

InkWell: A Creative Writer's Creative Assistant

Richard P. Gabriel
IBM Research
650 Harry Road
San Jose CA 95120
rpg@us.ibm.com

Jilin Chen*
Google
1600 Amphitheatre Parkway
Mountain View, CA 94043
chenjilin@gmail.com

Jeffrey Nichols*
Google
1600 Amphitheatre Parkway
Mountain View, CA 94043
jwnichols@google.com

ABSTRACT

InkWell is a writer's assistant—a natural language revision program designed to assist creative writers by producing stylistic variations on texts based on craft-based facets of creative writing and by mimicking aspects of specified writers and their personality traits. It is built on top of an optimization process that produces variations on a supplied text, evaluates those variations quantitatively, and selects variations that best satisfy the goals of writing craft and writer mimicry. We describe the design and capabilities of InkWell, and present an early evaluation of its effectiveness and uses with two established literary writers along with an experiment using InkWell to write haiku on its own.

Author Keywords

Natural language generation; creative writing; creativity; writing

ACM Classification Keywords

I.2.7 Natural Language Processing: Language generation

INTRODUCTION

<p>A Snowy Evening Deep in the dark— the power of snow walking in the deepness —InkWell</p>
--

InkWell is a writer's assistant designed to help (creative) writers augment their creativity by generating a variety of revisions of a given text using a synonym-based dictionary and a wide variety of soft constraints or “in-

fluences.” Because one¹ of the members of the InkWell team holds an MFA in Creative Writing, the set of constraints InkWell can handle embodies the kinds of thinking a poet or fiction writer would do—such things as the *music* of the words (“the bird's fire-fangled feathers dangle down” [1]); subtexts, moods, and connotations; subtle semantic differences created by the influence of a set of

* Author was at IBM Research when this work was done.

¹rpg

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.
C&C 2015, June 22–25, 2015, Glasgow, United Kingdom.

Copyright is held by the owner/author(s). Publication rights licensed to ACM.
ACM 978-1-4503-3598-0/15/06/\$15.00

DOI: <http://dx.doi.org/10.1145/2757226.2757229>

words; a detailed language-usage model; accurate semantic senses; orthographic characteristics of words; and the notion of a spectrum from very associative word choices to very dissociative—we call this *writerly thinking*.

While designing InkWell we expected that a writer would examine the revisions, select promising suggestions, use the revisions as triggers for a better draft, and repeat this process as many times as needed to come up with a good draft. When we tested the system with two very experienced creative writers, we found another, much deeper aspect of creative writing InkWell can augment—InkWell can be a tool to help writers understand what their drafts are teaching them. Moreover, as we worked with InkWell, we discovered ways it could (almost) become a creative writer itself.

WRITERLY THINKING

Producing a work of written art requires constant discovery and guessing—because all works of art are works of exploration and discovery. MFA programs teach writers how the *craft elements* in poetry, fiction, and creative nonfiction achieve their effects—how rhyme ties ideas and images together, how images can form metaphors, how enjambment² creates multiple meanings and surprise, etc. And through constant practice, writers find ways to permit their drafts to teach them what those texts need in order to become art. This is not exactly like scientific, technical, and academic writing, which are about the transfer of information—though there are aspects of exploration there too. Robert Boswell, a fiction writer, puts it like this [2]:

I have grown to understand narrative as a form of contemplation, a complex and seemingly incongruous way of thinking. I come to know my stories by writing my way into them. I focus on the characters without trying to attach significance to their actions. I do not look for symbols. For as long as I can, I remain purposefully blind to the machinery of the story and only partially cognizant of the world my story creates. I work from a kind of half-knowledge.

In the drafts that follow, I listen to what has made it to the page. Invariably, things have arrived that I did not invite, and they are often the most interesting things in the story. By refusing to fully know the world, I hope to discover unusual formations in the landscape, and strange desires in the characters. By declining to analyze the story, I hope to keep it open to surprise. . . . The world remains half-known.

...

There can be no discovery in a world where everything is known. A crucial part of the writing endeavor is to practice remaining in the dark.

²The continuation of a sentence without a pause beyond the end of a line of verse.

Even the most commercial professional writers work this way; Stephen King [3]:

You may wonder where plot is in all this. The answer... is nowhere... I believe plotting and the spontaneity of real creation aren't compatible... I want you to understand that my basic belief about the making of stories is that they pretty much make themselves. The job of the writer is to give them a place to grow.

INKWELL

InkWell is designed to serve these two purposes:

- assist creativity in writing
- mimic a specific writer

InkWell takes a *template*—a specification of original text annotated with which words are variable and characteristics of those words for InkWell to consider—and produces revisions. Other writing-related constraints are written either as local bindings in the template or as global parameters and soft constraints. The example in Figure 1 shows how a writer might express a template describing Robert Frost's "Stopping By Woods on a Snowy Evening" [4]. (The original text of the poem is in the Appendix.)

InkWell In Action

Let's start by looking at what InkWell can do with the Frost poem. We will focus on the last two stanzas—here are the originals:

He gives his harness bells a shake
To ask if there is some mistake.
The only other sound is the sweep
Of easy wind and downy flake.

The woods are lovely, dark, and deep.
But I have promises to keep
And miles to go before I sleep,
And miles to go before I sleep.

The following revision was created using conservative synonym search (don't traverse far from the originally specified words), Ernest Hemingway [5] as the writer to imitate, and a preference for short words. Here is that result (changes from the original are underlined):

He gives his harness bells a cloud
To ask if there is some spot.
The only other music is the illusion
Of worn wind and cold mound.

The logs are good, glad, and tight.
But I have words to keep
And miles to go before I sleep,
And miles to go before I sleep.

The second variant was created using wild synonym search (go far and wide), Walt Whitman [6] as the writer to sound like ("Leaves of Grass"), and also short words:

He gives his harness bells a whip
To ask if there is some end.
The only other show is the push
Of a couple of winds and a falling chip.

The woods are clean, murky, and vigorous.
But I have roots to preserve
And miles to go before I sleep,
And miles to go before I sleep.

The last example is like the Whitman one except InkWell was directed to sound a little like Adam Smith [7] and to be a lot wilder:

He gives his harness cacabels a scramble
To consult if there is some stumble.
The only other imposition is the affixation
Of counterbalanced levanter and puberulent pebble.

The cocuswoods are equitable, dishonorable, and redeemable.
But I have diminutives to number
And millimeters to go before I slumber,
And millimeters to go before I slumber.

In addition, InkWell was instructed to prefer long words and to work hard on rhymes.

These are two of the haiku InkWell wrote:

a man	a puddle of water:
steps out	as if the sun had changed
of the old woman	

Mimicking Writers

InkWell mimics writers by using techniques including these:

- match specified (or measured) Big Five personality traits and associated personality facets [8]; match basic human values as described by Schwartz [9] and Chen [10]
- match a writer's word choices: favored words, word music, word length, favored mood
- match writing patterns: *n-grams* (2-, 3-, 4-, and 5-grams); an *n-gram* is a series of *n* words in a row that has appeared in a naturally occurring, existing text.

Personality-based features (originally included to support work in persuasion campaigns) turned out to be useful for creative writers—for example, to make the text more agreeable by favoring words and phrases that exhibit that personality trait or to match the personality profile of a particular writer.

Assisting Creativity

InkWell assists creativity by using techniques including these:

- use conservative or wild synonym choice (*associative* versus *dissociative* writing)—search diameter, search relevance decay rate, preference for nearby words, preference for far away words, which synonym aspects to use (hypernyms, meronyms, etc)
- obey or defy *n-grams*—that is, use familiar or novel phrasing
- satisfy constraints such as word-length, alternative meanings, word rhythms
- favor rhymes and echoes (similar sounding words)
- select words based on ontology (concepts) or a cluster of word-centric concepts to favor or avoid
- mimic any specified personality profile
- take into account a mood specified by a construct called a *halo*

```

(with-personality-traits (*writer-big-five*)
 (with-global-constraints ((all-echo) (all-different))
  (with-pervasive-predicates (#'syllable-bonus-few)
   (bind ((w1 (<choose> verb-cognition :+sense [know certain]))
          (w2 (snow noun-substance :rhyme [though]))
          (w3 (<choose> adj :+sense [queer odd unusual weird demented stupid silly]))
          (w4 (or (year noun-quantity) (week noun-quantity) (month noun-quantity) (season noun-quantity)))
          (w5 (<choose> verb-motion :+sense [cause move back forth shake]))
          (w6 (<choose> noun-object :+sense [small fragment broken break whole flake]))
          (here (here adj)) (near (near adj)) (mile (mile noun-quantity pl)) (sleep (sleep verb))
          (woods (wood noun-plant pl :+sense [forest trees plants wooded area] :-sense [wood])))
    "Whose (ref woods) these are I (<choose> verb-cognition :sense think-sense) I (ref w1).
    His (house noun) is in the (village noun) though;
    He will not see me (stop verb gerund) (ref here :rhyme near)
    To (watch verb-perception) his (ref woods) (fill verb) up with (ref w2 :different w1 :rhyme w1).

    My (<choose> adj :sense little-sense) (horse noun-animal) must (<choose> verb-cognition :sense think-sense) it
    (ref w3 :echo w1)
    To (stop verb) without a (<choose> noun :sense farmhouse-sense) (binding near :rhyme w3)
    Between the (ref woods) and (frozen adj) (<choose> noun-object :sense lake-sense)
    The (<choose> ask -est :sense darkest-sense) (<choose> noun-time :sense evening-sense) of the
    (ref w4 :different w3 :rhyme w3).

    He gives his (harness noun) (bell noun pl) a (ref w5 :echo w3)
    To (ask verb) if there is some (mistake noun :rhyme w5).
    The only other (<choose> noun :sense sound-sense) is the (<choose> noun-act :sense arc-sense)
    Of (<choose> ADJ :sense easy-sense) (wind noun) and (<choose> ADJ :sense downy-sense)
    (ref w6 :different w5 :rhyme w5).

    The (ref woods) are (<choose> adj :sense lovely-sense), (<choose> adj :sense dark-sense), and
    (<choose> adj :sense deep-sense),
    But I have (promise noun pl) to (keep verb :rhyme sleep),
    And (ref mile) to go before I (ref sleep),
    And (ref mile) to go before I (sleep verb :different sleep :rhyme sleep)."))))

```

Figure 1. This is the internal template format used by Inkwell for “Stopping By Woods on a Snowy Evening.” It can be edited directly by users, but is usually created using our visual template creation tool. We show this format to demonstrate the functionality of Inkwell.

Every soft constraint has an associated weight, and thus any constraint can be inverted (negated): e.g. sound like a particular writer or sound like anyone but that writer, rhyme two words or avoid rhyming them, observe n-grams or deliberately violate them. InkWell produces any number of candidate revisions, and the writer can pick and choose revisions and wordings.

The notion of a *halo* is a good example of how InkWell was designed to mimic writerly thinking. A *halo* is a mood device: specify a set of words, and InkWell starts with each of those words and fans out along synonym arcs to other words. Where several of these wavefronts hit, those words are given more weight in the revision process. Looking at Frost’s poem, the line

The woods are lovely, dark, and deep

might be revised this way

The woods are bright, not very light, and high

when given this halo:

delighted, ebullient, ecstatic, elated, energetic, enthusiastic, euphoric, excited, exhilarated, overjoyed, thrilled, tickled pink, turned on, vibrant, zippy

and this way

The woods are sad, bad, and dead

when given this halo (and a preference for short words and strong rhyme):

affronted, belligerent, bitter, burned up, enraged, fuming, furious, heated, incensed, infuriated, intense, outraged, provoked, seething, storming, truculent, vengeful, vindictive, wild

Halos are an example of the strategy behind InkWell: as much as possible, use concrete words and their characteristics rather than reifications and abstractions. It’s of course possible to label the two halos above as *Anger* and *Happiness*, but the actual words tell the story more directly.

InkWell uses lots of data:

- Expanded Wordnet Synonym dictionary: ~160,000 words [11] [12]
- Expanded LIWC 2001 dictionary [17]—for personality analysis
- 20,000 most common words
- CMU Phonetic dictionary: algorithmically expanded to ~220,000 words [13]—for rhymes and echoes
- Stem dictionary: ~163,000 entries (+ Porter Stemmer + Lemmatization + Wordnet’s Morphy algorithm)
- n-grams: ~24m from general literature including the Google 2-grams [14] and the COCA 3-, 4-, and 5-grams [15]
- n-grams (including 1-grams) from writers; 100,000–1,000,000 per writer; currently there are about 70 writer samples to choose from (and supplying new ones is trivial)

N-grams are used to try to maintain some degree of familiarity and coherence. There are two sets of n-grams used: one from general literature (we use a filtered version of

the Google 2-grams [14] but use COCA for the 3-, 4-, and 5-grams [15]), and another set of n-grams from the writer chosen to be imitated. These constraints are expressed as global variables, with separate weights possible for each of the 2-, 3-, 4-, and 5-grams. In all there are 57 different constraints that can be imposed on the revision process (not all are described in this paper).

For the Whitman- and Hemingway-influenced revisions of the Frost poem, InkWell was told to conform to the target writer’s word choices and n-grams with writer 1-, 2-, 3-, and 4-gram weights of 20, 20, 30, and 40, respectively, and with a rhyme weight of 50 and an echo weight of 20. For the Adam Smith revision, Inkwell used -20, -30, and -40 for the writer 2-, 3-, and 4-gram weights, and 100 for both the rhyme and echo weights.

Synonyms and Word Senses

Synonyms in revision are tricky. It is not enough to know the part of speech or even the role of a word in a sentence—the exact sense of the word in question is essential. To revise the sentence “I like dogs” requires knowing whether by “dog” the writer means:

- *animal*: a member of the genus *Canis*
- *person*: a dull unattractive unpleasant girl or woman
- *person*: informal term for a man
- *person*: someone who is morally reprehensible
- *food*: a smooth-textured sausage of minced beef or pork, usually smoked
- *artifact*: a hinged catch
- *artifact*: metal supports for logs in a fireplace

Narrowing it down to an *animal* would characterize it well, but *person* and *artifact* still would leave it ambiguous. We have been working with definition-based word sets as a way to pluck good senses from synonym sets. We call these structures *senses*, and they form the most rapidly evolving parts of InkWell.

A sense is a data structure that can be thought of as a vector of words along with associated strengths. A set of words is supplied, and using each as a starting point, InkWell visits synonyms as specified by current synonym search settings (hypernms, hyponyms, etc) out to the current synonym-search depth, and adds those words with their decayed relevance values to the vector. If a word is already in the vector, the associated relevances are summed. The vector is normalized.

InkWell can use a sense to locate words. Each word (including all its synonym senses) in the dictionary is examined along with its definition (which is just a sentence or phrase). The words in the definition are compared to the vector, and a strength computed. A subset of the strongest candidates are selected. For example, if given a sense rooted in the words “sing” and “bird,” InkWell will find words like “sparrow,” “lark,” and “cockatoo.” A negative sense can be specified as well—a sense to avoid; a full sense can be a combination of positive and negative senses.

The use of senses is seen in Figure 1 where the expression

```
(wood noun-plant pl
  :+sense [forest trees plants wooded area]
  :-sense [wood])
```

means that the word chosen to mean “wood” should be in the semantic category *plant*, should be made plural when rendered, should be of the same sense as the words “forest,” “trees,” “plants,” “wooded area,” and not the same sense as “wood,” (that is, the material that trees are made of).

WordNet provides a variety of synonym types such as hypernyms, hyponyms, homonyms, meronyms, and variations on each. This complicates things. Which of these to consider is part of the specification InkWell uses to revise text.

On Templates

A template resembles a Lisp program whose body looks like text with parenthesized annotations. The text in Figure 1 is a template—it shows many of the mechanisms for specifying the text well enough to be effectively revised. Here are some of the highlights:

- items in parentheses are specifications of word choices; a word in the first position indicates the word whose synonyms are considered; <choose> indicates InkWell should find words that satisfy the specified word-sense description; subspecifications include:
 - the part of speech and semantic category
 - adjustments to the word when finally rendered; (*pl* for plural, for example)
 - words that should be the same, should be different, should rhyme, should echo
- words required to be the same are specified in the *bind* form
- local and global constraints are possible—short words, everything should echo, etc.
- personality traits to aim for

Creating complex templates requires tool assistance. InkWell has a template creation tool: it takes unadorned text and leads the writer through selecting adjustments, semantic categories, and the senses of words the writer intends. The tool is lightweight and visual—able to help novice writers create templates like the one in Figure 1 in less than five minutes. We are exploring ways to make this part of the process more automatic by integrating a parser (which helps only a little) and experimenting with determining the best sense for each word by looking for coherent assignments of sense in a text. For example, in a sentence like this there should be no question which sense of *dog* to use: “This dive restaurant serves amazing dogs slathered with messy, meaty chili.”

InkWell Flow

Each template along with all the specified constraints and parameters is compiled into an evaluation function which returns 0 when all constraints are satisfied. InkWell selects a set of candidate replacement words and phrases, and an optimization process then selects the combination of words and phrases that best satisfy the evaluation function. The optimization process uses simulated annealing. One nice characteristic of simulated annealing is that it requires neither gradients in the evaluation functions nor any explicit staging of

the order of choices. Just add constraints and let the relatively undisciplined SA process do its thing.

For example, to evaluate how close a set of word choices is to a specified personality profile, InkWell analyzes the proposed text to (computationally) determine its exhibited personality. This is then compared to the target, and either the current text is considered a step in the right direction or a step in the wrong direction. For the personality computation, InkWell builds on the work of Tal Yarkoni [16] and uses a method which computes LIWC scores [17] for the generated text, and uses a simple, learned classification function to produce Big Five scores. Because the analyzed texts are generally pretty short, the LIWC dictionary was expanded algorithmically using the WordNet synonym dictionary, a decay function, and semantic categories derived from the LIWC categories—this expanded the LIWC dictionary from about 2300 words to about 25,000 words. The computation of personality scores needs to be fast, which is accomplished by pre-computing as much of the LIWC categories as possible and using caches. InkWell also targets *values* [10], which are computed similarly.

The overall evaluation function can be computationally intense, which is appropriate for a deliberate, creative writing task, but perhaps not for real-time language generation. For example, the template shown in Figure 1 compiles to a Lisp function ~2000 lines long which is invoked 200,000 times during a typical optimization run. This means that the rhyming predicate is invoked ~3.6m times. This computation is feasible through the use of parallelism and many caches.

CASE STUDIES

We performed two case studies using InkWell. One was a writer study to see how well InkWell could serve the needs of (professional) creative writers. The other was to use InkWell as a fully automated haiku generator, and to determine whether it would produce decent haiku with recognizable aesthetics.

For the writer studies we met with two established writers—a fiction writer (“Julia”) and a poet (“Quinton”). One of the authors of this paper described and demonstrated InkWell to each writer (separately), and then engaged the writer in an exercise working with a draft of theirs.

The Fiction Writer

The fiction writer was first. Julia holds an MFA in Creative Writing (Fiction), and is well published in literary magazines. She also teaches writing to young adults. She was trained as a dancer and has an outgoing, extravagant personality. Much of her fiction is funny and fun. We began by inputting her text and creating a template.

Poets bartered commodities when they were available, amended soil with their waste when they were not. Composers filled reservoirs with orchestras, music crashing through the pipes into dry fountains and apartment sinks. Farmers made bibles of seeds and left them in every nightstand drawer, hotels and residences alike.

Initial Reactions

Julia’s first reaction was subdued.

The initial settings for the constraints and parameters were conservative (search to synonym depth 2, consider more general words, related words, and similar words), and thus the early revisions were close to the original. For instance, one was the following:

Poets bartered commodities when they were available, amended soil with their waste when they were not. composers filled reservoirs with orchestras, music crashing through the pipes into dry fountains and apartment sinks. Gods made promises of seeds and left them in every nightstand drawer, hotels and residences alike.

Julia’s reactions were “*kind of fun to add that stylistic layer or aesthetic*” and “*the sense... the relationship to reality has stayed the same... or is mine, still—that’s how it feels.*”

A Turning Point

At that point the researcher proposed making substantial changes to the settings. In particular, the synonym search was directed to look further from the original words and for more specific rather than more general synonyms. This yielded the following:

Poets bartered middlings when they were many, amended soil with their waste when they were not. Psalmists lined cisterns with string orchestras, music crashing through the lines into sound fountains and rooms sinks. Husbandmen made words of seeds and found them in every dresser drawer, courts and palaces alike.

This triggered Julia’s first strong reaction—a little scream and then: “*Oh, love it—so much better. Oh so interesting—so much more interesting. ‘Psalmists lined cisterns with string orchestras’—that’s as near the sense I was going for... only the language is much more specific. That’s fun. That looks fun to me. This is a rich mess.*” After this, the nature of the interaction changed—she was much more engaged, and suggested a number of experiments. Nevertheless, her goal was not what we anticipated. Instead of tweaking her draft, she was trying to understand or come to grips with the “structure” of her draft and the nature of its internal sense, consistency, and coherence—of the writing itself but also the world she was trying to create. Rather than seeking help with revision, she was seeking diagnostic entry points.

- “*I’m still determining the structure, the internal sense of it*”
- “*This is really interesting to me in terms of doing the ‘thinking’ work*”
- “*{the program helps figure out} what’s important in it, in the sense of it, where I’m creating a world that is very odd and quite intentional about the oddness, and you can get lost, really lost. It’s really good for honing... focusing in on what’s important.*”
- “*That’s the only way I can think of to do it—to see how many interesting iterations I’ve got, see what language sticks to my head. So I’m looking for velcro basically, and then, you know, I have the attachment piece here {points to head}*”.
- “*what’s... frightening about this is that it is teasing out some things—I suppose they are in there.*”

- “I can see how you would get very familiar with the little tide pools that the language, where they come from. operating behind there. Then you would be able to say ‘ok, I want to mess with it **this** way. . . .”
- “It’s kind of interesting to see what it changes, so it’s helping me understand what’s operating”

Julia was looking at multiple revisions (“iterations”), and because they were theoretically “close” to her original semantically—because the program looks at synonyms—the connotations and other triggers she was seeing were teaching her about what she had put on the page, and possibly how her thought processes were circling those nearby “tide pools.”

Imitating Writers

Julia was also interested in some of the wilder revisions—either when InkWell was directed to look far away from the original variable words or when specific writers were targeted for imitation. The following is a revision in which the King James Bible was targeted for imitation, along with the personality traits displayed there. We also tweaked that personality by aiming it more toward disagreeableness and extraversion:

Poets bartered numbers when they were perverted, choked clay with their scrap when they were not. Psalmists silenced washbasins with orchestras, euphony crashing through the phone lines into naked fountains and teaser baptistries. Tillers made bibles of rapeseeds and gave them in every medicine chest knee breeches, convents and convents alike.

. . . and this:

Poets bartered burdens when they were defenseless, hit sand with their piss when they were not. Psalmists cut cisterns with orchestras, euphony crashing through the pipes into outrageous fountains and teaser privies. Tillers made mournings of mockernut hickories and cut them in every vanity drawer, resorts and palaces alike.

Her comment regarding the influence of the Bible on the otherwise wild constraint settings was “I love it. So it takes the wildness and gives it a sort of imprimatur—something a little more solid. . . . And is it still mine, though? . . . We’ve run it through filters I didn’t create so much, though I ran it through my eyeball filter after that.” Her concern was that by using an aid such as InkWell, which alters her text under the influence of word-usage models of other writers, the resulting text might not be considered hers alone. She talked about how there is a similar effect when she reads another writer and there could be residual influences on what she writes after that, but that this seemed different from that.

“I know I wouldn’t have gotten [there] that fast with ‘hitting the sand with their piss’ or ‘choking the clay.’”

To end the session we switched to Ernest Hemingway as the writer to imitate.

Bards bartered picture shows when they were cleanable, doctored Kitty Litter with their scrap metal when they were not. Composers filled out washbowls with chamber orchestras, euphony crashing through the steam pipes into liberal arts fountains and flat public toilets. Planters made content words of gentleman’s-canes and called it a day them in every hope chest Jamaica shorts, holiday resorts and monasteries alike.

Julia was taken with the phrase “doctored Kitty Litter with their scrap metal” and remarked “That turns out to be one of

those fruitful little phrase arrangements. That is teaching me something about my writing.”

Julia’s Final Comments

Her final comment was as follows:

“So basically you’re kind of channeling different voices to see ‘can I mess with my syntax a little more, mess with the flavor, can I take vanilla or chocolate, my standard fare, my standard mannerisms in the written language, and mess with them?’ It works for my aesthetic, certainly, because it’s language driven. . . . and it could work for poets. I don’t know with straight fiction, although it might make it a whole lot better.”

Analysis

InkWell is a complex program, and there are many dials and sliders, and to warm things up, we started with conservative settings. She initially saw only an “interesting” level of value because the revisions were not far from her original.

Once she saw the first unusual or strange revision, she became more engaged, but her thought process was a surprise to us. Instead of primarily trying to harvest the revisions for phrasing to adopt or triggers for further writerly thinking, she used InkWell as a diagnostic tool or an instrument to explore her own thinking and process—she tried to make sense of what the revisions taught her about her draft and its consistency and coherence. Another way to put it is that she was trying to analyze her draft in a way similar to a “close reading.” A close reading is a deep analysis of a text, especially its use of language, in order to come to grasp its meaning and understand how it conveys that meaning. [18] [19] [20]

Only after that did she come to believe InkWell could help her “mess” with her text to explore alternative approaches—that is, to use InkWell as a direct writer’s assistant—or perhaps a writing partner.

The Poet

The poet was second. Like Julia, Quinton holds an MFA in Creative Writing, and is well published in literary magazines. Quinton has a book of poetry published—his manuscript won a poetry competition whose prize was publication. Quinton primarily writes formalist poems—fixed meter and rhyme scheme. His work is directed to observing small moments in urban settings. He is close to retirement age, served in the Navy for twenty years, is church-going, and is socially conservative. He is a lawyer.

Quinton wanted to work on a poem he had just started. The fragment was quite short:

The conic dome sluiced smartly on its tube
Is slung across its luggage cart and docked
in an aisle while its owner searches
a department for . . .

The poem fragment is describing small missiles on a dolly in a narrow passageway on a Navy ship. As with Julia, Quinton read the first couple of conservative revisions and began analyzing them—but this time it was to understand the revision

more than to understand his own draft. The following was one of the first revisions InkWell produced:

the conical dome flushed smartly on its thread
is slung across its grip barrow and docked
in an aisle while its owner gathers
a number for

For example, Quinton tried to make sense of the word “flushed”—how it could mean both “flush” as in clearing out with forced water as well as “flush” as in placed right up against a wall (“flush with the wall”). We continued with revisions, eventually using Herman Melville’s “Moby-Dick” [21] as the piece to imitate.

A Turning Point

Again, as with Julia, we hit a turning point or *aha!* moment when InkWell produced the following:

the conic dome flushed sprucely on its safety lock
is catapulted across its pocket pickup arm and docked
in an aisle while its lady of the house unlocks
a line of defense for

Quinton let out a laugh and said “*I love that. Anything that breaks you out of that earnest, that linear thinking.*” At this point he told the story of how the leader of a writing group he was in gave him the exercise to write the worst poem he could. Then after presenting that poem to the group, he was instructed to revise it to make it even worse. He said that this exercise, repeated for several poems, gave him some of the “strongest” poems he’s ever written. Quinton explained as follows:

“I think part of that was giving you permission to do things that are just wild, that you don’t think of as poetic and this program seems to be doing the same thing. You’ve got something in your head, you’re describing it in a certain way, and this is just showing you all these other ways of coming at this; and not only to find a word to describe what is in your head, but kind of to blow up your head, expand it so you can in a more objective way, you can go ‘wow, this whole thing that has nothing to do with what I was going after when I was starting out is actually kind of interesting.’ So you might end up going in a completely different direction, and writing a completely different poem that’s a hell of a lot more interesting than what you started out with. So I kind of like it as a generative tool.”

Quinton believed InkWell could serve as a relentless “Dean Young in a box.” Dean Young [22] has been described as a neo-surrealist, and his approach sometimes is to use surrealism to explore the imagination and to break down the border between the real and the unreal. Dean Young was one of Quinton’s teachers. Quinton described Dean Young as someone whose brain is wired as a dissociative engine, and that InkWell could help poets who wanted to write that way but were too cautious or reserved.

In the end Quinton asked when InkWell would be available for his use.

Analysis

Quinton went through a similar discovery process with InkWell to what Julia did: it took an unusual, strange, and unexpected revision to move him into a more engaged exploration of what InkWell could do. We then used a variety of writers to imitate, used both conservative and wild settings, and did some variations on the personality trait settings. Quinton used it both as an instrument to explore his own draft and as a generative aid for the sorts of disconnected jumps that are hard for him to do on his own. He always circled back to the idea that escaping “earnestness” and “linear thinking” was valuable to writers. The nature of exploring his own draft was less direct than Julia’s—he spent time trying to understand the mind that produced the revisions while Julia spent time trying to understand how what that mind produced revealed aspects of her own writing.

Both writers are experienced and well-trained. Both writers, when confronted with their first *aha!* revisions, displayed what could be described only as surprise and glee: Julia screamed and Quinton laughed.

InkWell as Poet

In our second study we wanted to see whether InkWell could compose decent (contemporary) haiku on its own using a fully automated, highly aleatoric process. Haiku is a good initial domain for InkWell composition experiments because syntactic sophistication is not required (haiku consist mostly of the juxtaposition of images), semantic linking of images is required, and their brevity means we can generate a lot of them to get a statistical view into InkWell’s abilities.

While pursuing this study we designed a simplified template mechanism and in the process, discovered a couple of techniques for improving generated text. Figure 2 shows one of the 54 simplified templates we used for generating Haiku; such a template compiles into the more verbose form already described. Almost all the templates we used were derived from Robert Hass’s translations of Bashō’s haiku [23]. Contemporary Japanese and English haiku have few syllabic constraints, but in later experiments we were able to approximate the popularly understood but not quite accurate 17-syllable constraint.

Each haiku template requires InkWell to choose all the non-trivial words based on sense specifications. We supplied a set of base senses to use—for example, the *change* base sense (not used in Figure 2) was constructed from the words “change,” “produce,” and “yield.” When a haiku is produced, each of its referenced base senses is augmented with the same randomly (and automatically) constructed other sense, as follows: a writing sample is randomly selected, and a sense is built from a randomly chosen contiguous segment of that text. For example, here is such a set of sense words selected from “Leaves of Grass”: “form,” “upright,” “death,” and “breast.” The constructed sense is then combined with each base sense using the formula $b \times B + (3 - b) \times C$ where $b = R(0.1, 1.0)$ and $R(x, y)$ produces a random float between x and y , B is the base sense, and C is the randomly constructed sense. This formula describes how each base sense and randomly

```
( (or a an) (young-insect noun-animal (mix-senses insect-sense-base ([larva young child] noun)))
 (comma) (return) this (intense-sense noun-time) (preposition?)
 (fall-sense noun-time (mix-senses season-sense-base ([season leaf fall] noun))) (return)
 (remain-sense adv) not (or a an)
 (old-insect noun-animal (mix-senses insect-sense-base ([adult old elder] noun))))
```

Figure 2. One of 54 Simplified Haiku Templates Used by InkWell

constructed topic sense are to be combined—a linear combination with random coefficients to vary the relative influences of the two component senses, with the randomly constructed sense having more influence than the base sense. For example, the combined senses of the sense words selected from “Leaves of Grass” and the senses specified in Figure 2 are based on the following words:

```
form upright death breast larva young child insect social bee animal
form upright death breast deep central
form upright death breast leaf fall season
form upright death breast still
form upright death breast adult old elder insect social bee animal
```

After the thematic senses are set, all the other InkWell constraint parameters are chosen randomly except writer n-grams, which were ignored because we wanted to see whether writer-specific traits would emerge. One of the 54 haiku templates is selected, and InkWell writes that haiku.

With this process we created about 1800 haiku. We looked at two questions: Are any of the haiku good? Do any of them conform to a recognizable aesthetic?

The first thing we did was to read them—that is, to see whether any of them were actually good haiku or at least good short poems. We observed a range of quality from the demented (*a hard tick, / this round-trip light time in round-trip light time / like mad not a whitebait*) to the breathtaking (*deep in the dark— / the power of snow / walking in the deepness*) to the demonically clever (*tuned adrenalin / my music, / a beat-boogied headful*).

Because the use of the same (randomly constructed) sense to augment all the base senses in the haiku templates tended to produce related word choices, many of the haiku were thematically coherent. We used a scoring system based on n-grams and music to measure aesthetics—when asked what kind of writing they like, most people would name authors, and some would mention language use. Scoring based on n-grams measures the density of recognized n-grams in a haiku. *Normal* scoring uses the n-grams from general literature, and so measures how conventional a haiku is; *writer* scoring uses the n-grams from a particular writer, and so measures how much like that writer a haiku is. *Music* scoring measures the density of rhymes and echoes in a haiku. Here are some of the haiku that stood out.

These were in the top fifty using a combination of *music* and *normal* scoring (musical but like ordinary prose):

```
back in the past—                on in the dark—
the start of the shit set piece    the length of the stone building
the ways of the world in the blue  breaking in the heat
```

These were in the top fifty using *Walt Whitman* scoring (sound like Walt Whitman):

```
a maverick
troops out
of the pride of California
```

Adam Smith:

```
a drop of water:
as if the bank had thought
```

King James Bible:

```
a man
passes out of
the pit
```

William Faulkner [24]:

```
a hell of pitch
as if the ring had exploded
```

Ernest Hemingway:

```
a trout
moves out of the bay

late in the afternoon
the glare of piss—
infernal machine sitting
in your seat
```

```
awake in the dark—
the edge of the water can
spread in your presence
```

```
scrupulous in the twilight—
the price of gold chases
the way of the world in power
```

```
time of life issue:
a bird of prey pulls up
out of the way, into the palm
```

```
short sight fog—
just enough to turn the face
of a man into a nag
```

```
a man
steps out of the head
```

```
even in the afternoon
the advantage of the mother
stills the ways of the world
```

And some of the best ones appeared at the end of scored lists: here are the not so *normal* and *unmusical*:

```
a dangerous work shift interruption—
producing dumplings, we cut nasturtiums
```

```
a crooked rag day—
by myself
dunking distracted sardines
```

```
a slow circadian rhythm set:
harmonizing oysters
broke fresh foods
```

Although these haiku might appear quite strange, they stand up well compared to highly respected surrealist haiku from the twentieth century, such as these by the French poet, Paul Éluard [25]:

```
The wind
hesitating
rolls a cigarette of air
```

```
The automobile is truly fast
four matured heads
move beneath it
```

Grammar

The haiku experiments revealed the need for mechanisms to get simple grammar right. We had always known that n-grams would be part of the solution, but we weren’t sure how to best accomplish that. We noticed right away that articles, prepositions, and determiners were sometimes wrong because the early templates explicitly mentioned them as constants. We discovered two mechanisms to address this. The first is an improved version of the `OR` annotation, which indicates explicit choices. The expression

```
... (or a an) noun ...
```

directs InkWell to choose the article that matches the selected noun. Some verbs take specific prepositions after them (like “run into”), so we added the form `(preposition)`,

which expands to an `or` that lists all possible prepositions. But this wasn't enough. In some cases InkWell needs to decide to not choose a word. We created the construct `<null-word>` to represent a position with no word in it. The form `(preposition?)` tells InkWell to choose either the correct preposition or no preposition at all; it expands to `(or <null-word> on in under ...)`. With these, this template

```
A dog (or is are) (or a an) animal.
```

resolves to *A dog is an animal.* and this one

```
Dogs (or is are) (or <null-word> a an) animals.
```

to *Dogs are animals.*

The use of these techniques in the haiku templates made ungrammatical haiku rare—crazy, perhaps, but grammatical.

RELATED WORK

Existing tools that we are aware of broadly seem to break down into two categories: teaching the writing process and assisting in the production of complex documents. Tools that teach writing focus on two areas, teaching young children to write (e.g. [26]) and assisting English as a Second Language learners (e.g., [27]). These systems either assist in the process of developing overall content and structure, or with very low-level details of grammar and language.

Other tools assist in the construction of “complex documents,” which are longer documents with lots of internal structure and many relations between different areas of the document. Mexica [28] is a system to help model the constraints of a complex document and to assist in writing based on a cognitive model of the writing process. There is great deal of work in assisting with the creation of narrative [29].

InkWell can be seen as supporting the Create phase of Shneiderman's genex model [30]. The Create phase is broken into three components: Explore, Compose, and Evaluate. Based on the results of our studies, Inkwell is currently useful for the Explore and perhaps the Evaluate components, and our eventual goal is to support Compose as well. Other work in supporting creativity for writers has examined scoring sentences for creativity based on a machine learned model [31].

Most of the work on Natural Language Generation (NLG) to this point has focused on planning the content, sentence structure, and sequencing of text. To the extent that style is considered in NLG systems, it's directed toward text generation for journalism or instruction manuals. Systems that produce stories are mainly concerned with planning and story coherence. Most of the NLG systems fall into two categories: rule-based systems and statistical systems. The former generally use handcrafted templates to generate text; the latter generally use statistical models to generate candidate utterances which are then checked for “validity,” typically using n-grams. The only NLG system that is directly related to InkWell is Personage by François Mairesse and Marilyn A. Walker [32]. It's a full NLG system which also is looking at producing text exhibiting Big Five personality traits.

POTENTIALLY IMPORTANT OBSERVATION

The InkWell project started as a simple revision program designed to match personality traits as part of a research system aimed at countering malicious persuasion campaigns—it did not start as a project about creativity. In fact, InkWell's strengths as a creative partner emerged as a surprise when we noticed that using wide synonym search parameters and negative weights yielded sort-of relevant but unexpected wording changes. One way of thinking about this is that creativity is not like a module in a system, nor is it something that can be designed for—what would its requirements be, what sort of specification would it have? Creativity is a property one notices in systems.

Only in hindsight is some modicum of explanation available: InkWell is creative because its 57 parameter settings are balanced against each other during optimization, so that creative choices are possible like selecting a not-as-directly relevant word that happens to rhyme very well with some others or offsetting some agreeability with some not-so-sensible n-gram conformance.

CONCLUSION

InkWell is a tool aimed at helping deeply creative people. It has many capabilities which combine to produce a variety of revisions of a template influenced by writer-craft elements, personality traits, and writer models. InkWell is complex, but so are the facets that go into a creative writer's thought processes. Through an initial pair of writer studies we learned that InkWell can be used for exploring alternative wordings and metaphors as well as for aiding a deeper analysis of the meaning and trajectory of a writer's project. With the haiku experiment we began to observe InkWell taking the role of a writer by choosing its own words (not simply finding synonyms), and using a common “sense” overlay to work toward coherence. We also found that the haiku produced could be partitioned after the fact into aesthetic groups based on similarity to other writers, music, and distance from “normal” texts. In this work, the notion of “sense” has proven to be the most intricate and important.

A remaining challenge is syntactic transformations. For example, everyone knows Hemingway tends to use short sentences and lots of “and”s—InkWell currently cannot transform sentence structure to match Hemingway. It can do only trivial transformations after optimization.

Creativity in writing takes a couple of ingredients: being prepared to notice, wide ranging and non-judgmental production of drafts, and a selection and revision process guided by aesthetics. In a sense it's like serendipity. With InkWell we've tried to explore writerly creativity in a creative way—by letting the program emerge from a haze of half-known ideas [2].

ACKNOWLEDGMENTS

Some of this work was supported by DARPA (W911NF-12-C-0028).

REFERENCES

1. Wallace Stevens, *Of Mere Being*, “The Palm at the End of the Mind: Selected Poems and a Play,” Vintage, 1990.

2. Robert Boswell, "The Half-Known World: On Writing Fiction," Graywolf Press. St. Paul, MN, 2008.
3. Stephen King, "On Writing," Pocket, New York, 2002.
4. Frost, Robert, "New Hampshire," Henry Holt, New York, 1923.
5. Ernest Hemingway, *The Complete Short Stories of Ernest Hemingway*, Scribner, 1998.
6. Walt Whitman, *Leaves of Grass*, 1855.
7. Adam Smith, *An Inquiry into the Nature and Causes of the Wealth of Nations*, 1776.
8. O. P. John & S. Srivastava, "The Big-Five trait taxonomy: History, measurement, and theoretical perspectives," in L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (Vol. 2, pp. 102–138), New York, Guilford Press, 1999.
9. S. H. Schwartz, "Basic human values: theory, measurement, and applications," *Revue Française de Sociologie*, 47(4), 2006.
10. Jilin Chen, Gary Hsieh, Jalal Mahmud, Jeffrey Nichols, "Understanding Individual's Personal Values from Social Media Word Use," in Proceedings of CSCW 2014, Baltimore, MD, February 15-19, 2014.
11. George A. Miller, "WordNet: A Lexical Database for English," *Communications of the ACM* Vol. 38, No. 11: 39–41, 1995.
12. Christiane Fellbaum (ed.) *WordNet: An Electronic Lexical Database*, Cambridge, MA: MIT Press, 1998.
13. <http://www.speech.cs.cmu.edu/cgi-bin/cmudict>.
14. <http://storage.googleapis.com/books/ngrams/books/datasetv2.html>.
15. Corpus of Contemporary American English, http://www.ngrams.info/download_coca.asp.
16. Tal Yarkoni, "Personality in 100,000 Words: A large-scale analysis of personality and word use among bloggers," *Journal of Research in Personality*, 44(3): 363–373, 2010.
17. Yla R. Tausczik and James W. Pennebaker, "The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods," *Journal of Language and Social Psychology*, 29(1) 24–54, 2010.
18. http://en.wikipedia.org/wiki/New_Criticism.
19. I. A. Richards, *Practical Criticism: A Study Of Literary Judgment*, Mariner Books, 1956.
20. C. K. Ogden and I. A. Richards, *The Meaning Of Meaning*, Mariner Books (reissue edition), 1989.
21. Herman Melville, *Moby-Dick: or, the Whale*. 1851.
22. [http://en.wikipedia.org/wiki/Dean_Young_\(poet\)](http://en.wikipedia.org/wiki/Dean_Young_(poet)).
23. Matsuo Bashō and Robert Hass (translator), "Matsuo Bashō: Poems," <http://poemhunter.com>, 2004.
24. William Faulkner, *Collected Stories*, Random House, 1950.
25. Jeffrey Johnson, "Haiku Poetics in Twentieth-Century Avant-Garde Poetry," Lexington Books/Rowman & Littlefield Publishing Group, 2011.
26. John Halloran, Eva Hornecker, Geraldine Fitzpatrick, Mark Weal, David Millard, Danilus Michaelides, Don Cruickshank, David De Roure, "The Literacy Fieldtrip: Using UbiComp to Support Children's Creative Writing," Proceedings of the 2006 Conference on Interaction Design and Children, DC '06, pp 17–24, 2006.
27. Yu-Chia Chang, Jason S. Chang, Hao-Jan Chen, Hsien-Chin Liou, "An automatic collocation writing assistant for Taiwanese EFL learners: A case of corpus-based NLP technology," *Computer Assisted Language Learning*, Vol 21 No 3, pp 283–299, 2008.
28. R. Pérez y Pérez and M. Sharples, "MEXICA: a computer model of a cognitive account of creative writing," *Journal of Experimental & Theoretical Artificial Intelligence*, Vol 13, pp 119–139, 2001.
29. Michael Mateas and Phoebe Sengers, "Narrative Intelligence," Working notes of the Narrative Intelligence Symposium, AAAI Fall Symposium Series, 1999.
30. Ben Shneiderman, "Creating creativity: user interfaces for supporting innovation," *ACM Transactions on Computer-Human Interaction*, Vol 7, pp 114–138, 2000.
31. Xiaojin Zhu, Zhiting Xu, Tushar Khot, "How Creative is Your Writing? A Linguistic Creativity Measure from Computer Science and Cognitive Psychology Perspectives," Proceedings of the Workshop on Computational Approaches to Linguistic Creativity, CALC '09, pp 87–93, 2009.
32. François Mairesse and Marilyn A. Walker, "Controlling User Perceptions of Linguistic Style: Trainable Generation of Personality Traits," *Computational Linguistics*, Volume 37 Issue 3, pp 455–488, 2011.

APPENDIX

Stopping by Woods on a Snowy Evening by Robert Frost

Whose woods these are I think I know.
His house is in the village though;
He will not see me stopping here
To watch his woods fill up with snow.

My little horse must think it queer
To stop without a farmhouse near
Between the woods and frozen lake
The darkest evening of the year.

He gives his harness bells a shake
To ask if there is some mistake.
The only other sound's the sweep
Of easy wind and downy flake.

The woods are lovely, dark and deep.
But I have promises to keep,
And miles to go before I sleep,
And miles to go before I sleep.