On the derivation and application of closure approximations of cellular automata models

Kale Davies

Thesis submitted for the degree of
Doctor of Philosophy
in
Applied Mathematics
at
The University of Adelaide
(Faculty of Engineering, Computer and Mathematical Sciences)

School of Mathematical Sciences

THE UNIVERSITY
of ADELAIDE

September 19, 2016
Contents

Signed Statement vii
Acknowledgements ix
Abstract xi

1 Introduction 1
  1.1 A brief summary of the thesis 3

2 Cellular automata: descriptions, background and definitions 7
  2.1 Literature review and background 7
    2.1.1 What is a cellular automata model? 7
    2.1.2 Conway’s Game of Life 10
    2.1.3 Elementary CA 12
    2.1.4 CA research 14
    2.1.5 Population density and moment closure 16
    2.1.6 Stochastic CA and site occupancy 18
    2.1.7 Kirkwood superposition approximation 20
  2.2 Defining our CA model 22
    2.2.1 CA model 22
    2.2.2 Discrete-time CA model 23
    2.2.3 Discrete-time Markov chains 24
    2.2.4 Continuous-time Markov Chain 25
    2.2.5 Continuous-time CA model 27
  2.3 Probability notation 28
    2.3.1 Probability of site occupancy 28
    2.3.2 Approximations to the probability of site occupancy 30
  2.4 A detailed summary of the thesis 31
3 Mean-field approximations and the assumption of independence
3.1 Review of mean-field theory
3.1.1 MF theory for population density
3.1.2 MF theory for site occupancy
3.2 Discrete-time CA model
3.2.1 Deriving a continuum approximation
3.2.2 Type A approximation
3.2.3 Type B approximation
3.3 Continuous-time CA model
3.3.1 Reduced errors in the continuous-time model
3.3.2 Ordinary and partial differential equation approximations
3.4 Analysis
3.4.1 Method of analysis
3.4.2 Analysis of approximations
3.5 The independence assumption
3.6 Summary

4 Formalisation of closures including existing approximations
4.1 Review of moment closure techniques
4.1.1 Local structure theory
4.1.2 CF–KSA method
4.2 Formalisation of closure methods
4.2.1 Closing the system
4.2.2 Symmetry and independence
4.2.3 Marginal probabilities and the law of total probability
4.3 Mean-field closure
4.3.1 Closure approximation
4.3.2 Rate of change functions
4.3.3 Total and marginal probabilities
4.3.4 Approximating site occupancy
4.4 Local structure closure
4.4.1 Closure approximation
4.4.2 Rate of change functions
4.4.3 Total and marginal probabilities
4.4.4 Approximating site occupancy
4.4.5 Preventing division by zero
4.4.6 Relation to MF closure ........................................ 103
4.5 KSA closure ...................................................... 104
  4.5.1 Closure approximation .................................... 105
  4.5.2 Rate of change functions ................................ 109
  4.5.3 Marginal probabilities .................................... 110
  4.5.4 Approximating site occupancy ......................... 112
  4.5.5 Comparison to CF–KSA .................................... 116
4.6 Analysis ......................................................... 117
  4.6.1 LS analysis ................................................ 118
  4.6.2 KSA analysis .............................................. 122
  4.6.3 Comparison of closure approximations ............... 122
  4.6.4 Proliferation-only results .............................. 130
4.7 Summary ....................................................... 131

5 Block closure approximations 133
  5.1 Introduction .................................................. 133
  5.2 Block mean-field closure ................................. 134
    5.2.1 Closure approximation ................................. 134
    5.2.2 Marginal probabilities ................................ 137
    5.2.3 Implementation ........................................ 139
    5.2.4 Symmetry of basis blocks ............................ 143
    5.2.5 Equivalent approximation using the MF closure .. 144
  5.3 Overlapping-blocks closure .............................. 147
    5.3.1 Closure approximation ................................. 147
    5.3.2 Marginal probabilities ................................ 149
    5.3.3 Implementation ........................................ 161
    5.3.4 Probabilistic interpretation ......................... 164
  5.4 Analysis ...................................................... 165
    5.4.1 Comparison of OB closures .......................... 176
  5.5 Summary ...................................................... 177

6 Extending the KSA closure 179
  6.1 Introduction ................................................ 179
  6.2 KSA3 closure ................................................ 180
    6.2.1 Closure approximation ................................. 181
    6.2.2 Rate of change equations ............................ 182
Signed Statement

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint award of this degree.

I give consent to this copy of my thesis, when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

I also give permission for the digital version of my thesis to be made available on the web, via the University’s digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

SIGNED: ....................... DATE: .........................
Acknowledgements

Throughout my education at the University of Adelaide, I have had the pleasure of interacting and working with so many wonderful people. I am appreciative to all of the staff, students, family and friends that have shaped my life positively throughout this time and whose influence has ultimately led to the completion of this thesis. I would like to take this opportunity to specifically thank a small number of people that have had a particularly large impact on my research and my life throughout my PhD.

I would like to thank Max Malacari, who has been my closest friend through both my undergraduate and postgraduate degrees. We have shared many incredible experiences and discussed countless ideas and issues together. I am grateful to have made such a close and reliable friend in my time at university.

I would like to thank my best friend and partner, Anna Kovac. She given me so much support throughout the majority of my PhD. She has taught me so many things in the relatively short time I have known her and reshaped the way I think about people and the world, making me a better person in the process. I am constantly thankful for the impact she has had on my life.

I am extremely grateful to Ed Green for his contributions to my work regarding singular perturbation problems. His assistance was fundamental to some early conclusions reached within my research which led to the bulk of the work produced in this thesis.

Finally, I would like to thank my supervisors, Benjamin Binder, Joshua Ross and Nigel Bean, for their assistance, support and patience. They were constantly available to give both personal and academic advice and went to massive efforts to ensure they had proofread my thesis with less than ideal time restrictions. I will greatly miss our weekly meetings which were always full of laughs and a highlight of my experience as a PhD student.
Abstract

We consider discrete agent-based models known as a cellular automata (CA). CA models are lattice-based systems, in which each position on the lattice, referred to as sites, may take on a number of states. The overall dynamics of the CA model are dictated by events which allow the states of lattice sites to change over time based on the states of neighbouring sites. As such, CA models are particularly good at modelling physical systems where the spatial distributions are of interest.

In particular, we are interested in exploring CA models that are derived from applications of cellular biology such as tissue and tumour growth or wound healing. As such, we consider a CA model with two possible states: occupied, referring to sites where a cell is present, and vacant, referring to sites where there is no cell present. We then describe two events for our CA model, based on dynamics that are generally present in such cell-based problems: motility and proliferation. A motility event will swap the states of two adjacent sites, corresponding to the movement of a cell from one position to another. A proliferation event will cause a previously vacant site to become occupied due to an interaction with an adjacent occupied site, corresponding to cell division.

Initially, each site is designated either an occupied or vacant state. Given a lattice with some initial condition, we wish to know the probability of each site being occupied at some later time. The CA model is a continuous-time Markov chain and as such, a solution to the model exists and can theoretically be obtained. Unfortunately, in order to obtain this solution, we must compute a corresponding system of differential equations based on the size of the state space of the Markov chain. The state space of the Markov chain is based on the number of different configurations of the states of the sites in the CA model, hence for a domain with X sites, we must numerically solve a system of $2^X$ differential equations. As the time taken to compute the solution to this system will increase with the number of equations in the system, it is not computationally viable to solve these problems, even when the total number of sites X is relatively small.
In order to solve the system, approximations are made, reducing the number of equations, resulting in a system of differential equations that are computationally manageable. The simplest existing approximation is known as mean-field (MF) theory. By applying MF theory, an assumption of independence is made between the state of each site. This reduces the system to only $X$ differential equations at the expense of no longer modelling the correlation between the states of each site, however we find that this yields inaccurate approximations in most cases. Some more sophisticated approximations have been considered in the literature in which some of the correlation between sites is modelled. These are also considered in this thesis.

We formalise an approach that allows us to derive a class of approximate systems based on a statistical technique called moment closure. Moment closure aims to describe approximations to statistical distributions using finitely many moments of a system (where in reality there are infinitely many moments). In the same manner, applying a closure approximation to a CA model reduces the system so that we need only consider a much reduced number of differential equations. While the concept of closure approximations has been applied to CA models, pre-existing approaches have had unaddressed issues that prevent them from performing optimally. Our formal approach outlines these issues as they appear by making a distinction between the differential equations governing the physical system and the differential equations governing the approximate system. Our approach is applied to MF theory, as well as the other pre-existing approximations.

After demonstrating the functionality of the formal closure approach, we begin to develop our own closure approximations. This allows us to develop and analyse new approximate systems with the goal of accurately approximating the motility and proliferation CA model. By exploring a variety of approximations, we come to understand the desirable and undesirable properties of the closures. By enforcing the desirable properties and rejecting closures with undesirable properties, we are able to restrict our search and develop an approach for deriving closure approximations. We demonstrate this approach, showing some interesting cases before drawing general conclusions about closure approximations when applied to CA models with different rules. Furthermore, we make conclusions about the performance of the closure approximations derived throughout the thesis when applied to the motility and proliferation CA model.