

3. 研究開発課題名 グローバルコンピューティング環境による汎用MCSCFソルバーの開発

3.1 代表研究者 日本電気株式会社 基礎研究所 主席研究員 高田 俊和

3.2 概要

スーパーコンピュータの性能向上に伴い、非経験的分子軌道法を軸とした分子シミュレーションがマテリアルサイエンスの分野において標準的な研究手段になりつつある。特にナノテクノロジーやバイオインフォマティクスの領域での中核的シミュレーション技術となると期待されている。このような状況の下、本プロジェクトでは、1) 化学反応のメカニズム解明に不可欠なMCSCF (多配置SCF) 計算プログラム、2) 多地点から同時に複数の研究者が計算結果を解析できるインターネット対応GUI、の自主開発を中心的課題として研究開発を行った。

3.3 研究開発実施内容

本プロジェクトでは、異機種並列分散処理に対応できる MCSCF 計算プログラムの自主開発を軸に、研究開発を行ってきた。研究題目としては、

- 1) 汎用MCSCFソルバー及びインターネットGUIの開発
- 2) 分散処理のための基本アルゴリズム開発
- 3) Internet上での性能評価

を設定し、NEC 基礎研究所、NEC 情報システムズ、NEC ソフトが実際のプログラム開発を担当した。協力研究員として、産総研、九工大、東海大から参加を頂き、GRID など異機種並列分散処理の技術動向や並列化プログラムの基本設計などに助言を頂くと同時に、開発したシステムの他地点からの動作確認に協力して頂いた。

MCSCF 法は、化学反応過程で起こる色々な電子配置の混合を正しく取り入れることのできる理論として確立されており、Gaussian、GAMESS などその計算機能を備えたプログラムは幾つか存在する。しかしながら、エネルギー表式に現れる分子軌道基底での2電子積分の生成が大変困難であるため、最新のスーパーコンピュータを以ってしても精々数10原子からなる比較的小さな分子系への適応に限られていた。本プロジェクトでは、特にPCクラスタなど分散メモリ型の並列計算機に着目し、1000原子軌道基底以上の計算を可能にする、MCSCFプログラムの開発を主たる目標とした。MCSCF法のプログラムは、1) 波動関数と全エネルギー計算部、2) 原子核に働く力を求めるエネルギー勾配計算部、からなるが、これらについて全て自主開発を行った。新規性は、1) 2電子積分の原子軌道基底から分子軌道基底への並列変換アルゴリズムの開発と、2) 2次の密度行列の分子軌道基底から原子軌道基底への変換によるエネルギー勾配計算の並列化、である。これらの実現により、当初の目的であった1000原子軌道の計算は、クロロフィル2量体(166原子、684電子、994軌道)で既に達成されている。要した計算時間は、33プロセッサのPCクラスタで10日である。世界記録を目指す大型計算では、その時々スーパーコンピュータを擁して数ヶ月に及ぶのが通例であり、本計算では極めて短時間に終了している。MCSCF法を含むポストHartree-Fock計算において、この規模の計算は報告されておらず、本システムの汎用性を示すものと自負している。本アルゴリズムでは、原子軌道の2乗で並列化できるので、1000軌道で100万台のプロセッサを備えた並列計算機に対応で

きることになる。現在、2000軌道を越える計算を計画中であり、今後のPCクラスタの進歩により更に大型分子のMCSCF計算が可能になると期待している。本システムを、現在生体高分子シミュレーションの分野で注目されているQM/MM (Quantum Mechanics/ Molecular Mechanics)と融合すれば、溶液中での化学反応について100原子からなるの系のシミュレーションが可能になると考えている。

インターネット対応GUIの開発についても、当初想定したインターネットを介しての多地点での計算結果の解析機能が実現されている。インターネットの著しい普及により、今後あらゆる分野のシミュレーションは、ブラウザで入力データを作成しサーバに送り、結果もブラウザ上に表示されるようになると考えられる。本GUIは、そのような環境を先取りして、将来のシミュレーションイメージを提示している。工夫した点は、多地点に分散している研究者が同一の画像を見ることを保証したことと、議論の対象を明確にするためにマーカーと呼ばれる指示棒を表示する機能を実現したことである。また、音声信号を送ることで、会話による検討を可能にしている。

MCSCF数値計算部とGUIは統合され、現在ひとつのシステムとして稼動している。これまでの述べてきたように、本プロジェクト開始時に想定した計算機能は基本的に全て達成されていると考えている。

3.4 題目別実施内容

(1) 汎用MCSCFソルバー及びインターネットGUIの開発 (高田)

NEC基礎研究所の高田、津田、NISの佐久間、平原、山本、岡本、NESの村瀬、中田と下記の開発項目について研究開発を実施した。

(a) MCSCFプログラムの基本設計及びCI行列要素計算プログラム開発 (高田、中田)

Symmetric Group Approach (SGA) 法により、MCSCF計算手法として多用されているCASSCF (Complete Active Space SCF)でのCI行列と一般化Fock行列の生成プログラムの基本設計とCI行列計算プログラムの開発を行った。

(b) MCSCF法における一般化Fock行列計算プログラムの開発 (中田)

2電子積分のダイレクト法による、一般化Fock行列生成プログラムの開発を行った。このプログラム開発においては、分子軌道の係数を求めるための繰り返し計算が必要であるが、その中間データを効率よく分散メモリ上に保存するように工夫し、計算時間の削減を実現した。

(c) ダイレクト2電子積分計算プログラムの開発 (村瀬)

2電子積分の原子軌道基底から分子軌道基底への変換を主記憶内で行うため、2電子積分の計算を特別な順序で計算できるように、積分計算プログラムを開発した。併せて、積分の事前評価による削除も行えるよう改良した。

(d) 分子軌道基底での2電子積分生成プログラムの開発 (高田、中田)

上記2電子積分計算プログラムを用いて、2電子積分の原子軌道基底から分子軌道基底への変換をローカルメモリー上で行うプログラムを開発した。

(e) MCSCF波動関数計算部の並列化性能向上とベンチマーク (中田)

PCクラスタによるベンチマークを行い、上記並列化アルゴリズムが有効であることを確認した。また、ATLASによるチューニングを行い、計算時間を約3分の1に短縮することができた。

(f) エネルギー勾配法のエネルギー表式計算プログラムの開発（高田、佐久間）

CASSCFで与えられるエネルギー表式を用い、原子核に働く力を計算するプログラムの基本設計を行った。

(g) エネルギー勾配法におけるカップリング係数行列生成プログラムの開発（佐久間）

CASSCF法により得られる1次及び2次の密度行列と2電子積分の原子核による微分との積和をとり、エネルギー勾配を計算するプログラムを開発した。

(h) エネルギー勾配計算部の並列化とベンチマーク（佐久間）

2次の密度行列の分子軌道基底から原子軌道基底への逆変換アルゴリズムを開発し、2電子積分の原子核による微分の計算との積和計算を効率よく並列化できたが、ベンチマークは今後の課題として残されている。

(i) 数値計算部の並列分散処理に伴うコア管理ユーティリティの開発（佐久間、中田、平原）

本システムはPC クラスタ等のローカルメモリーを搭載したコンピュータ上での稼働を想定しており、これらのプロセッサでの作業領域の確保は全て **Dynamic Core Allocation** をベースに行うよう設計した。そのユーティリティ関数を開発した。

(j) 数値計算部のパーツ化を図るためのデータセットの統一化（津田、中田、平原、山本）

大型プログラムシステムの開発において、最も困難なのはプログラム間のインターフェースの調整である。特に、引数の不整合によるバグが多数発生する。そこで、今回の開発では共通データを全てファイル経由で渡す方式を採用した。これにより、プログラム開発が容易になることが実証できた。それに必要なユーティリティ関数を開発した。

(k) 分子グラフィックス画像データ送信用インターネットツールの開発（岡本）

本システムではインターネット環境下での計算を想定しているため、送受信するデータ量の削減が必要である。そのための画像圧縮技術とソケット通信によるプロセス間通信を実現した。

(l) データ作成及び解析用インターネットツールの開発（津田、山本、岡本）

データ作成と解析用のインターネットブラウザを開発した。特徴のひとつは、多地点の研究者が同時に計算結果の解析を行えるように、同一の画面を送信する機能、表示画面にマーカーを共通表示する機能、音声通信機能を実現したことである。

(m) インターネット対応GUIの評価（山本、岡本）

協力研究員の協力を得て、本システムの多地点における動作確認を行った。

(2) 分散処理のための基本アルゴリズム開発（高田）

GRID など並列分散処理の将来動向や並列化の基本的考え方について、産業技術総合研究所の関口主任研究官と建部研究員から助言して頂いた。これらの助言をベースに、将来の異機種並列分散処理に対応できるように、プログラムの基本設計を行っている。また、数値計算部の並列化については、自主技術として以下の並列アルゴリズムの開発を行った。

(a) 2電子積分並列変換アルゴリズムの開発（高田、中田）

2電子積分の原子軌道基底から分子軌道基底への変換を分散メモリ上で並列化するアルゴリズムを開発し、波動関数の数値計算部にインストールした。

(b) 2次の密度行列の逆変換アルゴリズム開発（高田、佐久間）

2次の密度行列を分子軌道基底から原子軌道基底へ逆変換することで、2電子積分の原子

核の座標による微分との積和を分散メモリ上で並列処理するアルゴリズムを開発した。

(3) IMnet 上での性能評価 (高田)

協力研究員の協力を得て、本システムの多地点における動作確認を行い、当初想定した機能が実現できていることを確認した。

3.5 全体の総括と今後の課題

本プロジェクトでは、化学反応の計算に不可欠な MCSCF 法に着目し、そのプログラムの自主開発を中心課題として取り組んできた。MCSCF 法のプログラムは、既に Gaussian、GAMESS、MOLPRO など外国において数システム開発されているが、計算対象分子の大きさとして数 10 原子程度に限られていた。MCSCF 法におけるエネルギー表式が分子軌道基底での各種積分により表現されるため、膨大な数の 2 電子積分を原子軌道基底から分子軌道基底へ変換しなければならない。この変換がネックとなって、最新のスーパーコンピュータを以ってしても精々前述の規模の計算が行われているにすぎない状況であった。その結果として、MCSCF 法の適応範囲が限定され、化学反応機構の理論的解明が進まないという問題を提起していた。原子軌道の数の 4 乗に比例する 2 電子積分を、5 乗に比例する演算をかけて変換しなければならないとからである。このような状況から、本プロジェクトでは、1000 原子軌道基底を越える計算の実現を目指して、並列化アルゴリズムの開発を軸に高度なコンピュータ利用技術の構築を行ってきた。その結果、1) ポスト Hartree-Fock 法の大型分子への適応を妨げていた 2 電子積分の原子軌道基底から分子軌道基底への変換を効率的に行う並列化アルゴリズムと、2) 分子構造を最適化するために必要なエネルギー勾配計算において、2 電子積分の原子核座標による微分の計算を効率的に行う 2 次の密度行列の分子軌道基底から原子軌道基底に逆変換する並列化アルゴリズムを開発した。上記のプログラムで多用されている変換アルゴリズムの改良で効率的な並列化が実現できたのであるが、その発見の学問的意味は極めて大きいと自負している。このアルゴリズムをインストールして、クロロフィル 2 量体 (166 原子、684 電子、994 軌道) の計算を 33 プロセッサの PC クラスタで既に実現した。計算に要した日数は、10 日である。知る限りにおいて、ポスト Hartree-Fock 法で、この規模の計算は報告されていない。現在、2000 原子軌道を越える計算を計画中であり、積分変換の並列化アルゴリズムの開発により、MCSCF 法は本格的な応用計算の時代に入ると確信している。

本プロジェクトにおけるもうひとつの開発課題は、インターネットを介して MCSCF 計算を行うためのステアリングシステムの開発である。本事業の中心課題にネットワークの活用がうたわれていたので、多地点の研究者が計算結果についてインターネットを介して解析できるシステムの構築を試みた。これは、異なる部署に属する研究者が計算結果を見ながら議論することを可能にするシステムであり、どの分野であれ次世代のシミュレーションがインターネット環境下で行われることを考えると、世界に先駆けてそのシミュレーションシステムのイメージを提案するものである。また、IT のビジネス形態として ASP (Application Service Provider) が期待されているが、計算科学用ビジネスのブラウザとして使用することも想定している。これが更に進めば、学校教育における化学現象の体験的学習ツールとしても、利用される可能性もある。

結論として、MCSCF 法の数値計算部とインターネット対応 GUI については、当初の目標を上回る機能を実現できたと考えている。

今後の課題であるが、ふたつあると考えている。そのひとつは、プログラムの更なるチューニングである。チューニングの有効性を示す例をあげると、前述のクロロフィル2量体にはPCクラスタ（33CPU）を用い、MCSCFの1サイクルを処理するのに最初のバージョンでは数日を要していた。ATLAS (Automatically Tuned Linear Algebra Software)によるチューニングした結果、僅か1日に短縮された。特に、PCクラスタの場合、キャッシュのミスヒットをなくす最適化が、極めて有効であることが判明した。MCSCF 数値計算部の内、波動関数計算部は先に稼動し始めた分だけチューニングが施されているが、エネルギー勾配計算部については波動関数計算部の完成を待たないとデバッグできないという事情から、チューニングなどプログラムの整備が遅れている面は否定できない。今後数ヶ月をかけて、エネルギー勾配計算部のチューニングを行いたいと考えている。この作業が完了すれば、128CPUのPCクラスタでクロロフィル2量体程度の分子の構造最適化が1ヶ月以内に計算できると予想している。今回開発したプログラムは原子軌道の数の二乗で並列化できるので、クロロフィル2量体で100万プロセッサのシステムに対応できる。PCクラスタの進歩により一層の高速化は問題なく実現できると考えている。

もうひとつの課題は、プログラムの機能強化である。本プロジェクトで開発した数値計算部は、波動関数と原子核に働く力の計算を計算する機能のみである。量子化学計算の最大の利点は、波動関数が一旦求まれば、その波動関数を用いて静電ポテンシャルや双極子能率など色々な物理量を計算できることである。しかしながら、これらの計算機能を実現するには、相当のプログラム開発を必要とする。これまで非経験的分子軌道計算システムの自主開発を進めてきた経験に基づく反省として、1研究グループが全ての計算機能を担当するのは限られた人的資源では不可能であり、多くの研究者が自分のアイデアで計算機能を自由に付加できるような環境を整備することが不可欠と認識している。そのためには、公的研究機関がそのようなプラットフォームを提供することが強く望まれる。今から20年程前、アメリカにNRCC (National Resources for Computations in Chemistry) なる研究機関が、NSFの資金で設立された。その狙いは、計算化学の分野のアプリケーションプログラムを大学等との共同研究で開発しユーザに提供すると同時に、ハードウェアも整備することで簡便にユーザが計算できる環境をも提供することであった。NRCCはNSFの資金的な問題から数年で閉鎖となってしまったが、今こそそのような機関を日本国内にも設立すべきではないかと痛感している。今回の計算科学プロジェクトで開発された多くのプログラムを更に発展させることが今後の大きな課題であると考えているが、そのためのインフラストラクチャとして、JST主導によりNRCCのような研究機関の設置を検討して頂きたいと願う次第である。

3.6 研究開発実施体制

代表研究者氏名 高田 俊和

所属・役職 NEC 基礎研究所・主席研究員

(1) 汎用MCSCFソルバー及びインターネットGUIの開発

A. 参加研究者

高田俊和 NEC基礎研究所 主席研究員

MCSCFプログラムの基本設計及びCI行列要素計算プログラム開発

分子軌道基底での2電子積分生成プログラムの開発

エネルギー勾配法のエネルギー表式計算プログラムの開発
 津田健一郎 NEC 基礎研究所 主任
 数値計算部のパーツ化を図るためのデータセットの統一化
 データ作成及び解析用インターネットツールの開発
 中田一人 NECソフト 主任
 MCSCFプログラムの基本設計及びCI行列要素計算プログラム開発
 MCSCF法における一般化Fock行列計算プログラムの開発
 分子軌道基底での2電子積分生成プログラムの開発
 MCSCF波動関数計算部の並列化性能向上とベンチマーク
 数値計算部の並列分散処理化に伴うコア管理ユーティリティの開発
 数値計算部のパーツ化を図るためのデータセットの統一化
 村瀬匡 NECソフト 主任
 ダイレクト2電子積分計算プログラムの開発
 佐久間俊広 NEC情報システムズ 課長
 エネルギー勾配法のエネルギー表式計算プログラムの開発
 エネルギー勾配法におけるカップリング係数行列生成プログラムの開発
 エネルギー勾配計算部の並列化とベンチマーク
 数値計算部の並列分散処理化に伴うコア管理ユーティリティの開発
 平原幸男 NEC情報システムズ 課長
 数値計算部の並列分散処理化に伴うコア管理ユーティリティの開発
 数値計算部のパーツ化を図るためのデータセットの統一化
 山本章紀 NEC 情報システムズ 主任
 数値計算部のパーツ化を図るためのデータセットの統一化
 データ作成及び解析用インターネットツールの開発
 インターネット対応GUIの評価
 岡本直樹 NEC 情報システムズ 主任
 分子グラフィックス画像データ送信用インターネットツールの開発
 データ作成及び解析用インターネットツールの開発
 インターネット対応GUIの評価

B. 研究協力者 なし

C. 招聘研究協力者 なし

(2) 分散処理のための基本アルゴリズム開発

A. 参加研究者

高田俊和 NEC 基礎研究所 主席研究員
 2電子積分並列変換アルゴリズムの開発
 2次密度行列の逆変換アルゴリズムの開発
 中田一人 NEC ソフト 主任

2 電子積分並列変換アルゴリズムの開発

佐久間俊広 NEC 情報システムズ 課長

2 電子積分並列変換アルゴリズムの開発

B. 研究協力者

関口智嗣 産業技術総合研究所情報処理研究部門 主任研究官

建部修見 産業技術総合研究所情報処理研究部門 研究員

C. 招聘研究協力者 なし

(3) Internet 上での性能評価

A. 参加研究者

高田俊和 NEC 基礎研究所 主席研究員

山本章紀 NEC 情報システムズ 主任

岡本直樹 NEC 情報システムズ 主任

B. 研究協力者

柏木浩 九州工業大学情報工学科 教授

日向寺祥子 東海大学電子計算機センタ 講師

建部修見 産業技術総合研究所情報処理研究部門 研究員

C. 招聘研究協力者 なし

3.7 事業により得られた研究成果

(1) 外部発表等

別添 1 を参照。

(2) 成果プログラム等

プログラム名称 汎用 MCSCF プログラム CAS_AMOSS

機能概要 CASSCF 法に基づいて MCSCF 計算を行うプログラムシステムで、CAS 空間の制限は、開殻軌道の数 8 まで CAS 空間の軌道としては 16 軌道までである。用いる基底関数の数についての制限はプログラムアルゴリズム上ないが、主記憶と並列プロセッサ数との関係で決まる。これらの数値計算部をステアリングするインターネット対応 GUI も整備されている。

使用言語 Fortran, C, Perl, Java

サイズ MCSCF 数値計算部 312 K ライン
インターネット対応 GUI 53 K ライン

備考 プログラムは全て自主開発しているので、公開にあたっての著作権上の問題は基本的に無い。但し、高速化のため一部に PDS を使用しているので、ユーザはそれらを自分でダウンロードし用意する必要がある。また、一部チューニングが不十分の部分があるので、今後数ヶ月をかけてチューニングしたい。従って、一般的な公開は、このチューニング完了後が望ましい。また、NEC としてもビジネスへの活用を検討しているので、当面の間、覚書締結による希望者へのプログラム譲渡を基本にしたいと考えている。

(3) 特許出願記録

出願日 平成13年9月25日
発明の名称 分子積分並列計算方法
発明者 中田一人(25%)、村瀬匡(25%)、佐久間俊広(25%)、
高田俊和(25%)
出願人 科学技術振興事業団(50%)、日本電気(18%)、NECソフト(16%)、NEC
情報システムズ(16%)
整理番号

(4) 新聞記事、雑誌記事、テレビ報道等

なし

(5) 受賞等

なし

(6) ワークショップ等(主催分)

なし

別添1

(a) 原著論文

発表年	論文タイトル	掲載雑誌名 巻・号・ページ	著者名	整理番号
なし				

(b) 口頭・ポスター発表

発表年月日 開催場所	発表タイトル	学会等の名称 (予稿集名、掲載ページ)	発表者	整理番号
2001/03/05 早稲田大学	インターネット適応型 分子シミュレーション GUIの開発	インタラクシオン2001(論文集 P161)	岡本直樹、山本章紀、半田享、高田俊和	12/10A-3 発01
2001/03/16 つくば市	ネットワークコンピューティングに向かう計算化学	つくば WAN シンポジウム	高田俊和	12/10A-3 発02
2001/09/27 北海道大学	並列 CASSCF プログラムの開発	分子構造総合討論会 (講演要旨集 P764)	中田一人、村瀬匡、佐久間俊広、高田俊和	13/10A-3 発01
2001/09/27 山口大学	インターネット適応型 分子シミュレーション GUIの開発(2)	情報処理学会 第63回全国大会 (論文集(分冊1)P91)	岡本直樹、山本章紀、半田享、高田俊和	13/10A-3 発02