How to Do Things (To Texts) with Computers

Stephen Ramsay

February 27, 2020

Early on in my teaching career—while I was still in graduate school—a student in one of my Shakespeare classes came up to ask me a question that (I am embarrassed to say) puzzled me. I had just got done explaining to the class that I wanted them to write a critical essay on a Shakespeare play. I told them how long it should be, that it had to be double-spaced, listed some sources they might look at, and told them that in my infinite largesse, I had set the due date before (as opposed to after) spring break. I thought I had been quite clear about the requirements, but one student came up to me afterward looking very distraught. "I don't understand!" she said. "What are we supposed to *do*?"

Now, I was a quite inexperienced teacher at that time, but nonetheless, you might suppose that I had a snappy answer for her. I had written lots of critical essays at that point; surely I could explain how to do it in a way that goes beyond the rather banal matter of how to set the line spacing correctly. But I was a bit taken aback by the question, and I stammered something in response that really wasn't very useful from a pedagogical standpoint—"You know, write an essay about Shakespeare"—and moved on to my next class.

I wasn't happy with that answer; I knew that I had failed to take advantage of a "teaching moment," as they say. So I set myself the task of thinking long and hard about what it is I'm asking my students to do when I tell them to write a critical essay or give "a reading" of a work of literature—which, of course, is a matter very closely related to the important issue of what *I'm* doing when *I* do these things (another thing I should presumably know about).

Her question, of course, was a lot smarter than my answer. To begin with, she was asking a perfectly legitimate question. Whatever a critical reading is, it seems qualitatively different from good old-fashioned reading—the sort of thing you do on the beach or before drifting off to sleep at night. But the languge that surrounds that discourse ("Did you like it?", "What's it about?") are emphatically not the terms of literary-critical discourse. If you want to fail an English class, all you have to do is either engage in aesthetic rapture or recapitulate the plot. It's not that these are bad ways to talk about books, or even unsophisticated ways; if anything, they represent the natural ways. But the critical mode is different, and we're not born knowing how to do it.

My answer, moreover, was deeply misleading. She had asked what she was supposed to *do* with the text; I answered that she was supposed to do something *about* the text. Her question, whether she realized it or not, had struck upon the Greek root of the word *drama*—δράω, to do, or act—which seems a more appropriate way to talk about critical engagement (particularly with dramatic works). I had made it sound as if I wanted her to make an arrest ("you better do something about that Shakespeare character!").

I eventually came up with a way to explain what I wanted her to do (and what I've been doing these many years). I told her to study the play until she saw some non-obvious pattern, and then explain to me why I should see that pattern as well. It's a practical definition, but one that I think captures the essential difference between ordinary reading and critical reading. "I liked it" isn't a pattern; the plot is a pattern, but the bare "facts" of the plot are obvious. Critical reading doesn't explain what happened in *The Comedy of Errors*, but instead tries to show us that, for example, all the mistaken identities in the play (the pattern of identity, if you like) reveal the work to be both a modification of the traditional terms of Plautine farce and a commentary on the emergence of a distinct vision of the self in the English Renaissance. There are a finite number of "plots" in *The Comedy of Errors*; there are an infinite number of suggestive patterns and a correspondingly infinite number of possible explications.

Such maneuvers are not, of course, limited to literary criticism. Pattern and explanation inform the rhetoric of humanistic inquiry in all its forms. The historian doesn't tell us *that* Napoleon lost the battle of Waterloo, but that he lost it because the Duke of Welling-

ton was a talented coalition general who was able to create more robust communication lines across wider distances than his opponent. The art critic, likewise, doesn't enumerate the philosophers represented in Raphael's *School of Athens*, but shows us how the various groupings reflect the Renaissance vision of the organization of knowledge. Philosophy, perhaps more than any other discipline, exploits the power of pattern. Socrates wins his arguments not simply by saying that such a thing is right, but by demonstrating the patterns of his interlocutors' wrongness. The primacy of pattern also helps to explain why many believe that mathematics should be considered a humanities discipline, as it was in previous centuries. Few definitions capture its essence as well the one offered by the mathematician Keith Devlin: "mathematics is *the science of patterns*" (1).

I relate this story, and advance the framework which followed from it, because I believe it explains why that branch of digital humanities concerned with literary criticism and text analysis has largely failed to penetrate the mainstream of scholarly discourse in the humanities. In explaining itself—and indeed, in carrying out its own methodological project—it has put forth the same weak answer I gave to my student years ago. Against the hermeneutical injunction to do something with the text—discover patterns we might form into critical explanations—it has instead chosen to do something *about* the text. In its most extreme posture, it has purported to provide empirical evidence for why a particular reading might be good or bad (as in the case of John Burrows and D. H. Craig's attempts to find quantitative evidence for why Romantic dramas aren't as good as Renaissance ones). More often, it brings the computer's glorious ability to discover patterns to the water's edge of explication, but goes no further. Burrows and Craig believe the chief point to be drawn from their scatter plots and correlation matrices is that none of what they found contradicts the findings of earlier critics. Louis Milic believes that "The low frequency of initial determiners, taken together with the high frequency of initial connectives, makes [Jonathan Swift] a writer who likes transitions and made much of connectives" (Milic 1). Susan Hockey appears to offer such tautologies as among the principal virtues of computer-assisted work: "the computer is best at finding features or patterns within a literary work and counting occurrences of those features" (Hockey 66).

It is difficult not to read, in these refusals to engage fully with the hermeneutical process of pattern formation and explication, a subtle, inchoate desire to assuage fears of a mechanized literary criticism, or alternatively, a machinic instantiation of Wordsworth's famous quip: a dissection that amounts to a murder. Such fears, of course, go back (like computing in the humanities) to the days of punch card machines, when many of the anxieties concerning computers and automation were cathected onto the punch card itself: "Do not fold, spindle, or mutilate" (Lubar 44). Yet folding, spindling, and mutilating are at once the basis of computation (inasmuch as computation nearly always requires an alternate arrangement of the data) and the means by which we come to locate the patterns upon which reading depends. Why, then, this refusal to let computers lead us fully—without apology and without tentativeness—into interpretation and explanation?

Let's put that differently: What would computer technology, with its apparent disposition toward unerring processes and irrefragable answers, look like if it were loosed from the strictures of the irrefragable and allowed to become a bionic extension of our ability to deduce patterns and a launching pad for our bold attempts to explain those patterns? I believe we would have something that looks a lot less like the digital humanities' ironic alignment with experimental science, and a lot more like the implicitly humanistic methods of mathematics. An example will help to illustrate what I mean.

The mathematical discipline known as algebraic topology was inaugurated by the great Swiss mathematician Leonard Euler (1707–1783), who, in 1735, proposed a solution to what has come be known as the "Bridges of Königsberg" problem.
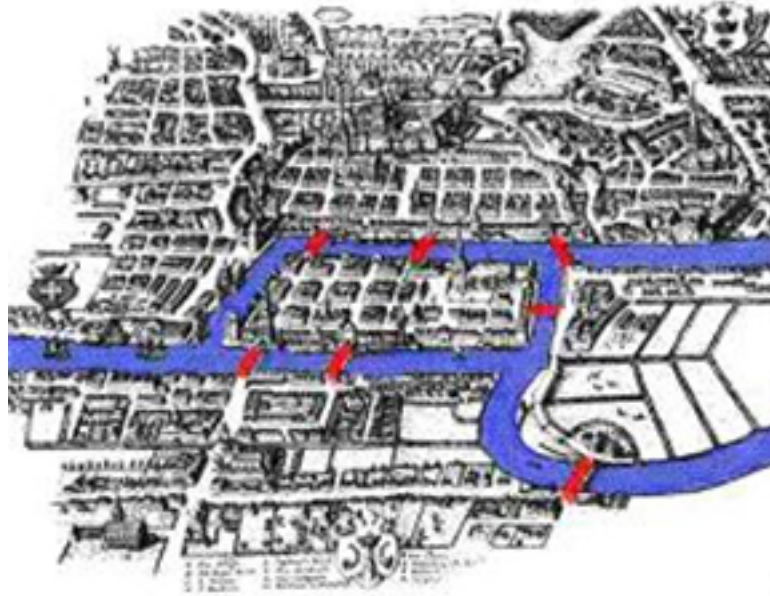
Königsberg, a small Prussian town during the eighteenth century, was divided into four sections (including one island) by the river Pregel.[1] The four regions were connected to one another by seven bridges, and the townspeople, who were fond of taking walks about the city on Sunday afternoons, wondered if it was possible to wander about the town crossing each bridge only once and end up back where you started (Rosen 693). The problem, it turns out, is not as easy as it looks—particularly if you want to *prove* that there is (or is not) a way to cross the bridges this way. The essay in which Euler proposes his solution, "Solutio Problematis ad Geometriam Situs Pertinentis" is one of the great masterpieces of mathematical exposition in part for the deftness with which Euler distills the complexity of the problem down to a simple problem in combinatorics.
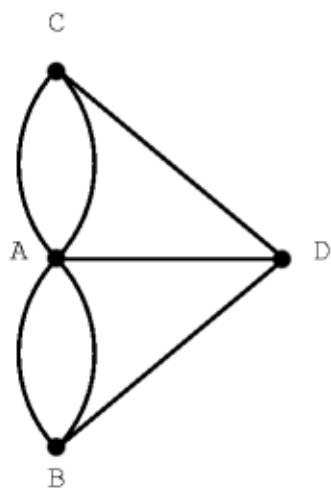
Euler begins by creating a simple schematic diagram of the regions which the river divides:

---

[1]Königsberg is now the Russian city of Kaliningrad on the Baltic Sea. See Trudeau, works cited.
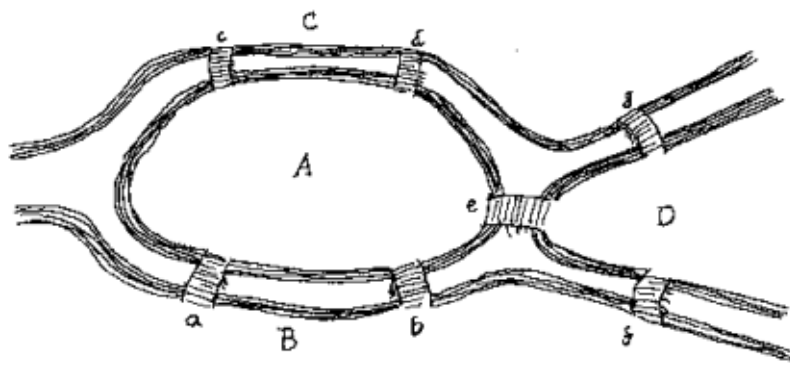
In a further attempt to remove all interference from the space of the problem, Euler collapses the regions and crossings into points and lines:

Such re-presentations of the data constitute one of the basic means by which we arrive at meaning—in literary criticism as surely as in mathematics. By distilling the problem down, Euler is able to see that the problem of the Königsberg bridges may be envisioned as a problem not of the bridges, but of the land masses which they connect. With his vision trained on the land masses, Euler is able to see the obvious fact that in order for someone to trace a continuous line through every point (without crossing over the same line more than once), every point would have to have a line leading toward it and a different line leading away from it—or, if there is more than one way onto the land, several pairs of such lines. The starting and end points may be an exception to this, but even in this case, the graph could only have (at most) two

points with an odd number of lines leading away from them.[2] The graph makes it clear that this is not the case with the bridges of Königsberg. The proposed journey is impossible (Biggs et al. 5). This result is further generalizable and is one of the basic theorems of graph theory: "a connected multigraph has an Euler circuit if and only if each of its vertices has even degree" (Rosen 478).

Euler's solution is an elegant one; despite the difficulty of the problem, once it is seen in a certain light (a certain arrangement), it strikes us with the force of the obvious, and thus obviates the need for a prohibitively exhaustive brute-force search through possible solutions. The real interest of Euler's solution for our purposes, however, lies in the fact that his method moved from one arrangement to another until a pattern enabled the insight necessary for a solution—an insight that was difficult for the pedestrian (or even the map reader) to see. However, the pattern itself is not the solution. It is instead the means by which Euler leads himself to the most useful interpretation of the problem.

In conceiving of the relationship between computers and humanistic study, it may be disingenuous to ask how we can create programs that can do literary criticism, philosophy, or historical analysis—not because we lack the elusive "strong AI" these tasks would presumably require, but because the question presupposes that doing these things is entirely a matter of interpretation. Interpretation is, to be sure, the *sine qua non* of humanistic inquiry, but it is only part of the process (and, we might say, the latter part). The other part is more serendipitous and ludic—closer to the following of a hunch or the formulation of a strategy. One wonders how much time Euler spent tracing his pencil over the map of Königsberg, trying this pattern and then that one, looking for something that would lead to the crucial moment of vision. We long for a machine that can give us the hermeneutical equivalent of the Euler circuit, but perhaps our efforts would be better directed toward creating a machine that can help us doodle. In other words, a machine that can assist at the moment when critical engagement leads to critical insight.

---

[2]Of course, the starting and end points could be the same, in which case every point would have to have an even number of lines leading away from it.

I've spent the last few years building such machines, and I'd like to demonstrate a more recent example, which, as luck would have it, doodles with a Shakespeare play.

It is customary, in interpreting Shakespeare's *Antony and Cleopatra*, to think of Alexandria in terms of license, nature, and femininity, and Rome in terms of stoicism, stability, and masculinity. With this binary overlay in place, many have come to regard the play as a political and erotic negotiation between two worlds personified by the Egyptian characters on the one hand, and the Roman characters on the other. It is a useful pattern, which has facilitated countless classroom discussions and yielded hundreds of critical articles.[3]

But what if we wrote a program that could automatically do to a Shakespeare play what Euler did to Königsberg:[4]

---

[3]This approach to the play goes back at least to Hazlitt's *Characters in Shakespeare's Plays* (1817). Harley Granville-Barker's *Prefaces to Shakespeare* (1927), in which the opposition of Rome and Egypt is put forth as the main organizing principle of the play's structure, might be said to inaugurate the modern discussion. See Granville-Barker, works cited.

[4]The graphs in this paper were generated using `StageGraph`, a custom-built program which uses `GraphViz`—an open source tool developed at AT&T Research—for graph layout and formatting.
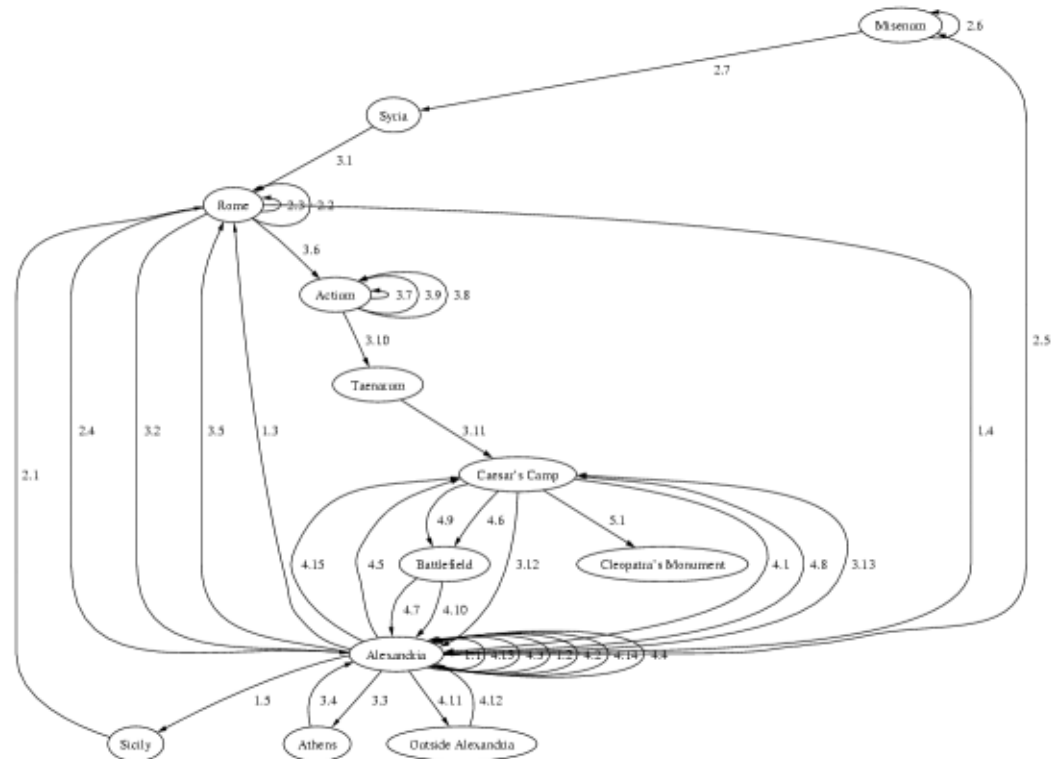
We arrive at something quite different from the obligatory maps of the ancient world which often accompany editions of this play. Unlike a geographical representation, this arrangement does not occur in metric space.[5] The lines in the graph (each of which is labeled with the appropriate act and scene numbers) represent the passage in the space of the drama from one location to the next, but the locations are more abstract. Scene locations which have more adjacent edges than others are said to be (in graph theoretical terms) vertexes of higher degree. Alexandria and Rome have, predictably, the highest degrees (with Alexandria displaying the highest number of adjacent scenes). Cleopatra's monument, which is both the last scene of the play and the only scene which takes place in that location, is a "pendant node" or vertex of degree one.

To read *Antony and Cleopatra* is to be aware that Rome and Alexandria are not the only settings in which the action occurs, but it is difficult to appreciate the range of locations in which the play occurs (Sicily, Misenum, Syria, Athens, Actium, Taenarum, and several places in and around Alexandria) and to perceive the amount of stage action they take up. The graph, by contrast, allows us not only to see the relationships among these other scenes more clearly, but to perceive them in terms of the play's general movement "between" Rome and Alexandria. Three scenes lead toward Rome (Sicily, Misenum, and Syria) and several away from Rome (Actium, Taenarum or "undisclosed," Caesar's Camp, and an indeterminate battlefield).[6] This, of course, does not correspond precisely to the chronology of the play, which for the reader, stands as the dominant organizing structure. There is, for example, no central Roman scene dividing the early events of the play from the Battle of Actium in the latter half of Act 3. The scene set in Sicily (2.1) passes to three subsequent scenes in Rome (2.2–4), but from there back to Alexandria (2.5) before passing through Misenum (2.6–7) and Syria (3.1) on the way to Rome again in 3.2.

The division of the play into scenes is, of course, the product of centuries of pattern-recognition beginning with Nicholas Rowe's 1709 edition of the play; the 1623 Folio has no act or scene division, nor even indications of setting. The play, in other words, comes to us as something already "patterned" by centuries of reading and commentary. In our case, the divisions have a purpose even more efficacious than what is required by the exigencies of reading or performance, since they constitute the data points of a representation that illustrates other patterns.

All of the scenes to which the graph draws our attention are sites of confusion and uncertainty among the characters and serendipitous changes in the action. Scenes from the earlier movement toward Rome are heavily laden with dramatic irony—a device used infrequently in this play, but significantly in these scenes and in Cleopatra's message to Antony reporting her death in 4.14. As 2.1 begins, Pompey is in Sicily making plans for war based on the knowledge that Antony remains in Egypt when Varrius reports as "most certain" the fact that "Mark Antony is every hour in Rome

---

[6]Most editions of the play follow Plutarch and represent the undisclosed scene as taking place in Taenarum.

/ expected" (ll. 28–31). 2.6 begins in Misenum with the certainty of war, but ends with the promise of a celebrated peace. 2.7 celebrates the agreement between Pompey and the triumvirate (not on the hill of Misenus, but displaced further "between" in a galley at sea). Throughout that scene we are given knowledge of various disjunctions between the actual and the apparent. All seems convivial, but we come to find out that Menas is plotting the murder of the triumvirate. Lepidus believes he is learning about the nature of crocodiles, when in fact he is merely hearing tautologies ("It is shaped, sir, like itself, and it is as broad as it hath breadth..." (ll.42–3)). Pompey believes he has the loyalty of Menas, but the latter reveals to us in an aside his intention to "never follow thy palled fortunes more" (l. 82).

The scenes corresponding to the nodes leading away from Rome, by contrast, emphasize the characters' full awareness of realities alternative to the apparent ones. No sooner has the triumvirate celebrated their renewed union, than we are taken to Sicily where Ventidius, having just avenged the death of Marcus Crassus, reminds us of the fragility of the first triumvirate (3.1). In Athens, Octavia becomes aware of the impossibility of stable relations: "A more unhappy lady, / If this division chance, ne'er stood between, / Praying for both parts...no midway / 'Twixt these extremes at all" (3.4.12–14, 18–19). Having chosen to engage Caesar on the indeterminate and insubstantial sea at Actium, Antony reverses course and leaves Enobarbus dumbfounded.

SCARUS: On our side, like the tokened pestilence
    Where death is sure. Yon ribaudred nag of Egypt—
    Whom leprosy o'ertake!—i'th' midst o'th' fight
    when vantage like a pair of twins appeared
    Both as the same—or, rather, ours the elder—
    The breeze upon her, like a cow in June,
    Hoists sails and flies.
ENOBARBUS: That I beheld.
    Mine eyes did sicken at the sight and could not
    Endure a further view.
SCARUS: She once being loofed,
    The noble ruin of her magic, Antony,
    Claps on his sea-wing and, like a doting mallard,
    Leaving the the flight in height, flies after her.

> I never saw an action of such shame.
> Experience, manhood, honour, ne'er before
> Did violate so itself.

> (3.10.9–24)

The scene in which Antony is most reduced and dissipated by defeat—"unqualitied with very shame" (3.11.44)—significantly occurs in an unspecified location.

To the oppositions of masculinity and femininity, nature and civilization, stoicism and license, we may therefore add the opposition of apparent reality to the transformative power of that tragic vision which perceives the contingency of reality.[7] This latter opposition serves to locate this apparently heterodox tragedy in the same general pattern of *Hamlet*, *King Lear*, and *Othello*. The world appears to the tragic protagonists as governed by a set of inviolate laws: daughters love their fathers, wives love their husbands, subjects obey their King, murders are always avenged.[8] The tragedy is set in motion by the knowledge that the world might be otherwise— an inexorable telos that exacts vision and knowledge at the price of freedom.

Each character necessarily inscribes a path that constitutes at least one, and possibly several, subgraphs of the overall topology of the drama. In order to capture this movement, we can ask the computer to generate the graph for the play, but to highlight every node in which a particular character appears at least once, and to draw a separate colored path denoting the passage of that character from one scene to the next. Here, for example, is Antony's path, which extends to nearly every location in the play:

---

[7]Susan Snyder has detected a more-or-less constant pattern of motion in the imagery of *Antony and Cleopatra* that harmonizes both with the notion of the real and the apparent and with the general movement from one state to the other: "Shakespeare has set images of solid fixity or speedy directness against images of flux and of motion unpurposive but beautiful to express kinetically the opposition of Rome and Egypt and, through their incompatibility, the nature of Antony's tragic dilemma" (114–115).

[8]Paul Yanchin maintains that through the characters of Antony, Cleopatra, Enobarbus, and Dolabella, *Antony and Cleopatra* "displayed 'absolutist loyalty' in such elaborate detail that absolutism's deeply conflicted dependence on 'sovereign subjectivity' was able to emerge into the consciousness of the members of its 1606–1607 audiences" (Yanchin 345).
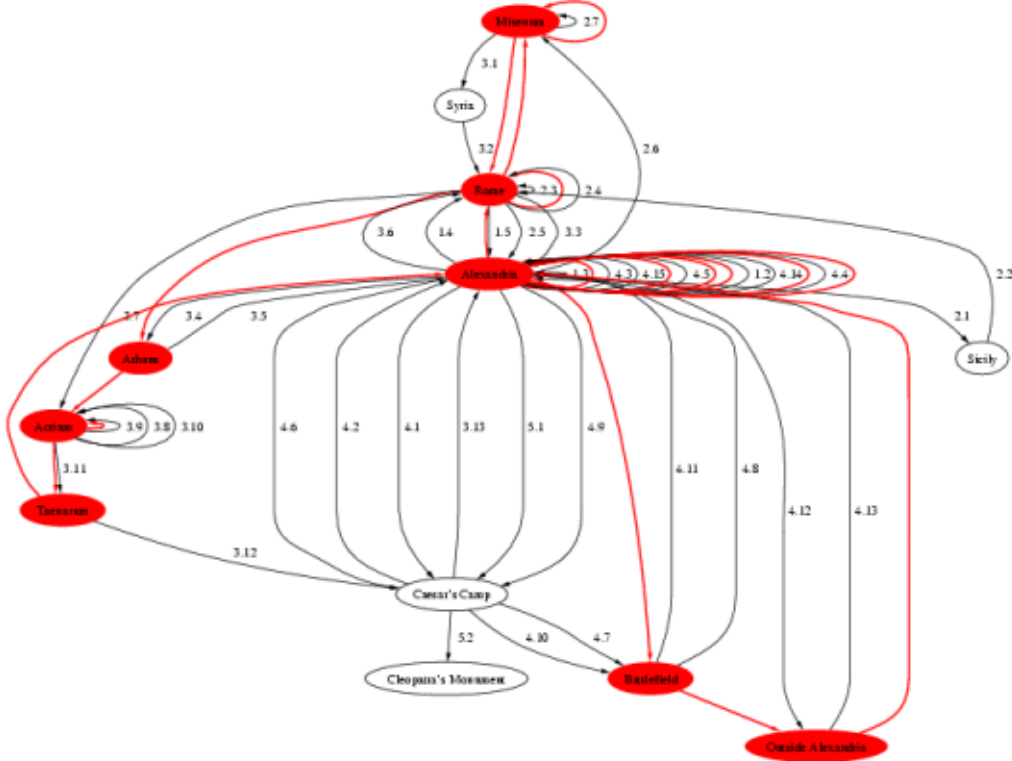
Cleopatra's path, by contrast, is confined to Egypt and Actium:
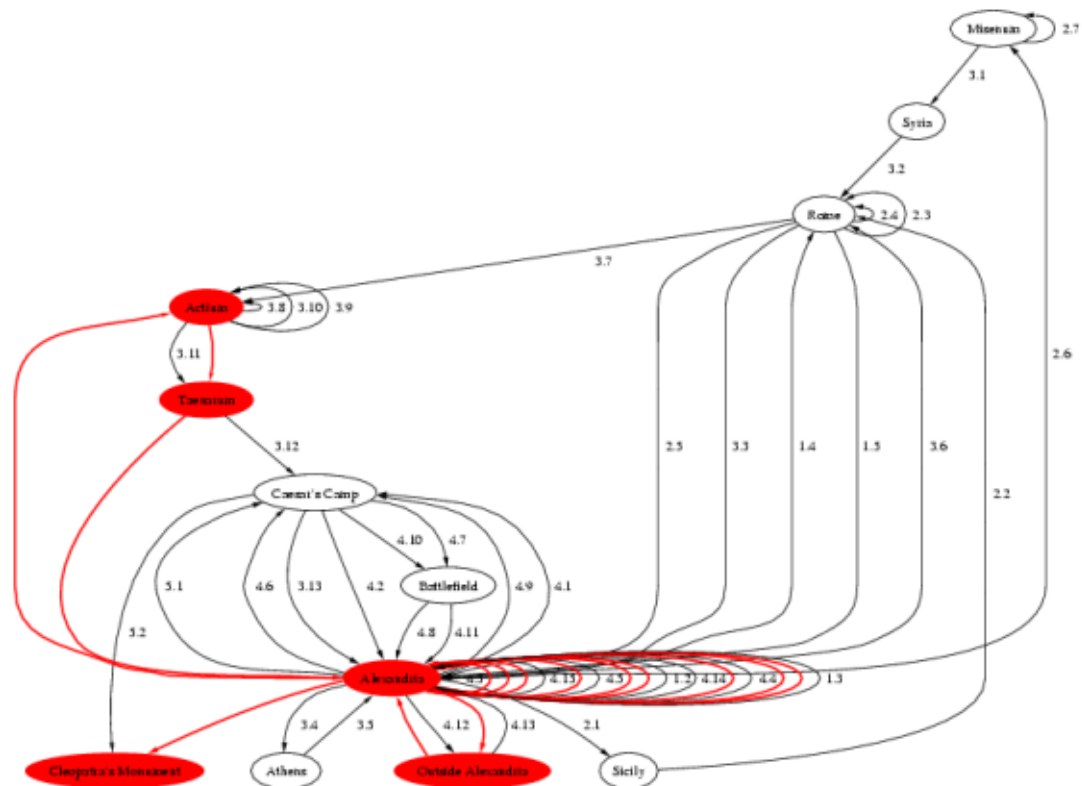
It is possible to superimpose the two paths by instructing the program to color Antony's path red, Cleopatra's yellow, and to color any node visited by both characters orange:
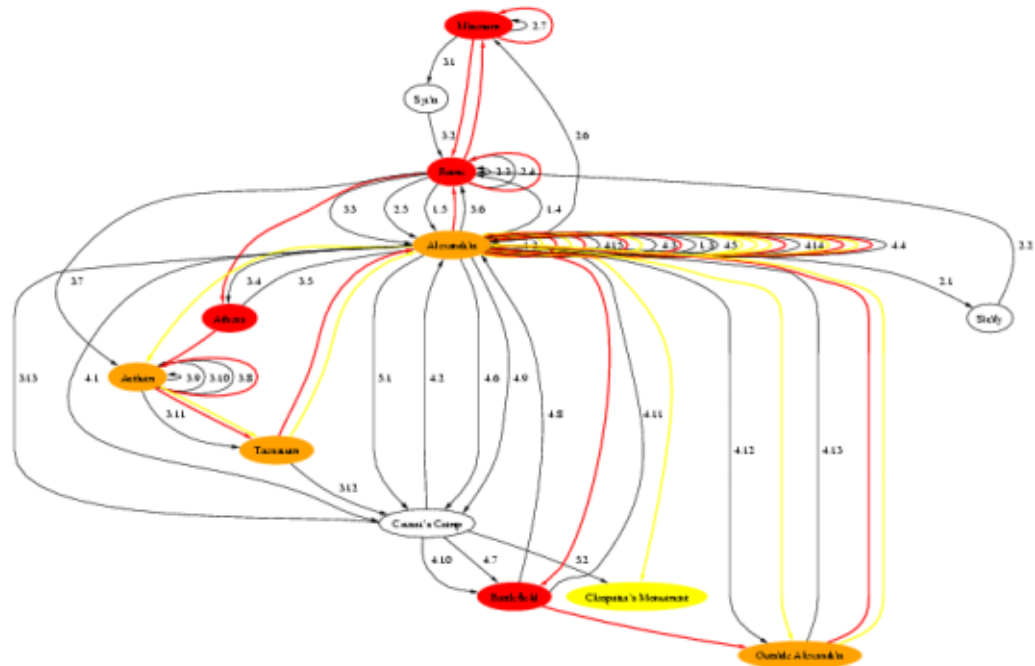
The confluence was technically visible in the previous graphs, but this one draws our attention to its singularity: Antony and Cleopatra share not only Alexandria, but the two liminal locations moving away from Rome; and besides these, another uncertain location "outside Alexandria." At first unaware of the status of the battle, Antony briefly exits the scene for a better view ("Where yond pine does stand / I shall discover all" (4.12.1–2). Like "Taenarum," this scene is a site of extraordinary confusion and rupture. Here Antony witnesses his own men going over to Caesar's army and subsequently accuses Cleopatra (who appears, but offers no response) of betrayal.

Where do we go with this? How about here: *Antony and Cleopatra* reveals itself as play in which the tragic pattern of apparent reality yielding to the unforseen forces of actuality is enacted not among characters moving within the local circumstances of a minor court, but in a series of events occurring on the grand stage of human history. The stable knowledge that precipitates the tragic awareness of fragility and contingency is located not merely in the psyche of the main characters (who are neither driven to madness nor paralyzed by inaction—the typical pattern in the major tragedies), but in the larger sweep of cultures and kingdoms. *Antony and Cleopatra*, which is both a history play and a tragedy, is above all a play about the tragic undercurrent of history itself.

It would be an absurdity to maintain that the graph which led us to the insight is itself an interpretation of the play or that it somehow proves that the reading is correct. The maneuvers which the program provokes simply reframe the play's intelligibility by creating alternative textual and graphical arrangements. Yet such rearrangements are nonetheless essential to the project of interpretation.

An objection presents itself. The processes of pattern formation which the graphing program performs could surely have been undertaken with pencil and paper, the computer merely adding speed (and perhaps a bit of showiness) to the process. The first half of this objection is accurate as far as it goes, though it might be noted that virtually any computational process can, at least in principle, be undertaken with pencil and paper. Adding twenty-digit numbers, preparing manuscripts, and drawing diagrams were all done with the simplest of writing instruments for centuries before the advent of computers. Even extremely difficult applications, such as

predicting the weather from simulative models, are not technically beyond the ambit of paper and brush. There is no inherent connection between the electro-mechanical device we call a computer and the formal abstraction we call an algorithm. Donald Knuth, author of the monumental *Art of Computer Programming*, considers it basic to the definition of an algorithm that "its operations must all be sufficiently basic that they can in principle be done exactly and in a finite amount of time by someone using pencil and paper" (Knuth 6). Even the most enthusiastic partisans of computer technology will have to concede that in the end, it all comes down to speed and automation.

It would be a grave mistake, however, to minimize the significance of speed and automation either as cultural commodities or as philosophical categories. Katherine Hayles, in a discussion of hypertext as a "self-organizing" system of signification, notes that speed of access is the defining feature of such systems:

> Confronted with the theory and practice of hypertextuality, many people insist that it is nothing new. After all, they say, *Paradise Lost* was published in print books with appendices and footnotes long before hypertext appeared on the scene. Nothing has changed with the hypertext version, these people argue, except speed of access. But for human memory, speed of access is crucial. It often makes the difference in whether self-organizing processes spontaneously emerge or not. A recondite reader may of course do for herself what the naive reader does when he repeats the sentence— mentally rehearse the footnote on page 497 while looking at the text on page 216 so that both are held in short-term memory together. This takes effort, however, and most readers will make it only occasionally if at all. By facilitating these juxtapositions, and especially by shortening the time it takes to make them, hypertext encourages self-organization. (Hayles 215)

Similar claims for the transformative power of speed could be made in numerous contexts. Satellite communications, email, television, and the World Wide Web have not merely added speed to the capable messenger on horseback. Our sense of the world has been radically altered by the automation of the slower systems.

Automation, which is often both the cause and the effect of speed, moves the task which a particular algorithm performs out of the realm of the problem space. Elementary school children are today capable of algorithmic calculations (such as long division) which left entire ancient civilizations baffled. They do this by employing "automatic" algorithms, which, because they are fast and performed with a minimum of effort, allow the student to focus on some more interesting problem. Because the graphing program produces graphs almost instantaneously and may be trained on any play without effort, we are able to use it to study aspects of Shakespearean drama which might simply have been too laborious to construct by hand. Whitehead's statement, though a bit grand, captured the basic sentiment at a time when computers were still in their infancy: "Civilization advances by extending the number of important operations which we can perform without thinking about them" (Whitehead 10).

Of course, someone (namely me) had to think quite a bit in order to get a computer to represent a Shakespeare play as a graph. The solution isn't quite as elegant, nor as useful as long division, which required a truly sublime act of human ingenuity to devise. But if humanities computing were to focus on facilitating the process of pattern formation—which is another way of referring to the process of critical engagement—it might one day produce something as transparently beautiful. Such a tool might be to humanistic study what the telescope or the particle accelerator is to science—not simply a study aid, but a means by which the field of inquiry itself is expanded and new interpretive valences unleashed. Who knows what we could do then?

WORKS CITED

Biggs, Norman L., E. Keith Lloyd, and Robin J. Wilson. *Graph Theory: 1736–1936*. Clarendon, 1977.

Devlin, Keith. *Mathematics: The Science of Patterns*. Holt, 1994.

Gillies, John. *Shakespeare and the Geography of Difference*. Cambridge UP, 1994.

Graville-Barker, Harley. *Prefaces to Shakespeare*. Sidgwick, 1927.

Hayles, Katherine N. *Cyberspace Textuality: Computer Technology and Literary Theory*. Edited by Marie-Laure Ryan, Indiana UP, 1999.

Hazlitt, William. *Selected Writings*. Oxford UP, 2009.

Hockey, Susan. *Electronic Texts in the Humanities*. Oxford UP, 2000.

Knuth, Donald E. *Fundamental Algorithms*. Vol. 1, *The Art of Computer Programming*, Addison, 1997.

Lubar, Steven. "'Do Not Fold, Spindle, or Mutilate:' A Cultural History of the Punch Card." *Journal of American Culture*, vol. 15.4, 1992, pp. 43–55.

Milic, Louis. *The Computer and Literary Style*. Edited by Jacob Leed, vol. 2, Kent Studies in English, Kent State, pp. 79–106.

Moretti, Franco. *Atlas of the European Novel: 1800–1900*. Verso, 1998.

Rosen, Kenneth H. *Discrete Mathematics and Its Applications*. 7th, McGraw, 2007.

Snyder, Susan. "Patterns of Motion in *Antony and Cleopatra*." *Shakespeare Survey*, vol. 33, 1981, pp. 113–22.

Trudeau, Richard J. *Introduction to Graph Theory*. Dover, 1994.

Whitehead, Alfred North. *An Introduction to Mathematics*. Oxford UP, 1958.

Yanchin, Paul. "Shakespeare's Politics of Loyalty: Sovereignty and Subjectivity in *Antony and Cleopatra*." *Studies in English Literature*, vol. 33, 1993, pp. 343–63.