Ode to a Keatsian Turn: Creating Meaningful and Poetic Instances of Rhetorical Forms

Tony Veale

Abstract Linguistic creativity requires 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 chapter describes a conceit-driven approach to computational poetry, in which metaphorical conceits are generated for a given topic and affective slant. Subtle inferences drawn from these conceits 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.

1 Introduction

Raymond Chandler saw the primary task of the "natural" writer as bridge-building, between "*what one wants to say*" and "*what one knows how to say*" [5]. The scholarly study of how best to bridge one's words and ideas is an ancient one [1], and rhetoricians have systematically identified and classified a wide variety of linguistic forms with which to give our meanings a persuasive force [13]. These rhetorical devices are so effective in the shaping and delivery of well-developed meanings that they can also lend our less substantial thoughts the unmerited appearance of

Tony Veale
School of Computer Science and Informatics, university College Dublin, Belfield D2, Ireland. e-mail: tony.veale@UCD.ie

solidity. This is not always a bad thing: a well-chosen rhetorical form can act as a scaffolding for an undeveloped idea, allowing it to take root and grow during subsequent elaboration. Nonetheless, just as good painters sometimes paint fakes, clever orators sometimes abuse rhetoric, to suggest profundity where there is shallowness, and sincerity where there is indifference.

One such rhetorical device is chiasmus, which takes its name from the cross-shaped Greek letter chi, or 'χ'. The name is apt, for chiasmus is the crossover repetition of words, meanings, images or syntactic structures in a text. It is a much-used device in the texts of the Bible — it is used in both old and new testaments — and in other ancient Hebrew and Greek texts. Indeed, one of the most widely-quoted examples of chiasmus is also nicely self-descriptive: "*Those that are first shall be last and those that are last shall be first*" (Matthew 19:30). Biblical uses of chiasmus combine a profundity of thought with a symmetry of form, and the effectiveness of this balancing act has not been lost on orators throughout the ages. Consider this use of chiasmus by Abraham Lincoln, the 16$^{\text{th}}$ president of the United States, "*I claim not to have controlled events, but confess plainly that events have controlled me*," and this use by its 35$^{\text{th}}$ president, John F. Kennedy, "*Ask not what your country can do for you, but what you can do for your country*". JFK's use of chiasmus at his inauguration in 1961 was both memorable and effective, allowing his words to hold a mirror, structurally speaking, to what he saw as a much-needed shift from selfishness to selflessness in modern times

The crisscross pattern of chiasmus is the linguistic equivalent of a tightly-laced boxing glove. When used effectively, with a substantial meaning to communicate, the surface crossover of linguistic content implies a dove-tailing of ideas at a deeper level. Chiasmus relies on repetition to drive home these ideas, though this duplication of content is not always superficial or obvious. James Joyce, for example, employed chiasmus to lend balance and symmetry to his use of imagery in *Dubliners*, but nowhere is his use of the form as structurally obvious as it is in either the Lincoln or Kennedy examples. The explicit repetition of words is a hallmark of chiasmus, as in this popular slogan of the American gun lobby, "*when guns are outlawed, only outlaws will have guns*" (though even here, the verb "*to outlaw*" is repeated as a noun). But one can implicitly repeat an idea in a chiastic crossover by referring instead to its opposite counterpart, as in these wise words from the Dalai Lama: "*In the practice of tolerance, one's enemy is the best teacher*" (here "*practice*" aligns with "*teacher*" while "*tolerance*" crosses over to "*enemy*").

These are aphoristic uses of the form, in which the chiasmus is designed to be noticed, just as the resulting epigrams and slogans are designed to be remembered. Yet this repetition with crossover can be just as effective even when it is not overtly noticed. Consider this use of chiasmus by the always quotable fashion-designer Karl Lagerfeld: "*Sweatpants are a sign of defeat. You lost control of your life so you bought some sweatpants*". If you feel the need to repeat yourself, as Lagerfeld does here with withering contempt, then repetition with crossover may subtly strengthen the logical force of your argument. Notice how Lagerfeld begins by asserting a causal link from *sweatpants* to *defeat*, and quickly follows this generalization by asserting a causal link in the opposite direction, from a loss of personal control

(defeat again) to the purchase of those very same sweatpants. His use of chiasmus suggests abductive and deductive reasoning, and shows us the same causal link from complementary perspectives, *effect* ← *cause* and *cause* → *effect*. The overriding impression that one is left with is that sweatpants are more than merely indicative of shame and demoralization; viewed through Lagerfeld's gimlet eye, they are one and the same thing.

By drawing our attention to superficial similarities and deep dissimilarities between what is expected and what is real, devices like chiasmus can pack a powerful ironic punch. Yet, though chiasmus offers a convenient vehicle for packing ironic insights into a structurally pleasing form, such devices can often be too easy to use, allowing one to fake the presence of cutting insight with little more than *cut-and-paste*. Consider the following exchange from the 1999 comedy *Mystery Men*, which concerns the misadventures of a group of wannabe superheroes with underwhelming powers. Mr. Furious has anger management issues, while the Sphinx's only power is an ability to torture syntax until it yields an apparent profundity.

> The Sphinx: He who questions training, only trains himself in asking questions. [. . .] Ah yes, work well on your new costumes my friends, for when you care for what is outside, what is inside cares for you. [. . .] Patience, my son. To summon your power for the conflict to come, you must first have power over that which conflicts you.
>
> Mr. Furious: Okay, am I the only one who finds these sayings just a little bit formulaic? "*If you want to push something down, you have to pull it up. If you want to go left, you have to go right.*" It's . . .
>
> The Sphinx: Your temper is very quick, my friend. But until you learn to master your rage . . .
>
> Mr. Furious: . . . your rage will become your master? That's what you were going to say. Right? Right?
>
> The Sphinx: Not necessarily.

That wonderful last line says it all: the Sphinx has hit on a successful formula for *mere generation*, to turn casual utterances into guru-like prognostications. His utterances appear deep, yet they are little more than superficial repetitions with crossover. One can imbue them with real meaning, of course, but it is clear that meaning takes a back seat to surface form in the Sphinx's need to appear wise and all-knowing. We laugh at the Sphinx because his formulaic use of rhetorical devices has made him necessarily predictable. This is the essence of a deterministic formula: it always produces the same outputs for the same inputs, making a weak demurral such as "*Not necessarily*" all the more risible.

Yet all rhetorical devices are formulas of a sort. It is their repeated utility in different contexts that makes them worthy of study by those who want to give their arguments a form that most effectively reflects their meaning. As an orator, the Sphinx is a one-trick pony; he is predictable not because chiasmus is always predictable, but because he is always predictable in his choice of chiasmus. As a device for inverting an opponent's argument, chiasmus has few equals, yet we tire quickly of any device that is used too often and with too little variety. Were the Sphinx to up his game, and use a wider variety of rhetorical forms to better convey an impression of mental agility, we might pay more attention to what he has to say. For the syntactic manip-

ulation of surface forms is actually a reasonable strategy for exploring the world of ideas. Words are often our only handle on subtle feelings and half-formed ideas, and the systematic manipulation of words can be an effective means of navigating the corresponding conceptual spaces (recall Goethe's maxim that "*words are often most useful precisely when ideas fail us*"). By searching for opportunities for chiasmus, the Sphinx is actually employing a simple form of dialectical reasoning. Given a thesis, he fabricates its structural antithesis, and then uses chiasmus to forge a synthesis of the two. The Sphinx is no Hegel, and he is certainly no Kant, but we must assume that he applies some aesthetic and semantic filters to his formulations. For he does not invert everything, but chooses to selectively invert theses whose antitheses appear structurally and semantically sound. A computer that modeled the generative abilities of the Sphinx would almost certainly be accused of mere generation. Yet its creator might validly reply, Sphinx-like, "*Not necessarily.*"

So what might distinguish a computer's best efforts at chiasmus from those of the Sphinx? Well, it would certainly help if it could display an appreciation of the different shades of meaning carried by related forms of the same word-concept. Consider Mae West's chiastic innuendo "*It's not the men in your life that counts, but the life in your men.*" Mae uses deliberate equivocation here, by employing the word "life" in two different senses — "life" as in personal life, and "life" as in zest and vigor. Between these two senses, Mae stakes out a third sense, her *sex "life"*. Equivocation like this is a form of trickery that often produces humor. Consider another humorous example of chiastic equivocation: "*Children in the back seats of cars sometimes cause accidents. Accidents in the back seats of cars sometimes cause children.*" This is more than syntactic manipulation for its own sake. The repeated use of "accidents" in two different senses — car accidents and accidental pregnancies — produces a pithy commentary on life's surprises, and gives the impression that the speaker has peeked behind the curtain of everyday language to glimpse a universal truth. Each of these examples relies on word play, but each also evokes an unspoken meaning that chimes with our experience of the world.

A computer can easily be programmed to scour a large text corpus for reversible chunks of language such as "*hardly working*" and "*working hard*", so as to generate countless examples of chiasmus in the egregious vein of "*working hard or hardly working?*". Yet this would surely be a poor investment of anyone's time. Even a more semantics-savvy generator, one capable of producing the political aphorism "*for society to prosper, prosperity should be socialized*" from the independent text chunks "*society to prosper*", "*prosperity should''* and "*be socialized*", is hardly worth the effort if all it can do is generate one instance of chiasmus after another. As Truman Capote once said of Jack Kerouac on hearing of the latter's frenetic stream-of-consciousness writing method, "*that's not writing, that's typing*." Even if such a system could generate instances of chiasmus of a quality deemed usable by a professional comedian, no professional would ever craft a whole act around a single rhetorical device. Chiasmus, like other conduits for linguistic creativity, should not be viewed as a party-trick. It should not be generated in bulk, nor sold by the yard. Party tricks are the province of chumps like the Sphinx, in whose hands they are glib generators of fakes rather than vehicles of self-expression.

Computer scientists are trained to embrace modularity, so it's tempting to imagine how a standalone chiasmus generator might later unite with generators of other creative forms to yield a comprehensively well-rounded system, such as one for generating poetry on a specific topic. But it is wishful thinking to imagine that a useful standalone generator might come first, rather like believing that five standalone fingers might later glom together to form a working hand. Any act of specific creativity, such as the generation of an apt chiastic form, must serve a larger creative goal that makes it apt — such as the creation of a poem with a specific purpose — and do so within a larger architecture for creativity that harnesses a diversity of knowledge sources. This architecture must coordinate the actions of many different components as they contribute to the same result, and provide a contextual focus for salient feelings, emotions and expectations. We describe such an architecture for poetry generation, one that meaningfully exploits rhetorical forms, in this chapter.

2 Shallow Heuristics and Deep Knowledge

It 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 meaning — it can also suggest a lasting and profound truth. Experiments reported in McGlone and Tofighbakhsh ([14], [15]) lend empirical support to this so-called *Keats heuristic*, the intuitive belief — named for John 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 and 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., to 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 to highlight nuances of a topic that are worthy of poetic expression (see [12]). *Stereotrope* uses a knowledge-base of conceptual norms to anchor its understanding of these metaphors, and though these norms are very much the stuff of banal clichés and stereotypes, such as that *dogs chase cats* and *cops eat donuts*, we show how *Stereotrope* finds and exploits corpus evidence to recast these banalities as witty, incisive and poetic insights.

But *Stereotrope* cannot operate without knowledge. Samuel Johnson famously 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 modeling of metaphor and other figurative devices have typically sought to imbue computers with the former (see [8]). 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. see [20], [21], [18], [22]). But the kind of knowledge demanded by a knowledge-hungry phenomenon such as metaphor is very different to the specialist "book" knowledge so beloved of Johnson. Metaphor demands knowledge of the quotidian world that we all tacitly share but rarely articulate, not even in the thoughtful definitions of Johnson's dictionary.

Fortunately, 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 intuitions of how these familiar entities are tacitly expected to behave. The authors of [20];[21] thus harvest large numbers of as-as-similes from the Web to build a stereotypical model of familiar ideas and their salient properties, while a similar approach is applied (albeit on a smaller scale) by [16] using Google's query completion service. David Fishelov ([11]) 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 stock of almost 8,000 features) from both as-as and like-similes. Poems construct affective conceits, but as shown in [24], 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.

Stereotrope's 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 fur-balls*. Following [22], we acquire inter-stereotype relationships from the Web, not by mining similes but by mining questions. As in [16], 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.

3 Generating Metaphors that are Affective *and* Effective

The Google n-grams ([4]) 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 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 analyze 3/4-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 $src(T)$. Thus:

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

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

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

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

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

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

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

However, we should focus only on apt conceits $M \in conceits(T)$ where:

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

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

The set $salient(T,M)$ identifies the properties and behaviors 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 (e.g. "you are such a baby!") or only the pleasant aspects (e.g. "you are my baby!"). The authors of [22] further show how n-gram evidence can be used to selectively project the salient norms of M onto T.

4 Once More, With *Feeling!*

As shown in [24], 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* must 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 simply model stereotypes as mere 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 stereotype-based description. Thus e.g. $cohere(lush) = \{green, fertile, humid, \ldots\}$ and $cohere(hot) = \{humid, spicy, sultry, arid, \ldots\}$. 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, *Stereotrope* trawls the Web for instances of the pattern "*as X and Y as*", and assumes for each X and Y pair that $Y \in cohere(X)$ and $X \in cohere(Y)$.

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 $feeling(humid) = \{\}$. Naturally, 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 is likely to evoke a feeling f in an audience if p suggests another property c that is known to evoke 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, then 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 $evoke(S)$ each 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*), but will also appreciate 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 the lens of 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 [6]. *Stereotrope* puts the contents of $evoke(T,M)$ to both of these uses: in the poem itself, it expresses feelings as a reaction to the metaphorical properties of T; and in an accompanying framing text, it cites these feelings as a reason 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 poem expresses the following feelings (generated using simple natural-language generation schemas):

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

5 Bridging Worlds With Phrasal 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*) = {*ambent, 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 perspectives 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 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 $compounds(\dots)$ 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 said to sample from the blend of *T and M*. As described in [7], and computationally modeled in various ways in [19], [17] and [22], 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 this fusion of inputs.

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 the set of stereotypes $compounds(typical(T) \cup typical(M))$ to suggest a wider range of descriptive possibilities than $compounds(typical(T)) \cup compounds(typical(M))$, the emergent compound descriptions that arise 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 could only emerge from the fusion of T and M together. Thus,

$$(9) \quad emergent(T,M) = compounds(salient(T,M)) \setminus compounds(typical(T)) \cup compounds(typical(M))$$

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?

6 The Keatsian Lathe: Re-Shaping the Banal as the Poetic

Insight requires depth. To derive original insights about the topic of a poem, 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 pseudo-logical 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 (see [2], [3]). To craft witty insights from innocuous generalities, a system must draw on an arsenal of logical mechanisms to frame its observations of the world in appealingly discordant ways

Attardo and Raskin (see [2], [3]) 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 be 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* seeks 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 brook contradictions. Nonetheless, uncontroversial views can be cleverly framed in such a way that they appear sound and contra-

dictory, 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 chapter: 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 $Xs < \underline{relation} > Ys$. A blend of both of these sources of knowledge can yield emergent oppositions that are not apparent in either 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: 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 $Xs < \underline{relation} > Ys$ 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 cliched 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* 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: "*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 an audience 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 concoct 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 (if 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 as a rhetorical question: "*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 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 also 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 pride? Reversing the order of the analogy significantly alters its larger import, suggesting that ostentatious wealth bears a lesson for us all.

7 Tying it all together: Poetry as a Service

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 as a coherent service. To recap, we have shown how affective metaphors are generated for a given topic, by building

on popular metaphors for that topic in the Google n-grams; 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 thru the lens of M. Given T, an apt 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, for $T = love$ and $M = fire$, since $salient(T,M)$ contains *hot* and the Google n-grams contains the 2-gram "*burn brightly*", this mix of elements may be rendered as "*No fire is hotter or burns more brightly*". It can also be rendered with the imperative "*Burn brightly with your hot love*", or the plea "*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 $< \underline{simile} >$, *Stereotrope* constructs one by comparing marriage to a collective; as the second-last line asks for an apt $< \underline{insight} >$, 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 metaphor interpretation to a sequence of poetic renderings of these elements.

Consider this poem, elaborated around 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 complements the stereotype that rivals are bitter, allowing *Stereotrope* to impose a thought-provoking opposition onto the banal norm *rivals suffer from jealousy*.

Stereotype emphasizes 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.

8 Quality Considerations

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. For instance, [23] 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 $typical(T)$. In addition, [24] 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, [23] also shows that copula metaphors of the form *T is an M* in the Google n-grams — the origins of $srcTypical(T)$ — are broadly consistent with the properties and affective profile of each stereotype T. 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 linked via 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, dubious matches often form part of a larger question such as "*do bats lay eggs?*", while 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: the authors of [22] show 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* ([9]). Moreover, these features can be reliably distributed to close semantic neighbors to overcome the problem of knowledge sparsity. The authors of [22] also 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 for 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 many of 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 albeit tacit relationships between two concepts. These n-grams obey the pattern $X < \underline{relation} > Y$, where X and Y are adjectives or nouns and $< \underline{relation} >$ 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 *WordSim-353* data set [10]. 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 in [10]. Though this is not the highest score achieved for this task, it is considerably higher than any that 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 functionality of *Stereotrope* in the guise of a public Web service, via this URL: `http://boundinanutshell.com/metaphor-magnet-acl`

We expect 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 [6] — 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.

References

1. Aristotle. 1991. *The Art of Rhetoric*. (translated and edited by Hugh Tancred-Jones). Penguin Classics.
2. 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.
3. Attardo, S., Hempelmann, C.F. and Di Maio, S. 2002. Script oppositions and logical mechanisms: Modeling incongruities and their resolutions. *Humor: International Journal of Humor Research*, 15(1):3-46.
4. Brants, T. and Franz, A. 2006. Web 1T 5-gram Version 1. *Linguistic Data Consortium*.
5. Chandler, R. 1953. The Simple Art of Murder. *Pearls Are a Nuisance*, Hamish Hamilton.
6. Colton, S., Goodwin, J. and Veale, T. 2012. Full-FACE Poetry Generation. In *Proceedings of ICCC 2012, the 3rd International Conference on Computational Creativity*, Dublin, Ireland.
7. Fauconnier, G. and Turner, M. 2002. The Way We Think. Conceptual Blending and the Mind's Hidden Complexities. Basic Books.
8. Fass, D. 1997. Processing Metonymy and Metaphor. *Contemporary Studies in Cognitive Science and Technology*. New York: Ablex.
9. Fellbaum, C. (ed.) 1998. *WordNet: An Electronic Lexical Database*. MIT Press, Cambridge.
10. 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.
11. Fishelov, D. 1992. Poetic and Non-Poetic Simile: Structure, Semantics, Rhetoric. *Poetics Today*, 14(1):1-23.
12. Lakoff, G. and Turner, M. 1989. *More than cool reason: a field guide to poetic metaphor*. University of Chicago Press.
13. Leith, S. (2012). *You Talkin' to Me? Rhetoric from Aristotle to Obama*. Profile Books.
14. McGlone, M.S. and Tofighbakhsh, J. 1999. The Keats heuristic: Rhyme as reason in aphorism interpretation. *Poetics*, 26(4):235-44.
15. McGlone, M.S. and Tofighbakhsh, J. 2000. Birds of a feather flock conjointly (?): rhyme as reason in aphorisms. *Psychological Science*, 11 (5): 424-428.
16. 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.
17. Pereira, F. C. 2007. *Creativity and artificial intelligence: a conceptual blending approach*. Walter de Gruyter.
18. Shutova, E. 2010. Metaphor Identification Using Verb and Noun Clustering. In *Proceedings of the 23rd International Conference on Computational Linguistics*, 1001-1010.
19. Veale, T. and D. O'Donoghue. 2000. Computation and Blending. *Cognitive Linguistics*, 11(3-4):253-281.
20. Veale, T. and Hao, Y. 2007. Making Lexical Ontologies Functional and Context-Sensitive. In Proceedings of the 46th Annual Meeting of Association of Computational Linguistics, Prague. Czech Republic.
21. Veale T. and Hao, Y. 2007. 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.
22. Veale, T. and Li, G. 2011. Creative Introspection and Knowledge Acquisition: Learning about the world thru introspective questions and exploratory metaphors. In *Proceedings of AAAI-2011, the 25th Conference of the Association for the Advancement of AI*, San Francisco, USA.
23. Veale, T. 2012. *Exploding the Creativity Myth: Computational Foundations of Linguistic Creativity*. London: Bloomsbury.
24. Veale, T. 2012. A Context-sensitive, Multi-faceted model of Lexico-Conceptual Affect. In *Proceedings of the 50th annual meeting of the Association of Computational Linguistics*. Jeju, South Korea.