Digitising is questioning, or else.

Willard McCarty
staff.cch.kcl.ac.uk/~wmccarty/

Long Room Hub Lecture Series
Trinity College Dublin
16 April 2008

> Vere, inquam, magistrum mirabilem mihi dedisti, qui a me factus me doceat, meque docens ipse nihil sapiat. Imo propter patientiam et obedientiam sui eum maxime amplector; cantabit enim mihi quando voluero, et nunquam de mei sensus tarditate commotus verberibus vel injuriis cruciabit.
>
> *Indeed, I say, you gave me a wonderful master, who made by me teaches me, and teaching me himself knows nothing. Rather I embrace him in submission and in obedience; truly he will play for me when I wish, and never will he torment, with floggings or blows, the thoughts that move slowly from me.*
>
> Odo Cluniacensis, *Opuscula de musica* (Migne 133, 762Dsq)

1. The problem

[SLIDE] Sixty-five years ago, on successive Friday evenings in the Physical Laboratory Lecture Theatre, Trinity College Dublin, physicist Erwin Schrödinger gave a three-part lecture on the curious failure of the physical sciences to answer the question, "What is life?", and the promise of a "new physics" to address that failure. Schrödinger was an intellectually brave man, who saw no other escape from the dilemma in which ignorance of the whole increasingly places the specialist than that "some of us should venture to embark on a synthesis of facts and theories, albeit with second-hand and incomplete knowledge of some of them – and at the risk of making fools of ourselves" (1943: 1). Now, a lifetime later, with the infirmity of literary computing rather than the impotence of physics in mind, I cannot think of a better way of putting before you the challenge we students of literature face from computing, nor can I think of a better response. To put the situation in a nutshell, by placing the question of method in the path of scholarship, computing offers us the chance to be similarly brave. I will return near the end of this lecture to a much greater intimacy in the parallel between us and Schrödinger. But for now let me just state the problem as I see it, which turns out to be not entirely dissimilar to his: why computing has been unable to do much for the interpretative operations that are central to the humanities and of all-consuming interest to scholars of literature and the arts.

2. Lessons of history, 1957-1989

[SLIDE] The relationship between computing and literary studies has had a problematic history. In 1966, when the first professional journal in the field, *Computers and the Humanities*, began, groundbreaking activity, which had begun some 20 years earlier, became suddenly visible across the disciplines. Early, dramatic successes in computing had stirred popular and scholarly interest in what computers were doing and might be able to do. [SLIDE] In 1957 Herbert Simon had made startling predictions for the subsequent decade: "Put it bluntly (hard now to shock)", he wrote in his lecture notes, "Machines think! learn! create!"[1] [SLIDE] In a similar vein, British computational linguist Margaret Masterman asserted in her contribution to *Freeing the Mind*, a series published in the *Times Literary Supplement* in 1962, that the computer was far more than the "menial tool" other contributors had described. It was, she declared, a "telescope of the mind" which like its astronomical namesake would soon "enlarge the whole range of what its possessors could see and do" and so change "their whole picture of the world" (38f) – words quoted approvingly four years later in a book on the automated analysis of text (Ellis and Favat 1970/1966: 126). [SLIDE] At the end of that decade, the editor of *The Economist* predicted that in the 1970s the computer would come into its own, perhaps even changing "the outlook of man". [SLIDE] A somewhat less restrained Canadian journalist had proclaimed five years earlier a "new age of miracles" [SLIDE] in terms strikingly similar to a 21st-century funding concept known as "the semantic web".

Battle-lines were being drawn over the significance of computing. Note, however, not the expected one, with enthusiastic promoters on one side and grumpy Luddites on the other, but the one between those, like Masterman, who focused on augmentation of human capacities, and those like Simon, whose interests in automation lead straight back [SLIDE] to Frederick Winslow Taylor's clerkish elevation of efficiency, the modern factory production-line, a leisure-class of passive consumers and the rest of modernism's children.

[SLIDE] Two technical assessments from the mid 1960s told a sobering story, however. In *Alchemy and Artificial Intelligence*, with the clock still ticking on Simon's confident predictions, Hubert Dreyfus diagnosed stagnation in AI research, pointing to unexpected difficulties which had followed the early successes. What had been assumed to be a difference of degree, between

[1] Simon 1957, later incorporated into Simon and Newell 1958a; the predictions, having encountered objections, were repeated with emphasis in Simon and Newell 1958b.

current and envisioned systems, had turned out to be a difference in kind with no known way across the discontinuities (1965: iii, 9). [SLIDE] The following year the devastating "black book" on machine translation, *Language and Machines: Computers in Translation and Linguistics*, came to more or less the same conclusion (ALPAC 1966: 32). As Yorick Wilks noted a few years later, "it had been clear for some time that the era of simple-minded MT was over" (1972: 4) and that a new, very different research paradigm was needed.

In the 1960s, among the beleaguered minority of literary scholars who dared to compute, the prevalent view of computing was, as it has remained, clerkish, Taylorian: "an invaluable assistant" to scholarship, as Susan Hockey has said, most usefully deployed to probe for textual surface-features and so to prompt the enquirer to reflect "on the methodology used to interpret the results" (2000: 84). It made scholarship more efficient, progress a matter of traversing what in 1965 poet and English professor Ephim Fogel called the "Vision-Actuality Interval", which however wearisome was, he thought, only an incremental matter of time, fortitude and steadily accumulating resources. Enough of these were in fact already to hand (Rommel 2004: 93). As with AI and machine translation, progress faltered because the problem was theoretical, not practical. A different, non-incremental idea of computing was required.

[SLIDE] A year later, in the first article of the inaugural issue of *Computers and the Humanities*, Louis Milic turned attention back to the ideals of scholarship. He argued that literary scholars, who are "involved primarily… with the mystery of the creative act", needed to reorganize their thinking for the new age by turning from the computer as plodding "mechanical clerk" to the possibilities of a creative instrument (1996: 6). He was not the only one saying such things.[2] [SLIDE] Perhaps the clearest statement came in 1978, when Susan Wittig, referring to Masterman's vision, pointed out that computing had undoubtedly allowed for improvements, making performance of old tasks more efficient and accurate, but it had not delivered on the vision's promise (211). She argued that inattention to theory had made literary computing vulnerable to covert influence by a positivistic "concept of text" derived ultimately from New Criticism (1978: 211). A computing without theory, Colin Martindale argued that same year, was no better than a method in search of a paradigm to direct and explain it (1978: 275f). Richard Bailey, quoting both Wittig and Martindale, declared that practitioners were blindly groping their way through criticism's past, with a time-lag of about 50 years (1978: 7).

[2] This part of the story is well told by Potter 1991: 403-7.

[SLIDE] Thus as early as the beginning of humanities computing's first professional journal, the whistle was blown on naïve literary computing. But, as with AI, imaginative scholars had opened the door to a problem that challenged the theoretical ground of its practice. The fact that this hugely difficult challenge was largely ignored in the humanities, given the relative ease of devising technical solutions to simple problems, can hardly be surprising, but ignoring it has meant that however excellent in kind, however supportive of interpretation, literary computing has tended silently to perpetuate the concept of text to which the whistleblowers objected. Despite the intention, declared two years ago at the Summit on Digital Tools in the Humanities, to "enable new and innovative approaches to humanistic scholarship" (Frischer et al. 2006), discussions of tool-building have likewise been preoccupied with features of software, continuing to allow old ideas their shadowy power, ignoring the discourse of criticism as it has moved on. Despite the ambitious aims of the Chicago Digital Humanities/Computer Science Colloquium, now in its third year, we find no sign that this challenge is recognized, although it is fundamental to the important question the Colloquium raises.[3]

No surprise, then, that a decade after Wittig, in the year that the text-analysis program TACT was released to the public, Rosanne Potter remarked that literary computing had "not been rejected, but rather neglected" by mainstream criticism – a complaint repeated many times since.[4] Numerous reasons for this neglect have been offered, but the fact remains that literary computing has had very little to say in response to critical discourse for the last half-century. In the preface to *Radiant Textuality* (2001), Jerome McGann noted its instrumental role in "the technical and pre-critical occupations" on which scholarship depends but its almost total absence from interpretative procedures (2001: xii).

It would be a mistake, however, to say that the fault lies solely with literary computing for its neglect of theory. For the past several decades most theorizing, shading into what Jonathan Culler has called "just plain 'theory'" (1997: 1), has offered few points of contact with the study of actual texts, and so few possibilities for coaxing computing practitioners out of their theoretical silence. In addition, the social division in universities that has separated the non-technically educated theorists and critics from the less institutionally

[3] For the Colloquium, see lucian.uchicago.edu/blogs/dhcs2008/; note esp. Dan Cohen's blog for Monday, 13 November 2006, at www.dancohen.org/category/computer-science/ (12 April 2008).

[4] Potter 1989: xvi; see also Potter 1991; Corns 1991; *Computers and the Humanities* 27.5-6, 1993; Opas and Rommel 1995; *Literary and Linguistic Computing* 18.2, 2003; Rommel 2004; Hoover 2007; McCarty 2008 (forthcoming).

privileged technical experts has impeded and in some countries continues greatly to impede progress.

Something has to be done, but what? In the article I have already cited, Martindale pointed out that both literary theory and literary computing "have strengths and weaknesses, but the striking thing is that the strengths of one are the weaknesses of the other. If the two were meshed," he suggested, "the weaknesses would largely be cancelled out." But rather than mesh, empirical and theoretical approaches have been taken up each in turn, each taken as the answer rather than as the answer's other half. Leonard Forster noted in his Presidential Address to the Modern Humanities Association in 1978 that as a result we get dogmatic abstractions, the criticism formed around them becoming what he called "a flight from literature". He recommended a "flexible pragmatism" analogous to the craftsman's, who selects now this tool, now that one to accomplish whatever task is at hand. Forster's metaphor is a good one, not because it privileges the scholarly task over the tool (the timid mantra of service-orientated computing) but because it places tools in the context of an *active interface* between craftsman and material.

The moral of the story is that neither task nor tool holds the secret. What's needed is attention to the craftsmanship, to the process and practice of that which criticism entails.

3. The matchmaker's tasks

[SLIDE] Literary computing can doubtless continue as the "invaluable assistant" to scholarship, following criticism wherever it goes and trying its best to be of service. But scattered results from literary computing, obviously better theory and the benefits of hooking up with the discourse of criticism suggest Martindale's Leibnizian marriage of theorist and empiric is not only possible but also holds great promise. The fact that literary computing remains vigorous, however ghettoized by specialist concerns into specialist periodicals (Corns 1991), suggests an underdeveloped rather than moribund research programme. The question is, what now must be done to realise the possibilities?

In *Humanities Computing* (2005) I took up part of the task by concentrating on the theoretical implications of computing as an analytical approach to the study of the humanities as a whole. I presented a negative epistemology, arguing that the primary function of computing is not to automate knowing but to identify what we somehow know but cannot adequately specify. Because computing gives us manipulatory power over the models we construct, we are able rapidly to close on that which cannot be formulated.

Thus we are confronted with our own quite specific ignorance of cultural artefacts and so are better equipped to question them. For literary studies this epistemology takes computing significantly beyond the standard view of an efficient but essentially mute and obedient handmaiden by challenging us in detail to account for the failure of any rigorously analytical try for a systematic order of things. But it takes us no further than the negative gift with positive consequences that lie somehow beyond what any such try can in principle do.

The situation in which we find ourselves suggests an analogy to the observational sciences. As Ian Hacking has argued for microscopy,[5] the fundamental problem raised by all observational instruments (including the telescope to which Masterman appeals) is that we "don't just peer" through them to newly visible objects that are as we see them to be, independently of the viewing. We must also "interfere" with the incoming data based on what we know of what we are trying to observe. We must *make sense with* these data, sometimes by intervening in the observational process, sometimes by altering the object of study. This we simply cannot do, or do well, without a good idea of what we are looking at. In literary studies such knowing interference is not, as in the sciences, so much a preliminary step toward consensus about the object in view as it is an ongoing, never-ending process. The literary object in view is hardly an object at all but the contingent, interactive, emergent outcome we wisely use a gerund to name: *reading*.

For centuries, of course, the codex book has functioned as such an observational instrument – I. A. Richards named it a "machine to think with" in 1926[6] – encouraging interpretative interference with the flow of language, even (e.g. in critical editions and commentaries) providing optional sequences of interfering moves. This is the book not only as metatheoretical statement but as analogue to firmware.[7] Computing foregrounds book-as-machine, especially in the design and construction of digital reference works. But "computing" is also significantly a gerund, not a name for an action or set of actions but a name for *acting*. It is, I argued in *Humanities Computing*, fundamentally a modelling machine. Hence its introduction into literary studies implicitly shifts emphasis from representation to intervening, and so implies that theorizing of text at the fundamental level of tool design and use is essential.

---

[5] Hacking 1983: 186-209. For the computer as microscope, see e.g. ALPAC 1966: 121, Gilmour-Bryson 1984: 11; for microscope and telescope, Denning 2007; more generally, Mahoney 2000: 31. Frege, echoing Leibniz, used the microscope as a metaphor for his *Begriffschrift* (notation of concepts), for which see Göranzon 1993: 44 and Crane 2003: 24.

[6] Richards 1926: 1; cf. McGann 2001: 54-7.

[7] On the critical edition as metatheoretical statement, see McGann 2001: 75-97.

If this is so, then much more than the epistemological question is at stake. To be brought to ask how we know what we somehow know but cannot represent computationally is a major step forward, [SLIDE] but it is preliminary to asking the ontological question Wittig raised in 1978 and McGann again in 2004: *what is text that it eludes all such representation* – that it can be, in McGann's words, "the hem of a quantum garment" (2004: 201)? Analytical literary computing tells us how to exploit the unavoidable difference between textual representation and reality, but it has nothing at all to say about what we choose to represent. Even if we agree (as we certainly should not) to limit the textual object of study to its verbal data, trouble starts with the context required for interpretation. The dominant consensus within a critical specialism may obscure the problem and often does. But we are warned of it by the crippling difficulties of infinite regress that the very idea of context appears to cause whenever anyone asks what exactly it is a promissory note for (Scharfstein 1989). Context, Jonathan Culler remarked, is merely more text and so appeal to it solves nothing (1988: 93f). But appealing to it, particularly if it is to be modelled computationally across the open domain of literature (or of real life) reveals how unsatisfactorily arbitrary and limiting the unspoken notion or any analytic formulation of it is. The problem of context is the problem of text. What is it?

By failing to ask the Wittig-McGann question, literary computing is confined to providing evidence for or against what we already know or suspect. It is strongly inhibited in its capacity to surprise. Providing evidence seems justification enough, but evidence becomes increasingly problematic as the volume of data exceeds the norm for critical practices formed prior to the exponential growth of online resources. As this volume increases, so does the probability of arbitrary choice, and so the ease with which any statement may be connected to any other. Good critics may do better scholarship by finding more of what they need; bad critics may be swiftly becoming worse ones more easily. The point, however, is that literary computing has thereby served only as mutely obedient handmaiden, and so done nothing much to rescue itself from its position of weakness, from which it can hardly deliver the benefits claimed for it by the faithful. It has done little to educate scholars methodologically.

There is, of course, no single answer to the Wittig-McGann question, because there are many kinds of text, many ideas of what to do with each kind and every reason to think that these kinds and ideas are limited only by human ingenuity. Given the renewed prominence that McGann's work has brought to the question, what can be done is to develop ways of asking it such that responses can be made in software.

An obvious starting point is with inherited tools of reference, e.g. lexicons, critical editions and commentaries, inferring from them the ideas of text they implement. To the degree this has been done, in aid of speculating about or designing a software equivalent, results suggest the prominent role of tacit uses in the social contexts of argument and in the building or maintaining of a social imaginary (McCarty 2004). I will return to the importance of this later. Results also illumine the primitive crudity of our software tools. Let us take a brief look at them to see what they can teach us.

4. What the tools say

[SLIDE] We know from unsatisfactory experience that none of these tools do very well with the Wittig-McGann question, but to do anything useful at all, they must afford a view of it that can be recovered.

Initially the answers we get back from existing tools are impoverished. [SLIDE] A concordancer, for example, implies that by "text" we mean a corpus informed by verbal correspondence of passages and by the words that collocate with whatever word is in focus. [SLIDE] Both relational database design and formal ontologies imply an instantiated set of concepts and their interrelations, and prior to these, well-defined perspectives of enquiry. [SLIDE] An annotation tool affords a view of text as the occasion for commentary. [SLIDE] A statistical analyzer yields a complex population of verbal clues to a literary style. These are all valid, even highly valuable aspects of text, but again, they are isolated and so isolating.

We can, however, greatly enrich what each has to contribute by considering their historical origins. [SLIDE] The most obvious to be explored is concordancing, a direct descendant of the late 12th or early 13th-century device invented to serve figural interpretation of the Bible, which once it achieved formal stability in the late 13th Century remained broadly the same until computing (McCarty 1993). The keyword-in-context format, devised in the 1950s to satisfy the needs and capabilities of automation, shifted focus from concordant passages of a text to shared collocates of a word, and so moved the principal domain of use from literary studies to corpus linguistics. Nevertheless the mechanized idea of semantic triangulation basic to the figural scheme remains implicit in the tool which that scheme articulates and so in the results the tool produces. It bears with it or more accurately implies a theory and compositional principle derived from the most influential text in the European tradition. [SLIDE] So also the tools and techniques of annotation imply a partial answer to the Wittig-McGann question. These have historical roots in ancient commentary practice, including manuscript glosses,

marginalia, free-standing notes and other forms of intertextuality, together with their social networks (McCarty 2004). [SLIDE] Relational database design and textual ontologies are similarly emergent from older practices of categorization and tabular layout beginning with ancient libraries, and more recently from the strong cultural predisposition toward discontinuous plurality. Lev Manovich's argument for the database as a symbolic form provides a starting point here (2001).

5. Turning to confront the context

[SLIDE] Writing a conceptual history of literary computing from its tools helps to give it a theoretical voice, but at best the exercise yields a semi-coherent miscellany with uncertain relationship to actual research. The result is worse than might be expected because the ideas of text we seek are partially in the tools, partially in unexpected uses of them (especially true of those tools not designed for the purpose) and, in the usual situation where more than one tool is used, partially in which tools are applied in what sequence. Lacking in a tradition of experimental work, under-educated technologically and so undervaluing or simply not seeing the mediation tools perform, researchers have tended to omit the kind of observations we need. Unsurprisingly, evidence from the scholarly record, in the rare instances in which it exists at all, is scattered through footnotes and asides in publications across many disciplines.

In any case, there is no whole for these parts to sum to, no great idea of text that may be assembled from the scattered fragments of its implementation. The point of asking the Wittig-McGann question is quite otherwise: to enable literary computing to make a great inductive leap from its mute servant's mimetic doldrums to an understanding of itself as a full participant in the interpretative operations of criticism.[8] To devise new tools without the benefit of that question has not and will not significantly increase the mildly helpful but severely cobbled abilities of literary computing no matter how much data accumulates.

McGann's own response to the question has been to argue for the reversal of perspective within criticism already implied by the Bakhtinian situating of text in an "immense, boundless world of others' words" (Bakhtin 1986: 143). The details of this response, including the online game *IVANHOE*, are best presented by his own writings, which are here taken as required reading and as a point of departure, to which I will return.[9] But that reversal of perspective

[8] For an early, primitive attempt see Smith 1989/1978.

[9] See McGann 2006, 2004 and 2001, also www.iath.virginia.edu/~jjm2f/online.html (6 April 2008).

is already inescapable given the problem of context, which itself seems inevitable once we free literary computing from the strictures of a knowledge jukebox to become a project for modelling literature. The fundamental role of modelling is itself an inevitable consequence of Mr Turing's universal machine (McCarty 2005: 170-2).

The term "modelling" is so polysemous that its meaning cannot be taken for granted, so I had better say what I mean by it. [SLIDE] Here I show the modelling relation between a formal or abstract system, such as a computer program, and a natural system or artefact, such as a poem. In *Humanities Computing* I argued for the analytical, mimetic kind that Clifford Geertz has called "modelling-of" (1993/1973: 93), which aims at refinement of the epistemological question, as noted earlier. Geertz distinguished this kind from its opposite, "modelling-for", a more or less creative realisation of an idea or design achieved through perfective, exploratory manipulation. (Design for a new airplane wing is a straightforward example.) [SLIDE] The Bakhtinian reversal, however, entails a different sort of modelling from either of those two, something that resembles modelling-for but begins without a pre-existing design, or at least not a consciously accessible one. It is a mapless modelling "forward", toward something that is not yet anything. Using the musicological term, I call it "improvisational modelling" to denote its moment-by-moment development in performance of an emergent potential. This sort of modelling is widely attested in the experimental sciences.[10] What it might be for text reflects, again, the Wittig-McGann question.

6. Bridging discourse

[SLIDE] It is a truism that asking questions is central to the humanities, and that good research leads from a worthy question to a better one. The Wittig-McGann question is certainly worthy, but it leaves us with the problem of how to reformulate it so that it may be asked in software.

Before we can even get properly started, that is, we must confront the gulf separating the language of criticism from the language of implementation. Happily this gulf is bridgeable. In fact collaborative projects in the digital humanities have for years negotiated it as a matter of course by developing common ways of talking about problems and objects that have different meanings for the various participants.[11] But although collaboration offers the

[10] Gooding, for example, focuses on the products of such modelling rather than the process; he names them *construals*, "flexible, quasi-linguistic messengers between the perceptual and the conceptual" (1986: 208) or "tentative representations of possible outcomes… continually constructed and revised to describe and communicate actual outcomes" (1992: 103).
[11] Galison 1997 s.v.; McCarty 1995: 121-9.

great benefit of other-mindedness, alone it is an inefficient and only partially effective means of furthering research that is fundamentally the result of two or more intersecting, interacting practices. Collaboration wherever possible needs to be internalized so that the interacting can occur at the speed of thought as well as at the pace of meetings. Hence the need for a bridging discourse.

The time-honoured approach for building a new discourse is to reach into older, better established fields for promising figures of speech and thought, then to assimilate them (McCarty 2005: 114-57). In each case a connection is established from the poorly understood phenomenon or system with which one is working to a better understood analogue elsewhere. [SLIDE] The analogy links relationships, not things: as A is to B (within one system) so C is to D (in another). Its strongest claim is that the two systems, however different, are *isotropic*, i.e. the same governing laws or principles apply in both. Hence a strong analogy not only holds up to examination and yields many insights, it also pulls the connected fields closer together by emphasizing similar processes operating in both. Each analogical connection must be probed for its actual benefits as well as cognitive trajectory, but because its yield may not be known for some time, the best anyone may be able to hope for is plausibility at the outset. Analogizing is conjectural. Considerable effort is required to maintain an analogy *as* a conjectural move and not blur it into an identity, especially when it appears greatly to simplify an intractable problem such as the one under consideration here. In other words, analogizing is as perilous as it is powerful.

In the present case what we are looking for is, in the words of a London improvisational musician, how one gets "from A to C when there is no B" (Bailey 1992: 136). If, that is, we begin, as readers do, with a text, and so with the question of how reading may be modelled, we need to bridge the gulf between Bakhtinian language and a design strategy for a computing system capable of implementing its outward-looking, improvisational trajectory.[12] One promising place to begin is with evolutionary biology, whose fundamental problem is precisely to answer the improvisational question for living systems. Thus we return to Erwin Schrödinger's lecture at Trinity in 1943 for the more intimate parallel I promised.

In his commentary on Schrödinger's project, theoretical biologist Robert Rosen has argued that by asking his question, illegitimate within the confines of ordinary science, Schrödinger diagnosed the fatally constricting path of

[12] Much closer to an actual language of design is Eco 1984: 3-43, but Eco sketches states and transitions between them rather than processes.

reductionist methods that had had such great influence on 20th-century thought. "[O]ur universes [of scientific discourse] are limited," Rosen declared, "not by the demands of problems that need to be solved, but by extraneous standards of rigor. The result is a mind-set of reductionism, of looking only downward toward subsystems, and never upward and outward" (2000: 2). What he does not say, but needs here to be said, is that the influence of scientific discourse on all others has been so great that this "mind-set of reductionism" has been ours as well. So also, analogically, is the alternative Rosen presents: a turn toward the quasi-teleological but non-deterministic idea of self-organizing systems, hence the ideas of complexity, emergence and autopoiesis (of which McGann makes extensive use) coming primarily from the culturally ascendant biological sciences, including biological anthropology.

Biology and its nearest neighbours (which, after all, still lie at a formidable conceptual distance from criticism) are not the only fields concerned with how more sophisticated systems arise from less sophisticated ones, however. Other likely candidates include anthropological linguistics and conversation analysis; improvisational musicology, including but not limited to studies of jazz; and the cognitive sciences, where for example the psychology of reading meets its neurological correlates. These are all promising sources for analogies.

7. An improvisational companion to criticism?

[SLIDE] Perhaps now it is worthwhile returning to Margaret Masterman's "telescope of the mind" to ask what sort of computational instrument might live up to the promise of enlarging "the whole range" of what we might see and do as critics and so change our "whole picture" of literature.

[SLIDE] The most imaginatively powerful attempt to date is *IVANHOE*, an online play-space in which participants intervene, change, add to and comment on the discourse field of a given cultural artefact.[13] The critical objective of the players is to explore in blog-like exchanges the possible worlds or imaginative trajectories of this artefact from an authorial "inner standing-point". Computational tools aid the interpretative play by managing communications and by visualising the interactions of players so as to stimulate their imaginations. Scope of play is constrained to the focal artefact, which players are assumed to know. Googling for whatever is permitted, but the game's tools do not aid or direct the search. *IVANHOE* is thus more closely analogous to a microscope than a telescope, but it is of Masterman's kind

[13] Although it is not scholarly in its immediate objectives, the online interactive game *Façade* holds some promise. See www.interactivestory.net/ and Rauch 2006. I am indebted to Matt Kirschenbaum for pointing me to this game.

nevertheless because it is built explicitly and self-consciously for looking outward from the artefact toward its manifold possibilities.

In the rationale for *IVANHOE*, McGann borrows extensively from theoretical biology and elsewhere, as I have suggested we need to do, but the analogies are rhetorical rather than computational. My research question is this: can we do more? Can we use these analogies to design modelling machines capable of finding connections from a given literary text to others, or can we adapt whatever software may exist, for example to simulate evolutionary or improvisational development? In 1989 Northrop Frye mentioned in passing the possibility that modelling such as I have described might be used to converge on fundamental structures of literature through systematic investigation of its recurring conventional units (1991: 6). Is this a realizable goal?

It's clear from what I have said that although such a modelling machine must be able to search all text in digital form, mere searching is not only insufficient but perilous without some kind of automated guidance. It's clear from the massiveness of the collection to be searched that only the most rudimentary scholarly metadata, if any at all, can be expected (though metadata generated by search-engines could perhaps be exploited to advantage).[14] It's clear that whatever the instrument does, it must be far more of a cognitively intimate companion than a bot, however semantic the web that gets searched. Searching will need to start from a reading, somehow specified, of a given text, produce results from the textual collection and learn from the reader's response, modifying both future and existing results according to what it learns. Hence, because the envisioned operations are massively combinatorial, they may well require more computing power than is easily available, at least now. They may be supercomputerish. Finally, it's evident that tools of some kind, perhaps like *IVANHOE* offering visual representations, will be needed so that the investigator can direct the machine more effectively and imagine more generously than otherwise.

The question of how to build such a thing is in essence the question of where the permeable, moving membrane is between reader and device, or to put the matter differently, how great a role computing can play in criticism. This is, in effect, the question of artificial intelligence, and so presumably a matter of keen interest for AI research. It is the most intimate, most promising encounter possible between literary theorist and literary empiric. But can it be done?

[14] I'm indebted to Dr John Keating (An Foras Feasa, Maynooth) for the suggestion concerning search-engines.

Works cited

ALPAC (Automatic Language Processing Advisory Committee). 1966. *Language and Machines: Computers in Translation and Linguistics*. Report 1416. Washington DC: National Academy of Sciences, National Research Council. www.nap.edu/catalog.php?record_id=9547 (7 April 2008).

Bailey, Derek. 1993/1992. *Improvisation: Its Nature and Practice in Music*. Cambridge MA: Da Capo Press.

Bailey, Richard W. 1989/1978. "The Future of Computational Stylistics". Paper for the Fifth International Meeting, Association for Literary and Linguistic Computing, King's College London, 15 December 1978. Rpt. in Potter 1989: 3-12.

Bakhtin, M. M. 1986. "From Notes Made in 1970-71". In *Speech Genres and other Late Essays*. Trans. Vern W. McGee. Ed. Caryl Emerson and Michael Holquist. 132-58. Austin TX: University of Texas Press.

Corns, Thomas N. 1991. "Computers in the Humanities: Methods and Applications in the Study of English Literature". *Literary and Linguistic Computing* 6.2: 127-30.

Crane, Tim. 2003. *The Mechanical Mind: A philosophical introduction to minds, machines and mental representation*. 2nd edn. London: Routledge.

Culler, Jonathan. 1988. *Framing the Sign: Criticism and its Institutions*. Oxford: Oxford University Press.

---. 1997. *Literary Theory: A Very Short Introduction*. Oxford: Oxford University Press.

Dreyfus, Herbert L. 1965. *Alchemy and Artificial Intelligence*. Paper P-3244. Santa Monica CA: Rand Corporation. www.rand.org/pubs/papers/P3244/ (23 March 2008).

Denning, Peter. 2007. "A New Interview with Peter Denning on the Great Principles of Computing". *Ubiquity* 8.22. www.acm.org/ubiquity/interviews/v8i22_denning.html (7 April 2008).

Eco, Umberto. 1984/1979. *The Role of the Reader: Explorations in the Semiotics of Texts*. Bloomington: University of Indiana Press.

Ellis, Allan B. and F. André Favat. 1970/1966. "From Computer to Criticism: An Application of Automatic Content Analysis to the Study of Literature". In *Science and Literature: New Lenses for Criticism*. Ed. Edward M. Jennings. 125-37. Garden City NY: Anchor Books.

Fogel, Ephim G. 1965. "The Humanist and the Computer: Vision and Actuality". *The Journal of Higher Education* 36.2: 61-8.

Forster, Leonard. 1978. "Literary Studies as a Flight from Literature?" *Modern Language Review* 73: xxi-xxxiv.

Frischer, Bernie *et al*. 2006. *Summit on Digital Tools for the Humanities: Report on Summit Accomplishments*. Charlottesville VA: Institute for Advanced Technology in the Humanities. www.iath.virginia.edu/dtsummit/ (7 April 2008).

Frye, Northrop. 1991. "Literary and Mechanical Models". In *Research in Humanities Computing 1. Selected papers from the 1989 ACH-ALLC Conference*. Ed. Ian Lancashire. 3-12. Oxford: Clarendon Press.

Galison, Peter. 1997. *Image and Logic: A Material Culture of Microphysics*. Chicago: University of Chicago Press.

Geertz, Clifford. 1993/1973. "Religion as a Cultural System". In *The Interpretation of Cultures: Selected Essays*. New York: Basic Books. Reprint: London: Fontana Press.

Gilmour-Bryson, Anne, ed. 1984. *Computer Applications to Medieval Studies*. Studies in Medieval Culture XVII. Kalamazoo MI: Medieval Institute Publications.

Gooding, David. 1986. "How do scientists reach agreement about novel observations?" *Studies in the History and Philosophy of Science* 17.2: 205-30.

---. 1992. "Putting Agency Back into Experiment". In *Science as Practice and Culture*. Ed. Andrew Pickering. Chicago: University of Chicago Press.

Göranzon, Bo. 1993. *The Practical Intellect: Computers and Skills*. Artificial Intelligence and Society. London: Springer Verlag.

Hacking, Ian. 1983. *Representing and Intervening: Introductory Topics in the Philosophy of Natural Science*. Cambridge: Cambridge University Press.

Hockey, Susan. 2000. *Electronic Texts in the Humanities. Principles and Practice*. Oxford: Oxford University Press.

Hoover, David. 2007. "The End of the Irrelevant Text: Electronic Texts, Linguistics, and Literary Theory". *Digital Humanities Quarterly* 1.2. www.digitalhumanities.org/dhq/vol/001/2/ (9 March 2007).

Mahoney, Michael S. 2000. "The Structures of Computation". *The First Computers – History and Architectures*. Ed. Raúl Rojas and Ulf Hashagen. 17-32. Cambridge MA: MIT Press.

Manovich, Lev. 2001. *The Language of New Media*. Cambridge MA: MIT Press.

Martindale, Colin. 1978. "Sit with the Statisticians and Commit a Social Science: Interdisciplinarity Aspects of Poetics". *Poetics* 7: 273-82.

Masterman, Margaret. 1962. "The Intellect's New Eye". *Freeing the Mind. Articles and Letters from The Times Literary Supplement during March-June, 1962*. 38-44. London: Times Publishing Company.

McCarty, Willard. 1993. "Handmade, Computer-assisted, and Electronic Concordances of Chaucer". *CCH Working Papers 3: 49-65.*

---. 2004. "Network with a Thousand Entrances: Commentary in an Electronic Age?" In *The Classical Commentary: Histories, Practices, Theory*. 359-402. Ed. Roy K. Gibson and Christina Shuttleworth Kraus. Leiden: Brill.

---. 2005. *Humanities Computing*. Basingstoke: Palgrave.

---. 2008 (forthcoming). "Response: Perspectives of Literary Theory". *Journal of Literary Theory* 2: 476-8.

McGann, Jerome. 2001. *Radiant Textuality*. Basingstoke: Palgrave.

---. 2004. "Marking Texts of Many Dimensions". In Schreibman, Siemens and Unsworth 2004: 198-217.

---. 2006. "Visible Langage, Interface, IVANHOE". In *The Scholar's Art: Literary Studies in a Managed World*. 148-71. Chicago: University of Chicago Press.

Merton, Robert K. and Elinor Barber. 2004. *The Travels and Adventures of Serendipity: A Study in Sociological Semantics and the Sociology of Science*. Princeton: Princeton University Press.

Milic, Louis T. 1966. "The Next Step". *Computers and the Humanities* 1: 3-6.

Opas, Lisa Lena and Thomas Rommel. 1995. "Introduction. Special Section: New Approaches to Computer Applications in Literary Studies". *Literary and Linguistic Computing* 10.4: 261-2.

Potter, Rosanne, ed. 1989. *Literary Computing and Literary Criticism. Theoretical and Practical Essays on Theme and Rhetoric*. Philadelphia: University of Pennsylvania Press.

---. 1991. "Statistical Analysis of Literature: A Retrospective on *Computers and the Humanities*, 1966-1990". *Computers and the Humanities* 25: 401-29.

Rauch, Jonathan. 2006. "Sex, Lies, and Videogames". *Atlantic Monthly* 298.4. www.theatlantic.com/doc/200611 (12 April 2008).

Richards, I. A. 1926. *Principles of Literary Criticism*. London: Routledge & Kegan Paul.

Rommel, Thomas. 2004. "Literary Studies". In Schreibman, Siemens and Unsworth 2004: 88-96.

Rosen, Robert. 2000. *Essays on Life Itself*. Complexity in Ecological Systems. New York: Columbia University Press.

Scharfstein, Ben-Ami. 1989. *The Dilemma of Context*. New York: New York University Press.

Schreibman, Susan, Ray Siemens and John Unsworth, eds. 2004. *A Companion to Digital Humanities*. Oxford: Blackwell. www.digitalhumanities.org/companion/ (9 March 2008).

Schrödinger, Erwin. 1992/1944. *What is Life? Physical Aspects of the Living Cell* with *Mind and Matter* and *Autobiographical Sketches*. Cambridge: Cambridge University Press.

Simon, Herbert A. 1957. Untitled lecture notes for a talk given to the Operations Research Society of America, 14 November 1957. *Herbert Simon Collection*, Lectures and Talks 1951-1991. http://diva.library.cmu.edu/webapp/simon/item.jsp?q=box00053/fld03599/bdl0001/doc0001/ (23 March 2008).

----- and Allen Newell. 1958a. "Heuristic Problem Solving: The Next Advance in Operations Research". *Operations Research* 6.1: 1-10.

-----. 1958b. "Reply: Heuristic Problem Solving". *Operations Research* 6.3: 449-50.

Smith, John B. 1989/1978. "Computer Criticism". In Potter 1989: 13-44.

Van Andel, Pek. 1994. "Anatomy of the Unsought Finding. Serendipity: Origin, History, Domains, Traditions, Appearances, Patterns and Programmability". *British Journal of the Philosophy of Science* 45: 631-48.

Wilks, Yorick Alexander. 1972. *Grammar, meaning and the machine analysis of language*. London: Routledge & Kegan Paul.

Wittig, Susan. 1978. "The Computer and the Concept of Text". *Computers and the Humanities* 11: 211-15.