

研究課題別事後評価結果

1. 研究課題名: システムバイオロジーのためのモデリング・シミュレーション環境の構築

2. 研究代表者名及び主たる研究参加者名(研究機関名・職名は研究参加期間終了時点):

研究代表者

富田 勝 (慶応義塾大学先端生命科学研究所 所長/環境情報学部 教授)

3. 研究実施概要

生物や細胞をシステムとしてとらえてその全体のしくみや振る舞いを理解する新しい学問分野は「システムバイオロジー」と呼ばれ、コンピュータシミュレーションが不可欠である。大規模化する細胞モデリングに対応した新しい統合環境ソフトウェアの開発を主な目的として、細胞シミュレーション統合環境(E-Cell IDE)の開発と、細胞シミュレーションのための要素技術の研究開発(計算アルゴリズムの基礎的研究、E-Cell システムの改良)を実施した。5年の研究期間を終えて主要な目的であった統合環境ソフトウェア、E-Cell IDE の開発をほぼ完了し安定版をリリースした。

(1)細胞モデリング・シミュレーション統合開発環境: E-Cell IDE の開発

E-Cell Integrated Development Environment (E-Cell IDE) は Microsoft Windows 上で動作する GUI アプリケーションであり、E-Cell3 をシミュレーションエンジンとして、モデリング環境および総合的なシミュレーション・解析環境をユーザに提供するためのソフトウェアプラットフォームである。

これによりモジュールの開発などのカスタマイズが可能な自由度の高い拡張性を持ち合わせつつ、効率の良いモデリングとシミュレーションが可能となり、その結果シミュレーション研究者だけではなく生化学・分子生物学の実験ラボ研究者のような一般的な生物学者にとっても幅広く使えるソフトウェアが完成するに至った。E-Cell IDE の主要な機能は大きく分けて以下の 4 つである。

- ・プロジェクト・モデル管理機能: モデルのリビジョン管理、シミュレーション条件・結果の保存、モデル構造の操作、SBML モデルのインポート・エクスポート
- ・パスウェイエディタ: モデルをパスウェイ・ダイアグラムとして構築、シミュレーション中の変化のアニメーション表示
- ・解析・デバッグ機能: システムバイオロジー研究でよく使われる解析ツールの提供(ロバスト解析、感度解析、分岐解析、パラメータ推定)、モデリング途中での不備・バグ等を自動的に指摘するデバッグ
- ・グリッドコンピューティングサポート: SGE(Sun Grid Engine)で構成されたバックエンドの計算クラスタに対して Globus Toolkit 経由で解析ジョブを送信、受信する機能

(2)細胞シミュレーションのための要素技術に関する研究

(a)シミュレーションアルゴリズムモジュール群の開発:

これまでに開発したマルチアルゴリズム手法に対応する、細胞シミュレーションに求められる確率論アルゴリズムとして、Gillespie、Tauleap、Langevin アルゴリズムの改良・実装を行い、これらのアルゴリズムをモデルの性質に応じて適切なものに自動的に切り替える機構を開発した。これにより、ユーザは本システムを用いることによって、最先端のシミュレーションアルゴリズムをプログラミングなしに自在に使えるようになった。

(b)ルールベースモデリング手法への対応:

組み合わせ論的な爆発が発生しうる複雑な複合体形成を含む生化学反応経路のモデリングとシミュレーションを行う上で必要となる、ルールベースのモデリング手法を E-Cell のオブジェクトモデルに適合する形で行えるようにするために必要な機構を米 The Molecular Sciences Institute (tMSI)の開発した Molecuizer をベースに研究し、E-Cell3 を拡張する形で実装した。このアルゴリズムを活用することにより、複合体形成などのイベ

ントをトリガーとして動的にパスウェイ経路を生成することで組み合わせ論的爆発により起こる問題を最小化することができるようになった。

(c) 共有メモリ型計算機に対応したマクロレベルでの物質拡散アルゴリズムの開発:

細胞マクロレベルでのタンパク質等の物質局在や拡散を取り扱うために、ガスキネティクス領域で用いられる格子気体法を発展させ3次元格子にも対応したアルゴリズムを開発した。これまでに、大規模なシミュレーションモデルを複数の区画に分割して並列計算を行わせる手法として共有メモリ型(SMP 型)のプロトタイプ開発と、PC用の汎用グラフィックスプロセッサ(GPU)を用いた反応拡散シミュレーションの高速化を行い、E-Cell3 コアのコードベースへのフィードバックを試みている。

(d) E-Cell システムの改良:

本研究課題で開発している統合環境ソフトウェア(E-Cell IDE)のシミュレーションエンジンや数値計算のためのコアアルゴリズムの実装については、旧バージョンの E-Cell3 を継続的に使用しているが、E-Cell IDE の開発過程において明らかとなった不具合等の修正やリファクタリングを継続的に実施しており、マイナーバージョンアップ版(version 3.2)としてリリースを行っている。E-Cell3 上で構築されたモデルによる研究は研究期間の全体にわたって実施されており、具体的な研究成果として、Toll 様受容体(TLR3/4)のモデルを用いたSFR(Signaling Flux Redistribution)の解析などが行われた。

4. 事後評価結果

4-1. 研究の達成状況及び得られた研究成果(論文・口頭発表等の外部発表、特許の取得状況等を含む)

システム生物学のためのモデリングおよびシミュレーションを自在に実行するためのシステムを構築し、さらに細胞シミュレーションに関する多くの要素技術を開発した。当初の計画にあった多数の機能の開発を達成し、十分に実用的なプロダクトとしての品質を達成したものをリリースしている。これはその拡張性や細胞シミュレーションに特化した計算アルゴリズム等の点で、世界的に見ても優位性の高い技術であり、今後のこの統合環境ソフトウェアを活用したシステムバイオロジー研究が進められると期待できる。知識ベース駆動型モデリング環境の構築については、計画の変更があったが、全体として計画はほぼ達成された。

E-Cell は ACT-JST による支援を受けて開発された時からの歴史がある。研究開始当初はシステムバイオロジーの研究が緒についたばかりであったが、今回の CREST プログラムによりシミュレーション統合環境としての E-Cell IDE がかなりレベルの高い状態でリリースされ、システムバイオロジーにおける重要なソフトウェアとしての位置づけを確立したといえる。

当初計画では想定していなかった新たな展開として米国 MSI の Molecuizer の導入がある。生物関連の複雑なモデリングにおける組み合わせの爆発がこれにより回避され、妥当なモデリングが容易になった。また、これを E-Cell に組み込む段階で、E-Cell 自身のブラッシュアップも行っている。そして米国で開発されたある標準に則ってマークアップ言語を開発したので国際的にも親和性のある言語になった。この他、本体側の研究遂行と並列して行われた試みとして、細胞シミュレーションの3次元可視化の研究を行い学会、メディア等でも注目を集めた。

ソフトウェアシステムの開発を行いつつ成果を論文にまとめるという作業は容易ではない。そのことを考えると本プロジェクトに関連する国際論文、国際学会での招待講演、国際学会でのポスター発表と世界に向けて大いに研究成果を発信している。E-Cell を中心とした発表によって、システムバイオロジーの研究分野の国際的なイニシアティブを継続的にとっており注目度も高く十分に評価できる。

E-Cell IDE を GPL ライセンスによって、ソースコードを完全に公開している。この分野において公開されているソフトウェアは数多くあるが、E-Cell IDE のように多機能で、かつ完成度の高いソフトウェアで、GPL ライセンスを採用している所はほとんど存在しない。従ってソフトウェアを公共財として担保するための手段として GPL ライセンスでの公開には他には無い高い優位性がある。

ソフトウェア開発の管理進捗から実験に至るまで多種多様の研究者を統括しなければならない研究テーマで

あり、研究チーム内の研究開発が適正に推進され十分なリーダーシップが発揮されたと判断できる。特に一般的に見ると難しい医工連携が円滑に実施されており、また大規模なソフト開発とメンテナンスについても頻繁な担当者間のコミュニケーションを行いスムーズな開発が行われている。研究費は主に統合環境ソフトウェアの開発費に当てている。このように公開して利用可能なレベルの高いソフトを開発するには不可欠であり、その結果、実用的で品質の高いGUIソフトウェアも開発しており予算執行は適切に行われた。

4-2. 研究成果の科学技術や社会へのインパクト、戦略目標への貢献

生体システムのシミュレーションソフトの中でも、**E-Cell IDE** はシステムバイオロジー研究に必要と考えられる多数の機能を盛り込みつつ、実用性の観点から高いレベルでの品質を提供するなど機能的に優れており使いやすい環境でリリースされた。特に先端的な細胞シミュレーションに関する要素技術の研究においては、決定論的モデルと、確率論的なモデルとが併存する細胞シミュレーションに親和性が高いアルゴリズム群を開発した。これに加え、今後のシステムバイオロジー研究において重要と考えられる、細胞内部でのタンパク質の拡散現象を物理シミュレーションの手法を活用して計算するための手法を、共有メモリ型の並列計算機と **GPGPU** に対応させる形で開発した。今後は、今回開発した技術をベースとした実用的なモデリング対象への展開が大いに期待できる。

E-Cell IDE を使うことによって、細胞シミュレーションに特化した計算アルゴリズムを情報技術分野の研究者だけではなく、生物系の大学院生のレベルまで幅広く、個人が持つPC環境において簡単にインストールして使うことができるようになった。その結果、すでに先端的な生物機能についての研究への適用もなされ優れた成果がいくつか挙がっている。

また、プラグイン機構を開発したことにより、高い拡張性を提供されており、研究プロジェクト毎に必要な新規機能の追加も比較的容易に行うことができる。これに加え、**E-Cell IDE** を構成する重要なソフトウェアコンポーネントは **GPL** を中心に選ばれており、**E-Cell IDE** も **GPL** ライセンスに基づきソースコードが完全に公開されている。

これらの点を鑑みると、本研究課題の成果として、世界的に見ても優位性が高いと言え、今後のこの統合環境ソフトウェアを活用したシステムバイオロジー研究がますます進められると期待できる。

システムバイオロジーという分野は新しく開拓された分野であり、今後、収穫が期待できる分野である。当初は小規模なシステム、モデルに関する理論的な研究が中心であったが、実用的なレベルでの統合環境ソフトウェアの構築がなされ、従来のソフトウェアでは実現できなかったウェットラボの生物学者であってもシミュレーションモデルを取り扱える高い操作性と拡張性を実現している。その結果、システムバイオロジーの研究を行うユーザーとの共同研究にも発展しつつある。このようにこの実用的な統合環境ソフトウェアをモデリング・シミュレーションの基盤として、ますます大規模化し、データドリブン化するシステムバイオロジー研究において、ウェット、ドライの研究者の融合研究が進展することが大いに期待される。本プロジェクトは「医療・情報産業における原子・分子レベルの現象に基づく精密製品設計・高度治療実現のための次世代統合シミュレーション技術の確立」という戦略目標にかなった計画であり、成果の科学的・社会的インパクトは大きい。具体的な例としては、シミュレーションとメタボローム解析による輸血用血液の保存法の開発があげられる。このことによって血液保存法開発のターゲットとすべき酵素反応段階や代謝への影響因子を予測することができ輸血用血液の使用可能期間を延ばすことが可能となると期待されている。

4-3. 総合的評価

システムバイオロジーのためのモデリング・シミュレーション統合環境ソフトウェアの開発を行い、システムバイオロジー研究に必要と考えられる多数の機能を盛り込みつつ実用性の観点からも高いレベルでの環境を提供したことを高く評価する。この成果については当研究領域の趣旨にてらして見ても、革新的なシミュレーション技術を実用的な統合環境アプリケーションとして広く研究者に提供するものであり、十分な成果が得られていると考えられる。

またシミュレーションモデルを構築・実証するためにはメタボローム解析が不可欠であるが、すでに細胞シミュ

レーションとメタボローム解析を組み合わせた取り組みを始めており今後の研究の発展が期待される。