Lexical Polymorphism and Word Disambiguation

Antonio Sanfilippo
SHARP Laboratories of Europe
Oxford Science Park
Oxford OX4 3AP, UK
antonio@sharp.co.uk

Abstract

We present an approach to lexical ambiguity where regularities about sense/usage extensibility are represented by underspecifying word entries through lexical polymorphism. Word disambiguation is carried out using contextual information gathered during language processing to ground polymorphic lexical entries.

1 Introduction

It is a well known fact about language knowledge that word usage can be easily adapted to context by manipulating syntactic and/or semantic properties of words. Consider the following examples:

- the verbs enjoy and bring exhibit distinct complementation options without consequent shift in sense usage, e.g.
  - *Mary enjoyed the book*
    Mary enjoyed reading the book
  - *Mary brought a bottle*
    Mary brought a bottle to the party
- move can express either locative or emotional change and bake conveys either change of state or creation when used with the same subcategorization frame, e.g.
  - *You'll have to move the car*
    The whole story had moved her profoundly
  - *John baked potatoes for dinner*
    She baked a cake to celebrate John's birthday
- swim describes either a process or telic eventuality and whistle expresses either sound emission or motion according to whether a directional argument is present or not, e.g.
  - *John swam for hours*
    John swam across the channel in ten minutes
  - *The wind whistled*
    The wind whistled round them

The ability to generate appropriate uses of words in context is often systematic, and therefore crucial in modelling the acquisition and representation of lexical knowledge. For example, the alternations discussed above can be generalized over discrete verb subclasses:[1]

- the NP/VP alternation found with enjoy can be extended to many psychological verbs with experiencer subject, such as *despise, fancy, hate, like, love, miss, prefer, regret*
- the directional PP alternation exemplified above by bring can be found with any verb of directed motion, e.g. *come, arrive, deliver, go, return, send*
- the locative vs. emotional change alternation found with move can be extended to most predicates of caused motion: *agitate, crash, cross, lift, strike, shake, unwind*
- the change of state vs. creation alternation observed with bake can be extended to many verbs of cooking and preparing: *boil, cook, fry, mix*
- the telic/atelic plus directional PP alternation observed above with swim can be extended to any verb of undirected motion: *drive, carry, float, push, run, walk*
- the sound-emission vs. motion and directional-PP alternation observed above with whistle can be extended to many verbs of sound emission: *buzz, hiss, rattle, thump, vroom, wheeze*

Traditional approaches to lexical representation are wont to model word usage extensibility by enumerating the distinct senses of a word along with its representative collocational properties. Word usage enumeration can be made to provide highly specialized lexical entries, but

- it fails to make explicit regularities about word usage extensibility which are necessary in promoting compactness in lexical description,
- it is at odds with our ability to create new word uses in novel contexts, and
- it generates massive lexical ambiguity.

Consequently, several attempts have been made to develop a more dynamic approach to lexical specification which provides a principled treatment of polysemy and can be used to model creative aspects of word use. For example, Pustejovsky (1991,1994) and Pustejovsky &

[1] See Levin (1993) and references therein for more details and other examples.

Boguraev (1993) propose an integrated multilayered representation of word meaning which incorporates salient aspects of world knowledge, e.g. purpose, origin, form and constituency properties are specified for nominals which refer to artifacts and natural kinds. This makes it possible to conflate different uses of the same word into a single *meta-entry* which can be extended to achieve contextual congruity using lexical rules (Copestake & Briscoe, 1992; Sanfilippo, 1994) or abductive reasoning (Hobbs *et al.*, 1993). The use of lexical rules or abductive reasoning provides a principled alternative to word sense enumeration in the treatment of polysemy and can be made to cater for novel uses of words. However, it is not clear whether these practices can address the question of lexical ambiguity efficiently as there is no known general control regime on lexical rules or abductive reasoning which would deterministically restrict polysemic expansion without preempting the generation of possible word uses.

A possible alternative to lexical rules or abductive reasoning is to use contextual information to guide word usage extension during language processing. For example Sanfilippo *et al.* (1994) propose that verb ambiguities due to multiple subcategorization options can be expressed using polymorphic lexical entries within a Typed Feature Structure formalism. A verb entry is assigned a type with subtype extensions describing all possible subcategorization options for the verb. Verb ambiguities can then be solved deterministically by using syntactic contextual information during language processing to ground underspecified verb entries. The goal of this paper is to explore extensions of this approach which make it possible to solve other kinds of lexical ambiguities.

2 Background

Following Sanfilippo *et al.* (1994), we assume an approach to lexical knowledge representation based on

- a version of HPSG (Pollard & Sag, 1994) implemented in Carpenter's ALE (Carpenter, 1992)
- Parson's event semantics (Parson, 1990) with sorted thematic roles (Sanfilippo, 1993; Sanfilippo & Poznański, 1992) and Pustejovsky's Qualia structure (Pustejovsky, 1991, 1994).

For example, the two uses of *enjoy* exemplified in (1) can be represented using the two feature structures in Figures 1-2.[2]

(1) a Mary enjoyed the book
b Mary enjoyed reading the book

[2]In representing typed feature structures, the following conventions are adopted:

- bold face fonts are used for types, and small caps for features,
- the type of a non-atomic feature structure is written inside the feature structure, at the top line,
- the semantics is linearized for ease of exposition, and
- boxed letters and integers are used to express reentrancy.

The semantic attribute of these two feature structures share the same predicate argument structure where the event described by the verb involves two participants:

- an objectual entity (obj) which is subject to intentional change (i-change), and
- a temporal entity (2 eve) which functions as the intentional source (i-source).

When *enjoy* is used as a subject equi verb (e.g. (1a)), the temporal entity which functions as the intentional source (2) is the index of the complement VP as shown in Figure 1. When *enjoy* is used as a transitive verb (e.g. (1b)), the same temporal entity corresponds to the index of the predicate in the telic role of the complement NP as shown in Figure 2.

Following Sanfilippo *et al.* (1994), we define a polymorphic type (**subjequi_OR_tv_espsych_synsem**) which subsumes the two uses of *enjoy* in (1) represented as in Figures 1-2, e.g.

```
subjequi_OR_tv_espsych_synsem sub
                    [subjequi_espsych_synsem,
                     tv_espsych_synsem].
```

To preempt spurious combinations, the constraints associated with this type indicate that the complement must be either an **np_synsem** or a **vp_synsem**, as shown in Figure 3.[3] In addition, the semantics of the verb is now only defined in the polymorphic type since both its subtypes share the same value. The lexical entry for *enjoy* is assigned this polymorphic type. Lexical type resolution consists in finding a maximal (terminal) instantiation for the polymorphic type using information available in the rule context. This is done by

- generating a list of resolving clauses from the compiled grammar which specify which terminal type can be inferred given some specific contextual information for each polymorphic type, e.g.[4]

```
solve_type(subjequi_OR_tv_espsych_synsem,
           [comp:np_synsem],
           tv_espsych_synsem).
solve_type(subjequi_OR_tv_espsych_synsem,
           [comp:vp_synsem],
           subjequi_espsych_synsem).
```

- attaching procedures to rules which carry out disambiguation through a table look-up on the list of resolving clauses using information about the head and daughter of a rule.

For example, the function of the ALE goal **solve_synsem** attached to the rule in Figure 4 is to ground the type of the head synsem (0) using the complement synsem (1) as the resolving information. This is done by

- extracting the type label of the two input feature structures,
- using the extracted type labels to fire a **solve_type** clause, and

[3]np_OR_vp_synsem is defined as a polymorhic type subsuming np_synsem and vp_synsem.

[4]See Benkerimi (in preparation) for more details.

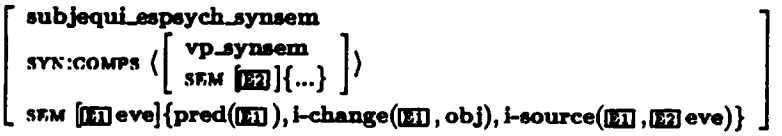


Figure 1: Lexical template for *enjoy*-type psych verbs when used as subject equi verbs.

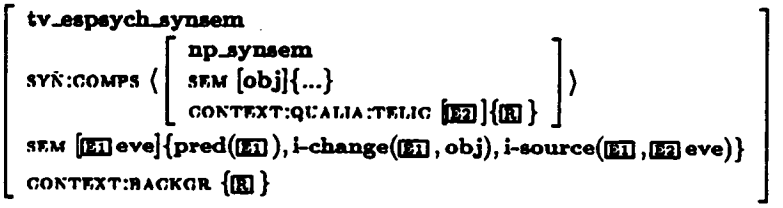


Figure 2: Lexical template for *enjoy*-type psych verbs when used as transitive verbs.

- turning the resolved type label into a feature structure which is reentrant with the input head synsem, as shown in the definition below.

```
solve_synsem(Synsem,Comp) if
        fs_to_type(Synsem,ST),
        fs_to_type(Comp,CT),
        solve_type(ST,[comp:CT],ResST),
        type_to_fs(ResST,Synsem).
```

For example, given the resolving clauses for **subjequi_OR_tv_espsych_synsem** above, the call in (2) will return as value the feature structure in Figure 1 for the resolved type **ResST**.

```
(2) solve_type(subjequi_OR_tv_espsych_synsem,
                [comp:vp_synsem],
                ResST).
```

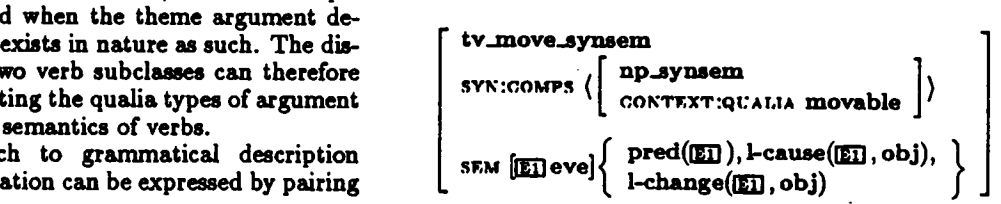
3 Extending the Treatment

To solve ambiguities due to multiple subcategorization options, the resolving information needed concerns complement type, e.g. **np_synsem**, **vp_synsem**. To solve other kinds of ambiguities by the same method we only need to establish which is the relevant resolving information. For example, the locative vs. emotional change alternation found with verbs such as *move, shake, etc.* as well as the change of state vs. creation alternation observed with verbs of cooking and preparing (e.g. *bake, boil*) can often be solved with respect to the qualia structure type of their arguments. With *move, shake, etc.*, the caused motion reading is obligatory when the direct object describes an entity which is movable and not sentient. With verbs of cooking and preparing, the creation reading occurs when the theme argument describes a composite entity which results from an artificial transformation of its constitutive elements, while the simple change of state is found when the theme argument describes an entity which exists in nature as such. The disambiguation of these two verb subclasses can therefore be implemented by relating the qualia types of argument nominals to the lexical semantics of verbs.

Within the approach to grammatical description adopted here, such a relation can be expressed by pairing qualia and thematic constraints since the lexical semantics of verbs is largely expressed in terms of thematic roles. More precisely, thematic roles are conceived of as cluster of properties which characterize aspects of eventualities. For example, following Pustejovsky (1991) and Asher & Lascarides (in press) we assume that both causation and change can be specified along the following dimensions so as to yield the thematic hierarchy in Figure 5:

- **locative** specifying (the causation of) motion, e.g. subj/obj of *put*
- **formal** specifying the creation and destruction of objects, e.g. subj/obj of *build*
- **matter** specifying (the causation of) changes in shape, size, matter and colour of an object, e.g. subj/obj of *paint*
- **intentional** specifying causation and change of the propositional attitudes of individuals, e.g. subj/obj of *amuse*

The lexical semantics of a verb such as *move* can thus be characterized using the thematic sorts **l-cause** and **l-change** (locative causation and change) for the subject and object roles when the verb expresses caused motion as in *move the car*, and **i-cause** and **i-change** (intentional causation and change) for the subject and object roles when the verb expresses caused emotion as in *her compassion moved him*. Each configuration of thematic roles is then related to a specific qualia type for the direct object, as shown below where **movable** is the qualia type for entities which are amenable to movement and **sentient** the qualia type of entities capable of acting intentionally.



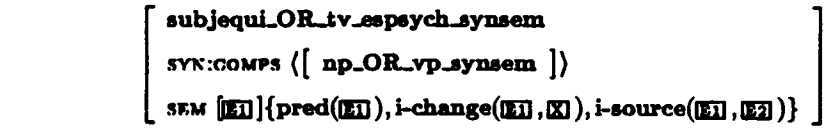
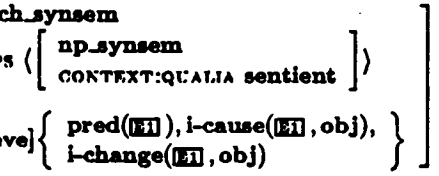
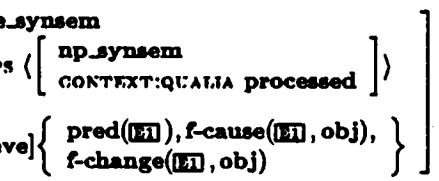
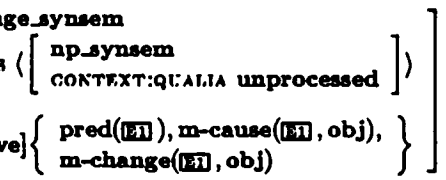


Figure 3: Polymorphic lexical type for verbs such as *enjoy*.



To characterise the lexical semantics of a verb of cooking and preparing, f-cause and f-change (formal causation and change) can be used for the subject and object roles when the verb expresses creation as in *bake a cake*, and m-cause and m-change (matter causation and change) for the subject and object roles when the verb expresses simple change of state as in *bake potatoes*. Formal causation and change are related to the qualia processed for the direct object pertaining to entities which result from an artificial transformation of some constitutive elements; matter causation and change are related to the qualia unprocessed for the direct object pertaining to entities which exists in nature as such, e.g.





As in the case of *enjoy*-type verbs, polymorphic types are declared which define ambiguous verbs such as *move* and *bake*:

(3) a `tv_move_OR_sepsych_synsem sub [tv_move_synsem, tv_sepsych_synsem].`
b `tv_create_OR_mchange_synsem sub [tv_create_synsem, tv_mchange_synsem].`

Now, resolving clauses can be derived for these two polymorphic types which perform disambiguation in terms of qualia types. For example, the resolving clauses in (4) — derivable automatically from the type definitions above — can be used for contextual disambiguation of verbs such as *bake* once the qualia type information is factored out from the rule context analogously to how complement type information is extracted from the rule context as shown in Figure 4.

(4)
```
solve_type(tv_create_OR_mchange_synsem,
           [comp_qualia:processed],
           tv_create_synsem).
solve_type(tv_create_OR_mchange_synsem,
           [comp_qualia:unprocessed],
           tv_mchange_synsem).
```

However, a question arises with respect to the disambiguation of verbs such as *move* if exactly the same procedure is adopted. Note in fact that automatic derivation of resolving clauses for the type `tv_move_OR_sepsych_synsem` would yield the clauses in (5). Assuming we are gathering qualia type information for the direct object from the rule context, the analysis of a phrase such as *move Mary* would yield two solutions since the qualia type of the direct object in this case is compatible with both movability and sentiency. This is somewhat undesirable as it is usually preferable to maintain the ambiguity rather than expanding it undeterministically.

(5)
```
solve_type(tv_move_OR_sepsych_synsem,
           [comp_qualia:movable],
           tv_move_synsem).
solve_type(tv_move_OR_sepsych_synsem,
           [comp_qualia:sentient],
           tv_sepsych_synsem).
```

An alternative is to condition the compilation of resolving clauses so that disambiguation would take place only when a deterministic choice can be made, e.g. when the direct object is non-sentient and movable:

(6)
```
solve_type(tv_move_OR_sepsych_synsem,
           [comp_qualia:inanimate_movable],
           tv_move_synsem).
solve_type(tv_move_OR_sepsych_synsem,
           [comp_qualia:movable_OR_sentient],
           tv_move_OR_sepsych_synsem).
```

The resolving clauses in (6) would force disambiguation when the theme argument expresses an entity which is movable but not sentient (`inanimate_movable`), but would leave the polymorphic type unresolved if the qualia of the theme argument is compatible with either movement, sentiency or both (`movable_OR_sentient`). In this case, another chance for disambiguation could be made available with respect to the qualia type of the subject.

4 Conclusions

The characterization of lexical knowledge within a lexicon suitable for real world NLP applications requires a mechanism which makes it possible to generate appropriate uses of words in context by manipulation of

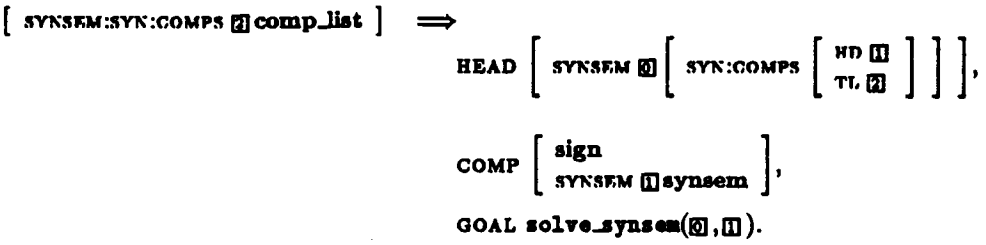


Figure 4: Aspects of the Head-Complement rule

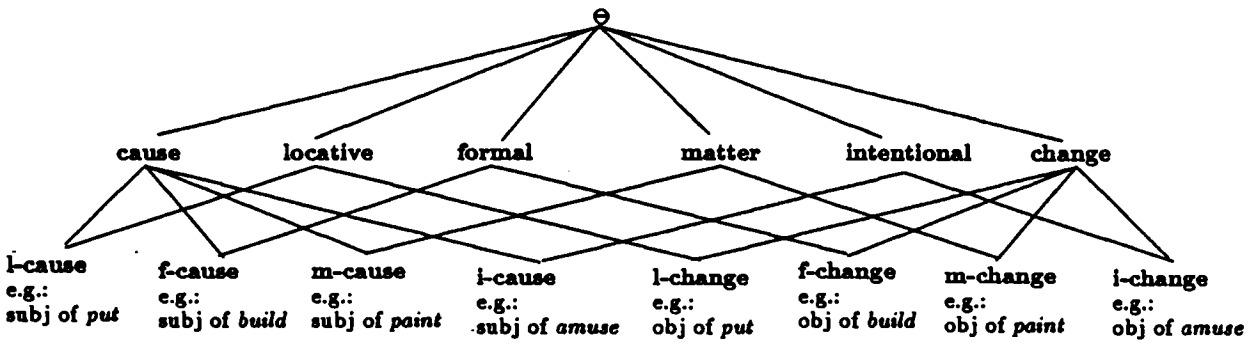


Figure 5: Top level fragment of the type hierarchy for thematic roles.

semantic and syntactic properties of lexical entries. According to our proposal, such a mechanisms is a type resolution device which uses information available in the context of grammar rules during language processing to ground polymorphic word entries. The resulting approach provides a general treatment of lexical disambiguation which capitalizes on the regularity of sense extensions to avoid undiscriminated generation of word uses during sentence processing.

References

Asher, N. and Lascarides A. (in press) Lexical Disambiguation in a Discourse Context. *Journal of Semantics.*

Carpenter, B. (1992) *ALE: The Attribute Logic Engine User's Guide.* Ms, Carnagie Mellon University.

Copestake A. & T. Briscoe (1991) Lexical Operations in a Unification-Based Framework. In J. Pustejovsky & S. Bergler (eds) *Logical Semantics and Knowledge Representation*, Springer-Verlag.

Hobbs, J., M. Stickel, D. Appelt and P. Martin (1993) Interpretation as Abduction. Artificial Intelligence 63, 69–142.

Benkerimi, K. (in preparation) Documentation Notes on the Implementation of Lexical Type Resolution Techniques. SHARP Laboratories of Europe, Oxford, UK.

Levin, B. (1993) *Towards a Lexical Organization of English Verbs.* Oxford University Press.

Pollard, C. and I. A. Sag (1994) *Head-Driven Phrase Structure Grammar.* CSLI & The University of Chicago Press.

Parsons, Terence (1990) *Events in the Semantics of English: a Study in Subatomic Semantics.* MIT press, Cambridge, Mass.

Pustejovsky, J. (1991) The Generative Lexicon. *Computational Linguistics*, 17(4).

Pustejovsky, J. (1994) Linguistic Constraints on Type Coercion. In P. St. Ditzier and E. Viegas (eds) *Computational Lexical Semantics*, CUP, in press.

Pustejovsky, J. & B. Boguraev (1993) Lexical Knowledge Representation and Natural Language Processing. *Artificial Intelligence* 63.

Sanfilippo, A. (1994) Word Knowledge Acquisition, Lexicon Construction and Dictionary Compilation. COLING-94, Kyoto, Japan.

Sanfilippo, A. (1993) LKB Encoding of Lexical Knowledge. In Briscoe, T., A. Copestake and V. de Paiva (eds.) *Default Inheritance within Unification-Based Approaches to the Lexicon.* Cambridge University Press.

Sanfilippo, A and Poznański, V. (1992) The Acquisition of Lexical Knowledge from Combined Machine-Readable Dictionary Sources. In *Proceedings of the 3rd Conference on Applied Natural Language Processing*, Trento.

Sanfilippo, A., K. Benkerimi and D. Dwehus (1994) Virtual Polysemy. COLING-94.