Monkeying Around with Text[1]

Terry Butler
University of Alberta

Abstract

An excursus through the pre-history of a modern commonplace—monkeys at their typewriters, typing the text of Shakespeare—leads to a fundamental question for literary computing: when are two versions of a textual work "the same"? Software which can process and align two texts is described, and preliminary results shown. Suggestions are offered as to the value of being able to see synoptically texts which are nominally the same but may vary a little or a lot in their verbal details. A dynamic editorial environment is also mocked up, wherein readers can engage the difference between two texts with a toolkit which can transform one text in the direction of another. Examples are drawn from Robinson Crusoe*, Dickens,* The Prelude*, and the Hardy Boys.*

KEYWORDS: random text generation, text display, editorial tools

Monkeys

> If an army of monkeys were strumming on typewriters
> they *might* write all the books in the British Museum.

So wrote Arthur Eddington, the gifted British physicist and science popularizer, in his 1927 Gifford Lectures *The Nature of the Physical World* (72). This conceit (which Eddington advanced only as a contrast to how much *more unlikely* was spontaneous order in a simple physical system) has become a hardy perennial; but Eddington was not the originator of it—not by a long shot. As we take a guided tour of sightings of this trope, we will frame a question which is important for textual criticism: when is one text *the same* as another? What degree of difference is acceptable, and in what contexts? Two computer environments will also be showcased: one which visualizes the differences between two versions of the "same text"; and another which lets the reader do something about the differences.

Eddington may well have met these monkeys through the French scientist Émile Borel, who wrote in "Mécanique Statistique et Irréversibilité":

> Concevons qu'on ait dressé un million de singes à frapper au hasard sur les touches d'une machine à écrire et que, sous la surveillance de contremaîtres illettrés, ces singes dactylographes travaillent avec ardeur dix heures par jour avec un million de machines à écrire de types variés. ("Parable of the Monkeys")

In his *Elements of the Theory of Probability* Borel refers to the probability that "thousands of monkeys, randomly typing on typewriters, will reproduce exactly the contents of the National Library" (105; tr. Freund 60). This probability is expressed numerically as 10 with a negative exponent of more than 1 trillion zeros; a number which expresses both the certainty that the monkeys will eventually succeed, and utterly dwarfs any present conceptions of the size of the universe, or the length of time it might continue to exist.

Many readers will be familiar with post-Eddington adaptations of the trope: "The Library of Babel" by Jorge Luis Borges; Isaac Asimov's "The Monkey's Finger" (1953); and "Epicac" by Kurt Vonnegut Jr. in his 1958 short story collection *Welcome to the Monkey House*. My own initiation into hands-on humanities computing came through the delightful Computer Recreation puzzles published in the *Scientific American*. Brian Hayes, in "A Progress Report on the Fine Art of Turning Literature into Drivel," showed us how to build a random text generator, and, even more interesting, how to create random texts that were inflected by the style of a particular author.

But our tour heads out in the other direction, looking at earlier expressions of this conceit. The underlying idea is that all possible works and all possible variations of them could be systematically generated. The idea is beautifully captured in "The Universal Library," a story by Kurd Lasswitz which was anthologized in Clifton Fadiman's *Fantasia Mathematica*. Written in 1901, the story teases us with the mental construct of a library which contains all possible all works (texts of finite length with every possible character in every possible position), then delightfully blows it up by showing that the works would fill a far greater volume than the known universe:

> "What?" said Mrs. Wallhusen. "You say everything will be in that library? The complete works of Goethe? The Bible? the works of all the classical philosophers?"

"Yes, and with all the variations of wording that nobody has thought up yet. You'll find the long lost works of Tacitus and their translations into all living and dead languages. … all forgotten and undelivered speeches in all parliaments, the official version of the Universal Declaration of Peace, the history of the subsequent wars …" (Lasswitz 239)

With computers, we can translate this vastness of bulk to the intractable time it would take to generate such a series of works. A computer simulator came up with this result on August 4, 2004: after 42,162,500,000 billion billion years, one of the "monkeys" typed, "VALENTINE. Cease toIdor: eFLP0FRjWK78aXzVOwm)-';8.t . . ." The first 19 letters of this sequence can be found in "The Two Gentlemen of Verona" ("Infinite Monkey Theorem," *Wikipedia*).

But the monkeys (and their typewriters) are merely the modernist guise under which this famous trope now appears. We can trace it back, to the seventeenth century English divine Archbishop Tillotson. In *Answer to the Epicurean System*, he argues,

How often might a Man, after he had jumbled a Set of Letters in a Bag, fling them out upon the Ground before they would fall into an exact Poem, yea or so much as make a good Discourse in Prose? And may not a little Book be as easily made by Chance, as this great Volume of the World? (10)

And the monkey's precursors were spotted by Lemuel Gulliver, in The Grand Academy of Lagado, when he saw the Literary Engine:

The Professor then desired me to observe, for he was going to set his Engine at Work. The Pupils at his Command took each of them hold of an Iron Handle, whereof there were fourty fixed round the Edges of the Frame, and giving them a sudden turn, the whole Disposition of the Words was entirely changed. He then commanded six and thirty of the Lads to read the several Lines softly as they appeared upon the Frame; and where they found three or four Words together that might make part of a Sen-

tence, they dictated to the four remaining Boys who were Scribes. (Swift 148)

Modern variants of the trope often task the monkeys with producing the works of Shakespeare, who so often acts (as in this case) as a strange attractor for cultural extremes. But Shakespeare himself is not uninvolved, as he mentions the famous random text generator of classical times:

> TITUS. ... And, come, I will go get a leaf of brass,
> And with a gad of steel will write these words,
> And lay it by: the angry northern wind
> Will blow these sands, like Sibyl's leaves, abroad,
> And where's your lesson, then? Boy, what say you? (*Titus Andronicus* 4.1.103-108)

Dante says the same:

> Even thus the snow is in the sun unsealed,
> Even thus upon the wind in the light leaves
> Were the soothsayings of the Sibyl lost. (*Paradiso* 33:64-66)

Henry Thurston Peck explains: "The most famous Sibyl in antiquity was that of Cumae in Campania … whom Virgil represents as being visited by Aeneas" ("Sibyllae"). In Book 3 of the *Aeneid*, the oracle at Delos explains to Aeneas to whom he must go for guidance:

> Arriv'd at Cumae, when you view the flood
> Of black Avernus, and the sounding wood,
> The mad prophetic Sibyl you shall find,
> Dark in a cave, and on a rock reclin'd.
> She sings the fates, and, in her frantic fits,
> The notes and names, inscrib'd, to leafs commits.
> What she commits to leafs, in order laid,
> Before the cavern's entrance are display'd:
> Unmov'd they lie; but, if a blast of wind
> Without, or vapors issue from behind,
> The leafs are borne aloft in liquid air,
> And she resumes no more her museful care,

> Nor gathers from the rocks her scatter'd verse,
> Nor sets in order what the winds disperse.
> Thus, many not succeeding, most upbraid
> The madness of the visionary maid,
> And with loud curses leave the mystic shade. (*Aeneid* 3.441-452)

The most interesting aspect of random texts is not producing them (we have the services of that modern monkey, the computer, to do it for us), but dealing with the results. Who is going to read these texts? Or, an even more urgent problem, who is going to *proof*-read them? Here is how it might go (courtesy Bob Newhart):

> You know the idea … if you put an infinite number of monkeys, at an infinite number of typewriters, they would type all the great books. Now, they are going to type a lot of gibberish, too. So they would have to hire guys to check the monkeys to see if they were turning out anything worthwhile. … Look, I've got something: "To be or not to be … that is the *gezortenblatt* …".

The earliest incarnations of the species exhibit a healthy suspicion about the permanence of texts. What has been made permanent by writing has unexpectedly been disordered by chance and time, and the meanings which will be taken away are not the right ones. It is to this problem that humanism first turned its critical attention, and which remains a central theme in present day debate about the status of text and versions, and of their electronic simulacra.

The Editorial Program

The problem of evidence in literary studies is that we have too much of it. A literary work is a conspiracy between author and collaborators (co-writers, editors, printers, publishers) to create a stream of words, conventionally printed on paper, which we read for delight and profit. The difficulty arises when we turn, with a scientific sense of duty and armed with the cudgels of historical method, to investigate the crime, and to uncover the details of the process. The neat fiction that we are tracing back to a single original artefact, which represents the author's idealized intention, is often exploded.

In many actual cases, we find that the authorial situation is complicated: authors some times are sloppy, uncertain, distracted, and devious; editors are meddlers, and become unauthorized revisers; friends are well-intentioned but over-zealous re-writers; scribes try their hand at composition; compositors pad out lines from memory or from spite; printers and publishers shuffle sheaves and mismatch pieces. Add to all this the constant erosion of the historical record which time imposes. We are left with a situation, for many important literary works, that the words of the text are uncertain, in significant ways. Among many important instances that could be adduced, we have (in English) much of the medieval canon, including the *Canterbury Tales*; *King Lear* and *Hamlet*; the works of inveterate revisers like Coleridge and Wordsworth; collaborative creations like *The Waste Land*, and anonymous/pseudonymous creations like the multiple varieties of *Piers Plowman*. (For a catalogue that shows these examples are by no means exceptions to a rule, see Thorpe 35-36).

The History of the Tree

The textual evidence for many literary works is various, damaged, and unclear; as a consequence, there has been an age-old drive to create editions which purify the work from the ravages of time, and the remove the errors and incompetencies of meddlers (a.k.a. previous editors). The prime requisite for success in this enterprise was, and still is, a well-formed critical judgement on the part of the editor. "The establishment of the text is, normally, the comprehensive act of textual criticism. It is to this act that the textual critic brings all of the accumulation of all his knowledge and all of his skill and all of his experience" (Thorpe 202). However, an uneasiness with the subjective nature of these judgements has led to repeated searches for an editorial "logic chopper": some definite, objective process which introduce certainty to the judgements (and forestall future meddling). Scholars have exerted their ingenuity in the search for something like a mechanical process through which to run the evidence, to tame and reduce it to a convincing and single narrative, and to enshrine the result with the aura of permanence.

Textual criticism, which is the scholarly procedure for restoring texts to their original form, was first undertaken by the Alexandrian grammarians. During the Renaissance, coupled with renewed interest in the records of antiquity, talented scholars such as Erasmus drew attention to the issue of the variety of readings in classical texts, without solving it. It

was at this time that we see the origins of the genealogical method of editing (the grouping of manuscripts into families). Famous editions of *Catullus* by Scaliger (1577) and *Horace* by Richard Bentley (1711) showed the potential of the process in the hands of brilliant operators, but the results as method were very preliminary, due to a lack access to manuscripts, a poverty of information about their provenance, and the primitive state of language study at the time.

The genealogical method uses reasoning about the variants in the text to establish chronological relationships between the physical objects which embody these texts (the manuscripts). These connections are read genealogically (in the sense of a family tree). Having early success with the texts of classical literature, there have been repeated efforts to systematize the process and to develop procedures which are "foolproof" and "scientific." Karl Lachmann's editions of the Greek New Testament and *Lucretius* (1850) were important for solidifying the genealogical method and giving it both prestige and wide diffusion (Kristeller). In the early twentieth century, A.W. Pollard and R.B. McKerrow developed the foundations for a critical bibliography suited to print materials.

But the complexity of the object of study (the written word) precludes a systematic process which will be satisfactory in every case. The copying of manuscripts is an inherently error-prone business, and there is no clear line (*post facto*) between "writer" and "reviser," so the conviction which may be established (on the probabilities) about the relationships between the readings in manuscripts cannot be transferred in its entirety to the texts carried by them.

A classic counterpoint to the scientism of the genealogically school came from A. E. Housman. The prefaces to his *Manilius* (1903) and *Juvenal* (1905) ridiculed via *reductio ad absurdum* the scientific blinkers which forced proponents of the stemmatic method to print nonsense which they claimed was authorized by the "best text" decision they had taken. Housman's critical brilliance shone as a counterpoint: the *textus receptus* had become merely a starting point for a *tour de force* of the informed and imaginative critic. ("Textual Criticism," *Encyclopædia Britannica*).

Further doubts surfaced in the 1920s, when Bédier and Quentin demonstrated that the stemmatic method is suspect when dealing with works in the medieval textual tradition. Quentin's "rule of iron," W.W. Greg's *Calculus of Variants* and Paul Maas *Textual Criticism* all attempted to impose rigorous formalism on the process, but scholars have found the principles unworkable and overly simplistic.

The ambitions of textual criticism move toward the methodology of the sciences, and promise a purification of error and the removal of inauthentic accretions. But textual editing is not like cleaning an ancient art object; the changes that have been introduced by multiple stages of copying cannot in principle be distinguished from normal linguistic processes: summarization, modernization, and so on. Because a written text is embodied in a physical object which rarely announces its provenance in an utterly unambiguous fashion, there is an inherent dangerous circularity as scholars argue from word to artefact and then back to the words themselves.

Texts and Authority

Stemmatics and other tools of textual criticism were developed to adjudicate between various versions of a work, where the original (in the author's own hand, or a copy convincingly close to it) was usually not available or determinable. Modern works might be thought to be less problematic, but a similar issue arises. When one or more editions or variations appeared during the author's lifetime, which should be preferred?

Fredson Bowers and Thomas Tanselle discussed the operative principles in a famously informed debate which included important contributions from Peter Shillingsburg and ultimately Jerome McGann. Attention was focussed on "authorial intention" as a means of adjudicating between multiple versions. However, as James McLaverty shows, the notion of "intention" itself is not incontestable, and principles which were proposed as a means of curbing the editor's arbitrary activities, in fact promote a new kind of subjectivity, inducing the present-day editor to speculate on authors' motives and intentions (in the psychological sense of the word). The present state of affairs is not in fact much advanced beyond the "rigorous eclecticism" which was counselled by Thorpe in his 1965 PMLA article "The Aesthetics of Textual Criticism" (revised as the first chapter of *Principles of Textual Criticism*, 1972). Thorpe candidly admits that aesthetic judgements are always involved in editorial decisions about which reading is "better" or "right." More recent summaries and editorial guidelines (Foulet and Speer) have agreed.

The role of the editor as final arbitrator has been increasingly questioned. Greetham imagines the textual evidence being assembled into "reading kits" (242). And Jerome McGann, whose "The Rationale of Hypertext" explicitly inserts itself into the Bowers/Tanselle tradition,

highlights the possibilities of the liberation of the text from the bounds of the codex. The opportunities provided by electronic texts have been hinted at, but rarely realized in anything more than a busy paste-up of multiple HTML text windows.

Showcasing textual difference

The textual potential of texts in electronic form was imagined by Todd Bender as early as 1976. That potential is only now being thoroughly explored. The eulogies which were pronounced in the 1990s over "text on the screen" were premature: they had before them only the incunabula of electronic textual systems. The pessimism which characterizes Birkerts's *Gutenberg Elegies* is now passé.

In dealing with multiple versions of a text, we can already do what Jack Stillinger dismissively imagines in *Coleridge and Textual Instability*: "Some scholars think we shall ultimately solve this problem with computers, using programs such as hypertext that will allow us to store and retrieve all the versions of a work at once, and ... create new version by synthesizing existing ones" (140). "A computer screen is not an appropriate interface for the primary act of reading" (Miall 52): the evidence which supported this judgement has changed, and the verdict need no longer be sustained. There are no longer any technical, practical or theoretical reasons to prevent the computer screen from being a good, or even ideal, forum for the first reading of literary text.

To test run the enriched environment which e-reading provides, I have developed two complementary e-text systems. The first, a dynamic text display system, puts two texts into a visual relationship, and lets the reader as "viewer" explore the large-scale similarities and differences. The second, a hands-on textual toolkit, guides the reader (now empowered as "reviser") to change one text into another, attending to the smallest, character-by-character details of the two texts.

Dynamic text display

The examples of the dynamic text display system shown here focus on the most extreme cases: what is the maximum amount of difference which we will tolerate, and still grant that two texts are (in some meaningful way) the same? The measures of adequacy for non-fiction précis and abstracts are practical: a good précis is one which serves the information needs of

the searcher—it guides the searcher to the longer text in cases where that longer text meets the information needs; or it provides the sought after information. What criteria are pertinent for summaries of literary texts? In what way is a short précis a meaningful representation of a story or a long novel? What is included, what excluded, and how do the resulting two texts relate to one another?

1. Masterpieces of World Literature

The examples are drawn from English language prose texts, in order that the summaries be not different from the original in language and genre. The first example is a comparison of Dickens's *A Christmas Carol* with the familiar "potted summary" originally produced by Frank Magill under the title *Masterplots*.

How does the summary (at 6,500 characters, only 4% as long) relate to its original? (Please refer to Figure 1.) The dynamic text display loads and visualizes the two texts. The longer text is above; *A Christmas Carol* strung out like Christmas lights along the top of the screen. The rectangular icons are proportional to the length of each paragraph. When a mouse hovers over, the start of the paragraph is shown; with a mouse click, the reader is given access to the full text.

Below it is Magill's summary. The two texts are aligned by a simple algorithm: for each paragraph of the second text, which paragraph of the first text is it most like? (A constraint is applied to prevent these really short texts from being strewn across the original text with no reference to order. A rolling average over several paragraphs is used to align the two texts.)

By browsing the sections of the Masterplots summary which are attracted to Dickens's text (closer together means more words in the corresponding paragraphs are the same), we can discover how the summary works, and where it fails. Proper names can act as attractors, and the frequency of the central names in this story preclude them from being distinctive: the summary attaches itself to the story through people and place names that match at particular points in the two texts. Distinctive verbs of action also draw the two texts together.

At a finer level of verbal detail, the texts are different: in fact, surprisingly different for a purported summary of a story. The summary is characteristically written with a preference for the general over the specific: adjectives such as "frightened," "skimpy," and "generous" cover one

or more specific incidents which Dickens presents without editorial comment. In the summary mood, setting and the varied voices of narrator and the characters are blurred. This is the reason that there is relatively little exact verbal matching between the two texts, and the summary remains mostly separated from its target.

This summary tells us what *A Christmas Carol* is about, but it is not (in any meaningful way) equivalent in its effects to the original. "As the time approached for closing the office on Christmas Eve, Scrooge's nephew stopped in to wish him a Merry Christmas"—Magill, not Dickens, to be sure.

2. Robinson Crusoe

With Robinson Crusoe, we have an even more extreme contrast. The Masterplot summary is less than 1% of the total length of the novel; what strategies does it adopt in order to try to provide an adequate précis? (Refer to Figure 2.) The patient amassing of detail, Defoe's single most noticeable narrative technique, is unavailable to the précis.

The Masterplot summary is at least as much about what we remember the novel to be about, as what it actually says: "Robinson Crusoe was a castaway on a desert island." But these words do not appear in Defoe's text. He only once has Crusoe refer to the "desert island," and that when imagining his relative happiness in having been deposited there: "'O happy desart!'" (139). And "castaway" as a noun is not in his text.

Another source of verbal difference arises because the summary has secularized the tale. Defoe's extensive dramatization of Crusoe's conversion, and his moral and spiritual struggles, have been reduced to conventional "daily devotional periods."

This summary, along with two other retellings of the tale which are available in the online system, have some similarities with one another. All three tend to cluster around certain episodes, and in fact, all tend to reuse the same words and phrases from the original. We can see the emergence of "hot spots" in the text: areas which attract our attention (as being memorable or characteristic of the work). The hot spots are also attractors for the choice of chapter headings, running heads, and illustrations (which are other forms of abstraction and summary which do not come directly from the author).

3. The Hardy Boys

The Hardy Boys example (Figure 3) illustrates another limiting case. This is a visualization of *The House on the Cliff*: two texts, by the same (putative) author, with the same title, on the same premise, but written then rewritten 60 years apart. Here, as the paragraph by paragraph display makes clear, are two very different texts. The 1959 rewrite (revised again in 1987) completely changes the story—not only different character names, new figures, and new plot elements, but also a reordering of activities and changed motivations (Dixon 1927, 1959).

Discussion in the popular press has emphasized the supposed "dumbing down" of the original series, written by a Canadian, Leslie MacFarlane, under the pen name of Franklin W. Dixon. But a comparison like this can also expose other strategies at work, including both a modern avoidance of out-of-favour cultural stereotypes, and a wilful wallowing in other kinds of stereotypes.

These two texts, visualized here as poles apart throughout their length, are the same text only in special bibliographic and cultural senses.

An Editorial Difference Engine

The visualization environment presented above relies upon a word-by-word comparison of the two rival texts: a high degree of similarity draws the corresponding data points close together; dissimilarity forces them apart. This synoptic view is helpful to discern broad patterns of similarity and difference between two versions of the same work.

A complementary strategy can also be imagined: what if the reader (à la Espen Aarseth's concept of ergodic literature) is put to work to explore the differences between two texts, indeed, to do more than explore it, in fact to resolve it? The computational concept of "edit distance" (Nesbit) can be used to underpin a dynamic editorial environment in which readers become doers, and remake one text in the direction of another.

This computing environment has been concocted in part as a response to suggestions made in Peter Shillingsburg's "Polymorphic, Protean, Reliable, Electronic Texts," where he says "all new editions, even scholarly ones, are new, different, and additional to the earlier editions" (36). His thinking about texts takes us away from the naive position that there is an single, ideal, encoded digital representation of any literary work; rather, we can focus on the digital text as an opportunity to access and manipulate the text, and thereby learn something about both the digital instantiation and the work that it embodies.

Edit distance is defined as a mathematical measure of the difference between two strings of symbols (Baeza-Yates; Bickel). Numerically is it the sum of the weights given to the operations required to transform string *t1* into *t2*:

t1	t2
All that glisters is not gold,	All that glitters is not gold,
Often haue you heard that told.	Often have you heard that told.

The Merchant of Venice (Folio 1, 1623): Act 2, Scene 7

If the editorial operations *insert* and *delete* are defined, each with a weight of 1, then the editing transformation to t1 into t2 is:

delete 's'; insert 't'; delete 'u'; insert 'v'

and the edit distance between the two texts is 4. (Note that the insert and delete pairs in both cases occur at the same point in the string.) The source text could be marked up accordingly:

t1	t2
All that glis^t^ters is not gold,	All that glitters is not gold,
Often hau^v^e you heard that told.	Often have you heard that told.

The Merchant of Venice (Folio 1, 1623): Act 2, Scene 7

The potency of the edit distance concept for literary versions is that the operations defined do not have to be all given the same weight. In the simple example above, it may be thought that "v" for "u" is an orthographic variation rather than a substantive difference: if so, that specific operation (in the context of two particular versions of a work) could be given an appropriately lower weight (for example, 0.2). A consequence of this is that a text with readings "glisters / have" is closer to t2 (ed=0.2) than a text with readings "glitters / haue" (ed=1.0).

The editorial difference engine (Figure 4) is a live editing environment for exploring text difference. The display shows two versions of a work in the left and right hand panes. In the middle is the toolkit, which empowers the reader/rewriter to transform the text on the left into the text on the right.

The two simple edit operations, *insert* and *delete*, are all that are needed to accomplish the transformation. However, a more intuitive result can be achieved if compound operations are permitted. *Replace* is defined as a deletion followed by an insertion at the same point; *transpose* is defined as deletion, a move of the insertion point, followed by an insertion of the deleted character.

When the user positions the cursor at a point in the text at which an edit operation can occur, the corresponding edit tools (represented here graphically by the familiar word processing icons) are enabled. Clicking on a tool performs the edit action. The text in the left pane is changed to show the result.

For example, highlighting "my" on line 479 in *The Prelude* 1805 pane activates the replace tool; clicking the tool changes the text in the left pane:

In this ~~my~~^our^ deep devotion. Fare thee well.

and the edit distance value between the texts will decrease. (If weights of 1 are given to each deletion and insertion of single character, the operation just performed reduces the edit distance by 5. But it might be more natural in this case to assign a maximum weight to the higher level operations, such as a single word replacement.)

Although the edit distance is a measure of the character by character difference between two strings, the implementation here allows for operations at a higher level of generality. Words as well as lines can be selected, in which case the corresponding edit changes can be enacted upon them. In this way, the lacuna of two lines in the 1805 pane can be rectified by placing the cursor on the blank line immediately below line 476, and clicking the icon for *insert*; the corresponding line from the 1850 version is inserted.

Even more complex transformations are also possible. String patterns, expressed as regular expressions, could be defined and given appropriate weights.

The system keeps track of a list of states of the text being transformed, allowing at all times the ability to undo an operation, or to step back to a previous stage of the process. Only valid operations are possible (ones which reduce the edit distance between the two texts). It is hoped that the reader will learn something about text as they consider in detail the nature of the differences and strategies for removing them. For example:

The most intense of Nature's worshippers,

becomes

The most assiduous of her ministers,

in the context of "Nature's temple." In light of Wordsworth's relation with Coleridge, who is at the literal level the subject of this passage, one might construe the revision of 1850 as both removing a description which might have given offence ("intense Nature worshipper") while simultaneously delivering a back-hand compliment ("assiduous") to a supposed intuitive poet.

Conclusion

These two electronic text environments are offered as playground equipment, with the hope they will prove interesting to students and scholars who wish to explore the potential of literary texts in electronic storage. The primary responsibility of a scholarly editor in the era of print was to weigh all the evidence and render a final verdict. This editors still must do, but the new web-based tools provide an opportunity to also give out all the evidence in accessible forms. It is to be hoped that this opportunity will soon be seen by editors as a responsibility.

Print editors struggled to square the circle when making editions of works which were repeatedly revised (Gill, "The Question of Text"; Parrish): how to provide access to all the evidence while at the same time providing a "single" text to print for reading purposes? Electronic environments can be a means to this end, especially if the reader is offered tools to cooperate in finding solutions.
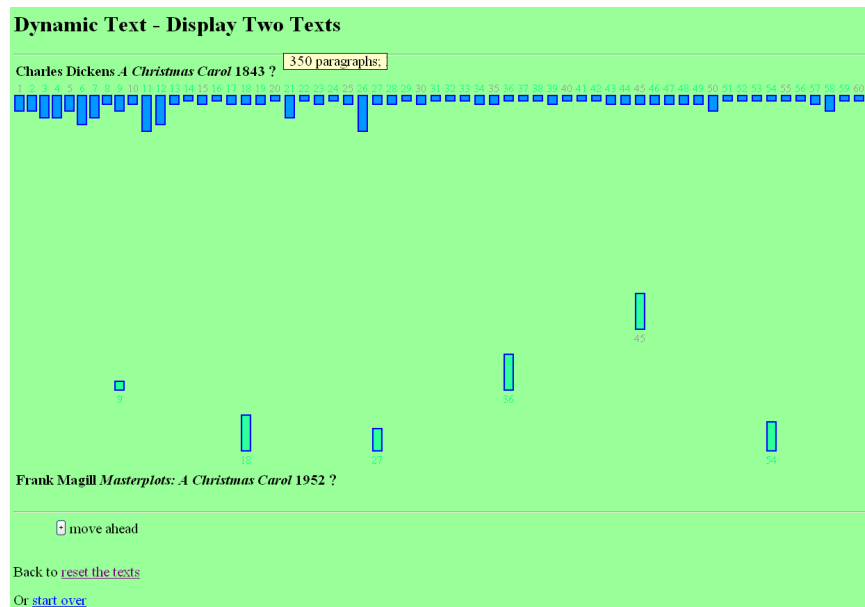
Figures

Figure 1: dynamic text display showing *A Christmas Carol*
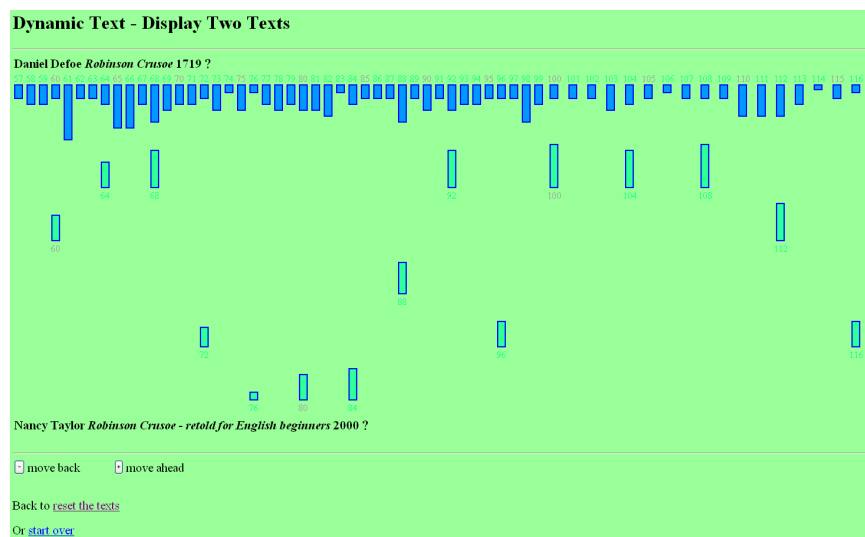with the Masterplots précis of it

Figure 2: dynamic text display showing *Robinson Crusoe* 1719
and a modern, simplified language retelling of the same story
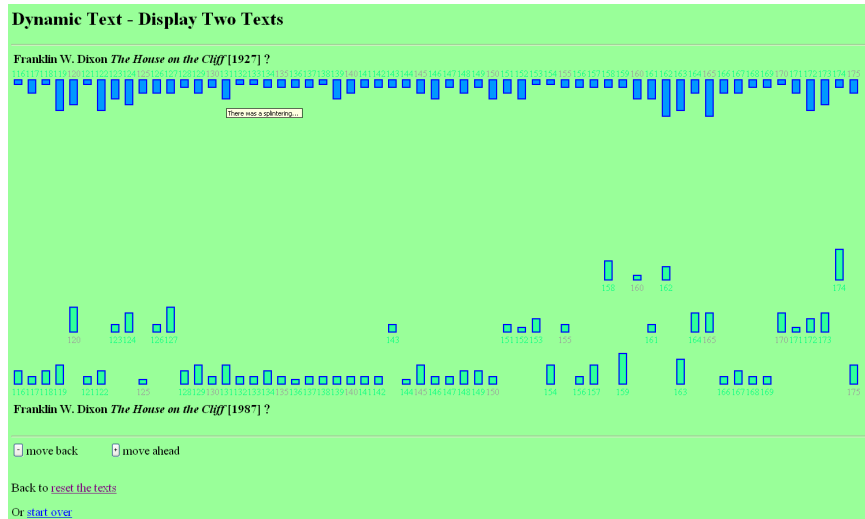
Figure 3: dynamic text display showing the first version
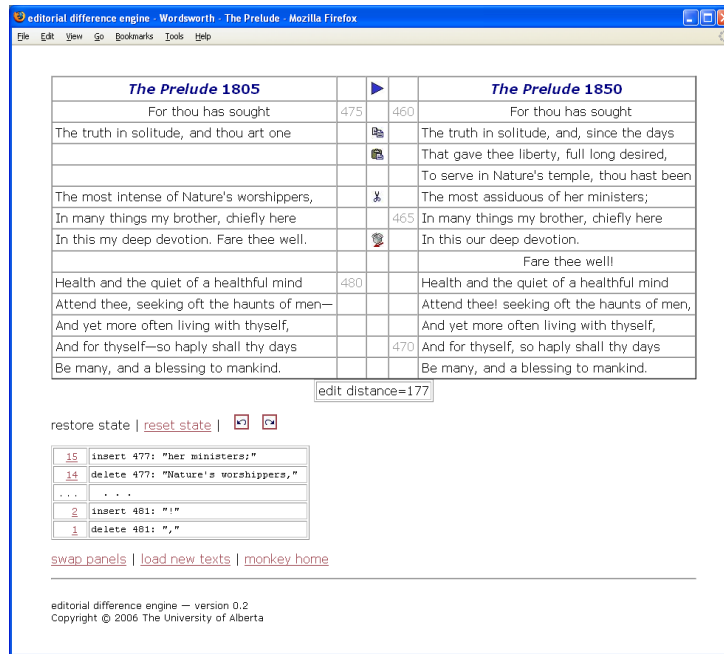and a later rewrite of the Hardy Boys story *The House on the Cliff*

Figure 4: the editorial difference engine, mocked up with
two versions of a passage from *The Prelude*

Notes

[1] [Editors' note] An earlier version of this article was published in *Computing in the Humanities Working Papers* (January 2007): <http://www.chass.utoronto.ca/epc/chwp/CHC2005/>.

Works Cited

Aarseth, Espen. *Cybertext: Perspectives on Ergodic Literature*. Baltimore: Johns Hopkins UP, 1997.

Alighieri, Dante. *Paradiso*. Trans. John D. Sinclair. Oxford: Oxford UP, 1939.

Asimov, Isaac. "The Monkey's Finger." *Startling Stories* Feb. 1953: 77-83.

Baeza-Yates, Ricardo A. "Algorithms for String Searching: A Survey." *ACM SIGIR Forum* 23.3-4 (1989): 34-57.

Bender, Todd K. "Literary Texts in Electronic Storage: The Editorial Potential." *Computers and the Humanities* 10 (1976): 193-199.

Bickel, Michael Allen. "Automatic Correction to Misspelled Names: A Fourth-Generation Language Approach." *Communications of the ACM* 30.3 (1987): 224-228.

Birkerts, Sven. *The Gutenberg Elegies: The Fate of Reading in an Electronic Age*. Boston: Faber and Faber, 1994.

Bowers, Fredson. "Greg's 'Rationale of Copy-Text' Revisited." *Studies in Bibliography* 31 (1978): 90-161.

Borel, Émile. *Élémentes de la Théorie des Probabilités*. Paris: Ablin Michel, 1950.

---. *Elements of the Theory of Probability*. Trans. John Freund. Englewood Cliffs, NJ: Prentice Hall, 1965.

---. "Mécanique Statistique et Irréversibilité." *J. Phys*. 5e.3 (1913): 189-196. [Reference given without specific page number on "Parable of the Monkeys" webpage, 13 Sept. 2006, <http://www.angelfire.com/in/hypnosonic/Parable_of_the_Monkeys.html>.]

Borges, Jorge Luis. "The Library of Babel [English translation]." 13 Sept. 2006. <http://jubal.westnet.com/hyperdiscordia/library_of_babel.html>.

Crasson, Aurèle. "Pour une Lecture Ouverte du Texte Électronique." *Literary and Linguistic Computing* 15.1 (Apr 2000): 109-113.

Defoe, Daniel. *Robinson Crusoe*. Oxford: Oxford UP, 1972.

---. *Robinson Crusoe*. Retold by Nancy Taylor. Penguin Readers Level 2. Upper Saddle River, NJ: Pearson Educational, 2000.

---. *Robinson Crusoe in Words of One Syllable*. Retold by Mary Godolphin. New York: McLoughlin Bros., 1882.

Dickens, Charles. *A Christmas Carol*. London: Chapman & Hall, 1843.

Dixon, Franklin W. *The House on the Cliff*. 1929. New York: Grosset & Dunlap, 1987.

Eddington, Arthur. *The Nature of the Physical World*. Cambridge: Cambridge UP, 1928.

Foulet, Alfred and Mary Blakely Speer. "A Historical Orientation." *On Editing Old French Texts*. Ed. Alfred Foulet and Mary Blakely Speer. Lawrence, KS: Regents Press of Kansas, 1979. 1-39.

Gill, Stephen. "The Question of Text." *Review of English Studies* 34 (May 1983): 172-190.

---. *William Wordsworth: The Prelude*. Cambridge: Cambridge UP, 1991.

Greg, W. W. *The Calculus of Variants: An Essay on Textual Criticism*. Oxford: Oxford UP, 1927.

---. "The Rationale of Copy-Text." *Studies in Bibliography* 3 (1950): 19-36.

Greetham, David C. *Theories of the Text*. Oxford: Oxford UP, 1999.

Hanley, Keith. "Crossings Out: the Problem of Textual Passage in The Prelude." *Romantic Revisions*. Ed. Robert Brinkley and Keith Hanley. Cambridge: Cambridge UP, 1992. 103-135.

Hayes, Brian. "Computer Recreations: A Progress Report on the Fine Art of Turning Literature into Drivel." *Scientific American* 249.3 (1983): 18-28.

Housman, A. E. "The Application of Thought to Textual Criticism." *Selected Prose*. Ed. John Carter. Cambridge: Cambridge UP, 1961. 131-150.

Huitfeldt, Claus. "Multi-Dimensional Texts in a One-Dimensional Medium." *Computers and the Humanities* 28 (1995): 235-241.

"Infinite Monkey Theorem." *Wikipedia, the Free Encyclopedia*. 13 Sept. 2006. <http://en.wikipedia.org/wiki/Infinite_monkey_theorem>.

Kristeller, Paul Oscar. "The Lachmann Method: Merits and Limitations." *TEXT: Transactions of the Society for Textual Scholarship* 1 (1981): 11-20.

Lasswitz, Kurt. "The Universal Library." Trans. Willy Ley. *Fantasia Mathematica; Being a Set of Stories, Together With a Group of Oddments and Diversions, All Drawn From the Universe of Mathematics*. Ed. Clifton Fadiman. New York: Simon and Schuster, 1958. 237-243.

Maas, Paul. *Textual Criticism*. Trans. Barbara Flower. Oxford: Clarendon P, 1958.

Magill, Frank N. *Masterpieces of World Literature in Digest Form*. [Originally published as Masterplots.] New York: Harper, 1952.

McGann, Jerome J. *A Critique of Modern Textual Criticism*. Chicago: U of Chicago P, 1983.

---. "The Rationale of Hypertext." *Electronic Text: Investigations in Method and Theory*. Ed. Kathryn Sutherland. Oxford: Clarendon P, 1997. 19-46.

McKenzie, D. F. *Bibliography and the Sociology of Texts*. Cambridge: Cambridge UP, 1999.

McLaverty, James. "The Concept of Authorial Intention in Textual Criticism." *The Library* 6.2 (1984): 121-138.

Miall, David S. "Rethinking English Studies: The Role of the Computer." *Humanities and the Computer: New Directions*. Ed. David S. Miall. Oxford: Clarendon P, 1990. 49-60.

The Monkey Shakespeare Simulator. 13 Sept. 2006 <http://user.tninet.se/ecf599g/aardasnails/java/Monkey/webpages/>. [Editors' note: the site may be accessed using the Internet Archive (23 Aug. 2007), <http://web.archive.org/web/20060207100939/http://user.tninet.se/~ecf599g/aardasnails/java/Monkey/webpages/index.html>]

Nesbit, John C. "The Accuracy of Approximate String Matching Algorithms." *Journal of Computer-Based Instruction* 13.3 (1986): 80-83.

Newhart, Bob. "An Infinite Number of Monkeys." *The Button Down Mind Strikes Back*. LP. Warner Bros., 1960.

"Sibyllae." *Harper's Dictionary of Classical Antiquities*. Ed. Henry Thurston Peck. 13 Sept. 2006 <http://daedalus.umkc.edu/sandbox/perseus/harper/page.10085.a.php>.

Parker, Hershel. *Flawed Texts and Verbal Icons: Verbal Authority in American Fiction*. Evanston, IL: Northwestern University Press, 1984.

Parrish, S. M. "The Whig Interpretation of Literature." *TEXT: Transactions of the Society for Textual Scholarship* 4 (1988): 343-350.

Poundstone, William. *The Recursive Universe: Cosmic Complexity and the Limits of Scientific Knowledge*. New York: William Morrow, 1985.

Quentin, Henri. *Essais de critique textuelle*. Paris: Picard, 1926.

Ross, Charles L. "The Electronic Text and the Death of the Critical Edition." *The Literary Text in the Digital Age*. Ed. Richard J. Finneran. Ann Arbor, MI: U of Michigan P, 1996. 225-231.

Internet Shakespeare Editions. University of Victoria. 15 Sept. 2006 <http://ise.uvic.ca/>.

Shakespeare, William. *The Complete Works of William Shakespeare*. Garden City, NY: Garden City Books, 1936.

Shillingsburg, Peter L. "Polymorphic, Polysemic, Protean, Reliable, Electronic Texts." *Palimpsest: Editorial Theory in the Humanities*. Ed. George Bornstein and Ralph G. Williams. Ann Arbor, MI: U of Michigan P, 1993.

---. *Scholarly Editing in the Computer Age: Theory and Practice*. Athens, GA: U of Georgia P, 1986.

Siskin, Clifford. "Working The Prelude: Foucault and the New History." *The Prelude*. Ed. Nigel Wood. Philadelphia: Open UP, 1993. 98-124.

Stillinger, Jack. "Textual Primitivism and the Editing of Wordsworth." *Studies in Romanticism* 28 (1989): 3-28.

---. *Coleridge and Textual Instability: The Multiple Versions of the Major Poems*. Oxford: Oxford UP, 1994.

Swift, Jonathan. *Gulliver's Travels and Other Writings*. Ed. Louis A. Landa. Boston: Houghton Mifflin, 1960.

Tanselle, G. Thomas. "The Editorial Problem of Final Authorial Intention." *Studies in Bibliography* 29 (1976): 167-211.

---. "Recent Editorial Discussion and the Central Role of Editing." *Studies in Bibliography* 34 (1981): 23-65.

Tetreault, Ronald. "New Models for Electronic Publishing." *Literary and Linguistic Computing* 16.1 (2001): 189-198.

"Textual Criticism." *Britannica Online*. 15 Sept. 2006 <http://www.britannica.com/eb/article-9108631>.

Thorpe, James. *Principles of Textual Criticism*. San Marino, CA: Huntington Library, 1972.

Tillotson, John. *Maxims and Discourses Moral and Devine: Taken from the Works of Arch-Bishop Tillotson, and Methodiz'd and Connected*. London, 1719. 10-11.

Vinaver, Ernest. "Principles of Textual Emendation." *Studies in French Language and Mediaeval Literature, Presented to Professor Mildred K. Pope, By Pupils, Colleagues, and Friends. 1939*. Freeport, NY: Books for Libraries P, 1969. 351-369.

Virgil. *Aeneid*. New York: P.F. Collier & Son, 1909.

Vonnegut, Kurt, Jr. "Epicac." *Welcome to the Monkey House*. New York: Delacourt, 1958. 268-75.

Wordsworth, Jonathan. "Revision as Making: The Prelude and its Peers." *Romantic Revisions*. Ed. Robert Brinkley and Keith Hanley. Cambridge: Cambridge UP, 1992. 18-42.

Wordsworth, William. *The Prelude: 1799, 1805, 1850*. Ed. Jonathan Wordsworth, et al. New York: Norton, 1979.

Afterword: *Dé-Partages*

Jean-Claude Guédon
Université de Montréal
jean.claude.guedon@umontreal.ca

Abstract

A conclusion to the special issue of TEXT Technology *on "Digital Humanities and the Networked Citizen." The diversity of studies that fall under the heading of "text technology" is well illustrated by the articles making up this special issue. However, one special theme recurs throughout: familiar classifications and well-established boundaries all seem to be dissolving or shifting. In one way or another, all the articles in this issue address the question of a changing "order of things" and they help us understand what some of the implications of these shifts may be and may mean. The network image also seems to run through all of these articles because, as a concept, a network allows simultaneously to link and to separate. However, one question remains that all these articles seem silently to point to: can atomistic entities adequately essentially anchor a networked structure? The article answers in the negative and suggests that a phonemic metaphor would be far better to address the mysterious conjunction of links and nodes.*

KEYWORDS: network, phoneme, technology, free software, Bourdieu, Foucault, identity

If anyone ever had any doubts about the variety of approaches that could fall within the scope of this journal, they will be laid to rest by the five articles that appear in this issue. Consider an array of texts that range from free software in India to the reshaping of citizenship in a networked world. Add to this considerations about the feminine body as a form of human(e) technology, and about the fabled monkeys typing away at Shakespeare's works. Spice up the whole dish with an intriguing exploration of inter-mediatic exchanges between video games and films. *Voilà!* as some of you might be wont to say in English.

It was, therefore, with some trepidation that I began approaching this challenge, feeling very much like that equally fabled blind man who touches the tail, the trunk and other parts of the elephant, except that a tinge of envy did inhabit my soul: he, at least, knew that he was identifying