

# 可変構造型デバイスを用いた神経回路網シミュレータの実現に関する検討

長名 保範 吉見 真聡 天野 英晴  
慶応義塾大学

## 概要

細胞内の各種物質の濃度を高精度かつ高速に測定する実験技術の進歩に伴い、定量的データに基づく数理的な生化学システムのモデルを用いた研究が盛んになっている。現在は多細胞系のシミュレーションは生物学的にも計算機の能力的にも困難とされているが、生物学的な困難さは今後 10 年程度の間には解決されるものと考えられる。本稿では、現在開発中の、FPGA を用いた生化学シミュレータをベースとした多細胞系、あるいは神経回路網をターゲットとしたシミュレータの可能性についての検討結果について述べる。

**Keywords :** 生化学シミュレータ, FPGA, ReCSiP, 多細胞系, 神経回路網

## A Feasibility Study on Reconfigurable Multi-cellular System Simulator

Yasunori Osana, Masato Yoshimi, and Hideharu Amano  
Keio University

### Abstract

Recent innovations in biological experiments enabled quantitative measurement of metabolites in cellular systems. These methods help development of mathematical modeling and simulation of cellular systems. Modeling and simulating multicellular system is difficult due to biological and computational reasons, but biologically, modeling multicellular systems will be possible in near future. In this paper, the possibility of an FPGA-based simulator with the capability to simulate multi-cellular systems is discussed.

**Keywords :** Biochemical Simulator, FPGA, ReCSiP, Multi-cellular System, Nerve System

## 1 背景

近年、質量分析計に代表されるような、細胞内に存在する各種の物質の濃度を高速かつ定量的に測定する技術が進歩し、実験によって蓄積されたデータをもとに生化学反応系の数理モデルを構築・シミュレーションすることが生命システムの解明に重要な役割を果たすようになりつつある。

現在はモデル・シミュレータともに空間のサイズを持たない、均一な系を想定してモデリング・シミュレーションを行うのが一般的である。これは、2次元や3次元の広がりを持つモデルを構築すること自体が生物学的に困難であることと、仮にモデルが与えられたとしても、シミュレーションを行うのに十分な計算能力を手に入れるのが困難で

あるという、ふたつの側面からの理由による。

他分野でモデリングとシミュレーションが重要な例として電子回路設計を挙げてみる。電子回路の設計においては設計の進捗段階に応じてさまざまなシミュレーションが行われており、LSIなどの大規模なシステム設計において各種のシミュレータはもはや欠くべからざる存在である。電子回路設計にまつわるツールとしては、図 1 に示すように、システムレベルでの設計に用いられるような高位記述言語処理系から、RTL やゲートレベルでのシミュレーションに用いられるシミュレータと記述言語、素子レベルでのシミュレーションに用いられるシステムと、デザインの前層に応じてさまざまなシミュレーション方式が用意されており、それぞれの方式は全く異なるものとなっている。

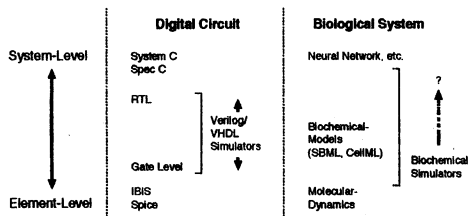


図1: システムモデルとシミュレーション方式

一方、生化学モデルにおいては、そのモデリング自体が新しい研究分野であることと、さまざまな理工学分野で従来から行われてきたさまざまなモデリング・シミュレーションに比べて問題がずっと複雑であることが、ツールやモデルの開発を困難にしている [1]。神経回路網をモデル化するニューラルネットワークのような高位のモデルや、タンパク質の作用を計算機を用いてシミュレートする分子動力学シミュレータのような下位のツールはさまざまな場面で用いられているが、その間には大きなギャップがあり、反応速度式を基盤として細胞内のシステムをモデリング・シミュレーションする生化学シミュレータや、システムを記述するための言語として SBML[2], CellML[3] などがその間をつなぐツールとして数多く開発されており、一定の成果を挙げているが、いずれも多細胞系のような大規模な系をシミュレーションするには十分な能力をもたない。

2次元、あるいは3次元の広がりを持つ系のモデリングやシミュレーションは困難であるが、神経細胞内での信号伝達物質としての  $Ca^{2+}$  の濃度変化を3次元モデルでシミュレーションした例 [4] などもあり、今後数年のうちに空間的な要素を含むモデルが作られるようになることが予想される。

生化学モデルはその複雑さから、並列化した際のスレッド間の通信が複雑になり、PC クラスタや分散共有メモリ型マルチプロセッサなど、旧来の並列システムでは十分な性能を得ることが困難であり、FPGA を用いてこれを高速化する目的で ReCSiP (REconfigurable Cell Simulation Platform)[5] が開発されている。現在の ReCSiP の実装では多くのシミュレータと同様に、空間的広がりを持たない、均一に混じり合っている系しか取り扱うことができないが、本稿ではこれを3次元、かつ多細胞な系に拡張する方式を提案し、多細胞系や神経回路網

のシミュレーションに適用する可能性について考察する。

## 2 現行の ReCSiP の構成

### 2.1 生化学シミュレーション

細胞内で発生する個々の化学反応について、反応に関わる分子の濃度を変数とし、反応速度を微分方程式 (反応速度式) で表すことによって、ある初期条件を与えられた場合の時系列の濃度変化を計算してシステムの挙動を知ることができる。反応速度式は反応機構によって決定され、反応速度式中の定数は各反応に固有のパラメータによって決定される。

図2に、反応系の例を示す。この反応系は  $S_1 \sim S_9$  の9つの物質と、 $R_1 \sim R_5$  の5つの反応で構成されており、 $R_1, R_4, R_5$  の3つの反応 (Reaction type A) と  $R_2, R_3$  の2つの反応 (Reaction type B) はそれぞれ同様の反応機構であるため、この反応系のシミュレーションを行うためにはふたつの反応速度式が解ければよい。

ReCSiP では、このような系のシミュレーションを行うために、図3のような回路を FPGA 上に構成して計算を行う。回路は、反応速度式を解くためのモジュールである Solver と、それを接続するスイッチから構成されており、データは Solver 内部のメモリに保持される。ある Solver が他の Solver に存在するデータを必要とする場合は、スイッチを経由してデータの転送を行うことで、複数の反応速度式で変数を行うことができる。

ReCSiP では、FPGA を PCI カード上に搭載し、初期条件の設定や時系列のデータ取得は PCI バス

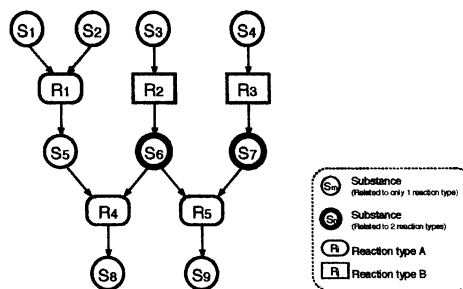


図2: 反応系の例

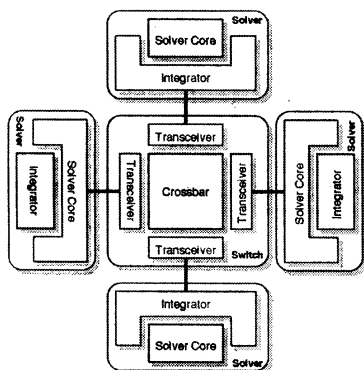


図 3: ReCSiP の回路構成

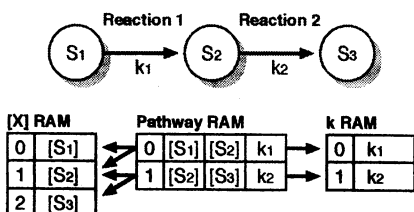


図 4: Pathway RAM による制御の概念図

経由でホスト PC から行うことを想定しており、リファレンスボードとして ReCSiP-2 Board を開発済みである [6]。

## 2.2 Solver の構成

ReCSiP では、個々の反応速度式を用いてシミュレーションを行うモジュールを Solver と呼ぶ。Solver は、数値積分を行うための積分モジュールと、反応速度式を解くための Solver Core と呼ばれるモジュールをそれぞれひとつ組み合わせる構成される。

Solver Core は、ひとつあるいは数種類の反応速度式を解くことができるモジュールで、いくつかの単精度浮動小数点演算器を組み合わせ、静的にスケジュールされたパイプラインによって処理を行う。

積分モジュールは Euler 法、Heun 法 (修正 Euler 法)、Runge-Kutta 法などが用意されており、精度・安定性や実行時間を考慮して選択することができる。これらはいずれも、内部に各分子の濃度を保持

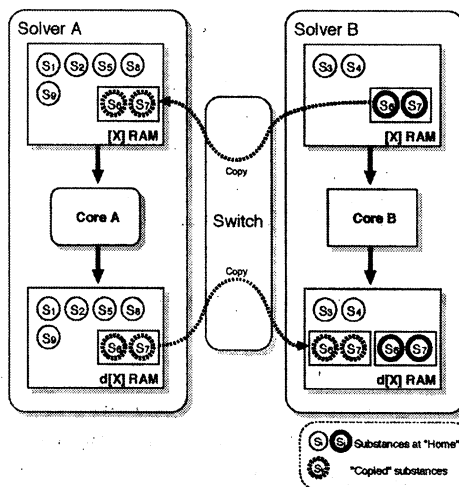


図 5: Solver への分子割り当て

する [X] RAM と、各時間刻みにおける濃度変化を保持する d[X] RAM のふたつのメモリを使いながら Solver Core へのデータ入出力を行うことで、時系列での濃度変化を数値的に求める処理を行う。[X] RAM や d[X] RAM へのアクセスなどは、ポイントの配列を格納したメモリ (Pathway RAM) によって個々の反応を表現し、これによってメモリへのアクセスを制御する (図 4) ことで、柔軟な処理を実現することができる [5]。Pathway RAM によって、反応速度式を解く順番は静的にスケジュールされ、時間刻みごとに同じ処理が繰り返される。

## 2.3 複数の Solver の利用

シミュレーション対象となる反応系に複数の反応機構が含まれており、それらをすべて解くことのできる Solver Core が存在しない場合、複数の Solver を接続して使用する必要が発生する。一連の反応式はひと組の連立常微分方程式として解かれているため、複数の Solver を用いる場合には、複数の Solver にまたがって利用される分子種については、データの一貫性を維持する必要がある。このような場合には、各分子ごとにいずれか一つの Solver を “Home” として定義し、他の Solver が計算にその濃度を必要とする場合には Solver 間転送スイッチを用いて該当する Solver の [X] RAM にデータを転送し、Home 以外の Solver で発生した濃度変

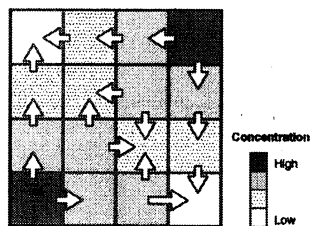


図 6: 拡散による物質の移動

化の大きさは Solver 間転送スイッチを通じて再び Home へと送り返され、そこで積分が行われる。

例として図 2 に示した反応系を解く場合の状況を図 5 に示す。 $S_6, S_7$  が共有される分子であり、それらの Home は Solver B であるので、まず Solver A の [X] RAM に濃度がコピーされ、Solver A の Core が処理を終了した後、Solver A で発生した  $S_6, S_7$  の濃度変化分が Solver B に転送され、そこで積分される。各 Solver の動作は、それぞれの Pathway RAM によって静的にスケジュールされているため、スイッチの動作も静的にスケジュールすることができ、Solver と同様に、転送元や転送先などをポインタ配列として格納したメモリによって制御される。3 つ以上の Solver が用いられる場合には、後述するように多ポートのクロスバススイッチを用いて、すべての Solver を相互に接続する。

### 3 3次元/多細胞系の取扱い

空間的に広がりを持つモデルの場合、前節で述べた反応速度式による濃度変化だけでなく、図 6 のように、対象のシステムを細かな区画に区切った上で濃度勾配と拡散方程式に従った物質の移動も計算することになる。さらに、細胞膜や核膜、および細胞内小器官の外側の膜などの生体膜が存在する場所では各種のチャネルなどを含めた物質の移動をも計算に含める必要がある。

#### 3.1 ReCSiP での取り扱い方式

空間を区切ってつくられた計算単位となる区画間での物質の移動は、図 7 に示すようにある種の反応として取り扱うことでひとつの大きな反応系として取り扱うことが可能であり、現在の ReCSiP

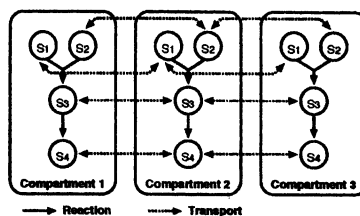


図 7: 物質の輸送と反応

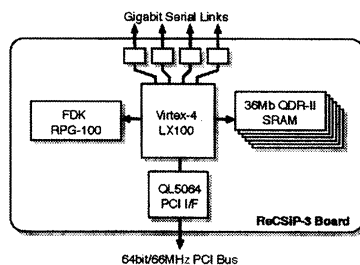


図 8: ReCSiP-3 ボードの概要

の実装においても不可能ではない。図ではすべての区画に同じ反応経路が描かれているが、これは区画ごとに特有の反応経路が存在する場合や組織を隔てる膜が存在する場合でも同じであって、膜による物質の輸送なども反応として扱えばよい。

拡散や膜輸送を取り扱うための Solver 数の増加は FPGA の面積増大によって十分に解決しうる範囲であると予測できるが、区画数に比例して起こる消費メモリ量の増大が問題となる。区画数は  $10^2$  から  $10^4$  程度のオーダーになると考えられ、現在の実装利用している FPGA の内部メモリだけではとうていすべてのデータを格納することができない。

これについては、将来的に FPGA 上のメモリサイズが拡大したとしても、商用デバイスではロジック-メモリのサイズ比はそれほど大きく変化しないものと考えられ、外部メモリを利用する方法の開発は必須事項であると考えられる。

#### 3.2 複数の FPGA を用いたシステム

区画数や、シミュレーション対象の系に含まれる細胞数が増加すればそれに比例して計算量も増大するため、十分な処理能力を得るためには複数の FPGA の使用が必須となる。

立方体の区画が連続するだけの場合には、ひとつの FPGA はいくつかの区画から構成される立方体あるいは直方体のブロックの処理を担当し、面で接する6つのほかのブロックの処理を行う FPGA に接続すればよい。

通常多細胞系のシミュレーションにおいては、細胞の表面全体が他の細胞と接するため、システム全体にわたって、各区画に隣接する6区画との相互接続性を保たねばならず、多数の FPGA を複数の基板に実装した場合にはボード間の相互接続が複雑になることが予想されるが、神経回路網のように、細胞と細胞の接触点が特定の部位に限られ、細胞の表面全体と外部との相互作用を考慮しなくてよい場合は、細胞単位でひとつの FPGA、あるいは基板でシミュレーションを完結させることで、FPGA 間、あるいは基板間のリンク数を少なく抑えることができる。

したがって、FPGA 間のリンクはただか6本あればよいことになる。近年の FPGA は毎秒数ギガビットに達する高速なシリアル伝送チャネルを多数持っており、これによって FPGA 間を接続することで、6本程度の双方向リンクならば容易に構築することができ、FPGA 間の相互接続については既に問題ないレベルであるといえる。

現在設計を進めている ReCSiP-3 ボード(図8)には、Serial ATA のケーブル/コネクタを利用した双方向 1Gbps 程度の差動シリアルリンクを数チャネル実装する予定であり、これによって多数のボードをホストの PCI バスを占有することなく高速に並列動作させることが可能になる。

### 3.3 ReCSiP の方式のメリット

ReCSiP が高い処理能力を発揮できるのは、FPGA 上に構成した複数の深いパイプラインを、相互の通信を含めて静的にスケジューリングして動かすためである。現在の実装では、単一の FPGA での処理だけを考慮した設計になっているが、複数の FPGA を用いた場合にも同様に通信を含めたすべての手続きを静的にスケジューリングして動作させることは可能である。この場合、ボード間でのクロック周波数のわずかなずれやデータの転送にかかる遅延時間を吸収するために、FPGA 内部では1クロックサイクルで済む転送処理に数クロックサイクルを要するなど、若干の処理時間の増加を伴うが、こ

のためのオーバーヘッドは多くても2~3%に収まると考えられ、性能に大きく影響するものではないといえる。

一方、ソフトウェアでこのような大規模な系のシミュレーションを行う場合には、並列システムを用いることが事実上必須であり、その際には FPGA を用いる場合と同様にノード間での変数の同期などのための通信が必要になる。しかし、ソフトウェアの場合には MPI などのより複雑な通信機構を用いることになり、オーバーヘッドが大きくなるため、より通信コストの低い FPGA による処理の優位性が増すと考えられる。

### 3.4 システムの実装形態

既に SRC-7[7] や BEE2[8] のように多数の FPGA を用いた科学技術計算向けの計算機が出現しているが、今後の FPGA あるいはそれに類する reconfigurable device の回路容量増大によって、より複雑なシミュレーションにも FPGA を適用できるようになるものと考えられる。

空間的広がりを持つ生化学モデルは細胞内のさまざまな器官や膜などを内包するため、モデルが非均質になり、細かな区画に区切った場合には各部分を処理するノードには非均質な処理が配分されることになる。しかし、これについては FPGA 内部の回路を、各 FPGA に割り当てられる区画の性質にあわせて個別に構成することで容易に対応することができるため、基板設計から上位のレイヤでは均質なシステム構成をとることができる。

基本的には図9の(a)のように、FPGA をボードに2次元メッシュ状に配置し、それを縦方向に結合し、(b)のようにシミュレーション対象の空間を各 FPGA に割り当てるのが妥当であろうと考えられる。さらに、各 FPGA ではこの空間を細かな区画に区切り、システム全体として(c)のようにひとつまたは複数の細胞モデルを構成する。システム全体にわたって均質な3次元メッシュを構成することが望ましいが、この構成をある程度のサイズで区切り、その間を高速なリンクで結合することで、一定のサイズの計算ノードを多数結合するような使い方をすることも可能であろう。

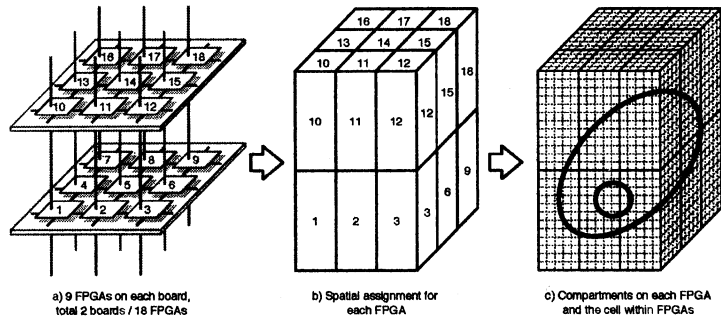


図9: システムの構成とFPGAへの空間割り当て

## 4 まとめ

本稿では大容量のFPGA、あるいはそれに類するプログラマブルデバイスを用いて、3次元的な構造をもつ生化学モデルをシミュレーションする可能性について議論した。結果として、(1)現在一般的に用いられているような連立常微分方程式を用いたモデルに偏微分方程式などによる物質の移動を加えても、FPGA等によるシミュレーションが可能であること、(2)静的なスケジューリングによる効率のよいパイプライン処理、Solver間通信は複数のデバイスを用いた場合にも問題なく行えること、(3)対象の系が空間的に不均一な構造をもっていたとしても、デバイスの柔軟性により、基板などを含めたハードウェアそのものは均一な構造をとることができること、などの基本的な点を確認することができた。

生物学的な側面からはモデルの構築、アーキテクチャ的な側面からはハードウェアの実装ができるようになれば、現在行われているニューラルネットワークなどの抽象化された知的情報処理モデルの研究と、具体的な生命現象の研究としての生化学シミュレーションの橋渡しをしていくことが可能になっていくと考えられる。

## 参考文献

- [1] Kouichi Takahashi, et al. Computational challenges in cell simulation: A software engineering approach. *IEEE Intelligent Systems*, pp. 64–71, Oct. 2002.
- [2] M. Hucka, et al. Evolving a lingua franca and associated software infrastructure for computational systems biology: The systems biology markup language (sbml) project. *IEE Systems Biology*, Vol. 1, No. 1, pp. 41–53, 2004.
- [3] Catherine M. Lloyd, Matt D.B. Halstead, and Poul F. Nielsen. CellML: its future, present and past. *Progress in Biophysics and Molecular Biology*, Vol. 85, pp. 433–450, Jun 2004.
- [4] Charles C. Fink, Boris Slepchenko, Ion I. Moraru, James Watras, James C. Schaff, and Leslie M. Loew. An image-based model of calcium waves in differentiated neuroblastoma cells. *Biophysical Journal*, Vol. 79, pp. 163–183, July 2000.
- [5] Yasunori Osana, Tomonori Fukushima, Masato Yoshimi, and Hideharu Amano. An FPGA-based acceleration method for metabolic simulation. *IEICE Trans. on Information and Systems*, Vol. E87-D, No. 8, pp. 2029–2037, Aug. 2004.
- [6] ReCSiP Project Team. ReCSiP project frontpage. <http://recsip.org/>.
- [7] SRC Computers Inc. SRC-7 reconfigurable general purpose computing system. <http://www.srccomp.com/>.
- [8] Chen Chang, John Wawrzynek, and Robert W. Brodersen. Bee2: a high-end reconfigurable computing system. *Design & Test of Computers*, Vol. 22, No. 2, pp. 114–125, Mar. 2005.