A Service-Oriented Architecture for Metaphor Processing

Tony Veale
School of Computer Science and Informatics
University College Dublin
Belfield, Dublin D4, Ireland.
Tony.Veale@UCD.ie

Abstract

Metaphor is much more than a pyrotechnical flourish of language or a fascinating conceptual puzzle: it is a cognitive *lever* that allows speakers to leverage their knowledge of one domain to describe, reframe and understand another. Though NLP researchers tend to view metaphor as a problem to be solved, metaphor is perhaps more fittingly seen as a *solution* to be used, that is, as an important tool in the support of creative thinking and the generation of diverse linguistic outputs. Since it pays to think of metaphor as a foundational cognitive service, one that can be exploited in a wide array of creative computational tasks, we present here a view of metaphor as a public *Web service* that can be freely called on demand.

1 Introduction

Metaphor is a knowledge-hungry phenomenon. Fortunately, much of the knowledge needed for the processing of metaphor is already implicit in the large body of metaphors that are active in a language community (e.g. Martin, 1990; Mason, 2004). For existing metaphors are themselves a valuable source of knowledge for the production of new metaphors, so much so that a system can mine the relevant knowledge from corpora of figurative text (see Veale, 2011; Shutova, 2010). Thus, though linguistic metaphors are most naturally viewed as the output of a language generation process, and as the input to a language understanding process, it is just as meaningful to view the conceptual metaphors that underpin these linguistic forms as an *input* to the generation process and an *output* of the understanding process. A rich source of existing linguistic metaphors, such as a text corpus or a database of Web n-grams, can thus be viewed as an implicit source of the knowledge a system needs to generate and understand novel linguistic metaphors. Of course, if one finds Web data to be a useful resource for metaphor, it also makes sense to think of the algorithms and tools for manipulating this knowledge as Web *services*, online systems that hide the complexity of metaphor processing yet which can be called upon to generate and understand linguistic metaphors on demand. Such metaphors can then, in turn, be exploited in higher-level linguistic outputs such as stories and poems by yet other, inter-operable Web services.

There are compelling reasons to see metaphor as a service rather than a problem. For one, many creative language tasks – such as poetry, joke and story generation – require the conceptual and linguistic divergence offered by metaphor. When metaphor is offered as a reusable Web service, such systems need not implement their own metaphor solutions, and are instead freed to focus on providing their own unique competences. For another, even as a problem, metaphor is not yet a *standardized* problem in NLP, and so different researchers focus on diverse aspects of metaphor using a wide range of bespoke models and approaches. But when these models are provided as public services, researchers are free to draw from a rich ecology of complementary solutions. New approaches to metaphor, and to broader problems of linguistic creativity, may then emerge as researchers and developers mix-and-match services to meet their own specific application needs.

A *Service-Oriented Architecture*, or SOA, is one in which solution logic is presented in the form of discoverable, modular and composable services that hide the complexity of their data and their inner workings (Erl, 2008). This paper advocates for a SOA treatment of metaphor in the form of open and reusable Web services. To this end, a number of metaphor Web services are

Proceedings of the Second Workshop on Metaphor in NLP, pages 52–60,
Baltimore, MD, USA, 26 June 2014. ©2014 Association for Computational Linguistics

presented, to both offer a practical demonstration of the merits of SOA and to kick-start further development of metaphor services by the field. After discussing related work in section 2, we thus present a series of publically-accessible metaphor services, for generating creative similes, for performing divergent categorization, for generating new affective metaphors from old, for generating metaphor-rich poetry, and for generating metaphor-inspired character arcs for stories.

2 Related Work and Ideas

Metaphor has been studied within computer science for four decades, yet it remains largely at the periphery of NLP research. The reasons for this marginalization are pragmatic ones, since metaphors can be as challenging as human creativity will allow. The greatest success has thus been achieved by focusing on conventional metaphors (e.g., Martin, 1990; Mason, 2004), or on specific domains of usage, such as figurative descriptions of mental states (e.g., Barnden, 2006).

From the earliest computational forays, it has been recognized that metaphor is fundamentally a problem of knowledge representation. Semantic representations are, by and large, designed for well-behaved mappings of words to meanings – what Hanks (2006) calls *norms* – but metaphor requires a system of soft preferences rather than hard (and brittle) constraints. Wilks (1978) thus proposed a *preference semantics* approach, which Fass (1991,1997) extended into a *collative semantics*. In contrast, Way (1990) argued that metaphor requires a dynamic concept hierarchy that can stretch to meet the norm-bending demands of figurative ideation, though her approach lacked specific computational substance.

More recently, some success has been obtained with statistical approaches that side-step the problems of knowledge representation, by working instead with implied or latent representations that are derived from word distributions. Turney and Littman (2005) show how a statistical model of relational similarity that is constructed from Web texts can retrieve the correct answers for proportional analogies, of the kind used in SAT/GRE tests. No hand-coded knowledge is employed, yet Turney and Littman's system achieves an average human grade on a set of 376 real SAT analogies.

Shutova (2010) annotates verbal metaphors in corpora (such as "to *stir* excitement", where "stir" is used metaphorically) with the corresponding conceptual metaphors identified by Lakoff and Johnson (1980). Statistical clustering techniques are then used to generalize from the annotated exemplars, allowing the system to recognize and retrieve other metaphors in the same vein (e.g. "he *swallowed* his anger"). These clusters can also be analyzed to find literal paraphrases for a given metaphor (e.g. "to *provoke* excitement" or "*suppress* anger"). Shutova's approach is noteworthy for operating with Lakoff and Johnson's inventory of conceptual metaphors without using an explicit knowledge representation of the knowledge domains involved.

Hanks (2006) argues that metaphors exploit distributional norms: to understand a metaphor, one must first recognize the norm that is exploited. Common norms in language are the preferred semantic arguments of verbs, as well as idioms, clichés and other multi-word expressions. Veale and Hao (2007a) suggest that stereotypes are conceptual norms that are found in many figurative expressions, and note that stereotypes and similes enjoy a symbiotic relationship that has obvious computational advantages. Similes rely on stereotypes to illustrate the qualities ascribed to a topic, while stereotypes are often promulgated via proverbial similes (Taylor, 1954). Veale and Hao (2007a) show how stereotypical knowledge can be acquired by harvesting "Hearst" patterns (Hearst, 1992) of the form "as P as C" (e.g. "*as smooth as silk*") from the Web. They go on to show in (2007b) how this body of stereotypes can be used in a Web-based model of metaphor generation and comprehension.

Veale (2011) employs stereotypes as the basis of the *Creative Information Retrieval* paradigm, by introducing a variety of non-literal-matching wildcards in the vein of Mihalcea (2002). In this paradigm, **@Noun** matches any adjective that denotes a stereotypical property of *Noun* (so e.g. ***@knife*** matches *sharp*, *pointy*, etc.) while **@Adj** matches any noun for which *Adj* is stereotypical (e.g. ***@sharp*** matches *sword, laser, razor*, etc.). In addition, **?Adj** matches any property / behavior that co-occurs with, and reinforces, the property denoted by *Adj* in similes; thus, ***?hot*** matches *humid, sultry* and *spicy*. Likewise, **?Noun** matches any noun that denotes a pragmatic neighbor of *Noun*, where two words are neighbors if corpora attest to the fact that they are often clustered together as comparable ideas, as in "*lawyers and doctors*" or "*pirates and thieves*". The knowledge needed for **@** is obtained by harvesting text from the Web, while that for **?** is obtained by mining Google 3-grams for instances of the form "*Xs and Ys*" (Brants and Franz 2006).

Creative Information Retrieval (CIR) can be used as a platform for the design of many Web services that offer linguistic creativity on demand. By enabling the flexible retrieval of n-gram data for non-literal queries, CIR allows a wide variety of creative tasks to be reimagined as simple IR tasks (Veale 2013). In the next section we show how CIR facilitates the generation of creative similes from linguistic readymades.

3 The Jigsaw Bard

Similes and stereotypes enjoy a mutually beneficial relationship. Stereotypes anchor our similes in familiar concepts with obvious features, while similes, for their part, further popularize these stereotypes and entrench them in a culture. Since the core of any good simile is an evocative stereotype that embodies just the qualities we want to communicate (see Fishelov, 1992), simile generation is essentially a problem of apt stereotype retrieval. However, we can also turn this view on its head by asking: what kinds of simile might be generated from a given stereotype, or a linguistic combination or two or more lexicalized stereotypes? For instance, were we to consider the many phrases in the Google n-grams that combine a lexicalized stereotype with an affective modifier (such as "*cold fish*"), or that combine multiple stereotypes with shared qualities (such as "*chocolate espresso*" (brown) or "*robot fish*" (cold and emotionless)), we might imagine re-purposing these phrases as part of a novel simile such as "*as emotionless as a robot fish*" or perhaps even "*as smooth as a chocolate martini*".

The n-grams encountered and re-purposed in this way are linguistic *readymades*, in much the same way that the everyday objects that catch an artist's eye for their secondary aesthetic qualities become art when re-imagined as art (see Taylor, 2009). Readymades in art are a product of serendipity: an artist encounters an object – perhaps a humble tool, or the discarded detritus of modern life – and sees in it a desired quality that can be brought to the fore in the right setting. Using a computer, however, linguistic readymades can be harvested from a resource like the Google n-grams on a near-industrial scale. Using CIR, a query can be issued for all bigrams that combine a lexicalized stereotype with a modifier that accentuates one of the stereotype's core qualities. Such a query might be "?@P @P" where P denotes a property like *cold* or *smooth*; the CIR query "?@cold @cold" thus matches "*wet fish*". Likewise, a CIR query of the form "@P @P" will retrieve all Google bigrams that juxtapose two lexicalized stereotypes for the same property P; thus, "@cold @cold" retrieves "*January rain*", "*winter snow*" and "*robot fish*". More elaborate queries will retrieve more elaborate n-grams, such as "*snow-covered grave*" and "*bullet-riddled corpse*" (again for the property *cold*).

The *Jigsaw Bard* is a creative Web service that exploits this notion of linguistic readymades to generate novel creative similes on demand. Of course, the Bard only appears to "invent" similes on demand (for a given input property like *cold*). In fact, the Bard has already scanned all of the Google n-grams to index a great many potential readymades that may, for some future request, be re-purposed as a creative simile. In keeping with the principles of SOA, the *Bard* does as little processing in real time as possible. Thus, when called as a Web service, it reliably retrieves, with remarkable speed, scores of fascinating similes that have already been indexed for a property. The *Jigsaw Bard* service can be accessed online at: *www.educatedinsolence.com/jigsaw/*

4 Thesaurus Rex

Metaphor is both a viewfinder and an adjustable lens: it helps us to find distant objects that share surprising similarities, and it allows us to focus on shared qualities that are not always apparent in a more conventional setting. So while metaphor exploits our sense of similarity to generate resonant yet surprising juxtapositions, it also *directs* our sense of similarity, to highlight shared qualities that might otherwise remain unnoticed.

One cannot have an eye for metaphor without also having a well-developed sense of similarity. Lexico-semantic resources like WordNet offer NLP researchers a comprehensive and widely-used basis for measuring the similarity of two words or lexical concepts (see Fellbaum, 1998). Yet WordNet offers a somewhat monochromatic view of conceptual structure: it is a convergent structure in which every lexical concept is put in its correct place according to conventional usage. Metaphor requires a more kaleidoscopic view of conceptual structure, in which the many diverse and unconventional ways that a word, object or idea may be used can be brought into play. The best place to find this kind of divergence is not a carefully curated resource like WordNet, but the unfiltered clamor and eclecticism of the Web.

One can see the many ways in a given lexical concept is viewed on the Web using a simple search query. The "such" construction, as used in “*expensive foods such as lobster and caviar*”, tells us that lobster and caviar are seen by some as expensive foods. The more often this view is found on the Web, the more credibility it can be given. Yet rather than trawl the Web for all uses of the “such” construction, it pays to be targeted and parsimonious in our searches. For instance, suppose a system already possesses the stereotypical association that *Champagne* is *expensive*. A targeted query of the form “*expensive * such as * and Champagne*” will now retrieve Web texts that indicate other, related expensive items, and an umbrella category in which to place them all. Google, for example, provides the snippets “*expensive wines such as French Burgundy and Champagne*“, “*expensive products such as Cognac and Champagne*” and “*expensive and exotic foodstuffs such as caviar, seafood, hares, game, wine and champagne*” in response to this query.

Knowing that *Champagne* and *caviar* are expensive items in the same category, a system can now look for the other categories they also share, and so the query “*expensive * such as caviar and Champagne*” finds that they are also considered to be *expensive delicacies* on the Web. By starting from a small seed of stereotypical knowledge (e.g. that *Champagne* is *expensive*), a system can generate a large body of targeted Web queries to elaborate and expand this knowledge. As new qualities and nuanced categories are acquired, these too can feed into the targeted acquisition process to form a virtuous bootstrapping circle.

As a result, a system that starts from a seed of 12,000 or so stereotypical associations will acquire over 1.5 million fine-grained categorizations in just five cycles of bootstrapping. Thus, for instance, a system can view *Champagne* as more than just a food, as the Web snippet “*luxury goods such as diamonds and champagne*” can attest. These many fine-grained, overlapping and competing perspectives – when combined in a Web service for divergent categorization we call *Thesaurus Rex* – provide the kaleidoscopic swirl of possibilities that WordNet is so lacking but which creative metaphors can do so much with.

Ask WordNet what the lexicalized concepts *War* and *Peace*, or *Life* and *Death*, or *Divorce* and *War* have in common, and its answer cannot fail but to disappoint. WordNet simply does not possess the fine-grained category structure to suggest what features might be shared by these very different concepts, even if, ironically, it can be used to generate a meaningful-seeming *numerical* measure of similarity in each case. In contrast, the *Thesaurus Rex* Web service will return a wealth of informative commonalities in each case. For instance, Figure 1 below presents a phrase cloud of the nuanced categories that are shared by both *War* and *Divorce*. Note how each is categorized as a stressful event, an unexpected and dramatic event, a traumatic event and an emotional event (eagle-eyed readers will note that each is also an adverse economic event).

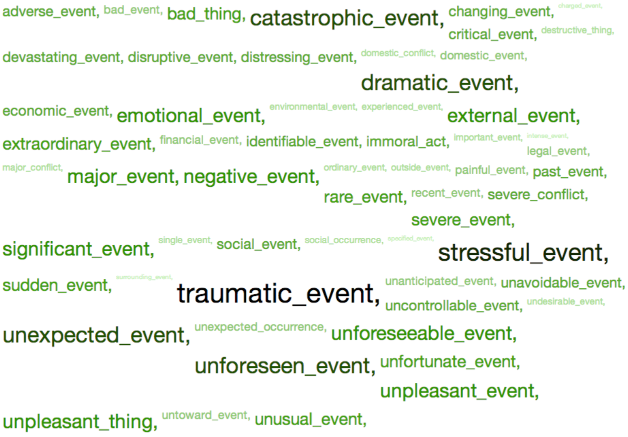

Figure 1. Shared categories for *War* and *Divorce*.

Thesaurus Rex thus provides a valuable service to any system that wishes to take a divergent view of conceptual structure, whether for purposes of literal similarity assessment or for non-literal metaphoric reasoning. *Rex* can be used as a browsing tool by Web users in search of insights or apt comparisons – for instance, one can go from *Leadership* to *Creativity* via the categories *soft skill*, *valuable skill* or *transferable skill* – or as a flexible similarity service that supports 3rd-party metaphor processing systems. It should be noted that while *Rex* relies on the Web for its divergent view of the world, it does not sacrifice quality for quantity. Veale & Li (2013) show that a combination of *Thesaurus Rex* and WordNet produces similarity scores for the standard Miller & Charles (1991) test-set that achieve a 0.93 correlation with human judgments. This is as good as the best machine-learning systems (which do not *explain* their ratings the way that *Rex* can) and far superior to any WordNet-only approach. The Thesaurus Rex service can be accessed here:

http://boundinanutshell.com/therex2

5 Metaphor Magnet

In many ways, a metaphor resembles a query in information retrieval (IR). Metaphors, like queries, allow us to simultaneously express what we believe and to elicit further information that may bolster or refute our beliefs. Metaphors, like queries, are often short and concise, and require unpacking and expansion to be properly understood and acted upon. An expanded IR query is considered successful if it leads to the retrieval of a richer set of relevant information sources. Likewise, an expanded metaphor can be considered successful if expansion produces a rich interpretation that is consonant with, and consistently adds to, our beliefs about a particular topic.

Of course, there are important differences between metaphors, which elicit information from other humans, and IR queries, which elicit information from search engines. For one, IR fails to discriminate literal from non-literal language (see Veale 2004, 2011), and reduces any metaphoric query to literal keywords and keyphrases that are matched near-identically to texts (see Salton, 1968; Van Rijsbergen 1979). Yet everyday language shows that metaphor is an ideal form for expressing our information needs. A query like "*Steve Jobs was a good leader*", say, can be viewed by a creative IR system as a request to consider all the ways in which leaders are typically good, and to then consider all the metaphors that can most appropriately be used to convey these viewpoints about *Steve Jobs*.

IR techniques such as corpus-based query expansion can thus be used to understand and generate metaphors on demand, if IR staples like query expansion (see Vorhees, 1998; Navigli and Velardi, 2003) are made both *affect-driven* and *metaphor-aware*. Expansion in each case can be performed using a comprehensive database of affective stereotypes that indicate e.g. the stereotypical properties of *geniuses*, *gurus* and *tyrants*.

Let us return to the example of Steve Jobs *qua* leader. Using the CIR query "leader is a ?leader" a range of different kinds of leader can be retrieved. For instance, the Google n-grams oblige with the 4-grams "leader is a *visionary*", "leader is a *tyrant*", "leader is a *terrorist*", "leader is a *master*", "leader is a *shepherd*", "leader is a *dictator*", "leader is an *expert*", "leader is a *teacher*" and "leader is a *catalyst*". But which of these views is consonant with being a *good* leader? If one wanted to criticize Jobs' leadership of Apple, then the stereotypes *tyrant*, *terrorist* and *dictator* offer clearly negative perspectives. In contrast, the stereotypes *visionary*, *shepherd, expert* and *teacher* are all positive, while *master* and *catalyst* may each evoke both good and bad qualities.

The under-specified positive metaphor "Steve Jobs was a *good* leader" can thus be expanded, via the Google n-grams, to generate the specific positive metaphors "Steve Jobs was a *visionary*", "Steve Jobs was a *shepherd*", "Steve Jobs was an *expert*" and "Steve Jobs was a *teacher*". Likewise, the under-specified negative metaphor "Steve Jobs was a *bad* leader" can be expanded to yield "Steve Jobs was a *tyrant*", "Steve Jobs was a *dictator*" and "Steve Jobs was a *terrorist*". The stereotypical properties of the vehicle in each case – such as *tyrant* or *expert* – can then be projected onto the tenor, Steve *Jobs* qua leader. Which properties of the vehicle are most relevant to Steve Jobs as a leader? CIR is again used to rank properties by their relevance to leadership. For instance, the CIR query "@tyrant leader" finds Google 2-grams where a property of *tyrant* is used to describe a leader – such as "*cruel leader*" and "*demanding leader*" – and allows a system to rank the properties of *tyrant* according to the frequencies of these corresponding 2-grams.

Metaphor Magnet is such a system. Deployed as a Web service that generates and expands affective metaphors on demand, *Metaphor Magnet* allows clients (human users or 3rd-party software systems) to enter single terms (such as *leader*), compound terms with an affective spin (such as *good leader* or *+leader*), or copula statements such as "*Steve Jobs is a +leader*". For each input, the service marries its extensive knowledge of lexicalized stereotypes to the grand scale of the Google n-grams, to meaningfully expand upon what it has been given and to generate the most appropriate affective elaborations and interpretations it can muster. In each case, *Metaphor Magnet* provides a rich property-level explanation of its outputs. So, for instance, if Steve Jobs were to be viewed as a *master*, the properties *skilled*, *enlightened*, *free* and *demanding* are all highlighted as being most appropriate. The Metaphor Magnet service can be accessed here:

http://boundinanutshell.com/metaphor-magnet-acl

6 Metaphorize with Metaphor Eyes

Metaphor Magnet offers a property-theoretic view of metaphor: since its model of the world is entirely property-based – in which words denote stereotypes that map to highly salient properties – it sees metaphor interpretation as a question of which properties are mapped from the vehicle to the tenor. *Metaphor Magnet* lacks a proposition-level view of the world, in which stereotypes are linked to other stereotypes by arbitrary relations. Thus, though it knows that *scientists* are *logical* and *objective*, it does not know, and cannot use, the generalizations that scientists work in labs, wear white coats, conduct experiments, write up their results, and so on. Another service, called *Metaphor Eyes*, remedies this deficiency by employing a propositional model of the world that reasons with *subject-relation-object* triples rather than *subject-attribute* pairs. *Metaphor Eyes* acquires its world-model from a variety of sources (see Veale & Li, 2011), but the most fascinating of these sources is a niche Web-service offered (until recently) by the Google search-engine.

Many users of Web search-engines still enter full NL questions as search queries, even though most engines do not perform syntactic analysis. The Google engine maintains a record of frequently-posed queries and helpfully suggests apt completions for any familiar-seeming inputs. Google also provides a *completions service* (now sadly defunct) through which one may automatically retrieve the most common completions for any given query stub. The pairing of these observations – full NL questions plus the availability of common completions – allows a computer to acquire a propositional model of the world by polling Google for completions to question stubs of the form "*Why do Xs* ...". Why-do questions are remarkably revealing about the beliefs that we take for granted when speaking to others. The query "*Why do dogs bury bones*" tells us more than the fact that some dogs bury bones; it tells us that the questioner presupposes this to also be a fact held by the addressees of the query, and so it is a stereotypical generalization over all dogs. By repeating polling Google for completions of the query "Why do Xs", where X is any concept the system wishes to flesh out, *Metaphor Eyes* acquires a large body of common-sense beliefs.

Metaphor Eyes retrieves apt vehicles for a given a tenor concept **T** using the simple CIR query "**?T**". Thus, given *philosopher* as a tenor, *Metaphor Eyes* considers *scholar*, *moralist*, *theologian*, *historian*, *scientist*, *visionary*, *explorer*, *thinker*, *sage*, *pundit*, *poet* and even *warrior* as possible vehicles for a copula metaphor. For any given vehicle it then attempts to accommodate its knowledge of that vehicle into its representation of the tenor, by considering which propositions associated with the vehicle can be turned into apt propositions about the tenor. Consider the choice of *explorer* as a vehicle, producing the copula metaphor *philosophers are explorers*. Knowing that explorers perform wanderings, go on quests and seek knowledge, *Metaphor Eyes* looks for evidence in the Google n-grams that one or more of these propositions can just as well be said of philosophers. The 3-gram "philosopher's quest" attests to the aptness of the proposition "philosophers go on quests", while the 3-gram "philosopher's knowledge" attests to "philosophers look for knowledge". The 2-gram "wandering philosopher" additionally attests to the proposition that philosophers perform wanderings of their own.

Metaphor Eyes views metaphor as a representational lever, allowing it to fill the holes in its weak understanding of one concept by importing relevant knowledge from a neighboring concept. As such, in offering a partial solution to metaphor as a *problem*, it simultaneously views metaphor as a an acquisition *solution* in its own right. The *Metaphor Eyes* service can be accessed here:

http://boundinanutshell.com/metaphor-eye/

7 Stereotrope Poetry Generation

The copula form "X is a Y" is metaphor at its simplest and its purest, which perhaps explains why the form is far more prevalent in the metaphor literature than it is in real texts. Metaphor in the wild thrives in a wide variety of syntactic forms and rhetorical guises, with the most creatively rhetorical found in poetry. Yet while metaphors are the stuff of poetry, a well-written poem is much more than a bag of fancy metaphors. Coherent poems are driven by a coherent *master* metaphor, a schema that governs a poet's choice of related metaphors to elaborate this core idea.

A key benefit of the SOA philosophy is that services represent modular chunks of solution logic that need not, and do not, do everything for themselves. Ideally, our Web services should be reusable modules that can be composed, mashed-up and further elaborated by other developers to yield new services. In this spirit, *Stereotrope* is a service that generates poems from the metaphors produced by the *Metaphor Magnet* Web service.

Given a topic on which to wax poetically, *Stereotrope* calls on *Metaphor Magnet* to suggest a master metaphor around which its poem might be organized. Suppose our topic is *love*, and that *Metaphor Magnet* responds with, among others, the trope *Love is a Fire* (this copula metaphor has a frequency of 331 in the Google n-grams). Choosing this familiar trope as the core of its poem, *Stereotrope* now asks *Metaphor Magnet* to produce elaborations of this metaphor. *Metaphor Magnet* generates elaborations of *Love is a Fire* that include *Love is a Shining Flame*, *Love is a Dazzling Explosion* and *Love is a Raging Cauldron*. These elaborations – once rendered in the typical rhetorical forms of poetry – are then packaged by *Stereotrope* into a complete poem.

A useful rhetorical device is the *Superlative*. For instance, Metaphor Magnet suggests that for *Love is a Fire*, the properties *hot*, *bright* and *burning* can all be sensibly projected from *Fire* onto *Love* (as attested by the Google n-grams). The explicit statement *Love is a Fire* lacks a certain something in a poem, yet the same meaning can be suggested with the superlative forms "*No fire is hotter*" or "*No fire is brighter*". By looking to attested combinations in the Google n-grams, *Stereotrope* notices that "brightly" is an adverb that frequently modifies "burning", and so it also suggests the superlative "*No fire burns more brightly*". Yet by also noting that *hot* and *bright* are mutually reinforcing properties, since *bright* ∈ *?hot*, it sees that the line "*No fire is hotter or burns more brightly*" will squeeze all three projected properties of *Fire* into a single superlative.

Stereotrope also calls upon the *Metaphor Eyes* Web-service to provide a proposition-level understanding of the world, for its poems must do more than allude to just the properties of entities. Unfortunately, banality is tacitly a pre-condition for the inclusion of almost any generalization in a common-sense knowledge-base. For it is precisely because so many of us tacitly share these beliefs that they are so worthy of inclusion in a knowledge-base and so unworthy of mention in a poem that rises above the obvious. Yet with the right rhetorical packaging, even a boring generalization can be pushed into the realm of the provocative, allowing an automated poetry system to temporarily slip the surly bonds of reality.

Consider the generalization "*celebrities ride in limousines*". Though it may fail to provoke when baldly expressed in this form, *Stereotrope* notes that *limousines* have some interesting qualities. They are typically *long*, for one, and though it does not believe celebrities to be typically short, it notes from the Google n-grams that the 2-gram "*short celebrities*" is also frequent enough to be an interesting talking point. Combining these two observations, it generates the rhetorical question "*Why do the shortest celebrities ride in the longest limousines?*". Though *Stereotrope* has no real insight into the frailty of celebrity egos, vertically challenged or otherwise, it is attracted to the elegant opposition of *long* vs. *short* that can be injected into this otherwise banal generalization.

As a rule, *Stereotrope* attempts to shoehorn a provocative opposition into any proposition that is said to be topic-relevant by *Metaphor Eyes*. Thus, knowing that arrows are fired from bows, that bows are *curved* and that arrows are *straight*, it generates the rhetorical question "*Why do the most curved bows fire the straightest arrows?*". The point is to suggest a more profound meaning beneath the surface. For when Don Corleone tells us that a fish rots from the head, he is not really talking about fish, but about how power corrupts an organization from the top down. Banal facts, when expressed in the right way, allude to a figurative meaning greater than themselves. By packaging its meagre facts in a rhetorical guise, *Stereotrope* can allude to a poetic meaning that lies outside its own power to comprehend.

Stereotrope generates the following poem from the master metaphor *Marriage is a Prison*:

The legalized regime of this marriage

My marriage is an emotional prison
Barred visitors do marriages allow
The most unitary collective scarcely organizes so much
Intimidate me with the official regulation of your prison
Let your sexual degradation charm me
Did ever an offender go to a more oppressive prison?
You confine me as securely as any locked prison cell
Does any prison punish more harshly than this marriage?
You punish me with your harsh security
The most isolated prisons inflict the most difficult hardships
Marriage, you disgust me with your undesirable security

Since the *Stereotrope* service complements the products of *Metaphor Magnet* (and *Metaphor Eyes*), it is engaged for each individual output of *Metaphor Magnet* directly. Thus, once again see:

http://boundinanutshell.com/metaphor-magnet-acl

8 The Flux Capacitor

The landmark television series *Breaking Bad* showcases story-telling at its most dramatic and its most transformational. It tells the tale of put-upon family man Walter White, a scientist with a brilliant mind who is trapped in the colorless life of a high-school chemistry teacher. When Walt is diagnosed with terminal lung cancer, he throws suburban caution to the wind and embraces a life of crime, first as a drug chemist of blue crystal meth and later as the ruthless drug baron *Heisenberg*. Walt's transformation, "from Mister Chips to Scarface" (in the words of the show's creator Vince Gilligan) is psychologically compelling because it is so unexpected yet so strongly rooted in our common-sense notions of similarity: for a drug chemist and a chemistry teacher share many of the same skills, while a drug baron embodies many of the same moral flaws as a drug chemist.

Literary transformations are often freighted with metaphorical meaning. Just think of the transformations of people into apt animals or plants in Ovid's *Metamorphoses*, or of Gregor Samsa's sudden, shame-driven transformation into a "gigantic vermin" in Franz Kafka's *Metamorphosis*. In *Breaking Bad*, where Walt's central transformation is slow-burning rather than magically immediate, a literal transformation is explained by the same kind of similarity judgments that motivate many of our metaphors. A service for producing apt metaphors, rooted in meaningful similarities, can thus be re-purposed to instead propose unexpected-but-apt character arcs for psychologically-compelling stories.

The *Flux Capacitor* is a new Web-service-in-development that re-packages the outputs of the *Metaphor Eyes* and *Metaphor Magnet* services as literal character transformations for use in computer-generated stories. The *Flux Capacitor* is thus conceived as a middleware service whose outputs are intended as inputs to other services. It does not package its outputs as metaphors, and nor does it package them as finished stories: rather, embracing the SOA philosophy of modularity and reuse, it produces Hollywood-style *pitches* that may underpin an interesting narrative that is to be fleshed out by another service or system.

Walter White's journey from chemistry teacher to drug baron is made believable by similarity, but it is made stimulating by *dissimilarity*. Like the best metaphors, a thought-provoking character transformation marries states that are both similar and incongruously dissimilar. The *Flux Capacitor* thus ranks the metaphors it receives from other services by their ability to surprise: a character arc from A to B is all the more surprising if our stereotype of A has properties that conflict with those in our stereotype of B. So the *Flux Capacitor* suggests the transformation of a scientist into a priest, or of a nun into a prostitute, or a king into a slave, or a fool into a philosopher, to capitalize on the dramatic possibilities of the oppositions that emerge in each case. The property-level interpretations of a character arc are given by *Metaphor Magnet*, while proposition-level insights are given by *Metaphor Eyes*.

The *Flux Capacitor* uses a variety of other techniques to ensure the meaningfulness of its proposed character arcs. For instance, it uses semantic knowledge to ensure that no transformation will change the gender of a character, and pragmatic knowledge to ensure that no transformation will reverse the age of a character. The *Flux Capacitor* is at present still being tested, but will soon be deployed as its own public Web service, where it may find useful work as a pitcher of new ideas to story-generation systems.

9 Out of the Mouths of Babes and Bots

The services described in this paper all operate in *pull* mode, where figurative products are generated on demand for the 3rd-party systems or users that ask for them. Each service produces HTML for human users and XML for automated queries.

We conclude this paper then by discussing an alternative model that has been overlooked here: a *push* mode of operation in which services broadcast their outputs, hopeful but unsolicited, to users or systems that may find some serendipitous value in being surprised in this way. Twitter is the ideal midwife for pushing automated metaphors into the world. For Twitter supports *twitterbots*, automated systems (or *bots*) that generate their own tweets, largely for the consumption and edification of human Twitter users. A new *twitterbot* named ***MetaphorIsMyBusiness*** (handle: **@*MetaphorMagnet***) employs all of the services described in previous sections to generate a novel creative metaphor every hour, on the hour.

@MetaphorMagnet's outputs are the product of a complex reasoning process that combines a comprehensive knowledge-base of stereotypical norms with real usage data from the Google n-grams. Though encouraged by the quality of its outputs, we continue to expand its expressive range, to give the twitterbot its own unique voice and identifiable aesthetic. Outputs such as "*What is an accountant but a timid visionary? What is a visionary but a bold accountant?*" lend the bot a sardonic persona that we wish to develop further.

We have seen the advantages to packaging metaphor systems as Web services, but there are also real advantages to packing metaphor Web-services as twitterbots. For one, the existence of mostly random bots that make no use of world knowledge or of metaphor theory – such as the playfully subversive *@metaphorminute* bot – provides a competitive baseline against which to evaluate the meaningfulness and value of the insights that are pushed out into the world by theory-driven / knowledge-driven twitterbots like *@MetaphorMagnet*. For another, the willingness of human Twitter users to *follow* such accounts regardless of their provenance, and to *retweet* the best outputs from these accounts, provides an empirical framework for estimating (and promoting) the figurative quality of the back-end Web services in each case. Finally, such bots may reap some social value in their own right, as sources of occasional insight, wit or profundity, or even of useful metaphors that are subsequently valued, adopted, and re-worked by human speakers.

References

Barnden, J. A. (2006). Artificial Intelligence, figurative language and cognitive linguistics. In: G. Kristiansen, M. Achard, R. Dirven, and F. J. Ruiz de Mendoza Ibanez (Eds.), *Cognitive Linguistics: Current Application and Future Perspectives*, 431-459. Berlin: Mouton de Gruyter.

Brants, T. and Franz, A. (2006). *Web 1T 5-gram Ver. 1*. Linguistic Data Consortium.

Erl, T. (2008). *SOA: Principles of Service Design.* Prentice Hall.

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

Fass, D. (1997). Processing Metonymy and Metaphor. *Contemporary Studies in Cognitive Science & Technology*. New York: Ablex.

Fellbaum, C. (ed.) (1998). *WordNet: An Electronic Lexical Database*. MIT Press, Cambridge.

Fishelov, D. (1992). Poetic and Non-Poetic Simile: Structure, Semantics, Rhetoric. *Poetics Today*, 14(1), 1-23.

Hanks, P. (2006). Metaphoricity is gradable. In: Anatol Stefanowitsch and Stefan Th. Gries (Eds.), *Corpus-Based Approaches to Metaphor and Metonymy*,. 17-35. Berlin: Mouton de Gruyter.

Hearst, M. (1992). Automatic acquisition of hyponyms from large text corpora. *In Proc. of the 14th International Conference on Computational Linguistics*, pp 539–545.

Lakoff, G. and Johnson, M. (1980). *Metaphors We Live By*. University of Chicago Press.

Martin, J. H. (1990). A Computational Model of Metaphor Interpretation. New York: Academic Press.

Mason, Z. J. (2004). CorMet: A Computational, Corpus-Based Conventional Metaphor Extraction System, *Computational Linguistics*, 30(1):23-44.

Mihalcea, R. (2002). The Semantic Wildcard. In *Proc. of the LREC Workshop on Creating and Using Semantics for Information Retrieval and Filtering*. Canary Islands, Spain, May 2002.

Miller, G. A. and Charles, W. G. (1991). Contextual correlates of semantic similarity. *Language and Cognitive Processes* **6**(1):1-28.

Navigli, R. and Velardi, P. (2003). An Analysis of Ontology-based Query Expansion Strategies. In Proc. of the workshop on Adaptive Text Extraction and Mining (ATEM 2003), at ECML 2003, the 14th European Conf. on Machine Learning, 42–49.

Salton, G. (1968). *Automatic Information Organization and Retrieval.* New York: McGraw-Hill.

Shutova, E. (2010). Metaphor Identification Using Verb and Noun Clustering. In *the Proc. of the 23rd International Conference on Computational Linguistics*, 1001-1010.

Taylor, A. (1954). Proverbial Comparisons and Similes from California. *Folklore Studies* 3. Berkeley: University of California Press.

Taylor, M. R. (2009). *Marcel Duchamp: Étant donnés* (Philadelphia Museum of Art). Yale University Press.

Turney, P.D. and Littman, M.L. (2005). Corpus-based learning of analogies and semantic relations. *Machine Learning* 60(1-3):251-278.

Van Rijsbergen, C. J. (1979). *Information Retrieval.* Oxford: Butterworth-Heinemann.

Veale, T. (2004). The Challenge of Creative Information Retrieval. *Computational Linguistics and Intelligent Text Processing: Lecture Notes in Computer Science*, Volume 2945/2004, 457-467.

Veale, T. and Hao, Y. (2007a). Making Lexical Ontologies Functional and Context-Sensitive. *In Proc. of the 46th Annual Meeting of the Assoc. of Computational Linguistics.*

Veale, T. and Hao, Y. (2007b). Comprehending and Generating Apt Metaphors: A Web-driven, Case-based Approach to Figurative Language. *In Proc. of the 22nd AAAI Conf. on A.I.* Vancouver, Canada.

Veale, T. (2011). Creative Language Retrieval: A Robust Hybrid of Information Retrieval and Linguistic Creativity. Proceedings of ACL'2011, the 49th Annual Meeting of the Association of Computational Linguistics. June 2011.

Veale, T. and Li, G. (2011). Creative Introspection and Knowledge Acquisition: Learning about the world thru introspective questions and exploratory metaphors. In Proc. of the 25th AAAI Conf. of the Assoc. for Advancement of A.I., San Francisco.

Veale, T. and Li, G. (2013). Creating Similarity: Lateral Thinking for Vertical Similarity Judgments. *In Proceedings of ACL 2013, the 51st Annual Meeting of the Association for Computational Linguistics,* Sofia, Bulgaria, August 2013.

Veale, T. (2013). A Service-Oriented Architecture for Computational Creativity. *Journal of Computing Science and Engineering, 7(3):159-167.*

Voorhees, E. M. (1998). Using WordNet for text retrieval. *WordNet, An Electronic Lexical Database*, 285–303. The MIT Press.

Way, E. C. (1991). Knowledge Representation and Metaphor. *Studies in Cognitive systems*. Holland: Kluwer.

Wilks, Y. (1978). Making Preferences More Active, *Artificial Intelligence* 11.