

Catalyzed Molecule Replication in an Artificial Chemistry

Thomas E. Portegys

Illinois State University
Campus Box 5150
Normal, Illinois, USA 61790-5150
portegys@ilstu.edu

Abstract

This paper instantiates an architecture for an artificial chemistry featuring continuous physics and discrete chemical reactions. A system of bonded complementary molecular strands replicates in the presence of a catalyst. The catalyst causes the strands to disengage; each strand subsequently replicates its missing complement by bonding to free atoms.

Introduction

Information only has meaning in the presence of an interpreting system, like a musical composition played on an instrument, or a message in a bottle found by a wanderer on a beach. In one sense, a living organism is an arrangement of elements interpreted by the rules of chemistry. The goal of artificial chemistry research is to examine this underlying substrate, and to identify its essential properties. A major impetus for artificial chemistry research is based on the belief that a sufficiently rich physics and chemistry is necessary for the development of artificial life (Bedau et al. 2000, Dittrich et al. 2001).

The focus in artificial chemistry is on processes that characterize chemical reactions: bonding, forces, kinetics, energy, catalysis, etc. The ways that atoms bond determine the shape and stability of molecules, which in turn has a lot to do with their functionality, including their capacity to bond with other molecules. The concept of bonding can be generalized to that of connectedness, which is at the heart of complex systems. Connections can of course take many forms: the weighted links in natural and artificial neural networks; the input-output sequences of attractor functions; cellular automata neighborhoods; and the scale-free networks that characterize the internet are other examples.

The attempt here is to combine a simple continuous Newtonian particle physics with the well-known and readily computable cellular automaton (CA) model as a chemistry architecture. Reaction rules for bonding and other actions are written in terms of CA neighborhood templates. Continuous physics allows molecular movements to occur en masse without the use of CA rules to generate displaced structures, and is more akin to real molecular motion.

This project is a system of reactions that permit bonded complementary molecular strands to replicate in the presence of a catalyst. The catalyst causes the strands to disengage; each strand subsequently replicates its missing complement by bonding to free atoms, a process reminiscent of DNA replication during cell mitosis.

The paper was primarily inspired by three recent efforts: Hutton's artificial chemistry Squirm3 (2002), Smith, Turney and Ewaschuk's JohnnyVon (2003), and my recent work (Portegys 2004), which it extends and improves.

Prior to these, Dittrich et al. (2001) compared a wide range of artificial chemistry approaches, including assembler automata (Rasmussen et al. 1990, Ray 1992, Adami and Brown 1994), Ono and Ikegami's autocatalytic membrane formation (1999), and lattice molecular systems (McMullin and Varela 1997), in which the atoms comprising a molecule map discretely to cellular space.

Squirm3 is a lattice molecular system in which mobile molecules self-replicate using available atoms in a 2D cellular space. Atoms have a fixed type and a variable state. Chemical reaction rules based on type, state and proximity determine the states and bonding status of atoms. Atoms move by jumping from cell to cell through simulated Brownian motion; however, atoms bound into molecules are largely immobile (although this has been since rectified in a later version (Hutton 2004)). Beginning with a soup of inert atoms exposed to state-disrupting cosmic rays, self-replication was shown to spontaneously occur.

In JohnnyVon, by contrast, T-shaped atoms called codons move through a continuous 2D space and interact exclusively through force fields; collisions do not occur. There are two types of codons, distinguished by their field "colors". Each codon is an automaton containing a set of rules governing its field strengths in response to signals represented by the proximity of other codon fields.

My recent artificial chemistry project, Maxwell, involved another example of a concerted set of bonding reactions that produces "foraging" behavior. Bonding is the most observable improvement in this project over Maxwell. In Maxwell, bonds (and thus molecules) were stiff and often created through an awkward "grapple"

reaction that attempted to instantaneously move an atom relative to another without regard to how it was bonded to other atoms. This project features spring-like bonds with variable strengths that smoothly move bonded atoms using forces. This also provides for molecular bending and tension. Another improvement involves discarding the fixed CA grid as a neighborhood metric in favor of “floating” neighborhoods centered over each reacting atom. This eliminates absolute position as a factor in reaction processes.

Description

This system is an artificial chemistry featuring atoms and molecules moving and colliding in a continuous manner in a viscous fluid filling a 2D space. A molecule consists of a bound set of atoms. Chemical reactions are mappings of discrete cellular configurations to parameterized actions on atoms. Actions allow atom creation and destruction, bonding and unbonding to make and break molecules, orientation, and state and/or type change. Time proceeds in discrete steps.

Atoms and molecules

Atoms are elementary particles possessing a type, state, mass, radius, elasticity, charge, orientation, position, and velocity. A molecule is a set of atoms connected by spring-like bonds. Each atom has a maximum of eight bonding sites corresponding to eight possible neighbors (see Reactions section below). A bond may be of variable strength. A bond exerts a force vector on a bonded atom to move it to the “correct” alignment as defined by the bonding site it is connected to. For example, if atom A bonds atom B on A’s “north” bonding site, A will exert a force on B to move it directly north of A at a specific distance. As with physical springs, the force magnitude is proportional to the distortion distance. Bonds are mutual: if A bonds to B then B also bonds to A. The structural integrity of a molecule is dependent on the cumulative effect of the bonded atoms that comprise it.

Forces resulting from bonds, charges and collisions result in continuous velocity and position updates of atoms according to Newtonian physics. Collisions can be variably elastic depending on the involved atoms. For this implementation, rotational momentum is not supported. The motion of bonded atoms can produce molecular bending and tension by distorting the bonds between them.

An atom is oriented in one of the eight compass directions. It may also be in a mirrored state to support symmetric reactions. An atom’s bonding sites are also relative to its orientation. The atom orientation attribute has no counterpart in real chemistry to my knowledge. However, it is quite useful in the reactions that I have modeled.

Space

The space is a 2D bounded rectangle filled with a viscous fluid that impedes the movement of atoms through it. Random atomic movement is accomplished by simulated Brownian motion.

Reactions

Reactions arise from the interaction of atoms. A floating 3x3 Moore neighborhood of cells is centered over each atom, which is considered the “source” atom for reactions within the neighborhood. A reaction specifies a configuration of atom types residing within the neighborhood. Cell residence is determined by whether the center of an atom falls within the cell. Since atoms have an orientation, their neighborhoods are oriented accordingly, as shown by the two atoms in Figure 1.



	NW	N	NE				
	W		E				
	SW	S	SE				
				W	NW	N	
				SW		NE	
				S	SE	E	

Figure 1: Oriented Cell Neighborhoods

A neighborhood cell matches on both atom type and state. A type match value is one of {*atom type*, *empty*, *occupied*, *ignore*}. The neighborhood evaluation is a conjunction of all 9 value matches. An *empty* value specifies that a cell must not contain any atom. This could be used for inhibitory control of a reaction. The *occupied* value matches any atom type. The *ignore* value positively matches any cell condition. A state match consists of {*atom state*, *ignore*}. This allows reactions to be optionally controlled by atom states as well as types.

A reaction consists of the following: {*neighborhood*, *action*, *target*, *next states*, *parameters*}. *Target* is the location of the cell where the action is directed. In the rare instance that the targeted cell contains more than one atom of the appropriate type, the reaction is applied to all of them if possible (it is not possible to bond to more than one from a given source bonding site, for example). The states of atoms involved in a reaction are updated by the next states values. Parameters apply to actions. These are: {*atom type*, *orientation*, *source and target bond sites*, *bond strength*}. The actions are:

- Create atom. Parameters: type and orientation of created atom.
- Destroy atom. Parameter: target atom type.
- Bond atoms. Parameter: target atom type, bonding sites and strength.
- Unbond atom. Parameter: bonding site.
- Orient atom relative to source atom's orientation. For example, if the source atom is oriented east, and it orients a target atom to its west, the atom will acquire a north orientation. Parameters: target atom type and orientation.
- Modify atom type. Parameter: target atom old and new types.
- Modify atom state.

Update cycle

1. Update positions and velocities of molecules based on accumulated forces.
2. Clear forces.
3. Compute new bond, charge, and collision forces.
4. Perform reactions.

Results

The problem involves an orchestration of cooperating reactions that permit bonded complementary molecular strands to replicate in the presence of a catalyst. The catalyst causes the strands to disengage; each strand subsequently replicates its missing complement by bonding to free particles. The double-stranded molecule is shown in Figure 2.

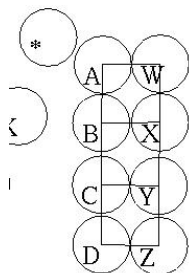


Figure 2: The Replicating Molecule

The atom marked by a '*' in the upper-left of the figure is a catalyst atom. This image was taken just before the catalyst bonded to the west of the 'A' atom to initiate the strand splitting process. Reactions allow a catalyst to initially bond with either with the 'A' or 'W' type of atom. The splitting process begins by passing the catalyst down the length of the strand, as shown in Figure 3.

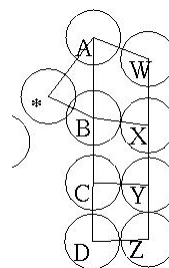


Figure 3: Beginning of Molecule Split

Type 'A-Z' atoms can have the following possible state values: {*free*, *bonded*, *unbonded*, *handoff*}. When situated in a double-stranded configuration, they are in the *bonded* state. When single stranded, they are in the *unbonded* state. Free 'A-Z' atoms are in the *free* state. The *handoff* state is used to coordinate strand splitting. For example, Figure 4 shows a reaction neighborhood that allows atom 'A' to unbind from the catalyst, atom 'B' having "acknowledged" the bond handoff by entering the *handoff* state after bonding to the catalyst.

Ignore	Ignore	Ignore
Catalyst	A (handoff)	Ignore
Ignore	B (handoff)	Ignore

Figure 4: Catalyst Handoff Neighborhood

Figure 5 shows the near-completion of the strand split reactions.

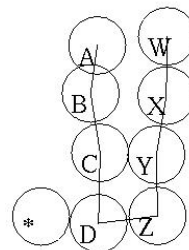


Figure 5: Conclusion of Splitting Reactions

After splitting, the two strands begin to replicate their complements by bonding to free atoms of the appropriate types. The bonding sequence proceeds from bottom to top. Figure 6 shows this process under way.

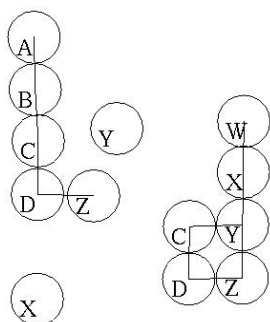


Figure 6: Replication by Bonding to Free Atoms

Experiments show that replication occurs quite readily in a mixture of replicants, free components and catalysts.

Conclusions

One of the aims of this project was to find a way to acquire some of the benefits of a continuous physical medium, such as that in which organic chemistry takes place, and also retain the computability benefits of a cellular automaton as a chemistry framework. For this iteration of the architecture, several important improvements were made, as previously mentioned. In addition, the code was cleaned-up considerably and user-programmability enhanced. Reactions were intentionally defined in declarative rather than procedural form in order to aid in evolvability. The reaction C++ class definition is given in Appendix 1.

The evolvability topic is important in that it relates to the question of how to choose a set of reactions that exhibit interesting or desirable behavior. One of the impressions I got after fastidiously coding the reactions for the replicating molecule and getting it to work was of how fragile and specific the reactions were – the smallest mistake or omission would ruin everything. Real chemistry supports a seemingly endless diversity of molecules and reactions, yet the computational cost to simulate it can be excessive (Zagrovic, et al. 2002); although using a 2D world will certainly lessen this cost. Possible evaluation metrics for an “interesting” chemistry might be:

1. Entropy. Interesting processes, including living ones, appear to exist “on the edge of chaos”.
2. Autocatalytic cycles. Many researchers believe these processes to be the key to the origin of life.
3. Semi-permeable membranes that create special reaction environments.
4. Replication.
5. Symmetry of more complex forms.

As for extensions of the model that may possibly foster more interesting chemical systems, such semi-permeable membranes, some of the possibilities are:

1. The use of variables to symbolize various quantities, such as target atom types and locations, would improve the model’s power and flexibility. An example of the use of variables can be seen in Squirm3 (Hutton 2004).
2. The use of a hexagonal 2D space would eliminate the awkward difference between diagonal and vertical/horizontal orientations.
3. The extra topological freedom of motion inherent in a 3D system should be explored. This would allow for some interesting structures like pores or stomata to exist.
4. Rotation and angular effects could be incorporated into the physics. This might allow en masse alignments of molecules to take place.

The open source C++/OpenGL code is available at www.itk.ilstu.edu/faculty/portegys/research.html

References

- Adami, C.; and Brown, C.T. 1994. Evolutionary learning in the 2D artificial life system *avida*. In Brooks, R.A. and Maes, P., editors, *Artificial Life IV*, Cambridge MA, MIT Press, 377-381.
- Dittrich, P.; Ziegler, J.; and Banzhaf, W. 2001. Artificial chemistries – A review. *Artificial Life* 7:2245-275
- Hutton, T. J. 2002. Evolvable Self-Replicating Molecules in an Artificial Chemistry. *Artificial Life* 8(4):341-356.
- Hutton, T. J. 2004. A Functional Self-Reproducing Cell in a Two-Dimensional Artificial Chemistry Molecules in an Artificial Chemistry. *Artificial Life 9 Proceedings*.
- McMullin, B. and Varela, F.J. 1997. Rediscovering computational autopoiesis. In Husbands, P. and I. Harvey, editors, *Fourth European Conference on Artificial Life*, pages 38-47, Cambridge, MA, MIT Press.
- Ono, N. and Ikegami, T. 1999. Model of self-replicating cell capable of self-maintenance. In Floreano, D., J.-D. Nicoud, and F. Mondana, editors, *Proc. Fifth European Conference on Artificial Life (ECAL'99)*, pages 399-406, Berlin, Springer.
- Portegys, T. E. 2004. An Evolvable Artificial Chemistry Featuring Continuous Physics and Discrete Reactions, *Artificial Life 9 Proceedings*.
- Rasmussen, S., C. Knudsen, R. Feldberg, and M. Hindsholm. 1990. *The coreworld: Emergence and*

evolution of cooperative structures in a computational chemistry. *Physica D*, 42:111-134.

Ray, T. S. An approach to the synthesis of life. In Langton, C. G., C. Taylor, J. D. Farmer, and S. Rasmussen, editors, *Artificial Life II*, pages 371-408, Redwood City, CA, Addison-Wesley.

Smith, A.; Turney, P.; and Ewaschuk, R. 2003. Evolvable Self-Replicating Molecules in an Artificial Chemistry. *Artificial Life* 9(1):21-40.

Zagrovic, B.; Snow, C.D.; Shirts, M.R.; and Pande, V.S. 2002. Simulation of Folding of a Small Alpha-helical Protein in Atomistic Detail using Worldwidedistributed Computing. *Journal of Molecular Biology*.

Appendix 1 – Reaction Structure

```
// Reaction.
// Pattern matching is done against a Moore (3x3)
// neighborhood of cells centered on a particle.
// Various reaction types determine changes to particles
// in the neighborhood.
class Reaction
{
    // Matching neighborhood particle types.
    static const int IGNORE_CELL;
    static const int EMPTY_CELL;
    static const int OCCUPIED_CELL;
    int types[3][3];

    // Matching neighborhood particle states.
    static const int IGNORE_STATE;
    int states[3][3];

    // Reaction type.
    int reactionType;

    // Reaction target location in neighborhood.
    int x,y;

    // Next states of reacting particles.
    int sourceState;
    int targetState;

    // Reaction parameters.
    int type;
    Orientation orientation;
    int sourceBond,targetBond;
    float bondStrength;
};
```