ \mathrm{F}: \mathbf{a} \end{bmatrix} \overset{<>}{\sqcap} \begin{bmatrix} \mathbf{t}' \\ \mathrm{F}: \mathbf{b} \end{bmatrix} = \bot \qquad \text{Conflicting non-defaults}$$

$$\begin{bmatrix} \mathbf{t} \\ \mathrm{F}: \top \end{bmatrix} \overset{<>}{\sqcap} \begin{bmatrix} \mathbf{t}' \\ \mathrm{F}: /\mathbf{b} \end{bmatrix} = \begin{bmatrix} \mathbf{t}' \\ \mathrm{F}: /\mathbf{b} \end{bmatrix} \qquad \text{Defeasible Modus Ponens}$$

$$\begin{bmatrix} \mathbf{t} \\ \mathrm{F}: \mathbf{a} \end{bmatrix} \overset{<>}{\sqcap} \begin{bmatrix} \mathbf{t}' \\ \mathrm{F}: /\mathbf{b} \end{bmatrix} = \begin{bmatrix} \mathbf{t}' \\ \mathrm{F}: \mathbf{a} \end{bmatrix} \qquad \text{Defeat of DMP}$$

$$\begin{bmatrix} \mathbf{t} \\ \mathrm{F}: /\mathbf{a} \end{bmatrix} \overset{<>}{\sqcap} \begin{bmatrix} \mathbf{t}' \\ \mathrm{F}: /\mathbf{b} \end{bmatrix} = \begin{bmatrix} \mathbf{t}' \\ \mathrm{F}: /\mathbf{b} \end{bmatrix} \qquad \text{Specificity/Penguin Principle}$$

The Telic Role of Artifacts



Enjoy

$$\left[\begin{array}{l} \textbf{coercing} \\ \text{SYN SUBCAT} : \left\langle \left[\begin{array}{l} \textbf{phrase} \\ \text{SYN:CAT} : \ \textbf{n} \\ \text{SEM} : \left[\begin{array}{l} \text{INDEX} : \boxed{\text{y}} \end{array}\right] \\ \text{QUALIA TELIC PRED} : \boxed{\text{P}} \end{array}\right] \right\rangle \\ \text{SEM} : \left[\begin{array}{l} \text{INDEX} : \boxed{\text{e}} \\ \text{LISZT} : \left\langle \left[\begin{array}{l} \text{PRED} : \ \textbf{\_enjoy\_rel} \\ \text{EVENT} : \boxed{\text{e}} \\ \text{ARG1} : \boxed{\text{x}} \\ \text{ARG2} : \boxed{\text{e}'} \end{array}\right], \left[\begin{array}{l} \text{PRED} : \boxed{\text{P}} \\ \text{EVENT} : \boxed{\text{e}'} \\ \text{ARG1} : \boxed{\text{x}} \\ \text{ARG2} : \boxed{\text{y}} \end{array}\right] \right\rangle \end{array}\right] \end{array}\right]$$

Enjoy the book

$$\left[\ \text{SEM}: \left[\begin{array}{l}\text{INDEX}: \boxed{\text{e}} \\ \text{LISZT}: \left\langle \left[\begin{array}{l}\text{PRED}: \ \textbf{\_enjoy\_rel} \\ \text{EVENT}: \boxed{\text{e}} \\ \text{ARG1}: \boxed{\text{x}} \\ \text{ARG2}: \boxed{\text{e}'}\end{array}\right], \right. \\ \qquad \left. \left[\begin{array}{l}\text{PRED}: \ \textbf{/read\_rel} \\ \text{EVENT}: \boxed{\text{e}'} \\ \text{ARG1}: \boxed{\text{x}} \\ \text{ARG2}: \boxed{\text{y}}\end{array}\right], \left[\begin{array}{l}\text{PRED}: \ \textbf{\_book\_rel} \\ \text{INST}: \boxed{\text{y}}\end{array}\right] \right\rangle\end{array}\right]\ \right]$$

$$enjoy(e, x, e') \wedge *read(e', x, y) \wedge book(y)$$

A Hybrid Approach: Representing Pragmatics (DICE)

The Rules:

- $A > B$: If A then normally B.
- `Goats don't read:` $goat(x) > \neg read(e, x, y)$
- `Elaboration:`
  $\langle \tau, \alpha, \beta \rangle \wedge subtype(e_\beta, e_\alpha) > Elaboration(\alpha, \beta)$

The Logic CE: Nonmonotonic $\approx\!\!|$

- DMP: one default applies,
  consequent nonmontonically inferred.
- Penguin Principle: conflicting defaults apply,
  consequent of most specific default inferred.
- Nixon Diamond: conflicting defaults apply,
  but none is most specific:
  no default conclusion inferred.
- $\beth(A, B)$ means $A \mathrel{|\!\!\approx} B$

Linking Lexicon to Pragmatics

KB_h: KB with all *-ed conditions removed.
(hard information only)

- `Defaults Survive:`
  $*\phi > \phi$
- `Discourse Wins:`
  $(*\phi \wedge \daleth(KB_h, \psi)) > \psi$

(29) John enjoyed the book.

a. $john(x) \wedge enjoy(e, x, e') \wedge book(y) \wedge$
$*read(e', x, y)$

b. $*read(e', x, y) > read(e', x, y)$
(`Defaults Survive`)

(33) The goat enjoyed the book.

a. $goat(x) \wedge enjoy(e, x, e') \wedge book(y) \wedge$
$*read(e', x, y)$

b. $*read(e', x, y) > read(e', x, y)$
(`Defaults Survive`)

c. $(*read(e', x, y) \wedge \sqsupset(KB_h, \neg read(e', x, y)) >$
$\neg read(e', x, y)$
(`Goats Don't Read, Discourse Wins`)

d. $\neg read(e', x, y)$
(Penguin Principle on b and c)

Some Things to Note

- Could have used a more general axiom than `Goats don't Read`.
- `Discourse Wins` is always going to be more specific than `Defaults Survive`.
- A lexical default is overridden only if there's enough monotonic information in the context to offer a conflicting alternative.

  So the discourse context must be 'informationally rich'.

Further Motivation for Having a Split between Lexical and Pragmatic Information

---

If all metonymy was handled by pragmatics, then:

- The Goat Rule would have to be really specific; how could it be derived?
  - `Enjoy Rule`
    $enjoy(e, e') \wedge object(e', x) \wedge book(x) > read(e')$
  - `The Goat Rule`
    $enjoy(e, e') \wedge object(e', x) \wedge agent(e', y) \wedge$
    $goat(y) \wedge book(x) >$
    $\neg read(e')$
- Spreading the load between
  pragmatics and the lexicon
  and having communication links
  allows us to 'loosen up' how we represent things.

Adjectives

- Don't want distinct lexical entries for *fast* in:
  - *fast typist*, *fast car*, *fast motorway*, and so on.
- Pustejovsky:
  Different sense for *fast* created via
  syntagmatic co-composition
  on a single entry for *fast*.
- Generalisation:
  An adjective like *fast* predicates over
  the telic role of the artifact.
- So DMP on `Defaults Survive`
  predicts
  - *fast car* is *car which drives fast.*
  - *fast typist* is *typist which types fast*
    (cf. *fast person..*)

Exceptions Triggered By Pragmatics!

(52) a. All the office personnel took part in the company sports day last week.
b. One of the typists was a good athlete, but the other was struggling to finish the courses.
c. The fast typist came first in the 100m.

Lexical reasoning:

- $typist(x) \wedge fast(e) \wedge *type(e, x)$

Pragmatic reasoning

- Assume DICE rules get *Elaboration*(52a-b,c).
- Then *The fast typist* must refer to a unique typist in (52a,b).
- The typists are differentiated only by athletic ability.
- So the uniqueness condition can be accommodated only if *fast* distinguishes their athletic ability.
- So $\sqsupset(KB_h, run(e,x))$
- So Penguin Principle yields $run(e,x)$.

Sentential Pragmatics vs. Discourse Pragmatics

(53) a. We went on holiday to a beach house last summer.
b. Everyone was so boring, because all they wanted to do was read.
c. Even Stephen King began a new novel just two days before we left.

(53c) in isolation of context:

- Lexical Processing: $*read > read$
- Pragmatics from KB_h: favours *write* because of WK that King is an author.
- Penguin Principle on `Defaults Survive` and `Discourse Wins` yields *write*.

(53c) in the Discourse Context

- Rules in DICE from KB'_h yield *Elaboration*(53b,53c).
- So (53c)'s event is a subtype of (53b)'s.
- So KB'_h gives *read* and DMP in DICE gives *read*.

(54) a. Germany entered the war.
b. Disraeli climbed the greasy pole of politics.
c. ?Disraeli entered the greasy pole of politics.

- Verbs of motion;
  arguments not of appropriate type.
- But the verbs in metaphorical meaning
  keep the properties that distinguish them from
  the other verbs (differentiae).
  - *enter* involves crossing from a
    (contextually-specified) nearby zone into a
    central zone.
  - Metaphorical *enter* keeps this.
    (54c) is odd, because you can't
    compute zones from *greasy pole of politics*.

Treating Metaphor

- You don't have underspecified types of arguments in motion verbs!
- You would lose lots of lexical information if you did this.
- Rather, you have lexical rules:
  - Metaphor lexical rule produces new entry with underspecified argument.
  - The result of this rule ensures that the differentiae remain the same.
  - But everything else is very underspecified, and has to be filled in by pragmatics.
  - You circumscribe what changes.
- Doing things this way allows you to keep conventionalised detail for motion verbs, but also gives the lexicon generative power.

Using Discourse To Interpret Metaphor

(55) John is a rock.

(56) ?Sam is a pebble.

(57) John is a rock.
But (compared to John) Sam is a pebble.

- Lexical Rule:
  Permits *rock* and *pebble* to take
  humans as argument.
- Metaphorical use of *rock* is ‘conventionalised’.
- It helps us compute what *pebble* means in (57).

Using Discourse Structure to Interpret *Pebble*

- Aim for maximum *Contrast* (SDRT).
  - Structure: *pebble* mapped to *rock*.
  - Semantics: *pebble* contrasts *rock*.
- So if *rock* means *reliable*,
  pebble means *unreliable*.
- Truth conditional revision triggered by *Contrast*.
- Infer *Sam is not a rock*
  (whatever *rock* means here).

Discourse Context can Trigger Metaphor!

(58) I have climbed up that greasy pole.

- I climbed a physical, greasy pole, because unification on CoL verb *climb* and **physical-space** *greasy pole* succeeds.

(59) a. I have always despised politics.
 b. But I have climbed to the top of that greasy pole.

- Anaphor *that greasy pole* must be identified with *politics*.
- So it's not a physical space.
- So unification on CoL verb and **physical space** *greasy pole* doesn't occur.
- Metaphorical *climb* requires a scale with top and bottom; career ladder of politics is appropriate.

Disambiguation in Discourse

---

(60) a. The judge asked where the defendant was.
b. The barrister apologised, and said he was at the pub across the street.
c. The court bailiff found him slumped underneath the bar.
c′. But in fact, the court bailiff found him slumped underneath the bar.

Hobbs:

- Abduce the ‘most likely’ sense of *bar* from rules representing:
  - World knowledge;
  - syntactic constraints;
  - compositional semantic constraints.

Statistics

- Use a corpus to create *neighbourhoods* for each word sense:
  - Group of words that occur most frequently within the vicinity of the word sense.
- Use this to guess which sense is being used in text.

Approaches to Disambiguation

Statistics:

- Ignores discourse structure.
- So can't distinguish (60a,b,c) vs. (60a,b,c$'$).

Abduction: Hobbs *et al.*

- Might work, if the discourse relations are assigned appropriate spatial constraints.
- But it's not clear how to do this,
  since some of them are default constraints,
  and these must be represented as
  weighted predicates to be circumscribed.

An Alternative: Avoid Incoherence

- From Context:
  expectation that
  the defendant isn't in the court
  when the event in (60b) occurs.
- From pragmatics:
  only candidate discourse relation
  for attaching (60c) to (60b) is *Narration*.
- Narrative can't violate expectations.
- So in spite of information in (60c),
  the bar isn't the courtroom bar,
  because otherwise the discourse is incoherent.

Things to Note

1. We need a model of rhetorical structure and the constraints the relations impose on coherent discourse.
2. There is knowledge conflict which is resolved.
3. Present theories of lexical disambiguation that take only world knowledge into account can't handle (60).
4. The above account used the *particular* coherence constraints imposed by narrative.

bar$_1$: courtroom bar.
bar$_2$: pub bar.

- Sentence Implicature:
  - `Bar Rule:` $FS_{bar} > FS_{bar_1}$
- Discourse Coherence Constraint:
  - If attaching β to α leads to
    at best weak discourse coherence,
    when a piece of information β' is assumed,
    then *whatever* the monotonic information
    about α and β,
    we try to infer $\neg\beta'$
    (so as to get better coherence).
- FS_{bar_1} leads to discourse incoherence.
- If we formalise the above constraint,
  we should be able to infer $\neg FS_{bar_1}$,
  in spite of the sentential implicature.

Formalising the Constraint

- **`Interpretation Constraint:`**

  (a) $(\langle\tau, \alpha, \beta\rangle \wedge \mathit{Info}_h(\alpha, \beta)$

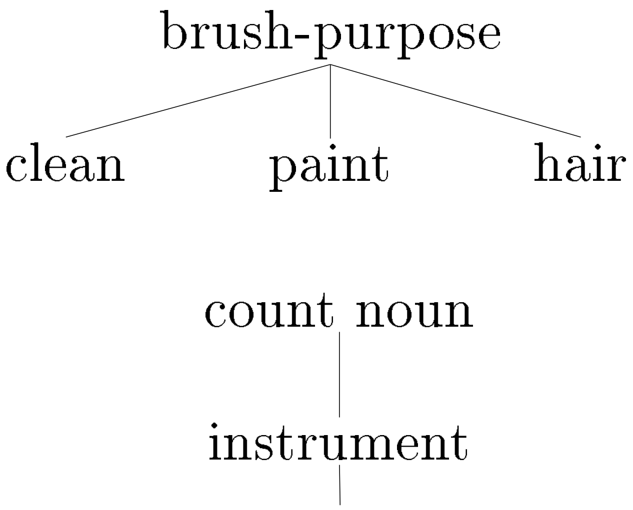
  (b) $\beth_{KB}(\beta', \mathit{weak}(\tau \cup \beta)))$

  (c) $> \neg\beta'$

For (60a,b,c), let: α be (60b)
β be (60c)
β' be FS_{bar_1}

- (a) holds.
- $\mathit{Info}_h(\alpha, \beta)$ entails FS_{bar}.
- FS_{bar_1} leads to discourse incoherence.
- So $\beth_{KB}(\beta', \mathit{weak}(\tau \cup \beta)))$ holds.
- So Interpretation Constraint Applies.
- It's more specific than the Bar Rule.
- By the Penguin Principle, we get $\neg FS_{bar_1}$.

(61) Nylon bristles make cheap brushes.



$$\begin{bmatrix} \textbf{brush} & \\ \text{qualia:telic:pred}: & \textbf{/brush-purpose} \end{bmatrix}$$

(62) The janitor learned how to use a brush.

(63) $brush(y) \wedge instrument(e, y) \wedge$
$agent(e, x) \wedge purpose(y, brush\text{-}purpose(e'))$

- ‘janitor’ > cleaning brush

The Interpretation Constraint and Sense Shift

(64) a. All the school staff learned how to paint.
  b. The janitor learned how to use a brush.

- Suppose purpose of brush is to clean.
- Then DICE yields *Narration*.
- Coherence Constraint: distinct common topic.
  - Topic is *school staff learn to do things*.
  - It's weak.
- So Interpretation Constraint verified when β' is $purpose(y, clean(e'))$.
- So Penguin Principle yields $\neg purpose(y, clean(e'))$.
- The Interpretation Constraint rules out every possible interpretation of *brush* except *paint*.

- So *paint* is inferred.
- So *Elaboration*(64a,64b) is inferred.

Zeugma

- Zeugma is the word pun effect produced, when you inappropriately link two things together.

  (65) Bank robbers and dam busters blow up banks.

- Presence of zeugma has been argued to justify multiple lexical entries, but this leads to problems:

(61) Nylon bristles make cheap brushes.

(66) Rembrandt and our janitor used a brush.

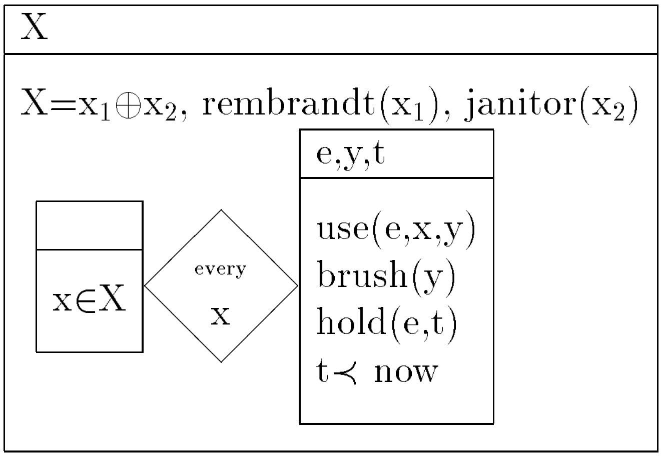
(67) Thighs are the most muscular part of the vertebrate hind limb.

(68) Bill likes mustard on his thighs, but John prefers sun tan lotion.

Why *Brush* in (66) isn't General

(66) Rembrandt and our janitor used a brush.

(66′)



World Knowledge >-rules

- 'Rembrandt' > 'paint brush'
  'Janitor' > cleaning brush.

Coherence of Coordination

We need to check coherence of coordination.
Do this by computing a discourse relation between the relevant propositions.

- *Parallel* with coherence constraints (Asher):
  - Common theme is **humans** use a painting or cleaning brush.
  - This is bad (Quantity)
  - So only weak coherence from World knowledge

The Interpretation Constraint

- **`Interpretation Constraint`**
  (a) $\langle \tau, \alpha, \beta \rangle \wedge Info(\alpha, \beta)$
  (b) $\sqsupset_{KB}(\beta', weak(\tau \cup \beta)))$
  (c) $\quad > \neg\beta'$

- Let β' be
  Rembrandt's brush is a paint brush and the janitor's brush is a cleaning brush.
- Then the Interpretation Constraint is verified.
- So there's irresolvable conflict between:
  - Interpretation Constraint
  - Rembrandt WK
  - Janitor WK
- (66) is odd because it invokes irresolvable *pragmatic* ambiguity.

Conclusion

- Lexical semantic information can be modelled using a typed feature structure formalism, extended to handle defaults.
- This allows us to represent default lexical semantic generalisations.
- But these defaults can be overridden by pragmatic information.

Further information and resources

- references in bibliography
- some pointers to other papers via
  `http://www-csli.stanford.edu/users/aac/`
- corpora, MRDs, etc etc
  `http://kinks.phil.cmu.edu/manning/statnlp.html`