Article:

Stell, J (2019) Qualitative spatial representation for the humanities. International Journal of Humanities and Arts Computing, 13 (1-2). pp. 2-27. ISSN 1753-8548

https://doi.org/10.3366/ijhac.2019.0228

to appear in Int. J. Humanities and Arts Computing

Qualitative Spatial Representation for the Humanities

John G. Stell   j.g.stell@leeds.ac.uk

School of Computing, University of Leeds, Leeds, UK

Abstract

'Qualitative spatial reasoning and representation' is a range of techniques developed in Artificial Intelligence to meet the need for a computational treatment of qualitative spatial relations. Examples of such relations include 'next to', 'overlapping', 'to the left of', 'separate from', 'including', and so on. These relations occur within the data found in the spatial humanities, but the computational techniques described here do not appear to have been used in connection with this context. While Geographical Information Systems (GIS) are widely used as a means of visualizing and exploring material in the spatial humanities, GIS technology is acknowledged to be ill-suited to information that is vague, uncertain, ambiguous, imprecise or having other qualities that in a scientific setting could be regarded as imperfections. In the humanities such 'imperfections' are of course important, and qualitative spatial relations are one source of data that challenges scientifically based GIS. This article reviews the origin of qualitative spatial reasoning and representation in A. N. Whitehead's mereotopology and argues for exploring how these methods could complement GIS as a computational technique in the humanities. Qualitative representation is applicable to modelling spatial arrangements in many domains, not just geographical space. This is demonstrated through an example of spatial relations in lines of printed text.

Keywords Spatial models; Qualitative spatial relations; Qualitative Spatial Reasoning and Representation; Mereotopology; Computational Topology; Spatial Humanities

1 GIS in Digital Humanities

The digital humanities use computational representations of data as a lens through which to consider human experience. The specifically spatial aspects of this process have become known as the 'spatial humanities'[1]. The development of spatial humanities has been driven partly by the availability of technologies that allow the processing of spatial data in the humanities on a scale that would simply not be possible by hand. In particular, the spatial humanities makes extensive use of Geographical Information Systems (GIS). The tools provided by GIS have proved valuable in the spatial humanities for exploring data, but there is a widely acknowledged mismatch between the capabilities of GIS and the processes of scholarly research in the humanities.

For example, Kemp[2] notes that ". . . many of the basic assumptions on which the technology was designed do not play well with the methods and information used in the Humanities." In a similar vein, Harris, Rouse, and Bergeron[3] point out that "The use of GIS as a lens to understand the geographical dimensions of the humanities raises questions about the biases, assumptions, and the silences in the technology that impinge upon the exploration of the spatial turn". Nowviskie[4] also acknowledges the contribution of digital geospatial technology in the humanities while discussing the challenges that its adoption presents to scholars.

GIS is by no means the only way of handling spatial information computationally but it does provide a valuable technology for spatial information that can be given a precise location. This might be the ability to 'pin-point' something on a map, or to describe a portion of space by drawing a hard boundary. In the humanities, much data is not of this form and cannot readily be fitted into the model underlying GIS. Of course not only the humanities has to deal with locations that are imprecise, ambiguous, unknown, or fuzzy. A description of such kinds of 'imperfection' is provided by Duckham, Mason, Stell, and Worboys[5] who are concerned with integrating diverse sources of geographic information.

The value and limitations of GIS in the spatial humanities are not in doubt. What is not clear is what other technologies might be used, either in conjunction with or as an alternative to, GIS. Kemp[6] identifies the difficulties with using GIS as a creative opportunity, a misalignment which "is the point at which insight and learning can take place, making the application of spatial reasoning and geographic technologies a new frontier in the humanities." To take advantage of this opportunity it is essential to be aware, as Harris, Corrigan, and Bodenhamer[7] put it, that ". . . there are more methods and approaches available to scholars to explore the spatial humanities than the very heavy emphasis on off-the-shelf software packages provided by GIS vendors". Examples of such methods and approaches identified by these authors include geovisualization and the 'Geospatial Web'. In another article, the Geospatial semantic web is placed in a central place for the humanities and is claimed potentially to form "the core of what constitutes a humanities GIS"[8].

Developing a vision for "Semantic space rather than geographical space" Harris, Corrigan, and Bodenhamer[9] highlight the role of text:

> "Text map transformations can reflect both absolute space based on Euclidean

> coordinate systems as well as the relative space of textual association based on place names, spatial rules, and geospatial markers that extract spatial relationships embedded in text and go beyond the strict cartographic map making that dominates current humanities use of GIS."

The mention of "spatial relationships embedded in text" is significant here. A river running through a town, a park on the border of a neighbourhood, an event taking place across a city, a church next to a cinema, are examples of spatial relationships which might appear in historical accounts, in literature, in the raw material or in an analysed form in sociology, anthropology, and so on. The direct representation of such relationships in GIS is not straightforward – if the location of the church and the cinema are known then GIS can readily tell us that they are next to each other, but if all that is known is this adjacency but not where either is located, we cannot place them on a map. The text can be added as text, but how is its semantic content represented?

One of the methods and approaches of potential value to humanities scholars beyond the 'off-the-shelf software packages' noted as limiting by Harris, Corrigan, and Bodenhamer[10] is the approach known in artificial intelligence as 'Qualitative Spatial Representation and Reasoning'[11]. The digital humanities is moving beyond using ready-made software and shows increasing use of coding as evidenced by resources such as The Programming Historian[12]. However, Qualitative Spatial Representation and Reasoning still seems unexplored. This approach is variously abbreviated as QSRR, QSR, and sometimes QSTR to include temporal as well as spatial representation. For simplicity I will use QSR throughout this article to include all these meanings. QSR has two significant features. One is that it deals with spatial relationships that are qualitative such as the examples above. The term 'qualitative' here distinguishes it from quantitative spatial models that translate space into numerical coordinates. However, 'qualitative' is not opposed to computational: the second significant feature of QSR is that it provides a qualitative space that can be represented digitally. QSR allows computation, not with numbers representing points, but with logical statements representing qualitative relationships. These qualitative relationships need to be explored in the next section before moving on to the spatial aspects of this view of data.

This article aims to place QSR within the context of the spatial humanities. In Section 2 a motivating example from a small portion of text is used to discuss the notion of binary relations and the idea of deducing additional information from given relationships is introduced. The following section considers the origins of QSR. This is important to understand because it is the connection with the space of human experience, as opposed to the idealized mathematical space of Cartesian coordinate geometry, that is one of the reasons for concluding that an evaluation of QSR as a potential tool for the spatial humanities is long overdue. Section 4 provides an overview of computationally inspired work in using qualitative relations to describe space. This looks at one particular approach, the Region-Connection Calculus (RCC), in some detail, as well as noting the existence of alternative ways of dealing with qualitative descriptions evidenced by intersection models. The well-known qualitative description of time given by the Allen relations is described in Section 5. This prepares the reader for Section 6 containing a small practical example of using qualitative relations to describe shape in spatially arranged text, such as poetry. The final section discusses some challenges and apparent disadvantages to the use of QSR.

2 Relationships

Qualitative spatial representation depends heavily on the use of relationships which found in common human experiences of space and which are modelled mathematically by relations. The term "relation" has a specific technical meaning in mathematics and parts of computer science. Givant[13] starts a two-volume treatise on the mathematical theory of relations thus: "Binary relations are a mathematical way of talking about relationships that exist between pairs of objects". The word "binary" here is because these particular relations model relationships that can occur between two things. To clarify the notion of a relation, consider the following extract from Wordsworth's *Guide to the Lakes*[14].

> At Dalemain, about three miles from Penrith, a Stream is crossed, called Dacre, which, rising in the moorish country about Penruddock, flows down a soft sequestered Valley, passing by the ancient mansions of Hutton John and Dacre Castle. The former is pleasantly situated, though of a character somewhat gloomy and monastic; and from some of the fields near Dalemain, Dacre Castle, backed by the jagged summit of Saddleback, and with the Valley and Stream in front of it, forms a grand picture. There is no other stream that conducts us to any glen or valley worthy of being mentioned, till you reach the one which leads you up to Airey Force, and then into Matterdale, before spoken of.

This extract mentions individual things by name, such as Penrith, Airey Force, and Dacre Castle. Several geographic features such as valleys, fields, and streams are mentioned as well as less tangible entities such as the "grand picture" and there is an implicit reference to the route taken, in “a Stream is crossed”. These particular individuals are stated to have certain relationships between them. Dalemain and Penrith have the relationship of being about three miles from each other. In another case the relationship of the route to the stream is that the former crosses the latter. A relationship between more than three things is given when, from the location of certain fields (thing 1), Dacre Castle (thing 2), the summit of Saddleback (thing 3), and the Valley and Stream (things 4 and 5) together form the grand picture mentioned (thing 6).

Although relationships between more than two things are commonplace, the process of building models does not actually need anything beyond the binary ones. To explain this consider the example of the spatial relationship of between. This involves three things and the text above describes Dacre Castle implicitly as being between “some of the fields near Dalemain” and the “summit of Saddleback”. There is no way to express this just in terms of binary relationships involving only the three individuals. However, by imagining a fourth more abstract entitity, namely the alignment or ‘between-ness’ of the three particular individuals, we can do so. Writing F, C, and S as abbreviations for the particular fields, for Dacre Castle, and for the summit of Saddleback respectively, the relationships can be modelled as in Figure 1.

C *is between* F *and* S

F *is first in alignment* FCS

C *is second in alignment* FCS

S *is third in alignment* FCS

Figure 1:

The single 3-way, or ternary, relationship between C, F, and S can be expressed as three 2-way, or binary, relationships involving C, F, and S, plus a fourth entity, called here AlignmentFCS.

This example can also illustrate the notion of reasoning about qualitative relationships. The text itself does not actually contain the word “between”. The text provides the information that from F the view shows C “backed by” S. From this we can deduce that C is between F and S. We can also deduce that someone on the “jagged summit of Saddleback” and looking in the direction of the fields would have been able to see Dacre Castle. We can also deduce that if we were at Dacre Castle and looking in the direction of Saddleback, the fields would be behind us.

These deductions are based on the semantics of the spatial relationships that are described in the text as well as the semantics of spatial relationships in the conclusions of the deductions. In linguistics and in natural language processing the semantics of spatial language has already been widely studied[15]. Spatial prepositions have provided with various semantics, and it can be challenging to account for the variety of meanings of a given preposition. For example, the English preposition "in", can signify several very different spatial situations depending on the context including the particular entities involved in the relationship. Zwarts[16] shows that some approaches to spatial semantics in linguistics make use of qualitative spatial representation. Spatial semantics on its own does not provide a means of making deductions; it provides a model of utterances in language but to go beyond this and to be able to relate one statement to another requires techniques from QSR enabling deduction.

In this particular example, it might be objected that the three entities can be located on a map and that all is needed to make deductions is to see what statements are true of the particular configuration by inspecting the geometry. There are, however, many situations in which spatial relationships are given in documents but it is not possible to locate the entities in a space other than that created by the relations themselves. Thus reasoning in the way illustrated above provides a capability beyond what might be achieved by plotting locations on a map.

Other more complex deductions from the relations might be made. Given that Dalemain is near to the fields, in addition to the three-way relationship involving the fields, the castle and the summit, we might deduce that standing at Dacre Castle and looking towards Dalemain we would have Saddleback roughly behind us. This is less straightforward, but still within the scope of qualitative representation. Modelling "near to" can involve a logic with three truth values[17].

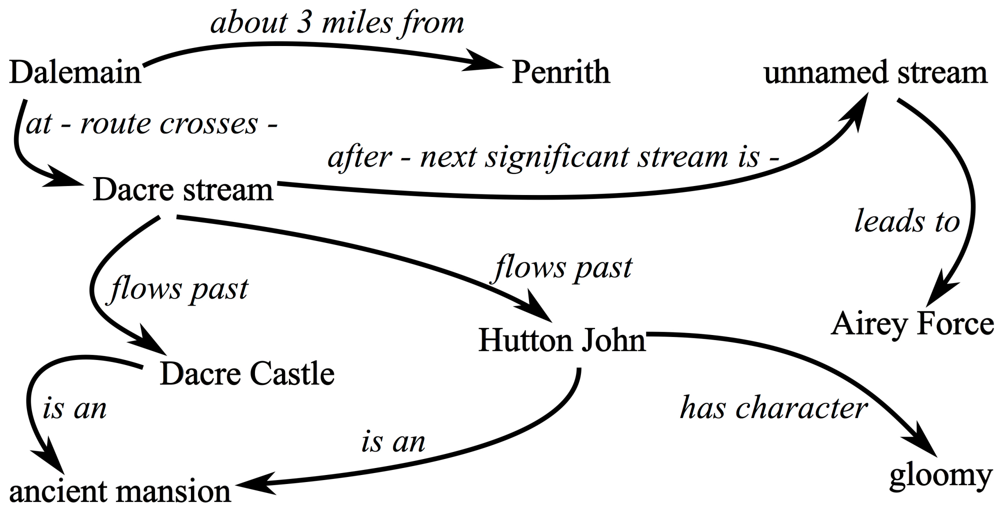

Figure 2:

Binary relationships in an extract from Wordsworth's *Guide to the lakes,* (1810).

One great advantage of binary relationships is the way they can be visualized in a network with the relationship indicated by a labelled arrow between the two participants. This leads to informal diagrams such as Figure 2, in which a number of different relationships appear between some things mentioned in the text. In general participants in relationships need not be specific individuals but can be types of entity, such as "ancient mansion", or qualities such as "gloomy". Diagrams such as the above have a long history[18], and there are numerous variants in which the syntax of the diagram needs to follow precise rules.

Instead of relationships between things in space, a radical alternative is that relationships might constitute space itself. For example, Smolin[19] advocates seeing

> "...the world around us as nothing but a network of evolving relationships. These relationships are not among things situated in space – they are among the events that make up the history of the world. The relationships define the space, not the other way round."

Smolin is proposing a space for theoretical physics, but the relevance of this view to radically novel conceptions of GIS has already been highlighted by Stell and Webster[20]. This kind of relational space might be better suited to the spatial humanities than conventional GIS, and it seems likely that the logical encoding of qualitative spatial relations would be one way to represent such a space digitally.

3 Varieties of Space

The limitations of GIS in the humanities are acknowledged, but is this inevitable with any mathematically based and computationally implemented representation of space? Giordano, Knowles, and Cole[21] take the view that

> "GIScientists can be unaware of the complexity of translating historical sources into the essentially mathematical, binary language of GIS, ... Humanists who view historical GIS as mere technical method may not realize that the disciplines of geography and GIScience give the tools of GIS their meaning, and that learning a few analytical techniques without learning about their intellectual context is unlikely to produce work of depth and substance."

This suggests a dichotomy. On the one hand the 'mathematical, binary world of GIS'; on the other the complexity of historical sources. The issue appears to be that the complexity of the scholarly material in the humanities cannot be captured adequately by the binary language of GIS. While GIS provides much of value, for humanists and other disciplines, it gives an impoverished impression of the variety of conceptual tools for representing spatial complexity in a computational way based in mathematics.

3.1 The Space of GIS

The space underlying GIS is built out of points. A point can be a location of a single entity; two points can demarcate the two ends of a straight line; a curved line on the surface of the Earth is approximated by many short straight lines, each specified by its end-points; areas are represented by their boundaries, which are sequences of lines described by the points at which direction changes. This point-line-polygon model is powerful and its computational grasp depends on the ability to specify points as pairs of numbers in a co-ordinate system. This translation of Euclidean plane geometry into numbers is due to Descartes in the 17th Century.

The conception of space as constructed out of points is a recurring theme in mathematical models of space. In the mathematisation of the plane into the space of all possible coordinates, areas are conceived as being sets of points and lines are again infinite collections of points. In higher dimensions the geometrical model continues: volumes are sets of points which make up each volume. The computational power of these models comes from Descartes, but the idea of points as the ultimate constituents of space appears in Euclid. Mathematics has less rigid conceptions of space, for example in topology, but here again the conventional approach is to assume a set of points and then to place additional structure on these atomic entities.

The ubiquity of points might lead us to the conclusion that these are inescapable features of any mathematical model of space. However the point-based world of the GIS is merely one of the assumed bases of spatial description that is rarely questioned by users of this technology.

3.2 The Space of Experience

The importance of human experience in the humanities makes it remarkable that the dominant digital spatial model, the one underlying GIS, is based on something outside human experience: the mathematical point. The ideal nature of points is noted by Simons: “no one has ever perceived a point, . . . whereas people have perceived individuals of finite extent.”[22]

The potential mismatch between mathematical space and experience is raised clearly by Whitehead:

> “We now know many alternative sets of axioms from which geometry can be deduced by the strictest deductive reasoning. But these investigations concern geometry as an abstract science deduced from hypothetical premises. In this enquiry we are concerned with geometry as a physical science. How is space rooted in experience?”[23]

Whitehead was not alone in questioning the connection between mathematical space and the perception and experience of the world. Contemporary with Whitehead, and dealing with similar concerns, was Jean Nicod (1893–1924) whose thesis “Geometry in the Sensible World”[24] acknowledged the role of sense data. Russell in the Preface[25] to this work, notes that Whitehead was interested in attaining the completed mathematical system from empirical data, whereas Nicod started in the reverse direction from empirical data to obtain a theory of space. Nicod[26] considers how places arise through changes in external and kinaesthetic data. I will not analyse the details of Nicod’s work as it has not led to a computational form. The work of Whitehead, on the other hand has done just so as will be discussed below.

Whitehead, in seeking to build mathematical space from experience cannot start with points. Instead he considers [27] “the physical relationship termed ‘extensive connection'" which may hold between spatio-temporal events. However Whitehead[28] clearly considered spatial entities separately from time and stated that the method of extensive abstraction being the same in principle in each case. Thus although Whitehead refers to events in general, he is also describing a spatial theory.

3.3 Regions and Connection

Whitehead's spatial theory concerns what we will call regions or portions of space. These regions have no properties except those that the theory specifies later, and the reader should beware of thinking of regions in a geographical sense – regions here are simply portions of space. The kind of regions described might simply be two-dimensional, say some abstraction of the surface of the Earth. Such regions could model spaces as extensive as a whole continent, or as small as a two-dimensional footprint left on a pavement, or even as tiny as the shadow of a particular ant at a particular time. Regions in this sense need have no natural description such as these and could include the surface delimited by certain arbitrary coordinates. Neither should it be assumed that regions are one-piece, a region may consist of many pieces. There are several different possibilities for what these regions are. Three dimensional, two dimensional in the plane, two dimensional on the surface of a sphere, one dimensional sets of intervals (which could also model time).

Whitehead thus takes regions rather than points as primitive and the spatial content of the theory is provided by a relation of "extensive connection" between regions. He explains[29] that extension is converse of part:

> ". . . the 'part' is an event which is extended over by the other event which is the 'whole.' "

Connection between regions is imagined to include overlapping (as in London overlaps England, or the region consisting of countries where French is an official language overlaps the region consisting of countries where English has this status). Connection also includes abutting or touching at the borders (the issue of whether the border is included does not affect this), for example the regions of England and Scotland are connected in this sense.

Whitehead does have points in his theory, but as abstractions from sets of successively smaller regions called abstractive sets. These define 'abstractive elements' and there is the intention[30] "to get hold of the class of abstractive elements that are in some sense the points of space". While the reconstruction of points in geometry from regions derived from physical experience is an interesting exercise, it is the regions alone and their connection that we concentrate on here. It turns out that regions with their relation of extensive connection have led to a computational approach to space to be described in the next section. Some of the main features of this can be seen already in Whitehead's writings.

One of these is the way of categorizing different forms of spatial relationship and initially a purely mereological, or part-whole, classification[31]

> "If an event A extends over an event B, then B is 'part of' A and A is a 'whole' of which B is a part. ...any two events A and B may have any one of four relations to each other, namely (i) A may extend over B, or (ii) B may extend over A, or (iii) A and B may extend over some third event C, but neither over the other, or (iv) A and B may be entirely separate. These alternatives can obviously be illustrated by Euler's diagrams as they appear in logical textbooks."

The notion of *junction* for what we would call 'connection' appears[32]

> "Two events have junction when there is a third event of which both events are parts, and which is such that no part of it is separated from both of the two given events"

and later[33] we find overlapping, and also the relationship now known as 'external connection', which can be imagined as regions touching only at a boundary

> "The relations of whole and part and of overlapping are particular cases of the junction of events. But it is possible for events to have junction when they are separate from each other; for example the upper and lower part of the Great Pyramid are divided by some imaginary horizontal plane."

Whitehead and others in the 1920s are by no means the only ones to critique the conventional mathematical approaches to space from the viewpoint of physical experience. More recently, for example, Poston[34] writes of a different kind of space:

> "Unlike the Euclidean plane … fuzzy spaces occur as objects of direct experience"

For the spatial humanities, however, the significant thing about Whitehead's theory, which is now known as a form of mereotopology[35], is that it has led to a computational representation.

4 Computational Mereotopology

The philosophically motivated inquiries of the 1920s were taken up by Clarke[36] in the 1980s who proposed an encoding of the properties of connection in formal logic. This was developed by Randell and Cohn[37], and in the many subsequent articles by Cohn and others cited in the overview by Cohn and Renz[38], into a well-established logical account of space in terms of regions and qualitative relations between regions

Cohn and Renz[39] motivate QSR by the need for humans to interact with GIS in a more flexible way. This echoes the awareness from the spatial humanities that GIS as currently implemented provides a rigid tool that is not well suited to the ways humans need to handle spatial information. They observe that “present-day GISs do not sufficiently support intuitive or common-sense oriented human-computer interaction”. Cohn and Renz go on to speculate “Arguably, the next generation GIS will be built on concepts arising from Naive Geography, wherein QSR techniques are fundamental” where Naive Geography refers to the approach advocated by Egenhofer and Mark[40]. The widespread use of QSR across several fields includes examples[41] from: robot navigation, computer vision, natural language processing, engineering design, and the specification of visual programming languages. More recent work includes processing video data to obtain quantitative information (such as locations of items in each video frame) and then QSR techniques are used to abstract from this numerical data to detect qualitative patterns of behaviour[42]. In this way the distinction between different events, such as between two people eating a meal and two people playing a card game, can be detected.

4.1 RCC as a description of space

In order to understand how RCC might be employed in the spatial humanities it is necessary to examine some of the details. From one viewpoint it provides a theory of space itself. Here, instead of the points, lines, etc of Euclidean space, there are just regions. Any two regions are either connected or not. Properties of the Connection relation constrain what space may be, in the same way that Euclid’s axioms constrain space. As with Euclid, the theory says nothing about what the regions are, only what properties they have.

Intuitively, regions are connected when they overlap or at least touch at their boundaries. In the process of describing the constraints on space and in working with the consequences of the axioms it is convenient to extend the vocabulary beyond the primitive notion of connection. Out of connection come a host of useful qualitative relations including part, overlap, tangential part, non-tangential part, and others.

Before considering the RCC in more detail, it is useful to be clear about what it can provide. A statement that two particular regions are spatially related in some way is not in itself a novelty in the spatial humanities and can readily be found in various encodings. For example Eide[43] considers translating text to maps and does use RDF, which is a standard format for linked data using triples. In the case of named regions the three elements of the triple would be the two regions and the fact that they are connected. What QSR offers is much more than such individual facts. It allows reasoning with relations, such as connection, which allow new facts to be derived from given ones – QSR provides is a way of creating new knowledge and establishing patterns out of this data.

The logic of RCC specifies various properties of connection, which it is not necessary to repeat in detail here. More can be found in the original articles and in what can be seen as a rational reconstruction[44] separating the mereology, that is the part-whole theory[45], from the topology. Besides connection the theory additionally assumes that we can talk about the union of two regions and the complement of a region. To give some intuition, it is helpful to think of the union just as the region formed by taking all the contents of the two regions together. The complement is everything outside the region, although a more technical account would need to be careful here.

Some of the properties of connection can be stated readily and the following serve as examples here. In the formal theory these are encoded in first order logic.

- Every region is connected to itself
- Every region, except the universe, is connected to its complement
- Every region, except the universe, has a region to which it is not connected.

this last property is somewhat controversial – it has the consequence that space is, in mathematical terminology, 'dense', or can always be repeatedly subdivided.

4.2 RCC as a language of spatial relations

Instead of describing space itself, in the model of Euclid's deductive geometry, RCC can be used as a language in which spatial relations between entities in the world can be stated. This yields a variety of systems (that is logical systems rather than computer implemented systems) of which the RCC8 is the most well-known.

The RCC8 is a classification of eight ways that two specific regions may relate to each other. Four of these are illustrated in Figure 3. The remaining four are: that the regions are equal; that they are separate (disconnected); that the first region contains the second as a non-tangential proper part; that the first region contains the second as a tangential proper part. The idea of categorizing relationships between things by means of a fixed set of possibilities appears frequently in QSR and not just for the RCC. Another example is seen in Allen's classification of time intervals in the next section, adapted subsequently for relationships between lines of text in Section 6.

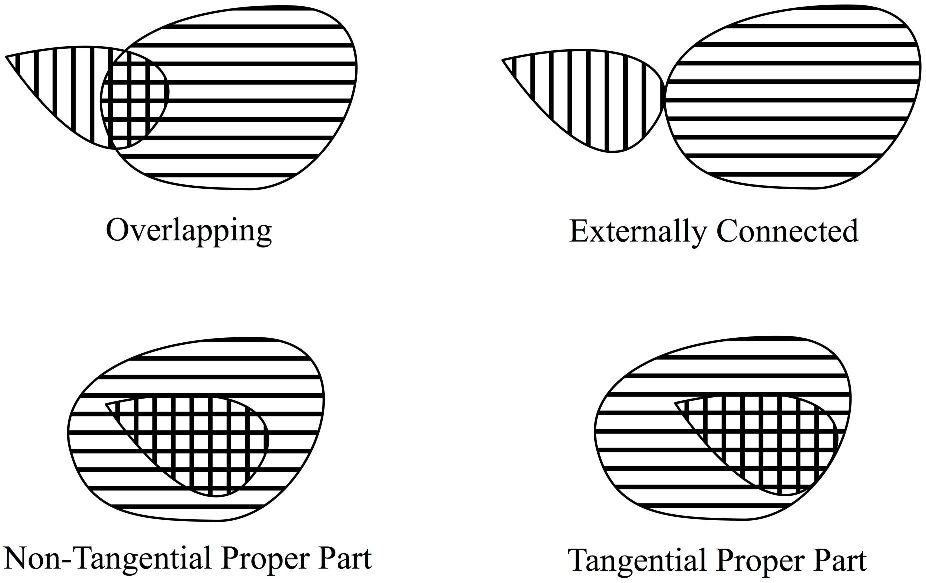

Figure 3:

Four ways a vertically shaded region can relate to a horizontally shaded one. These form four of the categories in the eightfold classification known as RCC8.

The illustrations in Figure 3 involve regions each consisting of a single part. However, an arbitrary region of the RCC may consist of many parts, leading to examples such as those shown in Figures 4 and 5. Some features of these more complex situations can be described using the logical apparatus of RCC. Others, such as the apparent sense in which parts of the two regions in Figure 5 enclose each other, may need extensions such as the idea of a convex hull operator which introduces some geometrical features into the topological theory.

All the relations in the RCC8 as well as many others can be reduced to statements about connection. Some of these statements are shown as examples in Table 1. For potential applications to the spatial humanities, the definitions in terms of connection would not be used directly by humanities scholars. They are included here because the ability to carry out deductions, such as those discussed above in Section 2, comes from the combination of these definitions with the axioms that specify the properties of connection in the RCC. These definitions and axioms, encoded in a formal logic, form the basis of the spatial semantics provided by the RCC.

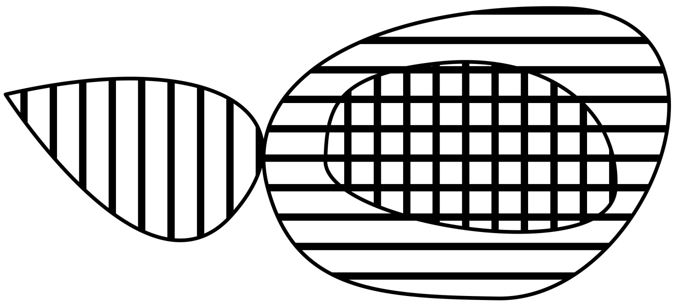

Figure 4:

A two-piece region (vertical shading) overlapping a one piece region

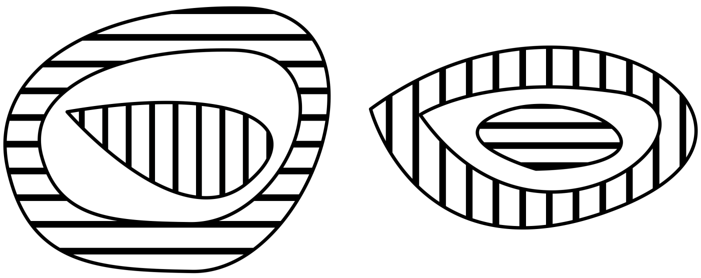

Figure 5:

Two two-piece regions which are disconnected. One is shaded vertically and one horizontally. Each region consists of two separate parts and one part of each region is enclosed by a part of the other region.

Table 1 Defining relationships in the Region-Connection Calculus from connection.

Relation of x to y	Definition in terms of connection and derivatives
part of	Every region connected to x is also connected to y
equal to	x is a part of y and y is a part of x
overlaps	There is a region which is part of x and also part of y.
partially overlaps	x overlaps y but neither x nor y is a part of the other
proper part	x is a part of y and y is not a part of x
externally connected	x is connected to y but they do not overlap
non-tangential proper part	x is a proper part of y and no region z is externally connected to both x and y
tangential proper part	x is a proper part of y and some region z is externally connected to both x and y

4.3 Intersection Models

The Region-Connection Calculus is only one aspect of QSR. Even for qualitative descriptions of regions of space in the plane there are alternatives to RCC. The work of Egenhofer and Franzosa[46], and their subsequent article[47], is entirely separate and different in conception from RCC, not being based in logic. However, it ends up with essentially the same eightfold classification as the RCC when considering certain two-dimensional regions in the plane. The fact that this qualitative classification arises in two different ways gives evidence of its value. The approach of Egenhofer and Franzosa can be described as 'intersection models' and, briefly, it categorizes spatial relationships between pairs of entities in terms of how aspects of each have or do not have non-empty intersections with each other. These aspects may, depending on the specific model, include the boundary, the interior, and the exterior. For example, with two externally connected regions in the plane, the interior of each intersects only the exterior of the other and not the boundary or the interior. By recording for each possible pair of intersections whether it is empty or not, a matrix of binary values is obtained. Pairs of regions yielding the same matrix are then seen as qualitatively the same.

The intersection models are cited by Yuan[48] who notes among techniques available from spatial analysis and modelling that "Topological relationships, such as overlaps and disjoint, are commonly used as the basis for spatial queries". The idea of using these matrices to determine what "qualitatively the same" should mean has been extended to relationships between many entities other than planar regions. Examples include directed line segments[49].

5 Qualitative Relations for Time and for Networks

Motivating examples for formal systems such as the RCC often use regions in geographical space, maybe planar regions, or three-dimensional volumes, or portions of the surface of the Earth. Qualitative relationships can be used to describe a much wider range of structures than the spaces these examples come from. Two specific cases are described in this section: relationships between intervals of time, and the type of space described by a network.

5.1 Qualitative Time

It is in the realm of time rather than space that mereotopological concepts are most readily found directly in the humanities. The work of Allen[50] provides a classification of ways in which intervals of time may relate qualitatively to each other. Drucker and Nowviskie[51] and Ryan[52] are among the authors in the humanities who cite Allen's work.

Two intervals of time may, for example, overlap, may be equal, may come one before the other, may be arranged just so that the end of one and the start of the other is their only point of coincidence, and so on. Drawn out on a line these time intervals appear as one-dimensional regions. The directional nature of time means that, for example, the qualitative relation of external connection in the RCC now comes in two distinct forms according as one region precedes or succeeds the other. In all, Allen dealt with thirteen qualitative relations between intervals as illustrated in Table 2.

Table 2 Relations between time intervals, redrawn from Allen's figure[53]

Relation	Example
before	`XXXX` `YYYY`
during	`XXXX` `YYYYYYYYYY`
meets	`XXXXX` `YYYYY`
overlaps	`XXXXXX` `YYYYYYY`

Relation	Example
equal	`XXXXXX` `YYYYYY`
finishes	`XXXX` `YYYYYYYYYY`
starts	`XXXX` `YYYYYYYYYY`

In Table 2 the examples use the convention that time flows from the past on the left to the future on the right. The time occupied by the first argument to the relation is shown by a row of `X`s on a line above a row of `Y`s in the next line indicating the second argument. For example, the relationship `X` finishes `Y` means that the interval `Y` starts before `X` but ends at the same time. Apart from the equal relation, each relation has an inverse which is different from the relation itself. This leads to 13 different qualitative relations between intervals.

Allen showed how reasoning could be performed with qualitative information by using a "transitivity table", also known as a composition table. The purpose of such a table is given the temporal relationship between a first and a second interval and between the second and a third, to list the possibilities for the relationships between the first and the third. Allen also discussed algorithms for processing networks of constraints. These consist of nodes, representing individual time intervals, and arcs representing the qualitative relationships between some pairs of these intervals. It is easy to construct networks of constraints between intervals that are impossible to realize, and Allen's article considers the detection of such inconsistent networks under certain conditions.

Time intervals can be seen as a one dimensional kind of space. In the next section this view is used in analyzing space in poetry. Here the lines have relationships that are described by Allen intervals.

5.2 Qualitative Space in Networks

Spatial concepts are often used as an organising tool for ideas and structures which are not spatial, or at least not spatial in the sense of geographical space. This is found in narratology where Ryan notes[54] “the representation of plot has inspired rhizomatic networks, trees, maps, Venn diagrams (the overlapping circles of set theory), iconic images, abstract shapes, . . . ”. Similarly, Mitchell[55] envisages using the spatial where “More specific spatial forms would display overlapping or intersecting patterns, some referring to principles of movement through the text, some governed by patterns of imagery or ideas that reflect authorial assumptions about world order”.

The computational representation of networks is often in terms of relations – that one node in the network may is related to another for some reason. Within a network of personal relationships for example, parts of the network may be related to other parts. One family may overlap another, two alliances may have no one in common, a group of individuals may be connected only to others within a specific group making them a non-tangential proper part. These kinds of relations between parts of networks can be treated in a similar way to the use of mereotopological systems for geographical space. This is achieved by taking the network to be a type of discrete space, as for instance in the work of Galton[56] and of Sindoni and Stell[57].

6 Qualitative Space in Text

In the digital humanities, the techniques of computational linguistics are used in the analysis of texts for projects involving both the subject matter and the language itself. The model of language underlying these techniques is that of a one-dimensional sequence of words. This means that the techniques are fundamentally challenged by written and printed language where the spatial layout is an important part of the meaning. This is particularly evident in much poetry where successive lines in a page may deliberately begin in different positions, or where gaps within lines are placed to manipulate whitespace in the page as a whole. In the more extreme forms of concrete poetry, words may not consists of letters arranged on a line but scattered in different orientations across the page. The history of poetry having a significant visual aspect is well-known[58]. Here, I will use some examples of the spatial form of poems to illustrate the use of qualitative representation, but there is no claim that these examples are an in depth evaluation of QSR within literary forms employing spatial means. The first example below is developed in sufficient detail to show that qualitative relations are capable of describing some spatial patterns in text.

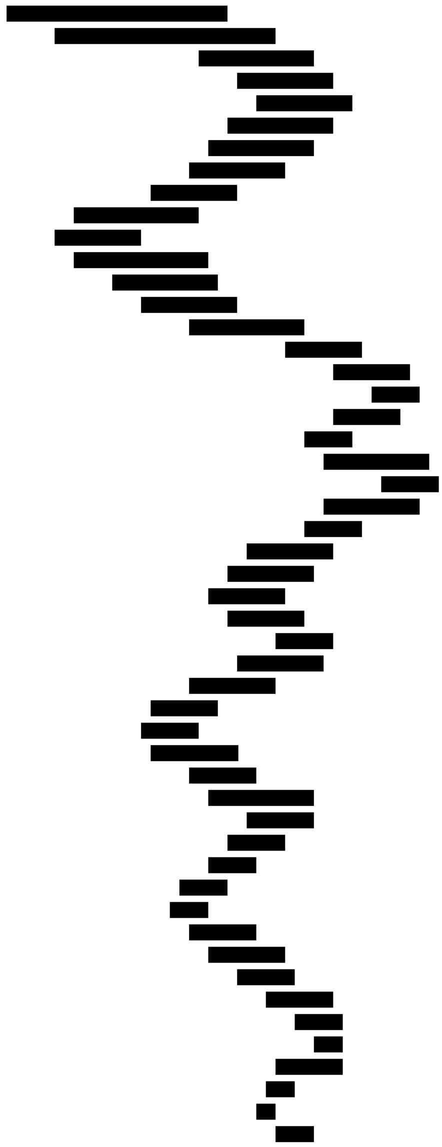

Figure 6:

Spatial arrangement of lines from Lewis Carroll's *Alice's Adventures in Wonderland,* (1866).

Figure 6 shows the way lines are arranged spatially in part of Lewis Carroll's *Alice's Adventures in Wonderland*[59] known as the mouse's tale. The image has been created simply to show the way the lines vary in horizontal position on the page. The original text includes other spatial features, especially the way that the text becomes smaller further down the page, which Figure 6 does not show. Figure 6 also ignores the boundary of the page, which can be highly significant in the appearance of the whole. Digital forms of this text can easily mimic some of the spatial structure. For example the Project Gutenberg version has a PDF of the text in which the mouse's tale is rendered in a fixed-width font using spaces to indent successive lines by differing amounts[60]. It is striking that different editions of this text present the mouse's tale in different ways. The variation between these is considerable which, even allowing for the constraints of typesetting such arrangements with the technology of the time, suggests that a key aspect of the shape is not its precise location. It seems rather that something "tale-like" or "wiggly" is needed and exactly where the lines are placed is not significant.

A description of the pattern of lines in Figure 6 in qualitative terms might be:

> *For several lines each line begins to the left of the start of the next one but ends before the next one ends. Then for several lines each line begins to the right of the start of the next one and ends after the end of the next one. There are several repetitions of this pattern of relationships: several where each line starts to the left of the next and then several where each line starts to the right of the next, with the endings as just described. Occasionally there are a few rare instances of other relationships: one line ending with the next but starting before it, or one line being all to the left of the next one.*

While it is necessary to clarify the way "*several*" is used, the description has as its core certain qualitative relationships between lines. These could be handled by using the Allen relations, not for time intervals but for lines, as shown in Figure 7 Conceptually these relations are just the Allen ones but with the wording rephrased. The two main relationships used in the above description are called in Figure 7 "*overlaps the start of"* and "*overlaps the end of".*

The qualitative description above of the relationships between lines in the mouse's tale could be represented computationally. An outline of what this would involve is as follows. The description contains the word "several" twice. Being more specific the first "several" could be specified as "normally around 5, but rarely only one or two" , whereas the second use would be consistent with "around 5". Although the original description is purely in words, and given the additional explanation of what "several" means, it is straightforward to determine whether a given sequence of relationships between successive lines fits this description or not. This would involve checking the only relationships in the sequence were *overlaps start of* and *overlaps end of* and a small number, say less than 10%, of the two relationships *all before* and *ends with but starts before.* Any sequence meeting those requirements could then modified to remove any relationships that were not overlaps of some kind. It would then need to be checked that the sequence had the pattern of several of one kind of overlap followed by several of the other and this alternation itself repeated several times. This

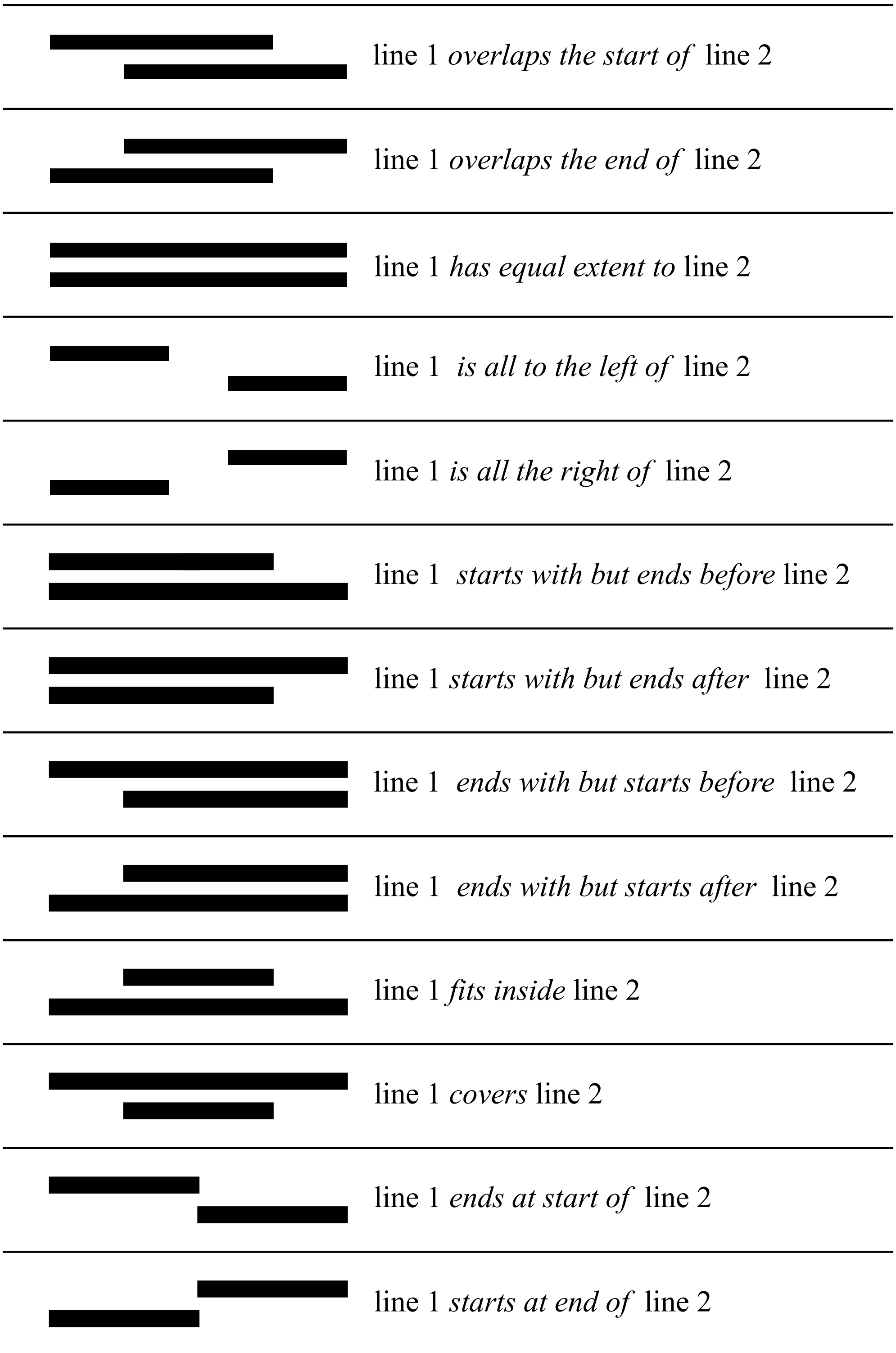

Figure 7: Allen’s thirteen qualitative relationships between time intervals expressed as relationships between lines.

is straightforward coding, taking for input a sequence of relationships and determining whether or not it fits the required pattern. The separate task of obtaining the sequence of relationships from an image of a page of printed text would also be well within the capabilities of current techniques; segmenting the image into lines and comparing coordinates of the start and end of each line with those of the next one. QSR works here as way of allowing the description of patterns in terms of relationships between lines. Although this depends critically for being useful on being able to obtain the relations from scanned text, this is a separate task and the numerical coordinates it involves can be isolated from describing patterns.

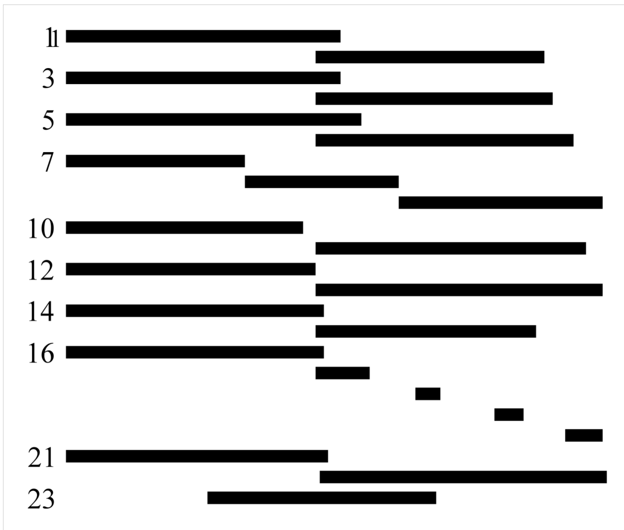

Figure 8:

Spatial arrangement of lines from Atkin's poem *Cannulation*.

The figurative and pictorial quality of the mosue tale is quite different from other examples of poetry. Consider the lines shown in Figure 8 which are derived from the layout of Atkin's poem *Cannulation*[61]. The spatial arrangement uses a number of evident devices. Lines 1,3,5 are left aligned and interspersed with line 2, 4, 6 which are also left aligned but positioned starting in the centre of the page. Similarly lines 10, 12, 14, 16 can be seen as a block of lines sharing a common left hand edge and these lines are interspered with 11, 13, 15 which are aligned similarly with 2, 4, 6. Lines 17—20 are evenly spaced and separate, descending to the right in the top to bottom reading direction. Lines 7—9 similarly span part of the page, the whole width in this case, and the three lines just overlap successively. Parts of the spatial structure are very clearly deliberate, especially the very short lines 17—20 which refer in the poem to both a process of dripping and of writing dots on letter 'i's which is echoed in the layout.

In *Cannulation* the relations are less easily described in terms only using the terminology of Figure 7. Key relationships are not just between one line and the next but between a line and the line next but one below it. The relationships are not only between pairs of individual lines but between groups of lines. For example the group of lines 2, 4, 6 relates to lines 11, 13, 15 in terms of both a similar horizontal placing and through relating to the groups of lines on their right, 1, 3, 5, and 10, 12, 14 respectively, in similar ways. This shows that working with qualitative relations in the humanities is not a matter of importing a ready-made method but will involve developing new sets of relationships depending on the task in hand.

7 Discussion

Drucker and Nowviskie[62] stress the formality and lack of ambiguity in the qualitative description of time. This raises perhaps the main obstacle to the exploration of the qualitative in a computational setting: the perception that the use of logic entails a framework of scientific rigidity that is unable to relate effectively to the ambiguous and nuanced aspects of the humanities. To some extent this appears to be a misapprehension. Ambiguity does in fact appear in Allen's work through the transitivity table described earlier. Allen's theory has no difficulty in recording ambiguity, simply by considering sets of relations rather than single relations. The relationship of one time to another may be unknown, but a set of various possibilities represents this uncertainty. Such a straightforward approach to ambiguity cannot, of course, be claimed to deal will all the many ways that this arises in the humanities. However, this example demonstrates that it may be too easy to reject the use of a formal approach to some aspects of qualitative relations.

Another way of handling uncertainly in QSR appears in an extension of RCC to deal with the case that the boundary of a region is uncertain, but can be roughly determined within some limits. The resulting "egg-yolk" regions[63] will only be capable of modelling some types of uncertainty. Nevertheless they show again that QSR can support a flexibility which needs to be explored within the context of the kinds of uncertainty and ambiguity encountered within the digital humanities.

When moving beyond the confines of classical Boolean logic we find that contradictions need not be problematic but can be used to represent important spatial concepts. The spatial notion of a boundary is a good example. Classically, the conjunction of the spatial relation 'inside' with its logical complement 'not inside' is a contradiction. Spatially there is nowhere that is both in a region and not in the same region. However, in an appropriate type of discrete space, that of a network, we find the natural underlying logic is not Boolean. There is a precise logical sense[64] in which the boundary of a region in this setting is those places that are both in the region and not in the region. Here, instead of an empty contradiction, we find the spatial concept of boundary.

The visual representation of qualitative spatial data is another challenge for work in this area. To draw two regions appears to offer a commitment as to their exact locations. However, if all we know is that they overlap we are already saying too much in any particular drawing. This tension between topological and geometric information is evident in the maps and diagrams of Moretti. In a response to Cerreti's review of his earlier book[65] Moretti writes[66] "The diagrams look like maps, yes, because they have been 'superimposed on a cartographic plane': but their true nature emerges unmistakably from the way I analyse them, which disregards the specificity of the various locations, to focus almost entirely on their mutual relations; which is indeed the way to read diagrams, but certainly not maps." The tension between interpreting spatial information presented cartographically in GIS and qualitative spatial information presented diagrammatically is considerable. Novel techniques for visualization are going to be needed here.

In conclusion, we have seen that Whitehead developed mereotopology from concerns about the space of experience as opposed to abstract science. This informal theory was later encoded in mathematical logic with a computational representation. By this means qualitative spatial representations can be handled computationally; given a database of facts about how regions are related to other regions, deduction can derive consequences which are implicit but not explicit in the data. This handling of qualitative relationships is not restricted to space, and Allen's intervals provide one qualitative approach to time. Neither is it the case that QSR can only handle regions that might be encoded as polygons in GIS. A specific example of QSR to describe spatial arrangement in a text has been given and there are many other opportunities for different spatial situations, including the discrete spaces of networks.

This article has provided sufficient evidence to motivate further explorations of the applications of QSR within the digital humanities. However, it needs to be stressed that such explorations will not be not simply a matter of using some existing software package; they need to involve the development of software in research projects involving interdisciplinary co-production of the necessary tools between humanities scholars and computer scientists.

Acknowledgements

This work was supported by a grant from The Arts and Humanities Research Council: AH/R006482/1, Space and Narrative in the Digital Humanities: A Research Network.

---

[1] D. J. Bodenhamer, J. Corrigan, and T. M. Harris, eds, *The spatial humanities. GIS and the future of humanities scholarship*, (Bloomington, 2010).

[2] K. K. Kemp, 'Geographic information science and spatial analysis for the humanities'. In Bodenhamer, Corrigan, and Harris, eds, *The spatial humanities.* 31–57.

[3] T. M. Harris, L. J. Rouse, and S. Bergeron, 'The geospatial semantic web, Pareteo GIS, and the humanities', in Bodenhamer, Corrigan and Harris, eds, *The spatial humanities,* 124–142.

[4] B. Nowviskie,. '"Inventing the Map" in the Digital Humanities (a Young Lady's Primer)'. *Poetess Archive Journal*, 2(1) (2010).

[5] M. Duckham, K. Mason, J. G. Stell, J. G and M. F. Worboys, 'A formal approach to imperfection in geographic information', *Computers, Environment and Urban Systems*, 25 (2001), 89–103.

[6] Kemp, 'Geographic information science and spatial analysis for the humanities'.

[7] T. M. Harris, J. Corrigan and D. J. Bodenhamer, 'Challenges for the spatial humanities: Toward a research agenda' in Bodenhamer, Corrigan, and Harris, eds, *The spatial humanities,* 167–176. Cited here at 171.

[8] Harris, Rouse, and Bergeron, 'The geospatial semantic web, Pareteo GIS, and the humanities', 127.

[9] Harris, Corrigan and Bodenhamer, 'Challenges for the spatial humanities', 172.

[10] Harris, Corrigan and Bodenhamer, 'Challenges for the spatial humanities'.

[11] A. G. Cohn and J. Renz, 'Qualitative spatial representation and reasoning', in F. van Harmelen, V. Lifschitz, and B. Porter, eds, *Handbook of knowledge representation*, (Amsterdam, 2008), 551–596.

[12] *The programming historian,* https://programminghistorian.org, last accessed 25 March 2019.

[13] S. Givant, *Relation Algebras, Volume 1,* (Berlin, 2017), xi.

[14] N. Mason, S. Stimpson and P. Westover, *Wordsworth's Guide to the Lakes. A Romantic Circles Electronic Edition.* (2015, original text 1810) https://www.rc.umd.edu/editions/guide_lakes/editions.2015.guide_lakes.1810.html, last accessed 12 Feb 2019.

[15] J. Zwarts, 'Spatial semantics: Modeling the meaning of prepositions', *Language and Linguistics Compass*, 11:e12241 (2017).

[16] Zwarts, 'Spatial semantics'.

[17] M. Duckham and M. F. Worboys, 'Computational structures in three-valued nearness relations', in D. Montello, ed., *Foundations of geographic information science international conference, COSIT 2001 Proceedings*, Lecture Notes in Computer Science 2205, (Berlin, 2001), 76–91.

[18] F. Lehmann, 'Semantic networks', *Computers & Mathematics with Applications,* 23 (1992), 1–50.

[19] L. Smolin, *Three roads to quantum gravity*, (New York, 2000), 96.

[20] J. G. Stell and J. Webster, 'Oriented matroids as a foundation for space in GIS', *Computers, Environment and Urban Systems*, 31 (2007), 379– 392.

[21] A. Giordano, A. K. Knowles and T. Cole, 'Geographies of the Holocaust', in A. K. Knowles, T. Cole and A. Giordano, eds, *Geographies of the Holocaust* (Bloomington, 2014), 1–17.

[22] P. Simons, *Parts. A study in ontology*, (Oxford, 1987), 42.

[23] A. N. Whitehead, *An enquiry concerning the principles of natural knowledge*, 2nd ed, (Cambridge, 1925), v.

[24] J. Nicod, *Foundations of geometry and induction*, (London, 1930).

[25] B. Russell, 'Preface', in Nicod, *Foundations of Geometry and Induction*, 5–9. Cited here at 7.

[26] Nicod, *Foundations of Geometry and Induction*, 112.

[27] A. N. Whitehead, *Process and reality: An essay in cosmology.* (Cambridge, 1929), 416.

[28] A. N. Whitehead, *The concept of nature,* (Cambridge, 1920), 79.

[29] Whitehead, *The concept of nature,* 58.

[30] Whitehead, *The concept of nature,* 85.

[31] Whitehead, *The concept of nature,* 75.

[32] Whitehead, *The concept of nature,* 76.

[33] Whitehead, *The concept of nature,* 77.

[34] T. Poston, *Fuzzy geometry,* (unpublished Ph.D. thesis, University of Warwick, 1971), iv.

[35] Simons, *Parts.*

[36] B. L. Clarke, 'A calculus of individuals based on 'connection'', *Notre Dame Journal of Formal Logic*, 22 (1981), 204–218.

[37] D. A. Randell and A. G. Cohn, 'Modelling topological and metrical properties in physical processes', in R. J. Brachman, H. J. Levesque, and R. Reiter, eds, *Proceedings of first international conference on principles of knowledge representation and reasoning,* (San Mateo, 1989), 357–368.

[38] Cohn and Renz, 'Qualitative spatial representation and reasoning'.

[39] Cohn and Renz, 'Qualitative spatial representation and reasoning', 553.

[40] M. Egenhofer and D. Mark, 'Naive geography', in A. Frank and W. Kuhn eds, *Spatial information theory: A theoretical basis for GIS*, Lecture Notes in Computer Science 988, (Berlin, 1995), 1–16.

[41] Cohn and Renz, 'Qualitative spatial representation and reasoning', 553–554.

[42] K. S. R. Dubba, A. G. Cohn, D. C. Hogg, M. Bhatt and F. Dylla, 'Learning relational event models from video', *Journal of Artificial Intelligence Research,* 53 (2015), 41–90.

[43] Ø. Eide, *Media boundaries and conceptual modelling: Between texts and maps,* (Basingstoke, 2015), 50.

[44] J. G. Stell, 'Boolean connection algebras: A new approach to the region-connection calculus', *Artificial Intelligence*, 122 (2000), 111–136.

[45] Simons, *Parts.*

[46] M. J. Egenhofer and R. Franzosa, 'Point-set topological spatial relations', *International Journal of Geographical Information Systems*, 5 (1991), 161–174.

[47] M. J. Egenhofer and R. Franzosa, On the equivalence of topological relations. *International Journal of Geographical Information Systems*, 9 (1995), 133–152.

[48] M. Yuan, 'Mapping text', in D. J. Bodenhamer, J. Corrigan, and T. M. Harris eds, *The spatial humanities,* 109–123.

[49] Y. Kurata and M. J. Egenhofer, 'The head-tail-body intersection for spatial relations between directed line segments', in M. Raubal et al., eds, *GIScience 2006, Proceedings,* Lecture Notes in Computer Science 4197, (Berlin, 2006), 269-286.

[50] J. F. Allen, 'Maintaining knowledge about temporal intervals', *Communications of the ACM*, 26 (1983), 832–843.

[51] J. Drucker and B. Nowviskie, 'Temporal modeling: Conceptualizations and visualizations of temporal relations for humanistic scholarship', presented at the joint international conference of The Association for Computers and the Humanities and The Association for Literary and Linguistic Computing, 2003, http://www3.iath.virginia.edu/time/reports/infodesign.doc last accessed 25 March 2019.

[52] M.-L. Ryan, 'Diagramming narrative', *Semiotica*, 165 (2007), 11–40.

[53] Allen, 'Maintaining knowledge about temporal intervals', figure 2.

[54] Ryan, 'Diagramming narrative', 18.

[55] W. J. T. Mitchell, 'Spatial form in literature: Toward a general theory', *Critical Inquiry*, 6(3) (1980), 539–567.

[56] A. Galton, 'The mereotopology of discrete space', in C. Freksa and D. Mark, eds, *Spatial information theory. Cognitive and computational foundations of geographic information science. International conference COSIT'99,* Lecture Notes in Computer Science 1661, (Berlin, 1999), 251–266.

[57] G. Sindoni and J. G. Stell, 'The logic of discrete qualitative relations', in E. Clementini, et al., eds, *13th International Conference on Spatial Information Theory. COSIT 2017,* Leibniz International Proceedings in Informatics 86, (2017), http://drops.dagstuhl.de/portals/extern/index.php?semnr=16045, article No. 11, last accessed 25 March 2019.

[58] Barker, N., *Visible Voices,* (Manchester, 2016).

[59] L. Carroll, *Alice's Adventures in Wonderland*, (London 1866), 37. From image at https://www.lib.umich.edu/online-exhibits/exhibits/show/-curiouser-and-curiouser----ex/the-mouse-s-tail-tale, last accessed 5 January 2019.

[60] Project Gutenberg, *Alice's Adventures in Wonderland by Lewis Carroll,* http://www.gutenberg.org/files/11/11-pdf.pdf, 12, last accessed 6 January 2019.

[61] P. Atkin, *Basic Nest Architecture*, (Bridgend, 2017), 56.

[62] Drucker and Nowviskie, 'Temporal modeling'.

[63] A. G. Cohn and N. M. Gotts, 'The 'egg-yolk' representation of regions with indeterminate boundaries', in P. A. Burrough and A. U. Frank, eds, *Geographical Objects with Indeterminate Boundaries*, (London, 1996), 171-187.

[64] F. W. Lawvere, 'Introduction', in F. W. Lawvere and S. H. Schanuel, eds, *Categories in continuum physics*, Lecture Notes in Mathematics 1174 (Berlin, 1986), 1–16.

[65] F. Moretti, *Atlas of the European novel 1800-1900*, (London, 1998).

[66] F. Moretti, *Graphs, maps, trees. Abstract models for a literary history*, (London, 2005), 54.