

INFORMATION-THEORETIC TELEODYNAMICS IN NATURAL AND ARTIFICIAL SYSTEMS*

Anthony F. Beavers and Christopher D. Harrison

afbeavers@gmail.com

The University of Evansville

<http://faculty.evansville.edu/tb2/>

Abstract

In this paper, we employ the method of computational philosophy, the use of computational techniques to aid in the discovery of philosophical insights that might not be easily discovered otherwise, to show that it is not unreasonable to suggest 1) that there are genuine teleological causes in nature, 2) that such causes can be computed from Newtonian (“push”) causation, provided that other architectural and environmental conditions are met, and 3) that it is possible to do so without recourse to semantics by taking an information-theoretic approach to measuring information flow in a system.

Keywords: information-theoretic teleodynamics; goal-oriented-behavior; dynamic associative networks; information theory; computation.

*Penultimate draft of a paper to appear in *A Computable Universe: Understanding Computation and Exploring Nature as Computation*, edited by Hector Zenil with a forward by Sir Roger Penrose, World Scientific, 2012.

1 Philosophical preliminaries

This volume addresses the question of computation in nature. To do so, we must define “computation,” since some definitions would appear to bias the case against natural computation and some not. Among several possibilities, we can consider computation from a conservative perspective, in which case “computation” might be defined as:

{Def 1} a procedure whereby tokens are manipulated according to the specifications of a formal language.

{Def 1} includes systems of math and logic, along with Turing computation more generally. It appears at first to exclude the possibility that a rock can “compute” sunlight into heat or that a sunflower can “compute” its motion when tracking a light source, except by mere extension of metaphor. These latter examples seem to need a more liberal definition of the term, perhaps “computation” as:

{Def 2} any nomological transformation of input into output.

Initially, {Def 2} would seem to include {Def 1}, but not the reverse. However, much hangs here on the meaning of the terms “token,” “formal language,” and “nomological.” If a nomological regularity can be interpreted as rule following in a formal game (in the sense described in Haugeland [20]), then the laws of science might represent a formal language, and if a token can be a physical particle of some sort or other, then these might amount to pieces in such a game. In this case, the natural world *is* a computational system, or collection of them, under both {Def 1} and {Def 2}.

Without getting into the fine details of argument concerning whether the conservative or liberal definition of “computation” is fairest to define the phenomenon, or even whether the two definitions really collapse into the same, one might still like to posit a distinction between digital and analogue computation. {Def 1} seems to require a digital system, rather than analogue, and {Def 2} would seem to include both. But if we follow the conflation of {Def 1} and {Def 2} suggested in the previous paragraph, then natural computation would be just as digital as computation in artificial systems.

The philosophy of computation is a bit of a dicey game, as we see here, depending on how one slices the phenomena and defines terms. Since so much hangs here precisely on such matters, “carving nature at its joints” does not seem meaningfully possible in this case. Some definitions treat the natural world as a computational system, some do not, and that pretty much is the end of the story. This situation, however, does not preclude us from learning a few lessons.

There are several issues one could explore in this regard, but in this paper, we wish to examine teleodynamic systems in particular, that is, systems that pursue a state and hence are partly describable as operating according to teleological or “pull” causation rather than the “push” causation of Newtonian physics. Our approach here will be to examine teleodynamics in artificial systems with the hope of showing how they could be instantiated in some natural systems.

One important note must be made early on: the term “teleodynamic” is borrowed here from Terrence Deacon [9] and belongs to a layered distinction between types of emergence he finds in nature, but that also applies to some computational systems. Though we borrow the term, we do not follow his specific usage, nor the details of his analysis here, which is interesting in its own right, but set to different purposes, namely to an understanding of emergence in natural systems. (See [10], [11], [12] and [13]). We are more interested in the question of computation in nature with regard to goal states, employing the emerging method of “computational philosophy” defined as the use of computational techniques to aid in the discovery of philosophical insights that might not be easily discovered otherwise. (See [7]). Nonetheless, it will be difficult not to intersect with Deacon in some places while disagreeing with him in others. In the interests of staying on task and saving space, we leave it to the reader to study the similarities and differences between his view and ours, though we do agree upfront that a teleodynamic system need not be conscious, at least in any rich sense of the term.

We, like Deacon, also respect the difference between teleodynamic and teleonomic systems. The latter are systems that can be *interpreted* as pursuing a goal state, whereas teleodynamic ones actually pursue a goal state. In the example above of a rock computing sunlight into heat, it would be difficult to describe this situation as teleodynamic. The rock does not pursue

heat; rather, its release of heat is fully explained by thermodynamic principles using “push” causation. The input here is light, the mechanism of action is molecular response according to nomological laws of thermodynamics, the output is heat. The case of the sunflower, however, lends itself to a different analysis. While it is true that the *internal* operations of the sunflower can also be described as operating according to the laws of thermodynamics, the fact that the flower follows the sun (or any other like light source) suggests something else. The sunflower seeks the sun, (though through the action of its internal mechanisms.) Thus, a full explanation of why the sunflower acts the way it does requires reference to the sun, for what sense can one make of the sunflower as a sun tracker were it not for the sun that it tracks? The important question here is whether the sunflower actually “desires” (read “is attracted to”) light, or whether it can merely be described that way. From our perspective, the former description is the correct one.

If we are correct, we need to explain how “push” causation can create a mechanism that operates according to “pull” causation. In Aristotelean terms, the question can be put this way: is it possible to describe a mechanism that is built up from “efficient” or “moving” causes that ultimately ends up with legitimate “final” causes? In this paper we hope to provide a computational model to serve as an existence proof that answers the question in the affirmative.

The philosophical significance of this paper, then, is thus established. If we are successful with this exercise, then perhaps physics (as defined in the period extending from Galileo to Newton) might have been too quick in throwing out all teleological causes. In suggesting such, we do not mean to vindicate Medieval physics by putting “entelechies” into inanimate objects. Rather, we wish to suggest that *some* objects with internal working parts can exhibit genuine teleological behavior. Falling weights from the Leaning Tower of Pisa may not qualify; but sunflowers and neural systems may be another matter. The significance of this paper to the question of computation in nature should, thus, also be established. If we can demonstrate a computational model of how a mechanism based on “push” causation can harness teleological causes, then we have reason to believe that there are genuine goal-oriented computational systems in the natural world, whether one should adopt as a definition of “computation” {Def 1} or {Def 2} above.

2 Framing the question along information-theoretic lines

Information-Theoretic Teleodynamics (ITT) is an approach to intelligence in both natural and artificial systems that uses the quantity and distribution of information to drive goal-oriented behavior without recourse to semantic content. By “intelligence” we mean simply massive adaptability to one’s changing environment, including the ability to be surprised by failed expectations, respond accordingly and act on the basis of desired ends. By invoking the “quantity and distribution of information” we mean to make a direct connection to Claude Shannon’s information theory, and, in particular, to his notion of information entropy [25]. Measuring information in this way to get goal-oriented behavior is what permits us to derive such behavior “without recourse to semantic content.” This is to say that we will not “look inside” a piece of information to determine what it means; rather the measure of information, insofar as it tracks similarity and difference (along with its spread) across a networked system, provides a way to modulate conditions of attraction and repulsion that pull a system in one direction and push it away from another, as we will see in the model below.

Generally, the notion of goal-oriented behavior is thought to include reference to semantic content in the conventional sense of language-like representations (perhaps thoughts) that “stand in” for states of affairs in the world. This is partly because goal-oriented behavior is often characterized as an act of imagining some future state, developing a plan to attain it, and then executing that plan. However, action-based semantics may provide another alternative. (See, for instance, [18] and [19].) Respecting the former, Floridi notes, “in the beginning, the proto-meanings of the symbols generated by an AA [artificial agent] are the internal states of that AA, which in turn are directly correlated to the action performed by the same AA” (p. 164 [18]). Further down the page he says that “The advantage of this approach is that the very first step in the generation of meaning is not in itself a semantic process, but rather an immediate consequence of an AA’s performance.... The internal states of the AA are excellent candidates for the role of non-semantic yet semantic-inducing resources.” Indeed, in an earlier study, Grim showed that signs can acquire meanings by coordinating signaling across artificial creatures in an agent-based model [19]. “Truth telling,” here, emerged

even in elementary models based solely on the interaction between agents, agent-environment relations, and simple mechanisms internal to each agent. No structured language was involved or needed. Furthermore, these agents responded positively to signals from other agents in a teleonomic (though *perhaps* not yet teleodynamic) struggle to survive.

Though we believe that both Grim and Floridi point in the right direction, we hope to go a little further with ITT. One way to address this goal is in terms of one of the “open problems in the philosophy of information” enumerated in chapter two of Floridi [18]. He asks, “Do information or algorithmic theories ... provide the necessary conditions for any theory of semantic information?” (p. 31 [18]). Setting algorithmic theories aside, our hypothesis is that it is possible to reduce semantic information to the quantification of information flow *provided that other conditions respecting the internal structure of an agent and its relations to its environment are met*. For our purposes, we will use a dynamic associative network (DAN), defined momentarily, to provide evidence that our hypothesis is correct, at least where goal-pursuing behavior is concerned; further experiments with computer models will be necessary to advance the case for a complete reduction.

As Beavers explained elsewhere [6], traditional connectionist network modeling (with Artificial Neural Networks or ANNs) begins with a fixed network structure (in terms of nodes and connections) and then proceeds to discover a set of weights to transduce signals (here information), transforming them from input to output. The techniques for determining the proper weight set vary, but the general strategy remains the same. It starts with a predefined structure and then sets the weights for connections. Other strategies are possible. One could, for instance, let information content determine the structure of the network, rather than using a predefined structure, adding nodes and connections wherever dictated by the data. McClelland and Rumelhart’s early Interactive Activation and Competition (IAC) Models (see [22] and [24]) use this approach along with contemporary network structures that are involved in social network analysis, citation networks, co-authorship networks, and so forth. The success of these networks has been sufficient to warrant the claim that they exhibit rudimentary intelligence and qualify as preliminary forms of artificial intelligence, if harnessed to be so.

Dynamic Associative Networks (DANs) use this latter strategy as well.

These networks are built up from the data to transform input into output in such a way that the output is predictive, that is, these models learn from experience to form associations based on some sort of statistical procedure that is implicitly determined by the model. (That is, no explicit statistical methods are used.) Originally conceived to show that one could get intentionality (in both the semantic and goal-oriented senses) from association [5] and implemented in various models over the past four years, DANs have been used in a variety of micro-world experiments to prove that they can exhibit low-level cognitive abilities, including: 1) object identification based on properties and context-sensitivity, 2) comparison of similarities and differences among properties and objects, 3) shape recognition of simple shapes regardless of where they might appear in an artificial visual field, 4) association across simulated sense modalities, 5) primary sequential memory of any seven digit number (inspired by Allen and Lange [1]), 6) network branching from one subnet to another based on the presence of a single stimulus, 7) eight-bit register control that could perform standard, machine-level operations as with Turing-style computational devices, and 8) rudimentary natural language processing based on a stimulus/response (i.e. anti-Chomskian) conception of language. Models 6 and 7 here serve as proof that we could, in principle, build a Turing machine inside of a DAN, thereby showing that some dynamic network structures are Turing-complete, and model 8 *may* be an early indicator of how to get rich semantics from an information-theoretic mechanism, but, again, more work is needed.

Unlike ANNs, DANs learn by adding nodes and connections wherever needed. In more recent applications, we are reintroducing dynamic weights based on entropy equations and thresholds to improve cognitive function. In our latest experiments, for instance, we have been working to isolate control parameters to transform ordinary databases into predictive mechanisms in order to create content-addressable memory.

In keeping with the spirit of Turing, DAN architecture is intended to avoid what Beavers has elsewhere identified as the “software seduction” [8]; “It is ... quite difficult to think about the code entirely in abstracto without any kind of circuit,” (p. 384 [27]) Turing wrote in his 1947 Lecture on the Automatic Computing Engine, suggesting that in a working machine there is no code, just hardware, and that really what computer code does is to configure circuitry within computing machinery to perform a particular

information processing task. This fact has led us to adopt the slogan “circuits not software,” when characterizing DANs and reframing John Haugeland’s “formalists’ motto.”: “If you take care of the syntax, the semantics will take care of itself” (p. 106 [20]). On our view, the right kind of circuit will provide the sufficient and necessary conditions for deriving semantic-respecting and goal-oriented behavior. The model described below is *part* of our effort to establish this claim. That said, we wish to be clear that we are not maintaining that only DANs can exhibit Information-Theoretic Teleodynamics. ANNs might as well—it has simply not been our project to experiment with them—and non-networked, biological systems and other natural systems could also.

3 Information entropy, the Inverse Relationship Principle and dynamic threshold values

Information “entropy” was originally used by Shannon as a measure of the uncertainty of a piece of information, the term being suggested to him by John von Neumann because of its isomorphism with entropy in thermodynamics, though whether or not this recommendation was fitting has been a matter of some debate. There is some overlap between the two conceptions nonetheless.

Floridi notes that the concept of information entropy is easily grasped when we think of information in terms of its ability to decrease our ignorance or in terms of a data deficit [17]. The toss of a coin has become the canonical example. If the coin is fair, then we cannot predict whether it will land heads or tails. Thus, tossing the coin stands to decrease our ignorance when it lands. If the coin is weighted, however, such that it will always land heads (and we know this fact), then we stand to learn nothing new when it lands. In Shannon’s terms, the information entropy in the toss of the fair coin is higher, that is, more uncertain, than that of the weighted coin.

Though Shannon famously claimed that his attempts to quantify information in engineering had little to do with semantic content [26], his word was not the last on the subject and has also been a matter of some debate.

Barwise and Seligman [3] noted semantic corollaries to Shannon in what has been identified as the Inverse Relationship Principle (IRP) [18]. IRP says that the more rare a piece of information is, the more informative it is, and vice versa. Thus, within the domain of one’s general knowledge of animals, in the phrase, “a four-legged animal that quacks,” the range of the term “animal” is larger than that of “four-legged,” and hence is less informative. The term “quacks” is less likely to occur and is thus the most informative. (Notwithstanding the fact that there are no four-legged animals that quack, it would not be surprising if the reader’s immediate reaction to the above phrase were to think of a duck, even though it does not have four legs, because very few animals (only ducks?) quack and many animals have four legs.)

Versions of IRP appeared before Barwise and Seligman (see, for instance, Wiener [28]), and the concept has led to informational paradoxes, as in the extreme cases of tautologies and contradictions (see Hintikka [21], and Bar-Hillell and Carnap [2]) and the suggestion that “gibberish” should be more meaningful than sensical information, since it is less frequent (see Dretske, p.42 [15]). Still, IRP proves useful as the foundation for an entropy equation that points in the direction of the quantification of semantic information, depending, of course, on how it is employed. Additionally, Floridi employs a concept of semantic-oriented information entropy based on IRP in his information ethics when dealing with contradictions, though he does not believe that it is possible to reduce semantic information to information theory [16].

Our DAN model allows the user to toggle between the simple formulae $1/n$ (IRP) or $1/n^2$, where n represents the number of nodes that feed into another node in a network. The result is that the more connected a node is, the less its connections weigh, making information that is more unique count more toward modulating network behavior. It should be apparent that by inverting this strategy, we can make information that is more typical, rather than less, exhibit more influence [6].

The chief advantage with this approach in general is that weights in our network are dynamically set on the fly rather than being determined by a training method, such as backpropagation. Furthermore, DANs, unlike ANNs, are not trained by finding weights to match a training set, but by the wiring schematic that gives structure to the network. This wiring structure can be easily modified, that is, nodes can be added, connected and removed,

without interfering with the overall cognitive performance of the network. No “retraining” is necessary when encountering new information; rather, the structure is modified, which, in turn, dynamically resets the weights according to the entropy formulae indicated above (and also dynamically resets the threshold values to be discussed momentarily).

$1/n$ and $1/n^2$ both produce interesting, but different results, and richer entropy equations are available to try. By reinterpreting DANs as complex adaptive systems, the nodes being reconceived as agents and the connections as interactions, for instance, we could adopt entropy measures as they are used in complexity theory, in which case $1/n$ provides just one component of a diversity measure. (See Page [23]). To integrate these richer entropy equations, we would have to employ them differently in the network, but their dynamic nature would be the same. A broader issue, however, is raised by the presence of several alternatives here concerning the determination of what the target behavior of the network should be. If $1/n$ and $1/n^2$ both produce intuitively interesting, though different results, which is the correct formula, if either? How do we know when we have things right? Finding an answer to this question, we believe, might be aided by testing for information entropy in free association tasks using human subjects, a study that we are just beginning and that will at least provide us with a target for testing even if it is not optimal.

We can summarize the above by considering a three-layered network, though there’s no reason to suspect that neat layers are necessary, save for the ease they afford in analysis. Some of the nodes in the first layer (A) connect to some of those in the second (B), and some of those in the second connect to some of those in the third (C), in keeping with the principles of DAN architecture. If the first node in A (A1), is connected to 13 nodes in B, then the entropy for each connection to B would be $1/13$ (.0769) or $1/13^2$ (.0059) depending on which entropy formula is used. To stay with the simpler entropy formula for a moment, if a node in B, say B1, receives only two connections from A, one valuing (.0769) and the other (.1667 for $1/6$), the total weight of B1 is simply the sum of the two dynamically-determined information-theoretic weights, or, in this case, .2435. After activation of all pertinent nodes in A, some or all of the nodes in B will thus be dynamically weighted.

Having thus calculated an information-theoretic weight for each node in

B, the task now becomes one of determining which of the nodes in B will be allowed to pass activation to C. To do this, we employ a threshold value that is dynamically-determined by the collective set of summed weights for the nodes in B. For the purpose of this demonstration, we use a simple average to set the threshold. (When addressing the question of computation in nature, it is important to realize that a simple average of the weights for nodes in a network layer can be determined and employed without violating localist principles, though it should also be said that this is difficult to do with an ANN and relatively simple with a DAN.) A host of other formulae are also available for this task, ranging from sigmoid functions to functions based in parametric statistics or various combinations thereof. The best way to proceed here, again, will depend on what we can learn from tests with human subjects and our tuned intuitions about what constitutes future-oriented, cognitive behavior.

Finally, before getting to the specific details and description of the model, it is worth noting that we are also exploring the possibility that some compound entropy/threshold formula might do the needed work. However, the “divide and conquer” strategy of treating each metric separately lends itself to a more intuitive understanding of how information transduction is performed in our network. We are also considering the possibility that more than one formula could be employed by a network that could dynamically decide which route to take in specific cases analogous to the interplay between the sympathetic and parasympathetic nervous systems in human beings.

4 Getting teleodynamics from Newtonian causation

Following Deacon [9], we agree that the addition of recursion and memory to a system are necessary to get teleodynamic behavior, though the situation is such that we can be theory-neutral on the metaphysical question of emergence. Settling the matter is neither here nor there when building a working, computational model.

Neither “recursion” nor “memory” are new to the language of networks, but a few preliminary comments should be made. Memory can be taken

here in a very minimal sense stripped of any cognitive interpretation; for our purposes, it is simply the ability of a system to detect when it is in a state that it has been in before. Memory in networks is not “stored” and “retrieved” as with conventional computers. Rather, it is implicitly embedded in the network structure itself, or, perhaps better, it *is* that structure coupled with the dynamics that permit particular parts of the circuit (network) to activate under certain conditions.

Recursion is intended here in the straightforward sense of recirculating information through the same network. To stay with the structure described in the previous section, layer A and layer C are made up of nodes representing the same features, with network processing going on between A and B, and, then again, between B and C. In our associative network, layer A is made up of nodes that represent various properties, layer B nodes that represent the objects that have those properties, and layer C again the properties in layer A. Activation passes from A to B and then B to C. Output values from C are then recirculated as the input values for A on each recursive iteration, letting these values define a trajectory for further activation.

The nodes in A represent the following characteristics which are listed below in the original order that they appear in the network. This is not incidental. The network was built by imagining that we were teaching a child about animals and, thus, reflects encounters that are more or less random as it would be in a real learning situation. It is worth noting in passing, that network performance remained intuitive throughout the entire engineering process, though this is a topic for another paper. That said, the nodes in A represent: small, furry, meows, barks, winged, flies, swims, finned, crawls, many-legged (more than four), two-legged, warm-blooded, live offspring, walks, four-legged, farm animal, oinks, moos, rideable, big, crows, woodland, hops, hoofed, curly, hisses, cold-blooded, scaled, quacks, coos, seafaring, spawns, large-mouthed, web-making, whiskered and feathered.

The nodes in B represent the following animals: dog, cat, bird, duck, fish, caterpillar, human, pig, cow, horse, rooster, rabbit, deer, retriever, poodle, mouse, snake, crocodile, lizard, turtle, dolphin, whale, ostrich, pigeon, shark, salmon, bass, house fly, cricket, spider, catfish and seal. The reader has, no doubt, noticed that we have included general classes of animals along with specific members. Thus, the list includes, for instance, bird, but also, duck, rooster, ostrich and pigeon. Doing so does not interfere with network perfor-

mance as long as differentiating properties appear in A that allow the network to distinguish specific instances from archetypes. Even without these differences, however, the network will still classify members with their archetypes. Proper nouns (such as “Rover”) could also function as animal names in the network as well, though we have not done so here.

Describing the complete wiring schematic is beyond the scope of this paper, but in the space allotted, we can provide some samples. Node A5 (winged) is connected to B3 (bird), B4 (duck), B11 (rooster), B23 (ostrich), B24 (pigeon), B28 (house fly) and B29 (cricket). Since A5 connects to 7 nodes in B, the weight on each connection is $1/7$ using IRP or .1428. Node A11 (two-legged) is connected to B3 (bird), B4 (duck), B7 (human), B11 (rooster), B23 (ostrich) and B24 (pigeon), which is 6 connections for a weighted value of $1/6$ using IRP or .1667. Simultaneous activation of A5 (winged) and A11 (two-legged) produces these weights in the B layer: B3 (bird .3095), B4 (duck .3095), B7 (human .1667), B11 (rooster, .3095), B23 (ostrich .3095), B24 (pigeon .3095), B28 (house fly .1428) and B29 (cricket .1428).

Having established weights for the appropriate nodes in B for the activation of A5 and A11, the network now calculates a dynamic threshold value, in this case .2500, determined by averaging the weights of all activated nodes in the B layer, though, again, another formula may prove better down the line. This allows activation from B3 (bird), B4 (duck), B11 (rooster), B23 (ostrich) and B24 (pigeon) to pass to the C layer. B7 (human), B28 (house fly) and B29 (cricket) are appropriately filtered out.

In turn, B3 (bird) is connected to C1 (small), C5 (winged), C6 (flies), C11 (two-legged), C12 (warm-blooded), C14 (walks), C21 (woodland) and C36 (feathered). Since B3 makes 8 connections, the entropy value is $1/8$ using IRP or .1250. The incoming weight of B3 was .3095, as we saw above, which then gets multiplied by the entropy value of .1250 to yield .0387 as the weighted output value of B3, which is then summed into each connection B3 makes to the C layer. In like manner, connections are made from B to C for B4 (duck), B11 (rooster), B23 (ostrich) and B24 (pigeon) to yield the following weights for nodes in the C layer:

C1	small	.1482
C5	winged	.1998
C6	flies	.1040
C7	swims	.0653
C11	two-legged	.1998
C12	warm-blooded	.1998
C14	walks	.1556
C16	farm animal	.0442
C20	big	.0515
C21	crows	.0442
C22	woodland	.0696
C29	quacks	.0309
C30	coos	.0344
C36	feathered	.1998

Table 1: Weights on the nodes in the C layer based on activation of the above.

The threshold value based on the average weights for the C layer is .1105, thereby allowing the items in bold above to pass the threshold while filtering the others out. Nodes that pass the threshold are then resubmitted to the input layer for recursion.

Before continuing further, a few observations should be made. Activation of winged and two-legged, returns the bird set (duck, rooster, ostrich and pigeon) along with the generic bird for output from the B layer to C. More interestingly, the output values for C partially complete the bird pattern. Activation of winged and two-legged attracts small, warm-blooded, walks and feathered. Note that at this point, flies has failed to be pulled into the list of appropriate characteristics, since ostriches and roosters do not fly. However, the next recursion cycle produces an interesting effect. Rooster and ostrich are slightly demoted (but only slightly), letting the generic bird with the duck and pigeon count for more. Typicality then wins out on the next recursion cycle over the uniqueness that is tracked by IRP, as flies is pulled into the list of relevant characteristics, further demoting the rooster and the ostrich. Further recursion after the third cycle produces no further re-evaluation of nodes, unlike in other networks that can take several hundred iterations to settle.

Suppose, however, that we want the rooster to rise and the other birds to be demoted. To do so, we must activate a property in A that has high entropy and belongs to the rooster. Unfortunately, using the IRP entropy formula still drags flies into the list of pertinent properties, and over the course of a few iterations the duck passes the rooster. However, using $1/n^2$ does the task.

Switching over to $1/n^2$ and activating A5 (winged) and A11 (two-legged) still picks out the same set of birds (bird, duck, rooster, ostrich, pigeon) as the network asks for activation of small, warm-blooded, walks and feathered, just as before. One iteration pulls flies in the list as well, as the bird, duck and pigeon climb just over the ostrich and the rooster. However, if we begin with A5 (winged), A11 (two-legged) and A21 (crows), the rooster climbs substantially over the other birds. Recursion in this case does not pull flies into the property set, but it does get farm animal as it should.

Generally, $1/n^2$ works better than $1/n$, at least using the averaging thresholding formula indicated above, though not always, and admittedly we run into the occasional counter-intuitive case. The network, for instance, is good at not letting a property set contain both two-legged and four-legged, but sometimes it will admit a set including both warm- and cold-blooded, indicating that we still have some work to do. Again, the situation is complicated by the fact that the threshold parameter also affects network performance. We are optimistic (and aren't we all!) that further analysis and experimentation with more advanced mathematics in light of test results from human trials will improve performance, but this does not mean that we are not seeing important theoretical results. For the purposes of this paper, *the way in which network activation pushes its way through a trajectory that simultaneously opens before it is particularly interesting, since the trajectory dynamically modifies itself in step with changing network activation and thereby continually sets up possibilities that the network will actualize given further stimulus.*

These possibilities allows us to conceive of the system as teleodynamic and not merely teleonomic. To see this, we can go back to Aristotle's notion of final (or teleological) causation. Aristotle does not think that every acorn will become an oak tree, since external influences might impede it. Squirrels sometimes eat acorns for food after all. Rather, when the acorn is considered according to its own internal principles, including its material (*hyle*) and form (*morphe*), if it becomes anything at all, it will be an oak tree and not, say, a

pig or a walnut tree [4].

Computational philosophy, again, is the use of computational tools to establish points of philosophical interest that might not be seen otherwise [7]. Even though this work is very much in progress, we do believe that we can tentatively claim on the basis of what we have said here that teleological causes can be computed in principle, given the right kind of structure, one in which information-theoretic quantities can be used to determine information flow through a network to produce activation patterns that march in step with our semantic sensibilities. We cannot say where this research will take us, only that the results are sufficiently interesting to warrant not setting the project aside at this time. Indeed, we see two possibilities for further projects that we believe will produce enlightening results. One is a project in teleodynamic game play in which we will engineer a network to play a formal game using information-theoretic principles and pattern matching rather than heuristic search. The other concerns enriching the communications strategies in Patrick Grim’s artificial life model [19], mentioned above, to see if we can build a community of artificial agents that can learn from each other and that might cross the boundary between simple signal processing to something closer to language use. (See the comments from Floridi [18] on action-based semantics above.)

5 Information-Theoretic Teleodynamics in natural systems

While it would seem obvious that there are teleodynamic systems in nature, since human beings seem to engage in goal-oriented, intentional behavior, the matter is far from settled. Dennett’s work on the “intentional stance” [14] and the literature surrounding it has been a part of rigorous debate for decades about whether people (and some animals) engage in genuine intentional behavior or whether such behavior is merely ascribed. We do not have space to rehearse the debate here, though in the terms of this paper, the question can be phrased to concern whether people and animals engage in genuine teleodynamic behavior or whether such is merely teleonomic, which is sufficient to establish that the question remains open. Is

teleonomic description a mere strategy that we use to cope with circumstances in a partially-ordered/partially-disordered world, or is there more to the story?

Hardcore reductionists of a variety of stripes, too, would like to suggest that the natural world can be fully understood solely in terms of “push” causation. We would like to challenge this picture with this simple model of “pull” causation that may provide an important part of the explanation about why a system behaves the way it does. Is such a system plausible in the natural world? Again, without suggesting that only DAN structures can yield such results, we believe that the answer to the question is affirmative. Though we are quite far from a complete explanation, the circuit here suggested is surprisingly simple, yet cognitively rich. A simple recurrent circuit in which properties map to objects which then map back to the properties that define them and that is controlled by the kind of electro-chemical dynamics one might find in real neural systems is not beyond the reach of nature, especially given the rich array of neural structures that are part of the human brain.

Of course, one need not stop here. If such behavior is explicable in information-theoretic terms, we might see it in a variety of structures, including those supporting information propagation in a group of human and/or non-human animals or even in wholly non-biological systems. We suspect that this latter possibility will interest the reader of this volume most, even though it has not been our immediate target here. We therefore invite the reader to speculate. Can communication (understood in the natural and not social scientific sense) and information processing in the natural world more generally be understood on the basis of Information-Theoretic Teleodynamics? We do not yet know, but this exercise might provide some support for the pursuit of worthy answers.

Acknowledgments

We would like to thank the following individuals for their helpful feedback on various phases of this project: Colin Allen, J. McKenzie Alexander, Aaron Bramson, Selmer Bringsjord, Luciano Floridi, Patrick Grim, Mirsad Hadzikadic, Derek Jones, Matthias Scheutz, Talitha Washington, Hector Zenil, and the attendees of a 2011 Institute for Advanced Topics in the Dig-

ital Humanities sponsored by the National Endowment for the Humanities (USA) and the Complex Systems Institute at the University of North Carolina, Charlotte.

References

- [1] C. Allen and T. Lange, Primary sequential memory: An activation-based connectionist model, *Neurocomputing*, 11(2-4):227-243, 1995.
- [2] Y. Bar-Hillel and R. Carnap, An outline of a theory of semantic information. Research Laboratory of Electronics, MIT, 1952.
- [3] J. Barwise and J. Seligman, *Information flow: The logic of distributed systems*, Cambridge, 1997.
- [4] A. Beavers, Motion, mobility and method in Aristotle's *Physics*: Comments on *Physics* II, 192b20-24, *Review of Metaphysics*, 42:357-374, 1988.
- [5] A. Beavers, Intentionality and association, The 3rd Annual Computing and Philosophy Conference, Oregon State University, Corvallis, Oregon, August 8th, 2003.
- [6] A. Beavers, Typicality effects and resilience in evolving dynamic networks, in FS-10-03, AAAI, 2010.
- [7] A. Beavers, Recent developments in computing and philosophy, *The Journal for General Philosophy of Science*, Springer Online First, 2011.
- [8] A. Beavers, Alan Turing: Mathematical mechanist, in S. B. Cooper and J. van Leeuwen (Eds.), *Alan Turing: His work and impact*, Elsevier, forthcoming.
- [9] T. Deacon, Emergence: The hole at the wheels hub, in P. Clayton and P. Davies (Eds.), *The re-emergence of emergence*, Oxford, 2006.
- [10] T. Deacon, Shannon-Boltzmann-Darwin: Redefining information. Part 1, *Cognitive Semiotics*, 1:123-148, 2007.

- [11] T. Deacon, Shannon-Boltzmann-Darwin: Redefining information. Part 2, *Cognitive Semiotics*, 2:167-194, 2008.
- [12] T. Deacon, What is missing from theories of information?, in P. Davies and N. Gregersen (Eds.), *Information and the nature of reality*, Cambridge, 2010.
- [13] T. Deacon, *Incomplete nature*, Norton, 2011.
- [14] D. Dennett, *The intentional stance*, MIT, 1987.
- [15] F. Dretske, *Knowledge and the flow of information*, MIT, 1981.
- [16] L. Floridi, Information ethics: On the philosophical foundation of computer ethics, *Ethics and Information Technology*, 1:37-56, 1999.
- [17] L. Floridi, *Information: A very short introduction*, Oxford, 2010.
- [18] L. Floridi, *The philosophy of information*, Oxford, 2011.
- [19] P. Grim, et. al., Making meaning happen, *Journal for Experimental and Theoretical Artificial Intelligence*, 16(4):209-243, 2004.
- [20] J. Haugeland, *Artificial intelligence: The very idea*, MIT, 1985.
- [21] J. Hintikka, Information, deduction and the a priori, *Noûs* 4(2):135-152, 1970.
- [22] J. McClelland and D. Rumelhart, An interactive activation model of the effect of context in perception, part I. Chip Report 91, Center for Human Information Processing, Univ. of California, San Diego, 1980.
- [23] S. Page, *Diversity and complexity*, Princeton, 2011.
- [24] D. Rumelhart and J. McClelland, An interactive activation model of the effect of context in perception, part II. Chip Report 95, Center for Human Information Processing, Univ. of California, San Diego, 1980.
- [25] C. Shannon, A mathematical theory of communication, *Bell System Technical Journal*, 27:379-423, 623-656, 1948.

- [26] C. Shannon and W. Weaver, *The mathematical theory of communication*, Univ. of Illinois, 1949.
- [27] A. Turing, Lecture on the automatic computing engine, in B. J. Copeland (Ed.), *The essential Turing*, Oxford, 2004.
- [28] N. Wiener, *The human use of human beings*, Houghton Mifflin, 1950.