A Survey of Dot Objects

James Pustejovsky
Computer Science Dept., Brandeis University
Waltham, MA. jamesp@cs.brandeis.edu

December 1, 2005

1 Theoretical Assumptions

1.1 Types in Generative Lexicon

Following standard assumptions in Generative Lexicon, the computational resources available to a lexical item consist of four levels: Lexical Typing Structure; Argument Structure; Event Structure; and Qualia Structure. Qualia Structure is viewed as expressing the componential aspect of a word's meaning and the meeting point of both argument and event structure. It is generally composed of the following attributes:

(1) a. FORMAL: the basic type distinguishing the meaning of a word;
b. CONSTITUTIVE: the relation between an object and its constituent parts;
c. TELIC: the purpose or function of the object, if there is one;
d. AGENTIVE: the factors involved in the object's origins or "coming into being".

The GL model defines a language for making types, where qualia can be unified to create more complex concepts out of simple ones. Following Pustejovsky (2001), the ontology divides the domain of individuals into three levels of type structure:

(2) a. NATURAL TYPES: Natural kind concepts consisting of reference only to Formal and Const qualia roles;
b. ARTIFICIAL TYPES: Concepts making reference to purpose, function, or origin.
c. COMPLEX TYPES: Concepts integrating reference to a relation between types.

For example, a simple natural physical object (3), can be given a function (i.e., a Telic role), and transformed into an *artificial type*, as in (4).

(3) $\begin{bmatrix} \textbf{physobj(x)} \\ \text{FORMAL} = \textbf{physform(x)} \end{bmatrix}$

(4) $\begin{bmatrix} \textbf{artifact\_obj(x)} \\ \text{FORMAL} = \textbf{physform(x)} \\ \text{TELIC} = \textbf{Pred(E,y,x)} \end{bmatrix}$

Artificial types (the "unified types" in Pustejovsky, 1995) behave differently from naturals, as they carry more information regarding their use and purpose. For example, the noun *sandwich* contains information of the "eating activity" as a constraint on its *Telic* value, due to its position in the type structure; that is, **eat(P,w,x)** denotes a process, **P**, between an individual **w** and the physical object **x**. It also reflects that it is an artifact of a "making activity".

(5) $\begin{bmatrix} \textbf{sandwich(x)} \\ \text{CONST} = \textbf{\{bread,...\}} \\ \text{FORMAL} = \textbf{physform(x)} \\ \text{TELIC} = \textbf{eat(P,w,x)} \\ \text{AGENTIVE} = \textbf{make\_activity(z,x)} \end{bmatrix}$

1.2 Canonical Dot Object Behavior

The third class of types is that of *complex types*. In this note, I wish to catalogue this class according to the containing elements. Complex types, such as *book* and *university* are given a unique status in GL, implemented as *dot objects* (Pustejovsky, 1995), in order to capture their properties of orthogonal (contradictory) inheritance (Pustejovsky and Boguraev, 1993).

Dot objects have a property that I will refer to as *inherent polysemy*. This is the ability to appear in selectional contexts that are contradictory in type specification.

(6) For a lexical expression, α, where $\sigma \sqcap \tau = \bot$:
a. $[\_\_\_]_\sigma$ X
b. $[\_\_\_]_\tau$ Y
are well-formed predications.

Here are the classic data on logical polysemy motivating dot object typing. The type specification for the underlined NP is given after each example.

(7) a. Mary doesn't believe the book. `info`

b. John sold his books to Mary. `physobj`

(8) a. Eno the cat is sitting on yesterday's newspaper. `physobj`

b. Yesterday's newspaper really got me upset. `info`

(9) a. Mary is in Harvard Square looking for the Bach sonatas. `physobj`

b. We won't get to the concert until after the Bach sonata. `event`

(10) a. I have my lunch in the backpack. `food`

b. Your lunch was longer today than it was yesterday. `event`

(11) a. The phone rang during my appointment. `event`

b. My next appointment is John. `human`

The apparently contradictory nature of the two senses for each pair actually reveals a deeper structure relating these senses, the *dot object*. For each sense pair, there is a relation which connects the senses in a well-defined way. In Pustejovsky (1994), I characterize this structure as a Cartesian type product of n types, with a restricted interpretation. The set of relations, $\{R_i\}$, can be seen as specialized type product operators, where the specific relation is built into the constructor itself:

(12) $\{R_i\} = \bullet_{R_1}, \bullet_{R_2}, \ldots, \bullet_{R_n}$

In this sense, it is similar to a record type, as pointed out by Mark Johnson, Stu Shieber, and Tim Fernando (p.c.). Similar remarks have been made by Robin Cooper (Cooper, 2005). In Pustejovsky (1998) and Asher and Pustejovsky (2001, 2005), however, it is argued that dot objects cannot be adequately modeled as conjunctive types. The $\bullet$-operator appears to create a unique entity, i.e., the reified relation between two dot elements.

2 Pseudo-Dot Objects

There are many lexical items which appear to have the canonical behavior of dot objects but are, in fact, introductions or exploitations (i.e., *coercions*) from a base type.

2.1 Animal*Food

anchovy bluefin bluefish catfish chicken codfish crawfish crayfish eel flounder grouper halibut hen herring lamb littleneck lobster octopus pheasant quail rabbit sardine scallop schrod scollop scrod shellfish squid striper swordfish trout

Note: Copestake and Briscoe (1995) rules of animal grinding can adequately account for these meanings. These do not satisfy the principle of inherent polysemy.

(13) a. We own a pet rabbit.
b. We had rabbit for dinner.

2.2 Animal*Artifact-Fur

beaver chinchilla mink

Note: Similar remarks hold for other products derived from the animal.

(14) a. I saw a beaver.
b. Mary wears beaver in the winter.

2.3 Container*Containee

bottle bucket carton cask crate flask jug keg kettle pail shovel shred spoon tablespoon teacup teaspoon thimble tub

Note: The above nouns are all typed as artificial entities, carrying a telic quale, τ; $e \otimes \tau$. The value of τ is HOLD, where $\lambda y \lambda x[hold(x, y)]$. Reference to the containee argument, y, is accomplished by a *qualia exploitation* rule on the telic quale (cf. Asher and Pustejovsky, 2005, Pustejovsky, 2006).

(15) a. Bob broke the bottle accidently.
b. Mary drank the whole bottle.

3 Dot Object Types

3.1 Act*Proposition

promise allegation lie

Note: These refer to both a speech act and a propositional artifact created by this act. Cf. Asher and Lascarides (2001).

(16) a. I heard John's quick promise from yesterday.
b. John's promise took months to realize.

3.2 State*Proposition

belief

Note: See Asher (1993).

(17) a. Nothing can shake John's belief.
b. John's belief is obviously false.

3.3 Attribute*Value

temperature weight height tension strength

Note: Classic Partee example. Intension/extension distinction may be a dot object variation.

(18) a. The temperature is 90.
b. The temperature is rising.

(19) a. I know John's weight.
b. John's weight is greater than mine.

3.4 Event*Information

lecture play seminar exam quiz test

(20) a. My lecture lasted an hour.
b. Nobody understood my lecture.

3.5 Event*Human

appointment

(21) a. Your next appointment is at 3:00 pm.
b. Your next appointment is a blonde.

3.6 Event*Music

concert

Note: Music is a representational artifact, which is itself a dot object, sound*information.

(22) a. The rain started during the concert.
b. The concert was confusing.

3.7 Performance*Music

sonata symphony song

(23) a. Sophie bought some Lerner and Lowe songs.
b. Sophie coughed during the song.

3.8 Event*Physical

lunch breakfast dinner

(24) a. My lunch lasted too long today.
b. I pack my lunch on Thursdays.

3.9 Information*Physical

book cd dvd dictionary diary email mail letter

Note: The canonical dot object type.

(25) a. Mary burned my book on Mahler.
b. Mary believes all of Chomsky's books.

3.10 Material*liquid

coffee tea

(26) a. John picked the coffee from the tree.
b. John drank the coffee in the cup.

3.11 Organization*(Information*Physical)

magazine newspaper journal

(27) a. The magazine fired its editor.
b. The cup is on top of the magazine.
c. I disagreed with the magazine.

3.12 Physical*(Attribute*Value)

money currency

Note: I'm not sure how this is reflected in the lexicalization of currency words, but I am convinced it has to be a dot.

3.13 Process*Result

construction depiction imitation portrayal reference rendering decoration display documentation drawing enclosure entry instruction design invention music obstruction pattern simulation illustration agreement approval recognition damage compensation contribution disbursal disbursement discount donation acquisition deduction endowment gift categorization classification grouping

(28) a. Linnaeus's classification of the species took 25 years.
b. Linnaeus's classification contains 3000 species.

3.14 Producer*Product

[[company_name]] (e.g., Honda, IBM, Microsoft)

Note: These involve intentional agents.

(29) a. Honda raised prices last week.
b. I used to drive a Honda.

3.15 Tree*Fruit (specialization of Producer*Product)

apple orange coffee

(30) a. We planted an orange last year.
b. Mary pealed an orange for breakfast.

3.16 Tree*Wood (specialization of Producer*Product)

oak elm pine

(31) a. We trimmed our oak last fall.
b. We used oak for our cabinets.

3.17 Sound*Information(*phys)

music

(32) a. I heard the music for hours.
b. Sophie can read music fluently.

4 References

Asher, N. (1993) *Reference to Abstract Objects in Discourse*, Kluwer Academic Publishers.

Asher, N. and A. Lascarides (2001) "Indirect Speech Acts", *Synthese* 128.

Asher, N. and J. Pustejovsky (2005) "Word Meaning and Commonsense Metaphysics". ms. Brandeis and University of Texas.

Buitelaar, P. (1998) *CoreLex: Systematic Polysemy and Underspecification*, Ph.D. Dept. of Computer Science, Brandeis University.

Pustejovsky, J. (1994) "Semantic Typing and Degrees of Polymorphism", in Martin-Vide *Current Issues in Mathematical Linguistics*, Elsevier, Holland.

Pustejovsky, J. (1995). *The Generative Lexicon*, Cambridge, MA: MIT Press.

Pustejovsky, J. (2001) "Type Construction and the Logic of Concepts", in P. Bouillon and F. Busa (eds.) *The Syntax of Word Meaning*, Cambridge University Press, Cambridge.

Pustejovsky, J., and P. Boguraev. (1993). Lexical knowledge representation and natural language processing. Artificial Intelligence 63:193-223.