Déjà Vu All Over Again

On the Creative Value of Familiar Elements in the Telling of Original Tales

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

Abstract

Movie studios have compelling reasons to love sequels. Familiar characters from successful films are valuable *properties* that come with large built-in audiences eager to pay for more. That such characters are commodities is beyond dispute, yet they are as much commodites for creative story-telling as for commercial film-making. Familiar characters come with pre-existing audiences and pre-existing audience *expectations*, and writers can exploit the latter to reduce exposition, establish *mise en scène*, create mood or motivate the use of genre tropes. Familiarity can also be abused for comic ends, to create narratives dense in references to other stories, worlds or genres. Post-modern irony thus abounds in stories that combine old characters in new, clever and perhaps even logically impossible ways. In this work we explore the value of a large knowledge-base of familiar characters within the plotting mechanics of the *Scéalextric* system, to quantify the extent to which familiarity can enhance or diminish our enjoyment of machine-crafted stories.

Tarzan vs. IBM

Jean-Luc Godard's first choice for the name of his seminal 1965 Sci-Fi film *Alphaville* was not the name of the film's dystopian cityscape, but *Tarzan Versus IBM*. While those characters are familiar in themselves – clichéd, even – the combination is unexpected, jarring and resonant. Godard's film is ultimately a bravura exercise in semiotics, and he uses his characters as much for their potential to signify as their potential to anchor a plot. So *Tarzan* signifies all that is natural, intuitive, vigorous and rooted in the world, while *IBM* signifies largely the opposite, that is, the mechanical, the logical and the tightly controlled. For Godard these two characters represent the antagonistic forces shaping France in the mid-60s, and so his film sets out to capture the battle between scientific technocracy and romantic freedom. But as his chosen signifiers raised obvious legal issues, Godard eventually chose two other familiar characters to carry his narrative. His *Tarzan*-substitute would be *Lemy Caution*, a grizzled detective in a popular series of hard-boiled novels, and his ersatz *IBM* would be a mix of *Wernher Von Braun*, *John Von Neumann* and *Robert Oppenheimer*, whom he blended into his nefarious scientist '*Nosferatu' Von Braun*.

Godard exploited a strategy that was cheekily articulated in Flann O'Brien's 1939 comic novel *At Swim Two Birds*:

"*Characters should be interchangeable as between one book and another. The entire corpus of existing literature should be regarded as a limbo from which discerning authors could draw their characters as required, creating only when they failed to find a suitable existing puppet. The modern novel should be largely a work of reference. Most authors spend their time saying what has been said before – usually said much better. A wealth of references to existing works would acquaint the reader instantaneously with the nature of each character.*"

William S. Burroughs (1963) famously used the corpus of existing literature as fodder for his *cut-up method*, wherein new texts were formed from old via the slicing, dicing and random re-splicing of text chunks. However, his approach lacks a certain finesse that does not allow creators to focus on niceties such as characterization or plot. What is needed for a story-generation process that respects these notions is a knowledge-base, to turn the *limbo* of existing characters into a well-structured inventory of foibles and affordances.

We use one such knowledge-base in this work, the *NOC Non-Official Characterization List* of Veale (2015a, 2016). The NOC provides copious detail on over 800 characters drawn from diverse genres and contexts that are not limited to the wholly fictional. Comprising 30,000 semantic triples the NOC affords story-tellers a level of lively detail for the construction of high-level scenarios and individual scenes that – we hypothesize – can transform a skeletal plot into a vivid narrative that is at once both novel and familiar. For Giora *et al*. (2004) argue that a complex stimulus – such as a headline, a joke or a story – is *optimally innovative* when it allows readers to contrast a creator's unconventional use of familiar elements, such as words and names, with their more conventional (or "salient") uses. Familiar elements, like familiar faces, always provoke a salient response even if this response is subsequently reassessed as incongruous (Giora, 1997). We use the NOC here to ensure that stories generated by our system are optimal innovations that flit along the thin line dividing the educated from the insolent.

Like *Tarzan* & *IBM*, our pre-existing characters are chosen for this potential to signify larger themes and ideas,

such as the conflict between *rich* & *poor*, *strong* & *weak* or *logical* & *illogical*. Fictional characters are liberally mixed with real or historical figures in pairings such as *Steve Jobs & Leonardo da Vinci*, *Frank Underwood & Richard Nixo*n or *Lex Luthor* and *Donald Trump*. But these characters can only come to life in stories with plots that put them to good use. These plots are crafted by a system named *Scéalextric*, which assembles skeletal plots from *action n-grams (*akin to the *Web n-grams* of Brants & Franz, 2006) that resemble the clickable segments of a toy slot-car racing track (Veale, 2016). '*Scéalextric*' is a portmanteau of '*scéal*', the Irish word for 'story', and *Scalextric*, a brand of toy racing kits. Our focus here is on conceptual integration – in the sense of Fauconnier & Turner (1998;2002) – to blend characters from the NOC with plots built using *Scéalextric*, to show how the former can elevate our appreciation of the latter. Importantly, we make the NOC and *Scéalextric* databases available to the community to support further research.

Mechanical story generation pre-dates the birth of the modern computer, so our story begins in the next section with a discussion of the relevant background to this work. We then present the NOC and its contents in greater detail, before exploring the *Scéalextric* model of plot generation. The NOC is then integrated with *Scéalextric* in a fashion that affords creativity from the highest to the lowest levels of a story, before the fruits of this integration are evaluated. The paper concludes with some indications of future work.

Related Work and Ideas

The 1920s was a fertile period for the structural analysis of narrative, producing analyses that would shape and reflect the future automation of the story creation process. In 1928 Vladimir Propp published his *Morphology of the Folktale*, offering an influential structuralist view into the familiar elements that dictate the structure of novel stories. Propp treated tales as akin to cocktails in a trendy bar; each may follow a different recipe and impart a distinctive taste, yet each is drawn from the same set of familiar ingredients. A folklorist and empiricist, Propp built his morphemic system of recurring story *functions* from a painstaking analysis of a corpus of Russian tales. His work influenced generations of folklorist analysis to come, but it has also found favour with computationalists who seek to model story-telling as an act of combinatorial creativity (see Gervás 2013; Veale, 2014). The character functions of Propp's morphology are not specific characters *per se*, but familiar archetypes of which memorable characters such as *Tarzan, Batman, Neo, Bond, Bourne, Loki, Vader* and so on are vivid instances.

While Propp's system is focused on analysis, 1928 also saw the publication of a more practical, generation-focused structualist approach in William Wallace Cook's *PLOTTO*. Cook was driven not by academic curiosity but by the need to produce novel tales of his own to a punishing schedule, sometimes writing a new book in a week. To systematize the generation of plots for each new tale, Cook compiled a corpus of ~1500 *master plots*, each comprising three parts or *clauses*. These plots are numbered and comprehensively categorized and cross-indexed to allow *PLOTTOist* writers

to quickly flesh a skeletal plot into a fuller outline. Though Cook's tripartite plot skeletons are steeped in melodrama, *Plotto* anticipates much of 60s/70s symbolic AI, especially that of the Schank & Abelson school (1977). Moreover, Cook's use of tripartite plot skeletons that can be simply clicked together still has much practical relevance today.

Yet it was not Cook's *Plotto* that would later find favour with Hollywood studios but Joseph Campbell's Propp-like comparative analysis of the heoric tales of world religions, with his 1949 book *The Hero With a Thousand Faces*. In those tales Campbell saw evidence of a protean *mono-myth* that imposes a single deep-structure on a great many heroic myths: "*A hero ventures forth from the world of common day into a region of supernatural wonder: fabulous forces are encountered and a decisive victory is won: the hero comes back from this mysterious adventure with the power to bestow boons on his fellow man*." Campbell's work has been acknowledged as a major influence in the creation of the *Star Wars* mythos by creator George Lucas, the success of which persuaded film studios to pursue the structuralist ideas of Propp and Campbell. So when Vogler (1984/1998) disseminated a far-reaching memo on Campbell's ideas, the schematic agenda became a commercial imperative.

Most AI systems that produce tales – as described in e.g. Meehan (1981), Turner (1994), Pérez y Pérez & Sharples (2004), Riedl & Young (2010) and Gervás (2013) – rely on abstract schematic structures in the mould of Propp, Cook, Campbell and Vogler: they model story-telling as a process of instantiating a core set of simple, reusable forms in new and diverse ways. Gervás *et al.* (2016) employed Proppian schemas to generate plot elements for musical theatre, the products of which were employed in a commercal musical. Veale (2014, 2016) gave Campbell's notion of the hero's journey a computational form in a generator of story *arcs* that models the transformation of a character over time. An arc specifies the start and end point of a character's journey but does not fully articulate a path between those points. In this paper we aim to meld a character with its arc in a beat-by-beat sequence of specific actions that lays out this path. While this sequence resembles a random walk more than a premeditated plan in the sense of Riedl & Young, the walk is sufficiently constrained to yield a coherent storyline.

Schematic structures offer typed slots that may only be filled by characters of the matching types. Cook's master plots use *A* and *B* to label slots for male and female fillers respectively, while Propp's and Campbell's schemas allow for a richer set of types to fill their slots. But the challenge – and the opportunity – for a story generator is to do more than fill slots with matching fillers. A good story requires a tight conceptual blend of characterization and plot, so that what the audience already knows about a given character can spill out of its local slot to colour the action as a whole. This is a challenge we address in the following sections, to allow the known affordances of a familiar character (what he/she wears, likes, dislikes, does for a living, etc.) to fully inform the instantiation and rendering of a plot action and thereby create a more memorable experience for the reader. We begin by taking inventory of these diverse affordances.

Persons of Interest

Godard reached into a grab-bag of cultural icons to pull out *Tarzan & IBM* as the ideal signifiers for his message. But when these proved legally cumbersome to use, he dipped in for another rummage, to now pull out *Lemy Caution* and *Wernher Von Braun* (not to mention *Robert Oppenheimer, John von Neumann* and *Nosferatu*). In other words, Godard used the body of shared cultural landmarks as a resource to drive his process of combinatorial creativity, in a cycle of selection, juxtaposition, rejection and repeated re-selection. He had his own artistic constraints to satisfy in each cycle: note, for instance, how the pairings of *Tarzan & IBM* and *Caution & Von Braun* each juxtapose a fictional creation with a real historical entity. Like the film as a whole, each pairing has one foot in reality and the other in pure fantasy. Different creators impose different constraints, and so our resource, a database of cultural icons, must be detailed and comprehensive if it is to satisfy many (if not all) of them.

Our NOC list will serve as this comprehensive resource. But what makes someone iconic enough to be worthy of representation in the NOC, and which aspects of this entity should the NOC aim to capture? The qualities that elevate a person into a cultural reference point are precisely those that seem to exist in a concentrated and exemplary form in that one person, yet which are so common in our shared experience as to be predicated of many others besides. Our ambitions for the NOC go beyond the representation of the simple adjectival qualities of an iconic person, and include their many affordances as complex, fully-realized entities in their own world, such as gender (*male* or *female*), locale (e.g. *New York, Tatooine*), style of dress, spouses or lovers, known enemies, apt vehicles, trademark weapons, relevant domains (e.g., *arts, science, politics*), semantic types (e.g., *politician, playboy*), fictive status (*fictional* or *real*), genres (e.g., *science fiction*), creators and screen actors (if fictive), typical activities (e.g., *building casinos, running political campaigns,*), political leanings (*left*, *right* or *moderate*) and group affiliations (e.g. *Tony Stark* belongs to *The Avengers* and *Eliot Ness* belongs to *The Untouchables*; *Darth Vader* belongs to the *Dark Side* and *Donald Trump* to the *GOP*).

The affective content of the NOC is intended to provide '*talking points*' in the sense of Veale & Hao (2008). That is, it codifies the qualities and behaviours that we humans naturally focus upon whenever we talk about a celebrity or compare one person to another. The NOC thus divides the adjectival qualities of each entity into positive and negative talking points. The former are those with a positive lexical affect, such as *resolute*, *wealthy* and *media-savvy*, the latter are those with an obvious negative affect, such as *evil*, *tight-fisted* and *devious*. The NOC provides at least three positive and three negative talking points for each of its 800+ entries, so that stories which are built around these entries can offer more than regurgitated clichés and instead offer nuanced qualities that motivate emergent inferences.

The NOC uses a standard frame format for its contents, allocating one frame per entity with its various slots and fillers as outlined above, and additional frames for those fillers that are themselves worthy of further elaboration.

Thus, for example, the NOC provides additional frames for one's clothing, weapons and vehicles of choice, with slots indicating the affordances of each (e.g. one *stabs* with a knife, one *drives* a Mercedes but *sails* on a yacht, and one wears a hat on one's *head* but wears shoes on one's *feet*). The NOC also associates a character's typical activities with apt locations (e.g. one *shops for shoes* in a *shopping mall* but *devises evil schemes* in an *underground lair*), while the taxonomic categories of each character are also organized into a type hierarchy. Importantly, the NOC is designed to be shared and modified in an open and cumulative fashion. It can be downloaded from *GitHub* as a convenient set of spreadsheets, containing approx. 30,000 semantic triples (see *https://github.com/prosecconetwork/*).

The NOC is sufficiently detailed to offer diverse points of overlap for a broad range of entities, real or fictional and historical or modern. Given any NOC character as a target of analysis, a system can quickly find potential sources of comparison amongst those that share a minimum number of overlapping qualities, such as in *domain* or *positive* and *negative* talking points. Or, as was the case with *Tarzan & IBM*, one can seek out antithetical entities that differ by virtue of opposing qualities (such as *boring & entertaining*, *smug & modest*, etc.) With an apt comparison or contrast in hand, generation then becomes a question of how best to frame the juxtaposition for rhetorical effect. The following rhetorical questions were generated by an autonomous bot named @*MetaphorMagnet* (Veale, 2015b) to squeeze into the 140-character message limit imposed by Twitter's API:

What if #TheEmpireStrikesBack were real? #HillaryClinton could be its #PrincessLeia: driven yet bossy, and controversial too

What if #TheNewTestament was about #AmericanPolitics? #MonicaLewinsky could be its #Lucifer: seductive yet power-hungry, and ruined too.

Like Godard's first choice of *Tarzan & IBM*, each of these pairings has one foot in reality and another in pure fantasy. In addition to the aptness of the pairings, it is this playful blurring of lines that contributes to the wit of each tweet. Indeed, it is the aptness of the comparison that justifies the blurring of the otherwise important boundary between fact and fiction. But one can go further than to say that Hillary and Leia, or Monica and Lucifer, are apt comparisons for one another, and suggest that they are also apt antagonists. After all, compelling stories are often built around conflict, and this conflict is most satisfying when it arises from the opposition of a well-matched protagonist and antagonist. What better way to ensure that two characters are different enough to be mutually antagonistic but similar enough to be well-matched than to require metaphorical equivalence? A story that pairs *Steve Jobs* and *Leonardo da Vinci* may be historically daft but it makes deft figurative sense, since each holds a comparable place in the public imagination: e.g., each is pioneering, artistic and far-seeing. The NOC allows one to find pairings that simultaneously make sense *and* nonsense, for stories to make readers think *and* smile.

Into the 'Woulds'

Our conversations about narrative often resort to metaphor (see e.g. McKee, 2010). Yorke (2013) views story-telling as a walk into a mysterious woods, while in keeping with Campbell's finding that heroic narratives often instantiate the *journey* schema, our most popular story metaphors treat narratives as having many of the properties of physical trajectories, such as pace and direction. We talk of plots as though they really could contain sudden twists and devious turns, of characters that go off the rails, of meandering tales that seem to go nowhere, and of tense, fast-paced stories that hurtle to nail-biting conclusions. As such, it is intuitive to think of plot as the track on which characters move forward or back, cross paths or occasionally collide. This track must be lain by narrative's author, of course, so that characters – and audience members, as fellow travelers – may move along it as the shape of the story dictates. The job of an automatic story-teller then is to lay tracks that can take its characters from a chosen start-state to a fitting end-state and so describe the journey as to reward our interest.

Scalextric model race-tracks aim to capture the drama of a real car race at toy scale, by providing kids with diverse track segments – at 1:500 scale – to build complicated and perhaps even treacherous tracks of long straights, hair-pin bends and tricky chicanes. We aim to do much the same for story-telling systems that use *Scéalextric* for their plots, by providing a diverse array of clickable plot segments. As in *Plotto*, we assume that each segment has three clauses or actions, making each segment an action *3-gram*. Consider, for instance, a plot segment comprising this action 3-gram:

A *flatter* B; B *promote* A; A *disappoint* B

A and B are placeholders that will be filled with specific protagonist & antagonist characters at the rendering stage. As for why A *flatters* B, this must be motivated by another segment that is clicked into place before this one, such as:

A *read about* B; B *impress* A; A *flatter* B

Likewise, what happens after A *disappoints* B may be this:

A *disappoint* B; B *humiliate* A; A *attack* B

Notice how each segment connects to others via a sharing of the first or last action, so the above three segments can be linked together, without repetition, to yield 7 scenes:

A *read about* B; B *impress* A; A *flatter* B; B *promote* A; A *disappoint* B; B *humiliate* A; A *attack* B

Diversity of generation is ensured when a multitude of plot segments, of three actions apiece, are available to a system. At present, *Scéalextric* provides 3000+ plot segments such as the above, collectively using over 800 action verbs. The collected segments form the equivalent of a textual n-gram language model whose outputs are plots, not sentences. By threading segments into a plot graph, a generator need only take a random walk in the graph to generate a logical plot. Before we consider how a system chooses a coherent start and end point for its walk, or how it determines the shape of a resulting plot, we first consider how a plot is rendered.

The above three-segment plot of 7 actions is satisfying in its way, yet it is just a skeleton, not a rendered story. To give these skeletons an idiomatic surface form, we provide one or more idiomatic templates for each of the 800+ verbs in the teller's repertoire. For instance, for *attack* we define:

A attacked B with all its strength. A launched a massive attack on B. A pounced on B. A threw itself into an attack on B. A launched a full-frontal attack on B. A clobbered B.

A mapping from skeletal plot to rendered narrative can be created by choosing randomly from the available idiomatic templates for each of the action verbs in the plot skeleton. We employ a similar mapping from plot verbs to pre-built prologue and epilogue texts. Any action verb that can serve as the first action of a narrative is linked to one or more opening texts, while any verb that can serve as the last action in a narrative is linked to one of more closing texts. For instance, these are the available prologues for *flatter*:

B's ego thrived on flattery. B's ego was a balloon inflated by flattery. B liked to be showered with flattery. B liked to be basted in the compliments of others.

A system can choose from the following epilogues for *kill*:

Well that was one way to deal with B! Well B won't be bouncing back from that in a hurry! So A extracts a biblical justice from B. In effect A went medieval on B's ass.

It follows that a system can begin a story at any plot action for which it possesses at least one prologue text, and end a story at any action that possesses at least one epilogue text. It is not the case then that the system first determines the starting and ending actions of a story and finds a sequence of other actions to coherently link them into a plot. Rather, if it provides apt prologue and epilogue texts for every verb in its inventory (which *Scéalextric* does, for all 800+ verbs) *any* linked sequence of plot n-grams will yield a valid plot.

We now turn the question of the causal connectives that connect successive actions, since the *twistiness* of a plot – which corresponds to the shape of a Scalextric track – is a function of how actions follow or defy causal expectations. The action B *promotes* A follows rather naturally from the prior action A *flatters* B since the goal of flattery is social gain, but the action A *disappoint* B is somewhat surprising once promotion is gained. When these are rendered we can thus expect a "so" connective to link A *flatter* B and *B promote* A and a "but" connective to link B *promote* A and A *disappoint* B. If inappropriate connectives are used a narrative will read as incoherent and confused, while if no connectives are used it may seem linear and uninteresting. Every *Scéalextric* 3-gram thus specifies an apt connective to link the first action to the second and link the second action to the third, in ways that respect a human reader's causal intuitions. A good story will be neither too linear – too many *so*'s – nor too twisty – too many *but*'s – but will contain a balance of both. Moreover, when two *but*'s are used in sequence, the second is rendered as *yet*; when two *so*'s are used in sequence, the second is rendered as *then*.

Tying It All Together With Metaphor

The rendering process is complete once the system chooses characters to fill the *A & B* slots in its idiomatic templates. A generic Aesop-inspired strategy can simply choose two random story-book animals to fill the roles of *A & B*, such as a *bear* and an *eagle*, a *dog* and a *cat,* or a *rat* and a *crow*. The following fully-rendered *Scéalextric* story employs all of the aforementioned steps and resources before choosing two animals at random, a *snake* and a *koala*, for its *A & B*:

0. If anyone was in need of supervision it was the *koala*.

1. So at first, the vigilant *snake* supervised the juvenile *koala*'s every activity.

2. But the *koala* could not reach the bar set by the *snake*.

3. So the *snake* considered the *koala* a loser.

4. Then the *snake* brutally beat the *koala*.

5. So the *koala* attacked the *snake* with all its strength.

6. But the *snake*'s trickery went unnoticed by the *koala*.

7. So the *snake* decorated the walls with the *koala*'s innards.

8. Then the *koala* assiduously curried favor with the *snake*.

9. But in the end the vigilant *snake* turned the juvenile *koala* into an indentured slave.

10. Thereafter the *koala* would wear the chains of a slave, but dreamt of choking the *snake* with those chains.

This 11-scene story is rendered from a plot skeleton of four connected action 3-grams – three actions apiece, with three shared overlapping verbs to connect them – that ekes out this chain of actions, with an added prologue and epilogue:

•—*are_supervised_by*—but→*fail_to_deliver_for*—so→
→*disappoint*—so→*are_beaten_by*—so→*attack*—but→
→*are_tricked_by*—so→*are_defeated_by*—so→
→*curry_favor_with* —but→ *are_enslaved_by*—•

As numbered above, scene 0 contains a fitting prologue for the opening action *are_supervised_by*, while scene 10 is an epilogue associated with the final action *are_enslaved_by*. Meehan's TALE-SPIN (1981) also used anthropomorphic animal fillers in the Aesopian tradition, yet such child-like conceits are suggestive of a toy world and a toy AI system. To open *Scéalextric* to the world of familiar human faces, we must integrate the NOC into the rendering process, not just by drawing *A & B* from the NOC but by integrating the affordances of the chosen characters into the rendering of actions, to add non-generic touches to an apt *mise en scène*. As noted previously, characters are paired on metaphorical grounds, so that *A & B* reside in different domains, genres or periods yet exhibit strong similarities, as is the case for *Jobs & Leonardo, Mahatma Gandhi & Obiwan Kenobi* or *Cicero & Obama*. Post-modern irony may also be used to assess figurative similarity via the NOC, so that *Lex Luthor*

and *Keysar Söze* are similar by virtue of having actor *Kevin Spacey* portray them both, while *Jor-El* and *Don Corleone* are similar because each was portrayed by Marlon Brando. Conversely, *Christian Bale* and *George Clooney* and *Adam West* are all well-matched as each has portrayed *Batman*.

A well-matched pairing of NOC characters should not be yoked to a plot with a random starting point; rather, the starting action – or at least one action in the initial segment – should befit the semantic types of the two characters. So if character *A* is a businessman but character *B* is a scientist we might expect A to *invest_in* B or B to *impress* A. If *A* is a businessman and *B* is a reporter, then B may *interview* A. We thus provide several thousand story seeds that link two semantic types, as given in the NOC, with a starting action. When the system then seeks a plot to connect its chosen *A & B*, it picks a matching story seed and then seeks any plot in which the seed action is found in the first plot segment. The following *Scéalextric* story pits *Frank Underwood* (of TV's *House of Cards*) against the very real *Richard Nixon*. Since Frank and Richard are both politicians, one story seed that links them both has the action *campaign_against*:

0. Richard Nixon and Frank Underwood were driven by very different political agendas.

1. So at first, Frank campaigned vigorously against Richard.

2. But Richard humiliated Frank by calling the sociopathic and ruthless Frank the Keyser Söze of wielding political power.

3. So the vindictive Frank hated Richard for being jowly, deceitful and secretive.

4. But Richard made a heartful appeal to Frank.

5. Then Frank's heart softened towards Richard.

6. So Frank forged a bond with Richard.

7. Then, Frank loyally sided with Richard in his struggles with John F. Kennedy.

8. But Richard underpaid Frank for his efforts plotting election strategies

9. So in the end the ruthless Frank cheated Richard out of a lot of money.

10. But those who cheat others have one fatal flaw: narcissism; it will be Frank's undoing.

Once again an 11-scene story is generated from a prologue, an epilogue and four plot segments of three actions apiece, where three duplicate actions overlap at the segment joins. The unrendered *Scéalextric* plot skeleton is as follows:

•—*campaign_against* —but→ *are_humiliated_by* —so→
hate —but→ *are_appealed_to_by* —so→ *are_moved_by*
—so→ *identify_with* —so→ *show_loyalty_to* —but→
are_underpaid_by —so→*cheat*—•

Notice the integration of specific NOC affordances in the renderings of 3, 7 and 8. In scene 3 Underwood's contempt for Nixon is vividly rendered with an appeal to the latter's negative talking points. Scene 7, A *shows loyalty to* B, is given a character-specific rendering that shows familiarity with the history of *Richard Nixon*, while the rendering of scene 8, A i*s underpaid by* B, is made specific to *Frank Underwood* by alluding to a typical activity from the NOC. When a rendering anchored in NOC details is available, it is always preferred to a generic idiomatic rendering.

Notice also the NOC-based rendering of scene 2, which does much more than simply integrate specific affordances from the NOC. Rather, following the Hollywood maxim of "*show, don't tell*" this scene employs a novel speech-act that is created on the fly to drive the plot forward. So the plot action A *is humiliated by* B is rendered not as a simple declaration that *Nixon* humiliates *Underwood* (a "tell") but as a speech act that explains how *Underwood* is humiliated (a "show"). The renderer creates these speech acts as they are needed, using the very same capacity for metaphor that allows it to pair *Nixon* with *Underwood* in the first place. But the speech act in scene 2 is not just metaphorical; it is also wryly ironic in its breaking of the "fourth wall." By comparing *Underwood* to the villainous *Keysar Söze* for apt story reasons, the renderer seems to be winking at the reader, for Nixon appears to know that actor Kevin Spacey portrayed both *Frank Underwood* in *House of Cards* and – spoiler alert! – *Keysar Söze* in *The Usual Suspects*. This is knowledge of the larger world that the NOC provides to the renderer, and as such it is grist for any metaphors and speech acts the renderer may need to generate as needed.

Empirical Evaluation

Two related approaches to story-generation were presented in the previous section. The first uses *Scéalextric* to build its plots around two randomly chosen Aesop-like animals, and renders this sequence of plot actions with a generic set of idiomatic mappings. The *koala* and *snake* narrative of the previous section exemplifies this baseline approach. The second approach uses the same plotting mechanism – clicking together overlapping segments to form a coherent whole – but uses the NOC to choose *human* characters, to choose the initial plot segment to suit the semantic types of these characters, and to opportunistically render individual plot actions using familiar associations from the NOC. The *Nixon* vs. *Underwood* story of the last section exemplifies this NOC-based, character-led approach. So each approach uses precisely the same plotting mechanism but employs different means to render its plots as polished surface texts.

We predict that plots rendered with familiar details from the biographies of its characters will be perceived as more vivid, more entertaining and even more dramatic than stories rendered using only generic surface forms. To test this prediction we generate 50 rendered instances of each kind of story and elicit ratings for each from human judges on the crowd-sourcing platform *CrowdFlower.com*. Each instance is generated following the mechansms described in the previous section, while the choice of which instances are shown to the judges is randomly determined. Although many factors influence a reader's enjoyment of a narrative – for example, whether an odious character gets his come-uppance, or whether a virtuous character finds her reward – we expect that these factors will balance themselves out in a random sampling of all the stories that can be generated.

Judges were not informed as to the mechanical provence of the experimental stimuli, but were simply told that each story was harvested from Twitter. 10 ratings were sought for 6 dimensions for each of the 50 NOC-based stories and 50 generic stories: *laughter* (how likely is this story to make someone laugh?); *entertainment* (how entertaining is this story?), *imagination* (does this story show evidence of an active imagination?), *vividness* (how memorable are the elements of this story?), *silliness* (how implausible is this story?) and *drama* (how eventful is this story?). Judges were shown just one story at a time and asked to rate just one dimension of each, on a scale of 1 (low) to 5 (high). A pool of judges was provided by CrowdFlower, allowing 10 ratings per stimulus to be averaged. We did not require all judges to rate all stimuli, so we report no measures of inter-annotator agreement. Note also that judges were not asked to directly rate the "creativity" of any stimulus, as notions of what constitutes creativity, and how to elicit numbers for those notions, vary significantly (see Jordanous, 2012).

So 10 independent human ratings were elicited for the 6 dimensions of the 50 NOC-based stories and the 50 generic stories, as outlined above. Judges were paid a small sum for each of their ratings, and scammers were eliminated in the usual way: a question requiring a non-random answer (e.g. "how many letters are in the word *<W>"*) was used to separate engaged users from disengaged cheats. To ensure that ratings for one dimension of a story did not influence a judge's ratings for other dimensions, judges were presented with just one story – and asked to rate just one dimension of that story – at a time. As shown in Veale & Alnajjar (2016), eliciting multiple ratings for the same stimulus in the same task unit can cause interference in the results, causing one dimension to influence a judge's rating of another. The mean ratings across stories and judges for 6 dimensions of each story type (and *All* taken together) are shown in Table 1. (Standard deviations are in parentheses.)

Table 1. *Mean ratings per dimension for each type of story*

Dimension	NOC-based	Generic	All
Laughter	**3.02** *(1.13)*	**2.33** *(1.13)*	2.67
Entertainment	**3.10** *(1.15)*	**2.91** *(1.12)*	3.00
Imagination	**3.43** *(1.04)*	**3.03** *(1.08)*	3.23
Vividness	**3.48** *(0.97)*	**2.90** *(1.05)*	3.19
Silliness	**3.47** *(1.04)*	**2.98** *(1.2)*	3.22
Drama	**3.53** *(0.98)*	**2.93** *(1.14)*	3.23

So we see significant improvements across all dimensions of evaluation, with the stories generated from metaphoric parirings of NOC characters that integrate familiar aspects of those characters into their renderings outperforming the stories that rely on generic characters and renderings. After Bonferroni correction is applied the improvements (*NOC* over *Generic*) remain significant at the *p* < .001 level for all dimensions except *Entertainment*, for which *p* < .01.

The most dramatic improvement can be seen, fittingly, in the dimension *Drama*. Though NOC-based and generic stories are each rendered upon a plot skeleton that is built in precisely the same way – save for the added caveat that NOC stories require an action in the initial plot segment to reflect the semantic types of the characters involved – the use of familiar characters and their vivid associations lend the actions of the plot a comic and exaggerated quality that appears to enhance the perceived eventfulness of the story. The prior expectations that readers bring to the NOC-based stories appear, in the main, to make actions and plot turns more engaging than when the same actions and turns are woven around the ephemeral characters of a more generic tale. As a finding about story-telling this is very old news, for scholars since Aristotle have recognized the importance of integrating character *and* plot to build a satisfying tale. It is nonetheless a welcome finding in an empirical context that brings complementary large-scale resources together for the purpose of automatically generating more engaging and entertaining tales on an industrial scale. The NOC list, which supports the automated creation of vivid metaphors, and *Scéalextric*, which turns the creation of story plots into a random walk in a strucrured forest of causal possibilities, work well together as a generator of interesting stories that achieve a more perfect union of character *and* plot.

Beyond Textuality: Multi-Modal Renderings

Reiter & Dale (2006) note that the generation of complex natural-language artefacts requires two levels of planning: *macro*-planning (what is it that I want to say?) and *micro*-planning (so how do I go about saying it?). This division of levels is found in our separation of plot generation and the subsequent rendering of this skeletal plot, in which actions are mapped to surface-level idiomatic forms. By defining more idiomatic templates for the 800+ verbs that make up *Scéalextric*'s plot segments, we provide a greater flexibility in rendering, allowing a story-generator to render its plots in more varied ways that read as fresh and unmechanical. To generate stories in German, French, Spanish or Klingon we need only provide the corresponding stock of idiomatic templates for the action verbs of the *Scéalextric* generator. But those alternate idiomatic templates are not restricted to textual encodings of spoken language, and may incorporate – or rely entirely upon – pictorial elements such as *Emoji*.

We can replace animal designators such as *koala* and *dog* with their corresponding Unicode characters when rendering the *A & B* fillers of generic story-lines, making our inventory of story animals co-extensive with that of animal Emoji in the Unicode standard. In this rendering of the action *are_bought_off_by*, in which *B* is a snake and *A* is a pig, each character is easily replaced with an Emoji:

In addition to providing textual idiomatic forms for each of the system's 800+ action verbs, we can also provide Emoji translations for each verb. The *rebus* principle allow us to use Emoji as both *pictograms* (images depicting ideas) and *sound images* (images depicting words that imply sounds) so that the above scene can be rendered entirely in Emoji:

An Emoji mapping for each of *Scéalextric*'s 800+ verbs is engineered manually, as this task requires some ingenuity. As shown above, a full Emoji rendering is presented side-by-side with its comparable text rendering (and a linguistic short-hand in parentheses), to allow readers to familiarize themselves with this new visual idiom at their own pace. Pure Emoji offers remarkable concision, allowing an entire story to be summarized in a single picture-only tweet:

Emoji serve largely at present as visual adornments for our textual renderings, rather like cute story-book illustrations. Though Emoji are not pictograms in a strong sense – they are far too ambiguous to serve reliably in this role – they nonetheless constitute a lexicon of visual ideas that reflects the collected interests of contemporary social media users. We plan to further explore the role of Emoji are proxies for the semantic primitives that comprise the semantic lexicon of a story-telling system, to achieve a stronger integration of plot, character *and* mental image in the tales that we tell.

Conclusions: Once More Unto The Breach

Scéalextric and the *NOC list* were each designed with the express purpose of supporting research in computational creativity that is practical, scalable *and* knowledge-driven. For each owes its genesis to the international student code-camps for which it was first created and from which each

has later grown in scale and complexity. Researchers may access either resource (and related code) in a public Github that is frequently updated: *github.com/proseconetwork*

But *Scéalextric* and the NOC must grow and evolve to remain relevant as comprehensive resources for research. For the NOC this means the inclusion of new cultural figures as they reach iconic status, while for *Scéalextric* this means tackling the various weaknesses of the *plot-as-path* approach as it now stands. For instance, the plotting mechanism currently assumes that each story has just two characters who move through a tale in parallel, whereas Campbell and Propp allow for a retinue of other characters to participate in the action. To address this shortcoming we will take a leaf from Plotto, which assumes that additional characters can be functions of the protagonist *A* (such as *F-A*, *father of A*) or antagonist *B* (such as *S-B*, *spouse of B*). In this way the supporting figures can be woven into the action as they are needed. As *Scéalextric* graduates from juggling *two* balls to juggling *many* at once, it can graduate to telling nuanced stories about real (or at least familiar) human characters of near-human-level complexity.

Acknowledgements

References

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

Burroughs, W. S. (963). The Cut Up Method. LeRoi Jones (ed.), *The Moderns: An Anthology of New Writing in America*. NY: Corinth Books.

Campbell, Joseph. *The Hero with a Thousand Faces*. Princeton: Princeton University Press.

Cook, William Wallace. (1928/1981). William Wallace Cook. *Plotto: The Master Book of All Plots*. Reprinted by Tin House Books (isbn: 9781935639183).

Fauconnier, G. & Turner, M. (1998). Conceptual Integration Networks. *Cognitive Science,* 22(2):133–187.

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

Gervás, P. (2013). Propp's Morphology of the Folk Tale as a Grammar for Generation. In *Proceedings of the 2013 Workshop on Computational Models of Narrative*, Dagstuhl, Germany.

Gervás, P., Hervás, R., León, C. & Gale, C.V. (2016). Annotating Musical Theatre Plots on Narrative Structure and Emotional Content. In *Proc. of the 7th International Workshop on Computational Models of Narrative*, Krakow, Poland.

Giora. R. (1997). Understanding figurative and literal language: The graded salience hypothesis. *Cogn. Linguistics* 8-3:183-206.

Giora, R., Fein, O., Ganzi, J., Levi , N.A. & Sabah, H. (2004). Weapons of Mass Distraction: Optimal Innovation and Pleasure Ratings. *Metaphor and Symbol **19**(2):115-141*.

Jordanous, A. (2012). A standardised procedure for evaluating creative systems: Computational creativity evaluation based on what it is to be creative. *Cognitive Computation* 4.3: 246-279.

McKee, R. (2010). *Story: Style, Structure, Substance, and the Principles of Screenwriting*. New York: Harper-Collins.

Meehan, J. (1981). *TALE-SPIN*. In Shank, R. C. and Riesbeck, C. K., (eds.), *Inside Computer Understanding: Five Programs plus Miniatures*. Hillsdale, NJ: Lawrence Erlbaum.

O'Brien, F. (1939/1998). *At Swim Two Birds*. Dublin, Ireland: Dalkey Archive Press (1998 reissue).

Pérez y Pérez, R. & Sharples, M (2004). Three Computer-Based Models of Storytelling: BRUTUS, MINSTREL and MEXICA. *Knowledge Based Systems Journal,* 17(1):15-29.

Propp, V. (1928/1968). *Morphology of the Folktale*. University of Texas Press (2nd edition; English translation by Laurence Scott).

Reiter, E. & Dale, R. (2006). Building Natural Language Generation Systems (Studies in Natural Language Processing). Cambridge University Press.

Riedl, M. O. and Young, R. M. (2010). Narrative planning: balancing plot and character. *Journal of Artificial Intelligence Research* 39.1, 217-268.

Schank, R. & Abelson, R. (1977). *Scripts, Plans, Goals and Understanding*. Psychology Press: New York, NY.

Turner, S.R. (1994). *The Creative Process: A Computer Model of Storytelling*, Hillsdale, NJ: Lawrence Erlbaum.

Veale, T. & Hao, Y. (2008). Talking Points in Metaphor: A Concise Usage-based Representation for Figurative Processing. In *Proc. of ECAI 2008, the 18th European Conference on Artificial Intelligence*. Patras, Greece.

Veale, T. (2014). Coming Good and Breaking Bad: Generating Transformative Character Arcs For Use in Compelling Stories. In *Proceedings of ICCC-2014, the 5th International Conference on Computational Creativity, Ljubljana, Slovenia, June 2014*.

Veale, T. (2015a). Fighting Words and Antagonistic Worlds. *In the Proceedings of the 3rd Workshop on Metaphor in NLP at the North American ACL (NAACL),* Denver, Colorado, June 5.

Veale, T. (2015b). Unnatural Selection: Seeing Human Intelligence in Artificial Creations. *Journal of General Artificial Intelligence*, 6(1):5-20.

Veale, T. & Alnajjar, K. (2016). Grounded for life: creative symbol-grounding for lexical invention. *Connection Science*, 28(2):139-154.

Veale, T. (2016). A Rap on the Knuckles and a Twist in the Tale: From Tweeting Affective Metaphors to Generating Stories with a Moral. In *Proceedings of the AAAI Spring Symposium on Ethical and Moral Considerations in Non-Human Agents*

Vogler, S. (1984/1998). *The Writer's Journey: Mythic Structure For Writers*. Studio City, CA: Michael Wiese Productions (a book treatment of Vogler's original 7-page memo from 1984).

Yorke J. (2013). *Into the Woods: A Five-Act Journey into Story*. London, UK: Penguin.