ELSEVIER

Future Generation Computer Systems 16 (1999) 259–271

FGCS
FUTURE GENERATION COMPUTER SYSTEMS

An empirical method for modelling and simulating some complex macroscopic phenomena by cellular automata

Salvatore Di Gregorio [a,*], Roberto Serra [b]

[a] *Dip. di Matematica, Università della Calabria, Arcavacata, I-87036 Rende (CS), Italy*
[b] *Centro Ricerche Ambientali Montecatini, I-48023 Marina di Ravenna, Italy*

Accepted 17 March 1999

Abstract

Novel parallel computing models sometime represent a valid alternative to standard differential equation methods in modelling complex phenomena. In particular, Cellular Automata (CA) provide such an alternative approach for some complex natural systems, whose behaviour can be described in terms of local interactions of their constituent parts. This paper illustrates an empirical method applied with interesting results in modelling and simulating some complex macroscopic phenomena.

While classical CA are based upon elementary automata, with few states and a simple transition function, in order to deal with macroscopic phenomena it is often necessary to allow a large number of different states a more complicated transition.

The notion of substate is introduced in the macroscopic case for decomposing the state of the cell. The values associated to substates can change in time either due to interactions among substates inside the cell (internal transformations) or to local interactions among neighbouring cells.

The internal transformations are treated in a way similar to ordinary difference equations.

The local interactions among cells can be often treated according to an algorithm for the minimisation of differences, which describes a tendency of conserved quantities to reach an equilibrium distribution.

A large class of complex macroscopic phenomena seem to satisfy the applicability conditions of such an empirical method; some of them are briefly reviewed.

Keywords: Cellular automata; Complex systems simulation; Modelling methods; Fluid-dynamics; Parallel computing.

1. Introduction

While in the past it was possible to investigate the behaviour of many complex phenomena only in particular cases, the developments in computing have allowed to enlarge the set of cases which can be carefully simulated. In particular, parallel computing is particularly interesting either for reasons of absolute performance or for reasons of cost/performance ratio; extending the range of its effective applicability is therefore an important task. In this paper we will discuss some cases where a model well suited for parallel computing, namely that of Cellular Automata (CA), represents an interesting alternative to classical differential equations methods for modelling complex natural phenomena [1].

CA describe systems whose global evolution may be exclusively described on the basis of local

* Corresponding author.
E-mail addresses: toti.dig@unical.it (S. Di Gregorio), rserra@cramont.it (R. Serra)

0167-739X/99/$ – see front matter
PII: S0167-739X(99)00051-5

interactions of their constituent parts [2,3] (Petitot [4] calls this property acentrism). Their computational universality [5] as well as many aspects of modelling have been widely investigated from a theoretical viewpoint [6,7]. Furthermore, CA are easily and naturally implementable on parallel computers and effectively exploit the power of parallelism.

Associating parallelism with acentrism, CA were one of the first Parallel Computing models. Conceived in the 1950's to investigate self-reproduction [8], CA have been used mainly for studying parallel computing methods and the formal properties of model systems [2]. However, during the 1980's, it has been shown that CA are effective in modelling physical systems, in particular fluid dynamics [9,10].

Our research group focused its attention on some macroscopic complex phenomena (lava flows [11,12], landslides [13,14], contamination of soils [15–17]), which can be modelled by CA which have a larger number of states in comparison with the most widely used CA models, like, e.g., lattice gases (a feature which is similar in some respects to lattice-Boltzmann models [18]); an empirical method emerged in those works and a parallel computing environment was developed for this type of applications [19,20].

Section 2 of this paper recalls the classical definition of homogeneous CA, while the Section 3 describes the main features of the method, highlighting the necessary (although not always sufficient) conditions for its application. Section 4 describes the algorithm for the minimisation of the differences among neighbouring cells, while the Section 5 briefly reviews some applications of the proposed approach to real world phenomena. The final section, as usual, is devoted to comments and conclusions.

2. Homogeneous cellular automata

We introduce here a definition which refers to the main class of CA, i.e., the homogeneous CA, which are widely used in modelling and simulation. Extensions need sometimes to be introduced, e.g., in order to deal with the boundaries of finite physical systems; a discussion can be found in [16].

A homogeneous CA can be intuitively considered as a d-dimensional Euclidean space, the cellular space, partitioned into cells of uniform size (e.g., with a square, cubic, hypercubic tessellation according to the dimension); to each cell an identical finite automaton, the so-called elementary automaton (EA), is associated. Input for each EA is given by the states of the EA of the neighbouring cells, where neighbourhood conditions are determined by a pattern invariant in time and constant over the cells. At the time t=0, EA are in arbitrary states and the CA evolves changing the state of all EA simultaneously at discrete times, according to the transition function of the EA.

Formally a CA $\boldsymbol{A}$ is a quadruple $\boldsymbol{A} = \langle Z^d, X, S, \sigma \rangle$, where

- Z^d is the set of cells identified by the points with integer co-ordinates in a d-dimensional Euclidean space;
- X, the neighbourhood index, is a finite set of d-dimensional vectors, which defines the set $N(X, i)$ of neighbours of cell $i =< i_1, i_2, \ldots\ldots i_d >$ as follows: let $X = \{\xi_0, \xi_1, \ldots\ldots \xi_{m-1}\}$ with $m = \#X$; then $N(X, i) = \{i + \xi_0, i + \xi_1, \ldots\ldots i + \xi_{m-1}\}$; ξ_0 is always the null vector;
- S is the finite set of states of the EA;
- $\sigma : S^m \to S$ is the deterministic transition function of the EA;

$C = \{c | c : Z^d \to S\}$ is the set of possible state assignments to $\boldsymbol{A}$ and will be called the set of configurations; $c(i)$ is the state of cell i. Let $c(N(X, i))$ be the ordered set of states of the neighbourhood of i. Then the global transition function τ is defined by

$$\tau : C \to C \mapsto [\tau(c)](i) = \sigma(c(N(X, i))).$$

It is to be noticed that when appropriate, the previous definition may be easily extended to different tessellations, e.g., hexagonal, triangular tessellation in a two-dimensional space, which can be easily reduced to the square one.

3. Internal transformations and local interactions

When natural macroscopic phenomena are considered, the cell corresponds usually to a portion of space: it is therefore natural to assume that the cellular space is at most three-dimensional, as in some cases symmetry or homogeneity considerations allow a reduction to two or one dimensions.

Although CA are inspired by acentrism, two global parameters must always be defined: the size of the

cell and the clock of the CA transition. Their choice deeply affects the form of the model, as they relate to the choice of a particular space–time scale.

The relevant state variables correspond to substates, and the permitted values of each substate belong to a finite set; continuous quantities might be approximated by discretizing them.

The set of possible states of the cell is given by the Cartesian product of the sets of substates, so $Q = Q1 \times Q_2 \times \ldots\ldots \times Q_n$, when n substates are considered. It could be useful in some cases to resort to a higher hierarchy, further dividing a substate in sub-substates.

The values of the substates are constant in the space occupied by the cell.

The change of each substate in the cell can be caused by internal transformations $T_1, T_2, \ldots T_p$ and/or interactions with the neighbourhood $I_1, I_2, \ldots, I_q$.

Internal transformations are defined as the changes in the values of the substates only due to interactions among substates inside the cell (e.g., the cooling of the lava inside the cell due to the substate 'lava temperature' and to the substate 'lava thickness' in the modelling and simulating lava flows) or due simply to the elapsing of the time (e.g., water loss by evaporation). That is, internal transformation are those which would take place if all the cells were independent of each other.

An advantage of internal transformations is that, due to the regular geometry of the cell, complicated evolution rules may often be simplified to manageable expressions.

A function $\sigma_{Ti} : S_{Ti1} \rightarrow S_{Ti2}$ is defined for each internal transition $T_i, 1 \leq i \leq p$, where $S_{Ti1}, S_{Ti2} \in \mathcal{P}(Q)$, which is the power set of Q.

The interactions with the cells belonging to the neighbourhood (local interactions) are described in terms of flows; in many interesting cases, this leads to the flow of some conserved quantity in the central cell towards or from its neighbours in order to reach equilibrium conditions, with an appropriate relaxation rate. Although this case can look very restrictive, in many cases the problem can be reduced to the minimisation of the differences of a certain quantity in the neighbourhood. The next section illustrates the algorithm adopted for the minimisation of the differences.

When this law is applied, the extent of non equilibrium conditions will be reduced. At each time step, the relaxation rate and the size of the CA time step must be properly chosen in order to describe the actual relaxations kinetics of the macroscopic system.

A function $\sigma_{Ij} : S_{Ij1^m} \rightarrow S_{Ij2}$ is defined for each interaction $I_j, 1 \leq j \leq q$, where $S_{Ij1}, S_{Ij2} \in \mathcal{P}(Q)$, the correspondent relaxation rates r_j are also fixed.

It is assumed here that the whole phenomenon can be described by sequentially calculating internal transformations and local interactions. This assumption is critic as it cannot generally be justified a priori, and it should be tested in the model validation phase; it often turns out to be justified only in a specific range of conditions.

Another solution would be completely parallel: starting from the values of the substates at the time t; computing for each substate Q_k the variations $\Delta k_1, \Delta k_2, \ldots \Delta k_p$, determined by the transformations and the flows $fk_1, fk_2, \ldots fk_q$, determined by the interactions; adding all the variations in order to obtain the new value of Q_k at the time $t + 1$. This last solution is depending on the values of the relaxation rates; inappropriate relaxation rates can give rise to big variations and subsequent inconsistency results, when the outflows are larger than the cell content.

4. The algorithm for the minimisation of the differences

An outline of the algorithm is given here in a simple version; the input data are the values at step t of a conserved variable q, which can be neither created nor destroyed. Let us focus for simplicity on a single cell of the automaton: $q[i], 1 \leq i \leq m - 1$ is the content of the ith cell of the neighbourhood. We will suppose that the total content of the central cell is the sum of two terms $p + q[0]$: p is the amount which can be distributed to the neighbouring cells; $q[0]$ is the content of the central cell which cannot be distributed (this distinction is necessary in some cases for physical reasons: for example, water flow in a porous medium can take place only if the water content exceeds a certain threshold; of course, if such a distinction is not required by the problem, then $q[0] = 0$). The flow from the central cell to the ith neighbouring cell will be denoted by $f[i], 0 \leq i < m$ where $f[0]$ is the part

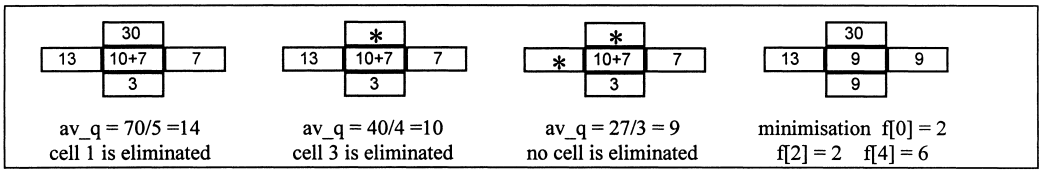

Fig. 1. Example of distribution.

of p which is not distributed. Therefore

$$p = \sum_{i=0}^{m-1} f[i] \tag{1}$$

Let $q'[i] = q[i] + f[i]$, $0 \leq i < m$ be the sum of the content of a neighbouring cell, plus the flow from the central cell, and let $q'\_\min$ be the minimum value for $q'[i]$, $0 \leq i < m$ in the neighbourhood; we will define an algorithm which determines the flows $f[i]$, $0 \leq i < m$, in such a way that the following expression is minimum:

$$\sum_{i=1}^{m-1} (q'[i] - q'\_\min) \tag{2}$$

This expression was chosen because it minimises the local differences (in a discrete acentric context). An alternative choice would be given by substituting $q'[0]$ instead of $q'\_\min$ in Eq. (2), but it can be shown that the application of such rule can give rise in some cases to distributions which do not minimise globally the differences.

The algorithm proceeds as follows.

The average is first computed:

$$\text{average} = \frac{\left(p + \sum_{i=0}^{m-1} q[i]\right)}{m} \tag{3}$$

a cell, whose q is greater than the average does not receive any flow from the central cell; so that cell must be eliminated from the distribution and from the computation of the new average.

The computation is iterated with the remaining cells, calculating the new average and eliminating cells with q greater than the new average; the iterations continue until there are no more cells to be eliminated; then the difference between $q[i]$ (for the not eliminated cell i) and the last calculated average represents the inflow to the neighbour cell i.

Let us consider the following example of a two-dimensional CA where $p = 10, q[0] = 7, q[1] = 30, q[2] = 7, q[3] = 13, q[4] = 3$; the flows $f[i]$ from the central cell are determined (Fig. 1).

4.1. The distribution procedure

The following Pascal-like procedure (Fig. 2) defines completely the algorithm; "eliminated[i]" is a logical variable, whose value is "true" when cell "i" is eliminated; "q_sum" specifies the partial sum of "q" of the cells; "new_control" is a logical variable true, when a new "average" must be calculated; "$f[i]$" is the outflow towards the neighbour cell "i", "count" contains the number of the remaining cells.

4.2. The theorem of the minimisation of the differences

Theorem 1. *The above procedure of distribution minimises the following expression:*

$$\sum_{i=1}^{m-1} (q'[i] - q'\_\min)$$

Proof. The average, computed in the distribution procedure, is such that:

$q[i] \geq \text{average}$ implies $q'[i] = q[i]$ $\quad f[i] = 0$
$q[i] < \text{average}$ implies $q'[i] = \text{average}$
$f[i] = \text{average} - q[i]$
$q'\_\min = \text{average}$

Any alteration of the values of the flows ($f''[i]$) implies that a new minimum value $q''\_\min$ is obtained in the neighbourhood such that $q''\_\min <$ average ($q''\_\min < q'\_\min$) because of Eq. (1) (increasing a flow decreases another one and vice versa). The theorem is demonstrated by Eq. (6), which is deduced from Eqs. (4) and (5). □

```
procedure distribution;.....................
begin
      for  i := 0 to m-1 do
            eliminated[i] := false;
      repeat
            new_control := false;
            q_sum := p;
            count := 0;
            for i := 0 to m-1  do
                  if not eliminated[i] then
                        begin
                              q_sum := q_sum + q[i];
                              count := count + 1
                        end;
            average:= q_sum / count;
            for i := 0 to m-1 do
                  if (q[i] > average) and (not eliminated[i]) then
                        begin
                              new_control := true
                              eliminated[i] := true
                        end
      until not new_control;
      for i := 0 to m-1 do
            if eliminated[i] then
                  f[i] := 0
            else f[i] := average - q[i];
end;
```

Fig. 2. Pascal-like procedure for the distribution algorithm.

$$\sum_{i=1}^{m-1}(q'[i]-q'\_\min)=\sum_{i=1}^{m-1}(q[i]+f[i]-q'\_\min)=\sum_{i=1}^{m-1}(q[i])+p-m(q'\_\min) \tag{4}$$

$$\sum_{i=1}^{m-1}(q''[i]-q''\_\min)=\sum_{i=1}^{m-1}(q[i]+f''[i]-q''\_\min)=\sum_{i=1}^{m-1}(q[i])+p-(q''\_\min) \tag{5}$$

$$\sum_{i=1}^{m-1}(q''[i]-q''\_\min)>\sum_{i=1}^{m-1}(q'[i]-q'\_\min) \tag{6}$$

5. Applications

Three models, which have been developed according to the empirical method previously illustrated, are described below. These models address phenomena of high practical and scientific importance. *Smart Cellular Interactive Automata for modelling the Rheology of Aetnean lava flows* (SCIARA), to be read '*shea'rah*'), *Smart Computational Innovative method for the Detection of Debris/mud flow path with Interactive Cellular Automata* (SCIDDICA), to be read '*she'ddrekah*') and SOIL concern, respectively, (i) Aetnean lava flows, (ii) debris/mud flows with large water content and (iii) bioremediation of phenol-contaminated soils. It is worth stressing that these models are not aimed at describing all the features of the corresponding general phenomena (lava flow, landslides, bioremediation) but are tailored to the specific cases at hand.

5.1. Lava flows

An important goal of the modern geology is the development of models to simulate the evolution of high hazard phenomena. Before the introduction of computational numerical methods, forecasting of lava flows was mainly limited to qualitative aspects, as solving analytically lava flow differential equations for real events is impossible, with the exception of very simple cases.

The use of computational numerical methods for the solution of differential equations governing lava flows is, however, very hard, as the lava may range rheologically from approximately Newtonian liquids to brittle solids.

The main assumptions in SCIARA are the following

- the drop in the temperature is modelled applying the irradiation equation inside the cell
- the effect of the temperature fall is modelled, decreasing the relaxation rate of the flows and introducing the adherence threshold
- the adherence is the lava thickness, below which there is no flow; its dependence upon the temperature is given by a simple inverse exponential relation
- the solidification effect is trivially computed by summing the value of the lava thickness to the altitude and zeroing the lava thickness [11,12].

The SCIARA model was developed for a particular type of lavas, the Aetnean lavas, with good results in comparison with other methods; it was later successfully applied in order to simulate the 1986 eruption of Volcano Piton de la Fournaise in the Reunion Island [21]. It is worth noting that the type of the lava of the Reunion Island is Hawaian, which is different from the basaltic "aa" type of Aetnean lavas.

5.1.1. The SCIARA model for lava flows

The CA model for lava flow is the sextuple

$$\boldsymbol{A}_{\text{sciara}} = \langle R, L, X, S, \sigma, \gamma \rangle$$

where

- $R = \{(x, y) | x, y \in \boldsymbol{N},\ 0 \leq x \leq l_x,\ 0 \leq y \leq l_y\}$ is the set of points with integer coordinates in the finite two-dimensional region, where the phenomenon evolves. N is the set of natural numbers.
- $L \subset R$ specifies the lava source cells, corresponding to the vents.
- The set X identifies the geometrical pattern of cells which influence the cell state change.

  They are, respectively, the cell itself and the "north", "south", "east" and "west" neighbouring cells:

  $$X = \{(0, 0), (0, 1), (0, -1), (1, 0), (-1, 0)\};$$

- The finite set S of states of the EA:

  $$S = S_{\text{a}} \times S_{\text{th}} \times S_{\text{T}} \times S_{\text{f}}^4$$

  where are substates
  S_{a} is related to the altitude of the cell
  S_{th} is related to the thickness of lava in the cell;
  S_{T} is the lava temperature, which effects the rheological resistance
  S_{f} is related to the flows (expressed as thickness) toward the four neighbourhood directions.
- $\sigma : S^5 \rightarrow S$ is the deterministic state transition for the cells in R.
- $\gamma : \boldsymbol{N} \times S_{\text{th}} \rightarrow S_{\text{th}}$ specifies the thickness variation by lava emission for the source cells of L at the step s, $\boldsymbol{N}$ are the steps of the CA.

At the beginning we specify the states of the cells in R, defining the initial configuration of the CA (initial values of substates S_{f} are null). At each step the function γ is applied to the cells in L, then the function σ is applied to all the cells in R.

The main mechanisms of the transition function concern the following internal transformations:
altitude variation by solidification
$\sigma_{T1} : S_a \times S_{\text{th}} \times S_{\text{T}} \rightarrow S_{\text{a}}$;
thickness variation by solidification
$\sigma_{T2} : S_{\text{th}} \times S_{\text{T}} \rightarrow S_{\text{th}}$;
temperature decrease
$\sigma_{T3} : S_{\text{th}} \times S_{\text{T}} \rightarrow S_{\text{T}}$;

and the following local interaction
lava flows
$\sigma_I : (S_{\text{a}} \times S_{\text{th}})^5 \times S_{\text{T}} \rightarrow S_{\text{f}}^4$
(where the algorithm for the minimisation of the differences is applied).

5.1.2. Examples of results concerning the SCIARA applications

In the present example the first part of the 1991–93 Aetnean eruption (which started in December 1991 and ended about the 20 January 1992) was simulated.

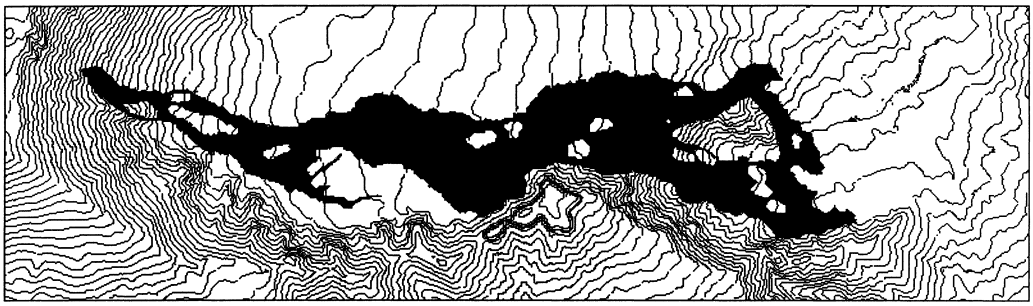

Fig. 3. The first 39 days of the 1991–93 Aetnean eruption.

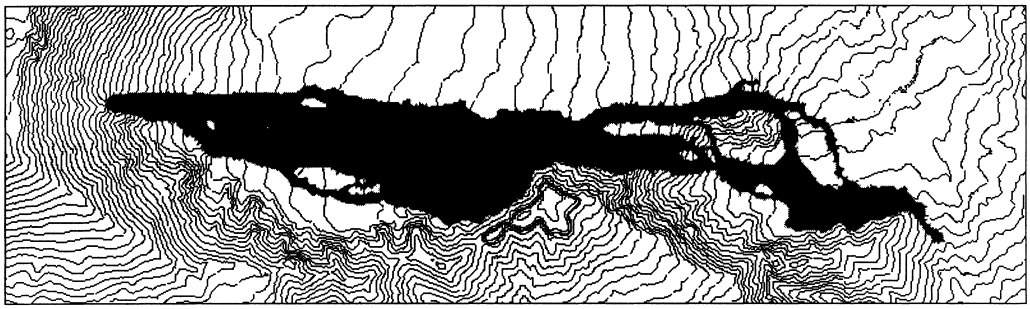

Fig. 4. SCIARA simulation of the first 39 days of the 1991–93 Aetnean eruption.

The eruptive story of that period is paroxystic, characterised by high average values of flow rate (18–25 mc); the simulation regards the first 39 days of the eruption [12].

The comparison between the real event (Fig. 3) and the SCIARA simulation (Fig. 4) can be considered satisfactory concerning geometric characteristics (thickness, area, form) and evolutionary history, despite the approximations in the input data. The comparison between the real and simulated events is slightly less satisfactory regarding the flow path and the surface covered by the flows.

5.2. Debris/mud flows

Some types of landslides, the debris/mud flows, can be considered as an acentric system, therefore they can be a promising CA application field. Their governing flow equations (e.g., Navier–Stokes equations for the debris flow) cannot be easily solved without making substantial simplifications. The problem is complicated by irregular ground topography and because debris/mud flows can change viscosity value by water loss.

SCIDDICA models fast flows, whose potential energy permits to run up obstacles. When the difference of run up height (potential energy) between the central cell and a neighbour is larger than a threshold, then flow is permitted. The threshold depends upon the water content of the debris; debris/mud solidification occurs below a value of water content.

The run up height is determined by the maximum jump of the debris inflows from the neighbours, considering also the loss of energy given by friction [12,14].

5.2.1. The CA model SCIDDICA for debris/mud flows

The CA model for lava flow is the quadruple

$$\boldsymbol{A}_{\text{sciddica}} = \langle R, X, S, \sigma \rangle$$

where

- $R = \{(x, y) | x, y \in \boldsymbol{N},\ 0 \leq x \leq l_x,\ 0 \leq y \leq l_y\}$ is the set of points with integer coordinates in

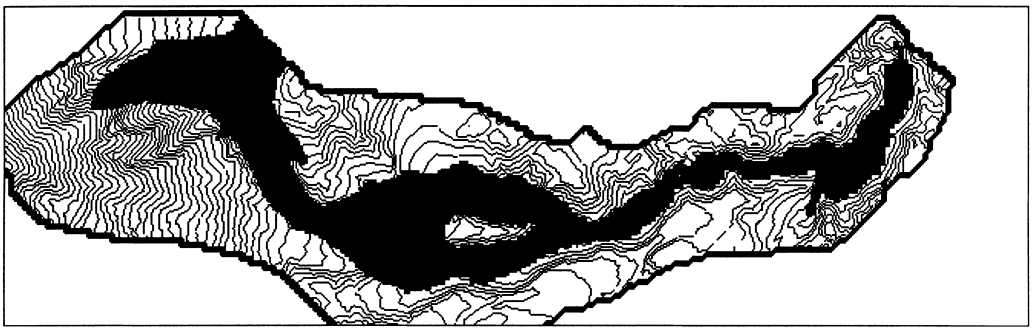

Fig. 5. The Mount Ontake landslide.

the finite two-dimensional region, where the phenomenon evolves. $\boldsymbol{N}$ is the set of natural numbers.

- The set X identifies the geometrical pattern of cells which influence the cell state change.

  They are, respectively, the cell itself and the "north", "south", "east" and "west" neighbouring cells:

  $$X = \{(0, 0), (0, 1), (0, -1), (1, 0), (-1, 0)\};$$

- The finite set S of states of the EA:

  $$S = S_a \times S_{th} \times S_w \times S_r \times S_f^4$$

  where are substates:
  S_a is related to the altitude of the cell
  S_{th} is related to the thickness of debris/mud in the cell
  S_w is the water content, which effects the rheological resistance
  S_r is related to the run up height of the cell debris (potential energy);
  S_f is related to the outflows (expressed as thickness) toward the four neighbourhood directions.
- $\sigma : S^5 \rightarrow S$ is the deterministic state transition for the cells in R.

At the beginning we specify the states of the cells in R, defining the initial configuration of the CA; the initial values of the substates S_{th} are null everywhere, except the area of moving mass, determined by such values; the values of the substates S_a represent the morphology after the detachment; the initial values of the substates S_r are equal to S_{th}. Initial values of substates S_f are zero. At each next step the function σ is applied to all the cells in R, so the time evolution of the CA is obtained.

The main mechanisms of the transition function concern the following internal transformations:
altitude variation by solidification
$\sigma_{T1} : S_a \times S_{th} \times S_w \rightarrow S_a$;
thickness variation by solidification
$\sigma_{T2} : S_{th} \times S_w \rightarrow S_{th}$;
water loss
$\sigma_{T3} : S_{th} \times S_w \rightarrow S_w$;
thickness variation by water loss
$\sigma_{T4} : S_{th} \times S_w \rightarrow S_{th}$;
run up determination
$\sigma_{T5} : S'^4_f \times S^5_r \rightarrow S_r (S'^4_f$ are the inflows);

and the following local interaction:
debris/mud flows
$\sigma_I : (S_a \times S_{th} \times S_r)^5 \times S_w \rightarrow S_f^4$
where the algorithm for the minimisation of the differences is applied.

5.2.2. Examples of results concerning the SCIDDICA applications

The Mount Ontake landslide was chosen as a case study because of the availability of many data [14].

A giant landslide (Fig. 5) was triggered on the slope of Ontake volcano (Japan) by an earthquake in 1984 and moved along the Denjo river at about 20–26 m/s over 9 Km, with a relief of 1625 m. The anomalous extension and speed is explained by the water saturated ground; it justifies that the debris water content is considered constant during the landslide.

The comparison between the real event and the SCIDDICA simulation (Fig. 6) can be considered very satisfactory in the second phase of the landslide, when the debris/mud flow reached the Denjo river.

Fig. 6. The SCIDDICA Mount Ontake landslide simulation.

5.3. Bioremediation application

An interesting application of the above principles has been developed to describe the remediation of contaminated soils by means of biotechnological methods. In particular, the so-called in-situ biostimulation process has been simulated with a CA model.

In-situ biostimulation is based upon the degrading properties of indigeneous microrganisms (recall that 1 g of soil may host perhaps 10^9 bacteria). In the case where the contaminant is an organic compound, or a mixture of organic compounds, it often happens that the indigeneous bacteria, which are already present in the polluted soil, possess the metabolic pathways which are required to degrade the contaminant. However, the process may not take place spontaneously due to the lack of some nutrient (e.g., oxygen), or to inadequate physical and chemical conditions (e.g., pH, moisture content). In these cases, the degrading activity may often be stimulated by providing the appropriate nutrients to the bacteria, directly in the soil.

The advantages of such an approach over conventional ones (like landfill disposal) are striking, both under the environmental and the economical viewpoint, and a very large market is expected to grow in the incoming years.

The remediation originates from the interaction of physical, chemical and biological phenomena, which have been modelled using macroscopic CA [15–17]. In particular, the description of the flow of an aqueous nutrient solution requires that two-phase flow in a porous medium be modelled.

In the two phase case (water and air) the water flow is influenced in a nonlinear way by the water content of the cell (called saturation); below a certain saturation, no flow takes place, while above that threshold water can flow in neighbouring pores. Water is affected both by surface and bulk forces (e.g., gravity): as long as saturation is small, the former prevail. In order to capture the main phenomenological features of this phenomenon, it has been assumed that two different kinds of water exist, the capillary water which is affected only by surface forces and the gravitational water which is affected only by volume forces.

The motion of the capillary water among neighbouring (macroscopic) cells has been described using the minimisation algorithm outlined above; the kinetic coefficient has been assumed to depend upon the water content of the cell, as suggested by data on the so-called relative permeability curves.

The model adjustable parameters have been determined by comparison with experimental data, using genetic algorithms. The agreement between the model and the observed outcomes (data from a pilot plant) has been very good even on flow rates varying by several orders of magnitude (a few examples of such data are given below for illustrative purposes, but see [16] for the data about fluid dynamics, and [17] for data about chemical and biological phenomena).

Further tests are currently under way on a wider set of experimental conditions.

5.3.1. The SOIL model

The CA model for lava flow is the sextuple

$$\boldsymbol{A}_{\text{soil}} = \langle R_3, L, X, S, \sigma, \gamma \rangle$$

where

- $R_3 = \{(x, y, z) | x, y, z \in \boldsymbol{N}, 0 \leq x \leq l_x, 0 \leq y \leq l_y, 0 \leq z \leq l_z\}$ is the set of points with integer coordinates in the three-dimensional finite

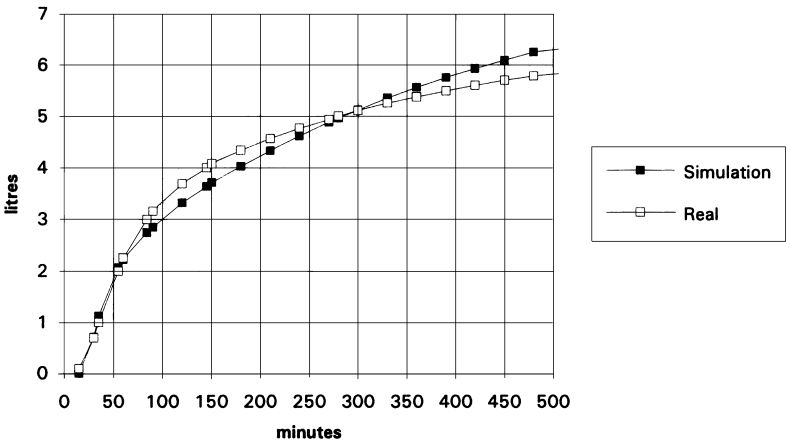

Fig. 7. Total percolated water.

region, where the phenomenon evolves. $\boldsymbol{N}$ is the set of natural numbers.

- $L \subset R_3$ specifies the upper cells, where contaminated water and nutrients are introduced.
- The set X identifies the geometrical pattern of cells which influence the cell state change.

  They are, respectively, the cell itself and the "up", "down", "north", "east", "west", "south" neighbouring cells:

$$X = \{(0,0,0), (0,0,1), (0,0,-1), (1,0,0), (0,-1,0)(0,1,0), (-1,0,0)\};$$

- The finite set S of states of the EA:

$$S = S_{\mathrm{w}} \times S_{\mathrm{ph\_w}} \times S_{\mathrm{ph\_s}} \times S_{\mathrm{f}}^6 \times S_{\mathrm{b}} \times S_{\mathrm{n}}$$

  where are substates:
  S_{w} is the water content of the cell;
  $S_{\mathrm{ph\_w}}$ is the phenol concentration in the cell water;
  $S_{\mathrm{ph\_s}}$ is the phenol concentration in the cell soil;
  S_{f} is related to the water flows toward the six neighbourhood directions.
  S_{b} is the biomass in the cell.
  S_{n} is the nutrient concentration in the cell water;
- $\sigma : S^7 \to S$ is the deterministic state transition for the cells in R_3.
- $\gamma : \boldsymbol{N} \times S_{\mathrm{w}} \times S_{\mathrm{ph\_w}} \times S_{\mathrm{w}} \times S_{\mathrm{ph\_w}} \times S_{\mathrm{n}}$ specifies the water emission with phenol or with nutrients for the source cells of L at the step s, $\boldsymbol{N}$ are the steps of the CA.

At the beginning we specify the states of the cells in R_3, defining the initial configuration of the CA; initial values of substates S_{f} are null. At each next step the function γ is applied to the cells in L, then the function σ is applied to all the cells in R_3.

The main mechanisms of the transition function concern the following internal transformations:
phenol transfer between soil and water
$\sigma_{T1} : S_{\mathrm{w}} \times S_{\mathrm{ph\_w}} \times S_{\mathrm{ph\_s}} \to S_{\mathrm{ph\_w}} \times S_{\mathrm{ph\_s}}$;
degradation of nutrients and phenol by biomass
$\sigma_{T2} : S_{\mathrm{b}} \times S_{\mathrm{n}} \times S_{\mathrm{ph\_s}} \to S_{\mathrm{b}} \times S_{\mathrm{n}} \times S_{\mathrm{ph\_s}}$
and the following local interactions:
water capillary flows
$\sigma_{I1} : S_{\mathrm{w}}^7 \to S_{\mathrm{f}}^6$
(where the algorithm for the minimisation of the differences is applied);
water gravitational flows
$\sigma_{I2} : S_{\mathrm{w}}^2 \to S_f$
(which rules the gravitational water flows from upper cell to the lower one).

5.3.2. Examples of results concerning the SOIL applications

Fig. 7 shows the behavior of the total percolated water during the first 8 h of a contamination experiment, comparing experimental results and model simulation results. The data concerning the long-term behavior (on a two-month period) indicate an even closer agreement, and those which describe the percolated water during the soil remediation phase indicate a good

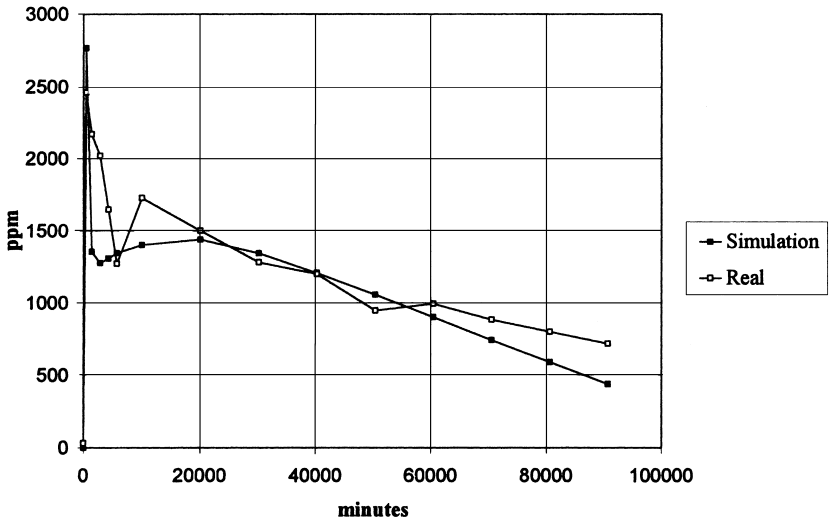

Fig. 8. Phenol concentration in drainage water.

agreement with experimental data, too. Therefore, the model appears capable of describing water percolation on widely separated time scales and slightly different operating conditions.

Fig. 8 shows a comparison (between model and experiment) concerning the concentration of phenol in percolated water, which also indicates a close agreement.

6. Conclusions

The minimisation algorithm is somehow similar to explicit finite difference methods for the integration of partial differential equations, but there is no problem of convergence because each cell cannot distribute flows larger than own content [22].

This empirical method involves, for each internal transformation or local interaction, the determination of relaxation rates, i.e., the introduction of problem-specific parameters, whose determination may be performed by applying optimization methods to minimize the difference between model results and experimental data. Genetic algorithms were effective for applications with several parameters [23,16].

A further consideration can be added: the decomposition of the complex macroscopic phenomenon in internal transformations and local interactions seems to have encouraged interdisciplinary cooperations and exchange of information, at least in the cases outlined above.

Furthermore, such a method permits to start with simple models, whose refinement can be performed in an incremental way, introducing other internal transformations and local interactions. This allows a careful monitoring of the model building phase by comparison between real phenomena and simulations.

The application of the empirical method here described is by no means limited to the examples of lava flow, landslides and bioremediation, but can deal with a much larger set of phenomena, like for examples, atmospheric pollution, pyroclastic flows, soil erosion: the limits to the application of this method have not yet been studied in depth.

Acknowledgements

The authors are grateful to many persons for the common researches, which have inspired this paper; M. Andretta, D. Barca, M. Cannataro, L. Carotenuto, G.M. Crisci, F. Mele, F.P. Nicoletta, R. Rongo, C. Siciliano, M. Sorriso-Valvo, W. Spataro, G. Spezzano, D. Talia, G. Trautteur, M. Villani.

This study has been conducted under the sponsorship of the UE Esprit project 24907, Colombo.

References

[1] T. Toffoli, N. Margolus, Cellular Automata Machines: A New Environment for Modeling, MIT Press, Cambridge, 1987.

[2] A.W. Burks, Essays on Cellular Automata, University of Illinois Press, Champaign. IL, 1970.

[3] S. Wolfram, Computation theory of Cellular Automata, Comm. Math. Phys. 96 (1984) 15–57.

[4] J. Petitot, Centrato/acentrato, in: Encyclopedia Einaudi, vol. 2, Einaudi, Torino, 1977, pp. 894–954 (in Italian).

[5] J.W. Thatcher, Universality in the Von Neumann Cellular model, in: A.W. Burks (Ed.), Essays on Cellular Automata, University of Illinois Press, Champaign. IL, 1970, pp. 103–131.

[6] N. Margolus, T. Toffoli, G. Vichniac, Cellular Automata supercomputers for fluid-dynamics modelling, Phys. Rev. Lett. 56(10) (1986) 1694–1696.

[7] G. Vichniac, Simulating physics with Cellular Automata, Physica D 10 (1984) 96–115.

[8] J. von Neumann, Theory of Self Reproducing Automata, University of Illinois Press, Champaign. IL, 1966.

[9] U. Frisch, B. Hasslacher, Y. Pomeau, Lattice-gas automata for Navier–Stokes equation, Phys. Rev. Lett. 56(14) (1986) 1505–1508.

[10] T. Toffoli, Cellular Automata as an alternative to (rather than an approximation of) differential equations in modeling physics, Physica D 10 (1984) 117–127.

[11] D. Barca, G.M. Crisci, S. Di Gregorio, F.P. Nicoletta, Cellular Automata for simulating lava flows: a method and examples of the Etnean eruptions, Transport Theory and Statistical Physics 23(1–3) (1994) 195–232.

[12] G.M. Crisci, A. Di Francia, S. Di Gregorio, F.P. Nicoletta, R. Rongo, W. Spataro, SCIARA.2: a cellular automata model for lava flow simulation, in: V.P. Glahn (Ed.), Proc. of the IAMG'97, Addendum, Barcelona, 1997, pp. 11–16.

[13] S. Di Gregorio, R. Rongo, F.P. Nicoletta, M. Sorriso-Valvo, W. Spataro, G. Spezzano, D. Talia, Landslide simulation by Cellular Automata in a parallel environment, in: M. Mango Furnari (Ed.), Proc. of 2nd International Workshop on Massive Parallelism: Hardware, Software and Application, World Scientific, Singapore, 1995, pp. 392–407.

[14] S. Di Gregorio, R. Rongo, C. Siciliano, M. Sorriso-Valvo, W. Spataro, Mount Ontake Landslide simulation by the Cellular Automata model SCIDDICA III, Physics and Chemistry of the Earth (A), 24 (2) (1999) 97–100.

[15] S. Di Gregorio, R. Rongo, R. Serra, W. Spataro, M. Villani, Simulation of water flow through a porous soil by a Cellular Automaton model, ACRI '96, Proc. of second Conference on Cellular Automata for Research and Industry, Springer, Berlin, 1997, pp. 79–88.

[16] S. Di Gregorio, R. Serra, M. Villani, Applying cellular automata to complex environmental problems, Theoretical Computer Science 217/1 (1999) 131–156.

[17] S. Di Gregorio, R. Serra, M. Villani, Simulation of soil contamination and bioremediation by a Cellular Automaton model, Complex Systems 11(1) (1998) 31–54.

[18] S. Succi, R. Benzi, F. Higuera, The lattice Boltzmann equation: a new tool for computational fluid dynamics, Physica D 47 (1991) 219–230.

[19] S. Di Gregorio, R. Rongo, W. Spataro, G. Spezzano, D. Talia, A parallel cellular tool for interactive modeling and simulation, IEEE Computational Science & Engineering 3 (1996) 33–43 .

[20] S. Di Gregorio, R. Rongo, W. Spataro, G. Spezzano, D. Talia, High performance scientific computing by a parallel cellular environment, Future Generation Computer Systems 12 (1997) 357–369.

[21] R. Rongo, W. Spataro, N. Villenueve, Lava flow simulation with SCIARA: the Reunion Island case, in: A. Buccianti, G. Nardi, R. Potenza (Eds.), Proc. of IAMG 98 Conference, Litografia Editrice, Naples, 1998, pp. 944–949.

[22] C. Murmura, Automi Cellulari e metodi di discretizzazione per lo studio della conduzione del calore, Laurea Thesis, Department of Mathematics, University of Calabria, 1996 (in Italian).

[23] S. Di Gregorio, R. Serra, M. Villani, Combining cellular automata and genetic search in complex environmental modelling, in: E. Pessa, M.P. Penna, A. Montesano (Eds.), Proc. of 3rd Systems Science European Congress, Ed. Kappa Roma, 1996, pp. 1127–1131.

Salvatore Di Gregorio was born in Castellammare del Golfo, Sicily, Italy, in 1948. He received the 'Laurea' degree in Physics from the University of Palermo in 1972 and continued studies and researches in Computer Science at the University of Naples and of British Columbia (Canada).

In 1978 he joined the University of Calabria with the appointment of the 'Computer Programming' course. Since 1985, he has been with the Department of Mathematics of the same university as associate professor of Foundations of Computer Science.

His research interests include parallel computing, in particular cellular automata and their implementation on parallel architectures, modelling and simulation of complex systems (lava flows, mud/debris flows, freeway traffic, soil contamination and bioremediation, etc.) by cellular automata.

He participated in Italian and European research programs; now he is involved in projects of the European Community, concerning parallel computing and its applications (COLOMBO) and environmental problems (RUNOUT for the landslide risk, FLOW for the volcanic risk).

Roberto Serra born in Bologna (Italy) in 1952, took his degree in physics, with honors, from the University of Bologna in 1977. He has then been research collaborator at Istituto Nazionale di Fisica Nucleare and University of Bologna (1978), application engineer at Hewlett Packard (1979), researcher at Tema (Eni group), where he became leader of the Complex Systems group in 1982. He has later been research area coordinator at Enidata for parallel computation, neural networks and genetic algorithms. After leaving the Eni group, he has been vice-director for R&D at Dida*El (1990) and then joined the Montedison group, where he has been IT general manager (1991–1995). Since 1995 he is the director of the Montecatini Environmental Research Centre in Ravenna.

Roberto Serra has been lecturer at the Universities of Parma (1990) and Milano (1997). Moreover, he has been member of committees of the European Union for designing research programmes and evaluating submitted projects. Dr. Serra co-authored and co-edited several books and has written many papers published on international journals and conference proceedings. He is vice president of *AI*IA* (Associazione Italiana per l' Intelligenza Artificiale) and member of the board of directors of SIREN (Società Italiana Reti neuroniche) and AIRS (Associazione Italiana per la Ricerca sui Sistemi). He is member of ACM, IEEE, New York Academy of Sciences.