

Overview of Cellular Automata Transforms

Written By
Olu Lafe, PhD
LAFE Technologies
28001 Chagrin Blvd., Suite 305
Beachwood, Ohio 44122

Cellular Automata Transforms, invented at LAFE® Technologies, are computational operations applied to the evolving field of cellular automata. Cellular Automata (CA) are dynamical systems in which space and time are discrete. The cells are arranged in the form of a regular lattice structure and each must have a finite number of states. These states are updated synchronously according to a specified local rule of interaction. For example, a simple two-state, one-dimensional cellular automaton will consist of a line of cells/sites, each of which can take value **0** or **1**. Using a specified rule (usually deterministic), the values are updated synchronously in discrete time steps for all cells. With a K -state automaton, each cell can take any of the integer values between **0** and $K - 1$. In general, the rule governing the evolution of the cellular automaton will encompass m sites up to a finite distance r away. We say the cellular automaton is a K -state, m -site neighborhood CA.

At LAFE® Technologies, we have discovered a smart approach to addressing the astronomical number of rules describing a large family of cellular automata that can be used for data encoding applications including data compression, data encryption and the modeling of complex processes. CAT involves the generation of information building blocks. Each building block has an associated cellular automata rule. A given data can be represented by a collection of these information building blocks. The size of the CA rules for the building blocks is smaller than that of the original data. The use of the information building blocks therefore results in a more compact representation (i.e., compression) of the data. Only the associated rules used for generating the building blocks need be stored or transmitted in the place of the original data. Furthermore, because there are more CAT rules than the age of the universe, the coding of any data using these rules is also an effective way of encrypting the data. A code breaker will have to employ the services of the most powerful computers, working longer than the known age of the universe to decipher a message that has been encrypted with CAT information building blocks.

Encryption

In LAFE® Encryption the plain text is used as the initial configuration of a cellular automaton. The CAT encryption scheme is based on a special family of cellular automata generated by our patented W-Rules. A large class of these rules is periodic because the initial state of the automaton is recovered after a finite number, T , of

evolutions. This class is suitable for a secret-key encryption system. Another class is invertible because every rule set, \mathbf{W} , has an exact inverse, \mathbf{InvW} , which can be used to recover the preceding state of the automaton. This class can be used as the basis of both secret-key and public-key cryptographic systems. CAT-based encryption can be made as secure as demanded by the task at hand. The size of the key varies exponentially with the neighborhood size, m , of the automaton. Common encryption tasks are secure for key lengths ranging from 64 bits ($m=3$) to 1024 bits ($m=7$). With a choice of $m=17$ the ensuing encryption key length exceeds one million bits. CAT encryption also produces the property desired for a good cryptographic system - the "avalanche effect". This occurs when a one-bit change of the key results in a significant change in the ciphertext using the same plaintext, and a one-bit change of the plaintext yields a significant change in the ciphertext using the same key.

Compression: Use of Bases Functions

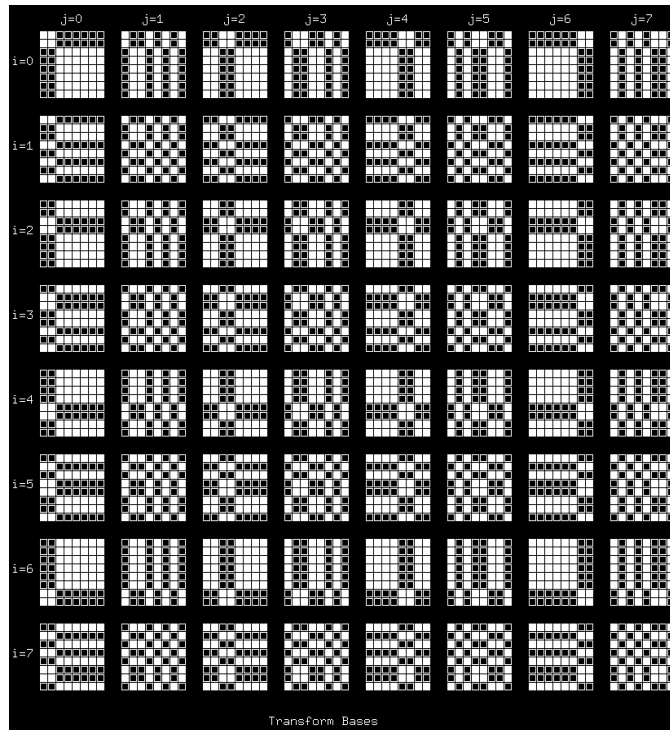
Given a data sequence f_i , all the CA transform techniques seek to represent the data in the form:

$$f_i = \sum_k c_k A_{ik}$$

in which c are transform coefficients, while A are the transform bases. The basic strategy for compressing data using CA is:

- Start with a set of CA gateway keys which produce basis functions A and its inverse B .
- Calculate the transform coefficients.
- For lossy encoding quantize the coefficients. In this approach the search is for CA bases (or base functions) that will maximize the number of negligible transform coefficients. The energy of the transform will be concentrated on the few of the retained coefficients. Ideally there will be a different set of CA gateway values for different parts of a data file. There is a threshold point at which the overhead involved in keeping track of different gateway values far exceeds the benefit gained in greater compression or encoding fidelity. In general, it is sufficient to "initialize" the encoding by searching for the one set of gateway keys with nice overall properties: *e.g.*, orthogonality, maximal number of negligible transform coefficients and predictable distribution of coefficients for optimal bit assignment. This approach is the one we will normally follow in most of CA data compression schemes. The encoding parameters include the gateway keys and the CA transform coefficients.

The figure below is the "binary-base" CAT building blocks that were reported by Lafe (1997).



Compression: W-Rule-based Transformation

We start with a given data (length N , maximum magnitude $K-1$) derived from text, documents, or digitized audio/video). The desire is to encode the data with a set of coefficients that is of a smaller length than N . In the W-Rule transformation approach we calculate a rule set W consisting of 2^m integer coefficients (each of maximum value $K-1$) is used to transform a base data B (of length N) (known to both the Coder and Decoder). W-Rule set is used to evolve the automata with the base data B as the Initial Configuration. The 2^m coefficients are stored or transmitted in the place of the original data. The compression ratio is $N/2^m$. There is a special "progressive" implementation of this approach that does not require a base data B of length N . Instead the initial configuration consists only of m input points. The computed output is used as a member of a next initial configuration for the subsequent computation. The first element is dropped and the new output is used as the m -th element). In this implementation, the compression achieved is $N/(m+2^m)$ since m additional values must be stored or transmitted to the decoder. The primary advantage of this technique is the ability to generate a large data set (i.e. large compression) given a relatively small number of encoding coefficients (2^m W-coefficients and m values for the initial configuration). The determination of the optimal values of the m coefficients is part of the encoding process.

Process Modeling

Many processes occurring in physics, chemistry, biology and information flow can often be described by Partial Differential Equations. The CA Transform approach to solving these PDEs is more direct than the traditional CA-based approach known as the lattice-gas method. We seek to use CA bases much in the same way traditional transforms, such as Fourier and Laplace transforms, are utilized. One advantage, in using CA Transforms, is the immense number of transform bases. The nice differential and integral properties of these bases can be exploited in obtaining robust solutions to PDEs.

Our principal objective, therefore, is not to model the microscopic process as done in the lattice-gas models. Rather, we want an efficient solver for the continuum equations. In the final analysis, except when called for by the resolution demands of the problem, we will not need millions of cells to solve a given PDE. The number of computational nodes in a CAT-based solution will normally be of the same order of magnitude as those required in traditional solvers based on finite differences and finite element methods. The way the CA differential operators are defined also provides us with robust transform tools for solving nonlinear equations.

Solution Process

The problem domain must be transformed into the cellular automata lattice space. If the domain is regular (*e.g.*, rectangular) the CA lattice space may be a simple discretized version of the physical problem domain. In that case the character of the partial differential equation remains unchanged. If the domain is irregular, the mapping will be more complex, and the nature of the PDE, and its associated initial/boundary conditions, may be different in the CA lattice space. In the following presentation it is assumed that the necessary transformation has been carried out and the PDE is the appropriate equation to be solved in the CA space.

For more detailed information on CAT please refer to Dr. Lafe's book: **Cellular Automata Transforms - Theory and Applications in Multimedia Compression, Encryption, and Modeling** circulated by Kluwer Academic Publishers, Boston/Dordrecht/London, 2000. The book is available at www.amazon.com and many neighborhood bookstores.