The Slot Grammar Lexical Formalism - McCord 2006

Sarah Schneider

Universität Tübingen
Seminar für Sprachwissenschaft

12th May 2016
Content

1. Overview
2. Overall Format
3. Index words
4. Sense frames
5. Parts of Speech
6. Slots
7. Inventory of slots
8. Inventory of slot options
9. Option tests
10. Features
11. Subject area tests
12. Sense names
13. Inflectional elements
14. Support word frames
1. Overview

- The Slot Grammar Lexical Formalism provides a rich descriptive language for lexical entries.
- Slots have two roles:
  - Indicate logical arguments
  - Act as grammatical relations
- Example Slots: subj, obj
  - Numerical parse scoring to arrive at most likely parse
- Slots help differentiating senses
- SLF is applicable to any language
- SLF can be used for
  - Building base lexicons for SG parsers
  - For building user lexicons for SG applications
- Parsing consists of filling the Slots
1. Overview

<table>
<thead>
<tr>
<th></th>
<th>Slots</th>
<th>Features</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Julia</td>
<td>subj(n)</td>
<td>noun pron</td>
<td></td>
</tr>
<tr>
<td>sent</td>
<td>top</td>
<td>verb</td>
<td></td>
</tr>
<tr>
<td>a</td>
<td>ndet</td>
<td>det indef</td>
<td></td>
</tr>
<tr>
<td>letter</td>
<td>obj(n)</td>
<td>noun</td>
<td></td>
</tr>
</tbody>
</table>
2. Overall Format

- A lexical entry consists of an index word followed by one or more elements
- access
  - < v obj
  - < n (p to)
3. Index Words

- Can be single or multiwords
- The element that gets inflected is the head and is indicated by the equals sign
  data = base < n
- If no word gets inflected, both can be head, but the rarer one should be chosen as it makes the look-up easier
- A multiword can contain the element num which stands for an arbitrary number
  = less than num < adj
4. Sense Frames

- Specifies the sense of the index word
- Members of the sense frame are
  - POS, Slot, Feature, Subject Area Test, Sense Score, Sense Name
- Best order of Sense Frame items:
  - Index Word < POS Slots Features Subject Area Test Score Sense Name
- = Attorney General < propn (p of) h (sa gov)
5. Parts of Speech

- n, propn, pron, num, v, modal, adj, adv, qual, det, prep, subconj, conj, infto, subinf, forto, thatconj
- Some are subcategories of normal POSs:
  - propn is a kind of noun
- Some are very special with only one or two words
6. Slots

- give \(<\ v\ obj\ iobj\)

- *Alice gave the book to Bob.*

- *Alice gave Bob the book.*

- obj and iobj are filled in different positions
  - The category of the filler may vary, but it fills the same slot
  - The logical role of the filler is the same - the book was given to Bob in both cases

- Slots correspond to the logical arguments of its senses

- Slots have a syntactic and a semantic nature

- A slot for a given word sense can be filled only once
6. Slots

- Slots can be obligatory or optional
  - optionality is default
- Obligatory slots are indicated by a 1
  - (obj1 n) (iobj n to)
- Complement vs. adjunct slots
  - specified in lexical entry vs. not specified
  - *Probably Alice gave Bob the book yesterday*
  - probably and yesterday are adjunct slots of the form vadv
7. Inventory of Slots

- Set of slots for a word sense is dependent on its POS
- Slots for verbs
  - subj, obj, iobj, pred, auxcomp, comp
  - subj needs not to be specified
  - comp
    - *Alice drove Betty to the store* vs. *Alice drove Betty crazy*
    - drive \(<\) v obj (comp lo a)
  - pred
    - in ESG only used by the word *be*
    - be \(<\) v (subj n v) pred
    - *Bob is a teacher*
  - auxcomp
    - mainly for modal verbs
    - *Bob may have been taken to the station*
7. Inventory of Slots

- **Slots for nouns**
  - The two noun slots are obj and nid
  - obj: can have different options, if no option is given explicitly, the default (n) is taken and filled by of-PPs
    - brother < n obj, construction < n obj
  - A noun sense frame may have several occurrences with different options
    - assignment < n obj(obj (p to) inf)
  - *The assignment of Bob to write the report*
  - nid: noun identifier: filled by proper nouns and numbers

- **Slots for adjectives**
  - aboj / obj in lexicon: several possible options
7. Inventory of Slots

- **Slots for adverbs**
  - avobj / obj in lexicon: several possible options

- **Slots for determiners and qualifiers**
  - no complement slots

- **Slots for prepositions**
  - objprep → is obligatory, but no 1 is needed
  - *in the house, in writing this program*

- **Slots for subordinate conjunctions (subconj)**
  - sscomp → obligatory
  - *if you write this program*

- **Slots for coordinate conjunctions (conj)**
  - lconj and rconj
  - *The cat sat on the mat and the dog lay on the floor*
7. Inventory of Slots

- **Slots for infinite to (infto)**
  - tocomp → complement slot
  - *to eat the chocolate*
  - to < ... < infto (tocomp binf)

- **Slots for subinf**
  - subinfcomp
  - in =order to < subinf (subinfcomp binf)

- **Slots for forto**
  - forsubj, forcomp
  - for < ... < forto (forsubj n) (forcomp inf)
  - *this would allow for you sister to be there earlier*

- **Slots for thatconj**
  - thatcomp
  - that < ... < thatconj (thatcomp bfin)
  - *he said that the chocolate was delicious*
8. Inventory of Slot Options

- The slots options are:
  - a, agent, aj, av, bfin, binf, dt, en, ena, fin, fina, finq, finv, ft, ger, gn, inf, ing, io, it, itinf, itthatc, itwh, lo, n, na, nen, nmeas, nop, nummeas, p, padj, pinf, pinfd, prflx, prop, pt, pthatc, pwh, qt, rflx, sc, so, thatc, v, wh

- Each option corresponds to a possible filler of the slot:
  - n → noun phrases
  - a → adjective phrases
  - p → prepositional phrases
  - ing → present participle

- Options might be shared across more slots, with different meanings

- most options apply to the verb slots obj and comp and to the obj slots nobj, aobj, avobj
8. Inventory of Slot Options - Example: n

- For verbs, n can be used with the slots subj, obj, iobj and comp → filled by noun phrases

- subj and obj → filled by two NP-like adjective phrases
  - *The very rich* → adjnoun
  - *The happiest* → superlative head adjective

- subj, obj and comp → filled by wh-verb phrases
  - *what he tried to find* | *whatever they want* → whnom

- *They elected her president of the association*

- Pred implicitly has an n option

- For noun POS, n can be the obj and is filled by of-PPs
  - *president of the association*

- For the adj POS, n can be filled by a noun phrase
  - *It is due me.*
9. Option Tests

- Slots options can have tests (OptionName $Test_1, Test_2, \ldots, Test_n$)
- At least one test must succeed
- Option test operators: $p$, $st$, $fe$, $f$, $of$, $nf$, $hd$, $sn$, $ph$, $wds$, $sa$, $ev$, $\&$, $|$, $^\wedge$
- Example: $p$:
  admonish $< v$ obj (p about against for)
- Example: $wds$
  (wds Word$_1$ Word$_2$ \ldots Word$_n$)
  (wds a bull in a china shop)
  $wds$ puts a requirement on all the words of the filler phrase
  $\rightarrow$ The argument words must match the filler phrase words exactly
10. Features - Grammatical Features

- Represent morphological or syntactical characteristics of the word
- Shared features: sg, pl, dual, cord, wh, whnom, ...
- Verb features: vfin, vfut ...
- Noun features: acc, cn, ...
- Adjective features: adjnoun, compar, ...
- Adverb features: locadv, compar, ...
- Determiner features: all, possdet, ...
- Qualifier features: pre, post, ...
- Subordinate conjunction features: assc, finsc, ...
- Preposition features: accobj, asprep, ...
10. Features - Semantic Features

- Also called semantic types or concepts
- Name set of things to determine the word sense, produce better parsers, choose target words in machine translations, ...
- $\text{Type}_1 \subset \text{Type}_2 \rightarrow h \subset \text{st\textunderscore anim}$
- $f < f_1 < f_2 < \ldots < f_n \rightarrow \text{man} < \text{male} < \text{human} < \text{animate}$
  The f’s need not be in order or related among one another
- A parser looks at each feature f, looks it up in the ontology and checks whether the option (e.g. g) is f or a superset of f
11. Subject Area Tests

The bat flew out of his hands

Subject areas are topics like entertainment or arts

The subject area stands for a set of entities in that topic
Example: mathematics → set of all mental activities, writings, discussion, ... in the mathematics domain

There can be more than one subject area test in a sense frame

- (sa Test): Test can be the name of the subject area, a flag or a boolean combination of tests (&, |, ^)
- (sa (| computers finance))

If the test fails, the sense frame will not be used in parsing
11. Subject Area Tests

- Flag Test
  \((fl \ Flag)\)
  \(\text{US} < \text{propn st\_country} (\text{sa} (\^ (fl ucseg)))\)

- Subject area tests can rule out, penalize or reward an option
  - Penalty: Sense frame is less likely to be the right one \(\rightarrow\) positive number
  - Reward: Sense frame is more likely to be the right one \(\rightarrow\) negative number
  - 0: Neither penalty nor reward

- \((\text{sa SubjAreaTest}_1 \text{ Penalty}_1 \text{ SubjAreaTest}_2 \text{ Penalty}_2 \ldots)\) \(\rightarrow\) \((\text{sa computers 2 finance})\)

- if(SubAreaTest\(_1\))
  Penalty\(_1\), Succeed
else if (SubAreaTest\(_2\))
  Penalty\(_2\), Succeed
12. Sense Names

- (sn SenseName)
  have < v (auxcomp ena) (sn have_perf) < ...
  auxiliary sense of have

- If no sense name is given, the index word and a number are taken as the name
  man < n h m < v obj1
  first: noun sense name: man1
  second: verb sense name: man2
13. Inflectional Elements

- Regular inflections of a word need not be given in the lexicon.
- Irregular forms are listed as index words:
  - ate < (ved eat) → ate as past tense of eat
  - made < (ven make) → made as past tense of make

- When the same word is citation form for one and inflected word for another word sense, both senses are shown in the entry.
- surprised
  - < adj (obj fin inf (p at about by))
  - < (veden surprise)
14. Support Word Frames

- *make use of, make reference to, take advantage of, ...*
- Information for support word frames are stored under the noun
  sup SupportWord SenseFrameForSupportWord
  `make use of`
  `use < ... < sup make v (obj1 (n use)) (p1 of)`
- The system automatically constructs a sense name, like `make_use` that shows support and index word