Extending Irregular Cellular Automata with Geometric Proportional Analogies

Diarmuid P. O'Donoghue, Emma-Claire Mullally

Department of Computer Science/National Centre for Geocomputation,
NUI Maynooth, Co. Kildare, Ireland.
Telephone: +353 1 708 3851
Fax: +353 1 708 3851
diarmuid.odonoghue@nuim.ie, www.cs.nuim.ie/~dod

1. Introduction

We exploit the similarity between irregular Cellular Automata (CA) and Geometric Proportional Analogies (GPA), as both involve manipulations of geometric objects (points, lines and polygons). We describe how each GPA effectively defines a CA-like transition rule and we adapt an algorithm (called *Structure Matching*) used for solving GPAs to solving CAs. Irregular CAs improve on regular CAs by allowing an irregular tessellation of the plane, while further extensions support transition rules that lie beyond the scope of traditional CA. We describe three facets of the resulting model; layered inferences, incremental structures and the merge operation. Examples describe how *structure matching* (Mullally *et al*, 2005) is used to update and enhance a topographic land-cover map.

2. Regular and Irregular Cellular Automata

Regular CA were conceived by Ulam and Von Neumann in the 1950's and consist of a regular grid of *cells*, each in one of a finite number of *states*. The state of a cell at time *t* is a function of two values. First, the cell's state at time *t-1* and secondly, the states of each of a finite number of *neighbourhood* cells at time *t-1*. Neighbouring cells are defined relative to the central cell and all cells have the same update rules (or transition rules). Irregular CA use an irregular tessellation of the 2D plane and are thus more easily applied to vector data. We also include *point* and *line* features in our irregular CA, as they are regularly found in GPAs and in vector data.

While irregular CA have been proposed (O'Sullivan, 2001), no standard means of describing neighbourhoods has emerged. We exploit the similarity between the transition rules of irregular CAs and geometric proportional analogies (GPA), applying the same predicate calculus representation (Gentner, 1983) of GPAs to our CA.

2.1 Knowledge Representation and Geometric Proportional Analogies

Both irregular CA and GPAs manipulate collections of geometric objects. GPAs are IQ-test type analogy problems involving collections of geometric objects, specified in the form A:B :: C:D (read as, A is-to B as C is-to D) – see Figure 1. The objective of these problems is to generate the missing information (D) given the three other pieces of information (A, B and C). The A & B pair specifies a transformation which must be applied to C, to generate the missing D.

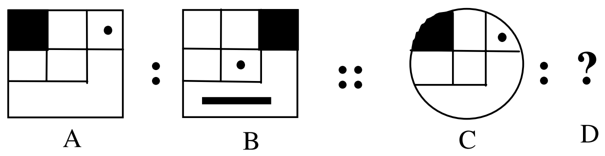

Figure 1. Apply the A-to-B transformation rule to C, thereby generating D.

Part A of Figure 1 defines the *before* part of a transition rule, while B defines the situation *after* applying the rule. In a typical GPA the A-to-B transformation will involve modifying, adding or removing information about the collection of objects introduced in A. Applying this transition rule to C allows us to generate the missing part D.

We describe parts A, B and C of a GPA using predicate calculus, thus part D can be generated by standard analogical reasoning models (i.e. Gentner, 1983, Keane *et al*, 1994; O'Donoghue *et al*, 2006). Parts A, B and C contain *points, lines* and *polygons* and each part corresponds to a CA neighbourhood. Areas that share a common boundary are neighbours to each other in a Voronoi spatial model (Gerevini & Renz, 2000) and neighbourhood topologies are described by binary relations including: *line-adjacent*, *point-adjacent* and *hasPoint*. Note these descriptions may be replaced by relations from the Region Connection Calculus (Chon, 1997) or DE9IM (Egenhofer & Herring, 1990) without affecting the remainder of the algorithm. Neighbourhoods additionally record the defining characteristics of each point, line and polygon, such as categorical information, which is recorded as unary attributes associated with the point, line or polygon.

3. The Structure Matching Algorithm

Structure Matching (O'Donoghue, 2006) is a multi-phase algorithm essentially combining Structure Mapping (Gentner, 1983), Attribute Matching (Bohan and O'Donoghue, 2000) and Copy with Substitution and Generation (CWSG) (Holyoak *et al*, 1994) processes – which we now describe.

i) Structure Mapping: Identifies the isomorphic 1-to-1 mapping between parts A and C. Structure mapping is a computationally expensive operation, being a variant of the Largest Common Sub-graph member of NP-Complete problems (Johnson and Garey, 1979). For efficiency reasons, we first apply a rough-cut filter that only allows structure mapping to proceed when the neighbourhood and the transition rule have the same number of cells in each of the states. A key output of this phase is the object-to-object alignment between the geometric objects of parts A and C.

ii) Attribute Matching: Ensures that all paired objects in the object-to-object alignment are in the same state (category or theme), as described by the objects attributes. For example, a building polygon can only be placed in correspondence with another building polygon and a text feature can only be aligned with another text feature, *etc*.

iii) CWSG – the Copy with Substitution and Generation algorithm generates the missing part D. This copies part B while substituting the mapped objects identified during the structure mapping phase.

4. Extended Cellular Automata

Many geo-spatial applications require inferences other than altering the state of a neighbourhoods central cell, adding and removing information associated with points, lines and cells, as well as changes to the topology of the neighbourhood itself. These modifications are supported by the *structure matching* algorithm as alterations to the neighbourhood definitions of each cell. However, it is important to place constraints on the (arbitrary) inferences that can be generated by the CWSG phase of *structure matching,* as it is too profligate and may destroy the integrity of the CA itself. The following three extensions overcome specific limitations identified in previous CA.

4.1 Neighbourhood Frequency

A typical map will have some neighbourhoods that occur very frequently (*e.g.* a house surrounded by a garden), while other neighbourhoods occur very infrequently. We highlight a power-law distribution in the frequencies with which neighbourhoods occur. So, a small number of neighbourhoods occur very frequently, while large numbers of other neighbourhoods will only be found once in any given map. Some categories of geographic object (e.g. roads and road-side) regularly have large numbers of neighbours requiring a large number of transition rules. But defining a transition rule for every possible neighbourhood can prove impractical. For example, a road-side's neighbourhood may contain over 60 polygons from 13 different categories, requiring an astronomical number of transition rules (Tobler, 1979) $S^N = 13^{60} = 6.8 * 10^{66}$ - even before topology is taken into account. We propose a solution to this problem, using the incremental structures described below.

4.2 Incremental Matching Method

As stated earlier, structure matching can be an expensive process, particularly for neighbourhoods with large number of objects. We now present an efficient strategy for identifying extensive structures (*eg* roads, railways and rivers) and irregular structures (*eg* hospitals, universities and schools). Our solution uses an incremental matching method (Keane *et al*, 1994) that involves two key aspects. First we identify a “root” neighbourhood that defines a starting point. Secondly, we identify a number of incremental neighbourhoods that can be iteratively added to that root or a previous incremental collection. Identifying a composite building can be achieved by identifying a “root” structure and iteratively adding connected buildings to the collection (Figure 3).

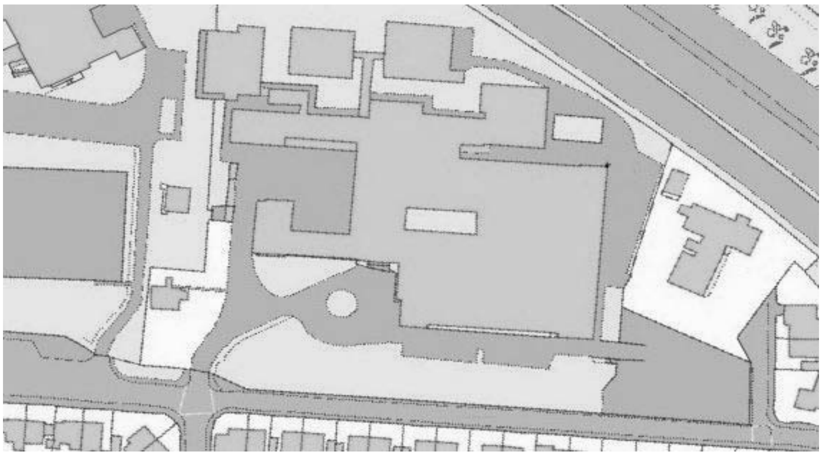

Figure 3. A composite building identified incrementally.

4.3 Layered States and Type 1 CA

Wolfram (1994) identifies four categories of CA, one of which (Type 2) converges to a unique steady state within a finite number of update cycles. Type 2 CA are particularly important as the start and final states of the CA may correspond to some situation in the real world. We ensure convergence to a final state by introducing additional states and by layering the inferences generated by *structure matching*. We define L extra states, in addition to the K initial cell states and refer to these L additional states as first-order states if they are derived directly from the initial K states. States belong to layer N if they can only be derived from the states in layer N-1. Additional states may correspond to sub-categorizations of the initial K states contained within a topographic map.

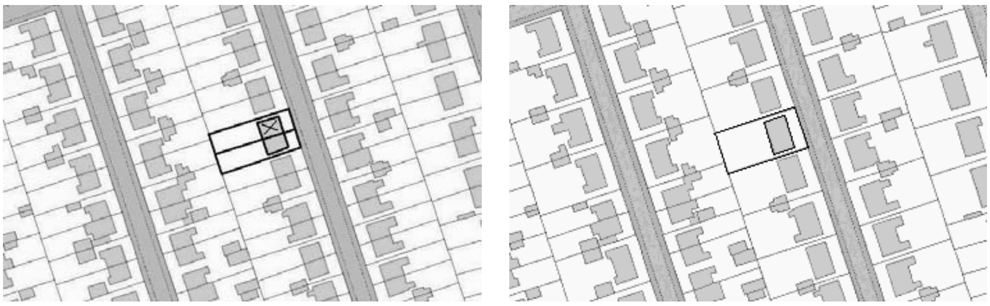

Figure 2. (left) Neighbourhood rule identifies a semi-detached dwelling. Generalisation by merging select neighbourhood polygons (right).

4.4 Merge Cells

The inferences of our extended CA may alter properties of points and lines as well as cells. As our neighbourhoods are essentially topological descriptions, modifications to the neighbourhood topology are easily achieved by changing the relations in the updated neighbourhood description.

Problematically, modifying or deleting cells can destroy the crucial contiguity of the CA, creating cells with incomplete neighbourhoods – such as that used by Shi and Pang (2000). To overcome this problem we use a *merge* operation which ensures that no boundary cells are ever accidentally created. This *merge* operation can be applied to points, lines and cells. Figure 2 illustrates how merge can be used for detail reduction and generalisation. Features may still be removed by simply merging them with existing cells without negative consequences and therefore reach the final solution state. However, no constraints are required on modifying or deleting existing points or lines. Furthermore, no constraints are required on inserting new points, lines or cells.

5. Conclusion

We exploit the similarity between irregular Cellular Automata (CA) and geometric proportional analogies (GPA), as both involve manipulations of specific configurations of geometric objects (involving points, lines and polygons). We adapt the knowledge representation and algorithm used for solving analogy problems to the domain of geometric information.

We introduce three extensions to the basic CA, overcoming some practical limitations on the application of CA to some problems types. First we introduce layered states that increase the number of states in the CA while ensuring it will converge to a stable state in a finite number of steps. Secondly, an incremental matching method addresses the problems of processing large structures that extend across many neighbourhoods. This incremental method is also useful in identifying irregular structures that are difficult to represent with standard transition rules. Finally, we describe some limitations that are placed on the modification and deletion of cells in the CA to maintain the crucial contiguity of the CAs cellular structure. These three methods have been successfully implemented to enhance the data contained within topographic maps.

6. Acknowledgements

Glen Hart and the Ordnance Survey Research Centre, Southampton, UK.

7. References

AG COHN, B BENNETT, J GOODAY, NM GOTTS, 1997 “Qualitative Spatial Representation and Reasoning with the Region Connection Calculus”, GeoInformatica.

A BOHAN, D. O’DONOGHUE, 2000 “A Model for Geometric Analogies using Attribute Matching”, AICS-2000 11th Artificial Intelligence and Cognitive Science Conference, Aug 23-25, NUI Galway, Ireland.

EGENHOFER, M.J., HERRING, J. 1990. “Categorizing binary topological relations between regions, lines, and points in geographic databases”, Fourth International Symposium on Spatial Data Handling, 803-813, Zurich, Switzerland.
GENTNER, D. 1983. “Structure-Mapping: A Theoretical Framework for Analogy”, Cognitive Science, 7, 155-170.
GEREVINI, A. RENZ, J. 2000. “Combining topological and size information for spatial reasoning”, Artificial Intelligence, 137, 1-42.
HOLYOAK K. J. NOVICK L. MELZ E. 1994 ”Component Processes in Analogical Transfer: Mapping, Pattern Completion and Adaptation”, in Analogy, Metaphor and Reminding, Eds. Barnden and Holyoak, Ablex, Norwood, NJ.
JOHNSON, GAREY, 1979, “Computers and Intractability”, WH Freeman, NY, USA.
KEANE, M. T. LEDGEWAY, T. DUFF, S. 1994.”Constraints on Analogical Mapping: A Comparison of Three Models”, Cognitive Science, 18, 387-438, 1994.
E. MULLALLY, D. O’DONOGHUE, 2005 “Geometric Proportional analogies in Topographic Maps”, 25^{th} SGAI International Conference on Innovative Techniques and Applications of Artificial Intelligence, Cambridge, England.
O'DONOGHUE, D. P. BOHAN, A. KEANE, M.T. 2006. “Seeing Things - Inventive Reasoning with Geometric Analogies and Topographic Maps”, New Generation Computing, Springer and Oshima, Vol. 24, (3) pp 267-288, 2006.
O'SULLIVAN, 2001. “Exploring spatial process dynamics using irregular cellular automaton models”. Geographical Analysis 33(1), 1–18.
RANDELL, D.A., CUI, Z. AND COHN, A.G. 1992 “A Spatial Logic Based on Regions and Connection.” Proc. 3rd Int. Conf on Knowledge Representation and Reasoning.
SHI W. AND PANG, M.Y.C. 2000. “Development of Voronoi-based Cellular Automata”, Int. J. Geographical Information Science, Vol. 14, No. 5, 455-474.
TOBLER, W. R. 1979 “Cellular Geography” in Philosophy in Geography, eds. S Gale, G Olsson (D. Reidel, Dordrecht) pp 379-386.
WOLFRAM, S. 1994.”Cellular Automata and Complexity”, AddisonWesley.

Biography

Dr. Diarmuid O’Donoghue is a lecturer in the Department of Computer Science and an associate of the National Centre for Geocomputation. His interests include cognitive modelling and spatial reasoning.