JAIST Repository
https://dspace.jaist.ac.jp/

Title	Particle System を用いた生化学反応シミュレーション
Author(s)	玉樹，真一郎
Citation	
Issue Date	2002-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/365
Rights	
Description	Supervisor:小長谷 明彦，知識科学研究科，修士

JAIST
JAPAN
ADVANCED INSTITUTE OF
SCIENCE AND TECHNOLOGY

Japan Advanced Institute of Science and Technology

修 士 論 文

particle system を用いた 生化学反応シミュレーション

指導教官　小長谷明彦 教授

北陸先端科学技術大学院大学
知識科学研究科 知識システム基礎学専攻

050056　玉樹真一郎

審査委員：　小長谷 明彦 教授　（主査）
佐藤 賢二 助教授
本多 卓也 教授
中森 義輝 教授

2002 年 3 月

Copyright © 2002 by Shinichiro Tamaki

目 次

1 序論 **1**

1.1 本論文の背景と目的 1

1.1.1 バイオインフォマティクスの現状 2

1.1.2 人工生命と複雑系 3

1.1.3 生化学シミュレーションの現状 5

1.1.4 モチベーション 6

1.2 本論文の構成 7

2 本論文に関する従来研究 **9**

2.1 はじめに 9

2.2 生化学反応による空間パターン形成 9

2.2.1 ショウジョウバエの胚発生 9

2.3 人工生命と複雑系 11

2.3.1 von Newmann の自己複製オートマトン 11

2.3.2 Chris Rangton の Adam Loop 12

2.4 シミュレーション手法 12

2.4.1 セルオートマトン 12

2.4.2 常微分方程式系 17

2.4.3 偏微分方程式系 20

2.4.4 ペトリネット 24

2.4.5 既存シミュレータの性質のまとめ 25

2.5 おわりに 26

3 Particle System **28**

3.1 はじめに 28

3.2 基本設計 28
3.2.1 化学反応のモデル化 28
3.2.2 Particle のデータ構造 29
3.2.3 反応の流れ 30
3.3 熱運動のモデル化 30
3.4 化学反応式のモデル化 30
3.5 化学反応の基本アルゴリズム 31
3.6 データ構造 32
3.7 アルゴリズム詳細及び実装と高速化 32
3.7.1 データ構造の詳細 33
3.7.2 ランダムな熱運動計算の詳細 34
3.7.3 化学反応アルゴリズムの詳細 35
3.8 シミュレーション例 36
3.9 おわりに 37

4 Particle System と微分方程式の比較 40
4.1 はじめに 40
4.2 反応確率とシステムの挙動 40
4.3 反応半径とシステムの挙動 41
4.4 反応確率と反応半径の関係 41
4.5 定量化 Particle System 42
4.5.1 概略 43
4.5.2 パラメータ決定 45
4.5.3 結論 47
4.5.4 微分方程式と一致しない場合 48
4.6 おわりに 50

5 Particle System のゆらぎ 52
5.1 はじめに 52
5.2 ゆらぎの例 52
5.3 Particle の個数とゆらぎ 52
5.3.1 結果 53
5.4 Particle の密度とゆらぎ 53
5.5 被覆率とゆらぎ 54

5.5.1 結果 55
5.5.2 考察 56
5.6 おわりに 56

6 結論 **60**
6.1 はじめに 60
6.2 Particle System 60
6.3 本論文の結論 61
6.3.1 Particle System と微分方程式系の一致する点 61
6.3.2 Particle System と微分方程式系の一致しない点：ゆらぎの評価 . . 61
6.3.3 Particle System の意義 62
6.4 今後の課題 62

A 大数の法則 **65**

図 目 次

1.1 バイオインフォマティクスの流れ [3] 3

2.1 ショウジョウバエの胚発生における体節パターン形成にかかわる遺伝子産物の濃度分布（[20], 図 2-2 より）：(a) *bicoid* と *nanos* 母系 mRNA が形成する勾配。(b) ギャップ遺伝子の発現の様子。 10

2.2 ショウジョウバエ初期胚にみられるペアルール遺伝子の発現パターン（[23], 図 19-41 より）：(a) 受精 2.5 時間後の胚。全体が 4 分割に色分けされているが、一番左の領域は *hunchback* が、その右隣の領域は *Kruppel* が発現している事を示している。(b) 受精 3.5 時間後の胚。ペアルール遺伝子のうち *even-skipped* 及び *fushi-tarazu* が反復して発現している。 11

2.3 SIMEGG によるショウジョウバエの胚発生シミュレーションの結果（[29], 図 8 より） 13

2.4 von Neumann の自己複製セルオートマトンにおけるセルの状態の分類（[12] より転載）：29 種類の状態は U で表される休止状態、20 種類の活性状態、及び 8 種類の潜像状態からなる。普通伝達状態と特別状態における興奮状態は「パルスが存在する」状態であり、合流状態などに制御されながら矢印の方向へとパルスを伝達していく。また、休止状態 U はパルスの入力により潜像状態へと遷移し、入力パルスにより 9 通りの活性状態へと遷移する。 14

2.5 von Neumann の自己複製セルオートマトン（[16] より）：格子上で 15 万の長さに延びたテープにセルの規則が実行される。自己複製の際は、まず組み立て用の腕が伸び、新しい生物の胴体が形を成し、ついには２つの同じパターンが出来上がる。 15

2.6 Adam Loop[16]：1 個のループから伸びる腕がもうひとつのループを複製し、複製が繰り返される事でループのコロニーを形成する。 16

2.7 Adam Loop における状態の時間変化 17

2.8 一次元セルオートマトンによる 4 つの分類 18
2.9 クリスラングトンによる 4 つのクラスへの分類：クラス 3 はクラス 2 とクラス 4 の間の一部の領域であり、少しでも左にずれれば情報は固まり、結果として生命は維持できない。右にずれれば情報があまりに自由に動きすぎその構造が維持不能になり、無秩序すぎるために生命は維持できない。 . 19
2.10 λ とセルオートマトンの複雑さの関係 [16]：縦軸は複雑さを示し、横軸は λ を表す。λ がある値の時にシステムの複雑さは鋭いピークを示している。また、λ を 0 から 1 へと増やしていくと、クラス 1→ クラス 2→ クラス 4→ クラス 3 とシステムの振舞いは変化していく。 20
2.11 e-cell システムを用いて構築したバーチャル自活細胞 21
2.12 e-cell のシステム概略図 22
2.13 e-cell によるシミュレーションの様子 23
2.14 チューリングパターンによるまだら模様 24
2.15 チューリングパターンによるシミュレーション 25
2.16 Visual Object Net++によるシミュレーションの様子 26

3.1 反応半径：ある Particle に注目した時、その Particle を中心とした半径 r 内に存在する Particle のみと反応可能。 29
3.2 Particle と仮想平面のデータ構造:仮想平面の格子はPSFieldUnit クラスのインスタンスで構成されており、Particle は位置により対応する PSFieldUnit クラスのインスタンスに接続される。 32
3.3 Particle の一次元双方向接続リスト 33
3.4 PSFieldUnit の構造 34
3.5 2 つのフェイズに伴うポインタ付け変え 34
3.6 結果例（$[A]_0 = [B]_0 = 10$）：横軸は time step を、縦軸は Particle の個数の初期個数に対する比を表している。10 回の試行結果のうち、一番大きいところで 6 個も途中結果に差が生じており、ゆらぎも大きい。 38
3.7 結果例（$[A]_0 = [B]_0 = 100$）：$[A]_0 = [B]_0 = 10$ に比べてよく収束している。 39
3.8 結果例（$[A]_0 = [B]_0 = 1000$）：$[A]_0 = [B]_0 = 100$ に比べてよく収束している。 39

4.1 反応確率を変化させた場合の A の個数の時間変化：横軸は time step 数を、縦軸は Particle の個数を表している。 41

4.2 反応半径を変化させた時の A の個数の時間変化：横軸は time step 数を、縦軸は Particle の個数を表している。 42
4.3 あるパラメータでの A の時系列と、反応確率を 2 倍、反応確率を $\frac{1}{4}$ 倍にした場合の A の時系列：2 本のプロットはほぼ一致しており、$r^2 \cdot k = Const.(0 < r, k \leq 1)$ が示される。 43
4.4 挙動が一致するための r と k の条件 44
4.5 P を用いたシミュレーション結果：縦軸は $\frac{[A]}{[A]_0}$ を、横軸は時間を表している。実線は微分方程式系による解を、破線は Particle System によるシミュレーションの 10 回試行を平均化した値を用いている。 47
4.6 Particle に覆われる仮想平面 48
4.7 モンテカルロ法による補正を行ったシミュレーション結果 49
4.8 仮想平面の大きさによる平衡点の変化：縦軸は平衡点における $\frac{[A]}{[A]_0}$ を、横軸は正方形である仮想平面の一辺の長さ L を表している。L を小さくしていくと、平衡点は正反応の方向にずれていく事が分かる。 51

5.1 結果例（$[A]_0 = [B]_0 = 10$） 54
5.2 結果例（$[A]_0 = [B]_0 = 100$） 55
5.3 結果例（$[A]_0 = [B]_0 = 1000$） 56
5.4 Particle の個数を変化させた場合のゆらぎ：10 回試行の平均をプロットしている。縦軸はゆらぎを、横軸は初期の Particle の個数を表している。Particle の個数 N を増やすにつれ、誤差は $\frac{1}{\sqrt{N}}$ となっている。 57
5.5 Particle の個数を一定にした時のゆらぎ：10 回試行の平均をプロットしている。縦軸はゆらぎを、横軸は正方形の仮想平面の一辺の長さ L を表している。L が 5 から 20 ぐらいまでの範囲において、ゆらぎは L に対して指数的に増加している。L が 20 より大きい場合はゆらぎの収束が悪いものの、一定の範囲内に収まっている。 58
5.6 充填率 59

表 目 次

2.1 SIMEGG のモデル式における変数 12
2.2 ペトリネットの定義 24
2.3 主要シミュレータの性質のまとめ 26

3.1 Particle の持つパラメータ 29
3.2 Particle の変数 33
3.3 PSFieldUnit の変数 33
3.4 例題に用いたパラメータ 37

4.1 ブリュセレーターモデルの持つパラメータ 44
4.2 パラメータ決定の例 46
4.3 平衡点のずれについての実験に用いたパラメータ 50

5.1 例題に用いたパラメータ 53
5.2 充填率の調査に用いたパラメータ 55

第 1 章

序論

本章では、遺伝子工学や生命科学からの視点と人工生命や複雑系からの視点を概説し、この研究での立場とモチベーションについて示す。

1.1 本論文の背景と目的

我々が生命という現象について理解するためには、in vitro や in vivo（生物学）から見いだされる決定論的な知見と in silico（計算科学）からもたらされる知見のすべてを総合しなければならない。生物学は情報処理技術や計算科学によって飛躍的に知見を得る速度を増したし、計算科学は生命の成り立ちを真似る事によって遺伝的アルゴリズムやニューラルネットワークなどの新たな手法を得た。遺伝子工学や生命科学の分野と人工生命や複雑系の分野を統合する事で、新たな知見が得られるかもしれないと考えられるのは至極当然の事のように思える。

近年における遺伝子工学や生命科学は、「生命の部品」としてのデータを遺伝子や遺伝子発現情報や生化学反応といった形で表してきた。詳細なデータが大量に発表される事により、部品を組み合わせる事により生命という現象全体を見渡そうという動きである。しかしながら人工生命や複雑系の分野からの視点では、「全体は部分の総和だ」という考え方自体に疑問が生じる。複雑系とは、部分の総和が全体にならない系を示す。一般的に、十分に複雑な系であれば、部分の総和は全体とイコールにはならないのである。例えば細胞 1 個を例にとったとして、細胞内で起こりうるすべての生化学反応を羅列できたとしても、細胞全体の振舞いを再現する事は不可能かもしれない。

一方、人工生命のアプローチは構成論的かつ複雑系の概念を重視したものであり、いかに簡単な規則から複雑な現象が産み出されるかについて追求している。もしある系が十分

に複雑であるならば、自己複製やカオスに代表される生命の本質とも言える現象が発生し得るのではないか、というスタンスである。簡単な規則から複雑な現象を作り出す場合、計算機は欠かせない道具であり、人工生命に携わる科学者は計算科学と情報処理工学を駆使して研究を行っている。

しかし人工生命や複雑系は純粋に生命現象の原理に焦点をあてたものであるから、現存する生物との照らし合わせ自体が困難である。ここで浮かんでくる問題は、いかに遺伝子工学や生命工学の分野と人工生命や複雑系の分野を統合していくかである。それぞれが生命という現象について研究しているにも関わらず、その間には歴然とした差が存在しているのである。そこでこの研究では、その双方の側面を取り込みバイオインフォマティクスを発展させる事を大きなモチベーションとする。

1.1.1 バイオインフォマティクスの現状

1980 年代後半に始まったゲノム計画に代表される遺伝子情報の網羅的な解析によって、現在のところ 70 種以上の生物の全ゲノム情報が解析、決定されている [1][21]。近年著しく発展を遂げているバイオインフォマティクス（情報生物化学）は、ゲノム情報の解析により蓄積された膨大かつ多種多様な生物情報を整理統合し、そこから新しい知見を見いだす事により、新しい産業や医療の創出を目的として体系化され [2]、現在最も注目される学問のひとつとなっている。

まず単独に生きる生物としてはじめてのバクテリアゲノムが決定され（1995 年）、翌年には初の真核生物として酵母のゲノムシーケンシングが終了した。そこから遺伝子領域を予測し、遺伝子産物の機能を指定する事は情報処理の問題となり、主な手法としてデータベースからのホモロジー検索やモチーフ検索が挙げられる。

シーケンシングによって得られた情報から様々な検索手法を用いて得られた知見は機能推定にとって大きな手がかりとなるが、これらの情報は一次元配列についての手がかりに過ぎない。実際には DNA もたんぱく質も立体的な形をした分子であり、立体構造を決める事が重要となる。立体構造の推定は、分子を構成する原子の原子間の力を基に動力学的に構造をシミュレートする分子動力学法（MD 法：Molecular Dynamics）や、分子を直接観測する X 線結晶解析や NMR などの手法が用いられている。ついで遺伝子の発現情報を DNA、mRNA、タンパク質レベルで解析する。

バイオインフォマティクスの重要性が叫ばれた当初は上記のような遺伝子の配列情報解析やタンパク質の構造情報解析及び発現情報の解析が主であったが、近年機能情報解析の重要性やが認識されつつあり、バイオインフォマティクスの取り扱う領域は拡大しつつあ

る [4]。

構造解析や発現情報解析などの「部品」の解析から「システム」の解析に移行する際に、発現情報解析と共に重要になるのは分子間相互作用である。現在は、分子間相互作用についての解析を進めると共に、生体内の代謝や転写制御などの経路を推定するネットワーク解析へとバイオインフォマティクスはその主たる目的を拡大している。将来的には分子間相互作用とネットワーク解析（このネットワークは遺伝子ネットワークと呼ばれる）がバイオインフォマティクスの主たる目的になると考えられており、配列解析とタンパク質構造解析と発現情報解析が主流である今日と対比する形で「ポストシークエンス解析の時代（post-sequence era）」と呼ばれる。図 1.1 にバイオインフォマティクスの流れについてまとめる [3]。

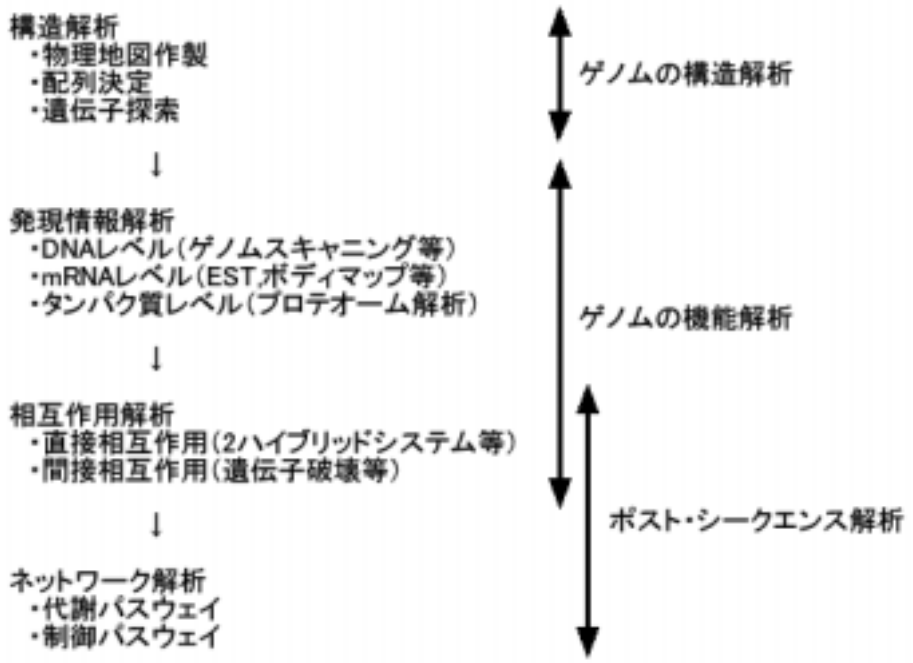


図 1.1: バイオインフォマティクスの流れ [3]

1.1.2 人工生命と複雑系

近年のバイオインフォマティクスが見せる目覚ましい進歩は、実在する生物の DNA の配列解析から始まる一連の小さな知識の積み上げによるものだと言える。その一方で、対象の本質的側面を数式によって表現し解析するという数理モデルによる研究が、生命現象

を解明するうえで着実に重要性を増してきた。神経細胞における電気的興奮の伝播や筋肉収縮の分子機構の技術などのミクロな対象から、発生過程でひとつの細胞が分化を起こし形態や機能が個々に作られるメカニズムの研究、神経細胞のネットワークから脳の働きの説明に至るまで、様々な分野で数理モデルが活躍している。

生命が内包する根本的な複雑さを数式や簡単な規則として抽出する研究は主に複雑系や人工生命の分野で行われており、生命現象そのものや生物社会におけるダイナミクスなどの現象をボトムアップ的にシミュレーションする事などが研究されている。ボトムアップ的な再現という方針は、人工生命の基本方針である。

人工生命の先駆者としてノイマン（J. von Neumann）を挙げる。ノイマンは 1930 年代にハンガリーから米国に移住した研究者であり、数学、物理学、計算機科学、経済学など幅広く活躍した [14]。ノイマンは有限オートマトン理論を展開し、計算機上で生命の根本原理の一つである自己複製を行う事を目的としたセルオートマトンを作り上げた。つまり初期のセルオートマトンの興味の中心は「抽象的な空間に配置され有限個の状態を持つセル群で自己と同じセル群を構成する」事にあった。この研究成果がシミュレーション手法の一つであるセルオートマトン理論の祖となり、一方で人工生命の走りともなった。人工生命の研究者はボトムアップによる創発や自己組織化を目的とし、以降、L-system や λ-gas 、遺伝的アルゴリズム、遺伝的プログラミングなどのシミュレーション手法が発表された。またそれらの手法は生命の形態形成や進化、自己組織化などに適用され、Chris Langton の”Adam loop”[7] や Karl Sims の”Artificial Evolution”[8]、Thomas S. Ray の”Tierra”[9] などの成果を残した。

これらの成果は、数理モデルが生命現象の根本を照らし出す事に部分的ながら成功している事を示唆しており、ひいては生命現象の再現を試みた時、人工生命や複雑系の分野から得られた知見を遺伝子工学や生命科学へ適用できる可能性も示唆している。

複雑系的生命科学　生命という現象についての学術的なアプローチは、大きく 2 つに分けられる。決定論的なアプローチと複雑系において用いられるような構成論的なアプローチである。

決定論的なアプローチとは、たとえば細胞内のシグナル伝達物質の濃度と遺伝子の発現についての関係を「If（シグナル分子の濃度があるしきい値より大）, then（ある遺伝子がオン、もしくはオフ）」という論理演算として解釈するものである。このように解釈された複数の論理式の組み合わせにより生命というシステムを記述しようとする [5]。しかしながら、生命という現象をミクロな if-then ルールに分割するこの方法とは別に、複雑系としての生命現象を研究しようとする方法が近年注目されている。

生命の普遍的性質を導く最低限度の性質を表現し、現存生物にこだわらず生命という現象を構築しようとするのが複雑系的生命科学である。ここではある世界を論理的に構築し、その中で何が普遍的であり必然であるかを明らかにしようとするものである。この際に世界を構築するにあたっては、現実に合わせようとして細部をごちゃごちゃと付け加えるのではなく、生命の理解につながる程度の最低限の仮定しか導入しない。

1.1.3 生化学シミュレーションの現状

近年バイオインフォマティクス分野で主に試みられているシミュレーションは、部分的な生化学反応である。たとえば細胞 1 個について、含まれる反応式を詰め込んだシミュレータを構築する事により、実験的に求められない問題の解をシミュレーション側から予測する事が可能になっている。

シミュレーションから実際の生物の特性を推し量る方法論は有用なものであるが、生物学的な知見が広がっていくにつれ、シミュレーションの手法自体もそれに従い改良を加える必要性が出てきた。

転写制御や細胞内シグナル伝達といった現象では、関与する分子数が少ない。したがって、既存の生化学のように物質を単純な濃度で取り扱う事ができない。また、発生の過程では化学物質の位置情報が非常に重要な役割を果たす事が知られている。加えて、たとえば発生の場合シグナル分子数はせいぜい 1000 個程度であり、大数の法則に従えば約 3 %のゆらぎが生じる。このゆらぎに対しても安定した挙動をシステムは示さなければならないし、一方でこのゆらぎによりそれぞれの細胞に変化が発生する場合もある。

こういった要請に答えられるシミュレータの構築は、ポストシークエンス時代に突入しつつある現代にとって非常に重要な課題である。また、こういったシミュレータの開発の際には、複雑系や人工生命における考え方も活用すべきであろう。

現在のところ、大半の生化学シミュレーションは常微分方程式系シミュレータを用いて行われている。化学反応系を表現する際に一般的な速度定数を伴った反応式をそのままシミュレーションに適用する事が可能である点がその原因であると思われるが、複雑系や人工生命の考え方に鑑みると不足な点が多い事も否めない。先述した通りに以下の三点についての再現性に欠いているからである。

- 濃度と解釈できない程に少ない反応物が関与する反応
- 位置情報の取り扱い
- ゆらぎ

一方、シミュレーションという方法論自体がバイオインフォマティクス研究者全体に広がりつつあり、以前は情報分野の研究者が行っていたシミュレーションを生物科学者も必要としている。これからのバイオインフォマティクス用シミュレータには以下の 2 点が要求される。

- 生化学反応特有の表現やシミュレートがしやすい
- 生物科学者にも扱いやすい

1.1.4 モチベーション

バイオインフォマティクスがより精緻なシミュレーションへと向かっていく今日、より精緻はシミュレータの構築の必要性が感じられる。そしてそのシミュレータは、遺伝子工学や生命科学によって得られた「部品」のデータと人工生命や複雑系における構成論的概念を両方サポートでき得るものでなければならない。特に、生化学現象の再現に特化し、トップダウン的に系の振る舞いを記述するのではなくボトムアップ的に系の振る舞いを再現可能なシミュレーションシステムの構築は、生命科学の視点からも複雑系や人工生命の視点からも多いに期待されるところであろう。

複雑系や人工生命の分野で用いられている手法や既存のシミュレーション手法や生化学現象の特徴を踏まえ、ポストシークエンス時代のシミュレーションを視野に入れた新シミュレーションシステムが持つべき特徴について考察した。結果以下のようなシミュレータの必要性を充足する事が求められているとの結論を下した。シミュレータは主に遺伝子ネットワークを含む反応系のシミュレーションに対して適用される事を想定し、

以下の特徴を持たせる事とする。

ボトムアップなシステム　生化学反応においてすべての反応は局所的に起こり、局所の総合として全体の反応が制御される事から、厳密なトップダウン制御による反応は存在しない。すべての反応は局所的に起こり、局所の総合として全体の反応が制御される。よって、本システムでは局所的な反応のみを扱い、系全体に対する制御を行わない。例えば「ある物質の濃度は 1 秒で半分になる」微分方程式を本システムでシミュレートする際は、局所的に「この物質はそれぞれ一秒の間に半分の確率で消える」という形でシミュレートする。ボトムアップという特徴は人工生命と同様であるが、本研究は定量的シミュレーションを行なう事も目的としている。

化学物質を濃度ではなく個々の粒子として表現するシステム　生化学反応において度々見られるような、濃度での取り扱いをした場合に齟齬が生じてしまうような系について精緻にシミュレートするために、化学物質を濃度ではなく個々の粒子で表現する。反応物質を離散かした事により、結果的にゆらぎが生じる事になる。個々の粒子は必ずしも「1 個」を表現する訳ではないが、ゆらぎなど粒子性が効果を表す現象には有効である。

位置情報を取り扱えるシステム　偏微分方程式では濃度分布という形で位置情報を取り扱うが、本システムでは濃度を取り扱わない。そこで、本システムでは、粒子によって表現した化学物質を仮想的な平面上に配置する事で粒子毎の位置情報を正確に管理する。また、局所的な化学反応の条件に位置情報を加える。粒子はタイムステップ毎に熱運動を繰り返す事でランダムに移動する。

定量的シミュレーション　現在主流である微分方程式系シミュレータは化学反応式を定量的に再現可能である。前述したようにゆらぎや位置情報を考慮しつつも、微分方程式系の暗黙の仮定である「反応系に関与する物質が十分に多く」「十分に撹拌されている」場合に微分方程式系と同じ振舞いを示すようなシステムを構成する。

そこで本論文は、これらの要求を踏まえた新シミュレーションシステム「Particle System」を提案する。

Particle System は、反応に関与する物質をすべて Particle（粒子）として計算機内に表現し、それらの Particle が計算機内の仮想的空間をランダムに運動する中で局所的に反応を制御する。反応を制御する要素は、ある Particle の周辺に存在する Particle の数及び事前に与えられた反応確率である。

本論文では、現在のところ最もシミュレーションに用いられている微分方程式系によるシミュレーションと Particle System とを比較し、系が十分撹拌され、かつ反応に関与する物質が十分にたくさん存在する場合に微分方程式系と Particle System が本質的に同様のシミュレーション結果を導く事を示す。また、反応に関与する物質の数が少ない状態での系のゆらぎを計測し、実際の生化学反応におけるゆらぎのシミュレーションを試みる。

1.2　本論文の構成

本章は、本論文の目的と構成について述べる。

第二章は、本論文に関する従来研究について述べる。第三章は Particle System のアルゴリズムや実装について述べる。第四章は Particle System によるシミュレーション結果

を分析する事により、微分方程式系と本質的に同じ振る舞いをしている事を示す。第五章はParticle System によるシミュレーションが示すゆらぎについて述べる。第六章は結論及び今後の課題を述べる。

第 2章

本論文に関する従来研究

2.1 はじめに

本章では、本研究に関する従来研究についてまとめる。

まず既存のシミュレーション手法について述べ、次に生化学反応についてのシミュレーションの特殊な性質やシミュレータに求められる能力についてまとめる。

2.2 生化学反応による空間パターン形成

本節では、生化学反応に見られる複雑な現象のひとつである空間パターン形成について、ショウジョウバエの胚発生を例にとり述べる。

2.2.1 ショウジョウバエの胚発生

ショウジョウバエの胚発生における体節形成のメカニズムについては、ショウジョウバエ自体の変異体作りとその解析が比較的容易である事から広く研究され、現在では可視的な発生以上をもたらす遺伝子についての同定がほぼ完了している [23]。胚の発生過程において、母性効果遺伝子である *becoid(bcd), nanos(nos), torso(tor)* とギャップ遺伝子の *Kruppel(Kr), knips(kni), hunchback(hb), tailless(tll), huckbein(hkb), giant(gt)* などからつくられたDNA結合タンパク質が胚の中に濃度勾配を作る事が知られており、これらの濃度勾配がペアルール遺伝子と呼ばれる一群の遺伝子の発現を制御し、七つの帯状の領域を作る事が知られている [19]。分節化の初期段階において胚の前極に *bicoid* 遺伝子が、後極には *nanos* 遺伝子の mRNA が局在し、これらの濃度勾配が分節化過程の初期段階に胚の

前後軸を形成するとされる [20]。これを位置情報物質として利用し、一連のギャップ遺伝子は互いに活性 · 抑制を行い、図 2.1 にみられるような局在パターンを形成する。

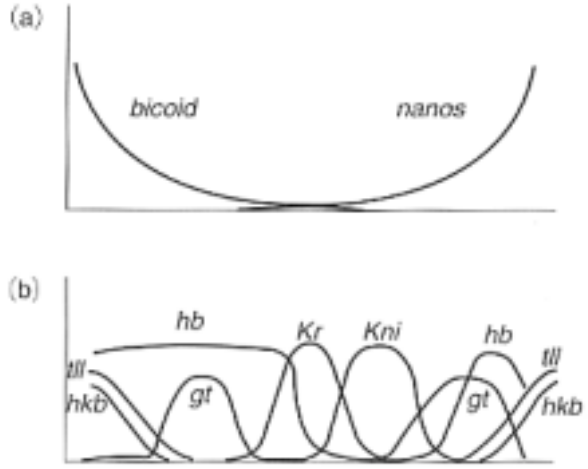


図 2.1: ショウジョウバエの胚発生における体節パターン形成にかかわる遺伝子産物の濃度分布（[20], 図 2-2 より）: (a) *bicoid* と *nanos* 母系 mRNA が形成する勾配。(b) ギャップ遺伝子の発現の様子。

このようにして形成された遺伝子群の濃度勾配の組み合わせにより特定のペアルール遺伝子の転写が制御され、さらにペアルール遺伝子間の負の非線形フィードバックによって、明瞭な境界をもったストライプが形成されうることが数理モデルによって示されている [5]。ペアルール遺伝子の発現例を図 2.2 に示す。

ショウジョウバエの胚発生シミュレーション　ショウジョウバエの胚発生についてのシミュレーション例として、SIMEGG[29] について示す。SIMEGG はショウジョウバエの胚発生についてのシミュレータであり、以下の偏微分方程式を用いてモデル化を行っている。以下の式における各変数の説明を表 2.1 に示す。

$$\frac{\partial U_i}{\partial t} = D_i \frac{\partial^2 U_i}{\partial x^2} + g \cdot U_i + f(\mathrm{U})$$

上式から、SIMEGG ではタンパク質の生成と拡散及び消滅をモデル化している。SIMEGG は未知のパラメータを遺伝的アルゴリズムを用いてパラメータ推定を行い、基本的に実際の生物学的実験で得られているデータよく一致したシミュレーション結果を残している。図 2.3 に SIIMEGG によるシミュレーションの結果を示す。

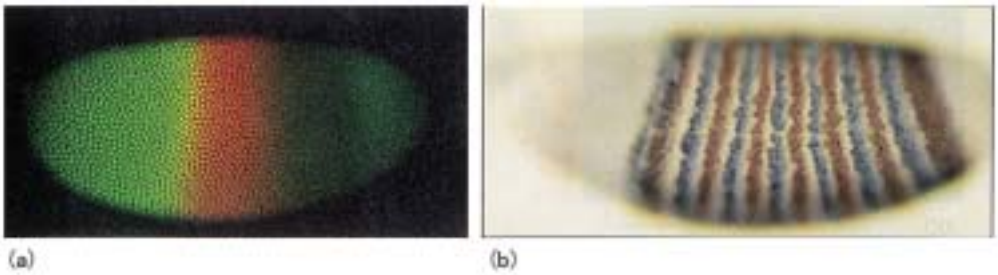


図 2.2: ショウジョウバエ初期胚にみられるペアルール遺伝子の発現パターン（[23], 図 19-41 より）：(a) 受精 2.5 時間後の胚。全体が 4 分割に色分けされているが、一番左の領域は *hunchback* が、その右隣の領域は *Kruppel* が発現している事を示している。(b) 受精 3.5 時間後の胚。ペアルール遺伝子のうち *even-skipped* 及び *fushi-tarazu* が反復して発現している。

2.3 人工生命と複雑系

本節では、人工生命や複雑系の視点から生命現象の再現を試みた例として von Newmann の自己複製オートマトンと Chris Rangton の Adam Loop について示す。

2.3.1 von Newmann の自己複製オートマトン

ハンガリー系アメリカ人の科学者であるフォンノイマン（John von Newmann）は、数学、物理学、計算機科学、さらに経済学の分野にまで及ぶ研究成果を残した。晩年にあたり生命現象の再現に興味を持ち、試行錯誤を繰り返したのちにセルオートマトンの基本概念を構築し、自己複製オートマトンについての草案を著した [14],[16]。

ノイマンの自己複製オートマトンは 2 次元の正方格子状のセル空間を基盤としており、各々のセルは 29 種類のいずれかをとる。そして自分自身及び辺で接する 4 個のセル（この近傍をノイマン近傍と呼ぶ）で相互作用を行う [11]。図 2.4 にノイマンの自己複製オートマトンにおけるセルの状態について示す。論理回路を模す事で情報の伝達を行なっている。

自己複製オートマトンは主に組み立てユニットとテープユニットとテープから成る。テープに書き込まれたデータの読み込みなどをテープユニットが行ない、そのデータをもとに自己複製を行なうのが組み立て制御ユニットである。図 2.5 に自己複製オートマトンの構成図を示す。

表 2.1: SIMEGG のモデル式における変数

変数	説明
U_i	タンパク質 i の密度
x	軸上での位置
t	時間
D_i	タンパク質の拡散係数
g	タンパク質の消滅係数
f	タンパク質生成係数
U	タンパク質の密度ベクトル

2.3.2 Chris Rangton の Adam Loop

ノイマンの自己複製オートマトンに影響を受けたラングトン（C. Rangton）は、自己を複製し続けるセルオートマトンを構成し、「Adam Loop」と名付けた。そのセルオートマトンはノイマン近傍をとり、8 つの状態を持つ。図 2.6 にコロニーが生成されていく様子を示し、図 2.7 に Adam Loop におけるセルの状態遷移の時間変化について示す。

2.4 シミュレーション手法

本節では、現在バイオインフォマティクスで用いられているシミュレーション手法について示す。また各々の手法に関して、その手法が提案されるに至った経緯を示し、生命や生化学反応のモデル化に対する試みがどのような変遷を経てきたかについて述べる。

2.4.1 セルオートマトン

セルオートマトンはノイマンとウラム（S.Ulam）によって基本的アイデアが形作られた。セルオートマトンの基本設計は、セル自身の状態とその周辺のセルの状態のみに依存して、セルが自律的に状態遷移を行うルール群を作る事である [17]。

- 空間的に離散である：セルオートマトンは空間的なセルの離散的な格子から成る。
- 時間的に離散である：各セルの値はある離散時間ステップのシーケンスによって構成される。

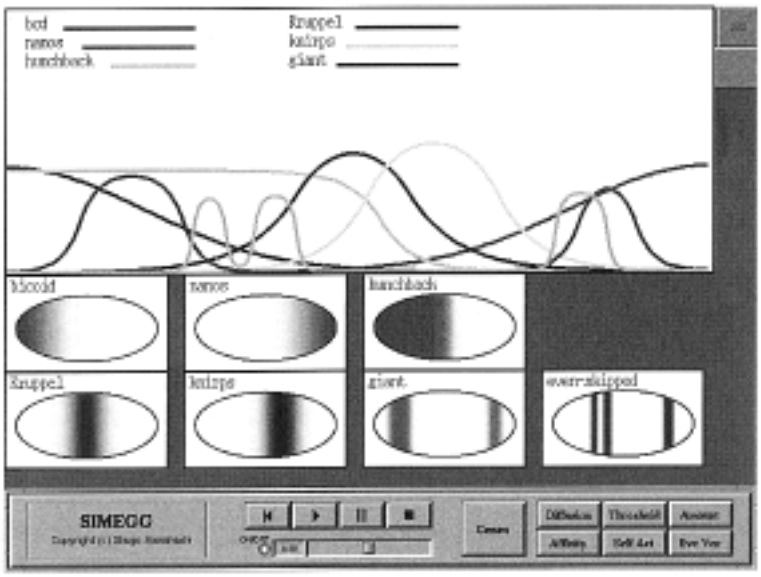


図 2.3: SIMEGG によるショウジョウバエの胚発生シミュレーションの結果（[29], 図 8 より）

- 離散状態をとる：各セルは有限の状態数を持つ。
- セルの均一性：すべてのセルは同じ状態遷移則に従い、規則正しく並べられている。
- 同期的な状態更新：すべてのセルの状態は、自身とその周辺のセルの状態に依存して、同期的に更新される。
- 決定論的な規則：各セルの状態は決定論的な規則に基づいて更新される。
- 空間的に局所的な規則の適用：書くセルに適用される規則はその近傍セルの局所的な状態にのみ依存する。
- 時間的に局所的な規則の適用：値あるいは状態更新のための規則は、ある固定数（通常は 1 ステップ前）先行する時間ステップ間の状態にのみ依存する。

同じ大きさの正方形（以下セルという）で区切られた碁盤の目のような面を想定した時、二次元セルオートマトンは以下のような定義をされる。

- 各セルは k 個の状態をとることができる。

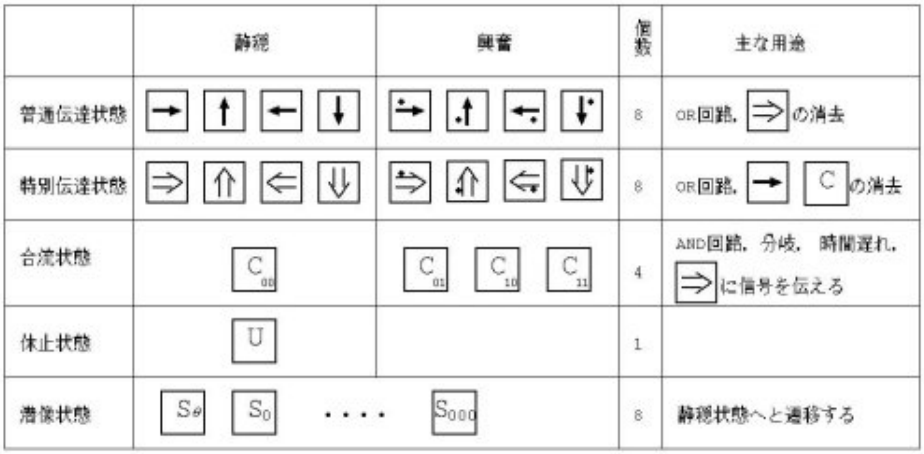


図 2.4: von Neumann の自己複製セルオートマトンにおけるセルの状態の分類（[12] より転載）：29 種類の状態は U で表される休止状態、20 種類の活性状態、及び 8 種類の潜像状態からなる。普通伝達状態と特別状態における興奮状態は「パルスが存在する」状態であり、合流状態などに制御されながら矢印の方向へとパルスを伝達していく。また、休止状態 U はパルスの入力により潜像状態へと遷移し、入力パルスにより 9 通りの活性状態へと遷移する。

- 次の時間のセルの状態は、現在の状態と隣り合うセルの状態だけの局所的な規則により決まる。
- できあがるパターンは初期のセルの状態と適用する規則によって決まる。

ライフゲーム

一般に知られている 2 次元セルオートマトンの例として、ライフゲームを挙げる。ライフゲームでは、多数の正方形のセルで区切ったコンピュータの画面上に、生きている状態と定義した生のセルを少数置くと増殖し始め、単純な初期の形と簡単な規則から予想もできない複雑で多彩なパターンが作り出される [14]。セルは生か死の 2 値の状態を持ち、1 個のセルとそのセルを取り巻く 8 個のセルの合計 9 個のセルの状態により次タイムステップのセルの状態が決まる。周囲のセルのうち生きているセルの個数を数え上げ、次の単純な規則を適用する。

- 隣りのセルが 2,3 個生きている場合は生き続ける。
- 隣りに生きているセルが 4 個以上または 1 個以下になった場合は死ぬ。

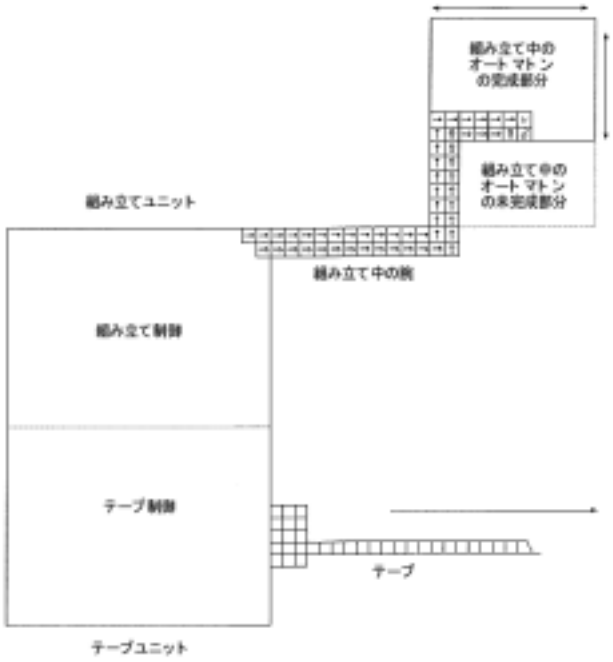


図 2.5: von Neumann の自己複製セルオートマトン（[16] より）：格子上で 15 万の長さに延びたテープにセルの規則が実行される。自己複製の際は、まず組み立て用の腕が伸び、新しい生物の胴体が形を成し、ついには２つの同じパターンが出来上がる。

- 死んだセルは隣に生きているセルがちょうど 3 個になった場合に生きかえる。

情報力学系としてのセルオートマトン

セルオートマトンの挙動について詳しく観察したウルフラム（S. Wolfram）は、挙動が 4 つのグループに分類できる事を示した。すなわち以下の 4 パターンである。図 2.8 に一次元セルオートマトンにおける分類について示す。

1. クラス１：すべてのセルの状態は均一となり、初期にもっていたパターンは消える。
2. クラス２：変化しないパターンまたは周期的に繰り返すパターンに落ち着く。
3. クラス３：非周期的でカオス的な振舞いをするパターンが現れる。
4. クラス４：パターンが消えたり、非周期的あるいは周期的なパターンとなったり、予測しがたい非常に複雑なパターンが現れる。

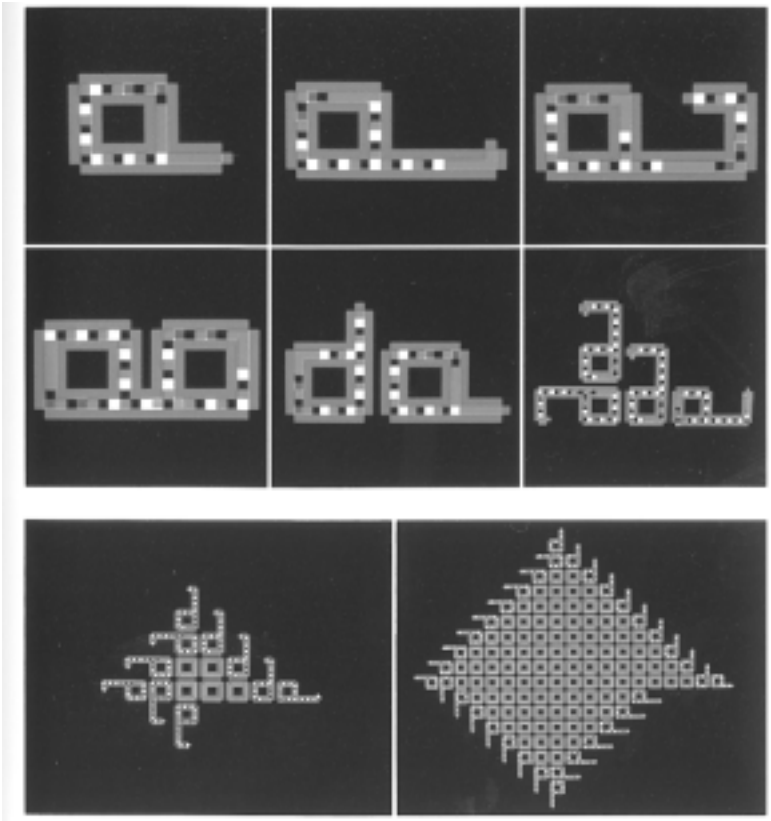


図 2.6: Adam Loop[16]：1 個のループから伸びる腕がもうひとつのループを複製し、複製が繰り返される事でループのコロニーを形成する。

クリスラングトン（Chris Rangton）はウルフラムの分類を拡張し、セルオートマトンのような複雑力学系における情報の動きを以下のように分類した。図 2.9 と対応する形で示す。

1. クラス１：固定、均一
2. クラス２：周期的
3. クラス３：カオス的
4. クラス４：複雑（その他）

図中の左側では情報が凍り、何もかもが生きられない。右に進むと結晶が生じるようなより柔軟性のある領域があるが、情報の移動は限られ、ここでも生命は維持できない。右

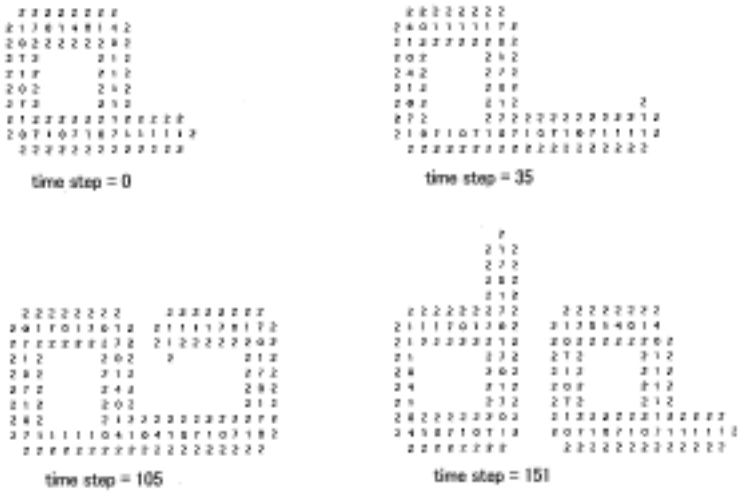


図 2.7: Adam Loop における状態の時間変化

端まで行ってしまうと、情報があまりに自由に動きすぎその構造が機能不能になり、あまりに無秩序なので生命は維持できない。中央部の「スイートスポット」と呼ばれる部分だけで、情報はその発信構造を維持できるほど安定し、発信できる程度にゆるやかである。ラングトンは「生命はそこにある」と主張した [16]。

またラングトンは、セルが K 個の状態を持ち、かつ遷移規則が N 個の近傍を参照している場合、以下の式から λ を算出し、λ からセルオートマトンの振舞いについて定量的に論じた [16],[13]。

$$\lambda = \frac{0\text{（死）以外の状態へと遷移する場合の数}}{K^N}$$

上式において、例えば 2 状態 3 近傍の遷移規則において 5 通りが 0（死）以外の状態を取りうる場合の $\lambda = \frac{5}{2^3} = \frac{5}{8}$ となる。ラングトンは λ に対応したセルオートマトンを生成し分析した結果、λ がある値を取った時にセルオートマトンの振舞いの複雑さが最大になる事を示した。図 2.10 に λ と複雑さの関係について示す。

2.4.2 常微分方程式系

化学反応系の振る舞いにおいて、反応物の濃度についてをトップダウン的にシミュレートできるモデルが微分方程式系によるモデルである。反応速度論によってもたらされた微分方程式を単純に数値計算によって解く手法であり、その汎用性は高い。

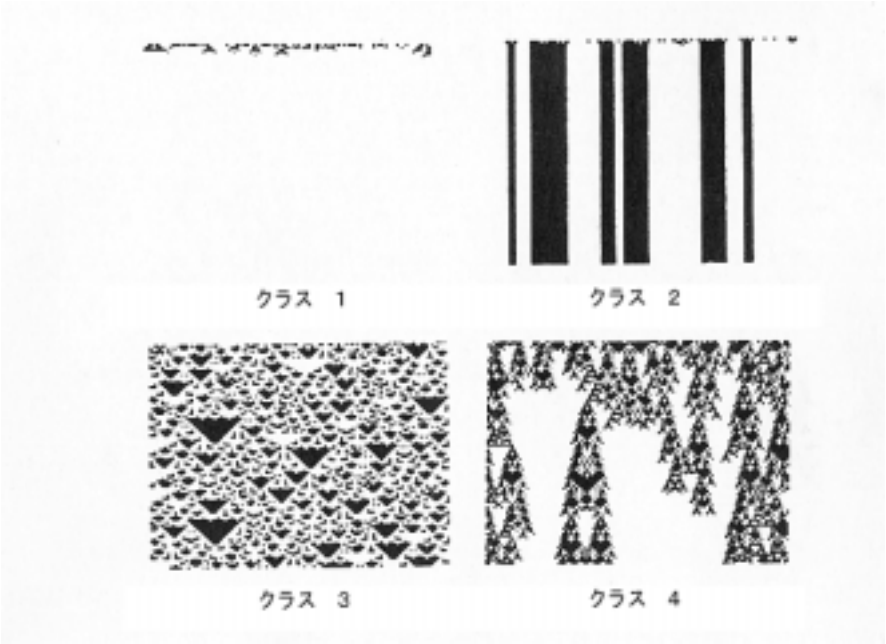


図 2.8: 一次元セルオートマトンによる 4 つの分類

現在のところ開発されている微分方程式系シミュレーションエンジンの中で、バイオインフォマティクス分野で頻繁に用いられているものとしては、gepasi[32]、e-cell[33] などが挙げられる。

ここで、例として酵素反応についてモデル化した Michaelis-Menten 式について示す。酵素 E が基質 S に結合して複合体 ES を作り、複合対の中で S が反応中間体を経て生成物 P となったあと、酵素 E と生成物 P が分離する、という過程を簡略化し、反応式を使って書くと次のようになる。

$$E + S \leftrightarrow ES \rightarrow E + P$$

定常状態時のこの反応を考慮した場合、P の時間変化 v は以下の式で書き表され、この式は Michaelis-Menten 式と呼ばれる。ただし V_{max} は基質 S の初期濃度を大きくした場合に得られる P の変化の最大速度であり、K_m はミカエリス定数と呼ばれる。

$$v = \frac{d[P]}{dt} = \frac{k_3[E]_0[S]_0}{K_m + [S]_0} = \frac{V_{max}[S]_0}{K_m + [S]_0}$$

e-cell

常微分方程式系をソルバーとした生化学シミュレータとして、e-cell について述べる。

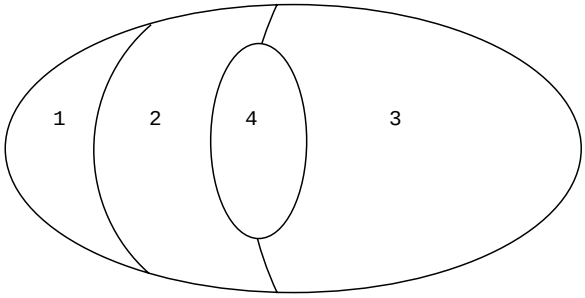


図 2.9: クリスラングトンによる 4 つのクラスへの分類：クラス 3 はクラス 2 とクラス 4 の間の一部の領域であり、少しでも左にずれれば情報は固まり、結果として生命は維持できない。右にずれれば情報があまりに自由に動きすぎその構造が維持不能になり、無秩序すぎるために生命は維持できない。

e-cell プロジェクトは細胞内の代謝をまるごとシミュレーションする事を究極の目的として 1996 年に発足した。1997 年には 127 個の遺伝子からなる架空の「バーチャル自活細胞」を、1999 年にはヒト赤血球細胞の全代謝をモデル化し「バーチャル赤血球細胞」を完成させた [22],[24]。

バーチャル自活細胞は、細胞全体をシミュレートすることの可能性を示すために、代謝経路、遺伝子発現系、膜輸送のモデルを融合し定義した仮想細胞モデルである（図 2.11 参照)。最小のゲノムセットを持っていると言われているマイコプラズマ菌（M. genitulium）を細胞モデルの参考にしたもので、アメリカの TIGR 研究所によって 95 年に同定された 500 弱の遺伝子の中から細胞としての自己維持のために必要最低限な 127 個の遺伝子のサブセットを定義し用いている。このバーチャル自活細胞は膜外からグルコースを取り込んでそれを解糖系によって分解しエネルギー（ATP）を生産する。また、細胞膜生成のためのリン脂質合成系を持ち、脂肪酸とグリセロールを取り込んでホスホチジルグリセロールを合成しこれが細胞膜となる。遺伝子発現のための転写機構（RNA ポリメラーゼなど）および翻訳機構（リボソームなど）を持ち、遺伝子からタンパク質を合成する。タンパク質は時間とともに自然分解するようにモデル化してあるので、タンパク質を作り続けないと細胞は死んでしまう。タンパク質を合成するにはエネルギーが必要で、そのためにはグルコースが必要となる。127 個の遺伝子のうち 20 個がトランスファー RNA、2 個がリボゾーム RNA をコードする RNA 遺伝子である。残りの 105 個のタンパク遺伝子のうち、7 つの遺伝子はマイコプラズマ菌の遺伝子リストに存在しなかったので大腸菌などの他の生物から拝借したものだとされている。

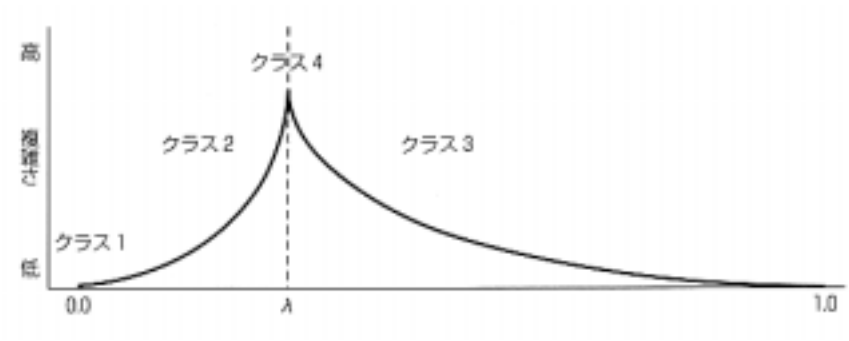


図 2.10: λ とセルオートマトンの複雑さの関係 [16]：縦軸は複雑さを示し、横軸は λ を表す。λ がある値の時にシステムの複雑さは鋭いピークを示している。また、λ を 0 から 1 へと増やしていくと、クラス 1→ クラス 2→ クラス 4→ クラス 3 とシステムの振舞いは変化していく。

e-cell システムの構成図を図 2.12 に、また e-cell によるシミュレーションの様子を図 2.13 に示す。

2.4.3 偏微分方程式系

偏微分方程式系は、常微分方程式系と同様に濃度を扱い、かつ位置情報を濃度勾配という形で扱える手法である。したがって、特に生化学反応などで空間パターンが出る場合などのシミュレーションに応用可能である。

チューリングモデル

偏微分方程式系によるシミュレーションの例として、チューリングモデルについて述べる。

チューリング（A.Turing）は生物学のみならずコンピュータ科学、人工知能および数学の部やにおいて大きく貢献した研究者で、チューリングモデルの他にもコンピュータ科学の基礎理論であるチューリングマシンを開発するなどした [18]。

チューリングは、生体内での化学反応と反応生成物質の拡散における拡散速度の差によって空間パターンが形成される事を示し、形態形成（morphogenesis）と言い表した [25]。

チューリングのモデルは、生体パターンの形成が、活性化（X）と阻害（Y）に働く二種類の化学物質の細胞間の反応と拡散に基づき起こるというものである。X は X 自身および Y の合成を濃度依存性に促進するものとする。つまり X の濃度が高いところほど、X

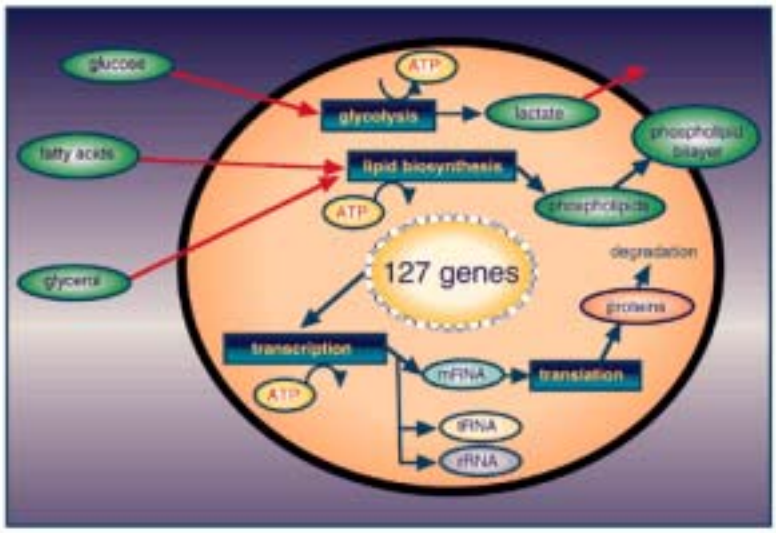


図 2.11: e-cell システムを用いて構築したバーチャル自活細胞

や Y の合成は盛んに発生する。一方 Y は X の合成の阻害剤であり、Y の濃度が高いところでは X の合成は抑制される。また、X と Y はともに拡散によって周囲に広がってゆくが、Y の拡散速度は X の拡散速度よりも大きいものとする。ここで、局所的に X の濃度が周囲より高い場所についてどのようなダイナミクスを形成するかについて以下に示す。

1. X の合成量は濃度が高いところほど大きいので、その場所での X の濃度が首位と比べてさらに高くなる。

2. 同じ場所で Y の濃度も増大するが、Y の拡散速度は X よりずっと大きいので、Y の大部分は周囲に拡散する。

3. 結果的に、周辺部では X はほとんど変化しない。一方 Y は拡散によって濃度が上昇するために、X の合成量は逆に減少する。

4. つまり、局所的に微少な濃度差が常に増幅されてしまう。

こうして拡散反応波によって作られたパターンは自律的に作られたものであり、かつほとんどすべての種類の動物の模様を作り出せる [27],[28]。

以下に、理想化された円環器官（continuous ring of tissue）つまり周期的な境界条件における反応拡散から、非対称な形態が生成されるというチューリングパターンの概要を紹介する。

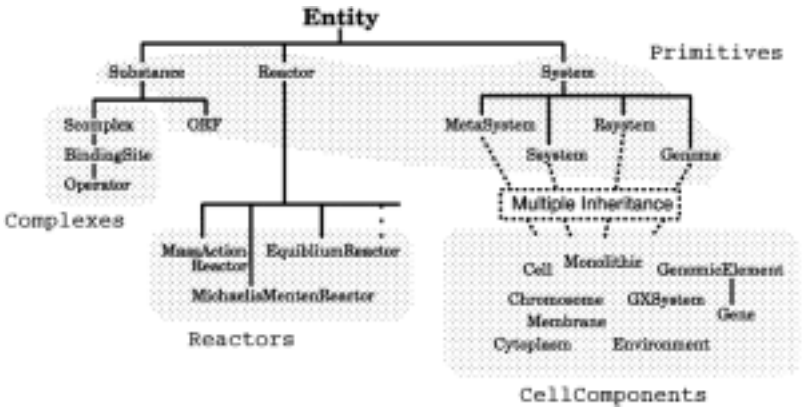


図 2.12: e-cell のシステム概略図

二つの拡散方程式系を図のようなセル（細胞）に分割された円環上で考える。ここで、半径 r は等しい（$r = \rho$）位置とする。この時の各セルの 2 種類の化学物質の濃度 $X(t,\theta), Y(t,\theta)$ は以下の式で表される。

$$\frac{\partial X}{\partial t} = a(X-h) + b(Y-k) + \frac{\mu'}{\rho^2}\frac{\partial^2 X}{\partial \theta^2},$$

$$\frac{\partial Y}{\partial t} = a(X-h) + b(Y-k) + \frac{\nu'}{\rho^2}\frac{\partial^2 Y}{\partial \theta^2}$$

ここで、μ', ν' はそれぞれ規格化された拡散係数で、以下とする。

$$\mu = \mu'\left(\frac{N}{2\pi\rho}\right)^2, \nu = \nu'\left(\frac{N}{2\pi\rho}\right)^2$$

この時 ν は μ の 10 倍以上に設定されるのが一般的であり、$\nu \gg \mu$ である。この拡散速度の差が空間パターンを生成する本質的な要因である。また、a, b, c, d は反応率である。これがチューリングモデルあるいは反応拡散モデルと呼ばれるものである。

このとき、一般解は以下のように示される。

$$X = h + \sum_{s=-\infty}^{\infty} (A_s \cdot \exp(p_s t) + B_s \cdot \exp(p'_s t)) \cdot \exp(i\theta),$$

$$Y = k + \sum_{s=-\infty}^{\infty} (C_s \cdot \exp(p_s t) + D_s \cdot \exp(p'_s t)) \cdot \exp(i\theta)$$

である。ここで i は虚数単位、p_s, p'_s は二次方程式

$$\left(p - a + \frac{\mu' s^2}{\rho^2}\right) \cdot \left(p - d + \frac{\nu' s^2}{\rho^2}\right) = bc.$$

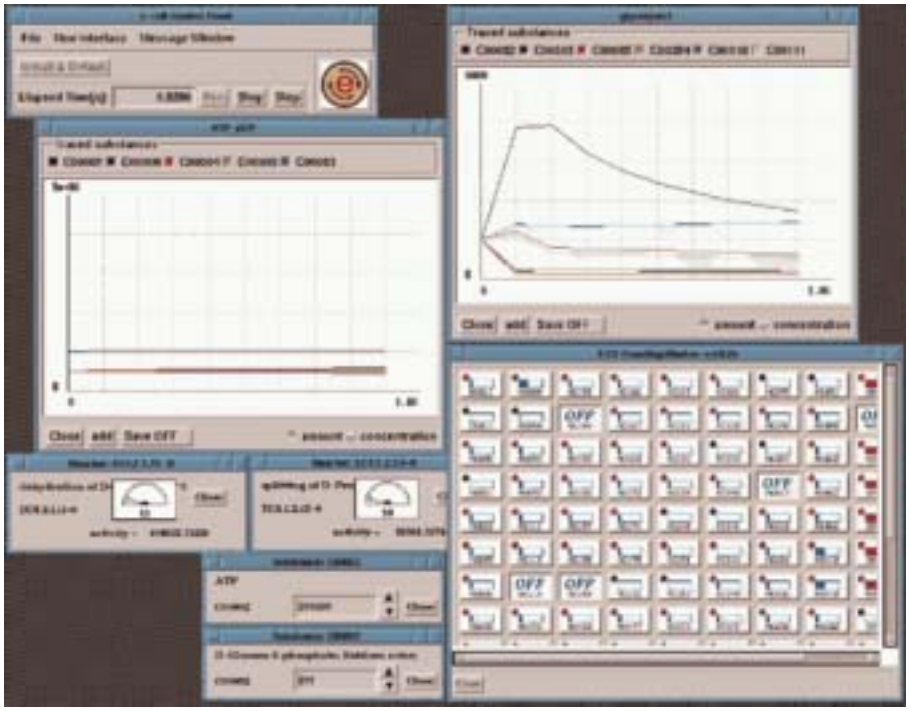


図 2.13: e-cell によるシミュレーションの様子

の解として求まる。また、

$$A_s\left(p_s - a + \frac{\mu' s^2}{\rho^2}\right) = bC_s, D_s\left(p_s - d + \frac{\nu' s^2}{\rho^2}\right) = B_s d.$$

であり、A_s, B_s, C_s, D_s は境界条件から決定される（チューリング論文では、フーリエ級数の利用により決まるとされる）。

この一般解をフーリエ級数の成分により分類し、円環内での解の漸近的挙動、すなわちこの円環内を伝播する化学物質の波で考えれば、停留状態と振動状態に分けられる。化学物質の濃度は、濃淡の波として初期の分布状態からしだいに拡散していく。

以上は線形系に対しての議論であったが、非線型項効果の考察もなされている。以下の式において、$u_{r,s}$ に-1 と+1 が等しい感覚で現れるように $h = 0.7$ としてこの値を計算し、(x, y) 平面に図示する（f が正の部分を黒く塗り潰す）と、図 2.14 のようなまだら模様が得られる。また、マインハートは 1995 年にチューリングモデルの基本式を様々に変更したシミュレーションの結果を集め発表した（図 2.15 参照）[26]。

$$f(x, y) = \sum u_{r,s} \exp\left[-\frac{1}{2}\left\{(x - hr)^2 + (y - hs)^2\right\}\right]$$

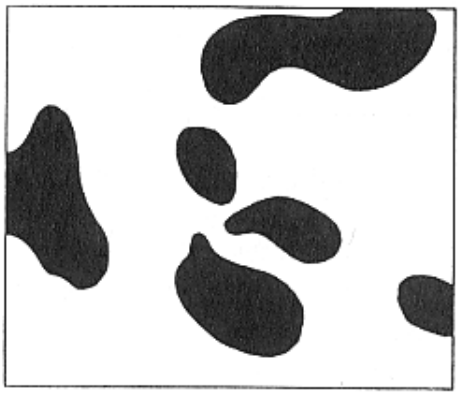


図 2.14: チューリングパターンによるまだら模様

表 2.2: ペトリネットの定義

ペトリネット $PN = (P, T, F, W, M_0)$, ここで
$P = \{p_1, p_2, \cdots, p_m\}$ は、プレースの有限な集合。
$T = \{t_1, t_2, \cdots, t_n\}$ は、トランジションの有限な集合。
$F \subseteq (P \times T) \cup (T \times P)$ は、アークの集合。
$W : F \to \{1, 2, 3, \cdots\}$ は、重み付け関数。
$M_0 : P \to \{0, 1, 2, 3, \cdots\}$ は、初期マーキング。
$P \cap T = \phi$ でかつ $P \cup T \neq \phi$ である。
特定の初期間ーキングが規定されていないペトリネットは N で表される。
初期マーキングが規定されているペトリネットは (N, M_0) で表される。

2.4.4 ペトリネット

ペトリネット（Petri Net）は 1962 年にドイツのペトリ（C.Petri）によって提案された。ペトリネットは有向グラフの一つであり、表 2.2 に示される定義中の要素から成る [31]。プレース及びトランジションと呼ばれるノードと、アークと呼ばれるプレースからトランジションもしくはトランジションからプレースへの接続によりペトリネットは成り立つ。プレースはトークンと呼ばれる物質の表現を格納し、トランジションは接続されているプレースから別プレースへとトークンを移動もしくは生成させる際に用いる条件を表す。

元々のペトリネットは離散モデルであったが、近年離散量と同時に連続量に関しても微分方程式系と同様に扱える Hybrid Petri Net[15] が開発され、より広範囲に渡って利用可能なシミュレーションツールとして注目を集めている。したがって、現在においてペ

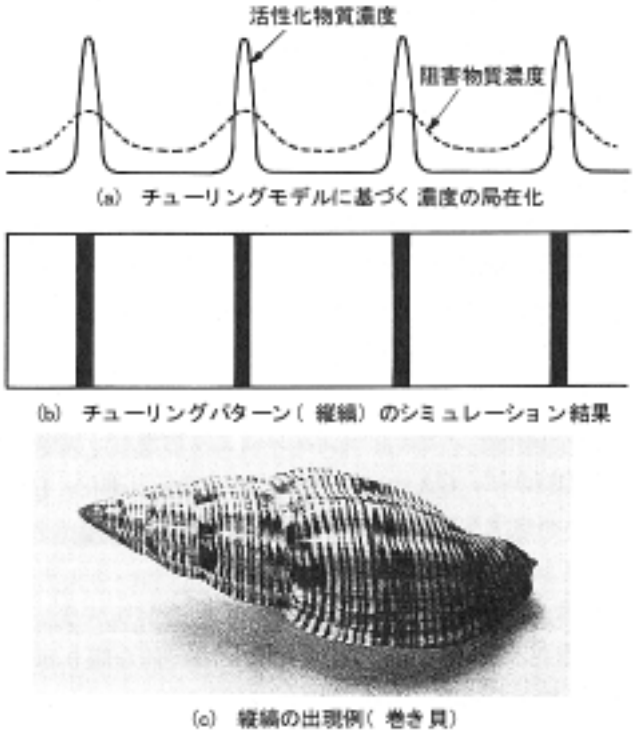


図 2.15: チューリングパターンによるシミュレーション

トリネットは可視化ツールとしての意味を求められる事が多い。生化学反応の記述がたやすく、理解もしやすいために特にバイオインフォマティクスでは有望。バイオインフォマティクス分野では、Visual Object Net++[30] などにより、生化学現象の記述やシミュレーションのしやすさなどに注力したシミュレーションツールが開発されている。Visual Object Net++によるシミュレーションの様子を図 2.16 に示す。

2.4.5 既存シミュレータの性質のまとめ

前小節までに示した主要なシミュレータの性質を以下の表 2.3 にまとめる。生化学シミュレーションに必要なシミュレータの能力をすべて満たしているものは無い。なお、ハイブリッドペトリネットは離散量と連続量の両方を扱えるものとし、連続値についての計算は微分方程式系のソルバーを用いているものとする。

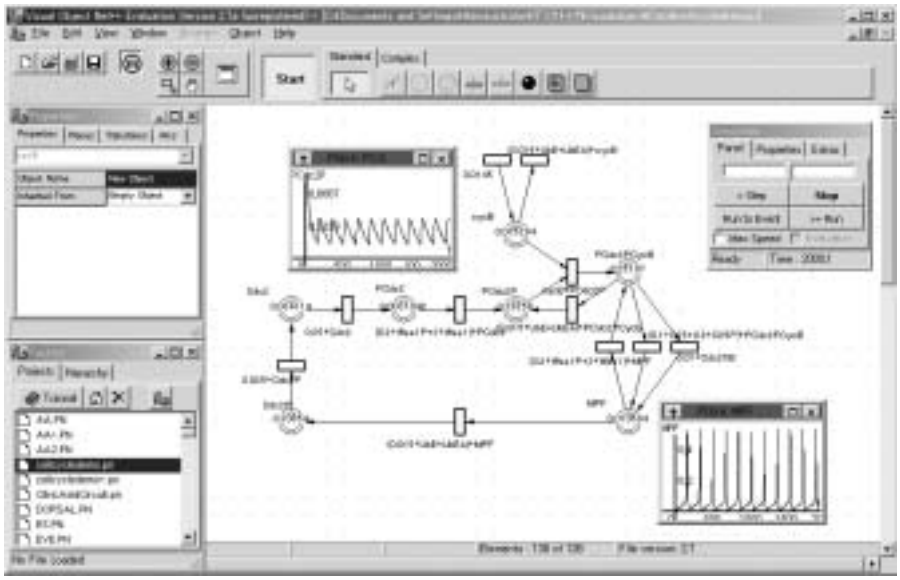


図 2.16: Visual Object Net++によるシミュレーションの様子

2.5 おわりに

本章では本論文に関する従来研究について以下の 3 点を示した。なお、この 3 つはそれぞれ「生物学からの視点」「構成論的生物学からの視点」「計算機工学からの視点」に対応している。

- 生化学反応の特徴
- 人工生命と複雑系

表 2.3: 主要シミュレータの性質のまとめ

シミュレータ	ゆらぎ	位置情報	個々の物質の表現	定量的計算
セルオートマトン	OK	OK	OK	NG
常微分方程式系	NG	NG	NG	OK
偏微分方程式系	NG	OK（濃度勾配）	NG	OK
Hybrid Petri net	NG	NG	OK	OK
Particle System	OK	OK	OK	OK

- シミュレーション手法

人工生命や複雑系の研究者が行なう生命現象のモデル化は、生命現象の本質を抽象的に再現しようと試みる。モデル化の際、そのモデルが実際の生物と比べて具体性に欠る事をいとわない。一方生物学者は実在する生物からデータを収集し、それぞれの部品としてのデータを総合する事で生命現象を説き明かそうとする。ポストシークエンス時代におけるシミュレーションの必要性に鑑みるに、両者の考え方を統合する事で新たな知見を見いだそうとする際に計算機工学の立場の人間は人工生命と生物学それぞおれの研究者に無理の無い形でシミュレータを提示する必要があるだろう。複雑系の本質であるところの「部分をシステム的に組み合わせた際にはじめて生まれる現象があり、その現象こそが対象の本質である」とする考え方を既存の複雑系における手法を参考にしながら取り込みつつ、生化学反応における化学物質の量や反応速度などの定量的なシミュレーションを可能にしなければならないだろう。

Particle System は、これらの特徴を取り込み（表 2.3 参照)、かつ定量的な化学反応をも再現し得るものである。ただし、現段階では単純な化学反応しか扱えていないが、複雑なネットワークへの拡張は容易である。さらに、ゆらぎを含んだ化学反応を扱えるという事は、シグナル伝達系などの非常に少数の分子が決定的な重要さを持つ反応を扱える可能性を持っている。これにより、定量的な予測力を持った人工生命的研究が可能となる事が期待される。

第 3章

Particle System

3.1 はじめに

本章では本論文で提案するシミュレーション手法 Particle System のアルゴリズムや実装についての詳細について説明する。本手法は反応物を Particle という離散的な単位に分割し、それらの位置情報を元にボトムアップ的な化学反応を行う事により反応系全体をシミュレーションするものである。

本章ではまず Particle System のアルゴリズムについて概略し、次に実装を含めて詳説する。また、計算結果の例について示し、Particle System が持つ「反応に関与する物質が多い場合は微分方程式系の解に近づく」「反応に関与する物質が少ない場合はゆらぎが大きくなる」という性質について確かめる。

3.2 基本設計

本節では、前節に示した設計方針を基本として作成した Particle System の基本設計について示す。

まず反応物のモデル化及びシミュレータ内での表現について述べ、次にシミュレーションの基本的な流れについて述べる。

3.2.1 化学反応のモデル化

生化学反応では、ある反応の生成物がその反応自体にかける負のフィードバックや自己触媒的反応がボトムアップ的に発生することにより、局所的に自己組織化が進行する場合

表 3.1: Particle の持つパラメータ

パラメータ名	説明
ID	各反応物にあらかじめ割り振られた ID
(X,Y)	平面上の位置

がある。このようなボトムアップ的な系の振る舞いを再現するために、トップダウン的に反応物の濃度を計算するのではなく、局所的な反応のシミュレーションを総合する事により全体の反応をボトムアップ的に構築する。

そのために、反応物を濃度として取り扱う事はせず、Particle（粒）として計算機内に再現し、それらが自分の周囲のみの情報から反応を起こす方式を採用した。つまり微分方程式系のように全体の濃度から一定時間後の濃度を算出するのではなく、Particle それぞれが自らの位置から一定の半径内に存在する Particle のみと反応を起こしうるとする。この時 Particle が反応しうる空間の半径を反応半径と定義する（図 3.1 参照）。

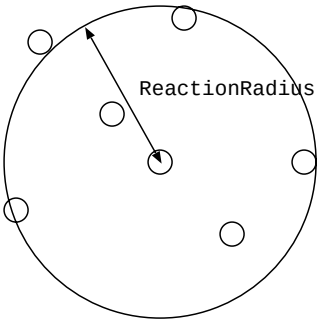


図 3.1: 反応半径：ある Particle に注目した時、その Particle を中心とした半径 r 内に存在する Particle のみと反応可能。

3.2.2 Particle のデータ構造

Particle は各々が表 3.1 に示されるパラメータを持つ。なお、本シミュレータは 2 次元平面上での反応をシミュレートする。

3.2.3 反応の流れ

反応はタイムステップ毎に 2 つのフェイズを交互に繰り返し行うことによって動的に再現される。2 つのフェイズとは「熱運動」及び「反応」である。以下に Particle System の簡単なアルゴリズムを示す。

1. Particle の初期配置（仮想的な平面を計算機内に再現し、その平面上に Particle を配置する）
2. 熱運動（各々の Particle を乱数を用いて移動）
3. 各々の Particle についての化学反応シミュレート
4. タイムステップを 1 増やし、2. に戻る

3.3 熱運動のモデル化

一般に熱運動は正規分布に乗るとされ、本シミュレーション手法でもランダムな熱運動の計算には正規分布を用いた。また、計算時間を短縮する目的のために time step 毎に $N(0,1)$ に従う正規分布の確率変数を用いた。

time step 毎にすべての Particle の熱運動を計算する事で、熱運動のフェイズは終了する。

3.4 化学反応式のモデル化

一般に化学反応の式は、速度定数を含め以下のように表される。

$$A + B \xrightarrow{k_1=2} C$$

本手法では、上記の式を以下のように解釈する。ここで確率 p は反応速度定数から推定するパラメータである。

$$A + B \xrightarrow{p_1=0.4} C$$

ある Particle を注目した時にその ID が”A” であった場合、周囲に”B” の ID を持つ Particle が存在すれば 40 %の確率で反応が進む。ただし k と p の値は同じではなく、k から p を決定するアルゴリズムに関しては 4 章で述べる。

また、”A” の周囲に”B” が n 個存在した場合は、”A” には n 回の反応の機会があると解釈し、以下の式を用いて反応確率 p_{all} を計算する。ただし、”B” が 1 個存在する時の反応確率を p_{one} とする。

$$p_{all} = \sum_{k=1}^{n} p_{one} \cdot (1 - p_{one})^{k-1}$$

したがって、周囲に反応可能な物質が大量にある場合、反応が進行する確率は限りなく 1 に近づく。

3.5 化学反応の基本アルゴリズム

熱運動フェイズの後、すべての Particle について化学反応をシミュレートする。化学反応は反応速度定数に制御され進行するが、本手法は注目した Particle の反応半径内に存在する Particle の個数により反応が制御される。

化学反応のチェックは以下のアルゴリズムによって行われる。多数の Particle が仮想平面上にランダムに位置していると仮定する。

1. すべての Particle の中からひとつの Particle を注目する。この注目された Particle について、化学反応をチェックする。
2. 注目された Particle に対して一定の半径内に位置する Particle のリストを作成する。
3. 化学反応式データの中から 1 つを選び、その反応が以下の条件を満たす場合にのみ反応を進ませる。すなわち、化学反応式の左辺に対応する Particle を消去し、右辺の生成物に対応する Particle を新たに生成する。
   (a) 反応式の左辺が注目された Particle 自身を含んでいる。
   (b) 反応式の左辺の全反応物が反応半径内に含まれている。
   (c) (a),(b) が満たされた時、反応確率により乱数を用いて反応するか否かを決定する。
4. 上のチェックをすべての反応式について行う。もし反応が成立した場合は次のステップに進む。
5. 前ステップの化学反応のチェックをすべての Particle について行う。終わったら 1. に戻る。

以上が簡略化した化学反応アルゴリズムの流れである。詳細については、プログラムの実装と高速化の節で述べる。

3.6 データ構造

本節ではシステムのデータ構造について示す。

システムが保持する仮想的平面は格子に切られており、それぞれの格子を表す「PSFieldUnit クラス」のインスタンスをシステムは保持する。また、Particle は「Particle クラス」として複数の Particle インスタンスが一次元双方向接続リスト構造を取る。たとえば平面上の $(3.2, 7.1)$ に位置する Particle は、仮想的平面中の格子 $(3, 7)$ を示すポインタの下に接続される。以下の図 3.2 にデータ構造について示す。

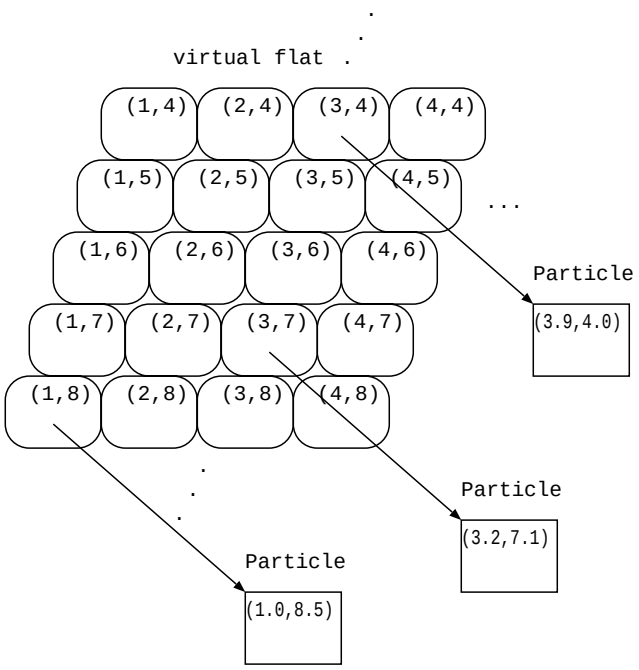


図 3.2: Particle と仮想平面のデータ構造：仮想平面の格子は PSFieldUnit クラスのインスタンスで構成されており、Particle は位置により対応する PSFieldUnit クラスのインスタンスに接続される。

3.7 アルゴリズム詳細及び実装と高速化

本節では、Particle System のシミュレーションアルゴリズムの詳細及び実装と高速化について示す。

表 3.2: Particle の変数

変数名	概略
`int id`	各反応物にあらかじめ割り振られた ID
`double x,y`	平面上の位置
`*nextP,*backP`	双方向リストの次の Particle と前の Particle を指すポインタ

表 3.3: PSFieldUnit の変数

変数名	概略
`int x,y`	平面上の位置
`*move_list,*reac_list`	2 つのフェイズに対応したポインタ

3.7.1 データ構造の詳細

PSParticle Particle を表すクラス`class PSParticle` は表 3.2 に示される変数を持つ。

Particle は位置によってグループ分けされ、それぞれのグループは一次元双方向接続リストとしてメモリ中に格納されている。図 3.3 にその様子を示す。

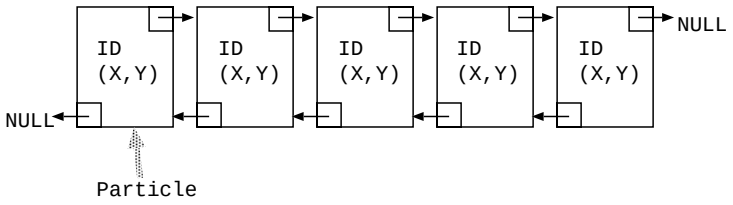


図 3.3: Particle の一次元双方向接続リスト

PSFieldUnit 格子を表すクラス `PSFieldUnit` は、仮想平面の格子毎にインスタンスが作成され、その格子内に存在する Particle の格納や取り扱いを行う。表 3.3 に`PSFieldUnit` の持つ変数について示す。

`PSFieldUnit` は Particle の一次元リストへのポインタを 2 つ持ち（図 3.4)、Particle の一次元リストはいずれかのポインタに接続される。2 つはそれぞれ「熱運動フェイズ用」と「化学反応フェイズ用」のポインタである。熱運動前の Particle は熱運動フェイズ用ポ

インタの下に接続されているが、熱運動が終了すると各々のParticleは化学反応フェイズ用ポインタに接続される。化学反応が終わると、Particleは再び熱運動フェイズに接続される（図3.5を参照)。この措置は「どのParticleが熱運動を終えたか」をチェックする手間を省くためのものである。

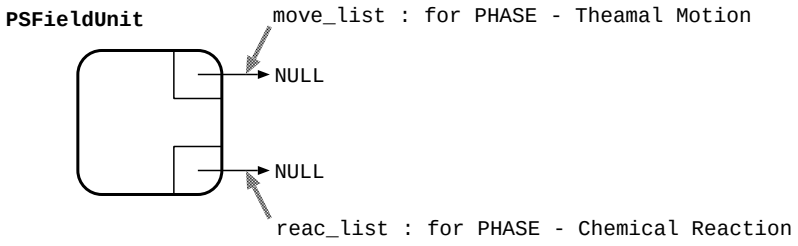


図 3.4: PSFieldUnitの構造

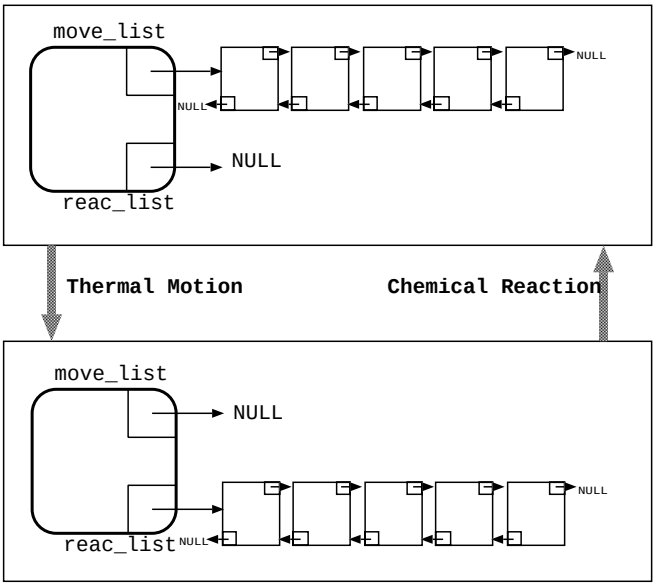


図 3.5: 2つのフェイズに伴うポインタ付け変え

3.7.2 ランダムな熱運動計算の詳細

本手法では、単位時間毎に$N(0,1)$に従う正規分布の確立変数を用いた。また、計算機において正規分布を算出する手法としてBox-Muller法[34]を用いた。

Box-Muller 法は 2 つの独立な一様乱数 U_1, U_2 から 2 つの互いに独立な標準乱数 Z_1, Z_2 を生成する手法で、計算機を用いれば比較的高速に計算ができる [34]。以下にその計算式を示す。

$$Z_1 = \sqrt{-2\log U_1}\cos 2\pi U_2,$$
$$Z_2 = \sqrt{-2\log U_2}\sin 2\pi U_2,$$

実装に際しては、容易に熱運動による位置の移動の大きさを容易に変化させられるよう、Box-Muller 法により算出した乱数に一定の重みを掛け合わせている。

3.7.3 化学反応アルゴリズムの詳細

以下に詳細な化学反応アルゴリズムを示す。なお、アルゴリズムのステップ中には計算の高速化のために用いたアルゴリズムも含めている。

1. ある Particle に注目する。
2. 反応半径内にある Particle の ID を調べ、ID 毎の個数を数え上げ、必要な個数だけそれぞれの Particle へのポインタを配列に保持する。
3. 反応式をランダムで一つ選び、前ステップで数え上げた ID 毎の個数と照らし合わせる事で、以下の条件が満たされているかを判定する。
   (a) 注目している Particle が反応式の左辺に現れている
   (b) 反応式の左辺が要求している個数の Particle が反応半径内に存在する
4. 前ステップで「反応可能」と判定された場合は、以下のアルゴリズムを用いて反応確率を算出する。なお、本論文内では 2 体反応のみを扱うものとする。
   (a) 反応式が進む確率を参照し、p に格納する。
   (b) 注目している Particle の ID ではない項の ID を持つ Particle の個数を $other$ とした時に、以下の算出式を用いて p を求める。$p = \sum_{n=1}^{other} p \cdot (1-p)^n$
   (c) 前ステップで求めた p を確率とし、反応を進める。
5. 前ステップで反応が進んだ場合は、以下のアルゴリズムを用いて Particle の ID 変更、Particle 自体の消去、生成を行う。ここで、成立した反応式の左辺の総 Particle 数を $left$、右辺の総 Particle 数を $right$ とする。

- $left - right > 0$、すなわち反応後に Particle が消える場合は、$left - right$ 個の反応に関与している Particle をメモリ上から消去する。
- $left - right < 0$、すなわち反応後に Particle が増える場合は、$right - left$ 個の Particle をメモリ上に生成し、格子クラスのインスタンスの熱運動フェイズ用ポインタに接続する。
- 前の 2 つのステップを行った後、反応に関与した Particle 及び生成した Particle の ID を反応式の右辺に出てくる反応物の ID にそれぞれ変更する。また、ID を書き換えたのみの Particle を熱運動フェイズ用ポインタに接続する事で再び反応を起こさないようにする。

6. 前ステップの操作をすべての Particle について行ったのち、熱運動フェイズに移行する。

3.8 シミュレーション例

表 3.4 で示したパラメータと以下の反応式を用いてシミュレーションを 10 回ずつ行った結果を図 3.6, 図 3.7, 図 3.8 に示す。図の縦軸には $[A]$ の個数を初期個数 $[A]_0$ で割った値を、横軸には time step 数をとっている。正反応の反応確率のほうが大きいために、$[A]$ は現象し、やがて平衡状態に落ち着く。なお、この実験では平衡状態におけるゆらぎを観測するために、反応確率を適宜修正している。

10 回試行を行なった結果、初期の Particle の個数が少ないほど平衡時におけるゆらぎが相対的に大きくなる事が確認された。

$$A + B \xrightarrow{p_1} C + D$$

$$A + B \xleftarrow{p_{-1}} C + D$$

平衡状態においては、Particle の初期個数が多いほど近い値に収束している。個数を増やしていくとより近い値に収束していき、微分方程式系と同様の計算になる。一方 Particle の数を少なくしていくと平衡状態時の値はゆらぎが大きくなる。

微分方程式系との比較や微分方程式系と同様の定量的なシミュレーションを Particle System で行う方法については 4 章で論じる。また Particle System のシミュレーション手法に由来するゆらぎについては 5 章で論じる。

表 3.4: 例題に用いたパラメータ

パラメータ名	以降で用いる記述
反応半径	$r = 1$
反応確率	$p_1 = p_{-1} \times 20$
仮想平面の大きさ	$S = 100 \times 100$
Particle の数	$[A]_0 = [B]_0 = (10, 100, 1000), \quad [C]_0 = [D]_0 = 0$

3.9 おわりに

本章では、Particle System のアルゴリズムや実装についての概要を示した。Particle System は局所的な情報のみに依存した反応を繰り返す事で系全体の振舞いを再現するボトムアップ的なアルゴリズムを持ち、反応に関与する物質に対応した Particle が多い場合には微分方程式系の解に近づき、少ない場合にはゆらぎが大きくなる。

微分方程式を用いて表現された生化学反応モデルは生化学シミュレーションにおいて主流であり、シミュレーションに定量的な精度が求められる場合が多い事から非常に有用性が高いとされている。Particle System はボトムアップなアルゴリズムから微分方程式系モデルの持つダイナミクスを再現している。一方微分方程式系ではカバーできない「ゆらぎ」のシミュレートについても Particle System はアルゴリズム自体に大きな変更を施す事なく成功している。反応に関与する Particle の数が少ない程ダイナミクスはゆらぎが大きくなる。ゆらぎは生化学反応において近年注目されている現象であることから、Particle System は定量的なシミュレーションとゆらぎを同時にシミュレートできる手法として有用性があると思われる。

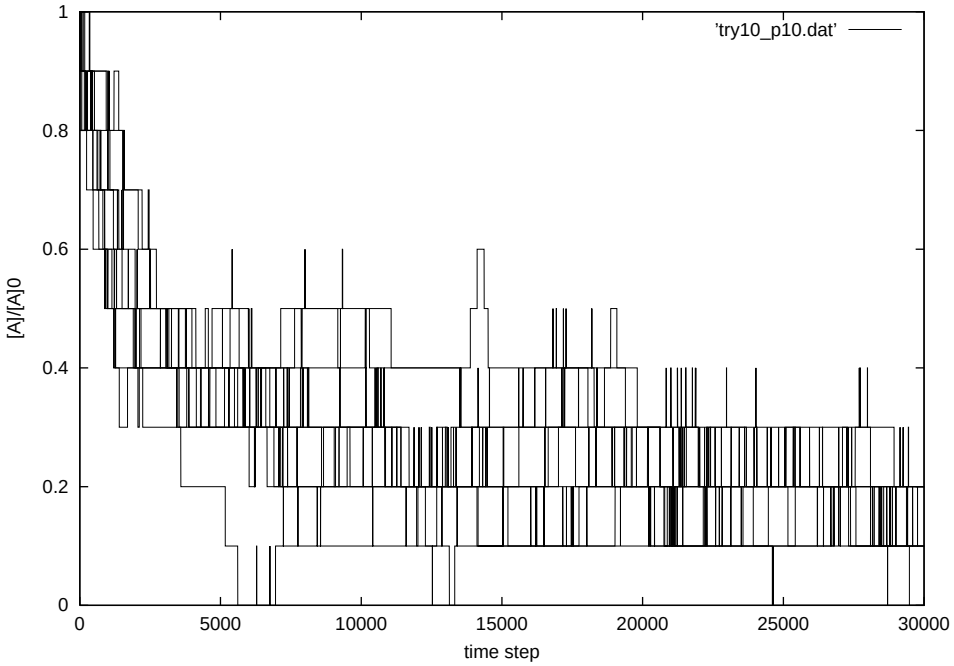


図 3.6: 結果例（$[A]_0 = [B]_0 = 10$）：横軸は time step を、縦軸は Particle の個数の初期個数に対する比を表している。10 回の試行結果のうち、一番大きいところで 6 個も途中結果に差が生じており、ゆらぎも大きい。

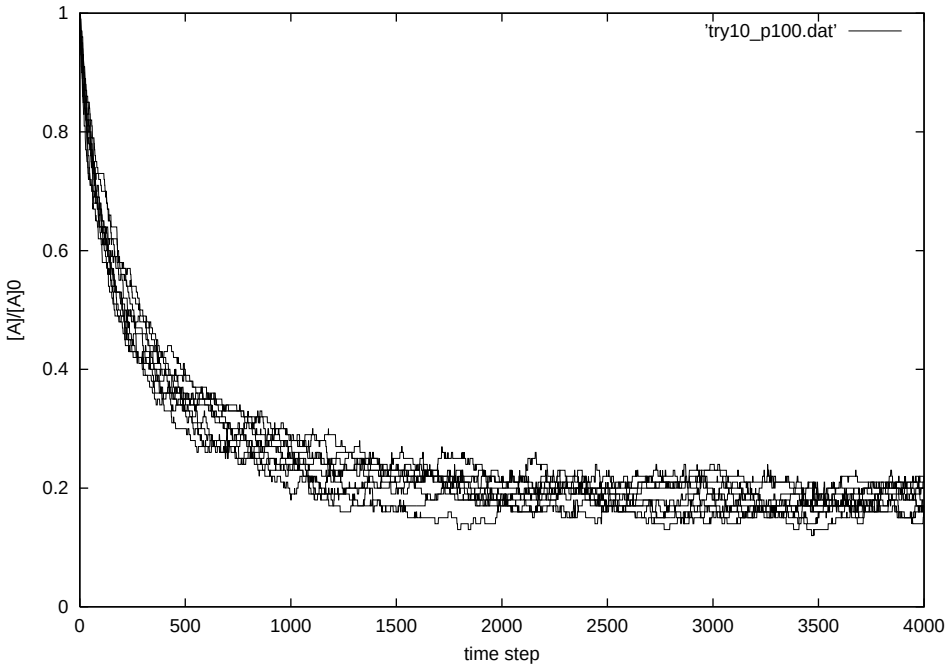


図 3.7: 結果例（$[A]_0 = [B]_0 = 100$）：$[A]_0 = [B]_0 = 10$ に比べてよく収束している。

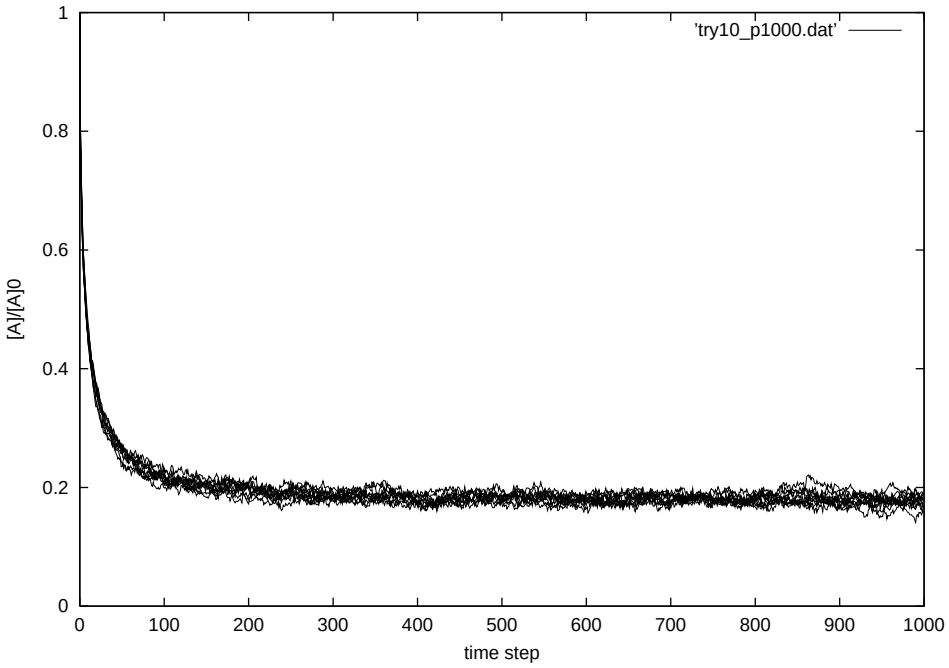


図 3.8: 結果例（$[A]_0 = [B]_0 = 1000$）：$[A]_0 = [B]_0 = 100$ に比べてよく収束している。

第 4章

Particle System と微分方程式の比較

4.1 はじめに

本章では、Particle System を用いた計算においてパラメータを様々に変化させた際のシステムの挙動を示し、微分方程式系と同等な性質を持つ事を示す。

系には十分にたくさんな反応物が介在し、かつ十分に撹拌されている状態を仮定した場合、Particle System は微分方程式系と同じ挙動を示す事が確認された。なお、上記の条件を満たさない場合の挙動については次章で論じる。

本章の実験に用いた反応系　本章の実験には、以下に示す簡単な化学反応を用いた。

$$A + B \leftrightarrow C + D$$

ただし、左辺から右辺への反応確率と右辺から左辺への反応確率の比を $(20:1)$ と設定している。この条件の反応式を微分方程式系で解くと、平衡時の $[A]$ は $0.1827...$ となる。

4.2 反応確率とシステムの挙動

Particle System のパラメータのうち、正反応の反応確率を変化させ、シミュレーションを行った。システムサイズは 100、反応半径は 0.4、初期個数は $A = B = 1000, C = D = 0$ とした。

Particle System のパラメータのうち、反応確率を変化させた際の計算結果について以下の図 4.1 に示す。

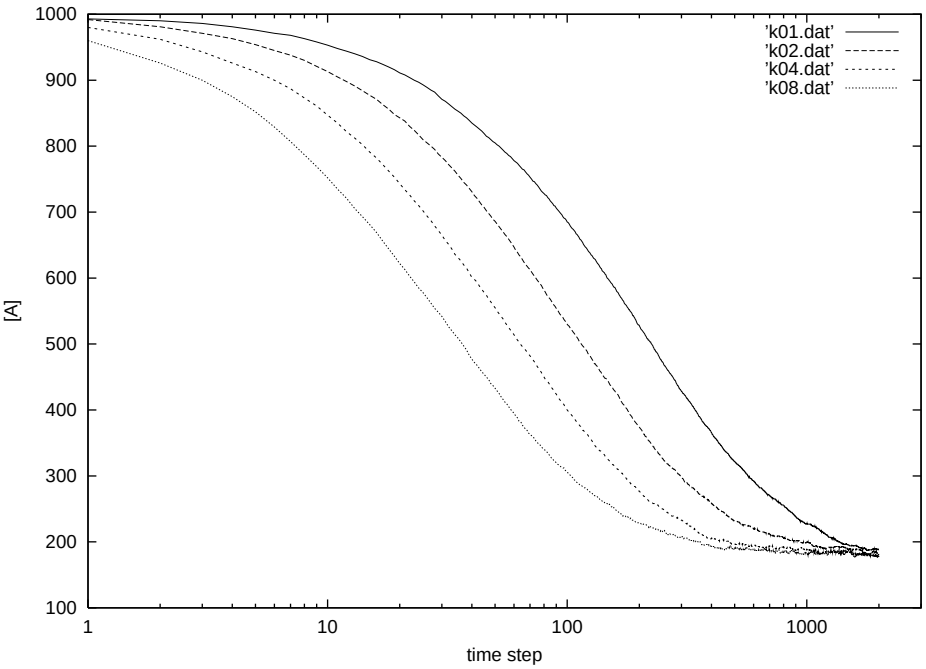


図 4.1: 反応確率を変化させた場合の A の個数の時間変化：横軸は time step 数を、縦軸は Particle の個数を表している。

図から分かるように、各反応確率における A がある値になるまでの時間が、反応確率と反比例している。

4.3 反応半径とシステムの挙動

Particle System のパラメータのうち、反応半径を変化させシミュレーションを行った。システムサイズは 100、反応確率は 0.4、初期個数は $A = B = 1000, C = D = 0$ とした。

Particle System のパラメータのうち、反応半径を変化させた際の計算結果について以下の図 4.2 に示す。

各反応半径における A がある値になるまでの時間が、反応半径と反比例している。

4.4 反応確率と反応半径の関係

Particle System のパラメータのうち、反応確率と反応半径の関係を調べるべく、反応半径を 2 倍にした時に反応確率を $\frac{1}{4}$ 倍にした場合でのシミュレーションを行った。

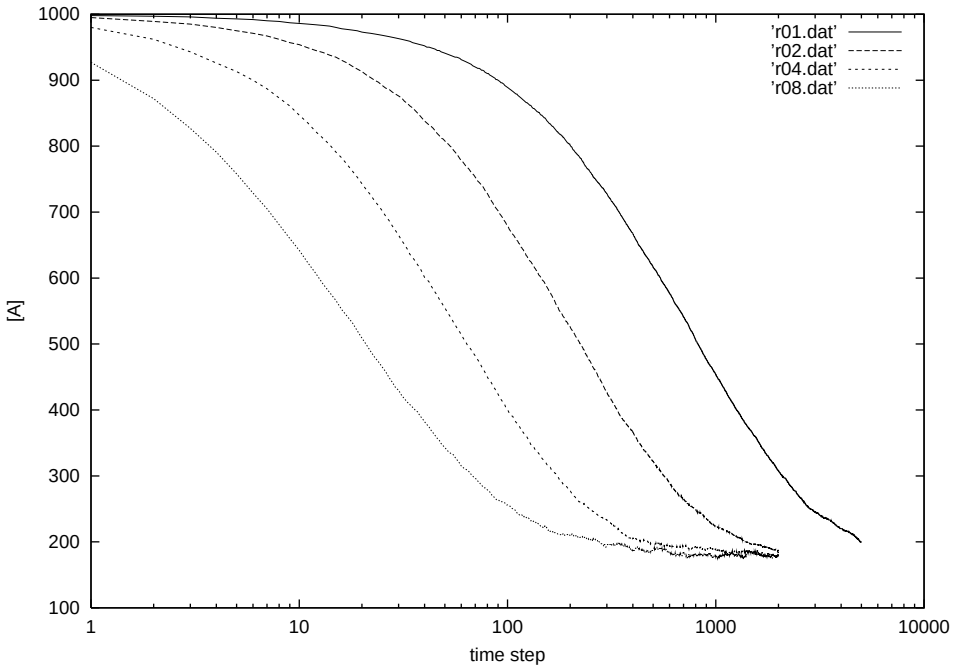


図 4.2: 反応半径を変化させた時の A の個数の時間変化：横軸は time step 数を、縦軸は Particle の個数を表している。

結果を以下の図 4.3 に示す。

反応半径を 2 倍にした時の結果と反応確率を $\frac{1}{4}$ 倍にした時の結果はきれいに一致した。この実験における設定は、Particle が認知できる領域の広さが 4 倍に設定し、反応確率を $\frac{1}{4}$ 倍に設定した事と等価である（図 4.4 参照)。Particle System によるシミュレーションが同じ振舞いを示す時の r と k の関係を式に表すと、以下のように示される。

$$r^2 \cdot k = Const.(0 < r, k \leq 1)$$

4.5 定量化 Particle System

生化学反応シミュレーションにおいて、微分方程式系によるモデル化が最も広く用いられている。微分方程式系によるモデル化は簡潔であるが、生化学反応シミュレーションへの適用の際には注意が伴う。一方 Particle System は微分方程式系の持つ弱点を克服してはいるものの、微分方程式系と同じ前提条件下においては、Particle System も同様のシミュレーション結果が出力できなければならない。

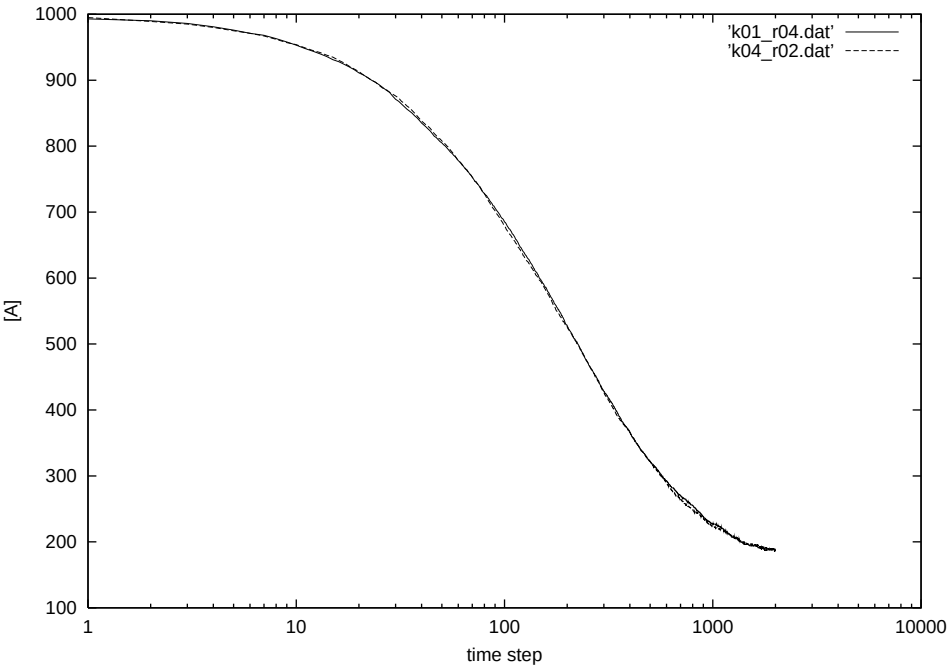


図 4.3: あるパラメータでの A の時系列と、反応確率を 2 倍、反応確率を $\frac{1}{4}$ 倍にした場合の A の時系列：2 本のプロットはほぼ一致しており、$r^2 \cdot k = Const.(0 < r, k \leq 1)$ が示される。

そこで本節では、常微分方程式系によって与えられた系に関して、反応に関与する物質が十分に多く撹拌されている状況下において Particle System でシミュレーションさせるためのパラメータ決定について論じる。微分方程式系と同じ前提条件下においては、Particle System も同様のシミュレーション結果が出力できなければならない。

4.5.1 概略

一般的な二体反応について、微分方程式系で記述された反応系を Particle System を用いてシミュレーションを行う事を目的とする。その際、以下の前提条件の下で、Particle System が微分方程式系によるシミュレーション結果とよく一致する事を示す。

- 系に関与する反応物が十分な数存在する。
- 系に関与する反応物は局在などが無く、一様に分布しているものとする。また、拡散のスピードは十分に早いものとする。

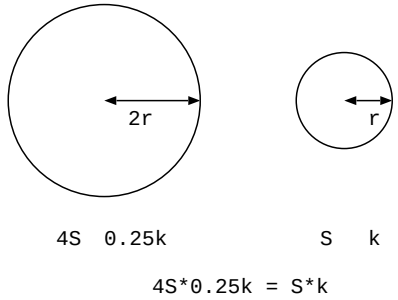


図 4.4: 挙動が一致するための r と k の条件

表 4.1: ブリュセレーターモデルの持つパラメータ

モデル	パラメータ名	以降で用いる記述
微分方程式系	初期濃度 [mol]	$[A]_0, [B]_0, [C]_0, [D]_0$
	反応速度定数	k_1, k_{-1}
Particle System	反応半径	r
	反応確率	p_1, p_{-1}
	仮想平面の大きさ	S
	1mol あたりの Particle	d
	1 タイムステップあたりの時間 [sec]	Δt

例として、微分方程式系モデルで示された以下の化学反応系を用いる。ただし、$[A]_0 = [B]_0 = A, [C]_0 = [D]_0 = 0$ とする。

$$A + B \xrightarrow{k_1} C + D$$

$$A + B \xleftarrow{k_{-1}} C + D$$

上記のモデルを、以下に示すような Particle System のモデルに変換する。

$$A + B \xrightarrow{p_1} C + D$$

$$A + B \xleftarrow{p_{-1}} C + D$$

この時、考慮しなければならないパラメータを表 4.1 に示す。

ここで「1mol あたりの Particle」とは、微分方程式中で 1mol として表現されている物質を、いくつの Particle に分割するかという事を意味している。例えば、微分方程式系で 2mol で表現されている物質を 8000 個の Particle で表現する場合、$d = \frac{2}{8000} = 4000$ となる。以上のパラメータを、微分方程式系によるシミュレーションの結果と Particle System によるシミュレーションの結果が一致するように決定する。ここで、前節までの説明より、Particle System においては、r と k について以下の関係が成り立つ。

$$r^2 \cdot k = Const.(0 < r, k \leq 1)$$

つまり、r が 2 倍、3 倍... となると、k は $\sqrt{2}$ 倍、$\sqrt{3}$ 倍... となる。そこで、以降 r を $r = 1$ と固定して考える事とする。また、S, d, t も固定して考える。ここで、微分方程式系の反応方程式における反応速度定数 k_1, k_{-1} に対して、以下の演算を行う事で Particle System のモデル化を行う事を考える。

$$p_1 = P \cdot k_1 \quad , \quad p_{-1} = P \cdot k_{-1}$$

上式の P を決定すれば、微分方程式系で表されたモデルを Particle System へ適用する事が可能になる。

4.5.2 パラメータ決定

本小節では、前小節の P を決定するアルゴリズムについて示す。仮想平面を各 Particle の反応半径による円の領域で塗り潰す事を考える。1 秒間に S を円で塗り潰すとすると、Δt 秒間には S のうち tS の領域を塗り潰せば良い。例えば 0.01 秒間には $S = 10000$ のうち $\Delta tS = 0.01 \cdot 10000 = 100$ の広さを塗り潰せば良い。Particle の個数を d とし、かつこの時の反応半径を R とすると、以下の式が成り立つ。

$$\pi R^2 \cdot d = \Delta tS$$

よって R は以下の式で求められる。

$$R = \sqrt{\frac{\Delta tS}{\pi d}}$$

ただし、本来 Particle はランダムに空間に配置されるので、S の全体を覆う事はできない。しかしここでは S の全体をカバーできると仮定して説明を続ける。上式から算出された R は事前に固定した r とは異なる値を持つ。ここで、もし R が r の 2 倍だとすると、反応確率への重み P は $r^2 \cdot k = Const.$ の関係から以下の式で算出できる。

表 4.2: パラメータ決定の例

パラメータ名	以降で用いる記述
反応半径	$r = 1$
仮想平面の大きさ	$S = 100 \times 100$
1mol あたりの Particle	$d = 1000$
1 タイムステップあたりの時間 [sec]	$\Delta t = 0.01$

$$P = \left(\frac{R}{r}\right)^2 = R^2 = \frac{\Delta t S}{\pi d}$$

上式で求められた P を用いて反応確率を補正すれば、微分方程式と同様の計算を Particle System を用いて可能である。

Particle System のパラメータを表 4.3 に固定した時の R は、以下のように求められる。

$$P = R^2 = \left(\frac{\Delta t S}{\pi d}\right) = \left(\frac{0.01 \cdot 100 \cdot 100}{\pi \cdot 1000}\right) = 0.031831...$$

ここで算出した P を用いて Particle System により計算した結果を以下の図 4.5 に示す。

同じモデルに関してのシミュレーションを行ったにも関わらず、計算結果はずれてしまっている。

P を算出する際に、反応半径を半径として持つ d 個の Particle により S の全体が被覆されているという前提条件があった。しかし実際は Particle はランダムに配置されているために、すきまが発生しているはずである（図 4.6 参照）。

ここで、正確な P を求めるためには、Particle が d 個あった時に仮想平面の何割の領域が塗り潰されるかをシミュレーションにより求める必要がある。このシミュレーションは単純なモンテカルロ法である。

そこで、表 4.3 の条件におけるモンテカルロシミュレーションを行い、実際に被覆されている領域は $S' = 9278.3...$ と求められた。そこでこの S' を'S の代りとして計算すると、$P' = 0.343052...$ となる。ここで求められた P' を用いて決定したパラメータを用い、Particle System によるシミュレーションを行った結果を以下の図 4.7 に示す。微分方程式系によるシミュレーション結果とほぼ一致している。なお、この実験では同様のシミュレーションを 20 回繰り返し、その平均をプロットしている。

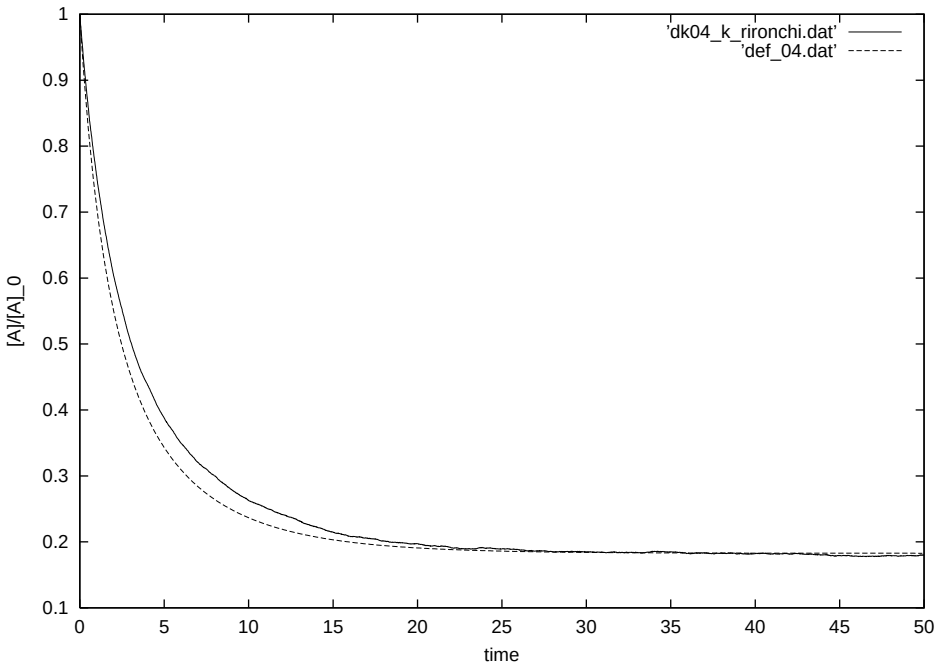


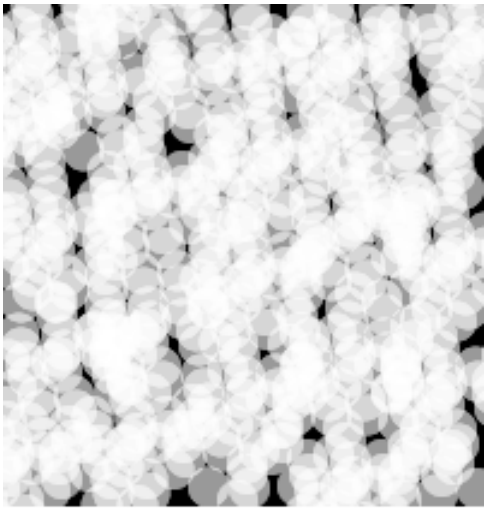
図 4.5: P を用いたシミュレーション結果：縦軸は $\frac{[A]}{[A]_0}$ を、横軸は時間を表している。実線は微分方程式系による解を、破線は Particle System によるシミュレーションの 10 回試行を平均化した値を用いている。

パラメータ推定アルゴリズムのまとめ　Particle System で微分方程式系と同様のシミュレーションを行う場合に必要になる P の決定アルゴリズムを以下にまとめる。

1. $S, \Delta t, d$ を固定する。また $r = 1$ とする。
2. 1. の条件の元でモンテカルロシミュレーションを行い、実際に被覆される領域の広さ S' を求める。
3. 前ステップで求めた S' を使って、$P = \frac{S'\Delta t}{\pi d}$ を求める。
4. 微分方程式系の反応式に含まれる各反応速度定数に P を乗じ、Particle System における反応確率とする。

4.5.3　結論

本小節では、微分方程式系と同様のシミュレーションを Particle System で行う場合のパラメータ推定について論じた。その結果、微分方程式系における反応速度定数から



※ 黒い領域が塗り残し

図 4.6: Particle に覆われる仮想平面

Particle System の反応確率を算出する式とモンテカルロシミュレーションを用いる事でパラメータを決定できる事が分かった。また、提案したパラメータ決定アルゴリズムを用いてシミュレーションを行った結果、Particle System は微分方程式系と同様のシミュレーションを行える事が示された。

4.5.4 微分方程式と一致しない場合

Particle System の目的のひとつに「反応物が希薄な状態に発生するゆらぎ」をシミュレートする点が挙げられるが、逆に反応物が非常に高い圧力下にある状態や高密度の状態をシミュレートすると、平衡点がずれる現象が起きた。

あるパラメータにおいて、Particle の数を保ったまま仮想平面の大きさを縮小していったところ、微分方程式系とのずれが確認された。密度が高くなればなるほど、平衡点が正反応平衡がずれる様子について、図 4.8 に示す。図 4.8 は横軸に仮想平面の正方形の一辺を、縦軸には平衡に達したと思われた時から 2000 ステップにおける $\frac{[A]}{[A]_0}$ の平均値をとっている。仮想平面が大きい場合は微分方程式系シミュレータによる平衡時の値とほぼ一致しているが、仮想平面が小さくなるほど平衡は正反応の方向にずれてしまっている。

このような現象が発生する理由について、以下のように予想した。このような現象が発生する状況は特殊な場合に限られているので、Particle System 自体の妥当性には問題は

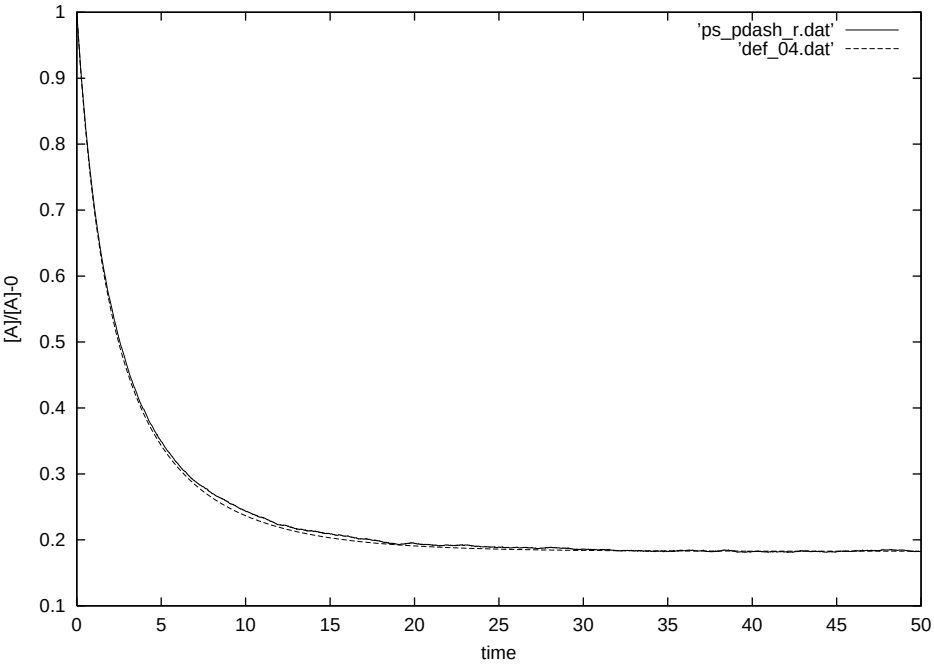


図 4.7: モンテカルロ法による補正を行ったシミュレーション結果

無いと思われるが、興味深い現象でもある。

1. Particle の密度が常に高いために、反応半径内には常に正反応についても逆反応についても反応可能なほどに Particle が存在している。

2. 反応する条件を以下の 2 点とする場合、1 点目の条件は常に満たされており、反応する量は 2 点目の条件によってのみ決定される事になる。

   (a) 反応半径内に反応式が進むのに十分な個数の Particle が存在している事。

   (b) 反応確率が満たされる事。

3. このような条件下では、正反応と逆反応の反応速度が Particle の個数による影響を受けにくい状態となる。この場合では、正反応に用いられる Particle である A, B の数が少なくなっても依然として正反応の反応速度が鈍らない現象が起こる。

4. よって、平衡が正反応の方向にずれてしまう。

表 4.3: 平衡点のずれについての実験に用いたパラメータ

パラメータ名	以降で用いる記述
反応式	$A + B \xrightarrow{p_1} C + D, \quad A + B \xleftarrow{p_{-1}} C + D$
反応半径	$r = 1$
反応確率	$p_1 = 0.4, p_{-1} = 0.02$
仮想平面の大きさ	$20 \times 20 \leq S \leq 120 \times 120$
Particle の数	$[A]_0 = [B]_0 = 1000, [C]_0 = [D]_0 = 0$

4.6 おわりに

本章では、Particle System が満たすべき重要な要求のひとつである「微分方程式系と同じ前提条件下において、微分方程式系と同様に定量的なシミュレーションが可能」というシミュレータの能力を満たしている事を示した。なお、本章では理化学研究所の吾妻竜三さんの研究を参考にさせていただいた。

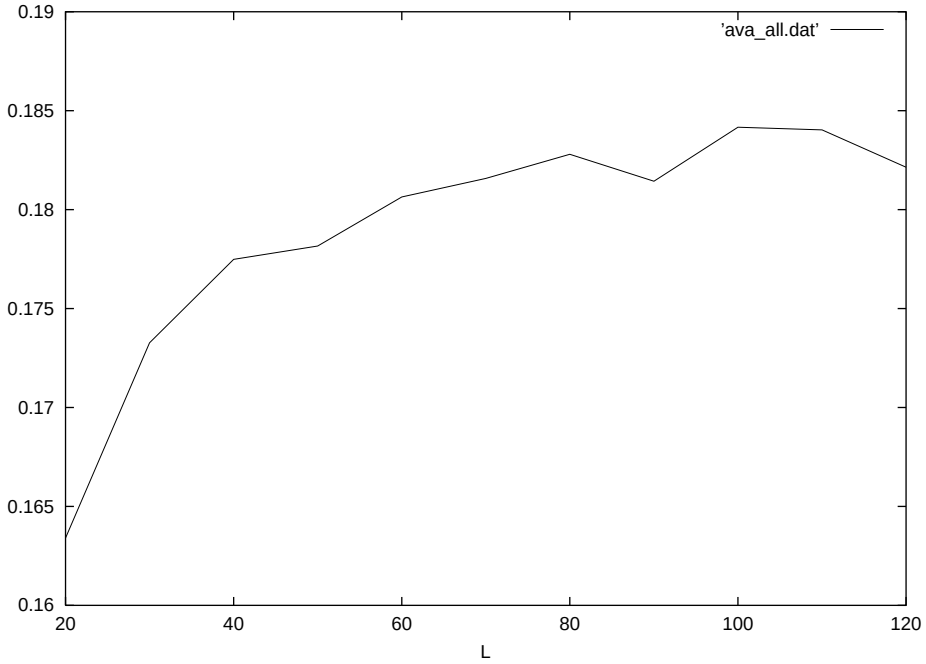


図 4.8: 仮想平面の大きさによる平衡点の変化：縦軸は平衡点における $\frac{[A]}{[A]_0}$ を、横軸は正方形である仮想平面の一辺の長さ L を表している。L を小さくしていくと、平衡点は正反応の方向にずれていく事が分かる。

第 5 章

Particle System のゆらぎ

5.1 はじめに

本章では、Particle System の計算結果のゆらぎについて示す。

まず Particle の個数とゆらぎの関係について示し、次に Particle 密度とゆらぎの関係について示す。加えて、前章のパラメータ決定アルゴリズムに用いたモンテカルロ法について検討する。

本章では「ゆらぎ」を評価する値として正規化した確率変数の標準偏差を用いている。一般的に、ある確率変数 X について $< X^2 > - < X >^2$ をゆらぎとしているが、本質的には同義である [35]。

5.2 ゆらぎの例

Particle の個数が少ないほどゆらぎは大きくなる。表 5.1 のパラメータで実験を行った際の結果を図 5.1, 5.2, 5.3 に示す。個数が多くなるにつれて平衡状態時のゆらぎ幅が小さくなっていくのが分かる。

5.3 Particle の個数とゆらぎ

Particle System の計算結果におけるゆらぎを計測するために、実験を行った。実験では以下の化学反応式を用いた。

$$A + B \leftrightarrow C + D$$

表 5.1: 例題に用いたパラメータ

パラメータ名	以降で用いる記述
反応半径	$r = 1$
反応確率	$p_1 = p_{-1} \times 20$
仮想平面の大きさ	$S = 100 \times 100$
Particle の数	$[A]_0 = [B]_0 = (10, 100, 1000), \quad [C]_0 = [D]_0 = 0$

ただし、左辺から右辺への反応の確率 $p_1 = 0.4$、右辺から左辺への反応の確率は $p_{-1} = 0.02$ とした。また、物質の初期設定は、A, B を変化させ、C, D は 0 で固定した。この条件下で、Particle を配置する仮想平面の大きさと A, B の初期設定個数を変化させ、各パラメータ毎に 10 回同様の反応を行った。その際、反応が十分に平衡状態に達してからの各タイムステップ毎に A の個数の平均と平均からの誤差の標準偏差を計測した。

5.3.1 結果

システムサイズを複数とり、Particle の数を変化させた時の結果を以下の図 5.4 に示す。

Particle の個数とゆらぎの関係を調べた実験により結果、ゆらぎの指標として用いた標準偏差の対数は Particle の個数の対数と比例関係を示した。これは大数の法則に則った振舞いであると考えられる（大数の法則の導出は付録参照のこと）。したがって、 Particle System は大数の法則が成り立つ範囲で妥当だという事が言える。

5.4 Particle の密度とゆらぎ

反応半径と仮想平面の大きさとの関係について調査すべく、Particle の個数と反応確率と反応半径を一定にしたままシステムサイズを変化させた場合のゆらぎについて実験を行なった。結果を 5.5 に示す。

考察 計算結果を 3 つの領域に分け、それぞれについて考察を加える。

$L < 5$ の場合、ゆらぎは 0.57 近辺の値を取る。このゆらぎは、Particle System の離散性と確率過程により発生する最低限のゆらぎであると考える事ができる。

$5 < L < 20$ の場合、ゆらぎは仮想平面の一辺の長さ L に対して対数的に増加している。

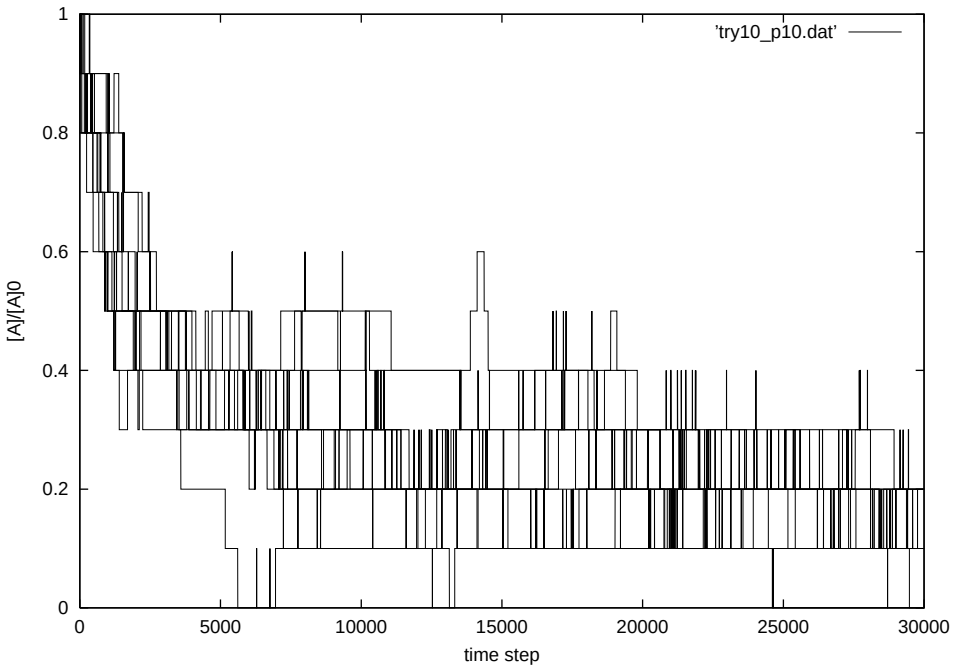


図 5.1: 結果例（$[A]_0 = [B]_0 = 10$）

$L > 20$ の場合、ゆらぎは収束が悪いものの、概してグラフは $5 < L < 20$ の延長ではなくある値に収束していると思われる。Particle System は溶液中での反応を想定した熱運動アルゴリズムを用いているため、溶液が希釈されると相対的に移動面積は小さくなる。よって Particle 同士の衝突数は少なくなり、分散の増大を制限していると考えられる。

5.5 被覆率とゆらぎ

本節では、パラメータ決定において用いたモンテカルロシミュレーションについての検討を加える。仮想平面を Particle の認識可能な領域で塗り潰した場合の面積の割合について調べた。

実験の手順を以下に示す。また、決定したパラメータを表 5.2 に示す。

1. Particle の個数 $d = 1000$ とし、1000 個の Particle がちょうど仮想平面の大きさ S を覆うように r を定めた。すなわち $\pi r^2 \cdot d = S$ であり、$r = \sqrt{\frac{S}{\pi \cdot d}}$ である。

2. 1. で用いた r を固定して、Particle の個数を変化させつつ、被覆率をモンテカルロ法により算出した。

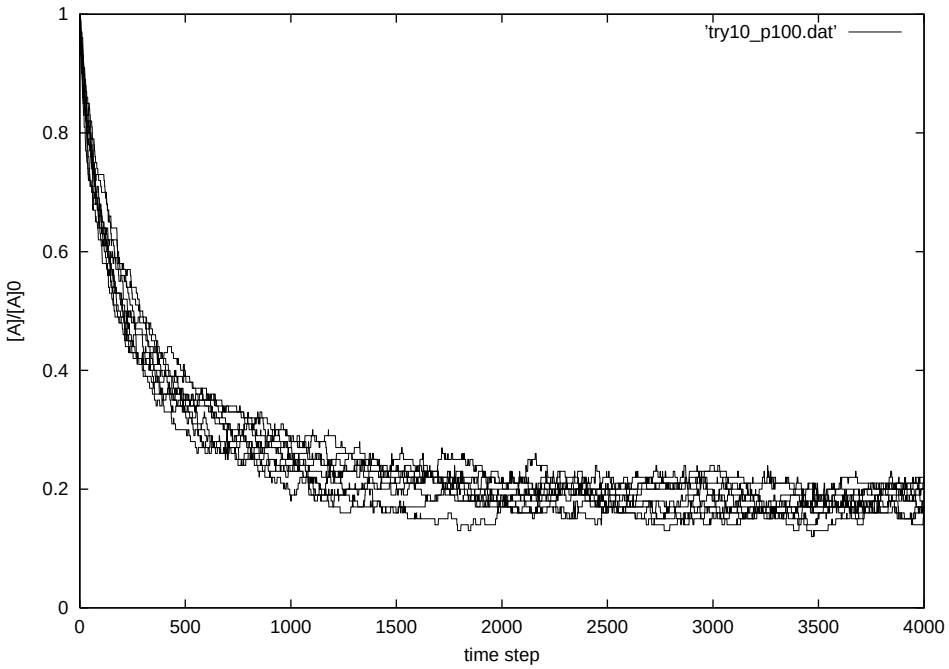


図 5.2: 結果例（$[A]_0 = [B]_0 = 100$）

5.5.1 結果

実験結果を図 5.6 に示す。

ちょうど 1000 個の Particle の時の充填率は約 0.59 であった。つまり仮想平面のうち約 41 %は、重複により被覆されなかった部分だったという事になる。約 2000 個の段階で被覆率が 90 %を越え、以降は 100 %に漸近した。

表 5.2: 充填率の調査に用いたパラメータ

パラメータ名	以降で用いる記述
反応半径	$r = 1.7481$
仮想平面の大きさ	$S = 100 \times 100$
1mol あたりの Particle	$d = 1000$

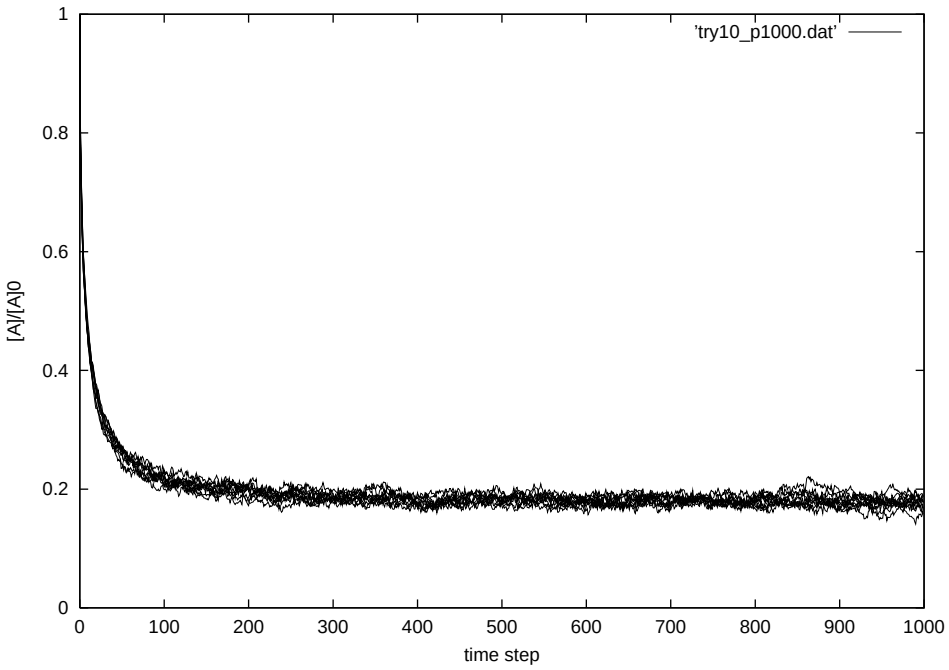


図 5.3: 結果例（$[A]_0 = [B]_0 = 1000$）

5.5.2 考察

微分方程式系の前提条件であるところの「反応に関与する物質が十分にたくさんある」とは、図 5.6 では 2000 個以上の Particle の部分が示す領域であると考えられる。この付近では微分方程式系と Particle System は同様の振舞いをするものと考えられ、またそれ以下になると離散化による揺ぎが顕著になると考えられる。

5.6 おわりに

本章では、反応に関与する Particle が少ない場合の Particle System のゆらぎについて実験した。

まず Particle の数を変化させながら実験したところ、ゆらぎは大数の法則に従い Particle の個数の平方根に比例する事が観測された。

また、反応半径を一定にしたまま仮想平面の大きさを変化させた場合のゆらぎについて計測したところ、十分に仮想平面が大きい場合、ゆらぎは上界を持つ事が示唆された。これは Particle System が溶液中での反応を想定した熱運動アルゴリズムを実装している

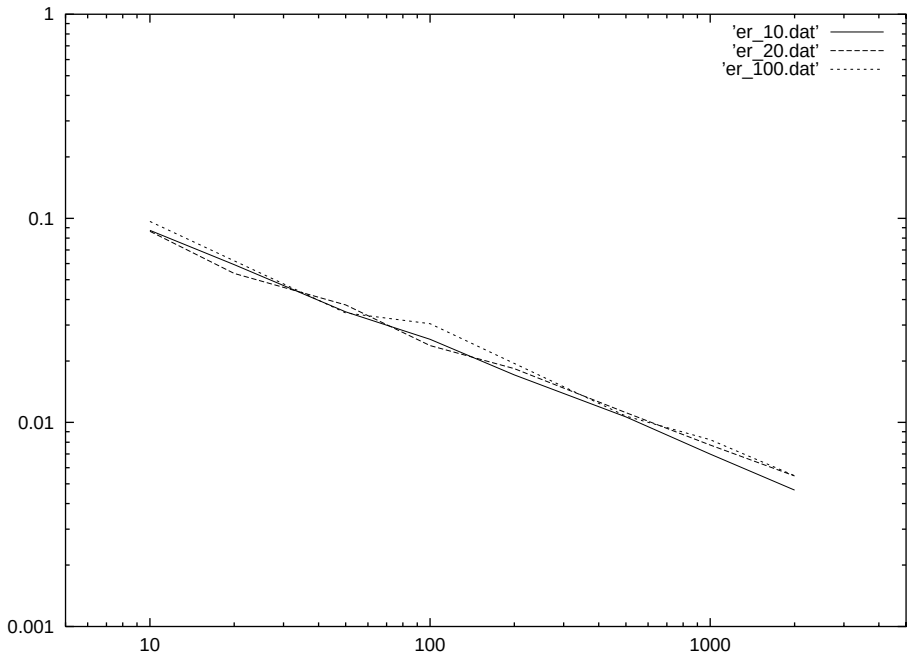


図 5.4: Particle の個数を変化させた場合のゆらぎ：10 回試行の平均をプロットしている。縦軸はゆらぎを、横軸は初期の Particle の個数を表している。Particle の個数 N を増やすにつれ、誤差は $\frac{1}{\sqrt{N}}$ となっている。

ために衝突回数が減る事が理由であると予想される。なお、本章では理化学研究所の吾妻竜三さんの研究を参考にさせていただいた。

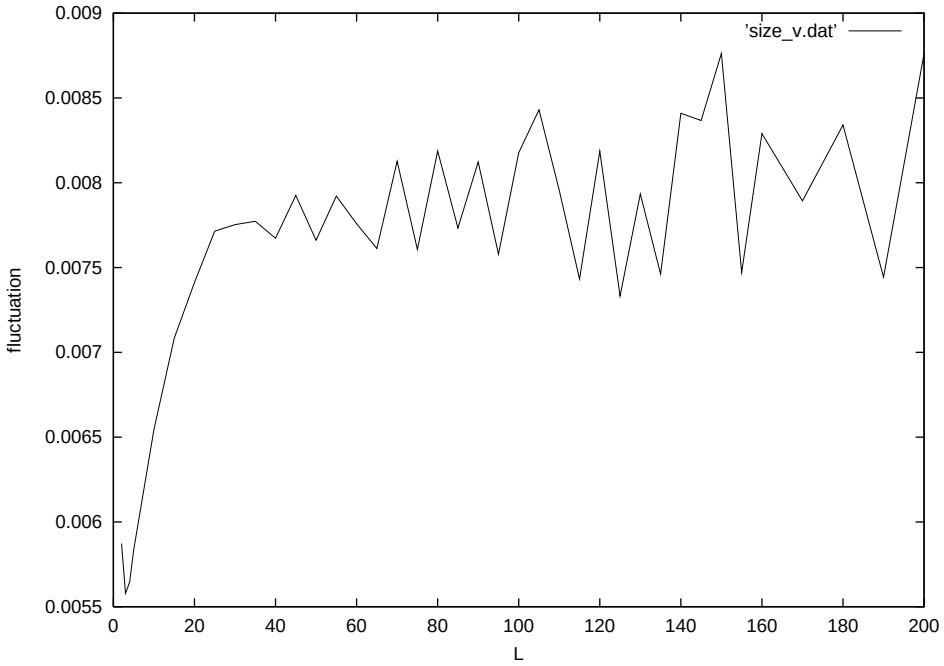


図 5.5: Particle の個数を一定にした時のゆらぎ：10 回試行の平均をプロットしている。縦軸はゆらぎを、横軸は正方形の仮想平面の一辺の長さ L を表している。L が 5 から 20 ぐらいまでの範囲において、ゆらぎは L に対して指数的に増加している。L が 20 より大きい場合はゆらぎの収束が悪いものの、一定の範囲内に収まっている。

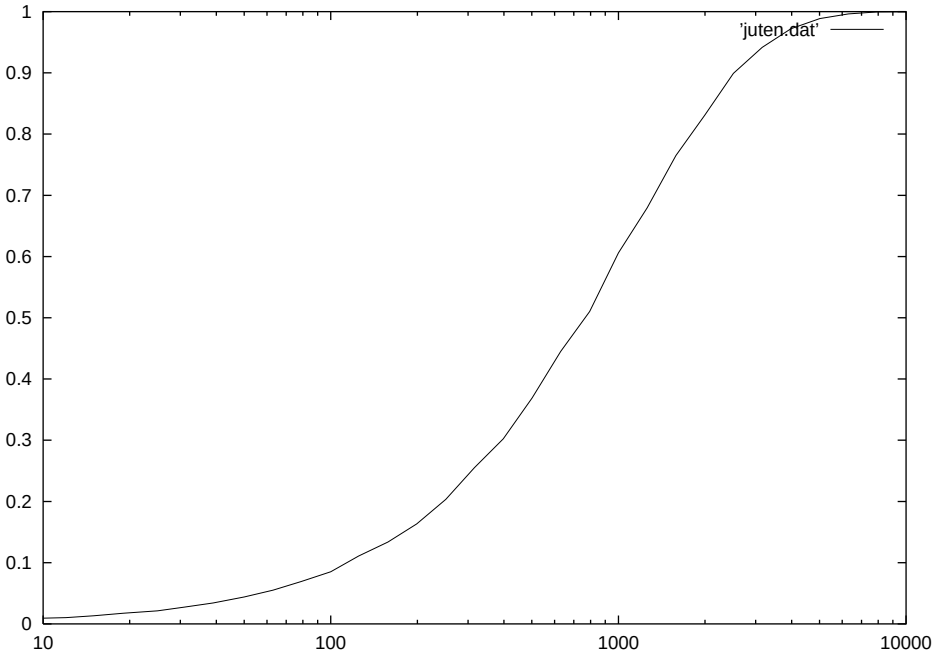


図 5.6: 充填率

第 6 章

結論

6.1 はじめに

本章では、本論文の結論について述べる。まず Particle System のアルゴリズムや実装についての結論を述べ、次に微分方程式系との比較について、最後に Particle System のゆらぎについて述べる。

6.2 Particle System

本論文では、以下の要求を満たす新しいシミュレーション手法 Particle System を提案した。

- ボトムアップなシステム
- 化学物質を濃度ではなく個々に表現するシステム
- 位置情報を取り扱えるシステム
- 定量的シミュレーションが可能なシステム

以上の要求を満たすために、システムは以下のような仕様を持つ。

- 化学物質を Particle （粒）として表現する。
- Particle は位置情報を持ち、仮想平面中に配置される。
- Particle は反応半径を持ち、反応半径内にある他の Particle のみを認識し、その情報のみをもとに化学反応を局所的に起こす。

6.3 本論文の結論

6.3.1 Particle System と微分方程式系の一致する点

現在生化学反応のモデルとして最も一般的な微分方程式系と同様に、本論文で提案したシミュレーション手法 Particle System においても以下の前提条件が保たれている場合における定量的なシミュレーションが可能である事を示した。

- 反応に関与する物質が十分にたくさんある事。
- 反応系は十分に撹拌されているものである事。

微分方程式系のモデルを Particle System のモデルに改変する事を考える。微分方程式系における反応速度定数から Particle System での反応確率を求める際、反応速度定数に重みづけをする定数 P を以下のアルゴリズムから求めた結果、微分方程式系と同様のシミュレーションが可能である事を示した。

1. Particle System の持つパラメータのうち、以下のパラメータを固定する。
   (a) 仮想平面の大きさ：S
   (b) 1mol に対応する Particle の個数：d
   (c) 1 time step の長さ：$\Delta t[sec]$
   (d) 反応半径：$r = 1$
2. 上記のパラメータを用いて、仮想平面を充填するモンテカルロ法を行い、d 個の Particle で実際に塗り潰せる仮想平面の領域を算定する。
3. S を実際に塗り潰せる領域の割合で割ったものを S' とし、$P = \frac{\Delta t S'}{\pi d}$ を算出する。
4. 3. で用いた P を微分方程式系モデルにおける各反応速度定数に乗じたものを反応確率として用いる。

6.3.2 Particle System と微分方程式系の一致しない点：ゆらぎの評価

Particle System は、反応に関与する物質が少ない場合に発生するゆらぎを評価した。

反応に関与する Particle の数とゆらぎの関係を調べたところ、大数の法則に則り、ゆらぎは Particle の個数の平方根に比例している事を示した。

また、反応に関与する Particle の仮想平面における密度が高い場合は、反応半径が大きくなるにつれ誤差が小さくなった。反応半径が大きい程 1 time step に参照する Particle の数が多くなる事が原因と考えられる。また反応確率が小さくなるにつれゆらぎが小さくなる事が観測され、この現象は反応確率が大きい程 1 time step で反応する Particle の数が多くなる事が原因と思われる。。

6.3.3 Particle System の意義

本論文で提案した Particle System は、生化学反応モデルのシミュレータとして主に使われている微分方程式系シミュレータのように定量的なシミュレーションと、人工生命や複雑系の分野で用いられているボトムアップ的で構成論的なシミュレーションの両方のアプローチが可能なシミュレータを目指した。計算コストは微分方程式系やセルオートマトンなど他のシミュレーション手法に比べると多少高くなるものの、定量的なシミュレーションがボトムアップ的な過程でなされていく様子は説得力を持つと思われる。

6.4 今後の課題

本研究の今後の課題について示す。

Particle System の実問題への適用　本研究はシミュレータの開発と基本的な評価に終始したが、実際に Particle System を用いて生化学反応シミュレーションを行う事により、シミュレータとしての真価を見極める必要がある。特に、複雑な化学反応ネットワークにおけるゆらぎの効果を評価することで、そのような系における新たな現象を発見する事が期待される。

アルゴリズム　アルゴリズムの面ではまだ足りない点が多数存在する。

- 二体反応以上の反応
- 酵素反応など特殊は反応
- 境界条件を考慮したパラメータ決定
- パラメータ決定の際に用いたモンテカルロ法の精度的な問題

高速化　Particle System のプログラムでは、主に以下の 3 点の計算コストが高かった。

- 距離の計算
- 乱数の処理
- ポインタ操作

以上の処理についての高速化を高める手法を調査、導入していく必要がある。

特に本研究における Particle System の実装では、プログラムの高速化のために反応半径の最大値を 1 に設定するなど、様々な制限を設けている。より使いやすいシミュレータにするためには、計算スピードは非常に重要な要素であるから、大きく改良する必要がある。例えば、粒子数の多い物質は Particle として再現するのではなく微分方程式系と同様に解き、粒子数の少ない物質に関しては Particle で再現するような、ハイブリッドなシステムが考えられる。与えられたモデルから自動的にゆらぎを算出し、連続量で計算するか Particle で計算するかを決定するようなアルゴリズムを作成できれば、可能かもしれない。

生物学者による評価　バイオインフォマティクスにおいて生物学者と情報科学者の協調が重要視されているように [4]、Particle System をアプリケーションとして生物学者が利用または評価できるように整備すべきである。この際、GUI などのユーザフレンドリーな側面を追加し、ネットワーク上の並列計算システムでの大規模シミュレーションに向けたオブジェクト化及びシステム化を進める事が重要点として挙げられる。

謝辞

まず最初に、私の研究活動に自由と的確な助言を与えて下さった指導教官小長谷明彦教授に、心から感謝いたします。

本研究をはじめる発端にあたり、明確な示唆とやる気を与えて下さった橋本敬教授に感謝いたします。研究にいきづまりが見えた時、可能性の道を開いて見せて下さった埼玉短期大学の仲隆助教授に感謝いたします。本研究の理論的な分析に関して、自らの研究成果を惜しみなく見せて下さった理化学研究所の吾妻竜三先生に感謝いたします。プログラムの実装や英語の指導を快くお受けいただいた小長谷研究室の Defago Xavier 助手に感謝いたします。懇意な指導と鋭い指摘を以て僕の研究を修士論文として成立させて下さった小長谷研究室の山本知幸助手に感謝いたします。

最後に、私の研究活動を内側から支えてくれた友人たちと、暖かい声をかけ続けてくれた家族に深く感謝いたします。

第 A 章

大数の法則

本章では本論文内で用いた大数の法則に関して述べる。まず正規分布と中心極限定理について述べ、チェビシェフの不等式から大数の法則を導く。

なお、本章は [36],[37],[38] を参考にした。

正規分布　正規分布は、ガウス（Gauss）によって発見された分布である。確率変数 X が正規分布にしたがうとき、その密度関数 $f(x)$ は以下の式で与えられる。

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad (-\infty < x < \infty)$$

この時、平均 $E[X] = 0$、分散 $V[X] = 1$ である。このような正規分布を $N(0.1)$ と記す。

また、$N(\mu, \sigma^2)$ にしたがう密度関数 $g(x)$ は以下のように表される。

$$g(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

中心極限定理　$X_i (i = 1, 2, ..., n)$ が互いに独立かつ同一の確率分布にしたがう確率変数列とし、$S_n = X_1 + X_2 + ... + X_n$ とする。また $E(X_i) = \mu, V(X_i) + \sigma^2 (< \infty)$ とするとき、以下の式が成り立つ。この定理を中心極限定理と呼ぶ。

$$P\left(\frac{S_n - n\mu}{\sqrt{n}\sigma} \leq x\right) \to \Phi(x)$$

二項分布に限らず、$X_i (i = 1, 2, ..., n)$ が互いに独立で同一の確率分布に従う場合、S_n の分布は常に正規分布に近づく。

チェビシェフの不等式　任意の確率密度関数の標準偏差 σ は、分布の幅または広がりの尺度である。すなわち、平均の近傍における確率密度の尺度である。σ が小さいときは、平均に近い値を得るかくりつが高く、σ が大きいときは、平均から離れた値を得る確率が σ に比例してより高くなる。この性質を表した式をチェビシェフの不等式と予備、平均からある限界を考えるとき任意密度関数のなす面積に上界を与える。したがって、確率分布が平均 $\bar{(}x)$ および標準偏差 σ をもてば、kx 標準偏差だけ平均からはずれた値を有する確率は $\frac{1}{k^2}$ 以下となる。以下に式を示す。

$$P\left(|x-\bar{x}| \geq k\sigma\right) \leq \frac{1}{k^2}$$

簡単に示すと、以下のようになる。平均を $\bar{(}x)$、標準偏差 σ とすると、$|x-\bar{x}| \leq k\sigma (k \geq 1)$ の範囲に含まれる変量 x の度数の全体に対する割合は $1-\frac{1}{k^2}$ 以上である。

大数の法則　$X_i (i=1,2,...,n)$ が互いに独立かつ同一の確率分布にしたがう確率変数列とし、$S_n = X_1 + X_2 + ... + X_n$ とする。この時 S_n の算術平均 $\bar{S_n}$ に関して $E(\bar{S_n}) = \mu, V(\bar{S_n}) = \frac{\sigma^2}{n}$ となる。この時標準偏差は $\frac{\sigma}{\sqrt{n}}$ となる。

この時、$k \gg \sqrt{n}$ であれば、チェビシェフの不等式より $\left|\bar{S_n} - \mu\right| \simeq 0$ となる。このことから、大きさ n の標本変量についての標本平均 $\bar{S_n}$ は、集める個数を大きくすれば、母集団の平均値にいくらでも近い値を取るようになる。この法則を大数の法則という。この時、表平均 $\bar{S_n}$ は平均 μ、標準偏差 $\frac{\sigma}{\sqrt{n}}$ にしたがう正規分布となる。

参考文献

[1] http://www.jsbi.org/

[2] http://bird.jst.go.jp/

[3] http://www.genome.ad.jp/

[4] http://www.gsc.riken.go.jp/

[5] 金子邦彦 複雑系のバイオフィジックス共立出版株式会社, 2001.

[6] 吉川研一, 非線型科学—分子集合体のリズムとかたち—, 学会出版センター, 1995.

[7] C. Langton, Artificial Life, Addison-Wesley, 1989.

[8] K. Sims, Evolving 3D Morphology and Behavior by Competition, R. Brooks and P. Maes (eds), Artificial Life IV, MIT Press, 28-39p, 1994.

[9] T.S.Ray, An Approach on the Synthesis of Life, Artificial Life II, Langton, C.G. et al. (eds.), Addison Wesley, 1991.

[10] J.J. tyson, R.J.Field, M.Burger, Oscillations and Travelling Waves in Chemical Systems John Wiley & Sons, NewYork, pp95, 1985.

[11] J. von Newmann, Theory of self-reproducing automata, edited and completed by A. Burks, University of Illinois Press, Champaign, IL, 1996.

[12] http://www.sfc.keio.ac.jp/ n96299sh/

[13] 服部桂, 人工生命の世界, オーム社, 1994.

[14] 加藤恭義, 光成友孝, 築山洋, セルオートマトン — 複雑系の自己組織化と超並列処理 — 秀英出版, 1998.

[15] 潮 俊光, ハイブリッドペトリネット, 計測と制御, 38, 169-175p, 1999.

[16] Steven Levy 著 服部桂 訳 人工生命 — デジタル生物の創造者たち — 朝日新聞社, 1996.

[17] 上田完次, 下原勝憲, 伊庭斉志, 人工生命の方法 — そのパラダイムと最前線 —, 工業調査会, 1995.

[18] 渡辺治, 米崎直樹, 計算論入門, 日本評論者, 1997.

[19] 小長谷 明彦, 遺伝子とコンピュータ, 共立出版, 2000.

[20] 北野 宏明, システムバイオロジーの展開 — 生物学の新しいアプローチ—, シュプリンガーフェアラーク東京, 2001.

[21] http://www.tigr.org/

[22] 富田勝, E-CELL：ゲノム情報に基づく細胞の再構築, 数理科学, 47-53p, NO.432, JUNE, 1999

[23] Bruce Alberts, Dennis Bray, Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts, Peter Walter, 邦訳 中村桂子, 藤山秋左夫, 松原謙一, Essencial Cell biology（日本語版）, 南江堂, 1999.

[24] Tomita, M., Hashimoto, K., Takahasi, K., Shimizu, T., Matsuzaki, Y., Miyoshi, F., Saito, K., Tanida, S., Yugi, K., Venter, J.C., Hutchison, C. E-CELL：Software environment for whole cell simulatioin Bioinformatics, Vol.15, No.1.

[25] A.M.Turing, The Chemical Basis of Morphogenesis Phil. Trans. R. Soc. Lond. B 237, 37-72(1952).

[26] H. Meinhardt, The Algorithmic Beauty of Sea Shells, 2nd enlarged edition, Springer, 1997.

[27] 近藤滋, キリンの斑論争と現代の分子発生学科学821-825p, Vol.67, No.11, Nov., 1997

[28] 近藤滋数理科学10-14p, NO.418, APRIL, 1998

[29] Shugo hamashita, Hiroaki Kitano, Simulation of Drosophila Embryogenesis, Artirficial life VI : proceedings of the sixth international conference on Artificial Life, p151-160, 1998.

[30] Hiroshi Matsuno, Atsushi Doi, Rainer Drath, Satoru Miyano, Genomic Object Net: Object Oriented Rpresentation of Biological Systems Genome Informatics No.11, 229-230p, 2000

[31] 村田忠男, ペトリネットの解析と応用近代科学社, 1998

[32] http://gepasi.dbs.aber.ac.uk/softw/gepasi.html

[33] http://www.e-cell.org/

[34] W.H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vetterling, Numerical Recipes in C （日本語版）訳 丹慶 勝市, 奥村 晴彦, 佐藤 俊朗, 小林 誠技術評論社, 1998

[35] 都築 卓司, なっとくする統計力学, 講談社, 1998.

[36] 戸川隼人, 計算機のための誤差解析の基礎, サイエンス社, 1989.

[37] 東京大学教養学部統計学教室 編, 自然科学の統計学, 東京大学出版会, 1998.

[38] 佐藤平八, 確率、統計、ランダム過程, 森北出版, 1990.