Towards a Scholarly Editing System for the Next Decades

Peter Robinson[*]

Institute for Textual Scholarship and Electronic Editing,
College of Arts and Law, University of Birmingham
Edgabaston, United Kingdom
p.m.robinson@bham.ac.uk
http://itsee.bham.ac.uk

Abstract. As almost all texts of interest to scholarly editors exist in multiple versions, efficient systems of collation are crucial. The first part of the paper explains desiderata for computer-based collation. The second part of the paper argues that the most radical impact of the digital revolution is to transform scholarly editing from the work of single scholars, working on their own on single editions, to a collaborative, dynamic and fluid enterprise spanning many scholars and many materials.

Key words: Collation, scholarly editing, collaboration, digital methods

1 Introduction: a short history

As so often happens when one is asked to give a talk, and then to write it into a paper: the talk you thought you wanted to give when you were asked, isn't quite the talk that you want to give when you come to the moment., and when you come to write it into a paper, it changes again. When I was first asked to come to the Brown symposium, and was told what it was about, my head was full of collation, and so the original topic of the talk was Towards a scholarly collation system for the next decades. This is actually quite an interesting thing to have in your head, as I hope this paper shows. However, in the months after I suggested this topic in early 2008, I became increasingly interested in the environment in which a collation system must exist. However, because scholarly collation is critical to any scholarly editing system, if it is to be of any use, scholarly collation will still figure largely.

First, let me establish my credentials, as some-one who might know some things about using computers to help scholars create useful collations of variant sources. I have been working on computer-assisted collation systems for

[*] This paper is based on the talk given at the 2008 International Sanskrit Computational Linguistics Symposium held at Providence Rhode Island in May 2008. I am grateful to Peter Scharf and the Symposium Organizers for the opportunity to attend the symposium, and to the National Science Foundation which supported the symposium.

over twenty years. It all started in 1985 when Ursula Dronke, then Vigfusson Reader in Old Icelandic at the University of Oxford, suggested that I might edit the Old Norse poetic narrative sequence Svipdagsmal: there were, she said, a few manuscripts about and it might be interesting. There turned out to be 44 manuscripts: very interesting indeed. Right then, I bought my first computer, purely for word-processing the thesis. However this computer – an Amstrad PCW – also had a Basic interpreter and I found myself fascinated with programming it. So I attended Susan Hockey's courses on computer programming with SNOBOL for the humanities at OUCS, along the way also making the acquaintance of Lou Burnard. I had already started making electronic transcripts of the manuscripts, initially to make concordances. Now I developed a dangerous ambition: to write a computer program to compare the manuscripts. This became what you might call Collate 0: some 1200 lines of SNOBOL, later SPITBOL, code. For those who think intelligent life began with the iPOD, the SNOBOL family were pattern-matching programs specifically designed to process text: rather like PERL, I think.

Collate 0 created an efficient collation of my Icelandic manuscripts. Even better, I think, I was able to translate the whole collation output into a relational database and use database tools to help me work out the relations between the manuscripts. I described this work in two articles published in Literary and Linguistic Computing in 1989-90 on the collation and analysis of Icelandic manuscripts [1]. Following on from this, in 1989 Susan Hockey and I won a grant to develop my rather crude collation program, from something which could only be used by me, on materials which were just-so, only in the basement of the Oxford University Computing Services building between 3 and 4 am in the morning, to a program which could be used by any scholar, on any text, anywhere, any time.

So, in 1990, I began writing what became Collate 2 [2]. This was designed for the Macintosh computer, then running the state-of-the-art System 7. Collate 2 is still, eighteen years later, in continuous use in at least eight major editing projects. It still runs in Classic mode, only, and I have stopped answering questions on when there will be a Windows version. However, now that Apple have stopped all support of Classic, we have to hoard our old computers so that we can carry on running Collate.

Here is a list, very far from comprehensive, of who is using, or has used, Collate: the asterisk indicates that the scholar is still using collate.

Prue Shaw: Dante's Monarchia [3]

Prue Shaw and others: Dante's Commedia*

INTF-Mnster/ITSEE Birmingham: Greek and Latin New Testament* [5][4]

Michael Stone: Armenian texts

Sid Reid: Conrad texts*

Eric Sneddon: Old French texts

Godfried Croenen: Old French Chronicles*

Tommy Wasserman: Gospel of Jude [6]

Wendy Phillips-Rodriguez: Sanskrit*

The author, Barbara Bordalejo, and others: the Canterbury Tales* [7]
Dorothy Severin/Fiona Maguire: the Spanish Cancioneros* [8]
John Kilcullen: William of Ockham [9]
Michael Bakker and others: Old Church Slavonic texts

I think this shows, rather nicely, that Collate achieved at least part of our aim: that it became one of that rather select group of humanities computing programs actually used by someone other than the person who wrote it. Indeed, even more rarely: it is used by people I have never met.

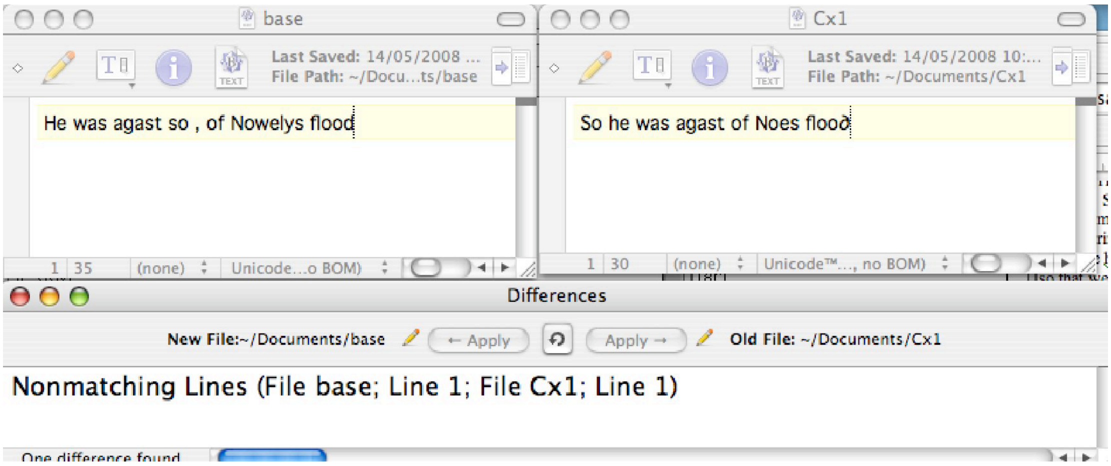
2 What scholarly collation programs must do

After all this time spent making collation programs, what have I learnt? We have learnt two principles which I think must be fundamental to any good collation program. Collate is built on these two principles, and hence its success. We have learnt one thing it does not do, which a scholarly collation program should do. And we have learnt that in a crucial respect the design of Collate was fundamentally flawed and we must do better.

The first of the two principles on which Collate is built was rather nicely put by a student of mine, who said it this way:

Scholarly collation is not Diff.

There are, of course, an immense number of text comparison programs out there, most of them with the venerable Diff algorithm lurking somewhere deep within them. Figure one is the result of one of these programs, working on variant texts of a single line.



base

Last Saved: 14/05/2008 ...
File Path: ~/Docu...ts/base

He was agast so , of Nowelys flood

1 35 (none) Unicode...o BOM)

Cx1

Last Saved: 14/05/2008 10:...
File Path: ~/Documents/Cx1

So he was agast of Noes flood

1 30 (none) Unicode™..., no BOM)

Differences

New File:~/Documents/base
Old File: ~/Documents/Cx1

Nonmatching Lines (File base; Line 1; File Cx1; Line 1)

One difference found

Fig. 1. The results of a 'Diff' comparison of one line of text in two sources.

All it does is tell us that there is a difference between these two lines. Well, thank you very much. Now, there are more sophisticated programs which will identify differences at the word level. And it is very tempting to think: all we need do is find one of these, tweak it a bit, and we will get the perfect collation. I can not say this strongly or loudly enough. Anyone who thinks that there will ever, ever be a collation program, a kind of superDiff, which will give scholars, for any text at all, exactly the collation they want is nave in the extreme. This does not stop people trying to write such a program and, even, funders giving them money to do this. But there never, ever, will be such a program.

There are two reasons for this. First, any kind of automated program can identify one thing, and one thing only: it can identify differences. Any scholar will tell you that differences are not variants. Depending on just how you are reading a text: some differences will be just, well, noise: only a few, perhaps a very few, are real variants, of real interest to real scholars. We scholars spend years looking at differences, filtering them according to our interests at the time, developing a sense of what is significant and what is not. The notion that somehow we can build this knowledge into some kind of reductive process is absurd.

The second reason is that even if you could, by some miracle, teach a machine exactly what is significant, what is not, you would still have to teach it to work out exactly what is the best way of presenting any given sequence of variants, at any particular moment. Let's take our sample two lines, to show what a scholarly collator has to do:

He was agast so , of Nowelys flood

So he was agast of Noes floo

Now, our first judgement is going to be: I think the comma in the first line, and the non-standard d on 'flood' in the second line are just 'noise'. So now I have these two lines to compare:

He was agast so of Nowelys flood

So he was agast of Noes flood

But, exactly how should I present this? I could present it as a single long phrase variant:

He was agast so of Nowelys] So he was agast of Noes

Or I could present it as three variants, as a sequence of addition/omission/replacement, thus:

He] So added

so] omitted

Nowelys] Noes

Or again as three variants, varying this, thus:

He] So he

so] omitted

Nowelys] Noes

Or, as two variants:

He was agast so] So he was agast

Nowelys] Noes

For me, the latter is much the best. Now, I could get the computer to generate any and all of the alternatives. But what I could never, ever, get it to do is decide which alternative, in every context, is the best.

When I think back on the decisions I took when designing Collate, the one decision which I think most absolutely right then, and right now, is this: that the scholar should at every point have the ability to intervene in the collation and fix the variation just the way he or she wants it. That is: the scholar can say: this spelling here is not a real variant, for example the two forms of ‘flood’ in this example. And the scholar can say: at this point, I want the variation shown exactly so, for example as a sequence of two variants in this line, not one or three. These two functions correspond to the ‘regularization’ and ‘set variants’ routines within Collate.

The one reason above all why Collate has achieved the success it has is because it has this faciltity: it allows the scholar to fix the collation exactly as he or she wants. This is why it is used by the New Testament group in Mnster, probably the most demanding collators anywhere. It still seems to me that this need is the most fundamental requirement of any scholarly collation program. Any program that does not offer this facility is just a dressed-up Diff program, a nice toy perhaps, but not a serious tool for serious use.

The second of the two principles on which Collate is built is this:

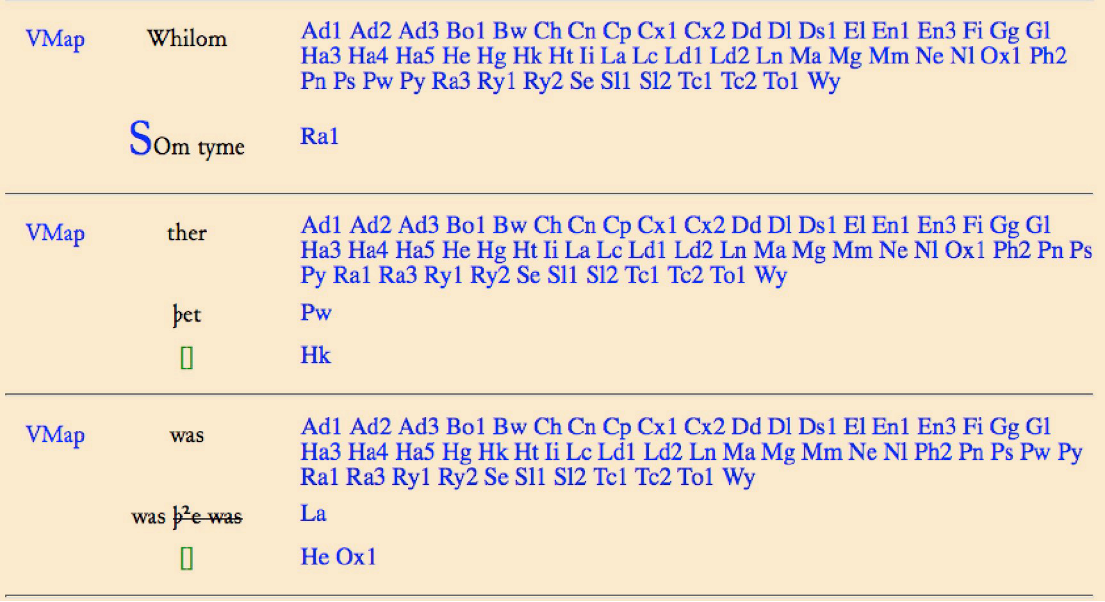
Collation is more than visualisation.

For most collation systems, the aim is to show the variation. Sometimes, they manage to show the variation in a rather beautiful form. But the problem is that they show it in one beautiful form only [1]. Again, one thing I did right with Collate, right back in the very beginning, was that I did not design the program so it could produce just one output. I designed it so that you could generate what you might call an intelligent output. Essentially, this distinguished all the different components of an apparatus the lemma, the variant, the witness sigil, and much more and allowed you to specify what might appear before and after each componetnt, and in what order the various components might appear. This gave exceptional flexibility to the Collate output. You could use Collate to generate complex print editions, using (or example) the EDMAC macros developed by Dominik Wujastyk and John Lavagnino; or you could turn out the apparatus in HTML, SGML or (lately) XML [11]. This allows you to make complex electronic editions, such as that in Figure 2.

This figure shows the collation for the fifty-four manuscripts of the first words of Geoffrey Chaucer’s Miller’s Tale [12]: beside each word and variant is a list of the manuscripts containing that word and variant. Notice that at this level, all the information about spelling variation is filtered out. You can see that spelling variation also, as shown in Figure 3.

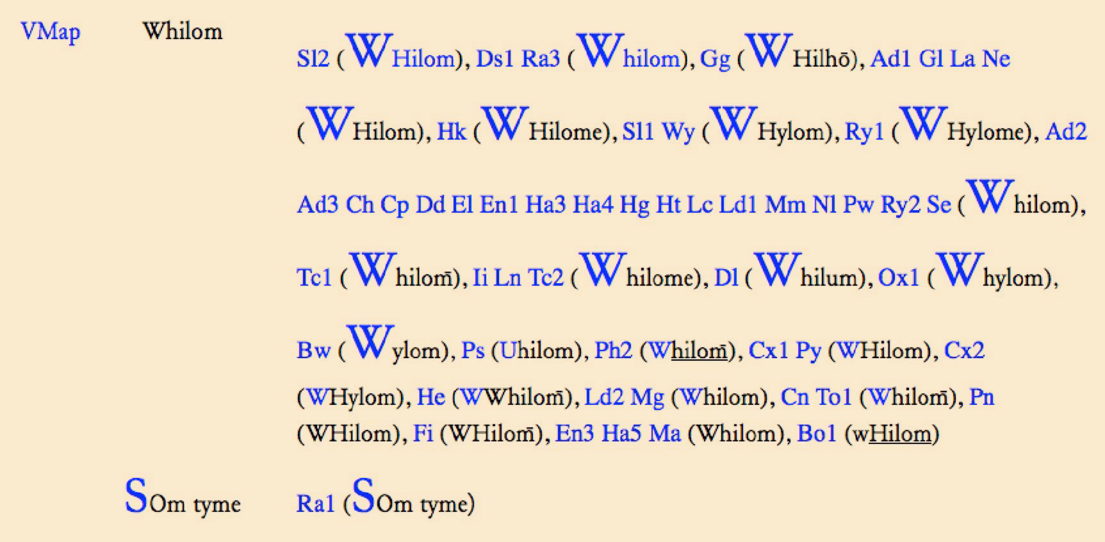
Alternatively, you could output the apparatus in a form ready for processing by an analysis program, to help you discover the relations among the witnesses. We have had great success in using evolutionary biology software to generate views of the relations between witnesses, as shown in Figure 4, for Dante’s *Monarchia* [3].

[1] For example, the excellent JUXTA program, developed in the University of Virginia [10]



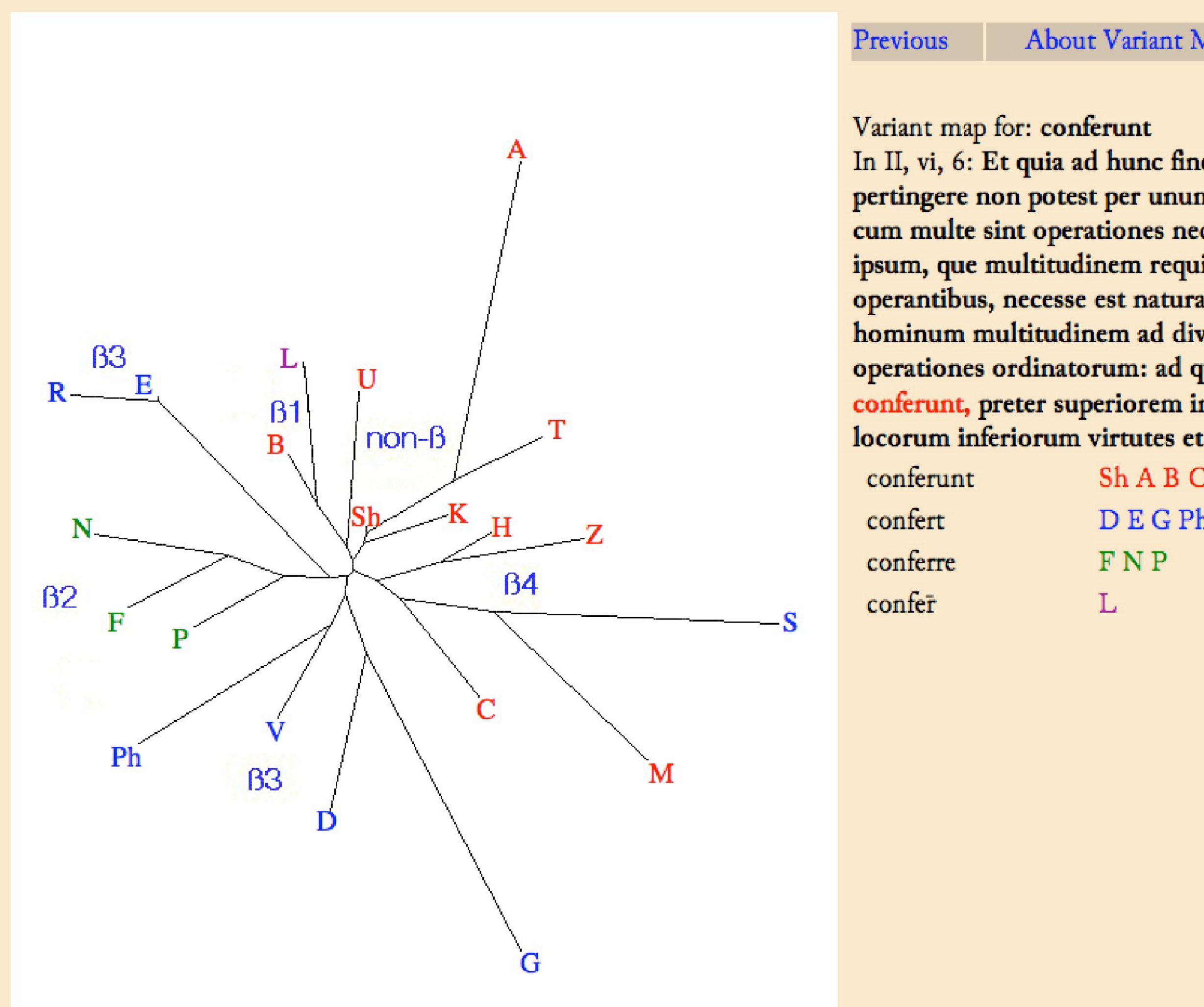
VMap	Whilom	Ad1 Ad2 Ad3 Bo1 Bw Ch Cn Cp Cx1 Cx2 Dd Dl Ds1 El En1 En3 Fi Gg Gl Ha3 Ha4 Ha5 He Hg Hk Ht Ii La Lc Ld1 Ld2 Ln Ma Mg Mm Ne Nl Ox1 Ph2 Pn Ps Pw Py Ra3 Ry1 Ry2 Se Sl1 Sl2 Tc1 Tc2 To1 Wy
	SOm tyme	Ra1
VMap	ther	Ad1 Ad2 Ad3 Bo1 Bw Ch Cn Cp Cx1 Cx2 Dd Dl Ds1 El En1 En3 Fi Gg Gl Ha3 Ha4 Ha5 He Hg Ht Ii La Lc Ld1 Ld2 Ln Ma Mg Mm Ne Nl Ox1 Ph2 Pn Ps Py Ra1 Ra3 Ry1 Ry2 Se Sl1 Sl2 Tc1 Tc2 To1 Wy
	þet	Pw
	□	Hk
VMap	was	Ad1 Ad2 Ad3 Bo1 Bw Ch Cn Cp Cx1 Cx2 Dd Dl Ds1 El En1 En3 Fi Gg Gl Ha3 Ha4 Ha5 Hg Hk Ht Ii Lc Ld1 Ld2 Ln Ma Mg Mm Ne Nl Ph2 Pn Ps Pw Py Ra1 Ra3 Ry1 Ry2 Se Sl1 Sl2 Tc1 Tc2 To1 Wy
	was ~~þ^{2}e was~~	La
	□	He Ox1

Fig. 2. The collation for fifty-four manuscripts of the first words of Geoffrey Chaucer's Miller's Tale.



VMap	Whilom	Sl2 (WHilom), Ds1 Ra3 (Whilom), Gg (WHilhō), Ad1 Gl La Ne (WHilom), Hk (WHilome), Sl1 Wy (WHylom), Ry1 (WHylome), Ad2 Ad3 Ch Cp Dd El En1 Ha3 Ha4 Hg Ht Lc Ld1 Mm Nl Pw Ry2 Se (Whilom), Tc1 (Whilom̄), Ii Ln Tc2 (Whilome), Dl (Whilum), Ox1 (Whylom), Bw (Wylom), Ps (Uhilom), Ph2 (Whilom̄), Cx1 Py (WHilom), Cx2 (WHylom), He (WWhilom̄), Ld2 Mg (Whilom), Cn To1 (Whilom̄), Pn (WHilom), Fi (WHilom̄), En3 Ha5 Ma (Whilom), Bo1 (wHilom)
	SOm tyme	Ra1 (SOm tyme)

Fig. 3. The collation for fifty-four manuscripts of the first words of Geoffrey Chaucer's Miller's Tale.



Variant map for: **conferunt**
In II, vi, 6: **Et quia ad hunc fin**
pertingere non potest per unun
cum multe sint operationes nec
ipsum, que multitudinem requi
operantibus, necesse est natura
hominum multitudinem ad div
operationes ordinatorum: ad q
conferunt, preter superiorem ir
locorum inferiorum virtutes et

conferunt	Sh A B C
confert	D E G Ph
conferre	F N P
confer̄	L

Fig. 4. The unrooted phylogram for the witnesses to Dante's *Monarchia*, generated directly from the computer-assisted collation.

I believe that a strong reason for Collate's success has been this flexibility of output. Put this together: the first thing Collate does well is to allow scholars to say: THIS is what the variation is. The second thing Collate does well is to allow scholars to say: I want the collation output in exactly this form. This is a very powerful combination, and any successor program that hopes to be as useful as Collate must include these capacities.

I said that we now know there is one thing Collate does not do, which we now think a collation program should do. Collate is extremely good at identifying variants of single words. It is very good in the basic identification of phrase variants, as a simple matter of saying 'this phrase in this manuscript corresponds with that phrase in that manuscript'. Look, for example, at this set of variants:

He was agast so

He was agast

So he was agast

He was so agast

He was agast and feerd

So was he agast

Now, Collate will do an excellent job of letting you see each of these phrases as variants of one another: in essence, of each phrase being a variant of each other, thus:

He was agast so] He was agast; So he was agast; He was so agast; He was agast and feerd; So was he agast

You can see, immediately, that there is something wrong here. The Collate presentation makes it look as if each variant is equally different from every other one. But clearly, this is not the case. You can see that the first, third and fourth differ only in the postion of the the word 'so'. Then, you can see that the second differs from these three by dropping the word 'so', while the fifth differs from the second by adding the phrase 'and feerd'. Finally, the sixth one has the word so in the same position as the third one, but changes the order 'he was' to 'was he'.

In fact, what we would like to see is something like Figure 5.

But Collate, as it is, just won't let you do this. A significant by-product of this is that Collate handles transpositions rather incompletely. Take this variation:

he was] was he

Collate identifies the phrase variation nicely. But Collate does not tell us what could be rather crucial information: that both witnesses do have the same two words, just in a different order.

What I have just described is known in genomic comparison systems as 'multiple progressive alignment': that is, you build up the picture of comparison piece by piece. [2] I think it is very important that a successor to Collate contains this facility.

[2] See the article by Matthew Spencer and Chris Howe on the use of multiple progressive alignment in evolutionary biology and its possible application to scholarly collation [13]

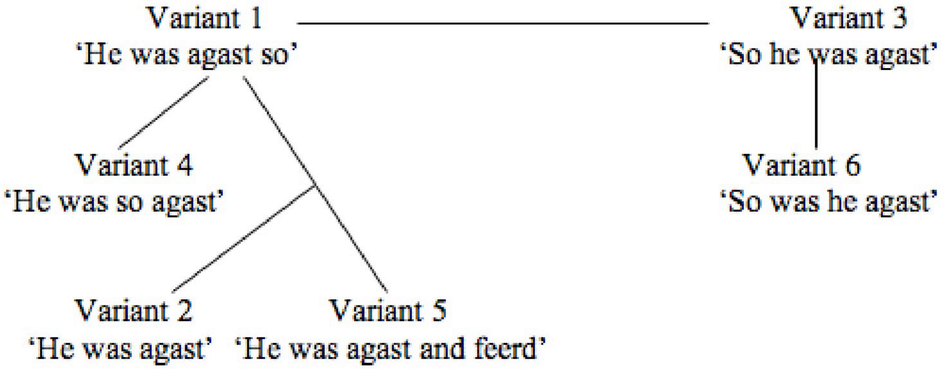


Fig. 5. Schema of the variants on this phrase, showing the relative closeness of each variant.

3 Towards collaborative scholarly editing

So far, the two good things Collate does and the one good thing it does not. I said too that we now realize that in one respect, the design of Collate is fundamentally flawed. It works like this. When I wrote Collate, I was a typical lone scholar, making an edition on my own. I found all the manuscripts, I transcribed them, I collated them, I edited the text. I did it all myself and I liked it that way. Further, every other editor I knew then worked the same way. So it is no surprise that the first and all subsequent versions of Collate were built with the same model in mind: that is, I ran the program on my computer; all the files were on my computer; everything was under my command.

At the time, I thought this was a virtue. But now, I can see that this is the biggest single defect of Collate. Quite simply, it was never designed as a collaborative tool and it is very awkward to use the program in a collaborative situation. This did not matter, at all, when scholars did not collaborate. But the single greatest effect of the digital revolution, in my opinion, is not that it is giving us wonderful digital libraries, with instant access to everything we want to see, or that it is giving us wonderful new publication possibilities, or that it offers all kinds of marvellous tools databases, analytic programs, and more. In my opinion, the single greatest effect of the digital revolution is that it is is empowering a new model of collaboration, and hence new modes of readership and study, among scholars, and between scholars and readers. Through well-constructed scholarly networks over the web, scholars and readers may not only look at materials: they may make them, annotate them, correct them, draw conclusions from them and then contribute to others their conclusions. Further, this may happen near-simultaneously: a library may contribute manuscript images in the morning; by midday a scholar has identified the text; by mid-afternoon a knowledgeable reader has transcribed it; in the evening, another scholar has collated this new transcript against other versions of the text, in other manuscripts. Nor is there any need for the scholars and readers to have any formal affiliation: they may be working far apart, without any project or other framework beyond common access to the web, shared interest and expertise.

The relevance of this to any large scale editorial project – indeed, to any editorial project at all – is obvious. Imagine that all the Sanskrit scholars of the world work in a single online workspace. In this workspace, some of you transcribe manuscripts; others of you collate the transcripts; others analyse the results of the collation. For some time I have been describing, in various articles and talks, what I have described as 'distributed, dynamic and collaborative editions'. The concept is not that there is a single system, a single set of software tools, which everybody uses. Instead, across the web we have a federation of separate but co-operating resources, all within different systems, but all interlinked so that to any user anywhere it appears as if they were all on the one server. To take the *Canterbury Tales* as an example: there might be transcriptions of the different manuscripts of the first line of the Tales available on different servers, made by different scholars, in New York, in Birmingham, in Utah. I can, any scholar can, access all these simultaneously: alongside images of the manuscripts, with collations, analyses, much else. Remarkably, we need very little to make this work: in fact, all we need are some basic agreements on how we name things, and on how we pass information between ourselves. We have much of this already, in the form of various metadata, encoding and web services protocols. We certainly have all the hardware and software we need, and more than enough people with all the skills needed to make this happen. I confess I do feel some frustration, with the various funding agencies who could set us well along the way by giving just a little bit of help. Instead, they appear to be fixated with Grand Single Solutions: I am thinking of the sort of projects funded by the NSF in America, and eScience in the UK, and by the Mellon Foundation, which show a predeliction towards grandiose projects emenating from prestigious institutions, Bamboo and SEASR being egregious instances of the genre [14][15]. Let me say this clearly, as most scholars seem afraid to say it: projects like these are vast wastes of time, effort and money. But actually, I don't think we need this funding – I think the idea is so good, and so obvious, that it will take off with very little or no funding.

I have, in other places, commented on what seems to me a very strange development [17][18][19]. Over the last fifteen years it has actually become harder for an ordinary scholar to create a high-quality scholarly edition in digital form. Indeed, it has become so much harder that a number of scholars and editorial projects have turned away from the digital medium: a development which really ought to alarm us. The answer to this flight from the digital is rather simple: we should make it as easy, or even easier, for a scholar to make a high-quality digital edition as it is to make a print edition. It seems to me quite absurd that, several decades into the digital revolution and with all the funding and all the effort lavished on it, we have not reached that point. We know that there are many scholars and readers with the interest and expertise to contribute greatly to the shared endeavour that is textual scholarship, in the editing of Sanskrit, of biblical materials, of medieval vernacular classics, everywhere there is text. The digital world should provide a space where any scholar with something useful to contribute may do so; where all may gain from the wealth of information so created.

You may ask: what am I doing to bring about this millennial vision? Several things. Firstly, I am a partner in an EU project named INTEREDITION, founded to bring about the creation of a supra-national infrastructure for digital textual scholarship [16]. The first workshop of INTEREDITION in September 2008 will address these issues. Even more immediately, the Birmingham and Munster partners in various New Testament editing projects are setting up what we call a Virtual Manuscript Room: a shared workspace in which all the many editors in these projects may work, as I have outlined it here. In our implementation of the Virtual Manuscript Room, one of the key factors will be a new version of our Anastasia software, redesigned so that she works directly from an XML database as the backbone of our live, interactive shared workspace. The first stage of this project has now received funding from the UK Joint Information Systems Committee [20]. Most immediately of all: since January a group of Dutch and German scholars have been working with me on what we call CollateX: the much-wished for successor to Collate.

I am suggesting here the future lies with a network of many servers, all holding different parts of an edition, with many other servers providing a range of services to the readers and scholars interested in this edition. Here, as a taste of what may come, is a tool we are making. We now have many manuscript images on the web. We now have many transcriptions of these same manuscripts on the web. The obvious thing to do is to align the text in the image with the text in the transcript. Here is a case where we have actually done exactly this, in the online version of Codex Sinaiticus [21], shown in Figure 6.

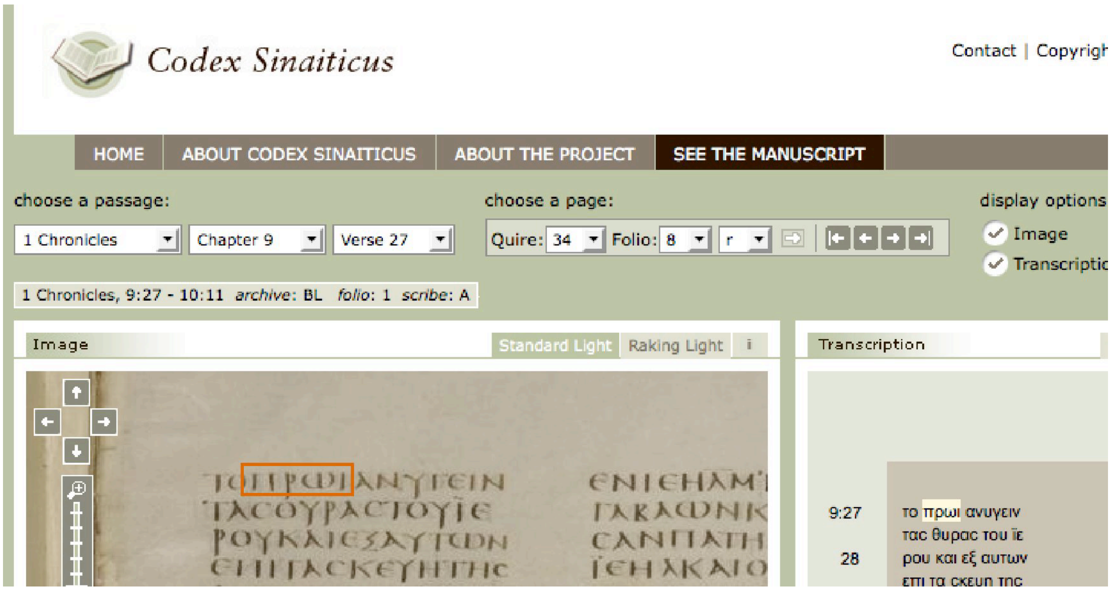
You can see that clicking on the word in the transcription highlights this word in the image.

We can imagine a vast number of potential research possibilities coming from this tool. For example: one could apply OCR software and pattern matching techniques to correlate every letter in the image with those in the manuscript, to explore questions of the characteristics of different scribal hands. One could also use automated analysis of the ruling and layout of the manuscript leaves to explicate its construction.

Perhaps best of all: tools like this will enable other people to do research we can not even imagine. For myself, I can not think of anything better.

References

1. "The Collation and Textual Criticism of Icelandic Manuscripts. (1): Collation. (2): Textual Criticism." Literary and Linguistic Computing 4 (1989): pp. 99-104, 174-181.
2. Collate. Computer Program. Release 1.0 April 1991; revised versions 1992-2008. Oxford: Computers and Manuscripts Project; Leicester: Centre for Technology and the Arts; Birmingham Institute for Textual Scholarship and Electronic Editing.



choose a passage: 1 Chronicles | Chapter 9 | Verse 27

choose a page: Quire: 34 | Folio: 8 | r

1 Chronicles, 9:27 - 10:11 *archive*: BL *folio*: 1 *scribe*: A

9:27 το πρωι ανυγειν
τας θυρας του ῑε
28 ρου και εξ αυτων
επι τα σκευη της

Fig. 6. Codex Sinaiticus on the web, showing the alignment of text and transcription

3. Shaw, Prue: Dante's Monarchia. Scholarly Digital Editions, Birmingham 2005, `http://www.sd-editions.com/Monarchia`, accessed 9 September 2008.
4. The Institute for New Testament Textual Research, Mnster, `http://www.uni-muenster.de/NTTextforschung/`, accessed 9 September 2008.
5. The Institute for Textual Scholarship and Electronic Editing, Birmingham, `http://itsee.bham.ac.uk`, accessed 9 September 2008.
6. Wasserman, T.: The Epistle of Jude: Its text and transmission. Coniectanea Biblica New Testament Series - CBNTS 43, Almqvist and Wiksell, 2006.
7. The Canterbury Tales Project, `http://www.canterburytalesproject.org`, accessed 9 September 2008.
8. The Cancioneros Project, `http://cancionerovirtual.liv.ac.uk/`, accessed 9 September 2008.
9. The William of Ockham Project, `http://www.britac.ac.uk/pubs/dialogus/ockdial.html`, accessed 9 September 2008.
10. The JUXTA project, `http://www.juxtasoftware.org/`, accessed 9 September 2008.
11. The EDMAC macors, `http://www.ucl.ac.uk/~ucgadkw/edmac/`, accessed 9 September 2008.
12. Robinson, Peter M. W. (ed.): The Miller's Tale on CD-ROM. Scholarly Digital Editions, Leicester 2004.
13. Spencer, M., Howe, C.: "Collating Texts Using Progressive Multiple Alignment", Computers and the Humanities (2004) pp. 253-270.
14. The BAMBOO project, `http://www.projectbamboo.org/`, accessed 9 September 2008.
15. The SEASR project, `http://seasr.org/`, accessed 9 September 2008.

16. The INTEREDITION project, http://interedition.huygensinstituut.nl/. accessed 9 September 2008.
17. Robinson, Peter M. W.: "Current issues in making digital editions of medieval texts–or, do electronic scholarly editions have a future?" Digital Medievalist 1.1 (2005) at http://www.digitalmedievalist.org/article.cfm?RecID=6
18. Robinson, Peter M. W.: "Where We Are with Electronic Scholarly Editions, and Where We Want to Be" Jahrbuch fr Computerphilologie Online at http://computerphilologie.uni-muenchen.de/ejournal.html, January 2004. In print in Jahrbuch fr Computerphilologie 2004, 123-143.
19. Robinson, Peter M. W.: "Current Directions in the Making of Digital Editions: towards interactive editions." Ecdotica 4 (2007) pp. 176-191.
20. The Virtual Manuscript Room, http://www.itsee.bham.ac.uk/projects/vmr/index.htm, accessed 9 September 2008.
21. Codex Sinaiticus on the web, http://www.codex-sinaiticus.net/en/manuscript.aspx, accessed 9 September 2008.