アルゴリズム 82－6
（2002．1．24）

Kinase Computing as Algorithmic Chemistry — Part A : Outline of the Computing Processes in Three-dimensional Artificial Chemistry

Jian-Qin Liu and Katsunori Shimohara

Human Information Science Labs., ATR-I, Kyoto, Japan

Abstract. Motivated by the desires of developing molecular computing algorithms for facilitating the emergence of proto-cell communication in artificial life, we have proposed a new algorithmic chemistry method based on kinase computing. In this report, we give an outline of our method and discuss its basic concepts. This is the fundamental work toward further studies of the kinase computing paradigm as a mathematical tool for artificial chemistry.

Keywords: Molecular Computing, Bioinformatics, Artificial Chemistry, Graph Rewriting.

算法的化学としてキナーゼコンピューティング --- Part A ：3次元人工化学の一つモデルの紹介

劉　健勤　下原勝憲

国際電気通信基礎技術研究所　人間情報科学研究所

あらまし　人工生命における原生細胞コミューニケシンズの創発のために分子計算のアルゴリズム開発を目指し、キナーゼコンピューティングによる算法的化学の手法を提案した。本論文ではその基本的な枠組を紹介し、関連するコンセプトについて議論する。さらに人工化学の数学ツールとしてキナーゼコンピューティング·パラダイズムの研究への基盤を築く。

キーワード:　分子計算、生命情報学、人工化学、グラフ書き換え系。

1 Introduction

To understand the complex behavior of artificial life in such systems as mesoscopic artificial chemistry by Peter Dittrich et al. [4], Tierra by Tom Ray [3], Avida by Christ Adami [5], and the cellular automata based system by Ono et al. [7], it is necessary to study the emergence phenomenon at different levels, ranging from molecules to organisms. Fenizio et al. on spontaneous proto-cells by planar graph [2] is significant for exploring systematic design methods of artificial chemistry systems. The successes made in proto-cells by Fenizio et al., Ono et al., and Y. Suzuki et al. have provided promising ways to study "organization" and derived functions [2] in artificial chemistry. To the best knowledge of the authors, for proto-cells' communication, which is crucial for self-regulation in living systems, there is still no rigorous and systematic analysis theory in the discipline of artificial chemistry. Based on the open problem for artificial chemistry labeled as A3.2 in [1], the purpose of this paper is to facilitate the emergence of stable proto-cells capable of communications within the artificial chemistry systems, based on the parallel computation theory with a reasonable coarse-fine degree of physical and biological faithfulness.

In the artificial chemistry paradigms that are used to study the emergence of proto-cells, three main methods for modeling and programming exist as: (1) The cellular automata: These methods have been widely applied, and related studies are intensive as well as extensive [6] , as Fenizio reported in [2]. However, from in the viewpoint of the information quantity, they have disadvantages and limitations exist because of "lack of the generating ability" [2] in the space aspect. Their information capacity is limited, as Fenizio pointed out in [2].
(2) The Euclidean space based methods: The Euclidean representation in these methods is convenient for operations, but in existing methods such as [2], the growth of space becomes n! [2]. It is not an "efficient" algorithm from the viewpoint of computation theory (here, "efficient" means the cost of computing is polynomial or lower than linear, according to the definitions in theoretical computer science). (3) Planar graph based methods: The planar graph based method by Fenizio et al. reduced the cost of the space growth to the linear cost n and is superior to those not done under "spatial conditions" [2]. This is one of the best methods in two dimensional artificial life (chemistry).

However, for exploring the supramolecular structure of artificial chemistry systems, we need to seriously consider the following aspects seriously: (1) In physical forms, the three dimensional space is the basis for cell configuration. The topological constraints have a strong influence on the spatial locations of the molecular objects within the operation processes. (2) In the three-dimensional space, embedding the wave effect into a particle-based artificial chemistry system is useful for going beyond the chemical limitations caused imposed by the neighborhood setting. (3) For studying the cell communication in a physical framework, which is the key to understanding how cells can work stably and collectively, interactions of coupled chain chemical reactions are the causes of increasing complexity. The counterpart of the interaction mechanism in formal systems can be regarded as rewriting on hypergraphs. (4) The quantitative rela-

tionship between the symbolic rules of object interactions and the spatial locations in artificial chemistry systems is tightly coupled. The result of dynamical evolution processes are determined by the balance of both of these two factors.

Accordingly, for a three-dimensional artificial chemistry with emergent communications, we still need to find new mathematical tools rather than the existing methods mentioned above. This leads to the differences between our artificial chemistry method [8] and the general evolution software in artificial life.

2 The Model of Artificial Chemistry

Basically, our model for artificial chemistry is defined as a construct form $< \Xi, \Upsilon, \varpi >$, where the object Ξ is the set of molecules, the interaction Υ is the set of updating rules for objects (e.g., the rewriting rules on hypergraphs or other formal representation equivalent to graph rewriting in computation theory), and population dynamics ϖ is the generating mechanism. The core of the algorithm is given in table 1 by a pseudo-C language representation.

2.1 Basic Concepts

Let L be the category represented by the construct $< Ob(L), Mor(X,Y), M >$, where Ob(L) is the set of hypergraph objects to be rewritten in graph rewriting systems, Mor(X,Y) is the morphism of X into Y (X, Y $\in$ Ob(L)), and M is the law of composition designed by rewriting rule set Qls in [8]. This graph rewriting system is formalized as the following construct: Wls $= < V, T, D, HGs, Z, PTs, Qls >$, where V is the alphabet set, T is the terminal set where T $\subset$ V, D is the set of {0, 1}, HGs is the set of hypergraphs $< HE, U >$ such that HE is the set of hyperedges in HGs, U is the set of vertexes (U $\subset$ V), Z is the set of symbolic (discrete) forms of local values of measures defined in the underlying space where Z $\subset$ V, PTs is the set of "combined flow", i.e., the subset of HGs that has inputs and outputs in U and contains hyperedges in HE with closure property with respect to interaction operations by Qls and owns at least two vertexes and one edge within its domain (set), and Qls is the set of rewriting rules on hypergraphs in HGs. Here, the condition for rewriting is within the conditions of "graph rewriting in topology" [8].

We define the "combined flow" set as the fundamental components executed by rewriting operations: PTs :: = PTs $\cup$ hyperedge with at least three vertexes. The rewriting process becomes: $PTs(hG) \rightarrow PTs(hG')$ s.t. the objects in Wls and the rules of Qls, where hG and hG' denote the hypergraphs in HGs, where the inputs and outputs of PTs located in the space defined in [8] must cover certain positions in one of the unit spheres. The rewriting processes can be divided as the following two classes:

(i) Rewriting for the set level (sub-population):

Qls = {Qinteraction, Qmodification}

s.t. α is the spatial constraint given in this paper,

where Qinteraction (Qi for short) is the operation to generate new hypergraphs by combining (interacting) existing hypergraphs by HR, Qmodification (Qm for short) is the operation to generate new subpopulations by VR, and α is the topological condition for rewriting.

(ii) Rewriting for the object level (single molecular level) that is equivalent to the representation (i) in the sense of function:

The modules and their corresponding flows are rewritten in the form of

PT $\rightarrow$ PT' s.t. g,

where g denotes the rewriting rule set and PT and PT' $\in$ the set of all flows. The rewriting rule set g adopted here consists of four rules: (1) Delete certain vertexes and related hyperedges from the existing ones. (2)Add new vertexes and/or hyperdges to the existing ones. (3)Add new hyperedges to the existing ones without changing the vertexes. (4) Produce new PTs by combining the existing PTs in the form of

PTs 0 (old) + PTs 1 (old) + ... + PTs n (old) $\rightarrow$ PTs (new).

"Graph rewriting in topology" initiated by us [8] is a kind of new rewriting system on hypergraphs under the topological constraints designed. It offers substantial information capacity in three-dimensional space and bridges the gap between parallel computing by graph rewriting and robustness mechanisms of unconventional computing with analog values. The system's features can be summarized as follows: (1) The derived grammar as R.E.: According to the proposition, for graph rewriting is equivalent to the grammar set that is R.E. (2) Computability: It is Turing computable, i.e., capable of universal computation. (3) Adaptation in the rules itself: the graph rewriting is expressed as: PTs (HGs(x)) $\rightarrow$ PT's (HGs'(y)), where x and y belong to the object set before and after rewriting. Here x and y are constrained by the homotopy-based operators given in later subsection. Therefore, the derived concrete rewriting rules will vary depending on the 3D spatial constraints where the objects are located and moved out. (4) P-cost for NP problem solving: The above model is capable of solving 3-SAT problems by p cost both in time and space.

TABLE 1 BRIEF PRESENTATION OF THE ALGORITHM

```
Set-Element-in-SET-from-Matrix(randomset,eleset,chemireactrule, spatcondn)
{ for (i=0; i<Lb,i++)
{
for (j=0, j<La; j++)
{
eleinsetax = elepos (ramdom1, random2, random3);
eleinsetbx = elepos (random4, random5, random6);
eleinsetemp = chemireact (eleinsetax, eleinsetbx);
updating-elepos (random7, raondom8, random9) = eleinsetemp;
} } }
```

This is for the object level of AlCh and the "object-to-object" is encapluated as the basic procedure.

```
Set-Level-Opeartion (..., HEmatrix, Vmatrix, ...)
for (i=0; i<Lb,i++)
{
for (j=0, j<La; j++)
{
cluster-center-calculation(..,delta,...)
neighborhood = neighborhood-calculation();
object-selection ();
VR-rewriting (neighborhood);
HR-rewriting (neighborhood);
population-generation ();
} }
```

This is at the level of graph rewriting.

2.2 Operations of the Model

Operations are carried out on the equivalent classes deduced by the topological constraints for the objects where the PTs are put into the population. Initialization for terminators is made for the objects, and chains of reactions are made on a random basis of normal distribution in 3D. The centers of the clusters that are equivalent classes for the rewriting operations are updated in the rewriting processes. The homotopy mapping condition is exerted on the VR, and the neighborhood is made on the basis of the wave effect for the new generating objects where the centers are randomly treated as the particles. The newly generated population is divided into different sub-populations, where the manifold features are kept by the Heegaard splittings for pushing into. Here, the manifold features are kept for the traits and sampled for the inputs and outputs of PTs on the unit spheres with the centers in nearest classes, and certain objects among them can be assigned with Gödel numbers.

Vertex Replacement Operations We explain the model's operations by using two classes:
(1)Topological (Homotopy) Condition for VR Rewriting
The elements (objects) in the neighborhood are located in different matrixes (arrays). Here, the operation is deterministic for constructive dynamics, and the object is randomly selected in the spatial domain under the constraint discussed in this paper. The processes for the object-level are encapsulated as a basic procedure (operator) for rewriting. Based on this OOP arrangement, the three rules are carried out for interactions of the reactions,and replacement of the vertexes and hyperedges among them is done under the spatial conditions of GRiT[8] and others in this paper. In brief, the above process for evolution of particle objects can be summarized as the matrix formular:

M (t+1) = M(t) Qv GR(t),

where M(.) is the matrix particle, Qv is the abstract operator for Q rules above, and GR (.) is the transition matrix that can be empirically obtained from the observation of the related samples. Here, evolution means the physical changes of particles in space.

(2)Wave Effect in Neighborhood for VR Operations (in Three-dimensional Rewriting)

The neighborhood is calculated based on the wave effect in three-dimensional space. Aside from the particles' interacting in the neighbordhood by contacting them each other, the wave effect for interacting is embedded into our model. Here the basic operation is defined as:

x @ y $\rightarrow$ z,

where x, y and z belong to the set of objects in artificial chemistry systems (molecules or more generalized form as particles).

Here, the objects fall into the space with the following parameter conditions of a domain based on the Schrödinger equation: (a) σ – the scatter area, where the new molecules enter; (b) ψ – the wave locations of the Schrödinger equation, i.e., the domain where the particles are constrainted; (c) Δ – the change in momentum, which indicates the factor affecting on the state of population dynamics; (d) r – the direction vector for the 3D space. The area where the objects move to is rotated as a fan face in the three-dimensional space.

We assume that the symmetry condition of the scattering of particles and the spin can be ignored under certain conditions, so the wave function is the solution obtained from the Schrödinger function, such that

$\psi \sim exp(ik * r) + f * exp(-ik' * ai) * exp(ikai * exp(ikr)/r + o(1/r2),$

and the scattering amplitude is

$f(\theta) = f * exp(-i\Delta ai)$

$\Delta = k' - k.$

Here, the positions in the three-dimensional space are constraints for the objects selected and generated in graph rewriting.

After initialization, the objects are clustered and centers are formed. We define the vertex in our graph rewriting model as:

LS-vertex is a vertex with vector

{ A', B', C'}, where A' – the name of the vertex, B' – the real value assigned to the vertex, C' – the measurement assigned to the vertex.

Accordingly, we construct a set of real valued points for the LS-vertexes as a set derived from the homotopy group, under the condition of constraints related to ς. Consequently, the mapping Hls from the vertex set in the hypergraph at one moment to the vertex set at another moment is quantitatively described in the topological space and constructed by the continuous mapping:

Hls: Xls * [0,1] $\rightarrow$ Yls,

where fls,gls: Xls $\rightarrow$ Yls denotes the continuous mapping,

and

for x $\in$ X , there exists that Hls (x, 0) = fls (x) and Hls (x, 1) = gls (x).

After the homotopy mapping of the space in which the local concentrations are defined is constructed, the output is obtained in the following form, according to its effect on the pathways:

fls (.) = fls (h1 * h2 * ... * hl (.)),

where the operation is derived by set Qls and executed on function objects. h1 (.), h2 (.), ..., hl (.) are precedent pathways, which consist of those PTs transited into the current PTs. Therefore, fls (.) generates a set of pathways that covers the input nodes as well as all of the routes from the input nodes. The indication measurement (e.g., q) defined in our rewriting model is valued in [0,1]. Here, the new objects are generated in positions within the above homotopy mapped subspace. The homopoty mapping has specified the spatial locations for the new generated objects of population dynamics. The neighborhood within the sigma wave face is used here.

After initialization as random values, the new objects generated fall into the partitions by Heegaard operations, where clusters can be made. The VR can be used for increasing or reduction.

Based on the thresholding for statistical selection, the space is divided into subspaces for allocating objects into the subsets where the cluster centers are updated. These objects are generated and exited as vertexes and put into the rewriting. This is one of the necessary topological constraints influencing on VR.

Interactions of PTs for HR Rewriting Let X be the topological space. With respect to x0, ... , xn $\in$ X, n > 1, the processes of the operations from Qls in the topological space lead to the following:

bn = s = (s1, ... , sn) $\in$ Rn,

where si > 0.

Here, the vector b corresponds to the vector in the topological space with dimension n,

(x1, x2, ... , xn) $\in$ R,

with the boundary in the space defined by a related homotopy measurement and the unitary sphere.

Based on the boundary mentioned above, we define

$< \sigma 1, \sigma 2 >$,

where $\sigma 1$ – the set of points that consists of real valued components corresponding to the vertexes in the combined flow, and $\sigma 2$ – the above-defined operator.

In this construct, the set that refers to the vertexes in the transition processes of the combined flows' rewriting is inferred as

$< \varsigma$', QLS s.t. $\varsigma >$,

where ς is the condition that those pathways can reach the unitary sphere. The main operations for mapping to the unitary sphere S(n) with dimension n include Qs for the points (LS-vetexes of combined flows) in the set that converge to the "net" with the definition of topology and that guarantee a related mapping

capable of covering all of the points on the unit spheres whose centers are in the Euclidean space. The cluster centers are for the centers of the unit spheres, and the hyperedges start from the objects located on the unit spheres and end on the unit spheres. HR is exerted for these hyperedges for the connections of the objects, which indicates continuing chemical reactions with stable circles. On the spheres, we obtain number $n \in N$, and within the spheres $r \in [0,1]$. The Gödel number is assigned by Gödelization with TM computability. The objects on the sphere are assigend by these. The homotopy group is defined based on the relation of homotopy mapping.

3 Conclusions

Based on graph rewriting, kinase computing has been applied as a new computing method for algorithmic (artificial) chemistry. Its advantages include its information capacity in three-dimensional space and parallel structures in computing. Future work will be focussed on the quantitative relation between the neighborhood size for interacting and the randomness degree (probability) of the Qs operations' selection. We are also making efforts toward the emergence of proto-cell communications by kinase computing.

References

1. Mark A. Bedau, John S. McCaskill, Norman H. Packard, Steen Rasmussen, Chris Adami, David G. Green, Takashi Ikegami, Kunihiko Kaneko, & Thomas S. Ray, Open problems in artificial life, Artificial Life, 6, 363-376, 2000.
2. Pietro Speroni di Fenizio, Peter Dittrich, and Wolfgang Banzhaf, spontaneous formation of proto-cells in an universal artificial chemsitry on a planar graph, J.Kelemen and P.Sosi'k (Eds.): ECAL 2001, LNAI 2159, pp.206-215, 2001.
3. Thoms S. Ray, An evolutionary approach to synthetic biology: Zen and the art of creating life, Artificial Life, 1 (1/2): 195-226. MIT Press.
4. Peter Dittrich, Jens Ziegler, Wolfgang Banzhaf, Mesoscopic analysis of self-evolution in an artificial chemistry, in: Proceedings of the Sixth Inetrnational Conference on Artificial Life, Christoph Adami, Richard K. Belew, hiroaki Kitano, and Charles E. Taylor, Editors, pp.95-103, MIT Press, Cambridge MA, 1998.
5. Christ Adami and Titus Brown, Evolutionary learning in the 2D artificial life system "Avida", Proc. of ALIFE IV, pp.377-381.
6. W.Banzhf, P.Dittrich, and B. Eller, Selforganization in asystem of binary strings with topological interactions. Physica D, 125:85-104,1999,
7. Naoaki Ono and Takashi Ikegami, "Artificial Chemistry: Computational Studies on the Emergence of Self- Reproducing Units.", In Proceedings of 6th European Conference on Artificial Life (ECAL01) (Eds. Keleman, J and Soki'k, P.), 186-195, 2001, Splinger, Prague, Czech Republic.
8. Jian-Qin Liu and Katsunori Shimohara, Graph Rewriting in Topology I: Operators and the Grammar, SIG-FAI-A101-4(9/21), Sept.21, 2001, pp.21-26.