Self-Reproduction of Dynamical Hierarchies in Chemical Systems

Bernd Mayer
Steen Rasmussen

SFI WORKING PAPER: 1998-05-038

SFI Working Papers contain accounts of scientific work of the authors and do not necessarily represent the views of the Santa Fe Institute. We accept papers intended for publication in peer-reviewed journals or proceedings volumes, but not papers that have already appeared in print. Except for papers by our external faculty, papers must be based on work done at SFI, inspired by an invited visit to or collaboration at SFI, or funded by an SFI grant.

©NOTICE: This working paper is included by permission of the contributing author(s) as a means to ensure timely distribution of the scholarly and technical work on a non-commercial basis. Copyright and all rights therein are maintained by the author(s). It is understood that all persons copying this information will adhere to the terms and constraints invoked by each author’s copyright. These works may be reposted only with the explicit permission of the copyright holder.

www.santafe.edu
Self-Reproduction of Dynamical Hierarchies in Chemical Systems

Bernd MAYER \textsuperscript{a} and Steen RASMUSSEN \textsuperscript{b,c}

\textsuperscript{a} Institute for Theoretical Chemistry and Radiation Chemistry, University of Vienna, UZAII, Althanstrasse 14, A-1090 Vienna, Austria.
bernd@asterix.msp.univie.ac.at; http://asterix.msp.univie.ac.at/local-link

\textsuperscript{b} EE5-5 MS D407 and T-CNLS MS B258, Los Alamos National Laboratory, Los Alamos, New Mexico, 87545, U.S.A., steen@lanl.gov.

\textsuperscript{c} Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico, 87501, U.S.A., steen@santafe.edu.

Abstract

In biological systems higher order hyperstructures seem to occur both in an intuitive and a formal sense. Starting at a molecular level of description we have: molecules, polymers, supramolecular structures, organelles, cells, tissues, organs, etc. But in models and simulations of these systems it has turned out to be difficult to produce higher order emergent structures from first principles. We demonstrate how monomers (first order structures) compose polymers (second order structures) which in turn can assemble into ordered, micellar (third order) structures, which in turn can self-reproduce as they catalyse the formation of additional amphiphilic molecules. Processes of this particular kind have probably been important for the origins of life.

Our molecular system is defined on a 2-D lattice and the dynamics is modeled as a discrete automaton. In this system all interactions (electromagnetic forces) are decomposed and communicated via propagating information particles. Each lattice site has an associated data structure where molecules are represented by information particles and their associated force fields (excluded volumes, kinetic energies, bond forces, attractive and repulsive forces) are decomposed and propagated as information particles as well. The propagation- and interaction rules are derived from Newton's Laws.

Based on this self-assembly and self-reproduction example it is possible to extract some of the principles involved in the generation of higher order (hyper-) structures and relate them to dynamical systems. An Ansatz for generating higher order structures in formal dynamical systems is given.

Introduction

One of the key steps in the origin of life and also one of the key elements in making a proto-cell is the emergence of a protecting 'shell' within which a more stable and controlled environment can be maintained [Schnur 1993]. Such a shell must be able to self-reproduce together with a templating molecular complex (a primitive genome) and a simple aggregate that is able to harness external energy (a proto-metabolism), presumably of chemical nature [Deamer 1997]. The most primitive, known self-reproducing shell-aggregate is a micelle. Luisi and his group pioneered this experimental work in 1990 [Bachmann et al. 1990] and have later elaborated on this scenario to include self-replicating RNA within a self-reproducing liposome yielding simultaneous core- and shell replication [Oberholzer et al. 1995] as well as specific shell-shell recognition [Berti et al. 1997]. The main topic of the present paper is to study the self-assembly of micelles and concomitant micelle self-reproduction in simulation using realistic models for the relevant physico-chemical interactions.

A step by step aggregation of molecular elements can lead to the emergence of novel functionalities as it must have happened in the processes that eventually lead to the origin of life. However, the same phenomenon is found in contemporary living systems as they also have clear functionalities to emerge at different levels:

molecules $\rightarrow$ organelles $\rightarrow$ cells $\rightarrow$ tissues $\rightarrow$ etc.

At each of these levels of description we can observe distinct properties which only have a meaning at this particular level and where each of the levels is generated by the levels below. A dynamical system organized in this manner defines a dynamical hierarchy [Baas et al. 1996]. The above identification of levels of course has many sub-levels. It should also be noted that it does not define a strict hierarchy. Communication can and does indeed occur between e.g. level 1 and level 5. In this paper we demonstrate how a 3-level dynamical hierarchy can be generated in a formal system and show how our Ansatz to do this also works in the general case for even higher order systems. A more detailed discussion of these issues can also be found in [Baas et al. 1996].

There are two significant reasons why it is not trivial to generate a dynamical hierarchy in a formal system: (i) It involves multilevel dynamics - that is simultaneous dynamics on several time- and length scales, which requires large computational resources. (ii) The natural, conceptual framework for such a system seems to be a set of interacting objects and not a closed form model as e.g. a differential equation system. How to form the higher order structures from the bottom up becomes simple in systems of interacting objects. For example,
starting with objects that are models of monomers it is trivial that the monomer objects can form polymers as they are combined into a string. Now the polymers can form membranes as they are aggregated in a particular fashion, and so forth. A systematic study of systems of interacting objects, including cellular automata, is a relatively recent scientific activity [Wolfram 1986, Langton et al. 1989].

It turns out that a formulation of a dynamical hierarchy can be made conceptually simple if the interacting objects are defined on a lattice. Furthermore, the most promising results on modeling macroscopic effects in chemical systems based on a fine grained system representation are based on lattice-type simulation methods [Chen et al. 1992, Ostrovsky et al. 1995, Covenev et al. 1996, Emerton et al. 1997]. Thus, we introduce a lattice gas style [Frisch et al. 1986], discrete field simulation concept [Rasmussen and Smith 1994, Baas et al. 1996, Mayer et al. 1997, Mayer and Rasmussen 1998] which allows to simulate micelle formation and self-reproduction as well as demonstrate some of the fundamental formal properties of a dynamical hierarchy.

Emergence of Dynamical Hierarchies

We have just discussed the notion of dynamical hierarchies in the context of chemical and biological systems. In order to understand them better, i.e. their use as well as their synthesis, it is important also to have formal systems in which such structures can be generated. In general when higher order structures occur, new properties arise at each level — for example through aggregation. This means that in this context we will be looking for objects or aggregates with new properties. An obvious question is then what new really means? This brings us into the basic discussion of emergence and the notion of an observer. For a general discussion of emergence and higher order structures we refer to [Baas 1993] where the concept of emergence and the relation between emergence and dynamics can be discussed. How this connects to dynamics we refer to [Rasmussen and Barrett 1995]. Let us just recall briefly a few basic notions [Baas 1993]. We consider families of objects or structures $S_r^1$ of first order

$$S_r^1 = S_r^1(f_{rs}, s_r, \tau_r), \ r, s = 1, 2, \ldots, n$$

where $s_r$ is the state of the object, $f_{rs}$ defines the object-object interactions, and $\tau_r$ is the local object time. In addition we need to define an update functional $U$ which schedules the object updates (e.g. parallel, random, event driven), which together with the interaction rules — given by $f_{rs}$ — defines the dynamics. Also the important notion of an observer $O^1$ needs to be introduced. With $O^1$ we can measure explicit system properties as for instance internal object states. The system dynamics may now generate a new structure $S^2$ through the interactions

$$S^2 = R(S_r^1), \ r = 1, 2, \ldots, n$$

where $R$ is the process that generates $S^2$. This is what we call a second order structure which may be subjected to a possible new observer $O^2$. Then we say that a property $P$ is emergent iff

$$P \in O^2(S^2), \text{ and } P \not\in O^2(S_r^1).$$

Clearly emergence depends on the observer in use which may be internal or external. An external observer can be the experimenter, but it can also be a mechanism (given by an algorithm) encoded in the system to detect patterns, regularities, aggregates, correlations, etc.

Part of the dynamics may actually be viewed as an internal observer. For example, each object will typically receive information about its neighborhood and act accordingly following an algorithm as an experimenter would have done from the outside. Also note that the emergent properties may be computable or non-computable [Baas 1993, Rasmussen and Barrett 1995].

The above process can be iterated in a cumulative, not necessarily a recursive, way to form higher order emergent structures which we shall call hyperstructures of e.g. order $N$:

$$S^N = R(S_{r_{N+1}}^{N-1}, S_{r_{N+1}}^{N-2}, \ldots).$$

It should be noted that the definition of an observation function is no more — or just as — arbitrary as the definition of the objects and their interactions. For more details we refer to [Baas 1993] and [Rasmussen and Barrett 1995] where the concept of emergence and the relation between emergence and dynamics are discussed.

<table>
<thead>
<tr>
<th>Level of Description</th>
<th>Molecular Structure</th>
<th>Emergent Property</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st Order</td>
<td>Lipophilic, Hydrophobic</td>
<td>Phase Separation, Phase Distribution</td>
</tr>
<tr>
<td>2nd Order</td>
<td>Polymers</td>
<td>Elasticity, Radius of Gyration</td>
</tr>
<tr>
<td>3rd Order</td>
<td>Micelles</td>
<td>Inside / Outside, Permeability, Self-Replication</td>
</tr>
</tbody>
</table>

![Figure 1: Dynamical Hierarchy in a Chemical System](image)

As we use this conceptual framework to interpret how self-assembling processes form micelles which again self-reproduce we obtain the following picture (see figure...
At each level of description the physical structures; the monomers (level 1), the polymers (level 2), and the micelles (level 3), 'carry', or more correctly, through their interactions, generate, properties that cannot be observed at the levels below. They are emergent properties. At the fundamental level 1 (water and monomers) we can e.g. observe the generation of water structures as well as phase separation between water and hydrophobic monomers. At the polymer level, level 2, we can e.g. observe elasticity. At the micelle level, level 3, we can observe an outside and an inside as well as permeability which does not have any meaning at the level below. Also, the micelles (the third order structures) can under certain conditions self-reproduce (assuming a distinct chemical reactivity at the water-membrane interphase [Oberholzer et al. 1995]) which none of the objects at the lower levels can. Thus, the interactions of the molecules (water and monomers) in this relatively simple, chemical system, generate higher order structures which carry non-trivial (e.g. life-like), emergent properties.

It should be stressed that the encoded functional properties of the basic, first order objects does not change during this process. It is only the context within which these objects are arranged that changes. Thus, the operational semantics of the information, the forces each object receives from its environment, is context sensitive. For example the accessible states for a hydrophobic monomer in bulk polar phase are distinctively different from the states in a non-polar phase and again different from the states of a hydrophobic monomer in an amphiphilic polymer [Baas et al. 1996, Mayer et al. 1997]. This fact defines a downwards causality as the higher order structures modulate or restrict the dynamics of the lower order structures by which they are made up. This phenomenon of observed downwards causality in dynamical hierarchies is related to the 'slaving principle' as originally suggested by Haken [Haken 1987].

The next section introduces the principles of a lattice gas style simulation concept, the Lattice Molecular Automaton (LMA) [Rasmussen and Barrett 1994, Mayer et al. 1997, Mayer and Rasmussen 1998], which is capable of generating higher order, chemical structures, as self-replicating micelles, in a simulation.

The Discrete Field Automaton Concept

The basic idea behind the discrete field automaton is to model both, matter and forces, as mediating information particles. Three main steps determine the molecular dynamics: (i) rules that propagate force information particles, (ii) rules that evaluate the received information together with the local state, and (iii) rules that move molecules on the lattice and transform the system into the next time-step. All rules are directly derived from the laws of physics.

Our simulation takes place on a square, 2-D lattice, but the general formalism holds for arbitrary lattice topologies as e.g. a triangular lattice [Rasmussen and Smith 1994, Baas et al. 1996, Mayer et al. 1997, Mayer and Rasmussen 1998]. The simulation objects (i.e. molecular entities) and vacuum are encoded as data structures, located at each site of the lattice (see figure 2A). A Boltzmann distribution of kinetic energies as well as potential energies (based on discrete force fields) are implemented to describe the molecular dynamics. Kinetic energies are distributed between colliding molecules following a hard sphere model conserving the overall momentum. The total potential energy \( V_{total} \) of our model system with \( n \) molecules on a lattice with \( q \) neighbors is described by:

\[
V_{total} = \sum_{i=1}^{n} \sum_{j=1}^{q} V_{dip}^{i,j} + V_{ind}.dip. + \sum_{i=1}^{n} \sum_{j=1}^{q} V_{coop}. \tag{5}
\]

These potential energy terms are implemented to account for specific physico-chemical properties of our molecular objects as e.g. dipoles, induced dipoles, hydrogen bond donor and acceptor sites, or polarizability volumes, all crucial parameters for the micelle formation.
in polar environment. This set of weak intermolecular interactions given in the above equation, commonly summarized as Van der Waals forces, has generally proven to be suitable for a description of macromolecular systems [Privalov et al. 1998, Brownman 1975] as well as responsible for the emergence of bulk phase phenomena as the structured hydrogen-bond network in water or the hydrophobic effect [Mayer et al. 1997]. Kinetic energy terms drive the molecular system into local minima on the potential energy hypersurface. The overall setup corresponds to a microcanonical ensemble, conserving mass, momentum and the total energy.

The dynamical system that defines our LMA is of the form
\[ \{S_r(t+1)\} = U\{S_r(t)\}, r = 1, \ldots, n. \] (6)
where \( S_r = S_r(f_{rs}, x_r, \tau_r) \),
\[ \] (7)
denotes the interacting objects defined on the 2-D, square lattice. Each object has an internal state \( x_r \), an object-object interaction function \( f_{rs} \) (which has its own state \( x_s \) as an argument together with the state(s) of the object(s) that it is interacting with \( x_s, s = 1, 2, \ldots, n \)), and local time \( \tau_r \). To generate the dynamics the object-object interactions have to be updated by an update functional \( U \), which is random sequential for this version of our LMA.

A data structure \( D_t^{(i,j)} \) at the lattice location \( (i,j) \), at time \( t \), denoting an object \( S_r \), e.g. vacuum if it is empty and a molecule if it is occupied - with \( k \) variables \( x_h, h = 1, \ldots, k \)
\[ D_t^{(i,j)} = (x_1^{(i,j)}(t), \ldots, x_k^{(i,j)}(t)), \] (8)
which means that the new value of a given variable at a given lattice site is a composed function of the variables at the site \((i,j)\) itself and of the variables \( h = 1, \ldots, k \) (where \( k \) defines the number of variables in the data structure) at the neighboring sites (in the four principal directions). For a more detailed discussion of the formal properties of such a dynamical system we refer to [Baas et al. 1996, Rasmussen et al. 1997].

A data structure \( D \) is defining all (lattice) objects through the \( q = 7 \) variables \( x_1 - x_7 \) (see figure 2A). These sets define the local object’s state space. Below is the list of variables associated with every lattice location (data structure) \((i,j)\) in the LMA:
\[ X_1 = \{ x_1 \}; \ x_1 \in N_0; \ \text{type-state}: \]
molecular types (including vacuum) at site \((i,j)\).
\[ X_2 = \{ x_{2,1}, \ldots, x_{2,8} \}; \ x_{2,1} \in N_0; \ \text{rec-type}: \]
molecular types (including vacuum) in the neighborhood of site \((i,j)\).
\[ X_3 = \{ x_{3,1}, \ldots, x_{3,4} \}; \ x_{3,1} \in Z; \ \text{send-state}: \]
outgoing force particles along lattice directions.
\[ X_4 = \{ x_{4,1}, \ldots, x_{4,8} \}; \ x_{4,1} \in Z; \ \text{rec-state}: \]
ingoing force particles from lattice directions.
\[ X_5 = \{ x_{5,1}, \ldots, x_{5,4} \}; \ x_{5,1} \in N_0; \ \text{kin-state}: \]
local kinetic energy at location \((i,j)\) in \( q \) directions.
\[ X_6 = \{ x_{6,1}, \ldots, x_{6,4} \}; \ x_{6,1} \in Z; \ \text{move-state}: \]
list of net energetic states (including potential and kinetic energies).
\[ X_7 = \{ x_{7,1}, \ldots, x_{7,8} \}; \ x_{7,1} \in \{0,1\}; \ \text{bond-state}: \]
maintain bonds within polymers.

The \text{type-state} defines vacuum and the molecular type at the monomer level, e.g. as water, a hydrophilic or a hydrophobic object. Associated with the \text{type-state} is the \text{send-state} list denoting the respective force field, e.g. describing dipoles and hydrogen bond sites. The information particles in the \text{send-state} list are propagated to the neighborhood and stored in the respective \text{rec-state} list. The \text{rec-type} stores the \text{type-state} entries from the neighboring molecules and is used to interpret the force particle data received in the \text{rec-state}. Thus the entries in the \text{send-state}, \text{rec-state} and \text{rec-type} lists are used to calculate the potential and respective forces following a Coulomb potential. The \text{kin-state} list holds the kinetic energy state of the object and the \text{bond-state} encodes the bonds formed within a polymer. The individual information particle propagation steps are schematically shown in figure 2B (propagation of ‘repellons’ to maintain excluded volume, of ‘bondons’ to keep the configuration of polymers) and 2C (propagation of force particles ‘forceons’ capable of representing the potential energy surface given in equation (5)). A water molecule is characterized by three hydrogen-bond sites, hydrophilic monomers show two
hydrogen bond sites (e.g. a COOH group) and all molecular objects furthermore interact based on various dipole contributions. Finally, the calculated potential energy as well as the kinetic energy determine the move - state of the object, i.e. which lattice site will be occupied in the next update cycle, considering constraints as the excluded volume and bonds between monomers in polymers.

It should be noted that as we include more details about the physics of the interactions it becomes possible to generate higher and higher order structures [Baas et al. 1996, Mayer et al. 1997]. For example, if we as here require that these objects need to be able to generate third order structures (micelles) which are able to self-reproduce and that the molecular interactions are based on known physical principles, then it cannot be done with much less object complexity (variable and interaction functions) than given here. A detailed discussion of the notion of object complexity and what it means in connection to a system’s ability to generate higher order (≥ 3) emergent structures is given in [Baas et al. 1996, Rasmussen et al. 1997].

The full LMA update cycle holds the following individual steps:
1. propagation of molecular types
2. construction of type-specific force fields
3. calculation of potential energies
4. calculation of the most proper move direction
5. random update of the individual objects transferring the system into the time step \( t + 1 \)

A detailed description of the update cycle is given in reference [Mayer and Rasmussen, 1998].

Simulation of Dynamical Hierarchies and Self-Reproduction in a Chemical System

A central molecule in biomolecular structure and dynamics is water. The strong polarity of water and concomitant ability to form hydrogen bonds allows to form large, hydrogen-bonded water networks that continuously break up and reform due to thermal noise. It is the entropic and enthalpic balance associated with joining these networks of water molecules that is responsible for the hydrophobic effect [Mayer et al. 1997]. An example of a water network generated by the LMA is shown in figure 3A, which is a detail of a 100 × 100 lattice.

When hydrophobic monomers are present in water we observe a phase separation of the molecular types (see figure 3B). The hydrophobic effect is responsible for this separation phenomenon, but it should be noted that the mechanics of the separation is as follows: The water - hydrophobic- and the hydrophobic-hydrophobic attractions are of the same magnitude in this simulation: The enthalpic loss due to hydrogen bond breakage is counteracted by various dipole interactions (dipole - induced dipoles between water and hyrophob. monomers and ind. dipoles - ind. dipoles between hydrophob. monomers) giving a comparable enthalpy for bulk water and the mixture. This is set according to experimental results on changes of enthalpy, entropy and resulting free energy of such mixtures. In our LMA model mixtures, the phase separation is generated by the complex dynamics of solvating hydrophobic surfaces and maintaining the hydrogen bond network, which results in an entropy-driven phase separation process [Mayer et al. 1997]. This process is commonly referred to as the hydrophobic effect, clearly an emergent property.

As monomers polymerize they form 'strings' or polymers. As we discussed in the previous section the dynamics of the polymer generates properties which are not observable at the level of the individual monomer. The polymer has its own (lower) diffusion constant, a radius of gyration, an elasticity constant, just to mention a few. In figure 3C we see how hydrophilic polymers stay solvated in water (contrary to hydrophobic objects) as they can participate in the ever changing water networks (as the hydrophilic model polymers can participate in the hydrogen bond network, too). When the polymers are amphiphilic (hydrophilic head and a hydrophobic tail) the hydrophobic effect together with the structure of the polymers generate a structurally well defined aggregate, a micelle, which we can identify as a third order structure (see figure 3D).

A micelle carries properties like - inside / outside, permeability, has another (much lower) diffusion constant
and defines a new interface chemistry. In fact, the particular chemical properties at the hydrophilic heads organized in the surface can hydrolyze one end of a hydrophobic polymer as demonstrated in the oleic acid / oleate system experimentally [Oberholzer et al. 1995]. The resulting polymer is amphiphilic. This amphiphilic polymer now seeks, with a certain probability, into the micelle that catalyzed it in the first place and as many such processes occur the micelle continues to grow until it becomes unstable and divides.

Thus, we have an autocatalytic self-reproduction of micelles as long as the system has a surplus of hydrophobic polymers which can be hydrolyzed. Figure 4A shows an initial condition of a mixture of access hydrophobic and some amphiphilic polymers in aqueous solution, where one micellar structure is formed. If one head group of a hydrophobic polymer faces the polar surface of the micelle (i.e. at least two hydrophilic head groups in our simulation) it is 'hydrolyzed', forming an amphiphilic polymer. After the hydrolysis of all the available hydrophobic polymers the initial micelles have replicated many times (see figure 4B). The dynamics of this process is reflected in the concentration of amphiphilic polymers which grows rapidly until it runs out of fuel (hydrophobic polymers). The kinetics of this process is shown in figure 4C.

This demonstrates how we can within our LMA concept (i): generate a dynamical hierarchy of chemical nature, which carries distinct properties at each level, (ii) including the ability for the generated third order structures to self-reproduce in an autocatalytic fashion.

**Conclusion and Outlook**

By defining a formal system of interacting (molecular) objects on a 2-D lattice we demonstrate how a dynamical hierarchy can be generated in simulation. All interactions are derived from first principles and are thus representing a simplified, but still realistic picture of the actual physico-chemical nature of the system. At each level of description (monomers, polymers, micelles) distinct emergent properties are being generated by the dynamics, which are also observable in the corresponding real system. The system's ability to generate higher order structures depends in a nontrivial way on the object complexity. The more details of the physics we include, the higher order structures the system can generate. We demonstrate how a generated third order structure, a micelle, can self-reproduce in an autocatalytic fashion, as has been shown in vitro.

This approach is neither limited to 2-D nor to generate third order structures. We are in the process of extending the simulation to 3-D and by including more details of the physics in the first order objects it is possible to have the system generate fourth or higher order emergence, e.g. allowing a templating molecule to cooperate with the self-reproducing micelles. This Ansatz has opened an avenue where we in a formal system can have objects interact with objects to generate higher order objects which in turn can interact with yet other objects which again can generate higher order objects, etc, without any principal limit.

This leaves us with many unanswered questions including: At which level of object complexity can we in a formal system generate a dynamical hierarchy which supports, say what corresponds to a proto-organism? Such a limit probably exists, since only a limited set of biological precursor molecules are in turn combined into a limited number of biomolecules, which in turn seem to constitute a universal 'tool kit' from which all cellular organelles are assembled. This is also in fine agreement with the existence of neutral networks for RNA secondary structures, and a very limited number of actual RNA shapes [Reidys et al. 1997], as well as the possibly limited number of actual folds of proteins [Rost et al. 1996].

**Acknowledgements**

B. M. thanks the Austrian Academy for Sciences for generous financial support within APART (Austrian Pro-
gramme for Advanced Research and Technology).

References


Langton Ch. G., Artificial Life, SFI Studies in the Science of Complexity (1989) - and later volumes about the Artificial Life activities.


