Complex Systems **2** (1988) 501–507

Decreasing Energy Functions and Lengths of Transients for Some Cellular Automata

Eric Goles*
Department of Mathematics, Engineering School, University of Chile, Casilla 170/3, Correo 3, Santiago, Chile

Andrew M. Odlyzko
AT&T Bell Laboratories, Murray Hill, NJ 07974, U.S.A.

Abstract. The work of Ghiglia, Mastin, and Romero on a "phase-unwrapping" algorithm gives rise to the following operation: for any undirected graph with arbitrary integer values attached to the vertices, simultaneous updates are performed on these values, with the value of a vertex being changed by one in the direction of the average of the values of the adjacent vertices. (When the average equals the value of a vertex, the value of the vertex is incremented by one, unless all the neighbors have the same value, in which case no change is made.) Earlier work of Odlyzko and Randall showed that iterating this operation always leads to a cycle of length one or two, but did not give a bound on how many iterations might be needed to reach such a cycle. This paper introduces a new "energy function" which does yield a bound for the transient. A novel feature of this energy is that it contains not only linear and bilinear terms, as is common, but also terms involving the minimum function.

1. Introduction

Let G be an undirected graph with vertices labelled $1, \ldots, n$, and suppose that for each i, an integer $x_i(0)$ is assigned at time 0 to vertex i. A series of synchronous updates is performed on these values with the rule that if $x_i(t)$ is the value of vertex i at time t, then

$$x_i(t+1) = \begin{cases} x_i(t) - 1 & \text{if } S_i(t) < 0, \\ x_i(t) & \text{if } x_j(t) = x_i(t) \text{ for all } j \in J_i, \\ x_i(t) + 1 & \text{otherwise,} \end{cases} \tag{1.1}$$

*Supported partially by AT&T Bell Laboratories and Fondo Nacional de Ciencias de Chile 86-87, Santiago, Chile.

where

$$\begin{aligned} S_i(t) &= \sum_{j\in J_i} x_j(t) - d_i x_i(t), \\ J_i &= \{j : \text{ vertex } j \text{ is connected to vertex } i\}, \\ d_i &= |J_i| = \text{ degree of vertex } i. \end{aligned} \tag{1.2}$$

This transformation was defined by Ghiglia, Mastin, and Romero [1] for certain graphs that arose in their work on "phase unwrapping" (determination of the phase of an analytic function from a set of values of the function at a discrete grid). Since the maximal value assigned to any of the vertices never increases, and the minimal value never decreases, it is clear that iterating the transformation leads to a configuration that repeats periodically. Ghiglia et al. observed that for their graphs, the transformation always ended up in a cycle of length 1 or 2. Their algorithm was in fact designed on the assumption that the cycle length is at most 2. Later, Brickell and Purtill (unpublished) defined the transformation for general graphs and conjectured that the period was always at most 2. This conjecture was proved in [6], by utilizing an "energy function" derived from that used in [4]. However, the proof of [6] did not provide an upper bound for the transient (the number of steps until the transformation enters a cycle), since the energy function used there was not strictly decreasing.

In this note we obtain an upper bound for the transient (theorem 2 in section 2). The most interesting aspect of this bound is the method used. The first results about periods of certain kinds of discrete iterations having to be 1 or 2 were obtained in [5] by rather cumbersome combinatorial methods. Other methods were developed later [2–4,7–11]. The methods in [4] relied on an "energy function;" i.e., a function associated to a configuration that could be shown to be bounded, yet was strictly decreasing at each iteration not in the cycle. The proof in [6] that the Ghiglia et al. iteration has period 1 or 2 relied on the use of the following energy function, adopted from those used in [4]:

$$E_0(t) = -\sum_{i,j=1}^{n} a_{ij} x_i(t) x_j(t-1), \tag{1.3}$$

where

$$a_{ij} = \begin{cases} 1 & \text{if } i \neq j \text{ but } j \in J_i, \\ -d_i & \text{if } i = j, \\ 0 & \text{if } i \neq j \text{ and } j \notin J_i. \end{cases} \tag{1.4}$$

This energy function has the property that $E_0(t+1) \leq E_0(t)$ for all t, but unlike the energy function of [4], equality sometimes holds for t not in the cycle, so that no upper bound for the transient can be obtained from it. On the other hand, this energy function involves only bilinear terms.

The energy function we will use here is slightly different, and is given by

$$\begin{aligned} E(t) &= \alpha E_0(t) - 2\sum_{i=1}^{n} \min(x_i(t), x_i(t-1)) \qquad (1.5) \\ &- 2\sum_{i=1}^{n}\sum_{j\in J_i} \min(x_i(t), x_j(t-1)) \\ &+ \sum_{i=1}^{n}(x_i(t) + x_i(t-1)), \end{aligned}$$

where

$$\alpha = 4\max_i d_i + 8. \qquad (1.6)$$

The novelty of our method is that we use the minimum function in the definition of (1.5), and that the terms with the minima in them appear to be essential. They allow us to show that $E(t+1) \leq E(t) - 1$ for all t not in the cycle (theorem 1), and this then leads to the upper bound of theorem 2 for the length of the transient. (As we discuss in section 3, this upper bound is probably not best possible.)

It is possible to bypass the use of the minimum function in (1.5). The function $E(t)$ of (1.5) was derived by studying an encoding (due to the first author and mentioned briefly in [6]) of the iteration (1.1) into a transformation of a graph with only 0 and 1 values, for which the methods of [4], for example, could be applied. However, the new graph is very large, and its size depends on the initial assignment of $x_1(0), \ldots, x_n(0)$. Thus this encoding does not show what is happening very clearly, and the bound for the length of the transient it gives is weaker than that of theorem 2.

After an early version of this paper was circulated, Poljak [7] obtained a bound for the length of the transient by using methods of convexity theory as well as some of the results of this paper. His bound (in the notation of theorem 2) is of the form $24ep^2 + 2$, so it is sometimes better and sometimes worse than the bound we obtain.

2. Main results

In this section we prove our main results, theorems 1 and 2.

Theorem 1. *For the transformation defined by (1.1) and the energy function (1.5), one has*

$$E(t+1) \leq E(t) - 1 \text{ for any } t \geq 1$$

unless $x_i(t+1) = x_i(t-1)$ for all i, in which case $E(t+1) = E(t)$.

Proof. Because of the symmetry $a_{ij} = a_{ji}$, a quick calculation shows that

$$\Delta E(t) := E(t+1) - E(t) = \sum_{i=1}^{n} \Delta_i(t), \qquad (2.1)$$

where

$$\begin{aligned}\Delta_i(t) = & - \alpha(x_i(t+1) - x_i(t-1))S_i(t) \\ & + 2\min(x_i(t), x_i(t-1)) - 2\min(x_i(t+1), x_i(t)) \\ & + 2\sum_{j\in J_i}\{\min(x_j(t), x_i(t-1)) - \min(x_j(t), x_i(t+1))\} \\ & + x_i(t+1) - x_i(t-1)\end{aligned} \tag{2.2}$$

with $S_i(t)$ defined by (1.2). What we will now show is that $\Delta_i(t) \leq 0$ for all i and t, and that $\Delta_i(t) \leq -1$ for at least one i if t is not in a cycle.

The first observation is that if

$$(x_i(t+1) - x_i(t-1))S_i(t) \neq 0, \tag{2.3}$$

then

$$(x_i(t+1) - x_i(t-1))S_i(t) \geq 1, \tag{2.4}$$

and so

$$\Delta_i(t) \leq -1,$$

because α (given by (1.6)) is large. To see that (2.3) implies (2.4), note that is $S_i(t) > 0$, then $x_i(t+1) = x_i(t) + 1$, so $x_i(t+1) \geq x_i(t-1)$, and an analogous argument applies when $S_i(t) < 0$.

The next observation is that if $x_i(t+1) = x_i(t-1)$, then $\Delta_i(t) = 0$. Hence it only remains to consider the case when $x_i(t+1) \neq x_i(t-1)$, but $S_i(t) = 0$. We consider two subcases:

Case a. $x_j(t) = x_i(t)$ for all $j \in J_i$. In this case $x_i(t+1) = x_i(t)$, and therefore

$$\begin{aligned}\Delta_i(t) &= 2\min(x_i(t), x_i(t-1)) - 2\,x_i(t) \\ &+ 2\sum_{j\in J_i}\{\min(x_j(t), x_i(t-1)) - \min(x_j(t), x_i(t))\} \\ &+ x_i(t) - x_i(t-1).\end{aligned}$$

If $x_i(t) \geq x_i(t-1)$, then, since $x_i(t+1) = x_i(t) \neq x_i(t-1)$, we must have $x_i(t) = x_i(t-1) + 1$, so

$$\min(x_j(t), x_i(t-1)) - \min(x_j(t), x_i(t)) \leq 0$$

for all $j \in J_i$, and therefore

$$\Delta_i(t) \leq 1 - 2 = -1.$$

If $x_i(t) < x_i(t-1)$, then $x_i(t+1) = x_i(t) = x_i(t-1) - 1$, so

$$\Delta_i(t) = -1.$$

Case b. $x_j(t) \neq x_i(t)$ for some $j \in J_i$. In this situation

$$x_i(t+1) = x_i(t) + 1$$

so, since $x_i(t+1) \neq x_i(t-1)$, we must have

$$x_i(t+1) \geq x_i(t-1) + 1$$

and therefore for each $j \in J_i$,

$$\min(x_j(t), x_i(t-1)) - \min(x_j(t), x_i(t+1)) \leq 0,$$

and for at least one $j \in J_i$, the difference above must be -1 or -2. Hence

$$2 \sum_{j \in J_i} \{\min(x_j(t), x_i(t-1)) - \min(x_j(t), x_i(t+1))\} \leq -2.$$

On the other hand,

$$\begin{aligned} 2 & \min(x_i(t), x_i(t-1)) - 2\min(x_i(t+1), x_i(t)) \\ + & \ x_i(t+1) - x_i(t-1) = 0 \text{ or } 1 \end{aligned}$$

so that $\Delta_i(t) \leq -1$ in this case also. ∎

We now turn to an investigation of the length of the transient. If we assume that the initial values satisfy $-p \leq x_i(0) \leq p$ for all i, then (1.3) and (1.5) show that $|E(t)| = O(\alpha p^2 e)$, and so the length of the transient is $O(\alpha p^2 e)$. One can improve this bound somewhat.

Theorem 2. *If $-p \leq x_i(0) \leq p$ for all i, then the iteration enters a cycle after*

$$\leq 30 \alpha \, p \, e \tag{2.5}$$

updates, where e denotes the number of edges in the graph.

Proof. We can clearly assume that the graph is connected, so that $e \geq n-1$. We first note that for all i

$$|S_i(t+1) - S_i(t)| \leq 2d_i. \tag{2.6}$$

If $S_i(t) > 2d_i$, then $x_i(t+1) = x_i(t) + 1$. Similarly, if $S_i(t) < -2d_i$, then $x_i(t+1) = x_i(t) - 1$. Furthermore, if $|S_i(t)| \geq 2d_i$, then $S_i(t+1) \leq |S_i(t)|$. The last observation implies that if $|S_i(t)| < 4d_i$, then $|S_i(t+1)| < 4d_i$ as well. ∎

We now show that after at most $2p+1$ iterations, all the $x_i(t)$ will be at most $4d_i$ in absolute magnitude. If $|S_i(u)| \geq 4d_i$, $|S_i(u+1)| \geq 4d_i$, ..., $|S_i(u+2p+1)| \geq 4d_i$, then by the above observations $S_i(u)$, ..., $S_i(u+2p+1)$ are all of the same sign, and when they are positive, we have

$$x_i(u+2p+1) = x_i(u+2p) + 1 = \cdots = x_i(u) + 2p_1,$$

which is impossible, as all the $|x_i(t)| \leq p$. The same argument applies, mutatis mutandis, when $S_i(u)$ is negative. We can therefore conclude that $|S_i(t)| \leq 4d_i$ for all i and all $t \geq 2p+1$.

We now proceed to the proof of the theorem. We have

$$\begin{aligned} |E(t)| &\leq \alpha |E_0(t)| + 4np + 2p \sum_{i=1}^{n} d_i \\ &= \alpha |E_0(t)| + 4p(n+e). \end{aligned}$$

Now

$$E_0(t) = -\sum_{j=1}^{n} x_j(t-1) S_j(t),$$

so for $t \geq 2p+1$,

$$|E_0(t)| \leq 4p \sum_{j=1}^{n} d_j.$$

Hence for $t \geq 2p+1$,

$$|E(t)| \leq 8\,\alpha\, pe + 4p(n+e).$$

The bound (2.5) of the theorem follows by an application of theorem 1. ∎

3. Concluding remarks

The constant 30 in the bound (2.5) of theorem 2 can easily be improved. However, it would be much more interesting to improve on the main term αpe. It is easy to show examples where the transient is at least proportional to n, the number of vertices, or to p, but this is probably not optimal. When the graph is a simple path, the bound of theorem 2 is of the form $O(pn)$, which may well be best possible. However, when the connectivity of the graph increases, one would not normally expect the maximal length of the transient to increase too much, if at all. It is even possible that a bound of the form $O(pn)$ might be valid for all graphs.

Some examples of assignments of values to simple paths that yield long transients are $(8,9,9,3,0,5,5,9,8)$, which has a transient of length 20 (this is known to be maximal for simple paths of length 9 with the $x_i(0)$ restricted to be between 0 and 9), and $(8,8,6,4,6,0,5,0,8,8,5)$, which has a transient of length 28. (In each case, the vector gives the values of the $x_i(0)$ for $i = 1, \ldots, n$, and vertex i is connected only to vertices j with $|i-j| = 1$.)

References

[1] D. C. Ghiglia, G. A. Mastin, and L. A. Romero, "Cellular automata method for phase unwrapping", *J. Optical Soc. America A*, 4 (1987) 267–280.

[2] E. Goles, "Dynamics of positive automata networks", *Theoretical Comp. Sci.*, **41** (1985) 19–32.

[3] E. Goles, "Antisymmetric neural networks", *Discr. Appl. Math.*, **13** (1986) 97–100.

[4] E. Goles-Chacc, F. Fogelman-Soulie, and D. Pellegrin, "Decreasing energy functions as a tool for studying threshold functions", *Discrete Appl. Math.*, **12** (1985) 261–277.

[5] E. Goles and J. Olivos, "Periodic behavior of generalized threshold functions", *Discr. Math.*, **30** (1980) 187–189.

[6] A. M. Odlyzko and D. J. Randall, "On the periods of some graph transformations", *Complex Systems*, **1** (1987) 203–210.

[7] S. Poljak, "Transformations on graphs and convexity", *Complex Systems*, **1** (1987) 1021–1033.

[8] S. Poljak and M. Sura, "On periodical behavior in societies with symmetric influences", *Combinatorica*, **3** (1983) 119–121.

[9] S. Poljak and D. Turzik, "Social influence models with ranking alternatives and local election rules", *Math. Social Sci.*, **10** (1985) 189–198.

[10] S. Poljak and D. Turzik, "On an application of convexity to discrete systems", *Discr. Appl. Math.*, **13** (1986) 27–32.

[11] S. Poljak and D. Turzik, "On pre-periods of discrete influence systems", *Discr. Appl. Math.*, **13** (1986) 33–39.