Integrating the lexicon and pragmatics

Ann Copestake
aac@csli.stanford.edu
CSLI, Stanford University

Alex Lascarides
alex@cogsci.ed.ac.uk
University of Edinburgh

Overview

Lecture 1

- Informal introduction

Lecture 2

- Use of typed feature structures
- Default unification
- Contrasting constructional polysemy and sense extension

Lecture 3

- Sense extension
- Discourse in Commonsense Entailment

Lecture 4

- Formalizing the interface between pragmatics and the lexicon
- Metaphor

Lecture 5

- Disambiguation
- Zeugma

Overview of lecture 1

- Main points of the theory
- Radical pragmatics and radical lexicalism
- Informal examples of how integrated theory works (details will be covered in subsequent lectures)

Lexical semantics and pragmatics: the theory

- Some phenomena require a theory which integrates the lexicon and pragmatics
- Lexical semantics is the interface between world knowledge and linguistic knowledge/processing
- Lexical (and other pure linguistic) processing is informationally encapsulated using minimal 'taxonomic' lexical semantic representations
- Pragmatic reasoning is open-ended on arbitrary world knowledge
- Lexicon proposes, pragmatics (may) dispose

Overview of formalization

- Lexical semantics can be represented using (typed) feature structures augmented with defaults and fully integrated into a syntactic theory (e.g. HPSG)
- Pragmatic representation integrated with Lascarides and Asher's DICE
- Pragmatic representation uses a nonmonotonic logic

Lexical Representation (simplified!)

$$\begin{bmatrix} \textbf{common-noun} \\ \text{ORTH} : \textbf{girl} \\ \text{SYN} : \textbf{nounsyn} \\ \text{SEM} : \begin{bmatrix} \text{INDEX} : \boxed{x} \begin{bmatrix} \text{GENDER} : \textbf{fem} \end{bmatrix} \\ \text{LISZT} : \left\langle \begin{bmatrix} \textbf{\_girl\_rel} \\ \text{COUNT} : \textbf{true} \\ \text{INST} : \boxed{x} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

```
< ORTH > = girl
< > = common-noun
< SEM LISZT FIRST > = _girl_rel
< SEM INDEX GENDER > = fem
< SEM LISZT FIRST COUNT > = true.
```

Much of the infomation is inherited from **common-noun**

Examples of types

$$\begin{bmatrix} \textbf{common-noun} \\ \text{ORTH} : \textbf{string} \\ \text{SYN} : \textbf{nounsyn} \\ \text{SEM} : \begin{bmatrix} \text{INDEX} : \boxed{x}\ \textbf{index} \\ \text{LISZT} : \left\langle \begin{bmatrix} \textbf{noun\_rel} \\ \text{COUNT} : \textbf{boolean} \\ \text{INST} : \boxed{x} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

$$\begin{bmatrix} \textbf{index} \\ \text{PERSON} : \textbf{per} \\ \text{NUMBER} : \textbf{num} \\ \text{GENDER} : \textbf{gen} \end{bmatrix}$$

Lexical Representation (cont.)

- Types are in bold font. Types encode generalizations
- Reentrancy (coindexation) is indicated by boxed letters or integers

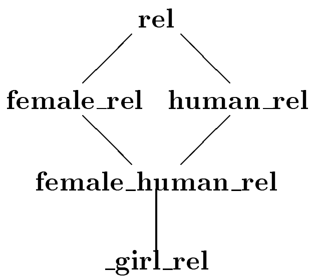
- Assume sign-based lexicalist theory, where feature structure (FS) includes orthography, syntax and semantics
- Only part of FS is shown, details are simplified
- Semantic representation has to support construction of logical form (MRS - equivalent to $\lambda x[girl(x)]$)
- Predicates corresponding to words are indicated by preceding underscore

Adding semantic constraints

$$\begin{bmatrix}\textbf{female\_rel} \\ \text{INST} : \begin{bmatrix}\text{GENDER} : (\textbf{neut} \vee \textbf{fem})\end{bmatrix}\end{bmatrix}$$

$$\begin{bmatrix}\textbf{human\_rel} \\ \text{COUNT} : \textbf{true}\end{bmatrix}$$

$$\begin{bmatrix}\textbf{female\_human\_rel} \\ \text{INST} : \begin{bmatrix}\text{GENDER} : \textbf{fem}\end{bmatrix}\end{bmatrix}$$




```
< ORTH > = girl
< > = common-noun
< SEM LISZT FIRST > = _girl_rel.
```

Overview of types

- Types label nodes in typed feature structures (TFS)
- Type hierarchy (bounded complete partial order) indicates which types can unify
- Types may introduce constraints, represented as TFSs
- Inheritance operates over hierarchy
- Multiple inheritance is possible
- Usually, inheritance is taken to be monotonic
- TFSs are used in HPSG — lexical semantic representation can use standard mechanisms

Hierarchy is motivated by needs of grammar, not real world knowledge representation

Default semantic constraints

$$\begin{bmatrix}\textbf{human\_rel} \\ \text{COUNT} : \textbf{/true}\end{bmatrix}$$

$$\begin{bmatrix}\textbf{female\_human\_rel} \\ \text{INST} : \begin{bmatrix}\text{GENDER} : \textbf{/fem}\end{bmatrix}\end{bmatrix}$$

```
< ORTH > = M\"adchen
<> = common-noun
< SEM INDEX GENDER > = neut
< SEM LISZT FIRST > = _M\"adchen_rel.
```

Overview of defaults

- Formalism introduced by Lascarides et al (1994/1996) as modified in Lascarides and Copestake (1995)
- Default information indicated by slash
- Defaults can be introduced by types
- Defaults are useful for morphology and syntax as well as semantics
- Discussed in detail in Lecture 2

Polysemy: some terminology

homonymy
bank:financial institution vs *bank*:mound

unpredictable polysemy
bank:tilt of plane vs *bank*:mound

regular polysemy
bank:to construct a *bank*:mound
bank:building used by *bank*:financial institution

'Dictionary' model of polysemy

bank 1 N financial institution

bank 2 N mound of earth

bank 3 N tilt of plane

bank 4 V to tilt a plane

...

- Relatedness of senses is not captured — meaning postulates are inadequate because:
  - senses of same word are not distinguished as such
  - grammatical effects
- Cannot list all potential usages
- Notion of discrete sense is dubious in many cases
- Cannot map senses between dictionaries
- NOT really the model found in dictionaries

Lexical rules

- Conventionally used for verb alternations, conversion, (derivational) morphology
- Map one sense into another
- Allow for morphological, syntactic and semantic effects

Lexical rules using typed feature structures:

- Use of types allows natural representation of classes
- Usually, little need be stipulated apart from input and output type

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

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

$$\begin{bmatrix} \textbf{creation-tverb} & \\ \text{ORTH}: \textbf{paint} & \\ \text{SYN}: & \begin{bmatrix} \text{CAT}: \textbf{v} \\ \text{SUBJ}: \textbf{NP}_{\boxed{x}} \\ \text{SUBCAT}: \langle \textbf{NP}_{\boxed{y}} \rangle \end{bmatrix} \\ \text{SEM}: & \begin{bmatrix} \text{INDEX}: \boxed{e} & \\ \text{LISZT}: & \left\langle \begin{bmatrix} \textbf{\_paint\_rel} \\ \text{EVENT}: \boxed{e} \\ \text{ARG1}: \boxed{x} \\ \text{ARG2}: \boxed{y} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

Floyd painted a picture

$$\begin{bmatrix} \textbf{benef-PP-verb} & \\ \text{ORTH}: \textbf{paint} & \\ \text{SYN}: & \begin{bmatrix} \text{CAT}: \textbf{v} \\ \text{SUBJ}: \textbf{NP}_{\boxed{x}} \\ \text{SUBCAT}: \langle \textbf{NP}_{\boxed{y}}, \textbf{PP(for)}_{\boxed{z}} \rangle \end{bmatrix} \\ \text{SEM}: & \begin{bmatrix} \text{INDEX}: \boxed{e} & \\ \text{LISZT}: & \left\langle \begin{bmatrix} \textbf{\_paint\_rel} \\ \text{EVENT}: \boxed{e} \\ \text{ARG1}: \boxed{x} \\ \text{ARG2}: \boxed{y} \end{bmatrix}, \begin{bmatrix} \textbf{benef\_rel} \\ \text{EVENT}: \boxed{e} \\ \text{ARG1}: \boxed{z} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

Floyd painted a picture for Seymour

The argument for radical pragmatics

1. Not all words/phrases can be conventionally/compositionally associated directly with limited number of specific meanings.

   (1) The Boston office called

2. Lexical rules do not allow for novel types of usage

   (2) The ham sandwich is at table 2

3. Reasoning about discourse situation is needed to resolve referent for novel cases
4. Since pragmatics is needed for interpretation in novel cases, it is unnecessary to assume any contribution of the lexicon in analogous established cases

BUT ...

If pragmatics only has access to logical form (i.e. denotations) there are problems:

1. Grammatical effects. E.g. *Boston office* in this context behaves syntactically as a group noun:

   (3) a The Boston office call every day (sg or pl agr possible in BrE)
   b The Boston office called. They are going on strike.
   c The Boston office called. ? It is going on strike.

2. Conventionalization. E.g. no metonymy from people to places

   (4) a The Boston office is built of brick.
   b * John is built of brick. (with the same meaning)

3. Language specificity

Therefore need both pragmatics and grammar/lexicon (or do away with distinction).

Radical pragmatics and time expressions

(5) I'll see you Tuesday

General pragmatic process allowing dates etc to appear as temporal locations? No, because:

1. Grammatical effects: *Tuesday* is adverbial above

   (6) * I'll see you Prague

2. Overgeneration

   (7) * I'll see you September

3. Dialect specificity ((5) is not good in many dialects of BrE)

   (8) a $ I'll see you Tuesday
   b £* I'll see you Tuesday

Sense extension

- 'Lexical rules' used to represent conventional aspects
- So new typed feature structures (TFSs) are generated
- Pragmatics may have to instantiate extended use further (also true for 'normal' words)
- Default information on output may be contextually or pragmatically overridden

Some examples of sense extension:

1. Noun-to-verb conversion
2. Grinding
3. Metaphor
4. Compound nouns

More terminology

nonce not recorded in the lexicon, but (usually) can be produced/interpreted by generative devices

institutionalised recorded in the lexicon as derived by a regular process

lexicalised either unrelated to other lexical entries or idiosyncratically augment or override regularly derived information

established covers **institutionalised** and **lexicalised**

Noun-to-verb conversion

Established examples:
shovel, *hammer*, *canoe*, *mound*, *sugar*, *tango*
Established patterns:
vehicle -> travel by vehicle
dance -> perform dance
paint-like substance -> apply substance
Productive use of established pattern:
skateboard, *Ronseal*
Nonce use, with non-established pattern:

(9) Max teapotted a policeman

(Clark and Clark, 1979) in context

(10) Max rubbed the back of the leg of the policeman with a teapot

(11) Have you vodkaed the water bottle?

(12) Have you rinsed the water bottle with vodka in order to dimish the unpleasant taste/odour that would result if it were left damp?

Noun-to-verb conversion (cont.)

- Lexical rules (or other such device) are minimally needed to get something which is syntactically a verb
- Lexical rules could give underspecified result
- Meaning of utterance in context can be much richer, but this is true of utterances in general

  (13) Bill lives close to Ben

  What distance is *close*? Depends on context.
- Meaning of verbs formed with established patterns can be understood in less marked contexts than completely nonce usages
- Established patterns should be represented by individual lexical rules inheriting from general rule
- Pragmatics should fill in the meaning from context

Defeasible aspects of noun-to-verb conversion

Established usages may have wider usage than predicted by rule:

(14) a Kim hammered the nail into the wall.

b Kim hammered the nail into the wall with his shoe.

c Kim hammered the nail into the wall. It took a long time because he had to use his shoe.

This (usually) doesn't happen with non-established uses:

(15) a Kim umbrellaed Sandy

b * Kim umbrellaed Sandy with a walking stick

(16) ... the sledgehammer simply swung out of nowhere, clotheslining him ...

Grinding

Count to mass for physical objects
Established examples:
turkey, *haddock*, *oak*, *mink*
Established patterns:
animal -> meat
animal -> skin/fur
tree -> wood
Productive use of established pattern:
elk, *ostrich*
General use (in context of Universal Grinder):

(17) After putting two tables through the grinding machine, there was table all over the floor

Grinding (cont.)

- 'Quantificational' uses of nouns as mass terms are probably different

  (18) There's more puddle than pavement

- Language specificity, some languages don't have conventional meat grinding, others mark it morphologically
- Nonce usages don't work cross-linguistically

  (19) We were up to our ankles in canal

- Hierarchy of lexical rules converting count to mass with established cases inheriting from general rule
- Meaning for general case is specialised for particular cases
- Discussed in more detail in Lecture 3

Metaphor

(20) a Kim is a sheep

b Kim is a wombat
He/*it thinks Prague is in Austria

c This audience is a flock of sheep

d ? This audience is a sheep

- Words in metaphorical usages keep some properties of their literal use
- Some metaphor patterns such as animal to human can be treated as established
- Metaphor lexical rule produces new entry of right type
- But semantics is very underspecified, and more information has to be filled in by pragmatics
- Presumably general cognitive processes influence conventionalisation

Compounds

Conventional aspects:

- Some compounds are established: *garbage man* / *dustbin man* (but not ?*rubbish man*)
- Others fall into regular patterns: *wickerwork chair*
- There are restrictions on compounds: ?*blacksmith hammer* / *blacksmith's hammer*

Pragmatic aspects:

- Strange relationships: *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)

Treatment of compounds

- Established uses have their own entries
- Hierarchy of schemata give possible interpretations, established entries usually inherit from schema
- Pragmatics may instantiate or override default interpretation

- Compound schemata are intermediate between lexical rules and grammar rules / HPSG schemata
- cf Construction grammar (Fillmore and Kay, in press) and Sag (1996) on relative clauses

Constructional polysemy

- A single sense assigned to a lexical entry is contextually specialised
- Implemented by underspecification
- Lexical entries may be multi-faceted
- Default information may be contextually or pragmatically overridden

Some examples of constructional polysemy:

1. General senses with specific subcases
2. Multiple aspects of meaning
3. Sense broadening
4. Adjectival modification
5. Implicit arguments

Reel

- Atkins and Levin (1992) investigated senses in traditional dictionaries and found many disparities
- We assume single general sense for *reel* as **container**, covering use for cotton reel, film reels etc

(21) Sometimes reels are made of plastic

- Contents are underspecified but could be specialized syntagmatically (e.g. *reel of film*) or pragmatically
- Sense extension may apply to specialized use only

(22) The mystery is only resolved in the final reel

Thesis, etc

- Dual nature: abstract entity and physical object
- These usages are not mutually exclusive

(23) a The books on the top shelf are about syntax
   b That thesis has thousands of pages and is quite unreadable

(we will discuss copredication in more detail in a later lecture)

$$\begin{bmatrix} \textbf{common-noun} \\ \text{ORTH} : \textbf{book} \\ \text{SYN} : \textbf{nounsyn} \\ \text{SEM} : \begin{bmatrix} \text{INDEX} : \boxed{x} \begin{bmatrix} \text{QUALIA} : \begin{bmatrix} \text{CST} : \textbf{repres.} \\ \text{FML} : \textbf{physobj} \\ \text{TLC} : \textbf{/read\_rel} \\ \text{AGV} : \textbf{/write\_rel} \end{bmatrix} \end{bmatrix} \\ \text{LISZT} : \left\langle \begin{bmatrix} \textbf{\_book\_rel} \\ \text{INST} : \boxed{x} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

Qualia structure

- Part of Pustejovsky's (1991, 1995) Generative Lexicon theory
- 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

Sense broadening

Basic sense denotes form and composition — modulated sense composition determined by (local) context:
cloud: *cloud of mosquitoes*, *dust cloud*

(24) a There are clouds in the sky — it's going to rain.
   b There was a cloud of flies round the cow.
   c The flies were pestering the horse — it swished its tail at the buzzing cloud.

$$\begin{bmatrix} \text{QUALIA} : \begin{bmatrix} \text{CST} : \textbf{phys\_cum / water-vapour} \\ \text{FML} : \textbf{amorphous} \end{bmatrix} \end{bmatrix}$$

Adjectival modification

- Meaning depends on the noun as well as the adjective:

(25) a fast road / car / driver / typist

- Pustejovsky argues that meaning arises by syntagmatic *co-composition*: adjectives select for particular qualia
- For example, *fast* always selects for an event
- *fast typist* means *typist who types fast*
- However, depends on context:

(26) There was a company sports day last week. One of the typists was a good athlete, but the other was very unfit. The fast typist came first in the 100m.

An account of adjective co-composition

Logical form for *fast typist*:

$$[x][\mathrm{typist}(x) \wedge \mathrm{fast}(e) \wedge /\mathrm{type}(e, x)]$$

Semantics and telic role for *typist*:

$$[x][\mathrm{typist}(x : \textsc{qualia telic} = /\mathrm{type})]$$

Semantics of basic type **adjective**:

$$[x][\text{adj-pred}(w) \wedge P(w, x)]$$

For **telic-adjectives**, P is instantiated by the telic predicate.

Grammar rules are not affected: e.g. with MRS semantic composition will involve appending LISZTs as usual

fast typist

$$\begin{bmatrix} \textbf{phrase} \\ \text{ORTH}: \langle \textbf{fast typist} \rangle \\ \text{SEM}: \begin{bmatrix} \text{INDEX}: \boxed{x} \begin{bmatrix}\text{QUALIA: TLC}: \boxed{P}\ \textbf{/\_type\_rel}\end{bmatrix} \\ \text{LISZT}: \boxed{1} \left\langle \begin{bmatrix}\text{PRED}: \textbf{\_typist\_rel} \\ \text{INST}: \boxed{x}\end{bmatrix}\right\rangle \oplus \\ \boxed{2}\left\langle \begin{bmatrix}\text{PRED}: \textbf{\_fast\_rel}\\ \text{INST}: \boxed{e}\end{bmatrix}, \begin{bmatrix}\text{PRED}: \boxed{P} \\ \text{EVENT}: \boxed{e} \\ \text{ARG1}: \boxed{x} \\ \text{ARG2}: \textbf{ind}\end{bmatrix}\right\rangle \end{bmatrix} \\ \text{DTRS}: \begin{bmatrix} \text{HEAD-DTR}: \begin{bmatrix} \textbf{common-noun} \\ \text{ORTH}: \langle \textbf{typist}\rangle \\ \text{SEM}: \begin{bmatrix}\text{INDEX}: \boxed{x} \\ \text{LISZT}: \boxed{1}\end{bmatrix}\end{bmatrix} \\ \text{COMP-DTRS HD}: \begin{bmatrix}\textbf{telic-adjective} \\ \text{ORTH}: \langle\textbf{fast}\rangle \\ \text{SEM LISZT}: \boxed{2}\end{bmatrix}\end{bmatrix}\end{bmatrix}$$

Implicit arguments

(27) a Kim bakes
b Kim baked on Tuesday
c Sixty potatoes were needed for the party. Kim was baking all afternoon.

Examples are interpreted as having an implicit (narrow-scope) existential bakee. By default, this is bread, cakes, pastries etc.
Similarly, intransitive *drink* implies alcoholic beverage, *eat* implies a meal (Fillmore, 1986).

$$\begin{bmatrix} \text{SYN}: \begin{bmatrix}\text{SUBCAT}: \langle \textbf{NP(opt)}\boxed{y}\,\text{SEM:LISZT}\langle\textbf{/meal\_rel}\rangle\rangle\end{bmatrix} \\ \text{SEM}: \begin{bmatrix}\text{LISZT}: \left\langle \begin{bmatrix}\textbf{eat\_rel} \\ \text{EV}: \boxed{e} \\ \text{ARG1}: \boxed{x} \\ \text{ARG2}: \boxed{y}\end{bmatrix}\right\rangle\end{bmatrix}\end{bmatrix}$$

Note: default may be pragmatically conditioned but 'Were you guzzling?' doesn't imply a meal, even though (presumably) **guzzle**$'$ implies **eat**$'$

Why not lexical rules?

Broadening and implicit arguments could be treated with lexical rules. But:

- danger of spurious ambiguity
- no change in sense
- default account is natural explanation of why context is necessary to get broader meaning

Summary of Terminology

- Homonomy vs polysemy
- Conventional vs pragmatic
- Lexicalized, established and nonce senses
- Metonymy and metaphor
- Constructional polysemy vs sense extension

Next lecture

1. Quick review of typed feature structures
2. Order Independent Typed Default Unification
3. Contrasting constructional polysemy and sense extension

Lecture 2

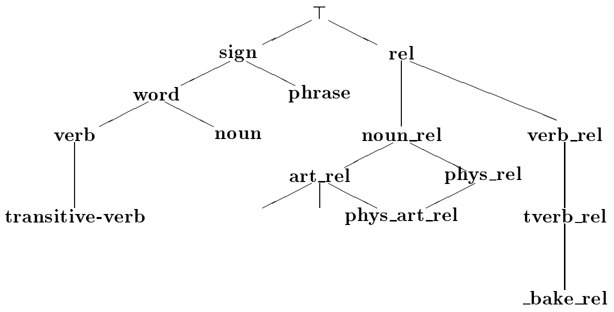
1. Quick review of typed feature structures
2. Order Independent Typed Default Unification
3. Contrasting constructional polysemy and sense extension

Typed feature structures

- Hierarchy of types is a bounded complete partial order
- Most general type is top ($\top$)
- $\sqsubset$ means "is more specific than"
- $\sqcap$ is join, $\sqcup$ is meet
- bottom ($\bot$) indicates inconsistency
- Typed feature structures only unify if types are compatible
- Overload operators:

  unification $\sqcap$
  generalisation $\sqcup$
  is subsumed by $\sqsubset$
  failure $\bot$

Example of a type hierarchy



Examples:

- **word** $\sqsubset$ **sign**
- **word** $\sqcap$ **sign** $=$ **word**
- **word** $\sqcap$ **tverb_rel** $= \bot$
- **art_rel** $\sqcap$ **phys_rel** $=$ **phys_art_rel**
- **art_rel** $\sqcup$ **tverb_rel** $=$ **rel**

Operations on typed feature structures

Where $\mathbf{u} \sqsubset \mathbf{t}$ and $\mathbf{a} \sqcap \mathbf{b} = \bot$:

$$\begin{bmatrix} \mathbf{t} \\ \mathrm{F} : \mathbf{a} \\ \mathrm{G} : \top \end{bmatrix} \sqcap \begin{bmatrix} \mathbf{u} \\ \mathrm{F} : \boxed{1} \\ \mathrm{G} : \boxed{1} \end{bmatrix} = \begin{bmatrix} \mathbf{u} \\ \mathrm{F} : \boxed{1}\,\mathbf{a} \\ \mathrm{G} : \boxed{1} \end{bmatrix}$$

$$\begin{bmatrix} \mathbf{u} \\ \mathrm{F} : \boxed{1}\,\mathbf{b} \\ \mathrm{G} : \boxed{1} \end{bmatrix} \sqsubset \begin{bmatrix} \mathbf{u} \\ \mathrm{F} : \boxed{1} \\ \mathrm{G} : \boxed{1} \end{bmatrix}$$

$$\begin{bmatrix} \mathbf{u} \\ \mathrm{F} : \boxed{1}\,\mathbf{b} \\ \mathrm{G} : \boxed{1} \end{bmatrix} \sqsubset \begin{bmatrix} \mathbf{u} \\ \mathrm{F} : \mathbf{b} \\ \mathrm{G} : \mathbf{b} \end{bmatrix}$$

$$\begin{bmatrix} \mathbf{t} \\ \mathrm{F} : \mathbf{a} \\ \mathrm{G} : \mathbf{b} \end{bmatrix} \sqcup \begin{bmatrix} \mathbf{u} \\ \mathrm{F} : \boxed{1}\,\mathbf{b} \\ \mathrm{G} : \boxed{1} \end{bmatrix} = \begin{bmatrix} \mathbf{t} \\ \mathrm{F} : \top \\ \mathrm{G} : \mathbf{b} \end{bmatrix}$$

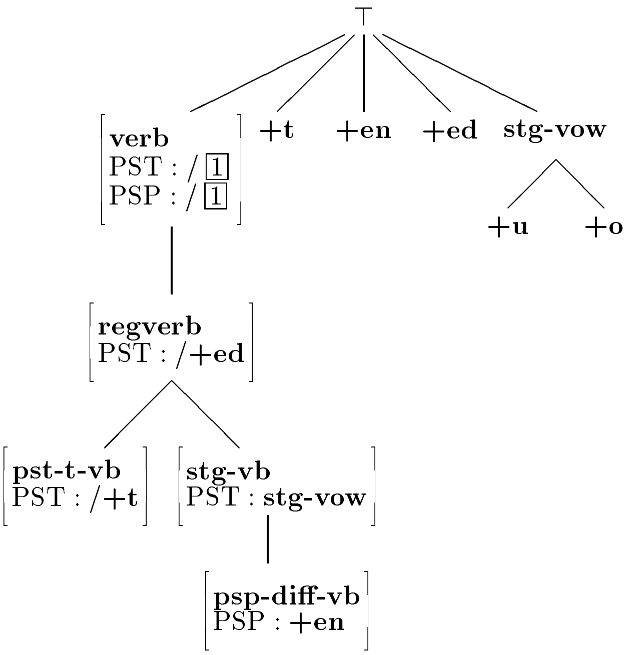
Order Independent Typed Default Unification

- Defaults are desirable both for capturing generalizations about the organization of the lexicon and grammar (abbreviatory defaults) and for interation with pragmatics (persistent defaults)
- Unification on TFSs is declarative and order-independent, similar properties desired for default unification
- Order-independence is a necessary property for defaults that interact with pragmatics because cannot prespecify order of unifications
- Compared to e.g. DATR situation is complicated by reentrancy

Previous work

- Kaplan (1987) — informal description
- Bouma (e.g. 1992)
- Carpenter (1991/1993) (and also Calder and van den Berg and Prüst) — asymmetric default unification as maximal incorporation as information
- various other definitions (overview in Copestake, 1991/1993)
- Young and Rounds (1993)
- Lascarides et al (1994/1996) — PDU
- Lascarides and Copestake (1995) — YADU

Lexical defaults



Desiderata

1. Non-default information can be distinguished from default information and is always preserved.
2. Default unification behaves like monotonic unification in the cases where monotonic unification would succeed.
3. Default unification returns a single result, deterministically.
4. Defaults can be given a precedence ordering such that more specific information overrides less specific information.
5. Default unification can be described using a binary, order independent (i.e., commutative and associative) operation.
6. Default unification never fails unless there is conflict in non-default information.

Carpenter's Asymmetric Default Unification: $\overset{<}{\sqcap}$

- LHS TFS is indefeasible (non-default), RHS is defeasible (default)
- Result is LHS plus maximal consistent information from RHS
- Credulous version ($\overset{<}{\sqcap}_c$) gives disjunction, skeptical ($\overset{<}{\sqcap}_s$) is the generalization of disjuncts

$$\begin{bmatrix} \mathrm{F} : \boxed{1} \\ \mathrm{G} : \boxed{1} \end{bmatrix} \overset{<}{\sqcap}_c \begin{bmatrix} \mathrm{F} : \mathbf{a} \\ \mathrm{G} : \mathbf{b} \end{bmatrix} = \left\{ \begin{bmatrix} \mathrm{F} : \boxed{1}\,\mathbf{a} \\ \mathrm{G} : \boxed{1} \end{bmatrix}, \begin{bmatrix} \mathrm{F} : \boxed{1}\,\mathbf{b} \\ \mathrm{G} : \boxed{1} \end{bmatrix} \right\}$$

$$\begin{bmatrix} \mathrm{F} : \boxed{1} \\ \mathrm{G} : \boxed{1} \end{bmatrix} \overset{<}{\sqcap}_s \begin{bmatrix} \mathrm{F} : \mathbf{a} \\ \mathrm{G} : \mathbf{b} \end{bmatrix} = \sqcup\left(\begin{bmatrix} \mathrm{F} : \boxed{1} \\ \mathrm{G} : \boxed{1} \end{bmatrix} \overset{<}{\sqcap}_c \begin{bmatrix} \mathrm{F} : \mathbf{a} \\ \mathrm{G} : \mathbf{b} \end{bmatrix} \right)$$

$$= \begin{bmatrix} \mathrm{F} : \boxed{1} \\ \mathrm{G} : \boxed{1} \end{bmatrix}$$

Carpenter's definition

Atomic feature structures

- $\pi \equiv \pi'$ indicates reentrancy (coindexation)
  e.g. $\mathrm{F} \equiv \mathrm{G}$ is equivalent to $\begin{bmatrix} \mathrm{F} : \boxed{1} \\ \mathrm{G} : \boxed{1} \end{bmatrix}$
- $\mathcal{P}(\pi) = \sigma$ means that the value on the path π is σ
  e.g. $\mathcal{P}(\mathrm{F}) = \mathbf{b}$ is equivalent to $\begin{bmatrix} \mathrm{F} : \mathbf{b} \end{bmatrix}$

Credulous Default Unification
$F \overset{<}{\sqcap}_c G = \{F \sqcap G' : G' \sqsupseteq G\}$
where G' is maximal such that $F \sqcap G'$ is defined

maximal is defined as being constructed from the greatest number of atomic feature structures

Another example of $\overset{<}{\sqcap}_c$

Where $\mathbf{u} \sqsubset \mathbf{t}$:

$$\begin{bmatrix}\mathbf{t}\\ \mathrm{F}:\mathbf{a}\\ \mathrm{G}:\top\end{bmatrix} \overset{<}{\sqcap} \begin{bmatrix}\mathbf{u}\\ \mathrm{F}:\boxed{1}\,\mathbf{b}\\ \mathrm{G}:\boxed{1}\end{bmatrix} = \left\{\begin{bmatrix}\mathbf{u}\\ \mathrm{F}:\mathbf{a}\\ \mathrm{G}:\mathbf{b}\end{bmatrix}, \begin{bmatrix}\mathbf{u}\\ \mathrm{F}:\boxed{1}\,\mathbf{a}\\ \mathrm{G}:\boxed{1}\end{bmatrix}\right\}$$

The atomic FSs in $\begin{bmatrix}\mathbf{u}\\ \mathrm{F}:\boxed{1}\,\mathbf{b}\\ \mathrm{G}:\boxed{1}\end{bmatrix}$ are:

- $\begin{bmatrix}\mathbf{u}\end{bmatrix}$
- $\begin{bmatrix}\mathrm{F}:\mathbf{b}\end{bmatrix}$
- $\begin{bmatrix}\mathrm{G}:\mathbf{b}\end{bmatrix}$
- $\begin{bmatrix}\mathrm{F}:\boxed{1}\\ \mathrm{G}:\boxed{1}\end{bmatrix}$

Order dependence

- Asymmetric default unification is non-commutative and non-associative
- For order independence we need a symmetric operation, where defaults are marked

$$\left(\begin{bmatrix}\mathrm{F}:\mathbf{c}\\ \mathrm{G}:\mathbf{c}\end{bmatrix} \overset{<}{\sqcap}_c \begin{bmatrix}\mathrm{F}:\mathbf{a}\\ \mathrm{G}:\mathbf{b}\end{bmatrix}\right) \overset{<}{\sqcap}_c \begin{bmatrix}\mathrm{F}:\boxed{1}\\ \mathrm{G}:\boxed{1}\end{bmatrix} = \begin{bmatrix}\mathrm{F}:\boxed{1}\,\mathbf{c}\\ \mathrm{G}:\boxed{1}\end{bmatrix}$$

$$\begin{bmatrix}\mathrm{F}:\mathbf{c}\\ \mathrm{G}:\mathbf{c}\end{bmatrix} \overset{<}{\sqcap}_c \left(\begin{bmatrix}\mathrm{F}:\mathbf{a}\\ \mathrm{G}:\mathbf{b}\end{bmatrix} \overset{<}{\sqcap}_c \begin{bmatrix}\mathrm{F}:\boxed{1}\\ \mathrm{G}:\boxed{1}\end{bmatrix}\right) = \begin{bmatrix}\mathrm{F}:\mathbf{c}\\ \mathrm{G}:\mathbf{c}\end{bmatrix}$$

$$\left(\begin{bmatrix}\mathrm{F}:\mathbf{c}\\ \mathrm{G}:\mathbf{c}\end{bmatrix} \overset{\frown}{\sqcap} \begin{bmatrix}\mathrm{F}:/\mathbf{a}\\ \mathrm{G}:/\mathbf{b}\end{bmatrix}\right) \overset{\frown}{\sqcap} \begin{bmatrix}\mathrm{F}:/\boxed{1}\\ \mathrm{G}:/\boxed{1}\end{bmatrix} = \begin{bmatrix}\mathrm{F}:\mathbf{c}/\boxed{1}\\ \mathrm{G}:\mathbf{c}/\boxed{1}\end{bmatrix}$$

$$\begin{bmatrix}\mathrm{F}:\mathbf{c}\\ \mathrm{G}:\mathbf{c}\end{bmatrix} \overset{\frown}{\sqcap} \left(\begin{bmatrix}\mathrm{F}:/\mathbf{a}\\ \mathrm{G}:/\mathbf{b}\end{bmatrix} \overset{\frown}{\sqcap} \begin{bmatrix}\mathrm{F}:/\boxed{1}\\ \mathrm{G}:/\boxed{1}\end{bmatrix}\right) = \begin{bmatrix}\mathrm{F}:\mathbf{c}/\boxed{1}\\ \mathrm{G}:\mathbf{c}/\boxed{1}\end{bmatrix}$$

Intuition behind YADU

1. Keep track of default information as set of atomic FSs (possibly mutually inconsistent, but consistent with indefeasible information) known as *tails*
2. Priority of information is indicated on elements of tail
3. Calculate default FS by working out which defaults currently 'win'
4. Add in defaults to non-default FS in priority order in a series of credulous unification operations and then generalize

Typed default feature structures: TDFS

TDFS has 3 parts: Indefeasible/Defeasible/Tail

Indefeasible is a normal TFS

Defeasible is a normal TFS, which must be subsumed by Indefeasible

Tail is a set of pairs of atomic TFSs and priority information

$$\begin{bmatrix}\mathbf{t}\\ \mathrm{F}:\mathbf{c}\\ \mathrm{G}:\mathbf{c}\end{bmatrix} / \begin{bmatrix}\mathbf{t}\\ \mathrm{F}:\boxed{1}\,\mathbf{c}\\ \mathrm{G}:\boxed{1}\end{bmatrix} / \{\langle \begin{bmatrix}\mathrm{F}:\boxed{1}\\ \mathrm{G}:\boxed{1}\end{bmatrix}, \mathbf{t}\rangle\}$$

$$\begin{bmatrix}\mathbf{t}\\ \mathrm{F}:\mathbf{c}/\boxed{1}\\ \mathrm{G}:\mathbf{c}/\boxed{1}\end{bmatrix}$$

- For a *basic* TDFS, defeasible structure and tails have equivalent information
- In general, defeasible structure is calculated from tails and indefeasible structure
- Priority depends on (root) type where information was introduced

Simple example of YADU

$$\begin{bmatrix}\mathbf{t}\\ \mathrm{F}:/\mathbf{a}\\ \mathrm{G}:/\mathbf{b}\end{bmatrix} / \begin{matrix}\{\langle\begin{bmatrix}\mathrm{F}:\mathbf{a}\end{bmatrix},\mathbf{t}\rangle,\\ \langle\begin{bmatrix}\mathrm{G}:\mathbf{b}\end{bmatrix},\mathbf{t}\rangle\}\end{matrix} \;\hat{\sqcap}\; \begin{bmatrix}\mathbf{t}\\ \mathrm{F}:\mathbf{c}/\boxed{\mathbf{1}}\\ \mathrm{G}:\mathbf{c}/\boxed{\mathbf{1}}\end{bmatrix} / \{\langle\begin{bmatrix}\mathrm{F}:\boxed{1}\\ \mathrm{G}:\boxed{1}\end{bmatrix},\mathbf{t}\rangle\}$$

Indefeasible part $= \begin{bmatrix}\mathbf{t}\\ \mathrm{F}:\mathbf{c}\\ \mathrm{G}:\mathbf{c}\end{bmatrix}$

New tail (information inconsistent with indefeasible part is removed) :

$$\langle\{\langle\begin{bmatrix}\mathrm{F}:\boxed{1}\\ \mathrm{G}:\boxed{1}\end{bmatrix},\mathbf{t}\rangle\}\rangle$$

Defeasible result

$$= \sqcup\left(\left(\begin{bmatrix}\mathbf{t}\\ \mathrm{F}:\mathbf{c}\\ \mathrm{G}:\mathbf{c}\end{bmatrix} \overset{<}{\sqcap}_{cs} \begin{bmatrix}\mathrm{F}:\boxed{1}\\ \mathrm{G}:\boxed{1}\end{bmatrix}\right)\right) = \begin{bmatrix}\mathbf{t}\\ \mathrm{F}:\boxed{1}\,\mathbf{c}\\ \mathrm{G}:\boxed{1}\end{bmatrix}$$

Final result

$$\begin{bmatrix}\mathbf{t}\\ \mathrm{F}:\mathbf{c}\\ \mathrm{G}:\mathbf{c}\end{bmatrix} / \begin{bmatrix}\mathbf{t}\\ \mathrm{F}:\boxed{1}\,\mathbf{c}\\ \mathrm{G}:\boxed{1}\end{bmatrix} / \{\langle\begin{bmatrix}\mathrm{F}:\boxed{1}\\ \mathrm{G}:\boxed{1}\end{bmatrix},\mathbf{t}\rangle\}$$

Example of YADU: specificity

$$\begin{bmatrix}\mathbf{u}\\ \mathrm{F}:/\mathbf{a}\\ \mathrm{G}:/\mathbf{b}\end{bmatrix} / \begin{matrix}\{\langle\begin{bmatrix}\mathrm{F}:\mathbf{a}\end{bmatrix},\mathbf{u}\rangle,\\ \langle\begin{bmatrix}\mathrm{G}:\mathbf{b}\end{bmatrix},\mathbf{u}\rangle\}\end{matrix} \;\hat{\sqcap}\; \begin{bmatrix}\mathbf{t}\\ \mathrm{F}:/\boxed{1}\\ \mathrm{G}:/\boxed{1}\end{bmatrix} / \{\langle\begin{bmatrix}\mathrm{F}:\boxed{1}\\ \mathrm{G}:\boxed{1}\end{bmatrix},\mathbf{t}\rangle\}$$

Where $\mathbf{u} \sqsubset \mathbf{t}$:

Indefeasible part $= \begin{bmatrix}\mathbf{u}\\ \mathrm{F}:\top\\ \mathrm{G}:\top\end{bmatrix}$

Specificity partition of new tail:

$$\langle\{\langle\begin{bmatrix}\mathrm{F}:\mathbf{a}\end{bmatrix},\mathbf{u}\rangle,\langle\begin{bmatrix}\mathrm{G}:\mathbf{b}\end{bmatrix},\mathbf{u}\rangle\},\quad \{\langle\begin{bmatrix}\mathrm{F}:\boxed{1}\\ \mathrm{G}:\boxed{1}\end{bmatrix},\mathbf{t}\rangle\}\rangle$$

Defeasible result

$$= \sqcup\left(\left(\begin{bmatrix}\mathbf{u}\\ \mathrm{F}:\top\\ \mathrm{G}:\top\end{bmatrix} \overset{<}{\sqcap}_{cs} \left\{\begin{bmatrix}\mathrm{F}:\mathbf{a}\end{bmatrix},\begin{bmatrix}\mathrm{G}:\mathbf{b}\end{bmatrix}\right\}\right) \overset{<}{\sqcap}_{cs} \begin{bmatrix}\mathrm{F}:\boxed{1}\\ \mathrm{G}:\boxed{1}\end{bmatrix}\right)$$

$$= \sqcup\left(\begin{bmatrix}\mathbf{u}\\ \mathrm{F}:\mathbf{a}\\ \mathrm{G}:\mathbf{b}\end{bmatrix} \overset{<}{\sqcap}_{cs} \begin{bmatrix}\mathrm{F}:\boxed{1}\\ \mathrm{G}:\boxed{1}\end{bmatrix}\right) = \begin{bmatrix}\mathbf{u}\\ \mathrm{F}:\mathbf{a}\\ \mathrm{G}:\mathbf{b}\end{bmatrix}$$

Example continued

Therefore:

$$\begin{bmatrix}\mathbf{u}\\ \mathrm{F}:/\mathbf{a}\\ \mathrm{G}:/\mathbf{b}\end{bmatrix} \hat{\sqcap} \begin{bmatrix}\mathbf{t}\\ \mathrm{F}:/\boxed{1}\\ \mathrm{G}:/\boxed{1}\end{bmatrix} =$$

$$\begin{bmatrix}\mathbf{u}\\ \mathrm{F}:/\mathbf{a}\\ \mathrm{G}:/\mathbf{b}\end{bmatrix} / \{\langle\begin{bmatrix}\mathrm{F}:\mathbf{a}\end{bmatrix},\mathbf{u}\rangle,\langle\begin{bmatrix}\mathrm{G}:\mathbf{b}\end{bmatrix},\mathbf{u}\rangle,\langle\begin{bmatrix}\mathrm{F}:\boxed{1}\\ \mathrm{G}:\boxed{1}\end{bmatrix},\mathbf{t}\rangle\}$$

Why we need tails

Suppose no tails, and:

$$\begin{bmatrix}\mathbf{u}\\ \mathrm{F}:/\mathbf{a}\\ \mathrm{G}:/\mathbf{b}\end{bmatrix} \hat{\sqcap} \begin{bmatrix}\mathbf{t}\\ \mathrm{F}:/\boxed{1}\\ \mathrm{G}:/\boxed{1}\end{bmatrix} = \begin{bmatrix}\mathbf{u}\\ \mathrm{F}:/\mathbf{a}\\ \mathrm{G}:/\mathbf{b}\end{bmatrix}$$

Then:

$$(\begin{bmatrix}\mathbf{u}\\ \mathrm{F}:/\mathbf{a}\\ \mathrm{G}:/\mathbf{b}\end{bmatrix} \hat{\sqcap} \begin{bmatrix}\mathbf{t}\\ \mathrm{F}:/\boxed{1}\\ \mathrm{G}:/\boxed{1}\end{bmatrix}) \hat{\sqcap} \begin{bmatrix}\mathbf{t}\\ \mathrm{F}:\mathbf{c}\\ \mathrm{G}:\mathbf{c}\end{bmatrix} = \begin{bmatrix}\mathbf{u}\\ \mathrm{F}:\mathbf{c}\\ \mathrm{G}:\mathbf{c}\end{bmatrix}$$

But

$$\begin{bmatrix}\mathbf{u}\\ \mathrm{F}:/\mathbf{a}\\ \mathrm{G}:/\mathbf{b}\end{bmatrix} \hat{\sqcap} (\begin{bmatrix}\mathbf{t}\\ \mathrm{F}:/\boxed{1}\\ \mathrm{G}:/\boxed{1}\end{bmatrix} \hat{\sqcap} \begin{bmatrix}\mathbf{t}\\ \mathrm{F}:\mathbf{c}\\ \mathrm{G}:\mathbf{c}\end{bmatrix}) = \begin{bmatrix}\mathbf{u}\\ \mathrm{F}:\mathbf{c}/\boxed{\mathbf{1}}\\ \mathrm{G}:\mathbf{c}/\boxed{\mathbf{1}}\end{bmatrix}$$

Definitions: Variants of Credulous Default Unification

Credulous Default Unification $\overset{<}{\sqcap}_{ca}$ on arbitrary sets of atomic FSs
$F_1 \overset{<}{\sqcap}_{ca} \{G_1, \ldots, G_n\} = \{F_1 \sqcap F_2\}$ where F_2 is the unification of a maximal subset of $\{G_1, \ldots, G_n\}$ such that $F_1 \sqcap F_2$ is defined

Credulous Default Unification $\overset{<}{\sqcap}_{cs}$ on Sets
Let $\mathcal{F}_1$ be a *set* of TFSs $\{F_1, \ldots, F_n\}$, and $\mathcal{G}_2$ a set of atomic TFSs. Then

$$\mathcal{F}_1 \overset{<}{\sqcap}_{cs} \mathcal{G}_2 = \{F_1 \overset{<}{\sqcap}_{ca} \mathcal{G}_2, \ldots, F_n \overset{<}{\sqcap}_{ca} \mathcal{G}_2\}$$

Definitions: Tails

The Specificity Partition of a Tail
$\mu_1, \ldots \mu_m$ is a Specificity Partition of T if:

$$T = \mu_1 \cup \ldots \cup \mu_m$$

and $\mu_1 = \{\langle \phi, t \rangle \in T : \forall \langle \phi', t' \rangle \in T, t' \not\sqsubset t\}$
$\mu_i = \{\langle \phi, t \rangle \in T : \exists \langle \phi', t' \rangle \in \mu_{i-1}$ where $t' \sqsubset t$ and
$\forall \langle \phi'', t'' \rangle \in T \setminus (\mu_1 \cup \ldots \cup \mu_{i-1}), t'' \not\sqsubset t\} 2 \leq i \leq m$

Projection of Tails
For tail T:

$$\wp_{fs}(T) = \{\phi : \langle \phi, t \rangle \in T\} \text{ and}$$
$$\wp_{t}(T) = \{t : \langle \phi, t \rangle \in T\}$$

Tail of basic TDFS
If $I/D/T$ is a basic TDFS then
$T = \{\langle \phi, t \rangle :$ where t is the root type on D and ϕ is an atomic FS such that $\phi \sqsupseteq D$ but $\phi \not\sqsupseteq I\}$

Definitions: YADU

If $F_{12} = I_{12}/D_{12}/T^{12} = I_1/D_1/T^1 \overset{<>}{\sqcap} I_2/D_2/T^2$

The Indefeasible Part:
$I_{12} = I_1 \sqcap I_2$

The Tail:
$T^{12} = (T^1 \cup T^2) \setminus Bot^{12}$
where Bot^{12} are all the elements of $T^1 \cup T^2$ where the atomic FS is incompatible with I_{12}. That is:

$$Bot^{12} = \{\langle F, t \rangle : \langle F, t \rangle \in T^1 \cup T^2 \text{ and } F \sqcap I_{12} = \bot\}$$

The Defeasible Part Assume T^{12} is divided into a specificity partition $\langle \mu_1, \ldots, \mu_n \rangle$. Then

$$D_{12} = \sqcup((I_{12} \overset{<}{\sqcap}_{cs} \wp_{fs}(\mu_1)) \overset{<}{\sqcap}_{cs} \ldots \overset{<}{\sqcap}_{cs} \wp_{fs}(\mu_n))$$

That is, D_{12} is the generalisation of the 'credulous default unification' of: I_{12}; the atomic FSs in μ_1; ...; the atomic FSs in μ_m, in that order.

Specificity

$$\begin{bmatrix} \mathbf{u} \\ \mathrm{F} : /\mathbf{a} \\ \mathrm{G} : /\mathbf{b} \end{bmatrix} \overset{<>}{\sqcap} \begin{bmatrix} \mathbf{t} \\ \mathrm{F} : /\boxed{1} \\ \mathrm{G} : /\boxed{1} \end{bmatrix}$$

Where $\mathbf{u} \sqsubset \mathbf{t}$:

$$I_{12} = \begin{bmatrix} \mathbf{u} \\ \mathrm{F} : \top \\ \mathrm{G} : \top \end{bmatrix} \sqcap \begin{bmatrix} \mathbf{t} \\ \mathrm{F} : \top \\ \mathrm{G} : \top \end{bmatrix} = \begin{bmatrix} \mathbf{u} \\ \mathrm{F} : \top \\ \mathrm{G} : \top \end{bmatrix}$$

$$\begin{aligned}
T^{12} &= (T^1 \cup T^2) \setminus Bot^{12} \\
T^1 &= \{\langle \begin{bmatrix} \mathrm{F} : \mathbf{a} \end{bmatrix}, \mathbf{u} \rangle, \langle \begin{bmatrix} \mathrm{G} : \mathbf{b} \end{bmatrix}, \mathbf{u} \rangle\} \\
T^2 &= \{\langle \begin{bmatrix} \mathrm{F} : \boxed{1} \\ \mathrm{G} : \boxed{1} \end{bmatrix}, \mathbf{t} \rangle\} \\
Bot^{12} &= \emptyset \\
T^{12} &= \{\langle \begin{bmatrix} \mathrm{F} : \mathbf{a} \end{bmatrix}, \mathbf{u} \rangle, \langle \begin{bmatrix} \mathrm{G} : \mathbf{b} \end{bmatrix}, \mathbf{u} \rangle, \langle \begin{bmatrix} \mathrm{F} : \boxed{1} \\ \mathrm{G} : \boxed{1} \end{bmatrix}, \mathbf{t} \rangle\}
\end{aligned}$$

Specificity partition of T^{12}:

$$\langle \{\langle \begin{bmatrix} \mathrm{F} : \mathbf{a} \end{bmatrix}, \mathbf{u} \rangle, \langle \begin{bmatrix} \mathrm{G} : \mathbf{b} \end{bmatrix}, \mathbf{u} \rangle\}, \quad \{\langle \begin{bmatrix} \mathrm{F} : \boxed{1} \\ \mathrm{G} : \boxed{1} \end{bmatrix}, \mathbf{t} \rangle\} \rangle$$

$$D_{12} = \sqcup((I_{12} \overset{<}{\sqcap}_{cs} \wp_{fs}(\mu_1)) \overset{<}{\sqcap}_{cs} \wp_{fs}(\mu_2))$$

$$= \sqcup\left(\left(\begin{bmatrix}\mathbf{u}\\ \mathrm{F}:\top \\ \mathrm{G}:\top\end{bmatrix} \overset{<}{\sqcap}_{cs} \left\{ \begin{bmatrix}\mathrm{F}:\mathbf{a}\end{bmatrix}, \begin{bmatrix}\mathrm{G}:\mathbf{b}\end{bmatrix} \right\}\right) \overset{<}{\sqcap}_{cs} \begin{bmatrix} \mathrm{F}: \boxed{1} \\ \mathrm{G}: \boxed{1}\end{bmatrix}\right)$$

$$= \sqcup\left(\begin{bmatrix}\mathbf{u}\\ \mathrm{F}:\mathbf{a} \\ \mathrm{G}:\mathbf{b}\end{bmatrix} \overset{<}{\sqcap}_{cs} \begin{bmatrix} \mathrm{F}: \boxed{1} \\ \mathrm{G}: \boxed{1}\end{bmatrix}\right)$$

$$= \begin{bmatrix}\mathbf{u}\\ \mathrm{F}:\mathbf{a} \\ \mathrm{G}:\mathbf{b}\end{bmatrix}$$

Hence:

$$\begin{bmatrix}\mathbf{u}\\ \mathrm{F}:/\mathbf{a} \\ \mathrm{G}:/\mathbf{b}\end{bmatrix} \overset{<>}{\sqcap} \begin{bmatrix}\mathbf{t}\\ \mathrm{F}:/\boxed{1} \\ \mathrm{G}:/\boxed{1}\end{bmatrix} =$$

$$\begin{bmatrix}\mathbf{u}\\ \mathrm{F}:/\mathbf{a} \\ \mathrm{G}:/\mathbf{b}\end{bmatrix} / \{\langle\begin{bmatrix}\mathrm{F}:\mathbf{a}\end{bmatrix}, \mathbf{u}\rangle, \langle\begin{bmatrix}\mathrm{G}:\mathbf{b}\end{bmatrix}, \mathbf{u}\rangle, \langle\begin{bmatrix}\mathrm{F}:\boxed{1}\\ \mathrm{G}:\boxed{1}\end{bmatrix}, \mathbf{t}\rangle\}$$

Conflicts without priority order

$$\begin{bmatrix}\mathbf{p}\\ \mathrm{F}:/\mathbf{a} \\ \mathrm{G}:/\mathbf{b}\end{bmatrix} \overset{<>}{\sqcap} \begin{bmatrix}\mathbf{q}\\ \mathrm{F}:/\boxed{1} \\ \mathrm{G}:/\boxed{1}\end{bmatrix}$$

Where $\mathbf{p}$ and $\mathbf{q}$ are in no specificity ordering, and $\mathbf{p} \sqcap \mathbf{q} = \mathbf{r}$, $\mathbf{a} \sqcap \mathbf{b} = \bot$ and $\mathbf{a} \sqcup \mathbf{b} = \mathbf{c}$

$$I_{12} = \begin{bmatrix}\mathbf{p}\\ \mathrm{F}:\top \\ \mathrm{G}:\top\end{bmatrix} \sqcap \begin{bmatrix}\mathbf{q}\\ \mathrm{F}:\top \\ \mathrm{G}:\top\end{bmatrix} = \begin{bmatrix}\mathbf{r}\\ \mathrm{F}:\top \\ \mathrm{G}:\top\end{bmatrix}$$

$$T^{12} = (T^1 \cup T^2) \setminus Bot^{12}$$

$$T^1 = \{\langle\begin{bmatrix}\mathrm{F}:\mathbf{a}\end{bmatrix}, \mathbf{p}\rangle, \langle\begin{bmatrix}\mathrm{G}:\mathbf{b}\end{bmatrix}, \mathbf{p}\rangle\}$$

$$T^2 = \{\langle\begin{bmatrix}\mathrm{F}:\boxed{1}\\ \mathrm{G}:\boxed{1}\end{bmatrix}, \mathbf{q}\rangle\}$$

$$Bot^{12} = \emptyset$$

So, specificity partition of T^{12} =

$$\langle\{\langle\begin{bmatrix}\mathrm{F}:\mathbf{a}\end{bmatrix}, \mathbf{p}\rangle, \langle\begin{bmatrix}\mathrm{G}:\mathbf{b}\end{bmatrix}, \mathbf{p}\rangle, \langle\begin{bmatrix}\mathrm{F}:\boxed{1}\\ \mathrm{G}:\boxed{1}\end{bmatrix}, \mathbf{q}\rangle\}\rangle$$

$$D_{12} = \sqcup((I_{12} \overset{<}{\sqcap}_{cs} \wp_{fs}(T^{12})))$$

$$= \sqcup\left(\left(\begin{bmatrix}\mathbf{r}\\ \mathrm{F}:\top \\ \mathrm{G}:\top\end{bmatrix} \overset{<}{\sqcap}_{cs} \left\{ \begin{bmatrix}\mathrm{F}:\mathbf{a}\end{bmatrix}, \begin{bmatrix}\mathrm{G}:\mathbf{b}\end{bmatrix}, \begin{bmatrix}\mathrm{F}:\boxed{1}\\ \mathrm{G}:\boxed{1}\end{bmatrix} \right\}\right)\right)$$

$$= \sqcup\left(\left\{\begin{bmatrix}\mathbf{r}\\ \mathrm{F}:\mathbf{a} \\ \mathrm{G}:\mathbf{b}\end{bmatrix}, \begin{bmatrix}\mathbf{r}\\ \mathrm{F}:\boxed{\mathbf{1}}\mathbf{a} \\ \mathrm{G}:\boxed{\mathbf{1}}\end{bmatrix}, \begin{bmatrix}\mathbf{r}\\ \mathrm{F}:\boxed{\mathbf{1}}\mathbf{b} \\ \mathrm{G}:\boxed{\mathbf{1}}\end{bmatrix}\right\}\right)$$

$$= \begin{bmatrix}\mathbf{r}\\ \mathrm{F}:\mathbf{c} \\ \mathrm{G}:\mathbf{c}\end{bmatrix} \text{ because } \mathbf{a} \sqcup \mathbf{b} = \mathbf{c}$$

Hence

$$\begin{bmatrix}\mathbf{p}\\ \mathrm{F}:/\mathbf{a} \\ \mathrm{G}:/\mathbf{b}\end{bmatrix} \overset{<>}{\sqcap} \begin{bmatrix}\mathbf{q}\\ \mathrm{F}:/\boxed{1} \\ \mathrm{G}:/\boxed{1}\end{bmatrix} =$$

$$\begin{bmatrix}\mathbf{r}\\ \mathrm{F}:/\mathbf{c} \\ \mathrm{G}:/\mathbf{c}\end{bmatrix} / \{\langle\begin{bmatrix}\mathrm{F}:\mathbf{a}\end{bmatrix}, \mathbf{p}\rangle, \langle\begin{bmatrix}\mathrm{G}:\mathbf{b}\end{bmatrix}, \mathbf{p}\rangle, \langle\begin{bmatrix}\mathrm{F}:\boxed{1}\\ \mathrm{G}:\boxed{1}\end{bmatrix}, \mathbf{q}\rangle\}$$

Well-Typed Unification

$$\begin{bmatrix}\mathbf{u}\\ \mathrm{F}:/\mathbf{a}\end{bmatrix} \overset{<>}{\sqcap} \begin{bmatrix}\mathbf{t}\\ \mathrm{F}:/\begin{bmatrix}\mathbf{v}\\ \mathrm{G}:\mathbf{b}\end{bmatrix}\end{bmatrix}$$

Where $\mathbf{u} \sqsubset \mathbf{t}$, and $\mathbf{a}$ is atomic

$$I_{12} = I_1 \sqcap I_2 = \begin{bmatrix}\mathbf{u}\\ \mathrm{F}:\top\end{bmatrix}$$

$$T^{12} = (T^1 \cup T^2) \setminus Bot^{12}$$

$$= \{\langle\begin{bmatrix}\mathrm{F}:\mathbf{a}\end{bmatrix}, \mathbf{u}\rangle, \langle\begin{bmatrix}\mathrm{F}:\mathbf{v}\end{bmatrix}, \mathbf{t}\rangle, \langle\begin{bmatrix}\mathrm{F}:\begin{bmatrix}\mathrm{G}:\mathbf{b}\end{bmatrix}\end{bmatrix}, \mathbf{t}\rangle\}$$

Therefore the specificity partition of T^{12} is $\langle\mu_1, \mu_2\rangle$, where:

$$\mu_1 = \{\langle\begin{bmatrix}\mathrm{F}:\mathbf{a}\end{bmatrix}, \mathbf{u}\rangle\}$$

$$\mu_2 = \{\langle\begin{bmatrix}\mathrm{F}:\mathbf{v}\end{bmatrix}, \mathbf{t}\rangle, \langle\begin{bmatrix}\mathrm{F}:\begin{bmatrix}\mathrm{G}:\mathbf{b}\end{bmatrix}\end{bmatrix}, \mathbf{t}\rangle\}$$

$$D_{12} = \sqcup(((I_{12} \overset{<}{\sqcap}_{cs} \wp_{fs}(\mu_1)) \overset{<}{\sqcap}_{cs} \wp_{fs}(\mu_2))$$

$$= \sqcup\left(\left(\left(\begin{bmatrix}\mathbf{t}\\ \mathrm{F}:\top\end{bmatrix} \overset{<}{\sqcap}_{cs} \left\{\begin{bmatrix}\mathrm{F}:\mathbf{a}\end{bmatrix}\right\}\right) \overset{<}{\sqcap}_{cs} \left\{\begin{bmatrix}\mathrm{F}:\mathbf{v}\end{bmatrix}, \begin{bmatrix}\mathrm{F}:\begin{bmatrix}\mathrm{G}:\mathbf{b}\end{bmatrix}\end{bmatrix}\right\}\right)\right)$$

$$= \sqcup\left(\begin{bmatrix}\mathbf{t}\\ \mathrm{F}:\mathbf{a}\end{bmatrix} \overset{<}{\sqcap}_{cs} \left\{\begin{bmatrix}\mathrm{F}:\mathbf{v}\end{bmatrix}, \begin{bmatrix}\mathrm{F}:\begin{bmatrix}\mathrm{G}:\mathbf{b}\end{bmatrix}\end{bmatrix}\right\}\right)$$

$$= \begin{bmatrix}\mathbf{t}\\ \mathrm{F}:\mathbf{a}\end{bmatrix}$$

Hence

$$\begin{bmatrix}\mathbf{u}\\ \mathrm{F}:/\mathbf{a}\end{bmatrix} \overset{<>}{\sqcap} \begin{bmatrix}\mathbf{t}\\ \mathrm{F}:/\begin{bmatrix}\mathbf{v}\\ \mathrm{G}:\mathbf{b}\end{bmatrix}\end{bmatrix} =$$

$$\begin{bmatrix}\mathbf{u}\\ \mathrm{F}:/\mathbf{a}\end{bmatrix} / \{\langle\begin{bmatrix}\mathrm{F}:\mathbf{a}\end{bmatrix}, \mathbf{u}\rangle, \langle\begin{bmatrix}\mathrm{F}:\mathbf{v}\end{bmatrix}, \mathbf{t}\rangle, \langle\begin{bmatrix}\mathrm{F}:\begin{bmatrix}\mathrm{G}:\mathbf{b}\end{bmatrix}\end{bmatrix}, \mathbf{t}\rangle\}$$

The Properties of $\overset{<>}{\sqcap}$

Lemma 1 *Tails are commutative.*

Lemma 2 *Tails are associative. That is:* $T^{(12)3} = T^{1(23)}$

Lemma 3 $\overset{<>}{\sqcap}$ *is a function. That is* $\mathrm{TDFS}_1 \overset{<>}{\sqcap} \mathrm{TDFS}_2$ *is a unique* TDFS.

Lemma 4 $\overset{<>}{\sqcap}$ *is commutative.*

Lemma 5 $\overset{<>}{\sqcap}$ *is associative.*

Theorem 1 $\overset{<>}{\sqcap}$ *is an order independent binary function.*

Theorem 2 The Typed Unification Property *Let* F_1, F_2 *be* TDFS*s. And suppose* $I_{12} \sqcap \wp_{fs}(T^1) \sqcap \wp_{fs}(T^2) \neq \bot$. *Then* $D_{12} = D_1 \sqcap D_2$.

Corollary 1 The Typed Unification Property for Basic TDFSs
If TDFS_1 *and* TDFS_2 *are basic* TDFS*s, and* $D_1 \sqcap D_2 \neq \bot$, *then*

$$D_{12} = D_1 \sqcap D_2$$

Converting TDFS to TFS

For non-persistent (e.g., purely lexical) defaults we convert a TDFS to a TFS simply taking the defeasible structure (since this is guaranteed to be more specific than the indefeasible structure):

$DefFill(I, D, T) = D$

For example:

$$DefFill(\begin{bmatrix}\textbf{pst-t-verb}\\ \mathrm{PST}:\top\\ \mathrm{PSP}:\top\end{bmatrix} / \begin{bmatrix}\textbf{pst-t-verb}\\ \mathrm{PST}:\boxed{1}+\mathbf{t}\\ \mathrm{PSP}:\boxed{1}\end{bmatrix})$$

$$= \begin{bmatrix}\textbf{pst-t-verb}\\ \mathrm{PST}:\boxed{1}+\mathbf{t}\\ \mathrm{PSP}:\boxed{1}\end{bmatrix}$$

BUT, conversion implies a hard interface at the point where defeasible information is made indefeasible.

Computational properties

Credulous default unification is worst-case factorial in number of atomic FSs. YADU involves a series of these operations.
BUT

- Partioning defaults helps
- Defeasible feature structure need only be calculated at an interface (display, conversion to indefeasible structure, pragmatics)
- Calculation of tails is set union plus low-cost check for compatibility
- For many real applications the overhead on ordinary unification is small because most of information is non-default

Constructional polysemy vs sense extension (1)

Constructional polysemy:

- A single sense assigned to a lexical entry is contextually specialised
- Implemented by underspecification
- Examples: *brush* for paint, hair, cleaning
  fast in *fast typist* vs *fast car*

Sense extension:

- 'Lexical rules' used to represent conventional aspects
- So new typed feature structures (TFSs) are generated
- Examples: *lacquer*:noun vs *lacquer*:verb
  lamb:animal vs *lamb*:meat

Hypothesis: difference in formalization corresponds to valid linguistic distinction

Constructional polysemy vs sense extension (2)

Compare:
sense modulation vs sense change (Cruse, 1986)
conceptual specification vs conceptual shift (Bierwisch, 1982)
vagueness vs ambiguity

If a single word can carry multiple meanings simultaneously, then a single FS must be able to cover all those meanings, and we must assume constructional polysemy.

(28) a Plastic handles are used to make all types of brush cheaply
b This thesis is 1000 pages long and totally unreadable

Sense extension is appropriate if there are (major) syntactic effects, analogous morphologically marked processes, clear directionality or semi-productivity.

Constructional polysemy vs sense extension (3)

Issue is not the same as whether underspecification can be used instead of lexical rules, because specialisation of a feature structure has global effect.

Co-predication is a better test than modification because of the possibility of unary sense extension rules applying to phrases

(29) a Newspapers which are owned by Murdoch are only good for lining cat boxes
b The south side of Cambridge voted Conservative

Co-predication and zeugma

(30) This thesis is 1000 pages long and totally unreadable

We have to avoid situations which are (more-or-less) impossible: the oddity of (31) does not mean that *teacher* is ambiguous between male and female

(31) ? The teacher was pregnant and had a beard

Also avoid sentences which would not make a coherent discourse

(32) ? This thesis is orange and totally unreadable

(33) ? Rembrandt and our janitor were both geniuses with a brush

Test is that absence of zeugma with copredication indicates constructional polysemy. Presence of zeugma does not indicate dual structures as it may be explained as a discourse effect (lecture 5).

Copredication continued

(34) a We need more fast typists and programmers

b Kim ate and enjoyed the cake

c ?? The ham sandwich wants a coke and has gone stale

d ?? The french fries is getting impatient and are getting cold

(35) That newspaper has been attacked by the opposition and publicly burned by demonstrators. (Nunberg, pc)

(36) * That newspaper is owned by a trust and is covered with coffee

Sense clines: Cruse (1986)

Conjunction of neighbouring sentences is OK in all cases, but ends of cline cannot be conjoined.

That newspaper is owned by a trust.
That newspaper is left of centre.
That newspaper supported the Labour Party at the last election.
That newspaper carries long articles about the internal struggles of the Labour Party.
That newspaper has obscure editorials.
That newspaper is full of metaphorical language.
That newspaper is full of long sentences.
That newspaper is full of spelling mistakes.
That newspaper is full of typographic errors.
That newspaper has an unreadable font.
That newspaper has lots of smudged type.
That newspaper is covered with coffee.

Quantificational tests

brush can refer to collection of different types of object: (37) could be true of two paint brushes and one hair brush

(37) The salesman sold me three brushes

book can refer either to copies or concept, but can't mix the two:

(38) Jeffrey has sold five books

thesis seems to primarily refer to conceptual contents rather than copies: (39) can't be true of three copies of the same thesis

(39) There are three theses on the table

(40) has to refer to three newspapers as institution and some arbitrary number of copies of each of their publications

(40) Three newspapers have been attacked by the opposition and publicly burned by demonstrators.

Sense extension combined with constructional polysemy

(41) a This chicken is corn-fed and healthy

b Corn-fed and inexpensive chicken is difficult to find

Assume that extended sense has an ORIGIN feature in its qualia structure which is available for modification. Also true of analogous underived words.

(42) a This calf is corn-fed and healthy

b Corn-fed and inexpensive veal is difficult to find

Constraints are imposed by discourse coherence.

(43) a ? We serve corn-fed and happy chicken

b ? We serve dense potatoes

Using agreement tests to see if reference is transferred

(44) a The hash browns is/*are getting impatient.

b The event of eating was enjoyed by John.

c Those hash browns *was/were enjoyed by John.

Summary

- TFSs for integrated syntax and semantics
- YADU: an order-independent unification operation on TDFS
- Constructional polysemy and sense extension involve different representation mechanisms and can be distinguished (to some extent) by ambiguity tests
- Explaining the results of ambiguity tests may require reasoning about discourse

Next lecture

- Sense extension and lexical rules
- DICE

Lecture 3: Part 1

- Sense extension
- Lexical rules
- Semiproductivity, blocking and unblocking

Sense extension

- Grammatical effects: count -> mass, verbing
- Comparison with morphology
- Established and non-established senses
- Language specificity

(45) a Sandy drank a bottle of whisky.
container -> contents
b Sandy drank a bottleful of whisky.

(46) a Kim ate some chicken.
animal -> meat
b That restaurant serves ostrich.

(47) a The ham sandwich has paid his check.
physical object -> associated person
b * The dark haired guy is in the microwave.

Sense extension (continued)

[Chester] serves not just country folk, but farming, suburban and city folk too. You'll see Armani drifting into the Grosvenor Hotel's exclusive (but exquisite) Arkle Restaurant and C+A giggling out of its streetfront brasserie next door. (Guardian Weekly)
manufacturer -> product + clothes -> wearer

Badger hams are a delicacy in China while mole is eaten in many parts of Africa (LOB corpus)

- Cross-linguistic effects: olive (fruit or tree) aceituna/aceituno
- Language specificity: Dutch compounds (e.g. *lamsvlees*), Eskimo has no conventionalised animal grinding (Nunberg and Zaenen, 1992)

'Lexical rules'

- Assuming account of lexical rules given in Copestake and Briscoe (1995)
- 'Lexical rules' are implemented as typed feature structures and can therefore form an inheritance hierarchy.
- Formally they are identical to unary compositional grammar rules and can therefore apply to phrases
- Sense extension rules are conventionalized, non-compositional (they can add meaning), and semi-productive

$$\begin{bmatrix} \textbf{lexrule} \\ 0 : \textbf{sign} \\ 1 : \textbf{sign} \end{bmatrix}$$

We write this as:

$$\textbf{lexrule} \\ \begin{bmatrix} \textbf{sign} \end{bmatrix} \rightarrow \begin{bmatrix} \textbf{sign} \end{bmatrix}$$

grinding

$$\begin{bmatrix} \textbf{common-noun} \\ \text{ORTH} : \boxed{0}\ \ \textbf{orth} \\ \text{SYN} : \textbf{nounsyn} \\ \text{SEM} : \begin{bmatrix} \text{INDEX} : \boxed{x} \\ \text{LISZT} : \boxed{1}\ \left\langle \begin{bmatrix} \text{PRED} : \boxed{P}\ \ \textbf{rel} \\ \text{COUNT} : \textbf{true} \\ \text{INST} : \boxed{x} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix} \rightarrow$$

$$\begin{bmatrix} \textbf{common-noun} \\ \text{ORTH} : \boxed{0} \\ \text{SYN} : \textbf{nounsyn} \\ \text{SEM} : \begin{bmatrix} \text{INDEX} : \boxed{y}\ \ \text{QU} : \begin{bmatrix} \text{ORIGIN} : \boxed{P} \end{bmatrix} \\ \text{LISZT} : \left\langle \begin{bmatrix} \text{PRED} : \textbf{stuff\_rel} \\ \text{COUNT} : \textbf{false} \\ \text{INST} : \boxed{y} \end{bmatrix}, \begin{bmatrix} \text{PRED} : \textbf{grind\_rel} \\ \text{ARG1} : \boxed{x} \\ \text{ARG2} : \boxed{y} \end{bmatrix} \right\rangle \oplus \boxed{1} \end{bmatrix} \end{bmatrix}$$

Hierarchies of lexical rules

Grinding: there was rabbit all over the road
Meat Grinding: Sandy ate rabbit
Fur/Skin Grinding: Sandy wears rabbit

Meat-grinding applied to *rabbit*:

$$\left[\begin{array}{l} \textbf{common-noun} \\ \text{ORTH} : \textbf{rabbit} \\ \text{SYN} : \textbf{nounsyn} \\ \text{SEM} : \left[\begin{array}{l} \text{INDEX} : \boxed{\text{y}} \ \ \text{QU} : \left[\begin{array}{l} \text{ORIG} : \textbf{\_rabbit\_rel} \\ \text{TELIC} : \textbf{/\_eat\_rel} \end{array}\right] \\ \text{LISZT} : \left\langle \left[\begin{array}{l} \text{PRD} : \textbf{meat\_rel} \\ \text{COUNT} : \textbf{false} \\ \text{INST} : \boxed{\text{y}} \end{array}\right], \right. \\ \left. \left[\begin{array}{l} \text{PRD} : \textbf{grind\_rel} \\ \text{ARG1} : \boxed{\text{x}} \\ \text{ARG2} : \boxed{\text{y}} \end{array}\right], \left[\begin{array}{l} \text{PRD} : \textbf{\_rabbit\_rel} \\ \text{COUNT} : \textbf{true} \\ \text{INST} : \boxed{\text{x}} \end{array}\right] \right\rangle \end{array}\right] \end{array}\right]$$

Sense extension and pragmatics

- We assume that lexical rules are also needed for 'ham sandwich' examples, because of grammatical effects and conventionalized aspects.
- Broad lexical rules define some limits to possible sense change and are analogous to Nunberg's (1978) transfer functions.
- Nunberg's (1978) cue-validity constraints are also necessary: can only use 'ham sandwich' when this uniquely identifies a customer etc. We assume this is part of pragmatics.

Sense extension and pragmatics (cont.)

- Narrow lexical rules are analogous to Nunberg and Zaenen's (1992) lexical licenses.
- Narrow lexical rules give an account of the usual reading. On our account, mass use of *rabbit* is ambiguous, but see approach to semi-productivity.
- Note: *calf* used for skin even though *veal* blocks its use for meat
- Meaning still isn't fully defined: **meat_rel** (as opposed to _meat_rel) is to be taken as covering animal flesh in general (e.g. don't distinguish *whitebait* and *rabbit*).

Phrasal 'lexical' rules

The south side of Cambridge voted Conservative

Three villages voted for the proposed ban on timber production
Three villages south of the river voted for the proposed ban on timber production
?Three villages built of stone voted for the proposed ban on timber production

Problems with assuming productivity

1. Narrow classes (Pinker, Levin) often have to be very small and are sometimes unintuitive
2. Exceptions due to blocking are (partially) formalizable but require detailed lexical semantic information
3. Rarity of some uses
4. Degrees of acceptability
5. Sense proliferation

Semi-productivity

- Semi-productivity is an issue of performance, not competence (Bauer 1983).
- Avoid obscurity: use the form which has highest frequency.
- Application of lexical rules is sensitive to both type and token frequency effects
- Frequency with which a given word form is associated with a particular lexical entry (i.e. sense or grammatical realization) is often highly skewed
- Frequency information could be utilized to resolve indeterminacies in both generation and interpretation

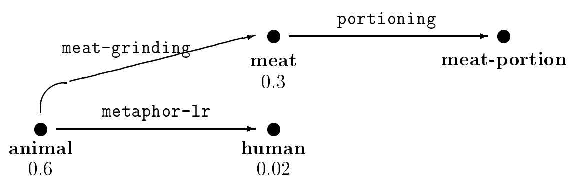
Representation

- Use covariation technique and construct FSM representing possible lexical rule applications (Meurers and Minnen, 1995)
- Entries include base and attested derived forms, each associated with a probability
- Lexical probabilities are acquired independently of the lexical rules used to create derived entries
- Topology of the FSM may be shared with other words, but probabilities are idiosyncratic

Lexeme for *rabbit*

$$\left[\begin{array}{l}\textbf{common-noun}\\ \text{ORTH} : \textbf{rabbit}\\ \text{SYN} : \textbf{nounsyn}\\ \text{SEM} : \left[\begin{array}{l}\text{INDEX} : \boxed{x}\\ \text{LISZT} : \left\langle\left[\begin{array}{l}\text{PRED} : \textbf{\_rabbit\_rel}\\ \text{COUNT} : \textbf{true}\\ \text{INST} : \boxed{x}\end{array}\right]\right\rangle\end{array}\right]\end{array}\right]$$

count-lexeme-fsm(animal(0.6), metaphor-lr(0.02), meat-grinding(0.3) ...)



Blocking

Blocking is an automatic consequence of the maxim to use highest frequency form: therefore generation of blocked forms is marked.

(48) That restaurant often serves venison/?deer.

Effect of the current proposal:

- Blocked forms will not be generated
- Blocked forms can be interpreted, if other possibilities fail
- Need formal account of pragmatic effects of unblocking

Unblocking

'Hot sausages, two for a dollar, made of genuine pig, why not buy one for the lady?'
'Don't you mean pork, sir?' said Carrot warily, eyeing the glistening tubes.
'Manner of speaking, manner of speaking,' said Throat quickly.
'Certainly your actual pig products. Genuine pig.' (Terry Pratchett, 1989. *Guards, Guards!*)

There were five thousand extremely loud people on the floor eager to tear into roast cow with both hands and wash it down with bourbon whiskey.
(Tom Wolfe, 1979. *The Right Stuff*)

In the case of at least one county primary school ... they were offered (with perfect timing) saute potatoes, carrots, runner beans and roast cow (Guardian, May 16 1990, article about mad cow disease)

Some aspects of lexical semantic representation NOT covered here

- Details of syntax / semantic interface
- Idioms
- Formalization of lexical rules
- Collocation

Conventionality and collocation

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

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.

Summary

- Sense extension is implemented using hierarchies of lexical rules
- Semi-productivity accounted for using frequencies
- Blocking may be treated using general mechanism for semi-productivity
- Unblocking has pragmatic effects