Selectional Restrictions
Selectional Restrictions

Introduction
Selectional Restrictions

Consider the two interpretations of:

I want to eat someplace nearby.

a) sensible:
   Eat is intransitive and “someplace nearby” is a location adjunct

b) Speaker is Godzilla
   Eat is transitive and “someplace nearby” is a direct object

How do we know speaker didn’t mean b) ?

Because the THEME of eating tends to be something edible
Selectional restrictions are associated with senses

• The restaurant serves green-lipped mussels.
  • THEME is some kind of food

• Which airlines serve Denver?
  • THEME is an appropriate location
Selectional restrictions vary in specificity

I often ask the musicians to *imagine* a tennis game.

To *diagonalize* a matrix is to find its eigenvalues.

Radon is an *odorless* gas that can’t be detected by human senses.
Representing selectional restrictions

Instead of representing “eat” as:

\[ \exists e, x, y \, \text{Eating}(e) \land \text{Agent}(e, x) \land \text{Theme}(e, y) \]

Just add:

\[ \exists e, x, y \, \text{Eating}(e) \land \text{Agent}(e, x) \land \text{Theme}(e, y) \land \text{EdibleThing}(y) \]

And “eat a hamburger” becomes

\[ \exists e, x, y \, \text{Eating}(e) \land \text{Eater}(e, x) \land \text{Theme}(e, y) \land \text{EdibleThing}(y) \land \text{Hamburger}(y) \]

But this assumes we have a large knowledge base of facts about edible things and hamburgers and whatnot.
Let’s use WordNet synsets to specify selectional restrictions

- The THEME of eat must be WordNet synset \{food, nutrient\} “any substance that can be metabolized by an animal to give energy and build tissue”

- Similarly
  - THEME of imagine: synset \{entity\}
  - THEME of lift: synset \{physical entity\}
  - THEME of diagonalize: synset \{matrix\}

- This allows
  - imagine a hamburger and lift a hamburger,

- Correctly rules out
  - diagonalize a hamburger.
Selectional Restrictions

Selectional Preferences
Selectional Preferences

• In early implementations, selectional restrictions were strict constraints (Katz and Fodor 1963)
  • Eat [+FOOD]

• But it was quickly realized selectional constraints are really preferences (Wilks 1975)
  • But it fell apart in 1931, perhaps because people realized you can’t eat gold for lunch if you’re hungry.
  • In his two championship trials, Mr. Kulkarni ate glass on an empty stomach, accompanied only by water and tea.
Selectional Association (Resnik 1993)

- **Selectional preference strength**: amount of information that a predicate tells us about the semantic class of its arguments.
  - *eat* tells us a lot about the semantic class of its direct objects
  - *be* doesn’t tell us much

- The selectional preference strength
  - difference in information between two distributions:
    - $P(c)$ the distribution of expected semantic classes for any direct object
    - $P(c|v)$ the distribution of expected semantic classes for this verb
  - The greater the difference, the more the verb is constraining its object
Selectional preference strength

- Relative entropy, or the Kullback-Leibler divergence is the difference between two distributions
  \[ D(P||Q) = \sum_x P(x) \log \frac{P(x)}{Q(x)} \]

- Selectional preference: How much information (in bits) the verb expresses about the semantic class of its argument
  \[ S_R(v) = D(P(c|v)||P(c)) = \sum_c P(c|v) \log \frac{P(c|v)}{P(c)} \]

- Selectional Association of a verb with a class: The relative contribution of the class to the general preference of the verb
  \[ A_R(v, c) = \frac{1}{S_R(v)} P(c|v) \log \frac{P(c|v)}{P(c)} \]
Computing Selectional Association

- A probabilistic measure of the strength of association between a predicate and a semantic class of its argument
  - Parse a corpus
  - Count all the times each predicate appears with each argument word
  - Assume each word is a partial observation of all the WordNet concepts associated with that word
  - Some high and low associations:

<table>
<thead>
<tr>
<th>Verb</th>
<th>Direct Object Semantic Class</th>
<th>Assoc</th>
<th>Direct Object Semantic Class</th>
<th>Assoc</th>
</tr>
</thead>
<tbody>
<tr>
<td>read</td>
<td>WRITING</td>
<td>6.80</td>
<td>ACTIVITY</td>
<td>-.20</td>
</tr>
<tr>
<td>write</td>
<td>WRITING</td>
<td>7.26</td>
<td>COMMERCE</td>
<td>0</td>
</tr>
<tr>
<td>see</td>
<td>ENTITY</td>
<td>5.79</td>
<td>METHOD</td>
<td>-0.01</td>
</tr>
</tbody>
</table>
Results from similar models

Ó Séaghdha and Korhonen (2012)

eat
food#n#1, aliment#n#1, entity#n#1, solid#n#1, food#n#2

drink
fluid#n#1, liquid#n#1, entity#n#1, alcohol#n#1, beverage#n#1

appoint
individual#n#1, entity#n#1, chief#n#1, being#n#2, expert#n#1

publish
abstract_entity#n#1, piece_of_writing#n#1, communication#n#2, publication#n#1
Instead of using classes, a simpler model of selectional association

• Model just the association of predicate \( v \) with a noun \( n \)
  (one noun, as opposed to the whole semantic class in WordNet)
• Parse a huge corpus
• Count how often a noun \( n \) occurs in relation \( r \) with verb \( v \):
  \[
  \log \text{count}(n,v,r)
  \]
• Or the probability:
  \[
  P(n|v,r) = \begin{cases} \frac{C(n,v,r)}{C(v,r)} & \text{if } C(n,v,r) > 0 \\ 0 & \text{otherwise} \end{cases}
  \]
Evaluation from Bergsma, Lin, Goebel

<table>
<thead>
<tr>
<th>Verb</th>
<th>Plaus./Implaus.</th>
</tr>
</thead>
<tbody>
<tr>
<td>see</td>
<td>friend/method</td>
</tr>
<tr>
<td>read</td>
<td>article/fashion</td>
</tr>
<tr>
<td>find</td>
<td>label/fever</td>
</tr>
<tr>
<td>hear</td>
<td>story/issue</td>
</tr>
<tr>
<td>write</td>
<td>letter/market</td>
</tr>
<tr>
<td>urge</td>
<td>daughter/contrast</td>
</tr>
<tr>
<td>warn</td>
<td>driver/engine</td>
</tr>
<tr>
<td>judge</td>
<td>contest/climate</td>
</tr>
<tr>
<td>teach</td>
<td>language/distance</td>
</tr>
<tr>
<td>show</td>
<td>sample/travel</td>
</tr>
<tr>
<td>expect</td>
<td>visit/mouth</td>
</tr>
<tr>
<td>answer</td>
<td>request/tragedy</td>
</tr>
<tr>
<td>recognize</td>
<td>author/pocket</td>
</tr>
<tr>
<td>repeat</td>
<td>comment/journal</td>
</tr>
<tr>
<td>understand</td>
<td>concept/session</td>
</tr>
<tr>
<td>remember</td>
<td>reply/smoke</td>
</tr>
</tbody>
</table>
Selectional Restrictions

Conclusion
Summary: Selectional Restrictions

• Two classes of models of the semantic type constraint that a predicate places on its argument:
  • Represent the constraint between predicate and WordNet class
  • Represent the constraint between predicate and a word

• One fun recent use case: detecting metonymy (type coercion)
  • Coherent with selectional restrictions:
    The spokesman denied the statement (PROPOSITION).
    The child threw the stone (PHYSICAL OBJECT)
  • Coercion:
    The president denied the attack (EVENT → PROPOSITION).
    The White House (LOCATION → HUMAN) denied the statement.

Pustejovsky et al (2010)