LEVEL OF DETAIL FOR GRAPHS: EQUIVALENCE RELATIONS AND PARTITIONS

John Stell

School of Computing, University of Leeds, Leeds, LS2 9JT, UK.

Corresponding Author's E-mail: j.g.stell@leeds.ac.uk

Abstract— Rough sets provide a theory of granularity, or level of detail, which has had numerous applications. This paper considers how the theory of rough sets may be extended to a theory of rough graphs. To do this it is necessary to develop notions of equivalence relation and of partition for graphs. This is accomplished by generalizing the account of rough sets within mathematical morphology based on the concept of graphical relations. The operation of converse on such relations is much less straightforward than for relations on sets, but is shown to provide a suitable notion of symmetry which can be used to define an analogue of equivalence relations for graphs. A correspondence between these equivalence relations and one notion of partition for graphs is established.

1 INTRODUCTION

1.1 Granularity in general

Granularity concerns the interaction between information and levels of detail. For example a passenger making a journey by air needs to know little about the route other than the start and end airports and the expected times of departure and arrival; a pilot on the same flight will clearly require information about the route containing much more detail. Events are often described at vastly different levels of detail for different purposes. For example, after a football match, simply knowing the final score is adequate for many people, others will want to know who scored and in which minute, while the team's coach or manager may make a minute analysis of sections of the game.

Such examples of granularity in everyday life already present many features important in the scientific study of the subject. Different tasks may require what is thought of as `the same information at different levels of detail', and much of the representation, processing and gathering of any kind of information is carried out within the context of one or more particular levels of detail.

The term granularity arises from the way in which a reduction of detail is often achieved by gathering together separate data items into indivisible granules. Granulation

is not understood solely as this clumping together process; data may also be omitted at a less detailed level. Thus granulating includes the selection of items to be gathered together as well as the gathering process itself.

Granularity can arise via intentional coarsening (reducing the available data to meet limitations of bandwidth or representational discrimination for example) or be forced upon us by limits of our measuring or sensing technologies. Deliberate changes to levels of detail are also of practical importance in integration. It may be necessary to reduce different data sources or data requests to a common level of detail in order to allow interoperation (see Li *et al*. 2005 for a road network example).

1.2 Some Theories of Granularity

Studies of information granularity in general include the work of Keet (2006) who identifies seven kinds of granularity and organizes these as the leaves of a taxonomic tree which classifies granularities. In Artificial Intelligence, Hobbs (1985) proposes a structure for reasoning under granularity in which granularity dependent theories are combined by `articulating theories' which mediate between them.

The theory of rough sets (Pawlak 1982, Orlowska 1998) can be seen as providing a particularly simple form of granularity, but nevertheless one which has found numerous practical applications. In rough set theory granulation is effected by imposing an equivalence relation on a set *X*, so that at the less detailed level we can no longer see the individual set elements but can only see the granules i.e. the equivalence classes. Every subset $A \subseteq X$ has two extreme approximations in terms of the cells in a partition. The upper approximation is the set of all cells having a non-empty intersection with A and the lower approximation is the set of all cells which are subsets of A. Abstract sets are inadequate for many modelling tasks, so it is natural to ask whether the rough set approach to granulation can be extended to richer structures, such as graphs and networks. Partition based granulations have been discussed for spatial data, (e.g. Bitner and Stell 1998) although strict partitions need modifying to allow overlapping boundaries. In (Stell 2007) it was shown how focussing on relations allows a framework in which mathematical morphology (in some sense a granular approach to image processing) includes rough sets as a special case, and in the present paper this work is developed further towards rough graphs.

1.3 Granularity for Graphs and Networks

Changes in cartographic representations in generalizing from large scale to small scale maps have been the subject of extensive research (Muller *et al*. 1995), but despite recent advances the process still cannot be entirely automated. There remain, however, significant problems in just considering networks, where the spatial aspect is topological rather than geometric. One example of network data for the UK is the Integrated Transport Network (ITN) provided by the (Ordinance Survey). This is the kind of data which motivates the theory presented here: given large networks how can we manipulate

versions of the same data at different granularities?
The reduction of level of detail in graphs is not a purely geographical problem. In many examples of very large networks in the visualization literature (Gansner *et al*. 2005, Kumar and Garland 2006), the main concern is a form of data mining or data discovery. This is a different kind of issue from many geo-networks where the structure and relative importance of elements the network is effectively known and the challenge is to present this structure to the end-user.

2 REQUIREMENTS FOR A THEORY OF ROUGH GRAPHS

2.1 The morphological account

We review briefly the correspondence between mathematical morphology and rough set theory from (Stell 2007). Given a relation *R* on a set *X*, for any subset $A \subseteq X$ we will use $A \oplus R$ to denote the dilation of *A* by *R*, that is the set $\{x \in X : \exists a \in \cdot aRx\}$. When *A* is a singleton, we will write $a \oplus R$ in place of $\{a\} \oplus R$. The erosion of *A* by *R* is denoted $R \ominus A$ and defined as $\{x \in X : \forall y \in X \cdot xRy \Rightarrow y \in A\}$.

Given a subset $A \subseteq X$ we can approximate *A* with respect to the relation *R* in two important ways:

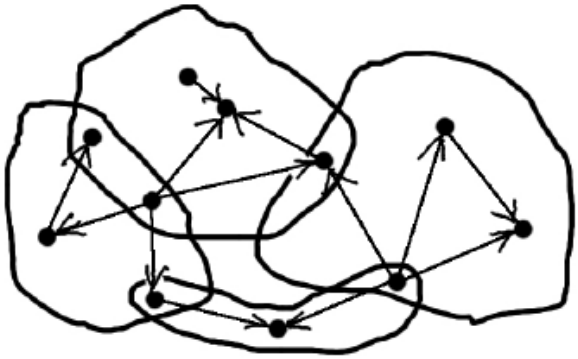
$$\textit{the closing: } R \ominus (A \oplus R) \subseteq X \,,$$

$$\textit{the opening: } (R \ominus A) \oplus R \subseteq X \,.$$

In the special case that *R* is an equivalence relation these are the upper and lower approximations from rough set theory. We can understand the opening as the union of the equivalence classes contained with *A*, and the closing as the union of the equivalence classes which intersect *A*.

2.2 Towards rough graphs

Partitioning a graph into disjoint parts is not appropriate for many applications. We might want to consider a road or railway network of nodes and links and partition it into pieces which are, for example, the responsibility are different maintenance companies. This suggests that partitions should allow cells to overlap in their boundary nodes as indicated in the following diagram.



However, there remain many options. For example, do we allow cells which only contain boundary nodes and do we permit edges in a cell which join only boundary nodes? Another possibility would be to restrict cells to containing connected subgraphs. There is unlikely to be a single correct way to proceed, and choices between these options will depend of specific applications. In order to construct a coherent theory, which can then be evaluated against applications, the following issues need to be tackled.

- What is a relation and an equivalence relation on a graph?
- What is a partition of a graph and is there a correspondence between partitions and equivalence relations?
- Can subgraphs be approximated by unions of cells in a partition in a way that relates to morphological operations on graphs?

Relations on graphs have been introduced in (Stell 2007), and the main points are included in Section 3 below along with some additional new results on these structures. To develop equivalence relations for graphs, analogues of transitivity and reflexivity are straightforward but symmetry depends on a notion of converse for graphical relations. Although a method for constructing the converse was identified in (Stell 2007) no explicit construction or properties of this operation were provided. These are presented in Section 4 and they are used in Section 5 to define equivalence relations for graphs and to establish a correspondence between these relations and one notion of partition for graphs.

3 RELATIONS IN GRAPHS

3.1 Graphs

The term `graph' will be used here to mean a directed graph where multiple edges and loops are permitted. These structures can be formalised in a number of ways. Commonly one has two sets, one of edges and one of vertices and two functions which assign to each

edge its source node and its target node. The approach taken here is different, and follows (Brown *et al*. 2006). This view of graphs differs from the usual one not in the graphs themselves, but in the morphisms between graphs. In particular a homomorphism is not required only to map edges to edges and nodes to nodes; we may have nodes mapped to edges and edges to nodes. The definition from (Brown *et al*. 2006) is as follows.

Definition 1. *A **graph** consists of a set G with source and target functions* $s,t:G\to G$ *such that for all* $g\in G$ *we have* $s(t(g))=t(g)$ *and* $t(s(f))=s(g)$.

The elements of *G* are the nodes and the edges of the graph, and the nodes can be distinguished as those elements *g* for which $s(g)=g=t(g)$. The subgraphs of *G* in this setting are those subsets of *G* which are closed under the source and target functions.

In modelling networks, there are two algebraic counterparts to the informal notion of the outside of a part of a network. The set of all subgraphs of a graph *G* has the structure of a bi-Heyting algebra. This can be thought of as analogous to the boolean algebra formed by the set of all subsets of a set. A key difference between subgraphs and subsets is that the former do not possess a Boolean complement. However they do have two weaker operations, one of which we will call the complement although strictly is only a pseudo-complement. The two operations arise from the existence of two kinds of elements in a subgraph: nodes and edges. The **complement** of $H,\neg H$, is obtained by taking the set-theoretic complement of the nodes of *H*, and including all edges between these nodes which exist in *G*. The **supplement** of $H,\sim H$, is obtained by taking the complement of the edges in *H*, and taking the nodes to be all the nodes in *G* which are endpoints of these edges.

3.2 RELATIONS ON GRAPHS

There are various possibilities for the notion of a relation a graph. Thinking in terms of sets *N* and *E* of nodes and edges, a natural concept would be two ordinary relations R_N and R_E on *N* and *E* subject to the constraint that $e_1 R_E e_2$ implies $m_1 R_N m_2$ and $n_1 R_N n_2$ where m_i and n_i are the source and target of e_i. This does not provide the appropriate notion on which to build a theory of rough graphs, as can be seen by considering the interaction between rough sets, and relations in mathematical morphology.

In mathematical morphology the significance of relations is due to the correspondence between relations on a set *X* and the join-preserving functions on the lattice $P(X)$ of subsets of *X*. A function $f:P(X)\to P(X)$ is called join-preserving when $f(A\cup B)=(fA)\cup(fB)$ and to such a function we can associate the relation *R* where

$xRy iff y \in fx$. It is well-known that relations on *X*, join-preserving functions on $P(X)$, functions $X \to P(X)$ and adjunctions on $P(X)$ are all equivalent concepts. For graphs, however, the above notion of relation fails to be equivalent to the notion of join-preserving function on the lattice of subgraphs of a graph. This can easily be seen by considering the simple example of a graph consisting of a single edge between two nodes.

The following notion of relation was introduced in (Stell 2007).

Definition 2. *A relation, R, on the set of elements of a graph G is* ***graphical*** *if for all* $g, h \in G$ *, it satisfies*

1. *if* $g R h$ *then* $gRs(h)$ *and* $gRt(h)$*, and*
2. *if* $s(g) R h$ *or* $t(g) R h$ *then* $g R h$.

We can use the erosion to give an alternative characterisation of graphical relations. The following lemma can be proved straightforwardly from the definitions.

Lemma 1. *Let* $G = (G, s, t)$ *be a graph and R a relation on the set G. Then the following conditions are equivalent:*

1. R is a graphical relation.

2. For every subset $A \subseteq G$:

(a) $s(A \oplus R) \subseteq A \oplus R$ *and* $t(A \oplus R) \subseteq A \oplus R$*, and*

(b) $(sA) \oplus R \subseteq A \oplus R$ *and* $(tA) \oplus R \subseteq A \oplus R$.

3. For every element $g \in G$ *: g 2 G:*

(a) $s(g \oplus R) \subseteq g \oplus R$ *and* $s(g \oplus R) \subseteq g \oplus R$*, and*

(b) $g \oplus R = \{sg, g, tg\} \oplus R$.

From the lemma it follows that for any arbitrary subset $A \subseteq G$ the dilation $A \oplus R$ by a graphical relation is not merely a subset but will be a subgraph of *G*. The same holds for the erosion.

Lemma 2. *For any graphical relation R on a graph G and any subset* $A \subseteq G$ *, the dilation* $R \ominus A$ *is a subgraph of G.*

Proof. Let $y \in R \ominus A$ and consider sy. If $syRz$ then yRz and so $z \in A$. Thus $sy \in R \ominus A$ and similarly $ty \in R \ominus A$. □

The proof of the theorem relating graphical relations to join-preserving functions has been outlined in [Ste07].

Theorem 3. *On any graph G the graphical relations form a monoid, with composition as for arbitrary relations and identity* $xIy \Leftrightarrow y \in \{s(x), x, t(x)\}$. *For any subgraph*

$H \leq G$ *and any graphical relation R we have a subgraph* $H \oplus R$ *and this assignment is a right action of the monoid of graphical relations on P(G). The assignment* $R \mapsto - \oplus R$ *is an isomorphism of the monoid of graphical relations with the monoid of join preserving functions on P(G) under function composition.*

4 Converse for Graphical Relations

4.1 Converse and negation

In a relation on a set the converse plays an essential role in the notion of equivalence relation. A relation is symmetric iff it is equal to its converse, and the converse if easily understood visually by reversing all the arrows in the relation. In the graph case, taking the converse of a graphical relation to be the converse of the underlying set-based relation on the elements will not work, as this does not in general produce a graphical relation. To see how the converse can be defined we need to recall the role of converse in mathematical morphology. The operations of dilation by the converse of *R*, written as $\oplus^*$, and of erosion by the converse, written $\ominus^*$ are given by

$$A \ominus^* R = \neg(\neg A \oplus R) \qquad R \oplus^* A = \neg(R \ominus \neg A) \ :$$

The significance of this is that the converse dilation and hence the converse of *R* is determined by the direct dilation and the boolean complement operation on the powerset *P*(*X*). The procedure is that from $\oplus$ we obtain the adjoint $\ominus$ and composing with the negation we obtain $\oplus^*$. This converse dilation operation determines the converse of the relation itself.

How then do we carry out this programme in the case of graphs? As has been pointed out in [Ste07], we can use the complement and supplement to define the converse dilation for graphs as

$$R \oplus^* A = \sim (R \ominus \neg A).$$

However, this formula does not provide an easy means of computing the converse on particular examples. The following section shows how to do this.

4.2 Explicit description of converse

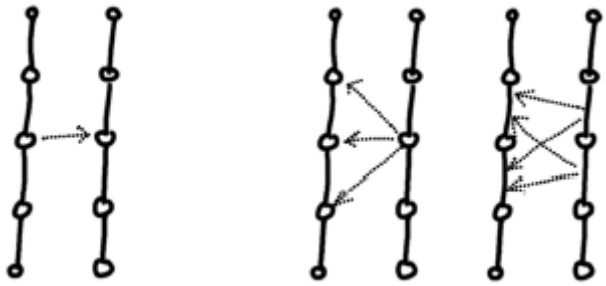
The next result allows us to construct the converse of a relation by breaking it down into smaller subrelations.

Lemma 4. *For graphical relations R and S on any graph G we have* $(R \cup S)^* = R^* \cup S^*$.

Proof. When *A* is a subgraph of *G*, $\sim((R \cup S) \oplus A) = \sim(R \ominus A) \cup \sim(S \ominus A)$. This

follows from $(R \cup S) \ominus A = (R \ominus A) \cap (S \ominus A)$ and $\sim (A \cap B) = \sim A \cup \sim B$. By putting $\neg A$ in place of *A* the result follows. □

Using this result we can use the definition of the converse to determine the converse of any relation in terms of the converses of the atomic relations out of which it is built. These atomic relation are those whose dilations map a single node or an edge with its two endpoints to a subgraph which is again of one of these forms. There are four cases to analyse, and just one of them is illustrated in the diagram below. On the left is a relation on a graph, in which just one node is related to another node. On the right is the converse, shown in two parts as drawing all the arrows on a single diagram obscures some of the detail. Taking the union of these two relations is the converse. Not all the arrows required in a graphical relation are drawn, as when an edge e_1 relates to and edge e_2 then e_1 relates also to the source and target of e_2. These additional arrows can easily be inferred from those given.



The results of the analysis lead to the following general description. For a graph *G* and *A* any set of elements of G, denote

$$A^+ = \{ g \in G \mid sg \in A or tg \in A \}.$$

If A consists of a single node then A^+ consists of that node together with any edges incident with the node. If *A* consists of a single edge then $A^+ = A$. We also denote by [*A*] the least subgraph of *G* which contains *A*, so

$$[A] = A \cup sA \cup tA.$$

We can specify a relation by saying for each element $a \in G$ to what subgraph *B* the element gets sent. Given this description for each $a \in G$, the converse will be the relation having all mappings of the form

$$\left\{ b \to \left[\{a\}^+ \right] \middle| b \in B^+ \right\}$$

Although the converse is not just reversing the arrows in the relation, there is a weaker

version of the fact that in a relation *R* on a set, *m R n* iff *n R* m*.

Lemma 5. *Let m, n be nodes of G and R a relation on G with converse R*.*

1. *If m R n then n R* m, and*

2. if m R n then there is some node k such that k R m where k is at most one edge away from n. That is either k = n or there is an edge of G between k and n.*

Conceptually, viewing the arrows in a relation as routes that may be followed, for any route in *R* you can always retrace your steps along *R**. However, in *R** you may not be able to retrace your steps along *R* but there will be somewhere near from where you can retrace along *R*, where the notion of nearness comes from the presence of edges in the graph.

The converse has substantially weaker properties than in the case of relations on sets, but it does have some interesting features such as the next lemma.

Lemma 6. *If I is the identity graphical relation on a graph G then dilation by the converse I* is the mapping* $A \mapsto \sim \neg A$, *and erosion by the converse is* $A \mapsto \neg \sim A$.

The significance of this is that $\sim \neg A$, called the **closure** of *A* consists of *A* together with all edges incident with *A* and all endpoints of such edges. Conceptually, this is *A* together with things near to *A*. The subgraph $\neg \sim A$ is called the **interior** of *A*, it corresponds to the elements of *A* not incident with any element outside *A*. The **boundary** of *A* is the set of nodes of *A* which are incident with edges outside *A*. This concept is used in defining partitions in the next section.

5 EQUIVALENCE RELATIONS

5.1 Properties of graphical relations

Having determined what should be meant by a relation on a graph from the mathematical morphology viewpoint we now consider what should be meant y an equivalence relation on a graph.

Definition 3. *A relation R on a graph G is*

reflexive *if* $I \subseteq R$,

symmetric *if* $R \subseteq R^*$,

transitive *if* $R;R \subseteq R$ *where ; denotes composition of relations.*

An equivalence relation on G is a relation which is reflexive, symmetric and transitive.

Definition 4. *A **partition** of a graph G is a set* $\prod$ *of subgraphs of G, called cells, such that*

1. *every node and every edge lies in at least one element of* $\prod$,
2. *any two distinct cells may intersect only in nodes which are in the boundaries of both the cells,*

3. *any node in the boundary of a cell K must be the endpoint of an edge of K.*

Note that we can have cells with no interior nodes; they can consist just of boundary nodes and edges between them.

Given an equivalence relation *R* on *G*, the analogues of the equivalence classes in the set case are the neighbourhoods $g \oplus R$ where *g* is an element of *G*.

Theorem 7. *Let R be an equivalence relation on G. Then the sets of nodes in the interiors of any two neighbourhoods are either equal or disjoint.*

Proof. Let *a* be a node in $(x \oplus R) \bigcap (y \oplus R)$. As $R \subseteq R^*$ we get $y R^* a$, and by the construction of the converse there is some *b* such that *b R y* and there is an edge between *a* and *b*. Since *a* is in the interior of $x \oplus R$, we must have *b* is in $x \oplus R$. So we have *x R b* and *b R y* which, as $R; R \subseteq R$, implies that $y \in x \oplus R$. But again using the symmetry of *R* we get that $y \oplus R \subseteq x \oplus R$. Similarly, $x \oplus R \subseteq y \oplus R$, so the intersection of the two neighbourhoods being non-empty implies that they are equal. □

This result shows that a graphical equivalence relation partitions a graph into cells which may only share nodes on their boundaries. The result is somewhat weak in that some nodes may not be in the interior of any neighbourhood as a neighbourhood can be purely boundary.

5.2 Relation from a partition

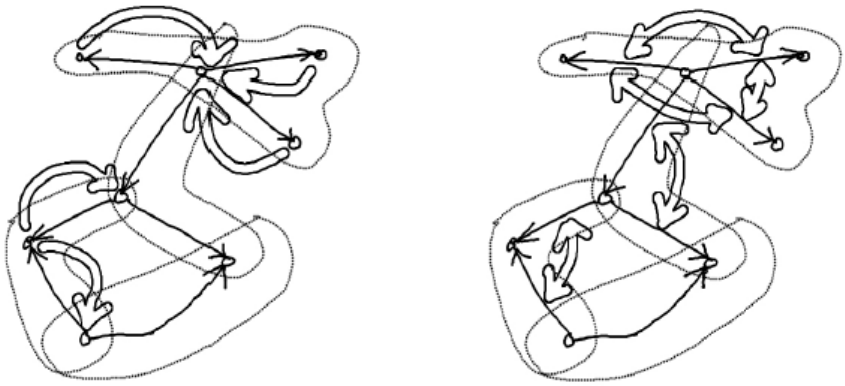
From a partition $\prod$, on a graph *G*, we can construct a relation which essentially induces the original partition by taking the cells to be the neighbourhoods.

For any cell *K*, let *B* be the set of boundary nodes of *K*, and let $I_B = \{(b,b) \in B \times B : b \in B\}$. Then the required relation is

$$R_\Pi = ((K - B) \times K) \bigcup I_B .$$

So every element which is not a boundary node is related to everything in *G*, and each boundary node is related only to itself.

The following diagram illustrates this. In the interests of clarity all the arrows in the relation from an element of the graph to itself are omitted and the node arrows and edge arrows are shown (as outline arrows) on two separate diagrams



Theorem 8. *The relation* R_{Π} *is a graphical equivalence relation.*

Proof. The relation is clearly graphical, and reflexivity is obvious. It is sufficient to show that the relation is symmetric and transitive when restricted to each cell. In the case of a cell which is purely boundary this is immediate, so we consider an arbitrary cell *K* which has at least one non-boundary element.

The transitivity is justified as follows. If *x R y R z* and $x \in K - B$ then $y \in K$ and $z \in K$, so *x R z*. If $x \in B$ x 2 B then we must have *x* = *y* = *z* so again *x z*.

The symmetry is obtained by analysis of cases. Suppose *x R y*; we have to show that *x R* y*. If $x, y \in K - B$ then we must have *y R x* and so *x R* y*. If $x, y \in B$ then we must have *x* = *y* so clearly *x R* y*. The only other possibility is that $x \in K - B$ and $y \in B$. In this case there will be an edge *e* having *y* as either source or target. We will assume without loss of generality that *y* = *s(e)*. Since *e R x*, we get *x R** s(*e*) as required. □

From this relation we can recapture the original partition by taking the maximal neighbourhoods, that is those not containing any smaller neighbourhood.

5.3 Upper and lower approximations

Given a graph *G* and a partition $\prod$, any subgraph $H \leq G$ has two extreme approximations in the usual rough-set style. The union of all the cells in $\prod$ which intersect *H* (the upper approximation), and the union of all the cells which are subsets of *H* (the lower approximation). These two can be understood morphologically for which the following lemma is helpful.

Lemma 9. *Let R be any relation on a set X which is reflexive and transitive. Then for any* $A \subseteq X$:

$$R \ominus (A \oplus R) = A \oplus R \text{ , and}$$

$$(R \ominus A) \oplus R = A \ominus R.$$

Using this it is easy to see that the upper approximation is the closing and the lower approximation is the opening, just as for rough sets.

6 CONCLUSIONS AND FURTHER WORK

The paper has shown that the morphological account of rough sets can be extended to rough graphs. To do this it has been necessary to understand the notion of converse for graphical relations in order to define symmetry for relations. Further work is needed to evaluate the various choices made in developing the theory by application to specific domains. Besides the idea of a graph as modelling a distribution or transport network, graphs can be used to model discrete spaces (Galton 1999). In such a context, it might be important to restrict partitions so that all the nodes in each cell are connected in the graph. It would be worthwhile examining whether additional conditions on a relation, beyond being an equivalence relation, could be found that correspond to such partitions.

REFERENCES

1. R. Brown, I. Morris, J. Shrimpton, and C. D. Wensley. Graphs of Morphisms of Graphs. Bangor Mathematics Preprint 06.04, Mathematics Department, University of Wales, Bangor, 2006.
2. T. Bittner and J. G. Stell. A boundary-sensitive approach to qualitative location. Annals of Mathematics and Arti_cial Intel- ligence, 24:93{114, 1998.
3. A. Galton. The mereotopology of discrete space. In C. Freksa and D. Mark, editors, Spatial Information Theory. Cognitive and Computational Foundations of Geographic Information Science. International Conference COSIT'99, volume 1661 of Lec- ture Notes in Computer Science, pages 251{266. Springer-Verlag, 1999.
4. E. R. Gansner, Y. Koren, and S. C. North. Topological _sheye views for visualizing large graphs. IEEE Trans. Vis. Comput. Graph., 11(4):457{468, 2005.
5. J. R. Hobbs. Granularity. In A. K. Joshi, editor, Ninth International Joint Conference on Arti_cial Intelligence, pages 432{435. Morgan Kaufmann, 1985.
6. C. M. Keet. A taxonomy of types of granularity. In 2006 IEEE International Conference on Granular Computing, pages 106{111. IEEE Press, 2006.
7. G. Kumar and M. Garland. Visual exploration of complex timevarying graphs. IEEE Transactions on Visualization and Computer Graphics, 12(5):805{812, 2006.
8. Y. Li, D. Blostein, and P. Abolmaesumi. Asymmetric inexact matching of spatially-attributed graphs. In L. Brun and M. Vento, editors, Graph-Based Representations in Pattern Recognition, volume 3434 of Lecture Notes in Computer Science, pages 253{262. Springer-Verlag, 2005.
9. J. C. Muller, J. P. Lagrange, and R. Weibel, editors. GIS and Generalisation: Methodology and Practice. Taylor and Francis, London, 1995.
10. Ordnance Survey. Integrated transport network. http://www.ordnancesurvey.co.uk/oswebsite /products/osmastermap/itn.
11. E. Or lowska, editor. Incomplete Information { Rough Set Anal- ysis, volume 13 of Studies in Fuzziness and Soft Computing. Physica-Verlag, Heidelberg, 1998.
12. Z. Pawlak. Rough sets. International Journal of Computer and Information Sciences, 11:341{356, 1982.

13. J. G. Stell. Relations in Mathematical Morphology with Applications to Graphs and Rough Sets. In S. Winter et al., editors, Conference on Spatial Information Theory, COSIT 2007, volume 4736 of Lecture Notes in Computer Science, pages 438{454. Springer Verlag, 2007.