Parallel Simulation of Asynchronous Cellular Automata Evolution *

Olga Bandman

Supercomputer Software Department
ICMMG, Siberian Branch Russian Academy of Sciences
Pr. Lavrentieva, 6, Novosibirsk, 630090, Russia
E-mail: bandman@ssd.sscc.ru

Abstract. For simulating physical and chemical processes on molecular level asynchronous cellular automata with probabilistic transition rules are widely used being sometimes referred to as Monte-Carlo methods. The simulation requires huge cellular space and millions of iterative steps for obtaining the CA evolution representing the real scene of the process. This may be achieved by allocating the CA evolution program onto a multiprocessor system. As distinct from the synchronous CAs which is extremely efficient, the asynchronous case of parallel implementation is stiff. To improve the situation we propose a method for approximating asynchronous CA by a superposition of a number of synchronous ones, each being applied to locally separated blocks forming a partition of the cellular array.

1 Introduction

The increase of computing power both of individual computers and of multiprocessor systems enhance the development of simulation methods for obtaining new knowledge about natural and technological processes. Usually, simulation of spatial dynamics in physics is performed by partial differential equations (PDE) solution. But in case when processes under simulation are nonlinear or have discontinuous behavior PDE are impuissant. Bright manifestation of the situation is kinetics of nano-systems, such that epitaxial growth on silicon crystal [1], autovawes and oscillations during the oxidation of carbon monoxide on catalyst surface [2, 3], where the direct modeling of possible movements of particles and their stochastic interactions in a discrete space is used. Due to the stochastic character of the processes the models are sometimes classified as "Random Selection Algorithms of Monte-Carlo methods" [4, 5], actually being asynchronous CA with probabilistic transition rules. It is clear, that very small size of real "particles", i.e. molecules or atoms, stipulate the necessity of huge size of the CA, and real speed of their movements requires large simulation time. Thus, the

* Supported by 1)Presidium of Russian Academy of Sciences, Basic Research Program N 14.15 (2006), 2) Siberian Branch of Russian Academy of Sciences, Integration Project 29 (2006)

capability of simulation is constrained by the performance of modern computers. The situation might be essentially improved by using multiprocessor supercomputers which are available for scientific community nowadays, but as distinct to the synchronous CA parallelization of asynchronous CA evolution is a hard task.

A natural way to achieve acceptable parallelization efficiency is to transform the given asynchronous CA into a synchronous one, which approximate its evolution. The advantages of such an approach is twofold. First, its implementation is faster, because of the decrease of random number generator use. Second, allocation of a CA onto many processors for parallel simulation becomes easier and more efficient. The idea has been already exploited for a particular case of surface reaction simulation [6]. Here we aim at the development of systematic method and present some results of its experimental investigation.

Apart from the Introduction and the Conclusion the paper contains three sections. The second section contains definitions of used concepts. In the third section asynchronous to block-synchronous transformation and its justification are given. The fourth section is dedicated to parallel implementation of of block-synchronous CA.

2 Formal Statement of the Problem

The class of CA under investigation is a mathematical model of the phenomena consisting of elementary actions of *particles.* A particle may be interpreted as a real atom or molecule. Elementary actions are mostly the following: adsorption of particles from the medium (gas) , sublimation, dissociation, diffusion, chemical reaction. The processes are stochastic, probability of each action being conditioned by physical parameters. The class of CA modeling the above processes differs from that of classical cellular automata in the following: 1) transition rules are probabilistic and deal not only with Boolean states, but also with integers and sometimes with symbols, 2) a single transition rule is allowed to update a group of cells at once, being a particular case of *substitution systems*)from [7] or [8], 3) the mode of operation is asynchronous, i.e. each time only one updating act is performed, the cells to be updated being randomly chosen.

A CA with the above features is further referred to as a *kinetic asynchronous CA*, being denoted as CA_α and represented by three concepts $CA = \langle A,M,\theta\rangle$, where A is a *state alphabet*, M – the set of elementary automata names, θ – a transition rule. There is no constraints imposed to the alphabet. As for the set of names, 2D Cartesian lattice $M = \{m : m = i, j, i = 0, \ldots, I; j = 0, \ldots, J\}$, is considered, m being used instead of (i, j) for short. A set $\Omega = \{a, m) : a \in A, m \in M\}$ forms a *cellular arrays*, where a pair (a, m) is called a *cell*, $a \in A$ and $m \in M$. On the set M naming functions $\varphi : M \to M$ may be defined. If $m_k = \varphi(m)$, then m_k is a neighbor of a cell named m.

A subset of cells

$$S(m) = \{(v_0, m), (v_1, \varphi_1(m)), \ldots, (v_q, \varphi_q(m))\} \tag{1}$$

form a *local configuration*, a cell (v_0, m) being referred to as its *reference cell*, the set of cells names in it

$$T(m) = \{m, \varphi_1(m), \dots, \varphi_q(m)\} \tag{2}$$

being called as an *underlying template.* Two local configurations $S(m)$ and $S'(m)$ with the same reference cells represent an elementary act of cellular array updating,

$$\theta(m) : S(m) \to S'(m), \quad T'(m) \subseteq T(m), \tag{3}$$

where $S'(m) = \{(u_k, \varphi_k(m)) : k = 0, 1, \dots, p, p \leq q$, is a next-state local configuration, whose cell states $u_k = f_k(v_0, v_1, \dots, v_q), \quad k = 0, 1, \dots, p$, are transition functions values.

Application of $\theta(m)$ to all cells of Ω_t transfers the cellular array into the next *global state* Ω_{t+1}, which is considered as an iteration. In CA$_\alpha$ this transition may be represented as a *transient sequence*

$$\sigma_\alpha(\Omega_t) = (\Omega_t, \dots, \Omega_{t+l\tau}, \dots, \Omega_{t+\mu\tau}), \tag{4}$$

where τ is a micro-step for one updating, and $\Omega_{t+\mu\tau} = \Omega_{t+1}$, $\mu = |M|$. All possible transient evolutions starting from Ω_t and ending at Ω_{t+1} constitute an ensemble $\gamma_\alpha(\Omega)$, whose cardinality is $|\gamma_\alpha(\Omega)| = \mu!$. The sequence $\Sigma(\Omega_0) = (\Omega_0, \dots, \Omega_t, \dots, \Omega_T)$ is referred to as an *evolution*, the set of all possible evolutions of a CA$_\alpha$ starting from an Ω is denoted as $\Gamma_{\mathtt{CA}_\alpha}(\Omega)$.

3 Approximation of an Asynchronous Ca by a Block-Synchronous One

Since synchronous CA are preferable for parallel implementation, there is a natural intention to transform a given CA$_\alpha$ into a synchronous one preserving the evolution of CA$_\alpha$. Unfortunately, there is no exact method known by now how to do this, hence, we make an attempt to obtain an approximate one. The idea used is to impose some order on the random choice of cells to be updated, making this in such a way as to bring no distortion in the evolution progress but only restricting the possible choice of state-transition sequences. Moreover, introducing synchronicity, one should be cautious for conservation behavioral correctness. It is most important because of the fact that in CA$_\alpha$ multicell updating is used, i.e. some cells are updated at once. The correctness condition (in [8] referred to as *noncontradictoryness*) requires that no two simultaneous acts of updating change the same cell state at the same time. Formally, the sufficient correctness condition is as follows.

$$T'(m) \cap T'(\varphi_l(m)) = \emptyset \quad \forall m \in M, \quad \forall l \in 1, \dots, q, \tag{5}$$

where $T'(m)$ and $T'(\varphi_l(m)$ are underlying templates for $S'(m)$ and $S'(\varphi_l(m))$ in (3). It is clear that CA$_\alpha$ are always correct because only a single (although a multicell one) is allowed at a time.

We shall say that a CA$_\beta$ approximates a CA$_\alpha$ if

$$\mathit{\Gamma}_\beta(\mathit{\Omega}) \subseteq \mathit{\Gamma}_\alpha(\mathit{\Omega}) \quad \forall \mathit{\Omega} \in A \times M. \tag{6}$$

and construct the approximation in the form of a block-synchronous CA (further denoted as CA$_\beta$) which operates as follows.

1. On $\mathit{\Omega}$ a set of partitions $\mathit{\Pi} = \{\mathit{\Pi}_1, \ldots, \mathit{\Pi}_k, \ldots, \mathit{\Pi}_b\}$ is defined as follows:

$$\mathit{\Pi}_k = \{B_k^1, \ldots, B_k^g, \ldots, B_k^G\}, \quad \bigcup_g B_k^g = \mathit{\Omega}, \quad \bigcap_g B_k^g = \emptyset, \quad G = |M|/b. \tag{7}$$

B_k^g having the underlying template

$$T_B(m_k) = \{m_k, \psi_1(m_k), \ldots, \psi_l(m_k), \ldots, \psi_b(m_k)\} \tag{8}$$

m_k being a reference cell name of a block $B_k^g \in \mathit{\Pi}_k$.

2. A transition $\mathit{\Omega}_t \to \mathit{\Omega}_{t+1}$ is divided into b steps, the resulting arrays forming a sequence:

$$\sigma_\beta(t) = (\mathit{\Omega}_t, \mathit{\Omega}_{t+t'}, \ldots, \mathit{\Omega}_{t+t'k}, \ldots, \mathit{\Omega}_{t+t'b}), \quad t' = \frac{t}{b}, \tag{9}$$

where on k-th step, $k = 1, \ldots, b$, $\theta(m)$ is applied synchronously to reference cells (v_k, m_k) of all blocks $B_k^g \in \mathit{\Pi}_k$, $g = 1, \ldots, G$.

3. Partitions $\mathit{\Pi}_k \in \mathit{\Pi}$ are processed in a random order, the ensemble γ_β of transient sequences in the transitions $\mathit{\Omega}_{t+t'k} \to \mathit{\Omega}_{t+t'(k+1)}$ having the cardinality $|\gamma_\beta| = b!$.

Theorem. A CA$_\beta = \langle A, M, \theta \rangle$ is an approximation of an AC$_\alpha$, if

$$T'(m) \subseteq T_B(m), \tag{10}$$

where $T'(m), T_B(m)$ are underlying templates of $\theta(m)$ and B_k^g, respectively.

Proof. To prove the Theorem it is sufficient to show that the relation $\gamma_\beta(\mathit{\Omega}) \subseteq \gamma_\alpha(\mathit{\Omega})$ holds for each iteration $\mathit{\Omega}_t \to \mathit{\Omega}_{t+1}$. The latter, according to (9), may be represented as a sequence of synchronous transitions $\mathit{\Omega}_{t+t'k} \to \mathit{\Omega}_{t+t'(k+1)}$ which also belong to the set of cellular arrays included in $\sigma_\alpha(\mathit{\Omega}(t))$ (4). It follows from two facts: 1) condition (10) of the Theorem provides the correctness condition (5) of the synchronous step , and 2) property (7) of CA$_\beta$ ensures that the result does not depend on the mode of operation. Moreover, in the sequence of synchronous steps the portion of next-state values used as arguments in functions (3), being equal to $\frac{\mu k}{q}$, increases with k in the similar way than it takes place in asynchronous case. So, the whole iteration result is equal to the result of an asynchronous iterative step, which proves the Theorem.

Taking into account the approximation concept (6), the approximation accuracy may be assessed only as the relation between the numbers Q_β of transient sequences $\sigma_\alpha(\mathit{\Omega}(t))$ encapsulated in a transition $\mathit{\Omega}(t) \to \mathit{\Omega}(t+1)$ of CA$_\beta$, and the total numbers of transient sequences in $\gamma_\alpha(\mathit{\Omega})$, which yields $\varepsilon = \frac{G!b!}{\mu!}$ How serious is the discrepancy from the true process under simulation depends on many factors and may be clarified only by a comprehensive experimental study.

Example 2. The most simple model of epitaxial growth on Silicon (Si) surface is a composition of two following actions: 1) absorption of Si-atoms from the gas with the probability p_a; 2) diffusion of the absorbed atoms over the surface. An atom diffuses to a neighboring cell if it has $n > 0$ neighbors $(n = 1, 2, 3, 4)$, whose states is less that that of its own. The probability of the diffusion act is $p' = 0,05^{4-n}$, and the choice among n possible directions to move to is equiprobable, so $p_d = p'/n$. The process may be described by an CA$_\alpha$= $\langle A, M, \theta\rangle$ where $A = \mathbf{N}$, $M = \{(i, j) : i = 0, \dots, I, j = 0, \dots, J\}$. A cell $(a, (i, j))$ corresponds to a site on a Si crystal surface, where the thickness of the adsorbed layer is equal to a atoms. The transition rule $\theta(i, j)$ is a superposition of ϑ_{ads} responsible for absorbtion, and $\vartheta_{diff}(i, j)))$ responsible for diffusion.

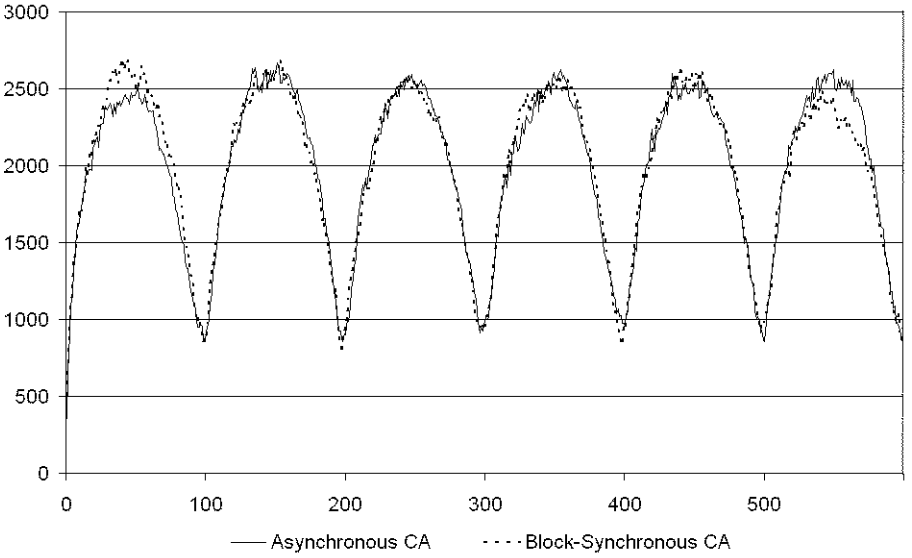

Fig. 1. The dependence $P(t')$ for AC$_\alpha$ simulating epitaxial growth and its approximation by a CA$_\beta$, $|M| = 200 \times 200$, $p_a = 0.2$, $t' = t/500$

$$\begin{aligned}\vartheta_{ads} &= \{(v_0, (i, j))\} \overset{p_a}{\to} \{(v_0 + 1, (i, j))\};\\ \vartheta_{diff} &= \{(v_0, (i, j)), (v_1, \varphi_1(i, j)), (v_2, \varphi_2(i, j)), (v_3, \varphi_3(i, j)),\\ &\quad (v_4, \varphi_4(i, j))\} \overset{p_d}{\to} \{(u_0, (i, j)), (u_1, \varphi_1(i, j)), (u_2, \varphi_2(i, j)),\\ &\quad (u_3, \varphi_3(i, j)), (u_4, \varphi_4(i, j))\},\end{aligned} \tag{11}$$

where

$$u_0 = \begin{cases} v_0 & if\ (\forall k : u_k \geq v_0)\ \wedge\ (rand > p_n),\\ v_k + 1 & if\ (u_k < v_0)\ \wedge\ (rand < p_n)\ \wedge\ (1/n \leq k/n \leq (k + 1)/n).\end{cases} \tag{12}$$

rand being a random number in the interval [0,1].

The simulation process shows the formation of islands of adsorbed atoms on the Si surface. One of the features under investigation is the dependence of total perimeter $P(t)$ of the islands on time. The perimeter P is computed as a number of cell pairs having different states. During the process this number exhibits oscillations, which are of interest for the researcher. In Fig.1 a few first waves produced both by CA$_\alpha$ and CA$_\beta$ of such oscillations are shown, CA$_\beta$ being obtained according to the above method with the block-size 3 × 3. The mean square error of approximation computed according the above experiment data is $E = 0.0412$. Moreover, CA$_\beta$ simulation is 1.5 times faster, than CA$_\alpha$.

4 Parallel Implementation of CA$_\alpha$

When simulating spatial dynamics on N processors the simulation space is divided into N parts, each allocated and processed in its own processor, the processors exchanging data each iteration. Time for transmitting a data package $T_{trans} = T_{lat} + VT_{bit}$, where T_{lat} is latency time, V-the amount of bits in the package, and T_{bit} - bit transmission time. From the relation $T_{lat} \gg T_{bit}$ it follows, that the exchange efficiency depends directly on package size. That is why parallel implementation of synchronous CA evolution, where all border cells states may be packed in one package, is extremely efficient, the speedup being close to N provided that

$$T_{computation} \gg T_{transmission} \tag{13}$$

Unfortunately, CA$_\alpha$ parallelization allows no package to be formed, because any delay in cell state transmission breaks the correctness condition (5). So, each state change on the border of the array allocated in a processor requires an exchange to be performed. It leads to the slow down instead of speeding up, because (13) cannot be reached even with very large arrays.

The situation is quite different for CA$_\beta$, because, border cells states of the results of synchronous steps may be transmitted in a package. The experimental results (Table 1), obtained by running the CA$_\beta$ simulation program on the cluster MVS-1000/128 of Siberian Supercomputer Center, show quite acceptable speedup for large enough cellular array size.

Table 1. Time T (min), speedup $S = T_N/T_1$, and efficiency $C = T_1/(NT_N)$ in performing 10^4 iterations of CA$_\beta$ simulation from the Example 1, array size being 6000 × 6000.

N	1	4	9	16	25
T	817.08	212.91	113.66	55.52	41.65
S	1	3.83	7.18	14.78	19.64
C	1	0.95	0.79	0.92	0.78

5 Conclusion

The problem of parallel simulation of kinetic asynchronous CA evolution is considered. It is shown that to make parallel implementation speedup acceptable it is necessary to approximate it by a a block-synchronous CA. An algorithm for constructing a block-synchronous approximation is given and approximation error is assessed. Experimental results are presented which show the approximation error to be admissible for probabilistic algorithms, and speedup of parallel implementation quite acceptable.

References

1. Neizvestny,I.G., Shwartz, N.L., Yanovitskaya, Z.Sh., and Zverev A.V. 3D-model of epitaxial growth on porous {111} and {100} Si surfaces. *Computer Physics Communications,* vol.147, pp.272-275, 2002.
2. Ziff, R.M., Gulari, E., Bershad Y. Kinetic phase transitions in irreversible surface-reaction model. *Physical Review Letters*, vol.56, pp.2553-2558, 1986.
3. Choppard B., Droz M. Cellular automata approach to nonequilibrium phase transition in a surface reaction model; static and and dynamic propereties. *Journ.of Physics.* ***A***. *Mathenatical and General*, vol.21, pp.205-211, 1988.
4. Elokhin V.I., Latkin E.I., MatveevA.V., and Gorodetskii V.V. Application of Statistical Lattice Models to the Analysis of Oscillatory and Autowave Processes on the Reaction of Carbon Monoxide Oxidation over Platinum and Palladium Surfaces.*Kinetics and Catalysis,* vol.44, N5, pp.672-700, 2003.
5. Makeev A.G. Coarse bifurcation analysis of kinetic Monte Carlo simulations: a lattice-gas model with lateral interactions. *Journ of chemical physics*, vol 117, N18, pp.8229-8240, 2002.
6. Nedea S.V., Lukkien J.J., Jansen A.P.J., and Hilbers P.A.J. Methods for parallel simulations of surface reactions. *arXiv:physics/0209017*, vol.1, N4, pp.1-8, 2002.
7. Wolfram S., A New Kind of Science. - Wolfram Media Inc., Champain, Ill., USA. 2002.
8. Achasova S., Bandman O., Markova V., Piskunov S. Parallel Substitution Algorithm. Theory and Application. - World Scientific, Singapore, 1994.