Complex Systems **5** (1991) 401–413

On Dynamical Properties of Neural Networks

Fernanda Botelho
Max Garzon
Department of Mathematical Sciences and
Institute for Intelligent Systems,
Memphis State University, Memphis, TN 38152, USA

Abstract. A transform is introduced that maps cellular automata and discrete neural networks to dynamical systems on the unit interval. This transform is a topological conjugacy except at countably many points. In many cases, it gives rise to continuous full conjugates, in which case the transform preserves entropy. The transform also allows transfer of many dynamical properties of continuous systems to a large class of infinite discrete neural networks (including cellular automata). For instance, it is proved that the network dynamics of very simple classes of neural networks, even with highly symmetric architectures, have chaotic regions of evolution (in the sense of existence of scrambled sets and configurations of arbitrarily large periods). These results raise the possibility of fully modeling parallel computability on *real-valued* dynamical systems by *discrete* neural networks.

1. Introduction

Certain types of neural networks may be regarded as discrete models of dynamical systems for phenomena hitherto deemed to be of a continuous nature [17, 18, 21], and in their own right. They are also well-established models of learning and cognition (see, for example, [22]). They are usually described by simple local rules (multiply-add-apply a squashing function) that, when applied to a network of neuron-like cells, give rise to very complex behavior. The study of the specific properties of their global behavior in terms of their local rules has become one of the most interesting and difficult problems in the field.

On the other hand, dynamical systems on continuous spaces, particularly euclidean spaces, have long been objects of study as mathematical systems, and their behavior is far better understood [8]. The purpose of this paper is to present a technique for studying the long-term behavior of neural networks as dynamical systems via their continuous counterparts on the unit interval. The main tool is a transform for neural networks regarded as dynamical systems on the Cantor set [9, 10, 12] to dynamical systems on the euclidean unit n-dimensional cubes. Roughly speaking, the main result is a metatheorem for transferring certain properties of real-valued dynamical systems to properties of neural networks.

Section 2 provides basic definitions of the two models. Section 3 presents the main technical results of the paper. Section 4 contains their proofs. Finally, section 5 points out some related issues and problems of interest for further research.

2. Definitions

The most common type of neural network is usually defined over a space of real weight lattices and sigmoid activation functions. Generally, however, implementations of these networks take the form of discrete approximations. For this and several other reasons pertaining to implementation, physical realization, and measurement limitations [12], we consider a different kind of *discrete* neural network. These networks have been considered before as generalizations of cellular automata [11, Theorem 2]. They also include generalizations of cellular automata [23] where the underlying architecture is homogeneous (can be coordinatized by a finitely generated group) and the activation functions identical for every node, although not necessarily on a euclidean grid.

We assume that the activation levels A are simply a finite set that possesses some algebraic structure that allows addition and multiplication. They also have a special activation 0. For convenience in the following definition, *cellular space* will refer to a pair (D, A) consisting of an activation set A and a countable (finite or infinite), locally-finite, arc-weighted digraph D on a vertex set V.

Definition 2.1. *An* activation function *is just a self-map $f : A \to A$ that fixes 0, that is, $f(0) = 0$. A* (discrete) neural network *is a triple $\mathcal{N} = \langle D, A, \{f_i\} \rangle$ consisting of a cellular space (D, A) (whose activation set A has an additive-multiplicative structure), and a family of activation functions f_i, one for each vertex i in D. The global dynamics of $\mathcal{N}$ is defined by equation* (1) *below.*

An assignment $x : V \to A$ of states to each cell i of a neural network is called a *total state* or *configuration*, and the set of all configurations is denoted $\mathbf{C}$. For example, O and $\mathbf{1}$ are configurations consisting of a 0 and a 1 at every cell, respectively. A pixel configuration e^k associated with node k is given by $e^k_i := \delta_{ki}$ (Kronecker's δ). The local dynamics of a neural network induces a global dynamics $T : \mathbf{C} \to \mathbf{C}$ given by

$$T(x)_i \;:=\; f_i\left(\sum_j w_{ij}x_j(t)\right), \tag{1}$$

for all cells $i \in V$.

Theorem 2.1. [10, Theorem 3.2] *A self-map $T : \mathbf{C} \to \mathbf{C}$ is realizable as an activation global dynamics of a neural network if and only if*

1. $T(O) = O$;
2. *T is continuous;*
3. *$T(e^k)$ has finite support for each pixel configuration e^k; and*
4. *$T = F \circ L$, where L is a linear self-map of $\mathbf{C}$ and F is strictly local.*

Here, $F : \mathbf{C} \to \mathbf{C}$ is strictly local if $F(x)_i = F(x_i e^i)_i$ for all x and i.

This result can be considered a generalization of earlier characterizations by Hedlund [15] and Richardson [20] of the self-maps of euclidean configuration spaces realizable by cellular automata in one and arbitrary finite dimensions, respectively.

The set of all configurations of a neural network with the product topology is a perfect, totally disconnected, compact metric space. A Cantor set is characterized as a topological space precisely by these conditions [16, p. 97]. Recall that the ternary Cantor set $\mathcal{C}$ is the set of points left in the unit interval $[0,1]$ of the real line after first deleting its middle third $(\frac{1}{3}, \frac{2}{3})$, and then continuing to delete the middle thirds of all remaining intervals *ad infinitum*. Under an encoding of this type, a neural network dynamics becomes a self-map

$$T : \mathcal{C} \to \mathcal{C}$$

of the middle-third Cantor set of the unit interval in the real line $\mathbf{R}$ satisfying certain conditions equivalent to those in Theorem 2.1. Therefore, in the remainder of this paper, suffice it to say that an infinite discrete neural network is simply a continuous dynamical system over the ternary Cantor set. Since the Cantor set topology is induced from the unit interval in the real line by the embedding obtained from its construction, henceforth denoted I, one might hope that dynamical properties of neural networks may be studied via this embedding. This is the approach taken in this paper.

The map ϕ defined below makes this program possible. Every $x \in \mathcal{C}$ can be expressed by a ternary expansion not containing the digit 1; that is,

$$x = \sum_{j \geq 1} x_j t_j \qquad t_j := \frac{2}{3^j},$$

where $x_j = 0$ or 1. Let $\phi : \mathcal{C} \to I$ be the map given by

$$\begin{aligned} \phi : \mathcal{C} &\to I \\ \phi(x) &= \sum_{j \geq 1} x_j b_j \qquad b_j := \frac{1}{2^j}. \end{aligned} \tag{2}$$

Thus a Cantor set point with finite (periodic) expansion is mapped by ϕ into a dyadic rational number (i.e., of form $p/2^q$) with finite (periodic) binary expansion. For example,

$$\begin{aligned} \phi(0) &= 0 \\ \phi(\tfrac{1}{3}) &= \phi(\tfrac{2}{3}) = \tfrac{1}{2} \\ \phi(\tfrac{1}{9}) &= \phi(\tfrac{2}{9}) = \tfrac{1}{4} \\ \phi(\tfrac{7}{9}) &= \phi(\tfrac{8}{9}) = \tfrac{3}{4}, \end{aligned}$$

and so forth. In general, ϕ successively "closes up" the gaps of C in the unit interval by piecewise affinely collapsing each of the middle-third intervals excised in the construction to the midpoint of arising subintervals of I, starting with collapsing $[\frac{1}{3}, \frac{2}{3}]$ to $\frac{1}{2}$. Hence ϕ is continuous everywhere. It is also two-to-one on endpoints (points with eventually constant expansion of 0s and 2s) and bijective elsewhere in the ternary Cantor set.

There is a way to parameterize the uncountably many right inverses of ϕ as follows. Let $\{I_n\}$ be some enumeration of the intervals successively deleted in the standard construction of the Cantor set. Enumerate accordingly the dyadic rationals of the unit interval. Identify the endpoints of I_n with 0 and 1 (say 0 := left and 1 := right). Two endpoints of an I_n (equal or distinct) will be called *adjacent*. (0 and 1 are considered adjacent as well.) Thus the family of inverses of ϕ are in one-to-one correspondence with binary expansions x of points in I. ψ_x will denote the inverse of ϕ obtained by making the choice corresponding to the nth digit x_n of x as the inverse of the corresponding dyadic point. Every choice of inverse images for finite dyadic rational points gives rise to a *right-inverse* ψ given by

$$\begin{aligned} \psi : I &\mapsto C \\ \psi\left(\sum_{j\geq 1} x_j b_j\right) &= \sum_{j\geq 1} x_j t_j\,. \end{aligned} \tag{3}$$

(Note ψ is well defined despite the fact that a real number x may have more than one binary expansion $x = \sum_{j\geq 1} x_j b_j$.)

Definition 2.2. *Given a right inverse ψ of ϕ, the ψ* transform *of a neural network $T : C \mapsto C$ into I is the self-map T_ψ of I given by the composition*

$$T_\psi : I \to I \tag{4}$$

$$T_\psi(x) = \phi(T(\psi(x)))\,. \tag{5}$$

Conversely, given a mapping $f : I \to I$ and a right inverse ψ of ϕ, one can define an inverse ψ transform *by the composition*

$$\begin{aligned} f_\psi : C &\to C \\ f_\psi(x) &= \psi(T(\phi(x)))\,. \end{aligned} \tag{6}$$

Likewise, one can define the ψ transform of a neural network T into S^1. It is the quotient map of T_ψ obtained by identifying the endpoints of I.

Hence the diagram

$$\begin{array}{ccc} & T & \\ \mathcal{C} & \longrightarrow & \mathcal{C} \\ \phi\downarrow & & \downarrow\phi \\ I & \longrightarrow & I \\ & T_\psi & \end{array} \tag{7}$$

commutes except maybe at those configurations belonging to an adjacent pair. Most of the results of this paper deal with the case where ϕ actually gives rise to a full semiconjugation of the two systems T and T_ψ for some choice of inverse ψ. This is not always possible—see Proposition 3.2 below.

Definition 2.3. *Let $\phi : \mathcal{C} \to I$ be a continuous mapping on a Cantor set $\mathcal{C}$. A neural network $T : \mathcal{C} \mapsto \mathcal{C}$ is* ϕ-continuous *if the ψ transform T_ψ into S^1 is continuous for every ψ.*

Some results on the question of what neural networks are ϕ-continuous can be obtained in terms of the following notions.

The well-known cellular automata are particular cases of neural networks [11, Theorem 2]. The configuration space of a binary neural network (in particular, a binary cellular automaton) can be regarded as a vector space over $GF(2)$ with cellwise addition, the cells being the vertices of the underlying lattice. *Linear* cellular automata are defined by convolutions of type

$$L(x)_i := \sum_{j \in N} m_j x_{i-j} \,,$$

where N is a finite subset containing cell $i = 0$ and defining the neighbors of the origin, and m_j are the *coefficients* of L. Examples in one-dimensional euclidean space are the *left shift*, the self-map of $\mathbf{C}$ given by

$$\sigma(x)_i = x_{i+1}$$

and the *right shift*, given by

$$\sigma^{-1}(x)_i = x_{i-1} \,.$$

For the sake of illustration, the encoding of $\mathbf{C}$ into the middle-third Cantor set $\mathcal{C}$ defined by the following homeomorphism will be used below. Given a doubly infinite binary configuration $c = (c_i)_{i \in \mathbf{Z}}$, let h_1 be the Cantor set point whose ternary expansion $(x_i)_{i \geq 1}$ is given by

$$\begin{aligned} x_i := h_1(c)_i \; &:= \; 2c_{i/2}, \quad i \text{ even} \\ &:= \; 2c_{-(i+1)/2}, \quad \text{else.} \end{aligned}$$

Thus h_1 conjugates every continuous self-map of $\mathbf{C}$ to a continuous self-map of $\mathcal{C}$ and vice versa. For instance, the *one-dimensional left-shift (right-shift) of I* is defined as the ψ transform σ_Ψ, where σ is the corresponding right shift on $\mathbf{C}$, and it will be denoted simply σ_1—see Proposition 3.2 below.

It is interesting to note that for every initial point with a finite expansion, iteration of σ_1 or σ_1^{-1} is eventually given by

$$\sigma_1(x) = \frac{1}{4}x$$

These maps are well known in the dynamical systems literature as *Baker maps*.

The foregoing encoding of $\mathbf{C}$ to $\mathcal{C}$ can be readily modified to an encoding of $\mathbf{C}$ to $\mathcal{C}\times\mathcal{C}$, then to the unit square $I\times I$ through ϕ componentwise. (In fact, this encoding was used in [14] to prove that *discrete* deterministic systems are capable of chaotic behavior.) One then obtains a two-dimensional ϕ transform. These alternatives will not be pursued in this paper since a lot more is known about dynamical systems on the line and circle than in higher dimensional euclidean spaces.

A point x is *n-periodic* with respect to a dynamics T if $T^n(x) = x$ for some $n > 0$, but $T^k(x) \neq x$ $(0 < k < n)$. The point is *eventually periodic* if some image $T^m(x)$ is periodic with respect to T. Other dynamical system terminology follows [8]. For instance, the *Sarkovskii order* is the ordering of the natural numbers by

$$\begin{array}{ccccccc}
3 & \rhd & 5 & \rhd & 7 & \rhd & \ldots \\
2.3 & \rhd & 2.5 & \rhd & 2.7 & \rhd & \ldots \\
2^2.3 & \rhd & 2^2.5 & \rhd & 2^2.7 & \rhd & \ldots \\
\vdots & & & & & & \\
2^n.3 & \rhd & 2^n.5 & \rhd & 2^n.7 & \rhd & \ldots \\
\vdots & & & & & & \\
\ldots & \rhd & \ldots 2^3 & \rhd & 2^2 & \rhd & 2\,.
\end{array}$$

This ordering of the natural numbers is relevant to the existence of admissible points of a given period in continuous self-maps of I or S^1 [8, Theorem 10.1]. The ordering is shown below to be relevant to the existence of periodic configurations of neural networks as well.

3. Results

The following properties of ϕ follow easily from the definitions in section 2.

Proposition 3.1.

1. *ϕ is continuous.*
2. *For any choice of x, ψ_x is continuous except possibly in some set of dyadic points.*
3. *T_{ψ_x} is Lebesgue integrable for any choice of x, and the value of the integral is independent of ψ.*
4. *A continuous ψ transform on I is uniquely determined by the value of its integral.*

The question of under what conditions T induces a continuous map of S^1 can be answered as follows.

Proposition 3.2. *If $T : \mathcal{C} \to \mathcal{C}$ is continuous, then the following conditions are equivalent:*

1. *T is ϕ-continuous.*
2. *T_ψ induces a continuous map of S^1 for some ψ.*
3. *T preserves adjacency (that is, if x, y are adjacent, so are $T(x), T(y)$).*
4. *T_ψ is independent of ψ.*
5. *Diagram (7) commutes (i.e., $T_\psi \circ \phi = \phi \circ T$) for some (every) ψ.*
6. *T and T_ψ are topologically semiconjugate for every ψ.*

In this case, T_ψ will simply be called the ϕ transform of T, and will be denoted T_ϕ.

One may wonder at this point about the existence of neural networks satisfying the conditions of Proposition 3.2. The following construction provides an infinite number of one-dimensional ϕ-continuous neural networks (in fact, cellular automata) for which the results below apply. We illustrate with a typical example.

Given two subintervals J, K of I, let $r_{J,K}$ map J affinely onto K after reflecting J about its midpoint (i.e., it maps the left end of J to the right end of K and vice versa). Let $r_J := r_{J,J}$. Let $r_{1,2} : \mathcal{C} \to \mathcal{C}$ be the self-map of $\mathcal{C}$ defined as follows. On $J_1 := [0, \frac{1}{9}]$, r acts as (the restriction of) r_{J_1} (to $\mathcal{C}$). Likewise on $J_3 := [\frac{2}{3}, \frac{7}{9}]$. But on the remaining set, r acts as $r_{[\frac{2}{9},\frac{1}{3}],[\frac{8}{9},1]}$ ($x < \frac{1}{2}$) or its inverse (otherwise). A picture shows immediately that $r_{1,2}$ preserves adjacent points, and hence its ϕ transform is continuous. For a ternary expansion of $x \in \mathcal{C}$, r flips every digit except the first two, where 01 and 11 are mapped to 11 and 01, respectively, but 00 and 10 are unchanged. (Thus r is not defined by a local rule *on the ternary Cantor set.*) Now it is necessary to prove that $r_{1,2}$ is in fact a cellular automaton.

Let T be the one-dimensional cellular automaton defined on 8 symbols $\{0 \equiv 000, 1 \equiv 001, 2 \equiv 010, \ldots, 7 \equiv 111\}$ as follows. Decode each Cantor set point via h_1^{-1} and then replace each bit b with $b11$ if it comes from the first or second digit of the expansion in the Cantor set, $b10$ otherwise (11 in the last two bits indicates a flip applied to the first bit; the other three—00, 01, and 10—indicate which action to take for the two exceptional cells). Let A be the set of symbols of form $*10$ or $*11$. The map T is defined by a *local* rule δ transforming each symbol in A exactly the same way $r_{1,2}$ does (modulo the encoding), and mapping every other state to 0 otherwise. Since T is defined by a local rule mapping O to O, T is a cellular automaton on 8 symbols. By construction, T can be conjugated to $r_{1,2}$ on $\mathcal{C}$ and therefore $r_{1,2}$ is indeed a cellular automaton. With little additional effort, it can be converted to a binary one-dimensional cellular automaton with a neighborhood of radius 6.

The continuity of T_ψ in S^1 allows one to transfer properties of real-valued dynamical systems to neural networks. For instance, a very interesting property of one-dimensional systems over the reals is Sarkovskii's Theorem [8, Theorem 10.1]. It simply says that if a continuous dynamics $T : \mathbf{R} \mapsto \mathbf{R}$ has a point of period p_0 then it has a point of every other period $p \trianglerighteq p_0$ in Sarkovskii's order.

It is known that even *linear* cellular automata may have an infinite number of forbidden periods. For instance, the linear map sum of 2 pixels $L = \sigma^0 + \sigma^m$ has no periodic points of period a power of 2 [7, remark 2.6]. This means that a Sarkovskii type statement does not hold for cellular automata in general. However, Sarkovskii's theorem can be extended to ϕ-continuous neural networks as follows.

Theorem 3.1. *If a neural network is ϕ-continuous and it has a configuration of period $p_0 \trianglerighteq 3$ in Sarkovskii's order, then it has points of every period $p \trianglerighteq 2p_0$, except possibly 2 or 4.*

In particular, the linear cellular automaton given by the pointwise sum $\sigma^0 + \sigma^m$ mentioned above is not ϕ-continuous.

The dynamical complexity of the iteration of a dynamical system f can be measured by its *topological entropy* [1]. Roughly speaking, entropy counts the number of asymptotically distinct orbits. The precise version used below follows a more dynamical-system approach due to Bowen [6].

Definition 3.1. *Let (X, d) be a compact metric space and $f : X \to X$ be a continuous self-map of X. Let n be a nonnegative integer and $\epsilon > 0$. A subset $E \subseteq X$ is called* (n, ϵ)-spanning *if for every $x \in X$ there exists $a \in E$ such that $d(f^i(x), f^i(a)) < \epsilon$, for every $i = 0, 1, \ldots, n-1$.*

The smallest cardinality of an (n, ϵ)-spanning set is here denoted $f(n_\epsilon)$, or just n_ϵ if f is understood. Since X is compact, n_ϵ is finite. It usually grows exponentially fast with n. The growth rate of n_ϵ is measured by $\limsup \frac{1}{n} \log n_\epsilon$.

Definition 3.2. *The* topological entropy *of f is given by*

$$h(f) := \lim_{\epsilon \to 0} \limsup \frac{1}{n} \log n_\epsilon .$$

The ψ transform also preserves a number of dynamically significant properties. For instance,

Theorem 3.2. *The ϕ transform preserves topological entropy.*

The results above can be generalized as follows. Let α be a dynamical property, that is, invariant under conjugacy. Say that α is *localizable* if, whenever a map has α, there exists an invariant, countable, and dense subset E of X such that the restriction of the map to the subspace $X - E$ does not have α. Otherwise α is said to be *nonlocalizable.*

Theorem 3.3. *Every nonlocalizable property α is invariant under an arbitrary transform ϕ from C onto a space X that is continuous, finite-to-one on, and a homeomorphism outside, a countable dense subset; that is, if a neural network T is ϕ-continuous, then T has α if and only if T_ϕ does.*

The following properties are nonlocalizable: sensitive dependence on initial conditions [8], chaoticity (in the sense of Li and Yorke's [19]), chain recurrence, topological mixing, and so forth.

4. Proofs

Proof of Proposition 3.1. The map ϕ is the restriction of the Cantor-Lebesgue function to the Cantor set; hence is it continuous since the topology in C is the topology induced from the unit interval. By deleting all pairs of adjacent points of the unit interval, the restriction of ϕ is still continuous, and thus ψ_x is also continuous except possibly at dyadic points. This proves (2). The third statement is an immediate consequence of (2). Now (4) follows by pulling back T_ψ by any ψ and recalling that, as a continuous map of $\mathbf{C}$, T is determined by its values in a dense subset. ■

Proof of Proposition 3.2. (1) $\Rightarrow$ (2). Obviously.

(2) $\Rightarrow$ (3). Two endpoints $a < b$ are adjacent iff $\phi(a) = \phi(b)$. Assume $T(\phi(a)) \neq T(\phi(b))$. In order to prove that T_ψ is discontinuous for every inverse ψ, assume $\psi\phi(a) = a$ (if it is b the argument is analogous). Since $T_\psi = \phi T \psi$, for any ψ, $T_\psi \phi(a) = \phi T(a)$. Let (x_n) be a decreasing sequence in I converging to $\phi(a)$ so that $\psi(x_n)$ converges to b. Since ϕ and T are continuous, $\phi T \psi(x_n)$ converges to $\phi T(b)$. Therefore T_ψ is discontinuous at $\phi(a)$.

(3) $\Rightarrow$ (1). It suffices to show that T_ψ is continuous at dyadic points for every ψ. Let y be a dyadic point with $\phi(a) = \phi(b) = y$ and $a < b$. Now

$$\lim_{x \to y^-} T_\psi(x) = \phi T(a) \, ,$$

and likewise

$$\lim_{x \to y^+} T_\psi(x) = \phi T(b) \, .$$

Since T preserves adjacent points, T_ψ is continuous at y, for every y.

(2) $\Rightarrow$ (4). It is enough to verify equality at dyadic points y. Let ψ_1, ψ_2 be two right inverses of ϕ. Since $\psi_1(y)$ and $\psi_2(y)$ are adjacent and T preserves adjacent points,

$$T_{\psi_1}(y) = \phi T \psi_1(y) = \phi T \psi_2(y) = T_{\psi_1}(y) \, ,$$

namely, T_ψ is independent of ψ.

(4) $\Rightarrow$ (5). Diagram (7) always commutes at non-endpoints. If a is an endpoint and $\psi(\phi(a)) = a$, applying ϕT proves the desired equality. Otherwise, choose ψ' with $\psi'(\phi(a)) := a$. Since T_ψ is independent of ψ, it follows that

$$\begin{aligned} T_\psi\phi(a) &= T_{\psi'}\phi(a)) \\ &= \phi T\psi'\phi(a) = \phi T(a)\,. \end{aligned}$$

Therefore diagram (7) commutes.

(5) $\Rightarrow$ (3). If $\phi(a) = \phi(b)$, then since the diagram commutes,

$$\begin{aligned} \phi T(a) &= T_\psi\phi(a) \\ &= T_\psi\phi(b) = \phi T(b)\,, \end{aligned}$$

namely, T preserves adjacency.

Finally, (5) and (6) are clearly equivalent. ∎

Proof of Theorem 3.1. This follows from the next two lemmas. ∎

Lemma 4.1. *Let T be a ϕ-continuous neural network.*

1. *If T has a periodic point of period t, then T_ϕ has a periodic point of period t or $t/2$.*
2. *If T_ϕ has a periodic point of period t, then T has a periodic point of period t or $2t$.*

Proof. If $T^t(x) = x$, then $T^t_\phi\phi(x) = \phi T^t\psi\phi(x) = \phi(x)$. Hence the period k of $\phi(x)$ divides t, say $t = qk$. Since $T^k(x)$ and x are adjacent, so are $T^{2k}(x)$ and $T^k(x)$. Hence $T^{2k}(x) = x$ or $T^{2k}(x) = T^k(x)$. In the first case, $k = t$ or $k = t/2$. In the second,

$$x = T^{(q-2)k}T^{2k}(x) = T^{(q-1)k}(x)\,,$$

which is impossible since x has period t and $(q-1)k < t$. The proof of (2) is similar and will be omitted. ∎

Lemma 4.2. [3] *Let $f : S^1 \to S^1$ be a continuous self-map of the unit circle.*

1. *If $|\deg(f)| > 1$, then f has periodic points of all periods, except only period 2 in the case $\deg(f) = -2$.*
2. *If f has a fixed point and a periodic point of period n_0, then it has periodic points of all periods $n \triangleright n_0$.*

Proof of Theorem 3.2. First prove that $h(T) \geq h(T_\psi)$. Every (n,ϵ)-spanning set $E_{(n,\epsilon)}$ for T of minimal cardinality yields a spanning (n,ϵ') set $\phi(E_{(n,\epsilon)})$ for T_ψ, where

$$|x-y| < \epsilon \Rightarrow |\phi(x) - \phi(y)| < \epsilon'\,.$$

In fact, given $x \in I$ there exists $a \in E_{(n,\epsilon)}$ such that

$$|T^j\psi(x) - T^j(a)| < \epsilon \qquad j = 0, 1, \ldots, n-1.$$

Therefore, for each $j = 0, 1, \ldots, n-1$,

$$|\phi T_\psi^j \psi(x) - \phi T_\psi^j(a)| \;=\; |T_\psi^j \phi(x) - T^j \phi(a)| < \epsilon'$$

Thus any minimal (n, ϵ')-spanning set $F_{(n,\epsilon')}$ for T_ψ satisfies $T(n_{\epsilon'}) \leq (n_\epsilon)$. It follows that $h(T_\psi) \leq h(T)$.

The converse follows from an application of the following result.

Theorem 4.1. [6, Theorem 17] *Let $(X, d), (Y, d')$ be compact metric spaces and $T : X \to X$, $s : Y \to Y$, and onto $\pi : X \to Y$ be continuous maps. If $\pi \circ T = s \circ \pi$, then*

$$h_d(T) \leq h_{d'}(s) + \sup_{y \in Y} h_d(T, \pi^{-1}(y)).$$

If moreover X and Y are compact, then $h_d(T) = h(T)$ and $h_{d'}(s) = h(s)$.

Taking $X := \mathcal{C}, Y := I$ and $\pi := \phi$, ϕ^{-1} has at most two points, so $h_d(T, \phi^{-1}(y)) = 0$ for all $y \in I$ and therefore $h(T) \leq h(T_\Psi)$. ∎

Proof of Theorem 3.3. Let E be an invariant, countable, and dense subset of X, so $\phi^{-1}(E)$ is also an invariant, countable, and dense subset D of $\mathcal{C}$. Since $\mathcal{C}$ is countable-dense homogeneous [2] (i.e., the homeomorphism group acts transitively on countable dense subsets), assume without loss of generality that D is the set of endpoints. Since α is nonlocalizable, $T|_{\mathcal{C}-D}$ satisfies α. Therefore, $T_\phi|_{X-E}$ also satisfies α by the dynamical invariance of α. ∎

5. Conclusions and open problems

A new tool, the ψ transform, has been introduced to transform neural networks into dynamical systems on the unit interval. This transform makes the two systems conjugate when the resulting system on I is continuous. This happens precisely when the transform is independent of the inverse ψ of ϕ. In this case the two dynamical systems share all dynamical properties—for instance, the same entropy and Sarkovskii property, and in general the same nonlocal properties. These results hold more generally for automata networks since they too can be regarded as, and can in many cases be simulated by, continuous self-maps of the Cantor set [10, Theorem 3.3]

A ψ tranform is an *exactly* computable self-map of the unit interval in a very precise sense, since real values can be encoded as configurations in some configuration space and computed by a local rule (easily implementable on a neural network) *in one step.* A question of interest is to find an inherent characterization (in the Cantor set or the unit interval) of dynamical systems that are computable in one step by a parallel computer (i.e., arise as ψ transforms of some neural network).

More generally, one may allow iteration of the neural network (or cellular automaton rule) transform a variable number of times depending on the input. (A fixed number of times reduces to the previous case since cellular automata are closed under composition.) A second problem is to find necessary and sufficient conditions for an $f : I \to I$ to be parallel computable.

A third question is how this notion compares to the notion of *sequential* computability for dynamical systems based on unit-cost real polynomial arithmetic introduced by [4].

Acknowledgments

This work of the first author was done with partial support from a Memphis State University Faculty Development grant. The work of the second author was done with partial support from the National Science Foundation grant CCR-9010985.

References

[1] A. Adler, A. Konheim, and M. MacAndrew, "Topological Entropy," *Transactions of the AMS*, **114** (1965) 309–319.

[2] R. Bennet, "Countable Dense Homogenous Spaces," *Fundamenta Mathematicae*, **74** (1971) 189–194.

[3] L. Block, J. Guckenheimer, M. Misiurewicz, and L. S. Young, "Periodic Points and Topological Entropy of One-dimensional Maps," pages 18–34 in (*Global Theory of Dynamical Systems*, edited by Z. Nitecki and C. Robinson (New York, Springer-Verlag, 1980).

[4] L. Blum, M. Shub, and S. Smale, "On a Theory of Computation over the Real Numbers: **NP**-completeness, Recursive Functions and Universal Machines," *Bulletin of the AMS*, **21** (1989) 1.

[5] F. Botelho and M. Garzon, "On Dynamical Properties of Cellular Automata," Memphis State University technical report 90-17.

[6] R. Bowen, "Entropy for Group Endomorphisms and Homogeneous Spaces," *Transactions of the AMS*, **153** (1971) 401–414

[7] R. Cordovil, R. Dilão, and A. Noronha Da Costa, "Periodic Orbits of Additive Cellular Automata," *Discrete Computational Geometry*, **1** (1986) 277–288.

[8] R. L. Devaney, *An Introduction to Chaotic Dynamical Systems* (Reading, MA, Addison-Wesley, 1986).

[9] S. P. Franklin and M. Garzon, "Global Dynamics in Neural Networks," *Complex Systems*, **3** (1989) 29–36.

[10] M. Garzon and S. P. Franklin, "Global Dynamics in Neural Networks II," pages 292–299 in *Proceedings of the 1st IEEE Annual Symposium on Parallel Distributed Computing*, Dallas, Texas (1989). Full version in *Complex Systems*, **4** (1990) 509–518.

[11] M. Garzon, "Cellular Automata and Discrete Neural Networks," *Physica D*, **45** (1990) 431–440.

[12] M. Garzon and S. P. Franklin, "Neural Computability II," submitted. Extended Abstract in volume I of *Proceedings of the 3rd International Joint Conference on Neural Networks*, Wash. D.C., (1989) 631–637

[13] M. Garzon and S. P. Franklin, "AMNIAC, a Universal Neural Network," forthcoming.

[14] P. Grassberger, "Chaos and Diffusion in Deterministic Cellular Automata," *Physica D*, **10** (1984) 52–58.

[15] G. A. Hedlund, "Endomorphism and Automorphism of the Shift Dynamical System," *Mathematical Systems Theory*, **3** (1969) 320.

[16] J. G. Hocking and G. S. Young, *Topology* (Boston, Addison-Wesley, 1969).

[17] J. J. Hopfield, "Neural Networks and Physical Systems with Emergent Collective Computational Abilities," *Proceedings of the National Academy of Sciences*, **79** (1982) 2554–2558.

[18] J. J. Hopfield, "Computing with Neural Circuits: A Model," *Science*, **233** (1986) 625–633.

[19] T. Li and J. Yorke, "Period Three Implies Chaos," *American Mathematical Monthly*, **82** (1975) 985–992.

[20] D. Richardson, "Tessellation with Local Transformations," *Journal of Computer and Systems Science*, **6** (1972) 373–388.

[21] P. Rujan, "Cellular Automata and Statistical Mechanical Models," *Journal of Statistical Physics*, **49** (1987) 139–232.

[22] D. E. Rumelhart and J. L. McClelland, editors, *Parallel Distributed Processing*, Volumes 1 and 2 (Cambridge, MIT Press, 1986).

[23] S. Wolfram, *Theory and Applications of Cellular Automata*, (Singapore, World Publishing Co., 1987).