Representing Lexical Polysemy

Ann Copestake
Stanford University, University of Cambridge and University of Stuttgart*
aac@csli.stanford.edu

Abstract

This paper gives an overview of work on a computational approach to representing lexically governed polysemy. The framework is illustrated with an account of the interpretation of compound nouns and with some examples of extended usages of words. I discuss how generative devices can be controlled to avoid an explosion in interpretations and in the number of lexical entries.

1 Introduction

This work arose out of an attempt to build large lexical knowledge bases (LKBs), during which it became apparent that existing computational and formal approaches were inadequate with respect to the representation of polysemy. In particular, it is not possible to build a satisfactory LKB based on word sense distinctions given in a conventional dictionary, because this will not reflect the relationship between word senses nor will it be compatible with an LKB constructed on the basis of a different dictionary's sense distinctions. Furthermore an LKB must incorporate generative devices in order to cover naturally occurring data, since it is not possible to list all potential usages. The current paper is based on the framework described in Copestake and Briscoe (in press). That paper (henceforth C+B) should be consulted for an overview of the literature and for many of the details of our own approach. Here I will briefly summarise the more important points and give more details about the use of probabilities to encode semi-productivity. I will also show that the representation techniques we introduced there can be applied to the interpretation of compound nouns. Lascarides (this volume) gives a formal description of the interface with pragmatics which was assumed in C+B but not worked out there.

I am only concerned with (semi-)regular polysemy, not homonymy (e.g. *bank*:financial institution vs *bank*:mound of earth) or irregular/unpredictable polysemy (e.g. *bank*:tilt of plane vs *bank*:mound of earth). I will make a distinction between *constructional polysemy* and *sense extension*. The former includes, for example, logical metonymy (Pustejovsky, 1991), while the latter term covers cases like *rabbit* meaning the animal or its meat or fur, and *violin* meaning the instrument or the musician. This distinction will be further motivated below.

I will also make a distinction between *nonce*, *institutionalised* and *lexicalised* terms, which more or less corresponds to the terminology in Bauer (1983). As there, I will use *established* to cover both institutionalised and lexicalised. In terms of the computational lexicon, these terms have a precise definition: nonce words and usages are not recorded in the lexicon at all, but (for the cases we are interested in) can be produced/interpreted by generative devices. Institutionalised senses are recorded in the lexicon, but all that needs to be stipulated about them is the regular process by which they were formed (e.g. that *retie* is a particular sense of *tie* with the prefix *re*). Finally, lexicalised senses will either be unrelated to other senses, or will have an entry in which some regularly derived information is either augmented or overridden.

The approach to lexical representation which I am assuming involves the use of typed default feature structures. Typed feature structures should be familiar from work within HPSG. The addition of defaults is less well understood but here I am assuming the variant of order independent default unification (Lascarides et al, in press) described by Lascarides and Copestake (1995). Luckily the technicalities of the definition will not be important, the main point is that we assume that defaults are not just confined to the lexicon, but have a role during syntagmatic processing, and in the interaction with pragmatics. This is discussed in detail by Lascarides (this volume): here I want to give an description of the lexical aspects of representation of polysemy using this framework.

2 Constructional polysemy

In C+B we used the term constructional polysemy to cover cases of polysemy where a single sense assigned to a lexical entry is contextually specialised. This can be represented using (default) inheritance. For example,

*The research reported in this paper was partially funded by the ESPRIT project ACQUILEX II grant to Cambridge University. Much of it was done in collaboration with Ted Briscoe and Alex Lascarides, but they are not responsible for the mistakes in this paper.

reel was taken to have a basic underspecified representation, which stated that it was a container artifact with the purpose of (un)winding, where the material wound is left largely unspecified. In a use such as *cotton reel*, the object of the (un)winding is specialised. We represent this using Pustejovsky's theory of lexical semantics (e.g. Pustejovsky 1991). Under this account, the representation of nouns includes a specification of their *qualia structure*, which encodes the form, constituency, agentive and telic (purpose) roles. Thus the telic role of the basic sense of *reel* would be only partially instantiated, but could become further instantiated contextually. However, physical differences between types of reel are treated as outside the domain of lexical semantics.

Our use of default inheritance allows us to extend this style of representation to examples such as *cloud*, which is treated as having a basic sense with a specified formal role, but a default constitutive role (water vapour). In phrases such as *cloud of mosquitos*, the default is overridden. Some further examples of constructional polysemy are discussed by C+B and Lascarides (this volume), including adjectival interpretation (*fast car* versus *fast typist*, Pustejovsky and Boguraev, 1993) and logical metonymy (Pustejovsky, 1991). What all these cases have in common is that conventional monotonic accounts can only assign a very underspecified representation, which is assumed to be instantiated by pragmatics. Clearly, without an account of the pragmatic component and the interface between it and syntax/semantics, theories which make this claim are difficult to falsify, and few formal accounts have been attempted (although see Hobbs et al (e.g. 1990)). But for all these examples, a pragmatic explanation runs into difficulties under standard assumptions about the limitations on the type of information the pragmatic component has access to. This is discussed in the references above. The alternative approach is to assume a relatively rich lexical semantic representation, such as qualia structure, but this has to be done in a representational framework which allows defaults, in order to allow pragmatics to override the lexically instantiated defaults. Lascarides (this volume) explains how the lexical / pragmatic interface may be formalised — since she also gives the more salient details of the C+B accounts of logical metonymy and adjectival interpretation, I will not repeat them here.

3 Sense extension

In previous papers we have argued for an approach to examples such as those in (1) which treats them as extensions from a base sense.

(1) a He bought two beers.
comestible substance -> conventional portion (portioning)
b He drank a bottle of whisky.
container -> contents

Sense extensions may have syntactic/grammatical effects: e.g. portioning (1a) converts mass to count. Some sense extensions are comparable in semantic effect to morphological processes. The 'container -> contents' example in (1b) is paralleled by suffixation with *-ful*. The same polysemy patterns are often found cross-linguistically, but may be accompanied by morphological effects in one language that are not mirrored in the other. For example, in English the same word is often used for a plant and for its fruit (e.g. *olive*), but in Spanish there is a gender distinction (e.g. *aceituna/aceituno*). Sense extensions often seem to fall into families, as in the examples in (2), where meat-grinding and fur/skin grinding can be regarded as conventionalised special cases of grinding.

(2) a Sandy likes to eat rabbit.
animal -> meat (meat-grinding)
b Sandy likes to wear rabbit.
animal -> fur/skin (fur/skin grinding)
c That stuff on the tarmac looks like rabbit.
physical object -> substance (grinding)

However languages differ in what sense extensions are possible: for example, Nunberg and Zaenen (1992) report that Eskimo has no conventionalised meat-grinding. Finally, some extended senses are established while others are not, as illustrated in (3):

(3) a That restaurant serves ostrich.
animal -> meat
b The ham sandwiches has paid his check.
physical object -> associated person
c [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

The essential points of our analysis are that sense extensions are treated by lexical rules which are similar in all respects to rules which encode derivational morphology except for the change in orthography/phonology.[1] Sense extension rules can inherit from one another. For example, the meat-grinding and fur/skin grinding rules illustrated above can be regarded as subtypes of the more general grinding rule. Because we use a fine-grained lexical semantic representation, we can be quite specific about the class of lexical items to which a rule is intended to apply. We sketch the grinding and meat-grinding rules in Figure 1 (see C+B for a full description). Note that the semantic effect of grinding is captured by an operator which applies to the predicate, while the specialisation of the meaning in meat-grinding is indicated in the qualia structure.

The motivation for making the distinction between constructional polysemy and sense extension is discussed

[1] The term lexical rule has come to seem something of a misnomer, but this is rather of peripheral interest here. I will also ignore the issues which arise when formalising lexical rules in a constraint based framework: for current purposes we can treat lexical rules as being similar to non-branching phrase structure rules or schemata (cf Riehemann's (1993) treatment of derivational morphology within HPSG).

grinding < lexical-rule

$$\left[\begin{array}{l}\textbf{lex-count-noun}\\ \text{ORTH} = \boxed{0}\\ \text{SYN} = \textbf{noun-cat}\\ \text{SEM PRED} = \boxed{3}\\ \text{QUALIA} = \textbf{physical}\end{array}\right] \rightarrow \left[\begin{array}{l}\textbf{lex-uncount-noun}\\ \text{ORTH} = \boxed{0}\\ \text{SYN} = \textbf{noun-cat}\\ \text{SEM PRED} = \textbf{grinding}'(\boxed{3})\\ \text{QUALIA} = \textbf{physical}\end{array}\right]$$

meat-grinding < grinding

$$\left[\text{QUALIA} = \textbf{animal}\right] \rightarrow \left[\text{QUALIA} = \textbf{edible\_substance}\right]$$

Figure 1: Grinding and meat-grinding

extensively in C+B and we will not repeat it here, but there are several reasons why simple inheritance is not adequate for lexical rule representation. Rules can feed each other: (3c) appears to be a case of this, and in a context such as (4c), below *three rabbits* could mean three portions of rabbit meat. Furthermore in order to describe the semantic effect of lexical rules it seems necessary to postulate operators such as **grinding′**, which cannot be satisfactorily represented using simple inheritance, because there is no (substantive) common core meaning to the senses which can be specialised. Simple inheritance is known to be inadequate for derivational morphology, for similar reasons, and since we would like to use the same mechanism for both derivation and sense extension this further argues for the use of lexical rules.

3.1 Semi-productivity

Lexical rules must be able to generate nonce senses to deal with examples such as those in (3). But, at least from a computational perspective, this leads to a considerable problem since lexical polysemy then becomes intractable. It is not even clear whether the lexicon can be regarded as finite, since sense extension rules can feed one another (again this is also an issue in derivational morphology). We described an approach to restricting productivity in C+B which I want to elaborate on here. We argued that in the absence of other factors, language users utilise frequency information to resolve indeterminacies in both generation and interpretation. Such a strategy is compatible with and may underlie the Gricean Maxim of Manner, in that language will be more easily interpretable if there is a tacit agreement not to utilise abnormal or rare means of conveying particular messages. We can model this aspect of language use as a conditional probability that a word form will be used in a specific sense: Prob(lexical-entry | word-form). We assume that probabilities are associated with all established forms, regardless of whether they can be treated as derived from other senses, and we make no claim that a derived sense will necessarily be less frequent than a basic one; for example for *turkey* our intuition is that the meat sense is more frequent than the basic animal sense.

For example, a word form such as *rabbit* can be associated with an entry like that illustrated in Figure 2, in which meat grinding is shown to be (hypothetically) more probable than grinding, meat grinding followed by portioning, or fur/skin grinding. The attribute LRS associated with the lexeme for *rabbit* records which combinations of lexical rules have been attested with what frequency in the experience of the language user. If we assume that speakers choose well-attested high-frequency forms to realise particular senses and listeners choose well-attested high-frequency senses when faced with ambiguity, then much of the 'semi-productivity' of lexical rules is predicted. For instance, we would predict that in the 'null' or a neutral context[2] (4a) will be interpreted as rabbit meat, and (4b) will be interpreted as animals. Less frequent but attested senses could be chosen when dictated by the context, as in (4c).

(4) a John prefers rabbit.
  b John wants three rabbits.
  c The diners ordered three rabbits.

Probability also plays a role in the application of lexical rules in novel usage. Under the current proposal, lexical rules are somewhat like 'redundancy' rules in that they can be used to construct appropriate signs for institutionalised senses of a word form, which will have a non-zero probability in the lexeme entry. However, in the situation where an interpretation for a novel usage is called for, an assessment of the relative probability of possible lexical rules would provide a means for adopting the most likely 'analogous' interpretation. For instance, interpreting an examples such as (5), the listener who had not experienced examples of any variant of the grinding rule with *guinea pig* would choose the rule with the highest probability given the semantic type of the noun.

(5) John prefers guinea pig

The probability of a lexical rule can be derived by comparing the number of lexemes to which the rule could apply where that sense is unattested, to those for which it is attested. The following formulation is due to Ted Briscoe:

$$\text{Prob(lexical-rule)} \approx \frac{\sum_N Prob(le_0|wf_0)}{N}$$

(where N = number of lexical entries which match the lexical rule input and le_0 and wf_0 denote the output lexical-entry and output word-form, respectively).

We regard our use of probabilities as being consistent with Bauer's claim that accounting for semi-productivity is an issue of performance, not competence (Bauer 1983:71f). The effect of the use of probabilities is that unattested senses will never be assumed unless no attested sense is possible, thus allowing interpretation of examples such as those in (3), while avoiding overgeneration. This improves on the control principle suggested in Copestake (1992), that lexical rules should only be applied if no interpretation was applicable which did not involve a lexical rule, since it allows for cases such as *turkey* without further stipulation. The two other control effects suggested in Copestake (1992) are both also superseded by the current proposal. One of these was to allow for blocking, which is discussed below. The other

[2]Following Lascarides (this volume) we can give a formal definition of what we mean by neutral context — it corresponds to a situation where there are no applicable pragmatic rules which affect the interpretation of the lexical item under consideration other than those licensed by the sentence itself.

$$\begin{bmatrix} \textbf{lexeme} \\ \text{BASE} = \begin{bmatrix} \text{SIGN} = \begin{bmatrix} \textbf{lex-count-noun} \\ \text{ORTH} = \textbf{rabbit} \\ \text{CAT} = \textbf{noun-cat} \\ \text{SEM} = \textbf{obj-noun-formula} \end{bmatrix} \\ \text{PROB} = \textbf{0.4} \end{bmatrix} \\ \text{LRS} = \left\langle \begin{bmatrix} \text{RULE} = \textbf{grinding} \\ \text{PROB} = \textbf{0.05} \end{bmatrix}, \begin{bmatrix} \text{RULE} = \textbf{meat-grinding} \\ \text{PROB} = \textbf{0.3} \end{bmatrix}, \begin{bmatrix} \text{RULE} = \textbf{meat-grdg+portioning} \\ \text{PROB} = \textbf{0.15} \end{bmatrix}, \begin{bmatrix} \text{RULE} = \textbf{fur/skin-grdg} \\ \text{PROB} = \textbf{0.1} \end{bmatrix} \right\rangle \end{bmatrix}$$

Figure 2: Lexeme for *rabbit*

was that more specific lexical rules should be preferred over more general ones, e.g. that meat-grinding should be preferred over the general grinding rule. We would expect that the probabilities derived by the formula above would make meat-grinding more probable than generic grinding, but the earlier proposal would have had the undesirable consequence that a specific sense extension rule would always be applied in preference to a more general one, even if it was attested in a much smaller number of cases.

3.2 Blocking

We are mainly interested in cases where a normally productive process does not apply to a lexeme because of the existence of a synonym. This is sometimes known as *preemption by synonymy* and is exemplified by (6).

(6) a ?Sam ate pig (pork)
  b ?Sam likes cow (beef)
  c 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*)

As illustrated by (6c), and discussed in detail in C+B and Briscoe et al (in press), blocked forms do sometimes occur, but their use is highly marked and carries extra implications compared to the unblocked form. In (6c), the use of *cow* rather than *beef* has distinctly negative connotations.

Preemption by synonymy could be explained simply by assuming that speakers will use higher frequency forms to convey a given meaning. Thus an extended meaning will not become conventionalised if a common synonym exists. This means that we do not have to stipulate a separate blocking principle in interpretation, since the blocked senses will not be attested or will have a very low frequency. Blocked forms can be interpreted however, when forced by context. And in generation, we assume that higher probability forms are preferred as a way of conveying a given meaning. However, this is not the complete story, since we have not accounted formally for the extra implicatures that the use of a blocked form conveys, nor have we allowed for the generation of blocked forms (apart from in the circumstances where the generator's lexicon omits the synonym). Both these problems require an account of the interface with pragmatics, though the latter is perhaps not serious for computational applications, since we are unlikely to want to generate blocked forms. However, I will leave this open here, since it is not primarily a lexical issue.

4 Compound nouns

The approach to compound noun interpretation proposed here draws on representation techniques we have used in other cases of constructional polysemy, but combines them with some of the machinery of the previous section, to allow for the establishment of compounds. It is well known that although many noun-noun compounds seem to fall into a limited number of classes (e.g. Levi, 1978), there are cases which can only be interpreted in context. Downing attests (7) in a context where there was a table already set with a glass of orange juice by three places and apple juice by the fourth:

(7) Please sit in the apple juice chair.

Here *apple juice chair* has a meaning which can be glossed as 'chair in front of a place setting with apple juice', but obviously this meaning cannot be listed in the lexicon. Even if a compound has an established interpretation, in context there may be another possibility. In (8), taken from Bauer (1983:86), *garbage man* can be taken to mean 'a man made out of garbage' by analogy with snowman:

> 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.

This has led to the suggestion that noun-noun compounds should be assigned a representation where the relationship between the two halves of the compound is left completely unspecified and further interpretation should be left to the pragmatic component (e.g. Bauer, 1983; Hobbs et al, 1990). There are however serious objections to this application of the pragmatic dustbin. Without further elaboration it gives no explanation of the observed classes, and the fact that the majority of compound nouns do behave in a semi-regular manner. Some compounds which should be possible on pragmatic grounds do not occur: for example **blacksmith hammer*, **dentist chair* and other such compounds appear unacceptable with the interpretation that the instrument is normally used by the someone with that occupation. The possessive construction is used instead (*blacksmith's hammer*, *dentist's chair*), although in German noun-noun compounds are possible for equivalent examples. Thus there are some linguistic constraints on compound interpretation, which have to be represented.

Even with a pragmatic account, some compounds, such as *garbage man*, must be explicitly listed as established, because in the 'null-context' the established interpretation is the only possible one. So either the pragmatic component has to contain lexical information[3] or the lexicon has to contain some compounds with their established meaning. Sentences containing such compounds would be ambiguous because the corresponding productively generated underspecified compound would still have to be available. And if established compounds are listed in the lexicon, then any generalisation about the behaviour of classes of compounds should also be accessible to the lexicon, since many established compounds have an interpretation that belongs to one of the standard patterns.

These observations make noun-noun compounds a good candidate for the use of defaults which persist beyond the lexicon, where a lexically instantiated default interpretation can be overridden by subsequent pragmatic processing, along broadly similar lines to the treatment of logical metonymy in C+B and Lascarides (this volume). A general schema for endocentric compound interpretation is shown in Figure 3, with an underspecified predicate, R, relating the indices of the constituents. Most compounds will instantiate one or more of the subschemata which inherit from this schema with the predicate relating the parts of the compound marked as persistently default. An example of a more specific schema is shown in Figure 4. The slash notation indicates a default value, i.e. the schema defeasibly specifies that the compounding predicate is **made-of-substance**.

It should be apparent that under this account it does not make sense to claim that compounds have some fixed number of possible interpretations, since some predicates are supertypes of more specific predicates, and there is not necessarily any limit on specialisation of schemata. Multiple schemata may apply to a particular instance of a compound. The account of semi-productivity of lexical rules described in the previous section, can be extended to compound interpretation, however, so massive ambiguity is avoided.

In Figure 5, I tabulate some examples of compounds taken from a newspaper corpus to illustrate this approach.[4]The sample was just collected for illustrative purposes and is much too small (about 100 tokens) to draw any firm conclusions, but it is worth noticing that the only example which I found that should probably be given a completely underspecified interpretation was *Mastermind chair*.[5] Most of the other examples are established, but nevertheless fall into some general classes. Some of the established examples are lexicalised, that is they would have lexical entries which would augment or override the productive interpretation: *deckchair*, for example, is no longer particularly associated with a ship's deck.

The structure below shows the result of instantiating the schema in Figure 4 with *wickerwork chair* (ignoring the substructure in *wickerwork*).

$$\begin{bmatrix} \textbf{lex-count-noun} \\ \text{SEM} = \textbf{wickerwork}(\boxed{4}) \wedge \textbf{chair}\ \boxed{6} \wedge /\textbf{made-of-substance}(\boxed{6}, \boxed{4}) \\ \text{QUALIA} = \textbf{artefact} \end{bmatrix}$$

In normal contexts, this interpretation will stand. However, since it is defeasible, it can be pragmatically overridden along the same lines as the examples discussed by Lascarides. Thus in a context such as (9), an alternative interpretation would be found, since the default interpretation is contradicted by the context.

(9) At school, everyone worked on crafts in groups round a big table, sitting on brightly coloured chairs. To make sure everyone could reach the materials, the groups used particular chairs: the wickerwork chairs were red plastic, for example.

There are many elaborations that could be made to this sketch. As it stands *apple juice chair* would be given the default interpretation

$$\text{chair}(y) \wedge \text{apple-juice}(x) \wedge \text{made-of-substance}(y, x)$$

This would be overridden by pragmatics, but it could be lexically excluded since physical entities are subcategorised according to physical state, so the rule could be set up to exclude solid entities being made of liquids, for example. But I make no claim for the adequacy of this particular classification of compounds since my aim here is just to illustrate that the use of the persistent default mechanism makes a classification possible in principle.

5 Conclusion

In this paper, I have described an approach to representing polysemy, which appears to have considerable promise for building LKBs, since by using these techniques we can get away from the 'dictionary' notion of a fixed number of discrete word senses. Our use of default inheritance and lexical rules gives the lexicon some generative capacity with which to interpret novel usages. The approach to defaults described here has been implemented in the ACQUILEX LKB system. From a computational perspective, the main difficulty in applying

[3]Note that it is unlikely that it is the combination of denotations with the underspecified predicate that has an established interpretation, since in BrE *rubbish* is the normal term rather than *garbage*, but *rubbish man* is not established.

[4]Not all of these examples are necessarily noun-noun compounds, but for the sake of this example I have included all the possible cases.

[5]Mastermind is a long-running quiz show on BBC television, where contestants are individually questioned while sitting in a particular black chair. The context in which the compound occurred was a discussion of bad TV reception and the full sentence was:

(8) It's surprising how many people put up with "snow", blips and two people sitting in the Mastermind chair.

This usage could be regarded as treating the TV program as a location, but this considerably weakens the notion of location, so it seems preferable to regard this example as one that should be pragmatically resolved, like *apple juice chair*.

compound-noun < binary-rule

$$\begin{bmatrix} \textbf{lex-noun} \\ \text{ORTH} = \boxed{1}, \boxed{2} \\ \text{SYN} = \textbf{noun-cat} \\ \text{SEM} = \boxed{3} \wedge \boxed{5} \wedge R(\boxed{6}, \boxed{4}) \\ \text{QUALIA} = \boxed{7}\ \textbf{nomqualia} \end{bmatrix} \rightarrow \begin{bmatrix} \textbf{lex-noun} \\ \text{ORTH} = \boxed{1} \\ \text{SYN} = \textbf{noun-cat} \\ \text{SEM} = \boxed{3}\ P(\boxed{4}) \\ \text{QUALIA} = \textbf{nomqualia} \end{bmatrix}, \begin{bmatrix} \textbf{lex-noun} \\ \text{ORTH} = \boxed{2} \\ \text{SYN} = \textbf{noun-cat} \\ \text{SEM} = \boxed{5}\ Q(\boxed{6}) \\ \text{QUALIA} = \boxed{7} \end{bmatrix}$$

Figure 3: General schema for endocentric noun-noun compounds

made-of-substance-schema < compound-noun

$$\begin{bmatrix} \textbf{lex-count-noun} \\ \text{SEM} = \boxed{3} \wedge \boxed{5} \wedge /\text{made-of-substance}(\boxed{6}, \boxed{4}) \\ \text{QUALIA} = \textbf{artefact} \end{bmatrix} \rightarrow \begin{bmatrix} \textbf{lex-uncount-noun} \\ \text{SEM} = \boxed{3}\ P(\boxed{4}) \\ \text{QUALIA} = \textbf{substance} \end{bmatrix}, \begin{bmatrix} \textbf{lex-count-noun} \\ \text{SEM} = \boxed{5}\ Q(\boxed{6}) \\ \text{QUALIA} = \textbf{artefact} \end{bmatrix}$$

Figure 4: A compound noun subschema

Examples	Type of x	Interpretation
armchair, wheelchair	artefact_physical_object	$P(x) \wedge$ /has-physical-part(y, x)
deckchair, camp chair dining room chair, hospital armchair	location	$P(x) \wedge$ /used-in-loc(y, x)
folding chair, rocking chair push-chair	event	$P(e) \wedge$ /p-pat(e, y)
cane chair, rattan chair wickerwork chair, wrought-iron chair	substance	$P(x) \wedge$ /made-of-substance(y, x)
Mackintosh chair	human	$P(x) \wedge$ /designed-by(y, x)
Mastermind chair	abstract	$P(x) \wedge R(y, x)$

Figure 5: Some examples of compounds involving *chair*

the approach to semi-productivity is collecting the probabilities, since this requires disambiguation of usages in corpora. This might be feasible, despite the rather disappointing results of most reported work on sense disambiguation since probabilities do not need to be exact, and the sense distinctions we are interested in are quite coarse grained compared to those made in most learners' dictionaries. Furthermore Harley (1994) suggests that better sense disambiguation is possible, given a good lexical database. However, this remains as future work. Further work would also be required to provide a computational account of the lexical / pragmatic interface, and a tractable implementation of the pragmatic component.

References

Bauer, L. (1983) *English word-formation,* Cambridge University Press, Cambridge, England.

Briscoe, E.J., A. Copestake and A. Lascarides (in press) 'Blocking' in P. St. Dizier and E. Viegas (ed.), *Computational Lexical Semantics,* Cambridge University Press.

Copestake, A. (1992) 'The representation of lexical semantic information', Doctoral dissertation, University of Sussex, Cognitive Science Research Paper CSRP 280.

Copestake, A. and E.J. Briscoe (in press) 'Semi-Productive Polysemy and Sense Extension', *Journal of Semantics,*

Harley, A. (1994) 'Cambridge Language Survey: Semantic Tagger', ACQUILEX-II working paper No 39: available from `cide@cup.cam.ac.uk`.

Hobbs, J.R., M. Stickel, D. Appelt and P. Martin (1990) 'Interpretation as Abduction', Technical Note No. 499, Artificial Intelligence Centre, SRI International, Menlo Park, CA.

Lascarides, A., E.J. Briscoe, N. Asher and A. Copestake (in press) 'Order Independent and Persistent Typed Default Unification', *Linguistics and Philosophy,*

Lascarides, A. and Copestake, A. (1995) 'A Logic for Order Independent Typed Default Unification', ACQUILEX-II working paper.

Levi, J. (1978) *The syntax and semantics of complex nominals,* Academic Press, New York.

Nunberg, G.D. and A. Zaenen (1992) 'Systematic polysemy in lexicology and lexicography', Proceedings of Euralex92, Tampere, Finland.

Pustejovsky, J. (1991) 'The generative lexicon', *Computational Linguistics,* **17(4)**, 409–441.

Pustejovsky, J. and B. Boguraev (1993) 'Lexical knowledge representation and natural language processing', *Artificial Intelligence,* **63**, 193–223.

Riehemann, S. (1993) 'Word formation in lexical type hierarchies', M.Phil thesis, University of Tübingen, Germany.