PRE-PRINT of: Dorin, A., & McCormack, J., "Self-Assembling Dynamical Hierarchies",
in *Proceedings of Alife 8*, Standish et al (eds), 2002, MIT Press

Self-Assembling Dynamical Hierarchies

Alan Dorin & Jon McCormack

School of Computer Science and Software Engineering
Monash University, Clayton, 3800, Australia
{aland, jonmc}@csse.monash.edu.au

Abstract

This paper addresses the open problem of assembling multi-levelled hierarchical structure. It presents a model of an infinitely-levelled, self-assembling dynamical hierarchy which arises from the interaction of geometric primary elements with a fixed complexity. A formal description of the presented hierarchy is derived. This quantifies the relative compression achieved by describing the system in terms of components of different organization. The relationship between properties of representations and those of physical objects is then discussed to support the view that at each level in the hierarchy presented, the components exhibit emergent properties not possessed by those at the levels below. It is concluded that these new properties are trivial and that such infinitely-levelled structures may be constructed easily. However since the definition of the problem in the literature admits such trivial possibilities, further discussion is required to ensure "interesting" emergent properties are clearly distinguished from those that are not.

Keywords: Self-assembly, Dynamical hierarchy, Observation and Representation, Artificial chemistry.

Introduction

Hierarchies are a useful way of understanding the organization of life (Nehaniv and Rhodes 2000, Baas 1994, Chaitin 1970, Simon 1962, Mirkin et al. 1996). Higher order biological organisms are constructed from atoms, molecules, organelles, cells and organs; hence, one aspect of relevance to constructing virtual organisms is the representation of this hierarchy. In software, a computer program specifies the primary components and their interactions. In contrast with the interactions of matter in the physical universe, computer programs deal exclusively with representations, the meaning of which is determined by observers. It is important to make this distinction, not only for the sake of clarity in discussions about virtual life, but also in defining hierarchies and their properties.

The notion that complex outcomes or behaviours may be arrived at through the interactions of simple building blocks is commonly held by artificial life researchers (Rasmussen et al. 2001a). Research has sought rules describing interactions between basic elements, in the hope that they will give rise to aggregates exhibiting new *emergent* properties not apparent in the primary elements themselves. These properties are understood to appear at, or even *define*, each level in the hierarchy — an idea discussed in more detail below.

Whilst *The Game of Life* (see e.g. Gardner 1970) and other cellular automata may yet provide a basis for constructing hierarchies with emergent properties, producing a multi-level hierarchy through self-assembly of primary units remains an open problem (Bedau et al. 2000). For a given framework, (Rasmussen et al. 2001a) propose that it may be impossible to extend the levels in a hierarchy, without adding to the complexity of the base units. This idea seems to run counter to the extreme view that it ought to be possible to derive complex global outcomes from simple local interactions.

Additional questions concerning hierarchies

The proposed relationship between the complexity of primary units and the number of hierarchical levels they may construct, raises a number of potentially interesting issues for artificial life. For example, how much complexity, if any, do the base units require to construct an extra level in the hierarchy (and how ought this be measured)? Is this amount independent of the order of the level being considered? Is there a threshold for the complexity of the basic building blocks, beyond which an infinitely levelled hierarchy may be achieved? Is the physical world limited in the number of hierarchical levels that are possible to arrange (and what is the evidence to support this)?

Representations, Hierarchies & Properties

Many systems in the artificial life literature are *computational/representational* rather than physical. The "building blocks" and "structures" discussed in these systems are commonly referred to *as if* they were material entities, even though they are only representations of those entities.[1] As authors of software, we must be certain our application of terms such as *hierarchy*, *complexity* and *property* are carefully considered. We propose in this paper, that one way to increase the rigour with which artificial life software is analyzed, is through the application of the principles of information theory (see Chaitin 1987). Information theory is particularly applicable to the study of computational artificial life, as this field fundamentally concerns patterns in information.

In this paper we describe a hierarchical structure of unlimited order, which self-assembles from primary units of fixed complexity. Each level in the potentially infinite hierarchy is shown to possess properties arising from the interactions of its components, which the lower level components do not themselves possess.

The reason for presenting this hierarchy is not to demonstrate that *any* hierarchy may be assembled from base units of fixed complexity, it is merely to show that hierarchies *do* exist that can be self-assembled in a manner that meets criteria specified in artificial life literature (Rasmussen et

[1] Although the machine on which the representations are manipulated is obviously physical.

al. 2001a). The model illustrated is uncomplicated, and is not directly related to any specific biological system. This allows us to illustrate that the current definitions for *hierarchy* and *property* need to be formalized in the context of artificial life. Formalization will assist us to define more clearly the relationship of software models to real biological systems.

The following sections discuss aspects of the simulation presented here that satisfy previously proposed criteria for hierarchies. These sections also address criteria recently proposed for multi-levelled, dynamical hierarchies constructed from components of fixed complexity that exhibit new properties at each level.

Related work

This paper addresses issues raised in (Rasmussen et al. 2001a, Rasmussen et al. 2001b, Gross and McMullin 2001) concerning self-assembly of hierarchical structures which model aspects of biology. The system below is a simple artificial physical/chemical system which bears resemblance to some of the more typical artificial chemistries documented in (Dittrich, Ziegler, and Banzhaf 2001) However, the present system is more closely derived from research in cellular automata (Gardner 1970, Langton 1986), and systems that link the concepts of cellular automata and artificial chemistry in order to study self-assembly (Dorin 2000). Other related research is detailed in the individual sections below to which it is most relevant.

The self-assembling hierarchy

This section describes a self-assembling hierarchical system that exhibits new properties at each level and does not require added complexity at the base level to achieve any number of additional levels. The number of levels that may be assembled is limited only by the amount of basic building material available in the model and the memory constraints imposed by the machine.

The basic elements of this system are equilateral triangles laid out on a planar, triangular grid (Figure 1). The entire system is updated simultaneously in discrete time steps. At each time step of the simulation each triangle may be shifted to a random, neighbouring, unoccupied location if one is available. The triangular lattice dictates the direction in which a triangle's vertices are oriented so that movement to a neighbouring cell includes a rotation by 180 degrees. This may also be viewed as a flip about a horizontal axis running through the triangle's centre. Triangles were selected for this model as they are the simplest regular polyhedra that can be used to tile a plane. Squares could equally well have been employed.

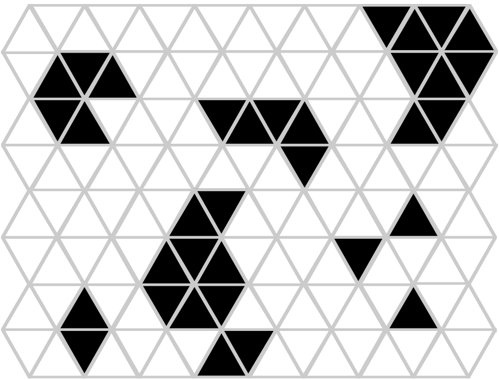

Figure 1 **A section of the triangular grid.**

After the movement stage of each time step, all triangles are examined to see if they neighbour any others. If they do neighbour another triangle, the two triangles will bond to one another with a fixed probability, *b*, established at the start of the simulation. If two neighbouring triangles are already bonded together at this time step, they will dissociate with probability, *d*, determined similarly.

During the movement stage of each time step, if the upper and leftmost triangle in a bonded aggregate is selected to move into a neighbouring location, the movement of all triangles in the whole is constrained in this direction also. This ensures that the aggregate is treated as a rigid body. Any planned movement that would cause an intersection between a member of the aggregate and an occupied cell on the grid is cancelled and the aggregate remains stationary for this time step.

The next section describes possible behaviours of the system.

Operation of the model

The probabilities *d* (dissociate) and *b* (bond) dictate the tendency of the triangles to form larger aggregates and for these to break apart after having been constructed. As long as *b* is non-zero, the chance that at least a single bond will join two aggregates increases in proportion to the length of the edge. That is, if many potential bonding sites are presented, the chance that at least one of these will link the two aggregates increases. Conversely, if the value of *d* is not unity, aggregates with internal structures made of closely packed triangles, presenting many internal edges for redundant bonding, are much less likely to dissociate than aggregates with narrow cross-sections.

Overall then, large, broad structures with redundant internal bonds will tend to develop, whilst smaller or long narrow structures will tend to break down. The extent to which each of these phenomena occur depends on the values of *b* and *d*.

Some elementary structures that may appear during a run of this model are illustrated in Figure 2 . It is apparent from Figure 3 that triangles may form larger triangles (or other shapes), and these may be assembled into larger triangles (or shapes) still. Thus, here is a form that may assemble itself from primary elements into larger and larger structures — a nested hierarchy. It should be clear from the description above that no additional information needs to be given to the individual triangles to have them continue to build a hierarchy of multiple levels. A measure of this hierarchical organization is given in the discussion below. This will be followed by a discussion of the new properties that arise at each level in the hierarchy.

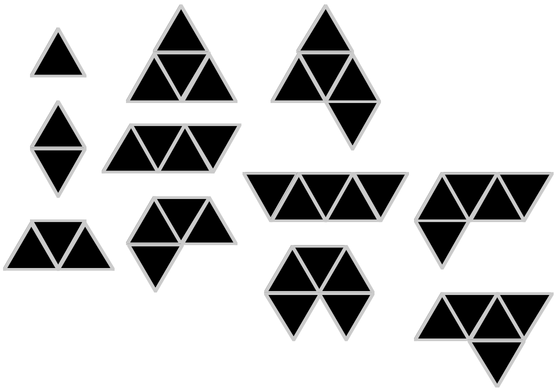

Figure 2: Sample shapes formed of primary elements.

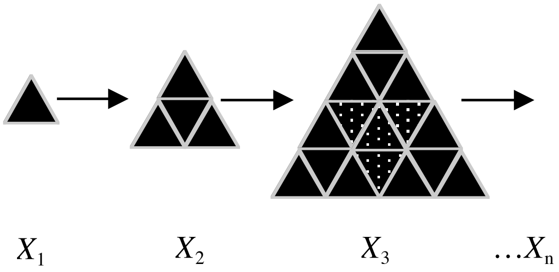

Figure 3: A hierarchical structure.

Identifying hierarchies

Hierarchies have been studied across a range of disciplines including Mathematics (Mirkin et al. 1996), General Systems Theory (Bertalanffy 1968 p.74, Simon 1962, Simon 1994 p.196), Information Theory (Boulton and Wallace 1970), in general biological terms (Polanyi 1968), and recently in Artificial Life (Nehaniv and Rhodes 2000). For the purposes of defining a hierarchy in this paper, we follow the description given by Baas in (Baas 1994). Specifically, for a given set of elements X, X is a *division hierarchy* (referred to commonly as a *nested-hierarchy*) if there is associated with it a system of levels $X_1, X_2, \dots X_n$, such that $X_n = X$, with each X_i related by a series of mappings:

$$X_1 < X_2 < \dots < X_n$$

That is, nested hierarchies involve levels that consist of, and contain, lower levels.

In the present situation, we are dealing with representational systems and it is therefore necessary that an appropriate way of defining and comparing hierarchical organization in this context be developed. The issue of concern here is the state of the variables being used to represent the properties of a system, and those of its components. This is fortunate as it allows us to find a measure of hierarchical structure or organization, which is difficult to find for real biological organisms or their artifacts (Chaitin 1970). Namely, we may specify the redundancy in a nested hierarchical structure, and thereby discover its levels and the collections of elements that are its components.

Information measure of the system

Let us take the triangular system X_3 illustrated in Figure 3 as an example to formally demonstrate the presence of a multi-level, nested hierarchy. A structure X of order n, written X_n, is composed of $4^{n-1}X_1$ primary elements (in this case simple triangles). If each primary element X_1 requires p bits to specify its position and orientation then, X_n requires $4^{n-1}p$ bits to specify as an aggregate of X_1's.

But, if X_n can be described in terms of the position and orientation of the 4 lower level elements X_{n-1} that compose it then, X_n requires $4p$ bits to specify in terms of X_{n-1}.

So, for n levels, X_n can be specified hierarchically in $4p$ (n-1) bits. Since, $4p(n-1) < 4^{n-1}p$ if $n > 2$, the hierarchical description is clearly more efficient than that obtained in terms only of the primary elements.

Ockham's razor may be paraphrased, "if two theories explain the facts equally well then the simpler theory is to be preferred". So the hierarchical scheme, because it requires less information (bits) to specify the aggregate's structure, is preferred (Wallace and Boulton 1968).

In practice, the triangles in the model must share edges to count as an aggregate. Even in continuous space the number of bits required to code a collection in terms of primary elements, is therefore substantially less than 4^{n-1}p. This holds because once a primary element is fixed in space using p bits, the location of others in the aggregate may be specified relatively. For the purposes of this example therefore, 4^{n-1}p may be considered as an upper limit or worst case. Similar constraints reduce the number of bits required to specify all levels of the hierarchy. Furthermore, since the model proposed here actually constrains triangles to lie on a regular lattice, the number of bits required to represent an aggregate is substantially less than that required to do so in continuous space. The principles may be shown to hold equally in continuous or lattice space however.

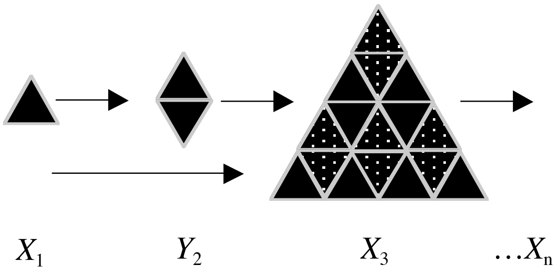

Figure 4: An alternative hierarchy to construct X_3

There may be more than one way of viewing a composite object as a hierarchy. For example, the structure X_3 in Figure 3 may also be seen in terms of components X_1 and Y_2 (Figure 4). We can obtain a measure of how succinctly this different way of viewing the system's components compresses the data using the procedure outlined above. If it results in a more concise description it is to be preferred over the decomposition given above.

Perhaps no compression is obtained in the hierarchy, for example, a trivial hierarchy in which an aggregate is made of two dissimilar components, *X* and *Y*, can also be specified in terms of these. There is no redundancy in the composition and hence no compression will be gained in the description. In this case there is no useful reason to view this aggregate as a nested hierarchy. One may as well accept that the aggregate is flat in its organization.

Whilst the example above is a hierarchy of triangles and larger triangles, of course it needn't be the case that the hierarchical structures formed by this model are perfectly regular, nor need it be the case that the same form be repeated at multiple scales. The form in Figure 5 serves as an example of a hierarchy with different shapes at each level. Of course there are a multitude of possible infinitely levelled hierarchies. The development of specific structures in any given run of the model is currently left to chance. That the space supports the development of such structures however remains clear.

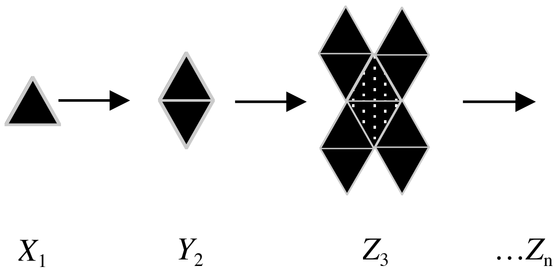

Figure 5: An alternative hierarchical structure.

Now that it has been shown that the structures above may in fact be specified hierarchically, it remains to be shown that each level of the hierarchy exhibits new emergent properties not found in the lower levels. This is the subject of the following section.

Identifying properties

Seeing is a theory laden enterprise
— Hansen (Hansen 1958)

In the physical world, a property is any observable aspect of an entity — an attribute, characteristic, feature, trait or aspect (Bealer 1999). For example, the wavelength of the light an object reflects is a property of that object, as are its length and mass.

In some artificial life literature, a property that arises through the interactions of many simple parts which do not themselves possess this property is labelled *emergent*. Hence, "A property that applies at a given level is emergent if it does not apply at any lower level" but with the proviso that "the specification of observable properties is somewhat arbitrary" (Rasmussen et al. 2001).

Let us make a few remarks regarding properties. Firstly, properties are *observed* by *people*. People overlook some things and are very good at detecting others. We deal with our environment by searching for certain kinds of pattern, ignoring, completely oblivious to, even incapable of grasping others (Dennett 1991, Tufte 1990). Technology may broaden the scope in which we search for patterns, by mapping them from unobservable domains to those which we may examine.

We may well describe a particular system in terms of many properties. However, the properties which stand the test of time are those that aid our understanding or enhance our ability to predict the way the system behaves. These properties help us to form a model of the system under observation.

There is some (perhaps misleading) sense of objectivity when one "observes" a property of the physical world and measures it. In the world of representation, in this case software-based universes, what does it mean to "observe a property"? A property in the virtual world may be anything an observer wishes it to be, as long as it can be distinguished from other properties. Any symbol or bit may stand for a property. It may indicate the presence or absence of some object, concept or ability. It may represent colour, or even beauty, yet it does not necessarily offer any predictive power about the properties of real word objects. It is just a signifier.

The concept of a *variable* in a computer program running on a digital computer and the different *values* it may take on are fundamentally based on patterns in the underlying digital machinery. The statement "variable X has the value 2" about a computing system is a different pattern in memory to the state representing "variable X has the value 3". The *observer* defines the *meaning* of a symbol or pattern in the machine. If it has any relation to the real world, this too is assigned by an observer.

If we wish to distinguish between a property "2" and a property "3", whether they are represented as values of a variable in a digital computer or in some other way, we need two different signifiers, one for each property. Hence, the number of properties which may be distinguished in a system of representations is limited by the number of discrete states it may enter. The properties of a collection of representations are determined by the kinds of relationships which those representations may be interpreted as participating in. The degree to which representations may interact of course depends on the number of ways in which they may be organized with respect to one another, which again, is determined by the number of states each may have.

To take a simple example, the "position" of a cell on a CA grid is, in the virtual world, simply a state of the data structure which represents it. If two cells are "neighbouring" this is *not* saying anything about their location in physical space, but is a comparison of their state variables which are used to *represent* position.

To say that an entity acquires a new property in the virtual/representational world (i.e. that a new property emerges) is to comment only on there being new relationships between the state variables used to represent the entity. The more state variables the entity has to describe it, the greater the number of properties it may be distinguished as having. To gain extra state variables therefore goes hand in hand with being able to distinguish new states, and therefore new properties[2].

Let us return now to the hierarchy illustrated in Figure 3. As the number of triangles in an aggregate increases, the aggregate does in fact gain new properties through this increase. Specifically, consider the observed property of a single, un-bonded triangle, "I may move as a single triangle". A new property of a bonded set of 4 triangles is, "I may move as a body of 4 triangles", and surprisingly enough, a set of 16 triangles has the property, "I may move as a body of 16 triangles".

Discussion

Trivially then, larger structures have properties which none of their components may be observed to have. There are many possible ways a large aggregate may be internally bonded. In order for a primitive element to remain attached to an aggregate it only requires a bond across one edge, even if it presents three edges as potential bonding sites. The property of moving as a rigid body of a certain size then emerges from the (bonding) relationships between the components. This may not seem very interesting, the properties of the structures are trivial, but they are properties nevertheless.

Hence, this is a self-assembling, infinitely-levelled hierarchy, which exhibits emergent properties at each level. Additionally, the primary elements do not require extra complexity in order to extend the number of levels in the hierarchy.

Conclusions

The example presented here is trivial — the hierarchy is not "interesting" (but it is a hierarchy), the new properties are not "interesting" either (but they are new properties). Yet, by the description in (Rasmussen et al. 2001) our system meets all the criteria to disprove the proposed ansatz. We expect that the authors of the ansatz did not have such a trivial model in mind when they proposed it, and that our model would not interest biologists in the slightest. However, this has not been our aim in this paper.

We propose that what is needed is a more formal definition of the kinds of behaviours we wish to see in our representations at each level of the hierarchy. What is also needed is a more considered discussion about the kinds of properties and relations between components which we would like to arise at each level in the hierarchy. Specifically, a definition of emergence that enables us to measure a property and determine whether or not it is emergent would be of benefit.

One way to do this might be to consider the systems under investigation in terms of their information content. Since representations on a computer are created and manipulated using bits, information theory is an objective way of comparing different biological theories, especially where they relate to hierarchies and the emergence of new properties.

Acknowledgments

The authors would like to thank John Crossley for his suggestions on some of the concepts discussed in the paper. Alan would also like to thank Mark Bedau for a discussion that initiated some of the ideas presented here.

References

Baas, N.A. 1994. Emergence, Hierarchies and Hyperstructures, *Proc. of Artificial Life III*, Addison-Wesley, 515–537.

Bealer, G. 1999. Property in *The Cambridge Dictionary of Philosophy*, 2nd edn., Audi (ed.), Cambridge Univ. Press, 751-752.

Bedau, M.A., et.al. 2000. Open Problems in Artificial Life, *Artificial Life,* **6**(4), MIT Press, 363–376.

Bertalanffy, Ludwig von, 1968. *General Systems Theory, Foundations, Development, Applications*, revised edition, George Braziller Inc.

Boulton, D.M. and Wallace, C.S. 1970. An information measure for hierarchic classification, *The Computer Jour-*

[2] This is not the *only* way new properties may arise. For example, new relationships may also arise between *existing* state variables as they take on new values within the range of expressible values they already possess. However, this does not alter the fact that the addition of new state variables is equivalent to the addition of new properties.

nal, **16**(3), 254–261.

Chaitin, G.J. 1987. Information Randomness & Incompleteness, Papers on Algorithmic Information Theory, World Scientific.

Chaitin, G.J. 1970. To A Mathematical Definition of 'Life', *ACM, SICACT News 4*, January 1970, 12-18.

Dennett, D.C. 1991. Real Patterns, *Journal of Philosophy*, 88, 27–51.

Dittrich, P., Ziegler, J., Banzhaf, W. 2001. Artificial Chemistries — A Review, *Artificial Life*, **7**(3), MIT Press, 225–275.

Dorin, A. 2000. Creating a Physically-based, Virtual Metabolism with Solid Cellular Automata, *Proc. Artificial Life 7*, Bedau, M. et al (eds), MIT Press, 13–20.

Gardner, M. 1970. Mathematical Games: The Fantastic Combinations of John Conway's New Solitaire Game 'Life', *Scientific American*, **223**(4), 120–123.

Gross, D., McMullin, B. 2001. Is it the Right *Ansatz*?, *Artificial Life*, **7**(4), MIT Press, 355–366.

Hansen, N.R., 1958. *Patterns of Discovery,* Cambridge Univ. Press.

Langton, C.G. 1986. Studying Artificial Life with Cellular Automata, *Physica* 22D, North-Holland, 120–149.

Mirkin, B., McMorris, F. R., Roberts, F. S., Rzhetsky, A., (ed's) 1996. Mathematical Hierarchies and Biology, *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, Vol. 37, Am. Math. Soc.

Nehaniv, C.L., Rhodes, J.L. 2000. The Evolution and Understanding of Hierarchical Complexity in Biology from an Algebraic Perspective, *Artificial Life*, Vol. 6, No. 1, MIT Press, 45–68.

Polanyi, M. 1968. Life's Irreducible Structure, *Science*, **160**(21) 1308–1312.

Rasmussen, S., Baas, N.A., Mayer, B., Nilsson, M., Olesen, W. 2001a. *Ansatz* for Dynamical Hierarchies, *Artificial Life*, **7**(4), MIT Press, 329–354.

Rasmussen, S., Baas, N.A., Mayer, B., Nilsson, M. 2001b. Defence of the *Ansatz* for Dynamical Hierarchies, *Artificial Life*, **7**(4), MIT Press, 367–374.

Simon, H.A. 1962. The Architecture of Complexity, *Proc. Am. Phil. Soc.*, **106**, 467–482.

Simon, H.A. 1994. *The Sciences of the Artificial*, second edition, MIT Press.

Tufte, E.R. 1990. *Envisioning Information*, Graphics Press.

Wallace, C.S. and Boulton, D.M. 1968. An information measure for classification. *The Computer Journal*, **11**(2), 185-194.