Less Rhyme, More Reason:

Knowledge-based Poetry Generation with Feeling, Insight and Wit

Tony Veale
Web Science & Technology Division, KAIST / School of Computer Science and Informatics, UCD
Korean Advanced Institute of Science & Technology, South Korea / University College Dublin, Ireland.
Tony.Veale@gmail.com

Abstract

Linguistic creativity is a marriage of form and content in which each works together to convey our meanings with concision, resonance and wit. Though form clearly influences and shapes our content, the most deft formal trickery cannot compensate for a lack of real insight. Before computers can be truly creative with language, we must first imbue them with the ability to formulate meanings that are worthy of creative expression. This is especially true of computer-generated poetry. If readers are to recognize a poetic turn-of-phrase as more than a superficial manipulation of words, they must perceive and connect with the meanings and the intent behind the words. So it is not enough for a computer to *merely generate* poem-shaped texts; poems must be driven by conceits that build an affective worldview. This paper describes a conceit-driven approach to computational poetry, in which metaphors and blends are generated for a given topic and affective slant. Subtle inferences drawn from these metaphors and blends can then drive the process of poetry generation. In the same vein, we consider the problem of generating witty insights from the banal truisms of common-sense knowledge bases.

Ode to a Keatsian Turn

Poetic licence is much more than a *licence to frill*. Indeed, it is not so much a licence as a *contract*, one that allows a speaker to subvert the norms of both language and nature in exchange for communicating real insights about some relevant state of affairs. Of course, poetry has norms and conventions of its own, and these lend poems a range of recognizably "poetic" formal characteristics. When used effectively, formal devices such as alliteration, rhyme and cadence can mold our meanings into resonant and incisive forms. However, even the most poetic devices are just empty frills when used only to disguise the absence of real insight. Computer models of poem generation must model more than the frills of poetry, and must instead make these formal devices serve the larger goal of meaning creation.

Nonetheless, is often said that we "eat with our eyes", so that the stylish presentation of food can subtly influence our sense of taste. So it is with poetry: a pleasing form can do more than enhance our recall and comprehension of a meaning – it can also suggest a lasting and profound truth. Experiments by McGlone & Tofighbakhsh (1999, 2000) lend empirical support to this so-called *Keats heuristic*, the intuitive belief – named for Keats' memorable line "Beauty is truth, truth beauty" – that a meaning which is rendered in an aesthetically-pleasing form is much more likely to be perceived as truthful than if it is rendered in a less poetic form. McGlone & Tofighbakhsh demonstrated this effect by searching a book of proverbs for uncommon aphorisms with internal rhyme – such as "*woes unite foes*" – and by using synonym substitution to generate non-rhyming (and thus less poetic) variants such as "*troubles unite enemies*". While no significant differences were observed in subjects' ease of comprehension for rhyming/non-rhyming forms, subjects did show a marked tendency to view the rhyming variants as more truthful expressions of the human condition than the corresponding non-rhyming forms.

So a well-polished poetic form can lend even a modestly interesting observation the lustre of a profound insight. An automated approach to poetry generation can exploit this symbiosis of form and content in a number of useful ways. It might harvest interesting perspectives on a given topic from a text corpus, or it might search its stores of common-sense knowledge for modest insights to render in immodest poetic forms. We describe here a system that combines both of these approaches for meaningful poetry generation.

As shown in the sections to follow, this system – named *Stereotrope* – uses corpus analysis to generate affective metaphors for a topic on which it is asked to wax poetic. *Stereotrope* can be asked to view a topic from a particular affective stance (e.g., view *love negatively*) or to elaborate on a familiar metaphor (e.g. *love is a prison*). In doing so, *Stereotrope* takes account of the feelings that different metaphors are likely to engender in an audience. These metaphors are further integrated to yield tight conceptual blends, which may in turn highlight emergent nuances of a viewpoint that are worthy of poetic expression (see Lakoff and Turner, 1989). *Stereotrope* uses a knowledge-base of conceptual norms to anchor its understanding of these metaphors and blends. While these norms are the stuff of banal clichés and stereotypes, such as that *dogs chase cats* and *cops eat donuts*. we also show how *Stereotrope* finds and exploits corpus evidence to recast these banalities as witty, incisive and poetic insights.

Mutual Knowledge: Norms and Stereotypes

Samuel Johnson opined that "Knowledge is of two kinds. We know a subject ourselves, or we know where we can find information upon it." Traditional approaches to the modelling of metaphor and other figurative devices have typically sought to imbue computers with the former (Fass, 1997). More recently, however, the latter kind has gained traction, with the use of the Web and text corpora to source large amounts of shallow knowledge as it is needed (e.g., Veale & Hao 2007a,b; Shutova 2010; Veale & Li, 2011). But the kind of knowledge demanded by knowledge-hungry phenomena such as metaphor and blending is very different to the specialist "book" knowledge so beloved of Johnson. These demand knowledge of the quotidian world that we all tacitly share but rarely articulate in words, not even in the thoughtful definitions of Johnson's dictionary.

Similes open a rare window onto our shared expectations of the world. Thus, the *as-as*-similes "as hot as an oven", "as dry as sand" and "as tough as leather" illuminate the expected properties of these objects, while the *like*-similes "crying like a baby", "singing like an angel" and "swearing like a sailor" reflect intuitons of how these familiar entities are tacitly expected to behave. Veale & Hao (2007a,b) thus harvest large numbers of *as-as*-similes from the Web to build a rich stereotypical model of familiar ideas and their salient properties, while Özbal & Stock (2012) apply a similar approach on a smaller scale using Google's query completion service. Fishelov (1992) argues convincingly that poetic and non-poetic similes are crafted from the same words and ideas. Poetic conceits use familiar ideas in non-obvious combinations, often with the aim of creating semantic tension. The simile-based model used here thus harvests almost 10,000 familiar stereotypes (drawing on a range of ~8,000 features) from both *as-as* and *like*-similes. Poems construct affective conceits, but as shown in Veale (2012b), the features of a stereotype can be affectively partitioned as needed into distinct *pleasant* and *unpleasant* perspectives. We are thus confident that a stereotype-based model of common-sense knowledge is equal to the task of generating and elaborating affective conceits for a poem.

A stereotype-based model of common-sense knowledge requires both features and relations, with the latter showing how stereotypes relate to each other. It is not enough then to know that cops are tough and gritty, or that donuts are sweet and soft; our stereotypes of each should include the cliché that *cops eat donuts*, just as *dogs chew bones and cats cough up furballs*. Following Veale & Li (2011), we acquire inter-stereotype relationships from the Web, not by mining similes but by mining questions. As in Özbal & Stock (2012), we target query completions from a popular search service (Google), which offers a smaller, public proxy for a larger, zealously-guarded search query log. We harvest questions of the form "*Why do Xs <relation> Ys*", and assume that since each relationship is presupposed by the question (so "why do bikers wear leathers" presupposes that everyone knows that *bikers wear leathers*), the triple of subject/relation/object captures a widely-held norm. In this way we harvest over 40,000 such norms from the Web.

Generating Metaphors, N-Gram Style!

The Google n-grams (Brants & Franz, 2006) is a rich source of popular metaphors of the form *Target is Source*, such as "politicians are crooks", "Apple is a cult", "racism is a disease" and "Steve Jobs is a god". Let *src(T)* denote the set of stereotypes that are commonly used to describe a topic T, where commonality is defined as the presence of the corresponding metaphor in the Google n-grams. To find metaphors for proper-named entities, we also analyse n-grams of the form *stereotype First [Middle] Last*, such as "*tyrant* Adolf Hitler" and "*boss* Bill Gates". Thus, e.g.:

src(racism) = {*problem, disease, joke, sin, poison, crime, ideology, weapon*}

src(Hitler) = {*monster, criminal, tyrant, idiot, madman, vegetarian, racist, …*}

Let *typical*(T) denote the set of properties and behaviors harvested for *T* from Web similes (see previous section), and let *srcTypical*(T) denote the aggregate set of properties and behaviors ascribable to T via the metaphors in *src*(T):

$$(1) \quad srcTypical\,(T) = \bigcup_{M \in src(T)} typical(M)$$

We can generate conceits for a topic *T* by considering not just obvious metaphors for *T*, but *metaphors of metaphors*:

$$(2) \quad conceits(T) = src(\mathrm{T}) \cup \bigcup_{M \in src(T)} src(M)$$

The features evoked by the conceit *T as M* are given by:

$$(3) \quad salient\,(T,M) = [srcTypical(\mathrm{T}) \cup typical(\mathrm{T})] \cap [srcTypical(\mathrm{M}) \cup typical(\mathrm{M})]$$

The degree to which a conceit *M* is apt for *T* is given by:

$$(4) \quad aptness(\mathrm{T, M}) = \frac{|salient(T, M) \cap typical(\mathrm{M})|}{|typical(\mathrm{M})|}$$

We should focus on apt conceits $M \in conceits(T)$ where:

$$(5) \quad apt(\mathrm{T, M}) = |salient(T,S) \cap typical(\mathrm{M})| > 0$$

and rank the set of apt conceits by *aptness*, as given in (4).

The set *salient* (*T,M*) identifies the properties / behaviours that are evoked and projected onto *T* when *T* is viewed through the metaphoric lens of *M*. For affective conceits, this set can be partitioned on demand to highlight only the *unpleasant* aspects of the conceit ("you are such a baby!") or only the *pleasant* aspects ("you are my baby!"). Veale & Li (2011) futher show how n-gram evidence can be used to selectively project the salient norms of *M* onto *T*.

Once More With Feeling

Veale (2012b) shows that it is a simple matter to filter a set of stereotypes by affect, to reliably identify the metaphors that impart a mostly positive or negative "spin". But poems are emotion-stirring texts that exploit much more than a crude two-tone polarity. A system like *Stereotrope* should also model the emotions that a metaphorical conceit will stir in a reader. Yet before *Stereotrope* can appreciate the emotions stirred by the properties of a poetic conceit, it must model how properties reinforce and imply each other.

A stereotype is a simplified but coherent representation of a complex real-world phenomenon. So we cannot model stereotypes as simple sets of discrete properties – we must also model how these properties cohere with each other. For example, the property *lush* suggests the properties *green* and *fertile*, while *green* suggests *new* and *fresh*. Let *cohere*(p) denote the set of properties that suggest and reinforce *p*-ness in a stereotye-based description. Thus e.g. *cohere*(*lush*) = {*green, fertile, humid, dense, …*} while *cohere*(*hot*) = {*humid, spicy, sultry, arid, sweaty*, …}. The set of properties that coherently reinforce another property is easily acquired through corpus analysis – we need only look for similes where multiple properties are ascribed to a single topic, as in e.g. "as *hot and humid* as a jungle". To this end, an automatic harvester trawls the Web for instances of the pattern "*as X and Y as*", and assumes for each X and Y pair that Y ∈ *cohere*(X) and X ∈ *cohere*(Y).

Many properties have an emotional resonance, though some evoke more obvious feelings than others. The linguistic mapping from properties to feelings is also more transparent for some property / feeling pairs than others. Consider the property *appalling*, which is stereotypical of *tyrants*: the common linguistic usage "feel appalled by" suggests that an entity with this property is quite likely to make us "feel appalled". Corpus analysis allows a system to learn a mapping from properties to feelings for these obvious cases, by mining instances of the n-gram pattern "feel P+ed by" where P can be mapped to the property of a stereotype via a simple morphology rule. Let *feeling*(p) denote the set of feelings that is learnt in this way for the property *p*. Thus, *feeling*(*disgusting*) = {*feel_disgusted_by*} while *feeling*(*humid*} = *{}*. Indeed, because this approach can only find obvious mappings, *feeling*(p) = *{}* for most *p*.

However, *cohere*(p) can be used to interpolate a range of feelings for almost any property *p*. Let *evoke*(p) denote the set of feelings that are likely to be stirred by a property *p*. We can now interpolate *evoke*(p) as follows:

$$(6) \quad evoke(p) \; = \; feeling(p) \cup \bigcup_{c \,\in\, cohere(p)} feeling(c)$$

So a property *p* also evokes a feeling *f* if *p* suggests another property *c* that evokes *f*. We can predict the range of emotional responses to a stereotype *S* in the same way:

$$(7) \quad evoke(S) \; = \bigcup_{p \,\in\, typical(S)} evoke(p)$$

If *M* is chosen from *conceits*(T) to metaphorically describe *T*, the metaphor *M* is likely to evoke these feelings for *T*:

$$(8) \quad evoke(T, M) \; = \bigcup_{p \,\in\, salient(T,\, M)} evoke(p)$$

For purposes of gradation, *evoke*(p) and *evoke*(S) denote a *bag* of feelings rather than a *set* of feelings. Thus, the more properties of *S* that evoke *f*, the more times that *evoke*(S) will contain *f*, and the more likely it is that the use of *S* as a conceit will stir the feeling *f* in the reader. *Stereotrope* can thus predict that both *feel disgusted by* and *feel thrilled by* are two possible emotional responses to the property *bloody* (or to the stereotype *war*), and also know that the former is by far the more likely response of the two.

The set *evoke*(*T, M*) for the metaphorical conceit *T is M* can serve the goal of poetry generation in different ways. Most obviously, it is a rich source of feelings that can be explicitly mentioned in a poem about *T* (as viewed thru *M*). Alternately, these feelings can be used in a *meta*-text to motivate and explain the viewpoint of the poem. The act of crafting an explanatory text to showcase a poetry system's creative intent is dubbed *framing* in Colton *et al*. (2012). The current system puts the contents of *evoke*(*T, M*) to both of these uses: in the poem itself, it expresses feelings to show its reaction to certain metaphorical properties of *T*; and in an accompanying framing text, it cites these feelings as a rationale for choosing the conceit *T is M*. For example, in a poem based on the conceit *marriage is a prison*, the set *evoke*(*marriage, prison*) contains the feelings *bored_by*, *confined_in, oppressed_by, chilled_by* and *intimidated_by*. The meta-text that frames the resulting poem expresses the following feelings (using simple NL generation schema):

> "*Gruesome marriage and its depressing divorces appall me. I often feel disturbed and shocked by marriage and its twisted rings. Does marriage revolt you?*"

Atoms, Compounds and Conceptual Blends

If linguistic creativity is chemistry with words and ideas, then stereotypes and their typical properties constitute the periodic table of elements that novel reactions are made of. These are the descriptive atoms that poems combine into metaphorical mixtures, as modeled in (1) … (8) above. But poems can also fuse these atoms into nuanced compounds that may subtly suggest more than the sum of their parts.

Consider the poetry-friendly concept *moon*, for which Web similes provide the following descriptive atoms:

typical(moon) = {lambent, white, round, pockmarked, shimmering, airless, silver, bulging, cratered, waning, waxing, spooky, eerie, pale, pallid, deserted, glowing, pretty, shining, expressionless, rising}

Corpus analysis reveals that authors combine atoms such as these in a wide range of resonant compounds. Thus, the Google 2-grams contain such compounds as "pallid glow", "lambent beauty", "silver shine" and "eerie brightness", all of which can be used to good effect in a poem about the moon. Each compound denotes a compound property, and each exhibits the same linguistic structure. So to harvest a very large number of compound properties, we simply scan the Google 2-grams for phrases of the form "ADJ NOUN", where ADJ and NOUN must each denote a property of the same stereotype. While ADJ maps directly to a property, a combination of morphological analysis and dictionary search is needed to map NOUN to its property (e.g. *beauty* → *beautiful*). What results is a large poetic lexicon, one that captures the diverse and sometimes unexpected ways in which the atomic properties of a stereotype can be fused into nuanced carriers of meaning. Compound descriptions denote compound properties, and those that are shared by different stereotypes reflect the poetic ways in which those concepts are alike. For example, *shining beauty* is shared by over 20 stereotypes in our poetic lexicon, describing such entries as *moon*, *star*, *pearl*, *smile, goddess* and *sky*.

A stereotype suggests behaviors as well as properties, and a fusion of both perspective can yield a more nuanced view. The patterns "VERB ADV" and "ADV VERB" are used to harvest all 2-grams where a property expressed as an adverb qualifies a related property expressed as a verb. For example, the Google 2-gram "glow palely" unites the properties *glowing* and *pale* of *moon*, which allows *moon* to be recognized as similar to *candle* and *ghost* because they too can be described by the compound *glow palely*. A *ghost*, in turn, can *noiselessly glide*, as can a *butterfly*, which may *sparkle radiantly* like a *candle* or a *star* or a *sunbeam*. Not every pairing of descriptive atoms will yield a meaningful compound, and it takes common-sense – or a poetic imagination – to sense which pairings will work in a poem. Though an automatic poet is endowed with neither, it can still harvest and re-use the many valid combinations that humans have added to the language trove of the Web.

Poetic allusions anchor a phrase in a vivid stereotype while shrouding its meaning in constructive ambiguity. Why talk of the *pale glow* of the moon when you can allude to its *ghostly glow* instead? The latter does more than evoke the moon's paleness – it attributes this paleness to a supernatural root, and suggests a halo of other qualities such as *haunting*, *spooky*, *chilling* and *sinister*. Stereotypes are dense descriptors, and the use of one to convey a single property like *pale* will subtly suggest other readings and resonances. The phrase "ghostly glow" may thus allude to any corpus-attested compound property that can be forged from the property *glowing* and any other element of the set *typical(ghost)*. Many stereotype nouns have adjectival forms – such as *ghostly* for *ghost*, *freakish* for *freak*, *inky* for *ink* – and these may be used in corpora to qualify the nominal form of a property of that very stereotype, such as *gloom* for *gloomy*, *silence* for *silent, or pallor for pale*. The 2-gram "inky gloom" can thus be understood as an allusion either to the *blackness* or *wetness* of *ink*, so any stereotype that combines the properties *dark* and *wet* (e.g. *oil*, *swamp*, *winter*) or *dark* and *black* (e.g. *crypt*, *cave*, *midnight*) can be poetically described as exhibiting an *inky gloom*.

Let *compounds*(…) denote a function that maps a set of atomic properties such as *shining* and *beautiful* to the set of compound descriptors – such as the compound property *shining beauty* or the compound allusion *ghostly glow* – that can be harvested from the Google 2-grams. It follows that *compounds(typical(S))* denotes the set of corpus-attested compounds that can describe a stereotype *S*, while *compounds(salient(T, M))* denotes the set of compound descriptors that might be used in a poem about *T* to suggest the poetic conceit *T is M*. Since these compounds will fuse atomic elements from the stereotypical representations of both *T* and *M*, *compounds(salient(T, M))* can be viewed as a blend of *T* and *M*. As described in Fauconnier & Turner (2002), and computationally modeled in various ways in Veale & O'Donoghue (2000), Pereira (2007) and Veale & Li (2011), a "blend" is a tight conceptual integration of two or more mental spaces. This integration yields more than a mixture of representational atoms: a conceptual blend often creates emergent elements – new molecules of meaning – that are present in neither of the input representations but which only arise from the fusion of these representations.

How might the representations discussed here give rise to emergent elements? We cannot expect new descriptive atoms to be created by a poetic blend, but we can expect new compounds to emerge from the re-combination of descriptive atoms in the compound descriptors of *T* and *M*. Just as we can expect $compounds(typical(T) \cup typical(M))$ to suggest a wider range of descriptive possibilities than $compounds(typical(T)) \cup compounds(typical(M))$, we say:

$$\text{(9)}\quad emergent(\mathrm{T}, \mathrm{M}) = \{p \in compounds(salient(T, M)) \mid p \notin compounds(typical(T)) \wedge p \notin compounds(typical(M))\}$$

In other words, the compound descriptions that emerge from the blend of *T* and *M* are those that could not have emerged from the properties of *T* alone, or from *M* alone, but can only emerge from the fusion of *T* and *M* together.

Consider the poetic conceit *love is the grave*. The resulting blend – as captured by *compounds(salient(T, M))* – contains a wide variety of compound descriptors. Some of these compounds emerge solely from the concept *grave*, such as *sacred gloom*, *dreary chill* and *blessed stillness*. Many others emerge only from a fusion of *love* and *grave*, such as *romantic stillness*, *sweet silence*, *tender darkness, cold embrace, quiet passion* and *consecrated devotion*. So a poem that uses these phrases to construct an emotional worldview will not only demonstrate an understanding of its topic and its conceit, but will also demonstrate some measure of insight into how one can complement and resonate with the other (e.g., that darkness can be tender, passion can be quiet and silence can be sweet). While the system builds on second-hand insights, insofar as these are ultimately derived from Web corpora, such insights are fragmentary and low-level. It still falls to the system to stitch these into its own emotionally coherent patchwork of poetry. What use is poetry if we or our machines cannot learn from it the wild possibilities of language and life?

Generating Witty Insights from Banal Facts

Insight requires depth. To derive original insights about the topic of a poem, say, of a kind an unbiased audience might consider witty or clever, a system needs more than shallow corpus data; it needs deep knowledge of the real world. It is perhaps ironic then that the last place one is likely to find real insight is in the riches of a structured knowledge base. Common-sense knowledge-bases are especially lacking in insight, since these are designed to contain knowledge that is common to all and questioned by none. Even domain-specific knowledge-bases, rich in specialist knowledge, are designed as repositories of axiomatic truths that will appear self-evident to their intended audience of experts.

Insight is both a process and a product. While insight undoubtedly requires knowledge, it also takes work to craft surprising insights from the unsurprising generalizations that make up the bulk of our conventional knowledge. Though mathematicians occasionally derive surprising theorems from the application of deductive techniques to self-evident axioms, sound reasoning over unsurprising facts will rarely yield surprising conclusions. Yet witty insights are not typically the product of an entirely sound reasoning process. Rather, such insights amuse and provoke via a combination of over-statement, selective use of facts, a mixing of distinct knowledge types, and a clever packaging that makes maximal use of the Keats heuristic. Indeed, as has long been understood by humor theorists, the logic of humorous insight is deeply bound up with the act of framing. The *logical mechanism* of a joke – a kind of pseudological syllogism for producing humorous effects – is responsible for framing a situation in such a way that it gives rise to an unexpected but meaningful incongruity (Attardo & Raskin, 1992; Attardo *et al*., 2002). To craft witty insights from innocuous generalities, a system must draw on an arsenal of such logical mechanisms to frame its observations of the world in appeallingly discordant ways.

Attardo and Raskin view the role of a logical mechanism (LM) as the engine of a joke: each LM provides a different way of bringing together two overlapping scripts that are mutually opposed in some pivotal way. A joke narrative is fully compatible with one of these scripts and only partly compatible with the other, yet it is the partial match that we, as listeners, jump to first to understand the narrative. In a well-structured joke, we only recognize the inadequacy of this partially-apt script when we reach the punchline, at which point we switch our focus to its unlikely alternative. The realization that we can easily duped by appearances, combined with the sense of relief and understanding that this realization can bring, results in the AHA! feeling of insight that often accompanies the HA-HA of a good joke. LMs suited to narrative jokes tend to engineer oppositions between narrative scripts, but for purposes of crafting witty insights in one-line poetic forms, we will view a script as a stereotypical representation of an entity or event. Armed with an arsenal of stereotype "scripts", *Stereotrope* will seek to highlight the tacit opposition between different stereotypes as they typically relate to each other, while also engineering credible oppositions based on corpus evidence.

A sound logical system cannot not brook contradictions. Nonetheless, uncontroversial views can be cleverly framed in such a way that they appear sound *and* contradictory, as when the columnist David Brooks described the Olympics as a "peaceful celebration of our warlike nature". His form has symmetry and cadence, and pithily exploits the Keats heuristic to reconcile two polar opposites, *war* and *peace*. Poetic insights do not aim to create real contradictions, but aim to reveal (and reconcile) the unspoken tensions in familiar ideas and relationships. We have discussed two kinds of stereotypical knowledge in this paper: the property view of a stereotype *S*, as captured in *typical*(*S*), and the relational view, as captured by a set of question-derived generalizations of the form *Xs <relation> Ys*. A blend of both these sources of knowledge can yield emergent oppositions that are not apparent in either source alone.

Consider the normative relation *bows fire arrows*. Bows are stereotypically *curved*, while arrows are stereotypically *straight*, so lurking beneath the surface of this innocuous norm is a semantic opposition that can be foregrounded to poetic effect. The Keats heuristic can be used to package this opposition in a pithy and thought-provoking form: thus compare "*curved bows fire straight arrows*" (so what?) with "*straight arrows do curved bows fire*" (more poetic) and "*the most curved bows fire the straightest arrows*" (most poetic). While this last form is an overly strong claim that is not strictly supported by the stereotype model, it has the sweeping form of a penetrating insight that grabs one's attention. Its pragmatic effect – a key function of poetic insight – is to reconcile two opposites by suggesting that they fill complementary roles. In schematic terms, such insights can be derived from any single norm of the form *Xs <relation> Ys* where X and Y denote stereotypes with salient properties – such as *soft* and *tough, long* and *short* – that can be framed in striking opposition. For instance, the combination of the norm *cops eat donuts* with the clichéd views of cops as *tough* and donuts as *soft* yields the insight "*the toughest cops eat the softest donuts*". As the property *tough* is undermined by the property *soft*, this may be viewed as a playful subversion of the *tough cop* stereotype. The property *toughness* is can be further subverted, with an added suggestion of hypocrisy, by expressing the generalization as a rhetorical question: "*Why do the toughest cops eat the softest donuts*?"

A single norm represents a highly simplified script, so a framing of two norms together often allows for opposition via a conflict of overlapping scripts. Activists, for example, typically engage in tense struggles to achieve their goals. But activists are also known for the slogans they coin and the chants they sing. Most slogans, whether designed to change the law or sell detergent, are catchy and uplifting. These properties and norms can now be framed in poetic opposition: "*The activists that chant the most uplifting slogans suffer through the most depressing struggles*". While the number of insights derivable from single norms is a linear function of the size of the knowledge base, a combinatorial opportunity exists to craft insights from pairs of norms. Thus, "*angels who fight the foulest demons* *play the sweetest harps*", "*surgeons who wield the most hardened blades wear the softest gloves*", and "*celebrities who promote the most reputable charities suffer the sleaziest scandals*" all achieve conflict through norm juxtaposition. Moreover, the order of a juxtaposition – positive before negative or vice versa – can also sway the reader toward a cynical or an optimistic interpretation.

Wit portrays opposition as an inherent part of reality, yet often creates the oppositions that it appears to reconcile. It does so by elevating specifics into generalities, to suggest that opposition is the norm rather than the exception. So rather than rely wholly on stereotypes and their expected properties, *Stereotrope* uses corpus evidence as a proxy imagination to concocts new classes of individuals with interesting and opposable qualities. Consider the Google 2-gram "*short celebrities*", whose frequency and plurality suggests that shortness is a noteworthy (though not typical) property of a significant class of celebrities. *Stereotrope* already possesses the norm that "*celebrities ride in limousines*", as well as a stereotypical expectation that *limousines* are *long*. This juxtaposition of conventions allows it to frame a provocatively sweeping generalization: "*Why do the shortest celebrities ride in the longest limousines*?" While *Stereotrope* has no evidence for this speculative claim, and no real insight into the status-anxiety of the rich but vertically-challenged, such an understanding may follow in time, as deeper and subtler knowledge-bases become available for poetry generation.

Poetic insight often takes the form of sweeping claims that elevate vivid cases into powerful exemplars. Consider how *Stereotrope* uses a mix of n-gram evidence and norms to generate these maxims: "*The most curious scientists achieve the most notable breakthroughs*" and "*The most impartial scientists use the most accurate instruments*". The causal seeds of these insights are mined from the Google n-grams in coordinations such as "*hardest and sharpest*" and "*most curious and most notable*". These n-gram relationships are then be projected onto banal norms – such as *scientists achieve breakthroughs* and *scientists use instruments* – for whose participants these properties are stereotypical (e.g. *scientists* are *curious* and *impartial*, *instruments* are *accurate*, *breakthroughs* are *notable*, etc.).

Such claims can be taken literally, or viewed as vivid allusions to important causal relationships. Indeed, when framed as explicit analogies, the juxtaposition of two such insights can yield unexpected resonances. For example, "*the most trusted celebrities ride in the longest limousines*" and "*the most trusted preachers give the longest sermons*" are both inspired by the 4-gram "*most trusted and longest.*" This common allusion suggests an analogy: "*Just as the most trusted celebrities ride in the longest limousines, the most trusted preachers give the longest sermons*". Though such analogies are driven by superficial similarity, they can still evoke deep resonances for an audience. Perhaps a sermon is a vehicle for a preacher's ego, just as a limousine is an obvious vehicle for a celebrity's ego? Reversing the order of the analogy significantly alters its larger import, suggesting that ostentatious wealth bears a lesson for us all.

Tying it all together in *Stereotrope*

Having created the individual pieces of form and meaning from which a poem might be crafted, it now falls to us to put the pieces together in some coherent form. To recap, we have shown how affective metaphors may be generated for a given topic, by building on popular metaphors for that topic in the Google n-grams; shown how a tight conceptual blend, with emergent compound properties of its own, can be crafted from each of these metaphors; shown how the feelings evoked by these properties may be anticipated by a system; and shown how novel insights can be crafted from a fusion of stereotypical norms and corpus evidence.

We view a poem as a summarization and visualization device that samples the set of properties and feelings that are evoked when a topic *T* is viewed as *M*. Given *T*, an *M* is chosen randomly from *conceits*(T). Each line of the text renders one or more properties in poetic form, using tropes such as simile and hyperbolae. So if *salient(T, M)* contains *hot* and *compounds(salient(T, M))* contains *burn brightly* – for *T=love* and *M=fire*, say – this mix of elements may be rendered as "*No fire is hotter or burns more brightly*". It can also be rendered as an imperative, "*Burn brightly with your hot love*", or a request, "*Let your hot love burn brightly*". The range of tropes is best conveyed with examples, such as this poetic view of *marriage* as 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
O Marriage, you disgust me with your undesirable security

Each poem obeys a semantic grammar, which minimally indicates the trope that should be used for each line. Since the second-line of the grammar asks for an apt <*simile*>, *Stereotrope* constructs one by comparing *marriage* to a *collective*; as the second-last line asks for an apt <*insight*>, one is duly constructed around the Google 4-gram "*most isolated and most difficult*". The grammar may also dictate whether a line is rendered as an assertion, an imperative, a request or a question, and whether it is framed positively or negatively. This grammar need not be a limiting factor, as one can choose randomly from a pool of grammars, or even evolve a new grammar by soliciting user feedback. The key point is the pivotal role of a grammar of tropes in mapping from the properties and feelings of a metaphorical blend to a sequence of poetic renderings of these elements.

Consider this poem, from the metaphor *China is a rival*:

No Rival Is More Bitterly Determined

Inspire me with your determined battle
The most dogged defender scarcely struggles so much
Stir me with your spirited challenge
Let your competitive threat reward me
Was ever a treaty negotiated by a more competitive rival?
You compete with me like a competitively determined athlete
Does any rival test more competitively than this China?
You oppose me with your bitter battle
Can a bitter rival suffer from such sweet jealousies?
O China, you oppress me with your hated fighting

Stereotypes are most eye-catching when subverted, as in the second-last line above. The Google 2-gram "*sweet jealousies*" catches *Stereotrope*'s eye (and ours) because it up-ends the belief that *jealousy* is a *bitter* emotion. This subversion nicely complements the sterotype that *rivals* are *bitter*, allowing *Stereotrope* to impose a thought-provoking opposition onto the banal norm *rivals suffer from jealousy*.

Stereotype emphasises meaning and intent over sound and form, and does not (yet) choose lines for their rhyme or metre. However, given a choice of renderings, it does choose the form that makes best use of the Keats heuristic, by favoring lines with alliteration and internal symmetry

Evaluation

Stereotrope is a knowledge-based approach to poetry, one that crucially relies on three sources of inspiration: a large roster of stereotypes, which maps a slew of familiar ideas to their most salient properties; a large body of normative relationships which relate these stereotypes to each other; and the Google n-grams, a vast body of language snippets. The first two are derived from attested language use on the web, while the third is a reduced view of the linguistic web itself. *Stereotrope* represents approx. 10,000 stereotypes in terms of approx. 75,000 stereotype-to-property mappings, where each of these is supported by a real web simile that attests to the accepted salience of a given property. In addition, *Stereotrope* represents over 50,000 norms, each derived from a presupposition-laden question on the web.

The reliability of *Stereotrope*'s knowledge has been demonstrated in recent studies. Veale (2012a) shows that *Stereotrope*'s simile-derived representations are balanced and unbiased, as the positive/negative affect of a stereotype T can be reliably estimated as a function of the affect of the contents of *typical(T)*. Veale (2012b) further shows that *typical(T)* can be reliably partitioned into sets of positive or negative properties as needed, to reflect an affective "spin" imposed by any given metaphor M. Moreover, Veale (ibid) shows that copula metaphors of the form *T is an M* in the Google n-grams – the source of *srcTypical*(T) – are also broadly consistent with the properties and affective profile of each stereotype T. So in **87%** of cases, one can correctly assign the label *positive* or *negative* to a topic T using only the contents of *srcTypical*(T), provided it is not empty.

Stereotrope derives its appreciation of feelings from its understanding of how one property presupposes another. The intuition that two properties X and Y that are found in the pattern "*as X and Y as*" evoke similar feelings is supported by the strong correlation (**0.7**) observed between the positivity of X and of Y over the many X/Y pairs that are harvested from the web using this acquisition pattern.

The "fact" that *bats lay eggs* can be found over 40,000 times on the web via Google. On closer examination, most matches form part of a larger question, "*do bats lay eggs*?" The question "*why do bats lay eggs?*" has zero matches. So "*Why do*" questions provide an effective superstructure for acquiring normative facts from the web: they identify facts that are commonly presupposed, and thus stereotypical, and clearly mark the start and end of each presupposition. Such questions also yield *useful* facts: Veale & Li (2011) shows that when these facts are treated as features of the stereotypes for which they are presupposed, they provide an excellent basis for classifying different stereotypes into the same ontological categories, as would be predicted by an ontology such as WordNet (Fellbaum, 1998). Moreover, these features can be reliably distributed to close semantic neighbors to overcome the problem of knowledge *sparsity*. Veale & Li demonstrate that the likelihood that a feature of stereotype A can also be assumed of stereotype B is a clear function of the WordNet similarity of A and B. While this is an intuitive finding, it would not hold at all if not for the fact that these features are truly meaningful for A (and B).

The problem posed by "bats lay eggs" is one faced by any system that does not perceive the whole context of an utterance. As such, it is a problem that plagues the use of n-gram models of web content, such as Google's n-grams. *Stereotrope* uses n-grams to suggest insightful connections between two properties or ideas, but if these n-grams are mere noise, not even the Keats heuristic can disguise them as meaningful signals. Our focus is on relational n-grams, of a kind that suggests deep tacit relationships between two concepts. These n-grams obey the pattern "X <rel> Y", where X and Y are adjectives or nouns and <*rel*> is a linking phrase, such as a verb, a preposition, a coordinator, etc. To determine the quality of these n-grams, and to assess the likelihood of extracting genuine relational insights from them, we use this large subset of the Google n-grams as a corpus for estimating the relational similarity of the 353 word pairs in the Finklestein *et al*. (2002) WordSim-353 data set. We estimate the relatedness of two words X and Y as the PMI (pointwise mutual information score) of X and Y, using the relational n-grams as a corpus for occurrence and co-occurrence frequencies of X and Y. A correlation **of 0.61** is observed between these PMI scores and the human ratings reported by Finklestein *et al*. (2002). Though this is not the highest score achieved for this task, it is considerably higher than any than has been reported for approaches that use WordNet alone. The point here is that this relational subset of the Google n-grams offers a reasonably faithful mirror of human intuitions for purposes of recognizing the relatedness of different ideas. We thus believe these n-grams to be a valid source of real insights.

The final arbiters of *Stereotrope*'s poetic insights are the humans who use the system. We offer the various services of *Stereotrope* as a public web service, via this URL:

http://boundinanutshell.com/metaphor-magnet

We hope these services will also allow other researchers to reuse and extend *Stereotrope*'s approaches to metaphor, blending and poetry. Thus, for instance, poetry generators such as that described in Colton *et al*. (2012) – which creates topical poems from fragments of newspapers and tweets – can use *Stereotrope*'s rich inventories of similes, poetic compounds, feelings and allusions in its poetry.

Summary and Conclusions

Poets use the Keats heuristic to distil an amorphous space of feelings and ideas into a concise and memorable form. Poetry thus serves as an ideal tool for summarizing and visualizing the large space of possibilities that is explored whenever we view a familiar topic from a new perspective. In this paper we have modelled poetry as both a product and an expressive tool, one that harnesses the processes of *knowledge acquisition* (via web similes and questions), *ideation* (via metaphor and insight generation), *emotion* (via a mapping of properties to feelings), *integration* (via conceptual blending) and *rendering* (via tropes that map properties and feelings to poetic forms). Each of these processes has been made publicly available as part of a comprehensive web service called *Metaphor Magnet*.

We want our automated poets to be able to formulate real meanings that are worthy of poetic expression, but we also want them to evoke much more than they actually say. The pragmatic import of a creative formulation will always be larger than the system's ability to model it accurately. Yet the human reader has always been an essential part of the poetic process, one that should not be downplayed or overlooked in our desire to produce computational poets that fully understand their own outputs. So for now, though there is much scope, and indeed *need*, for improvement, it is enough to know that an automated poem is anchored in real meanings and intentional metaphors, and to leave certain aspects of creative interpretation to the audience.

Acknowledgements

This research was supported by the WCU (World Class University) program under the National Research Foundation of Korea (Ministry of Education, Science and Technology of Korea, Project no. R31-30007).

References

Attardo, S. and Raskin, V. 1991. Script theory revis(it)ed: joke similarity and joke representational model. *Humor: International Journal of Humor Research,* **4**(3):293-347.

Attardo, S., Hempelmann, C.F. & Di Maio, S. 2002. Script oppositions and logical mechanisms: Modeling incongruities and their resolutions. *Humor: Int. J. of Humor Research,* **15**(1):3-46.

Brants, T. & Franz, A. 2006. Web 1T 5-gram Version 1. *Linguistic Data Consortium*.

Colton, S., Goodwin, J. and Veale, T. 2012. Full-FACE Poetry Generation. *In Proc. of ICCC 2012, the 3rd International Conference on Computational Creativity*. Dublin, Ireland.

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

Fauconnier, G. & Turner, M. 2002. *The Way We Think. Conceptual Blending and the Mind's Hidden Complexities*. Basic Books.

Fellbaum, C. (ed.) 2008. *WordNet: An Electronic Lexical Database*. MIT Press, Cambridge.

Finkelstein, L., Gabrilovich, E., Matias, Y., Rivlin, E., Solan, Z., Wolfman, G. and Ruppin, E. 2002. Placing Search in Context: The concept revisited. *ACM Transactions on Information Systems, 20(1):116-131*.

Lakoff, G. & Turner, M. 1989. *More than cool reason: a field guie to poetic metaphor*. University of Chicago Press.

McGlone, M.S. & Tofighbakhsh, J. 2000. Birds of a feather flock conjointly (?): rhyme as reason in aphorisms. *Psychological Science* **11** (5): 424–428.

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

McGlone, M.S. & Tofighbakhsh, J. 1999. The Keats heuristic: Rhyme as reason in aphorism interpretation. *Poetics* **26**(4):235-44.

Özbal, G. and C. Strapparava. 2012. A computational approach to automatize creative naming. In *Proc. of the 50th annual meeting of the Association of Computational Linguistics*, Jeju, South Korea.

Pereira, F. C. 2007. *Creativity and artificial intelligence: a conceptual blending approach*. Walter de Gruyter.

Shutova, E. 2010. Metaphor Identification Using Verb and Noun Clustering. *In Proceedings of the 23rd International Conference on Computational Linguistics*, 1001-10.

Veale, T. & D. O'Donoghue. 2000. Computation and Blending. *Cognitive Linguistics*, 11(3-4):253-281.

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

Veale T. & Hao, Y. 2007b. Comprehending and generating apt metaphors: a web-driven, case-based approach to figurative language. In *Proceedings of AAAI'2007, the 22nd national conference on Artificial intelligence*, pp.1471-1476.

Veale, T. & Li, G. 2011. Creative Introspection and Knowledge Acquisition: Learning about the world thru introspective questions and exploratory metaphors. In Proc. of AAAI'2011, the 25th Conference of the Association for the Advancement of AI.

Veale, T. 2012a. *Exploding the Creativity Myth: Computational Foundations of Linguistic Creativity*. London: Bloomsbury.

Veale, T. 2012b. A Context-sensitive, Multi-faceted model of Lexico-Conceptual Affect. In *Proc. of the 50th annual meeting of the Association of Computational Linguistics*. Jeju, South Korea.