Interpretation as Optimization: Constitutive Material Adjectives

Partee has recently argued, on the basis of NP splits in Slavic, that there are no privative adjectives. Privative adjectives are really just subsective adjectives plus coercion [Partee, 2007]. This is a powerful and interesting proposal, but unfortunately the details of the proposed coercion process have yet to be explored in great detail. Partee suggests that Optimality Theory [Prince & Smolensky, 2004][Smolensky & Legendre, 2006] may be of some use in articulating this proposal, but as of yet no such proposal has been formally articulated. I propose an account of the interpretation of constitutive material adjectives, which is a subset of privative adjectives, using optimization.

Constitutive material adjectives include adjectives such as *stone*, *wooden*, *velveteen*, *paper*, etc. When composed with some nouns these adjectives behave intersectively, as in *stone monument*, but in other contexts, such *stone lion*, they entail that nothing in the derived extension is a member of the extension corresponding to the noun (i.e. they behave as if they are privative adjectives). Partee has suggested that adjectives of this sort are really subsective adjectives with some mechanism for coercion [Partee, 2009]. The coercion mechanism is triggered when a construction would violate the Non-Vacuity Principle.

(1) **Non-vacuity Principle** (NVP). In any given context, try to interpret any predicate so that both its positive and negative extension are non-empty. [Kamp and Partee, 1995]

When this violation occurs, the extension of the noun is coerced to include everything that is a representation of the extension of the noun (i.e. *lion* is coerced to "representation of a lion"). After coercion composition may proceed as if the adjective had been a subsective adjective all along. This process of coercion is interesting, but it has yet to be articulated in great detail. In an attempt to reflect on the nature coercion, I propose that we follow Partee's original suggestion[Partee, 2009] and proceed using optimization, specifically Stratified Domination Hierarchies [Tesar & Smolensky, 1998], as the general architecture.

The proposal employs Stratified Domination Hierarchies in which there are two strata: (1) a stratum corresponding to the adjective, and (2) a stratum corresponding to the noun. For these (lexical) strata, defeasible entailments serve as constraints. Defeasible entailments are functions, $c(x)$, that map entities in the domain of possible entities to $[0, 1]$. The specific defeasible entailments contributed by a lexical item are determined in the lexical entry. For example, a defeasible entailment of *lion* is $\text{ANIMATE}'(x)$, which maps animate entities to 1 and all others to 0.

We want to derive interpretations. I will assume that an interpretation is a set $I \subseteq S$, where S is the set of all possible entities. This interpretation is derived according to the following formula.

$$I_n = argmax_{i \in I_{n-1}} c(i)$$

The interpretation of a lexical item is derived by taking the *argmax* over a defeasible entailment and then using the supplied output as an input for the next defeasible entailment (where we always start with S as the first input). While the defeasible entailments in a given stratum are unordered, they are associated with a value, p, called it's strength. Stronger (higher p-valued) defeasible entailments are more likely to be selected during sampling. Sampling is a process that randomly selects defeasible entailments from a lexical entry on the basis of their p-values. Each defeasible entailment selected in this way decreases the likelihood that a further defeasible entailment will be selected for the given lexical item. This means that the interpretation of a given lexical item is the result of the above formula over the defeasible entailment selected by sampling.

This framework can naturally be extended to include more than one lexical item. However, we must determine how the lexical strata are ranked. Does the noun stratum outrank the adjective stratum or vice versa? Let us assume an axiom that is a notion of full interpretation and a definition for privative adjectives (including constitutive material adjectives).

(2) **Full Interpretation**. Lexical items (and their defeasible entailments) must contribute to the interpretation of a structure.

(3) **Privative Construction**. A construction, $\langle \alpha, \beta \rangle$, is a privative construction iff some $c \in \sigma(\alpha)$ conflicts with some $c' \in \sigma(\beta)$.

Full Interpretation requires that a defeasible entailment make a contribution to a structure. A defeasible entailment makes a semantic contribution to a structure if it is a monotonic, strictly decreasing function. For example, Full Interpretation would be violated if we had two conflicting defeasible entailments such as $\textsc{Animate}'(x)$ and $\neg\textsc{Animate}'(x)$. In any order that you evaluate these functions, the first one to be evaluated is the only one that has an influence on the overall result. If you take a set of entities and apply $argmax$ to the output of $\textsc{Animate}'(x)$, you will be given a set of entities that are animate. Now, if you take this set of entities and feed it to $\neg\textsc{Animate}'(x)$, and take $argmax$, you will get the same set of entities back. They are all assigned a value of 0, but 0 is still the maximum value assigned by the function.

As for the definition of a privative construction, *stone lion* (and other constitutive material adjectives) would be a privative construction because *stone* would contribute the defeasible entailment $\neg\textsc{Animate}'(x)$ which would conflict with the the defeasible entailment $\textsc{Animate}'(x)$ contributed by *lion*.

Taken together, we can conclude that regardless of which lexical stratum ranks above the other, we will have a violation of Full Interpretation, given the definition of a privative construction. To resolve this conflict we may invoke coercion, as suggested by Partee. The notion of coercion that is assumed is as follows:

(4) **Coercion (Conflict Resolution)**. Coercion resolves necessary violations of Full Interpretation by *resampling* the defeasible entailments belonging to the lower ranked lexical stratum.

If the noun stratum is ranked above the adjective stratum, then optimization guarantees that the winning candidate is in the extension of the noun and maximize over the compatible defeasible entailments supplied by the adjectives (continuously resampling without replacement when another conflict arises). If the adjective stratum is ranked above the noun stratum, then optimization guarantees that the winning candidate is in the extension of the noun and maximize over the compatible defeasible entailments supplied by the noun.

For constitutive material adjective the most accessible reading is the one in which the adjective ranks above the noun. For example, in the case of *stone lion*, with the adjective stratum ranking about the noun stratum, the winning candidate will be the possible entity that is the most lion-like entity (which is given by the optimization over (re)sampled defeasible entailments) in the extension of stone things (i.e. a statue of a lion, a stone figurine of a lion, etc.). However, another reading is predicted where the noun ranks above the adjective. This reading is typically less accessible, but may be exemplified by the metaphorical use of *paper lion* (among others). Resampling essentially turn a privative construction into a subsective construction but discarding the defeasible entailments that give rise to a semantic conflict.

The proposed architecture is quite general and can be a applied to a wide variety of adjectival classes. It has the advantage of being somewhat clearer about the possible nature of coercion. Coercion amounts to resampling the defeasible entailments of the lower ranked stratum in response to violations of Full Interpretation. In addition, it has the advantage of having the Non-Vacuity Principle "built-in". Optimization always chooses a winner, which means that there will always be something in the positive extension and Full interpretation requires that not everything be in the positive extension (in which case the involved function would not be a monotonic, strictly decreasing function). Furthermore, this proposal permits constitutive material adjectives to always have the same lexical entry/denotation regardless of the content of the noun with which they compose. The construction *stone monument* simply is not a privative construction, there are no semantic conflicts. Interpretation, in this case, can proceed according to prescribed formula for evaluation without need for resampling.

References

[Partee, 2007] Partee, B. (2007). Compositionality and coercion in semantics: The dynamics of adjective meaning. *Cognitive foundations of interpretation*, (pp. 145–161).

[Partee, 2009] Partee, B. (2009). Formal semantics, lexical semantics and compositionality: The puzzle of privative adjectives. *Philologia*, (7), 11–24.

[Prince & Smolensky, 2004] Prince, A. & Smolensky, P. (2004). Optimality theory: Constraint interaction in generative grammar.

[Smolensky & Legendre, 2006] Smolensky, P. & Legendre, G. (2006). *The harmonic mind*. MIT Press.

[Tesar & Smolensky, 1998] Tesar, B. & Smolensky, P. (1998). Learnability in optimality theory. *Linguistic Inquiry*, 29(2), 229–268.