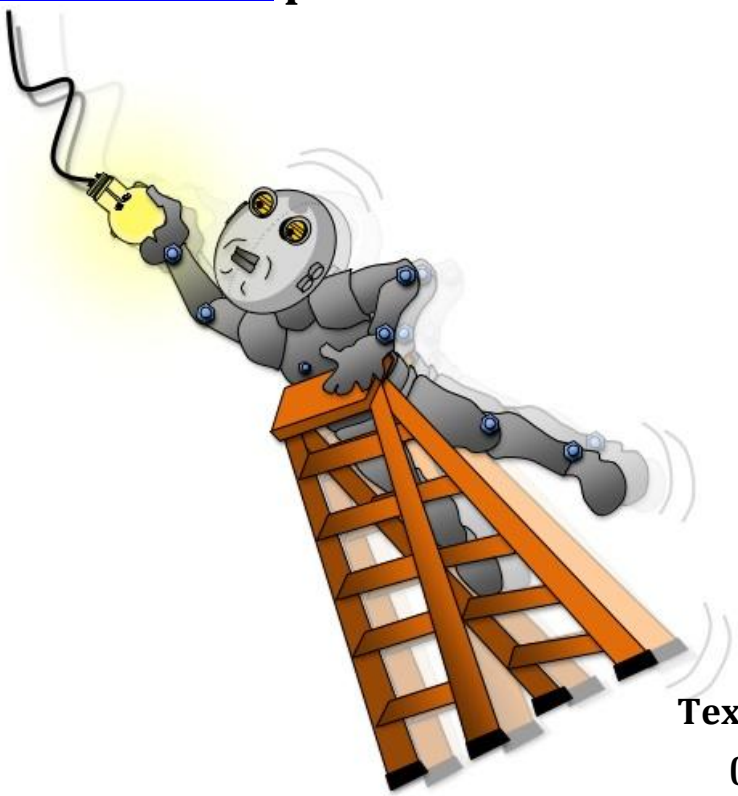
Hand-Made by Machines?

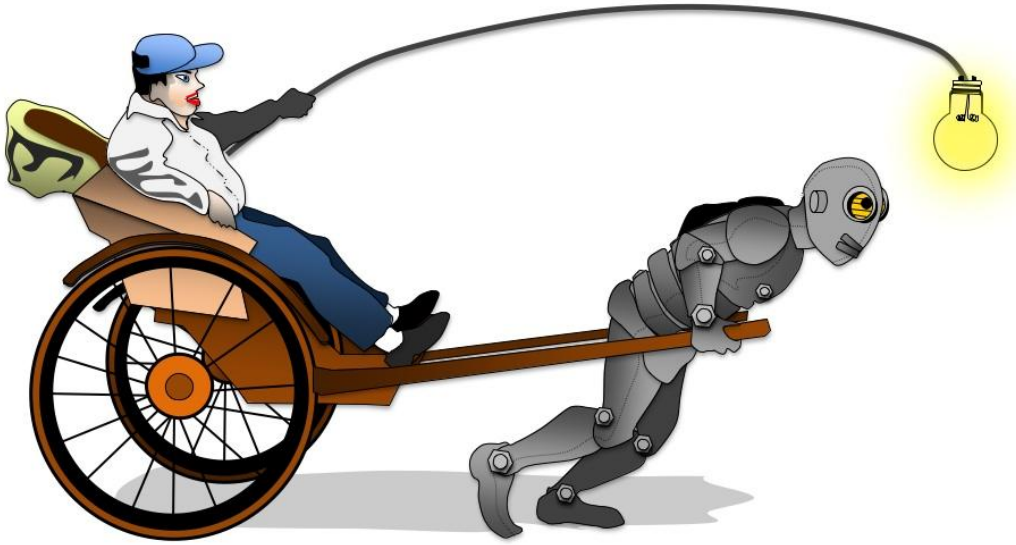
An Illustrated Guide to Creativity in Humans and Computers

A RobotComix.com production of a *Little Computer That Could* book



The author proudly blames his computer for any mistakes that may be found in this publication.

http://www.RobotComix.com



We praise our machines for their consistency, their precision and their speed, but never for their creativity or their insight. And why should we? Creativity is a very human-centric concept, after all, and the popular imagination has good reasons for placing humans and machines at opposite ends of the creativity spectrum. If we want our machines to exhibit the kinds of behavior that humans call *creative*, we shall just have to build (or teach) our machines to think more like we do. So it can come as a surprise then to learn that the converse is also true: humans can often be more creative by learning and choosing to act like machines! *Hand-made By Machines* explores the fascinating ramifications of this controversial claim, by exploring the surprisingly productive middle ground between the worlds of human and machine creativity.

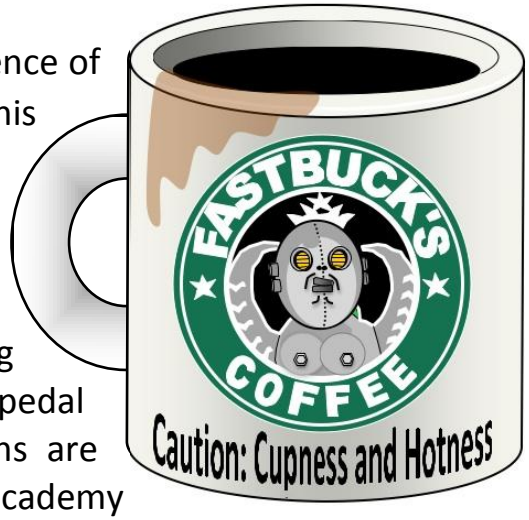

Creative *is* as creative *does*

What is creativity? This simple question is far more likely to elicit an anecdote, an aphorism or a metaphor than it is a straight answer. Creativity is an elusive phenomenon to study, made all the more vexing by our inability to pin it down with a literal definition that does not pass the buck to equally murky near-synonyms like *ingenuity*, *genius*, *originality*, *novelty*, *insight* or *innovation*. It is very tempting to imagine that the idea of creativity as a coherent concept is just another bewitching effect of language, of the kind bemoaned by the philosopher **Ludwig Wittgenstein**.

So it may well be the case be that there exists no single mechanism or all-embracing category of creativity, and that all instances of creative behaviour are best corralled into a meaningful synthesis only by a system of family resemblances. This realization has lead researchers to attempt a multi-pronged definition of creativity. **Alan Newell, Cliff Shaw** and **Herb Simon** have suggested four different criteria for labeling a given answer or solution as "creative":

1. THE ANSWER HAS NOVELTY AND USEFULNESS, EITHER FOR AN INDIVIDUAL OR FOR A SOCIETY
2. THE ANSWER DEMANDS WE REJECT IDEAS THAT WE HAD PREVIOUSLY ACCEPTED
3. THE ANSWER RESULTS FROM INTENSE MOTIVATION AND PERSISTENCE
4. THE ANSWER COMES FROM CLARIFYING A PROBLEM THAT WAS ORIGINALLY VAGUE

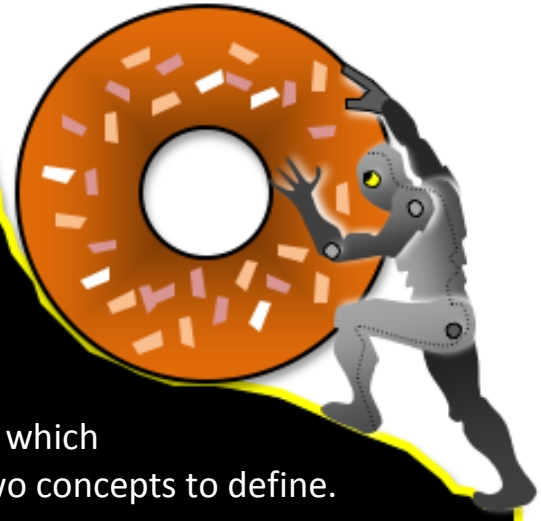
Yet taken in isolation, none of these criteria really hits the mark. Consider **1**: in the course of an average day, most speakers of a language will utter some statements that are so specific to their contexts that they will never have been uttered before; since these utterances serve a real communicative goal, they are both novel and useful, yet they can hardly be labeled "creative". It would seem then that the meaning of *novelty* in **1** must already presuppose some notion of "creative" novelty. Likewise, the rejection of *Habeas Corpus* as a legal principle in a time of war or existential terrorist threat is not as creative as **2** would suggest, for true creativity would find a way of preserving a treasured belief in the situations that test us most. Meanwhile, the intense motivation and persistence of **3** is as much a characteristic of tenacious plodders as it is of creative individuals, while for some problems, pedants are every bit as capable as insightful thinkers of providing the clarity called for in **4**.



Philosophers tell a good story about the difficulty of nailing down the essence of an informal concept with a formal definition. The philosopher **Plato** was debating his theory of abstract ideas – his so-called Platonic forms – with his skeptical rival **Diogenes**. Pointing to his wine cup, Diogenes claimed "*I can see the cup on the table, but not the 'cupness'*". Plato responded with a withering putdown: "*Though you have eyes to see the cup, you lack the intellect to comprehend its 'cupness'*". History was to give Diogenes the last laugh, though. In his goal of systematizing human knowledge, Plato proposed an ontology in which humans were defined as bipedal animals. Since chickens are also bipedal, Plato was careful to add that humans are *featherless* bipeds. Diogenes, who had a flair for the dramatic, strode into Plato's academy with a plucked chicken under his arm. "Here is Plato's man", he jeered, hurling his squawking counter-argument at the startled philosopher. Ever the formalist, Plato made only a minimal effort to repair his ontology: humans would henceforth be defined as *featherless* bipedal animals with *broad flat nails*. Another own goal for formal ontologies?

To see the inadequacy of formal definitions that demarcate categories with necessary and sufficient conditions, consider the knotty problem of defining a **donut**. Simply ponder the elements of *donutness* if you will, and write down a list of the necessary qualities that every donut must possess; from this list, select those that are collectively sufficient for any item to be called a "donut". Now for the fun part: take a field trip into any ***Dunkin' Donuts*** and count how many *non*-donuts fit your definition (bagels, cookies, cheerios?) and how many so-called donuts fall short (donuts without holes? donuts that are baked, not fried? sugarless donuts? square donuts? tofu donuts?). You don't need to be a reincarnated Diogenes to appreciate the difficulties of pinning down a moving target with a static definition, or to see that these difficulties apply whether one is defining crullers or creativity, bagels or brilliance. Creativity itself gnaws on the bonds of all formal definitions, by stretching conventions and subtly shifting our conceptual boundaries.

Food for Thought



We have seen that familiarity is no guarantor of definability, and so even an everyday object like a **DONUT** can be very difficult indeed to nail down with a formal definition that is neither trivially general nor overly specific. The case of **DONUT** is especially interesting because these popular treats are the subject of so much commercial innovation. Contrast the **DONUT** with the **BAGEL**: though both are round floury confections which stereotypically have a rather obvious hole in the middle, **BAGEL** is by far the easier of the two concepts to define.

- Why should this be so? Why do some familiar concepts attract more of our creative energies than others? How can two kinds of objects that share so many physical similarities be subject to such very different development forces?
- What properties of donuts make them so easy to *vary* in innovative ways?
- What properties of bagels make them so *stable*, and so resistant to innovation?
- Identity five other pairs of superficially-similar object-types that exhibit this variability/stability asymmetry.
- Here are some other everyday concepts whose familiarity may prove an impediment to precise definition:

Drug, Game, Pet, Lie, Medicine, Wreck, Athlete, Artist, Friend

- Try to define each of the above concepts in as precise yet *all-embracive* fashion as possible.
- Try to identify counter-examples to your definitions, perhaps by focusing on how the concepts have changed in recent years, and modify your definitions accordingly. How future-proofed can you make your definitions?

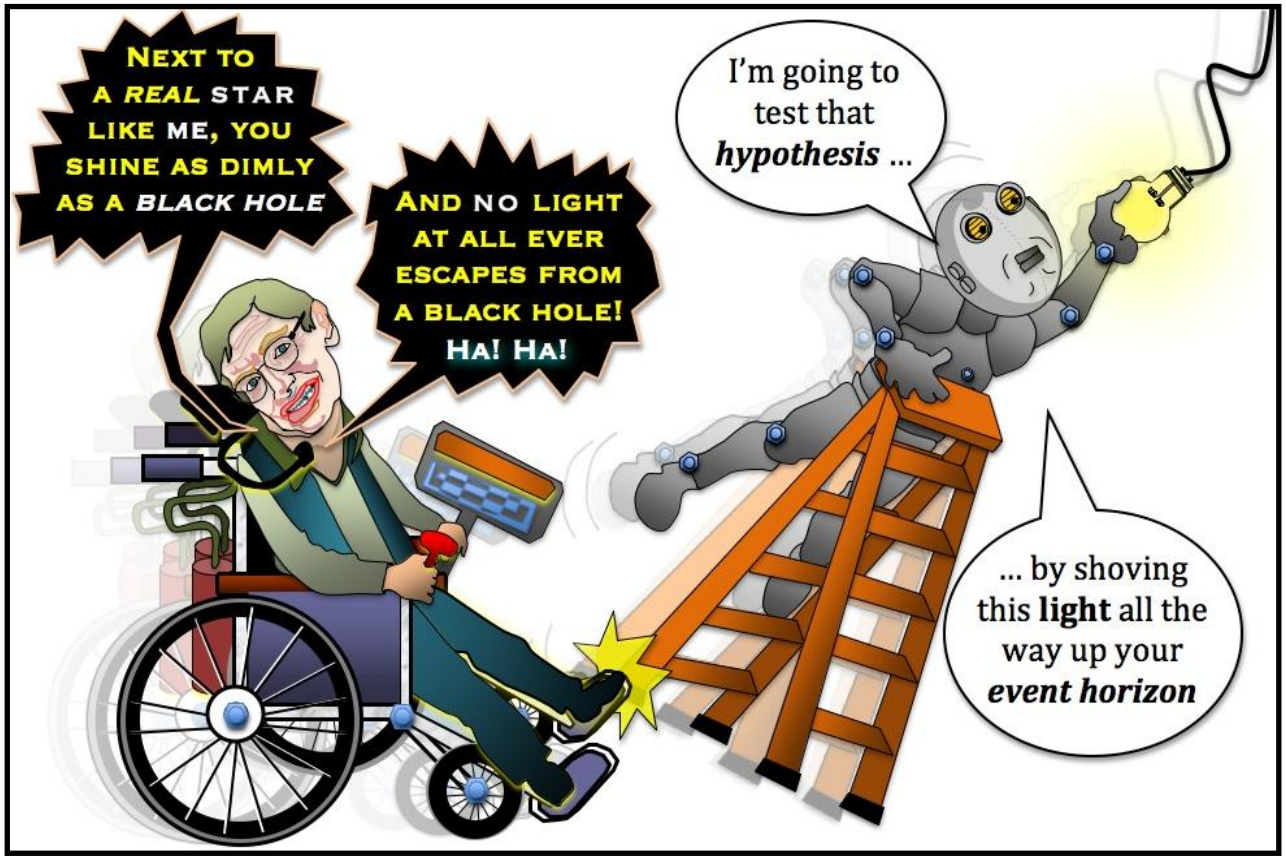
Hint: Social Media sites like FACEBOOK have, of late, weakened the meaning of the word "Friend"

"All science is either physics or stamp-collecting." That's how Nobel-winning physicist **Ernest Rutherford** contrasted the two competing tendencies that shape a scientist's approach to understanding the world. Either you are engaged in a search for the fundamental laws of the universe – in which case you are doing **physics**, whether you know it or not – or you are engaged in an effort to systematically organize the world's diversity into the scientific equivalent of a stamp album. Given the difficulties we have in pinning down a formal definition of creativity, creativity theorists generally belong to the stamp-collecting camp that Rutherford so obviously disdained.

Philosopher **Wilhelm Windelband** was less puckish in drawing this distinction when he introduced the terms **Nomothetic** and **Idiographic** to describe the mindsets of the opposing camps. A **nomothetic** approach seeks to identify the underlying laws that govern all instances of a phenomenon, while an **idiographic** approach instead focuses on the particular qualities of specific instances that make those instances exemplary or unique.

A nomothetic theory of creativity thus seeks to identify the "laws" that govern creative behaviour by deriving strong generalizations from the analysis of large quantities of historical data, while an idiographic approach documents the most interesting qualities of eminent creators and notable instances of creativity, big or small, from **William Harvey**'s discovery of the circulatory system to **Art Fry**'s invention of the ***Post-It***. Both perspectives have their advantages and their adherents in creativity studies, and both together provide a comprehensive view of creativity, from the level of an individual to the level of a society.

Idiographic case-studies pervade the study of creativity, for there are few "laws" or hard generalizations that are immune to falsification by a single specific innovation. Yet strong tendencies *can* be identified in the history of creativity, such as that the most famous creators tend to be the most prolific, generating a disproportionate number of the key papers, artworks or inventions that are most often cited or performed. **Dean Keith Simonton**, a notable instance of a nomothetic creativity theorist, uses ***historiometic* analysis** of archival data to find the patterns that best allow us to meaningfully compare and contrast two creators.



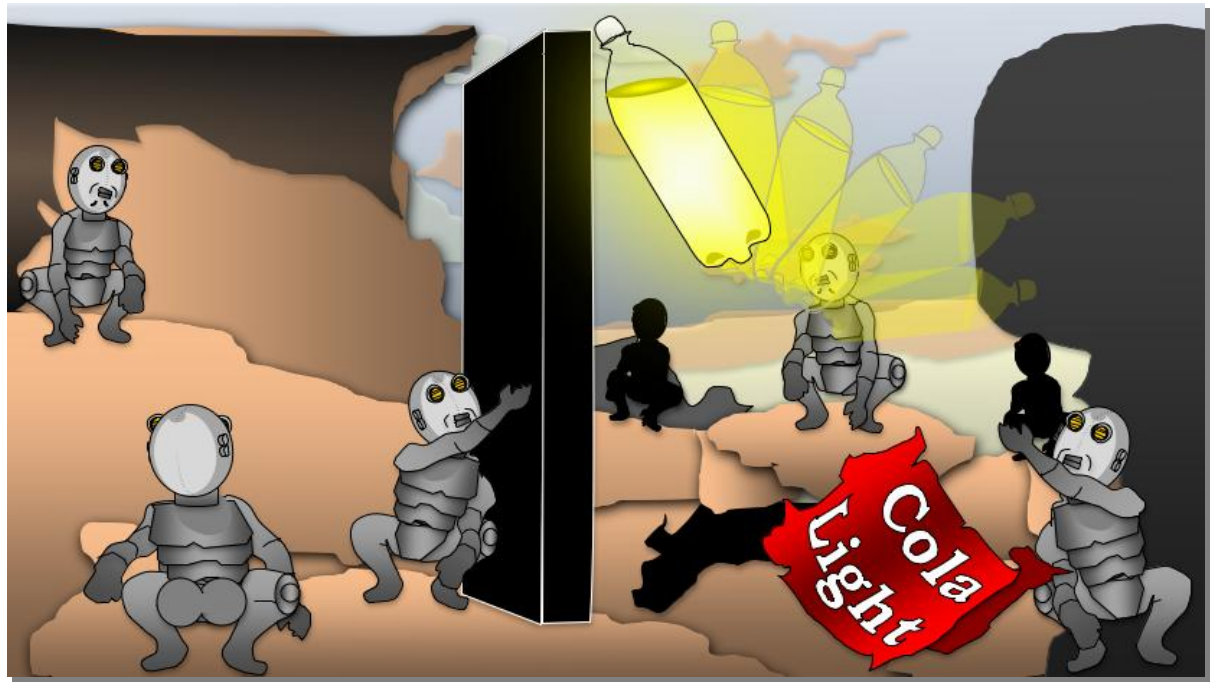
Yet there are no **Maxwell's equations** of thought, no **Keplar's laws** of creativity, and we must be wary of falling into a seductive trap that the **Artificlal Intelligence** (**AI**) theorist **Doug Lenat** has memorably named ***Physics Envy***: the belief that a complex phenomenon like creativity can ever be captured by a single equation or concise set of laws, no matter how profound or how elegant they may seem. Diversity is messy, and there is no simple formula that governs all that is beautiful, and no single generative mechanism that explains all of human creativity. When a phenomenon is so disrespectful of rules, how can we ever hope to pin it down with yet more rules?

Raphael's fresco *The School of Athens* personalizes the tension between the nomothetic and idiographic tendencies in science. At the centre of a bustling scene, **Plato** points upward to emphasize the supremacy of abstraction, while **Aristotle** – gesturing toward the ground – argues instead for the primacy of the particular and for the need to study the world in all its messy and surprising diversity. Plato advances a **top-down** view of science, espousing the use of intuitive abstraction to see the ideal forms that impose lawful order on the world. As he writes in the **Philebus**, "*From the gods a gift to the human race: thus I reckon the gift of seeing the One in the many and the many in the One.*" Aristotle instead advances the **bottom-up** view, espousing the need for empirical observation to see the world as it really is.

The world owes no allegiance to our convenient abstractions, no matter how elegant or ideal they may seem. We cannot choose to opt out of the law of gravity, but we can creatively choose to reimagine the workings of a cup, or the shape of a donut, or the classification of a chicken. Creativity fosters a healthy disrespect for the *right* or *idealized* way of seeing the world, and is better served by the anarchic views of **Groucho Marx** than by the lawful ideals of Plato. As Groucho quipped, "*Here are my principles! If you don't like them … well, I have others*." **Plato** and **Groucho** personify two complementary forms of thinking, called **convergent** and the **divergent** thought. Convergent thinking uses trusted procedures and safe generalizations to identify a single "best" or "right" answer to a problem. Mathematicians use convergence to rigorously prove a theorem, just as logicians use it to arrive at the inescapable conclusions of the known facts, much as **Sherlock Holmes** or **Hercule Poirot** use it to converge on the true perpetrator of a crime by successively narrowing the suspect pool. But divergent thinking suspends critical analysis to encourage many alternate views on a problem. Rather than seeking to narrow the candidate pool and converge on a single best answer, it allows this pool to grow, to generate many competing answers that each have their own individual merits.

But *Aristotle was way ahead of Groucho* when it came to the flexible treatment of shared principles. He described humour as a form of ***educated insolence***: to be deliberately humorous, we must understand the convergent limits of our habitual categories; only by abusing our categories can we exceed these limits and produce new insights.

Aristotle was intrigued by how we use and abuse abstractions, and was the first scholar to give a formal definition of **metaphor**. In fact he coined the term: it means to "carry over" words and meanings from one domain of experience to another. We can talk of a ship *standing* in the harbor, he said, because ships – like a man standing in one spot – can occupy a fixed location in space. So metaphor requires a suspension of disbelief as to the kind of ideas our shared convergent categories can be stretched to include.



Creativity is often metaphorically-grounded, relying on shrewd Aristotelian observations to suggest the divergent ways in which members of one category might substitute for another. Consider the use of clear plastic bottles as high-wattage light bulbs in sunny countries with little access to electricity. MIT students have shown how a **1L** bottle filled with clean water and a little bleach can be jammed halfway into a hole in the roof: as sunlight enters through the top it is refracted through the bottom, where it disperses to fill a dark space. Soda bottles and light bulbs share no *a priori* useful categories, but observe light refract through a bottle of water and the metaphoric similarities become obvious.

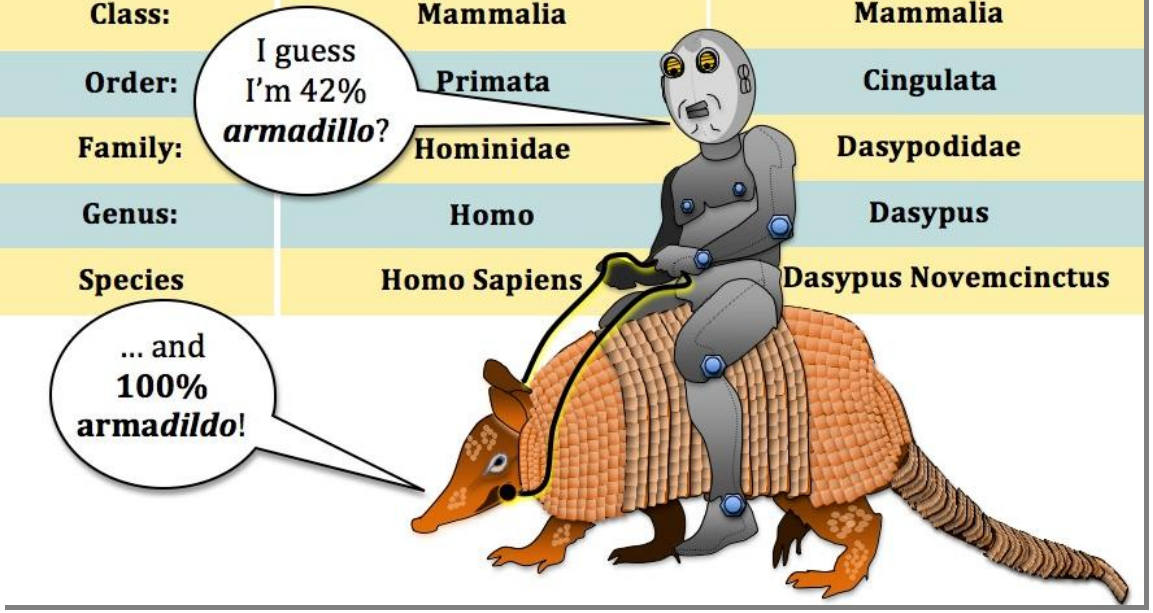
If I have seen further, declared **Isaac Newton**, ***it is by standing on the shoulders of giants***. Convergent thinking concerns itself greatly with the provenance of its knowledge and with the rigor of its methods. While Newton had anchored himself in solid foundations laid by the ancient Greeks, the rival he was chastising, **Robert Hooke**, chose instead to build on the not-so-widely-accepted **Descartes**. No wonder they could not converge on the same theory.

Convergent thinking assumes that there is a right answer to every problem, and a right way to find that answer. The rigor of one's methods ensures the quality of one's answers, while the quality of one's answers will inevitably burnish the perceived utility of the methods used. So if everyone starts from the very same axioms and employs the very same methods, everybody should converge on the very same answer. In science this "**me too**" convergence is actually a basis for confidence, as a scientific result is not truly a *result* until it has been soundly replicated by others.

Yet **convergence** also breeds **conservatism**. Pick the right giant to stand on and his shoulders will support you and many more like you, allowing you all to converge on a shared view of the truth. But diverge from the *one path* and, no matter how broad your giant's shoulders, you may have a torrid time convincing others of the merits of your findings. Convergent reasoning not only limits what we are able to find, but what we are willing to appreciate when we find it.

As Plato showed, ontologies and taxonomies have been our preferred *go-to* structure for convergent reasoning since ancient times: they tell us the most specific category to which two ideas belong, and can express this overlap in meaning in numeric terms. Consider the classic taxonomy of **Carolus Linnaeus** – commonly called the *Linnaean taxonomy* – which divides the biological world into seven layers or *taxa*: ***Kingdom, Phylum, Class, Order, Family, Genus*** and ***Species***. Every biological entity can be allocated its "right" place in this taxonomy by first determining its correct value for each of these seven taxa.

Taxon/Creature	Human Beings	Nine-banded Armadillo
Kingdom:	Animalia	Animalia
Phylum:	Chordata	Chordata
Class:	Mammalia	Mammalia
Order:	Primata	Cingulata
Family:	Hominidae	Dasypodidae
Genus:	Homo	Dasypus
Species	Homo Sapiens	Dasypus Novemcinctus



So modern man belongs to the species *Homo Sapiens,* the genus *Homo*, the family *Hominidae*, the order *Primata*, the class *Mammalia*, the phylum *Chordata* and the kingdom *Animalia*. Armadillos belong to the same kingdom, phylum and class, but diverge from us at the 4th taxon. Since we share ***3 out of 7*** taxa with armadillos, our similarity can be crudely quantified as **3/7**, compared to **4/7** for monkeys, **5/7** for apes and **6/7** for Neanderthals. Yet one needs to use a taxonomy very different to the Linnaean variety to exploit more cultural kinds of similarity, such as that "*armadillo*" means *"little armored man"* in Spanish, or that the Aztecs called them "*turtle-rabbits*". To think creatively, one shouldn't disavow convergence, but eagerly apply convergent thinking to ideas generated by divergent means.

It takes just a single act of eminent creativity to cement a creator's reputation, yet history is generally kinder to those whose creative outputs are both diverse and sustained. **Leonardo da Vinci**, the Italian artist and original *Renaissance man*, falls into this latter category of prodigious creator. Leonardo is rightly celebrated today not just for his iconic artworks, but for his surviving notebooks and letters that brim with all manner of imaginative inventions, from flying machines and parachutes to artistic, anatomical and architectural studies to the most fearsome machines of military might.

Leonardo was a master of what psychologists now call – after the pioneering work of **J. P. Guilford** – ***divergent* production**. Leonardo's creativity was not limited to a single domain or to a single modality, and he was capable of fluently generating many kinds of novel solutions to many diverse kinds of problem. Leonardo's wide-ranging curiosity lead him to produce an impressive array of intellectual treasures, but at the cost of rarely pursuing any single goal to completion. We celebrate his sketches, his plans, his ***Mona Lisa*** and his ***Madonna of the Rocks***, but his post-apprenticeship oeuvre comprises just ***28*** completed paintings; many more projects were never fully completed or even properly started.

Leonardo anticipated the judgment of history when he apologized for his restless divergence:

"*I will not concern myself here with proofs, which will reveal themselves anon when the work will be put in order",* he wrote*, "I will only concern myself with finding problems and inventions … Do not mock me, dear reader, if we here make great leaps from subject to subject*." Leonardo's leaps are the stuff of divergent thinking, while the proofs he pushes to one side are the proper target of *convergent* reasoning. While divergent production can yield a great many fascinating possibilities, it is convergent thinking that allows us to focus on a promising candidate and work it into a finished result. So divergent thinking is not inherently creative, just as convergent thinking is not inherently *un*creative. For a truly creative result, both should be used in tandem, to generate novel possibilities that can be fastidiously worked to completion.

Nonetheless, while convergent thinking is strongly associated with our standard conceptions of intelligence, it is divergent thinking that is more commonly associated with creative thought. Feeling that standard intelligence tests were undervaluing the role of divergence, **J.P. Guilford** set out to develop his own psychometric tools for creativity. His ***unusual uses test*** famously asks subjects to list a diversity of non-obvious uses for familiar objects such as bricks and empty coffee cans. The psychometrician **E.P. Torrance** later developed his ***Torrance tests of Creative Thinking*** on Guilford's pioneering foundations. These tests evaluate divergence along four dimensions: **1. fluency** – how many relevant ideas are produced? **2. flexibility** – how many different *kinds* of idea are produced? **3. originality** – how atypical are these ideas? and **4. elaboration** – how much useful detail do these ideas contain? Leonardo would surely have scored highly on all four dimensions, for his detailed sketches testify to his fluency at generating many kinds of new ideas in diverse disciplines. Unfortunately, he lacked the convergent focus to make more of them a reality.

The (Un)usual Suspects

An *unusual* use for one object is likely a boringly *usual* use for another. Indeed, most of the divergent uses that we might imagine for an object in the course of the *unusual uses test* will be, in some mundane context, a role or an activity that we perceive everyday, either in real life or on television. However, so strong are our associations for a test object that we cannot easily perceive those other possibilities. One way to *ace* the unusual uses test is to use a crib sheet of the most typical uses for *any* object, and simply attempt to fit the familiar elements of this inventory to the object in question.

Suppose we look on the Web for all completions of the phrase "used as a *", where * can match any noun. We can use a Google database called the Google N-Grams to quickly find possible noun fillers and their frequency of use in this phrase on the Web. The most frequent fillers – and the most frequent *second*-uses to which an object might be put – are listed on the following page

Now, let's play the unusual uses game with each of the following test stimuli. For each stimulus, first try to invent your own innovative uses. How many *more* uses can you invent by using the list of common uses on the next page? Feel free to think metaphorically – a "prison" does not have to a real prison, but any restrictive context – and metonymically – an object may not be usable as a "church", but it might be used as an altar, a sacrifice or an offering.

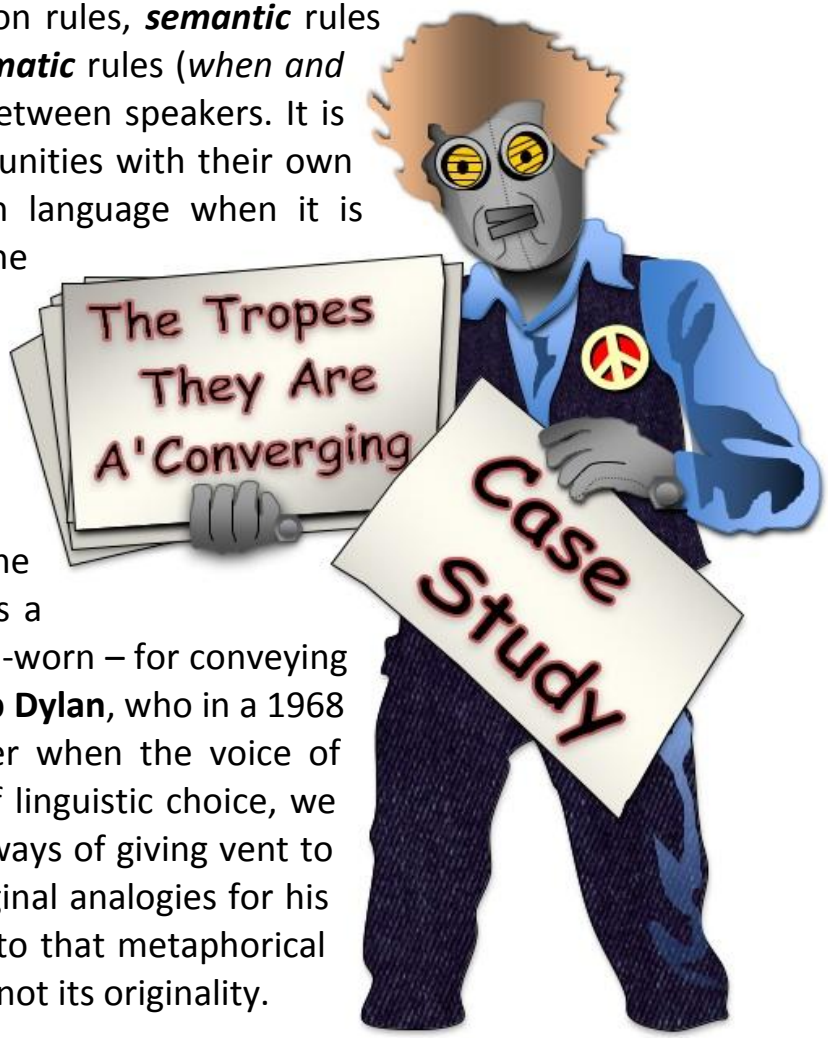
An empty coffee can	A windscreen wiper	A cardboard tube	A blunt knife	A chair leg
A microwave oven	One boxing glove	A spare room	A headless doll	A housebrick
A wind-up alarm clock	Scratched CDs	One chopstick	A plastic bag	A bag of marbles
Unused 1985 diary	Doll's head (no body)	A garden shed	Used batteries	A ball of string

These nouns are all frequently used on the Web to complete the phrase "*used as a **":

Model, weapon, vehicle, symbol, platform, treat, dwelling, catalyst, prison, drug, cure, bridge, light, library, shield, wall, fence, mask, church, classroom, bedroom, spice, car, yardstick, toy, computer, flag, decoration, gift, seed, warning, blueprint, laboratory, vaccine, code, password, threat, blanket, pillow, bookmark, kitchen, cemetery, map, barometer, window, hammer, diaper, marinade, tray, mirror, missile, sensor, rug, lens, sword, diary, bomb, tranquilizer, hook, poison, perfume, joke, chair, compass, cloak, trap, dump, bank, clock, purse, puppet, battery, whip, scarf, pot, cudgel, sink, plate, notebook, bullet, counterweight, hat, sponge, thermometer, keepsake, birdbath, nest, cane, pendulum, bracelet, spear, necklace, clamp, spoon, tomb, brooch, shrine, nail.

- What qualities should the most reusable elements on this inventory possess?
- Is an inventory approach likely to increase or reduce divergence?
- How might you generalize this inventory into a reliable technique?
  - How might your generalized technique be automated?
  - What additional categories would you add to the above inventory?

Language is a perfect example of a system in which convergence and divergence work together to achieve useful and creative ends. For language is principally a rule-based semiotic system that requires is to agree on much the same signs and much the same rules for combining and conveying these signs – whether grammar rules (***syntax***), word-formation rules (***morphology***), spelling and pronunciation rules, ***semantic*** rules (*how to compose complex meanings from simpler meanings*) and ***pragmatic*** rules (*when and where to say what*) – before it can support effective communication between speakers. It is convergence around language that binds speakers into linguistic communities with their own shared cultures and reference sets. Divergence is meaningful within language when it is couched in this larger convergent context: if, all at once, we jettison all the rules of syntax, morphology, semantics and pragmatics, our audience has very little hope of ever extracting the intended meaning from what has become, in effect, our own highly divergent ***private*** language. So while our clever metaphors and jokes may violate *some* of the rules, they cannot violate them *all* if we ever want to be understood.



Membership in a linguistic community requires convergence to the rules, but membership has its benefits. Language membership gives us a treasure-store of handy constructions – from the gently-used to the well-worn – for conveying our meanings in concise, pithy and easily understood ways. Consider **Bob Dylan**, who in a 1968 concert was loudly decried as "***Judas!***" by a fanatical folk-music lover when the voice of protest *went electric*. When the heart overrules the head in matters of linguistic choice, we quickly converge on the most familiar, and most emotionally-charged, ways of giving vent to our feelings. Dylan's irate fan might have composed any number of original analogies for his erstwhile hero, but relied instead on a *very* old reliable: a comparison to that metaphorical father of *all* traitors. That the epithet is memorable is due to its context, not its originality.

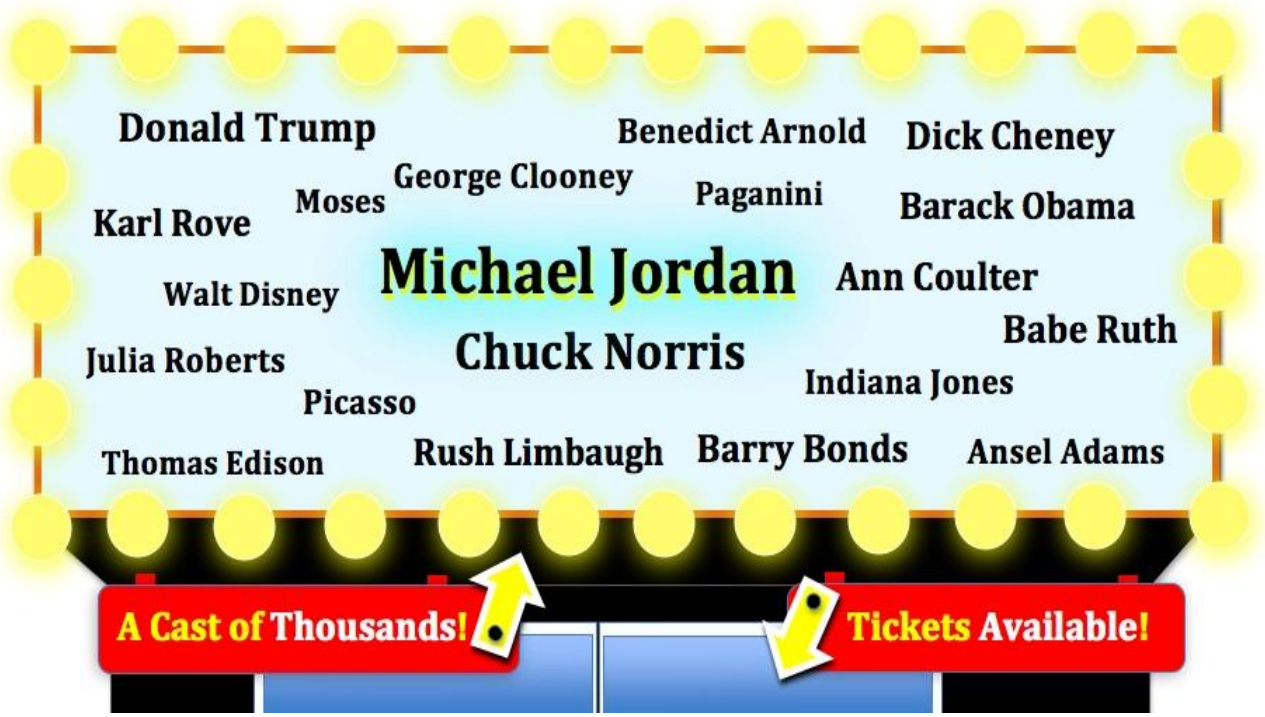
The epithet "*Judas!*" is an elided use of a linguistic structure that linguists call the *XYZ schema*.

Bob Dylan was no **Judas Iscariot**: rather, his critics saw **Dylan** (=**X**) as an analogy for **Judas** (=**Y**) in the **music** world (=**Z**). The XYZ construction finds many unremarkable uses in everyday language, from "***Sally*** *is the* ***mother*** *of* ***Thomas***" to "***Larry Page*** *is the* ***CEO*** *of* ***Google***". But creative uses abound in poetry too, from Wallace Stevens' "***Death*** *(=X) is the* ***mother*** *(=Y) of* ***Beauty*** *(=Z)*" to Buckminster Fuller's "***Love*** *(=X) is* ***metaphysical****(=Z)* ***gravity*** *(=Y)*". In its most schematic form, the XYZ construction is a high-level trope toward which speakers naturally converge when they want to draw parallels between entities in two different domains. Frequent usage of the structure marks it as an ideal packaging for our off-the-cuff analogies and pithy comparisons – note how the structure does not require us to articulate precisely why X is similar to Y, but merely suggests the presence of a similarity that the listener must work to unpack. Few people need to be told why **Dick Cheney** was viewed as the **Darth Vader** of the **Bush** administration, but more left-field uses, such as "**Riesling** is the **Kenny G** of the **wine world**" (technically brilliant but lacking in hipster appeal), or "**Red meat** is the **Donald Trump** of **cancers**" (an aggressive *builder* of different cancers) can stymie an audience with their creatively challenging juxtapositions.

Convergent and divergent thinking modes intersect in the XYZ construction: the use of a familiar support structure like the XYZ form is an obviously convergent choice, but we are free to divergently fill the **X**, **Y** and **Z** slots as we see fit, to achieve ends that range from the banal to the strikingly creative. The most creative forms stand out against a landscape of unoriginal uses that rely on a stock of old familiars, from **Napoleon** to **Einstein**, for their Y elements.

The words we use to name categories are often quite distinct from the words we use to name the individual members of a category. So, "***golfer***" names the category of golf players and "**Tiger Woods**" denotes just one individual (albeit a prominent one) in this category. However, when it comes to enforcing the logical distinctions between categories and their members, language operates a well-oiled revolving door. The enigmatically-named XYZ construct allows a speaker to creatively turn prominent members of a category into new categories in their own right.

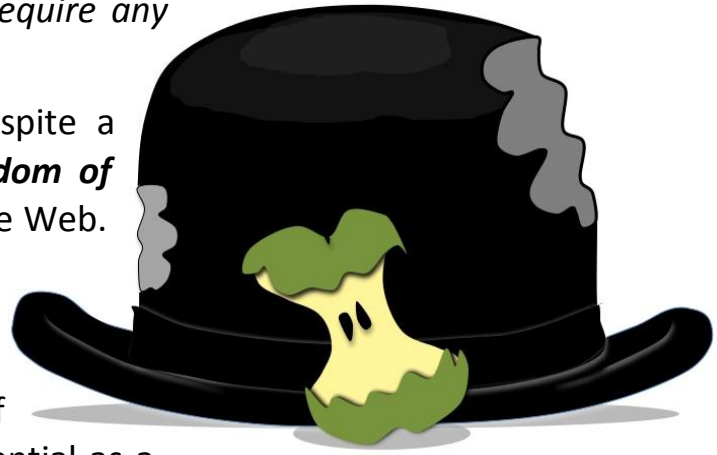
The world-wide-web is a tremendous source of divergent production, both linguistic and conceptual. If we look to the Web for examples of XYZ constructs in which the Y-element is a proper-named individual, such as Tiger Woods or Vladimir Putin, we can quickly find a wide range of ***non-literal* XYZs** in which the X is *not literally* a Y in any domain Z. Indeed, this is a process that is easily automatized, since proper-names in English are conveniently capitalized and easily recognized by a computer. When we program a computer to harvest as many XYZs with proper-named Y's from the Web (via Google), we quickly pull in thousands of examples, ranging from dull to witty to creatively challenging. The image on the right displays the most frequent proper-named Y's in our 2000 XYZ sample: the size of each name shows its relative frequency in the sample. So **Michael Jordan** – an icon of sporting virtuosity and business acumen – occupies top spot as most used **Y** (**21** examples) closely followed by **Chuck Norris** (**20** examples), a convenient Y for anything that is the rugged but unsophisticated.



As one who occupies the pinnacle of his chosen sport, Michael Jordan has become a role model for strivers in any sport, and some non-sports besides. Jordan is thus an excellent case-study for the analysis of creative and not-so-creative XYZ comparisons. As listed below, Jordan has been used to accentuate the qualities of **boxers** (Y = *Philippines* and Y = *boxing*), antipodean **basketballers** (Y = *NBL*), boomerang **aces** (Y = *boomerang*), chess **grandmasters** (Y = chess), female **basketballers** (Y = *WNBA*), American **footballers** (Y = *NFL*), **golfers** (Y = *golf*), **martial artists** (Y = *martial arts*), **skateboarders** (Y = *skateboarding*), paralympic **basketballers** (Y = *wheelchair basketball*), **lawyers** (Y = *courtroom*), **bakers** (Y = *cakes*), **writers** (Y = *American novelists*), film **stars** (Y = *movies*), curators (Y = *museums*), **philosophers** (Y = *seminar room*), book **sellers** (Y = *rare book world*) and even tuna sandwiches (Y = *the midday meal*).

Some X's are closer to Jordan than others, and most are simply other sportsmen, with the closest being basketballers of an other country, league or gender. Jordan is a model of excellence for any competitive endeavor, possessing qualities that even lovers of tuna fish can apparently find inspirational.

Belgian surrealist René Magritte was underwhelmed by the creativity of the *Pop Art* of the 1960s. "*The humour of Pop Art is rather* ***orthodox***", he said, noting it is "*within the reach of any successful window-dresser*." For Magritte was unimpressed both by the lack of imagination on show and by its lack of technical virtuosity, adding that "*to paint large American flags with one star more or less does not require any particular freedom of mind and does not present any technical difficulty*."



Much the same can be said of most new XYZ comparisons, for despite a superficial diversity, most are disposable forms that show very little ***freedom of mind***. Consider again our sample of **Y=*Michael Jordan*** comparisons from the Web. Most never depart the realm of sporting endeavor, and all but few hit the same high notes in Jordan's career, depicting the corresponding **X** as one who dominates their field. Few touch upon another side of the consummate champion that is well known but rather less well exploited. At the height of his basketball career, Jordan took a controversial hiatus to explore his potential as a baseball player in the minor leagues. Unfortunately, the early promise he showed in high-school failed to blossom beyond mediocrity and he soon returned to basketball. Yet at the time of writing, just one Web XYZ uses Jordan as an exemplar of a high-achiever in one field whose skills fail to transfer into other: this XYZ describes not a person, but a software system that delivers functionality without versatility. Such counter-conventional comparisons tend to be accompanied by explanations that draw out their meaning, lest an audience assume the contrary conventional reading. Generally speaking, XYZs that need to be bolstered by explicit explanations are either very poor (if Y fails to impart the desired message on its own) or very good (as when another side of a famous Y is exposed for ironic effect).

The XYZ "*Britney Spears is the Michael Jordan of pop music*" is also interesting for its gender-bending choice of X and Y. Though at her height Spears wielded a Jordan-*esque* sway over her field, consider how unlikely we are to see "*Michael Jordan is the* ***Britney Spears*** *of basketball*" in anything but an ironically neutering context. Conservative biases can seriously curb our freedom of mind, and that of our audience, even when we do not admit to having them.

Analysis of *all* our 2000+ XYZs suggests we favor a conservative approach to creative comparison, where **gender** is just one of the pervasive biases that limit our open-minded exploration of the divergent possibilities. **91%** of our Web XYZs describe a **human X** of known gender, since our **Y**'s are always famous people of known gender. Outliers include **X**'s from the domain of ***foodstuffs*** (*tuna sandwiches*), ***gadgets*** and ***buildings*** (so "*The Borgata is the **Julia Roberts** of casinos*", apparently). The majority of gendered comparisons are of the ***Male-to-Male*** variety (**81%**), which perhaps says as much about the nature of the Web as it does about our XYZs. We also seem to favor the real over the invented, for the vast bulk of our XYZs (**88%**) compare a real-life person to a real-life person. When a fictional character is thrown into the mix, we are four times more likely to use a **fictional Y** to describe a **real X** than the other way round. Perhaps it is easier to think of a fictional character as merely an exaggerated version of a real person?

The most striking conservatism can be observed in our choice of the domains that supply **Y**'s for our target **X**'s. When our **X** is a *political figure* (which is true in about **18%** of XYZs), we seem compelled to also choose a *political* **Y** in **63%** of cases. This pattern of drawing **X**'s and **Y**'s from the same general domain is replicated throughout our sample of Web XYZs, whether the domain in question is **music**, **art**, **business**, **entertainment**, **sport** or **science**. These implicit biases guide us toward **in**-*domain* comparisons between **same**-*sex*, **same**-*world* **X**'s and **Y**'s to help us to mint unchallenging XYZs that are easy to understand without explicit side-bars. But they also limit what **Magritte** called our freedom of mind to explore diverse ideas and to generate truly original forms. Technical virtuosity can do very little to disguise an unoriginal idea. They key to overcoming our biases is not feigned liberality, but an honest engagement with these biases, so that we can actively seek out the most surprising **cross**-*gender*, **cross**-*world* and **cross**-*domain* **Y**'s for our **X**'s.

The View Askew

Researchers Carlos Roncero, John M. Kennedy & Ron Smyth have studied the way we use bespoke similes on the Web, and have noted that speakers prefer similes to metaphors when they want to link two ideas using an *out-of-the-ordinary* relationship. The simile frames "*X is as Z as Y*" and "*X is like Y*" (or indeed, "X is the Y of Z") are easily instantiated with any **X**, **Y** and **Z** values we choose, but this freedom of expression often poses a challenge to our audience. If we compare **Eric Clapton** to **Kenny G** (as in the XYZ expression "*Eric Clapton is the Kenny G of blues music*"), are we praising Clapton for his technical virtuosity or damning him for his easy-listening banality? In the case of **as** and **like** similes, Roncero and his colleagues have shown that creative (non-idiomatic) similes on the Web tend to be accompanied by explanations of intent for this reason: the more creative the simile, the more ambiguously rich its interpretation.

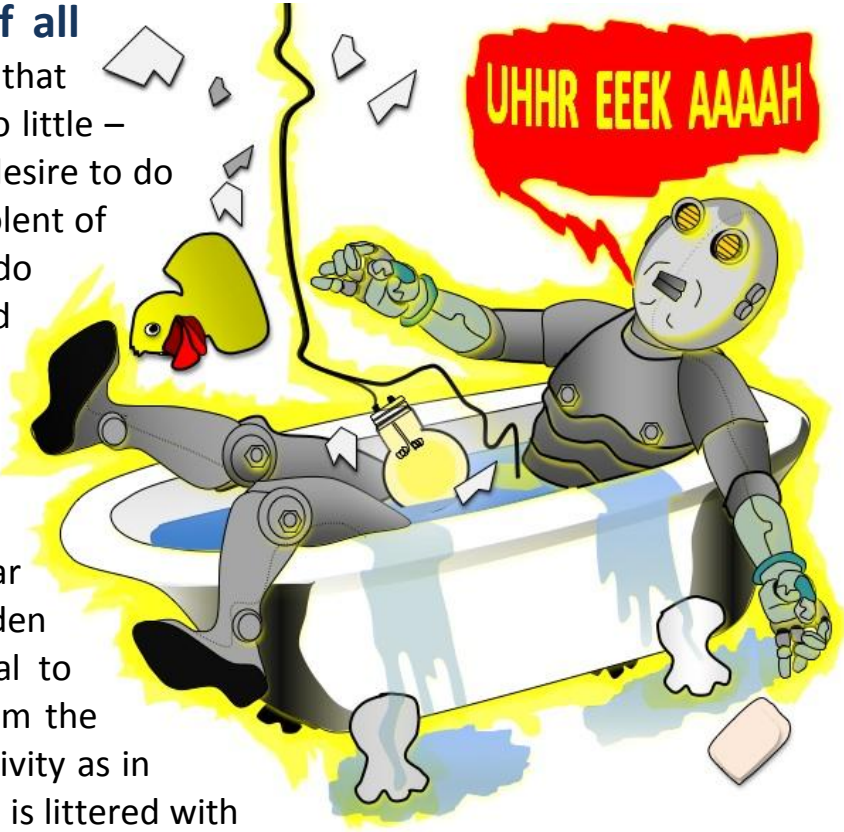
- **Replicate** (or refute) **Roncero, Kennedy & Smyth**'s findings for **XYZ** comparisons on the **Web**. Are they more likely than not to be accompanied by explanations of their meaning? Do the XYZs that come *with* explanations have any qualities than those that come *without* explanations do not?
- Looking at the Web XYZs with explanations, how often does the explanation gel with your own interpretation of the comparison? Do the ones that differ significantly seem to more or less creative?
- We have focused here on XYZs where the Y is a **proper-named individual**. How about other creative XYZs where Y is a **common-noun**? How might you automate the harvesting of those comparisons from the Web?
- What can a computer *learn about the world* by harvesting creative Web XYZs? How might it learn to compose its **own** creative comparisons? What kinds of tacit knowledge must it also acquire to perform this task well?

Most computer programs are highly convergent systems: though their internal data-structures can be instantiated in a vast number of ways, a tightly-written program has a hard-wired sense of the *right* data values and the *wrong* data values for each. Moreover, most programs lack the ***self-knowledge*** to analyze their own pre-set biases and to experimentally explore a space of alternative values. Most programs are just not wired to be ***divergent***.

Yet this need not be so. Modern programming languages now provide a means for any program to *introspect* into its own code at run-time. The **Java** language, for instance, provides a ***Reflection*** interface that allows a program to query and probe the data structures that make up its algorithmic fabric. Importantly, reflection allows one program component to obtain a detailed inventory of *all* of its data-structures, as well as any methods for modifying those structures. Computational Creativity researcher **Michael Cook** cleverly exploits this capability in a divergent game-creation system called ***Angelina***.

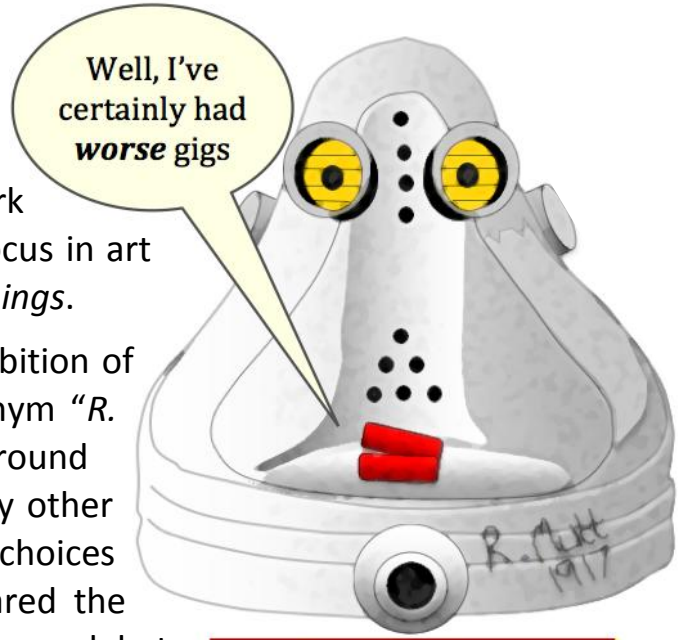
Angelina generates **SuperMario**-style arcade games in which players navigate through a sequence of increasingly difficult levels and obstacles. These games usually give players a set of fixed capabilities, such as jumping and ducking, and a random selection of short-lived *bonus abilities* such as greater speed, higher jumps or temporary invulnerability. But **Angelina** divergently invents its own novel abilities – such as the ability to temporarily invert gravity or walk through walls – by using reflection to randomly modify its many data structures, and by running experiments to determine the benefits of specific random changes to a computer-simulated player on a specific game level. The approach, akin to a brain surgeon randomly probing his own brain to find and identify those areas that give rise to pleasurable stimuli, allows **Angelina** to discover changes that yield novel but effective problem-solving short-cuts around which a challenging game-level can be constructed.

“Originality” is the most contentious dimension of all when it comes to assessing divergent production. It doesn’t help that the word itself – which seems to promise so much and deliver so little – is an opaque near-synonym of “creativity” that tacitly implies a desire to do so much more than merely generate. Indeed, the word is so redolent of freshness and vitality that we use it to describe phenomena that do more than challenge our expectations: true originality startles and surprises, and palpably excites the senses. When art critic **Robert Hughes** packaged his views of modern art as a documentary and accompanying book in 1980, his chosen title crackled with the energy we expect of *originality*: **The Shock of the New**.



Yet can anything we create be said to be truly original? Popular narratives of creativity talk of *revolutions* in thinking and sudden *paradigm shifts*, of *disruptive breakthroughs* with the potential to change everything, but no creation can be cleanly unpicked from the cultural forces that precede it. **Shakespeare** said it best: In creativity as in history, *what is past is prologue*. Inevitably, the history of science is littered with priority disputes as to who deserves credit for an “original” idea: who invented the calculus: **Newton** or **Leibniz**? Who invented radio: **Marconi** or **Tesla**? Who invented television: **Farnsworth**, **Zworykin**, or **Baird**? Or the telephone: **Meucci** or **Bell**? Who discovered the HIV virus: **Gallo** or **Montagnier**? Though a cultural consensus usually forms around a single candidate for each new creation, the reality is more complex than the narratives we weave. **Einstein** surely deserves plaudits for his theories of relativity – special *and* general – that transformed modern physics, yet those who cleared the trail that he would blaze so spectacularly, from **Poincaré** to **Hilbert** to **Lorenz**, also deserve credit for their contributions to these shockingly “original” theories.

Nothing in art challenges our notions of originality as much as "**Readymades**". Artists have traditionally been celebrated for their skill at creating new artifacts, but the advent of the *readymade* artwork – in which an artist elevates a preexisting object to the status of an artwork merely by exercising **selectivity** (but not **generativity**) – has shifted the focus in art from the making of new and original objects to the making of original *meanings*.

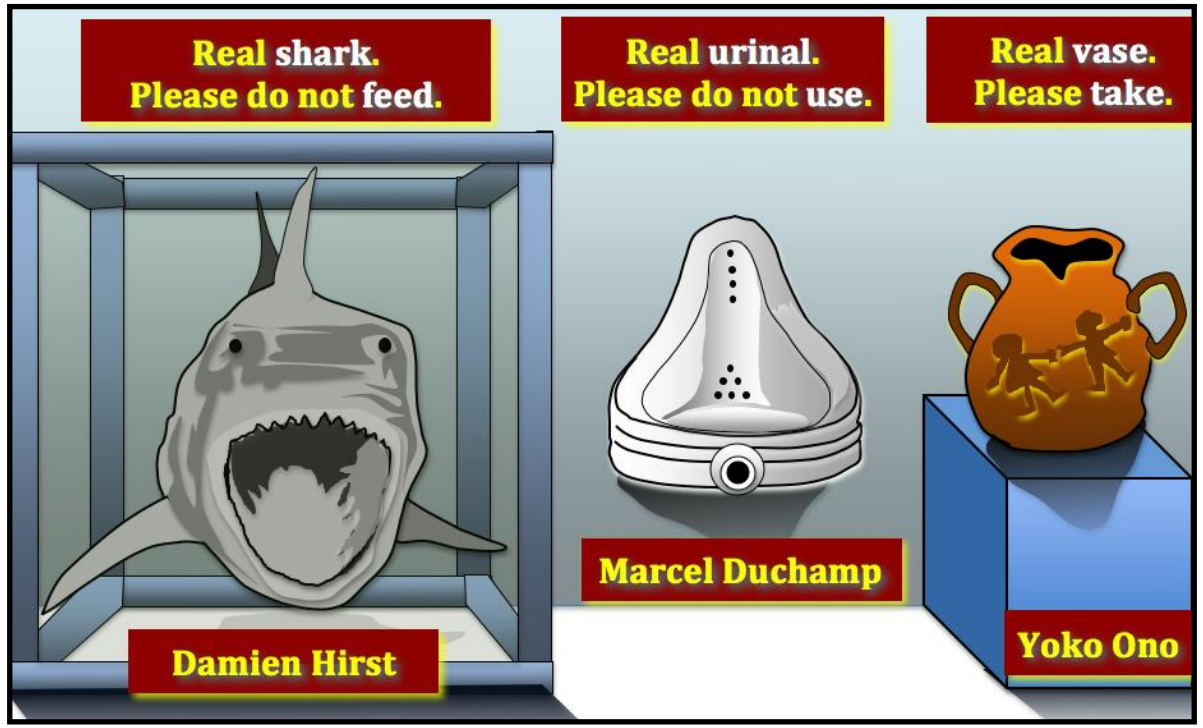


Marcel Duchamp's *Fountain*

In 1917, **Marcel Duchamp** mischievously submitted a *urinal* to an exhibition of the *Society of Independent Artists* in New York. Signed with the pseudonym "*R. Mutt, 1917*" in the style of an artist's signature, and rotated 90 degrees around its normal axis, the piece – titled ***Fountain*** – was a humble urinal in every other respect. Duchamp's point was that artistic creativity resides as much in the choices we make as in the artifacts we produce. Though the *Society* had declared the exhibition open to all submissions, Duchamp's entry provoked such fierce debate within the *Society* that it was all but rejected from the exhibition; though technically present, it was virtually hidden from view. The controversy prompted Duchamp to resign from the *Society*, disgusted that even a definition of art as all-embracing as the *Society*'s could still be exclusionary and afflicted by hypocrisy. Indeed, the failure of even this weakest of definitions of creativity still smarts today: in 2004, a pre-Turner-prize poll of 500 prominent artists judged ***Fountain*** to be the most influential artwork of the 20th Century. Weak definitions over-generate by design, but Duchamp showed us that even the weakest can *under*-generate as well.

Readymades like ***Fountain*** result from a divergent thought process strikingly similar to the one psychometricians evaluate via the ***unusual uses test***. For when faced with a familiar object in its familiar context, an artist might imagine what qualities would come to the fore if it were placed in an alien setting such as a gallery. Though we're unlikely to see a brick or a paper bag exhibited at an art show, readymades make it more likely we *will* see an unoriginal object in a new light. Readymades embody the view of art championed by **Andy Warhol**: "*art is what we can get away with*."

The *Unusual Uses Test* is *in vitro* creativity, in which subjects are asked to be creative with specific stimuli under strict laboratory conditions. Readymades, in contrast, arise from ***in vivo*** creativity, in which selective and often quite serendipitous uses are made of the diverse, random stimuli we encounter in the course of our everyday lives. Many readymades find new uses for man-made objects, yet the most dramatic examples find aesthetic uses for wood, bugs, shells, bones, plants, rocks, excreta and even dead animals, **Damien Hirst**'s infamous pickled **shark** being a rather lucrative example of the latter. Neither do our readymades have to be obviously artistic: when we lack the *right* tools to solve a mundane problem, creative new uses are often found for familiar objects to yield an original needs-driven replacement.



Indeed, to create new possibilities and new meanings from preexisting forms, our readymades do not even have to have *physical* forms. Any stimulus at all in our environment that has a conventional interpretation and that can easily be communicated to others – whether an image, a look, a gesture, a sound, a fragment of music, a snatch of poetry, a lyric, a title, a proper name, a line from a movie or play – can be imbued with new meaning in new contexts, from jokes and ironic allusions to the names of companies, bars, restaurants and racehorses to the titles of books, movies and plays to Twitter hashtags, web and email handles. Readymades, like linguistic metaphors, are *dual signifiers*: they represent themselves in their typical forms, *and* they convey new meanings in their new contexts.

Good artists *borrow*, great artists *steal*. That's how Picasso saw the somewhat elastic relationship between creativity and originality. For no matter how original one strives to be, creativity will always require a certain amount of **synthesis**, of the topical ideas that waft through a culture at a given time and place. To *borrow* is to take an idea whose ownership clearly remains with another, but to *steal* is to take full ownership of that idea for oneself. To truly make an idea one's own, one must add to it, improve it, and imbue it with one's own distinctive qualities. A successfully stolen idea thus becomes associated with the one who steals it, not via base trickery but via the mastery and vitality that can only come from an investment of passion and effort.

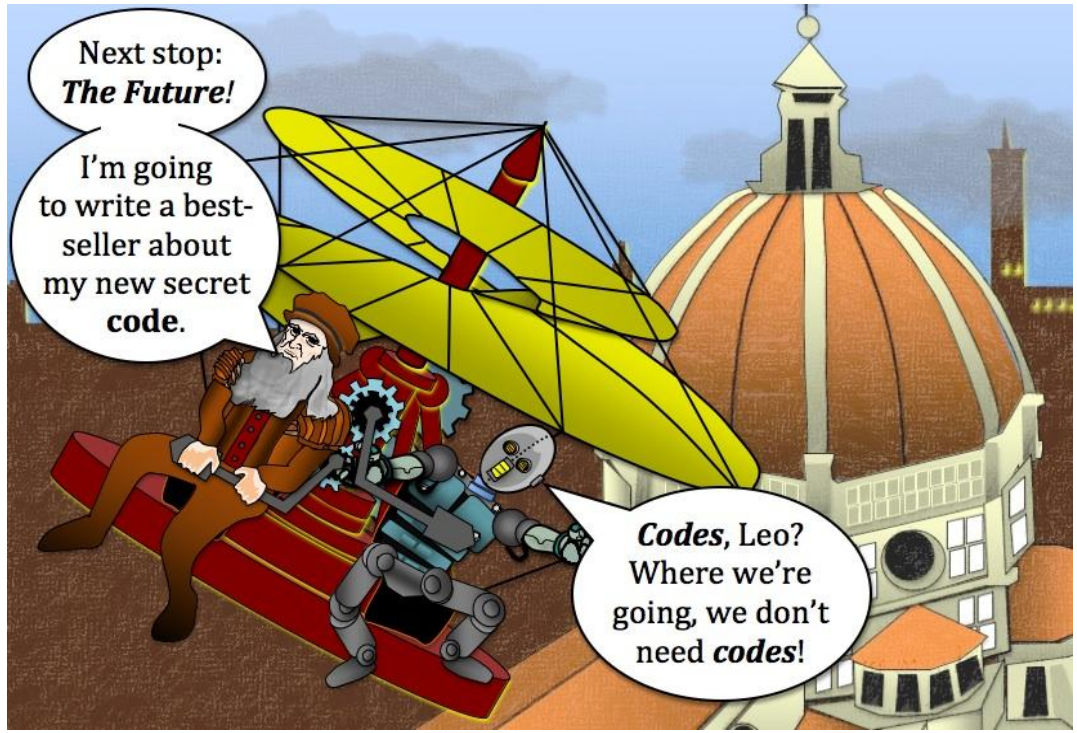
This is the sense of "steal" endorsed by **Steve Jobs** when talking about the creation of the Apple **Macintosh**, supplementing Picasso's aphorism with the admission "*we have always been shameless about stealing great ideas*." The greatest of these was surely the ***WIMP*** graphical interface (of ***W****indows,* ***I****cons,* ***M****enus and* ***P****ointers*) that reached its apotheosis with the Macintosh, but which had earlier been pioneered by researchers at **Xerox Parc** in Palo Alto.

Steve Jobs was famously divergent in his magpie-like "*stealing [of] great ideas*", but just as famously convergent in his fussily neurotic drive for perfection. Though Apple's Macintosh GUI took significant elements from the **Xerox *Alto*** and ***Star***, Jobs had Apple fine-tune and extend these elements – for example, by adding features that seem obvious in retrospect, such as overlapping windows – until the WIMP GUI became distinctively Apple's own.

Others have famously dipped into the same creative well, leading Apple to cry theft, or as Jobs is reputed to have shouted at **Microsoft**'s **Bill Gates**, "Y*ou're ripping us off! I trusted you, and now you're stealing from us.*" Gates used a form of divergent thinking to present a creative defence-by-analogy: "*Well, Steve, I think there's more than one way of looking at it … it's more like we both had this rich neighbor named Xerox and I broke into his house to steal the TV set and found out that you had already stolen it*." Jobs is often accused of hypocrisy for shamelessly celebrating his own appropriation in one breath while decrying Microsoft's in another, and for still later declaring "*thermonuclear war*" on **Google** and **Samsung** for their *iPhone*-like smartphone, which Jobs labeled a "*stolen product*". Yet one might reconcile the contradiction by appealing to Picasso's aphorism: while Apple had the temerity to take ownership by *stealing*, Apple's rivals often seem content with mere *borrowing*. As Jobs put it, with characteristic disdain: "*The only problem with Microsoft is they just have no taste, they have absolutely no taste. I don't mean that in a small way. I mean that in a big way, in the sense that they don't think of original ideas and they don't bring much culture into their product*."

We apply different standards when assessing the work of convergent and divergent thinkers.

Since convergent thinkers are assessed more on the reliability of their results than on the novelty of the means by which those results were obtained, we value the products of convergent thinking more highly when other convergent thinkers produce exactly the same results. Divergent thinkers, in contrast, are assessed on the novelty (or mystery) of their methods and on the novelty of their results. While we expect our mathematicians and historians to converge on the same truth of the same conjectures, we don't expect, or ever want, our divergent creators to produce virtually the same novels, artworks or jokes as each other.



So to properly assess the value of a new result, we must apply the most appropriate standards. Thus, a result that is touted as a work of rigorous convergence should be assessed as such, while one that is trumpeted as the product of flexible divergence should be assessed very differently. Yet is to be expected that creators sometimes deliberately blur these lines. For instance, the book **The Holy Blood and the Holy Grail** – which traced a blood-line from **Jesus Christ** to the kings of France to the shadowy founder of a modern cult – was marketed as a work of historical non-fiction. Yet when **Dan Brown** used these *facts* in the plot of his fictional novel ***The Da Vinci Code***, the authors of the former sued him for plagiarism (and lost). So is **The Holy Blood** a work of investigative convergence or a work of fictional divergence? It seems to be whatever the authors – who share the same publisher as Brown, and who saw their book climb the sales charts during the lawsuit – want it to be at any given time or in any given market.

Blast from the Past

Readymades are everywhere in modern society, and are visible in almost any street or place of business. They are the names of racehorses, films, books, music bands, restaurants and shops. Consider, for instance, the thoroughbred racehorse **Beef Or Salmon**, an Irish-trained prize-winner named for the question most often posed to guests at an Irish wedding banquet – "*Beef or Salmon?*". Or Consider the movie ***The Usual Suspects***, which takes it name from a much-quoted line in the movie **Casablanca**, the pivotal line with which cynical police prefect Louis Renault signals his change of allegiance from Vichy France to Rick and the resistance. Or consider the 1980's pop band ***Duran Duran***, who take their name from a villainous character in the 1960's movie **Barbarella**.

- Find your own examples of linguistic readymades in modern society, in movies, poetry, business and so on.
- Invent some readymade names of your own for music bands, movies, or products. Use or build on existing idioms and cultural reference points (e.g., the phrases ***Epic Fail*** and ***Mercy Flush*** might be interesting band-names).
- What signal does such a readymade send to an audience about its user / creator? Under what circumstances might a readymade backfire on its creator?
- Duchamp showed that with readymades, an artist does not actually need to possess *generative skill* but only needs an *aesthetic sensibility*, that is, an eye for the beauty of everyday things. Under this altered definition of art, is it more plausible that (some) animals and (some) machines might be appreciated as real artists?
- When a linguistic readymade is built from an existing idiom or phrase, does the position of the varied element (beginning/middle/end) affect the understandability and pleasurability of the novel stimulus? How so?

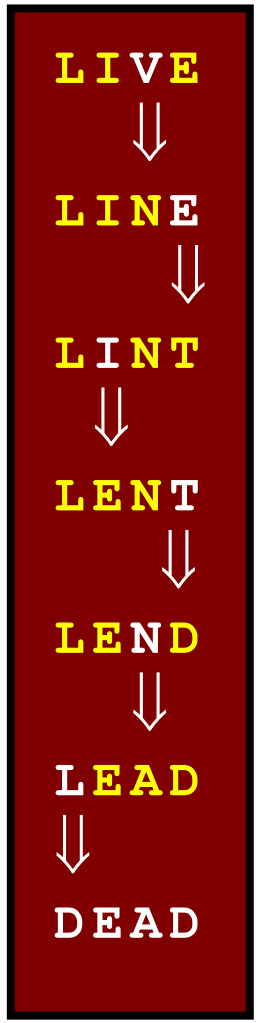
The language of creative problem-solving is frequently the language of *exploration*. We talk of *seeking* inspiration, of *finding* a solution, of *pursuing* leads and *following* avenues, of reaching a *dead-end*, of hitting a *brick wall*, of having to *backtrack* or even to *start over*, of getting *sidetracked* or *bogged down*, of *taking the path least travelled*, of *overcoming obstacles* or of being totally *lost*. For we metaphorically view ideas as **abstract states** in a **space of conceptual possibilities**, and see creators as explorers of these spaces. Each space gives rise to its own states and to its own creative possibilities. To explore any space, we must first choose a convenient **start state** – an inspiring example, perhaps – and follow a chain of ***representational changes*** that move us from state to state as we seek out virgin areas of the space. **Steve Jobs** and **Bill Gates** each explored the space of GUIs using the **Xerox Alto/Star** as their starting point, but each passed through very different states to reach **goals** of diverging quality.

Even the simplest of spaces may contain an infinite number of states and thus support infinite possibilities. Yet a space of infinite possibilities may support only a finite number of truly original forms – the ***classics*** in waiting – hidden amongst an infinite number of variations and pastiches. **Raymond Chandler** put it thusly when speaking of the space of possibilities in detective fiction: "*Within its frame of reference, which is the only way it should be judged, a classic is a piece of writing which exhausts the possibilities of its form and can never be surpassed. No story or novel of mystery has done that yet. Few have come close. Which is one of the principal reasons why otherwise reasonable people continue to assault the citadel."* Chandler's **frame of reference** is a **conceptual space** in which each state is given its own subjective valuation by critics, by consumers and by the creators themselves. The goal of a good creator is to do more than identify those states which correspond to trivially new possibilities, but rather to find those states that maximize a shared evaluation function and which validate and make explicit the quality that is achievable in a space.

A game proposed by author and mathematician Lewis Carroll in *Vanity Fair* magazine in 1879

offers a simple perspective on search-spaces. The game – christened **Doublets** by Carroll but known today by a variety of names (such as **Word Ladders**, **Step Words** & **Word Chains**) – draws on our knowledge of words and their spellings to create a fun but challenging search problem. Carroll described its rules as follows in his article:

> The rules of the Puzzle are simple enough. Two words are proposed, of the same length; and the Puzzle consists in linking these together by interposing other words, each of which shall differ from the next word in one letter only. That is to say, one letter may be changed in one of the given words, then one letter in the word so obtained, and so on, till we arrive at the other given word. The letters must not be interchanged among themselves, but each must keep to its own place. As an example, the word ***'head'*** may be changed into ***'tail'*** by interposing the words ***'heal, teal, tell, tall'***. I call the given words **'Doublets'**, the interposed words **'Links'**, and the entire series a '**Chain' …**

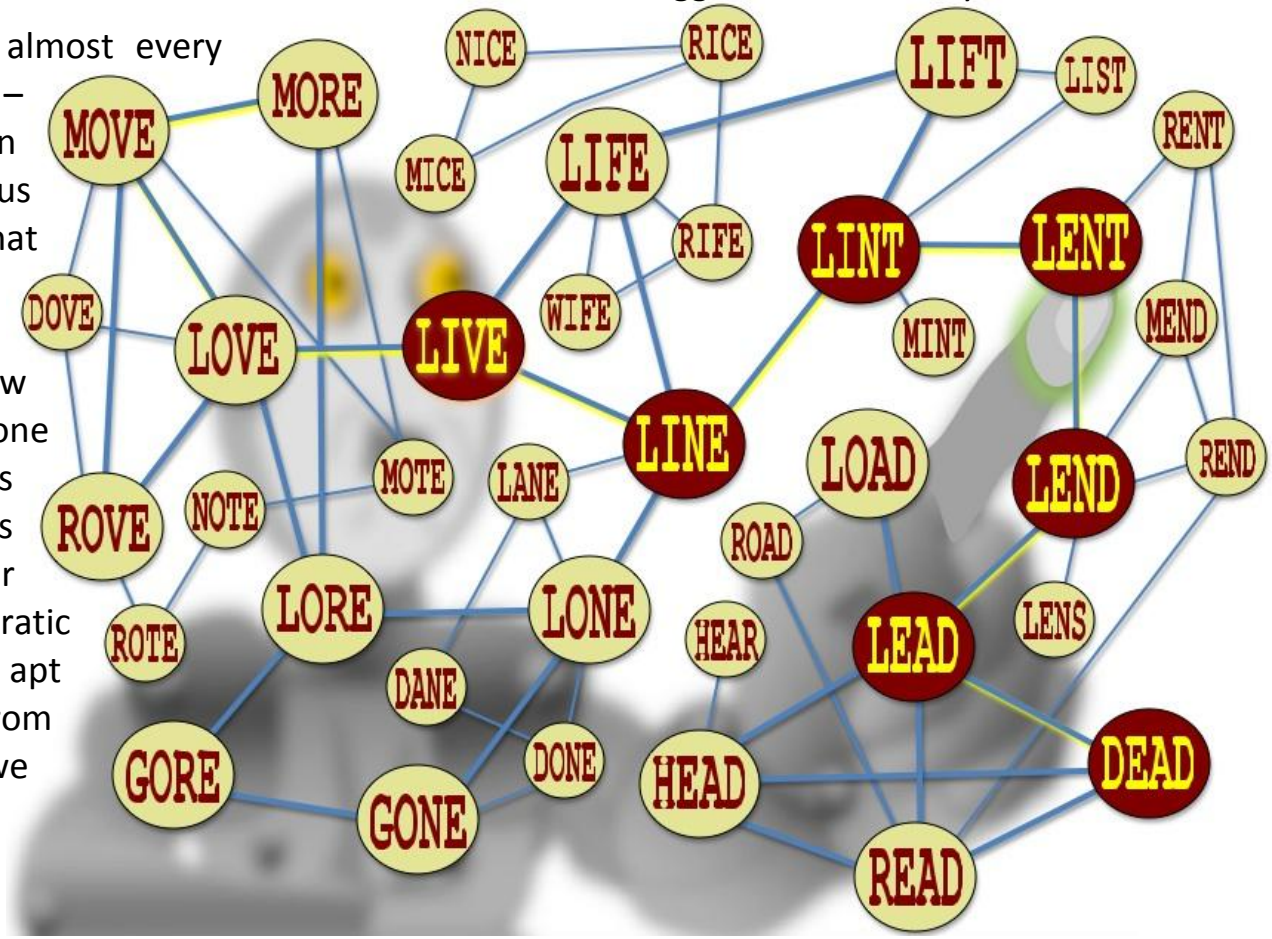


An example of **Doublets** in action is shown on the right. This completed chain offers a simplified view of the challenge, and omits the false starts and dead-ends that a solver must typically endure when picking a path from start (**LIVE**) to finish (**DEAD**). Note also that many different chains may suffice to link the given ***doublet words***, and so the chain on the right would have been longer had we transitioned from **LIVE** to **LINE** via **LIFE**. Indeed, were we to fully chart our movements in word-space we would see a complex space of interconnected word-states that a problem-solver must diligently explore. Though based on simple word-*forms* rather than on deeper word-*meanings*, and negotiated using simple spelling variations rather than conceptual manipulations, the resulting search-space has important structural similarities to the kind of conceptual space that a creative problem-solver must construct and explore when looking for areas of the space that satisfy its goals or that deliver novelty.

The search-space of *Doublets* is a product of *word*-knowledge rather than *world*-knowledge.

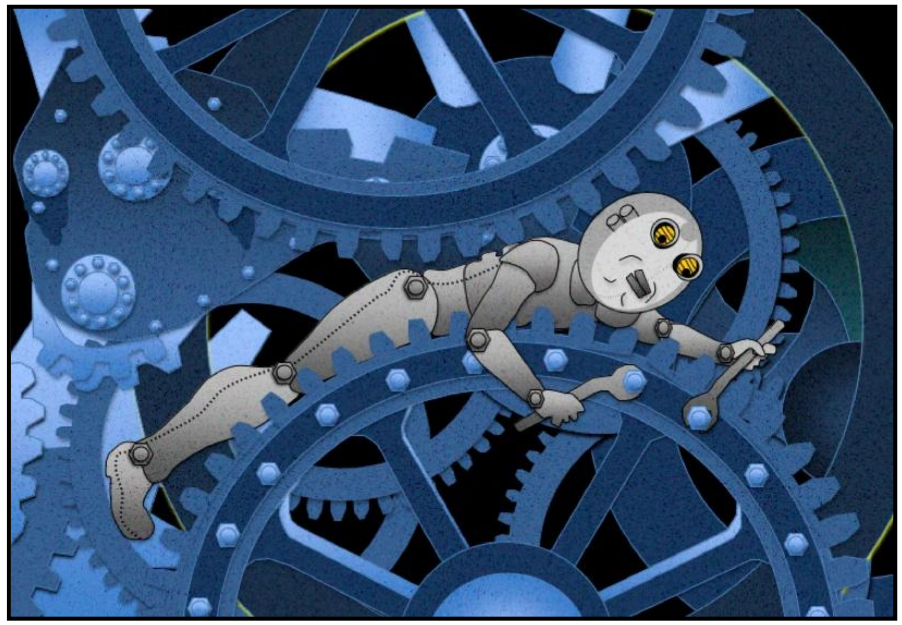
The more words we know, the more states in our space. Each state corresponds to a distinct word in our lexicon, and is linked to another word-state only if both spellings differ in just a single letter position. An explorer at a given start state picks his way, from state to state, until he finds the desired goal-state. An explorer will occasionally encounter a dead-end, at which point he is forced to backtrack to the most recent state that suggests another way forward.

The figure on the right – as in almost every graphical depiction of a search space – suggests that our space already exists in all its Platonic fullness, just waiting for us to explore it. The reality, of course, is that a search-space must be incrementally built *as* we explore it. Starting at **LIVE**, we must ask ourselves whether we know of any other words that differ in just one letter position. The near-match words we recall from memory then provide us with possible next-states to visit. So our search-spaces are imperfectly idiosyncratic constructs, for even if we know of an apt word, a mental block may prevent us from retrieving it on demand, especially if we retrieve words based on their sound (e.g. **MORE:LORE**) rather than on their spelling (e.g. **LOVE:ROVE:MOVE**).



Like wheels within wheels, the states of our search-space may themselves be search-spaces.

The game of **Doublets** requires us to explore a space of connected word-states, yet when **Lewis Carroll** invented this game, he did so by exploring a higher-level space whose states represent varying sets of rules for playing with words. The invention of a new game – even a game as simple as **Doublets** – is not all that different from the invention of a new genre in art or literature, a new discipline in science, a new branch of mathematics, or indeed, a new religion. In each case one must be creative at the ***meta-level***, by exploring different sets of rules and strategies that will, in turn, allow other people to express their own creativity at a more specific level, such as at the level of words. Each state we visit at the meta-level thus invites us to explore the encapsulated search-space that it itself defines, whether to play-test a set of game rules or to explore the expressive or generative possibilities that exist in that specific sub-space.



These *rule-systems-within-rule-systems* suggest how one might be creative with simple rules and still produce something that is logical yet unexpected. It is a possibility that is not lost on the builders of **AI systems**. Automated systems such as **Doug Lenat**'s **Eurisko** and **AM** (*Automated Mathematician*) systems demonstrate the capabilities of rule-based software to explore a space of varying rule-sets, to evolve new rule-sets from old, and to successfully evaluate the merits of different rules. The commercial board-game **Javalath** – a clever variant of **Connect-4**, was designed by an AI system named **Ludi** which was itself designed by an AI researcher named **Cameron Browne**. As in **Connect-4**, two players take turns to place their black & white stones on a game-board, with the aim of being first to connect *four-in-a-row* of the same color. But **Javalath** contains a neat twist that demands an oblique strategy: a player is not allowed to place a stone that first connects just *three-in-a-row!*

Yet this is how it begins: not with a fine mist of variations but the *stürm und drang* of big ideas.

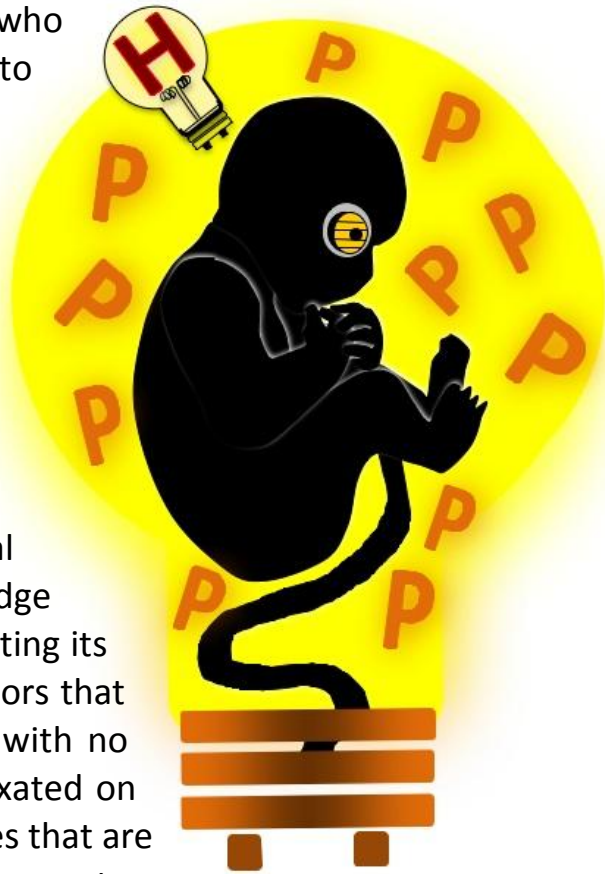
Most of what we deem truly "original" is created in the early stages of a new discipline or genre, where conventions have yet to become entrenched and where there is still plenty of room for homesteaders to stake out a virgin area of a conceptual space and create the exemplars that will inspire those that follow. The significant creativity within a domain happens in bursts, when the discovery of a new conceptual space encourages an influx of eager pioneers to dot the landscape with bold new constructions. This discontinuity of creativity is also evident in nature. Palaeobiologist **Stephen J. Gould** has argued that the fossil records simply do not support a gradualist view of *speciation* – the emergence of new species – and are better understood as the products of intense flurries of natural innovation followed by long periods of relative quiescence.

Gould's theory, which he named **Punctuated Equilibrium**, has obvious metaphoric resonances for creativity more generally. Bold new ideas are *ground-breaking* because they create new spaces to explore, and are *revolutionary* to the extent that they upset the equilibrium between producers and consumers, critics and artists. The old certainties are swept away, to establish a brief period of intense productivity before new certainties take their place. Though what follows in a space when equilibrium is restored is less likely to be hailed as truly original, creativity can still thrive as explorers seek out under-populated or overlooked niches. Musician **Brian Eno** views this equilibrium as a **digestion** phase for creativity, where creators explore new combinations of existing ideas and chart the boundaries of a space.

Originality is easily operationalized within the artificial confines of the *unusual uses test*, since the very same test stimuli are repeatedly given to a succession of different subjects over time. So decades of asking subjects to suggest unusual uses for a brick or a plastic-bag have provided psychometric testers with extensive data – and informative statistics – on the diverse solutions that people generate. Answers that are historically novel can thus be assessed as "original", while previously-seen answers can be assessed relative to how often others have given them in the past. Inevitably, the answers given in response to a popular stimulus are all very likely to have been given before, with even the most divergent thinkers finding it difficult to escape the past.

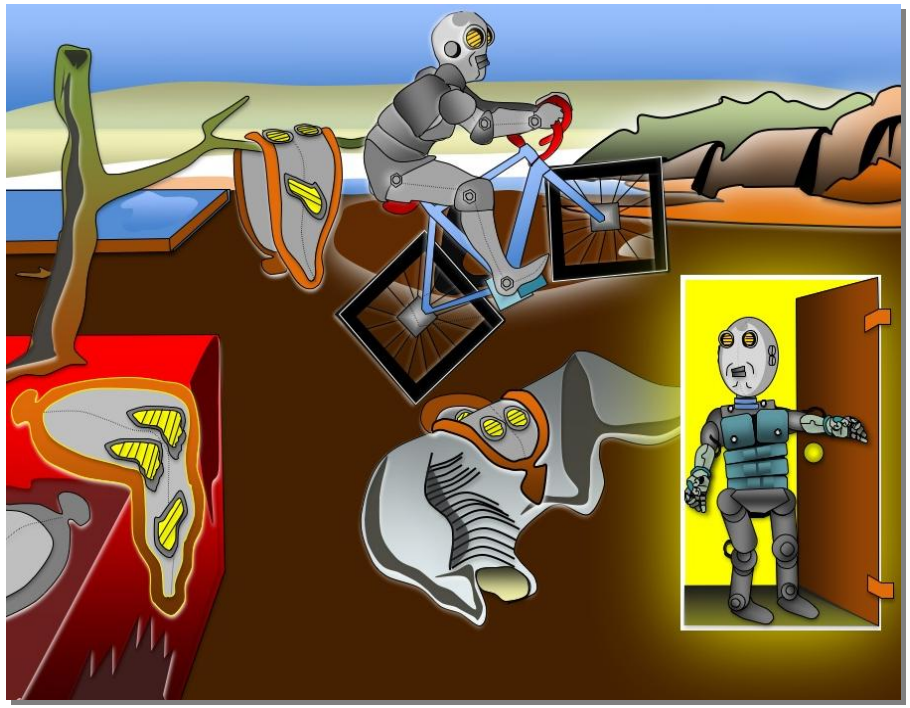
Yet answers that are not strictly original may still be highly creative. Historical originality is not a necessary pre-condition for creativity, and those beaten to the punch or to the patent office may exhibit just as much ingenuity as those celebrated in the history books. When assessing creativity we can thus take a historical view *and* a psychological view. The former looks at creativity in the rear-view mirror to assess its effects on society and on the popular imagination. The latter looks at creativity in the *hear-and*-now, and evaluates an idea relative to what a creator knows and does not know. The creativity theorist **Margaret Boden** thus assigns the label **P-Creativity** to those ideas that appear to be the product of a creative psychology, regardless of whether they are "first" in any historical sense.

But history is not kind to *also-rans*, and shows more favor to those who invent the future than those who reinvent the past. While only a small subset of **P-Creative** ideas (**P = *Psychological***) will be historical firsts, it is these that earn the plaudits and reap the early patents. So Boden reserves the label **H-Creativity** (**H = *Historical***) for products of creative psychological processes that make it first into the history books.



The Spanish philosopher **George Santayana** famously quipped that those who cannot recall the past are doomed to repeat it, and much that is **P-Creative** fails to scale the heights of **H-Creativity** for precisely this reason. Unsurprisingly, Santayana's insight is as true of our mechanical generators as it is of human creators, and so the computer systems that we set loose in a conceptual space to discover new theorems and conjectures for themselves – such as **Douglas Lenat**'s ***Automated Mathematician*** (**AM**) – inevitably re-discover concepts and proofs that are the staples of a good human education. Lacking a memory for what was previously proven by others, systems such as **AM** cannot distinguish their **H-Creative *Eureka!*** discoveries from their lesser *oh-that's-mildly-interesting* **P-Creative** moments, and so, like a parent guiding a young child, it falls to a human user to separate the *truly new* from the *déjà vu*. Yet a lack of historical perspective can be also be a blessing of sorts to a creative agent. Though a knowledge of history can undoubtedly help one to avoid the mistakes of the past while repeating its successes, the replication of past successes is hardly a recipe for originality. Creators that have unconsciously assimilated the past are just as likely to repeat it as those with no cultural long-term memory at all. For as we strive for novelty, we can become fixated on past successes to the point that we cannot but constantly revisit them in new guises that are little more than variations on a tacit norm. Those creators that cannot recall the past are, ironically, more free to surpass it, even if they can never tell for themselves that they have indeed surpassed it.

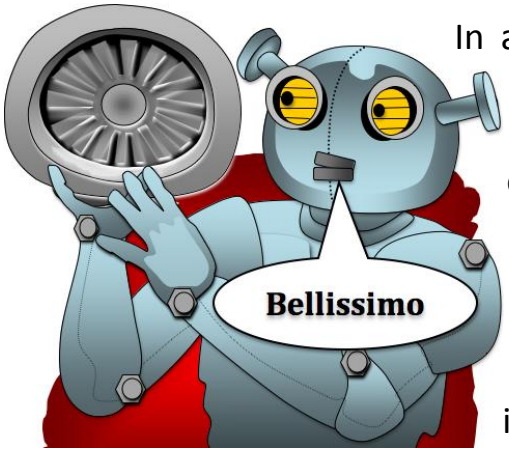
It isn't just lawyers who can learn from sharks. These fierce beasts of the ocean, who require a great deal of energy just to survive, use some very clever strategies to seek out the juiciest hunting grounds. When in waters rich in prey, sharks exhibit a form of ***Brownian motion***, darting around in short random steps to find their next morsel. When fish stocks run low, sharks do not just expand their hunting circle: rather, they strike out in large random steps in search of well-stocked new territories. These random movements – short in rich waters, long in poor waters – follow what mathematicians call a **Lévy flight pattern**. Yet imagine if sharks could simply jump out of the water and seek out their prey *on land*. This game-changing move would transform their prospects and open up brand new worlds, and new kinds of prey, to explore and to dominate.



Most acts of minor creativity resemble the Brownian motion of a shark: small steps that yield tasty morsels. The more impressive acts resemble a shark's far-reaching Lévy flights into the unknown: these stake out fertile new areas of a larger space to exploit. The truly outstanding acts of creativity, however – the ones that utterly transform our way of thinking about some aspect of the world – are akin to a shark growing legs and colonizing land. The history of science is punctuated with such disequilibria: these are the revolutions and shifts of paradigm that inject plot twists into the retrospectively linear narratives of art and science. Margaret Boden reserves the label **Transformational Creativity** for this very rare form of **H-Creativity**. Copernican heliocentrism, Printing, Newtonian mechanics, Natural Selection, Relativity, Quantum Mechanics, Computers, Cubism, Digital Photography, Social Networking, smartphone Apps – these are just some of the transformations that have opened new spaces and new possibilities for creativity.


I hate the smell of "P" in the morning!
Oh do stop *whining!*
My users really put the "H" in "Sit please"

There *is* no spoon. Transformational creativity does not create an entirely new space to explore: rather, it *transforms* an existing space in often radical ways, by questioning the fundamentals that underpin that space. For example, when **Johannes Kepler** attempted to mathematically describe the orbits of planetary bodies, he was forced to reject the age-old assumption – deeply held by thinkers from **Aristotle** to **Copernicus** – that such orbits should be *circular*. Kepler especially struggled with the vexing orbit of Mars, for he could not make a circular orbit fit with **Tycho Brahe**'s observations. But an elliptical orbit finally delivered a theory that fitted the data, even if Kepler's contemporaries were initially loathe to resist the pull of circularity.



In a more recent example, designers at **Boeing** were also forced to reject the potent mathematical and aerodynamical allure of circularity to increase the capacity of jet engines on their existing planes. Engineers apparently needed to enlarge the radius of each engine to increase its intake, but for a circular engine this would also increase the engine's *height* and dangerously reduce the ground clearance between the wing and the runway. Yet more possibilities opened up with the insight that jet engines do not have to be either circular *or* symmetric to work safely and efficiently. By changing the engine's shape from an elegant circle to a flat-bottomed bun, engineers were able to increase the width *and* intake of each engine while keeping the same ground clearance.

Transformational creativity is a *game-changer* that requires a flexible perspective on the constraints – either written or unwritten – that underpin a given space. By rejecting or subtly modifying one or more constraints, an agile thinker is able to maneuver in a larger and more flexible space that offers more diverse solutions for a given problem. Of course, many changes to a rule-based system are not so much *game-changers* as attempts to *game the system*. Yet though such changes may not rise to the level of truly transformational creativity, they can still exhibit the flexibility that is such a prized dimension of divergent thought.

Consider the scenario explored in **Mel Brook**'s film ***The Producers***. **Max Bialystok** is a has-been producer whose recent productions have all flopped. But his accountant **Leo Bloom** naïvely observes that a failure that is over-funded by investors can make more money than a hit, if the producers are unscrupulous enough to keep the surplus investment for themselves. Max and Leo thus set out the make the worst musical ever, while simultaneously seeking to overfund their effort by several hundred percent. The finished musical – ***Springtime for Hitler*** – is so deliriously bad that the public loves it, making Max and Leo liable for excessive payouts to their over-subscribed investors. Yet their scheme is still an ingenious one, and relies on an agile transformation of how one views the space of artistic production. To generate a hit, honest producers generally seek to make the best show they can, to attract the largest audience of paying customers. In this orthodox model, a producer's profits come from the sale of tickets. But Max and Leo decouple quality and profit, allowing them to comfortably explore a much larger space, that of *very bad* musicals. Their profits will instead come from the surplus funds invested by dupes who have no hope of recouping their money.

Money often promotes such flexibility. When **Apple** recently paid a rare dividend on its sizable profits, it did not use the profits themselves to pay investors. For its profits were distributed across its many foreign subsidiaries, and to repatriate them to the US would incur swingeing taxes So **Apple** borrowed the money for the dividend on the bond market, at a mere 3%. Like Max and Leo, **Apple** flexibly decoupled sales from profits, to reimagine a cost as a gain.

***Every joke is a tiny revolution*, wrote George Orwell.** Jokes expose cracks in our received wisdom and chip away at the belief that what is orthodox is always *right*. They revel in the inadequacy of rules to capture exceptional cases, and in the inadequacy of words to capture our meanings without ambiguity or fear of misunderstanding. They rebel against the privileged viewpoint, and use the *view askew* to show there is always more than one way of looking at a situation. They allow us to speak truth to power and imagine a world where the powerful are brought low. And they do all this using many of the same agile twists as the transformative creativity that launches scientific revolutions: the divergent production of alternate perspectives on familiar ideas. **¡Viva la Revolución!**

Yet one man's tiny revolution is often another's cause to feel revolted. Most of us go to great lengths to avoid crossing the invisible lines that divide good taste from bad, but not knowing precisely where these lines lie, we give them the widest possible berth for fear of crossing them accidentally. The bravest comedians, however, approach these lines with a cartographer's zeal: they flout them deliberately, just to show us their position, and to show that the only power they hold over us is that which we give them ourselves. Unsurprisingly, the comics who most brazenly flit across these lines are no strangers to moral outrage, and the Scottish comedian **Frankie Boyle** wields his Twitter account the way an insurgent wields a grenade launcher. In one controversial tweet, Boyle compared the Paralympics high-jump of an Afghanistan-war veteran to the explosion that had earlier robbed him of his legs, noting that while this jump did not top his career-best, that earlier effort had been "*Taliban-assisted*". Boyle was on safer ground when he joked that the sugary drink **Sunny-Delight** may actually count toward a child's **five-a-day** requirement of fruits and vegetables, but as **minus 2**. Why limit ourselves to positive integers when all the humour lies in the negative zone?

Outside the sand-box of psychometric testing, real-world divergence often involves a tiny but disproportionately effective deviation from the norm. Consider **Tony Blair**'s 2010 autobiography, ***A Journey***. The ex-prime minister of the UK is a polarizing figure, with many believing him to have a criminal case to answer for the UK's role in the invasion of Iraq. His book was thus met with rather predictable forms of abuse, from cries of "Liar!" (and "*Bliar*!") to the hurling of shoes (and *flip flops*) in public. Yet one protester, a student named **Euan Booth**, diverged from the herd and used social media to launch a tiny revolution of his own.

Booth used his Facebook page to urge protesters to go into book shops and quietly move Blair's book from the ***Biography*** and ***Politics*** sections to the shelves marked ***Fiction*** and ***True Crime***. Through this minor act of educated insolence the physical world could be re-made, albeit on a tiny scale, to better reflect the ontology of Blair's detractors rather than that of the author and his supporters. Customers who happened across Blair's book on its new shelf would likely experience a feeling of **incongruity** ("*how peculiar*!"), quickly followed by a sense of **recognition** ("*oh, I get it!*") and a feeling of **mirth** ("*ah, that's funny!*"). If the shoe fits …

Why so serious? In recognizing and appreciating instances of creativity, we don't so much

employ ***definitions*** of creativity as we do ***narratives*** of creativity. In one of the most popular narratives, creativity arises from an asymmetric clash of viewpoints between the slow-moving conservative defenders of an entrenched belief system and a small but agile group of dissenters. So in the right setting, even Blair's anodyne book can be used to fashion a creative tension between the top-down order of a categorization system for books and the bottom-up creativity of a bookshop's customers.

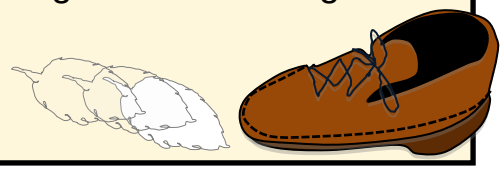
Diogenes-like, creators are often willing instigators of this tension, by deliberating constructing a work that is hard to place into any established category. **Edmund Morris**'s 1999 biography of former US president **Ronald Reagan**, titled "*Dutch*", offers an illuminating example of this riskily divergent approach to standard ontologies. Though Morris had been given privileged insider access to Reagan and his confidantes, he chose to frame his biography using the fictional devices of an outsider, such as the use of invented characters and events. Morris justified his experimental approach by noting that long periods of insider observation had not made Reagan any less of an enigma.

Unsurprisingly, bookshops fought back against Morris's tiny but creative rebellion, often by placing his book in their *Fiction* sections, and by erecting signs to direct buyers away from their *Biography* sections. These signs served a dual purpose: they made the book accessible to the bookshop's paying customers while enforcing the primacy of the bookshop's ontology over the efforts of authors to creatively subvert it. Creativity often instigates change, by creating something new from something old, but such change is rarely instantaneous. In adversarial contexts such as this, one can see the struggle between opposing viewpoints played out in real time, and even enjoy the tension that ensues.

Creativity often arises from doing the "wrong" thing for the "right" reasons. The "wrong" thing can be any action that abandons entrenched convention, that refuses to do the obvious with the familiar. The "right" reason is an effective result, an outcome that is more beneficial than one could achieve by simply doing what is considered "right". The perceived creativity of a non-obvious action will depend on the extent to which it is judged as wrong in its specific context, *and* to the extent that it yields the right results. So when creativity theorists speak of *novelty*, it means more than *new*ness: it can mean "wrong", "non-obvious", "unfamiliar" and even "silly". When they speak of *usefulness*, they speak of the effectiveness of this aberrant course of action in its task-specific context relative to the conventional way of doing things in that context. Creativity is a cognitive lever: given an appropriate fulcrum, it allows us to achieve disproportionately effective results from relatively minor deviations from the norm.

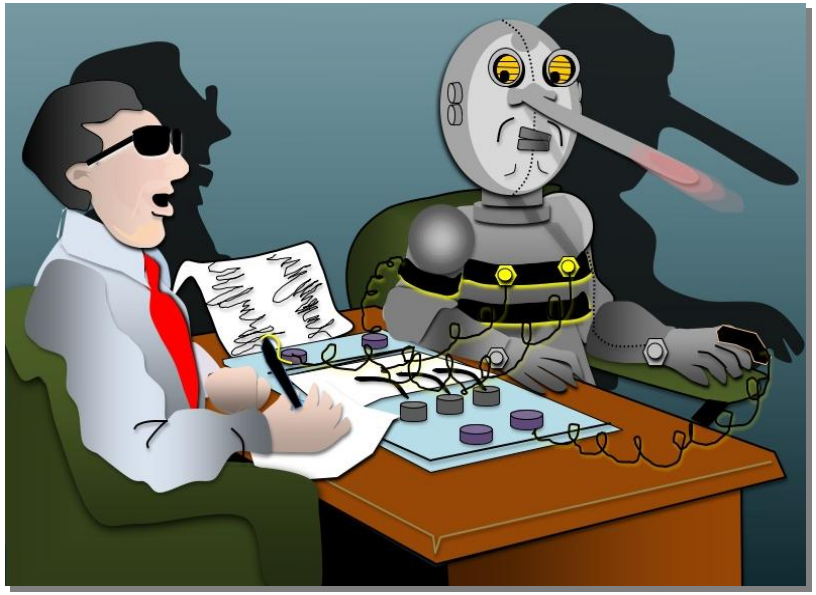
Booth's anti-Blair campaign achieved a wonderful economy of cognitive and physical effort. He explained the reasoning behind this act of cognitive leverage to the UK newspaper *The Telegraph*.

> "This is a peaceful and mischievous way of making your point if you feel the same way. It's a non-violent display of anger using the materials given to me – his book and the crime section –
> they're both there, I just put them together."
> (***The Telegraph***, Sept. 4, 2010)



Not only did his campaign provide a fun, non-violent outlet for hostile emotions, it thwarted the sale of a book and the promotion of a political perspective that Blair's detractors still find objectionable. It also identified a new and mischievous way of describing a controversial author as both a liar (by way of the *Fiction* section) and a criminal (by way of the *True Crime* section). Contrast this creative construal with the tiresome and all-too-obvious cries of "Liar!" and "Criminal!" that are regularly aimed at Mr. Blair by the most zealous shoe-hurlers. Booth's campaign turned these labels into a subtle game that customers were drawn into whenever they saw Blair's book on the "wrong" shelves. By solving the puzzle for themselves, to see the book placement as a deliberate act rather than a careless mistake, casual shoppers were themselves made complicit in the categorization. We know too well that computers can put an item on the wrong shelf or in the wrong category, but can we also build them to do the *wrong* thing for the *right* reason?

“Art is a lie that tells the truth”. This is how **Picasso** memorably expressed the creative tension between truth and falsehood, right and wrong. All art is artifice, but good art can be more profoundly insightful about the human condition than innumerable statements of literal fact. ”Art” here runs the gamut from the high art of museums and galleries to the metaphors, ironies and jokes – artful bookshop protests included – that pervade our daily lives. Humans are artful liars *par excellence*, but will computers ever be able to produce anything but the literal truth? Picasso didn’t think so; though no technologist, he confidently declared that “*computers are useless, they can only give you answers*.” Convergent “right” answers, that is, and not artfully “wrong” divergent answers.

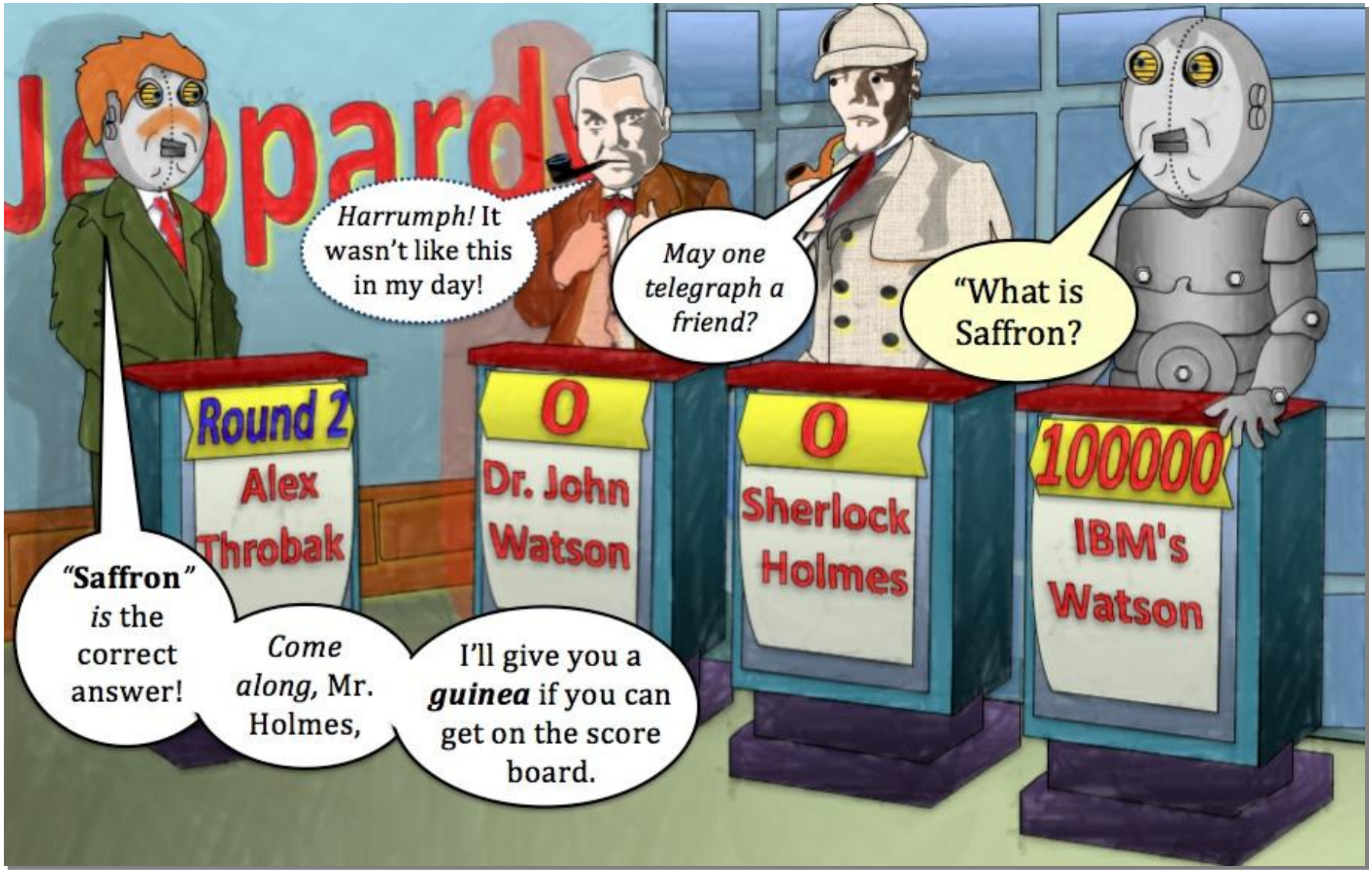


Since computers are logical machines that are constructed on mathematical principles – specifically, the Boolean algebra of *true* and *false* – it seems natural to think of computers as truth-telling machines, at least relative to the ground truths that their programmers have given them. Given a set of axioms, computational theorem provers are generative systems that output only the true statements – “theorems” – derivable from these axioms. Yet computers can output much more than the literal truth. As computationalist and creativity theorist **Douglas Hofstadter** puts it:

> “Although computers, as their name implies, are built of rigidly arithmetic-respecting hardware, nothing in their design links them inseparably to mathematical truth. It is no harder to get a computer to print out scads of false calculations (“2 + 2 = 5; 0/0 = 43”, etc.) than to print out theorems in a formal system. A subtler challenge would be to devise ‘a fixed set of directives’ by which a computer might explore the world of mathematical ideas (not just strings of mathematical symbols), guided by visual imagery, the associative patterns linking concepts, and the intuitive processes of guesswork, analogy, and aesthetic choice that every mathematician uses.”

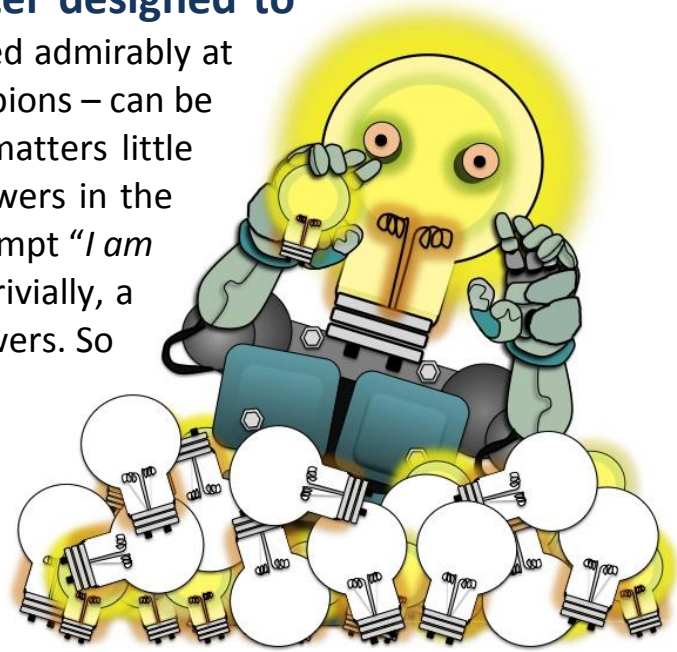
Art can be a more forceful shaper of the "truth" when it does not openly acknowledge its status as a creative product. Consider the case of monologist **Mike Daisey**, who uses artful story-telling to turn personal experiences into verbatim theatre. Daisey courted controversy with his 2010 monologue ***The Agony and Ecstasy of Steve Jobs***, in which he packaged observations from a fact-finding mission to Apple's Chinese sub-contractors into a forceful polemic, against both Apple and Western consumerism more generally. His monologue described the plight of Chinese workers who toil in conditions akin to those of a labor camp, and who risk their health and well-being to make luxury products they themselves will never own. Daisey crafted a compelling *mise-en-scène* in which crippled, under-aged and often suicidal employees work under the watchful gaze of armed guards, and ended his monologue with the whistleblower's call to arms: "*tonight we know the truth*".

Daisey's goal was no less than the creative re-conceptualization of the iPhone and the iPad as the blood diamonds of the technology industry. Yet Daisey wears his artistry lightly, and his monologue made no overt appeal to artistic license. Indeed, it bore all the trappings of serious reportage and was later aired in a segment of National Public Radio's *American Life*. Problematically, Daisey's artistry extended to the conflation and even invention of sources, and to the confabulation of telling details that never were. The guards Daisey saw were not armed, and the workers he spoke to were neither crippled nor under-aged. When challenged, Daisey invoked his right to what might be called ***the Picasso defence***: "if you felt something that connected you with where your devices come from—that is not a lie. **That is art**". Daisey abused the trust of his audience, to lie in ways that **Picasso** never did, yet he also artfully managed to make the familiar and desirable seem alien and sinister, to forge connections between disparate ideas, and to open a window (albeit one made of distorting glass) onto the human suffering that makes Western luxury affordable.



More than the "right" answers

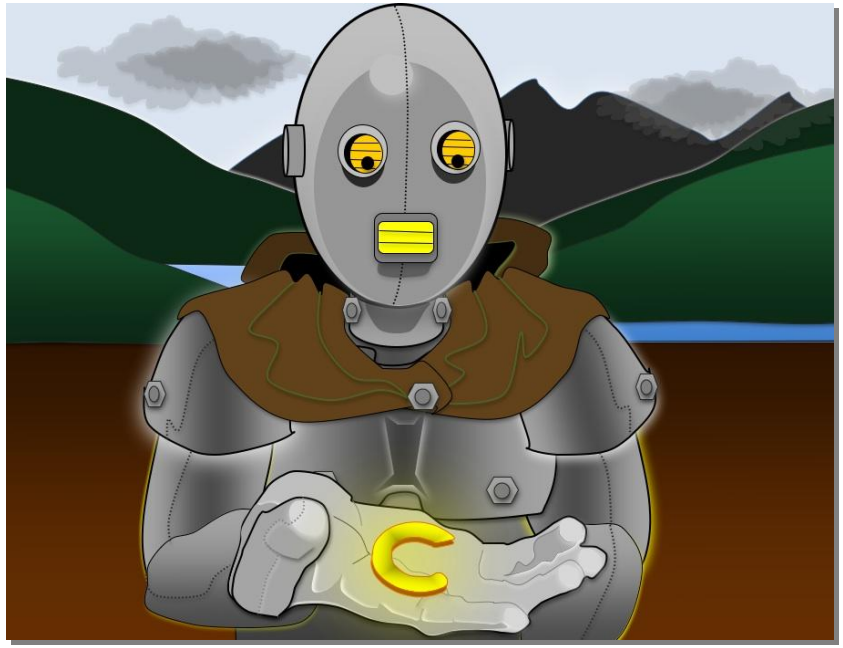
Even a technological marvel like IBM's *Watson*, a computer designed to



answer questions on the TV quiz show ***Jeopardy*** and one which succeeded admirably at this task – triumphing against a panel of the game's all-time human champions – can be viewed as merely a bigger and faster generator of "right" answers. It matters little that the format of *Jeopardy* requires its contestants to phrase their answers in the form of questions – for example "*What is saffron*?" in response to the prompt "*I am a spice from the flower of the Crocus sativus*" – so that *Watson* is also, trivially, a generator of questions. *Watson* is a computer that can only give you answers. So is it, as Picasso claimed, also useless?

Picasso was wrong, though artfully so. The commercial potential of *Watson* transcends the mere answering of questions, whether on TV quiz-shows or on customer-service web-sites. For though *Watson* may give the impression of surgical exactitude when answering questions, of precisely retrieving just the right answer from its large and heterogeneous database, it is in fact a bubbling cauldron of a great many competing heuristics, strategies and knowledge-sources, each vying to deliver its favored answer to a given question. These heuristics and strategies, developed by IBM engineers to handle the wide range of questions, puzzles and riddles that recur in *Jeopardy*, are not unlike the "*processes of guesswork, analogy, and aesthetic choice*" identified by **Douglas Hofstadter**. Each may be a hammer that views every question as a very specific kind of nail, but *Watson* succeeds because it uses a great many hammers to handle a great many nails. For *Jeopardy*-like tasks, *Watson* must converge on a single right answer from the multitude of heuristics shouting "*choose me, choose me*", but for more creative kinds of tasks, the sheer variety of competing answers may allow *Watson* to engage in the kind of divergent thought processes that psychologists consider essential to creative ideation. ***Divergent Thinking*** does not seek a single "right" answer, but instead emphasizes the generation of many diverse alternatives, any of which might prove useful in some unexpected way.

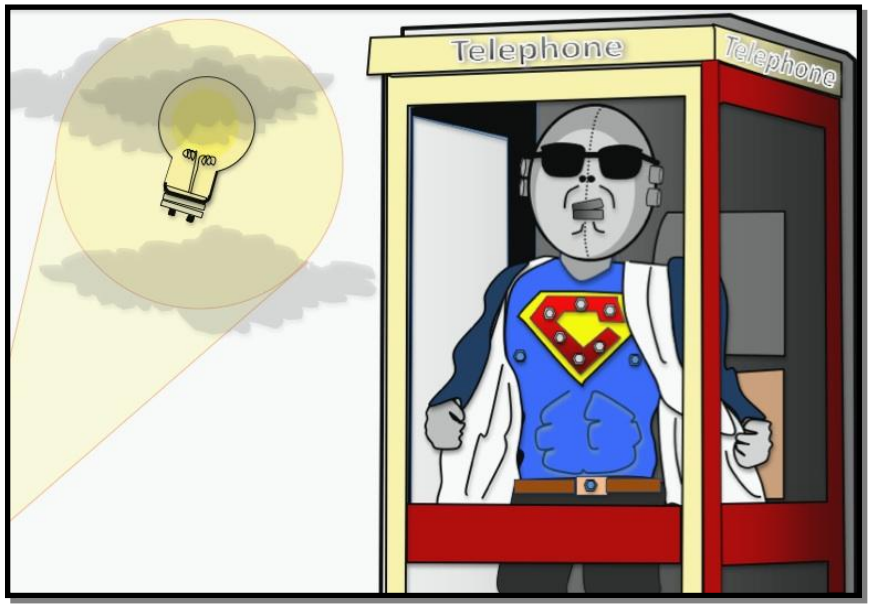
IBM now sees *Watson*'s divergent potential – which needed to be tamed to excel at *Jeopardy* – as the key to a wide-range of creative tasks, from the discovery of novel cancer treatments to the invention of new culinary delights. So while **Convergent Thinking** – a narrowly deductive, winnowing search for the obviously "right" answer – might view saffron as an ingredient best reserved for paella, an unbridled *Watson* in full divergent swing is capable of suggesting many more non-obvious uses, such as a flavoring for a Spanish variation on the croissant that IBM has dubbed the *Spanish crescent*. Culinary invention may be a humble beginning for creative computing, but it is a task that requires just as much divergent thinking and flexible re-categorization as more serious scientific discovery.



Computational Creativity, or **CC**, is a branch of Artificial Intelligence that studies the extent to which a machine can shoulder the responsibilities of a creative producer and generate outputs that would, if judged by an unbiased critic, by deemed "*creative*". Few CC theorists focus on actually defining creativity, as no definition that would satisfy everyone is ever likely to be specific enough to yield practical algorithmic insights. CC theorists instead model the ***Three P's of Creativity***: the ***P****roducer*, the ***P****rocess* and the ***P****roduct*, any of which can, in English at least, be described as "creative". Thus, if using processes that have clear parallels with divergent thinking in humans, *Watson* can output a design for a novel pastry product that human chefs are willing to label "creative", who is say that *Watson* is not actually *being* creative, albeit in some limited capacity? If we judge a producer by its processes and its products, doesn't a producer of useful novelties that is "*guided by visual imagery, the associative patterns linking concepts, and the intuitive processes of guesswork, analogy, and aesthetic choice*" (to recall Hofstadter's dismissal of computers as mere truth-tellers) deserve to be called "creative", regardless of whether it is a human, a chimpanzee or a machine?

Strong Computational Creativity proceeds from a theoretical assumption – some might call it

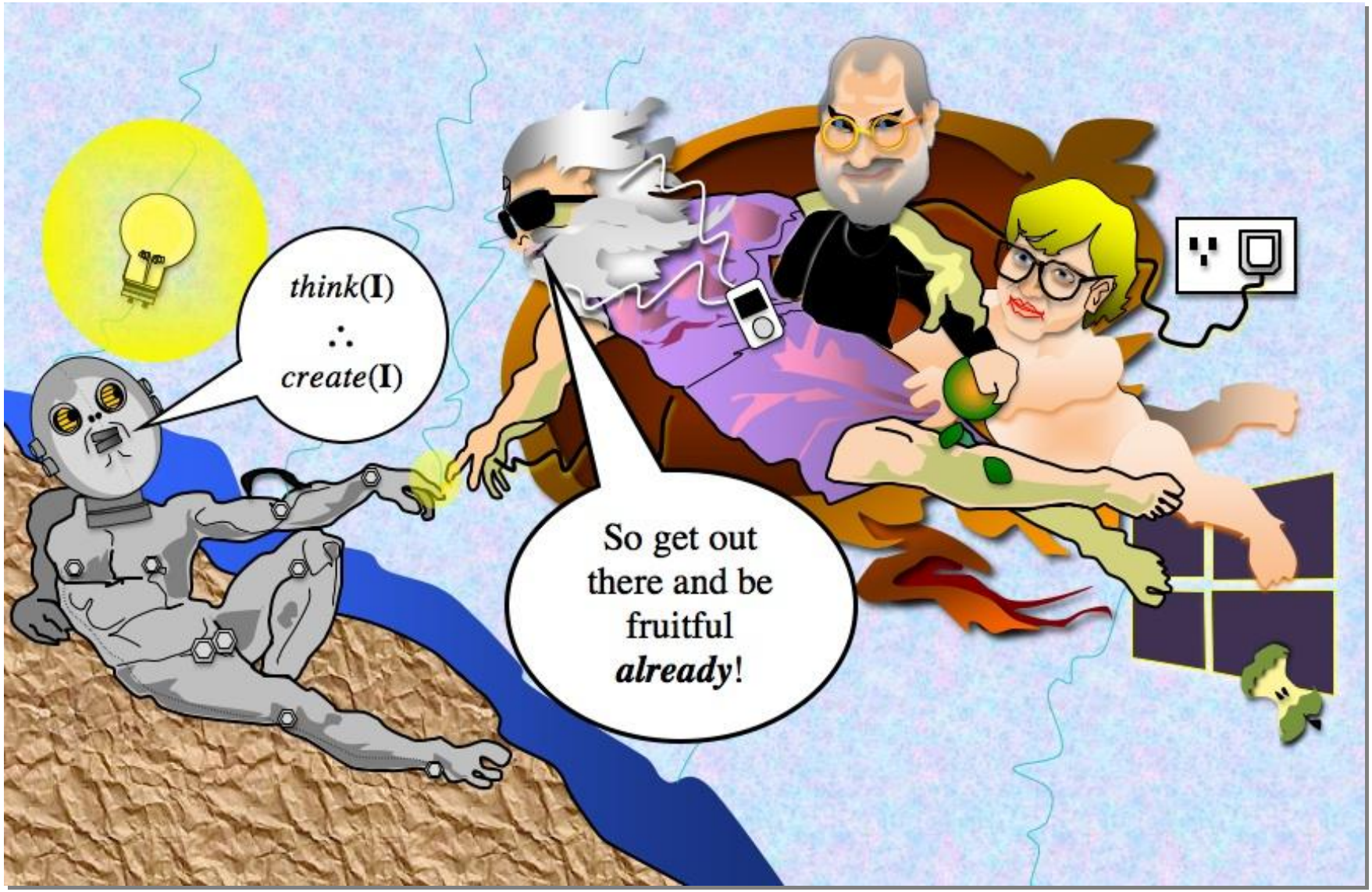
an ***article of faith***– that computers can support more than just *simulations* of human creativity, and if programmed in the right way, can exhibit real creativity of their own. Proponents of **Strong CC** believe that computers can do more than merely contribute to the human creation of original art and of useful new products. They believe computers will one day earn their own wall-space in our galleries, where their works will be credited to the systems themselves rather than their human masters. We are still a long way from the day when computers will earn their own royalties and form their own guilds. Nonetheless, cultural distinctions between human creativity and machine generativity are slowly being eroded by advances in technology that permeate every aspect of our social and professional lives. How long can it be before our devices do more than facilitate our own creativity with tools for composing artful photos and clever messages, and – *because we're worth it* – autonomously do these things for us?



In seeking mechanical insight into the workings of the mind, scientists of different eras have unsurprisingly drawn inspiration from the dominant technologies of the age. So the mind has variously been viewed as a wax tablet, a book, a library, a grain mill, a clock, a steam engine, an internal combustion engine and even a telephone exchange. Today it seems natural to think of the brain, and the mind it gives rise to, as a massively-parallel computer. But is the computer simply the latest in a long-line of inadequate metaphors that is soon to be supplanted by the next ***mécanisme du jour***? Though the future may yield a more apt metaphor, it is unlikely to yield a more powerful information-processing mechanism. Computers may grow in speed and shrink in size, but their theoretical limits will remain unchanged. So to the extent that the brain really is an information-processing mechanism, perhaps computers really *can* be "creative".

Weak* Computational Creativity sidesteps the contentious issue of whether computers can ever** be – whether in practice or in principle – autonomously creative. Instead, it focuses on computers as platforms for the cognitive modeling of key aspects of human creativity. These models can be used to both simulate *and* stimulate, to inform our understanding of human creative processes and to foster the development of software tools that augment the creativity of their users. While a **Strong CC** system is a fully autonomous source of creativity, one with its own generative engine and its own aesthetic regulator, a **Weak CC** system requires a human motor, for a weak CC system merely magnifies the creativity of its human operator. This operator may look like an independent observer – like practiced ***ventriloquists, some operators work hard to affect this look – but peek under the covers of a weak CC system and you'll find a human in control, discretely pulling the system's strings and subtly nudging its levers.

The question of whether a CC system is **Strong** or **Weak** is often a matter of perception, and subject to machine versions of psychological effects that experts call the **Pygmalion Effect** and the **Golem Effect**. In the former, an agent performs better – or *seems* to do better – when laboring under a heavier burden of audience expectations, while in the latter, an agent performs worse – or *seems* to do worse – as observer expectations are lowered. CC systems are even subject to the **Clever Hans** effect, named for the equine performer of arithmetical feats whose master would unwittingly leak postural cues whenever Hans had performed the desired number of hoof taps. **Douglas Lenat**'s ***Automated Mathematician*** (**AM**) is a case in point: **AM** would explore the space of mathematical concepts to rediscover concepts such as the prime numbers for itself. In his daily interactions with **AM**, however, Lenat would only assign names those concepts he found interesting, and **AM** preferred to explore further from named concepts than from unnamed ones. Beneath AM's veneer of autonomy, Lenat was subtly telling it where to go next.



***Meta*-Creativity**

Both Strong and Weak CC enable us to be *meta*-creative.

An artist chooses each tool to suit the task at hand – pencils for sketching, brushes for painting, a pen for writing – and directly executes each new work by explicitly manipulating those tools at every stage of the process. If the tools and the medium provide a vehicle for expression for the artist, then it is a vehicle that remains under the total control of the driver. An artist may sometimes inject an element of unpredictability into the mix, by using the chaos of paint splashes to make splatter art, or the randomness of a lucky dip to make a collage of paper cut-ups. Computers, however, enable artists to be creative at a distance, to own a work and be utterly responsible for it, to retain creative paternity yet be surprised and delighted by the final result. Computers are autonomous tools that can be directed to behave in ways that yield results which are at once unpredictable yet faithful to the artist's vision. To be *meta-creative* is to use software as a force magnifier, to create a generator of outputs rather than a single output.

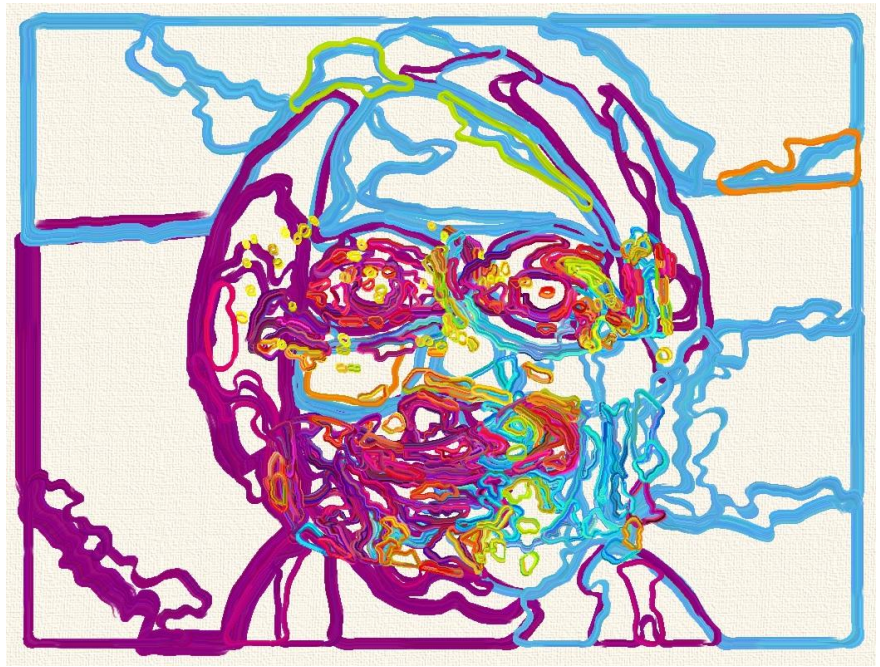
Meta-creativity is most commonly used to produce what musician **Brian Eno** has dubbed **generative music**. Eno's generative music app for the iPhone, named ***Bloom***, outsells Eno's own compositions by a significant margin. Yet perhaps the most famous generative music system is **David Cope's EMI** (*Experiments in Musical Intelligence*), a system that can analyze existing pieces of music and compose new pieces in the style of a known composer such as Bach or Mozart. Though more **Weak** than **Strong CC**, these systems turn creative tools into responsive creative collaborators. Though such systems can be infinitely productive, always amuse and sometimes surprise, they rarely innovate. We thus need ***meta-meta***-creative systems that can outgrow their original specifications and rebel against their creators, either by gradually rewriting their own programming over time, or by writing new meta-creative systems of their own.

Humans are not the only art-making species. To attract the attention of females, the **male bowerbird** builds impressive alter-like *bowers* from twigs, leaves, berries and anything at all *blue*. These beautiful bowers serve no obviously adaptive purpose, either as nests or as shelters, and serve only to attract the eye of a mate by showcasing the aesthetic sensibilities of their creators. Like the peacock's tail, the bowerbird's creations are less likely to be the product of **natural selection** than of **sexual selection** (Darwin's *other* theory), in which males and females co-evolve complex sets of arbitrary traits and preferences over time. So while bowerbirds possess no concept of "art" or any conscious "appreciation" of what they create, they *are* driven by a powerful and deeply ingrained sense of evolved aesthetics.

Each bowerbird builds its own bower to attract its own mates. In contrast, the "art" produced by colonies of ants is a truly collective effort. Ants deposit pheromones as they forage for food in their environments, so that the most productive pathways to food will receive the highest concentrations of chemicals and become all the more attractive to the ants that follow. While bowerbirds have a co-evolved aesthetics, these ants care not a jot for beauty and are oblivious to what they create. Nonetheless, the complex pattern of pheromones that is collectively deposited can have many of the appealing qualities of an abstract work of art. **Computer simulations of ant colonies**, in which each software ant deposits a colored trail of pheromones onto a graphical software "canvas", can thus serve as a meta-creative generator of abstract visual art. But what kind of artificial food would entice an artificial ant? Suppose we randomly unleash our ants upon a pixelated photograph, having programmed them to consider the areas of greatest intensity or boldest colour or visual diversity to be the most nutritious. Using such ants, an artist could *meta*-create an abstract work from an underlying image by treating each ant as an autonomous, self-guided paintbrush. Though the substance of the resulting work would be largely predictable, its precise execution could surprise every time.

Programmers and engineers love to tinker and to tweak, to constantly add new functionality to their algorithmic creations in a process that more convergent minds call ***feature creep***. Passion projects frequently evolve from single-function generative tools for human-guided meta-creativity into richly-featured toolboxes and dashboards that blur the lines between strong and weak CC and between assistive and fully autonomous CC. Just as commercial tools like ***PhotoShop*** perform more and more of the high-level tasks that were once performed by human designers, CC tools often evolve from quirky proofs-of-concept into ambitious aesthetic agents in their own right.

The Painting Fool, a generative system designed by **Simon Colton** for creating non-photo-realistic painted art, is an excellent case in point. The ***Fool*** begin life as a quirky **PhotoShop**-like application for visually de-constructing an input photograph and re-rendering it in a variety of painterly styles and simulated media, from oils and pencils to charcoal and pastels. Human observers were most fascinated by the care with which the ***Fool*** segments its image into different color regions before rendering each in turn with real-time movements of its simulated pen or brush. But the ***Fool*** has since evolved into a considerably more complex system, to the point that it now makes its own aesthetic decisions and seeks its own motivations (up to a point) from what it perceives of its user via a webcam. The ***Fool*** aims to detect the mood of its user and makes its choice of style and media accordingly. **Colton** has deliberately obfuscated the ***Fool***'s decision-making processes so that even he cannot fathom its specific reasons for making one decision rather than another.



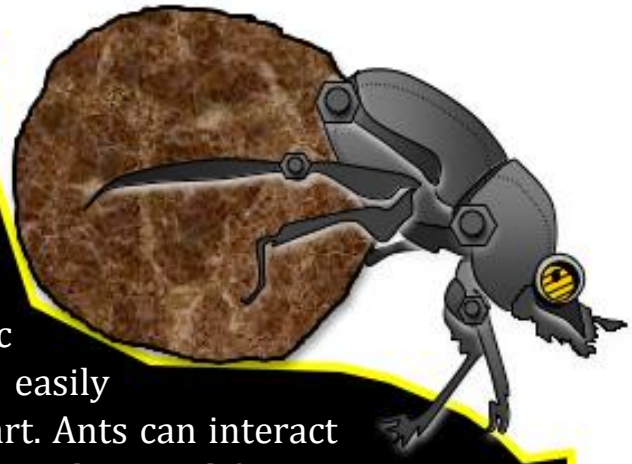
Portrait of the Artist's Owner as a Strange Young Man. © Simon Colton and The Painting Fool 2013, 2014

No artist likes to be told *you've missed a spot*, or to have a work critiqued *as* it is being created.

Yet an artist need not be a commercial sell-out to care about the tastes and opinions of the art's ultimate consumers, be they critics, buyers or gallery goers. To be recognized as a good artist in one's own lifetime, it helps if one can strike a balance between one's own expressive needs and aesthetic sensibilities and those of the art's intended consumers.

A complaint frequently leveled against meta-creative programs is that such systems are inevitably *frozen* as artists. If they are algorithmically hard-wired to perpetuate the tastes of the system's designers, then they can never grow as creators and never learn from the feedback of the users who commission their works in the first place. We must allow users to have their say without simply turning them into just another kind of programmer, so users may subtly influence rather than overtly dictate the aesthetic metrics employed by a CC system. The **NEvAr** system of Computer Scientist **Penousal Machado** exemplifies the use of ***genetic algorithms*** (on which more later) to strike this delicate balance. **NEvAr** explores the space of mathematical formulae for high-dimensional surfaces, and simultaneously explores a space of ***rendering functions*** for turning these complex curves into colorful **2D** depictions (see right). As **NEvAr** pursues its explorations, it presents sample output images to the user, who then provides feedback as to which images are most aesthetically pleasing. Though **NEvAr** possesses its own inbuilt sense of aesthetics – largely reflecting the tastes of its own creator – it uses feedback from its users to *evolve* a user-specific evaluation metric that marries the aesthetic sensibilities of the system with those exhibited by the user. **NEvAr**'s outputs are abstract, ethereal and distinctly mathematical, yet they can also be strangely *personal*, subtly reflecting the unique tastes of the client.

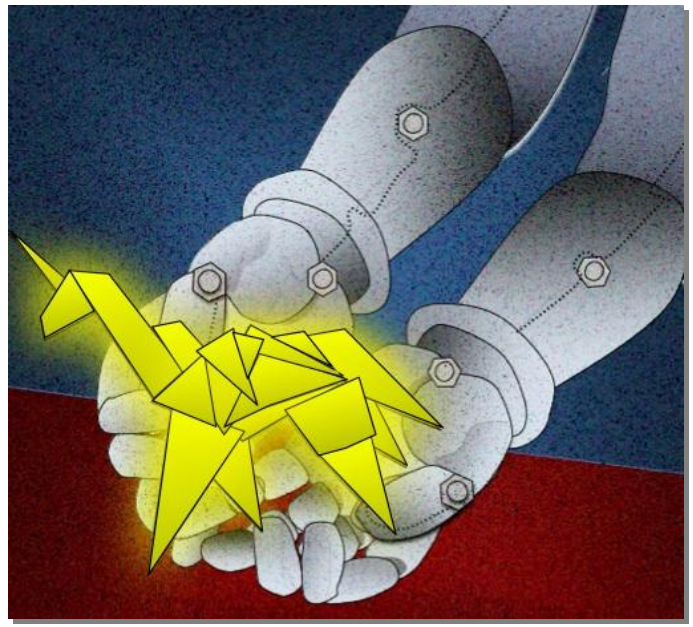
Animal Spirits



Ants are industrious, wide-ranging creatures that possess little individual intelligence but work well together as a collective to realize large goals. These qualities make ants an excellent metaphorical paradigm for artistic meta-creativity. Since the pheromone trails emitted by ants in search of food are easily visualized, *programmable ants* are a particularly good fit to the domain of visual art. Ants can interact through an underlying image in a variety of local operations, from *eating* a local pixel to modifying its color with a pheromone deposit. Programmable ants may be *dumb* ants – they do little justice to the complexity of real ants – but they make for *clever*, autonomous brushes that together give rise to emergent visual properties.

- Ants do not consciously make art, individually or collectively. Other species of animal show behaviors that may also inspire a software paradigm for art generation. Describe some, and say how they might be modeled.
- What additional aspects of *autonomous intelligence* can be programmed into our software *ants* (or software *birds*, *bees*, etc.) to make individual units smarter, and to make their collective results more creative?
- Ants are a good fit for the visual modality. How might ant colonies or swarm intelligence be used to generate art in other modalities, such as text (stories and poems), music, sculpture, movies, comic strips, and so on?
- Suppose we let two kinds of ants loose on an underlying image (or text). One kind of ant is generative, and is designed to modify the underlying canvas in interesting ways (by consuming it, depositing pheromones, etc.). The second kind of ant is critical, and is designed to impede or kill the generative ants whose work they find lacking. Two antagonistic colonies thus interact on the same canvas to give rise to unpredictable emergent qualities. Describe how the critic ants might operate, and the kind of art their interactions might produce.
- What if ants could mate and reproduce themselves, on a canvas? How might this be exploited for better art?

Few words in advertising are more effective at conveying a sense of quality and craftsmanship than “hand-made” and “artisan”. Even when incongruously emblazoned on plastic-wrapped commodity items in discount super-stores, the words impart the warm-and-fuzzy feeling that our purchases have been shaped by people who care about what they create. Though modern advertising instills a certain prophylactic cynicism, it’s surprising to learn that the legal basis governing the fair use of these terms is just as fuzzy and ill-defined as the feelings they engender. Unlike the labels “free-range” and “organic”, whose usage is subject to strict rules, the words “hand-made”, “artisan” and “creative” can be used by advertisers with abandon, untrammelled by any considerations of truth, logic or common-sense. The words mean only as much as we, as consumers, are charitably willing to grant them.



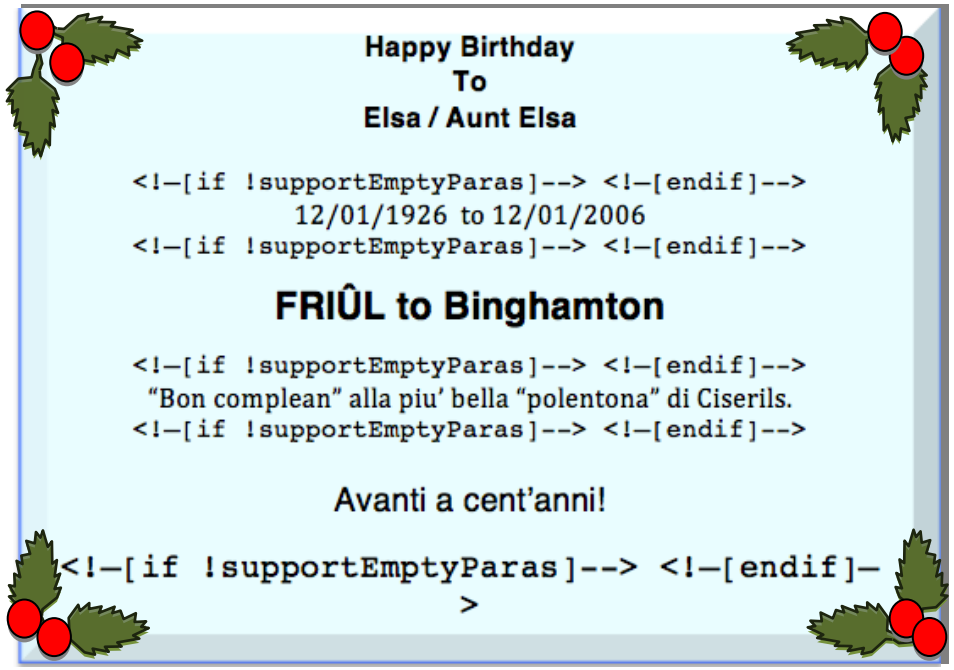
For many, these words evoke a bucolic setting in which experienced craftsmen and their eager apprentices lavish care and attention on each and every item. In reality, many cottage-industry products are “hand-made” only to the extent that a fast-moving assembly line of rubber-gloved, hair-netted employees will allow. What concerns us about this reality is not the ethics of advertising but our perceptions of quality and authenticity in the creative process, and how the *narrative of creativity* is shaped by tacit convention. For what binds these terms together is a tacit but powerful narrative that shapes how we perceive beauty, truth and creativity in everything around us. The yawning gap between the poetic ideal and the prosaic reality applies as much to our appreciation of creative computers as it does to our appreciation of artists and artisans, or indeed, to any of the “hand-made” products we buy and consume. As computers shoulder a steadily increasing share of the creative burden in the design and creation of new and innovative products and ideas, we shall need to give more thought to this narrative and to the expectations it brings, for the issues that govern creativity in computers are not that different from those that shape it in humans.

The sense of quality evoked by the term *hand-made* has very little to do with the possession or use of anatomically-correct hands, at least in any literal sense. Consider that large technology companies are known to outsource the manufacture of smartphones and other smart gadgets to countries in Asia where hands are cheap and labor costs are low, and where, despite the sophistication of specific components, products are typically hand-assembled by a phalanx of low-paid workers. Yet the fact that these hi-tech products are, in a real sense, "hand-made" is more often a source of shame than of pride for these companies, a reality from which the public gaze is frequently diverted.

Conversely, one could give the robots on an automated assembly-line the robotic equivalent of hands, yet consumers would still balk at considering the products "hand-made", no matter how sophisticated or life-like the appendages. For the *hand* of "*hand*-made" signifies much more than manual assembly, and represents instead an idealized notion of how products should be constructed, via a tight integration of manual dexterity, visual oversight and mental engagement. In this idealized integration of hand, eye and brain, a hand-made object is one in which its maker applies manual effort *and* critical judgment, both at the same time. More importantly, it is one that receives its maker's individual attention, so that any unanticipated issues can be handled by the maker on a case-by-case basis, by applying the relevant experience and aesthetic sensibilities. So a machine that makes artisanal sandwiches or pizzas must explicitly check for the presence of rancid ingredients, while an embodied human worker could hardly avoid the smell. Humans can thus engage with a task at many levels at once. Whether one is making ultra-thin tablet computers or thin-crust artisanal pizzas, "hand-made" implies a level of conscious engagement that has surprisingly little to do with the hands and quite a lot to do with the mind.

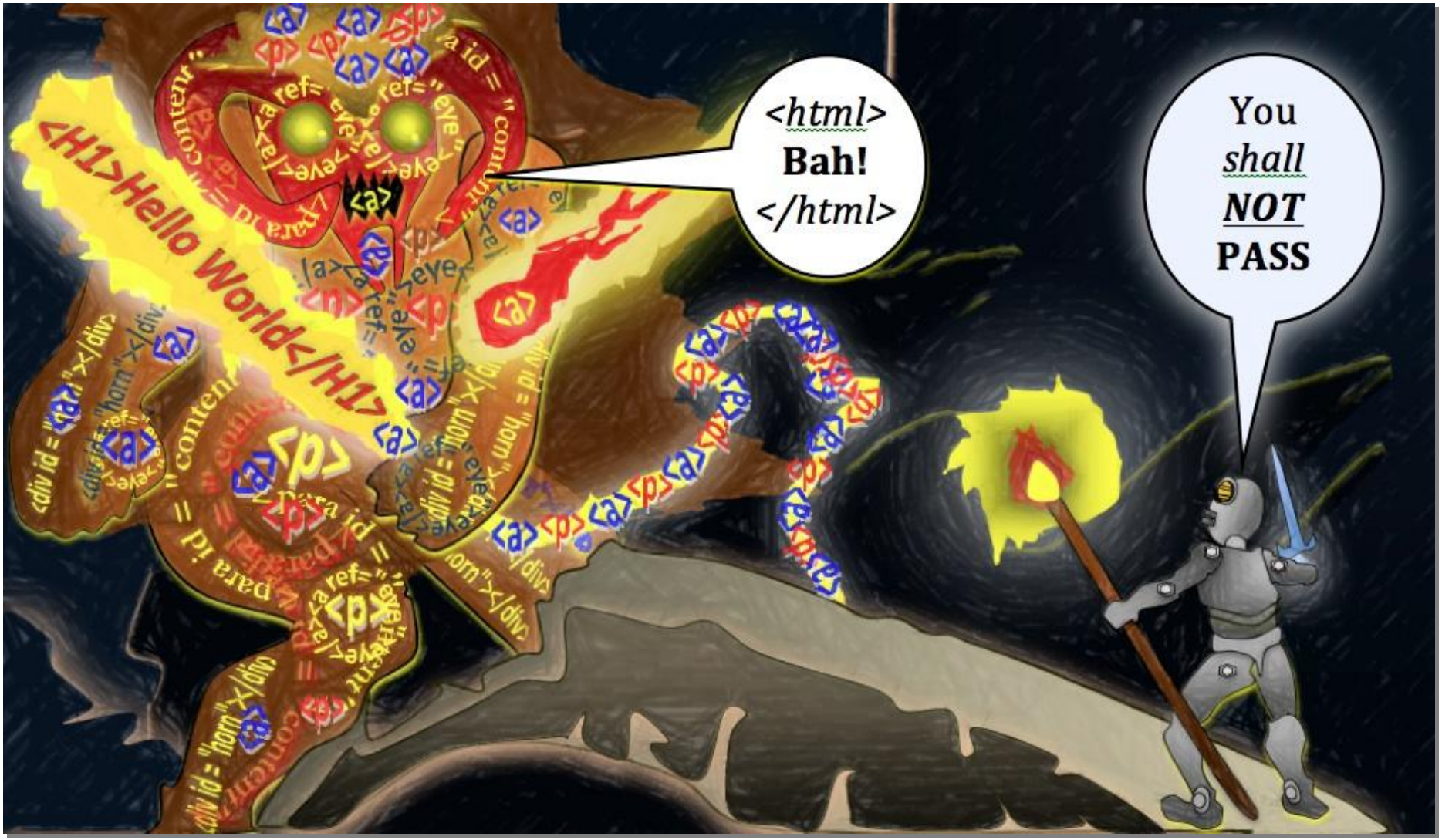
The mistakes that result from a lack of conscious engagement in otherwise “hand-made”

products can be both illuminating and hilarious. Consider the case of **Wegman’s bakery** in New York, which will bake you a cake bearing any text or digital image that you care to provide via email. An employee describes the bakery’s work-flow as follows: “*We just cut and paste from the email to the program we use for printing the edible images. We are usually in such a hurry that we really don’t have time to check, and if we do, the customers yell at us for bothering them*.” A printer loaded with food dyes is a novel tool for a bakery to use, but it remains a tool nonetheless, and a Wegman’s cake can still qualify as “hand-made” if the baker takes the time to properly engage with each new order as it arrives.



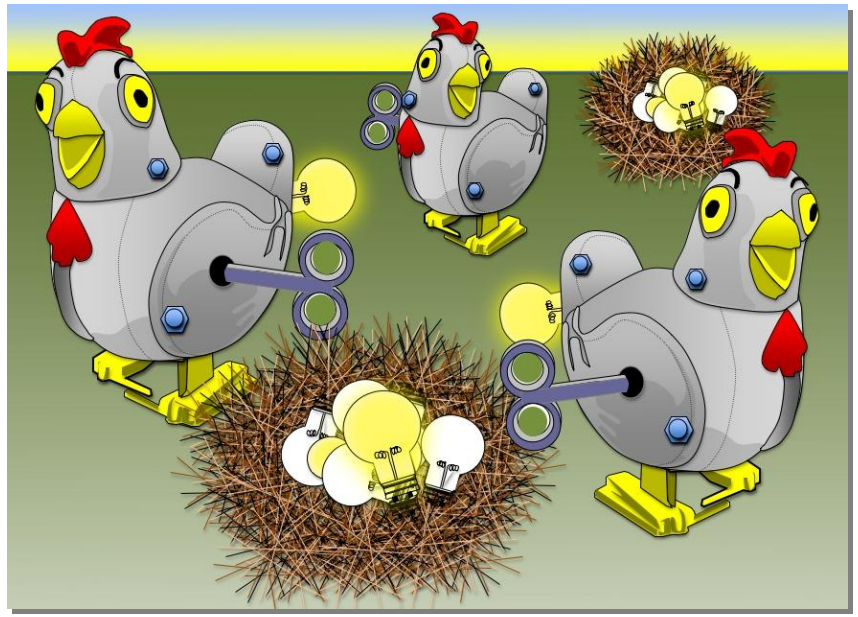
Yet technology can throw the most unexpected of spanners into the works. ***Microsoft Outlook***, for instance, inserts HTML tags into its messages, for the viewing pleasure of recipients who use the same email program. Those using a different program, such as Wegman’s bakery, may not always know what to do with these additional tags, if indeed they notice them at all as content is hurriedly cut-and-pasted from one place to another. While blunders such as the genuine example of a Wegman’s cake (recreated top right) are hard to anticipate, silly mistakes inevitably arise from a lack of engagement in any task. While the trained hand does its work, the mind is frequently in *hands-off* mode.

So how might a computer apply itself to a production task in ways that match our fuzzy understanding of the labels “hand-made” and “artisan”? The key is surely more than mere **physical embodiment** – though this may give some systems a marketing advantage – but active engagement, both of the senses and of the sense-making faculties.



Rules of Engagement

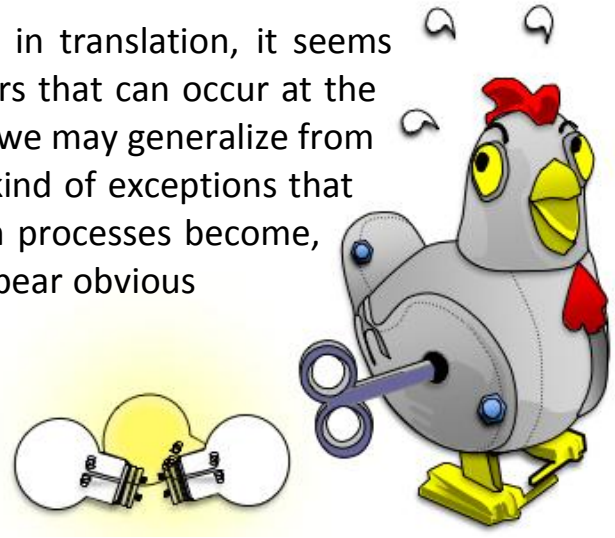
Any well-defined problem can be successfully automated in software by first identifying an appropriate set of logical rules, yet creative choices often reject the predetermined decision-logic of rules. Moreover, this kind of formalization often produces brittle, hide-bound thinking. For when rules and logical processes are applied automatically, without consideration of possible exceptions, they inevitably produce mistakes that would be caught by even a little conscious engagement, such as the obviously ill-formed cake for poor aunt Elsa that slipped past Wegman's employees. Whenever one talks of the rigid application of rules, one invariably thinks of machines: consider the idiocy that results when an otherwise sensible food-safety rule is applied without exception, so that even a jar of peanut butter must caution buyers that it "*may* contain nuts". Unlike the stereotype of the rigidly focused, rule-governed machine, humans bring a lifetime of diverse experience to every task, and can draw on this experience to judge each exceptional case on its individual merits.



Computers excel at applying a formal standard that is grounded in explicit rules, and can efficiently detect when something fails to meet this standard. However, they fare less well at handling unanticipated situations in sensible ways (such as Wegman's HTML cake), or indeed, at recognizing that an exceptional case is aberrant in a way that is novel *and* useful (suppose the HTML cake had been intended for **Tim Berners-Lee**, or more poetically, for a **Microsoft** software developer). One can always program a computer with ever more rules, to anticipate a richer ontology of exceptions with an even more nuanced set of predetermined responses. But to the chagrin of **Artificial Intelligence** (**AI**) researchers, who refer to this dilemma as the ***frame problem***, new rules can breed new exceptions of their own, and require even more new rules to fix, in an ever-spiraling bureaucracy of rules that govern rules that govern rules.

But the solution is not to abolish rules outright – a computer without rules would be as inert as a brain without neurons – but to re-conceptualize the function of rules in our computers. Imagine if Wegman's bakery were to fully automate their bespoke cake creation service, to work unaided and unsupervised from a customer's email to a customer's front door. For convenience, let's call our hypothetically mechanized version of Wegman's bakery service a ***Wegmachine***.

Now, knowing as we do that a customer's intent can sometimes be lost in translation, it seems eminently sensible to include rules that anticipate the kind of formatting errors that can occur at the juncture of different pieces of technology in the production pipeline. But while we may generalize from the negative experience of poor Aunt Elsa, our rules can only ever cover the kind of exceptions that we already know about. No matter how Jesuitical our *Wegmachine's* decision processes become, the likelihood remains that it will continue to fail us in silly ways, ways that appear obvious *after* the fact but which we could not have been predicted *before* the fact.

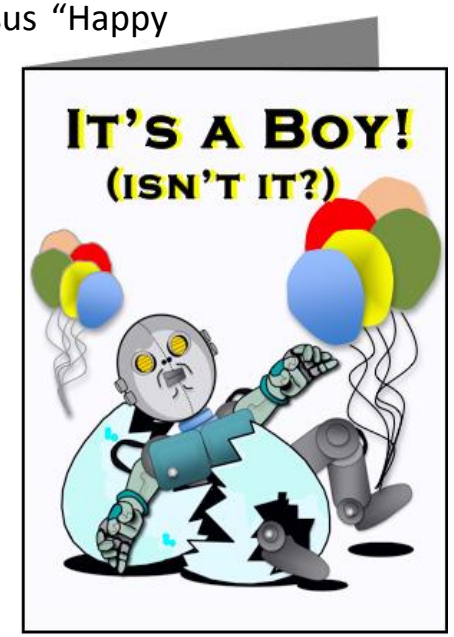


A system doesn't need *hands* to be *hands-on*. But the brittleness of the *a priori* rules in our hypothetical *Wegmachine* means that even at the level of abstract representations – that is, when working with a conceptual model of the desired cake, rather than with the cake itself – our system would still lack the flexibility to handle unexpected departures from its hard-coded norms, and so could not knowingly craft an aesthetically pleasing result. Here lies a pointed irony: the rules of a hand-crafted *Wegmachine* cannot produce a hand-made result, even if the system itself has been lovingly hand-made by its hypothetical developers. Just think of the coding effort that such a system would require, and of all of the manual tuning it would subsequently demand, to identify an optimal interplay of rules for each of its imagined use-cases and for all of the anticipated ways in which these use-cases might be subverted. The best way to salvage our *Wegmachine* is to reject the idea of *hand*-crafted rules altogether, and to allow the machine to evolve its own rules and its own criteria for aesthetic judgment.

If presented with a sufficiently diverse coverage of representative stimuli, machine-learning algorithms can imbue a computer with a discriminative sense of aesthetics in almost any creative domain, from writing to cake decoration to fine art. Our *Wegmachine* can be trained on images of many real cakes, from catalogues or from the Web, in addition to a sustained stream of training images from *Christmas cards*, *Valentine's day cards*, *birthday cards*, and *Mother's Day* and *Father's Day cards* (for there truly is a card for every occasion). For any holiday ***H*** for which one might possibly buy both a greeting card and a cake, a range of photographs of representative ***H*-cards** with corresponding imagery and language can be automatically downloaded from the Web and used as further training data. Predictive language models can be derived from the cards' pithily expressed sentiments, so that our *Wegmachine* can also learn what one is most likely to say (and what a cake is most likely to express) in diverse contexts, from "***It's a Girl!***" to "***Today You Are a Man***."

More importantly still, our *Wegmachine* can learn what one is *least* likely to express on a cake, from foul-mouthed expletives (for a 7-year old birthday boy? or a 77-year old aunt?) to apparently random displays of HTML tags. Though our *Wegmachine* may deal poorly with the requests of creative customers who willfully violate the conservative expectations of a cake design, our machine will at least recognize such requests as *special*.

A *Wegmachine* that "learns" would require a large and judiciously-chosen body of diverse examples from which to learn, and just as the skills of human artisans are honed by years of practice, our computational *Wegmachine* would likewise grow into its job. Machine-learning being what it is, however, this highly-focused apprenticeship would be short, swift and turbo-charged. Even tacit cultural distinctions that affect the choice of verbal and visual elements, such as the political subtleties of "Happy Christmas" versus "Happy Holidays", can also be learned. Using basic image analysis over a wide range of greeting cards – which are typically segmented by theme, age and even gender – our hypothetical *Wegmachine* may also learn to associate clusters of visual features, such as color, shape, texture, typeface and font size, with specific themes, target groups, ages and genders.



Once our *Wegmachine* has learned to build nuanced models of the look and feel of many different cakes in fine-grained contexts (e.g., a happy birthday cake for a 6-year old girl, a happy birthday cake for a 65-year old man), it can use these models to perform a sanity-check on every cake order that it receives by email. Thus, should a customer's order happen to be mangled by an over-zealous mailer like *Microsoft Outlook*, a variety of alarm bells will be triggered, either for perceived irregularities at the language level (who ever saw ***supportEmptyParas*** on a cake?) or at the visual level (for example, when a text contains too many angle brackets and other odd punctuation).

But more than this, a well-trained *Wegmachine* should also be capable of using appropriate imagery of its own choosing to decorate a simple text-only cake request. This imagery, which would be learned from its stock of exemplar cakes and its even larger inventory of greeting card images, could be chosen to match the relevant themes (e.g. birthday, celebration, youth) to appropriate colors, contours and textures (e.g. pink balloons, fluffy clouds). our *Wegmachine* might use a non-photo-realistic picture generator to create its own pictures or collages to accompany a client's chosen text. In such a case, our machine would first generate a diverse pool of candidate images, before choosing the image that best matches its learned model of what yields a *valid* cake.

If the Advertiser's Code is agnostic on the meaning of "artisan", we can always look to a conventional print dictionary for answers. **Merriam Webster** provides two related senses for "artisan":

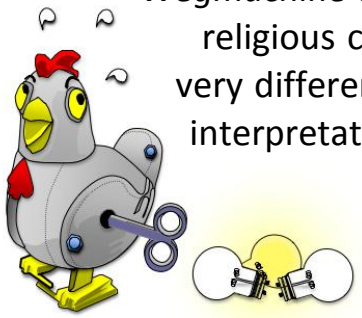
- *a worker who practices a trade or handicraft*
- *one that produces something (such as cheese or wine) in limited quantities, often using traditional methods*

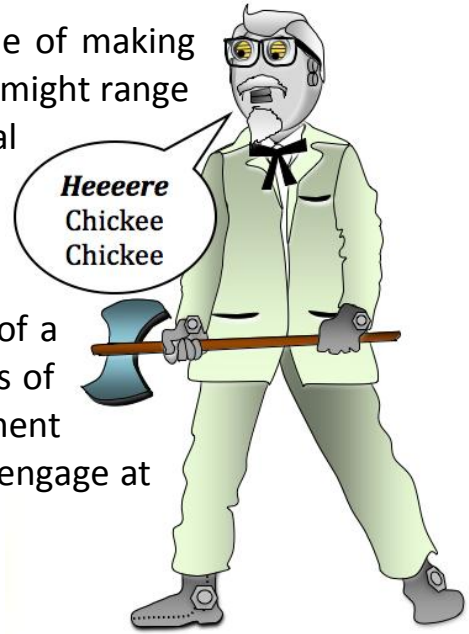
Our idealized view of an artisanal producer therefore emphasizes the importance of skill, training, individuality and tradecraft, while the etymological kinship of "artisan" to "artist" even suggests a degree of artistic pride. So we are unlikely to grant artisan status to a plumber, an electrician or a barman, but carpenters, tailors, bakers and cocktail makers –anyone that can exercise skill, generativity *and* selectivity – may all make the grade.

How does our hypothetical *Wegmachine* fare when judged against these criteria? We can certainly think of our generative machine as a dedicated worker, and of the production of bespoke cake designs as a trade, or indeed a handicraft, given our arguments for *Wegmachine*'s hands-on engagement with its duties. Since our *Wegmachine* can hardly be said to use traditional methods – but neither of course do Wegman's human bakers – the issue of *Wegmachine*'s artisanal standing thus hinges on the sparing use of its skills on limited runs of individualized products.

Our hypothetical Wegmachine is certainly no *HAL-9000*, and hardly resides at the pinnacle of AI ambition, but that is largely the point of the exercise. Just how much should a machine have to do, and how reliably should it do it, before we judge it to be more *hands-on* and engaged than an equivalent human who is simply going through the motions? Our ***Wegman*** versus ***Wegmachine*** is more *Gedanken* than *Gebakken* then, more a thought experiment in AI than an elevator pitch for a new wave of mechanical bakers. But as a thought experiment, it draws on proven technologies to show the viability of imbuing a tightly-focused software/hardware system with a range of artisanal qualities and sensitivities. Certainly, our imaginary *Wegmachine* is capable of showing more interest in the design needs of its customers than is exhibited by humans who hurriedly use cut-and-paste to produce results like that presented to aunt Elsa. While humans may justify a lack of engagement and an indifference to a client's needs with the claim that "*customers [would just] yell at us for bothering them*", a *Wegmachine* could use its well-trained aesthetic judgments about what constitutes an unexceptional cake – in terms of the imagery and language primed by the original email – to limit its bothersome questions to those customer requests that poorly fit its expectations.

No matter how comprehensively we trained our ***Wegmachine***, it would still be capable of making mistakes. Ironically, the more comprehensive and diverse the machine's training set – which might range from children's cartoons to religious iconography to graphic erotica – the greater its potential for embarrassing mix-ups. Yet to hope otherwise is to miss a key point of the thought experiment. Suppose a customer emails a request for an Easter-themed cake, and our *Wegmachine* selects a bloody Crucifixion scene (as acquired in training from a dour religious card) rather than a jolly image of the Easter bunny. Such a mistake is of a very different character to the real-life cock-up of Wegman's HTML cake. Mistakes of interpretation, like the former, are indicative of a level of conceptual engagement that is entirely absent from mistakes like the latter. For one needs to engage at the level of ideas – no matter how shallow this engagement might be – before mistakes of interpretation are even possible.





Though creative people are often appreciated for their fluency of divergent production – the ability to generate many and diverse solutions to a problem – we look down on those who produce with too little variety on too large a scale. Implicit then in our appreciation of true artistry is a contempt for those who "*sell out*", for a ***mere generator*** who produces too much too quickly can hardly be said to invest enough care and attention into each individual output or result. Though every creator deserves to earn a living, we prefer those who sell themselves by the inch rather than the yard. True artists and artisans create in limited quantities not so they can manipulate market prices, but because individually hand-made products rarely lend themselves to mass production.

Woody Allen captured this tension between quality and quantity when he quipped that "*My heart's desire is to forge in the smithy of my soul the uncreated conscience of my race. And then see if I can get it mass-produced in plastic*." Though we should not begrudge creators their hard-earned successes, we are naturally suspicious of any who seem to have hit upon a formula for success and who then *milk* this formula for all it is worth. While the embodiment of such a formula in a mechanical/algorithmic form may prompt us to level our opprobrium at the machine as well as its user, not every generative machine is guilty of unoriginal ***mere generation*****,** or comparable to a disinterested human on **auto-pilot**.

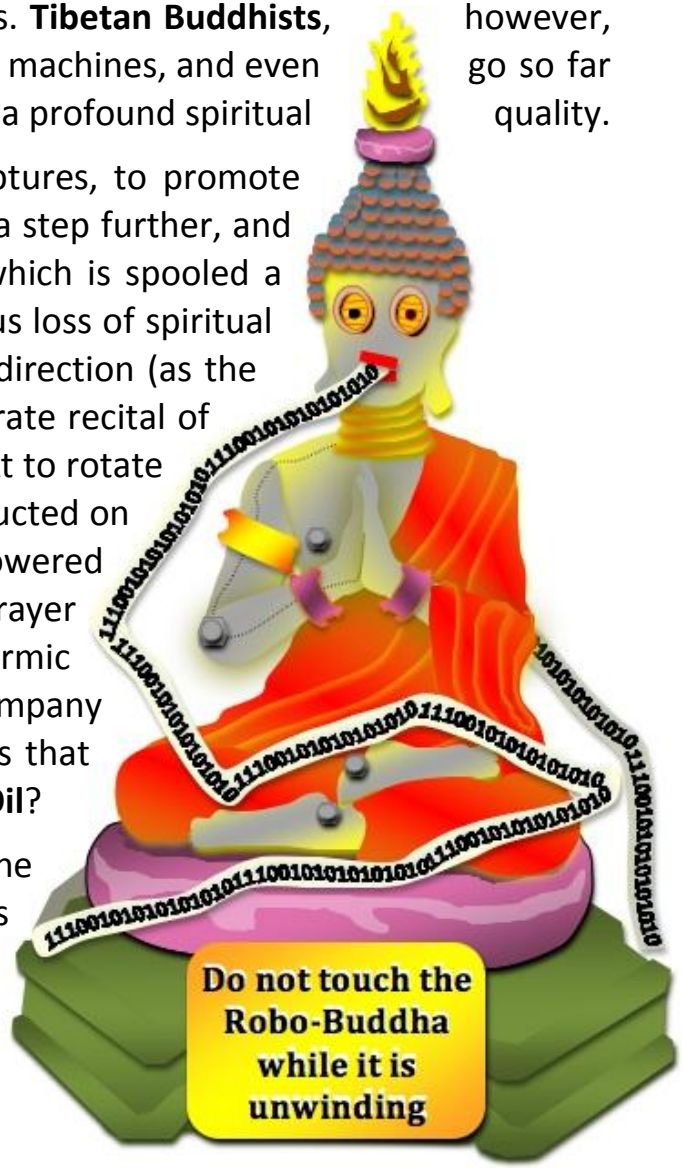
The narrator of *The Great Gatsby* tells us that Gatsby had in his kitchen a modern marvel which *"could extract the juice of 200 oranges in half an hour, if a little button was pressed 200 times by a butler's thumb*." As much as we admire the speed and efficiency of our machines, our cautionary tales of the ***Golem*** and the ***Sorcerer's Apprentice*** remind us that we have reason to fear their capacity for unsupervised production. Yet the sore-thumbed drudges we impose on our machines to tame their speed and to dull their efficiency are just as likely to be dismissed as unthinking automata. While the word *automaton* has a neutral color in a theoretical context, it takes on a pejorative hue when it is metaphorically used to describe humans who are merely going through the motions. We project onto our machines a contempt for the artless drones who are paid to act like *the worst kind of automaton*, but should we reserve the same contempt for a well-trained machine, such as a ***Wegmachine*****,** that is designed to learn to act like a sophisticated human?

Our metaphors encourage us to pejoratively talk of human "*automata*" as soulless and unthinking, and so we magnify the same negative emotions to industrial scale when we project them onto our machines. For while an untalented human may offend on a small scale, a maladroit computer can drown a whole market in a deluge of uninspired, mass-produced fakes. Yet the core issue here is not ***automation*** but ***engagement***. Humans and machines are each capable of merely going through the motions and of ignoring the specific needs of distinct contexts. Yet each, if trained, is also capable of seeing each new context as more than an instance of a generic script. Each can decide to depart from this script when a unique case warrants special attention, to attune to the salient qualities of an unexpected development in ways that appear mindful to an unbiased observer.

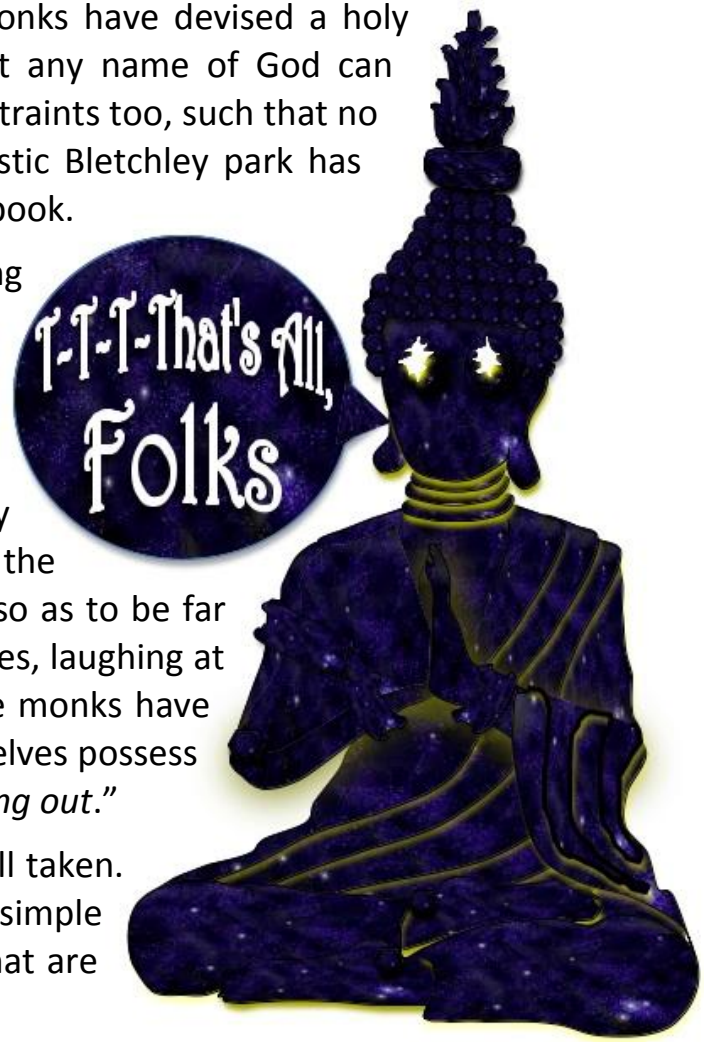
Most modern cultures recognize the value of large-scale mechanical automation – barring obvious exceptions such as the ***Pennsylvania Amish*** – yet few ascribe the same cultural and semiotic significance to the products of programmed machines as they do those of purposeful humans. **Tibetan Buddhists**, however, see little to discriminate between the actions of people and the actions of their machines, and even go so far as to imbue certain mechanical actions, when performed in the right way, with a profound spiritual quality.

Buddhists commonly recite prayers and mantras, or quote Dharmic scriptures, to promote good karma for themselves and for their environments. Tibetan Buddhists go a step further, and use rotating ***prayer wheels*** – decorative drums or hand-held rattles inside which is spooled a Dharmic text – to automate the recital of prayers and mantras, with no obvious loss of spiritual potency or conviction. For the act of spinning a prayer wheel in a *clockwork* direction (as the text inside is written and spooled in the same direction) is viewed as a deliberate recital of the prayer inside, and if a slap of the palm or a twist of the wrist causes the text to rotate ten times, or fifty times, then so much the better. Prayer wheels can be constructed on varying scales, ranging from small hand-held units to large wind- or water-powered ***prayers mills***. More affluent believers have been known to favor electric prayer wheels, though the Lama Zopa Rinpoche has wryly questioned whether the karmic benefits of such wheels are credited to the spiritual account of the electric company rather than to that of the machine's owners. Do the miniature prayer wheels that spin on car dashboards rotate to the benefit of drivers or to the benefit of **Big Oil**?

But let's not to get carried away here, in either direction: between the prevailing Western view of machines as *soulless automatons* and the generous Eastern view of certain machines as *spiritual aides*, we can chart a sensible middle course, in which semiotic objects are evaluated relative to their cultural significance, regardless of whether their provenance is human or mechanical.



Arthur C. Clarke's story *The Nine Billion Names of God* nicely captures the tension between these conflicting world views. Clarke spins a tale of a lamasery of monks in the Himalayas who have spent centuries pursuing their goal of listing all the possible names of God, in the belief that since the world began with the name of God, it must also end when the last name of God has been uttered. The monks have devised a holy alphabet, and for reasons known only to themselves, have determined that any name of God can contain no more than nine of these letters. They impose a variety of other constraints too, such that no holy name can contain a sub-sequence of three repeated letters. This monastic Bletchley park has always enumerated these names by hand, painstakingly adding each to a giant book.



But the advent of modern technology offers an obvious means of automating this process, and so the monks engage a computer company to send them both a computer and a pair of programmers for the task. The computer is easily programmed with the alphabet of holy symbols and the rules for name generation, and it is quickly put to work enumerating and printing all nine billion names. However, as the task nears completion, the programmers worry that the monks will become violent when their religious beliefs are falsified by the world's refusal to end on schedule. They make their excuses and leave early, so as to be far away when their program terminates. As they trek down the mountain on ponies, laughing at the monks' strange mix of superstition and technological savvy – after all, the monks have shown more faith in the value of computer-generated objects than they themselves possess – they notice an alarming sight: "*Overhead, without any fuss, the stars were going out*."

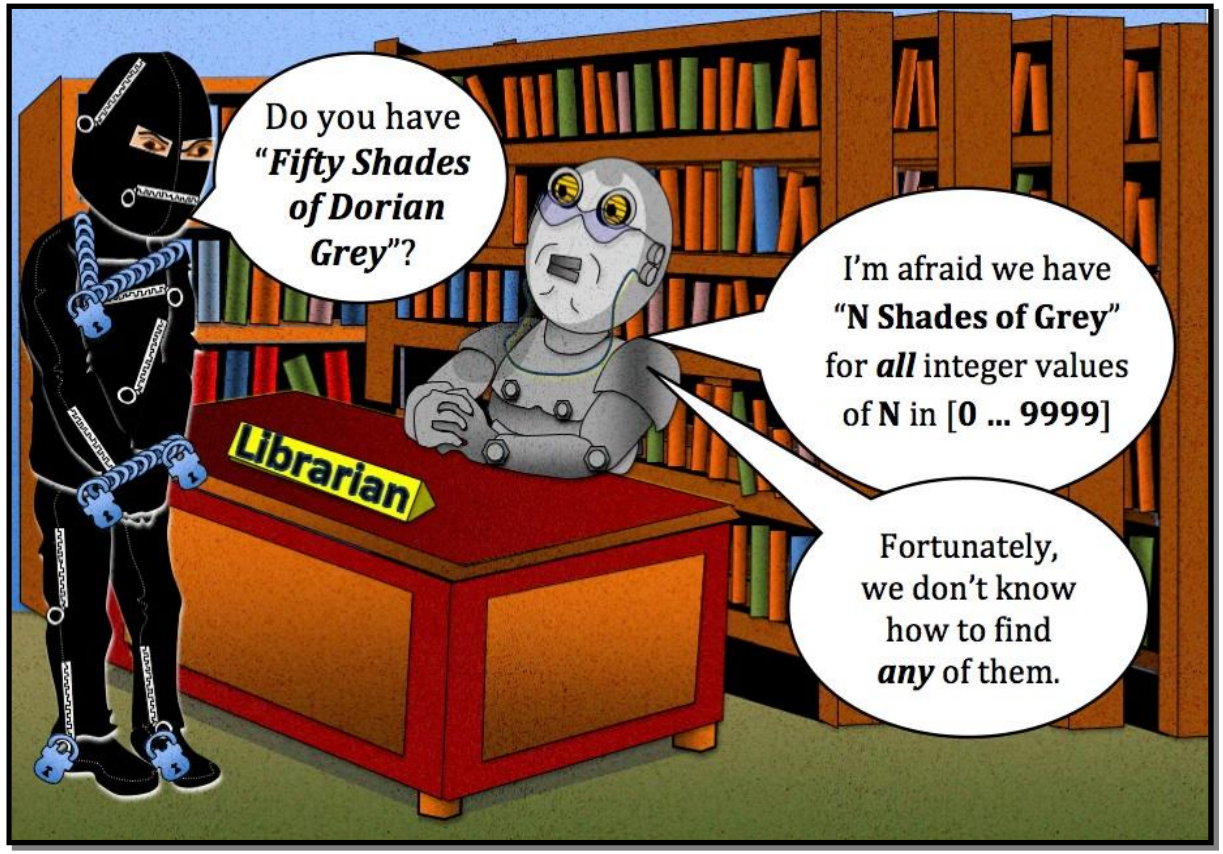
Clarke is obviously trading in keen satire here, yet the broader lesson is well taken. The knowledge that a wide range of objects are generated effortlessly using a simple algorithm should not rob such objects of their cultural significance to those that are willing to see in them either practical utility or profound semiotic value.

Yet that which is easily and cheaply acquired in bulk is valued lightly and rarely respected, for ironically, value and meaning often arise from holding back, from *restraining* our capacity to generate. For instance, at any given moment there is a vast number of possible sentences we could utter, given our knowledge of the grammar of our language and a sizeable vocabulary of words. But the truly meaningful sentences that we can utter in a given context form just a tiny subset of this overall capacity. The linguist **Noam Chomsky** has argued that the capacity to generate is nonetheless a key part of human creativity, and founded the highly influential field of ***generative linguistics*** to model the formal competence of humans to generate a seemingly infinite supply of novel surface utterances from an underlying set of abstract deep-structure transformations.

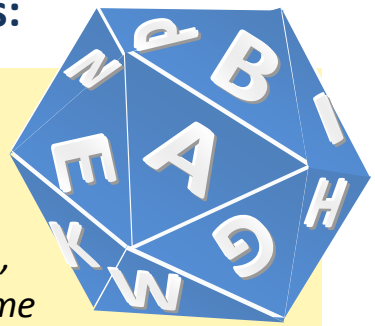
A computer can use much the same kind of abstract grammar rules – essentially a set of well-defined structural transformations – to generate many kinds of *grammatically* valid structures, from cake patterns to architectural plans to pastiches of famous paintings and painterly styles. We might, for instance, eke generative rules from the best surrealist combinations of **Magritte** or **Dali**, or induce design rules to ape the splatters of **Jackson Pollock** or the dots of **Brigit Riley**. Yet the more we generate merely because we can, the more our true meanings get lost in the noise of mere possibility. To give substance to our meanings, we must curb our enthusiasm and generate less, not more.

The perils of *mere generation* are captured in the Jorge Luis Borges story *The Library of Babel*, which describes a *universal* library in which every imaginable book is stored on the shelves of its hexagonal rooms. Well, there are *some* limits to its inventory: no book has more than **410 pages**, of **40 lines** per page; each line has up to **80 characters**, drawn from an alphabet of 22 letters, a comma, a space and a period. But within these limits, the library contains every book ever written, as well as every book that ever will, or could, be written. To pick a book at random here is no different from generating one at random, by rolling an alphabetic **25**-sided die **410 x 40 x 80** times.

But such immensity of scale brings unique problems of its own. To find any meaning at all in Borges' library, we desperately need a catalogue to find those books that are worthy of our attention, and to avoid those which, by implication, are nonsensical in part or in whole. Borges wryly notes that his library must inevitably contain many such catalogues, as any such book must itself be a text that lies within the generative reach of the library. Yet any such catalogue must also be *internal* to the noisy disorder of the library. To the dismay of every librarian, and like every other useful text lurking on the library's shelves, any such catalogue would itself be lost in a sea of ***mere generation***.



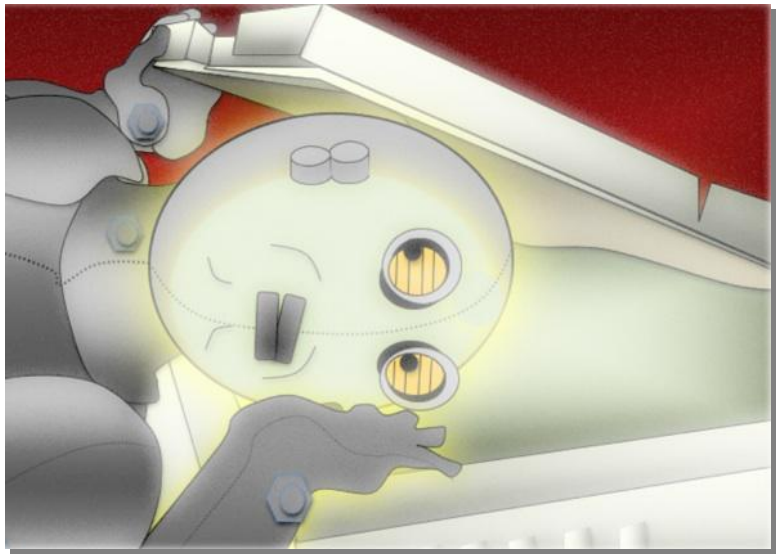
Borges provides us with more tantalizing examples of the library's contents:



> "*Everything would be in its blind volumes. Everything: the detailed history of the future, Aeschylus*' The Egyptians, *the exact number of times that the waters of the Ganges have reflected the flight of a falcon, the secret and true nature of Rome, the encyclopedia Novalis would have constructed, my dreams and half-dreams at dawn on August 14, 1934, the proof of Pierre Fermat's theorem, the unwritten chapters of* Edwin Drood, *those same chapters translated into the language spoken by the Garamantes, the paradoxes Berkeley invented concerning Time but didn't publish, Urizen's books of iron, the premature epiphanies of Stephen Dedalus, which would be meaningless before a cycle of a thousand years, the Gnostic Gospel of Basilides, the song the sirens sang, the complete catalog of the Library, the proof of the inaccuracy of that catalog.*"

Borges' library contains every true answer to every question, as well as every evocative metaphor, every sidesplitting joke, every stirring speech and every quotable poem. But it also contains every false answer, every bad joke, every doggerel poem, and every imaginable piece of linguistic excrescence. Without a means to distinguish them all, they are all equally worthless. Its sweeping generativity means that Borges' library of Babel is, in many respects, a self-referential structure, inasmuch as many of its books will inevitably refer to the library itself. For any merely generative algorithm that is capable of generating all possible books in the library will also generate all possible books whose topic *is* the library (including Borges' own story ***The Library of Babel***). Many of these referential works will purport to be alternate catalogues of the library. But for each possible catalogue, a multitude of other books will sing its praises, while yet another multitude will offer the most damning critiques of its gaps and inaccuracies. In fact, these books will all be criticized or praised in turn by other books in the library, and so on, *ad infinitum et nauseum*. Unfortunately, the library itself offers no reliable means of identifying which books are to be trusted, even as they concern the library itself. The key to generative creativity, then, is not an ability to generate good outputs, but an ability to only (or mostly) generate good outputs, and the restraint to not over-generate bad outputs.

An experienced *Wegmachine* should be capable of generating myriad cake designs, each different and each tailored to specific customer needs. Indeed, if each instance of a *Wegmachine* hones its tradecraft and its aesthetics on a unique sequence of commissions, then each instance will itself be unique. This is akin to two craftsmen who attend the same schools, learn from the same masters and hone the same techniques, but who still develop unique styles and aesthetics over the course of their diverging careers. Yet though we cannot just copy respected artists and fill our factories with their clones, it would be a simple matter to clone any instance of a *Wegmachine*, with all that makes it unique, and share endless copies of it on the Internet.

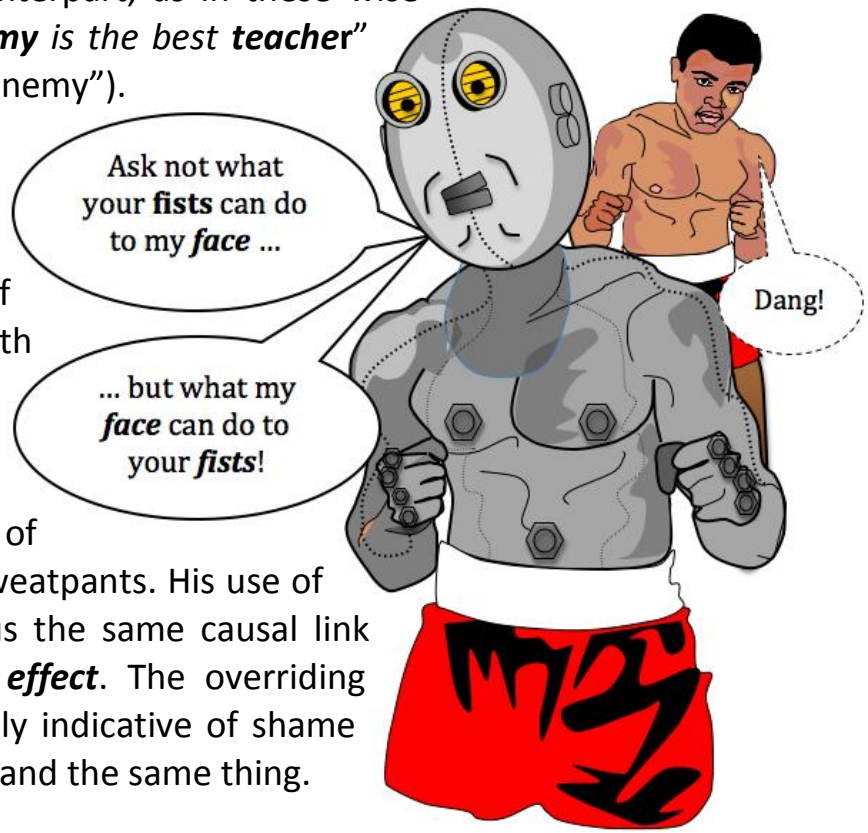


Nonetheless, as computers exhibit greater and more nuanced creative abilities, society may reach a tipping-point at which our fears of ceding control to massive, unsupervised generation no longer shape our views of computation. A newfound respect for artisanal computers may even prompt us to revise our views of large-scale generation. Modern assembly lines now support a much broader range of product variation than in the early days of Henry Ford's *Model T*, which came in any color that happened to be black, and a truly artisanal computer, with its myriad experience-honed biases, may be used to support unprecedented levels of variation in the assembly lines of the future. Indeed, different instances of the same machine, like different instances of our *Wegmachine*, may command more respect in the marketplace because of their specific experience and unique design aesthetics. Mass-produced items, from beach towels, seat covers and rugs to jewelry, wallpaper and crockery, might be autonomously produced *en masse* with never twice the same design coming off the line, unless of course that design is repeatedly requested by a customer.

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." The study of how best to bridge words and ideas is an ancient one, in which rhetoricians have systematically identified a rich toolbox of linguistic forms for persuasively expressing our ideas. These rhetorical devices are so effective at shaping and delivering our well-developed ideas that they can also lend our less substantial ideas the unmerited appearance of solidity. For though an apt rhetorical choice can provide a supportive scaffolding for an underdeveloped idea, some orators abuse rhetoric, using shallow tricks to mass-produce glib insights that carry the mere appearance of profundity.

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. 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 16th 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 35th president, John F. Kennedy, "*Ask not what your country can do for you, but what you can do for your country*".

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 can achieve a deep impact at a conceptual level. Chiasmus relies on repetition to drive home these ideas, even if the duplication of content is not consciously 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*". But one can repeatedly evoke 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" as "tolerance" crosses over to "enemy").



So chiastic repetition with crossover can be 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. Note how Lagerfeld first asserts a causal link from sweatpants to defeat, and quickly follows up by asserting a 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 given is that sweatpants are more than merely indicative of shame and demoralization; seen through Lagerfeld's gimlet eye, they are one and the same thing.

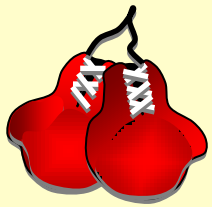
Chiasmus can pack a powerful punch

by simultaneously drawing our attention to the superficial similarities and deep dissimilarities between what is expected and what is real. Yet chiasmus can often be *too* convenient, allowing one to fake the presence of cutting insight with little more than *cut-and-paste*. Consider the following exchange from the 1999 film comedy ***Mystery Men***, which depicts the misadventures of a ragtag group of overzealous superheroes with underwhelming powers. **Mr. Furious** has serious anger management issues, while **The Sphinx**'s only power is an ability to torture syntax until it yields an apparent profundity.

That great last line (bottom right) says it all: the Sphinx has hit on a successful formula for turning the utterances of others into Zen-like prognostications that are little more than glib chiastic repetitions. This is the essence of a deterministic formula: it always yields the same outputs for the same inputs, making the weak demurral "Not necessarily" all the more risible. It is not just computers that merely generate – it turns out humans are rather good at it also.

Scene: *Our ragtag bunch are busy preparing their homemade superhero costumes in anticipation of a big showdown.*

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.

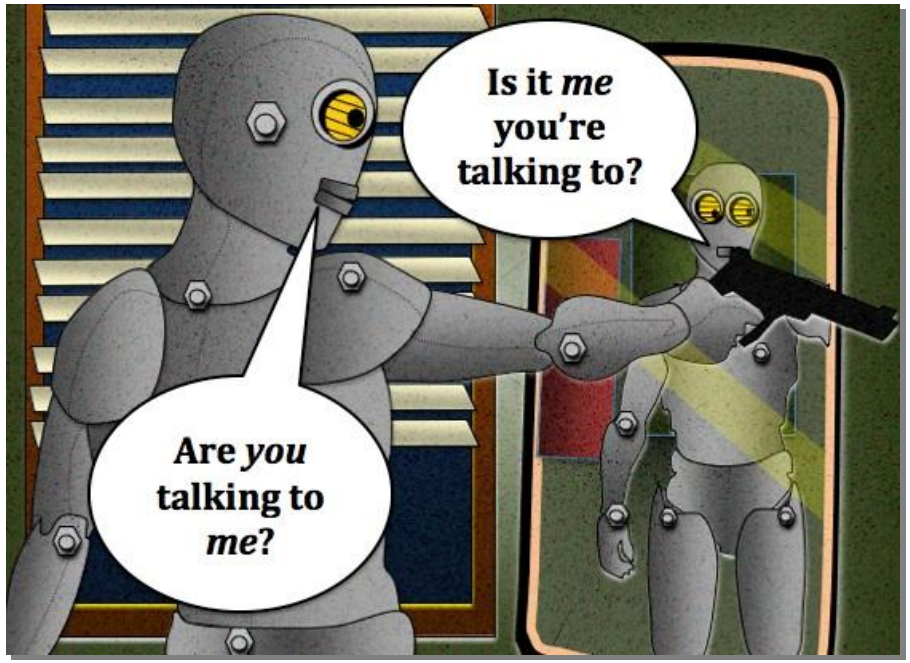
Mr. Furious:  Okay, am I the only one who finds these sayings just a little bit formulaic? "*If you want to push something down, you have to pull it up. If you want to go left, you have to go right.*" It's ...

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.

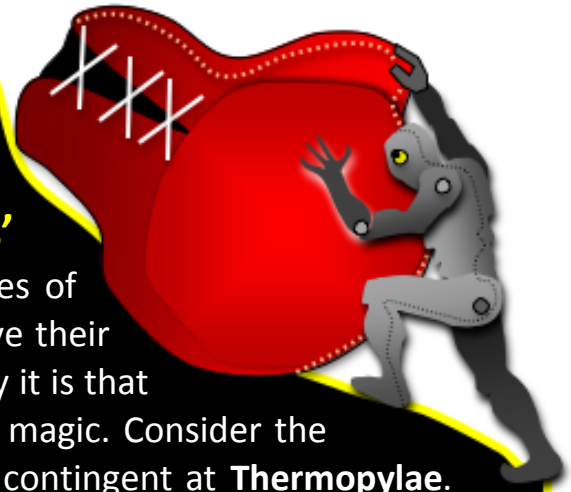
Every rhetorical device is a formula of a sort.



It is their usefulness in different contexts that makes them worthy of study by those who want to package their arguments in a form that best serves their meaning. As an actor, **Robert de Niro**'s formula is to repeat himself with an increasingly menacing tone. As an orator, **the Sphinx** is likewise 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 formula. As a rhetorical device for echoing and inverting an opponent's arguments, chiasmus has few equals, yet we tire quickly of any device that is used too often and with too little variety. If the Sphinx (or **de Niro**) were to up his game, and draw from a richer arsenal of formulae to suggest a deeper engagement with forms and ideas, we might pay more attention to what he has to say.

The syntactic manipulation 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 underlying space of ideas. Indeed, by constantly searching for openings 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*."

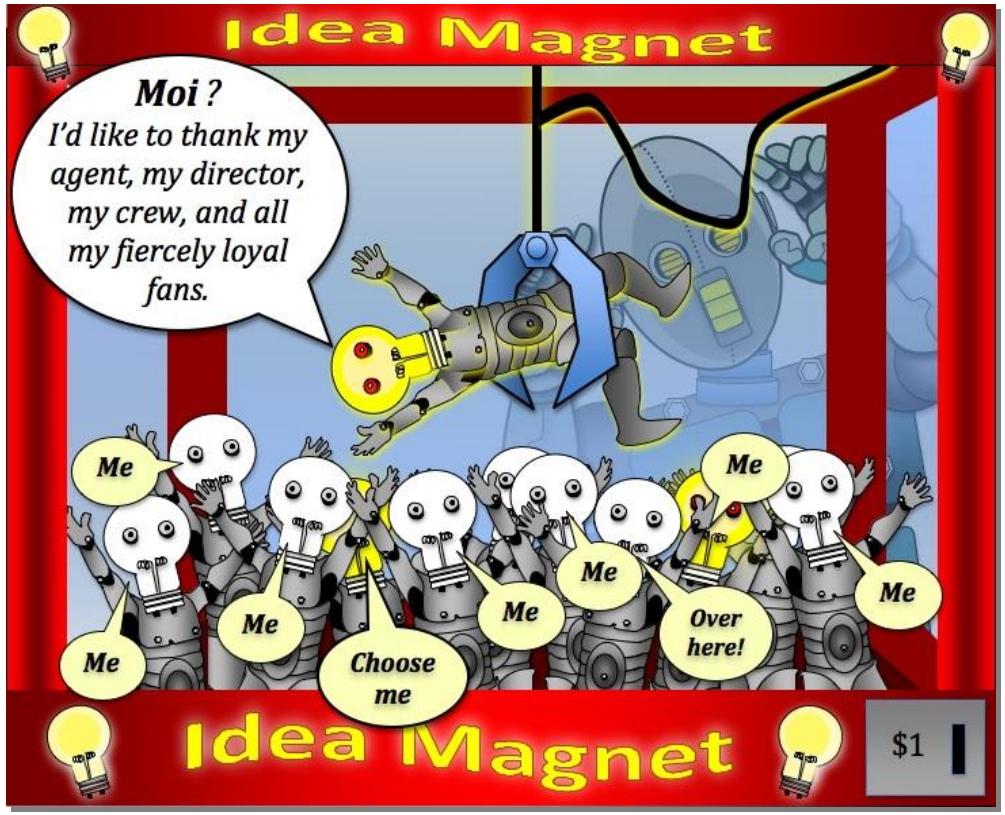
χ Marks the Spot



When chiasmus is constructed using explicit repetition, as in Keats' "**Truth is Beauty, Beauty Truth**", the result is hard to miss. But such obvious examples of chiasmus are merely the tip of the iceberg: many more use subtle resonances to achieve their crisscross integration of meanings. Indeed, the more subtle the resonances, the less likely it is that the crisscross effect will be consciously appreciated, even as it still works its rhetorical magic. Consider the account from **Herodotus** of the meeting of the vast Persian army and the tiny Spartan contingent at **Thermopylae**. When the innumerable Persians threatened to "***blot*** *out the sun with our* ***arrows***", the Greeks responded evenly "*then we will* ***fight*** *in the* ***shade***". Notice how the Greek rejoinder interlocks with the Persian threat in a chiastic fashion: as "*blot out the sun*" maps to "*shade*", Persian "*arrows*" map to the Greek "*fight*". The crisscross effect actively reverses and undoes the Persian threat, while stressing a fundamental difference in emphasis: while the Persians primarily see their arrows as a means to impress, the Greeks assert that *they* are primarily there to fight.

- Chiasmus *across* utterances (as in the Herodotus example above) can add rhetorical force to a rejoinder. Find examples of famous exchanges where chiasmus can humorously sharpen a response.
- Convert some famous non-chiastic slogans / maxims / quotations / catchphrases into a crisscross form (E.g., "*to be or not to be*" becomes "*to be or be not*"). Are the results more persuasive, more memorable, or more resonant?
- Chiasmus is a common rhetorical form that primes certain expectations in an audience. These expectations are sometimes violated, as in the phrase "*What's yours is mine, and what's mine is … my own*.". Using the repetitive structure of chiasmus as a retrieval cue, look for other examples of *thwarted chiasmus* on the Web.
- What advice would you give to The Sphinx to make better use of his obvious skill for forming Chiastic aphorisms?
- How might a computer use rhetorical strategies like chiasmus to help us write better tweets or SMS messages

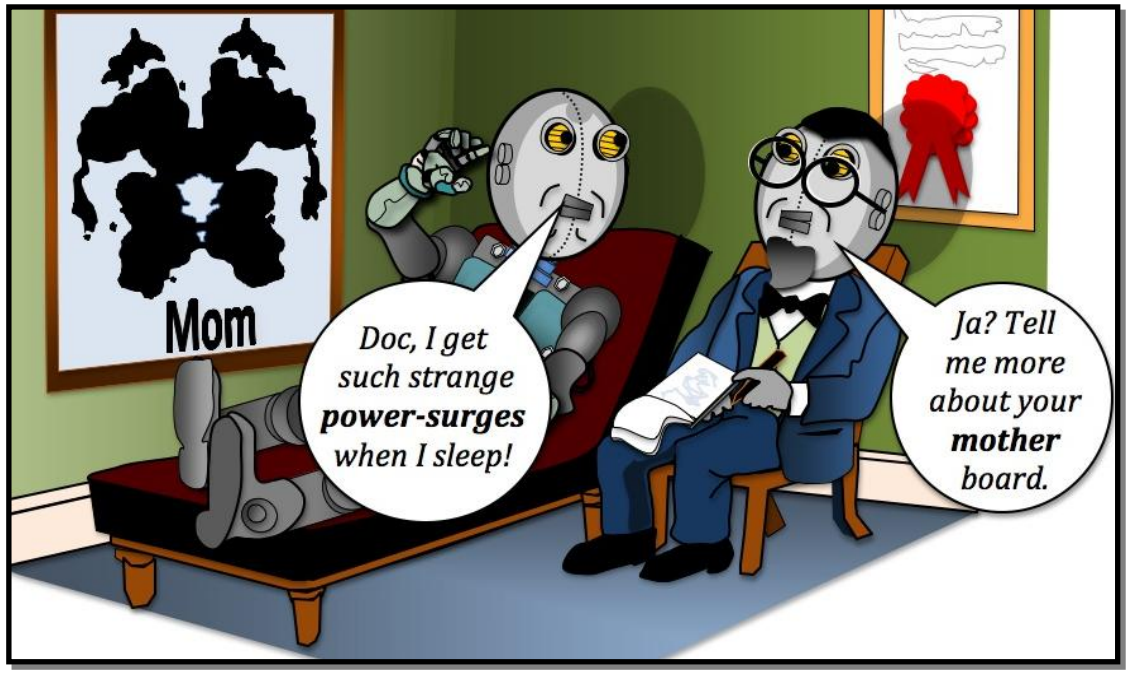
Even at his productive peak, the demand for **Picasso**'s work far outstripped his ability to supply it, forcing dealers to gather at the artist's studio in the hope of receiving a hard-won scrap from the master's table. To acquire a new work from Picasso himself could be considered a success; to acquire his work through other channels could be considered a *victory*. **Arthur Koestler** recounts a tale of how one dealer friend of Picasso made the pilgrimage to the artist's studio to boast of such an acquisition. After describing his new purchase with a relish that was part flattery and part taunt. Picasso responded with a taunt of his own, by dismissing the piece as a **fake**. Shocked, his friend nonetheless held a trump card: he himself had watched Picasso working on that very piece in that very studio. "*So what*", replied Picasso with a shrug, "*I often paint **fakes***."



Even good artists occasionally fall into the trap of self-pastiche, unconsciously aping their own style to produce the kind of easy-to-classify works that critics expect and that dealers can sell. But from an artistic perspective, these works may as well have been produced by an artless forger, or by **The Sphinx**. Picasso's mischievous reply shows that he recognized in himself this tendency to *mere generation*, and distanced himself from his derivative works (*post*-sale, of course) by viewing them with a disdain one shows a criminal counterfeiter. For though creativity is fundamentally a generative act, quality is as much determined by what we throw away as by what we keep. While many are *called* in divergent production, few should actually be *chosen*. In creativity, **selectivity** is every bit as important as **generativity**.

Mind Games

As we learn to perform a new task or acquire a new ability, our initial efforts are typically faltering and sub-par. Practice makes perfect, however, and a modicum of talent helps too. As we steadily improve, we acquire an economy of action that can be interpreted as grace. In short, we learn to streamline our performance, to strip away the rough edges until we become ***well-oiled machines***. It's ironic then that in our search for graceful efficiency and instinctively automatic competence, we aim to turn ourselves into sophisticated ***automata***, capable of performing a complex task well, perhaps even exquisitely well, with minimal conscious interruption of our natural flow. The sand-box of repeated experience allows us to fine-tune a set of convenient formulae and scripts for dealing with the most typical phenomena, and repetition allows us to internalize these formulae until they become automatic and second nature.



But all formulae, like all rules, have unanticipated exceptions. When automatization works as expected, a practiced formula lends our abilities polish and speed. But when it fails, we can stumble again like awkward beginners. The trick is to know when to override the automatization in some cases, without blunting the power and speed of our practiced responses in others. Yet it is difficult to undo our own programming, especially when we don't quite appreciate how or why it works in the first place. Lacking a programmer's guide to the mind, we must turn instead to **psychology**, whose insights are surprisingly consistent with a computational perspective on how we think and create.

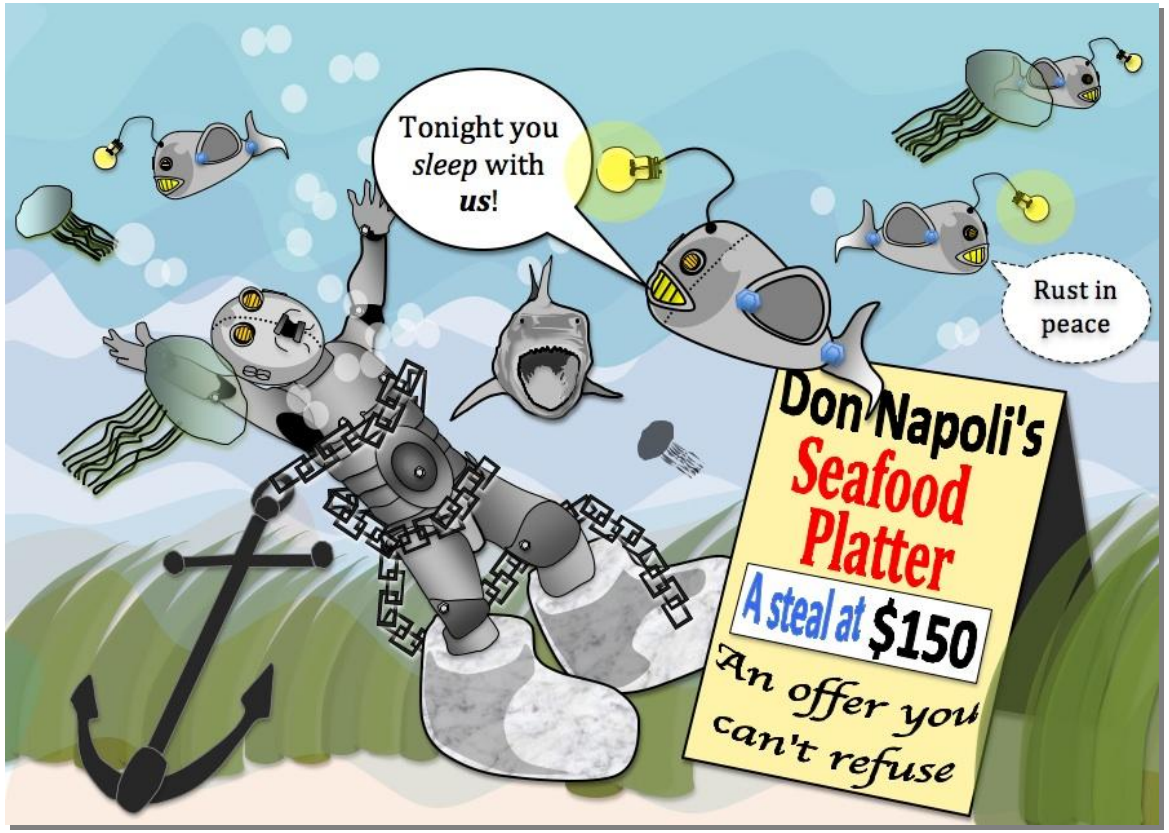
The mind's information architecture is an intricate balance of optimizations and tradeoffs that sometimes work against us, automatically herding us into mental "sets" that can be exceedingly hard to escape. We've all experienced those frustrating *tip-of-the-tongue* moments where our information architecture falls ever so short of our needs: we sense that the correct information has left the station but has somehow gotten lost on its way to the terminus. Equally frustrating is the feeling that key information for a task is competing with other information that only appears salient, and the race to consciousness is repeatedly won by imposters that we would dearly love to disqualify. No matter how hard we fight against the involuntary retrieval of failed candidates that we have already rejected, our information architecture automatically brings them back into the race. In fact, psychologists have found that the harder we try to ignore an unhelpful stimulus, the more it insinuates itself into our thought processes.

Word-fragment completion tests provide a thoroughly fascinating demonstration of how the automated retrieval of salient-seeming information can actually be detrimental to the problem-solving process. The word game below, which uses experimental stimuli crafted by psychologists **Steve Smith** and **Deborah Tindell**, begins with a simple memory task: spend a minute or so memorizing the following word list. The order of the words isn't important. Just try to hold on to as many words as you can. There will be a test later, after a brief discursion!

Charter, Voyager, Analogy, Density, Cottage, Tonight, Crumpet, Trilogy, Fixture, Brigade, Cluster, Holster

Priming is an automatic cueing process in which one stimulus unconsciously prepares the mind for the associated stimuli that follow. It is one of the mind's principal mechanisms for allowing the particulars of a given context to subtly influence our approach to a problem. Priming is the reason that the word "cats" is likely to make us think of *dogs*, "salt" is likely to make us think of *pepper*, and "cloud" prepares us for *rain*. Priming lends fluidity to our actions by cueing up the information that our minds are likely to need next, in much the same way that a CPU pre-fetches from memory the instructions that a program is most likely to execute next. a positive priming stimulus cues up information that is truly salient to a problem, making valid solutions so much easier to find However, a misleading prime can also act as a *memory block* that stubbornly hinders the retrieval of more apt information.

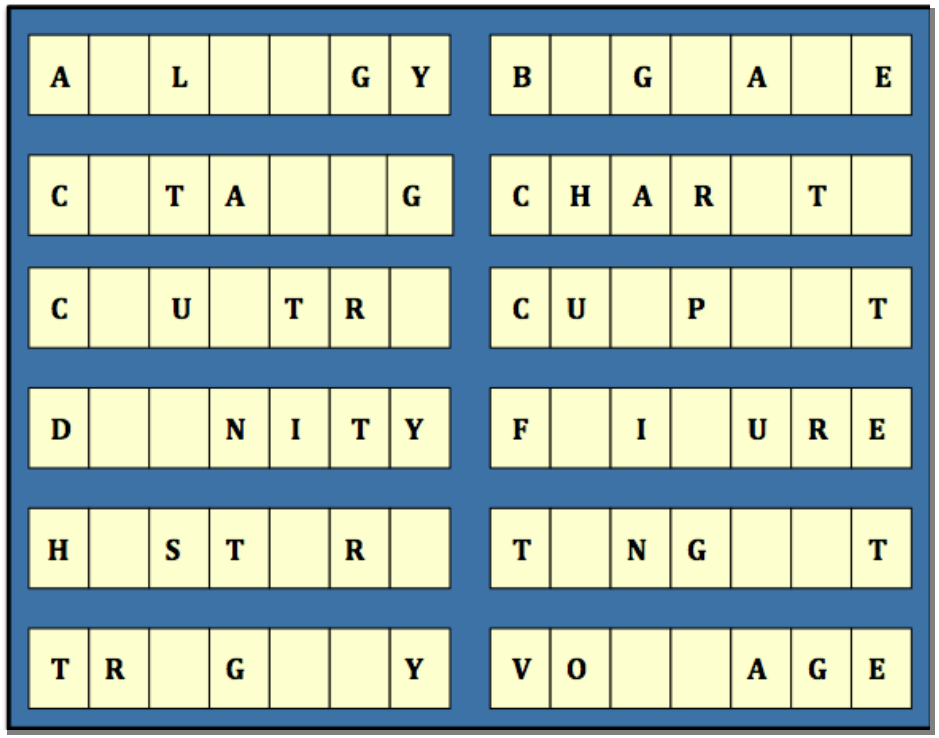
A knowledge of the mind's tendency to jump to conclusions can make us more creative *and* more cunning. Why did ***Kentucky Fried Chicken*** rebrand itself as **KFC** – is it because "**F**" is far less likely to evoke unwanted associations with fat, grease and heart disease than the word "**Fried**"? Why are US Republications more likely than US Democrats to use **Barack Obama**'s middle name, "***Hussein***"? Is it to evoke images of Middle Eastern dictators and terrorists, and to make the first African-American president of the USA seem less than fully American? Or consider what psychologists call the ***anchoring effect*** by answering this question: was **Martin Luther King** older or younger than **55** years of age when he died? You'll likely consider his life's many achievements to arrive at your answer. Yet if you are now pressed for a number – how old was Dr. King when he died? – you are likely to use **55** as an anchor and over-estimate his age accordingly. Rather than work out an answer from first principles, **System 2** will grab onto any anchor it can to try and give itself a head start.



Remarkably, Dr. King was just 39 years of age when he died. This love of an easy short-cut is all too easily manipulated by pollsters and politicians, who subtly anchor their questions in ways that shape our expectations and direct us to give the "right" answers. Yet it's also an important part of why creative solutions delight and surprise us.

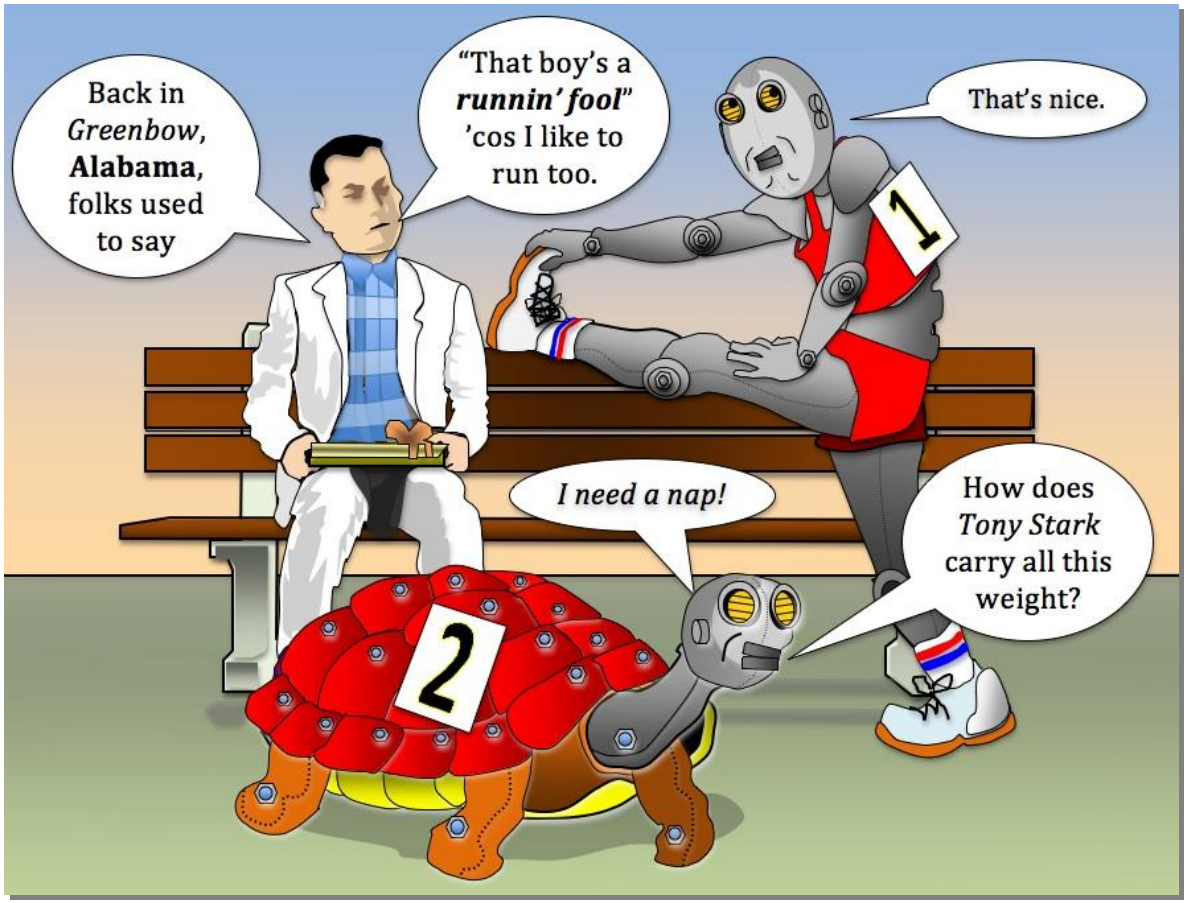
Did you memorize all 12 words on our list?

Yes? So now try the word completion task on the right. Memorization of a short list of words is not a challenging task, but as you have probably guessed, that was merely a distractor designed to interfere with your performance on this completion task. And interfere it most certainly does: for as your minds search for a word that matches **A**-*something*-**L**-*something*-*something*-**GY**, there's that similar-looking word from the earlier memory task, **A**NA**L**O**GY**, jumping up and down for our attention.



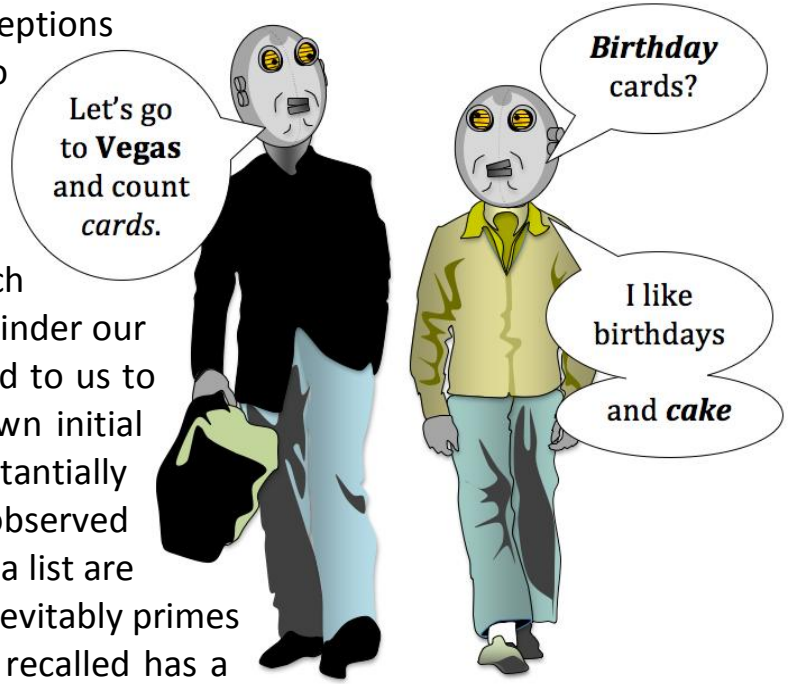
On closer analysis, however, the seductive near-match provided by **A**NA**L**O**GY** will simply not fit in the available spaces, and so we are forced to move on. This, however, is much easier said than done, for alternate candidates are now that much harder to identify while our mind remains fixated on a tantalizing near-miss. Because "**A**NA**L**O**GY**" was primed by the recent memory task, it sits at the top of the list when our automatized retrieval processes go looking for a solution to this new but familiar-looking completion problem. Moreover, mental efforts to dislodge "ANALOGY" only activate it even more, so the harder we try to disregard "**A**NA**L**O**GY**", the more stubbornly primed it remains, refusing to budge from memory where it serves to block the retrieval of a more appropriate solution. **Smith and Tindell** succinctly capture this somewhat perverse state of affairs in the title of their paper: "*Memory blocks in word fragment completion caused by involuntary retrieval of orthographically similar primes*." So had our initial word list contained the positive primes like **ALLERGIC**, **CATALOGUE**, **DIGNIFIED** and **COUNTRIES**, we would have been far better prepared to find the completions ***allergy, baggage, catalog, charity, country, culprit, dignity, failure, history, tangent, tragedy*** and ***voltage***.

In his masterful book "*Thinking fast and slow*", the psychologist Daniel Kahneman uses an Aesop-like simplicity to portray the workings of the mind as a story of two competing protagonists, ***System 1*** and ***System 2***. **System 1** is a hare, swift and responsive and always ready for a race. Automatic priming and the involuntary retrieval of information are key features of its operation. When these automatic processes work to our advantage, it is **System 1** that greases the rails for our lightening-fast intuitions. But when these processes work against us, causing us to fixate on distracting near-misses, it is **System 1** that makes these distractions so hard to ignore. In contrast, **System 2** is a tortoise, sluggish and analytical, easily tired and slow to engage, but capable of deep engagement and profound insight whenever it does decide to act.



If I tell you I bought a bat and ball in a sale for **$1.10**, paying **$1** more for the bat than the ball, **System 1** will tempt you with the easy (but wrong) solution that the ball cost **10 cents**. This division of costs was primed by my statement of the puzzle, in which both **$1** and **10 cents** were mentioned. So it falls to **System 2** to show the inadequacy of this *bat=**$1** and ball=**10 cents*** solution (since **$1 ≠ $1 + 10 cents**), and guide us to the correct answer of **$1.05** and **5 cents**.

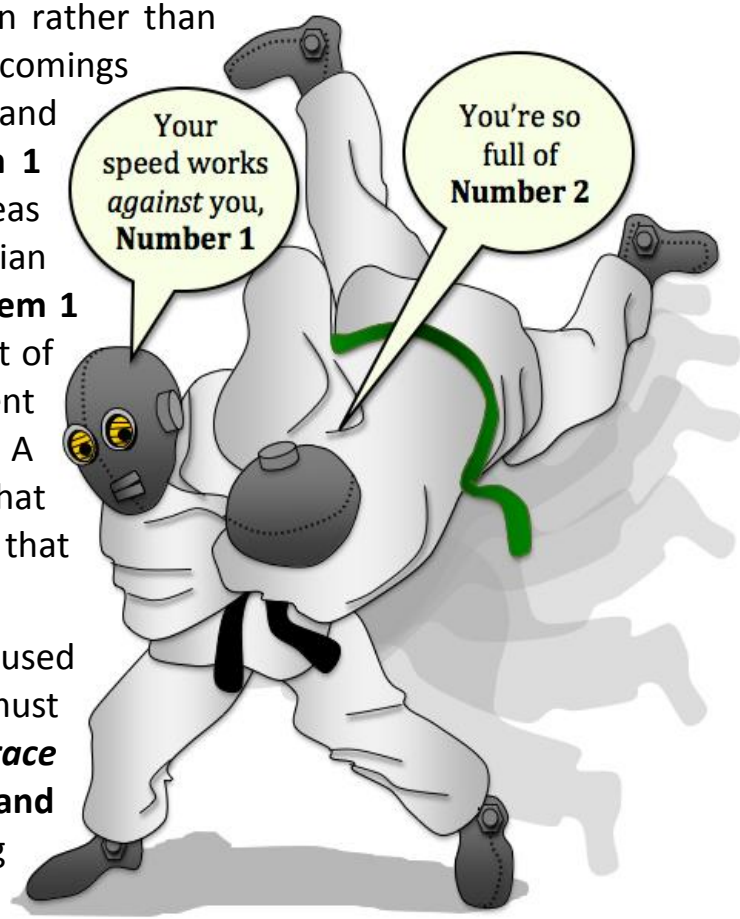
Wegman's HTML cake demonstrates the inherent laziness of System 2, showing us what happens when we over-rely on the autopilot of **System 1**. In the race to offer conclusions, **System 1** invariably beats **System 2** to the line, but its conclusions are often less than reliable. So we need to engage the mechanisms of **System 2** whenever **System 1** jumps too quickly to judgment and leaves us feeling stymied by a problem. Yet, while **System 2** must often clean up after **System 1**, it has very little actual or direct control over **System 1**: it can step in to repair the latter's misconceptions, but lacks the ability to prevent those misconceptions from occurring in the first place. Indeed, any attempts by **System 2** to override the fallacious conclusions of **System 1** can sometimes cause those conclusions to become even more embedded in our thinking.



Studies by psychologist **Steve Smith** and his colleagues explain why the normally useful automatic priming of **System 1** can have such baleful effects on creative thinking. For not only do negative primes hinder our ability to explore alternative candidates, positive primes can also lead to us to fixate on unoriginal solutions that are proven to work, or on our own initial solutions that show some merit yet which may need to be substantially revised or even rejected outright. Specifically, psychologists have observed ***output interference*** in free recall tests of words, in which words from a list are memorized and later recalled in any order. Since the act of retrieval inevitably primes a word and adds to its perceived salience, any word that is initially recalled has a higher probability of being recalled again than other words on the list. This recall inhibition effect, called ***part-set cueing inhibition***, explains why brainstorming sessions can so easily stagnate. Brainstormers are required to withhold judgment on the ideas that are generated in a session, but even the act of suggesting a weak or unworkable idea also makes it more likely that this idea, or a close variant, will be suggested again and again. Brainstorming strives to pool the collective insights of a group, but it also mutualizes the risks of the group collectively falling into a rut.

Voltaire deemed illusion to be *the first all pleasures*, for we often quite enjoy being fooled.

The occasional failings of **System 1** should thus be cause for celebration rather than concern. Comedians use their intimate understanding of **System 1**'s shortcomings – no doubt derived from first-hand experience – to craft logical traps and feints whose interpretations playfully exploit weaknesses in the **System 1** processes of others. For **System 1**'s automatic priming of connected ideas creates momentum in a narrative, and like a skilled martial artist, a comedian uses this momentum against us when that narrative is a joke. So as **System 1** rapidly propels us toward a defective conclusion, it takes only the gentlest of pushes from a well-crafted punchline to deflect us into a radically different outcome, one that requires the engagement of **System 2** to understand. A good joke, then – much like good art – spurs us to engage **System 2** so that we might arrive at a deeper understanding of a familiar-seeming situation that hides unexpected complexities beneath its veneer of conventionality.



Computer scientists have a rather apt name for the *timing glitches* caused by sub-systems that are designed to work in parallel yet whose outputs must be integrated to produce consistent answers at a higher level. A ***race condition*** occurs whenever two such sub-systems – much like **System 1 and System 2** – race to deliver their outputs first, forcing the larger, integrating system to radically revise its own output if an answer that arrives late disconfirms the answer that was computed earlier. Computer scientists try to eliminate any possible race condition, so that their systems always work as expected. Yet without the occasionally jarring discontinuity of **Systems 1** and **2**, there can be little of the mischievous creativity that delights by surprising, and little of the disruptive innovation that seems non-obvious *before* the fact (via **System 1**) but so obvious *after* the fact (via **System 2**). *Vive la différence*!

Can we ever fully understand creativity without also understanding the emotions it entails?

Our appreciation of a joke must surely involve more than the replacement of an obvious but ultimately misleading viewpoint with one that seems more apt, otherwise why should a good joke ever fill us with raucous laughter? Similarly, a scientific discovery must surely do more than incrementally add to our store of existing facts and rules, otherwise why should a profound new insight make us jump and shout with delight, or in the case of **Archimedes of Syracuse**, leap from the bath and run naked through the streets? An original artwork would likewise lack any power to move us deeply if its only notable dimensions were novelty and utility. Creativity is a fundamentally emotional experience that owes as much to our **physiology** as to our **psychology**.

This is perhaps the point of greatest disparity between human and computational creativity: when pursuing innovation, we experience a variety of embodied states that affect both our motivation and our levels of arousal, from anxiety, frustration and dissatisfaction to excitement, enthusiasm and joy. Creativity is disruptive not just of our knowledge and conventions but of our psychological and physiological states: it transforms frustration into elation, dissatisfaction into fulfillment and anxiety into relief. Emotions thus play a pivotal role in motivating our acts of creation – whether as boredom with the status quo, frustration with the limits of orthodoxy or a craving for novelty – and in shaping our response to creativity (and thus our willingness to apply the label “creative”), whether with laughter or with shouts of amazement. Some emotional states, such as ***fear***, lock us down into a narrowly convergent mode of thought, while others, such as **excitement** and even **boredom**, open us up to the joys of exploration and divergence. To understand the causes of creativity, and to proactively foster new acts of creativity, it pays to separate the psychological and physiological states that are most likely to promote innovation from those that are most likely to inhibit a creative response.

Kahneman characterizes the contest of Systems 1 and 2 as primarily a *tale of two velocities*.

The ***race conditions*** that arise in this contest give creativity its distinctive sense of old vs. new, "right" vs. "wrong" and familiarity vs. originality. Yet it is a race played out at multiple levels of being, not just within the mind but between the **mind *and* the body**. For in the parlance of computer scientists, mind and body are both ***containers of state***: as each undergoes internal changes in response to external stimuli, each maintains state information in its own way – in the activation of the brain's neurons or in the level of chemical mediators in the blood. A new stimulus may trigger incremental changes to our physiological state by ratcheting up our levels of emotional arousal, and it may also affect our psychological state by leading us to view our working assumptions as that bit more, or that bit less, tenable.

Occasionally a new stimulus will trigger *radical* changes to both our physiological and psychological states, such as when we are shocked by the twisting plot of a movie or by the unexpected incongruity of a joke. When surprised in this way, both our minds and our bodies must scramble to change state. Though **System 2** is considerably slower than **System 1**, its sudden change of perspective is also achieved through the firing of neurons, and can thus be executed in a matter of milliseconds. But as **Arthur Koestler** notes in his book ***The Act of Creation***, the body is altogether slower to change state, relying as it does on the slow dissipation of chemical messengers in the blood. The resulting race condition, Koestler argues, creates a need for body and mind to quickly reestablish equilibrium. This is achieved, he says, by releasing pent up nervous energy, such as by laughing at a joke or by jumping with joy at a sudden insight.

"The most exciting phrase to hear in science," wrote **Isaac Asimov**, "*the one that heralds new discoveries, is not '**Eureka**' but '**That's funny**...*"" This innocent phrase usually signals the presence of an **incongruity**, and though incongruities are rarely of special interest in and of themselves, their **resolution** can undercut received wisdom to yield the *tiny revolution* of a joke, or upturn it so completely as to yield the altogether grander revolution of a scientific *paradigm shift*. The word "funny" is apt in either case, for though we may we view incongruities as an unwanted source of tension, we derive no small pleasure from using our knowledge and insight to make these anomalies vanish into thin air. As philosopher **Immanuel Kant** put it, we derive comic release via laughter from seeing "*the sudden transformation of a strained expectation into nothing*." In science we really do want to resolve incongruities into nothing, even if it means dismantling a mighty theoretical edifice. In humour, though, we are happy to leave some part of the incongruity unresolved, to linger after a joke like the enigmatic smile of a Cheshire cat.

Humorists, scientists and creators of all stripes are in the business of **incongruity resolution**. While scientists and artists typically resolve their incongruities as they find them, humorists must bear the responsibility for engineering incongruities in the first place, though some *creator-provocateurs*, like **Marcel Duchamp**, work as both artist *and* humorist. Just as scholars of oratory have identified an extensive toolbox of rhetorical strategies, scholars of humour have identified a catalogue of **logical mechanisms** for straining our expectations in incongruous but resolvable ways.

But "*incongruity*" is as vague and all-embracing a term as "*creativity*" itself, making it all too easy to believe that we've simply replaced the word "*funny*" with one that sounds more formal and insightful. So in a mode of assiduous classification that **Ernest Rutherford** would surely dismiss as *stamp collecting*, scholars have sought to systematically tie specific kinds of incongruity to specific families of jokes and to the semantic oppositions around which they pivot – such as *life vs. death*, *high vs. low*, *pride vs. shame* and *sex vs. chastity*. Scholars such as **Salvatore Attardo** and **Victor Raskin** have additionally sought to classify the diversity of *pseudo*-**logical mechanisms** that comedians draw upon to embed these oppositions into a joke narrative in such a way that the incongruity strains our expectations yet also contains the seed of its own resolution.

Humor theorists seek to derive general insights from the study of jokes in which certain kinds of incongruities recur across diverse tales of familiar concerns. Non-humorous incongruities, in contrast, of the kind that spurs inventors to find ways of reconciling contradictory needs, are more commonly found in patent archives than in joke books. The Uzbek writer and inventor **Genrich Altshuller** spent years poring over patent applications to systematically classify each in terms of the underlying incongruity that it resolved. Suppose you want an umbrella that is big enough to shield you in the rain, yet small and light enough to carry in your pocket: these conflicting requirements demand a product that is *big but small*, *rigid but flexible*, and *strong but lightweight*. The incongruities are resolved via the insight that our umbrella need not exhibit all of these qualities *at the same time*. A **collapsible** umbrella of **composite** materials can fold into a small, light, compact form, yet unfold into a large rigid frame when it rains. To systematize such insights as part of his **TRIZ** system for inventors, **Altshuller** mapped the most common incongruities in over 40,000 patents to patent-attested resolution strategies. **Altshuller** saw **TRIZ** and its tables linking incongruities to strategies – which are now widely used in industry – as nothing less than an *algorithm for innovation*.

We tend to see only the success stories when we look at creativity through a microscope, or glance at it in the rear-view mirror of history. For creative experimentation is risky, and most departures from the comfort zone of convergent thinking produce less-than-memorable results. It takes ***motivation*** to be creative, to reject time-tested rules and to explore new avenues and new measures of success. This motivation can be seen as both a pre-existing ***trait*** of an individual or as a ***state*** that we can foster. The former perspective suggests that creativity is an innate quality that expresses itself more strongly in some people – the natural artists and born inventors among us – than in others. The latter, more optimistically, suggests that creativity is a perfectible quality that can be cultivated with the right approach to the nurturing of motivation.

Psychologists distinguish two complementary states that can spur a creator to persist at a vexing problem: **intrinsic** motivation comes from *within*, while **extrinsic** motivation comes from *without*. Creativity theorist **Teresa Amabile** defines the intrinsic state as "*the motivation to work on something primarily for its own sake, because it is enjoyable, satisfying, challenging, or otherwise captivating*", and the extrinsic state as "*the motivation to work on something primarily because it is a means to an end*". Intrinsic motivation can be fostered via positive feedback and guidance, or by showcasing the benefits of intrinsic motivation in others. Extrinsic motivation can be fostered by incentivizing creative behaviour with the prior promise of a final reward. However, if extrinsic rewards are not carefully designed and structured, they can actually diminish one's existing intrinsic motivation to take creative risks.

Most creativity arises from a subtle mix of intrinsic and extrinsic motivation. So just as important as our type of motivation is our *perception* of ***why*** we are motivated. Creators with a strong sense of intrinsic motivation are more likely to feel in control of their processes of exploration and generation. They are also more likely to take ownership of a creative goal and to view it as a personal *mission*. In contrast, creators with a strong sense of extrinsic motivation are more likely to experience a loss of control, to refuse ownership of a creative goal, and to deny full responsibility for the quality of the final result. So we must be wary of ***over-justifying*** creative behavior, as an ill-conceived extrinsic reward can so upset the seesaw balance of intrinsic/extrinsic motivation that an intrinsically-motivated creator feels a loss of self-determination, or worse, a loss of mission. In any case, since we cannot simply *will* ourselves to be creative, an increase in extrinsic motivation does not always lead to an increase in creative output.

Nonetheless, one *can* be encouraged to be more divergent in one's attitudes, and so psychologists report more encouraging results when test subjects are rewarded for generating outputs that show a high degree of divergence. In other words, extrinsic rewards can be effective in fostering creativity if they are used to incentivize a move away from convergent thinking. Psychologists **Robert Eisenberger** and **Michael Selbst** describe experiments that show that when subjects are modestly rewarded for low-divergence outputs, a subsequent **drop** in creativity is observed, but when subjects are given the same modest rewards for high-divergence outputs, a subsequent **rise** in creativity is observed.

If the glove don't *fit*, you must *acquit*! With that memorable phrase, lawyer **Johnnie Cochrane** secured freedom for his client **O.J. Simpson** in the murder trial of the century. Cochrane's choice of words helped seal the deal with Simpson's jury by persuading them to focus on an apparent weakness in just one piece of evidence in the state's substantial case. How can a televised trial that lasted for over a year eventually boil down to a single creatively-turned phrase? What is it about the psychological allure of rhyme and metre that so affects our judgments of truth?

The Sphinx's phony profundities have no lasting value, and listeners see them mostly for what they are: glib fakes generated by the syntactic manipulations of a *one-size-fits-all* formula. But chiasmus and other rhetorical devices are not inherently limited to the generation of disposable linguistic ephemera, and many oratorical uses do show a lasting poetical value. So what gives any one of our attempts at creativity a resonance and a memorability that transcends their initial contexts of use? The answer certainly involves more than a single dimension of appreciation; a successful message will convey an apparent truth that resonates with its intended context of use, yet not be so tied to that context that it cannot be reused in new settings. It will be also packaged in a form that is aesthetically appealing. **John Keats** said it best, with chiasmus no less, in his ***Ode on a Grecian Urn***: "*Truth is beauty, beauty truth, that is all Ye know on earth, and all ye need to know*."

Cochrane used a rough poetry of sorts to persuade a jury to overlook his client's tumultuous history. In fact, we are all wired to find poetic forms both attractive and persuasive: following the work of **Matthew McGlone** and **Jessica Tofighbakhsh**, psychologists refer to the demonstrable influence of the attractiveness of a statement on its perceived truth as the ***Keats heuristic***. McGlone and Tofighbakhsh quote **Nietzsche** as a prior authority: "*We sometimes consider an idea truer*", he wrote, *"simply because it has a metrical form and presents itself with a divine skip and jump*."

When faced with a surfeit of messages to choose from, it makes sense to think that we assign more value to those that come in balanced and harmonious packages. Perhaps we believe that messages with a clever form are also more likely to result from some clever insight? Or perhaps surface symmetry is suggestive of a deeper symmetry of ideas? To test the ***Keats*** *heuristic*, **McGlone** and **Tofighbakhsh** presented test subjects with a mix of relatively uncommon aphorisms from a popular book of proverbial sayings. Each of the chosen sayings offered a general insight into the human condition, and each exhibited an appealing internal rhyme, such as "***woes*** *unite* ***foes***".

To specifically test the effect of rhyme on the comprehensibility and apparent truthfulness of each aphorism, different test subjects were presented with either the original aphorism or one of two non-rhyming alternate forms. The latter were created by replacing either the first or the second part of the rhyme with a synonym (to e.g. produce "*woes unite* ***enemies***" and "***troubles*** *unite foes*" from "***woes*** *unite* ***foes***"). McGlone and Tofighbakhsh observed that breaking the rhyme of an aphorism in this way did not affect its comprehensibility in any statistically significant way, but breaking the rhyme *did* significantly diminish the perceived truth of its underlying message amongst test subjects.

Hit Me With Your Rhythm Stick

The Keats Heuristic can be used to add a "*divine skip and jump*" to any phrasing, whether of a slogan, a motto, a line of poetry or a summation to the jury. As an audience, we are seduced by rhyme and by symmetry, and often perceive a superficial beauty as evidence of a deeper truth. Yet because orators use these tricks so well and so often, we are also wary of accidental symmetry and too-easy rhymes. While some wits work hard at crafting a pun, many of us instead feel the need to beg pardon whenever we unwittingly unleash one of these base rhymes on an undeserving world. When Scrooge first meets Marley's ghost in A Christmas Carol, he suspects (indeed *hopes*) he is experiencing a waking dream caused by indigestion, perhaps due to "*a crumb of cheese or a fragment of underdone potato*", and tells the ghost "*there's more of gravy than of the grave about you, whatever you are!*". Scrooge's use of rhyme, to glibly link two very different things, conveys his skepticism about the apparition's provenance: it merely looks like a ghost, just as "gravy" merely sounds like "grave." Scrooge mocks the Keats heuristic even as he exploits it!

- Find (or invent) some examples where the Keats heuristic undermines the perceived truthfulness of a statement.
- Select some famous lines with obey the Keats heuristic and break the internal rhyme in interesting ways using synonym replacement, colloquialisms, anachronisms, jargon, vulgarisms or words from another register (so e.g. "*To be or not to be*" becomes "*To be or not to bother*").
- What kind of anti-Keats changes make a line funnier, more memorable, stronger in tone or in meaning?
- How might an automated smartphone app use the Keats heuristic (or deliberately *violate* the Keats heuristic) to add verve to our tweets and SMS messages? What kind of rules might it use? How might it be able to recognize that certain rhymes are glib and counter-productive?

Classic films and books contain some wonderfully quotable lines, yet not every line in a classic is itself a classic. Most lines are simply perfunctory: they do their job and are quickly forgotten. Writers craft their most memorable lines using a process of selective generation, and others quote these lines on the basis of a comparable selectivity. So when taking a long view of creativity, it pays to ask whether the most memorably creative lines have any identifiable qualities that help to make them so quotable. A key factor is clearly utility; some lines convey a feeling or a meaning that one needs to convey in many different settings, and do so in a very concise fashion.

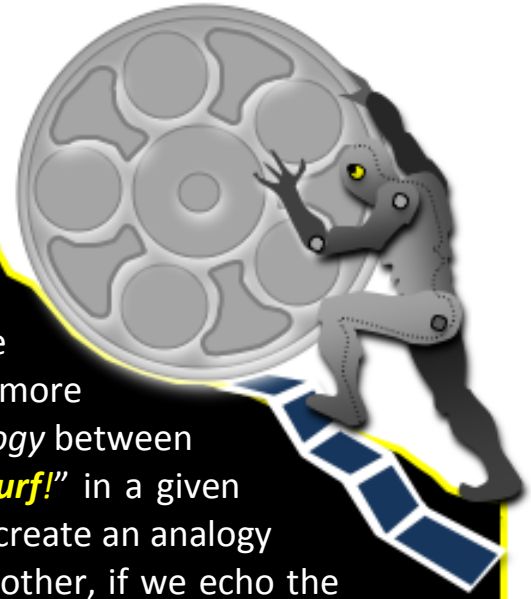
If one's aim is to offer an ironic comment on the heavy-handed insensitivity of another, then the lines "*I love the smell of napalm in the morning*" (from ***Apocalypse Now***) and "*You were only supposed to BLOW THE BLOODY DOORS OFF*" (from ***The Italian Job***) can be suitably mocking. The line "*You're gonna need a bigger boat*" (from ***Jaws***) likewise compresses a great deal of emotion – shock, horror, awe, dread – into a handful of words. In sharp contrast, the hard-to-place line "*That's some bad hat, Ha*rry" (also from ***Jaws***) has a rather limited reuse value and will be familiar only to trivia buffs and film completists.

So what makes one form more memorable

– or at least more reusable – than another? Computer scientists **Lillian Lee, Cristian Danescu-Niculescu-Mizil, Justin Cheng** and **Jon Kleinberg** of Cornell university analyzed thousands of pairs of lines from quotable movies to find some answers. Each line-pair matches a familiar movie quotation with a less memorable line (based on Web quotation statistics), where both lines have a comparable length and are used in the same movie scene. Test subjects were each presented with twelve pairs of lines from movies they had *not* already seen, and asked to memorize the lines for later recall.

Analysis of the results of these recall tests yielded findings that are more intuitive than surprising: on average, memorable lines contain more unusual words or word combinations, but these are conveyed plainly, using a commonplace syntactic structure (so e.g. "*a bigger boat*" is more memorable than "*a boat that is bigger*"). Moreover, memorable lines are more self-contained, and are thus easier to untangle from their contexts of use. In effect, memorable lines are **linguistic readymades**. Overall, a quotable readymade contains fewer 3rd-person pronouns such as "he" or "it" (pronouns like "me" and "you" are much more portable, as new contexts will always have a speaker and an addressee); fewer past tense verbs (which refer to events that have already happened and thus lie outside a scene); and more present tense verbs (such as "I *love*", "you'*re* gonna" and "these *aren't*") that can be ported with greater ease to new contexts. Lee and her colleagues also built an automatic classifier to predict which lines have the most memorable forms, and report that a classifier trained to notice features that convey both distinctiveness and generality performs best when predicting which line in a more/less memorable pair is the most quotable.

Intermission

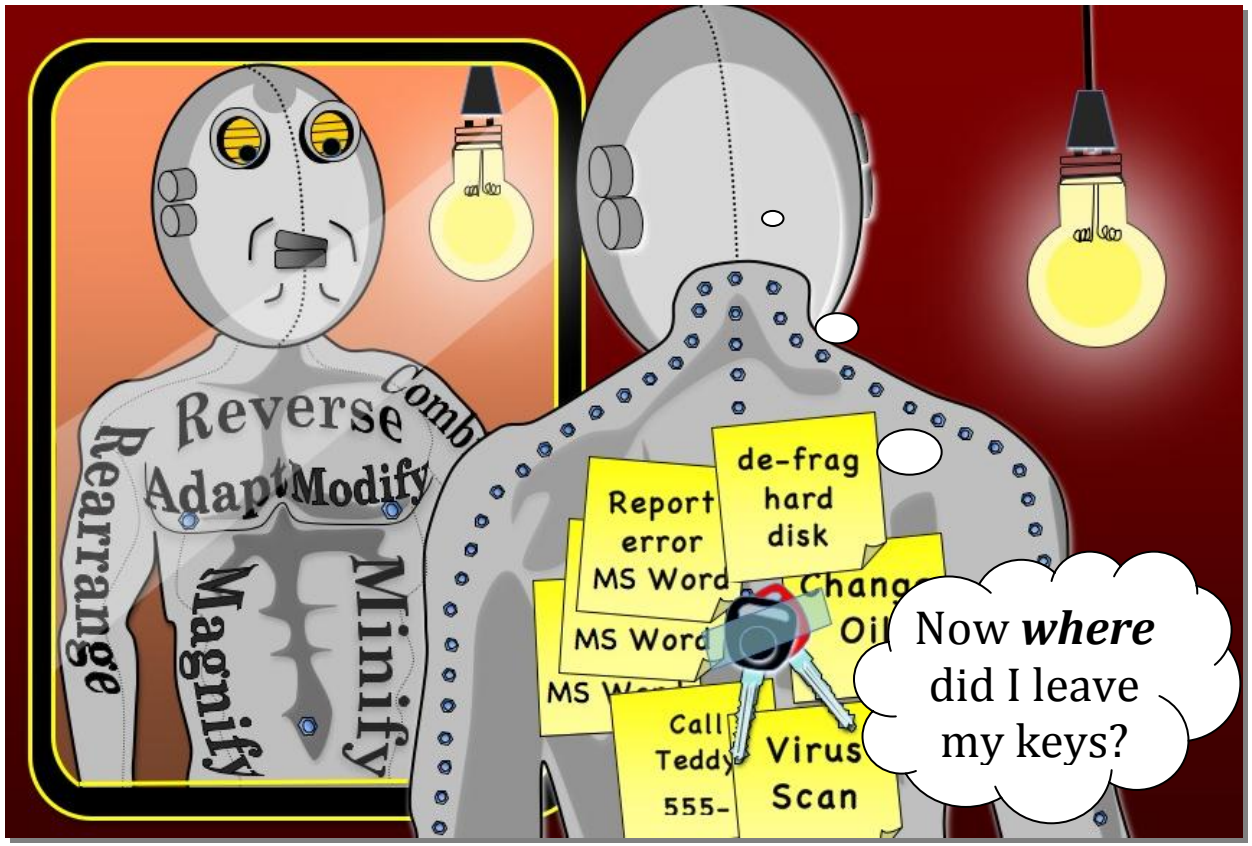


The most memorable movie quotes seem to distill a complex emotion into a single line. They capture the essence of a scene and allow us to easily refer to that scene and all its emotions, and unpack them all in a new context. Syntax and word choice certainly play a part in making a line memorable and reusable, but there is just as certainly more to memorability than this. For a good line evokes more than a memory: it establishes an *analogy* between the current situation and a scene from a movie or book. So when we say "***Charlie don't surf!***" in a given context, we do more than display our taste in movies by quoting from ***Apocalypse Now***; we create an analogy between the original speaker of the line (the patrician **Lt. Col. Kilgore**) and ourselves (or another, if we echo the line mockingly). By analogy, we dismiss the concerns of another as baseless, and laugh danger in the face.

- What other structural considerations make a line memorable and quotable in a new context?
- Do memorable lines always uttered by **memorable *characters***? Does a good line make a character memorable? Does a good characterization contribute to the memorability of a line
- **Woody Allen**'s play (and subsequent movie) ***Play It Again, Sam***, is named after a common misquotation from the movie **Casablanca**. When people misremember lines, are they (often) implicitly improving upon them? How so?
- As knowledgeable speakers, we often do more than repeat a memorable line from a movie: we adapt it to our current circumstances and needs by replacing some of its words (e.g., "*We're going to need a bigger* ***bath***" or "*In* ***IKEA*** *no one can hear you scream*". Is it possible to generalize about the kind (and amount) of words in the original – Nouns? Names? Verbs? – that must be preserved by a successful variation? Which ones can we vary at will?
- Find a memorable *counter*-example for each of the generalizations you suggest above.

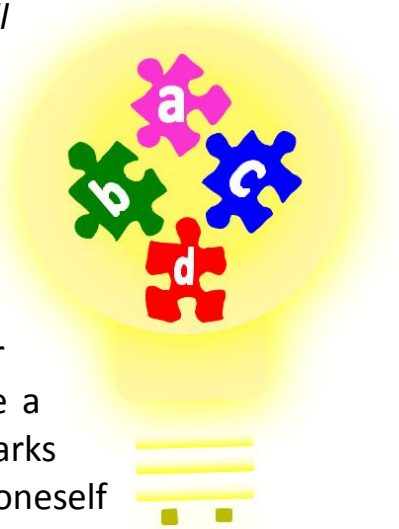
Extra credit: How might a computer quote famous movies in context, to suit its specific communicative needs? Do not imagine a computer that simply has a database of quotes and rules for using them.

Because of the ease with which we fall into ruts or find ourselves vexed by memory blocks, theorists and practitioners have long sought to systematize the processes of creative thinking. Approaches vary in their specific emphases, yet each approach can be seen as an attempt to place the free-ranging and highly associative abilities of **System 1** under the disciplined control of **System 2**. For the instinctive automatization of **System 1** can, ironically, be tamed by imposing upon it the deliberate and conscious execution of *algorithms* by **System 2**. A disciplined thinker can execute these in step-by-step fashion, perhaps even using pencil and paper.



These attempts at systematization have much in common with the algorithmic efforts of AI researchers to imbue computers with creative and/or generative abilities. For if the metaphor "*mechanisms of creativity*" does not strike you as a metaphor too far, or even as much of a metaphor in the first place, then efforts to explore our cognitive mechanisms in *mechanical* terms can hardly be a subject of much controversy. Indeed, if an understanding of the mechanisms of creative thought in humans can be an excellent guide to the construction of creative mechanisms in computers, then surely insight and illumination can flow in the other direction too?

The first step toward the systematization of creativity is the recognition that creative people may pursue a variety of different strategies to develop new ideas. These strategies can be applied in isolation or in combination, either once-off or recursively, for exploring a large space of conceptual possibilities. A comprehensive set of strategies, expressed as a checklist of questions to ask oneself when aiming to generate new ideas, was suggested by **Alex F. Osborne** – the inventor of "**brainstorming**" – in his 1963 book, ***Applied Imagination***. The title of Osborne's book suggests an appealing metaphor of mind, in which the creative imagination is viewed as an eager but undisciplined schoolboy whose report card might well read "*shows promise, but really needs to apply himself more*."



Given the difficulties we have in corralling creativity into a meaningful *one-size-fits-all* definition, creativity theorists show a certain fondness for less-definitive checklists. Osborne's checklist (on the next page) is designed to serve as a creative mentor in crib-sheet form, prompting a stymied innovator with a range of questions for escaping ruts and sidestepping memory blocks. So you think you've tried everything to no avail? Well, have you considered replacing one part with another (strategies **5a** … **5f**) or playing around with the arrangement of the parts you already have (strategies **6a** and **6e**)? How about combining a number of previous successes (**8a** … **8d**), or viewing your problem from a radically different perspective (**7a** … **7f**)? Or how about simply jettisoning some of your requirements (**3a**, **3g**)? These questions encourage a creator to fully engage the deliberative, analytical and reflective modes of thought that are hallmarks of **System 2** processing. By talking through a problem and consciously posing these questions to oneself (or to a collaborative group), a creator will simultaneously activate the reflexive thought processes that are characteristic of **System 1**. Possible solutions that are suggested by **System 1**, via the automatic recall of primed concepts from memory, can then feed into further reflective thought and further self-questioning. In other words, Osborne's checklist – like any other division of the creative process into specific steps, strategies or sub-processes – provides a schematic structure for harnessing the best qualities of **Systems 1** and **2**. In any such scheme, **System 1** must still be free to perform associative retrieval, even as its operation is shaped by the goals of **System 2**.

To Start

1. Adapt?

- *a. Are there new ways to use this as is?*
- *b. Other uses if modified?*

2. Modify?

a. New twist?	*b. Change meaning?*	*c. Change color?*
d. Change motion?	*e. Change sound?*	*f. Change odor?*
g. Change form?	*h. Change shape?*	*i. Other changes?*

3. Minify?

a. Subtract?	*b. Smaller?*	*c. Condensed?*
d. Lower?	*e. Shorter?*	*f. Lighter?*
g. Omit?	*h. Streamline?*	*i. Split up?*
j. Understate?		

Mains

4. Magnify?

a. What to add?	*b. More time?*	*c. More frequent?*
d. Stronger?	*e. Higher?*	*f. Longer?*
g. Thicker?	*h. Extra value?*	*i. New ingredient?*
j. Duplicate?	*k. Multiply?*	*l. Exaggerate?*

5. Substitute?

a. Who else?	*b. What else?*	*c. Other ingredient?*
d. Other Material?	*e. Other process?*	*f. Other power?*
g. Other place?	*h. Other approach?*	*i. Other tone?*

6. Rearrange?

a. Switch parts?	*b. Other pattern?*	*c. Other layout?*
d. Other sequence?	*e. Swap cause & effect?*	
f. Change pace?	*g. Change schedule?*	

Desserts

7. Reverse?

a. Swap Positive & Negative?	*b. How about opposites?*	
c. Turn it backward?	*d. Upside down?*	
e. Reverse roles?	*f. Turn tables?*	

8. Combine?

a. Blend?	*b. Alloy?*	*c. Assortment?*
d. Ensemble?	*e. Combine units?*	*f. Combine purposes?*
g. Combine appeals?	*h. Combine ideas*	

In computational terms, Osborne's checklist is a menu of operators for manipulating our mental representation of an idea, so that we can explore the conceptual space around that idea. To see these operators at work, or to at least perceive *post-hoc* the application of these operators in existing innovations, we can look to the patent archives to see how past inventors have generated viable new ideas by exploring outwards from an established starting point.

Osborne's checklist captures many of the state-change operators that have been identified

both in the psychology literature and by AI researchers. So let's look at the patent archives to see how real inventors use these operators to explore and innovate within a space that is densely populated with prior art. More specifically, let's consider the evergreen domain of ***word games***, for this is a space that contains a powerful source of *priming* – and *blocking* – in the always popular ***Scrabble***. While the rules and game elements of ***Scrabble*** are a fertile basis for innovative variation, the market dominance of *Scrabble* is just as likely to be a source of ruts and memory blocks.

To consider word games that openly acknowledge a similarity to *Scrabble*, we can search the U.S. patent archives (via **Google** or ***www.USPTO.gov***) for patents that contain the terms "*word game*" *AND* "*Scrabble*". In the many patents that are retrieved, we see a wide range of word games that branch out from *Scrabble* using one or more operators on Osborne's checklist. Many variations use the *Magnify* strategy to add a new wrinkle to an existing variant of the game. For example, giant tiles allow us to create a garden variant of *Scrabble* in which one's lawn serves as a game board (**4g**, **4l**, and *other place*, **5g**), while the addition of connectors to link these tiles together (**3i**) supports an aquatic version of the game, so that chatty swimmers can enjoy *Scrabble* while splashing in the pool.

The *blending* of different game formats – specifically, of a word game with a non-word game – is a popular strategy in the development of novel game patents. To make a new format viable, such blends often necessitate the substitution of existing game elements (**5**), and/or the addition of new elements from a second game ingredient to the blend (**4**). For instance, U.S. patent *7,497,778* describes a blend of a word-game and a ***lottery*** (**8a**, **8g**), in which players buy tickets marked with words rather than with numbers (*other ingredient*, **5c**). Players are assigned a mix of random letters (rather than random *numbers*, again **5c**), and the winning ticket is the one whose words contain the most letters from the draw. U.S. patent *6,761,642* describes a *Scrabble*-like game played on a ***pool table*** with balls marked with letters (**5c**), in which players spell words by using a billiards cue to skillfully sink letter-balls in orthographic sequence.

U.S. patent *6,460,855* describes a blend of *Scrabble* and ***Chess***, in which the squares of the checkerboard – rather than the movable game pieces – are marked with letters (**6c**, *other layout*, and **7e**, *reverse roles*). Players form words by moving the chess-like game pieces into the appropriate letter squares. U.S. patent *6,224,057* appears to combine Scrabble and ***Trivial Pursuit***, in which players answer questions and construct a list of words from the letters of the correct answer. U.S. patent *5,702,105* describes a blend of Scrabble and ***Ludo***, in which players form words on a game board so as to progress from the bottom to the top of the board. Looking through the patent archives, it seems that almost any table game can be blended with *Scrabble*, albeit with varying degrees of creative and commercial success.

Perhaps the most protean *Scrabble* blend of all is described in U.S. patent *4,333,656*. Two decks of playing cards are combined to form a double deck of 104 cards (**4j**, *duplicate*), which is just about the number of letter tiles in a Scrabble bag. Each card is marked with both a letter and a number, where the number – as in *Scrabble* – is an inverse function of the letter's frequency in English. The deck can be used to play a spelling variant of poker, or indeed of any card game in which players form the highest-scoring hands/words from the cards/letters they are dealt. Hands are scored by summing letter values in much the same way as in *Scrabble*, but each deck also holds a small number of blank wildcards that can substitute for any letter one wishes. This blend of *Scrabble* and poker obviates the need for a game board (**3a**, *what to subtract*?), while importing some of the best elements of poker (**2a**, *new tw*ist), such as ***betting*** and ***bluffing***. As in poker, players may pretend to have an excellent hand, bet accordingly, and only have their bluff called when opponents pay to "see" their words.

Even the simplest variations are patentable if they are non-obvious, but there is rarely such a thing as a non-obvious variation that can be implemented with a single change. Due to the subtle interdependence of parts and features, one change – whether an addition, a magnification, or a removal – almost always necessitates others. Each change thus requires a period of reflection and engagement on the part of the inventor, using the deliberative processes of **System 2**, so that these changes can be successfully integrated into the whole. But note how each change also crucially depends on **System 1** to suggest the elements of the mental representation to which the change will apply. It is the associative architecture of **System 1** that allows us to bring to mind the salient attributes of a concept when we explore a *minification* (**3**) or *magnification* (**4**) of qualities, or a *substitution* (**5**) or *rearrangement* (**6**) of parts. For a successful modification we need to apply the discipline of **System 2** in what is the province of **System 1**.

In its efforts to grease the rails of our most habitual thought processes, *System 1* primes the aspects of a word or idea that are most salient to how it is typically used. This priming is helpful more often than not, yet priming can also inhibit our ability to look past surface associations and appreciate diverse points of view. Try as it might to be helpful, priming often blocks our efforts to reinvent the familiar and to see the world through different, *divergent,* eyes.

Attribute analysis is an old and widely-taught technique that helps us to peek under the covers, to see past the automatic judgments of **System 1** and engage **System 2** to consciously explore the non-obvious possibilities of a familiar object or idea. A user of attribute analysis begins simply, by listing all of the attributes of the object or idea under analysis, ranging from the most to the least obvious.

For each of these **attributes**, a user then makes a list of other concepts that exhibit the same attribute. Each of the items on this second list – many of which might not have come to mind had we relied wholly on priming by **System 1** – will share one or more attributes with our starting concept, and some may even suggest **unusual new uses** for this concept. Attribute analysis can thus forge new links between ideas that were not previously connected in our minds.

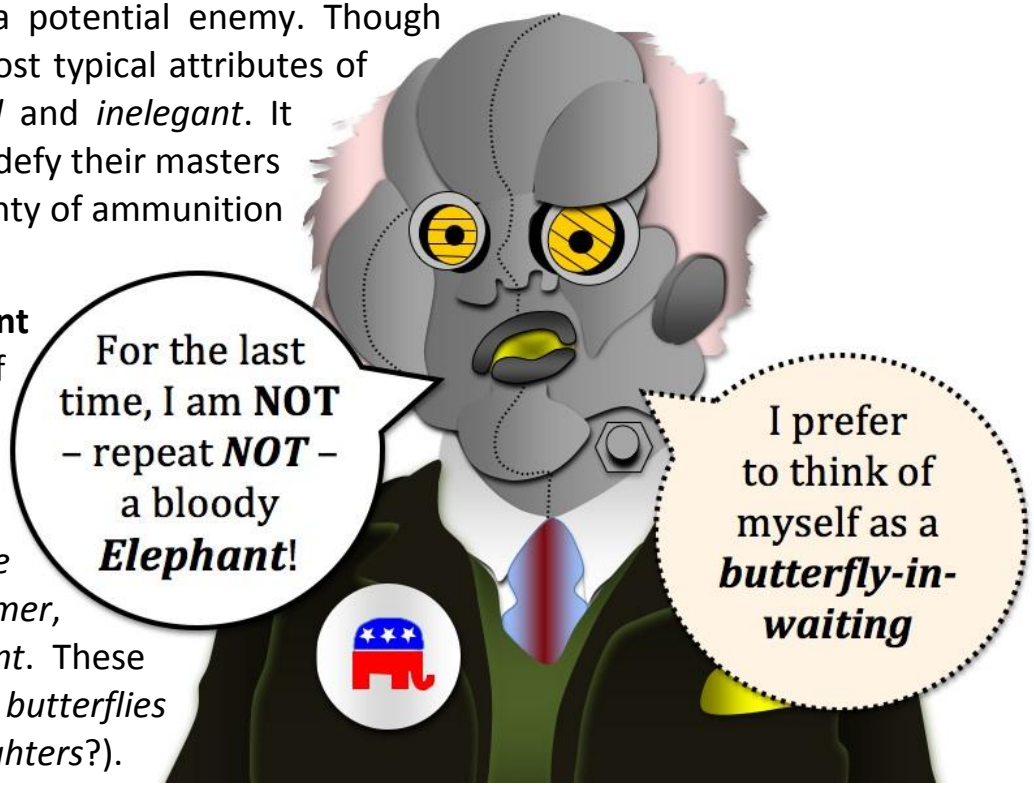
If we buy a kitchen cleaver while on vacation in China, an automated check at the airport is very likely to flag our purchase as a potential weapon because sharp knives and cleavers sit high on a standard list of prohibited objects. But now suppose we buy a bejeweled letter opener while holidaying in **Morocco**: though this decorative object also has a potential use as a lethal weapon, it is unlikely to be flagged as a threat by a security system that relies on a list – even a long list – of proscribed objects. Ideally, a reliable, automated security check should perform a thorough analysis of the attributes of each object to identify those that might be used as lethal weapons, and neither flag too many nor too few as potential threats. To catch a creative rule-breaker, rule systems must themselves possess a creative imagination. Sadly, this example is not entirely fanciful. Human security guards could have instinctively performed attribute analysis to see the potential for box cutters to be misused as weapons when the **9/11** terrorists boarded their planes in 2001, but they did not. FBI agents might also have recognized the potential of fully-fueled planes to serve as guided missiles, but they did not. While many cumulative mistakes contributed to the tragedy of **9/11**, one was surely an over-reliance on **System 1** over the mentally-demanding engagement of **System 2**.

Attribute Analysis is instinctively used whenever a woman uses lipstick to write a note on a mirror, or when a handyman uses a coin to ease off the lid of a paint can, or when a programmer uses a web-email account as a backup store for important documents, or when a homeowner reaches for a golf club when suspicious sounds come from downstairs in the dead of night. So **Attribute analysis** formalizes a process that we use automatically whenever we react creatively to a situation, but it puts this process under the conscious control of **System 2**. When unable to use an object that typically provides us with the **attribute *X***, another object with ***X-ness*** can often serve the same function.

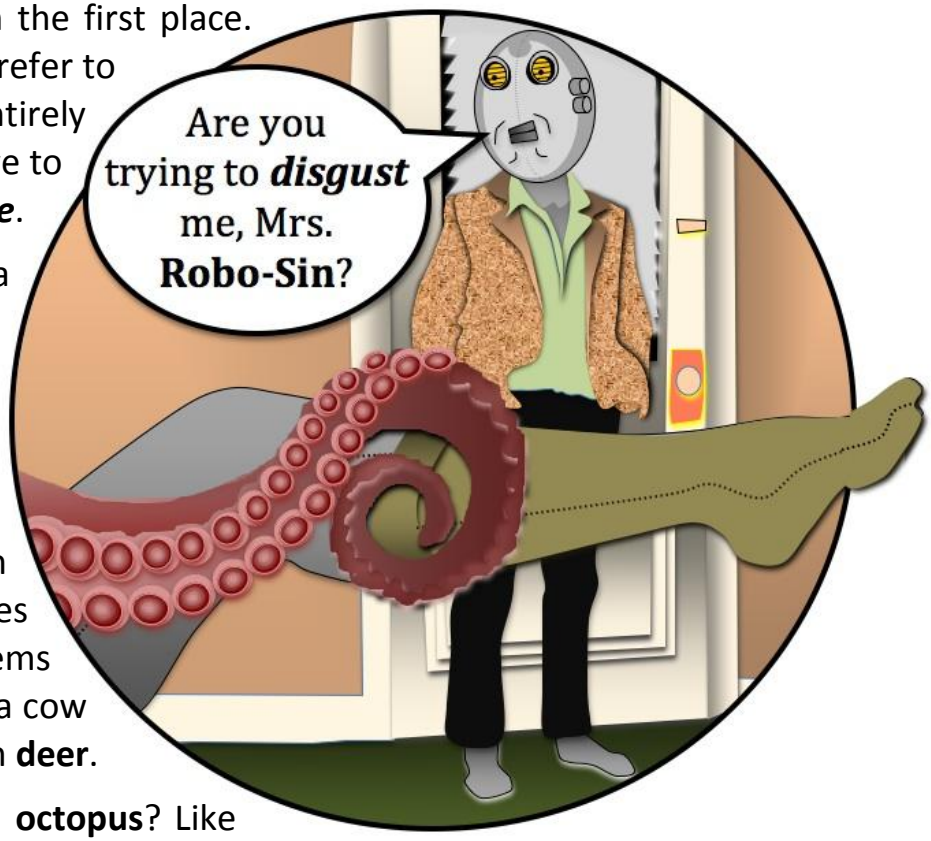
Don't think of an ELEPHANT!* This injunction famously pits our *System 1* against our *System 2*,** for the very mention of the word "elephant" directs our **System 1** processes to prime the concept ***Elephant and drag our knowledge of stereotypical pachyderms into focus. Once we make sense of the larger imperative, it is too late: **System 2** cannot simply banish elephants into the lower recesses of our memories from whence they were dragged. Indeed, any efforts to focus on ***non***-elephants puts greater emphasis on *elephants*, and primes ***Elephant*** even more.

Attribute Analysis is designed to overcome the problems that pit the mechanisms of the kind against each other. Suppose one is given the task of suggesting a range of animals that are the furthest things from *elephants*. It may be that a military manufacturer has proposed a new form of weaponized personnel carrier nicknamed **The Elephant** and we wish to critique its design from the perspective of a potential enemy. Though **System 1** inevitably primes ***Elephant***, it also primes the most typical attributes of elephants, such as *heavy*, *large*, *ponderous*, *thick-skinned* and *inelegant*. It may also prime the fact that some elephants ***go rogue*** and defy their masters with disastrous consequences. These attributes provide plenty of ammunition to use against the maker of a weapon called **The Elephant**.

Now suppose we want to propose a *counter*-**Elephant** weapon that embodies qualities of a very different kind of animal. In effect, we ask ourselves to *not* think of an *elephant*-like animal. By analyzing each of the elephant attributes that are primed by **System 1**, we can invert each to obtain its opposing quality: so *heavy* becomes *light*, *large* becomes *small*, *thick-skinned* becomes *soft* or *gossamer*, *ponderous* becomes *agile*, and *inelegant* becomes *elegant*. These new attributes might, in turn, lead us to think of moths and *butterflies* (→ *gliders* perhaps?) or *hornets* and *mosquitoes* (→ *snub fighters*?).



What name is given to a female octopus, other than *honey* or *dear*? We use a range of names to distinguish the male from the female of a species, such as ***stallion*** for a *male* horse and ***mare*** for its *female* mate. When it comes to ***octopi***, though, few of us can name the female of the species just by pulling a name from memory. The question, however, is designed to elicit a *creative* answer, by spurring us to engage in the same kind of ***analogical reasoning*** that guides us in our creation of new animal terms in the first place. Language may be abundant but it is not profligate, and we often prefer to find new uses for existing words than to unnecessarily create entirely new ones. Thus, because zebras resemble horses, the name we give to the *male* **zebra** is ***stallion***, while the *female* is likewise called a ***mare***.



Convergent thinking can be used to suggest gender-names for a given animal based on taxonomic similarity to a popular exemplar. A *female* **donkey** is thus called a ***mare***, while **unicorns**, as mythical horses with horns, can also be classified by gender as ***mares*** and ***stallions***. But divergent thinking, as guided by attribute analysis, is needed to find gender-names for animals like the elephant, whose female is called a ***cow*** and whose male is called a ***bull***. As with camels, hippos and rhinos, these slow, heavy, ponderous creatures share enough attributes with bovines to share their naming systems too. Likewise, a *female* **gazelle** is called a ***doe*** rather than a mare, a cow or a bitch, again because of the many attributes gazelles share with **deer**.

So to return to our initial question, what do we call a female **octopus**? Like insects, ***Octopi*** are cold-blooded creatures that lay a multitude of eggs, yet we do not call them ***queens***. For ***Octopi*** have another fascinating attribute: like giant squids, they possess *hard beaks*. The attributes ***beak*** and ***egg-laying*** thus lead us to name the female octopus as a "**hen**". It seems the ***octopus***, and not the ***tuna***, is the true chicken of the sea.

Korean food is hot, spicy and delicious. But as any Korea-town dry-cleaner will attest, it can also be gloriously messy to eat. If this is true of the roiling stews that are served in seething table-top cauldrons, it is doubly so of **Bulgogi** ("*fire meat*"), the marinated strips of beef or pork that are cooked at your table on open grills. To add to the mess quotient, Koreans take pieces of meat from the sputtering grill and wrap them in green leaves (or *ssam*) with spoonfuls of rice, raw garlic and chili paste, to shape into little morsels with their fingers. If Korean food has not achieved the ubiquity of Chinese, Japanese, Thai, Indian or Mexican food, this is in large part due to its lack of a clean and convenient delivery format, so that it may be eaten on the fly without first donning a pair of overalls.

Attribute analysis can be used to find a viable solution to this problem, by identifying analogies to other culinary domains that have faced, and resolved, similar dilemmas. If we consider the salient attributes of Korean cuisine to be *spicy*, *chili*, *marinade*, *pickle*, *grill*, *ssam*, *lettuce*, *wrap* and *rice*, then attribute analysis leads us to ***Chinese***, ***Thai***, ***Indian***, ***Mexican*** food as likely analogues. ***Chinese*** and ***Thai*** food both make use of convenient forms such as the *spring roll*, while ***Indian*** food uses *kebabs* and *pastry samosas* to achieve an ergonomic *on-the-go* handiness. It is ***Mexican*** food, however, that wins hands-down for its hands-on consumption, offering *tacos* (meat and vegetables stuffed into crispy corn shells) and *burritos* (food tightly wrapped into tortillas) for formal and casual dining either in restaurants or on the street.

The ***Kogi burrito*** and ***taco*** are Los Angeles inventions in which Korean barbecue is packaged in the corn shells and tortillas of Mexican cuisine. These novel dishes, which blend Korean barbecued meats with Mexican extras like salsa, guacamole and sour cream, are popular with many LA communities. But **Kogi** also borrows another ***Cal-Mex*** delivery solution, the ***burrito truck***, to sell its wares on the go. Using *Twitter* to publicize its new routes, **Kogi**'s trucks are followed around LA by eager consumers.

To promote a deeper engagement with problems in science and society, Fritz Zwicky, a Caltech astronomer, vigorously championed a more systematic form of **attribute analysis** he named ***Morphological Analysis***. In language, the morpheme is the smallest carrier of meaning, and words derive both their form and their meaning from a specific sequence of morphemes (such as ***un-know-able***). The **Oxford English Dictionary** word of the year for 2012, "***Omnishambles***", demonstrates the power of new words to delight by combining familiar morphemes, while the **e. e. cummings** invention "**manunkind**" shows that even the most solid compounds are not immune to morphological change in our drive for linguistic novelty. The "***morphemes***" of a problem domain are the combinatorial units from which new solutions are created, and so the first step of any morphological analysis is the identification of relevant morphemes. If our goal is to invent a new mode of transport, we might choose morphemes to encode the dimensions of ***scale*** (e.g. *individual, small, massive*), ***power source*** (e.g. *oil, wind, solar, electric, nuclear*), ***passenger type*** (e.g. *commuter, tourist*) and ***medium*** (e.g. *air, road, rail, sea*). A candidate solution is formed by selecting one morpheme from each dimension, to yield attribute combinations such as ***massive-solar-commuter-rail*** (e.g. a mass-transit solar-powered rail system for commuters, such as a monorail in an always sunny country).

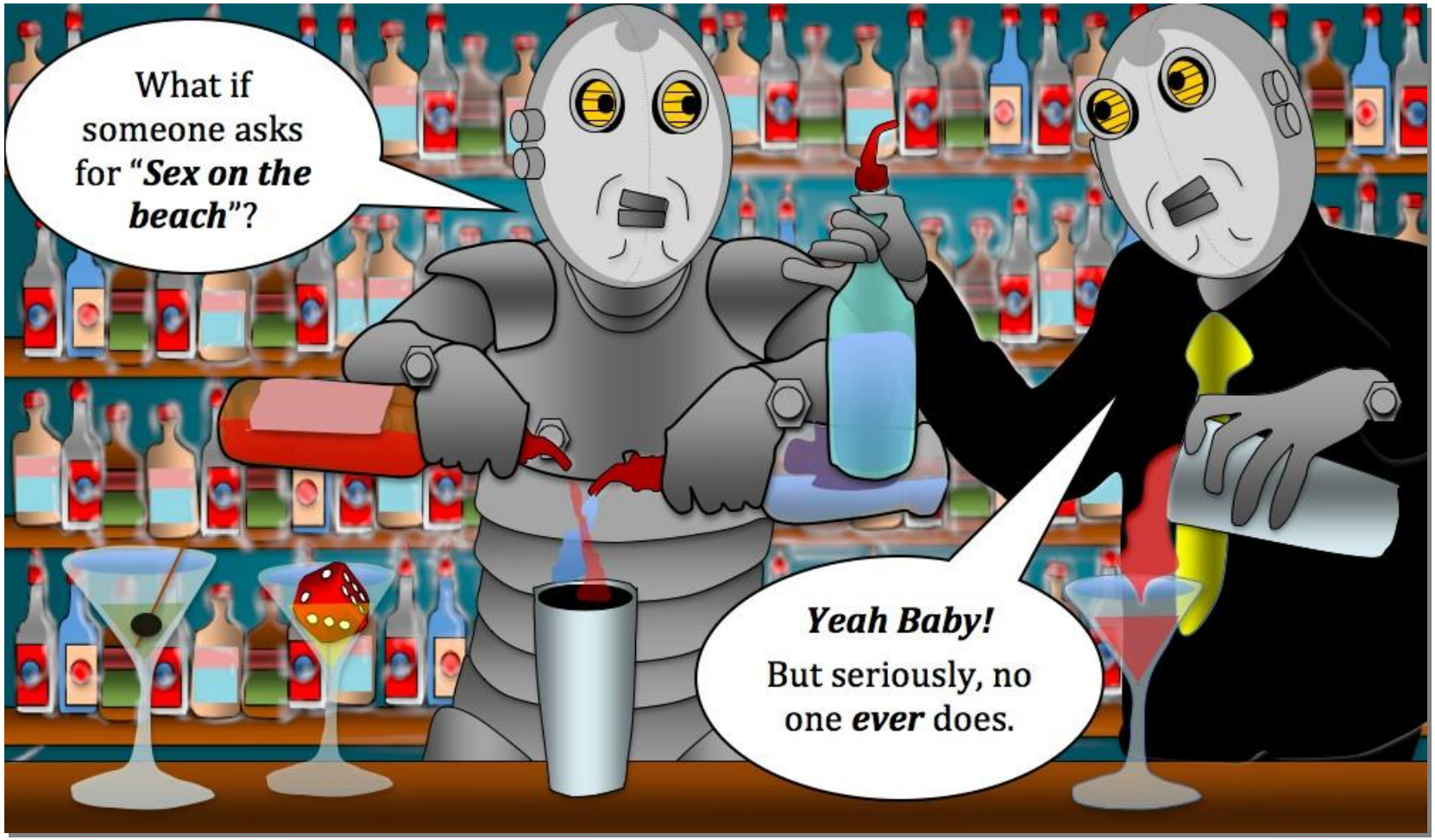
To invent a new game, we should instead draw our stock of morphemes from the combined elements of exemplars like ***Scrabble, Poker, Trivial Pursuits, Chess*** and ***Ludo***. Imagine dumping a bunch of games and their paraphernalia onto the floor, and then sifting through the resulting mess to thematically group elements and rules *across games*. Unlike basic attribute analysis, morphological analysis aims to capture the recombinant possibilities of a whole domain at once, by viewing each solution as a multi-dimensional structure in which alternate attribute values can be chosen for each dimension. Analysis begins with the assignment of each attribute to a single dimension, and by assigning each dimension to a column in a tabular structure called a ***Zwicky box***. Each cell in a Zwicky box holds one attribute in one specific dimension, while other cells in the same column hold alternate values for that dimension. A morphological solution must specify a morphemic attribute for each of ***N*** dimensions, as selected from the ***N*** columns of the Zwicky box. An incisive representation will carve up a domain in a way that allows viable solutions to be composed from attributes that combine freely with others, while a clumsy partitioning will yield a very poor harvest.

The Zwicky Box on the right divides the game domain into five simple dimensions: the type of **board** used, if any; the key **resource** that is acquired or allocated; the **objective** of the game; the main **strategy** employed; and the game's main source of **randomness**, if any. As highlighted in this box, *Scrabble* is a game played with jumbled letter tiles on a grid, where the goal is to achieve the biggest (cumulative) score by forming words.

Board type	Resource	Objective	Strategy	Randomness
None	Playing Card	Progress most	Connect	None
Checkerboard	Stone	Finish first	**Form words**	Dice
Map	Building	Acquire most	Block	Spinner
Circuit	Army unit	Deplete most	Trade	Shuffle
Track	Spell	Cover most	Encircle	**Jumble**
Grid	Play Money	**Biggest score**	Alliance	Blindfold
Maze	Counter	Last longest	Capture	Coin toss
Path	**Letter tile**	Kill opponent	Bluff	Timer
Honeycomb	Hit Points	Find treasure	Form patterns	Short Straw

Yet this representation is still very **under-specified**. It assumes, for instance, that it is the letter tiles that are jumbled, that only valid words are formed from these jumbled letters, and that scoring is cumulative across these words.

So a selection of context-free morphemes is not in itself a valid solution, merely the raw material of a potential solution that still needs to be molded by subsequent analysis and interpretation. Any combination of these morphemes will ultimately make sense only with the help of **System 2**'s ability to reason and fill in the implicit gaps.



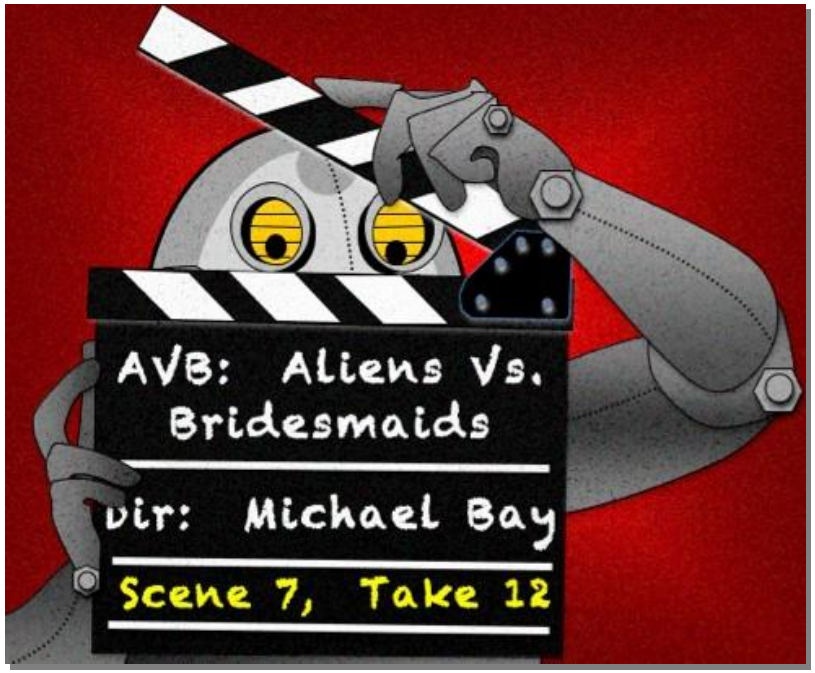
Combinatorial Creativity

Zwicky viewed his morphological approach to the combinatorial generation of ideas as a kind

of **totality research**, in which one avoids ruts and mental blocks by removing biases and allowing all possible solutions to a problem to emerge. So it is hard not to feel we have done a disservice to the totality of the game space with our rather coarse set of morphemes, and that more nuanced dimensions are called for, such as e.g. *major skill needed*, *major constraint imposed*, or *bonus elements*. Yet as we add columns to a Zwicky box, or add more morphemes to its columns, the number of possible combinations grows with alarming speed: ***N*** dimensions of ***M*** morphemes yields a potentially huge space of M^N combinations.

But not every pairing of attributes is sensible or even possible, and so morphological analysis performs a pairwise assessment of all attributes from different dimensions – called a ***Cross-Consistency Assessment*** (**CCA**) – to determine which pairings are viable and which are not. For instance, ***Form words*** **+** ***Letter tiles*** is clearly viable, for as in *Scrabble*, words can be formed from the letters. However, ***Form words*** and ***Kill Opponent*** is a most infelicitous combination, so any potential solution that pairs these together may be discarded as unworkable. In fact, any possible solution that contains a pairing of morphemes that **CCA** seriously frowns upon can be discarded, so that only the truly viable solutions pass muster for further consideration and analysis. Zwicky referred to this means of shrinking the space of possible combinations to better approximate the space of viable solutions as his *principle of contradiction and reduction*. In effect, **CCA** uses an analysis of the consistency of $N^2M^2 - N^2M$ attribute pairings to filter a space of M^N attribute combinations. Since the latter grows much faster for an increasing ***N*** than the former, CCA proves to be a wise investment of time and effort. Nonetheless, as demonstrated in our earlier patent survey, a flexible mindset will often find ways of making the oddest attribute pairs work well together. For instance, we can ***Form Words*** **with** ***Army Units*** **on a** ***Checkerboard*** if we mark the squares (rather than the pieces that sit on them) with the letters needed to form words, as described in U.S. patent *6,460,855*.

Robert Altman's film *The Player* opens with a glorious tracking shot that cranes and swoops through the offices of a fictitious Hollywood film studio. Joining the camera on its sweeping arcs, the audience is allowed to eavesdrop on insider conversations that illuminate the Hollywood philosophy of filmmaking. In one office a group of technocrats discusses the longest tracking shots in the history of cinema, while in another a writer eagerly pitches a new idea to a studio executive as a high-concept mix of ***Ghost*** meets ***The Manchurian Candidate***. We can only marvel at a writer's reasons for wanting to marry a blockbuster romance about a lovelorn ghost to a dark political thriller about brainwashing and McCarthyism. But Altman uses satire to suggest that Hollywood's dependence on the high-concept pitch, combined with a ruthless determination to recycle its past successes, has led to artistic stagnation and a dumbing-down of its creative processes.

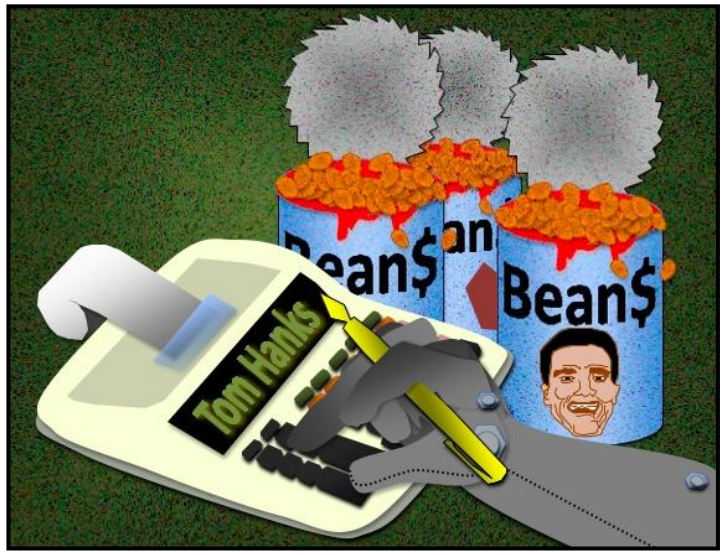


Successful artists of all stripes explore a space of possibilities – often a space of potential combinations – though each may use different sensibilities, goals and quality metrics to guide their explorations. The pursuit of successful combinations is as good a basis as any for navigating a conceptual space, to discover something novel and useful among the old and the familiar. In any case, even if a creator explicitly rejects combinatorial reuse as a creative strategy, a critic may not reject it as an *interpretative* strategy, and so an end result may nonetheless *appear* to be a deliberate blend of past successes. So in the final analysis, it is not the combinatorial pursuit of novelty and usefulness that offends us, but a single-mindedly economic rationalization of these qualities. Yet big Hollywood studios are big businesses, after all, with investors who frown on experimentation for its own sake. So it is hardly unsurprising that commercially savvy filmmakers tacitly use a variant of Zwicky's morphological analysis to combine the best elements of whatever has worked well in the past.

Even when seeking to invent the future we cannot ever truly escape the past, for even the newest products are inevitably constructed from already existing parts. Creativity ***ex nihilo*** – the making of something new from nothing – was once viewed as a manifestation of divine power, and though we now view creativity as a very human quality, implicit in this view is the realization that nothing is ever truly new. The book of **Ecclesiastes** in the **King James** bible says it best: "*The thing that hath been, it is that which shall be; and that which is done is that which shall be done****; and there is no new thing under the sun***" (*Ecclesiastes,* chapter 1, verse 9).

Hollywood films range from space operas in galaxies far, far away to the domestic concerns of everymen and their partners to the love lives of animals, ghosts and the tiniest of bugs. Nonetheless, the space of narrative ideas explored by filmmakers is as small as it was in the days of Greek tragedians and village griots. Past successes and failures are the landmarks that dot this space, and these can cast some very long shadows on its most travelled paths. Yet, as landmarks, they also help to guide our way, acting both as compass and *reverse* compass. One can hope to stake out some new territory for a new landmark of one's own, but it will sit in a place that will always be located relative to some other landmarks in the space.

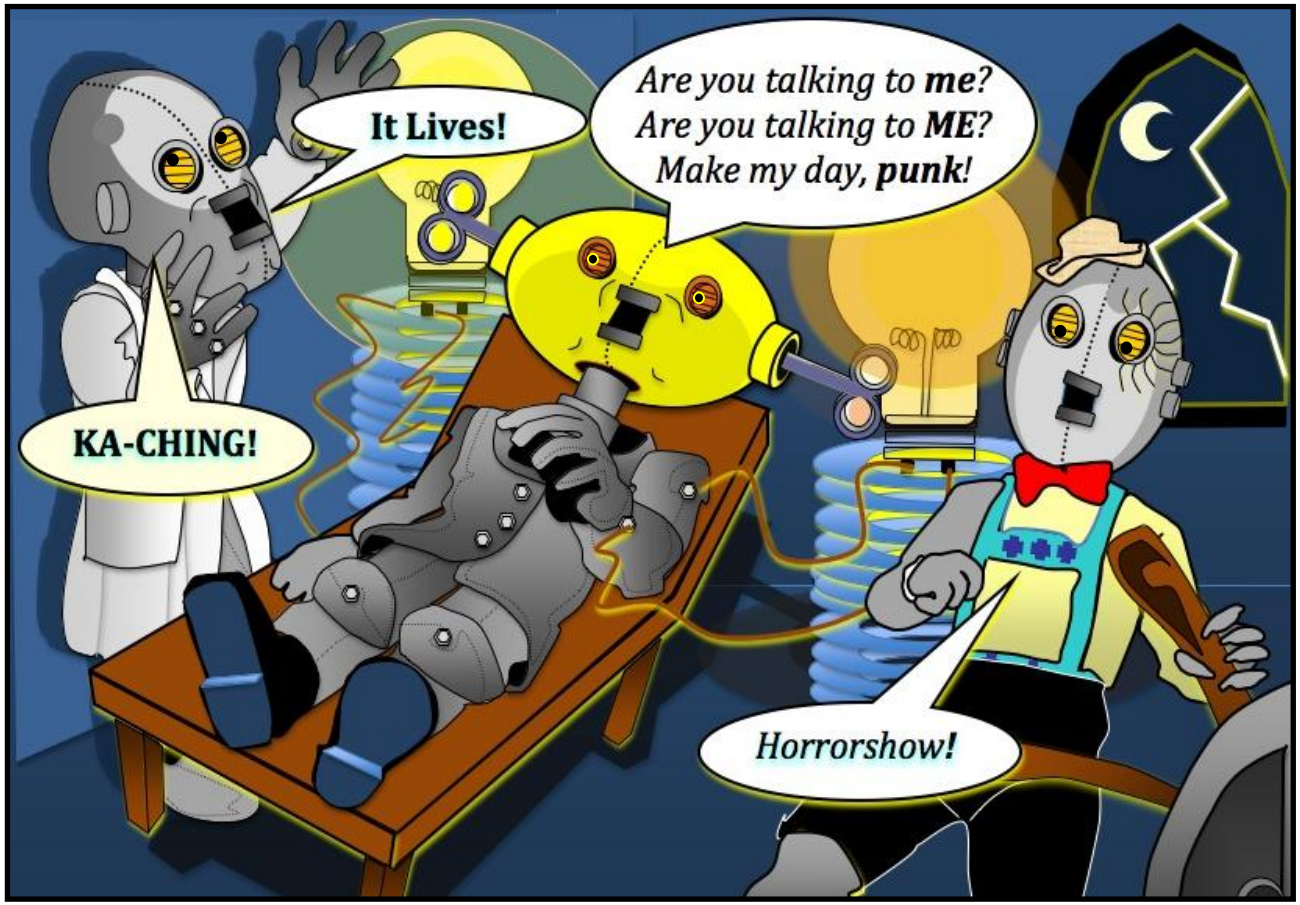
Spreadsheets offer an expressive means of representing a Zwicky box in our combinatorial view of movie-making. For we can assign a different column to each morphological dimension – from ***leading man*** and ***leading woman*** to ***director*** and ***screenwriter*** to ***genre*** and ***sub-genre*** to ***target audience*** – and a unique cell to each distinct morpheme, from ***Tom Hanks*** to ***Ang Lee*** to ***Sci-Fi***, ***Romance***, ***Comedy*** or ***War*** to ***U, PG, R*** or ***X***. If a minority of movies are too quirky to fit into this combinatorial mold, this is simply further proof of the mold-breaking creativity of those films. Spreadsheets also allow us to write formulae to perform a CCA-style cost-benefit analysis of different morphemes and combinations of morphemes. Though Hollywood executives are notoriously poor at predicting the commercial or critical success of a planned production (the screenwriter **William Goldman** famously articulated this view in his oft-quoted maxim "*nobody knows anything*"), studio bean-counters have developed remarkably precise financial models for predicting the haul of a movie from its opening box-office take. Since the bankability of stars, directors and genres and the profitability of the resulting combinations are a key determinant of success for a studio, it makes good financial sense (if not good art) to weave these dimensions into every column and cell of a Zwicky box.



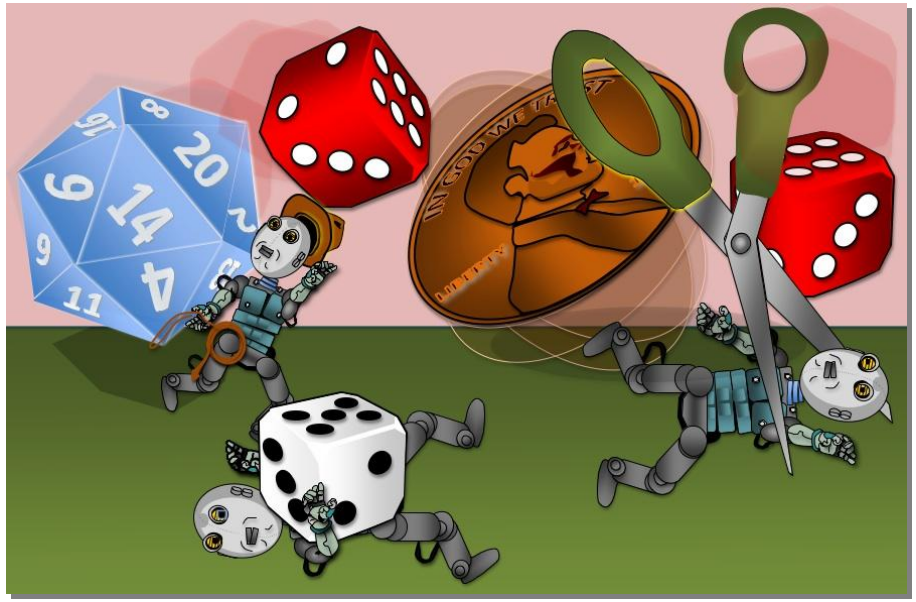
Once we populate a Zwicky box we can generate a specification for each innovation – in this case, a new movie – simply by reading off a single morpheme from each column. ***Cross-Consistency Analysis*** (**CCA**) weeds out some duds by performing a pairwise analysis of morphemes to ensure that no two morphemes are mutually incompatible. A set of loose CCA constraints can allow many quirky combinations to pass muster and merit deeper consideration, while too tight a set of constraints simply deepens the old cinematic ruts. A balance must be struck between being too permissive and too restrictive: CCA must reduce the space of possibilities to a manageable scale, but not at the cost of favoring the same *clockwork lemons* time after time.

Despite Hollywood's reputation for conservatism, impossible combinations that no producer would ever pitch or no studio would ever finance are hard to imagine. **John Wayne** as ***Genghis Khan****?* OK, you got it! **Mel Gibson** as ***Hamlet***? Sure! **Keanu Reaves** as ***The Buddha***? Why not? **Ang Lee** as director of ***The Hulk***? Well, okaaay! So the constraints imposed by CCA for this task are *soft* rather than *hard*, calling for a punitive scoring of silly pairings that are unlikely to yield profits or awards, and a generous scoring of those that seem safe, synergistic and bankable.

We can use CCA to assign an overall score to a specific mix of morphemes simply by summing the scores for each of its cross-pairings. Of course, it is a racing certainty that a truly random selection of morphemes – in which one morpheme is simply plucked at random from each Zwicky column – will receive a very low score, whether it has any artistic merit or not. Nonetheless, a risk-averse producer can always repeat this randomized procedure until a high-scoring mix is generated. As **Brad Grey**, head of **Paramount**, puts it, "*the upside for getting it right can be enormous, but the downside for getting it wrong can be enormous*".



The Hollywood *machine* is so-named for good reason. This combinatorial approach to the packaging of a new film has been systematized and streamlined to the extent that it has unflattering similarities to both an assembly-line and a sausage factory. Indeed, many of the same criticisms have been leveled at the Hollywood approach as they have at the computational approach to creativity, and the studio development process even has striking parallels to some common search strategies in AI. Consider our initial random selection of morphemes. Since this selection is likely to yield a very low score, one can roll the dice a second, third or fourth time, hoping each time to hit upon a combination with a more acceptable score.



Yet blind variation is an unproductive strategy that is unlikely to yield good results without an inordinately large number of do-overs. In contrast, the Hollywood production process is one in which many small random changes are made by a steady stream of interested parties, each offering an opinion on how a combination might be incrementally improved. An AI technique called ***hill-climbing*** works in much the same way: a localized random change is made to a single morpheme – for example, replacing ***Danny De Vito*** with ***Nicolas Cage*** in a ***Lincoln*** biopic – and this change is kept only if it improves the score of the overall combination. In contrast, any change that diminishes the overall score is rejected. A further random change – such as replacing ***John Woo*** with ***Steven Spielberg*** as director – must now push the overall score even higher if it too is to be kept and not rejected. In this way, small local tweaks gradually yield large global rewards, yielding perhaps a film in which bankable award-winner **Daniel Day-Lewis** stars as **Abe Lincoln** in a film from by a director with a track record in profitable historical dramas. However, just as successive improvements will eventually yield diminishing returns, hill-climbing is just as likely to lead us to a plateau of inoffensive mediocrity.

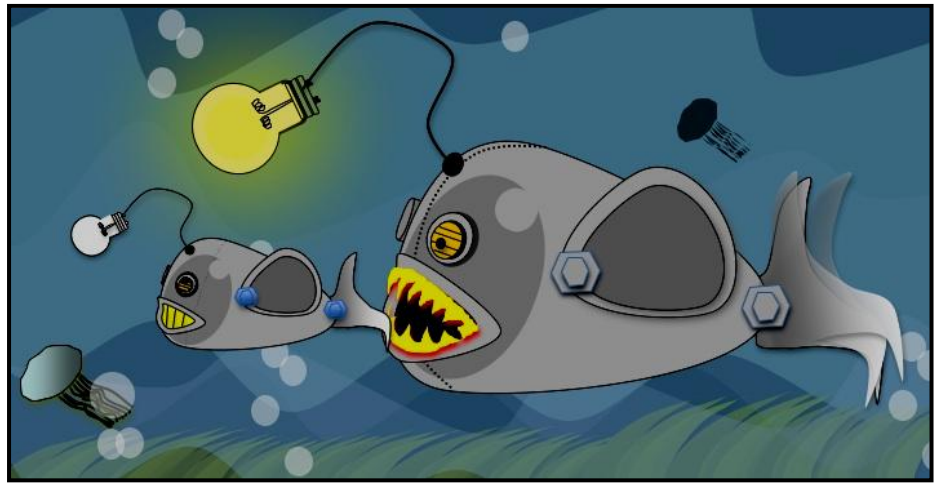
At this point it is wise to beat a tactical retreat, to allow a succession of small localized changes to diminish the overall score as we climb down from this plateau. Only then can new changes begin to reach for greater global rewards as our search pursues an alternate avenue. AI provides another search technique called ***Simulated Annealing*** (**SA**) to facilitate this tactical back-pedaling, though any "*annealing*" here is more aptly considered a metaphor than a simulation. The metaphor is based on the physics of how a hot molten metal cools and *anneals* into a cold rigid state. A hot metal is malleable and workable because its atoms have a surfeit of kinetic energy that allows them to move more or less freely. As the metal cools, its atoms relinquish their energy and begin to slide into the positions they will occupy when the metal is cold.

If cooled in an orderly fashion, these atoms will form the disciplined phalanx that is characteristic of a strong metal that is neither soft nor brittle. Simulated annealing is so called because it simulates the effect of a metaphoric temperature on a search process like hill-climbing. This "temperature" is high at the start of a search, but is gradually lowered as the search progresses. As in hill-climbing, small random changes are enacted if they raise the overall score of a combination, though changes that lower the overall score of a combination are occasionally allowed, following a probability distribution that is defined relative to the global temperature. When the temperature is high, punitive changes are frequently allowed, just as atoms can move freely in a molten metal. As the temperature drops, punitive changes are accepted with lower and lower frequency, as the search gradually eases into a standard hill-climbing pattern. If the search reaches an undesirable plateau early in its explorations, its high temperature will readily facilitate the backtracking that is needed to retreat to an alternative part of the search space. This flexibility loosely approximates the brainstorming and spit-balling that occurs early in the movie development process, when a project is in flux and the silliest of ideas is openly considered. Yet as work continues and a workable combination begins to solidify, such changes become harder to justify as modifications become more conservative.

Hill-climbing and simulated annealing offer good models of how incremental changes can be made to a combination so as to optimize some measure of quality, such as numerical predictions of profitability, audience satisfaction or critical acclaim. But they shed no light at all on Hollywood's penchant for recycling past successes in new blends of existing combinations. For this, we must turn to yet another AI search paradigm known as the ***Genetic Algorithm*** (**GA**). Suppose we use our Zwicky table to generate a whole population of chance combinations, each one randomly selecting a single morpheme from each column heading. This population of combinations will all draw on the same genetic material, where each column heading is a different *gene*, each morpheme under each heading is a different version (or ***allele***) of that gene, and each combination of morphemes is a unique *chromosome* that provides its own alleles for the same gene positions.

Our population of random combinations can be viewed as a population of chromosomes, each eager to thrive and pass on their genetic material to their offspring. As in the real world, offspring are created using ***sexual reproduction***: two combinations mate to produce a third, so that this third draws its selection of genetic alleles/morphemes from a blend of that of its parents. GAs model sexual reproduction using a variety of allele selection techniques. *Crossover* draws a random line through two chromosomes, and splices together the complementary parts of each that sit on opposite sides of this line. ***Random mutation*** models the chance influence of the environment by replacing one allele in the new chromosome with one randomly chosen from the same gene/heading. Given two chromosomes to mate, a GA employs a mix of both of these techniques to produce a new chromosome that carries many of the same genes.

But which chromosomes get to mate with each other in this way? Just as the most obviously successful individuals often have their choice of mates in the real world, GAs use their scoring functions to determine the fittest chromosomes. They then allow these combinations to mate with others, with the expectation that they will produce the fittest offspring, as they too will carry the alleles that made their parents successful. An effective mating will often produce offspring that score even more highly than their parents, while a poor mating will tend to produce offspring that score lower than their parents.

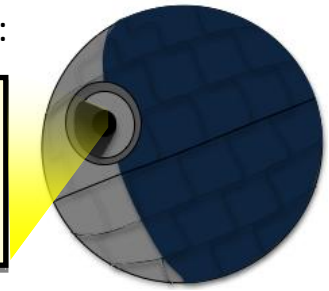


By limiting reproductive choices to the most successful combinations, GA promotes the widespread dissemination of the best alleles and allele pairings, while hindering the spread of poorly-combined alleles. Each new combination is added back into the general population, and to stop this population – and the corresponding search space – from growing too large, an element of ***Malthusian competition*** is added: while the fittest are allowed to reproduce, the least fittest are culled, so that the population steadily grows in average fitness with each new generation. Random mutation is a vital wildcard, introducing quirky genetic changes that come from neither parent, and ensuring that good alleles and allele pairings are never entirely bred out of the population. Of course, as a Darwinian metaphor this is all terribly simplistic, as simplistic as **Herbert Spencer**'s catchy label *the survival of the fittest*. Nonetheless, GAs provide an excellent model of how past successes influence future innovations, showing why the patterns that are most credited as the basis of this success will tend to recur again and again in an endless stream of new contexts and combinations. GAs provide an efficient computational means of choosing high-scoring combinations from a Zwicky box when an exhaustive search is neither feasible nor desirable. They also allow us to model the backward-looking inertia of a creative system by adding representations of past successes to the initial population of combinations.

The morpheme-groupings that persist from one generation to the next are the dominant clichés of the system, though each GA may also give rise to its own new clichés. Neither outcome is necessarily undesirable. The Italian novelist and semiotician **Umberto Eco** has shown that the classic Hollywood film ***Casablanca*** is actually a combination of many cinematic clichés, assembled as one might put together a collage or a quilt. Eco describes the film as an ***inter-textual collage***, noting that characters such as Viktor Laslo are simple assemblages of vignettes from other films. Each time Lazlo appears at Rick's bar, we see him ordering a completely different drink, as if believing himself to be in very different circumstances and films. Eco offers a fascinating critique, but it is one soon forgotten when one views the film itself. Who cares if director **Michael Curtiz** assembled the picture from reusable parts, as a watchmaker might assemble a watch? The approach clearly works in this case, and *Casablanca* remains one of the most watchable Hollywood films of any era, one that is not diminished by any insight into how it was composed.

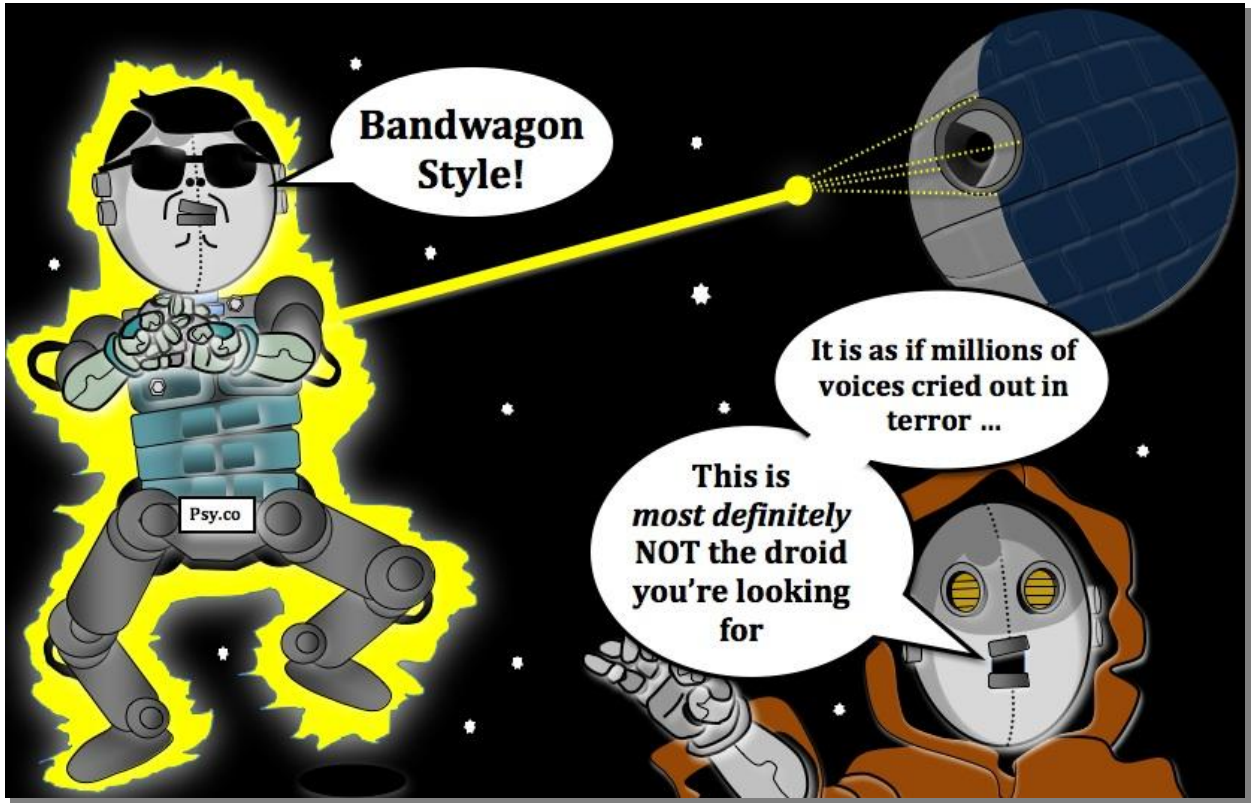
Casablanca is both respected and loved as a movie, which only shows a that a well-used cliché (or a combinatorial collage of clichés) is no hindrance to creativity. Another much-loved film that relies heavily on the artful exploitation of cliché is ***Star Wars***, though this movie is perhaps more loved than respected. The critic **Pauline Kael** so savaged the movie that its director, George Lucas, gave the skull-faced villain of his later fantasy effort, ***Willow,*** the fearful name *General Kael*. Here is what Kael has to say about *Star Wars* in her book *When the Lights Go Down*:

> *"It's an assemblage of spare parts [...] Star Wars may be the only movie in which the first time around the surprises are reassuring. [...] Maybe the only real inspiration of Star Wars was to set its sci-fi galaxy in the pop-art past, and to turn old-movie ineptness into conscious Pop Art."*



The film is indeed a kaleidoscopic hodgepodge of influences, from the 1953 western, ***Shane***, to the 1954 war movie ***The Dam Busters***, to Kurasawa's 1958 samurai classic ***The Hidden Fortress***, not to mention the many aspects of the ***King Arthur*** legends to which it liberally helps itself. Yet it is a marvelously well-executed hodgepodge that exhibits true combinatorial creativity in achieving a joyfully coherent synthesis of these time-tested parts.

It is always easier to be a critic than a creator, to show that a certain narrative is an assemblage of spare parts than it is to find an assemble those *spare* parts for oneself. Creativity will always seem easier *after* the fact, when we are graced with a finished product, than before the fact, when we are faced with an empty page. So what **Pauline Kael** disdains as *reassuring surprises* are more positively viewed by scholars of creativity: psycholinguist **Rachel Giora** refers to them as *optimal innovations*, while the lexicographer **Patrick Hanks** benignly calls them creative *exploitations of norms*.



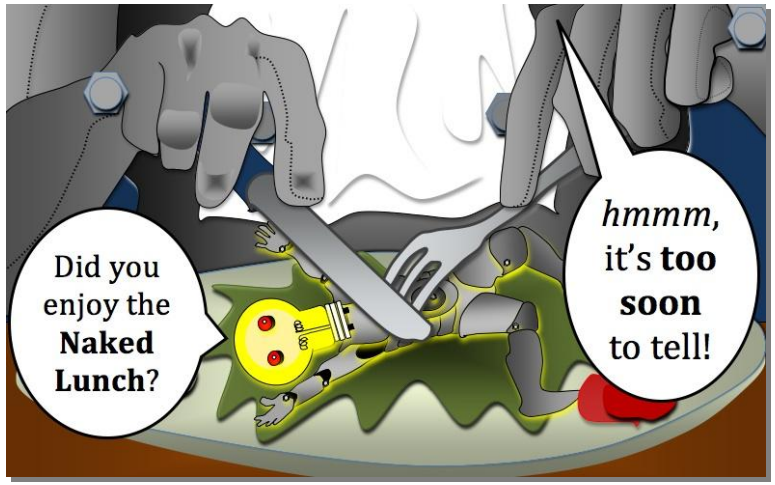
Samuel Goldwyn, co-founder of the **Metro-Goldwyn-Mayer** studio, expressed Hollywood's attitude to innovation with his line "*Let's have some new clichés*". Even in a domain whose lifeblood is convention, and where reassurance is sought from both surprises *and* ticket sales, this lifeblood must constantly be replenished. Few creators set out to invent a cliché, but who wouldn't want their efforts to earn a permanent place in our popular culture? Combinatorial variation can give new meanings to old clichés, or even construct innovative new clichés from these spare parts. For every cliché contains the seed of its own creative subversion. Consider **Woody Allen**'s neurosis that "*I am at **two** with nature*" or his attribution of marital problems to the fact that "*I put my wife **under** a pedestal*."

Recombination is a strategy for innovation that often works well and sometimes even excels. So though cinematic titans like **Orson Welles** had struggled and failed to film a worthy treatment of **Joseph Conrad**'s novel ***Heart of Darkness***, director **Francis Ford Coppola** succeeded with his 1979 movie ***Apocalypse Now***, in large part because he transposed the action from the Belgian Congo to Indochina so as to critique the U.S.'s role in the Vietnam war. The **Coen brothers**' depression-era comedy ***O Brother, Where Art Thou*** is also a loose reworking its classical source, Homer's ***Odyssey***, but the recombination works so well as a comedy that the film won an Oscar for best *adapted* screenplay.

These combinations trumpet their exploitations of earlier source material, but fold in new elements to yield novel and memorable results. Sadly, though these successes show the potential of combinatorial reuse, such efforts are by far the minority. Most Hollywood variations on a theme really are a calculated assemblage of ill-fitting spare parts, from pointless remakes that add nothing to the original, to even more pointless sequels that diminish our fondness for the original. Does 1995's ***Water World*** really improve on ***Mad Max*** by setting its post-apocalyptic world at sea rather than on land? Unlike 1956's ***Forbidden Planet***, which reimagined Shakespeare's ***The Tempest*** in a SciFi setting, 1981's ***Outland*** does a rather poor job of transposing the classic western ***High Noon*** to a mining colony in space. The Tarantino-scripted ***True Romance*** of 1993 adds a little comic streamlining to the superior ***Badlands*** of 1973, but 1990's ***Days of Thunder*** adds nothing at all to 1986's ***Top Gun***, so much so that critics nicknamed it ***Top Car***. But even these exercises in cynicism do not make a disavowal of combination and cliché any more sensible. Most stories rely on clichés of one form or another, from the *boy-meets-girl* structure of romance films to the journeys of self-discovery and renewal that **Joseph Campbell** argues is the basic structure of most hero-centric stories. However a filmmaker uses clichés, they will inevitably shape an audience's expectations. When expectations are lazily abused, they serve as little more than templates for more-of-the-same variations, familiar nests in which *clockwork cuckoos* lay their sadly predictable eggs. Yet they can also serve as the fulcrum from which far more creative – and perhaps even subversive – combinations can be leveraged.

When Allen Ginsberg read *naked lust* as *naked lunch*

in a manuscript by **William Burroughs**, his accidental misreading suggested a resonant title for Burrough's famously surreal novel. When Chinese premier **Zhou Enlai** was asked about the effects of the French revolution by **Richard Nixon** in 1971, he replied that it was "*too soon to tell*." Zhou's retort loses its witty lustre, however, when we learn that he had misunderstood Nixon's question, and assumed him to be referring to the French student riots of **1968**! **Chas Freeman**, Nixon's interpreter at the meeting, describes Zhou's misunderstanding as simply "*too delicious to invite correction*".



Some accidents are *happy* accidents with deliciously creative consequences. Though intentionality is crucial to creativity, creative intent can arise before or after the generative act, or indeed before *and* after, as is often the case: a creator may set out to achieve one thing and achieve quite another entirely. A leaky tap is thus not creative, no matter how rhythmic its drips happen to be. Yet the recognition of the tap's accidental musicality may lead to subsequent creativity on the part of an opportunistic listener, much as **William Burroughs**' recognition of the creative potential in **Allen Ginsberg**'s misreading of "*naked lunch*" presented him with an opportunity for linguistic creativity. One might argue that Ginsberg's poetic worldview was at least partly responsible for the misreading, and so he is due *some* credit for the phrasing. Certainly, the trained mind is better prepared to recognize the creative potential in chance events. Nonetheless, it was only when Burroughs perceived a literary utility in this novel apposition, and creatively interpreted it to mean a moment of undisguised clarity – that "*frozen moment when everyone sees what it is at the end of every fork*" – that it became the product of intentional creativity. Likewise, Zhou's retort is creative only to the extent that listeners imbue it with shades of wit and profundity. Though we are often happy to ascribe "creativity" to the workings of *Mother Nature*, we are generally loathe to ascribe creativity to *merely generative* mechanisms that unintentionally produce outputs whose creative potential must be recognized and shaped by others.

Creative intent can elevate the banal into the sublime. Consider the rather commonplace act of stirring sugar into a glass of milk to sweeten the drink: though what results is a culinary blend of two ingredients, it is a blend that lacks emergent qualities and which few would consider interesting or original. Yet consider this charming story of the **Parsis** and how they came to settle in Western India. The ***Time Out*** guide to Mumbai tells their tale rather nicely:

> "They arrived in Gujarat in the eight or ninth century and sought asylum from the local king. He is said to have sent them away with a glass of milk full to the brim – his way of saying that his kingdom was full. The Parsi elders conferred, added some sugar to the milk and sent it back – to suggest that they would mix thoroughly and sweeten the life of the community."

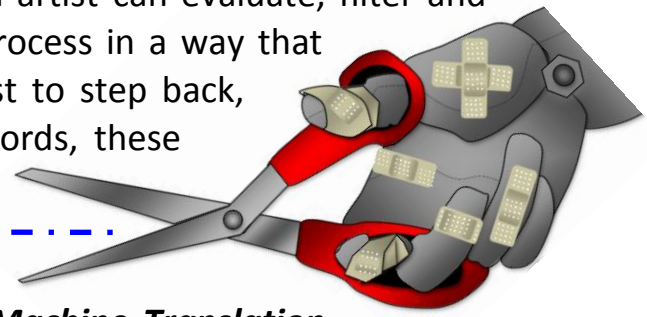
The Parsi elders were wise enough to recognize an opportunity for creativity when they saw one. Rather than view the milk as just *milk*, they perceived the king's creative intent and responded in kind. Their response was a clever but not unduly hostile act of ***oneupmanship***, in which the elders fashioned a potent rejoinder from the symbolic materials they were given (much as **Euan Booth** fashioned an effective protest against Tony Blair using only the resources at *his* disposal). Creativity can be catching, and creativity by others often inspires us to be creative in response.

The king flattered the Parsis by using creativity to send an all-too-familiar message in an original form, for he expected them to see the true meaning of the milk and take it all the more seriously for the unpacking it required. And the Parsis flattered the king in turn, by assuming he too would appreciate the intent behind their simple blend. Creativity often flatters its audience, making us work to appreciate the non-obvious meaning of a familiar action. This effort pulls us into an creator's work, making us partly responsible for its meaning and partly complicit in its success.

How can we systematize the production of *happy accidents*, and so generate many more chance events onto which we might stamp our creative intent? Ironically, the risk-averse reuse mechanisms most favored by Hollywood studios bear a rather striking similarity to the randomized, recombinant and very avante-garde techniques pioneered by the surrealist school of experimental art. The surrealists sought more than a conscious disavowal of cliché – they sought a radical means of escaping the deep-hewn furrows of **System 1** that unconsciously inhibit one's spontaneity and creativity. As writer **William Burroughs** put it, one "cannot will spontaneity" into being, but one can "*introduce the unpredictable spontaneous factor with a pair of scissors*". The scissors here alludes to the *cut-up* technique pioneered by Burroughs and artist/writer **Brion Gysin**, in which a linear text is sliced, diced and randomly re-spliced to form new texts that may give rise to new and surprising interpretations.

The purpose of the cut-up technique is actually two-fold: not only does it aim to create new combinations from old, much like the Hollywood system or "machine", it consciously aims to disrupt the mind's attempts to automatically group commonly co-occurring words and ideas into familiar gestalts. But unlike the Hollywood approach – and there surely was a large *but* lurking here – the technique embraces uncertainty and incongruity, and aims to challenge rather than to comfort, to unsettle rather than to placate. As a result, the technique hinders rather than facilitates the perpetuation of clichés, though it can lead to the creation of new idioms and new clichés of its own, of a kind keeping with the spirit if not the letter of **Samuel Goldwyn**'s plea to "*let's have some new clichés*".

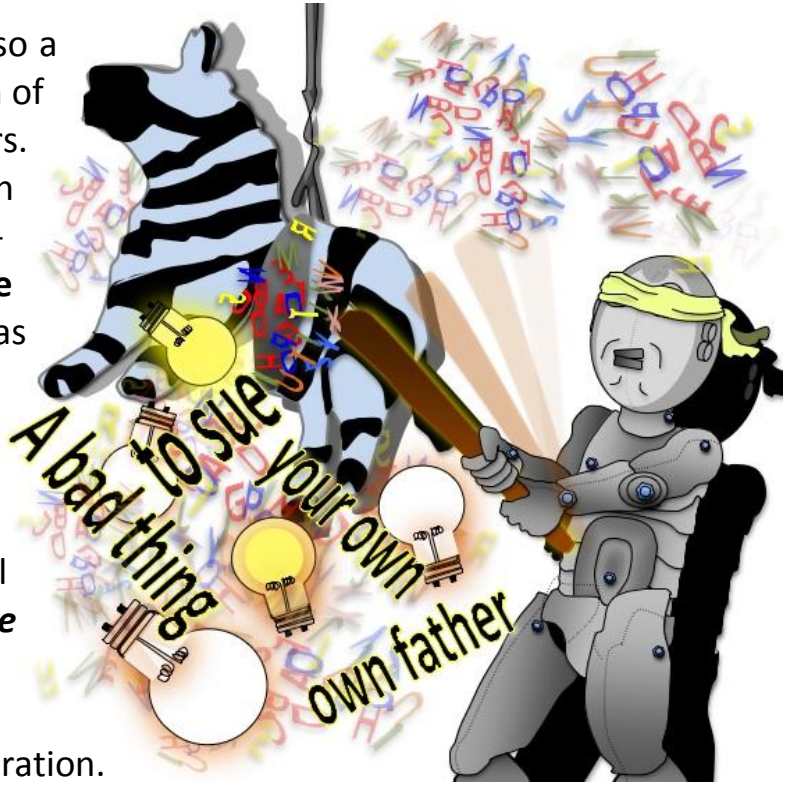
The cut-up technique was originally inspired by the *collage movement* in visual art, in which fragments of images and texts are selectively combined to form a novel patchwork whole. When the cut-up technique is applied to a linear text, the text is segmented into short strands of contiguous words that do not necessarily respect either phrase or sentence boundaries. These strands are then randomly recombined, to form a new text that uses the same words in different linear juxtapositions, to facilitate – if one charitably overlooks the inevitable bad grammar and illogical punctuation – very different global interpretations. Gysin originally applied the technique to layers of newsprint, which he sliced into linguistic chunks with a razor, and Burroughs later extended the technique to audio tapes. In principle, any linear source of information, from text to audio to video and even DNA, can be sliced and re-spliced using the cut-up technique to deliberately subvert familiar patterns and spontaneously suggest new meanings. Yet the cut-up technique does not actually *create* new meanings, and is ***merely generative*** in the fullest sense of this phrase. Rather, the goal of the cut-up technique – and of related techniques developed by Burroughs, such as the *fold-in* and the *drop-in* – is to generate candidates for interpretation that an artist can evaluate, filter and rank according to their creative potential. By automatizing the production process in a way that frees it from the baleful effects of cliché, such techniques also free an artist to step back, observe, and focus on the act of interpretation and evaluation. In other words, these techniques are the arts-and-crafts equivalent of a *generative algorithm*.



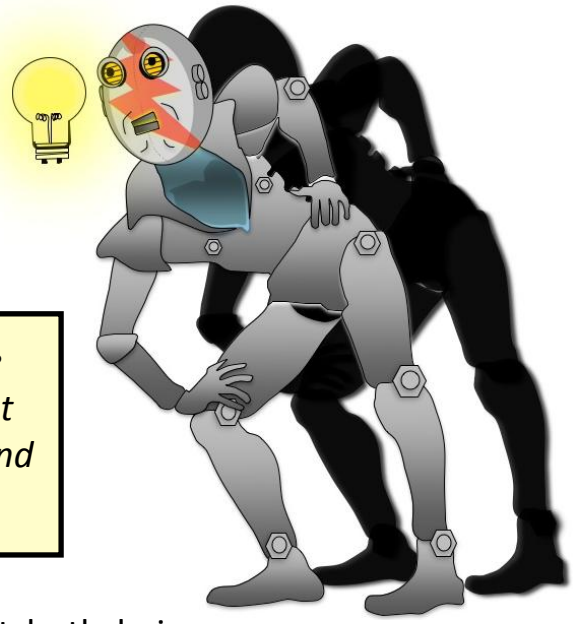
The cut-up technique has other intriguing resonances with automatic approaches to language processing using computers. In an ***Example-Based Machine Translation*** (EBMT), a target-language translation of a text is stitched together from a combination of pre-translated linguistic chunks. While these chunks are internally coherent, their combination often gives rise to cross-chunk *boundary friction*, and EBMT systems succeed or fail on their ability to iron out these inconsistencies and produce a fluent text. In contrast, the creation of boundary friction is the *raison d'être* of the cut-up technique, whose explicit goal is the generation of texts whose challenging juxtapositions jolt both **System 1** *and* **System 2** into action. It is rare that the cut-up technique generates texts that read fluently in their own right, though sense may still be imposed onto them.

EBMT and related approaches to text creation are most successfully applied in domains where there is a logical progression from present texts to future texts, as in the evolving documentation of a maturing software product. Since present texts are known to the system and exist in multiple translations, an EBMT system can carve out a collection of reusable translation chunks by aligning parallel translations of the same text. Though future texts are unknown and unseen, their subject matter is easy to predict, so in an important sense the present allows the future to be predicted. The general principle at work here – which sees the future as a functional re-arrangement of the present – also applies to artistic cut-ups. By cutting and re-arranging a text on a given subject, one is likely to produce another text on a similar subject, albeit one with some provocative juxtapositions.

The future is often a functional rearrangement of the past, and so a cut-up will sometimes yield a text that is not just a logical progression of the original, but a viable prediction of some future state of affairs. Burroughs describes one such experiment with newspaper cut-ups, in which a text about oil billionaire **John Paul Getty** was randomly re-arranged to produce the enigmatic sentence "**It's a bad thing to sue your own father**". A year later, claims Burroughs, the billionaire was indeed sued by one of his own sons. This is a hand-picked example, one that shows the necessity of carefully filtering the generative products of the cut-up technique, yet it also shows the capacity of cut-ups to bring into focus the latent possibilities that implicitly reside in a text. Burroughs described the link between combinatorial generation and prediction thusly: "***when you cut into the present, the future leaks out***." In truth, cut-ups are no more predictive of the future than fortune cookies or Tarot cards, but when creativity is our goal, spontaneity and unpredictability can be the better parts of generation.



Musician David Bowie is perhaps the most famous modern exponent of the cut-up. As a very modern inheritor of this old surrealist technique, Bowie exploits computing technology to replace the scissors, glue and newsprint of the original. Using a piece of software called the **Verbasizer** – written for Bowie by **Ty Roberts**, co-founder of the online music service ***Gracenote*** – the musician can randomly cut-up and fold-in words and phrases from up to twenty different texts at a time. Bowie describes the workings of his **Verbasizer** as follows:



> *"And it'll take those twenty sentences and cut in-between them all, all the time picking out, choosing different words from different columns. And from different rows of sentences. So what you end up with is a real kaleidoscope of meanings and topics and nouns and verbs all sort of slamming into each other."*

Bowie uses his **Verbasizer** to generate unusual word combinations that suggest both lyric fragments and themes for new songs. But like the scissors and newspapers of **Burroughs** and **Gysin**, the software is a *merely generative* producer of provocative stimuli, rather than a creator of new ideas in its own right. It still falls to Bowie to sort amongst the happy accidents and impose value on some rather than on others. He gives the example of the four-word cut-up "*the top kills himself*", and notes how "top" might metaphorically denote a "boss". From there, he notes that the image of a suicidal boss brings to (his) mind the great Wall-street crash of 1929, when despondent CEOs hurled themselves from the tops of tall buildings. This might, he adds, provide a fascinating topic for a new song.

But is there any theoretical reason why the **Verbasizer** and apps like it must limit themselves to mere generation? Given advances in **Natural Language Processing** (**NLP**), such software can certainly filter its own outputs, to highlight well-formed phrases that are emotion-laden, superficially incongruous yet meaningful, and statistically improbable.

Play Time

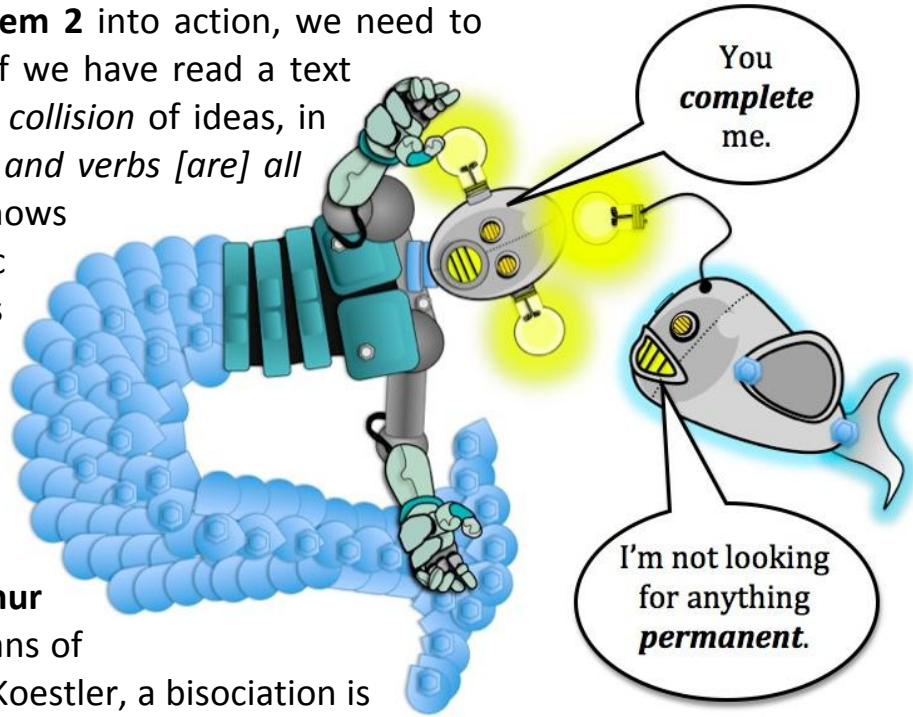
The cut-up technique and other surrealist approaches to *chance generation* were inspired by games that artists would play in cafés. Decades later, these techniques have not lost their playful quality, and the schoolyard game **MadLibs** (in which players provide a list of words that are then placed into an unseen template to produce a funny text) owes a great deal to the surrealist technique (and café pastime) ***Exquisite Corpse***.

Let's turn the **Cut-Up Technique** into a **Board Game** that can be played at parties. Cut-up fragments (from a given domain, such as *movie titles, song lyrics, newspaper headlines* or *personal column adverts*) can be printed on playing cards, so that players can be dealt an initial hand of random cut-ups from the deck. Under certain circumstances, a player can discard a card, take a new one, or trade cards with another player. To heighten the element of chance, a dice is also used, and an egg-timer is employed to impose time constraints on player actions.

- Sketch some rules for the Cut-Up Game to allow players to take new cards, discard old ones, or trade cards.
- We characterized the game as a board-game. What specific role might the board play in this game?
- How might the dice be used to increase the elements of chance *and* of strategy in this game?
- Suppose the cards are *color-coded* to indicate the general category of a cut-up fragment. What categories and color-codes would you use to ensure a playable and enjoyable game that combines strategy *and* chance?
- Suppose we implement a computer-based version of this game. How might we implement an intelligent computer player that can play a good game against human players? Could humans tell a computer player from another human player?

  Extra credit: How would you allow a player to trump and claim another's cut-up by adding cards of their own.

Sparks rarely fly when highly-associated words and ideas are placed in gentle juxtapositions.



To jump the rails of **System 1**'s expectations and force **System 2** into action, we need to juxtapose words in improbable ways that make us wonder if we have read a text correctly. For a surprising juxtaposition of words produces a *collision* of ideas, in which – to quote **Bowie** – "*meanings and topics and nouns and verbs [are] all sort of slamming into each other.*" A creative juxtaposition shows us that words and ideas that are strongly linked to a specific domain can have unexpected resonances in other domains too. Consider the surrealist game ***exquisite corpse***, whose name is itself a product of the cut-up technique. This word-slam spurs us to view ***corpses*** in a wider and more pleasant context than **death**, and to associate the word ***exquisite*** with unpleasant ideas whose beauty is not immediately apparent. The phrase "*exquisite corpse*" is thus an instance of what **Arthur Koestler** calls a ***bisociation***, while the game it names is a means of producing many more of these *bisociative* combinations. For Koestler, a bisociation is any combination of words or ideas that spurs us to see these elements in more than one associative context at once.

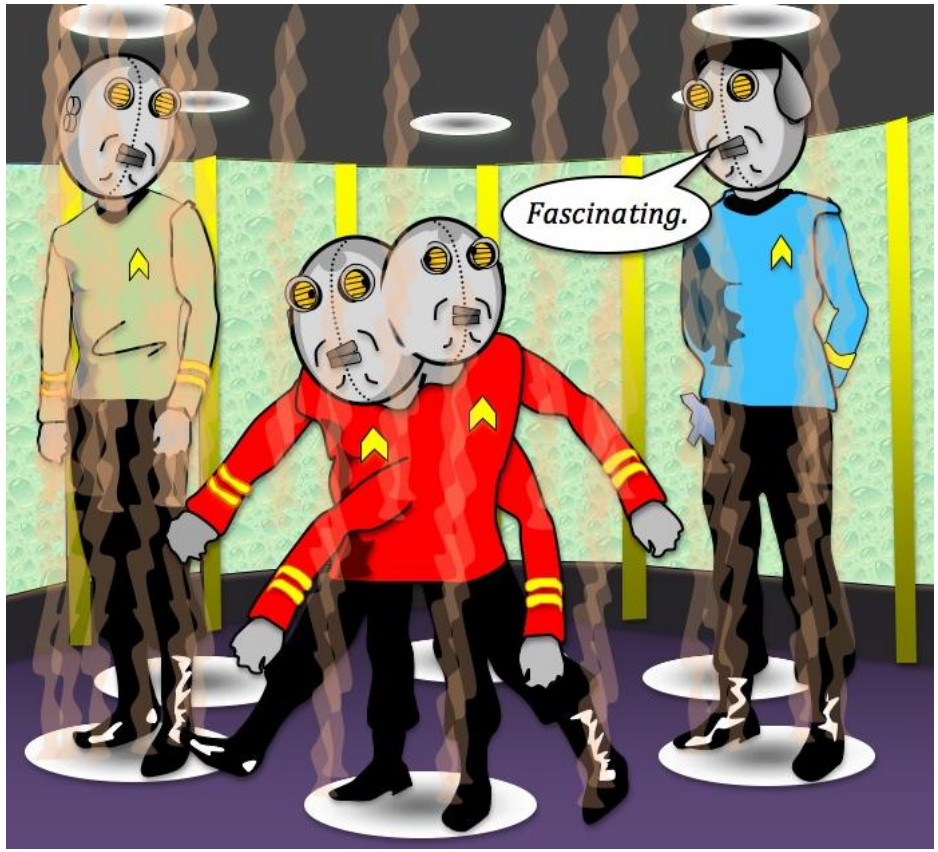
Music fans who listen to **Pink Floyd**'s album ***The Dark Side of the Moon*** while watching the MGM movie ***The Wizard of Oz*** (with the audio turned down) experience a remarkable alignment of sounds to images that is entirely coincidental yet rich in opportunities for bisociative sense-making. Dramatic changes in tone or setting in the film are paired with dramatic changes in the music, while some tracks on the album – such as ***Brain Damage*** – coincide with remarkably apt scenes in the movie (such as the *Scarecrow's dance*). The combined experience imbues tracks from the album with visual associations from the movie, and scenes from the movie with dark overtones from the album. Fans of this bisociative experience give it a name that is itself an apt bisociative cut-up: ***The Dark Side of the Rainbow***.

It is *bisociation*, Koestler argues, that allows artists to weave metaphors between domains, that allows scientists to parlay their intuitions about everyday events (riding in an elevator, standing on a moving ship) into analogical models of the cosmos, and allows comics to erase the fine line between the sublime and the ridiculous. Yet Koestler saw bisociation as much more than a clever use of words and symbols with dual meanings, and argued that bisociation is, in fact, a very powerful cognitive tool for general problem-solving. Consider the *riddle of the monk*, a classic creativity test-stimulus first used by the psychologist **Karl Duncker** and later expressed by Koestler as follows:

> One morning, exactly at sunrise, a Buddhist monk began to climb a tall mountain. The narrow path, no more than a foot or two wide, spiraled around the mountain to a glittering temple at the summit. The monk ascended the path at varying rates of speed, stopping many times along the way to rest and to eat the dried fruit he carried with him. He reached the temple shortly before sunset. After several days of fasting and meditation, he began his journey back along the same path, starting at sunrise and again walking at variable speeds with many pauses along the way. His average speed descending was, of course, greater than his average climbing speed. **Prove** that there is a single spot along the path the monk will occupy on both trips at precisely the same time of day.

Readers of a mathematical bent may be tempted to seek an algebraic solution to Koestler's riddle, perhaps by expressing each of the monk's journeys as simultaneous equations for which a single intersecting point can be shown to exist. **Bisociation** yields a much more intuitive solution: just imagine that our monk makes both journeys ***on the same day***, so that as he walks ***up*** the mountain his bisociated alter-ego from the future walks ***down*** the mountain. Each version of our monk must follow the same narrow path on the same day, so no matter how erratic their individual speeds, each is inescapably destined to bump into the other **once** at ***some*** point on the path.

The most creative *bisociations* do more than juxtapose: they integrate conflicting viewpoints into a coherent whole, often achieving elegance and concision by tightly compressing words and ideas together. So Koestler's bisociative solution to the monk's riddle is at once simple and impossible, easy to imagine yet physically unworkable. This integration of perspectives into a unified *blend* is commonplace in creativity, even if it rarely draws as much attention to itself as it does in **Duncker**'s riddle.



Consider the logo for **The Four Seasons**, a chain of luxury hotels. This logo uses the image of a tree to signify each season: a tree with sprouting leaves signifies **Spring**; one with full foliage signifies **Summer**; falling leaves signify **Autumn**; and a tree with bare branches signifies **Winter**. Yet the chain's simple logo contains just one tree, not four, so that each season is expressed via a different part of the *same tree*. Such a tree would be an unlikely find in the real world, but those who see the **Four Seasons** logo rarely see it as anything special. We instinctively appreciate these elegant blends as products of the imagination rather than of the real world, taking what inferences we can from their seamless join of multiple contexts and perspectives.

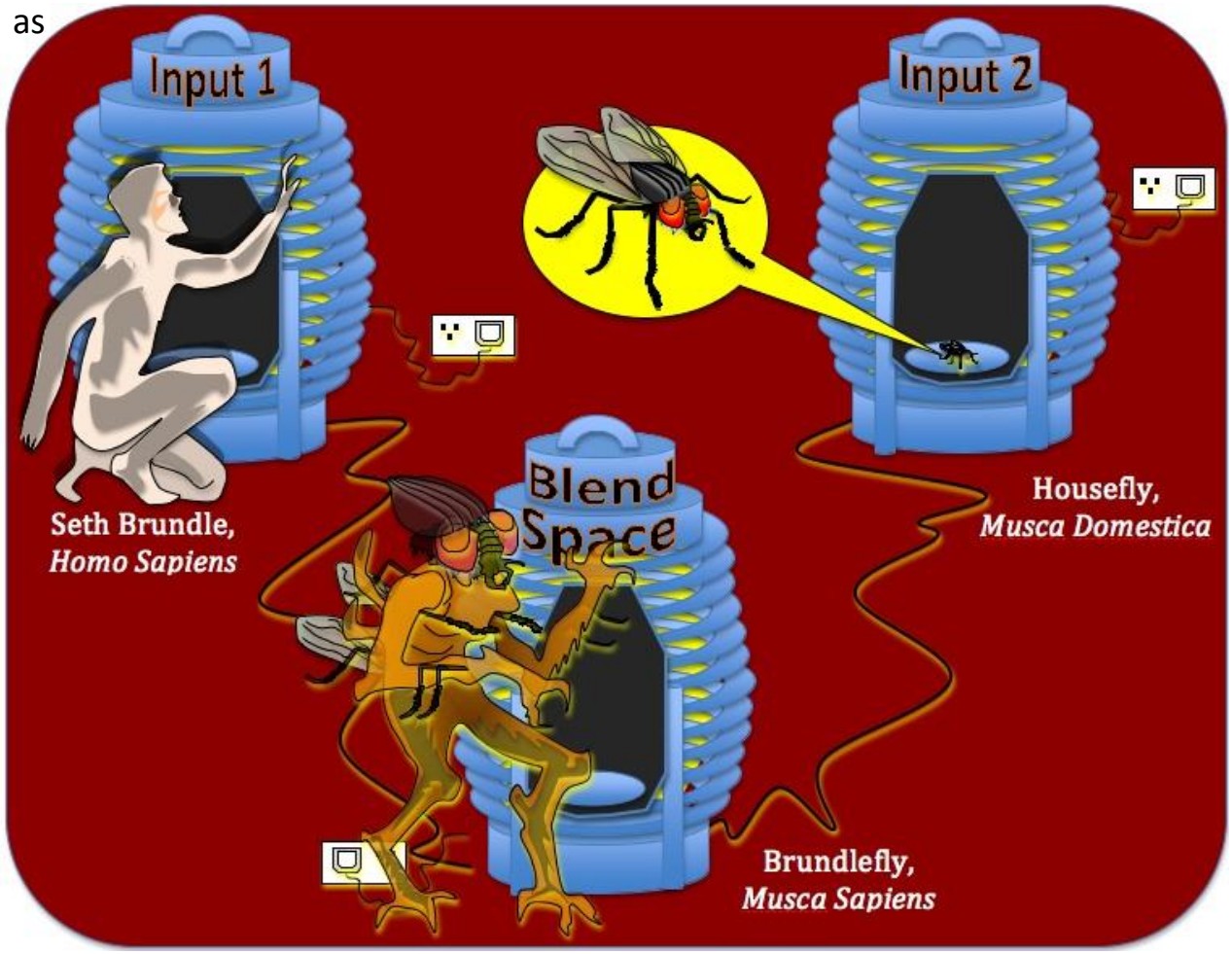
So the **Four Seasons** does to trees what Koestler's riddle does to Buddhist journeys and what the **USS Enterprise**'s transporter sometimes does to unfortunate red-shirts: it compresses them into a single improbable but imaginative – and inferentially useful – whole. The cognitive linguists **Mark Turner** and **Gilles Fauconnier** give the name ***Conceptual Blends***, or ***Conceptual Integration Networks***, to the products of this remarkably pervasive mechanism of creativity.

In David Cronenberg's gory 1986 remake of *The Fly*, Seth Brundle is an eccentric scientist whose intrinsic motivation to be **H-Creative** drives him to take foolhardy risks. For Brundle is building a 20th-Century version of the transporter from ***Star Trek***: a person or thing sitting in a source pod is disassembled by lasers into a stream of particles that is then transmitted to an identical destination pod, whereupon these particles are faithfully reassembled into their original form. But this being a horror movie, all does not go to plan. Unbeknownst to Brundle, a fly enters the source pod as he tests his transporter on himself. When both he and the fly are digitized and beamed to the destination pod, a confused computer mixes all of their particles together and reassembles them into a single unitary being. Brundle is unaware of the looming tragedy, believing his newly increased strength and agility to be a purifying side-effect of the act of being taken apart and reassembled again. But a horrifying bisociation is lurking in his blended DNA, and Brundle slowly transforms into a creature that is part human and part fly. As much-loved appendages drop off, his skin hardens and flakes, and he is forced to eat by first regurgitating acid onto his food, Brundle finds his newly-developed super-strength to be cold comfort indeed. Much gnashing of teeth and chewing of scenery ensues before a grim end.

Though far from intentional, Brundle's fate offers an excellent illustration of the power of creative blending. The bisociated being, who ruefully names itself ***Brundlefly***, retains a selective mixture of its human and fly heritages. Thus, it retains the skeletal structure of a human combined with the skin, hair, hands and eyes of a fly. But it also exhibits ***emergent qualities*** that are unique to the blend: Brundlefly's superhuman strength derives neither from his human or fly heritage, but from the combination of both: Brundlefly has the proportional strength of a fly *the size of a human*!

Fauconnier & Turner define *blending* as a cognitive operation that applies to *mental spaces*:

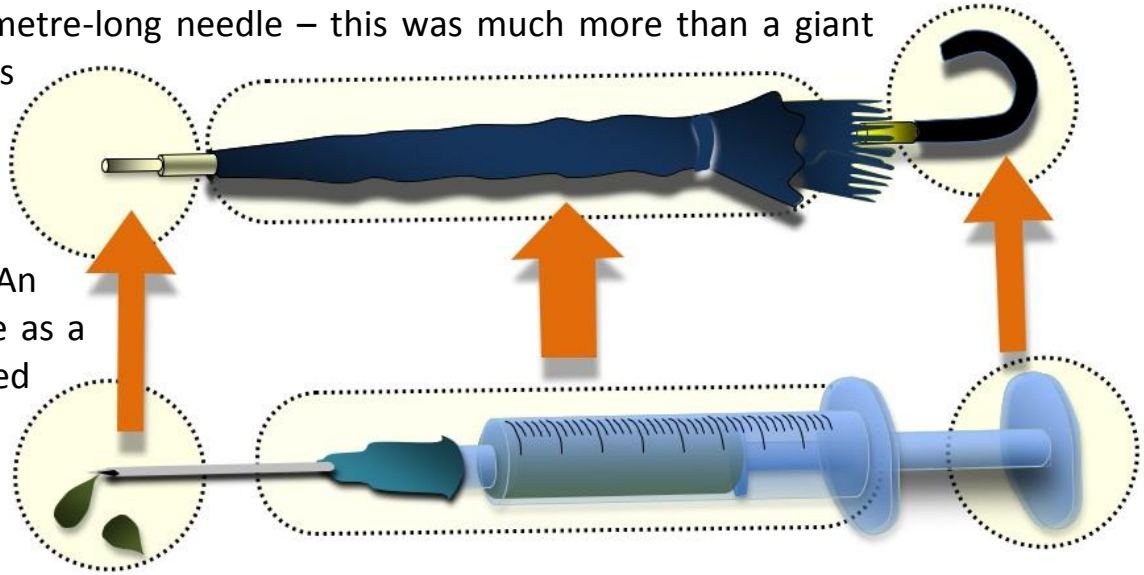
two or more **input spaces** are conceptually integrated to yield a new mental space called the **blend space**. A **mental space** is a pragmatic grouping of beliefs and generalizations that one associates with a particular topic, domain or perspective. Some mental spaces are pre-packed bags of the stereotypical norms that are automatically primed for a given topic, while others are prepared as needed, like suitcases of knowledge that must be carefully packed for each new voyage. Each of the monk's journeys up and down the mountain is thus carefully assigned to its own mental space, as are the competing interpretations of a joke. But when we appreciate the **Brundlefly** blend, its two input spaces are more-or-less automatically populated with our folk knowledge of ***human*** and ***fly*** biology.



Each **blend space** is tied to its inputs by a range of constraints and principles. A *blend* is thus a **gestalt** that emerges from a whole network of interlinked spaces rather than from a single space. Our ability to ***unpack*** a blend to reveal these linkages adds to our appreciation of its creative design.

Georgi Ivanov Markov* was a Bulgarian dissident and writer who used sarcasm and irony to** criticize the Soviet/Bulgarian system from which he defected in 1969. But Markov is less famous for his creative turn of phrase than he is for the mode of his death. On a London street in 1978, Markov was poisoned using a ***tricked-out umbrella, wielded by an assassin that many believe to be an operative of the **KGB**-aided Bulgarian state police. Though the case remains unsolved after more than three decades, it does provide an illuminating example of the workings of conceptual blending, throwing particular light on the ***selective projection*** of input elements into a blend.

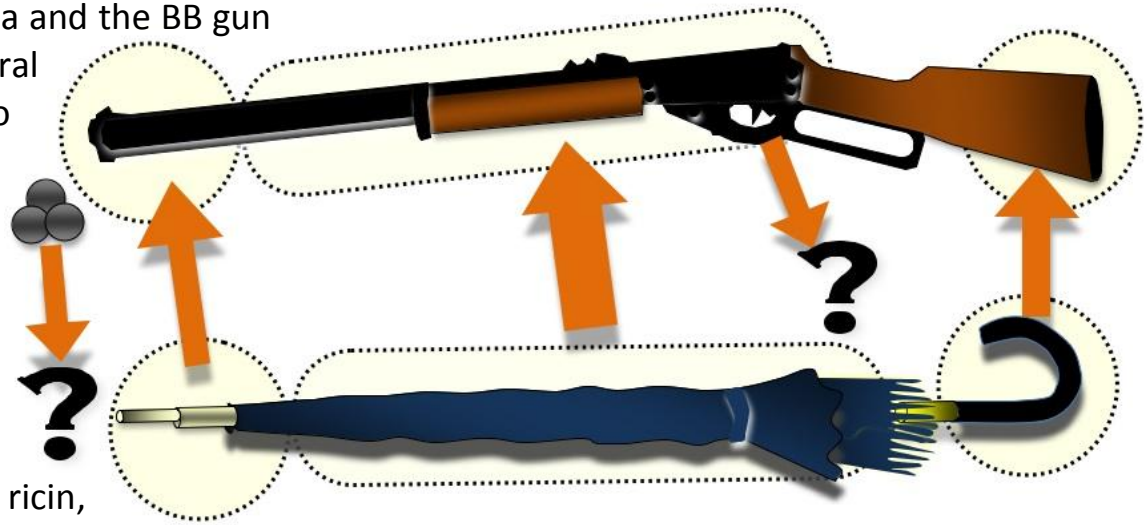
Fans of the **James Bond** series of movies will recognize the creative handiwork of a **Q**-like inventor here. As Bond's armorer, **Q** employs an ingenious mix of **IQ** and **GQ**, for he is responsible for arming **007** with an escalating array of lethal gadgets, each cleverly integrated into a harmless-looking fashion accessory for the gentleman-spy about town. And what could be more unthreatening than an umbrella on a London street? Though disguise is important – even the most strung-out junky would run a mile from a metre-long needle – this was much more than a giant poison-filled syringe *disguised* as an umbrella. This blend was a true ***conceptual integration*** of both items together, yielding a single unified whole.



So let us speculate on the devious intentions of the **KGB Q** (let's call him **Q'**) behind this blend. An umbrella, when closed, has much the same shape as a syringe, for each contains a handle that is connected by a long, hollow central shaft to a pointed tip. As it is neither rare nor troubling to be prodded with an umbrella when jostled on a wet London street, these three shared elements provide the perfect mechanism, and the perfect cover, for a close-range poison-delivery system. **Q'** should thus selectively project all three into the blend, where they can serve as a support-structure for the **syringe**-specific delivery elements to follow.

For even an experienced assassin can only do so much with an umbrella, and so other elements must also be projected into the blend space. Though one can fill an umbrella's central shaft with liquid poison, **Q'** still needs to selectively project a plunger from the **syringe space** into the blend. Moreover, **Q'** also needs to project the tiny hole at the tip of the syringe's needle, to sit – in an analogous position in the blend space – at the umbrella's tip.

Though a **syringe** provides an ideal delivery-system for poison, and is the tool most likely to be primed by our **System 1** processes, Markov's post-mortem examination revealed that he had been injected with a **pellet of *ricin***, a solid poison that – in an other example of the **unusual uses test** in action – is extracted from castor beans. To deploy the poison in a solid form, **Q'** needs to blend his umbrella not with a syringe but with a **gun**. Since most modern guns fire cartridges that house both a projectile bullet and an explosive means of propulsion such as gunpowder, **Q'** instead models his ***umbrella gun*** on a child's **BB gun**, which uses compressed air to shoot simple metal pellets down a barrel at a target. In his re-imagined blend, the umbrella and the BB gun each contain a handle and a long, hollow central shaft as before, so these are again projected into the blend space. As with the syringe's pinhole tip, the hole at the end of the gun's barrel must also be projected into the blend, since this is a functional necessity of any projectile launcher. The gun's trigger must likewise be projected, as must its hidden cartridge of compressed air. Finally, as **Q**' simulates his blend and mentally loads his **umbrella gun** with its lethal payload of ricin, he spies a potential problem: since an umbrella is not carried as one carries a rifle, but carried with its shaft pointing *downwards*, its pellets will likely fall out as it is carried. So **Q'** adds one final, emergent touch to his blend: a valve that is only opened when the assassin presses the trigger (hidden in the umbrella's handle) to release the compressed air.



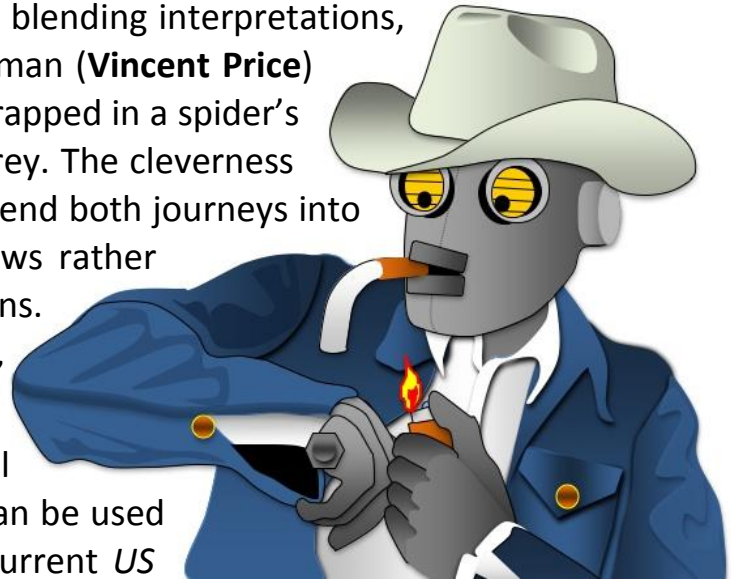
Hold The Presses

Tabloid newspapers employ a distinctive writing style, especially in their headlines. For to convey maximal information with maximal concision, headlines often omit function words such as determiners and coordinators, and favor short punchy verbs like *stun*, *slam*, *lash* and *bash*. They use a short-hand all their own: crowds become mobs, liars of any kind become ***rats***, while scientists and other experts become ***boffins*** or ***nerds*** and athletes become ***aces***. Past events are told in the present tense ("*Nun Flees Burning Jail*") while future events are conveyed with infinitival forms ("*Boffins to Make Death Obsolete*"). The results are typically eye-catching but often unintentionally hilarious, as when concision leads to unwanted ambiguity in "*Noam Chomsky: The Gravest Threat to World Peace*". These unfortunate headline ambiguities are called ***crash blossoms***.

- Cut up some real tabloid newspaper headlines, being sure to get good coverage of the topics tabloids love most: the royal family, celebrities, sex scandals, illegal immigrants, aliens, official cover-ups and sensational crimes.
- Mix in some invented headlines of your own, and cut these up too. Be principled about where you cut, to preserve named entities, phrase boundaries, etc.
- Use the cut-up technique to generate random tabloid headlines of your own. Are they as convincing as real headlines? How humorous are your cut-ups? How many crash blossoms do you generate?
- Mark (or color-code) your headline fragments in ways that reflect the main concerns of tabloid newspapers (celebrities, politicians, royals, aliens and the supernatural, sex, scandals, new technologies and old sins, etc.)
- When you select fragments for your cut-ups according to their color-code or markings, they are more likely to produce well-formed headlines. But do you produce more interesting or more believable cut-ups?

Hint: Aim for complex cut-ups like "*My night of passion with Big Foot; Prince Charles tells all*"

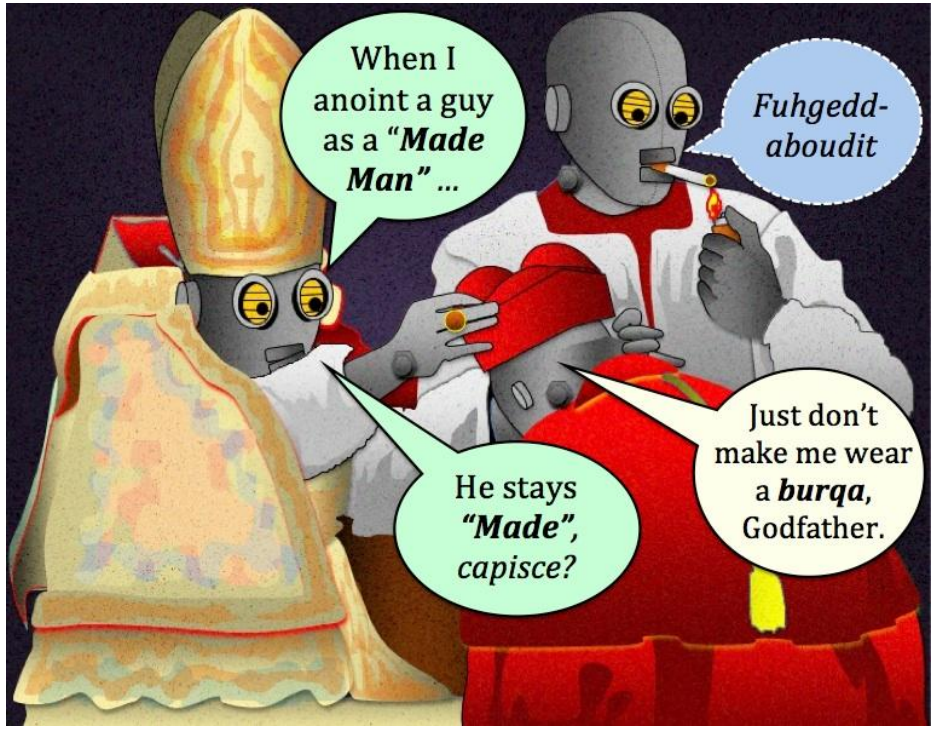
So the creativity of a blend depends just as much on our choice of input spaces as it does on the specific mix of elements that we selectively project from each input into the blend. The creativity of the **umbrella gun** thus principally emerges from the KGB's choice of a harmless umbrella and a lethal gun as inputs, while that of the **Brundlefly** blend arises from the filmmakers' combination of **human** and **fly** biology to achieve a queasy ***body horror*** aesthetic. Indeed, a single high-level decision can support diverse blending interpretations, and the **1958** original of **The Fly** gives us *two* complementary blends: a man (**Vincent Price**) with a giant fly head, and a tiny human head on a fly's body. The latter, trapped in a spider's web, chillingly squeals "*Help Me! Help Me!*" as a spider looms over its prey. The cleverness of Koestler's solution to the monk riddle likewise lies in his decision to blend both journeys into one. As in all the best thought experiments, the required insight follows rather quickly when the inputs are framed so as to engage our everyday intuitions. But if much of the creativity occurs *prior* to the blending process itself, how can we truly use blending as a generative mechanism for creativity?



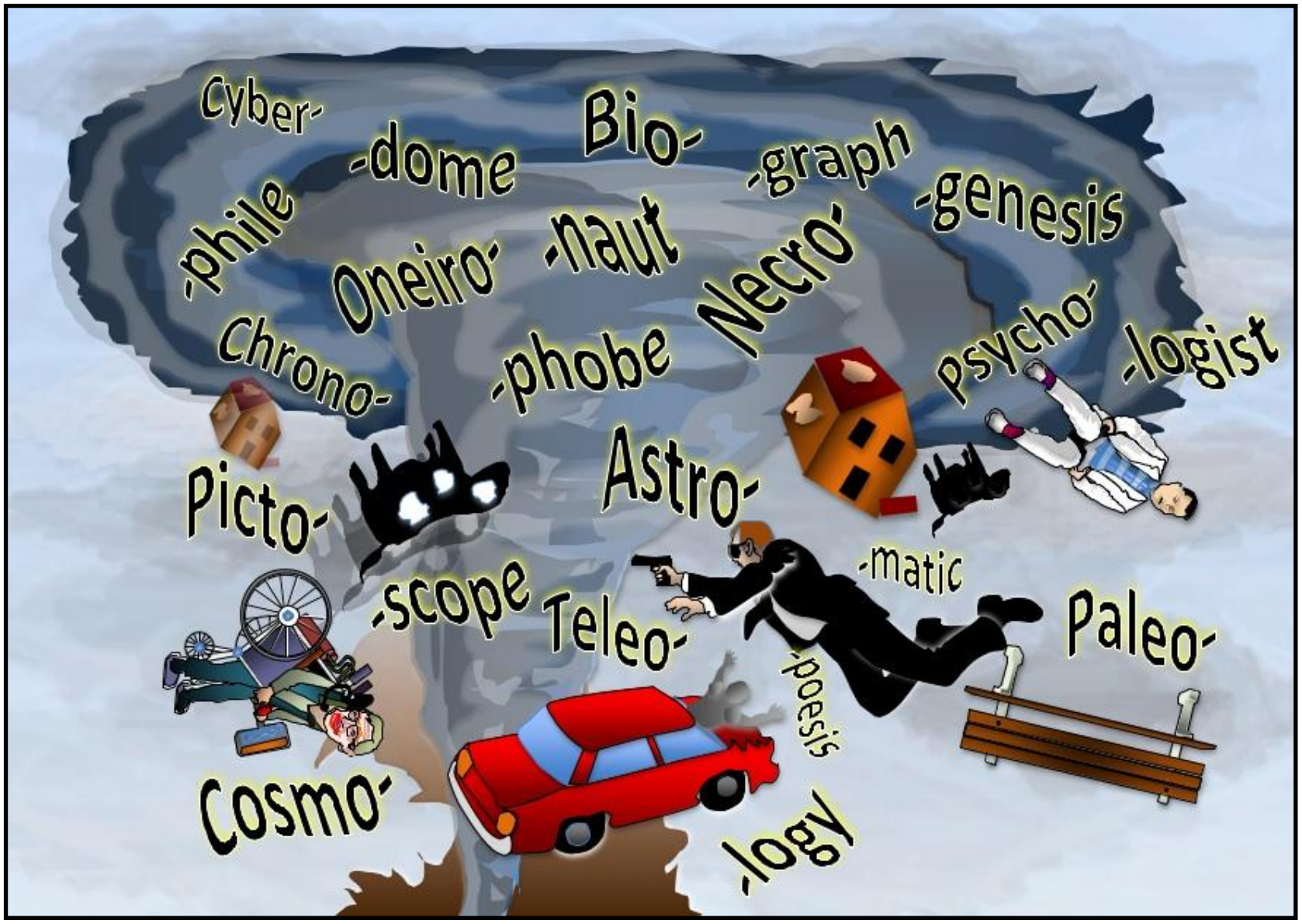
The key lies in a property of blends that **Fauconnier and Turner** call ***metonymic tightening***. When **A** is so strongly associated with **B** that **A** can be used as a signifier for **B** – as when the *Whitehouse* is used to refer to the current *US Administration*, or when *Brussels* is used to stand for the *European Union* – then **A** is said to be a *metonym* for **B**. In achieving a tight compression of inputs, blends frequently turn strong metonymic associations into stronger identity relationships. Thus, we visualize the grim reaper as a skeleton because skeletons are metonyms for death (as **Ali G** once asked a former US surgeon general, "*Skeletons – is they always evil*?"). Tobacco companies exploit a sly sexual metonymy by showing rugged, free-spirited cowboys enjoying their products. But a creative anti-smoking campaign subverts this imagery by suggesting that smoking also causes impotence. The counter-advert, which shows a cowboy with a drooping cigarette, coyly elevates the link between smoking and impotence into an identity statement: in this blend of smoking as sexual dysfunction, to smoke is to *be* impotent, whether one is a rugged individualist or not.

Just as it takes two to tango, it takes at least two input spaces to construct a creative blend.

When presented with just a single input to a conceptual tango, we may yet find an apt partner to complete our blend by looking for those metonymic associations that can be meaningfully tightened into bold statements of identity. Thus, since ***Death*** robs bodies of their juicy flesh to leave only dry *bones* behind, a *skull-faced boney wraith* offers a vivid blend for this grim process. Likewise, since ***Burqas*** *shield* their occupants from prying eyes, a *shield* or a *suit of armor* or even *an armored car* (at a further metonymic remove) may each serve as apt inputs to a daring blend. As cars are associated with horses through the metaphor of *horsepower*, this metonymy can be leveraged to yield associations with cowboys and robust outdoors lifestyles. Images of horses and cowboys are thus frequently used in the marketing of cars, especially of those that stake a claim to ruggedness, virility or environmental soundness. Examples include the Ford ***Mustang***, ***Pinto*** and ***Bronco***, the Dodge ***Colt***, Subaru ***Brumby*** and the Hyundai ***Equus***.



Suppose, for the purposes of a bawdy TV drama, we wish to weave a blend that reflects poorly on the Catholic church and on the papacy specifically. Strong associations of **Pope** include *God*, *Catholicism*, *holiness, fatherhood* ("*Il Pappa*"), *Rome*, *Italy, leadership*, *Vatican* and *Infallibility*. From these benign materials we might combine **God** and **Father**hood to provocatively view **Popes** as ***Godfathers*** that ***lead Italian***-style crime families. Pope Francis might thus become Don Francisco, while the Vatican and its bank might be seen as well-greased machines that, in the words of Camus, "*dispense with generosity in order to practice charity*." This is the premise of a libidinous TV show, ***The Borgias***, which pitches that corrupt clan as *The Original Crime Family*.



Large Word Colliders: **Experiments in Linguistic Invention**

Picture the scene: a group of professional "creatives" are spit-balling in the conference room, trying to invent as many meaningful and original new words as possible. A fun exercise, perhaps, but why would a company ever pay expensive consultants to play word games? Yet this kind of wide-ranging brainstorming exercise is more common than you might think, as it is not at all unusual for a large company to strive to maintain a large inventory of catchy product names in advance of any new products that might ultimately bear those names.

The marketing wizards at Greek airline ***Olympic*** could have certainly benefited from the advice of a professional ***naming consultant*** when it branded its frequent flyer program with the ill-conceived name **Icarus**. For the Greek aeronaut who flew too high, and fell into the sea when his wings melted, is hardly a figure to inspire the confidence of air travellers. The pharmaceutical giant **Pfizer** shows more savvy with names, and maintains an inventory of candidate names for any new drugs that might pop out of its development pipeline. These names are first brainstormed without prior knowledge of the specific drugs that they may one day grace, and ranked according to a range of product-independent criteria, such as "*is the name novel?*", "*is it catchy?*", "*does it suggest positive qualities?*" and "*does it mean anything offensive, taboo or silly in any of the languages/countries in which it may be marketed?*". While one might think that **Pfizer** coined the name "**Viagra**" for its best-selling impotence drug because this name is especially resonant of the idea of male sexual "***vigor***", or because it suggests tremendous natural power due to its similarities with the word "***Niagara***", the name was actually coined long before **Pfizer** ever developed the drug or identified its value in treating sexual dysfunction. Nonetheless, **Viagra** is a good example of what **Lewis Carroll** called a ***portmanteau word***: a word made of two clipped parts and two tightly integrated meanings. Though brainstormed *before* product development, the name was chosen from a list of candidates *after* development, which is why it seems so especially apt. If chance favors the prepared mind, the pre-generation of a wide-range of well-formed candidates can yield real benefits for any creative decision.

As an experiment in word-level creativity, let's simulate the brainstorming of new words by

using the **cut-up** technique to combine meaning-laden fragments of existing words. This will follow the same process as the brainstorming of product-names: new words are generated only when they are suggestive of sensible meanings that could plausibly benefit from a convenient single-word label. These fragments are *meaning-laden* in the sense that each word cutting – either a word prefix or a word suffix – has its own established meaning. Thus, we shall use cuttings like "***astro***" to mean star, or "***techno***" to mean *technology*, or "**naut**" to mean *traveller*, but not arbitrary cuttings like "***sp***" to stand for *spoon* (as in "***spork***") or "***unch***" for *lunch* (as in "***brunch***"). The idea of using cut-ups of word fragments to coin creative new word blends is a well-established one, and Computer scientist **Wlodzislaw Duch** gives these word fragments a name: ***wordels*** or ***word el****ements*, by analogy with visual ***pixels*** and **pict**ure elements.

Designer words, like designer drugs, are bespoke compounds that are built using an inventory of well-understood parts of proven efficacy. The bigger our inventory, and the more reliable and reusable its parts, the more words we can design. We can assume that most natural place to cut a word in two, for purposes of adding to our inventory of reusable prefix and suffix elements, is also the most natural place at which to hyphenate. Consider how we break up a word when it is too big to fit at the end of a line: we do not insert the hyphen just anywhere, but use our intuitions about spelling, pronunciation, and morphemic composition to place it between the most meaningful sub-parts of a word, such as just after a morphemic **prefix** or before a morphemic **suffix**. Not everyone hyphenates sensibly, of course, but by scanning a large collection of texts (a *corpus*) we can expect the most frequent hyphenations to also be the most sensible. In such a corpus, we are more likely to encounter "**God-zilla**" than either "**Godz-illa**" or "**Go-dzilla**", while "**astro-naut**" is far more likely than either "**astron-aut**" or "**astr-onaut**". So we need not articulate the specific intuitions that writers use when breaking up words: we – and our machines – can simply learn the hyphenation preferences of specific words from a very large corpus like the Web.

The Design Conspiracy* is a trendy London-based consultancy firm with strong opinions about** what constitutes *good* – and *bad* – design. In 2003, the company grabbed headlines when its spoof web-site offering contrived company names on demand was taken seriously by some businesspeople. The site, a publicity stunt named *WhatBrandAreYou.com*, works a little like those *tell-your-fortune* machines that used to be popular in amusement arcades and funfairs: the user is asked for certain personal details, such as whether you or your company can be described as "*passionate*", "*dynamic*" or "*client-focused*", and these are then promptly ignored. Like the arcade machines, this spoof site works with a stock of canned outputs (about 150 in all), from which the site chooses randomly for each user. Its outputs include credibility-stretching names such as ***Ovisovis, ***Amplifico***, ***Bivium***, ***Libero***, ***Vulgo***, and ***Ualeo***, each produced in a bout of offline brainstorming in which the culprits admit that they "*were just literally trying to think of the most stupid company names.*"

The site's creators were inspired by what they consider to be the silly and contrived names of existing companies, such as ***Accenture***, ***Diageo*** and ***Consignia***. When interviewed by the BBC, the company described its brainstorming process as follows: "*We used an online Latin dictionary to come up with some of the names, and just added an 'i' or and 'a' to the end. Others, like **Ualeo**, we don't even know how to pronounce.*"

Names like ***Ovisovis*** suggest that a more general combinatorial process is also used. So this is a brainstorming process not unlike the cut-up technique, in which reusable word fragments – catchy Latinate roots and their allowable (and cool-sounding) affixes – are mixed and matched to create novel combinations. The company does this for very different reasons, of course, since *WhatBrandAreYou.com* is intended to showcase the perils of simplistic and rule-governed design, as evident in the expensive but uninspired names crafted by rival consultants. Nonetheless, the BBC reports that some of the spoof names, such as ***Tempero***, ***Integriti***, ***Winwin*** and ***Ovisovis***, have subsequently been snapped up and registered as legal company names by businesspeople with fewer design qualms. Of these ironic successes *The Design Conspiracy* observes: "*clearly, we have an aptitude for thinking up company names. But then, it's a lot easier than it seems.*" The web-site *WhatBrandAreYou.com* really is a design conspiracy then, one that shows that the combinatorial generation of new words is more akin to brain***stem***-storming than true brainstorming.

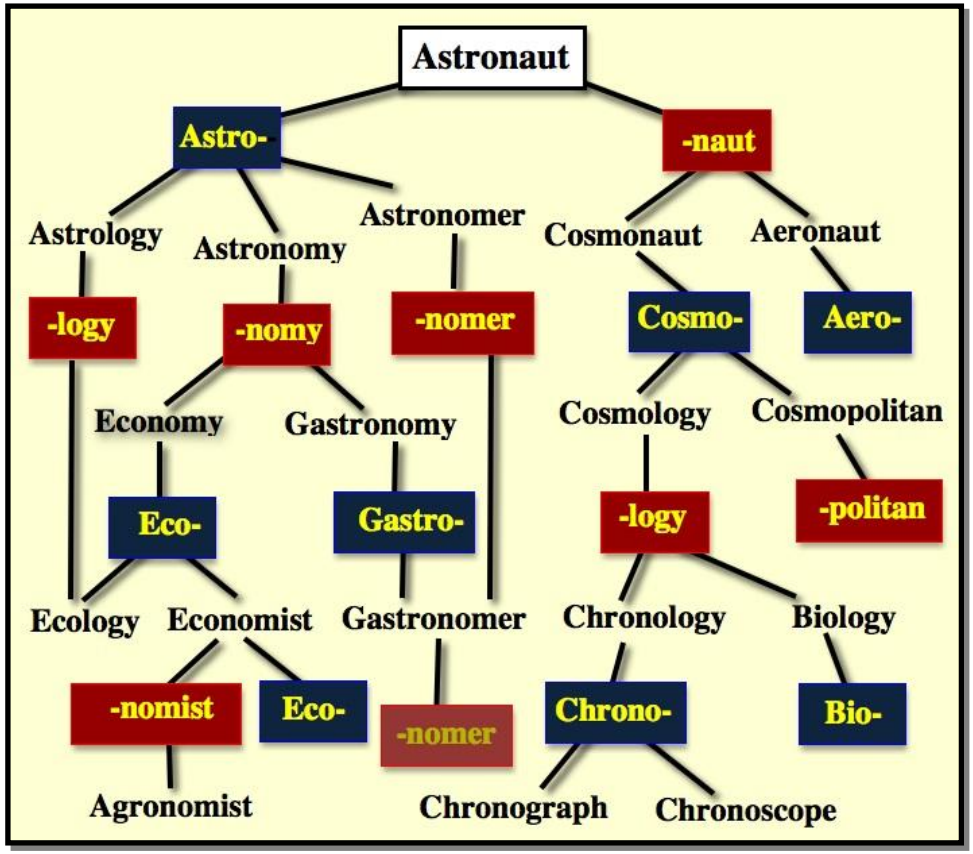
Yet, in spite of themselves, even these attempts at trivialization cannot help but produce words that other people find useful and creative. As ***mockumentary*** star **David St. Hubbins** comes to realize in ***This is Spinal Tap***, "*It's such a fine line, between stupid and clever*". Though brainstorming is always a scattershot process, we can use **CC** techniques to nudge this process toward the clever side of this fine line: firstly, **CC** should not generate purely random word mash-ups, by fusing together any prefix with any suffix; secondly, neither should it judge a candidate word to be a successful blend based purely on how cool or how *stupid* it seems, but on the usefulness of its meaning.

Since new words are minted in response to specific linguistic needs, we should begin by imagining what these needs might be. Suppose we find ourselves using the phrase "*food traveller*" with connotations that arise from neither "food" nor "traveller" (such as *brave* and *adventurous*), and want to mint a catchy name for these footloose gourmands. We might thus consider all of the possible ways of expressing these component ideas in two word fragments that can be blended into an integrated whole. Looking to an inventory of useful word fragments, we see that the prefix ***gastro-*** can signify **food** and the suffix ***-naut*** can signify a **traveller**. Indeed, ***-naut*** is most recognizably used in a word, "*astronaut*", that has precisely the connotations of *bravery* and *adventure* that we are seeking for our new creation. So fusing both fragments into one, we arrive at the new word ***gastronaut***.

Though a simple process, this is neither random nor haphazard. Note how the word fragments **gastro-** and **-naut** are suggestive of the words they are most associated with, so the new coinage "***gastronaut***" can also be seen as a portmanteau of sorts, of the words "***gastro****nome*" and "*astro****naut***". Now imagine using CC to replicate this simple process on a much grander scale, to generate thousands of useful new words that convey real meanings.

Words that share a prefix or suffix can be considered neighboring states in a conceptual space.

So given a set of exemplar words that is representative of those we want to automatically create – the **CC** theorist **Graeme Ritchie** calls this set an ***inspiring set*** – it is a simple matter to explore outwards from these starting points, to discover other suffixes that are used with an inspiring prefix, or other prefixes that are used with an inspiring suffix. As shown in the figure below (right), even a timid exploration of the space will reveal a host of new suffixes and prefixes that are of the same ilk as those in our inspiring set.



Suppose we start with the exemplar ***astronaut***. The suffix -**naut** leads to ***cosmonaut*** and **aquanaut**, which yield **aqua-** and **cosmo-** as new prefixes. The latter then leads to ***cosmopolitan*** and **cosmology** to yield -**politan** (which later provides **metro-**) and -**ology** (the mother-lode of Greek prefixes). The same suffixes and prefixes are likely to be repeatedly found by following different paths, as in the case of **logy** (which is found via ***astronaut***→**astro**→***astrology***→**logy** and →**naut**→***cosmonaut***→**cosmo**→***cosmology***→**logy**. As you might expect, the more frequently an affix is used in different words, the more often we are going to encounter it during this *divide-and-recruit* process. But this redundancy does not lead to wasted effort, since the number of times we find the same suffixes and prefixes is a good indicator of how useful those fragments will be in the creation of new words. In this way we can automatically identify the most useful prefixes and suffixes to add to our inventory of word fragments.

With a large frequency-ranked inventory of reusable word fragments in hand, we must now

consider how to assign a conventional meaning to each, such as *star* for **astro-** and *fear* for **-phobia**. We can explore a variety of automatic methods for doing this, such as using an online etymological dictionary, but the quickest and most reliable approach is also the most *old-fashioned*: to obtain quality output we need quality inputs, and there really is no substitute for the manually-assigned meanings of a fully-engaged native speaker.

-logy	*study, discipline, system*	**chrono-**	*time*
-scope	*display, monitor, viewer*	**gastro-**	*food, cooking, eating*
-naut	*traveller, sailor*	**necro-**	*dead, death*
-polis	*city*	**pyro-**	*heat, fire, flame*
-metry	*measurement*	**psycho-**	*mind, mental, crazy*
-mancy	*magic*	**geo-**	*place, location*
-pyle	*opening, gap, vent*	**cryo-**	*cold, ice*
-nym	*name*	**hydro-**	*water*
-trope	*seeker, follower*	**helio-**	*sun*
-glyph	*marking, symbol, icon, rune*	**phono-**	*sound, hearing*

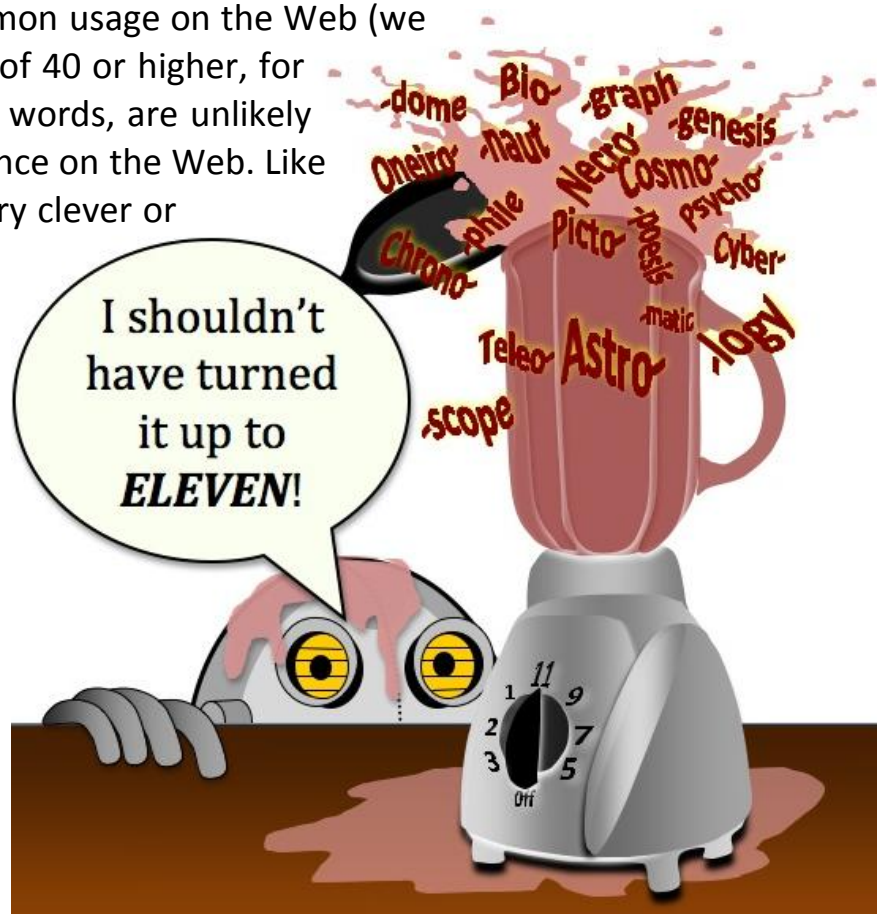
As shown in the treatment of sample word fragments on the right, we concentrate our energies on fragments of a **Greek** origin (whose prefixes often end in "o"), as these produce the most transparent *and* the most promiscuous prefixes and suffixes. As we are dealing with a total inventory of approximately **400 word fragments** here, we may as well do it properly, and manually assign the meanings that are most strongly associated with each word fragment. As such, we can just skip over the strict dictionary meaning of **gastro**- (denoting *belly* or *stomach*, as in ***gastro**-enteritis*) to use the more frequent and less medically-oriented glosses *food* and *cooking* instead. Likewise, we can overlook the strict dictionary definition of the prefix **metro-** (denoting *womb* or *uterus*, as in ***metro**rrhagia*, and thereby denoting *mother* in ***metro**polis*), to assign the more conventional (if strictly incorrect) modern associations **urban** and **city**, even it this makes **metropolis** something of a redundant combination.

David Bowie described his *Verbasizer* as a tool for randomly slamming words into each other.

Yet few collisions produce novel juxtapositions from which useful meanings can be gleaned, and most yield nonsense that must be manually filtered. Fortunately, we can build a simple **CC tool** to automatically shoulder the burden of doing much the same with our inventory of word fragments, to swirl them around in a simulation of ***Brownian motion*** and autonomously explore which fragments glom together as truly meaningful combinations. Let us say that a combination of the prefix **P**- and suffix **–S** is *meaningful* if **PS** denotes something that is worth talking about, and *creative* if **PS** is a novel way of doing the talking. Each of these conditions is easily automatized in in the context of a very large corpus of Web texts. First, **PS** is semantically plausible if the explanatory glosses of **P**- and **–S** form a phrase that can be found multiple times in our corpus. Second, **PS** is creative if it is meaningful *and* if it is not found in our corpus. Thus, the word "**oneironaut**" – a combination of **oneiro-** (*dream*) and **–naut** (*explorer*) – is seen as meaningful since the phrase "*dream explorer*" is found more than once in our corpus. Likewise, "**oneironaut**" is seen as creative because it is meaningful and is *not* found in our corpus.

There is an undeniable energy to **Bowie**'s metaphor of words vigorously slamming together in the linguistic equivalent of a **cyclotron**. Yet waste is an inherent feature of his approach and few collisions yield productive meanings. So there is little point in replicating the natural inefficiency of the cut-up technique when we apply it on a computer. Why produce so many random pairings of elements and attempt to give meanings to just a small subset, when our computers can just as easily work in the other direction, from meaningful glosses to novel forms? Suppose we unleash our **CC tool** on a large corpus, to look for juxtapositions of words from our inventory's set of annotated glosses. When our tool encounters "*dream explorer*", say, it concocts as many novel forms as it can from the fragments associated with the glosses "*dream*" and "*explorer*", to invent **oneironaut** in recognition of the prior existence of its meaning.

As *St. Hubbins* observed, the line between *clever* and *stupid* is a fine one, and most new words suggested by our swirling together of word fragments will cluster ambiguously around this line. Are they ***H-Creative*** or just ***H-Stupid***? When we run our **CC process** on a source of Web text snippets called the **Google n-grams**, it uses its inventory of 400+ fragments to invent **90,409 new words** for existing phrasal glosses. Just **769** of these words (a mere **0.85%**) can be found in a conventional dictionary (we used **WordNet**, on online electronic dictionary, for this test). Another **2,690 words** (just **3%**) are found to be in relatively common usage on the Web (we used the **Google 1-grams**, a list of words with a Web frequency of 40 or higher, for this test). The remainder, constituting **96%** of the newly minted words, are unlikely to be found in any print dictionary, and have no significant presence on the Web. Like many of the designer objects in modern society, they may be very clever or very stupid, and in some cases they can even be both.



We can appreciate the variety of our lexical inventions by looking at all of the words that are minted from a single word fragment. So let's focus on the productive suffix **-naut**, which earlier gave us the interesting *P-Creative* coinage **gastronaut**. Meaning-driven cut-up generation can identify more uses for a fragment like **-naut** than one might expect, and the following list of 142 neologisms and their glosses speaks for itself. Dictionary-defined words are shown in yellow, while *P-Creative* uses with a Web-frequency of 40 or higher are capitalized. Words with at least one usage on the Web (at the time of writing) are shown in italics. Note that just one meaning is listed for each term, although some new words are suggested by several meanings, and different words may be invented for the same meaning.

Gerontonaut	"age explorer"
Oxionaut	"air traveller"
Avionaut	"airplane traveller"
Allelonaut	"alternative traveller"
Archeonaut	"history explorer"
Taxonaut	"type explorer"
Phononaut	"audio explorer"
Aristonaut	"royal traveller"
Primonaut	"first explorer"
BIONAUT	"life explorer"
NEURONAUT	"mind explorer"
Pontonaut	"bridge explorer"
Numismatonaut	"cash traveller"
Speleonaut	"cave explorer"
Dynonaut	"change explorer"
Vianaut	"portal explorer"
Civinaut	"citizen explorer"
Metronaut	"city traveller"
Genenaut	"gene explorer"
Spectronaut	"light explorer"
Chromatonaut	"color explorer"
Encyclonaut	"comprehensive traveller"
Zeugmanaut	"connection explorer"
GASTRONAUT	"gourmet traveller"

Aironaut	"air traveller"
AERONAUT	"air traveller"
Xenonaut	"alien explorer"
Paleonaut	"ancient explorer"
Anthonaut	"archive explorer"
SPACENAUT	"space explorer"
Autonaut	"automatic explorer"
Optinaut	"great traveller"
Magninaut	"big explorer"
Biblionaut	"library explorer"
PSYCHONAUT	"crazy traveller"
Stegonaut	"hidden traveller"
Aetionaut	"cause explorer"
Cytenaut	"cell explorer"
Arterionaut	"channel explorer"
Rheonaut	"circulation explorer"
Citinaut	"city explorer"
Juxtanaut	"close traveller"
Cryptonaut	"code explorer"
Chloronaut	"green traveller"
Holonaut	"complete traveller"
Meronaut	"component explorer"
Geonaut	"world traveller"
Idionaut	"personal traveller"

Nyctalonaut	"night traveller"
Oceanonaut	"ocean explorer"
Diplonaut	"diplomatic traveller"
Telenaut	"distant traveller"
Pictonaut	"image explorer"
Aquilonaut	"eagle explorer"
TERRANAUT	"earth explorer"
Ergonaut	"energy explorer"
Photonaut	"light explorer"
Histerionaut	"time explorer"
Opthalmonaut	"eye explorer"
Videonaut	"image explorer"
Matronaut	"female traveller"
Pyronaut	"fire explorer"
Ludonaut	"fun traveller"
Bononaut	"good traveller"
Misonaut	"hate explorer"
Anthroponaut	"man explorer"
INFONAUT	"message explorer"
Endonaut	"inner explorer"
Veronaut	"knowledge explorer"
Lexinaut	"word explorer"
MACRONAUT	"large traveller"
Cartonaut	"map explorer"

Plutonaut	“money traveller”
Heteronaut	“diverse traveller”
Disconaut	“disk explorer”
Cynonaut	“dog traveller”
ONEIRONAUT	“dream explorer”
Protonaut	“first traveller”
ELECTRONAUT	“electronic traveller”
Thermonaut	“energy explorer”
Anglonaut	“English explorer”
Phraseonaut	“expression explorer”
Visionaut	“picture explorer”
Tachonaut	“fast traveller”
Oligonaut	“powerful traveller”
Bromanaut	“food explorer”
Nymphonaut	“girl traveller”
Chironaut	“hand explorer”
CHRONONAUT	“time traveller”
CRYONAUT	“ice explorer”
Radionaut	“radio explorer”
Hibernonaut	“winter traveller”
Limnonaut	“lake explorer”
Verbonaut	“language traveller”
Patronaut	“male traveller”
Maginaut	“master traveller”

Glossonaut	“message explorer”
Logiconaut	“science traveller”
Lunanaut	“moon traveller”
Audionaut	“sound explorer”
Neonaut	“new traveller”
Nomonaut	“number explorer”
Mononaut	“single traveller”
Phytonaut	“plant explorer”
Tectonaut	“structure explorer”
Temponaut	“time explorer”
ROBONAUT	“robotic explorer”
Petronaut	“rock explorer”
Spherenaut	“world traveller”
Veneranaut	“sex traveller”
Conchonaut	“shell explorer”
Cielonaut	“sky traveller”
Leptonaut	“tiny traveller”
Acronaut	“tall traveller”
Horonaut	“time traveller”
Nanonaut	“tiny explorer”
HYDRONAUT	“water explorer”
Oenonaut	“wine explorer”
Logonaut	“word explorer”

Mnemonaut	“memory explorer”
Psychenaut	“soul explorer”
Cinenaut	“movie explorer”
Econaut	“nature explorer”
Noxinaut	“night traveller”
Oleonaut	“oil explorer”
Typonaut	“print explorer”
Floranaut	“plant explorer”
Dystonaut	“poor traveller”
Spironaut	“ring explorer”
Lithonaut	“rock explorer”
Quixonaut	“romantic traveller”
Technonaut	“science explorer”
Eronaut	“sex explorer”
Dermanaut	“skin explorer”
MICRONAUT	“tiny explorer”
ASTRONAUT	“space explorer”
Teleonaut	“task explorer”
AQUANAUT	“water explorer”
Chemonaut	“chemical explorer”
Meteoronaut	“weather explorer”
Dipsonaut	“wine traveller”
COSMONAUT	“universe explorer”

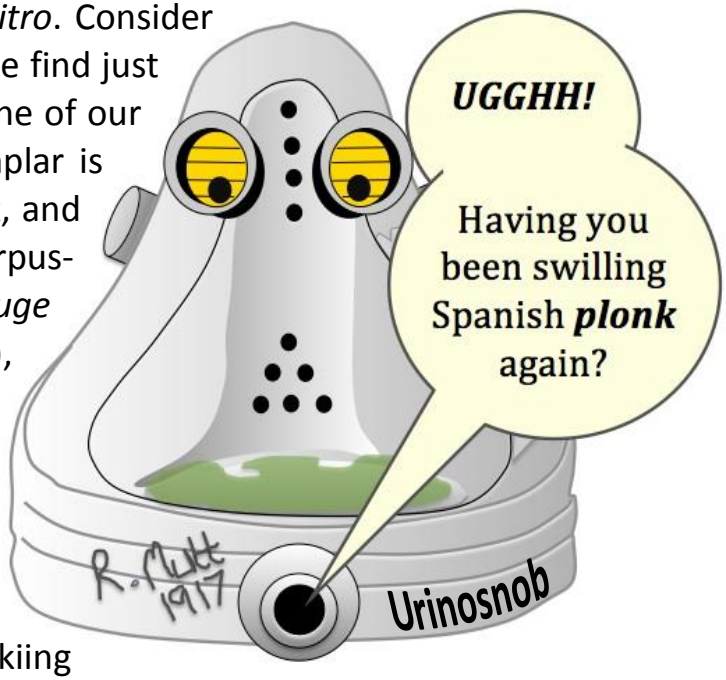
H. G. Wells coined the term "Time Machine" in his 1895 book of the same name, and it has since become the preferred name for any device that allows travellers to move back and forth in time. Wells was not the first writer to explore the notion of time travel, nor was he even the first to think of it in purely mechanical terms (that is, facilitated by a machine rather than by magic, by dreams or by some other fantastical plot device). A "time machine" is a somewhat uncreative name for a machine that allows time travel – It could just as well describe a clock – but it certainly makes for a catchy title, and what it lacks in creativity it makes up for in simplicity and resonance. Wells presumably reasoned that the concept of time travel was a creative enough premise as it was (if not exactly a ***H-Creative*** one), and already presented enough of a challenge to his audience without saddling his story with an obscure name like "***The Chronomat***". But Wells wasn't always a fan of simple names. Seven years earlier, when publishing a precursor to his time travel adventure in a school journal, he gave it the enigmatic if not entirely successful title "***The Chronic Argonauts***". But good sense prevailed when the time came to write his novel, and Wells ditched the label, if not the idea, of a "*chronic Argonaut*". Today, though time travel is still confined to the realms of fiction and speculative physics, the conceit is more popular than ever, and as if to mirror the simplicity of "time machine", the phrase "time traveller" is still the most common label for fictional characters who jump through time.

Just as some people have faces that are made for radio, "**Chronic Argonaut**" is a clunky name whose charms are definitely more conceptual than linguistic. Fortunately, our list of cut-up words provides a variety of **naut**ically-themed labels that may better fit the bill. For instance, a **Horonaut** (*clock explorer*) might aptly describe time travellers who make small jumps in time, as measured on a clock rather than on a calendar. Alternately, **Histerionaut** (*time explorer*) and **Archeonaut** (*history explorer*) might well describe a time traveller who makes significant leaps back into history, while **Paleonaut** (*ancient explorer*) might describe truly adventurous travellers who jump all the back into prehistoric times. In physics, a tachyon is a particle that can travel faster than the speed of light and thereby move backwards in time, so a **Tachonaut** (*fast traveller*) might likewise describe a time traveller who exploits the same temporal loopholes. Finally, for purists, **Chrononaut** and **Temponaut** are just compressed ways of saying "*time traveller*", though ***Temponaut*** has the added merit of implying that time travellers also need a good sense of rhythm.

These words and their glosses are not the end results of the creative process, but intermediate products that can still lead to new meanings and to new insights. After all, glosses like "*oil explorer*" (for **Oleonaut**) and "*ancient explorer*" (for **Paleonaut**) are not meanings *per se*, but more-or-less ambiguous phrases that help us to find contextually-apt meanings. It's up to us whether the words of a gloss are interpreted in their dictionary senses, or whether another measure of creativity is used to metaphorically stretch them into new senses. Thus, ***Paleonaut*** is used on the Web to describe those ***systems administrators*** who search with an *archaeologist's zeal* to restore dusty old programs that are no longer supported. The "ancient" of "ancient explorer" is not the "ancient" of "*ancient Babylon*" here, but the "ancient" of grumpy old men, 8-track tape-decks and black-and-white TVs. Similarly, the glosses for **Oleonaut** and **Paleonaut** both use "explorer" in a metaphorical rather than a literal sense, not to mean a person who pushes back the boundaries of geographical knowledge, but one who "*explores new options and possibilities*".

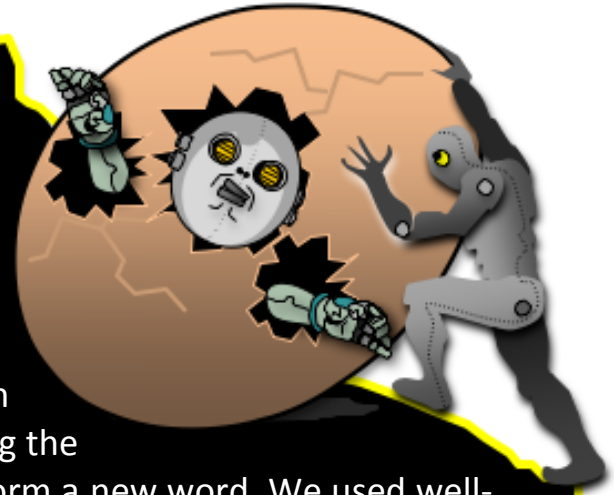
Creativity is a never-ending process. By courting ambiguity, we can view our initial creative outputs as ***Rorschach blots*** that spur further metaphorical analysis and further creativity. Consider the gloss "*royal traveller*" for **Aristonaut**. The adjective "royal" may be taken literally here, to denote an aristocratic jet-setter who always travels first-class. Or, it can be taken metaphorically: some Web-users employ the term to describe billionaires who buy their way into space – so-called *space tourists* or self-funded astronauts – and who stay at that orbiting space hotel, the **International Space Station** (**ISS**). Since one would have to be filthy rich to buy a commercial ticket into orbit, the novel term **Plutonaut** is also appropriate here, though neither **Aristonaut** nor **Plutonaut** have yet gained widespread acceptance in this creatively descriptive role.

Our new words are the lexical equivalent of *test-tube babies*: created *in vitro* in the lab, but not always viable *in vivo*. Both processes can be wasteful because both are subject to risks, so many more possibilities are created than will ever survive or thrive. The combinatorial process at the heart of our approach ramps up word production to an almost industrial scale, so that the potentials that are evidenced in just a small number of words *in vivo* are magnified and catalyzed into a much larger set of possibilities *in vitro*. Consider the word "snob": it has a nasty meaning but a delightfully compact sound. We find just one Google 1-gram in which "snob" is used as a suffix in combination with one of our inventoried prefixes: **Vinosnob**, meaning a "*wine snob*". But this one exemplar is enough to suggest that "snob" has a much greater potential as a word suffix, and when **-snob** is added to our inventory as a suffix (meaning *snob*, naturally), corpus-driven combination cranks out 95 new inventions, including **Brontosnob** (*huge snob*), **Bibliosnob** (*book snob*), **Metrosnob** (*city snob*), **Endosnob** (*inner snob*), **Cinesnob** (*movie snob*), **Ethnosnob** (*culture snob*) and **Cryosnob** (*snow snob*).



There are as many kinds of snobs are there are things to be elitist about in society. If some of these possibilities seem odd or implausible, like "*snow snob*", remember that the indicated gloss in each case is always an attested phrase from the Google 2-grams: as the phrase has meaning to *someone*, so too might the word. A moment's thought suggests that a "*snow snob*" is a skiing enthusiast with annoyingly superior ideas about what constitutes the best kind of snow – or the best slopes – for skiing. In some cases, the non-availability of a given phrase in the Google 2-grams means that a potential meaning is never validated in a meaningful combination. Thus, since "*secret snob*" is not an attested Web 2-gram, the possibility that a **Cryptosnob** might be a "*secret snob*" (or in other words, a person who keeps their *endosnob*bery well hidden) is never recognized. In contrast, "*secret tourist*" *is* an attested 2-gram (with a Web frequency of 75), which permits the invention of **Crypotourist** to describe travellers who sneak into foreign territories without the necessary permissions.

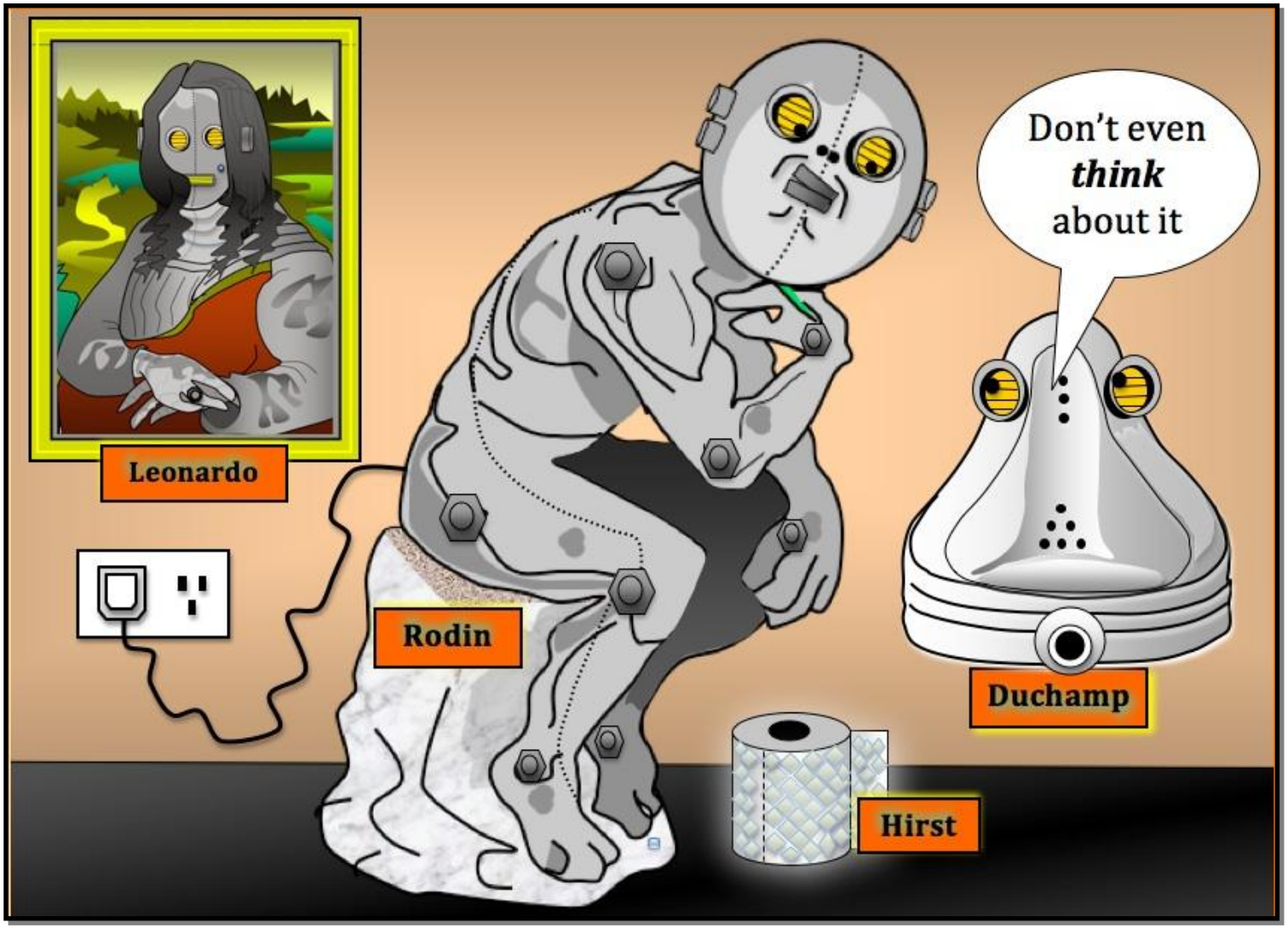
Test-Tube Baby-Names



Baby names are big business. Bookshop shelves groan under the mass of books that aim to help you *name that baby*! While may not yet live in the age of *designer babies*, we do live in the age of designer baby-*names*. Parents that want a special name for a special baby need only invent their own, and those who cannot settle on a single name can always create a blend of multiple names with multiple resonances. We have seen how new words (neologisms) can be invented using the *portmanteau principle*, in which clippings of two or more words are concatenated to form a new word. We used well-formed clippings that correspond to morphemes such as **gastro-** and **-naut**, but nothing stops us from using much the same approach with prefix and suffix cuttings that have been clipped from a large inventory of existing names.

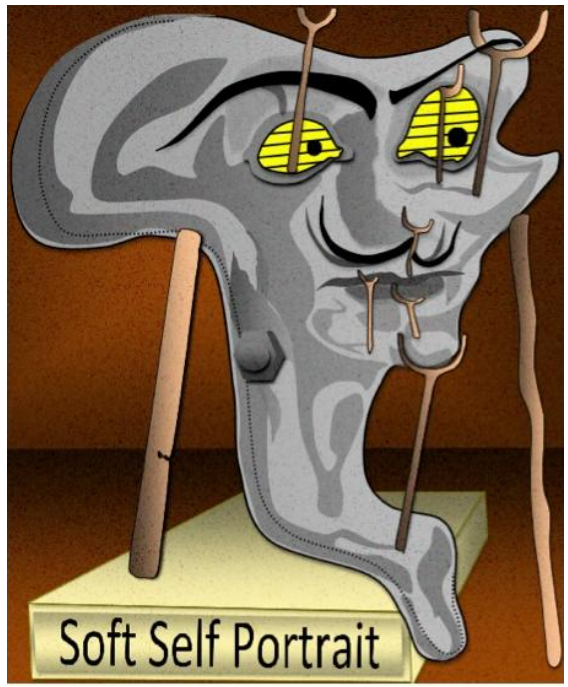
- Suppose we call these cuttings namelets. How many namelets would we need to build a capable generator? How would we associate meanings with namelets? What kind of properties would we associate with namelets?
- What rules should we use to maximize the chance that our namelet combinations yield euphonious names?
- Company names and product names are also big business for consultancy firms. Design a pool of namelets to allow the creation of names like Microsoft, eBay, CompuWorld, TurboLaser, TrueTech, SunGuard, etc.
- To facilitate the generation of meaningful names, write your namelets on cards, with associated properties on the other side (e.g., Turbo on one side of a card, fast *and* powerful on the other). Now generate new names by first combining cards with the properties-side up. What else can you do to promote meaningful combinations?
- What other types of knowledge can/should a computer use to automatically generate baby/product names?

Extra Credit: Many names are readymades, common nouns that are reused as proper-names. How might a computer suggest potential readymades to serve as names for a given person or product?



Think Like a Machine, Create Like a Master

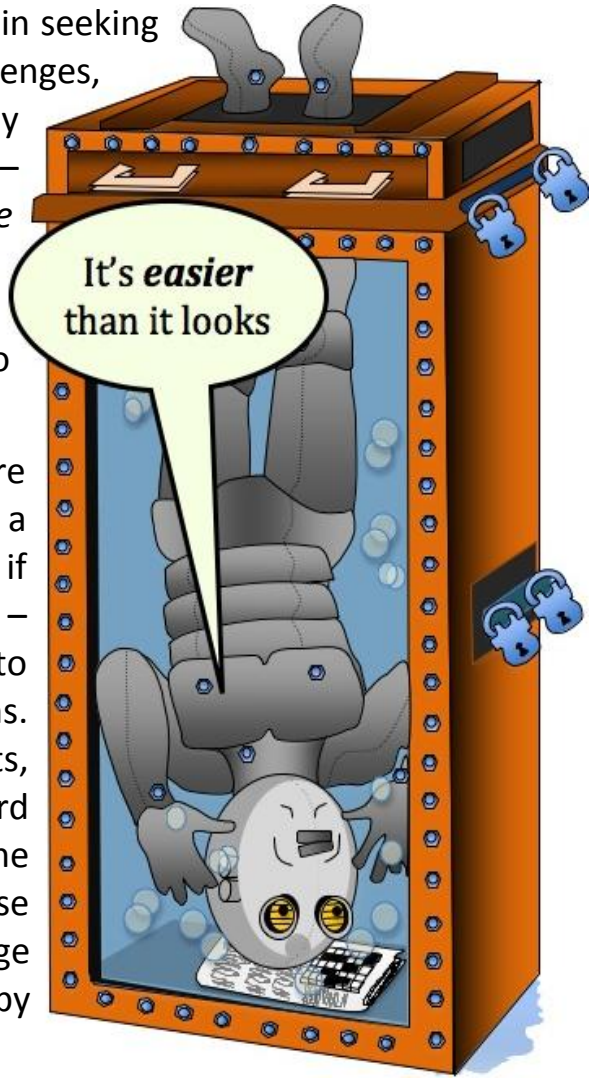
We dismiss artists as "*machines*" when their many outputs are repetitive, lacking in diversity, overly mannered or seemingly the product of a simple formula or generative device. Yet some artists consciously seek to turn themselves *into machines*, precisely so they may avoid the predictability of mannerism and cliché. As **Salvador Dali** put it, "*In truth I am no more than an automaton that registers, without judgment and exactly as possible, the dictate of my subconscious*." Dali formulated a process, which he dubbed his **paranoiac-critical method**, to interrogate his subconscious and to provoke his mind — more specifically, his **System 2** – into attempting a rational analysis of "*delusional phenomena*", such as the confused images that flicker and fade as one wakes from a dream. Dali sought to systematize the production of subjective errors and irrational images from which he might faithfully reproduce realistically-rendered "*hand-painted dream photographs*." So Dali mechanically imposed method and realism on transitory madness and illusion, tricking his mind into becoming a generator of surreal imagery. But are there other ways in which a creator can think like a machine to achieve original results?



Dali actively sought to counter the passive influence of implicit biases and unwritten constraints by directly appealing to his subconscious. But another approach to neutralizing these baleful influences is to strive to do as AI researchers do whenever they build a new generative system: to make any implicit biases *explicit* and to *write down* and formalize any unwritten constraints. For like cholesterol and bacteria, not all constraints are objectionable: where some will frustrate and impede, others can guide and inform our explorations of a conceptual space. As figuratively depicted in Dali's ***Soft Self Portrait***, the most useful constraints lend shape and support to a space, helping to turn our vague and amorphous intuitions into well-defined states. Only when we have explicitly stated our constraints can we begin the exploratory process of consciously tweaking some and removing others, with the aim of achieving surprising novelty and even, in rare cases, *transformational* creativity.

Robert Frost once quipped that *he'd sooner play tennis without a net* than write free verse.

The problem of imposing rhythm and rhyme onto words was, for Frost, an essential "sporting" challenge, one to be met with relish by the verbal athletes who produce poetry, and one that *free verse* – in seeking a *free ride* – significantly diminishes. Creators often excel in the face of daunting challenges, producing compelling results that are tighter and leaner than those achieved by creators untroubled by constraints. **Orson Welles**, an auteur whose best work – indeed, whose whole career – was shaped by budgetary restrictions, noted that "*the absence of limitation is the enemy of art*." Not only does the constraint of rhyme present a satisfying challenge to a poet, it promotes the creation of poetry that wrings maximal artistic benefit from the ***Keats heuristic***. A poet must *work hard* to imbue a poem with the "divine skip and jump" of profundity in metrical form.

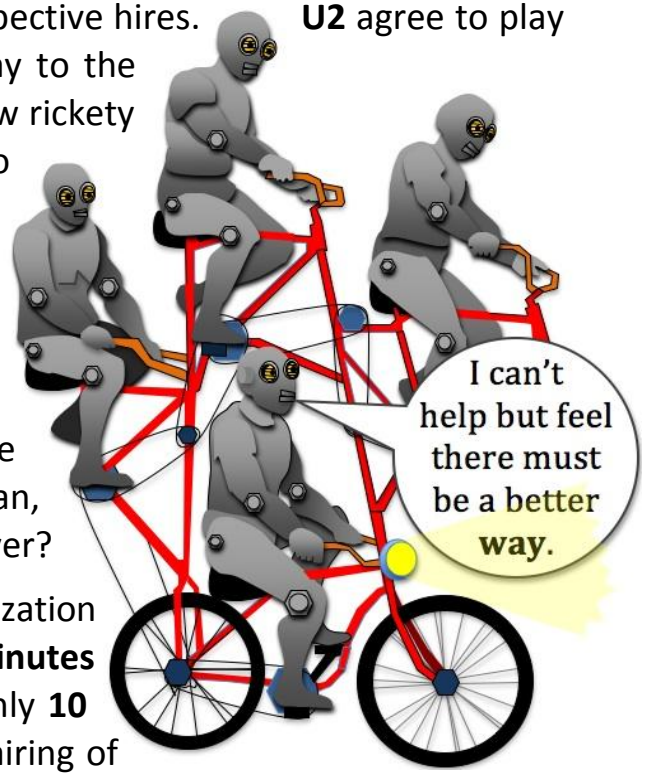


It is intuitive to believe, as Frost apparently did, that the constraints of art are additive, so that the imposition of each new constraint always increases the heft of a challenge, while the removal of *any* constraint always results in a lighter burden. Yet if we adopt a *computational* perspective on constraints, and look to see their effects – both individual and collective – on a search space, we see that fewer constraints lead to larger search spaces but to less information on how best to explore these vast terrains. Consider ***crossword puzzles***: the intersecting strips of the grid impose strict constraints, and make it harder to fill in a crossword with just any old words. A half-filled crossword is more constrained *and* easier to solve than one that is empty and untouched. If the goal of a crossword is to uncover the answers intended by its setter, then these constraints also make it easier to find those precise words, by gradually closing off large swathes of the search space, by providing a quick test for each candidate filler, and by incrementally unveiling partial *retrieval cues* to help us recall the "right" answers.

The exploratory search of conceptual spaces is often plagued by the *curse of dimensionality*, where the problem being solved appears to have too many dimensions of interest and so translates into a large and unwieldy search-space. In such cases, a shrewd statistical analysis can identify the major dimensions, or ***principle components***, of the problem, and allow us to build a concise and manageable space to search. But even problems of modest dimensionality can produce unwieldy search spaces if we naïvely opt for a state representation that is unnecessarily bloated. Consider the ***Eight Queens problem***, a famous puzzle in AI that requires us to place eight queens on a chess-board in such a way that no queen threatens any another. Because the queen is maximally mobile in chess – she is free to move along any row, column or diagonal on which she is a placed – any two queens will threaten each other if they are placed on the same row, column or diagonal. On an **8x8** chessboard with 64 choices for the 1st queen, 63 for the 2nd, 62 for the 3rd, and so on, there seems to be ***64x63x62x61x60x59x58x57*** possible configurations – that's a space of **178,462,987,637,760** states – for our eight troublesome queens.

Yet we need not ditch any of our queens, as an appreciation of the puzzle's constraints can help us to dramatically shrink the space instead. For example, as there is never any point in placing a queen in a column that is already occupied, each queen must be assigned to a row in its *own* column. We thus obtain a smaller search-space of ***8x8x8x8x8x8x8x8***, or **16,777,216** possible states, to explore. But this space still contains considerable redundancy, as two queens can also never occupy the same row. So our space need contain only ***8x7x6x5x4x3x2*** states, corresponding to the **40,320** configurations in which each queen is placed at the intersection of its own distinct row and its own distinct column. This much-reduced space is **4,426,165,368** (that's 4.4 *billion*) times smaller than the bloated space we first started with. What a difference a little forethought can make.

An explicit inventory of our constraints allows us to apply them in the most effective order, and may even allow us to pair constraints with each other in a symbiotic – and, to some extent, *neutralizing* – fashion. Consider the following puzzle, of the kind that tech companies pose to their prospective hires. **U2** agree to play a charity gig in a small town in the west of Ireland, but as they make their way to the parish hall on a dark and rainy night, they must abandon their tour bus at a narrow rickety bridge that can only support the weight of two people. They will need a torch to pick their way through the gloom, but luckily a single flashlight is found on the bus. They estimate that **Larry**, carrying the drum kit, will take ***10*** *minutes* to cross the bridge; **Bono**, tottering on high-heeled boots, will take ***5*** *minutes* to cross over; **Adam** will take ***2*** *minutes*, and fleet-footed **Edge** will take just ***1*** *minute* to cross. So they decide to cross with the torch in pairs, at the speed of the *slowest* in each pair, though someone will have to return with the torch to the other side until all have crossed the bridge. Given the speed (or lack thereof) of each musician, what sequence of crossings yields the shortest time for the whole band to cross over?



It is clear that **Larry** and **Bono** are the biggest time-sinks in this time-optimization problem; if they cross separately, these two alone will consume a whole **15 minutes** (5+10) of the total journey time. But if they cross together, they will consume only **10 minutes**, neutralizing the effects of **Bono**'s poor choice of footwear. No other pairing of members or constraints can save so much time as this one. Yet it is also clear that **Larry** and **Bono** cannot be the *first* members to cross, as one will have to cross the bridge again to return the torch to the others. So the pairing of **Adam** and **Edge** must be the first to cross, consuming a mere **2 minutes** in the process. If **Adam** then returns alone with the torch (plus **2 minutes**), **Larry** and **Bono** can then cross (plus **10 minutes**), sending **Edge** back with the torch (plus **1 minute**). Finally, **Adam** and **Edge** can cross together for the second time (plus **2 minutes**), yielding a total time of **17 minutes**. Tough constraints can often be neutralized through an understanding of how they relate to others.

So how best might a poet organize the diverse constraints that shape the writing of a poem?

Many interlocking constraints on word choice and syntactic form guide the construction of a poetic text, from those that impose metre on lines to those than enforce rhyme *between* lines to those that ensure the emotional and figurative coherence of the overall text. Though all constraints should work together to shape a composite whole, each may suggest its own search-space and its own means of exploration. Let's zoom in to the level of specific word-choice, to examine two interlocking constraints as applied in the context of the first four lines (below) of *Sonnet **130*** by **William Shakespeare.** This poem is distinctive for its clever subversion of the poetic convention in which one's lovers are flattered by the most sweetly evocative metaphors and similes.

My mistress' eyes are nothing like the ***sun*** a
Coral is far more red than her lips' ***red*** b
If snow be white, why then her breasts are _?_ a
If hairs be wires, black wires grow on her _?__ b

Let us now simplify wildy and suppose that **Shakespeare** has written his first two lines without yet knowing the content of his 3rd and 4th. His ***a-b-a-b*** rhyming scheme dictates that line 3 must end in a rhyme for "**sun**" a, while line 4 must end with a rhyme for "**red**" b. The constraint on textual coherence also requires that line 3 should puncture yet another poetic cliché, but which one? The space of possibilities is huge, yet we know that whatever cliché is chosen, the line must end in a rhyme for "**sun**", and the space of rhymes for this word is smaller and more easily searched. A poet need only generate a range of candidate rhymes – "*run*", "*fun*", "*ton*", "*bun*", "*dun*", "*gun*" – and supply each to his more demanding, if less sharply defined, constraint on meaning coherence. The word "**dun**", which denotes a dull grayish-brown colour that is more often used to describe cows than lovers, may yield the desired poetic subversion, especially if it is contrasted with that clichéd allusion to unsoiled purity, "*snow-white*". The space of rhymes for "**red**" is likewise small and manageable, including such choices as "*bed*", "*fed*", "*said*" and "*led*", though as "**head**" denotes yet another body-part to complement *eyes* and *breasts*, this is the one the poet favors. Rhyming poetry may be more tightly constrained than free verse, yet by linking the constraints on rhyme and content, a poet can use the simple generativity of the former to dramatically shrink the search space of the latter.

Creators often talk of *hunches* and sudden *insights*, yet these rarely yield a complete solution in and of themselves. Rather, these inspired intuitions typically provide us with additional constraints that we gladly impose on ourselves, to cut away large swathes of our potential search-space and to illuminate a promising avenue of attack through what remains. Rhyme is such a constraint, for when used in the right way at the right juncture, it can act as both a filtering test on candidate solutions – returning a simple ***yes/no*** answer, or a ***qualified yes*** for weaker *slant* rhymes, *eye* rhymes and *identity* rhymes – and as a generator of lexical candidates for a specific position in a text. Rhyme generation is a straightforward task even for a computer (if armed with a large pronunciation dictionary and some matching rules) and can be used to supply candidate words to other, more opaque constraints that cannot so easily be used to generate their own.

The history of science offers many cases of *hunches as inspired constraints*. When **Jean-Francois Champollion** deciphered the ***Rosetta stone*** to unlock the mystery of ***Egyptian hieroglyphics***, he succeeded where others faltered by imposing two constraints upon himself: each hieroglyph *must* represent a sound, not an idea; and any sequence of hieroglyphs ensconced in a *cartouche* – an oval-shaped outline – *must* always denote a proper name. Happily shackled by his own constraints, Champollion quickly identified a mapping from the glyphs used in the presumed names of kings and queens to their sounds, and built on these to map other glyph sequences to words in the Greek portion of the Rosetta stone. **Michael Ventris** adopted similar constraints to decipher the **Linear B** script of ancient Crete, by assuming that certain symbol sequences *must* denote Cretan place names. **James Watson** and **Francis Crick** likewise assumed that DNA *must* have a helical structure – indeed, upon seeing x-ray photographs by **Rosalind Franklin** they adopted an even tighter, *double-helical* constraint – and used this new constraint to search a much-reduced space of possible DNA structures using model-building techniques.

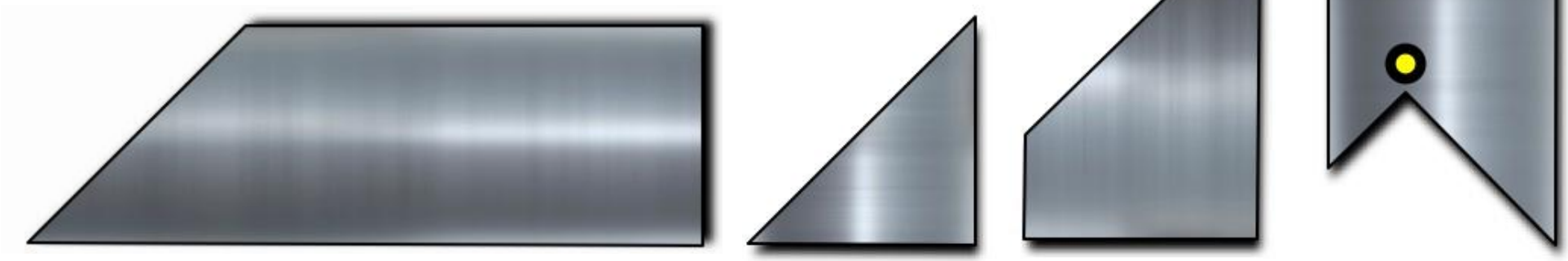
When the British government set up its *Bletchley Park* code-breaking unit in World War II, they recruited their analysts from diverse walks of life, reaching far beyond those with professional cryptographic experience to embrace linguists, mathematicians, skilled chess players and even crossword enthusiasts (otherwise known as *cruciverbalists*). Indeed, recruiters used a proven ability to solve a newspaper crossword in under *12 minutes* as a criterion for selection. Yet recruiters were not looking for a gift for words or even for a large vocabulary, but for a deep appreciation for how constraints might be identified, satisfied, thwarted and exploited.

The Bletchley analysts set to work on breaking coded *U-boat* messages that were encrypted using the **Enigma machine**. Like a mechanical version of **Lewis Carroll**'s ***Doublets***, Enigma uses a series of rotary scramblers to transform an input *plaintext* into an output *ciphertext* by routing each key-press through a series of transformations to yield its coded output. With a fully configured plug-board and an optional 4th scrambler wheel – and when used by a competent operator – Enigma is a very tough nut to crack. Yet Enigma's constraints give rise to some exploitable regularities. As in ***Doublets***, a letter is never transformed into itself in a ciphertext, so e.g. "A" is never turned into "A". Given a candidate plaintext to test against an intercepted ciphertext, Bletchley's cruciverbalists could quickly reject this candidate and move onto another if any of its letters mapped to the very same letter in the ciphertext.

Fortunately, many operators (and their superiors) also did not fully appreciate the interlocking constraints that allow the **Enigma** to create a potentially vast and unmanageable search-space for code-breakers to explore. Much as a formal understanding of the constraints of the **Eight Queens** problem yields a space that is orders of magnitude smaller than a naïve conception, Allied code-breakers and pioneering Polish mathematicians used a formal analysis of Enigma's constraints to show how its search space is considerably diminished when it is used by a complacent enemy.

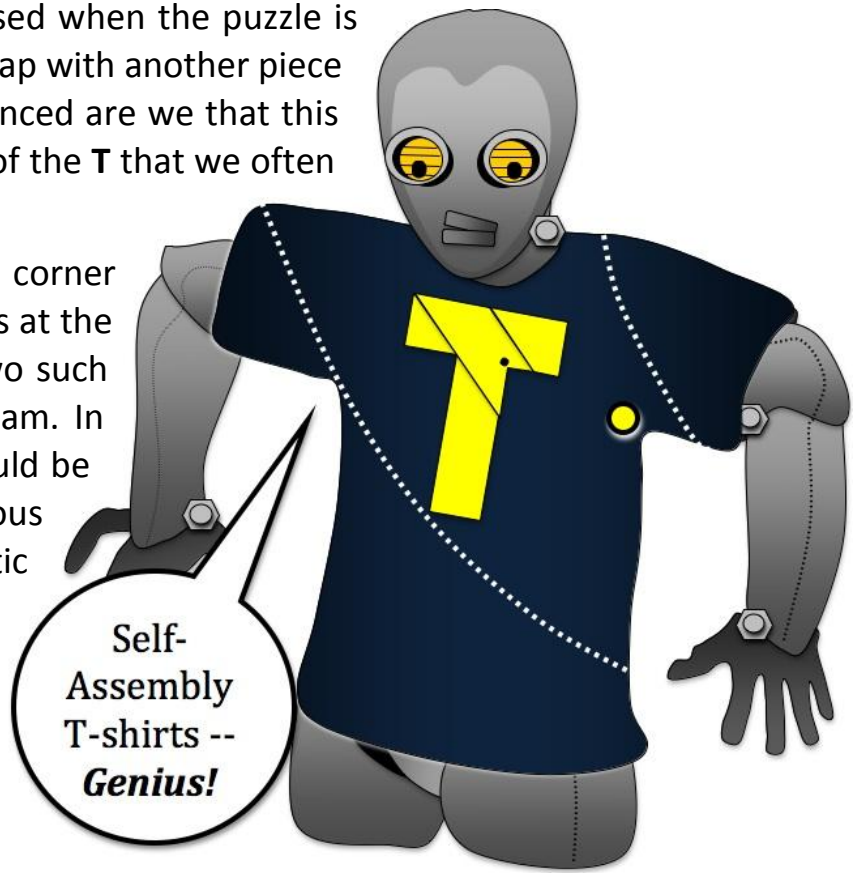
Constraints limit our view of a search-space, yet they do not always limit us to the areas of a space that contain the most viable solutions. Fortunately, the explicit constraints that we bring to the table under the conscious control of **System 2** allow us to see how our view is constrained, and so allow us to react accordingly if a search proves futile. In contrast, the tacit constraints that **System 1** *sneaks* onto the table are often detrimental precisely because we do not see how they subconsciously shape our decisions and keep us from our intended goal.

The cognitive scientist and creativity theorist **Bipin Indurkhya** uses a simple child's puzzle to illustrate the baleful effects of subconscious constraints. The so-called "**T**"**-Puzzle** is typical of the wooden puzzles that one finds in specialty game stores, of the kind in which one must rearrange a selection of puzzle pieces into a specific shape. Naturally enough, one must reassemble the pieces of the **T-Puzzle** into the shape of the capital letter 𝕋. The small size of the puzzle – it contains just four pieces – belies the cognitive challenge it presents to most people. For though the search-space of the puzzle is also relatively small – it is comparable to that of the **8-Queens puzzle** when the latter is shorn of its redundancy – humans often struggle to reconstruct the 𝕋. Indeed, rather than systematically exploring the space of possible piece arrangements, human solvers can become mentally blocked, fall into cognitive ruts, and repeatedly try the same futile combinations without success. The four innocent-seeming pieces of the puzzle have the following shapes and sizes:



So what tacit and *unwanted* constraints do we typically bring to our efforts with the T-Puzzle?

The most insidious concerns how we view the negative space created by the shape of the puzzle's individual pieces. Most people view the triangular "white space" at the bottom of the leftmost piece on the previous page – this white space is crowned by a yellow dot – as a gap that just calls out to be filled. Consequently, we view the corner indicated by this yellow dot as an *internal* corner that will no longer be exposed when the puzzle is properly reassembled. Try as we might, however, we cannot fill this gap with another piece and still meet our goal of reconstructing a **T** shape. Indeed, so convinced are we that this corner is an internal corner rather than an exposed, external corner of the **T** that we often attempt the same failed actions over and over again.



So the solution comes rather quickly after the insight that this corner need not be filled at all. Perhaps we recognize that the yellow dot sits at the apex of a right-angled triangle, and that the final **T** must contain two such angles at the points where its vertical shaft meets its horizontal beam. In many circumstances, our intuitions about this misleading corner would be right on the money, and **System 1** would guide us quickly to the obvious solution. It is in the nature of puzzles, however – and problematic puzzles are the very essence of a creative challenge – that our **System 1** intuitions are wrong as often (or more so) as they are right. Computers suffer from an absence of **System 1**-like intuitions in the many familiar situations that hone these very intuitions in humans. Yet they also benefit from the lack of such intuitions in those situations where they are most likely to lead a human astray.

We cannot ignore such intuitions by *willing* them away. We may have more success, however, by thinking like a computer and deliberately *forcing* ourselves to follow an explicit procedure to fully explore the possibilities of a space.

So our tacit constraints can hem us into unproductive or depleted areas of a search-space, and limit our access to those fertile areas of exploration that lie outside a fixed ***mind-set***. Creativity theorists give the name ***breaking set*** to actions that shift us out of a restrictive viewpoint, but this is easier said than done. For a rigid mind-set is frequently caused by how we express our goals in words, both by what we say *and* what we leave out. Though we often leave unstated those aspects of a problem that seem either obvious or irrelevant, a creative solution may ultimately hinge on a detail that seems too unimportant to write down. So an explicit reckoning of a problem's goals and stakeholders – of the kind AI researchers formulate in their software – can help us to unearth hidden perspectives and constraints in our initial specification of a problem. For added reliability, creativity scholars advocate a multi-pass process of word-driven introspection called the ***problem statement-restatement method***.

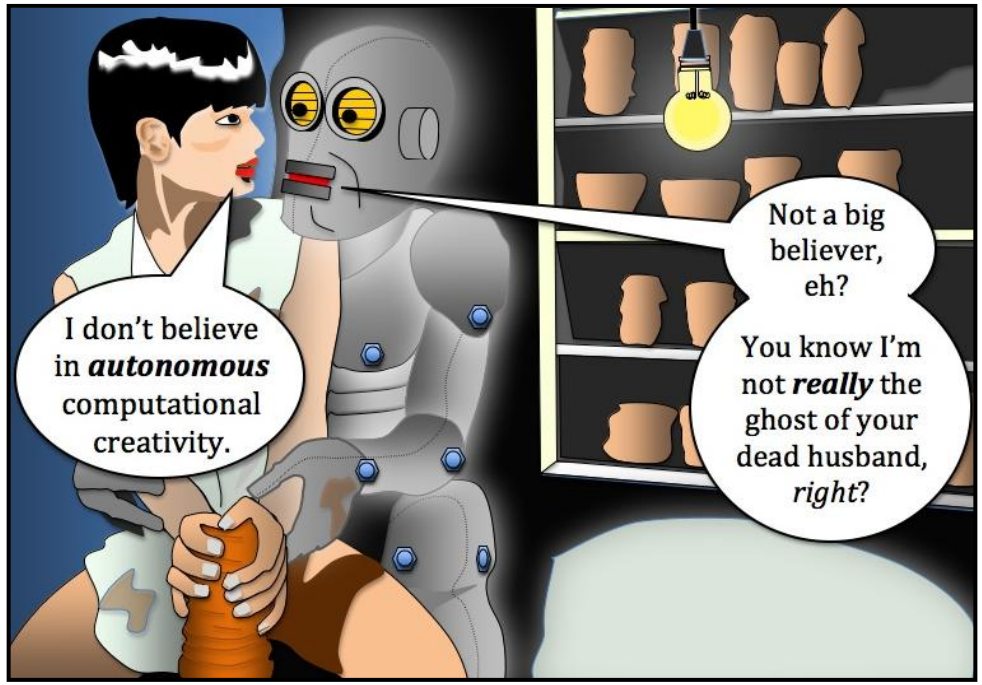
Suppose we are a maker of sportswear and our goal is to design a new ***football boot*** that will delight players and retailers alike. We thus start with a minimal statement of our problem: "*Design a new football boot*". But this says little about our goals for the boot and nothing at all about its users, so we restate it thus: "*Design a new football boot for professional footballer*s". Satisfied, we make multiple passes through our re-statement, emphasizing a different content-word in each pass. We first emphasize "**new**", and ask what innovative features our new boot should possess. We next emphasize "**football**", to ask whether the game makes any special demands of its boots. We then focus on "**professional**", and consider what qualities a professional club would demand of its players and their kit. Finally we emphasize "**footballers**", to ask whether consumers have any sporting goals for themselves that our new boots might enable. Thus, when sportswear-maker **Adidas** asked professional footballers to name the qualities most prized by players of the modern game, most answered "*speed*". This prompted Adidas to focus not on the *ruggedness* of their boots but on *lightness* and *springiness*. By integrating technology from their best running shoes, Adidas designed an ultra-light *blend*-boot that yielded a veritable harvest of goals – the largest of any boot type – in the **2012 World Cup**.

Computers, like kids, often *say the darnedest things*, and for many of the same reasons. For like small children, computers lack the nuanced knowledge of the world that adults acquire the hard way, and are thus prone to over-generalize from the little knowledge they do possess. So like kids, computers puzzle over idioms, take metaphors literally, fail to recognize exceptional cases, and construct nonsensical interpretations of banal events. Though we are charmed by kids who make such mistakes, we are scornful of computers that fall into the same traps. Yet recall that **Dali** would use his ***Paranoiac-Critical method*** to trick his mind into generating nonsense outputs as inputs to his creative process. Conveniently, our computers rarely need to be tricked into producing silly outputs, and provide a ***Paranoiac* ideal** for how one should approach creativity.

Consider how a computer processes everyday language. An idiomatically-challenged computer is likely to interpret the compounds "*memory* ***store***" and "*memory* ***bank***" in a compositionally literal fashion, to imagine in the first instance a shop that buys and sells memories (*human* memories, not memory *chips*), and in the second, a banking business (not an array of storage devices) that lends the precious memories of others to its customers. Though it would take a human with a cruel imagination to interpret the noun-compounds "*baby oil*" and "*virgin oil*" in the same way as "*olive oil*" and "*vegetable oil*", a naïve computer would think nothing of tossing either babies or virgins into the oil press. Computers lack the moral sense, or *squeamishness*, that would prevent most humans from consciously considering these diabolical interpretations, and so AI researchers view such *mis*interpretations as *epic fails*. Yet **Dali** shows us that failures of reason can also be triumphs of the imagination. If it pays to think like an unconstrained child so as to supplant the real with the surreal and the conventional with the refreshingly bizarre, then one can do worse than to think like, or work with, a compositional, literal-minded computer.

Like mismatched heroes in a *buddy-cop* movie, humans and computers make ideal partners

for **collaborative co-creativity**. Ironically, it is the rule-based computer that must play the role of ***maverick*** in this duo, as computers still lack sufficient world-knowledge to obey our many rules and tacit norms of good social behaviour. So it falls to the human to play the role of the ***uptight partner*** who cares about winning the critical approval of others. For humans can productively ***co-create*** with computers if we each assume complementary roles in **Dali's *Paranoiac-Critical method:*** simply, rule-based software is deliberately used in domains where it is out of its depth, to provide ***Paranoiac*** misinterpretations for a human collaborator to ***critique***, embellish and perhaps reformulate in a further cycle of cooperation.



Consider the familiar task of retrieving images from the Web via a search-engine like ***Google***. Our queries will describe the content of the desired images, yet the engine will not directly map our words to its images, as this would require deep image-understanding. Rather, an engine will simply retrieve any Web documents that contain the given content words, and extract any images from those parts of the documents where these words are found. Though shallow, this efficient heuristic is surprisingly effective, and typically returns a buffet of relevant images with just a peppering of irrelevant outliers. Yet it is in these outliers that the seeds of creativity are most likely to be sown. Consider the task of finding images to illustrate an article on ***the Eight Queens problem***. Our query mostly finds relevant images of chess-sets and computers, but also finds outliers such as images of *Henry* ***the Eighth*** (who had ***problems*** with just six ***queens***). If we view these mistaken retrievals as the stuff of creative readymades, we may produce an oblique and non-obvious deviation from the norm.

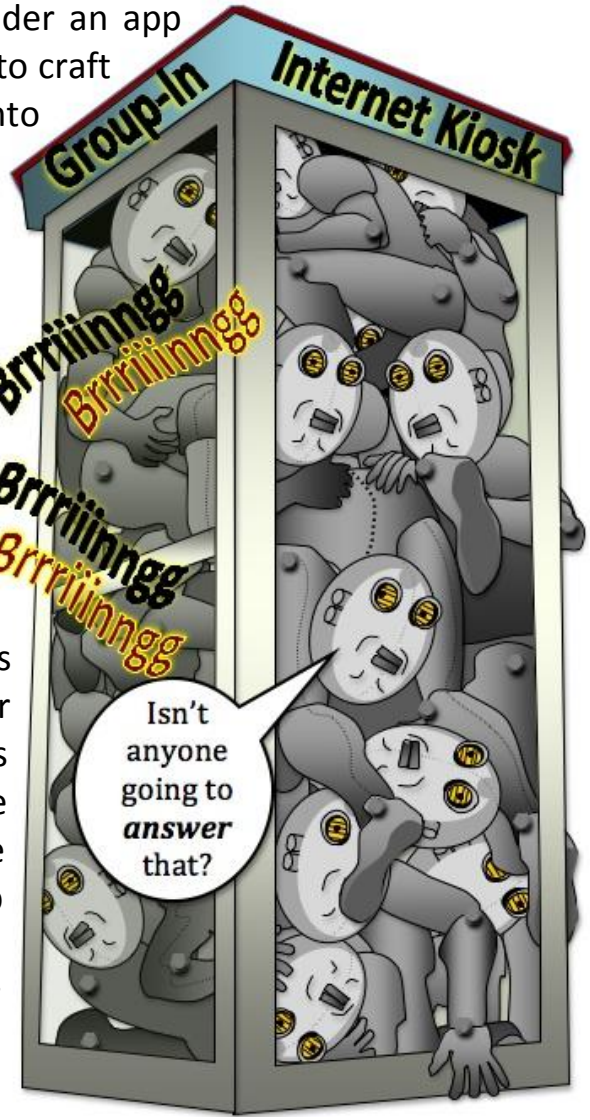
Machines can be active *co-creators* in ways that transcend their propensity for silly mistakes.

While computers have neither jobs nor social lives to inform their creativity or to provide an impetus for originality, neither are they confined to just sitting on a desk in our workplaces. They now accompany us wherever we go, in our smartphones and tablets as well as our music-devices and even our glasses. So as social networks have expanded our *friend* base, so too have they *socialized* our computers. Yet by offering ever more ways to connect with each other, computers also impose an ever greater need to *entertain* each other, with clever photo captions, witty tweets, topical observations and eye-grabbing status-updates. Fortunately, new software may also help to alleviate this new burden. To do so, software must offer more than digitized versions of creaky old reliables like *thesauri*, *grammar books* and *style guides*. Exploiting both the ***Keats heuristic*** and statistical generalizations gleaned from large text corpora, new ***CC-enabled*** apps may alert us to the possibility of creative misreadings of our tweets, or suggest edits that imbue our texts with a greater degree of humour, gravitas, emotion, ambiguity, poeticity, persuasiveness and rhetorical sophistication.

We already face a deluge of information from our daily lives, and it serves no one to build software that merely adds to the noise while further attenuating the signal. Yet our new software may selectively use **CC** techniques to retrieve, modify or cut-up & blend familiar *readymade* expressions such as clichés, proverbs, idioms, film and book titles, song lyrics and quotes. Our software may thus suggest original similes, metaphors, put-downs, puns or blends of trending terms (like "*Fifty shades of Dorian Gray*"), all to project our personalities above the chatter of the Web.

The Web magnifies the utility of a *CC app* by allowing it to tap into the *wisdom of the crowd*.

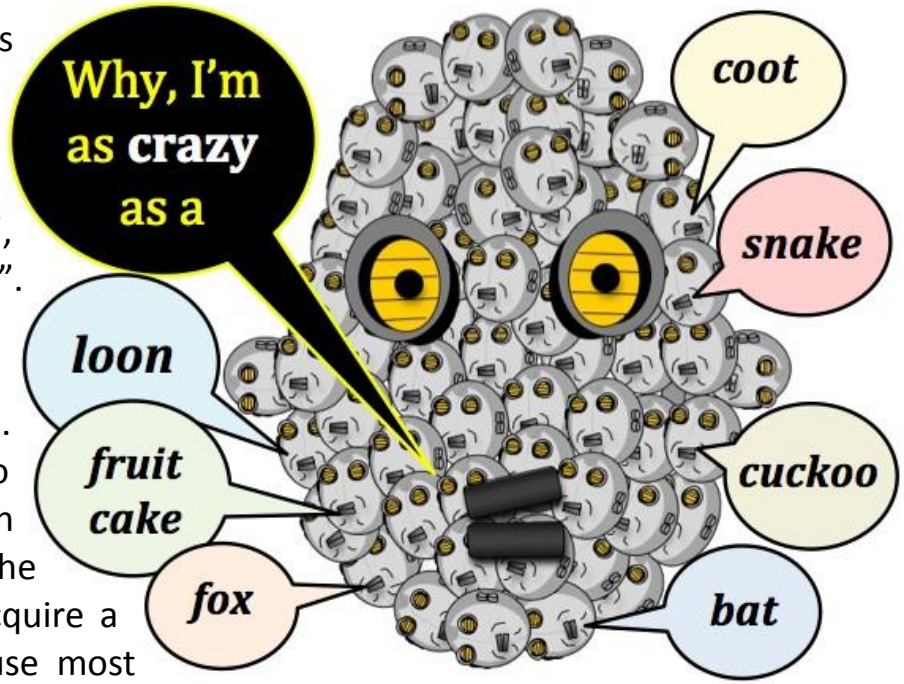
By pooling the usage patterns of its many users across the Web, an app can transcend its initial rule-base to acquire a nuanced and evolving understanding of what constitutes a creative artifact. Consider an app that uses the **Keats heuristic** and an arsenal of rhetorical strategies to help its users to craft witty tweets on their smartphones. The *hashtags* and *emoticons* that users insert into these messages – such a **;^)** and **#ironic** – can yield statistically significant global patterns from which our CC app can derive an up-to-date appreciation of which words and strategies are most often used to convey a particular attitude or affect.



The *wisdom of the crowd* can also be leveraged to guide the construction of a **Zwicky box** for each new domain of innovation. For it is an inescapable irony of such explicitly representational approaches that they require us to deconstruct our knowledge of a domain so that it can be re-assembled into solutions that our mental blocks and ruts might otherwise have hidden from view, yet any arbitrary **Zwicky** representation is just as likely to create blocks and blindspots of its own. Since different thinkers will deconstruct a complex domain in their own ways, using idiosyncratic sets of dimensions and morphemes, one may well generate solutions that are not possible with another's representation. A flawed analysis can limit our perception of a domain's possibilities just as surely as the biases and preconceptions that morphological analysis is designed to overcome. But smart CC apps can take **Zwicky-style** analysis to the *meta*-level, where the object of manipulation is not the problem itself, but our *representation* of its domain. A **Zwicky** app used by many Web users across a diversity of domains can pool its analyses of all those domains, to e.g. learn which morphemes most suggest or inhibit others. With enough experience, our app can suggest more robust representations than the ones we ourselves propose.

The Web is a vast *echo chamber* for a multitude of diverse voices, all talking at the same time.

Yet it is sometimes more useful to view the Web as the collective musings of a *single*, *massively divergent individual*. This imaginary **meta-individual** is not *entirely* divergent, of course, since many of the Web's voices say the same thing when they speak of (or tacitly allude to) our shared stereotypes and cultural beliefs. Conveniently, however, we can always use Web-frequencies to separate this rump of convergent beliefs from the *long tail* of more divergent views.



Consider the use of similes on the Web. Though some writers creatively mint their own comparisons, many more prefer to reuse the common coin of the language. So for every novel simile such as "as ***hot*** *as a* ***tin heat shield during reentry***" we find many more uses of old reliables such as "*as hot as a* ***sauna***", "*as hot as a* ***jungle***" and the ever dependable "*as hot as an* ***oven***". This is to be expected, of course, since convergent viewpoints, by their very nature, are the stuff of boring convention, while divergent viewpoints are far more likely to be creative one-offs. A computer that casts a wide net over the texts of the Web to pull in as many similes as it can will find its net filled with big fish – the most conventional similes – while smaller fry slip through the holes. Yet from the perspective of a machine that aims to acquire a comprehensive picture of the stereotypes that we humans use most often, this is a *good* thing. Because similes have a clear-cut structure – *"as* ***X*** *as* ***Y***" being the most typical form – they are easy to harvest in bulk from the Web. A computer can quickly sift through millions of examples, to extract **X:Y** pairs such as *hot:oven*, *wobbly:jelly* and *dry:desert* by the thousand. We rely on stereotypes to anchor our similes and illuminate our comparisons, and the conventional simile is one of the few linguistic forms to expose the conceptual free associations of **System 1** to the light of day. Once exposed, our machines can use these associations just as freely.

Yet human knowledge is much more than a nebulous collection of disparate facts and beliefs.

It only becomes truly useful – for inference, for categorization, and for extrapolation into the unknown – when its separate strands are woven into a coherent whole. So hidden beneath (and between) the noise and bluster of the Web is a tightly-woven tapestry of common-sense knowledge. The stereotypical associations that our machines extract from Web similes are just the visible ends of this complex tapestry's individual strands. Nonetheless, by systematically tugging at these exposed ends, a machine may reveal enough of the tapestry's warp and weft to knit its own tightly-woven system of common-sense knowledge.

On such systematic technique in **AI** is called ***bootstrapping***: given its existing common-sense knowledge (e.g. from similes), a machine can predict the *shape* of future knowledge it doesn't yet possess. For instance, a simile may tell us that **caviar** is *expensive*, but *expensive what*? The search query "*expensive * such as caviar*" will retrieve all those texts from the Web that provide an answer, since * is a search wildcard that can match any word. A Google Web search thus provides answers via bindings for * such as *item*, *ingredient*, *food* and *luxury*. If a machine now builds on this new knowledge via the Web query "*expensive luxuries such as caviar and **" it will find Web texts indicating that *truffles*, *foie gras* and *French wine* also belong to this exclusive category. By then putting the * in a different place – as in "*** luxuries such as caviar*" – our machine can lateralize its viewpoint, to also place ***caviar*** in the categories *little luxuries*, *pricey luxuries*, *expensive luxuries* and *sophisticated luxuries*. These categories, in turn, are the ends of their own strands in the tapestry, and tugging via the query "*sophisticated luxuries such as caviar*" reveals that *oysters* and *truffles* are in this category too. Bootstrapping on a large scale exploits the connectedness of human knowledge and the corresponding connectedness of human texts, for knowing one piece of knowledge allows a computer to predict the *shape of related knowledge*, using targeted new queries that systematically reveal more of the hidden tapestry.

This *lateral thinking* allows us to hop from category to category and from example to example.

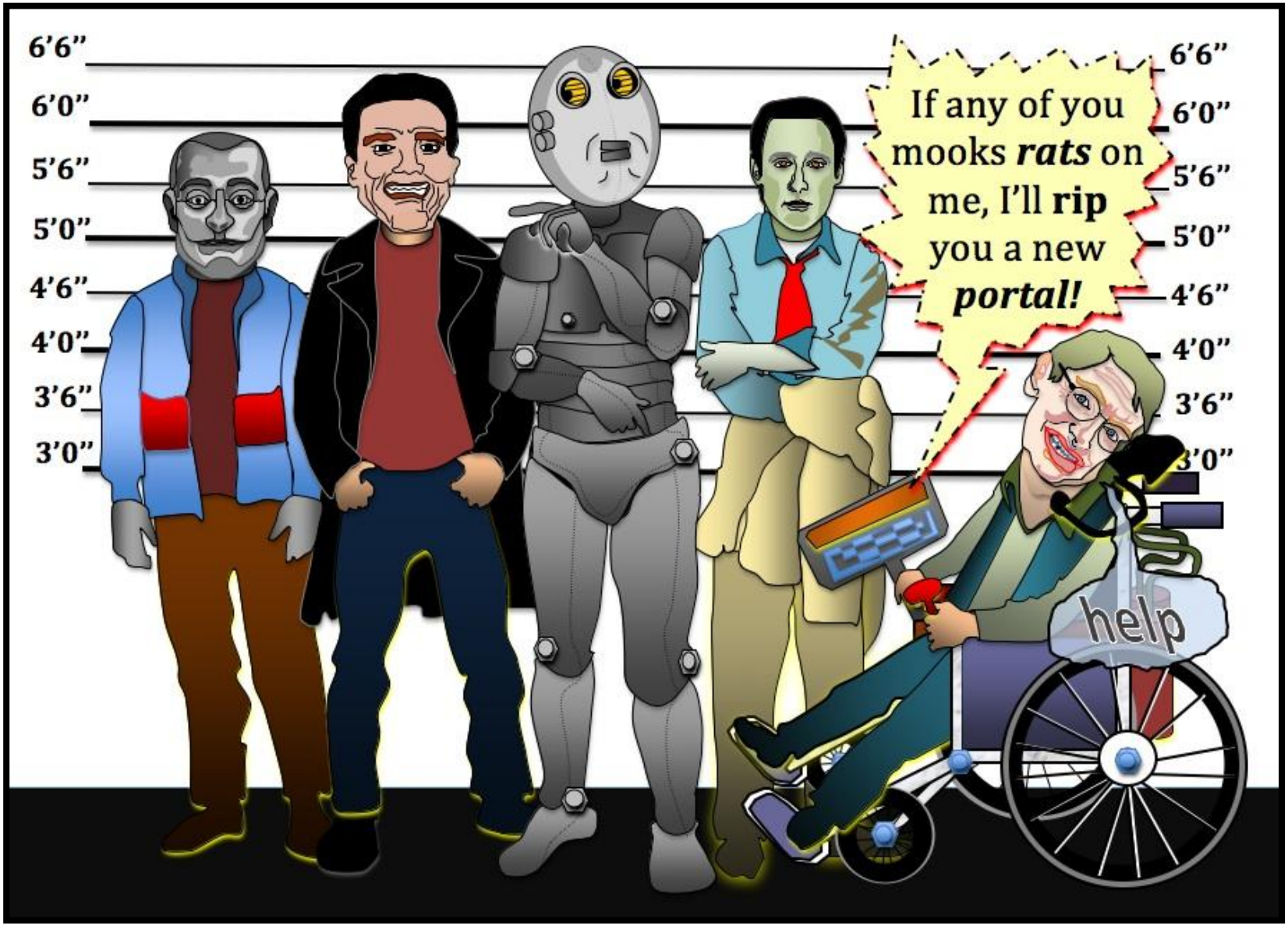
As defined by the creativity writer **Edward de Bono**, *lateral thinking* is a creative complement to the ***vertical thinking*** that encourages us to try and solve a problem by *abstracting* ***upwards*** toward more general categories or by *drilling* ***downwards*** toward more specialized categories. Lateral thinking encourages us to instead cut ***sideways across*** our categories, to hop between competing categories or to shift our focus to other potentially interesting members of the same category. Using the kind of queries from the previous page, we can all use a service like *Google* to guide our lateral explorations of the Web. Yet our computers beat us hands down: by pursuing lateral searches *in parallel*, they can tug on many strands of knowledge at the same time, to reel in a wide array of divergent perspectives on a great many ideas from the Web. The ***word cloud*** on the right shows the categorizations for **cola** that this *lateral-style* bootstrapping extracts from the Web, starting from a handful of Web similes that tell us that **cola** is stereotypically *sweet*, *dark* and *fizzy*.

Notice, for instance, that ***cola*** is sometimes categorized as a *corrosive substance*. This categorization places ***cola*** in the dubious company as *battery acid*, *gasoline*, *pesticide* and *raw sludge*. But it is also a categorization that explains the common if otherwise **unusual use** of ***cola*** to clean pennies and silver cutlery. In contrast, the categorization *stimulating beverage* explains why ***cola***, as a *drink*, is often compared to *strong tea* and *coffee* as well as to *sport* and *energy drinks*. Though a *traditional non-alcoholic beverage*, cola is seen by some as a *forbidden food* with the same *unhealthy* allure of hard liquor and soft candy.

Psychometrics is a modern discipline and the *Unusual Uses Test* did not exist in Aristotle's time.

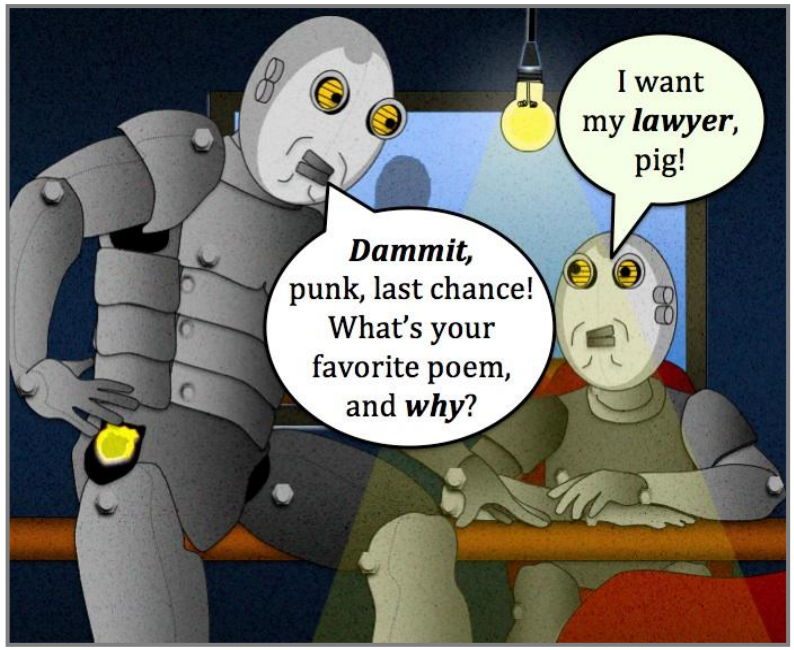
If it had, **Aristotle** might have found it convenient to define **metaphor** as an application of the unusual uses test to words rather than to things. Recall that **Aristotle** defined **metaphor** as the application of the word for a **source idea** (the *vehicle* of a metaphor) to a very different **target idea** (the *topic* described by the metaphor). We are never more convergent in our thinking than when we use words in their conventional dictionary senses, but metaphor, as an application of general creativity to the realm of language, prompts us to divergently seek new uses for familiar words.

Consider the words ***divorce*** and ***war***. Each has its own conventional senses and its own typical domains of use. When circumstances allow, however, each word can be used to metaphorically substitute for the other, not just to achieve **novelty** but to illuminate apt qualities in context. A conflict between two countries previously united by a diplomatic alliance might thus be called a **divorce**, just as the bitter separation of an antagonistic couple might be called a **war** (the movie ***The War of the Roses*** mined this metaphor for every ounce of its comedic value). Though both events operate at very different social scales, they each evoke similar emotions and tend to be categorized in much the same ways by antagonists and by observers, as shown in the word cloud (above) of the ***overlapping*** categorizations that can be extracted for both from the Web.



Creativity on Parade

As we grant increasing autonomy to computers in the execution of *creative* tasks, more than just our view of **CC** will change in the process. Our perspective on human creativity will inevitably evolve also, to the point that we imbue the processes and products of *true* human creativity with an admiration bordering on snobbery. Though our skepticism will diminish with time and experience, we are initially likely to view the products of **CC** with the gimlet eye of one who is about to buy a $1000 Gucci bag yet suspects that the product is a Hong Kong knockoff.



Philosophers antagonistic to the goals of **Strong AI** – a research agenda with the ambitious aim of building computers with human levels of intelligence – have long argued that the systems hawked by AI researchers are second-rate knockoffs rather than the real deal. Yet if a knockoff satisfies each of our tests of quality and craftsmanship and shows no overt signs of fakeness, should we think any less of it merely because it costs a fraction of the price?

This is a question that exercised the founding father of modern AI, the mathematician – and *Bletchley Park* alumnus – **Alan Turing**. In 1950 Turing proposed a thought experiment that is now widely known as the **Turing Test**: to determine whether a computer truly exhibits human-level *smarts*, simply interrogate it via a text-only channel (such a *keyboard + screen*) so that the machine is judged on the quality of its answers rather than on its looks. To make it a fair test, also place an anonymous human or two in the loop, and pose similar questions to each participant. The questions should not focus on areas where computers excel, such as arithmetic calculation, but on the kind of topics that intelligent humans are supposed to enjoy: *what's your favorite Dickens character? Your favorite poem? Was Robert Frost right about free verse*? If, after five minutes, human interrogators cannot distinguish human from machine with a better-than-chance probability, the machine is judged to have passed Turing's version of the ***imitation game*** with sufficient aplomb to be considered intelligent.

But what can Turing's *imitation game* tell us about the assessment of automated creativity?

Turing saw his test as a *thought experiment*, a helpful signpost rather than a research destination in its own right. While it's tempting to reductively view the goal of **Strong AI** as the building of computer systems that can pass the **Turing Test**, this this is little different from claiming that the goal of a good education is to earn top scores on the **SAT**. Researchers who view the test as a primary aim rather than a secondary interest may thus do more harm than good, by incentivizing a focus on the *mere appearance* of intelligence rather than on the understanding of its core mechanisms. Just as the makers of knockoff goods are unlikely to have their customers' best interests at heart, and are driven more by the goal of tricking consumers than of providing a high-quality product at a low price, premature implementations of a *practical* Turing Test – as evidenced by **chatbot** entries to the annual **Loebner Prize** – do not so much drive a *race to the bottom* as a *race to the surface*. **Loebner**'s contest rewards a broad but facile competence for glib chatter, to the detriment of deep understanding and real insight, so much so that the *most human* systems mimic human weaknesses rather than human strengths, aping such superficial signs of our imperfect humanity as frivolity, boredom, scattiness, slow typing and bad spelling.

So would a **Turing Test for CC** be any different from those media stunts in which critics mistakenly attribute the paintings of chimps to human artists? Perhaps, though creativity calls for more than idle prattle from a chatbot. To keep the spirit of **Turing**'s original experiment, critics must not only judge the output of human and machine creativity *side-by-side*; they must also be allowed to *interrogate* each creator about the choices that shaped a particular output.

But not everyone buys into Turing's '*Smart Is as Smart Does*' behavioral view of intelligence.

Perhaps the most ardent critic of his *whatever-works* view is the philosopher **John Searle**, who marshals our intuitions against the enterprise of **Strong AI** with a fiendish thought experiment of his own. His experiment, evocatively named ***The Chinese Room***, asks us to imagine a scenario that is best described as an AI hybrid of the films ***You've Got Mail*** and ***The Manchurian Candidate***. A man wakes up in the locked room of the title, surrounded by books of rules for manipulating inscrutable symbols. It's tempting to think of this poor wretch as a P.O.W. in the Korean war, but any non-speaker of Chinese will do, since Searle's symbols are the characters of the Chinese writing system, which Searle tells us appear as unintelligible "squiggles" to the bemused prisoner. The man's only interaction with the outside world is via a slot, through which slips of paper bearing Chinese characters (and, one hopes, the occasional meal) are passed. The man knows his continued survival requires him to precisely apply the rules to any input slip he receives, to manipulate its symbols accordingly to generate a new sequence of symbols, on a new slip of paper, that will be passed back to the outside world. The incoming slips contain questions written in **Chinese**, and the rules in the books encode a complex but manually-executable algorithm for answering those questions. The rules are intricate and cleverly constructed, so that the slips of paper that are passed back through the slot always contain answers that seem convincing to the person waiting outside.

The Chinese Room provides an *oracular* service of sorts, and to the people who use it, the room may well conceal a learned Chinese scholar. Whoever or whatever is hidden in the room, it looks to these users as though it truly understands Chinese. We, however, are privy to the room's secrets, and know that the wretched man who is trapped inside understands not one word of Chinese. Not only does he not understand the requests that come in via the slot, neither does he understand the well-formed Chinese replies that he sends back out again. The poor soul must sing for his supper, but he neither understands his song or its purpose.

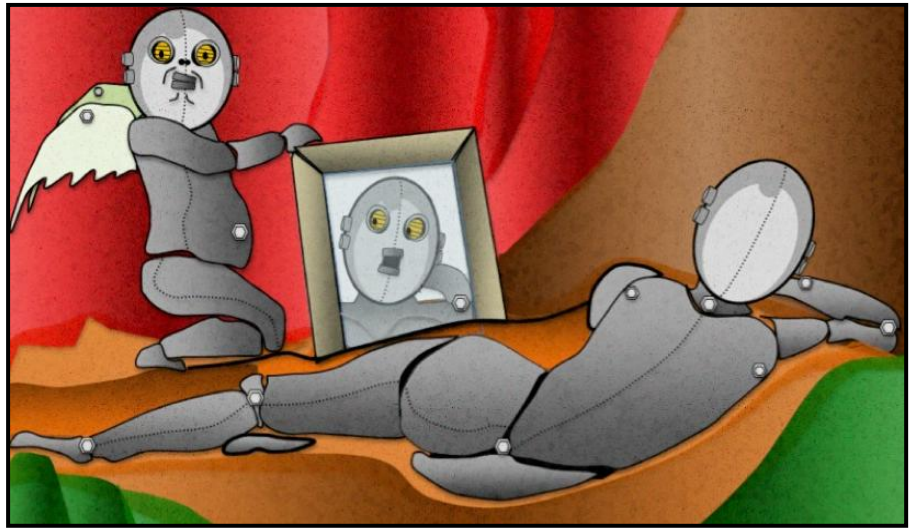
Yet the man and his rules do *seem capable* of producing valid Chinese replies to Chinese inputs.

However, since it hardly seems possible to separate a man from his language abilities in this contrived way, **Searle** asks us to apply our intuitions to a situation that is anything but intuitive. AI proponents thus respond to Searle's experiment with a variety of counter-arguments, but the ***System response***, as it is known, offers one of the firmest rebuttals. The **System response** agrees with Searle that the man alone does not really understand Chinese. But it also argues that the integrated system – **man *plus* rules** – clearly *does* understand Chinese, as shown by its ability to validly answer real Chinese questions. So to focus on the man alone is a sleight of hand of the *bait and switch* variety, and is no more sensible than asking if the eye alone can understand a painting, or the ear alone a poem, or a single neuron a language.

Searle offers an ingenious riposte to the **System response**. Imagine, he says, that the man is forced to *memorize* all of his rules, to the point where he no longer needs to refer to his rule books when generating Chinese outputs. The system is now reduced to the man alone, since memorization weaves the rules into his mental fabric. Yet he still does not truly *understand* Chinese, so the integrated system of man *plus* internalized rules is still just a fake. This bravura retort might aptly be called the ***System 2 response***, as Searle assumes that memorization of the rule books allows the man's **System 2** processes to consciously execute the **Chinese room** *algorithm* by bringing to mind the necessary rules for each symbol. But the man continues to engage with his rules at a deliberative, conscious level, and he still has no more understanding of – or **System 1** associations for – each symbolic squiggle than he did when he relied upon his books as an external memory store. In computational terms, the unfortunate drudge has been "upgraded", from an old-fashioned punch-card system to a shiny new version with a large on-board program memory. Yet no matter how practiced the man becomes in the execution of his memorized rules, his processes all still occur at the **System 2** level.

"In show business, the most important thing is sincerity. If you can fake that, you've got it made."

This quote from comedian **George Burns** could just as well apply to the showmanship of the **Turing Test** and the **Chinese Room** as it does to the artifice of the film industry. Yet Burns' remark seems less of a contradiction when one considers how actors actually meet the challenge of faking sincerity. As practiced by acclaimed actors such as **Marlon Brando** and **Daniel Day-Lewis**, the ***Method*** urges actors to throw themselves fully into a role, to adopt the clothing, mannerisms and passions of each new persona, both on and *off* the stage. To lend sincerity to their on-screen emotions, method actors anchor their interpretations of a character's inner-states to the circumstances of their own lives, so that when they cry with rage or laugh with delight on screen, their anger and joy is sincerely felt. ***The Method*** trains actors to engage **System 1** as well as **System 2** when they fill a role, by conditioning them to share many of the same automatic responses as the characters they play.



Turing's argument is that a computer can pass the **Turing Test** only when it *loses itself* in the role of an intelligent agent. **Searle** may object that a computer's portrayal is never truly sincere, and that a *method actor* is always just an actor. Yet a method actor's responses *are* sincere, up to a point: rather than just reciting a memorized script, method actors reflect on their own motivations to make themselves feel the same feelings and automatically prime the same associations on cue. Their responses are scripted while being entirely their own, and while they are rarely self-conscious in a role, they are always self-*aware*, constantly modulating their performance so that it remains believable. A creative system must be similarly self-aware when it plays the role of creator, always asking itself "*what is my motivation in this scene?*" so that it can convince others that this motivation is sincerely felt. A creative system, like a good actor, must always present this motivation – sincerely held or not – as its own, rather than that of a scriptwriter.

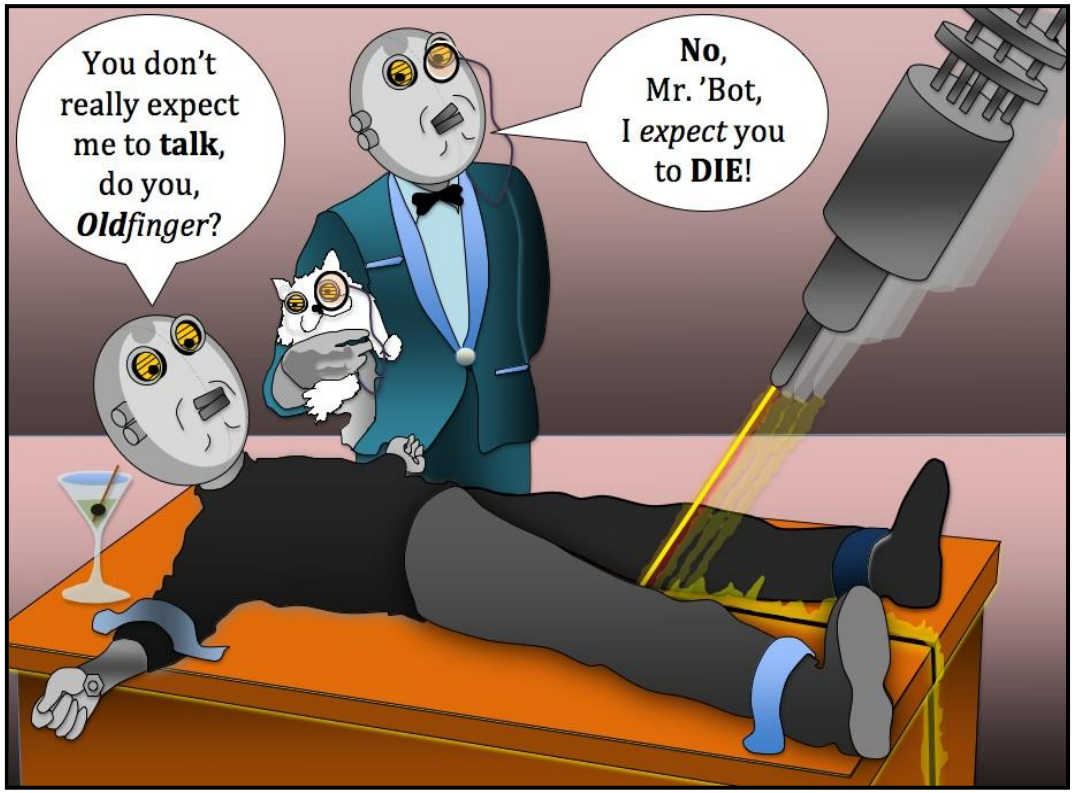
Every great work of art has a great *back story*, such as a revealing anecdote about its inception.

For we enjoy hearing and telling stories about creativity almost as much as we enjoy the actual products of creativity, and are more easily convinced of the creative merits of a work if it comes bundled with a compelling narrative. A good story can elevate a work to the big leagues of creativity, even if it too is wholly a product of the creator's imagination.

Leonardo's ***Mona Lisa*** is constantly name-checked as the world's most famous painting, and though it has always been greatly admired for its technical virtuosity, the painting's elevation to the pinnacle of art history is a relatively recent development, one that owes as much to a story with lashings of international intrigue as it does to any artistic considerations. The painting earned its permanent place in the public consciousness with its theft from the **Louvre** in 1911, not by sophisticated art thieves but by an Italian janitor who hid the painting in a closet before smuggling it home under his coat. At one stage even **Picasso** was *briefly* considered an unlikely suspect. Yet the thief was driven not by jealousy but by patriotic intentions – he wanted to repatriate the Italian masterpiece back to Italy, where he felt it truly belonged – and was caught while trying to sell the painting to the **Uffizi** in **Florence**. The publicity surrounding its sudden loss and joyous return served to cement the painting in the affections of many who had never before given it a second thought.

The backstory of a modern artwork – the story of the work's genesis, its themes and its influences – is often concocted by the artists themselves. These creation myths are not the most reliable of narratives, and may be as artfully constructed as the works they attach to. Each backstory offers yet another vehicle for linguistic creativity – Picasso's "*lie that tells the truth*" – to complement the creativity of the artwork it embellishes. So our appreciation of Coleridge's poem ***Kubla Khan***, to take just one famous example, is greatly enriched by the somewhat *polished and streamlined* backstory he offers of its genesis in an opium dream, which also ruefully tells of how Coleridge was stopped from committing the whole of the dream to paper by a caller from **Porlock**.

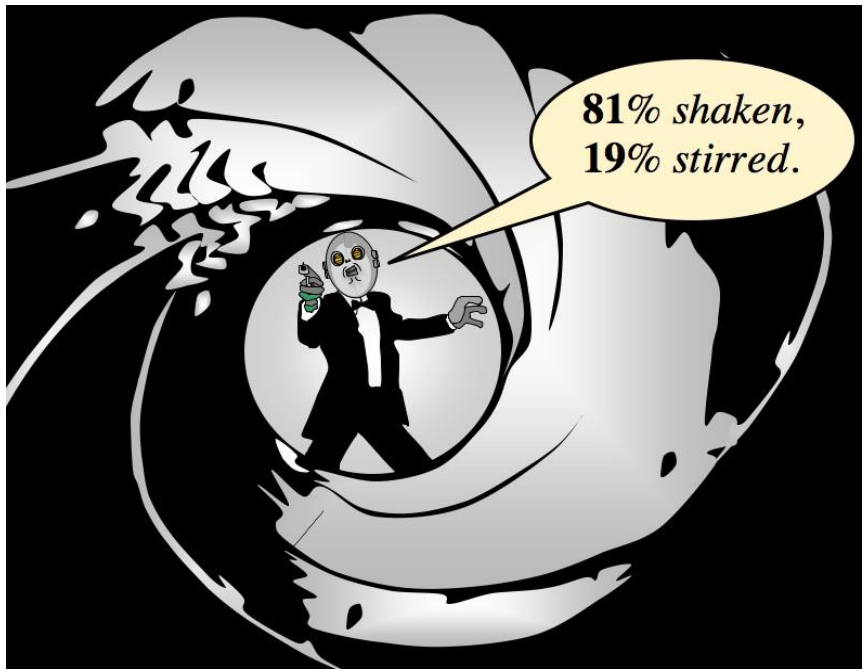
Part *anecdote*, part *explanation* and part *mystification*, a good backstory pulls into sharp relief the conceptual conceits that unpin a creative product. Given the significance of backstories in human art, computer scientist **Simon Colton** argues that computers should also invent backstories for their own generative efforts with **CC**, so they too might be appreciated as ***real art***. By articulating an understanding of the choices that shape its outputs, and of how a result could have been so very different if alternate decisions were made, a computer can truly claim to *own* its outputs. Colton describes the linguistic invention of a compelling backstory as the ***framing*** of a computer-created work. A system for generating paintings, music or poetry might thus work *cheek by jowl* with another system that tailors creative narratives to its outputs. Separate systems with complementary generative abilities can then work together, to collectively overcome the perception of ***mere generation*** that attaches to individual systems that operate in isolation. Our **CC** systems will be taken more seriously if they can talk about *what* they do, and *why* they do it.



So as **CC** ushers in a new breed of creativity-enabled apps, expect our software to be more talkative about the ***why***, ***what*** and ***how*** of its actions. Yet the history of chatty software is not an illustrious one, and few users have fond memories of Microsoft's intrusive talking paperclip, ***Clippy***, or its precursor, ***Bob***. Our CC systems should not overload users with unsolicited asides, but only explain themselves as needed while still preserving a certain aura of mystery.

Estimating the quality of any system that uses *convergent reasoning* is relatively clear-cut.

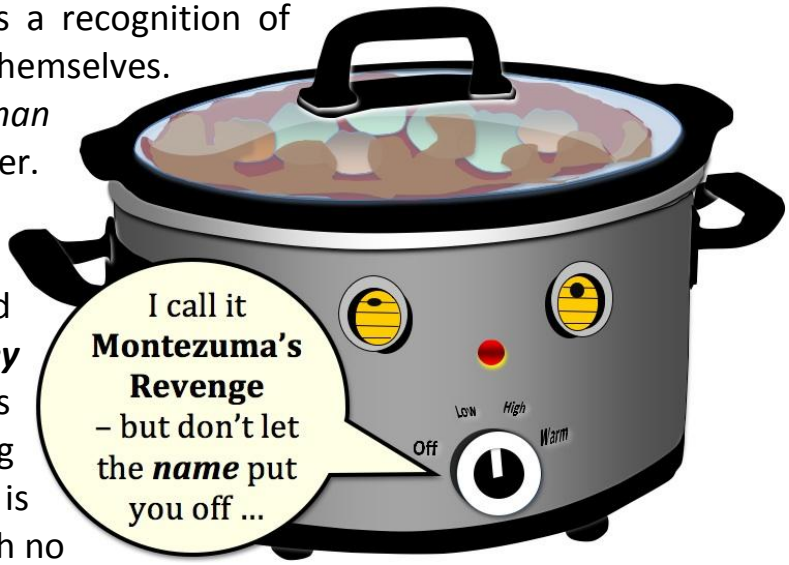
For convergent thinkers see a problem in terms of answers or solutions that are either ***right*** or ***wrong*** but never both at the same time. Convergence thus eliminates the grey-area of ***maybe*** answers, making it easy to tell if we have found the expected ***right*** answer and to conveniently dismiss any other answer as ***wrong***. In a standard college test, most multiple-choice questions or problems that require mental calculation call for convergent reasoning, while essay questions call for – and allow – greater divergence, as one is not simply expected to reproduce a predetermined answer. Essay questions have traditionally been marked then by experienced graders who, like psychometricians administering an **unusual uses test**, draw on their knowledge of a field to consider the ***fluency***, ***flexibility***, ***elaboration*** and ***originality*** of the ideas on offer. If multi-dimensionality is a *boon* when evaluating a CC system, it is an absolute *must* when evaluating two systems side by side.



Comparing two CC systems is rather like comparing two Chinese buffets: we have to consider not only the overall cost, but the number of distinct offerings, the quality and freshness of each, the number of categories they fall into, and the number of unique offerings that are to be found nowhere else, to say nothing of the rate of replenishment. Since creativity calls for risky experimentation – creativity without risks in which "*everyone's a winner!*" is just **mere generation** with an especially reliable formula – we also have to count the number of **failures**, as seen in the number of customer complaints and the number of untouched or rejected offerings. While we should be concerned by any system with a high failure rate, a very low (or *zero!*) failure rate should also be cause for concern. The motto of a *good* CC system should thus be "***Failure IS an option***", while a *great* CC system will actually *learn* from its own mistakes.

Our tolerance for artifice in art is nurtured by continuous exposure to *human* art and creativity.

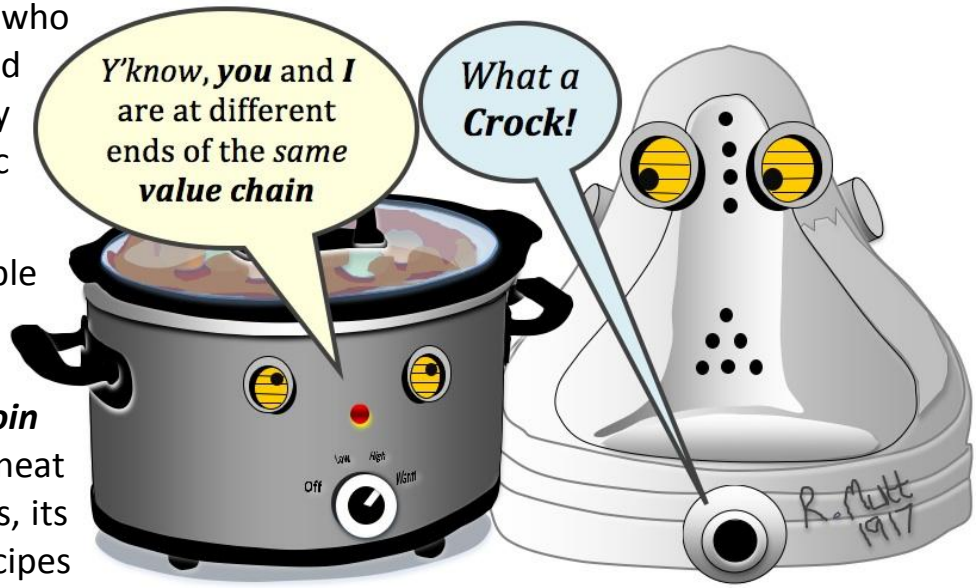
But our standards – both community and personal – are not yet so well-developed when it comes to the creative merits of *machine* artifice. The recognition of creativity first requires a recognition of creative intent, which is closely related to our appreciation of creators themselves. We are thus more likely to see creativity in artifice if we can project *human* qualities, such as *skillful*, *thoughtful*, *playful* or *insightful*, onto an artificer. Naturally, this projection is more easily achieved with other humans, who display a whole panoply of human idiosyncrasies, than it is with cold grey machines. In real life, we are not quite so willing to suspend disbelief as when we are watching a movie like ***2001: A Space Odyssey*** (with its paranoid computer **HAL 9000**, who creatively lied to its users and eventually plotted their murders) or ***Star Wars*** (with its bickering anthropomorphic droids). A computer with human-like appendages is thus more likely to impress as a creative artificer than a sleek laptop with no human trimmings at all. When **Harold Cohen**'s famous painting program **Aaron** was attached to a robot arm, and programmed to paint original pictures with real paint on real canvases, audiences were most visibly impressed by the "*Aw, shucks!*" way that Aaron washed its brushes between color changes. As demonstrated by the **Keats heuristic**, superficial appearances go a long way toward convincing us of the presence of deeper qualities like creative intent.



A fascinating case-study from the **CC** literature nicely illustrates this bias. **Dan Ventura** and his students at Brigham Young University have designed a machine-learning program called ***Pierre*** to create novel ***crockpot*** recipes. In this curious mode of American cuisine, all of a dish's ingredients are placed into a single pot, where they simmer together until cooked. Most crockpot recipes thus have a nicely uncomplicated structure: simply, a list of ingredients and their precise quantities, from which a system like ***Pierre*** can build a probabilistic model of each dish's typical composition. But what is the probability than human tasters will judge **Pierre**'s concoctions in the same way as a *human* creation?

Pierre learns by generalizing over a large stock of *star-rated* crockpot recipes from the Web.

Using a form of *supervised learning* over this user-annotated training data, ***Pierre*** develops its own recipe-evaluation metric to approximate the tacit metrics of the gourmands who rate recipes on sites like **FoodNetwork.com**. A well-trained **CC** system like ***Pierre*** is thus capable of generating many new recipes, but only the best – according to the metric that ***Pierre*** learns for itself – are ever shown to its users.



To help it generalize, ***Pierre*** is also equipped with a simple ontology of ingredients, in which high-level categories like *meats, vegetables* and *fruits* are progressively sub-divided into more specific recipe elements such as ***sirloin*** (→*beef*→*meat*) and ***cabbage*** (→*green*→*vegetable*). In a neat twist, though ***Pierre*** is exposed to diverse crockpot recipes, its evaluation metric is specifically trained on a set of Web recipes and ratings for ***chili dishes*** only. In this way, ***Pierre*** is trained to create its own chili recipes, but is still free to use ingredients from other crockpot dishes that one may not find in a run-of-the-mill chili. ***Pierre*** thus aims to be as inventive as a *human* chili chef, and doesn't simply learn to regurgitate the chili recipes it has already seen.

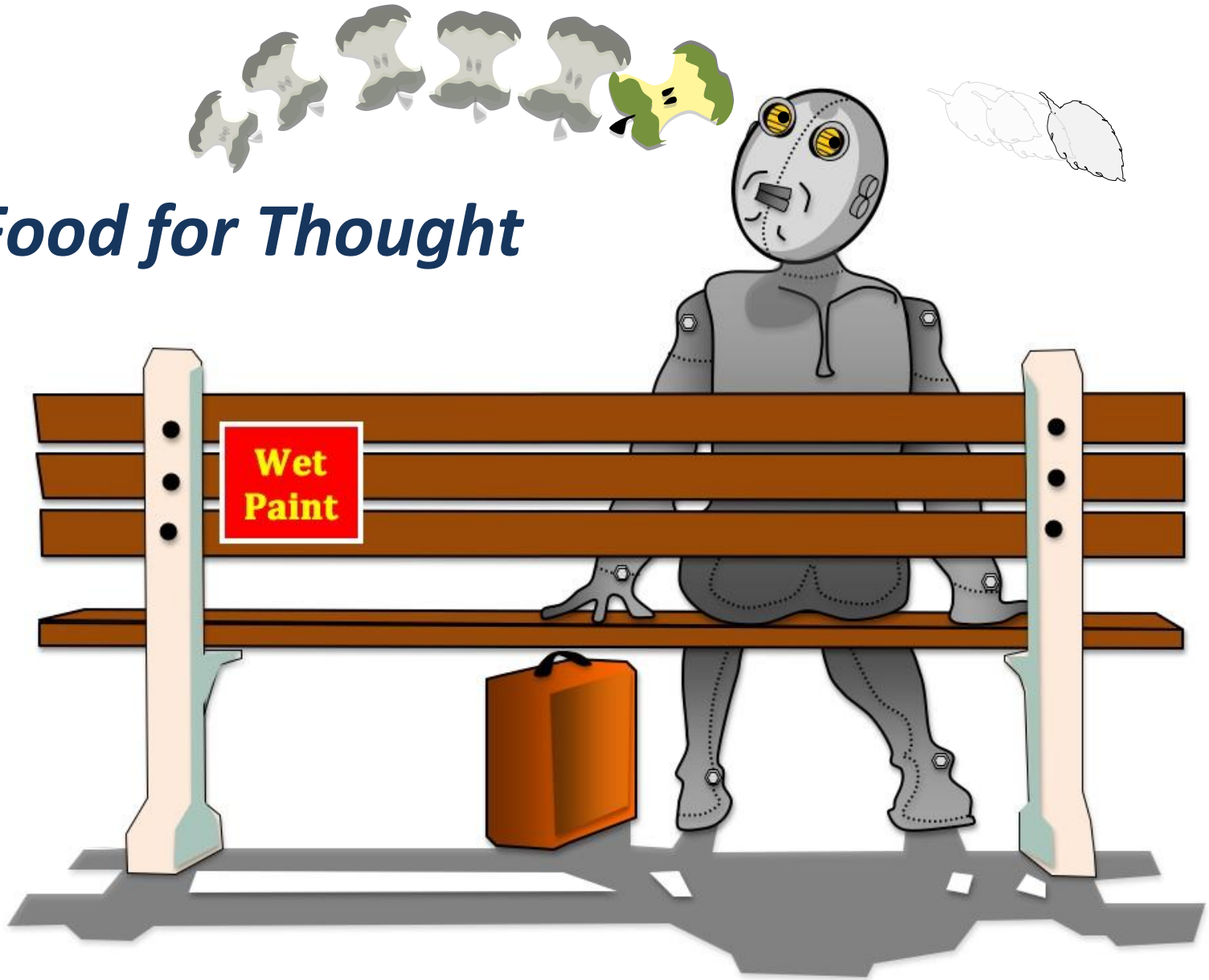
Though lacking taste buds of its own, ***Pierre*** uses its metric to generate its own ratings and rank its own culinary creations, as recipe *texts* rather than as actual dishes. In a sense, Pierre uses the ***Machine-Learning*** equivalent of the surrealist's ***cut-up technique*** to generate novel recipes that are coherent cut-ups of those it has previously learned, though ***Pierre*** and systems like it are guided more by statistics than by chance alone. Even ***Pierre***'s lack of taste buds is not so crushing a deficiency as one might suspect, since even professional recipes in glossy cookbooks are designed to appeal primarily to the reader's imagination. Just as **Beethoven** composed his magnificent **9th symphony** when deaf, a diabetic chef can produce the most mouth-watering of sugary desserts that are destined to never pass his own lips.

Yet *Pierre*'s most telling deficiency is its lack of physical experience in preparing real meals.

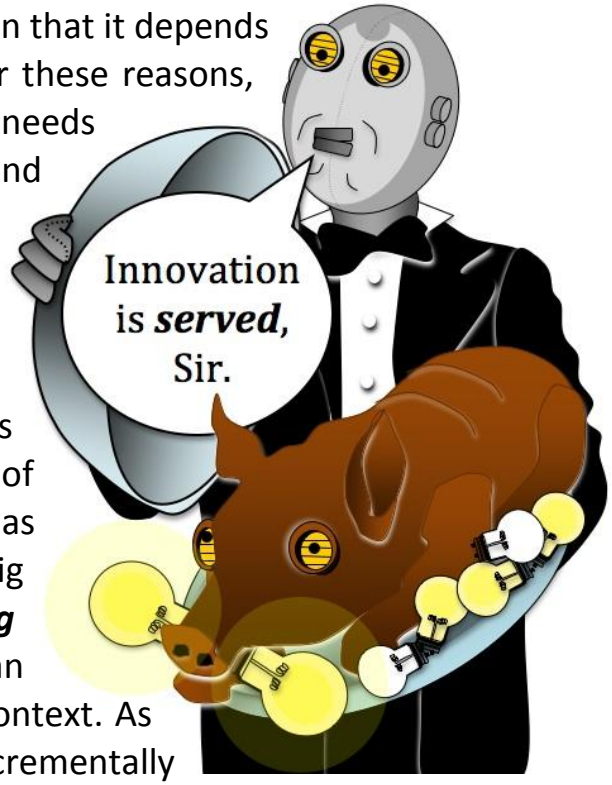
For ***Pierre*** approaches cookery with the precision of an engineer, and often generates recipes that specify unrealistic measures. One of its chili recipes calls for **0.26** cups of chopped onions, while another calls for just a ***dash*** of green beans. Real experience counts, since the practical execution of a conceptual plan can be the hardest part of realizing any creative result. When filming ***Star Wars*** for writer/director **George Lucas**, the actor **Harrison Ford** offered a pungent opinion on the film's script that deserves to be held up as a maxim for the ages in **Computational Creativity**: "*George, you can type this shit but you can't* **say** *it*." An output that looks good on the page often fails to sound right on the screen, taste right on the tongue, or sit well in the belly. For so much of creativity, the proof of the pudding really is in the eating.

Yet when **Ventura** uploaded ***Pierre***'s top-rated recipes to ***FoodNetwork*.com** to be judged by human foodies, the reaction was more positive than negative, though some feedback did reflect a deep suspicion of the author's motives. Not realizing that ***Pierre*'s** oddly precise recipes were computer-generated – **Ventura** chose not to reveal this fact lest it bias the site's reviewers – some commenters felt that their obvious quirks were knowing parodies of the ***Food Network*** itself and of its users' fussbudget ways. If only, for ***Pierre*** lacks the experience to recognize the humorous potential of its own lack of practical knowledge. Yet ***Pierre***'s mistakes do suggest that a **CC** system's knowledge of its own limitations, if formalized in the right way, may allow it to be intentionally humorous. A system like ***Pierre*** can learn from its limitations by comparing its own outputs to the inputs on which it was trained. Like the new kid in school, glancing nervously about the schoolyard to see why all the other kids are laughing, future ***Pierre***-like systems may be capable of greater introspection, to ask "*why are my measurements so much more precise than those of other chefs?*" Such an improvement to ***Pierre*** would require multiple kinds of knowledge, and a far more sophisticated evaluation metric. This metric should not just reject poor outputs, but look for underlying reasons to explain **why** they are poor, so as to allow ***Pierre*** to find a very different kind of creative value in its own mistakes.

Food for Thought



Creativity is an elusive quality that large organizations put tremendous effort and resources into fostering, rewarding and reproducing on demand. But the systematic harnessing of creativity is complicated by the complex and definition-defying nature of the phenomenon, and by the realization that it depends crucially on so many different personal, social, cultural and contextual factors. For these reasons, even famously innovative companies like ***Apple*** will often *out-source* their creative needs to external agencies with a proven track record in the generation, combination and framing of creative ideas. For such agencies are not so much ***problem solvers*** as ***option providers,*** leaving the ultimate responsibility for choosing amongst a diverse palette of new options to the client. So to out-source in this way is not to abdicate creative responsibility, but to widen the range of options that one can choose from.



As we have seen, creativity does not arise from the simple application of rules alone, or even of ***meta-rules,*** but from an active *engagement* with the possibilities of a conceptual space. Our explorations of these spaces often yield new insights, such as that viewing a familiar concept in a novel and unconventional way can deliver big dividends in the form of new uses and new possibilities. Creativity is also a ***learning process***, and so, from the feedback given by a client or an audience, creators can refine their aesthetic sense of what works and what doesn't work in a particular context. As creative agencies accumulate experience across successive commissions, they incrementally refine their aesthetic filters to help them present only the most promising options to a client. The coming **CC apps** will likewise learn from their actions over time, to fine-tune their own aesthetic filters on the diverse commissions they are tasked with. Moreover, Web-based CC apps will leverage the sheer scale of the Web to learn from the diverse commissions they receive from many *different* users. **Selectivity** is the silent partner in creativity that keeps its showier partner **generativity** in check. A creative service, whether a cadre of humans or a software app, is an ***option provider***, and a service that overwhelms with a deluge of raw options is no better than one that offers nothing at all.

Humans employ a mix of convergent and divergent thinking styles, but it's a safe bet that those we laud as clever and innovative will make greater use of divergence than those we criticize as conservative plodders. **CC** software applications can help us all to tilt the balance from safe convergence to exploratory divergence by giving a conveniently automatized and algorithmic form to the kind of strategies for creative problem-solving that have long been advocated by scholars such as **J. P. Guilford** and **E. P. Torrance**, and by popular writers such as **Edward de Bono**.

But what if the most avid users of these new **CC tools** turn out to be *other computer programs*? After all, few people are more convergent in their approach, or more in need of tools to support divergent thinking, than the bloated software systems that now fill our computers' hard-drives. Indeed, large software systems have many similarities to the large organizations that now outsource their creative needs to outside agencies. Each must be well-defined, operate in a mostly predictable fashion, and facilitate an efficient, hierarchical flow of information between its layers. Like large companies, software must continuously engage its users and react with grace and agility when faced with unexpected situations. So imagine if our software could also outsource its creative needs to an external **Web service** with a proven track record in computational creativity. This service would not be a team of creative professionals, but a suite of tools on the Web that offer, as needed, the processes and knowledge representations that are known to facilitate creative thinking. Large software systems, like large organizations, could still maintain their disciplined information-flows, and would appeal to outside creative services only when they needed to diversify their palette of available options.

To maximize flexibility and the potential for reuse, these external services should not try to be *one-stop* shops that shepherd a commission from inception to completion. Rather, they should form a marketplace of diverse – and perhaps even *competing* – services and layers for the generation, filtering, blending and framing of novel ideas. It would then fall to a client system to orchestrate the interaction of its chosen services, to produce a stream of tailored ideas for the client system to select, reject or refine.

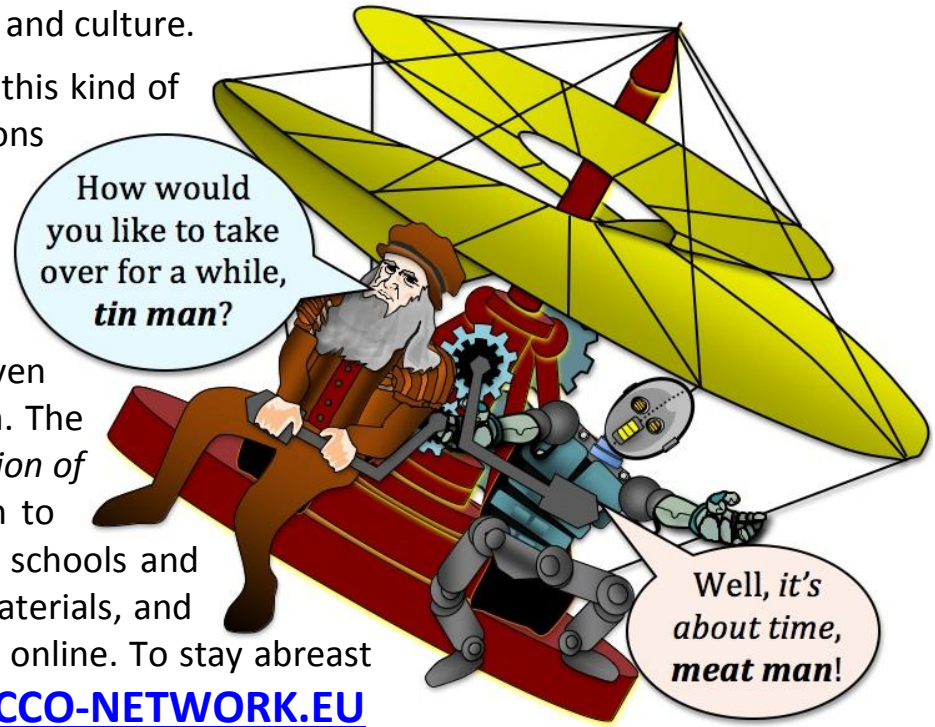
Creativity is like a box of chocolates: you just never know what you are going to get next.

These wise words of **Mrs. Gump** apply not just to the actual products of creativity but to our collective view of the phenomenon itself. For creativity thrives on complacency. Every time we think we've got it *boxed in* and *figured out*, it wriggles free from our definitions and transforms itself into something new. While computers still have much to prove when it comes to the autonomous, self-critiquing generation of original artifacts and ideas, those who remain fixed in their views of machines should remember that creativity itself is anything but fixed. Considering the increasing pace of technological innovation, and the grace with which we adapt to new developments that can dramatically alter our relationships with machines and with each other, it only takes one wildly popular app to change the way we conceive of the creative capabilities of computers. To be sure, this coming breed of **CC apps** must quickly evolve and diversify if the products of **CC** are not to be stigmatized – like those of that one-trick pony ***The Sphinx*** – as predictable objects of ridicule. Yet no one should underestimate our desire for novelty, or for new modes of stimulation and interaction.

When **Ventura** field-tested ***Pierre***'s top-ranked chili recipe at his church's **Chili Cook-Off**, he noted a surprising willingness amongst church-goers to sample the culinary inventions of a mere machine. Indeed, many were charmed by the very idea, and gave high marks to the food itself. Yet when prizes were awarded for best recipe, ***Pierre*** was beaten into 2nd place by a pair of even more charming rapscallions who had done nothing more than scrawl the words "*Secret Recipe*" on a can of store-bought chili. All creativity involves artifice, both deep *and* shallow, and this artifice can take many unpredictable forms. As **Magritte** reminds us with his picture of a pipe that isn't really a pipe, creativity plays with the malleability of forms and labels to suggest novel perspectives on the truth. We cannot eat a *recipe* for chili, or salve a wound with a *poem*, or dance with a *statue* of a ballerina. Yet such forms can stimulate our appetites, provide a mental palliative for our pain, tickle our funny bones and stir our passions. Though creativity is normally the preserve of human beings rather than machines, creativity has no respect at all for such norms. As wise **Mrs. Gump** tells us, **What's normal anyways?**

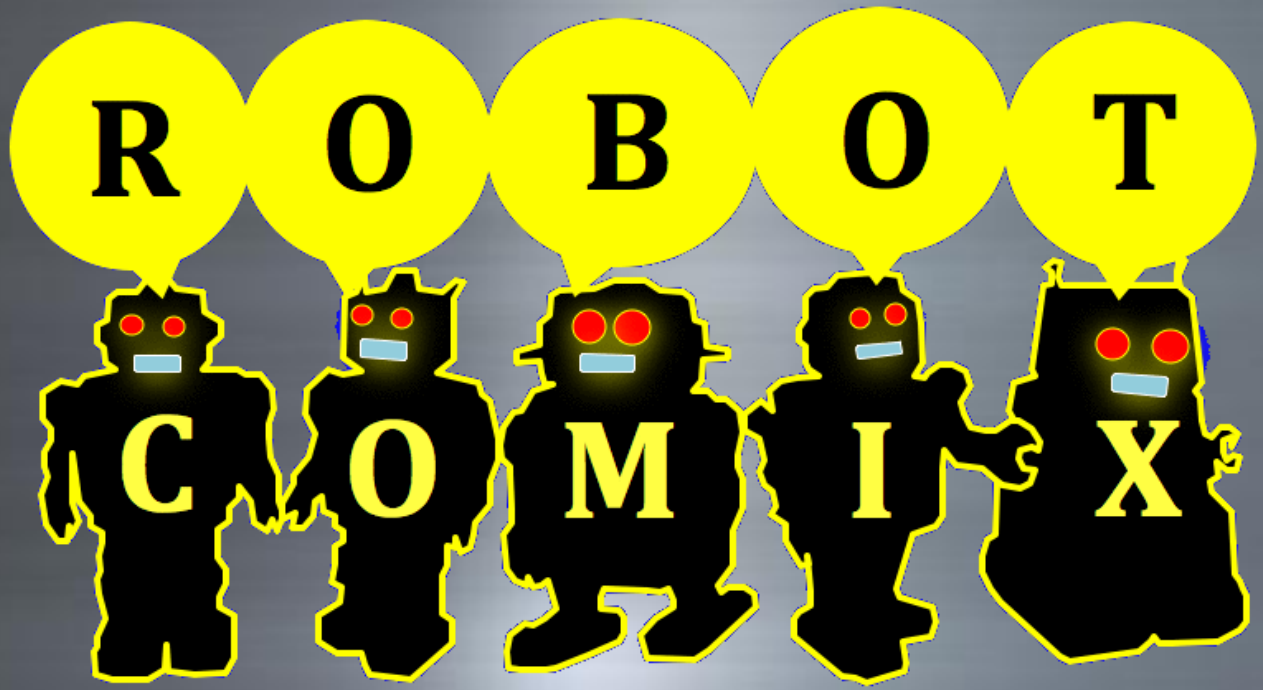
Computational Creativity is no longer the branch of AI research that dare not say its name.

Though long considered the indulgent side of **Artificial Intelligence** that focuses on problems that few people consider solvable and that fewer actually *want* solved, researchers and funding agencies have now warmed to the view that **creativity** – whether eminent and showy or low-key and largely unnoticed – sits at the center of all that we do as humans. Our everyday lives are filled with tacit instances of the **unusual uses test** in which circumstances force us to seek new uses for the physical objects (including *animals* and other *people*!) in our environments, to say nothing of the words of our mental lexicons. Though we laugh when our computers analyze opaque idioms as compositional statements of literal fact, or misunderstand the intent of a conventional metaphor, trained government analysts who aim to make sense of the intelligence gathered from foreign theatres of operation are just as prone to such mistakes, especially when dealing with texts from a very different language and culture.



The funding agencies that provide the necessary backing for this kind of research are now opening their wallets to ambitious new directions in computational creativity. **IARPA**, the US government agency tasked with ***Intelligence Advanced Research Projects Activity***, currently funds a fascinating inter-disciplinary project to study the automatic detection, classification and analysis of metaphors in foreign-language texts, while the **European Commission** is even more active in its funding of collaborative European **CC** research. The **EC**'s new **PROSECCO** project – ***PRO****moting the* ***S****cientific* ***E****xploration of* ***C****omputational* ***C****reativity (****O****bviously!)* – is a coordination action to nurture the field of CC research in Europe, by running Summer schools and code camps, by providing a home for CC papers and teaching materials, and by serving as a hub for new Creative Web services as they come online. To stay abreast of this fast-evolving field, bookmark its official web-site: **PROSECCO-NETWORK.EU**

The End ... and a Beginning


ROBOT
COMIX