Pathways to Creativity in Lexical Ontologies

Tony Veale

Department of Computer Science,
University College Dublin, Belfield, Dublin 6, Ireland.
Email: Tony.veale@UCD.ie
WWW: http://www.cs.ucd.ie/staff/tveale/home/

Abstract. Language is a highly creative medium, and lexicalized ontologies like WordNet are rich in implicit evidence of the conceptual innovations underlying lexical inventiveness. We argue that WordNet's overt linguistic influences make it far more conducive to the development of creative thinking systems than other, more formalized conceptual ontologies like Cyc.

1 Introduction

Creativity is a vexing phenomenon to pin down formally [1], which is perhaps why we tend to think of it in largely metaphoric terms. For example, creativity is often conceived as a form of mental agility that allows gifted individuals to make astonishing mental leaps from one concept to another [2]. Alternately, it is popularly conceived as a form of lateral thinking that allows those who use it to insightfully cut sideways through the hierarchical rigidity of conventional categories [3]. Common to most of these metaphors is the idea that creativity involves recategorization, the ability to meaningfully move a concept from one category to another in a way that unlocks hidden value, perhaps by revealing a new and useful functional property of the concept. For example, psychometric tests such as the Torrance test of creative thinking [4] try to measure this ability with tasks that, e.g., ask a subject to list as many unusual and interesting uses of old tin cans as possible.

The ad-hoc nature of creativity is such that most ontologies, perhaps all ontologies, do not and can not provide the kinds of lateral linkages between concepts to allow this kind of inventive recategorization. Instead, ontologies tend to concentrate their representational energies on the hierarchical structures that, from the lateral thinking perspective, are as much a hindrance as an inducement to creativity. This is certainly true of WordNet [5], whose *isa* hierarchy is the most richly developed part of its lexical ontology, but it is also true of language-independent ontologies like Cyc [6], which are rich in non-hierarchical relations but not of the kind that capture deep similarity between superficially different concepts. It is connections like these that most readily fuel the recategorization process.

However, because WordNet is an ontology of lexicalized concepts, it necessarily captures much of the lexical creativity evident in everyday language. Often, this word-use is a reflection of deeper recategorization processes at the conceptual level. We argue that if we can identify and extract this evidence using automatic or semi-automatic means, we then have a basis for augmenting WordNet with the lateral connections from which novel creative pathways can be constructed.

Petr Sojka, Karel Pala, Pavel Smrž, Christiane Fellbaum, Piek Vossen (Eds.): GWC 2004, Proceedings, pp. 220–225.

2 Polysemy versus Homonymy

Polysemy is a form of lexical ambiguity in which a word has multiple related meanings. The form of polysemy that interests us most from a creativity perspective is function-transforming polysemy, which reflects at the lexical level the way concepts can be extended to fulfill new purposes. For instance, English has a variety of words that denote both animals and the meat derived from them (e.g., *chicken, lamb, cod*), and this polysemy reflects the transformation potential of animals to be used as meat.

If we can identify all such instances of function-transforming polysemy in WordNet, we can generalize from these a collection of pathways that allow a system to hypothesize creative uses for other concepts that are not so entrenched via polysemy. For example, WordNet defines several senses of *knife*, one as an *{edge-tool}* used for cutting and one as a *{weapon}* used for injuring. Each sense describes structurally similar objects (sharp flat objects with handles) with a common behavior (cutting) that differ primarily in function (i.e., slicing vs. stabbing). This polysemy suggests a generalization that captures the functional potential of any other *{edge-tool}*, such as *{scissors}* and *{shears}*, to also be used as a *{weapon}*. More formally, for every polysemous sense pairing $<\omega_1, \omega_2>$ with immediate hypernyms $<h_1, h_2>$, we can create a category subsumption entailment $h_1(x) \rightarrow h2(x)$ if h_2 is a broader category that h_1, which is to say, if h_2 has more descendent hyponyms than h_1. Since *{weapon}* is a broader category than *{edge-tool}*, we can infer that other edge-tools may be used as weapons too, but conversely, we do not infer that all weapons are potential edge-tools. In effect, the generalization represents an inductive hypothesis that it is the sharp edge in a tool that allows it to be used as a weapon.

3 Identifying Creativity-Supporting Polysemy in WordNet

It is crucial that our generalization process be able to distinguish polysemy from homonymy – another form of ambiguity in which the multiple senses of a word are not related – since WordNet's synset representation does not explicitly mark either phenomenon.

True polysemous relationships can be recognized using a variety of automatic approaches. In the top down approach, cousin relations [5,7] are manually established between concepts in the upper-ontology to explain the systematicity of polysemy at lower levels. For instance, once a connection between *{animal}* and *{food}* is established, it can be instantiated by words with both an animal and food sense. However, this approach is limited by the number of high-level connections that are manually added, and by the need to list often copious exceptions to the pattern (e.g., *mate* the animal partner, and *mate* the berry drink, are merely homonyms; the latter is not derived from the former). Conversely, in the bottom-up approach, systematic patterns are first recognized in the lower ontology and then generalized to establish higher-level connections [8,9,10]. For instance, several words have senses that denote both a kind of music and a kind of dance (e.g., *waltz, tango, conga*), which suggests a polysemous relationship between *{music}* and *{dance}*.

Both of these approaches treat polysemy as a systematic phenomenon best described at the level of word families. However, while such a treatment reveals interesting macro-tendencies in the lexicon, it does little to dispel the possibility that homonymy might still operate on the micro-level of individual words (as demonstrated by the size of the exception

list needed for the first approach). We thus prefer to use an evidential case-by-case approach to detecting polysemy, connecting a pair of senses only when explicit local taxonomic evidence can be found to motivate a connection. This evidence can take many forms, so a patchwork of heuristic detectors is required. We describe here the three most interesting of these heuristics.

The coverage of each heuristic is estimated relative to that achieved by the *cousins* collection of 105 regular polysemy noun-sense groupings that are hand-coded in WordNet [7]. Over-generation is estimated relative to the overlap with the cousins exception list [7], which permits us to also estimate the accuracy of each heuristic.

Explicit Ontological Bridging: a sense pair $<\omega_1, \omega_2>$ for a word ω can be linked if ω_1 has a hypernym that can be lexicalized as M-H and ω_2 has a hypernym that can be lexicalized as M, the rationale being that ω_2 is the M of ω_1 and ω_1 is the H of ω_2. E.g., the word *olive* has a sense with a hypernym *{fruit-tree}*, and another with the hypernym *{fruit}*, therefore M = *fruit* and H = *tree*. *(Coverage: 12%, Accuracy: 94%).*

Hierarchical Reinforcement: if $<\alpha_1, \alpha_2>$ and $<\beta_1, \beta_2>$ are sense pairs for two words α and β where α_1 is a hypernym of β_1 and α_2 is a hypernym of β_2, then $<\alpha_1, \alpha_2>$ reinforces the belief that $<\beta_1, \beta_2>$ is polysemous, and vice versa. For example, *herb* denotes both a plant and a foodstuff in WordNet, and each of these senses has a hyponym that can be lexicalized as *sage*. *(Coverage: 7%, Accuracy: 12%).*

Cross-Reference: if $<\omega_1, \omega_2>$ is a sense pair for a word ω and the WordNet gloss for $\omega_2$explicitly mentions a hypernym of ω_1, then ω_2 can be seen as a conceptual extension of ω_1. For instance, the railway-compartment sense of *diner* mentions *restaurant* in its gloss, while another sense actually specifies *{restaurant}* as a hypernym. This suggests that the railway sense is an extension of the restaurant sense that uses the later as a ground for its definition. *(Coverage: 62%, Accuracy: 85%).*

These heuristics are very effective at arguing for polysemy on the local merits of individual words. However, for every creatively-useful instance of polysemy like *knife* (*{weapon}* versus *{edge-tool}*), there is an unhelpful instance like *capsule* (*{space-vehicle}* versus *{medicine}*), for one cannot meaningfully reuse aspirin-capsules as spacecraft, and vice versa. At present, we manually filter those instances of polysemy (almost 50%) from the set produced by the above heuristics whenever structural and behavioral properties are not preserved between senses.

4 Types of Ontological Creativity

The polysemy relationships that can be extracted from WordNet are merely the residue of past creativity by the language community. However, new creative insights can be generated by generalizing from these entrenched precedents, to either broaden existing categories and admit new members not previously considered eligible, or to re-categorize members of existing categories under different branches of the ontology.

Category Broadening: Imagine we want to broaden the WordNet category *{weapon}*. The members of this category can be enumerated by recursively visiting every hyponym of the category, which will include *{knife}, {gun}, {artillery}, {pike}*, etc. But by traversing polysemy links as well as *isa* relations, additional prospective members can be reached and admitted on the basis of their functional potential. Thus, the polysemy of *knife* causes not

only *{dagger}* and *{bayonet}* but *{steak_knife}* and *{scalpel}* to be visited. Stretching category boundaries even further, the generalization *edge_tool(x)* → *weapon(x)* allows the category *{edge_tool}* to be subsumed in its entirety, thereby allowing *{scissors}, {axe, ax}, {razor}* and all other sharp-edged tools to be recognized as having weapon-like potential.

Category broadening is a very revealing process, not only about the functional potential of everyday objects, but also about the inevitable gaps in an ontology like WordNet. For instance, the category *{apparel, clothing, clothes}* can be broadened to admit baseball gloves, anklets, metal helmets, furs and animal skins, while the category *{medicine, medication}* can be broadened to admit toiletries and oleoresins, and the category *{food}* can be broadened to admit a variety of potentially edible substances, some too disgusting to list here.

Category Hopping: Imagine, following the Torrance test, we want to move the concept *{coffee_can}* to a new category that will offer a functional perspective on how to effectively reuse old tin cans. The existing WordNet categories that house *{coffee_can}* can be enumerated by recursively visiting each of its hypernyms in turn, which will include *{can, tin_can}*,*{container}* and *{artifact}*. Now, each of these hypernyms is a potential point of departure to another category if, as well as traversing *isa* relations, we use polysemy relationships to slip from one rail of the ontology to another. WordNet defines *{coffee_can}* as a hyponym of *{can, tin_can}*, and from here a leap can be made to *{steel_drum, drum}*, since both are hyponyms of *{container}* whose glosses further specify them as kinds of *metal container*. From *{steel_drum, drum}* there exists a polysemy link to *{tympan, membranophone, drum}*, a non-container artifact which WordNet defines as a hyponym of *{percussion_instrument}*. This chain of reasoning, from *{coffee_can}* to *{tin_can}* to *{steel_drum}* to *{tympan, membranophone, drum}*, supports the creative insight that allows an old tin can to be used a musical drum, and central to this insight is the polysemy of *drum*. In general, polysemy supports creativity by providing just one very important link in the recategorization chain. A dog collar can be fashionably reused as a necklace because the polysemy of *collar* links *{collar}* to *{choker, collar}*. We can meaningfully think of jewelry as a piece of fine-art (and thus consider exhibiting it in a gallery) because of the polysemy of *gem* that links *{gem, jewel}* to *{gem, treasure}*. Likewise, we can think of photography as a fine art because photograph and art collide via the polysemy of *mosaic*, *vignette* and *scene*.

5 Creativity, Utility and Similarity

Some recategorizations will exhibit more creativity than others, largely because they represent more of a mental leap within the ontology. We can measure this distance using any of a variety of taxonomic metrics [11], and thus rank the creative outputs of our system. For instance, it is more creative to reuse a coffee can as a *{percussion_instrument}* than as a *{chamberpot, potty}*, since like *{tin_can}* the latter is already taxonomized in WordNet as a *{container}*. Any similarity metric (called σ, say) that measures the relative distance to the lowest common hypernym will thus attribute greater similarity to *{coffee_can}* and *{potty, chamberpot}* than to *{coffee_can}* and *{tympan, drum, membranophone}*. This allows us to measure the creative distance in a recategorization from α to γ as $1 - \sigma(\alpha, \gamma)$.

Of course, distance is not the only component of creativity, as any recategorization must also possess some utility to make it worthwhile (e.g., there is a greater distance still between tin cans and fish gills, but the former cannot be sensibly reused as the latter). In other words,

a creative product must be unfamiliar enough to be innovative but familiar enough to be judged relative to what we know already works. This is the paradox at the heart of ontological creativity: to be creative a recategorization must involve a significant mental leap in *function* but not in *form*, yet typically (e.g., in WordNet), both of these qualities are ontologically expressed in the same way, via taxonomic structure. This suggests that taxonomic similarity σ must be simultaneously maximized (to preserve structural compatibility) and minimized (to yield a creative leap).

Fortunately, polysemy offers a way to resolve this paradox. If a creative leap from α to γ is facilitated by a polysemous link from $<\beta, \gamma>$, the sensibility of the leap can be measured as $\sigma(\alpha, \beta)$ while the creativity of the leap can be measured as $1 - \sigma(\alpha, \gamma)$. The value of a creative product will be a function of both distance and sensibility, as the former without the latter is unusable, and the latter without the former is banal. The harmonic mean is one way of balancing this dependency on both measures:

$$value(\alpha, \gamma) = 2\sigma(\alpha, \beta)(1 - \sigma(\alpha, \gamma))/(1 + \sigma(\alpha, \beta) - \sigma(\alpha, \gamma))$$

Other variations on this formula can be used to give greater of lesser weight to the roles of sensibility and distance in determining the value of a creative insight.

6 Concluding Observations

The ideas in this paper have now been implemented in a computational system called *Kalos* (a Greek word connoting beauty through fitness of purpose [3]). A collection of 25 different polysemy detectors (of which 3 were described here) achieve 96% of the coverage offered by WordNet's own cousin relations, at a precision of 85%. In our pilot study, we focused on the subset of these polysemous relations that connect artifactual noun senses, where this subset is hand-filtered to yield 991 instances of behaviour-preserving, function transforming polysemy. Generalizing from these instances and performing a second phase of hand-checking to filter out spurious hypotheses, we are left with 454 inter-category subsumption hypotheses. These generalizations are a powerful addition to WordNet's upper and middle ontologies, facilitating a creative flexibility in determining category membership that is useful to a variety of applications, from creative writing tools to text understanding systems.

References

1. Wiggins, G. Categorizing Creative Systems, in: Proc of the 3rd Workshop on Creative Systems, IJCAI'03, Acapulco, Mexico. (2003).
2. Hutton, J.: Aristotle's Poetics. Norton, New York (1982).
3. de Bono, E. Parallel Thinking. Viking Press: London (1994).
4. Torrance, E. P. The Torrance Tests of Creative Thinking. Scholastic Testing Service. Bensonville, Illinois. (1990).
5. Miller, G. A.: WordNet: A Lexical Database for English. Communications of the ACM, Vol. 38 No. 11 (1995).
6. Lenat, D., Guha, R. V.: Building Large Knowledge-Based Systems. Addison Wesley (1990).
7. WordNet documentation. `http://www.princeton.edu/~wn/` (2003).

8. Peters, W., Peters, I., Vossen, P. Automatic sense clustering in EuroWordNet. In: Proc of the 1st international conference on Language Resources and Evaluation. Spain. (1998).
9. Peters, I., Peters, P. Extracting Regular Polysemy Patterns in WordNet. Technical Report, University of Sheffield, UK. (2000).
10. Peters, W., Peters, I. Lexicalized Systematic Polysemy in WordNet. In the proceedings of the 2nd international conference on Language Resources and Evaluation. Athens. (2000).
11. Budanitsky, A., Hirst, G. Semantic Distance in WordNet: An experimental, application-oriented evaluation of five measures. In: Proc of the Workshop on WordNet and Other Lexical Resources, North-American chapter of ACL. Pittsburgh. (2001).
12. Lakoff, G.: Women, Fire and Dangerous Things. Uni. of Chicago Press: Chicago (1987).