Chapter 9: Ambiguity Resolution

9.1 Selectional Restrictions

Word senses can be related in the different ways based on the object classes they describe.

- Some senses are disjoint: that is no object can be in both classes in the same time: DOG1 (sense of *dog*) and CAT1 (sense of *cat*).
- Other senses are subclasses of other senses: class DOG1 will be subclass of class MAMMAL1, and subclass of class PET1 (house pets).
- Other senses will overlap, such as: MAMMAL1 and PET1.
- All this knowledge can play a role in semantic disambiguation

Ambiguity Resolution

9.1 Selectional Restrictions

- The subset relation defines an abstraction hierarchy on the word senses.
- This relation is very important as it allows restriction to be state in terms of very broad classes.

For instance:

- adjective *purple* makes sense if it is modifying a physical object. It does not make sense: *purple ideas* or *purple event.*
- Adjective *precise* makes sense modifying *an idea* or *action*
- Adjective *infortunate* makes sense modifying *event* or *situation.*

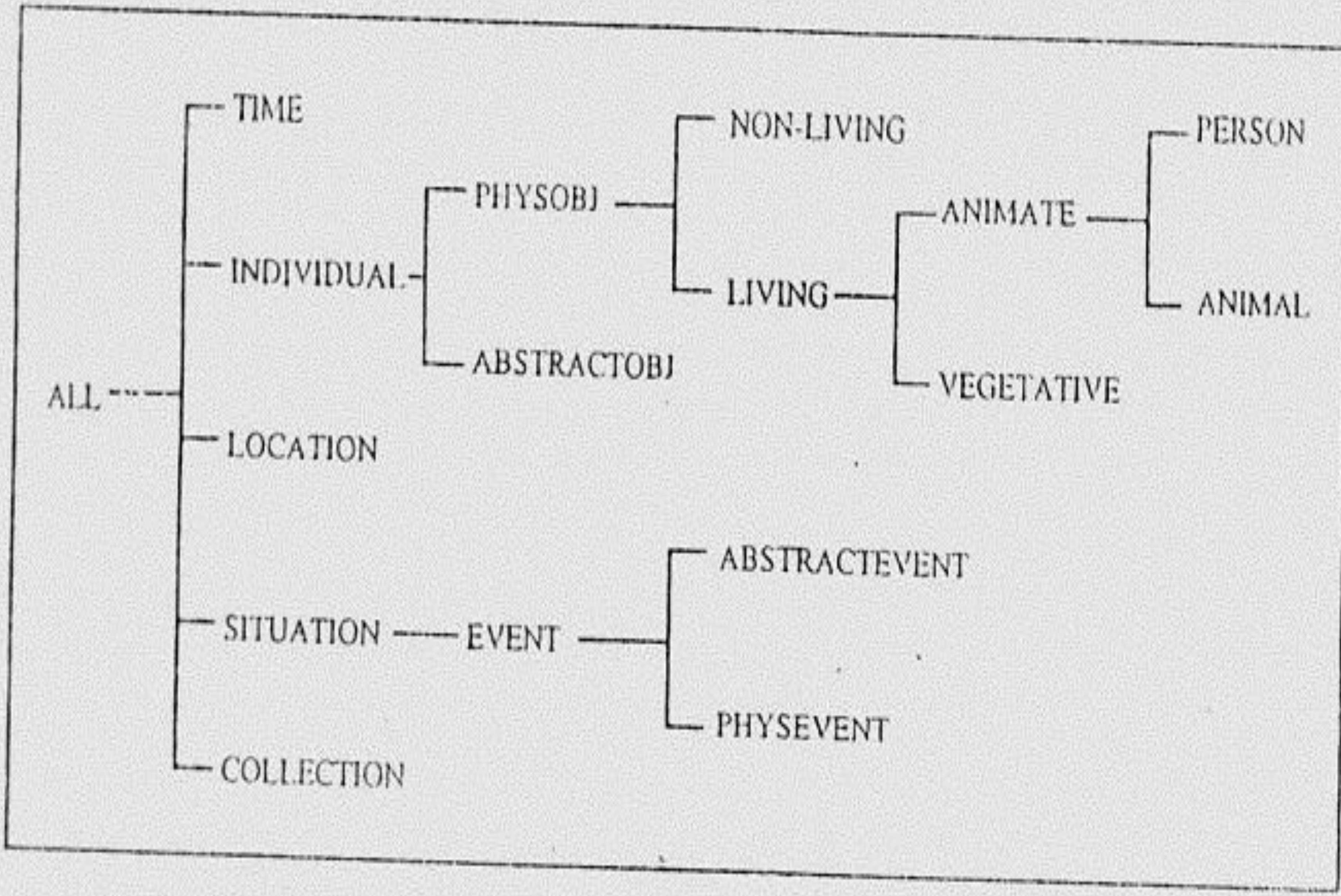


Figure 9.1: A word sense hierarchy

Ambiguity Resolution

9.1 Selectional Restrictions

Figure 9.1 shows a fragment of the top of type hierarchy that is useful for natural language. note that hierarchies need not be tree structures, that is, senses may have multiple super-type.

Example:

- MALE and FEMALE apply at level ANIMATE/VEGETATIVE
- ANIMATE and VEGETATIVE combine with these subclasses across the subclass LIVING…

Consider verb ***read.*** It has two principal arguments: the agent and the theme. The agent, which must be an object capable of reading (for something of *type PERSON*)

Ambiguity Resolution

9.1 Selectional Restriction

The theme must be an object that contains text (book, newspaper…).

To introduce a new type for handling correctly the verb ***read*:**

TEXTOB under NOLIVING, TEXTOB is a superset of BOOK, ARCLE/TEXT

Example:

- the noun *dishwasher has two senses;* either a machine (DISHWASH/MACHI) or a person (DISHWASH/PERS).
- The noun *article* can be a paper (ARTICL/TEXT) or a part of speech (ARTICLE1).

Ambiguity Resolution

9.1 Selectional Restriction

These senses are in figure 9.2. Since these two words are ambiguous, the sentence *The dishwasher read the article* may have four distinct semantic meanings, but only one reading makes sense, namely:

(READS1 [AGENT < THE d1 DISHWASH/ PERS >]

[THEME < THE p1 ARTICLE/ TEXT >])

The semantic interpreter can perform this form of disambiguation by using *selectional restriction.*

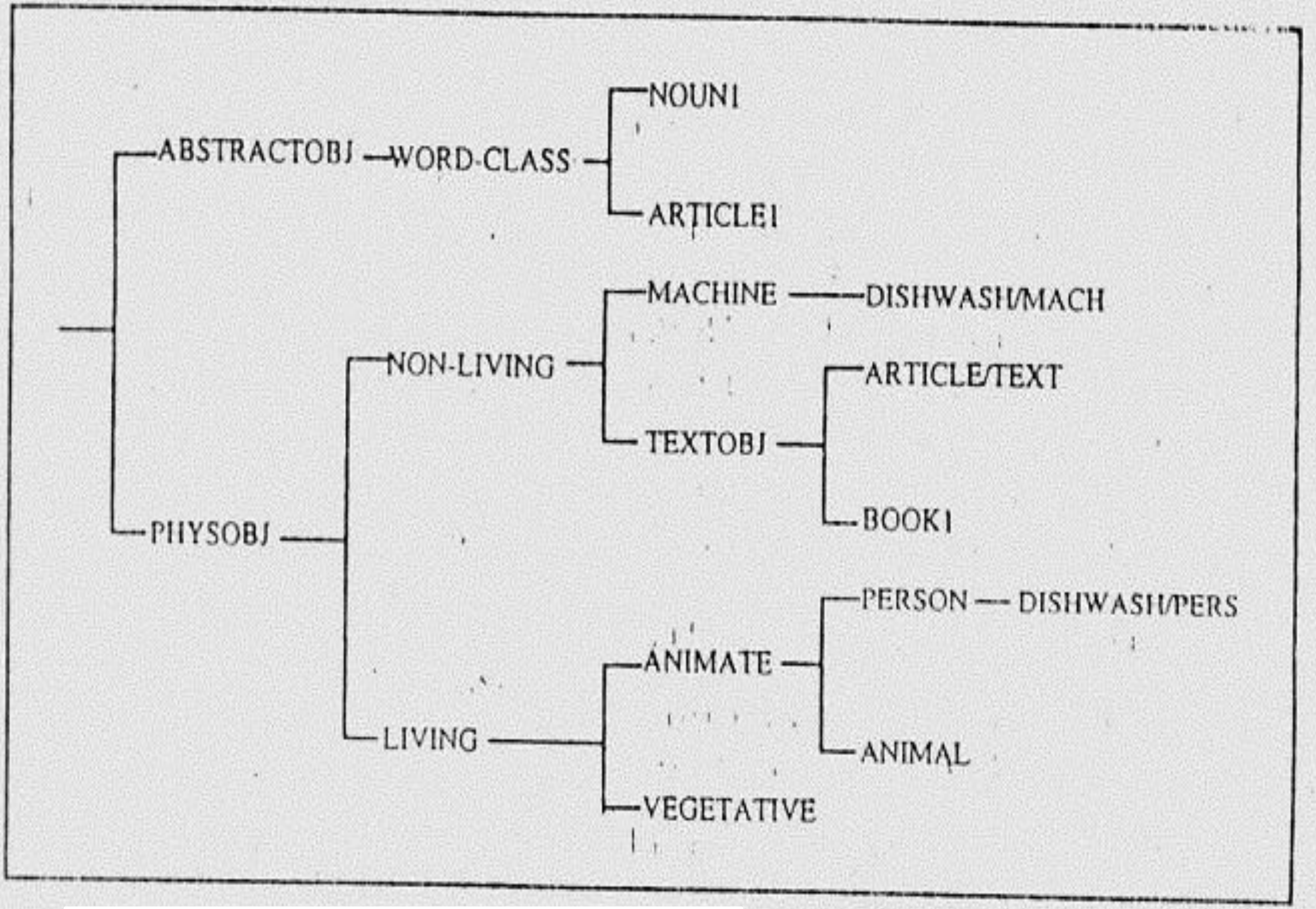


Figure 9.2: A fragment of the hierarchy

Ambiguity Resolution

9.1 Selectional Restriction

Example: The logical form of the sentence *The dishwasher read the article* before applying any *selectional restriction:*

(READS1 r1 [AGENT < THE d1 {DISHWASH/ MACH1 DISHWASH/ PERS } >] [THEME < THE p1 {ARTICLE/ TEXT ARTICLE1 }])

Unpacking the notation, the unary and binary relations are found:

(READS1 r1) ({DISHWASH/ MACH1 DISHWASH/ PERS}d1)

({ ARTICLE/ TEXT ARTICLE1} p1)

(AGENT r1 d1) (THEME r1 p1)

Ambiguity Resolution

9.1 Selectional Restriction

The allowable combinations can be viewed as a *constraint satisfaction problem.*

The selectional restrictions of READS1 are expressed as follows:

(AGENT READS1 PERSON)
(THEME READS1 TEXTOBJ)

For (AGENT r1 d1) to be valid, d1 must be person. Thus the unary constraints on d1 can be simplified from

({DISHWASH/ MACH1 DISHWASH/ PERS } d1) to (DISHWASH/ PERS d1).

Similarly, the interpretation of p1 is simplified to

(ARTICLE/ TEXT p1).

Ambiguity Resolution

9.1 Selectional Restriction

By transferring these constraints back into the logical form, we end up with a single unambiguous reading, as desired.

Note that the verb read has two senses READ1 and say READ2 as a form of understanding a person's intentions, as *Jill can read John's mind.* The selectional restrictions for READ2 might be

(AGENT READS2 PERSON)
(THEME READS2 MENTAL – STATE)

With the additional sense, the initial logical form of *The dishwasher read the article* is:

((READS1 READS2) r1[AGENT <THE d1{DISHWASH/ MACHI DISHWASH/PERS} >] [THEME < THE p1{ARTICLE/ TEXT ARTICLE}>])

Ambiguity Resolution

9.1 Selectional Restriction

This additional ambiguity does not effect the final result, because the READ2 requires a M.ENTAL-STATE as a THEME.

- We also need extend this technique to pronouns, proper noun, adjectives.

Example 1: Proper name: John might be MALE, that is animate object.

Unknown name might just default to having proper name INDIVIDUAL

Example 2: The pronoun: SHE1 should be a subclass of FEMALE
IT1 would be anything but PERSON

Ambiguity Resolution

9.1 Selectional Restriction

Example 3: for adjective: using the state variable representation and new thematic relation MOD.

+ ***happy dishwasher;*** instead of using predicate-argument form (HAPPY1 d1), we use the unary relation (HAPPY- STATE h1) and binary relation (MOD h1 d1).

Example 4: the set relations derived from the sentence ***The happy dishwasher read the paper*** would be:

(READS1 r1) ({DISHWASH/CHI DISHWASH/PERS} d1)

({ARTICLE/TEXT ARTICLE1} p1) (HAPPY-STATE h1)

(AGENT r1 d1) (THEME r1 p1) (MOD h1 d1)

Ambiguity Resolution

9.1 Selectional Restriction

The selectional restriction for ***happy dishwasher*** would be:

(MOD HAPPY - STATE ANIMATE)

HAPPY – STATE must modify an animate object.

To explore the constraint satisfaction algorithm in a little more detail in figure 9.3.

As example, consider running this algorithm on the sentence:

The dishwasher read the article.

Initialization Step

Assign types(variable$_i$) to the list of possible senses for variable i.

Iteration Step

Iterate through each binary relation (rel variable$_1$ variable$_2$):

1. For each sense$_1$ in types(variable$_1$):
   a. find all selectional restrictions (rel sense$_1$ sense$_2$) where sense$_2$ intersects with some sense in types(variable$_2$).
   b. If none found, remove sense$_1$ from types(variable$_1$).
2. Eliminate from types(variable$_2$) any sense that did not match at least one restriction in step 1.

Termination Step

If any changes were made to the types of the variables in the last iteration, then perform the iteration step once again.

Otherwise, if type(variable$_i$) is empty for any i, then fail.

Figure 9.3: A simple constraint satisfaction algorithm

Ambiguity Resolution

9.1 Selectional Restriction

The initial step produces the following types:

type (r1) = READS1 , READS2
type (p1) = ARTICLE/ TEXT , ARTICLE1
type (d1) = DISHWASH/ PERS , DISHWASH/ MACH1

- Interation step (first time)

There are binary relations:

(AGENT r1 d1) and (THEME r1 p1)

For (AGENT r1 d1), we iterate through the senses of r1: READ1 and READ2:

Ambiguity Resolution

9.1 Selectional Restriction

+ READS1 – we find selectional restriction (AGENT READS1 PERSON); PERSON matches only DISHWASH/ PERS (with result DISHWASH/ PERS)

+ READ2 – we find selectional restriction (AGENT READS2 PERSON) and PERSON matches DISHWASH/ PERS (with result DISHWASH/ PERS)

Thus the type (d1) becomes (DISHWASH/ PERS) , that is, DISHWASH/ MACH1 has been eliminated because it can not satisfy any binary constraint.

Ambiguity Resolution

9.1 Selectional Restriction

For (THEME r1 p1), we iterate through the senses of r1

+ READS1 – we find selectional restriction (THEME READS1 TEXTOBJ), TEXTOBJ matches ARTICLE/ TEXT (with result ARTICLE/ TEXT)

+ READS2 –we find no matching selectional restriction, that is, (THEME READS2 MENTAL - STATE) can not be satisfied. Thus type (r1) becomes (READS1). READS2 is eliminated and type (p1) becomes (ARTICLE/ TEXT) because ARTICLE is eliminated.

since changes we made, we iterate again.

Ambiguity Resolution

9.1 Selectional Restriction

- Interation step (second time)

For (AGENT r1 d1), only one sense of r1 remains

+ READS1 – we find selectional restriction

(AGENT READS1 PERSON)

For (THEME r1 p1)
+ READS1 –we find selectional restriction

(THEME READS1 TEXTOBJ)

Since no change we made this time, we are don. The final types are:

Ambiguity Resolution

9.1 Selectional Restriction

type (r1) = READS1
type (p1) = ARTICLE/ TEXT
type (d1) = DISHWASH/ PERS

Selectional restrictions are also very useful for further refining the type of unknown object.

Example: *He read it*

Assuming just the READS1 sense of the verb. The logical form of *He read it* would be:

(READS1 r3 [AGENT(PRO i1 HE1)][THEME(PRO n1(IT1 n1))])

The unary and binary constraints on the objects are: (READS1 r3), (AGENT r3 i1), (THEME r3 n1), (MALE i1), (IT1n1)

Ambiguity Resolution

9.1 Selectional Restriction

From sense (READS1 r3) □r3. r3 and (AGENT r3 i1) □ i1□he (from [AGENT(PRO i1 HE1)]) and (MALE i1) then type of he will be constrained to be of type MALE – PERSON (is intersection of MALE and PERSON).

From sense (READS1 r3) > r3, and (THEME r3 n1) □ n1, and THEME(PRO n1(IT1 n1))] □ IT, that is the type of it will be constrained to be a TEXTOBJ (is intersection IT vaøTEXTOBJ)

Thus, after applying the selectional restrictions, the logical form of the sentence would be :

(READS1 r3 [AGENT(PRO i1 (&(MALE i1) (PERSON i1))]

[THEME(PRO n1(&(IT1 n1)(TEXTOBJ n1)))])

Ambiguity Resolution

9.2 Semantic Filtering Using Selectional Restrictions

- Two ways that selectional restrictions can be added to a parser: sequential model and incremental model.

An incremental model

Consider the sentence *He booked a flight to the city for me.*

PPs to be attached to either VPs or NPs.

PP - *to the city* may modify the verb *booked* or noun *flight.*

PP- for me may modify noun *city* or verb *booked.*

There are five ways these possibilities into a legal syntactic structure, however there is only one plausible reading, that would be: ***flight to the city*** **and** ***booked for me***

Ambiguity Resolution

9.2 Semantic Filtering Using Selectional Restrictions

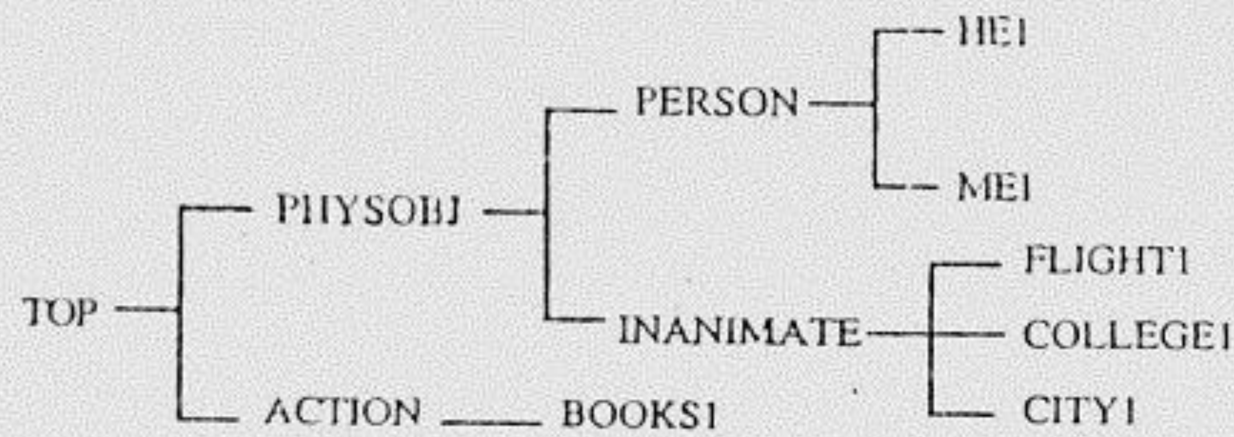
The selectional restriction is implemented on verb *booked*, nouns *flight* and *city.* The selectional restrictions for the sentence *He booked a flight to the city for me,* that would be:

(AGENT BOOKS1 PERSON1)
(THEME BOOKS1 FLIGHT1)
(BENEFICIARY ACTION1 PERSON1)
(DESTINATION FLIGHT1 CITY1)
(NEARBY PHYSOBJ PHYSOBJ)
(NEARBY ACTION PHYSOBJ)

1. (S SEM ?semvp) → (NP SEM ?semsubj) (*VP SUBJ ?semsubj SEM ?semvp*)
2. (VP SUBJ ?semsubj SEM ?semv) → (*V[_none] SUBJ ?semsubj SEM ?semv*)
3. (VP SUBJ ?semsubj SEM ?semv) →
   (*V[_np] SUBJ ?semsubj OBJ ?semnp SEM ?semv*) (NP SEM ?semnp)
4. (NP VAR ?v SEM (PRO ?v (?sempro ?v))) → (*PRO SEM ?sempro*)
5. (NP VAR ?v SEM <?semart ?v ?semcnp>) → (ART SEM ?semart) (*CNP SEM ?semcnp*)
6. (CNP VAR ?v SEM (?semn ?v)) → (*N SEM ?semn*)
7. (CNP SEM (& ?semcnp ?sempp)) →
   (*CNP VAR ?v SEM ?semv*) (PP PRED + ARGVAR ?v SEM ?sempp)
8. (VP SEM (& ?semvp ?sempp) →
   (*VP VAR ?v SEM ?semvp*) (PP PRED + ARGVAR ?v SEM ?sempp)
9. (PP PRED + ARGVAR ?v1 SEM (?semp ?v1 ?semnp) →
   (*P SEM ?semvp*) (NP SEM ?semnp)

Head features for S, VP, NP, CNP: VAR

Grammar 9.4: A small grammar allowing PP attachment ambiguity



a	(ART AGR 3s SEM INDEF1)
booked	(V SUBCAT _np VFORM past SUBJ ?subj OBJ ?obj SEM (& (BOOKS1 *) (AGENT * ?subj) (THEME * ?semobj)))
city	(N AGR 3s SEM CITY1)
college	(N AGR 3s SEM COLLEGE1)
flight	(N AGR 3s SEM FLIGHT1)
for	(P PFORM for SEM BENEFICIARY)
he	(PRO AGR 3s SEM HE1)
me	(PRO AGR 1s SEM ME1)
near	(P PFORM near SEM NEARBY)
the	(ART AGR {3s 3p} SEM THE)
to	(P PFORM to SEM DESTINATION)

Figure 9.5: A small lexicon and word sense hierarchy

Ambiguity Resolution

9.2 Semantic Filtering Using Selectional Restrictions

Given the grammar 10.4, consider the bottom up chart parser on the sentence *He booked the flight to the city for me.*

- Without semantic filtering parser finds five different interpretations and generates 52 constituents on the chart.
- With semantic filtering parser finds one interpretations and generates 33 constituents.

Consider the first constituent parser suggested by the parser that is rejected by semantic filtering:

(VP SEM (BOOKS1 V258) [AGENT ?semsubj]
[THEME < INDEF1 V260(FLIGHT1 V260) >]
[DESTINATION<THE V263(CITY V263)>])
VAR v258 SUBJ ? Semsubj)

Ambiguity Resolution

9.2 Semantic Filtering Using Selectional Restrictions

There are constituents, such as: VP- *book the flight*, *PP- to the city*. It is rejected because it violates the selectional restrictions on the DESTINATION predicate, that is (DESTINATION BOOKS1 CITY1) is not matched any selectional restriction.

Ambiguity Resolution

9.3 semantic Networks

Semantic networks ease the construction of the lexicon by following inheritance of the properties.

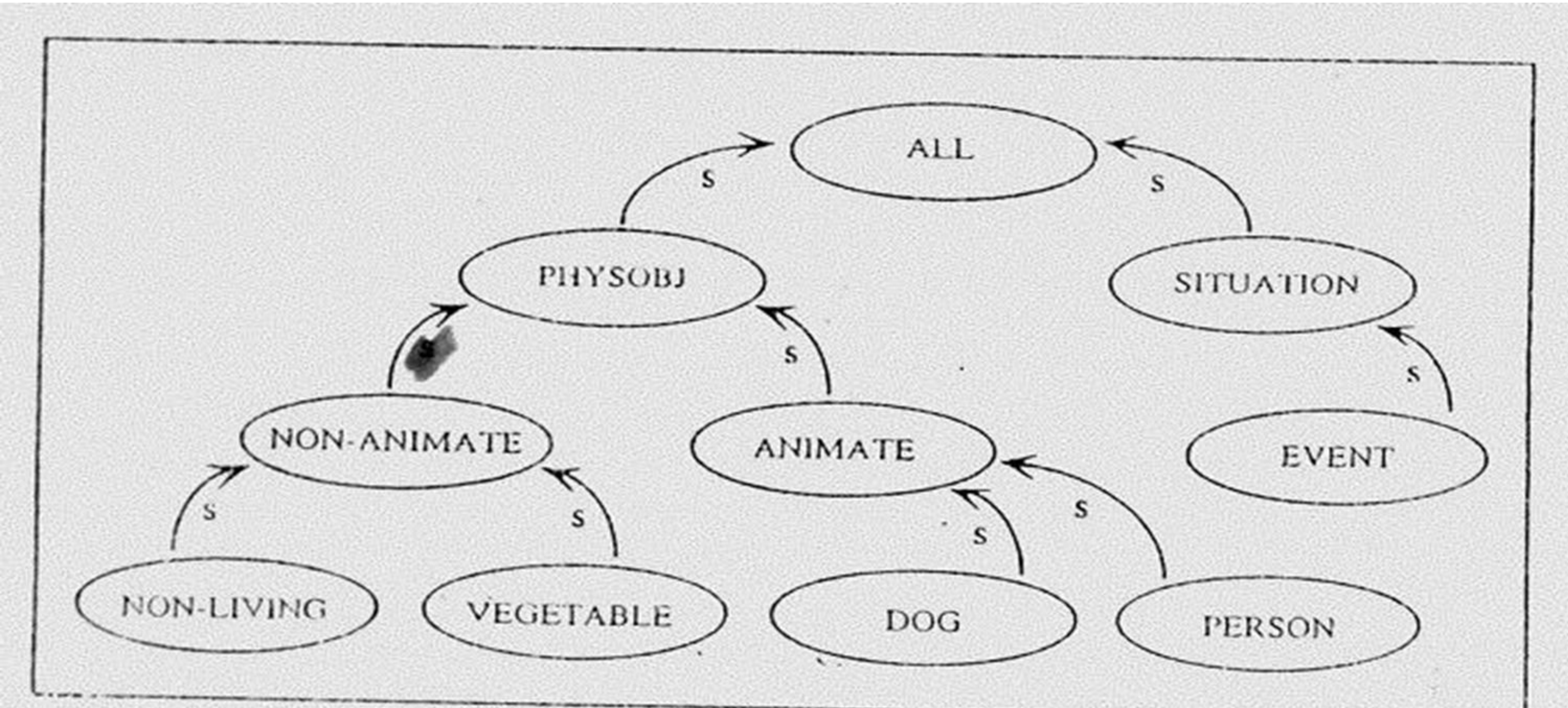


Figure 9.6: Part of a type hierarchy

Ambiguity Resolution

9.3 semantic Networks

- In figure 9.6, the s arc indicates the subtype relationship.
- The Selectional restrictions for semantic relations can be in a network form using arcs.

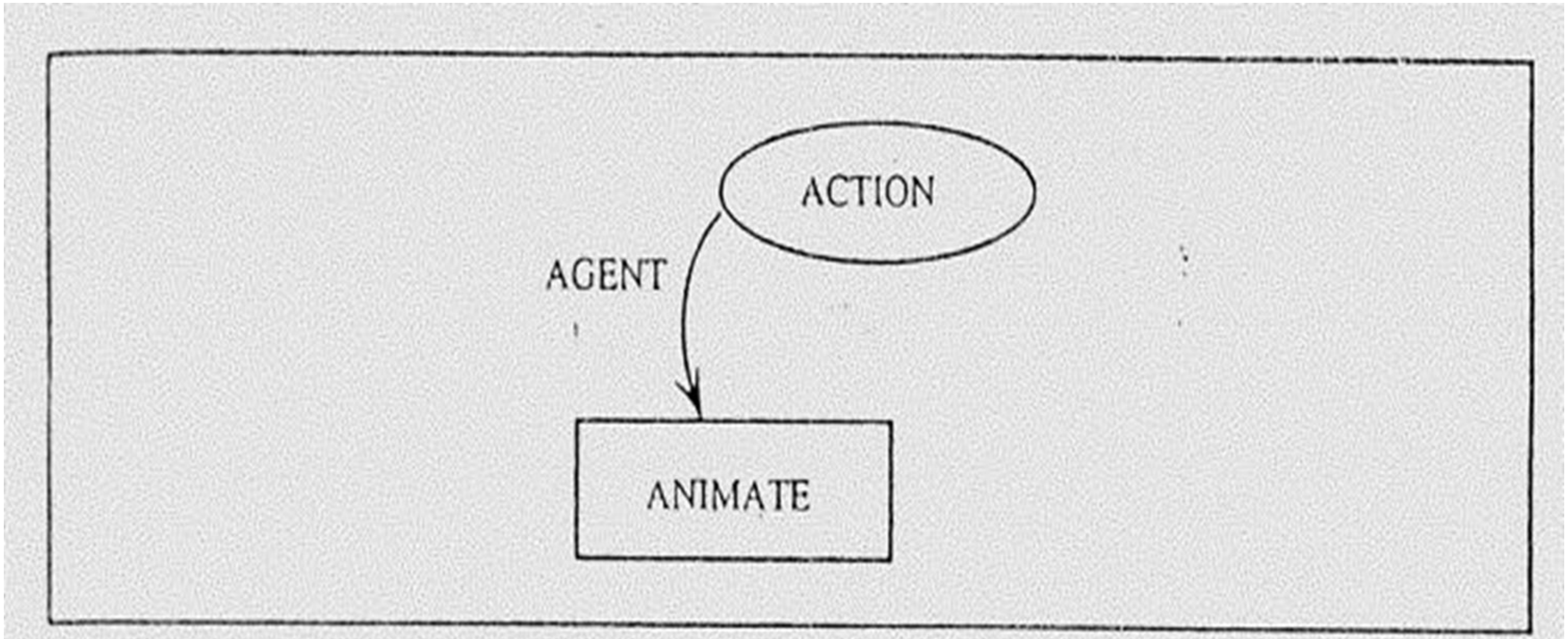


Figure 9.7: All actions have an animate agent

Ambiguity Resolution

9.3 semantic Networks

Figure 9.7 introduces here is a new node type, an extential node, depicted by a square, which represents a particular value.

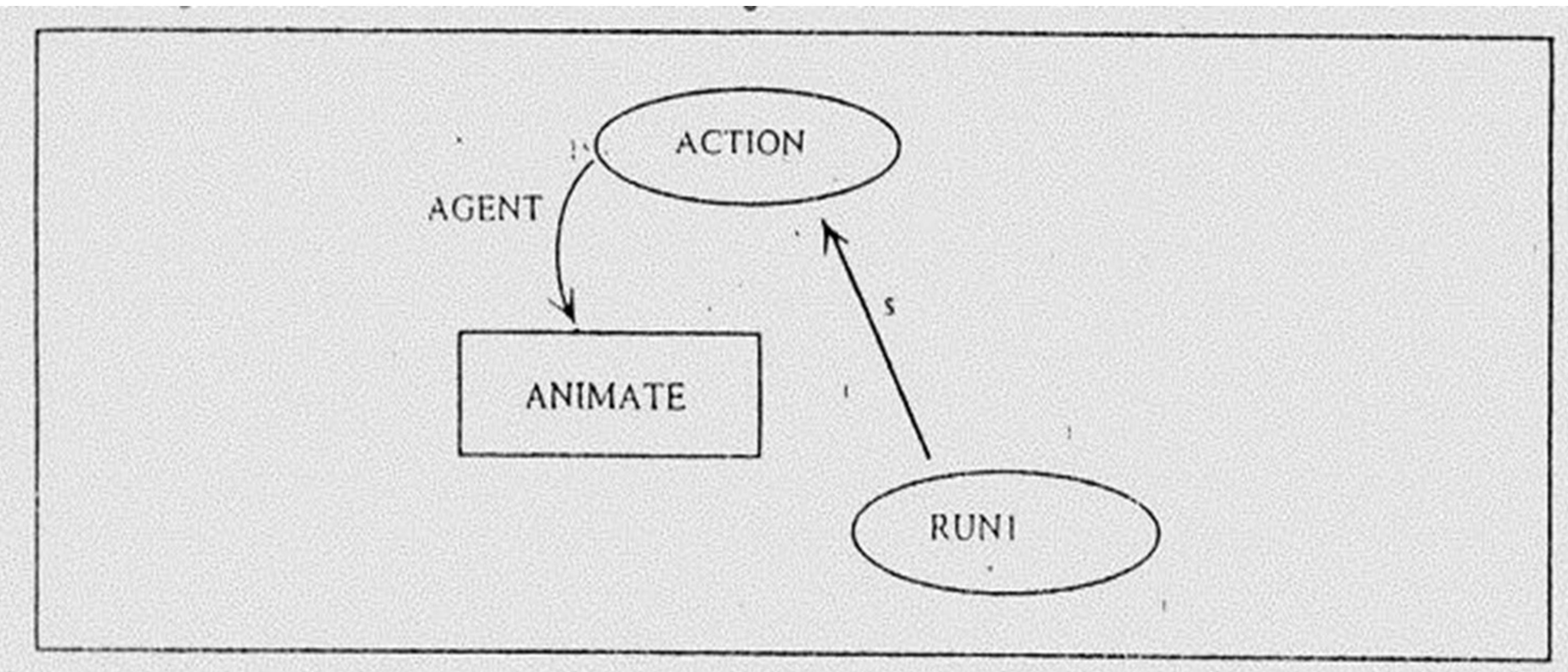


Figure 9.8: A network showing inheritance of roles

Ambiguity Resolution

9.3 semantic Networks

- An important property of semantic networks is the inheritance of the properties.

- Given the network is shown in the figure 9.8, the action class RUNS1 would inherit the property that every instance has an AGENT role filled by an ANIMATE object.

- Inheritance hierarchies are extremely useful for expressing selectional restrictions across abroad classes of verbs.

- Figure 9.9 shows the selectional restriction for a set of verb senses that are subslasses of ACTION.

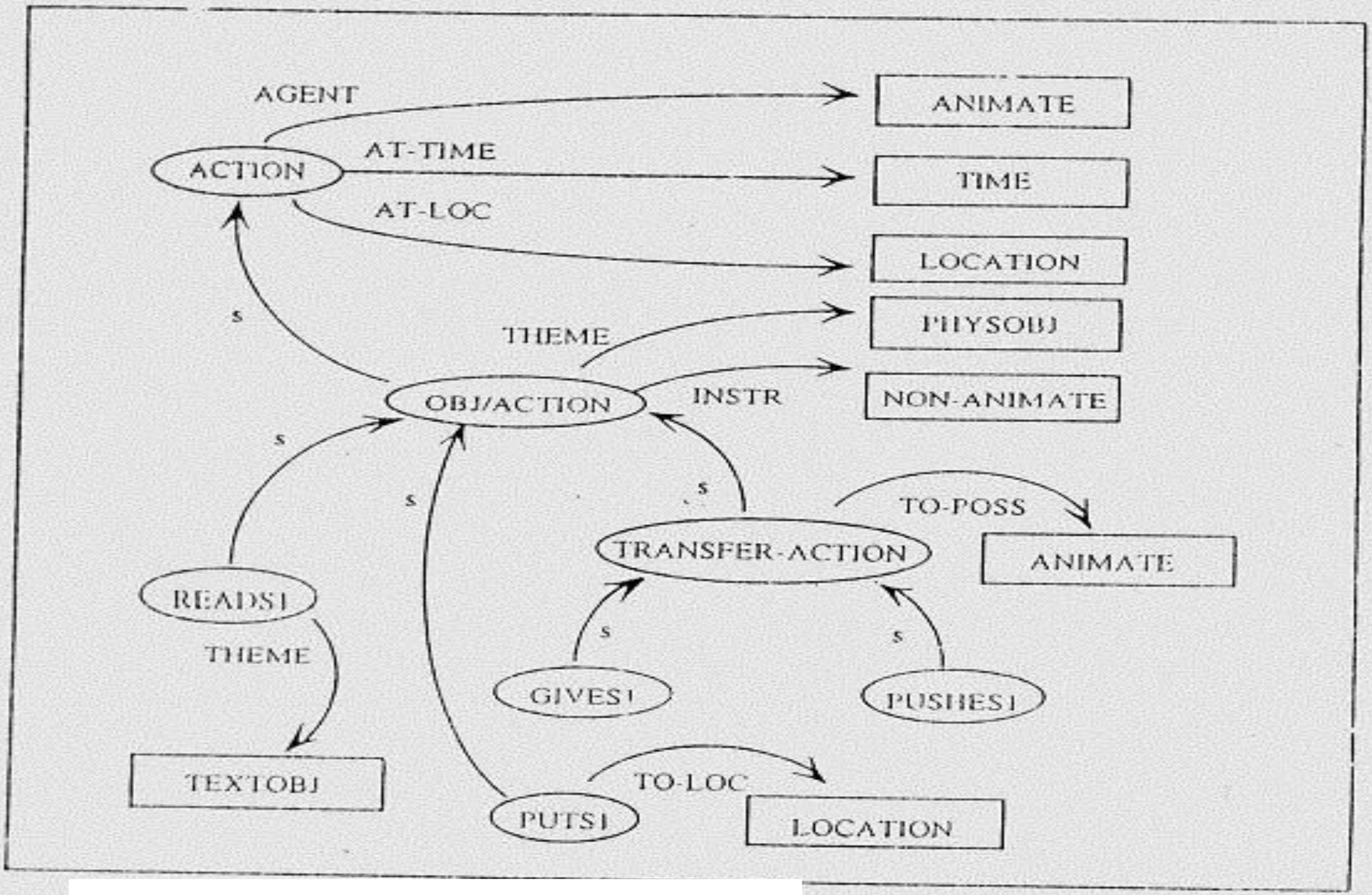


Figure 9.9: Action hierarchy with roles

Ambiguity Resolution

9.3 semantic Networks

In the figure 9.9, using the inheritance mechanism, we can see the action class TRANSFER-ACTION alows semantic relations AGENT, AT-TIME, AT-LOC inherited form the class ACTION. THEME and INSTR inherited form the class OBJ/ACTION. The case TO-POSS is explicitly defined .for TANSFER-ACTION.

- Another important is part of hierarchy in which objects are related to their subparts (figure 9.10).

Example: *The desk drawer* (the drawer is a part of the desk)

- *The man's head (*the head is part of the man)

- *The handle of the drawer* (the handle is a part of the drawer)

Ambiguity Resolution

10.3 semantic Networks

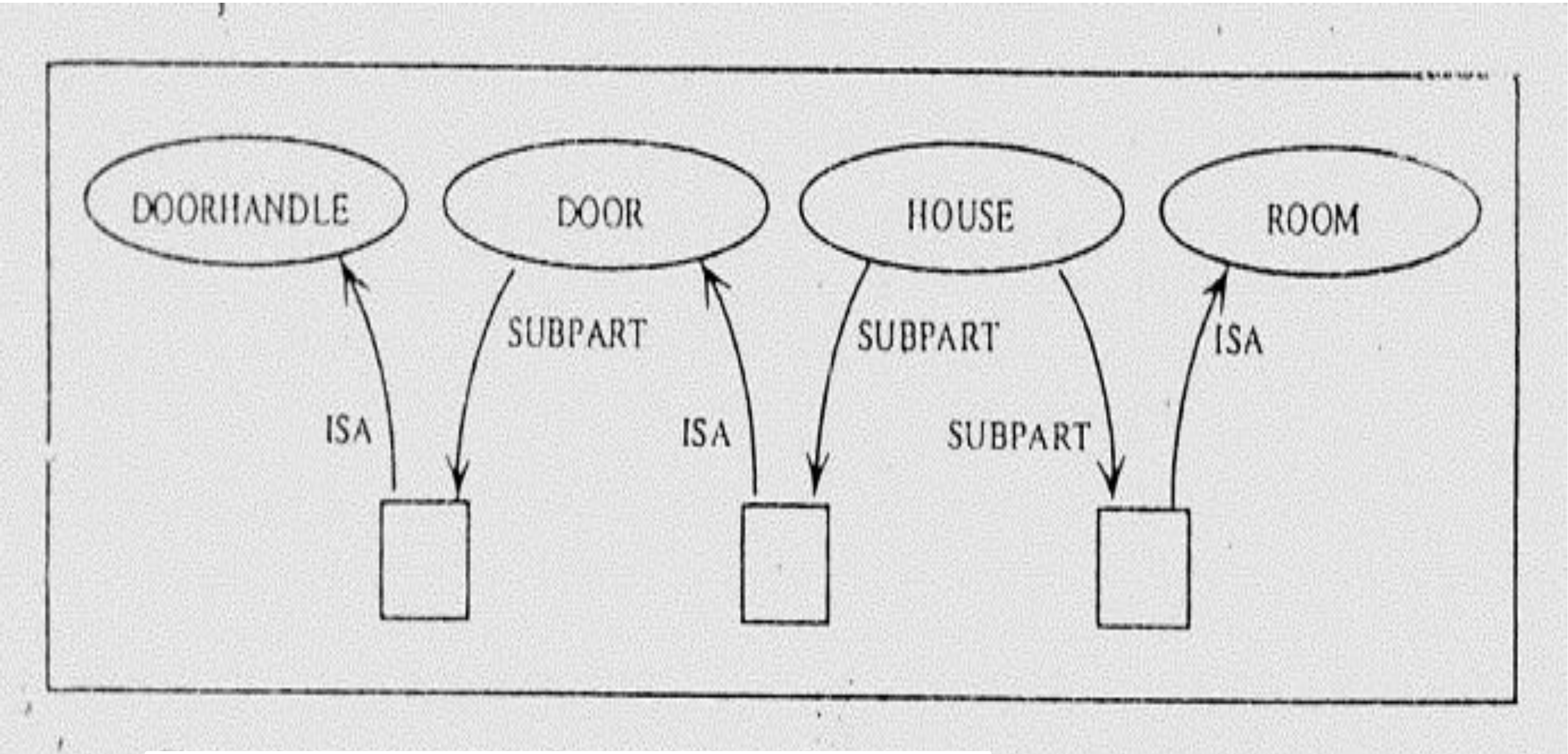


Figure 9.10: Some subpart relationships

Ambiguity Resolution

9.4 Statistical Word Sense Disambiguity

- Selectional restrictions provide only a coarse classification of acceptable and unacceptable form, many cases of sense ambiguity cannot be resolved.
- To better model human processing, more predictive techniques must be developed that give a preference for the common interpretation of senses over rarer senses. Thus, the way to use is statistic technique.
- The simplest techniques are based on simple unigram statistics. Given a suitable labeled corpus. We collect information on usage of the different senses of each word.
- ***Example:*** there are 5845 uses of the word bridge.

  5651 uses of STRUCTURE1
  194 uses of DENTAL – DEV37

Ambiguity Resolution

9.4 Statistical Word Sense Disambiguity

Given this data, we would guess that bridge occurs in the STRUCTURE1 sense every time and has 97% (5651 times/5845 times).

We would like to do much better than this by including some effect of context.

Consider the rare sense DENTAL–DEV37, it occurs very rarely in the entire corpus. But in the certain texts (dentistry or orthodontics), it will be the most common sense of the word.

- It is concerned with word collocations.

Collocation: that is what words would tend to appear together. We may consider bigram probabilities, trigrams or large groups, say five surrounding words. The amount of text examined for each word is called the window.

Ambiguity Resolution

9.4 Statistical Word Sense Disambiguity

- To adapt part-of speech-tagging techniques to use word senses rather than syntactic categories.

    + To need a corpus of words tagged with their senses.

    + Then we could compute unigram, bigram statistics (the probability that word w has sense s).

- Estimating the probability of the senses of a word w relative of a window of the word in the text centered on w.

- Given a window size of n centered on the word w , the words in the window are indicated as follows:

$$W_1 \ W_2 \cdots W_{n/2} \ \underline{W} \ Wn_{/2+1} \cdots W_{n-1}$$

Ambiguity Resolution

9.4 Statistical Word Sense Disambiguity

We want to compute the sense s of word w that masimazes the formula:

$$\text{PROB}(\, w/S \mid w_1 \;\; w_2 \ldots w_{n/2} \;\; w \;\; w_{n/2+1} \ldots w_{n-1} \,)$$

We rewrite the formula by using Baye's rule and then make independence assumptions, the formula becomes:

$$\frac{\text{PROB}(\, w_1 \ldots w_{n-1} \mid w/S \,) * \text{PROB}(\, w/s \,)}{\text{PROB}(\, w_1 \ldots w_{n-1} \,)}$$

$\text{PROB}(\, w_1 \ldots w_{n-1} \,)$ is not change for each sentence.

Assuming that each word w_i appear independently of other words in the window.

Ambiguity Resolution

9.4 Statistical Word Sense Disambiguity

$$\text{PROB}(w_1 \ldots W_{n-1} \mid w/S) = \prod_{i=1,n-1} \text{PROB}_n(w_i \mid w/s)$$

$\text{PROB}_n(w_i \mid w/S)$ is the probability that word w_i occurs in a n-word window centered on word w in sense S.

The best sense S will be the one that maximizes the formula:

$$\text{PROB}(w/s) * \prod_{i=1,n-1} \text{PROB}(w_i \mid w/S)$$

$$\text{PROB}_n(w_i \mid w/s) = \frac{\text{Count (\#times wi in a window centered on w/S)}}{\text{Count (\#times w/S is the center of a window)}}$$

Given the data in figure 9.11, we will find a sense for the word *bridge,* by using the window size 11 words in the corpus with 10.000.000 words

	with *bridge*/ STRUCTURE1	with *bridge*/ DENTAL-DEV37	in any window
teeth	1	10	300
suspension	200	1	2000
the	5500	180	500,000
dentist	2	35	900
total occurrences	5651	194	501,500

Figure 9.11: The counts for the senses for *bridge* in a hypothetical corpus

Ambiguity Resolution

9.4 Statistical Word Sense Disambiguity

Given the data in figure10.11, we get the following estimates:

$PROB_n$ (teeth/bridge/ STRUCTURED) = 1/ 5651 = 1.77 * 10^{-4}

$PROB_n$ (teeth/bridge/ DENTAL – DEV37) = 10/ 194 = 0.052

$PROB_n$ (suspension/bridge/ STRUCTURE) = 200/ 5651 = 0.35

$PROB_n$ (suspension/bridge/DENTAL – DEV37) = 1/194 = 5.15 * 10^{-3}

$PROB_n$ (the/bridge/ STRUCTURE1) = 5500/ 5651 = 0. 97

$PROB_n$ (the/bridge/ DENTAL – DEV37) = 180/194 = 0.93

$PROB_n$ (dentist/ bridge / STRUCTURE1) = 2/5651 = 3.54.10^{-4}

$PROB_n$ (dentist/ bridge / DENTAL – DEV37) = 35/194 = 0.18

$PROB_n$ (bridge/ STRUCTURE1) = 5651 / 501500 = 0.113

PROBn (bridge/ DENTAL – DEV37) = 194/ 501500 = 3.87 * 10^{-4}

$PROB_n$ (the/bridge/STRUCTURE1)*PROB(bridge/STRUCTURE1)= 0.97* 0.113 = 0.109

$PROB_n$ (the/bridge/ DENTAL – DEV37) * PROB (bridge/ DENTAL – DEV37) = .93 * 3.87 * 10^{-4} = 3.6 * 10^{-4}

Ambiguity Resolution

9.4 Statistical Word Sense Disambiguity

The context independent probabilities of the word senses are easily estimated:

PROB (bridge/ STRUCTURE1) = 5651/501500 = **0.113**

PROB(bridge/ DENTAL – DEV37) = 194/501500 = **3.878*10-4**

Note that the probability estimates for the senses in the window that contains the word *the* are very similar to the no-context estimate:

$PROB_n$ (the/bridge/ STRUCTURE1)* $PROB_n$ (bridge/ STRUCTURE1)

= 0.97*0.113 = **0.109**

$PROB_n$ (the/bridge/ DENTAL – DEV37)* PROBn (bridge/ DENTAL – DEV37) = 0.93* 3.87 * 10^{-4} = **3.6* 10-4**

Ambiguity Resolution

9.4 Statistical Word Sense Disambiguity

It is content words, like *teeth* in this example that has the most dramatic effect. For instance:

$PROB_n$(dentist/bridge/STRUCTURE1)*PROB(bridge/STRUCTURE1)

$= 3.54.10^{-4} * 0.113 = \mathbf{4 * 10^{-5}}$

$PROB_n$(dentist/bridge/DENTAL-DEV37)*PROB(bridge/DENTAL-DEV37) $= 0.18 * 3.87 * 10^{-4} = \mathbf{6.97 * 10^{-5}}$

Of course, with a larger window, there are many more chances for content words that strong effect the decision.

Example: *The dentist put a bridge on my teeth*

The words *teeth* and *dentist* together in the same window combine to strongly prefer the rare sense of the word *bridge*.

Ambiguity Resolution

9.4 Statistical Word Sense Disambiguity

In the fact, the estimate for the sense DENTAL-DEV37 would be 3.6*10-6, considerably greater than the estimate of 7.08*10-7 for STRUCTURE1.

Collocations and Mutual Information

In the area uses collocations, which measure how likely two words are to co-occur in a window of text. One way to compute such a measure is to consider a correlation statistic (where n is the window size).

$$Cn\ (w/\ S, w') = \frac{\text{PROB (w/ S \& w' are in the same window)}}{\text{PROB(w/S in the window)*PROB(w' in the window)}}$$

Ambiguity Resolution

9.4 Statistical Word Sense Disambiguity

If K is the number of windows in the corpus, then each of the probabilities above could be:

Count (#times event occurs in window)/K

After substituting such estimates in for each probability uses in Cn (w/ S, w'), simplifying we get the formula:

$$Cn(w/S, w') = \frac{K * Count(\#times\ w/\ S\ \&\ w'\ co\text{-}occur\ in\ window)}{Count(\#times\ w/S\ in\ window * Count(\#times\ w'\ in\ window)}$$

In our sample corpus K is 10^7. Base on the date in figure 9.11, the estimates for Cn are as follows:

Ambiguity Resolution

9.4 Statistical Word Sense Disambiguity

$PROB_n$(bridge/ STRUCTURED,teeth)=(10^7 *1)/(5651*300)= 5.9

$PROB_n$ (bridge/DENTAL-DEV37,teeth)=(10^7 *10)/(194*300)=171.9

$PROB_n$(bridge/ STRUCTURED, suspension) =(10^7 *200)/(5651*2000)=17.7

$PROB_n$ (bridge/DENTAL-DEV37, suspension) = (10^7 *1)/(194*2000)=2.5

$PROB_n$(bridge/ STRUCTURED, the) =(10^7 *5500)/(5651*500.000)= 1.94

$PROB_n$ (bridge/DENTAL-DEV37, the) = (10^7 *180)/(194*500.000)=1.84

$PROB_n$(bridge/ STRUCTURED, dentist) =(10^7 * 2)/(5651* 900)= 3.9

$PROB_n$ (bridge/DENTAL-DEV37, dentist) = (10^7 * 35)/(194*900)=200

Ambiguity Resolution

9.4 Statistical Word Sense Disambiguity

- To better distinguish statistics based on ratios, work in this area is often presented in terms of the log of ratio.
- For word ratios as described in this section, this measure is called the mutual information of the two words and is written as In (w_1, w_2).

$$In\ (w_1, w_2) = log\ C_n(w_1, w_2)$$

For example involving the two senses of *bridge*, the mutual information statistics are:

I_3 (bridge/STRUCTURE1, teeth) = 1.77

I_3 (bridge/DENTAL-DEVS7, teeth) = 5.14

I_3 (bridge/STRUCTURE1, the) = 0.66

Ambiguity Resolution

9.4 Statistical Word Sense Disambiguity

Note that words that have no association with each other and co-occur together according to chance will have a mutual information number close to zero, if words are anticorrelated, that is, they co-occur together at a rate less than chance, then the mutual information number will be negative.

EXERCISE OF CHAPTER 9

1) Using the Cn function described in section 9.4 compute the score of each of the senses of the word *bridge* in the five-word window *"the suspension bridge the construction"*.

2) Extend the grammar, lexicon, sense hierarchy and selectional restriction given in section 9.2 as necessary to appropriately interpret the following sentences:

   He gave the book to the college

   He knows the route to the college

3) The technique for disambiguation in section 9.4 was based only on the probability of the binary relations. How might you extend this to account for unary relations as well ?. Describe your algorithm in detail and show it would operate given the sentence:

He painted the suspension bridge at night