

Torbjørn Helvik

Dynamical Systems of Interacting Units: Information Transport and Higher Order Structures

Doctoral thesis
for the degree of philosophiae doctor

Trondheim, November 2005

Norwegian University of
Science and Technology
Faculty of Information Technology, Mathematics and Electrical
Engineering
Department of Mathematical Sciences



NTNU

Norwegian University of
Science and Technology

Doctoral thesis
for the degree of philosophiae doctor

Faculty of Information Technology, Mathematics and Electrical
Engineering
Department of Mathematical Sciences

©Torbjørn Helvik

ISBN ISBN 82-471-7738-2 (printed ver.)
ISBN ISBN 82-471-7737-4 (electronic ver.)
ISSN 1503-8181

Doctoral Theses at NTNU, 2006:1

Printed by Tapir Uttrykk

Abstract

How can one model, characterize and analyse systems consisting of a large number of interacting units? This thesis addresses various related aspects of this issue using dynamical systems theory. The thesis is constituted by four papers presenting the main findings as well as an introductory part giving the context of the work.

In Paper I, we use evolutionary game theory to study a social dilemma where a group can obtain a collective benefit, but only if a threshold on the number of members willing to cooperate at a personal cost is met. The replicator dynamics of our model has interesting features, including catastrophic events. We also generalize the analysis to other n -player games.

In Paper II, we introduce several new concepts and principles for describing and studying hierarchical multilevel systems with interactions both between groups and among individual units. We illustrate these concepts through defining a class of systems called higher order cellular automata, and discuss new phenomena and patterns of behaviour found in these systems.

In Paper III, we introduce a local information measure for one-dimensional lattice systems in order to characterize the coherent structures emerging from the dynamics of higher order cellular automata as well as ordinary cellular automata. A further objective is to investigate to what extent information can be viewed as a local quantity in such systems. We demonstrate the applicability of our information measure to these problems.

Paper IV is a continuation of Paper III. We show that local information is a locally conserved quantity in the important class of surjective cellular automata and provide bounds on the information transport.

The novel concepts and findings of the thesis may lead to an improved understanding of various phenomena found in systems of interacting units, and in particular systems with a multilevel dynamics.

Preface

This thesis is submitted in partial fulfilment of the requirements for the degree of philosophiae doctor (PhD) at the Norwegian University of Science and Technology (NTNU). The research was funded by the Research Council of Norway, and was carried out at the Department of Mathematical Sciences, NTNU, except for nine months during fall 2003 and spring 2004, when I visited the Complex Systems Group at the Department of Physical Resource Theory at Chalmers University of Technology in Göteborg, Sweden.

I would like to express my gratitude for the support and guidance of my advisor, Nils A. Baas, and of my two mentors at Chalmers, Kristian Lindgren and Mats G. Nordahl. Moreover, I am grateful to the Department of Physical Resource Theory for allowing me to stay there, to Kristian for teaching me about information theory, and to Mats for sharing perspectives on physics and on science in general. I would also like to thank Lars Bach for interesting discussions about games, ecology, and evolution, and Christian Skau and Jeffrey Steif for help with the world of σ -algebras. Finally, I have really appreciated the support of my parents, my grandparents, and, particularly, my dear Christine.

Trondheim, November 2005
Torbjørn Helvik

List of Publications

I. The Evolution of n -player Cooperation — Threshold Games and ESS Bifurcations

Lars A. Bach, Torbjørn Helvik and Freddy B. Christiansen (2005), *Journal of Theoretical Biology*, In press (doi:10.1016/j.jtbi.2005.06.007).

II. Higher Order Cellular Automata

Nils A. Baas and Torbjørn Helvik (2005), *Advances in Complex Systems*, Vol. 8, Nos. 2–3, pp. 169–192.

III. Local Information in One-Dimensional Cellular Automata

Torbjørn Helvik, Kristian Lindgren and Mats G. Nordahl (2004), in *Cellular Automata: ACRI 2004 proceedings*, Vol. 3305 of *Lecture Notes in Computer Science*, eds. P. M. A. Soot, B. Chopard and A. G. Hoekstra, Springer-Verlag, Heidelberg, pp. 121–130.

IV. Continuity of Information Transport in Surjective Cellular Automata

Torbjørn Helvik, Kristian Lindgren and Mats G. Nordahl (2005), Submitted for publication.

Contents

1	Introduction	1
1.1	Background	1
1.2	Overview	2
1.3	Outline	6
2	Threshold games	7
2.1	Game theoretical background	7
2.1.1	Classical game theory	7
2.1.2	Evolutionary game theory	8
2.1.3	Multi-player evolutionary games	10
2.2	Threshold-type games	10
2.2.1	Public Goods games	10
2.2.2	The three-player threshold game	11
2.2.3	Games with general cost and benefit structures	13
2.2.4	General conclusions	13
2.2.5	Comments on spatial structure	14
2.2.6	Outlook	15
3	Cellular automata	17
3.1	Definition and basic results	18
3.1.1	Definition	18
3.1.2	Surjective and reversible CA	19
3.1.3	Measure-theoretic and ergodic properties	20
3.1.4	Conserved quantities	21
3.1.5	References to further aspects of CA theory	22
3.2	Coherent, propagating structures in CA	22
3.2.1	Space-time diagrams	22
3.2.2	Formal approaches	24
3.3	Applications of CA	24
3.3.1	Reversible CA with conserved quantities	25
3.3.2	Models of excitable media	26

3.3.3	Models of traffic flow	27
3.3.4	Applications in computer science	27
3.3.5	Artificial life	28
4	Higher order dynamical systems	29
4.1	Background and motivation	29
4.1.1	Hyperstructures	30
4.1.2	An <i>ansatz</i> for dynamical hierarchies	31
4.1.3	Other models	33
4.2	Higher order dynamical systems	33
4.3	Higher order cellular automata	35
4.3.1	2-Dynamics	35
4.3.2	2-Morphology	38
4.3.3	Experiments and results	39
4.3.4	Discussion and future directions	40
5	Local information and information transport	43
5.1	Background	44
5.1.1	Shannon entropy	44
5.1.2	Entropy and information in physics	45
5.1.3	Entropy conservation and reversibility	47
5.1.4	Microscopic entropy	47
5.1.5	Localized information	48
5.2	Local information	49
5.3	Information transport	50
5.3.1	The information current	50
5.3.2	Main results	52
5.3.3	Discussion	52
5.4	Correlation, structure and self-organization	53
5.5	Future developments	55
	Bibliography	57
	Paper I: The Evolution of n-player Cooperation — Threshold Games and ESS Bifurcations	
	Paper II: Higher Order Cellular Automata	
	Paper III: Local Information in One-Dimensional Cellular Automata	
	Paper IV: Continuity of Information Transport in Surjective Cellular Automata	

Chapter 1

Introduction

Virtually all of the major unsolved problems in science today concerns complex, self-organizing systems, where vast numbers of components interact simultaneously, with each shift in one agent influencing the other.

–Steven Strogatz, cited in *Nature*, Vol. 421, p. 782

1.1 Background

Current investigations in many parts of science, and in particular in biology, involve systems which are composed of a large number of interacting parts. These parts are often of different types, and both their internal dynamics and their coupling to other parts can be highly complex and nonlinear. However, even when the parts themselves are comparatively simple, their collective behaviour can be extremely complicated. A classical example is the human brain [138].

Due to the complexity of these systems, even remotely realistic mathematical models of them are not regarded as feasible. It should, however, be kept in mind that constructing overly detailed models rarely is the purpose of mathematical modelling [97]. Rather, the goal when modelling a physical system, process or phenomenon is to find a model that is simple enough to be analysed, yet retains enough of the essential mechanisms of the physical system to provide insights and predictions about it. Furthermore, also “toy models” — illustrative models not originating from any particular systems — can provide insight into physics and biology. For instance, through such models one can show that a particular set of mechanisms is sufficient to generate some given phenomenon [139], such as spiral waves in a homogenous medium [169] or the evolution of cooperation in animal populations [3].

There exist many useful mathematical models of systems with locally interacting units. This includes spatial stochastic processes such as interacting particle systems [96], statistical mechanics models such as the Ising model [74], as well as continuous and discrete dynamical systems such as coupled oscillators [85] and cellular automata [48, 167]. Often, these models are surprisingly simple, having only one type of interaction, a regular interaction topology such as a lattice, and a single type of unit having a few internal states. The reason for this simplicity is at least twofold. Firstly, even such simple mathematical systems can exhibit exceedingly complex behaviour, and generate complicated, global structures from random, non-structured initial configurations [14, 31, 48, 158, 168]. Such phenomena are often called self-organization [2, 43]. Secondly, the spatial aspect and possibly nonlinear interactions often make even quite basic models hard to analyse, and their long-term behaviour sometimes impossible to determine [168]. In order to improve the understanding of systems with interacting units, both new classes of models and techniques for their study are needed.

1.2 Overview

This thesis is concerned with dynamical systems of interacting units. A dynamical system is a pair consisting of a state space, with each point representing a state of the system, and a fixed rule governing the time evolution of these points. The time can be continuous or proceed in discrete steps.

The thesis is constituted by four papers presenting the main results of the research as well as an introductory part consisting of five chapters. The introductory part provides background on method and related work, summarizes the main results and contains discussions about implications, limitations and possibilities for further work, see Sect. 1.3. Below, we provide a non-technical overview of the problems addressed in the papers and a summary of our main results and contributions.

Paper I

A motivating question for the work presented in this paper is how one can model and analyse situations where there is competition both between different groups of individuals and within each group. This is a common situation in both human societies and biological systems. Typically, the performance of a group depends on the willingness of group members to cooperate. However, for each individual within the group there is a temptation to “free-ride” — to have the rest of the group do the work and still reap the benefits of the effort. Game theory is the natural framework in which to formalize and study such social dilemmas.

In Paper I we study a simplified situation where there is no direct conflict between groups. Rather, a similar effect is obtained by introducing a threshold on the number of cooperators needed to earn the group a reward. Our basic objectives are:

- Investigation of the prerequisites for cooperation in such situations, and the dependence on two parameters: the magnitude of the reward and the personal cost of cooperation.
- To establish general mathematical results facilitating the analysis of situations where cooperative and competitive relations collide.

To model the situation, we construct a three-player game which is an instance of a general class of games we call threshold games. We analyse the dynamics of the game using standard concepts from evolutionary game theory. In particular, this includes the replicator equation, which is a continuous dynamical system governing the evolution of the fraction of cooperators in a large population. Our main findings and results are summarized below.

- (1) We show that threshold games feature bifurcations. This means that small changes in the parameter values can lead to catastrophic events, such as a sudden eradication of cooperation.
- (2) In connection with population biology, we demonstrate that fragmentation of habitats may facilitate cooperation.
- (3) We prove a general result which provides a scheme for calculating the stable states of a game with n players for any cost and reward structure.
- (4) We illustrate that in contrast with two-player games, n -player games can have rich structures of stable states and bifurcations. This has implications for modelling.

A possible ecological implication of our findings is that minor environmental changes might have particularly large effects on species dependent on individually inaccessible and aggregated resources, such as large prey items.

Paper II

Consider again the question about how to analyse situations where there is competition both between groups and within each group. This is an example of a situation where:

- (A) There are two structural levels of units, such that each unit at the second level is an aggregation of units at the level below.
- (B) There are interactions at both levels.

Systems which fit these two criteria, possibly with more than two levels, are ubiquitous at all scales in biology and in human societies as well as in technical systems. In biology, this hierarchical organization principle leads to systems which are stable, reliable, and adaptable, yet never in equilibrium. This provides a strong

impetus for the investigation of hierarchical dynamics by general mathematical models, in order to elicit the source of this advanced functionality. However, little research has been conducted in this area.

In Paper II we introduce several new concepts and principles for describing and studying hierarchical, multilevel dynamics. The concepts are based on Baas's general hyperstructure framework. We refer to systems constructed using these principles as higher order dynamical systems. A working hypothesis is that our higher order concepts enables the construction of dynamical systems with more advanced functionality. The objectives of the paper are:

- To outline the principles behind higher order dynamical systems.
- To illustrate the concepts by providing a detailed formalism for a class of higher order dynamical systems based on cellular automata (CA), which are discrete systems consisting of a lattice of locally interacting "cells" with internal states.
- Investigation of the scope for interesting and novel behaviour found in these systems.

The spatio-temporal phenomena we are interested in are hard to discover using analytic tools, due to the nonlinearity and discreteness of the system. Hence, the primary method used for investigation of higher order CA in Paper II is experiments by computer simulations. In addition to the new conceptual framework, the main contributions of the paper can be summarized as follows.

- (1) We present several examples of interesting phenomena and patterns of behaviour found in higher order CA.
- (2) We demonstrate the applicability of the principles to modelling.
- (3) Based on our observations, we identify several fundamental issues pertaining to the nature of higher order systems.

Paper III

A feature which we have observed widely in higher order CA, and which also is commonly observed in ordinary CA, is the emergence of persistent particle-like "objects" which can traverse the lattice and sometimes interact. The quintessential example of such objects is the Gliders found in the well-known "Game of Life" CA. An essential characteristic of these phenomena is the locality with respect to the lattice. However, in dynamical systems theory the entire lattice configuration is treated as a single point in state space. Consequently, traditional concepts from this discipline are not suitable in order to characterize and analyse the emergence and dynamics of particle-like objects. Rather, concepts that are genuinely local seems required. This is one of the motivations for the work presented in Paper III.

In Paper III we introduce and study local information (LI), which is a function associating an information quantity to each point of the one-dimensional lattice.

Concepts such as information and entropy have been widely used to characterize correlation and structure in lattice systems, but hitherto almost exclusively in a global sense. An important property of information is that it provides a universal approach to describing different systems. Additionally, it is a conserved quantity in important cases.

Another motivation for the investigation of LI is to provide a more solid foundation for the use of information based concepts in various physical systems. A fundamental question is to what extent information can be considered as a local quantity in physical systems. Cellular automata are relevant in this context, because these systems embody many of the characteristics of microscopic physical systems, and have been used to explore the relation between the microscopic and the macroscopic level of description.

The objectives of the paper are:

- To introduce an information quantity associated with each lattice point.
- To study the behaviour of this quantity under CA dynamics.
- To illustrate the applicability to characterizing self-organization, such as the emergence of particle-like objects in CA.

In this paper we focus the formal investigations on reversible CA, which is an important class for modelling, since microscopic physics is reversible. An important question regarding the interpretation of information as a physical quantity is whether local information is a locally conserved quantity under iteration by suitable CA. When this is the case, information that disappears from one cell appears nearby in the next time step.

The main contributions of the paper are:

- (1) We introduce LI, and argue that LI is a valid and natural measure of the information content of a single cell of the lattice.
- (2) Through examples and numerical approximations we demonstrate the applicability to characterizing and detecting particle-like phenomena in CA.
- (3) We sketch a proof showing that LI is locally conserved for reversible CA.

A detailed proof and further results are found in Paper IV.

Paper IV

In Paper IV we further investigate the mathematical properties of local information and transport of LI in cellular automata. We focus on the class of surjective one-dimensional CA, which comprises exactly those CA for which the Shannon entropy always is globally conserved. The class includes reversible CA as a special case. The main objectives of the paper are:

- To investigate whether LI is a locally conserved quantity for surjective CA.
- To establish rigorous results on the maximum speed of propagation of information, as well as other properties of information flow for various classes of CA.

The main results are:

- (1) We prove that local information is locally conserved for any surjective CA, under a condition on spatial homogeneity.
- (2) We provide and argue for a definition of an information current. The current gives the flow rate of LI between each pair of neighbouring cells.
- (3) We find bounds on the maximum transport rate of information in the lattice, both on average and locally.

Our results suggest that LI may be a useful and important concept for understanding the dynamical basis of the self-organization observed in higher order CA, as well as in ordinary CA. This includes both detection of structure by numerical means and rigorous quantification of this structure. There is also ample opportunity for rigorous investigation of local information for further classes of CA and for other dynamical systems of interacting units.

1.3 Outline

Chapter 2 gives an overview of relevant concepts from evolutionary game theory, presents our threshold games, and provides a discussion of results, implications, and spatial aspects. In Chapter 3, we define cellular automata and discuss their mathematical properties as well as the variety of particle-like objects found in CA. We also give an overview of important applications. Chapter 4 concerns higher order dynamical systems, and in particular higher order cellular automata. We define this class of dynamical systems and summarize the results of our investigations. The first part of Chapter 5 is a review of the role of information in dynamical systems theory and in physics. We then introduce our local information measure, state our main results on information transport in CA, and discuss applications to the characterization of spatio-temporal structure in lattice systems. Finally, Papers I to IV are included after the bibliography.

Chapter 2

Threshold games

The theory of many person games may seem to stand to that of two-person games in the relation of sea-sickness to a headache.

–William D. Hamilton [62, p. 151]

Outline

Section 2.1 introduces the basic concepts and results from evolutionary game theory applied in Paper I. This includes previous work on multi-player games. In Section 2.2, we first discuss Public Goods games and their role in studying the evolution of cooperation. We then introduce one of the main topics of Paper I, namely our three-player threshold game, and consider its relation to Public Goods games. We discuss the novel behaviour found in threshold games as well as generalizations and possible ecological implications.

2.1 Game theoretical background

2.1.1 Classical game theory

Game theory can be defined as the study of mathematical models of conflict and cooperation between intelligent rational decision-makers [114]. Modern game theory is often said to originate with the seminal book of von Neumann and Morgenstern [162]. The term *game* refers to any social situation involving two or more parties. The parties involved in a game are called *players*. Each player has a set of possible strategies and the *utility*, or *payoff*, of each player depends on the strategies

implemented by all players. A game is most conveniently represented in the *normal form* [162]:

Definition 2.1. A *normal-form game* Γ is a collection

$$\Gamma = (N, (S_i)_{i \in N}, (u_i)_{i \in N}),$$

where N is a nonempty set, each S_i is a nonempty set, and each u_i is a function

$$u_i : \prod_{j \in N} S_j \rightarrow \mathbb{R}.$$

The set N represents the players and S_i is the collection of *pure strategies* player i can select from. The function u_i gives the resulting payoff to player i as a function of the strategies selected by all players. In general, a player can employ a *mixed strategy*, which is to play $s_k \in S_i$ with probability p_k . For a finite game, a strategy x_i is then a point in the standard n -simplex Δ^n , with $n = |S_i| - 1$. A strategy is *strictly mixed* if at least two pure strategies have positive probabilities.

The fundamental solution concept in classical game theory is that of a *Nash equilibrium* [117]. A strategy profile

$$\mathbf{x} \in \prod_{i \in N} \Delta^{|S_i|-1}$$

is a Nash equilibrium if no player can improve its payoff by altering its strategy, given that the strategies of all other players remain fixed. Each finite game has at least one Nash equilibrium, possibly involving mixed strategies [117]. As an example, the unique Nash equilibrium of the game of rock-paper-scissors is for both players to use each of the three options with equal probability. Note that the use of the Nash equilibrium as a solution concept is based on the fundamental assumption that all players are rational and have unlimited computational resources. The validity of this assumption is questionable for many interactions among humans and certainly inadequate for interaction among animals.

2.1.2 Evolutionary game theory

Evolutionary game theory came into existence with the work of Maynard-Smith and Price [109]. Their goal was to explain why conflicts among animals from species which possess lethal weapons often are settled by displays rather than all-out fighting. In classical game theory players have strategy sets from which they choose. In evolutionary game theory, however, *species* have strategy sets, and each individual uses a fixed strategy in all interactions. This strategy, or trait, is assumed to be imprinted in the genes of the individual, and the success of the strategy determines the individual's reproductive potential. Thus, rationality is irrelevant in evolutionary game theory. The interaction between two individuals is described by a symmetric

two-player normal-form game, called the *stage game*. See [72] and [164] for general introductions to evolutionary game theory.

There are two central concepts in evolutionary game theory. The first is that of an *evolutionarily stable strategy* (ESS), which is a strategy stable against invasion by any mutation. Let the stage game have n strategies, and define $u(x; y)$ as the expected payoff to a player using the mixed strategy $x \in \Delta^{n-1}$ against a player using the mixed strategy $y \in \Delta^{n-1}$.¹ Then x is an ESS if [108, 109]:

$$u(x; x) \geq u(y; x) \quad \forall y \in \Delta^{n-1} \quad (\text{Nash equilibrium}), \quad (2.1a)$$

$$u(y; x) = u(x; x) \Rightarrow u(x; y) > u(y; y) \quad \forall y \neq x \quad (\text{stability condition}). \quad (2.1b)$$

Thus, x is a best reply against itself, and if y is an alternative best reply, then x is better against y than y is against y . A game can have more than one ESS, but there is a strong restriction. For a mixed strategy $x \in \Delta^n$, define the *support* $\text{supp}(x)$ of x as the set of pure strategies played with non-zero probability. The Bishop-Cannings theorem states that if x and y both are ESSs for a two-player game, then $\text{supp}(x) \not\subseteq \text{supp}(y)$ [15].

The second central concept is the *replicators dynamics*, which yields a dynamical system governing the evolution of the distribution of pure strategies in the population. The basic mechanism is that pairs of individuals repeatedly are drawn at random from a large population to play the stage game. Each individual reproduces with a rate proportional to its average payoff resulting from these confrontations. Let x_i denote the fraction of the population playing the pure strategy i , and let $x(t) = (x_1(t), \dots, x_n(t))$ be the state of the population. The state can be considered as the mixed strategy employed by the *population* as a whole. The replicator equation obtained from the above description is given by [153]:

$$\dot{x}_i = x_i \left(u_i(x) - \sum_{j=1}^n x_j u_j(x) \right), \quad i = 1, \dots, n, \quad (2.2)$$

where $u_i(x)$ is the average payoff resulting from always playing strategy i . If x is an ESS, then x is an asymptotically stable equilibrium point of the replicator equation [71].

Evolutionary game theory has been widely applied in modelling the natural world. In addition to the replicator dynamics, also other game dynamics, not necessarily involving reproduction, have been constructed. For instance, in the imitation dynamics described by Weibull [164] strategies are transmitted within the population through imitation of more successful individuals. Such alternative dynamics facilitate the use of evolutionary game theory in other areas, such as social science.

¹Equivalently, $u(x; y)$ is the expected payoff of a random individual drawn from a population in which only pure strategies are played and with a distribution given by x , against a random individual from a population with pure strategy distribution given by y .

2.1.3 Multi-player evolutionary games

When an interaction involves more than two individuals, the stage game modelling it can be chosen as a symmetric N -player game. However, such games have rarely been used in the ecological context. Rather, interactions involving more than two individuals are commonly either modelled as a set, or tournament, of two-player games [3, 26] or as a “playing-the-field”-game such as the sex ratio game of Maynard-Smith [108], where the entire population plays in the same game. Some notable exceptions where proper multi-player games are constructed are [20], [60] and [113].

In an N -player evolutionary game a number N of individuals are repeatedly drawn at random from the population to play the stage game. The notions of ESS and replicator dynamics were first generalized to such N -player games by Palm [124]. However, only recently have systematic studies of N -player evolutionary games been conducted [25, 27]. The notion of an ESS for a symmetric N -player game is a straightforward extension of the two-player case. Denote by $u(x; y)$ the expected payoff to an individual playing the mixed strategy $x \in \Delta^{n-1}$ when the other $N - 1$ players play the mixed strategy y . Then $x \in \Delta^{n-1}$ is an ESS if

$$u(x; \varepsilon y + (1 - \varepsilon)x) > u(y; \varepsilon y + (1 - \varepsilon)x) \quad (2.3)$$

for all $y \in \Delta^{n-1}$ and sufficiently small $\varepsilon = \varepsilon(y)$. This is equivalent to requiring $u(x; y) > u(y; y)$ for all y in some neighbourhood of x , except at $y = x$ [27]. The replicator equation for an N -player game has a form identical to (2.2). Bukowski and Miękisz [27] have proved the following useful result:

Theorem 2.1. *In a symmetric N -player game with two strategies, x is an ESS if and only if x is an asymptotically stable equilibrium point of the replicator dynamics.*

Broom et al. [25] observed that for N -player games with $N > 2$, the Bishop-Cannings theorem is no longer valid. For a three-player game, the support of an ESS can be contained in the support of another ESS, and for four-player games two ESSs can have identical supports.

2.2 Threshold-type games

2.2.1 Public Goods games

In both natural and social systems the actions of each single individual often affect the benefits acquired by all individuals. There are numerous situations where so-called free-riders take advantage of others cooperating for a common good [66]. A much studied game theoretic model of this situation is the N -player Prisoner’s Dilemma game, or *Public Goods game* [78]. In this game, each of the players has the option to contribute some resource c to a common pool. The pool is then increased by some factor $p < N$, and redistributed equally among all players. The outcome is tragic —

the only ESS is to contribute nothing to the common pool. However, cooperation can be promoted by introducing other mechanism such as punishment [21], the need to maintain a good reputation [110], or optional participation [69].

Motro [113] has studied some extensions of the Public Goods game with reference to biological situations. He considered a group of N players performing a common task which involves cooperation, such as food foraging in a patchy environment. Each player has two strategies: cooperate (C) or defect (D). All players achieve some benefit $f(m)$, which only depends on the number m of cooperators. However, cooperation also inflicts an additional, personal, cost c , for instance due to increased risk of predation in the food foraging scenario. Thus, the total payoffs are $f(m)$ for defectors and $f(m) - c$ for cooperators. Motro analysed this game for three classes of benefit functions f : linear, convex and concave. He found that an ESS consisting of a mix of both strategies only can exist for a concave benefit function. These types of games are called N -player Snowdrift games in [38]. Note that the standard Public Goods game fits Motro's framework with a linear f .

2.2.2 The three-player threshold game

One motivation for the work presented in Paper I is that in various natural situations a benefit function with constant second (discrete) derivative may fail to capture the nature of the interaction. Therefore, a more general analysis would be desirable. In particular, one can consider situations where obtaining some advantage is an all-or-nothing event. A natural situation exemplifying this is the group hunting of certain predators, such as the African wild dog, where the prey is either caught or not. The collective effort of individuals in the groups ensures that large prey can be caught that no single individual could ever capture [50]. A minimum number of hunters seems required in order to capture large prey items as the fatigue of the prey cannot be provoked without a joint effort of a number of individuals [34]. Another situation where obtaining the benefit is an all-or-nothing event is a competition among two or more groups where only the winning group receives a reward.

In Paper I we introduce a three-player threshold game as a model of the situation where obtaining the benefit is a binary event and where a minimum number of cooperators is required for success. In the game, each player has the choice of cooperating (C) or defecting (D). Cooperation incurs a cost c , and all three players obtain a reward of $r > c$ if a threshold of two cooperators is met. The payoff structure is summarized in Table 2.1. The benefit function f is in this case given by $f(0) = f(1) = 0$ and $f(2) = f(3) = r$. This game contrasts with the Public Goods game in one important aspect: there is one situation where cooperation is preferable to defection, namely if exactly one other player cooperates. As a consequence, a population with a certain fraction of cooperators will be stable if the cost is not too high. We show this by constructing the replicator dynamics.

Denote by x the fraction of cooperators in the population, and define the

	CC	CD	DD
C	$r - c$	$r - c$	$-c$
D	r	0	0

Table 2.1: The payoff to ego as a function of the strategy employed and the strategies of the two other players in the three-player threshold game.

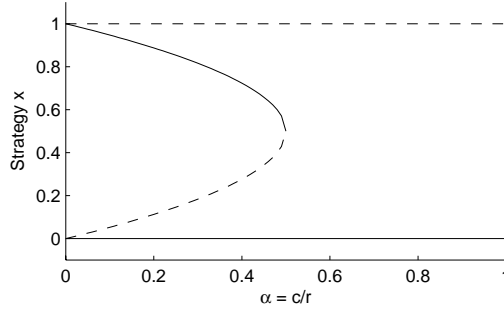


Fig. 2.1: The bifurcation diagram for the replicator system (2.4). Solid-drawn lines represent stable equilibrium points and dashed lines unstable equilibrium points.

dimensionless variable $\alpha = \frac{c}{r}$. The replicator equation for the system is, with a suitable time scale, given by

$$\dot{x} = x(1-x)(2x(1-x) - \alpha). \quad (2.4)$$

It follows from Theorem 2.1 that the ESSs of the game are exactly the asymptotically stable equilibrium points of the replicator dynamics. A bifurcation diagram showing the location of all stable and unstable equilibrium points as a function of α is shown in Fig. 2.1. A striking feature of the system is the existence of a saddle-node bifurcation at $\alpha = \alpha^* = \frac{1}{2}$. For $\alpha < \alpha^*$, the game has two ESSs: a mixed strategy with a fraction $x^+ = \frac{1}{2} + \frac{1}{2}\sqrt{1-2\alpha}$ of cooperators, and the pure strategy with only defectors. Their basins of attraction are separated by the unstable equilibrium point $x^- = \frac{1}{2} - \frac{1}{2}\sqrt{1-2\alpha}$. For $\alpha > \alpha^*$, always defecting is the only ESS.

Consider a situation where the cost and reward parameters are non-constant in time. If the replicator dynamics has a significantly faster time scale than the changes in c and r , we can assume that the state $x(t)$ of the system always is an ESS. This means that if α is slightly less than α^* and the system is in the mixed ESS, then small changes in c or r can cause a catastrophic event as α increases beyond α^* and cooperation vanishes entirely. A subsequent reduction of α will not suffice to reestablish cooperation, since a population consisting of only defectors also is an ESS

for the species. Even when α is returned to a value lower than α^* before $x(t)$ reaches 0, the system will still end up in the basin of attraction of the defective state, and cooperation will vanish.

2.2.3 Games with general cost and benefit structures

In Paper I we also generalize the analysis, and describe how to calculate the ESSs of an N -player game with a general benefit function and a cost of cooperation possibly depending on the number of cooperators.

Definition 2.2. *Let the game $\Gamma(N, S, \mathbf{r}, \mathbf{c})$, with $\mathbf{c} = (c_1, \dots, c_N) \in [0, \infty)^N$ and $\mathbf{r} = (r_0, r_1, \dots, r_N) \in [0, \infty)^{N+1}$, have N players, each with the strategy set $S = \{C, D\}$. The payoff is defined as following: If k players play C , these will obtain $r_k - c_k$ and the remaining $N - k$ players will obtain r_k .*

Let Δ denote the forward difference operator. The mixed ESSs are found to be the roots of the polynomial

$$g_{\mathbf{r}, \mathbf{c}}(x) = \sum_{k=0}^{N-1} \binom{N-1}{k} (\Delta r_k - c_{k+1}) x^k (1-x)^{N-1-k}, \quad (2.5)$$

also satisfying $\frac{d}{dx} g_{\mathbf{r}, \mathbf{c}}(x) < 0$. This allows the construction of evolutionary games with rich structures of internal saddle-node bifurcations where pairs consisting of an ESS and an unstable equilibrium point are created.

As an example, consider a six-player game with two thresholds, one at two cooperators and one at five cooperators. Let \mathbf{r} be given by $r_0 = r_1 = 0$, $r_2 = r_3 = r_4 = 1.4$, and $r_5 = r_6 = 2.4$, and let the cost be c for any number of cooperators. The function $g_{\mathbf{r}, c}(x)$ from (2.5), with c denoting the vector (c, c, \dots, c) , is given by

$$g_{\mathbf{r}, c}(x) = 5x(1-x)(x^3 + 1.4(1-x)^3) - c. \quad (2.6)$$

Using the results of Paper I, it is straightforward to construct a diagram showing the ESSs of the game for each value of the cost c , see Fig. 2.2. For costs in the approximate interval $0.363 < c < 0.420$ the system has two mixed ESSs, which we can denote by x_H and $x_L < x_H$. If the state of the system is x_H and the cost increases beyond 0.420, the degree of cooperation will decrease and the system will settle in x_L . State x_H will not be reinstalled if the cost subsequently decreases below 0.420 — to obtain this the cost has to be reduced below 0.363.

2.2.4 General conclusions

The three-player threshold game and games with more general benefit functions and costs are natural generalizations of N -player Public Goods games obtained by relaxing the requirement of a linear benefit function. An N -player Public Goods

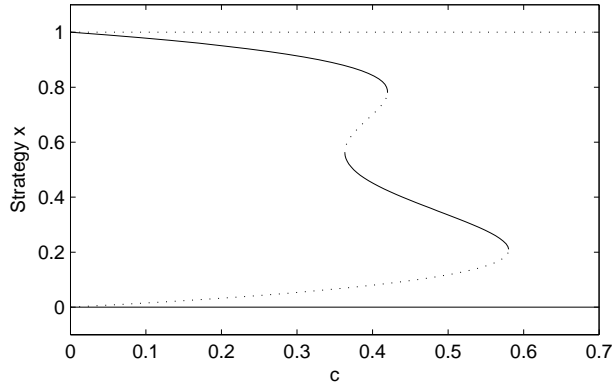


Fig. 2.2: The bifurcation diagram of the six-player double threshold game with $g_{r,c}(x)$ in (2.6). Solid lines are ESSs and dotted lines are unstable equilibrium points of the replicator dynamics.

game only has one ESS, namely a pure population of defectors. In contrast, our examples of threshold games have shown rich structures of mixed ESSs and ESS bifurcations. If the parameters defining the game are allowed to vary at a slow time scale, catastrophic events can occur. Here, the fraction of cooperators in the population shows sudden transitions among discrete plateaus, see Figs. 2.1 and 2.2.

In natural ecological systems, fluctuations in environmental conditions makes variations in the cost and benefits associated with cooperative behaviour commonplace. In light of our results for the threshold game and related games, species depending on collaborative endeavours can be prone to sudden drastic events, such as altered behavioural characteristics or even extinctions. This description will for instance apply to group foragers that depend on individually inaccessible and aggregated resources, such as large prey items.

Broom et al. [25] showed that the Bishop-Cannings theorem no longer is valid for games with more than two players. Our analysis of the threshold game demonstrates the practical importance of this result. For a two-player game with two strategies the coexistence of two ESSs are possible, but both must be pure strategies. Games with N players can however have several mixed ESSs for the same parameter values and hence a much richer dynamics.

2.2.5 Comments on spatial structure

Our basic threshold game model is a mean field model which does not take spatial aspects into account. In natural systems, however, there is often some spatial stability resulting in local and repeated interactions. In such settings several mechanisms

are known to facilitate cooperation, including direct reciprocity [3, 157], indirect reciprocity due to reputation [94], punishment [21], and kin selection [61]. Therefore, it would be interesting in the future to study versions of the threshold game where spatial structure and repeated interactions are taken into account. There are several ways in which the threshold game could be modified to obtain this. One option is to use the replicator equation, but to introduce a probability w of the group persisting from one interaction to the next, and augment the set of possible strategies to include reactive strategies, see, e.g., [21]. A different approach is to consider spatial games where individuals occupy the points of a lattice [101, 121, 122]. Each individual plays repeatedly against its four neighbours, and successful strategies propagate either through invasion, imitation, or learning. With deterministic updating, the model is a two dimensional cellular automaton, which is a class of dynamical systems considered in Chap. 3. Spatial Public Good games have been considered by Brandt et al. [23]. We have conducted preliminary studies of a spatial version of the basic threshold game, and obtained promising results. In particular, the system has exhibited a wide range of behaviours for different parameter values.

Space may also be introduced indirectly by considering metapopulation models [63, 95, 143]. These are applicable when there are several subpopulations in discrete patches due to a fragmented habitat. The dynamics within each patch is treated separately, but with occasional dispersal though extinctions and subsequent invasions from other patches. In Paper I, we discuss how such dynamics can facilitate the propagation of cooperative behaviour because subpopulations in which cooperation is prominent have greater emigration potential.

2.2.6 Outlook

The basic threshold game is quite simple, yet embodies two fundamental principles, namely simultaneous interaction among multiple players and a nonlinear relationship between the degree of cooperation and the resulting benefit. Therefore, it may serve as a prototypical model on which to base further investigations of the scope for cooperation in various situations. Threshold games may, for instance, be a suitable setting in which to explore the link between group size and cooperation [19, 134]. We are currently investigating the connections between group size, threshold size and average payoff in the population.

The work presented in Paper I grew out of discussions about game theoretical modelling of situations with competition both within groups and between groups. Such models would be natural instances of higher order dynamical systems, which is the topic of Chap. 4. Investigation in this direction is left as a future problem, but we give some further indications about the relation between threshold games and higher order systems at the end of Chap. 4.

Chapter 3

Cellular automata

*Teknikken kan skabe
en verden som klapper,
når menneskene bare
vil trykke på knapper.*

*Og nu er der endnu
en nyhed på trapperne:
en automat
som kan trykke på knapperne.*

–Piet Hein

Outline

This chapter provides background material on the dynamical systems known as cellular automata. These are central in Papers II, III and IV. In Section 3.1, we define cellular automata, introduce notation used in the subsequent chapters, and review the main results relevant for our work. Section 3.2 focuses on a key aspect of cellular automata phenomenology, namely coherent localized structures. In Section 3.3, we illustrate the relevance of cellular automata to physics and biology through several examples of applications.

3.1 Definition and basic results

3.1.1 Definition

Let \mathcal{A} be some finite alphabet, and let \mathbb{Z}^d be the d -dimensional lattice. A *configuration* of the lattice is a map $x: \mathbb{Z}^d \rightarrow \mathcal{A}$ assigning a *state* to each lattice point. Denote the state at position v of \mathbb{Z}^d by x_v . For $A \subset \mathbb{Z}^d$, let x_A be the configuration x restricted to A . Equip \mathcal{A} with the discrete topology, and the space $\mathcal{A}^{\mathbb{Z}^d}$ of all configurations of the lattice with the corresponding product topology. This makes $\mathcal{A}^{\mathbb{Z}^d}$ a compact metric space. A metric generating the topology is given by

$$d(x, y) = 2^{-n}, \quad \text{where } n = \max\{m \geq 0 : \|v\|_\infty < m \Rightarrow x_v = y_v\}.$$

The action of \mathbb{Z}^d on itself by translation induces a *shift action* on the configuration space $\mathcal{A}^{\mathbb{Z}^d}$. The shift by v is defined by $\sigma^v(x)_u = x_{v+u}$.

Definition 3.1. A d -dimensional cellular automaton (CA) is a continuous map $F: \mathcal{A}^{\mathbb{Z}^d} \rightarrow \mathcal{A}^{\mathbb{Z}^d}$ that commutes with all shifts.

That F commutes with all shifts means that $F \circ \sigma^v = \sigma^v \circ F$ for all $v \in \mathbb{Z}^d$. One can generalize the definition of a CA by allowing any monoid \mathbb{M} in place of \mathbb{Z}^d [129] or by allowing any subshift of finite type in place of $\mathcal{A}^{\mathbb{Z}^d}$ (as advocated by Boyle and Kitchens [22]). The pair $(F, \mathcal{A}^{\mathbb{Z}^d})$ is a topological dynamical system.

Write the cardinality of a set A as $|A|$. The following theorem is a simple yet fundamental result about CA first proved by Hedlund [70].

Theorem 3.1. The map $F: \mathcal{A}^{\mathbb{Z}^d} \rightarrow \mathcal{A}^{\mathbb{Z}^d}$ is a CA if and only if there exists a finite set $N \subset \mathbb{Z}^d$ and a map $f: \mathcal{A}^{|N|} \rightarrow \mathcal{A}$ such that $F(x)_v = f(\sigma^v(x)_N)$.

The theorem follows from the result that any continuous map on a compact space is uniformly continuous. We will usually define the cellular automata we study by stating the pair N and f . The map f is called the *block map* or *rule* of the CA, and N is called the *neighbourhood*. The elements of the lattice \mathbb{Z}^d are traditionally called *cells*. See Fig. 3.1 for an illustration.

In light of Theorem 3.1, the dynamics of a CA can be regarded as local interaction among the cells. The interaction is defined by the rule f , and all cells update their states synchronously. This property makes CA particularly suitable for describing processes consisting of a large collection of simple, locally interacting units.

We now introduce some notation for *one-dimensional CA*. Since the lattice points are the integers in this case, we denote them by the symbols $i, j \in \mathbb{Z}$, rather than by v and u . It is convenient to choose the neighbourhood N of a one-dimensional CA as the smallest possible interval $\{-l, \dots, r\}$ containing 0. The non-negative integers l and r are called the right and left radii of the CA respectively. For $I = \{i, i+1, \dots, j\}$, we

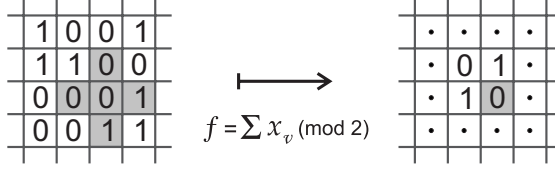


Fig. 3.1: An illustration of a two-dimensional CA with $\mathcal{A} = \{0, 1\}$, neighbourhood $N = \{v : \|v\|_1 \leq 1\}$, and block map f given by $f(x_N) = \sum_{v \in N} x_v \pmod{2}$. The states of all cells are updated synchronously. A small part of the lattice is shown.

write $x_I = x_i^j$, and for $J = \{i, i + 1, \dots\}$, we write $x_J = x_i^\infty$. For any $n \geq 1$, the block map f can be extended in a natural way to a map $f_n : \mathcal{A}^{l+r+n+1} \rightarrow \mathcal{A}^{n+1}$ by putting

$$f_n(x_{-l}^{r+n}) = (f(x_{-l}^r), f(x_{-l+1}^{r+1}), \dots, f(x_{n-l}^{r+n})). \tag{3.1}$$

We will omit the subscript n , and write f for the block map applied to a block of any length. The inverse block map f^{-1} is defined by

$$f^{-1}(y_1^n) = \{x_{1-l}^{n+r} \in \mathcal{A}^{n+l+r} : f(x_{1-l}^{n+r}) = y_1^n\}. \tag{3.2}$$

We denote the set of all non-empty blocks of symbols from \mathcal{A} by \mathcal{A}^+ . For $a \in \mathcal{A}$, a^∞ denotes the sequence $x \in \mathcal{A}^{\mathbb{Z}}$ with $x_i = a$ for all i .

Wolfram [167] has introduced a convenient way of enumerating the set of all one-dimensional CA with $\mathcal{A} = \{0, 1, \dots, k - 1\}$ and $N = \{-l, \dots, r\}$. For each $0 \leq m < k^{l+r+1}$, let $\varphi_m \in \mathcal{A}^{l+r+1}$ be the k -ary representation of m . For instance, with $l = r = 1$ and $k = 2$, we have $\varphi_3 = 011$. The rule number R_f of the rule $f : \mathcal{A}^{l+r+1} \rightarrow \mathcal{A}$ is

$$R_f = \sum_{m=0}^{k^{l+r+1}-1} f(\varphi_m) \cdot k^m. \tag{3.3}$$

We will use this rule numbering systems when referring to the 256 *elementary CA* [167], which are the CA on $\{0, 1\}^{\mathbb{Z}}$ with neighbourhood $N = \{-1, 0, 1\}$.

3.1.2 Surjective and reversible CA

A cellular automaton F is *surjective* if for each $y \in \mathcal{A}^{\mathbb{Z}^d}$ there exists an $x \in \mathcal{A}^{\mathbb{Z}^d}$ such that $F(x) = y$. In this case, x is called a preimage of y . A CA is *injective* if each $y \in \mathcal{A}^{\mathbb{Z}^d}$ has at most one preimage, and *bijective* if each $y \in \mathcal{A}^{\mathbb{Z}^d}$ has exactly one preimage. Surjective CA will be central in Papers III and IV, since these are the CA that preserve information in a global sense. Reference [39] reviews the known relationships among injectivity, surjectivity and bijectivity of a CA F and of F when restricted either to the

set of all periodic configurations or all configurations that are finite with respect to some quiescent state.¹

The following fundamental result for one-dimensional CA was proved by Hedlund [70].

Theorem 3.2. *Let F be a one-dimensional CA with block map $f : \mathcal{A}^{l+r+1} \rightarrow \mathcal{A}$. The following are equivalent:*

- (1) F is surjective.
- (2) For all $x_1^n \in \mathcal{A}^+$, we have $|f^{-1}(x_1^n)| = |\mathcal{A}|^{l+r}$.
- (3) For all $x \in \mathcal{A}^{\mathbb{Z}}$, we have $|F^{-1}(x)| \leq |\mathcal{A}|^{l+r}$.
- (4) For all $x \in \mathcal{A}^{\mathbb{Z}}$, the set $F^{-1}(x)$ is countable.

From the equivalence of (1) and (3) in the theorem, it follows that any injective one-dimensional CA also is surjective, and hence bijective. This is true in any dimension d [112, 115]. Moreover, the inverse map of any bijective CA is also a CA. This was first noted by Richardson [132], and follows since the inverse of any continuous map on a compact space is continuous. For this reason, bijective CA are often called *reversible*. Deciding whether a given CA is surjective or reversible is not straight-forward, nor is finding the inverse CA. An efficient algorithm for one-dimensional CA which uses finite automata is described in [149]. In dimensions $d \geq 2$, surjectivity and reversibility are undecidable properties [80].

The seminal paper by Hedlund [70] contains many further useful results on surjective and reversible CA. See [156] for a review of reversible CA.

3.1.3 Measure-theoretic and ergodic properties

The topology of $\mathcal{A}^{\mathbb{Z}}$ is generated by the the collection of all cylinder sets of the form $\text{Cyl}(a_i^{i+n}) = \{x \in \mathcal{A}^{\mathbb{Z}} : x_i^{i+n} = a_i^{i+n}\}$, where $i \in \mathbb{Z}$, $n \geq 0$, and all $a_k \in \mathcal{A}$. The collection of cylinder sets is also a sub-algebra that generates the Borel σ -algebra \mathcal{B} of $\mathcal{A}^{\mathbb{Z}}$. A *probability measure* μ on $(\mathcal{A}^{\mathbb{Z}}, \mathcal{B})$ is defined by assigning a probability $\mu(\text{Cyl}(a_i^{i+n}))$ to each cylinder set in a consistent way, see §27 in [126]. We will usually write this probability $\mu(a_i^{i+n})$, thus letting a_i^{i+n} represent both the symbol block of length $n+1$ and the cylinder set. The measure μ is *shift-invariant* if it satisfies $\mu(\sigma^{-1}(B)) = \mu(B)$ for all $B \in \mathcal{B}$, and *ergodic* if additionally all $B \in \mathcal{B}$ satisfying $\sigma^{-1}(B) = B$ have measure $\mu(B) = 0$ or $\mu(B) = 1$.

It is often convenient to consider the measure μ as defining a discrete stochastic process $(X_n)_{n \in \mathbb{Z}}$, $X_n \in \mathcal{A}$, with joint distributions given by $\text{Prob}(X_i^j = a_i^j) = \mu(a_i^j)$. In this case, μ is called the Kolmogorov measure of the process. A Bernoulli measure is a measure for which the coordinate random variables X_i are all independent and identically distributed. The uniform Bernoulli measure $\bar{\mu}$ gives equal probability to

¹A state $e \in \mathcal{A}$ is said to be quiescent with respect to F if $f(e, e, \dots, e) = e$.

each block, i.e. $\bar{\mu}(a_1^n) = |\mathcal{A}|^{-n}$ for all $a_1^n \in \mathcal{A}^n$. By a *uniformly random* configuration of \mathbb{Z} , we mean a configuration generated by the stochastic process associated with $\bar{\mu}$. The conditional probability $\mu(a_0 | a_{-n}^{-1}) = \mu(a_{-n}^0) / \mu(a_{-n}^{-1})$ is the probability that $X_0 = a_0$ given that $X_{-n}^{-1} = a_{-n}^{-1}$.

A cellular automaton F can be considered to act on measures. Let μ^0 be a probability measure on $\mathcal{A}^{\mathbb{Z}}$, and let $\mu^1 = F(\mu^0)$ be defined by $\mu^1(B) = \mu^0(F^{-1}(B))$ for each $B \in \mathcal{B}$. The block probabilities of μ^1 can be calculated from the relation

$$\mu^1(y_0^n) = \sum_{x_{-l}^{n+r} \in f^{-1}(y_0^n)} \mu^0(x_{-l}^{n+r}). \quad (3.4)$$

More generally, set $\mu^t = \mu^0 \circ F^{-t}$. A measure μ^0 is called F -invariant if $\mu^1 = \mu^0$. It follows from the Krylov-Bogolioubov Theorem [163, Cor. 6.9.1] that for every CA F there exists an F -invariant measure. However, non-trivial measures of this kind are in general hard to find. An invariant measure for surjective CA is given by the following corollary to Theorem 3.2.

Corollary 3.1. *The uniform Bernoulli measure is F -invariant if and only if F is surjective.*

For μ^0 a shift-invariant measure on $\mathcal{A}^{\mathbb{Z}}$, denote the metric entropy² of the shift σ on $(\mathcal{A}^{\mathbb{Z}}, \mathcal{B}, \mu^0)$ by $h(\mu^0)$. This quantity is often called the *spatial entropy* in the context of CA [167]. The spatial entropy is invariant under iteration by a surjective one-dimensional CA [99].

Theorem 3.3. *Let μ^0 be a shift-invariant probability measure, and $F : \mathcal{A}^{\mathbb{Z}} \rightarrow \mathcal{A}^{\mathbb{Z}}$ a surjective CA. Then, $h(\mu^1) = h(\mu^0)$.*

For non-surjective CA, $h(\mu^1) \leq h(\mu^0)$ [99], so the entropy is always non-increasing. When μ is F -invariant, also the metric entropy $h_\mu(F)$ of F is defined. In connection with CA, this is sometimes called the *temporal entropy*, to separate it from the spatial entropy. Directional entropies defined for arbitrary directions in the two-dimensional space-time lattice have been considered by Milnor [111].

3.1.4 Conserved quantities

Cellular automata that satisfy various conservation laws are important in applications. Pomeau [130] and Hattori and Takesue [68] have investigated CA with conserved quantities defined by a local density function $E : \mathcal{A}^m \rightarrow \mathbb{R}$. They defined E to be conserved by a CA F if for each $N \in \mathbb{N}$ and periodic configuration $x \in \mathcal{A}^{\mathbb{Z}}$ with period N , the function

$$\Phi_N(x) = \sum_{i=0}^{N-1} E(x_{i+1}^{i+m})$$

²The metric entropy is also known as the Kolmogorov-Sinai entropy, as the measure theoretic entropy, or merely as the entropy.

satisfies $\Phi_N(Fx) = \Phi_N(x)$. Hattori and Takesue [68] constructed procedures to find such conserved quantities, and found the expressions for a current and a corresponding continuity equation the quantity will obey. Pivato [128] has proved the following result for a CA F possessing a conserved quantity with $m = 1$: for any $r \in [\min_{a \in \mathcal{A}} E(a), \max_{a \in \mathcal{A}} E(a)]$ there is an F -invariant and shift-invariant measure μ_r satisfying $\int E(x_0) d\mu_r = r$.

An elegant way to construct reversible two-dimensional CA with conservation laws is the space partitioning technique introduced by Margolus [105]. The lattice is partitioned into 2×2 squares in two different ways, such that a square from the first partition intersects exactly four squares from the second partition. The local rule is now a permutation $f : \mathcal{A}^4 \rightarrow \mathcal{A}^4$, so that each square is updated independently. The two partitions are used on alternate time steps. This scheme can be recast as an ordinary CA, and makes it particularly easy to construct CA conserving the number of each symbol. In particular, Margolus [105] used it to construct the billiard ball CA, which is a computationally universal, reversible CA inspired by Fredkin and Toffoli's billiard ball model of reversible computation [52].

3.1.5 References to further aspects of CA theory

Much theoretical work has been done on one-dimensional CA, and less on CA in dimensions $d \geq 2$. One reason is that a wide array of the techniques used to analyse one-dimensional CA are unavailable in higher dimensions. Here, we point out some literature surveying the theory of one-dimensional CA.

An exposition of the topological dynamics of CA is provided by [88], while the survey article by Blanchard et al. [16] also covers measure-theoretic aspects. Three formal classification schemes of one-dimensional CA are discussed in [87]. See [118] for applications of finite automata theory to analysing CA. Two recent reviews of CA theory and history aimed at computer scientists are [81] and [137].

3.2 Coherent, propagating structures in CA

An important property of many CA is the ability to generate and support persistent, localized structures that can traverse the lattice and sometimes interact. Many types of such coherent structures exist, and different terms are used to describe them based on the nature of the phenomenon. This includes particles [64], defects [58], gliders [107], solitons [125] and waves [165].

3.2.1 Space-time diagrams

A wide range of stable, propagating structures were first found in the famous CA known as the "Game of Life" [14]. These range from the simple Glider to large, persistent structures with advanced functionalities [56, Chs. 20-22]. Berlekamp et al.

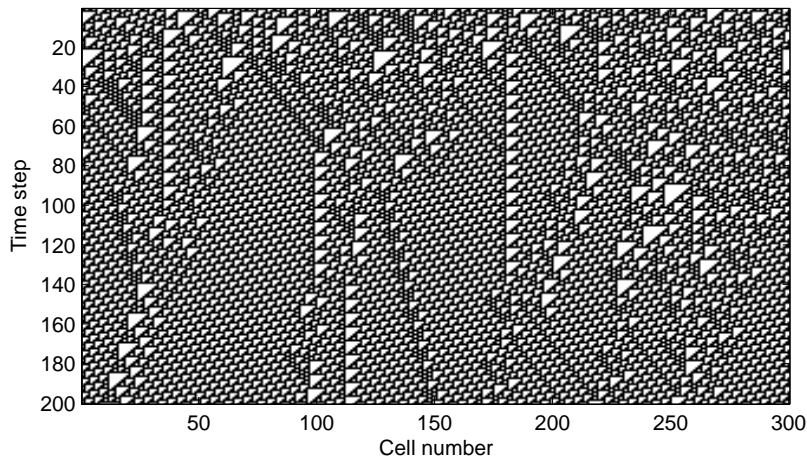


Fig. 3.2: A space-time diagram with 300 sites for ECA rule 110, starting from a uniformly random initial configuration. The lattice configurations produced by the subsequent iterations of the CA are displayed. The symbols 0 and 1 are represented by white and black squares respectively.

[14] have proved that the Game of Life is computationally universal by using signals, logic gates, and memory units constructed from structures stable in a quiescent background of zeros.

The computational study of CA was initiated by Wolfram [167], who conducted an extensive study of the elementary CA (ECA). In particular, he discovered the richness of the behaviour of CA through constructing space-time diagrams. A space-time diagram is a time series of the configuration of a finite part of the lattice, plotted with states represented by suitable colours. Often, periodic configurations are used, and a single period is plotted. An example of a space-time diagram is found in Fig. 3.2. Based on the typical behaviour from random initial configurations, Wolfram informally divided CA into four complexity classes [168]. Most interesting is his class four, which consists of rules generating complex localized structures which sometimes are long-lived. CA exhibiting behaviour of this type are epitomized by ECA rule 110, which is the CA used to generate the diagram in Fig. 3.2. Wolfram's work demonstrated that particles and particle interactions are widespread phenomena in CA dynamics, and his observations led to a major resurgence of interest in cellular automata.

3.2.2 Formal approaches

A formal approach to classifying and finding coherent structures in CA was conceived by Hanson and Crutchfield [64]. They realized that the regular background pattern seen in space-time diagrams formally can be described in terms of a regular language invariant under the action of the CA. A regular language is a language that can be recognized by a finite state automaton, and is equivalent to a sofic shift, see [9]. Intervals of the lattice whose configurations conform with the regular language are called regular domains. The regular domains can be filtered out of the space-time diagram using a finite state transducer, leaving only the deviations. These are called domain boundaries or embedded particles. These particles can subsequently be classified and their interactions investigated, see [65] and [73] for examples. The time evolution of some CA may after an initial transient be described merely in terms of the movements of the embedded particles.

The sofic shifts characterizing the regular domains can have zero or positive topological entropy, and the embedded particles can traverse the lattice at constant speed or perform random movements dependent on the local context. An example of the latter case is elementary CA rule 18. This local map of this rule is defined by $f(1,0,0)$ and $f(0,0,1)$ being 1, and f of the other six neighbourhood configurations being 0. See the left panel of Fig. 3.3 for a space-time diagram generated by rule 18. The regular domains are defined by the sofic shift having an odd number of zeros between every pair of ones. The domain boundaries, or defects as they are often called for this and similar rules, are blocks of the form $10^{2n}1$ for $n \geq 0$. The defects diffuse and annihilate pairwise upon collision. The right panel of Fig. 3.3 shows a filtered space-time diagram. Grassberger [58] first discovered the defects in rule 18, and Lind [98] conjectured that their movements constitute random walks. This was later proved by Eloranta and Nummelina [45]. The source of the randomness is the random initial configuration. Eloranta [44] has rigorously studied and constructed cellular automata with particles exhibiting biased random walks with capabilities of annihilation and coalescence.

3.3 Applications of CA

As pointed out by several authors [51, 106, 156, 159], cellular automata possess by construction many of the characteristics of physical systems. This includes space, time, locality of interaction, and a finite propagation speed of information [106]. Furthermore, by a suitable choice of rule, a CA can be made microscopically reversible and made to satisfy appropriate conservation laws. For these reasons, it is not surprising that many physical systems and other systems where local interactions are essential have been successfully modelled by CA. Moreover, as pointed out by Vichniac [159], CA can also be studied as discrete dynamical systems that embodies general physical ideas rather than simulating specific phenomena. Indeed, several

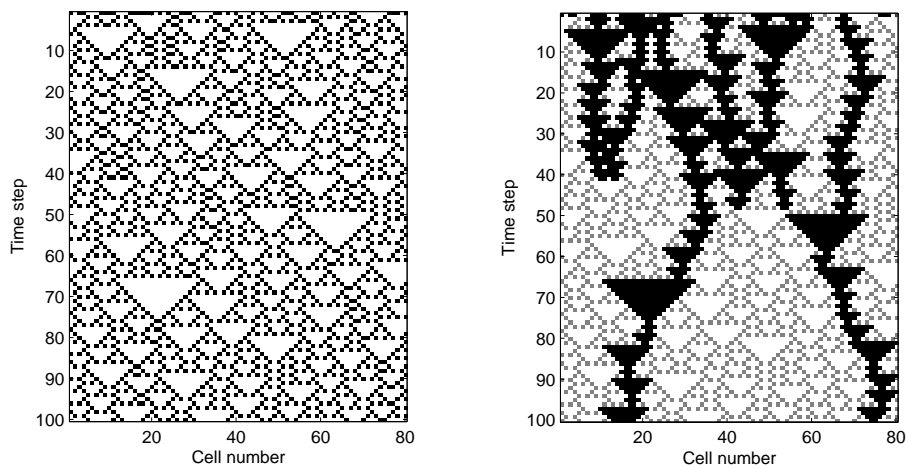


Fig. 3.3: Left: a space-time diagram with 80 sites for ECA rule 18. Right: a filtered version of the same space-time diagram showing the embedded particles.

dynamical phenomena found in CA have analogues in physics. In particular, this includes the emergence of coherent structures at different scales. Cellular automata also have several parallels to computation — they are digital and exact, and there exist reversible CA which are computationally universal [106]. CA thus provide a useful link between physics and computation [105]. The discreteness of cellular automata enables *exact* simulations of models, rather than approximations which may be prone to numerical errors.

The rest of this chapter provides an overview of some important applications of cellular automata. Some of these models may be interesting as bases for higher order cellular automata, see Chap. 4, or as systems in which to study information flow, see Chap. 5.

3.3.1 Reversible CA with conserved quantities

The most well-known application of CA is the study of fluid flows by lattice gas automata. The first lattice gas automaton was the HPP model introduced by Hardy et al. [67]. This is a reversible CA with $\mathcal{A} = \{0, 1\}^4$. Each component of \mathcal{A} represents the presence or absence of a particle moving in one of the four lattice directions. The update rule first moves each particle one position, then resolves the effects of collisions. For instance, if both a west-going and an east-going particle end up at coordinate i , they collide. The result is a deflection of both particles, and the new state at i is a north-going and a south-going particle. Both the number of particles

and the momentum are conserved quantities in the HPP model. However, due to the insufficient degree of rotational symmetry of the lattice, the dynamics is anisotropic at all scales. This led Frisch et al. [53] to introduce the FHP model, which is a lattice gas on a triangular lattice possessing hexagonal symmetry. The FHP model provides in the macroscopic limit the Navier-Stokes equation, and allows for the simulation of many hydrodynamical phenomena, such as the complicated interface between two fluids in the Rayleigh-Taylor instability [135] and flows in porous media [30]. The FHP model has some stochastic update rules used in resolving certain collisions. For details on lattice gas automata, see [133]. Dab et al. [37] have extended lattice gas models to include several types of particles and chemical reactions taking place at the lattice points.

A further application of reversible CA with conserved quantities to physics is simulation of Ising spin dynamics. The model introduced by Creutz [36] is a two-dimensional CA with alphabet $\mathcal{A} = \{\uparrow, \downarrow\} \times \{0, 1, 2, 3\}$, where the first term is the spin (up or down), and the second term is called the momentum. A spin will flip if the momentum at the site can compensate the energy change caused by the flip, such that the total energy is conserved. An alternative CA model of Ising spin dynamics is the Q2R model introduced by Vichniac [159]. When executed on finite lattices, both the Q2R CA and Creutz's CA are microcanonical models of Ising spin systems, since the total amount of energy in the system is constant. The models have the advantage over traditional Monte Carlo algorithms [120] of not needing access to random numbers. They can also be used to study various aspects of the Ising model, such as the thermal conductivity [136].

A more fundamental approach to using CA to investigate physics has been pursued by Takesue [150, 151, 152], who studied several one-dimensional reversible CA with locally conserved quantities. His goal was to find the characteristics of rules which realize thermodynamic behaviour, and through this investigate under which conditions statistical mechanics can be applied to a system. Takesue interpreted one of the conserved quantities of the CA as energy, and examined how well the behaviour of the CA complied with that predicted by thermodynamics. For some CA he obtained good results for the canonical distribution [150], Fourier's law of heat conduction [152] and other characteristics of thermodynamical systems. He found that thermodynamic behaviour is intimately related to the number and dynamical features of locally conserved quantities.

3.3.2 Models of excitable media

Excitable media are spatial systems which have the ability to propagate waves without damping, but where each unit after the passing of a wave needs a refractory time before it again can become excited. An example is chemical reactions such as the Belousov-Zhabotinskii reaction [169]. A classical CA model of excitable media is the Greenberg-Hastings CA [59], which can be defined in an arbitrary dimension

d with alphabet $\mathcal{A} = \{0, 1, 2\}$. The lattice coordinates represent cells which can be excited (1), tired (2) or rested (0). An excited cell is always tired at the next time step, and a tired cell becomes rested. A rested cell becomes excited if at least one of its $2d$ closest neighbours is excited. For this model there exist some rigorous results on large time behaviour [40, 146]. Later, more sophisticated CA models have been introduced, better reproducing the behaviour of real excitable media, such as curvature dependence of propagation speeds and wavefront propagation into partially rested cells, see for instance [165]. Reference [48] reviews CA models of excitable and oscillatory media, as well several other biologically motivated CA which arise in models of population biology, developmental biology and neurobiology.

3.3.3 Models of traffic flow

In recent years, modelling various aspects of traffic flow by cellular automata has become increasingly popular. The simplest CA that can be used to model traffic is elementary CA rule 184. The symbol 1 represents a car and the symbol 0 an empty space. The dynamics can be described in terms of the following rule: each car moves one space to the right if this space is empty, otherwise the car does not move. Even in this exceedingly simple model one can observe some of the phenomena of real traffic, such as traffic jams moving backwards.

The basic model for simulating traffic by CA was introduced by Nagel and Schreckenberg [116]. It is based on ECA rule 184 but includes acceleration and stochastic effects. The model can be analysed analytically in special cases. Later, more sophisticated models have been introduced, such as the modified model by Knospe et al. [82], which is able to reproduce the three phases (free-flow, synchronized, and stop-and-go) observed in real single lane traffic. Purely deterministic CA models have also been considered. For example, Boccaro and Fuk s [17] have investigated a family of CA which generalize rule 184.

3.3.4 Applications in computer science

Cellular automata can be considered as abstract models of parallel computers. The input data is represented by the initial configuration of the lattice, and the CA transforms this to an output configuration representing the output of the computation. CA are known to have strong computation capabilities. For example, ECA rule 110 is capable of universal computation [33]. There also exist CA with the ability to simulate all other CA of the same dimension, see [123] for details. The computational capabilities of CA hinge on the existence of particles and interactions among them. For instance, Steiglitz et al. [147] showed how computations can be embedded in one-dimensional CA by using the soliton CA introduced by Park et al. [125]. G acs [54, 55] has shown how to construct CA which reliably can perform arbitrarily large computations when each cell can perform an error with

a positive probability. Toffoli and Margolus [155] have investigated the possibility of building microscopic, parallel hardware designed to implement CA, in order to obtain exceptionally large computational power and efficient microscopic simulation of systems that can be modelled by CA.

3.3.5 Artificial life

Many of the coherent structures observed in cellular automata seem to take on a “life of their own”. Examples include many of the complex oscillating or traversing local structures in the Game of Life. Indeed, Berlekamp et al. [14, p.849] speculated that “it’s probable, given a large enough Life space, initially in a random state, that after a long time, intelligent, self-reproducing animals will emerge and populate some parts of the space.” Due to the capabilities of CA for self-organization, these systems are much used by researchers within the artificial life community [91, 131]. The goal of this discipline is to unravel the profound organizational principles and rules allowing the existence of living organisms. In the words of Langton [92], artificial life seeks to “locate life-as-we-know-it within the larger picture of life-as-it-could-be”.

Of particular importance is the property of self-reproduction. In fact, cellular automata were originally introduced by von Neumann [161] to serve as an idealized structure in which to study the logical conditions for self-reproduction. He wanted to abstract from natural biological self-reproduction the logical form of the process, independent of its material representation in any physio-chemical form. Von Neumann managed to construct a two-dimensional CA with $|\mathcal{A}| = 29$ and $|N| = 5$ in which a particular initial configuration was self-replicating. Von Neumann’s CA was also capable of universal computation. By not requiring universal computation, Langton [90] was later able to construct a CA with $|\mathcal{A}| = 8$ in which a certain sheathed loop structure replicates itself in 151 iterations of the CA. In both these CA models, the lattice must have a particular initial configuration for self-replication to occur. In contrast, Chou and Reggia [31] have created a two-dimensional CA where self-replicating loops emerge from a random initial configuration.

Chapter 4

Higher order dynamical systems

The history of life is a history of the emergence of new organizational grades, and their subsequent diversification.

–Leo W. Buss [28]

Outline

In Section 4.1, we outline the motivation for introducing higher order dynamical systems, and discuss background material such as Baas's hyperstructure concept. Section 4.2 describes the principles of higher order dynamical systems, and in Section 4.3, we present our work on higher order cellular automata. The presentation includes an illustrative example, the main results of our experiments, and an outlook.

4.1 Background and motivation

Evolution seems to favour systems that are built up by subunits on several levels, such that each unit on level n is an aggregation of units on level $n-1$. This hierarchical organizational principle is ubiquitous in biology, from the microscopic level to the macroscopic level. For instance, atoms form molecules, molecules form macromolecules, then organelles, cells, tissues, organs, organisms and ecosystems. The units on each level in the hierarchy obtain their properties by virtue of interactions among units on lower levels. However, units of higher order often have new properties, not found in lower level units. This phenomenon is known as emergence. Emergence in a dynamical system can be regarded as the creation of a new and higher level of organization, containing new types of units with new properties. The

formalization and philosophical aspects of emergence remain a highly debated topic, see e.g. [46].

Two mechanisms are vital in generating robust hierarchical structures with advanced functionalities: evolution [144] and self-organization or self-assembly [131]. The latter type of processes are essential in molecular biology and in bridging living and non-living matter. Additionally, self-organization allows a limited amount of heritable information to code for exceedingly complex processes, such as brain functions. Furthermore, the construction of new materials by self-assembly is an emerging technique in nano technology. It is a basic problem to give a framework in which to study hierarchical structures with assembly of new levels, emergence of new properties from collective dynamics, and interactions between and within levels. A step toward this goal is Baas's hyperstructure concept, which will be described in Sect. 4.1.1.

It has been pointed out by several authors, e.g. [49, 131, 138], that classical dynamical systems often not are quite suitable in order to construct models of how to generate various kinds of organization, such as biological structures. However, dynamical system models will still be useful in modelling systems with a hierarchical organization, as long as this structure is properly taken into account. For this reason, some general principles seem to be needed in order to facilitate the construction of general, flexible classes of dynamical systems with a hierarchical interaction structure. By studying such dynamical systems one can also hope to illuminate questions such as what properties that are typical and possible for systems with hierarchical organization, and why evolution favours such systems. The current chapter concerns higher order dynamical systems, which is a concept that addresses these issues. Higher order dynamical systems are related to hyperstructures, which we discuss first.

4.1.1 Hyperstructures

The hyperstructure concept was introduced by Baas [4] in 1994 as a framework in which to study structures with multiple levels, and in a fruitful way combine hierarchies, self-organization, and emergence. Recently, Baas [5] has elaborated the notion and introduced a new kind of structure, called Abstract Matter. We here give a brief overview of the hyperstructure framework, using the notation from [4].

The framework has three fundamental components. The first is the primitive objects or units, which can be of physical or abstract nature. The second component is some kind of observational mechanism (Obs) that observes, describes and evaluates the objects. The observational mechanisms can be internal in the system or external. Thirdly, we have interactions (Int) among the objects. The interactions use the properties detected by Obs.

A family S_i of objects, with i in some index set, together with specified properties and interactions defines a process or construction. The stable outcome or result of

this process, in some cases the attractor for the system, is written as $R(S_i, \text{Obs}(S_i), \text{Int})$. The process may generate a family of second order objects S_j^2 . In the reformulated version of hyperstructures presented in [5], the interactions are bonds which directly “bind” families of primitive objects to produce second order objects. The new family of second order objects has itself a set of observations, called Obs^2 . The observations Obs^2 may equal, overlap, or be disjoint from Obs . Baas calls a property P of S_j^2 emergent if $P \in \text{Obs}^2(S_j^2)$, but $P \notin \text{Obs}(S_i^1)$. Thus, a property of S_j^2 is emergent if it is not found in any of the constituents of S_j^2 . There are also interactions Int^2 between the second order objects. These second order interactions can either occur as a consequence of first order interactions, or they can be introduced externally using the new emergent properties.

The formalism can be iterated to define objects of third and higher orders. An N th order object is a result of interactions on all lower levels: $S^N = R(S^{N-1}, S^{N-2}, \dots, S^1)$. The resulting structure, consisting of objects on different levels together with their interactions and properties, is called a hyperstructure. It should be emphasized that a hyperstructure not merely is the result of a recursive procedure. Objects at higher levels obtain new properties, and the new interactions and observations at higher levels may or may not be meaningful at lower levels. Overlapping objects are allowed, and there may be interactions between objects at different levels. See the illustration in Fig. 4.1. Furthermore, hyperstructures allow for both “upwards” and “downwards” causation, in that interactions on higher levels may cause changes in interactions on lower levels. Note also that the choice of observables dictates what higher order objects that can be generated.

The formalism is a conceptual one, and more formal rigour and details are required in each specific case. One way to become more specific with the constructions is to interpret them within category theory, see [5] and [7]. Baas [4] states that “it is our conviction that complexity often takes the form of a hyperstructure”.

4.1.2 An *ansatz* for dynamical hierarchies

Rasmussen et al. [131] have studied self-assembly in a molecular dynamics model, and interpreted the result of the construction within the framework of hyperstructures. The primitive objects in the model are monomers which are hydrophilic or hydrophobic. The dynamics is given by a molecular dynamics lattice gas, which is an extension of the lattice gas CA described in Sect. 3.3.1. All molecular interactions, including force fields, are modelled by mediating particles.

When the monomers are equipped with enough structure, the model is able to generate three levels of organization with emergent properties on each level. The data structures of the monomers are internal observers. On the second level one has polymers with elasticity, and on the third level micelles with an inside/outside as well as permeability.

In the model, the second order interactions are generated from a composition of

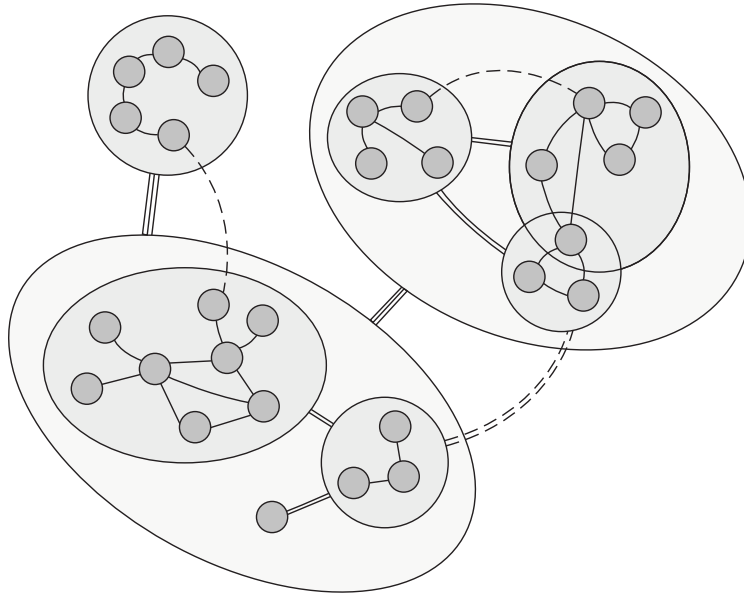


Fig. 4.1: An illustration of a third order hyperstructure. Circles represent objects of increasing order, and lines between them represent interactions. Dashed and solid-drawn lines represent different types of interactions, and double and triple lines denote interactions on higher levels.

first order interactions, as only the monomers interact directly. Thus, the meaning of an interaction will depend on the context in which it occurs. The model also exhibits downward causation, since the dynamics of the monomers become more restricted once they form a polymer. The authors state that “the local use — or interpretation — of the different kinds of communicated information defines the operational semantics of the information”, and further that “in general, new hyperstructural levels support new means of communication, both within new levels and between old and new. This is why the object complexity is bound to increase as more hierarchical levels are to be generated.” The observation that higher object complexity is needed to support different means of communication leads to the main issue in the paper, which is a new *ansatz*¹ for dynamical hierarchies [131]:

Given an appropriate simulation framework, an appropriate increase of the object complexity of the primitives is necessary and sufficient for generation of successively higher-order emergent properties through aggregation.

¹An *ansatz* is a hypothesis taken to be true but acknowledged to be unproven that is used to reach further conclusions [131].

This contradicts what the authors call the complex systems dogma, namely that a common minimal simplicity underlies all emergent structures [131].

4.1.3 Other models

Related to Baas's hyperstructure model is the work of Ehresmann and Vanbremeersch [41, 42], who within the frame of category theory [103] have developed a model for hierarchical autonomous systems, called memory evulative systems. In their model a system is, at each time, modelled by a category whose objects represent the objects of the system and whose morphisms represent interactions. The hierarchical structure is modelled by the colimit operation, such that an object on level $N + 1$ is the colimit of a pattern of objects at level N . A particularly important phenomenon occurs when two different patterns have the same colimit. Then, due to the composition property of categories, new *complex links* not induced by any interactions on the lower level will be formed at level $N + 1$. In the language of [5], the colimit construction is a particular type of bond structure. See [6] for a comparison between hyperstructures and memory evulative systems. A further general model of complex systems based on category theory is the universal dynamics introduced by Mack [104].

Fontana and Buss [49] have developed an artificial chemistry model based on λ -calculus for the formation of self-generating robust hierarchical systems. The objects in the model are λ -expressions which can act on each other according to the laws of the calculus, and hence form new expressions. With the model, Fontana and Buss obtained three levels of organization, with coexistence of self-maintaining organizations on the highest level. Eigen and Schuster [43] introduced the notion of a hypercycle, which is an interrelated hierarchy of cyclic reaction networks. They considered the emergence of hypercycles as a possible mechanism for the creation of biological hierarchies. Also these systems can be viewed as instances of hyperstructures [4, 5].

4.2 Higher order dynamical systems

By the term higher order system we mean in general a system where we lift attention from a set X to the power set $\mathcal{P}(X)$ of subsets of X , as well as higher power sets, viz. $\mathcal{P}^2(X)$, $\mathcal{P}^3(X)$, and so on. The power set process is fundamental in mathematics. For instance, fractals are objects in $\mathcal{P}(X)$, and topologies and σ -algebras are elements of $\mathcal{P}^2(X)$. Power sets are also the fundamental building block for hyperstructures, where collections and patterns of objects obtain new properties as a consequence of interactions among them [5].

In Paper II, we outline the principles behind higher order dynamical systems and illustrate the concept by defining a class of systems called higher order cellular automata. The motivation behind introducing higher order dynamical systems is the need for a framework in which to construct and study dynamical systems

of interacting units where the interactions in various ways are hyperstructured. Furthermore, a basic tenet is that allowing for higher order effects in a system may lead to new and interesting types of dynamical behaviour. For instance, in Paper I we demonstrate that n -player evolutionary games with two strategies can have a significantly richer dynamical behaviour than such games with only two players can. Thus, looking at a collective interaction within a group can lead to new types of behaviour compared to looking merely at a set of binary interactions.

We take as the basis a system of locally interacting units, and look at which higher order extensions that can be introduced. A system of this type can be characterized by the three parts of which it is composed:

- (1) The units or agents themselves.
- (2) An interaction structure defined through the neighbourhoods of each unit.
- (3) A set of local interaction rules governing the dynamics.

A basic higher order extension is to let *groups* or aggregates of units become new entities with their own dynamical behaviour and interactions. This corresponds to the central mechanism of hyperstructures, namely emergence of second order objects, see Sect. 4.1.1. The groups are called second order units, or 2-units for short. Continuing the process one may consider 3-units, being groups of 2-units, and so on. The introduction of higher order units gives the system a certain multilevel structure, which we call a *higher order morphology*. A basic example of 2-units and 3-units is the formation of molecules from atoms and macro-molecules from molecules. A further example is the formation of families from individuals and communities from families.

It is important to note that the existence of second and higher order units is taken as predefined, and is not supposed to follow from the properties of the primitive units. This is also the case for internal states of higher order units and interactions among the higher order units. These properties are thus emergent, but the mechanisms leading to this emergence are typically beyond the scope of the dynamical system model. Still, the properties must be taken into account, as they may have a significant impact on the behaviour we wish to model.

The power set process can also be applied to the other components of the basic system, i.e. to the local rules and the interaction neighbourhoods. This produces 2-rules and 2-neighbourhoods which constitute a *higher order dynamics* — a dynamics on the dynamics. A typical example of a 2-rule is a collection of k basic rules together with a process that governs exactly which rule that is applied at each time step. This can be interpreted as the units having different modes of operation. Allowing for 2-rules and 2-neighbourhoods adds complexity to the primitive units in a natural way, and corresponding to the ansatz for dynamical hierarchies this will lead to facilitation of structures of higher morphological order. When a unit u is a component of a 2-unit v , it is natural to let the 2-unit v affect u indirectly through in some way controlling u 's 2-rule. This corresponds to the downward causation discussed in connection with the

molecular self-assembly model described in Sect. 4.1.2, where the existence of higher order units led to restrictions in the dynamics of lower levels.

We say that a system has a *higher order structure* if it has a higher order morphology and/or a higher order dynamics. It is then a higher order system. We use the term higher order dynamical system when considering a higher order system as a formal dynamical system respecting the higher order structure. The motivation for introducing higher order dynamical systems comes from biological and social networks where relations among the units depend on their positions within a larger structure, and where groups of units obtain new properties and both influence and are influenced by their constituents.

4.3 Higher order cellular automata

In Paper II, we introduce higher order cellular automata as a particular class of higher order dynamical systems. The objectives are to illustrate the concept and to explore how the existence and properties of various higher order structures influence the dynamical behaviour of the system. There are many reasons for the choice of CA as an underlying dynamical system. Firstly, as illustrated in Chap. 3, the behaviour of CA is extremely rich, and CA provide good models of many physical phenomena. Therefore, CA have been extensively studied in the context of self-organization and complexity, see references in Chap. 3. Furthermore, the behaviour of CA is easy to visualize through space-time diagrams, and CA can be exactly simulated on a computer.

We only define second order cellular automata, which we call 2-CA for short, but the formalism can readily be extended to CA of higher orders than two. A 2-CA can have one or two morphological levels and one or two dynamical levels. We consider first the case of a 2-CA with a 2-dynamics, and then a 2-CA with both a 2-dynamics and a 2-morphology. The notation used in connection with CA is introduced in Sect. 3.1.1.

4.3.1 2-Dynamics

Denote by 2-DCA a 2-CA that has no 2-morphology². The process of constructing a 2-DCA is in some sense merging a collection ordinary CA into a single dynamical system. Recall that a CA on $\mathcal{A}^{\mathbb{Z}^d}$ can be defined by a neighbourhood N and a rule $f: \mathcal{A}^{|N|} \rightarrow \mathcal{A}$. For a 2-DCA, one has a collection $\mathcal{M} = \{(N_1, f_1), \dots, (N_k, f_k)\}$ of pairs of possible neighbourhoods and rules. The 2-dynamics consists of \mathcal{M} together with a 2-transition rule ϕ that governs the local change of rule and neighbourhood for each cell $v \in \mathbb{Z}^d$. In the language of hyperstructures [5], the collection \mathcal{M} is a *bond* of rules and/or neighbourhoods.

²The term 2-DCA is not used in Paper II.

Definition 4.1. Let \mathbb{Z}^{d_1} be the d_1 -dimensional lattice and \mathcal{A}_1 a finite alphabet. Let $\mathcal{M} = \{(N_1, f_1), \dots, (N_k, f_k)\}$, where for each i :

- N_i is a finite subset of \mathbb{Z}^{d_1} containing the origin.
- f_i is a map $\mathcal{A}_1^{|N_i|} \rightarrow \mathcal{A}_1$.

Lastly, let $\phi = (\phi_i)_{i=1}^k$ be a collection of maps $\phi_i : \mathcal{A}_1^{|N_i|} \rightarrow \mathcal{M}$. The quadruple $\langle d_1, \mathcal{A}_1, \mathcal{M}, \phi \rangle$ defines a 2-DCA.

A 2-DCA is a dynamical system on the space $(\mathcal{A}_1 \times \mathcal{M})^{\mathbb{Z}^{d_1}}$. Each cell $v \in \mathbb{Z}^{d_1}$ is assigned a state in \mathcal{A}_1 as well as a pair in \mathcal{M} consisting of an active neighbourhood and an active rule. These values are updated synchronously for all cells, using the following scheme:

- (1) Denote the current pair in \mathcal{M} assigned to v as (N_j, f_j) .
- (2) Define $X = \sigma^v(x)_{N_j} \in \mathcal{A}_1^{|N_j|}$, so that X is the configuration of the current neighbourhood.
- (3) The new state of v becomes $f_j(X)$.
- (4) The new rule and neighbourhood of v become $\phi_j(X)$.

With the Tychonoff topology on $(\mathcal{A}_1 \times \mathcal{M})^{\mathbb{Z}^{d_1}}$, the 2-DCA is a compact and continuous topological dynamical system. Definition 4.1 is a concise, but equivalent, reformulation of the definition found in Paper II. In particular, the set \mathcal{M} is used to indirectly represent the 2-neighbourhood $\mathcal{N} = \{N_i : 1 \leq i \leq k\}$ and the 2-rule $\mathcal{R} = \{f_i : 1 \leq i \leq k\}$ used in Paper II. The following is an example of the construction of a 2-DCA.

Example 4.1. Let F be a one-dimensional 2-DCA defined by the dimension $d_1 = 1$, the alphabet $\mathcal{A}_1 = \{0, 1\}$, the set $\mathcal{M} = \{(N, f_1), (N, f_2)\}$, with $N = \{-1, 0, 1\} \subset \mathbb{Z}$ and rules

$$f_1(x_{-1}, x_0, x_1) = \min(x_{-1}, x_0, x_1) \quad \text{and} \quad f_2(x_{-1}, x_0, x_1) = \max(x_{-1}, x_0, x_1),$$

and by the 2-transition rule

$$\phi_1(x_{-1}, x_0, x_1) = \begin{cases} f_2 & \text{if } x_{-1} + x_0 + x_1 = 1, \\ f_1 & \text{otherwise.} \end{cases}, \quad \phi_2(x_{-1}, x_0, x_1) = \begin{cases} f_1 & \text{if } x_{-1} + x_0 + x_1 = 2, \\ f_2 & \text{otherwise.} \end{cases}$$

Thus, there is only one possible neighbourhood, namely $N = \{-1, 0, 1\}$, and two possible rules. These comprise a 2-rule $\{f_1, f_2\}$. The 2-DCA F is a dynamical system on the space $(\{0, 1\} \times \{f_1, f_2\})^{\mathbb{Z}}$. A space-time diagram generated by F is found in Fig. 4.2.

Let F_1 be the CA with block map f_1 and F_2 the CA with block map f_2 . Then, for any configuration x not being equal to 0^∞ or 1^∞ we have

$$\lim_{t \rightarrow \infty} F_1^t(x) = 0^\infty, \quad \text{and} \quad \lim_{t \rightarrow \infty} F_2^t(x) = 1^\infty.$$

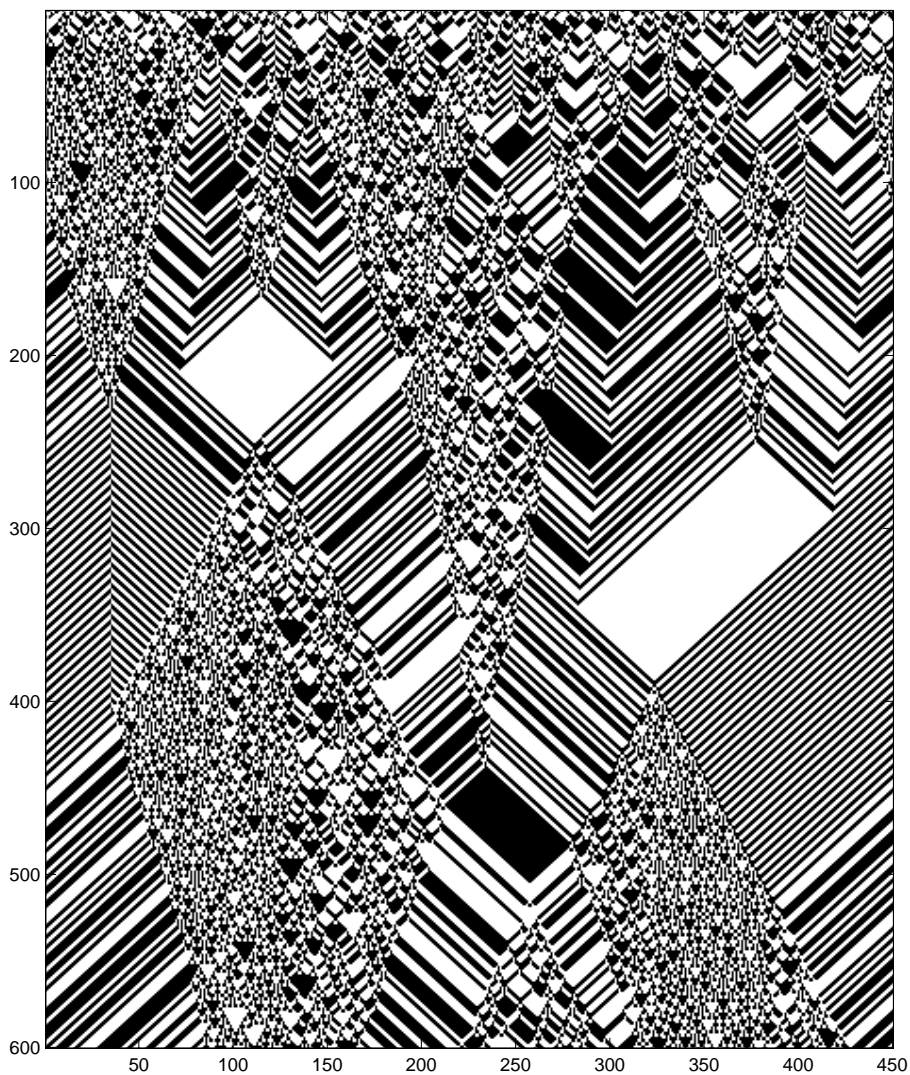


Fig. 4.2: A space-time diagram generated by the 2-DCA from Example 4.1, showing the evolution of a periodic initial configuration with period 450. The cell states are shown, with white representing 0 and black representing 1.

Thus, both F_1 and F_2 are rather simple CA. The 2-DCA, however, has an exceedingly more complicated behaviour. From space-time diagrams we observe the emergence of several types of domains, whose internal dynamics range from that of a shift map to one generating apparently chaotic patterns. Each domain typically first expands, then shrinks and disappears as neighbouring domains of other types take over.

In Paper II we prove that every 2-DCA is topologically conjugate to an ordinary CA. Thus, the set of all non-trivial 2-DCA³ can be considered as a subset of the collection of all CA. However, the above formalism is more natural for describing this class of CA, in the same way that the formalism of partitioned CA is more natural for describing CA such as Margolus's billiard ball model, see Sect. 3.1.4. Moreover, we will be interested in studying the relations between a basic CA rule f and various 2-DCA based on it.

4.3.2 2-Morphology

We have seen that a 2-DCA is an aggregation of a collection of CA into a single dynamical system through a 2-transition map ϕ . This is also the case for a 2-CA with two morphological levels, but in this case the system additionally has a second level which also behaves like a CA. The objects on this level are called 2-cells, and are interpreted as aggregations of cells that have become new units with their own properties.

The 2-cells are represented by a lattice \mathbb{Z}^{d_2} , and their constituent cells by a map $M: \mathbb{Z}^{d_2} \rightarrow \mathcal{P}(\mathbb{Z}^{d_1})$, where $\mathcal{P}(\mathbb{Z}^{d_1})$ denotes the set of all finite subsets of \mathbb{Z}^{d_1} . This gives each 2-cell u an *inner structure*. Furthermore, each cell v has an *outer structure* consisting of the 2-cells in which v is a component. The 2-cells have their own internal states from a set \mathcal{A}_2 , as well as their own neighbourhood N_2 and update rule f_2 . In the definition presented below we require that each 2-cell contains the same number c of cells, and that each cell is a component of exactly e 2-cells. Although this requirement limits the generality of the definition, it simplifies the notation considerably.

Definition 4.2. *Let d_1, \mathcal{A}_1 and \mathcal{M} be as in Definition 4.1. Furthermore, let \mathbb{Z}^{d_2} be a d_2 -dimensional lattice, \mathcal{A}_2 a finite alphabet, N_2 a finite subset of \mathbb{Z}^{d_2} containing the origin, and M a map $M: \mathbb{Z}^{d_2} \rightarrow \mathcal{P}(\mathbb{Z}^{d_1})$, such that for some $c \in \mathbb{N}$ and $e \in \mathbb{N}$,*

$$|M(u)| = c \quad \forall u \in \mathbb{Z}^{d_2}, \quad \text{and} \quad \left| \{u \in \mathbb{Z}^{d_2} \mid v \in M(u)\} \right| = e \quad \forall v \in \mathbb{Z}^{d_1}.$$

Lastly, let f_2 be a map $f_2: \mathcal{A}_2^{|N_2|} \times \mathcal{A}_1^c \rightarrow \mathcal{A}_2$ and $\phi = (\phi_i)_{i=1}^k$ be a collection of maps $\phi_i: \mathcal{A}_1^{|N_i|} \times \mathcal{A}_2^e \rightarrow \mathcal{M}$. The tuple $\langle d_1, d_2, \mathcal{A}_1, \mathcal{A}_2, M, N_2, f_2, \mathcal{M}, \phi \rangle$ defines a 2-CA.

A 2-CA is a dynamical system on the space $(\mathcal{A}_1 \times \mathcal{M})^{\mathbb{Z}^{d_1}} \times \mathcal{A}_2^{\mathbb{Z}^{d_2}}$. Each lattice point $u \in \mathbb{Z}^{d_2}$ represents a 2-cell, and is assigned a state in \mathcal{A}_2 . The inner structure map M

³Non-trivial in the sense that $|\mathcal{M}| > 1$.

binds the lattices together. The values of each cell v are updated in the same way as for a 2-DCA. However, now the 2-transition rule ϕ also takes into account the states of the 2-cells v is a component of. The state of each 2-cell u is updated by f_2 , using the neighbourhood N_2 and the states of $M(u)$.

4.3.3 Experiments and results

A guiding question to ask when exploring higher order dynamical systems is what impacts various higher order extensions have on the behaviour of the underlying systems. There are many subclasses of questions one can ask about 2-CA related to this issue. Firstly, one can study particularly interesting instances of 2-CA and their mathematical characteristics. Secondly, one can fix a CA F and look at various higher order extensions of it, and how they relate to F and to each other. Similarly, one can fix a collection \mathcal{M} of CA and look at various 2-transition rules and, possibly, 2-morphologies. Thirdly, one can look at the collection of all 2-CA and how this class of dynamical systems compares to ordinary CA and other types of discrete dynamical systems. When looking at a particular group of 2-CA, for instance those based on a certain collection \mathcal{M} , one can both consider what properties these 2-CA typically have, in some appropriate statistical sense, or what properties that exist within the group.

To obtain an impression of the impacts of various higher order structures, we have looked at many 2-CA using computer simulations and space-time diagrams. Some general principles we followed during the exploration were to limit the complexity of the rules involved, to take as starting points CA rules known to produce interesting behaviour, and to look at classes of 2-CA having, in a wide sense, plausible transition rules from a physics perspective. Some of the interesting and illustrative 2-CA we found are presented in Paper II. We summarize our observations here, and refer to Paper II for examples.

Several of the 2-CA we discovered exhibit interesting and, to our knowledge, novel phenomena and behaviour. We have observed many different types of coherent structures, including both localized structures propagating at fixed speeds, localized structures behaving in a more irregular way, as well as the emergence of distinct domains with different internal behaviours. It appears that a 2-DCA is most likely to exhibit such complex behaviour when at least one of the constituent CA also supports coherent structures. Therefore, equipping CA known to have interesting characteristics with various higher order extensions could be a viable way to construct new interesting dynamical systems. Nevertheless, also when merging CA which themselves produce no coherent structures, one can obtain 2-DCA with such properties. The 2-DCA in Example 4.1 is an example of two trivial CA that merge into a 2-DCA with a highly non-trivial behaviour. Overall, our investigations imply that it in general is hard to predict whether a 2-DCA will produce interesting behaviour or not, based merely on knowledge of \mathcal{M} and ϕ .

A novel phenomenon observed in several 2-DCA is the one called *crystallization* in Paper II. Crystallization occurs when a system undergoes a transition from a disordered apparently chaotic regime to an ordered regime with a much lower temporal entropy through gradual expansion of regular areas. A regular area is a part of the lattice where the configuration is more regular than on the surrounding lattice. This regularity must be robust to influences from neighbouring cells, so that it is retained during evolution.

Concerning the relations among 2-DCA based on the same collection \mathcal{M} of ordinary CA, we cannot draw any strong conclusions from the experiments. In many cases, the range of behavioural characteristics found within a group of 2-DCA with the same \mathcal{M} is very wide. But as previously stated, if one of the rules in \mathcal{M} exhibit coherent structures, there is a propensity for the 2-DCA to have this property too. However, the coherent structures which are produced are often of dissimilar nature for different 2-transition rules. Overall, similarity of \mathcal{M} does not seem to be a universally good measure of proximity in the space of all 2-DCA of a certain size, in the same way that similarity of block maps is a poor measure of proximity in CA-spaces.

For 2-CA with 2-morphologies there are more parameters involved than for 2-DCA, and consequently we have conducted less exploration of this class of systems. However, we have observed that the coupling geometry can have a fundamental impact on the dynamical features of the higher order system.

4.3.4 Discussion and future directions

Our explorations of 2-CA have so far mainly been based on visual aspects of space-time diagrams displaying evolution from uniformly random initial configurations. We have made several interesting observations of novel types of complexities and phenomena. Nevertheless, it can be argued that although such space-time diagrams provide good qualitative information about the dynamical features of a given CA or 2-CA, potentially important quantitative aspects remain concealed. For instance, conserved quantities as well as computational properties are not easily detectable from space-time diagrams. Therefore, our observations should be seen as an impetus for more quantitative or formal investigations of the questions raised in Sect. 4.3.3.

Cellular automata, as well as 2-CA, are in general difficult both to analyse and classify, due to the discreteness of the construction. One viable approach to study both the scope for self-organizing complex behaviour in the class of 2-CA, and the behavioural characteristics of a given 2-CA, is investigation by various concepts adapted from information theory. In particular, this includes the notion of local information, which we introduce and use to study information transport in ordinary CA in Papers III and IV. Since 2-DCA constitute a subclass of CA, local information can be applied directly to this class of systems. A future challenge is to extend the concept to 2-CA with 2-morphologies in a way which takes the higher order structure into

account. For details on local information and its relevance to CA we refer to Chap. 5.

A specific task in the future is to undertake a closer study of crystallization and its causes, including a formal definition of the phenomenon. We believe that also here the concept of local information and information transport may be useful. There is probably smaller and more regular flow of information within the regular areas than in the rest of the lattice. Additionally, one can apply the theory of formal languages to describe and classify the expanding regular areas created during crystallization. This is an approach similar to that Hanson and Crutchfield [64] used to characterize embedded particles in CA, see Sect. 3.2.2.

There are several extensions and variations of the 2-CA framework presented in Paper II that would be interesting to study. Firstly, the framework can be extended to accommodate additional morphological and dynamical levels, more than one type of primitives, and more than one lattice at each level. A further variant inspired by biological systems is to allow for a slower time scale on the second morphological level. Furthermore, by introducing rules that create new 2-cells when the local conditions are suitable, one can utilize the framework to investigate self-assembly.

The 2-CA model presented in Paper II is intended mainly as a conceptual mathematical model, and is probably not directly suited for the mathematical modelling of systems with a hierarchical structure. An exception is that certain 2-DCA might be adequate models of systems of interacting units having different modes of behaviour or an adaptable neighbourhood structure. Nevertheless, we believe the underlying principle of constructing dynamical systems by using 2-units, 2-rules and 2-neighbourhoods are applicable in other settings. These “higher order principles” are certainly also applicable to systems with stochastic dynamics and more general agent based systems. Furthermore, the principles may be useful in connection with the large recent interest in complex networks and systems whose interaction topologies are complex networks [119, 148].

In connection with the work presented in Paper I, the higher order CA formalism can be used to construct spatial, multilevel games based on our threshold game. The basic objects will represent individuals and second order objects will represent groups or packs. One might also include third order objects representing biotopes. It is natural to let the threshold on the number of cooperators be implicitly given by competition among neighbouring groups, rather than being fixed. The dynamics can, for instance, be of imitation type [122]. We believe that such models might be relevant and useful in many settings.

Chapter 5

Local information and information transport

Now, confusion, like the correlative term order, is not a property of material things in themselves, but only in relation to the mind which perceives them.

–James C. Maxwell, “Diffusion”, Encyclopædia Britannica, 1878

Outline

Section 5.1 provides an extensive review of the concepts of entropy and information, their role in dynamical systems theory and in physics, and attempts to construct microscopic versions of the entropy. In Section 5.2, we introduce our local information measure, and Section 5.3 presents the main results of our work on the continuity of information transport in CA. Section 5.4 concerns the applicability of local information to the characterization and detection of correlations and coherent structures in lattice dynamics. Finally, in Section 5.5, we point out some future directions. Much of the notation used in the chapter is presented in Sections 3.1.1 and 3.1.3.

5.1 Background

5.1.1 Shannon entropy

The field of information theory originates in Shannon's seminal 1948 paper [141]. He introduced the notion of the information of a message in order to analyse transmission and compression of data. The *Shannon information* associated with an event A with probability p is $I(A) = -\log p$, where the logarithm is taken with base 2. The information $I(A)$ can be interpreted as the information gained when observing A . The function I satisfies the basic requirements that the information gained by observing a set of independent events should be additive, and that observing an event with probability one should convey no information. The *Shannon entropy* $H(X)$ of a discrete random variable X with n possible values is the expected information gained by observing its outcome, that is

$$H(X) = - \sum_{i=1}^n p_i \log p_i. \quad (5.1)$$

The convention that $0 \cdot \log 0 = 0$ is used, so that events with zero probability can be ignored. The entropy is often interpreted as the a priori uncertainty about the outcome of X . Note that the entropy $H(X)$ only depends on the *distribution* of X . Information and entropy are measured in bits. One bit is the information that can be stored in a binary variable.

The entropy $H(X_1, \dots, X_n)$ of a collection X_1, X_2, \dots, X_n of random variables is the entropy of their joint distribution. Note that $H(X, Y) \leq H(X) + H(Y)$, with equality only if X and Y are independent random variables. Of particular interest is the case when the variables constitute a discrete stationary stochastic process $(X_i)_{i \in \mathbb{Z}}$ with values in a finite set \mathcal{A} . The Shannon *entropy rate* of the process is defined as

$$h(\mu) = \lim_{n \rightarrow \infty} \frac{1}{n} H(X_1, X_2, \dots, X_n). \quad (5.2)$$

Here, μ denotes the Kolmogorov measure of the process, see Sect. 3.1.3.

If the random variables X and Y are statistically dependent, knowledge about the outcome of Y will reduce the uncertainty about X . The residual uncertainty about X is the conditional entropy $H(X|Y) = H(X, Y) - H(Y)$. The entropy rate can be written as a limit of conditional entropies [35],

$$h(\mu) = \lim_{n \rightarrow \infty} H(X_0 | X_{-n}, \dots, X_1) = - \lim_{n \rightarrow \infty} \sum_{x_0^n} \mu(x_0^n) \log \mu(x_0 | x_{-n}^{-1}). \quad (5.3)$$

In light of this relation, a natural interpretation of $h(\mu)$ is as the average information gained by reading a new symbol in a sequence generated by μ when remembering all past symbols. In this way, all correlations in the stochastic process are taken into

account. An upper bound for the entropy rate of a process on \mathcal{A} is $\log|\mathcal{A}|$. This is attained only for the uniform Bernoulli measure.

Information, entropy and related concepts have proved to be very useful also outside the field of communication theory, and many applications and further developments exist. The metric entropy is an extension of $h(\mu)$ to any measure-preserving dynamical system [83]. The metric entropy of the measure-preserving dynamical system $(\mathcal{A}^{\mathbb{Z}}, \mu, \sigma)$ is equal to the Shannon entropy rate $h(\mu)$. A topological analogue of $h(\mu)$ is the topological entropy introduced by Adler et al. [1]. Both the topological entropy and the metric entropy are important conjugacy invariants in ergodic theory. Positive entropies are commonly used to detect and quantify chaotic behaviour in dynamical systems. Physically, positive metric entropy can be interpreted as a transport of information from small length scales to larger length scales, since finer and finer aspects of the initial condition become visible as the system is iterated [142]. An example of a system exhibiting such behaviour is the logistic map $f(x) = 4x(1 - x)$ on $[0, 1]$. Other areas of mathematics where concepts from information theory are important are the theory of fractals, large deviation theory, statistical estimation, and coding theory. Also in physics these concepts play a prominent role.

5.1.2 Entropy and information in physics

Thermodynamics deals with description of systems such as gases in terms of macroscopic variables, such as temperature, volume and pressure. The *macrostate* of the system is given by the values of these variables. There are a vast number of microscopic configurations, or microstates, of the system corresponding to each macrostate. The entropy concept was first introduced into thermodynamics by Clausius [32]. He defined the change in entropy of a system as $dS_{\text{th}} = \delta Q / T$, where δQ is the amount of heat transferred to the system if the system follows a reversible path. This leads to a state variable S_{th} , called the thermodynamic entropy, which is defined for systems in thermal equilibrium¹. The second law of thermodynamics states that the entropy is non-decreasing in an isolated system.

Boltzmann [18] was the first to investigate entropy at the microscopic scale. He suggested that the thermal equilibrium of a system is the most *probable* macrostate, and found for the entropy the formula $S_{\text{th}} = k_{\text{B}} \ln W$ [18], where W is the number of microscopic configurations consistent with the macroscopic quantities and $k_{\text{B}} = 1.38 \cdot 10^{-23} \text{ JK}^{-1}$ is the Boltzmann constant.

Boltzmann's works led to the development of statistical mechanics, which deals with description of systems in terms of ensembles of microstates consistent with the macrostate. An ensemble is mathematically a probability measure on the phase space Γ , which is the space of all possible microstates. For instance, for an N particle gas the phase space is the $6N$ -dimensional space where each point specifies the position

¹A system is in thermal equilibrium when the macroscopic variables have ceased to change with time.

and momentum of each particle. When the state space is continuous, the measure is often represented by a probability density ρ . The trajectory of each point in phase space is determined by Hamilton's equations. The value of ρ is constant on each trajectory, and volumes in phase space are conserved, see, e.g., [127]. The evolution of ρ is governed at the microscopic scale by the deterministic and reversible Liouville equation,

$$\frac{\partial \rho}{\partial t} = - \sum_{j=1}^N \left(\frac{\partial H}{\partial \mathbf{p}_j} \frac{\partial \rho}{\partial \mathbf{r}_j} - \frac{\partial H}{\partial \mathbf{r}_j} \frac{\partial \rho}{\partial \mathbf{p}_j} \right). \quad (5.4)$$

Here, \mathbf{r}_j and \mathbf{p}_j are respectively the position and momentum of the j -th particle and H is the Hamiltonian.

Gibbs [57] introduced a new notion for entropy, namely the Gibbs H function,

$$H_G = \int_{\Gamma} \rho(x) \ln \rho(x) dx.$$

Note that when the phase space is finite, this expression is equal to the expression for Shannon entropy in (5.1), except for the sign and the base of the logarithm. The quantity H_G is invariant under evolution by the Liouville equation, due to volume preservation [77]. Furthermore, H_G has the property that for a system in thermal equilibrium, $-k_B H_G \leq S_{\text{th}}$ for all measures on Γ consistent with the macroscopic variables, with equality for the Gibbs distribution [77]. Von Neumann [160] later extended the definition of entropy to quantum physics.

The expressions of Boltzmann and Gibbs inspired Shannon when he created the theory of communication [8], but the analogue between Shannon's entropy and the H -function remained merely semantic until the concept of information was introduced into statistical physics by Brillouin [24] and Jaynes [75, 76]. The new insight was the interpretation of entropy as the lack of information about the actual microscopic configuration of the system. This interpretation yields a direct connection between information and entropy, namely that 1 bit = $k_B \ln 2$ units of entropy.

The concept of information has also been pivotal in exploring the relation between physics and computation [102]. In the 1950s it was believed that making a measurement on a system increases the entropy of the system. As a consequence, von Neumann [161] and others speculated that each logical operation performed in a computer at temperature T must dissipate at least $k_B T \ln 2$ joules of energy, and thereby increase entropy by $k_B \ln 2$. This was later proved to be wrong by Landauer [89], who showed that only irreversible logical operations require energy dissipation. In particular, memory erasure in a computer is irreversible, and therefore feeds entropy to the environment of at least $k_B \ln 2$ for each bit of information erased. Bennett [11] showed that all computations could be performed using only reversible logical operations. Therefore, computation in itself does not require dissipation of energy or entropy increase in the environment. More recently, quantum computation

and quantum information have become major topics of investigation, see [13] for a review.

The concept of information plays an increasingly important role in physics, for instance in the thermodynamics of black holes [10]. Several researchers have suggested that information should be viewed as a fundamental physical quantity. Wheeler [166] introduced the slogan “It from Bit”, and maintained that “every it — every particle, every force, even the spacetime continuum itself — derives its function, its meaning, its very existence entirely — even if in some contexts indirectly — from the apparatus-elicited answers to yes-or-no questions, binary choices, bits.”

5.1.3 Entropy conservation and reversibility

As stated in the previous section, it follows from the reversibility of Hamiltonian dynamics and the Liouville equation that the Gibbs H function H_G is an invariant of the dynamics. However, when we start with a system out of equilibrium, the evolution is *irreversible* at the *macroscopic* scale. Furthermore, S_{th} for the equilibrium state the system will settle into will be larger than $-k_B H_G$ (where H_G is value of the Gibbs H function associated with the initial state). This paradox that has led to much confusion about the nature of irreversibility and its relation to different types of entropy and the second law of thermodynamics [77].

Let us consider an example that illustrates the resolution of the paradox. We have two separated volumes of different gases at the same pressure and temperature. These are then mixed, and the system approaches the equilibrium state consisting of a homogenous mixture. In this state one cannot detect that the gases initially were separated. However, this *is* detectable at the microscopic scale, since if one could observe all positions and momenta exactly, one could calculate all previous configurations by applying the Hamiltonian equations. Thus, information is not lost, but rather transferred from variables of the system that can be observed macroscopically to subtle correlations among the particles, which are macroscopically out of reach. The increase of the thermodynamic entropy is in a sense not related to the system itself, but rather by the limited ability of the observer to detect all microscopic variables. The Gibbs H function, however, takes correlations into account, and is therefore invariant. By introducing suitable entropy measures corresponding to which variables that can be observed, the dissipation of information can be measured exactly, see, e.g., [8].

5.1.4 Microscopic entropy

A conceptual problem related to entropy is its definition in terms of the probability distribution over all possible microstates, while a physical system at each time is found in a single microstate. In the words of Bennett [12, p. 936], “the absence of a microscopic quantity corresponding to entropy is a nuisance both practically and

conceptually”.

For many models in statistical mechanics, such as the Ising model [74], the microstates are configurations of a d -dimensional lattice, such that $\Gamma = \mathcal{A}^{\{-n, \dots, n\}^d}$. In this case, it is convenient to consider the entropy *per site* in the limit of an infinitely large system. In the one dimensional case, $-H_G$ per site is equal to the Shannon entropy rate defined in (5.2). For such systems of symbol strings, Bennett [12] suggested the use of algorithmic complexity [29, 84, 145] to construct a microstate function with average almost equal to the entropy H_G . The algorithmic complexity of a symbol string is the length of the shortest binary computer program that prints the string and then halts. The approach was later developed by Zurek [170], who showed that the average algorithmic complexity of the microstates in a thermodynamical ensemble μ is only negligible larger than $h(\mu)$ [170, Th. 4.1].

An alternative approach applicable to systems where each microstate can be coded into a symbol sequence $x \in \mathcal{A}^{\mathbb{Z}}$, is to use the internal statistics of x to approximate the entropy rate [100]. For a shift-invariant measure μ , the *empirical measure* ν_x is well defined μ -a.e. This is the measure with block probabilities given by

$$\nu_x(a_1^n) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=0}^{N-1} 1_{\text{Cyl}(a_1^n)}(\sigma^{-k}x). \quad (5.5)$$

The idea is to take $H(\nu_x)$ as the entropy of x . For any ergodic measure μ , Birkhoff's ergodic theorem ensures that $\nu_x = \mu$, μ -a.e., so in this case almost all microstates contain all information about the ensemble. Furthermore, it is then clear from Theorem 3.3 that $H(\nu_x) = H(\nu_{Fx})$ μ -a.e. for any surjective one-dimensional CA F , so the entropy of the sequence is preserved under evolution by the CA.

5.1.5 Localized information

There are many fundamental conserved quantities in physics, such as mass, charge, and energy. The flows of these quantities obey local continuity equations. For instance, in fluid dynamics the continuity equation for mass takes the form

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot (\rho \mathbf{u}),$$

where ρ is the density and \mathbf{u} is the fluid velocity. The term continuity does not refer to continuity in the topological sense, nor to continuity as the opposite of discreteness, but rather to “continuity of existence”.

In the previous section we discussed how entropy can be assigned to microstates that are coded as symbol sequences. An important question is whether we can introduce a *local* version of the microscopic entropy that is locally conserved under reversible dynamics, in the sense that it obeys a continuity equation involving an information current. Since the entropy depends on the degree of correlation in the

system, it is not obvious that this can be done. This approach was first proposed in the context of CA by Toffoli [154]. He considered perturbations around uncorrelated equilibria of particle conserving reversible CA, such as lattice gas automata. He looked at the one-site entropy and proved that the surplus or deficit of this quantity is transported as an additive conserved quantity.

5.2 Local information

One objective of the work presented in Papers III and IV is to explore to what extent information can be considered as a local quantity in dynamical systems of interacting units. In order to achieve this, we first define a local version of the Shannon entropy, called *local information*. The local information quantity is based on the interpretation of the Shannon entropy rate of a measure μ on $\mathcal{A}^{\mathbb{Z}}$ as the average information gained by reading the next symbol in a sequence x when knowledge of all previous symbols in some direction is assumed. The assumption of this knowledge is necessary, because the correlations in a sequence can be arbitrary long, and all correlations must be taken into account.

Definition 5.1. *Let μ be a probability measure on $\mathcal{A}^{\mathbb{Z}}$. The left local information with respect to μ is given by*

$$S_L(x; \mu) = - \lim_{n \rightarrow \infty} \log \mu(x_0 | x_{-n}^{-1}). \quad (5.6)$$

Furthermore, the left local information at coordinate i of $x \in \mathcal{A}^{\mathbb{Z}}$ is defined by $S_L(x; i; \mu) = S_L(\sigma^i x; \mu)$.

The function $S_L(x; i; \mu)$ is local in the sense that it measures the gain of information resulting from observing the symbol at a single position i of x . The following theorem establishes that S_L is well defined. For the proof, as well as the proofs of the other theorems stated in the chapter, we refer to Paper IV.

Theorem 5.1. *The sequence $(\log \mu(x_0 | x_{-n}^{-1}))_{n \geq 0}$ of functions in $L^1(\mu)$ converges μ -almost everywhere and in $L^1(\mu)$. Consequently, $S_L(x; \mu) \in L^1(\mu)$.*

A basic requirement for a local information quantity is that its average should equal the Shannon entropy rate. This is true for S_L , since from (5.3) and L^1 convergence it follows that $\int S_L(x; \mu) d\mu = h(\mu)$, where the integral is taken over $\mathcal{A}^{\mathbb{Z}}$. The local information $S_L(x; \mu)$ depends on the measure μ , thus knowledge of μ is required to calculate S_L . However, in the case of a shift-invariant measure, the left local information can in principle be recovered with probability one from the sequence $x_{-\infty}^0$. This is achieved by considering the empirical measure ν_x defined in (5.5).

Theorem 5.2. Let μ be a shift-invariant measure on $\mathcal{A}^{\mathbb{Z}}$ and ν_x the empirical measure generated by x . Then

$$S_L(x; \mu) = S_L(x; \nu_x) \quad \mu\text{-a.e.}$$

Analogously to the left local information, we define the *right local information* as

$$S_R(x; \mu) = - \lim_{n \rightarrow \infty} \log \mu(x_0 | x_1^n). \quad (5.7)$$

The right local information obviously has the same convergence properties as the left information, and for all results valid for the left information, there will be corresponding results for the right information. Note, however, that while the left and right information have the same expectation for all shift-invariant measures, they are not equal, nor do they in general have the same probability distribution.

5.3 Information transport

For S_L to be a reasonable measure of the information found at a single position, it should be *locally* conserved under reversible dynamics. Papers III and IV investigate this in the case where the dynamics is given by a cellular automaton. As illustrated in Sect. 3.3, CA have many of the properties characterizing physical systems, such as local interactions, and provide good models of many physical systems. Studies of local information and information transport can also provide insight into the behaviour of CA and quantitative measures of the structure which is built up. We return to these aspects in Sect. 5.4.

5.3.1 The information current

The goal is to show that the local information S_L satisfies a continuity equation involving an information current $J_L(x; i; \mu)$:

$$\Delta_t S_L = -\Delta_i J_L. \quad (5.8)$$

The operator Δ is the forward difference operator, so explicitly the terms are

$$\begin{aligned} \Delta_t S_L(x; i; \mu^t) &= S_L(F(x); i; \mu^{t+1}) - S_L(x; i; \mu^t), \\ \Delta_i J_L(x; i; \mu^t) &= J_L(x; i+1; \mu^t) - J_L(x; i; \mu^t). \end{aligned} \quad (5.9)$$

Since we consider time steps of length one, $J_L(x; i; \mu^t)$ can be interpreted as the information flow from position $i-1$ to position i generated by applying the CA.

In order to define the expression for the current we need two new notions. For a semi-infinite sequence $x_{-\infty}^i$, define $Z(x_{-\infty}^i)$ as the set of all semi-infinite sequences that have the same image and the same tail as $x_{-\infty}^i$:

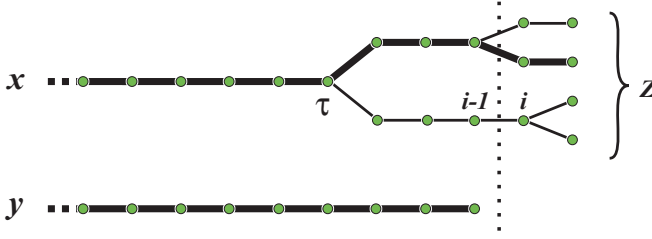


Fig. 5.1: An illustration of $Z = Z(x_{-\infty}^{i+r-1})$ and $\tau = \tau(x_{-\infty}^{i+r-1})$. In this case $r=2$, $|Z|=4$ and $\tau = i-4$. The circles represent symbols in \mathcal{A} . These are connected by lines to semi-infinite sequences which all map to the same $y_{-\infty}^{i-1}$ and coincide with x at all $j \leq \tau$.

Definition 5.2. For $x \in \mathcal{A}^{\mathbb{Z}}$ and a surjective CA F , define the sets $Z(x_{-\infty}^i)$ as

$$Z(x_{-\infty}^i) = \left\{ z_{-\infty}^j : f(z_{-\infty}^j) = f(x_{-\infty}^i) \text{ and } \exists j \leq i \text{ such that } z_{-\infty}^j = x_{-\infty}^j \right\}.$$

It follows from the results of [70] that $|Z(x_{-\infty}^i)|$ is bounded for any given surjective CA F . Furthermore, define $\tau(x_{-\infty}^i)$ as the largest index less than $i-r$ for which all sequences in $Z(x_{-\infty}^i)$ coincide (recall that r is the right radius of F):

Definition 5.3. For $x \in \mathcal{A}^{\mathbb{Z}}$, define $\tau(x_{-\infty}^i) \in \mathbb{Z}$ as

$$\tau(x_{-\infty}^i) = \max_j \left\{ j : j < i-r, \text{ and } z_{-\infty}^j = x_{-\infty}^j \forall z_{-\infty}^j \in Z(x_{-\infty}^i) \right\}.$$

We are now ready to define the information current. The integer τ and the set Z used in the definition are illustrated in Fig. 5.1.

Definition 5.4. Let F be a surjective one-dimensional CA with right radius r , and μ a measure on $\mathcal{A}^{\mathbb{Z}}$. Put $Z = Z(x_{-\infty}^{i+r-1})$ and $\tau = \tau(x_{-\infty}^{i+r-1})$. Define the left information current at coordinate i of x with respect to μ and F as

$$J_L(x; i; \mu) = -\log \mu(x_{\tau+1}^{i-1} | x_{-\infty}^{\tau}) + \log \sum_Z \mu(z_{\tau+1}^{i+r-1} | x_{-\infty}^{\tau}). \quad (5.10)$$

As for S_L , we write $J_L(x; \mu)$ for $J_L(x; 0; \mu)$. Theorem 5 in Paper IV ensures that $J_L \in L^1(\mu)$. In Paper IV we also describe a way of decomposing $J_L(x; \mu)$ into

$$J_L(x; \mu) = J_L^+(x; \mu) - J_L^-(x; \mu), \quad (5.11)$$

with $J_L^+, J_L^- \geq 0$, such that J_L^+ has a natural interpretation in terms of information flowing to the right between coordinates -1 and 0 , and J_L^- in terms of information flowing to the left. The component $J_L^-(x; \mu)$ is the additional information gained about the continuation x_1^∞ when knowing $x_{-\infty}^0$ and observing the configuration $(Fx)_{-\infty}^0$. This gain is measured using a relative entropy, also known as a Kullback-Liebler distance [86]. The component $J_L^+(x; \mu)$ is the information gained by observing $x_{\tau+1}^{-1}$ when $x_{-\infty}^{\tau}$ as well as $y_{-\infty}^{-1}$ is known, compare with Fig. 5.1.

5.3.2 Main results

The main results are two theorems that ascertain the correctness of the continuity equation in two different cases. In the case of a reversible CA, no requirements on the measure are necessary.

Theorem 5.3. *Let F be a reversible one-dimensional CA, and μ a probability measure on $\mathcal{A}^{\mathbb{Z}}$. Then $\Delta_t S_L(x; i; \mu) + \Delta_i J_L(x; i; \mu) = 0$ for all $i \in \mathbb{Z}$ μ -a.e.*

The second theorem states that the continuity equation also is valid for surjective CA, under the assumption that the measure is shift-invariant.

Theorem 5.4. *Let F be a surjective one-dimensional CA, and μ a shift-invariant probability measure on $\mathcal{A}^{\mathbb{Z}}$. Then $\Delta_t S_L(x; i; \mu) + \Delta_i J_L(x; i; \mu) = 0$ for all $i \in \mathbb{Z}$ μ -a.e.*

Corresponding results are of course also true for right local information. The proofs of both theorems are based on considering each sequence x as the outcome of a discrete stochastic process, and using results from martingale theory and uniform integrability.

A class of CA where the information flow has a particularly simple form is the class of permutative CA [70]. A one-dimensional CA F is right permutative if the map $a \mapsto f(x_{-m}^{t-1} a)$ is a permutation of \mathcal{A} for each $x_{-m}^{t-1} \in \mathcal{A}^{m+r}$. All right permutative CA are surjective.

Corollary 5.1. *Let μ be any shift-invariant measure on $\mathcal{A}^{\mathbb{Z}}$ and $F : \mathcal{A}^{\mathbb{Z}} \rightarrow \mathcal{A}^{\mathbb{Z}}$ a right permutative CA with right radius r . Then, μ -almost everywhere,*

$$S_L(Fx; i; \mu \circ F^{-1}) = S_L(x; i+r; \mu). \quad (5.12)$$

In particular, if $r = 0$, then $S_L(F^t x; i; \mu^t) = S_L(x; i; \mu^0)$ for all $t \geq 0$, so that the local information is locally constant.

5.3.3 Discussion

Our results show that the continuity equation is valid for reversible CA for all measures. This is a clear indication that the function S_L is an appropriate local information measure in a spatially extended system. Furthermore, the continuity equation is also valid for surjective CA, when the measure is shift-invariant.

There are although some limits to the locality of information when compared to physical locally conserved quantities such as mass. Firstly, the quantity $S_L(x; i; \mu)$ is not calculable from the local state x_i only, rather it depends on the semi-infinite context $x_{-\infty}^{i-1}$. This feature is however necessary to obtain a locally conserved information measure, because the correlations that are build up in the sequence during iteration can involve arbitrary many cells. However, there is a *choice* between the left and the right context, and this results in the existence of two independently

conserved local information quantities, S_L and S_R . These can behave quite differently. For instance, for a right permutative but not left permutative CA with $r = 0$, $S_L(x; i; \mu)$ is invariant under F for each i , while the dynamics of S_R is unaffected by right permutativity.

In Paper IV we also consider how far information can flow in a single iteration of the CA. It is true that $|J_L(x; i; \mu)| \leq \sum_{k=i}^{i+r-1} S_L(x; k; \mu)$, so there cannot flow more information to the left than the amount of information in the length r interval $[i, i + r - 1]$. Regarding flow of information to the right, we prove that for invertible CA, $J_L(x; i; \mu) < \sum_{k=i-\tilde{r}}^{i-1} S_L(x; k; \mu)$, where \tilde{r} is the right radius of the inverse CA. For surjective CA, $i - \tilde{r}$ is replaced by $\tau + 1$, which is not bounded. Thus, for surjective CA, local information can in one time step propagate farther than the radii of the CA, but only to the right. However, in light of the interpretation of J_L^- and J_L^+ , only flow of left information to the left is relevant to communication across the lattice. We refer to Paper IV for further discussion of this issue.

5.4 Correlation, structure and self-organization

In Sect. 5.1.3, we discussed how thermodynamical processes are macroscopically irreversible, because information is transferred to correlations among a large number of particles. The same phenomenon takes place in surjective CA when we consider the individual cells as the analogue of the particles. As an example we can consider evolution by the one-dimensional CA with neighbourhood and rule given by

$$N = \{0, 1\}, \quad \text{and} \quad f(x_0, x_1) = x_0 + x_1 \pmod{2}. \quad (5.13)$$

Let the initial measure be Bernoulli with $\mu^0(1) = \frac{1}{4}$, such that all coordinate random variables initially are independent. Since this CA is right permutative, the left local information is constant. But on the other hand, any initial Bernoulli-measure converges in the weak* sense in Cesàro mean to the uniform Bernoulli measure [98]. That is, for all $k \geq 1$ and all finite blocks $a_1^k \in \mathcal{A}^k$, we have

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=0}^{n-1} \mu^t(a_1^k) = \frac{1}{|\mathcal{A}|^k}. \quad (5.14)$$

The reason that both these results can be valid is that the information in the system is transferred to correlations over longer and longer distances as time progresses. Define the left local information conditioned on a finite number n of cells as $S_L^n(x; \mu) = -\log \mu(x_0 | x_1^- \dots x_n^-)$. Figure 5.2 displays the exact values of $S_L^n(x; i; \mu^t)$ for $0 \leq t < 20$ and various values of n for a part of the lattice in an evolution by the CA. As can be seen, the convergence of S_L is slow due to long correlations even for small times t . We also observe that the convergence is rather non-uniform with respect to position.

The details of the correlations in measures μ^t generated by CA have been studied by Lindgren [99], who looked at the information found in correlations at different

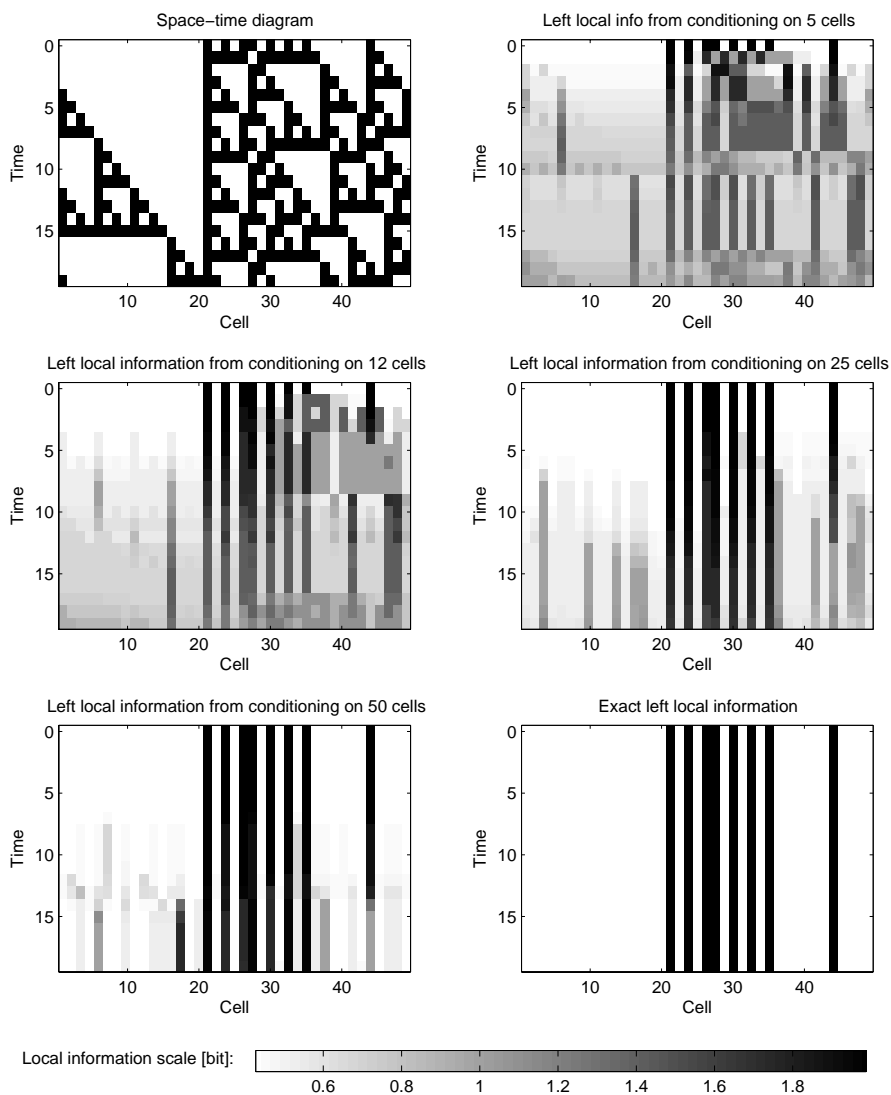


Fig. 5.2: The left local information from conditioning on finite number of cells, calculated during an evolution of the CA defined in (5.13). The initial measure μ^0 is a Bernoulli measure with $\mu^0(1) = \frac{1}{4}$. A part of a larger lattice is shown.

lengths, a concept introduced by Eriksson and Lindgren [47]. The information k_m in correlations of length m can be written in terms of local information conditioned on a finite number of cells as $k_m = \int (S_L^m - S_L^{m-1}) dx \geq 0$. With $k_0 = \log |\mathcal{A}| - H(X_0)$, called the density information, we obtain

$$\log |\mathcal{A}| = h(\mu) + \sum_{m=0}^{\infty} k_m.$$

Thus, the total information capacity of each lattice point is decomposed into entropy, density information, and correlations at different lengths. This is a macroscopic approach, in the sense that it only looks at expectancy values of S_L . Through the use of local information it could be possible to characterize and study the correlations developing in the system at the *microscopic* scale. For instance, even though $k_m \geq 0$ for each m , the quantity $S_L^m(x, \mu) - S_L^{m-1}(x; \mu)$ can vary wildly with m within a single sequence x , and take both signs. Also, one can investigate to what degree the correlations are homogenous in the system, or vary with position i .

One can also use local information to study self-organization and the emergence of coherent structures in the dynamics of CA and 2-CA. In Paper III we sketch how numerical estimates of S_L^n can be used to detect such structures. In particular we do this for ECA rule 18, described in Sect. 3.2.2. The defects in the lattice configuration can be singled out, because they contain more local information than the background pattern of the surrounding lattice. This will be the case for most localized coherent structures in CA, provided that they do not appear in larger numbers. To detect the structures, only a single sequence $F^t(x)$ from the evolution by F is necessary. A related approach is the local statistical complexity recently introduced by Shalizi et al. [140].

5.5 Future developments

It would be interesting in the future to investigate both numerically and analytically the flow of information in CA that have been studied in other contexts. This includes both the elementary CA and some of the CA related to physics and biology described in Sect. 3.3, such as the reversible CA with plausible thermodynamic behaviour investigated by Takesue [150], CA for simulating Ising dynamics [36, 159], and, in connection with computation, the billiard ball CA [105].

As pointed out at the end of Chap. 4, local information can be applied directly to 2-DCA, since these 2-CA are conjugate to ordinary CA. This is also the case for some basic instances of 2-CA with 2-morphologies. Nevertheless, it would be more interesting to apply local information in a way that takes higher order structures into account, and enables measurement of information transport both within each level and between the levels. This is left at a future problem.

In Paper IV we establish the existence of a continuity equation for one-dimensional surjective CA, and find a natural expression for the information current. We are

currently investigating the possibility of extending these results to non-surjective and probabilistic CA. For a non-surjective CA, $h(\mu^t)$ is in general a decreasing function of t . This involves a loss of local information, and a continuity equation would have to take this loss into account. Probabilistic CA are CA where the block map is stochastic [93], and can be regarded as a generalization of ordinary CA. The stochasticity may lead to an increase in the average local information during iteration — a further aspect which must be taken into account by a continuity equation. The continuity equation may also be extended to reversible and surjective CA in dimension two and higher.

One can also look at local information and transport of local information for other types of spatially extended dynamical systems. In particular, this includes coupled map lattices, which are similar to CA except that the local states are continuous variables, see, e.g., [79]. Extensions of the formalism to other systems might bring us closer to addressing fundamental issues relating to information transport and conservation in physical systems and the dynamical basis of self-organization.

Bibliography

- [1] ADLER, R., KONHEIM, A., AND MCANDREW, M. 1965. Topological entropy. *Trans. Amer. Math. Soc.* 114, 61–85.
- [2] ASHBY, W. R. 1962. Principles of the self-organizing system. In *Principles of Self-Organization*, H. von Foerster and G. W. Zopf, Eds. Pergamon Press, 255–278.
- [3] AXELROD, R. AND HAMILTON, W. D. 1981. The evolution of cooperation. *Science* 211, 1390–1396.
- [4] BAAS, N. A. 1994. Emergence, hierarchies and hyperstructures. In *Artificial Life III, Santa Fe Studies in the Sciences of Complexity, Proc. Volume XVII*, C. G. Langton, Ed. Addison-Wesley, Reading, MA, 515–537.
- [5] BAAS, N. A. 2004. Abstract Matter. Preprint NTNU.
- [6] BAAS, N. A., EHRESMANN, A., AND VANBREMEERSCH, J.-P. 2004. Hyperstructures and memory evolutive systems. *Int. J. Gen. Syst.* 33, 5, 553–568.
- [7] BAAS, N. A. AND EMMECHE, C. 1997. On emergence and explanation. *Intellectica* 2, 25, 67–83.
- [8] BALIAN, R. 2005. Information in statistical physics. *Stud. Hist. Philos. Mod. Phys.* 36B, 2, 323–353.
- [9] BÉAL, M.-P. AND PERRIN, D. 1997. Symbolic dynamics and finite automata. In *Handbook of formal languages, Vol. 2*. Springer, Berlin, 463–505.
- [10] BEKENSTEIN, J. 1973. Black holes and entropy. *Phys. Rev. D* 7, 2333–2346.
- [11] BENNETT, C. H. 1973. Logical reversibility of computation. *IBM J. Res. Develop.* 17, 6, 525–532.
- [12] BENNETT, C. H. 1982. The thermodynamics of computation – a review. *Int. J. Theor. Phys.* 21, 12, 905–940.
- [13] BENNETT, C. H. AND SHOR, P. W. 1998. Quantum information theory. *IEEE Trans. Inform. Theory* 44, 6, 2724–2742.

- [14] BERLEKAMP, E. R., CONWAY, J. H., AND GUY, R. K. 1982. *Winning Ways for your Mathematical Plays*. Vol. 2. Academic Press. Chapter 25.
- [15] BISHOP, T. AND CANNINGS, C. 1976. Models of animal conflict. *Adv. Appl. Prob.* 8, 616–621.
- [16] BLANCHARD, F., KURKA, P., AND MAASS, A. 1997. Topological and measure-theoretic properties of one-dimensional cellular automata. *Physica D* 103, 86–99.
- [17] BOCCARA, N. AND FUKS, H. 1998. Generalized deterministic traffic rules. *Int. J. Mod. Phys. C* 9, 1, 1–12.
- [18] BOLZMANN, L. 1877. Über die Beziehung zwischen dem zweiten Hauptsatze der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung, respektive den Sätzen über das Wärmeleichgewicht. *Wiener Berichte* 76, 373–435.
- [19] BONACICH, P., SURE, G. H., KAHAN, J. P., AND MEEKER, R. J. 1976. Cooperation and group size in the n -person prisoner's dilemma. *J. Conflict Resolut.* 20, 687–705.
- [20] BOYD, R. AND RICHERSON, P. J. 1988. The evolution of reciprocity in sizable groups. *J. Theor. Biol.* 132, 337–356.
- [21] BOYD, R. AND RICHERSON, P. J. 1992. Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethol. Sociobiol.* 13, 3, 171–195.
- [22] BOYLE, M. AND KITCHENS, B. 1999. Periodic points for onto cellular automata. *Indag. Math. (N.S.)* 10, 4, 483–493.
- [23] BRANDT, H., HAUERT, C., AND SIGMUND, K. 2003. Punishment and reputation in spatial public goods games. *Proc. R. Soc. Lond. B* 270, 1099–1104.
- [24] BRILLOUIN, L. 1956. *Science and Information Theory*. Academic Press, New York.
- [25] BROOM, M., CANNINGS, C., AND VICKERS, G. T. 1997. Multi-player matrix games. *Bull. Math. Biol.* 59, 5, 931–952.
- [26] BROOM, M., CANNINGS, C., AND VICKERS, G. T. 2000. Evolution in knockout conflicts: The fixed strategy case. *Bull. Math. Biol.* 62, 451–466.
- [27] BUKOWSKI, M. AND MIĘKISZ, J. 2004. Evolutionary and asymptotic stability in symmetric multi-player games. *Int. J. Game Theory* 33, 41–54.
- [28] BUSS, L. W. 1987. *The Evolution of Individuality*. Princeton University Press, Princeton, NJ.
- [29] CHAITIN, G. 1966. On the length of programs for computing finite binary sequences. *Journal of the ACM* 13, 547–569.
- [30] CHEN, S. Y., DIEMER, K., DOOLEN, D., EGGERT, K., FU, C., GUTMAN, S., AND TRAVIS, B. J. 1991. Lattice gas automata for flow through porous media. *Physica D* 47, 1–2, 72–84.

- [31] CHOU, H.-H. AND REGGIA, J. A. 1997. Emergence of self-replicating structures in a cellular automata space. *Physica D* 110, 3–4, 252–276.
- [32] CLAUSIUS, R. 1854. Über eine veränderte Form des zweiten Hauptsatzes der mechanischen Wärmetheorie. *Poggendorffs Annalen* 93, 480–506.
- [33] COOK, M. 2004. Universality in elementary cellular automata. *Complex Systems* 15, 1, 1–40.
- [34] COURCHAMP, F. AND MACDONALD, D. W. 2001. Crucial importance of pack size in the African wild dog *lycaon pictus*. *Anim. Conserv.* 4, 169–174.
- [35] COVER, T. M. AND THOMAS, J. A. 1991. *Elements of information theory*. Wiley Series in Telecommunications. John Wiley & Sons Inc., New York.
- [36] CREUTZ, M. 1986. Deterministic Ising dynamics. *Annals of Physics* 167, 62–72.
- [37] DAB, D., LAWNICZAK, A., BOON, J., AND KAPRAL, R. 1990. Cellular-automaton model for reactive systems. *Phys. Rev. Lett.* 64, 2462–2465.
- [38] DOEBELI, M. AND HAUERT, C. 2005. Models of cooperation based on the Prisoner's Dilemma and the Snowdrift game. *Ecology Letters* 8, 7, 748–766.
- [39] DURAND, B. 1998. Global properties of 2d-cellular automata. In *Cellular Automata and Complex Systems*, E. Golèz and S. Martinez, Eds. Kluwer.
- [40] DURRETT, R. AND STEIF, J. E. 1991. Some rigorous results for the Greenberg-Hastings model. *J. Theoret. Probab.* 4, 4, 669–690.
- [41] EHRESMANN, A. C. AND VANBREMEERSCH, J. P. 1987. Hierarchical evolutive systems: A mathematical model for complex systems. *Bull. Math. Bio* 49, 1, 13–50.
- [42] EHRESMANN, A. C. AND VANBREMEERSCH, J. P. 1996. Multiplicity principle and emergence in MES. *J. Syst. Anal. Model. Simulat.* 26, 81–117.
- [43] EIGEN, M. AND SCHUSTER, P. 1979. *The Hypercycle: A principle of natural self-organization*. Springer-Verlag, Berlin.
- [44] ELORANTA, K. 1994. The dynamics of defect ensembles in one-dimensional cellular automata. *J. Statist. Phys.* 76, 5–6, 1377–1398.
- [45] ELORANTA, K. AND NUMMELINA, E. 1992. The kink of cellular automaton Rule 18 performs a random walk. *J. Statist. Phys.* 69, 5–6, 1131–1136.
- [46] EMMECHE, C., KÖPPE, S., AND STJERNFELT, F. 1997. Explaining emergence: towards an ontology of levels. *J. Gen. Philos. Sci.* 28, 83–119.
- [47] ERIKSSON, K.-E. AND LINDGREN, K. 1987. Structural information in self-organizing systems. *Physica Scripta* 35, 388–397.

- [48] ERMENTROUT, G. B. AND EDELSTEIN-KESHET, L. 1993. Cellular automata approaches to biological modeling. *J. Theor. Biol.* 160, 1, 97–133.
- [49] FONTANA, W. AND BUSS, L. 1994. ‘The arrival of the fittest’: Toward a theory of biological organization. *Bull. Math. Biol.* 56, 1–64.
- [50] FRAME, L. H., MALCOLM, J. R., FRAME, G. W., AND LAWICK, H. V. 1979. Social organization of African wild dogs (*lycaon pictus*) on the Serengeti Plains, Tanzania, 1967–1978. *Zeitschrift für Tierpsychologie* 50, 225–249.
- [51] FREDKIN, E. 1990. Digital mechanics: An informational process based on reversible universal CA. *Physica D* 45, 254.
- [52] FREDKIN, E. AND TOFFOLI, T. 1982. Conservative logic. *Int. J. Theor. Phys.* 21, 219–253.
- [53] FRISCH, U., HASSLACHER, B., AND POMEAU, Y. 1986. Lattice-gas automata for the Navier-Stokes equation. *Phys. Rev. Lett.* 56, 1505–1508.
- [54] GÁCS, P. 1986. Reliable computation with cellular automata. *J. Comput. System Sci.* 32, 1, 15–78.
- [55] GÁCS, P. 2001. Reliable cellular automata with self-organization. *J. Statist. Phys.* 103, 1–2, 45–267.
- [56] GARDNER, M. 1983. *Wheels, Life, and Other Mathematical Amusements*. W. H. Freeman and Company, New York.
- [57] GIBBS, J. W. 1902. *Elementary principles of statistical mechanics*. Yale University Press.
- [58] GRASSBERGER, P. 1983. New mechanism for deterministic diffusion. *Phys. Rev. A* 28, 3666–3667.
- [59] GREENBERG, J. M. AND HASTINGS, S. P. 1978. Spatial patterns for discrete models of diffusion in excitable media. *SIAM J. Appl. Math.* 34, 3, 515–523.
- [60] HAIGH, J. AND CANNINGS, C. 1989. The n -person war of attrition. *Acta Appl. Math.* 14, 59–74.
- [61] HAMILTON, W. D. 1963. The evolution of altruistic behavior. *Am. Nat.* 97, 354–356.
- [62] HAMILTON, W. D. 1975. Innate social aptitudes of men: an approach from evolutionary genetics. In *Biosocial anthropology*, R. Fox, Ed. Malaby Press, London, 133–155.
- [63] HANSKI, I. 1999. *Metapopulation Ecology*. Oxford University Press, Oxford, UK.
- [64] HANSON, J. E. AND CRUTCHFIELD, J. P. 1992. The attractor-basin portrait of a cellular automaton. *J. Statist. Phys.* 66, 5–6, 1415–1462.

- [65] HANSON, J. E. AND CRUTCHFIELD, J. P. 1997. Computational mechanics of cellular automata: an example. *Physica D* 103, 169–189.
- [66] HARDIN, G. 1968. The tragedy of the commons. *Science* 162, 143–148.
- [67] HARDY, J., POMEAU, Y., AND DE PAZZIS, O. 1973. Time evolution of two-dimensional model system. I. invariant states and time correlation functions. *J. Math. Phys.* 14, 1746–1759.
- [68] HATTORI, T. AND TAKESUE, S. 1991. Additive conserved quantities in discrete-time lattice dynamics. *Physica D* 49, 295–322.
- [69] HAUERT, C., MONTE, S. D., AND SIGMUND, K. 2002. Replicator dynamics in optional public goods games. *J. Theor. Biol.* 218, 187–194.
- [70] HEDLUND, G. A. 1969. Endomorphisms and automorphisms of the shift dynamical system. *Math. Systems Theory* 3, 320–375.
- [71] HOFBAUER, J., SCHUSTER, P., AND SIGMUND, K. 1979. A note on evolutionarily stable strategies and game dynamics. *J. Theor. Biol.* 81, 609–612.
- [72] HOFBAUER, J. AND SIGMUND, K. 1998. *Evolutionary games and population dynamics*. Cambridge University Press.
- [73] HORDIJK, W., SHALIZI, C. R., AND CRUTCHFIELD, J. P. 2001. Upper bound on the products of particle interactions in cellular automata. *Physica D* 154, 3–4, 240–258.
- [74] ISING, E. 1925. Beitrag zur Theorie des Ferromagnetismus. *Z. Physik* 31, 253–258.
- [75] JAYNES, E. T. 1957a. Information theory and statistical mechanics. *Phys. Rev.* 106, 620–630.
- [76] JAYNES, E. T. 1957b. Information theory and statistical mechanics. II. *Phys. Rev.* 108, 171–190.
- [77] JAYNES, E. T. 1965. Gibbs vs Boltzmann entropies. *Am. J. Phys.* 33, 391–398.
- [78] KAGEL, J. H. AND ROTH, A. E. 1995. *The Handbook of Experimental Economics*. Princeton University Press, Princeton, NJ.
- [79] KANEKO, K. 1993. *Theory and applications of coupled map lattices*. John Wiley & Sons, New York.
- [80] KARI, J. 1990. Reversability of 2D cellular automata is undecidable. *Physica D* 45, 379–385.
- [81] KARI, J. 2005. Theory of cellular automata: A survey. *Theor. Comput. Sci.* 334, 3–33.

- [82] KNOSPE, W., SANTEN, L., SCHADSCHNEIDER, A., AND SCHRECKENBERG, M. A. 2000. Towards a realistic microscopic description of highway traffic. *J. Phys. A* 33, 48, L477–L485.
- [83] KOLMOGOROV, A. N. 1958. A new metric invariant of transient dynamical systems and automorphisms in Lebesgue spaces. *Dokl. Akad. Nauk. SSSR* 119, 861–864.
- [84] KOLMOGOROV, A. N. 1965. Three approaches to the quantitative definition of information. *Prob. Info. Trans.* 1, 1–7.
- [85] KOPELL, N. AND ERMENTROUT, G. 1990. Phase transitions and other phenomena in chains of coupled oscillators. *SIAM J. Math. Anal.* 50, 4, 1014–1052.
- [86] KULLBACK, S. AND LEIBLER, R. A. 1951. On information and sufficiency. *Ann. Math. Statistics* 22, 79–86.
- [87] KURKA, P. 1997. Languages, equicontinuity and attractors in cellular automata. *Ergodic Theor. Dynam. Syst.* 17, 2, 417–433.
- [88] KURKA, P. 2001. Topological dynamics of cellular automata. In *Codes, systems, and graphical models (Minneapolis, MN, 1999)*. IMA Vol. Math. Appl., vol. 123. Springer, New York, 447–485.
- [89] LANDAUER, R. 1961. Irreversibility and heat generation in the computing process. *IBM J. Res. Develop.* 5, 3, 183–191.
- [90] LANGTON, C. G. 1984. Self-reproduction in cellular automata. *Physica D* 10, 1–2, 135–144.
- [91] LANGTON, C. G. 1986. Studying artificial life with cellular automata. *Physica D* 22, 120–149.
- [92] LANGTON, C. G. 1989. Artificial life. In *Artificial Life*, C. Langton, Ed. Addison-Wesley, 1–47.
- [93] LEBOWITZ, J. L., MAES, C., AND SPEER, E. R. 1990. Statistical-mechanics of probabilistic cellular automata. *J. Statist. Phys.* 59, 1–2, 117–170.
- [94] LEIMAR, O. AND HAMMERSTEIN, P. 2001. Evolution of cooperation through indirect reciprocity. *Proc. R. Soc. Lond. B* 268, 745–753.
- [95] LEVINS, R. 1969. Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bull. Entomol. Soc. Am.* 15, 237–240.
- [96] LIGGETT, T. 1985. *Interacting Particle Systems*. Grundlehren der mathematischen Wissenschaften, vol. 276. Springer.
- [97] LIN, C. C. AND SEGEL, L. A. 1988. *Mathematics Applied to Deterministic Problems in the Natural Sciences*. Classics in Applied Mathematics, vol. 1. SIAM. Chapter 1.

- [98] LIND, D. A. 1984. Applications of ergodic theory and sofic systems to cellular automata. *Physica D* 10, 36–44.
- [99] LINDGREN, K. 1987. Correlations and random information in cellular automata. *Complex Systems* 1, 529–543.
- [100] LINDGREN, K. 1988. Microscopic and macroscopic entropy. *Phys. Rev. A* 38, 4794–4798.
- [101] LINDGREN, K. AND NORDAHL, M. G. 1994. Evolutionary dynamics of spatial games. *Physica D* 75, 292–309.
- [102] LLOYD, S. 2000. Ultimate physical limits to computation. *Nature* 406, 1047–54.
- [103] MAC LANE, S. 1971. *Categories for the working mathematician*. Number 5 in Graduate Texts in Mathematics. Springer.
- [104] MACK, G. 2001. Universal dynamics, a unified theory of complex systems. emergence, life and death. *Commun. Math. Phys.* 219, 141–178.
- [105] MARGOLUS, N. 1984. Physics-like models of computation. *Physica D* 10, 81–95.
- [106] MARGOLUS, N. 1993. A bridge of bits. In *Workshop on Physics of Computation - PhysComp'92*. IEEE Computer Society Press, 253–257.
- [107] MARTÍNEZ, G. J., ADAMATZKY, A., AND MCINTOSH, H. V. 2006. Phenomenology of glider collisions in cellular automaton Rule 54 and associated logical gates. *Chaos Solitons Fractals* 28, 1, 100–111.
- [108] MAYNARD-SMITH, J. 1982. *Evolution and the Theory of Games*. Cambridge University Press, Cambridge.
- [109] MAYNARD-SMITH, J. AND PRICE, G. 1973. The logic of animal conflict. *Nature* 246, 7716–7718.
- [110] MILINSKI, M., SEMMANN, D., AND KRAMBECK, H. J. 2002. Reputation helps solve the ‘tragedy of the commons’. *Nature* 415, 424–426.
- [111] MILNOR, J. 1988. On the entropy geometry of cellular automata. *Complex Systems* 2, 357–386.
- [112] MOORE, E. F. 1962. Machine models of self reproduction. In *Proc. Sympos. Appl. Math.* Vol. 14. American Mathematical Society, Providence, RI, 17–33.
- [113] MOTRO, U. 1991. Co-operation and defection: Playing the field and the ESS. *J. Theor. Biol.* 151, 145–154.
- [114] MYERSON, R. B. 1997. *Game Theory: Analysis of Conflict*. Harvard University Press, Cambridge, MA.
- [115] MYHILL, J. 1963. The converse to Moore’s garden-of-Eden theorem. *Proc. Am. Math. Soc.* 14, 685–686.

- [116] NAGEL, K. AND SCHRECKENBERG, M. A. 1992. A cellular automaton model for freeway traffic. *Journal de Physique I* 2, 12, 2221–2229.
- [117] NASH, J. F. 1951. Noncooperative games. *Annals of Mathematics* 54, 289–295.
- [118] NASU, M. 1978. Local maps inducing surjective global maps of one-dimensional tessellation automata. *Math. Systems Theory* 11, 327–351.
- [119] NEWMAN, M. E. J. 2003. The structure and function of complex networks. *SIAM Review* 410, 167–256.
- [120] NEWMAN, M. E. J. AND BARKEMA, G. T. 1999. *Monte Carlo Methods in Statistical Physics*. Oxford University Press.
- [121] NOWAK, M. A. AND MAY, R. 1992. Evolutionary games and spatial chaos. *Nature* 359, 826–829.
- [122] NOWAK, M. A. AND SIGMUND, K. 2000. Games on grids. In *The Geometry of Ecological Interactions*, U. Dieckmann, R. Law, and J. Metz, Eds. Cambridge Studies in Adaptive Dynamics, vol. 1. Cambridge University Press, Chapter 8.
- [123] OLLINGER, N. 2002. The quest for small universal cellular automata. In *Proc. of ICALP 2002, 29th Internat. Colloq. on Automata, Languages and Programming*. Lecture Notes in Computer Science, vol. 2380, Springer, Berlin, 318–329.
- [124] PALM, G. 1984. Evolutionarily stable strategies and game dynamics for n -person games. *J. Math. Biol.* 19, 329–334.
- [125] PARK, J. K., STEIGLITZ, K., AND THURSTON, W. P. 1986. Soliton-like behavior in automata. *Physica D* 19, 423–432.
- [126] PARTHASARATHY, K. 1977. *Introduction to probability and measure*. MacMillan Co. of India Ltd.
- [127] PETRINA, D. Y., GERASIMENKO, V. I., AND MALYSHEV, P. V. 2002. *Mathematical foundations of classical statistical mechanics*, Second ed. Advanced Studies in Contemporary Mathematics, vol. 8. Taylor & Francis, London.
- [128] PIVATO, M. 2002. Conservation laws in cellular automata. *Nonlinearity* 15, 6, 1781–1793.
- [129] PIVATO, M. AND YASSAWI, R. 2002. Limit measures for affine cellular automata. *Ergodic Theor. Dynam. Syst.* 22, 4, 1269–1287.
- [130] POMEAU, Y. 1984. Invariants in cellular automata. *J. Phys. A* 17, 415–418.
- [131] RASMUSSEN, S., BAAS, N. A., MAYER, B., NILSSON, M., AND OLESEN, M. W. 2001. Ansatz for dynamical hierarchies. *Artificial Life* 7, 329–353.
- [132] RICHARDSON, D. 1972. Tessellations with local transformations. *J. Comput. System Sci.* 5, 373–388.

- [133] RIVET, J.-P. AND BOON, J. P. 2001. *Lattice gas hydrodynamics*. Cambridge Nonlinear Science Series, vol. 11. Cambridge University Press, Cambridge.
- [134] ROBERTS, G. 1996. Why individual vigilance declines as group size increases. *Anim. Behav.* 51, 1077–1086.
- [135] ROTHMAN, D. H. AND KELLER, J. M. 1988. Immiscible cellular-automaton fluids. *J. Statist. Phys.* 52, 3–4, 1119–1127.
- [136] SAITO, K., TAKESUE, S., AND MIYASHITA, S. 1999. Transport phenomena at a critical point: Thermal conduction in the Creutz cellular automaton. *Phys. Rev. E* 59, 3, 2783–2794.
- [137] SARKAR, P. 2000. A brief history of cellular automata. *ACM Computing Surveys* 2000, 80–107.
- [138] SCOTT, A. 1995. *Stairway to the Mind*. Springer-Verlag.
- [139] SHALIZI, C. R. 2005. Methods and techniques of complex systems science: An overview. Forthcoming in *Complex Systems Science in Biomedicine*, T. S. Deisboeck and J. Y. Kresh, Eds. Springer. Available at <http://arxiv.org/abs/nlin.A0/0307015>.
- [140] SHALIZI, C. R., HASLINGER, R., ROUQUIER, J.-B., KLINKNER, K. L., AND MOORE, C. 2005. Automatic filters for the detection of coherent structure in spatiotemporal systems. To be published in *Phys. Rev. E*. Available at <http://arxiv.org/abs/nlin.CG/0508001>.
- [141] SHANNON, C. E. 1948. A mathematical theory of communication. *Bell System Tech. J.* 27, 379–423, 623–656.
- [142] SHAW, R. 1981. Strange attractors, chaotic behavior, and information-flow. *Zeitschrift für Naturforschung* 36A, 80–112.
- [143] SLATKIN, M. 1974. Competition and regional coexistence. *Ecology* 55, 128–134.
- [144] SMITH, J. M. AND SZATHMÁRY, E. 1995. *The Major Transitions in Evolution*. Oxford University Press.
- [145] SOLOMONOFF, R. J. 1964. A formal theory of inductive inference. *Information and Control* 7, 1–22.
- [146] STEIF, J. E. 1995. Two applications of percolation to cellular automata. *J. Statist. Phys.* 78, 5–6, 1325–1335.
- [147] STEIGLITZ, K., KAMAL, I., AND A., W. 1988. Embedding computation in one-dimensional automata by phase coding solitons. *IEEE Trans. Comput.* 37, 2, 138–145.
- [148] STROGATZ, S. H. 2001. Exploring complex networks. *Nature* 410, 268–276.

- [149] SUTNER, K. 1999. Linear cellular automata and de Bruijn automata. In *Cellular Automata: A Parallel Model*, M. Delorme and J. Mazoyer, Eds. Kluwer.
- [150] TAKESUE, S. 1987. Reversible cellular automata and statistical mechanics. *Phys. Rev. Lett.* 59, 2499–2502.
- [151] TAKESUE, S. 1989. Ergodic properties and thermodynamics behavior of elementary reversible cellular automata. I. Basic properties. *J. Statist. Phys.* 56, 3, 371.
- [152] TAKESUE, S. 1990. Fourier's law and the Green-Kubo formula in a cellular-automaton model. *Phys. Rev. Lett.* 64, 252–255.
- [153] TAYLOR, P. AND JONKER, L. 1978. Evolutionarily stable strategies and game dynamics. *Math. Biosci.* 40, 145–156.
- [154] TOFFOLI, T. 1988. Information transport obeying the continuity equation. *IBM J. Res. Develop.* 32, 1, 29–36.
- [155] TOFFOLI, T. AND MARGOLUS, N. 1987. *Cellular Automata Machines*. MIT Press, London.
- [156] TOFFOLI, T. AND MARGOLUS, N. H. 1990. Invertible cellular automata: a review. *Physica D* 45, 1–3, 229–253.
- [157] TRIVERS, R. 1971. The evolution of reciprocal altruism. *Q. Rev. Biol.* 46, 33–57.
- [158] TURING, A. M. 1952. The chemical basis of morphogenesis. *Phil. Trans. R. Soc. Lond. B* 237, 37–72.
- [159] VICHNIAC, G. 1984. Simulating physics with cellular automata. *Physica D* 10, 96–115.
- [160] VON NEUMANN, J. 1955. *Mathematical Foundations of Quantum Mechanics*. Princeton University Press, Princeton, NJ.
- [161] VON NEUMANN, J. 1966. *Theory of Self-Reproducing Automata*. University of Illinois Press, Urbana, Illinois.
- [162] VON NEUMANN, J. AND MORGENSTERN, O. 1944. *Theory of Games and Economic Behaviour*. Princeton University Press, NJ.
- [163] WALTERS, P. 1982. *An Introduction to Ergodic Theory*. Number 79 in Graduate Texts in Mathematics. Springer.
- [164] WEIBULL, J. 1995. *Evolutionary Game Theory*. MIT Press, Cambridge, MA.
- [165] WEIMAR, J. R., TYSON, J. J., AND WATSON, L. T. 1992. Third generation cellular automaton for modeling excitable media. *Physica D* 55, 3–4, 328–339.

-
- [166] WHEELER, J. 1988. *Information, Physics, Quantum: The Search for Links, in Complexity, Entropy and the Physics of Information*. Addison-Wesley, Redwood City, CA.
- [167] WOLFRAM, S. 1983. Statistical mechanics of cellular automata. *Rev. Modern Phys.* 55, 601–644.
- [168] WOLFRAM, S. 1984. Universality and complexity in cellular automata. *Physica D* 10, 1–35.
- [169] ZHABOTINSKII, A. M. AND ZAIKIN, A. N. 1970. Concentration waves propagating in two-dimensional liquid-phase self-oscillating systems. *Nature* 225, 535–537.
- [170] ZUREK, W. 1989. Algorithmic randomness and physical entropy. *Phys. Rev. A* 40, 4731–4751.

Paper I

The Evolution of n -player Cooperation — Threshold Games and ESS Bifurcations

Lars A. Bach, Torbjørn Helvik and Freddy B. Christiansen (2005)

Journal of Theoretical Biology, In press.

doi:10.1016/j.jtbi.2005.06.007

The Evolution of n -player Cooperation — Threshold Games and ESS Bifurcations

Lars A. Bach^{1,2}, Torbjørn Helvik³, and Freddy B. Christiansen²

¹Department of Theoretical Ecology
Lund University
SE-223 62 Lund, Sweden

²Department of Ecology and Genetics
University of Aarhus
DK-8000 Aarhus C, Denmark

³Department of Mathematical Sciences
Norwegian University of Science and Technology
NO-7491 Trondheim, Norway

Abstract

An evolutionary game of individuals cooperating to obtain a collective benefit is here modelled as an n -player Prisoner's Dilemma game. With reference to biological situations, such as group foraging, we introduce a threshold condition in the number of cooperators required to obtain the collective benefit. In the simplest version, a three-player game, complex behaviour appears as the replicator dynamics exhibits a catastrophic event separating a parameter region allowing for coexistence of cooperators and defectors and a region of pure defection. Cooperation emerges through an ESS bifurcation, and cooperators only thrive beyond a critical point in cost-benefit space. Moreover, a repelling fixed point of the dynamics acts as a barrier to the introduction of cooperation in defecting populations. The results illustrate the qualitative difference between two-player games and multiple player games and thus the limitations to the generality of conclusions from two-player games. We present a procedure to find the evolutionarily stable strategies in any n -player game with cost and benefit depending on the number of cooperators. This was previously done by Motro [1991. Co-operation and defection: playing the field and the ESS. *J. Theor. Biol.* 151, 145–154] in the special cases of convex and concave benefit functions and constant cost.

Keywords: n -Player game; Prisoner's Dilemma; cooperation; bifurcation; hysteresis

1 Introduction

Game theoretical analysis has been widely applied in evolutionary theory. The canonical metaphor for the dilemma arising when cooperative and competitive relations between individuals collide is the Prisoner's Dilemma [16, 21, 24, 32]. The Prisoner's Dilemma game can either be represented as a two-player game or more generally as an n -player game, sometimes termed the 'public goods game' [6]. The n -player game describes a situation in which several individuals can cooperate to achieve a common benefit, which in turn is shared among both the cooperators and the (free riding) defectors in the group. An often used 'tragic' sociological metaphor — the tragedy of the commons — describes the overexploitation of a grassland resource shared by village farmers [18]. Indeed this metaphor represents the rational outcome of an n -player game. The verbalized public goods game illustrates a situation where the social agents contribute some amount of resource (money) to a common pool, which is then increased with some factor (interest rate). Subsequently the total quantity is distributed among all players regardless of their individual contribution [19]. In its pure form the outcome is identical to the tragedy of the commons. However, cooperation can be promoted by evoking further mechanisms such as e.g. punishment [7, 15], or optional participation either in groups [19] or in iterated two-player games [4].

A crucial notion to analytically describe the properties and ultimate outcome of a conflict or a game situation is the Evolutionarily Stable Strategy (ESS). Roughly, a strategy is evolutionarily stable if no alternative strategy can invade it (see [1, 21, 24, 27]). In evolutionary theory, however, the ESS analysis is mainly applied in two-player games only (but see e.g. [6, 19]).

As in the public goods game the per capita benefit is often implicitly assumed to increase in a linear fashion with the number of cooperators, i.e. the effect of cooperation is additive. In an n -player game, Motro [25] relaxed this assumption and investigated two further classes of strictly increasing benefit functions describing the benefit obtained by group members as a function of the number of cooperators. For linear (additive) and convex (superadditive) benefit functions only the two trivial ESSs of full defection or full cooperation appear, depending on whether the additional benefit obtained by a single member switching to cooperative behaviour exceeds the individual cost of cooperation [25]. However, for concave (subadditive) benefit functions there may exist a polymorphic ESS consisting of both strategies (or alternatively a homogenous population of mixed strategies). Hence, under such condition some degree of cooperation in the population can be expected [25].

In various natural situations the assumption of the benefit function having a monotonic derivative, let alone being linear, sometimes fails to capture the nature of the interaction. Animals that are dependent on very aggregated resources that require joint effort to handle and process, may face a reality best represented by a threshold scenario. In such case the benefit represents a single discrete all-or-nothing event, thus the benefit function becomes a step function. In other words, the benefit function

may be characterised by a discrete transition between a plateau of low or zero benefit and a plateau of high benefit that is reached only if sufficiently many group members cooperate. A conspicuous natural situation complying with this model is the group hunting of certain predators, as for example the African wild dog (*Lycaon pictus*) [5, 11–14]. The collective effort of individuals in the groups ensures that large prey can be caught that no single individual could ever capture [14]. Moreover, the hunting success seemingly depends on the propensity to cooperate among group members. A minimum number of hunters seems required in order to capture large prey items as the fatigue of the prey cannot be provoked without a joint effort of a number of individuals [13]. As the prey is either captured or not, a game representation should account for the binary outcome.

Territory defence is another example of a benefit function which should be represented by a threshold scenario since the result of the joint defence is of a binary nature in the sense that the territory will either be lost or maintained. Game theoretical considerations has arisen from the observation that in female groups of lions (*Panthera leo*) certain individuals seem to consistently refrain from contributing to the common territory defence [20, 29].

In this paper we wish to investigate the simplest possible game allowing for a step-wise or threshold relation between the proportion of cooperators in a group and the benefit obtained. In a three-player game we introduce a threshold level of two cooperators that has to be exceeded in order to obtain the benefit. We describe this game in Section 2, and show that even this simple case yields complex evolutionary behaviour as the system exhibits hysteresis around critical parameter values separating a regime with stable coexistence of defectors and cooperators and a regime of pure defection. In Section 3 we generalize Motro’s work and describe a procedure to find all ESSs for any benefit function. We then proceed to find conditions for when the dynamics is similar to that of our three player threshold game. In Section 4 we discuss the implications and biological relevance of our findings.

2 Single group and mean field models

Initially, assume we have a group of three players. Each player has the choice of cooperating (C) or defecting (D). If at least two players cooperate, all three players will receive a benefit of r . Otherwise, all players will receive no benefit. Additionally, cooperating bears some cost c , which is inflicted whether or not the benefit is achieved. The payoff to ego is summarized by the table:

	CC	CD	DD
C	$r - c$	$r - c$	$-c$
D	r	0	0

We see that the pure strategy of always defecting is a Nash equilibrium while the pure strategy of cooperating never is. Let the mixed strategy x be to cooperate with

probability x , and defect with probability $1 - x$. Assume that two of the players play strategy x . If player A plays strategy y , its expected payoff is given by the function:

$$W(y; x) = rx^2 + y(2rx(1 - x) - c) = rx^2 + y \cdot g(x). \quad (1)$$

According to classical game theory, there is a *mixed* Nash equilibrium when this payoff is independent of y . In our case this is when $g(x) = 0$. This is the case only for the two x values

$$x_- = \frac{1}{2} - \frac{1}{2} \sqrt{1 - 2\frac{c}{r}}, \quad x_+ = \frac{1}{2} + \frac{1}{2} \sqrt{1 - 2\frac{c}{r}}. \quad (2)$$

Thus, there are only mixed Nash equilibria when $r \geq 2c$.

The same game can also be regarded as an evolutionary game. Here, the setting is a large population where triples of players are randomly selected and play the game. A strategy x is said to be evolutionarily stable if it is resistant to all invading strategies. That is, if any alternative strategy y is played by a sufficiently small fraction of the population, x always does better than y . The notion of ESS was introduced in [23] for two-player games. We use the natural extension to n -player games defined in [8], namely the condition that

$$W(x; \varepsilon y + (1 - \varepsilon)x) > W(y; \varepsilon y + (1 - \varepsilon)x) \quad (3)$$

for all $y \in [0, 1]$, $y \neq x$ and ε smaller than some $\varepsilon(y)$. By (1) and Taylor expansion we have

$$\begin{aligned} W(x; \varepsilon y + (1 - \varepsilon)x) - W(y; \varepsilon y + (1 - \varepsilon)x) \\ &= (x - y)g(x + \varepsilon(y - x)) \\ &= (x - y)g(x) - \varepsilon(x - y)^2 g'(x) + O(\varepsilon^2). \end{aligned} \quad (4)$$

Hence, for a mixed strategy x to be an ESS it must satisfy $g(x) = 0$ (Nash equilibrium) and additionally $g'(x) < 0$. Consequently, x_+ in (2) is an ESS while x_- is not. The Nash equilibrium where players always defect is also an ESS, since always defecting in this case is the unique best reply to itself. Note that while having a mixed ESS and a pure strategy ESS coexisting for the same parameter values is impossible for a two-player game, it is not in a three-player game [8].

The replicator equation for the system determines the dynamics of the game when players reproduce proportionally to their achieved payoff. The equation is given by [21]:

$$\dot{x} = x(f(x) - \bar{f}). \quad (5)$$

Here, $f(x)$ is the fitness of cooperators and \bar{f} is the mean fitness of the total population. In our case the exact form is given by

$$\dot{x} = x(1 - x)g(x). \quad (6)$$

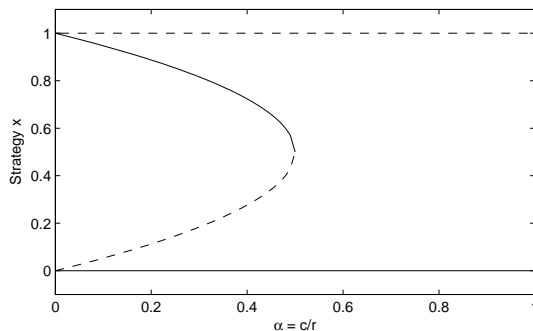


Fig. 1: The bifurcation diagram for the replicator system (6). Fixed points are plotted versus the parameter $\alpha = \frac{c}{r}$. Solid-drawn lines represent stable fixed points. These are ESSs for the system. Dashed lines represent unstable fixed points.

The fixed points of this dynamical system are the same as the Nash equilibria of the group game, along with the point $x = 1$, since pure populations always will be fixed points of the replicator system. It is easily checked that the stable fixed points are those corresponding to ESSs for the group game. We can thus represent the dynamics of the game by the bifurcation diagram in Fig. 1. The diagram shows that we have a saddle-node bifurcation at $\alpha = \frac{c}{r} = \frac{1}{2}$. Here, one stable and one unstable fixed point are created as α decreases. Consequently, in the case that $\alpha < \frac{1}{2}$ the system will settle in state x_+ if the initial fraction of cooperators is higher than x_- . Otherwise, the system will settle in a state consisting entirely of defectors. The last outcome is always the case when $\alpha > \frac{1}{2}$. Note that this leads to a hysteresis effect. If the system initially is in the cooperating state x_+ and α is increased beyond $\frac{1}{2}$, the system will settle in the state of pure defection. However, a subsequent decrease of α will not suffice to reestablish the cooperative state due to the fact that $x = 0$ also is an ESS. This effect is discussed in Section 4. The dynamics of threshold games with more than three players are similar, see Example 6.

3 General conditions for threshold type dynamics

Considering the general case, we wish to find a general procedure to understand the dynamics of an n -player evolutionary game as well as obtain conditions under which the game will have dynamics equal to the dynamics of our three player threshold model. Assume we have a group of N players (capital N is hereafter used to indicate a fixed group size). Denote by r_k the benefit to each group member if k players opt to cooperate and $N - k$ players defect. Furthermore, assume that cooperation bears an additional cost of c_k when k of the players cooperate, that is, the cost is specified by the vector $\mathbf{c} = (c_1, \dots, c_N)$. Define $\mathbf{r} = (r_0, \dots, r_N)$ as the vector containing the benefits.

It may be natural that r_k is non-decreasing and c_k is non-increasing in k , but we do not require this. This game is formally defined as

Definition 1. Let the game $\Gamma(N, \mathbf{r}, \mathbf{c})$, with $\mathbf{c} \in [0, \infty)^N$ and $\mathbf{r} \in [0, \infty)^{N+1}$, have N players, each with the strategy set $S = \{C, D\}$. The payoff is defined as following: If k players play C , these will obtain $r_k - c_k$ and the remaining $N - k$ players will obtain r_k .

The game with constant cost considered by Motro [25] is referred to as $\Gamma_c(N, \mathbf{r}) = \Gamma(N, \mathbf{r}, (c, c, \dots, c))$. Before proceeding we introduce a useful terminology.

Definition 2. The forward difference operator Δ is defined on the sequence \mathbf{r} by

$$\Delta r_k = r_{k+1} - r_k$$

for $0 \leq k \leq N - 1$. Its higher iterates are defined recursively by

$$\Delta^m r_k = \Delta^{m-1} r_{k+1} - \Delta^{m-1} r_k$$

for $0 \leq k \leq N - m$.

Note that $\Delta r_k - c_{k+1}$ is the additional payoff ego will obtain by switching from D to C if exactly k other players cooperate. Denote by $W_{\mathbf{r}, \mathbf{c}}(y; x)$ the expected payoff to ego playing strategy y while the remaining $N - 1$ players are playing strategy x . Equivalently, the remaining $N - 1$ players might be drawn at random from a large population of which a fraction x always plays C and the rest plays D . In both cases the number of players in the group playing C will be binomially distributed. Consequently, $W_{\mathbf{r}, \mathbf{c}}(y; x)$ is given by

$$W_{\mathbf{r}, \mathbf{c}}(y; x) = \sum_{k=0}^{N-1} \binom{N-1}{k} x^k (1-x)^{N-1-k} r_k + y \cdot \mathbf{g}_{\mathbf{r}, \mathbf{c}}(x), \quad (7)$$

where

$$\mathbf{g}_{\mathbf{r}, \mathbf{c}}(x) = \sum_{k=0}^{N-1} \binom{N-1}{k} x^k (1-x)^{N-1-k} (\Delta r_k - c_{k+1}). \quad (8)$$

The function $\mathbf{g}_{\mathbf{r}, \mathbf{c}}(x)$ is a polynomial defined on the interval $[0, 1]$. It is called the *gain function* because $\mathbf{g}_{\mathbf{r}, \mathbf{c}}(x)$ is interpreted as the expected increase in payoff ego will gain if playing C rather than D . As argued in the previous section, a strictly mixed strategy x can only be a Nash equilibrium if $\mathbf{g}_{\mathbf{r}, \mathbf{c}}(x) = 0$.

For the constant-cost game $\Gamma_c(N, \mathbf{r})$ Motro [25] found the possible ESS for the special case of all $\Delta^2 r_i$ having the same sign. In this case, the gain function is strictly monotonic, as implied by Eq. (17). We will refer to the gain function of $\Gamma_c(N, \mathbf{r})$ as $\tilde{\mathbf{g}}_{\mathbf{r}, c}(x)$.

Example 3. Consider the game where the cooperators share a fixed cost c , that is, $\mathbf{c} = c\sigma$ with $c \in [0, \infty)$ and $\sigma = (1, \frac{1}{2}, \frac{1}{3}, \dots, \frac{1}{N})$. The function $\mathbf{g}_{\mathbf{r}, \mathbf{c}}(x)$ is then given by

$$\mathbf{g}_{\mathbf{r}, c\sigma}(x) = \tilde{\mathbf{g}}_{\mathbf{r}, c}(x) + c(1 - \gamma(x)),$$

where $\tilde{g}_{\mathbf{r},c}(x)$ corresponds to the game $\Gamma_c(N, \mathbf{r})$ and

$$\gamma(x) = \frac{1 - (1-x)^N}{Nx} \quad \text{and} \quad \gamma(0) = 1.$$

In $\Gamma_c(N, \mathbf{r})$ with $r_k = r_0 + kc$, $k = 1, 2, \dots, N$, we have $\tilde{g}_{\mathbf{r},c}(x) \equiv 0$ (i.e., all strategies are Nash equilibria). With cost sharing and

$$r_k = r_0 + c \sum_{i=1}^k \frac{1}{i},$$

we again have $g_{\mathbf{r},c}(x) \equiv 0$. A particular Nash equilibrium may thus exist with a quite moderate increase in reward in the cost-sharing game, as compared the constant-cost game.

We will now describe a general procedure to locate the Nash equilibria and ESS of the game defined by \mathbf{r} and \mathbf{c} . These properties of the game are decided by the function $g_{\mathbf{r},c}(x)$ alone. In the proposition below, Cases (1)–(3) are the common cases while (4) contains special cases which occur if the derivative $g'_{\mathbf{r},c}(x)$ is zero or, for the endpoints, $g_{\mathbf{r},c}(x)$ is zero.

Proposition 4. Consider the function $g_{\mathbf{r},c}(x)$ defined in (8). All symmetric Nash equilibria (NE) and all ESS to the game $\Gamma(N, \mathbf{r}, \mathbf{c})$ in Definition 1 are given by the following cases:

- (1) If $g_{\mathbf{r},c}(0) < 0$, then $x = 0$ is an ESS.
- (2) If $g_{\mathbf{r},c}(1) > 0$, then $x = 1$ is an ESS.
- (3) If $g_{\mathbf{r},c}(x) = 0$, then x is a NE. If additionally $g'_{\mathbf{r},c}(x) < 0$, then x is also an ESS.
- (4) x is an ESS in the special cases:
 - (a) $0 < x < 1$, $g_{\mathbf{r},c}(x) = 0$, $g'_{\mathbf{r},c}(x) = 0$ and there is an $m \in \mathbb{N}$ such that $g_{\mathbf{r},c}^{(k)}(x) = 0$ for $k \leq 2m$ and $g_{\mathbf{r},c}^{(2m+1)}(x) < 0$.
 - (b) $x = 0$, $g_{\mathbf{r},c}(0) = 0$ and there is an $m \in \mathbb{N}$ such that $g_{\mathbf{r},c}^{(k)}(0) = 0$ for $k < m$ and $g_{\mathbf{r},c}^{(m)}(0) < 0$.
 - (c) $x = 1$, $g_{\mathbf{r},c}(1) = 0$ and there is an $m \in \mathbb{N}$ such that $g_{\mathbf{r},c}^{(k)}(1) = 0$ for $k < m$ and $(-1)^{m+1} g_{\mathbf{r},c}^{(m)}(1) < 0$.

The analysis is easily done graphically by plotting the function $g_{\mathbf{r},c}(x)$. Note that $g_{\mathbf{r},c}(x) = 0$ is satisfied at no more than $N - 1$ values of x , since $g_{\mathbf{r},c}(x)$ is a polynomial of degree $N - 1$. Also note that $g_{\mathbf{r},c}(0) = \Delta r_0 - c_1$ and $g_{\mathbf{r},c}(1) = \Delta r_{N-1} - c_N$. The proof of Proposition 4 is found in the appendix, as are the proofs for the propositions stated below.

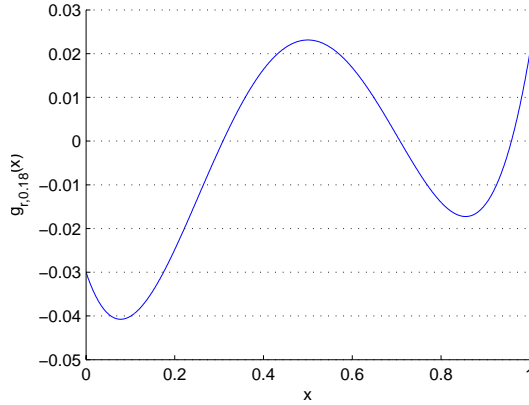


Fig. 2: The function $g_{\mathbf{r},c}(x)$ for $c=0.18$ and \mathbf{r} from Example 5. We see that one mixed ESS is present, and located at $x=0.708$. Also, $x=0$ and $x=1$ are ESSs.

Example 5. Consider the five player game $\Gamma_c(5, \mathbf{r})$ with $r_0 = 0, r_1 = 0.15, r_2 = 0.3, r_3 = 0.7, r_4 = 0.8, r_5 = 1$. The function $g_{\mathbf{r},c}(x)$ is plotted for $c = 0.18$ in Fig. 2. We have the following ESS for various values of c :

Range of c	ESS
$c < 0.139$	$x = 1$
$0.139 < c < 0.15$	one mixed ESS, $x = 1$
$0.15 < c < 0.163$	$x = 0, x = 1$
$0.163 < c < 0.2$	$x = 0$, one mixed ESS, $x = 1$
$0.2 < c < 0.203$	$x = 0$, one mixed ESS
$0.203 < c$	$x = 0$

Hence this five-player case generates the following ESS-bifurcations: at $c = 0.139, 0.163$ and 0.203 saddle-node bifurcations emerge, and at $c = 0.15$ and 0.2 transcritical bifurcations separate the regions of distinct evolutionary dynamics.

Example 6. Consider a threshold game $\Theta_{r,c}(N, M)$ with N players, where a threshold of M cooperators brings a reward of r per player and the cost of cooperating is constant and equal to c . The benefit is thus $r_k = 0$ for $k < M$ and $r_k = r$ for $k \geq M$. In this case,

$$\tilde{g}_{\mathbf{r},c}(x) = r \binom{N-1}{M-1} x^{M-1} (1-x)^{N-M} - c. \quad (9)$$

For $1 < M < N$, this function has a single maximum in $[0, 1]$, located at $\tilde{x} = \frac{M-1}{N-1}$. Define

$$\Upsilon_{N,M} = \binom{N-1}{M-1} \left(\frac{M-1}{N-1} \right)^{M-1} \left(\frac{N-M}{N-1} \right)^{N-M}. \quad (10)$$

The bifurcation diagram for the system will look qualitatively similar to that in Fig. 1. For $\frac{c}{r} < \gamma_{N,M}$ the system has a mixed ESS x_+ and an unstable mixed NE $x_- < x_+$. For $\frac{c}{r} > \gamma_{N,M}$ the pure strategy of defecting is the only ESS.

Finally, we state some general conditions under which the dynamic of a game $\Gamma_c(N, \mathbf{r})$ is equal to that of our three player threshold game presented in Section 2.

Proposition 7. Consider the game $\Gamma_c(N, \mathbf{r})$. If

- (1) $\Delta^3 r_i \leq 0$ for $0 \leq i \leq N-3$
- (2) $\Delta^2 r_0 > 0$ and $\Delta^2 r_{N-2} < 0$

then there exists a $c^* > \max\{\Delta r_0, \Delta r_{N-1}\}$ such that

- (1) For $c < \Delta r_{N-1}$ and for $c > c^*$, there are no strictly mixed ESS
- (2) For $\Delta r_{N-1} < c < c^*$ there exists a single strictly mixed ESS $x = x_+$.
- (3) For $\Delta r_0 < c < c^*$ there exist a NE $x = x_-$ with $x_- < x_+$.
- (4) Additionally, $x = 1$ is an ESS for $c < \Delta r_{N-1}$ and $x = 0$ is an ESS for $c > \Delta r_0$.

Thus, when the conditions in the proposition are satisfied we have the same situation as in our three player threshold game. When the cost c is high the only ESS is $x = 0$ and when the cost is low the only ESS is $x = 1$. The last case is not present in $\Theta_{r,c}(3, 2)$ of Section 2 since there $\Delta r_{N-1} = 0$. In addition $x = 0$ is an ESS coexisting with x_+ in the interval $\Delta r_0 < c < c^*$.

Note that the while Proposition 7 gives a sufficient condition, it is not necessary. For a three player game ($N = 3$), however, the second condition in Proposition 7 is both necessary and sufficient. From the proof of Proposition 7 it follows that the first condition in the proposition has an alternative formulation.

Proposition 8. The first condition in Proposition 7 can be replaced with:

- (1) The polynomial

$$\sum_{k=0}^{N-2} \binom{N-2}{k} x^k (1-x)^{N-2-k} \Delta^2 r_k$$

has at most one zero in $(0, 1)$.

Let this zero be \tilde{x} . Then the critical value c^* is given by $\tilde{g}_{\mathbf{r},0}(\tilde{x})$.

4 Discussion

Besides the interest in relaxing the assumptions of linear or strictly sub- or super-linear relations in the theory of group games, there are biological observations suggesting to go beyond such premises. The described extensions to threshold games

allow a natural relation between the per capita costs and benefit and the number of cooperators, and indeed this scenario turned out to exhibit a complex evolutionary dynamics (Figs. 1 and 2). The multiplayer threshold game we have studied yields dynamics not possible for two player games, in particular the appearance of a catastrophic event at a critical parameter value c^* separating very different ESS profiles.

The three-player game is the simplest possible threshold game and yet sufficient to yield complex dynamics. Above the critical cost c^* is a parameter region where only defection is stable, and below this point two alternative attracting states appear. One of these is an ESS comprising both cooperators and defectors and the other is a trivial ESS with pure defection. The cooperative state may disappear as a drastic event when the cost increases above the critical cost (for $\alpha > \frac{1}{2}$ in Fig. 1). Moreover, the system exhibits hysteresis behaviour. By this we mean that a system in the mixed state will experience a sudden transition and settle in the defecting state for gradually decreasing reward values or gradually increasing costs making the system traverse a critical line in the cost–benefit space. However, a subsequent complete recovery of parameter values allowing for cooperation will not recover the cooperative state. The population is trapped in the attractive basin of the defecting state.

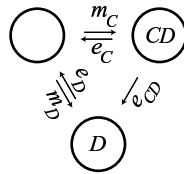
Such dynamics suggest the prediction that populations thriving near a critical point in parameter space can show drastic transformations among polymorphic states with cooperation or collapse into the trap of the fully defecting state. Subtle changes in environmental conditions may therefore induce a regime shift as the cooperative state suddenly collapse and disappear. A mere recovery of the original environmental conditions will then be insufficient to reinstall cooperation and mechanisms external to the game are required. Group foragers often depend on individually inaccessible and aggregated resources, such as large prey items. Such feeding strategies require a population density that make group behaviour feasible. The populations may accordingly be prone to complex evolutionary dynamics, as for instance rapid extinctions, when exposed to minor environmental changes that affect the cost–benefit parameters and/or population density. Unless, of course, alternative and individually accessible resources are available.

Initial evolution of cooperation from the defecting state needs alternative mechanisms as, for instance, kin selection or repeated interactions among individuals. However, when the cost–benefit ratio is low the attracting region of defection is small (Fig. 1), and this in turn facilitates the exit from the defection trap by stochastic fluctuations. Indeed, cooperation can gain foothold when, due to local fluctuations, the fraction of cooperators rises above a certain level (that is, above the repelling state x_* in the three-player game).

This scenario offers an alternative mechanism for the initial evolution of cooperation from the defecting state, in that stochastic fluctuations may originate in a metapopulation structure of the species. Cooperative group behaviour increases the productivity of the local population, because resources are increased beyond those that are exploitable by the individual. A subpopulation with cooperative behaviour

will thus have a higher probability of inoculating an empty environmental patch, as compared to a defective population in a similar environment.

Consider an exceedingly simple model of this situation made in the tradition of Slatkin [10, 28]. Three kinds of patches may exist: empty patches, patches with defectors only, and patches with defectors and cooperators. The frequency of these are P_0 , P_D , and P_{CD} , $P_0 + P_D + P_{CD} = 1$. Cooperators can only invade empty patches, and this occurs at the rate $m_C P_{CD}$. After invasion defectors invade so rapidly that we can assume it to occur immediately. Defectors can invade empty patches both from defecting patches and from cooperating patches, and we model the rate as $m_D Q_D$, where $Q_D = w_D P_D + w_{CD} P_{CD}$ and $w_D < 1$. Extinction of all individuals in defecting and cooperating patches occurs at the rates e_D and e_C , respectively, and in addition we allow extinction of cooperators from cooperating patches at the rate e_{CD} . The model may be summarized as



described by the equations:

$$\dot{P}_0 = -(m_C P_{CD} + m_D Q_D) P_0 + e_C P_{CD} + e_D P_D, \quad (11)$$

$$\dot{P}_D = m_D Q_D P_0 + e_{CD} P_{CD} - e_D P_D, \quad (12)$$

$$\dot{P}_{CD} = m_C P_{CD} P_0 - (e_C + e_{CD}) P_{CD}. \quad (13)$$

A trivial equilibrium with no cooperators always exists:

$$\tilde{P}_0 = \frac{e_D}{m_D w_D + e_D} \quad \text{and} \quad \tilde{P}_D = \frac{m_D w_D}{m_D w_D + e_D}, \quad (14)$$

and cooperators can invade this equilibrium when $e_C + e_{CD} < m_C \tilde{P}_0$. The invasion condition then becomes

$$1 + w_D \frac{m_D}{e_D} < \frac{m_C}{e_C + e_{CD}}. \quad (15)$$

The colonization rate of cooperators must thus be somewhat higher than their extinction rate for invasion to proceed. The extinction rate e_{CD} therefore must be small, which requires a polymorphic equilibrium in the patch and a defection trap, that is not too large. The requirement is, however, moderate, because the addend to one on the left side is expected to be small. The defecting patches are supposed to have a comparatively low population size because of fewer available resources. This causes their emigration potential (parameter w_D) to be low, and their colonization ability of empty patches (parameter m_D) will be low because of fewer resources. Finally their extinction rate (e_D) is higher. In other words, such an ecological scenario illustrates, in terms of the invasion condition (15), how increasing group facilitation

makes it increasingly likely that a more extensive attractive basin of defection can be surmounted and cooperation can gain foothold

Scenarios similar to the threshold model may apply to other levels of organisation of adaptive biological systems. Aggregates of cells sometimes produce generic biochemical compounds, which are required in order to trigger a beneficial event. A specific local concentration has to be reached in order to increase cell proliferation and hence fitness. The onset of aberrant neovascularisation in tumour growth has been modelled as an evolutionary game process in which groups of cells compete for attracting the budding of existing blood vessels [2, 30]. More precisely, certain threshold levels of vascular growth factor (VGF), a blood vessel stimulating compound, are required in order to attract the extension of newly formed blood vessels [9]. A model of neovascularisation occurring as a discrete event in the among-cell competitive process with threshold conditions showed similarly complex patterns in the evolutionary proliferation dynamics [2].

Previous models have shown that mechanisms such as kin selection and reciprocation may become important in small and/or spatially stable groups and could hence facilitate cooperation [17, 31]. Small groups tend to accumulate high degrees of local relatedness, which increase the potential for kin selection and thereby cooperation (especially in polygynous species such as e.g. Lions). Repeated interaction with neighbouring individuals may also favour cooperation due to reciprocation as for example in cells situated in solid tissue, which interact repeatedly with neighbouring cells [3]. However, mechanisms promoting cooperation have been most thoroughly investigated in evolutionary games with only two players as in the well known repeated games [1], games with punishment and reputation [22], or spatially structured games [26]. Such effects would most likely also promote cooperation in multiple player games. Nevertheless, the purpose of our model was to investigate the scope for cooperation and evolutionary dynamics in the multiple player threshold game and related games without invoking additional mechanisms already known to promote cooperation. Due to the general definitions of the cost and benefit this analysis may offer more suitable scenarios of the evolutionary dynamics of group interactions compared to previous models. However, it should be noted that this type of analysis is still based on certain assumptions that preclude additional realistic features potentially affecting group behaviour, such as e.g. instantaneous information on the other players' strategies and a corresponding conditional decision making.

One should note that the type of behaviour we have found in the threshold game, i.e. bifurcations and hysteresis, is only possible in games with at least three players. In fact, in two-player games a mixed ESS cannot coexist with a pure ESS for the same parameter values [8]. This suggests that modelling scenarios with a set of two player games instead of an n -player game may cause the model to miss important aspects of the dynamics. The exclusive use of two player games is however a common practice, and should perhaps be exerted with caution.

Future work include plans to extend the threshold model to include a hierarchical

structure with competitive relations among groups with repeated interactions.

Acknowledgements

This work was supported by the Danish Natural Science Research Council (LAB with Grant no. 21-03-0382) and by the Research Council of Norway. We thank Per Lundberg for helpful comments on an earlier version of the manuscript. Dedicated to the memory of Michael Svenstrup 1970–2002.

Note Added in Proof. The authors would like to clarify a statement made in the discussion. When stating that ESS bifurcations in two-strategy games only are possible in games with at least three players, we refer to internal bifurcations of saddle node type. Transcritical bifurcations altering the stability of monomorphic fixed points of the replicator equation or continuously transforming monomorphic ESSs to polymorphic ESSs can occur in two-player games. Hence, the difference between two and n -player games is specifically that only the latter permits the emergence of new internal ESSs, paired with new unstable fixed points. This property allows a rich variety of distinct ESS profiles, as illustrated by Example 5. Inferentially, hysteresis behaviour in cost-benefit parameter space involving polymorph states of the population is only possible with games involving at least three players, whereas with two-player games hysteresis is limited to the population shifting among the monomorphic states.

A Proofs

Proof of Proposition 4. A strategy x is an ESS if (3) is satisfied. Using the definition of W from (7) and a Taylor expansion around x the requirement for an ESS can be written as

$$\sum_{k=0}^{N-1} \frac{\varepsilon^k}{k!} g_{r,c}^{(k)}(x)(y-x)^{k+1} < 0 \quad (16)$$

for all $y \in [0, 1]$, $y \neq x$ and ε smaller than some $\varepsilon(y)$. Since ε can be chosen arbitrarily small, only the first term of the sum having $g_{r,c}^{(k)}(x) \neq 0$ has to be taken into account.

For the pure strategies, (16) shows that $x = 0$ is an ESS if $g_{r,c}(0) < 0$ and that $x = 1$ is an ESS if $g_{r,c}(1) > 0$. If $g_{r,c}(0) > 0$, (7) shows that $y = 1$ is the unique best response to $x = 0$, so in this case $x = 0$ is not a NE. Similarly, $x = 1$ is not a NE if $g_{r,c}(1) < 0$.

For mixed strategies x , we see from (16) that if $g_{r,c}(x) \neq 0$ or if an even integer k is the lowest such that $g_{r,c}^{(k)}(x) \neq 0$, then the inequality (16) cannot be satisfied for all y because the term $(y-x)^{k+1}$ can be both positive and negative. However, if an odd integer k is the lowest such that $g_{r,c}^{(k)}(x) \neq 0$, then the left hand side is always negative if $g_{r,c}^{(k)}(x) < 0$ and always positive otherwise.

The special cases for $x = 0$ and $x = 1$ follows in the same way. \square

A useful property of $\tilde{g}_{r,c}(x)$ that we will use in the next proof is:

Proposition 9.

$$\tilde{g}'_{r,c}(x) = (N-1) \sum_{k=0}^{N-2} \binom{N-2}{k} x^k (1-x)^{N-2-k} \Delta^2 r_k \quad (17)$$

This is proved by Motro in [25].

Proof of Proposition 7. By Proposition 9 it follows that

$$\frac{d^2}{dx^2} \tilde{g}_{r,c}(x) = (N-2)(N-1) \sum_{k=0}^{N-3} \binom{N-3}{k} x^k (1-x)^{N-3-k} \Delta^3 r_k$$

Thus if $\Delta^3 r_i \leq 0$ for all i then $\tilde{g}_{r,c}(\tilde{x})$ is concave and has at most one critical point \tilde{x} . When additionally $\Delta^2 r_0 = \tilde{g}'_{r,c}(0) > 0$ and $\Delta^2 r_{N-2} = \tilde{g}'_{r,c}(1) < 0$, this \tilde{x} will exist and be a maximum. Let $c^* = \tilde{g}_{r,0}(\tilde{x})$. For a given value of c , the maximal value $\tilde{g}_{r,c}$ takes is $c^* - c$. Thus $\tilde{g}_{r,c} = 0$ has no solutions when $c > c^*$.

When $c \leq c^*$, there is a solution to $\tilde{g}_{r,c}(x) = 0$ to the right of \tilde{x} if and only if $\tilde{g}_{r,c}(1) = \Delta r_{N-1} - c \leq 0$. Due to the concavity of $\tilde{g}_{r,c}$ this solution is an ESS. A solution to the left of \tilde{x} will not be an ESS, and statements (1), (2) and (3) follows. Statement (3) follows directly from Proposition 4. \square

Proof of Proposition 8. The statement is another way to express the condition that $\tilde{g}_{r,c}$ has exactly one critical point. As in the proof of Proposition 7, c^* is given by $\tilde{g}_{r,0}(\tilde{x})$. \square

References

- [1] Axelrod, R. and Hamilton, W. D. The evolution of cooperation. *Science*, 211:1390–1396, 1981.
- [2] Bach, L. A., Bentzen, S., Alsner, J., and Christiansen, F. B. An evolutionary-game model of tumour-cell interactions: possible relevance to gene therapy. *Eur. J. Cancer*, 37:2116–2120, 2001.
- [3] Bach, L. A., Sumpter, D. J. T., Alsner, J., and Loeschcke, V. Spatial evolutionary games of interaction among generic cancer cells. *Jour. Theor. Med.*, 5(1):47–58, 2003.
- [4] Batali, J. and Kitcher, P. Evolution of altruism in optional and compulsory games. *J. Theor. Biol.*, 175:161–171, 1996.
- [5] Boesch, C. Cooperative hunting in wild chimpanzees. *Anim. Behav.*, 48:653–657, 1994.

- [6] Boyd, R. and Richerson, P. J. The evolution of reciprocity in sizable groups. *J. Theor. Biol.*, 132:337–356, 1988.
- [7] Boyd, R. and Richerson, P. J. Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethol. Sociobiol.*, 13(3):171–195, 1992.
- [8] Broom, M., Cannings, C., and Vickers, G. Multi-player matrix games. *Bull. Math. Biol.*, 59(5):931–952, 1997.
- [9] Carmeliet, P. and Jain, R. K. Angiogenesis in cancer and other diseases. *Nature*, 407:249–257, 2000.
- [10] Christiansen, F. B. and Fenchel, T. M. *Theories of Populations in Biological Communities*. Springer Verlag, Berlin, Heidelberg, New York, 1977.
- [11] Courchamp, F., Clutton-Brock, T., and Grenfell, B. Multipack dynamics and the allee effect in the African wild dog, *lycaon pictus*. *Anim. Conserv.*, 3:277–285, 2000.
- [12] Courchamp, F., Grenfell, B., and Clutton-Brock, T. Impact of natural enemies on obligate cooperators. *OIKOS*, 91(2):311–322, 2000.
- [13] Courchamp, F. and Macdonald, D. W. Crucial importance of pack size in the African wild dog *lycaon pictus*. *Anim. Conserv.*, 4:169–174, 2001.
- [14] Frame, L. H., Malcolm, J. R., Frame, G. W., and Lawick, H. V. Social organisation of African wild dogs (*lycaon pictus*) on the Serengeti Plains, Tanzania, 1967-1978. *Zeitschrift fur Tierpsychologie*, 50:225–249, 1979.
- [15] Gardner, A. and West, S. A. Cooperation and punishment, especially in humans. *Am. Nat.*, 164:354–356, 2004.
- [16] Gintis, H. *Game Theory Evolving*. Princeton University Press, 2000.
- [17] Hamilton, W. D. The evolution of altruistic behavior. *Am. Nat.*, 97:354–356, 1963.
- [18] Hardin, G. The tragedy of the commons. *Science*, 162:143–148, 1968.
- [19] Hauert, C., Monte, S. D., and Sigmund, K. Replicator dynamics in optional public goods games. *J. Theor. Biol.*, 218:187–194, 2002.
- [20] Heinsohn, R. and Packer, C. Complex cooperative strategies in group-territorial African lions. *Science*, 269:1260–1262, 1995.
- [21] Hofbauer, J. and Sigmund, K. *Evolutionary games and population dynamics*. Cambridge University Press, 1998.
- [22] Leimar, O. and Hammerstein, P. Evolution of cooperation through indirect reciprocity. *Proc. R. Soc. London Ser. B*, 268:745–753, 2001.
- [23] Maynard-Smith, J. The theory of games and the evolution of animal conflicts. *J. Theor. Biol.*, 47:209–221, 1974.

- [24] Maynard-Smith, J. *Evolution and the Theory of Games*. Cambridge University Press, Cambridge, 1 edition, 1982.
- [25] Motro, U. Co-operation and defection: Playing the field and the ESS. *J. Theor. Biol.*, 151:145–154, 1991.
- [26] Nowak, M. A. and May, R. Evolutionary games and spatial chaos. *Nature*, 359:826–829, 1992.
- [27] Riechert, S. E. and Hammerstein, P. Game theory in the ecological context. *Annu. Rev. Ecol. Syst.*, 14:377–409, 1983.
- [28] Slatkin, M. Competition and regional coexistence. *Ecology*, 55:128–134, 1974.
- [29] Svenstrup, M. and Christiansen, F. B. Biker's dilemma: Cooperation between foes. *Selection*, 1-3:147–152, 2000.
- [30] Tomlinson, I. P. Game-theory models of interactions between tumour cells. *Eur. J. Cancer*, 33(9):1495–1500, 1997.
- [31] Trivers, R. The evolution of reciprocal altruism. *Q. Rev. Biol.*, 46:33–57, 1971.
- [32] von Neumann, J. and Morgenstern, O. *Theory of Games and Economic Behaviour*. Princeton University Press, NJ, 1944.

Paper II

Higher Order Cellular Automata

Nils A. Baas and Torbjørn Helvik (2005)

Advances in Complex Systems, Vol. 8, Nos. 2–3, pp. 169–192.

Higher Order Cellular Automata

Nils A. Baas and Torbjørn Helvik

Department of Mathematical Sciences
Norwegian University of Science and Technology
NO-7491 Trondheim, Norway

Abstract

We introduce a class of dynamical systems called Higher Order Cellular Automata (HOCA). These are based on ordinary CA, but have a hierarchical, or multi-level, structure and/or dynamics. We present a detailed formalism for HOCA and illustrate the concepts through four examples. Throughout the article we emphasize the principles and ideas behind the construction of HOCA, such that these easily can be applied to other types of dynamical systems. The article also presents new concepts and ideas for describing and studying hierarchical dynamics in general.

Keywords: Hierarchies; cellular automata; dynamical systems; hyperstructures.

1 Introduction

The study of systems of a large number of interacting agents is currently of foremost importance in many parts of science. Of special interest is the relation between the local interactions among agents and the resulting global behavior - the microlevel versus the macrolevel dynamics. Frequently, even very simple microlevel dynamics leads to highly complex macrolevel dynamics. This phenomenon, often referred to by the term self-organization, is usually very hard to analyze mathematically. Some classes of standard mathematical models of such complex systems have been much studied in the past, among them cellular automata [12, 30], random boolean networks [1], coupled map lattices [16], systems of interacting oscillators [18, 24] and stochastic interacting particle systems [21].

An aspect that is missing in all these standard models is that of hierarchical organization. This is the background for the present paper. Indeed, complex systems organized in a hierarchy of some kind are ubiquitous in science, and particularly in biology. To cite Arthur Koestler [17]: “We do not know what forms of life exist, but we

can safely assume that whenever there is life, it is hierarchically organized.” We feel that the aspect of hierarchical organization ought to be addressed in general mathematical models. The hope is that by studying such abstract models we can start to unravel the capabilities and general consequences of the hierarchy as an organizational principle for interacting systems. Advancement in this area would result in better understanding of the prominence of hierarchical organization in nature and also how it can be utilized e.g. in engineering. The abstractness of mathematics is necessary to ensure that a framework will accommodate different fields and enable us to draw conclusions across fields.

In this paper we present several concepts towards a framework for interacting systems with any type of hierarchical organization. These include the notions of higher order systems, morphologies and dynamics. However, we only give an informal discussion of the general notions. Rather, the main concern of the article is to illustrate the higher order principles by applying the concepts to the class of dynamical systems known as cellular automata (CA). We explore all natural types of higher order extensions in this setting, and give a detailed construction of a new class of dynamical systems called higher order cellular automata (HOCA). This is an interesting class of systems in its own regard. Through four detailed examples of HOCA we explore the impact of various higher order dynamics and morphologies on the behavior of the system.

The choice of cellular automata as an illustrative system has several reasons. Firstly, CA has been extensively used as standard models of complexity and cooperative behavior [6, 30]. Secondly, CA exhibit a wealth of interesting behavior and complex phenomena. Thirdly, their implementation on a computer is straightforward, thus making their behavior easy to visualize. Lastly, their simplicity eliminates the risk of implemental artifacts, such as round-off errors. The motivation for introducing HOCA comes from networks where cells often aggregate into dynamical units, several rules come into play and the neighborhood structure may vary. One may speculate whether networks in the brain work this way [27]. Our treatment of HOCA gives clear indications of how higher order dynamical systems should be introduced in other contexts.

In the next section we present the general notions and principles related to higher order systems. These will lay the foundations for the construction of HOCA. In Section 3 we briefly describe ordinary CA. We then extend the definition of CA to the case where there are underlying second order structures, and we see how this influences the dynamics through several examples. In the last example we sketch an application where a second order majority vote CA is used. We also discuss possible extensions of the HOCA formalism.

2 Higher Order Systems

Mathematically speaking, going from a basic system to a higher order (HO) system is lifting attention from a set X to the power set $\mathcal{P}(X)$ of subsets of X . A spatial dynamical system of interacting agents can be characterized by the three parts of which it is composed: (1) the agents or entities, (2) an interaction structure defined through the neighborhood of each agent and (3) a set of local interaction rules governing the dynamics. We will take this as our basic system X and look at possible higher order extensions. For instance, instead of single agents being the only units of interaction, also *groups* or aggregates of agents become natural units with their own dynamical behavior and interactions. We will call these groups second order agents, or 2-agents for short. Continuing the process one may consider 3-agents, being groups of 2-agents and so on. This power set process applied to the agents gives the system a certain multilevel structure, which we call a *higher order morphology*. A familiar example of 2-agents and 3-agents are the formation of molecules from atoms and macro-molecules from molecules. A similar example from a different field is the formation of families from individuals and communities from families.

The power set process can also be applied to the other components of the basic system, i.e., to the local rules and the interaction neighborhoods of the agents. This produces 2-rules and 2-neighborhoods which constitute a *higher order dynamics* - a dynamics on the dynamics. For instance, differentiation of biological cells can be considered as a 2-rule coming into play. We say that a system has a *higher order structure* if it has a HO morphology and/or a HO dynamics. It is then a higher order system. We use the term HO dynamical system when considering a HO system as a formal dynamical system respecting the HO structure. Higher order systems are instances of hyperstructures, which is a concept introduced by Baas in [2], and further developed in [3] and [4] where a general framework is presented (see also [5]). Hyperstructures are often created through a dynamical process of interactions which stabilizes in a multilevel structure. The present paper is more focused on state dynamics than previous work.

Relating to higher order systems there are several interesting approaches and classes of questions:

- (1) Given a system with a higher order structure, through which kind of dynamical process did the HO structure come into existence? That is, how was it generated from first principles? [25]
- (2) How does the existence and properties of a HO structure influence the dynamical behavior of the system?
- (3) How can one design a HO system to perform a given task?

We will in this paper be concerned with (2). Thus, we assume that our system, which is based on a CA, has an underlying second order morphology and/or a second order dynamics which is predefined.

3 Cellular Automata

A CA is a discrete dynamical system defined on a lattice $\mathcal{L} = \mathbb{Z}^d$, for some dimension d . The elements of the lattice \mathcal{L} are traditionally called *cells*. Each cell takes a *state* in a finite set \mathcal{A} , called the alphabet. A *configuration* \mathbf{a} of the lattice is the assignment of a state to each cell. The configuration is formally a map $\mathbf{a} : \mathcal{L} \rightarrow \mathcal{A}$. The collection of all possible configurations is written $\mathcal{A}^{\mathcal{L}}$. The state of cell $v \in \mathcal{L}$ is denoted by $\mathbf{a}(v)$. If S is a subset of \mathcal{L} , $\mathbf{a}(S)$ denotes the configuration \mathbf{a} restricted to S .

Let $N = \{v_0, v_1, \dots, v_{n-1}\}$ be a finite subset of \mathcal{L} containing the origin. The set N is called the *neighborhood* set, and defines the interaction-structure of the cellular automaton. The dynamics of the CA is governed by a map $f : \mathcal{A}^n \rightarrow \mathcal{A}$ which updates the state of each cell based on the configuration of its neighborhood. The map f is often called the local *rule* of the CA. The quadruple $(\mathcal{L}, N, \mathcal{A}, f)$ defines a d -dimensional CA [22].

An example is the CA with $\mathcal{L} = \mathbb{Z}$, $N = \{-1, 0\}$, $\mathcal{A} = \{0, 1\}$ and $f : \{0, 1\}^2 \rightarrow \{0, 1\}$ given by

$$f(a_{-1}, a_0) = a_{-1} + a_0 \pmod{2}. \quad (1)$$

This is a simple one dimensional CA. Figure 1 shows a time series of an evolution of this CA. The lattice \mathbb{Z} runs horizontally and the time steps run downwards. Such diagrams are called space-time diagrams, and are the main tool for obtaining information about the global qualitative behavior of a CA.

The local transition map f generates a global CA-map $F : \mathcal{A}^{\mathcal{L}} \rightarrow \mathcal{A}^{\mathcal{L}}$, defined by

$$(F\mathbf{a})(u) = f(\mathbf{a}(u + v_0), \mathbf{a}(u + v_1), \dots, \mathbf{a}(u + v_{n-1})). \quad (2)$$

Here, $u + v_i$ means v_i translated by the vector u (or vice versa). If we give the set \mathcal{A} the discrete topology and $\mathcal{A}^{\mathcal{L}}$ the corresponding product topology, then $F : \mathcal{A}^{\mathcal{L}} \rightarrow \mathcal{A}^{\mathcal{L}}$ is a continuous map. Conversely, if G is an arbitrary continuous and shift-commuting map from $\mathcal{A}^{\mathcal{L}}$ to $\mathcal{A}^{\mathcal{L}}$, then there is some local transition map g that generates G in the above sense [14, Thm. 3.1 and 3.4]. The dynamics is given by using the local transition map f to update the states of all cells synchronously. One may also consider finite CA, where \mathcal{L} is a circle graph or a cartesian graph product of circle graphs, as

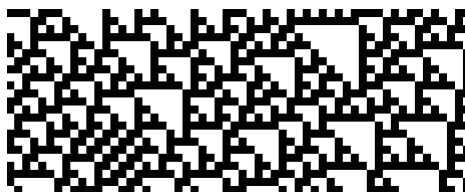


Fig. 1: The behavior of the CA defined in (1). A white square represents the state 0 and a black square the state 1.

well as CA on other graph structures. The study of various CA and classes of CA has generated much interest in the past.

A common classification of cellular automata is the one introduced by Wolfram in [29]. This is an *empirical* classification based on the qualitative features of typical evolutions from random initial configurations. Wolfram's four classes are:

- (1) Evolution leads to a homogenous state.
- (2) Evolution leads to a set of separated simple stable or periodic structures.
- (3) Evolution leads to a chaotic pattern.
- (4) Evolution leads to complex localized structures, sometimes long-lived.

The CA defined in Eq. (1) is a typical member of class (3). Later, several formal classification schemes have been proposed, see e.g. [8, 11, 19]. Some of these formalize Wolfram's intuition while others are based on different mathematical or computational concepts. We will refer to Wolfram's classes later in the article.

4 Second Order Cellular Automata

We now introduce and define a second order CA, or 2-CA, step by step. We start by describing the dynamical second order extensions, and then proceed to the morphological second order extension.

Let \mathcal{L}_1 be a lattice. The lattice points represent the cells of the 2-CA. Let $\mathcal{A}_1 = \{0, 1, \dots, k_1 - 1\}$ be a finite set called the first level alphabet. Each cell has a state in \mathcal{A}_1 . This gives the first level state configuration \mathbf{a}_1 :

$$\mathbf{a}_1 : \mathcal{L}_1 \rightarrow \mathcal{A}_1. \quad (3)$$

The subscript 1 generally denotes the first morphological level.

4.1 2-dynamics

In an ordinary CA, all cells have identical neighborhoods, given by the set N . In a 2-CA, however, each cell can have one of several possible neighborhoods. The set of possible neighborhoods is the same for all cells, and is generated from a central collection of neighborhoods in the following sense. Let \mathcal{N} be a finite collection of finite subsets of \mathcal{L}_1 , all containing the origin. Neighborhoods are assigned to cells by the map

$$\mathbf{n} : \mathcal{L}_1 \rightarrow \mathcal{N}. \quad (4)$$

We call \mathbf{n} the first level neighborhood configuration. The neighborhood that is used to update a cell v is $\mathbf{n}(v)$ translated by v , written $(v + \mathbf{n}(v)) \subset \mathcal{L}_1$. The collection \mathcal{N} is called a *2-neighborhood*. Later we will define a 2-transition map ϕ that governs the local change of neighborhood.

The possible neighborhoods N in the 2-neighborhood \mathcal{N} may have different sizes. In this case, we must define different local rules for differently sized neighborhoods. In general, we will introduce one *or more* rules for each neighborhood size through the notion of a 2-rule. Define the set $M \subset \mathbb{N}$ as the set of cardinalities in \mathcal{N} :

$$M = \{\text{Card}(N) : N \in \mathcal{N}\}. \quad (5)$$

Then define the 2-rule $\mathcal{R} = \{f_1, f_2, \dots, f_k\}$ as a set of maps such that

- (1) Each map is of the form $f_i : \mathcal{A}_1^m \rightarrow \mathcal{A}_1$ for some $m \in M$
- (2) For each $m \in M$ there is at least one map $f_j \in \mathcal{R}$ with domain \mathcal{A}_1^m .

The maps in \mathcal{R} are the possible local rules of the cells. It is convenient to partition the sets \mathcal{N} and \mathcal{R} according to M in the following way. Let

$$\mathcal{N}_m = \{N \in \mathcal{N} : \text{Card}(N) = m\}, \quad (6)$$

$$\mathcal{R}_m = \{f \in \mathcal{R} : f : \mathcal{A}_1^m \rightarrow \mathcal{A}_1\}. \quad (7)$$

Rules are assigned to the cells by a map

$$\mathbf{r} : \mathcal{L}_1 \rightarrow \mathcal{R}. \quad (8)$$

We call \mathbf{r} the first level rule configuration. Obviously, the rule of a given cell has to be *compatible* with the neighborhood of that cell. That is, we require that $\mathbf{r}(v) \in \mathcal{R}_m$ if and only if $\mathbf{n}(v) \in \mathcal{N}_m$.

Finally we define the local 2-transition map ϕ that alters the local neighborhood and rule during system evolution. For each $m \in M$ define a map

$$\phi_m : \mathcal{A}_1^m \times \mathcal{N}_m \times \mathcal{R}_m \rightarrow \mathcal{N} \times \mathcal{R} \quad (9)$$

that satisfies the obvious requirement that for each set of arguments the rule and neighborhood components of the image are compatible. Define the 2-transition map ϕ as the family $(\phi_m)_{m \in M}$. Through the 2-transition map the system is given a 2-dynamics. The 2-dynamics is a process on the 2-neighborhoods and 2-rules. To complete the picture it is also possible to introduce 2-states, being families of states from \mathcal{A} . However, we will not pursue this here. An example of a 2-CA with a 2-rule is given in Sec. 5.1, and an example of a 2-CA with a 2-neighborhood in Sec. 5.2.

4.2 2-morphology

In the previous section we equipped the CA with a 2-dynamics. We will now concentrate on the higher order morphological extension obtained by grouping the cells into 2-cells. This is called a 2-morphology. A higher order CA will in general consist of a number N of levels, each consisting of a collection of disjoint lattices. In this paper we

only consider in detail second order CA with one lattice on each of two levels. Moreover, we will only allow a 2-dynamics on the first level. Nevertheless, the extension to general HOCA of arbitrary order will be evident.

Let \mathcal{L}_2 be a lattice, where each lattice point represents a group of cells which we consider as a functional unit. We call these 2-cells, but “organs” would also be an appropriate moniker. A 2-cell should be considered as containing a number of cells. This collection of cells is called the *inner structure* of the 2-cell, and is given by a map

$$M: \mathcal{L}_2 \rightarrow \mathcal{P}(\mathcal{L}_1), \quad (10)$$

where $\mathcal{P}(\mathcal{L}_1)$ denotes the set of all *finite* subsets of \mathcal{L}_1 . It is convenient to define

$$M^*: \mathcal{L}_1 \rightarrow \mathcal{P}(\mathcal{L}_2) \quad \text{by} \quad M^*(v) = \{u \in \mathcal{L}_2 | v \in M(u)\}. \quad (11)$$

Thus, $M^*(v)$ gives the collection of 2-cells in which cell v is a component. This collection is called the *outer structure* of v .

Generally, the outer structure of a cell can consist of one, zero or more than one 2-cells. However, in this paper we will only consider the case where each cell resides in exactly one 2-cell and all 2-cells contain the same number of cells, say c . Mathematically, this is expressed as $\text{Card}(M^*(v)) = 1$ for all $v \in \mathcal{L}_1$ and $\text{Card}(M(u)) = c$ for all $u \in \mathcal{L}_2$. This restriction makes the formulae more transparent.

Let $\mathcal{A}_2 = \{0, 1, \dots, k_2 - 1\}$ be a finite set called the second level alphabet. Each 2-cell has a state in \mathcal{A}_2 . This gives the second level state configuration $\mathbf{a}_2: \mathcal{L}_2 \rightarrow \mathcal{A}_2$. The second level has no 2-dynamics, but has a CA dynamics defined by a neighborhood $N_2 \subset \mathcal{L}_2$ containing the origin and a transition map

$$f_2: \mathcal{A}_2^{\text{Card}(N_2)} \times \mathcal{A}_1^c \rightarrow \mathcal{A}_2. \quad (12)$$

The term \mathcal{A}_1^c represents the states of the cells $M(u)$ contained in the 2-cell u . However, a 2-cell will generally contain many cells, and f_2 will only take some mean value of their states as an input. See for instance Examples 3 and 4 in Secs. 5.3 and 5.4.

A final question to answer is how the dynamics of the second level should influence the cells. We have chosen to modify the maps ϕ_m to take also as an input the value of the 2-cell the given cell belongs to. Thus, alter equation (9) to

$$\phi_m: \mathcal{A}_1^m \times \mathcal{A}_2 \times \mathcal{N}_m \times \mathcal{R}_m \rightarrow \mathcal{N} \times \mathcal{R}. \quad (13)$$

A change in 2-cell value will influence the rules and neighborhoods of the contained cells in the next time step. The structure of a 2-CA with a 2-morphology is illustrated in Fig. 2.

A brief note on terminology: we have opted to use the terms 2-rule and 2-neighborhood to denote the families of rules and of neighborhoods the 2-dynamics are acting on. Thus, it should be kept in mind that these terms do not refer the rule and neighborhood set of the 2-cells.

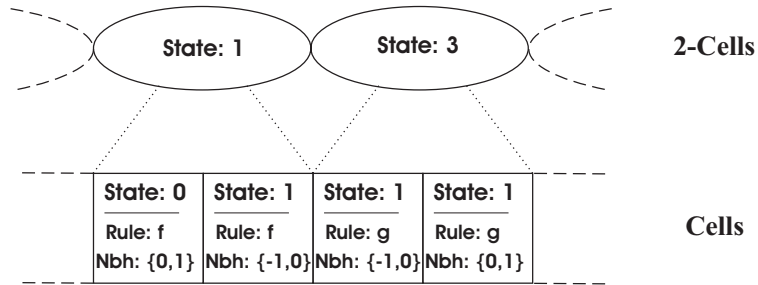


Fig. 2: A illustration of the structure of a 2-CA with $\mathcal{L}_1 = \mathcal{L}_2 = \mathbb{Z}$, $\mathcal{A}_1 = \{0, 1\}$, $\mathcal{A}_2 = \{0, 1, 2, 3\}$, $\mathcal{R} = \{f, g\}$, $\mathcal{N} = \{-1, 0, 0, 1\}$ and $c = 2$. Each 2-cell contains two cells, which here are adjacent.

4.3 Global dynamics

Although we are mainly interested in studying the state dynamics of the cells, we have to define the 2-CA as a dynamical system on a larger structure to keep track of rule and neighborhood data. Denote by X the product set

$$X = \mathcal{A}_1^{\mathcal{L}_1} \times \mathcal{N}^{\mathcal{L}_1} \times \mathcal{R}^{\mathcal{L}_1} \times \mathcal{A}_2^{\mathcal{L}_2}. \quad (14)$$

Analogously to ordinary CA, the local transition maps ϕ and f_2 generate a global transition map F . This map must be defined as a map on the product space X defined above. Let $\pi_{\mathbf{a}_1}$, $\pi_{\mathbf{n} \times \mathbf{r}}$ and $\pi_{\mathbf{a}_2}$ be the projection maps from X to $\mathcal{A}_1^{\mathcal{L}_1}$, $\mathcal{N}^{\mathcal{L}_1} \times \mathcal{R}^{\mathcal{L}_1}$ and $\mathcal{A}_2^{\mathcal{L}_2}$ respectively. Moreover, let $\mathbf{n}_v, \mathbf{r}_v$ and M_v^* be shorthand for $\mathbf{n}(v)$, $\mathbf{r}(v)$ and $M^*(v)$. Then, given a configuration $\langle \mathbf{a}_1, \mathbf{n}, \mathbf{r}, \mathbf{a}_2 \rangle$ of the lattices, we define the global map

$$F: X \rightarrow X \quad (15)$$

componentwise by the three equations (16a) to (16c) below. The first equation says that the state of each cell v is updated by its current rule applied to its current neighborhood configuration:

$$(\pi_{\mathbf{a}_1} F \langle \mathbf{a}_1, \mathbf{n}, \mathbf{r}, \mathbf{a}_2 \rangle)(v) = \mathbf{r}_v(\mathbf{a}_1(v + \mathbf{n}_v)). \quad (16a)$$

The second equation states that the rule and neighborhood of each cell v are updated by the 2-transition rule using the states of its current neighborhood and associated 2-cell:

$$(\pi_{\mathbf{n} \times \mathbf{r}} F \langle \mathbf{a}_1, \mathbf{n}, \mathbf{r}, \mathbf{a}_2 \rangle)(v) = \phi_{\text{Card}(\mathbf{n}_v)}(\mathbf{a}_1(v + \mathbf{n}_v), \mathbf{a}_2(M_v^*), \mathbf{n}_v, \mathbf{r}_v). \quad (16b)$$

Finally, the last equation says that the state of each 2-cell u is updated by f_2 applied to its current neighborhood configuration and the states of the cells contained in u :

$$(\pi_{\mathbf{a}_2} F \langle \mathbf{a}_1, \mathbf{n}, \mathbf{r}, \mathbf{a}_2 \rangle)(u) = f_2(\mathbf{a}_2(u + N_2), \mathbf{a}_1(Mu)). \quad (16c)$$

If we endow $\mathcal{A}_1, \mathcal{A}_2, \mathcal{N}$ and \mathcal{R} with the discrete topologies and X with the corresponding product topology, then X will by Tychonoff's theorem (see e.g. [23]) be a compact topological space. It is not hard to show that F is a continuous map since $\mathcal{A}_i, \mathcal{R}, \mathcal{N}$ and all $N \in \mathcal{N}$ are finite sets. In formal terms, this makes the pair (X, F) a compact, topological dynamical system.

4.4 Relation to ordinary CA

As the alert reader already will have realized, it is easy to prove that a 2-CA with a 2-dynamics but no 2-morphology is conjugate to an ordinary CA. This is because the rule and neighborhood information can be coded as a part of the state of each cell, and the local rule f can be modified to act as the 2-transition rule. We state this as an observation and give a constructive proof.

Observation. A 2-CA with a 2-dynamics, but no 2-morphology, is conjugate to an ordinary CA.

Proof. Denote the 2-CA by $D = (\mathcal{L}_1, \mathcal{N}, \mathcal{A}_1, \mathcal{R}, \phi)$. We construct a conjugate CA $C = (\mathcal{L}, \mathcal{N}, \mathcal{A}, f)$. Let $\mathcal{L} = \mathcal{L}_1$ and $\mathcal{A} = \mathcal{A}_1 \times \mathcal{N} \times \mathcal{R}$. Write $\mathcal{N} = \{N_1, N_2, \dots, N_l\}$, and let $N = \bigcup N_i$ be the neighborhood of the CA. The union is taken regarding the N_i as subsets of \mathcal{L}_1 . Note that $0 \in N$ since $0 \in N_i$ for all i , and write $N = \{0, v_1, \dots, v_n\}$. We now define the local map f of the CA:

$$f : (\mathcal{A}_1 \times \mathcal{N} \times \mathcal{R})^{n+1} \rightarrow \mathcal{A}_1 \times \mathcal{N} \times \mathcal{R}. \quad (17)$$

Denote the current state of cell v_i as $\langle \mathbf{a}_{v_i}, \mathbf{n}_{v_i}, \mathbf{r}_{v_i} \rangle$. Put

$$f(\langle \mathbf{a}_0, \mathbf{n}_0, \mathbf{r}_0 \rangle, \langle \mathbf{a}_{v_1}, \mathbf{n}_{v_1}, \mathbf{r}_{v_1} \rangle, \dots, \langle \mathbf{a}_{v_n}, \mathbf{n}_{v_n}, \mathbf{r}_{v_n} \rangle) = \langle \mathbf{r}_0(\mathbf{a}(\mathbf{n}_0)), \phi(\mathbf{a}(\mathbf{n}_0), \mathbf{n}_0, \mathbf{r}_0) \rangle. \quad (18)$$

Note that f only uses the \mathbf{n} and \mathbf{r} values of cell 0 and only the \mathbf{a} values of the cells contained in \mathbf{n}_0 to calculate the next state (recall that $\mathbf{n}_0 \subseteq N$ by construction). The identity map on $(\mathcal{A}_1 \times \mathcal{N} \times \mathcal{R})^{\mathcal{L}_1}$ is a topological conjugacy between the 2-CA D and the CA C . \square

This means that 2-CA with a 2-dynamics constitute a subclass of ordinary CA. However, this fact does not make this class of 2-CA uninteresting to study using our formalism. Firstly, our way of expressing the system is more natural for these types of CA. Secondly, we are mainly interested in studying the differences between a basic CA rule f and various HO extensions of it, to see what effects the extensions have on the dynamics. Thirdly, the HO framework could also be useful in the modeling of various systems. See for instance Example 4 in Sec. 5.4. Mathematically, CA with 2-dynamics seems to be an interesting subclass of ordinary CA.

In the case where the 2-CA has a 2-morphology it will in general not be conjugate to any ordinary CA due to the lack of a homogenous structure in the system. Exceptions are cases with a very regular morphology map M , as, for instance, in the example displayed in Fig. 2.

5 Examples

Although the definition of a 2-CA presented in Sec. 4 is rather general, we will keep the examples fairly basic. In particular, we will only consider one-dimensional 2-CA and mainly introduce only a single second order extension in each example. We base the examples on the subclass of CA known as totalistic CA [29]. For these CA, the value of the local transition map f only depends on the sum of the inputs and not the exact neighborhood configuration. This property considerably reduces the number of parameters involved. According to Wolfram, the subclass of totalistic CA seems to exhibit behavioral characteristics of all types of CA [29].

5.1 Example 1 — Dynamic rule 2-CA

In this example the 2-CA has a 2-rule consisting of two different local rules. During time evolution, the rule used to update each given cell will vary between these rules. Which rule that is applied is governed by a 2-transition map ϕ .

Let $\mathcal{L}_1 = \mathbb{Z}$ and $\mathcal{A}_1 = \{0, 1\}$. There is only one possible neighborhood - the standard radius two neighborhood. That is,

$$N = \{-2, 1, 0, 1, 2\} \tag{19}$$

and $\mathcal{N} = \{N\}$. Furthermore, define $\mathcal{R} = \{f_a, f_b\}$, where both f_a and f_b are maps $\mathcal{A}_1^5 \rightarrow \mathcal{A}_1$. In this example we let both f_i be totalistic CA rules. Thus, they are each defined by six parameters in \mathcal{A}_1 in the following sense. Let a_0, a_1, \dots, a_5 be elements of \mathcal{A}_1 , and let f_a be given in terms of these parameters by

$$f_a(x_{-2}, x_{-1}, x_0, x_1, x_2) = a_{\Sigma(x)}, \quad \text{where } \Sigma(x) = \sum_{i=-2}^2 x_i. \tag{20}$$

Addition is performed regarding 0 and 1 as elements of \mathbb{Z} . In the same way, let f_b be given by b_0, b_1, \dots, b_5 .

Finally, we have to define the 2-transition map $\phi : \mathcal{A}_1^5 \times \mathcal{R} \rightarrow \mathcal{R}$. We do this in a totalistic way too, by introducing two parameters c_a and c_b with $0 \leq c_i \leq 5$. Let

$$\phi(x_{-2}, x_{-1}, x_0, x_1, x_2, f) = \begin{cases} f_a & \text{if } f = f_b \text{ and } \Sigma(x) = c_b, \\ f_b & \text{if } f = f_a \text{ and } \Sigma(x) = c_a, \\ f & \text{otherwise.} \end{cases} \tag{21}$$

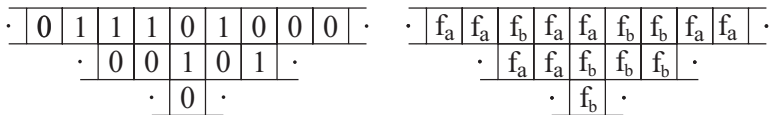


Fig. 3: Some steps in an evolution of the 2-CA described in Example 1 with parameters defined in (22) and (23). The state configuration is shown to the left and the corresponding rule configuration to the right.

Thus, the 2-transition map changes the local rule when the sum of the neighborhood configuration is exactly c_a or c_b , depending on what the current rule is. To give a specific example, let

$$\begin{aligned} a_0 a_1 \dots a_5 &= 001100, \\ b_0 b_1 \dots b_5 &= 010000, \end{aligned} \tag{22}$$

and

$$c_a = c_b = 3. \tag{23}$$

Figure 3 gives an example of an evolution of this system. The first line of the left hand side shows the initial state configuration of an interval of the lattice, and the first line of the right hand side shows the initial rule configuration of the same interval. Subsequent lines show the states after one and two iterations. Only the states of the central cells are shown in these lines since only these states can be computed from the initial values that are given.

The first two panels of Fig. 4 show space time diagrams for the ordinary totalistic CA with rules as in (22). Time runs downwards. Periodic boundary conditions are applied, creating finite CA defined on circle graphs. The number of cells is 293 and the evolution is cut off after 400 iterations. The initial configurations are uniformly random both for states and rules. The four subsequent panels of Fig. 4 show space time diagrams for the 2-CA introduced above with 2-rule given by a_i and b_i in (22) and various values of c_a and c_b . Note that these diagrams only show the state configurations and not the corresponding rule configurations. Generally, the space time diagrams for the rule configurations look similar to those for the state configurations. For instance, for the third and for the sixth panel the rule configuration locally settles into a periodic pattern of low period when the state configuration does.

As can be seen, both rule 001100 and rule 010000 yield a chaotic time evolution, and Wolfram placed them in class 3 [29]. However, by combining them into 2-rules through various choices of c_a and c_b , one obtains a broad spectrum of possible behaviors. Some of these 2-rules, such as the one in the last panel, have a ordered behavior (Wolfram class 2) while others, such as the one in the second last panel, have a complex behavior (comparable to Wolfram class 4).

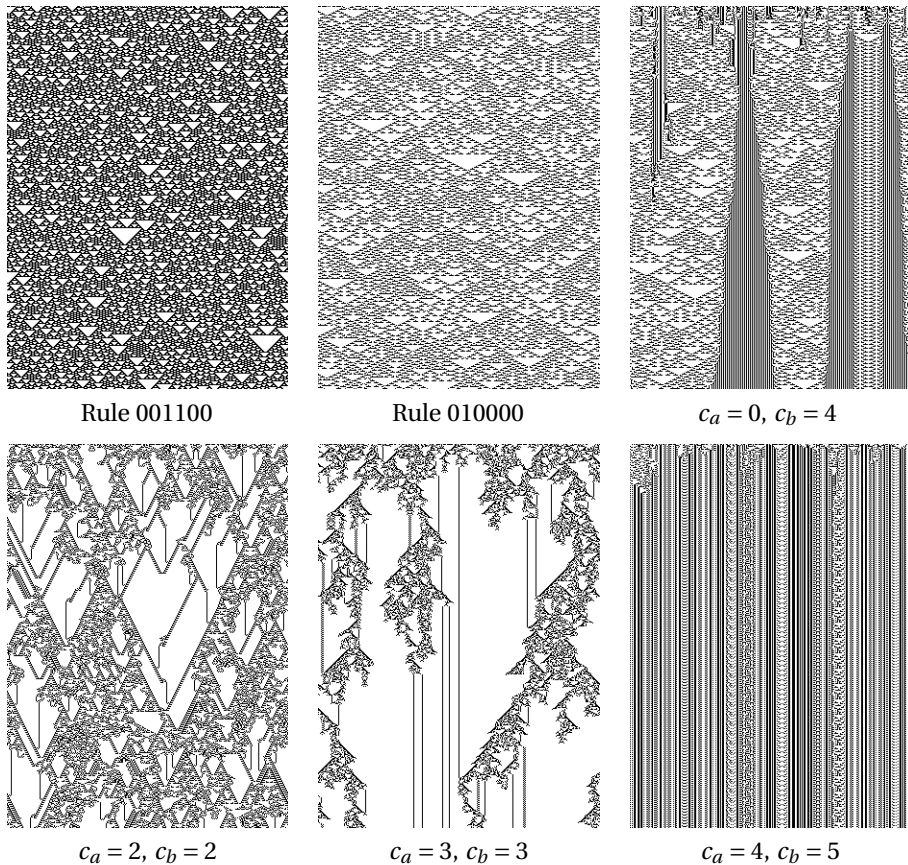


Fig. 4: The first two panels: Space-time diagrams for ordinary totalistic CA with local rules as given. The last four panels: Space-time diagrams for the 2-CA in Example 1 with 2-rule given in (22) and various values of c_a and c_b .

5.2 Example 2 — Dynamic neighborhood 2-CA

In this example the 2-CA has a 2-neighborhood consisting of several connected neighborhoods of distinct sizes. We will describe a special class of such 2-CA that is a natural extension of ordinary totalistic CA.

Let the 2-CA be one-dimensional (that is, $\mathcal{L}_1 = \mathbb{Z}$). For $m \in \mathbb{N}$ define the sets $N_m \subset \mathcal{L}_1$ as

$$N_m = \{-m, -m+1, \dots, m\}. \quad (24)$$

We now define a class of 2-CA called equitotalistic dynamic-neighborhood 2-CA as follows

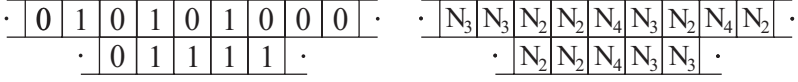


Fig. 5: One time step in an evolution of the equitotalistic dynamic neighborhood 2-CA with parameters defined in (27). The state configuration is shown to the left and the corresponding neighborhood configuration to the right.

Definition 1. An equitotalistic dynamic-neighborhood 2-CA is given by the integer parameters $k, r = (r_{\min}, r_{\max}), s = (s_{\min}, s_{\max})$ and $a = (a_i)$, for $0 \leq i \leq (k-1)(r_{\max}+1)$, representing:

- (1) The alphabet $\mathcal{A}_1 = \{0, \dots, k-1\}$.
- (2) The 2-neighborhood $\mathcal{N} = \{N_{r_{\min}}, \dots, N_{r_{\max}}\}$.
- (3) The 2-rule $\mathcal{R} = \{f_{r_{\min}}, \dots, f_{r_{\max}}\}$ where $f_m : \mathcal{A}_1^{2m+1} \rightarrow \mathcal{A}_1$ is defined by

$$f_m(x_{-m}, \dots, x_m) = a_{\Sigma_m(x)}, \quad \text{with } \Sigma_m(x) = \sum_{i=-m}^m x_i. \quad (25)$$

- (4) The 2-transition map $(\phi_{r_{\min}}, \dots, \phi_{r_{\max}})$, where ϕ_m is defined by

$$\phi_m(x_{-m}, \dots, x_m, N_m, f_m) = \begin{cases} (N_{m+1}, f_{m+1}) & \text{if } \Sigma_m(x) < s_{\min} \text{ and } m < r_{\max}, \\ (N_m, f_m) & \text{if } \Sigma_m(x) < s_{\min} \text{ and } m = r_{\max}, \\ (N_m, f_m) & \text{if } s_{\min} \leq \Sigma_m(x) \leq s_{\max}, \\ (N_m, f_m) & \text{if } \Sigma_m(x) > s_{\max} \text{ and } m = r_{\min}, \\ (N_{m-1}, f_{m-1}) & \text{if } \Sigma_m(x) > s_{\max} \text{ and } m > r_{\min}. \end{cases} \quad (26)$$

Intuitively, the 2-transition map expands neighborhoods when their sum is below s_{\min} and contracts neighborhoods when their sum is larger than s_{\max} . However, the neighborhood radius is kept between r_{\min} and r_{\max} . The local rules used for each neighborhood size are all totalistic and given by the same a_i -parameters. Let us consider a specific example. Put

$$k = 2, \quad r = (2, 4), \quad s = (2, 6), \quad (27)$$

$$a_0 a_1 \dots a_9 = 0101010101.$$

The three local rules f_2, f_3 and f_4 are merely the sum modulo 2 rules for the appropriately sized neighborhoods. Figure 5 gives an example of one time step of this system.

Looking at a large scale space time diagram of this 2-CA reveals an intriguing phenomenon. An example of a space time diagram for the states along with the corresponding diagram for the neighborhood configuration is shown in Fig. 6. An interval

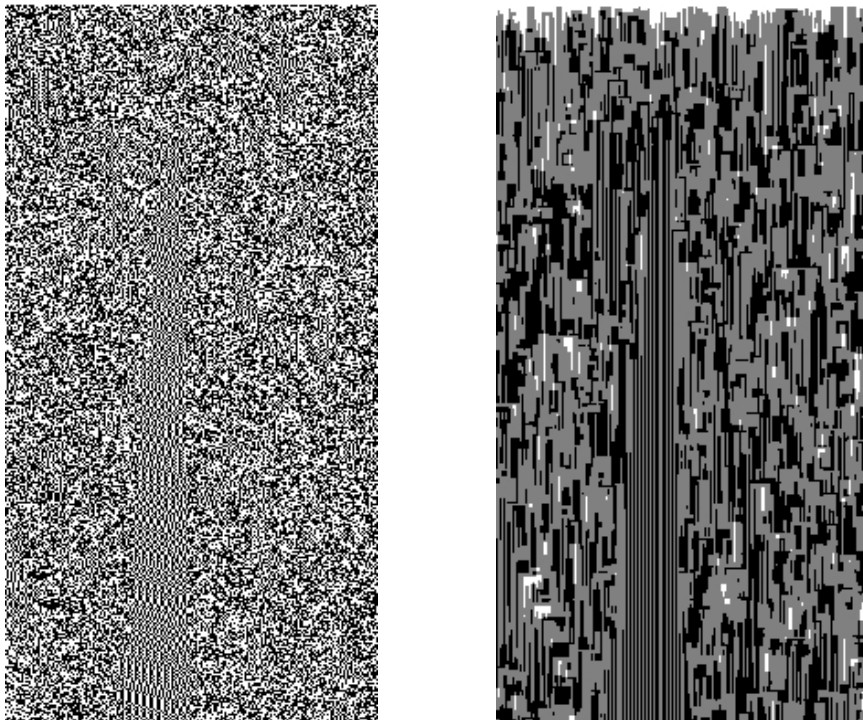


Fig. 6: Crystallization in the equitotalistic dynamic neighborhood 2-CA with parameters given in (27). The state configuration is shown to the left and the neighborhood configuration to the right. For the neighborhood configuration, white represents N_2 , gray represents N_3 and black represents N_4 .

of 204 cells from a larger lattice is displayed. The initial radii of the cells are r_{\min} . As can be seen, the behavior is initially very disordered, but after some time an area in the middle begins to stabilize into a more regular pattern. This area gradually expands and if we continue the evolution it will, together with other such areas which may appear, eventually cover the entire lattice. We call this phenomenon *crystallization*. Also note that the neighborhood configuration of the crystallized area becomes fixed. To formalize the notion of crystallization would involve defining a crystallized area as a group of cells whose configuration is more regular than that of the surrounding lattice. Furthermore, this regularity must be robust to influences from neighboring cells so that it is retained during evolution. Lastly, the regular area must have a tendency to grow. Regularity can be formalized for instance by using regular languages or by the notions of local information and information flow introduced in [15].

In computer experiments we have also observed crystallization in many other

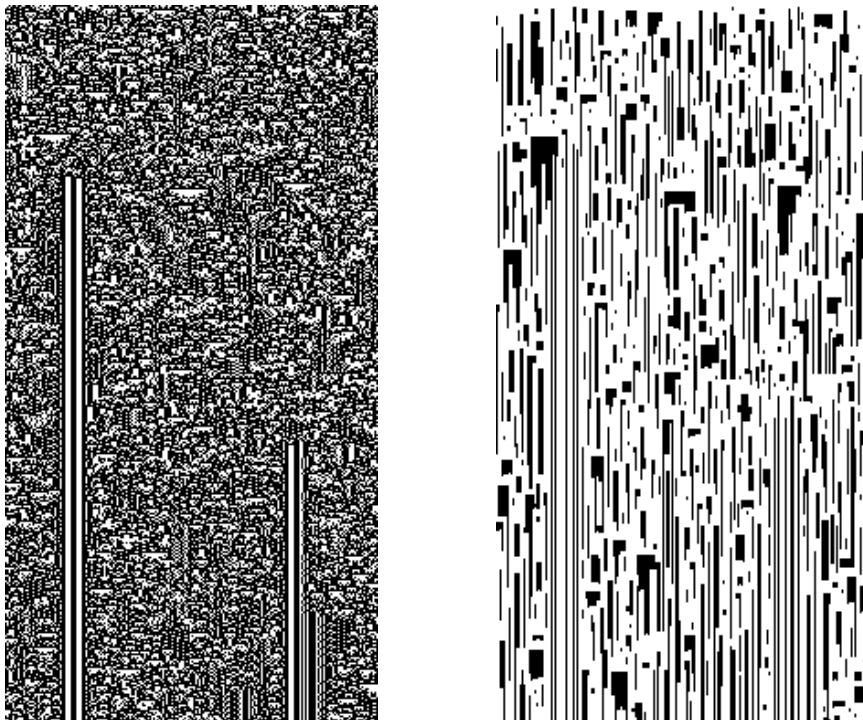


Fig. 7: Crystallization in the equitotalistic dynamic neighborhood 2-CA with parameters given in (28). The state configuration is shown to the left and the neighborhood configuration to the right. For the neighborhood configuration, white represents N_2 and black represents N_3 .

equitotalistic dynamic neighborhood 2-CA. One further example is shown in Fig. 7. Here, the regular areas that appear are more ordered. The parameters of this 2-CA are

$$k = 2, \quad r = (2, 3), \quad s = (1, 6), \quad (28)$$

$$a_0 a_1 \dots a_7 = 01011011.$$

Crystallization even occurs in some of the 2-CA with dynamic rules described in Example 1, for instance in the one in the third panel of Fig. 4.

We study one further example of an equitotalistic dynamic neighborhood 2-CA. In this case, let $k = 3$ and let an ordinary totalistic CA with radius 3 be defined by the parameters

$$a_0 a_1 \dots a_{14} = 020201121221202. \quad (29)$$

This CA has a disordered behavior, but the CA with radius 2 given by the first 11 parameters a_0 to a_{10} exhibits several localized, periodic structures and interactions among

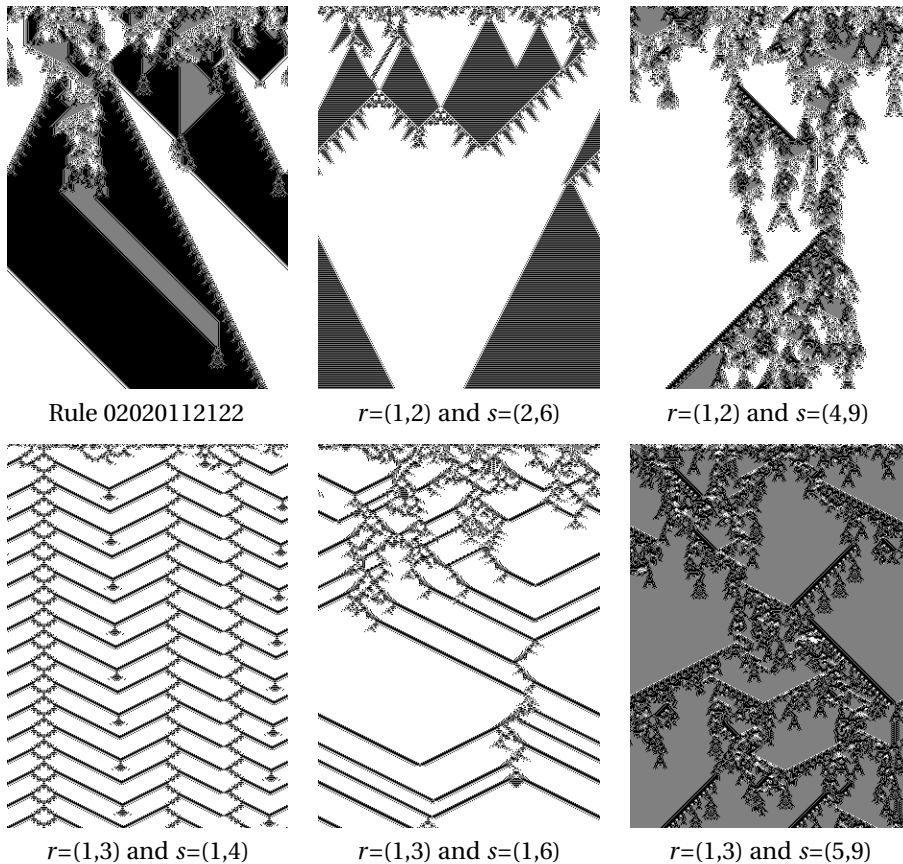


Fig. 8: First panel: A totalistic ordinary CA. Subsequent panels: Space time diagrams for equitotalistic dynamic-neighborhood CA with rule parameters in (29) and as given.

these. See the space time diagram in the first panel of Fig. 8. In the other panels are shown various equitotalistic dynamic neighborhood CA based on the rule in (29). The state configurations are shown. As can be seen in Fig. 8 all these extensions have a complex behavior, but of quite distinct flavors.

5.3 Example 3 — A two-level 2-CA

In this example the 2-CA has a 2-morphology. It has no 2-neighborhood, but must necessarily have a 2-rule. To make the example as illustrative and specific as possible we define it directly as a finite 2-CA with periodic boundary conditions. That is, the lattices are circle graphs rather than the group \mathbb{Z} . To be able to add elements of the

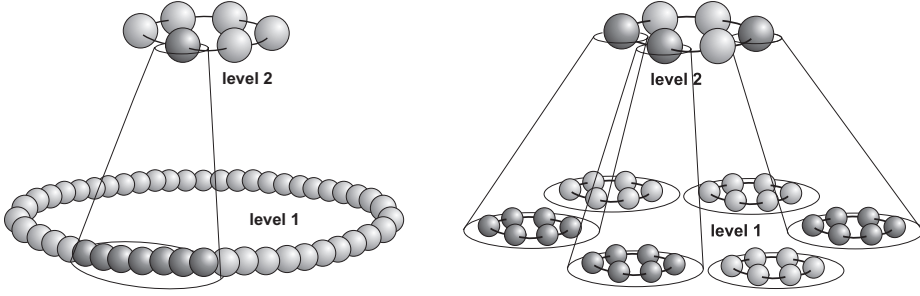


Fig. 9: Two possible morphological structures for Example 3. Left: All cells are connected in a large ring ($\mathcal{L}_1 = \mathbb{Z}_{589}$). Right: only the cells within each 2-cell are connected ($\mathcal{L}_1 = \coprod \mathbb{Z}_{19}$).

lattices we regard an n -cell lattice as the group $\mathbb{Z}/n\mathbb{Z}$, often written \mathbb{Z}_n .

Let $\mathcal{L}_1 = \mathbb{Z}_{589}$ and $\mathcal{L}_2 = \mathbb{Z}_{31}$, and note that $589 = 31 \cdot 19$. The 2-morphology is defined through the map M in the following way

$$M(u) = \{19u, 19u + 1, \dots, 19u + 18\}. \quad (30)$$

Thus, each 2-cell contains 19 adjacent cells (i.e., $c = 19$). An illustration of this two level structure is found in the left panel of Fig. 9.

Furthermore, let $\mathcal{A}_1 = \mathcal{A}_2 = \{0, 1\}$ be the alphabets of the two levels. Let $\mathcal{N} = \{-2, -1, 0, 1, 2\}$ be the first level neighborhood and $N_2 = \{-1, 0, 1\}$ be the second level neighborhood. Let the 2-rule $\mathcal{R} = \{f_a, f_b\}$ of the system consist of two totalistic rules, exactly as in Example 1. Consequently, f_a is given by the parameters $a_0 \dots a_5$ and f_b by $b_0 \dots b_5$.

The rule f_2 that updates the states of the 2-cells is of the form $f_2 : \mathcal{A}_2^3 \times \mathcal{A}_1^{19} \rightarrow \mathcal{A}_2$. We chose to let f_2 only depend on the parity of sum of the 19 cell values and depend in a totalistic way on the neighborhood. Consequently, f_2 is defined by eight parameters $c_{i,j}$ with $0 \leq i \leq 3$ and $j \in \{0, 1\}$, such that

$$f_2((y_{-1}, y_0, y_1), (x_1, \dots, x_{19})) = c_{i,j}, \quad (31)$$

where $i = \sum y_i$ and $j = \sum x_i \pmod{2}$.

Finally, we have to consider the 2-transition map ϕ , which now is of the form $\phi : \mathcal{A}_1^5 \times \mathcal{A}_2 \times \mathcal{R} \rightarrow \mathcal{R}$. We chose to let ϕ only depend on the value of the 2-cell, such that $\phi : \mathcal{A}_2 \rightarrow \mathcal{R}$, and simply put

$$\phi(0) = f_a \quad \text{and} \quad \phi(1) = f_b. \quad (32)$$

As a specific example, put

$$\begin{aligned} c_{0,0} \dots c_{3,0} &= 0010, \\ c_{0,1} \dots c_{3,1} &= 1110, \end{aligned} \quad (33)$$

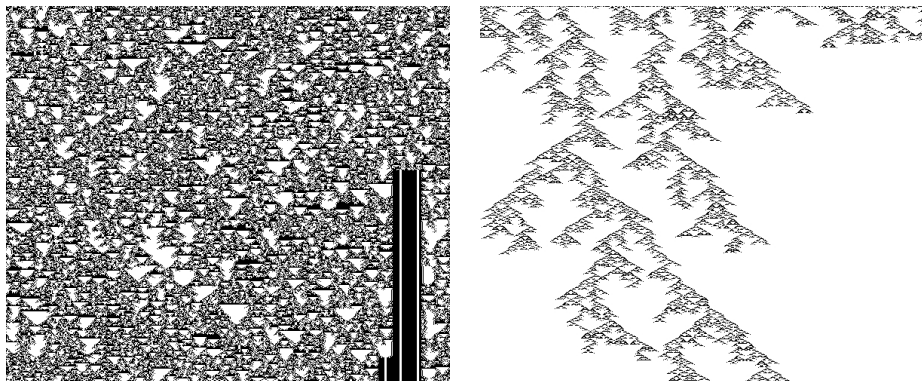


Fig. 10: Space-time diagrams for level 1 of the 2-CA from Example 3. Rule parameters for the left diagram is given in (34) and for the right diagram in (35).

and

$$a_0 \dots a_5 = 010111, \quad b_0 \dots b_5 = 001010. \quad (34)$$

A space time diagram of this 2-CA is shown in the left panel of Fig. 10. Initially the behavior is disordered, but gradually cells settle in a stable pattern. All 19 cells in a given 2-cell settle at the same time. If the evolution is continued the cells of all 2-cells will eventually settle. However, this will for most initial configurations take several thousand iterations. In the situation displayed in Fig. 10 the cells contained in 2-cells 28 and 29, out of a total of 31, settle after 220 iterations and the cells contained in 2-cell 27 settle after an additional 250 iterations. Neither the rule 010111 nor the rule 001010 considered as ordinary totalistic CA yields chaotic behavior. Another example is given by retaining the $c_{i,j}$ parameters and putting

$$a_0 \dots a_5 = 001100, \quad b_0 \dots b_5 = 010000. \quad (35)$$

A space-time diagram for this system is shown in the right panel of Fig. 10. Notice that the 2-rule here is exactly the same as in Example 1 in Sec. 5.1.

It would be natural to allow each level in a multilevel HOCA to consist of a disjoint union of lattices rather than merely a single lattice. We will exemplify this in the current setting and assess the effect this structural change has on the dynamical behavior of the system.

In our system we now let each 2-cell consist of a closed ring of cells. We still have $\mathcal{L}_2 = \mathbb{Z}_{31}$, but the first level lattice is now expressed as

$$\mathcal{L}_1 = \coprod_{i=1}^{31} \mathbb{Z}_{19}, \quad (36)$$

where \coprod means disjoint union. The map M is the same. This morphological structure is illustrated in the right panel of Fig. 9.

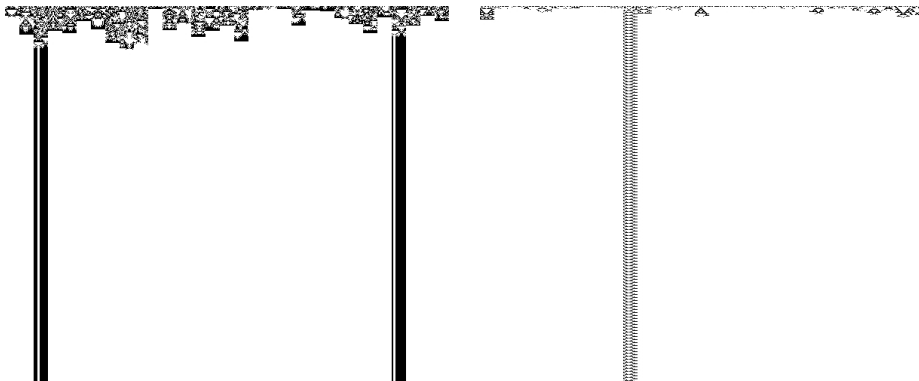


Fig. 11: Space-time diagrams for level 1 of the 2-CA from Example 3, with disjoint \mathcal{L}_1 as given in (36). The parameters are the same as for the diagrams in Fig. 10.

Figure 11 shows space-time diagrams for the same systems as in Fig. 10 but now with the first level interaction structure modified to one disjoint lattice per 2-cell. As can be seen in the diagrams, the state configuration now rapidly settles into an almost fixed configuration.

5.4 Example 4 — A higher order majority vote CA

While the previous examples were purely mathematical, we end this chapter by outlining an application of our HO-concept by developing a 2-CA inspired by a real situation. Here, 2-cells are aggregates of cells which play an important part in the dynamics, and yet must be given a priori as their existence is not related to this dynamics. The example is based on the popular majority vote CA (e.g., [9, 10]).

The majority vote CA is a simple model of the opinion on a certain issue spreading in a population. The cells represent individuals. There are k alternative opinions, represented by the CA alphabet $\mathcal{A} = \{0, \dots, k - 1\}$. The neighborhoods represent social connections (friends), and the CA rule is constructed such that each individual adjusts her opinion to better match those of her friends. This rule may be deterministic or stochastic.

The simplest deterministic majority vote CA is given by a one-dimensional lattice $\mathcal{L} = \mathbb{Z}$, two possible opinions, $\mathcal{A} = \{0, 1\}$, a neighborhood $N = \{-r, \dots, r\}$, a parameter d satisfying $0 \leq d \leq 2r$ and the transition rule $g : \mathcal{A}^{2r+1} \rightarrow \mathcal{A}$ given by

$$g(x_{-r}, \dots, x_r) = \begin{cases} 1 & \text{if } \Sigma(x) > d, \\ 0 & \text{otherwise.} \end{cases} \quad (37)$$

Here $\Sigma(x)$ denotes the sum of the neighborhood configuration. A sensible choice is to

put $d = r$, such that each individual changes opinion to the most common one in her neighborhood.

Now, let us introduce a HO extension. It is natural to assume that in addition to her friends each individual is influenced by the general opinion in some social organization she partakes in. Let us call this her workplace. Even though she does not know everybody at work personally, she will perceive the general opinion and might adjust to it. We assume that the influence from the workplace may lead to a possible bias in her reactions to the opinions of her friends. In particular, if she is currently in consensus with a significant majority at her workplace, it takes one more friend to sway her.

This situation can be modeled in the following way using a HOCA. Take as a basis the majority vote CA described above, and introduce a second morphological level consisting of 2-cells representing workplaces. Let $\mathcal{L}_2 = \mathbb{Z}$ and let the inner structure map M be of the same form as in Example 3, illustrated to the left in Fig. 9. That is, each 2-cell consists of c adjacent cells. Thus, most individuals have friends with whom they also work. If we wanted less individuals to have friends from work we should make M more 'fuzzy'.

Let $\mathcal{A}_2 = \{0, U, 1\}$ be the set of states a workplace can be in, where 0 and 1 means a strong majority is of opinion 0 or 1 and U means no strong majority is present. In the following, let $c = 10$ and $r = 2$ (radius of first level neighborhood). We assume there is no direct communication between different workplaces, and model this by simply putting $N_2 = \{0\}$ as the second level neighborhood set. Furthermore, define $f_2 : \mathcal{A}_2 \times \mathcal{A}^{10} \rightarrow \mathcal{A}_2$ by

$$f_2(a, x_1 \dots x_{10}) = \begin{cases} 1 & \text{if } \Sigma(x) > 6, \\ 0 & \text{if } \Sigma(x) < 4, \\ U & \text{otherwise.} \end{cases} \quad (38)$$

We introduce a 2-rule $\mathcal{R} = \{g_0, g_U, g_1\}$ simply by letting g_0 be as in equation (37) with $d = 3$, g_U the same rule with $d = 2$ and g_1 with $d = 1$. We let 2-transition rule ϕ depend only on the value of the 2-cell such that $\phi : \mathcal{A}_2 \rightarrow \mathcal{R}$ and put

$$\phi(i) = g_i. \quad (39)$$

Note that the influence from the workplace has small delay since rather than being accessed directly by the cell the value of the 2-cell changes which local rule that is used by the cell.

We have simulated this model with 200 cells. At the top of Fig. 12 is shown a space time diagram of the ordinary radius 2 majority vote CA from (37) with $d = 2$. The initial configuration \mathbf{a} is uniformly random. As can be seen, the dynamics very quickly reaches a fixed point. The middle panel shows evolution from the same \mathbf{a} when the HO extension described above is applied. The initial rule configuration \mathbf{r} and second level state configuration \mathbf{a}_2 are calculated from \mathbf{a} . In the bottom panel the regular inner structure map M described above is replaced by a random assignation of cells

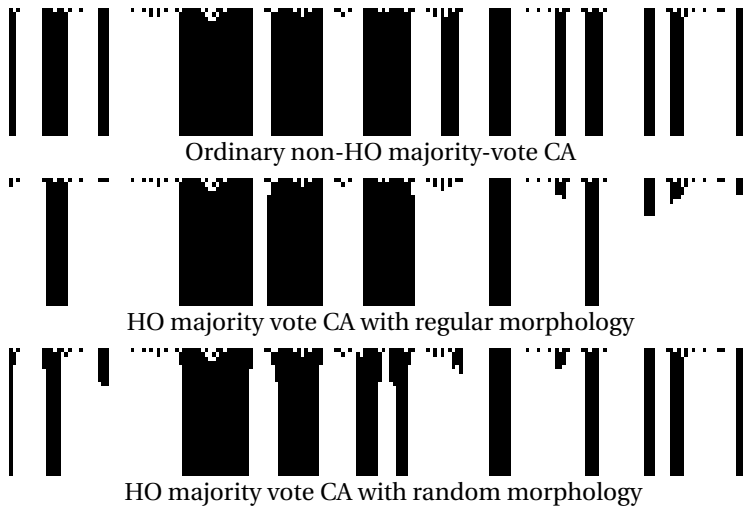


Fig. 12: Space time diagrams for various majority vote systems starting from the same initial configuration.

to 2-cells, so that an individual's set of friends is unrelated to her workplace. For instance, the members of the first 2-cell are cells 8, 41, 43, 49, 85, 115, 127, 142, 146 and 196. No other component of the rule is altered.

As can be seen from the figure, the introduction of a HO structure has a small but noticeable effect on the dynamics. In the case of regular morphology some of the small groups of people with opinion 1 disappear. This happens due to individuals in these groups adjusting to their workplaces. The same effect can be seen in the case with random morphology, but to a lesser extent. Thus, in the two level situation global mixing through workplaces leads to a smaller degree of local consensus compared to the case of local workplaces. The same trend is observed quite consistently also in other runs.

A possible higher order extension that only consists of a 2-dynamics, would be to let individuals modify their set of friends in reaction to some events. This will naturally lead to a 2-neighborhood. For instance, an individual may stop caring about the opinion of a friend if at some time all other friends and herself is of the opposite opinion. Technically, this is modeled by introducing the following 2-transition map: whenever only a single cell v in the neighborhood of cell u has a given value, the neighborhood of u is changed not to include v anymore. This introduces a memory into the system. A possible extension of the 2-morphology would be to let individuals partake in more than one group. Moreover, the majority vote model generally yields more interesting results in two dimensions or in a stochastic setting. Nevertheless, we feel that the present example provides a good illustration of the applicability the

HO-concept and more particular HOCA to modeling.

We should remark that for more advanced and detailed spatial models more flexible and less static structures than CA would probably be required. However, this would involve sacrificing mathematical tractability.

6 Extensions and Generalizations

There are many natural extensions of 2-CA. Some are evident, and some not so evident. We do not give details here, but briefly summarize a number of generalizations below.

- Allowing 2-cells to contain different numbers of cells, and modifying the second level dynamics to accommodate this.
- Allowing cells to be in more than one 2-cell or no 2-cell and modifying the map ϕ to accommodate this.
- Allowing a second order dynamics also on the second level. This involves defining sets \mathcal{N}_2 and \mathcal{R}_2 and a map ϕ_2 in a way parallel to the approach described for the first level.
- Allowing for any number of levels. Generally, an n -cell $u \in \mathcal{L}_n$ will be a collection of $n-1$ -cells in \mathcal{L}_{n-1} defined by a morphology map M_n . The dynamics are extended to take this structure into account.
- Rather than having a single lattice on each level it is natural to allow for a disjoint union of sublattices, written $\mathcal{L}_n = \coprod_k \mathcal{L}_{n,k}$. This was applied in Example 3 in Sec. 5.3.
- Exchanging the lattices with circle graphs to obtain finite HOCA, analogous to finite CA. This will be necessary when performing computer simulations.
- Allowing for general interaction graphs G_n on each level instead of restricting the focus to regular lattices. This will produce higher order graph CA.
- One can easily imagine synthesizing 2-CA in a dynamical way. This could involve extending our scheme by adding maps that cluster cells into new organs when the local conditions are suitable as well as maps that remove cells from 2-cells and incorporate cells into existing 2-cells. This approach might be useful for illuminating questions of type (1) in the list in Sec. 2.

Finally, as stated in the introduction the HO concept is applicable to general dynamical systems defined on spatial structures such as lattices or graphs, not only to cellular automata. For example, one can construct higher order sequential dynamical systems (HO-SDS), e.g. [7], and higher order dimer automata, see [26], along the same lines as we have done for CA.

7 Conclusions and Prospects

In this article we have introduced our higher order concept by equipping the class of dynamical systems known as cellular automata with a higher order dynamics as well as a higher order morphology. The first case gives the system a meta-dynamics where the update rule and input domain varies locally and is governed by a 2-transition function. The second case gives the system an additional level consisting of 2-cells, which are families of cells. In [4] this is put into a more general framework of higher order systems and structures. The present paper illustrates the meaning of dynamical hierarchies in this context and how they might work.

We have studied four classes of 2-CA and made several interesting observations. Firstly, it seems that the HO approach is a viable way to easily construct rule schemes leading to complex behavior. An interesting question is whether complex behavior, in a sensible definition of the term, is more common within the class of 2-CA than in ordinary CA. Our results indicate that new types of complexities are being created. Secondly, new phenomena like crystallization is commonly observed in 2-CA. One may speculate whether 2-CA dynamics is more “regularizing” than ordinary CA dynamics in some sense which has to be made precise. Thirdly, Example 3 shows that the coupling geometry of the system has a fundamental impact on the resulting behavior. In Example 4 we demonstrated the applicability of the HO concept in modeling.

There are many paths of research that would be interesting to pursue in future work with 2-CA. Firstly, to study the relation between an ordinary CA and various HO extensions of it. Which aspects of the extensions have the most significant impact on the behavior of the system? Furthermore, to construct some reasonable formal classification of 2-CA, and relate the classes to rule parameters. Unfortunately, cellular automata are generally hard to analyze and classify due to the extreme discreteness of their construction, and 2-CA will be no easier to handle. One possible approach to attack the above challenges and problems is comparison by suitable entropy concepts and notions of complexity in dynamical systems. In particular, the notion of local information, which was introduced and used to study information flow in ordinary CA in [15], may be extended to HOCA. It could also be useful to study localized structures, invariant subshifts and various notions of attractors [13, 20]. A computer science oriented approach would be to consider computational complexities (e.g., [28]).

Initially, the motivation of the work is to gain insight into how the introduction of various higher order structures into a system affects the dynamical behavior of the system. Is there any rule of thumb on what typically happens? On a larger timescale, the work may become useful for applications to distributed systems with a hierarchical structure, such as traffic and communications networks, social and population dynamics and biological systems such as the brain. Furthermore, we believe that HO dynamical systems may become very useful in the modeling and simulation of new materials, for example, with a hierarchical structure. Such materials now seem to be

constructible with modern nanotechnological methods.

We would like to end this article by encouraging the reader to take her favorite system, put 2 in front of everything, and see what happens in the new 2-world!

Acknowledgements

This research has been partly funded by the Research Council of Norway.

References

- [1] Aldana, M., Coppersmith, S., and Kadanoff, L. P. Boolean dynamics with random couplings. In *Perspectives and problems in nonlinear science*, pages 23–89. Springer, New York, 2003.
- [2] Baas, N. A. Emergence, hierarchies and hyperstructures. In Langton, C. G., editor, *Artificial Life III, Santa Fe Studies in the Sciences of Complexity, Proc. Volume XVII*, pages 515–537, Reading, MA, 1994. Addison-Wesley.
- [3] Baas, N. A. Self-organization and higher order structures. In Schweitzer, F., editor, *Self-Organization of Complex Structures: From Individual to Collective Dynamics*, pages 71–81. Taylor & Francis, 1997.
- [4] Baas, N. A. Abstract Matter. Preprint NTNU, 2004.
- [5] Baas, N. A., Ehresmann, A., and Vanbremeersch, J.-P. Hyperstructures and memory evolutive systems. *Int. J. Gen. Syst.*, 33(5):553–568, 2004.
- [6] Badii, R. and Politi, A. *Complexity: Hierarchical Structures and Scaling in Physics*. Cambridge University Press, Cambridge, 1997.
- [7] Barrett, C., Mortveit, H., and Reidys, C. Elements of a theory of simulation. II: Sequential dynamical systems. *Appl. Math. Comput.*, 107(2-3):121–136, 2000.
- [8] Culik II, K. and Yu, S. Undecidability of CA classification schemes. *Complex Systems*, 2:177–190, 1988.
- [9] Durrett, R. and Steif, J. E. Fixation results for threshold voter systems. *Ann. Probab.*, 21(1):232–247, 1993.
- [10] Eloranta, K. Voter dynamics in deterministic cellular automata. *Front. Pure Appl. Prob.*, 8:51–58, 1996.
- [11] Gutowitz, H. A hierarchical classification of CA. *Physica D*, 45:136, 1990.
- [12] Gutowitz, H., editor. *Cellular Automata: Theory and Experiment (Special Issue of Physica D)*. MIT Press, 1991.
- [13] Hanson, J. E. and Crutchfield, J. P. The attractor-basin portrait of a cellular automaton. *J. Stat. Phys.*, 66:1415–1462, 1992.

- [14] Hedlund, G. A. Endomorphisms and automorphisms of the shift dynamical system. *Math. Systems Theory*, 3:320–375, 1969.
- [15] Helvik, T., Lindgren, K., and Nordahl, M. G. Local information in one-dimensional cellular automata. In Sloot, P. M. A., Chopard, B., and Hoekstra, A. G., editors, *Cellular Automata*, volume 3305 of *Lecture Notes in Computer Science*, pages 121–130, Heidelberg, 2004. Springer-Verlag.
- [16] Kaneko, K. *Theory and applications of coupled map lattices*. John Wiley & Sons, New York, 1993.
- [17] Koestler, A. *Janus. A Summing Up*. Hutchinson & Co. Ltd., 1978.
- [18] Kopell, N. and Ermentrout, G. Phase transitions and other phenomena in chains of coupled oscillators. *SIAM J. Math. Anal.*, 50(4):1014–1052, 1990.
- [19] Kurka, P. Languages, equicontinuity and attractors in cellular automata. *Ergod. Th. & Dynam. Syst.*, 17(2):417–433, 1997.
- [20] Kurka, P. and Maass, A. Limit sets of cellular automata associated to probability measures. *J. Stat. Phys.*, 100(5-6):1031–1047, 2000.
- [21] Liggett, T. *Interacting Particle Systems*, volume 276 of *Grundlehren der mathematischen Wissenschaften*. Springer, 1985.
- [22] Milnor, J. On the entropy geometry of cellular automata. *Complex Systems*, 2:357–386, 1988.
- [23] Munkres, J. R. *Topology*. Prentice Hall, 2nd edition, 2000.
- [24] Pérez, C. J., Corral, A., Díaz-Guilera, A., Christensen, K., and Arenas, A. On self-organized criticality and synchronization in lattice models of coupled dynamical systems. *Int. J. Mod. Phys. B*, 10:1111–1151, 1996.
- [25] Rasmussen, S., Baas, N. A., Mayer, B., Nilsson, M., and Olesen, M. W. Ansatz for dynamical hierarchies. *Artificial Life*, 7:329–353, 2001.
- [26] Schöfisch, B. and Haderer, K. Dimer automata and cellular automata. *Physica D*, 94(4):188–204, 1996.
- [27] Scott, A. *Stairway to the Mind*. Springer-Verlag, 1995.
- [28] Sutner, K. On the computational complexity of finite cellular automata. *J. Comput. System Sci.*, 50:87–97, 1995.
- [29] Wolfram, S. Universality and complexity in cellular automata. *Physica D*, 10:1–35, 1984.
- [30] Wolfram, S. *Cellular Automata and Complexity*. Addison-Wesley, 1994.

Paper III

Local Information in One-Dimensional Cellular Automata

Torbjørn Helvik, Kristian Lindgren and Mats G. Nordahl (2004)

In *Cellular Automata: ACRI 2004 proceedings*, Vol. 3305 of *Lecture Notes in Computer Science*, eds. P. M. A. Sloot, B. Chopard and A. G. Hoekstra, Springer-Verlag, Heidelberg, pp. 121–130.

Local Information in One-Dimensional Cellular Automata

Torbjørn Helvik^{1,2}, Kristian Lindgren^{2,*}, and Mats G. Nordahl³

¹Department of Mathematical Sciences
Norwegian University of Science and Technology
NO-7491 Trondheim, Norway

²Department of Physical Resource Theory
Chalmers University of Technology and Göteborg University
SE-41296 Göteborg, Sweden

³Innovative Design
Chalmers University of Technology and Göteborg University
SE-41296 Göteborg, Sweden

Abstract

A local information measure for a one-dimensional lattice system is introduced, and applied to describe the dynamics of one-dimensional cellular automata.

1 Introduction

Cellular automata (CA) are spatially extended dynamical systems. They have been widely used as models of spatial dynamical processes [31]. They have also been used as simple systems in which to study the phenomenon of self-organization, i.e., the ability of a system to build up structure from homogeneous initial conditions.

In this paper we use concepts from information theory to study the dynamics of cellular automata. The field of information theory was founded by Shannon in his 1948 paper [23]. The Shannon entropy introduced in [23] is a global quantity that measures the average information content per symbol in symbol sequences generated from a given distribution, when all correlations are taken into account.

*This work has been supported by PACE (Programmable Artificial Cell Evolution), a European Project in the EU FP6-IST-FET Complex Systems Initiative.

Since the work of Boltzmann and Gibbs, entropy has been used in physics as a measure of the disorder in a system. The connections between physics and information theory are strong. For instance, in the framework of Jaynes [13], the entropy of a system with many degrees of freedom quantifies the uncertainty regarding the microstate of the system. A number of physicists, going back to Wheeler's vision of "It from Bit", have argued that the concept of information should be used to describe fundamental physical processes (e.g., [19, 29]). One area where information theory has been particularly important is in the thermodynamics of black holes (e.g., [1]).

In numerical and mathematical studies of self-organizing spatially extended systems, the microstate of the system may often be known in great detail or exactly. In this case, it is useful to define local versions of entropy or information to characterize the structure that arises in the time evolution [4, 5]. In [5], such local quantities were utilized to describe structure arising at different length scales in continuous spatiotemporal systems. Information flow in spatial dimensions and between length scales was also discussed. A number of other authors have also investigated information transport in spatially extended dynamical systems [14, 17, 21, 22, 25, 27].

If the time evolution is reversible, microscopic entropy is a conserved quantity. This is the case both for classical systems viewed in terms of phase space distributions, and in quantum mechanics. An important question is whether we can introduce a local version of the microscopic entropy that is *locally* conserved. In other words, we would like to define a corresponding local information flow so that the system obeys a local continuity equation. Since the information or entropy depends on the degree of correlation in the system, it is not immediately obvious that this can be done.

Several authors have however suggested that such an approach may be possible. This was first proposed in the context of CA by Toffoli [25], who showed the local continuity of information flow in a restricted setting of perturbations around uncorrelated equilibria of particle conserving reversible CA, such as lattice gas automata. The continuity equation for information flow in 1D CA of this paper, which takes correlations into account, was first formulated in [6].

The goal of this paper is to provide a mathematically rigorous foundation for these statements in the case of one-dimensional reversible and surjective CA rules. In Sect. 2 we introduce a local information measure for a 1D lattice system and discuss its properties. In Sect. 3 we define the information flow and derive the local continuity equation.

2 Local Information

The Shannon entropy measures the average information gained by observing a new symbol from an infinite string being read in some direction, given knowledge of the past. The intent of introducing a local information quantity is to measure exactly

how much information is contained in each symbol. However, the correlations in a symbol sequence can in general be arbitrarily long. Consequently, it is impossible for information to be completely localized. That is, the local information at a particular position i cannot always be correctly computed merely by looking at the configuration in a finite neighbourhood of i . The natural approach is therefore to define the local information as a limit which converges to a local analogue of the Shannon entropy as more and more distant neighbours are taken into account, and which can be computed locally when the correlations are finite in extent. The speed of convergence depends on the typical length of correlations in the system, which could be quantified as in [9].

In the following, \mathcal{A} is a finite set and $\mathcal{A}^{\mathbb{Z}}$ is the space of all bi-infinite sequences of symbols from \mathcal{A} . For a sequence $x \in \mathcal{A}^{\mathbb{Z}}$, we use the notation x_i^{i+k} for the length $k+1$ block $(x_i, x_{i+1}, \dots, x_{i+k})$. A probability measure μ on $\mathcal{A}^{\mathbb{Z}}$ is defined by a consistent set of block probabilities $\mu(a_i \dots a_{i+n})$ for all finite blocks $a_i \dots a_{i+n}$ of symbols in \mathcal{A} and positions $i \in \mathbb{Z}$.

We define the local information in the following way. Let μ be a probability measure on $\mathcal{A}^{\mathbb{Z}}$. Define the left local information at position i of $x \in \mathcal{A}^{\mathbb{Z}}$, conditioned on n past symbols, as

$$S_L^n(x; i) = -\log \mu(x_i | x_{i-n}^{i-1}) . \quad (1)$$

Then define the *left local information* at position i by $S_L(x; i) = \lim_{n \rightarrow \infty} S_L^n(x; i)$, and define the quantities S_R^n and S_R in the same way, but conditioned to the right instead, using $\mu(x_i | x_{i+1}^{i+n})$. Finally, define the local information S as $\frac{1}{2}(S_L + S_R)$.

The quantity $S_L^n(x; i)$ is the information gained from the symbol at position i when only knowledge of the n left symbols are assumed. If μ is Markov there is an n such that $S_L^n = S_L$. In other cases, correlations are not finite and no such n will exist. The information $S_L(x; i)$ depends on the measure μ . However, in an ergodic case, the correct probabilities of finite blocks can be recovered with probability one from x by estimating the block frequencies. In a numerical simulation, this would normally be the case.

A position i in a sequence has high local information if its symbol is unexpected given knowledge of already observed symbols. We can illustrate this using two examples of CA dynamics. For some CA, the time evolution can after an initial transient be described in terms of domains of some background pattern together with moving domain boundaries. The background pattern typically consists of blocks from a regular language which is invariant under the CA (e.g., [10]). The domain boundaries are local configurations which are often periodic and propagate with fixed velocity (often called particles), or move at random depending on the local context. They may be destroyed, transformed or created during collisions. In terms of local information, the background pattern generated by the CA will have low information, common particles will have higher information, rarer particles will have even higher and extremely rare events will have very high information.

An example is elementary CA rule 110, using Wolfram's rule numbering system [30], which generates a background pattern of large spatial and temporal periodicity and a large diversity of particles. A space-time diagram and a numerical estimate of the local information for rule 110 is shown in Fig. 1. As can be seen, local information acts as a filter for the space-time diagram and quantifies the structure that is built up.

An example that shows a different type of spatial structure is elementary CA rule 18. Here, domain boundaries (kinks) perform random walks and annihilate pairwise upon collision [3, 8]. The background pattern is in this case the regular language consisting of all blocks without consecutive 1's. The CA acts as the additive rule 90 on the background pattern. A space-time diagram for rule 18 along with the corresponding local information is found in Fig. 1. Notice that although the background in this case has non-zero entropy it is still filtered out by the local information. For both rule 110 and rule 18, the probability measure at each time step was estimated from the block frequencies for blocks of length 14 in a CA run with a sequence of length 10^7 . The initial measures were iid product measures with the probability of a 1 being respectively 0.5 and 0.2.

For the definition of local information to make sense, we need to check that it converges with n . The convergence of the local information S_L^n to S_L follows immediately from results by Khinchin (Lemma 7.2 and Lemma 7.7) given in connection with a proof of the entropy theorem in [15]. When applied in our setting, these prove the existence of the local information in the following sense: Fix $i \in \mathbb{Z}$. For all n , $S_L^n(\cdot; i) \in L^1(\mu)$, and $S_L^n(\cdot; i) \rightarrow S_L(\cdot; i)$ almost everywhere and in the L^1 -norm. By symmetry, the same is true for $S_R^n(\cdot; i)$.

Another requirement of a local information measure is that in the case of a translation invariant measure μ the average of the local information should equal the Shannon entropy h . Write $H(X|Y)$ for the conditional entropy of the random variable X given Y . We can then show that $E[S_L^n] = H(x_0|x_{-n}^{-1})$ and $E[S_L] = h$. The same is true for S_R and S .

The first statement is proved by the calculation

$$\int_{\mathcal{A}^{\mathbb{Z}}} S_L^n d\mu = - \sum_{x_{-n}^0} \mu(x_{-n}^0) \left(\log \frac{\mu(x_{-n}^0)}{\mu(x_{-n}^{-1})} \right) = H(x_0|x_{-n}^{-1}) . \quad (2)$$

The second statement then follows from the L^1 convergence of local information and the fact that $H(x_0|x_{-n}^{-1}) \rightarrow h$ [2, Ch. 4.2]. The equality for S_R and S follows since $H(x_0|x_{-n}^{-1}) = H(x_0|x_1^n)$.

3 Information Flow

We now consider the time evolution of the local information introduced in the previous section. In this paper, we primarily discuss reversible CA rules (e.g., [26]), where

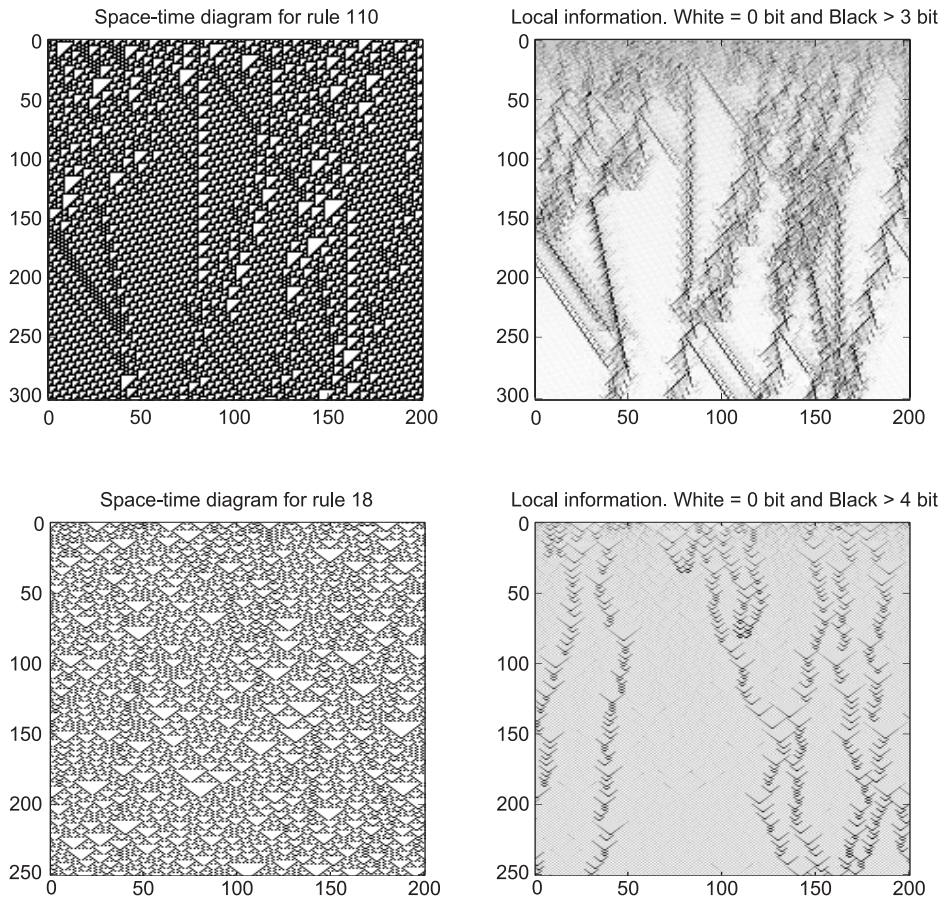


Fig. 1: Space-time diagrams and the corresponding local information S^{14} for elementary CA rules 110 and 18.

the global CA mapping on bi-infinite sequences has a unique inverse. Since microscopic classical dynamics in physics is reversible, this is a very important class of CA for modelling physical phenomena. Some well-known physical CA models are lattice gas automata [7] and CA for simulating spin systems such as the Q2R rule [28]. Simple nearest-neighbour reversible CA have in particular been studied by Takesue (e.g., [24]). The elementary reversible CA rule 26 in [24] can be considered as the simplest known system with plausible thermodynamic behaviour when one interprets its locally conserved quantity as energy.

We also consider the more general class of surjective cellular automata, where the global mapping is onto. In this case, the number of preimages of any infinite

sequence is bounded [11]. Since there are interesting examples of surjective rules among the elementary CA (both additive rules, and more complicated examples such as rule 30), but only trivial examples of reversible rules, simple illustrative examples of CA behaviour are often from this class. For surjective CA, and thus also for reversible CA, the global Shannon entropy (as well as all Rényi entropies) is preserved in the time evolution [16, 18]. For non-surjective rules, where infinite sequences typically have an infinite number of preimages, the entropy decreases in time reflecting the shrinking phase space.

Our aim is to show that the local information introduced above obeys a local continuity equation under the time evolution of a CA, or in other words, that there is an analogue of the continuity equations which apply to locally conserved physical quantities such as charge or particle density. We show how a flow J_L that satisfies a continuity equation can be constructed. We consider the case of a reversible CA F , but the construction can also be carried out for surjective CA (see [12] for details).

Fix $x \in \mathcal{A}^{\mathbb{Z}}$, and let $S_L(t; i) = S_L(F^t(x); i)$. We will show that there is a well-defined function $J_L(t; i) = J_L(F^t(x); i)$ such that

$$\Delta_t S_L + \Delta_i J_L = 0 \quad . \quad (3)$$

Here, Δ is the forward difference operator, so that $\Delta_t S_L = S_L(t+1; i) - S_L(t; i)$ and $\Delta_i J_L = J_L(t; i+1) - J_L(t; i)$. With these definitions, $J_L(t; i)$ is the information flow from position $i-1$ to position i in iteration $t+1$ of the CA. Since the flow is associated with a change between two time steps (just like other currents in physics which involve a time derivative), it will naturally involve the state and measure of the CA at $t+1$ as well as t , even though it can be expressed solely in terms of quantities defined at t ,

In the following, we consider a reversible CA F with left radius m , right radius r and local map $f: \mathcal{A}^{m+r+1} \rightarrow \mathcal{A}$. The inverse of a reversible CA is also a CA [20]. Let the inverse CA \tilde{F} of F have local map \tilde{f} with left radius M and right radius R . Fix $x \in \mathcal{A}^{\mathbb{Z}}$, and let $y = F(x)$.

In general, the measure μ changes in time due to the cellular automaton time evolution, unless we happen to start from an invariant measure for the CA. Let μ^0 be a measure on $\mathcal{A}^{\mathbb{Z}}$ and let $\mu^t = F^t(\mu^0)$. The evolution of μ^t is given by the standard relation (e.g., [18])

$$\mu^{t+1}(y_{i-n}^i) = \sum_{f^{-1}(y_{i-n}^i)} \mu^t(x_{i-n-m}^{i+r}) \quad (4)$$

where $f^{-1}(y_{i-n}^i)$ is the set of all blocks x_{i-n-m}^{i+r} that map to y_{i-n}^i under f . For simplicity of notation, we consider the information flow at spatial coordinate 0 in the first iteration of F , but this can trivially be changed to an arbitrary position i and an arbitrary time step t to $t+1$.

We also define the joint measure ν of two consecutive time steps as the measure on $(\mathcal{A} \times \mathcal{A})^{\mathbb{Z}}$ defined by

$$\nu(x_{-n}^0, y_{-n}^0) = \mu^0(\{z_{-n-m}^r \in \mathcal{A}^{m+n+r+1} \mid z_{-n}^0 = x_{-n}^0 \text{ and } f(z_{-n-m}^r) = y_{-n}^0\}) \quad . \quad (5)$$

It is easy to show that ν actually is a measure, and furthermore that ν is translation invariant if μ^0 is translation invariant. Note that by summing over all possible y_{-n}^0 or x_{-n}^0 we obtain from the definition $\nu(x_{-n}^0) = \mu^0(x_{-n}^0)$ and $\nu(y_{-n}^0) = \mu^1(y_{-n}^0)$.

From the definition of local information we have

$$\Delta_t S_L(0) = \lim_{n \rightarrow \infty} (-\log \mu^1(y_0 | y_{-n}^{-1}) + \log \mu^0(x_0 | x_{-n}^{-1})) . \quad (6)$$

Adding and subtracting the same terms, we can express this change as

$$\begin{aligned} \Delta_t S_L(0) &= \lim_{n \rightarrow \infty} \left(\log \frac{\nu(x_0, y_0 | x_{-n+M}^{-1}, y_{-n}^{-1})}{\mu^1(y_0 | y_{-n}^{-1})} - \log \frac{\nu(x_0, y_0 | x_{-n}^{-1}, y_{-n+m}^{-1})}{\mu^0(x_0 | x_{-n}^{-1})} + \gamma_n \right) \\ &= \lim_{n \rightarrow \infty} \left(\log \frac{\nu(x_{-n+M}^0 | y_{-n}^0)}{\nu(x_{-n+M}^{-1} | y_{-n}^{-1})} - \log \frac{\nu(y_{-n+m}^0 | x_{-n}^0)}{\nu(y_{-n+m}^{-1} | x_{-n}^{-1})} + \gamma_n \right) \end{aligned} \quad (7)$$

where

$$\gamma_n = -\log \frac{\nu(x_0, y_0 | x_{-n+M}^{-1}, y_{-n}^{-1})}{\nu(x_0, y_0 | x_{-n}^{-1}, y_{-n+m}^{-1})} \quad (8)$$

converges to zero almost everywhere as $n \rightarrow \infty$. This can be proved using martingale methods. See [12] for details.

We can now write

$$\begin{aligned} \Delta_t S_L(0) &= \lim_{n \rightarrow \infty} (-\log \nu(y_{-n+m}^0 | x_{-n}^0) + \log \nu(y_{-n+m}^{-1} | x_{-n}^{-1}) \\ &\quad + \log \nu(x_{-n+M}^0 | y_{-n}^0) - \log \nu(x_{-n+M}^{-1} | y_{-n}^{-1})) . \end{aligned} \quad (9)$$

Let us show that two of these terms have well-defined limits. The other two can be resolved in the same way. The first term can be written as

$$-\lim_{n \rightarrow \infty} \log \nu(y_{-n+m}^0 | x_{-n}^0) = -\lim_{n \rightarrow \infty} \log \mu^0(x_1^r \in B | x_{-n}^0) \quad (10)$$

where $B \subset \mathcal{A}^r$ is the set of extensions to the right of x_{-n}^0 that are compatible with the image sequence y_{-n}^0 . That is, $B = \{z_1^r \in \mathcal{A}^r | f(x_{-m-r}^0, z_1^r) = y_{-r}^0\}$. The right hand side of (10) converges because $\mu^0(x_1^r | x_{-n}^0)$ converges for each block x_1^r . This follows from the convergence of local information. If $r > 0$ there is at least one block for which this conditional probability is non-zero, otherwise x will have infinite local information at some position. If $r = 0$ we define $\mu^0(x_1^0 \in B | x_{-n}^0)$ to be 1. If we denote by Z_0 the set of all preimages $z_{-\infty}^r$ of $y_{-\infty}^0$ that satisfies $z_{-\infty}^0 = x_{-\infty}^0$, (10) becomes

$$-\lim_{n \rightarrow \infty} \log \nu(y_{-n+m}^0 | x_{-n}^0) = -\log \sum_{Z_0} \mu^0(z_1^r | x_{-\infty}^0) . \quad (11)$$

Using the reversibility of the CA, we have analogously that

$$\lim_{n \rightarrow \infty} \log \nu(x_{-n+M}^0 | y_{-n}^0) = \lim_{n \rightarrow \infty} \log \mu^1(y_1^R \in C | y_{-n}^0) \quad (12)$$

where $C \subset \mathcal{A}^R$ is the set of extensions of $y_{-\infty}^0$ to the right that are compatible with the preimage $x_{-\infty}^0$. By the same argument as above the right hand side converges. We can also express $\lim_{n \rightarrow \infty} \log v(x_{-n+M}^0 | y_{-n}^0)$ purely in terms of x and μ^0 . In fact, one can prove that

$$\lim_{n \rightarrow \infty} \log \mu^1(y_1^R \in C | y_{-n}^0) = \log \frac{\sum_{Z_0} \mu^0(z_{-R+1}^R | x_{-\infty}^{-R})}{\sum_Z \mu^0(z_{-R+1}^R | x_{-\infty}^{-R})} \quad (13)$$

where Z_0 is as above and Z is the set of *all* preimages of $y_{-\infty}^0$. Since F is reversible and R is the right radius of the inverse, all preimages agree on $z_{-\infty}^{-R}$, but might be different to the right of coordinate $-R$.

Equations (9), (11) and (13) now enable us to define the locally conserved information flow

$$J_L(i) = J_L^+(i) - J_L^-(i) \quad (14)$$

where

$$J_L^-(i) = -\log \sum_{Z_0} \mu^0(z_i^{i+r-1} | x_{-\infty}^{i-1}) , \quad (15)$$

$$J_L^+(i) = -\log \frac{\sum_{Z_0} \mu^0(z_{i-R}^{i+r-1} | x_{-\infty}^{i-R-1})}{\sum_Z \mu^0(z_{i-R}^{i+r-1} | x_{-\infty}^{i-R-1})} . \quad (16)$$

Here, Z are the preimages of $y_{-\infty}^{i-1}$ and Z_0 are those preimages that satisfies $z_{-\infty}^{i-1} = x_{-\infty}^{i-1}$. Both $J_R^+(i)$ and $J_R^-(i)$ are non-negative.

By using these definitions, (9) can be written as the continuity equation

$$\Delta_t S_L + \Delta_i J_L = 0 \quad (17)$$

which is valid at every i and t and for all measures μ^0 , not necessarily translation invariant.

The expressions J_L^- and J_L^+ have natural information theoretic interpretations. This is more readily apparent if the flow is written in terms of quantities defined both at t and $t+1$:

$$J_L(1) = J_L^+(1) - J_L^-(1) = \lim_{n \rightarrow \infty} (-\log v(x_{-n+M}^0 | y_{-n}^0) + \log v(y_{-n+M}^0 | x_{-n}^0)) . \quad (18)$$

The quantity $-\log v(y_{-n+M}^0 | x_{-n}^0)$ is the additional information needed to uniquely specify y_{-n+M}^0 when knowing x_{-n}^0 . The definitions above can be viewed as a choice of boundary conditions in the definition of the flow. In the case of $\Delta S_L(1)$, we only want the contribution from information flow across the boundary between 0 and 1 and not the boundary at $-n$. Due to the condition on the left boundary of the finite strings (knowing m more symbols of x than of y , see Fig. 2) the left part of y_{-n+M}^0 is uniquely determined by x_{-n}^0 , and we avoid the unwanted effect of an information flow across

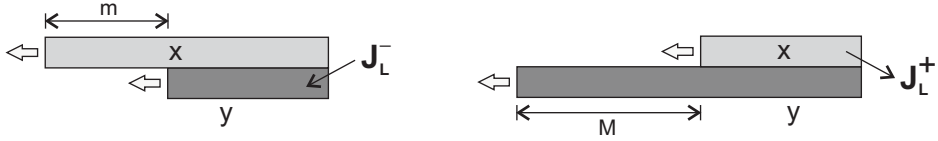


Fig. 2: Left: The information $J_L^-(i)$ needed to determine $y_{-\infty}^{i-1}$ which is not found in $x_{-\infty}^{i-1}$ flows to the left across the boundary. Right: The information $J_L^+(i)$ in $x_{-\infty}^{i-1}$ which is not present in $y_{-\infty}^{i-1}$ leaks out to the right across the boundary.

the left border. Therefore, the remaining information needed to determine the right part of y_{-n+m}^0 must come from the coordinates to the right of 0. Consequently, a flow $J_L^-(1)$ from coordinate 1 to 0 is induced. The limit $n \rightarrow \infty$ must be taken to ensure that all information contained in $x_{-\infty}^0$ about its left prolongation is taken into account.

The quantity $-\log v(x_{-n+M}^0 | y_{-n}^0)$ is the additional information needed to uniquely specify x_{-n+M}^0 given knowledge of y_{-n}^0 . Once again the left boundary conditions ensure that the left part of x_{-n+M}^0 is uniquely determined, see Fig. 2. However, not all information about the right part of x_{-n+M}^0 is found in y_{-n}^0 . This missing information leaks out to the coordinates to the right of 0 inducing the flow $J_L^+(1)$. Technically, the boundary conditions are obtained by introducing a term γ_n which vanishes in the limit, see (7).

By an argument similar to that used for reversible CA, the existence of a continuity equation can also be extended to a larger class of CA than the reversible CA. The crux of the argument used to find a continuity equation for reversible CA is that all preimages of y_{-n}^0 agree on a central block. This is also the case for those surjective CA where almost all $y \in \mathcal{A}^{\mathbb{Z}}$ have exactly one preimage. These are the CA with $M(F) = 1$ in the notation of Hedlund [11]. In this case $J_L^-(i)$ has the same expression, but in $J_L^+(i)$ R will be changed to the stopping time τ , which is the smallest integer such that all preimages of $y_{-\infty}^{i-1}$ agree to the left of coordinate $i - \tau$. We will present the details of this and other cases as well as examples in [12].

4 Conclusions

We have studied local information quantities S_L and S_R that measure the local information density of a one-dimensional lattice system. These are useful for detecting and quantifying structure in a lattice configuration.

An information flow J_L has been introduced such that a continuity equation $\Delta_i S_L + \Delta_i J_L = 0$ holds under iteration of a one-dimensional reversible CA. We have discussed how the CA generates two different types J_L^+ and J_L^- of information flow satisfying $J_L = J_L^+ - J_L^-$. Moreover, we have sketched how an expression for J_L may be found in

other cases by choosing boundary conditions in a suitable way.

The continuity equation is a fundamental property of information transport in reversible systems. But we expect local information in cellular automata to have further interesting properties. In particular, a continuity equation in physics can be viewed as a constraint, rather than an equation that determines the dynamics of the system. In a similar way, one may expect information flow to have different dynamic characteristics in different CA. We have seen examples of CA building up high information in local configurations. Other CA might show diffusive behaviour and smear out initial inhomogeneities in local information. Such properties would give much information about the dynamical behaviour of the CA.

References

- [1] Bekenstein, J. Black holes and entropy. *Physical Review D*, 7:2333–2346, 1973.
- [2] Cover, T. M. and Thomas, J. A. *Elements of information theory*. Wiley Series in Telecommunications. John Wiley & Sons Inc., New York, 1991.
- [3] Eloranta, K. and Nummelina, E. The kink of cellular automaton Rule 18 performs a random walk. *Journal of Statistical Physics*, 69(5-6):1131–1136, 1992.
- [4] Eriksson, K.-E. and Lindgren, K. Structural information in self-organizing systems. *Physica Scripta*, 35:388–397, 1987.
- [5] Eriksson, K.-E., Lindgren, K., and Månsson, B. Å. *Structure, Context, Complexity, and Organization*. World Scientific, Singapore, 1987.
- [6] Eriksson, K.-E., Lindgren, K., and Nordahl, M. G. Continuity of information flow in discrete reversible systems. Chalmers preprint, 1993.
- [7] Frisch, U., Hasslacher, B., and Pomeau, Y. Lattice-gas automata for the Navier-Stokes equation. *Physical Review Letters*, 56:1505–1508, 1986.
- [8] Grassberger, P. Chaos and diffusion in deterministic cellular automata. *Physica D*, 10:52–58, 1984.
- [9] Grassberger, P. Toward a quantitative theory of self-generated complexity. *International Journal of Theoretical Physics*, 25(9):907–938, 1986.
- [10] Hanson, J. E. and Crutchfield, J. P. Computational mechanics of cellular automata: an example. *Physica D*, 103:169–189, 1997.
- [11] Hedlund, G. A. Endomorphisms and automorphisms of the shift dynamical system. *Mathematical Systems Theory*, 3:320–375, 1969.
- [12] Helvik, T., Lindgren, K., and Nordahl, M. G. Local information in discrete lattice systems. Forthcoming.
- [13] Jaynes, E. T. Information theory and statistical mechanics. *Phys. Rev. (2)*, 106:620–630, 1957.

- [14] Kaneko, K. Lyapunov analysis and information flow in coupled map lattices. *Physica D*, 23:436–447, 1986.
- [15] Khinchin, A. I. *Mathematical Foundations of Information Theory*. Dover Publications Inc., New York, N.Y., 1957.
- [16] Lindgren, K. Correlations and random information in cellular automata. *Complex Systems*, 1:529–543, 1987.
- [17] Lindgren, K., Eriksson, A., and Eriksson, K.-E. Flows of information in spatially extended chemical dynamics. To appear in Proceedings of ALife 9, 2004.
- [18] Lindgren, K. and Nordahl, M. G. Complexity measures and cellular automata. *Complex Systems*, 2:409–440, 1988.
- [19] Lloyd, S. Computational capacity of the universe. *Physical Review Letters*, 88(23), 2002. 237901.
- [20] Richardson, D. Tessellations with local transformations. *Journal of Computer and Systems Sciences*, 5:373–388, 1972.
- [21] Schreiber, T. Spatio-temporal structure in coupled map lattices: two-point correlations versus mutual information. *Journal of Physics. A.*, 23(8):L393–L398, 1990.
- [22] Schreiber, T. Measuring information transfer. *Physical Review Letters*, 85:461–464, 2000.
- [23] Shannon, C. E. A mathematical theory of communication. *Bell System Tech. J.*, 27:379–423, 623–656, 1948.
- [24] Takesue, S. Reversible cellular automata and statistical mechanics. *Physical Review Letters*, 59:2499–2502, 1987.
- [25] Toffoli, T. Information transport obeying the continuity equation. *IBM J. Res. Develop.*, 32(1):29–36, 1988.
- [26] Toffoli, T. and Margolus, N. H. Invertible cellular automata: a review. *Physica D*, 45(1-3):229–253, 1990.
- [27] Vastano, J. A. and Swinney, H. L. Information transport in spatiotemporal systems. *Physical Review Letters*, 60(18):1773–1776, 1988.
- [28] Vichniac, G. Simulating physics with cellular automata. *Physica D*, 10:96–115, 1984.
- [29] Wheeler, J. A. *Information, Physics, Quantum: The Search for Links, in Complexity, Entropy and the Physics of Information*. Addison-Wesley, Redwood City, CA, 1988.
- [30] Wolfram, S. Universality and complexity in cellular automata. *Physica D*, 10:1–35, 1984.
- [31] Wolfram, S., editor. *Theory and Application of Cellular Automata*. World Scientific, 1986.

Paper IV

Continuity of Information Transport in Surjective Cellular Automata

Torbjørn Helvik, Kristian Lindgren and Mats G. Nordahl (2005)

Submitted for publication.

Continuity of Information Transport in Surjective Cellular Automata

Torbjørn Helvik¹, Kristian Lindgren², and Mats G. Nordahl³

¹Department of Mathematical Sciences
Norwegian University of Science and Technology
NO-7491 Trondheim, Norway

²Department of Physical Resource Theory
Chalmers University of Technology and Göteborg University
SE-41296 Göteborg, Sweden

³Department of Applied Information Technology
Chalmers University of Technology and Göteborg University
SE-41756 Göteborg, Sweden

Abstract

We introduce a local version of the Shannon entropy in order to describe information transport in spatially extended dynamical systems, and to explore to what extent information can be viewed as a local quantity. Using an appropriately defined information current, this quantity is shown to obey a local conservation law in the case of one-dimensional reversible cellular automata with arbitrary initial measures. The result is also shown to apply to one-dimensional surjective cellular automata in the case of shift-invariant measures. Bounds on the information flow are also shown.

1 Introduction

A number of authors have suggested that information should be viewed as a fundamental physical quantity, starting with the vision of “It from Bit” of Wheeler [28] and the fundamental work on the thermodynamics of computation by Landauer [14] and Bennett [1]. More recently, quantum computation and quantum information have become major topics of investigation, and the issue of whether information is conserved in black holes has been a topic of considerable debate.

Information theory also has a close relation to the foundations of statistical mechanics. One example of this is the information theoretic formulation of statistical

mechanics introduced by Jaynes [11], where entropy is viewed as a measure of the ignorance of the actual microstate of the system. Other authors have considered how information theory and computation theory can be used to define an entropy for individual microstates in spatially extended systems [17, 30].

In a microscopic view, information or entropy quantified in terms of the Gibbs H-function is a conserved quantity due to Liouville's theorem. This statement is a global conservation law. A natural question to consider is to what extent this statement has a local analogue in spatially extended dynamical systems. This article explores this question in the context of one-dimensional reversible or surjective cellular automata. Precise statements of the notion of conservation of information, and possible extensions of this formalism to other systems, could provide a more solid foundation for the use of information based concepts in different physical systems.

To be able to consider information as a local quantity we first introduce a local version of the Shannon entropy. In a one-dimensional system, the local information is defined in terms of the conditional probability of a local state given its left or right infinite context (in one dimension, these provide two separate locally conserved quantities). Information can only be completely localized in a system without correlations, and any measure of information needs to take correlations into account. The measure we introduce is localized to the extent that correlations allow, and reduces to a completely local quantity when correlations vanish. However, even with correlations present, this quantity does obey a local continuity equation with an appropriately defined information current.

In the context of cellular automata, local conservation of information was first proposed by Toffoli [24], who derived a continuity equation for information transport in the case of small perturbations around the uncorrelated equilibrium states of particle conserving reversible cellular automata, such as lattice gases. Here we investigate how these concepts can be applied to more general classes of dynamical systems and to arbitrary measures, and how they can be given a rigorous formulation. In this article we only consider one-dimensional systems; systems in higher dimensions will be addressed in future work.

We first consider reversible cellular automata, i.e., cellular automata where the cellular automaton mapping has an inverse, so that all infinite configurations have exactly one preimage. Reversible cellular automata have been used as algorithms for simulating physical systems, e.g, for microcanonical simulations of spin systems (e.g., [26]), simulations of fluid dynamics using lattice gas automata [4, 8], and simulations of chemical reactions [2]. They have also been studied as simple dynamical systems in their own right, in particular in order to provide illustrative examples of fundamental issues in statistical mechanics [22, 23]. For one-dimensional reversible cellular automata, we show local conservation of information for any initial measure, including measures without shift-invariance.

We also consider the more complicated case of surjective cellular automata, where the global mapping is finite-to-one [9]. In this case, local conservation of information

flow is shown for all shift-invariant measures. A particularly simple class of surjective cellular automata are the permutative rules, and for these we show that the information flow always has the simple form of information being shifted to the right or to the left.

The aim of the article is to explore exactly to what extent information can be viewed as a local quantity in spatially extended systems. The main results show that important aspects of locality remain also in systems with correlations. We also give examples which illustrate the limits of locality in the formalism.

The rest of this article is organized as follows. Section 2 contains background material on shift spaces and cellular automata. In Sect. 3 we introduce a local measure of information and show that it is well-defined. Section 4 contains the main results of the paper. We first define the information current, and prove that information is locally conserved for one-dimensional reversible cellular automata for arbitrary initial measures. We then extend this result to surjective cellular automata in the case of shift-invariant measures. In Sect. 5, we give an information theoretic interpretation of the current, and provide bounds on the information flow. We also characterize the information flow in permutative cellular automata, and study some examples that illustrate the limits of locality. Section 6 contains conclusions and a discussion.

2 Preliminaries

2.1 The shift space

In this paper we study dynamical systems on the space $\mathcal{A}^{\mathbb{Z}}$ of all bi-infinite symbol sequences over a finite set \mathcal{A} . For $x \in \mathcal{A}^{\mathbb{Z}}$ we write $x = (x_i)_{i \in \mathbb{Z}}$. The length $j - i + 1$ block $(x_i, x_{i+1}, \dots, x_j)$ of symbols from \mathcal{A} will be written compactly as x_i^j . Likewise, $x_{-\infty}^i = (\dots, x_{i-1}, x_i)$. The *shift map* σ is defined on $\mathcal{A}^{\mathbb{Z}}$ by $\sigma(x)_i = x_{i+1}$.

The set \mathcal{A} is equipped with the discrete topology and $\mathcal{A}^{\mathbb{Z}}$ with the corresponding product topology, making $\mathcal{A}^{\mathbb{Z}}$ a compact metric space. The topology is generated by the collection of all *cylinder sets* of the form $\text{Cyl}(a_i^{i+n}) = \{x \in \mathcal{A}^{\mathbb{Z}} : x_i^{i+n} = a_i^{i+n}\}$, where $i \in \mathbb{Z}$, $n \geq 0$ and all $a_k \in \mathcal{A}$. The collection of cylinder sets also forms a sub-algebra that generates the Borel σ -algebra \mathcal{B} of $\mathcal{A}^{\mathbb{Z}}$.

A probability *measure* μ on $(\mathcal{A}^{\mathbb{Z}}, \mathcal{B})$ is defined by assigning a probability $\mu(\text{Cyl}(a_i^{i+n}))$ to each cylinder set in a consistent way, see [27, §0.2]. We will usually write this probability $\mu(a_i^{i+n})$, thus letting a_i^{i+n} represent both the symbol block of length $n + 1$ and the cylinder set. It is often convenient to consider the measure μ as defining a discrete, stochastic process $(X_n)_{n=-\infty}^{\infty}$, $X_n \in \mathcal{A}$, with joint distributions given by $\text{Prob}(X_i^j = a_i^j) = \mu(a_i^j)$. In this case, μ is called the Kolmogorov measure of the process. A Bernoulli measure is a measure for which the coordinate random variables X_i are all independent and identically distributed. The conditional probability $\mu(a_0 | a_{-n}^{-1}) = \frac{\mu(a_{-n}^0)}{\mu(a_{-n}^{-1})}$ is the probability that $X_0 = a_0$ given that $X_{-n}^{-1} = a_{-n}^{-1}$. The measure μ is *shift-*

invariant if it satisfies $\mu(\sigma^{-1}(B)) = \mu(B)$ for all measurable subsets $B \subseteq \mathcal{A}^{\mathbb{Z}}$. Equivalently, μ is shift-invariant if $\mu(a_i^{i+n})$ only depends on the symbols a_k and not on the starting position i . When μ is shift-invariant, the expectation $E[f]$ of any measurable function f on $\mathcal{A}^{\mathbb{Z}}$ satisfies $E[f] = E[f \circ \sigma]$.

A shift-invariant measure μ is *ergodic* if for all measurable sets $B \in \mathcal{B}$ satisfying $\sigma^{-1}B = B$ we have $\mu(B) = 0$ or $\mu(B) = 1$. For $x \in \mathcal{A}^{\mathbb{Z}}$, define the empirical measure ν_x generated by x as the measure having block probabilities

$$\nu_x(a_1^n) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=0}^{N-1} 1_{\text{Cyl}(a_1^n)}(\sigma^{-k}x) , \quad (1)$$

where 1_B is the characteristic function of the set B . For any shift-invariant measure μ , the limit exists μ -a.e. for each block a_1^n and ν_x is ergodic μ -a.e. Every shift-invariant measure μ can be decomposed into a generalized convex combination of ergodic measures, in the sense that for any $B \in \mathcal{B}$,

$$\mu(B) = \int \nu_x(B) d\mu(x) . \quad (2)$$

2.2 Cellular automata

One-dimensional cellular automata (CA) are discrete dynamical systems on $\mathcal{A}^{\mathbb{Z}}$ that commute with the shift σ .

Definition 1. A cellular automaton $F : \mathcal{A}^{\mathbb{Z}} \rightarrow \mathcal{A}^{\mathbb{Z}}$ is a dynamical system that can be defined by non-negative integers l, r and a map $f : \mathcal{A}^{l+r+1} \rightarrow \mathcal{A}$, such that

$$(Fx)_i = f(x_{i-l}, x_{i-l+1}, \dots, x_{i+r}) \quad \forall i \in \mathbb{Z} . \quad (3)$$

By the left and right *radii* of F we mean the smallest such integers l and r for which there is a block map f that generates F . The block map is often called the CA rule. It is easy to see that F is continuous and shift-commuting (that is, $\sigma \circ F = F \circ \sigma$). Conversely, any continuous and shift-commuting map from $\mathcal{A}^{\mathbb{Z}}$ to $\mathcal{A}^{\mathbb{Z}}$ is a cellular automaton [9, Thm. 3.1 and 3.4].

Example 1. Let $\mathcal{A} = \{0, 1\}$, and denote by F_1 the simple CA on $\mathcal{A}^{\mathbb{Z}}$ defined by the radii $l = 0$ and $r = 1$ and the block map $f : \mathcal{A}^2 \rightarrow \mathcal{A}$ given by $f(x_0, x_1) = x_0 + x_1 \pmod{2}$. The global map F_1 can be written as $F_1(x) = x + \sigma(x)$, where addition is coordinate-wise and modulo 2.

For any $n \geq 1$ the block map f can be extended in a natural way to a map $f_n : \mathcal{A}^{l+r+n+1} \rightarrow \mathcal{A}^{n+1}$ by putting

$$f_n(x_{-l}^{r+n}) = (f(x_{-l}^r), f(x_{-l+1}^{r+1}), \dots, f(x_{n-l}^{r+n})) . \quad (4)$$

We will omit the subscript n and write f for the block map applied to a block of any length. We also write $f(x_{-\infty}^i)$ for f applied to the semi-infinite sequence $x_{-\infty}^i$.

A particular class of CA are the *reversible* ones. These have the property that F has an inverse map, so that each bi-infinite sequence $y \in \mathcal{A}^{\mathbb{Z}}$ has exactly one preimage under F (that is, for each $y \in \mathcal{A}^{\mathbb{Z}}$ there is a unique $x \in \mathcal{A}^{\mathbb{Z}}$ that satisfies $F(x) = y$). The inverse map of a reversible CA is always itself a CA, as was first pointed out by Richardson [20]. The inverse CA does not necessarily have the same radii as F . See [25] for a review of reversible CA.

Example 2. Denote by F_2 the reversible CA on $\{0, 1, 2\}^{\mathbb{Z}}$ having radii $l = 0, r = 1$ and block map given by $f(10) = f(11) = f(12) = 0$, $f(01) = f(20) = f(22) = 1$ and $f(00) = f(02) = f(21) = 2$. The preimage x of a given $y \in \mathcal{A}^{\mathbb{Z}}$ is found by the following procedure. If $y_i = 0$ we must have $x_i = 1$. If $y_i = 1$ then $x_i = 2$ unless $y_{i+1} = 0$, in which case $x_i = 0$. Finally, if $y_i = 2$ then $x_i = 0$ unless $y_{i+1} = 0$, in which case $x_i = 2$. The inverse CA \tilde{F}_2 also has $l = 0$ and $r = 1$, but a different block map \tilde{f} .

A more general class of CA are the *surjective* ones. That is, CA F such that all $y \in \mathcal{A}^{\mathbb{Z}}$ have at least one preimage. There are interesting examples of surjective rules among one-dimensional CA with $|\mathcal{A}| = 2$ and radii 1, but only trivial examples of reversible rules. The class of surjective CA includes the much studied *linear*, or *additive*, CA, which are CA that can be written as polynomials of shift maps with addition modulo $|\mathcal{A}|$, e.g. [10]. The CA F_1 from Example 1 is linear. Below we recall some special properties of surjective CA. Call $x \in \mathcal{A}^{\mathbb{Z}}$ bi-transitive if all finite blocks occur an infinite number of times in x on both sides of coordinate 0. The following result is well known and was initially proved in [9].

Lemma 1. A CA F is surjective if and only if all finite blocks have the same number of pre-images under f (that is, for all $n \geq 1$ and $y_1^n \in \mathcal{A}^n$ there are exactly $|\mathcal{A}|^{l+r}$ blocks $z_{1-l}^{n+r} \in \mathcal{A}^{l+r+n}$ that satisfy $f(z_{1-l}^{n+r}) = y_1^n$). Furthermore, there is a constant $M(F) \leq |\mathcal{A}|^{l+r}$ such that each bi-transitive $x \in \mathcal{A}^{\mathbb{Z}}$ has exactly $M(F)$ preimages.

Notice that while each finite block always has $|\mathcal{A}|^{l+r}$ preimages, bi-infinite sequences can have less. Having $M(F) = 1$ is a necessary but not sufficient condition for reversibility.

Example 3. Denote by F_3 the surjective CA on $\{0, 1\}^{\mathbb{Z}}$ with radii $l = 2, r = 0$ and block map $f(x_{-2}, x_{-1}, x_0) = x_0 + x_{-1}x_{-2} \pmod{2}$. It has $M(F_3) = 1$ but is not reversible as, e.g., the sequence x with $x_i = 0$ for all i has two preimages.

For surjective CA one can define some useful coefficients known as *Welch coefficients*. Let $x_1^n \in \mathcal{A}^n$ with $n \geq l + r$. A compatible right extension of x_1^n of length m is a collection $B \subset \mathcal{A}^m$ such that for each $z_1^m \in B$, the $(n + m - l - r)$ -block $f(x_1^n z_1^m)$ is the same. Define the integer $R(F)$ as the maximal number of elements in any compatible right extension of any length m and of any block x_1^n of length $n \geq l + r$. Define

compatible left extensions and $L(F)$ in the same way. The coefficients $L(F)$ and $R(F)$ are finite, and there is a simple relation between $M(F)$, $R(F)$ and $L(F)$, namely that $L(F) \cdot M(F) \cdot R(F) = |\mathcal{A}|^{l+r}$ [9, Th. 14.9]. For the CA from the examples, $R(F_1) = R(F_3) = 1$ while $R(F_2) = 3$ and $L(F_1) = L(F_2) = 1$ while $L(F_3) = 4$.

A surjective CA F is called right *permutative* if $R(F) = 1$, left permutative if $L(F) = 1$ and bipermutative if both these coefficients are one. The name permutative was first coined by Hedlund [9], and originates in the fact that F is right permutative if and only if the map $a \mapsto f(x_{-l}^{r-1}a)$ is a permutation of \mathcal{A} for each $x_{-l}^{r-1} \in \mathcal{A}^{l+r}$. The class includes all linear CA over $\mathcal{A}^{\mathbb{Z}}$ when \mathcal{A} has prime cardinality, as well as more complex CA such as rule 30 (using Wolfram's rule numbering system [29]).

3 Local Information

In this paper, we use the word information to mean information in the sense of Shannon [21]. The Shannon information gained by observing an event with probability p is defined as $\log \frac{1}{p}$, where the logarithm is taken with base 2. The *entropy* of a discrete random variable g with n possible values is the expected information gained by observing its outcome. We write $H(g) = -\sum_{i=1}^n p_i \log p_i$. Information and entropy are measured in *bits*. A bit is the amount of information that can be stored in a single binary variable. The entropy of a random variable is thus the average number of binary variables needed to store its outcome, using an optimal code.

Associated to a shift-invariant measure μ on $\mathcal{A}^{\mathbb{Z}}$ are the block entropies $H_n(\mu)$, $n \in \mathbb{N}$, which are the entropies of the probability distributions $\mu(a_{-n}^n)$ on \mathcal{A}^n . The entropy of μ is defined by $h(\mu) = \lim_{n \rightarrow \infty} \frac{H_n(\mu)}{n}$. The natural interpretation of h is as the average information per symbol when all correlations are taken into account. That is, the average information gained by observing a new symbol when reading the sequence in some direction and remembering all past symbols. Additionally, h is also the Kolmogorov-Sinai entropy of the measure-preserving dynamical system $(\mathcal{A}^{\mathbb{Z}}, \mathcal{B}, \mu, \sigma)$. It is well known that h also can be expressed as a limit of conditional entropies. The conditional entropy $H(g|f)$ is the average information gained by observing the outcome of g when the outcome of f is known. The entropy $h(\mu)$ can be written as

$$h(\mu) = \lim_{n \rightarrow \infty} H(X_0 | X_{-n}^{-1}) = - \lim_{n \rightarrow \infty} \sum_{a_{-n}^0 \in \mathcal{A}^{n+1}} \mu(a_{-n}^0) \log \mu(a_0 | a_{-n}^{-1}) . \quad (5)$$

This forms the basis for our definition of local information.

The intent of introducing a local information quantity is to measure exactly how much information that is located at each position of an infinite symbol sequence generated by some stochastic process. However, the correlations in such symbol sequences can in general be arbitrarily long. Consequently, it is impossible for information to be completely localized. That is, the local information at a particular position

i cannot always be computed merely by looking at the symbol at i , nor at the configuration in a finite neighbourhood of i . The natural approach is therefore to define the local information as a limit which converges to a local analogue of the Shannon entropy as more and more distant neighbours are taken into account, and which can be computed locally when the correlations are finite in extent. While the Shannon entropy is limited to shift-invariant measures, we can define left local information for any measure.

Definition 2. Let μ be a measure on $\mathcal{A}^{\mathbb{Z}}$. The left local information at coordinate i of $x \in \mathcal{A}^{\mathbb{Z}}$ with respect to μ is given by

$$S_L(x; i; \mu) = - \lim_{n \rightarrow \infty} \log \mu(x_i | x_{i-n}^{i-1}) . \quad (6)$$

The following theorem ensures that the left local information with respect to μ is a well-defined function on the probability space $(\mathcal{A}^{\mathbb{Z}}, \mathcal{B}, \mu)$.

Theorem 1. For each $i \in \mathbb{Z}$, $-\log \mu(x_i | x_{i-n}^{i-1})$ converges μ -almost everywhere and in $L^1(\mu)$. Consequently, for each fixed measure μ and $i \in \mathbb{Z}$,

$$S_L(x; i; \mu) \in L^1(\mu) . \quad (7)$$

We will often use the intuitive notation $-\log \mu(x_i | x_{-\infty}^{i-1})$ for $S_L(x; i; \mu)$. When discussing local information as a function on the probability space $(\mathcal{A}^{\mathbb{Z}}, \mathcal{B}, \mu)$ we always mean local information with respect to μ . The theorem ensures that S_L has finite expectation. Indeed, from L^1 convergence and (5) it follows that $E[S_L(x; i; \mu)] = h(\mu)$ for all i in the shift-invariant case. Using local information we can also define local or coordinate-wise left entropies for all measures as $h_i(\mu) = E[S_L(x; i; \mu)]$, $i \in \mathbb{Z}$.

The quantity $-\log \mu(x_i | x_{i-n}^{i-1})$ is the information gained from the symbol at position i when only knowledge of the n left symbols is assumed. If, and only if, μ is Markov there is a fixed n such that $S_L(x; i; \mu) = -\log \mu(x_i | x_{i-n}^{i-1})$. In other cases, correlations are not finite and no such n will exist. The speed of convergence depends on the typical length of correlations in the system, which could be quantified as in [6].

The validity of Theorem 1 follows directly from a more general result in probability theory. We first introduce the necessary background. This theory will also be used in the proofs of our main results in Sect. 4. Let (X, \mathcal{B}, μ) be a probability space. The function $E[f | \mathcal{F}]$, with \mathcal{F} a sub- σ -algebra of \mathcal{B} , is defined up to a set of measure zero by being measurable with respect to \mathcal{F} and having the same expectation as f over any $B \in \mathcal{F}$. The function is called the *conditional expectation* of f given \mathcal{F} . Of particular interest to us is $E[1_A | \mathcal{F}]$, with $A \in \mathcal{B}$. This function is commonly written as $\mu(A | \mathcal{F})$ and called the conditional probability of A . If \mathcal{F} is generated by the finite partition β of X , then $\mu(A | \mathcal{F})(x) = \sum_{B \in \beta} \frac{\mu(A \cap B)}{\mu(B)} 1_B(x)$. The function $\mu(A | \mathcal{F})$ satisfies $\mu(A | \mathcal{F}) > 0$ μ -a.e. on A [12, Lemma 3.1.2]. For α a partition of X , the *conditional information* of α given \mathcal{F} is defined by

$$I_{\alpha | \mathcal{F}}(x) = - \sum_{A \in \alpha} \log \mu(A | \mathcal{F})(x) \cdot 1_A(x) . \quad (8)$$

For two partitions α and β of X , the conditional information satisfies $I_{\alpha \vee \beta | \mathcal{F}} = I_{\alpha | \mathcal{F}} + I_{\beta | \alpha \vee \mathcal{F}}$.

A *filtration* of a probability space is an increasing family of sub- σ -algebras $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \dots \subset \mathcal{B}$. Let \mathcal{F}_∞ be the smallest σ -algebra containing all \mathcal{F}_n , $\mathcal{F}_\infty = \sigma(\cup_n \mathcal{F}_n)$. A *supermartingale* relative to $\{\mathcal{F}_n\}$ is a stochastic process $(g_n : n \geq 0)$ such that (1) g_n is \mathcal{F}_n -measurable, (2) $g_n \in L^1(\mu) \forall n$ and (3) $E[g_{n+1} | \mathcal{F}_n] \leq g_n$. Condition (3) can be checked by assuring that for each $B \in \mathcal{F}_n$,

$$\int_B g_{n+1} d\mu \leq \int_B g_n d\mu . \quad (9)$$

The martingale convergence theorem states that any non-negative supermartingale g_n converges a.e. to a function g which is \mathcal{F}_∞ -measurable and finite a.e. Based on the martingale theorem, the following result is proved e.g. in [12].

Lemma 2. *Let $\{\mathcal{F}_n\}$ be a filtration of (X, \mathcal{B}, μ) . For any finite partition α*

$$\lim_{n \rightarrow \infty} I_{\alpha | \mathcal{F}_n} = I_{\alpha | \mathcal{F}_\infty} \quad \mu\text{-a.e. and in } L^1(\mu) . \quad (10)$$

We now proceed to prove Theorem 1 using this lemma.

Proof of Theorem 1. Fix an $i \in \mathbb{Z}$. Define \mathcal{F}_n as the σ -algebra generated by the coordinate random variables from $i - n$ to $i - 1$: $\mathcal{F}_n = \sigma(X_j : i - n \leq j \leq i - 1)$, and define α as the partition $\{\text{Cyl}(a_i) : a_i \in \mathcal{A}\}$. Then

$$-\log \mu(x_i | x_{i-n}^{i-1}) = I_{\alpha | \mathcal{F}_n}(x) . \quad (11)$$

Convergence a.e. of $-\log \mu(x_i | x_{i-1}^{i-n})$ to $I_{\alpha | \mathcal{F}_\infty}$ follows directly from Lemma 2. Note that $\mathcal{F}_\infty = \sigma(X_i : i \leq 0)$. \square

The local information $S_L(x; i; \mu)$ depends on the measure μ , thus knowledge of μ is required to calculate S_L . However, in the case of a shift-invariant measure μ the left local information at position i of x can in principle be recovered with probability one from $x_{-\infty}^i$ only. This is achieved by considering the empirical measure ν_x obtained from the frequencies of finite blocks in $x_{-\infty}^i$. If μ is ergodic, then $\nu_x = \mu$ a.e. However, even when μ is only shift-invariant it suffices to look at the local information with respect to ν_x .

Theorem 2. *Let μ be a shift-invariant measure on $\mathcal{A}^{\mathbb{Z}}$ and ν_x the empirical measure generated by $x_{-\infty}^i$. Then*

$$S_L(x; i; \mu) = S_L(x; i; \nu_x) \quad \mu\text{-a.e.}$$

Proof. The result is true since the infinite history determines with probability one which ergodic component of μ x is generated by. More precisely, Lemma 8.6.2. in [7] states that for a $B \in \sigma(X_i : i \geq 0)$,

$$\mu((X_0, X_1, \dots) \in B | \sigma(X_{-\infty}^{-1})) = \nu_x((X_0, X_1, \dots) \in B | \sigma(X_{-\infty}^{-1})) \quad \mu\text{-a.e.} \quad (12)$$

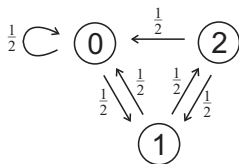


Fig. 1: An automaton that defines a Markov measure μ on $\{0, 1, 2\}^{\mathbb{Z}}$ with the property that $S_L(x; i; \mu)$ and $S_R(x; i; \mu)$ have different distributions.

Recall that the local information at position 0 is a conditional information $I_{\alpha|\mathcal{F}}$ on the form (8) with $\alpha = \{\text{Cyl}(a_0) : a_0 \in \mathcal{A}\}$ and $\mathcal{F} = \sigma(X_{-\infty}^{-1})$. Therefore, the theorem follows for the case $i = 0$ from (12). This generalizes directly to any arbitrary coordinate i . \square

The choice of conditioning on the left history of coordinate i when defining local information is arbitrary. We can also define the *right local information* at coordinate i of x with respect to μ as

$$S_R(x; i; \mu) = - \lim_{n \rightarrow \infty} \log \mu(x_i | x_{i+1}^{i+n}) . \quad (13)$$

The right local information obviously has the same convergence properties as the left information, and all results we show for the left information will have corresponding results for the right information. Note, however, that although the left and right information have the same expectation for all shift-invariant measures, they are not equal nor do they in general have the same probability distribution. This is exemplified by the Markov measure on $\{0, 1, 2\}^{\mathbb{Z}}$ defined by the automaton in Fig. 1. We will see in the next section that both the left and the right information are fundamental locally conserved quantities under iteration by a one-dimensional surjective CA. One could also take the average of left and right information and the difference as the fundamental conserved quantities.

4 Information Transport

In this section we investigate the transport of local information in a one-dimensional system generated during the time-evolution of a surjective cellular automaton. We show that the local information satisfies a continuity equation involving an information current J_L , and supply an expression for this current. Recall that a CA F is a dynamical system on $\mathcal{A}^{\mathbb{Z}}$ that maps each $x \in \mathcal{A}^{\mathbb{Z}}$ to a $y \in \mathcal{A}^{\mathbb{Z}}$ by simultaneously updating the symbol at each position by a local block map f . Additionally, F can be considered to act on measures. Let μ^0 be a measure on $\mathcal{A}^{\mathbb{Z}}$. The measure $F(\mu^0) = \mu^0 \circ F^{-1}$ gives the joint distributions of the stochastic process $(Y_i)_{i \in \mathbb{Z}}$ with $Y_i = f(X_{i-l}^{i+r})$ when

$(X_i)_{i \in \mathbb{Z}}$ has joint distributions given by μ^0 . Denote $\mu^0 \circ F^{-1}$ by μ^1 and, more generally, set $\mu^t = \mu^0 \circ F^{-t}$. The block probabilities of μ^1 can be calculated from

$$\mu^1(y_0^n) = \sum_{z_{-l}^{n+r} \in f^{-1}(y_0^n)} \mu^0(z_{-l}^{n+r}) . \quad (14)$$

It is well known that $h(\mu^1) = h(\mu^0)$ whenever F is surjective and μ^0 is shift-invariant (if F is non-surjective, this relation is replaced by $h(\mu^1) \leq h(\mu^0)$). An elementary proof is found in [16]. Our goal is to prove the much stronger result that the local information in fact obeys a local continuity equation under the time evolution of the CA. This equation is an analogue of the continuity equations which apply to locally conserved physical quantities such as charge or particle density. The existence of a continuity equation for information flow in reversible systems was first proposed by Toffoli [24], though it was shown only in the case of small perturbations around uncorrelated equilibrium states of particle conserving reversible cellular automata, such as lattice gas automata.

A continuity equation for the local information is an equation of the form

$$\Delta_t S_L + \Delta_i J_L = 0 , \quad (15)$$

where $J_L(x; i; \mu)$ is the information current. The operator Δ is the forward difference operator, so that

$$\begin{aligned} \Delta_t S_L(x; i; \mu^t) &= S_L(F(x); i; \mu^{t+1}) - S_L(x; i; \mu^t) , \\ \Delta_i J_L(x; i; \mu^t) &= J_L(x; i+1; \mu^t) - J_L(x; i; \mu^t) . \end{aligned} \quad (16)$$

With these definitions, $J_L(x; i; \mu^t)$ can be interpreted as the information flow from position $i-1$ to position i generated by applying the CA. Note that the local information of $F(x)$ is taken with respect to a different measure than x , unless μ^t happens to be invariant for the CA.

Before we present the formula for J_L and the main results we need two definitions. These are illustrated in Fig. 2. For a semi-infinite sequence $x_{-\infty}^i$, define $Z(x_{-\infty}^i)$ as the set of all semi-infinite sequences that have the same image and the same tail as $x_{-\infty}^i$:

Definition 3. For $x \in \mathcal{A}^{\mathbb{Z}}$ and a surjective CA F , define the sets $Z(x_{-\infty}^i)$ as

$$Z(x_{-\infty}^i) = \{z_{-\infty}^i : f(z_{-\infty}^i) = f(x_{-\infty}^i) \text{ and } \exists j \leq i \text{ such that } z_{-\infty}^j = x_{-\infty}^j\} .$$

Note that $|Z(x_{-\infty}^i)| \leq R(F)$ for all x by the definition of the Welch coefficient $R(F)$. Define $\tau(x_{-\infty}^i)$ as the largest index less than $i-r$ for which all sequences in $Z(x_{-\infty}^i)$ coincide (recall that r is the right radius of F):

Definition 4. For $x \in \mathcal{A}^{\mathbb{Z}}$, define $\tau(x_{-\infty}^i) \in \mathbb{Z}$ as

$$\tau(x_{-\infty}^i) = \max_j \{j : j < i-r, \text{ and } z_{-\infty}^j = x_{-\infty}^j \forall z_{-\infty}^i \in Z(x_{-\infty}^i)\} .$$

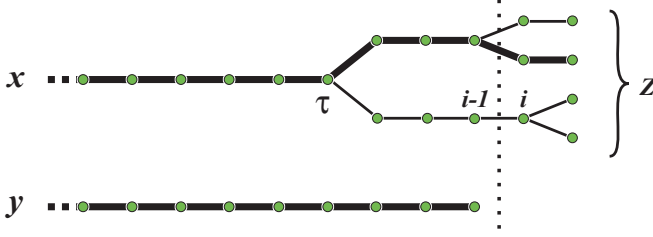


Fig. 2: An illustration of $Z = Z(x_{-\infty}^{i+r-1})$ and $\tau = \tau(x_{-\infty}^{i+r-1})$. In this case $r = 2$, $|Z| = 4$ and $\tau = i - 4$. The circles represent symbols in \mathcal{A} . These are connected by lines to semi-infinite sequences which all map to the same $y_{-\infty}^{i-1}$ and coincide with x at all $j \leq \tau$.

We are now ready to define the information current.

Definition 5. Let F be a surjective one-dimensional CA with right radius r , and μ a measure on $\mathcal{A}^{\mathbb{Z}}$. Put $Z = Z(x_{-\infty}^{i+r-1})$ and $\tau = \tau(x_{-\infty}^{i+r-1})$. Define the left information current at coordinate i of x with respect to μ and F as

$$J_L(x; i; \mu) = -\log \mu(x_{\tau+1}^{i-1} | x_{-\infty}^{\tau}) + \log \sum_Z \mu(z_{\tau+1}^{i+r-1} | x_{-\infty}^{\tau}) . \quad (17)$$

The quantities $\mu(z_{\tau+1}^{i-1} | x_{-\infty}^{\tau})$ are defined as $\lim_{n \rightarrow \infty} \mu(z_{\tau+1}^{i-1} | x_{\tau-n}^{\tau})$, in agreement with the definition of local information. Since J_L is constructed entirely from conditional probabilities of this type, an analogue to Theorem 2 yields

$$J_L(x; i; \mu) = J_L(x; i; \nu_x) \quad \mu\text{-a.e.} \quad (18)$$

Note that $\tau < i - 1$ by the requirement that $\tau(x_{-\infty}^i) < i - r$ included in Definition 4. Also note that in the case of a reversible CA, the existence of an inverse CA ensures that τ is bounded. In particular, let \tilde{r} be the right radius of the inverse CA. Then $\tau \geq i - 1 - \tilde{r}$ unless $\tilde{r} = 0$, in which case $\tau = i - 2$. For non-reversible CA, τ is in general unbounded but always finite. Some information theoretic considerations elucidating why (17) is a natural form for the information current is presented in Sect. 5.

It remains to show that $J_L(x; i; \mu)$ is well defined as a function on (X, \mathcal{B}, μ) . Since τ is finite by definition, this will follow from almost everywhere convergence of all conditional probabilities involved in J_L . Such convergence is ascertained by the following lemma.

Lemma 3. For any measure μ on $\mathcal{A}^{\mathbb{Z}}$,

$$\mu(\{x : \lim_{n \rightarrow \infty} \mu(a_{-k}^{-1} | x_{-n}^{-k-1}) \text{ exists for all } k \geq 0 \text{ and all } a_{-k}^{-1} \in \mathcal{A}^k\}) = 1 .$$

Proof. Fix k and a_{-k}^{-1} . Define the measurable functions $f_n = \mu(a_{-k}^{-1} | x_{-k-n}^{-k-1})$. The process $(f_n : n \geq 0)$ is a martingale with respect to the natural filtration $\mathcal{F}_n = \sigma(X_i : -k - n \leq i \leq$

$-k-1$). Furthermore, $E[f_n]$ is bounded, so by the martingale convergence theorem f_n converges a.e. By subadditivity of the measure, the set

$$A_k = \{x : \lim_{n \rightarrow \infty} \mu(a_{-k}^{-1} | x_{-k-n}^{-k-1}) \text{ does not exist for some } a_{-k}^{-1} \in \mathcal{A}^k\} .$$

has measure 0. By countable subadditivity, $\mu(\bigcup_{k \geq 0} A_k) = 0$, and the result follows. \square

It is also the case that $J_L(x; i; \mu) \in L^1(\mu)$. This is stated in Theorem 5 in Sect. 5.3.

We now proceed to present Theorems 3 and 4, which are the main results of the paper. The conclusions of both theorems are identical, namely the validity of the continuity equation $\Delta_t S_L(x; i; \mu) + \Delta_i J_L(x; i; \mu) = 0$, with components defined in (16). The first theorem states that for reversible CA the equation is valid for all initial measures.

Theorem 3. *Let F be a reversible one-dimensional CA, and μ a measure on $\mathcal{A}^{\mathbb{Z}}$. Then $\Delta_t S_L(x; i; \mu) + \Delta_i J_L(x; i; \mu) = 0$ for all $i \in \mathbb{Z}$ μ -a.e.*

In the more complex case of a general surjective CA, the requirement of μ being shift-invariant is necessary to ensure the validity of the continuity equation.

Theorem 4. *Let F be a surjective one-dimensional CA, and μ a shift-invariant measure on $\mathcal{A}^{\mathbb{Z}}$. Then $\Delta_t S_L(x; i; \mu) + \Delta_i J_L(x; i; \mu) = 0$ for all $i \in \mathbb{Z}$ μ -a.e.*

The theorem is true for all shift-invariant measures, so it is not required that μ gives positive probability to all finite blocks. Example 4 in Sect. 5 shows that the continuity equation as defined above can fail to be valid if μ is not shift-invariant and F is surjective without being reversible.

Note that if one of the theorems is valid for a CA F together with an initial measure μ^0 , then the continuity equation will be satisfied at all time steps of the iteration by F .

Proof of Theorem 3. We first show that it is sufficient to prove the theorem in the case of $r = 0$. Here, and in the rest of the proof, we look at the initial measure μ^0 and its image μ^1 .

Assume that Theorem 3 is valid for CA with $r = 0$, and let F have right radius r . There exist a CA G with $r = 0$ such that $F = \sigma^r \circ G$. We have $S_L(Fx; i; \mu^1) = S_L(Gx; i+r; \mu^1)$, since $F(\mu^0) = G(\mu^0) = \mu^1$. Write $\tau_1 = \tau(x_{-\infty}^{i+r-1})$, $\tau_2 = \tau(x_{-\infty}^{i+r})$, $Z_1 = Z(x_{-\infty}^{i+r-1})$ and $Z_2 = Z(x_{-\infty}^{i+r})$. Using the formula for J_L we obtain

$$\begin{aligned} S_L(Gx; i+r; \mu^1) &= -\log \mu^0(x_{i+r} | x_{-\infty}^{i+r-1}) \\ &\quad -\log \mu^0(x_{\tau_1+1}^{i+r-1} | x_{-\infty}^{\tau_1}) + \log \sum_{Z_1} \mu^0(z_{\tau_1+1}^{i+r-1} | x_{-\infty}^{\tau_1}) \\ &\quad + \log \mu^0(x_{\tau_2+1}^{i+r} | x_{-\infty}^{\tau_2}) - \log \sum_{Z_2} \mu^0(z_{\tau_2+1}^{i+r} | x_{-\infty}^{\tau_2}) . \end{aligned} \quad (19)$$

Since $\tau_1 \leq i-2$ and $\tau_2 \leq i-1$ by definition, we can write

$$\log \mu^0(x_{\tau_2+1}^{i+r} | x_{-\infty}^{\tau_2}) = \log \mu^0(x_{\tau_2+1}^i | x_{-\infty}^{\tau_2}) + \log \mu^0(x_{i+1}^{i+r} | x_{-\infty}^i) \quad (20)$$

and

$$\begin{aligned} -\log \mu^0(x_{i+r}|x_{-\infty}^{i+r-1}) - \log \mu^0(x_{\tau_1+1}^{i+r-1}|x_{-\infty}^{\tau_1}) = \\ -\log \mu^0(x_{\tau_1+1}^{i-1}|x_{-\infty}^{\tau_1}) - \log \mu^0(x_i|x_{-\infty}^{i-1}) - \log \mu^0(x_{i+1}^{i+r}|x_{-\infty}^i) . \end{aligned} \quad (21)$$

Substituting (20) and (21) into (19) gives the correct continuity equation for F .

For the rest of the proof we assume that F has right radius $r = 0$, and left radius $l \geq 0$. We look only at coordinate $i = 0$. This leads to no loss of generality. Call the inverse CA \tilde{F} and let \tilde{F} have left radius \tilde{l} and right radius \tilde{r} . Sequences at time $t = 0$ are generally denoted by x or z and sequences at time $t = 1$ by y .

We first define the joint measure ν of two consecutive time steps (considering the CA as a channel with memory, this usage of the term joint measure is the same as in information theory). Let ν be the measure on $(\mathcal{A} \times \mathcal{A})^{\mathbb{Z}}$ defined by the block probabilities

$$\nu(x_i^j, y_i^j) = \mu^0(\{z_{i-l}^j \in \mathcal{A}^{j-i+l+1} | z_i^j = x_i^j \text{ and } f(z_{i-l}^j) = y_i^j\}) . \quad (22)$$

It is easy to show that ν actually is a measure and that ν is shift-invariant if μ^0 is shift-invariant. We will need the following lemma.

Lemma 4. *Let ν be a measure on $\mathcal{A}_1^{\mathbb{Z}} \times \mathcal{A}_2^{\mathbb{Z}}$, where each \mathcal{A}_i is a finite set. Let $(x, y) \in \mathcal{A}_1^{\mathbb{Z}} \times \mathcal{A}_2^{\mathbb{Z}}$. Then there is a $g \in L^1(\nu)$ such that for any $k \in \mathbb{Z}$*

$$\lim_{n \rightarrow \infty} \nu(x_0, y_0 | x_{-n-k}^{-1}, y_{-n}^{-1}) = g \quad \nu\text{-a.e.} \quad (23)$$

Proof. Let $(X_i, Y_i)_{i=-\infty}^{\infty}$ be the stochastic process corresponding to the measure ν . Fix a k . We can write $\nu(x_0, y_0 | x_{-n-k}^{-1}, y_{-n}^{-1}) = I_{\alpha | \mathcal{F}_n^k}$ with $\alpha = \{\text{Cyl}(a_0, b_0) : a_0, b_0 \in \mathcal{A}\}$ and

$$\mathcal{F}_n^k = \sigma(X_i, Y_j : -n-k \leq i \leq -1, -n \leq j \leq -1) . \quad (24)$$

Almost everywhere convergence of $I_{\alpha | \mathcal{F}_n^k}$ to $I_{\alpha | \mathcal{F}_\infty^k}$ follows from Lemma 2. However, \mathcal{F}_∞^k is the same for any k :

$$\mathcal{F}_\infty^k = \sigma\left(\bigcup_n \mathcal{F}_n^k\right) = \sigma(X_i, Y_j : i, j \leq -1) = \mathcal{F}_\infty . \quad (25)$$

The result follows. \square

Let $y = F(x)$. From the definition of local information we have

$$\begin{aligned} S_L(x; 0; \mu^0) &= -\lim_{n \rightarrow \infty} \log \mu^0(x_0 | x_{-n}^{-1}) \\ &= \lim_{n \rightarrow \infty} \left(\log \frac{\nu(x_0, y_0 | x_{-n}^{-1}, y_{-n+l}^{-1})}{\mu^0(x_0 | x_{-n}^{-1})} - \log \nu(x_0, y_0 | x_{-n}^{-1}, y_{-n+l}^{-1}) \right) \\ &= \lim_{n \rightarrow \infty} \log \frac{\nu(y_{-n+l}^0 | x_{-n}^0)}{\nu(y_{-n+l}^{-1} | x_{-n}^{-1})} - \lim_{n \rightarrow \infty} \log \nu(x_0, y_0 | x_{-n}^{-1}, y_{-n+l}^{-1}) \\ &= -\log \nu(x_0, y_0 | x_{-\infty}^{-1}, y_{-\infty}^{-1}) \end{aligned} \quad (26)$$

by virtue of Lemma 4 and the fact that $v(y_{-n+l}^0|x_{-n}^0) = v(y_{-n+l}^{-1}|x_{-n}^{-1}) = 1$ for all $n > l$, since y_{-n+l}^0 in this case is uniquely determined by x_{-n}^0 through the local map f and likewise for y_{-n+l}^{-1} and x_{-n}^{-1} . A similar treatment of $S_L(y; 0; \mu^1)$ yields

$$\begin{aligned} S_L(y; 0; \mu^1) &= \lim_{n \rightarrow \infty} \left(\log \frac{v(x_0, y_0 | x_{-n+l}^{-1}, y_{-n}^{-1})}{\mu^1(y_0 | y_{-n}^{-1})} - \log v(x_0, y_0 | x_{-n+l}^{-1}, y_{-n}^{-1}) \right) \\ &= \lim_{n \rightarrow \infty} \log \frac{v(x_{-n+l}^0 | y_{-n}^0)}{v(x_{-n+l}^{-1} | y_{-n}^{-1})} - \log v(x_0, y_0 | x_{-\infty}^{-1}, y_{-\infty}^{-1}) . \end{aligned} \quad (27)$$

When taking the difference $\Delta_t S_L$ the last term is canceled out, so

$$\Delta_t S_L(x; 0; \mu^0) = \lim_{n \rightarrow \infty} \log v(x_{-n+l}^0 | y_{-n}^0) - \lim_{n \rightarrow \infty} \log v(x_{-n+l}^{-1} | y_{-n}^{-1}) .$$

We claim that

$$- \lim_{n \rightarrow \infty} \log v(x_{-n+l}^{-1} | y_{-n}^{-1}) = J_L(x; 0; \mu^0) . \quad (28)$$

A verification of this will conclude the proof. We show it through a sequence of transformations. Firstly,

$$\log v(x_{-n+l}^{-1} | y_{-n}^{-1}) = \log v(x_{-n+l}^{-\bar{r}-1} | y_{-n}^{-1}) + \log v(x_{-n+l}^{-1} | x_{-n+l}^{-\bar{r}-1}, y_{-n}^{-1}) . \quad (29)$$

The first term on the right hand side is zero, since $x_{-n+l}^{-\bar{r}-1}$ is uniquely determined by y_{-n}^{-1} through the local map \tilde{f} of the inverse CA. For the second term one can prove that

$$\lim_{n \rightarrow \infty} \log v(x_{-n+l}^{-1} | x_{-n+l}^{-\bar{r}-1}, y_{-n}^{-1}) = \lim_{n \rightarrow \infty} \log v(x_{-n+l}^{-1} | x_{-n+l}^{-\bar{r}-1}, y_{-n}^{-1}) \quad (30)$$

by the same technique used in the proof of Lemma 4. Furthermore, for any events A, B and C in a probability space it is true that $v(A|BC) = \frac{v(C|AB)v(A|B)}{v(C|B)}$. Let $A = x_{-n+l}^{-1}$, $B = x_{-n+l}^{-\bar{r}-1}$ and $C = y_{-n}^{-1}$. Then $v(C|AB) = 1$. Thus,

$$\log v(x_{-n+l}^{-1} | x_{-n+l}^{-\bar{r}-1}, y_{-n}^{-1}) = \log \mu^0(x_{-n+l}^{-1} | x_{-n+l}^{-\bar{r}-1}) - \log v(y_{-n}^{-1} | x_{-n+l}^{-\bar{r}-1}) . \quad (31)$$

By the definition of v , the last term can be written as

$$- \log v(y_{-n}^{-1} | x_{-n+l}^{-\bar{r}-1}) = - \log \sum_{Z(x_{-\infty}^{-1})} \mu^0(z_{-n}^{-1} | x_{-n+l}^{-\bar{r}-1}) . \quad (32)$$

Substituting (32) into (31) and taking the limit $n \rightarrow \infty$ we arrive at the equation for $J_L(x; 0; \mu^0)$ presented in Definition 5. □

For general surjective CA, there is no inverse CA and in general several possible preimages. As a consequence, the proof of Theorem 4 requires a different approach.

Proof of Theorem 4. By the same argument as in the proof of Theorem 3 it suffices to prove the theorem for CA with right radius $r = 0$. As before, we only look at coordinate $i = 0$ and consider the initial measure μ^0 and its image μ^1 .

Let $y = F(x)$, and define

$$q(x) = S_L(y; 0; \mu^1) - S_L(x; 0; \mu^0) + J_L(x; 1; \mu^0) - J_L(x; 0; \mu^0) . \quad (33)$$

Our goal is to prove that $q(x) = 0$ μ^0 -a.e. This is equivalent to $E[|q|] = 0$. To prove this we will introduce a sequence q_k of approximations of q which are measurable with respect to finite parts of the history. To this end we make several definitions.

Define the following equivalence relation on \mathcal{A}^{l+n+1} for $n \geq 0$:

$$x_{-l}^n \sim z_{-l}^n \text{ iff } f(x_{-l}^n) = f(z_{-l}^n) \text{ and } x_{-l}^{-1} = z_{-l}^{-1} . \quad (34)$$

That is, two blocks in \mathcal{A}^{l+n+1} are equivalent if they have the same image under f and agree on the first l coordinates. Denote the equivalence class containing z_{-l}^n by $[z_{-l}^n]$. For an $x \in \mathcal{A}^{\mathbb{Z}}$ we will in particular look at the equivalence classes $[x_{-k-l}^{-1}]$ for $k \geq 1$. There is a close relationship between $[x_{-k-l}^{-1}]$ and $Z(x_{-\infty}^{-1})$, namely that for each $k \geq 1$ we have the inclusion

$$Z(x_{-\infty}^{-1}) \supseteq \{x_{-\infty}^{-l-k-1} z_{-l-k}^{-1} : z_{-l-k}^{-1} \in [x_{-l-k}^{-1}]\} . \quad (35)$$

Recall that $\tau(x_{-\infty}^{-1})$ is largest index such that all sequences in $Z(x_{-\infty}^{-1})$ agrees on and to the left of $\tau(x_{-\infty}^{-1})$. Therefore, for all $k \geq -\tau(x_{-\infty}^{-1}) - 1$ equation (35) is an *equality*. This will be essential. Define $\tau^k(x)$ for $k \geq 1$ as the analogue of τ obtained when considering $[x_{-k-l}^{-1}]$ rather than $Z(x_{-\infty}^{-1})$:

$$\tau^k(x) = \max_j \{j \leq -2, \text{ and } z_{-k-l}^{-1} \in [x_{-k-l}^{-1}] \Rightarrow z_{-k-l}^j = x_{-k-l}^j\} . \quad (36)$$

Note that $\tau^k(x) = \tau(x_{-\infty}^{-1})$ iff $k \geq -\tau(x_{-\infty}^{-1}) - 1$. We now define finite versions of the current $J_L(x; 0; \mu^0)$ and of $q(x)$. Write τ^k for $\tau^k(x)$ and put for $k \geq 1$

$$J_L^k(x) = -\log \mu^0(x_{\tau^k+1}^{-1} | x_{-k-l}^{\tau^k}) + \log \sum_{z_{-k-l}^{-1} \in [x_{-k-l}^{-1}]} \mu^0(z_{\tau^k+1}^{-1} | x_{-k-l}^{\tau^k}) \quad (37)$$

and

$$q_k(x) = -\log \mu^1(y_0 | y_{-k}^{-1}) + \log \mu^0(x_0 | x_{-k-l}^{-1}) + J_L^{k+1}(\sigma x) - J_L^k(x) . \quad (38)$$

It is straightforward to check that $q_k(x) \rightarrow q(x)$ a.e. using the properties of $[x_{-k-l}^{-1}]$ and τ^k discussed above. Define $K(x) = -\tau(x_{-\infty}^{-1}) - 1$. Then $K(x)$ is finite and for all $k \geq K(x)$ we have $\tau^k(x) = \tau(x_{-\infty}^{-1})$ and equality in (35). Consequently, the sum over $[x_{-k-l}^{-1}]$ in (37) is the same as the sum over Z in the the expression for J_L . Thus, for $k \geq K(x)$ the only difference between $J_L^k(x)$ and $J_L(x; 0; \mu^0)$ is the length of the history that is conditioned on. Almost everywhere convergence follows from Lemma 3.

We can write

$$\int |q| d\mu^0 \leq \int |q_k| d\mu^0 + \int |q - q_k| d\mu^0 . \quad (39)$$

The result will follow if we can prove that both integrals on the right hand side converge to zero. In order to do this we investigate a particular random process $(g_n)_{n \geq 0}$ on $\mathcal{A}^{\mathbb{Z}}$. The interest in this process is due to the relationship $q_k \circ \sigma^k = \log g_{k-1} - \log g_k$ which will be established later. Define the measurable functions g_n for $n \geq 0$ by

$$g_n(x) = \frac{\mu^1(y_0^n)}{\mu^0([x_{-l}^n])}, \quad \text{where } y_0^n = f(x_{-l}^n) . \quad (40)$$

Here, $\mu^0([x_{-l}^n])$ means $\sum_{z_{-l}^n \in [x_{-l}^n]} \mu^0(z_{-l}^n)$. We will prove that the process $(g_n)_{n \geq 0}$ is a supermartingale with respect to a filtration that we now will describe.

Let \mathcal{D}^n be the partition of $\mathcal{A}^{\mathbb{Z}}$ defined by the equivalence relation (34) on \mathcal{A}^{l+n+1} . That is, the elements of \mathcal{D}^n are the sets $P_{[u]} = \{x : x_{-l}^n \in [u]\}$ for all equivalence classes $[u]$ of \mathcal{A}^{l+n+1} . Let $\mathcal{F}_n = \sigma(\mathcal{D}^n)$. From the definition of the equivalence relation it is obvious that g_n is measurable with respect to \mathcal{F}_n (in fact, \mathcal{F}_n is the σ -algebra generated by g_n). We have to show that $\mathcal{F}_n \subseteq \mathcal{F}_{n+1}$ for all n . This follows if we can show that the partition \mathcal{D}^{n+1} is a refinement of \mathcal{D}^n . Consider a general element $P_{[w_{-l}^n]}$ of \mathcal{D}^n . We claim that

$$P_{[w_{-l}^n]} = \bigcup_{z_{-l}^n \in [w_{-l}^n]} \left(\bigcup_{a \in \mathcal{A}} P_{[z_{-l}^n a]} \right) . \quad (41)$$

If $x \in P_{[w_{-l}^n]}$, then $P_{[x_{-l}^{n+1}]}$ clearly is a member of the double union. Conversely, if x is in some $P_{[z_{-l}^n a]}$ in the union, then $x_{-l}^{-1} = z_{-l}^{-1} = w_{-l}^{-1}$ and $f(x_{-l}^n) = f(z_{-l}^n) = f(w_{-l}^n)$. Thus, $x \in P_{[w_{-l}^n]}$. The claim follows, and $(\mathcal{F}_n : n \geq 0)$ is a filtration.

To prove that g_n is a supermartingale with respect to this filtration we show that $E[g_{n+1} | \mathcal{F}_n] \leq g_n$. Since each sub- σ -algebra \mathcal{F}_n is finite it suffices to show $\int g_{n+1} d\mu^0 \leq \int g_n d\mu^0$ over any $P_{[u]} \in \mathcal{D}^n$. We find that

$$\int_{P_{[u]}} g_n d\mu^0 = \sum_{x_{-l}^n \in [u]} \mu^0(x_{-l}^n) \frac{\mu^1(y_0^n)}{\mu^0([x_{-l}^n])} = \mu^1(y_0^n) . \quad (42)$$

For g_{n+1} we split the integral into cylinder sets where g_{n+1} is constant:

$$\begin{aligned} \int_{P_{[u]}} g_{n+1} d\mu^0 &= \sum_{x_{-l}^n \in [u]} \sum_{a \in \mathcal{A}} \int_{\text{Cyl}(x_{-l}^n a)} g_{n+1}(x) d\mu^0 \\ &= \sum_{x_{-l}^n \in [u]} \sum_{a \in \mathcal{A}} \mu^0(x_{-l}^n a) \frac{\mu^1(y_0^n f(x_{n-l+1}^n a))}{\mu^0([x_{-l}^n a])} \\ &= \sum_{b \in \mathcal{A}} \psi_b \cdot \mu^1(y_0^n b) , \end{aligned} \quad (43)$$

where

$$\psi_b = \sum_{x_{-l}^n \in [u]} \sum_{a \in \mathcal{A}} \frac{\mu^0(x_{-l}^n a)}{\mu^0([x_{-l}^n a])} \cdot 1_{\{f(x_{n-l+1}^n a) = b\}}(x) . \quad (44)$$

We claim that for each b the quantity ψ_b is equal either to 0 or to 1. Fix a b . First note that all blocks $x_{-l}^n a$ in the double sum in (44) that satisfy $f(x_{n-l+1}^n a) = b$ must generate the same equivalence class $[x_{-l}^n a] = [v]$. Conversely, each block $z_{-l}^{n+1} \in [v]$ is an element of the double sum, since it will satisfy $z_{-l}^n \in [x_{-l}^n] = [u]$. Consequently, $\psi_b = 1$ by summation of the fractions and cancelation. The exception is the case where no $x_{-l}^n \in [u]$ can be extended with one symbol to the right such that the new block maps to $y_0^n b$ under f . For such b , $\psi_b = 0$. We can conclude that

$$\int_{P_{[u]}} g_{n+1} d\mu^0 \leq \sum_{b \in \mathcal{A}} \mu^1(y_0^n b) = \mu^1(y_0^n) = \int_{P_{[u]}} g_n d\mu^0 . \quad (45)$$

It is easy to prove that $E[|g_0|] \leq |\mathcal{A}|^{l+1} < \infty$, and this finalizes the proof that g_n is a supermartingale. Furthermore, g_n is non-negative, so by the martingale convergence theorem g_n converges a.e. to a $g \in L^1(\mu^0)$.

We claim that $g \geq 1$ a.e. In fact, it follows from the formula for $\mu^1(y_0^n)$ in (14) that $g_n \geq 1$ for all n . Since g is positive, $\log g_n$ will converge a.e. to $\log g$. In particular $(\log g_n - \log g_{n-1}) \rightarrow 0$ a.e. A careful examination of the expression $\log g_n - \log g_{n-1}$ reveals that

$$\log g_{n-1}(x) - \log g_n(x) = q_n(\sigma^n x) . \quad (46)$$

We sketch the transformations necessary to show this. Firstly,

$$\log g_{n-1}(x) - \log g_n(x) = -\log \frac{\mu^1(y_0^n)}{\mu^1(y_0^{n-1})} + \log \frac{\mu^0([x_0^n])}{\mu^0([x_0^{n-1}])} .$$

The first term is equal to $-\log \mu^1((\sigma^n y)_0 | (\sigma^n y)_{-n}^{-1})$. The second term can be written as

$$\log \frac{\mu^0(x_{-l}^{\tau_2})}{\mu^0(x_{-l}^{\tau_1})} + \log \sum_{z_{-l}^n \in [x_0^n]} \mu^0(z_{\tau_2+1}^n | x_{-l}^{\tau_2}) - \log \sum_{z_{-l}^{n-1} \in [x_0^{n-1}]} \mu^0(z_{\tau_1+1}^{n-1} | x_{-l}^{\tau_1}) ,$$

where $\tau_1 = \tau^n(\sigma^n x) + n \leq n-2$ and $\tau_2 = \tau^{n+1}(\sigma^{n+1} x) + n + 1 \leq n-1$. By adding and subtracting $\log \mu^0(x_{-l}^n)$ to/from the first term, it is transformed to

$$\log \frac{\mu^0(x_{-l}^{\tau_2})}{\mu^0(x_{-l}^{\tau_1})} = \log \mu^0(x_n | x_{-l}^{n-1}) + \log \mu^0(x_{\tau_1+1}^{n-1} | x_{-l}^{\tau_1}) - \log \mu^0(x_{\tau_2+1}^n | x_{-l}^{\tau_2}) ,$$

which are the three remaining terms in $q_n(\sigma^n x)$.

We now proceed to show $L^1(\mu^0)$ convergence of $\log g_n$ using uniform integrability. A family $(f_n)_{n \geq 0}$ of measurable functions is said to be uniformly integrable if

$$\lim_{M \rightarrow \infty} \sup_n \int (|f_n| - M)^+ d\mu = 0 . \quad (47)$$

If $(f_n)_{n \geq 0}$ is an uniformly integrable family and $\lim_{n \rightarrow \infty} f_n(x) = f(x)$ a.e., then ([5, Sec. 1.14])

$$\lim_{n \rightarrow \infty} \int |f - f_n| d\mu \rightarrow 0 .$$

We claim that $(\log g_n)_{n \geq 0}$ is an uniformly integrable family. Note first that

$$\log g_n(x) > t \Leftrightarrow \mu^0([x_{-l}^n]) < 2^{-t} \cdot \mu^1(y_0^n) . \quad (48)$$

Define $A_{n,t} \subset \mathcal{A}^{l+n+1}$ as $A_{n,t} = \{x_{-l}^n : \mu^0([x_{-l}^n]) < 2^{-t} \cdot \mu^1(f(x_{-l}^n))\}$. We obtain

$$\begin{aligned} \mu^0(\{\log g_n > t\}) &= \sum_{x_{-l}^n \in A_{n,t}} \mu^0(x_{-l}^n) \\ &\leq \sum_{x_{-l}^n \in A_{n,t}} 2^{-t} \mu^1(f(x_{-l}^n)) \\ &\leq 2^{-t} \sum_{x_{-l}^n \in \mathcal{A}^{l+n+1}} \mu^1(f(x_{-l}^n)) = 2^{-t} \cdot |\mathcal{A}|^l \end{aligned} \quad (49)$$

for all n . A simple application of Fubini's theorem yields

$$\begin{aligned} \sup_n \int (|\log g_n| - M)^+ d\mu^0 &= \sup_n \int_M^\infty \mu^0(\{\log g_n > t\}) dt \\ &\leq |\mathcal{A}|^l \int_M^\infty 2^{-t} dt = 2^{-M} \cdot |\mathcal{A}|^l . \end{aligned} \quad (50)$$

Thus, uniform integrability is satisfied and $\lim_{n \rightarrow \infty} E[|\log g - \log g_n|] = 0$. By shift-invariance

$$\lim_{n \rightarrow \infty} E[|q_n|] = \lim_{n \rightarrow \infty} E[|q_n| \circ \sigma^n] = \lim_{n \rightarrow \infty} E[|\log g_{n-1} - \log g_n|] = 0 . \quad (51)$$

Hence, the first integral on the right hand side in (39) converges to zero. Regarding the second integral on the right hand side, we know that $\lim_{n \rightarrow \infty} q_n = q$ a.e. Thus, if we can prove that $(q_n)_{n \geq 0}$ is itself a uniformly integrable family then we are done. From (46) and the fact that $\log g_n \geq 0$ for all n it follows that

$$\{|q_n| \circ \sigma^n > M\} \subseteq \{\log g_{n-1} > M\} \cup \{\log g_n > M\} . \quad (52)$$

We can conclude that

$$\sup_n \int (|q_n| - M)^+ d\mu^0 = \sup_n \int (|q_n \circ \sigma^n| - M)^+ d\mu^0 \leq 2^{-M+1} \cdot |\mathcal{A}|^l , \quad (53)$$

and the result follows. \square

A corresponding continuity equation can also be written for right local information. The right variants of the set Z and variable τ are defined by

$$\begin{aligned} Z(x_i^\infty) &= \{z_i^\infty : f(z_i^\infty) = f(x_i^\infty) \text{ and } \exists j \geq i \text{ such that } z_j^\infty = x_j^\infty\} , \\ \tau(x_i^\infty) &= \min_j \{j : j > i - l, \text{ and } z_j^\infty = x_j^\infty \forall z_i^\infty \in Z(x_i^\infty)\} . \end{aligned}$$

Put $Z = Z(x_{i-l}^\infty)$ and $\tau = \tau(x_{i-l}^\infty)$, and define the right information current at coordinate i of x with respect to μ and F by

$$J_R(x; i; \mu) = \log \mu(x_i^{\tau-1} | x_\tau^\infty) - \log \sum_Z \mu(z_{i-l}^{\tau-1} | x_\tau^\infty) . \quad (54)$$

Then, $J_R(x; i; \mu)$ satisfies the continuity equation

$$\Delta_t S_R(x; i; \mu) + \Delta_i J_R(x; i; \mu) = 0 \quad (55)$$

at all $i \in \mathbb{Z}$ μ -a.e. This is proved under the condition that F is reversible or the conditions that F is surjective and μ is shift-invariant in the same way as Theorem 3 or Theorem 4.

5 Further Aspects of Information Transport

5.1 Information theoretic interpretation

We now explain why the expression for J_L presented in Definition 5 is natural from an information theoretic perspective. We describe a way of decomposing $J_L(x; i)$ into

$$J_L(x; i) = J_L^+(x; i) - J_L^-(x; i) , \quad (56)$$

with $J_L^+, J_L^- \geq 0$ such that J_L^+ has a natural interpretation in terms of information flowing to the right between coordinates $i-1$ and i , and J_L^- in terms of information flowing to the left. Here, and in the rest of the section, we omit μ from the notation in J_L and S_L when considering an arbitrary, but fixed, measure μ .

First recall the definition of $Z(x_{-\infty}^i)$ and define

$$Z_0(x_{-\infty}^i) = \{z_{-\infty}^i \in Z(x_{-\infty}^i) : z_{-\infty}^{i-r} = x_{-\infty}^{i-r}\} . \quad (57)$$

We will consider the set $Z_0(x_{-\infty}^{i+r-1})$, which consists of the one sided infinite sequences that have the same image as $x_{-\infty}^{i+r-1}$ and coincide with $x_{-\infty}^{i+r-1}$ up to index $i-1$. In Fig. 2, $Z_0(x_{-\infty}^{i+r-1})$ consists of the two uppermost sequences. Define J_L^+ and J_L^- at coordinate $i = 0$, with $\tau = \tau(x_{-\infty}^{r-1})$, $Z = Z(x_{-\infty}^{r-1})$ and $Z_0 = Z_0(x_{-\infty}^{r-1})$, as

$$J_L^-(x; 0) = -\log \sum_{Z_0} \mu(z_0^{r-1} | x_{-\infty}^{-1}) , \quad (58)$$

$$J_L^+(x; 0) = -\log \frac{\sum_{Z_0} \mu(z_{r+1}^{r-1} | x_{-\infty}^r)}{\sum_Z \mu(z_{r+1}^{r-1} | x_{-\infty}^r)} . \quad (59)$$

It is straightforward to confirm that $J_L^-(x; 0)$ and $J_L^+(x; 0)$ are non-negative and satisfy (56).

We first examine J_L^- . Using the joint measure ν defined in (22) we can write

$$J_L^-(x; 0) = - \lim_{n \rightarrow \infty} \log \frac{\nu(x_{-n}^{-1}, y_{-r}^{-1})}{\nu(x_{-n}^{-1})} = - \log \nu(y_{-r}^{-1} | x_{-\infty}^{-1}) . \quad (60)$$

The equation states that $J_L^-(x; 0)$ is the information gained by observing y_{-r}^{-1} when having knowledge of $x_{-\infty}^{-1}$. This is what one should expect. Indeed, since $x_{-\infty}^{-1}$ is known the semi-infinite sequence $y_{-\infty}^{-r-1}$ is uniquely determined by the CA map. Hence, all uncertainty about $y_{-\infty}^{-1}$ is with respect to y_{-r}^{-1} , and this uncertainty comes from lack of knowledge about the continuation x_0^∞ of $x_{-\infty}^{-1}$. The quantity $-\log \nu(y_{-r}^{-1} | x_{-\infty}^{-1})$ is thus the further information about the continuation x_0^∞ found in $y_{-\infty}^{-1}$ but not in $x_{-\infty}^{-1}$. This information has been transported from x_0^∞ , and consequently a corresponding contribution J_L^- to the information current is generated (note that since the time step is one, the information flow per iteration is equal to the information current). The contribution J_L^- consist of information flowing in the negative x -direction.

Considering J_L^+ , we can write

$$\begin{aligned} J_L^+(x; 0) &= - \lim_{n \rightarrow \infty} \log \frac{\sum_{Z_0} \mu(z_{-n}^{r-1})}{\sum_Z \mu(z_{-n}^{r-1})} \\ &= - \lim_{n \rightarrow \infty} \log \frac{\nu(x_{-n}^{-1}, y_{-n+l}^{-1})}{\nu(x_{-n}^r, y_{-n+l}^{-1})} = - \log \nu(x_{r+1}^{-1} | x_{-\infty}^r, y_{-\infty}^{-1}) . \quad (61) \end{aligned}$$

Thus, $J_L^+(x; 0)$ is the information gained from observing x_{r+1}^{-1} when $x_{-\infty}^r$ as well as $y_{-\infty}^{-1}$ is known. Since $y_{-\infty}^{-1}$ is known, the preimage $x_{-\infty}^{r-1}$ is determined up to the set $Z(x_{-\infty}^{r-1})$. This is illustrated by Fig. 2, where x_{r+1}^{-1} must be one of the ‘‘branches’’ to the right, but it is not decidable from $y_{-\infty}^{-1}$ which one. However, which member of $Z(x_{-\infty}^{r-1})$ that $x_{-\infty}^{r-1}$ actually is will, with probability one, be determined by the continuation y_0^∞ of $y_{-\infty}^{-1}$. Therefore, the information $-\log \nu(x_{r+1}^{-1} | x_{-\infty}^r, y_{-\infty}^{-1})$ flows to the right and is found to the right of coordinate -1 in y . The contribution J_L^+ thus consists of information flowing in the positive x -direction. Note that for non-surjective CA the sets $Z(x_{-\infty}^{r-1})$ will not be finite and which element of $Z(x_{-\infty}^{r-1})$ that is the correct $x_{-\infty}^{r-1}$ will not be determined by y_0^∞ . As a consequence, some information about the preimage will be lost in this case.

5.2 Permutative cellular automata

For the class of right permutative CA the information dynamics has a particularly simple form. Recall that F is right permutative if $R(F) = 1$. Equivalently, F is right permutative if $|Z(x_{-\infty}^i)| = 1$ for all pairs x and i . Therefore, for right permutative CA (58) gives $J_L^-(x; i; \mu) = - \log \mu(x_i^{i+r-1} | x_{-\infty}^{i-1})$ and (59) gives $J_L^+ \equiv 0$. This result and the continuity equation yield the following corollary to Theorem 4.

Corollary 1. *Let μ be any shift-invariant measure on $\mathcal{A}^{\mathbb{Z}}$ and $F: \mathcal{A}^{\mathbb{Z}} \rightarrow \mathcal{A}^{\mathbb{Z}}$ a right permutative CA with right radius r . Then, μ -almost everywhere,*

$$S_{\mathbb{L}}(Fx; i; \mu \circ F^{-1}) = S_{\mathbb{L}}(x; i + r; \mu) . \quad (62)$$

In particular, if $r = 0$, then $S_{\mathbb{L}}(F^t x; i; \mu^t) = S_{\mathbb{L}}(x; i; \mu^0)$ for all $t \geq 0$ so that the local information is locally constant. As a result of Corollary 1, the distribution of local information will also remain unchanged. Below we state this as a separate result.

Corollary 2. *Let μ^0 be any ergodic measure on $\mathcal{A}^{\mathbb{Z}}$ and $F: \mathcal{A}^{\mathbb{Z}} \rightarrow \mathcal{A}^{\mathbb{Z}}$ a right permutative CA. Then, for any measure ν being a linear combination of the measures μ^t , $t \geq 0$, the random variable $S_{\mathbb{L}}(x; i; \nu)$ has the same distribution.*

Proof. Write $\nu = \sum_{t=0}^{\infty} \alpha_t \mu^t$, with $\sum \alpha_t = 1$, and let B be an arbitrary measurable subset of $[0, \infty)$. Since every CA conserves ergodicity, all μ^t are ergodic and satisfy $\mu^t = \nu_x \mu^t$ -almost everywhere, where ν_x is the empirical measure. Hence, by decomposition

$$\begin{aligned} \nu\{x: S_{\mathbb{L}}(x; i; \bar{\mu}_n) \in B\} &= \sum_{t=0}^{\infty} \alpha_t \mu^t(\{x: S_{\mathbb{L}}(x; i; \nu) \in B\}) \\ &= \sum_{t=0}^{\infty} \alpha_t \mu^t(\{x: S_{\mathbb{L}}(x; i; \mu^t) \in B\}) . \end{aligned}$$

The last equality follows by Theorem 2. To show that each term of the sum is the same, let

$$T = \{z: S_{\mathbb{L}}(F^t(z); i - rt; \mu^t) = S_{\mathbb{L}}(z; i; \mu^0) \forall t \geq 0\} .$$

Then $\mu^0(T) = 1$ by Corollary 1 and the fact that F^t is right permutative for all $t > 1$ if F is right permutative [9]. Therefore, for any $t \geq 0$,

$$\begin{aligned} \mu^t(\{x: S_{\mathbb{L}}(x; i; \mu^t) \in B\}) &= \mu^0(G^{-t}(\{x: S_{\mathbb{L}}(x; i; \mu^t) \in B\})) \\ &= \mu^0(G^{-t}(\{x: S_{\mathbb{L}}(x; i; \mu^t) \in B\}) \cap T) \\ &= \mu^0(\{z: S_{\mathbb{L}}(z; i + rt; \mu^0) \in B\}) . \end{aligned}$$

The last equality follows by the definition of T . □

Note that even though the behaviour of the local information is very simple in the case of permutative CA, the sequence μ^t of measures generated by a linear CA under iteration is quite complicated. Block probabilities and the structure of correlations in the system varies widely with t [16]. On the other hand, for many bipermutative CA large classes of initial measures weak* converge in Cesàro mean to the uniform Bernoulli measure. That is,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=0}^{n-1} \mu^t(a_1^k) = \frac{1}{|\mathcal{A}|^k} \quad (63)$$

for all $k \geq 1$ and finite blocks $a_1^k \in \mathcal{A}^k$. This was first proved for the linear CA $F = \sigma + \sigma^{-1}$ on $\{0, 1\}^{\mathbb{Z}}$ with μ^0 a Bernoulli measure by Lind [15]. It has later been extended to a larger subclass of the permutative CA and classes of measures by Pivato and Yassawi [18, 19], see also [3]. These include all Markov measures with full support.

For the uniform Bernoulli measure $\bar{\mu}$ the local information has a uniform distribution: $S_L(x; i; \bar{\mu}) = \log |\mathcal{A}|$ for all x and i . We can use the result on Cesàro convergence to demonstrate that convergence of a sequence $(\mu_n)_{n \geq 0}$ of measures to a limit measure μ in the weak*-topology does not, in any sense, mean that $S_L(x; i; \mu_n)$ converges to $S_L(x; i; \mu)$. Indeed, let F and μ^0 be any combination of a CA and an ergodic measure that such that (63) is valid, and put $\mu_n = \frac{1}{n} \sum_{t=0}^{n-1} \mu^t$. Then μ_n converges to the uniform Bernoulli measure, but by Corollary 2 all $S_L(x; 0; \mu_n)$ have the same non-uniform probability distribution on \mathbb{R} . The reason that the distribution of $S_L(x; i; \mu_n)$ can remain unchanged even though $\mu_n \rightarrow \bar{\mu}$, with $\bar{\mu}$ uniform Bernoulli, is that local information takes *all* correlations in the system into account while the weak* topology only considers finite blocks. In this sense local information yields a different, more microscopic, view of the system than the weak* topology does.

We now use Corollary 1 to demonstrate the necessity of the condition on the measure in Theorem 4. The following example is a construction where the continuity equation $\Delta_I S_L + \Delta_I J_L = 0$ fails to be valid, due to a lack of shift-invariance.

Example 4. Let F be the CA $\sigma^{-1} \circ F_1$, with F_1 from Example 1. Note that F is right permutative, has right radius 0 and local map f given by $f(x_{-1}, x_0) = x_{-1} + x_0 \pmod{2}$. Let μ^0 be the uniform Bernoulli measure on $\{0, 1\}^{\mathbb{Z}}$, except that x_0 always is zero. Formally, each coordinate random variable X_i is independent and distributed according to

$$\mu^0(X_i = 1) = \begin{cases} \frac{1}{2} & \text{for } i \neq 0 \\ 0 & \text{for } i = 0 \end{cases}. \quad (64)$$

Therefore, for all $x \in \mathcal{A}^{\mathbb{Z}}$ we have $S_L(x; i, \mu^0) = 1$ for $i \neq 0$ and $S_L(x; 0, \mu^0) = 0$. However, we claim that $S_L(y; i, \mu^1) = 1$ for all i .

Each sequence $y \in \mathcal{A}^{\mathbb{Z}}$ has two preimages under f , call them z and w . These have the property that z_i and w_i always are different, $z_i = 1 - w_i$. Assume that $z_0 = 0$. We obtain

$$\mu^0(z_j^k) = \begin{cases} 2^{j-k} & \text{if } j \leq 0 \leq k \\ 2^{j-k-1} & \text{otherwise} \end{cases}, \quad \mu^0(w_j^k) = \begin{cases} 0 & \text{if } j \leq 0 \leq k \\ 2^{j-k-1} & \text{otherwise} \end{cases}. \quad (65)$$

The local information $S_L(y; i, \mu^0)$ is the limit $n \rightarrow \infty$ of

$$-\log \mu^1(y_i | y_{i-n}^{i-1}) = -\log \frac{\mu^0(z_{i-n-1}^i) + \mu^0(w_{i-n-1}^i)}{\mu^0(z_{i-n-1}^{i-1}) + \mu^0(w_{i-n-1}^{i-1})}. \quad (66)$$

In all the three cases $i < 0$, $i = 0$ and $i > 0$ inserting the probabilities from (65) gives $-\log \mu^1(y_i | y_{i-n}^{i-1}) = 1$ for all n .

Now assume that $\Delta_t S_L + \Delta_i J_L = 0$ at all i μ^0 -a.e. Since F is right permutative and has $r = 0$, Corollary 1 states that $J_L = 0$ μ^0 -a.e. However, this makes $S_L(y; 0, \mu^1) = 0$, which is not satisfied for the image of any $x \in \mathcal{A}^{\mathbb{Z}}$.

An alternative way to appreciate that $S_L(y; 0, \mu^1) = 1$ for this system is to realize that $y_{-\infty}^{-1}$ does not give any information about x_{-1} . Therefore, even though $x_0 = 0$ with certainty, $\mu^1(y_0 | y_{-\infty}^{-1}) = \frac{1}{2}$. Note that once y_0 is observed, we will have perfect knowledge of $x_{-\infty}^{-1}$. Thus, information about the preimage that is not contained in the tail of y is made available at some position in the sequence.

The information need not appear at a single position as it did in Example 4. A case illustrating this would be to let $\mu(X_i = 1) = \frac{1}{4}$ for all $i \geq 0$ and $\mu(X_i = 1) = \frac{1}{2}$ for $i \leq 0$. Then the correct preimage will be learnt gradually from observing y_0, y_1, y_2, \dots since the fraction of 1's in the preimage block x_{-1}^n will converge either to $\frac{1}{4}$ or $\frac{3}{4}$. In this case, the continuity equation will in general not be satisfied at any $i \geq 0$, but will be an increasingly better approximation as i increases.

Recall that the conserved information quantities S_L and S_R were obtained by taking all correlations in the symbol sequences into account. This was achieved by considering respectively all symbols to the left of or to the right of the coordinate in question as known. This choice of a frame of reference will influence how the information is distributed in the system. Due to the simple behaviour of local information in bipermutative CA these CA serve well to illustrate this phenomenon.

Consider, for instance, the CA on $\{0, 1\}^{\mathbb{Z}}$ defined by the radii $l = r = 1$ and local rule

$$f(x_{-1}, x_0, x_1) = x_{-1} + x_0 + x_1 \pmod{2} . \quad (67)$$

This bipermutative CA is rule 150 in Wolfram's numbering system [29]. Let the initial measure μ^0 be Bernoulli with a very small probability for a 1, say $\mu(1) = 2^{-10}$. Assume that $i = 0$ is the only coordinate in the interval $-100 \leq i \leq 100$ initially having $x_i = 1$. Figure 3 shows the configurations of the interval $-50 \leq i \leq 50$ for all iteration up to time $t = 50$. Coordinate $i = 0$ initially has 10 bits of left and right local information, $S_L(x; 0; \mu^0) = S_R(x; 0; \mu^0) = 10$. Let $y = F^t(x)$. An observer which knows the left history will by observing $y_{-t} = 1$ learn that $x_0 = 1$ and gain 10 bits of information. However, from each of the subsequent symbols the observer will gain only $-\log(1 - 2^{-10})$ bits of information. This is in agreement with Corollary 1. On the other hand, an observer knowing the right history will gain the 10 bits of information by observing that $y_t = 1$.

Thus, the question about where in the pattern the information generated by the unlikely event $\{x_0 = 1\}$ is located at time t cannot be answered without also taking into account which frame of reference an observer has. Knowledge of either the right or left history must be assumed, and these two cases give different answers about the location.

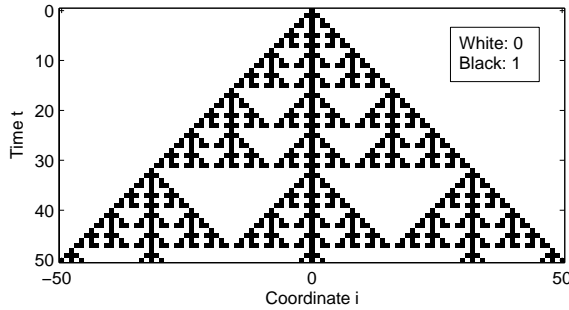


Fig. 3: The evolution of the symbol sequence with a single 1 located at $i = 0$ under CA rule 150. The left local information from the initial 1 is located at the left boundary of the expanding pattern while the right local information is located at the right boundary.

5.3 Bounds on the current

In addition to providing intuition, the splitting of J_L into $J_L^+ - J_L^-$, with components defined in (58) and (59), yields some additional information about the current. Firstly, if the right radius $r = 0$ then $J_L^- \equiv 0$ and $J_L \geq 0$. Hence, left local information only flows to the right in this case. On the contrary, having left radius $l = 0$ does not result in $J_L \leq 0$. This is shown by the following example

Example 5. Consider the CA F_2 from Example 2. Let μ be Bernoulli and x be any sequence with $x_{-2}, x_{-1}, x_0 = 001$. The formula for J_L gives

$$J_L(x; 0; \mu) = \log \left(1 + \frac{\mu(2)^2 - \mu(0)^2}{\mu(0)} \right).$$

So J_L has the same sign as $\mu(2) - \mu(0)$.

For right local information the situation is opposite, if $l = 0$ then $J_R \leq 0$.

Secondly, we can obtain bounds for the information current. We first show that the amount of information that flows from coordinate $i - 1$ to i is limited by the amount of information available in the intervals $[\tau + 1, i - 1]$ and $[i, i + r - 1]$. Let $\tau = \tau(x_{-\infty}^{i+r-1})$. Then

$$-\sum_{k=i}^{i+r-1} S_L(x; k) \leq J_L(x; i) \leq \sum_{k=\tau+1}^{i-1} S_L(x; k). \quad (68)$$

The first inequality follows since from (56), (58) and monotonicity of $\log x$, $J_L(x; i) \geq -J_L^-(x; i) \geq \log \mu(x_i^{i+r-1} | x_{-\infty}^{i-1})$. The second inequality follows directly from the definition of J_L .

The second inequality may seem surprising because the sum starts at $\tau + 1$ rather than at $i - l$. This suggests that left local information in a single iteration can be transported a distance farther to the right than the left radius l . However, a perturbation of one symbol in the initial configuration can only propagate a distance l per time step (often called the rightwards speed of light). The appearance of $\tau + 1$ as the lower limit in the sum therefore warrants a closer examination.

For reversible CA the existence of an inverse CA ensures that $\tau + 1 \geq i - \bar{r}$, where \bar{r} is the right radius of the inverse CA. Therefore, the distance over which information can flow in a single iteration is uniformly bounded for a given reversible CA.

For surjective non-reversible CA τ is in general unbounded. In the following discussion, assume that $r = 0$, since this case gives the maximal flow of left local information to the right. We look at coordinate $i = 0$. The appearance of $\tau + 1$ in (68) is related to the interpretation of J_L^+ as the information gained by observing $x_{\tau+1}^{-1}$ when the image $y_{-\infty}^{-1}$ as well as the history $x_{-\infty}^r$ is known. The second inequality in (68) is an equality if and only if the additional knowledge of $y_{-\infty}^{-1}$ leads to no reduction in information gain compared to knowledge of only $x_{-\infty}^r$. From (59) this is equivalent to having $\sum_Z \mu(z_{\tau+1}^{-1} | x_{-\infty}^r) = 1$. However, since $|Z| \leq R(F)$ the sum is rarely close to this magnitude, particularly when $|\tau|$ is large. A further argument that large information flows are unprobable is the observation that $J_L(x; i; \mu) > s$ requires $\mu(x_{\tau+1}^{-1} | x_{-\infty}^r) / \sum_Z \mu(z_{\tau+1}^{-1} | x_{-\infty}^r) < 2^{-s}$, so $x_{\tau+1}^{-1}$ must be a very unlikely continuation of $x_{-\infty}^r$ to generate a large current s . We illustrate these considerations with the following example.

Example 6. Let the surjective CA F on $\{0, 1, 2\}^Z$ be defined by the radii $l = 1$, $r = 0$ and local function f given by $f(10) = f(11) = f(22) = 0$, $f(12) = f(20) = f(21) = 1$ and $f(00) = f(01) = f(02) = 2$. We demonstrate how unlikely events in this system can generate information flows over large distances in a single iteration.

Let μ be Bernoulli with a low probability $p = \mu(2)$ for the symbol 2 occurring and $q = \mu(0) = \mu(1) = \frac{1-p}{2}$. Although p is small, long blocks of successive 2's will occur at some points. Assume that $x_n^{-1} = 22\dots 2$ while $x_{-n-1}^{-1} = 00$ and $x_0^3 = 0000$. The set $Z(x_{-\infty}^{-1})$ is illustrated to the left in Fig. 4 using the representation introduced in Fig. 2. For p small, the quantity $\sum_Z \mu(z_{\tau+1}^{-1} | x_{-\infty}^r)$ is much larger than $\mu(x_{\tau+1}^{-1} | x_{-\infty}^r)$, since the two other elements in $Z(x_{-\infty}^{-1})$ consist only of 0's and 1's. It follows from (59) that J_L^+ is large and an increasing function of the number n of 2's. Eq. (17) yields for $1 \leq k \leq n$,

$$J_L(x; -n + k; \mu) = \log \left(1 + 2 \left(\frac{q}{p} \right)^k \right) \approx 1 + k \log \left(\frac{q}{p} \right). \quad (69)$$

The current increases approximately linearly in k . The result is that only approximately $-\log q$ bits of information remains at each coordinate $-n \leq i \leq -1$ while the surplus information is transported to the right of $i = -1$. Figure 5 displays the values of S_L at $t = 0$ and $t = 1$ as well as the values of J_L , all for the choice $n = 7$ and $p = 0.1$.

Most of the information is accumulated at position $i = 1$. The reason is found by considering the set $Z(x_{-\infty}^{-1})$, which is illustrated to the right in Fig. 4. Observing the

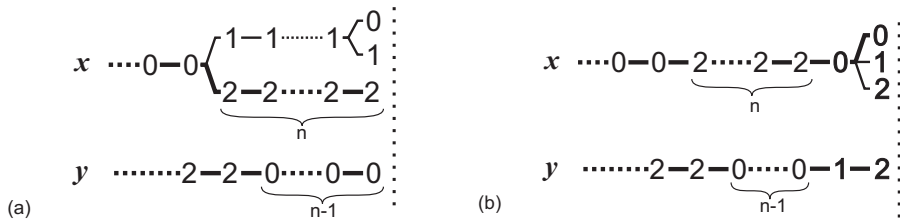


Fig. 4: An illustration of the sets (a) $Z(x_{-\infty}^{-1})$ and (b) $Z(x_{-\infty}^1)$ for the CA in Example 6 when $x_{-n}^{-1} = 22\dots 2$ and $x_{-n-2} = x_{-n-1} = x_0 = x_1 = 0$.

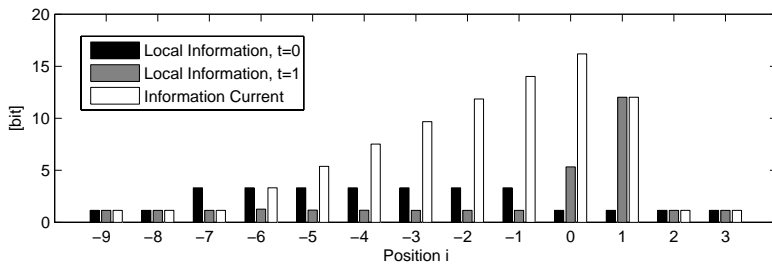


Fig. 5: The distribution of local information S_L and the corresponding information currents $J_L(x; i; \mu)$ for the situation in Example 6 with $n = 7$ and $p = 0.1$. The information is accumulated at coordinates $i = 0$ and $i = 1$ at time $t = 1$.

value $y_1 = 2$ while knowing $y_{-\infty}^0$ establishes that the actual preimage was the one containing the large block of 2's. This was highly improbable and a high local information results.

Finally, note that since F is left permutative, the transport of right local information always occurs at the rightwards speed of light, $J_R(x; i; \mu) = S_L(x; i - 1; \mu)$.

The possibility that $J_L(x; i) > -\log \mu(x_{i-1}^{i-1} | x_{-\infty}^{i-1})$ can be better appreciated by recalling that local information S_L is defined with respect to an infinite frame of reference, namely knowledge of the left history. Therefore, a permutation arbitrary far to the left of i can in principle alter the conditional probability $\mu(x_i | x_{-\infty}^{i-1})$ and hence the local information here. Contrary to this, the propagation of a perturbation in the initial configuration is propagation of information in a different sense. In this case the propagation consists of the symbols at an increasing number of coordinates deviating from some reference symbols, and clearly no frame of reference is needed to detect the deviation of a given symbol.

We can compare the results above to a situation that involves communication between two parts of the lattice. Consider an observer A who knows the initial con-

figuration $x_{-\infty}^0$ of the negative part of the lattice. How much information about the continuation x_1^∞ can A gain by observing the configurations $(F^k x)_{-\infty}^0$ for $0 < k \leq t$? This question can be answered by using the concept of relative entropy, or Kullback Liebler distance [13]. The relative entropy of a posterior measure μ with respect to a prior measure μ_0 satisfying $\mu \ll \mu_0$ is defined as

$$D(\mu||\mu_0) = \int_X \log \frac{d\mu}{d\mu_0} d\mu , \quad (70)$$

where $\frac{d\mu}{d\mu_0}$ is the Radon-Nikodym derivative. The quantity $D(\mu||\mu_0)$ is non-negative and is interpreted as the Shannon information gained by going from the prior to the posterior.

The posterior is in our case expressed in terms of the joint measure ν^t of all times $0 \leq k \leq t$ obtained as a straightforward generalization of ν defined in (22). For $x \in \mathcal{A}^{\mathbb{Z}}$ define the measure μ_x^t on $\sigma(X_i : i > 0)$ through the block probabilities

$$\mu_x^t(z_1^n) = \nu^t(z_1^n | x_{-\infty}^0, (F x)_{-\infty}^0, \dots, (F^t x)_{-\infty}^0), \quad n \geq 1 . \quad (71)$$

Thus, $\mu_x^t(z_1^n)$ is the probability that $x_1^n = z_1^n$ given knowledge of the symbols at all coordinates $i \leq 0$ up to time step t . The information A gains by time t is hence given by the relative entropy $D(\mu_x^t||\mu_x^0)$. The following relations are valid

Proposition 1. *The measures defined in Eq. (71) satisfies*

$$D(\mu_x^t||\mu_x^0) \leq \sum_{i=1}^{rt} S_L(x; i; \mu) , \quad (72)$$

$$D(\mu_x^1||\mu_x^0) = J_L^-(x; 1; \mu) . \quad (73)$$

In particular this means that A during the first t iterations of F cannot gain more information about $x_{-\infty}^0$ than the left local information initially located within the interval $[1, rt]$. Conversely, if B is an observer knowing x_0^∞ and observing the symbols at $i \geq 0$ for times $1 \leq k \leq t$ his information gain about $x_{-\infty}^{-1}$ would be bounded by $\sum_{i=-1t}^{-1} S_R(x; i; \mu)$.

Proof. Define $B_x \subseteq \mathcal{A}^{rt}$ as $B_x = \{z_1^{rt} : f^k(x_{-kr}^0, z_1^{kr}) = f(x_{-kr}^{kr}) \text{ for } 1 \leq k \leq t\}$. Both results follow from

$$\begin{aligned} D(\mu_x^t||\mu_x^0) &= \lim_{n \rightarrow \infty} \sum_{z_1^n} \mu_x^t(z_1^n) \log \frac{\mu_x^t(z_1^n)}{\mu_x^0(z_1^n)} \\ &= \lim_{n \rightarrow \infty} \sum_{z_1^n} \mu(z_1^n | x_{-\infty}^0, z_1^{rt} \in B_x) \log \frac{\mu(z_1^n | x_{-\infty}^0, z_1^{rt} \in B_x)}{\mu(z_1^n | x_{-\infty}^0)} \\ &= -\log \mu(\{z_1^{rt} \in B_x\} | x_{-\infty}^0) \\ &\leq -\log \mu(x_1^{rt} | x_{-\infty}^0) . \end{aligned}$$

□

We now move on to determine bounds on the *average* information flow generated by a surjective one-dimensional CA.

Theorem 5. *Let $J_L(x; i; \mu)$ be the information current with respect to a surjective CA F and a measure μ . Then, for each $i \in \mathbb{Z}$, $J_L(x; i; \mu) \in L^1(\mu)$. Furthermore, if μ is shift-invariant then $E[J_L]$ satisfies the relationship*

$$-rh(\mu) \leq E[J_L] \leq \log R(F) - rh(\mu) \quad (74)$$

The term $\log R(F)$ on the right hand side in (74) is related to the interpretation of J_L^+ as the information about which member of $Z(x_{-\infty}^{r-1})$ that $x_{-\infty}^{r-1}$ is. Since $|Z(x_{-\infty}^{r-1})| \leq R(F)$, the average of this information cannot exceed $\log R(F)$. The term $-rh(\mu)$ is related to J_L^- .

Proof. We look at coordinate $i = 0$. The current can be written as

$$J_L(x; 0; \mu) = -\log \frac{\mu(x_{\tau+1}^{r-1} | x_{-\infty}^r)}{\sum_Z \mu(x_{\tau+1}^{r-1} | x_{-\infty}^r)} + \log \mu(x_0^{r-1} | x_{-\infty}^{r-1}) \quad (75)$$

where the first term is non-negative and the second term is non-positive. To bound the integral of the first term we divide $\mathcal{A}^{\mathbb{Z}}$ into the sets $T_k = \{x : \tau(x_{-\infty}^{r-1}) = k\}$ for $k \leq -2$. Furthermore, we wish to subdivide each T_k though an equivalence relation similar to that defined in (34). Define the following relation on $\mathcal{A}^{|k|+2r+l-1}$:

$$x_{k+1-l-r}^{r-1} \sim z_{k+1-l-r}^{r-1} \text{ iff } f_{(k+1-l-r)}^{(r-1)} = f_{(k+1-l-r)}^{(r-1)} \text{ and } x_{k+1-l-r}^k = z_{k+1-l-r}^k . \quad (76)$$

Then, transfer the equivalence relation to T_k through

$$x \sim z \text{ iff } x_{k+1-l-r}^{r-1} \sim z_{k+1-l-r}^{r-1} \quad (77)$$

We denote the equivalence classes of T_k by $P_{k,j}$ with j in some finite index set. Furthermore, for each $P_{k,j}$ denote the corresponding equivalence class of $\mathcal{A}^{|k|+2r+l-1}$ by $\bar{P}_{k,j}$. Each $\bar{P}_{k,j}$ has at most $R(F)$ members.

For each j there is a set $P_{k,j}^- \in \sigma(X_i : i \leq k-l-r)$ of histories $x_{-\infty}^{k-l-r}$ such that

$$P_{k,j} = P_{k,j}^- \bigcap \bigcup_{\bar{P}_{k,j}} \text{Cyl}(z_{k+1-l-r}^k) . \quad (78)$$

If $|\bar{P}_{k,j}| = R(F)$ then $P_{k,j}^- = \mathcal{A}^{\mathbb{Z}}$, but otherwise $P_{k,j}^-$ can be a subset of $\mathcal{A}^{\mathbb{Z}}$. For instance, let F be the CA from Example 6 and let x have $x_i = 1$ for all i . Then $\tau(x_{-\infty}^{r-1}) = -2$ and the set $P_{-2,j}$ containing x satisfies $P_{-2,j} = \{z : z_i \neq 0 \text{ for } i \leq -3\}$ and $|\bar{P}_{-2,j}| = 2$.

Using the subdivision, we can write

$$\begin{aligned} - \int_{\mathcal{A}^{\mathbb{Z}}} \log \frac{\mu(x_{\tau+1}^{r-1} | x_{-\infty}^r)}{\sum_Z \mu(x_{\tau+1}^{r-1} | x_{-\infty}^r)} d\mu = \\ \sum_{k=-\infty}^{-2} \sum_{P_{k,j} \in T_n} \int_{P_{k,j}^-} \Psi_{k,j}(x_{-\infty}^{r-l-r}) \left(\sum_{\bar{P}_{k,j}} \mu(z_{\tau+1-l-r}^{r-1} | x_{-\infty}^{r-l-r}) \right) d\mu(x_{-\infty}^{r-l-r}) , \quad (79) \end{aligned}$$

where

$$\Psi_{k,j}(x_{-\infty}^{r-l-r}) = - \sum_{\tilde{P}_{k,j}} \frac{\mu(z_{r+1}^{r-1}|x_{-\infty}^r)}{\sum_{\tilde{P}_{k,j}} \mu(z_{r+1}^{r-1}|x_{-\infty}^r)} \log \frac{\mu(z_{r+1}^{r-1}|x_{-\infty}^r)}{\sum_{\tilde{P}_{k,j}} \mu(z_{r+1}^{r-1}|x_{-\infty}^r)} . \quad (80)$$

The function $\Psi_{k,j}(x_{-\infty}^{r-l-r})$ is for each $x_{-\infty}^{r-l-r}$ the entropy of a discrete random variable with at most $R(F)$ different outcomes. Therefore,

$$\Psi_{k,j}(x_{-\infty}^{r-l-r}) \leq \log R(F) .$$

By using the inequality we obtain from (79) that

$$- \int_{\mathcal{A}^{\mathbb{Z}}} \log \frac{\mu(x_{r+1}^{r-1}|x_{-\infty}^r)}{\sum_Z \mu(x_{r+1}^{r-1}|x_{-\infty}^r)} d\mu \leq \log R(F) \sum_{k=-\infty}^{-2} \sum_{T_k} \mu(P_{k,j}) = \log R(F) . \quad (81)$$

Considering the second term in (75),

$$\int_{\mathcal{A}^{\mathbb{Z}}} \log \mu(x_0^{r-1}|x_{-\infty}^{-1}) d\mu \geq -r \log |\mathcal{A}| .$$

Therefore,

$$E[|J_L(x; 0; \mu)|] \leq r \log |\mathcal{A}| + \log R(F) < \infty ,$$

so $J_L(x; 0; \mu) \in L^1(\mu)$. The second statement follows since for μ shift-invariant,

$$\int_{\mathcal{A}^{\mathbb{Z}}} \log \mu(x_0^{r-1}|x_{-\infty}^{-1}) d\mu = -r h(\mu) .$$

□

From Theorem 5 one can also construct a uniform bound on the average information current. For any surjective CA and any measure μ , the following relation is valid:

$$\frac{|E[J_L(x; i; \mu)]|}{(l+r) \log |\mathcal{A}|} \leq 1 \quad (82)$$

It is worth noting that the higher the average value of the function Ψ from (80) is, the larger the transport of left local information to the left. This average value is maximal when each element in almost all $Z(x_{-\infty}^{r-1})$ have the same conditional probability $\mu(z_{r+1}^{r-1}|x_{-\infty}^r)$. This is in particular the case for the uniform Bernoulli measure $\bar{\mu}$, and we close this section by looking at information transport for this measure.

It is well known that $\bar{\mu}$ is invariant for all surjective CA. This is a direct consequence of the fact that all finite blocks have the same number of preimages. With $\bar{\mu}$ there are no correlations in the system, and $S_L(x; i; \bar{\mu}) \equiv S_R(x; i; \bar{\mu}) \equiv \log |\mathcal{A}|$. The information currents J_L and J_R are consequently also constant, but these depend on the radii and the Welch coefficients $L(F)$, $M(F)$ and $R(F)$. Using (17) and (54) we obtain

$$J_L(x; i; \bar{\mu}) \equiv \log R - r \log |\mathcal{A}| , \quad (83)$$

$$J_R(x; i; \bar{\mu}) \equiv -\log L + l \log |\mathcal{A}| . \quad (84)$$

By using the relation $L \cdot M \cdot R = |\mathcal{A}|^{L+r}$ we obtain the interesting consequence

$$J_R - J_L \equiv \log M . \quad (85)$$

Thus, the right local information flows with a greater or equal velocity to the right than left local information. Furthermore, the sum of the velocity of left information *to the left* and right information *to the right* is constant, and only depends on $M(F)$, the number of preimages that almost all bi-infinite sequences possess. The higher M is, the higher the potential for information transport. The choice of radii decides how the potential is allocated to transport of information to the right and to the left.

6 Conclusions and Discussion

We have studied the quantity $S_L(x; i; \mu)$, called left local information. S_L measures the information at each position of a bi-infinite symbol sequence with respect to the probability measure μ . A position i has high left local information if the symbol at i is unlikely given knowledge of its infinite context to the left. We have also defined the corresponding right local information S_R . In Theorem 1 we have shown that the local information is well-defined and that $S_L(x; i; \mu) \in L^1(\mu)$ with $E[S_L]$ equal to the entropy $h(\mu)$ of the measure. For shift-invariant measures we proved in Theorem 2 that $S_L(x; i; \mu)$ in principle can be calculated from knowledge of x alone, since $S_L(x; i; \mu) = S_L(x; i; \nu_x)$ μ -a.e. with ν_x denoting the empirical measure.

The main concern of the paper has been to investigate transport of local information in the time evolution of a cellular automaton F . In Sect. 4, we have introduced an information current $J_L(x; i; \mu)$ such that the continuity equation $\Delta_t S_L + \Delta_i J_L = 0$ holds. Theorem 3 showed this for the case of one-dimensional reversible cellular automata with no restrictions on the measure. In Theorem 4 we extended this result to general surjective rules, where the global CA mapping is finite-to-one, under the condition of μ being shift-invariant. For a surjective but non-reversible rule and a measure which is not shift-invariant, the continuity equation of the form above may fail to be valid, as demonstrated in Example 4.

We have also given an information theoretic interpretation of the current, and shown bounds for the information flow.

The fact that the local information is a locally conserved quantity for *all* measures under iteration of any reversible CA is a clear indication that the function S_L is an appropriate local information measure in a spatially extended system. However, we still need to consider the fact that information is not a strictly local quantity when correlations are present, and that it depends on the choice of context. In one dimension, this is illustrated by the fact that both the left and right local information are locally conserved, and in general different (as seen, e.g., for bipermutative CA). We have also given other examples which illustrates the limits of locality when correlations are present.

We are currently investigating how the continuity equation can be extended to other classes of cellular automata. In particular this includes non-surjective and probabilistic CA in dimension one as well as CA in dimension two and higher. For non-surjective or non-deterministic systems a continuity equation must take loss and production of information into account. One can also look at local information and transport of local information for other types of spatially extended dynamical systems, in particular coupled map lattices. Extensions of the formalism to other systems will also bring us closer to addressing fundamental issues relating to information transport and conservation in physical systems.

Local information could also be used to detect and quantify structure in a lattice configuration or to investigate the structural properties of a measure. For instance, the variability of S_L , such as the standard deviation of its distribution, says something about how the information is distributed in the system. Furthermore, the convergence properties of $-\log \mu(x_i | x_{i-n}^{i-1})$ give information about correlations in the system.

The continuity equation is a fundamental property of information transport in reversible systems. But we expect local information in cellular automata to have further interesting properties. In particular, a continuity equation is a constraint, rather than an equation that determines the dynamics of the system. One may expect information flow to have different dynamic characteristics in different cellular automata. In particular it should be investigated whether some systems allow a description of the dynamics of information separate from the underlying dynamical system, which would provide an additional argument for viewing information as a fundamental physical quantity.

References

- [1] Bennett, C. Logical reversibility of computation. *IBM J. Res. Develop.*, 17(6):525–532, 1973.
- [2] Dab, D., Lawniczak, A., Boon, J., and Kapral, R. Cellular-automaton model for reactive systems. *Phys. Rev. Lett.*, 64:2462–2465, 1990.
- [3] Ferrari, P., Maass, A., Martínez, S., and Ney, P. Cesàro mean distribution of group automata starting from measures with summable decay. *Ergodic Theory Dynam. Systems*, 20(6):1657–1670, 2000.
- [4] Frisch, U., Hasslacher, B., and Pomeau, Y. Lattice-gas automata for the Navier-Stokes equation. *Phys. Rev. Lett.*, 56:1505–1508, 1986.
- [5] Gänsler, P. and Stute, W. *Wahrscheinlichkeitstheorie*. Springer Verlag, 1977.
- [6] Grassberger, P. Toward a quantitative theory of self-generated complexity. *Int. J. Theor. Phys.*, 25(9):907–938, 1986.
- [7] Gray, R. M. *Probability, random processes, and ergodic properties*. Springer-Verlag, New York, 1988.

-
- [8] Hardy, J., Pomeau, Y., and de Pazzis, O. Time evolution of two-dimensional model system. I. invariant states and time correlation functions. *J. Math. Phys.*, 14:1746–1759, 1973.
- [9] Hedlund, G. A. Endomorphisms and automorphisms of the shift dynamical system. *Math. Syst. Theory*, 3:320–375, 1969.
- [10] Ito, M., Osato, N., and Nasu, M. Linear cellular automata over Z_m . *J. Comput. System Sci.*, 27(2):125–140, 1983.
- [11] Jaynes, E. T. Information theory and statistical mechanics. *Phys. Rev.*, 106:620–630, 1957.
- [12] Keller, G. *Equilibrium states in ergodic theory*, volume 42 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 1998.
- [13] Kullback, S. and Leibler, R. A. On information and sufficiency. *Ann. Math. Statistics*, 22:79–86, 1951.
- [14] Landauer, R. Irreversibility and heat generation in the computing process. *IBM J. Res. Develop.*, 5(3):183–191, 1961.
- [15] Lind, D. A. Applications of ergodic theory and sofic systems to cellular automata. *Physica D*, 10:36–44, 1984.
- [16] Lindgren, K. Correlations and random information in cellular automata. *Complex Systems*, 1:529–543, 1987.
- [17] Lindgren, K. Microscopic and macroscopic entropy. *Phys. Rev. A*, 38:4794–4798, 1988.
- [18] Pivato, M. and Yassawi, R. Limit measures for affine cellular automata. *Ergodic Theory Dynam. Systems*, 22(4):1269–1287, 2002.
- [19] Pivato, M. and Yassawi, R. Limit measures for affine cellular automata II. *Ergodic Theory Dynam. Systems*, 24(6):1961–1980, 2004.
- [20] Richardson, D. Tessellations with local transformations. *J. Comput. System Sci.*, 5:373–388, 1972.
- [21] Shannon, C. E. A mathematical theory of communication. *Bell System Tech. J.*, 27:379–423, 623–656, 1948.
- [22] Takesue, S. Reversible cellular automata and statistical mechanics. *Phys. Rev. Lett.*, 59:2499–2502, 1987.
- [23] Takesue, S. Fourier’s law and the Green-Kubo formula in a cellular-automaton model. *Phys. Rev. Lett.*, 64:252–255, 1990.
- [24] Toffoli, T. Information transport obeying the continuity equation. *IBM J. Res. Develop.*, 32(1):29–36, 1988.
- [25] Toffoli, T. and Margolus, N. H. Invertible cellular automata: a review. *Physica D*, 45(1-3):229–253, 1990.

- [26] Vichniac, G. Simulating physics with cellular automata. *Physica D*, 10:96–115, 1984.
- [27] Walters, P. *An Introduction to Ergodic Theory*. Number 79 in Graduate Texts in Mathematics. Springer, 1982.
- [28] Wheeler, J. A. *Information, Physics, Quantum: The Search for Links, in Complexity, Entropy and the Physics of Information*. Addison-Wesley, Redwood City, CA, 1988.
- [29] Wolfram, S. Universality and complexity in cellular automata. *Physica D*, 10:1–35, 1984.
- [30] Zurek, W. H. Algorithmic randomness and physical entropy. *Phys. Rev. A*, 40:4731–4751, 1989.