Deriving Metonymic Coercions from WordNet

Sanda M. Harabagiu
Artificial Intelligence Center
SRI International
333 Ravenswood Ave, Menlo Park, CA 94025
harabagi@ai.sri.com

Abstract

This paper presents a method for deriving metonymic coercions from the knowledge available in WordNet. Two different classes of metonymies are inferred by using (1) lexico-semantic connections between concepts or (2) morphological cues and logical formulae defining lexical concepts. In both cases the derivation of metonymic paths is based on approximations of sortal constraints retrieved from WordNet.

This novel method of inferring coercions validates the related knowledge through coreference links. As a result, metonymic coercions are potentially useful for the recognition of coreferring entities in information extraction systems.

1 Problem description

The pervasive phenomenon of metonymy raises a problem for the interpretation of real-world texts. Metonymies are figures of speech in which, according to the literature definition from (Lakoff and Johnson, 1990), "one entity is used to refer to another, that is related to it". Characteristic of a metonymic reading of a textual expression is the fact that the satisfaction of sortal constraints guides the coercion to related knowledge.

The comprehensive account of the semantics of meaning transfers presented in (Nunberg, 1995) indicates that coercions need to be embedded in a conceptual and lexico-semantic space, ideally provided by a linguistic knowledge base. Nunberg also notes that coercions are licensed by pragmatic circumstances, specifically pertaining to the Gricean principles (Grice, 1975).

In this paper, we revisit the notion of metonymy and address the computational aspects of its resolution in the context of the relational semantics provided by the recently released WordNet 1.6 lexical database (*www.cogsci.princeton.edu/~wn*). Following the lessons learned from the WordNet-based inference of Gricean implicatures, reported in (Harabagiu et al., 1996), a novel methodology of producing metonymic paths was devised.

The coercions combine WordNet relations with semantic information derived from conceptual definitions. In WordNet (Miller, 1995) synonym words are structured in *synsets*, underlying a linguistic concept. Every synset is associated with a *gloss*, representing a textual definition, that can be translated in a *logical form* following the notation introduced in (Hobbs, 1986-1). This formalism, used in the implementation of `TACITUS` (Hobbs, 1986-2), accommodates a large variety of discourse inferences and, moreover, provides an elegant manner of localizing ambiguities, as was shown in (Bear and Hobbs, 1988).

Conceptual support from linguistic knowledge bases was already considered in the implementation of several metonymy resolution systems (e.g. (Markert and Hahn, 1997), (Fass, 1991) (Hobbs, 1986-2)), but none of these systems provided with more inferential flexibility than the typical coercion classes formulated by Lakoff (Lakoff and Johnson, 1990). We propose here a metonymy resolution approach that accounts for an open class of coercions. Similarly to Nunberg (Nunberg, 1995) and more recently to Markert and Hahn (Markert and Hahn, 1997), we find metonymy and nominal reference resolution to be two interacting processes; therefore, the proposed computational model validates metonymies through coreference links.

2 Classes of metonymic coercions

Stallard proposed in (Stallard, 1993) a distinction between two kinds of metonymy: (1) *referential metonymy*, in which the referent of a nominal predicate argument requires coercion and (2) *predicative metonymy*, featuring the coercion of the predicate usually corresponding to a verbal lexicalization. In his study, Stallard focuses on metonymic inferences required by a specific performative context, characterized by wh-questions and imperatives. His formalization of referential and predicative metonymies is based on the logical form readings of utterances from the `DARPA ATIS` (Air Travel Information Service) domain (MADCOW, 1992), a question-answering database about commercial air flights, comprising questions of the form:

`(Q1) Which wide-body jets serve dinner?`
`(Q2) Which airlines fly from Boston to Denver?`
The `ATIS` domain is characterized by a pre-established formal system of categories and relations onto which the utterances must be mapped. In this domain it is known that only flights fly or serve meals; thus, both `(Q1)` and `(Q2)` can only be understood metonymically. The interpretation of the `ATIS` utterances is performed in the logical language imposed by the implementation in the `DELPHI` system (Bobrow et al., 1991). In this framework, a question is translated into a LISP-expression:

```
(wh x S (and (P1 x) (P2 x))
```

interpreted as a query for all members of `S` (the semantic class of the wh-NP) that satisfy both `P1` (defined as the modifiers of the wh-NP) and `P2` (the predicate of the clause). For exemplification, `(QLF1)` and `(QLF2)` represent the logical form translations of `(Q1)` and `(Q2)`[1]:

```
(QLF1) (wh x flights
          (and (exists y jets
                 (and (aircraft-of x y)
                      (wide-body y))
                 (serve flight-of
                     x meal-of dinner)))
```

```
(QLF2) (wh x airlines
          (exists y flights
              (and (airline-of y x)
                   (fly flight-of y
                        orig-of Boston
                        dest-of Denver))))
```

In the case of `(QLF1)`, the coercion relation `aircraft-of` maps between flights and the aircrafts they are on, whereas in the case of `(QLF2)`, the coercion relation `airline-of` translates the connection between airlines and flights. Although both `(QLF1)` and `(QLF2)` have an interpolated quantifier for flights, which is not specified in the utterance, the difference comes from the position of the variable: in the case of `(QLF1)`, the interpolated variable `x` is the wh-variable, whereas in the case of `(QLF2)`, the interpolated variable `y` is part of the description that should be returned by the query. Stallard notes that this is the crux of the referential/predicative distinction of metonymies.

In `(QLF1)`, the NP-argument `jets` does not represent the domain of the wh-variable, but `flight` does, thus indicating a metonymic reference and deriving a referential metonymy reading. In contrast, in `(QLF2)` the domain of the wh-variable coincides with the semantic type from the utterance (i.e., `airlines`), but the *subject-argument* of the `fly` predicate is replaced by the coercion `flights`. This prompts Stallard to state that "predicative metonymy can be loosely thought of as coercion of a predicate place, rather than that of the argument NP itself".

[1] `(Q1)`, `(Q2)`, `(QLF1)` and `(QLF2)` are borrowed from (Stallard, 1993)

We argue that this definition is dependent on two factors: (1) the specific logical transformation imposed by the `DELPHI` implementation, which does not cover forms of utterances other than wh-questions, and (2) the availability of thematic role relations and coercions tailored specifically for the `ATIS` domain. This characterization of referential/predicative metonymies is not applicable when processing different genres of text, operating in different domains and, thus needing a different knowledge representation. An example of a metonymy resolution system using a more general representation is reported (Markert and Hahn, 1997).

In their model of metonymy inference, Markert and Hahn employ a two-tiered conceptual and semantic test: conceptual checks identify well-formed role chains between a pair of syntactically linked concepts, and then semantic checks distinguish whether these chains mirror literal or metonymic relationships. To perform these checks, they use (1) a concept hierarchy $\mathcal{C}$ with a taxonomic relation $isa_{\mathcal{C}}$ and (2) a set of relation names $\mathcal{R}$, containing labels of all conceptual roles of the elements from $\mathcal{C}$, hierarchically organized by $isa_{\mathcal{R}}$. For the `ATIS` domain, we can assume $\mathcal{C}$ = {`AIRLINE, JET, MEAL, PASSENGER, PILOT, ...`} and $\mathcal{R}$ = {*has-aircraft, has-flight, property-of, serves, fly-from, fly-to, ...*}.

To be able to parse texts, Markert and Hahn devised a system that grants a syntactic link between two concepts x and y if there is an acyclic path of relations $r_i \in \mathcal{R}$ and concepts $c_j \in \mathcal{C}$ such that each r_i is a conceptual role of c_{i-1}, with $range(r_i) = c_i$ and $c_0 = x \wedge (c_n \ isa_{\mathcal{C}*} \ y \vee y \ isa_{\mathcal{C}*} \ c_n)$.

Following established classifications (Lakoff and Johnson, 1990), Markert and Hahn predefine some of the relations from $\mathcal{R}$ as metonymic. These relations are {*has-part, part-of, produced-by, contained-in, made-of*}. Thus, a metonymy is recognized whenever one of the relations r_i from a path is metonymic.

In this framework, the interpretation of the relation between x =`airline` and y =`dinner` in `(Q1)` is rendered by the *Path-1*: c_0=company (with x=airline$-$*isa*$\rightarrow$company), r_1=*has-part*, c_1= employee (with flight-attendant$-$*isa*$\rightarrow$ employee and *range* (serve, `SUBJECT`) = flight-attendant), r_2=*serve*, c_3=meal (with y=dinner$-$*isa*$\rightarrow$meal). Similarly, the interpretation of the conceptual relatedness between x=`airline` and y=`Boston` in `(Q2)` is rendered by the *Path-2*: c_0=airline, r_1=*has-flight*, c_1=scheduled-flight, r_2=arrive-at, c_3=city (with y=Boston$-$*isa*$\rightarrow$city).

The presence of r_1=*has-part* in Path-1 and of r_1=*has-flight* in Path-2 indicates that the two paths correspond to metonymic readings. Relation *has-flight* is not among the predefined metonymic relations considered by Markert and Hahn, but clearly would need to be so, to classify Path-2 as metonymic. Moreover, as a distinction on the predicate coercions, Path-1 contains the relation r_2=*serve*, identical to the wh-predicate of `(Q1)`, indicating that it is not a predicative metonymy. The referential metonymy from Path-1 is determined by the metonymic relation r_1=*has-part*, coercing `airline` to `flight attendants`. In contrast, Path-2 has relation r_2=*arrive-at* that coerces the wh-predicate `fly-to` from `(Q2)`.

Markert and Hahn do not analyze the semantics of the metonymic paths, but instead distinguish referential and predicative metonymies only in the anaphoric cases. Considering that expression A is a metonymic coercion of the concept B, they assume that in the predicative case A should be available for reference resolution, whereas in the referential case, only B should be so. To be able to assess the availability for reference resolution, they search for the presence of A and B in the list of forward-looking centers of the previous sentences (thus using the functional centering framework defined in (Grosz et al., 1995)).

A similar path-finding methodology for deriving metonymies was used by the *met** system (Fass, 1991), in which connections between the sense frames of textual concepts are retrieved from a lexicon of the size of 500 word senses. These paths are then classified against a small set of predefined metonymic inference rules, and form the grounds for the figurative interpretation of textual expressions. In *met** there is no support for distinctions between predicative and referential metonymies, since coercions are possible from any concept in a text. The appeal of this implementation stems from the fact that it uses word sense frames, inspired by the structure of dictionary entries and shows that the paths retrieved from such a knowledge representation can be used to identify classical forms of metonymy. This indicates that metonymy resolution can be performed by processing knowledge from lexical dictionaries, and therefore WordNet 1.6 is a suitable candidate.

A different methodology of deriving coercions was implemented in `TACITUS` (Hobbs et al., 1993). Whenever sortal constraints are violated, explicit arguments are replaced with coercion variables, related to the explicit arguments by generic relations. The coercion is devised when the generic relation subsumes some predicate that is brought forward by the abductive interpretation of the text. In the case of `(Q1)`, argument *jets* is replaced with a coercion variable k which is expected to satisfy the subject-constraints of the verb `serve`. The abductive interpretation of `(Q1)` brings forward a reasoning path showing that variable k may be coerced to any subsumer of concept *person*. Such a subsumer is synset {`steward, flight attendant`}, having the gloss (*an attendant on an airplane*). This gloss translates the generic relation between *jets* (a hyponym of airplanes) and variable k to the predicate *on*, cued by the prepositional relation *attendant–on→airplane*. The interpretation of this prepositional relation is produced when it is matched against WordNet-based classes of prepositional attachments collected from large treebanks, following the methodology described in (Harabagiu, 1996). For this case, the *on*-prepositional relation attaches the place of work to the worker, thus giving meaning to the coercion of `jets` into `flight attendants`.

Although the coercions derived in `TACITUS` do not distinguish the predicative or referential cases, they present a different method of building metonymic paths. By incorporating this unification-based mechanism of producing coercions with a lexical path-finder working on WordNet, a novel way of deriving metonymies is made possible. It has the advantage that it relies only on approximations of sortal knowledge, as indirectly available from the WordNet database, and it does not need full-fledged abductions to be able to return metonymic paths.

3 Metonymic paths

The process of deriving metonymic paths from WordNet consists of three distinct phases: (1) the identification of sortal constraints that need to be satisfied during the interpretation of nominal expressions, (2) the retrieval of related knowledge that complies with the sortal restrictions, and (3) the validation of coercions against anaphoric expressions from the sentences following the processed sentence. The first two phases rely on access to semantic information available in (1) the relational semantic encoded in WordNet (e.g., *hypernyms*, *is_part*, *is_member*, *is_stuff*, *entail* or *pertaymym*) spanning synsets or words encoded in the database and (2) the semantic of the synset definitions (known as glosses). To be able to have computational access to the gloss semantic, synset definitions have been translated into logical formulae inspired by notation proposed in (Hobbs, 1986-1) and implemented in `TACITUS`.

Based on the davidsonian treatment of action sentences, in which events are treated as individuals, every gloss is transformed in a first-order predicate formula for which (1) verbs are mapped in predicates *verb(e,x,y)* with the convention that variable e represents the eventuality of that action or event to take place, x represents the subject of the action, and y represents its object (in the case of intransitive verbs, y is not attached to a predicate, whereas in the case of bitransitive verbs, y is assumed to range over both the direct and indirect object); (b) nouns are mapped into their lexicalized predicates and (c) modifiers have the same argument as the predicate they modify. Prepositional attachments are indicated by preposition-predicates ranging over the pair of arguments of the predicates they attach. For example, the gloss of synset {**airline, airline business, airway**} is (*a commercial enterprise that provides scheduled flights for passengers*) and has the following logical form transformation (LFT):

[enterprise(x)&commercial(x)&provide(e,x,y)& &flight(y)&scheduled(y)&for(e,p)&passenger(p)]

Characteristic of LFTs is the fact that the gloss genus is always the first predicate, rendering the LFT a formula of the form **[genus(x)&differentia(y)]**. Gloss geni are accessed repeatedly during the derivation of metonymic paths, and thus they need to be easily accessible.

The first two phases of the metonymic inference relies also upon lexico-semantic relations determined by derivational morphology, specifically the links between verbs and their nominalizations. Relations between verb and noun synsets that have elements with common morphological roots have been added to the database, classifying them as (a) the action nominalization; (b) the result (or object) of the action or (c) the agent (or subject) of the action. For example, verb propose and noun proposal refer to the same action. A *nominalization(act)* link was established between verb synset {propose, project} and the sense of proposal glossed as (*the act of making a proposal*). A second *nominalization(result)* link was established between the same verb synset and the sense of proposal glossed as (*something proposed*).

Phase I: Approximation of the sortal constraints. A nominal N is interpreted as literal or figurative depending on whether its sortal constraints to syntactically linked verb V are satisfied or not. Sortal information can be found in:

- (i) the LFT(g), where g stands for the gloss of any sense i of V (hence V_i) or any of its hypernyms;
- (ii) LFT(e), where e represents an example from g;
- (iii) LFT(c), where c represent those glosses where N and V co-occur (and the sense of V is unknown).

To access the sortal constraints of V implicitly encoded in WordNet, we collect all expressions from LFT(g) or LFT(e) such that:

(1) they contain a predicate $verb_i(e_1^i, x_1^i, x_2^i)$, representing either (a) V_i or (b) one of its hypernyms or (c) one of the geni from the LFTs of V_i or its hypernyms;

(2) they also contain any predicate that is (a) a subject, (b) an object and/or (c) a prepositional attachment to $verb_i$ in the same LFT.

When all predicates are conceptualized in the respective LFTs, such expressions have the form:

$$S_i = verb_i(e_1^i, x_1^i, x_2^i) \ \&\ subject^i(x_1^i) \ \&\ object^i(x_2^i) \ \& \\ \&\ \prod_j (prep_j^i(e_1^i, y_j^i) \ \&\ noun_j^i(y_j^i))$$

The sortal information for subjects and objects of V_i is:

$Subject_i(\mathtt{V}) = \bigcup_k subject_k^i(x_{1,k}^i)$, where $subject_k^i$ is:
(a) the subject from some S_i $(= S_i^k)$ and
(b) is the hypernym of any other subject (from another $S_i^{l \neq k}$) that belongs to the same WordNet hierarchy.

$Object_i(\mathtt{V}) = \bigcup_{k'} object_{k'}^i(x_{1,k'}^i)$, where $object_{k'}^i$ is:
(a) the object from some S_i $(= S_i^{k'})$ and
(b) is the hypernym of any other object (from another $S_i^{l' \neq k'}$) that belongs to the same WordNet hierarchy.

Similarly, for each prepositional attachment determined by a preposition prep, we define the sortal information:

$Noun_i(\mathtt{V,prep}) = \bigcup_{k''} noun_{k''}^i(x_{1,k''}^i)$ where $noun_{k''}^i$ is:
(a) attached to $verb_i$ in some S_i $(S_i^{k''})$ through **prep** and
(b) is the hypernym of any other noun attached through **prep** (in another $S_i^{l'' \neq k''}$), when both nouns belong to the same WordNet hierarchy.

Expressions S'_w are collected from the LFT(c). The semantic sense k of V in S'_w is selected as a result of the fact that the similarity measure between S'_w and the collection $\{S_i\}$ is maximal when $i = k$. The similarity measure between two expressions is defined as:

$$Similarity(S'_k, S_i) = Sim(subject, S'_k, S_i) + \\ + Sim(object, S_w, S_i) + \textstyle\sum_j Sim(prep_j, S'_k, S_i)))$$

where for $role \in \{subject, object, \{prep_j\}\}$ we have:
(i) $Sim(role, S'_k, S_i) = 1$ if the conceptualizations of $role(S'_k)$ and $role(S_i)$ belong to the same hierarchy
(ii) $Sim(role, S_w, S_i) = 0$ when either $role(S'_k)$ $role(S_i)$ are not defined or
iii) $Sim(role, S'_k, S_i) = -1$ otherwise.

Finally, considering the set operator:

$$S_1 \bigoplus_h S_2 = \{e \mid e \in S_1 \cup S_2 \text{ and there is no other } e' \in S_1 \cup S_2 \text{ such that } e' \text{ is a } hyeprnym \text{ of } e\}$$

the sortal constraint approximations are defined as:

- □ $Subject\_Sort_i(\mathtt{V}) = Subject_i(\mathtt{V}) \bigoplus_h \bigcup_q subject_q^i$, where $subject_q^i$ is the subject from some S'_w in which the sense of V is i;
- □ $Object\_Sort_i(\mathtt{V}) = Object_i(\mathtt{V}) \bigoplus_h \bigcup_{q'} object_{q'}^i$, where $object_{q'}^i$ is the object from some S'_w in which the sense of V is i;
- □ $Prep\_Sort_i(\mathtt{V,prep}) = Noun_i(\mathtt{V,prep}) \bigoplus_h \bigcup_{q''} noun_{q''}^i$, where $noun_{q''}^i$ attaches to the sense i of V through **prep** in some S'_w.

Phase II: Sorts satisfaction and expansion of coercions. Sorts satisfaction amounts to (1) the recognition of the sense of V in the text and (2) a search for any element from $Role\_Sort_i(\mathtt{V})$ across all concepts semantically related to all WordNet senses of N (given that N has the same *role* in the text).

The recognition of sense i of V is based also on the maximal value of similarity between S_s, the expression retrieved from the LFT of the text, and $\{Sort_i(\mathtt{V})\}$, where $Sort_i(\mathtt{V})$ is defined as:

$$Sort_i(\mathtt{V}) = Subject\_Sort_i(\mathtt{V}) \ \& \ Object\_Sort_i(\mathtt{V}) \ \& \\ \& \prod_j \ Prep\_Sort_i(\mathtt{V}, prep_j)$$

The satisfaction of $Role\_Sort_i(\mathtt{V})$ is a search for any element from this set along (1) all *synonyms*, (2) all *hypernyms* and (3) all `LFT` *geni* for each WordNet sense of `N`. If this search is successful, we rule that `N` had a literal reading. Otherwise, we need to build metonymic paths to be able to access the related knowledge. Two distinct ways of deriving metonymic paths have been developed.

Lexico-semantic paths. The codification of meronymic relations in WordNet determines the consideration of lexico-semantic paths composed of *isa* and at least one *is_part*, *is_member*, or *is_stuff* relations (or their reverses) as a means of deriving coercions. Implementing 22.29% of the semantic connections between noun concepts as meronyms, WordNet 1.6 sets an acceptable level of granularity for a knowledge representation needed to derive metonymic information.

Lexico-semantic metonymies retrieve concepts $C_m \in Role\_Sort_i(\mathtt{V})$ that (1) are linked through a meronymic relation to any WordNet sense of `N` (or one of its hypernyms or geni); (2) morphologically or idiomatically indicate meronymic relations to `N`, or (3) represent predicates from the `LFT` of `N` (or its hypernyms) having thematic roles for the same verb that is the genus of the `LFT`. The general form of the lexico-semantic paths is determined to be $Path_{LS}$ = (C_0=`N`,r_1,C_1,r_2,...,C_m), with at least for one i, r_i is a WordNet meronymic relation. The case when a triplet (C_{j-1},r_j,C_j) is part of an `LFT` extends the classical metonymies *object-for-agent* and *product-for-producer*.

The WordNet concepts that morphologically cue meronymic relations are those synsets containing collocations of such lexemes as **unit** (e.g., administrative unit, army unit), **system** (e.g., exhaust system, file system), **part** (e.g. body part, academic department), **group** (e.g., jazz group, pressure group), or other words that form the same hierarchies as **part** to **member**. Similarly, concepts containing in their glosses idioms like **''a group of''** or **''part of''** cue meronymic relations to the gloss genus.

Morpho-logical paths. Frequently, `N` is a nominalization of a V_N, and thus `LFT`(V_N) brings forward related semantic information. Moreover, the nominalization relations indicate the role of `N` in the `LFT`(V_N): a **nominalization(result)** link corresponds to an *object* role, whereas a **nominalization(agent)** link relates to a *subject* role of `N` in the `LFT`(V_N). This entails the interchangeability of the genus of `LFT(N)` with $verb(e,x_1,x_2)\&noun(x_i)$ (with i=1 if role=subject and i=2 if role=object) when *verb* is the predicate representing V_n and *noun* is the predicate for `N` or any of its geni (in the hierarchies of the sense of `N` morphologically related to V_N). The resulting logical form transformations are denoted as `LFT'(N)`.

In this case, we can compute the *Similarity*(Text,`LFT'(N)`) and pick the *Role(s)* for which it is maximal and incorporate the corresponding `LFTs` in the most similar `LFT'(N)`, producing the final coercion. Morpho-logical paths are sequences of three kinds of steps: (1) relatedness based on morphological relations (i.e., `N`−**nominalization**→V_N), (2) ad-hoc weighted abduction based on similarity between text roles and the logical forms of the hypernyms of `N` (i.e. `LFT'(N)`−*Similarity*(text)⊨`LFT`(*Roles*)), and (3) unification of similar logical expression (i.e., `LFT'`(V_N)−unification(`LFT`(text)→N_c.

Morpho-logic paths exploit morphological links and overlaps in the `LFTs` and resolve predicative metonymies (by bringing into play additional verbal predicates). In contrast, lexico-semantic paths resolve referential metonyies.

Phase III: *Anaphora validation.* Metonymic paths produce the expected coerced knowledge if they bring forward concepts that corefer with nominals from the successive sentences. There are three tests for the validation of coercions through anaphora. They determine whether there is a concept in a lexico-semantic path or a predicate in a morpho-logical path that (1) is identical, (2) is a hypernym, or (3) is a genus of one of the nominal expressions from the following sentences.

4 A case study

The processing associated with the derivation of metonymic paths is exemplified on a text presented in the manual defining the coreference task for the DARPA-sponsored MUC 7 competition (Hirshman and Chinchor, 1997):

(S1) The White House sent its health care proposal to the Congress yesterday.
(S2) Senator Dole said the administration's bill had little chance of passing.

The approximation of the sortal constraints of **V=send** from `(S1)` determines:

- (a) the selection of sense i=2 from the eight senses encoded for verb **send** in WordNet 1.6, due to greater similarity between its sorts and the roles from `(S1)`.
- (b) $Subject_2$(send)={{person,individual,someone}}; $Object_2$(send)={{communication},{physical object}}; $Prep\_Sort_2$(send,to)=$Prep\_Sort_2$(send,from)= ={{person,individual,someone},{address}}.

The search space for the sortal constraints is determined by the `LFT` of `(S1)`:

`LFT(S1)=[`**White_House**(x_1)&**send**(e_1,x_1,x_2)& &**proposal**(x_2)&**health_care**(x_2)&**to**(e_1,x_3)& &**Congress**(x_3)&**yesterday**(e_1)`]`

across the synonyms, hypernyms and geni of both WordNet senses of `White House`, the three senses of noun `proposal` and proper noun `Congress` respectively.

The WordNet search for the satisfaction of sortal constraint identifies `communication` as the hypernym of sense 1 and 2 of `proposal` and `person` as the genus of {`legislature,law_makers`}, the hypernym of `Congress`, thus indicating that in `(S1)`, `proposal` and `Congress` have literal meaning. Noun `proposal` as a nominalization is a candidate for morph-logic path derivations as well. The search for sortal constraints for the role of *subject* is not successful, requiring the inference of metonymic paths. A lexico-semantic path is derived, linking sense 1 of `White House` to `person`, the genus of synset `administration`:

Path1: **White_House**–*isa*→**government_department**–
–*isa*→**administrative_unit**–*morphological_cue*→
→**administration**–*genus*→**person**

The meronymic morphological cue from *Path1* is consistent with the meronyms encoded in WordNet 1.6, because we have:

▷ (a) government_department–*is_part*→administration
as an immediate inference from *Path1*
▷ (b) administration–*is_part*→government
as a semantic relation encoded in WordNet 1.6
▷ (c) government_department–*is_part*→government
as a semantic relation encoded in WordNet 1.6

Moreover, the anaphoric validation of *Path1* is possible because nominal `administration` is present in `(S2)`.

In the case of $Object_2$(send) a morpho-logical path accounts for the coerced knowledge. Nominalization `proposal` is the result of the action expressed by synset `V'`={ propose,project}. Integrating predicate *proposal* as an object in the `LFT(V')` we obtain:

`LFT'`(proposal)=present(e_2, y_1, y_2)&proposal(y_2)&
&for(e_2, y_3)&consideration(y_3)

When computing the similarity with `(S1)`, we obtain $Prep\_Sort_2$(send,to) as the role candidate to be incorporated in `LFT'`(proposal). This is enforced by the `LFT` of synset {`motion,question`}, a hyponym of sense 1 of proposal. The gloss of {`motion,question`} is (*a proposal for action made to a deliberative assembly for discussion and vote*), producing the `LFT`:

`LFT`(motion)=proposal(p)&for(p,d)&discussion(x)&
&vote(x)&to(p,a)&assembly(a)

in which present(e_2, y_1, y_2)&proposal(y_2) may substitute proposal(p)whereas `discussion` can be replaced by its hypernym `consideration`. Furthermore, since `Congress` is a hyponym of `assembly` and the filler of $Prep\_Sort_2$(send,to) in `(S1)`, we can unify `LFT`(motion), `LFT'`(proposal) and `LFT`(legislature) and obtain the final coercion as:

proposal(p)⇒present(e_2, x_1, x_2)&proposal(x_2)&
&for(e_2, x_3)&consideration(3_3)&to(e_2, x_2)&
&person(x_2)&make(e_3, x_2, l)&law(l)

Path2 is shown in its entirety in Figure 1. The anaphoric validation is shown to link `bill` from `(S2)` to `law`, since law is a genus of the first WordNet sense of bill. *Path2* brings forward two new actions (indicated by e_2 and e_3), which accounts for its classification as a predicative metonymy.

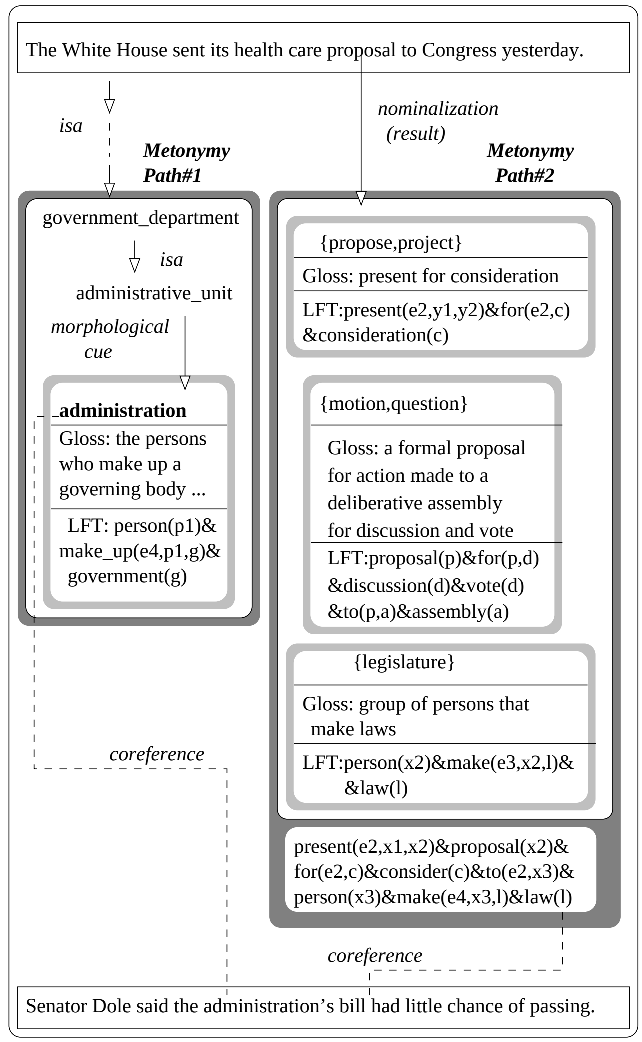

Figure 1: Example of metomymic paths

The usage of the `LFT` of a hyponym of `proposal` is a consequence of modeling Gricean principles via lexical chains from WordNet. The relevance maxim is enforced whenever as many lexical chains as possible can be retrieved between previously activated concepts and novel information. When we process the sort satisfaction of the object role, concept `legislature` is already activated and it satisfies the $Prep\_sort_2$(send,to) constraints. The same thematic role is represented in the `LFT` of `motion`, thus enlarging the number of lexical paths between `proposal` and `legislature` and increasing the relevance of `legislature` in the context of coercing knowledge for the object `proposal`. The final unification reinforces the relevance of `legislature`, congruent with the coreferential link between `bill` and `law`.

5 Conclusions

This paper proposes a method of deriving metonymic coercions by using the information available in WordNet 1.6. The method does not presuppose the availability of sortal constraints, but rather builds approximations of sortal information from the knowledge implemented in WordNet. It uses two distinct approaches for knowledge coercion: one that relies on lexico-semantic information, another based on morphological links and unifying logic formulae (`LFT`s) inferred from conceptual definitions.

To evaluate this methodology of deriving metonymic coercions, a test set of 20 *New York Times* articles were parsed by FASTUS (Appelt et al., 1993) and used in conjunction with their coreference keys, as provided by the MUC test data. There were 1261 nominal expressions, distributed in four classes as illustrated in Table 1. A percentage of 23% nominals were anaphoric, out of which almost 74% had literal meaning. In approximating the sortal constraints, sorts for 68.2% of nominals were returned. When the sorts were not satisfied by searches through WordNet, metonymic paths were derived. As Table 1 indicates, 68.9% of these paths were lexico-semantic (denoted with r=referential metonymies) and 31.1% were morpho-logical (denoted with p=predicative metonymies).

Nominal Type	Literal	Metonymy
Bare	36 (16.7%)	r=3 p=8
Definite	75 (34.8%)	r=7 p=4
Indefinite	6 (2.7%)	r=9 p=5
Proper Nouns	98 (45.5%)	r=11 p=2
TOTAL	215 (74.1%)	49 (16.8%)

Table 1: Distribution of metonymic anaphorae

The evaluation of path-validating anaphorae against coreference keys resulted in a precision rate of 76% and a recall of 83%. These results indicate that we need to experiment with different similarity measures. We also found that the metonymies have a significant contribution as knowledge sources for coreference resolution. 43% of the coreference keys were accounted by mere string matches, 3.1% by synonyms encoded in WordNet, 3.7% by hypernyms and 16.3% by links made possible through coercions.

References

Douglas E. Appelt, Jerry R. Hobbs, John Bear, David Israel, Megumi Kameyama, and Mabry Tyson. The SRI MUC-5 JV-FASTUS Information Extraction System. In *Proceedings of the Fifth Message Understanding Conference (MUC-5)*, 1993.

John Bear and Jerry R. Hobbs. Localizing Expressions of Ambiguity. In *Proceedings of the Second Conference on Applied Natural Language Processing, Association for Computational Linguistics*, pages 235–242, 1988.

Richard Bobrow, Ron Ingria and David Stallard. The Mapping Unit Approach to Subcategorization. In *Proceedings of the Speech and Natural Language Workshop*, 1991.

Dan Fass. *met** : A Method of Discriminating Metonymy and Metaphor by Computer. In *Computational Linguistics*, 17(1):49-90, 1991.

Paul J. Grice. Logic and Conversation. In P. Cole and J. Morgan, editors, *Syntax and Semantics Vol.3: Speech Acts*, pages 41–58. Academic Press, New York, 1975.

Barbara J. Grosz, Aravind K. Joshi and Scott Weinstein. A Framework for Modeling the Local Coherence of Discourse. In *Computational Linguistics*, 12:(3), 1995.

Sanda Harabagiu, Dan Moldovan and Takashi Yukawa. Testing Gricean Constraints on a WordNet-Based Coherence Evaluation System. In *Working Notes of the AAAI-96 Spring Symposium on Computational Approaches to Interpreting and Generating Conversational Implicature*, pages 31–38, 1996.

Sanda Harabagiu. An Application of WordNet to Prepositional Attachment. In *Proceedings of the 34st Annual Meeting of the Association for Computational Linguistics, ACL-96*, pages 361–362, 1996.

Lynette Hirshman and Nancy Chinchor. MUC-7 Coreference Task Definition, Version 3.0., 1997.

Jerry R. Hobbs. Discourse and Inference. Unpublished manuscript, 1986-1.

Jerry R. Hobbs. Overview of the TACITUS Project. In *Computational Linguistics*, 12:(3), 1986-2.

Jerry R. Hobbs, Mark E. Stickel, Douglas E. Appelt and Paul Martin. Interpretation as Abduction. In *Artificial Intelligence*, 63:69–142, 1993.

George Lakoff and Mark Johnson. Metaphors We Live By. Chicago University Press, Chicago, IL, 1990.

MADCOW Committee (Hirshman, Lynette et all). Multi-Site Data Collection for Spoken Language Corpus. In *Proceedings of the Speech and Natural Language Workshop*, 1992.

Katya Markert and Udo Hahn. On the Interaction of Metonymies and Anaphora. In *Proceedings of the 15th International Joint Conference on Artificial Intelligence (IJCAI-97)*, Nagoya, Japan, 1010–1015.

George A. Miller. WordNet: a Lexical Database for English. In *Communications of the ACM*, Vol.38, No.11:39–41, 1995.

Geoffrey Nunberg. Transfers of Meaning. In *Journal of Semantics*, 12:109-132, 1995.

David Stallard. Two Kinds of Metonymies. In *Proceedings of the 31st Annual Meeting of the Association for Computational Linguistics, ACL-93*, pages 87–94, 1993.