Learning Containment Metaphors

Sushobhan Nayak (snayak@iitk.ac.in)
Amitabha Mukerjee (amit@iitk.ac.in)
Indian Institute of Technology Kanpur
Kanpur, UP 208016 India

Abstract

We present a computational approach that traces the developmental process, from containment image schemas to metaphors, in four phases: a) perceptual discovery of image schemas, b) associating perceptual arguments and the relation with linguistic units, c) discovering a linguistic structure encoding the schema, and finally d) enriching the semantics of the schema via extended language usage (via a corpus). In the first three phases, we use no prior knowledge about either the perceptual or language domains; in the corpus analysis, we use the WordNet ontology. Our input is an animation based on the Heider-Simmel video, together with a small corpus of transcribed commentaries. From the image sequence, we cluster the visual angle subtended by a landmark, and find that one cluster reflects containment. This is then correlated with the sentences from the adult commentaries uttered contemporaneously with containment situations, yielding strong object-nouns and relation-preposition associations. For discovering linguistic constructs, we use no knowledge of grammatical category or syntax but find recurring patterns using the approach of (Solan, Ruppin, Horn, & Edelman, 2002). Knowing the units involved, we can identify several phrasal patterns (e.g. "X moved into", "in the Y"). We then search a corpus with the "in the Y" schema to identify container words. We find that the most common class involving containment is location (66%), followed by group membership (20%), time, and cognition (17% each). These may be thought of as language-based non-spatial enrichments for the image schema.

Keywords: containment metaphor; grounded concepts; selectional preference

Introduction

Containment metaphors arise in infancy and may help organize the adult conceptual system (Mandler, 2010). The earliest structures(*image schema*) may arise initially in perception, but are then enriched by language in several ways, and extended to various non-spatial categories. Thus, a sentence such as "I put a lot of energy *into* washing the windows." reflects the schema ACTIVITIES ARE CONTAINERS in the influential Conceptual Theory of Metaphor (CTM) (Lakoff & Johnson, 1980).

While such extensions of the initial spatial schema become conventionalized in a linguistic group, they retain the grounding. So while starting with the final text is not very useful, the grounded interpretation gives it much more flexibility. A computational study of the process would a) suggest mechanisms for understanding this process, and b) may itself be useful computationally - e.g. by providing an interpretation via simulation using the original grounding.

While much computational work has been done on metaphors, there appears no work that attempts such a vertical sweep from the initial perceptual schema to a language corpus. The emphasis within the NLP paradigm has been on identifying and analyzing metaphors. The earliest *rule-based* attempts - e.g. (Fass, 1991) were based on hand-coded knowledge and metaphors were identified as a violation of selectional restrictions in a given context (e.g. "my car drinks gasoline").

Other approaches use syntactic and co-occurrence statistics across large corpora to identify metaphors. We may call these attempts *corpus-driven*; work here may include Shutova, Sun, and Korhonen (2010) who demonstrate metaphor paraphrasing using noun-verb clustering, or Kintsch (2000) who effectively uses Latent Semantic Analysis to interpret metaphors like "My lawyer is a shark". Cormet (Mason, 2004) is able to find mappings given separated datasets for two domains, e.g. it finds LIQUID $\rightarrow$ MONEY once provided with LAB and FINANCE specific corpora to train from. Corpus-based approaches keep the metaphor mapping implicit, i.e. while the system can identify many metaphorical usages, the source domain has no grounding. Even distinguishing source from target domain is difficult, e.g. TIME co-occurs more often with SPEND than MONEY. Also, due to a primary reliance on verbs, it becomes difficult to treat ontological metaphors like CONTAINER that are more preposition dependent. Most importantly, purely linguistic approaches are hard to extend - e.g. container metaphors may invoke other attributes of the schema (e.g. 'stir excitement', or "the idea jumped out").

A third category of work, which we may call *embodied modeling* (Narayanan, 1997), is more cognitively motivated. A model is learned from a tagged training set simulating premotor cortical representation of movement (Bailey, 1997); this is then mapped to other domains to interpret metaphoric usage such as "India releases the stranglehold on business". The embodied approach is appealing and elegant, but is hard to scale up because new training data for learning the schema have to be manually created, and the syntactic structures require knowledge of the language.

In this work, we present a grounded model where initial image schemas are discovered from untagged video, and are then associated with textual commentary without using any prior language models. We focus on n-gram models, and on discovering merged paths through the sentence graphs (Solan et al., 2002). Once we have a basic construct, we can enrich the schema by exposing the learner to lots of language situations, which is simulated here by considering the 1-million word Brown corpus.

Motivation

This work combines ideas of metaphorical extension from the seminal work of Lakoff and Johnson (1980) together with the developmental ideas from Mandler (2010). Both suggest a strong role for spatiality in adult conceptual structures. Containment is discriminated by infants by the age of 2.5 months, and becomes "accessible" by 5.5 months, when it is used for multiple activities including visual and manual exploration (Spelke & Hespos, 2002). This may imply the presence of a mental structure incorporating arguments like a container and at least one trajector, and a function that given a configuration, accepts it as an instance of containment. This structure, which may be called an initial image schema, is eventually mapped to language, when containment is acquired before support (IN before ON before UNDER). This acquisition reflects an awareness not only of the preposition, but also for the linguistic argument structure that maps the image schema. But after this point, linguistic usage adapts the concept in ways that are specific to the linguistic-cultural context (Spelke & Hespos, 2002). Extensions emerge involving new structures that transfer the relationship to new domains, not only in language, but also in thought. Over increasing exposure, many of these extensions become conventionalized, many of which are listed in the CTM corpus.

To get a baseline check, we compiled 85 containment metaphors from Lakoff and Johnson (1980); of these, 65 involve prepositions in/into/out (IN a lot of trouble, INTO the century); the remaining 20 involve verbs explode, erupt, fill or adjectives full, empty - which profile other aspects of the containment schema (fullness, enclosed-ness etc).

Given this picture of the metaphor acquisition, extension, and conventionalization process, our goal is to try to model this embodied developmental process computationally, right upto the point where language affects and changes the image schema. In this process, we would like to minimize the domain knowledge available to the system; we assume only a large set of statistical learning tools, and a preference for smaller explanatory structures. Of course, we cannot model many important factors like social, interactive aspects.

The next section focuses on how an uninformed agent, with a capability for statistical learning, may acquire the containment schema as a cluster in its sensory space. The following sections discuss the discovery of linguistic units (and *n*-grams), the discovery of linguistic constructions associated with containment, and finally, the mapping to a large corpus.

Learning Containment from Perception

Linguistic concepts are cognitively characterized in terms of *image schemas*, which are schematized recurring patterns from the embodied domains of force, motion, and space(Langacker, 1987; Lakoff & Johnson, 1980). The precise structure of an image schema remains quite unclear, with different authors using differing characterizations. In this work, we take an image schema to consist of two related structures. First is the list of arguments which participate in the associated relation or activity. The other is a characterization of the situation in terms of a function defined over some feature space, so that situations satisfying this function may

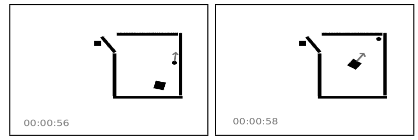


Figure 1: *Multimodal input: 2D video "Chase"*: Three shapes, big-square(**[BS]**), small-square(**[SS]**) and circle(**[C]**) interacting with each other and the static box(**[box]**).

be considered as instances of the image schema. We wish to learn such an image schema given a simple video as input (Figure 1), where three objects - a big square (**[BS]**), a small square (**[SS]**) and a circle (**[C]**) are moving around, interacting with each other and going in and out of a static **[box]** via a **door**. Though the objects deform a bit while rotating, and also occasionally overlap, it is relatively straightforward to segment them.

The linguistic database consists of a co-occurring narrative with 36 descriptions of the video. These narratives exhibit a wide range of linguistic variation in terms of linguistic focus, lexical choice and construction. In an earlier work in learning prepositions and nouns from the same multi-modal data, we used dynamic bottom-up attention to correlate objects seen in the video with their linguistic counterparts(*nouns*) (Mukerjee & Sarkar, 2007). In this work, we further consider the subset of utterances which have a temporal overlap with the frames during which a containment situation prevails.

Acquiring Containment Prepositions

In spatial reasoning, there have been several attempts at defining spatial relations involving continuum measures defined over different geometric features on object pairs. Regier (1996), a seminal work in preposition grounding, uses angle measures and a connectionist network to correlate videos and prepositions. The work, however, is limited in the sense that Regier uses videos annotated with single words like IN, OUT, THROUGH etc. while we hope to learn these schemas by clustering the untagged video. Also, because his videos are tagged with prepositions, he never has to work to discover the preposition; we have to discover these units from the unconstrained unparsed narrative. Mukerjee and Sarkar (2007) use the same dataset as ours, but use a measure based on visual proximity - the *Stolen Voronoi Area* - to cluster space using Kohonen Self Organising Maps (SOMs). We initially tried these two approaches and find that in an unsupervised clustering task (*k*-means with 6 classes), these earlier models do not work well for distinguishing the inside and outside of irregular (L- or U-shaped) containers (1st row, Fig 2). In a supervised scenario they show good results training with sophisticated neural-nets over multiple epochs, but our goal is to try not to use supervision data.

Another feature implicated in place learning in animals is *visual angle* (Rolls et al., 1999) - the angle subtended by a landmark on the retinal image. We attempted to improve on the previous features by using a single feature - - the total angle subtended by a landmark at the object position. With this measure, we find that when the resulting feature space is clustered, one of the clusters works quite well for identifying the IN-schema. Computing this feature involves computing the angle that the landmark, **[box]**, would subtend at each point in the space; the result is measured and clustered using k-Means ($k = 6$). We can see in Fig 2 (bottom row left) that one cluster completely covers what may be thought as the inside of **[box]**, whereas the the outside is graded between a number of clusters. If we accept this as a characterization for an image schema for containment, then the distribution of visual angle in this cluster will serve to represent this relation. To test whether this model, learned from the single **[box]** shape, really represents the *category* of containment relations, we generalize and evaluate it over a number of other shapes. The results of applying the same learned distribution to three novel shapes is shown in Fig 2(bottom row). We find that regions with varied levels of 'IN-ness' have been separately grouped, validating our choice of features. While for closed convex shapes the measure has a clear demarcation of 'inside'(360° angle), it gives a more graded assessment for open figures as well, such as the open-top square in Fig 2(2nd row, 4th fig).

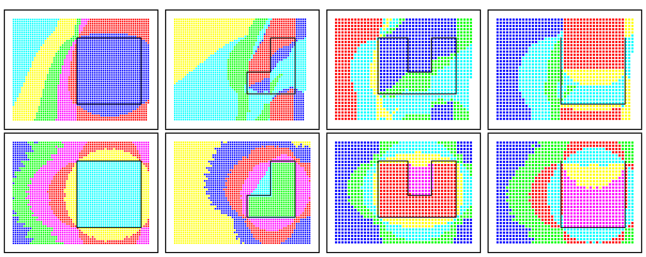


Figure 2: *k-Means Clustering of Space*. a) Voronoi and Angle features (top row) and b) Visual Angle feature (bottom row). The inside of all the containers has been clearly identified as a separate cluster only in the latter case.

At this stage, the system computes the visual angle subtended at an object position, for a landmark (the box or some other shape). The two arguments participating in this computation are the container and the trajector, though the system does not use these terms; these relationships are implicit in the feature computation. If the visual angle falls within the distribution associated with containment (the *IN-cluster*), it is accepted as an instance of this relation. Thus the system has both the arguments, and the acceptance function characterizing the image schema, This acquisition is pre-linguistic, from perceptual data alone. When a pre-discovered object, say **[BS]**, lies in IN-cluster, we have the argument structure {**[BS]** IN **[box]**} or IN(**[BS]**,**[box]**). After learning the unit IN, we can attempt to map this perceptual argument structure to linguistic syntactical structure, thereby discovering aspects of syntax.

Linguistic Elements Describing Containment

We now detail the discovery of linguistic elements pertaining to containment through correlation of the commentary with the acquired schemata. We restrict our analysis to utterances occurring while **[BS]** is in the *IN-cluster* w.r.t. **[box]**. Since the commentaries are unconstrained and there is no syntactic information, every uttered word is a possible label. Given an uttered unit w_i, the probability that it refers to schema s_j, is given by:

$$P(s_j|w_i) = \frac{P(w_i|s_j)P(s_j)}{P(w_i)} \propto \frac{P(w_i|s_j)}{P(w_i)} = A^r_{ij}$$

This metric, the *relative association*(A^r_{ij}), is prone to give erroneous results for infrequent units, while working well for high frequency words. For example, it gives an association value of 1 for a word that has been uttered only once in the whole commentary. To counter this trend, we also subscribe to mutual information between states s_j and words w_i, which eliminates the possibility of uninformative rare words being assigned a high score. The word-object association is then estimated using the product of mutual information of word w_i and state s_j with their joint probability,

$$A^m_{ij} = Pr(w_i, s_j) \log \frac{Pr(w_i, s_j)}{Pr(w_i)Pr(s_j)}$$

where A^m_{ij} is the *mutual association*. We use this measure because if $W(= \cup_i w_i)$ and $S(= \cup_i s_i)$ are two random variables then their Mutual Information $I(W,S)$ would be

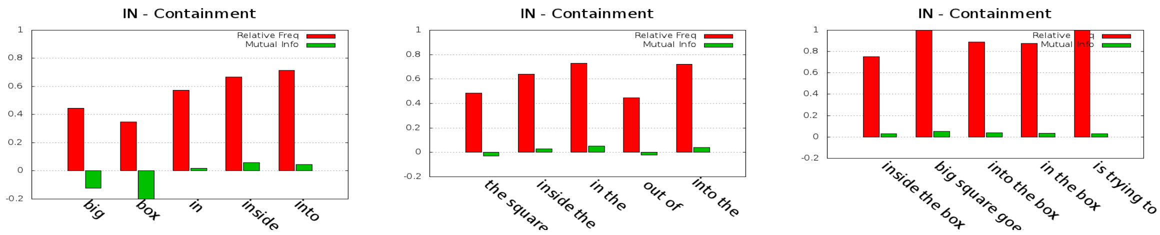
$$I(W,S) = \sum_i \sum_j Pr(w_i, s_j) \log \frac{Pr(w_i, s_j)}{Pr(w_i)Pr(s_j)} = \sum_i \sum_j A^m_{ij}$$

A^m_{ij} is, thus, the contribution of each word object pair. The results are shown in Fig 3. Notice that `in`, `inside` and `into` emerge as the three dominant monograms (their frequencies in the containment subset are 28, 26 and 15 of 1100 words).

Before moving onto syntax discovery, we observe that the nouns corresponding to the three objects in the video, had been acquired earlier using a attentional correlation model (Mukerjee & Sarkar, 2007). These will be used in the next section: "big square" for **[BS]**, "little square" for **[SS]**, and "circle" for **[C]**.

Deriving Syntactical Structure

Discovering syntactic structures of a containment preposition like IN will enable us to discover other labels for objects participating in containment. For example, if we know that only object A with label l_A is in container B with label l_B, the *perceptual* arguments of IN are {trajector:A, container:B}. Now, suppose at the same time we hear the utterance: 'l_A goes into l_C'. If we know the linguistic argument structure associated with IN, then we can have a high confidence from this single instance that l_C is most likely another label for the container B. Once internalized, this process would help the agent recreate context in a novel discourse, by simulating the action and



(a) Monogram association

(b) Bigram association

(c) Trigram association

Figure 3: *Association of words with the containment cluster*. Both the association measures, as used previously for noun-label learning, are shown. The **red** bars indicate the *Relative Frequency measure* while the **green** bars are for the *Mutual Information Measure*.

identifying probable trajectors and landmarks via the syntactic argument structure. In the example above, if we don't know what 'goes' refers to, some idea of this may be formed by realizing that it is something that *A* and *B* may be participating in. This goes along with Siskind (1996)'s ***constraining hypotheses with partial knowledge***: " When learning word meanings, children use partial knowledge of word meanings to constrain hypotheses about the meanings of utterances that contain those words."

We start our discovery of syntactic structure by analyzing bi- and tri-gram correlations, which are associated with the same metric as in the monogram case. We observe that the prominent bi-grams are `inside the`, `into the` and `in the`. The tri-grams that emerge are `inside the box`, `in the box` and `into the box`. These associations could help learn the label of not only the containment schema, but also the container itself. Note that the attention based model for learning nouns cannot learn the container/box, which is never dynamically salient. Thus, it's label is not known. However it is prominent in these containment sentences, and discounting the frequent word `the` in trigrams such as "{ `inside | in | into`} `the box`", we may associate `box` with the **[BOX]**, treating it as a label for the container. We also note the presence of other fragments (`big square goes` and `is trying to`) as prominent trigrams; but these are present only in this context, and will be diluted as the agent looks at other containment situations. But despite some glimmers, the n-gram approach is not very illuminating regarding the construction encoding for containment.

A richer model of syntactic structure has been developed over many years by Edelman and his group, implemented as the tool ADIOS (Solan et al., 2002). In this approach, a graph called Representational Data Structure (RDS) is constructed from the morphologically segmented input sentences. It then repeatedly scans and modifies the RDS in an attempt to merge edges and come up with a more compact description of the input. In the process, a number of syntactic clusters and constructions are discovered without any prior knowledge of grammatical categories. Examples of some of the patterns found are:

$$\begin{bmatrix} the \begin{bmatrix} ball \\ door \\ box \\ square \end{bmatrix} \\ \\ [circle] \end{bmatrix} \rightarrow \begin{bmatrix} move \\ came \\ got \end{bmatrix} \rightarrow into$$

$$\begin{bmatrix} in \\ inside \\ into \end{bmatrix} \rightarrow the \rightarrow box$$

Clearly, these structures are discovering some relevant patterns for containment, including the group `in | into | inside` which was also observed in the trigram model. The noun cluster - with `box` and `door` is also an impressive discovery. But most effective is the fact that `box` is identified as a label appearing after the IN while the trajector appears before.

One of the uses of this structure would be the discovery of synonyms. For example, consider the sentence `large square moves into the box`, which is being uttered as the agent perceptually knows that **[BS]** is moving into IN-cluster, activating schema IN(**[BS]**,**[box]**). By aligning this with the linguistic input, `large square` is learned as a referent for **[BS]**, i.e. a synonym for "big square". Also, an unit appearing frequently in the trajector position in the discourse is "it", which has no constant referent; it is possible now to discover that this is a unit which may refer to multiple objects. Further analysis may reveal that "it" is strongly correlated to the most recently observed trajector - and thus our system may begin its journey of understanding anaphora, another computationally promising domain in language.

Metaphorical Mappings from Language Usage

We have alluded to learning metaphorical maps through association before. We are in a position where we have grounded concepts of agents taking part in an event, and the concept of containment, through a mapping for linguistic element IN. Consider the following occurrences:

1. What state is the project IN? `IN-STATE (project, state)`
2. He did it IN three minutes. `IN-TIME (he, 3min)`

Here, the abstract concepts of STATE, TIME etc. are understood in terms of a container, thanks to their syntactic association with linguistic instances of "IN a BOX (container)", for which the learner has a physical basis. It's therefore prudent to assume that metaphorical concepts would occur in similar lexico-syntactic environments in language usage. We have demonstrated the ability of the agent in discovering the 'container' through ADIOS; the question we would like to address now is, would it also be able to discover the object of containment in a novel context with sentences of myriads of different structure? We investigated it by running ADIOS on IN-containing sentences of Brown corpus. We find that it can, indeed, distinguish 'containers' from other elements of a sentence, as evidenced by the following pattern:

$$in \rightarrow the \rightarrow \begin{bmatrix} building & war & fight & car \\ group & death & woods & cellar \end{bmatrix}$$

While discovery of 'container' elements would be the first evidence of mapping of abstract concepts to the 'container', the mappings would be prominent only if further evidence abounds in language, so that learning due to false evidence is minimized. The *propensity* of a concept to be described as a 'container' can be gauged through the regularity of its occurrence in the object position of IN. In literature, **selectional preference (SP)**(Resnik, 1993) is used abundantly to measure regularities of a verb w.r.t. the semantic class (*subject, object* etc.) of its argument. It has been used previously for word-sense disambiguation(Resnik, 1993) and metaphor interpretation(Mason, 2004). While it has only been used for finding verb-preferences, we will adapt it to include prepositional preferences, so that we are able to learn containment metaphors. While verbs have different syntactic relations like *verb-object* or *subject-verb*, the prepositions we are considering, have only one relation to the trailing noun, that of Object of Preposition (*pobj*)(according to the Stanford Parser(Marneffe & Manning, 2006)). Therefore, the formulation from Resnik (1993) is slightly modified. The *selectional association* of class c for predicate p(IN) is defined as:

$$A(p,c) = \frac{1}{S(p)} P(c|p) \log \frac{P(c|p)}{P(c)}$$

where,

$$S(p) = D(P(c|p)||P(c)) = \sum_c P(c|p) \log \frac{P(c|p)}{P(c)}$$

and,

$$P(c|p) = \frac{freq(p,c)}{freq(p)} = \frac{\sum_{w \epsilon c} \frac{count(p,w)}{classes(w)}}{freq(p)}$$

where $count(p,w)$ is the number of time word w occurred, and $classes(w)$ is the number of classes it belongs to.

Table 1: Selectional association strength of different classes

Class	SA	Class	SA
location	0.658	act	0.058
group	0.201	artifact	0.077
time	0.175	object	0.055
cognition	0.164	food	-0.030
state	0.145	animal	-0.042

WordNet (Feinerer & Hornik, 2011) is our knowledge-base for class c. WordNet was developed as a system that would be consistent with the knowledge acquired over the years about how human beings process language. To represent the early learner's limited concept-repository, only top level classes of WordNet are considered. We use the Brown Corpus as the sample space to determine the selectional preferences. All the sentences involving the containment concepts, i.e. all 21,480 sentence-parts with words `in/into/inside` were extracted. The sentences with IN were converted to the functional form of IN(*pobj*) in a rather simple way: the first occurrence of a noun after IN in the tagged corpus was assigned to the concept. For example, the sentence fragment `into a hot cauldron` is converted to IN(cauldron).

The most occurring 'container' words were `world, way, order, years, case, states` etc. The resultant associations are shown in Table 1. Notice that Location class has the highest association for container schema, activating a LOCATIONS ARE CONTAINERS mapping. Group class also has a strong association to containers, representative of the notion that groups or teams are visualized as containers (`in a group, in a team`). Time, Cognition and State also show high associativity, while Food and Person class demonstrate a significantly negative mapping. The negative numbers only represent the weakness of mapping, and should not be treated as repudiating existence of the same. The association measures only demonstrate that some mappings are used more in language, and consequently, are stronger in our cognition than others. In fact, in the original metaphor list(Lakoff & Johnson, 1980), the most prominent mappings to container are those of Cognition(15%), State(14%), Location(7.3%), Group(8.6%), Time(5.4%) and Act(4.8%), somewhat representative of their strength acquired from the whole corpus. Similarly, the least occurring classes in the list too are Plant(0.3%), Animal(0.3%) and Food(1%). Some examples of sentences from both the Brown corpus and metaphor list are presented below:

- STATE/COGNITION AS A CONTAINER
  - Meredith began falling in love. (Brown)
  - We're IN a mess. (Lakoff & Johnson, 1980)
- TIME AS A CONTAINER
  - We're well into the century. (Lakoff & Johnson, 1980)
  - There comes a time in the lives of most of us when we want to be alone. (Brown)
- LOCATION AS A CONTAINER

– If you've travelled in Europe a time or two , ... (Brown)
– ...and begin to feel at home in the capitals of Europe...(Brown)

- ACT AS A CONTAINER
  – How did you get into window-washing as a profession? (Lakoff & Johnson, 1980)

Conclusion

In this work, we proposed a plausible method for a primitive artificial agent with multi-modal input handling and feature extraction capability, to internalize linguistic concepts of containment. Containment metaphors are a primary way in which humans understand abstract concepts of state/emotion etc. and it's therefore necessary for a cognitive system to be able to do so if it's to acquire human level intelligence. Through a grounded system and selectional preference of IN, we also showed how the model can internalize conventional metaphors in vogue in English.

In the work, we have used some novel ideas, like that of modifying selectional preference to include prepositional bindings, which is, to our knowledge, unexplored in literature. Also, through simple methods of spatial feature extraction, unsupervised clustering and word-cluster association, we have been able to extract the idea of containment.

Nonetheless, this work can benefit from much improvisation. The image schema for containment that we provide, is a very crude one. While it's possible to extend it to more general scenarios, as has been demonstrated, we haven't investigated its limitations on non-convex shapes. That, however, hardly undermines our contribution here since in non-convex shapes, even adults find it difficult to separate regions into inside/outside, and it largely becomes an outcome of some context. Furthermore, our aim was to show that even with so meagre a sense-world (that of one video and associated commentaries), an artificial agent can get some semblance of human-like notions of containment and employ them to inculcate metaphorical mappings in its system. Instead of 81 seconds of learning, however, the human learner has days and months and years of exposure, and clearly this can lead to the construction of extremely rich and diverse schemata.

While we have handled concepts that easily fall into a container mould due to usage of IN, we have not been able to model other container metaphors that have different attributes, like the 20 out of our list of 85 that depend on verbs and adjectives related to substance. Simulating all that is indeed a Herculean task in this limited set-up. We intend to look into these aspects and also, to matters of finding more mappings using WordNet's synsets (at present we are using only the lexical files – synset usage would lead to discovery of more maps, but it would also make the process noisy) in our future work.

References

Bailey, D. (1997). *A computational model of embodiment in the acquisition of action verbs.*

Chalnick, A., & Billman, D. (1988). Unsupervised learning of correlational structure. In *Proc. 10th Annual Cogsci Conf.* (pp. 510–516).

Fass, D. (1991). Met*: A method for discriminating metonymy and metaphor by computer. *Computational Linguistics*, *17*(1), 49-90.

Feinerer, I., & Hornik, K. (2011). wordnet: Wordnet interface [Computer software manual]. Available from `http://CRAN.R-project.org/package=wordnet` (R package version 0.1-8.)

Hill, J. A. C. (1983). A computational model of language acquisition in the two-year old. *Cognition and Brain Theory*, *6*, 287–317.

Kintsch, W. (2000, July). Metaphor comprehension: A computational theory. *Psychonomic Bulletin and Review*, 257-266.

Lakoff, G., & Johnson, M. (Eds.). (1980). *Metaphors we live by*. University of Chicago Press.

Langacker, R. (Ed.). (1987). *Foundations of cognitive grammar I: Theoretical prerequisites*. Stanford University Press.

Mandler, J. M. (2010). The spatial foundations of the conceptual system. *Language and Cognition*, *2*(1), 21-44.

Marneffe, B. M. Marie-Catherine de, & Manning, C. D. (2006). Generating typed dependency parses from phrase structure parses. In *Proceedings of LREC 2006.*

Mason, Z. J. (2004, March). Cormet: a computational, corpus-based conventional metaphor extraction system. *Comput. Linguist.*, *30*, 23–44.

Mukerjee, A., & Sarkar, M. (2007). Grounded acquisition of containment prepositions. In *Proc. ICON.*

Narayanan, S. (1997). *Knowledge-based action representations for metaphor and aspect (karma)*. Unpublished doctoral dissertation, CS, UC Berkeley.

Regier, T. (1996). *The human semantic potential: Spatial language and constrained connectionism*. Bradford Books.

Resnik, P. S. (1993). *Selection and information: A class-based approach to lexical relationships* (Tech. Rep.).

Rolls, E., et al. (1999). Spatial view cells and the representation of place in the primate hippocampus. *Hippocampus*, *9*(4), 467–480.

Shutova, E., Sun, L., & Korhonen, A. (2010). Metaphor identification using verb and noun clustering. In *Proc. of COLING* (pp. 1002–1010).

Siskind, J. (1996). A computational study of cross-situational techniques for learning word-to-meaning mappings. *Cognition*, *61*, 39-91.

Solan, Z., Ruppin, E., Horn, D., & Edelman, S. (2002). Automatic acquisition and efficient representation of syntactic structures. In *Proc. of NIPS.*

Spelke, E., & Hespos, S. (2002). Conceptual development in infancy: The case of containment. *Representation, memory, and development: Essays in honor of Jean Mandler*, 223–246.