

Qualitative Spatial Representation for Material Documents

Mark Bell 

The National Archives, Kew, UK

John G. Stell 

University of Leeds, UK

Abstract

This paper proposes the use of Qualitative Spatial Representation (QSR) to model the complex relationships of paper documents and books, both in their physical and digital forms. We begin with motivating examples from The National Archives' holdings, and literature, to describe the myriad ways something as simple as text on paper can be manifested, and introduce appropriate terminology. The primary focus is on the challenges of working with documents comprised of multiple pages, and understanding the relationships between the pages themselves, text items which span more than one page. We end by presenting case studies using Allen Intervals to represent relationships between physical documents, their digitised representations, and the text on the page.

2012 ACM Subject Classification Computing methodologies → Spatial and physical reasoning; Applied computing → Document management and text processing; Applied computing → Arts and humanities

Keywords and phrases Qualitative spatial representation, QSR, spatial text, document understanding, information extraction, Allen interval relations, archives

Digital Object Identifier 10.4230/LIPIcs.COSIT.2026.3

Funding *John G. Stell*: Research supported by the UK's Economic and Social Research Council (ES/W003473/1) and the US National Science Foundation (NSF#2225179).

1 Introduction

The aim of this paper is to explore the role of Qualitative Spatial Representation (QSR) in the context of information within material documents, such as books and other media based on flat surfaces that can support text whether handwritten or printed.

A conventional book consists of a sequence of pages which follow each other in a straightforward linear order. This sequence in a physical book need not be followed by the reader, who can jump randomly from one page to another, or methodically move backwards and forwards in studying some aspect of the contents. However, the structure of the book itself still appears fundamentally as a sequence, despite our ability to explore in a non-linear fashion. The idea of a text as a space to explore has a form of parallel in the work of Mocnik and Fairbairn [24] who consider the ways maps can be enabled to support story telling. We argue that the interaction between spatial placement on a page and the spatial location of the page among other pages leads to a relatively neglected aspect of qualitative spatial information which is not as limited as it might initially appear, and which has practical applications. We present applications to information extraction from archives of material documents which also exist as digitised images, and applications to other areas too.

Today, books that exist as material copies are frequently experienced digitally as images or other forms of data on a screen. Texts that are “born digital” may only become physical copies as a derived form when printed out and do not initially exist as material objects. In this paper our concern is with material documents, motivated mainly by archival collections but also by contemporary uses of books in which the material is essential, such as “artists’



© Crown copyright. Re-used under the terms of the Open Government Licence v3.0; licensed under Creative Commons License CC-BY 4.0

17th International Conference on Spatial Information Theory (COSIT 2026).

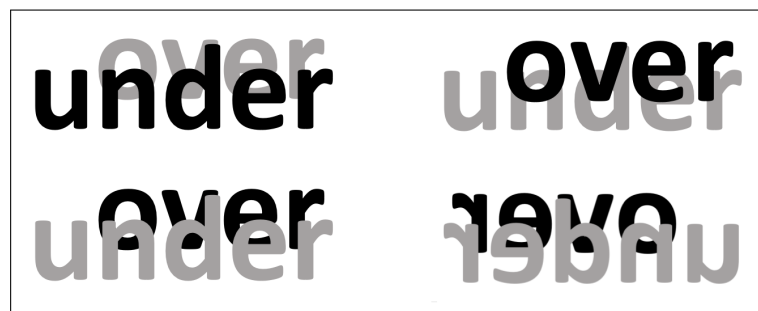
Editors: Sabine Timpf, Gabriele Filomena, Armand Kapaj, Rui Zhu, Nicholas Giudice, and Ed Manley; Article No. 3; pp. 3:1–3:19



Leibniz International Proceedings in Informatics
LIPIcs Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

3:2 Qualitative Spatial Representation for Material Documents

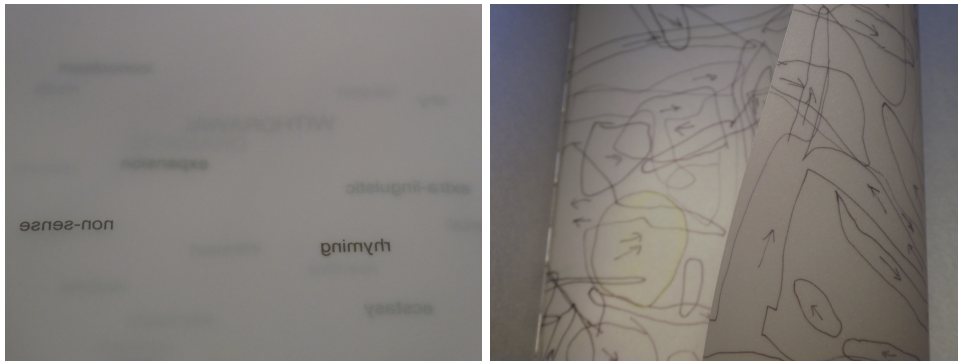
books” and some experimental literature. The digital is, however, a key feature of our investigation. In the archive which provides specific examples we discuss below, there is a very substantial body of digitised versions of material documents. As we explain in detail in the paper, the correspondence between the material originals and the digitised versions is spatially complex. This happens because the originals often depend for their interpretation on qualitative spatial relationships involving the material arrangement. For example, that some text or mark is located on the other side of the surface from another. This “other side of” in the context of a leaf in a book or other writing surface is a feature of the space of material documents that is the topic of this paper. It presents a practical challenge and is spatially complex because a digital document, such as a pdf of images of pages, may not explicitly record such spatial structure but it may be possible to deduce the structure from images.



■ **Figure 1** Spatial Relations between apparent layers. In the lower right is over over or under under?

We are mainly focussing our attention on books and documents as sources of information for the reader or viewer. This is significant because some spatial aspects depend critically on there being a point of view, or an observer experiencing the document in a particular orientation. This would not be the case if we only wanted to describe the structure of a book as a three-dimensional construction independent of any special viewpoint. The importance of the viewer can be explained in connection with Figure 1. In the figure are four pairs of words, one grey and one black. Looking at the top row first, the black word appears “on top of” the grey word in each case. We can imagine the words printed on transparent sheets with the sheet carrying the black word uppermost – that is closest to the imagined observer. In the bottom row on the left the grey word is on top. If we had transparent sheets arranged to show the example at the top right, we could obtain the lower left example by swapping the top and bottom sheets. Now consider the lower right example. Is the black word on top of the grey one, or is the grey one on top? There are only two possible answers and either can be argued for.

If we think in terms of transparent sheets bearing the words, then the lower right example would be obtained by turning over, as a single thing, the two sheets from the upper right example. The black word is now physically underneath the grey word. The alternative answer could be justified by saying that the marks constituting the black word need to be viewed “from the front”, otherwise it is not really a word, and similarly for the grey word. Thus even though to one observer the grey marks might be uppermost, there is an implied position which is actually “from the other side” and from there the black word would be closest to the observer.



■ **Figure 2** material layers and almost transparency. On the right from [10, p73] with a detail from Mehretu's *Migration Direction Map*, 1996, ink on Mylar. On the left a succession of pages printed on paper that allows a degree of transparency from [4].

Our position in this paper is that either of the answers about which word is “on top” and which is “underneath” is possible and will depend on the particular situation being modelled. Maybe both are needed simultaneously, in which case a formal model needs to incorporate at least one observer as a component in the model. It should be noted that the example here is not unrealistic as our potential applications include examples such as those illustrated in Figure 2 where pages are not opaque. Galleries, archives, libraries and museums have collections which may include artists’ books, experimental literature, or, as in the right of Figure 2, catalogues showing drawings made on transparent media that are reproduced on sheets that can be viewed from both sides. Such examples might need to be digitised or might exist within corpora which need to be searched. Search criteria might well need to allow a user to find items where something is “on top of” something else.

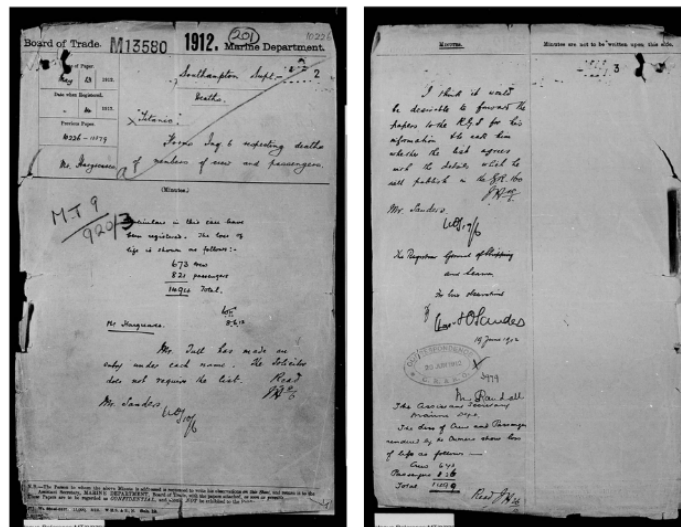
To give another example to show how relationships across different surfaces are less than straightforward, consider whether two pages “have the same layout”, whether page numbers are “in the same place” on different pages, and whether two pages “coincide exactly”. The notion of “the same as” leads to well-known philosophical conundrums involving identity, such as “the ship of Theseus”. In our context the issues relate more to geometrical concepts, such as two triangles being “the same” if one can be superposed on the other, which is itself by no means straightforward. Consider a book which has an opening showing two pages in which the page numbers are both positioned in the upper right corner of the page. That is the lefthand page has the page number close to the page edge that is attached in the binding, and the righthand page has the page number close to the edge that is free to move and parallel to the spine of the book. This is different from the conventional traditional layout in which layout on the left is the mirror image of layout on the right, but it does occur in many examples from the 20th century onwards. In one sense the page numbers appear “in the same place” on the two pages. In geometrical terms, by translating the lefthand page to superpose it on the righthand one the positions of the page numbers would coincide. However, if the book were closed, so the left and right pages were in contact, then the page numbers would not coincide. This again, is not an artificial example because experimental literature such as *House of Leaves* [11] discussed in Section 5.2, includes cases where the reader is expected to imagine a space formed in a closed book through a succession of flat parts on adjacent pages. Here the coincidence of things on the left and right halves of an opening closed up is part of the semantic content of the book itself.

The structure of the paper is as follows. In Section 2 we present the situation that has been the main motivation for this study. This involves the combination of archival collections, information extraction, and digital images which are needed as part of the process of extracting information from the original documents. Having given motivating examples, we use Section 3 to review related work including document understanding, text where spatial arrangement conveys meaning and the use of Qualitative Spatial Representation to describe these spatial arrangements. At the end of Section 3 we introduce the variety of material structure of physical documents but we leave more detailed exploration of this area until later. Section 4 presents detailed case studies showing how spatial arrangement and material structure are important for understanding documents and extracting information from them. The key examples use specific documents from The National Archives. In Section 5 we consider the role of folds, tunnels, slits, and other holes in material documents. The section also suggests how the formal theory of qualitative spatial relationships might be developed to meet some of the challenges that we have highlighted. The paper ends with overall conclusions and topics for future work.

2 Archival Motivation

The National Archives is the archive of the UK government and holds around 1000 years of history. Although the vast majority of the holdings are only available in their physical form, around 10 million documents have been digitised. With the advent of handwritten text recognition (HTR) technology, and large language models, it is becoming possible to transcribe these documents at scale for computational analysis. In addition to this, initiatives such as IIF are encouraging and making it easier for institutions to publish digitised collections online and making their textual content searchable. The reality of working with archival documents is that they are highly variable. Not only in terms of the content and layout of the pages themselves, but how they are stored, bound, and digitised. A case in point is catalogue reference MT 9/920/C at The National Archives, described as “Titanic. MT 9. Marine correspondence including SOS telegram from the Titanic, 1912”. This is available to download at <https://discovery.nationalarchives.gov.uk/details/r/C5907412> as a 5 part PDF file. In total the PDFs contain 618 pages consisting of digitised images of a physical, paper file. The first image is of a manilla folder, which is then followed by a cover sheet (Figure 3) which details the contents and a section for minutes, which appear to capture notes and directions for onward processing of the information contained in the file. On the front page the minutes take up the lower two thirds of the page, while on the back they are separated into two columns one labelled “minutes” while the right hand one is clearly labelled “Minutes are not to be written on this side”. What follows are several lined sheets of paper, including an introductory note and lists of passengers and crew on the ship. These lists consist of multiple columns which are laid out across two pages - starting on the back of one sheet, and going across to the front of the next. Each two page table is preceded by a page detailing the role of the people named in the table (passengers, crew, stewards, restaurant staff etc.). The tables for crew and passengers contain different columns. The remainder of the images are a mix of the aforementioned minute sheets, tabular information, correspondence on letter headed paper, envelopes, telegrams, handwritten notes, typed interview transcripts, pages with text cut from a different page glued to them, newspaper clippings, typed reports, bound instruction booklets, maps, and a few more manilla envelopes.

This is just one item of millions which have been digitised and are in theory ripe for exploitation by digital humanists with computational tools. Recent experience of the Material Culture of Wills project, suggests the practice is not so easy. Probate copies of wills are held



■ **Figure 3** A sheet used for capturing minutes from the Titanic Correspondence file (MT 9/920/C).

in large bound volumes which were digitised from microfilm over 20 years ago. While the scribes wrote the wills in a continuous fashion, one will continuing immediately after the previous one, digital access is in the form of one PDF per will. The ordering of the wills in the catalogue is by date of probate not page ordering in the volume. Reconstructing the images of a volume of wills in their original order, with start and ends of each will marked and linked to its catalogue entry, was achieved through a combination of cryptographic hashes, text extraction, layout analysis, and manual intervention. It is clear from this example that digitisation was undertaken with providing access to individual wills in mind with no metadata retained to connect one page to the next, and that even in these structurally simple documents, layout has semantic meaning. If working with images of pages in bound volumes is a challenge at scale, what hope do we have with files such as the Titanic correspondence with all their variety?

It is not feasible to re-digitise millions of documents with much more detailed metadata about their material origin. Even if it were, it is less than obvious exactly what needs to be recorded and what structure might be appropriate. This raises the same issue as is found in working with the existing digitised versions. Experts in the documents may understand how different indentations, groupings of names, locations of signatures or marginal notes, for example, all play a part in extracting meaning. This knowledge, however, has to be communicated to computational experts who themselves need to describe the relevant patterns to process the data. What is needed is a language to express the spatial structure of the material in a way suitable for archival and other applications. This paper is a step towards such a language.

3 Related work

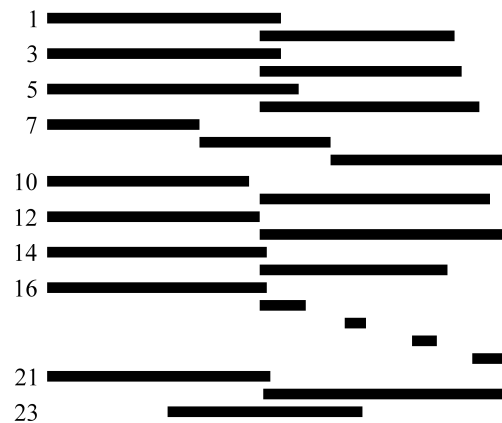
3.1 Document Understanding

The field of Document Layout Understanding is focused on identifying and interpreting structural aspects of documents. The DocLayNet dataset defines 11 distinct classes of structural element: namely caption, footnote, formula, list-item, page-footer, pageheader,

picture, section-header, table, text, and title. Layout recognition is also a key component of automated text recognition pipelines which translate segmented lines or words from an image into text. Another research area covering document layout is Document AI [19] which refers to the techniques for automatically reading, understanding, and analyzing business documents. A survey of papers [22] published at ICDAR, the bi-annual International Conference for Document Analysis and Recognition, held since 1991, identified three main tasks being addressed by researchers: pre-processing for image quality, identification of regions with homogeneous content, and textual content recognition. Within the second category several sub-tasks were identified: text line segmentation, baseline detection, text detection, symbol detection. The authors further summarised the application of deep learning technology as mapping inputs (e.g., a digitised page) to outputs (page, region, table, text-line, bounding box, transcribed text, image patch, image), with the algorithms generally classifying at the pixel level. Wang et al. [30] build a layout graph where nodes are elements extracted from PDF metadata, and edges are created between nearest neighbours in each direction (left, right, up, down). A graph convolutional network is then trained to classify each element as one of the 11 classes of DocLayNet or the 5 of PubLayNet, another layout analysis dataset. The GraphDoc dataset [8] contains 80000 page images with associated layout graphs depicting spatial (left, right, above, below) and logical (parent, child, sequence, reference) relationships between page elements (e.g. paragraphs, diagrams) found in scientific literature. Kim et al. [20] introduce the Document Understanding Transformer which extracts structured data from complex documents such as invoices and receipts for downstream tasks such as Information Extraction, and Visual Question Answering (where the task is to answer natural language questions about the content of an image). While relationships between elements have been used as features to classify elements, or documents, by type, to the best of our knowledge we have found no evidence of documents being represented by those relationships. Furthermore, the literature is focused on layouts of individual pages rather than multiple pages. Two data standards used in the humanities, the International Image Interoperability Framework (IIIF) and Text Encoding Initiative (TEI), are able to represent certain aspects of layout. The IIIF annotation standard stores bounding box coordinates of text on the page, and digitised images are held in a sequential order for viewing, but there is currently no concept of two sides of a sheet within the standard (although a working group has been set up to discuss this). TEI includes metadata tags for the cover of a book, front matter (such as contents), the main body, and back matter, and also has the concept of a group for material which are related in some way. It can also represent boundaries between pages and lines, chapters, paragraphs, and many of the concepts we may use to describe documents and the text they contain. It does not capture spatial relationships between elements on a page though.

3.2 Spatial Text

Text, whether handwritten, typed, or printed, flows in a direction dictated by the language; left to right (western), right to left (arabic), down the page (chinese). It also has orientation on the page, and from any particular perspective could be the right way around, upside down, at an angle, or even backwards (da Vinci famously wrote with a mirror). How much text fits on a page is dictated by the size of the page. A letter may cover both sides of a single sheet of paper, tabular information may continue on every page of a ledger, or its columns may span each double page spread. We may logically wish to think of a piece of text occupying a single page, even if physically it spans multiple sheets. Equally the text on the reverse side of a leaf may be contextually related but not a continuation of the text on the front, or it may



■ **Figure 4** Line structure from Atkin's poem *Cannulation*. Image source [29, p23].

be completely unrelated beyond sharing the same piece of material. There is also a temporal element to text. A typed document may be heavily annotated with handwritten comments and corrections. It was common in centuries past for replies to letters to be written between the lines of the original, or even perpendicular to them, to save paper (cross-written letters).

Written language is read not only from screens but also from physical media such as paper. In much communication, the layout of the words and other building blocks of language is a significant part of the meaning. A particularly obvious example is concrete poetry [5], where the placement of the words on the page can be as important as choice of the words themselves. Poetry in which words are arranged to form pictorial images has a long history [17, 16, 5]. Well known examples include George Herbert's *Easter Wings*, poems in the form of altars, and the arrangement of words to imitate the shape of the tail of a mouse in *Alice in Wonderland* by Lewis Carroll.

Beyond poetry with a specifically pictorial function, the layout of poetry is often deliberate to connect with ideas in the wording. The lines in *Cannulation* by Polly Atkin [3] include the words:

the self made small, importing little one dot iot iota drip.

but the layout presents this in lines 16 to 20 of the text of the poem (see Figure 4) as follows:

```
the self made small, importing little
                                one dot
                                    iot
                                        iota
                                            drip.
```

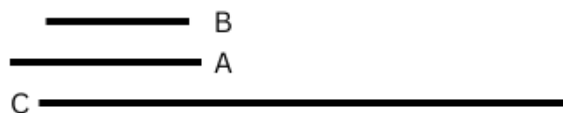
Here the succession of short lines progressing to the right relates to the process of successive dripping or measured puncturing. This layout is part of the poem itself in as fundamental a way as the choice of words themselves.

This kind of example concerns spatial relationships on a single page. The use of space by the author, and what matters to the reader or viewer, is not the exact coordinate location of words, but qualitative spatial relationships. In this context such relationships include the way that two words on successive lines appear with the later one further to the right than the earlier one but with a clear space vertically between the end of the first and the start of the second. For this spatial situation Allen Interval Relations have been found useful, as we explain next.

3.3 Allen Interval Relations

It has been suggested [29] that the Allen interval relations [2] can be used to record spatial relationships between lines of text. This is a different application area from the original context of intervals of time described by Allen. The Allen relations can be situated within a larger body of work on qualitative spatial representation. This area is concerned with general relationships such as “next to”, or “inside”, or “touching only at the borders”. These relationships have had multiple applications to different kinds of spatial information, as explained in the survey [9]. The relevance of some of this work to the Digital Humanities is discussed in [29].

When representing layout with digital methods the position of text on the page is usually expressed in terms of pixel coordinates which may describe a rectangular bounding box, or a complex polygon, enclosing the text. While it may initially seem more appropriate to use a two dimensional QSR for rectangles, in practice most neighbouring bounding boxes will be disconnected - generally an element of text occupies its own discrete space. Bounding boxes can instead be represented as a pair of Allen intervals covering the vertical and horizontal extremities. Any element may be a part of multiple intervals: the horizontal span of a glyph will be part of a word, which is part of a line. Vertically the line is part of a paragraph, which is part of a page. We have a preference for the Extended version of Allen Intervals proposed by Renz [27], as we are able to express degree of overlap, although often only the degree from the perspective of one interval is important rather than both. In Figure 5, interval A overlaps both B and C in standard Allen terminology. Under the Extended algebra, the most part of A overlaps with both B and C, but while the most part of B also overlaps with A, only the least part of C overlaps with A, or $mol(A,C)$ using the Extended Allen notation. The situation will dictate which is important so we separate degree of overlap for each interval from the Allen relationship. In this case we may have $mo?(A,C)$, which indicates that mol and mom (Most Part overlaps Most Part) are equivalent.



■ **Figure 5** Degree of overlap between intervals.

Other work using extended Allen relations can be seen in [6]. Unlike the present paper, the work described there is focussed on relationships between elements on a single page. Here we are dealing with the more complex case of relationships that may hold between entities on different pages.

3.4 Physical Documents

The physical structure of documents, whether handwritten manuscripts or printed books, or more generally forms such as scrolls, rolls, etc, is already a highly developed area of study. For example, see [15] for just one aspect of printed books, and [23] for the picture-scroll in China. However, the interaction of physical structure and information extraction, especially in combination with the need to process information computationally using existing images of the documents appears to have received relatively little attention. This situation is our primary motivation, as explained in Section 2.

The complexities of elaborate documents with multiple folding parts, loosely inserted pieces, pockets containing smaller documents, pages with holes, creases, etc are too extensive to address in full detail in this paper. It is even more challenging when such documents have been digitised in the past in ways that mean it is unclear how to reconstruct the original physical arrangement from a collection of digital images. We are able to illustrate some of the issues in a detailed case study of archival material in Section 4. Some of the challenges that occur with folds, cuts and holes are discussed in the subsequent section. We also use Section 5 to present a proposal for how formal theories of qualitative spatial relations might be developed to support the representation of relationships that occur within a space that is not rigid but can be modified, for example by turning the pages of a book.

4 Case Studies

Following analysis of a selection of digitised documents from TNA's online catalogue, and discussions with records experts, we have identified several common patterns for multi-page documents, in terms of both content and materiality, which we will model using QSR. We have not included rolled volumes in our analysis, as they are not turnable pages. Some of these patterns can be seen in Figure 6.

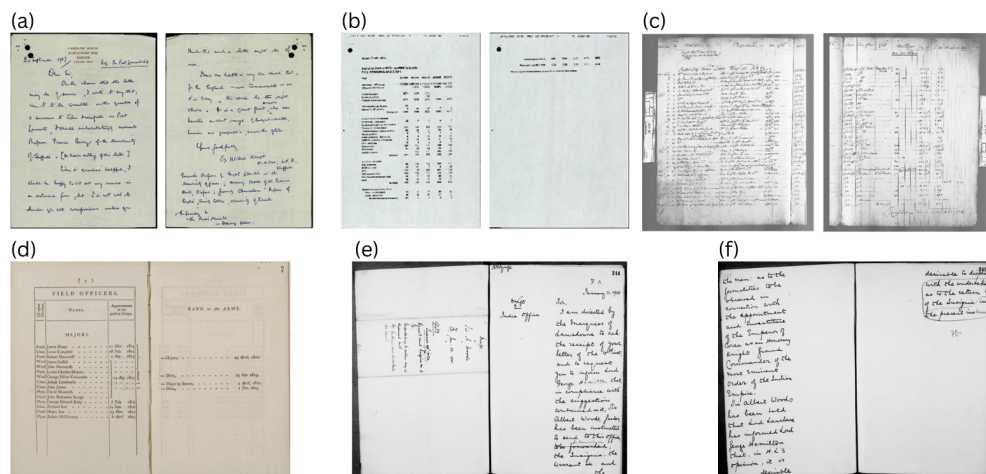


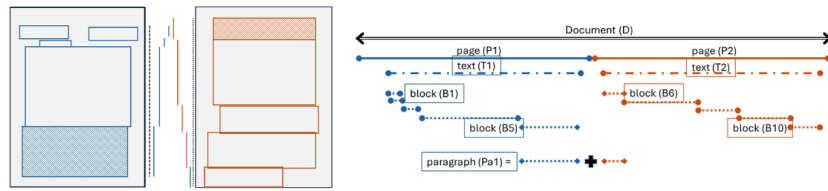
Figure 6 Selected double page layouts: (a) letter on two sides of sheet; (b) table continuing onto next sheet; (c) table with columns across two pages; (d) complex tabular information across two sheets; (e) open notebook pages, descriptive label at right angles, followed by text aligned to right and left hand margin note; (f) open notebook pages, opposite alignments leaving central margins.

Content patterns:

- Text, continues to next page (Figure 6(a))
- Text, related on next page but not continuous
- Tabular information, continuous over page (Figure 6(b))
- Tabular information, continuous across pages (Figure 6(c))
- Formatted text (e.g. indented), continuous over page
- Re-used material (e.g. reply to letter in line)

Material patterns:

- Single sheet
- Single sheet, folded
- Multiple sheets bound together (e.g. by treasury tag, or folder)



■ **Figure 7** Left: A letter covering two sides of the same sheet, represented by bounding boxes. The vertical sides of the boxes form a set of intervals; Right: Allen interval representation of the document, pages, text, and text blocks.

- Bound book, horizontal open (left-right, right-left)
- Bound book, vertical open (bottom-top, top-bottom)
- Bound volume, contents originally individual sheets

4.1 Single sheet, continuous text

The document in Figure 6(a) is a letter written on two sides of a sheet of paper. We can model this document with bounding boxes around discrete blocks of handwritten text (e.g. the date, each paragraph, the signature). Between the two page images we have represented the vertical edges of the bounding boxes as a set of Allen intervals. The dashed and dotted lines represent the text on each page.

Figure 7 visualises Allen intervals for the entire document, each page, the entire text of each page, and each block of text identified by the bounding boxes. Additionally we create a new interval (Pa1) which represents a concatenation of the two shaded blocks in Figure 7.

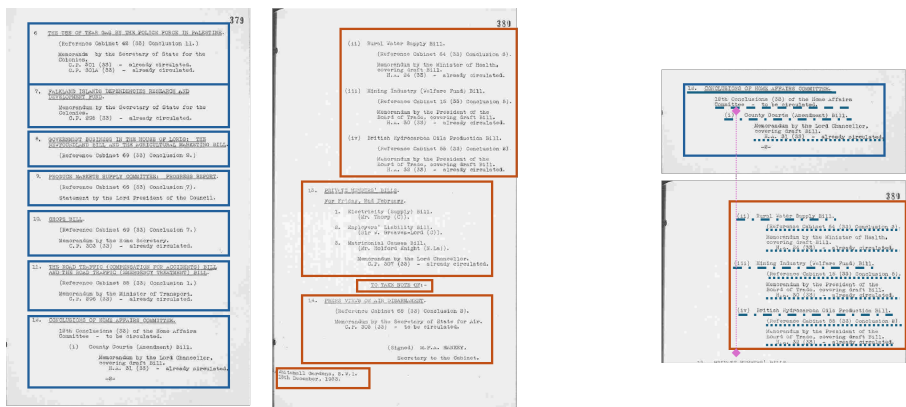
The following example relations can be described between these intervals:

- startedBy(D, P1);
- finishes(P2, D);
- contains(P1, T1);
- contains(P2, T2);
- startedBy(T1, B1);
- finishes(T1, B5);
- finishes(T2, B10);

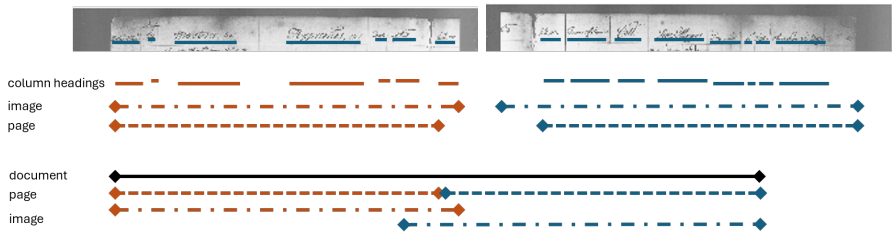
Physically the first handwritten line on page 1 begins just below the two holes in the sheet, but logically it is the beginning of the letter. So we could logically state `startedBy(P1, T1)` instead of `contains(P1, T1)`. Blocks B5 and B6 are physically separate on different sides of the page, one at the bottom, one at the top. In terms of the text on the page they are a single paragraph, only broken up because of limitations dictated by the size of the paper. Logically, B5 and B6 make a paragraph, Pa1, which overlaps both pages, giving us `overlaps(Pa1, P1)` and `overlaps(Pa1, P2)`.

4.2 Continuing indentation

Figure 8 shows indented text covering two pages. Each section, marked by bounding boxes, has multiple levels of indentation. The content of the final block of page 1 continues over to the first block of page 2. There is continuity in the indentation so the text on page 2 starts justified to the right, and then unindents after the block ends. Physically the indentation is



■ **Figure 8** A document where indentation continues over the page. The indentations do not accurately line up, as indicated by the vertical dotted line.



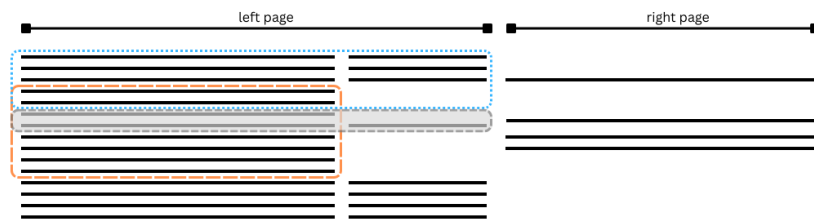
■ **Figure 9** Column headings of a table which go across two pages.

not consistent between the two pages, with items (ii-iv) starting noticeably further to the right than item (i), as seen in the right hand image, but logically the four items all start together.

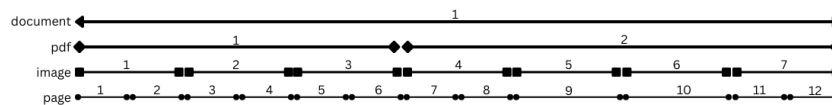
4.3 Tabular layouts

Figure 9 shows the columns headings of a table which spans across two pages. This representation shows the need for the digital as well as the physical as there is some overlap between the images with the first column on the second page appearing in both images.

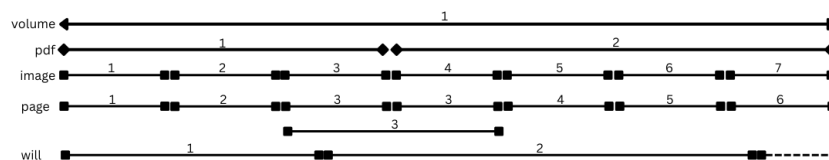
A more complex tabular form is depicted in Figure 6(d). On the left hand page is a three column table. The table is a list of people, falling under the category of FIELD OFFICERS – a table which continues over several double pages of the book – and a sub-category of MAJORS. Names are listed in the order in which they were appointed to the rank, and are grouped by the use of brackets when more than one person was appointed on the same day. They are further grouped according to what appears on the right hand page. The title is RANK IN THE ARMY, which does not always appear on the right hand page when there is an entry. When the right hand page is completely blank it has been digitised separately to the left hand page. The entries in the right hand page are aligned with the centre of the outermost brackets on the left hand page. In addition, we note the word Ditto in some entries which means to take the same value as the entry immediately above. In the first case of Ditto, there is a noticeable space between it and the entry above. As a final complication, the sections of the book are catalogued by rank so Majors appear in a different PDF to the Captains, and so on.



■ **Figure 10** Complex tabular information where entries which share attributes (e.g. a date) are bracketed together. The text on the right hand page is not in the table but is aligned with its entries.



■ **Figure 11** Multi-PDF document represented as intervals.



■ **Figure 12** Volume of wills in PDFs represented as intervals.

The two pages are depicted as grouped intervals in Figure 10, with the groupings inferred from the bracketing represented by boxes around the appropriate intervals. We see that an interval may be a member of more than one group, with the consequence of group membership being dictated by the spatial relationships with other items on the same page, or the next one. The document itself is modelled as intervals in Figure 11 to illustrate the complex relationships between the physical and digital forms. Most images contain a double open page, while some show only one. The equivalent diagram for the wills mentioned in the introduction is shown in Figure 12. Extra complexity is introduced by the duplication of images between consecutive PDFs. We have both *Finishes*(Page3, PDF1) and *Starts*(Page3, PDF1), but *Finishes*(Image3, PDF1) and *Starts*(Image4, PDF2). The second Will overlaps both PDFs and images 3-6 but neither starts nor finishes any.

To represent the relationship of one will to the pages of the original book there is another technique besides giving an Allen relations between the will and each one of the pages. As each will is a continuous section of the book, without any gaps or interruptions, the relationship can be represented by a single Allen relation between the will and the sequence of pages in which the will appears. For example, in Figure 12 Will2 is found in pages 3–6. Thinking of this sequence of pages as a single interval denoted *Extent2*, the relationship *Contains*(*Extent2*, Will2) holds.

As well as capturing the logical flow of one page to the next we must also represent the physical relationships between pages to enable sense-making in the digital form. When documents are digitised they may be captured as double or single pages. When sifting

through digital images it is not always clear whether two images are of front and back of the same sheet, or whether they could be viewed side by side in their physical form. The pages of reporter style notebooks are flipped upwards, depending on the preferences of the owner, the writing on the upper page may be upside down or the same way up as the lower page when photographed. We have also encountered instances of standard opening notebooks where the book has been rotated through 90 degrees for digitisation due to the orientation of the writing. A page may therefore be represented by three orientations: natural reading, physical, and digitised. However, the orientation of the text to the page may vary in some documents. Qualitative Direction Calculi [26] such as the STAR algebra can represent relative orientation of a page to a viewer but further reasoning is required to derive the optimal orientation for a specific purpose. These details matter both to the person viewing search results in the image, and to an automated text recognition process. Another common pattern we see is for a space to be left on the side closest to the fold in the centre of a book when writing, as seen in Figure 6 e-f. This space acts as a margin, although it is not specifically marked as such. Any text in this space should be treated as marginalia, additional but related to the main text. The spatial relationship between text and marginalia alternates by page, depending if it is on the left or right on the fold.

5 Material Structures and Relationships

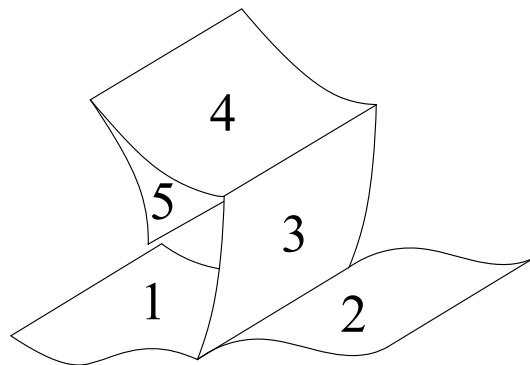
On a single page there are relationships between lines, between blocks of text, between the margins and blocks of text, etc. In a document with multiple pages all of those relationships are important on each page, but in addition there are relationships between the two sides of one leaf, as well as between leaves that share a common edge. There are relationships that depend on the viewer and on temporary configurations: two surfaces may be in contact when a book is closed and consequently be simultaneously visible when the book is open. However, once our books include pieces that fold out, two surfaces that can be seen simultaneously when the book is open need not be in contact when the book is closed. The complexity of the topic precludes a comprehensive account at this stage, but we draw the reader's attention to several ways that physical structure can impinge on the interpretation of information.

5.1 Folding

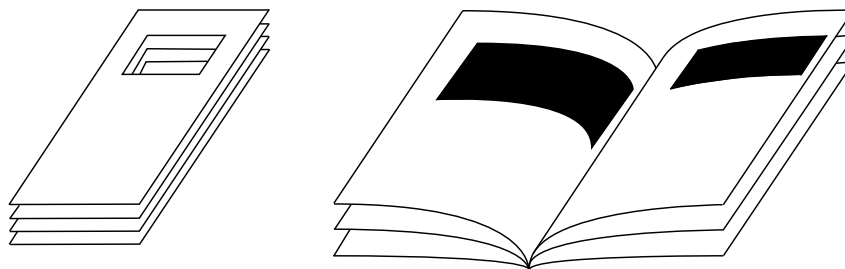
Consider maps that fold up. All of the sections on one side of the map are simultaneously visible when that side is uppermost to the viewer, but only a limited number of the sections would be in contact when the map is folded up. In the case of a folding map there are usually both horizontal and vertical folds and associated adjacency relations between sections in both dimensions.

There is a large literature on aspects of folding, both historical [14] and for practical applications in computational origami. However, most of this work is in quite a different direction to ours. For example, Justin [18] discusses questions such as the number of different ways a sheet such as a folding map can be folded up, or in a simpler case, the number of ways to fold up a sequence of panels arranged in a linear sequence when unfolded. It is notable that folding is of importance in archives in several ways such as the use of folding as a security device. This happens where a complex sequence of folds was used that would be almost impossible to restore if a document were intercepted and unfolded en route.

Consider Figure 13, which shows diagrammatically a book with two leaves (1 and 2 in the diagram) and a fold-out made up of three parts (3, 4, 5 in the diagram). Unlike the considerations in [18] about the combinatorics of folding up the 3-panel leaf 3-4-5, we



■ **Figure 13** Diagram of part of a book with a fold out in three panels.

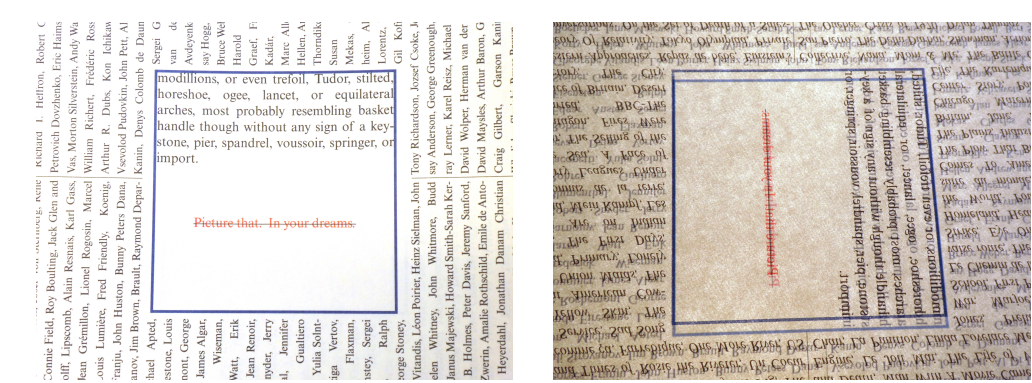


■ **Figure 14** Books exhibit spatial relationships beyond those on a single page. A succession of pages with holes produces a tunnel when stacked together; an opening seen by a reader as opposing pages becomes two touching surfaces when closed. The reader can imagine tunnels without physical holes when regions on touching surfaces match up through several successive leaves.

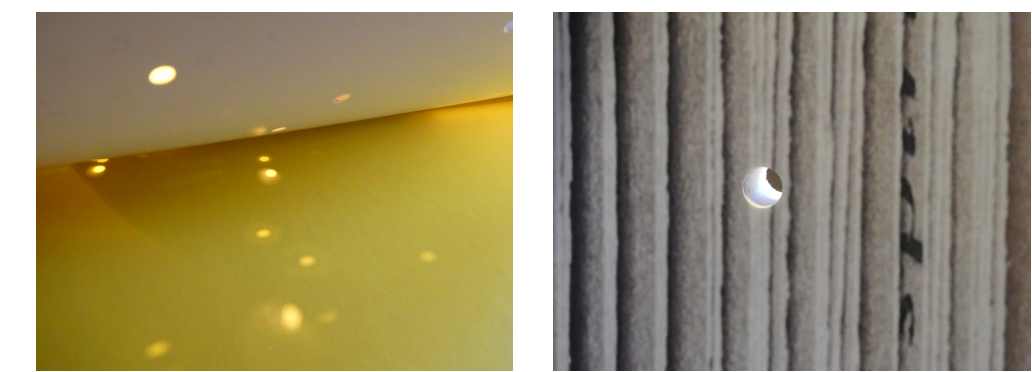
generally assume a given way to fold it (which may be determined by the design of the book or historical precedence). One way in which we diverge from most studies of folding is that we are concerned with not just the material leaves and panels, but also with what may be present on the surfaces. Another is that qualitative relations have not been explored in this setting, whereas computational origami is generally concerned with geometric precision and complex three-dimensional forms.

5.2 Conceptual Holes

In *House of Leaves* by Michael Danielewski [11] the spatial structure of a book is essential to the content. Blocks of text in different typefaces are used in such a way that the reader is not following a one-dimensional progress through a sequence of words, but is required to navigate a network in which parts of one page link to parts of another. In places the text in a part of one page corresponds on the opposite side of the same piece of paper bearing the following page, to the same text repeated but transformed in a reflection. This provides the effect of the text on the second page functioning as the view of the text on the first page but seen from the opposite side, see Figure 15. This example of a relationship across two successive pages on a single leaf is extended further in the book as it forms part of a sequence of square blocks of text over several pages. This sequence, when the book is closed, makes up a tunnel through the leaves. The spatial structure of the tunnel is related to part of the narrative involving a tunnel itself. Here, words relate spatially not just to words on the same page, but to words elsewhere in a three-dimensional stack of pages.



■ **Figure 15** In *House of Leaves* [11] a sequence of pages have the two sides printed with what you would see if the paper were transparent. Although the practicality of printing this exactly has been challenging, the spatial relationship which is significant for the reader is the intended one rather than its physical realisation.



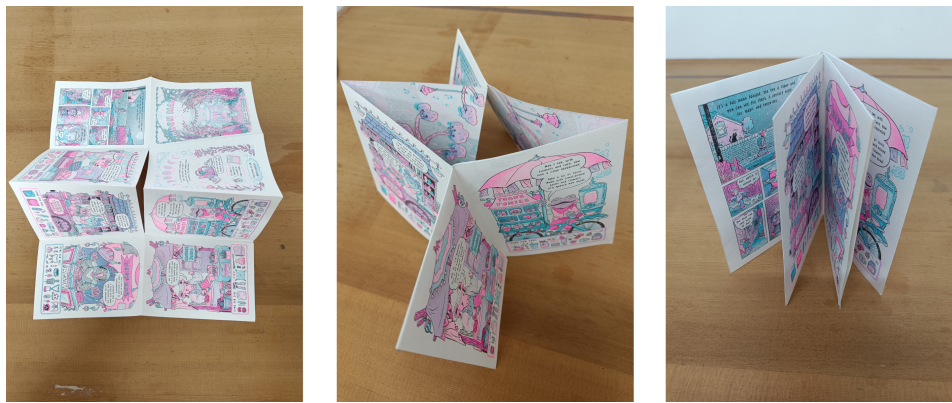
■ **Figure 16** Unlike the tunnel in *House of Leaves*, these are physical holes. Examples from the dust jacket of *Holes* [7] (left), and *Inscription: Journal of Material Text* (right).

5.3 Cuts and Physical Holes

Unlike the tunnel that the reader imagines just discussed, books and documents can contain actual holes in their structure as shown in Figure 16. In other examples, a slit in a sheet allows a document to be folded up and read as a sequence of pages in a way that would otherwise be impossible. An example is shown in Figure 17, where the document is both a flat A3 poster on one side when unfolded, and an eight page zine on the other side when folded up. This type of slit in a sheet also appears in practical geographical examples in the map foldings patented by Falk [25].

5.4 Towards a formal theory

Documents involve a multitude of entities such as pages, leaves, openings, fold-outs, etc which make up a space where other entities such as individual lines of text, blocks of text, page numbers, illustrations, etc can be located. The categories of located entities within documents can be extended to encompass holes, creases, tears, edges of pages, and even missing pages. All these entities have relationships between them, for example two pages may constitute a single leaf or be opposite each other when a book is open. It may be possible to simultaneously view a page and some matter that folds out. Two blocks of text



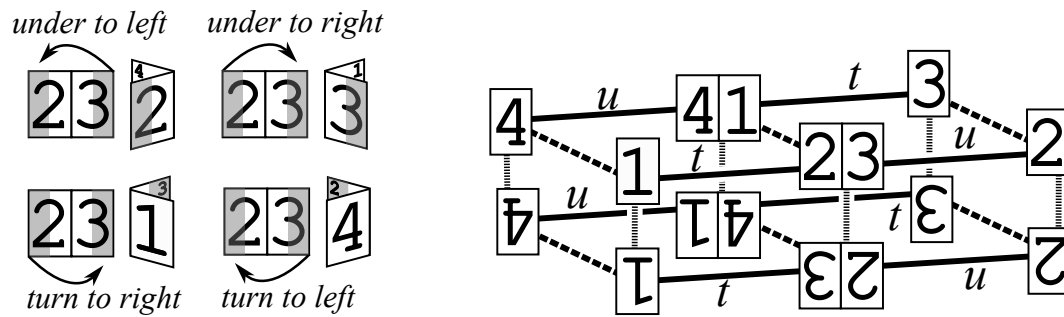
■ **Figure 17** A single sheet with a slit folded into an 8-page zine: *MOON MART*, Kristyna Baczynski. More elaborate foldings with multiple slits are used for maps, one example being the Falk maps.

on different pages might be in contact once a book is closed. A comprehensive ontology giving a systematic account of all possible entities and relationships in this area would be a challenging project well beyond the scope of this paper. To illustrate some of the issues and to suggest a possible form for a more general study we consider a material space made up of four pages with a single fold. Imagine a greetings card for example. See Figure 18 which is explained below

The challenge for recording relationships between entities located within any material document space is that, unlike many situations with qualitative relationships, the configuration of the space can change and also may need to be described with respect to an observer. These are two different issues. For example, a book may be open at a particular page and turning the page changes the configuration in a way that can affect what surfaces are in contact. On the other hand, in a single sheet of paper, a hole might be on the left or on the right depending on the observer if there is no convention about what left and right mean. Both changes to configuration and changes to observer position will affect qualitative relationships.

To explore a specific case, consider the 4-page document in Figure 18. To make the example concrete, we imagine the document lying flat on a table in front of the reader. If the document is open and showing pages 2 and 3, as on the left of the figure, there are four possible ways of turning a page that bring us to a new configuration. The conventional action of turning a page to move forwards in a book is labelled in the figure as “turn to left”. The turn in the opposite direction is labelled “turn to right”. In such a simple document we can also imagine picking it up and turning a page underneath the document so the edge of the turned page moves away from the viewer and not towards them.

In addition to these four kinds of turning, which change the configuration of the document, there are also geometrical transformations in which the document is manipulated as a whole. These all come down to flipping about one axis or another (the line of the fold or orthogonally to it). Rotating by half a turn is the result of flipping about one axis and then about the other. Figure 18 shows some of these transitions between 12 different configurations and positionings of the document. Only some transitions are shown for clarity. The vertical lines with very short dashes all show rotations through half a turn. The diagonal broken lines all show flipping the document about the fold. The solid lines show turns (labelled *t*) and under turns (labelled *u*).



■ **Figure 18** Space of a four-page document.

Now consider how qualitative spatial relationships are affected by some of the transitions between states shown in Figure 18. Look at the two grey regions shown occupying parts of pages 2 and 3 in the left of the figure. We have seen examples where the alignment of regions when a book is closed is significant. We have also seen in Figure 2 examples where paper is used that allows printing on one side to be read on the other side. It might thus be necessary to represent that the grey regions are actually in contact in some configurations, whereas in other ways of folding the document they can be “on top of each other” but not actually touching. When the document is unfolded and laid flat, they exhibit a third relationship.

The situation here appears different from regions which move and change relationships in examples such as two lakes becoming larger and merging together. Here, although the regions can move with the whole document, it is the space that they occupy that changes. To develop a formal theory that can accommodate this kind of change seems to need a way of representing how the relations in an appropriate calculus change along the transitions between different configurations. This suggests modelling the interaction between transitions and spatial relationships by having the space of configurations as a structure which acts (a term used here in its algebraic sense) on the algebra of relations; each transition being associated to a way of transforming the relations among themselves.

6 Conclusion

We have demonstrated that archival collections are very complex and have proposed that QSRs such as Allen Relations provide a logical framework for representing the relationships between the pages in documents, and the text which appears on them. A logical representation leads inevitably to the ability to reason over the documents based on the arrangement and layouts. The challenge is one of interpretation. While we can use qualitative representations to reason that one page follows the previous one, the examples we have presented require some level of knowledge to understand. We know intuitively if columns on one page share headings, and meaning, with those on the previous page. That the indentation of text is continuing from a previous page, and that indentation has an organisational purpose, implying parent child relations between blocks of text. The archivist knows about how paper was reused in times past and replies to letters were written at some angle – 45, 90, 180 degrees – to the original. Somehow we need a translation between the expert view of a document and the qualitative representation. Large language models certainly now provide an interface between human expression and computer interpretable language. Perhaps the framework and terminology of TEI is a starting point, which when combined with derived spatial relations, and a language model, could provide us with the tools to begin to work with large scale digitised collections such as we find in archives.

References

- 1 Ernest W. Adams. *Surfaces and Superposition*. CSLI Publications, Stanford, California, 2001.
- 2 J. F. Allen. Maintaining knowledge about temporal intervals. *Communications of the ACM*, 26:832–843, 1983. doi:10.1145/182.358434.
- 3 P. Atkin. *Basic Nest Architecture*. Seren, 2017.
- 4 Fia Backström. *COOP a-script*. Primary Information, Brooklyn, New York, 2016.
- 5 N. Barker. *Visible Voices*. Carcanet, 2016.
- 6 Mark Bell, Eirini Goudarouli, John Stell, and Jane Winters. *Qualitative Space in Digital Diplomats*, pages 175–196. Böhlau Verlag, Vienna, 2025. doi:10.7788/9783412530921.175.
- 7 Roberto Casati and Achille Varzi. *Holes and other superficialities*. MIT Press, Cambridge, Massachusetts, 1994.
- 8 Yufan Chen, Ruiping Liu, Junwei Zheng, Di Wen, Kunyu Peng, Jiaming Zhang, and Rainer Stiefelwagen. Graph-based document structure analysis. In *The Thirteenth International Conference on Learning Representations, ICLR 2025, Singapore, April 24–28, 2025*. OpenReview.net, 2025. URL: <https://openreview.net/forum?id=Fu0aggezN9>.
- 9 A. G. Cohn and J. Renz. Qualitative spatial representation and reasoning. In F. van Harmelen, V. Lifschitz, and B. Porter, editors, *Handbook of knowledge representation*, pages 551–596. Elsevier, 2008. doi:10.1016/S1574-6526(07)03013-1.
- 10 Suzanne Cotter, editor. *Julie Mehretu: A Transcore of the Radical Imaginatory*. Museum of Contemporary Art, Australia, Sydney, Australia, 2025.
- 11 Mark Z. Danielewski. *House of Leaves*. Pantheon Books, New York, the remastered full-color edition, 2000.
- 12 Johanna Drucker. *The Century of Artists' Books*. Granary books, New York City, 1995.
- 13 Dennis Duncan and Adam Smyth, editors. *Book Parts*. Oxford University Press, 2019.
- 14 Michael Friedman. *A History of Folding in Mathematics. Mathematizing the Margins*, volume 59 of *Science Networks. Historical Studies*. Birkhäuser, 2018.
- 15 P. Gaskell. *A new introduction to bibliography*. Oxford University Press, 1972.
- 16 S. Gross. The word turned image: Reading pattern poems. *Poetics Today*, 18:15–32, 1997.
- 17 J. Hollander. *Vision and Resonance. Two senses of poetic form*. Oxford University Press, 1975.
- 18 Jacques Justin. Aspects mathématiques du pliage de papier. In H. Huzita, editor, *Proceedings of the First International Meeting of Origami Science and Technology, Ferrara*, pages 263–277, 1989.
- 19 Sotirios Kastanas, Shaomu Tan, and Yi He. Document AI: A Comparative Study of Transformer-Based, Graph-Based Models, and Convolutional Neural Networks For Document Layout Analysis. *arXiv e-prints*, page arXiv:2308.15517, August 2023. doi:10.48550/arXiv.2308.15517.
- 20 Geewook Kim, Teakgyu Hong, Moonbin Yim, JeongYeon Nam, Jinyoung Park, Jinyeong Yim, Wonseok Hwang, Sangdoo Yun, Dongyoon Han, and Seunghyun Park. Ocr-free document understanding transformer. In *Computer Vision – ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part XXVIII*, pages 498–517, Berlin, Heidelberg, 2022. Springer-Verlag. doi:10.1007/978-3-031-19815-1_29.
- 21 T. Y. Kong and A. Rosenfeld. Digital topology: a comparison of the graph based and topological approaches. In G. M. Reed, A. W. Roscoe, and R. F. Watcher, editors, *Topology and Category Theory in Computer Science*. Clarendon Press, Oxford, 1991.
- 22 Francesco Lombardi and Simone Marinai. Deep learning for historical document analysis and recognition—a survey. *Journal of Imaging*, 6(10), 2020. doi:10.3390/jimaging6100110.
- 23 Shane McCausland. *The Art of the Chinese Picture-Scroll*. Reaktion Books, London, 2023.
- 24 Franz-Benjamin Mocnik and David Fairbairn. Maps telling stories? *The Cartographic Journal*, 55(1):36–57, 2018.

- 25 Mark Monmonier. Folding, unfolding. In *Patents and Cartographic Inventions: A New Perspective for Map History*, pages 105–135. Springer International Publishing, Cham, 2017. doi:10.1007/978-3-319-51040-8_4.
- 26 Jochen Renz and Debasis Mitra. Qualitative direction calculi with arbitrary granularity. In Chengqi Zhang, Hans W. Guesgen, and Wai-Kiang Yeap, editors, *PRICAI 2004: Trends in Artificial Intelligence*, pages 65–74, Berlin, Heidelberg, 2004. Springer Berlin Heidelberg. doi:10.1007/978-3-540-28633-2_9.
- 27 Jochen Renz, Reinhold Rauh, and Markus Knauff. Towards cognitive adequacy of topological spatial relations. In Christian Freksa, Wilfried Brauer, Christopher Habel, and Karl Friedrich Wender, editors, *Spatial Cognition II, Integrating Abstract Theories, Empirical Studies, Formal Methods, and Practical Applications*, Lecture Notes in Computer Science, pages 184–197. Springer, 2000. doi:10.1007/3-540-45460-8_14.
- 28 Barry Smith. Mereotopology: A theory of parts and boundaries. *Data and Knowledge Engineering*, 20(3):287–303, 1996. doi:10.1016/S0169-023X(96)00015-8.
- 29 J. G. Stell. Qualitative spatial representation in the humanities. *International Journal of Humanities and Arts Computing*, 13(1–2):2–27, 2019.
- 30 Jilin Wang, Michael Krumdick, Baojia Tong, Hamima Halim, Maxim Sokolov, Vadym Barda, Delphine Vendryes, and Chris Tanner. A graphical approach to document layout analysis. In Gernot A. Fink, Rajiv Jain, Koichi Kise, and Richard Zanibbi, editors, *Document Analysis and Recognition - ICDAR 2023*, pages 53–69, Cham, 2023. Springer Nature Switzerland. doi:10.1007/978-3-031-41734-4_4.