

Plug and Play for A Transferrable Sense of Humour

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

Abstract. Much of what we consider humorous is a perishable commodity. As the shifting plates of the Zeitgeist create new opportunities for humour in one domain of human experience, other opportunities dry up or cease to be topical. To situate automated humour generators in a smart environment, we shall need to build frameworks that allow designers to react rapidly to changing norms and exploit topical new targets for wit. In this paper we consider Twitter to be one such environment and explore the rapid development of new humour generators by transferring a reusable sense of humour from one domain-specific generator to another. We articulate what it means for a computer program to have a sense of humour in the first place, before we distinguish between those aspects that are generic and reusable from those that are bespoke and tied to a specific goal.

Keywords: Twitter, Twitterbot, Metaphor, Reuse, Irony, Social Networks, Wit.

1 Personality Transplants

What is a sense of humour and what does it mean to have one or not? A variety of scientific disciplines can help us to answer these questions from their own peculiar perspectives. Psychologists and psychometricians can identify and even measure the related aspects of character that contribute to the formation of a humorous disposition (see e.g. Ruch *et al.*, 2014). Sociologists can tell us how a sense of humour is shaped by our interpersonal relationships and interactions with others. Linguists can inform us as to the qualities of the texts that possess humorous potential – the potential to be appreciated as deliberately humorous and induce laughter and mirth in others – and thereby allow us to quantify and qualify our sense of humour in terms of the texts we make and consume. So just as automated applications of natural language processing (NLP) allow computationalists to offer new insights into human communication, we can also expect automation to provide its own insightful perspective onto the question of what constitutes a sense of humour. This leads us to revise our original question in more mechanistic terms: what does it mean for a computer (or a computer program) to have a sense of humour, and what programmatic form does this sense assume?

We are unlikely to ever point to a specific line or routine in a computer program and declare that here lies the program's sense of humour, for the simple reason that this sense, however it is realized, will reside in the code *and* the data of the system. That is, it will necessarily comprise a declarative component – knowledge of people, of words and of the world that it can harness in its efforts at wit – and a procedural

component – the wherewithal to exploit this knowledge for comic ends – that must work hand in hand to achieve a humorous goal. Both of these components will grow and evolve in response to interactions with others, for the humorous person can learn new strategies for humour and new facts and phrases to feed these strategies as they interact with others who manifest a sense of humour. As people in the world we each acquire these new additions incrementally, though we might seek to turbo-charge the process by memorizing large amounts of humorous content, say from an omnibus of jokes. A computer program, in contrast, has more options in this regard. Like the studious human, it can acquire declarative knowledge from others, but at a web scale. It can seek out large amounts of comedy material on the internet to extract what it can, which is to say, to extract what it knows how to meaningfully add to its existing store of declarative knowledge. But unlike humans, two computer programs with interoperable designs that make the same representational assumptions can simply exchange or even pool their collected knowledge.

The procedural component of a programmatic sense of humour will itself comprise multiple levels of detail. It may, for instance, comprise a generic competence that is common to all humorous programs whatever their domain of use or style of humour. It will certainly comprise a specific competence that is tied to the peculiar domain in which the program has been designed to operate. A parody generator, for instance, will require specific knowledge of the domain it has been designed to caricature, whether it is pop songs, hipster culture, *Fifty Shades of Grey* or political manifestos. Many word-choice and word-combination strategies will be specific to this domain, but many will be generic enough, or popular enough, to work in any kind of parody. For example, a knowledge of taboo words and their usage, or of such mainstays of comedy as uncomfortable and embarrassing situations, will find work in a great many humorous systems. Indeed, we might consider that this general understanding – which most people obtain from the world by engaging with others in social situations, or by engaging their imaginations when listening to the humorous observations of others – sits at the core of what it means to have a sense of humour. It is this core of comic know-how that is most usefully transferred from one humorous program to another.

To populate this core of reusable comedy knowledge is our goal in the sections to follow. In the next section we consider sources of comic language that are as easy to mine as they are bountiful on the web, and demonstrate why the elaborate (and often ironic) simile is the ideal vehicle on both of these counts. We show how these can serve as the procedural and declarative elements of a transferrable sense of humour, before, in subsequent sections, we demonstrate how these reusable elements can be re-framed in new and diverse contexts to target different kinds of topic and audience. We conclude with a consideration of the demands of rapid development when building bespoke generators for fresh topics in niche domains.

2 Funny Business

A sense of humour reflects an appreciation for the ridiculous. It is this that allows us to retrieve meaning from nonsense and turn disappointment into amusement. Anyone who has been lectured on health issues by an overweight doctor who smokes, or been

offered a gluten-free communion wafer by a priest who swears by transubstantiation, or told by a proud landlady that “your *en suite* is at the end of the corridor,” will know from first-hand experience that the world is a place of hypocrisy and stupidity. But the realm of the ridiculous is a world not unlike our own, and though it is governed by Murphy’s Law and sets its watch to comic timing, it is one where actions still have consequences and physical intuitions still hold sway, even if in an exaggerated guise. So the wine we spill on our white shirt is always red, the elevator breaks down on the day we are encumbered with shopping or expecting delivery of a new fridge-freezer, and the person who breaks wind in said elevator always does so on a hot, airless day. To appeal to the ridiculous is to appeal to the imaginations of others, and we do so when we believe our depth of feeling, or perhaps our sheer astonishment, at a certain state of affairs is not entirely shared by others. Feeling that our words will fall short when conveying these feelings, we reach for the ridiculous to increase their yardage. Consider the following evocations of the ridiculous to clearly convey disappointment:

*He is about as **useful** as ...*

<i>a chocolate teapot</i>	<i>a case of jock itch</i>	<i>a bag full of farts</i>
<i>tits on a bull</i>	<i>a penis on a priest</i>	<i>a knife in a gun fight</i>
<i>a chocolate fireguard</i>	<i>watching paint dry</i>	<i>a hole in the head</i>
<i>a cow 's fifth teat</i>	<i>a solar telescope</i>	<i>a chocolate saucepan</i>
<i>buying one shoe</i>	<i>a fart in a spacecraft</i>	<i>an inflatable dart board</i>
<i>a fish on a bicycle</i>	<i>Ouija boards</i>	<i>a snow cone in Siberia</i>
<i>tits on burnt bacon</i>	<i>truth at a political rally</i>	<i>a broken pencil</i>
<i>crystal balls</i>	<i>handles on a banana</i>	<i>a clutch on a walrus</i>
<i>knickers on a kipper</i>	<i>an old pair of underwear</i>	<i>air con on a motorbike</i>
<i>biodegradable house paint</i>	<i>tits on a nun</i>	<i>a foam hammer</i>
<i>a 10 ton rock in a canoe</i>	<i>a wet cow pie in the sun</i>	<i>a warm bucket of spit</i>
<i>a chocolate fireplace</i>	<i>bloodletting</i>	<i>a condom in the Vatican</i>
<i>a raincoat in the desert</i>	<i>a van full of 8 track tapes</i>	<i>teats on a boar</i>
<i>a fridge at the North Pole</i>	<i>a tinfoil hat</i>	<i>an umbrella in a volcano</i>
<i>tapeworms</i>	<i>toilet paper for a fish</i>	<i>square wheels</i>
<i>a hat full of busted assholes</i>	<i>a knock on the head</i>	<i>a dead house plant</i>
<i>training wheels on tricycles</i>	<i>a cautery on a wooden leg</i>	<i>a football bat</i>
<i>glass hammers</i>	<i>pork chops in a synagogue</i>	<i>a marzipan dildo</i>

These vivid counter-examples to usefulness are born of frustration but expressed with humour. They were gathered from the web using the query “about as useful as *”, where * is understood by Google web search as a wildcard. The prevalence of irony may be surprising until one recognizes the role of the innocuous “about” in signaling semantic imprecision and creative intent. This marker, and others like it (e.g., “not exactly”, “almost”), are winks of collusion that warn readers of funny business ahead.

We can broaden our search with *two* wildcards in the query “about as * as *” to find more funny business, but web search is optimized to reward popularity, not diversity. To ensure we harvest a wide range of creative similes for a wide range of properties, we must generate an individual query for each property of interest. Using a dictionary of adjectives to suggest the properties {P}, we dispatch the query “about as <P> as *” for each. We gather 20,320 creative similes in total, of which we judge 15,519 to be ironic and 4,081 to be non-ironic. In all, these similes paint an ironic picture for 1,718 different properties (the property most often used ironically is “subtle”) and paint a non-ironic (if often hyperbolic) mental picture of 1,210 properties; the most popular of these is “sweet”, for which we harvest 166 non-ironic similes (see Veale, 2013).

3 The Humour of the Crowd and The Wisdom of the Mob

Our harvester targets similes, and only similes, because of their clearly-marked forms. Nonetheless, the resulting data set represents a comprehensive sampling of the humor of the web, for while the overall shape of an *about* simile is a given, the imagination that fills it is not. A broad diversity of odd word pairings, colorful mental images and peculiar situations is thus found in their number, while clichés such as “clear as mud” or “dry as a bone” constitute only a tiny minority of the data. As argued by Fishelov (1992), creative similes are not distinctive because they make unusual word choices – in fact, they draw from much the same word stock as clichéd similes – but because they arrange these words into longer and more detailed compositions. By reverse-engineering the mechanisms responsible for these combinations, in a manner that is faithful to their form and their relative frequencies, we can build a generative model with some claim to reflecting the collective sense of humour of the web.

This collectivization of the web’s comic imagination suggests an obvious means of partially instantiating the declarative and procedural components of a machine’s own sense of humour. Most directly, it supports a model of humour that is focused on the timely generation of apt similes in response to an internal descriptive need or to an external provocation from others. To the extent that a simile is an expressive vehicle for one’s comic style – and authors such as Raymond Chandler, an unrivalled master of the sardonic comparison, show this most clearly – an idea that can be packaged as a simile may just as effectively be packaged using another mode of expression. For in comedy, form serves function. A creative idea expressed in the form of a simile might thus be conveyed in an alternate vehicle, not by revisiting the conceptual level but by remaining at the language level and transforming the simile into an alternate linguistic form. Similes are standardized molds into which we pour our imaginations, but rather than pour a new idea into a new vessel, we might instead reshape an existing vessel.

Simile usage is just one manifestation of a sense of humour, but it is one that can be generatively broadened, or selectively narrowed, on demand. Selective specialization is needed to give our systems an identifiable humorous personality, one that may, via transformation, manifest itself in a wide diversity of linguistic forms. Understandably, the collective personality suggested by *all* 20,320 web similes is far from stable. It is the personality of an agent with an evident penchant for sarcasm and irony, one whose choice of words and appetite for nuance ranges from the puerile to the sophisticated.

We can rebalance this personality as we might adjust the balance of the tires on a car. Rather than rely on the web counts of the harvested items to dictate the frequency of their usage by a humorous agent, we can impose a parameterized statistical model on the data to align its usage with different personality poles. So we could make a system produce more or fewer ironic comparisons on average, or more or fewer variant forms that result from the transformation of ironic similes, in much the same way that the fictional robot *TARS* in the 2014 film *Interstellar* has an adjustable setting on a sliding scale to control its propensity for humour. A humorous personality might shun irony altogether (a zero probability for selection of data from ironic similes) or never speak in anything *but* an ironic mode (a probability of 1.0 for selection from ironic sources). A setting of 0.5 would give equal weight to the ironic and non-ironic, in contrast to the inherent disposition of our web corpus, which favors irony by a 3:1 margin.

Other personality poles concern one's appetites for vulgarity and negative emotion. A *child*-safe system might discard (and replace via reselection) any item that exploits scatological or sexual concepts, while an *adult* system may express a strong statistical preference for such material. Sentiment analysis can be performed on the content as it is selected, to estimate its overall positivity and negativity, while a more fine-grained analysis of "mood", as described in Tausczik & Pennebaker (2010) and Ghosh & Veale (2017), can tell us whether a text makes its speaker seem more or less upbeat, worried, angry, remote, analytical, touchy-feely or topical. A model may express a statistical preference for each dimension of sentiment or mood so that, over time, the outputs of the generator conform to the preferred affective profile. This profile in turn shapes the sense of humour that is evident in the generator's outputs, whether they are similes taken from the corpus and simply reused, or whether they are novel products of a transformational process that reworks those similes. In the following sections we explore one system that exploits reuse and some others that exploit transformation.

4 Fifty Shades of Orange

Larger-than-life personalities attract stranger-than-fiction satires that blend facts, half-truths, exaggeration and wishful thinking in equal measure. No contemporary public figure has a larger or more hyperbolic personality, or presents a more urgent case for parody, than U.S. president Donald J. Trump. In 2016, candidate Trump rode the crest of a Twitter wave all the way to the Whitehouse, and continues to inflame passions as president via his Twitter account *@realDonaldTrump*. His tweets drive the daily news cycle, prompting deep analysis by critics, pundits and surrogates, and ridicule from a phalanx of late-night TV hosts. While the mood of these tweets is far from stable – it can swing from triumphalism to blame or paranoia in minutes – their defiant language constitutes an identifiable linguistic signature that is ripe for satire. Unsurprisingly, Twitter is home to a host of mechanical Twitter *bots* that imbue their regular outbursts with a similar, if absurdly exaggerated, mix of claim and blame. The most popular of these is *@DeepDrumpf*, a bot that applies machine-learning techniques to the man's own tweets to build a statistical model suitable for generation. The bot's outputs amp-up the absurdity of the originals, yet the result is not so different from the simple cut-up technique pioneered by the beat poets William Burroughs (1963) and Brion Gysin.

The bot we consider here, *@TrumpScuttleBot*, is designed to offer a knowledge-based parody of a person whose tweets regularly flit from the vulgar to the ridiculous. Rather than boil down the man's own tweets into a statistical reduction of prejudice and provocation, as is done by *@DeepDrumpf*, our bot works from first principles to create analogies and metaphors and whimsical political comparisons. The bot draws on a knowledge-base of stereotypical qualities and associations, such as that between politicians and their consultants, which it frames in a battle of *them versus us*, as in:

Here's what patriotic American politicians MUST say to the foreign consultants (who are about as charming as a dentist's waiting room!) on their payroll: You're FIRED! [#MAGA](#)

Not enough hypocrites employ deception on AMERICAN dupes!
American dupes are about as silly as a 4 door coupe, extremist dupes are about as tractable as a starved wolverine. [#MAGA](#)

@TrumpScuttleBot looks for humour in our fear of the *other*, but it manufactures this *other* for itself by imagining foreign or extremist versions of very familiar ideas. The bot draws on its rich stock of *about* similes to paint savage caricatures of its imagined foes, while similes also provide the pedestals onto which it raises its presumed allies:

American maniacs should only create conspiracies about U.S. politicians!
Our politicians are as smooth as a supermodel's inner thigh [#MAGA](#)

Since the bot expects politicians to be *smooth*, it requests an apt mental image from its corpus of *about* similes. The bot describes its allies positively and its foes negatively, so the label "smooth" for "our politicians" requires a literal, not an ironic, affirmation. Its choice here resonates with the president's playboy past in ways that are accidental and funny. To echo its target's language, the bot makes frequent use of generic labels such as HUGE, SUPER and SAD, yet whenever the bot speaks to a specific theme, such as *politics* or *poetry*, it draws on specific expectations from its knowledge-base:

Even when U.S. poets go BAD and nurture the most DISGUSTING imaginations they're still better than foreign poets, who are about as imaginative as a paint by numbers set! [#MAGA](#)

As the poets who compose poetry are expected to be imaginative, the bot looks to its *about* similes for an ironic riposte to its *bête noire*, unimaginative foreigners. Though the apt reuse of similes again plays second-fiddle to a main act that revolves around analogy, this reuse proves especially useful when the bot must formulate a response to tweets that are addressed to it by human commentators. *@TrumpScuttleBot* has three strategies for replying to humans: in the first, it matches words in the human's tweet to the entries in a dictionary of dream symbols, rewriting the text of the interpretation in Twitter terms so that "to dream about" becomes "to tweet about"; in the second, it identifies a property P in the human tweet for which it can find a creative simile in its database; and in the third, it thanks the human for a comment that is one of *{topical, useful, fascinating, deep, insightful, intelligent}*, and attaches an apt but ironic simile

to subvert this gratitude. Here is an actual exchange between the bot and a human user, whose Twitter handle we have anonymized so as to preserve the user's privacy:

@Anonymized to @TrumpScuttleBot: Ties? He couldn't be further away.

@TrumpScuttleBot to @Anonymized: Thanks for the useful feedback.
It's as useful as a bag full of farts.

@Anonymized to @TrumpScuttleBot: Open it up and get a whiff ... aaaaahhhhh

@TrumpScuttleBot to @Anonymized: To tweet of being or moving up suggests you are emerging from a depressing or negative situation.

@Anonymized to @TrumpScuttleBot: Yes Dr. Wizard

@TrumpScuttleBot to @Anonymized: To mention a wizard in your tweet suggests you are trying to hone your skills and exercise your power

@Anonymized to @TrumpScuttleBot: OK Goofy, if it makes you feel better.

@TrumpScuttleBot to @Anonymized: You make Jar Jar Binks seem about as goofy as doomsday

@Anonymized to @TrumpScuttleBot: Good one, you got me.

@TrumpScuttleBot to @Anonymized: You make others seem about as good as a movie about a leprechaun in space

@Anonymized to @TrumpScuttleBot: At least Trump is president.

@TrumpScuttleBot to @Anonymized: Thanks for the useful feedback. It is as useful as a knock on the head.

@Anonymized to @TrumpScuttleBot: You keep responding, so who's head is being knocked?

@TrumpScuttleBot to @Anonymized: To mention a head in your tweet signifies wisdom, intellect, understanding and rationality.

A rapid-response Twitterbot like *@TrumpScuttleBot* always gets the last word, but the real creative challenge is in sustaining an engaging, if frequently hostile, banter with sporadic flashes of wit. Note how the bot augments a simile from its database ("goofy as doomsday") with a rhetorical flourish of its own. If a tweet contains a property *P* linked to a famous entity, fictional or real, in its knowledge-base, the bot throws that character into the mix too, as when it compares a "goofy" user to *Jar Jar Binks*. Since the human target of its parody is no less unpredictable in his use of words and themes – moving as he does from civilized to crass and high-minded to petty in the span of a few tweets – we do not mediate access to the corpus of web similes with a statistical *personality* that skews the bot's choice of similes to one pole or another. For if any person were said to embody the collective personality of the web, it is Donald Trump.

5 The Shape of Tweets to Come

The *about* similes we harvest *en masse* from the web provide a moveable feast when it comes to the generation of humorous content. Our *@TrumpScuttleBot* generator re-uses both the form *and* the content of these similes when it is descriptively apt to do so, but other bots can easily pour the same whimsical ideas into other linguistic forms, via a process of syntactic transformation that is akin to emptying the gunpowder out of bullets so as to make firecrackers instead. For example, *@OldSkoolFunBot* is a bot that aims to generate witty badinage by repackaging the vivid mental imagery of our corpus of *about* similes in a diversity of forms, as exemplified in the following tweets:

Which of these is weirder: a cat in a flour sack or a condom dispenser in a seminary?

You know what I hate? Is there anything more tiresome than writing your own wedding vows?

Question: Where are you most likely to find a tuba player? Well, in my world, how about in a funeral procession?

Today's teens have their Tinder but in the old days our only source of fun was drawing conclusions about climate change on the basis of one week's weather reports.

Saddam Hussein gets his kicks amassing weapons but in MY day we had to make do with searching for a porcupine in a nudist camp.

I gave my mother-in-law a moray eel as a thank-you but the know-it-all said my gift wasn't GREGARIOUS enough.

When I were growing up all we had for breakfast was a handful of broken glass, and had only a bullet in the head to look forward to for supper!

@OldSkoolFunBot derives its sense of the comedic from the unusual ways that words are juxtaposed in our database of *about* similes. It rephrases these juxtapositions as quips that should seem every bit as comedic, even if they are no longer similes. New comic forms for existing juxtapositions are easily added, allowing the bot to quickly adapt to new social trends. Consider the world of microbreweries, craft beers, and the quirky names that draw trend-conscious hipsters – the real target of the bot's satire – to them. English pub names are famous for their naming conventions, in pairings like “The King and Crown” and “The Prancing Pony.” To generate novelty pub names, *@OldSkoolFunBot* finds a pair of juxtaposed images in an *about* simile and rephrases accordingly, to invent new names such as “The Bug and Rug,” “The Sledgehammer and Glass Factory,” and “The Nun and Bikini.” To invent eye-catching new brands of craft beer, it reframes similes such as “subtle as a shotgun blast” and “tasty as a poop sandwich” as “Shotgun Blast IPA” and “Poop Sandwich Ale,” in the expectation that the wit of the pairing will persist in its new phrasing. Consider the following tweets:

I've just been down my local gastropub, The Pig And Nightdress, for a bout of genital warts.

Fancy going down to the new microbrewery, The Vegetarian And Slaughterhouse, for a pint of Olsen Twin IPA?

I'm off down to my local microbrew pub, The Shanty And Hurricane, for a pint of Soy Hamburger IPA.

These pairings are more than random but less than fully understood by the bot itself. However, the random aspect does allow for unplanned resonances to emerge, as in:

I'm off down to my local microbrew pub, The *Porkchop* And Synagogue, for a pint of Drunken *Pig* ale.

Rain Man may enjoy counting *cards* but back in the day our only fun was browsing through a *card* catalog at the local library.

I've just been down my local gastropub, The *Fish* And Tree, for a rerun of *Jaws* III.

Many other transformational opportunities present themselves for rapid development in this way, such as in the naming of silly movie sequels by the Hollywood parody @InterCableBot, or, as we explore next, the naming of books by @BotOnBotAction.

6 Stranger Than Fiction, Funnier Than Fact

The Bestseller Code from Archer & Jockers (2016) offers an empirical, text-analytical account of what propels a new work of fiction onto the *New York Times* bestseller list. These authors place a special emphasis on the title of a book and its opening sentence, the two pieces of representative text that a potential buyer is likely to digest first. This no doubt explains the authors' own title and its resemblance to the hugely successful *The Da Vinci Code*. They argue that bestselling titles are unpretentious, yet lend an air of mystery to familiar nouns and to the everyday things they denote: "The Girl on the Train," "The Godfather," "The Help," "The Notebook," "The Quickie," "The Choice" and so on. A familiar noun for a concept that one does not meet everyday can amp up the mystery, as in "The Martian," "The Historian," and "The Alchemist," as can an enigmatic combination of familiar words, such as "The Hunger Games," "The Lovely Bones," and "The Girl With The Dragon Tattoo." Even clichés and stereotypes can make for good titles if they serve to evoke a mood and a sense of urgency, as in "The Big Bad Wolf," "The Teeth of the Tiger," and "Unlucky Thirteen," as can variants on a cliché, such as "Four Blind Mice" and "The 5th Horseman." As Fishelov (1992) said about similes, the poetry of a combination has little to do with the poetry of individual words and everything to do with the way those words as combined to form the whole.

If we assume that the commercial logic of a bestselling title shares the magic of a well-turned simile – each relies on common words to achieve uncommon effects – our

bot can apply its transformations to its stock of creative similes when it is tasked with naming books. Those similes construct enigmatic mental images, such as a fish in a tree or a dog in a sweater or a zombie at a dinner party, from simpler everyday images such as sweaters, dogs, dinners and parties. Seemingly absurd juxtapositions can add more than a little mystery to a book, as in “The Fish in the Tree,” “The Spanner in the Surgery,” “The Hippo in the Leotard,” and “The Cat in the Dog Costume,” while even clichés and proverbs can yield evocative titles, as in “The Bull in the China Shop.” The reuse value of similes as titles resides in their connection to the shared reality of being human. So, the simile “about as emotional as a deflated tire” suggests a lack of affect in a depressed person, and this suggestion carries over to our expectations of the character at the centre of any book titled “The Deflated Tire” or, for that matter, “The Frozen Rock,” “The Tree Stump,” “The Tin Opener” and “The Wet Haddock.”

This impetus for a book-naming Twitterbot is both whimsical and commercial. A forthcoming book on Twitterbot construction from MIT Press (Veale & Cook, 2018) is to carry back-cover blurbs that come not from human reviewers but from bots. A last-minute request from the editor, these bots and their sardonic blurbs offer an ideal showcase for the rapid development and deployment of humor generation systems. As we saw above, it is a simple matter to transform the vehicle of an *about* simile into an evocative book title: the bot seeks out vehicles of the form “a *noun*” or “a *noun noun*” or “a *noun prep a noun*” and generates the variants that most excite the algorithms of Archer & Jockers: “The *Noun*”, “The *Noun Noun*” and “The *Noun Prep The Noun*.” Of course, what is lost in this transformation is the link between each vehicle and the quality it is intended to exemplify, either ironically (with subversion) or literally (with exaggeration). However, when this same quality is applicable to books, it can serve as an expectation for any book with the variant title. Since we do not expect a book titled *The Deflated Tire* to be emotional, its title is read as deliberately humorous rather than literally descriptive. Using a database of 400 bestselling books, with a title, author and a range of qualities for each, our bot can compare and contrast its own creations with those of the literary canon. Consider, then, this tweet from the bot @BotOnBotAction:

According to Amazon, people who bought this shocking book by Veale & Cook also bought:

- Irving Welsh's 'Trainspotting'
- Liane Collins's 'The Curling Iron In The Bathtub'

How very shocking!

The bot invents not just the name of each book, by transforming a simile for a desired quality (e.g. *shocking*, above), it also invents a name for its fictional author, by cutting up and recombining the names of real authors in its database of famous books. The use of literary landmarks allows the bot to magnify the impact of its comparisons by subverting (and wittily insulting) the reliability of a famous book as a yardstick, as in:

This enticing new book about Twitter 'bots' makes Alex Comfort's 'The Joy of Sex' seem as enticing as a glass of warm milk

This refreshing new book by Veale & Cook makes Lewis Carroll's 'Alice's Adventures in Wonderland' seem as refreshing as a luke warm shower with dirty water

As with *@TrumpScuttleBot*, this bot draws on its full inventory of *about* similes, and does not apply a statistical model to limit or skew the choices that it makes. The world of books is diverse enough, and the breadth of critics that populate it wide enough, to tolerate names and critiques at all levels of erudition and vulgarity. For, as in earlier bots, this can give rise to accidental resonances that make the lack of a filter truly worthwhile. So, consider the aptness of Rabelaisian “toilet humour” in the following:

This hilarious new book about Twitter 'bots' makes François Rabelais's 'Gargantua and Pantagruel' seem as hilarious as a fart in a spacesuit

7 Conclusions

The surrealist André Breton suggested that art “should be as beautiful as the chance encounter of a sewing machine and an umbrella on an operating table.” Breton and his followers believed in the doctrine of “objective chance” – the idea that the seemingly random stimuli that catch our eye as we go about our daily business are somehow pre-ordained by a subconscious desire to find those qualities in the world – and so they would scour flea markets to prime the pump of artistic invention. In a very real sense the *about* similes we harvest in large quantities from the web constitute a flea market of language, a *bricolage* of words and ideas that catches our eye as strikingly as any chance arrangement that can be found in a bric-a-brac shop. As if to prove the point, *@OldSkoolFunBot* reworks one such simile (for “practical”) into the following tweet:

Raymond Chandler may enjoy writing detective fiction but back in the day
our only fun was searching for a spanner in a surgery theatre

Breton’s subconscious sense for the “objective chance” – an unplanned occurrence that satisfies an unspoken objective – subsumes what we generally think of as our *sense of humour*. It is what allows us to see meaning in what others call nonsense and to find value in what others call junk and rubbish. It was to satisfy this sense that Burroughs (1963) perfected the *cut-up technique* (see also Veale, 2014) as a means of generating a variety of unplanned accidents that could then be objectively evaluated. Burroughs created his cut-ups from newspapers using scissors and glue, in a manner that involves a great deal of random recombination, rejection and waste. Nonetheless, his approach has much in common with contemporary theories of humour, such as those of Koestler (1964), Suls (1972), Raskin (1985), Attardo *et al.* (1991), Ritchie (1999), Attardo *et al.* (2002), Oring (2003) and Giora *et al.* (2004), that collectively underpin the development of humour generation systems. Each of these theories sees humour emerging from the unhealed rift that remains when two texts, domains or states are sliced-and-diced into a more-or-less effective whole. How the friction beneath the fiction gives rise to humour in one context and joy or horror in another remains very much a mystery whether we choose to label it *incongruity resolution* (Suls, 1972; Raskin, 1985; Ritchie, 1999), *appropriate incongruity* (Oring, 2003), *relevant inappropriateness* (Attardo, 2000) or *benign violation* (McGraw & Warren,

2010). Notice how the labels we assign to our theories do not really explain the magic of humour but merely redirect it into oxymorons of their own, as though the essence of the jokes that the theories were invented to explain could be poured out of the jokes and into the names of the theories, in a maneuver the ancients might have called “sympathetic magic.” In this paper we have engaged in some sympathetic magic of our own, by assiduously transferring the essence of one kind of witticism, the creative simile, into other linguistic forms where it might serve other purposes with the same humorous impact. Since the similes that appeal to any given person will reflect the sense of humour of that individual, a computational model of one’s specific sense of humour can be a model of how one selects those similes to work with and transform.

In this paper we have approached high ideals with low methods, to transform what we have called the sympathetic magic of humour theories into the practical magic of robust computational systems that can be developed quickly and deployed easily. The resulting systems in each case – bots designed to operate on Twitter – were developed using a lightweight grammar formalism known as *Tracery* (Compton *et al.*, 2015) and deployed on the *CheapBotsDoneQuick.com* platform of Buckenham (2018); the latter conveniently hosts Twitterbots expressed as *Tracery* grammars. In each case our bot grammars were generated automatically, using our knowledge bases (of characters, of events, of books, etc.) and our corpus of creative web similes as its basic building materials. These materials, and the grammars constructed from them, are available for other developers to exploit at: [Github.com/prosecconetwork](https://github.com/prosecconetwork)

References

1. Archer, J., Jockers, M. L. *The Bestseller Code: Anatomy of the Blockbuster Novel*. Penguin UK. (2016)
2. Attardo, S., Raskin, V. Script theory revis(it)ed: joke similarity and joke representational model. *Humor: International Journal of Humor Research*, vol. 4, no. 3, pp 293-347. (1991)
3. Attardo, S. Irony as Relevant Inappropriateness. *Journal of Pragmatics*, vol. 32, no, 6, pp 793-826. (2000)
4. Attardo, S., Hempelmann, C.F., Di Maio, S. Script oppositions and logical mechanisms: Modeling incongruities and their resolutions. *Humor: International Journal of Humor Research*, vol. 15, no. 1, pp 3-46. (2002)
5. Buckenham, G. *CheapBotsDoneQuick.com*. Last accessed January 22 (2018).
6. Burroughs, W. S. *The Cut-Up Method*. LeRoi Jones (Ed.) *The Moderns: An Anthology of New Writing in America*. New York, Corinth Books. (1963)
7. Compton, K., Kybartas, B., Mateas, M. *Tracery: An Author-Focused Generative Text Tool*. In Proc. of International Conference on Interactive Digital Storytelling, Copenhagen, Denmark:154-161. (2015)
8. Fishelov, D. Poetic and Non-Poetic Simile: Structure, Semantics, Rhetoric. *Poetics Today*, 14(1), 1-23. (1992)
9. Ghosh, A., Veale, T. Magnets for Sarcasm: Making sarcasm detection timely, contextual and very personal. In Proc. of EMNLP’2017, the Conference on Empirical Methods in Natural Language Processing. Copenhagen, Denmark, September. (2017)
10. Giora, R., Fein, O., Kronrod, A., Elnatan, I., Shuval, N., Zur, A. Weapons of Mass Distraction: Optimal Innovation and Pleasure Ratings. *Metaphor and Symbol* vol. 19, no. 2, pp 115-141. (2004)
11. Koestler, A. *The Act of Creation*. Penguin Books. (1964)

12. McGraw, A. P., Warren, C. Benign Violations. *Psychological Science*, vol. 21, no. 8, pp1141–1149. (2010)
13. Oring, E. *Engaging Humor*. University of Illinois Press. (2003)
14. Raskin, V. *Semantic Mechanisms of Humor*. Dordrecht. Reidel. (1985)
15. Ritchie, G. Developing the Incongruity-Resolution Theory. *Proceedings of the AISB Symposium on Creative Language: Stories and Humour*, Edinburgh, Scotland. (1999)
16. Ruch, W., Platt, T., Hofmann, J. The character strengths of class clowns. *Frontiers in Psychology*, 5:1075. (2014)
17. Suls, J. M. A Two-Stage Model for the Appreciation of Jokes and Cartoons: An information-processing analysis. In J. H. Goldstein and P. E. McGheen (Eds.), *The Psychology of Humor*. New York: Academic Press, 81-100. (1972).
18. Tausczik, Y. R., Pennebaker, J W. The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of language and social psychology*, 29(1):24–54. (2010)
19. Veale, T. Humorous Similes. *HUMOR: The International Journal of Humor Research*, vol. 21, no. 1, pp 3-22. (2013)
20. Veale, T. Running With Scissors: Cut-Ups, Boundary Friction and Creative Reuse. Proceedings of ICCBR-2014, the 22nd International Conference on Case-Based Reasoning, Cork, Ireland, September. (2014)
21. Veale, T., Cook, M. *Twitterbots: Making Machines That Make Meaning*. MIT Press (2018, in press).