# Using the CoSMoS approach to study Schelling's Bounded Neighbourhood Model

Ali Afshar Dodson[1,2], Susan Stepney[1,2], Emma Uprichard[3], and Leo Caves[1,4]

[1] York Centre for Complex Systems Analysis, University of York, YO10 5DD
[2] Department of Computer Science, University of York, YO10 5DD
[3] Centre for Interdisciplinary Methodologies, University of Warwick, CV4 7AL
[4] Department of Biology, University of York, YO10 5DD

**Abstract.** The basic CoSMoS process concerns the design, implementation, and use of a simulation built from scratch. However, the CoSMoS approach may be tailored and adapted for other styles of use. Here we describe how it has been applied to analyse and re-engineer an existing simulation, that of Schelling's Bounded Neighbourhood Model. We find that using a principled approach to the analysis of an existing simulation facilitates formalisation of the model and reimplementation of the simulation. In the process, several ambiguities in implementing a simulation from the model were revealed. This highlights the importance of formalising a model for clarity and reproducibility in simulation studies, and also for providing new avenues for exploration of, and insight into, the factors influencing its emergent behaviour.

## 1 Introduction

The Complex System Modelling and Simulation (CoSMoS) approach [2] has been designed for the purpose of developing and using a simulation as a scientific instrument [5]. It provides a guide for modelling and simulating complex systems, and incorporates verification and validation throughout. It provides a structure for the development and use of simulations in an interdisciplinary endeavour between scientists who study a particular domain (the *domain scientists*), and software engineers who construct simulations to facilitate the study of that domain (the *simulation engineers*).

The CoSMoS approach has already been successfully used in a number of studies including auxin transport canalisation [9], environment orientation [11], immunology [1] and cancer systems biology [6]. Despite this specific use, CoSMoS can also be tailored and adapted for other uses. For example, it can be used to design bio-inspired algorithms [4], and to reverse engineer models from implementations [3]. Here we describe how it has been used to formalise, reimplement, and analyse an existing third party model and simulation: Schelling's Bounded Neighbourhood Model [15] of segregation.

The structure of the rest of the paper is as follows:

- §2 summarises the relevant parts of the CoSMoS approach.

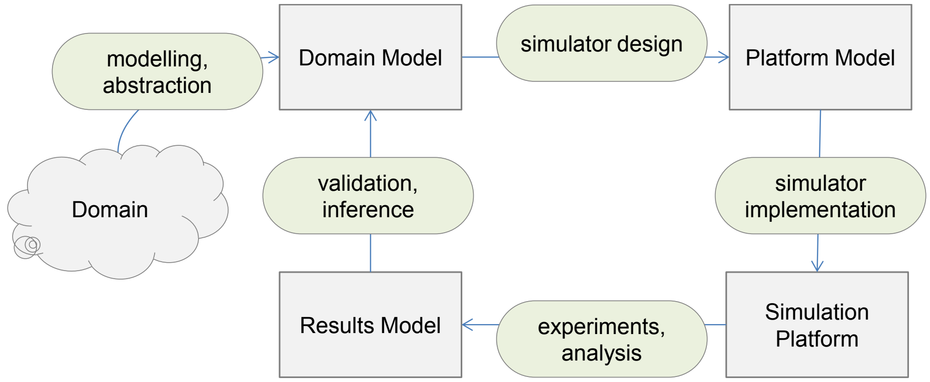


**Fig. 1.** The generic Complex Systems Modelling and Simulation (CoSMoS) approach. Products are shown in rectangles; activities are shown in rounded boxes; the entire approach (products and activities) takes place within a specified research context.

– §3 introduces Schelling's Bounded Neighbourhood Model.
– §4 uses CoSMoS concepts to analyse Schelling's model, formalise it, and then reimplement it using the standard CoSMoS approach.

## 2 The CoSMoS approach

The CoSMoS process is described in detail in [2]. It comprises a set of products (models, software, arguments), including [2, p13]:

**Research Context** : captures the overall scientific research context of the simulation development project, including the motivation for the research, the questions to be addressed by the Simulation Platform, and the requirements for validation and evaluation.

**Domain** : the domain scientist's view of the subject of simulation; for example, a real-world system that is the subject of scientific research, or an engineered system that is the subject of engineering research and design.

**Domain Model** : distils appropriate aspects of the Domain into explicit domain understanding. The Domain Model focuses on the scientific understanding; no simulation implementation details are considered.

**Platform Model** : comprises specification, design and implementation models for the Simulation Platform, based on the Domain Model and research context. Any emergent properties captured in the Domain Model that are hypothesised to arise from lower level interactions are removed from the Platform Model, to ensure that the desired result is not explicitly coded into the simulations.

**Simulation Platform** : encodes the Platform Model into a software and hardware platform with which simulation experiments can be performed.

**Results Model** : encapsulates the understanding that results from simulation experiments: the Simulation Platform behaviour, results of data collection and observations of simulation runs. Note that the way that the Domain Model captures the relevant understanding of the domain, via experiments,observation, and theory, is mirrored by the way that the Results Model captures understanding of the simulation experiments.

By following the CoSMoS process an important distinction is made, not only between the model and the software that implements the model in the simulation, but also between the results from the simulation and any results from the 'real world' domain.

## 3 Schelling's Bounded Neighbourhood Model

Schelling's 1971 paper "Dynamic Models of Segregation" [15] is a seminal work in the field of Computational Social Science and is often used as a classic example of what agent based modelling (ABM) can offer to social sciences and urban studies. The models of segregation presented therein have been the focus of a number of studies, using a variety of different techniques. Schelling used agent based models to explore ideas of segregation in heterogeneous populations. He developed models to examine levels of segregation between two populations, using simplified ideas of social interaction. His work has been described as the first ABM [7]. It has also been called the first simulation of an *artificial* society [14].

Schelling describes two distinct models, a Spatial Proximity model [15, p149] and an aspatial Bounded Neighbourhood model [15, p167]. In both models, agents have a general preference for their own type (*like*) over the other type (*unlike*), and move to maintain a favourable ratio. Schelling found that mixed populations are impossible to maintain, with populations eventually segregating, even when the desire for like over unlike is quite weak.

The vast majority of the work after Schelling focuses on the grid-based Spatial Proximity Model in both one [13] and two [8] dimensions. Here we focus on the less well studied aspatial Bounded Neighbourhood Model [15].

The Bounded Neighbourhood Model has a heterogeneous population of two agent types, with some agents inside and some outside an aspatial *neighbourhood*. It considers the flow of agents into, and out of, the neighbourhood. Each agent calculates its *happiness*, based on the current ratio of the number of the other type, compared to the number of its own type, inside the neighbourhood. This ratio is compared to the agent's individual fixed *tolerance* of such a ratio. If its tolerance exceeds the ratio, it is deemed *happy*; if its tolerance is below the ratio, it is deemed *sad*. Sad agents inside the neighbourhood leave, while happy agents outside enter.

Figure 2 shows Schelling's analysis of the resulting population flows, given a total population of 100 $W$s and 50 $B$s, where the agents of a given type have a uniform distribution of tolerances, from 2 (willing to tolerate twice as many unlike agents as like agents) down to 0 (unwilling to tolerate any unlike

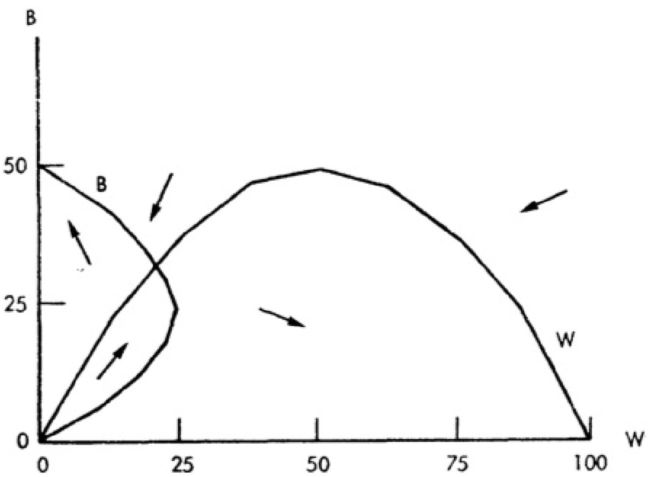


**Fig. 2.** Schelling's results (from [15])

agents). The change of population numbers inside the neighbourhood for any given current number is shown by the arrows. In some regions both $W$ and $B$ agents leave, in some both enter, and in some one type enters and one leaves. The resulting equilibrium state is tipped to either all $B$ or all $W$. This is despite the fact that half of each population is actually willing to tolerate a (limited) majority of unlike agents: the migration of the least tolerant changes the ratios enough that even the most tolerant find themselves sufficiently outnumbered that they too leave. The same result is found with different population numbers, and different tolerance schedules.

The popularity of Schelling's ideas led to criticisms, most notably Yinger [16], who argues that there is no consideration of a number of factors such as economics and social mobility. Similarly, Massey [12] argues that not accounting for environmental/spatial factors and using a homogenous environment means the models are too simplistic to have any relation to reality. If the model were being developed today, this could be a fair criticism, since the input of experts in the field is an important part of the modelling process. However, it is important to recognise the purpose and scope of Schelling's simple model. Schelling demonstrated that a simple desire for non-minority status in a population could lead to highly segregated systems.

Recent work has nevertheless attempted to apply Schelling's models to 'real-world' situations [10]. Such attempts, without acknowledging the limitations of the models, could lead to invalid conclusions. We apply the CoSMoS approach retrospectively to Schelling's Bounded Neighbourhood Model [15, p167], in order to formalise and re-implement it, and to analyse the consequences of its assumptions through simulation.

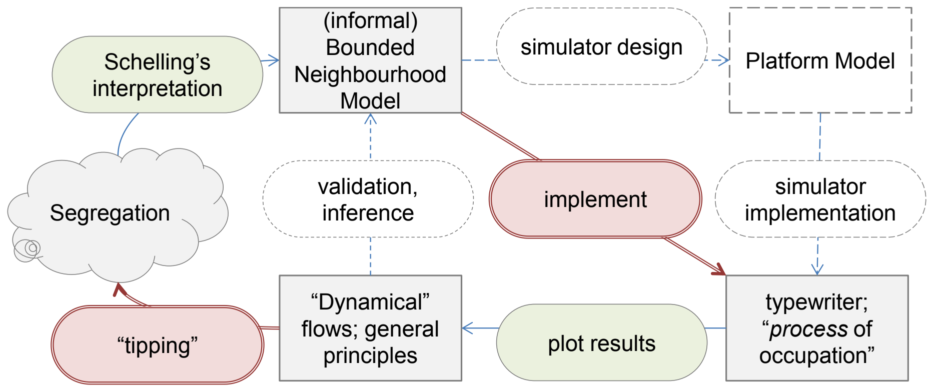


**Fig. 3.** Schelling's process in developing the Bounded Neighbourhood Model. The dashed boxes show the CoSMoS products and activities not present in Schelling's work, whilst the double edged boxes and lines are alternative steps employed.

Whilst our work is an attempt to gain a better understanding of Schelling's model, his interpretation is necessarily retained. *Because of this, and the inherent difficulty of relating models to reality, we make no attempt to relate the model back to reality.* Instead, results from the simulation are used to formulate questions about Schelling's model.

## 4 Schelling's development through a CoSMoS lens

Here we analyse Schelling's process in building his model in terms of the CoSMoS approach (figure 3). There are points where Schelling's approach differ from the CoSMoS approach. We need to analyse these differences to determine their implications for his model, the simulation results, and subsequent interpretations. Any shortcomings are not, necessarily, the fault of Schelling; the work presented here is an attempt to 'lift' Schelling's model to current computational modelling standards.

Firstly, and most importantly, Schelling's first step from the domain (Segregation) to the Domain Model (informal Bounded Neighbourhood Model) is Schelling's personal interpretation [15, p143], rather than being based on observational and experimental data from the Domain. Whilst this may be acceptable (but not advised) for models in which the simulation engineer is also a domain expert, Schelling's background was mathematical, rather than sociological. CoSMoS requires the Domain Model to be built in collaboration with domain scientists, or, at the very least, validated by them.

Secondly, Schelling's presentation of the Domain Model is informal, and contains ambiguities that require resolution before a Platform Model or Simulation can be developed. For example, when talking of the movement of sad agents, Schelling states "some will move" in order of tolerance, but does not specify

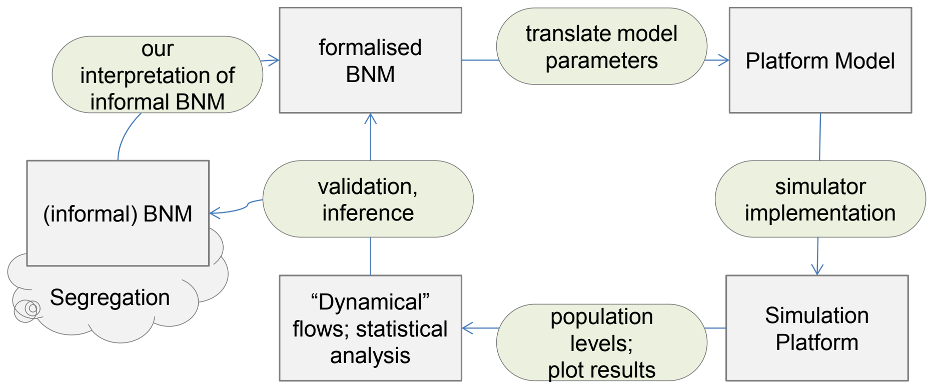


**Fig. 4.** An application of the CoSMoS process in developing a simulation of Schelling's Bounded Neighbourhood Model. Note the outputs of the process feed back into the interpretation, rather than the real world.

more precisely which ones (which type?), or how many (one? all?). Which agents move at each step needs to be formalised before a computational simulation can be developed, and the decisions published, for clarity and reproducibility.

Next, Schelling proceeds directly from this informal Domain Model to an simulation, without passing through an equivalent of the CoSMoS Platform Model. In this case the Domain Model is very simple, and it might be thought that no Platform Model is necessary. However, some of the ambiguities identified in the Domain Model became clear only by attempting to formalise them within a Platform Model (order of execution of operations, for example). If proceeding directly to implementation, such ambiguities would be resolved only in the code, leaving their resolution (and existence) opaque to the general reader. Note that in Schelling's case, there was no computer-based simulation: the work was done with pencil and paper, or on a typewriter. Hence these ambiguities were resolved in Schelling's head, leaving their resolution even more opaque.

Finally, Schelling takes his results and then attempts to apply them to a real world event (neighbourhood tipping) [15, p181]. Because of the extreme distance of his Domain Model from the real world Domain of population segregation, and lack of any validation from social scientists, it is difficult to defend attempts to relate his results back to reality. This argument is used by a number of authors who attack his models as missing essential components of reality [16], [12].

## 5 Reimplementing the Bounded Neighbourhood Model using CoSMoS

Having analysed Schelling's model, and his (assumed) development approach, we are now in a position to reimplement the model, in the form of a simulation, suitable for further experimentation. We use the basic CoSMoS approach to do

**Algorithm 1:** Agent movement rule, one timestep

**if** $\exists a_n \in E * \wedge \tau(a_n) \geq R_n$ (there is a happy agent outside) **then**
  move $a_n$ with max $\tau$ to $E$ (move happiest agent inside)
**else if** $\exists a_n \in E \wedge \tau(a_n) \leq R_n$ (there is a sad agent inside) **then**
  move $a_n$ with min $\tau$ to $E*$ (move saddest agent outside)

so, modified to accommodate the fact that we are using only literature, not domain experts, in order to build the models and simulation.

### 5.1 Domain Model

Following the CoSMoS approach, we refined Schelling's formulation into a formalised Bounded Neighbourhood Model, specifying mathematically what is meant by terms such as 'environment', 'agent location', 'agent type', 'ratio', 'tolerance', 'happiness', and agent movement. This activity highlighted the key parameters of the model and made them explicit, removing ambiguities that could arise from the language of the model [Afshar Dodson *et al*, forthcoming *a*].

Note here that we effectively took Schelling's informal model as our Domain, and formalised it by examining the text in detail, and by using social science knowledge (of author EU).

### 5.2 Platform Model and Simulation Platform

Once explicitly defined, the parameters were encoded in the Platform Model. This process highlighted a few more ambiguities, this time in the actual algorithm used to iterate through generations of movement in and out of the neighbourhood. These were resolved by specifying the details in high-level pseudo-code (for example, algorithm 1), using the mathematical notation defined in the Domain Model. Here we had to make specific choices (such as, the happiest agent moves); highlighting the choice allows us to make other choices (such as, a random happy agent moves) to investigate the effect of the choice on the behaviour of the system [Afshar Dodson *et al*, forthcoming *b*].

The Platform Model was used as the specification for a NetLogo ABM implementation.

### 5.3 Simulation Experiment Results

Given the re-implemented Simulation Platform, it was possible to conduct a number of experiments on the model.

In the first iteration, the validity of the formalised model was tested against the original model results. We initialised the model to each of the possible starting conditions (number of agents of each type) and plotted the resulting flows on a 2D plane (figure 5). The comparison shows the simulation reproduces

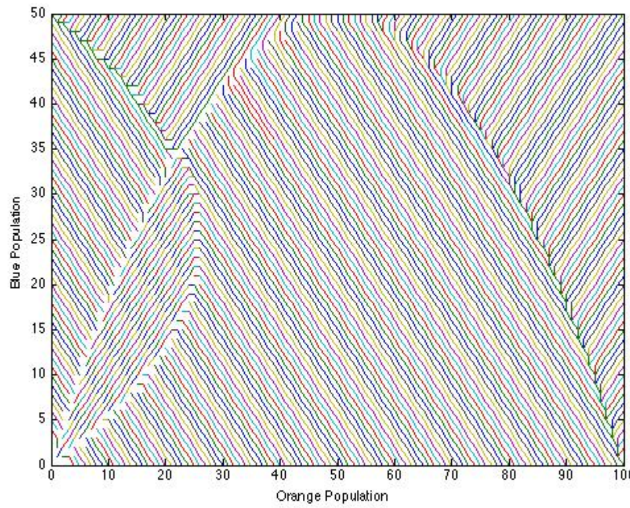


**Fig. 5.** To validate the simulation, its output and Schelling's results (figure 2) were compared. For each initial condition the resulting population flow is plotted. The simulation results clearly show Schelling's boundaries, with the flows following Schelling's results.

Schelling's flows (figure 2), providing evidence that it can be considered a valid simulation of Schelling's Bounded Neighbourhood Model.

Given a valid formalisation and Simulation Platform, we could then perform further simulation experiments, and to explore the assumptions underlying the original model, and determine how the resulting behaviour depends on those assumptions. We have discovered that the model is surprisingly robust to quite drastic changes in the underlying assumptions: it is very difficult to change find a model that does not inevitably result in segregation [Afshar Dodson *et al*, forthcoming *b*].

## 6 Summary and Conclusions

We have applied the basic CoSMoS approach in order to develop a formalised model and validated simulation of Schelling's Bounded Neighbourhood Model, suitable for use as the basis for further simulation experiments to gain a better understanding of the factors underlying its emergent properties.

This work demonstrates that the concepts from the CoSMoS approach can be employed to interrogate a previously developed model and simulation. We have analysed Schelling's original model through the CoSMoS approach, identified products and activities missing or glossed over, and explained why these are necessary for a reproducible simulator, even in such a seemingly simple model as this. Our intention is not to criticise Schelling's approach; his work represents the very earliest steps in agent-based social simulation, and he was running experiments on a typewriter or using pencil and paper, resolving ambiguities on the fly. However, it is telling that few if any subsequent authors implementing his model in a computational setting have noted the many ambiguities in Schelling's original description.

One of the key advantages to using the CoSMoS approach is that it breaks down the making and shaping of a simulation into distinct parts of the modelling process. Importantly, it does this without losing the depth and breadth of expertise that may be required at different parts of the process, whilst at the same time allowing for the model to developed iteratively, and thereby refined and tested over time. By focusing on Schelling's Bounded Neighbourhood Model, we have shown how the CoSMoS approach can be usefully deployed to re-develop past models.

We have deliberately made no attempt to relate this particular model to reality, yet the CoSMoS approach may be especially interesting for modelling social phenomena more generally. Social behaviours are emergent properties of an underlying complex system. The crux in the interpretation of social simulation is what factors influence the emergent features. Such factors span issues of approximation, implementation and potential coding errors. Thus, it is essential that models and their implementations are formalised and carefully described, in order that their results and interpretation may be fully reproduced and tested. These steps are necessary to put simulation studies (in general, not just in the social domain) into the realm of science. Principled model construction approaches, such as CoSMoS, offer a promising way forward.

In redeveloping Schelling's model, we have also, by implication, made a first step towards an ambitious longer term project that involves using CoSMoS as way of rethinking how we might methodically and systematically (re)explore causal mechanisms from which the social emerges.

## References


1. K. Alden, J. Timmis, P. S. Andrews, H. Veiga-Fernandes, and M. C. Coles. Pairing experimentation and computational modeling to understand the role of tissue inducer cells in the development of lymphoid organs. *Frontiers in immunology*, 3(172), 2012.
2. P. S. Andrews, F. A. C. Polack, A. T. Sampson, S. Stepney, and J. Timmis. The CoSMoS Process, Version 0.1: A Process for the Modelling and Simulation of Complex Systems. Technical Report YCS-2010-453, University of York, 2010.
3. Paul S. Andrews and Susan Stepney. Using CoSMoS to reverse engineer a Domain Model for Aevol. In *Proceedings of the 2014 Workshop on Complex Systems Modelling and Simulation, New York, USA, 2014*. Luniver Press, 2014.
4. Paul S. Andrews, Susan Stepney, Tim Hoverd, Fiona A. C. Polack, Adam T. Sampson, and Jon Timmis. CoSMoS process, models, and metamodels. In *Proceedings of the 2011 Workshop on Complex Systems Modelling and Simulation, Paris, France, August 2011*, pages 1–13. Luniver Press, 2011.

5. Paul S. Andrews, Susan Stepney, and Jon Timmis. Simulation as a scientific instrument. In *Proceedings of the 2012 Workshop on Complex Systems Modelling and Simulation, Orleans, France, September 2012*, pages 1–10. Luniver Press, 2012.
6. J. Bown, P. S. Andrews, Y. Denni, A. Goltsov, M. Idowu, F. A. C. Polack, A. T. Sampson, M. Shovman, and S. Stepney. Engineering Simulations for Cancer Systems Biology. *Current Drug Targets*, 13(12):1560–1574, 2012.
7. J. M. Epstein. Why model? *Journal of Artificial Societies and Social Simulation*, 11(4):12, 2008.
8. M. Fossett. Ethnic preferences, social distance dynamics and residential segregation. *Presented at the Annual Meetings of the American Sociological Association*, March 1999.
9. P. Garnett, S. Stepney, F. Day, and O. Leyser. Using the CoSMoS Process to enhance an executable model of auxin transport canalisation. In *Proceedings of the 2010 Workshop on Complex Systems Modelling and Simulation, Odense, Denmark, August 2010*, pages 9–32. Luniver Press, 2010.
10. E. Hatna and I. Benenson. The Schelling model of ethnic residential dynamics: Beyond the integrated - segregated dichotomy of patterns. *Journal of Artificial Societies and Social Simulation*, 15(1):6, 2012.
11. Tim Hoverd and Susan Stepney. Environment orientation: an architecture for simulating complex systems. In *Proceedings of the 2009 Workshop on Complex Systems Modelling and Simulation, York, UK, August 2009*, pages 67–82. Luniver Press, 2009.
12. D. S. Massey and N. A. Denton. *American apartheid: segregation and the making of the underclass*. Harvard University Press, 1993.
13. R. Pancs and N. J. Vriend. Schelling's spatial proximity model of segregation revisited. *Journal of Public Economics*, 91:1–24, 2006.
14. R. Sawyer. *Social Emergence, Societies as Complex Systems*. Cambridge University Press, 2008.
15. T. C. Schelling. Dynamic Models of Segregation. *Journal of Mathematical Sociology*, 1:143–186, 1971.
16. J. Yinger. *Closed doors, opportunities lost: the continuing costs of housing discrimination*. Russel Sage Foundation, 1997.