Chapter 1

# A molecular network rewiring rule that represents spatial constraints

**Hideaki Suzuki**
National Institute of Information and Communications Technology
hsuzuki@nict.go.jp

Network artificial chemistry (NAC) is an approach that uses a mathematical graph to emulate molecular interaction in a solvent. To emulate molecules' movement in a 3D space, rewiring rules for NAC's edges must be designed to enable the edges to imitate the encounter relations between molecules or atomic clusters. Our research formulated the 'network energy' representing this constraint and rewired the NAC graph to minimize required energy. Experimental results for the NAC rewiring are compared with a hard-sphere random walk simulation.

## 1.1 Introduction

Chemical reactions are the most fundamental processes proven to cause self-organization and emergence of life. Basic reaction rules are simple (or at least, clarified by modern physico-chemistry), and yet they have enough power to bring about self-assembly of molecules to create such highly functional molecules as ribosomes.

Self-assembly and self-organization are driving forces that create higher structures or functions in the bio-chemical system, and the ability to make self-

assembly/self-organization of molecules happen in a computer would potentially enable creation of highly functional creatures or programs in an artificial medium.

Artificial chemistry (AChem) [9, 8] is a research approach whose goal is to actualize such phenomena by emulating chemical processes in a computer. Based on a relatively strong comparison to real chemical processes, an AChem model can be steadily improved by incorporating real properties of chemicals one by one. Thus, AChem provides us with a more reliable way of modeling bio-molecular systems than the conventional methodology ALife, which depends on human designers to invent models.

Since AChem was first proposed by Fontana *et al.* [9, 10], research on it has mainly focused on mathematical implementation of the *chemical reactions*. Several ways of representing molecules/reactions in a computer or of regarding computational operations as chemical processes have been presented [17, 24, 7, 15, 14, 18]. By contrast, another important environmental factor, the *spatial structure* for the reactions, has attracted relatively little attention, and only a few models are so far known to the community: Ono *et al.* and McMullin's lattice models [13, 11, 12] and Speroni's triangular graph model [16].

Network Artificial Chemistry (NAC) [19, 20, 21, 22, 23] is a new method of artificial chemistry that represents molecular interactions purely with a mathematical graph. Being void of coordinate information for each node, a NAC graph can, in principle, have an edge (denoting interaction) between any pair of nodes (denoting particles). In a real bio-chemical system, however, molecular activities such as collisions and reactions are governed by physical (spatial) constraints, so the tight emulation of those activities with a NAC graph naturally requires the graph to conform to some spatial constraints restricting its topology. This Rapid Communication's method focuses on formulating the NAC graph so it conforms to the constraints. We developed a unified formula for newly introduced network 'energy' and derived a rewiring rule for NAC edges from the criterion of energy minimization.

The organization of the paper is as follows. After formulating the NAC rewiring rule with the network energy in Section 1.2, we conduct NAC simulations in Section 1.3 and discuss the results in Section 1.4. Some concluding remarks are given in Section 1.5.

## 1.2 The Rewiring Rule

### 1.2.1 Network Energy

We disregarded the bonding energy of the NAC edges and considered only the network's spatial constraint energy, $E_\mathrm{s}$, which is defined as the sum of the node degree's diversity term ($\mu$ term) and the second shortest path term ($\nu$ term):

$$E_\mathrm{s} = \frac{\mu}{N}\sum_n |k_n - \overline{k_n}|^\sigma + \frac{\nu}{N}\sum_{<mn>}\left\{\left(\frac{k_m k_n}{\overline{k}^2}\right)^\gamma (l_2)_{mn}\right\}^\alpha, \qquad (1.1)$$

where $\sum_n$ represents the summation of all the nodes in the network, $N$ is the number of nodes, $k_n$ is the degree of node $n$, $\overline{k_n}$ is the node's expected degree (which can vary from node to node), $\sum_{<mn>}$ represents the summation of all the adjacent node pairs, $\overline{k}$ is the average degree of all the nodes, and $(l_2)_{mn}$ is the length of the second shortest path between adjacent node pair $mn$.

Eq. (1.1) was chosen from preliminary experiments and analyses on NAC rewiring that compared the NAC graph's dynamics with the hard-sphere's random walks in a 3D space. Among several candidate formulas for network energy, Eq. (1.1) was the best in that it represents the degree of the network's unnaturalness for a spatial graph, gives edge joining/cutting probabilities similar to those given from the hard-sphere random walks, makes a graph that has the same topological properties as a hard-sphere encounter network (a graph wherein edges occur between contiguous hard sphere pairs) [23], and can be calculated locally. Some of these properties are discussed later in more detail.

Roughly speaking, minimizing the $\mu$ term makes the degrees of the nodes uniform, and minimizing the $\nu$ term shortens the edges' second shortest paths, which are usually short for a hard-sphere encounter network.

### 1.2.2 The Procedure

Using the above constraint energy the NAC graph is rewired by conducting the following two operations at each time step:

**(1)** **[Edge joining]** One terminal node A in the network is randomly chosen, another terminal node B in the network with path length $l$ from A is randomly chosen, and the edge between A and B is joined with the acceptance probability $P(\Delta E_{\rm s})$, which is calculated from $\Delta E_{\rm s}$, an increment of the network energy caused by the joining.

**(2)** **[Edge cutting]** An edge $\overline{\mathsf{AB}}$ in the network is randomly chosen and then cut with the acceptance probability $P(\Delta E_{\rm s})$ calculated from $\Delta E_{\rm s}$, an increment of the network energy caused by the cutting.

Path length $l$ is chosen in proportion to the "free path function" defined as

$$P_{\rm fp}(l) = \exp(-l/\lambda) \tag{1.2}$$

within the range of $l \leq L_{\max}$. $\lambda$ is the "mean free path" in the NAC and the larger the $\lambda$ is, the more a node tends to encounter (create an edge) the farther node. $P(\Delta E_{\rm s})$ is calculated from the *modified* Metropolis criterion:

$$P(\Delta E_{\rm s}) = \begin{cases} \kappa \exp(-\beta \Delta E_{\rm s}) & \text{if } \Delta E_{\rm s} \geq 0 \\ \kappa & \text{if } \Delta E_{\rm s} < 0, \end{cases} \tag{1.3}$$

where $\beta$ is the inverse temperature appearing in Arrhenius's law of chemical reaction velocity theory [3, 25]. A similar criterion was also used in the Monte

Carlo experiments on network rewiring by Burda *et al.* [5] and Berg *et al.* [4], who used energy different from that specified by Eq. (1.1).

When actually calculating $\Delta E_{\rm s}$, we replaced $\sum_{<mn>}$ in Eq. (1.1) with the partial sum $\sum'_{<mn>}$, the summation for the neighboring edges around the target edge $\overline{\mathsf{AB}}$. Here, the 'neighboring edges' include edges directly connected to nodes $\mathsf{A}$ and $\mathsf{B}$ and edges constituting the shortest or the second shortest path between $\mathsf{A}$ and $\mathsf{B}$ for joining or cutting $\overline{\mathsf{AB}}$, respectively. This replacement makes possible the *local* evaluation of an energy increment and, consequently, dramatically reduces the computational cost of the simulation. It was confirmed by preliminary experiments that $\Delta E'_{\rm s}$ calculated from $\sum'_{<mn>}$ is in good agreement with $\Delta E_{\rm s}$ calculated from $\sum_{<mn>}$. Moreover, to minimize the computational cost, we also limited the length of the second shortest paths by $l_2 \leq L_{\max}$.

## 1.3 NAC Experiments

Starting from an initial random network with $N$ nodes and average degree $\overline{k}$, we conducted a Monte Carlo experiment on NAC using the presented rewiring rule and compared the properties of the obtained networks to those of an encounter network created by a hard-sphere random walk simulation in a three-dimensional space [23].

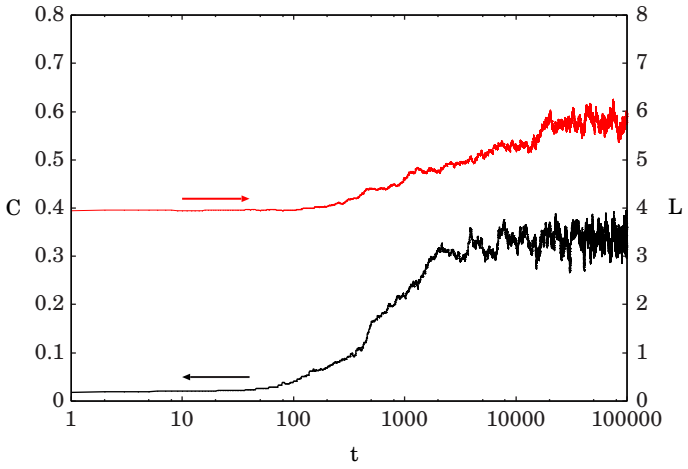


**Figure 1.1**: Temporal change of the cluster coefficient ($C$) and average path length ($L$) during a typical run of the NAC rewiring simulation. Initial values of $C$ and $L$ are equal to theoretical values for a random graph ($C_{\rm rand} \sim 0.018$ and $L_{\rm rand} \sim 3.95$ [2]). Here and in subsequent figures, the parameter values are taken to be $N = 200$, $\overline{k} = 4$, $\lambda = 3$, $L_{\max} = 15$, $\mu = 0.01$, $\sigma = 4$, $\nu = 0.02$, $\gamma = 0.1$, $\alpha = 2$, $\kappa = 0.5$, and $\beta = 1500$.

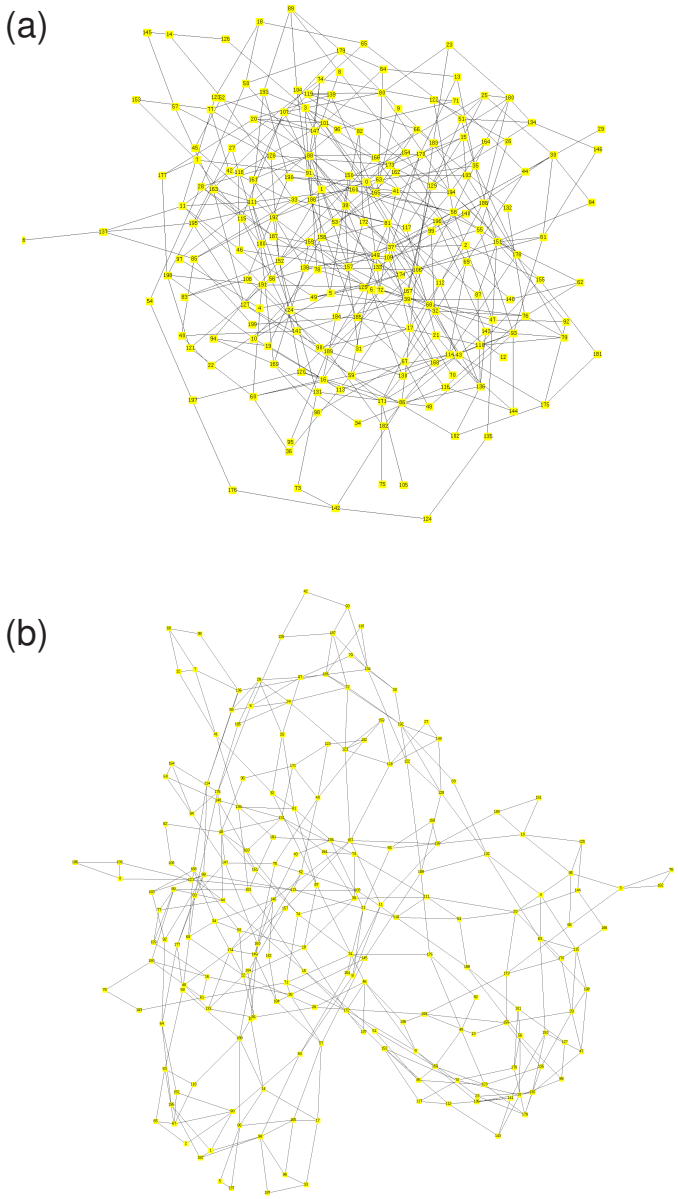


**Figure 1.2**: Two-dimensional drawing of (a) initial random NAC graph and (b) NAC graph after eighty thousand time steps of rewiring operations, obtained by the run in Fig. 1.1. Graphs have average degree (a) 4.0 or (b) 3.91 and are drawn with the same rubberband algorithm of the commercial software, aiSee [1].

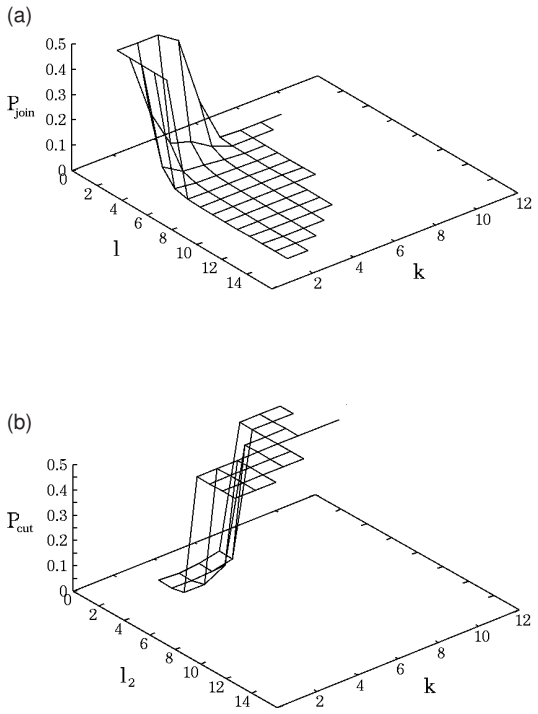


**Figure 1.3**: (a) Edge joining probability $P_{\text{join}}$ as a function of shortest path length $l$ and degree $k$ of a terminal node, and (b) edge cutting probability $P_{\text{cut}}$ as a function of the second shortest path length $l_2$ and the degree $k$ of a terminal node. Results were derived from Eq. (1.3) substituted with the average values of $\Delta E_{\text{s}}$ that are statistically calculated during the NAC simulation run in Figs. 1.1 and 1.2.

The first properties we compared were the cluster coefficient ($C$), average path length ($L$), and the graph's compactness. According to Figs. 1.1 and 1.2, after a sufficient number of rewiring operations, the graph has much larger $C$ than a random graph ($C \sim 18.3 \cdot C_{\text{rand}}$), slightly larger $L$ than a random graph ($L \sim 1.47 \cdot L_{\text{rand}}$), and a compact structure (not separated into loosely connected subgraphs). This is never a trivial result. If we want to have only large $C$ and $L$, we could have the energy represented by a negative weighted sum of $C$ and $L$ (note that $L$ cannot be evaluated locally) and rewire a graph so that the sum might be minimized. However, an example network that minimizes such energy is obviously one that is separated into several complete subgraphs with $C = 1$ and $L = \infty$, which means that it is difficult to augment $L$ while maintaining the entire connectivity of the graph. The presented rewiring rule solves this problem, and the final graph has a large $L$ (Fig. 1.1) and a compact structure (Fig. 1.2) at the same time. The $C$ and $L$'s final values shown in Fig. 1.1 are quantitatively

in good agreement with the hard sphere's encounter network where a typical value of $C$ is about ten to twenty times larger than $C_{\rm rand}$, that of $L$ is about 1.6 times larger than $L_{\rm rand}$, and topological compactness is assured when hard spheres are packed tightly enough in a 3D space.

The second property we compared was the graph's tangledness. As can be seen in Fig. 1.2, the presented rewiring causes a network to have a much smaller number of edge crossing points than the initial random network has. A similar property is observed in an encounter network of hard spheres in a 3D space that has no long edges between spatially distant sphere pairs.

Our final assessment measured the edge joining and cutting probability as a function of a few network parameters. Figure 1.3(a) obtained from the NAC simulation says that the edge joining probability between a pair of nodes (let's call them A and B), $P_{\rm join}$, is large when the shortest path length $l$ between A and B is small and the degree $k$ of A or B is small. This suggests that the minimization of $E_{\rm s}$ tends to cause a pair of nodes to encounter each other when they are close to each other ($l$ ∼small) and there is a large open surface area for the nodes ($k$ ∼small). Figure 1.3(b), on the other hand, says that the cutting probability of edge $\overline{\rm AB}$, $P_{\rm cut}$, is large when the second shortest path length $l_2$ between A and B is large or the degree $k$ of a terminal node A or B is large. Likewise, the results in Fig. 1.3 are in qualitative agreement with those from the random walk simulation [23].

## 1.4 Discussion

Since, in a real bio-chemical system, molecules/atomic clusters are quasi-uniformly distributed in the 3D Euclid space, their encounter network has some topological properties in common with a regular graph. The rewiring operations presented in this paper cause a graph to have some kind of regularity. A graph resulting from these operations is located between a regular network (with $C \gg C_{\rm rand}$ and $L \gg L_{\rm rand}$) and a small-world network (with $C \gg C_{\rm rand}$ and $L \sim L_{\rm rand}$) within the famous Watts and Strogatz scenario [26]. Several reasons are given for this regularity.

First, because of the free path function Eq. (1.2), edges tend to occur between pairs of nodes that are close together, which obviously increases the number of clusters in the graph and increases $C$. (The same mechanism was suggested by Davidsen et al. who pointed out the small-world property of a human acquaintance network [6].)

Second, the uniformity of node degrees required by the $\mu$ term's minimization is a distinctive feature of a regular network.

Finally, the minimization of the $\nu$ term deletes an edge with large $l_2$ (the second shortest path length). This clearly increases $L$ because cutting an edge with $l_2$ changes the length of the shortest path between the edge's terminal node pair from one to $l_2$. In addition, an edge with large $l_2$ tends to have a number of crossing points with other edges in the 2D drawing of a graph and is

likely to augment the graph's tangledness. This is also reduced by the $\nu$ term minimization.

## 1.5 Conclusion

We defined the network energy of the NAC graph that represents spatial constraint in the Euclid space, and formulated an edge-rewiring rule that minimizes the energy. From the NAC simulation using the formulated rewiring rule, it was found that the rule gives some regularity to a graph and causes a graph to have intermediate properties that is between those of a regular network and those of a small-world network. Also, it was shown that the obtained properties of a NAC graph are the same as or similar to those of an encounter network obtained by the random walk simulation that emulates the Brownian motion of hard spheres in a 3D Euclid space.

Analyzing the NAC graph's properties with the statistical mechanical methods used in [5, 4] is a problem to be tackled in the future. Incorporating the edge bonding energy $E_{\mathrm{b}}$ or node properties/categorization and emulating such realistic phenomena as pseudo-lattice formation of water or micelle/vesicle formation by amphipathic molecules are also interesting problems for the future.

## Acknowlegements

This study was supported in part by Doshisha University's Research Promotion Funds, "Academic Frontier, Intelligent Information Science", in Japan.

## Bibliography

[1] aiSee: Commercial software for visualizing graphs with various algorithms such as rubberband. http://www.aisee.com/

[2] Albert, R., Barabási, A.L.: Statistical mechanics of complex networks. Reviews of Modern Physics **74** (2002) 47

[3] Barrow, G.M.: Physical chemistry. Chapters 15-17. McGraw-Hill Education (1988)

[4] Berg, J., Lässig, M.: Correlated random networks. Phys. Rev. Lett. 89(22) (2002) 228701

[5] Burda, X., Correia, J.D., Krzywicki, A.: Statistical ensemble of scale-free random graphs. Physical Review E 64(4) (2001) 046118

[6] Davidsen, J., Ebel, H., Bornholdt, S.: Emergence of a small world from local interactions - modeling acquaintance networks. Physical Review Letters **88**(12) (2002) 128701

[7] Dittrich, P., Banzhaf, W.: Self-evolution in a constructive binary string system. Artificial Life **4** (1998) 203–220

[8] Dittrich, P., Ziegler, J., Banzhaf, W.: Artificial chemistries-a review. Artificial Life **7** (2001) 225-275

[9] Fontana, W.: Algorithmic chemistry. In: Langton, C.G. et al. (eds.): Artificial Life II: Proceedings of an Interdisciplinary Workshop on the Synthesis and Simulation of Living Systems (Santa Fe Institute Studies in the Sciences of Complexity, Vol. 10), Addison-Wesley (1992) 159–209

[10] Fontana, W., Buss, L.W.: 'The arrival of the fittest': toward a theory of biological organization. Bull. Math. Biol. **56** (1994) 1-64

[11] Madina, D., Ono, N., Ikegami, T.: Cellular evolution in a 3D lattice artificial chemistry. In: Banzhaf, W., Christaller, T., Dittrich, P., Kim, J.T., Ziegler, J. (eds.): Advances in Artificial Life (7th European Conference on Artificial Life Proceedings), Springer-Verlag, Berlin (2003) 59–68

[12] McMullin, B., Groß, D.: Towards the Implementation of Evolving Autopoietic Artificial Agents. In: Kelemen, J., Sosik, P. (eds.): Advances in Artificial Life (6th European Conference on Artificial Life Proceedings), Springer-Verlag, Berlin (2001) 440-443

[13] Ono, N., Ikegami, T.: Artificial chemistry: computational studies on the emergence of self-reproducing units. In: Kelemen, J., Sosik, P. (eds.): Advances in Artificial Life (6th European Conference on Artificial Life Proceedings), Springer-Verlag, Berlin (2001) 186-195

[14] Ono, N., Suzuki, H.: String rewriter that allows the maintenance of different types of self-replicators. In: Proceedings of the Fifth International Conference on Humans and Computers (HC-2002) (2002) 173-178

[15] Speroni di Fenizio, P., Banzhaf, W.: Stability of metabolic and balanced organisations. In: Kelemen, J., Sosik, P. (eds.): Advances in Artificial Life (6th European Conference on Artificial Life Proceedings), Springer-Verlag, Berlin (2001) 196-205

[16] Speroni di Fenizio, P., Dittrich, P., Banzhaf, W.: Spontaneous formation of proto-cells in an universal artificial chemistry on a planar graph. In: Kelemen, J., Sosik, P. (eds.): Advances in Artificial Life (6th European Conference on Artificial Life Proceedings), Springer-Verlag, Berlin (2001) 206-215

[17] Suzuki, H.: Evolution of self-reproducing programs in a core propelled by parallel protein execution. Artificial Life **6** N.2 (2000) 103–108

[18] Suzuki, H., Ono, N.: Universal replication in a string rewriting system. In: Proceedings of the Fifth International Conference on Humans and Computers (HC-2002) (2002) 179-184

[19] Suzuki, H.: Spacial representation for artificial chemistry based on small-world networks. In: Pollack, J., Bedau, M., Husbands, P., Ikegami, T., Watson, R.A. (eds.): Proceedings of the Ninth International Conference on the Simulation and Synthesis of Living Systems (Artificial Life IX) (2004) 507-513

[20] Suzuki, H.: Network artificial chemistry – molecular interaction represented by a graph. In: Bedau, M., Husbands, P., Hutton, T., Kumar, S., Suzuki, H. (eds.): Workshop and Tutorial Proceedings of the Ninth International Conference on the Simulation and Synthesis of Living Systems (Alife IX) (2004) 63-70

[21] Suzuki, H.: Mathematical folding of node chains in network artificial chemistry. In: Sixth International Workshop on Information Processing in Cells and Tissues (IPCAT), Proceedings. Held in York, UK (2005) 52-68

[22] Suzuki, H.: Mathematical folding that constructs data-flow computers in network artificial chemistry. In: The 8th European Conference on Artificial Life (ECAL) Workshop Proceedings CD-ROM. Canterbury, UK (2005)

[23] Suzuki, H., Ono, N.: Statistical mechanical rewiring in network artificial chemistry. In: The 8th European Conference on Artificial Life (ECAL) Workshop Proceedings CD-ROM. Canterbury, UK (2005)

[24] Suzuki, Y., Tanaka, H.: Chemical evolution among artificial proto-cells. In: Bedau, M.A. et al. (eds.): Artificial Life VII: Proceedings of the Seventh International Conference on Artificial Life, MIT Press, Cambridge (2000) 54–63

[25] Vemulapalli, G.K.: Physical chemistry. Chapters 23-33. Prentice-Hall Inc. (1993)

[26] Watts, D.J., Strogatz, S.H.: Collective dynamics of 'small-world' networks. Nature **393** (1998) 440-442