

CSC485 A3

Michael Dzamba - 994803806

November 2, 2009

1 Verb complements and gap features

- i. The student preferred to leave.
 - ii. The student persuaded the teacher to leave.
 - iii. The student promised the teacher to leave.
 - iv. The student expected the teacher to leave.
- A) For each of the sentences (ii)-(iv) above, state each role received from the verbs by the NPs *the student* and *the teacher*
- i. 'The student' is the agent of 'preferred' and of 'leave'.
 - ii. 'The student' is the agent of 'persuaded'.
'the teacher' is the theme of 'persuaded' and the agent of 'leave'.
 - iii. 'The student' is the agent of 'promised' and of 'leave'.
'the teacher' is the theme of 'promised'.
 - iv. 'The student' is the agent of 'expected'.
'the teacher' is the theme of 'expected' and the agent of 'leave'.
- B) The grammar provided below is quite minimal in order to provide only those features necessary to ensure the appropriate interpretation of NPs with respect to semantic roles. Additional argument feature structures can be easily added to support larger lexicon and more complex sentences, but they were not because they were not specifically requested in the assignment and I did not want to clutter the grammar. ϵ strings were used to store the SEM role of each gap NP, and the tree structure along with these gaps should be sufficient to distinguish the semantic role of each NP. Also there is an assumption made that if a feature relation copies a feature from another node in the parse tree, then at one point that feature

relation must fill in that copy, if not the parse is assumed to fail.

$S \rightarrow NP VP$	$[< VP \text{ GAP SEM } > = < NP \text{ SEM } >]$
$VP \rightarrow V VP_{to}$	$\left[\begin{array}{l} < V \text{ TV } > = \text{SG} \\ < VP \text{ GAP SEM } > = < VP_{to} \text{ GAP SEM } > \end{array} \right]$
$VP \rightarrow V NP VP_{to}$	$\left[\begin{array}{l} < V \text{ TV } > = \text{BI} \\ < V \text{ INF-USE-OBJ } > = \text{TRUE} \\ < VP_{to} \text{ GAP SEM } > = < NP \text{ SEM } > \end{array} \right]$
$VP \rightarrow V NP VP_{to}$	$\left[\begin{array}{l} < V \text{ TV } > = \text{BI} \\ < V \text{ INF-USE-OBJ } > = \text{FALSE} \\ < VP_{to} \text{ GAP SEM } > = < NP \text{ SEM } > \\ < NP \text{ GAP SEM } > = < VP \text{ GAP SEM } > \end{array} \right]$
$VP_{to} \rightarrow NP V_{to}$	$[< VP_{to} \text{ GAP SEM } > = < NP \text{ GAP SEM } >]$
$NP \rightarrow \epsilon$	$[< NP \text{ SEM } > = < NP \text{ GAP SEM } >]$
$NP \rightarrow \text{DET } N$	$[< NP \text{ SEM } > = < N \text{ SEM } >]$
$\text{DET} \rightarrow \text{'the'}$	
$N \rightarrow \text{'student'}$	$[< SEM > = \text{'student'}]$
$N \rightarrow \text{'teacher'}$	$[< SEM > = \text{'teacher'}]$
$V \rightarrow \text{'preferred'}$	$[< TV > = \text{SG}]$
$V \rightarrow \text{'persuaded'}$	$\left[\begin{array}{l} < TV > = \text{BI} \\ < INF-USE-OBJ > = < \text{TRUE} > \end{array} \right]$
$V \rightarrow \text{'promised'}$	$\left[\begin{array}{l} < TV > = \text{BI} \\ < INF-USE-OBJ > = < \text{FALSE} > \end{array} \right]$
$V \rightarrow \text{'expected'}$	$[< TV > = < \text{SG} >]$
$V_{to} \rightarrow \text{'to leave'}$	

C) For the parse tree please see attached diagram.

2 Playing with WordNet

- A) An example of the deepest leaf is " with a depth of 18 from the entity root node. An example of the shallowest leaf is " with a depth of 2 from the entity root node. The ratio of leaf nodes to total nodes in the tree rooted at the entity node is ≈ 0.797 , but when considering all nodes and not just those rooted at the entity the ratio is ≈ 0.776 .
- B) In the table below the scores of *path_ssimilarity* are called on each pair using word net. A important thing to note is that in the Charles and Miller rankings the rankings are in the range of [0,5] while word net similarities are in the range of [0,1]. Two of the most different rankings are the pairs 'lad-wizard' and 'furnace-stove'. The similarity of 'lad-wizard' seems to be

relatively much higher in word net then in the human trials. This may be due to an additional hypernym-hyponym connection that is not commonly thought of or used in modern english. The similarity of 'furnace-stove' is relatively much lower then what the human trials provided, this may be due to a lack of more direct relation between the two synsets in word net. In any case, addition or subtraction of more hypernym-hyponym relations may not solve these problems. This is because if every edge is considered to be the same weight as all others this may cause problems with their frequency and usage in modern english. For example although a direct hypernym-hyponym relation may exist between node A and B this does not mean that the relation is common knowledge and that the human trials would reflect this. Also a long distance relation may exist between two nodes A and B, that is very frequently used in modern english and could bias the nouns into having a lower score in human trials.

Noun-pair	WN score $\in [0, 1]$	Miller Score $\in [0, 5]$	Miller Rank	WN Rank
car-automobile	1.0	3.92	1	$\in [1, 4]$
gem-jewel	1.0	3.84	2	$\in [1, 4]$
magician-wizard	1.0	3.50	7	$\in [1, 4]$
midday-noon	1.0	3.42	8	$\in [1, 4]$
journey-voyage	0.50	3.84	3	$\in [5, 11]$
boy-lad	0.50	3.76	4	$\in [5, 11]$
coast-shore	0.50	3.70	5	$\in [5, 11]$
asylum-madhouse	0.50	3.61	6	$\in [5, 11]$
bird-cock	0.50	3.05	11	$\in [5, 11]$
tool-implement	0.50	2.95	13	$\in [5, 11]$
brother-monk	0.50	2.82	14	$\in [5, 11]$
bird-crane	0.25	2.97	12	
lad-brother	0.20	1.66	15	$\in [13, 18]$
crane-implement	0.20	1.68	16	$\in [13, 18]$
coast-hill	0.20	0.87	21	$\in [13, 18]$
shore-woodland	0.20	0.63	23	$\in [13, 18]$
monk-slave	0.20	0.55	24	$\in [13, 18]$
lad-wizard*	0.20	0.42	26	$\in [13, 18]$
coast-forest	0.17	0.42	25	19
monk-oracle	0.13	1.10	18	$\in [20, 21]$
glass-magician	0.13	0.11	28	$\in [20, 21]$
cemetery-woodland	0.11	0.95	19	$\in [22, 23]$
forest-graveyard	0.11	0.84	22	$\in [22, 23]$
furnace-stove**	0.10	3.11	9	$\in [24, 25]$
food-fruit	0.10	3.08	10	$\in [24, 25]$
chord-smile	0.091	0.13	27	26
noon-string	0.083	0.08	30	27
food-rooster	0.063	0.89	20	28
journey-car	0.056	1.16	17	29
rooster-voyage	0.042	0.08	29	30

3 Corpus based disambiguation

B) Where $v = \text{swam}$, $n = \text{whales}$ and $p = \text{onto}$

$$Pr(v, n, p, a) = Pr(v) \cdot Pr(n) \cdot Pr(a|v, n) \cdot Pr(p|a, v, n)$$

$$cl(v, n, p, n2) = \begin{cases} N & \text{if } p = \text{of} \\ \arg \max_{a \in \{N, V\}} Pr(v, n, p, a) & \text{otherwise} \end{cases}$$

$$= \begin{cases} N & \text{if } p = \text{of} \\ \arg \max_{a \in \{N, V\}} Pr(a|v, n) \cdot Pr(p|a, v, n) & \text{otherwise} \end{cases}$$

$$c() = 42, c(n, \text{true}) = 14, c(n, p, \text{true}) = 3, c(v) = 10, c(v, \text{true}) = 8, c(v, p, \text{true}) = 3$$

$$Pr(\text{true}|n) = \begin{cases} \frac{c(n, \text{true})}{c(n)}, & c(n) > 0 \\ 0.5, & \text{otherwise} \end{cases} \quad Pr(\text{true}|v) = \begin{cases} \frac{c(v, \text{true})}{c(v)}, & c(v) > 0 \\ 0.5, & \text{otherwise} \end{cases}$$

$$= \frac{14}{42} = \frac{1}{3} \quad \quad \quad = \frac{8}{10} = \frac{4}{5}$$

$$Pr(p|\text{true}, n) = \begin{cases} \frac{c(n, p, \text{true})}{c(n, \text{true})}, & c(n, \text{true}) > 0 \\ \frac{1}{|P|}, & \text{otherwise} \end{cases}$$

$$= \frac{3}{14}$$

$$Pr(p|\text{true}, v) = \begin{cases} \frac{c(v, p, \text{true})}{c(v, \text{true})}, & c(v, \text{true}) > 0 \\ \frac{1}{|P|}, & \text{otherwise} \end{cases}$$

$$= \frac{3}{8}$$

$$Pr(V|v, n) \cdot P(p|V, v, n) \approx \frac{Pr(\text{true}|v)}{Pr(\text{true}|n) + Pr(\text{true}|v)} \cdot Pr(p|\text{true}, v)$$

$$= \frac{\frac{4}{5}}{\frac{4}{5} + \frac{1}{3}} \cdot \frac{3}{8} = \frac{9}{34} \approx 0.265$$

$$Pr(N|v, n) \cdot P(p|N, v, n) \approx \frac{Pr(\text{true}|n)}{Pr(\text{true}|n) + Pr(\text{true}|v)} \cdot Pr(p|\text{true}, n)$$

$$= \frac{\frac{1}{3}}{\frac{4}{5} + \frac{1}{3}} \cdot \frac{3}{14} = \frac{15}{238} \approx 0.0630$$

Thus $cl(v, n, p, n2) = V$, the preferred attachment is to the Verb.

- C) i. For the example, 'placed seals onto', there are 11 occurrences of the word 'placed' in the text, but there are 0 occurrences of the word 'placed' in text that overlaps with a non-ambiguous propositional phrase gathered in training data. The backoff formulas are used only when the denominator is of 0 value, thus $Pr(p|true, v) > 0$ even though $c(v, true) = 0$ (by back-off), but $c(v) > 0$ and so $Pr(true|v) = 0$ and hence the verb attachment probability is 0.

$$\begin{aligned}
Pr(v, n, p, V) &= Pr(v) \cdot Pr(n) \cdot Pr(V|v, n) \cdot Pr(p|V, v, n) \\
&\approx Pr(v) \cdot Pr(n) \cdot \frac{Pr(true|v)}{Pr(true|n) + Pr(true|v)} \cdot Pr(p|true, v) \\
&= Pr(v) \cdot Pr(n) \cdot \frac{0}{Pr(true|n) + Pr(true|v)} \cdot Pr(p|true, v) = 0
\end{aligned}$$

For the example 'advanced whales for', there are 5 occurrences of the word 'advanced' in the text, and there are 3 occurrences of the word 'advanced' in distinct training examples, but there are 0 occurrences of the word 'advanced' with a propositional phrase 'for' attached to it in the training data. This causes the backoff not to be used when calculating $Pr(true|v)$ resulting in $Pr(p|true, v) = 0$ which causes the verb attachment probability to be 0.

$$\begin{aligned}
Pr(v, n, p, V) &= Pr(v) \cdot Pr(n) \cdot Pr(V|v, n) \cdot Pr(p|V, v, n) \\
&\approx Pr(v) \cdot Pr(n) \cdot \frac{Pr(true|v)}{Pr(true|n) + Pr(true|v)} \cdot Pr(p|true, v) \\
&= Pr(v) \cdot Pr(n) \cdot \frac{Pr(true|v)}{Pr(true|n) + Pr(true|v)} \cdot \frac{c(v, p, true)}{c(v, true)} \\
&= Pr(v) \cdot Pr(n) \cdot \frac{Pr(true|v)}{Pr(true|n) + Pr(true|v)} \cdot \frac{0}{c(v, true)} = 0
\end{aligned}$$

- ii. Using the smoothed estimates from 4.2.2 instead of those in 4.2.1, this will change one of the verb attachment probabilities from 0 to non-zero from above. This is because $Pr(p|true, v) \neq 0$ when $c(v, p, true) = 0$ in the example 'advanced whales for' using estimates from 4.2.2. In the case of 'placed seals onto' this will not change the fact that the verb attachment probability is zero, since the term causing this is $Pr(true|v) = 0$ and the calculation of this estimate is not changed in 4.2.2.
- iii. The cases for when the back-off formulas to be used in $Pr(true|a)$ are when $c(a) = 0$, when this happens then $c(a) = 0 \rightarrow c(a, true) = 0 \rightarrow c(a, p, true) = 0$, thus back-off formulas are used in both parts of the equation when $c(a) = 0$. If $c(a) \neq 0$ then to get a non-zero probability $c(a, true) > 0$ and if this is the case then no back-off formulas are used.