A probabilistic automata network epidemic model with births and deaths exhibiting cyclic behaviour

1994 J. Phys. A: Math. Gen. 27 1585

(http://iopscience.iop.org/0305-4470/27/5/022)

J. Phys. A: Math. Gen. 27 (1994) 1585–1597. Printed in the UK

A probabilistic automata network epidemic model with births and deaths exhibiting cyclic behaviour

Nino Boccara†‡, Kyeong Cheong‡ and Mark Oram‡
† DRECAM-SPEC, CE Saclay, 91191 Gif-sur-Yvette Cedex, France
‡ Department of Physics, University of Illinois, Chicago, IL 60607-7059, USA

Received 17 December 1993, in final form 4 January 1994

Abstract. A probabilistic automata network model for the spread of an infectious disease in a population of moving individuals is studied. The local rule consists of two subrules. The first one, applied synchronously, models infection, birth and death processes. It is a probabilistic cellular automaton rule. The second, applied sequentially, describes the motion of the individuals. The model contains six parameters: the probabilities p for a susceptible to become infected by contact with an infective; the respective birth rates b_s and b_i of the susceptibles from either a susceptible or an infective parent; the respective death rates d_s and d_i of susceptibles and infectives; and a parameter m characterizing the motion of the individuals. The model has three fixed points. The first is trivial, it describes a stationary state with no living individuals. The second corresponds to a disease-free state with no infectives. The third and last one characterizes an endemic state with non-zero densities of susceptibles and infectives. Moreover, the model may exhibit oscillatory behaviour of the susceptible and infective densities as functions of time through a Hopf-type bifurcation. The influence of the different parameters on the stability of all these states is studied with a particular emphasis on the influence of motion which has been found to be a stabilizing factor of the cyclic behaviour.

1. Introduction

This paper discusses a general epidemic model which takes into account the infection of susceptibles by contact with infectives, birth of susceptibles from either a susceptible or infective parent, death of susceptibles and infectives and motion of the individuals.

Our model is formulated in terms of automata networks (Goles and Martínez 1990) which describe the local character of the infection process more correctly. An automata network is a graph with a discrete variable at each vertex which evolves in discrete time steps according to a definite rule involving the values of neighbouring vertex variables. The vertex variables may be updated sequentially or synchronously.

Automata networks are discrete dynamical systems, which may be defined more formally as follows.

Let $G = (V, E)$ be a graph, where V is a set of vertices and E a set of edges. Each edge joins two vertices that are not necessarily distinct. An automata network, defined on V, is a triple $(G, Q, \{f_i | i \in V\})$, where G is a graph on V, Q a finite set of states and $f_i : Q^{|U_i|} \to Q$ a mapping, called the local transition rule associated with vertex i. $U_i = \{j \in V | \{j, i\} \in E\}$ is the neighbourhood of i, i.e. the set of vertices connected to i, and $|U_i|$ denotes the number of vertices belonging to U_i. The graph G is assumed to be locally finite, i.e. for all $i \in V$, $|U_i| < \infty$.

In our model the set V is the two-dimensional torus $\mathbf{Z}_L^2$, where $\mathbf{Z}_L$ is the set of integers modulo L. A vertex is either empty or occupied by either a *susceptible*, i.e. an individual

0305-4470/94/051585+13$19.50

who is not infected but who is capable of contracting the disease and becoming infective; or an *infective*, i.e. an individual who is capable of transmitting the disease to susceptibles.

The evolution of these two populations is governed by the following rules:

(i) Susceptibles become infective by contact, i.e. a susceptible may become infective with a probability p if, and only if, it is in the neighbourhood of an infective. This hypothesis neglects latent periods, an infected susceptible becomes immediately infective.

(ii) Susceptibles and infectives may die with respective probabilities d_s and d_i. In this case, they are removed from the lattice and the site they occupied becomes vacant. This assumption states that death is equally likely among each group of individuals, which means, in particular, that these two parameters are supposed to be independent of age and the length of time an individual has been infective.

(iii) Susceptibles and infectives may give birth at a neighbouring empty site to a susceptible with respective probabilities b_s and b_i. That is, we assume that all newborns are susceptibles and have only one parent.

(iv) The time unit is the time step. During one time step, the preceding rules are applied after the individuals have moved on the lattice according to the following specific rule.

(v) An individual selected at random may move to a vertex also chosen at random. If the chosen vertex is empty the individual will move, otherwise the individual will not move. The set in which the vertex is randomly chosen depends on the range of the move. To illustrate the importance of this range, we considered two extreme cases. The chosen vertex may be either one of the four nearest neighbours or be any vertex of the graph. These two particular types of move will be called, respectively, *short-* and *long-range* moves. If N is the total number of sites of $\mathbf{Z}_L^2$, and c the total density of individuals, mcN individuals, where m is a real positive number, are sequentially selected at random to perform a move. This sequential process allows some individuals to move more than others. Since an individual may only move to an empty site, the parameter m represents the average number of *tentative* moves per individual during a unit of time.

This model is an automata network with a mixed transition rule. That is, at each time step, the evolution results from the application of two subrules. The first subrule models infection, birth and death processes. It is a three-state cellular automaton rule applied *synchronously*. The second one specifies the motion of the individuals. It is applied *sequentially*. Both subrules are probabilistic and translation invariant, i.e. they do not depend upon the vertex i.

2. Mean-field approximation

The mean-field approximation ignores space dependence and neglects correlations. It assumes that the probability that either a susceptible or an infective occupies a lattice site is proportional to the density of the corresponding population. In lattice models with local interactions, quantitative predictions of such an approximation are not very good, but, for the epidemic model described in the preceding section, since the second subrule represents a process that destroys the correlations created by the first subrule, if m tends to ∞, the mean-field approximation becomes exact.

Let $S(t)$ and $I(t)$ denote the densities at time t of, respectively, susceptibles and infectives. We have

$$\begin{aligned} S(t+1) &= S(t) + F_1(S(t), I(t)) \\ &= S(t) + (1 - S(t) - I(t))\, f(b_s S(t) + b_i I(t)) - d_s S(t) - (1 - d_s) S(t) f(p I(t)) \end{aligned} \tag{1}$$

$$I(t+1) = I(t) + F_2(S(t), I(t)) = I(t) + (1-d_s)S(t)f(pI(t)) - d_iI(t) \tag{2}$$

where the function f is defined by

$$f(x) = 1 - (1-x)^z. \tag{3}$$

z is the number of neighbouring vertices of a given vertex. $z = 4$ for the square lattice. The expression for $f(x)$ is straightforward to derive. If x is the probability of either 'at time t, a susceptible is infected by an infective located at a specific neighbouring site', i.e. $x = pI(t)$; or 'at time t, a susceptible or an infective gives birth to a susceptible at a specific neighbouring site', i.e. $x = b_sS(t) + b_iI(t)$, then $(1-x)^z$ is the probability that such an event does not occur, and, finally, $1-(1-x)^z$ is the probability that such an event occurs at any neighbouring site. Note that, within the framework of this approximation, the interaction terms are not bilinear as in most models (Bailey 1975, Waltmam 1974, Anderson and May 1991). Non-bilinear interactions have recently been shown to exhibit very different dynamic behaviour (Hethcote and van den Driessche 1991).

The fixed points are the solutions of the equations

$$(1-S-I)f(b_sS + b_iI) - d_sS - (1-d_s)Sf(pI) = 0 \tag{4}$$

$$(1-d_s)Sf(pI) - d_iI = 0. \tag{5}$$

These fixed points are stable if the absolute value of the eigenvalues λ_1 and λ_2 of the Jacobian matrix

$$\mathbf{J} = \begin{bmatrix} 1 + \partial F_1/\partial S & \partial F_1/\partial I \\ \partial F_2/\partial S & 1 + \partial F_2/\partial I \end{bmatrix}$$

are less than 1.

Since $f(0) = 0$, $I = 0$ is a solution of equation (5). In this case, equation (4) can have two different solutions, either $S = 0$ or $S = S_0$ such that

$$(1-S_0)f(b_sS_0) - d_sS_0 = 0. \tag{6}$$

The solution $(0, 0)$ always exists. It is a stable equilibrium if $-1 < \lambda_1(0,0) < 1$, i.e.

$$4b_s - d_s < 0. \tag{7}$$

Since $0 < d_i < 1$, the eigenvalue $\lambda_2(0,0) = 1 - d_i$ is positive and less than 1.

The solution $(S_0, 0)$, which characterizes a disease-free state, exists if the trivial fixed point $(0, 0)$ is unstable. The eigenvalues of $\mathbf{J}(S_0, 0)$ are

$$\lambda_1(S_0, 0) = 1 - f(b_sS_0) + (1-S_0)b_sf'(b_sS_0) - d_s$$

$$\lambda_2(S_0, 0) = 1 + 4p(1-d_s)S_0 - d_i.$$

When $(S_0, 0)$ exists, it can be verified that $\lambda_1(S_0, 0)$ is positive and less than 1. Therefore, $(S_0, 0)$ is stable if

$$4p(1-d_s)S_0 - d_i < 0. \tag{8}$$

When the disease-free state is stable, a non-zero initial density of infectives will eventually go to zero. If we assume that this initial density $I(0)$ is small, then, at $t = 1$, from equation (2) it follows that

$$I(1) - I(0) = (4(1 - d_s)pS(0) - d_i)I(0) + O(I^2(0)) \tag{9}$$

i.e.

$$I(1) - I(0) > 0 \qquad \text{if } S(0) > d_i/[4(1 - d_s)p].$$

An epidemic occurs if the initial density of the susceptible population is larger than a threshold value equal to $d_i/(1 - d_s)zp$ (figure 1). This threshold theorem was first established by Kermack and McKendrick (1927) using an epidemic model formulated in terms of a set of three differential equations.

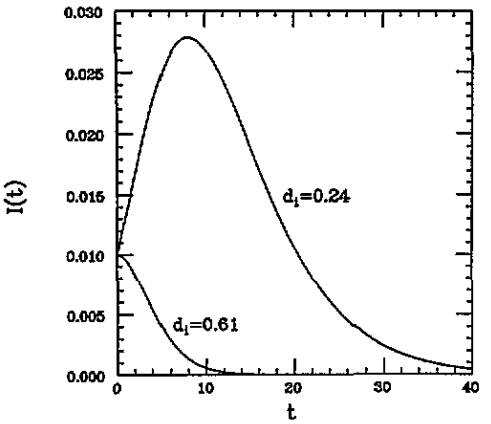

Figure 1. Mean-field approximation. Time evolution of the density of infectives. $I(0) = 0.01$, $S(0) = 0.59$, $z = 4$, $p = 0.3$, $b_s = 0.051$, $b_i = 0.001$, $d_s = 0.2$, $d_i = 0.24$ ($S(0) > d_i/(1 - d_s)zp$) and $d_i = 0.61$ ($S(0) < d_i/(1 - d_s)zp$).

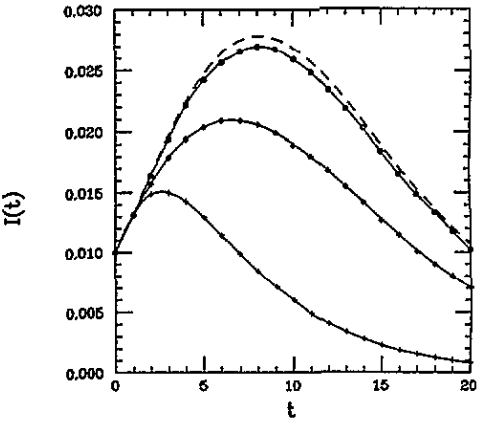

Figure 2. Time evolution of an epidemic for different values of m in the case of short-range moves. $I(0) = 0.01$, $S(0) = 0.59$, $p = 0.3$, $b_s = 0.051$, $b_i = 0.001$, $d_s = 0.2$, $d_i = 0.24$; 200×200 lattice. Each point represents the average of 20 experiments: $+$, $m = 0$; $\diamond$, $m = 5$; $\circ$, $m = 100$. The broken curve corresponds to the mean-field approximation.

If $I \neq 0$, equations (4) and (5) may have another solution which will be denoted (S^*, I^*). This fixed point characterizes an endemic state. It exists when condition (8) is violated, i.e. when the disease-free state is unstable. I^* goes to zero at the bifurcation point as $4p(1 - d_s)S_0 - d_i$.

The expression for the Jacobian matrix $\mathbf{J}(S^*, I^*)$ is rather complicated and it is numerically easier to study the stability of this fixed point. An interesting feature of the model is that (S^*, I^*) may lose its stability and a limit cycle will become stable through a Hopf bifurcation.

Within the mean-field approximation (equations (1)–(2)), our model has three equilibrium points and exhibits three bifurcations.

(i) For all the values of the five parameters of the model—p, b_s, b_i, d_s and d_i—the trivial fixed point (0, 0) always exists.

(ii) $(S_0, 0)$, which characterizes a disease-free state, exists if, and only if, $(0, 0)$ is unstable. The stability of $(0, 0)$ is transferred to $(S_0, 0)$ when $4b_s - d_s = 0$.

(iii) (S^*, I^*), which characterizes an endemic state, exists if, and only if, $(S_0, 0)$ is unstable. The transfer of stability from $(S_0, 0)$ to (S^*, I^*) occurs when $4p(1-d_s)S_0 - d_i = 0$. This bifurcation is similar to a second-order phase transition, the role of the order parameter being played by the density of infectives I^* which tends continuously to zero at the transition point.

(iv) Finally, the system exhibits a Hopf bifurcation when (S^*, I^*), which is a spiral node, loses its stability. A numerical study shows that, if four among the five parameters that characterize the model have fixed values, the stable oscillatory state is favoured if either p or d_i is increased or if b_i, b_s or d_s is decreased.

3. Simulations

In our simulations, the emphasis is on the influence of motion. To discuss size effects, we have considered $L \times L$ lattices for values of L ranging from 100 to 1000.

To show the general character of our model, we shall describe:

(i) the approach of the fixed point $(S_0, 0)$ when it is stable;

(ii) the second-order phase transition from the endemic state to the disease-free state; and

(iii) the stable oscillatory behaviour.

3.1. The approach of $(S_0, 0)$

This model generalizes a simpler automata network epidemic model studied recently by Boccara and Cheong (1992) in which the parameters b_s, b_i and d_s are neglected. This model exhibits a similar behaviour. Figure 2 shows the influence of the parameter m on the time evolution of an epidemic for short-range moves. As m increases the density of infectives as a function of time tends to the mean-field result. Figure 3 shows that, to the precision of our measurements, size effects are negligible if $L \geqslant 200$. Similar results have been obtained for different values of m up to $m = 100$.

3.2. Transition endemic state → disease-free state

The influence of the motion of the individuals on this transition has been studied in detail by Boccara and Cheong (1993) in the case of a simpler automata network model in which only infection and recovery were taken into account. In the case of the present model we have studied the critical behaviour in the vicinity of the transition point as a function of various parameters.

Figure 4 is a log–log plot of variation of the density of infectives I^* as a function of $d_i^c - d_i$ when p, b_s, b_i, d_s and m are held constant. For $m = 0$ (no moves), our model is a two-dimensional probabilistic cellular automaton and, as expected, we found that the critical exponent β_{d_i}, defined by

$$\beta_{d_i} = \lim_{d_i^c - d_i \to 0^+} \frac{\log I^*}{\log(d_i^c - d_i)} \tag{10}$$

is equal to the value of β for two-dimensional directed percolation (Bease 1977).

As shown by Boccara and Cheong (1992, 1993), the motion of the individuals favours the spread of the disease. As far as the phase transition is concerned, compared with the

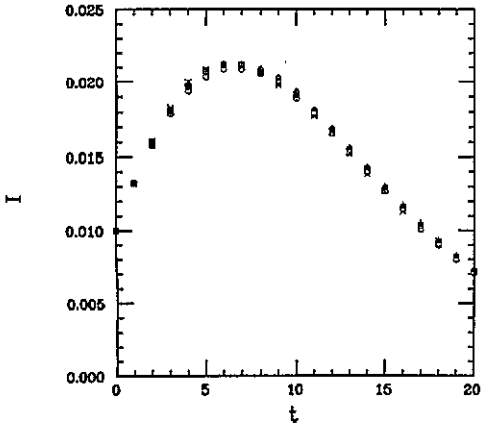

Figure 3. Time evolution of an epidemic for different lattice sizes in the case of short-range moves: $I(0) = 0.01$, $S(0) = 0.59$, $p = 0.3$, $b_s = 0.051$, $b_i = 0.001$, $d_s = 0.2$, $d_i = 0.24$, $m = 5$. Each point represents the average of 20 experiments: ×, $L = 100$; ○, $L = 200$; □, $L = 300$; ◇, $L = 500$; +, $L = 800$.

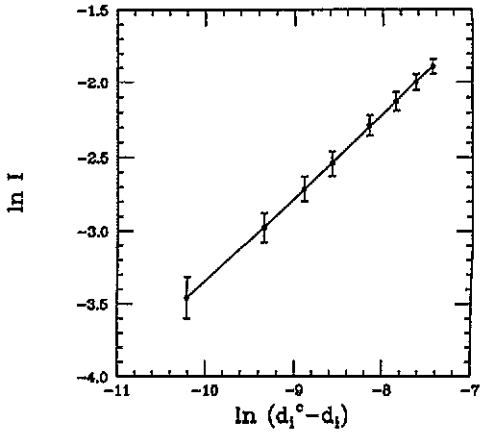

Figure 4. Log–log plot of I^* as a function of $d_i^c - d_i$ for given values of p, b_s, b_i, d_s and m. Here $p = 0.01$, $b_s = 0.4$, $b_i = 0.1$, $d_s = 0.3$ and $m = 0$. The critical value of d_i is 0.011 487 and $\beta_{d_i} = 0.568 \pm 0.05$. Typical error bars are represented. Lattice size is 200 × 200.

control parameters p, b_s, b_i, d_s and d_i, m is not really different in type, except that it changes the range of the interactions and, consequently, the critical exponent β_m, defined by

$$\beta_m = \lim_{m - m_c \to 0^+} \frac{\log I^*}{\log(m - m_c)} \qquad (11)$$

when p, b_s, b_i, d_s and d_i are held constant, will vary with m_c, going from the directed percolation value for small m_c—typically less than 10—to the mean-field value for large m_c—i.e. larger than 100 (Boccara and Cheong 1993). Figure 5 represents a log–log plot of I^* as a function of $m - m_c$. The values of the parameters held constant are such that $m_c = 0.2028$; it is, therefore, not surprising to find $\beta_m = 0.582 \pm 0.05$, in agreement with the directed percolation value.

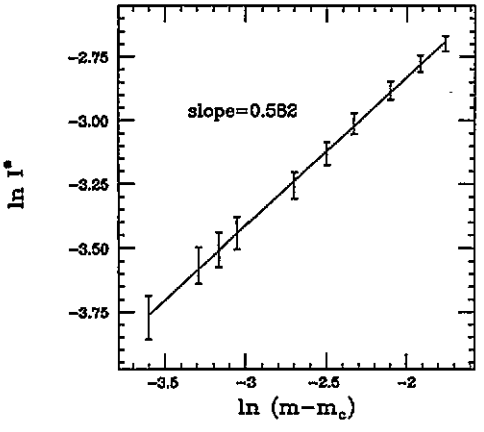

Figure 5. Log–log plot of I^* as a function of $m - m_c$ for given values of p, b_s, b_i, d_s and d_i. Here $p = 0.9$, $b_s = 0.143$, $b_i = 0.0001$, $d_s = 0.001$ and $d_i = 0.15$. The critical value of m is 0.2028 and $\beta_m = 0.582 \pm 0.05$. Typical error bars are represented. Lattice size is 1000 × 1000.

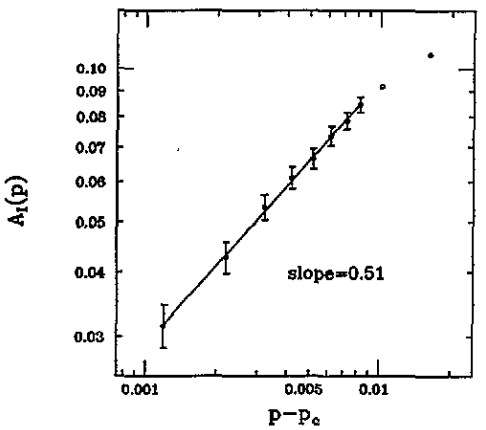

Figure 6. Log–log plot of A_I as a function of $p-p_c$ for given values of b_s, b_i, d_s, d_i and m. Here $b_s = 0.143$, $b_i = 0.0001$, $d_s = 0.001$, $d_i = 0.15$ and $m = 0.2$ for long-range moves. The critical value of p is 0.5338 and $\beta_p^A = 0.51 \pm 0.06$. Typical error bars are represented. Lattice size is 1000×1000.

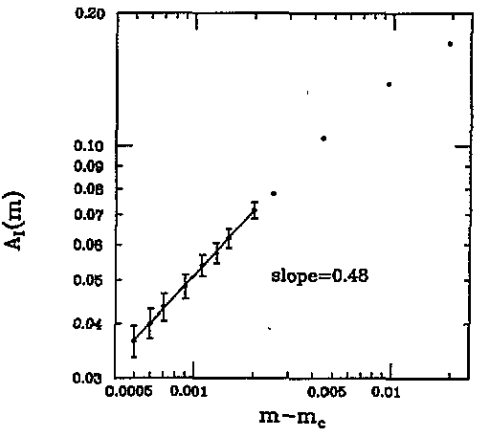

Figure 7. Log–log plot of A_I as a function of $m-m_c$, in the case of long-range moves, for given values of p, b_s, b_i, d_s and d_i. Here $p = 0.9$, $b_s = 0.143$, $b_i = 0.0001$, $d_s = 0.001$, and $d_i = 0.15$. The critical value of m is 0.0005 and $\beta_m^A = 0.48 \pm 0.07$. Typical error bars are represented. Lattice size is 1000×1000.

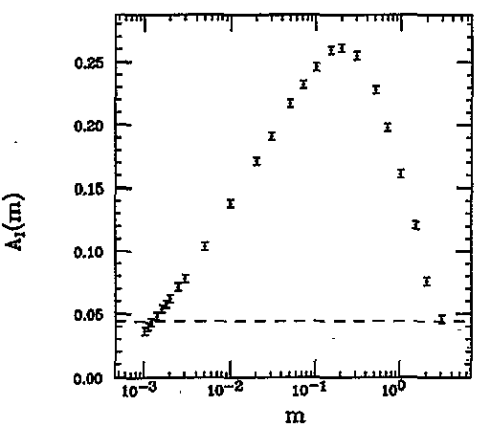

Figure 8. Variation of A_I as a function of m, in the case of long-range moves, for given values of p, b_s, b_i, d_s and d_i. Here $p = 0.9$, $b_s = 0.143$, $b_i = 0.0001$, $d_s = 0.001$, and $d_i = 0.15$. Lattice size is 1000×1000.

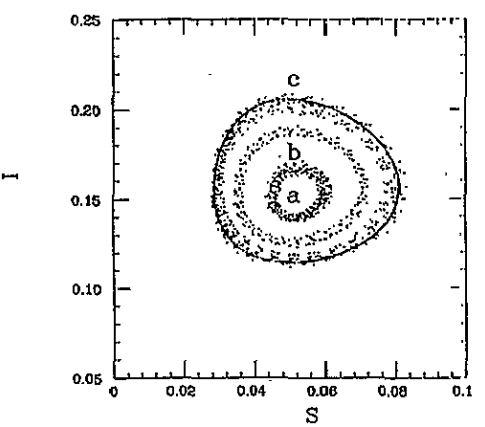

Figure 9. Variation in the size of the cycle as a function of m, in the case of short-range moves, for given values of p, b_s, b_i, d_s and d_i. Here $p = 0.9$, $b_s = 0.143$, $b_i = 0.0001$, $d_s = 0.001$, and $d_i = 0.15$. (a) $m = 20$, $L = 1000$, (b) $m = 50$, $L = 500$, (c) $m = 200$, $L = 200$. The full curve is the mean-field cycle.

3.3. The stable oscillatory behaviour

The most interesting feature of our model is the existence, after a transient depending upon the initial densities of susceptibles and infectives, of oscillatory behaviour for both densities corresponding to a stable limit cycle. This oscillatory behaviour could model, for instance, the periodic recurrence of high densities of infectives as occurs in certain infantile diseases. These oscillations are the manifestation of a collective behaviour that is still the subject of much controversy. In an extensive paper Chaté and Manneville (1992) studied a large

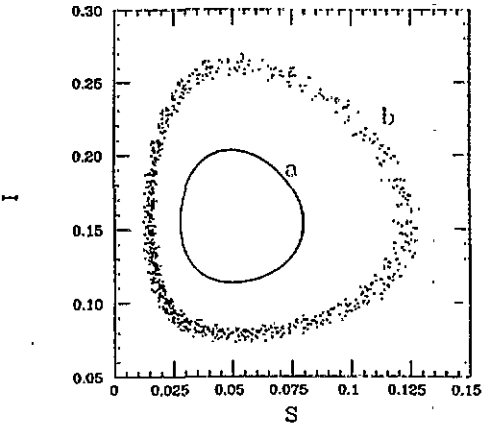

Figure 10. Larger (b) than mean-field (a) cycle obtained for $m = 100$, $L = 200$. Here again $p = 0.9$, $b_s = 0.143$, $b_i = 0.0001$, $d_s = 0.001$, and $d_i = 0.15$.

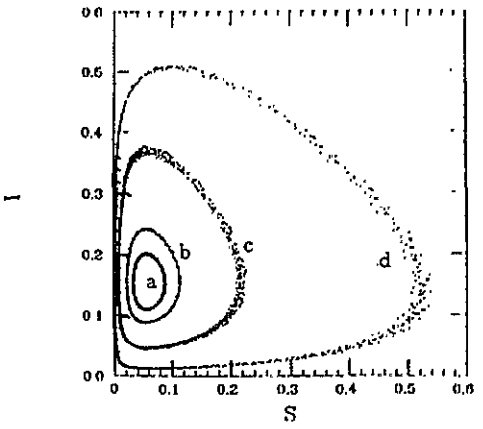

Figure 11. Variation in the size of the cycle as a function of m, in the case of long-range moves, for given values of p, b_s, b_i, d_s and d_i. Here $p = 0.9$, $b_s = 0.143$, $b_i = 0.0001$, $d_s = 0.001$, and $d_i = 0.15$. (a) $m = 2$, $L = 1000$, (b) $m = 1$, $L = 1000$, (c) $m = 0.1$, $L = 1000$. At this scale, the $m = 3$ cycle (the smallest cycle) is indistinguishable from the mean-field cycle.

class of cellular automaton rules for space dimensionalities ranging from two to six. Their results confirm the usual belief that no non-trivial collective behaviour exists for spatial dimensionalities less that four. In our case, oscillatory behaviour has indeed been found for a space dimension equal to two, but this result does not contradict the conclusion of Chaté and Manneville since this collective behaviour has not been found for $m = 0$, i.e. for a standard probabilistic cellular automaton but only for a site-exchange cellular automaton. To observe oscillatory behaviour the m has to be greater than a minimum value—typically 1 for short-ranges moves and 0.001 for long-range ones.

In the vicinity of the Hopf bifurcation, we have studied the amplitude A_I of the oscillating density of infectives. This quantity goes continuously to zero at the bifurcation point. Its behaviour may be characterized by critical exponents similar to the critical exponents defined by (10) and (11). We have determined the two following exponents

$$\beta_p^A = \lim_{p - p_c \to 0^+} \frac{\log A_I}{\log(p - p_c)} \tag{12}$$

$$\beta_m^A = \lim_{m - m_c \to 0^+} \frac{\log A_I}{\log(m - m_c)}. \tag{13}$$

As shown in figures 6 and 7, these two exponents are close to 0.5 as for a standard Hopf bifurcation. These results show once more that there is no fundamental difference between m and the other control parameters.

As a function of m, the amplitude A_I does not vary monotonically (figure 8). The amplitude presents a maximum for a certain value of m close to 0.2 for long-range moves. The largest cycle is not obtained as m goes to ∞, but for a finite non-zero value of m. This phenomenon is also illustrated in figures 9, 10, and 11 which show the influence of motion on the cyclic behaviour. As m decreases from a rather large value ($m = 200$ for short-range

and $m = 3$ for long-range), the size of the cycle, in the (S, I)-plane, first increases then decreases and goes to zero at the bifurcation point.

To show how the correlation is broken as m increases, we have determined analytically the expression for the Hamming distance between two configurations as a function of m for short- and long-range moves (see appendix). These expressions show that, in particular, the mean-field behaviour is approached as $1/m$ for short-range moves and as e^{-m} for long-range moves.

Acknowledgments

One of us (NB) greatly benefited from numerous stimulating discussions with Pablo Ferrari, Roger Bidaux and David Mukamel.

Appendix. Diffusive motion on a lattice as a mixing process

While short-range moves are clearly diffusive moves on a two-dimensional lattice, long-range moves may be viewed as diffusive moves on an infinite-dimensional lattice. To help clarify the mixing process that results from the motion of the individuals, it might be worthwhile to characterize the mixing process as a function of m and the lattice dimensionality d.

Consider a random initial configuration $C(0)$ of random walkers with density c on a d-dimensional torus $\mathbf{Z}_L^d$. Select sequentially mcN walkers at random and move them to a neighbouring site also selected at random if, and only if, the randomly selected neighbouring site is vacant. *At random* means that all possible choices are equally probable. $N = L^d$ is the total number of sites of the torus and m, which is the average number of tentative moves per random walker, is a non-negative real number. Let $C(mcN)$ denote the resulting configuration.

To characterize the mixing process, we shall consider the Hamming distance

$$d_H(C(0), C(mcN)) = \frac{1}{N}\sum_{j=1}^{N}(n(0, j) - n(mcN, j))^2$$

where $n(0, j)$ and $n(mcN, j)$ are, respectively, the occupation numbers of site j in $C(0)$ and $C(mcN)$, as a function of the density c , the parameter m and the space dimensionality d.

If m is very large, $C(0)$ and $C(mcN)$ are decorrelated and the Hamming distance, which is the average value over space of

$$(n(0, j) - n(mcN, j))^2 = n(0, j) + n(mcN, j) - 2n(0, j)n(mcN, j)$$

is equal to $2c(1 - c)$.

Since, in our simulations, the initial configurations are random, only averages over all random walks and all initial configurations with a density c of random walkers are meaningful quantities.

We shall prove that, when N tends to ∞, the average reduced Hamming distance

$$\delta = \frac{d_H(C(0), C(mcN))}{2c(1 - c)}$$

depends on m and d but not on c.

To simplify the notation, consider a one-dimensional torus, and let $P(mcN, j)$ be the probability that the site j is occupied after mcN tentative moves—since only a fraction of these moves are effective. $P(mcN, j)$ is, therefore, the average value $\langle n(mcN, j)\rangle_{\mathrm{rw}}$ over all the possible random walks starting from the same initial configuration. In order to find the evolution equation of $P(mcN, j)$, consider the set of configurations in which $n(mcN, j-1)$, $n(mcN, j)$, and $n(mcN, j+1)$ have fixed values whereas, for all $i \neq j-1, j, j+1$, $n(mcN, i)$ takes any value with the restriction that the total density is equal to c. If k is an integer less than eight the binary representation of which is $x_1x_2x_3$, let $p_k(mcN)$ denote the probability of a configuration such that

$$n(mcN, j-1) = x_1 \qquad n(mcN, j) = x_2 \qquad n(mcN, j+1) = x_3.$$

Then

$$P(mcN+1, j) = \frac{1}{2cN}(p_1(mcN) + p_4(mcN)) + \left(1 - \frac{1}{cN}\right) p_2(mcN) + \left(\frac{1}{2cN}\right)(p_3(mcN) + p_6(mcN)) + \frac{1}{cN} p_5(mcN) + p_7(mcN).$$

Since

$$P(mcN, j-1) = p_4(mcN) + p_5(mcN) + p_6(mcN) + p_7(mcN)$$

$$P(mcN, j) = p_2(mcN) + p_3(mcN) + p_6(mcN) + p_7(mcN)$$

$$P(mcN, j+1) = p_1(mcN) + p_3(mcN) + p_5(mcN) + p_7(mcN)$$

we finally obtain the following discrete diffusion equation

$$P(mcN+1, j) = P(mcN, j) + \frac{1}{2cN}(P(mcN, j+1) + P(mcN, j-1) - 2P(mcN, j)). \tag{A1}$$

In the case of a d-dimensional torus, we would have obtained

$$P(mcN+1, j) = P(mcN, j) + \frac{1}{2dcN}\left(\sum_{i \mathrm{nn} j} P(mcN, i) - 2dP(mcN, j)\right) \tag{A2}$$

Where the summation runs over the $2d$ nearest-neighbours of j. In the case of a complete graph, in which any pair of sites is a neighbouring site, we have

$$P(mcN+1, j) = P(mcN, j) + \frac{1}{c(N-1)N} \sum_{k=1}^{N} (P(mcN, k) - P(mcN, j)). \tag{A3}$$

Equation (A2) may also be written

$$P(mcN+1, j) = \left(1 - \frac{1}{cN}\right) P(mcN, j) + \frac{1}{2dcN} \sum_{i \mathrm{nn} j} P(mcN, i) \tag{A4}$$

showing that the problem of cN random walkers on a d-dimensional torus is equivalent to the problem of one random walker which may move to any single neighbouring site with probability $1/2dcN$ or not move with probability $1 - 1/cN$.

Let us first solve equation (A3) which is much simpler. Since

$$\sum_{k=1}^{N} P(mcN, k) = cN$$

(A3) may be written

$$P(mcN + 1, j) = \left(1 - \frac{1}{c(N-1)}\right) P(mcN, j) + \frac{1}{(N-1)}. \tag{A5}$$

Since (A5) involves only one site—which is typical of mean-field-type equations—its solution is straightforward. We find

$$P(mcN, j) = \left(1 - \frac{1}{c(N-1)}\right)^{mcN} (n(0, j) - c) + c. \tag{A6}$$

Averaging over all random walks starting from the same initial configuration $C(0) = \{n(0, j) | j = 1, 2, \ldots, N\}$, the Hamming distance is

$$d_H(C(0), C(mcN)) = \frac{1}{N} \sum_{j=1}^{N} (n(0, j) + P(mcN, j) - 2n(0, j)P(mcN, j))$$

i.e.

$$d_H(C(0), C(mcN)) = 2c - 2\frac{1}{N} \sum_{j=1}^{N} n(0, j)P(mcN, j).$$

Replacing $P(mcN, j)$ by its expression, we have

$$n(0, j)P(mcN, j) = \left(\left(1 - \frac{1}{c(N-1)}\right)^{mcN} (n(0, j) - c) + c\right) n(0, j)$$

$$= \left(1 - \frac{1}{c(N-1)}\right)^{mcN} (n(0, j))^2 - \left(1 - \frac{1}{c(N-1)}\right)^{mcN} cn(0, j) + cn(0, j)$$

and, taking the average over space, we obtain

$$\langle n(0, j)P(mcN, j)\rangle_{sp} = \left(1 - \frac{1}{c(N-1)}\right)^{mcN} \langle (n(0, j))^2\rangle_{sp} - \left(1 - \frac{1}{c(N-1)}\right)^{mcN} c^2 + c^2$$

where $\langle f(j)\rangle_{sp}$ denotes the average over the space of $f(j)$. Since $\langle (n(0, j))^2\rangle_{sp} = \langle n(0, j)\rangle_{sp} = c$, letting N go to ∞, we finally obtain $(2c - 2c^2)(1 - e^{-m})$, i.e.

$$\delta = 1 - e^{-m}.$$

Consider now equation (A1), and define the Fourier transform of $P(mcN, j)$. We have

$$P(mcN, j) = \frac{1}{\sqrt{N}} \sum_{k=1}^{N} \widehat{P}(mcN, k) e^{2i\pi kj/N}.$$

Replacing this in (A1) yields

$$\widehat{P}(mcN + 1, k) = \left(1 - \frac{2}{cN} \sin^2\left(\frac{\pi k}{N}\right)\right) \widehat{P}(mcN, k)$$

and, therefore,

$$\widehat{P}(mcN, k) = \left(1 - \frac{2}{cN} \sin^2\left(\frac{\pi k}{N}\right)\right)^{mcN} \widehat{P}(0, k).$$

The average Hamming distance over all initial configurations is

$$\frac{1}{N} \sum_{j=1}^{N} \langle (P(0, j) + P(mcN, j) - 2P(0, j)P(mcN, j)) \rangle_{ic}$$

$$= 2c - 2\frac{1}{N} \sum_{k=1}^{N} \langle \widehat{P}(0, k)\widehat{P}(0, -k) \rangle_{ic} \left(1 - \frac{2}{cN} \sin^2\left(\frac{\pi k}{N}\right)\right)^{mcN}$$

where $\langle f(j) \rangle_{ic}$ denotes the average over all initial configurations of $f(j)$. Since

$$\langle \widehat{P}(0, k)\widehat{P}(0, -k) \rangle_{ic} = \frac{1}{N} \sum_{j_1=1}^{N} \sum_{j_2=1}^{N} \langle P(0, j_1)P(mcN, j_2) \rangle_{ic} e^{2i\pi k(j_1 - j_2)/N}$$

we find

$$\frac{1}{N} \sum_{k=1}^{N} \langle P(0, k)P(0, -k) \rangle_{ic} \left(1 - \frac{2}{cN} \sin^2\left(\frac{\pi k}{N}\right)\right)^{mcN}$$

$$= \frac{1}{N^2}(Nc + N(N-1)c^2) + (c - c^2)\frac{1}{N} \sum_{k=1}^{N-1} \left(1 - \frac{2}{cN} \sin^2\left(\frac{\pi k}{N}\right)\right)^{mcN}.$$

Finally, in the limit $N \to \infty$, the average Hamming distance over all random walks and initial configurations is equal to

$$2c(1 - c)\left(1 - \int_0^1 e^{-2m \sin^2 \pi q} \, dq\right)$$

i.e. the reduced Hamming distance δ is equal to

$$1 - e^{-m} I_0(m)$$

where I_0, which is given by

$$I_0(m) = \frac{1}{\pi} \int_0^{\pi} e^{m \cos u} \, du$$

is the modified Bessel function of the first kind of order zero. Since, for large values of its argument, I_0 behaves as $e^m/\sqrt{2\pi m}$, δ approaches the unity, as m tends to ∞, as $1/\sqrt{m}$.

Equation (A2) could be solved in a similar way and, in particular, we would find that δ approaches the unity, as m tends to ∞, as $1/\sqrt{m^d}$.

For small m, it is easy to verify that, for all d, δ behaves as m.

References

Anderson R M and May R M 1991 *Infectious Diseases of Humans, Dynamics and Control* (Oxford: Oxford University Press)

Bailey N T J 1975 *The Mathematical Theory of Infectious Diseases and its Applications* (London: Charles Griffin)

Bease J 1977 *J. Phys. C: Solid State Phys.* **10** 917–24

Boccara N and Cheong K 1992 *J. Phys. A: Math. Gen.* **25** 1–25

—— 1993 *J. Phys. A: Math. Gen.* **26** 3707

Chaté H and Manneville P 1992 *Prog. Theor. Phys.* **87** 1–60

Goles E and Martínez S 1990 *Neural and Automata Networks. Dynamical Behavior and Applications* (Dordrecht: Kluwer)

Hethcote H W and van den Driessche P 1991 *J. Math. Biol.* **29** 271–87

Kermack W D and McKendrick A G 1927 *Proc. R. Soc.* A **115** 700–21

Waltman P 1974 *Deterministic Threshold Models in the Theory of Epidemics* (Heidelberg: Springer)